PROCEEDINGS OF SCIENCE

Global Spline Fit (GSF) 2024

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

[a] *Fakultät Physik, Technische Universität Dortmund, Otto-Hahn Str. 4a, 44221, Dortmund, Germany*

[b] *Karlsruhe Institute of Technology, D-76021, Karlsruhe, Germany*

[c] *Institute of Physics, Academia Sinica, 11529, Taipei, Taiwan*

E-mail: hans.dembinski@tu-dortmund.de, ralph.engel@kit.edu, anatoli@gate.sinica.edu.tw, kfujisue@gate.sinica.edu.tw

The Global Spline Fit (GSF) is a data-driven parametrization of cosmic-ray flux and 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. 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. In this study, an updated GSF is shown, including datasets from the last seven 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.

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.

PoS(UHECR2024)087

1. Introduction

Cosmic rays have been observed across more than eleven orders of magnitude in energy through various experiments. At energies below roughly 1 PeV, cosmic rays are directly observed using satellites and balloon-borne instruments, enabling precise measurements of the flux of individual elements. In contrast, at energies above roughly 10 TeV, cosmic rays are indirectly observed through the detection of extensive air showers at ground level. The air shower observatories have reported the all-particle flux, the flux of cosmic rays within a given mass range, and the moments of the mass of cosmic rays. Each observational method covers a specific energy range, and their overlapping energy regions collectively span over eleven decades in energy. Each experiment has inherent uncertainties in its measurements. In particular, in the high-energy region where indirect observations are employed, there is an energy scale uncertainty of approximately 10–20%.

The *Global Spline Fit* (GSF) [1] is a data-driven model of cosmic-ray flux and mass composition covering the energy range from 1 GeV to 10^{11} GeV. The GSF parametrizes observational data without assuming power-law shapes or rigidity-dependent cutoffs, while simultaneously cross-calibrating the energy scales of different experiments. The covariance matrix of this joint fit represents the uncertainties of the model, which typically follow those of the experimental data. The GSF model provides a smooth and mathematical representation of experimental knowledge on cosmic-ray flux and mass composition, and hence allows for computing other observables — such as the nucleon flux for atmospheric lepton fluxes — together with their associated uncertainties.

Since the original GSF work in 2017, numerous new observational results from both direct and indirect measurements have been published. In this study, we present an updated GSF model that incorporates observational data from the last seven years. We evaluate the mutual compatibility of the updated and previous models and demonstrate the impact of the newly included observational data on the cosmic-ray flux and mass composition.

2. Global Spline Fit methodology

The fitting procedure follows the same methodology as the original work [1]. The cosmic-ray flux is divided into four mass groups, each spanning approximately equal intervals in logarithmic mass ($\ln A$), based on the mass sensitivity of air-shower measurements. The most abundant element in each mass group is designated as the leading element L of that group, with $L \in \{p, \mathrm{He}, \mathrm{O}, \mathrm{Fe}\}$. The oxygen and iron groups also include sub-leading elements. At low energies, where individual elements are directly measured, these elements are represented by individual splines. Outside this range, their flux $J_i(R)$ is taken to be proportional to the leading element flux $J_L(R)$:

$$J_i(R) = f_{iL} \times J_L(R), \tag{1}$$

where f_{iL} is a constant flux ratio. This assumption is empirically supported by direct observations, which show approximately constant flux ratios between neighboring elements as a function of rigidity. The contribution of sub-leading elements amounts to roughly half of the flux for the oxygen group and more than half for the iron group. Thus, these additional contributions must be explicitly considered when transitioning to the high-energy region, where air-shower measurements encompass the entire mass group.

The flux of a leading element $J_L(R)$ is parameterized as a smooth curve using a linear combination of B-splines

$$J_L(R) = \sum_k a_{Lk} \times b_k\big(\ln(R/\mathrm{GV})\big) \times (R/\mathrm{GV})^{-3}, \tag{2}$$

where spline amplitudes a_{Lk} are free parameters fitted to data, $b_k\big(\ln(R/\mathrm{GV})\big)$ are standard cubic B-Splines defined over a common vector of knots, and $(R/\mathrm{GV})^{-3}$ is a shaping factor to stabilize the numerical computation. The knot positions form a regular grid in $\ln(R/\mathrm{GV})$, typically with an interval of 0.5.

