Probing multiparticle production properties of the first ultra-high energy cosmic ray-Air interaction in Extensive Air Showers with low muon content

Miguel Alexandre Martins*
Laboratório de Instrumentação e Física Experimental de Partículas (LIP) - Lisbon, Av. Prof. Gama Pinto 2, 1649-003 Lisbon, Portugal and
Instituto Superior Técnico (IST), Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal
(Dated: December 5, 2021)

The energy spectrum of hadronically interacting particles and the production cross-section of neutral pions of the first p-Air interaction are constrained by measurements of the slope, Λ_μ, of the probability density function of the number of muons at the ground level, $P(\ln N_\mu)$, in Extensive Air Showers with low muon content. We show that precise measurements of Λ_μ in mixed composition scenarios can be achieved, within current experimental uncertainties, provided the number of events is large enough and $P(\ln N_\mu)$ is corrected for the cosmic ray flux dependence on the primary energy.

We propose the reconstruction of $P(\ln N_\mu)$ through its correlation with the expected number of active pads in the fiducial area of RPCs of MARTA stations at 450 m from the shower axis. Under ideal conditions, the reconstruction of $P(\ln N_\mu)$ was successful at $E_0 = 10^{18.5}$ eV and zenith angle $\theta = 40°$, despite the electromagnetic contamination. However, a threshold effect prevented a completely unbiased measurement of Λ_μ, and of the first and second moments of $P(N_\mu)$. Results achieved with RPCs over the SD-750 array of the Pierre Auger Observatory can be reproduced with 7 MARTA stations in a minimal hexagonal configuration.

I. INTRODUCTION

Ultra-High Energy Cosmic Rays (UHECR) are relativistic charged particles, either protons or light nuclei, produced and accelerated by astrophysical objects at energies greater than $E_0 = 10^{18}$ eV. Their origin is extra-galactic [1] and their energy spectrum scales like $E_0^{-\gamma}$ [2] before being highly suppressed by source exhaustion or propagation effects, like the GZK cut-off [3, 4] and the giant dipole resonance [5]. The energy, arrival direction and mass composition of UHECRs are determined through measurements of the Extensive Air Showers (EAS) they produce in the upper regions of the Earth's atmosphere. In fact, UHECRs offer a unique opportunity to probe the nearby Universe and improve our description of hadronic interactions at centre-of-mass energies of $\sqrt{s} \sim 100$ TeV.

After the first UHECR-Air interaction, about 75% of the primary energy is carried by hadronically interacting particles, and the shower behaves as two weakly coupled cascades: the electromagnetic cascade and the hadronic cascade. The former is mostly fed by the decay of neutral pions arising from the first UHECR-Air interaction, while in deeper shower generations, muon decay, photo-pion production and low energy hadronic interactions also couple these cascades. The hadronic component is mostly composed of charged pions, kaons and light baryons, which below their critical energy ξ_c, mostly decay into muons measured at the ground level. Thus, these leptons probe high energy hadronic interactions, while the shower-to-shower fluctuations of their number correlate with the primary composition [6].

The development of EAS is simulated with Monte Carlo packages, such as CORSIKA [7] and CONEX [8], which combines Monte Carlo simulations with low energy cascade equations to enhance computational efficiency. In both cases, high energy interactions are simulated with phenomenological hadronic interaction models, which extrapolate from accelerator data, to energies and kinematic regimes beyond the reach of human-made colliders. In fact, at ultra-high energies, differences across models dominate the uncertainty in the moments of the shower observables used to infer the primary composition [9]. Hence, the interpretation of the primary in terms of its mass composition inherits the uncertainties of the hadronic interaction models, whose accurate description relies, in itself, on the primary nature. To break this degeneracy the number of shower observables independent of $X_{\max}$ must be increased.

Measurements of $X_{\max}$ show that, at the highest energies, the primary is heavier than proton [10], pointing towards unconsidered astrophysical sources or phenomena in UHECR propagation, or towards an incorrect modelling of the cascade. Moreover, data from the Pierre Auger Observatory shows that simulations underestimate the muon content of EAS by $26 - 43\%$ for $E_0 > 10^{16}$ eV, despite correctly predicting its shower-to-shower fluctuations [11]. This either points towards a small increase in the fraction of energy carried by hadronically interacting particles over all shower generations or to a large increase in the first generation, possibly due to new exotic phenomena at the highest energies, such as Lorentz Invariance Violation (LIV) [12], or the formation of Quark-Gluon Plasma (QGP) [13].

The Pierre Auger Observatory measures EAS to determine the flux, primary energy, arrival direction and mass composition of UHECRs. It comprises an hexagonal array of 1 661 surface detector stations spaced by 1.5 km over 3 000 km^2, surrounded by 27 fluorescence telescopes in 4 different sites [14]. The observatory's hybrid design allows for important cross-calibrations and high quality EAS reconstructions. Surface detector (SD) stations consist of cylindrical water Čerenkov detectors (WCDs) filled with 12 tons of pure water and three 9 in. photomultiplier tubes (PMTs) viewing the water volume from above. These, capture the Čerenkov light produced by relativistic charged shower particles and amplify the resulting electric signal, which is the sum of the signals produced by the muonic and electromagnetic components. SD stations are calibrated *in situ* using

* miguel.a.j.s.martins@tecnico.ulisboa.pt

the atmospheric muon flux and analysed for time and spatial correlations to discern physics events from background and random coincidences [14, 15]. The regular array attains full efficiency at $E_0 = 10^{18.5}$ eV, while the Infill (SD-750), a denser array with 61 water tanks spaced by 750 m, measures EAS with 100% efficiency for primary energies $> 10^{17.5}$ eV [16]. Fluorescence detectors are described in [14].

Recently, a set of upgrades known as AugerPrime was implemented to better measure the shower components individually [17], along with R&D projects such as MARTA. MARTA (Muon Array with RPCs for Tagging Air showers) is a hybrid detector which combines data from RPCs (resistive plate chambers) with calorimetric measurements of WCDs to measure the muon component of EAS directly [18]. This detector comprises 4 RPC units, for a total detection area of $\sim 7\,\mathrm{m}^2$. Each RPC is composed of three 2 mm thick resistive plates under a high-voltage, separated by 1 mm gaseous volumes, embedded in a uniform electric field. Within the gaseous volumes, ionising shower particles trigger avalanches of electrons, inducing an electric signal on the readout plane. The readout is segmented into 64 individual pads for optimal spatial resolution and the signal pick-up electrodes are physically separated from the gas chambers, for high-voltage insulation and gas tightness. The multi-gap gaseous design provides a time resolution of a few nanoseconds. Importantly, to achieve a purer measurement of the muon content of EAS, the RPCs are enclosed by a 20 cm thick concrete structure and placed underneath WCDs of the SD-750 array, which trigger the MARTA stations. Considering a water depth of $h = 1.2$ m, this corresponds to a vertical mass overburden of ~ 170 g cm^{-2}, partially shielding the RPCs from electromagnetic contamination.

This text is organised as follows. In Section II measurements of the slope of the tail of the muon number distribution are used to constrain the energy spectrum of hadrons and the forward production cross-section of neutral pions of the first p-Air interaction in EAS with low muonic content. Additionally, the possibility of performing these measurements is assessed within current experimental uncertainties for different mixed composition scenarios. In Section III we develop a procedure to reconstruct the muon number distribution using the MARTA engineering array, and compute the resolution and bias induced by the reconstruction on the features of this distribution, for different primary energies and zenith angles. In Section IV conclusions are drawn.

II. ACCESSING THE FIRST INTERACTION IN EAS WITH LOW MUON CONTENT

The hadronic sector of EAS is probed through its decay into muons, which carry information about the hadronic activity in all shower generations, hampering the access to the Cosmic-Ray-Air interaction. However, in [19] it was shown that shower-to-shower fluctuations of the muon content of EAS are dominated by fluctuations of a modified fraction of energy contained in the hadronic sector of the first interaction in proton induced showers, α_1, while deeper shower generations contribute to the overall muon scale.

The origin of the shower-to-shower fluctuations of the muon content of EAS can be understood in terms of the Heitler-Mathews model [20]. This model assumes that: the electromagnetic component is fed by the decay $\pi^0 \to \gamma\gamma$; all hadrons decay into muons at their critical energy ξ_c, at a fixed generation g_c; muons do not decay; in every interaction of the cascade the multiplicity is fixed and energy is equipartitioned among particles of the same generation. Using these assumptions and neglecting the binding energy of nucleons, the average number of muons in EAS, $\langle N_\mu \rangle$, for primaries with mass number A and energy E_0, reads

$$\langle N_\mu \rangle = A \left(\frac{E_0}{A\xi_c} \right)^\beta = A^{1-\beta} \mathcal{C} E_0^\beta \ , \ \beta \equiv \frac{\ln m}{\ln m_{\rm tot}}, \quad (1)$$

where m is the multiplicity of hadronically interacting particles and $m_{\rm tot}$ is the total multiplicity.

In reality, the multiplicity of each interaction varies and energy is shared unevenly among final state particles, resulting in fluctuations of the muon content of EAS. In fact, assuming all hadrons decay into muons at an average critical generation g_c, and letting m_g be the average multiplicity of generation g we have

$$N_\mu = \prod_{g=1}^{g_c} m_g. \quad (2)$$

The multiplicity is a realisation of a random variable with a probability density function $P(m)$, average m and dispersion $\sigma(m)$. Assuming that realisations of the multiplicity are independent, $\sigma(m_g) = \sigma(m)/\sqrt{M_{g-1}}$, where M_{g-1} is the number of hadronically interacting particles in generation $g-1$. Thus, fluctuations of the multiplicity in deeper shower generations are exponentially suppressed and those of the first interaction dominate the fluctuations of N_μ via Eq. (2).

