PROCEEDINGS OF SCIENCE

Prospects of a composition-enhanced search for large-scale anisotropy with the Pierre Auger Observatory

Edyvania Emily Martins[a,b,*] **for the Pierre Auger Collaboration**

[a] *Karlsruhe Institute of Technology, Institute for Experimental Particle Physics, Karlsruhe, Germany*
[b] *Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), Buenos Aires, Argentina*
[c] *Observatorio Pierre Auger, Av. San Martín Norte 304, 5613 Malargüe, Argentina*
Full author list: https://www.auger.org/archive/authors_2024_11.html

E-mail: spokespersons@auger.org

The Pierre Auger Observatory has been collecting data for over 19 years, reaching more than 135 000 km^2 yr sr of accumulated exposure, with the surface detectors spread over 3000 km^2. A remarkable discovery is the large-scale dipole structure with a total amplitude of 7.4% for energies above 8 EeV. The observed modulation in right ascension has a statistical significance of 6.8σ. The dipolar pattern in the events with energies between 8 and 16 EeV has a statistical significance of over 5σ. The Pierre Auger Collaboration has also reported an increase in the dipole amplitude with energy. This anisotropy is understood to be of extragalactic origin, as the maximum of the dipolar component is located $\sim 115°$ away from the Galactic Center. In the same energy range, the observed evolution of the depth of maximum shower development with energy indicates a progression towards heavier composition of cosmic rays with increasing energy. This contribution presents a novel approach to a search for composition-enhanced large-scale anisotropy. On the one hand, lighter events have higher rigidity than their heavier counterparts with similar energy; therefore, their trajectories are less affected by magnetic fields. The expected effect is a higher anisotropy in the arrival direction of a subset of events with smaller mass and charge than the anisotropy in the overall flux of cosmic rays. On the other hand, the attenuation length is distinct for each mass group, leading to different horizon of cosmic ray sources for each of them. Under a source-agnostic model, we investigate the dipole amplitude as a function of rigidity. Using a simulation library, we analyze the possibility of measuring a separation in total dipole amplitude between two sub-populations distinct in mass of the Auger Phase I dataset.

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

[*] Speaker

ISSN 1824-8039. Published by SISSA Medialab. https://pos.sissa.it/

PoS(UHECR2024)102

1. Introduction

The Pierre Auger Observatory in Malargüe, Argentina, is the largest facility for detecting cosmic rays (CR). Over 1600 water Cherenkov detectors form the surface detector (SD) covering an area larger than 3000 km^2. Such a large detection area counteracts the steeply falling rate of events with increasing energy. In the ultra-high-energy regime, above 1×10^{18} eV, the rate of events is roughly one cosmic ray per squared kilometer per year. The SD operates nearly 100 % of the time, overlooked by four sites where fluorescence telescopes are housed. The 27 telescopes operate on clear, moonless nights, corresponding to a duty cycle of nearly 15 %. While the observation of air showers using the fluorescence detector (FD) allows for the direct measurement of the atmospheric depth where the maximal size of a shower is reached, $X_{\max}$, its reduced operation cycle limits the available statistics. Interested in the mass composition of the flux of ultra-high-energy cosmic rays (UHECR) the Auger Collaboration has continuously improved the inference of $X_{\max}$, a quantity closely related to the CR atomic mass A, via measurements with the SD.

The high-quality data collected in the Pierre Auger Observatory resulted in the discovery [1] of a dipole structure in the arrival direction of UHECR. This anisotropy pattern requires a large number of events to significantly distinguish this dipolar structure from an isotropic background. After a decade of data collection, the observed events with energy above 8 EeV allowed for the discovery and precise description of this dipole, which has an amplitude of 7.4% and points $\sim$115° away from our Galaxy center. With over 19 years of data-taking, the established dipolar anisotropy in right ascension has now reached a significance of 6.8σ when considering all events with energies above 8 EeV [2, 3]. The increased statistics also allow for the evaluation of the large-scale anisotropy in finer energy ranges, leading to the observation of a similar dipole pattern with a statistical significance of 5.7σ when only considering energies between 8 and 16 EeV. Albeit the steeply decreasing number of events at higher energy ranges does not yet allow for a description of the dipole pattern with comparable significance, a trend of increasing amplitude of the dipole anisotropy has been verified [4]. Additionally, the direction of this dipolar structure consistently points away from our Galaxy center in all evaluated energy ranges above 4 EeV. The current interpretation is that the most energetic cosmic rays originate outside our Galaxy.

