

## Observational constraints on cosmic-ray escape from ultra-high energy accelerators

---

**Antonio Condorelli,<sup>a,\*</sup> Quentin Luce,<sup>b</sup> Sullivan Marafico,<sup>a</sup> Jonathan Biteau<sup>a</sup> and Olivier Deligny<sup>a</sup>**

<sup>a</sup>*Université Paris-Saclay, CNRS/IN2P3, IJCLab, Orsay, France,*

<sup>b</sup>*Karlsruhe Institute of Technology, Institute for Experimental Particle Physics (ETP), Karlsruhe, Germany*

*E-mail:* [condorelli@ijclab.in2p3.fr](mailto:condorelli@ijclab.in2p3.fr)

The energy spectrum and mass composition of ultra-high energy cosmic rays inferred at the Pierre Auger Observatory are used to derive a benchmark scenario for the emission mechanisms at play in extragalactic accelerators as well as for their energetics and for the abundances of elements in their environments. Assuming a distribution of sources following the density of stellar mass, the gradual increase of the cosmic ray mass number observed on Earth from 2 EeV up to the highest energies is shown to call for nuclei accelerated up to an energy proportional to their electric charge and emitted with a hard spectral index. In addition, the inferred flux of protons down to 0.6 EeV is shown to require for this population a spectral index significantly softer than that of heavier nuclei. This is consistent with in-source interactions that shape the energy production rate of injected charged nuclei differently from that of the secondary neutrons escaping from the confinement zone. Together with the inferred abundances of nuclei, these results provide constraints on the radiation levels in the source environments. Within this scenario, an additional component that falls off steeply with increasing energy up to the ankle feature is necessary to make up the all-particle flux in the sub-ankle energy range.

\*\*\* 27th European Cosmic Ray Symposium - ECRS \*\*\*

\*\*\* 25-29 July 2022 \*\*\*

\*\*\* Nijmegen, the Netherlands \*\*\*

---

\*Speaker

## 1. Introduction

The origin of ultra-high energy cosmic rays (UHECRs) is currently unknown. Their identification is based mostly on catching a potentially interesting class of astronomical objects in the UHECR arrival directions. An observation remains elusive, although new investigations have allowed for broad statements to be made. Above  $\approx 8$  EeV [1], an anisotropy at large scales has been discovered, the intensity and direction of which are consistent with assumptions drawn from sources distributed in a similar way to extragalactic matter [1, 2]. A correlation between UHECR arrival directions and the flux distributions of massive, star-forming, or active galaxies within 200 Mpc provides evidence for anisotropy at higher energies ( $\gtrsim 40$  EeV) [3, 4]. Overall, these results suggest that UHECRs are primarily of extragalactic origin, at least above the so-called ankle energy ( $\approx 5$  EeV).

The observed energy spectrum and chemical composition of UHECRs on Earth are the result of emission processes that include acceleration mechanisms, losses and escape from source environments, as well as propagation effects. Differently from, and complementing to anisotropies, these two observables provide constraints, helping to infer the features of the acceleration processes, the energetics of the sources, and the abundances of elements in the source environments. Following this idea, various investigations have set the groundwork for a generic scenario that broadly replicates the observations. [5–7].

The intensity of the different nuclear components at the sources, generally assumed as standard candles and uniformly distributed in a comoving volume, is supposed to drop off at the same magnetic rigidity to explain the steady increase with energy of mass number  $A$  observed on Earth [8, 9]. This agrees to the underlying assumption that electromagnetic processes accelerate particles to a maximum energy proportional to their electric charge  $Z$ . The abundance of nuclear elements is observed to be dominated by intermediate-mass ones, ranging from He to Si, accelerated to  $E_{\max}^Z \approx 5Z$  EeV and leaving from the source environments with a very hard spectral index  $\gamma$ .

