Testing UHECR origin hypotheses using deep learning

Oleg Kalashev

INR, Moscow

JCAP 11 (2020) ; arXiv:2105.06414

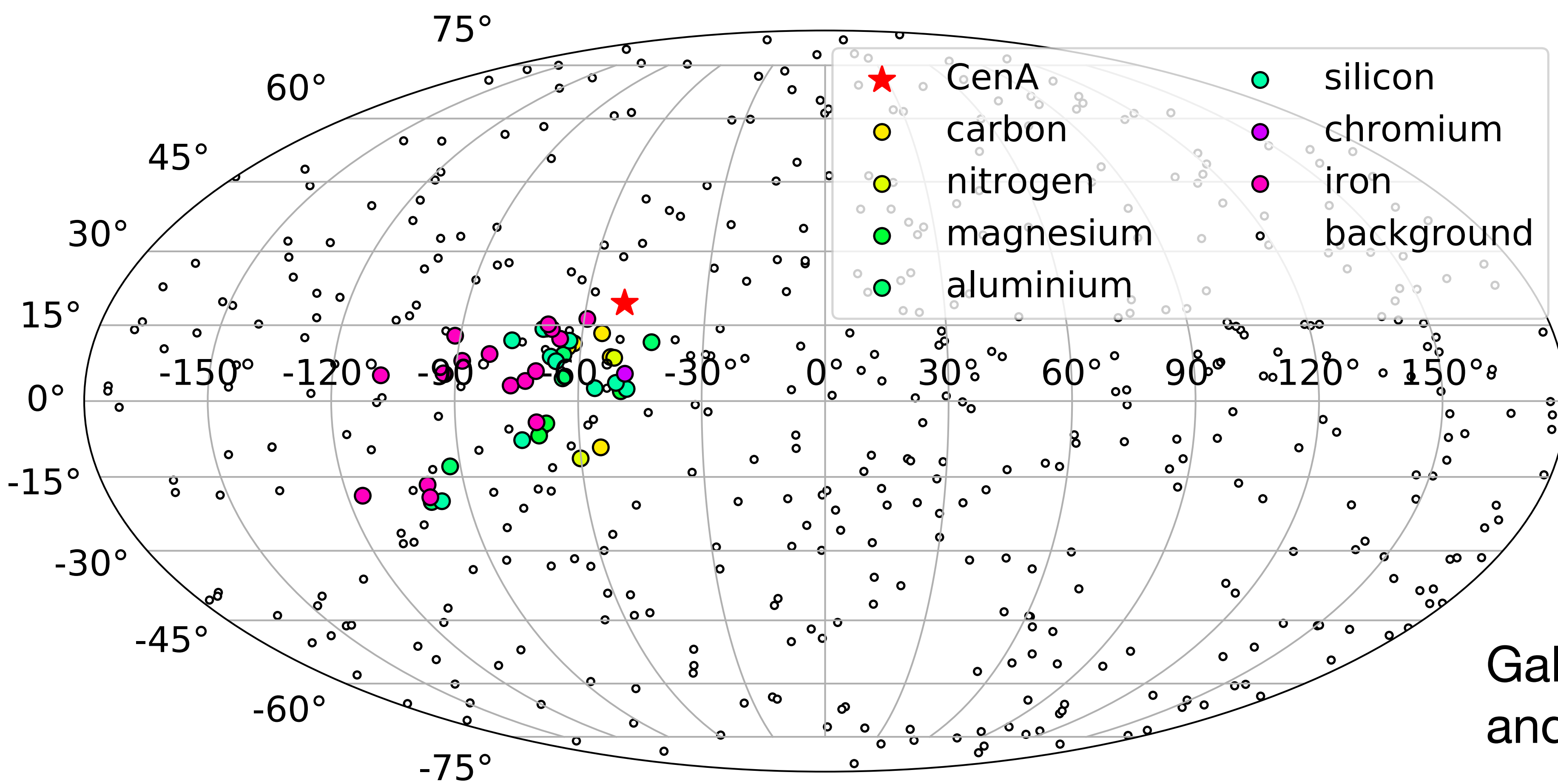
O.K., Maxim Pshirkov, Mikhail Zotov

Advanced Computing in Particle Physics, June 8-9, 2021

Problem

study capability of future orbital UHECR detectors like K-EUSO to reveal anisotropy

- null-hypothesis: isotropic flux
- alternative hypothesis: (subclass of) AGN as UHECR sources
  η - fraction of the events from the nearest source; few nearest candidates, galactic magnetic field model
- observables: $\theta_i, \phi_i, i = 1..N$ for cosmic rays with energy above 57 EeV



The case of $N_{\mathrm{UHECR}} = 500$ and 9% events coming from Cen A.

If CR are mostly heavy nuclei Deflection angles in GMF: tens of degrees

Galactic MF model by R. Jansson and G. R. Farrar 2012

Injected spectrum:
M. Kachelrieß et al 2017

Alternative Hypothesis

M. Kachelrieß, O. Kalashev, S. Ostapchenko, D.V. Semikoz “A minimal model for extragalactic cosmic rays and neutrinos,” PRD 96, 083006 (2017); arXiv:1704.06893.

Basic assumptions:

- UHECRs are accelerated by (a subclass of) AGN
- the energy spectra of nuclei after the acceleration phase follow a power-law with a rigidity-dependent cutoff: $j_{\rm inj}(E) \propto E^{-\alpha} \exp[-E/(ZE_{\rm max})]$
- the CR nuclei diffuse first through a zone dominated by photo-hadronic interactions, and then they escape into a second zone dominated by hadronic interactions with gas.

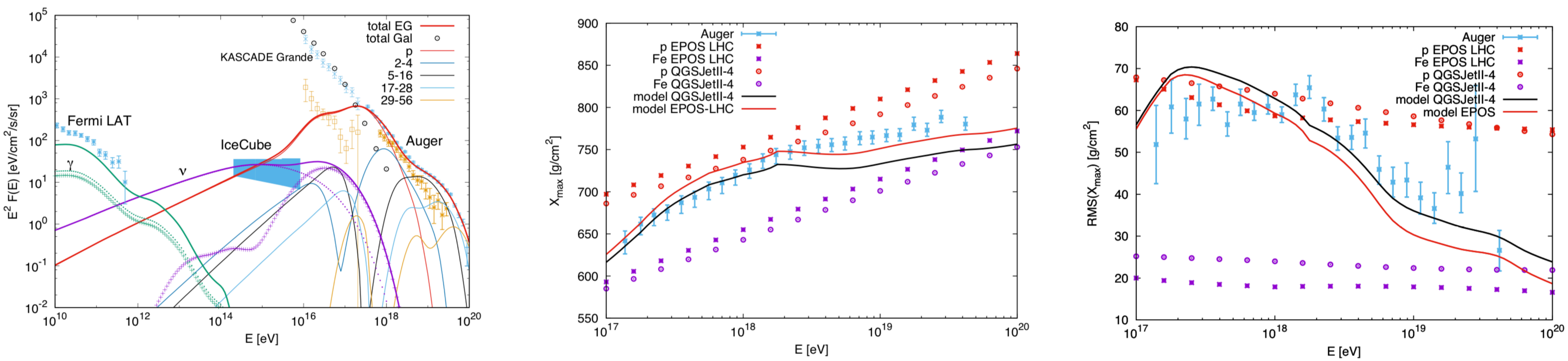
The model matches:

- Pierre Auger data on the total CR flux, the mean EAS maximum depth $X_{\rm max}$ and its width RMS($X_{\rm max}$) above $\sim 10^{17}$ eV
- HE neutrino flux measured by IceCube

One of the consequences: there should be a source of UHECRs within $\sim$ 20 Mpc

Alternative Hypothesis

M. Kachelrieß, O. Kalashev, S. Ostapchenko, D.V. Semikoz “A minimal model for extragalactic cosmic rays and neutrinos,” PRD 96, 083006 (2017); arXiv:1704.06893.



