Origins of extragalactic cosmic ray nuclei by contracting alignment patterns induced in the galactic magnetic field

M. Erdmann*, L. Geiger, D. Schmidt, M. Urban, M. Wirtz

Ill. Physikalisches Institut A, RWTH Aachen University, Otto-Blumenthal-Str., Aachen 52056, Germany

A R T I C L E   I N F O

Article history:
Received 23 July 2018
Revised 19 October 2018
Accepted 28 November 2018
Available online 29 November 2018

Keywords:
ultra-high energy cosmic rays
Composition
Sources
Magnetic fields

A B S T R A C T

We present a novel approach to search for origins of ultra-high energy cosmic rays. These particles are likely nuclei that initiate extensive air showers in the Earth's atmosphere. In large-area observatories, the particle arrival directions are measured together with their energies and the atmospheric depth at which their showers maximize. The depths provide rough measures of the nuclear charges. In a simultaneous fit to all observed cosmic rays we use the galactic magnetic field as a mass spectrometer and adapt the nuclear charges such that their extragalactic arrival directions are concentrated in as few directions as possible. Using different simulated examples we show that, with the measurements on Earth, reconstruction of extragalactic source directions is possible. In particular, we show in an astrophysical scenario that source directions can be reconstructed even within a substantial isotropic background.

© 2018 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license.
(http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

In the past decade, research on ultra-high energy cosmic rays has advanced greatly. The key to progress are large-area observatories for extensive air showers equipped with modern detection techniques [1,2]. First, the arrival distribution of cosmic rays exhibits a significant departure (> 5σ) from an isotropic distribution by a large-scale dipole structure [3]. Second, the cosmic ray energy spectrum features a high-energy cut-off [4,5]. Third, the abundance of cosmic rays with larger nuclear masses increases with energy [6,7]. The origin of cosmic rays, however, remains a burning research question. Although interesting candidate sources exist [8,9], clear evidence for point sources are still to be found.

Extrapolating from lower-energy cosmic ray measurements, ultra-high energy cosmic rays are likely protons or charged nuclei. Thus, deflections of cosmic rays in cosmic magnetic fields complicate searches for point sources by displacing charged particles from their original directions [10–16]. For coherent magnetic fields, this seeming disadvantage is balanced by energy-ordered patterns in the arrival directions of e.g. protons which, in principle, allow cosmic rays from a single direction to be distinguished from an isotropic cosmic ray background. These alignment patterns are similar to the patterns typically found in spectrometers analyzing particle momenta (e.g. [17]). Fig. 1a shows a simplified example for protons at different energies. Note that prior knowledge of the magnetic field is not required here as the arrival directions and the energy arrangement together provide information on the direction and magnitude of the field.

If the sources of ultra-high energy cosmic nuclei are within a few ten megaparsecs, the dominant deflections of cosmic rays are most likely to appear in the galactic magnetic field rather than in extragalactic fields [18–22]. From numerous Faraday rotation and synchrotron emission measurements, parameterizations of the galactic field have been constructed featuring a highly inhomogeneous field with coherent and turbulent components [23–27]. Thus, using the galactic field as a magnetic spectrometer is much more challenging compared to a laboratory experiment. Nevertheless, the field parameterizations enable detailed predictions for the deflection of cosmic particles as a function of their arrival directions, energy and charge [28–32]. Analyses of the precision of the field parameterizations indicate that - for sufficiently large cosmic ray rigidities at least, i.e., energy divided by charge - parameterizations agree at a level sufficient for cosmic ray data analyses [33,34].

To add to the challenge, a point source scenario of a cosmic ray accelerator with a mixed composition will change the energy arrangement in the arrival directions. Fig. 1b shows a simplified example of cosmic rays with the same energies as before. However, instead of protons (Fig. 1a), a mixed composition of a proton, a helium and a carbon nucleus is used, assigning the highest energy to the carbon (Z = 6). In this case, the highest energy

* Corresponding author.
E-mail address: erdmann@physik.rwth-aachen.de (M. Erdmann).

https://doi.org/10.1016/j.astropartphys.2018.11.004
0927-6505/© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)
A cosmic ray appears far away from the source as its deflection is largest owing to the large charge. A search method for origins of cosmic nuclei including corrections for such galactic magnetic field deflections has been presented in Erdmann et al. [35].

In order to exploit such mixed-composition alignment patterns induced by the galactic field, the cosmic ray charge needs to be estimated as well. Here, measurements of the air shower depth in the atmosphere provide at least a rough estimate of the cross section and thus of the nuclear mass, or the nuclear charge, respectively [7].

In this paper, we present a novel method to decompose the distribution of cosmic ray arrival directions in terms of mixed-composition alignment patterns that arise from the galactic magnetic field. Using a global fit, we minimize the number of common cosmic ray directions outside the galactic magnetic field while simultaneously adjusting the cosmic ray charges. Optimization of fit parameters is performed through backpropagation technique known from training of neural networks.

This paper is structured as follows: First, we describe the basic analysis strategy for mixed-composition alignment patterns. Then we explain the fitting technique in detail. We present the main features of the method, initially using a 1-dimensional scenario, before extending to the two dimensions on the surface of a sphere. Using a simulated astrophysical scenario, we show that point sources can be identified by correcting for the galactic magnetic field even in the presence of isotropic background. Finally, we present our conclusions.

