Updates on the harmonic-space cross-correlation power spectrum analysis: towards ICRC2025

Marta Bianciotto[1,2]
Antonella Castellina[2,3]
Armando di Matteo[2]

[1] Università degli studi di Torino
[2] INFN Sezione di Torino
[3] INAF-OaTo

Auger Meeting Italia, Torino
February, 2025

UNIVERSITÀ DEGLI STUDI DI TORINO

INFN

Introduction & motivations

- The **harmonic-space cross-correlation power spectrum** (XC) could be advantageous with respect to other methods to search for correlations between UHECRs and a catalog of source candidates[1]
- The goal of our analysis is to test the XC and compare it with the **test statistics** (TS)[2] and the **auto-correlation** (AC, also known as angular power spectrum, does not require a catalog)[3]

[1] Urban et al. (2021); Tanidis et al. (2022); Tanidis et al. (2023); Urban et al. (2024)
[2] $\max_{\psi,f} \mathrm{TS}(\psi, f, E_{\min} = 32\ \mathrm{EeV})$ as in Auger + TA (UHECR 2022)
[3] $C_\ell := \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$

Introduction & motivations

Every function $\Phi(\hat{n})$ over the celestial sphere ($\hat{n} = (\alpha, \delta)$) can be expressed in terms of **spherical harmonics**

$$\Phi(\hat{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n})$$

as a function of **harmonic coefficients**

$$a_{\ell m} = \int_{4\pi} Y_{\ell m}(\hat{n})^* \Phi(\hat{n})\, d\Omega$$

where $d\Omega = d\alpha\, d \sin\delta$
and ℓ is the degree of anisotropy over angular scale $\sim 180^\circ/\ell$.

In case of full sky coverage, $a_{\ell m} = \sum_{i(\text{ev})} \frac{Y_{\ell m}(\hat{n}_i)^*}{\omega(\hat{n}_i)}$, where $\omega(\hat{n}_i)$ is the weight for each event

Introduction & motivations

The cross-correlation method is based on the following concepts:

UHECR flux

$$\Phi(\hat{n})^{\mathrm{CR}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n})$$

Galaxy density/flux

$$\Phi(\hat{n})^{\mathrm{GAL}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} b_{\ell m} Y_{\ell m}(\hat{n})$$

$\longrightarrow$ if UHECRs come from galaxies, $a_{\ell m} \propto b_{\ell m}$

$\longrightarrow$ harmonic-space cross-correlation power spectrum S_ℓ[4]

$$\boxed{S_\ell := \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a^*_{\ell m} b_{\ell m}}$$

[4] Urban et al. (2021); Tanidis et al. (2022); Tanidis et al. (2023); Urban et al. (2024)

Catalogs of data & sources

After having extensively studied the cross-correlation method on full-sky simulations and public datasets (see GAP-2024-030), we apply it on the **new full-sky dataset** employed in the contribution for the UHECR 2024 Symposium:

→ **Auger dataset**: events with $E \geq 32$ EeV detected until December 2022

→ **TA dataset**: $E \geq 39.96$ EeV detected until May 2024

And we test the correlation between the data and two **source catalogs**:

1. **Lunardini catalog**: a catalog of nearby galaxies with a high star formation rate, denoted as starburst galaxies (SBGs)
2. **2MASS catalog**: a catalog based on the Two Micron All Sky Survey (2MASS), which considers all IR galaxies

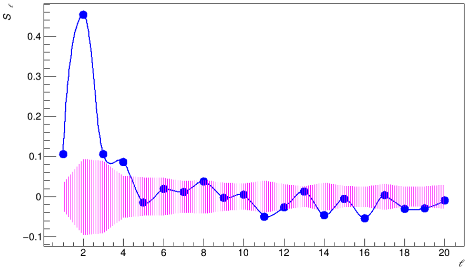
---

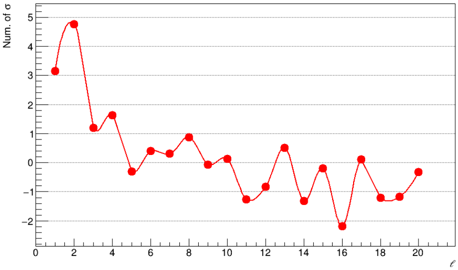
Local Group galaxies (D < 1 Mpc) are excluded from the source catalogs

Weights are applied to galaxy fluxes, with attenuations based on Auger combined fit (EPOS LHC 1st minimum) as in Auger ApJ 2022

Results for the cross-correlation

- Fixed energy threshold (48 EeV)
- **Cross-correlation with Lunardini, comparison with isotropic band:**



