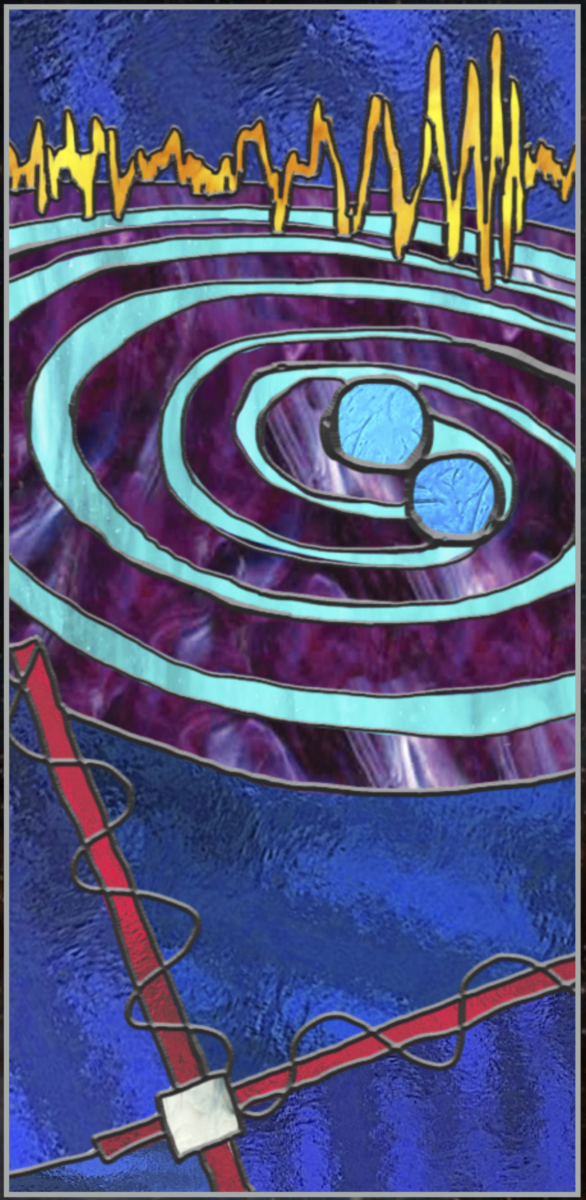
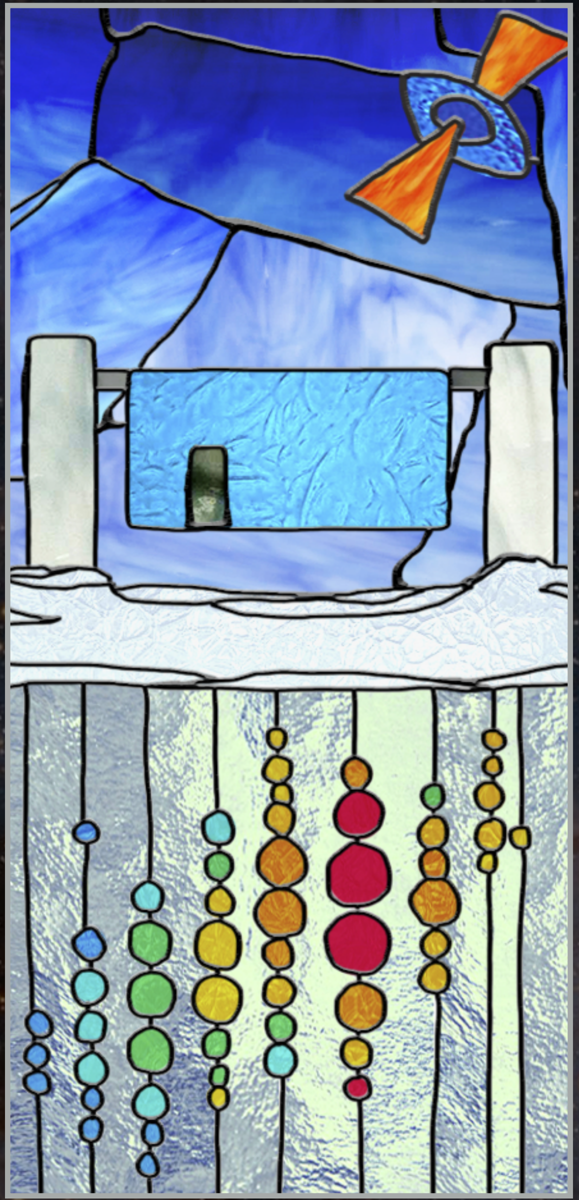
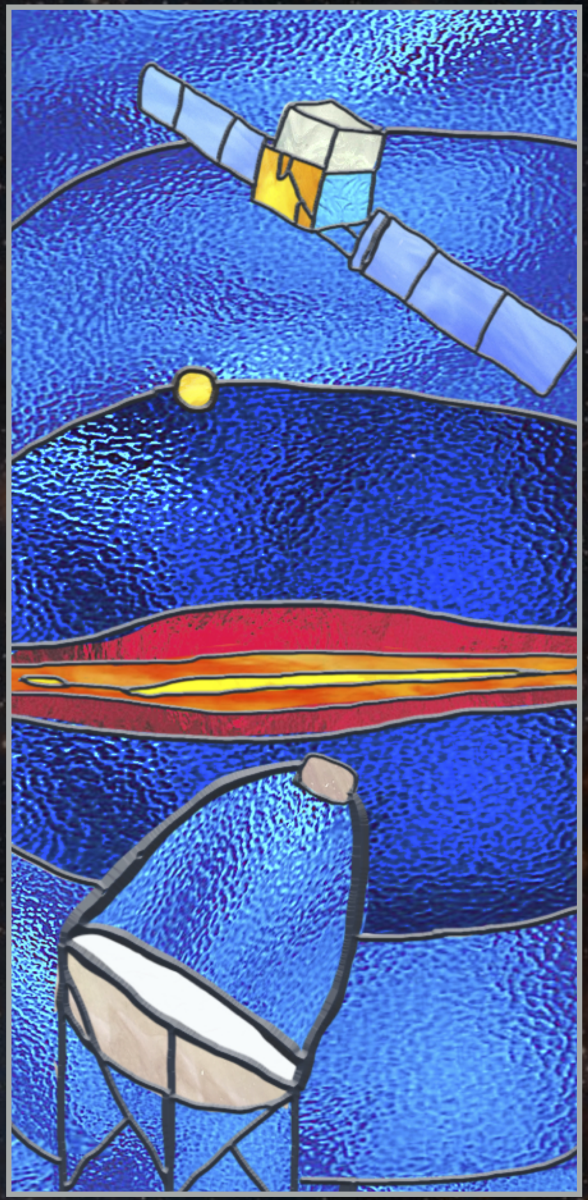






Multi-Messenger Data Analyses: Selected Topics

Markus Ahlers

Niels Bohr Institute

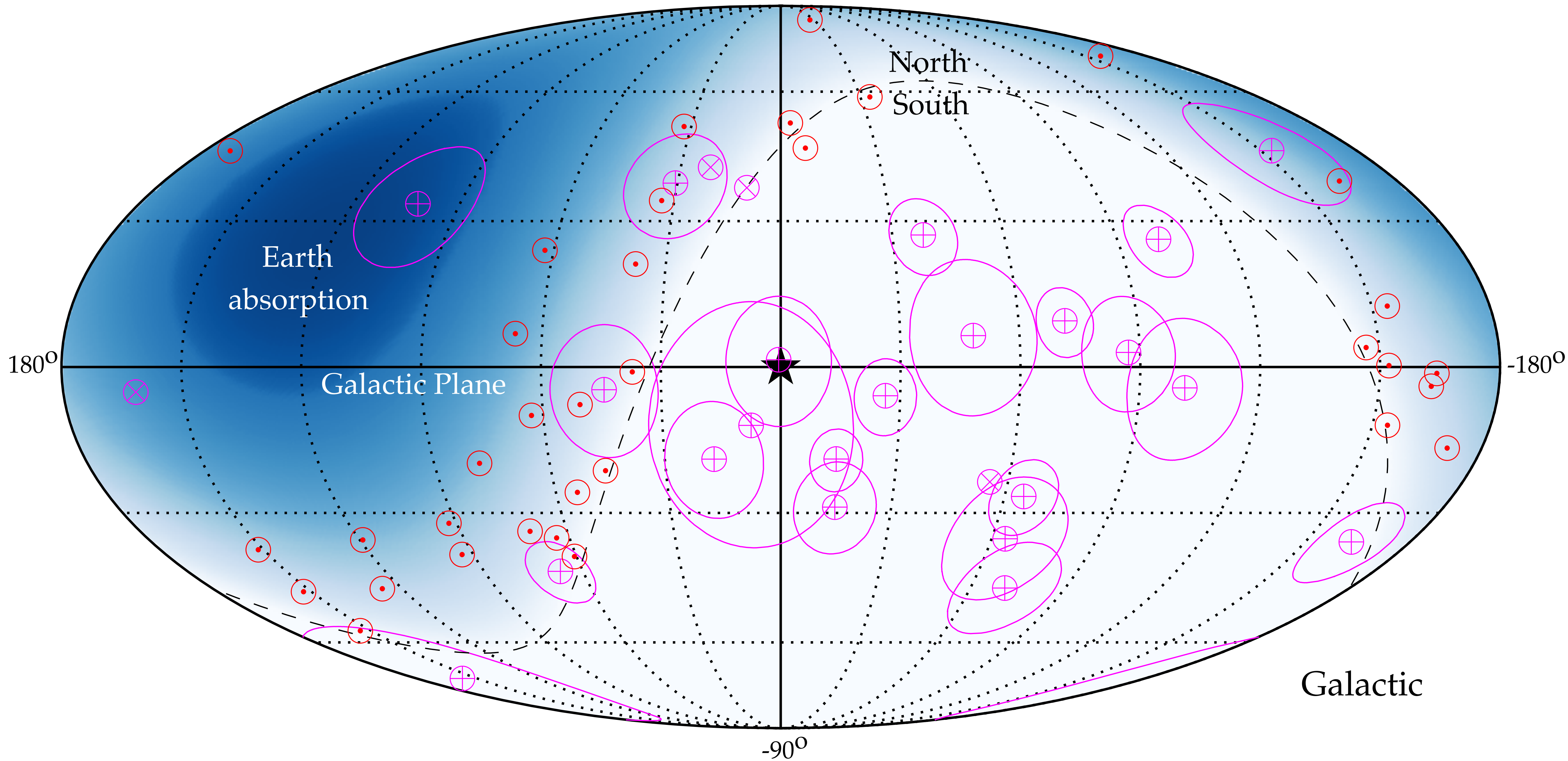
Georges Lemaître Chair 2023

VILLUM FONDEN

KØBENHAVNS UNIVERSITET

Status of Neutrino Astronomy

Most energetic neutrino events (HESE 6yr (magenta) & $\nu_\mu + \overline{\nu}_\mu$ 8yr (red))



No significant steady or transient emission from known Galactic or extragalactic high-energy sources, but **several interesting candidates.**

Statistical Hypothesis Tests

Typical problem in physics and astronomy:

You have collected data with your experiment or observatory and want to test a theory (**signal hypothesis** H_1).