2. Basic strategy

For each cosmic ray, the projection of the arrival direction outside the galaxy to its observed arrival direction on Earth is determined by the galactic field. Here, we rely on a field parameterization to reflect the true galactic field to a sufficient extent.

Our basic idea is to fit a data set of observed cosmic rays to originate from similar directions outside the galaxy by varying the charges within their experimental uncertainties. The objective function of the fit minimizes the distances between cosmic ray arrival directions outside our galaxy and simultaneously constrains the measurements on Earth, namely the cosmic ray energy, arrival direction and composition. To obtain the significance of a measurement, the clustering strength obtained in the data is compared to scenarios of isotropic arrival directions. Alternatively, we use the final value of the objective function of the fit.

Initially, we explain the idea using a simplified 1-dimensional scenario. In analogy to momentum measurements \( p = -Z e B r \) of particles with a charge \( Z e \) by the radius of curvature \( r \) in a magnetic field \( B \), we use a displacement from the original direction \( s \) by energy- and charge-dependent translations to arrive at direction \( p \):

\[
p = s + \frac{Z e}{E}
\]

Here, the cosmic ray energy is denoted by \( E \) and the charge by \( Z \). The ratio \( R = E/Z \) is usually referred to as rigidity. As a numerical example we allow for cosmic rays with energies between \( E = 1 \ldots 10 \) in arbitrary units, and charges between \( Z = 0 \ldots 1 \) (also in arbitrary units), such that the displacement is limited to between \( 0 \leq p - s \leq 1 \).

In Fig. 2a, we show the mixed-composition alignment pattern of Fig. 1b together with the source position of the cosmic rays. In the two histograms below, we show different levels of potentially available information to reconstruct the true source position. If no information on the true cosmic ray charges is available, the range of possible source positions is limited by the maximum charge (Fig. 2b). For a low-energy cosmic ray of \( E = 1 \), the displacement may be as great as \( p - s = 1 \) if \( Z_{\text{max}} = 1 \). While for cosmic rays with \( E = 3 \) the maximal possible displacement is \( p - s = 1/3 \). In the remaining interval the reconstructed source position follows a uniform probability distribution. Such estimates can be improved by using probability distributions obtained from measurements of the average shower depths instead of the intervals.

In Fig. 2c, we show the best possible case of having charge estimates for the individual cosmic rays available from measurements of the shower depth. Here, each cosmic ray provides an estimate of the reconstructed source position which can be improved by a weighted average of all estimated source positions.

Below, we will also study cosmic ray deflections on the 2-dimensional surface of a sphere. To this end, we will first analyze the effects of a horizontal magnetic field before investigating deflections in the galactic magnetic field.

3. Techniques of the fit

For each cosmic ray \( i \) we observe the arrival direction \( \hat{p}_i \) and its energy \( E_i \) on Earth. We also have one of the above-mentioned charge estimates \( Z_i \) (compare Fig. 2).

A transformation prescription \( T \) is used to predict for each given pair \((\hat{s}_i, Z_i)\) of the estimated original cosmic ray direction \( \hat{s}_i \) and the estimated cosmic ray charge \( Z_i \) the cosmic ray directions \( \hat{p}_i \) on Earth:

\[
\left( \begin{array}{c} \hat{s}_i \\ Z_i \end{array} \right) \Rightarrow T(\hat{s}_i, Z_i, E_i) \Rightarrow \hat{p}_i
\]

We optimize the directional parameters \( \hat{s}_i \) and charge parameters \( Z_i \) according to the objective functions explained below. In doing so, we cluster the arrival directions \( \hat{s}_i \) outside the galaxy while requiring consistency between predicted \( \hat{p}_i \) and measured
arrival directions \( p_i \) as well as between estimated charges \( \hat{Z}_i \) and measured shower depths \( X_{\text{max}} \). For the fitting technique we use backpropagation as implemented in TensorFlow [37]. For the optimization we use the concept of gradient descent.

3.1. Transformations and parameters

In the course of our study, we will increase the complexity of the scenario under investigation, and adapt the transformations and corresponding parameters accordingly.

One-dimensional transformations. For our 1-dimensional studies, we use as the transformation \( T \) the translation \( T = \hat{Z} + Z/E \) presented in Eq. (1). In the fit, the directional parameters \( \hat{Z}_i \) and \( \hat{p}_i \) are scalar values, and the charge parameters \( \hat{Z}_i \) are real numbers.

Two-dimensional transformations on a sphere. On the surface of the sphere, the directional parameters \( \hat{Z}_i \) and \( \hat{p}_i \) are 3-dimensional unit vectors with \((x, y, z)\)-components, and the fitted charge parameters \( \hat{Z}_i \) are real numbers.

