PROCEEDINGS OF SCIENCE

Global Spline Fit GSF-2025 – An update of the data-driven model of the cosmic-ray flux and its mass composition

Hans Dembinski,[a] **Ralph Engel*,**[b,c] **Anatoli Fedynitch**[d] **and Kozo Fujisue**[d]

[a] *Fakultät Physik, Technische Universität Dortmund, 44221 Dortmund, Germany*
[b] *Karlsruhe Institute of Technology, Institute for Astroparticle Physics, Germany*
[c] *Karlsruhe Institute of Technology, Institute of Experimental Particle Physics, Germany*
[d] *Institute of Physics, Academia Sinica, 11529 Taipei, Taiwan*

The Global Spline Fit (GSF) is a data-driven parameterization of the cosmic-ray flux and corresponding mass composition. It combines direct and indirect measurements of cosmic-ray flux of individual elements from 1 GeV to 10^{11} GeV, considering their uncertainties and possible difference in the energy scale of the measurements. At lower energies, the flux is corrected to the local interstellar spectra using the individual data-taking periods of the experiments. The systematic energy scale uncertainty for each experiment is treated as a nuisance parameter and minimized jointly with other model parameters, thus, matching the flux from indirect measurements above the knee to direct measurements below the knee region. Since the original work was presented in 2017, many new measurements have been published by both direct and indirect cosmic-ray experiments. This presentation describes an updated GSF model, referred to as GSF-2025, which includes datasets from the last eight years. We assess the mutual compatibility and demonstrate the impact of the newly added data on the all-particle flux and mass composition over 11 decades in energy. It is shown that a much better understanding of the energy region around the knee is obtained.

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

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(ICRC2025)248

1. Introduction

The number of high-quality measurements of the all-particle flux of cosmic rays and also of high-statistics data sets of fluxes of individual elements or mass groups has increased over the last two decades considerably. This has motivated the authors of [1] to develop a first purely data driven parametrization of the overall cosmic ray spectrum and its composition, called Global Spline Fit (GSF), in 2017. Having a self-consistent parametrization of the world data of cosmic-ray fluxes, including the corresponding data-derived uncertainty bands, is very valuable for a multitude of applications ranging from identifying regions of inconsistent data or of lack of information, over the calculation of secondary particle fluxes in the atmosphere, to interpreting the observed flux and composition features in terms of astrophysical models and theories.

Several important challenges have to be addressed to obtain an agnostic description of the data: (i) The different data sets cover only a fraction of the cosmic ray spectrum and there are energy ranges without detailed composition data. (ii) Different measurements are characterized by different sys. uncertainties, in particular, of the overall energy scale of the data sets. (iii) Low-energy data are available for individual elements and sometimes even for several isotopes of one element, high-energy data are available only for mass groups. (iv) Inconsistent data sets of the existing world data have to be identified and, if needed, excluded. (v) Trying to implement as much as possible an agnostic approach, no assumptions on the shape of the spectra should be used as external constraints. This necessarily leads to a certain dependence of the results on the choice of smallest scale of a structure in the spectrum. (vi) At low energy, solar modulation should be corrected for to obtain the local interstellar spectrum (LIS).

The GSF approach is one implementation of solutions to these challenges based on combining piecewise spline fits for the individual elements of the cosmic ray spectrum. These individual spline fits are, by construction, connected smoothly to reproduce the flux data. Thanks to obtaining the parameters of the spline functions by a global minimization, also uncertainties and correlations (encoded in the covariance matrix) can be given for the obtained flux model and used in applications of it.

2. Methods

2.1 Model Construction

We adopt the Global Spline Fit (GSF) framework introduced in 2017 for a data-driven modeling of cosmic-ray spectra. In this approach, all cosmic-ray nuclei are partitioned into four mass groups ($L = 1 \ldots 4$) – proton (p), helium (He), the oxygen group (O*), and the iron group (Fe*) – roughly dividing the logarithmic mass scale into four equal parts. Each group is represented by a leading element (indicated with a star) whose flux $J_L(R)$ is modeled as a smooth function of rigidity R. We use a flexible spline parametrization: specifically, $J_L(R)$ is expanded in a basis of cubic B-splines defined on a grid in $\ln R$. A scaling factor R^{-3} is applied to each spline series to improve numerical stability. The four group-spline functions together span a rigidity range from about 1 GV up to 3×10^{11} GV, covering the entire energy spectrum from GeV to ultra-high energies.

