# Anisotropies of ultra-high energy particles in cosmic magnetic fields

For the attainment of the academic degree of

**Doctorate in Science**

from the KIT-Fakultät für Physik of the
Karlsruhe Institut of Technologie (KIT)
and the Instituto de Tecnología "Prof. Jorge A. Sábato" of the
Universidad Nacional de San Martín (UNSAM)

accepted

**Dissertation**

of

**M. Sc. Luca Deval**

from Aosta

Day of the oral examination: 07. 02. 2025
Referee: Prof. Dr. Ralph Engel
Co-referee: Dr. Esteban Roulet
Supervisors: Dr. Michael Unger

# Anisotropies of ultra-high energy particles in cosmic magnetic fields

Zur Erlangung des akademischen Grades eines

**Doktors der Naturwissenschaften (Dr. rer. nat.)**

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
und des Instituto de Tecnología "Prof. Jorge A. Sábato" der
Universidad Nacional de San Martín (UNSAM)

genehmigte

**Dissertation**

von

**M. Sc. Luca Deval**

aus Aosta

Tag der mündlichen Prüfung: 07. 02. 2025
Referent: Prof. Dr. Ralph Engel
Korreferent: Dr. Esteban Roulet
Betreuer: Dr. Michael Unger

# Anisotropies of ultra-high energy particles in cosmic magnetic fields

Tesis presentada para optar por el título de

**Doctor en Astrofísica**

de la KIT-Fakultät für Physik del
Karlsruhe Institut of Technologie (KIT)
y del Instituto de Tecnología "Prof. Jorge A. Sábato" de la
Universidad Nacional de San Martín (UNSAM)

aceptada

**Disertación**

de

**M. Sc. Luca Deval**

nacido en Aosta

Fecha de la defensa oral: 07. 02. 2025
Director: Prof. Dr. Ralph Engel
Co-director: Dr. Esteban Roulet
Colaboradores: Dr. Michael Unger

# Contents

# Abstract

Ultra-high-energy cosmic rays (UHECRs) are the most energetic particles observed in the Universe, and their arrival directions provide essential information about their origin and propagation. Anisotropies in the UHECR sky, observed as deviations from isotropy, are expected to be influenced by deflections caused by both the Galactic Magnetic Field (GMF) and Extragalactic Magnetic Field (EGMF).

In this work, simulations are performed to study the impact of magnetic deflections on the arrival directions of UHECRs, incorporating both the coherent and turbulent components of the GMF, as well as the influence of the EGMF on observed anisotropic structures. A likelihood-based analysis is used to quantify the cross-correlation between the GMF-deflected arrival directions of UHECRs and a source model, given by a catalog of starburst galaxies, providing insights on the deviation from isotropic expectation when angular deflections due to the GMF are included. Additionally, the intermediate-scale anisotropies in the UHECR sky are investigated to place constraints on the strength of the EGMF. Systematic scans over the signal fraction and smearing angle related to the EGMF demonstrate that scenarios with moderate random deflections by the magnetic fields and low signal fractions successfully reproduce the observed data. Conversely, parameter configurations with high signal fractions and small EGMF deflections result in significant overdensities that deviate from observations. Comparisons between different astrophysical models, including variations of the GMF, further highlight the crucial role of magnetic fields in shaping the UHECR sky and thus determine the properties of these fields.

This study offers a statistical framework for exploring the cross-correlation between UHECR anisotropies and magnetic fields, proposing constraints on the strength of such fields by identifying smearing angles for which clustering of events is diluted to match the observed anisotropic features in the sky.


The material presented in this thesis is the result of my original work. To improve readability and flow, tools such as *DeepL* and *ChatGPT* have been utilized.

## Zusammenfassung

Ultrahochenergetische kosmische Strahlen (UHECRs) sind die energiereichsten Teilchen, die im Universum beobachtet werden. Ihre Ankunftsrichtungen liefern wichtige Informationen über ihren Ursprung und ihre Ausbreitung. Anisotropien am UHECR-Himmel, die als Abweichungen von der Isotropie auftreten, werden von Ablenkungen der UHECR Teilchen im galaktischen (GMF) und extragalaktischen (EGMF) Magnetfeld beeinflusst.

In dieser Arbeit werden Simulationen durchgeführt, um den Einfluss magnetischer Ablenkungen auf die Ankunftsrichtungen von UHECRs zu untersuchen. Dabei werden sowohl die kohärente als auch die turbulente Komponente des GMF sowie der Einfluss des EGMF auf beobachtete anisotrope Strukturen berücksichtigt. Eine Likelihood-basierte Analyse wird verwendet, um die Kreuzkorrelation zwischen den durch das GMF abgelenkten Ankunftsrichtungen der UHECRs und einem Quellenmodell, Als Quellmodell wird hierbei ein Katalog von Starburst Galaxien verwendet. Diese Analyse liefert Einblicke in die Abweichung von der isotropen Erwartung, wenn die Ablenkung durch das GMF berücksichtigt werden. Zusätzlich werden Anisotropien mittlerer Winkelgröße am UHECR-Himmel untersucht, um Einschränkungen für die Stärke des EGMF abzuleiten. Systematische Scans über den Signalanteil und den von der EGMF verursachten Winkelverschmierung zeigen, dass Szenarien mit minimalen Beiträgen des Feldes und geringen Signalanteilen die beobachteten Daten erfolgreich reproduzieren. Im Gegensatz dazu führen Parameterkonfigurationen mit hohen Signalanteilen und unzureichenden EGMF-Ablenkungen zu signifikanten Überdichten, die von den Beobachtungen abweichen.

Vergleiche zwischen verschiedenen astrophysikalischen Modellen, einschließlich Variationen des GMF, unterstreichen zusätzlich die entscheidende Rolle der Magnetfelder bei der Gestaltung des UHECR-Himmels. Diese Studie bietet einen statistischen Rahmen, um die Kreuzkorrelation zwischen UHECR-Anisotropien und Magnetfeldern zu erforschen. Sie schlägt Einschränkungen für die Stärke solcher Felder vor, indem sie die Bereiche für die Winkelstreuung identifiziert, in denen die Häufung von Ereignissen verringert wird, um die beobachteten anisotropen Merkmale am Himmel zu beschreiben..

# Resumen

Los rayos cósmicos de ultra alta energía (UHECRs) son las partículas más energéticas observadas en el Universo. Sus direcciones de llegada proporcionan información esencial sobre su origen y propagación. Se espera que las anisotropías en el cielo de los UHECRs, observadas como desviaciones de la isotropía, estén influenciadas por las desviaciones causadas tanto por el Campo Magnético Galáctico (GMF) como por el Campo Magnético Extragaláctico (EGMF).

En este trabajo se realizan simulaciones para estudiar el impacto de las desviaciones magnéticas en las direcciones de llegada de los UHECRs, incorporando tanto los componentes coherentes como turbulentos del GMF, así como la influencia del EGMF en las estructuras anisotrópicas observadas. Se utiliza un análisis basado en verosimilitud para cuantificar la correlación cruzada entre las direcciones de llegada de los UHECRs desviadas por el GMF y un modelo de fuentes, específicamente el catálogo de galaxias Starburst, proporcionando información sobre la desviación respecto a la expectativa isotrópica cuando se incluyen los efectos del GMF. Además, se investigan las anisotropías a escalas intermedias en el cielo de los UHECRs para establecer límites a la intensidad del EGMF. Escaneos sistemáticos sobre la fracción de señal y el ángulo de dispersión relacionado con el EGMF demuestran que los escenarios con contribuciones mínimas del campo y bajas fracciones de señal reproducen con éxito los datos observados. Por el contrario, configuraciones de parámetros con altas fracciones de señal y desviaciones insuficientes del EGMF generan sobredensidades significativas que se desvían de las observaciones.

Las comparaciones entre diferentes modelos astrofísicos, incluidas variaciones del GMF, muestran además que los campos magnéticos juegan un papel crucial en la determinación de la distribución de los UHECRs en el cielo. Este estudio ofrece un marco estadístico para explorar la correlación cruzada entre las anisotropías de los UHECRs y los campos magnéticos, proponiendo límites a la intensidad de dichos campos mediante la identificación de ventanas angulares donde la agrupación de eventos se diluye para ajustarse a las características anisotrópicas observadas en el cielo.

# Chapter I

# Introduction

Ultra-high-energy cosmic rays (UHECRs) are the most energetic particles ever detected, with energies exceeding $10^{18}$ eV. Despite decades of observations, the origin, composition, and propagation of these extreme particles remain active areas of scientific investigation. UHECRs offer valuable insights into some of the most energetic astrophysical processes, such as those occurring in active galactic nuclei (AGN), starburst galaxies (SBGs), and gamma-ray bursts (GRBs). However, the deflection of UHECRs during their propagation, due to cosmic magnetic fields, complicates efforts to reconstruct their trajectories and identify their sources.

The propagation of UHECRs is significantly influenced by both Galactic and Extragalactic magnetic fields. The Extragalactic magnetic field (EGMF) remains poorly constrained due to its low strength and the challenges of isolating its effects from other astrophysical processes. The EGMF is expected to be an unordered random magnetic field. Therefore, it introduces a rigidity-dependent angular smearing effect on observed UHECR arrival directions, which can dilute anisotropies and affect event clustering. Understanding the interplay between EGMFs and UHECR propagation is crucial for interpreting the observational data. In addition to the EGMF, the Galactic magnetic field (GMF) plays an important role in the deflection of charged particles. The GMF, with its complex structure of coherent and turbulent components, influences the path of UHECRs as they travel through the Galaxy, further complicating the identification of their extragalactic origin by introducing an additional smearing as well as a coherent displacement of the arrival direction of UHECRs at Earth.

In this thesis, a study on UHECR anisotropies will be presented as they provide a powerful method for probing both the properties of cosmic magnetic fields and the

distribution of cosmic ray sources. Observations from the Pierre Auger Observatory and the Telescope Array Project have revealed evidence of anisotropic structures, such as overdensities in specific regions of the sky, which suggest potential correlations with nearby extragalactic sources. By analyzing these anisotropies, it is possible to place constraints on the strength of the EGMF and GMF, while also identifying candidate sources for the emission of UHECRs.

Chapter 2 of this thesis provides an overview of the key observables related to UHECRs, along with the processes that govern their acceleration. It also discusses current detection techniques, with a particular focus on the Pierre Auger Observatory. Additionally, this chapter introduces the deflections of both EGMF and GMF on the propagation of charged particles. In chapter 3, an overview of the simulations developed for in this study is presented, along with an analysis of how variations in the GMF influence the propagation of charged particles. This sets the foundation for the cross-correlation analysis between the arrival directions of charged particles, influenced by the GMF, and the Starburst Galaxy source catalog which is discussed in detail in chapter 4. In Chapter 5, a study of intermediate-scale anisotropies in the UHECR sky is presented, examining regions in the sky that deviate from isotropic expectations. This analysis is conducted on both observed data and simulations, which incorporate the angular deflections caused by the GMF and EGMF contributions. The findings are used to investigate the sensitivity of current UHE data to set constraints on the strength of these magnetic fields.

# Chapter II

# Ultra-high energy cosmic rays

Cosmic ray science is a broad field deeply interconnected with various branches of physics. Until the mid-20th century, it provided the sole means for scientists to explore high-energy particle phenomena. Even today, cosmic rays remain invaluable messengers from the Universe, as they can be accelerated to energies that surpass those achieved by the Large Hadron Collider (LHC) at CERN by a factor of 10 million.

Several experiments such as KASCADE [1], KASCADE-Grande [2], the Telescope Array Project [3], the IceCube neutrino Observatory [4] and the Pierre Auger Observatory [5] detected over the years cosmic rays after their propagation from their origins to Earth. Once such energetic particles reach Earth, they interact with the atmosphere generating a cascade of billion of particles, that are detected at Earth employing ground-based detectors. Among the key observable that are obtained through the analysis of such data the attention is focused on the energy spectrum and the arrival direction.

## 2.1. Key observables

### 2.1.1. Energy spectrum

The energy spectrum offers a quantitative measure of the particle flux as a function of energy. It is commonly expressed as the differential flux per unit of energy e.g. [6]:

$$J(E) = \frac{\mathrm{d}^4 N}{\mathrm{d}E\,\mathrm{d}A\,\mathrm{d}\Omega\,\mathrm{d}t}, \tag{2.1}$$

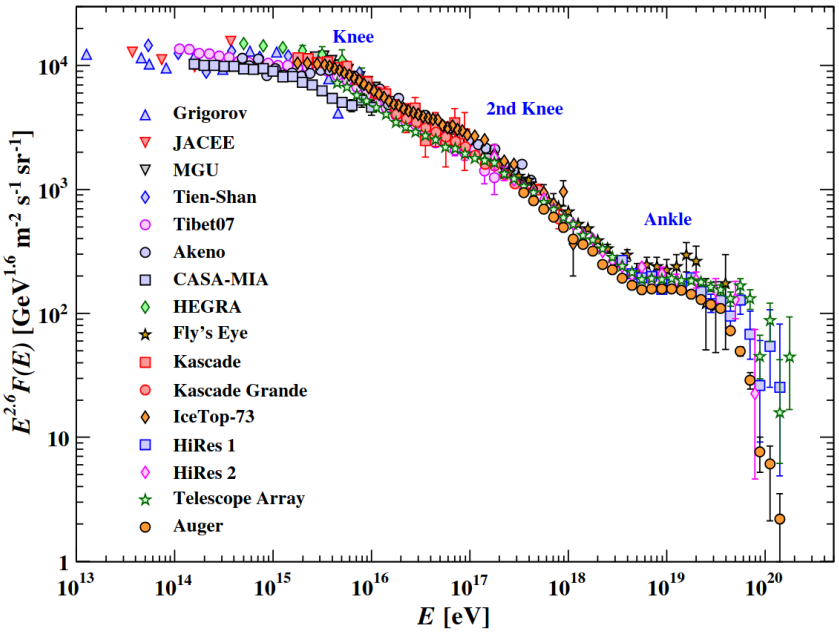



**Figure 2.1:** Observed cosmic ray spectrum obtained incorporating data from different experiment [7].

where $J(E)$ is the energy spectrum, $N$ represents the number of particles, $E$ is the particle energy, $A$ is the detector area, $\Omega$ is the solid angle of observation, and $t$ is the observation time. Understanding the energy spectrum is fundamental for interpreting the results of UHECR studies, as it reflects not only the nature of cosmic ray sources but also the effects of propagation and interaction processes.

The spectrum spans many orders of magnitude in energy and the particle flux exhibits a steep decrease with increasing energy, typically following a power-law form $J(E) \propto E^{-\alpha}$, where $\alpha$, the spectral index, quantifies the steepness of the spectrum. The changes in the spectral index correspond to distinct features that provide insights into the origins and propagation mechanisms of cosmic rays. Key features, that can be seen in Fig. 2.1, where observations related to several experiments are displayed, include:

- The Knee: At energies around $3 \times 10^{15}$ eV, the spectral index hardens, transitioning from $\alpha \sim 2.7$ to $\alpha \sim 3.0$. This feature, known as the knee, is commonly associated with the transition towards the heavier Galactic elements [8].

- The Second Knee: Near $10^{17}$ eV, another change in the spectral index is observed. This transition is thought to mark the diminishing influence of Galactic sources and the increasing dominance of extragalactic contributions [9].

- The Ankle: Around $5 \times 10^{18}$ eV, the spectrum flattens, indicating the transition to a fully extragalactic origin for UHECRs. This feature likely reflects a transition between two different extragalactic components, with the one above the ankle requiring powerful sources with harder spectrum [10].

- The Flux Suppression: At the highest energies, approximately 100 EeV, the spectrum shows a significant suppression. This feature might be attributed to energy losses, in which case it is referred to as the Greisen-Zatsepin-Kuzmin (GZK) cutoff [11,12], due to interactions with cosmic microwave background photons during propagation and/or to the maximum energy that can be achieved by the sources (see Sec. 2.4.1).

### 2.1.2. Arrival direction

Alongside the energy spectrum, another important observable that carries relevant information concerning UHECRs is their arrival direction. Measuring the arrival direction distribution of such particles could give important insights regarding their propagation and their origins. The main studies on the arrival directions of UHECRs have been conducted with the intention of analyzing a possible deviation from isotropically distributed events at different energy ranges see e.g. [13,14]. The first aspect that indicates a deviation from isotropy can be seen in large-angular-scale anisotropies, that are described in the following section.

#### Large-angular-scale anisotropy

The large-scale anisotropy of the arrival direction of cosmic rays has been studied using 19 years of data collected by the Pierre Auger Observatory [15]. For energies above the full-efficiency threshold of the Observatory, a combined Fourier analysis has been employed [16]. This approach is sensitive to two components of the dipole anisotropy: the equatorial component ($d_\perp$), inferred from Right Ascension (R.A.) modulations, and the North-South component ($d_z$), derived from azimuthal variations. Together, these components allow for a three-dimensional reconstruction of the dipolar component defined as [16]:

$$d_\perp \simeq \frac{r_1^\alpha}{< cos\delta >},\ d_z \simeq \frac{b_1^\phi}{cos(l_{obs}) < sin\theta >},\ \alpha_d = \varphi_1^\alpha,\ \delta_d = \arctan\left(\frac{d_z}{d_\perp}\right) \tag{2.2}$$

where ($\alpha_d$,$\delta_d$) indicates the dipole direction in Equatorial coordinates, $r_1^\alpha$ indicates the amplitude of the event rate modulation of order k=1, $< cos\delta >\sim 0.7814$ indicates the average cosine of the events, $< sin\theta >\sim 0.6525$ indicates the mean sine of the zenith of the events and $l_{\rm obs} = -35.2°$ indicates the latitude of the Pierre Auger Observatory. The analysis indicates that for $E \geq 8$ EeV, the statistical significance of the dipolar modulation in right ascension (R.A.) reaches $6.8\sigma$. The dipole is observed to be located approximately 120° away from the Galactic center, as presented in Fig. 2.2. This position, along with its orientation pointing away from the Galactic center, suggests an extragalactic origin for the sources. At high energies, charged particles experience smaller deflections due to their increased rigidity, which results in their arrival directions more closely reflecting their sources. Consequently, a dipole pointing away from the Galactic center implies that the sources are not concentrated within the Galaxy but are distributed in extragalactic space. Moreover, the Galactic magnetic field provides valuable insights into the propa-

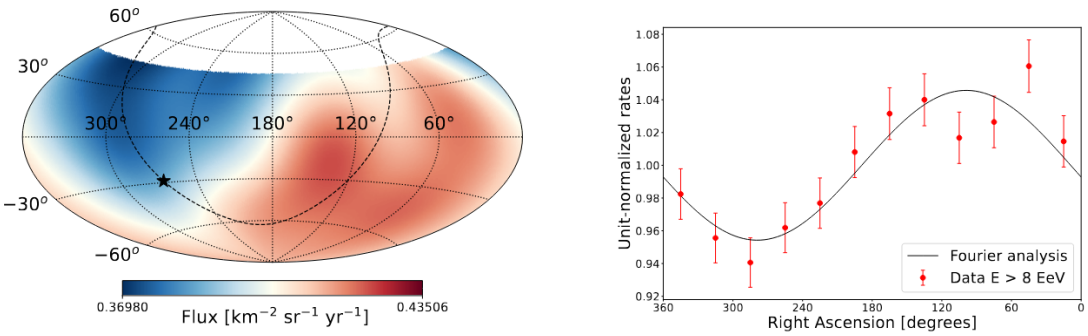



**Figure 2.2:** Left: Flux obtained above 8 EeV presented with a Fisher smoothing. The Galactic plane is indicated with a dashed line while the position of the Galactic center is indicated with a black star. Right: Normalized rates in R.A. presented with red dots. The black line indicates the result of the Fourier analysis.

gation of cosmic rays. The observed dipole direction, pointing away from the Galactic center, indicates that the influence of the GMF on the dominant cosmic ray population is negligible in this case. This suggests that the particles have sufficient rigidity to minimize deflections, maintaining the large-scale anisotropy associated with extragalactic sources. Combined with the dipole's amplitude, these observations suggest that the primary contributions to the cosmic ray flux originate from extragalactic sources, where the effects of the GMF are minimal.

A similar analysis has been conducted for different energy thresholds, as summarized in Fig. 2.3, which shows the evolution of the dipole's direction and amplitude with increasing energy. The observed increase in the dipole amplitude with energy can be

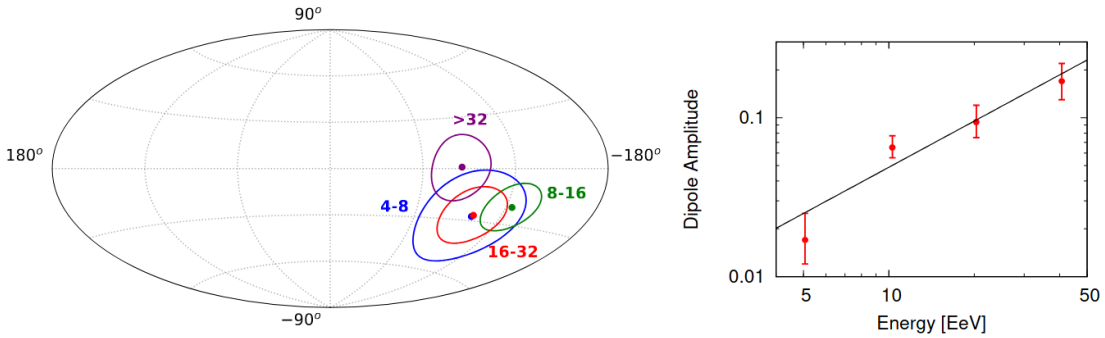



**Figure 2.3:** Left: Direction of the dipole for different energies. Right: Evolution of the amplitude of the dipole with energy.

attributed to a significant contribution from nearby sources that are unevenly distributed. Additionally, at higher energies, the mean rigidity ($R = E/Z$) of particles reaching Earth increases, resulting in smaller deflections. This reduced deflection leads to larger dipolar amplitudes. Another indication that the large-angular-scale provides regarding the possible extragalactic origin on UHECRs is related to the reconstruction of the equatorial component of the dipole for low energies at 0.03 EeV. Fig. 2.4 illustrates that

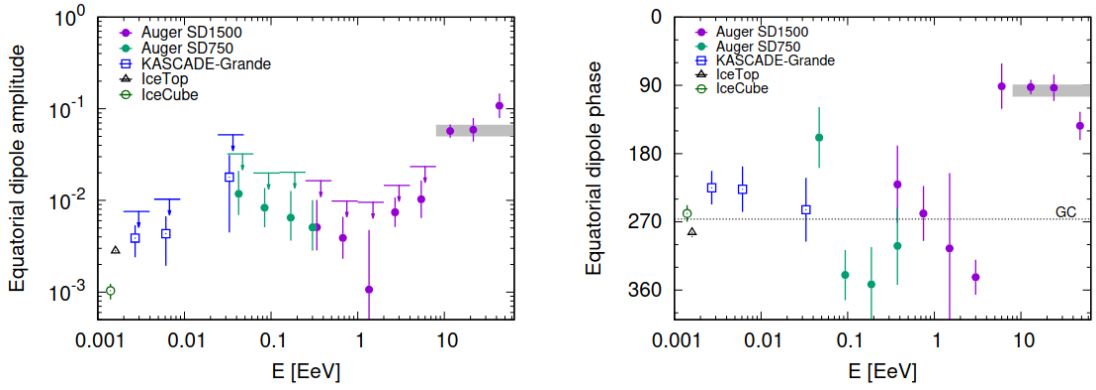



**Figure 2.4:** Left: Evolution of the amplitude of the dipole with the energy. Right: Evolution of the phase of the dipole with energy.

also in this case the amplitude of the dipole increases with energy and its phase presents a shifting in direction from a direction close to the Galactic center to a direction that points away from the Galactic center [17].

Investigations into the potential origins for UHECRs can be extended to intermediate angular scales, particularly by examining the high-energy end of the spectrum. Adjusting the energy threshold to higher values may reveal important details about the

presence of distinct structures and clustered events in the observed flux. These features, which deviate significantly from isotropy, can provide valuable insights into the origins of these particles see e.g. [18–20].

### Intermediate-angular-scale anisotropy

To investigate the arrival directions of UHECRs, the Pierre Auger Collaboration conducted an analysis above 32 EeV with more than 2500 events, collected between 2004 and 2020, with three methods [21]:

**Blind search:** a blind search over the overall sky is conducted performing a scan in the angular window starting from 1° to 30° with a step of 1°. Similarly, a scan over the energy is performed from 32 EeV to 80 EeV with a step of 1 EeV. For every step of the scan the difference between the observed events and the isotropic expectation is computed. The discrepancy between data and isotropic expectation is presented in Fig. 2.5 where the localized excess in the Centaurus region is observed with a pre-trial significance of $5.4\sigma$ for an energy threshold of 41 EeV and an angular window of 24°. The post-trial significance of such excess is 3% which indicates the low significance of such excess.

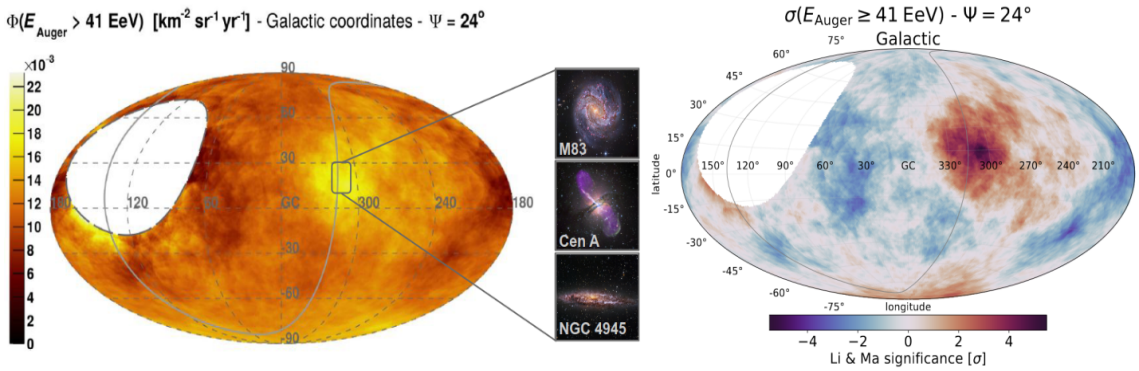



**Figure 2.5:** (a): Observed flux at energies above 41 EeV in Galactic coordinates with a top-hat angular smoothing of 24°. The Centaurus region is highlighted together with the sources contained in it. (b): Local significance computed employing events above 41 EeV. The highest excess can be visibile in the Centaurus region once a top-hat smoothing of 24° is applied [21].

**Correlation with prominent sources and structures:** similarly to the blind search in the overall sky, also regions of major interest have been investigated. Large structure such as the Galactic and supergalactic plane have been monitored with no significant

deviation from isotropy. Although, as shown in Fig. 2.6, an excess has been detected in the Centaurus region, whereas the targeted search was performed around the position of the core of the Centaurus A radio galaxy, a source that has been often proposed as main source for UHECR due to its close distance and physical properties, with a significance of $3.9\sigma$.

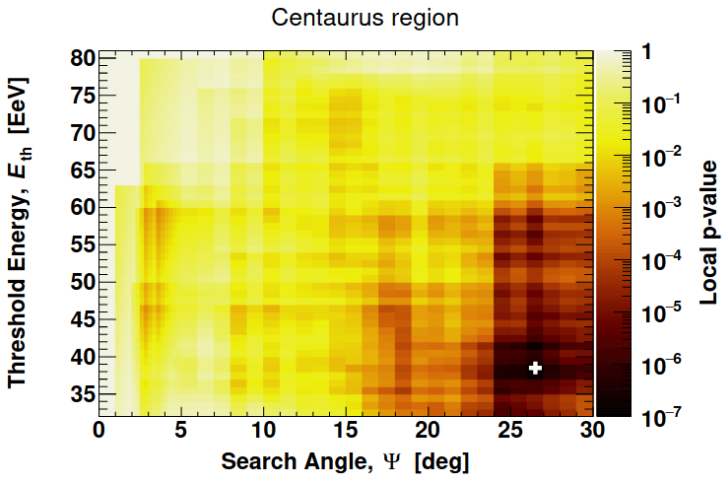



**Figure 2.6:** Local p-value related to the blind search conducted in the Centaurus region as a function of the top-hat search radius and energy threshold. The minimum value of the p-value is displayed as white cross [21].

**Correlation with astrophysical source catalogs:** alongside the blind search conducted on several regions of the sky, also a possible correlation between data and source catalogs has been analyzed. The source catalogs included in the analysis are the 2MASS [22], which maps galaxies in the near-infrared that cover almost the entire sky, Starburst Galaxies (SBGs) [23], X-ray emitting Active Galactic Nuclei (AGNs) [24] and jetted AGNs [25]. A likelihood ratio (test statistic, TS) has been employed to determine if a flux map constructed with different catalog sources can provide a better description of the observed flux when compared to an isotropic expectation. The largest deviation from isotropy has been reported for the SBG catalog with a significance of $4\sigma$ above the energy threshold of 38 EeV. Fig. 2.7 shows the energy dependence of the TS for the different source catalogs when compared to isotropy. It is easy to see that the SBG catalog provides the largest deviation from isotropy at every energy considered. A similar behavior is observed for the Centaurus region that contains the closest sources such as NGC4945, M83 and Centaurus A. A more detailed description of the likelihood ratio and correlation with source catalog will be provided in Chap. 4. It is important to note that the effect

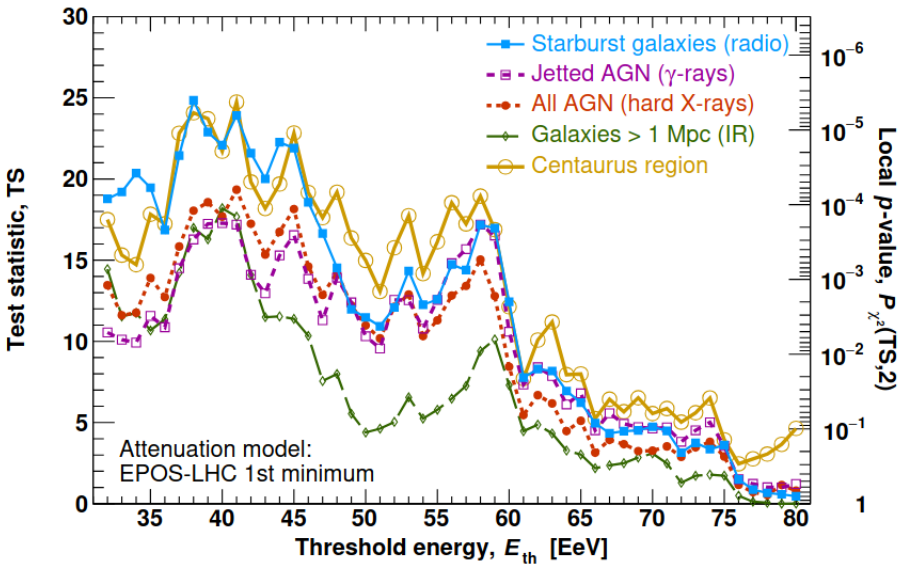



**Figure 2.7:** Evolution of the test statistic as a function of the energy threshold for the source catalogs employed in the analysis [21].

of the Galactic magnetic field has not been included in this analysis, and no coherent deflections have been considered. This topic will be addressed and analyzed further in Chap. 3.

The detailed observations and comprehensive analyses of cosmic ray data have shown several significant properties, notably, the measured flux of cosmic rays appears to diverge from the isotropic expectations. This divergence suggests that there is a considerable contribution from extragalactic sources, hinting at the complex and diverse nature of the environments from which these particles originate. In the following section, a brief overview of the techniques used to obtain the observed data, along with a discussion on the mechanisms that can generate high-energy cosmic rays, will be presented.

## 2.2. Detection of ultra-high energy cosmic rays

When UHECRs enter the Earth's atmosphere, they interact with air nuclei at altitudes of tens of kilometers, initiating extensive air showers (EAS). These cascades consist of secondary particles, which include electrons, positrons, photons, muons, and hadrons. The development of an air shower is governed by the energy and composition of the primary cosmic ray, along with its interactions with atmospheric elements. Through the detection and study of EAS, it is possible to infer key properties of the primary particles, such as their energy, mass, and arrival direction.

### 2.2.1. Extensive air shower

When cosmic rays interact with the atmosphere, the resulting air shower can be categorized into three interconnected components. The electromagnetic component primarily consists of electrons, positrons, and photons, which are produced through electromagnetic cascades. The hadronic component includes hadrons, kaons, and pions, which are generated during high-energy nuclear interactions. Lastly, the muonic component comprises muons and muonic neutrinos, originating from the decay of charged pions and kaons. These components collectively describe the complex nature of air showers initiated by cosmic ray interactions. The development of the three components in the atmosphere can be described using the vertical shower depth, defined as:

$$X_h = \int_h^\infty \rho(h)\,\mathrm{d}h \tag{2.3}$$

where $h$ represents the height above sea level. When a particle initiates an air shower by entering the atmosphere at an angle $\theta$, the shower depth can be expressed as a function of this angle: $X = X_h/\cos(\theta)$.

