Effects of the Galactic Magnetic Field on the UHECR Correlation Studies with Starburst Galaxies

Ryo Higuchi1, Takashi Sako2, Toshihiro Fujii1,4, Kazumasa Kawata2, and Eiji Kido1

1 Astrophysical Big Bang Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan; ryohiguchi@riken.jp
2 Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan
3 Osaka Metropolitan University, 3-3-138 Sugimoto Sumiyoshi-ku, Osaka-shi, 558-8585, Japan
4 Nambu Yoichiro Institute of Theoretical and Experimental Physics, 3-3-138 Sugimoto Sumiyoshi-ku, Osaka 558-8585, Japan

Received 2022 August 26; revised 2023 March 22; accepted 2023 March 22; published 2023 June 2

Abstract
We estimate the biases caused by the coherent deflection of cosmic rays due to the Galactic magnetic field (GMF) in maximum likelihood analyses for searches of ultrahigh-energy cosmic-ray sources in the literature. We simulate mock event data sets with a set of assumptions for the starburst galaxy source model, coherent deflection by a GMF model, and mixed-mass composition. We then conduct a maximum likelihood analysis without accounting for the GMF in the same manner as previous studies. We find that the anisotropic fraction $f_{\text{ani}}$ is estimated to be systematically lower than the true value. We estimate the true parameters that are compatible with the best-fit parameters that were reported, and find that except for a narrow region with a large anisotropic fraction and a small separation angular scale, a wide parameter space is still compatible with the experimental results. We also develop a maximum likelihood method that takes the GMF model into account and confirm in Monte Carlo simulations that we can estimate the true parameters within a 1σ contour under the ideal condition that we know the event-by-event mass and the GMF.

Unified Astronomy Thesaurus concepts: Cosmic ray astronomy (324); Cosmic ray sources (328); Cosmic rays (329); Ultra-high-energy cosmic radiation (1733)

1. Introduction
Cosmic rays (CRs) are high-energy nuclei that exist throughout the universe. Specifically, ultrahigh-energy cosmic rays (UHECRs) with energies of about 100 EeV ($10^{20}$ eV) are observed, but their origin is not yet known. There are two leading experiments that observe UHECRs: the Telescope Array (TA) experiment (Kawai et al. 2008; Sagawa 2020; it is located in the U.S.A., 39.3 deg N, 112.9 deg W) and the Auger experiment (Aab et al. 2015a; it is located in Argentina, 35.2 deg S, 69.5 deg W), which cover the sky in the northern and southern hemispheres, respectively. Thanks to a large number of UHECR events observed through these experiments in the past decade, the arrival directions of UHECRs are found to be anisotropic in the intermediate angular ($\sim 10^6$ to $\sim 20^6$) scale (Abreu et al. 2012; Abbasi et al. 2014; Aab et al. 2015b), which is believed to give us keys to knowing the UHECR origins. Due to the energy loss through the photopion production or photoneutron dissociation, most UHECRs cannot propagate more than 30 to 100 Mpc (Greisen 1966; Zatsepin & Kuzmin 1966; the GZK limit). Consequently, as astronomical candidates of UHECR sources, it is natural to consider nearby extragalactic high-energy objects. Previous studies have investigated the correlation between the arrival directions of UHECRs and their source candidates (Abreu et al. 2007, 2010; He et al. 2016; Aab et al. 2018; Abbasi et al. 2018). Especially, recent studies (Abreu et al. 2010; Aab et al. 2018; Abbasi et al. 2018) try to explain the arrival direction of UHECRs by the weighted sum of events originating from sources and isotropic backgrounds (see also the review in Batista et al. 2019). In these studies, the flux of UHECRs (CR flux model) is composed of the source-originated flux (source-flux model) and the isotropic backgrounds (isotropic flux model). Based on a catalog of source candidates, the source flux model is constructed as a superposition of the Gaussian-smearred angular distributions of an individual point source. Previous studies introduced the anisotropic fraction $f_{\text{ani}}$ as the fraction of the source contribution to the total CR flux and the separation angular scale $\theta$ as the scale of the Gaussian smearing. The separation angular scale $\theta$ is considered to reflect the deflections and scattering by the Galactic and extragalactic magnetic fields (EGMF). The parameters ($f_{\text{ani}}$, $\theta$) are searched to best fit the observed CR angular distribution (see Section 3 for details).

One of the possible source-flux models suggested by previous studies is that UHECRs originate from nearby starburst galaxies (SBGs). Aab et al. (2018) investigated the correlations between UHECRs observed by the Auger experiment and the CR flux models constructed with nearby extragalactic high-energy objects such as SBGs, active galactic nuclei, and gamma-ray bursts. They reported that the best correlation was found above 39 EeV with the SBG source-flux model (SBG model), and the best-fit anisotropic fraction $f_{\text{ani}}^{\text{Auger}}$ and the separation angular scale $\theta^{\text{Auger}}$ were estimated to be 9.7% and 12.29, respectively (we call this best-fit model the Auger best-fit model).