- *How can you judge if the hypothesis is correct or wrong?*
- *How does the alternative hypothesis* (**null hypothesis** H_0) *look like?*
- *How confident can you be that your conclusions are correct?*
- In most cases there is a chance that your decision is wrong:
  - You decided that H_1 is **correct**, but it is actually **wrong**. (type I error)
  - You decided that H_1 is **wrong,** but it is actually **correct.** (type II error)

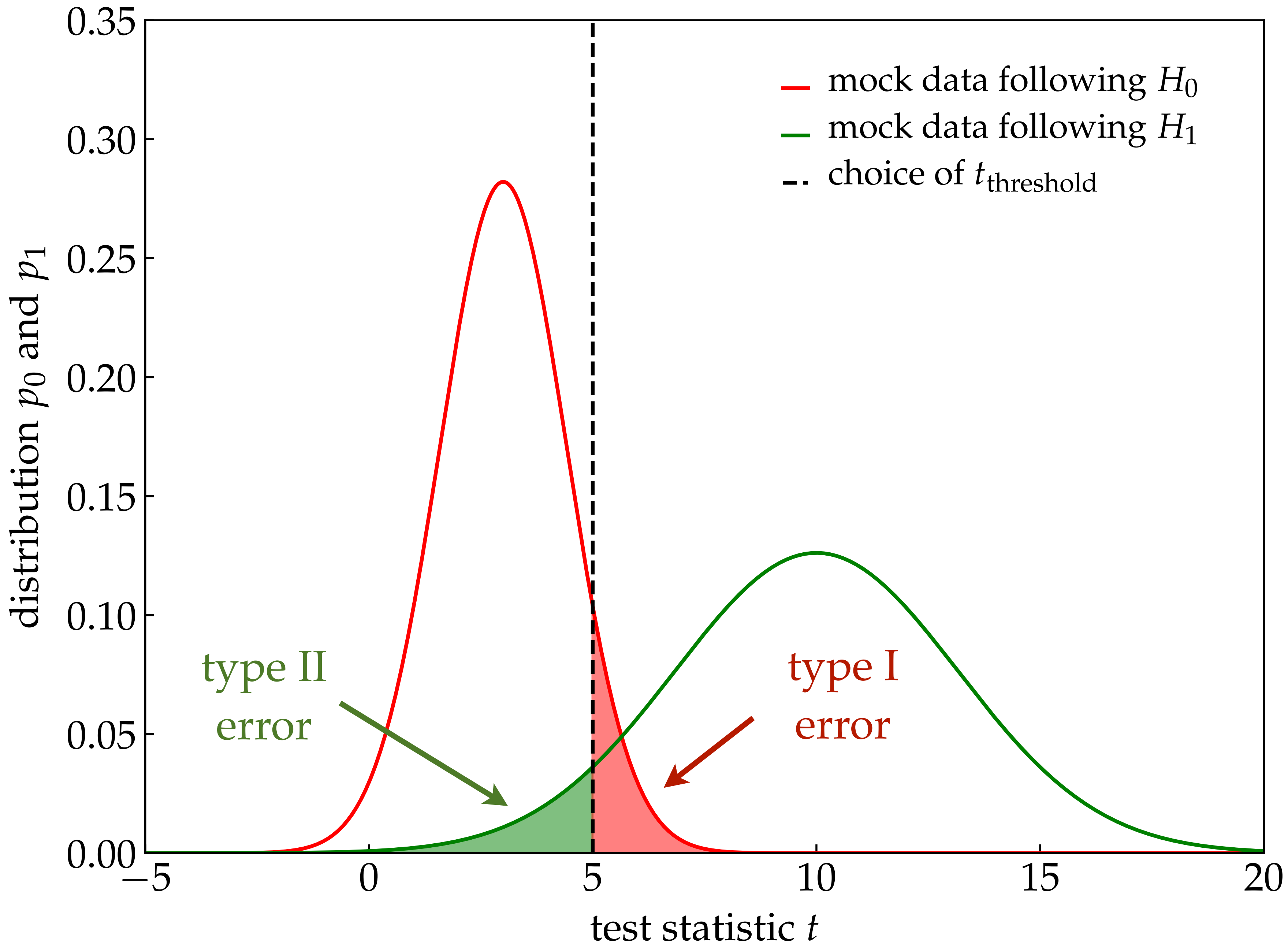
Statistical Hypothesis Tests

- A **statistical hypothesis test** is based on a quantity called **test statistic** that allows us to quantify the degree of confidence that your decision was right or wrong.

- A useful test statistic:
  - is **sensitive** to the signal hypothesis H_1 (that's a must!)
  - is **efficiently calculable** (*e.g.* fast calculation on your computer)
  - has a **well-known behaviour** for data following the null hypothesis H_0

- If we apply the statistical test to the observed data we can quantify the *"false positive"* (type I) and *"false negative"* (type II) errors by comparing to the expected test statistic distribution, p_0 and p_1, of data following background (H_0) and signal (H_1) hypothesis, respectively.

Statistical Hypothesis Tests



In a hypothesis test we have to choose a **threshold test statistic value** to either reject or accept the hypothesis.

Statistical Hypothesis Tests

- **significance** (α):

probability that background creates outcome with $t_{\rm thr}$ *or larger:*

$$\alpha = \int_{t_{\rm thr}}^{\infty} {\rm d}t\, p_0(t)$$ (type I error)

  - **Note:** It is a convention that t increases for a more "signal-like" outcome. If not, just define a new test statistic $t' \equiv -t$.

- **power of test** $(1-\beta)$:

probability that signal creates outcome with $t_{\rm thr}$ *or less:*

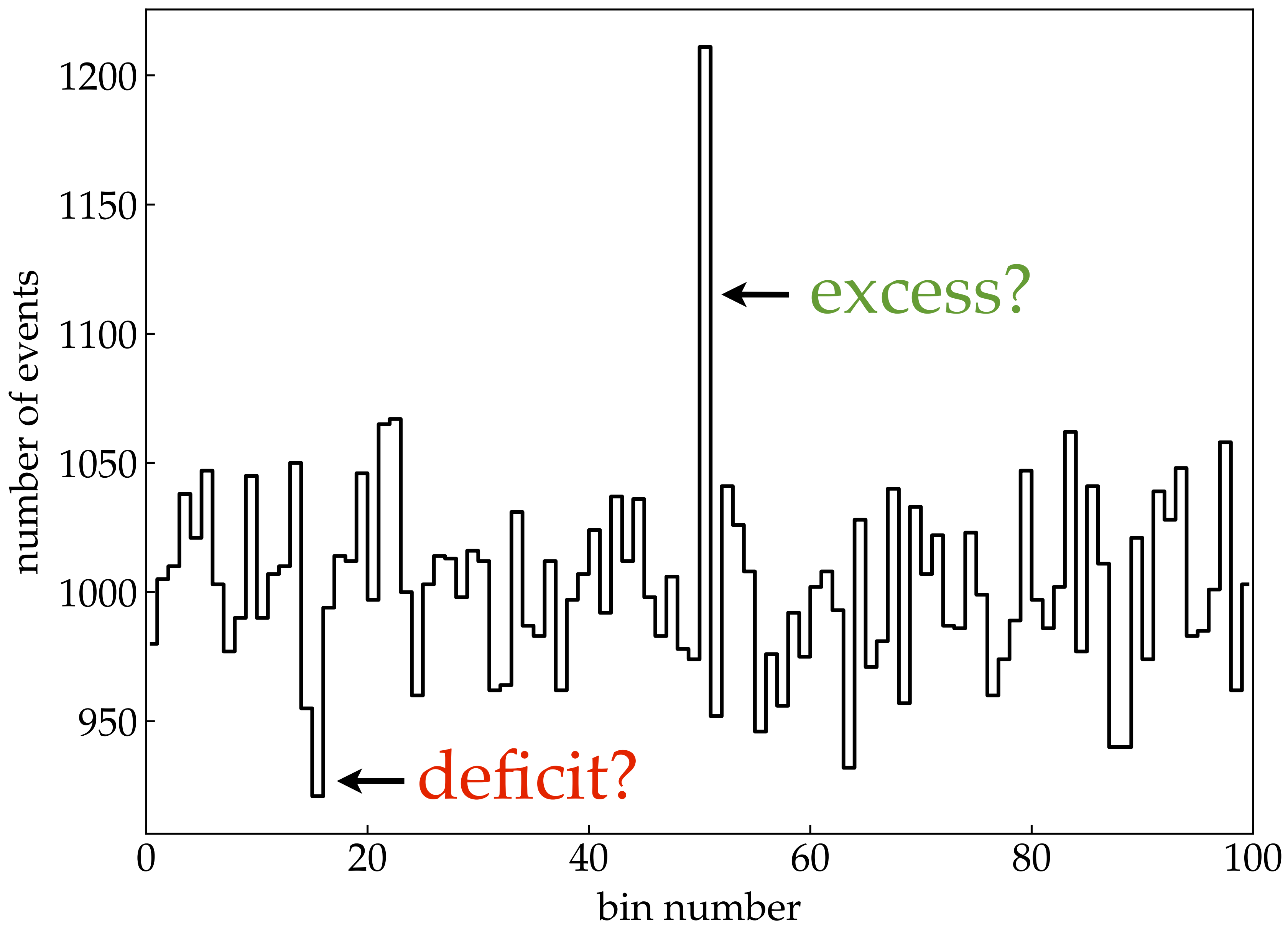
$$\beta = \int_{-\infty}^{t_{\rm thr}} {\rm d}t\, p_1(t)$$ (type II error)

Statistical Hypothesis Tests

- A good statistical test will have good "separation" of p_0 and p_1 to allow to minimize type I/II errors. Separation from background allows to quantify significance of event excesses:
  - **discovery** (in particle physics): $\alpha = 5.7 \times 10^{-7} \quad (5\sigma)$
  - **evidence** (in particle physics): $\alpha = 2.7 \times 10^{-3} \quad (3\sigma)$
- Often, we want to estimate the performance of a statistical test **prior to a measurement** by simulations. We can determine this by tuning the signal strength, *e.g.* the **IceCube experiment** uses:
  - **discovery potential**: $\alpha = 5.7 \times 10^{-7} \qquad \beta = 0.5$
  - 90 % **sensitivity level**: $\alpha = 0.5 \qquad \beta = 0.1$

Example: Excess in Binned Data

- Consider data ($N_{\rm tot}$ "events") distributed in $N_{\rm bins}$ bins.
- **Question:** Is there an excess or deficit in the data?