The air shower can be initiated by photons and electrons, resulting in a purely electromagnetic air shower (EM), or by protons and heavier nuclei, which creates a hadronic air shower initiated by nuclei. An overview of the modeling of EAS is presented in Fig. 2.8 where the development of electromagnetic and hadronic showers are displayed.

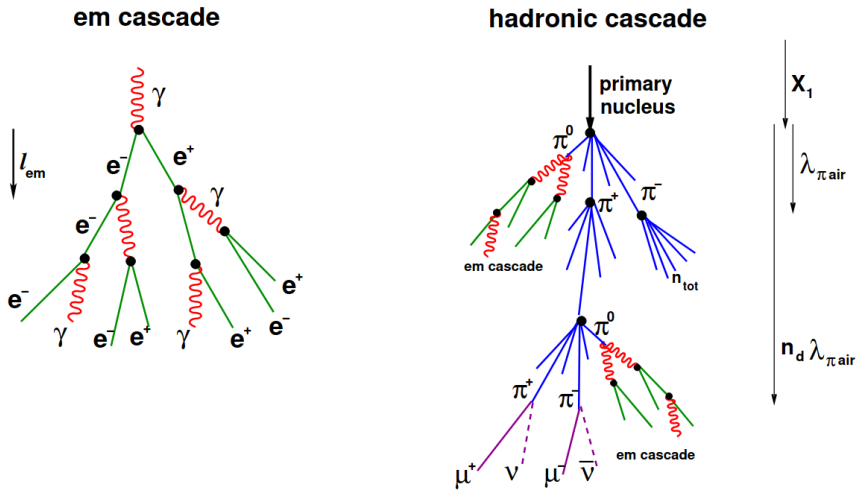



**Figure 2.8:** Schematic view of the development of the air showers when initiated by EM component (left) and nuclei (right) [6].

In the case of a purely EM shower, simple modeling of the cascade can be achieved using the Heitler model [26]. When high-energy photons interact with the atmosphere, they produce electron-positron pairs ($e^+e^-$) through interactions with atomic nuclei. The resulting electrons and positrons undergo bremsstrahlung interactions, emitting photons. Other processes such as ionization and excitation interactions become significant only when particle energies fall below the critical energy, $E_c \simeq 84\,\text{MeV}$.

In order to describe these interactions, it is assumed that after a characteristic interaction length $\ell_{\text{em}}$, the particles produced in the processes through

$$\gamma \rightarrow e^+e^- \text{ and } e^\pm \rightarrow e^\pm\gamma \tag{2.4}$$

equally share the energy. Under this assumption, the number of particles in the shower doubles at each stage, while their average energy halves. After $n$ interactions during the development of the shower, the number of particles produced is $N = 2^n$, and their energy is $E = E_0/N$, where $E_0$ is the energy of the primary particle. This geometrical growth continues until the particle energy falls below $E_c$, at which point ionization and excitation interactions become significant leading to energy dissipation. Therefore, it is expected that the shower reaches its maximum development when all particles reach $E_c$. At this limit, the maximum number of particles produced in the shower can be expressed as:

$$N_{\text{max}} \simeq \frac{E_0}{E_c} \tag{2.5}$$

This is an important relation that connects the maximum amount of particles produced to the initial energy of the primary particle. The maximum number of particles can be used also to quantify the max depth of the shower as follow:

$$X_{\text{max}} = X(N_{\text{max}}) = X_0 \ln\left(\frac{E_0}{E_c}\right) \tag{2.6}$$

where $X_0$ is the electromagnetic radiation length in the air ($X_0 \simeq 37\ \text{g/cm}^2$).

This simplistic model shares some similarities with the more complex hadronic showers, although several differences arise due to hadronic interactions as presented in Fig. 2.8. After interacting with the atmosphere, the nuclei that initiate the air shower produce a large number of secondary particles, predominantly pions. The $\pi^0$ decays into photons, that initiate an EM sub-shower, while charged pions re-interact or decay into muons and neutrinos. The interplay between decays and re interactions, as well as the continuous feeding of the EM component by the hadronic core, results in a more

complex longitudinal development.

The depth of the maximum for proton-initiated showers can be estimated as:

$$X_{\max} \simeq X_1 + X_0 \ln\left(\frac{E_0}{2n_{\text{tot}}E_c}\right)$$

where $X_1$ is the depth of the first interaction given by the proton-air cross-section, and the second term accounts for the development of EM sub-showers initiated by photons, with energies $E_\gamma \simeq E_0/2n_{\text{tot}}$, from $\pi^0$ decays where $n_{\text{tot}}$ indicates the total multiplicity mostly dominated by pions of all three charges. This expression highlights the dependence of $X_{\max}$ on features of hadronic interactions, such as the inelastic cross-section and the multiplicity of secondary particles. Moreover, the depth of shower maximum for cosmic rays is influenced by the primary mass, with heavier nuclei resulting in shallower $X_{\max}$ values compared to protons. This is because the energy per nucleon is lower in heavier nuclei, leading to less penetrative sub-showers. The relationship between $X_{\max}$ and the primary mass is given by $X^A_{\max}(E) \simeq X^{\text{p}}_{\max}(E/A)$, where larger mass numbers lead to smaller $X_{\max}$ values. The superposition model approximates the shower development for heavy nuclei as a sum of A proton showers, implying that the average $X_{\max}$ for a nucleus is similar to that of protons with a lower energy per nucleon. Fluctuations in $X_{\max}$ also differ for protons and heavy nuclei, with proton showers exhibiting larger variance due to the first interaction point, whereas heavy nuclei show smaller fluctuations as the variance from individual proton showers averages out. Additionally, the superposition model is an approximation, as nucleons within a nucleus are not independent. The variance in $X_{\max}$ for mixed mass compositions is larger, requiring careful interpretation of measurements. Predictions on the air shower properties can be performed using simulation codes such as AIRES [27] or CORSIKA [28], in which different models for treating the hadronic interaction can be used such as QGSJET [29], EPOS [30] and SIBYLL [31].

### 2.2.2. Observation of air showers

When the cascade of secondary particles reaches the ground, it is detected using ground-based detectors. These detectors are designed to study the secondary particles generated during the cascade in order to gather information about the origin, composition, and properties of the primary cosmic ray that entered the atmosphere.

Several techniques can be used to detect information regarding the properties of the air shower: one consists of arranging particle detectors on a sparse grid at ground level

to measure the footprint of the shower at ground. An example of such particle detector is the water Cherenkov detector (WCD), composed by a cylindrical tank filled with pure water and equipped with photomultiplier tubes (PMTs) to detect Cherenkov radiation produced by the charged particles passing through water. The Cherenkov radiation is produced since the charged particles that reach the ground, travel faster than the speed of light in water, emitting Cherenkov light [32]. This radiation is detected by the PMTs and the information regarding the signal amplitude and timing are used to obtain information about the particle density and arrival direction of the air shower. By analyzing the arrival times of the shower front at different WCDs, it is possible to reconstruct the arrival direction of the primary cosmic ray. On the other hand, the energy of the shower is estimated by comparing its lateral distribution of particles to simulations. However, this technique has a disadvantage: the energy estimation carries a larger uncertainty compared to other methods because it relies on indirect measurements and comparisons with simulations rather than a direct measurement of the energy deposited in the atmosphere. A visualization of the WCD is presented in Fig. 2.9. Another technique relies on the fact that the atmosphere acts as a calorimeter, where the energy of the air shower can be measured through the observation of fluorescence light. When charged particles in the air shower travel through the atmosphere, they excite nitrogen molecules, which then emit ultraviolet fluorescence light. This light is collected by large mirrors in the fluorescence telescopes and focused onto PMTs, which record the intensity of the light as a function of time. The recorded light profile enables the reconstruction of the air shower's longitudinal development, including the determination of the depth of the shower maximum which carries important information concerning the primary particle's mass and energy. Moreover, the fluorescence light is proportional to the energy deposited in the atmosphere, providing a direct measurement of the shower's energy. Although FDs provide precise energy and composition measurements, they have significant limitations. Their operation is restricted to specific conditions—moonless nights with clear skies—leading to a duty cycle of approximately 10%. Furthermore, accurate measurements require detailed knowledge of the atmospheric conditions to correct for light attenuation, scattering, and other atmospheric effects [35]. Compared to an array of WCDs, the reconstruction of air showers using fluorescence light detection reduces the reliance on simulations for energy estimation and allows for a direct measurement of the depth of shower maximum.

Alongside the observations of the fluorescence light and Cherenkov light, also the radio observation can be used to gain information about the air showers. After the initiation of the shower in the atmosphere, the cascade of particles can emit electromagnetic

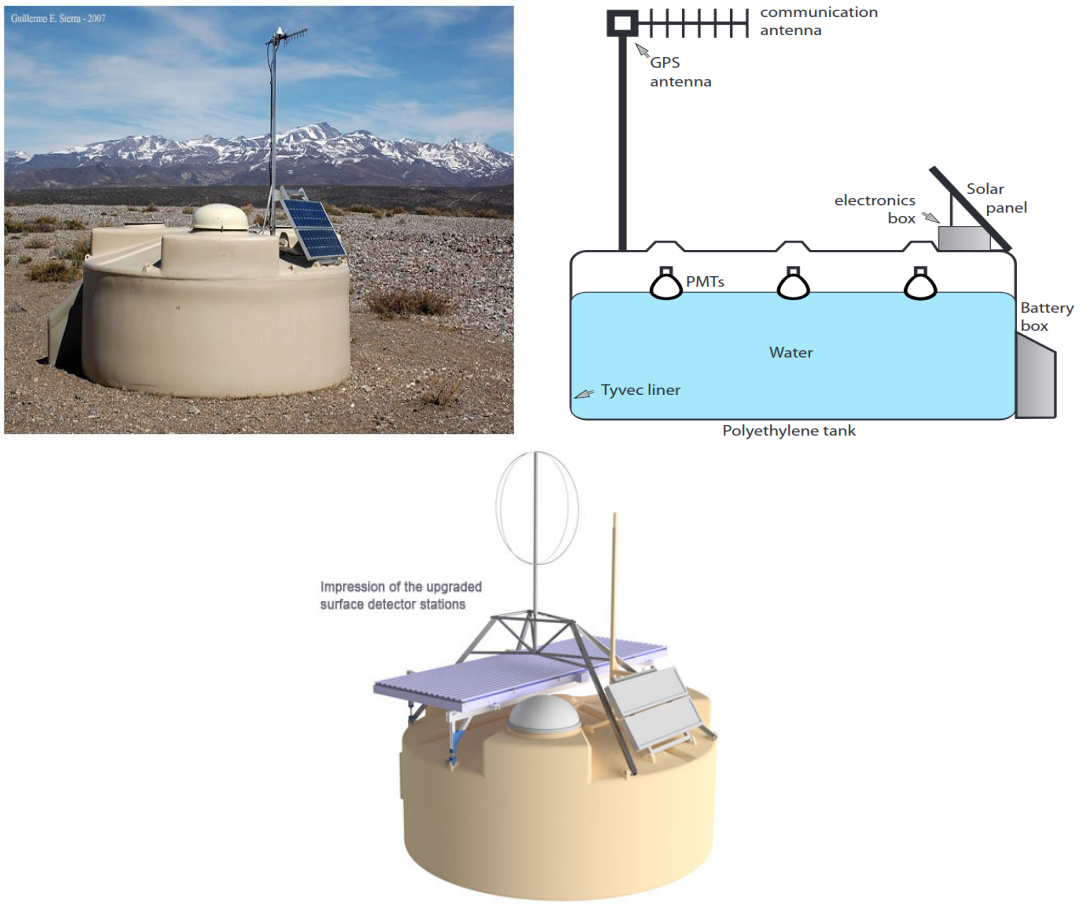



**Figure 2.9:** Visualization of the WCD from the outside (left) and a cross-section where the position of the PMTs is displayed (right) [33]. Schematic overview of the updated SD station with the installation of the scintillator and radio antenna (bottom) [34].

radiation, primarily in the radio frequency range, due to the acceleration of charges in the shower, especially at high energies. By using large antenna arrays, the radio detectors are able to capture the pulses from air showers with high precision. This method has several advantages over traditional detection techniques, including the ability to detect air showers at very large distances and to measure the energy of cosmic rays directly through the radio signal's characteristics.

However, as all methods have inherent limitations, a more effective detection strategy is achieved by combining the several techniques. This hybrid detection approach is employed at the Pierre Auger Observatory, whose data have been used in this thesis. Therefore, a brief overview of the Observatory, along with its detectors, is presented in the next section.

## The Pierre Auger Observatory

The Pierre Auger Observatory, the world's largest facility for studying ultra-high-energy cosmic rays, is located in Malargüe, Argentina, at an altitude of 1400 meters above sea level. Spanning an area of 3000 km$^2$, the observatory comprises a hexagonal array of 1660 WCDs, as shown in Fig. 2.10. These detectors achieve an angular resolution of approximately 1°, enabling precise reconstruction of the arrival directions of cosmic rays. The energy of the primary cosmic ray is determined through cross-calibration with events observed by both the surface detector (SD) and fluorescence detector (FD) systems.

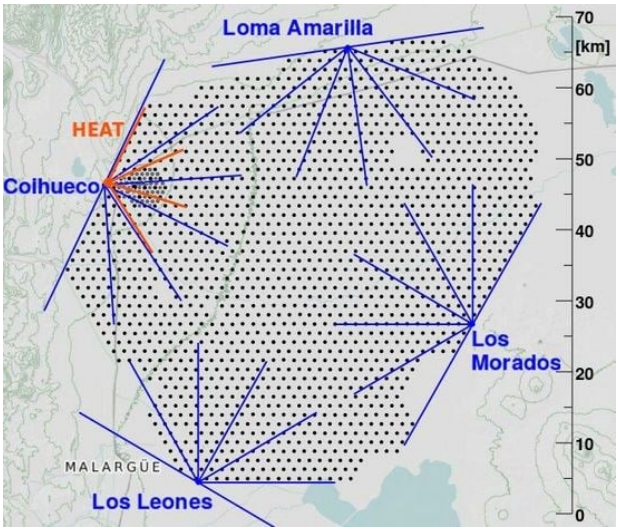



**Figure 2.10:** A schematic overview of the Pierre Auger Observatory, showing the arrangement of the water-Cherenkov detectors (WCDs) and the field of view of the fluorescence telescopes [34].

The observatory also features 27 fluorescence telescopes distributed across five buildings (Coihueco, Loma Amarilla, Los Morados, Los Leones, and HEAT) [36]. The fields of view of these telescopes are represented as blue and red lines in Fig. 2.10. Overlooking the SD array, these telescopes facilitate the simultaneous detection of air showers by both systems. Events observed concurrently by the FD and SD are critical for cross-calibrating the energy scale since it reduces the systematic uncertainty on the energy scale to approximately 14% [37]. A visualization of a fluorescence telescope is shown in Fig. 2.11, illustrating its key components, including the large mirrors and PMTs that enable this method of detection.

The FD directly measures the longitudinal profile of the air shower by capturing the

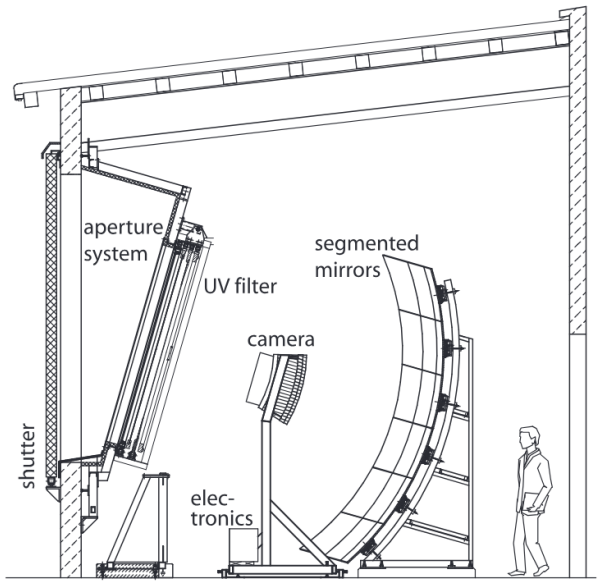



**Figure 2.11:** Visualization of the fluorescence telescope employed in the Pierre Auger Observatory alongside its keys component used for detection of air showers [36].

fluorescence light signal using the triggered PMTs in its camera. This signal is combined with timing information and the reconstructed location of the shower core, which is obtained from at least one SD station. The energy deposited along the shower track ($\mathrm{d}E/\mathrm{d}X$) is modeled using a Gaisser-Hillas function [38]. From this fit, it is possible to determine the depth of shower maximum ($X_{\mathrm{max}}$) and the energy deposited in the atmosphere. This method achieves an energy resolution of better than 10%, with a systematic uncertainty of 14% on the energy scale. Further details on energy reconstruction with the FD can be found in [36, 39, 40]. In addition to the SD and FD systems, the Pierre Auger Observatory includes several other detection systems designed to study specific aspects of EAS and enable multi-hybrid event detection.

Notable additions to the previous systems, that can be seen in Fig. 2.9, are:

- Auger Engineering Radio Array (AERA), which measures radio emissions from EAS as they propagate through the atmosphere. AERA consists of 153 radio antennas spread across 17 km$^2$. This system provides additional insights into the energy scale and the depth of the shower maximum by correlating its observations with those of the FD. Such cross-correlations improve the accuracy of energy and shower profile measurements.

- The Auger Muons and Infill for the Ground Array (AMIGA) system focuses on

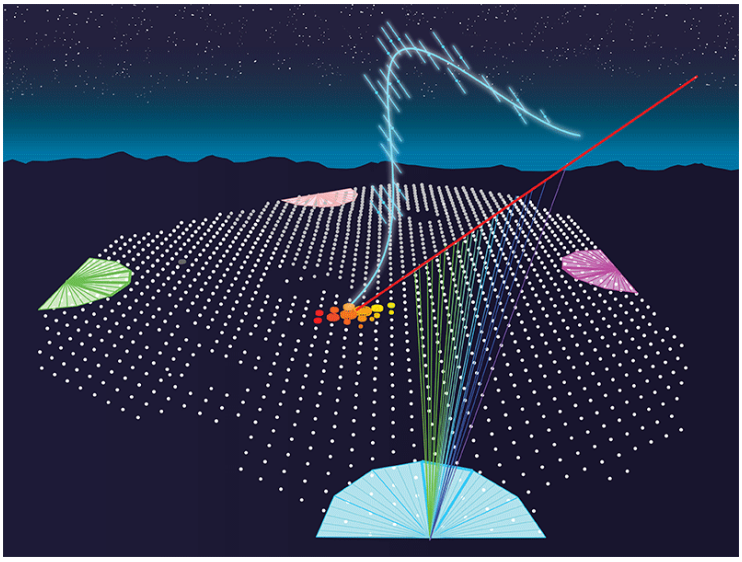


**Figure 2.12:** Schematic view of a hybrid event observed at the Pierre Auger Observatory. The red line shows the evolution of fluorescence light over time, while the yellow/red color coding represents the detection time of secondary particles at the ground by the SD stations. The reconstructed profile of $X_{max}$ is shown in blue [41].

detecting muons within air showers. It comprises underground scintillators dedicated to measuring muons and a denser array of SD stations with a spacing of 433 meters, complemented by the 750m Infill array designed to enhance the detection capabilities in regions of higher station density. By distinguishing muons—which penetrate deeper into the ground—from the electromagnetic components detected by surface stations, AMIGA allows precise monitoring of EAS composition. Additionally, the denser array extends the observatory's sensitivity to lower-energy cosmic rays.

These additions characterized the upgrade of the Pierre Auger Observatory, called *AugerPrime*, where the stations in the SD array are now equipped with a plastic scintillator, a radio antenna, and small photomultiplier tubes coupled with new readout electronics. The upgraded electronics enhance the precision of data acquisition, ensuring better synchronization and efficiency. AugerPrime is designed to operate until at least 2035, providing long-term, high-precision measurements of cosmic ray properties such as energy, shower depth, and arrival direction, contributing to the ongoing study of UHECRs.

## 2.3. Sources of ultra-high cosmic rays

The data collected by the Pierre Auger Observatory has provided valuable insights into the potential sources of UHECRs. Analyses of the large-scale dipole anisotropy and correlations with extragalactic source catalogs suggest an extragalactic origin for UHECRs, although the exact sources remain uncertain.

One hypothesis for their origin lies in top-down models, which propose that UHECRs arise from the decay of super-heavy relic particles. However, recent measurements of neutrino and photon fluxes have effectively ruled out these models [42, 43]. Instead, bottom-up models, which consider that UHECRs are accelerated within powerful astrophysical sources, currently offer the most consistent explanation for their origins.

### 2.3.1. One-shot and stochastic acceleration

In bottom-up scenarios, one-shot acceleration involves particles gaining energy directly from electric fields. While such processes are effective in Earth-based accelerators, large-scale electric fields in intergalactic environments are typically neutralized by surrounding particle plasmas. Nonetheless, specific environments such as black hole magnetospheres, white dwarfs, or neutron stars with strong magnetic fields can sustain the conditions necessary for acceleration.

Another prominent mechanism for UHECR acceleration is stochastic acceleration, driven by interactions with magnetic fields in astrophysical structures [44, 45]. The second-order Fermi acceleration model, introduced in 1949 [46], describes particles undergoing repeated scatterings off moving magnetized clouds with velocity $\beta$ (in natural units). Head-on collisions, resulting in energy gain, are more frequent than tail-on collisions. The average energy gain per interaction scales as:

$$\left\langle \frac{\Delta E}{E} \right\rangle \propto \beta^2,$$

giving the process its characteristic name.

Second-order Fermi acceleration, however, is inefficient for achieving ultra-high energies due to the low velocities of magnetized clouds ($v \sim 10\,\mathrm{km/s}$) and long intervals between collisions ($\sim 10^7$ s). To overcome these limitations, a more efficient model known as diffusive shock acceleration (DSA), or first-order Fermi acceleration, has been proposed. In this mechanism, particles interact between the unshocked upstream medium and the shocked downstream medium of supersonic shock fronts.

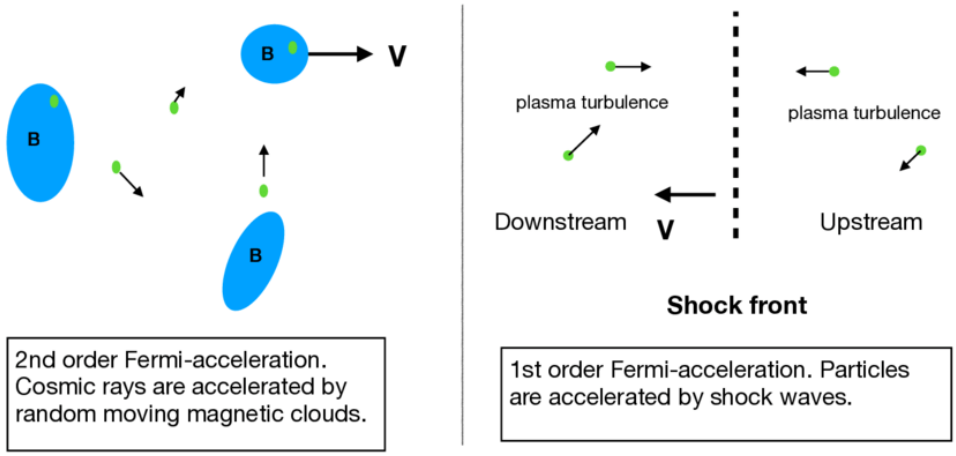



**Figure 2.13:** Illustration of second-order and first-order Fermi acceleration mechanisms in astrophysical media and shock waves [47].

Unlike the second-order model, the energy gain per interaction in DSA scales directly as $\beta$, the velocity of the shock front relative to the speed of light:

$$\left\langle \frac{\Delta E}{E} \right\rangle \propto \beta.$$

This direct proportionality to $\beta$ reflects the systematic energy gains achieved during repeated crossings of the shock front. Head-on collisions dominate due to the nature of the flow, consistently increasing particle energy with each crossing. This efficiency makes DSA a first-order process, capable of accelerating particles to ultra-high energies. Stochastic acceleration also naturally produces a power-law energy spectrum for the accelerated particles:

$$N(E) \propto E^{-p},$$

where the spectral index $p$ depends on the compression ratio of the shock front. Fig. 2.13 illustrates the differences between second-order and first-order acceleration mechanisms in astrophysical environments, highlighting the role of shocks in driving efficient particle acceleration.

### 2.3.2. Energy spectrum from Fermi acceleration

Stochastic acceleration is a strong candidate for UHECR production due to its ability to produce a power-law energy spectrum at injection, consistent with the observed spectrum on Earth. The energy gain after $n$ collisions can be expressed as:

$$E_n = E_0(1+\zeta)^n, \tag{2.7}$$

where $E_0$ is the initial energy and $\zeta = \Delta E/E$. The number of collisions required to reach energy $E$ is given by

$$n = \frac{\ln\left(\frac{E}{E_0}\right)}{\ln(1+\zeta)}. \tag{2.8}$$

The number of particles, $N$, that can reach energies above $E$ is proportional to $(1-P_{\text{esc}})^n/P_{\text{esc}}$, where $P_{\text{esc}}$ is the escape probability from the medium. This relationship leads to

$$N(>E) \propto \frac{1}{P_{\text{esc}}}\left(\frac{E}{E_0}\right)^{-\alpha} \sim E^{-\alpha}, \tag{2.9}$$

where

$$\alpha = \frac{\ln\left(\frac{1}{1-P_{\text{esc}}}\right)}{\ln(1+\zeta)}.$$

The resulting differential spectrum is:

$$\frac{\mathrm{d}N}{\mathrm{d}E} \propto E^{-\alpha-1} \sim E^{-\gamma}. \tag{2.10}$$

For non-relativistic shocks, the expected spectral index is $\gamma = 2$, while for relativistic shocks in first-order acceleration, $\gamma \sim 2.3$ [48].

An important consideration concerning the Fermi acceleration is related to the limit on the highest reachable energy, as explained in [49], that can be derived to be:

$$E_{\text{max}} \leq \frac{3}{20}\frac{u_1}{c}ZeB(u_1 T_A) \tag{2.11}$$

where $u_1$ is the speed of the shock, $Z$ the atomic number of the particle, $B$ the magnetic field and $T_A$ the lifetime of the accelerator.

It is important to notice that this relation shows also limitation of the life-time of the accelerator. Considering a supernova remnant (SNR) with a mass $M \sim 10M_{\text{sun}}$, a velocity $u_1 \sim 10^9$ cm/s and a duration of the free streaming phase of $\sim 10^3$ years, the

maximum acceleration energy is estimated as in [49] to be of the order of:

$$E_{\max} = 2.4Z \times 10^5 \text{ GeV} \tag{2.12}$$

This results in $\sim 10^{14} - 10^{15}$ eV for protons and in $\sim 10^{17}$ eV for heavier elements. Part of the observed spectrum might be described by such sources although the UHE part of the observed spectrum requires sources to accelerate particles to energies above $10^{20}$ eV.

### 2.3.3. Hillas criterion

Accelerating cosmic rays to energies as high as $10^{20}$ eV was introduced as the minimum requirements for accelerators, often referred to as the "Hillas condition" [50]. According to this condition, a key requirement for particle acceleration to ultra-high energies is confinement; particles can remain in the acceleration region as long as their Larmor radius is smaller than the size of the accelerator. Therefore, the maximum energy that can be reached, $E_{\max}$, in a source with a characteristic size $R$ and magnetic field strength $B$, is given by:

$$E_{\max} = \eta^{-1} \beta_{\text{sh}} BR\Gamma,$$

where $\beta_{\text{sh}}$ is the shock velocity in units of the speed of light, $c$, $\eta$ is a parameter that defines the acceleration efficiency, with $\eta = 1$ corresponding to the maximum possible efficiency, and $\Gamma$ represents the Lorentz factor of the motion, which is typically $\Gamma \sim 10-50$ in AGN jets and $\Gamma \sim 10 - 1000$ in GRBs. Fig. 2.14 shows the size and magnetic field strength necessary for possible sources candidates to the acceleration of UHECRs at energies around $10^{20}$ eV. Iron (blue) and proton (red) lines indicate the BR product beyond which confinement of protons and iron nuclei with energy $10^{20}$ eV are possible for outflows with velocity $\beta_{\text{sh}}$. Under this assumptions, several objects could meet the selection criteria for acceleration such as neutron stars and Galaxy clusters.

However, satisfying the confinement condition alone does not guarantee successful acceleration to ultra-high energies. An additional critical requirement for UHECR sources is an adequate energy budget to account for the observed diffuse UHECR flux. This condition is often expressed in terms of the luminosity density $Q$, which must be sufficiently large—estimated to be $\sim 6 \times 10^{44}$ erg Mpc$^{-3}$ yr$^{-1}$. Achieving such a high luminosity density requires a balance between the luminosity emitted by individual sources ($L_{\text{CR}}$) and their number density. This condition can be satisfied either by a large population of weakly emitting sources or by a smaller population of highly luminous

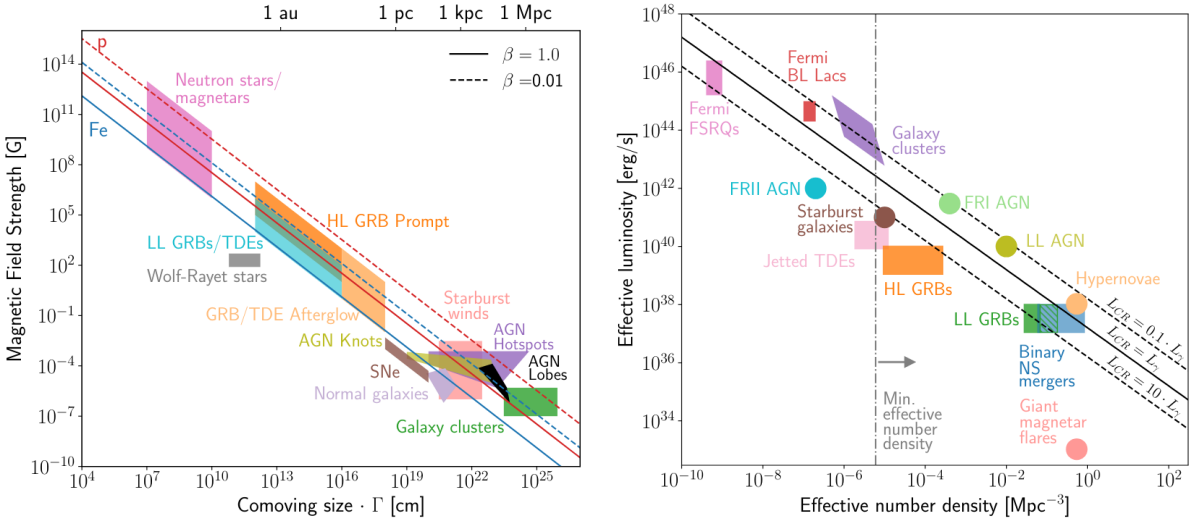



**Figure 2.14:** Characteristic size and magnetic field strength of different source candidate for the acceleration of particle to energies around $10^{20}$ eV. Red and blue lines indicate the acceleration of iron and protons with velocity $\beta_{sh}$(left). The same source candidates are displayed under the constraint related to the luminosity density (right). In both cases, the sources that lie below the line do not meet the criteria [51].

sources, provided the combined contribution is sufficient to account for the observed energy density. In Fig. 2.14 the relation between number density and luminosity is displayed. In this case, since the UHECRs luminosity is not directly measured the $\gamma$-ray luminosity ($L_\gamma$) is used as a proxy. Also in this case, possible source candidates can be determined, such as AGNs and SBGs.

It is important to note that under the Hillias criterion and the limitation related to the maximum energy scales linearly with the charge number indicating a Peter's cycle [52] where a possible source candidate has a max energy $\propto ZE_{max}$ to which a given particle can be accelerated. Moreover, the maximum energy achievable by the source candidates might be a possible factor, together with the propagation effects (GZK effect), to the suppression of the energy spectrum observed at Earth.

Based on the constraints established by the Hillas criterion, several potential candidates emerge as viable sources for cosmic ray acceleration. Among these, an overview of the starburst galaxies will be presented, as they will form the central focus of this study.