The constraints on the characteristics of the sources are actually discussed in [10]. To accomplish so, data from the Pierre Auger Observatory beyond 0.63 EeV ( $10^{17.8}$  eV) were taken into account. However, in contrast to other approaches, only the proton spectrum is used in the energy range between 0.63 EeV and 5 EeV. This method allows us to reconstruct the extra-galactic component without introducing any ad hoc nuclear components at the lowest energies, to represent what is generally believed to be the upper end of the galactic component, in the combined fit of both energy spectrum and mass composition. Because of both observational systematics and theoretical uncertainties, such modeling alters the reconstruction of the extragalactic proton component with biases intrinsic in the choices made.

## 2. Generic model of UHECR production with in-source interactions

The main assumptions of the benchmark astrophysical model employed here can be found in [6]. The non-thermal processes responsible for accelerating the different types of particles are modeled using power-law spectra proportional to the charge  $Z$ , such as  $E_{\max}^Z = ZE_{\max}$ , while an exponential suppression is used to describe the end of the acceleration process in the absence of any strong indication from theory. The sources are believed to accelerate diverse proportions of

nuclei, which are represented by five stable ones: hydrogen ( $^1\text{H}$ ), helium ( $^4\text{He}$ ), nitrogen ( $^{14}\text{N}$ ), silicon ( $^{28}\text{Si}$ ), and iron ( $^{56}\text{Fe}$ ). The ejection rate per comoving volume unit and per energy unit of nucleons is usually modeled as

$$q_p(E) = q_{0p} \left( \frac{E}{E_0} \right)^{-\gamma_p} f_{\text{supp}}(E, Z_p), \quad (1)$$

Using  $Z_p = 1$ , which is a single ejection rate  $q_{0p}$ , spectral index  $\gamma_p$ , and suppression function  $f_{\text{supp}}(E, Z_p)$  for both escaping protons and neutron decay protons. In this case,  $E_0$  is arbitrarily fixed to 1 EeV. Generically, the ejection rate of nuclei with mass number  $A_i$  is likewise modeled as

$$q_{A_i}(E) = q_{0A_i} \left( \frac{E}{E_0} \right)^{-\gamma_{A_i}} f_{\text{supp}}(E, Z_{A_i}), \quad (2)$$

For helium, nitrogen, silicon, and iron, there is a single spectral index  $\gamma_A$  and four separate reference ejection rates  $q_{0A_i}$ . The suppression function used for nucleons and nuclei is the same as in the reference case of [6].

$$f_{\text{supp}}(E, Z) = \begin{cases} 1 & \text{if } E \leq E_{\text{max}}^Z, \\ \exp(1 - E/E_{\text{max}}^Z) & \text{otherwise.} \end{cases} \quad (3)$$

The maximum acceleration energy is expected to be proportional to the electric charge of each element,  $E_{\text{max}}^Z = ZE_{\text{max}}$ , with the five species having a single free parameter  $E_{\text{max}}$ .

The ejection rate for protons in this approach accounts for both the accelerated ones up to  $E_{\text{max}}$  and those produced by escaping neutrons, resulting in an energy of the order of  $E/A$  inherited from the nuclei of energy  $E$ . The maximal energy reached by this population of secondary protons is then  $E_{\text{max}}^Z/A \simeq E_{\text{max}}/2$ , which is not represented in the rigidity-acceleration scheme modeled by Equation (3). Replacing  $E_{\text{max}}$  with  $E_{\text{max}}/2$  in  $q_p(E)$  would test the extreme situation in which all ejected protons are photodissociation by-products. Modeling of sources and their environments is required for further characterization of the balance between the population of accelerated protons from the initial abundance in the source environment and that of secondaries from nuclear cascades.

