PROCEEDINGS OF SCIENCE

Prospects for constraining transient sources of UHECRs with arrival-direction data

Teresa Bister[a,b,*] **and Jonathan Biteau**[c,d]

[a] *Nationaal instituut voor subatomaire fysica (NIKHEF), Science Park 105, 1098 XG Amsterdam, The Netherlands*

[b] *IMAPP, Radboud University, Houtlaan 4, 6525 XZ Nijmegen, The Netherlands*

[c] *Université Paris-Saclay, CNRS/IN2P3, IJCLab, Bât. 100 et 200, 15 rue Georges Clémenceau, F-91405 Orsay, France*

[d] *Institut Universitaire de France (IUF), France*

E-mail: teresa.bister@ru.nl

Models in which ultra-high energy cosmic-ray (UHECR) sources are transients that occur proportionally to the star-formation rate of each galaxy have recently been shown to roughly reproduce the intermediate-scale overdensities observed in the UHECR arrival directions at $E \gtrsim 40\,\mathrm{EeV}$. In this contribution, we explore the prospects for a likelihood-based fit of such models to event-level data from the Pierre Auger Observatory. Specifically, we evaluate the sensitivity to constrain the transient emission rate as well as the time spread and blurring induced by turbulent magnetic fields in the local cosmic web. Considering also deflections by the Galactic magnetic field in realistic simulated scenarios, we investigate the ability to correctly identify the underlying magnetic-field model and quantify its impact on the parameter reconstruction.

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)070

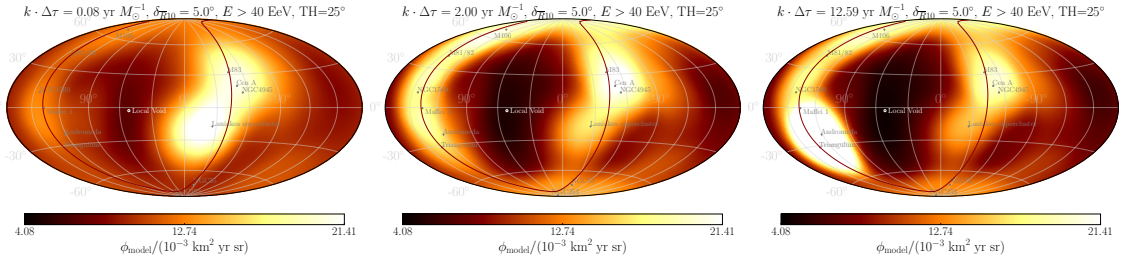
1. Introduction

Recently, the authors of [1] have shown that a model in which ultra-high-energy cosmic-ray (UHECR) sources are transients, with a rate proportional to the star-formation rate (SFR) of every galaxy, can reproduce the observed intermediate-scale anisotropies in the UHECR flux at energies $E \gtrsim 40\,\mathrm{EeV}$. Such a model could provide an explanation for the observed correlation of the UHECR arrival directions with a catalog of nearby starburst galaxies [2, 3]. In these proceedings, we explore the prospects for constraining the parameter space of such a transient model using the event-level arrival-direction data of the Pierre Auger Observatory. To do this, we first present the model in section 2, followed by the method in section 3. The sensitivity of the method is then explored in section 4, and a conclusion is given in section 5.

2. Model for the origin of UHECRs

The model [1] is based on the assumption that UHECR sources are of transient nature and can be hosted in any galaxy i in proportion to its star-formation rate, SFR_i. The flux-limited sample of $\sim 400,000$ galaxies with SFR values from [4] is utilized, in combination with a correction factor $c(d_i, b_i)$ that depends on the distance d_i and Galactic latitude b_i of the galaxy. Additionally, galaxy cloning in the Galactic plane region beyond 11 Mpc, an isotropic component following the full-sky SFR density beyond 350 Mpc, and a correction for magnetic trapping inside galaxy clusters with strong magnetic fields are used; see [1] for details.