Likelihood Function

- **Likelihood function** $\mathscr{L}(\boldsymbol{\theta}\,|\,\mathbf{x})$ for data vector $\mathbf{x}$ and parameter vector $\boldsymbol{\theta}$.
- Assuming(!) Poisson statistics in our bins:

$$\mathscr{L}(\boldsymbol{\mu}\,|\,\mathbf{x}) = \prod_{i=1}^{N_{\rm bins}} \frac{1}{x_i!}\mu_i^{x_i}e^{-\mu_i}$$

- **Background hypothesis** ("no signal"):

$$\mu_i = \mu_{\rm bg} = {\rm const}$$

- **Signal hypothesis** ("signal (excess or deficit) in bin 1"):

$$\mu_i = \begin{cases} \mu_{\rm sig} + \mu^*_{\rm bg} & i = 1 \\ \mu^*_{\rm bg} & {\rm else} \end{cases}$$

- Note that, in general: $\mu_{\rm bg} \neq \mu^*_{\rm bg}$

Maximum-Likelihood

- For (mathematical) convenience $\mathscr{L} \rightarrow \ln \mathscr{L}$ ("log-likelihood"):

$$\ln \mathscr{L}(\boldsymbol{\mu} \,|\, \mathbf{x}) = \sum_{i=1}^{N_{\rm bins}} (x_i \ln \mu_i - \mu_i) + {\rm const}$$

- In general, maximum of $\mathscr{L}$ (or $\ln \mathscr{L}$) can only be determined numerically.
- *This example is easy enough to solve analytically.*
- maximum $\hat{\mu}_{\rm bg}$ of **background hypothesis** is at:

$$\hat{\mu}_{\rm bg} = \frac{N_{\rm tot}}{N_{\rm bins}}$$

- maximum $(\hat{\mu}^*_{\rm bg}, \hat{\mu}_{\rm sig})$ of **signal hypothesis** is at:

$$\hat{\mu}^*_{\rm bg} = \frac{N_{\rm tot} - x_1}{N_{\rm bins} - 1} \qquad \hat{\mu}_{\rm sig} = x_1 - \hat{\mu}^*_{\rm bg}$$

Maximum-Log-Likelihood Ratio

- We now define the test statistic λ as the **maximum likelihood ratio**:

$$\lambda(x) \equiv -2\ln\frac{\mathscr{L}(\hat{\mu}_{\rm bg},0\,|\,\mathbf{x})}{\mathscr{L}(\hat{\mu}^*_{\rm bg},\hat{\mu}_{\rm sig}\,|\,\mathbf{x})}$$

- After some algebra using the solutions for $\hat{\mu}^*_{\rm bg}$, $\hat{\mu}_{\rm bg}$ and $\hat{\mu}_{\rm sig}$:

$$\lambda(x) = 2x_1\ln\left(\frac{N_{\rm bins}}{N_{\rm tot}}x_1\right) + 2(N_{\rm tot}-x_1)\ln\left(\frac{N_{\rm bins}}{N_{\rm tot}}\frac{N_{\rm tot}-x_1}{N_{\rm bins}-1}\right)$$

- For $\hat{\mu}^*_{\rm bg} \simeq \hat{\mu}_{\rm bg}$ we can simplify this to:

$$\lambda(x) \simeq 2x_1\left(\ln x_1 - \ln\hat{\mu}_{\rm bg}\right) - 2\left(x_1 - \hat{\mu}_{\rm bg}\right) \qquad \hat{\mu}_{\rm bg} = \frac{N_{\rm tot}}{N_{\rm bins}}$$

Maximum-Log-Likelihood Ratio

- Example python notebooks can be found here:

`https://github.com/mahlers77/KSETA2023`

- You can use *Google Colaboratory* to execute the python notebook:

`https://colab.research.google.com`

- If you want to explore the exact solution of the maximum likelihood TS for binned data, have a look at this example:

`max_LH_example.ipynb`

- In the interest of time we will go straight to real IceCube data:

`IceCube_allsky.ipynb`

Example: IceCube 10-yr Data

- We can devise a simplified version of this analysis by a binned maximum likelihood test with an energy threshold.
- Analysis of IceCube PS data from '08-'18 : `IceCube_allsky.ipynb`
- We can estimate the background expectation $\mu_{\rm bg,i}$ via **RA scrambling**:

$$\mathrm{RA} \to \mathrm{RA} + \theta_{\rm rnd}$$

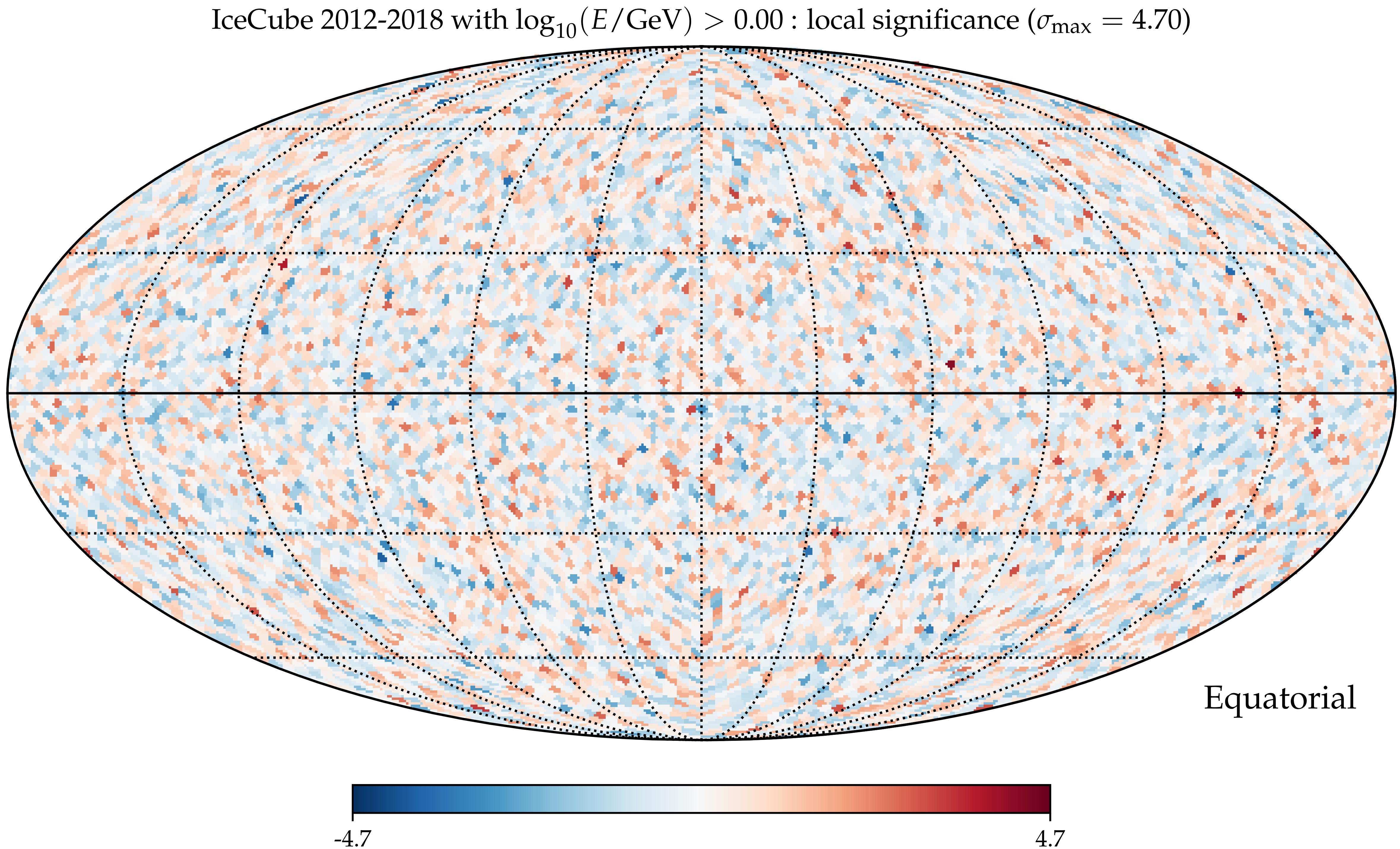
- The maximum log-likelihood ratio for a **point-source in each pixel** i can be approximated as:

$$\lambda_i = -2\ln\frac{\mathscr{L}_{0,i}}{\mathscr{L}_i} \simeq 2x_i\left(\ln x_i - \ln \mu_{\rm bg,i}\right) - 2\left(x_i - \mu_{\rm bg,i}\right)$$

- Using **Wilks' theorem** the p-value and significance S (units of σ) is:

$$p = 1 - \mathrm{erf}(\sqrt{\lambda_i/2}) \qquad S = \sqrt{\lambda_i}$$

Example: IceCube 10-yr Data



IceCube_allsky.ipynb

Unbinned Max-Likelihood

- The previous maximum-likelihood test required that we binned events.
- We can incorporate the uncertainty of the event reconstruction in an **unbinned maximum likelihood** test for n_s signal events:

$$\mathcal{L}(n_s, \boldsymbol{\theta}\,|\,\mathbf{x}) = \prod_{i=1}^{N_{\rm tot}} \left[\frac{n_s}{N_{\rm tot}} S_i(\boldsymbol{\theta}) + \left(1 - \frac{n_s}{N_{\rm tot}}\right) B_i\right]$$

- For instance, in a neutrino analysis with uncertainties of event location $\boldsymbol{\Omega}$ and energy E described by probability distributions w_i, the **signal** and **background** weights are:

$$S_i(\boldsymbol{\theta}) = \int \mathrm{d}\Omega \int \mathrm{d}E\, w_i(\Omega, E) T_{\rm sig}(\Omega, E, \boldsymbol{\theta})$$

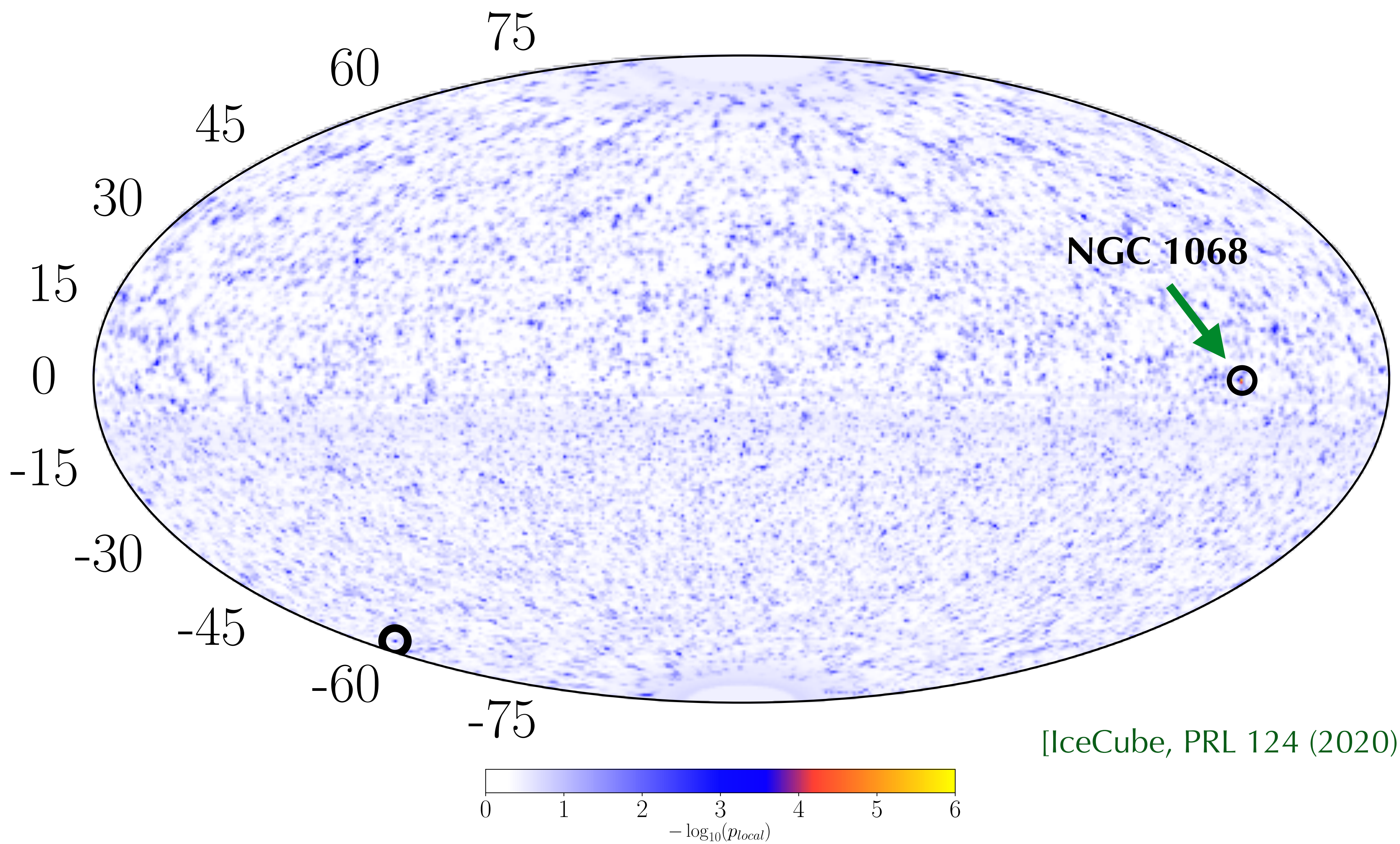
$$B_i = \int \mathrm{d}\Omega \int \mathrm{d}E\, w_i(\Omega, E) T_{\rm bg}(\Omega, E)$$

Unbinned Max-Likelihood

- Normalized signal template $T_{\rm sig}(\Omega, E, \boldsymbol{\theta})$:
  - point-source analysis: $\boldsymbol{\theta} = \{\gamma_{\rm src}, \mathbf{n}_{\rm src}\}$
  - stacking of M (identical) PSs: $\boldsymbol{\theta} = \{\gamma_{\rm src}, \mathbf{n}_1, \ldots, \mathbf{n}_M, \mathbf{w}_1, \ldots, \mathbf{w}_M\}$
  - extended sources: $\boldsymbol{\theta} = \{\gamma_{\rm src}, \mathbf{n}_{\rm src}, \sigma_{\rm src}\}$

- Normalized background template $T_{\rm bg}(\Omega, E)$:
  - background of **atmospheric muons and neutrinos** with spectrum $E^{-3.7}$ (conventional) and $E^{-2.7}$ (prompt) components; azimuthally symmetric
  - derived from **MC simulations** or from **background scrambling**, *i.e.* randomized arrival times (corresponding to right ascension scrambling at Southpole).