Fluctuations of the energy sharing among final state particles also induce fluctuations of the muon content of EAS. Considering only fluctuations arising from the first interaction, $g = 1$, and assuming that each of the m_1 resulting hadronic subshowers, carrying a fraction $x_i = E_i/E_0$ of the primary energy, decays into $\langle N_\mu(E_i) \rangle = \mathcal{C} E_i^\beta$ muons, we can write, for proton induced showers

$$N_{\mu,1}(E_0) = \mathcal{C} E_0^\beta \sum_{i=1}^{m_1} x_i^\beta = \langle N_\mu(E_0) \rangle \, \alpha_1, \quad (3)$$

where we defined

$$\alpha_1 \equiv \sum_{i=1}^{m_1} x_i^\beta. \quad (4)$$

This variable is a modified fraction of energy carried by hadronically interacting particles, that accounts for both multiplicity and energy fluctuations through β.

By recursively defining α_g, for generation g, we can write

$$N_\mu = \langle N_\mu \rangle \prod_{g=1}^{g_c} \alpha_g = \alpha_1 \omega \quad \text{with} \quad \omega = \langle N_\mu \rangle \prod_{g=2}^{g_c} \alpha_g. \quad (5)$$

Now, for $g = 2$ there are m_1 subshowers contributing to α_2, so that its fluctuations with respect to α_1 are suppressed by $\sim 1/\sqrt{m_1}$. Proceeding recursively, it is clear that the

distribution of α_g in deeper shower generations becomes increasingly narrow. Hence, we expect that fluctuations of N_μ are dominated by those of α_1. On the other hand, through Eq. (5) the actual muon scale depends on all shower generations.

This reasoning was thoroughly tested and verified in fully simulated proton induced EAS, as discussed in [19]. The remaining results presented in this section where recently published in [21].

A. Constraining the energy spectrum of hadrons of the first p-Air ultra high energy interaction

The correlation between shower-to-shower fluctuations of α_1 and N_μ was verified using 10^6 proton initiated showers with primary energy $E_0 = 10^{19}$ eV and zenith angle $\theta = 67°$, with CONEX simulations, using the post-LHC hadronic interaction models: EPOS-LHC [22], QGSJET-II.04 [23] and SIBYLL 2.3d [24], to simulate high energy interactions, and FLUKA [25, 26] for low energy hadronic interactions. The number of muons reaching the ground, N_μ, was measured at 1400 m above sea level, the average altitude of the Pierre Auger Observatory [14].

The values of α_1 and N_μ were computed for each event, and their distributions over the ensemble of showers built. Fig. 1 shows joint distribution of α_1 and N_μ, along with their correlation factor $\rho_{\mu,\alpha} = \mathrm{cov}(\alpha_1, N_\mu)/\sigma(N_\mu)\sigma(\alpha_1)$ for the EPOS-LHC model.

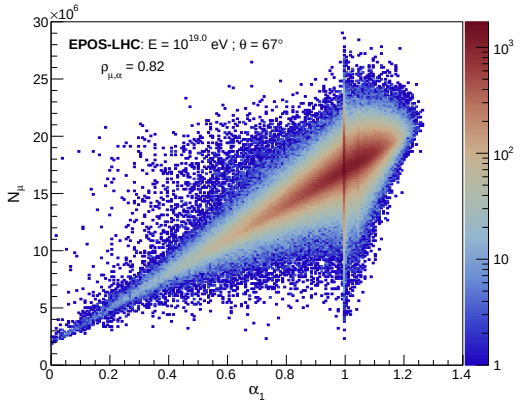

FIG. 1. Correlation between the distributions of α_1 and N_μ over an ensemble of $\sim 10^6$ proton induced EAS, with primary energy $E_0 = 10^{19}$ eV and zenith angle $\theta = 67°$. Simulations were performed with CONEX, using the hadronic interaction model EPOS-LHC. The correlation factor between the two quantities can be read off the upper left corner.

The strong positive correlation between α_1 and N_μ is evident from $\rho_{\mu,\alpha} \simeq 0.82$. The non-vanishing y-intercept is due to photo-pion production. The vertical structure along $\alpha_1 \simeq 1$, corresponds to diffractive events in which the incident proton scatters quasi-elastically of the air nucleus. Given the low energy transfer, the scattered proton is still highly energetic and hence able to yield large multiplicities of hadronically interacting particles in subsequent interactions. Thus, N_μ becomes more sensitive to deeper interactions locally breaking its correlation with α_1. A similar less pronounced feature for $\alpha_1 \simeq 1.05$, can also be observed. Weather it has physical meaning or whether it is an artefact, is not yet known.

Fig. 2 shows the p.d.fs of α_1 and $\ln N_\mu$. These distributions display a left low tail corresponding to events with low muon content, where an appreciable fraction of the primary energy is carried by the electromagnetic sector of the first p-Air interaction. The sharp peak at $\alpha_1 \simeq 1$ is due to diffractive events. The steepness of the low tail of both distributions is model dependent, along with the moments of the α_1 and $\ln N_\mu$ probability densities. These differences stem from the different physical laws employed by the hadronic interaction models.

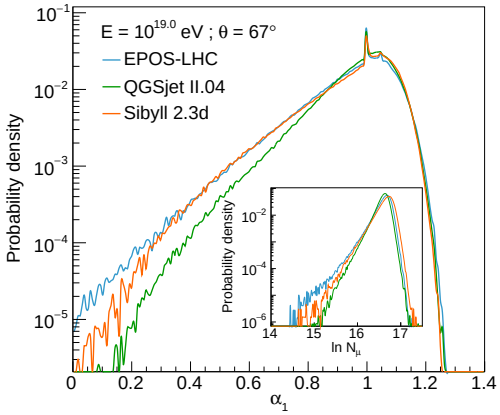

FIG. 2. Distributions of α_1 and $\ln N_\mu$ (inset) over an ensemble of $\sim 10^6$ proton induced EAS, with primary energy $E_0 = 10^{19}$ eV and zenith angle $\theta = 67°$. Simulations were performed with CONEX, using the hadronic interaction models EPOS-LHC (blue), QGSJET-II.04 (green) and SIBYLL 2.3d (orange).

The low tails of the aforementioned distributions were fitted to exponential functions of the form $y = \mathcal{C}_\alpha \exp(\alpha_1/\Lambda_\alpha)$ and $y = \mathcal{C}_\mu \exp(\ln N_\mu/\Lambda_\mu)$, with $\mathcal{C}_\alpha$, Λ_α, $\mathcal{C}_\mu$ and Λ_μ free parameters. The fit range was selected by fixing its upper limit at a given distance from the maximum of the distributions and decreasing its lower bound until the vertical distance between the real distribution and the fit differed by more than 5%. This criterion has the advantage of being independent of the model. Additionally, we verified that small variations of the fit window, did not compromise the qualitative results shown in this section.

The fit was performed with the method of least squares for all hadronic interaction models, yielding values of Λ_α for each model incompatible with each other. Thus, they characterise, in a model dependent way, the energy spectrum of hadrons of the first shower generation. The values of Λ_μ for each model are also incompatible with each other.

The correlation between the distributions of α_1 and $\ln N_\mu$ was exploited to investigate the connection between Λ_μ and Λ_α. To do so, we perturbed the tail of the α_1 distribution below $\alpha_1 = 1$ by re-sampling the ensemble of events and selecting pairs of $(\ln N_\mu, \alpha_1)$ with a probability density function $p(\alpha_1) \propto \mathrm{e}^{(\alpha_1-1)/\delta\Lambda_\alpha}$, where $\delta\Lambda_\alpha$ parametrises the perturbation. Thus, the exponential slope of the α_1 distribution is effectively changed by $\Lambda_\alpha^{-1} \to \Lambda_\alpha^{-1} + \delta\Lambda_\alpha^{-1}$, while keeping its non-exponential features. Hence, through its correlation with α_1, the slope of the tail of $\ln N_\mu$ is also changed $\Lambda_\mu^{-1} \to \Lambda_\mu^{-1} + (\delta\Lambda'_\mu)^{-1}$. The tails of the re-sampled α_1 and $\ln N_\mu$ distributions were then fitted with the criterion discussed previously, yielding pairs of $(\Lambda_\mu, \Lambda_\alpha)$ for each perturbation $\delta\Lambda_\alpha$. These perturbations are small and assume that the hadronic properties of the interaction remain

unchanged, covering a limited region around the nominal $(\Lambda_\mu, \Lambda_\alpha)$ values. Moreover, this study does not emulate any physical process, which would ultimately change other multiparticle production properties in p-Air interactions.

By varying $\delta\Lambda_\alpha$, we can parametrise curves in the $(\Lambda_\mu, \Lambda_\mu)$ space for each model, which allow the conversion of measurements of Λ_μ into values of Λ_α as shown in Fig. 3. Solid circles indicate the nominal pair $(\Lambda_\mu, \Lambda_\alpha)$. Also displayed in Fig. 3 are the fitted nominal and modified distributions of α_1 and $\ln N_\mu$ for a specific perturbation.