## Starburst Galaxies

Starburst galaxies are characterized by intense star formation activity often indicated by infrared luminosities that are more than 10 times higher than those of normal galaxies. In the central region of a starburst galaxy, stellar winds and supernova explosions heat

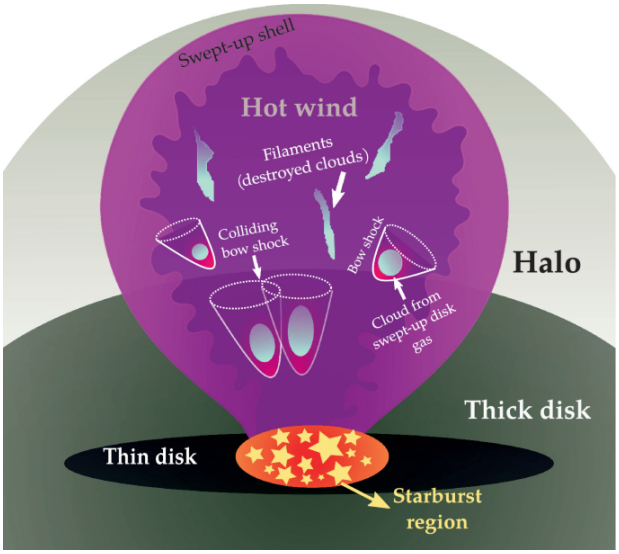



**Figure 2.15:** Diagram of a starburst galaxy: the central region (red) drives a galactic superwind into the surrounding halo, where shock fronts are formed [53].

a cavity to temperatures of approximately $10^8$ K and resulting in powerful, magnetized outflows, which could serve as sites for high-energy particle acceleration [54]. The central region of an SBG, that can be schematically seen in Fig. 2.15, can only accelerate cosmic rays up to energies of $10^{15}$ eV, although the cosmic rays could be further accelerated by shock waves in the galactic winds. Different results have been obtained regarding the maximum energy that can be reached in this scenario ranging between $\sim 10^{18}$ eV up to $\sim 10^{20}$ eV [54]. Another plausible scenario regarding the acceleration of particles in the SBG environment is that UHECR acceleration occurs during the frequent stellar explosions which are related to the high rate of star formation, i.e. in this scenario, SBGs are characterized by a high flux of UHECRs since they are the most likely hosts of candidate sources such as GRBs or young magnetars.

## 2.4. Transport of cosmic rays

The astrophysical sources capable of accelerating particles to energies as high as $\sim 10^{20}$ eV are typically located at distances of several Mpc or more from Earth. As a result, the propagation of UHECRs from their possible sources to Earth is a complex process influenced by several factors, including the expansion of the universe, energy loss mechanisms, and interactions with background photon fields. These processes determine the observed UHECR spectrum, composition, and arrival directions, providing important information concerning their origins.

### Expansion of the Universe

The propagation of UHECRs occurs over cosmological distances, making the expansion of the universe a significant factor. The comoving distance traveled by a UHECR is affected by the Hubble parameter $H(z)$, which describes the rate of expansion at a given redshift $z$. The relationship between the velocity of recession $v$, the Hubble constant $H_0$, and the distance $d$ to a galaxy at $z = 0$ is given by:

$$v = H_0 \times d, \tag{2.13}$$

where $H_0$ is the present-day value of the Hubble parameter. At higher redshifts, $H(z)$ evolves as:

$$H(z) = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}, \tag{2.14}$$

where $\Omega_m$ ($\sim 0.315$) and $\Omega_\Lambda$($\sim 0.685$) represent the matter density and dark energy density parameters, respectively [55]. This evolution affects the travel time and energy losses of UHECRs as they propagate through the expanding universe.

The energy $E$ of a cosmic ray decreases over time due to the redshift effect, which can be described as:

$$E(t) = E_0\frac{1}{1+z}, \tag{2.15}$$

where $E_0$ is the initial energy of the cosmic ray at the source, and $z$ is the redshift of the source. Adiabatic expansion of the universe leads to energy loss as UHECRs' wavelengths stretch in proportion to the cosmic scale factor. This effect is directly tied to the redshift and contributes to the overall decrease in energy following:

$$\lambda^{-1} = -\frac{H(z)}{c} \tag{2.16}$$

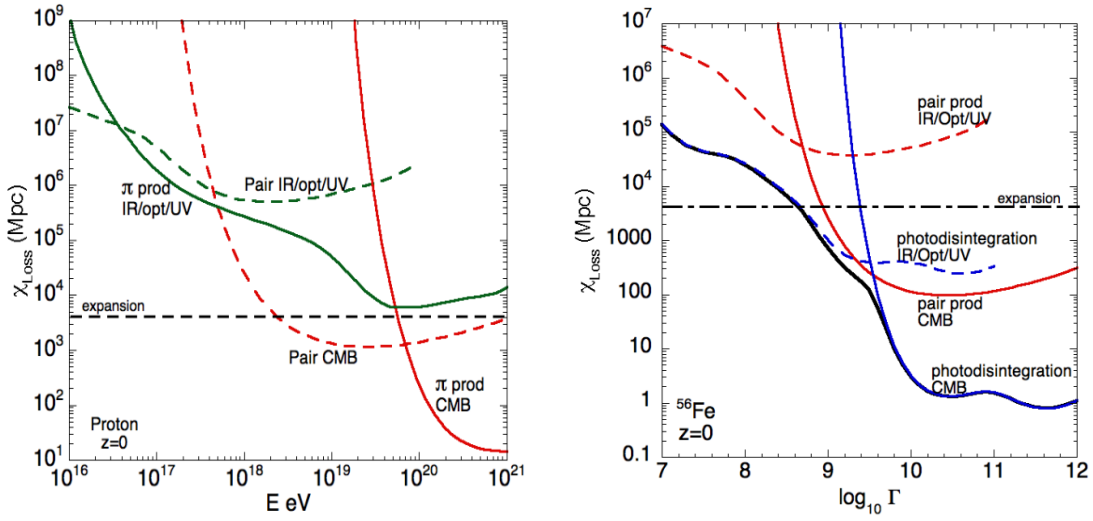



**Figure 2.16:** Evolution of the attenuation length for protons (left) and iron (right) as a function of the energy. Different energy loss processes (adiabatic expansion, pair production and pion production) are considered and displayed together with the photon background [56].

In Fig. 2.16 the energy loss length related to the adiabatic expansion is shown as a dashed dark line for both protons and iron nuclei. Considering the energy range between $10^{17}$ and $10^{22}$ eV, the effects of the energy loss related to the adiabatic expansion are relevant for low energies while higher energies are characterized by other interaction with photon fields such as the Cosmic Microwave Background (CMB) and the Extragalactic background light (EBL).

### 2.4.1. INTERACTION WITH BACKGROUND PHOTON FIELDS

UHECRs traveling in the intergalatic medium can interact with background photons of many different origins covering a wide spectrum. For the energies considered in this thesis, the attention is focused on interaction with the EBL and CMB.

#### EXTRAGALACTIC BACKGROUND LIGHT

The EBL is a diffuse radiation field that spans a wide range of wavelengths, from ultraviolet (UV) to infrared (IR). This radiation is primarily composed of the accumulated light from all stars and galaxies, as well as contributions from dust-reprocessed starlight and AGN. The EBL spectrum is characterized by two distinct peaks:

1. The Optical-UV Peak: Dominated by direct starlight, this peak corresponds to the wavelengths where stars emit most of their radiation. The spectral energy distribution (SED) is shaped by the stellar population and star formation history of galaxies.

2. The Infrared Peak: Caused by dust re-emission, this peak occurs at longer wavelengths and carries information about the amount of dust and the efficiency of dust reprocessing in galaxies.

The relative intensity of these peaks varies depending on the type of galaxy population and their evolutionary state.

Despite its astrophysical importance, the EBL is notoriously difficult to measure due to several observational and theoretical challenges. One of the main issue is the contamination from zodiacal light, which is sunlight scattered by interplanetary dust within our solar system. Zodiacal light is bright and spatially variable, often exceeding the faint EBL signal, making it challenging to separate the two components with precision. This necessitates the use of detailed models of the interplanetary dust distribution, but uncertainties in these models propagate into EBL estimates. Given these difficulties, several theoretical models have been proposed to estimate the EBL spectrum [57–59]. These models, based on galaxy evolution and star formation histories, exhibit significant variability due to differing assumptions about cosmic star formation and dust content. Moreover, the evolution of the EBL with redshift adds another layer of complexity. Since the EBL is the cumulative light from sources over cosmic time, its intensity reflects not only the current galaxy population but also their past contributions. Accurately modeling this evolution requires a detailed understanding of how galaxies and their stellar populations evolve, which remains an active area of research.

### Cosmic microwave background

Alongside the EBL, UHECRs can interact also with the CMB which is a pervasive photon field with a blackbody spectrum peaking at $\sim 1\,\mathrm{meV}$ ($T \sim 2.7\,\mathrm{K}$). The CMB photons are the dominant source of interactions for protons with energies exceeding 2 EeV through pair production and photopion production processes. In Fig. 2.16, the energy loss length for pair production (red dashed line) and photo-pion production (solid red line) decreases significantly above this energy threshold. This behavior reflects the Greisen–Zatsepin–Kuzmin (GZK) effect [11, 12], where UHECRs lose energy due to interactions with CMB photons, producing pions and other secondary particles. The

interaction length for pair production, which dominates at lower energies, is also shown in Fig. 2.16. This process occurs as UHECR protons interact with CMB photons to produce electron- positron pairs, leading to gradual energy losses during propagation. These interactions play an important role in the observed spectrum of UHECRs and in limiting the maximum distance from which high-energy cosmic rays can reach Earth.

The interaction processes of UHECRs with these photon fields are:

- Pair production: High-energy protons interact with low-energy photons in the CMB or EBL, producing electron-positron pairs:

$$p + \gamma \rightarrow p + e^+ + e^-. \tag{2.17}$$

  This process is very efficient for protons, and nuclei, with energies above $\sim 10^{18}$ eV, resulting in gradual energy losses.

- Photopion production: Protons with energies above the GZK threshold ($\sim 5 \times 10^{19}$ eV) interact with CMB photons, producing pions:

$$p + \gamma \rightarrow \Delta^+ \rightarrow p + \pi^0 \quad \text{or} \quad n + \pi^+. \tag{2.18}$$

  This process causes a steep suppression in the UHECR spectrum at the highest energies, known as the GZK cutoff.

- Photodisintegration: High-energy nuclei interact with CMB or EBL photons, causing them to lose nucleons (protons or neutrons). This interaction can be expressed as:

$$A + \gamma \rightarrow (A - 1) + N, \tag{2.19}$$

  where A is the mass number of the nucleus, and N is a nucleon. This process is particularly relevant since it reduces the energy and mass of the nuclei during propagation, effectively altering the observed energy and composition of UHECRs at Earth.

The combination of these factors shapes the observed spectrum and composition of UHECRs. The GZK suppression, for instance, imposes a limit on the distance from which the highest-energy UHECRs can reach Earth ($\sim 100$ Mpc). This horizon is complemented by the effects of energy losses and interactions with background photon fields, which also influence the secondary particles, such as neutrinos and gamma rays, produced during propagation.

### 2.4.2. Cosmic magnetic fields

During their journey to Earth, cosmic rays encounter magnetic fields that significantly influence their propagation. Two key contributors to these magnetic fields are the Galactic Magnetic Field (GMF) and the Extragalactic Magnetic Field (EGMF). These fields alter cosmic ray trajectories through deflections, which depend on the charge, energy, and magnetic field structure. The extent and nature of these deflections are highly model-dependent, with variations in magnetic field models leading to different interpretations of cosmic ray origins and propagation.

In the following sections, an overview of the current understanding of the GMF and EGMF is presented, highlighting their roles in cosmic ray propagation and their impact on the observed properties of cosmic rays.

#### Galactic magnetic field

After being accelerated by astrophysical sources capable of meeting the Hillas criterion, cosmic rays propagate through intergalactic space and eventually enter the Galaxy. Within the Milky Way, the GMF becomes the dominant factor influencing the trajectories of charged particles. This magnetic field is highly structured, with variations in strength and direction across the Galaxy. It is strongest in dense regions, particularly along the spiral arms associated with star-forming activity. To measure the GMF a combination of observational data is used including [55, 60–64]:

- Synchrotron radiation: emission from relativistic electrons spiraling in the magnetic field provides insights into the field's strength and direction perpendicular to the line of sight.
- Faraday rotation measures: the rotation of the plane of polarization of radio waves due to the magnetic field offers information on the field strength and orientation parallel to the line of sight.
- Dust polarization: dust grains aligned with the magnetic field emit polarized thermal radiation in the far-infrared wavelengths. Observations of polarized emission from dust offer valuable insights into the magnetic field's orientation and intensity, particularly in dense regions of the Galaxy.

Despite advances in observational techniques, uncertainties remain in modeling the GMF due to the difficulty of inferring the three-dimensional structure from the two

dimensional data available. The complexity of the field alongside the limitations in data coverage and the position of the Earth in the Galaxy induce large uncertainties which lead to varying model predictions, influencing the understanding of the propagation of charged particles. The GMF is generally considered to consist of the following components [65, 66]:

- Large-scale regular field (coherent field): the coherent component represents the large-scale, organized magnetic field that follows the spiral arms of the Galactic disk and large-scale poloidal and toroidal components. This component is primarily responsible for the significant, systematic deflections of charged particles as they traverse the Galaxy, bending their trajectories in predictable ways over large distances.

- Turbulent field (isotropic random): the turbulent field is characterized by rapid variations over short distances and introduces small-angle deflections to the paths of charged particles, leading to diffusion. These random scatterings cause the particles to experience deviations from their otherwise smooth trajectories, adding complexity to the overall behavior of particles deflected inside the Galaxy.

- Striated field (ordered random): the striated field can be seen as an intermediate state between fully turbulent and fully ordered. It might introduce anisotropies in the propagation of charged particles, as it can focus or scatter particles along specific directions [67]. It can enhance or suppress deflections in certain regions, depending on the alignment of the striated magnetic structures with the incoming charged particle.

A visualization of the three components of the GMF can be seen in Fig. 2.17 where the differences on the observables are displayed.

Charged particles propagation in the GMF are influenced by the particle's rigidity, defined as $R = \frac{E}{Z}$, where $E$ is the energy and $Z$ is the charge of the particle. Particles with lower rigidity experience larger deflections, complicating the task of tracing their origins and reconstructing their trajectories.

One of the most popular models used to quantify GMF deflections is the `JF12` model proposed by Jansson & Farrar [65,66]. A top view of its component can be seen in Fig. 2.18 where on the left panel the disk model for the regular component is shown while on the right panel a superposition of the regular and random component is shown. Inside the radius of 20 kpc, it can be seen that the strength of the field is not constant along

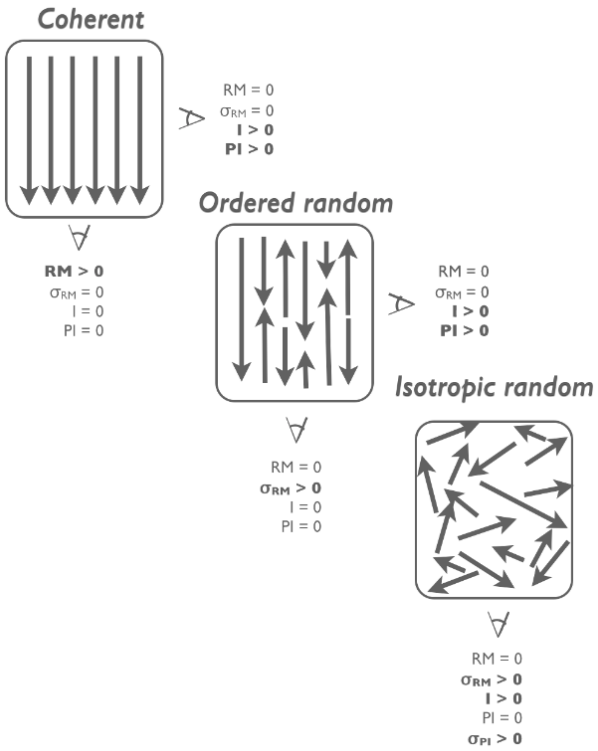



**Figure 2.17:** Illustration of the magnetic field components. The three observables for these magnetic fields are total intensity I, polarized intensity PI and rotation measure RM [67].

the spiral arm and differ between every spiral arms from 2-4 µG. When the turbulent component is introduced major differences can be observed especially in the sudden direction changes together with the general structure of the strength along the spiral arms. Newer models for the GMF field have been proposed by Unger & Farrar [68] in which they incorporated the improved catalogs with higher resolution for the Faraday rotation measurements alongside the updated sky surveys [55] for synchrotron emission providing enhanced detail for modeling the small-scale structures. In addition, they studied different field parametrization and axillary models of thermal and cosmic ray electron models, resulting in eight variations called `base`, `cre10`, `expX`, `nebCor`, `neCL`, `spur`, `synCG` and `twistX`. In Fig. 2.19, a top-view of the coherent component of the `neCL` model is compared to the `JF12` model. In the case of the `neCL` model several differences can be seen in the amplitude and structure of the spiral arms due to the improvement in the data resolution together with the effect of using a fixed-arm spiral disk (`JF12` model) compared to a Fourier spiral disk field (`neCL` model) which results in a smooth description of the field without sudden discontinuities.

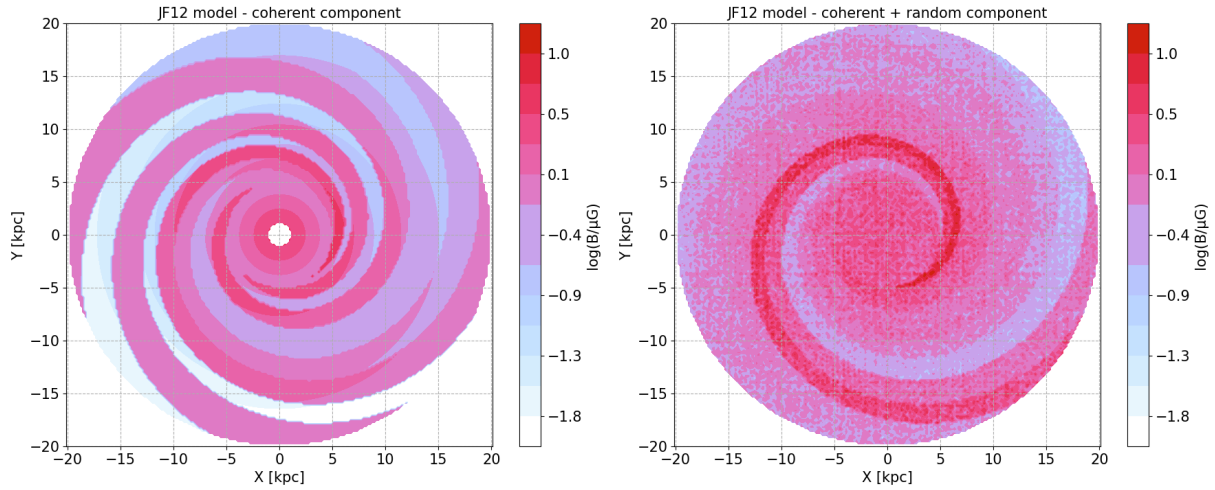



**Figure 2.18:** Top view of the coherent component of the `JF12` model (left) alongside the top view of the turbulent component (right). The color scale indicates the strength of the field when considering a coherent length of 60 pc.

## Extragalactic magnetic field

The Extragalactic Magnetic Field (EGMF) is the magnetic field present in the extragalactic medium, which permeates the space between galaxies and galaxy clusters. Unlike the GMF, which is confined to the Milky Way, the EGMF affects cosmic ray propagation over cosmological distances, playing a crucial role in the trajectory and energy of UHECRs.

The EGMF is thought to originate from several astrophysical processes, including the dynamo action in large-scale structure formation, shocks in galaxy clusters, and the interaction of high-energy particles with surrounding matter. It may also have originated from a primordial magnetic field created in the early Universe during events like inflation or phase transitions, which was later amplified by astrophysical processes to produce the observed fields. While its exact strength and configuration remain uncertain, models predict that the EGMF is a heterogeneous field, with regions of stronger magnetic fields found in galaxy clusters and voids, and weaker fields in the more diffuse regions of the intergalactic medium.

Measuring the EGMF is highly challenging due to its extremely low intensity and the difficulty of isolating its effects from other astrophysical processes. Nevertheless, several observational and theoretical constraints on the EGMF have been proposed and are summarized in Fig. 2.20. These constraints are derived using various methods, including CMB anisotropies, $\gamma$-ray propagation, and UHECR observations, as well as theoretical predictions.

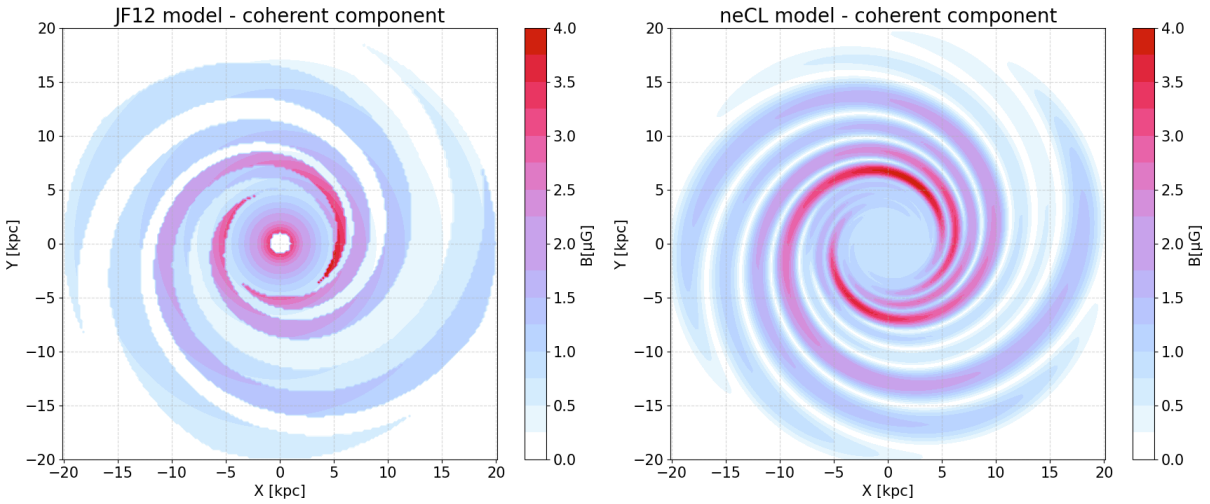



**Figure 2.19:** Top view of the coherent component of the `JF12` model (left) alongside the top view of the coherent field in the `neCL` model (right). The color scale indicates the the strength of the field when considering a coherent length of 60 pc.

Constraints on the EGMF are derived from various observational and theoretical considerations, each providing insights into its possible strength and coherence scale. Observations of the CMB power spectrum impose an upper limit on the primordial EGMF, which would have originated in the early Universe before recombination. Such fields could perturb the CMB anisotropies, leaving detectable traces. Recent analyses, including those from Fermi-LAT and CMB experiments, have constrained the EGMF strength to below $B < 2.8\,\mathrm{nG}$. This limit, represented by the blue dashed line in Fig. 2.20, reflects the boundary where magnetic field effects on large coherence lengths become measurable through CMB power spectra. Gamma-ray observations provide complementary constraints. High-energy $\gamma$-rays from astrophysical sources like blazars interact with the EBL, producing electron-positron pairs. These pairs are subsequently deflected by the EGMF, resulting in time delays and angular broadening of the observed $\gamma$-ray profiles. Analysis of these effects suggests a lower limit on the EGMF strength of $B > 3 \times 10^{-7}\,\mathrm{nG}$ to allow sufficient deflections. This limit, depicted as the pink dashed line in Fig. 2.20, demonstrates the utility of $\gamma$-ray propagation as a probe for weak intergalactic magnetic fields. UHECRs observations also can be employed to propose constraints on the EGMF. The EGMF influences UHECR propagation, affecting their trajectories and causing anisotropies in their observed arrival directions. Constraints on the EGMF strength are linked to specific source populations, such as starburst galaxies (SBGs) and active galactic nuclei (AGNs), under the assumption that these populations

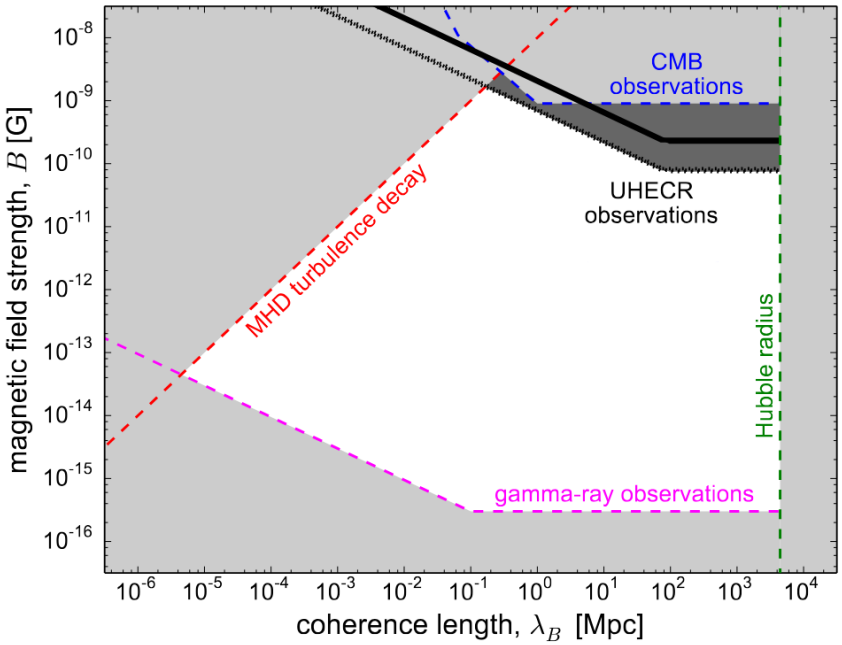



**Figure 2.20:** Limits on the EGMF strength as a function of the coherence length [69]. The shaded region represents excluded parameter space based on different observational and theoretical constraints.

contribute to observed UHECR overdensities. The resulting constraints are represented by the black solid line in Fig. 2.20 [69]. These UHECR-based limits are subject of this thesis and will be discussed in detail in Chap. 5.

Theoretical considerations further constrain the EGMF. Turbulence decay models suggest that primordial magnetic fields generated in the early Universe decay over time, imposing an upper limit on the field strength as a function of coherence length. This constraint is illustrated by the red dashed line in Fig. 2.20. Additionally, the Hubble limit suggests that magnetic fields cannot exhibit coherence lengths larger than the observable Universe's size ($\sim 10^4$ Mpc), as shown by the green dashed line in the same figure. Together, these observational and theoretical constraints offer a comprehensive framework for understanding the EGMF, delineating its strength and coherence scale while highlighting the interplay between different astrophysical processes and cosmic magnetic fields.

The strength of the EGMF remains an open question, with significant variability in theoretical and observational predictions regarding its magnitude [70]. Despite these uncertainties, the EGMF is widely believed to influence the deflection of charged particles as they propagate from astrophysical sources to Earth. Given the complexity of the

deflection processes, these effects are often simplified using a random walk approximation.

In this framework, UHECRs are considered to propagate in the non-resonant regime, which occurs when the particle energy, $E$, is above the critical energy $E_c = ZeBl_c$. The critical energy delineates two distinct regimes:

- the resonant diffusion regime, where lower-energy particles experience substantial deflections comparable to the Larmor radius $r_L$ which is smaller than the coherence length
- the non-resonant regime, where higher-energy particles undergo smaller deflections. In the latter case, after traversing a coherence length $l_c$, the deflection angle is given by $\theta = l_c/r_L$

When considering the non-resonant regime, the charged particles with high energy undergo numerous small-angle scatterings due to interactions with the EGMF. These cumulative scatterings result in a broadening of the particle beam, manifesting as an angular spread of arrival directions around their astrophysical sources. The root-mean-squared angular deflection can be mathematically described following [6,71]:

$$\theta_{\rm rms} = 25^\circ Z \sqrt{\frac{l_c}{1\,\mathrm{Mpc}}} \sqrt{\frac{d}{1\,\mathrm{Mpc}}} \frac{B}{1\,\mathrm{nG}} \frac{1\,\mathrm{EeV}}{E} \tag{2.20}$$

This expression relates the beam-widening, or spread, to parameters like the root-mean-square field strength $B$ of the EGMF, the coherence length $l_c$ , the source distance $d$, and the rigidity $R = E/Z$ of the charged particle. Additionally, the beam-widening depends on the square root of the distance to the source, meaning that sources farther from Earth contribute to a larger angular spread of arrival directions (similar estimation can be found in [72]).

# Chapter III

# Simulation of UHECR propagation and deflections in magnetic fields

The main focus of this chapter is to highlight the effects of the GMF on charged particles and to incorporate these effects into the likelihood analysis with the intention of quantifying the influence of magnetic fields on source identification. The workflow used to include the effects of magnetic fields, which underlies the structure of the following chapters, is presented in Fig. 3.1. First, a spectrum of nuclei is injected based on a specific source catalog. These particles are propagated to the edge of the Galaxy, where their arrival directions are modified due to deflections by the GMF, which incorporates source emission, flux weights, and GMF deflection effects. Then, the database of simulated events is combined with an isotropic background, computed by employing events isotropically distributed that follow the observed spectrum, with varying signal fractions to account for contributions from the source catalog and isotropic events. Realizations are then generated from this combined dataset, which are used as inputs for the likelihood analysis. A comprehensive overview of the simulation setup employed to construct simulated datasets is provided, along with an examination of the effects of the GMF on UHECR propagation. This includes a focus on the deflections that modify the observed arrival directions and their implications for interpreting cluster-like structures in the sky. The simulated datasets, which include a combination of signal, where magnetic field effects are introduced, and isotropic background, are utilized in a likelihood analysis that will be detailed in Chap. 4.

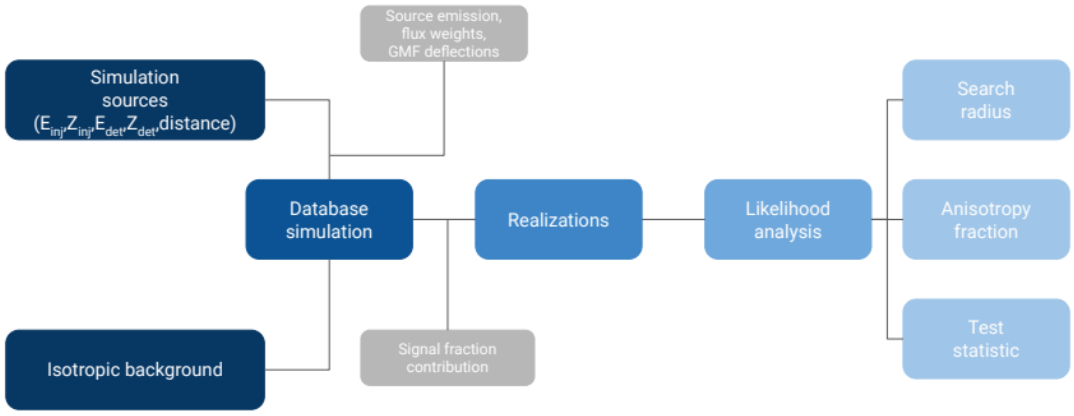


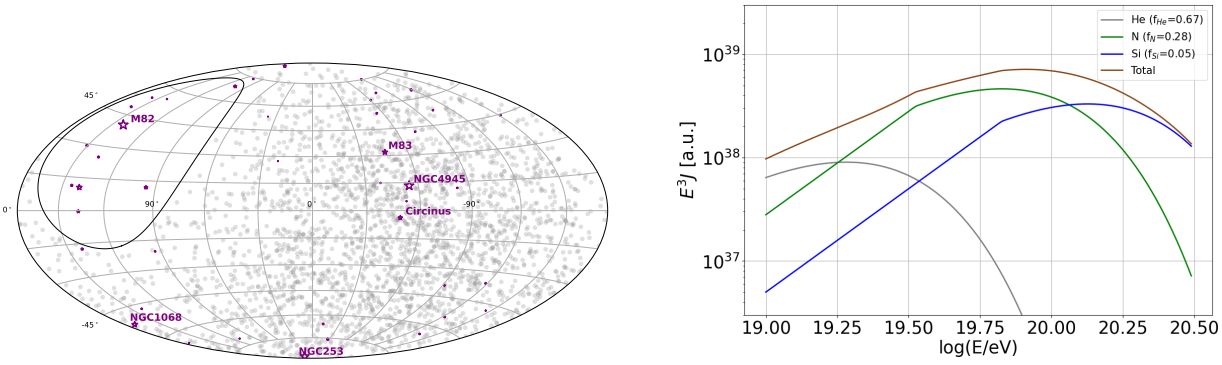

**Figure 3.1:** Work flow of the analysis that highlights the main components of the analysis and structure of the following chapters.

## 3.1. SIMULATIONS OF THE ASTROPHYSICAL SCENARIO

The work conducted in this thesis aims to provide a phenomenological description of the data published by the Pierre Auger Collaboration above 32 EeV, covering the period from 2004 to 2020. The public data has been used to model the simulations, alongside the results reported in [21]. In Fig. 3.2, the public dataset is presented as a scatter plot in Galactic coordinates, where the white area bounded by a black curve is outside the field of view of the experiment. The sources displayed in Fig. 3.2 as stars, with sizes proportional to their flux weight reported in [21], correspond to the Starburst Galaxies (SBGs) catalog [23]. This catalog has shown the highest deviation from isotropy in the likelihood analysis conducted by the Collaboration and has been employed to model the signal contribution in the simulations.