The fluxes of individual elements J_j are determined from direct measurements (e.g. HEAO and AMS-02 data) and fitted using individual splines (lower left panel of Fig. 1). At energies

PoS(ICRC2025)248

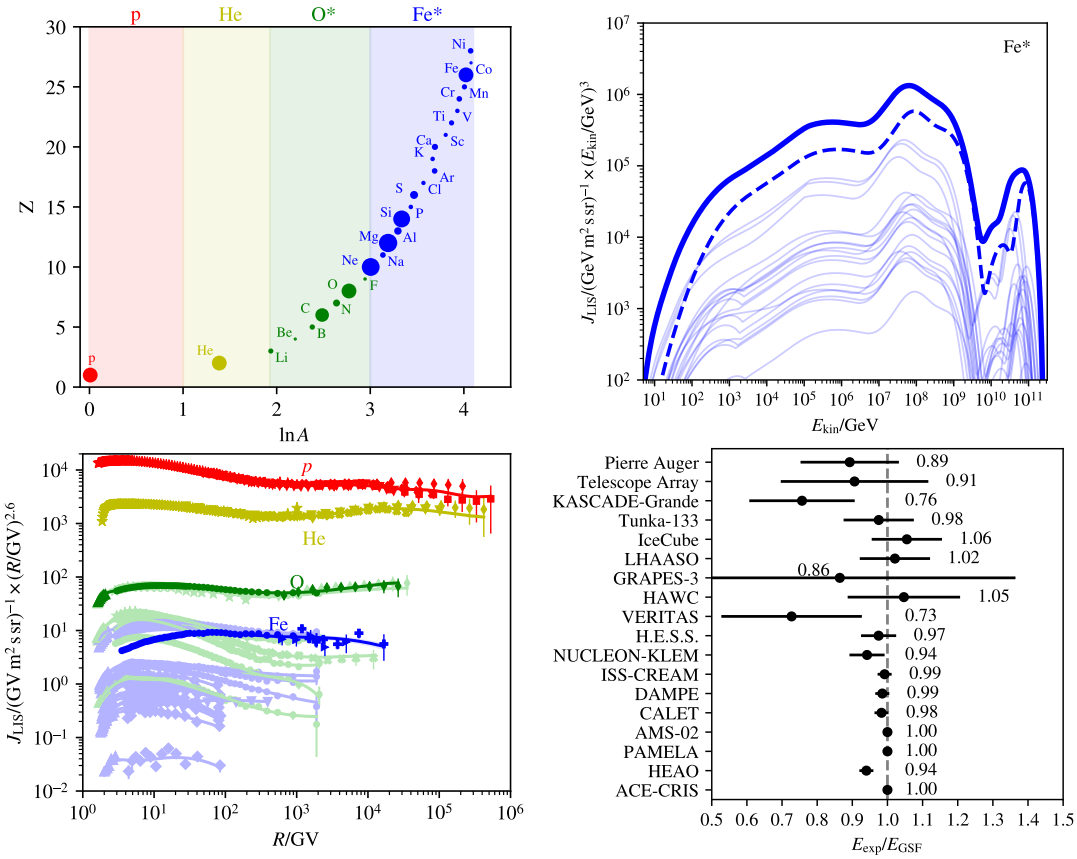

Figure 1: Top left: GSF mass groups, each named after a leading element (starred), divide $\ln A$ into four parts. Circle areas indicate flux ratios $J_j(R)/J_L(R)$. Top right: Iron group flux (thick line), elemental iron (dashed), and subleading elements (thin lines). Bottom left: Direct measurements of proton, helium, oxygen, and iron fluxes; unlabeled points show subleading elements. Bottom right: GSF energy scale adjustment factors; values < 1 indicate underestimated energy. See Tab. 1 for data sources.