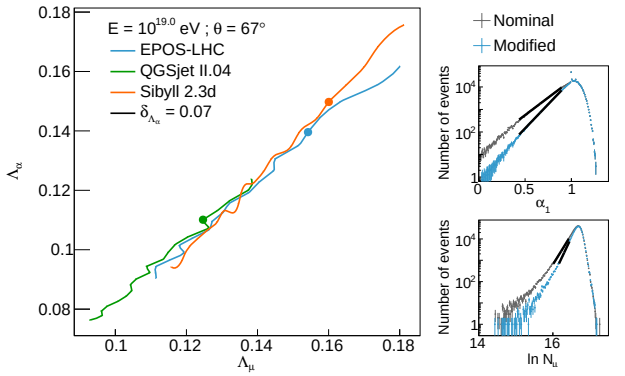

FIG. 3. Conversion curves between Λ_μ and Λ_α, for the hadronic interaction models EPOS-LHC (blue), QGSJET-II.04 (green) and SIBYLL 2.3d (orange). The solid lines show how Λ_μ changes if Λ_α is changed. Solid dots represent the nominal Λ_α and Λ_μ values. Upper inset: Nominal and modified distributions of α_1. Bottom inset: Nominal and modified distributions of $\ln N_\mu$.

There is a monotonic relation between Λ_α and Λ_μ for all models, with a residual model dependence of $\delta_{\Lambda_\alpha} = 7\%$. We have also applied the reverse re-sampling procedure, where perturbations $\delta\Lambda_\mu$ were directly applied to the tail of the distribution of $\ln N_\mu$ and the distribution of α_1 is affected by the selection of pairs of $(\ln N_\mu, \alpha_1)$ with $p(\ln N_\mu) \propto \mathrm{e}^{\ln N_\mu/\delta\Lambda_\mu}$. Although the results were similar, the model dependence increased due to a rotation of the calibration curves around the nominal values $(\Lambda_\mu, \Lambda_\alpha)$, stemming from the imperfect correlation between α_1 and N_μ.

We conclude that Λ_μ probes Λ_α, constraining the energy spectrum of hadronically interacting particles of the first p-Air interaction in showers with low muon content. Note that, the above *ad-hoc* considerations do not bring insight into the origin of the universality of the calibration curves across models, which is dependent on the functional form of the perturbation to the distribution of α_1. In future works, the parameters of the simulations will be changed *a priori*, inducing a physical change in Λ_α, and thus on Λ_μ.

A few remarks are in need. By the superposition principle, for primaries with mass number A the distribution of $\ln N_\mu$ is a convolution of A independent $\ln N_\mu$ distributions for proton induced showers. Thus, fluctuations of the muon content of EAS induced by primaries heavier than proton are narrower by a factor of $\sim 1/\sqrt{A}$, hiding their the tail of these distributions underneath the distribution of $\ln N_\mu$ in proton induced EAS. While this means that the conversion from Λ_μ to Λ_α is only valid for proton primaries, it also means that Λ_μ could still be measured in a mixed composition scenario, provided it includes proton primaries and the fit range is judiciously chosen. Moreover, we have considered a well defined primary energy for the entire ensemble of events, while in reality they spread over an energy bin, which distorts the distribution of $\ln N_\mu$ via the dependence of N_μ and the cosmic ray flux on the primary energy. Finally, we have not considered uncertainties in the reconstruction of the primary energy and muon content of EAS.

B. Constraining the energy spectrum of neutral pions in ultra-high energy p-Air interactions

To first order, the fraction of energy contained in the electromagnetic sector of the first interaction, reads $f_{\mathrm{em}} \simeq 1 - \alpha_1$, and about 90% of f_{em} is carried by neutral pions.

By isospin symmetry, the number and energy carried by the bulk of neutral and charged pions is roughly the same, so the energy imbalance toward π^0's in showers with low muon content is likely due to fast π^0's. These are connected with forward production and leave less energy in the hadronic shower available to produce muons. Thus, it is natural to study the connection between the distribution of $\ln N_\mu$ and the forward region of the energy spectrum of neutral pions, in showers with low muon content. In fact, we considered the fraction of energy taken by the leading (most energetic) neutral pion, in the laboratory frame, x_{L}, instead of the full energy spectrum. The reason for this choice is two-fold: on the one hand, in deep inelastic scattering, the leading particle takes a significant portion of the available initial state energy, so the features of the forward region of the full energy spectrum should be dominated by the leading particle; and on the other, considering a spectrum of neural pions per event, rather than a simple scalar, breaks the one-to-one relation with N_μ needed to perform the re-sampling method described earlier.

We verified that the correlation factor between x_{L} and N_μ is $\rho \sim -0.6$ for all hadronic interaction models, and thus weaker than the correlation between α_1 and $\ln N_\mu$. This might be due to events where there is an energy imbalance towards the electromagnetic sector, yielding low N_μ, but where most of the energy is taken by more than one electromagnetic particle, yielding low x_{L} values. Furthermore, this correlation is further broken by the impact of deeper shower generations. Nevertheless, the fraction of energy carried by the leading neutral pion of the first interaction plays a dominant role in muon production.

The forward region of the energy spectrum of the leading neutral pion, in the laboratory reference frame, is not exponential. However, by weighing each bin of $\mathrm{d}\mathcal{N}/\mathrm{d}x_{\mathrm{L}}$ events by x_{L} we get a high exponential tail which was fitted, using the method of least squares, to an exponential function of the form $y = \mathcal{C}_\pi \exp(x_{\mathrm{L}}/\Lambda_\pi)$, where $\mathcal{C}_\pi$ and Λ_π are free parameters. Note that this definition of Λ_π differs from the one used in [21], where the un-weighted distribution of x_{L} was fitted. The fit region was chosen so that only 5% deviations from a pure exponential were allowed. As expected, the values of Λ_π for the different models are incompatible with each other, since the hadronic interaction models employ different physical laws.

The connection between Λ_μ and Λ_π was verified using the already described re-sampling procedure. The high tail of the distribution of x_{L} was perturbed above its maximum, x^*_{L}, by selecting pairs of values $(x_{\mathrm{L}}, \ln N_\mu)$ with probability $p(x_{\mathrm{L}}) \propto \mathrm{e}^{(x^*_{\mathrm{L}} - x_{\mathrm{L}})/\delta\Lambda_\pi}$. Thus, the slope of the high tail of

the distribution of $x_{\rm L}$ weighted by $x_{\rm L}$ is changed by $\Lambda_\pi^{-1} \to \Lambda_\pi^{-1} + \delta\Lambda_\pi^{-1}$. Via its correlation with $x_{\rm L}$, the slope of the tail of the $\ln N_\mu$ distribution is also changed, with little effect on the rest of the distribution. These perturbations are not an attempt to reproduce any specific physical process, but rather a change in the frequency of events with high $x_{\rm L}$, to study its influence on Λ_μ.

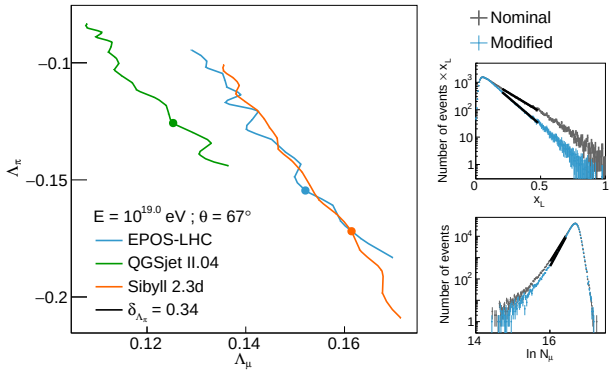

FIG. 4. Conversion curves between Λ_μ and Λ_π, for the hadronic interaction models EPOS-LHC (blue), QGSJET-II.04 (green) and SIBYLL 2.3d (orange). The solid lines show how Λ_μ changes if Λ_π is changed. Solid dots represent the nominal Λ_π and Λ_μ values. Upper inset: Nominal and modified distributions of $x_{\rm L}$ weighted by $x_{\rm L}$. Bottom inset: Nominal and modified distributions of $\ln N_\mu$.

The tails of the distributions arising from this re-sampling procedure were then fitted, as shown in the insets of Fig. 4 for the EPOS-LHC model. As expected, the softer the spectrum of the leading neutral pions, the steeper the tail of the distribution of $\ln N_\mu$ as there is more energy available to produce muons. By varying $\delta\Lambda_\pi$, we parametrise calibration curves for each model, through which measurements of Λ_μ can be converted to values of Λ_π as shown in Fig. 4. Filled circles indicate the nominal pair $(\Lambda_\mu, \Lambda_\pi)$.

There exists a monotonic relation between Λ_μ and Λ_π for all models, with a remnant model dependence of $\delta_{\Lambda_\pi} \sim 34\%$. This increased model dependence with respect to the conversion between Λ_μ and Λ_α may stem from the weaker correlation between $x_{\rm L}$ and N_μ. Nevertheless, a thorougher study where the full spectrum of neutral pions is changed *a priori* will be performed using full Monte Carlo simulations.

Although the conversion from measurements of Λ_μ to Λ_π is manifestly model dependent, one can still constrain the forward region of the energy spectrum of neutral pions, hence its production cross section in p-Air interactions. Thus, Λ_μ could be used to exclude exotic physics phenomena affecting the production of neutral pions at the highest energies.