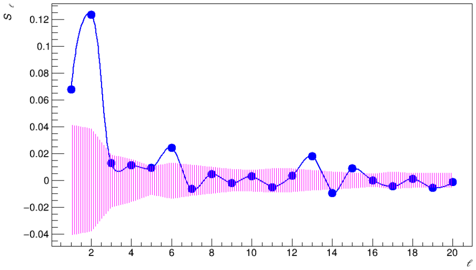


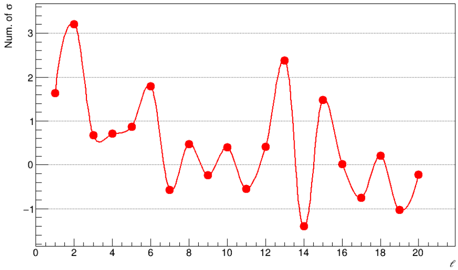
$$\hookrightarrow \boxed{S_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a^*_{\ell m} \cdot b_{\ell m}}$$

$$\hookrightarrow \boxed{N_\sigma = \frac{S_\ell - \langle S_\ell \rangle_{\text{iso}}}{\sigma(S_\ell)_{\text{iso}}}} \; \langle S_\ell \rangle_{\text{iso}} = 0$$

Results for the cross-correlation

- Fixed energy threshold (37 EeV)
- **Cross-correlation with 2MASS, comparison with isotropic band:**



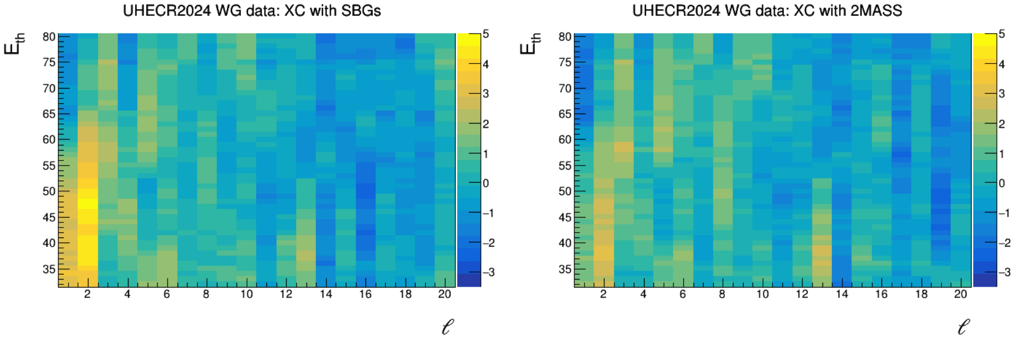


$\hookrightarrow$ $$S_\ell = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a^*_{\ell m}\cdot b_{\ell m}$$

$\hookrightarrow$ $$N_\sigma = \frac{S_\ell - \langle S_\ell\rangle_{\text{iso}}}{\sigma(S_\ell)_{\text{iso}}} \quad \langle S_\ell\rangle_{\text{iso}} = 0$$

Results for the cross-correlation

- Scan in energy thresholds 32-80 EeV
- **Cross-correlation with Lunardini (left) and 2MASS (right):**



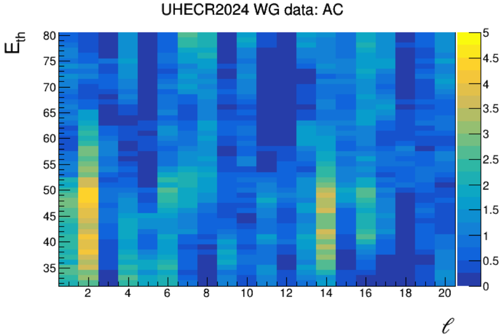
Maximum significance:

Cross-correlation, Lunardini pre-trial = 4.8σ ($\ell = 2$, $E_{\rm th} = 48$ EeV), post-trial (1-tailed) = 3.4σ

Cross-correlation, 2MASS pre-trial = 3.2σ ($\ell = 2$, $E_{\rm th} = 37$ EeV), post-trial (2-tailed) = 1.0σ

Comparison with the auto-correlation

- Scan in energy thresholds 32-80 EeV
- **Auto-correlation:**



Maximum significance:

Auto-correlation pre-trial $= 4.4\sigma$ ($\ell = 2$, $E_{\rm th} = 48$ EeV), post-trial (2-tailed) $= 2.8\sigma$

Combining auto- and cross-correlation power spectra

Motivations

- According to Urban et al. (2023), combining the auto- and cross-correlation power spectra could lead to reach detection levels of 3σ or more for single multipoles at the largest scales