As our transformation \( T \) we use the rotation matrix \( M(\delta) \) around the \( z \)-axis targeting perpendicular to the galactic plane, thereby varying only the longitude \( l \) coordinate of the cosmic ray, while keeping its latitude \( b \) constant. The rotation angles \( \delta \) depend on the cosmic ray energies in units of EeV and their charges

\[
\delta_i(\hat{Z}_i, E_i) = -2 \frac{\hat{Z}_i}{(E_i/E_{\text{eV}})} .
\]

Here, the factor 2 serves to displace heavy nuclei by several 10 degrees. The transformations \( T = M(\delta) \) in Eq. (2) are then calculated according to

\[
\hat{p}_i = M(\delta_i(\hat{Z}_i, E_i)) \cdot \hat{Z}_i .
\]

Deflections in the galactic magnetic field. Finally, we use a representation \( L \) of the galactic magnetic field to predict the arrival directions on Earth in Eq. (2) with \( T = L \) (see Section 5):

\[
\hat{p}_i = L(\hat{Z}_i, E_i) .
\]

This study is also performed on the surface of a sphere. The cosmic ray energies are in units of EeV, and their fitted charges are real numbers.

3.2. Objective function

The objective function to guide the fit (2) consists of several terms which aim at concentrating cosmic ray arrival directions prior to transformation while preserving the observed quantities, namely the arrival directions on Earth and the shower maximum.

Spatial distances between predicted and observed directions. The quality of the initial parameters, direction \( \hat{Z}_i \) and charge \( \hat{Z}_i \), in Eq. (2), can be judged by means of a comparison to the observed direction \( p_i \). To this end, we define an objective term evaluating the quality of the initial parameters for all \( N \) cosmic rays by:

\[
D = \frac{1}{N} \sum_i \|p_i - \hat{p}_i\|^2 .
\]

Differences between predicted and observed charges. To relate individual cosmic ray charge estimates \( \hat{Z}_i \) to measurements of the shower depth of maximum \( X_{\text{max}} \) we introduce an objective term which takes into account the asymmetry of the \( X_{\text{max}} \)-distribution that results from shower-to-shower fluctuations. This asymmetric distribution can be described by Gumbel functions \( G(A_i, E_i) \) [36]. For charge numbers \( \hat{Z}_i \geq 2 \) we assume \( A_i \approx 2\hat{Z}_i \). To cover the case of the proton \( (A_i = 1, \hat{Z}_i = 1) \), we use \( \hat{Z}_i = Z_i^2 \) for charge numbers \( \hat{Z}_i < 2 \). To enable smooth gradients, we combine both functions by multiplication with the sigmoid function. We construct the additional objective term \( Q \) by demanding that the mean squared error of the \( X_{\text{max}} \) values follow a \( \chi^2 \)-distribution with \( N \) degrees of freedom:

\[
Q = \left( \frac{1}{N} \sum_i \frac{(X_{\text{max},i} - \mu_i)^2}{\text{Var}(G(A_i, E_i))} - 1 \right)^2 .
\]

Here, the sum is taken across all \( N \) cosmic rays, \( \mu_i = \arg \max X_{\text{max}} G(A_i, E_i) \) denotes the most probable \( X_{\text{max}} \) value and \( \text{Var}(G(A_i, E_i)) \) the Gaussian approximation of the variance for either the left or the right tail of the Gumbel distribution.

Clustering: spatial distances to common original directions. For the 1-dimensional case, in order to demand the cosmic rays to originate from similar directions \( \hat{Z}_i \) we use the concept of \( k \) nearest neighbors. The direction \( \hat{z}_i \) of each cosmic ray should be close to the average direction \( \langle \hat{Z}_i \rangle \) of its \( k \) nearest neighbors \( (\text{knn}) \), which is calculated with:

\[
\langle \hat{Z}_i \rangle = \frac{1}{k} \sum_{\text{knn}(\hat{Z}_i)} \hat{Z}_j .
\]

The objective term to demand clustering of original cosmic ray directions reads:

\[
C = \frac{1}{N} \sum_i \|\hat{Z}_i - \langle \hat{Z}_i \rangle\|^2 .
\]

On the sphere we follow a slightly different approach. To force clustering, we apply a mean squared distance objective between all original cosmic ray directions of the form \( \|\hat{Z}_i - \hat{Z}_j\|^2 \).

We only allow clustering of neighboring cosmic rays by additionally multiplying with an elliptical shaped weight factor \( \epsilon_{ij} \) around each cosmic ray \( i \), where the major axis is aligned with the elongated structures caused by the magnetic field. This particular shape reduces contributions from cosmic rays other than those originating from the same source. The elliptical weight factor reads \( \epsilon_{ij} = \cos^2(\gamma_0) \) where \( \gamma_0 \) denotes the angular distance between cosmic ray source estimations \( \hat{z}_i \) and \( \hat{z}_j \). We use \( \gamma = 4.3 \) along the major axis to achieve a weight drop to 0.1 at \( \gamma = 40 \) deg, and \( \gamma = 470 \) along the minor axis for a weight drop to 0.1 at \( \gamma = 4 \) deg. The complete cluster objective thus reads:

\[
C = \frac{1}{N} \sum_i \sum_j \epsilon_{ij} \cdot \|\hat{Z}_i - \hat{Z}_j\|^2 \sum_i \sum_j \epsilon_{ij} \]

(10)

Total objective function of the fit. The fit is driven by an objective function consisting of the objective terms (6), (7), and depending on the study (9) or (10). Their relative weights are adjusted by the hyperparameters \( \lambda_Q \) and \( \lambda_C \):

\[
J = D + \lambda_Q Q + \lambda_C C .
\]

(11)

We investigated the influence of the hyperparameters on the performance of the fit using different scenarios. In general, larger values of \( \lambda_Q \) led to a better reconstruction of the charge. However, the measurement accuracy of \( X_{\text{max}} \) yielding the charge reconstruction is rather limited compared to the accuracy of the measured cosmic ray directions.