The model matches:

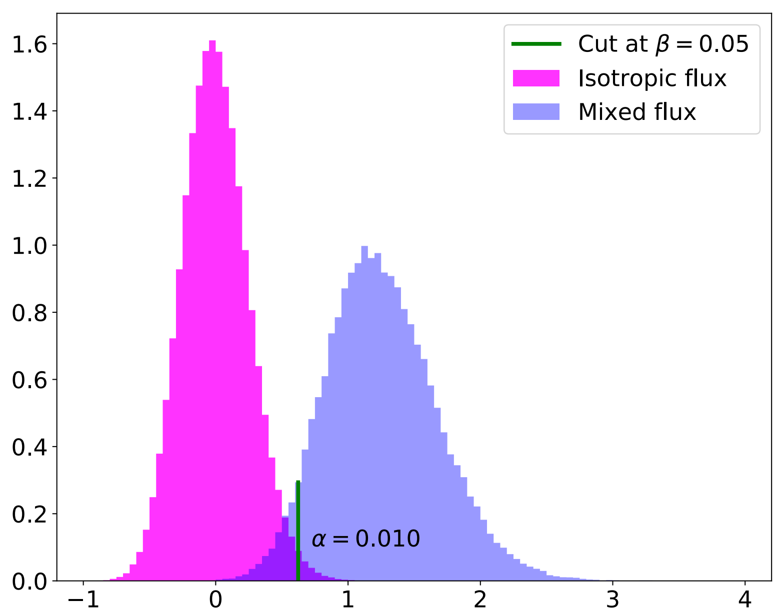
- Pierre Auger data on the total CR flux, the mean EAS maximum depth X_{max} and its width RMS(X_{max}) above $\sim 10^{17}$ eV
- HE neutrino flux measured by IceCube

One of the consequences: there should be a source of UHECRs within $\sim$ 20 Mpc

Problem

study capability of future orbital UHECR detectors like K-EUSO to reveal anisotropy

- null-hypothesis: isotropic flux
- alternative hypothesis: (subclass of) AGN as UHECR sources
  η - fraction of the events from the nearest source; few nearest candidates, galactic magnetic field model
- observables: $\theta_i, \phi_i, i = 1..N$ for cosmic rays with energy above 57 EeV
- aim: find optimal test statistic $\xi = f(\theta_i, \phi_i)$ to reject null-hypothesis if alternative hypothesis is true



Type I and II errors:

find threshold ξ_{th}

probability

II : false negative $\quad \beta = \int_{-\infty}^{\xi_{th}} p_{alt}(\xi) d\xi = 0.05$

I : false positive $\quad \alpha = \int_{\xi_{th}}^{\infty} p_{iso}(\xi) d\xi \leq 0.01$

Find minimal nearest source event fraction η to reject isotropy hypothesis

How to choose $\xi = f(\theta_i, \phi_i)$

- **Use angular power spectrum** $\xi = f(C_l)$
  - **e.g.**
    - IceCube, Auger [Hülss, Wiebusch, ICRC2007; Aab et al. JCAP 06 (2017) 026]:

$$\xi \propto \sum_{\ell=1}^{\ell_{max}} \left(\frac{C_\ell - \langle C_{\ell,\mathrm{iso}} \rangle}{\sigma_{l,iso}} \right)^2$$

    - O.K., Pshirkov, Zotov 2018

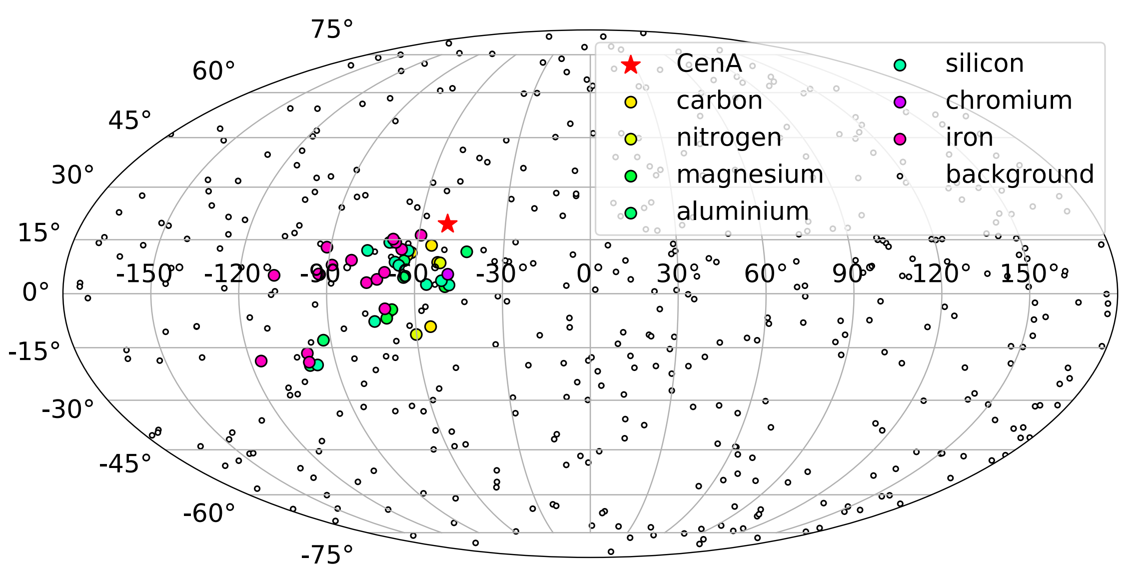
$$\xi \propto \sum_{\ell=1}^{\ell_{max}} \frac{C_\ell - \langle C_{\ell,\mathrm{iso}} \rangle}{\sigma_{l,iso}}$$

- **Use other function of direct observables**
  - **e.g.**
    - function using expected arrival direction density maps (TA correlation with LSS analysis)
    - functions obtained using machine learning ML (this talk)
      - use ML to build classifier which discriminates samples derived assuming null or alternative hypothesis
      - use the classifier output as test statistic ξ

Arrival directions map calculation

O.K., Phirkov, M. Zotov 2018

0. K-EUSO is expected to register from 120 to 500 CRs at $E \gtrsim 57$ EeV in 2 years
   [M. Casolino *et al*. PoS (ICRC2017) 368]
1. Assume EECRs propagate from a source to the Galaxy in the ballistic regime, so that deflections from the direction to the source are $\leq 2°$.
2. Take a spectrum at the source and obtain a spectrum after propagating to the Galaxy: `TransportCR` code
   [O. Kalashev, E. Kido, J. Exp. Theor. Phys. 120 (2015) 790; arXiv:1406.0735]
3. Propagate CRs from the boundary of the Galaxy to the Earth with `CRPropa3`
   [R. Alves Batista *et al*. JCAP 05 (2016) 038; arXiv:1603.07142]
   assuming the Jansson–Farrar (2012) GMF model (actually backtrack)



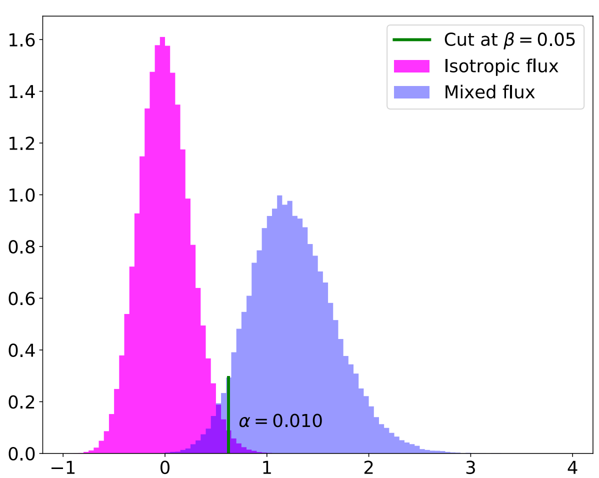
The case of $N_{\rm UHECR} = 500$ and 9% events coming from Cen A.

Previous study

O.K., Phirkov, M. Zotov 2018

0. K-EUSO is expected to register from 120 to 500 CRs at $E \gtrsim 57$ EeV in 2 years
   [M. Casolino *et al*. PoS (ICRC2017) 368]
1. Assume EECRs propagate from a source to the Galaxy in the ballistic regime, so that deflections from the direction to the source are $\leq 2°$.
2. Take a spectrum at the source and obtain a spectrum after propagating to the Galaxy: `TransportCR` code
   [O. Kalashev, E. Kido, J. Exp. Theor. Phys. 120 (2015) 790; arXiv:1406.0735]
3. Propagate CRs from the boundary of the Galaxy to the Earth with `CRPropa3`
   [R. Alves Batista *et al*. JCAP 05 (2016) 038; arXiv:1603.07142]
   assuming the Jansson–Farrar (2012) GMF model (actually backtrack)
4. Study large-scale anisotropy of CRs with the angular power spectrum.

$$\xi = D \propto \sum_{\ell=1}^{\ell_{max}} \frac{C_\ell - \langle C_{\ell,\mathrm{iso}} \rangle}{\sigma_{l,iso}}$$