In parallel to the evaluation of the CR arrival direction distribution, the Collaboration has continuously reported the moments of the $X_{\max}$ distribution, a key ingredient in interpreting the mass composition of UHECR. This composition is commonly disentangled into contributions from four or five representative elements (see Refs. [5–7]). The observed moments of the $X_{\max}$ distribution indicate a mixed and progressively heavier composition of UHECRs.

Considering the increase of the dipole amplitude and the changing, mixed composition of UHECR with energy, a method is developed to search for the signature of composition dependency on the dipole structure. Acknowledging the extragalactic origin of the observed anisotropy, one can expect composition-related effects in the production, acceleration, and/or propagation of UHECR to affect the distribution of arrival directions. Under a source-agnostic model, we investigate the dipole amplitude as a function of rigidity. Using a simulation library, we analyze the possibility of measuring a separation in total dipole amplitude between two A-distinct sub-populations of the Auger Phase I dataset.

2. Simulation library

We aim to investigate signatures of a mass dependence of the observed large-scale anisotropy by dividing the dataset into two composition-distinct subpopulations. Taking a conservative approach, we define the subpopulations according to simulation expectations rather than a data-driven approach. In doing so, we can avoid an *a posteriori* penalization of the results due to excessive scanning of the parameters, which is particularly important considering the small amplitude of the measured large-scale anisotropy.

The simulation library adopted mimics the characteristics of the measured data, such as the spectrum and the mean composition. While the spectral features are well-defined [8, 9], the precise contribution of different nuclei to the overall flux is more uncertain. The composition of CRs and its evolution with energy are particularly important once a composition-informed selection of events is aimed at. To that end, we introduce a four-component continuous, empirical model of the mass composition of CRs based on Auger data [6], but extended to higher energies. The *extended Auger mix*, the composition model adopted in this analysis, also incorporates a mass-ordering of the fractional contributions to the overall CR flux. It is composed of four nuclei components, which are approximately equally spaced in $\ln A$, namely proton, indicated by p (or H), helium (He), nitrogen (N), and iron (Fe).

The library of simulated events contains the primary cosmic ray information and the expected detector response. We employ the EPOS-LHC hadronic interaction model [10] in CORSIKA-simulated events [11] that lead to the predicted detector response and consequent measurable quantities via the $\overline{\text{Off}}\underline{\text{line}}$ framework [12]. Higher level quantities, such as $X_{\max}$ and the relative-to-proton-shower muon number, R_μ, were obtained in this analysis with the shower universality method [13, 14]. In Fig. 1, we can observe the spectrum and composition distributions adopted in the simulation library, with each colored line indicating representative nuclei. In Fig. 2, we evaluate the good compatibility between the number of measured (red) and simulated (black) events. In both figures, the gray-shaded regions indicate the energy ranges of interest for the large-scale anisotropy analysis: 4 to 8 EeV, 8 to 16 EeV, 16 to 32 EeV, and above 32 EeV.

3. A model of composition dependency of the dipole amplitude

In an approach consistent to the definition of the mass composition model, we aim to minimize the number of assumptions while reproducing well the measured dipole amplitude. Instead of probing the source distribution and considering the propagation of CR towards Earth, which requires the evaluation of the effects of magnetic fields and interactions with radiation, we opt for a semi-empirical approach.

As reported in [2, 4], the energy dependence of the dipole amplitude is $d(E) = d_{10}\ (E/\,10\,\text{EeV}))^{\beta}$, with $d_{10} = 0.049 \pm 0.009$ and $\beta = 0.97 \pm 0.21$. Inspired by it, the parameterization of the dependence of the total dipole amplitude d on rigidity $R = E/Ze$, is defined by the following relation,

$$d(E, Z) = d_R \left(\frac{E/\text{EeV}}{Ze}\right)^{\beta_R} \qquad \text{or} \qquad d(R) = d_R (R/\text{EV})^{\beta_R}, \tag{1}$$

where Z is the nuclear charge (Z = 1, 2, 8, 26, for proton, helium, oxygen, and iron nuclei, respectively), and where d_R and β_R are the two model parameters. In this model, we additionally