Therefore, we chose the term \( D \) of the distance loss as the dominant term and downscale the influence of the term \( Q \) of the charge reconstruction. We also observed that the hyperparameter \( \lambda_C \), which controls the influence of the clustering strength \( C \) at source positions, should be subordinate in comparison. As typical values, we found \( \lambda_Q = 0.1 \) and \( \lambda_C = 0.01 \) to provide the best performance in both charge and source reconstruction.
4. Benchmark distributions of fits to mixed-composition patterns

Initially, we benchmark the fit procedure using 1-dimensional transformations as presented in Eq. (1). We then extend to the 2-dimensional sphere of the cosmic ray sky using similar linear transformations of the longitude while keeping the latitude constant (see Eq. (4)). We use simulated mixed-composition data as will be explained below.

4.1. One-dimensional transformations

As our data set for the 1-dimensional case, we use pseudo cosmic rays with uniformly distributed energies between $E = 1 \ldots 10$ (arbitrary units) and charges with $Z = c/26$ for $c = 1, \ldots, 26$. Estimating the nuclear mass by $A = 2c$, we assign a value of the shower depth $X_{\text{max}}$ using the above-mentioned Gumbel functions $G(A, E)$.

In Fig. 3a we show the direction of a single source using the star symbol, and its $N = 10$ cosmic rays observed on Earth at directions $p_i$ using bars. The displacements were calculated using the transformation $T = Z_i/E_i$ in Eq. (1). Cosmic ray energies $E_i$ are represented by the magnitudes of the bar symbols, and the true charges $Z_i$ are coded with the color scale.

For the fit, we assign $Z_i = 0.5$ to the initial charge values in Eq. (2) and calculate the initial original directions $\hat{s}_i$ by applying the inverse transformations $\hat{s}_i = p_i - Z_i/E_i$ to the measured directions $p_i$.

In Fig. 4a, we show for each cosmic ray the predicted source direction $\hat{s}$, as a function of the number of fit iterations. After 200 iterations, the fit converged to the true source direction $s$, which is represented by the star on the right.

In each iteration of the fit, the average predicted direction $\langle \hat{s} \rangle_i$ of the $k$ nearest neighbors of cosmic ray $i$ is first calculated using Eq. (8). In the example shown we use $k = N = 10$. With the cluster term (9) the fit is guided to modify each predicted direction $\hat{s}_i$ to move towards the average predicted direction $\langle \hat{s} \rangle_i$.

To modify the predicted direction $\hat{s}_i$ and still correctly predict the observed cosmic ray direction $p_i$ on Earth, as required in the distance term (6), the fit needs to adjust the cosmic ray charge $Z_i$ accordingly. The charges are thus varied within the constraints of the measured shower depths $X_{\text{max}}$ using the charge term (7). In this way, consistency of the predicted source direction and adjusted cosmic ray charges with the measured arrival directions, energies, and shower depths is achieved.

To enable a comparison with the single-source scenario presented thus far, we also discuss scenarios of isotropic cosmic ray arrival. Isotropic arrival of $N = 10$ cosmic rays can be interpreted as a scenario with $m = N$ sources each emitting a single cosmic ray. In Fig. 3b we show as an example the directions of the 10 sources using stars, and their cosmic rays observed on Earth at directions $p_i$ using bars.

In Fig. 4b, we show for each cosmic ray the predicted source direction $\hat{s}_i$ depending on the fit iterations. The true source directions are shown using stars on the right, where the color code relates each cosmic ray with its respective source.

Again, for each cosmic ray $i$ the fit averages the predicted direction $\langle \hat{s}_i \rangle$ of the $k = 10$ nearest neighbors (8), and attempts to concentrate the predicted directions according to the cluster term (9). However, to modify the predicted source direction $\hat{s}_i$, the predicted cosmic ray direction $p_i$ on Earth needs to be preserved in the distance term (6). This requires an adjustment of the predicted cosmic ray charge $Z_i$ according to the charge term (7).
In the balance of the different objective terms (11), the fit is driven to split the predicted directions into several clusters with low cosmic ray occupancy as can be seen in Fig. 4b.

We also evaluate the precision in the reconstruction of source directions and cosmic ray charges, again starting with the single-source scenario. In Fig. 5a, we show the reconstruction quality of the source direction. The red solid histogram represents for each cosmic ray the difference of the reconstructed and true source directions $\Delta s = \hat{s}_i - \hat{s}_t$. To provide a general measure, we average over 100 scenarios with 1 source emitting $N = 10$ cosmic rays. The resolution in the source direction is rather good ($\sigma = 0.024$).

