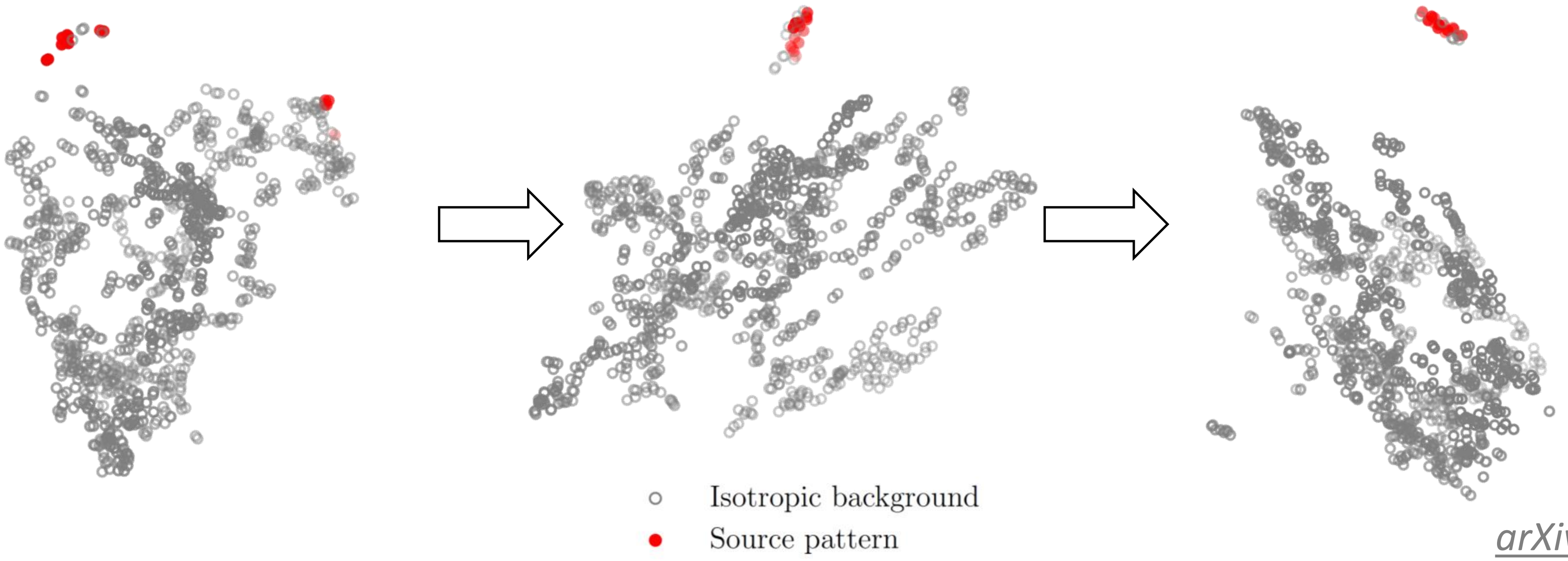
Identification of Patterns in Cosmic-Ray Arrival Directions using **Dynamic Graph Convolutional Neural Networks**

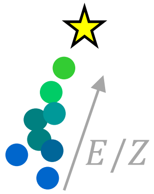
Teresa Bister, Martin Erdmann, Jonas Glombitza
Niklas Langner, Josina Schulte, Marcus Wirtz



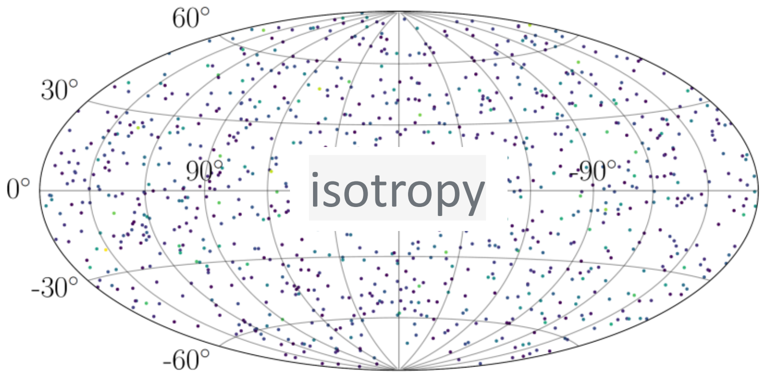
arXiv:2003.13038

Motivation

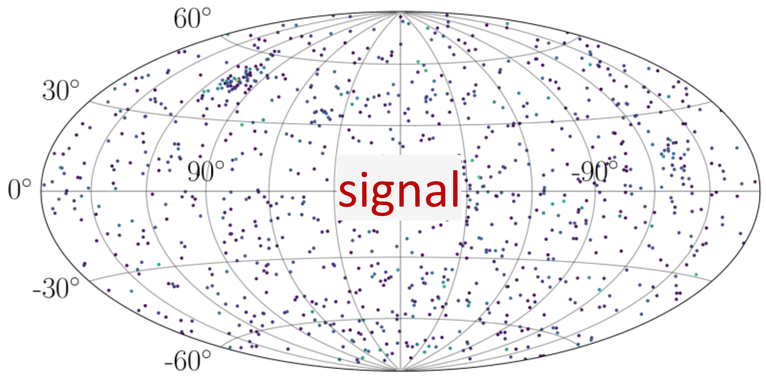
Ultra-high energy cosmic rays from source → Passage through GMF → **Energy- and charge**-dependent deflection → **Pattern** in arrival directions



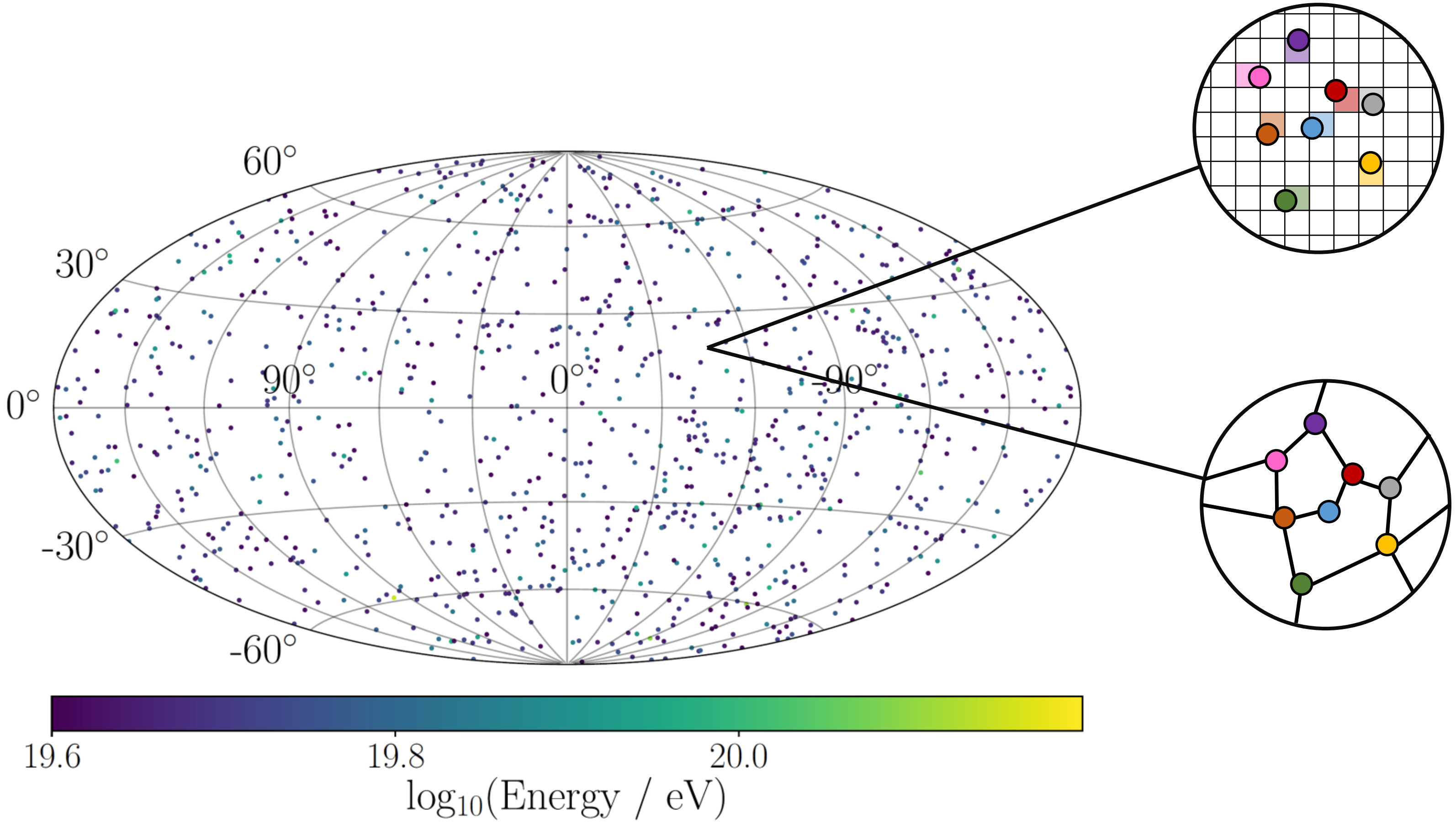
- Identify patterns to identify sources
- Pattern recognition task
- Use **convolutional neural networks**



Classification task



Approach

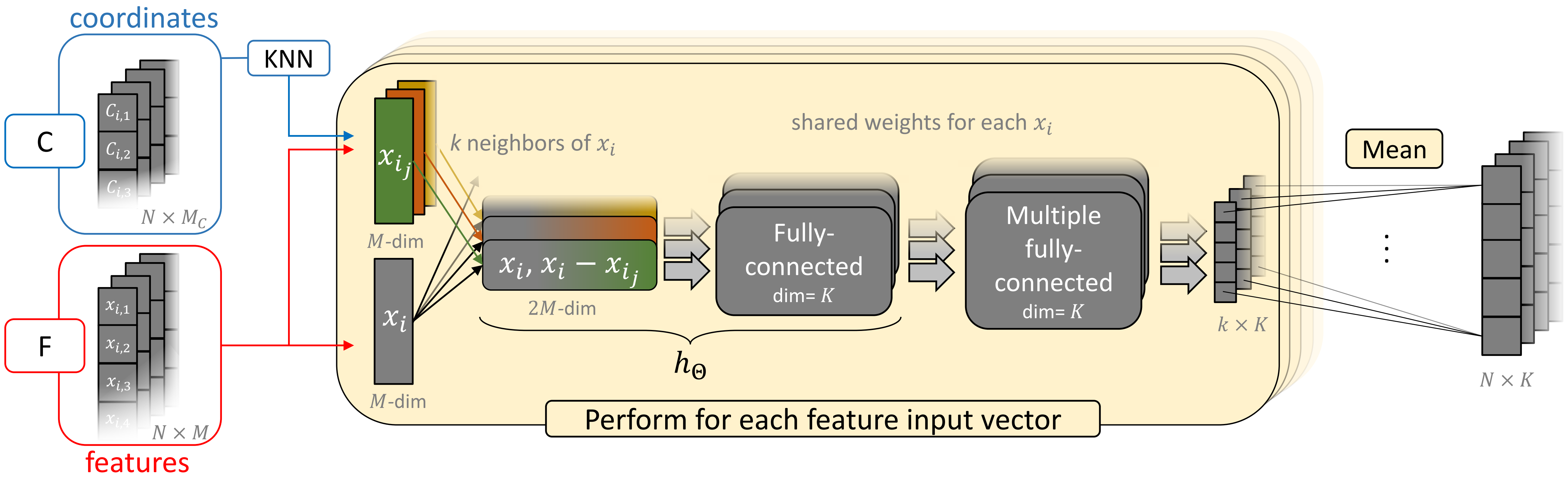


Continuously distributed on **sphere**

- ↯ Not well-suited for classical pixel-grid-based CNNs
- ➞ **Use Graph Convolutional Neural Network!**