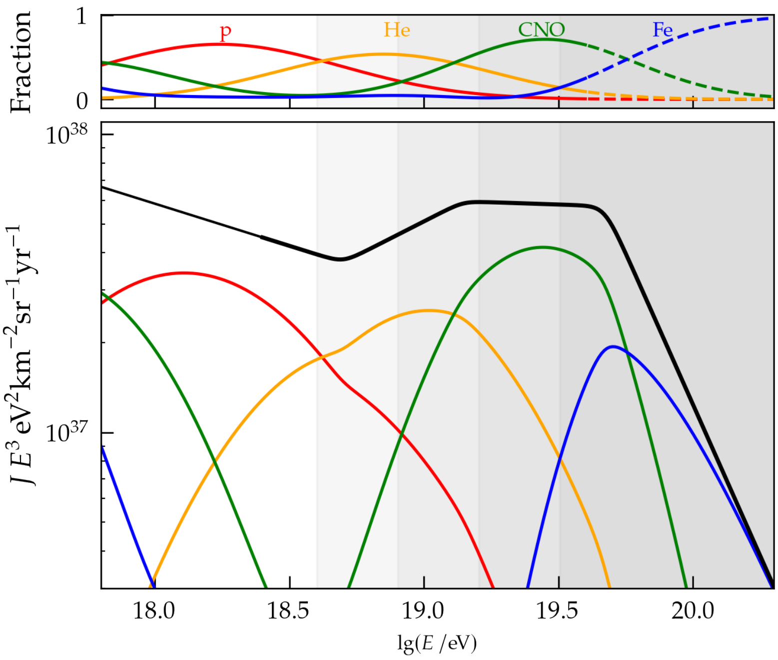

Figure 1: Spectrum and contributions from each of the four components considered in the extended Auger mix. In the top panel, the fractional contributions of four representative CR primaries as a function of energy are given. The dashed lines indicate the region of extrapolation of the measurements, i.e., where the low statistic in the hybrid measurements prevents a composition estimate. In the bottom panel, the fitted spectrum and the contribution from each of the four components, color-coded as in the upper panel, are shown.

limit the dipole amplitude to a maximal upper-value $d_{\max}$, so that

$$d(R) = \begin{cases} d_R (R/\mathrm{EV})^{\beta_R} & ;\text{ if } d(R) < d_{\max}, \\ d_{\max} & ;\text{ otherwise.} \end{cases} \tag{2}$$

The introduction of the parameter $d_{\max}$ is related to the origin of the dipole structure. In the scenario where a single source is responsible for the anisotropy $d_{\max} = 3$, while in the case of many sources contributing to the overall anisotropy $d_{\max} \simeq 1$ (see Ref. [15]).

The parameters d_R and β_R can be estimated so that the model, under the described composition scenario, agrees with the measured total dipole amplitude. For the parameter estimation, we consider the three energy bins with $E > 8$ EeV since the dipole component in the energy bin between 4 and 8 EeV is not significant despite the large number of events. Additionally, the dipole amplitude in that energy range could be affected by the contribution from a low-energy source population needed to explain the flux below the ankle energy E_{ankle}, $\lg(E_{\text{ankle}}/\mathrm{eV}) \approx 18.7$. We compute a goodness-of-fit to evaluate how well the parameterizations reproduced the three data points. From the map of reduced chi-squared, χ^2, on the left side of Fig. 3, we conclude that the parameters that allow for the best compatibility with measurements are $d_R = 0.0018$ and $\beta_R = 2.1$, adopted hereafter.

Combining the composition model (Fig. 1) with the rigidity-dependent dipole amplitude depicted on the left side of Fig. 3, we obtain the total dipole amplitude shown on the right side of that figure. The colored lines are given by multiplying the component fraction for that energy by

