Mapping distributions of production variables of UHECR-air interactions onto the (N_μ, $X_{\max}$) space

Lorenzo Cazon,[a] **Ruben Conceição,**[b,c] **Miguel Alexandre Martins**[a,*] **and Felix Riehn**[d,e]

[a]*Instituto Galego de Física de Altas Enerxías (IGFAE),*
Rúa de Xoaquín Díaz de Rábago, s/n, Campus Vida, Universidade de Santiago de Compostela, 15705,
Santiago de Compostela, Galicia, Spain

[b]*Departamento de Física, Instituto Superior Técnico (IST), Universidade de Lisboa, Av. Rovisco Pais 1,*
1049-001 Lisbon, Portugal

[c]*Laboratório de Instrumentação e Física Experimental de Partículas (LIP) - Lisbon, Av. Prof. Gama Pinto*
2, 1649-003 Lisbon, Portugal

[d]*Technische Universität Dortmund, August-Schmidt-Straße 4, 44221 Dortmund, Germany*

[e]*Ruhr-Universität Bochum, Universitätsstraße 150, 44801 Bochum, Germany*

E-mail: miguelalexandre.jesusdasilva@usc.es

Ultra-high energy cosmic rays impinging on the Earth's atmosphere offer a unique opportunity to probe hadronic interactions at the highest energies through measurements of the extensive air showers they generate. In this work, we introduce a semi-analytic universal mapping between new macroscopic variables describing the first ultra-high-energy proton-air interaction and the shower-to-shower values of the depth of the shower maximum, $X_{\max}$, and the shower muon number. The distributions of these new production variables are highly dependent on the hadronic interaction model. These are combined into the primary interaction variables (α_1, ξ_1) which determine most of the features of the joint distribution of $X_{\max}$ and N_μ. We characterize this universal causal relation and show that it can be used to constrain the shape of the distributions of multi-particle production variables in proton-air interactions beyond current accelerator capabilities.

7th International Symposium on Ultra High Energy Cosmic Rays (UHECR2024)
17-21 November 2024
Malargüe, Mendoza, Argentina

---

*Speaker

1. Introduction

The interaction of cosmic rays with air nuclei at ultra-high energies produces hundreds of secondary hadrons in phase-space regions not covered by human-made accelerators. These secondaries further interact producing macroscopic cascades of particles, known as Extensive Air Showers, which partially retain the information on hadron production in the primary interaction.

Key observables of Extensive Air Showers are: the depth of the shower maximum, X_{max}, and the number of muons at the ground, N_μ. The connection between these observables and hadronic interactions is established using Monte Carlo simulations. These employ phenomenological hadronic interaction models that extrapolate from accelerator data and provide an inconsistent description of air-shower data [1, 2].

The proton-air cross-section can be estimated from the distribution of X_{max}, via the exponential fluctuations of the depth of the primary interaction point, X_1 [3, 4]. In turn, the moments of the distribution of N_μ contain information about the history of hadron production through $\langle N_\mu \rangle$ - underestimated by hadronic interaction models [1], the energy spectrum of hadrons in the primary interaction through $\sigma(N_\mu)$ [5, 6] and on the energy dependence of the inclusive cross section for the production of neutral pions in the primary interaction [7]. Finally, the latter cross-section can be probed as a function of the hadronic activity of the primary interaction, via a scan in X_{max} [8].

In this contribution, we develop a semi-analytic correspondence between the observables N_μ and X_{max} and a set of new multi-particle production variables of the primary interaction. These variables are motivated and described in Section 2. Their distributions are shown to be highly dependent on the hadronic interaction model and offer a natural description of different regions of the joint distribution of N_μ and X_{max}. These can be combined into two variables which are mapped onto X_{max} and N_μ. The precision and accuracy of this mapping are discussed in Section 3. Examples of the inverse map are presented in Section 4, showing the potential to constrain hadronic interactions at ultra-high energies using the joint distribution of N_μ and X_{max}.