In Fig. 5b, the red solid histogram also shows the reconstructed cosmic ray charges compared to the true charges $\Delta Z = \hat{Z}_i - \hat{Z}_t$. Once again, we average over 100 scenarios. The resolution in the cosmic ray charge amounts to $\sigma_Z = 0.15$ for $0 \leq Z \leq 1$ used in the fit. For charges within $c = 1 \ldots 26$ this implies a resolution of $26 \times \sigma_Z = 4$ which is superior when compared to the information contained in the $X_{\text{max}}$ measurement which has a logarithmic mass or charge dependence ($Z = A/2$), respectively. Note, however, that the true transformation $T$ in Eq. (1) must be known for this charge reconstruction.

For the isotropic scenario, neither the reconstruction of the source directions nor of the charges work at a satisfactory level. In Fig. 5a, the blue dashed histogram shows the resolution in the source direction averaged over 100 isotropic scenarios. In Fig. 5b, we also show the corresponding charge resolution.

In Fig. 6a, the red histogram shows the values of the objective function (11) after the fit for 100 single-source scenarios, and the blue histogram shows the objective values of 100 isotropic scenarios. For better visibility, we normalized the final objective values to the largest objective value observed in all 200 scenarios. In this comparison the single-source scenario can be well distinguished from isotropic scenarios.

We also investigate the dependence on the number $N$ of cosmic rays. In Fig. 6b, the red solid histogram shows the normalized values of the objective function (11) after the fit to $N = 100$ cosmic rays from a single source for 100 scenarios, and the blue dashed histogram shows the objective values of 100 isotropic scenarios with 100 cosmic rays. With increased statistics, the signal scenarios can be even better distinguished from isotropic arrival.

Finally, we study a mixed scenario with a source with $N = 50$ cosmic rays which is overlayed with 50 cosmic rays from an isotropic scenario. As $k$ we used the value $k = N = 100$, so the fit tries to find one single source. In Fig. 6c, the red solid histogram shows the values of the normalized objective function (11) for the mixed scenario, and the blue dashed histogram shows the objective values of 100 isotropic scenarios. In spite of the large background, half of the signal scenarios can be distinguished from isotropic arrival.

In Fig. 7 we show the influence of the charge information resulting from measurements of the cosmic ray shower depth $X_{\text{max}}$ for 100 mixed scenarios with 50% signal and 50% isotropic
4.2. Spherical mixed-composition patterns

In this step, we extend our benchmark studies to the two-dimensional arrival directions of cosmic rays on the surface of a sphere. Before we investigate a realistic astrophysical scenario in chapter 5, we examine data sets of pseudo-cosmic rays to verify the capability of the fit concept on the sphere.

As our transformation $T$ in Eq. (2) in the fit, we use the matrix $M(x)$ in Eq. (4) for rotations around the $z$-axis which is perpendicular to the galactic plane. We vary only the longitude $I$ coordinate of the cosmic ray while keeping its latitude $b$ constant. The rotation angle is energy- and charge-dependent according to (3).

In our simulations, we use uniformly distributed cosmic ray energies between $E_{\text{min}} = 40$ EeV and $E_{\text{max}} = 100$ EeV. For the charges we assign a uniform distribution between protons ($Z = 1$) and iron ($Z = 26$). This leads to rotation angles between $\delta = 0.02 \ldots 1.3$ rad. Again, we use the Gumbel functions $G(A, E)$ to assign values for the shower depth $X_{\text{max}}$.

Fig. 8a shows an example arrival map for $m = 10$ sources each emitting $N = 10$ cosmic rays. The symbol sizes indicate the cosmic ray energies, and the charges are denoted by the color code.

As the initial parameters for the fit in Eq. (2), for the predicted cosmic ray arrival directions $\vec{s}_i$ outside the galaxy we use the observed arrival directions on Earth $\vec{s}_i = p_i$, and for the charges $Z_i$ we apply Bayes’ theorem to the shower depth $X_{\text{max}, i}$ assuming a flat composition prior between $Z_i = 1 \ldots 5$.

In Fig. 9a we show the result of the 2-dimensional fit for this source scenario. The stars again denote the source directions, and the circular symbols present the predicted arrival directions prior to the coherent deflections. After $\sim 10^6$ fit iterations on the sphere, the vast majority of the $N = 100$ cosmic rays has been assigned correctly to their source directions and are almost not visible owing to the large clustering strengths at the sources.

A small fraction of the cosmic rays remains isolated or was allocated to the wrong sources; this is attributable to a combination of the initial charge assignment and the algorithm for evaluating original directions of neighboring cosmic rays. Once the allocation to a group of nearest cosmic rays is incorrect, it appears to be difficult for the fit to reassign the cosmic ray to another group of neighbors.

In Fig. 10a we show the reconstructed charges of the cosmic rays as a function of their true charges. A linear correlation of the reconstructed with the true charges is visible for the majority of the cosmic rays. The two outliers can be identified in Fig. 9a, one in the upper left in an equilibrium between two source positions, and one at longitude 80 deg and latitude 5 deg misidentified with a different source due to the high assigned charge.

