

Constraining properties of cosmic ray sources across the energy spectrum

Igor Vaiman

Advisor: Carmelo Evoli

In collaboration with: Philipp Mertsch¹, Michael Unger²

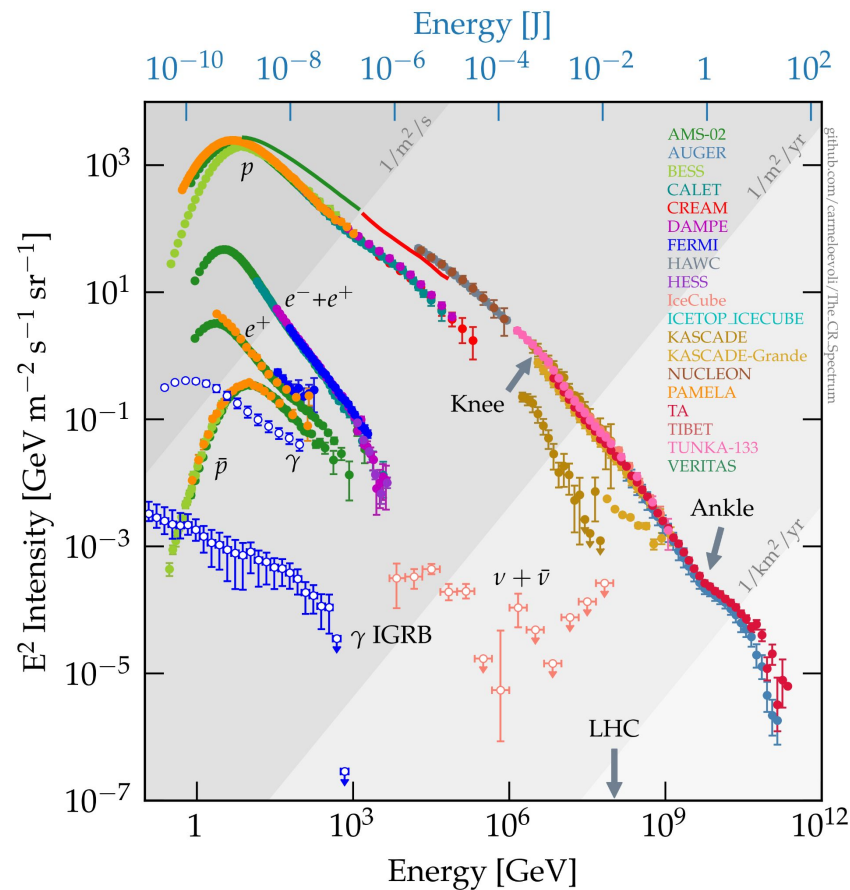


¹ Institute for Theoretical Particle Physics and Cosmology, RWTH Aachen University, Germany

² Institute for Astroparticle Physics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

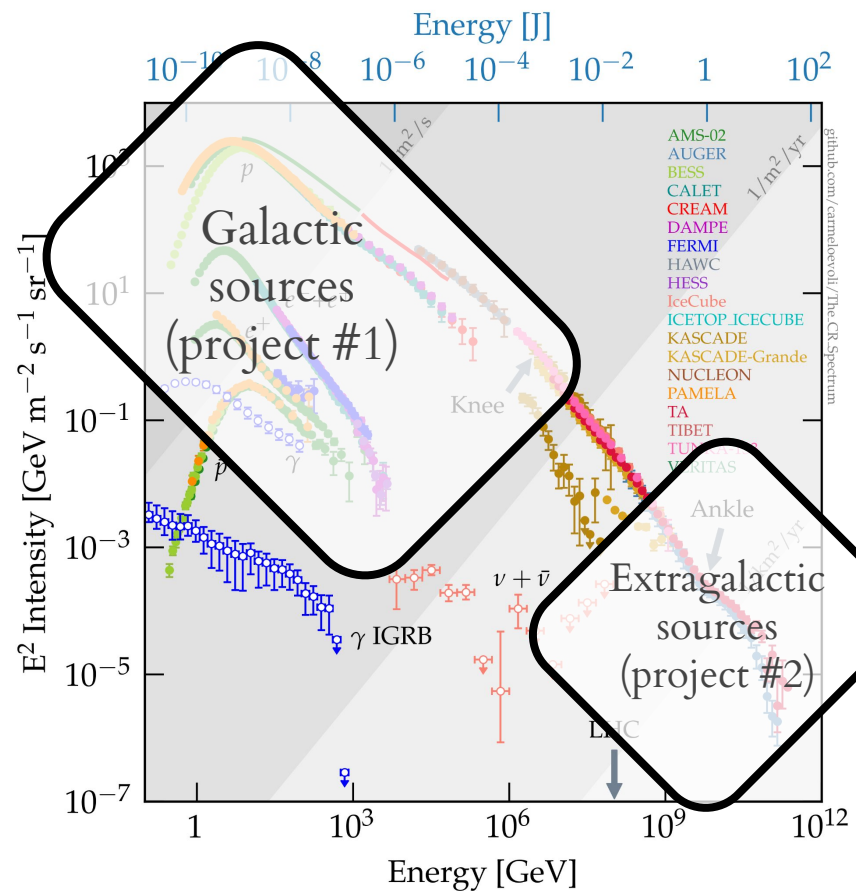
Context

- Cosmic rays are observed for 100+ years, yet their sources and propagation are not fully understood
- As charged particles, they experience multiple types of losses and are deflected by magnetic fields → the inverse problem is highly degenerate
- Multimessenger observations might provide additional constraints



Context

- Cosmic rays are observed for 100+ years, yet their sources and propagation are not fully understood
- As charged particles, they experience multiple types of losses and are deflected by magnetic fields → the inverse problem is highly degenerate
- Multimessenger observations might provide additional constraints

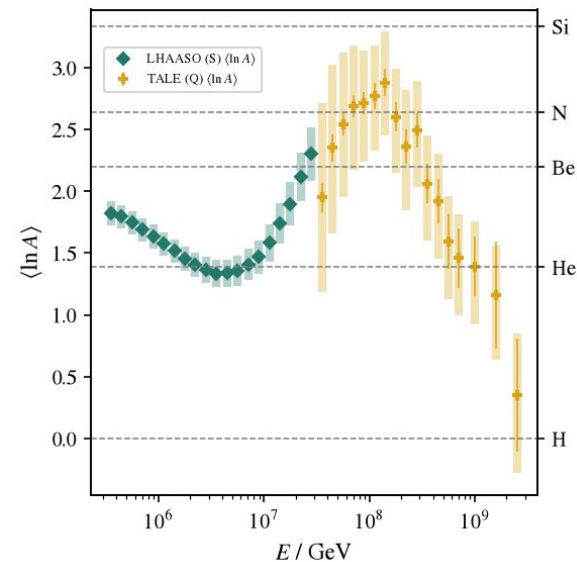
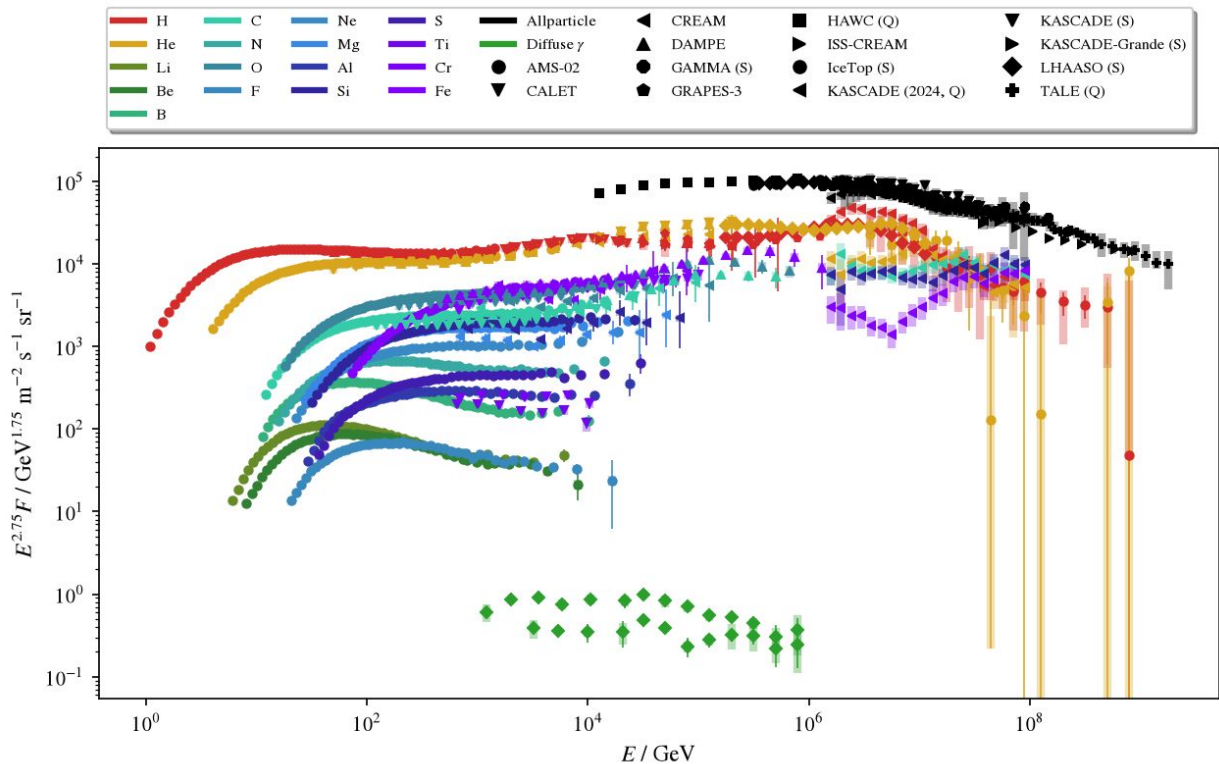