2. New set of primary interaction variables

About 70% of the fluctuations of the relative muon content of EAS, $R_\mu = N_\mu / \langle N_\mu \rangle$, are determined by fluctuations of a variable of the primary interaction, α_1, defined by [5]:

$$\alpha_1 = \sum_{i=1}^{m_{had}} x_i^{0.93}, \tag{1}$$

where i runs over m_{had} hadronically interacting particles (therefore excluding $e^\pm$, γ, η and π^0), each carrying a fraction of the primary energy x_i, in the lab. frame. This variable is the combined expected muon yield for each hadronically interacting secondary of the first interaction, given in the Heitler-Matthews framework [9]. The distribution of α_1 is similar to the distribution of the fraction of energy in the hadronic sector of the first interaction: α_{had}.

Similarly, the shower-to-shower fluctuations of $\Delta X_{max} = X_{max} - X_1$, are connected to stochastic fluctuations of the energies carried by secondaries of the primary interaction. Secondary neutral pions decay into two photons, starting the electromagnetic cascade, while other hadrons further interact, producing more neutral pions that continuously feed the electromagnetic component of the

shower. Within the Heitler-Matthews framework, and for a primary proton with energy E_0, such estimator of $\Delta X_{\max}$ from the primary interaction reads

$$\xi_1 = \ln\left(\frac{E_0}{2\xi_c^e}\right) + \left[C_0 + (\omega - 1)\ln\left(\frac{E_0}{1\,\mathrm{PeV}}\right)\right]\alpha_{\mathrm{had}} + \omega\sum_{i=1}^{m_{\mathrm{had}}} x_i \ln x_i + \sum_{j=1}^{m_{\mathrm{EM}}} x_j \ln x_j, \qquad (2)$$

where $\xi_c^e \simeq 85\,\mathrm{MeV}$ is the electron critical energy [10], m_{EM} is the multiplicity of neutral pions and C_0 and ω are free parameters. These encode the evolution of the interaction cross-section and multiplicity with the projectile's energy in each interaction. Their values averaged over the hadronic interaction models, are $C_0 = -0.04^{+0.15}_{-0.30}$ and $\omega = 0.59^{+0.06}_{-0.05}$. These uncertainties include the dispersion between hadronic interaction models.

We can define the new multi-particle production variables

$$\zeta_{\mathrm{had}} \equiv -\sum_{i=1}^{m_{\mathrm{had}}} x_i \ln x_i \quad \text{and} \quad \zeta_{\mathrm{EM}} \equiv -\sum_{j=1}^{m_{\mathrm{EM}}} x_j \ln x_j. \qquad (3)$$

They characterize how evenly energy is distributed among secondaries of the primary interaction as a function of their multiplicity. Their values vanish for elastic primary interactions and are maximal for the equipartition of energy among secondaries in the respective sector. Their maximum value grows logarithmically with multiplicity. The distributions of these variables are discussed in Section 2.1 and are bounded energy conservation.

The correspondence between the first interaction variables α_1 and ξ_1, and the shower observables $(N_\mu, X_{\max})$ is studied with a set of 10^6 proton-induced Conex v7.50 [11, 12] simulations with primary energy $E_0 = 10^{19}\,\mathrm{eV}$ and zenith angle $\theta = 67°$, using the high-energy hadronic interaction model Epos-lhc [13]. Particles above $E_{\mathrm{th}} = 0.005 \times E_0$ and their interactions are tracked individually. Below this value, the longitudinal shower profile is obtained by numerically solving cascade equations. The ground level was set to $1\,400\,\mathrm{m}$ a.s.l, the average height of the Pierre Auger Observatory [14], corresponding to a vertical depth of $X_{\mathrm{gr}} = 880\,\mathrm{g\,cm^{-2}}$. The number of muons at this height and above $1\,\mathrm{GeV}$, N_μ is read from the muon profile. The value of $X_{\max}$ is taken from a Gaisser-Hillas fit to the longitudinal profile of all charged particles.

The correlation between α_1 and N_μ was studied in [5]. The correlation between $\lambda_r \xi_1$ and $\Delta X_{\max}$ is shown in the left panel of Figure 1. The distribution of residuals $R_X = \Delta X_{\max} - \lambda_r \xi_1$ is shown in the right panel of the same figure, along with the distributions of ξ_1 and $\Delta X_{\max}$, for the hadronic interaction models Epos-lhc, QGSjet -II.04 [15] and Sibyll2.3d [16]. The bias in the estimation of $\Delta X_{\max}$, $\langle R_X \rangle$, and the asymmetric resolution, $\sigma_\pm$, are also shown.