As a measure of clustering strength we use the so-called top-hat counting [38]. For each cosmic ray we count the number of cosmic rays within a radial distance of 5 deg and include the initiating cosmic ray in the count. In Fig. 11 the red solid histogram shows the count for each of the $N = 100$ cosmic rays. The majority of the cosmic rays yield a cluster strength of 10, as expected, when they are correctly allocated to their sources. The above-mentioned cosmic ray which was incorrectly allocated to another source is visible at $N_i = 11$, and also the cosmic ray that remained in an equilibrium between two sources is visible at $N_i = 1$.

For comparison, we also present the results of a fit to the isotropic scenario shown in Fig. 8b. In Fig. 9b we show the predicted original directions of the cosmic rays prior to their deflections. The fit attempts to concentrate the predicted directions as required in the objective function (11), however, with the result of many small clusters spread around the sphere. These small clusters remain vertically aligned (same longitude band) as the fit is only capable to perform horizontal rotations.

In Fig. 11 the blue dashed histogram shows the clustering strength by the top-hat counting explained above. Compared to the $m = 10$ source scenario, only small clusters are formed. Therefore, the source scenario can be easily separated from isotropic scenarios, and top-hat counting can serve to evaluate the significance of a signal scenario.
In Fig. 10b, we also show the reconstructed charges of the cosmic rays for the isotropic scenario. Here, a weak correlation between the true charges $Z$ and the reconstructed charges $\hat{Z}$ is found which results from the $X_{\text{max}}$ constraint of the charge objective in Eq. (7).

5. Cosmic ray sources by correcting for galactic magnetic field deflections

Finally, we apply the fit to correct for cosmic ray deflections in the galactic magnetic field. As our simulated astrophysical scenario, we use $m = 4$ sources each emitting $N = 25$ cosmic rays and an additional 900 cosmic rays following an isotropic distribution. The energy spectrum follows [4] above a lower-energy cut-off of $E > 40$ EeV. With a total of 1000 cosmic rays above this energy threshold, we yield cosmic ray statistics compatible with actual experiments.

We use an energy-independent mixed composition between proton and oxygen ($Z = 1, \ldots, 8$) that follows an uniform charge distribution. Also here, we use the Gumbel functions $G(A, E)$ to assign shower depths $X_{\text{max}}$ to each cosmic ray. The upper charge limit avoids low cosmic ray rigidities $R = E/Z$ which, depending on the arrival direction, may lead to large non-ballistic deflections [33].

Cosmic ray deflections in the galactic magnetic field are performed using the regular JF12 field [24] encoded in rigidity-dependent magnetic lenses [39,40]. The lenses are based on HEALPix coordinates using 1 deg$^2$ resolution [41]. In order to include a turbulent magnetic field component, we add a Gaussian smearing of the arrival directions with a standard deviation of

$$\sigma = 0.5 \frac{Z}{E/\text{EeV}} \text{rad.}$$

(12)

In Fig. 11 we show the arrival directions of the cosmic rays on Earth where the circular symbols denote the cosmic rays, the sizes represent the cosmic ray energies, and the color codes denote the charges. The gray shaded areas indicate cosmic rays originating from the sources which are marked by star symbols.

In our fit we again use the regular field of the JF12 parameterization [24] to calculate cosmic ray deflections. To address cosmic ray deflections in the fit, we replaced the lenses by a deep neural network $L$ which serves to predict magnetic field deflections.
in Eq. (5). For a given cosmic ray direction \( \hat{s}_i \) outside the galaxy, energy \( E_i \) and charge \( \hat{Z}_i \), the network outputs the most likely arrival direction on Earth, thus providing the transformation \( T = L \) in Eq. (2). We verified that the network provides a good representation of the lenses above cosmic ray rigidities \( R = 1 \) EV, and interpolates well with respect to continuous rigidity and arrival direction values.

To take into account turbulent field components or possible uncertainties in the regular field, we adapted the transverse size of the elliptical search region to 4 deg; details can be found around Eq. (10) in Section 3.2 above.

For the fit, initial values of the predicted cosmic ray arrival directions \( \hat{s}_i \) outside the galaxy and the charges \( \hat{Z}_i \) in Eq. (2) are required. We set \( \hat{s}_i \) to the observed arrival directions \( \hat{s}_i = p_i \). To estimate the cosmic ray charges we apply Bayes' theorem to the shower depth \( X_{\text{max}, i} \), assuming a flat composition prior between \( \hat{Z}_i = 1 \ldots 5 \).

Fig. 13 shows the fit results of the original arrival directions \( \hat{s}_i \) for all cosmic rays. Most of the signal cosmic rays (colored symbols) cluster in regions close to their original sources.

Isotropic background events (gray shading) tend to cluster as well, as required by the fit (10), where some directions of the sky are more populated than others. The latter effect results from the directional dependent transparency of the galactic magnetic field model. Some of the signal events are attracted by clusters of background events and therefore miss their original source directions. Conversely, the source regions contain some background cosmic rays that are pulled towards the high clustering region.

To quantify a significance for fit results of source scenarios with \( m = 4 \) sources, we investigate the distribution of top-hat
counts within 5°. To account for the above-mentioned directional dependent transparency of the galactic magnetic field model, we show the top-hat counts \( N_r \) relative to the average top-hat counts expected for isotropic distributions \( \langle N_{r,iso}^s \rangle \) at the same arrival direction, i.e., the same HEALPix coordinate at 1° resolution.