The spectrum of UHECR nuclei as they escape from each galaxy is assumed to follow a power law with an exponential cut-off, the parameters of which are determined by a fit to the measured spectrum and shower-depth distributions at Earth, see [1]. The injection rate S_i of each galaxy i is given by $S_i = k \frac{\mathrm{SFR}_i}{c(d_i,b_i)}$, where k represents the *burst rate* per SFR unit, in $M_\odot^{-1}$, a parameter that is assumed to be common to all sources and descriptive of the nature of the transients. Another important parameter to consider is the time spread due to intervening magnetic fields, which delays and stretches the time it takes for UHECRs to reach Earth. As there is only limited knowledge about the extent, strength, and coherence length of the magnetic fields in the intergalactic medium and our local environment, we adopt a generic approach rather than modeling these magnetic fields in detail. The total time spread due to all the intervening magnetic fields is described by the parameter $\Delta\tau$. As shown in [1], this time spread is expected to be dominated by local structures, in particular the Local Sheet in which the Milky Way resides, so that we can neglect the distance dependence of the time spread beyond tens of Mpc. At smaller distances, $\Delta\tau$ is a quadratic function of the distance of the source [5]. The combination of $k \cdot \Delta\tau$ then shapes the contribution of galaxies to the flux at Earth: for small $k \cdot \Delta\tau$, nearby galaxies on average do not contribute to the flux as the observer on Earth is likely between two bursts. For large $k \cdot \Delta\tau$, the emission is almost continuous and the flux is dominated by nearby galaxies. This is visualized in Fig. 1. By comparing the measured flux on the sphere with the model maps obtained with different values of $k \cdot \Delta\tau$, constraints on that quantity can then be determined.



(a) small $k \cdot \Delta\tau$ **(b)** intermediate $k \cdot \Delta\tau$ **(c)** large $k \cdot \Delta\tau$

Figure 1: Modeled median arrival flux at $E > 40$ EeV for different values of $k \cdot \Delta\tau$ with a top-hat blurring of 25°. For small $k \cdot \Delta\tau$ values, the Laniakea supercluster dominates the flux, while for larger $k \cdot \Delta\tau$ values more nearby galaxies are dominant. These nearby galaxies are predominantly found in the Local Sheet, i.e., along the supergalactic plane shown as a dark-red line.

3. Method

To find the best model to describe the data, we compare the measured counts in a given arrival direction, $\text{data}_{E,\text{pix}}$, and the modeled counts, $\text{pdf}_{E,\text{pix}}$, both binned in equal-area healpy pixels [6] (nside=32) and in energy $E \in (10^{19.5}\,\text{eV}, 10^{19.6}\,\text{eV}..., 10^{20.2}\,\text{eV})$, using the following likelihood function:

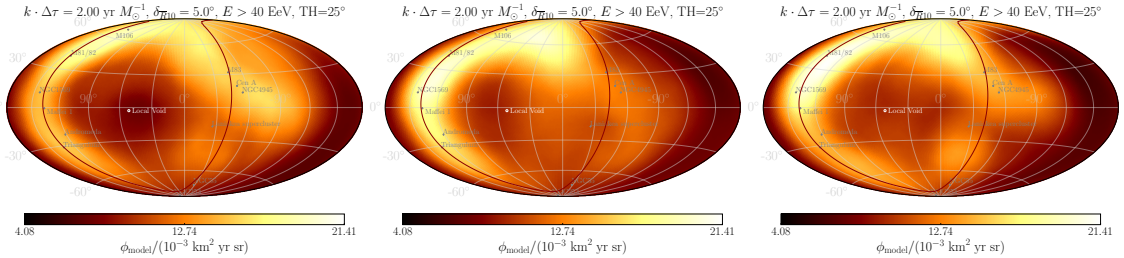
$$\log \mathcal{L} = \sum_E \sum_{\text{pix}} \text{data}_{E,\text{pix}} \log(\text{pdf}_{E,\text{pix}}(k \cdot \Delta\tau, \delta_{\overline{R}_{10}}, \alpha, \text{GMF})). \tag{1}$$

Notably, compared to previous likelihood analyses of the arrival-direction data using models based on catalog sources [2, 3], the likelihood is calculated in energy bins instead of scanning over an energy threshold. As was shown in [7], such a differential approach increases the sensitivity to energy-dependent effects and reduces the penalty factor.