### 3.1.1. MODELING COSMIC RAY EMISSION

As discussed in Sec. 2.1.2 and shown in Fig. 2.7, the highest significance has been observed at 38 EeV, which it will be used as the lower limit of the simulations throughout this work. The signal contribution has been modeled by assuming that the emitted flux at the source $J_{\rm inj}$ as a function of the injected energy $E_{\rm inj}$ and the mass number $A_{\rm inj}$ of each representative element is a power-law with a cutoff function $f_{\rm cut}$, which is applied when the maximum energy $E_{\rm cut} = Z_A R_{\rm cut}$ is reached:




**(a)** **(b)**

**Figure 3.2:** (a): Events detected by the Pierre Auger Observatory from 2004 to 2020. The dataset includes vertical and inclined events with an energy ranging from 32 EeV up to 165 EeV. Together with the events the SBG catalog is shown with purple stars. (b): Injected energy spectrum at the source. The contributions of H (red curve), He (grey curve), N (green curve) and Si (blue curve) are obtained using equation 3.1 and 3.2 together with the best-fit parameters presented in [73].

$$J_{\text{inj}}(E_{\text{inj}}, A_{\text{inj}}) = J_0 \cdot a_A \cdot \left(\frac{E_{\text{inj}}}{10^{18}\,\text{eV}}\right)^{-\gamma} f_{\text{cut}}\left(\frac{E_{\text{inj}}}{Z_A R_{\text{cut}}}\right) \tag{3.1}$$

The broken exponential cutoff function $f_{\text{cut}}$ is defined as in [73]:

$$f_{\text{cut}}\left(\frac{E_{\text{inj}}}{Z_A R_{\text{cut}}}\right) = \begin{cases} 1 & Z_A R_{\text{cut}} > E_{\text{inj}} \\ \exp\left(1 - \frac{E_{\text{inj}}}{Z_A R_{\text{cut}}}\right) & Z_A R_{\text{cut}} \leq E_{\text{inj}} \end{cases} \tag{3.2}$$

Here, $\gamma$ is the spectral index, and $R_{\text{cut}}$ is the maximum rigidity of the source. $J_0$ is the flux normalization factor. $Z_A$ is the charge number of the injected species with mass number $A_{\text{inj}}$, and $a_A = a(A_{\text{inj}})$ is the fraction of particles of that species. In Fig. 3.2b, the injected spectrum is shown where the elements employed in the analysis are displayed with a solid colored line obtained using a spectral index $\gamma = -1$. The fractions employed and the rigidity cut-off ($\log_{10}(R_{\text{cut}}/\text{V})$=18.68) are based on the results presented in [73].

### 3.1.2. MODELING COSMIC RAY INTERACTION

To model the propagation processes that a charged particle undergoes as it travels from its origin to Earth, simulations were conducted using `CRPropa3` [74]. To describe the various interactions of protons and nuclei with the EBL (see Sec. 2.4.1), the Gilmore model [57] is employed, which provides an accurate representation of the photon density

at different redshifts.

Previous studies, such as [73,75], have demonstrated that the differences in interaction models can lead to varying predictions for cosmic ray propagation. These studies provide the basis for understanding the uncertainties introduced by different theoretical frameworks. Beyond the interactions with the EBL, UHECRs also interact with the CMB. The CMB photons are the dominant source of interactions for protons with energies above 2 EeV. The energy loss length for pair production and photo-pion production decreases significantly above this energy threshold as shown in Fig. 2.16. This behavior reflects the Greisen–Zatsepin–Kuzmin (GZK) effect, where UHECRs lose energy due to interactions with CMB photons, producing pions and other secondary particles. These interactions play an important role in the observed spectrum of UHECRs and in limiting the maximum distance from which high-energy cosmic rays can reach Earth.

For the cosmological parameters and background fields, the standard assumptions from `CRPropa3` are used, which include a flat universe model.

### Simulation Details

The propagation simulations are carried out for a wide range of parameters, including multiple distances, injected energies, and mass numbers. Specifically:

- **Injected Energy**: UHECRs are simulated with initial energies that range between $10^{18}$ eV and $10^{21}$ eV. The injected energies follow a power-law spectrum proportional to $E^{-1}$,as implied by the best-fit spectral index reported in [73].
- **Mass Numbers**: The simulations are performed for four representative elements based on the elements listed in [73]. These elements range from protons to heavier nuclei such as Silicon, providing a broad representation of the potential composition of UHECRs.

For each combination of injected energy, mass number, and source in the SBG catalog [23], a total of $10^6$ particles are injected. This results in a substantial number of injected particles per source and per element considered. Such a large sample size ensures that statistical fluctuations are minimized. Furthermore, the construction of numerous datasets, combined with the large number of particles simulated, ensures that the resulting datasets are not heavily correlated.

After including propagation effects and considerations of energy and mass composition, the observed spectrum detected at Earth, for the signal contribution, is presented in Fig. 3.3.

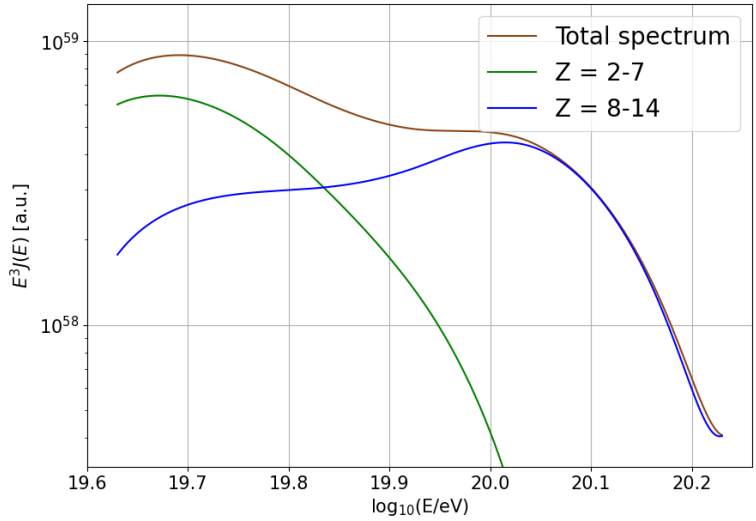



**Figure 3.3:** Observed spectrum from SBG sources obtained using the best-fit parameters listed in [73], evaluated at the top of the Earth's atmosphere in arbitrary units. The solid lines indicate contributions from different elements, while the brown line represents the total spectrum.

The section of the spectrum of major interest for this thesis corresponds to energies above 38 EeV. In this range, the Helium component becomes negligible, and the spectrum is primarily composed of heavy elements such as Nitrogen and Silicon. These heavy elements dominate the spectrum up to energies around 120 EeV, where Silicon particles become the sole significant contributors.

This distinction is important since heavy elements experience substantial deflections under the influence of the GMF at these energy levels. The following section analyzes the propagation of high-energy charged particles through the GMF, focusing the attention on its impact on their arrival directions.

## 3.2. TRANSPORT OF UHECR IN MAGNETIC FIELDS

An important aspect that characterizes the propagation of charged particles is their interaction with magnetic fields. During their propagation towards Earth, cosmic rays interact with magnetic fields such as the GMF, which influences their trajectories by inducing additional deflections. These deflections are highly model-dependent, and changes in the GMF model can lead to different results and interpretations.

The GMF models discussed in this thesis consists of three main components, as outlined in Sec. 2.4.2. The first component is the coherent component, which accounts for large-scale deflections occurring over long distances. The second is the turbulent component, characterized by significant variations over short distances. Lastly, the third component is the striated field, which represents an intermediate state between fully ordered and fully turbulent conditions.

Among the three main components of the GMF, the effects of varying the coherent and turbulent components are studied by constructing different configurations of the turbulent field and examining how changes to the coherent component impact the resulting deflection patterns. This approach provides a more comprehensive understanding of how each component influences cosmic ray propagation through the Galaxy, while also allowing for an estimation of the Galactic variance when employing different configurations within the same model.

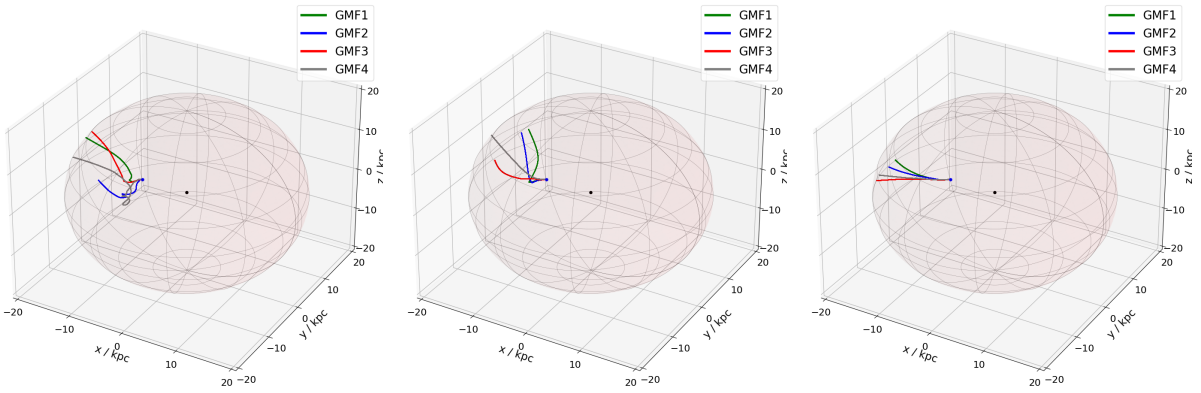
### 3.2.1. DEFLECTIONS IN THE TURBULENT COMPONENT OF THE GALACTIC MAGNETIC FIELD

As mentioned previously, the turbulent component induces random deflections of charged particles over short distances. These random deflections can be analyzed within the framework of the same general GMF model to quantify the effects of varying this component [76].

A few examples of the effects of this variation are shwon in Fig. 3.4. The figures (panels 3.4a, 3.4b, 3.4c) highlight the backtracking (see Sec. 3.2.5) of charged particles through the `JF12` model, which consists of a coherent field combined with different realizations of the turbulent field, each characterized by a distinct seed. The different GMF configurations are represented by colored solid lines. Backtracking was performed for a single direction in the sky, with particle rigidities of 5, 10, and 20 EV, to illustrate how deflection patterns change with increasing energy.

From left to right, it is evident that the differences between the realizations of the same GMF model are substantial at low energies, where charged particles experience large deflections across the sky. At these energies, the arrival directions of cosmic rays are less correlated with their actual sources. However, at 20 EV, the differences between the four realizations of the turbulent field reduce to just a few degrees[1]. This demonstrates the decreasing influence of the turbulent field at higher energies, where cosmic rays

[1]A Nitrogen particle with the given rigidities corresponds to charged particles with approximate energies of ~35, 70, and 140 EeV, respectively.




(a) (b) (c)

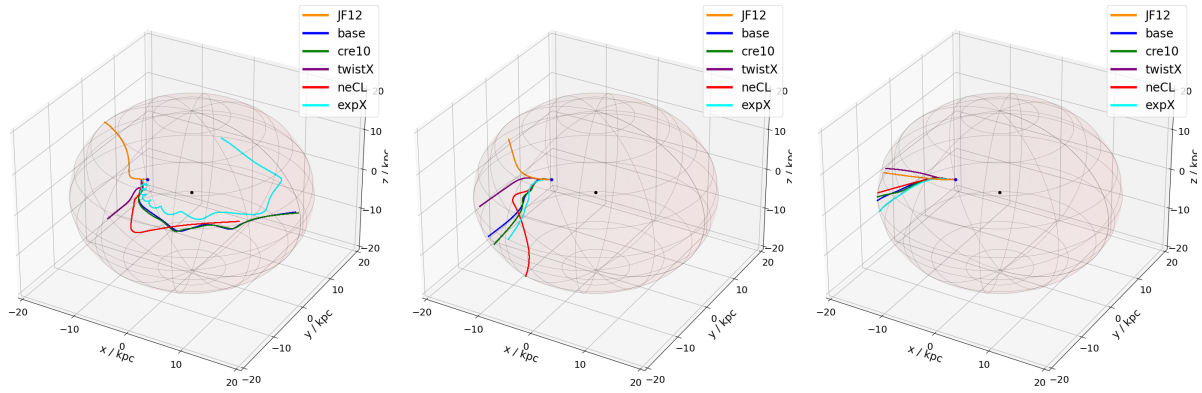
**Figure 3.4:** Back-tracked particle trajectories when different realizations of the turbulent component of the `JF12` model are used. The trajectories related to the different realizations are displayed with a solid colored line. The backtracked particles have rigidities 5 EV (a), 10 EV (b), and 20 EV (c).

are less deflected and their trajectories more closely align with their original directions. The behavior shown in the three rigidity bands suggests that the differences between realizations are substantial only at lower energies but negligible at higher energies, suggesting that the Galactic variance should play a minor role when the variation only affects the turbulent component.

### 3.2.2. Deflections in the coherent component of the Galactic magnetic field

In addition to the effects introduced by variations in the turbulent field, variations in the coherent component of the GMF also play a crucial role in determining the propagation of charged particles. The coherent component, which is responsible for large-scale, systematic deflections, is influenced by the choice of the GMF model, as different models incorporate varying assumptions and datasets for their construction.

Panels 3.5a, 3.5b, and 3.5c illustrate the coherent field's effects on particle trajectories for the same rigidity bands (5, 10, and 20 EV) employed in the turbulent field study. In this example four distinct GMF models are compared: the widely used `JF12` model and three newer models proposed in [68]. These newer models differ significantly in their modeling approaches and the data used for their construction, providing a broader representation of the possible outcomes of charged particle deflections in the presence of varying coherent field configurations. To highlight potential differences, the models




**(a)** **(b)** **(c)**

**Figure 3.5:** Back-tracked particle trajectories when different coherent fields are used. The coherent fields considered are the `JF12` model alongside the newer models presented in ( [68]). The rigidities are 5 EV (a), 10 EV (b), and 20 EV (c).

chosen are the `twistX` model, which shares similar features with `JF12` regarding overall flux magnification with the `base` and `cre10` models that exhibit substantial variations in flux magnification. These four models were chosen as they give the largest variety in the flux maps (see Sec. 3.2.3) At low rigidity the deflections in the coherent fields appear very different: all the models largely differ from `JF12` while `cre10` and `base` appear to have a similar behavior. The differences between the models are reduced when higher rigidities are considered: above 20EV the backtracked position is similar between the models and they do not drastically differ from `JF12`. This is expected since particles with high energies are less deflected in the GMF and they directly point closer to their origin.

### 3.2.3. MEAN DEFLECTION ANGLE

Another important piece of information that can be used to understand the differences between the different GMF coherent fields relies on the mean deflection angle. The deflection angle $\theta_{\text{def}}$ is defined as:

$$\theta_{\text{def}} = \arccos\left(\frac{u \cdot v}{|u|\,|v|}\right) \tag{3.3}$$

where $u$ represents the cosmic ray's inferred direction at the Galaxy's edge while $v$ represents its observed arrival direction at Earth. Therefore thanks to the parameter $\theta_{\text{def}}$, it is possible to quantify the actual deviation in the path of a cosmic ray particle

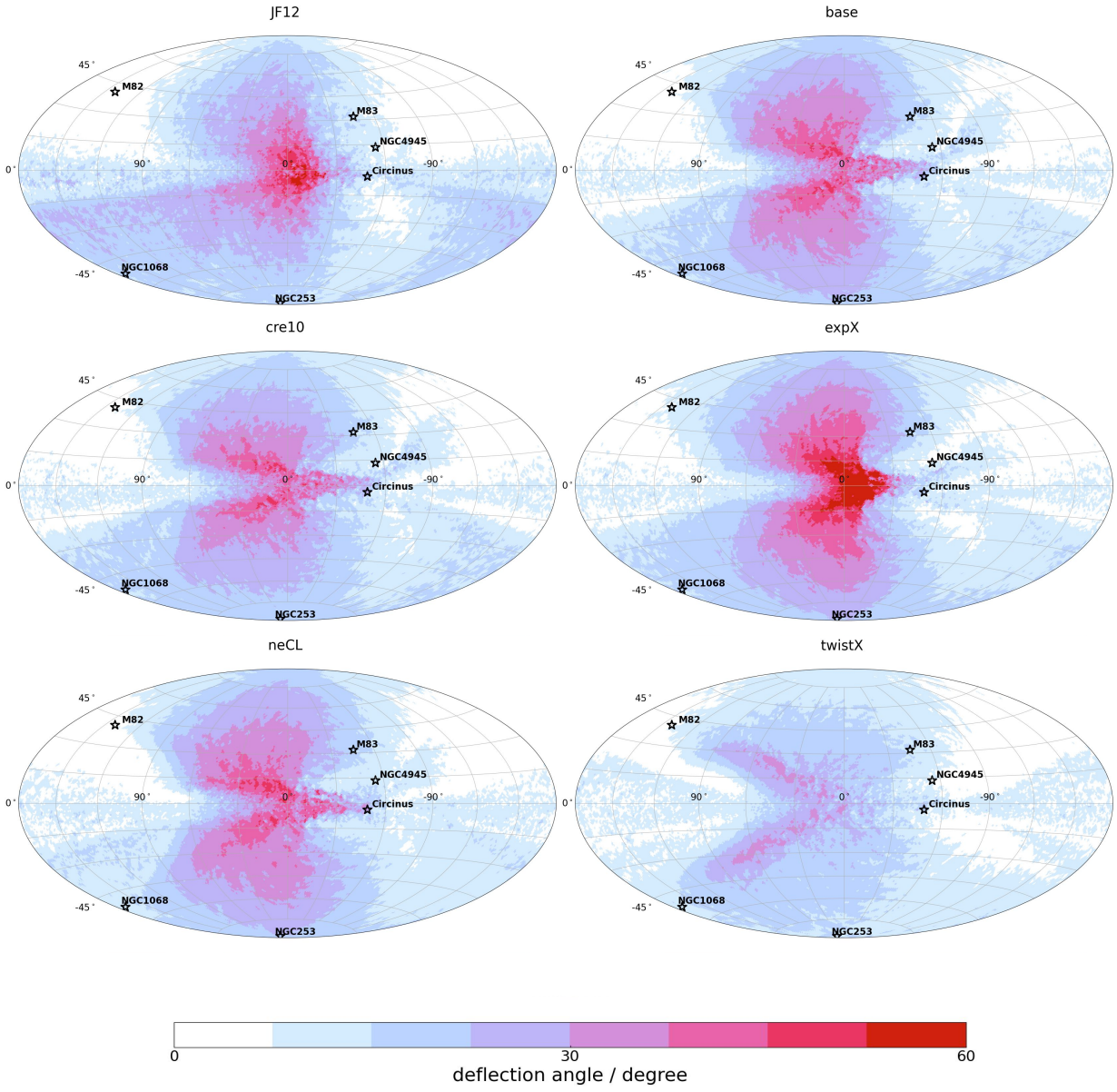



**Figure 3.6:** Comparison of mean deflections under the assumption of different model for the description of the models of the coherent GMF. The `JF12` model is shown in the top panel and it can be compared with the newer models from UF23 [68]. The rigidity is $2 \times 10^{19}$ V.

due to the GMF by comparing the particle's inferred direction at the edge of the Galaxy (before it encounters significant GMF influence) with its arrival direction at Earth, taking into account the bending effect of the GMF along its trajectory. The comparison between several coherent field models previously employed in the backtracking examples above is shown in Fig. 3.6. As can be seen, the models display qualitatively the same deflection pattern but they differ in terms of amplitude and behavior. The main differences between the newer models and `JF12` might be found in the way they are modeled, for example, the `nebCor` model shows larger deflections because the scale of the magnetic field is computed using polarized synchrotron data instead of Faraday's rotation data while the `expX` model has stronger deflections around the Galactic center because it is constructed with a six times larger poloidal field strength than the other models. This information shows how relevant the uncertainties of the GMF might be on the arrival direction of charged particles and also the importance of employing several models to investigate possible differences among the effects in the coherent field.

### 3.2.4. Magnetic fields in previous analyses

In the analysis of arrival direction correlations that initially revealed the overdensities around the catalog sources (see Sec. 2.1.2), the arrival directions are modeled using a circular spreading around the source position. This spreading is described by a von Mises-Fisher distribution [77], $F(x|\mu,\kappa)$, which is the equivalent of a two-dimensional Gaussian on the surface of a sphere. It is given by:

$$F(x|\mu,\kappa) = \frac{\kappa}{4\pi \sinh(\kappa)} e^{\kappa \mu \cdot x} \tag{3.4}$$

where $x$ indicates the directional vector and $\mu$ the mean direction of the distribution. The concentration parameter $\kappa$, in the small angle approximation [78], characterizes the width of the distribution and is related to the root-mean-square (rms) spreading angle $\theta_{\rm rms}$ by $\kappa = 1/\theta_{\rm rms}^2$. In the analysis conducted in [21] a constant smearing angle for all sources was assumed, contrary to the expectation, that it should depend on the square-root of the distance of the source. Moreover, the center of the distribution is not displaced to avoid dependency on the coherent deflections. These assumptions are made with the intention to avoid a dependency from a specific model of the GMF. Although this approximation of the deflections has some limitation since the effects of the GMF are not only represented by a constant blurring. If coherent deflections are taken into account, and if they are greater than the spreading, the image of the

source could be displaced several degrees from its original position. Moreover, rigidity deflection patterns, where high-energy events are located close to the nominal position of the source, while low-energy events are deflected further away from their original location are expected to be observed.

Alongside the contribution from the GMF, also the effects of the EGMF have been incorporated in the 2D Gaussian centered around the source position. One of the positive aspect of approximating the EGMF as a von-Misses Fisher distribution is linked to the fact that the Gaussian distribution is independent from the rigidity which is unknown in data, although this approximation has its limitation as well since the deflections of high energy events largely differ from particle types since the deflections in the EGMF are $\propto 1/R$. A possible approach on simulation that introduces a rigidity dependent spreading is presented in Sec. 3.2.5.

### 3.2.5. Magnetic fields in this thesis

#### GMF in this thesis

In this study, GMF deflections are included in the analysis by utilizing a technique known as lensing, as described in [79]. The lensing method consists of tracing the trajectories of antiprotons in a backward direction, from a given direction in the sky at the Earth out to the boundary of the Galaxy. By modeling the behavior of antiprotons under the influence of the GMF, this technique permits accounting for the deflections that charged particles experience as they propagate through the Galaxy if the rigidity is known.

The initial positions in the sky from which the antiprotons are tracked are determined using a pixelization scheme provided by the `Healpy` package [80]. Specifically, the angular resolution used corresponds to a `nside` parameter of 1024, which leads to a high-resolution map of the celestial sphere with more than $10^6$ pixels. The backtracking process employs a logarithmic step size of 0.02 in $\log(R/V)$, where $R$ represents the rigidity of the particles, spanning a range from $\log(R/V) = 18.00$ to $\log(R/V) = 21.00$. This choice ensures sufficient resolution in rigidity space to accurately capture the influence of magnetic fields on the trajectories of charged particles. Additionally, a coherence length of $l_c$=60 pc is adopted for the backtracking. This parameter is chosen to account for realistic deflections of particles as they propagate through the Galaxy. Together, these choices provide a high-resolution framework for determining the backtracked positions under the influence of the GMF.

Each pixel represents a small region of the sky, and the backtracking procedure

begins from the center of each pixel. This approach ensures that the method accurately captures the potential GMF-induced deflections for any possible arrival direction of cosmic rays. The information obtained throughout the lensing is displayed in Fig. 3.7 and the interpretation of the effects displayed is the following [81]: a bright, highly magnified region on the sky map indicates that cosmic rays arriving from that direction are more likely to reach Earth, suggesting an amplification effect by the GMF. These magnified areas correspond to regions where the deflections caused by the GMF are such that more cosmic rays are directed toward Earth. On the other hand, a dark, demagnified region reflects a reduced chance of detection, implying that there exists few or in extreme cases no path from the source to Earth. If a source lies within a heavily demagnified region, it may be nearly or entirely obscured, meaning that few, if any, cosmic rays originating from that source would be observed on Earth. This scenario underlines the importance of understanding GMF-induced deflections to accurately predict cosmic ray arrival directions.

The models used in this analysis, such as `JF12` and the newer models presented by UF23 [68], exhibit considerable variability in terms of transparency. For example, several models—such as the `base`, `cre10`, `expX`, `nebCor`, `neCL`, `spur`, and `synCG` model—show significant demagnification effects for many primary cosmic ray sources in the SBG catalog. A particularly interesting case is the galaxy NGC253, which is located in a region that appears almost completely opaque in most of these models. The near-total demagnification of this region suggests that cosmic rays originating from NGC253 are highly unlikely to be detected due to the strong deflection and scattering effects imposed by the GMF.

This finding has critical implications for cosmic ray source identification, as it reveals that certain sources, despite being included in source catalogs like the SBG catalog, may contribute very little—if anything at all—to the observed UHECR flux [82]. If the location of these sources is in heavily demagnified regions, this means that their cosmic ray signatures are effectively "hidden" from detection on Earth. Given that the position of a source in the sky plays a crucial role when the deflections caused by the GMF are taken into account, it becomes relevant to quantify the expected contribution of the primary sources in the catalog. The contribution of a single source must consider both the magnification effects induced by the GMF and the exposure of these sources to Earth's detectors.

To determine the contribution of the primary sources to the observed UHECR flux, simulations of the expected event distributions have been constructed using a signal fraction of $f_{\mathrm{sgn}} = 15\%$ (see Chapter 4.1 for detailed information regarding the selection

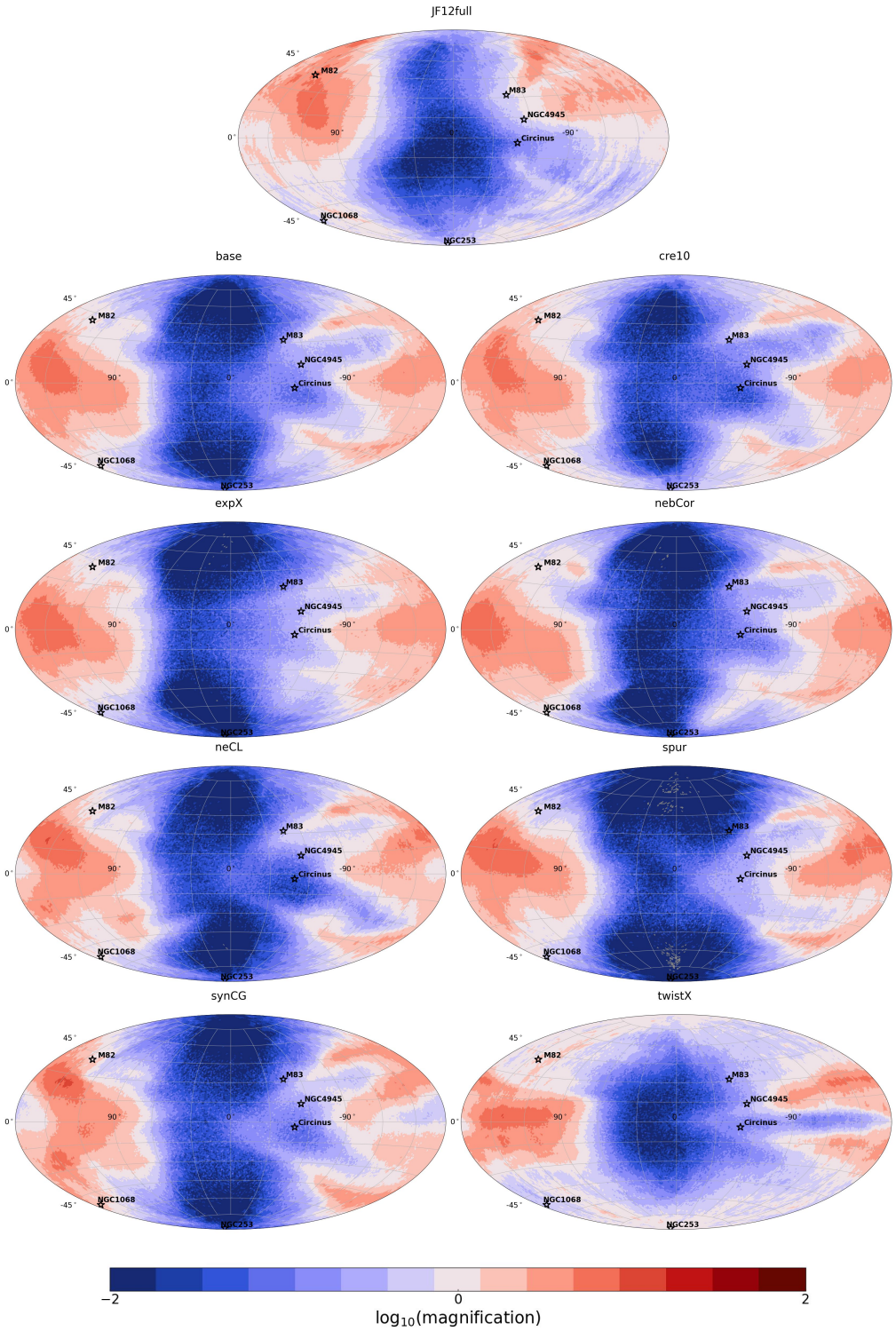



**Figure 3.7:** Magnification effects employing an energy spectrum $\propto E^{-1}$.

of this signal fraction) and a corresponding background contribution of $f_{\mathrm{bkg}} = 1 - f_{\mathrm{sgn}}$.

Each of the 1000 simulations was constructed by combining signal and background events, where the signal represents contributions from astrophysical sources in the catalog, and the background represents isotropic events. For each simulation, the expected contribution from individual sources in the catalog was computed, taking into account the relative flux of the sources. These fluxes were weighted using the values reported in the Lunardini catalog [23].

The results of these calculations are summarized in Fig. 3.8, which displays the relative contribution of each source, weighted by its flux, to the total UHECR flux for comparison. To provide an interpretation of the source contributions in the context of the Pierre Auger Observatory, the exposure of the experiment (i.e., the field of view of the detector) was applied providing an estimation of the expected source contributions observable from Earth.

### Flux (de-)magnification

One of the key sources analyzed is NGC4945, located in the Centaurus region. NGC4945 is a prominent source in the SBG catalog, and it plays an important role in UHECR source studies due to its relative proximity and high contribution to the overall observed flux. When different models of the GMF are applied, the overall contribution of NGC4945 to the flux remains relatively stable, with only minor variations between models. This suggests that the deflections induced by the GMF in this part of the sky do not significantly modify the observed flux of events originating from this source. A possible explanation for this behavior can be found by examining the magnification maps of the Centaurus region (see Fig. 3.7). In these maps, the region appears to have a consistent magnification pattern across all models, resulting in minimal change to the observed flux from NGC4945 regardless of the GMF model used.

However, a different behavior is observed for the second-brightest source in the SBG catalog, NGC253. This galaxy, located in the southern sky, shows a large shift in its contribution to the overall flux when GMF models other than `JF12` are used. In the `JF12` model, the southern sky, particularly the region where NGC253 is located, contains several areas with significant magnification. These regions allow more cosmic rays from NGC253 to reach Earth, making this galaxy the second most dominant contributor to the observed UHECR flux in simulations with the `JF12` model as GMF model. In contrast, when other GMF models are applied —in eight out of nine of the new models—the southern region, including the area around NGC253, becomes largely demagnified.

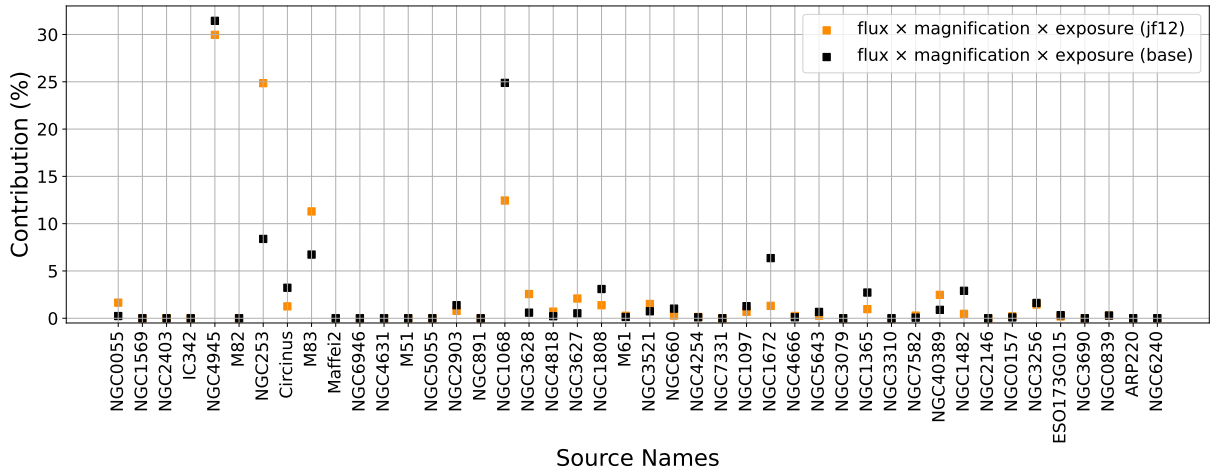



**Figure 3.8:** The expected source contribution employing the best-fit parameters obtained in Chapter 4.1. The two sets of points represent the contributions of sources from the Lunardini catalog, ordered by distance from closest to furthest, under two different GMF models. Black dots show the contributions when the `base` GMF model is used, while orange dots represent the `JF12` model. For several prominent sources, such as M82, the contribution is negligible since the source is well outside the Pierre Auger Observatory field of view (exposure).