Example: IceCube 10-year



[IceCube, PRL 124 (2020) 5]

IceCube "all-sky" point-like source search:
each location tested for an excess!

Trial Correction

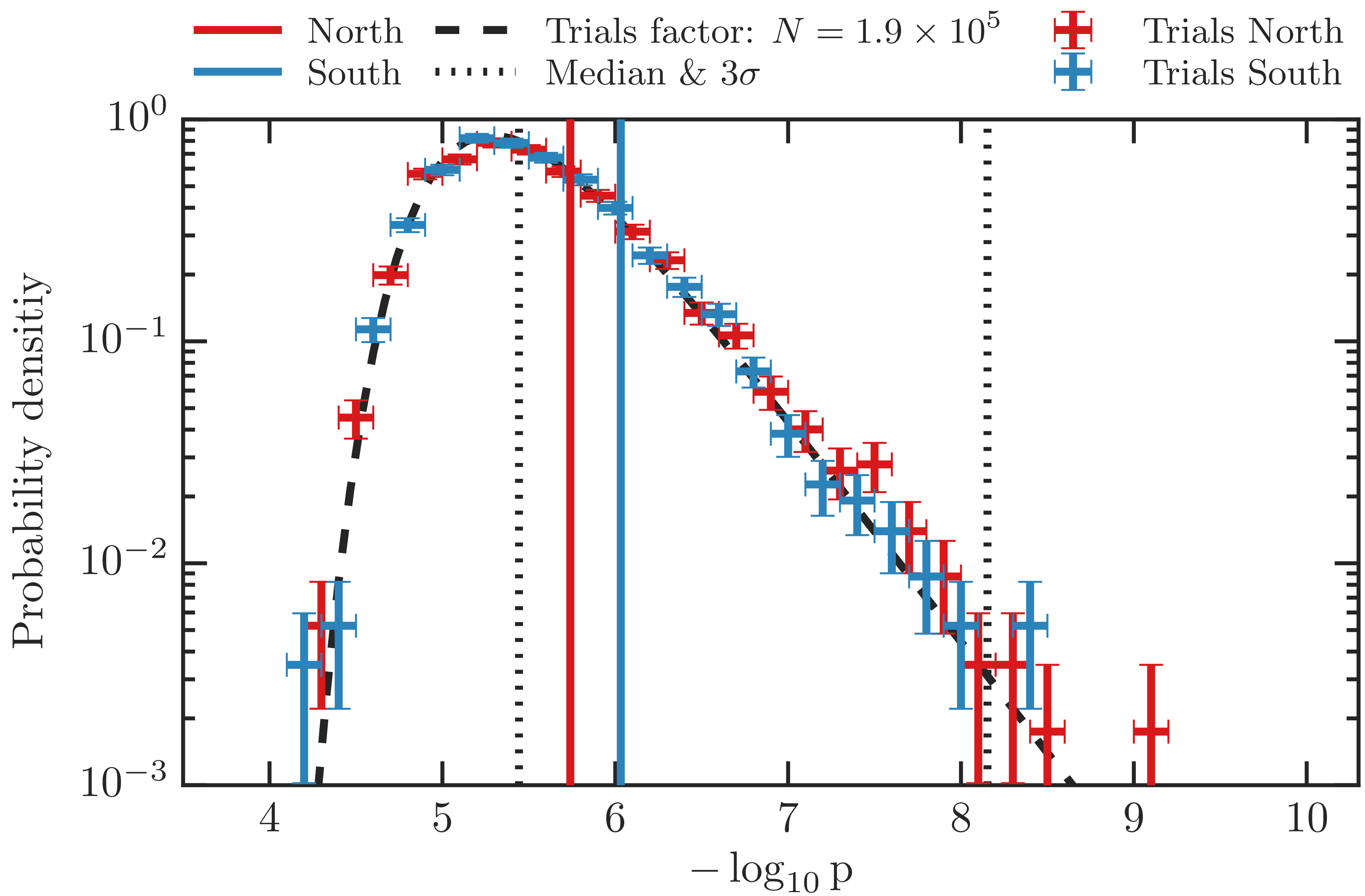
- What happens if we want to find a signal not just in the first bin but in any of the $N_{\rm bins}$ bins?
- We can simply repeat the test over all bins and identify the bin with minimal p-value $p_{\rm min}$.
- **Problem:** There are many bins ("hypothesis") and we have to account for the fact that there can be a chance fluctuation in the local p-values.
- If $N_{\rm bins}$ are independent of each other (as in our example) then we can define a post-trial p-value as:

$$p_{\rm post} = 1 - \underbrace{(1 - p_*)^{N_{\rm bins}}}_{\text{background probability}} \simeq N_{\rm bins} p_*$$

- Number of independent "trials" ($N_{\rm trials}$) is often difficult to estimate.

Example: IceCube 7-year

- Trial factor: $N_{\text{trials}} \sim N_{\text{bins}} \sim \mathcal{O}(10000)$
- IceCube procedure: choose maximal p_{local} in sky map as a **new test statistic** and compare against maximal p_{local} of randomly generated maps.



[IceCube ApJ 835 (2017) 2, 151]

Binomial Test

- Consider a sorted list of p-values $p_i \leq p_j$ $(i < j)$. Trial-corrected p-value is:

$$p_1' = 1 - (1 - p_1)^N$$

- We can ask, how likely it is that the background yields **at least two p-values** with $p \leq p_2$:

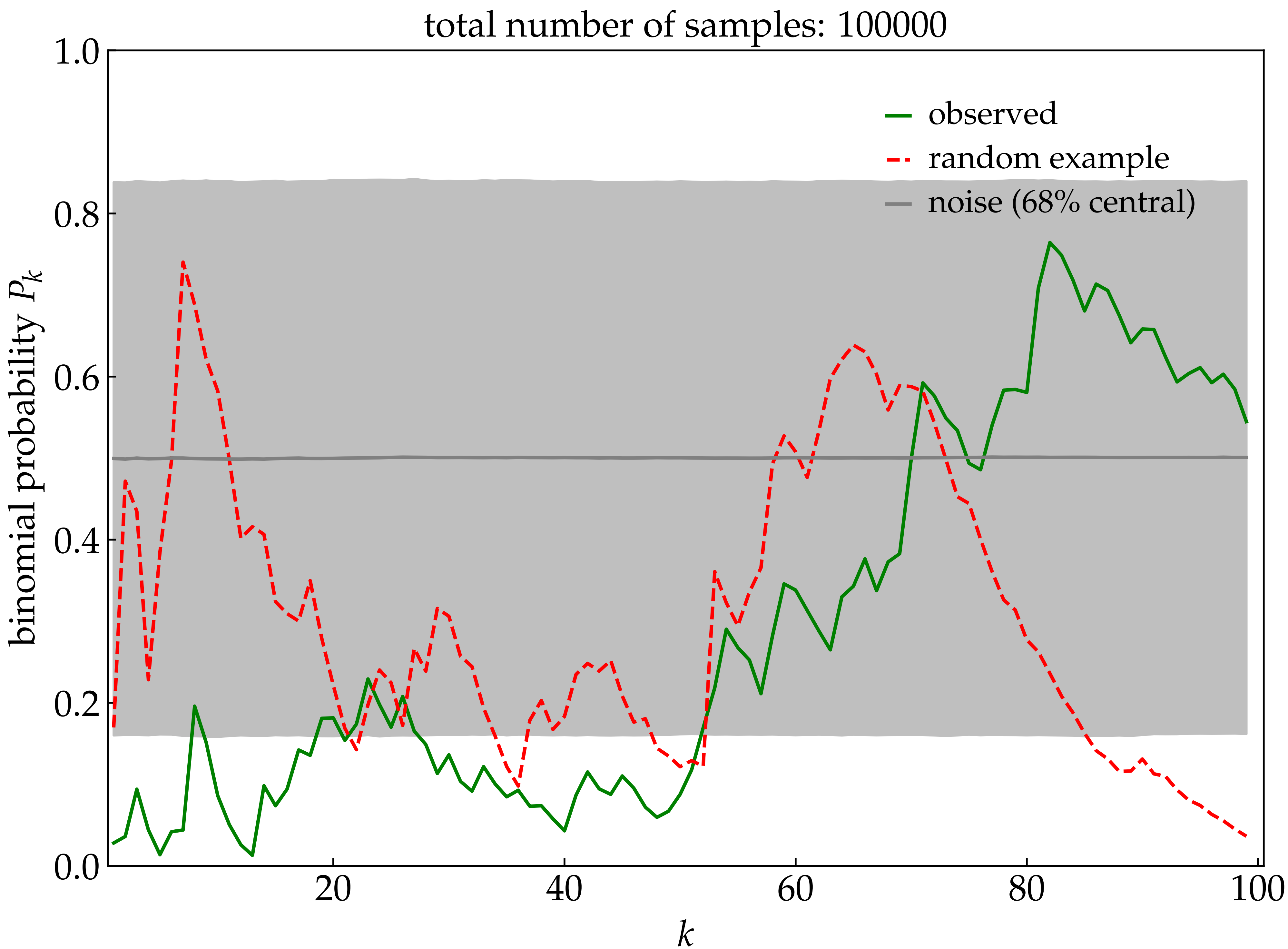
$$p_2' = 1 - (1 - p_2)^N - Np_2(1 - p_2)^{N-1}$$

- In general, the background probability for **at least k p-values** with $p \leq p_k$:

$$p_k' = 1 - \sum_{n=0}^{k-1} \binom{N}{n} p_k^n (1 - p_k)^{N-n} \simeq 1 - \frac{\Gamma(k, Np_k)}{\Gamma(k)}$$