The TA experiment studied the correlation between the UHECR arrival directions observed in the northern sky and the source-flux model with SBGs using the best-fit parameters reported by Aab et al. (2018). It concluded that the UHECR arrival directions are compatible with both the isotropic distribution ($f_{\text{ani}} = 0$%) and the Auger best-fit model (Aab et al. 2018) with current statistics (Abbasi et al. 2018). While the nearby SBGs belong to the group of attractive candidates for the UHECR sources, it is surprising that even this best-fit source model can explain only $\sim 10\%$ of the observed UHECRs. This question motivated us to study possible biases...
that cause an underestimate of \( f_{\text{ani}} \) and to estimate a realistic constraint of \( f_{\text{ani}} \) deduced from the observations.

The isotropic scattering approximation in previous studies does not reflect the actual structure of the GMF, which deflects UHECR trajectories in a certain direction (coherent deflection). Current analyses treating the coherent deflection as a part of isotropic scattering may result in a smaller \( f_{\text{ani}} \) and a larger \( \theta \) than the true values. In this study, we call these systematic effects on the parameter estimation caused by the GMF “the GMF bias.” To consider the GMF bias (in addition to introducing the GMF model), we take the following two components into account: (1) The dependence on the arrival direction. The coherent deflection around the Galactic center (GC) and the Galactic plane (GP) is much larger than in other regions of the sky. This dependence also affects analyses whose samples are divided into the northern sky (TA experiment) and the southern sky (Auger experiment). Generally, UHECRs observed in the southern sky should be more affected by the GMF than those in the northern sky. These effects, caused by the limitation of the sky coverage of each experiment, are not evaluated in the previous studies. (2) The dependence on the rigidity \( R = p_c/|Ze| \sim E/|Ze| \). Here \( p_c \), \( E \), and \( Ze \) represent the particle momentum, energy, and electric charge, respectively. The approximation is valid in the ultrarelativistic regime considered here. Because the deflection angle is proportional to \( 1/R \) of each CR, the magnetic field effect is strongly coupled with the energy spectrum and the mass composition. The Auger experiment suggests that the mass composition of UHECRs is higher at higher energy (Aab et al. 2017; Batista et al. 2019; Heinze & Fedynitch 2019).

In this study, we investigate the GMF bias in previous studies by applying a commonly used GMF model (Jansson & Farrar 2012a, 2012b) by means of Monte Carlo (MC) simulations. We generate mock event data sets assuming the true parameters \( (f_{\text{ani}}, \theta_{\text{true}}) \), taking coherent deflections by the GMF and a mass model into account (a mixed-mass spectrum model proposed in Heinze & Fedynitch 2019). For some samples of mock events, we demonstrate that the event arrival directions are apparently displaced from the real source directions. The size of the displacements strongly depends on the direction of the sky and the rigidity. Then, we applied a maximum likelihood analysis to the mock events in the same manner as in previous studies. In order to focus on the bias in the previous analysis (Aab et al. 2018; Abbasi et al. 2018), we fix the source-flux model to be the SBG model. Based on the analyses of the mock data sets, we discuss the biases in the parameter estimation separately in the northern and southern hemispheres and in the all-sky regions. We also develop an analysis technique to reduce the GMF bias.

2. Mock Data-set Production

The mock event data sets are generated under a set of assumptions (the SBG model, the GMF model, and the mass-dependent energy spectrum) with the flow shown in Figure 1. The construction of the CR flux models before and after considering the GMF deflection is described in Sections 2.1 and 2.2, respectively. To generate mock events taking the coherent deflections caused by the GMF into account, one needs to assign a rigidity \( R \), i.e., energy and mass, to each event. We use a mixed-mass assumption to reflect a realistic situation as described in Section 2.3. Results with single-mass assumptions are summarized in Appendix A.1. In all cases, we generate 1000 data sets, each of which contains 4000 mock events across the whole sky (all-sky data set). To compare the data sets with the observed UHECRs by the TA and Auger experiments, we select in Section 2.4 the north-sky and south-sky data sets from the all-sky data sets, taking the sky coverage of each experiment into account.

2.1. CR Flux Model that Originated from the SBG Model

With the source model fixed to the SBG model, the CR flux model is constructed with the assumption of a two-parameter set \( (f_{\text{ani}}, \theta) \). In Aab et al. (2018) and Abbasi et al. (2018), the CR flux model \( F_{\text{org}}(\mathbf{n}, \theta) \) is determined as the superposition of the von Mises-Fisher function (Fisher 1953; the Gaussian distribution on the sphere) of each source,

\[
F_{\text{org}}(\mathbf{n}, \theta) = \frac{\sum_i f_i \exp(\mathbf{n} \cdot \mathbf{n}/\theta^2)}{\int_{4\pi} \sum_i f_i \exp(\mathbf{n} \cdot \mathbf{n}/\theta^2) d\Omega}.
\]

Note that \( i \) indicates each SBG, and \( \mathbf{n}_i \) and \( f_i \) mean its direction and relative flux (contribution from each source), respectively. Table 1 is the list of SBGs in our SBG model defined in Aab et al. (2018), which contains the values of \( f_i \) and \( \mathbf{n}_i \). The relative flux of SBGs \( f_i \) in Table 1 is determined by their continuum radio flux. The source directions and relative contributions in Table 1 are visualized in Figure 2. We can see that most SBGs are located along the super-galactic plane (SGP) and the top-4 contributions of SBGs dominate with \( \sim 60\% \) of the total flux.