In this model, the flux is parameterized as a function of rigidity. Observations originally expressed in terms of energy — such as air-shower measurements — are related to the model via $J(E) = J(R)\mathrm{d}R/\mathrm{d}E$. The all-particle flux or flux of a given mass group is obtained by summing the contributions of all relevant elements. In this updated version, we also incorporate measurements of the mean logarithmic mass $\langle \ln A \rangle$, calculated as $\langle \ln A \rangle = (\sum_i \ln A_i \times J_i(E))/(\sum_i J_i(E))$, where $\sum_i$ runs over all elements.

The effect of energy-scale offsets in experiments is accounted for by introducing nuisance parameters $z_E = (\tilde{E} - E)/E$, where $\tilde{E}$ is the energy reported by each experiment. If $J(E)$ denotes the true differential flux at the true energy E, the flux measured by an experiment is given by

$$\tilde{J}(\tilde{E}) = J(E)\frac{\mathrm{d}E}{\mathrm{d}\tilde{E}} = J\left(\frac{\tilde{E}}{1+z_E}\right)\frac{1}{1+z_E}. \tag{3}$$

The residuals between these measurements and the GSF model are then used to determine both the best-fit spline amplitudes a_{Lk} and energy-scale offsets z_E. This is accomplished by minimizing

$$F' = \sum_i F_i + \sum_i \left(\frac{z_{E,i}}{\sigma_{E,i}}\right)^2, \tag{4}$$

where F_i represents the sum of squared residuals between the data points of experiment i and the model. These residuals incorporate statistical and systematic uncertainties, including correlations, while the term $\left(z_{E,i}/\sigma_{E,i}\right)^2$ penalizes large deviations from each experiment's nominal energy scale. In doing so, the global fit balances the need to reduce data-model discrepancies against the requirement that any shifts in energy scale remain compatible with each experiment's reported uncertainties.

3. Datasets

Since the original GSF model was introduced in 2017, numerous direct and indirect cosmic-ray measurements have been published. In particular, new satellite data now span energies from tens of GeV up to the PeV range, overlapping with ground-based measurements and expanding the available data. We have updated our data set accordingly, incorporating results published in the last seven years, including some from the *Cosmic-Ray Data Base* (CRDB) [2].

To highlight the impact of these additions on cosmic-ray flux and mass composition, we constructed five data sets. *Data Set 1* serves as the baseline, while *Data Set 2*, *Data Set 3*, *Data Set*

Table 1: Datasets used for the GSF 2024 model. A circle (○) indicates that an experiment's measurements are included in a given dataset, while labels (a), (b), (c), (d), and (e) denote specific dataset variations. **ACE-CRIS:** Li, B, C, N, O, F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, and Ni; datasets marked with (a) exclude Fe data. **HEAO:** Be, B, C, N, O, F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar, K, Ca, Sc, Ti, V, Cr, Mn, Fe, Co, and Ni; datasets marked with (b) exclude Be, B, C, N, O, F, Ne, Na, Mg, Al, Si, S, and Fe data. **PAMELA:** H and He, plus B and C. **AMS-02:** H, He, Li, Be, B, C, N, O, F, Ne, Na, Mg, Al, Si, S, and Fe. **CALET:** H, He, B, C, O, Fe, and Ni; datasets marked with (c) exclude B, C, O, and Fe. **DAMPE:** H, He, and a combined H + He flux at high energies (where separate proton and helium data are unavailable). **CREAM I+III:** H and He. **ISS-CREAM:** H. **NUCLEON-KLEM:** H, He, C, O, Ne, Mg, Si, Fe, and the all-particle flux; datasets marked with (d) exclude Ne, Mg, Si, and Fe. **GRAPES-3:** p. **H.E.S.S.:** Fe. **VERITAS:** Fe. **HAWC:** H + He and the all-particle flux. **LHAASO:** All-particle flux and $\langle \ln A \rangle$. **IceCube:** All-particle flux and flux of each mass group (listed twice to indicate multiple data samples). **KASCADE-Grande:** Flux of a light-mass group (treated as H + He) and a heavy-mass group (treated as O and Fe). **TA:** All-particle flux. **Auger:** All-particle flux and flux of each mass group. Datasets marked with (e) use the same configuration as the previous GSF model [43].