The Pearson correlation between $\lambda_r \xi_1$ and $X_{\max}$ indicates that over $50\,\%$ of $X_{\max}$ fluctuations are explained by $\lambda_r \xi_1$. The vertical line at $\sim 815\,\mathrm{g\,cm^{-2}}$ reflects highly elastic primary interactions with minimal energy transfer. In these cases, variability in $\Delta X_{\max}$ primarily stems from fluctuations in deeper shower stages. We found that primary-interaction fluctuations alone explain 65% of $\Delta X_{\max}$ variability, with $\lambda_r \xi_1$ capturing 80% of these fluctuations. Fluctuations in deeper shower interactions account for the remaining 35% of the variance of $X_{\max}$. Thus, the distribution of $\Delta X_{\max}$ is largely shaped by ξ_1, linking its shape to the primary interaction's secondary-particle spectra. This causal connection is largely independent of the hadronic interaction model.

The spread in biases in estimating $\Delta X_{\max}$ from ξ_1 is $\pm 5\,\mathrm{g\,cm^{-2}}$, and represents a systematic uncertainty due to the hadronic interaction model. The fast decoupling of the electromagnetic
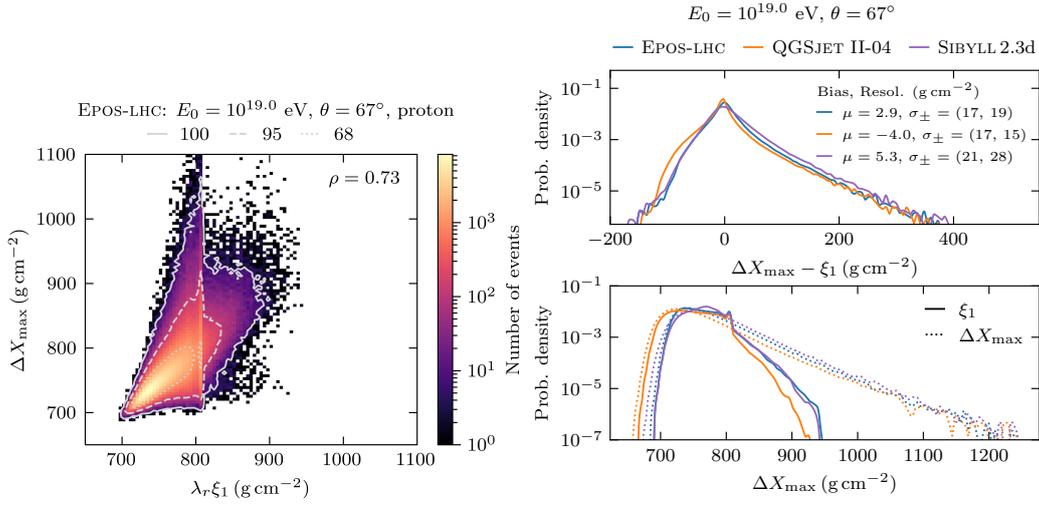
Figure 1: Left panel: Correlation between $\lambda_r \xi_1$ and $\Delta X_{\max}$ and their Pearson correlation coefficient ρ for Epos-lhc. Contours encompassing 68% (dotted), 95% (dashed) and 99.7% (solid) of the events are represented in white. Right upper panel: Distribution of the residuals $\Delta X_{\max} - \lambda_r \xi_1$, their bias, μ, and asymmetric resolution, $\sigma_\pm$. Right lower panel: Distributions of $\Delta X_{\max}$ and $\lambda_r \xi_1$. Figures on the right panels refer to simulations produced with the hadronic interaction models Epos-lhc, QGSjet -II.04 and Sibyll2.3d.

cascade ensures the primary interaction sets the $\Delta X_{\max}$ scale, captured by $\langle \xi_1 \rangle$ [17]. The right tail of $\Delta X_{\max}$ arises from highly elastic primary interactions, where the leading particle traverses multiple interaction lengths before re-interacting. These lengths follow an exponential distribution, and their fluctuations are independent of the stochastic energy partitioning in the primary interaction. As a result, ξ_1's predictive power in the diffractive limit decreases and the right resolution $\sigma_+ \sim 21\,\mathrm{g\,cm^{-2}}$ exceeds the left resolution $\sigma_- \sim 18\,\mathrm{g\,cm^{-2}}$. Nonetheless, a reduced hadronic-model dependence of the probabilistic shower response to ξ_1, $p(\Delta X_{\max}, |, \xi_1)$, would still permit the estimation of the distribution of ξ, $p(\xi_1)$ from that of $\Delta X_{\max}$. We are currently quantifying the hadronic-model dependence of such a response.