outside of this range, where no data are available, the relative abundance with respect to the leading element, $w_{Lj} = J_j/J_L$, is kept constant in rigidity and determines the average composition of that group. These ratios are shown in the top left panel of Fig. 1 for a rigidity outside of the direct measurement range. This grouping strategy – assigning the elements to four groups of roughly equal size in $\ln A$ – is common practice and reflects the fact that air-shower experiments have relatively coarse resolution in mass (approximately constant resolution in $\ln A$). It provides a reasonable approximate description of composition: neighboring elements tend to show nearly constant flux ratios (in R) in direct observations. The total all-particle flux $J_{\mathrm{all}}(E)$ at energy E is obtained by summing contributions from all groups (and all elements therein), converting each group's rigidity-dependent flux to energy space. Notably, the flux of a heavy-element group (e.g. iron group) can be significantly higher than the flux of its leading element alone, as shown in the top right panel of Fig. 1.

PoS(ICRC2025)248

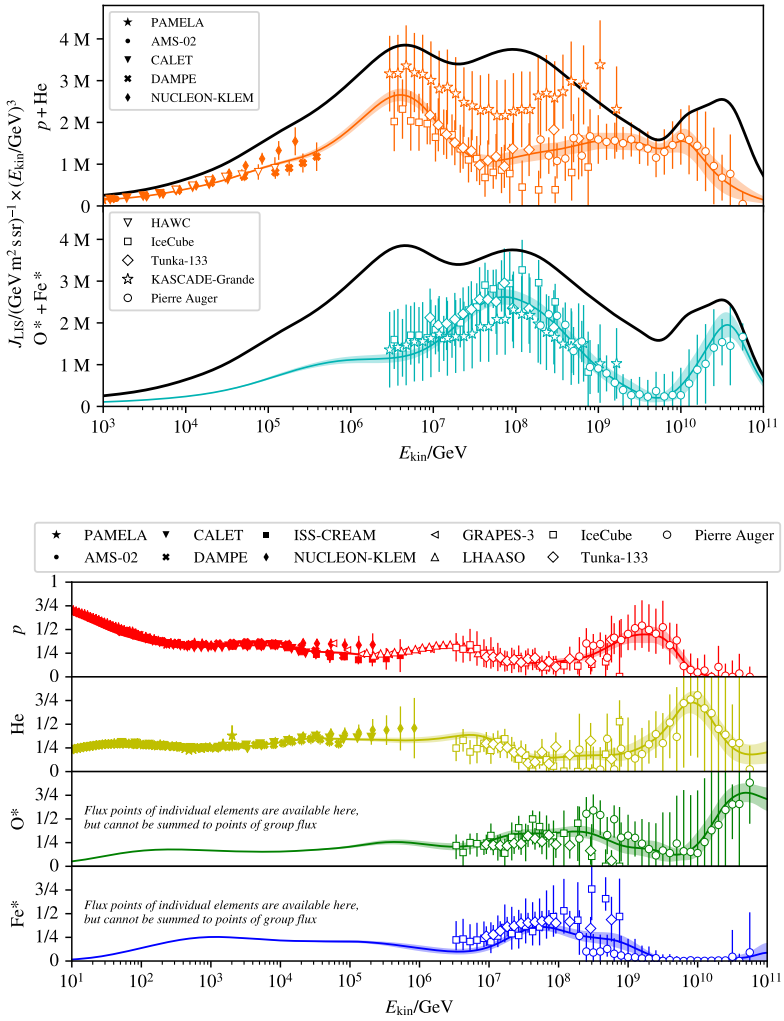

Figure 2: Top: Scaled particle flux per energy interval from direct and indirect measurements (see Tab. 1) for the sum of elements in a light (proton + helium) and heavy (other elements) mass group; the total flux (thick solid, black) is also shown for reference. Bottom: Fraction of flux per energy interval from direct and indirect measurements for the four mass groups with leading elements proton, helium, oxygen, and iron. Error bars indicate the quadratic sum of statistical and systematic uncertainties. Lines and error bands represent the GSF fit. All data points are adjusted to the common energy scale.