Theory

- C_ℓ and the S_ℓ are not independent measurements, so their covariance matrix $\mathcal{M}$ has to be taken into account:

$$\mathcal{M} = \begin{pmatrix} \mathrm{Cov}(C_\ell, C_\ell) & \mathrm{Cov}(S_\ell, C_\ell) \\ \mathrm{Cov}(C_\ell, S_\ell) & \mathrm{Cov}(S_\ell, S_\ell) \end{pmatrix}$$

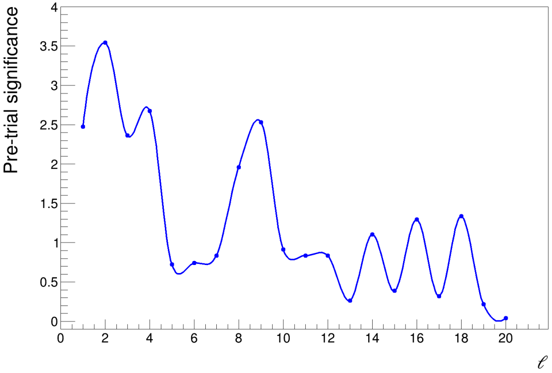
- The combined significance can then be calculated as:

$$\mathcal{S} = \sum_\ell \begin{bmatrix} C_\ell - \langle C_\ell \rangle_{\mathrm{iso}} \\ S_\ell - \langle S_\ell \rangle_{\mathrm{iso}} \end{bmatrix}^T \mathcal{M}^{-1} \begin{bmatrix} C_\ell - \langle C_\ell \rangle_{\mathrm{iso}} \\ S_\ell - \langle S_\ell \rangle_{\mathrm{iso}} \end{bmatrix}$$

where $\mathcal{S}$ follows a χ^2 distribution with 2 degrees of freedom

Results of the combined analysis

- **Dataset:** Public, Auger (ApJ 2022) + TA(ApJL 2014)
- **Catalog:** Lunardini



Maximum significance:

Auto- and cross-correlation pre-trial $= 3.6\sigma$ ($\ell = 2$, $E_{\rm th} = 47$ EeV), penalized $= 2.9\sigma$

Summary of results

Dataset	**Method**	**Post-trial σ**	ℓ	$\mathbf{E_{th}}$ **[EeV]**
UHECR 2024	XC (Lunardini)	3.4	2	48
	XC (2MASS)	1.0	2	37
	AC	2.8	2	48
ICRC 2023	XC (Lunardini)	3.2	2	47
	XC (2MASS)	1.1	2	38
	AC	2.3	2	47
Public dataset	XC (Lunardini)	1.2	2	47
	XC (2MASS)	< 1	2	44
	AC	1.6	2	47
	XC (Lunardini) + AC	2.9	2	47

Conclusions & discussion

Conclusions

- In general, the cross-correlation method appears to be more sensitive than the auto-correlation and slightly less sensitive than the test statistics ($\sim 4.5\sigma$)
- It could be interesting to study the cross-correlation method, maybe in combination with other methods, such the test statistics, and include it in future analysis
- To keep in mind: results shown in Urban et al. (2021); Tanidis et al. (2022); Tanidis et al. (2023); Urban et al. (2024) are not so relevant for our case scenario (SBGs), since their assumptions are different (energy threshold, composition, GMF...)

What's next?

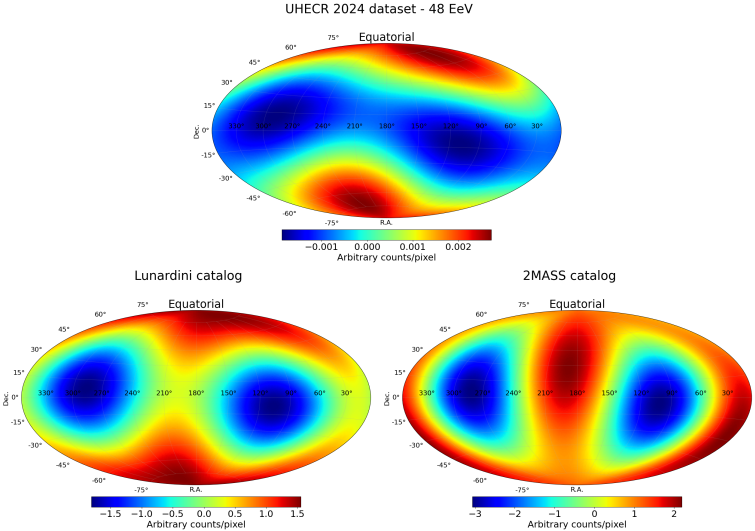
- The results will be presented at ICRC2025 in the context of the Auger-TA working group on arrival directions
- Apply to new dataset the analysis combining the auto- and cross-correlation power spectra, which looks promising

Thank you!