EdgeConv[1] **Layer**

$$h_\Theta^a\left(x_{i,c}, x_{i_j,c}\right) = \sum_{c=1}^{M} \theta_c^a\, x_{i,c} + \sum_{c=1}^{M} \theta_c'^a\left(x_{i_j,c} - x_{i,c}\right) \qquad a = 1 \dots K \text{ (number of filters)}$$

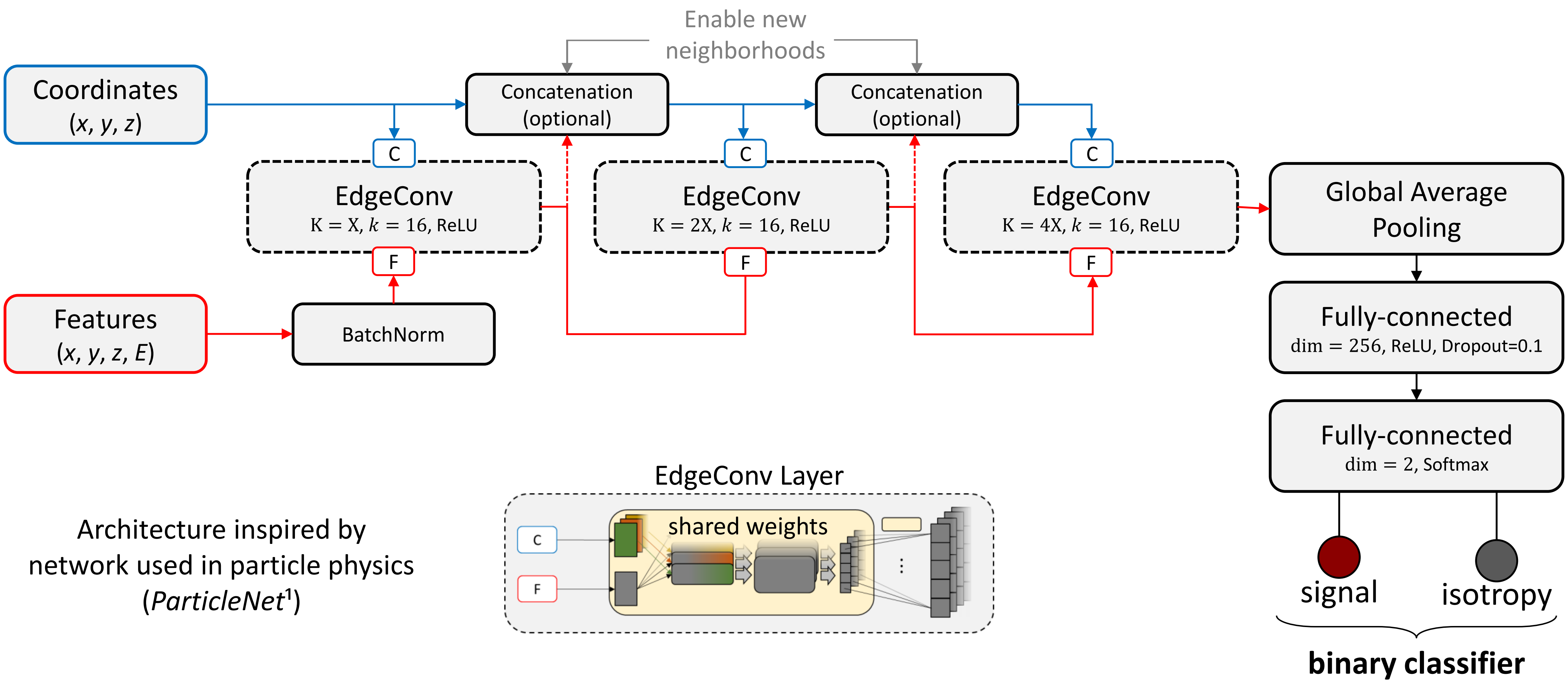


- **Weight sharing** for each pair of cosmic rays
- **Mean** over neighbors j

} **permutation invariance**

[1] https://arxiv.org/abs/1801.07829

Network Architecture

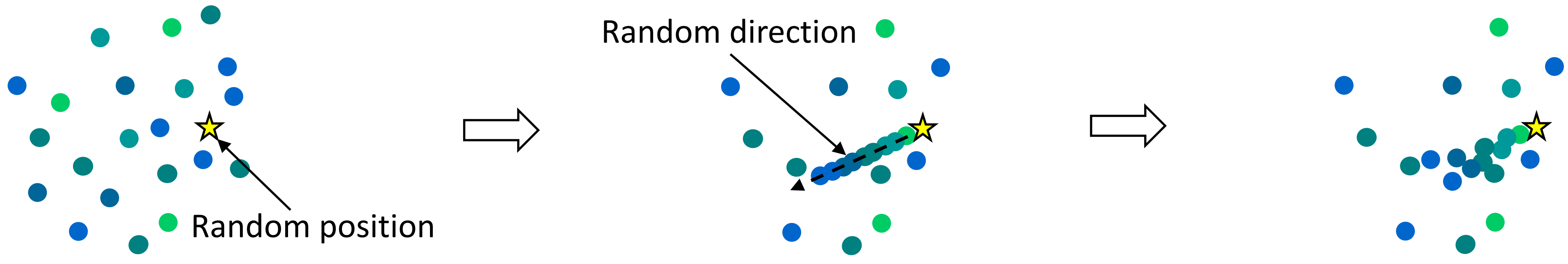


Architecture inspired by network used in particle physics (*ParticleNet*[1])

[1] https://arxiv.org/abs/1902.08570

Simulation of a Single Source

Simplified scenario: **one source pattern** of N_S cosmic rays + isotropic background



Isotropy

1. **Coherent deflection**
   Rotation in random direction with rotation angles

$$\delta_{\mathrm{coh}}(R = E/Z) = \frac{D}{R/EV}\,\mathrm{rad}$$

2. **Turbulent deflection**
   Scattering according to Fisher distribution of width

$$\sigma_{\mathrm{turb}}(R = E/Z) = \frac{T}{R/EV}\,\mathrm{rad}$$

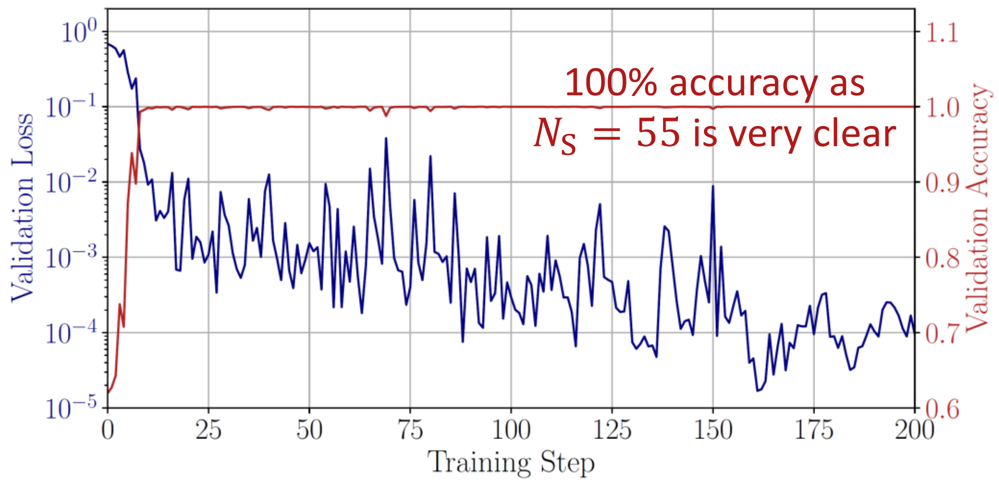
Training

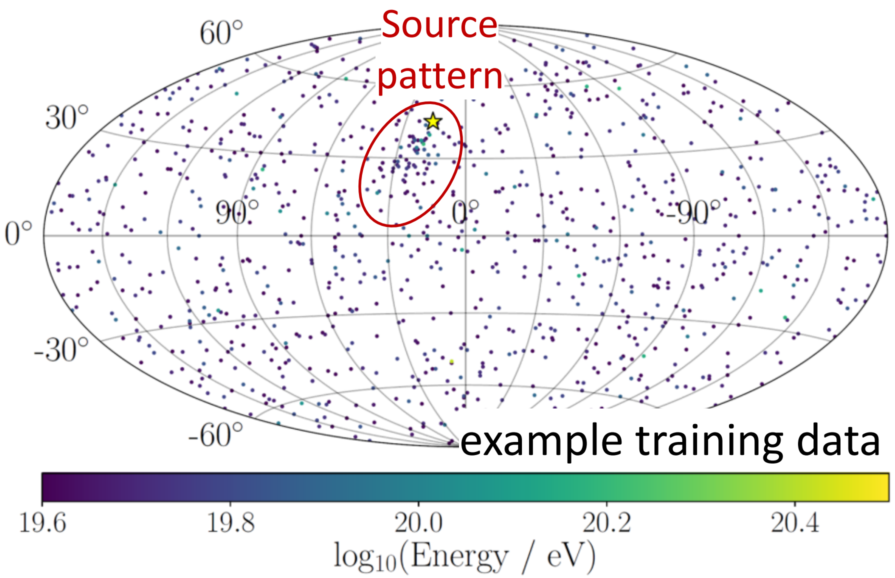
- **1000 cosmic rays** with $\boldsymbol{E > 40}$ **EeV**, spectrum similar to measurements of Pierre Auger Observatory
- Simulate on the fly during training → **no overfitting**
- Train on **strong multiplets** and let the network generalize

Composition	Turb. deflection T	Coherent deflection D	Source CRs
Pure Helium	50% of JF12 maximum in train values from JF12 in validation	Typical values from JF12 but larger than turbulent	55

EdgeConv dims	Loss	Optimizer	Concatenation
16/32/64	Categorical cross entropy	Adam	No

Typical training

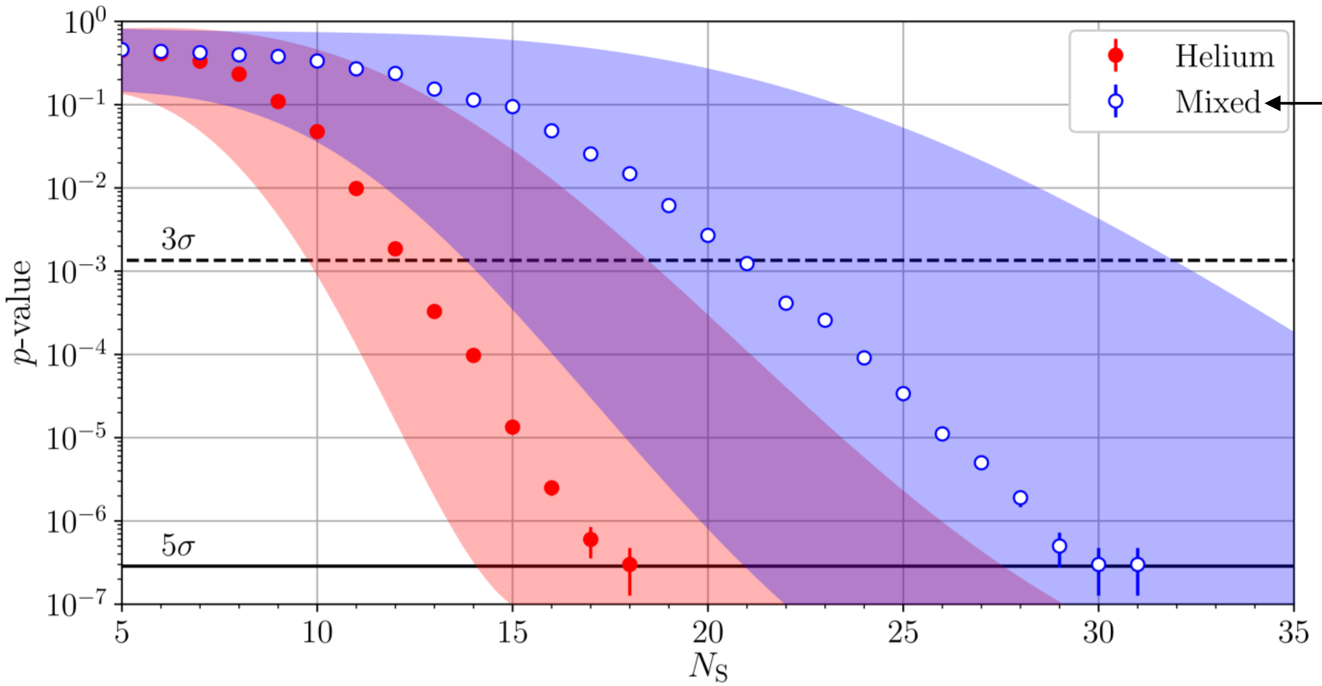




Sensitivity

Analyze cosmic-ray skies simulated using position-dependent deflection strengths from JF12