3. Combined fit to direct and indirect data

The GSF model is fitted to a broad collection of cosmic-ray measurements, including both direct observations of individual elemental spectra and indirect air-shower observations of aggregate fluxes and composition. All-particle energy spectra from air-shower arrays are compared to the sum of all group fluxes $J_{\rm all}(E)$. In cases where experiments report partial fluxes (such as the light-component p + He or heavy-component C–Fe), the model predictions are formed by summing the appropriate group contributions (e.g. $J_p + J_{\rm He}$ for the light group) (Fig. 2 top). Composition measurements given as fractional abundance (fractions of total flux in a mass group) are also incorporated by

PoS(ICRC2025)248

comparing fractional group fluxes at the corresponding energies (Fig. 2 bottom). In this manner, the GSF simultaneously fits all available datasets – each with potentially different energy ranges and reporting formats – in one consistent framework.

Statistical and systematic uncertainties of each data point are combined in quadrature and used to weight the fit (effectively constructing a χ^2 or likelihood function). Importantly, direct experiments typically have well-calibrated energy (rigidity) scales, whereas indirect experiments often have sizable systematic uncertainties in energy calibration. The GSF method addresses this by treating each experiment's energy-scale offset z_E as a nuisance parameter to be determined by the fit. In practice, an energy shift corresponds to adjusting the data's energy axis by a factor $(1 + z_E)$, and the flux by $(1+z_E)^{-1}$. Allowing small z_E adjustments (within the reported calibration uncertainty of each experiment) enables an automatic cross-calibration of all data sets on a common energy scale. This approach generalizes earlier efforts that aligned spectra "by hand" – here the shifts are obtained algorithmically as part of the global fit [2–4]. The penalty term $\sum_j (z_{E,j}/\sigma_{E,j})^2$ is added to the total χ^2 to constrain each shift $z_{E,j}$ within the experiment's uncertainty $\sigma_{E,j}$. This simultaneous fitting of flux and energy-scale parameters significantly reduces systematic normalization differences between experiments, leading to a more coherent all-particle spectrum.

The model has on the order of 300 free parameters (about 20 B-spline coefficients for each of 4 groups, about 5 coefficients for each of subleading elements, plus a modest number of energy-offset parameters) to be determined from the data. We perform an iterative, hybrid fitting procedure to efficiently handle this high-dimensional optimization. In essence, the method alternates between adjusting the flux shapes (linear step) and adjusting the experimental energy scales (non-linear step) to find a consistent minimum. Finally, a joint fit of all parameters together is performed starting from the previous solution to ensure that a true global minimum is reached and to compute the covariance matrix of the best-fit parameters. The covariance (obtained from the inverted Hessian matrix at the minimum) provides uncertainty estimates for each fitted component, which can be propagated to uncertainty bands on the flux predictions. The outcome of this procedure is a smoothly varying, data-driven cosmic ray flux model that spans many decades.

3.1 Data Selection and New Data Sets

The complete list of cosmic-ray data sets used in the 2025 update of the Global Spline Fit (GSF) is provided in Tab. 1. Several new high-energy direct measurements have been incorporated, including elemental spectra from CALET, DAMPE, NUKLEON, and ISS-CREAM. For many elements, the latest AMS-02 data supersede earlier measurements from HEAO and ACE-CRIS. This replacement is motivated by the significantly higher precision of AMS-02 and the limited statistical impact of the older data in the global fit. In cases where AMS-02 and HEAO data were in tension, AMS-02 was given preference. In the energy range near the knee, recent results from HAWC, GRAPES-3, and in particular LHAASO have substantially improved the observational coverage, bridging the gap between direct and indirect measurements of the light-component flux. At ultra-high energies, we have incorporated updated measurements of the energy spectrum and mass composition from the Pierre Auger Observatory.

A new feature in this GSF update is the inclusion of measurements of the moments of logarithmic mass $\ln A$. Specifically, we include the first moment published by LHAASO and the first and second moments from Auger. While moments of $\ln A$ are not entirely independent of mass fraction

PoS(ICRC2025)248

Table 1: Experimental data used in the GSF-2025 fit. Altitude h and atmospheric depth X are approximate and intended for orientation. The last two columns show the systematic uncertainties on energy scale and flux normalization. Legend: a) published systematic uncertainties used; b) uncertainty of flux fractions in mass groups increased by 0.1; c) uncertainty of mass group fractions taken as the range across hadronic interaction models; d) 0.2 for light/heavy flux, 0.1 for total flux. Atmospheric depths are based on [4].