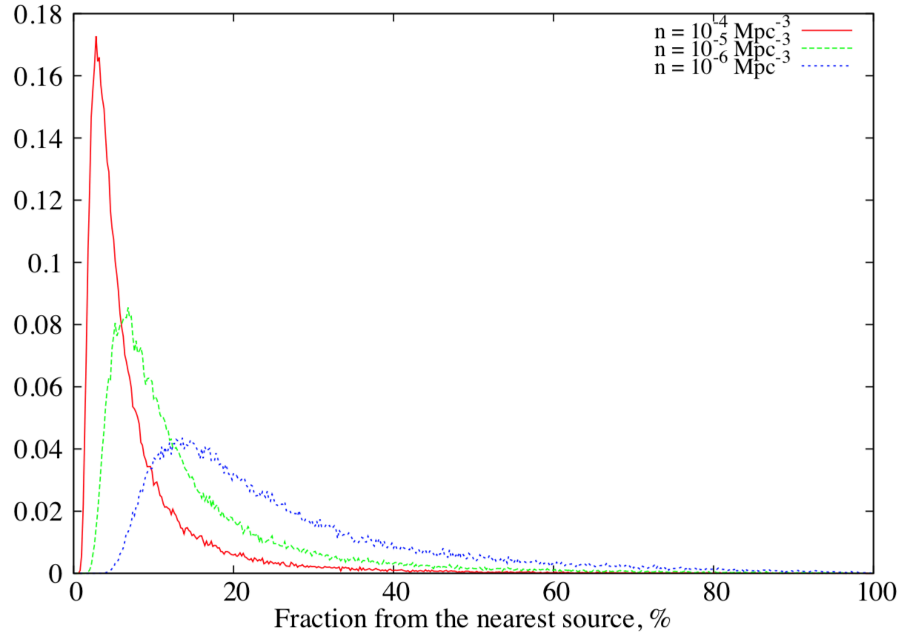
Minimal nearest source event fraction in % to reject isotropy hypothesis:

N_{UHECR}	100	200	300	400	500
NGC 253	17	12	10	8	7
Cen A	21	14	12	10	9
M82	26	18	14	12	11
M87	29	20	16	14	12
Fornax A	19	13	11	9	8

Previous study

O.K., Phirkov, M. Zotov 2018

Interpretation

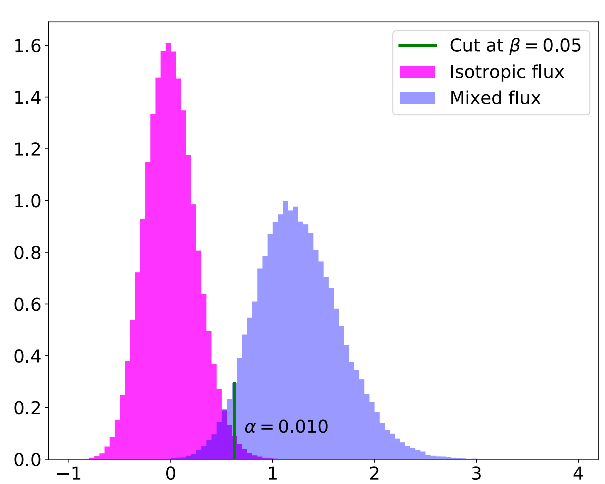


n, Mpc^{-3}	Closest	Second closest
10^{-4}	5.2	1.8
3×10^{-5}	7.5	2.7
10^{-5}	10.6	3.9
3×10^{-6}	15.0	5.0
10^{-6}	20.9	6.3

Average from-source flux fraction in % assuming identical sources

4. Study large-scale anisotropy of CRs with the angular power spectrum.

$$\xi = D \propto \sum_{\ell=1}^{\ell_{max}} \frac{C_\ell - \langle C_{\ell,\mathrm{iso}}\rangle}{\sigma_{l,iso}}$$



Minimal nearest source event fraction in % to reject isotropy hypothesis:

N_{UHECR}	100	200	300	400	500
NGC 253	17	12	10	8	7
Cen A	21	14	12	10	9
M82	26	18	14	12	11
M87	29	20	16	14	12
Fornax A	19	13	11	9	8

ML-based study

0. K-EUSO is expected to register from 120 to 500 CRs at $E \gtrsim 57$ EeV in 2 years
   [M. Casolino *et al*. PoS (ICRC2017) 368]
1. Assume EECRs propagate from a source to the Galaxy in the ballistic regime, so that deflections from the direction to the source are $\leq 2^\circ$.
2. Take a spectrum at the source and obtain a spectrum after propagating to the Galaxy: `TransportCR` code
   [O. Kalashev, E. Kido, J. Exp. Theor. Phys. 120 (2015) 790; arXiv:1406.0735]
3. Propagate CRs from the boundary of the Galaxy to the Earth with `CRPropa3`
   [R. Alves Batista *et al*. JCAP 05 (2016) 038; arXiv:1603.07142]
   assuming the Jansson–Farrar (2012) GMF model (actually backtrack)

4. Construct $\xi = f(\theta_i, \phi_i)$ using machine learning (ML)

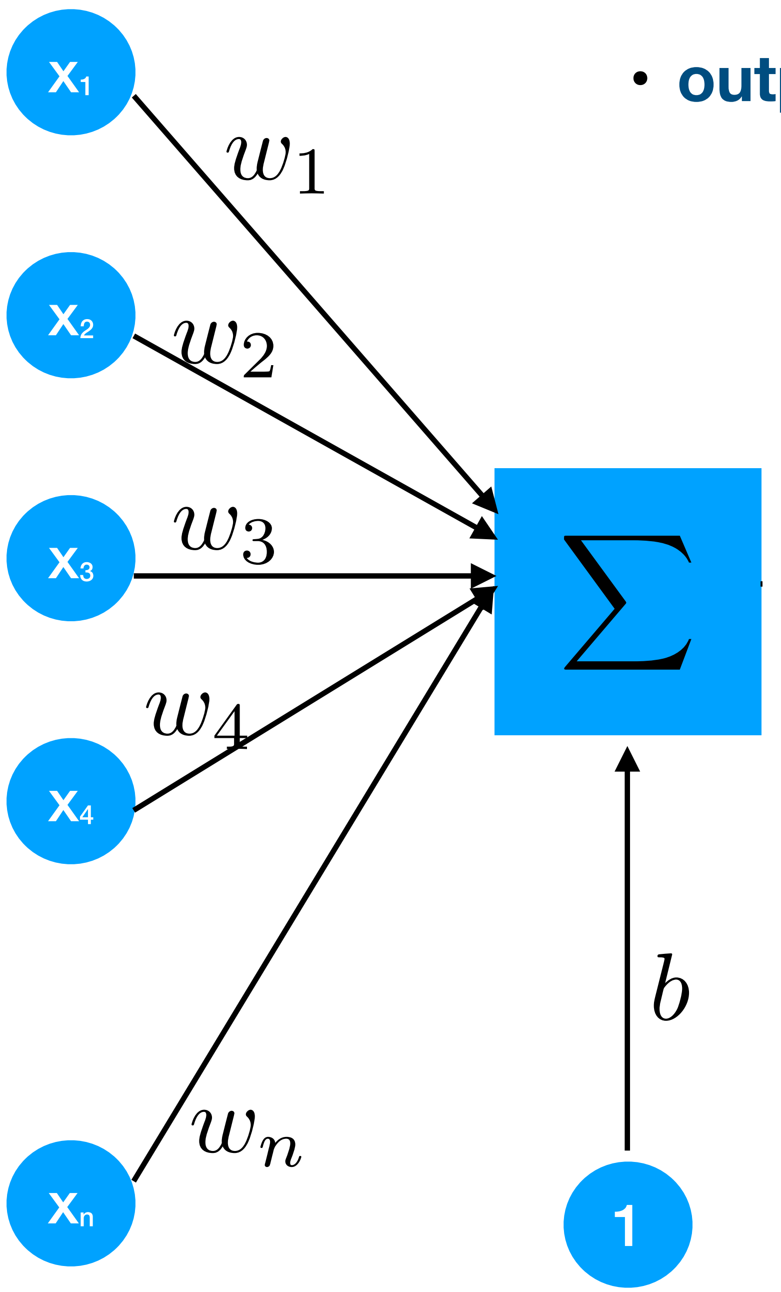
   - use ML to build classifier which discriminates samples derived assuming null or alternative hypothesis
   - use the classifier output as test statistic