C. Measurement of Λ_μ

Measurements of $X_{\rm max}$ are compatible with heavier primaries at the highest energies, for which the $\ln N_\mu$ distribution is an A-fold convolution of $\ln N_\mu$ distributions for proton. Moreover, the energy spectrum of cosmic rays is a power law in the primary energy, on which N_μ depends via Eq. 1. Thus, distributions of $\ln N_\mu$ over a bin $\mathrm{d}E$ are distorted. Finally, the measured $\ln N_\mu$ distribution is convoluted with the finite reconstruction resolution of N_μ and E_0.

The distribution of $\ln N_\mu$ for an ensemble of showers with $E_0 \pm \mathrm{d}E$ can be viewed as superposition of distributions of $\ln N_\mu$ shifted according to each fixed E_0 and increasingly suppressed by $E_0 \propto E_0^{-\gamma}$. Thus, to extract Λ_μ, the distortion induced by this effect must be mitigated by plotting the distribution of $\ln(N_\mu/E_0^\beta)$, which is independent of the shower energy.

Realistic mixed composition scenarios were studied using CONEX simulations of EAS with primary energy $E_0 = 10^{18.7}$ eV and zenith angle $\theta = 67°$, using different primaries. High energy hadronic interactions were handled with SIBYLL 2.3c, as in [21]. The mixed composition is achieved by combining proton, helium, nitrogen, iron and photon primaries in different proportions. Additionally, each distribution of $\ln N_\mu$ was smeared by 20% to emulate current uncertainties in the reconstruction of E_0 and N_μ [1]. The distribution of $\ln N_\mu$ for p:He:N:Fe in proportion 2:1:1:0, along with a photon contamination of 0.5 % [28] is plotted in Fig. 5. The shaded area corresponds to the fit region.

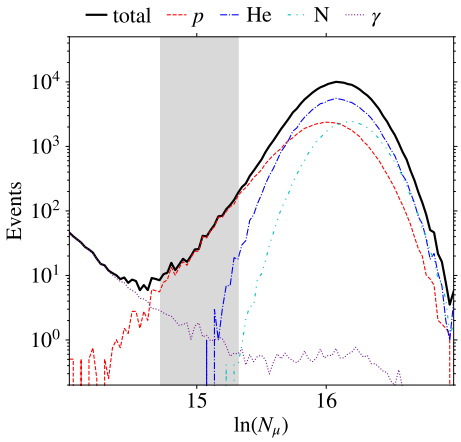

FIG. 5. Distribution of $\ln N_\mu$ over an ensemble of EAS initiated by a mixture of p:He:N:Fe in proportion 2:1:1:0 with primary energy $E_0 = 10^{18.7}$ eV and zenith angle $\theta = 67°$, simulated with CONEX, using SIBYLL 2.3c. The distributions corresponding to proton, helium and nitrogen are shown in red, blue and cyan, respectively, along with a photon contamination shown in purple. The total distribution is shown in black. A Gaussian smearing of 20% was applied to each distribution. Plot taken from [21].

From Fig. 5 we see that the photon contamination affects only the low end of the tail of the $\ln N_\mu$ distribution, while He contamination affects the high tail region. The impact of heavier primaries is negligible since fluctuations of their muon content are suppressed by $1/\sqrt{A}$. Thus, the fit region was chosen to avoid the mentioned contaminations, where the total and proton distributions $\ln N_\mu$ coincide. The increased width of all distributions stems the 20% Gaussian smearing.

The relative accuracy of the measurement of Λ_μ with respect to a pure proton case, $\delta_\mu = 1 - \Lambda_\mu^{\rm mixed}/\Lambda_\mu^{\rm p}$, was determined for different mixed composition scenarios, along

[1] Indeed the current resolution attained by the Pierre Auger Observatory of 17% [27]

with the minimal number of events required to distinguish between hadronic interaction models, i.e., $\delta_\mu = 20\%$ [29]. Results are shown in Tab. I.

Scenario	Model	$\mathcal{N}_{\rm tail}$	$\mathcal{N}_{\rm total}$
1:1:1:1	QGSJet-II.04	21	1 564
	EPOS-LHC	29	1 926
	Sibyll 2.3c	30	1 667
1:2:1:0	QGSJet-II.04	32	7 056
	EPOS-LHC	36	5 505
	Sibyll 2.3c	33	4 411
1:6:2:0	QGSJet-II.04	205	385 776
	EPOS-LHC	132	136 212
	Sibyll 2.3c	123	78 482

TABLE I. Number of events within the fit range, $\mathcal{N}_{\rm tail}$, and in the total distribution of $\ln N_\mu$, $\mathcal{N}_{\rm total}$, required to measure Λ_μ with a precision of $\delta_\mu = 0.2$, for distinct composition scenarios of the form p:He:N:Fe for the hadronic interaction models EPOS-LHC, QGSJet-II.04 and Sibyll 2.3c. Adapted from [21].

Regardless of the composition scenario, the bias in the measurement of Λ_μ was negligible [30]. Moreover, even in extreme mass composition scenarios, it is still possible to measure Λ_μ with the required accuracy provided the number of events is large enough and that Λ_μ is extracted from the distribution of $\ln(N_\mu/E_0^\beta)$, assuming a Gaussian experimental resolution that does not hide the tail of the distribution of $\ln N_\mu$.

For the composition compatible with $X_{\rm max}$ measurements, the precise determination of Λ_μ requires about 5000 events. For the energy range where the detectors of the regular array of the Pierre Auger Observatory have full efficiency, i.e., above $10^{18.5}$ eV, this is still a difficult number to reach. This number could be lowered if the reconstruction resolution was improved by better disentangling the electromagnetic and muonic components in EAS, as proposed by AugerPrime upgrades, and MARTA. Moreover, the latter engineering array is set to be placed underneath stations of the Infill array of the Pierre Auger Observatory, which allow for the measurement of showers at $E_0 = 3 \times 10^{17}$ eV, with 100% efficiency, where the cosmic ray flux is about 10^3 larger than that at the threshold energy of the regular array. Therefore, a precise measurement of Λ_μ could be achieved faster and at energies where the forward region of the energy spectrum of neutral pions is already being measured in proton-lead collisions at the LHCf [31]. Thus, explore the capabilities of the MARTA engineering array to reconstruct the distribution of $\ln N_\mu$ and extract Λ_μ.

III. RECONSTRUCTION OF THE MUON NUMBER DISTRIBUTION IN PROTON INDUCED EAS WITH THE MARTA ENGINEERING ARRAY

To reconstruct the muon number distribution using the MARTA engineering array, about 6 000 full 3-D CORSIKA v7.7410 simulations of proton induced EAS with primary energy of $E_0 = 10^{17.5}$ eV and $\theta = 30°$, were run with EPOS-LHC, and Fluka for low energy hadronic interactions. Vertical showers were preferred for this study to reduce particle losses, since the effective detection area presented by RPCs to showers with zenith angle θ is suppressed by $\cos\theta$. Besides, the reconstruction of inclined showers is too involved for the purposes of our analysis. The $10^{1.5}$ decrease in the primary energy with respect to the previous section allows for a 10^4-fold increase in the cosmic ray flux, while ensuring a 100% trigger efficiency of the SD-750 array, without saturating it. In fact, taking $E_0 = [10^{17.4}, 10^{17.6}]$ eV and $\theta \in [25°, 35°]$, about 75 events a month fall within a unitary hexagonal cell with spacing $d = 750$ m. Thus, from a statistical point of view, a precise measurement of Λ_μ for $\mathcal{N} \sim 5\,000$ events could be achieved in less than 6 years with 7 MARTA stations. Furthermore, using $\sim 100\,000$ CONEX simulations with EPOS-LHC we verified that Λ_μ can still be converted into values of Λ_α an Λ_π at $E_0 = 10^{17.5}$ eV and $\theta = 30°$.

Auger events used for the reconstruction of the muon number distribution based on the information retrieved by the RPCs were simulated with the Offline framework [14, 32], while the physical response of the RPCs was simulated with Geant4 [33]. The trigger of MARTA stations was provided by the SD water tanks. In fact the trigger hierarchy of the surface array of the Pierre Auger Observatory discerns physical events from the background of atmospheric muons and random coincidences by looking at spatial and time correlations between SD stations [15]. Triggered stations of a physical event are known as candidate stations, while non-triggered ones are flagged as silent.

The core of Auger events was chosen at random and uniformly within a 2.25 km^2 square tile. In a first iteration, the reconstruction was conducted with RPCs in all stations of the Infill array, before considering an engineering array of 7 stations: a central one surrounded by its 6 first neighbours. We assumed a 100% detection efficiency for the RPCs and neglected systematic uncertainties of the reconstruction of the shower geometry and primary energy. Moreover, N_μ denotes the true number of muons measured at the ground level, i.e., 1400 m above sea level, with $E_\mu > 0.2$ GeV.

A. Impact of selecting a fiducial RPC area in electromagnetic contamination

Despite the vertical mass overburden of $\sim 170\,{\rm g\,cm^{-2}}$, the RPCs of MARTA stations are subjected to electromagnetic contamination coming from highly energetic electrons closer to the shower axis; from the decay of muons which completely traverse the water tanks; and from electrons hitting the lateral surface of the WCDs with a track length inside the water volume short enough to reach the RPCs. Since RPCs register the number of hits, provided their spread in time is larger than the dead time of the detector, regardless of the flavour of the charged particle, we defined a fiducial area where the number of hits is a good estimate of the number of muon hits. The fiducial area is the set of pads whose slant mass, assuming a planar shower front, is larger than $120\,{\rm g\,cm^{-2}}$ (the vertical mass overburden) [18].