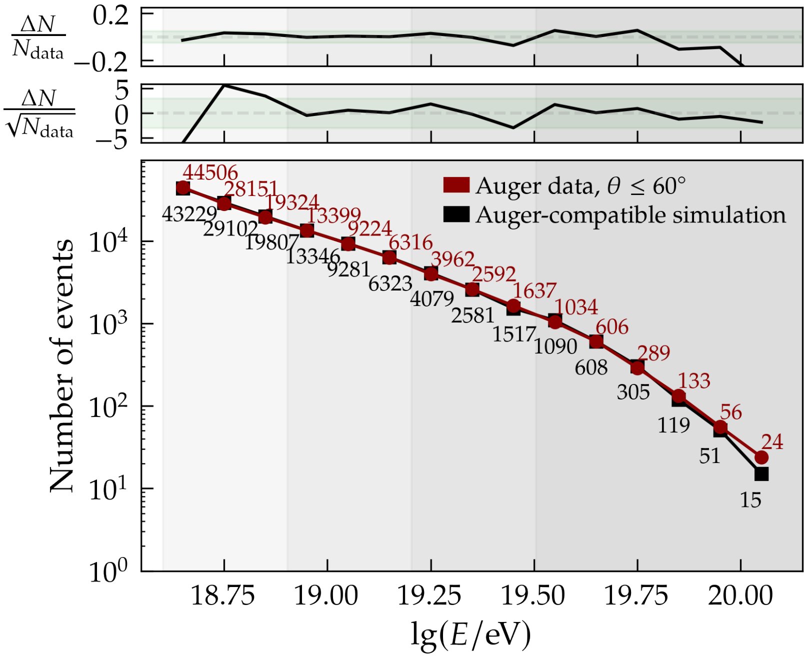

Figure 2: Number of events in Auger-compatible simulation library and in the measured dataset of events with zenith angle $\theta \leq 60°$. For the simulation library, E stands for the simulated energy, while for the Auger dataset, E represents the measured energy. Gray bands indicate the large energy bins used in the large-scale anisotropy searches. The top panels indicate the difference in the number of events of the measured and simulated data sets relative to the number of measured events (*top*) or to the square root of that number (*mid*). The green bands indicate a 5% (*top*) or 3σ (*mid*) deviation.

its corresponding dipole amplitude. Adding the contributions from each component results in the expected total dipole amplitude, shown in black. In the measurements, however, we obtain the dipole amplitude as a single observation in each energy bin. To translate the dipole amplitude model, a continuous function of energy, into the equivalent binned value, we have to consider the spectrum of the CR flux. We thus calculate the weighted average for the dipole amplitude in each energy bin, represented here by the gray horizontal. The weights are computed by the cumulative distribution function of the spectrum function, defined in Ref. [8].

With the model introduced above that reproduces the dipole amplitudes in the three highest energy bins well and the Auger-compatible simulation library described in the previous section, we can probe the composition signature on the large-scale anisotropy. The separation of the two sub-populations, **light** and **heavy**, will be defined in the next section through a scan in the Universality-reconstructed mass estimator $\ln A$. The mass estimator is obtained as a function of X_{max} and R_μ . While the λ parameter "weights" how much the X_{max} from a nucleus-initiated shower deviates from the proton expectations, the β parameter does so for the R_μ quantity [16],

$$-\lambda \ln A = X_{\mathrm{max}} - X^{\mathrm{p}}_{\mathrm{max}} \,, \qquad \text{and} \qquad \beta \ln A = \ln R_\mu - \ln R^{\mathrm{p}}_\mu \,. \tag{3}$$

By evaluating the simulation library, the parameters λ and β could also be estimated presenting a