Using toy ANN to build $\xi(C_l)$

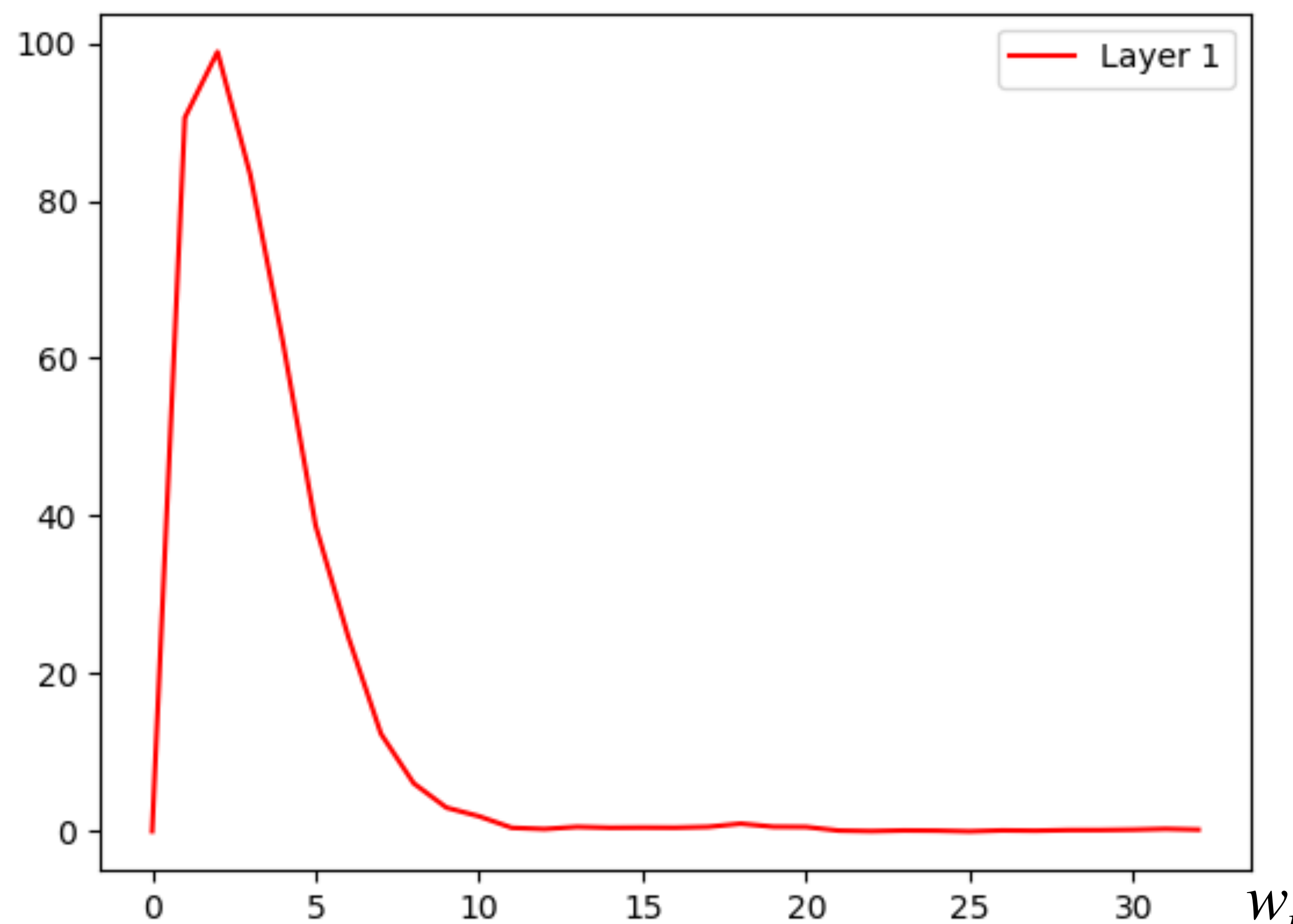
- **input features:**

$$x_i = \frac{C_l - \langle C_{l,iso} \rangle}{\sigma_{l,iso}}, l \leq 32$$

- **output:** **0 - for isotropic, 1 - for mixture**



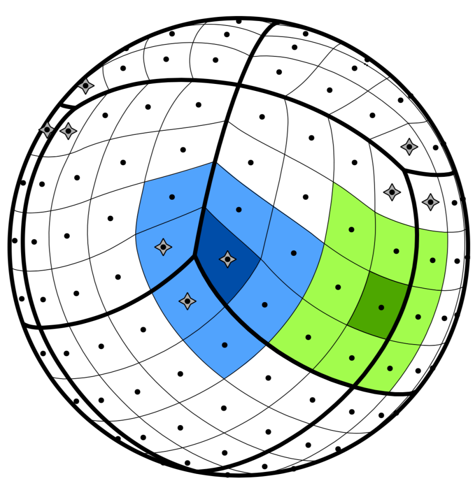
Learned weights



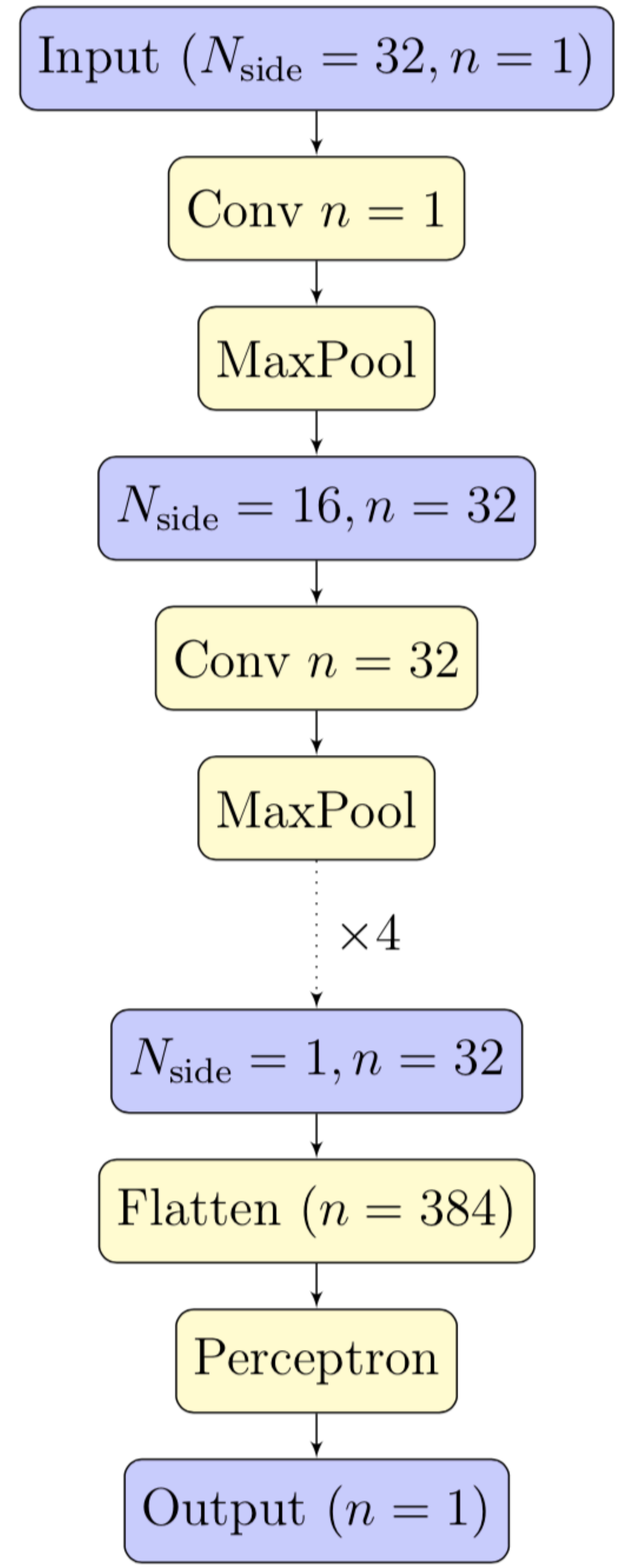
Convolutional NN on sphere

Several implementations exist:

- *Taco S. Cohen et al 2018*
- *Nathanaël Perraudin et al 2018*
- *N. Krachmalnicoff et al 2019 (used in this work)*



CNN on HEALPix map



- **input features: HEALPix map**
- **target output: 0 - for isotropic, 1 - for mixture**
- **training set consisting of 100K samples generated in mixture model distributed uniformly in log η and 100K samples generated from isotropic model**