2.1 Marginal distributions

The information in the spectra of secondaries of the primary interaction that is mapped onto N_μ and $X_{\max}$ is condensed in the multi-particle production variables α_{had}, ζ_{had} and ζ_{EM}. The distributions of these variables are shown in Figure 2 for three hadronic interaction models. These distributions are highly dependent on the hadronic interaction models. Due to their sensitivity to differences in hadron production mechanisms, their data-driven probing would constrain hadronic interactions at the highest energies. The peaks at $\zeta_{\mathrm{had}} = \zeta_{\mathrm{EM}} = 0$ correspond to diffractive/elastic primary interactions, while their bulks correspond to high hadronic (ζ_{had}) and electromagnetic (ζ_{EM}) activities.

Since these new production variables are produced from the energy spectra of secondaries, they can be readily measured in collider experiments, within the phase space covered by the specific detectors. A detailed comparison of the distributions of these production variables across hadronic interaction models and rapidity regions is being conducted.
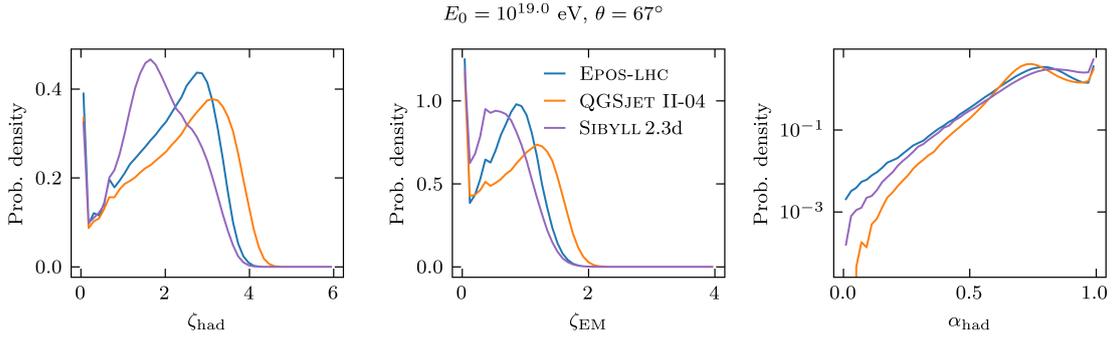
$E_0 = 10^{19.0}$ eV, $\theta = 67°$

Figure 2: Distributions of ζ_{had} (left), ζ_{EM} (middle) and α_{had} (right) for 10^6 proton-air interactions simulated with CONEX at $E_0 = 10^{19}$ eV and $\theta = 67°$, using the hadronic interaction models EPOS-LHC, QGSJET -II.04 and SIBYLL2.3d.

2.2 Variables on the N_μ-X_{max} plane

Figure 3 shows the variables ζ_{had}, ζ_{EM} and α_1 on the (R_μ, X_{max}) plane.

EPOS-LHC: $E_0 = 10^{19.0}$ eV, $\theta = 67°$
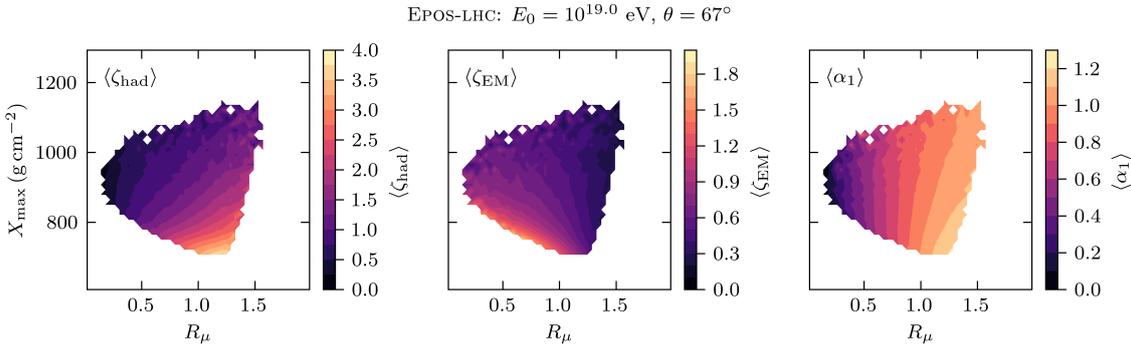
Figure 3: Values of ζ_{had} (left), ζ_{EM} (middle) and α_1 (right) over the joint distribution of (R_μ, X_{max}), for 10^6 proton-air interactions simulated with CONEX at $E_0 = 10^{19}$ eV and $\theta = 67°$, using the hadronic interaction models EPOS-LHC.