The pdf of the modeled arrival directions depends on several parameters:

- the parameter for the burst rate normalization and time spread, $k \cdot \Delta\tau$, described above;
- a von-Mises-Fisher blurring parameter, $\delta_{\overline{R}_{10}}$, defined at rigidity $R = 10\,\text{EV}$. The blurring applied to the models shrinks as $1/R$, as expected for deflections in a turbulent magnetic field;
- an anisotropic contribution of the flux from the catalog sources, $\alpha \in [0; 1]$, with an isotropic counterpart, $1 - \alpha$, that could e.g. come from another source population. This parameter is inspired by previous studies from the Pierre Auger Collaboration [2, 3], but since the model used in this work aims to describe the total SFR density in the Universe (instead of only the contribution from nearby starburst galaxies in the earlier works), α is assumed to be large and may even be fixed to one.
- a model of the Galactic magnetic field (GMF).

The effect of different GMF models is visualized in Fig. 2. We use the `JF12` model [8] as well as the `UF23` models [9], both in combination with the `JF12` random magnetic field [10] tuned to the updated Planck data [11] ("Pl"). For the `JF12` model, we additionally compare models using no turbulent field ("reg") and the original stronger `JF12` turbulent model ("full").



(a) `JF12` **(b)** `UF23-base` **(c)** `UF23-twistX`

Figure 2: Modeled arrival flux at $E > 40$ EeV for the intermediate value of $k \cdot \Delta\tau = 2\,\mathrm{yr}\,M_\odot^{-1}$ shown in Fig. 1b), including additionally deflections in different models of the Galactic magnetic field.

Compared to the case without GMF shown in Fig. 1, the overdensity regions are displaced by the coherent deflections in Fig. 2, but still aligned with the supergalactic plane. However, the strongest overdensity in the Centaurus region is diluted especially for the `UF23` models for the chosen value of $k \cdot \Delta\tau$.

4. Expected Sensitivity

To test the sensitivity of the method, 500 mock simulations are generated. Each contains events according to a model with $k \cdot \Delta\tau = 2\,\mathrm{yr}\,M_\odot^{-1}$, $\delta_{\overline{R}_{10}} = 5°$, $\alpha = 1$ and varying GMF models. The events are drawn randomly following the energy-dependent event counts and the arrival-direction dependence of the exposure of the Pierre Auger Observatory during its Phase 1 [3].

Magnetic-field model sensitivity: As a first test, we estimate how well a simulated "true" GMF model underlying the mock simulations can be identified by the method. In Fig. 3, the test statistic is shown for each combination of simulated true GMF model and GMF model used in the reconstruction. Here, the test statistic is defined as $\mathrm{TS} = 2(\log\mathcal{L}(k\cdot\Delta\tau, \delta_{\overline{R}_{10}}, \alpha, \mathrm{GMF}) - \log\mathcal{L}(\alpha = 0))$, where the likelihood is maximized over the first three parameters for each GMF model used in the reconstruction.

The lighter diagonal of the matrix shown in Fig. 3 highlights that the correct GMF model always reaches the largest TS value compared to all other models, but not always very distinctively. Confusion is minimal when the simulated truth does not include the GMF (simulated truth = `none`). It is also possible to distinguish `JF12` models (with different random fields) from `UF23` models. However, the level of confusion between the different `UF23` models is higher. Only the `UF23-twistX` model shows enough differences to the others to be reasonably identified within uncertainties. This study leads to the conclusion that it is not necessary to evaluate, in the data analysis, all the GMF models presented in Fig. 3. Instead, a reasonable choice would be to use only no GMF, `JF12-Pl`, `UF23-base` and `UF23-twistX`.