Evoli, '20

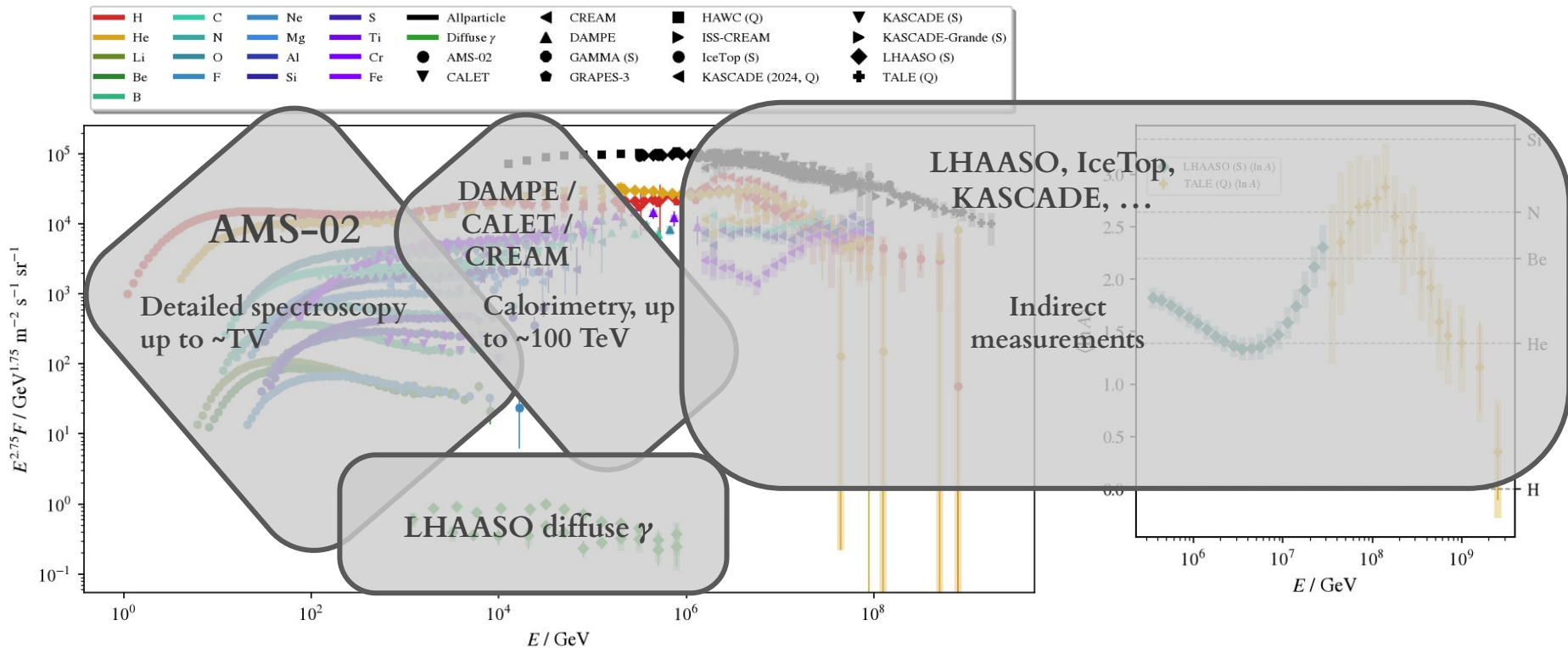
Project #1

Galactic source populations implied by spectrum measurements below and across the knee

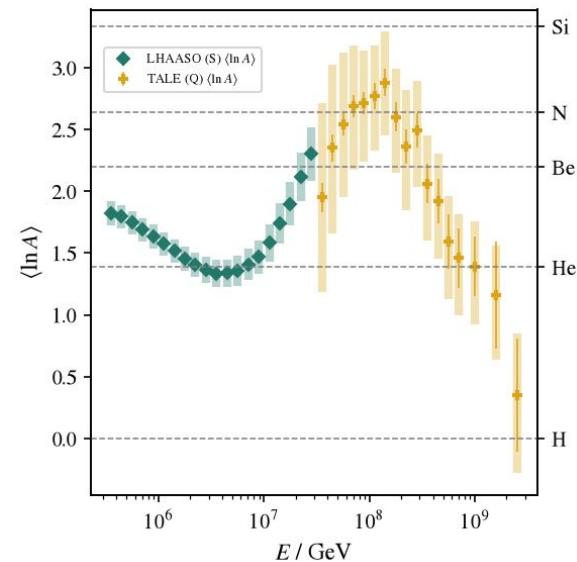
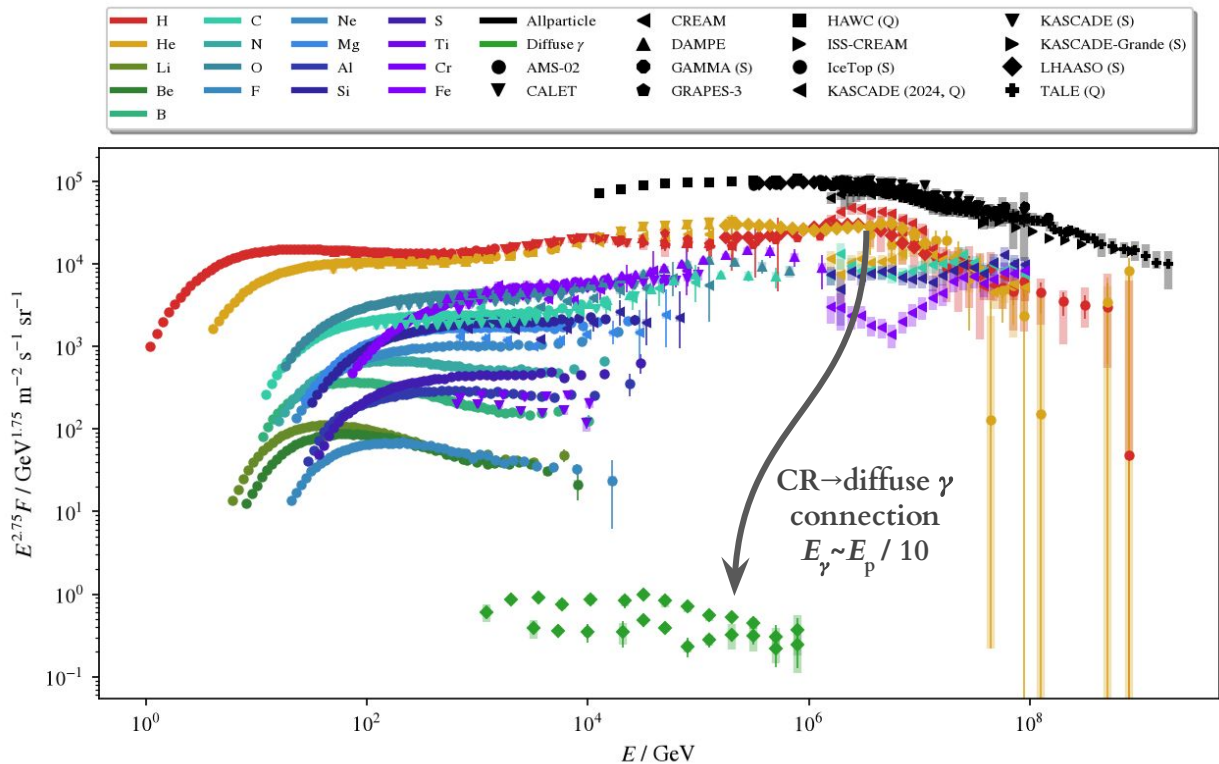
Context: cosmic ray landscape from GeV to PeV



Context: cosmic ray landscape from GeV to PeV



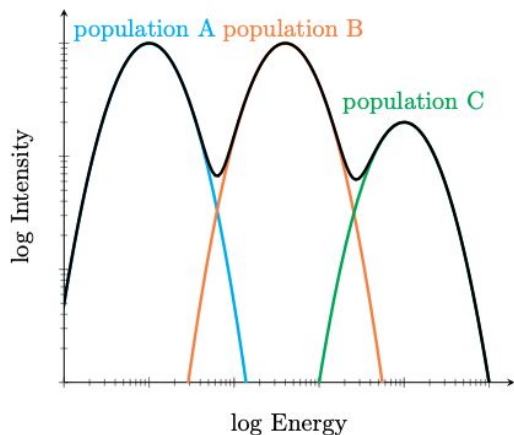
Context: cosmic ray landscape from GeV to PeV



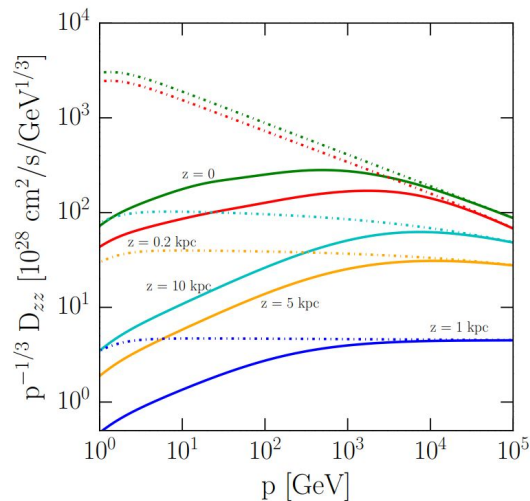
Spectral feature interpretation

Source populations

Propagation effect



Bumps = dominance regions of different accelerator populations; “ankles” do not obey $\mathcal{R}$ scaling