The nearly vertical gradient of ζ_{had} indicates that shallower, muon-enriched showers stem from primary interactions with higher hadronic activity. The gradient of ζ_{EM} indicates the shallowest showers for each value of R_μ exhibit high electromagnetic activity. Deeper showers correspond to diffractive events (high α_1) or interactions producing few highly energetic neutral pions (low α_1). Regardless of X_{max}, muon-enriched showers correspond to high α_1 (high hadronic activity), while muon-depleted showers correspond to low α_1.

Thus, the multi-particle production variables show clear gradients in the (R_μ, X_{max}) plane, directly linking the primary interaction to distinct regions of the joint distribution of these shower observables. This connection is not broken by the uncorrelated fluctuations of X_1.

3. Direct mapping

For pairs of values (α_1, ξ_1), the probabilistic response of the shower $p(R_\mu, X_{\max} \mid \alpha_1, \xi_1)$, is such that

$$p(R_\mu, X_{\max}) = \int p(R_\mu, X_{\max} \mid \alpha_1, \xi_1) p(\alpha_1, \xi_1) \, d\alpha_1 \, d\xi_1 \,. \tag{4}$$

The response function $p(R_\mu, X_{\max} \mid \alpha_1, \xi_1)$ also includes the fluctuations of X_1 and the effects of muon attenuation and muon profile truncation by the ground level. That is, effects beyond the production physics that connects the first interaction variables (α_1, ξ_1) with the shower observables.

Examples of the response function $p(R_\mu, X_{\max} \mid \alpha_1, \xi_1)$ for particular values of α_1 and ξ_1 can be found in Figure 4. This figure was produced using the simulation library defined in Section 2. The light-grey lines are the contours of the relative number of produced muons.
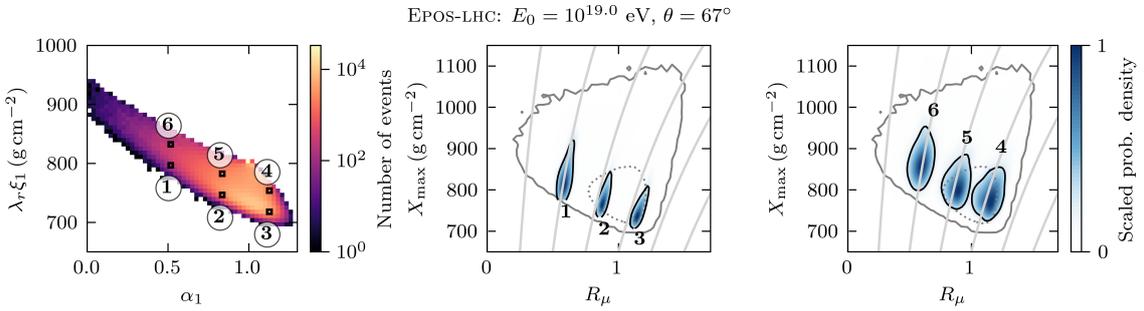

Figure 4: Left panel: joint distribution $p(\alpha_1, \lambda_r \xi_1)$ with highlighted bins at 6 positions (α_1^k, ξ_1^k). Middle and right panels: direct shower response function $p(R_\mu, X_{\max} \mid \alpha_1, \xi_1)$ for the each value of (α_1^k, ξ_1^k). The 68% contour is represented by solid black lines. Light grey lines are the contours of a constant number of produced muons. Figure produced using the simulation library described in Section 2.