For each species  $j$ , the differential energy production rate per comoving volume unit of the sources, which is directly connected to their differential luminosity, is consequently  $\ell_j(E, z) = E^2 q_j(E) S(z)$ , where  $S(z)$  reflects the redshift evolution of the UHECR luminosity density. For convenience, the quantity  $\ell_j(E, z)$  is referred to as the differential energy production rate. The bolometric energy production rate per comoving volume unit at redshift  $z$ , on the other hand, is calculated as  $\mathcal{L}_j(E, z) = S(z) \int_E^\infty dE' E' q_j(E')$ . We report its average value in a volume spanning  $z_{\text{min}} - z_{\text{max}}$  as  $\bar{\mathcal{L}}_j(E) = \int_{z_{\text{min}}}^{z_{\text{max}}} dz \left| \frac{dt}{dz} \right| \mathcal{L}_j(E, z) / \int_{z_{\text{min}}}^{z_{\text{max}}} dz \left| \frac{dt}{dz} \right|$ , where  $t(z)$  is the lookback time.

The density of baryonic matter is used to trace the evolution of the UHECR luminosity density throughout cosmic time. This is supposed to follow star mass density, which is well approximated by a constant out to redshift  $z = 1$  [e.g. 11]. As in the benchmark case of [6], such a constant evolution is expected to hold to a first approximation of  $z_{\text{max}} = 2.5$ . The local Universe, on the other hand, has an overdensity because the Milky Way, like most galaxies, is part of a group of galaxies which is embedded in the Local Sheet [12]. We use the overdensity correction factor calculated by [13], which is defined as a function of distance  $r$  and is effective up to 30 Mpc,

$$\frac{\delta\rho(r)}{\bar{\rho}} = 1 + \left( \frac{r}{r_0} \right)^{-\alpha}, \quad (4)$$

with  $r_0 = 5.4 \text{ Mpc}$  and  $\alpha = 1.66$ . The evolution of the UHECR luminosity density is then represented by the equation  $S(z) = \delta\rho/\bar{\rho}$ , with  $z_{\text{max}} = 2.5$  and  $z_{\text{min}} = 2 \times 10^{-4}$ . The minimal redshift, which corresponds to the Local Group limit ( $r \simeq 1 \text{ Mpc}$ ), precludes any divergence in Equation (4) and effectively removes very-closeby galaxies that would otherwise dominate the UHECR sky, which contradicts observations [4].

In this scenario, we are neglecting the effect of extra-galactic magnetic field on particle propagation, therefore this can thus be termed one-dimensional.

The current all-particle energy spectrum  $J(E)$  derives from the integration of all sources' contributions across lookback time, with redshift playing a role:

$$J(E) = \frac{c}{4\pi} \sum_{A,A'} \iint dz dE' \left| \frac{dt}{dz} \right| S(z) q_{A'}(E') \frac{d\eta_{AA'}(E, E', z)}{dE}. \quad (5)$$

The relationship between cosmic time and redshift is derived from the cosmological concordance model,  $(dt/dz)^{-1} = -H_0(1+z)\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$ , where  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  is the Hubble constant at present time,  $\Omega_m \simeq 0.3$  is the density of matter (baryonic and dark matter) and  $\Omega_\Lambda \simeq 0.7$  is the dark-energy density;  $\eta_{AA'}(E, E', z)$  is the amount of particles detected on Earth with energy  $E$  and mass number  $A$  released by sources with energies  $E' > E$  and mass numbers  $A' > A$  and describes the energy losses and spallation processes. In practice, for a given redshift  $z_0$  source emitting a nuclear species  $A_0$  at energy  $E_0$ , the corresponding  $\eta_{AA'}(E, E', z)$  function is tabulated in bins of  $A'$  and  $E'$  by propagating a large number of emitted particles ( $O(10^7)$  particles) by using the SimProp package [14]. By repeating the simulations for different  $z_0$ ,  $A_0$ , and  $E_0$  values, the entire  $\eta_{AA'}(E, E', z)$  function is tabulated as a 5D histogram, returning the fractions sought. SimProp accounts for pair production, photo-pion production, and photodissociation off the photon fields of interest, which are those from cosmic-microwave-background (CMB) radiation and infrared photons from extragalactic background light (EBL), which is the radiation produced in the Universe since the formation of the first stars. The CMB radiation can be described as a black-body spectrum with a redshift-dependent temperature  $T(z) = T_0(1+z)$ , with  $T_0 = 2.725 \text{ K}$ . The EBL is less well known, particularly in the far infrared and at high redshifts. For the benchmark scenario explored below, we utilise the model of [15], which is consistent with the minimal intensity level derived from galaxy counts as well as indirect measurements derived from gamma-ray absorption at multi-TeV energies [e.g. 16, for a recent review]. Equation (5) can be used, with these ingredients, to evaluate the all-particle flux from the various contributions of each individual nuclear component on the condition of assigning values to the five ejection rates  $q_{0A_i}$ , the two indices  $\gamma_p$  and  $\gamma_A$ , and the maximum energy  $E_{\text{max}}$ . These eight parameters are fitted to the data applying the method described in the following section.