Bumps = changes in propagation mechanism, source of turbulence, etc; $\mathcal{R}$ -dependent features; discrimination by secondary/primary ratios

Not considered here:

- Stochasticity of sources (bumps = random fluctuations of the realization w.r.t. mean)
- Intrinsic source spectral features (depends on the conditions → washed out when integrating over the population)

 [Mertsch, '25](#)

 [Stall, Loo, Mertsch, '25](#)

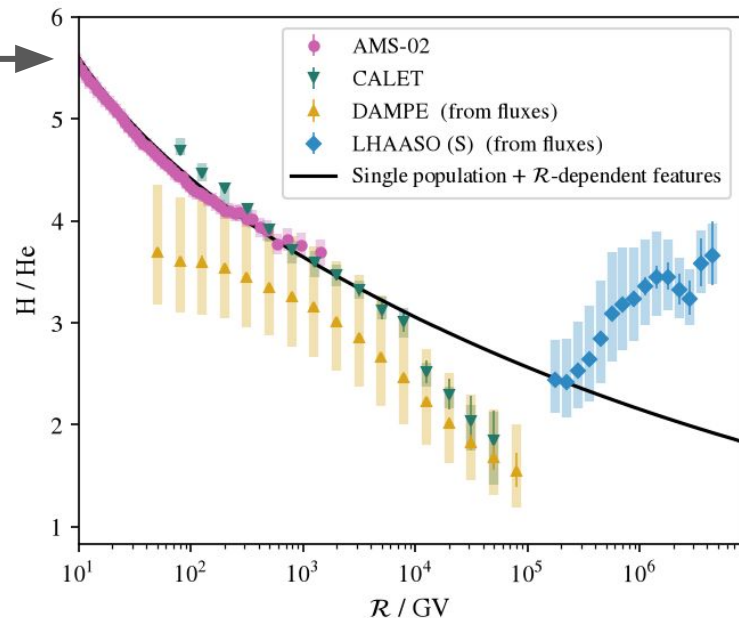
 [Evoli, Aloisio, Blasi, '19](#)

A minimal model of cosmic ray spectra

What is the simplest self-consistent description of the data in GeV-PeV range?

- Single population models are firmly excluded

Attributing *all* features to propagation (e.g. CRINGE model) is attractive, but is ruled out by clear non-rigidity-scaling at >100 TV



[IV, Evoli, Espinosa Castro, '25](#)

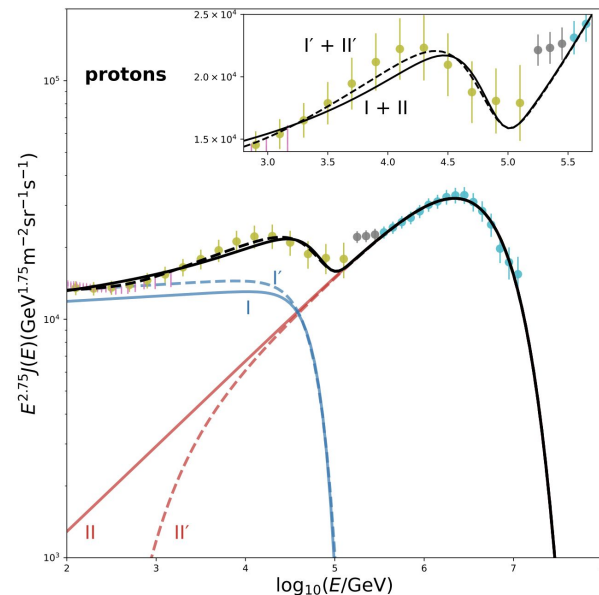
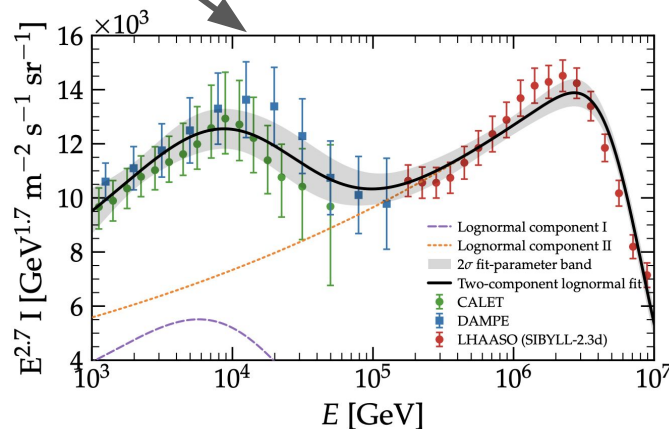
[Prevotat et al, '25](#)

[Schwefer, Mertsch, Wiebusch, '23](#)

A minimal model of cosmic ray spectra

What is the simplest self-consistent description of the data in GeV-PeV range?

- Single population models are firmly excluded
- Two-population models are viable at least for proton and Helium spectra



Aharonian, Zhang, '26

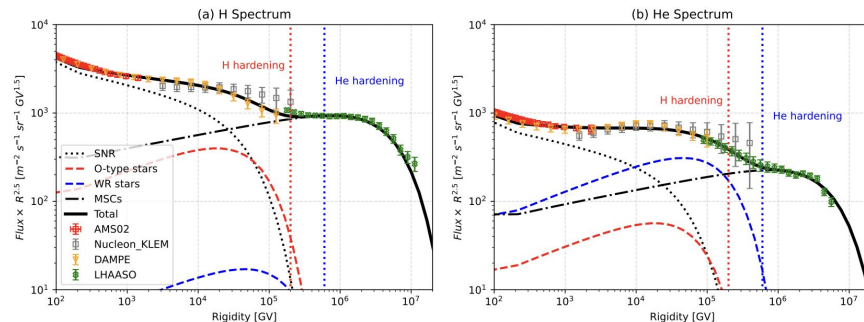
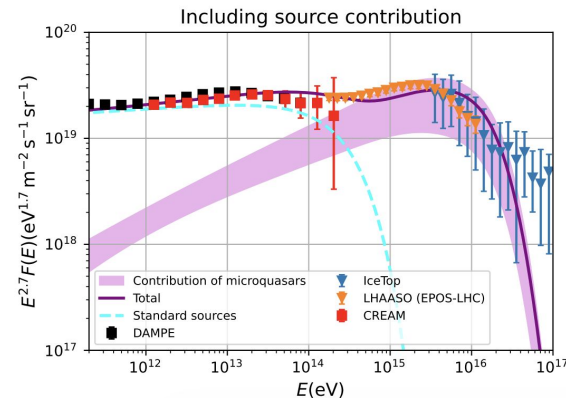
Evoli, '26

A minimal model of cosmic ray spectra

What is the simplest self-consistent description of the data in GeV-PeV range?

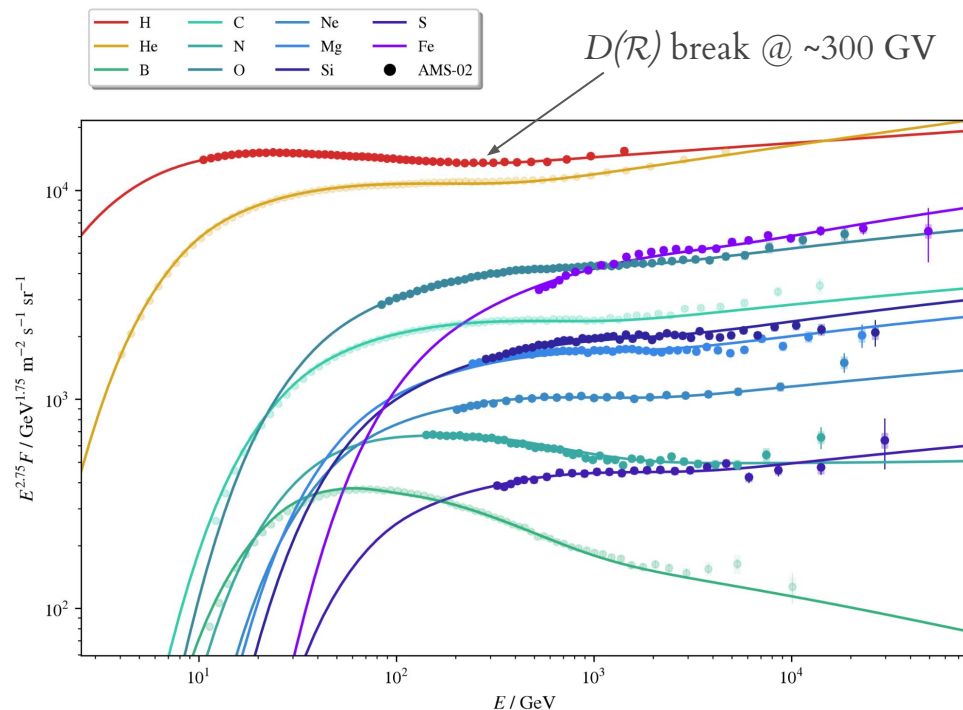
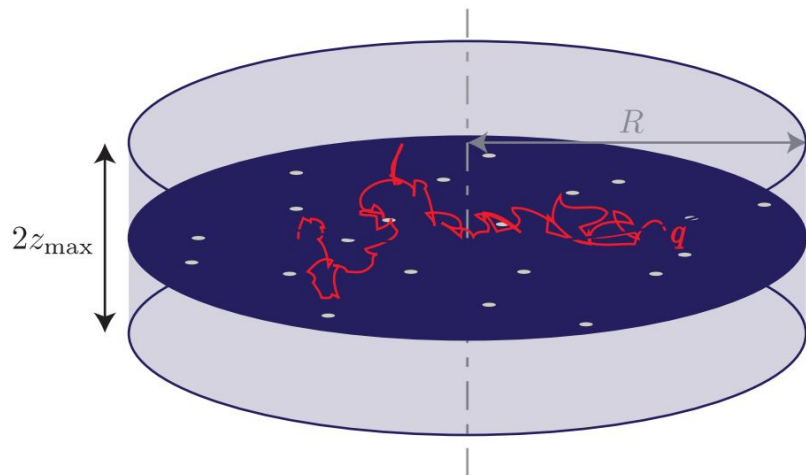
- Single population models are firmly excluded
- Two-population models are viable at least for proton and Helium spectra
- Regular galactic SNR seem to have trouble accelerating CRs to the knee, and another kind of sources is required: microquasars, massive star clusters, superbubbles, rare powerful SNRs, ...