Minimal from-source event fraction in % to reject isotropy hypothesis

Source	Method	50	100	200	300	400	500
NGC 253	APS	24	17	12	10	8	7
	CNN	12	7	4.5	3.67	3	2.6
Cen A	APS	28	21	14	12	10	9
	CNN	16	11	7	5.67	5	4.4
M 82	APS	36	26	18	14	12	11
	CNN	20	12	7	6	4.75	4.2
M 87	APS	38	29	20	16	14	12
	CNN	22	14	9	8	6.25	5.2
Fornax A	APS	28	19	13	11	9	8
	CNN	16	9	6	5	4.5	3.8

using

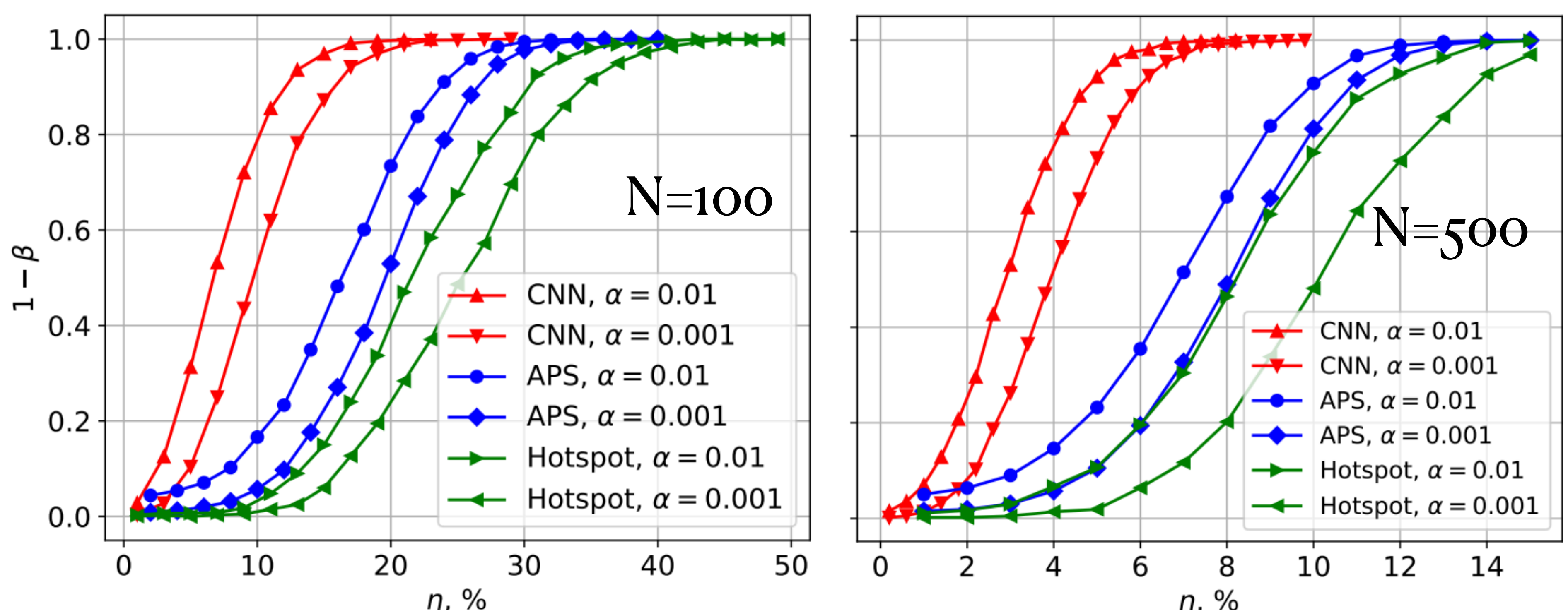
$$\xi = D \propto \sum_{\ell=1}^{\ell_{max}} \frac{C_\ell - \langle C_{\ell,\text{iso}} \rangle}{\sigma_{l,iso}}$$

single TS

using $\xi_{cnn} = f_{CNN}(\{\theta_i, \phi_i\})$ **given by the output of convolutional NN on HEALPix grid**

$$\alpha = 0.01\,, \beta = 0.05$$

multiple CNN classifiers optimized to each source



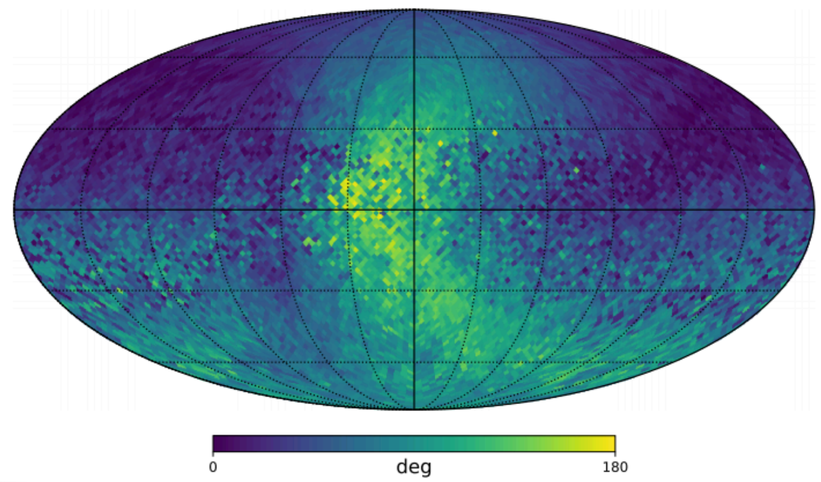
source M87

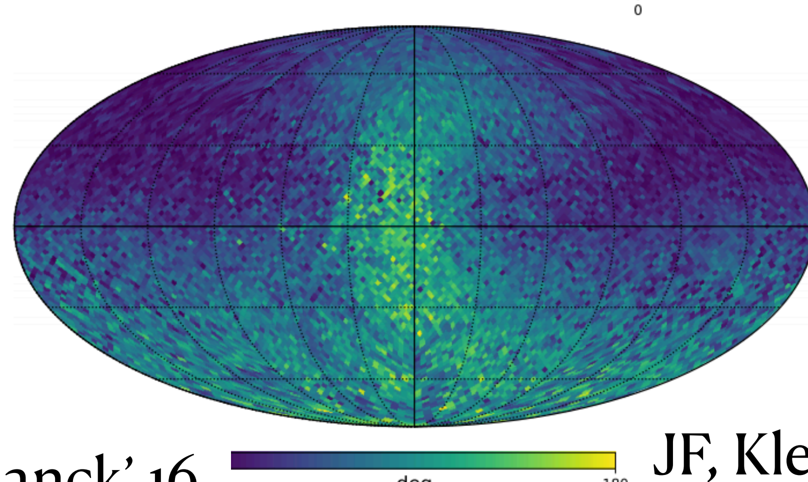
Model dependence

galactic magnetic field

Deflections of 100 EeV Fe nuclei:

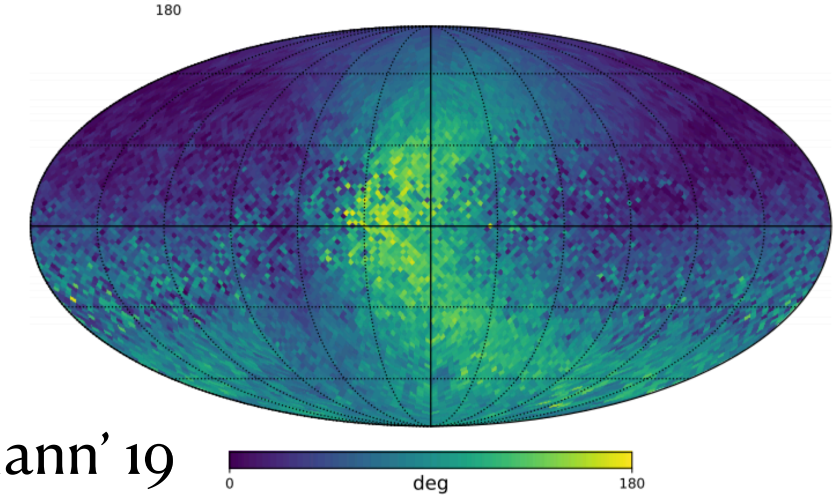
Jansson–Farrar '12

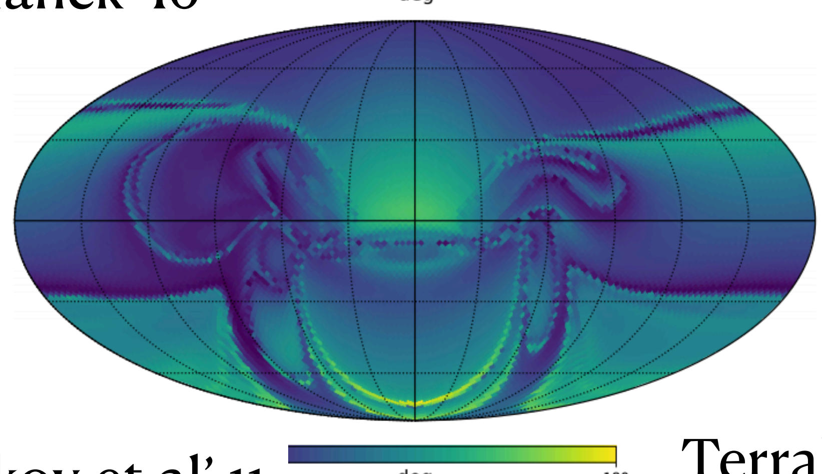




JF Planck' 16

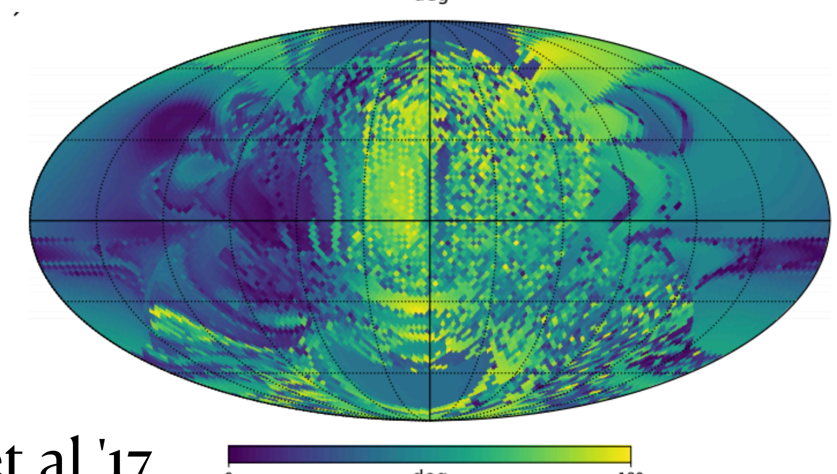
JF, Kleimann' 19





Pshirkov et al' 11

Terral et al '17



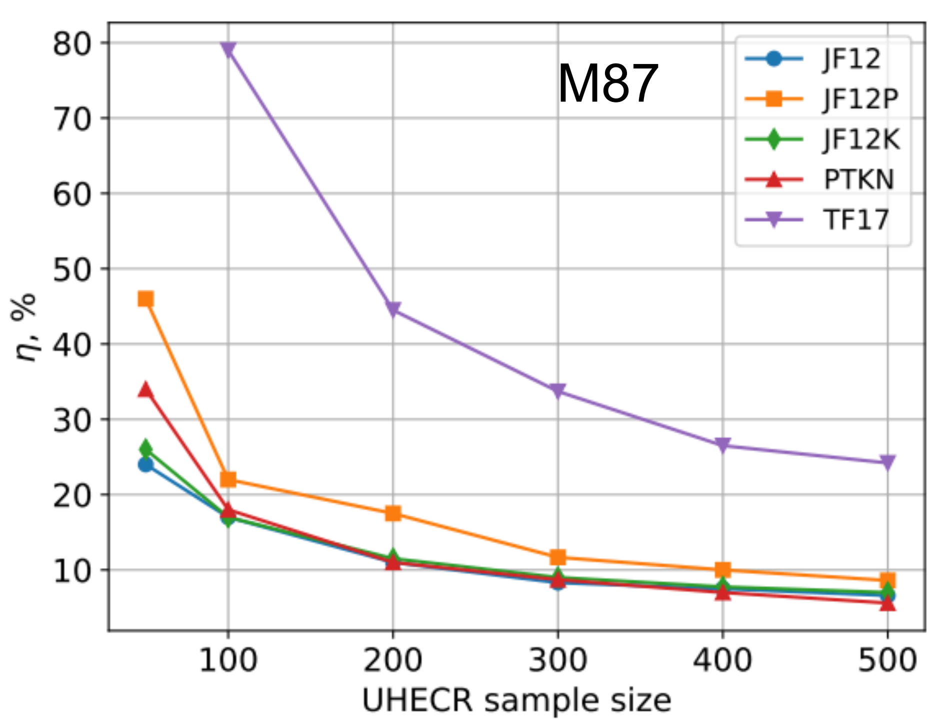
Model dependence

galactic magnetic field

Method:

Train model using JF'12 samples.

Stop training when accuracy on Pshirkov' 11 test set is not improving anymore

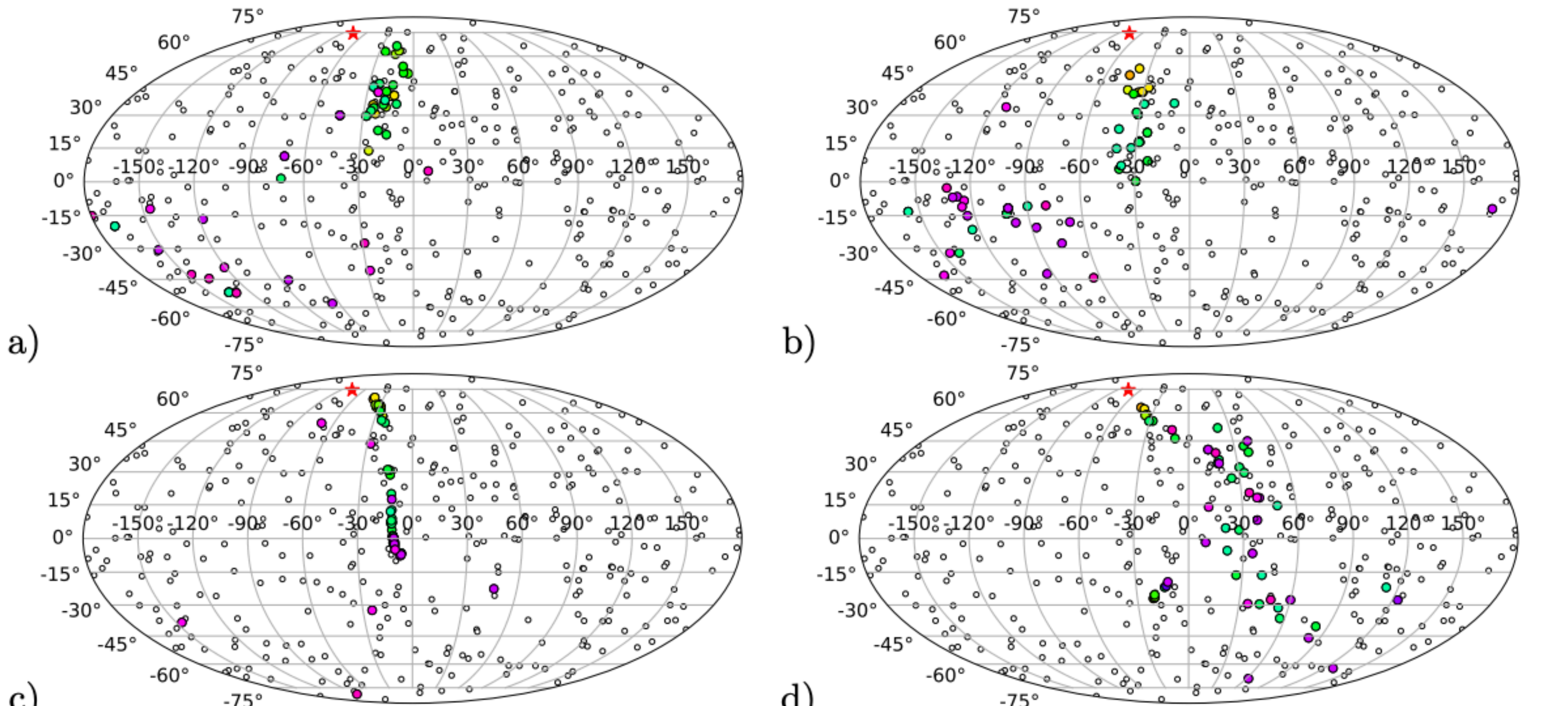


Source	GMF	50	100	200	300	400	500
NGC 253	JF12	14	10	6	5	3.75	3.6
	JF12P	20	13	8	6.7	5.5	5.2
	JF12K	14	9	5.5	4.3	3.75	3.4
	PTKN	28	16	9.5	8	6.5	5.8
	TF17	64	42	24	18.3	14.25	12
Cen A	JF12	16	12	7.5	6	5.25	4.8
	JF12P	22	14	10	8	6.75	5.8
	JF12K	16	11	8	6.3	5.5	4.8
	PTKN	20	12	8	6	5.25	4
	TF17	40	25	19	13.7	12	10.4
M 82	JF12	20	14	8.5	6.3	5.5	4.6
	JF12P	26	17	10	8	6.5	6.2
	JF12K	24	15	9	7.3	6	5
	PTKN	20	12	7	5.3	4.25	3.8
	TF17	30	20	12.5	9.7	7.75	6.2
M 87	JF12	24	17	11	8.3	7.5	6.6
	JF12P	46	22	17.5	11.67	10	8.6
	JF12K	26	17	11.5	9	7.75	7
	PTKN	34	18	11	8.7	7	5.6
	TF17	–	79	44.5	33.7	26.5	24.2
Fornax A	JF12	16	11	7.5	6	4.75	4.6
	JF12P	26	17	10.5	8.3	7	6.6
	JF12K	16	10	6.5	5	4.5	3.8
	PTKN	26	16	10	7.7	5.5	5.4
	TF17	58	38	18	15	10.75	9.8