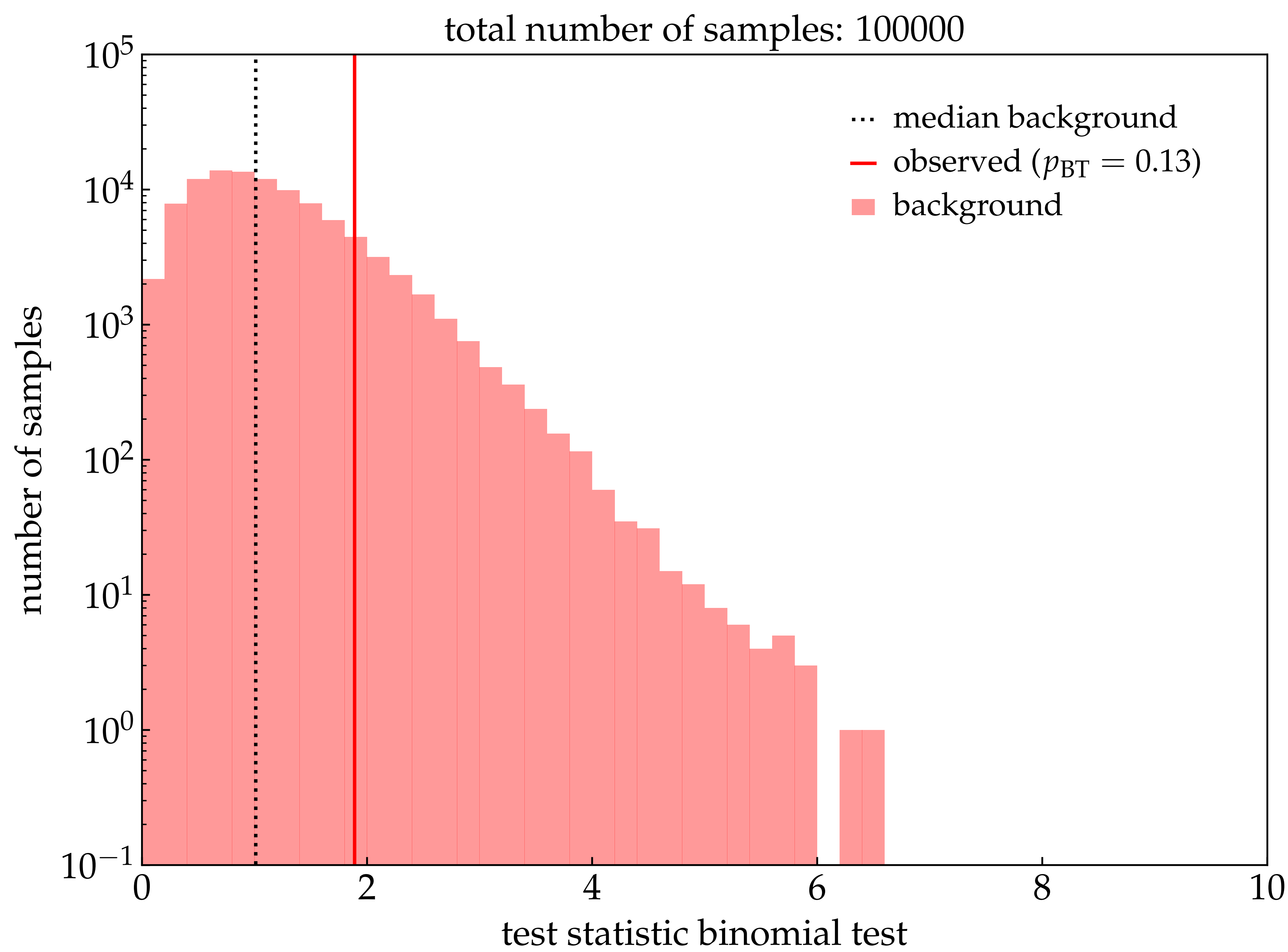
- We can define yet another **test statistic** as : $t \equiv \min\{p_k'\}$ ("binomial test").

Example: IceCube



IceCube_allsky.ipynb

Example: IceCube



`IceCube_allsky.ipynb`

Galactic Cosmic Rays

- *Standard paradigm*: Galactic CRs accelerated in supernova remnants

[Baade & Zwicky'34]
[Ginzburg & Sirovatskii'64]

- diffusive shock acceleration:

$$n_{\rm CR} \propto E^{-\Gamma}$$

- rigidity-dependent escape from Galaxy:

$$n_{\rm CR} \propto E^{-\Gamma-\delta}$$

- Arrival directions of cosmic rays are scrambled by magnetic fields.

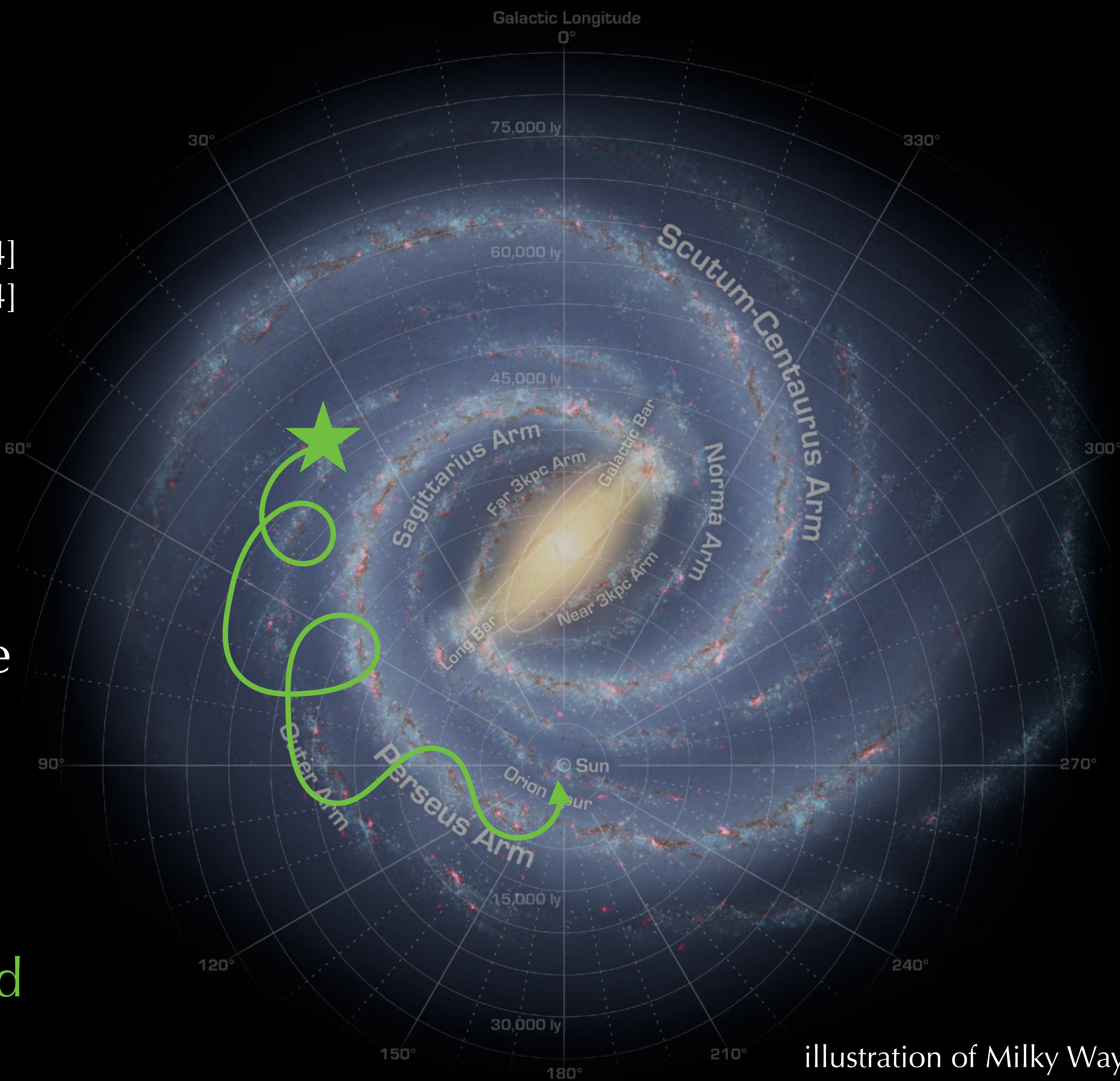
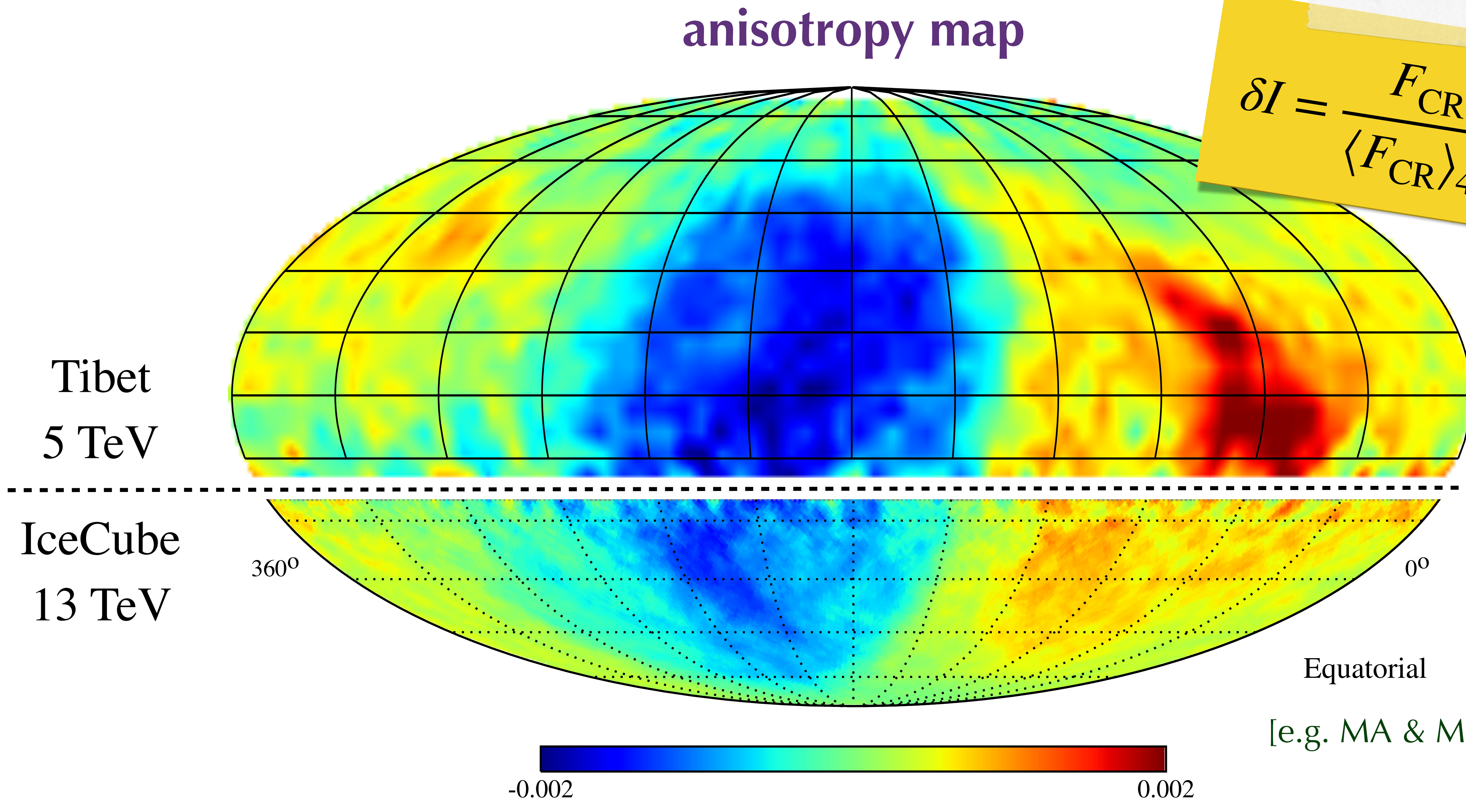


illustration of Milky Way
[Credit: NASA]

Galactic Cosmic Rays Anisotropy

Cosmic ray anisotropies up to the level of **one-per-mille** at various energies

(Super-Kamiokande, Milagro, ARGO-YBJ, EAS-TOP, Tibet ASγ, IceCube, HAWC)