Backup slides

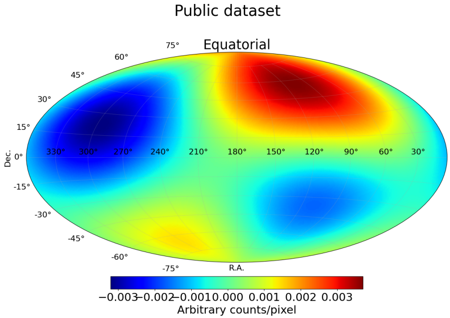
Results for the cross-correlation

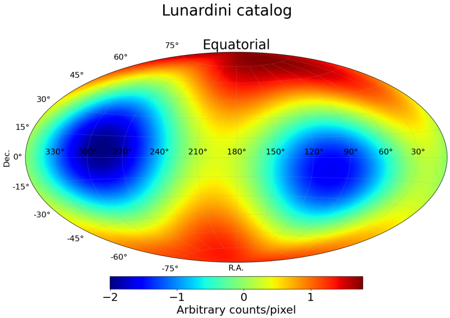
Quadrupole ($\ell = 2$) visual representation:



Combining dipole & quadrupole of the cross-correlation

- **Sum of S_ℓ ($\ell = 1, 2$), fixed energy threshold (47 EeV)**
- **Dataset:** Public, Auger (ApJ 2022) + TA(ApJL 2014)
- **Catalog:** Lunardini





Maximum significance:

Dipole & quadrupole pre-trial $= 4.0\sigma$, post-trial (1-tailed) $= 3.0\sigma$

Introduction

Harmonic space

Every function $\Phi(\hat{n})$ over the celestial sphere $(\hat{n} = (\alpha, \delta))$ can be expressed in spherical harmonics

$$\Phi(\hat{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n})$$

as a function of coefficients

$$a_{\ell m} = \int_{4\pi} Y_{\ell m}(\hat{n})^* \Phi(\hat{n}) \, d\Omega$$

where $d\Omega = d\alpha \, d \sin \delta$
and ℓ is the degree of anisotropy over angular scale $\sim 180^\circ/\ell$.

N.B. If we have full sky coverage $a_{\ell m} = \sum_{i(\mathrm{ev})} \frac{Y_{\ell m}(\hat{n}_i)^*}{\omega(\hat{n}_i)}$, where $\omega(\hat{n}_i)$ is the weight for each event → Auger + TA

Motivations

Why harmonic space?

- Spherical harmonics are a convenient basis for an expansion on a sphere because they are orthonormal and linearly independent
- The angular power spectrum describes the angular scales ($180°/\ell$) of anisotropy in a rotationally invariant way and multipole moments of each order (ℓ) are separated by the same angular distance
- Dipole ($\ell = 1$) and quadrupole ($\ell = 2$) amplitudes are not much affected by coherent magnetic deflections and are attenuated by turbulent magnetic deflections only by a factor $e^{\frac{-\ell^2 \theta_{\mathrm{turb}}^2}{2}}$

Why cross-correlation?

- Cross-correlation (XC) is more sensitive to small-scale angular anisotropies than the standard auto-correlation (AC) used in previous works
- XC has higher S/N_ℓ ratio than AC if optimal weights are used

AC & TS

Auto-correlation

The angular power spectrum is the average $a_{\ell m}^2$ as a function of ℓ

$$C_\ell := \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} |a_{\ell m}|^2$$

- The power in mode ℓ is sensitive to variations over angular scales of $180°/\ell$
- C_ℓ provides a quick and sensitive method to test for anisotropy and to determine its magnitude and characteristic angular scale(s)

This method is called AC (auto-correlation) because only a_{lm} from CRs data are considered

2022 test statistics

The TS used in Auger + TA (UHECR 2022) is the following

$$\mathrm{TS}(\psi, f, E_{\mathrm{min}}) = 2 \ln \frac{L(\psi, f, E_{\mathrm{min}})}{L(\psi, 0, E_{\mathrm{min}})}$$

where ψ is the angle, f is the signal fraction and E_{min} is the energy threshold

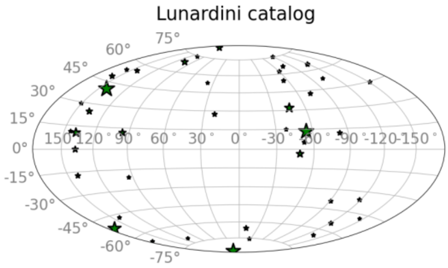
If a single E_{min} is considered the TS is a χ^2 with 2 degrees of freedom and the p-value is