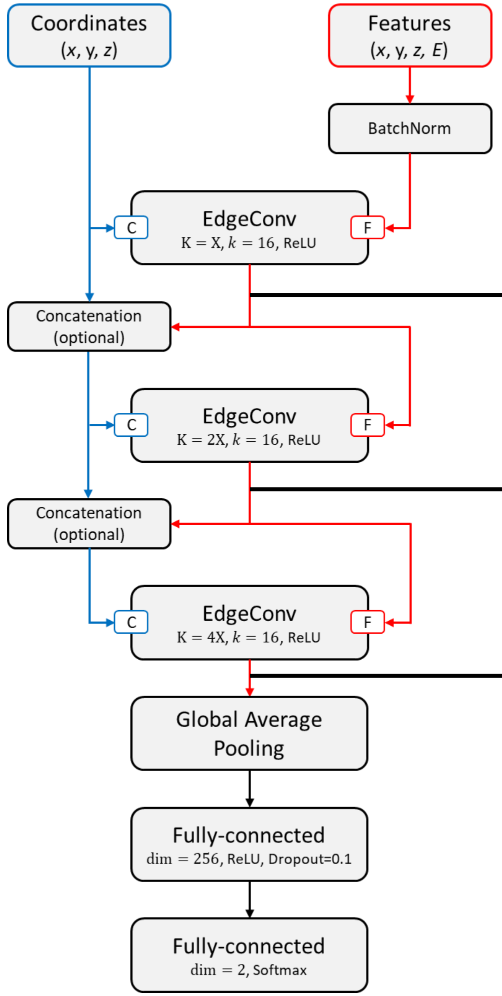
- Determine the **'signal'-output** of
  - 10^3 cosmic-ray skies for **varying N_S** (x_{sig})
  - 10^7 **isotropic** cosmic-ray skies
- Calculate the relative amount of **'signal'-outputs from isotropy** $\geq x_{\mathrm{sig}}$



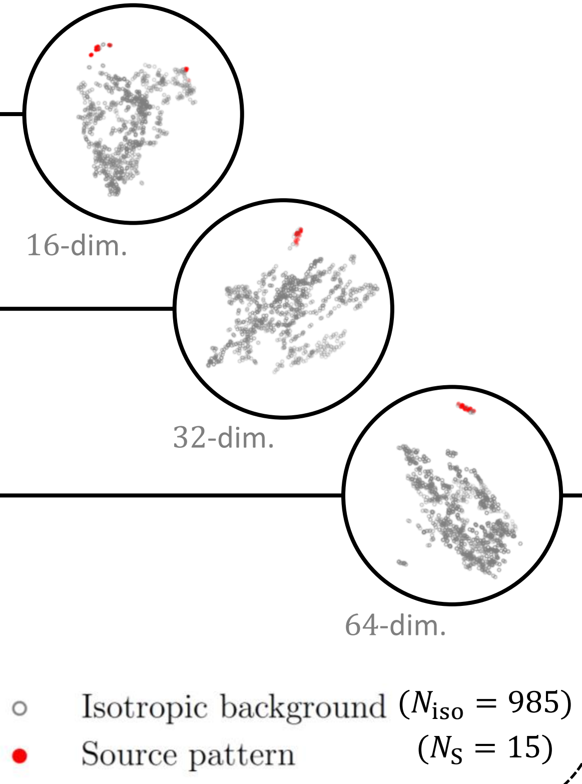
15% H
45% He
40% C, N, O

Works great, but what does it do?

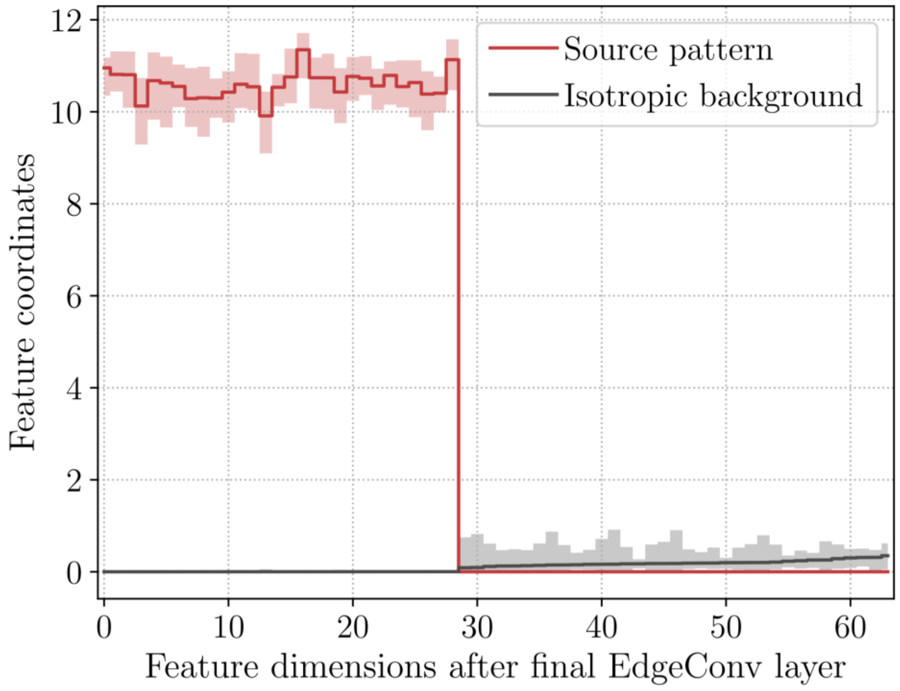
The Network‘s Working Principle



Visualize **high-dimensional point clouds** using t-SNE:

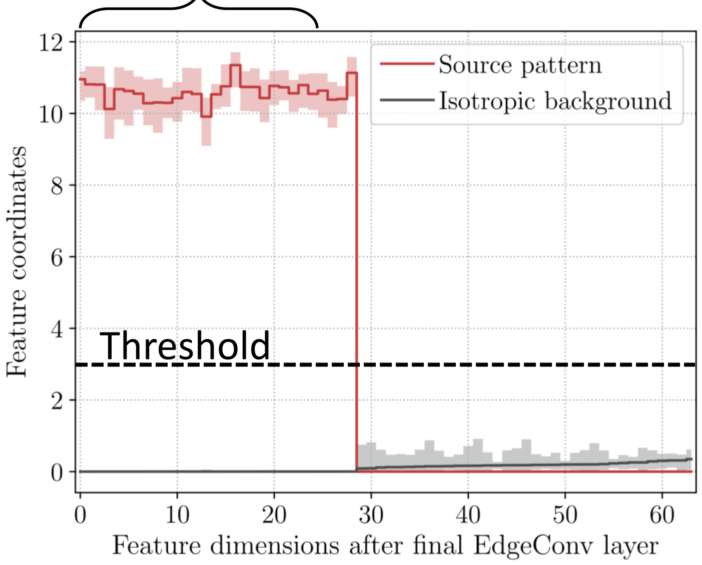


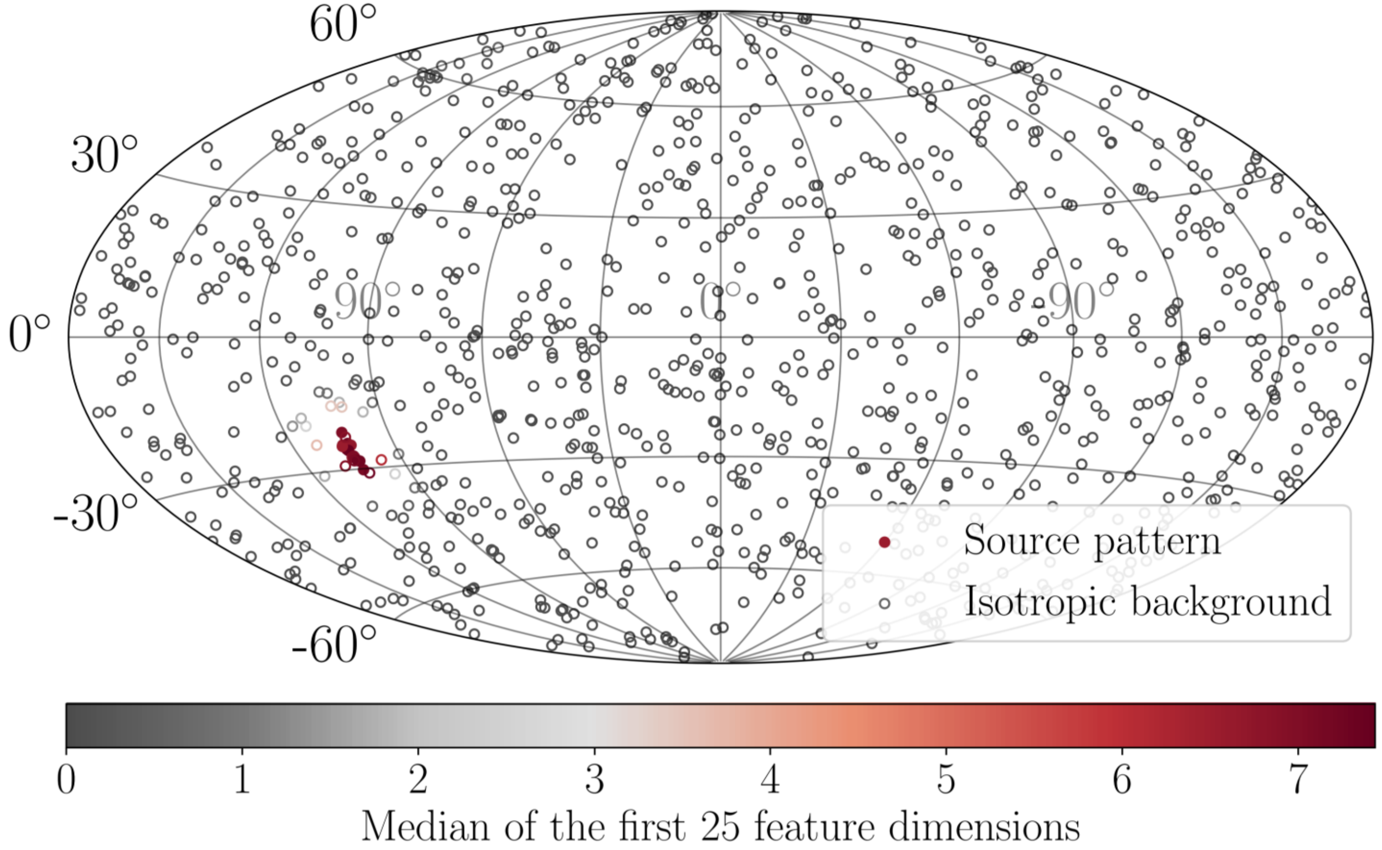
One cosmic-ray sky with $\boldsymbol{N_S = 15}$
(Shaded: 68% percentile)