Parameter reconstruction: Since it is highly unlikely that any of the available GMF models is exactly true, it is important to quantify with mock simulations how well the model parameters are reconstructed when the GMF model is not the same as the simulated true one. The results of the

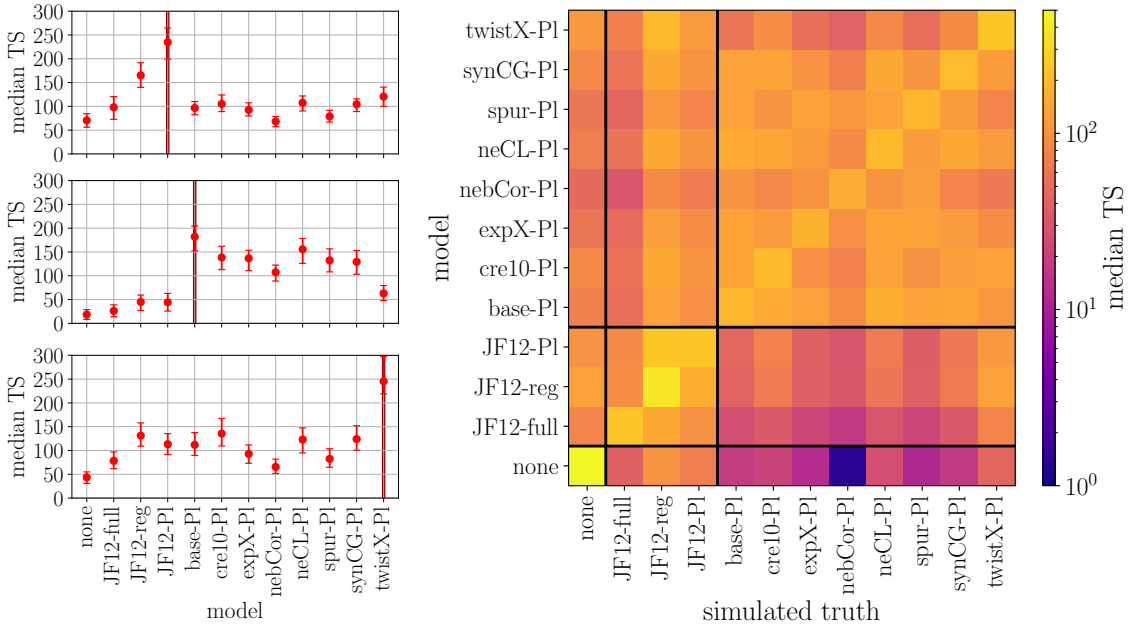


Figure 3: Confusion between GMF models: on the *left*, the median and inner 68% test statistic over 500 mock simulations is shown when using a different GMF models in the reconstruction. The three examples are shown for different simulated true GMF models as indicated by the vertical line in each panel. Using the same GMF model in the simulation and reconstruction leads to the largest TS, as expected. On the *right*, an overview using more GMF models is shown, where the color represents the median TS. The plots on the left represent individual rows of the plot on the right.

study are shown in Fig. 4, for the case of no GMF in the mock simulations on the left-hand side, and for a simulated model following `UF23-synCG` on the right-hand side.

When using the correct GMF model (reconstruction model = simulated truth), the parameters underlying the mock simulations are reconstructed without visible bias. When using a slightly different GMF model (for example another `UF23` model when the simulated truth is `UF23-synCG`), $k \cdot \Delta\tau$ is still reconstructed without sizable bias. However, the blurring parameter, $\delta_{\overline{R}_{10}}$, is reconstructed with a positive bias to compensate for the mismatching deflections. If the GMF model differs substantially from the simulated truth (e.g., no GMF in the reconstruction when the simulated truth is `UF23-synCG`), then $k \cdot \Delta\tau$ may be over- or underestimated, and the blurring is again positively biased. In addition, the anisotropic fraction α shows a negative bias. A correlation between α and $\delta_{\overline{R}_{10}}$ was also seen in previous analyses [2, 3] and is expected: if the data are not a perfect realization of the model, increasing the blurring parameter or the fraction of isotropic events has a similar effect of smoothing the model anisotropy.