### 3. Combined fit to energy-spectrum and mass-composition data

The expected spectrum represented by Equation (5) is dependent on various unknown features that characterize acceleration processes, source surroundings, and source energetics. Constraints from the measured energy spectrum and mass composition can contribute in inferring these properties. We will use, on the one hand, the all-particle energy spectrum inferred from these data [18], and on the other, the distributions of the slant depth of maximum of shower development

(hereinafter called  $X_{\max}$ ), which is the best up-to-date proxy of the primary mass of the particles [9, 19]. Using the hadronic-interaction generators to represent the development of the showers, the  $X_{\max}$  distributions allow for statistical inference of the energy-dependent mass composition. Here, two hadronic-interaction generators are considered: EPOS-LHC [20] for the benchmark scenario and Sibyll2.3c [21] as an alternative, both of which are the most recent and best describe the data. As a result, by combining the all-particle energy spectrum and the abundances of the various mass components  $J_j(E)$ , the different elemental spectra as a function of energy can be calculated. The energy threshold here taken into account is the nominal one of the  $X_{\max}$  detection mode at ultra-high energies, namely  $\approx 0.63$  EeV ( $10^{17.8}$  eV). This threshold, incidentally, allows us to investigate the energy spectrum of the ankle feature. A Gaussian likelihood fit is used to establish the best match between the observed spectrum and that expected from Equation (5) from a set of proposed parameters  $\Theta$ . Assuming that the transition to extragalactic UHECRs has already occurred above 5 EeV,<sup>1</sup> the equivalent probability term,  $L_J$ , is calculated above that threshold. The model is fitted to the  $X_{\max}$  data using the multinomial distribution, which represents how likely it is to observe  $k_{mx}$  events out of  $n_m$  with probability  $p_{mx}$  in each  $X_{\max}$  bin in each energy bin  $m$ . The probabilities  $p_{mx}$  are calculated by modeling the  $X_{\max}$  distributions in terms of the parameters  $\Theta$  using generators of hadronic interactions. The corresponding likelihood,  $L_{X_{\max}}$ , is considered as the product of energy bins greater than 5 EeV. The last contribution to the likelihood comes from the sub-ankle proton component,  $L_{J_p}$ , which is calculated by weighting the all-particle spectrum below 5 EeV with the proton abundance  $f_p(E)$  from the  $X_{\max}$  distributions. The statistical uncertainties in  $\hat{J}_p(E)$  are dominated by those in  $f_p(E)$ , and they are accounted for using a Gaussian likelihood fit.