Model dependence

galactic magnetic field

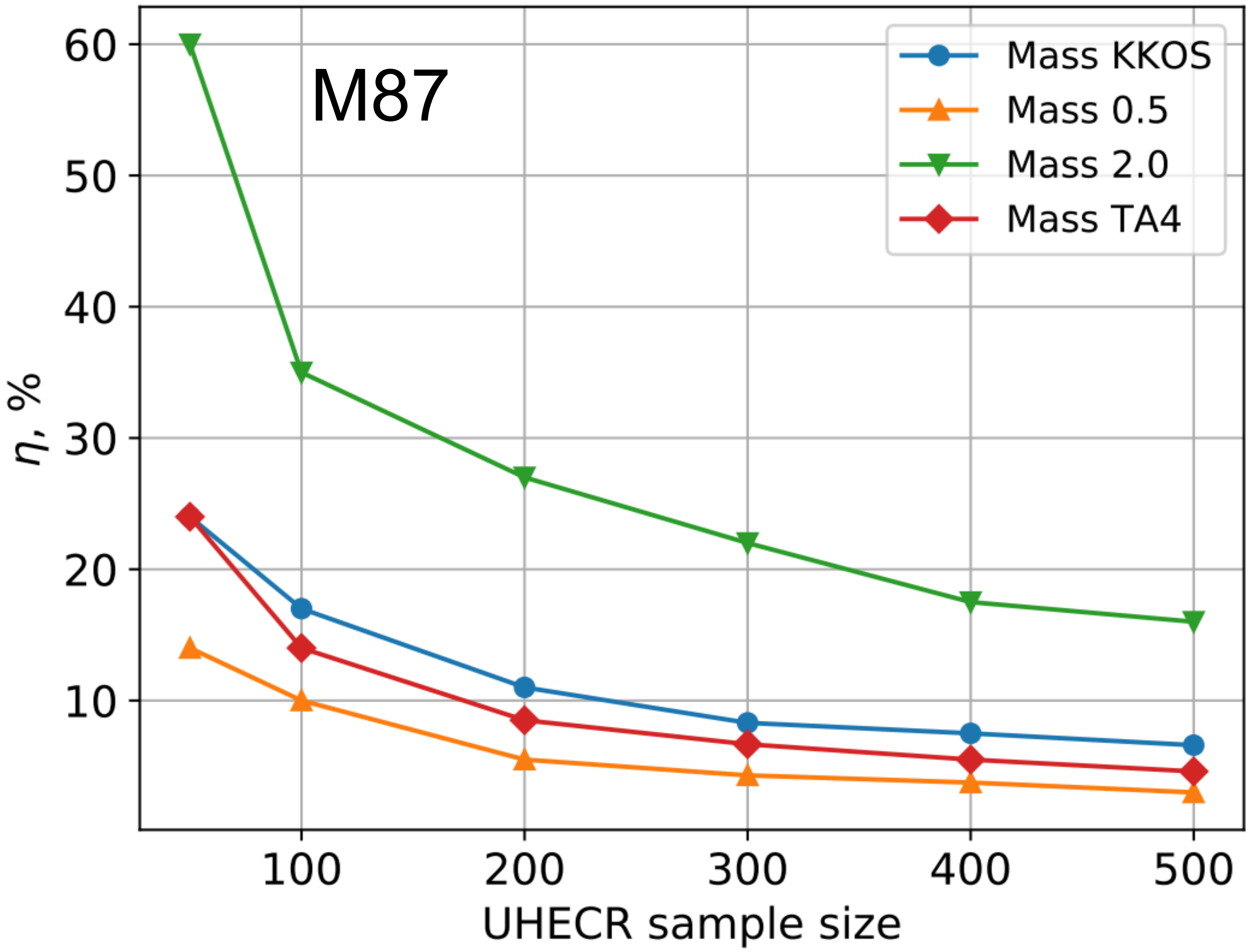
400 UHECRs with 14% coming from M87:



a) JF12P, b) JF12K, c) PTKN'17, d) TF17

Model dependence

mass composition



Applying original model to shifted mass composition

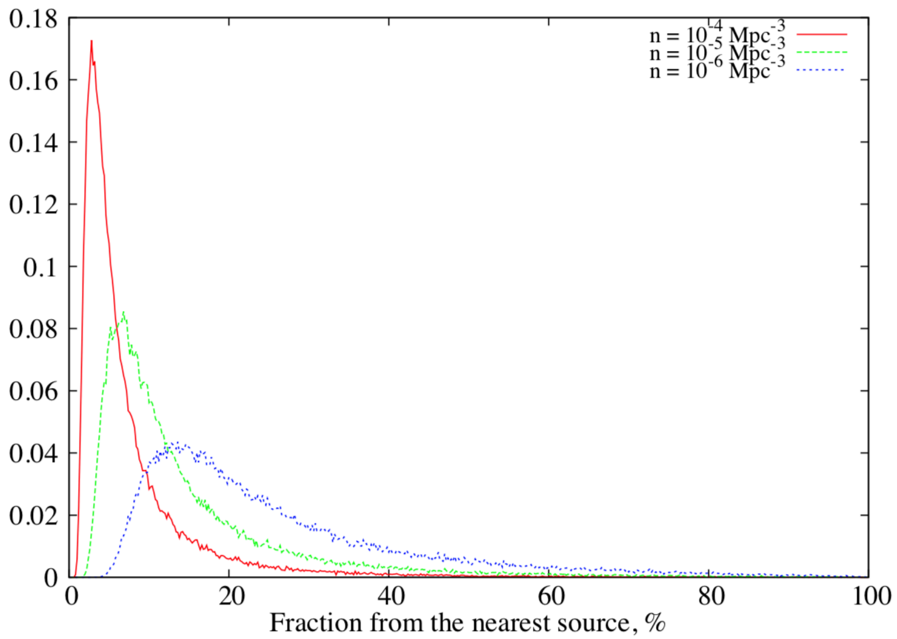
0.5: 2 times more light

2: 2 times heavier

TA4: TA 10 year composition ICRC’19 p:He:N:Fe = 57 : 18 : 17 : 8

Model dependence

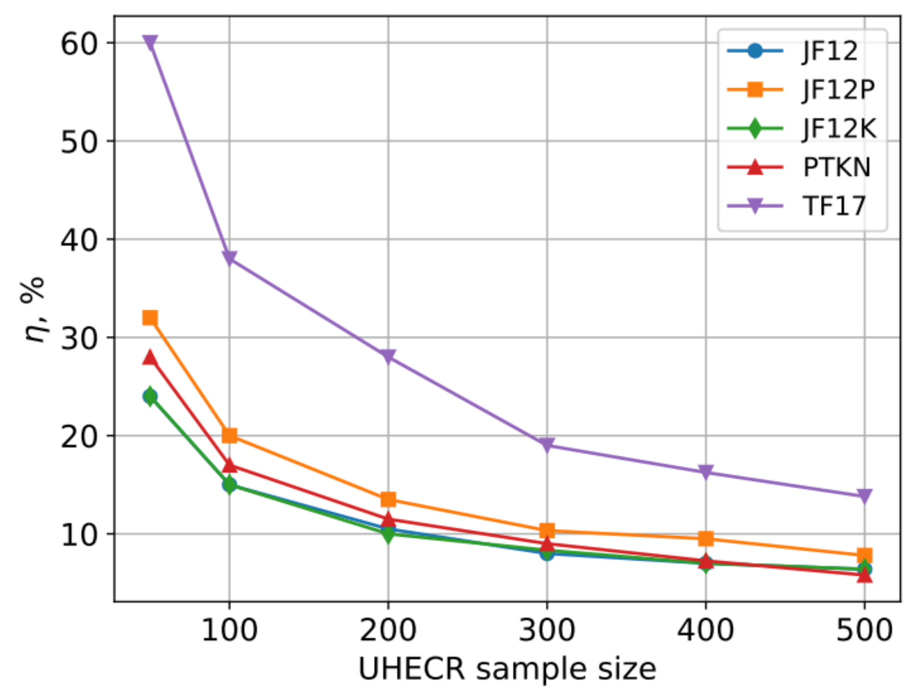
admixture from a 2-nd source



n, Mpc^{-3}	Closest	Second closest
10^{-4}	5.2	1.8
3×10^{-5}	7.5	2.7
10^{-5}	10.6	3.9
3×10^{-6}	15.0	5.0
10^{-6}	20.9	6.3

Average from-source flux fraction in % assuming identical sources

CenA	JF12	16	12	7.5	6	5.25	4.8
	JF12P	22	14	10	8	6.75	5.8
	JF12K	16	11	8	6.3	5.5	4.8
	PTKN	20	12	8	6	5.25	4
	TF17	40	25	19	13.7	12	10.4
CenA+M87 CenA:M87 = 3:1	JF12	24	15	10.5	8	7	6.4
	JF12P	32	20	13.5	10.33	9.5	7.8
	JF12K	24	15	10	8.33	7	6.4
	PTKN	28	17	11.5	9	7.25	5.8
	TF17	60	38	28	19	16.25	13.8



Application of the test statistics based on classifiers trained on Cen A to samples containing an admixture of events from another source.