Experiment	Detector type	h [km]	X [g cm^{-2}]	$\sigma_{\rm sys}[E]/E$	$\sigma_{\rm sys}[J]/J$
ACE-CRIS [5, 6]	satellite	–	–	–	a)
HEAO [7]	satellite	–	–	0.02	a)
PAMELA [8, 9]	satellite	–	–	–	a)
AMS-02 [10–13]	satellite	–	–	–	a)
CALET [14–19]	satellite	–	–	0.02	a)
DAMPE [20–23]	satellite	–	–	0.02	a)
ISS-CREAM [24]	satellite	–	–	0.02	a)
NUCLEON-KLEM [25]	satellite	–	–	0.05	a)
GRAPES-3 [26]	surface array	2.2	600	0.50	a)
H.E.S.S. [27]	Cherenkov telescope	1.8	795	0.05	a)
VERITAS [28]	Cherenkov telescope	1.3	860	0.20	a)
HAWC [29, 30]	surface array	4.1	540	0.16	a)
LHAASO [31, 32]	optical + surface	4.4	520	0.10	a)
IceCube [33, 34]	surface array	3.0	680	0.10	a), b)
Tunka [35, 36]	optical	0.68	938	0.10	0.10
KASCADE-Grande [37]	surface array	0.11	1022	0.15	d)
TA [38, 39]	optical + surface	1.4	845	0.20	0.10
Auger [40–45]	optical + surface	1.4	845	0.14	a), c)

data, we find that these measurements provide additional constraints predominantly at higher energies, where mass group fraction data are sparse. In the overlapping energy region, the inclusion of $\ln A$ moments has negligible impact on the fit.

4. Results

The main result of the GSF model is shown in Fig. 3. We observe excellent consistency between direct measurements, with all data sets agreeing within their quoted uncertainties. Notably, the proton spectrum is now experimentally constrained from the lowest rigidities to the highest energies observed, owing to recent data from DAMPE, ISS-CREAM, and LHAASO. The helium flux is also well determined across the full energy range of direct measurements and the crossing of p and He fluxes is well established.

The least certain components remain the oxygen and iron groups near knee energies. Between the knee and the ankle, previously noted discrepancies between IceCube, Auger, and KASCADE-Grande persist, largely unchanged since the 2017 GSF analysis. These tensions are not unexpected, given the differing hadronic interaction models employed in the interpretation of air-shower data.

We encourage indirect detection experiments—such as TALE, LHAASO, and Auger—to provide results interpreted using multiple hadronic models to facilitate model-agnostic comparisons and improve compatibility across datasets.

At ultra-high energies, the GSF fit continues to be dominated by data from the Pierre Auger Observatory. The inclusion of recent composition and moment measurements has led to a further reduction in the uncertainties of the high-energy mass composition.

In summary, the GSF-2025 model provides a coherent and comprehensive description of the global cosmic-ray flux and composition data, serving both as a tool for interpretation and a reference