An example of the SBG model \( F_{\text{org}}(\mathbf{n}, \theta = 10^6) \) is presented in Figure 3. As expected from Table 1 and Figure 2, a small number (~10) of sources dominates the distribution. The normalized CR flux model \( F_{\text{norm}} \) is defined as the weighted sum of the SBG model \( F_{\text{org}} \) and the isotropic flux model \( F_{\text{iso}} \),

\[
F_{\text{norm}}(\mathbf{n}, f_{\text{ani}}, \theta) = f_{\text{ani}} F_{\text{org}}(\mathbf{n}, \theta) + (1 - f_{\text{ani}}) F_{\text{iso}}
\]

\[
F'_{\text{org}}(\mathbf{n}, \theta) = \int_{4\pi} F_{\text{org}}(\mathbf{n}, \theta) d\Omega = 1/4\pi.
\]

2.2. Flux Mapping

At this stage, the flux model does not include the GMF deflections. To include them, we use a back-propagation technique using the CR propagation code CRPropa3 (Batista et al. 2016). We adopt the JF12 model (Jansson & Farrar 2012a, 2012b) as the GMF model. We calculate the trajectories of antiprotons of energy \( E \) emitted from the Earth to a sphere of 20 kpc radius from the GC (the galaxy sphere). The position of the Earth is defined to be 8.5 kpc away from the GC, following the JF12 model. These trajectories represent the trajectories of particles with the same rigidity that can arrive on the Earth through the Galactic sphere. The trajectory of a higher-mass particle with the charge \( Ze \) is replaced with the trajectory of a proton with the rigidity \( R = E/Ze \).

Based on the CR trajectories obtained through the back-propagation, we convert the CR flux model on the galaxy sphere into that on Earth through the GMF model (flux mapping). We define the original CR flux model as \( F_{\text{org}}(\mathbf{n}_{\text{org}}, \theta) \), where \( \mathbf{n}_{\text{org}} \) indicates the direction on the galaxy sphere. The conversion from the directions on the Earth \( \mathbf{n}_{\text{earth}} \) to those on the galaxy sphere \( \mathbf{n}_{\text{org}} \) of particles with rigidity \( R \) is expressed as

\[
\mathbf{n}_{\text{org}} = \mathbf{A}_{\text{BT}}(\mathbf{n}_{\text{earth}}, R),
\]

where \( \mathbf{A}_{\text{BT}} \) is the back-propagation matrix and \( R = E/Ze \) and \( E = Ze \).

2.3. Parameter Estimation in the GMF Model

The parameter estimation in the real data sets is carried out using the maximum likelihood method. The pixel values \( \lambda_{\text{pix}} \) for each pixel are given by

\[
\lambda_{\text{pix}} = \frac{1}{\sigma_{\text{pix}}} \int_{\text{sky}} \left( \frac{1}{2} \ln Z + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \left( \frac{\mathbf{n} \cdot \mathbf{n}}{\theta^2} \right)^2 \right) d\Omega.
\]

where the number of events in each direction \( i \) is \( N_i \), the number of trials in each direction is \( n_i \), and the sky coverage fraction is \( \Delta \Omega \).

The parameter estimation is carried out separately in the northern and southern hemispheres and in the all-sky regions. In order to focus on the bias in the parameter estimation, we compare the results with the previous analysis, which was carried out without considering the GMF effect.
where $A_{\text{BT}}$ indicates the conversion function. As Liouville’s theorem tells that the flux value along each CR trajectory remains constant (Bradt & Olbert 2008), we can determine the CR flux on Earth $F_{\text{earth}}$ as

$$F_{\text{earth}} = F_{\text{org}} n_{\text{earth}} (n_{\text{org}}, \theta).$$

Because of the nearby source contributions, the photonuclear interaction during intergalactic propagation is not taken into account in this study. Some examples of $F_{\text{earth}}$ for different $R$ based on $F_{\text{org}}$ in Figure 3 are shown in Figure 4. As shown in Figure 4 (top left), at the highest rigidity ($R = 100$ EV), the GMF does not affect the CR flux. As the rigidity $R$ decreases, the peaks of CR flux around NGC1068 and NGC253 become displaced from the true source directions (top right panel in Figure 4) for $\log(R/\text{EV}) = 1.5$. At lower rigidity (bottom left panel in Figure 4 for $R = 10$ EV), the displacements become larger and the peak around NGC4945 splits along the GP. Through visual inspections, it is clear that the GMF bias is stronger in the southern hemisphere.