Expected performance: To get an idea of the expected performance of the method on the data, another set of mock simulations is generated, which this time is not within the parameter space of the transient model. Instead, the mock simulations are based on the previous finding that UHECR data are well explained when part of the flux comes from a catalog of nearby starburst galaxies (SBGs). In [7], a combined fit to spectrum, shower-depth distributions, and arrival directions measured by

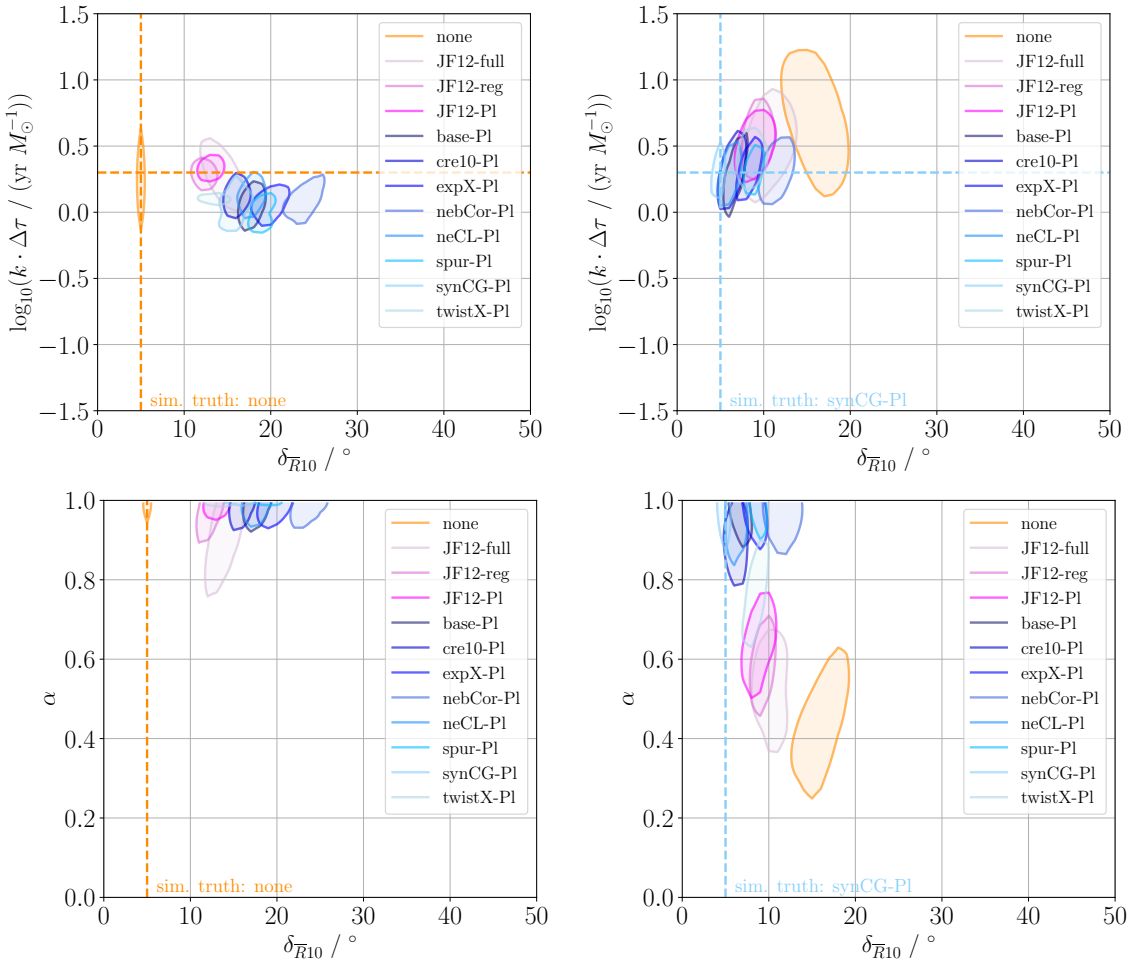


Figure 4: Parameter reconstruction using different magnetic field models on 500 mock simulations with no GMF (*left*) and the `UF23-synCG` (*right*) GMF model. The parameters underlying the mock simulation are indicated by the dashed lines. The contours contain 50% of the distribution calculated via Gaussian kernel-density estimation.