Fig. 6 shows the average lateral distribution function, LDF, of the number of total, muon and electromagnetic hits on the full and fiducial RPCs areas. For comparison, the LDF of the total, muonic and electromagnetic signals measured by candidate WCDs are also shown. All particles reaching the RPCs are counted as hits. The grey band for $r \lesssim 100$ m corresponds to the distance before which the LDF of the number of muon hits stops being monotonic.

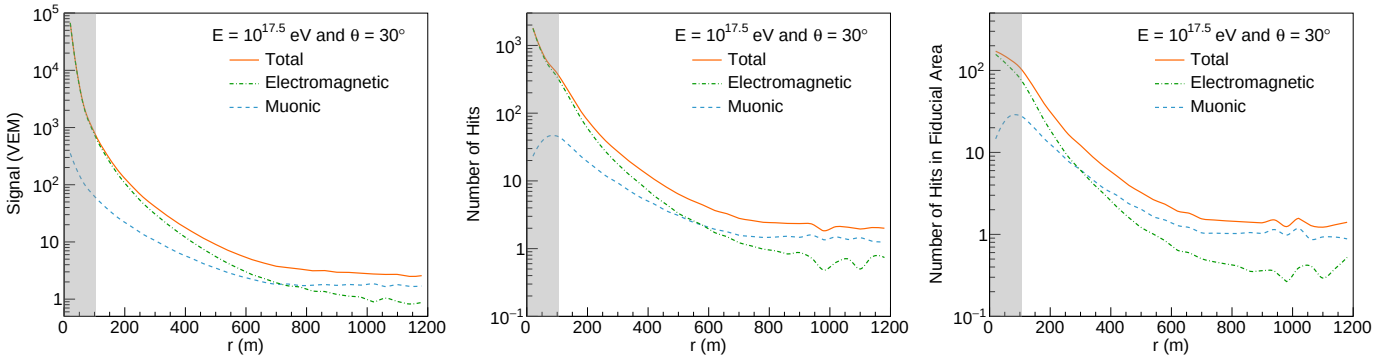

FIG. 6. Left panel: average LDFs of the total (orange), electromagnetic (green) and muonic (blue) signals. Middle panel: average LDFs of the total, muonic and electromagnetic number of hits on the entire RPC area. Right panel: average LDFs of the total, muonic and electromagnetic number of hits on the fiducial area of the RPC. The ensemble of $\sim 6\,000$ CONEX simulations is for proton initiated showers with $E_0 = 10^{17.5}$ eV and $\theta = 30°$, using EPOS-LHC, and the Auger event is simulated with Offline.

This value is approximate and excludes unclear phenomena close to the shower axis.

Although, on average, electromagnetic particles produce a dimmer signal than muons, the electromagnetic signal dominates the measured signal by WCDs for $r < 700\,\mathrm{m}$, since these are not shielded. Far form the shower axis, the total signal stabilises at ~ 3 VEM, due to positive fluctuations induced by the trigger threshold. Using the full RPC area, the absorption of the electromagnetic component within the WCDs is apparent, and the muon hits dominate the total hits for $r > 550\,\mathrm{m}$. The selection of the fiducial area further improves the purity of the measure of the muonic component, dominating the total hits above $r > 300\,\mathrm{m}$. In fact, at $r = 450\,\mathrm{m}$, the distance at which the normalisation of the LDF is extracted for the SD-750 array, about 60% of the simulated hits in the fiducial area of the RPCs are produced by muons.

The reconstructed variable closer to the number of simulated muon hits on the RPCs is the total number of reconstructed hits in the fiducial area corrected for pile-up effects. However, corrections for pile-up effects such as the ones described in [34], must be further studied and are not yet implemented in the Offline trunk, neither is a calibration between the number of muon hits and the deposited charge on the RPCs, which could improve the purity of the measurement. These improvements are left for a second iteration of this work. Instead, we approximate the number of muons reaching the RPCs by the number of active fiducial pads (large dead-time limit) and study the impact of the residual electromagnetic contamination through a direct comparison with the results obtained with the simulated muon hits on the RPCs. In fact, we verified that there is a correlation of $\rho \sim 0.93$ between the number of muon hits and active pads in the fiducial area of each station, indicating that the latter quantity is a good proxy to the number of muon hits. The verified correlation is broken by the residual electromagnetic contamination and by multiple muon hits on the same pad close to the shower axis where the density of muons is larger than the spatial resolution of the RPCs.

B. Average Muon LDF

The proxy to N_μ is the normalisation of the Muon Lateral Distribution Function (MLDF), $\rho_\mu(r)$, since its steepness varies little for showers with fixed energy and zenith angle. In Fig. 7 we compare $\langle \rho_\mu(r) \rangle$ with the average LDFs of the number density of muons whose trajectories cross the WCDs, the number density of muon hits on the RPCs, $\langle N_\mu^{\mathrm{Hits}}(r) \rangle / A_{\mathrm{RPC}}^{\mathrm{eff}}$, and the number density of active fiducial pads of the RPCs, $\langle M(r) \rangle / A_{\mathrm{fid}}^{\mathrm{eff}}$. The effective areas of the WCDs and RPCs for $\theta = 30°$ read: $A_{\mathrm{tank}}^{\mathrm{eff}} \simeq 11.0\,\mathrm{m}^2$, $A_{\mathrm{RPC}}^{\mathrm{eff}} \simeq 6.7\,\mathrm{m}^2$ and $A_{\mathrm{fid}}^{\mathrm{eff}} \simeq 5.0\,\mathrm{m}^2$.

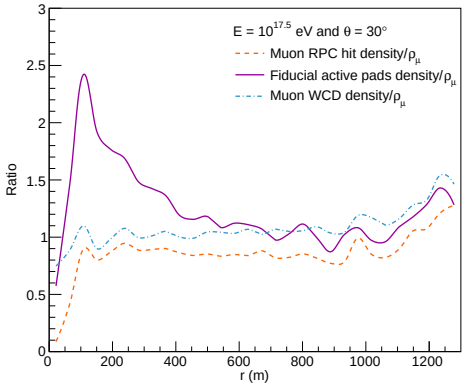

FIG. 7. Ratio between the average LDFs of the number density of muon hits on the RPCs (orange), of active fiducial pads (purple) and of muons crossing the WCD (blue), and the real MLDF. Both figures were produced using 1 000 CORSIKA simulations of proton initiated showers with $E_0 = 10^{17.5}$ eV and $\theta = 30°$, with EPOS-LHC. The corresponding Auger events were simulated and reconstructed using the Offline framework.

As a base case, the average LDF of the number density of simulated muons crossing the WCDs is very close to the true MLDF, except for the suppression close to the shower axis, and the positive fluctuations above $r > 1000\,\mathrm{m}$. Besides the suppression close to the shower axis and positive fluctuations, $\langle N_\mu^{\mathrm{hits}}(r) \rangle / A_{\mathrm{RPC}}^{\mathrm{eff}} \sim 0.85 \times \langle \rho_\mu(r) \rangle$ for $r \in [100, 1000]\mathrm{m}$. This number is expected since, on average, muons with a kinetic energy of $E_{\mathrm{min}} = 390\,\mathrm{MeV}$ do not traverse the WCDs. The residual electromagnetic contamination on $\langle M(r) \rangle$ is prominent for $150\,\mathrm{m} < r < 600\,\mathrm{m}$. Note that for $r < 150\,\mathrm{m}$, where the density of shower particles is much larger than the spatial resolution of the RPCs, $\langle M(r) \rangle$ saturates since the RPCs only have 256 pads.

From the above considerations we expect that the overall scale of the LDF of the number of active fiducial pads of the RPCs, $M(r)$, is mostly sensitive to fluctuations of the muon content of EAS at the distances where the normalisation of

the LDF is measured. The impact of the residual sensitivity to fluctuations of the electromagnetic component are studied by a direct comparison with the scale of the LDF of the number of muon hits on the RPCs, $N_\mu^{\rm Hits}(r)$.

C. Reconstruction of the event-by-event LDF

For each event, N_μ is inferred by fitting $M(r)$ and determining its value at an optimal radial distance from the shower axis, $M(r_{\rm opt})$. Base cases are established by fitting the LDF of the reconstructed signal measured by WCDs, $S(r)$, in VEM, and $N_\mu^{\rm Hits}(r)$. A modified Nishimura-Kamata-Greisen (NKG) fit function was used [35]:

$$\mathrm{LDF}(r) = \mathrm{LDF}(r_{\rm opt}) \left(\frac{r}{r_{\rm opt}}\right)^{\beta} \left(\frac{r + r_{\rm scale}}{r_{\rm scale} + r_{\rm opt}}\right)^{\beta+\gamma} \quad (6)$$

where $r_{\rm scale} = 700\,\mathrm{m}$. The parameter β controls the overall shape of the LDF and γ its steepness far from the shower axis, where the muon component dominates. Both shape parameters depend on the zenith angle and energy of the shower. Furthermore, $r_{\rm opt}$ is the radial distance least dependent on the functional shape of the LDF, which we fix to $r = 450\,\mathrm{m}$ for the SD-750. The fit to $S(r)$ is usually done is several stages, however we implemented a simpler single stage version.