The high degree of correlation between α_1 and R_μ, and between ξ_1 and $\Delta X_{\max}$, guarantees that the response of the shower to each (α_1, ξ_1) is narrow. Moreover, the numbered points in the α_1-ξ_1 space keep their relative positions on the R_μ-$X_{\max}$ space. The slanted smear of the response function is due to the fluctuations of X_1 and the increased muon attenuation flux in shallower showers. The overall variance in (α_1, ξ_1) corresponds to 75% of the two-dimensional variance in $(R_\mu, X_{\max})$, establishing the sub-dominate role of the rest of the shower in shaping $p(R_\mu, X_{\max})$. These observations hold qualitatively regardless of the hadronic interaction models.

Since the shape of $p(\alpha_1, \xi_1)$ is partially preserved in $p(R_\mu, X_{\max})$, we are currently assessing the hadronic-model dependence of the shower response. Provided this dependence is reduced, the measured distribution of $(R_\mu, X_{\max})$ can be used to estimate the shape of $p(\alpha_1, \xi_1)$, ultimately constraining the energy spectra of secondaries of the primary interaction, in kinematic regimes untenable in current human-made colliders.

4. Inverse mapping

The distribution $p(\alpha_1, \xi_1)$ could be estimated by folding in the response $p(R_\mu, X_{\max} \mid \xi_1, \alpha_1)$ together with the response of the detector to fit a measured distribution of $p(R_\mu, X_{\max})$. However, looking directly at the inverse mapping $p(\alpha_1, \xi_1 \mid R_\mu, X_{\max})$ is instructive.

Examples of the inverse response function $p(\alpha_1, \xi_1 \mid R_\mu, X_{max})$ for 6 bins in the joint distribution $p(R_\mu, X_{max})$ can be found in Figure 5.
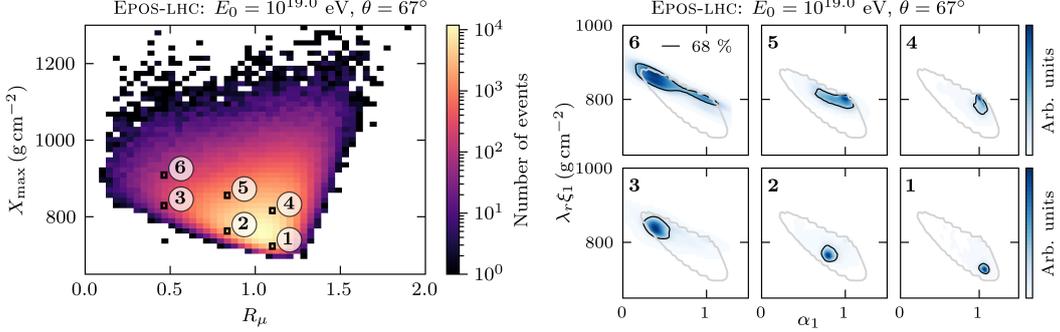

Figure 5: Left panel: joint distribution $p(R_\mu, X_{max})$ with highlighted bins at 6 positions (R_μ^k, X_{max}^k). Right panels: inverse shower response function $p(\alpha_1, \lambda_r \xi_1 \mid R_\mu, X_{max})$ for the each value of (R_μ^k, X_{max}^k). The 68% contour is represented by solid black lines. Light grey lines are the contours of a constant number of produced muons. Figure produced using the simulation library described in Section 2.

Disjoint regions on the plane R_μ-X_{max} are mapped onto distributions in (α_1, ξ_1) with small overlapping and low dispersion compared to the phase-space covered by $p(\alpha_1, \xi_1)$. Therefore, there is a precise correspondence between the joint distribution of R_μ and X_{max} in the energy spectra of secondaries of the primary interaction, which can be exploited. The dependence of this mapping on the hadronic interaction model is under investigation.

5. Conclusions

We introduced new multiparticle production variables, ζ_{had}, ζ_{EM}, and α_{had}, derived from the spectra of secondary particles in the primary interaction. Combined into the estimator ξ_1, these variables capture 80% of the stochastic fluctuations in hadron production in the primary interaction and explain over 50% of ΔX_{max} fluctuations. The variable ξ_1 provides a robust estimate of the X_{max} scale with a model dependence of only $5 \, \mathrm{g \, cm^{-2}}$, independent of X_1. The causal relation between ξ_1 and X_{max} is largely model-independent.