Individual Classification of Cosmic Rays

Take **median** of first 25 dimensions and classify via **threshold** of 3





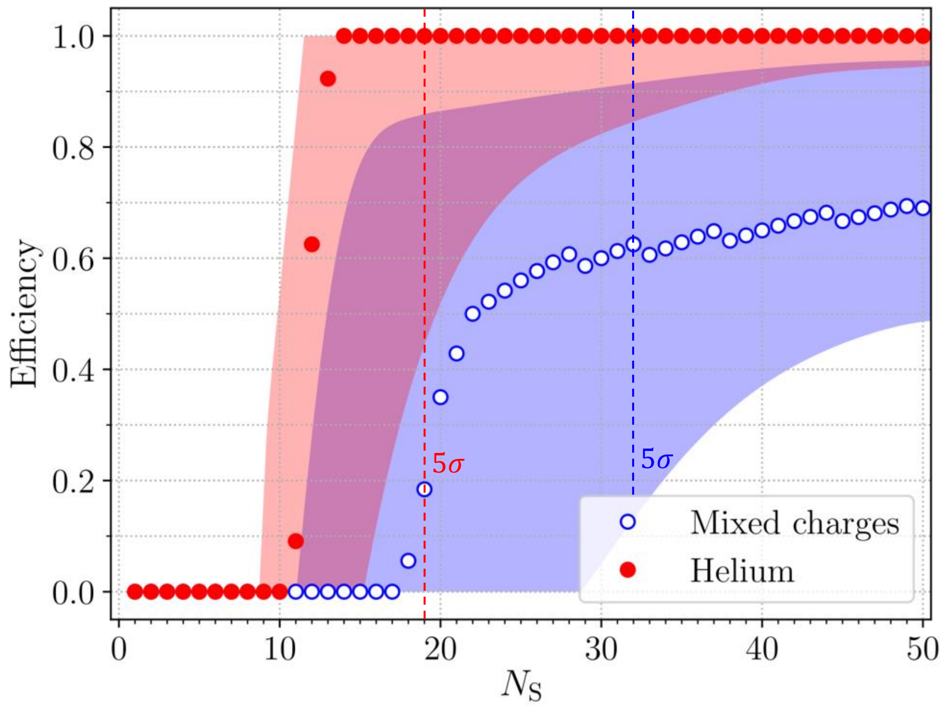
Individual Classification of Cosmic Rays

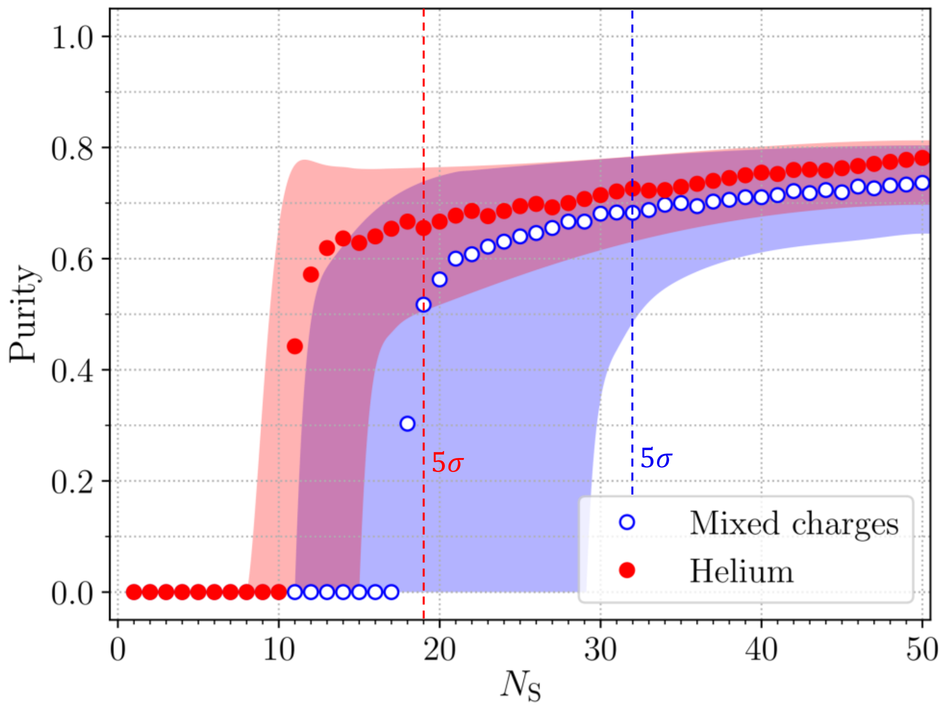
Efficiency:

$$\epsilon = \frac{\#\ \text{correctly identified signal cosmic rays}}{\#\ \text{signal cosmic rays}}$$

Purity:

$$\rho = \frac{\#\ \text{correctly identified signal cosmic rays}}{\#\ \text{identified signal cosmic rays}}$$

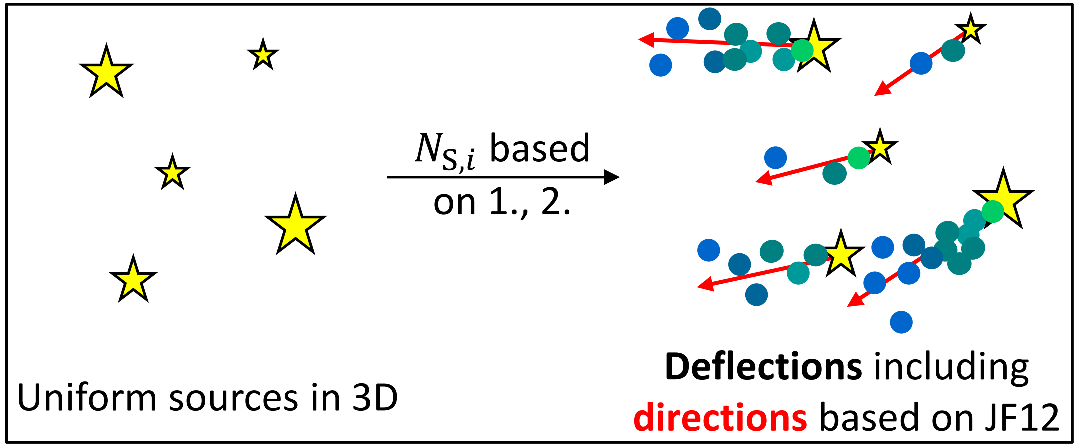




(of 1000 cosmic rays)

Multiple Sources: Simulation & Training

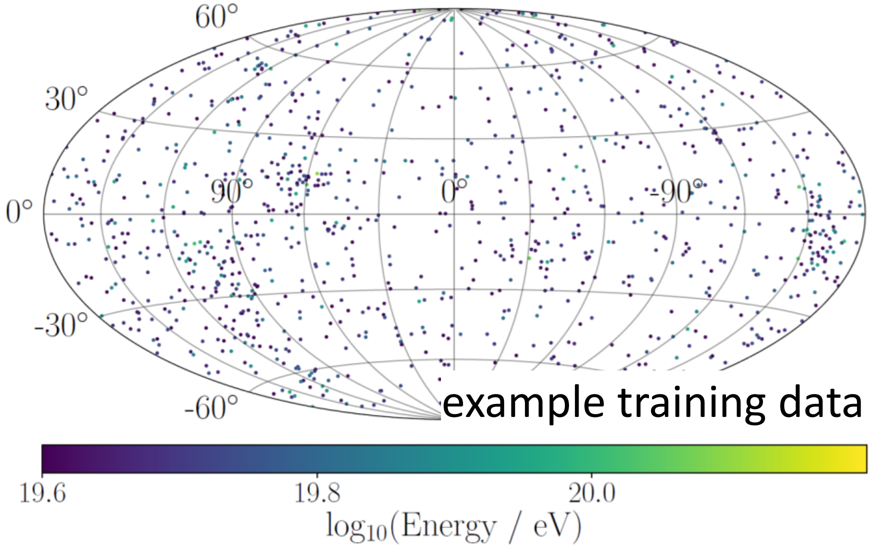
- Realistic **3D universe**: **uniformly distributed** and identical sources, accounting for:
  1. **Geometrical effect** on fluxes: $f_i \propto d_i^{-2}$
  2. **Interactions**, e.g. with photon fields
- Use parameters from **Auger Combined Fit**[1]
- Only free parameter: **source density ρ_S**
- **Apply deflection** based on JF12



[1] https://arxiv.org/abs/1612.07155

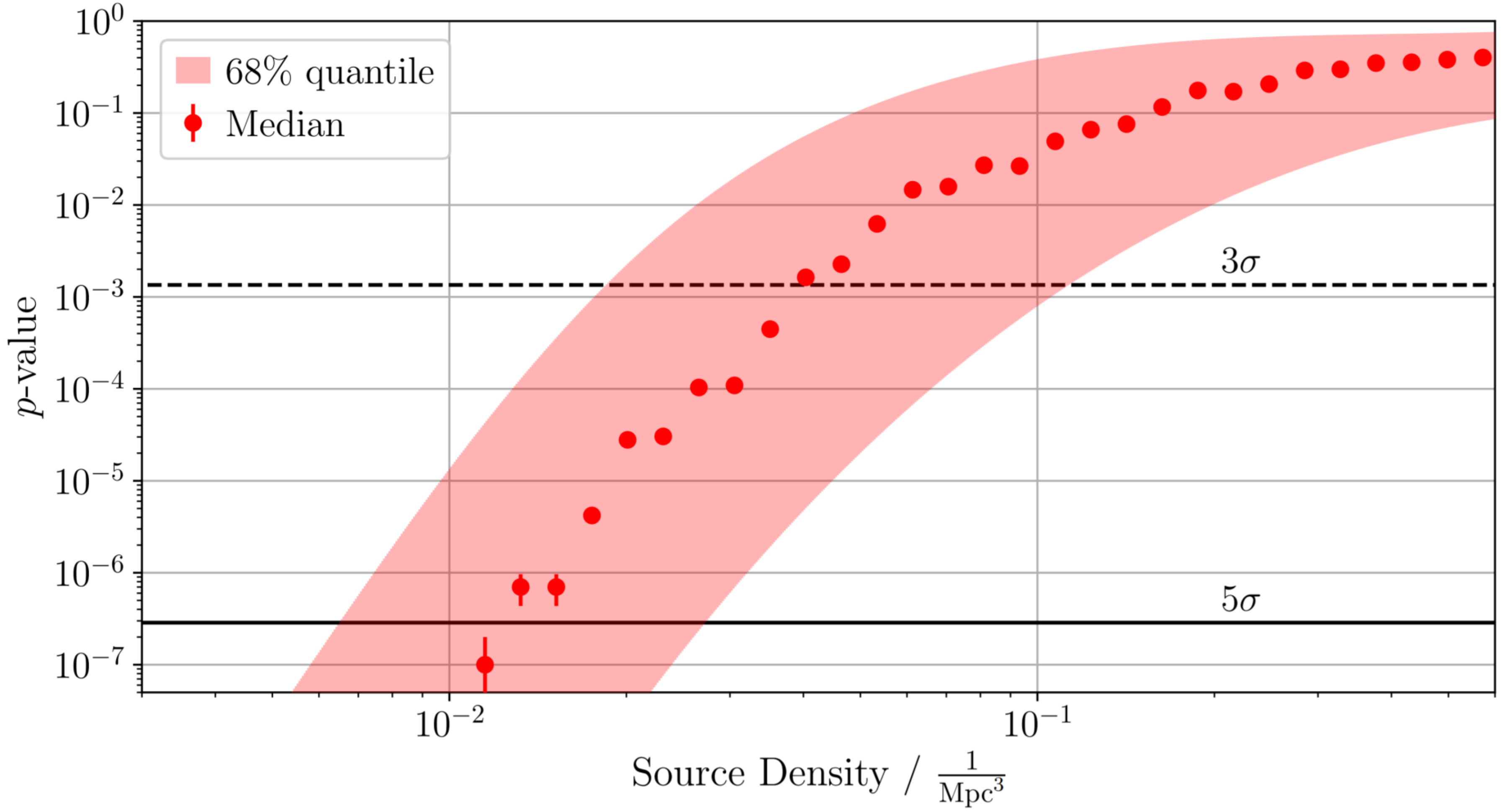
Adjust network to new simulation:

- EdgeConv-dimensions: 64/128/256
- **Concatenation: Yes**
- Train on $\rho_S = 10^{-3}/\mathrm{Mpc}^3$
- **Deflection** strengths and directions **maps** from JF12 **randomly rotated** and used for all cosmic rays to achieve global consistency



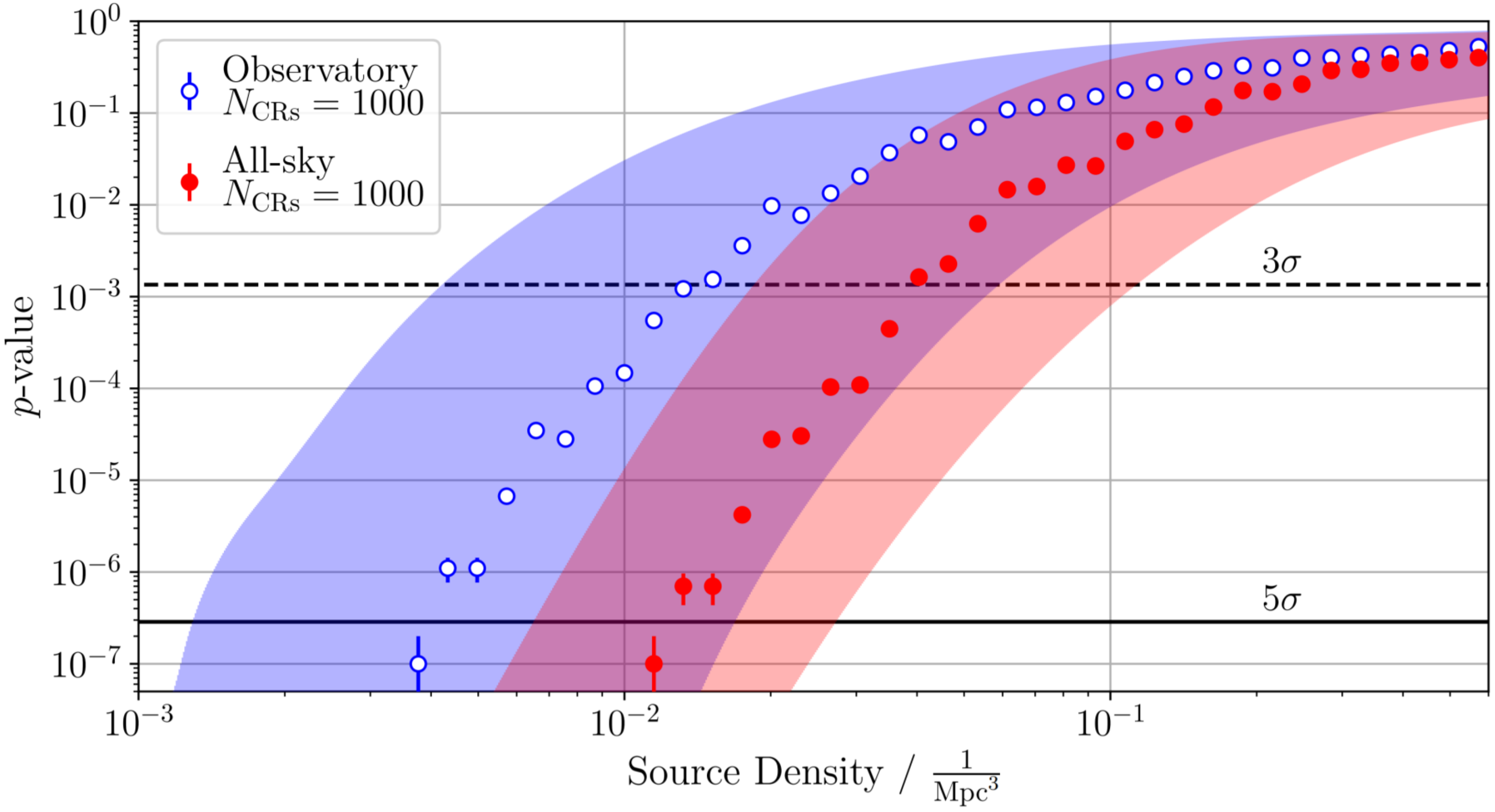
example training data

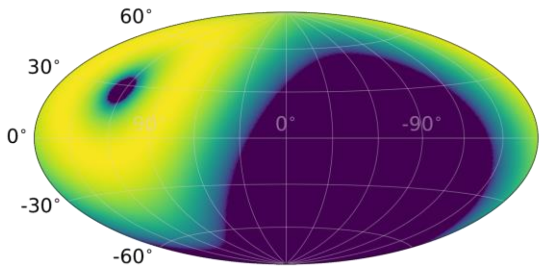
Sensitivity



5σ at $\rho_S \approx 1.2 \cdot 10^{-2}/\mathrm{Mpc}^3$

Sensitivity: Effect of Limited Sky Coverage





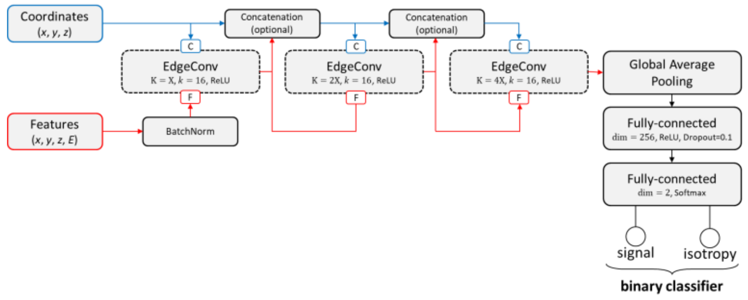
Examplary sky coverage of observatory in galactic coordinates

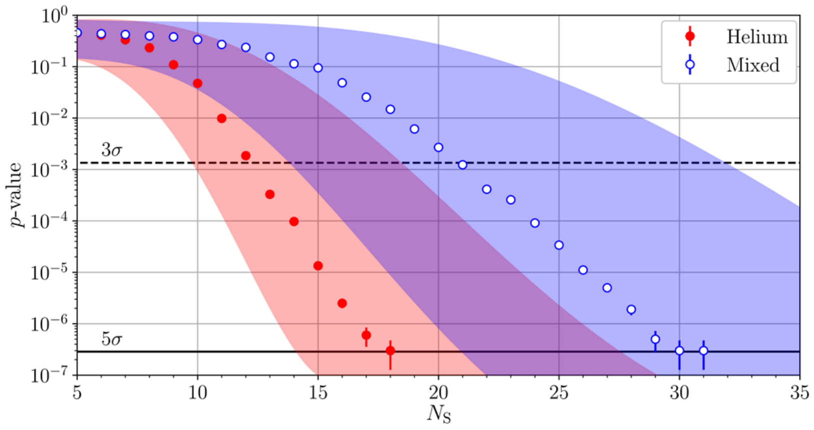
Conclusion

- (Dynamic) GCNN well-suited for **sparse cosmic-ray data**, efficiently using all available information
- Network automatically performs **clustering**

- High **sensitivity for a single pattern** of one strong source (5σ at $f_{\mathrm{sig}} = 1.9\%$)
  - ➞ Possibility to **identify individual source cosmic rays**
- Identify the **global structure** of simulated universe (5σ at $\rho_{\mathrm{S}} \approx 1.2 \cdot 10^{-2}/\mathrm{Mpc}^3$)
  - ➞ **Constrain source density**

➞ **Great potential** for data from current observatories ($N_{\mathrm{CRs}} \sim 1000$)



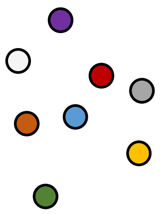


Backup

Dynamic Graph Convolutional Neural Network

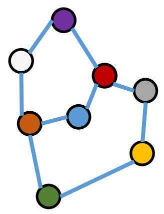
Input: Point Cloud of *N* points of dimension *M*

Our task: 4D points (*x*, *y*, *z*, *E*)



k=2

Build graph by connecting points x_i to k nearest neighbors x_{i_j}



Perform 'EdgeConv'[1] operation

$a = 1 \dots K$ (number of filters)

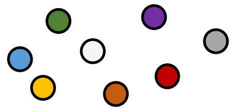
$$h^a_\Theta\left(x_{i,c}, x_{i_j,c}\right) = \sum_{c=1}^{M} \theta^a_c \, x_{i,c} + \sum_{c=1}^{M} \theta'^a_c \left(x_{i_j,c} - x_{i,c}\right)$$