[Diesing, '23](#)
[Zhang, Yu, '26](#)
[Kaci et al. '25](#)
[Qiu, Lin, Yang, '26](#)



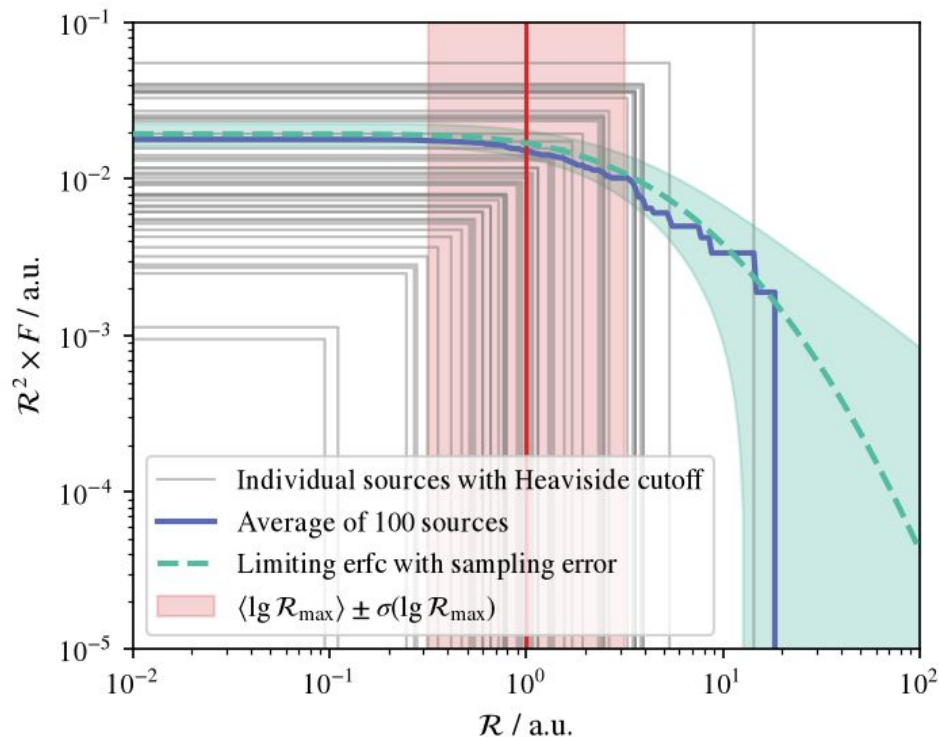
Simulating cosmic ray propagation

- Classical picture: resonant scattering on turbulent B $\rightarrow$ pitch-angle diffusion $\rightarrow$ spatial diffusion $\rightarrow$ 1D approximation along z
- Propagation effects = $D(\mathcal{R})$ dependence



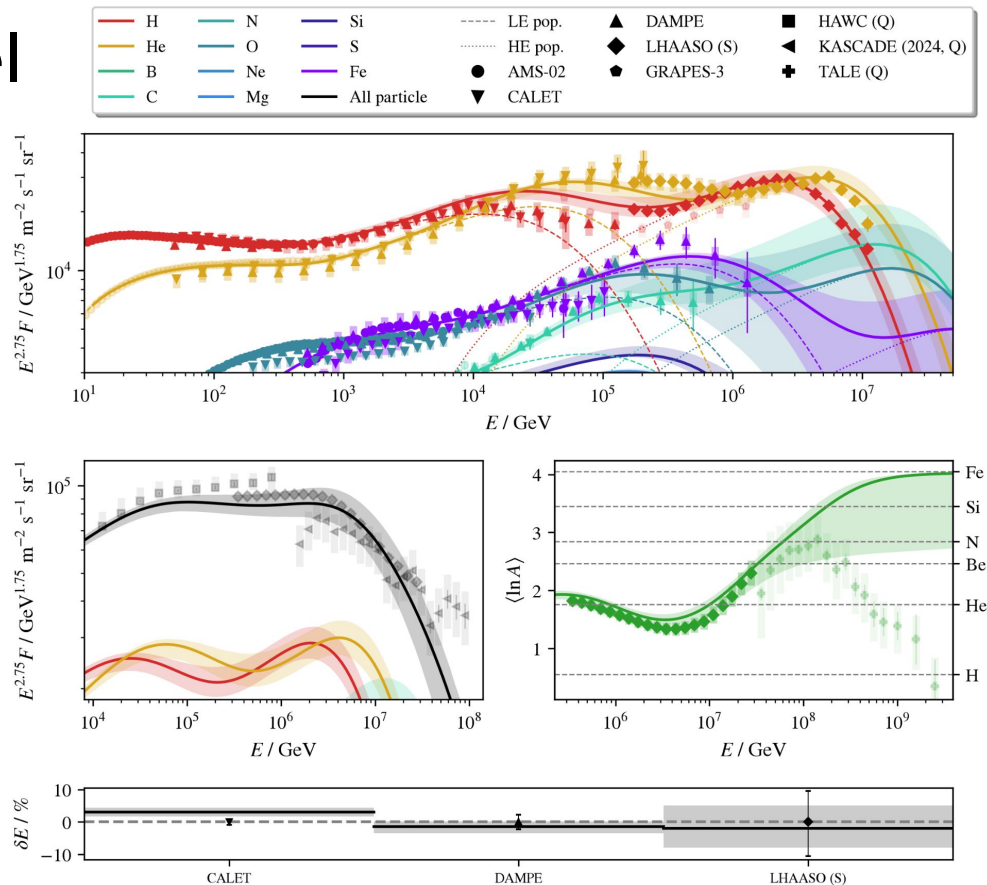
Population cutoff

- Assume individual sources to have sharp cutoffs distributed lognormally
- Then, effective injection has a erf-shaped cutoff
- Sharpness of the cutoff is connected to the variability of sources
- For Galactic SNR $\sigma \approx 0.5-0.6$, dominated by variance of E_{SN}
- Naive $\exp(-R/R_{\text{max}})$ cutoff corresponds roughly to $\sigma \approx 0.3-0.4$



Results: minimal model

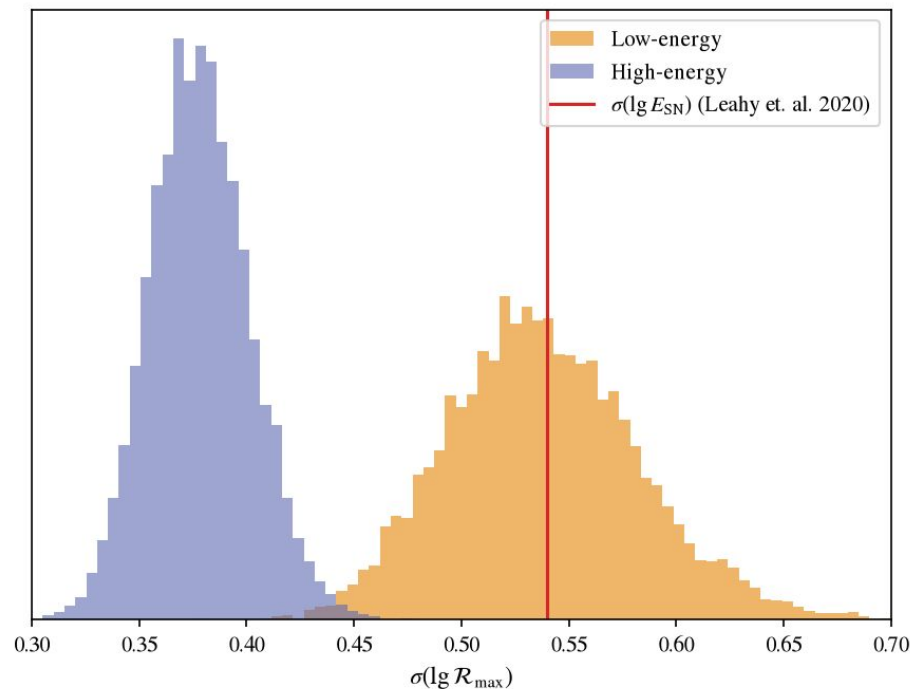
- Two-population model reproduces all data in GeV-PeV range reasonably well
- Low-energy population is constrained in the GeV range by AMS-02 and extrapolated up with the population cutoff
- High-energy population enters at ~ 100 TeV and dominates the knee
- Allparticle observables are not included in the fit, but reproduced up to ~ 30 PeV