Experiment	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5
ACE-CRIS [4, 5]	(a)	(a)	○	(a)	(a)
HEAO [6]	(b)	(b)	○	(b)	(b)
PAMELA [7, 8]	○	○	○	○	○
AMS-02 [9–12]	○	○		○	○
CALET [13–18]	(c)		○	(c)	(c)
DAMPE [19–21]	○		○	○	○
CREAM I+III [22]		○			
ISS-CREAM [23]	○		○	○	○
NUCLEON-KLEM [24]	(d)		○	(d)	(d)
GRAPES-3 [25]	○		○	○	○
H.E.S.S. [26]	○	○	○	○	○
VERITAS [27]	○	○	○	○	○
HAWC [28, 29]	○	○	○	○	○
LHAASO [30]	○	○	○		○
IceCube [31, 32]	○	○	○	○	○
Tunka [33, 34]	○	○	○	○	○
KASCADE-Grande [35]	○	○	○	○	○
TA [36, 37]	○	○	○	○	○
Auger [38–42]	○	○	○	○	(e)

4, and *Data Set 5* represent variations on this baseline. Table 1 summarizes each data set and the experiments it includes.

Data Set 1 serves as the baseline for GSF 2024 and includes all modern data, except for certain measurements that display tension with AMS-02. Specifically, iron data from ACE-CRIS [4], boron/carbon/oxygen/iron data from CALET [13, 14, 17], and iron-group data from NUCLEON-KLEM [24] are excluded, as indicated in Table 1 by symbols (a), (c), and (d). In addition, HEAO measurements are used only for elements not covered by AMS-02, marked by symbol (b).

Data Set 2 illustrates the effect of newly added measurements in the tens-of-GeV to few-PeV range. Here, we remove the hydrogen and helium data from ISS-CREAM, CALET, DAMPE, NUCLEON-KLEM, and GRAPES-3 included in *Data Set 1*. Instead, we adopt the combined CREAM-I and CREAM-III hydrogen and helium fluxes [22] used in the 2019 GSF model (GSF 2019) [43].
Data Set 3 explores the impact of excluding AMS-02 entirely. It reintroduces the ACE-CRIS iron data, the CALET boron/carbon/oxygen/iron data, the NUCLEON-KLEM iron-group data initially excluded in *Data Set 1*, and all HEAO measurements, while omitting AMS-02 altogether.
Data Set 4 is derived from *Data Set 1* by removing the LHAASO measurements.
Data Set 5 is identical to *Data Set 1* except that Auger data are replaced with those used in GSF 2019 [38, 39], indicated by symbol (e) in Table 1.

4. Results and discussion

Figure 1 presents the fitted results for *Data Set 1*, yielding a χ^2 of 1034 with 1072 degrees of freedom, which indicates a good fit. The fitted results for all five GSF 2024 data sets, along with the previous GSF 2019 model, are shown in Fig. 2. Overall, the updated GSF 2024 model reproduces the well-known features of the cosmic-ray spectrum and exhibits smaller uncertainties, reflecting the improved precision of recent measurements. In particular, it shows the dip near 5 TeV, knee near 5 PeV, low-energy ankle around 10 PeV, second knee at about 100 PeV, ankle around 8 EeV, instep at roughly 20 EeV, toe near 50 EeV, and an approximately rigidity-dependent cutoff at ultra-high energies qualitatively resembling a Peters cycle [44].

Above $\sim$ 10 TeV, the all-particle flux of GSF 2024 is higher than that of GSF 2019, primarily because of the newly included data below the PeV scale, as seen by comparing Data Sets 1, 2, and 3. The model also incorporates the recent observations of flux hardening near 300 GV and softening around 1 TV in both proton and helium components. The right panel of Fig. 2 highlights the influence of LHAASO and the updated Auger data: LHAASO contributes notably to the light-mass component near the knee, while the new Auger data lead to a steeper drop in proton flux below the ankle.

Figure 3 shows the nucleon flux plotted per nucleon energy, which is relevant for atmospheric lepton flux calculations. The new features in the proton and helium fluxes, such as the bump around 10 TeV per nucleon and dip around a few hundred TeV per nucleon, are also reflected in the nucleon flux. These features may be expected to manifest in the resulting atmospheric lepton flux.