This means that cosmic rays from NGC253 are scattered away from Earth, resulting in a significant reduction in its contribution to the overall flux. As a result, NGC253 is now barely visible in the UHECR flux under these GMF models. The shift in its contribution is a direct consequence of the demagnification effect observed in these newer GMF models. This reduction in the contribution of NGC253 has notable consequences for the visibility of other sources. In particular, other sources that were faint with the `JF12` GMF model may become more prominent. One of these sources is NGC1068, which was not largely contributing to the simulations based on the `JF12` model. However, when the `base` model is used for the GMF, NGC1068 becomes one of the most dominant sources, responsible for approximately 25% of the total UHECR flux. This shift highlights how changes in the GMF can modify the perceived importance of certain sources in the sky. NGC1068's increased visibility in the new GMF models suggests that it may play a much larger role in UHECR observations than previously thought.

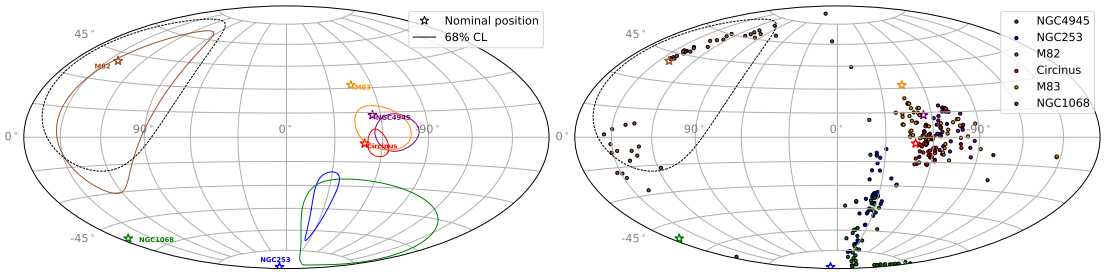
### Coherent deflections

It has become clear that the nominal position of cosmic ray sources in the sky is important when the effects of the GMF are accounted for in the arrival directions of UHECR.

The cosmic rays originating from these sources undergo substantial deflections as they propagate through the Galaxy due to interactions with the GMF. These deflections lead to noticeable deviations in their arrival directions which make it harder to trace them back to their sources. The degree of deflection is determined by the rigidity of the cosmic rays and the structure of the GMF, which varies across different regions of the Galaxy. For cosmic rays of similar rigidity, their deflection paths might follow comparable patterns, resulting in a clustering of events in a localized region in the sky. This clustering effect is particularly relevant in the analysis of UHECR anisotropies, as the deflections produce a displaced "image" of the source in the sky.

This shift, induced by the GMF, creates what might be interpreted as a "false" source location, leading to incorrect conclusions about the cosmic ray source if magnetic deflections are not properly taken into account. In some cases, the deflections might move the UHECR events away from the true source, while in others, a cluster of events could be wrongly attributed to an unrelated source simply due to proximity to the sky. This can lead to a situation where sources are mistakenly associated with regions where no actual source exists or where real sources go undetected because their UHECR flux is deflected away from their true position.

The deflection pattern and resulting image displacement are highly model-dependent. Different models of the GMF, such as the `JF12` and newer models, predict varying degrees of deflection for UHECRs passing through different regions of the Galaxy. As a result, the expected deflected positions of sources can vary significantly between models, leading to differing interpretations of where a cluster of UHECR events might originate. In Fig. 3.9, a visualization of these effects is displayed. The nominal position of the source is indicated by a colored star while the area in which the source image might lie is displayed by a solid contour with the same color code. This area is obtained computing the spread of the deflected events with a convexhull with a 68% confidence level (see Fig. 3.9a). The deflected events used to determine the area are displayed in Fig. 3.9b, where the JF12 is employed to compute the deflections of events coming from the main sources in the SBG catalog. An important behavior that can be observed concerns the sources in the Centaurus region: the events coming from M83 are deflected in the nominal position of NGC4945 which might amplify an overdensity in this region, while the events coming from NGC4945 are deflected towards the Galactic plane possibly misplacing the position of observed clusters. The same reasoning might be applied to Circinus which lies on the Galactic plane, its events are not deflected far from the source but they might increase the event clustering in the misplaced position driven by the deflected events originating from NGC4945. A crucial observation concerns the sources located




(a) (b)

**Figure 3.9:** (a): Region of interest (ROI) for the deflected position of the source computed with a 68% confidence level. The main sources are displayed with a star while the ROI is displayed with a solid line with the same color scheme. (b): Events following the deflection pattern of the `JF12` model. The events are injected with a $E^{-1}$ spectrum at the source and deflections are computed for a turbulent and coherent field. The events are displayed with a colored dot related to their source.

in the southern part of the sky. Specifically, the events originating from the galaxy NGC253 show significant deflections, approximately 90° away from their source position, towards the Galactic center. Similarly, events originating from another prominent source, NGC1068, display a different pattern of deflection, where the deflections occur around the position of NGC253 itself.

These deflection patterns introduce significant challenges in interpreting the cosmic rays anisotropy. If NGC253 is not in a region of demagnification, its contribution will show as a clustering of events in the southern sky, potentially leading to an accumulation of events in this region. However, such clustering is not visible in the current data, raising questions about the actual source of these events.

On the other hand, if NGC253 lies in a demagnified region of the sky, its contribution to the observed events would be negligible or close to zero. In this case, the events observed in the southern region may not be linked directly to NGC253. Instead, these events could have originated from a different source such as NGC1068, and due to strong deflections, they are now observed in new positions in the sky. This scenario complicates the source identification process, as the observed arrival directions may not correlate well with the true locations of their cosmic ray sources.

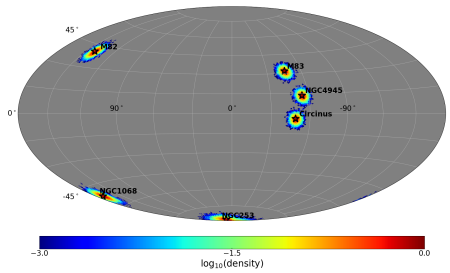
### Effective description of the EGMF

As previously mentioned, the effects of the EGMF have been modeled through a 2D Gaussian on the sphere, although a more realistic approach, that can be employed in

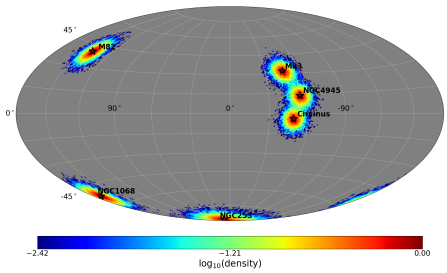
simulation, is to introduce the rigidity dependence that is expected for magnetic field deflections. Incorporating a rigidity-dependent magnetic field deflection is feasible within the presented modeling of the astrophysical scenario because it includes the overall simulation of UHECR from sources to Earth. This implies the knowledge of the detected energy together with the particle type allowing a prediction for the rigidities of the final arrival events based on their source location. For rigidity-dependent deflection, the deflection angle is given by [71,72]:

$$\theta = \theta^* \sqrt{\frac{D_{\text{source}}}{D^*}} \frac{R^*}{R_{\text{event}}} \tag{3.5}$$

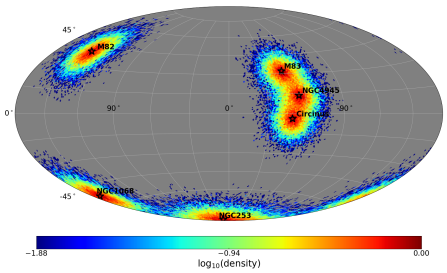
where the $D^*$ is set at 3.47 Mpc which is the distance of NGC4945, the brightest source in the SBG catalog (if the exposure of the Auger experiment is taken into account), $R^*$=40/7 EV which is chosen based on the consideration that the spectrum at the energies in the analysis (events above $\simeq$ 40 EeV) is dominated by Nitrogen (Z=7). The parameter $D_{\text{source}}$ indicates the distance of the source while $R_{\text{event}}$ is related to the rigidity of the individual observed event. The parameter $\theta^*$ is a free parameter in the analysis that is varied to quantify the differences between several smearing. The impact of adding an extra smearing component to UHE events is shown in Fig. 3.10, where an initial smearing angle, $\theta^*$, of 5, 10, and 20 degrees (presented in the top, left and right panels respectively) is applied to six sources in the SBG catalog for Nitrogen events at 40 EeV ($R$ = 40/7 EV). It is easy to see that this smearing produces a cluster of events near the nominal positions of the sources, composed of high-energy events that remain almost unaffected by the EGMF, allowing them to point directly to their source. In contrast, lower energy events are significantly deflected, spreading across portions of the sky farther from the nominal source (or further away) positions and defining the broader distribution. The area covered by the distributions obtained employing Eq.3.5 are similar since the sources are characterized by approximately the same distance. In this analysis, Eq.3.5 will be employed for an effective description of the smearing of the arrival direction in the turbulent EGMF.




**(a)**




**(b)**




**(c)**

**Figure 3.10:** Deflection patterns of UHE events originating from five of the brightest sources in the SBG catalog. The top panel shows the deflections that charged particles undergo when a smearing angle of $\theta^* = 5°$ is used. The left and right panels show the deflection patterns related to the same events generated from the source sample when a smearing angle $\theta^* = 10°$ and $\theta^* = 20°$ are used.

# Chapter 3. Simulation of UHECR propagation and deflections in magnetic fields

# Chapter IV

# Cross-correlation analysis

The propagation of UHECRs is governed by a complex relation of physical processes and astrophysical environments, including interactions with magnetic and photon fields. These effects significantly influence the observed arrival directions of cosmic rays, complicating the efforts to trace their origins back to their sources. As discussed in the previous chapter, the GMF deflections play a critical role in shaping the arrival distribution of UHECRs.

In this chapter, deflection and (de-)magnification effects are introduced into the SBG model to evaluate whether the reported correlation between UHECR arrival directions and SBG sources holds under more realistic astrophysical assumptions. A likelihood analysis is performed to quantify the impact of the GMF on source identification and to assess its influence on the observed anisotropy patterns. Using the SBG catalog as a primary example, the analysis evaluates the correlation between simulated arrival directions and observed data. The results provide insight into the effects of the magnetic fields on the anisotropy patterns expected from UHECR sources.

## 4.1. Compatibility of simulated scenarios with data

To simulate the expected outcomes of the likelihood analysis conducted in [21] while including the magnetic field deflections previously described, multiple datasets are created where the signal contribution is systematically adjusted. In each dataset, the number of events used as signal is determined as $f_{ani} \times N_{data}$, where $f_{ani}$ represents the anisotropic fraction of the total events, and $N_{data}$ is the total number of events above the energy threshold of 38 EeV, set at 1621 for this analysis.

As shown in Fig. 3.1, it is important to note that the events associated with the signal contribution are taken from the simulated arrival directions after applying the effects of the GMF. These events are randomly selected from all simulated events originating from sources in the SBG catalog. This means that sources initially outside the observable range, such as those that would not contribute without accounting for the GMF due to their position being located inside the exposure blind spot, may contribute to the observed events once the GMF effects are introduced. This consideration is also influenced by the magnification and demagnification region shown in Fig. 3.7: a source that sits in the exposure blind spot might become relevant also because it sits in a magnified region therefore if relevant structures such as clustering of events are observed at the edge of the exposure blind-spot or in regions close to it, it might be related to contributions of sources that can be explained once the coherent and turbulent deflections are taken into account. An example of simulated events that underwent GMF deflections can be seen in Sec. 4.2.2, where the impact on the likelihood analysis when different models of the GMF are used is further described.

## Likelihood analysis

For each of the realization that represents the Auger dataset, the same maximum likelihood analysis is applied as was used in [21] to determine the best cross-correlation with astrophysical catalogs (see Sec. 2.1.2). The maximum likelihood analysis compares two hypotheses: the null hypothesis ($H^0$) versus the alternative hypothesis ($H^1$). The null hypothesis is defined as an isotropic flux distribution such as:

$$\mathrm{n}^{H_0}(u) = \frac{\omega(u)}{\sum_i \omega(u_i)} \tag{4.1}$$

where $\omega(u)$ is the direction exposure, following [83], of the array and $i$ is the pixel index for a sphere of bins of equal area obtained with `HEALPix` [80] using an `nSide`=64. A visual representation of the exposure is shown in Fig. 4.1a. The alternative hypothesis is modeled as the sum of the isotropic component with a component related to the source catalog. $H^1$ is defined as:

$$\mathrm{n}^{H_1}(u) = (1-\alpha) \times n^{H_0}(u) + \alpha \times \frac{\sum_j s_j(u,\theta)}{\sum_i \sum_j s_j(u_j,\theta)} \tag{4.2}$$

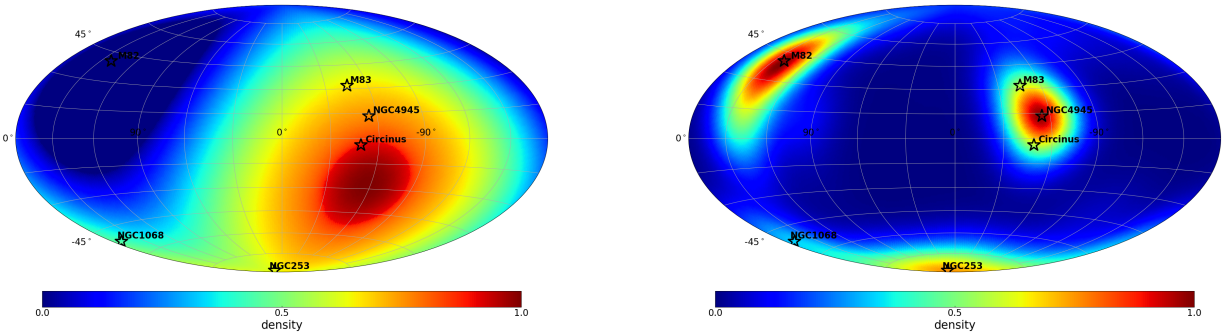
where the contribution of the sources in the catalog ($s_j(u, \theta)$) is modeled as a von Misses-Fisher distribution (Eq. 3.4) centered at the position of the source with smearing angle $\theta$. A visualization of the second term in Eq.4.2 (source contribution) is shown in Fig. 4.1b with a smearing angle $\theta = 15°$. In this particular sky map the exposure is not yet applied to show the effect of using a fixed angle around the nominal position of the sources. As previously mentioned this method has limitations, which led to changes in $H^1$, that have been analyzed and explained in Sec. 3.2.1-3.2.2. To quantify the strength of the anisotropic model versus the isotropic one, the hypotheses are combined in the test statistic which is defined as:

$$\mathrm{TS} = 2 \sum_i k_i \times \ln \frac{n^{H_1}}{n^{H_0}} \tag{4.3}$$

where $k_i$ is the number of events in direction $i$ in the simulated dataset. The likelihood analysis and the TS employed to quantify the strength of the alternative hypothesis against the null hypothesis relies on the Wilks theorem which states that the distribution of the likelihood ratio test statistic, under the null hypothesis, asymptotically follows a $\chi^2$ distribution with degrees of freedom equal to the difference in the number of parameters between the null and alternative models, if $H^1$ and $H^0$ are "nested models", i.e. as in this case since $H^1$ equals $H^0$ for $\alpha = 0$. This result holds when the sample size is large and is widely used in hypothesis testing to approximate the significance of TS [84]. Using Wilks' theorem, the $\chi^2$ distribution can be used to determine the p-value associated with the observed TS, therefore quantifying the probability of observing the given TS under the assumption of isotropy. A low p-value would indicate that the observed data are unlikely described by $H^0$, which provides evidence for anisotropy. Another important remark is related to the number of pixels used for constructing the null and alternative hypothesis: `nside` has been chosen as 64 which corresponds, under the relation:

$$\text{pixel size} = \sqrt{(4\pi/\texttt{Npix}) \cdot (180°/\pi)^2} \tag{4.4}$$

where $\texttt{Npix} = 12 \cdot \texttt{nside}^2$, to an angular resolution of 0.92° that aligns with the resolution of the Auger experiment ($\simeq 1°$). This choice yields a refined angular resolution together with a large number of pixels that creates a small area in the sky suitable for detecting variations between isotropic and anisotropic flux distributions.




(a) (b)

**Figure 4.1:** (a): Exposure density function for the Pierre Auger Observatory in Galactic coordinates (b): Alternative hypothesis with 2D-Gaussian applied to all the sources in the SBG catalog weighted by their flux.

## 4.2. Impact of the GMF on likelihood analysis

The likelihood analysis conducted on simulations intends to identify realizations in which the best-fit parameters, namely: best-fit signal fraction ($\alpha$), best-fit search radius ($\theta$), and TS value (TS), match the results observed for the data in [21]. To assess whether a simulated realization is compatible with the reference values, Monte-Carlo simulations were conducted: these simulations generate multiple realizations based on the best-fit hypothesis presented by the Auger Collaboration and by analyzing these realizations, a $1\sigma$ range was established for the TS values, which defines a threshold for compatibility. The resulting interval, TS$\in$[13,34], is used as a criterion to determine whether the TS values of simulated realizations fall within the expected range, indicating statistical agreement with the Auger best-fit parameters. This information is combined in Fig. 4.2. The figure shows the best-fit parameters obtained from the likelihood analysis of simulated datasets. Every point is a realization obtained by drawing 1621 events (same as observed in the data) and fitting for the best anisotropy fraction and search radius. This particular realization includes a signal contribution of 15%, and it illustrates how the simulated best-fit parameters are distributed compared to the constraints set by the Auger Collaboration. The simulated realizations cover a broad range of values, pointing out the variability in the simulations. To determine if a simulated dataset is compatible with the Auger results, a selection criterion is applied as follows: only realizations that fall within the uncertainty bounds of all the two reported best-fit parameters are considered compatible.

Moreover, as previously mentioned, compatibility in terms of signal fraction and

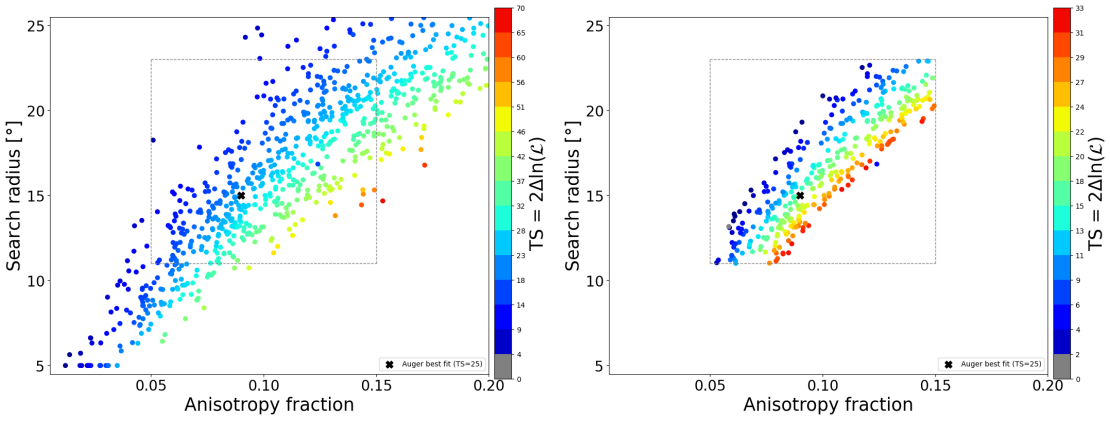



**Figure 4.2:** Simulated datasets with a signal contribution $f_{\mathrm{sgn}} = 15\%$ for the `JF12` model. Every point corresponds to the best-fit parameters obtained with the likelihood analysis of one realization of an Auger data set. The color code indicates the value of the TS while the grey lines indicate the errors on the best-fit parameters reported in [21]. In panel 4.2a, all simulated datasets are displayed while in panel 4.2b only the datasets with compatible best-fit parameters are shown.

search radius implies a high likelihood of compatibility in the TS as well. Therefore, a third selection is applied to the simulations, reducing the number of realizations that satisfy all three criteria. This multi-parameter selection provides simulations that match all individual best-fit parameters and should align with the Auger dataset. The left panel in Fig. 4.2 shows the effect of the third selection which reduces the number of realizations allowed to be considered valid when the TS constraints are taken into account. By performing this analysis across different signal contributions, it is possible to identify the configuration that maximizes the fraction of simulated datasets agreeing with the Auger findings. This approach provides insights into which signal strengths correspond to the highest compatibility with Auger's reported best-fit parameters, highlighting the behavior of the analysis when the strength of the source contribution is increased.

The results of this analysis are presented in Fig. 4.3, which shows the outcomes for various signal contributions together with multiple realizations of the turbulent component of the `JF12` model. Testing different levels of signal contribution allows considerations on how the signal strength affects compatibility with Auger results while employing multiple GMF realizations show the impact of magnetic field fluctuations on the results of the analysis. It is worthwhile noting, that the due to the three-dimensional selection in ($\alpha$, $\theta$, TS) within one standard deviation, a perfect match of data and simulation would imply that a fraction of $(0.683)^3 = 0.319$ of the realization is selected.

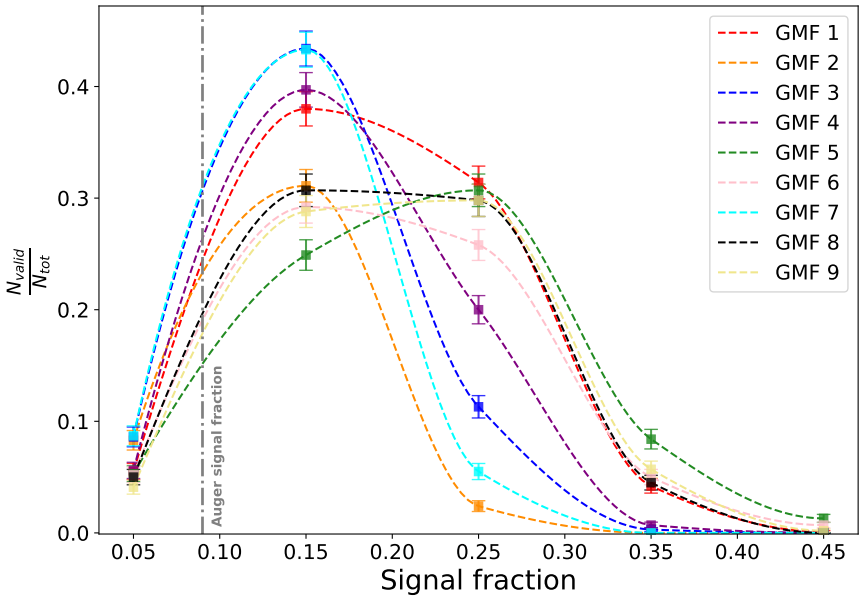



**Figure 4.3:** Fraction of simulated datasets in agreement with the Auger findings when a selection on best-fit anisotropy fraction, best-fit search radius, and TS is applied. The dashed lines represent the variation of the fraction of valid realizations obtained employing the likelihood analysis under different realizations of the GMF turbulent component. The grey line indicates the best-fit anisotropy fraction reported in [21].

This is indeed about the maximum fraction observed in simulations, indicating that the simulations describe the data reasonably well.

### 4.2.1. Impact of the turbulent field variations

As discussed in Sec. 3.2.1, it is important to account for potential variations of the same GMF model to estimate the effects of the Galactic variance. By incorporating different realizations of the turbulent component, this analysis can determine if the likelihood ratio remains stable when variations in the GMF model are introduced. This consideration is important since it reveals whether the TS results show sensitivity to changes in the GMF model's configuration.

Since the events for the construction of the simulated datasets are characterized by energies above 38 EeV (see Sec. 2.1.2) and the spectrum at high-energy is dominated by heavy elements such as Nitrogen and Silicon, the resulting rigidity range reduces the

variations between the different configurations of the GMF as shown in Fig. 3.4. This aspect can also be seen in Fig. 4.3 where several configurations share the same behavior and similar agreement in the signal fraction that maximizes the number of compatible simulated datasets.

Minor differences are present when configurations are compared; for example, the number of realizations that are in agreement with the Auger data can vary by 10%. Additionally, the signal fraction that maximizes agreement with observations ranges from 15% to 25%, reflecting some variability in the derived best signal contribution. This variability can be explained by the intrinsic randomness of the turbulent magnetic field, which introduces extra fluctuations that manifest as distinct behaviors across configurations. These differences show the influence of the turbulent field's stochastic nature and the Galactic variance in the correlation between simulated datasets and the observed data.

In addition to the differences caused by variations in the turbulent field, the analysis shows a general overestimation of the signal fraction necessary to have simulated datasets in agreement with observed data. As previously noted, the anisotropy fraction reported in the Auger analysis is around 9%. However, when accounting for variations in the turbulent magnetic fields, the signal fraction that maximizes agreement with data ranges from 15% to 25%. Thus, the simulations suggest that the best-fit fraction $\alpha$ underestimates the true signal fraction of 9% and this aspect may be explained by the injected spectrum and mass composition above the energy threshold of the analysis ($E > 38$ EeV). Since here it is assumed that the source properties are following the combined fit of spectrum and composition [73]), the integral flux is composed of around 60% of Nitrogen and 40% Silicon (see Fig. 3.3). These heavier elements, when they undergo the effects of the GMF, experience significant deflections away from their origins and often appear as part of the isotropic background. This suggests that a fraction of what is injected as a signal in the simulations might appear as an isotropic background due to magnetic deflections. As a result, the signal contribution in the simulations must be increased to reproduce the level of anisotropy observed in the Auger data.

### 4.2.2. Impact of the coherent field variations

The results of the likelihood analysis are influenced not only by the turbulent component of the GMF but also by the choice of the GMF model itself. Different models of the GMF can produce different deflection patterns for charged particles, which impacts the interpretation of anisotropy and the inferred source contributions. In addition to

established models such as JF12, newer GMF models, such as the UF23 model ensemble, offer alternative descriptions that may better capture the complexities of the Galactic field.

The introduction of newer models did not change the structure of the likelihood ratio itself; however, several differences emerged in the components contributing to the TS. One of the main differences is illustrated in Fig. 4.4, which shows the density of events deflected by magnetic fields under a different new GMF model (base and twistX model), alongside the previously used model based on JF12. These variations in event density underline how different GMF models can impact the interpretation of deflection patterns and the overdensities that characterize the TS. The event densities show significant differences in deflection patterns when comparing the newer GMF models with the JF12 model. One of the most important differences is related to the southern region of the sky, where sources like NGC253 contribute to a clustering of events in the JF12 model. This clustering effect, however, is reduced when other models, such as the base model, are applied. This variation can be understood by examining the magnification map shown in Fig. 3.7. In the JF12 model, certain sources, including NGC253, fall within regions of high magnification. This means that cosmic rays originating from these regions have a stronger contribution to the observed flux. In contrast, in the base model, these same sources are located in demagnified regions. Therefore, their contribution to the observed flux is minimal, as the cosmic rays coming from that region have no valid deflection patterns to reach Earth and be observed. These differences underline that the positioning of sources relative to magnified or demagnified regions in the magnetic field significantly influences event clustering and the resulting anisotropy signal.

Another significant difference consists in the interpretation of observed event clusters. As shown in Fig. 4.4, overdensities appear in the Centaurus region as well as in the southern portion of the sky. Considering the previously discussed magnification effects, these overdensities may be linked to potential sources in the SBG catalog. In the Centaurus region, sources such as NGC4945, M83, and Circinus are possible sources that contribute to the observed clustering. However, examining the specific source contributions in Fig. 3.8 reveals that NGC4945 and M83 are the primary contributors to the overall flux, with Circinus playing a minor role. The same reasoning can be applied to the sources located in the southern region of the sky: among the possible contributors to the overdensity, such as NGC253 and NGC1068, NGC253 has a large contribution only if the JF12 model is used while NGC1068 starts to be relevant only when the base model is employed as GMF model. These differences highlight how the interpretation of sources contributing to sky overdensities can vary significantly depending on the chosen GMF

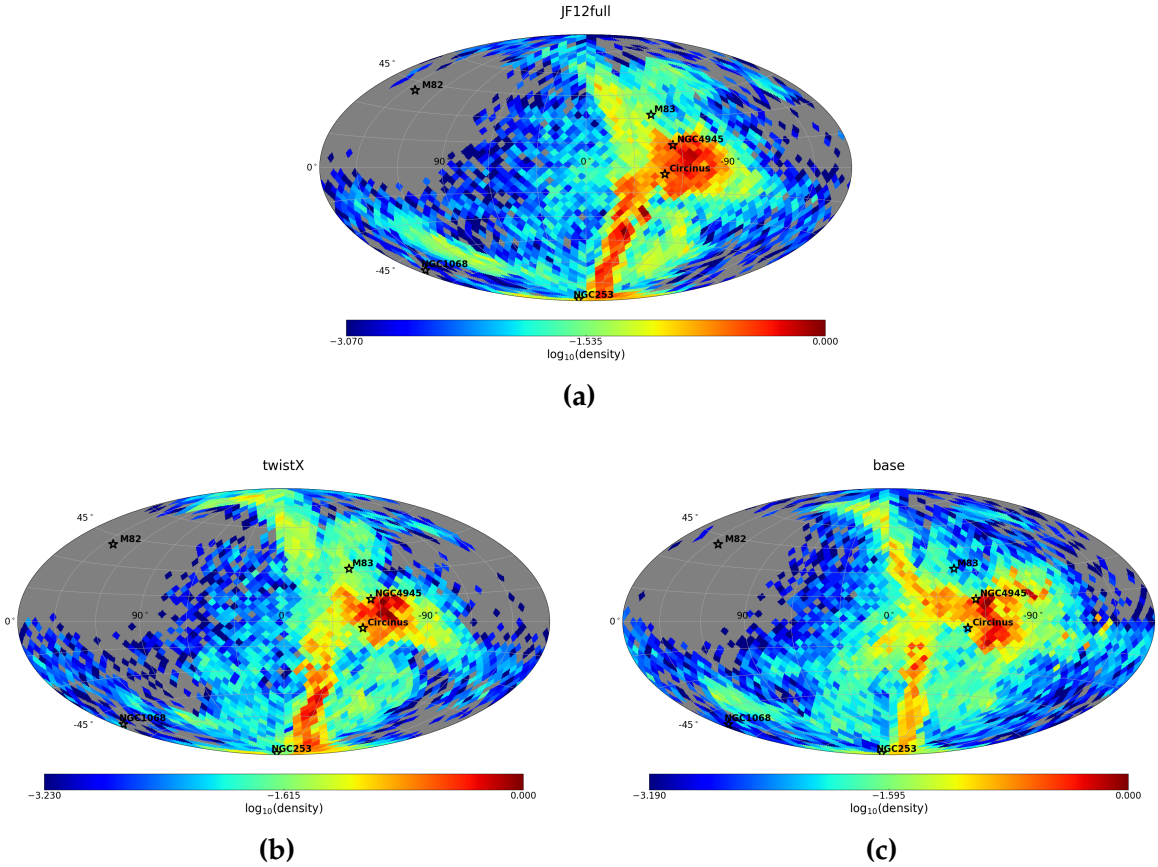



**Figure 4.4:** Deflection patterns of UHE events originating from the sources in the SBG catalog. The top panel shows the deflections that charged particles undergo when the `JF12` model with both turbulent and coherent components is used. The left and right panels show the deflection patterns related to two new models of the UF23 ensemble (`base` and `twistX`) that underline several differences in the amplitude of the density of the events in the Centaurus region together with the diluted overdensity in the southern portion of the sky.

model.