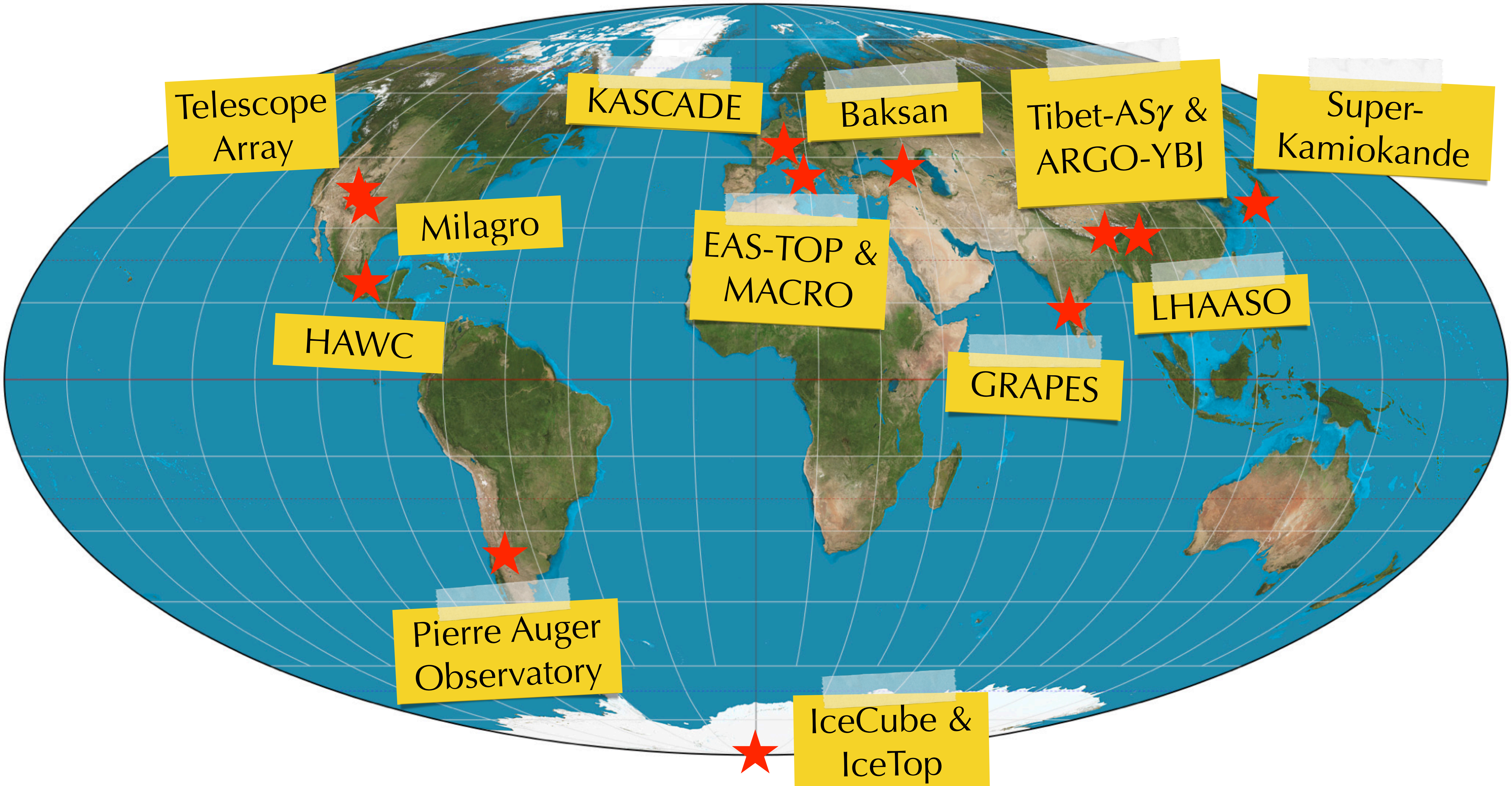
$$\delta I = \frac{F_{CR}}{\langle F_{CR} \rangle_{4\pi}} - 1$$

[e.g. MA & Mertsch'16]

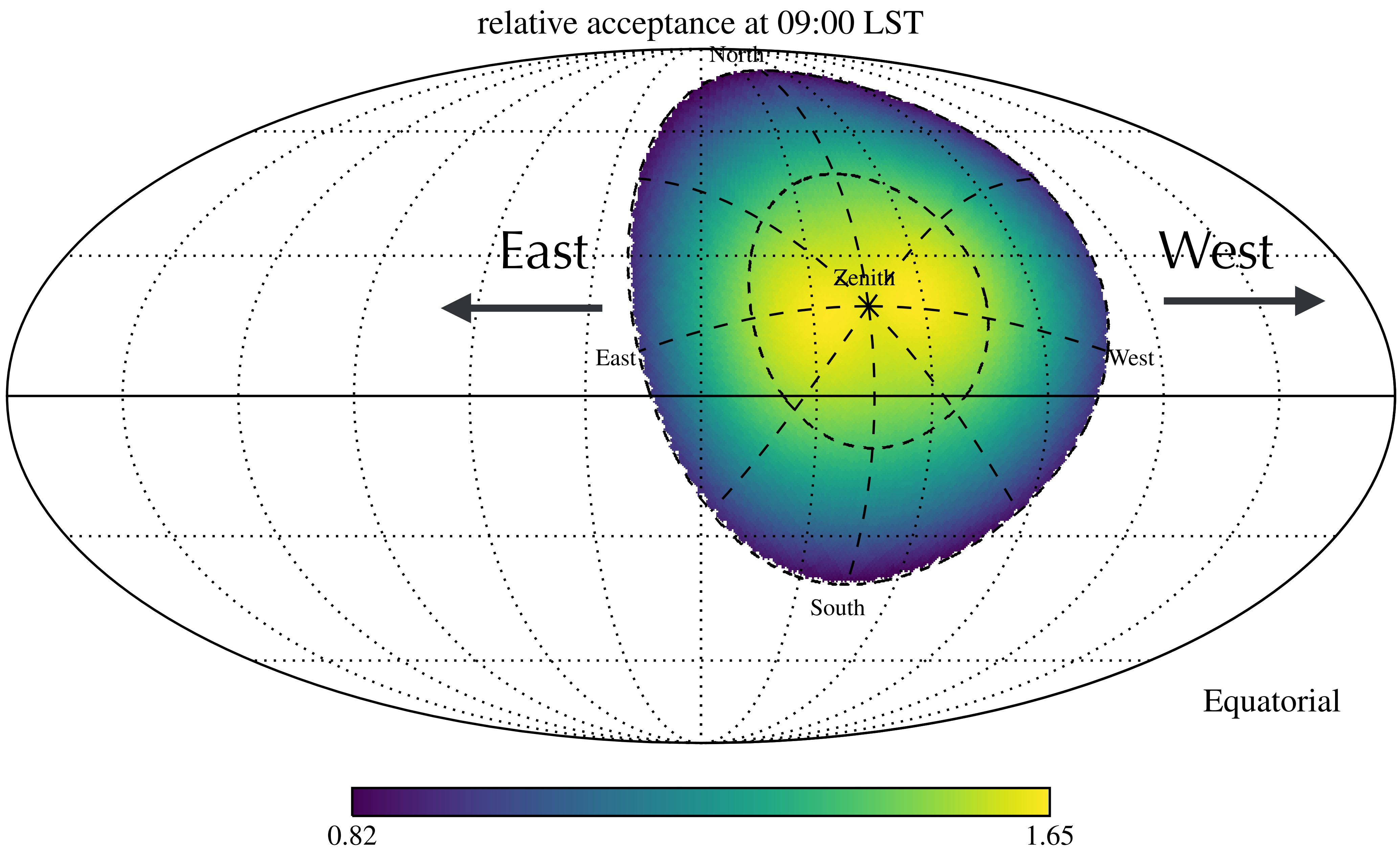
Galactic Cosmic Rays Anisotropy

Cosmic ray anisotropies up to the level of **one-per-mille** at various energies

(Super-Kamiokande, Milagro, ARGO-YBJ, EAS-TOP, Tibet ASγ, IceCube, HAWC)

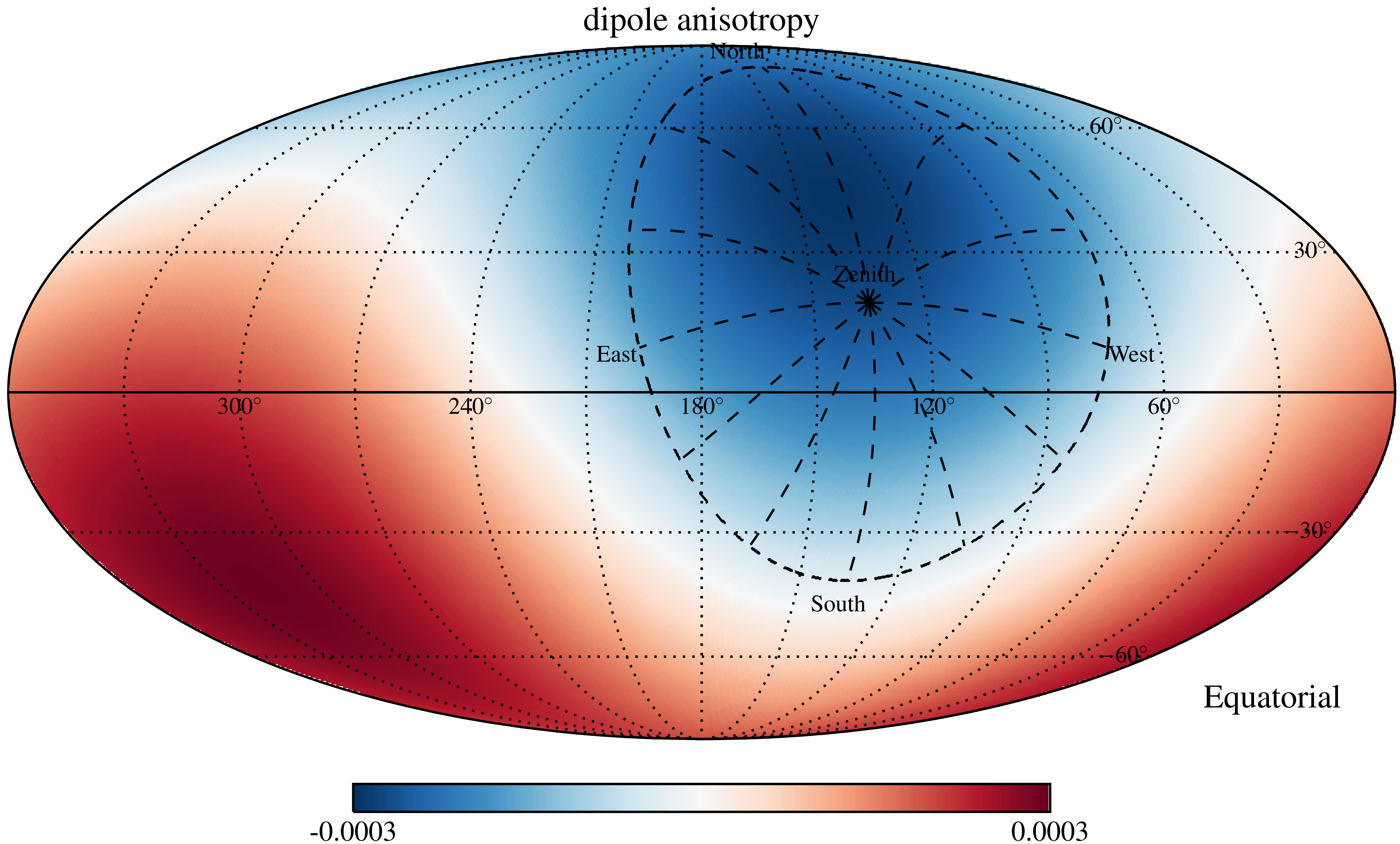


Ground-Based Observations



Field of View (FoV) of ground-based detector (*e.g.* HAWC at geographic latitude 19°) sweeps across the Sky over 24h.