To estimate $S(r)$, β and γ were written as a polynomial in $\log_{10} S(450)$ and $\sec\theta$ according to [36]. The shape parameters for $M(r)$ and $N_\mu^{\rm Hits}(r)$ were extracted from a χ^2 fit to their average LDFs for fixed E_0 and θ. The fit range was limited to avoid saturation close to the shower axis for $r \lesssim 100\,\mathrm{m}$, and positive fluctuations for $r \gtrsim 1000\,\mathrm{m}$. Freeing the shape parameters for each event or finding a suitable parametrisation of β and γ is left for a second iteration of this work.

The normalisations $S(450)$, $M(450)$ and $N_\mu^{\rm Hits}(450)$ were extracted from a likelihood fit to the corresponding LDFs, considering the trigger probability of SD stations and saturation effects. To implement this fitting method, we note that a measurement is a realisation of a random variable with a piecewise p.d.f according to its value, and compute, for each station, the probability of a measured signal, given the expected signal for a set of parameters of the NKG fit function, $\mathcal{L}_i$. The likelihood is obtained via $\mathcal{L} = \prod_i \mathcal{L}_i$. By varying the fit parameters, their most likely value is found by minimising $-\ln\mathcal{L}$ with the `MINUIT` package.

To estimate $S(r)$, the measured signals were converted into an effective number of particles, n, as described in [37], and the corresponding likelihood function reads

$$\mathcal{L}_{\rm SD} = \prod_{i=1}^{n_{\rm cand}} P(n_i,\mu_i) \prod_{j=1}^{n_{\rm sat}} P_{\rm sat}(n_j,\mu_j) \prod_{k=1}^{n_{\rm zero}} P_{\rm zero}(n_k,\mu_k), \quad (7)$$

where $\mu = \mu(S(450),\beta,\gamma)$ is the expected number of particles according to Eq. (6), i runs over $n_{\rm cand}$ candidate unsaturated stations, j over the $n_{\rm sat}$ candidate saturated stations and k over the $n_{\rm zero}$ silent stations. Here $P(n_i,\mu_i)$ is a Poisson distribution, for unsaturated candidate stations. For saturated signals, either the signal is recovered [38], and it follows a Gaussian whose σ takes into account the recovery uncertainty, or the signal can not be recovered, and the p.d.f reads $P_{\rm sat}(n_i,\mu_i) = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{n_i-\mu_i}{\sqrt{2}\sigma}\right)\right]$, where $\mathrm{erf}(x)$ denotes the error function. For silent stations, we consider the probability that the measured effective number of particles below $n_{\rm th} = 3$ [36], $P_{\rm zero} = \sum_{i=1}^{n_{\rm th}} P(n_i,\mu_i)$.

Since MARTA stations are triggered by WCDs, we set their signal to 0 for silent stations and took into account the trigger probability in their likelihood fit. Furthermore, saturated stations were excluded from the fit, as there is no algorithm to recover them. The likelihood to estimate $M(r)$ thus reads

$$\mathcal{L}_{\rm M} = \prod_{i=1}^{n_{\rm cand}} P(m_i,\mu_i) \prod_{k=1}^{n_{\rm zero}} P_{\rm zero}(n_k,\mu_k), \quad (8)$$

where m_i is the measured number of active fiducial pads of candidate station i, $\mu_i = \mu_i(M(450),\beta,\gamma)$ is the expected number of active fiducial pads. $P_{\rm zero}(n_k,\mu_k)$ is the same as before, and $P(m_i,\mu_i)$ is a Poisson distribution.

To fit $N_\mu^{\rm Hits}(r)$, we have not considered trigger effects, so that the likelihood function to maximise is a product of the Poisson p.d.fs followed by the number of muon hits on the RPCs of each event station.

D. Reconstruction of the muon number distribution with the full SD-750 array

The values of $S(450)$, $M(450)$ and $N_\mu^{\rm hits}(450)$ were plotted against N_μ, scaled by 10^{-5}. Events with saturated stations were excluded. In the end, 5 105 of the original ensemble of 6 000 proton initiated events with $E_0 = 10^{17.5}$ eV and $\theta = 30°$, simulated with CORSIKA and EPOS-LHC were kept and used to produce the density plots shown in Fig. 8. The correlation factor and a χ^2 linear fit with equation $y = m(N_\mu/10^5) + b = \tilde{m}N_\mu + b$ are also displayed in each plot of the same figure.

It is apparent, that the correlation between $M(450)$ and N_μ is better than the correlation achieved with $S(450)$, indicating that $M(450)$ is more sensitive to fluctuations of the muon component. The residual electromagnetic contamination may explain the non-vanishing y-intercept of $M(450) = \tilde{m}N_\mu + b$. The sensitivity to fluctuations of N_μ is not greatly improved by considering directly the number of muon hits on the RPCs. Finally, the flat shape of all density plots is indicative of a threshold effect.

Using the linear regressions from Fig. 8, the number of muons of each shower was reconstructed using $N_\mu^{\rm rec} = (X(450) - b)/\tilde{m}$, where $X(450)$ can be $M(450)$ or $N_\mu^{\rm Hits}(450)$. The resolution and bias of the reconstruction were extracted from a Gaussian fit the distribution of residuals $r_\mu = \left(N_\mu^{\rm rec} - N_\mu\right)/N_\mu$. Despite the good quality of the fit, the presence of tails towards high r_μ values confirms the threshold effect. A positive bias of about 5% is induced by the reconstruction, for both $M(450)$ and $N_\mu^{\rm Hits}(450)$. The reconstruction based on $M(450)$ induces a resolution of $\sigma_{\rm rec} = 32\%$, which is larger than that based on $N_\mu^{\rm Hits}(450)$, $\sigma_{\rm rec} = 23\%$. Moreover, the tail of distribution of $\ln N_\mu$ can be reconstructed with a resolution of $\sigma_{\rm rec} = 41\%$ using the physically accessible variable $M(450)$.

The reconstructed distributions of $\ln N_\mu$ from $M(450)$ and $N_\mu^{\rm Hits}(450)$, were built and the slope of their low tail, $\Lambda_\mu^{\rm rec}$, was extracted by fitting the entire distribution to a

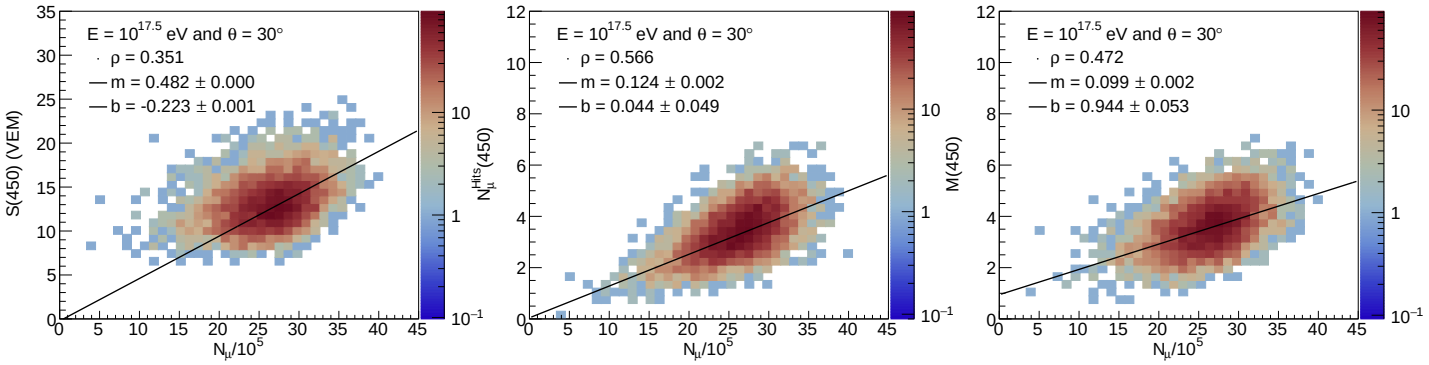

FIG. 8. Left panel: Correlation between the distributions of $S(450)$ and N_μ. Middle panel: Correlation between the distributions of $N_\mu^{\rm Hits}(450)$ and N_μ. Right panel: Correlation between the distributions of $M(450)$ and N_μ. Linear fits of the form $y = m(N_\mu/10^5)+b$ are shown in black. Fit parameters and correlation factors are displayed in the top left of each panel. Auger events were reconstructed with Offline from the ensemble of 5 105 CORSIKA simulations of EAS with $E_0 = 10^{17.5}$ eV and $\theta = 30°$, using EPOS-LHC.

convolution of an exponential with slope Λ_μ^{-1} and a Gaussian, as shown in Fig. 9. The true and reconstructed slopes Λ_μ and $\Lambda_\mu^{\rm rec}$ are displayed in the same figure.

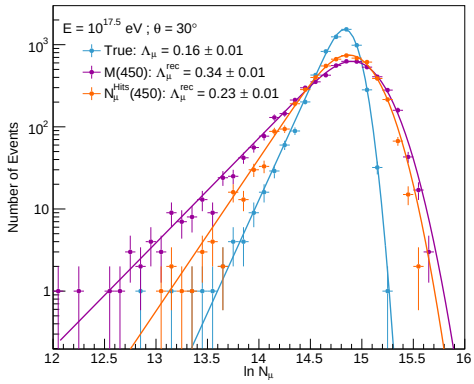

FIG. 9. Fits to the true (blue) and reconstructed distributions of $\ln N_\mu$ from the linear conversion of $M(450)$ (purple) and $N_\mu^{\rm Hits}(450)$ (orange) over an ensemble of 5 105 unsaturated proton induced events with primary energy $E_0 = 10^{17.5}$ eV and $\theta = 30°$. Showers were simulated with CORSIKA, using EPOS-LHC, and the Auger event simulated with Offline.