The behavior of the simulated datasets under the hypothesis of a new GMF model is presented in Fig. 4.5. Although the overall spread of the data points remains consistent between the analysis conducted with the `JF12` model and the one conducted with the `cre10` model, there is a noticeable change in the slope of the data points. The transition from realizations with relatively low TS values to those with comparable or even higher TS values is visibly steeper in the `cre10` model. Additionally, this model produces fewer realizations that exceed the upper limit of the TS selection criteria. These differences can be attributed to the considerable demagnification of NGC253 (and magnification

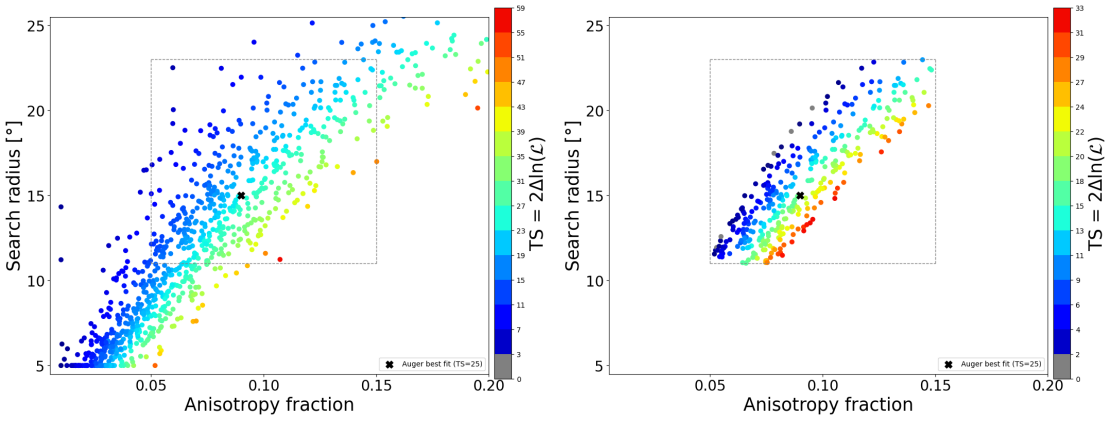



**Figure 4.5:** Simulated datasets with a signal contribution $f_{\mathrm{sgn}} = 15\%$ for the `cre10` model. Every point corresponds to the best-fit parameters obtained with the likelihood analysis of one realization of an Auger data set. The color code indicates the value of the TS while the gray lines indicate the errors on the best-fit parameters reported in [21]. In panel 4.5a, all simulated datasets are displayed while in panel 4.5b only the datasets with compatible best-fit parameters are shown.

of NGC 1068) as previously mentioned. As a result, the primary contribution is now concentrated in the Centaurus region, aligning well with the observations, while no other contributions seem to be relevant. Similar qualitative changes are observed for the other models of the UF23 ensemble.

Fig. 4.6 illustrates the overall behavior of the TS as a function of the signal contribution for the new models. Similar to the findings with the `JF12` model, a comparable behavior is shown when comparing the anisotropy fraction reported by the Auger Collaboration with the signal fraction required to align with the observations. As observed in the `JF12` model, the signal contribution must be increased due to the presence of heavy elements in the high-energy spectrum above 38 EeV. However, quantifying the precise range of the required signal contribution in this case proves to be challenging due to the differences between the GMF models. Generally, it is clear that the signal fraction required to obtain a good agreement with the data is above 10%, but each model shows a distinct behavior as the signal fraction varies. This variability is largely attributable to the nature of the new GMF models: while they have similarities in fundamental aspects, they differ significantly in their construction and the data used in their development (see [68]). These structural differences lead to unique behaviors across models when adjusting the signal fraction therefore the interpretation of the results needs to be carried out model

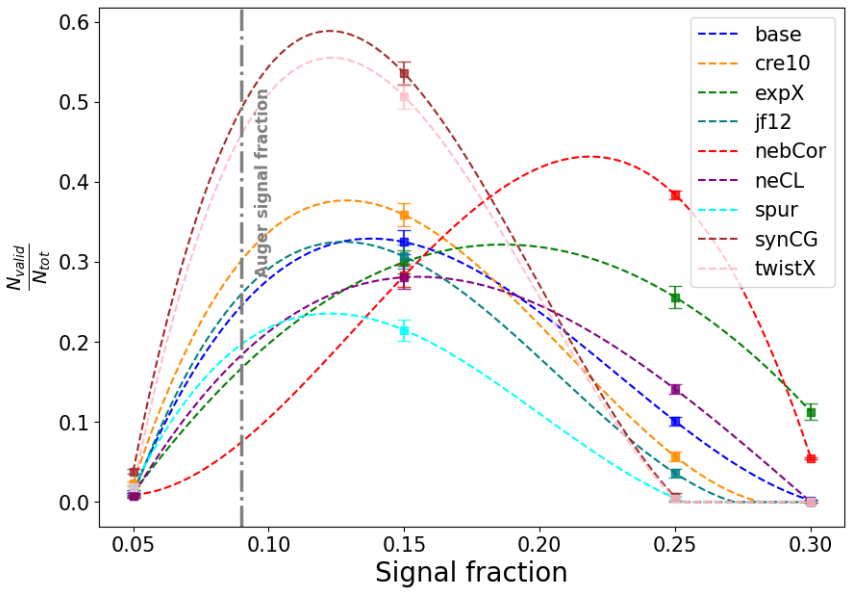



**Figure 4.6:** Fraction of simulated datasets in agreement with the Auger findings when a selection on best-fit anisotropy fraction, best-fit search radius and TS is applied. The dashed lines represent the behavior of the likelihood analysis under different realizations of the GMF models. The gray line indicates the best-fit anisotropy fraction reported in [21].

by model. For example, if the `twistX` model in Fig. 4.6 is used (pink line), its behavior is similar to the one shown in Fig. 4.3 where different realizations of the `JF12` model are presented, this is due to the fact that the `twistX` model shares several similarities with the `JF12` model such as a similar structure of the magnification region around the brightest sources in the SBG catalog together with a similar pattern in the mean deflection angles as shown in Fig. 3.7 and Fig. 3.6 respectively. Applying the same reasoning to the **`expX`** model (green curve), which has larger deflection angles around the Galactic center due to its strong poloidal field component, underlines why a higher signal fraction is necessary to align with observational data. This is a direct consequence of the interpretation already presented concerning the deflections of heavy elements that can appear as an isotropic background. The signal fraction required to match observations in the **`expX`** model is nearly double that of the `twistX` model and the highest among the models analyzed. This finding aligns with expectations, as the pronounced deflections in the **`expX`** model effectively dilute the anisotropic signal, imposing a larger injected signal fraction to achieve comparable results.

In conclusion, the analysis using simulated cosmic ray events in which GMF deflections are taken into account demonstrates that an agreement of model parameters,

specifically signal fraction, search radius, and TS values, can be found when the SBG source model is used as model for the signal contribution (see also [85]). This holds even when different configurations of the turbulent GMF component are applied within a single GMF model, indicating that Galactic variance in the turbulent field does not substantially impact the TS's stability. Moreover, the analysis extended to newer GMF models, such as `expX`, `base`, and `twistX`, underlines the effects of refined modeling adjustments based on recent data. These models include modifications that better align Galactic components and correct certain overmagnification issues observed in the `JF12` model alongside its overestimation of the turbulent field ( [66,86]), particularly relevant to high-energy source distributions and deflections of high-energy charged particles. While compatibility with observed data can still be obtained under these newer models, the results indicate a dependency on the specific GMF model chosen. Thus, while an overall agreement with the observations can be found, the amount of signal fraction required may vary across models due to differences in cosmic ray deflection patterns and field configurations, underlining the importance of model dependency in the interpretation of cosmic ray anisotropy at ultra-high energies. Furthermore, results obtained by employing different models for the coherent magnetic field alongside different realizations of turbulent fields suggest that a contribution from the EGMF is not required to achieve compatibility with findings reported by the Auger Collaboration. This topic will be explored in more detail in the next chapter.

# Chapter V

# UHECR intermediate-scale anisotropies

The likelihood analysis conducted on simulated datasets above 38 EeV showed that an agreement in terms of search radius, anisotropy fraction, and test statistic can be found with the observations. A crucial component in understanding the correlation between UHECR arrival directions and their potential sources lies in the detection of localized excesses in the UHECR sky. These intermediate-scale anisotropies offer an opportunity to identify regions where cosmic ray events deviate from an isotropic distribution, potentially pointing to the presence of individual sources or extended structures that contribute to the observed UHECR flux.

Intermediate-scale anisotropies can provide key insights into the nature of UHECR sources. They could reflect the presence of point-like astrophysical sources or more extended cosmic structures, such as galaxy clusters or superclusters. Detecting these anisotropies would not only offer insights into specific regions of the Universe where UHECRs might be produced but also provide valuable information on their angular extent and position as a function of rigidity, adding crucial findings on cosmic ray propagation models, particularly the role of magnetic fields.

The Pierre Auger Collaboration has investigated the presence of intermediate-scale anisotropy in the UHECR sky, as reported in [21]. In this analysis, a cluster of UHECR events with local significance of $5.4\sigma$, corresponding to a post-trial p-value of 3%, was observed in the Centaurus region at Galactic coordinates (l, b) = (305.4°, 16.2°). This result was achieved by systematically scanning the sky and comparing, at each position, the number of observed events within a specified circular region to the expected event count under an isotropic event distribution. This analysis identifies localized excesses, which may suggest the presence of astrophysical sources or other possible structures in

the UHECR arrival directions.

## 5.1. Structures in the arrival direction of UHECRs

The UHECR events observed above 38 EeV, according to the likelihood analysis, show a consistent deviation from isotropy, suggesting that high-energy cosmic ray events are not randomly distributed across the sky. Instead, they form distinct patterns in their arrival directions, indicating the presence of structured regions. The importance of such overdensities will be presenteed in the following sections.

### 5.1.1. Identification of structures in the arrival direction of UHECRs

Together with the deviation from isotropy, the Pierre Auger Collaboration also determined the position in the sky where the highest deviation from isotropically distributed events is placed. This result has been achieved by scanning the overall sky and for every position computing the Li-Ma significance [87] where the events observed in a given circular window are compared to the events observed if they were isotropically distributed. The significance is computed following:

$$S_{\mathrm{LM}} = \sqrt{2\left(N_{\mathrm{on}} \cdot \ln\left(\frac{(1+\alpha)N_{\mathrm{on}}}{\alpha(N_{\mathrm{on}}+N_{\mathrm{off}})}\right) + N_{\mathrm{off}} \cdot \ln\left(\frac{(1+\alpha)N_{\mathrm{off}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right)} \tag{5.1}$$

Here $N_{\mathrm{on}}$ is the number of events within the circular window, and $N_{\mathrm{off}}$ is the number of events outside this window. The parameter $\alpha$ represents the ratio of the expected background in the circular window region to that in the surrounding area, accounting for differences in the exposure, for the case of isotropically distributed events. The circular window is defined by a top-hat function with an angular radius $\Psi$, indicating the window's size. To compute the Li-Ma local significance, the top-hat function is parametrized as:

$$\mathrm{TH}(\theta) = \begin{cases} 1 & \text{if } \theta \leq \Psi \\ 0 & \text{if } \theta > \Psi \end{cases} \tag{5.2}$$

where $\theta$ represents the angular distance between the center of the window and the position of the event, while $\Psi$ denotes the angular radius, which determines the size of the search region. This parametrization leads to a sharp cutoff, which could result

in a sudden transition from the signal region to the background region. Additionally, such a sharp boundary between search regions may cause "overcounting" in cases where circular regions overlap. This limitation of the top-hat function is worth noting, as it might impact regions where the count of observed events or the expected count from an isotropic distribution is low. Angular sizes from 1° to 30° with a step of 1° were scanned over by the Pierre Auger Collaboration to determine the optimal value of $\Psi$ that maximizes the local significance. This optimal value was found to be $\Psi_{\text{Li-Ma}} = 24°$. In addition to the angular size, the energy range from 32 EeV to 80 EeV, with steps of 1 EeV, was also scanned to find the energy that maximizes the local significance. The optimal value for the energy was found at $E_{\text{Li-Ma}}$ at 41 EeV and it will employed as energy threshold in the following analysis. In this work, a similar approach is used to determine the overdensity in the Centaurus region. Specifically, a binomial distribution is employed to quantify the local significance, in contrast to the Li-Ma significance. This approach considers the probability of observing $N$ events in a given circular region, under the assumption of isotropic arrival directions. Here, $N_{\text{tot}}$ represents the total number of observed events, and $p$ is the isotropic probability that an event falls within the circular region (top-hat). The probability $p$ is calculated as the ratio of the integrated exposure within the region to the total exposure.

$$P(X \geq N \mid N_{\text{tot}}, p) = \sum_{k=N}^{N_{\text{tot}}} \binom{N_{\text{tot}}}{k} p^k \, (1-p)^{N_{\text{tot}}-k} \tag{5.3}$$

To convert the binomial probabilities into significance values expressed in units of standard deviations ($\sigma$), the inverse cumulative distribution function (CDF) of the standard normal distribution is used. In Fig. 5.1, a comparison between the significance values obtained using the binomial method and the Li-Ma method is shown. The comparison is performed for the full angular scan at `nside` = 64. As it can be seen, there is a significant, but small difference between the two methods. On average, the significance derived from the binomial distribution is approximately $0.1\sigma$ smaller than that obtained using the Li-Ma method. For the subsequent analyses, the significance derived from the cumulative binomial probability will be used.

### 5.1.2. Local significances in the UHECR sky in the Auger data

Using the binomial approach, a blind search is conducted over the UHECR sky, with the circular window varying between 1° and 80° in steps of 1°. The probability of observing

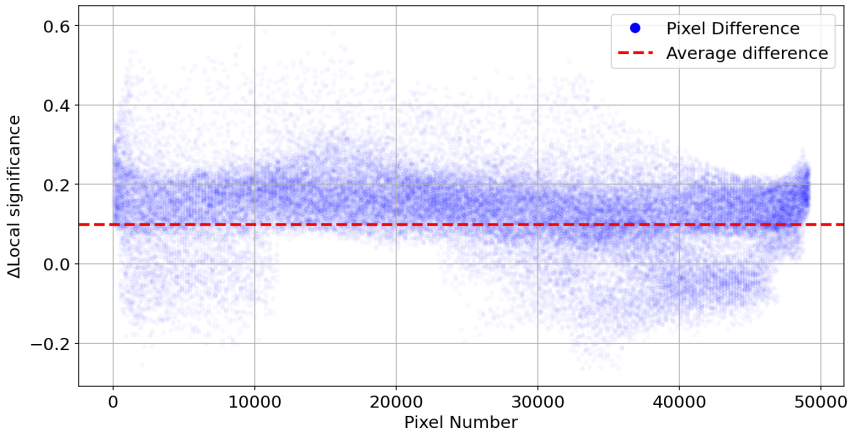
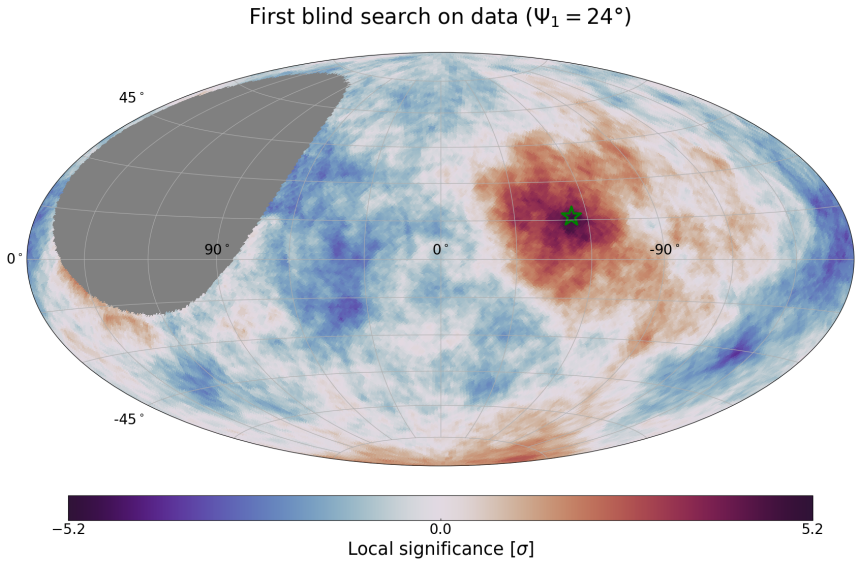


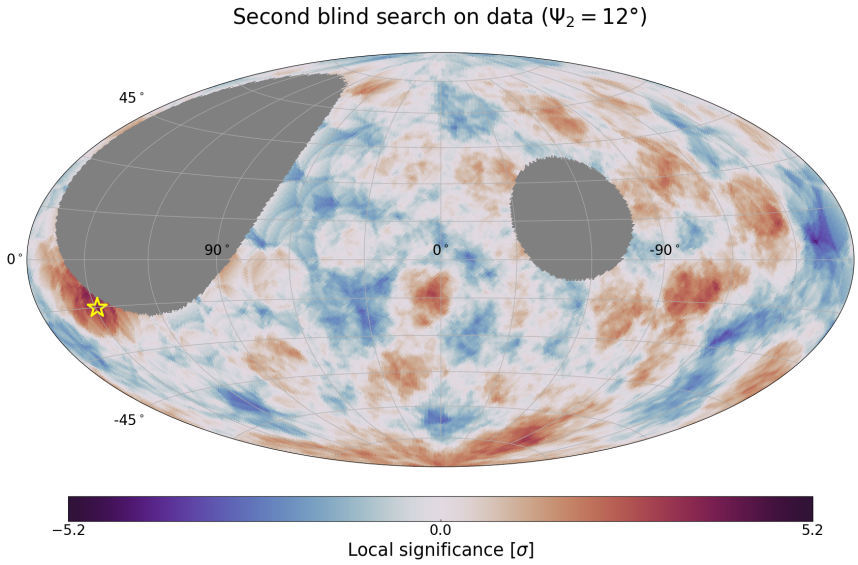

**Figure 5.1:** Difference in the values pixel by pixel for the computation of the local significance between the Li-Ma method and binomial approach. On the $x$ axis the pixel number is shown for a `HEALPix` grid of `nside`=64. The average difference is shown with a dashed red line.

a deviation from isotropy is then computed using Eq.5.3. This method also allows for the determination of the maximum angular size that returns the highest value of local significance. The first blind search yields a local significance of $\sigma^{\text{loc}}_{1,\text{obs}} = 5.2\sigma$ with a top-hat search radius amplitude of $\Psi_{1,\text{obs}} = 24°$ located at $(l, b)_{1,\text{obs}} = (305.4°, 16.2°)$, which shows a good agreement with the values found by the Pierre Auger Collaboration [21]. After determining the relevant information concerning the local significance, such as its amplitude $\sigma^{\text{loc}}_{1,\text{obs}}$, its position $(l, b)_{1,\text{obs}}$, and the radius of the search region $\Psi_{1,\text{obs}}$, a second blind search is conducted following the same approach. However, in this case, the region of interest defined by the search window of amplitude $\Psi_{1,\text{obs}}$, centered on the position where the highest local significance was detected, is excluded. This exclusion of the first local significance is done with the intention of identifying any potential additional structures in the UHECR sky, which may only become relevant if the most significant structure, which dominates the sky, is removed. The second blind search yields $\sigma^{\text{loc}}_{2,\text{obs}} = 3.4\sigma$, which represents the value of the second highest local significance, additionally, the position of this hotspot in the sky is located at $(l, b)_{2,\text{obs}} = (149.1°, -15.1°)$, with $\Psi_{2,\text{obs}} = 12°$ that refers to the radius of the search region that maximizes the deviation from isotropy.

An overview of the described method alongside the results of the first blind search can be seen in Fig. 5.2. The results of the first blind search are shown in panel 5.2a where the position of the highest local significance is highlighted with a green star. Similarly, panel 5.2b illustrates the appearance of the overall sky after excluding the circular




**(a)**




**(b)**

**Figure 5.2:** (a): The results related to the first blind search where the location of the first overdensity is indicated by a green star while the amplitude of the search window is indicated by $\Psi_1$. (b): The results related to the second blind search where the location of the second overdensity is indicated by a yellow star while the amplitude of the search window is indicated by $\Psi_2$

region around the position of the highest local significance. In this case, the radius that maximizes the second local significance is used to compute the overall skymap.

The analysis of the data indicates that, if a second overdensity is observed in the sky, it is associated with a smaller top-hat search radius ($\Psi_{2,\mathrm{obs}} < \Psi_{1,\mathrm{obs}}$). This suggests that, once the largest overdensity is excluded, only smaller-scale structures contribute to the local significance, requiring a narrower search window. Furthermore, the local significance $\sigma^{\mathrm{loc}}_{2,\mathrm{obs}}$ is approximately $3.4\sigma$ which indicates that when a blind search that excludes the major overdensity, the events in the remaining portions of the sky are distributed almost isotropically.

When the sequential blind search analysis is applied to simulations that incorporate GMF deflections, it should ideally reveal similar features to those observed in the data such as an overdensity in the Centaurus region and additional minor structures consistent with isotropy once the primary clustering area is excluded. The directional scan of the simulations is preformed above the same energy threshold as found for the data, E=41 EeV, with $N_{\mathrm{tot}} = 1274$ events above this energy. The analysis of these simulated datasets is presented in the following section, where the behavior of the second overdensity is analyzed in details exploring also the correlation with the EGMF extra-smearing component.

### 5.1.3. Local significances in the UHECR sky in simulations

Due to the introduction of GMF deflections in the arrival directions of UHECR used to construct the simulated datasets, several differences appear when the simulations are compared to the data. The main parameters of the two scans are shown in Fig. 5.3 where a single realization of 1274 events, with a signal contribution of 15% that maximizes the likelihood analysis, is considered for the model employed in the study. It can be seen that the simulations predict a large overdensity in the the Centaurus region but it is displaced, in comparison to the one observed in data, by the GMF deflections that deviate the arrival directions of particles originating from sources in that region towards the Galactic plane as shown in Fig. 3.9b. Furthermore, the deflection patterns cause events from multiple sources to converge in the same area. For instance, events originating from M83 are deflected toward the nominal position of NGC4945, along with contributions from some events from Circinus A, leading to a larger overdensity than observed in the data in this region. This clusterization effect caused by the deflections induced by the GMF was expected since the deflected events from different sources might cover the same area inducing a clustering of events that is not observed in data.

This results in a larger local significance when the observed events are compared to the isotropic expectations. Moreover, in Fig. 5.3, it is possible to observe that when the positions of the first and second local significances obtained from simulations (star marker) are compared to those from the data (square marker), the position of the first overdensity may be recovered among the eight new models. However, no agreement is found for the second overdensity, indicating that a major structure remains present in the sky.

### Intermediate-scale structures in simulations

The primary difference between simulations and data, although, appears in the southern region of the sky: in this area, a second prominent overdensity is observed in the simulations, which is absent in the actual data. The overdensity that arises in the southern region has a similar strength compared to the one obtained throughout the first blind search. The behavior of the average values for the parameters related to the first blind search under different GMF model assumptions are illustrated in Fig. 5.4. Similarly the average values for the parameters related to the second blind search are presented in Fig. 5.5. The values obtained from simulations for both overdensities tend to be larger for the structures in the sky. As mentioned, this outcome is expected for the overdensity in the Centaurus region, as several sources have their events deflected into the same area of the sky. However, the exact value of $\sigma_1^{\rm loc}$ depends on the chosen signal strength (here 15%), as it will be further discussed in section 5.2. The second cluster, on the other hand, is influenced by different sources, depending on the model used as it will be discussed in the following section. Moreover, the second cluster is found under search window radii of $\Psi_2 \sim 15^{\circ}$. The values obtained for the second angular window that characterizes the second hotspot in the sky do not largely differ from the values related to the first overdensity indicating that the GMF deflections enhance the detection of clustered events within limited regions of the sky, generating structures that significantly deviate from isotropy, even when the highest overdensity from the first scan is excluded. This suggests that, in the case of simulations, excluding the most prominent structure does not lead to a distribution compatible with isotropy; instead, it reveals significant deviations from an isotropic event distribution. This phenomenon is consistently observed across all realizations of the GMF. The cause of this behavior can be explained by looking at the events originating from NGC253, the second brightest source in the starburst galaxy catalog when the exposure of the Pierre Auger Observatory is taken into account. These events are deflected far from the source due to the GMF and are redirected into a small

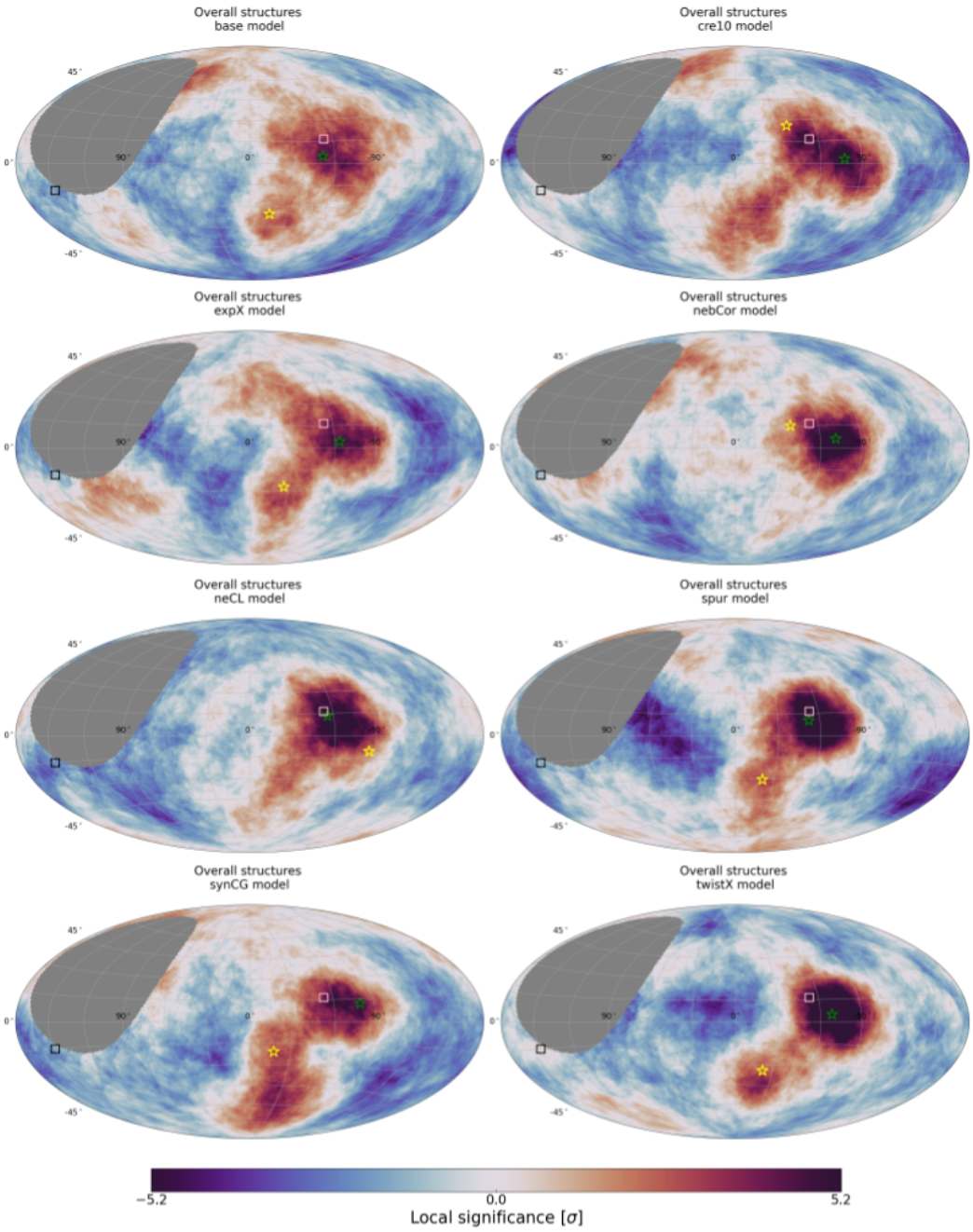



**Figure 5.3:** Overview of the sequential blind search applied on simulation. The realizations have a signal contribution of 15%, which maximizes the likelihood analysis, and deflection patterns obtained within the UF23 GMF model ensemble. The first and second overdensities are shown with green and yellow stars for the simulations and purple and black squares for the data. The top-hat smoothing angle employed is $\Psi = 24°$.

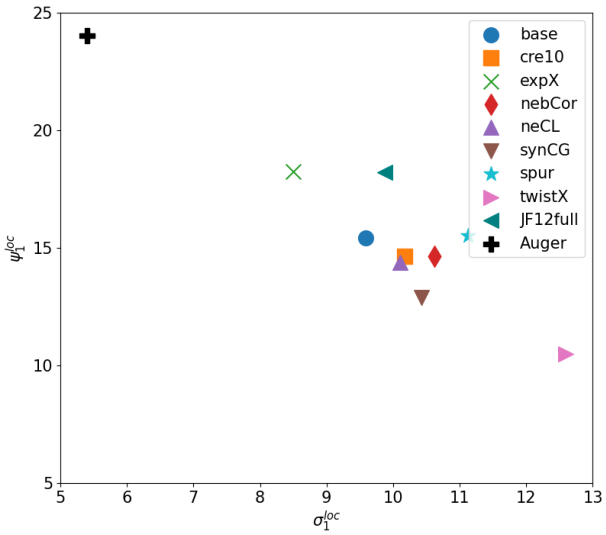



**Figure 5.4:** Average values of the parameters related to the first blind search obtained on simulations under the assumption of different GMF models. The simulated datasets have been constructed with a signal contribution from the sources in the Lunardini catalog of 15% with no EGMF component. The results of the data are displayed as a black cross.

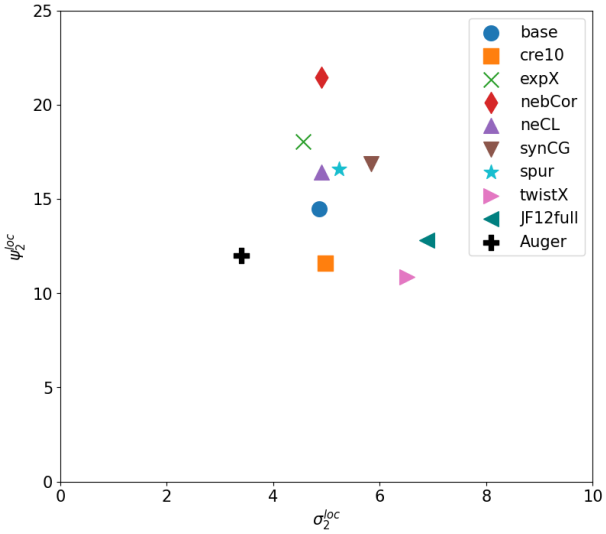



**Figure 5.5:** Average values of the parameters related to the second blind search obtained on simulations under the assumption of different GMF models. The simulated datasets have been constructed with a signal contribution from the sources in the Lunardini catalog of 15% with no EGMF component. The results on data are displayed as a black cross.

region of the sky. This deflection leads to the formation of a dense cluster of events, which dominates the sky and produces a strong deviation from isotropy, even after the removal of the highest overdensity region related to NGC4945 and M83. Models like `JF12` and `twistX`, which align in terms of flux amplification (as shown in Fig. 3.7 for the region around NGC253), place this source in a magnified area, likely contributing to the formation of narrow clustered regions in the sky. Conversely, models such as the `base` and `expX`, in which NGC253 is located in a demagnified region, still exhibit an overdensity in the southern sky not observed in the data. This may be explained by the deflection patterns of another source, NGC1068, whose events are deflected toward the direction of NGC253 and, at the lowest energies, follow the same deflection direction as the events from NGC253. When NGC253 contributes minimally to the observed flux, NGC1068's contribution rises to about 25%, making it the second brightest source in the SBG catalog and creating a prominent overdensity in the region. A visual comparison of the differences in the estimation of local significance properties is provided in Fig. 5.6. The panels illustrate how the average local significance, computed over 1000 realizations with a 15% signal contribution, is influenced by deflections calculated using the newer models. For example, one of the main differences between the choice of a GMF model can be seen when the `base` model is assumed as the GMF model: a secondary overdensity seems to be absent once the primary one is excluded from the overall sky. In contrast, under the `twistX` model, a secondary structure can be easily identified in the southern region of the sky similar to what is observed in the `JF12` model.

### 5.1.4. EGMF effects on local significances

Since simulations reveal a discrepancy with the data regarding the positions and strengths of sky overdensities, an additional smearing component has been introduced to address the overestimation of $\sigma_2^{\rm loc}$. This component accounts for potential deflections induced by the EGMF on charged particles. Simulated events are deflected by the EGMF following Eq.3.5, where the deflection is determined by the particle's rigidity and source distance. The final event position after deflection is derived using the parameter $\theta$, which sets a random position from a von Mises-Fisher distribution with $k = 1/\theta^2$.