2.3. Generation of Mock Data Sets

To quantitatively discuss the GMF biases in Section 3, we generate mock events as follows. We adopt a best-fit function and the parameters given in Heinze & Fedynitch (2019) based on the observed UHECRs from the Auger experiment (Aab et al. 2017). In Heinze & Fedynitch (2019), the energy spectrum of each mass $A$ at the source is assumed by the following function $J_A$, and fitting was performed for the mass composition observed on Earth,

$$J_A(E) = J_{A_{\text{cut}}} (E, Z_A, R_{\text{max}}) n_{\text{evol}} (z) \left( \frac{E}{10^9 \text{ GeV}} \right)^{-\gamma}. \quad (5)$$
The cutoff function $f_{\text{cut}}$ is given as

$$f_{\text{cut}} = \begin{cases} 1 & (E < Z_A R_{\text{max}}) \\ \exp \left(1 - \frac{E}{Z_A R_{\text{max}}} \right) & (E > Z_A R_{\text{max}}). \end{cases}$$

Because we only focus on the nearby sources, the redshift-evolution term $n_{\text{redvol}}(z)$ is approximated to be 1. We also assume that the mass composition observed on Earth and the composition at the source are the same. The fractions of elements are defined as $f_A = J_A / \sum_A J_A$ at 10 EeV in Heinze & Fedynitch (2019). We adapt the best-fit parameters from Heinze & Fedynitch (2019) as $\gamma = -0.80$ and $R_{\text{max}} = 1.6$ EV. We also adapt the values of $f_A$ as $(^1\text{H}, ^4\text{He}, ^{14}\text{N}, ^{28}\text{Si}, ^{56}\text{Fe}) = (0.0, 82.0, 17.3, 0.6, 2.0 \cdot 10^{-2})[\%]$. According to this mass fraction and to the spectra, we determine the mass and energy, i.e., $R$, of each mock event. The energy $E$ of a mock event is randomly sampled from the spectrum, and the rigidity $R$ of the event is calculated through the formula $R = E / Z_e$. We adapt the minimum energy $E_{\text{min}} = 40$ EV according to previous studies (Aab et al. 2018, $E_{\text{min}} = 39$ EeV). Using the selected rigidity $R$, we determine the arrival direction of the event based on the CR flux model defined in Section 2.2.

An example of the distribution of the mock event arrival directions is provided in Figure 5. The distribution in Figure 5 is similar to the distribution of pure-carbon case (Figure 10 in Appendix A.1). Although we can see clustering around M82 and NGC4945, the centers of the distributions are displaced from the source directions. The events that originated from NGC1068 and NGC253 are mostly deflected. This suggests that the real UHECR distribution cannot be reproduced with a single isotropic smearing, and the deflections by the GMF depend on the arrival directions.

$$\omega(\delta) \propto \cos(a_0) \cos(\delta) \sin(\alpha_m) + \alpha_m \sin(a_0) \sin(\delta).$$

$$\alpha_m = \begin{cases} 0 & (\xi > 1) \\ \pi & (\xi < -1) \\ \cos^{-1}(\xi) & (-1 < \xi < 1) \end{cases}$$

$$\xi = \frac{\cos(\theta_m) - \sin(a_0) \sin(\delta)}{\cos(a_0) \cos(\delta)}.$$
where $\theta_m$ is the maximum zenith angle and $a_0$ is the latitude of the experimental site. We adopt the latitude $a_0 = 39.3^\circ \pm 35.2^\circ$ and the maximum zenith angle $\theta_m = 55^\circ$ (60°) for the TA (Auger) experiment. From the all-sky data sets, we randomly select mock events with the probability of each sky coverage. Out of the 4000 mock events in each data set, approximately 1000 mock events each are selected by the TA and Auger coverage. We define the data set selected by the sky coverage of TA (Auger) as the north-sky (south-sky) data set.

### 3. Analysis

In order to investigate how much GMF deflections affect the estimated parameters, we conduct the same maximum likelihood analysis as in Aab et al. (2018) on the mock data sets. We test two hypotheses for the CR flux models. One is a flux with nonzero $f_{\text{ani}}$ with the SBG model ($F_{\text{norm}}$), and the other is the isotropic flux ($F_{\text{iso}}$), i.e., $f_{\text{ani}} = 0$. The test statistics $TS$ are calculated as the log-likelihood ratio,

$$TS = 2 \ln (L(F_{\text{norm}})/L(F_{\text{iso}})).$$

The likelihood of each model $L(F)$ is given as

$$L(F) = \prod_{\text{CR}} \int_{4\pi} F(n) \omega(n) \, d\Omega,$$

where $F$, $\omega(n_{\text{CR}})$, and $n_{\text{CR}}$ are the normalized CR flux model, the sky coverage of each experiment (Equations (7)), and the arrival directions of the observed UHECRs, respectively.

By scanning the set of parameters $(f_{\text{ani}}, \theta)$, the best-fit parameters that maximize the $TS$ in Equation (10) are determined.