the Pierre Auger Observatory was performed. The authors found that a model where ∼ 17% of the flux at 40 EeV comes from the SBG catalog (where the relative contribution increases with the energy due to propagation effects), in combination with no coherent GMF model and a blurring parameter of $\delta_{\overline{R}_{10}} = 23°$, describes the data better than a homogeneous and isotropic model by 4.5σ. Based on those best-fit parameters, again 500 mock simulations are generated following both the event numbers and exposure of the Pierre Auger Observatory.

The TS values obtained by applying the transient model to SBG mock simulations are shown in orange in Fig. 5. For comparison, the blue line shows the TS values obtained by applying the SBG model to measured data instead.[1] The test-statistic values obtained by applying the transient model to the SBG simulations are in the same range as those obtained by applying the SBG model

[1] Fig. 5 shows only the arrival-direction test statistic defined in eq. 3.5 of [7]. Note that the definition of the arrival-direction likelihood in [7] is mathematically equivalent to the likelihood used in this work (eq. 1). Only the underlying source models differ.

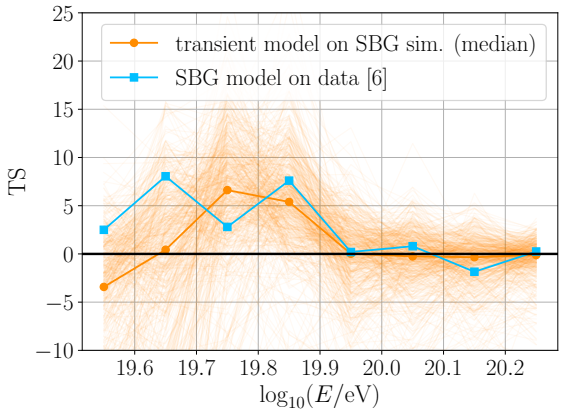


Figure 5: Test statistic as a function of energy (the sum over all energy bins then gives the total TS). A transient model without GMF is evaluated on mock simulations assuming that 17% of the UHECR flux at 40 EeV comes from nearby SBGs, as shown in orange. The thick lines denotes the median and the thin lines individual simulations. The test statistic obtained using instead the SBG model on arrival-direction data from the Pierre Auger Observatory is shown in blue (from [7], Fig. 9).

to the data. We can therefore expect the transient model to describe the data as well as, if not better than, the model that includes only the 40 brightest SBGs. The evaluation of the transient model on the data would also lead to a more complete explanation of the observed correlation with galaxies with high star-formation rates.

The reconstructed parameters for the 500 SBG mock simulations are shown in Fig. 6. The blurring used in the mock simulations can be reconstructed without sizable bias (even when a coherent magnetic field model is used), albeit with worse resolution than for the transient mock simulations shown above. The reconstructed anisotropic fractions include the value $\alpha = 1$ but show a large spread, suggesting that this parameter may be underconstrained. We also tested fixing $\alpha = 1$ in the transient model, which leads to a small reduction of the total TS value by about 1.5 units and to an improved resolution on the blurring parameter. This suggests that 100% of the data may be explained by the transient model presented here, including not only the 17% contribution from bright nearby SBGs but also the 83% isotropic background at 40 EeV that is not explained by the SBG model.

Finally, we conducted tests on the number of energy bins to include in the analysis. If we extend the energy range from $> 10^{19.5}$ eV (~ 32 EeV) to $> 10^{19.2}$ eV (~ 16 EeV), as used in [7], the total TS value remains approximately constant when using the SBG mock simulations. In order not to miss parts of the signal when applying the presented analysis to data [7], it may be beneficial to use a minimum energy of 16 EeV. At even lower energies the dipole becomes significant [12], which may or may not originate from the same source class. Compatibility of the best-fit transient model with the dipole could be tested in a second step.