Results: cutoff shape and source variability

Cutoff shape $\rightarrow R_{\max}$ variance

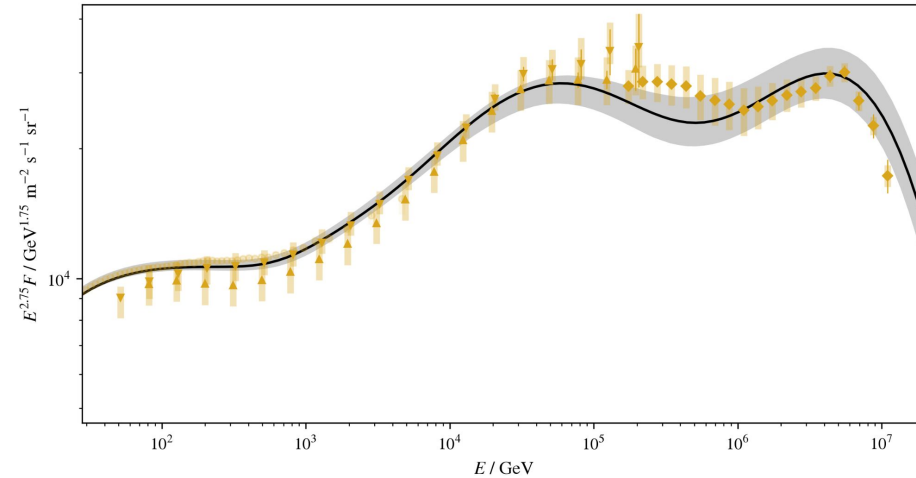
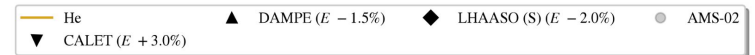
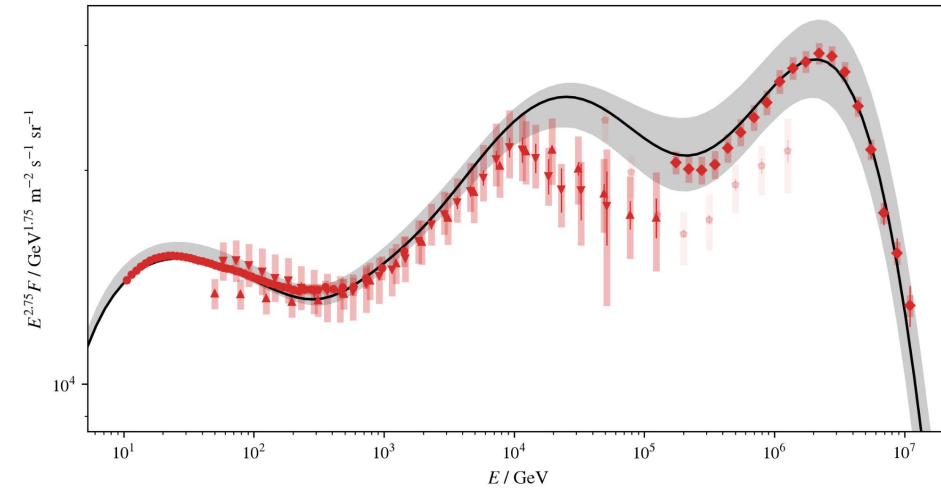
- HE: quite uniform sources, e.g. a few powerful accelerators
- LE: consistent with E_{SN} variance, although intrinsic cutoff shape and environment variance need to be taken into account



[Leahy, Ranasinghe, Gelowitz, '20](#)

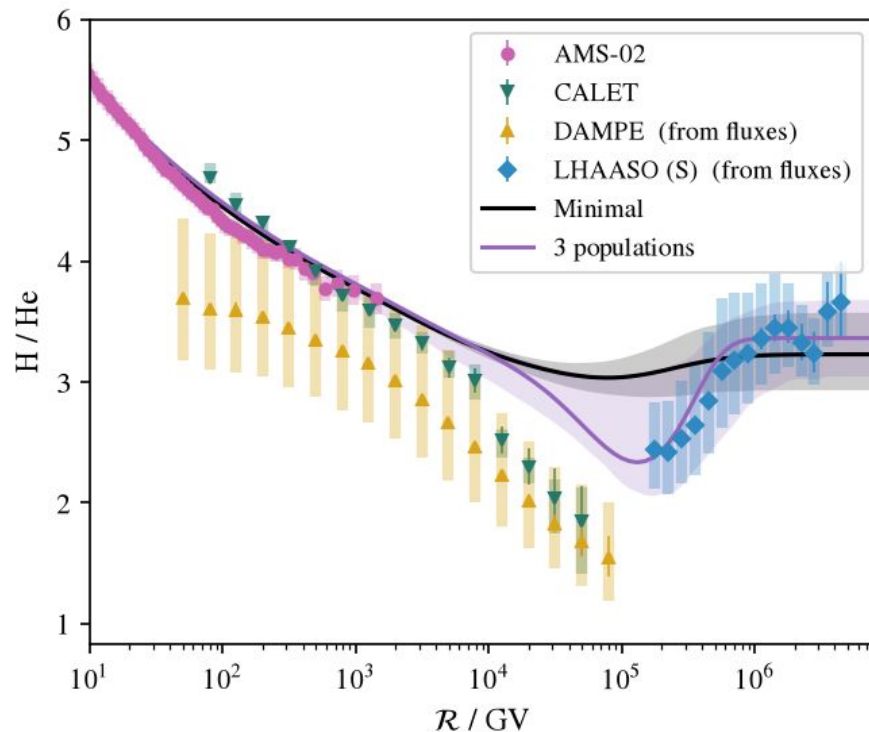
Minimal model shortcomings

- In 10–100 TeV range, H and He are over- and underpredicted, although within (correlated) systematics
- If true physical effect, this requires another population to provide non- R -scaling features in that range



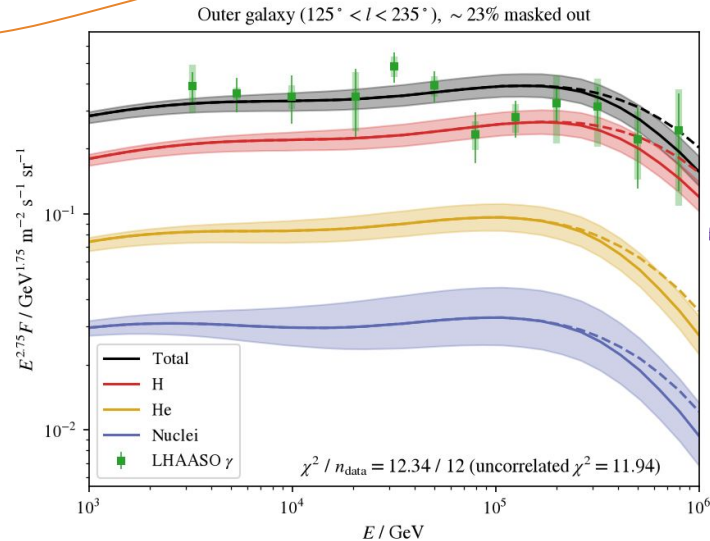
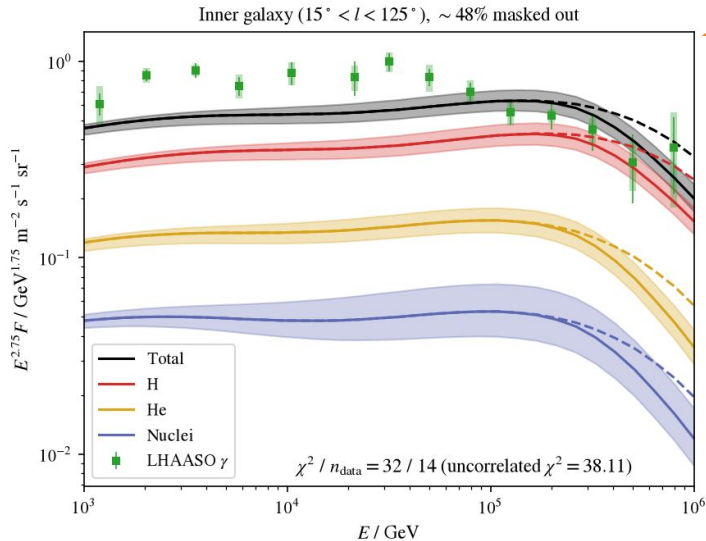
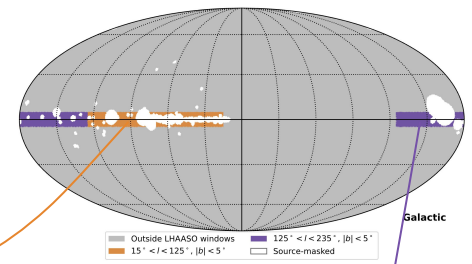
Proton-over-Helium diagnostic

- p/He ratio directly probes non-rigidity dependent structures in the spectra (e.g. population crossover)
- At the same time, some systematics cancel out in ratios, providing a more precise measurement
- 10-100 TV dip in the p/He ratio is a smoking gun of the more complicated interpretation



Diffuse γ -ray implications

- We use **HERMES** code to LOS-integrate γ -ray emissivity from CR-gas interactions and produce diffuse γ -ray spectrum, assuming homogeneous CR density
- Outer galaxy is well reproduced, while there's clear discrepancy in the inner one, not explainable by a simple rescaling



[Dundovic et al, '21](#)

[Espinosa Castro et al, '25](#)