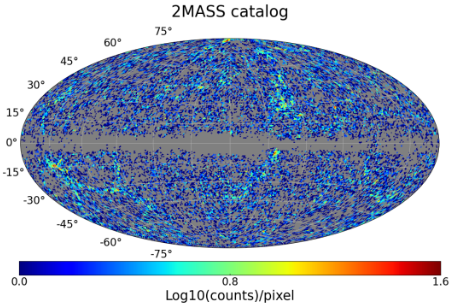
$$p = \frac{e^{\frac{-\mathrm{TS}}{2}}}{2}$$

which can be easily converted in number of σ

Catalogs

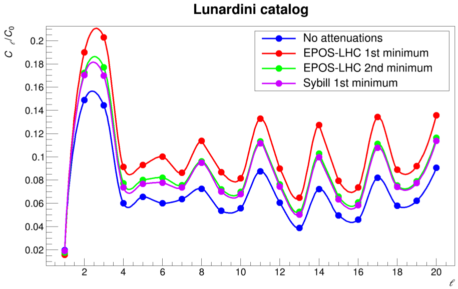
Source catalogs

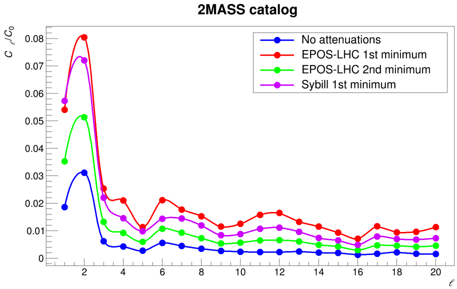




AC with catalogs

- $C_\ell = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} |b_{\ell m}|^2$, where $b_{\ell m}$ refers to the two catalogs





- For multipoles greater than $\ell = 3$ for the Lunardini catalog and $\ell = 2$ for the 2MASS catalog it is difficult to detect anisotropies even if there is a positive cross-correlation between data and catalog

Datasets

Energy cross-calibration

Cross-calibration of energies using events arriving in the part of the sky visible to both:

$$E_{\mathrm{Auger}} = E_0 \, e^{\alpha} \left(\frac{E_{\mathrm{TA}}}{E_0} \right)^{\beta}$$

where $E_0 = 10\,\mathrm{EeV}$, $\alpha = 0.157$ and $\beta = 0.949$

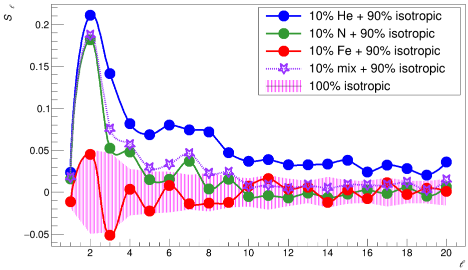
Comparison between public & ICRC2023 WG datasets

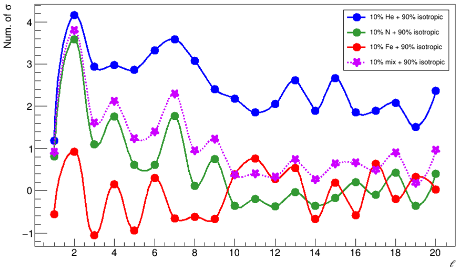
Dataset	**Experiment**	**Number of events**	E_{th} **[EeV]**
UHECR 2024 WG data	Auger	2936	32
	TA	XXX	40.2
ICRC 2023 WG data	Auger	2936	32
	TA	404	40.2
Public data	Auger	2635 (2040)	32 (44.58)
	TA	72	57

Analysis on simulations

Analysis on simulations

- **Composition**: 10% pure compositions (He, N, Fe) + 90% isotropic background & 10% mixed (28.60% He + 69.05% N + 2.35% Fe) + 90% isotropic background
- **Catalog**: Simulated with Lunardini, **reconstructed with Lunardini**
- **GMF model**: JF2012 regular



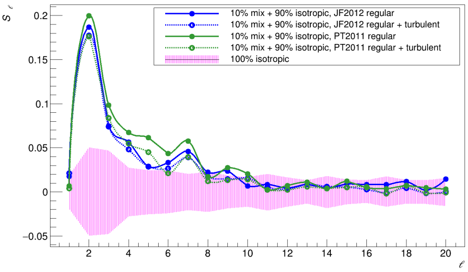


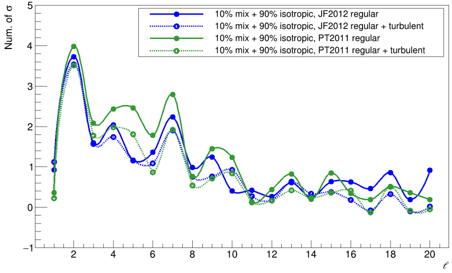
$\hookrightarrow$ $$S_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a^*_{\ell m} \cdot b_{\ell m}$$