### 4. Results

We show the distribution of the best-fit parameters for 1000 mock event data sets in Figures 6. The different panels show the results for the different true parameters $(f_{\text{ani}}^{\text{true}}, \theta^{\text{true}})$ at the top of each panel and with the gray star in each panel. From the left to right column, the true parameter $f_{\text{ani}}^{\text{true}}$ is given as 20%, 40%, and 60%. From the top to the bottom row, the true parameter $\theta$ is 10°, 20°, and 30°. The black, blue, and red contours indicate the 68% and 95% percentile containment for the all-sky, north-, and south-sky data sets, respectively. The black cross, blue circle, and red triangle show the most frequent values $(f_{\text{ani}}, \theta)$ for the all-sky, north-, and south-sky data sets, respectively. The distributions of the best-fit parameters are smoothed with a kernel-Gaussian distribution. The best-fit parameter $(f_{\text{ani}}^{\text{Auger}}, \theta^{\text{Auger}}) = (9.7\% \pm 12.9^\circ)$ in Aab et al. (2018) is shown as the black triangle.
for the distributions for the all-sky, north-sky, and south-sky data sets, respectively. The black cross, blue circle, and red triangle show the most frequent values $q_{\text{ani}}$ of the best-fit parameters of the all-sky, north-sky, and south-sky data sets, respectively.

The distributions of the all-sky, north-sky, and south-sky data sets do not agree with each other, especially for a higher anisotropic fraction $f_{\text{ani}}$ true, (the right panels). When the anisotropic fraction $f_{\text{ani}}$ true is larger and separation angular scale $\theta$ true is smaller (the four upper right panels), the estimated separation angular scale $\theta$ becomes larger due to the deflection of the GMF. When the anisotropic fraction $f_{\text{ani}}$ true is smaller (left panels), the distribution of the all-sky, north-, and south-sky data sets are similar due to the low contrast between the SBG model and the isotropic backgrounds.

Focusing on the north–south difference, the GMF affects the results of south-sky data sets more than north-sky data sets. This can be explained with two reasons: first, the GMF deflections are larger around the GC. In the rest of the sky, the GMF deflection is larger in the Galactic South (see also Figure 11 in Jansson & Farrar 2012a). NGC253 and NGC1068, which contribute to the south-sky data sets, are located in the region of the sky in which the GMF deflections are large. In any parameter set $(f_{\text{ani}}^{\text{true}}, \theta^{\text{true}})$, it is found that the most frequent values of the anisotropic fraction $f_{\text{ani}}$ for the south-sky data sets (red triangles) are largely underestimated, namely below 50% of the true ones $f_{\text{ani}}^{\text{true}}$ (gray stars). Regardless of the true parameters $(f_{\text{ani}}^{\text{true}}, \theta^{\text{true}})$, the distributions for the south-sky data sets include the best-fit parameters $(f_{\text{ani}}^{\text{Auger}}, \theta^{\text{Auger}}) = (9.7\%, 12.9\, \text{deg})$ in Aab et al. (2018) is shown as the black triangle. Note that the mock event data sets are generated with a set of assumptions for the source model (SBG), the GMF model (the JF12), and the mass-composition model (Heinze & Fedynitch 2019).

Figure 7. Examples of the distributions of the best-fit parameters estimated from the mock event data sets. Red contours show the 68 and 95% percentile containments of the best-fit parameters. The red triangle shows the most frequent value of the best-fit parameters $q_{\text{ani}}$ of the mock event data sets. The best-fit parameter $(f_{\text{ani}}^{\text{Auger}}, \theta^{\text{Auger}}) = (9.7\%, 12.9\, \text{deg})$ in Aab et al. (2018) is shown as the black triangle. Note that the mock event data sets are generated with a set of assumptions for the source model (SBG), the GMF model (the JF12), and the mass-composition model (Heinze & Fedynitch 2019).

5. Discussions

5.1. The Uncertainty of the GMF Models

In this section, we discuss the uncertainty effect on the GMF models. To test the effects caused by the uncertainty of the halo components of the JF12 model, we conduct the same analysis, but change the halo components in the model within $1\sigma$ uncertainty, generate the mock event data sets, and repeat the analysis. It is found that the uncertainty of the halo components
We generate 4000 south-sky data sets in the same manner as in Section 2.3, but each has the same number of events (894 events) as was used in the analysis of Aab et al. (2018). These mock event data sets are analyzed in the same manner, and the best-fit parameters are obtained. Examples of the distributions of 4000 best-fit parameters are shown in Figure 7. Because of the GMF bias and the statistical fluctuation regardless of the true parameters \( f_{\text{true}} \) and \( \theta_{\text{true}} \), marked by the gray star, the estimated parameters tend to be distributed around the Auger best-fit parameters \( f_{\text{Auger}} \) and \( \theta_{\text{Auger}} \), marked by the black triangle within the 68% or 95% containment levels. We classify the true parameters \( f_{\text{true}} \) and \( \theta_{\text{true}} \) according to whether \( (f_{\text{Auger}}, \theta_{\text{Auger}}) \) is contained in the 68% or 95% contours. Figure 8 shows the result of this classification. From Figure 8, except in the bottom right corner, a wide range of parameters is still compatible with \( (f_{\text{Auger}}, \theta_{\text{Auger}}) \). Considering the GMF effect and the mass-dependent energy spectrum, a large contribution of SBGs to the UHECR flux is still possible.