The distributions of these variables depend strongly on the hadronic interaction model, making them sensitive probes of hadron production mechanisms. Notably, these variables can be measured with current accelerator data within available phase space.

By combining ξ_1 with α_1—the estimator of the relative number of muons, N_μ—we build a probabilistic shower response for (N_μ, X_{max}): $p(R_\mu, X_{max} \mid \alpha_1, \xi_1)$. This response has low dispersion, with fluctuations of (α_1, ξ_1) accounting for 75% of the variance in (N_μ, X_{max}). Thus the effects of the fluctuations from deeper generations, fluctuations of X_1, and X_{max}-dependent muon attenuation are subdominant.

Currently, we are quantifying the hadronic-interaction-model dependence of this response. If found minimal, the distribution of (N_μ, X_{max}) could constrain the energy spectra of secondaries produced in ultra-high-energy hadronic interactions in kinematic regions beyond current accelerator capabilities.

References

[1] J. Albrecht et al., *The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider*, *Astrophys. Space Sci.* **367** (2022) 27 [2105.06148].

[2] PIERRE AUGER collaboration, *Testing hadronic-model predictions of depth of maximum of air-shower profiles and ground-particle signals using hybrid data of the Pierre Auger Observatory*, *Phys. Rev. D* **109** (2024) 102001 [2401.10740].

[3] THE PIERRE AUGER COLLABORATION collaboration, *Measurement of the Proton-Air Cross Section at $\sqrt{s}$=57 TeV with the Pierre Auger Observatory*, *Phys. Rev. Lett.* **109** (2012) 062002.

[4] A. Abdul Halim et al., *Studies of the mass composition of cosmic rays and proton-proton interaction cross-sections at ultra-high energies with the Pierre Auger Observatory*, in *Proceedings of 38th International Cosmic Ray Conference — PoS(ICRC2023)*, vol. 444, p. 438, 2023, DOI.

[5] L. Cazon, R. Conceição and F. 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*, *Phys. Lett. B* **784** (2018) 68 [1803.05699].

[6] 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** (2021) 152002 [2102.07797].

[7] L. Cazon, R. Conceição, M.A. Martins and F. Riehn, *Constraining the energy spectrum of neutral pions in ultra-high-energy proton-air interactions*, *Phys. Rev. D* **103** (2021) 022001 [2006.11303].

[8] L. Cazon, R. Conceição, M.A. Martins and F. Riehn, *Proton-air interactions at ultra-high energies in muon-depleted air showers with different depths*, *Phys. Lett. B* **859** (2024) 139115 [2406.08620].

[9] J. Matthews, *A Heitler model of extensive air showers*, *Astropart. Phys.* **22** (2005) 387.

[10] T.K. Gaisser, R. Engel and E. Resconi, *Cosmic Rays and Particle Physics: 2nd Edition*, Cambridge University Press (6, 2016).

[11] T. Bergmann, R. Engel, D. Heck, N. Kalmykov, S. Ostapchenko, T. Pierog et al., *One-dimensional hybrid approach to extensive air shower simulation*, *Astroparticle Physics* **26** (2007) 420.

[12] T. Pierog et al., *First results of fast one-dimensional hybrid simulation of EAS using CONEX*, *Nucl. Phys. B Proc. Suppl.* **151** (2004) 159 [astro-ph/0411260].

[13] T. Pierog, I. 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** (2015) 034906 [1306.0121].

[14] PIERRE AUGER COLLABORATION collaboration, *The Pierre Auger Cosmic Ray Observatory*, *Nucl. Instrum. Meth. A* **798** (2015) 172 [1502.01323].

[15] S. Ostapchenko, *Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model*, *Phys. Rev. D* **83** (2011) 014018 [1010.1869].

[16] F. Riehn, R. Engel, A. Fedynitch, T.K. Gaisser and T. Stanev, *Hadronic interaction model SIBYLL 2.3d and extensive air showers*, *Phys. Rev. D* **102** (2020) 063002.

[17] S. Ostapchenko and M. Bleicher, *Constraining pion interactions at very high energies by cosmic ray data*, *Phys. Rev. D* **93** (2016) 051501 [1601.06567].