As a result, the model likelihood is provided by  $L = L_J L_{X_{\max}} L_{J_p}$ . The goodness-of-fit is measured using a deviance,  $D$ , which is defined as the negative log-likelihood ratio between a given model and the saturated model that represents the data:

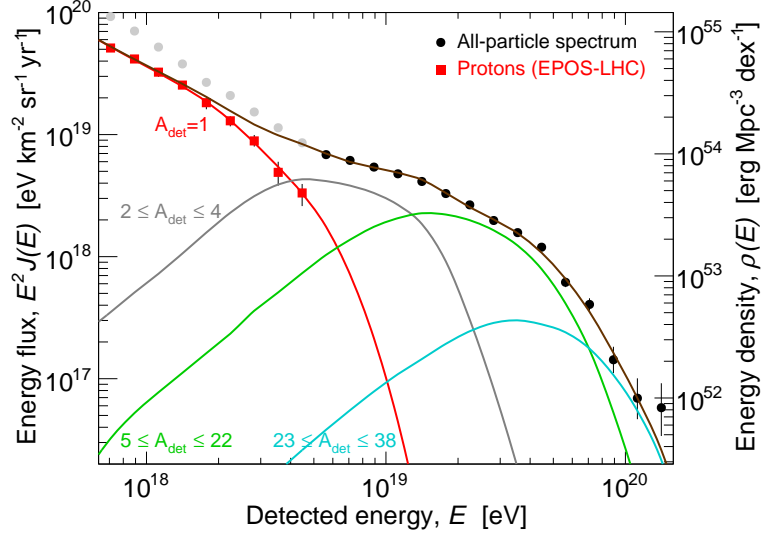
$$D = -2 \log \frac{L_J}{L_J^{\text{sat}}} - 2 \log \frac{L_{X_{\max}}}{L_{X_{\max}}^{\text{sat}}} - 2 \log \frac{L_{J_p}}{L_{J_p}^{\text{sat}}}. \quad (6)$$

The three different contributions are referred to as  $D_J$ ,  $D_{X_{\max}}$  and  $D_{J_p}$  respectively.

## 4. Results

Fitting the model to the data using the benchmark scenario provided in Section 2, with free spectral indices for both nucleons and nuclei ( $\gamma_p \neq \gamma_A$ ), provides the parameters and deviance shown in Table 1 of [10] using EPOS-LHC to interpret  $X_{\max}$  data. Table 1 of [10] includes the results predicted under the assumptions of a proton component dominated by neutron escape (proton maximum energy of  $E_{\max}/2$ ), no local overdensity (widely used uniform distribution), and a shared spectral index across the five species ( $\gamma_p = \gamma_A$ ). The value of  $E_{\max}$ , derived by the drop in the nuclear components at  $E_{\max}^Z$ , indicates that the suppression of the spectrum is due to a combination of the cut-off energy at the sources for the heavier nuclei and energy losses *en route*, as found in [1]. The nuclei's spectral index,  $\gamma_A$ , is defined by the increase in average mass with energy, which is virtually monoelemental, in order to mimic the  $X_{\max}$  distributions as accurately as possible.

<sup>1</sup>We evaluated that a moderate increase in the value at which the transition to extragalactic UHECRs is considered to be complete would not have any effect on the final results.

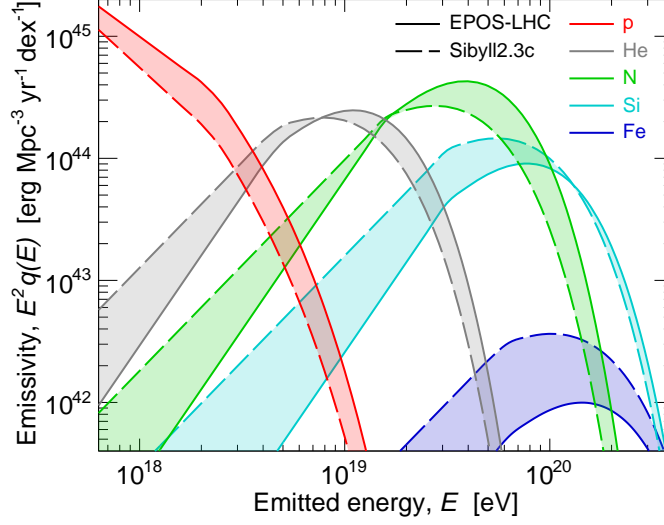


**Figure 1:** Energy flux at Earth as a function of energy, as modeled by the best-fit parameters for the benchmark scenario. The all-particle spectrum and proton component are shown as black circles and red squares, respectively [18]. Lighter points are not included in the fit. The best-fit components obtained for five detected mass groups are displayed with solid colored lines, as labeled in the Figure. The energy flux of the heaviest mass group, with detected mass number in 39–56, is below the range of interest.