Given the large resolution induced by the reconstruction, the value of $\Lambda_\mu^{\rm rec} = 0.34 \pm 0.01$, obtained from $M(450)$ is incompatible with the true slope at 18σ. A positive bias of 5% in $\langle N_\mu \rangle$ renders it incompatible with the true value at $\sim 12\sigma$. Moreover, the reconstructed relative fluctuations of N_μ, $\sigma(N_\mu)/\langle N_\mu \rangle$, read 0.34, twice the value of the physical fluctuations. These results are not considerably improved by removing the electromagnetic contamination. Thus, the reconstruction of the distribution of $\ln N_\mu$ for showers with $E_0 = 10^{17.5}$ eV and $\theta = 30°$ using RPCs in the entire SD-750 array, does not allow for an unbiased estimation of Λ_μ, $\langle N_\mu \rangle$ nor of $\sigma(N_\mu)/\langle N_\mu \rangle$ due to a threshold effect. The impact of the electromagnetic contamination in these measurements appears to be sub-dominant.

The robustness of the reconstruction procedure was verified by increasing the shower energy and zenith angle. Thus, an ensemble of $\sim 1\,000$ proton initiated showers with $E_0 = 10^{18.5}$ eV and $\theta = 40°$ was simulated with CORSIKA v7.7410, using the post-LHC hadronic interaction model EPOS-LHC. Each event was simulated 5 times, with different core positions, to mitigate statistical fluctuations of the reconstructed distributions and obtain Auger events without saturated stations for each shower. The larger zenith angle was chosen to further mitigate the electromagnetic contamination, while keeping a reasonable effective RPC area. The higher primary energy was chosen to avoid the threshold effect noted in Fig. 8, without saturating the entire Infill array. In this way, the average number of candidate stations tripled for the 10-fold increase in energy and 10° in θ, better fixing the shape of the LDF.

The reconstruction procedure was integrally repeated. The obtained density plots correlating N_μ with $S(450)$, $M(450)$ or $N_\mu^{\rm Hits}(450)$ are shown in Fig. 10, along with a linear regression. Events with saturated stations or abnormally large $-\log \mathcal{L}_{\rm SD}$ were discarded.

The improvement in the correlation between N_μ and the different reconstruction variables is apparent. Furthermore, the more symmetrical shape of the density plots around the linear regression, reflects the reduction of the threshold effect. In particular, the correlation between $M(450)$ and N_μ is now relatively high with $\rho \simeq 0.69$, and the y-intercept of the linear regression is stays incompatible with zero. In the absence of electromagnetic contamination, the correlation is the strongest, although the positive y-intercept may indicate some bias in the determination of $N_\mu^{\rm Hits}(450)$.

Using $M(450)$, the resolution induced by the reconstruction of the tail of the distribution of $\ln N_\mu$ was improved, now reading 21%, while the overall reconstruction resolution reads, $\sigma_{\rm rec} = 17\%$. Furthermore, the reconstruction bias dropped to 3.5%. In the absence of electromagnetic contamination, the tail of the same distribution is reconstructed with a resolution of 17%. A remnant bias of 3.5% was observed, indicating that the reconstruction procedure could be further improved using algorithms dedicated to treat any residual threshold effects.

The true and reconstructed distributions of $\ln N_\mu$ were built for the ensemble of $\sim 3\,000$ unsaturated events and fitted as shown in the inset of Fig. 11. The values of $\Lambda_\mu^{\rm rec}$ for both reconstruction variables are closer to the true value of Λ_μ and to each other, despite the bias of 6σ. The reconstruction bias, lead to an overestimation of $\langle N_\mu \rangle$ by 3%. The relative fluctuations of N_μ read 24%, which is still larger than the physical fluctuations of 17%. The improvement by using directly the muon hits on the RPCs is not significant.

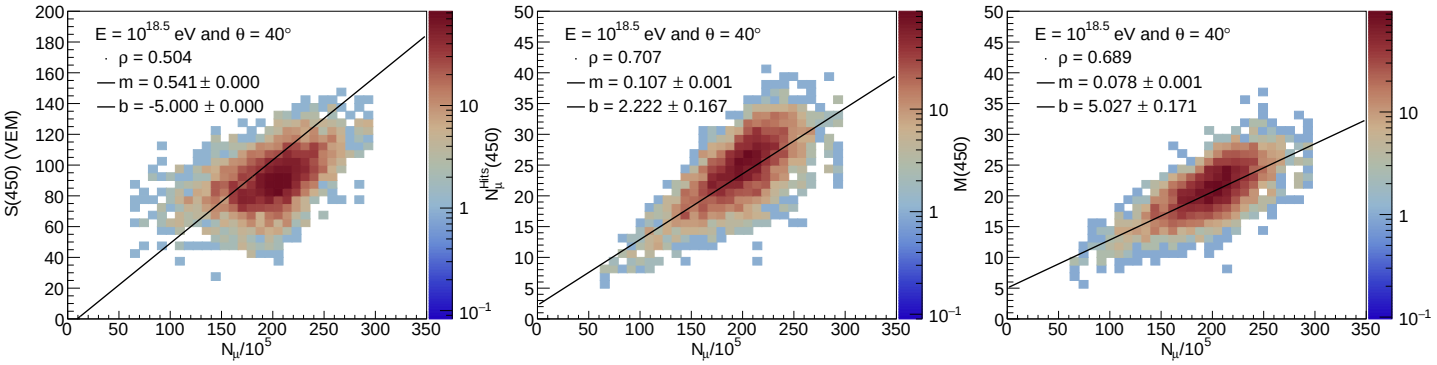

FIG. 10. Left panel: Correlation between the distributions of $S(450)$ and N_μ. Middle panel: Correlation between the distributions of $N_\mu^{\rm Hits}(450)$ and N_μ. Right panel: Correlation between the distributions of $M(450)$ and N_μ. Linear fits of the form $y = m(N_\mu/10^5)+b$ are shown in black. Fit parameters and correlation factors are displayed in the top left of each panel. Auger events were reconstructed with Offline from the ensemble of 5 105 CORSIKA simulations of EAS with $E_0 = 10^{18.5}$ eV and $\theta = 40°$, using EPOS-LHC.

E. Reconstruction of the muon number distribution with an engineering array of 7 MARTA stations

Having confirmed the robustness of the reconstruction prescription for showers with $E_0 = 10^{18.5}$ eV and $\theta = 40°$, we restrict the number of stations of the SD-750 with RPCs to 7 in a unitary hexagonal cell: one central station surrounded by its six closest neighbours. The shower core is simulated inside the engineering array. Using an ensemble of $\sim$ 1000 CORSIKA showers, where high energy interactions are simulated with EPOS-LHC, the reconstruction procedure was repeated. The reconstructed and true fitted distributions of $\ln N_\mu$ are shown in Fig. 11.

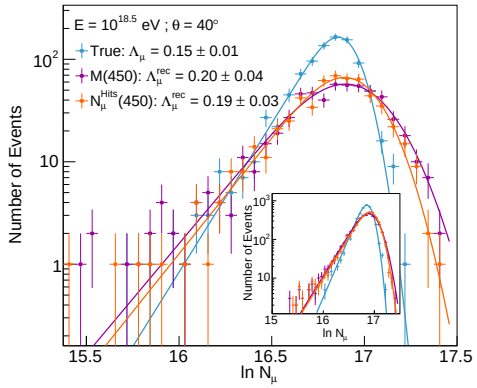

FIG. 11. Fits to the true (blue) and reconstructed distributions of $\ln N_\mu$ from the linear conversion of $M(450)$ (purple) and $N_\mu^{\rm Hits}(450)$ (orange) over an ensemble of $\sim$ 1 000 unsaturated proton induced events with primary energy $E_0 = 10^{18.5}$ eV and $\theta = 40°$, using 7 stations of the MARTA engineering array. Inset: Reconstructed distributions of $\ln N_\mu$ with the entire SD-750 array. Showers were simulated with CORSIKA, using EPOS-LHC, and the Auger event simulated with Offline.

The resolutions induced by the reconstruction using $M(450)$ and $N_\mu^{\rm Hits}(450)$ are $\sigma_{\rm rec} = 21\%$ and $\sigma_{\rm rec} = 18\%$. This means that the resolution worsens by $\sim 25\%$ with respect to the one obtained when using the full SD-750 array. However, the values of $\Lambda_\mu^{\rm rec}$, $\langle N_\mu \rangle$ and $\sigma(N_\mu)/\langle N_\mu \rangle$ achieved with the engineering array are very similar to the ones obtained with the full Infill array. This shows that the shape of the LDF is mostly fixed by stations of the engineering array, provided it contains the shower core, while the remaining stations play a sub-dominant role, tuning the resolution and bias of the reconstruction. The tail reconstruction is not significantly improved in the absence of electromagnetic contamination. We also verified that the quality of the reconstruction worsens by considering events with a core outside the cell of MARTA stations, since the fit to the LDF includes as little as 1 MARTA station.