5. Summary

We present an updated version of the Global Spline Fit (GSF), a data-driven parametrization of cosmic-ray flux and mass composition that now incorporates measurements from the past seven years. Overall, the well-established spectral features seen in the previous model are confirmed, with smaller uncertainties reflecting the precision of recent data. Notably, newly observed hardening and softening in the proton and helium fluxes are reproduced, leading to corresponding structures in the nucleon flux per nucleon energy.

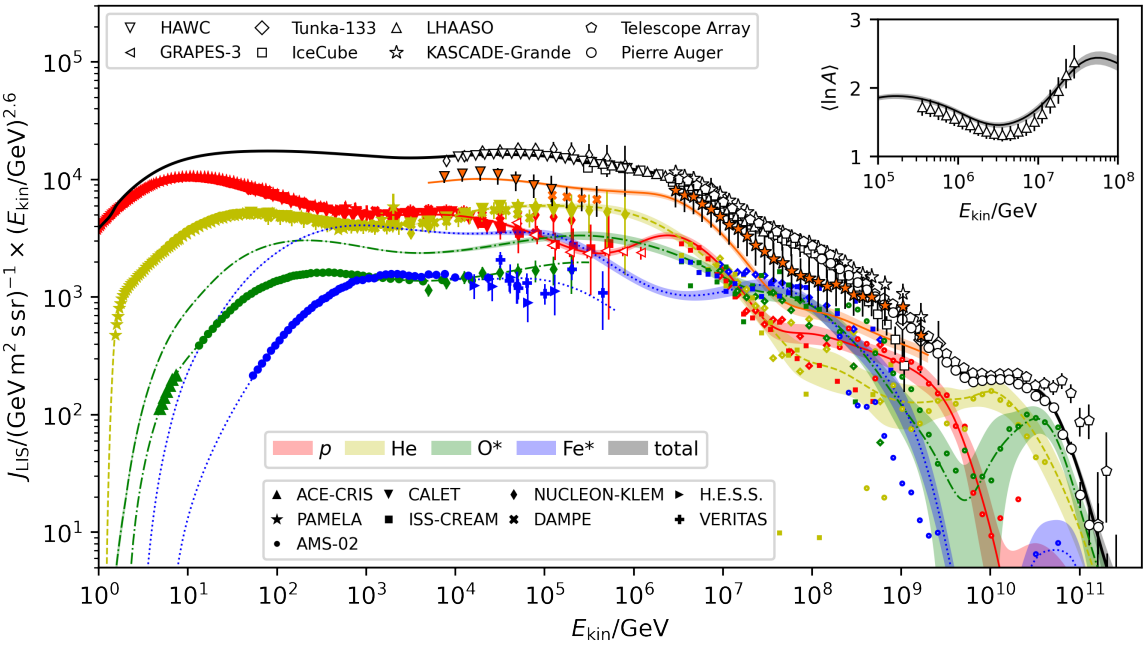

Figure 1: The cosmic-ray flux from the GSF 2024 fitted with Data set 1: all-particle flux (black thick solid line), proton flux (red solid line), helium flux (yellow dashed line), oxygen group flux (green dash-dotted line), and iron group flux (blue dotted line). The orange solid line represents the flux from the combination of proton and helium. The bands around the lines indicate one standard deviation regions. The elemental fluxes of oxygen and iron are also shown. The data points represent measurements after the energy scaling described in the text. The error bars represent quadratic sum of statistical and systematic uncertainties. The inset panel shows the mean logarithmic mass from the GSF 2024 (black solid line) and data points from LHAASO [30].

6. Acknowledgments

KF and AF acknowledge the support from the Academia Sinica Grand Challenge Seed Grant under Grant No. AS-GCS-113-M04.