The fit's solution thus consists on setting a hard index for nuclei so that the contributions of each element mix as little as possible: high-energy suppression imposed by the cut-off beyond  $E_{\max}^Z$  and low-energy suppression via the hard index  $\gamma_A$ . This effect, however, does not apply to protons, which persist in an energy range where a mixing of elements is required. The best-fit value of  $\gamma_p$  is significantly softer than that of  $\gamma$ . The addition of  $\gamma_A$  improves the fit of the data down to 0.63 EeV, with a total deviance  $D = 236.8$  compared to 862.7 in the case of  $\gamma_p = \gamma_A$ , therefore the addition of this extra free parameter is sufficiently justified. On the other hand, using local overdensity to trace the source distribution improves the deviation significantly while it has a minor impact on the best-fit parameters.

The balance determining the intensity of each component is reported in terms of energy production rates  $\tilde{\mathcal{L}}$  greater than 0.63 EeV. The result is displayed in Figure 1, which shows the contributions of each nuclear component to the observed energy flux and energy density. The energy production rates necessary at the sources to power the observed energy flux are shown as a function of energy in Figure 2 for the various primary mass groups. The solution is illustrated in Figure 1, While the contribution of nuclei is highest at  $\approx 5 \times 10^{44}$  erg Mpc $^{-3}$  yr $^{-1}$  dex $^{-1}$ , that of protons is increasing up to  $\approx 10^{45}$  erg Mpc $^{-3}$  yr $^{-1}$  dex $^{-1}$  when going down in energy. Extrapolations of the results below 0.63 EeV, on the other hand, are unsafe since the functional form employed in Equation (1) may no longer hold depending on the features of the source environments that determine the nuclear cascade. The total energy production rate, integrating above 0.63 EeV, is found to be  $(10.8 \pm 0.4) \times 10^{44}$  erg Mpc $^{-3}$  yr $^{-1}$ .

The reduced deviance is divided into three terms, according to Equation (6). The fit of the spectra leads to an acceptable fit ( $D_J + D_{J_p} = 37.3$  for  $N_J + N_{J_p} = 24$  points). As in [1], the



**Figure 2:** Mass-dependent energy production rate at the sources as a function of energy, as constrained by the best-fit parameters for the benchmark scenario. The dashed lines illustrate a variation of the hadronic interaction model, with EPOS-LHC and Sibyll2.3c shown as solid and dashed lines, respectively.

$X_{\max}$  sector has a worse deviance (value  $D_{X_{\max}} = 199.5$  for  $N_{X_{\max}} = 109$  points). In general, the benchmark scenario has a fit quality quite similar to those obtained by considering only data beyond 5 EeV. Considering the variation of Equation (1) that consists in substituting  $E_{\max}$  for  $E_{\max}/2$ , it was found that the main changes concern the spectral index of protons, which becomes  $\gamma_p \simeq 2.5$ , while the deviance does not change between the two scenarios.