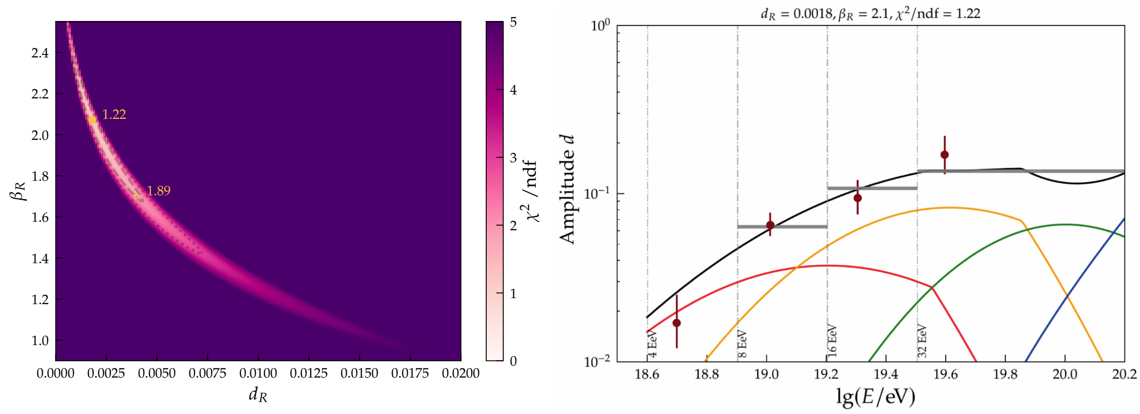

Figure 3: *Left*: maps of reduced χ^2 for the possible parameters in Eq. (1). The χ^2 is a goodness-of-fit evaluation of these parameters to describe the measured dipole amplitudes from Ref. [4] in each energy bin. The expected dipole amplitude in each energy bin is computed via Eq. (1), and the contribution from each component follows the composition fraction as described in the EPOS extended Auger mix, weighted by the spectrum function. The maximum dipole amplitude considered here is $d_{\max} = 3$. Contours denote points with $\chi^2/\mathrm{ndf} = 2$ (dashed line) and with $\chi^2/\mathrm{ndf} = 3$ (dotted line). A full circle denotes the best parameters for the EPOS extended Auger mix composition, while the empty star denotes an edge point, falling just within the $\chi^2/\mathrm{ndf} = 2$ contour region with $d_R = 0.004$ and $\beta_R = 1.7$. *Right*: contribution to the dipole amplitude by each component: proton (red), helium (orange), CNO (green), and iron (blue). Summing all contributions, the total dipole amplitude is shown in black. Horizontal gray bars correspond to the spectrum-weighted total dipole amplitude in each large energy bin. Dark red markers represent the measured dipole amplitude as in Ref. [4]. The parameters that allow for the best description of the data, denoted with a circle in the left-side graph, are used in this realization and correspond to $d_R = 0.0018$ and $\beta_R = 2.1$.

small energy dependence, as also found in Ref. [16].

$$\lambda = 16.45 + 0.38 \times \lg(E/10^{19}\mathrm{eV}) \ , \tag{4}$$
$$\beta = 0.013 + 0.006 \times \lg(E/10^{19}\mathrm{eV}) \ . \tag{5}$$

In the next section, we compute the expected separation between those populations in terms of their dipole amplitude by assuming that the dipole direction is the same in all subsets of data. Although an oversimplification, this prevents us from making assumptions about the (classes of) sources of UHECR.

4. Analysis method and prospective results

The two subsets of interest are defined to maximize the difference in their dipole amplitudes. All events with $\ln A$ below a threshold value $\ln A^{\mathrm{thr}}_{\mathrm{light}}$ constitute the *light population*, and all the events with $\ln A$ above another threshold value $\ln A^{\mathrm{thr}}_{\mathrm{heavy}}$ constitute the *heavy population*. No event can pertain to both populations, while some events do not pertain to either. This also means $\ln A^{\mathrm{thr}}_{\mathrm{heavy}} > \ln A^{\mathrm{thr}}_{\mathrm{light}}$.

To quantify the separation, we evaluate the standardized mean difference, SMD, between the expected $d(R)$ for each population. The SMD is computed with

$$SMD = \frac{|d_{\text{light}} - d_{\text{heavy}}|}{\sqrt{\sigma^2_{\text{light}} + \sigma^2_{\text{heavy}}}}, \tag{6}$$

with $\sigma_{\text{light/heavy}}$ representing the uncertainty of such amplitude. We use the statistical uncertainty as a proxy for the uncertainty in $d(R)$, computed as $\sigma_{\text{pop.}} = \sqrt{2/N_{\text{pop.}}}$. This approach differs by a nearly-constant factor from the quantity in the standard analysis [2, 4], which does not affect the maximization of the SMD. Again, for each pair of possible threshold $\ln A$ values, the composition of the defined subpopulation is used to compute the corresponding expected dipole amplitude, and then the SMD is evaluated. The best set of threshold values for this approach in a given energy bin is the pair of values that maximizes this quantity.

Once the light and heavy populations are defined for each energy, we can evaluate the composition therein. Then, making use of Eq. (1) we can compute the amplitude in each population and evaluate how distinct they are via Eq. (6). In Fig. 4, we present the expected dipole amplitude for the light and heavy populations. Also shown are the measurements of the total dipole amplitude reported in [4]. Additionally, the horizontal lines indicate the expected dipole amplitude for a single-component CR beam for proton (red) or iron (blue). According to the prescription described here, we can expect a separation in terms of the dipole amplitude for two A-distinct populations with the Auger Phase I data. The SMD ranges between 2 and 3 in all evaluated energy bins.