Issues with Reconstructions

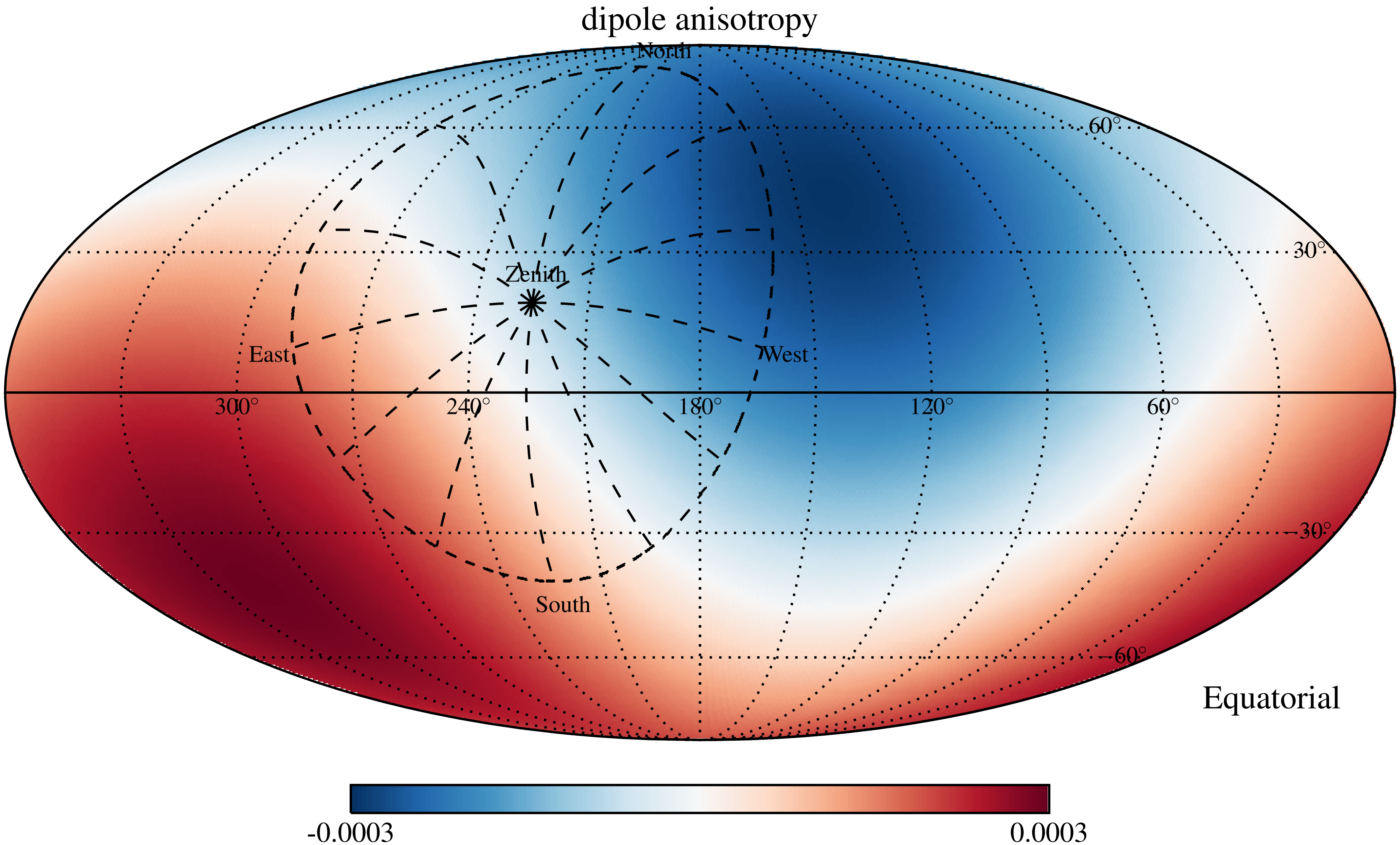


True CR dipole is defined by amplitude A and direction (α, δ).

Observable dipole is projected onto equatorial plane: $A' = A\cos\delta$

[Iuppa & Di Sciascio'13; MA *et al.*'15]

Issues with Reconstructions

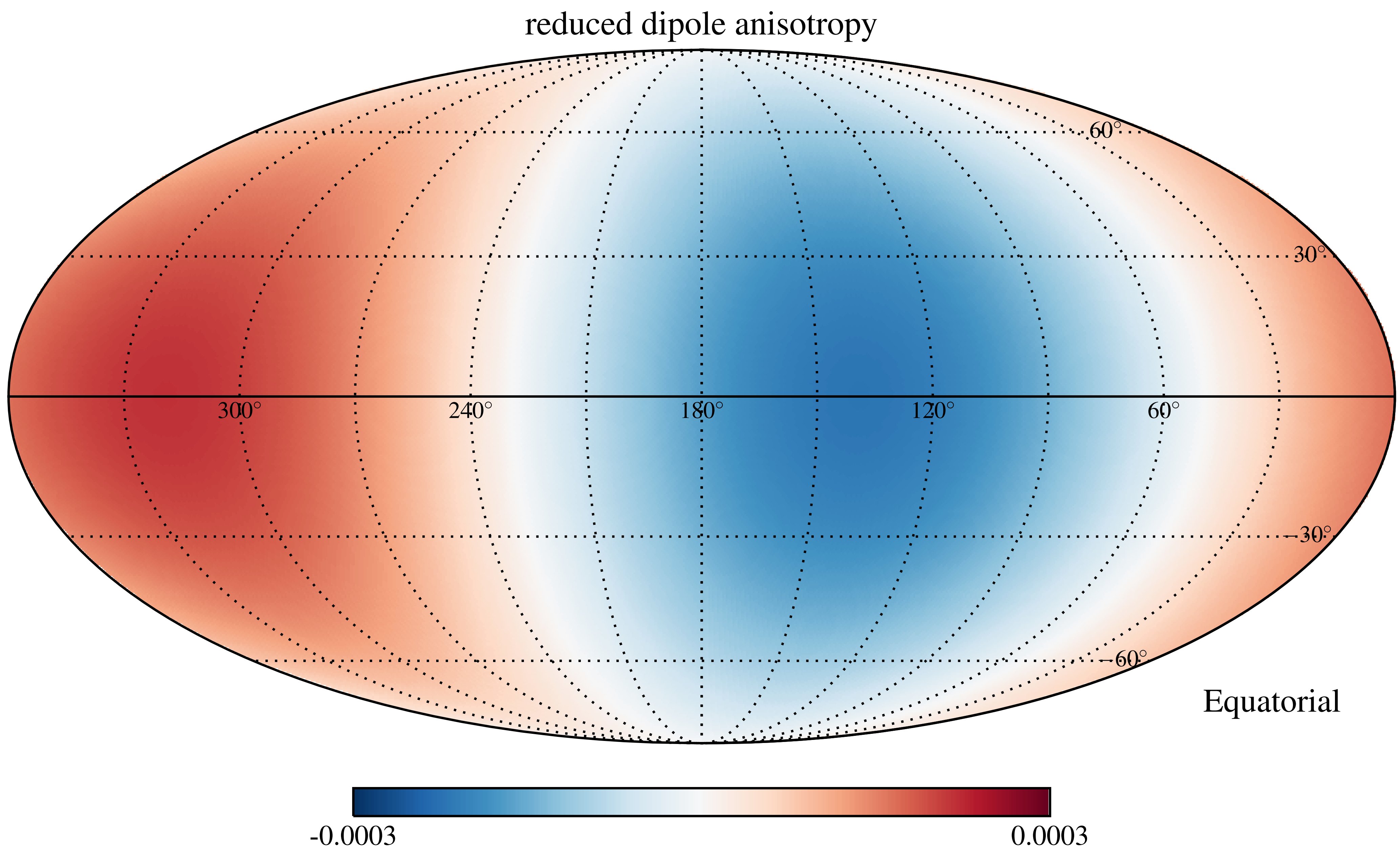


True CR dipole is defined by amplitude A and direction (α, δ).

Observable dipole is projected onto equatorial plane: $A' = A \cos\delta$

[Iuppa & Di Sciascio'13; MA *et al.*'15]

Issues with Reconstructions

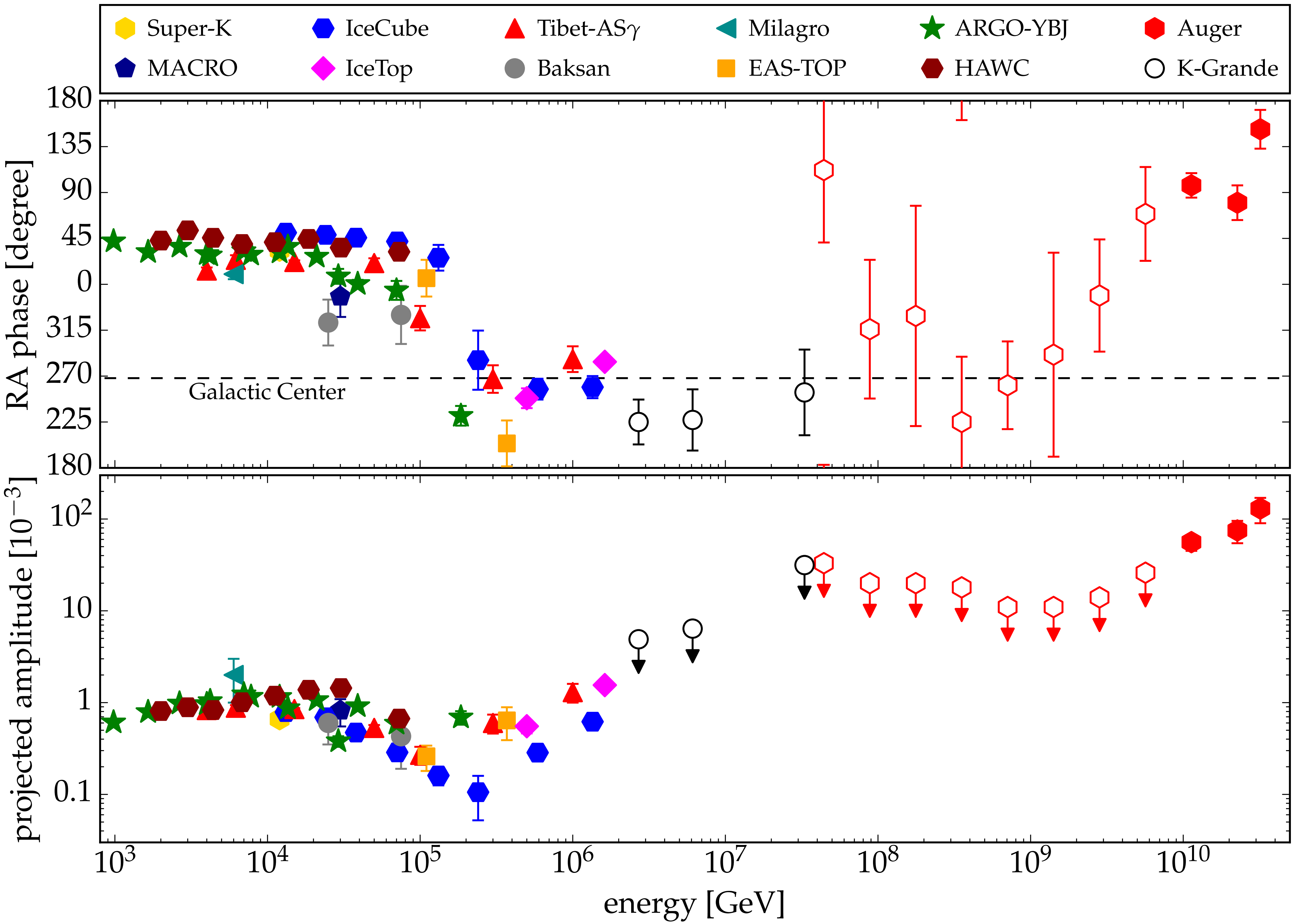


True CR dipole is defined by amplitude A and direction (α, δ).