Conclusions

- CR spectra measurements imply at least two Galactic source populations
- There are good candidates for both populations; CR spectrum carries information about both their intrinsic and population properties
- The minimal model is sufficient, but already hints at further complexity
- Data interpretation is mostly limited by systematics and their usual treatment is flawed
- Diffuse γ -ray data probe possible non-local features of the CR spectrum

Project #2

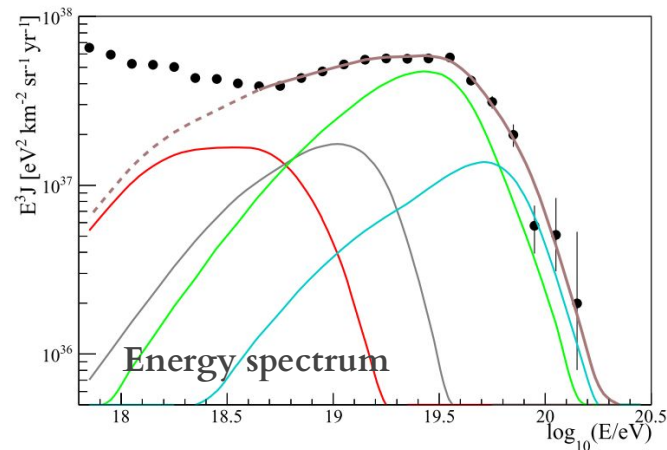
Inferring properties of ultra-high-energy cosmic-ray sources with Information Field Theory

Context: UHECR spectrum and composition fit

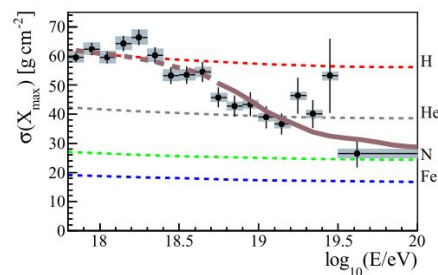
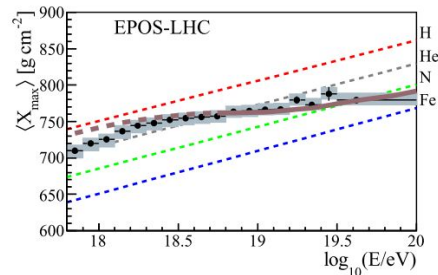
- **Sources:** diffuse, isotropically & homogeneously distributed
- Injection **spectrum:**

$$J \propto E^{-\gamma} f_{\text{cut}}(E/ZR_{\text{cut}})$$

- **Propagation:** 1D Monte Carlo simulation with adiabatic, pair production, and photomeson losses + photodisintegration
-  Consistent description of spectrum and composition
-  Some challenges for the interpretation...



Depth of shower maximum distribution ~ composition

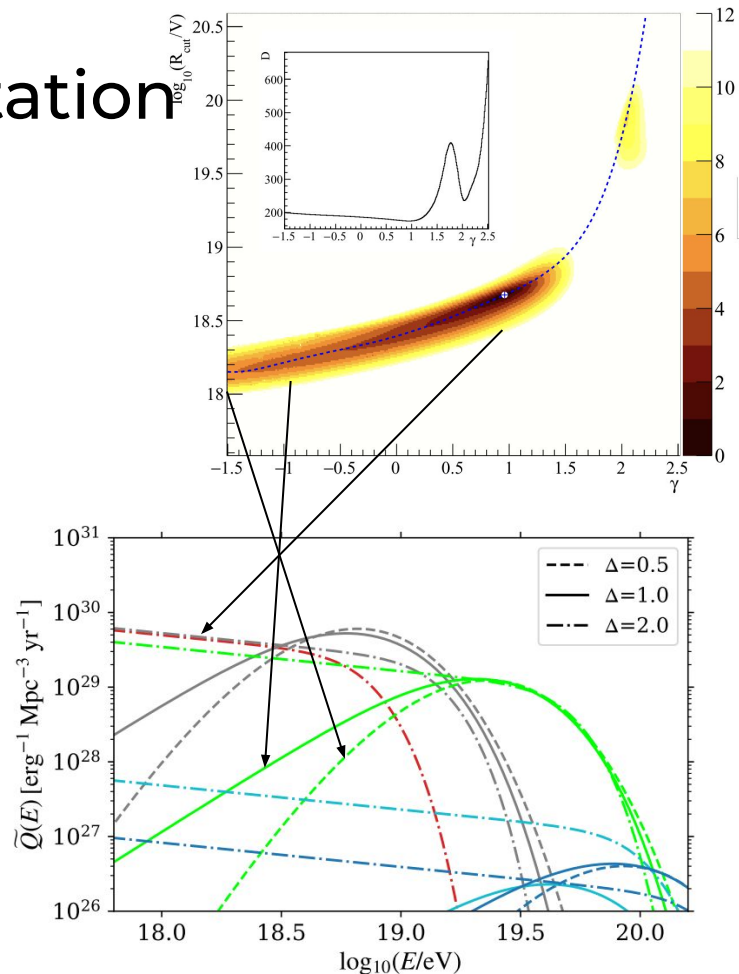


Context: combined fit interpretation

- Best-fit spectrum is very hard: $\gamma \cong 1$ originally, down to $\gamma \lesssim -2$ with more recent datasets
- The spectral index and cutoff rigidity are highly degenerate between each other...
- ... and with the cutoff shape!

$$f_{\text{cut}}(E) = \exp\left(1 - \max\left(1, E/E_{\text{cut}}\right)\right)$$

$$f_{\text{cut}}(E) = \text{sech}\left(\left(E/E_{\text{cut}}\right)^\Delta\right)$$



Motivation

The standard spectrum parametrization is

-  broadly theoretically motivated
-  well-established

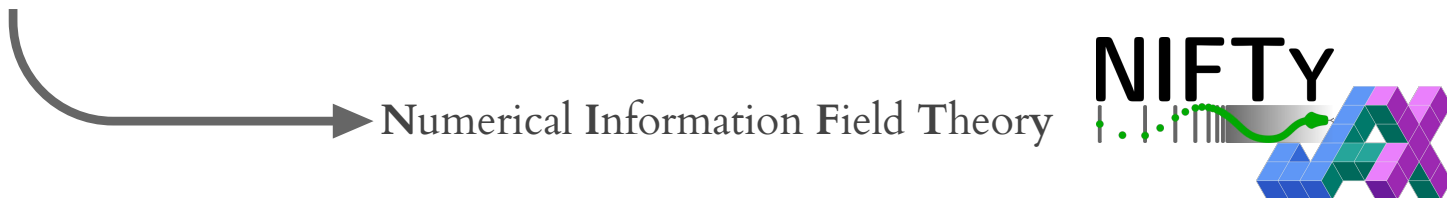
...but

-  yields uninterpretable results
-  parameter space is highly degenerate
-  in some cases doesn't describe the data at all

Let us look for an alternative way to specify the signal shape!

Information Field Theory

- *Theoretically*, IFT combines Bayesian statistics in its information-theoretic form with methods from quantum and statistical field theories to infer continuous quantities (fields)
- *Practically*, it's a Bayesian framework for efficient and statistically sound signal reconstruction
- *Technically*, it has inspired a few state-of-the-art Gaussian Variational Inference algorithms and self-calibrating setups

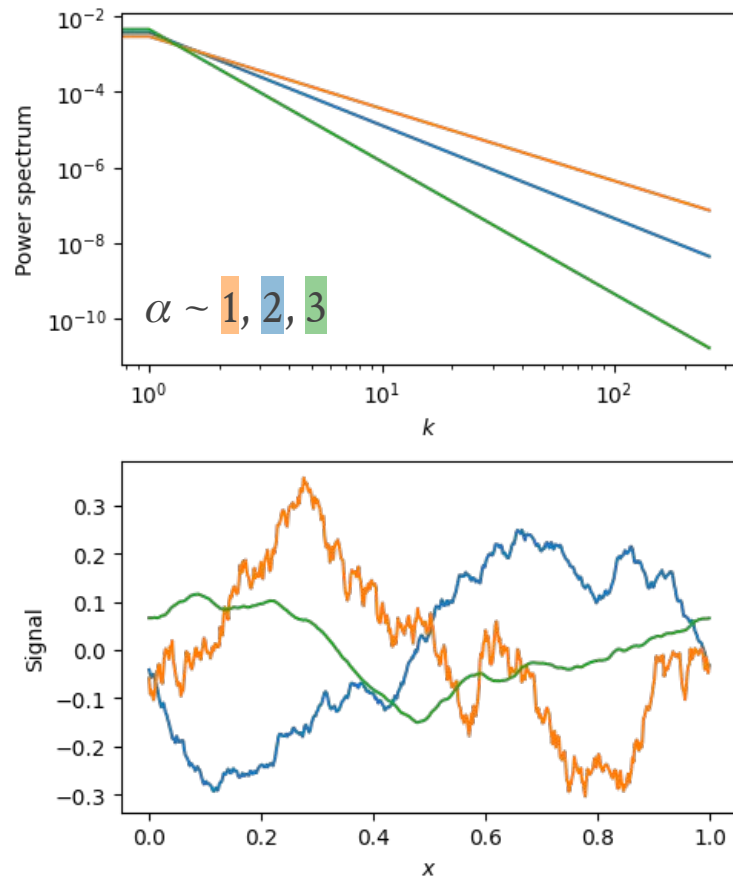


 [Enßlin, '19](#)

 [Edenhofer et al, '24](#)

Generic signal model

- Isotropic and homogeneous arbitrarily-dimensional Gaussian field (1D in our case)
- Fully defined by power spectrum (Wiener–Khinchin)
- Power spectrum captures our apriori expectation on signal's roughness / smoothness
- Ideally, prior power spectrum would be chosen from theoretical considerations, but can also be fitted during inference (self-calibration)
- Observed signal (likelihood) breaks the assumptions of isotropy and homogeneity!