$\hookrightarrow$ $$N_\sigma = \frac{S_\ell - \langle S_\ell \rangle_{\text{iso}}}{\sigma(S_\ell)_{\text{iso}}}$$ $\langle S_\ell \rangle_{\text{iso}} = 0$

Analysis on simulations - Regular *vs* turbulent GMF

- **Composition**: Mixed composition
- **Catalog**: Simulated with Lunardini, **reconstructed with Lunardini**
- **GMF model**: All models





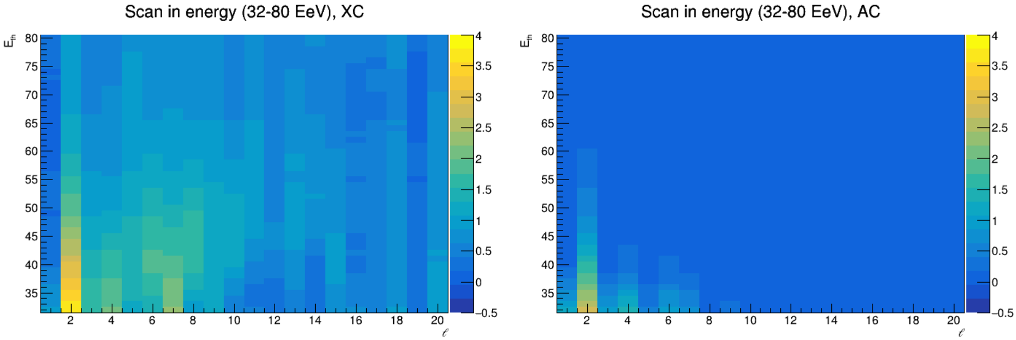
$$\hookrightarrow \boxed{S_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a^*_{\ell m} \cdot b_{\ell m}}$$

$$\hookrightarrow \boxed{N_\sigma = \frac{S_\ell - \langle S_\ell \rangle_{\rm iso}}{\sigma(S_\ell)_{\rm iso}}} \quad \langle S_\ell \rangle_{\rm iso} = 0$$

Analysis on simulations - Regular GMF

Cross-correlation with Lunardini catalog (left) and autocorrelation (right):

- **Composition**: Mixed composition
- **Catalog**: Simulated with Lunardini, **reconstructed with Lunardini**



Maximum significance:

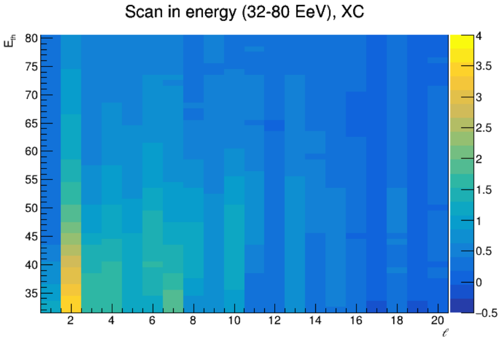
Cross-correlation pre-trial = 3.7σ ($\ell = 2$, $E_{\mathrm{th}} = 32$ EeV), post-trial (1-tailed) = 1.9σ

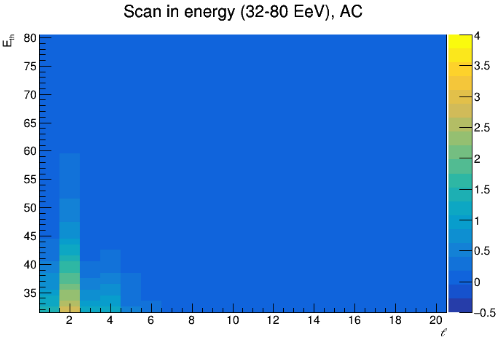
Auto-correlation pre-trial = 2.7σ ($\ell = 2$, $E_{\mathrm{th}} = 32$ EeV), post-trial (2-tailed) $< 1\sigma$

Analysis on simulations - Regular & turbulent GMF

Cross-correlation with Lunardini catalog (left) and autocorrelation (right):

- **Composition**: Mixed composition
- **Catalog**: Simulated with Lunardini, **reconstructed with Lunardini**





Maximum significance:

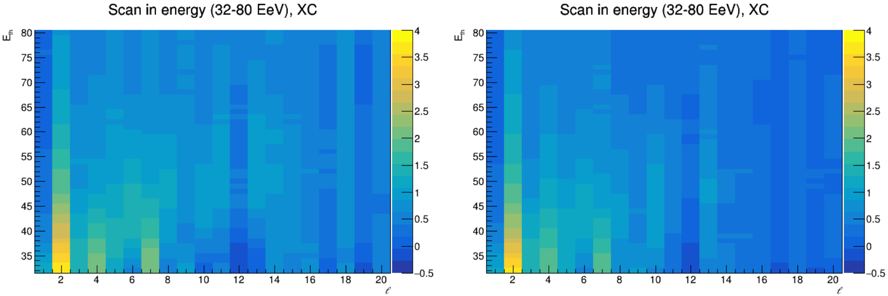
Cross-correlation pre-trial $= 3.5\sigma$ ($\ell = 2$, $E_{\rm th} = 32$ EeV), post-trial (1-tailed) $= 1.6\sigma$

Auto-correlation pre-trial $= 2.7\sigma$ ($\ell = 2$, $E_{\rm th} = 32$ EeV), post-trial (2-tailed) $< 1\sigma$

Analysis on simulations - 2MASS catalog

Cross-correlation with 2MASS catalog; regular GMF (left) and regular & turbulent GMF (right):

- **Composition**: mixed composition
- **Catalog**: simulated with Lunardini & reconstructed with 2MASS



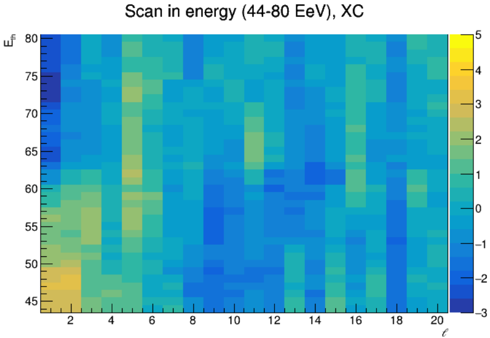
Maximum significance:

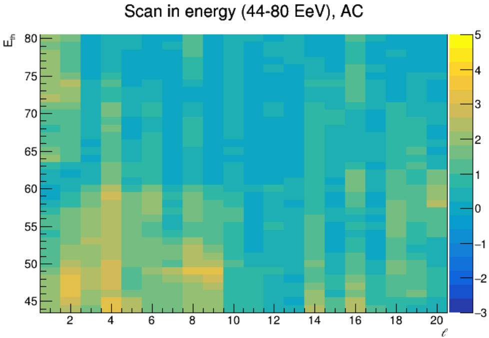
- Regular GMF: pre-trial $= 3.7\sigma$, post-trial (1-tailed) 2.0σ ($\ell = 2$, $E_{\rm th} = 32$ EeV)
- Regular & turbulent GMF: pre-trial $= 3.6\sigma$, post-trial (1-tailed) 1.7σ ($\ell = 2$, $E_{\rm th} = 32$ EeV)

Analysis with the public dataset

Analysis on public data

Cross-correlation with Lunardini catalog (left) and autocorrelation (right):





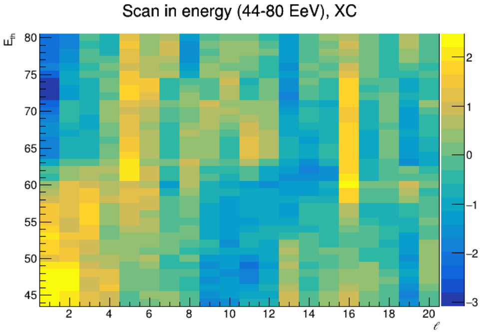
Maximum significance:

Cross-correlation pre-trial $= 3.2\sigma$ ($\ell = 2$, $E_{\rm th} = 47$ EeV), post-trial (1-tailed) $= 1.2\sigma$

Auto-correlation pre-trial $= 3.3\sigma$ ($\ell = 2$, $E_{\rm th} = 47$ EeV), post-trial (2-tailed) $= 1.6\sigma$

Analysis on public data - 2MASS catalog

Public dataset, reconstructed with 2MASS:



Maximum significance:

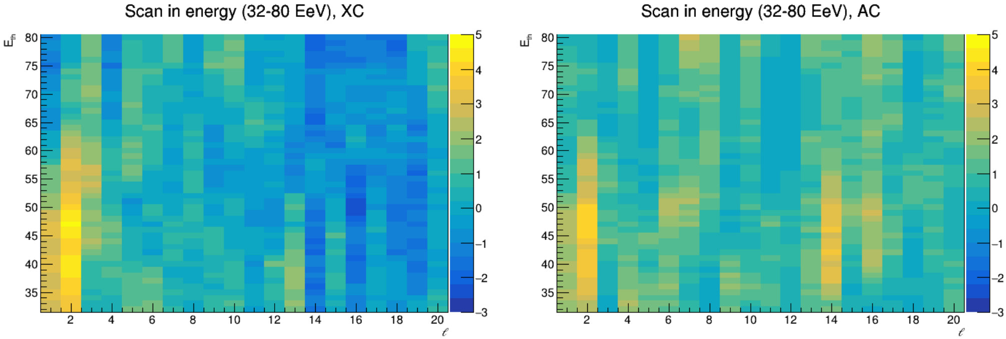
- Cross-correlation: pre-trial $= 2.5\sigma$, post-trial (1-tailed) $< 1\sigma$ ($\ell = 1$, $E_{\text{th}} = 44$ EeV)

Uncertainties in calibration (affecting $\ell = 1$ and $\ell = 2$) not taken into account

Analysis with the ICRC2023 dataset

Analysis on ICRC2023 WG data

Cross-correlation with Lunardini catalog (left) and autocorrelation (right):



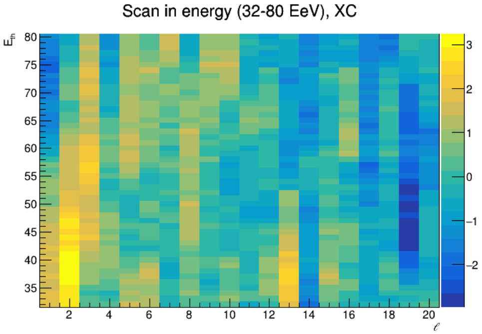
Maximum significance:

Cross-correlation pre-trial $= 4.6\sigma$ ($\ell = 2$, $E_{\rm th} = 47$ EeV), post-trial (1-tailed) $= 3.2\sigma$

Auto-correlation pre-trial $= 4.0\sigma$ ($\ell = 2$, $E_{\rm th} = 47$ EeV), post-trial (2-tailed) $= 2.3\sigma$

Analysis on ICRC2023 working group data - 2MASS catalog

ICRC 2023 dataset, reconstructed with 2MASS:



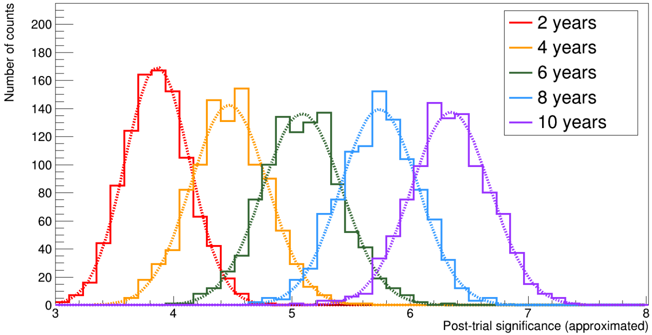
Maximum significance:

- Cross-correlation: pre-trial = 3.2σ, post-trial (1-tailed) = 1.1σ ($\ell = 2$, $E_{\text{th}} = 38$ EeV)

How many more years of TA?

How many more years of TA to reach observation level?

For the cross-correlation method, the post-trial significance for 2/4/6/8/10 more years of TA data:

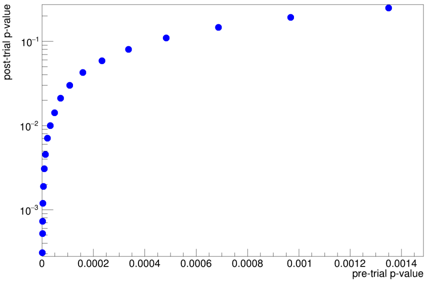


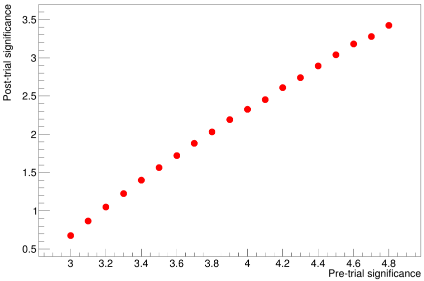
→ With ~ 6 more years of TA, we have the 50% of probability to observe 5σ post-trial using the cross-correlation method

How many more years of TA to reach observation level?

Cross-correlation

- Pre-trial to post-trial conversion was done with an extrapolation:





$$\sigma_{\text{post-trial}} = 13.39 \cdot \log_{10}(\sigma_{\text{pre-trial}}) + 1.51$$