In Fig. 14a, we show for the fitted source scenario presented in Fig. 13 the top-hat ratio \( N_r/\langle N_{r,iso}^s \rangle \) for each cosmic ray using the red solid histogram. To enable a comparison with purely isotropic arrival directions, we also show the fit result of a single isotropic realization using the blue dashed histogram. For the source scenario, a population of cosmic rays is visible exceeding \( N_r/\langle N_{r,iso}^s \rangle = 5 \) (black dashed line). This population corresponds mostly to signal cosmic rays and a few background events which were attracted to the same clusters.

As a statistical measure to discriminate between isotropic and anisotropic scenarios, we use the number of cosmic rays that exceed the threshold \( N_{crit} (N_r/\langle N_{r,iso}^s \rangle > 5) \) after the fit. In Fig. 14b, we show the number of cosmic rays that exceed this threshold for 50 signal scenarios using the red solid histogram, and its median by the red vertical line. These signal scenarios again contain \( m = 4 \) random source positions and 25 signal cosmic rays per source, overlaid with 900 cosmic rays which follow an isotropic distribution.

For comparison, the blue dashed histogram represents the expectations for 500 isotropic realizations. We find that a single isotropic set exceeds the median expectation for our astrophysical scenario. Thus, as the expected sensitivity to distinguish clustering from an astrophysical scenario with 10% signal fraction from 4 sources from clustering resulting from isotropic arrival distributions we report a p-value of \( p = 2 \cdot 10^{-3} \) corresponding to 2.9 standard deviations.

For the astrophysical fit presented in Fig. 13, we also investigate the reconstruction quality of the source directions and of the cosmic ray charges. To enhance signal cosmic rays, we use only cosmic rays with top-hat counts above isotropic expectations \( N_r/\langle N_{r,iso}^s \rangle > 5 \). In Fig. 15a we show the top-hat ratio of signal (background) cosmic rays using the colored (gray) symbols, whereby the symbol size follows the top-hat ratio. Also shown are the true source directions (star symbols). Regions of strong clustering are primarily found close to the sources.

Fig. 15b shows reconstructed cosmic ray charges in comparison to their true charges. While many of the charges of signal cosmic rays correlate with their true charges, background cosmic rays often receive the wrong charges in order to join clusters.

6. Conclusions

In this paper we presented a new method for estimating the source directions of ultra-high energy cosmic rays by means of a simultaneous fit to all cosmic rays. Initially, the source directions and charges of the particles are estimated from measurements of their arrival direction on Earth, energy and shower depth. Assuming that some directions of origin are not isotropically distributed, but are instead concentrated in the direction of point sources, the estimated directions of origin and charges of the particles are iteratively improved within the fit. For this, we need information at least about the approximate local deflection directions of the galactic magnetic field, which we take from a parameterization of Faraday rotation and Starlight polarization measurements. The
exact strength of the magnetic field is less important than the
direction, because it can be compensated by a shift in the charge
measurements.

The initially analyzed case of a 1-dimensional shift shows that
the reconstruction of the sources is possible by the fit. The case
also shows the challenge when signal particles and background
particles are aligned and partially overlap. Here, half of the signal
scenarios could still be distinguished from isotropic scenarios. In
the 2-dimensional spherical case with deflections along the lon-
gitudinal coordinate, the fit allocates most particles correctly to their
sources. After the fit, directions of particle origins cluster with a
magnitude that does not occur in isotropic cases.

We tested an astrophysical scenario by simulating \( N = 1000 \)
different nuclei traversing the galactic magnetic field, where \( m = 4 \)
sources serve as sources of 10% of the cosmic rays and 90% of
them follow an isotropic distribution. Here, too, we observe strong
clustering of particle origins around the expected source directions.
The clusters consist mainly of cosmic rays which were correctly at-
located to their sources and, to a lesser extent, of particles of the
isotropic admixture which are pushed by the fit towards the clus-
ter. In purely isotropic scenarios, however, large clustering strength
as observed in source scenarios is suppressed, such that source
scenarios can be clearly separated from isotropic arrival distribu-
tions.

Overall, the fit method provides a novel approach for decoding
the arrival distribution of ultra-high energy cosmic rays in terms of
mixed-composition alignment patterns generated by deflections in
the galactic magnetic field. It enables identification of cosmic ray
directions outside our galaxy with a high occupancy and therefore
candidate directions of cosmic ray point sources.

Acknowledgments

We wish to thank Roger Clay for his valuable comments on the
manuscript. This work is supported by the Ministry of Innovation,
Science and Research of the State of North Rhine-Westphalia, and
the Federal Ministry of Education and Research (BMBF).