## 5. Discussion

The results help to explain the origin of protons below the ankle energy and the hard spectra of nuclei above this energy. The component of protons in this scenario is extragalactic in origin and exponentially suppressed above the ankle energy, while heavier nuclei slowly take over to the highest energies via a rigidity-dependent maximum-energy scenario. A softer spectral index for this population of light primaries reflects the fact that protons are not suppressed when they lose energy at the same rate as heavier nuclei. Remarkably, such behavior qualitatively corresponds to scenarios of in-source interactions in which abundant fluxes of neutrons originated by accelerated charged particles interacting with the bath of photons permeating the sources escape unimpeded from the electromagnetic fields. Our conclusions provide a solid base for constraining the spectral indices  $\gamma_p$  and  $\gamma_A$ , as well as the energetics of the sources and the abundances of elements in their surroundings. There are two reasons for the constraints' robustness. First, by taking into account only the proton spectrum between 0.63 EeV and 5 EeV, we can rebuild the extragalactic component as directly and accurately as possible, with no "interference" from the expected composition of the supposed end of the galactic component. Second, by adding a local overdensity to the widely used uniform distribution of sources in a comoving volume, we can match the prediction with the data on a statistical basis characterized by  $D/\text{ndf} \simeq 237/125$  i.e. a reduced deviance on the order of that found in [6] for data strictly above the ankle. The study presented here significantly enhances the

description of the data, particularly in the mass-composition sector. Future observations will thus provide evidence to either confirm the role of in-source interactions in the interpretation of UHECR data, by detecting such a sub-dominant component below the ankle and its nature, either galactic or extra-galactic, with the upgraded instrumentation of the Pierre Auger Observatory in the next years [22].

## References

- [1] Aab, A. & Pierre Auger Collaboration , *Science* **357** (2017), 1266
- [2] Aab, A.& Pierre Auger Collaboration , *ApJ* **868** (2018), 4
- [3] Aab, A.& Pierre Auger Collaboration , *ApJ* **L29** (2018), 853
- [4] Abreu, P. & Pierre Auger Collaboration, *ApJ* **935** (2022), 170
- [5] Aloisio, R., Berezhinsky, V. & Blasi, P. , *JCAP* **10** (2014), 020
- [6] Aab, A. & Pierre Auger Collaboration , **4** pp. 038 (2017), [Erratum: *JCAP* 03, E02 (2018)]
- [7] Guido, E. & Pierre Auger Collaboration , *PoS. ICRC2021* **311** (2021).
- [8] Abraham, J. et al., *Phys. Rev. Lett.* **104** (2010), pp. 091101
- [9] Aab, A. et al. , *Phys. Rev. D* **90** (2014), 122006
- [10] Luce, Q., Marafico, S., Biteau, J., Condorelli, A. & Deligny, O. , *ApJ* **936** (2022), 62
- [11] Madau, P. & Dickinson, M. , *ARA&A* **52** (2014), pp. 415-486
- [12] McCall, M. , *MNRAS*. **440** (2014), 405-426
- [13] Condon, J., Matthews, A. & Broderick, J. , *ApJ* **872** (2019), e148 .
- [14] Aloisio, R., Boncioli, D., Grillo, A., Petrera, S. & Salamida, F. , *JCAP* **10** (2012), pp. 007
- [15] Gilmore, R., Somerville, R., Primack, J. & Dominguez, A. , *MNRAS*. **422** pp. 3189 (2012)
- [16] Pueschel, E. & Biteau, J. , *ArXiv E-prints*. arXiv:2112.05952 (2021,12)
- [17] Aab, A. et al. , *NIM A*. **798** (2015), pp. 172-213
- [18] Abreu, P. & Others *EPJC* **81** (2021), pp. 966
- [19] Bellido, J. & Pierre Auger Collaboration , *PoS. ICRC2017* **506** (2018)
- [20] Pierog, T., Karpenko, I., Katzy, J. Yatsenko, E. & Werner, K. , *PhRvC*. **92**, 034906 (2015)
- [21] Riehn, F., Dembinski, H., Engel, R., Fedynitch, A., Gaisser, T. & Stanev, T. , *PoS. ICRC2017* **301** (2017)
- [22] Castellina, A & Pierre Auger Collaboration , *EPJ Web Conf.*, **210** (2019)