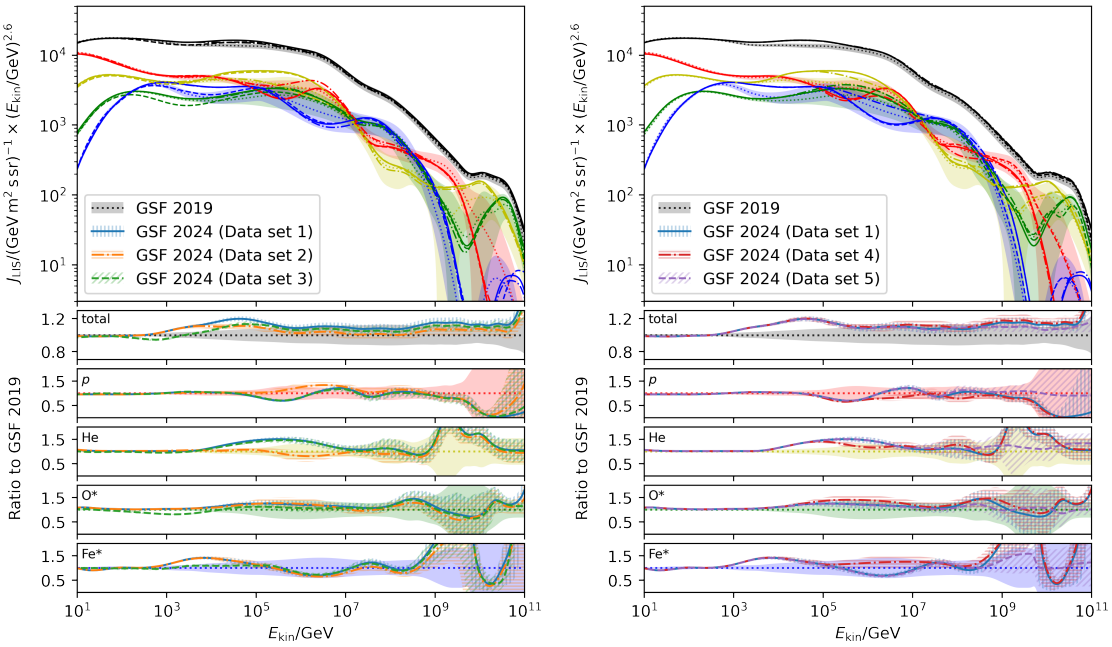

Figure 2: Comparison of cosmic-ray flux between the GSF 2019 model and GSF 2024 models. In both left and right panels, the dotted and solid lines represent GSF 2019 and GSF 2024 with Data Set 1, respectively. The dash-dotted (dashed) lines indicate GSF 2024 with Data Set 2 (Data Set 3) in the left panel and GSF 2024 with Data Set 4 (Data Set 5) in the right panel. The top panels show the cosmic-ray fluxes for all particles (black), proton (red), helium (yellow), the oxygen group (green), and the iron group (blue), along with the error bands of GSF 2019. The lower panels display the ratio of the cosmic-ray fluxes of GSF 2024 to those of GSF 2019 for all particles (second panel), proton (third panel), helium (fourth panel), the oxygen group (fifth panel), and the iron group (sixth panel). The filled bands represent the standard deviation regions of GSF 2019, while the hatched regions indicate those of the GSF 2024 models.

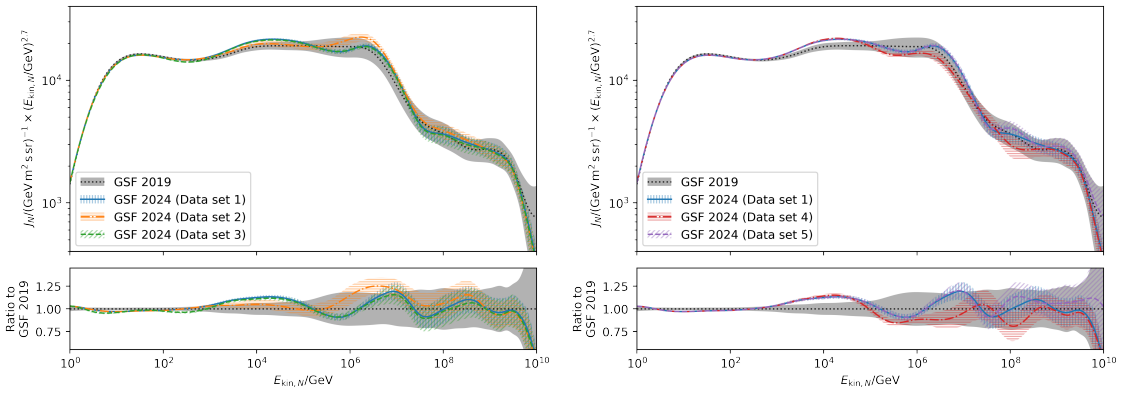

Figure 3: Comparison of nucleon flux per nucleon energy between the GSF 2019 model and GSF 2024 models. The notation follows the same convention as Fig. 2