UHECR spectrum model

$$Q_A(E) = \underbrace{\frac{Q_{A,0}}{Z}}_{\text{Per-element normalization}} \underbrace{\frac{dN}{dV dt d\mathcal{R}}}_{\text{Baseline signal}} = \frac{Q_{A,0}}{Z} \times \underbrace{B(\mathcal{R})}_{\text{Nonparametric shape modifier}} \times \underbrace{10^{\text{GF}(\lg \mathcal{R})}}_{\text{log-space Gaussian field; power spectrum } P(k) \propto Ak^{-\alpha}}$$

Per-element **normalization**, the shape of the spectrum is universal

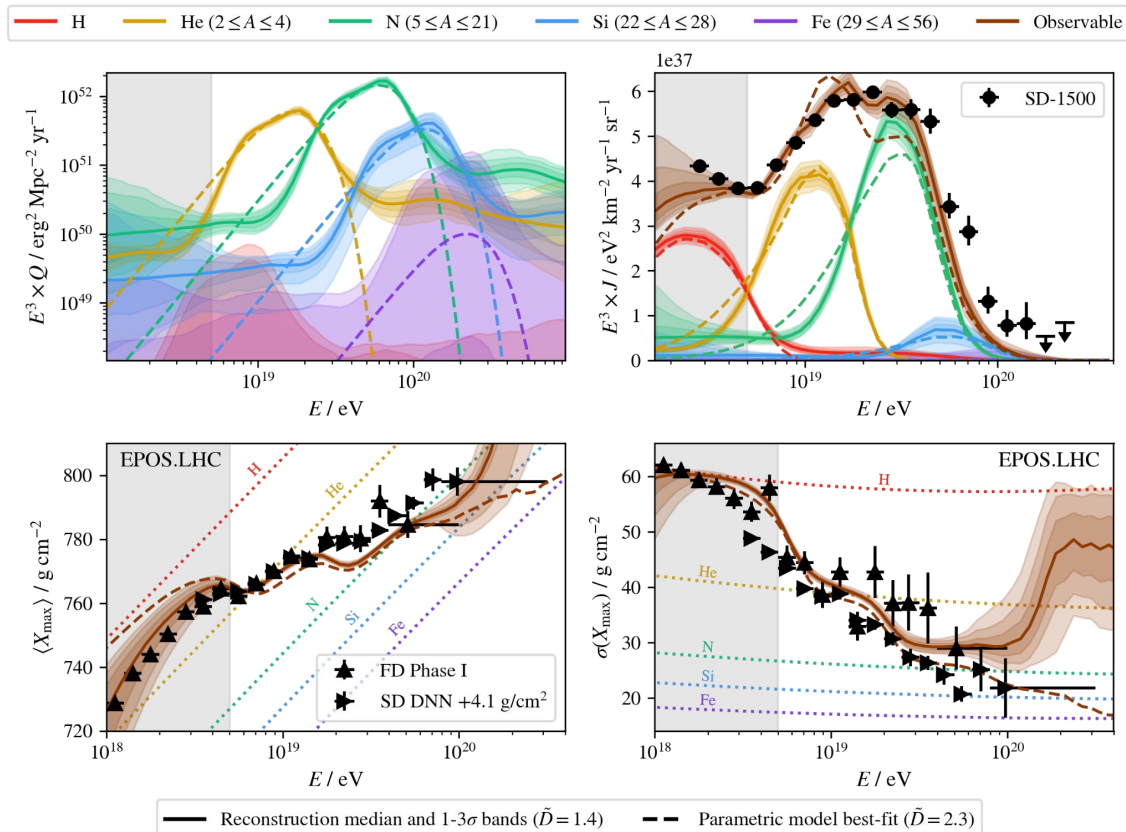
Baseline signal, deviations from which are modeled nonparametrically; can be constant or fitted

Nonparametric **shape modifier** — log-space Gaussian field; power spectrum $P(k) \propto Ak^{-\alpha}$

Parametric alternative: $Q_{A,0} \times (E/E_0)^{-\gamma} \times \text{sech} \left((E/Z\mathcal{R}_{\text{cut}})^\Delta \right)$

Reconstruction results: reference case

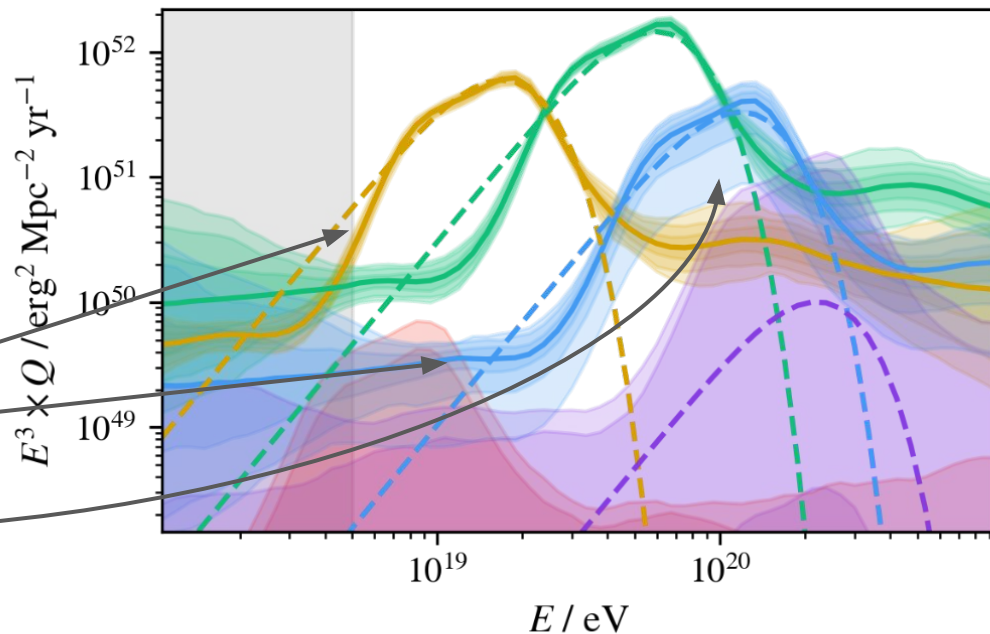
- We fit latest FD & SD DNN $X_{\max}$ measurements with EPOS.LHC & nominal E & $X_{\max}$
- Gaussian field hyperparameters:
 - $\alpha = 3$ — relatively smooth shape
 - $A = 3$ — a priori ± 3 orders of magnitude deviation from the baseline
- Deviance per n.d.o.f. improvement $2.1 \rightarrow 1.4$ w.r.t. parametric model



Reconstruction results: spectral features

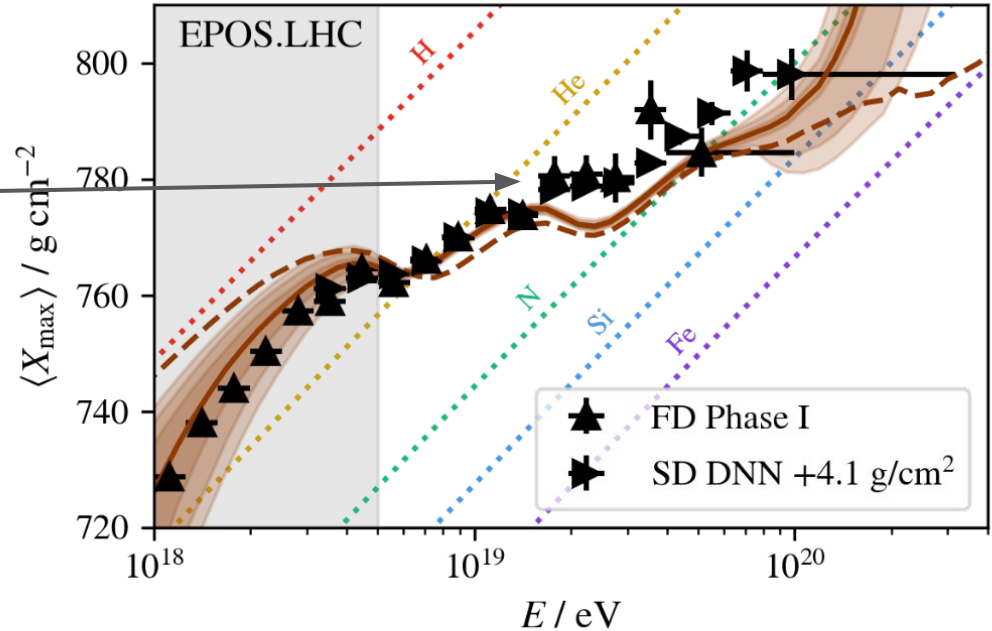
“Bump” in the spectrum is reconstructed, but with substantial differences w.r.t. the parametric case:

- Very hard $\gamma \ll 0$ spectra are localized to half a decade / factor of 3 in $\mathcal{R}$
- Soft $\gamma \sim 3$ low-energy tail reconstructed
- $\gamma \sim 2$ “shelf” between two suppressions