For example, with a fixed parameter $\theta^* = 5°$, a Nitrogen nucleus at 40 EeV originating from NGC4549 would exhibit an angular dispersion of approximately 5°. In contrast, if the same particle were emitted by NGC3690 (distance $D_{\rm source} = 49$ Mpc), the deflection would follow a broader distribution with $\theta \approx 19°$. A similar analysis applies to Silicon, another dominant element in this energy range. The energy spectrum from 38 EeV to

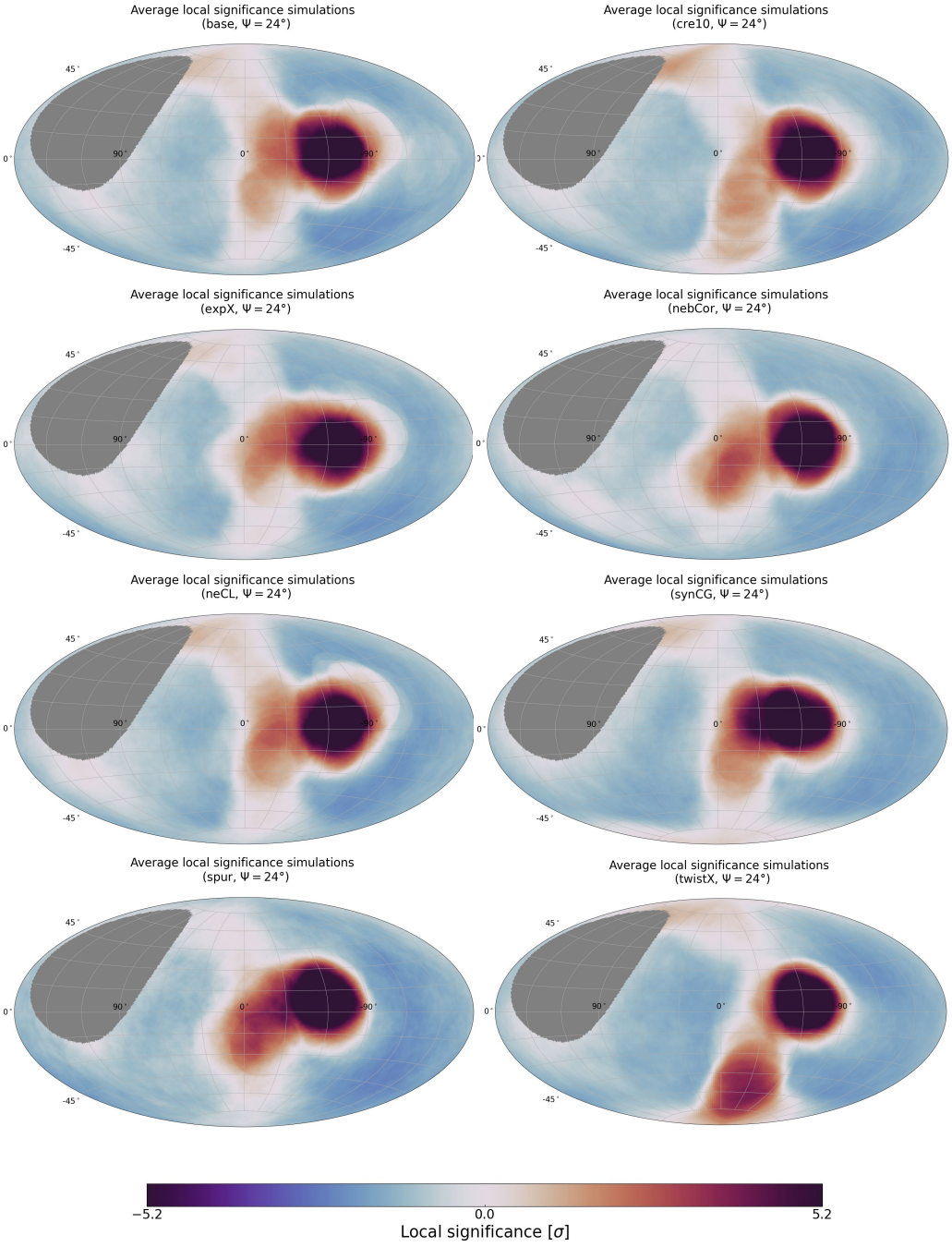



**Figure 5.6:** Average local significance obtained on simulation constructed with a 15% signal contribution under the assumption of different model for the GMF. The smearing angle used to top-hat smoothing is the same as reported by the Pierre Auger Collaboration to have an easier comparison with data ($\Psi = 24°$). The sky maps illustrate the differences between the models and how the clustering structures might be interpreted and correlated to sources.

60 EeV is dominated by Nitrogen, which gradually decreases as Silicon contributions increase above this range. For a high-energy scenario dominated by Silicon particles, an energy of 100 EeV can be used. The particle rigidity, $R^* = 100/14$, implies that an initial $\theta^* = 5°$ would correspond to a deflection angle $\theta \approx 4°$ for a source at $D^* = 3.47$ Mpc, suggesting minimal deflection at close distances for high energy particles. However, the deflection increases to around 16° for the same particle emitted from a source 49 Mpc away, indicating a significantly wider spread on the sky for faraway sources.

The simulated events experience deflection by the EGMF before they reach the edge of the Galaxy, where they are then subjected to the deflections caused by the GMF. The inclusion of this additional smearing component, simulating the influence of EGMF, adds randomness to the trajectories of charged particles as they travel through extra-galactic space. This extra component affecting the deflection is expected to reshape the clustering patterns observed in the sky.

The primary expected outcome of adding EGMF-induced deflections is a reduction in the observed strengths of local significance for clusters of events. By introducing an additional smearing component, the dispersion in the arrival directions of the events is increased. The added smearing spreads out what would otherwise be narrower clusters of events, therefore diluting their local overdensities when they arrive on Earth. As a result, the previously observed localized overdensities in the sky are softened, with individual event positions spread out by varying degrees depending on their rigidity and distance of origin.

Once the particles encounter the GMF at the Galaxy's edge, their trajectories are further modified, undergoing additional deflections based on the characteristics of the GMF. This sequential deflection process—first by the EGMF and then by the GMF—contributes to the final overall angular spread observed in the arrival directions. The overlapping effects of these magnetic fields are likely to reduce any initially concentrated clustering, making the distribution of arrival directions appear less clusterized than it would if only the GMF effects were taken into account.

In this way, the combined deflections from the EGMF and GMF increase angular dispersion, decreasing local clustering and reducing observed cosmic ray anisotropies. This adjustment aligns simulations more closely with observation by diminishing discrepancies from isotropic event distributions when regions of major overdensities are excluded.

Moreover, introducing an additional smearing component also affects the intensity of the primary sky overdensity. By increasing event dispersion, this deflection reduces clustering, spreading initially concentrated events into broader regions. Therefore, local

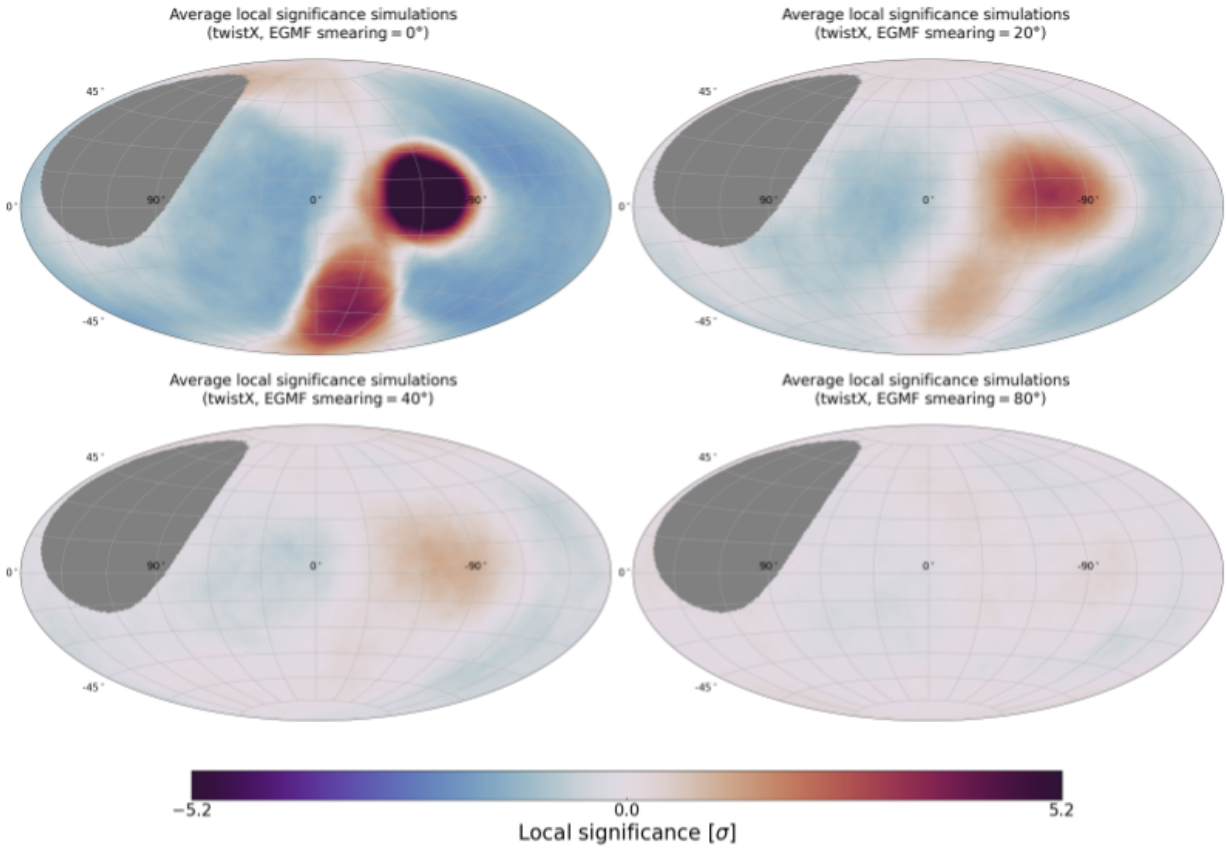



**Figure 5.7:** Effects of the EGMF smearing component on events that underwent GMF deflections, analyzed within the `twistX` model. The panels illustrate that the impact of the EGMF smearing consists on a blurring and dilution of overdensities across the sky, starting from a smearing contribution of 0° and increasing to 80°, at which point compatibility with isotropy is recovered.

significance values associated with $\sigma_1^{\text{loc}}$ are expected to decrease as smearing intensifies. Larger smearing diminishes the clustering that defines overdensities, potentially lowering local significance to isotropic levels. In extreme cases, sufficient dispersion may create a UHECR sky that is indistinguishable from isotropic event distributions.

These considerations are illustrated in Fig. 5.7, which displays the effects on simulated datasets with the introduction of the EGMF smearing component and its impact on the local significances. Following this reasoning, if the simulations behave as expected when the additional smearing component is introduced, it would be possible to set constraints on the strength of the EGMF. This can be achieved by calculating the field strength under the assumptions that a specific GMF is affecting the trajectories of charged particles and that SBG sources are responsible for contributing to the simulated signal, using Eq. 2.20

where the deflection angle is directly tied to the field strength as given by the relation:

$$\sqrt{\frac{l_c}{\mathrm{Mpc}}}\frac{B}{\mathrm{nG}} = \frac{\theta}{25^\circ}\cdot\frac{1}{Z}\cdot\frac{E}{\mathrm{EeV}}\cdot\sqrt{\frac{\mathrm{Mpc}}{d}} \tag{5.4}$$

An upper limit on the EGMF strength can be estimated based on the idea that, with a sufficiently strong EGMF, the cosmic ray events responsible for forming overdensities in the sky would be deflected over a wider area. This increased dispersion would dilute the previously observed local significance, eventually spreading the events to the point where they are indistinguishable from an isotropically distributed event scenario. Therefore, when the smearing effect of the EGMF is strong enough to eliminate observable clustering, the EGMF strength reaches a threshold beyond which no anisotropic patterns remain, setting an upper limit on the field's strength.

Similarly, a lower limit on the EGMF strength can be estimated by considering that, once the region containing $\sigma_1^{\mathrm{loc}}$ is excluded, the remaining events should have arrival directions consistent with isotropy. This observation helps define the minimum field strength required for the EGMF to sufficiently dilute the cluster of events responsible for the secondary overdensity, ensuring that the distribution of the remaining events appears to be compatible with events distributed isotropically. Thus, by determining the EGMF strength needed to erase this clustering effect, a lower bound on the field's strength can be established.

## 5.2. Constraints on the EGMF

After the qualitative comparison of the simulations and data given in the previous section, an in-depth analysis of sky overdensities, examining how they behave under the influence of various GMF models. Additionally, the analysis explores a potential correlation between the signal fraction used in the simulations and the smearing parameter that reflects the strength of the EGMF with the intention of establishing constraints for such a field's strength.

To study the behavior of $\sigma_1^{\mathrm{loc}}$ and $\sigma_2^{\mathrm{loc}}$ when an extra smearing component is applied to the trajectory of the events before introducing GMF deflections, the parameter $\theta^*$ is varied from 0° to 80° in step of 2°. Then, simulated datasets are constructed for each configuration of signal fraction, ranging from 0.1 to 0.9 with a step of 0.06.

The variation of both signal contributions to the simulated datasets and EGMF smearing might underline a correlation between these two parameters, which might influence

the behavior of the local significances in the overall sky. It is important to remember that all considerations of the behavior of the main quantities related to the local significances rely on the assumption of a precise model for the GMF.

Under the assumption of `twistX` as the model for the GMF, an overview of the behavior of both local significances is presented in Fig. 5.8. When no EGMF contribution is applied (panel 5.8a), it becomes evident that simulated scenarios with higher signal contributions show extreme values for both the primary and secondary overdensities. This outcome is expected because, when clusters of events form due to GMF deflection patterns, increasing the signal contribution in a simulated dataset will amplify these event clusters within localized regions of the sky. As a result, the distribution shows large deviations from isotropy, with pronounced overdensities forming in areas influenced by the signal. When a smearing contribution of $\theta^* = 20^\circ$ is applied (see panel 5.8b), the added randomness in the arrival directions of events becomes visible. In particular, simulations with low signal contributions start to resemble the behavior expected from isotropy, represented by the gray marker. This gray marker (and shaded gray region) is derived from applying the sequential blind search to isotropically distributed events. Since no clustering is expected in isotropic distributions, the values of both local significances should be similar, resulting in the nearly symmetric shape of the gray region. Due to the same randomness in the arrival directions, simulations with higher signal contributions also begin to resemble the observations, indicating that a correlation between the signal contribution and the smearing angle can be determined. As shown in panels 5.8c and 5.8d, when larger contributions from the EGMF are applied, simulations with low signal contributions begin to closely resemble an isotropic event distribution, where no significant anisotropies are observed across the sky. This isotropic behavior is preserved even in simulations with an almost pure signal contribution, where the added EGMF smearing effectively disperses any cluster in the arrival directions of the events, resulting in an overall sky distribution with minimal deviation from isotropy. Thus, with stronger EGMF influences, the anisotropy structures present in the sky due to the GMF effects are masked making them indistinguishable from a truly isotropic scenario.

The behavior of the second overdensity, under the same configuration previously described, underlines a similar trend. As the EGMF contribution is increased, the mean values $\langle \sigma_2^{\mathrm{loc}} \rangle$ decreases and shows agreement with an isotropic sky. This can be seen by computing the average value of the second hotspot present when the EGMF component is zero, which yields $\langle \sigma_2^{\mathrm{loc}} \rangle = 6.5\sigma$, indicating the presence of several clusters and a significant discrepancy from isotropy.

Adding the EGMF blurring to the events drastically reduces this second overdensity,

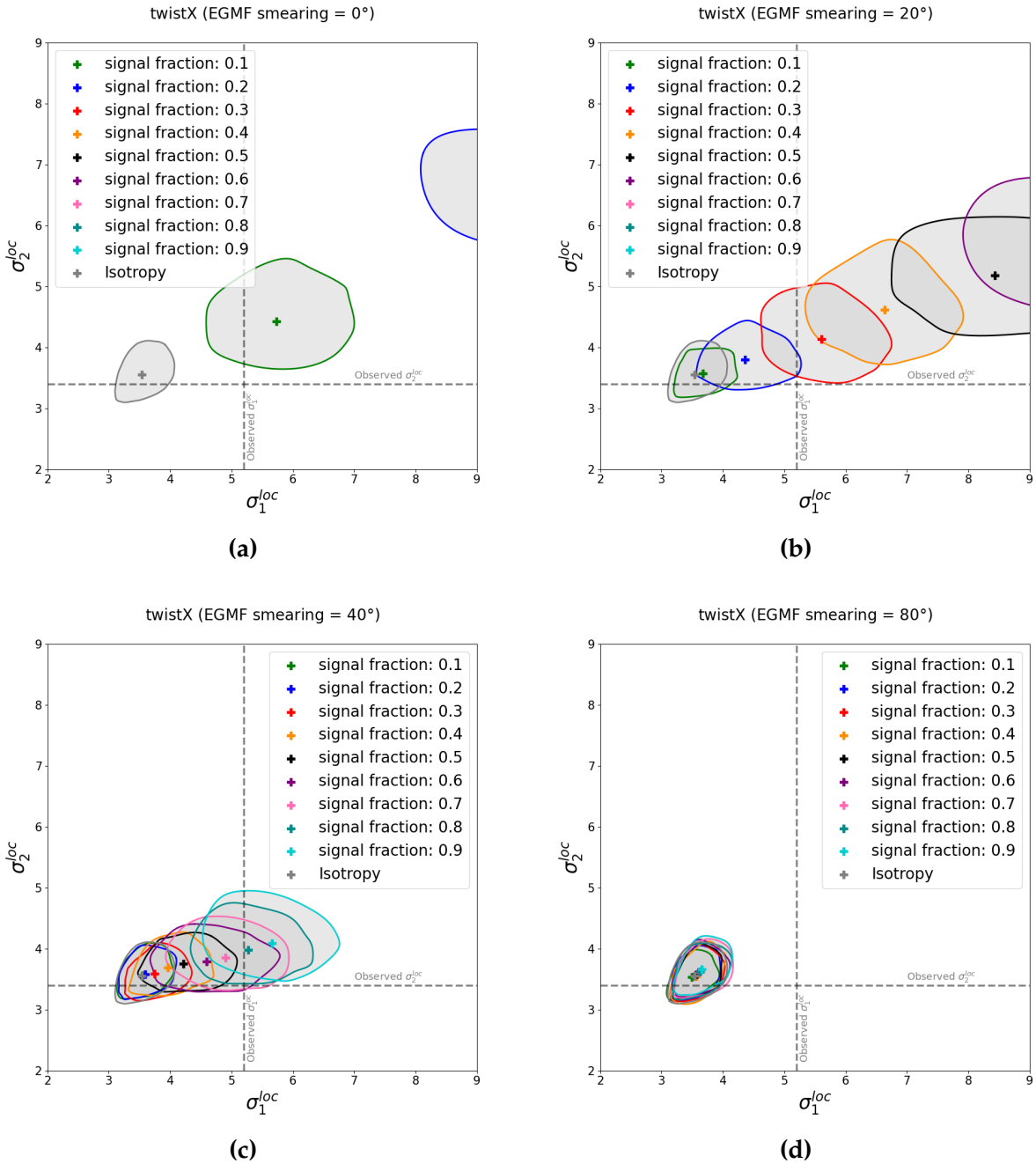



**Figure 5.8:** Behavior of the distribution center is analyzed with varying signal contributions in the simulated datasets and increasing EGMF contributions. The colored crosses represent the average values of both $\sigma_1^{\text{loc}}$ and $\sigma_2^{\text{loc}}$. The same color scheme is used to present the 68% confidence level contours, calculated using a convex hull method. In this case the simulated events are obtained assuming the `twistX` model for computing the deflections in the GMF.

bringing its value to $\langle \sigma_2^{\rm loc} \rangle = 3.7\sigma$ with a 20° smearing contribution, where the isotropic expectation is $\langle \sigma_2^{\rm loc,\,iso} \rangle = 3.39\sigma$. This indicates that agreement with the observation can also be found also for the second hotspot. A complete agreement with isotropy is observed when a contribution of 80° is introduced, yielding $\langle \sigma_2^{\rm loc} \rangle = 3.4\sigma$. Therefore, introducing the EGMF can reduce the strength of the second overdensity, which is not observed, to values compatible with the observations. It is important to note that the EGMF contribution should be strong enough to dilute the second overdensity, as this clustering of events is not observed, but it must also preserve the first overdensity at a value that remains comparable to the observations.

### 5.2.1. Angular size properties

The simulated datasets highlight major differences from the observed data in terms of $\langle \sigma_1^{\rm loc} \rangle$ and $\langle \sigma_2^{\rm loc} \rangle$. These differences also influence the top-hat search radius, which quantifies the window in which the deviation from isotropy is computed. The observed data is characterized by a large top-hat search radius required to detect a sizable deviation from isotropy. However, this behavior might differ in simulations, as they are characterized by induced clustering of events related to the GMF deflections. In fact, this is emphasized if the EGMF smearing is not taken into account, as the clustering of events is not prevented when the simulated events are deflected into one region of the sky. Similarly, if no EGMF component is introduced, even after the exclusion of the major overdensity, clusters of events will still be present in certain regions of the sky, reducing the amplitude of the search window required to observe a deviation from isotropy. These considerations can be seen in Fig. 5.9a, which shows that, on simulated datasets without EGMF smearing component, the general behavior of the angles defining the search windows is similar ($\Psi_1 \sim \Psi_2 \sim 15°$). This indicates that clustering of events can be present in the sky, resulting in deviations from isotropy that are observed even with relatively small window amplitudes. When the EGMF is introduced, as shown in Fig. 5.9b, the window amplitude required to observe a major overdensity increases to around 20°. This reflects the added randomness in the particle arrival directions, diluting clustering and indicating that a larger search windows are necessary to detect deviations from isotropy. With stronger EGMF contributions, the search window for the second overdensity decreases, suggesting that smaller angles are more effective for detecting marginal isotropy deviations. This trend is evident in Fig. 5.9c and Fig. 5.9d, where $\Psi_1$ and $\Psi_2$ correspond to progressively smaller angles, indicating an overall isotropic sky except for small regions of notable deviations.

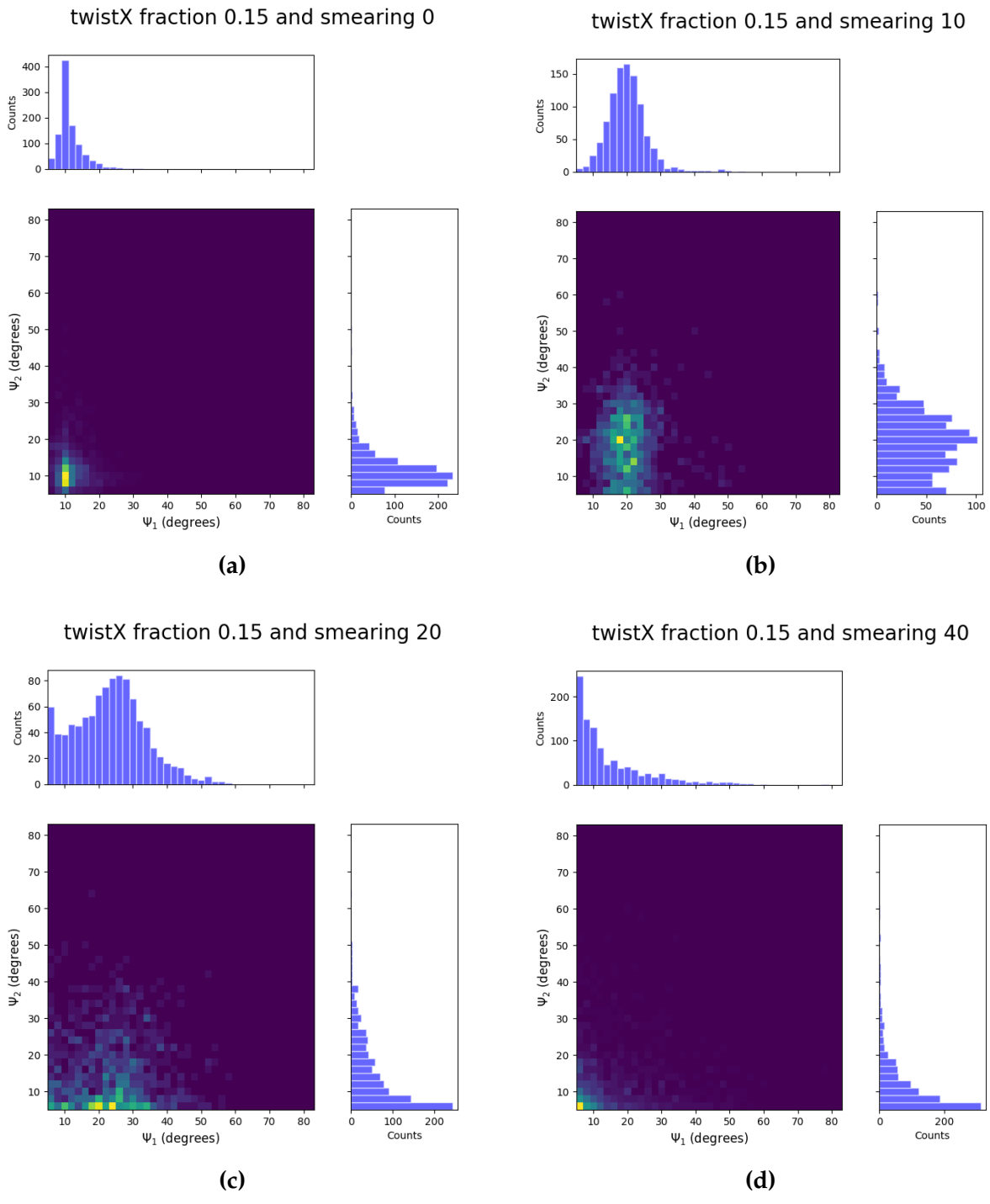



**Figure 5.9:** Behavior of the first and second top-hat search radii under different contribution of the EGMF. The number of counts corresponding to the values obtained in simulations are displayed on top side for $\Psi_1$ and on the right side for $\Psi_2$.

The search radii behavior aligns with the EGMF effects shown in Fig. 5.7 and Fig. 5.8. Without EGMF, significant clustering is observed in both blind searches, resulting in similar window amplitudes, as the remaining sky still exhibits anisotropy. Including the EGMF reduces the clustering of events and shifts the average local significance values towards values compatible with isotropy. As a consequence, progressively smaller search radii are necessary to detect deviations.

The analysis of the overdensity indicates that four main parameters influence the understanding of the properties of the UHECR sky: $\Psi_1$, $\Psi_2$, $\sigma_1^{\mathrm{loc}}$, and $\sigma_2^{\mathrm{loc}}$. These variables, which describe the variations in the effects of the EGMF, can be systematically used to quantify the strength of the EGMF smearing component, by defining an upper and lower limit for this field.

### 5.2.2. Evaluation of constraints on cosmic magnetic fields

Blind searches performed on simulations reveal the presence of significant structures in the UHECR sky that strongly deviate from isotropy. Assuming a rigidity-dependent smearing, the clustering of events can be diluted while preserving the critical features of the anisotropy.

Using the previously estimated local significances and the angular search windows that maximize these overdensities for data, limits on the strength of the EGMF can be derived, under the assumption that all other considerations on GMF, source model and composition are correct. These limits are based on the reasoning that the lower bound corresponds to the minimum angle required to dilute the clustering in regions where such features are not observed in the data, while the upper bound represents the EGMF strength above which the simulations are more isotropic than the data. To estimate these constraints, a likelihood-based method is employed. The likelihood is constructed to compare simulations affected by an additional angular smearing approximating the influence of the EGMF to the observed data. This comparison is facilitated by evaluating the free parameters obtained through systematic scans against the corresponding values derived from data. The probabilty distribution of the correlated quantities ($\sigma_1^{\mathrm{loc}}$, $\Psi_1$) and ($\sigma_2^{\mathrm{loc}}$, $\Psi_2$) alongside the values derived from data is illustrated in Fig. 5.10. The red cross marks the observed values. In this example, a good agreement between simulations and data was found assuming the configuration $f_{\mathrm{sgn}} = 0.1$ and $\theta^* = 0$ and the `base` model. This agreement indicates that a low EGMF contribution is sufficient to reproduce the observed clustering behavior. The observed consistency can be attributed to differences among models. For instance, while the `base` model does not predict a second prominent

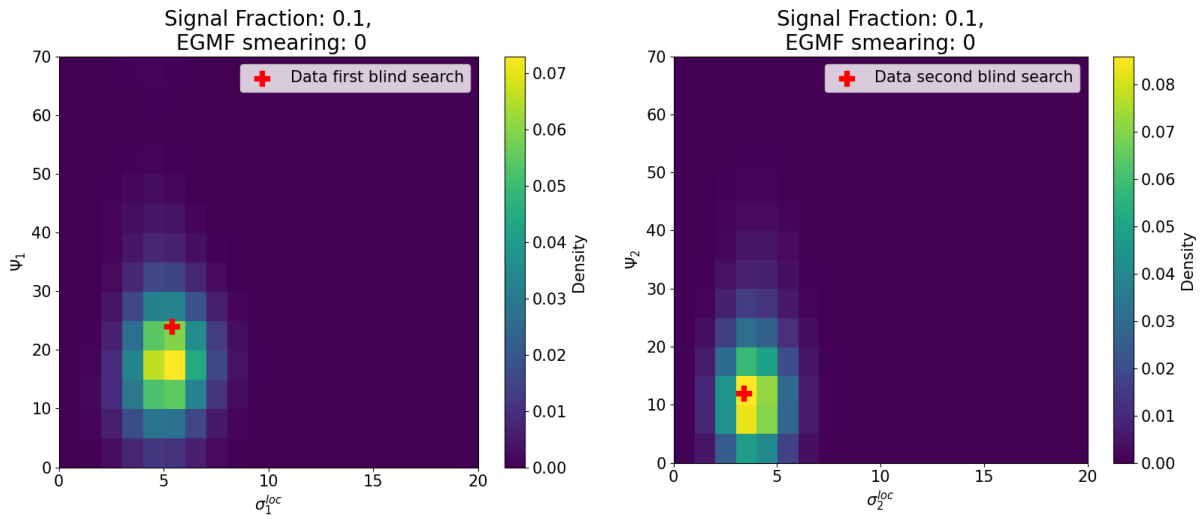



**Figure 5.10:** Behavior of the correlated quantities $(\sigma_1^{\mathrm{loc}}, \Psi_1)$ and $(\sigma_2^{\mathrm{loc}}, \Psi_2)$ compared to observed data values under the assumption of the `base` model for $f_{\mathrm{sgn}} = 0.1$ and $\theta^* = 0$. The red cross represents data values, identifying the points requiring further analysis in the likelihood framework.

overdensity, other models such as `JF12` do. As a result, configurations with minimal EGMF contributions can still align well with data. This trend is further supported by Fig. 5.8, where the mean local significances show a good agreement with observations, even at low EGMF contributions.

To provide a direct comparison, the impact of a strong EGMF contribution is examined in Fig. 5.11. In this configuration ($f_{\mathrm{sgn}} = 0.6$ and $\theta^* = 10$), the significance of both overdensities are significantly larger than those observed in data. This behavior suggests that, as expected, parameter configurations characterized by a strong signal fraction, unbalanced by sufficient EGMF contributions, are disfavored. A similar pattern emerges in Fig. 5.8, where configurations with signal fractions exceeding 20% reach a local significance below $9\sigma$ only for $\theta^* = 20°$. These findings suggest that agreement of simulations with the `base` model with data can be achieved in configurations with low EGMF contributions and low signal fractions, as well as in cases where strong signal fractions are adequately compensated by the EGMF.