### 5.3. Maximum Likelihood Analysis Method with the CR Flux Models on the Earth

To calculate the likelihood (Equation (11)) and TS (Equation (10)), we use the original CR flux model \( F_{\text{org}} \) instead of the CR flux model on the Earth \( F_{\text{earth}} \) (Equation (2)). This causes the GMF bias in the parameter estimations. To reduce the GMF bias in the previous parameter estimation, it is necessary to replace \( F_{\text{org}} \) with \( F_{\text{earth}} \) in Equation (2). Note that this analysis is valid only when the GMF and mass-dependent spectrum models are correct. In other words, we need to test a set of assumptions together. We rewrite Equation (2) as follows:

\[
F'_{\text{norm}}(n, f_{\text{ani}}, \theta, R) = f_{\text{ani}} F'_{\text{earth}}(n, \theta, R) + (1 - f_{\text{ani}}) F_{\text{iso}}.
\]

\[
F'_{\text{earth}}(n, \theta, R) = \frac{F_{\text{earth}}(n, \theta, R)}{\int_{4\pi} F_{\text{earth}}(n, \theta, R) d\Omega} = 1/4\pi.
\]

Here, \( F_{\text{earth}}(n, \theta, R) \) is obtained using Equation (4). Thus, we can rewrite the CR flux models from the sources \( F'_{\text{earth}}(n, f_{\text{ani}}, \theta, R) \) as

\[
F'_{\text{earth}}(n_{\text{CR}}, f_{\text{ani}}, \theta, R_{\text{CR}}) = \frac{F_{\text{org}}(A_{BT}(n_{\text{CR}}, R_{\text{CR}}), \theta)}{\int_{4\pi} F_{\text{org}}(A_{BT}(n, R), \theta) d\Omega}.
\]

The numerator \( \int_{4\pi} F_{\text{org}}(A_{BT}(n, R), \theta) d\Omega \) in Equation (13) is derived by integrating \( F_{\text{earth}} \).

We conduct the improved maximum likelihood analysis following Equation (12) to the same data sets as in Section 4. Note that the new analysis is carried out assuming that we
know the event-by-event mass of each mock event. Figure 9 illustrates the results in the same manner as Figure 6, but for the estimation following Equation (12). In all cases, the analysis improves the estimates of the parameters \((f_{\text{ani}}, \theta)\). Specifically, the GMF bias is reduced when the separation angular scales \(\theta\) are small. For a larger separation angular scale, there is still a significant difference between the estimated parameters and the true parameters \((f_{\text{ani}}^{\text{true}}, \theta_{\text{true}})\). Because the GMF bias caused by the regular component of the GMF is effectively reduced within the 1\(\sigma\) contour, the origins of the dispersion should be the statistical fluctuation. Future observations with large statistics are expected to reduce the dispersion. In this analysis, we assumed a perfect event-by-event energy and (rigidity) resolution, but the effect of realistic resolutions will be discussed in a future publication.

### 5.4. Discussions of the Random GMF and EGMF

In this section, we refer to the components of the magnetic fields that are not fully considered in this study. To reveal the effects of coherent deflection by the GMF, we only focus on the regular component of the GMF. We assume that the separation angular scale \(\theta\) includes both the EGMF and a random component of the GMF (random GMF). Aab et al. (2018) and Abbasi et al. (2018) also included regular components of the GMF. In general, the deflections by the EGMF and the random GMF also should have a rigidity dependence. Bray & Scaife (2018) suggested that the upper limit to the EGMF lies at the nano-Gauss scale. Although the upper limit to the EGMF is smaller than that of the GMF, the distance between each source and the Earth is much larger than the radius of our galaxy. We need to consider this distance dependence (see also Anchordoqui 2019). A random component of the GMF has an arrival-direction dependence that is the same as for the regular component of the GMF. Pshirkov et al. (2013) investigated the random deflections in the GMF. They suggested that the random component of the GMF deflects 40 EeV protons by less than 1°–2° in most of the sky and ∼5° along the GP.

Although the physically correct description of random components of the GMF and EGMF is important, it is beyond the scope of this study. The effect of random components in the GMF and EGMF should also have an arrival-direction dependence and rigidity dependence. We will take them into account in a future realistic model.
6. Summary

We estimate the biases caused by the coherent deflections due to the Galactic magnetic field in searches for UHECR sources in the literature. We generated mock event data sets with a set of assumptions for the source model (Aab et al. 2018), coherent deflection by the GMF model (Jansson & Farrar 2012a, 2012b), and a mass-composition model (Heinze & Fedynitch 2019), and conducted a maximum likelihood analysis on the data sets neglecting the GMF in the same manner as in previous studies. Our main results are listed below.