⋮

Additional fully-connected layers

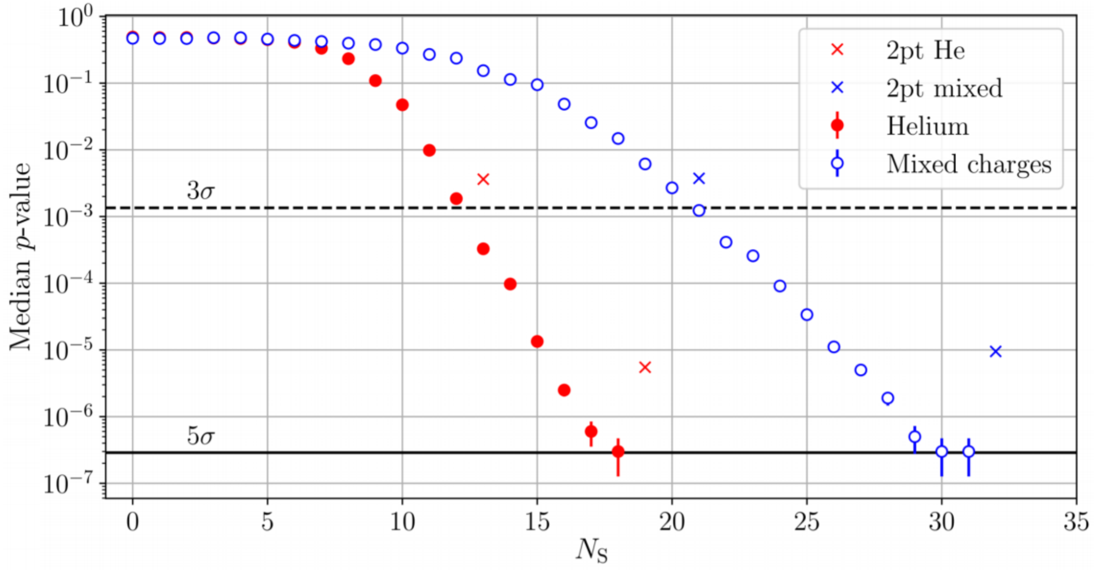
Mean

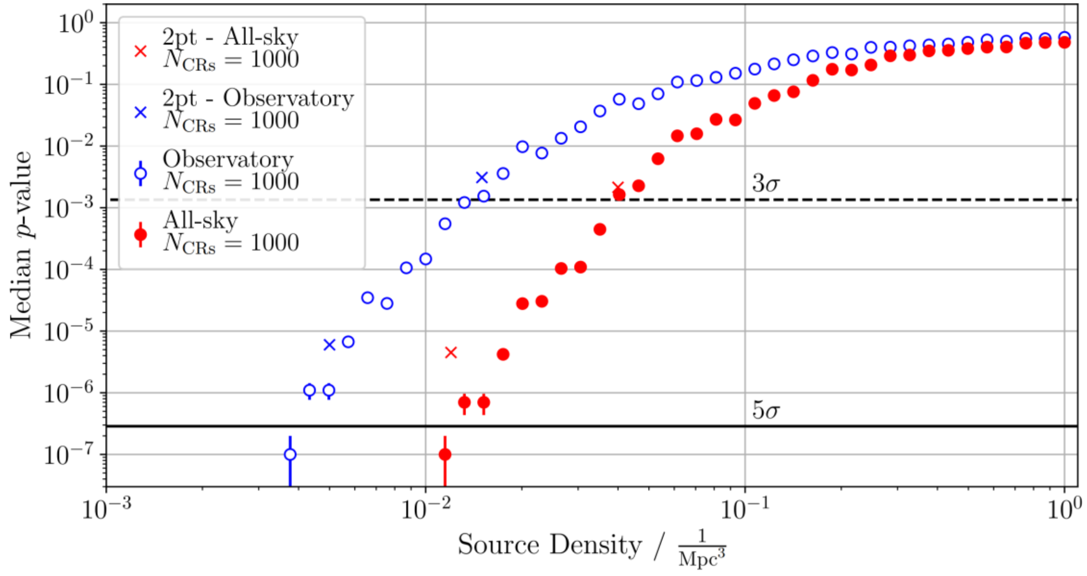
N points in *K* dimensions



[1] https://arxiv.org/abs/1801.07829

Comparison to 2pt auto-correlation





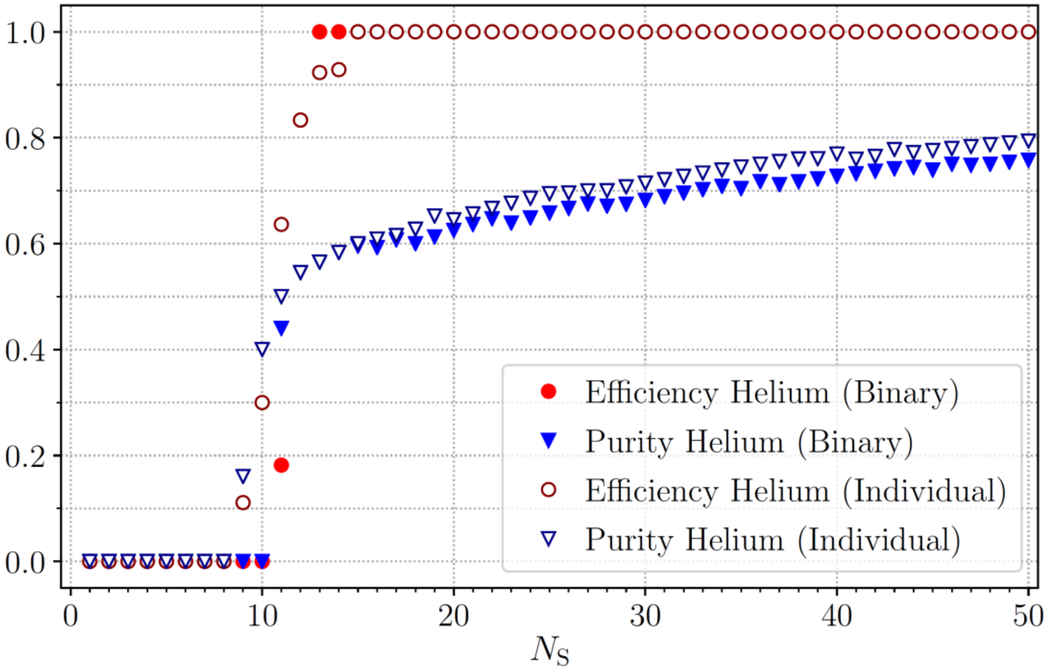
Conservative 2pt-*p*-values:

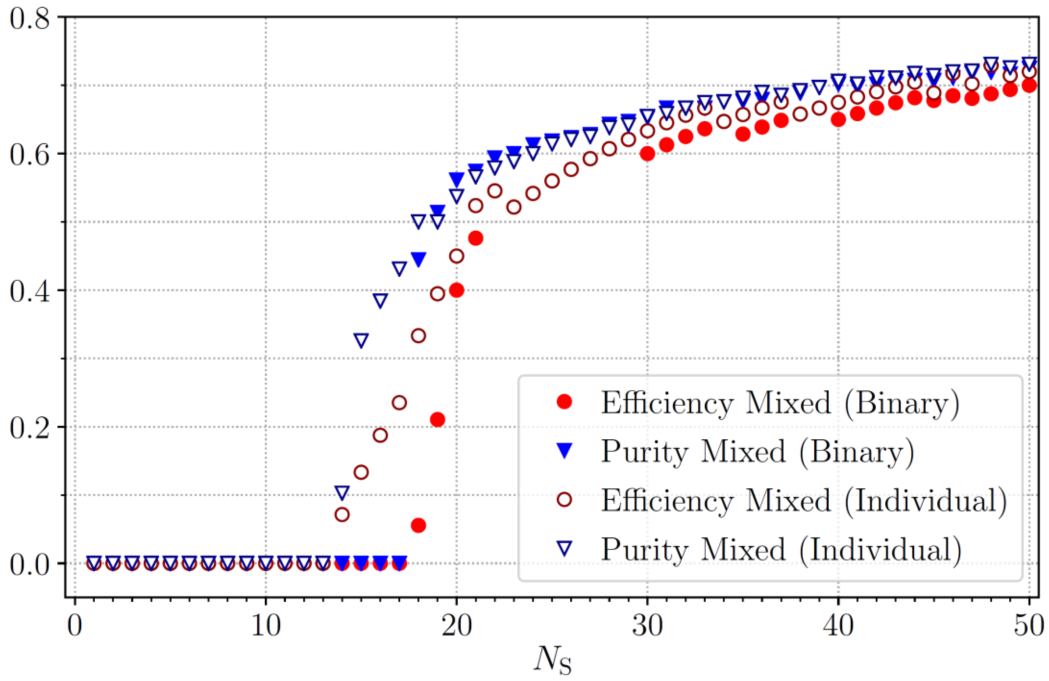
- Calculate using 180 angle-bins with 1000 skies
- Chose smallest value from all bins for each sky
- Calculate median of those 1000 values

Individual Classifier

Comparison of strategies to achieve classification of individual cosmic rays

1. 'Binary': Binary cosmic-ray sky classifier with threshold
2. 'Individual': Network explicitly trained to classify individual CRs (1000 outputs)

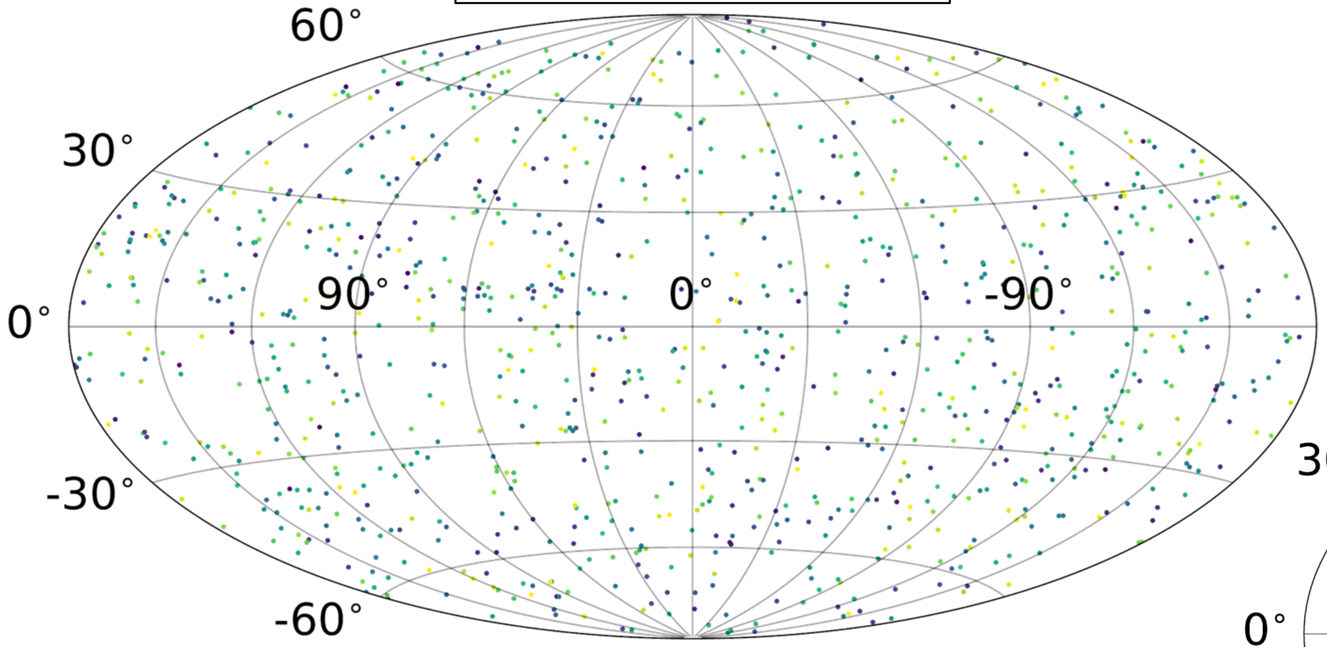




Representative skymaps

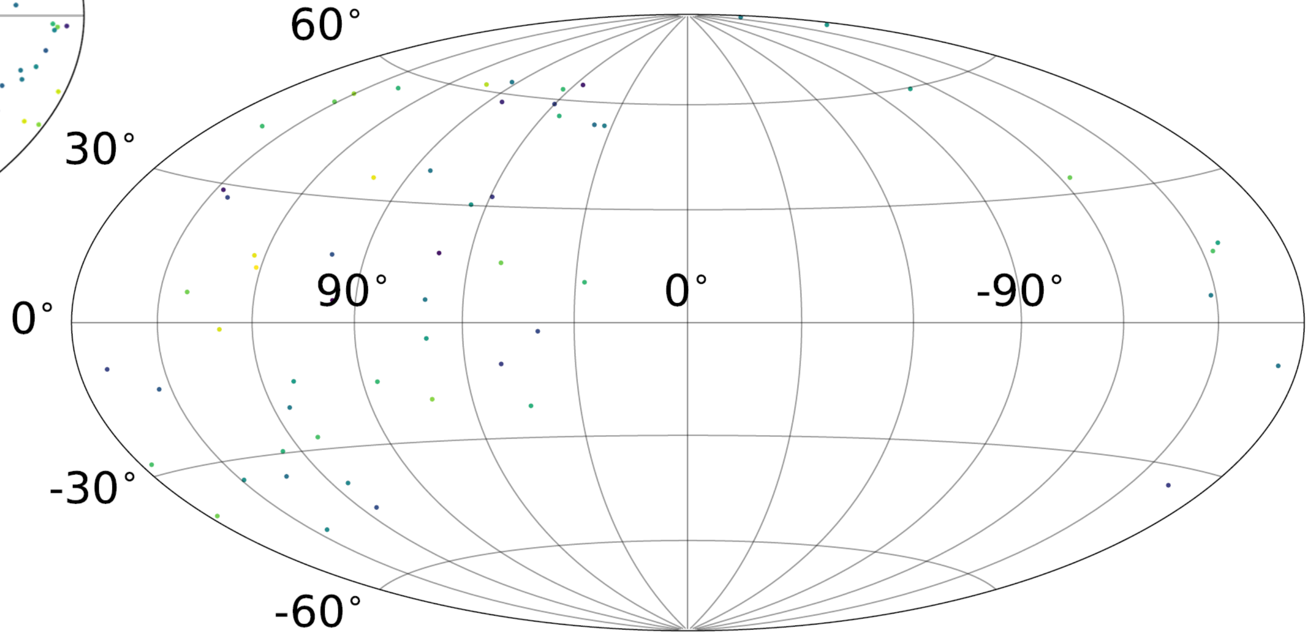
(Cosmic-ray skies that result in the median network-response of a given source density.)