Towards Universal Test Stats

Single classifier for multiple sources

Source	GMF	50	100	200	300	400	500
Cen A	JF12	22	14	9.5	7.3	6.25	5.6
	JF12P	24	16	11	8.3	7	5.6
	JF12K	24	15	9.5	7.67	6.5	5.8
	PTKN	22	14	9	7	5.25	4.4
	TF17	42	27	18.5	15	11	9.4
M 87	JF12	28	19	12	11	9.5	7.8
	JF12P	58	36	26.5	22	19.5	11.8
	JF12K	30	20	13	11.3	9	8
	PTKN	32	20	12	9.67	8.75	6
	TF17	–	95	60.5	39.7	31.23	26.4
Fornax A	JF12	18	13	8	6.3	6	5.2
	JF12P	32	19	13	10.3	9	8
	JF12K	18	12	7.5	6	5	5
	PTKN	36	21	14	10	8.5	8.4
	TF17	70	41	24	17	15.5	15.2

universal model

Source	GMF	50	100	200	300	400	500
Cen A	JF12	16	12	7.5	6	5.25	4.8
	JF12P	22	14	10	8	6.75	5.8
	JF12K	16	11	8	6.3	5.5	4.8
	PTKN	20	12	8	6	5.25	4
	TF17	40	25	19	13.7	12	10.4
M 87	JF12	24	17	11	8.3	7.5	6.6
	JF12P	46	22	17.5	11.67	10	8.6
	JF12K	26	17	11.5	9	7.75	7
	PTKN	34	18	11	8.7	7	5.6
	TF17	–	79	44.5	33.7	26.5	24.2
Fornax A	JF12	16	11	7.5	6	4.75	4.6
	JF12P	26	17	10.5	8.3	7	6.6
	JF12K	16	10	6.5	5	4.5	3.8
	PTKN	26	16	10	7.7	5.5	5.4
	TF17	58	38	18	15	10.75	9.8

source specific model

Results of tests of the CNNs trained simultaneously for three sources (Cen A, M 87 and Fornax A) assuming only one of them is contributing to a large-scale anisotropy of the flux.

“Source independent” test stats

universal model

Source	GMF	50	100	200	300	400	500
NGC 253	JF12	36	22	15.5	11.67	10.25	8.2
	JF12P	76	48	31.5	25	22	15.8
	JF12K	32	20	12.5	9.7	8.25	6.4
	PTKN	–	88	63.5	48	41.5	26.8
	TF17	–	75	61.5	42.3	34	29.8
M 82	JF12	50	32	20	15	13.25	10.4
	JF12P	50	32	18.5	15.3	12.75	10.2
	JF12K	50	32	20	15.3	12.25	10.8
	PTKN	32	22	13	10.3	9	7.4
	TF17	48	28	18	13.7	12	10

source specific model

Source	GMF	50	100	200	300	400	500
NGC 253	JF12	14	10	6	5	3.75	3.6
	JF12P	20	13	8	6.7	5.5	5.2
	JF12K	14	9	5.5	4.3	3.75	3.4
	PTKN	28	16	9.5	8	6.5	5.8
	TF17	64	42	24	18.3	14.25	12
M 82	JF12	20	14	8.5	6.3	5.5	4.6
	JF12P	26	17	10	8	6.5	6.2
	JF12K	24	15	9	7.3	6	5
	PTKN	20	12	7	5.3	4.25	3.8
	TF17	30	20	12.5	9.7	7.75	6.2

Results of tests of the CNNs trained simultaneously for three sources (Cen A, M 87 and Fornax A) assuming only one of them is contributing to a large-scale anisotropy of the flux.

NGC 253 and M 82 were

not involved in training.

Angular power spectrum based stat.

N_{UHECR}	100	200	300	400	500
NGC 253	17	12	10	8	7
Cen A	21	14	12	10	9
M82	26	18	14	12	11
M87	29	20	16	14	12
Fornax A	19	13	11	9	8

Towards real experiment

Problems:

- nonuniform exposure
- finite angular resolution
- finite energy resolution

Way to solve:

- weighted map generation
- map smearing
- binning and smearing in energy

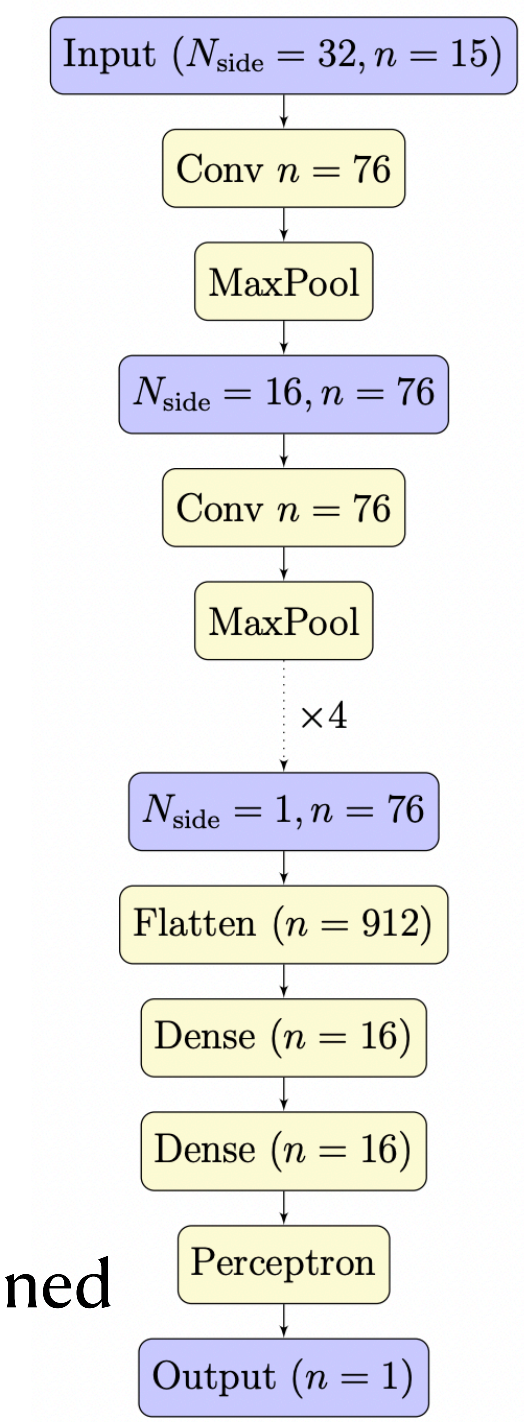
Binned map generation: each (E,Z) pair contributes to several neighbour energy bins with weights w_i given by:

$$w_i = \int_{\lg E_i}^{\lg E_i + \Delta_b} N(E, \Sigma_{\lg E}) \qquad \text{where} \qquad \Sigma_{\lg E} \simeq \frac{1}{\ln 10} \frac{\Delta E}{E} \quad \text{is energy resolution}$$

Towards real experiment

Test statistics performance on binned maps assuming $\Delta E/E = 0.2$

Source	Method	50	100	200	300	400	500
NGC 253	BCNN	12	8	5	4	3.25	3.2
	CNN	12	7	4.5	3.67	3	2.6
Cen A	BCNN	16	10	7	5.67	4.75	4.2
	CNN	16	11	7	5.67	5	4.4
M 82	BCNN	20	14	8.5	6.67	5	4.4
	CNN	20	12	7	6	4.75	4.2
M 87	BCNN	22	15	11.5	7	6.25	5.4
	CNN	22	14	9	8	6.25	5.2
Fornax A	BCNN	16	10	6	5	4.5	3.8
	CNN	16	9	6	5	4.5	3.8



Model optimised for binned map with $\Delta_b = 1/20$

Conclusions

- **ANNs provide efficient universal way to define test statistic as function of direct observables. In case of arrival directions interpretation, convolutional NN on HEALPix grid can be used**
- **pros:**
  - **optimal (trained to discriminate two hypothesis)**
  - **easy to apply to nonuniform exposure data (e.g. TA and Auger)**
- **cons:**
  - **not easy to interpret (see however perceptron sample)**

Backup Slides

Observable spectra calculation

Photo-hadronic and hadronic interactions

Injection power-low $\alpha = 1.5$

Maximal energy $E_{max} = 6$ EeV

Evolution: AGN with $LogL_X = 43.5$

Photon temperature $T = 850$ K

Interaction depth $\tau_0^{pp} = 0.035$

Interaction depth $\tau_0^{p\gamma} = 0.29$

Diffusion:

$\delta_{Ap} = 0.5$

$\delta_{A\gamma} = 0.77$