References

[1] T. Abu-Zayyad, et al., The telescope array collaboration, Nucl. Instrum. Methods A 689 (2013) 87, doi:10.1016/j.nima.2012.05.070.
[2] A. Aab, et al., The pierre auger collaboration, Nucl. Instrum. Methods A 798 (2015) 172, doi:10.1016/j.nima.2015.06.058.
[3] A. Aab, et al., The pierre auger collaboration, Science 357 (6357) (2017) 1266, doi:10.1126/science.aan4338.
[4] J. Abraham, P. Abreu, M. Aglietta, et al., The pierre auger collaboration, Phys. Lett. B 685 (2010) 239.
[5] T. Abu-Zayyad, R. Aida, M. Allen, et al., The telescope array collaboration, Astrop. J. 13 (2013) 768, L1.
[6] A. Aab, P. Abreu, M. Aglietta, et al., The pierre auger collaboration, Phys. Rev. D 90 (2014) 122005.
[7] A. Aab, P. Abreu, M. Aglietta, et al., The pierre auger collaboration, Phys. Rev. D 90 (12) (2014) 122008.
[8] R.I. Abbasi, M. Abe, T. Abu-Zayyad, et al., The telescope array collaboration, Astrophys. J. 790 (2014) L21.
[9] A. Aab, et al., The pierre auger collaboration, Astrophys. J. 853 (2018) L29, doi:10.3847/2041-8213/aaada6.
[10] T. Stanek, Astrophys. J. 479 (1997) 290.
[11] D. Harari, S. Mollerach, E. Roulet, JHEP 0002 (2000) 035.
[12] D. Harari, S. Mollerach, E. Roulet, F. Sanchez, JHEP 0203 (2002) 045.
[13] G. Golup, D. Harari, S. Mollerach, E. Roulet, Astropart. Phys. 32 (2009) 260.
[14] G. Giacinti, M. Kachelriess, D.V. Semikoz, G. Sigl, JCAP 1008 (2010) 036.
[15] G. Golup, D. Harari, S. Mollerach, E. Roulet, JCAP 1107 (2011) 006.
[16] G. Giacinti, M. Kachelriess, D.V. Semikoz, G. Sigl, Phys. Rev. D 85 (2012) 113502.
[17] P. Abreu, et al., [pierre auger collaboration], Astropart. Phys. 35 (2012) 354, doi:10.1016/j.astropartphys.2011.10.004.
[18] K. Dolag, D. Grasso, V. Springel, I. Tkachev, JETP Lett. 79 (2004) 583.
[19] D. Fauvet, E. Faletti, Astron. Phys. Rev. D 83 (2010) 123503.
[20] R. Durrer, A. Neronov, Astron. Astrophys. Rev. 21 (2013) 62.
[21] S. Hackstein, F. Vazza, M. Brueggen, G. Sigl, A. Dundovic, MNRAS 462 (4) (2016) 3660.
[22] J.D. Bray, A.M.M. Scaife, Astrophys. J. 861 (2018) 3.
[23] M.S. Pshirkov, P.G. Tinyakov, P.K. Kronberg, K.J. Newton-McGee, Astrophys. J. 738 (2011) 192.
[24] R. Jansson, G.R. Farrar, Astrophys. J. 757 (2012) 14.
[25] M.S. Pshirkov, P.G. Tinyakov, F.R. Urban, MNRAS 436 (2013) 2326.
[26] R. Jansson, G.R. Farrar, Astrophys. J. 761 (2012) L11.
[27] M.C. Beck, A.M. Beck, R. Beck, K. Dolag, A.W. Strong, P. Nielaba, JCAP 1605 (05) (2016) 056, doi:10.1088/1475-7516/2016/05/056.
[28] G.R. Farrar, Comptes Rendus Physique, 2014, arXiv:1405.3680, 15339.
[29] A. Keivani, G.R. Farrar, M. Sutherland, Astropart. Phys. 61 (2014) 47.
[30] M. Erdmann, G. Müller, M. Urban, 34th Int. Cosmic Ray Conf., The Hague, The Netherlands, PoS (ICRC2015) 557 (2015), arXiv:1508.03737.
[31] G.R. Farrar, N. Aalal, D. Khurana, M. Sutherland, 34th Int. Cosmic Ray Conf., The Hague, The Netherlands, PoS (ICRC2015) 560 (2015), arXiv:1508.04530.
[32] G.R. Farrar, M.S. Sutherland, arXiv:1711.02730 [astro-ph.HE].
[33] M. Erdmann, G. Müller, M. Urban, M. Wirtz, Astropart. Phys. 85 (2016) 54, doi:10.1016/j.astropartphys.2016.10.002.
[34] M. Unger, G.R. Farrar, arXiv:1707.02339 [astro-ph.GA].
[35] M. Erdmann, G. Müller, M. Urban, M. Wirtz, 35th Int. Cosmic Ray Conf., 2017, Busan, South Korea, PoS (ICRC2017), 518.
[36] M. De Domenico, M. Settimo, S. Riggi, E. Bertin, JCAP 107 (2013) 050, doi: 10.1088/1475-7516/2013/07/050.
[37] M. Abadi et al., arXiv:1601.04467 https://www.tensorflow.org.
[38] D.E. Alexandreas, D. Berley, S. Biller, G.M. Dion, J.A. Goodman, T.J. Haines, C.M. Hoffman, E. Horch, NIMPA 328 (1993) 570.
[39] H.P. Breit, M. Erdmann, P. Schiffer, D. Walz, T. Winchen, Astropart. Phys. 54 (2014) 110.
[40] A. Dundovic, M. Erdmann, R.A. Batista, et al., JCAP 05 (2016) 38.
[41] K. Gorski, E. Hivon, A. Banday, et al., Astrophys. J. 622 (759) (2005) 771.