1000 CRs, no exposure



$10^{-1}/\mathrm{Mpc}^3$

57 CRs, exposure

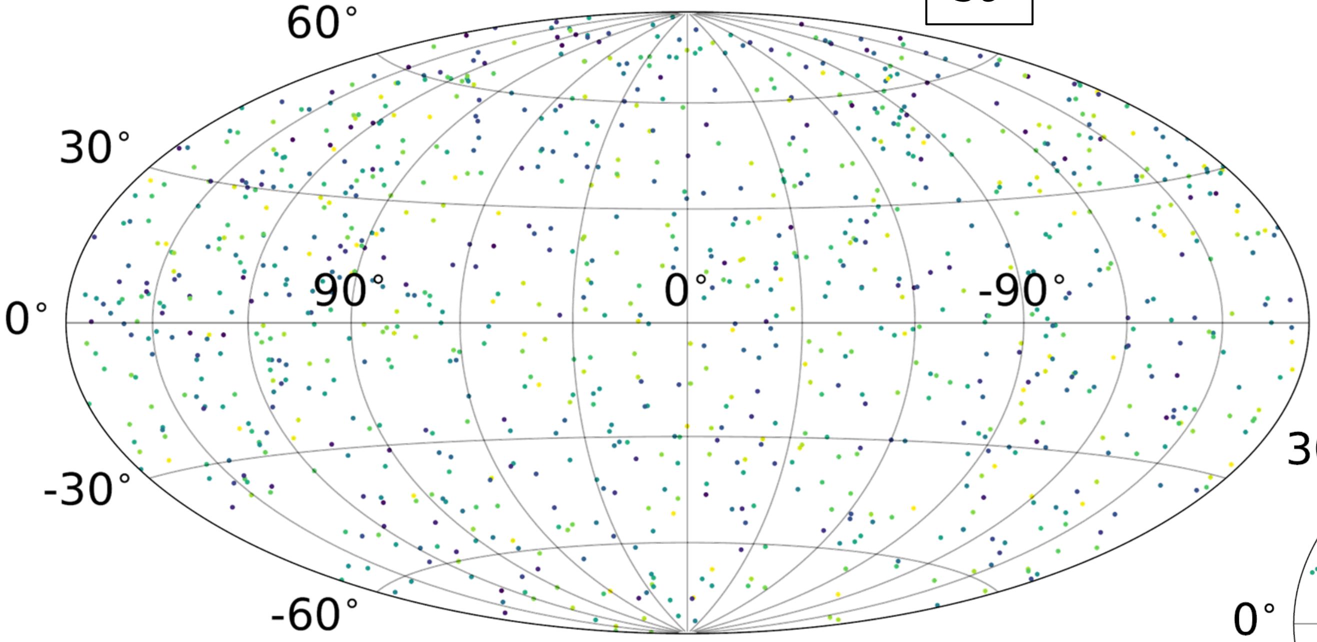


Representative skymaps

(Cosmic-ray skies that result in the median network-response of a given source density.)

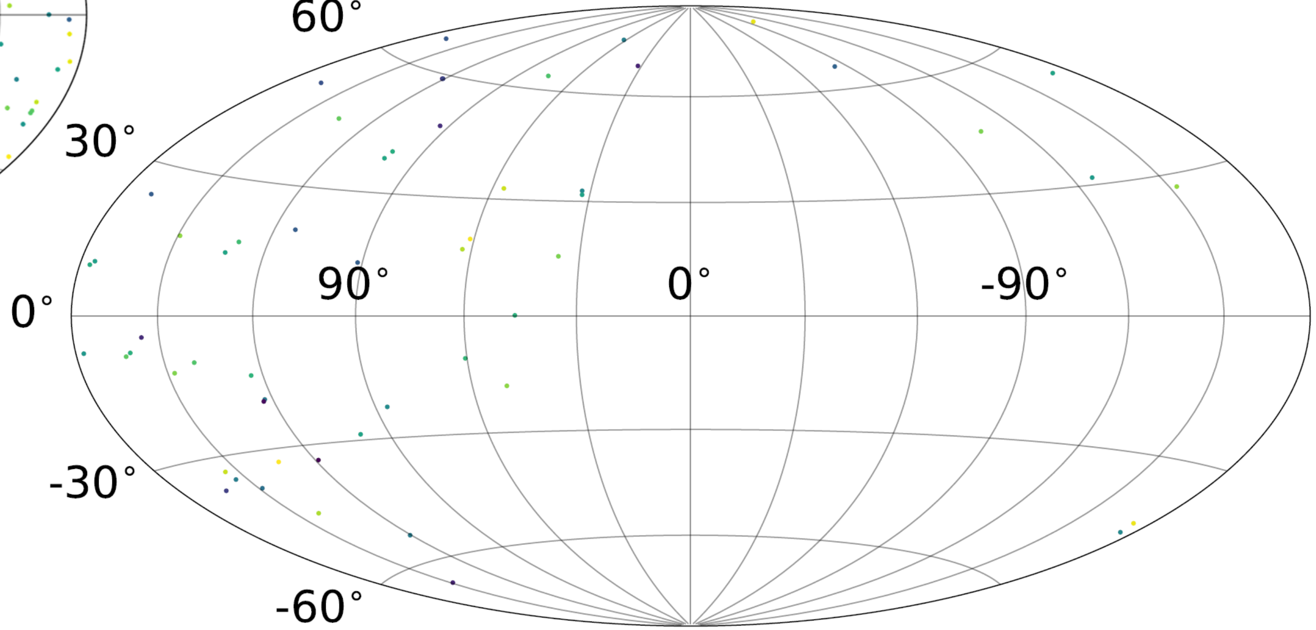
1000 CRs, no exposure

5σ



$10^{-2}/\mathrm{Mpc}^3$

57 CRs, exposure

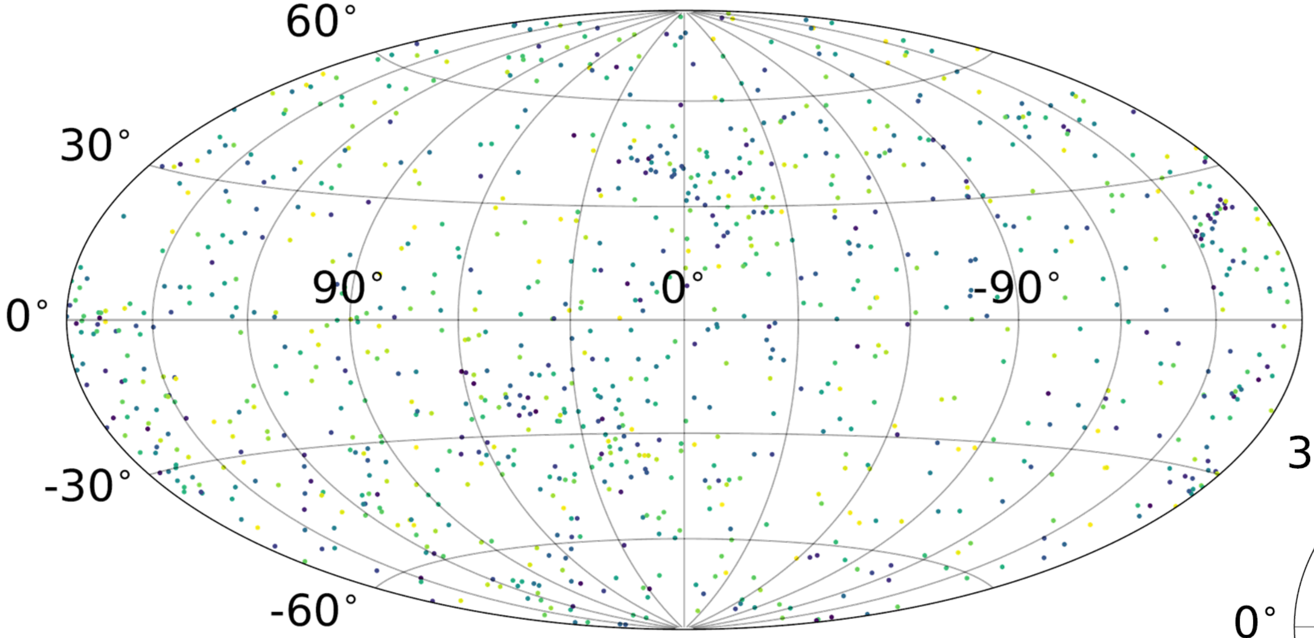


Representative skymaps

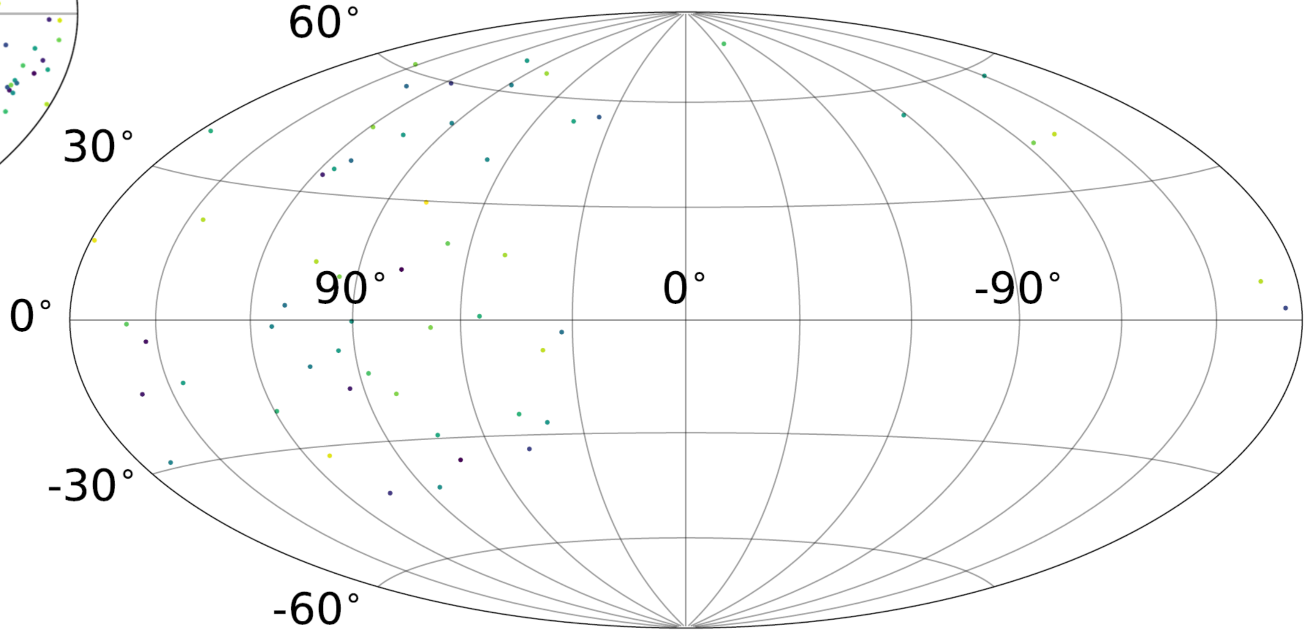
(Cosmic-ray skies that result in the median network-response of a given source density.)