1. The distributions of the estimated parameters \(f_{\text{ani}}, \theta\) are displaced from the true parameters \(f_{\text{ani}}^{\text{true}}, \theta^{\text{true}}\). This confirms the existence of the GMF bias.
2. The distributions of the estimated parameters \(f_{\text{ani}}, \theta\) in the all-sky, north-sky, and south-sky data sets do not agree with each other. The directional or sky dependence of the GMF bias is also confirmed.
3. We find that the estimated \(f_{\text{ani}}\) is systematically reduced by more than 50\% in the south-sky data sets.
4. We search for the true parameters \(f_{\text{ani}}^{\text{true}}, \theta^{\text{true}}\) that are compatible with the best-fit parameters reported in Aab et al. (2018), taking the number of events used in their study into account. Except for the narrow region with the large anisotropic fraction \(f_{\text{ani}}^{\text{true}}\) and small separation angular scale \(\theta^{\text{true}}\), a wide parameter space is still compatible with the experimental result within the 95\% confidence limit.
5. We develop a maximum likelihood analysis that takes the GMF deflections into account and confirm that the parameters would be correctly estimated within the 1\(\sigma\) contour under the ideal condition that we know the event-by-event energy and mass of each UHECR event and the GMF structure.

Note that again this study is conducted under a specific set of assumptions: the source model (Aab et al. 2018), the magnetic field model (Jansson & Farrar 2012a, 2012b), the energy spectrum, and the mass composition (Tsunesada et al. 2017; Heinze & Fedynitch 2019).

Although these models and the assumptions are to be tested and updated regularly, the technique in Section 5.3 can be applied to future models and updated observational data sets. The extension of the TA and Auger experiments (Castellina & Pierre Auger Collaboration 2019; Abbasi et al. 2021, TA \(\times 4\) and AugerPrime) and next-generation UHECR observation (Hörandel 2021) will play an important role. The improvement of the GMF and CR propagation models also leads us to more realistic source searches (Boulanger et al. 2018).

We thank the members of the Telescope Array collaboration for fruitful discussions. We are grateful to Peter Tinyakov, Anatoli Fedynitch, and Federico Urban for fruitful discussions and revisions. This work was supported by JSPS KAKENHI grant Nos. JP19J11429 and JP19KK0074 and the joint research program of the Institute for Cosmic Ray Research (ICRR) at the University of Tokyo. E.K. is grateful for support from the “Pioneering Program of RIKEN for Evolution of Matter in the Universe (E-EMU).”

Appendix A

Analysis with a Single-mass Assumption

A.1. Mock Data Sets with a Single-mass Assumption

As single-mass assumptions, we test the pure-proton, He, C, Si, and Fe cases. We fix the energy spectrum as a broken power law with spectral indexes \(\gamma = -2.69\) \((E < 10^{18.1}\text{ EeV})\) and \(\gamma = -4.63\) \((E > 10^{18.1}\text{ EeV})\), which is as reported by the TA experiment (Tsunesada et al. 2017). In the same manner as in Section 2.3, we choose the arrival direction of the anisotropic event based on the generated CR flux models \(F_{\text{true}}\). Examples of the distribution of the mock event arrival directions with the single-mass assumption and with \((f_{\text{ani}}^{\text{true}}, \theta^{\text{true}}) = (100\%, 10^\circ)\) are shown in Figure 10. For the light masses of proton and helium, for instance, the distributions are similar to that of Figure 3, which means that the GMF bias is small. On the other hand, for the higher masses, the distortion due to the GMF is significant. In the pure-Fe assumption, the clustering of events around the top four contributing SBGs (M82, NGC4945, NGC1068, and NGC253) are not seen.

A.2. Estimated Parameters in Previous Studies with Single-mass Assumption

The results of the likelihood analysis for single-mass models are shown in Figures 11–15. In the pure-proton (Figure 11) and pure-He (Figure 12) cases, the most frequent best-fit values \((f_{\text{ani}}^{\text{true}}, \theta)\) fall near the true values \((f_{\text{ani}}^{\text{true}}, \theta^{\text{true}})\), which means that the GMF bias is small. However, when the separation angular scale \(\theta^{\text{true}}\) is larger, the dispersion of the estimated parameters, especially for \(f_{\text{ani}}\), becomes larger. This dispersion gives an intrinsic statistical uncertainty in the \(f_{\text{ani}}\) estimation. In the pure-C case (Figure 13), a discrepancy in the distributions between the north-sky and south-sky data sets can be seen. This tendency becomes larger for the heavier single-mass cases. In the pure-Si and pure-Fe cases, the most frequent value of \(f_{\text{ani}}\) in the south-sky data sets becomes 0\% with any values of the true parameters (Figures 14 and 15). Although both are separated from the true parameters, the distributions of north-sky data sets are closer to those of the all-sky data sets in any case. The single and dominant source contribution of M82 and smaller deflection by the GMF in the northern sky can explain this tendency.
Figure 10. Examples of the distribution of the mock event arrival directions with the single-mass assumptions $(f_{\text{true}}^{\text{true}}, \theta_{\text{true}}) = (100\%, 10')$ and 4000 events over the all-sky region.)
Figure 11. Same as Figure 6, but for the pure-proton case.
Mass model: pure-He assumption, GMF model: JF12, Source model: SBG