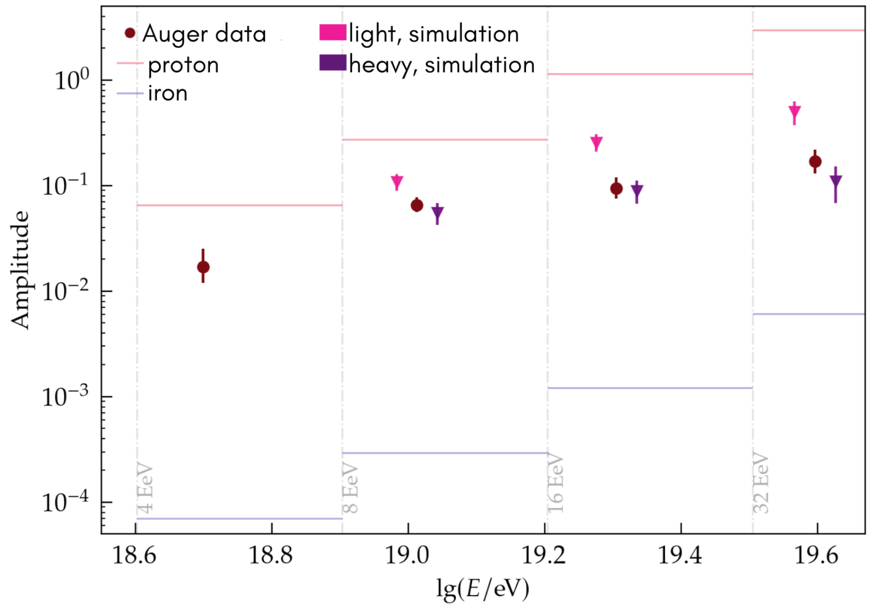

Figure 4: Expected dipole amplitude for two mass-distinct populations. The light (pink) and heavy (purple) populations were defined according to the threshold values that maximize the difference in terms of the dipole amplitude. The latest measurements [2] of the dipole amplitudes are shown in dark red circles. For reference, the expected dipole amplitude for a single-component CR beam is shown for proton (red) and iron (blue).

5. Summary

The proposed approach to describe the rigidity dependence of the dipole amplitude, under the assumption of the extended Auger mix composition model, allowed us to evaluate the discovery potential of a composition-informed anisotropy with Auger Phase I data. A separation in terms of the dipole amplitude in all energy bins for two A-distinct populations can be expected. Since this approach does not rely on a scan of the parameters on measured data, no subsequent penalization is needed. We remark that different directions of the dipole in each population, not evaluated here, can contribute to the separation of the dipolar anisotropies of two A-distinct subsets of data. Other composition-informed analyses on the anisotropy of UHECR are under development, as reported in Ref. [17].

[1] The Pierre Auger Collaboration *Science* **357** (2017) 1266.

[2] The Pierre Auger Collaboration *The Astrophysical Journal* **976** (2024) 48.

[3] The Pierre Auger Collaboration *PoS* **UHECR2024** (2024) 008.

[4] The Pierre Auger Collaboration *PoS* **ICRC2023** (2023) 252.

[5] B. Bortolato et al. *Physical Review D* **108** (2023) 022004.

[6] The Pierre Auger Collaboration *PoS* **ICRC2023** (2023) 438.

[7] The Pierre Auger Collaboration *JCAP* **2024** (2024) 022.

[8] The Pierre Auger Collaboration *Physical Review D* **102** (2020) 062005.

[9] The Pierre Auger Collaboration *PoS* **ICRC2021** (2021) 324.

[10] T. Pierog et al. *Phys. Rev. C* **92** (2015) 034906.

[11] D. Heck et al. *Forschungszentrum Karlsruhe Report FZKA* **6019** (1998) .

[12] S. Argirò et al. *Nucl. Instrum. Meth. A* **580** (2007) 1485.

[13] M. Stadelmaier et al. *Phys. Rev. D* **110** (2024) 023030.

[14] The Pierre Auger Collaboration *PoS* **UHECR2024** (2024) 117.

[15] D. Harari et al. *Physical Review D* **92** (2015) 063014.

[16] M. Stadelmaier , *On air-shower universality and the mass composition of ultra-high-energy cosmic rays*, Ph.D. thesis, KIT and UNSAM, 2022.

[17] The Pierre Auger Collaboration *PoS* **UHECR2024** (2024) 068.

PoS(UHECR2024)102