PoS(ICRC2025)248

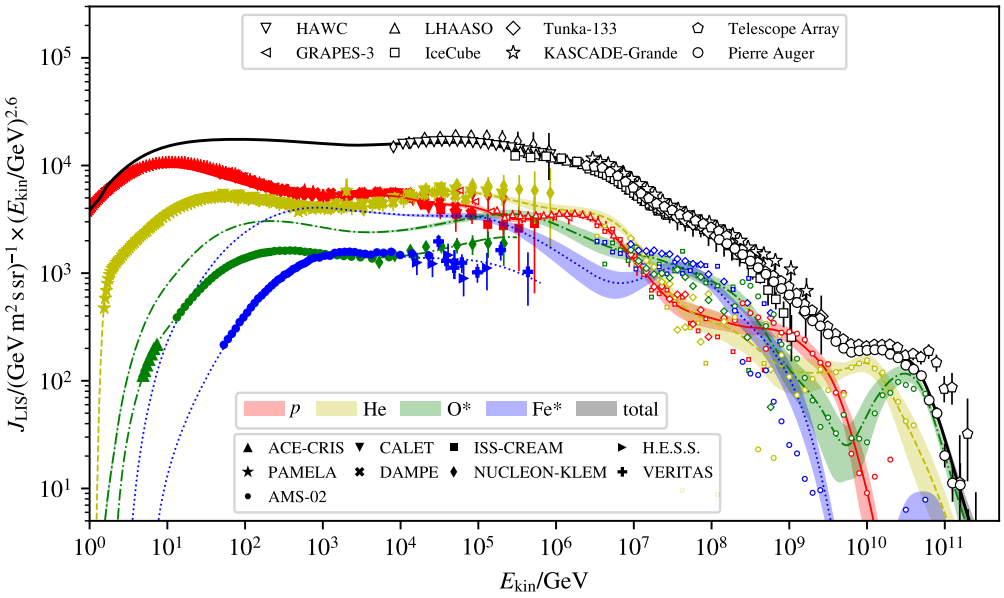

Figure 3: Scaled cosmic-ray flux per energy interval from direct and indirect measurements (see Tab. 1). Solid markers indicate direct measurements of the leading elements in each mass group; open markers represent indirect air-shower observations. Error bars denote the quadratic sum of statistical and systematic uncertainties, except for indirect mass-group data, where they are omitted for clarity. All data points are adjusted to a common energy scale E, consistent with Fig. 1. Solid lines and shaded bands show the GSF prediction and its uncertainty. Note that the fluxes of elemental oxygen and iron are lower than the total flux of their respective mass groups, as discussed in the text.

for future theoretical and experimental studies.

Acknowledgments AF and KF acknowledge support from Academia Sinica (Grant No. AS-GCS-113-M04) and the National Science and Technology Council (Grant No. 113-2112-M-001-060-MY3). RE has been supported by the German Ministry of Education and Research (BMBF) grant no 05A23VK4.

References

[1] H. P. Dembinski, R. Engel, A. Fedynitch, T. Gaisser, F. Riehn, and T. Stanev, PoS **ICRC2017**, 533 (2018), arXiv:1711.11432 [astro-ph.HE] .

[2] J. R. Hoerandel, Astropart. Phys. **19**, 193 (2003), arXiv:astro-ph/0210453 [astro-ph] .

[3] T. K. Gaisser, Astroparticle Physics **35**, 801 (2012).

[4] T. K. Gaisser, T. Stanev, and S. Tilav, Front. Phys.(Beijing) **8**, 748 (2013), arXiv:1303.3565 [astro-ph.HE] .

[5] T. T. von Rosenvinge *et al.*, Astrophys. J. **770**, 117 (2013).

[6] G. A. de Nolfo *et al.*, Adv. Space Res. **38**, 1558 (2006), arXiv:astro-ph/0611301 .

[7] J. J. Engelmann, P. Ferrando, A. Soutoul, P. Goret, and E. Juliusson, Astron. Astrophys. **233**, 96 (1990).

[8] O. Adriani *et al.* (PAMELA), Science **332**, 69 (2011), arXiv:1103.4055 [astro-ph.HE] .

PoS(ICRC2025)248

[9] O. Adriani *et al.*, Astrophys. J. **791**, 93 (2014), arXiv:1407.1657 [astro-ph.HE] .

[10] M. Aguilar *et al.* (AMS), Phys. Rept. **894**, 1 (2021).

[11] M. Aguilar *et al.* (AMS), Phys. Rev. Lett. **126**, 041104 (2021).

[12] M. Aguilar *et al.* (AMS), Phys. Rev. Lett. **127**, 02101 (2021), [Erratum: Phys.Rev.Lett. 127, 159901 (2021)].

[13] M. Aguilar *et al.* (AMS), Phys. Rev. Lett. **130**, 211002 (2023).

[14] O. Adriani *et al.* (CALET), Phys. Rev. Lett. **128**, 131103 (2022), arXiv:2204.00845 [astro-ph.HE] .

[15] O. Adriani *et al.* (CALET), Phys. Rev. Lett. **129**, 101102 (2022), arXiv:2209.01302 [astro-ph.HE] .

[16] O. Adriani *et al.* (CALET), Phys. Rev. Lett. **130**, 171002 (2023), arXiv:2304.14699 [astro-ph.HE] .

[17] O. Adriani *et al.* (CALET), Phys. Rev. Lett. **125**, 251102 (2020), arXiv:2012.10319 [astro-ph.HE] .