To quantify the strength of the EGMF required to reproduce the observed data, the deviance between data and simulation is calculated for a wide range of combinations of $f_{\mathrm{sgn}}$ and $\theta^*$. The deviance serves as a measure to relate the probability values obtained

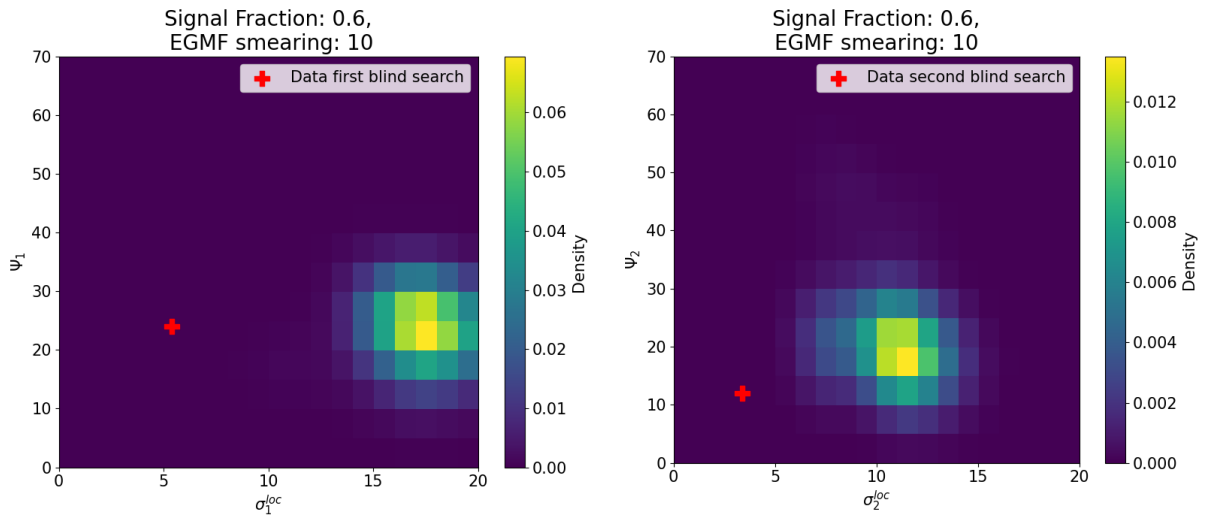



**Figure 5.11:** Behavior of the correlated quantities ($\sigma_1^{\mathrm{loc}}$, $\Psi_1$) and ($\sigma_2^{\mathrm{loc}}$, $\Psi_2$) for $f_{\mathrm{sgn}} = 0.6$ and $\theta^* = 10$ under the assumption of the `base` model. Overdensities exceed observed data values, indicating the impact of strong signal contributions without adequate EGMF compensation.

in simulations to those observed in the data. It is defined as:

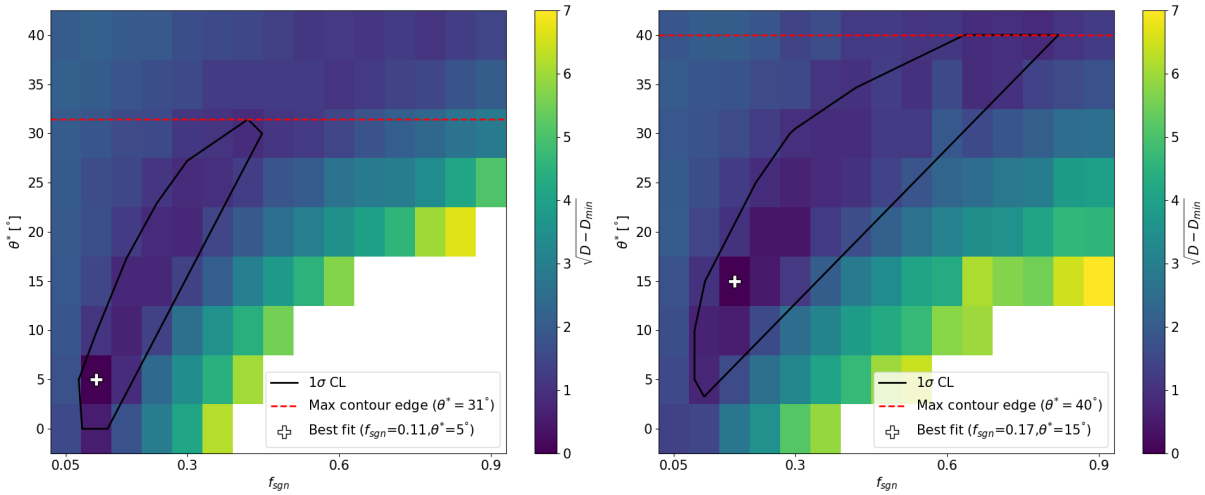
$$\sqrt{D - D_{\mathrm{min}}} = \sqrt{-2\ln(\mathcal{L}_1) - 2\ln(\mathcal{L}_2) - 2\ln(\mathcal{L}_{\mathrm{min}})}, \tag{5.5}$$

where $\ln(\mathcal{L}_1)$ and $\ln(\mathcal{L}_2)$ correspond to the log-probability density values at the positions of the red crosses shown in the left and right panels of Fig. 5.10 and Fig. 5.11, respectively, i.e. $\ln(L_i) = \ln(P(\sigma_i^{\mathrm{data}}, \psi_i^{\mathrm{data}})$, where $P$ is the joint probability density from simulation for overdensity $i$ and it is evaluated at the values observed in the data. By iterating this calculation over various combinations of signal contributions ($f_{\mathrm{sgn}}$) and EGMF-induced smearing angles ($\theta^*$), the minimal EGMF strength required to maximize agreement with the observed data can be determined. The value of $\ln(\mathcal{L}_{\mathrm{min}})$, representing the minimal log-likelihood, is subtracted to interpret the results in terms of a pseudo-standard deviation relative to the minimum.

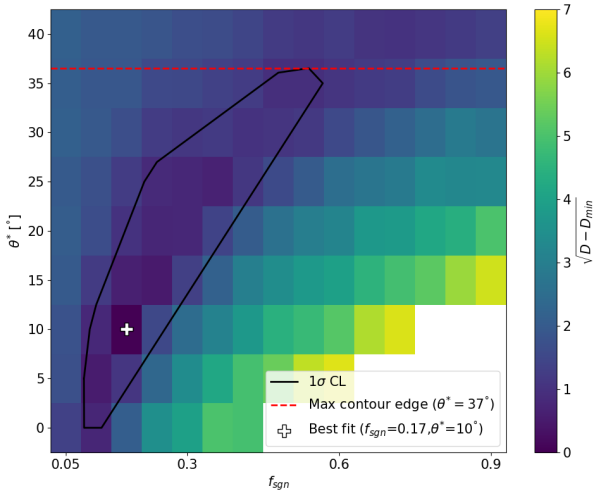
As previously discussed, adopting a different GMF model modifies the deflections of charged particles, which impacts the behavior of the best-fit minimum and the overall pseudo-standard deviation. To estimate these differences, the analysis was repeated using three models. Due to large number of realizations considered for every combination of $f_{\mathrm{sgn}}$ and $\theta^*$, the calculations are computationally very demanding. Therefore, three models (`base`, `JF12` and `nebCor`) have been considered. Given that the information

obtained throughout the maximum likelihood analysis conducted in Chap. 4 underlines similar values for the signal fraction that maximizes the agreement with the observation, the `base` model has been chosen as one sample for the UF23 ensemble. Similarly, since the `nebCor` model shows the largest signal fraction required to achieve a good agreement with the observations, they are expected to span the range of uncertainties. The results, shown in Fig. 5.12, present the values of $\sqrt{D - D_{\mathrm{min}}}$ alongside the $1\sigma$ contour level shown as black line. The color scale is saturated at $7\sigma$ and parameter combinations resulting in a worse description of the UHECR sky appear as white. This region is characterized by a high signal fraction without sufficient compensation from the EGMF. As expected, when the EGMF strength is insufficient, major structures appear in the sky, leading to substantial deviations from isotropy for large signal fraction. Beyond this point, increasing the signal contribution yields similar behavior, as the maximum deviation has already been reached. On the other hand, when a strong contribution by the EGMF is not compensated by a sufficient signal contribution, isotropy is reached and similarly as the maximum deviation, once isotropy is reached, additional smearing cannot add more isotropy to an already isotropic sky. This region is characterized by values of $\sqrt{D - D_{\mathrm{min}}}$ around $3\sigma$. This limit also represents the limit on the y-axis in Fig. 5.12.

The white cross marks the best-fit minimum, indicating the parameter values that shows the best agreement with the observations. As expected, variations in the GMF models' deflection patterns result in distinct best-fit minima, underlining the differences between the models. For the `base` model, the best-fit signal fraction is 11%, with a smearing angle of 5°. In contrast, the `JF12` model exhibits a similar signal fraction of 17% but with a larger smearing angle of 15°. This difference in the smearing angles is linked to the prominent cluster in the `JF12` model, caused by deflected events from NGC253 in the magnified southern sky region (see Fig. 3.7). To align with observations, this requires greater compensation from the EGMF. Similarly to the `JF12` model, the `nebCor` model also presents a best-fit signal fraction of 17%. However, its best-fit smearing angle is approximately 10°. This can be explained by insights from Chapter 4, where the `nebCor` model is found to be one of those requiring the largest signal contribution to achieve compatible best-fit parameters with the findings reported in [21]. As a result, a higher signal fraction for the `nebCor` model was anticipated. Unlike the `JF12` model, the `nebCor` model does not exhibit significant overdensities in the southern sky, as shown by the average local significance (see Fig. 5.6). Therefore, the best-fit smearing angle is sufficiently large to account for the increased signal fraction while remaining smaller than that of the `JF12` model.




**(a)** base model **(b)** JF12 model




**(c)** nebCor model

**Figure 5.12:** Behavior of the deviance as a function of the signal fraction ($f_{\mathrm{sgn}}$) and the EGMF smearing angle ($\theta^*$). The white region represents the largest deviations from the minimum. The best minimum is indicated with a white cross while a $1\sigma$ contour level is indicated with a solid black line. The red line indicates the $1\sigma$ bound on the smearing angle.

The best-fit minima provide insights into the EGMF's strength. In fact, a lower limit could be proposed based on the minima, but the $1\sigma$ contour reveals that configurations with strong EGMF and high signal fractions can also achieve good agreement with the observations. This emphasizes the difficulties in establishing a definitive lower limit for the EGMF strength, as different combinations of smearing angle and signal fraction can yield comparable results [69].

On the other hand, an upper limit on the EGMF strength can be derived by analyzing the properties of the $1\sigma$ contour. In Fig. 5.12, the red dashed line marks the maximum smearing angle where the deviance remains within the chosen confidence level, representing the largest deflection consistent with maintaining $\sqrt{D - D_{\rm min}}$ within acceptable limits. This upper limit is also influenced by the GMF model used. The results for the upper limits on the EGMF strength across different models are summarized in Table 5.1. These constraints on the EGMF strength is determined using the $1\sigma$ threshold level, as it

| GMF model | Best-fit $\theta^*$[°] | Max smearing angle [°] | B strength [nG$\left(l_c/1\ \mathrm{Mpc}\right)^{-1/2}$] |
|---|---|---|---|
| `base` | 5 | 31 | 3.9 |
| `JF12` | 15 | 40 | 5.1 |
| `nebCor` | 10 | 37 | 4.6 |

**Table 5.1:** Upper limits at 68% C.L. on the additional EGMF-related angular smearing compatible with the data. The limits have been computing employing Eq.5.4.

provides a clearer characterization of the parameter space. As it can be seen in Fig. 5.12, higher confidence levels, such as $3\sigma$, fail to isolate a distinct region where the deviance consistently stays below the threshold. At these higher levels, configurations with large signal fractions and strong EGMF contributions become indistinguishable from those with low signal fractions and minimal EGMF effects [88].

The results presented rely on several key assumptions, particularly those related to the GMF models and the source catalog used to determine the signal contribution. Previous studies on GMF models have demonstrated that the choice of model can significantly influence the final findings. Variations in the GMF model used to compute the deflections of charged particles result in different clustering patterns of events in the sky. These variations, in turn, affect the derived parameter spaces when computing the variables that characterize the likelihood analysis yielding different limits for the strength of the field. Another important consideration is related to the choice of the source catalog. The SBG catalog, which assigns flux weights based on 1.4 GHz radio emission,

heavily impacts the results, as powerful sources shape event distributions more than weaker ones [23,69]. These flux-weighted contributions can create distinct sky structures influenced by GMF effects [89]. Discrepancies in the location and strength of the second hotspot across GMF models suggest limitations in the SBG catalog, highlighting the need to revise flux weights or explore alternative models to better match observations.

Thus, the study illustrates that the GMF uncertainty does not allow to set a firm lower limit on the EGMF strength as claimed in [88]. Moreover, the approach conducted in the analysis reflects that the derived limits can not have a confidence level that is higher than the evidence of anisotropy itself, in contrast with [69]. A comparison to existing limits shown in Fig. 2.20 underlines that the intermediate-scale anisotropy of UHECRs can provide important constraints in the B-$l_c$ plane.

# Chapter VI

# Conclusions

This thesis focuses on the understanding of the distribution of the arrival directions of ultra-high-energy cosmic rays (UHECRs), particularly how they are influenced by deflections caused by both Galactic and Extragalactic magnetic fields (GMF and EGMF). The analysis reveals that the GMF plays a crucial role in shaping the arrival directions of UHECRs. The moderate level of anisotropy observed in the data excludes the presence of large flux contributions from individual sources, unless cosmic rays are also significantly affected by the EGMF.

A detailed simulation framework was developed to analyze the effects of the GMF on UHECRs. As a first application, this simulation framework was applied to simulate the expected signal if UHECR are produced by starburst galaxies (SBGs). The study demonstrated that even when accounting for deflections, the likelihood of the arrival directions matches the observational data well in agreement with [85]. A signal fraction, search radius, and test statistic were found to be compatible with observations, even when varying the turbulent GMF configuration within a single model. This result suggests that the Galactic variance in the turbulent component of the GMF does not significantly impact the statistical outcomes. Additionally, simulations were extended to include newer GMF ensembles, which highlighted the importance of model dependencies. The compatibility with the observational data was maintained across all tested models, although for all the GMF variations studied, it was found that the signal fraction derived from the simulated datasets underestimates the true signal fraction. This is due to the cosmic ray composition at UHE, assumed to be a mixture of Nitrogen and Silicon, as suggested by air shower data. This increase indicated that heavy elements dominating the UHECR spectrum at the energies under study were significantly deflected by

the GMF, leading to isotropically distributed events. Notably, the signal fraction varied with the choice of GMF model, emphasizing the critical influence of GMF uncertainties when interpreting the anisotropy of UHECRs.

Furthermore, the analysis explored the sky structures at intermediate-angular scales. The study demonstrated, that the significance and angular size of intermediate-scale anisotropy is highly sensitive to the EGMF strength, characterized by a rigidity and distance-dependent smearing angle. When low signal fractions are assumed in simulations, the deflections in the EGMF could sufficiently dilute any clustering, resulting in agreement with isotropically distributed events. The analysis identified correlations between the signal contribution and the EGMF strength. Larger signal contributions were found to be compensated by sufficiently strong EGMF, leading to a dilution of structures and agreement with observational data. Conversely, when the signal fraction was high and the EGMF smearing angle was low, large deviations from the isotropic distribution were observed, making such scenarios less likely.

The best agreement of the sky maps, simulated with certain signal fractions and EGMF smearing strengths, with the observations was identified through a likelihood analysis that incorporated both GMF and EGMF effects. The best-fit parameters were derived by finding the minimum deviance between data and model. An upper limit on the maximum EGMF smearing angle could be derived at 68% confidence level ($1\sigma$). Differences among GMF models were found to affect the interpretation of EGMF strength, with the largest limit found for the `JF12` model, where the 68% confidence level corresponds to $B \sim 5.1 \cdot (l_c/1\ \mathrm{Mpc})^{-1/2}$ nG. The analysis further emphasized the importance of key assumptions, such as the validity of the GMF models, the spectrum composition of UHECRs at high energies, and the choice of the SBG catalog.

Overall, the study illustrates the complex interplay between the GMF and EGMF in shaping the arrival directions of UHECRs. While valuable constraints on EGMF strength have been obtained, the analysis also underlines the challenges posed by model dependencies, statistical limitations, and assumptions about the properties of source catalogs. The statistical precision of this analysis will be further enhanced with future data from experiments like AugerPrime, which is expected to continue collecting data for the next decade. Additionally, future projects such as GCOS [90] and POEMMA [91] will provide a broader view of cosmic ray sources and magnetic fields. These advancements, coupled with datasets offering higher statistics, will offer new insights into UHECRs, significantly improving the understanding of their anisotropy and the role of magnetic fields in shaping their arrival directions.

# Bibliography


[1] A. Haungs *et al.*, "The KASCADE Cosmic-ray Data Centre KCDC: Granting Open Access to Astroparticle Physics Research Data," *Eur. Phys. J. C*, vol. 78, no. 9, p. 741, 2018.

[2] W. Apel *et al.*, "The KASCADE-Grande experiment," *Nucl. Instrum. Meth. A*, vol. 620, no. 2, p. 202, 2010.

[3] H. Sagawa, "Highlights from the Telescope Array Experiment," *SciPost Phys. Proc.*, vol. 13, p. 041, 2023.

[4] R. Abbasi *et al.*, "IceTop: The surface component of IceCube. The IceCube Collaboration," *Nucl. Instrum. Meth. A*, vol. 700, p. 188, 2013.

[5] A. Aab *et al.*, "The Pierre Auger Cosmic Ray Observatory," *Nucl. Instrum. Meth. A*, vol. 798, p. 172, 2015.

[6] S. Mollerach and E. Roulet, "Progress in high-energy cosmic ray physics," *Progress in Particle and Nuclear Physics*, vol. 98, p. 85, 2018.

[7] C. Patrignani *et al.*, "Review of particle physics," *Chinese Phys.*, vol. 40, p. 100001, 2016.

[8] J. R. Hoerandel, "Models of the knee in the energy spectrum of cosmic rays," *Astropart. Phys.*, vol. 21, p. 241, 2004.

[9] T. Abu-Zayyad *et al.*, "The Knee and the Second Knee of the Cosmic-Ray Energy Spectrum," *arXiv e-prints 1803.07052*, 2018.

[10] A. Aab *et al.*, "Measurement of the cosmic-ray energy spectrum above $2.5\times10^{18}$ eV using the Pierre Auger Observatory," *Phys. Rev. D*, vol. 102, no. 6, p. 062005, 2020.

[11] K. Greisen, "End to the cosmic ray spectrum?," *Phys. Rev. Lett.*, vol. 16, p. 748, 1966.

[12] G. T. Zatsepin and V. A. Kuzmin, "Upper limit of the spectrum of cosmic rays," *JETP Lett.*, vol. 4, p. 78, 1966.

[13] A. Aab *et al.*, "Searches for Anisotropies in the Arrival Directions of the Highest Energy Cosmic Rays Detected by the Pierre Auger Observatory," *Astrophys. J.*, vol. 804, no. 1, p. 15, 2015.

[14] A. di Matteo *et al.*, "Full-sky searches for anisotropies in UHECR arrival directions with the Pierre Auger Observatory and the Telescope Array," *PoS*, vol. ICRC2019, p. 439, 2020.

[15] A. Aab *et al.*, "Observation of a Large-scale Anisotropy in the Arrival Directions of Cosmic Rays above $8 \times 10^{18}$ eV," *Science*, vol. 357, no. 6537, p. 1266, 2017.

[16] A. A. Halim *et al.*, "Large-scale Cosmic-ray Anisotropies with 19 yr of Data from the Pierre Auger Observatory," *Astrophys. J.*, vol. 976, no. 1, p. 48, 2024.

[17] A. Abdul Halim *et al.*, "An update on the arrival direction studies made with data from the Pierre Auger Observatory," *PoS*, vol. ICRC2023, p. 252, 2023.

[18] C. Ding *et al.*, "The Imprint of Large Scale Structure on the Ultra-High-Energy Cosmic Ray Sky," *Astrophys. J. Lett.*, vol. 913, no. 1, p. L13, 2021.

[19] T. Bister and G. R. Farrar, "Constraints on UHECR Sources and Extragalactic Magnetic Fields from Directional Anisotropies," *Astrophys. J.*, vol. 966, no. 1, p. 71, 2024.

[20] D. Harari, S. Mollerach, and E. Roulet, "Anisotropies of ultrahigh energy cosmic ray nuclei diffusing from extragalactic sources," *Phys. Rev. D*, vol. 92, no. 6, p. 063014, 2015.

[21] P. Abreu *et al.*, "Arrival Directions of Cosmic Rays above 32 EeV from Phase One of the Pierre Auger Observatory," *Astrophys. J.*, vol. 935, no. 2, p. 170, 2022.

[22] J. P. Huchra *et al.*, "The 2MASS redshift survey—description and data release," *Astrophys. J.*, vol. 199, no. 2, p. 26, 2012.

[23] C. Lunardini *et al.*, "Are starburst galaxies a common source of high energy neutrinos and cosmic rays?," *JCAP*, vol. 10, p. 073, 2019.

[24] W. H. Baumgartner *et al.*, "The 70 month Swift-BAT all-sky hard X-ray survey," *Astrophys. J.*, vol. 207, no. 2, p. 19, 2013.

[25] Fermi-LAT Collaboration, "The third catalog of hard Fermi-LAT sources (3FHL)," *Astrophys. J.*, vol. 232, no. 2, p. 18, 2017.

[26] W. Heitler, *The Quantum Theory of Radiation*. Oxford, UK: Oxford University Press, 3rd ed., 1954.

[27] S. J. Sciutto, "AIRES: A system for air shower simulations," *arXiv e-prints: 9911331*, 1999.

[28] D. Heck *et al.*, "CORSIKA: A Monte Carlo code to simulate extensive air showers," 1998. Report No. FZKA-6019.

[29] S. Ostapchenko, "Monte Carlo treatment of hadronic interactions in enhanced pomeron scheme: Qgsjet-iii model," *Phys. Rev. D*, vol. 83, no. 1, p. 014018, 2011.

[30] T. Pierog and K. Werner, "EPOS LHC: A new model for high-energy hadronic interactions," *Phys. Rev. C*, vol. 92, no. 3, p. 034906, 2015.

[31] T. Pierog *et al.*, "Sibyll 2.3: A new hadronic event generator for cosmic ray air showers," *EPJ C*, vol. 75, no. 9, p. 1, 2015.

[32] P. A. Cherenkov, "Visible light from clear liquids under the action of gamma radiation," *C.R. (Dokl.) Acad. Sci. URSS*, vol. 2, p. 451, 1934.

[33] K.-H. Kampert and A. A. Watson, "Extensive air showers and ultra high-energy cosmic rays: a historical review," *EPJ H*, vol. 37, no. 3, p. 359, 2012.

[34] A. Aab *et al.*, "The Pierre Auger Observatory Upgrade - Preliminary Design Report," 2016. arXiv:1604.03637.

[35] J. Abraham *et al.*, "A Study of the Effect of Molecular and Aerosol Conditions in the Atmosphere on Air Fluorescence Measurements at the Pierre Auger Observatory," *Astropart. Phys.*, vol. 33, p. 108, 2010.

[36] J. Abraham *et al.*, "The fluorescence detector of the pierre auger observatory," *Nucl. Instrum. Meth. A*, vol. 620, no. 2–3.

[37] A. Aab *et al.*, "Reconstruction of events recorded with the surface detector of the Pierre Auger Observatory," *JINST*, vol. 15, no. 10, p. P10021, 2020.

[38] T. K. Gaisser and A. M. Hillas, "Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers," in *International Cosmic Ray Conference*, vol. 8 of *International Cosmic Ray Conference*, p. 353, 1977.

[39] A. Aab *et al.*, "Data-driven estimation of the invisible energy of cosmic ray showers with the Pierre Auger Observatory," *Phys. Rev. D*, vol. 100, no. 8, p. 082003, 2019.

[40] H. M. J. Barbosa *et al.*, "Determination of the calorimetric energy in extensive air showers," *Astropart. Phys.*, vol. 22, p. 159, 2004.

[41] C. Pérez de los Heros and T. Terzić, "Cosmic Searches for Lorentz Invariance Violation," *Lect. Notes Phys.*, vol. 1017, p. 241, 2023.

[42] M. Kachelriess, "The rise and fall of top-down models as main UHECR sources," in *20th Rencontres de Blois on Challenges in Particle Astrophysics*, p. 215, 2008.

[43] J. Abraham *et al.*, "Upper limit on the cosmic-ray photon flux above $10^{19}$ eV using the surface detector of the Pierre Auger Observatory," *Astropart. Phys.*, vol. 29, p. 243, 2008.

[44] L. O. Drury, "An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas," *Rep. Prog. Phys.*, vol. 46, p. 973, 1983.

[45] R. D. Blandford and D. Eichler, "Particle acceleration at astrophysical shocks: A theory of cosmic ray origin," *Phys. Rep.*, vol. 154, p. 1, 1987.

[46] E. Fermi, "On the origin of the cosmic radiation," *Phys. Rev.*, vol. 75, p. 1169, 1949.

[47] H. Che and G. P. Zank, "A Brief Review on Particle Acceleration in Multi-island Magnetic Reconnection," *J. Phys. Conf. Ser.*, vol. 1332, no. 1, p. 012003, 2019.

[48] A. Achterberg *et al.*, "Particle acceleration by ultrarelativistic shocks: theory and simulations," *Mon. Not. R. Astron. Soc.*, vol. 328, no. 2, p. 393, 2001.

[49] T. K. Gaisser, R. Engel, and E. Resconi, *Cosmic Rays and Particle Physics*. Cambridge, UK: Cambridge University Press, 2016.

[50] A. M. Hillas, "The Origin of Ultrahigh-Energy Cosmic Rays," *Ann. Rev. Astron. Astrophys.*, vol. 22, p. 425, 1984.

[51] R. Alves Batista *et al.*, "Open Questions in Cosmic-Ray Research at Ultrahigh Energies," *Front. Astron. Space Sci.*, vol. 6, p. 23, 2019.

[52] B. Peters, "Primary cosmic radiation and extensive air showers," *Nuovo Cim.*, vol. 22, no. 4, p. 800, 1961.

[53] G. E. Romero, A. L. Müller, and M. Roth, "Particle acceleration in the superwind of starburst galaxies. part ii: Small-scale processes," *Astron. Astrophys.*, vol. 616, p. A57, 2018.

[54] G. E. Romero, A. L. Müller, and M. Roth, "Particle acceleration in the superwinds of starburst galaxies," *Astron. Astrophys.*, vol. 616, p. A57, 2018.

[55] Planck Collaboration, "Planck 2018 results. VI. Cosmological parameters," *Astron. Astrophys.*, vol. 641, p. A6, 2020.

[56] D. Allard, "Extragalactic propagation of ultrahigh energy cosmic-rays," *Astropart. Phys.*, vol. 39, p. 33, 2012.

[57] R. C. Gilmore *et al.*, "Semi-analytic modeling of the EBL and consequences for extragalactic gamma-ray spectra," *Mon. Not. Roy. Astron. Soc.*, vol. 422, p. 3189, 2012.

[58] A. Domínguez *et al.*, "Extragalactic background light inferred from AEGIS galaxy-SED-type fractions," *Mon. Not. R. Astron. Soc.*, vol. 410, no. 4, p. 2556, 2011.

[59] F. W. Stecker, M. A. Malkan, and S. T. Scully, "Intergalactic photon spectra from the far ir to the uv lyman limit for $0 < Z < 6$ and the optical depth of the universe to high energy gamma-rays," *Astrophys. J.*, vol. 648, p. 774, 2006.

[60] R. Beck and R. Wielebinski, "Magnetic fields in galaxies," in *Planets, Stars and Stellar Systems: Volume 5: Galactic Structure and Stellar Populations* (T. D. Oswalt and G. Gilmore, eds.), p. 641, Springer, 2013.

[61] J. L. Han, "Observing interstellar and intergalactic magnetic fields," *Annu. Rev. Astron. Astrophys.*, vol. 55, p. 111, 2017.

[62] M. S. Pshirkov, P. G. Tinyakov, P. P. Kronberg, and K. J. Newton-McGee, "Deriving global structure of the galactic magnetic field from faraday rotation measures of extragalactic sources," *Astrophys. J.*, vol. 738, no. 2, p. 192, 2011.

[63] M. S. Pshirkov, P. G. Tinyakov, and F. R. Urban, "Mapping UHECRs deflections through the turbulent galactic magnetic field with the latest RM data," *Mon. Not. Roy. Astron. Soc.*, vol. 436, p. 2326, 2013.

[64] A. Korochkin, D. Semikoz, and P. Tinyakov, "The coherent magnetic field of the milky way halo, local bubble and fan region," *Astronomy Astrophysics*, 2024.

[65] R. Jansson and G. R. Farrar, "A New Model of the Galactic Magnetic Field," *Astrophys. J.*, vol. 757, p. 14, 2012.

[66] R. Jansson and G. R. Farrar, "The Galactic Magnetic Field," *Astrophys. J. Lett.*, vol. 761, p. L11, 2012.

[67] T. R. Jaffe, "Practical Modeling of Large-Scale Galactic Magnetic Fields: Status and Prospects," *Galaxies*, vol. 7, no. 2, p. 52, 2019.

[68] M. Unger and G. R. Farrar, "The Coherent Magnetic Field of the Milky Way," *Astrophys. J.*, vol. 970, no. 1, p. 95, 2024.

[69] J. D. Bray and A. M. M. Scaife, "An upper limit on the strength of the extragalactic magnetic field from ultra-high-energy cosmic-ray anisotropy," *Astrophys. J.*, vol. 861, no. 1, p. 3, 2018.

[70] R. Durrer and A. Neronov, "Cosmological Magnetic Fields: Their Generation, Evolution and Observation," *Astron. Astrophys. Rev.*, vol. 21, p. 62, 2013.

[71] D. Harari, S. Mollerach, and E. Roulet, "Signatures of galactic magnetic lensing upon ultrahigh-energy cosmic rays," *JHEP*, vol. 02, p. 035, 2000.

[72] A. Achterberg, Y. A. Gallant, C. A. Norman, and D. B. Melrose, "Intergalactic propagation of uhe cosmic rays," *arXiv:astro-ph/9907060*, 1999.

[73] A. Aab *et al.*, "Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory," *JCAP*, vol. 04, p. 038, 2017.

[74] R. Alves Batista *et al.*, "CRPropa 3.2 — an advanced framework for high-energy particle propagation in extragalactic and galactic spaces," *JCAP*, vol. 09, p. 035, 2022.

[75] E. Guido *et al.*, "Combined fit of the energy spectrum and mass composition across the ankle with the data measured at the Pierre Auger Observatory," *PoS*, vol. ICRC2021, p. 311, 2021.

[76] P. Mertsch, "Test particle simulations of cosmic rays," *Astrophys. Space Sci.*, vol. 365, no. 8, p. 135, 2020.

[77] R. A. Fisher, "Dispersion on a sphere," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 217, no. 1130, p. 295, 1953.

[78] A. Dundović and G. Sigl, "Anisotropies of Ultra-high Energy Cosmic Rays Dominated by a Single Source in the Presence of Deflections," *JCAP*, vol. 01, p. 018, 2019.

[79] H.-P. Bretz *et al.*, "Parsec: A parametrized simulation engine for ultra-high energy cosmic ray protons," *Astropart. Phys.*, vol. 54, p. 110, 2014.

[80] A. Zonca *et al.*, "Healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python," *Journal of Open Source Software*, vol. 4, no. 35, p. 1298, 2019.

[81] G. R. Farrar, N. Awal, D. Khurana, and M. Sutherland, "The Galactic Magnetic Field and UHECR Optics," *PoS*, vol. ICRC2015, p. 560, 2016.

[82] T. Bister, G. R. Farrar, and M. Unger, "The Large-scale Anisotropy and Flux (de)magnification of Ultrahigh-energy Cosmic Rays in the Galactic Magnetic Field," *Astrophys. J. Lett.*, vol. 975, no. 1, p. L21, 2024.

[83] P. Sommers, "Cosmic ray anisotropies and the location of the sun," *Astropart. Phys.*, vol. 14, no. 4, p. 271, 2001.

[84] S. S. Wilks, *Mathematical Statistics*. Wiley, New York, 1962.

[85] R. Higuchi *et al.*, "Effects of the Galactic Magnetic Field on the UHECR Correlation Studies with Starburst Galaxies," *Astrophys. J.*, vol. 949, no. 2, p. 107, 2023.

[86] R. Adam *et al.*, "Planck intermediate results.: XLII. Large-scale Galactic magnetic fields," *Astron. Astrophys.*, vol. 596, p. A103, 2016.

[87] T. P. Li and Y. Q. Ma, "Analysis methods for results in gamma-ray astronomy.," *Astrophys. J.*, vol. 272, p. 317, 1983.

[88] A. van Vliet, A. Palladino, A. Taylor, and W. Winter, "Extragalactic magnetic field constraints from ultrahigh-energy cosmic rays from local galaxies," *Mon. Not. Roy. Astron. Soc.*, vol. 510, no. 1, p. 1289, 2021.

[89] S. Mollerach and E. Roulet, "Ultrahigh energy cosmic rays from a nearby extragalactic source in the diffusive regime," *Phys. Rev. D*, vol. 99, no. 10, p. 103010, 2019.

[90] J. R. Hörandel *et al.*, "GCOS - The Global Cosmic Ray Observatory," *PoS*, vol. ICRC2021, p. 027, 2021.

[91] A. V. Olinto *et al.*, "The POEMMA (Probe of Extreme Multi-Messenger Astrophysics) observatory," *JCAP*, vol. 06, p. 007, 2021.

## Acknowledgments

I would like to express my gratitude to my supervisors, Ralph Engel, Michael Unger, and Esteban Roulet, for their guidance and support throughout my PhD. Their expertise and constructive feedback have been invaluable in shaping this work and improving my skills as a researcher. I am also thankful to Thomas Fitoussi for his valuable insights and advice during our discussions, which greatly enhanced this research. His support and perspective were particularly helpful in addressing key challenges. A special thanks to Marta for her encouragement and steady support throughout this process. Her understanding and support have been a constant source of help during this journey. During my time in Bariloche, I benefited greatly from the support of Juan Gonzales, who provided assistance both with work-related matters and in helping me adapt to the new environment. I am also grateful to the friends I made during my stay there, whose companionship made that period both productive and enjoyable. Lastly, I extend my appreciation to colleagues and collaborators who contributed to this work through discussions, shared knowledge, and support. Their input and cooperation have been an important part of this thesis.