$10^{-3}/\mathrm{Mpc}^3$

1000 CRs, no exposure



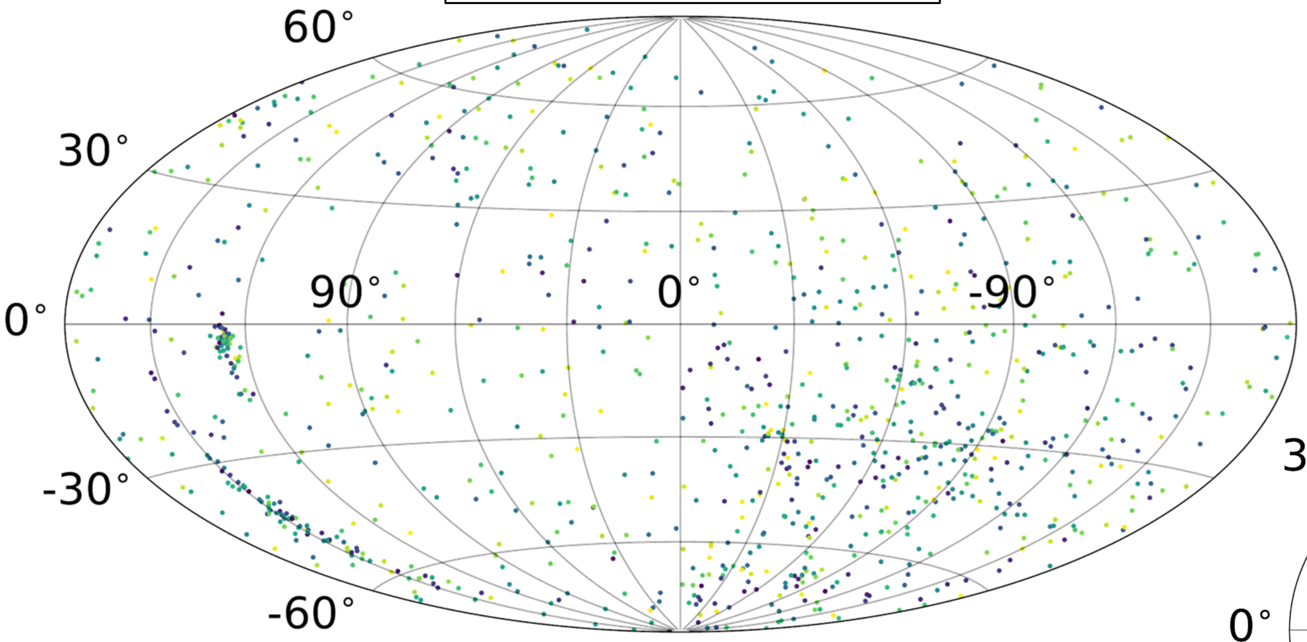
57 CRs, exposure



Representative skymaps

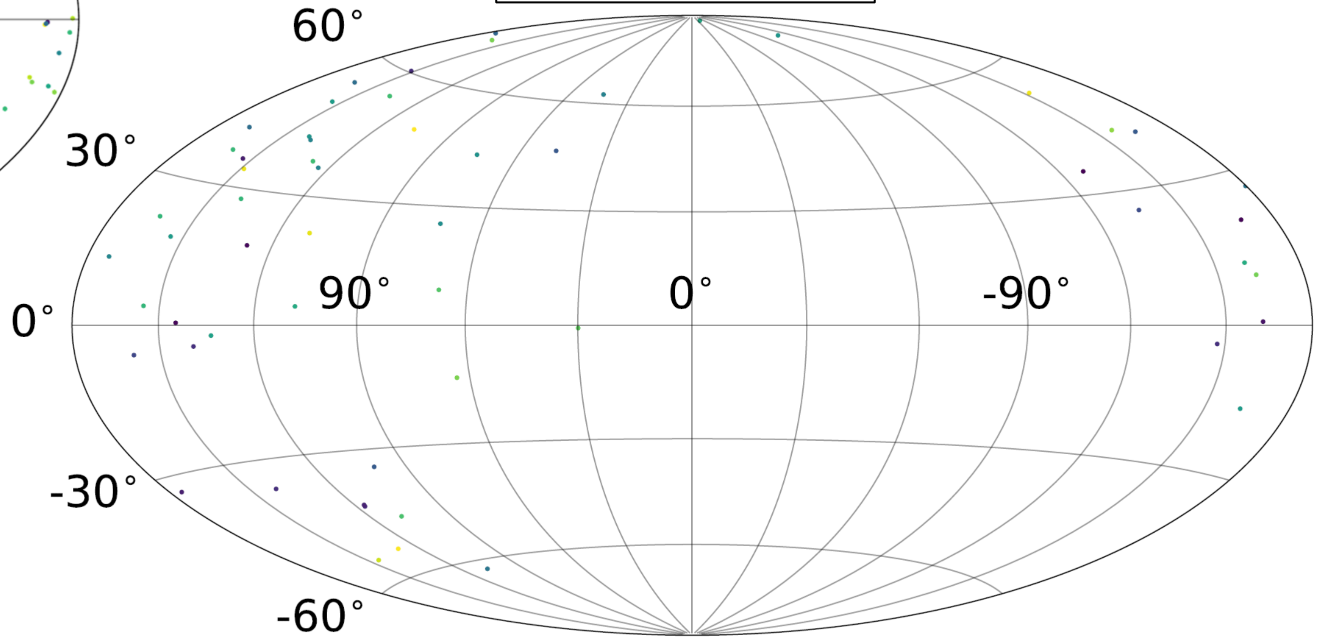
(Cosmic-ray skies that result in the median network-response of a given source density.)

1000 CRs, no exposure



$10^{-4}/\mathrm{Mpc}^3$

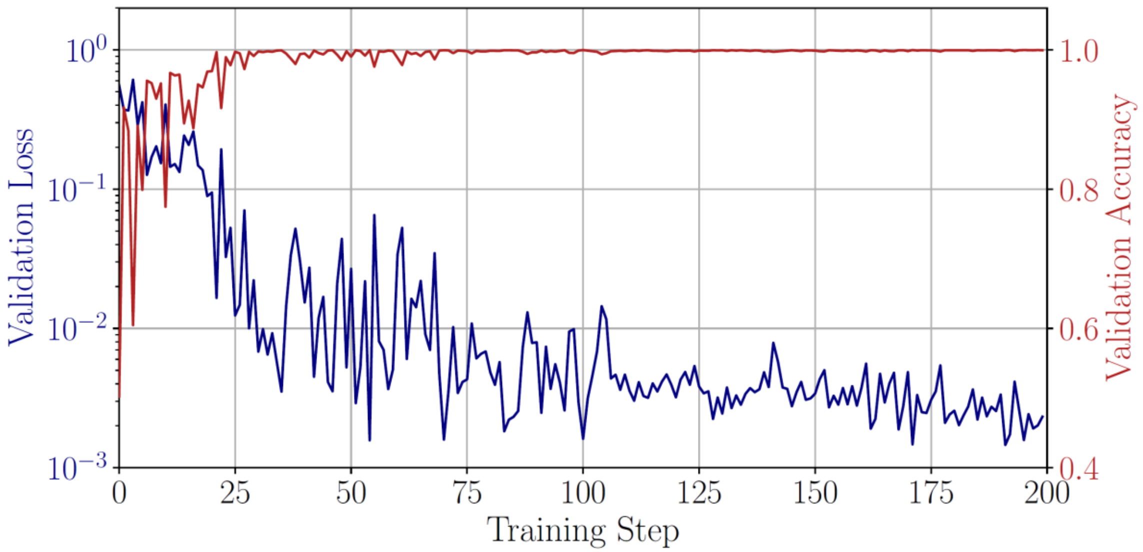
57 CRs, exposure



Training on simulation of multiple sources

- Based on Auger (parameters from Combined Fit): 1000 cosmic rays with $E > 40$ EeV,
- Simulate on the fly during training → no overfitting
- Train on $\rho_S = 10^{-3}/\mathrm{Mpc}^3$
  or $\rho_S = 10^{-2}/\mathrm{Mpc}^3$ (with exposure)
- Deflection strengths and directions maps from JF12 randomly rotated and used for all cosmic rays
- Turbulent deflection: 50% of maximum during training

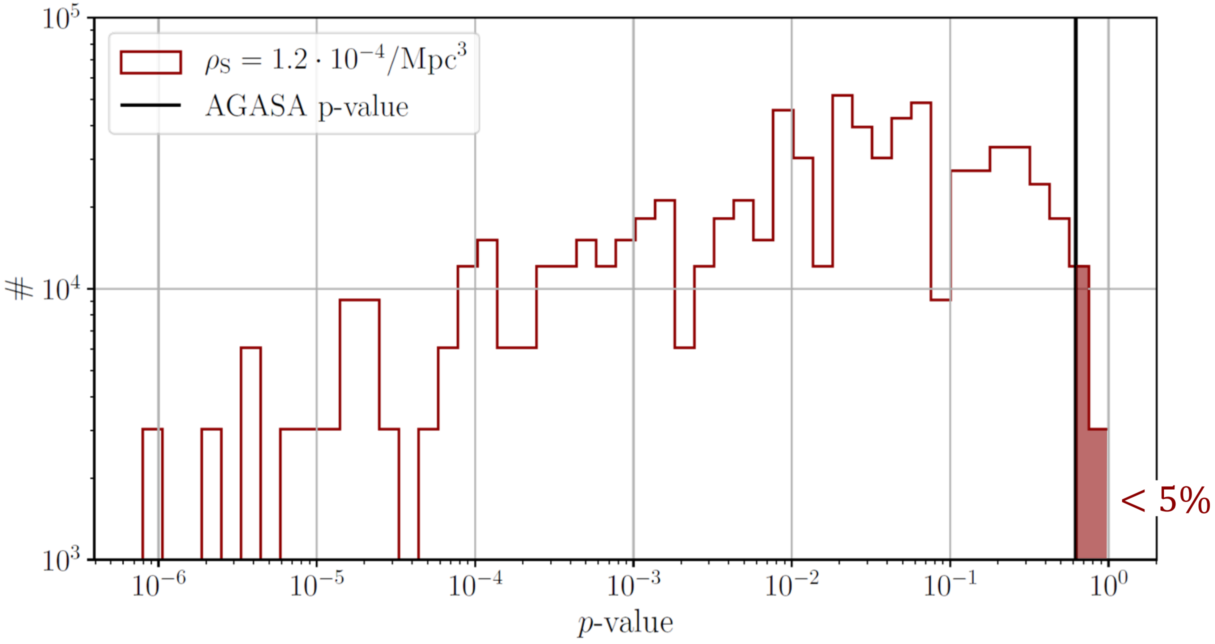
EdgeConv dims 64/128/256	Loss Categorical cross entropy
Optimizer Adam	Concatenation Yes

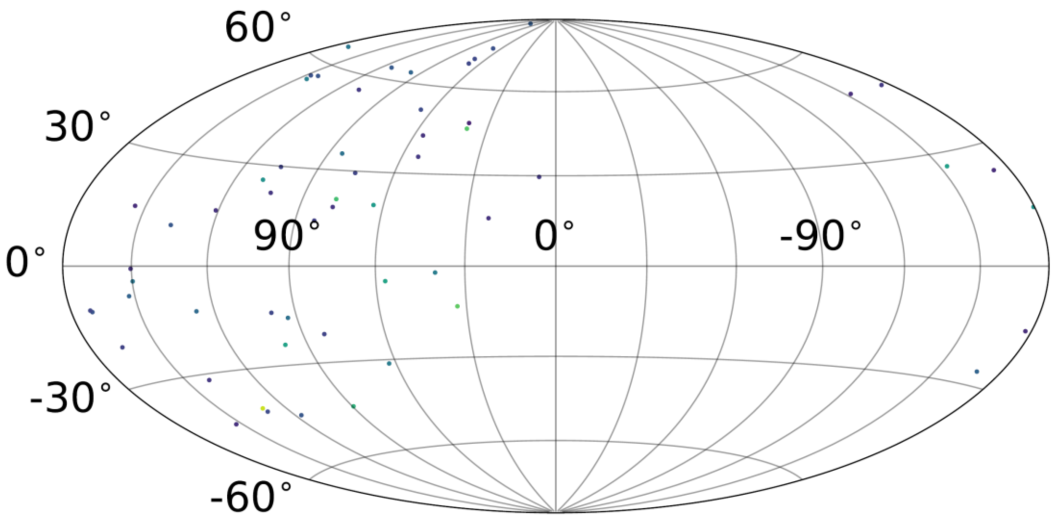


Source density limit from AGASA data

Akeno Giant Air Shower Array (**AGASA**): cosmic ray observatory, in operation from 1990 to 2007

1. Calculate ***p*-value** using equatorial scrambling as isotropy
   $\rightarrow \mathbf{p = 0.82}$
2. Take ***p*-value distribution** for varying ρ_S and determine the **probability of** $\mathbf{p \geq 0.82}$





public AGASA dataset

At $\boldsymbol{\rho_S^{95} = 2 \cdot 10^{-3}/\mathrm{Mpc}^3}$: $\mathrm{p} \geq 0.82$ for 5%

$\rightarrow$ Lower limit of ρ_S at 95% confidence