[18] O. Adriani *et al.* (CALET), Phys. Rev. Lett. **129**, 251103 (2022), arXiv:2212.07873 [astro-ph.HE] .

[19] O. Adriani *et al.* (CALET), Phys. Rev. Lett. **126**, 241101 (2021), arXiv:2106.08036 [astro-ph.HE] .

[20] Q. An *et al.* (DAMPE), Sci. Adv. **5**, eaax3793 (2019), arXiv:1909.12860 [astro-ph.HE] .

[21] F. Alemanno *et al.* (DAMPE), Phys. Rev. Lett. **126**, 201102 (2021), arXiv:2105.09073 [astro-ph.HE] .

[22] F. Alemanno *et al.* (DAMPE), Phys. Rev. D **109**, L121101 (2024), arXiv:2304.00137 [astro-ph.HE] .

[23] F. Alemanno *et al.* (DAMPE), Phys. Rev. Lett. **134**, 191001 (2025), arXiv:2412.11460 [astro-ph.HE] .

[24] G. H. Choi *et al.*, Astrophys. J. **940**, 107 (2022).

[25] N. Gorbunov *et al.*, Adv. Space Res. **64**, 2546 (2019), arXiv:1809.05333 [astro-ph.IM] .

[26] F. Varsi *et al.* (GRAPES-3), Phys. Rev. Lett. **132**, 051002 (2024).

[27] F. Aharonian *et al.* (H.E.S.S.), Phys. Rev. D **75**, 042004 (2007), arXiv:astro-ph/0701766 .

[28] A. Archer *et al.* (VERITAS), Phys. Rev. D **98**, 022009 (2018), arXiv:1807.08010 [astro-ph.HE] .

[29] J. A. Morales-Soto *et al.* (HAWC), PoS **ICRC2021**, 330 (2021), arXiv:2108.04748 [astro-ph.HE] .

[30] A. Albert *et al.* (HAWC), Phys. Rev. D **105**, 063021 (2022), arXiv:2204.06662 [astro-ph.HE] .

[31] Z. Cao *et al.* (LHAASO), Phys. Rev. Lett. **132**, 131002 (2024), arXiv:2403.10010 [astro-ph.HE] .

[32] Z. Cao *et al.* (LHAASO), (2025), arXiv:2505.14447 [astro-ph.HE] .

[33] M. G. Aartsen *et al.* (IceCube), Phys. Rev. D **100**, 082002 (2019), arXiv:1906.04317 [astro-ph.HE] .

[34] M. G. Aartsen *et al.* (IceCube), Phys. Rev. D **102**, 122001 (2020), arXiv:2006.05215 [astro-ph.HE] .

[35] V. V. Prosin *et al.*, Nucl. Instrum. Meth. A **756**, 94 (2014).

[36] N. M. Budnev *et al.*, Astropart. Phys. **117**, 102406 (2020), arXiv:2104.03599 [astro-ph.HE] .

[37] S. Schoo *et al.* (KASCADE-Grande), PoS **ICRC2015**, 263 (2016).

[38] R. U. Abbasi *et al.* (Telescope Array), Astrophys. J. **865**, 74 (2018), arXiv:1803.01288 [astro-ph.HE] .

[39] D. Ivanov (Telescope Array), PoS **ICRC2019**, 298 (2020).

[40] J. Bellido (Pierre Auger), PoS **ICRC2017**, 506 (2018).

[41] A. Aab *et al.* (Pierre Auger), Phys. Rev. Lett. **125**, 121106 (2020), arXiv:2008.06488 [astro-ph.HE] .

[42] P. Abreu *et al.* (Pierre Auger), Eur. Phys. J. C **81**, 966 (2021), arXiv:2109.13400 [astro-ph.HE] .

[43] A. Abdul Halim *et al.* (Pierre Auger), PoS **ICRC2023**, 438 (2023).

[44] A. Castellina (Pierre Auger), PoS **ICRC2019**, 004 (2021), arXiv:1909.10791 [astro-ph.HE] .

[45] A. Abdul Halim *et al.* (Pierre Auger), Phys. Rev. D **111**, 022003 (2025), arXiv:2406.06319 [astro-ph.HE] .

PoS(ICRC2025)248