IV. CONCLUSIONS

The low tail of the energy spectrum of hadronically interacting particles of the first p-Air interaction is characterised by a slope Λ_α which can be constrained by measurements of the slope, Λ_μ, of the muon number distribution in EAS with low muon content. Similarly, the forward region of the energy spectrum of neutral pions arising from the first p-Air interaction is characterised by a slope Λ_π, which is correlated with Λ_μ. Thus, measurements of Λ_μ in EAS with low muon content, constrain the forward production cross-section of neutral pions in ultra-high energy p-Air interactions. These results hold from ultra-high to LHC energies. Furthermore, Λ_μ can be accessed in mixed composition scenarios, since the dominant mass composition of showers with low muon content is proton. Even in extreme composition scenarios, Λ_μ can be measured within current experimental uncertainties, provided the number of events is large enough and the muon number distribution is corrected for its energy dependence.

From the expected number of active fiducial pads of the MARTA engineering array at $r = 450$ m from the shower axis, we were able to successfully reconstruct the muon number distribution for showers with $E_0 = 10^{18.5}$ eV and zenith angle $\theta = 40°$, under ideal conditions and despite the residual electromagnetic contamination. However, a threshold effect prevented completely unbiased measurements of Λ_μ, the average value of the muon content of EAS and of its relative fluctuations. We established that 7 stations of the MARTA engineering array in a minimal hexagonal configuration can reproduce the bias and resolutions achieved with the entire SD-750 array, provided the shower core falls within the engineering array. The bias and resolution of the reconstruction deteriorate with a 10-fold decrease in the primary energy, where the cosmic ray flux is larger. Given the area of the RPCs and their sensitivity to fluctuations of N_μ, more sophisticated analysis are necessary to mitigate the observed threshold effects. Lastly, combining the information of MARTA and WCD detectors could further improve our results and mitigate systematic uncertainties.

[1] A. Aab *et al.*, "Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18} ev," Science **357**, 1266–1270 (2017), https://science.sciencemag.org/content/357/6357/1266.full.pdf.

[2] A. Aab *et al.* (The Pierre Auger Collaboration), "Measurement of the cosmic-ray energy spectrum above 2.5×10^{18} eV using the Pierre Auger Observatory," Phys. Rev. D **102**, 062005 (2020).

[3] Kenneth Greisen, "End to the Cosmic-Ray Spectrum?" Phys. Rev. Lett. **16**, 748–750 (1966).

[4] G. T. Zatsepin and V. A. Kuzmin, "Upper limit of the spectrum of cosmic rays," JETP Lett. **4**, 78–80 (1966).

[5] Roberto Aloisio, "Acceleration and propagation of ultra-high energy cosmic rays," Progress of Theoretical and Experimental Physics **2017** (2017), 10.1093/ptep/ptx115, 12A102, https://academic.oup.com/ptep/article-pdf/2017/12/12A102/22075738/ptx115.pdf.

[6] Karl-Heinz Kampert and Michael Unger, "Measurements of the cosmic ray composition with air shower experiments," Astroparticle Physics **35**, 660–678 (2012).

[7] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, "CORSIKA: A Monte Carlo code to simulate extensive air showers," (1998).

[8] M. Alekseeva, T. Bergmann, V. Chernatkin, R. Engel, D. Heck, N. Kalmykov, S. Ostapchenko, T. Pierog, and K. Werner, "Extensive Air Shower Simulation Program CONEX: Matching Monte Carlo and Numerical Methods," in *29th International Cosmic Ray Conference (ICRC29), Volume 7*, International Cosmic Ray Conference, Vol. 7 (2005) p. 139.

[9] A. Aab *et al.*, "Direct measurement of the muonic content of extensive air showers between 2×10^{17} and 2×10^{18} eV at the Pierre Auger Observatory," The European Physical Journal C **80**, 751 (2020).

[10] A. Aab *et al.* (Pierre Auger Collaboration*), "Depth of maximum of air-shower profiles at the Pierre Auger Observatory. ii. Composition implications," Phys. Rev. D **90**, 122006 (2014).

[11] A. Aab *et al.* (Pierre Auger Collaboration), "Measurement of the Fluctuations in the Number of Muons in Extensive Air Showers with the Pierre Auger Observatory," Phys. Rev. Lett. **126**, 152002 (2021).

[12] Sidney Coleman and Sheldon L. Glashow, "High-energy tests of Lorentz invariance," Phys. Rev. D **59**, 116008 (1999).

[13] Luis A. Anchordoqui, Haim Goldberg, and Thomas J. Weiler, "Strange fireball as an explanation of the muon excess in Auger data," Phys. Rev. D **95**, 063005 (2017), arXiv:1612.07328 [hep-ph].

[14] A. Aab *et al.*, "The Pierre Auger Cosmic Ray Observatory," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **798**, 172–213 (2015).

[15] J. Abraham *et al.* (Pierre Auger), "Trigger and Aperture of the Surface Detector Array of the Pierre Auger Observatory," Nucl. Instrum. Meth. A **613**, 29–39 (2010), arXiv:1111.6764 [astro-ph.IM].

[16] Pedro Abreu *et al.* (Pierre Auger), "The energy spectrum of cosmic rays beyond the turn-down around 10^{17} eV as measured with the surface detector of the Pierre Auger Observatory," (2021), arXiv:2109.13400 [astro-ph.HE].

[17] Antonella Castellina (Pierre Auger), "AugerPrime: the Pierre Auger Observatory Upgrade," EPJ Web Conf. **210**, 06002 (2019), arXiv:1905.04472 [astro-ph.HE].

[18] P. Abreu *et al.*, "MARTA: a high-energy cosmic-ray detector concept for high-accuracy muon measurement," The European Physical Journal C **78**, 333 (2018).

[19] Lorenzo Cazon, Ruben Conceição, and Felix Riehn, "Probing the energy spectrum of hadrons in proton air interactions at ultrahigh energies through the fluctuations of the muon content of extensive air showers," Physics Letters B **784**, 68–76 (2018).

[20] J. Matthews, "A Heitler model of extensive air showers," Astroparticle Physics **22**, 387–397 (2005).

[21] Lorenzo Cazon, Ruben Concei ç ao, Miguel Alexandre Martins, and Felix Riehn, "Constraining the energy spectrum of neutral pions in ultra-high-energy proton-air interactions," Phys. Rev. D **103**, 022001 (2021).

[22] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider," Phys. Rev. C **92**, 034906 (2015).

[23] S. Ostapchenko, "Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model," Phys. Rev. D **83**, 014018 (2011).

[24] Felix Riehn, Ralph Engel, Anatoli Fedynitch, Thomas K. Gaisser, and Todor Stanev, "Hadronic interaction model Sibyll 2.3d and extensive air showers," Phys. Rev. D **102**, 063002 (2020), arXiv:1912.03300 [hep-ph].

[25] Alfredo Ferrari, Paola R. Sala, Alberto Fasso, and Johannes Ranft, "FLUKA: A multi-particle transport code (Program version 2005)," (2005), 10.2172/877507.

[26] T.T. Böhlen *et al.*, "The FLUKA Code: Developments and Challenges for High Energy and Medical Applications," Nuclear Data Sheets **120**, 211–214 (2014).

[27] Felix Riehn, "Measurement of the fluctuations in the number of muons in inclined air showers with the Pierre Auger Observatory," PoS **ICRC2019**, 404 (2019).

[28] A. Aab *et al.*, "Search for photons with energies above 1018eV using the hybrid detector of the Pierre Auger Observatory," Journal of Cosmology and Astroparticle Physics **2017**, 009–009 (2017).

[29] Lorenzo Cazon, Ruben Conceição, Miguel Martins, and Felix Riehn, "Probing the high energy spectrum of neutral pions in ultra-high energy proton-Air interactions," PoS **ICRC2019**, 226 (2020), arXiv:1908.09668 [hep-ph].

[30] Lorenzo Cazon, Ruben Conceição, and Felix Riehn, "An opportunity to measure the high-x region of π^0 production in UHE proton-Air interactions with Auger," (2019), auger internal note GAP-2019-001.

[31] O. Adriani *et al.* (LHCf Collaboration), "Measurements of longitudinal and transverse momentum distributions for neutral pions in the forward-rapidity region with the LHCf detector," Phys. Rev. D **94**, 032007 (2016).

[32] S. Argiro *et al.*, "The Offline Software Framework of the Pierre Auger Observatory," Nucl. Instrum. Meth. A **580**, 1485–1496 (2007), arXiv:0707.1652 [astro-ph].

[33] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506**, 250–303 (2003).

[34] A.D. Supanitsky, A. Etchegoyen, G. Medina-Tanco, I. Allekotte, M. Gómez Berisso, and M.C. Medina, "Underground muon counters as a tool for composition analyses," Astroparticle Physics **29**, 461–470 (2008).

[35] Koichi Kamata and Jun Nishimura, "The Lateral and the Angular Structure Functions of Electron Showers," Progress of Theoretical Physics Supplement **6**, 93–155 (1958), https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.6.93/5270594/6-93.pdf.

[36] Alexander Schulz, "Measurement of the Energy Spectrum of Cosmic Rays between 0.1 EeV and 30 EeV with the Infill Extension of the Surface Detector of the Pierre Auger Observatory," (2012), auger internal note GAP-2012-136.
[37] Daniela Mockler, *Measurement of the Cosmic Ray Spectrum with the Pierre Auger Observatory*, Ph.D. thesis, Karlsruher Institut für Technologie (KIT) (2019).
[38] Marco Aglietta, Ivan De Mitri, Silvio Maglio, Simone Maldera, Ioana C. Mariş, Daniele Martello, Gianni Navarra, and Markus Roth, "Recovery of Saturated Signals of the Surface Detector," (2008), auger internal note GAP-2008-030.