References

[1] H. Dembinski *et al.*, *PoS* **ICRC2017** (2018) 533.
[2] D. Maurin *et al.*, *Eur. Phys. J. C* **83** (2023) 971.
[3] C. de Boor, *A Practical Guide to Splines. Springer Verlag*, New York, Heidelberg, Berlin, 1978.
[4] K. A. Lave *et al.*, *Astrophys. J.* **770** (2013) 117.
[5] G. A. de Nolfo *et al.*, *Adv. Space Res.* **38** (2006) 1558.
[6] J. J. Engelmann *et al.*, *Astron. Astrophys.* **233** (1990) 96.
[7] O. Adriani *et al.*, *Science* **332** (2011) 69.
[8] O. Adriani *et al.*, *Astrophys. J.* **791** (2014) 93.
[9] M. Aguilar *et al.*, *Phys. Rep.* **894** (2021) 1.
[10] M. Aguilar *et al.*, *Phys. Rev. Lett.* **126** (2021) 041104.
[11] M. Aguilar *et al.*, *Phys. Rev. Lett.* **127** (2021) 021101.
[12] M. Aguilar *et al.*, *Phys. Rev. Lett.* **130** (2023) 211002.
[13] O. Adriani *et al.*, *Phys. Rev. Lett.* **125** (2020) 251102.
[14] O. Adriani *et al.*, *Phys. Rev. Lett.* **126** (2021) 241101.
[15] O. Adriani *et al.*, *Phys. Rev. Lett.* **128** (2022) 131103.
[16] O. Adriani *et al.*, *Phys. Rev. Lett.* **129** (2022) 101102.
[17] O. Adriani *et al.*, *Phys. Rev. Lett.* **129** (2022) 251103.
[18] O. Adriani *et al.*, *Phys. Rev. Lett.* **130** (2023) 171002.
[19] Q. An *et al.*, *Sci. Adv.* **5** (2019) eaax3793.
[20] F. Alemanno *et al.*, *Phys. Rev. Lett.* **126** (2021) 201102.
[21] F. Alemanno *et al.*, *Phys. Rev. D* **109** (2024) L121101.
[22] Y. S. Yoon *et al.*, *Astrophys. J.* **839** (2017) 5.
[23] G. H. Choi *et al.*, *Astrophys. J.* **940** (2022) 107.
[24] V. Grebenyuk *et al.*, *Adv. Space Res.* **64** (2019) 2546.
[25] F. Varsi *et al.*, *Phys. Rev. Lett.* **132** (2024) 051002.
[26] F. Aharonian *et al.*, *Phys. Rev. D* **75** (2007) 042004.
[27] A. Archer *et al.*, *Phys. Rev. D* **98** (2018) 022009.
[28] J. A. Morales-Soto *et al.*, *PoS* **ICRC2021** (2022) 330.
[29] A. Albert *et al.*, *Phys. Rev. D* **105** (2022) 063021.
[30] Z. Cao *et al.*, *Phys. Rev. Lett.* **132** (2024) 131002.
[31] M. G. Aarsten *et al.*, *Phys. Rev. D* **100** (2019) 082002.
[32] M. G. Aarsten *et al.*, *Phys. Rev. D* **102** (2020) 122001.
[33] V. V. Prosin *et al.*, *Nucl. Instrum. Meth. A* **756** (2014) 94.
[34] N. M. Budnev *et al.*, *Astropart. Phys.* **117** (2020) 102406.
[35] S. Schoo *et al.*, *PoS* **ICRC2015** (2015) 263.
[36] R. U. Abbasi *et al.*, *Astrophys. J.* **865** (2018) 74.
[37] D. Ivanov *et al.*, *PoS* **ICRC2019** (2019) 298.
[38] A. Castellina *et al.*, *PoS* **ICRC2019** (2017) 004.
[39] J. Bellido *et al.*, *PoS* **ICRC2017** (2018) 506.
[40] A. Aab *et al.*, *Phys. Rev. Lett.* **125** (2020) 121106.
[41] P. Abreu *et al.*, *Eur. Phys. J. C* **81** (2021) 966.
[42] O. Tkachenko *et al.*, *PoS* **ICRC2023** (2023) 438.
[43] F. Schröder, *PoS* **ICRC2019** (2019) 030.
[44] B. Peters, *Nuovo Cim* **22**, (1961) 800.