5. Conclusion

In this study, we examined a possible explanation for the observed flux of UHECRs, in which UHECRs are emitted by transient events that occur in proportion to the star-formation rate of their

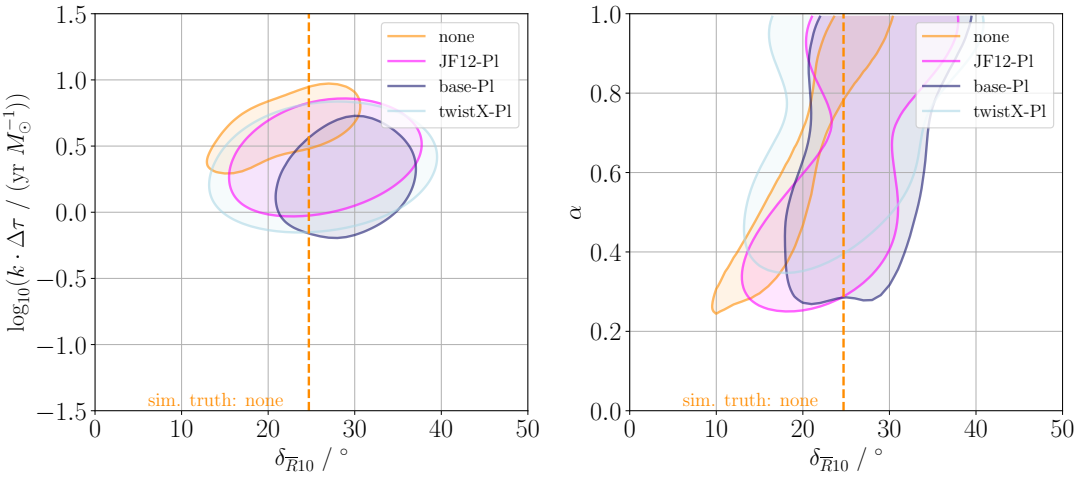


Figure 6: Parameter reconstruction using different GMF models for simulations following the best-fit SBG model without GMF deflections from [7]. The blurring of the mock simulation is indicated by the dashed lines. The contours contain 50% of the distribution calculated via Gaussian kernel-density estimation.

host galaxies. We propose a likelihood method to test the model on event-level data measured at the Pierre Auger Observatory. We have investigated the reliability of the parameter estimation, the impact of the Galactic magnetic field model, and the expected performance on the data. We have shown that this method can be used to constrain the contribution of transient sources to the UHECR flux, and to obtain information about the interaction of UHECRs with Galactic and extragalactic magnetic fields. Unlike previous work that estimated the possible contribution of limited catalogs of nearby candidate sources, this model could explain the entire UHECR flux at once. The observed over- and under-densities on the UHECR sky would then be naturally explained by the spatial distribution of the sources and the rate of the transients. By constraining this rate with the proposed method, information on the possible types of transients responsible for UHECR production can be obtained, enabling a thorough evaluation of transients as sources of UHECRs.

Acknowledgements: We thank Michael Unger for providing the lenses to model the deflections in the `UF23` [9] magnetic field models.

References

[1] S. Marafico *et al.*, ApJ **972** (2024) 4
[2] A. Aab *et al.* [Pierre Auger Collaboration], ApJL **853** (2018) L29
[3] P. Abreu *et al.* [Pierre Auger Collaboration] ApJ **935** (2022) 170
[4] J. Biteau, ApJS **256** (2021) 15
[5] A. Achterberg *et al.*, arXiv:astro-ph/9907060 (1999)
[6] A. Zonca *et al.*, J. Open Source Softw. **4** (2019) 35
[7] A. Abdul Halim *et al.* [Pierre Auger Collaboration], JCAP **01** (2024) 022
[8] R. Jansson and G. Farrar, ApJ **757** (2012) 14
[9] M. Unger and G. Farrar, ApJ **970** (2024) 95
[10] R. Jansson and G. Farrar, ApJL **761** (2012) L11
[11] R. Adam *et al.* [Planck Collaboration], A&A **596** (2016) A103
[12] A. Abdul Halim *et al.* [Pierre Auger Collaboration], ApJ **976** (2024) 48