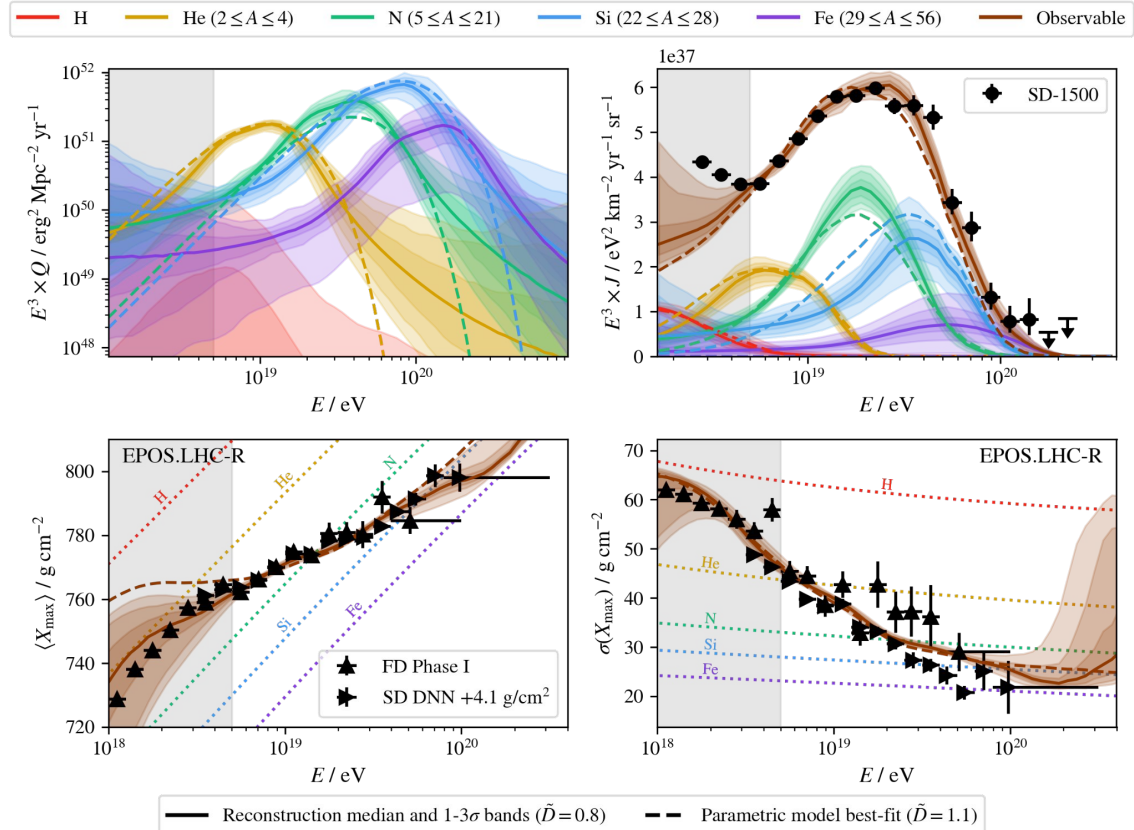
Reconstruction results: shortcomings

- Despite the flexible injection spectrum model, the spectrum and composition still cannot be described consistently
- Genuine tension from e.g. inadequate hadronic interaction model



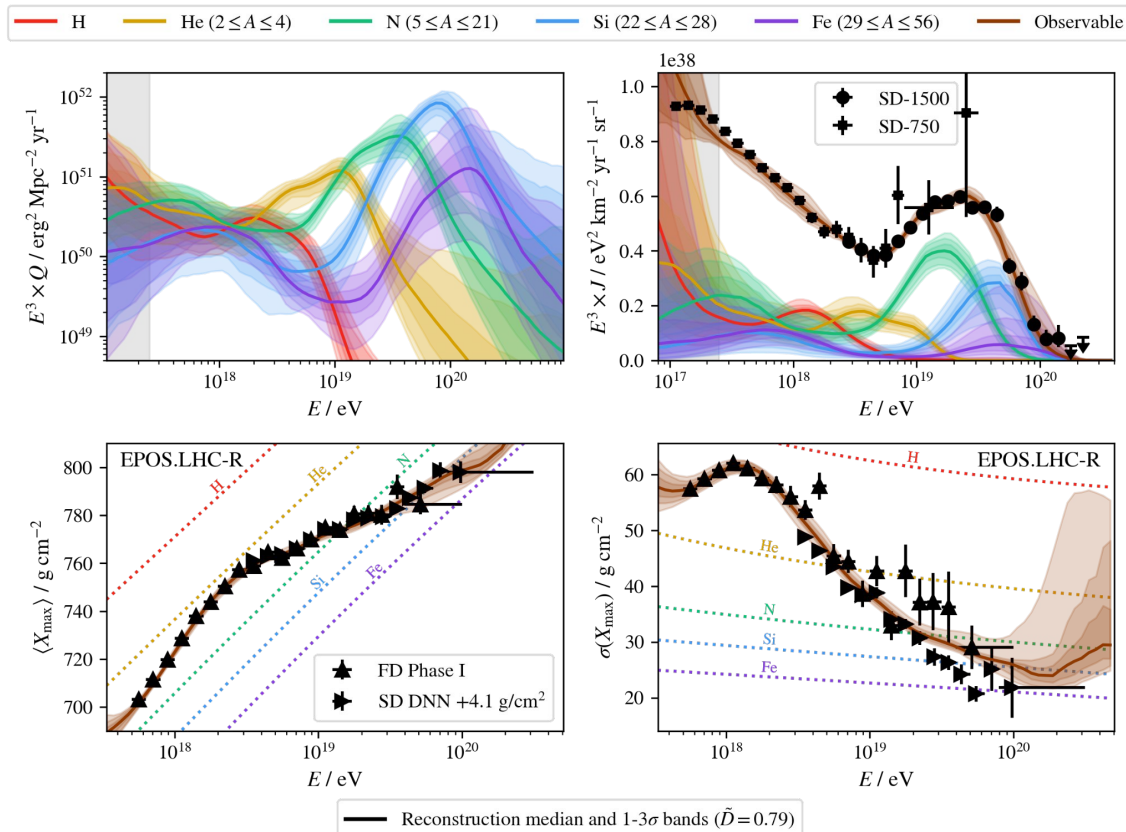
Reconstruction results: EPOS.LHC-R

- New(er) EPOS.LHC-R predicts significantly deeper showers $\rightarrow$ same data imply heavier composition
- Tensions between spectrum and composition data are mostly gone!
- Nonparametric reconstruction is closer to the parametric baseline, although most qualitative features remain
- A similar effect is achievable by considering systematic uncertainties in E and $X_{\max}$ scale



Reconstruction results: deviation from $\mathcal{R}$ -scaling

- As in the previous project, multiple source populations lead to deviations from $\mathcal{R}$ -scaling of spectral features
- Instead of adding the populations manually, we incorporate the $\mathcal{R}$ -independent part of the signal in the reconstruction
- We can then recover the populations a posteriori from the data-driven reconstruction



Conclusions

- Standard parametric models for the UHECR injection emissivity are spherical cows: broadly theoretically motivated, but possibly used beyond their applicability
- Nonparametric method based on the Information Field Theory offers a flexible alternative that is easier to interpret and distinguish $\pi/\mathcal{L}$ -dominated signal regions
- The reconstruction shows rich and interpretable phenomenology
- The reconstruction is stable with respect to main hyperparameters
- Further extensions into the spatial dimension(s) might resolve the UHECR source spectrum evolution with distance → look for spatially resolved populations

Outlook

- The two main projects are being finalized and prepared for publication
- Other ongoing projects at different stages:
 - UHECR mass composition methodology and systematics
 - Phenomenological implications of the ankle shape
 - Galactic magnetic field signatures in the UHECR spectrum

Publications

- Evoli, C., Vaiman, I., Petrera, S., & Salamida, F. (2026). Updated air-shower X_{max} moment parametrizations for UHECR composition with latest hadronic interaction models. [Astroparticle Physics, 179, 103239](#)
- Vaiman, I., Evoli, C., & Espinosa Castro, L. E. (2025). Revisiting the Cosmic Ray knee: new insights from LHAASO. [Proceedings of 39th International Cosmic Ray Conference — PoS\(ICRC2025\), 148](#)
- Pierre Auger Collaboration. (2026). Depth of maximum of air-shower profiles above $10^{17.7}$ eV measured with the fluorescence detector of the Pierre Auger Observatory. [Physical Review D, 114\(4\)](#)

Conferences and seminars

- International Cosmic Ray Conference (ICRC) 2025, Geneva, Switzerland, 14–24 July 2025
 - *Revisiting the Cosmic Ray knee: new insights from LHAASO*
- Rencontres de Moriond 2026: Very High Energy Phenomena in the Universe (VHEPU), La Thuile, Italy, 15–22 March 2026
 - *Astrophysical interpretation of ultra-high-energy cosmic-ray measurements at the Pierre Auger Observatory* (on behalf of the Pierre Auger Collaboration)
- KIT IAP Seminar, Karlsruhe, Germany, 7 May 2026
 - *Reconstruction of UHECR source spectra with Information Field Theory*
- 29th European Cosmic Ray Symposium (ECRS), Santiago de Compostela, Spain, 29 June–3 July 2026
 - *Anisotropy of ultra-high-energy cosmic-ray arrival directions and its astrophysical interpretations at the Pierre Auger Observatory* (on behalf of the Pierre Auger Collaboration)
 - *Characterizing sources of ultra-high-energy cosmic rays with Information Field Theory*
- TeV Particle Astrophysics (TeVPA), Tendō, Japan, 30 August–4 September 2026
 - *A minimal two-population model of cosmic ray spectra at TeV–PeV energies*

Thank you!