Figure 12. Same as Figure 11, but for the pure-He case.
Figure 13. Same as Figure 11, but for the pure-C case.
Mass model: pure-Si assumption, GMF model: JF12, Source model: SBG

Figure 14. Same as Figure 11, but for the pure-Si case.
Figure 15. Same as Figure 11, but for the pure-Fe case.
Appendix B
Analysis with the PT11 Model

For an independent comparison with the JF12 model, we also refer to the Pshirkov & Tinyakov 2011 model (PT11) in this study (Pshirkov et al. 2011). The mock event data sets are generated in the same manner as in Section 2.3, except for the GMF model. Figure 16 shows the distributions of the best-fit parameters in the same manner as Figure 6. Although there is a quantitative difference, both results show the same tendency due to the GMF bias.

Figure 16. Same as Figure 6, but for the mock event data sets generated with the PT11 model and the mixed-mass assumption.
ORCID iDs
Ryo Higuchi @ https://orcid.org/0000-0008-1058-216X
Takashi Sako @ https://orcid.org/0000-0002-2589-8561
Toshihiro Fujii @ https://orcid.org/0000-0003-2401-504X
Kazumasa Kawata @ https://orcid.org/0000-0001-6332-2005
Eiji Kido @ https://orcid.org/0000-0001-7278-3049

References
Aab, A., Abreu, P., Aglietta, M., et al. 2015a, NIMPA, 798, 172
Aab, A., Abreu, P., Aglietta, M., et al. 2015b, ApJ, 804, 15
Aab, A., Abreu, P., Aglietta, M., et al. 2017, JCAP, 2017, 038
Aab, A., Abreu, P., Aglietta, M., et al. 2018, ApJL, 853, L29
Abbasi, R., Abe, M., Abu-Zayyad, T., et al. 2021, NIMPA, 1019, 165726
Abbasi, R. U., Abe, M., Abu-Zayyad, T., et al. 2014, ApJL, 790, L21
Abbasi, R. U., Abe, M., Abu-Zayyad, T., et al. 2018, ApJL, 867, L27
Abreu, P., Aglietta, M., & Aguirre, C. 2007, Sci, 318, 938
Abreu, P., Aglietta, M., Ahlers, M., et al. 2012, ApJS, 203, 34
Abreu, P., Aglietta, M., Ahn, E. J., et al. 2010, APh, 34, 314
Anchordoqui, L. A. 2019, PhR, 801, 1
Batista, R. A., Bécaut, J., Bustamante, M., et al. 2019, FrASS, 6, 1
Batista, R. A., Dundovic, A., Erdmann, M., et al. 2016, JCAP, 2016, 038

Boulanger, F., Enßlin, T., Fletcher, A., et al. 2018, JCAP, 2018, 049
Bradt, H., & Olbert, S. 2008, Astrophysics Processes (Cambridge: Cambridge Univ. Press)
Bray, J. D., & Scaife, A. M. M. 2018, ApJ, 861, 3
Castellina, A. & Pierre Auger Collaboration 2019, in European Physical Journal Web of Conf. 210, Ultra High Energy Cosmic Rays 2018 (UHECR 2018), ed. I. Lhenry-Yvon et al., 06002
Fisher, R. 1953, RSPSA, 217, 295
Greisen, K. 1966, PhRvL, 16, 748
He, H. N., Kusenko, A., Nagataki, S., et al. 2016, PhRvD, 93, 043011
Heinze, J., & Fedynitch, A. 2019, ApJ, 873, 88
Hörandel, J. R. & the GCOS Collaboration 2021, ICRC (Berlin), 395, 027
Jansson, R., & Farrar, G. R. 2012a, ApJ, 757, 14
Jansson, R., & Farrar, G. R. 2012b, ApJL, 761, L11
Kawai, H., Yoshida, S., Yoshii, H., et al. 2008, NuPhS, 175, 221
Pshirkov, M. S., Tinyakov, P. G., Kronberg, P. P., & Newton-McGee, K. J. 2011, ApJ, 738, 192
Pshirkov, M. S., Tinyakov, P. G., & Urban, F. R. 2013, MNRAS, 436, 2326
Sagawa, H. 2020, JInst, 15, C09012
Sommers, P. 2001, APh, 14, 271
Tsunesada, Y., AbuZayyad, T., Ivanov, D., et al. 2017, ICRC (Busan), 301, 535
Zatsepin, G. T., & Kuz’min, V. A. 1966, JETP, 4, 78