Observable dipole is projected onto equatorial plane: $A' = A \cos\delta$

[Iuppa & Di Sciascio'13; MA *et al.*'15]

Dipole Anisotropy



Example: Pierre Auger >8 EeV

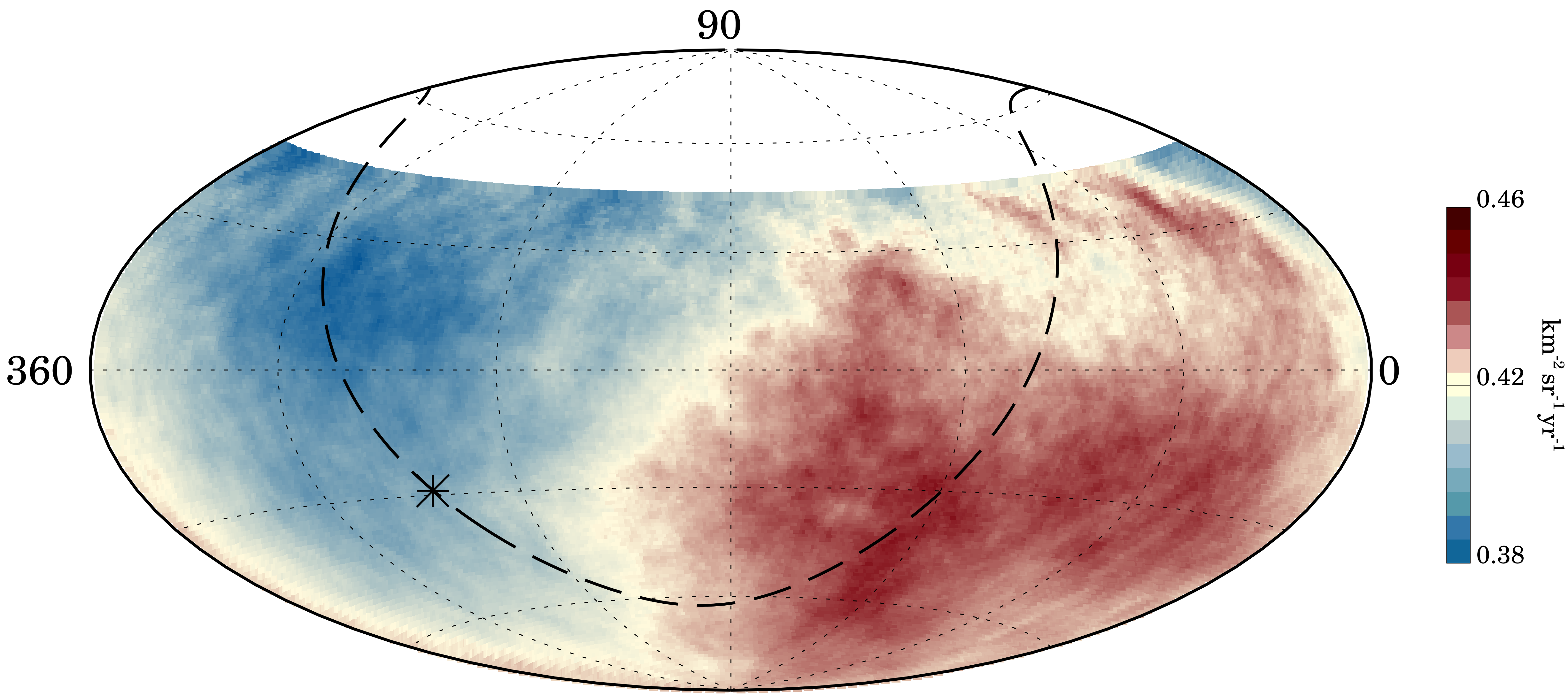
Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18} eV

The Pierre Auger Collaboration*†

Cosmic rays are atomic nuclei arriving from outer space that reach the highest energies observed in nature. Clues to their origin come from studying the distribution of their arrival directions. Using 3×10^4 cosmic rays with energies above 8×10^{18} electron volts, recorded with the Pierre Auger Observatory from a total exposure of 76,800 km^2 sr year, we determined the existence of anisotropy in arrival directions. The anisotropy, detected at more than a 5.2σ level of significance, can be described by a dipole with an amplitude of $6.5^{+1.3}_{-0.9}$ percent toward right ascension $\alpha_d = 100 \pm 10$ degrees and declination $\delta_d = -24^{+12}_{-13}$ degrees. That direction indicates an extragalactic origin for these ultrahigh-energy particles.

[Pierre Auger, Science 357 (2017) 6357]

Example: Pierre Auger >8 EeV



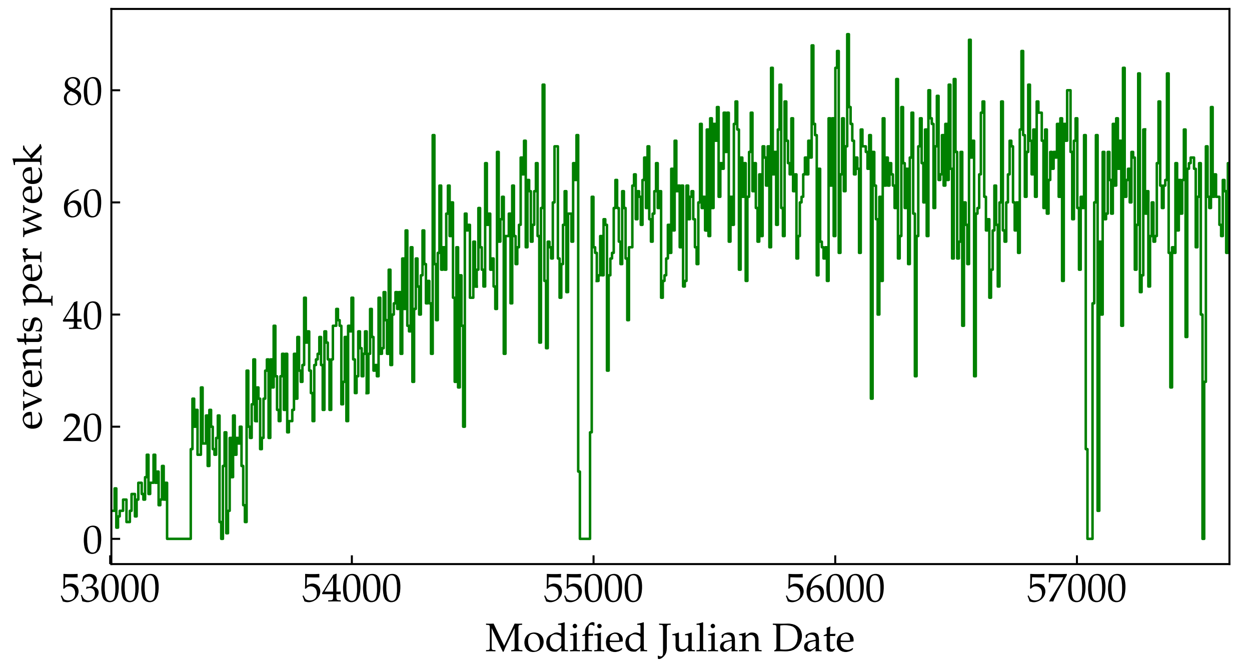
[Pierre Auger, Science 357 (2017) 6357]

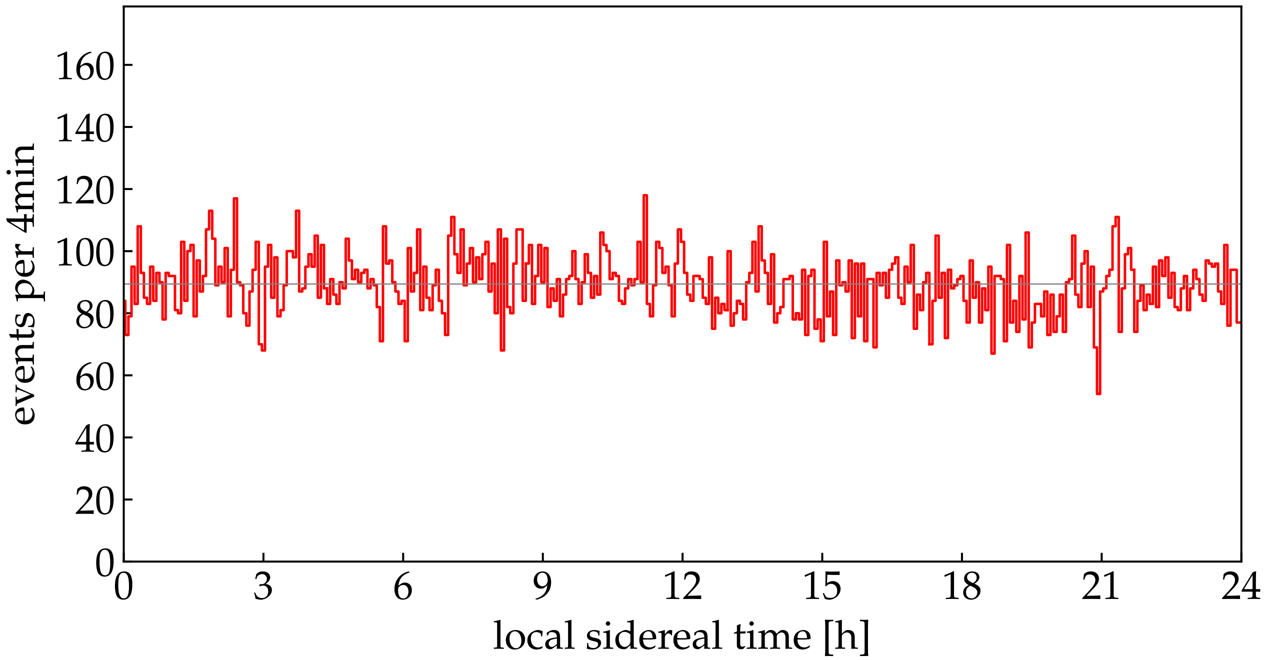
Energy [EeV]	Dipole component d_z	Dipole component $d_\perp$	Dipole amplitude d	Dipole declination δ_d [°]	Dipole right ascension α_d [°]
4 to 8	-0.024 ± 0.009	$0.006^{+0.007}_{-0.003}$	$0.025^{+0.010}_{-0.007}$	-75^{+17}_{-8}	80 ± 60
8	-0.026 ± 0.015	$0.060^{+0.011}_{-0.010}$	$0.065^{+0.013}_{-0.009}$	-24^{+12}_{-13}	100 ± 10

Reconstruction

- data has **strong time dependence**
  - detector deployment/ maintenance
  - atmospheric conditions (day/ night, seasons)
  - power outages, etc.
- **local anisotropy** of detector:
  - non-uniform geometry
- two analysis strategies:
  - **Monte-Carlo & monitoring** (systematic limited)
  - **data-driven LH methods** (statistics limited)

Example: Auger data > 8 EeV





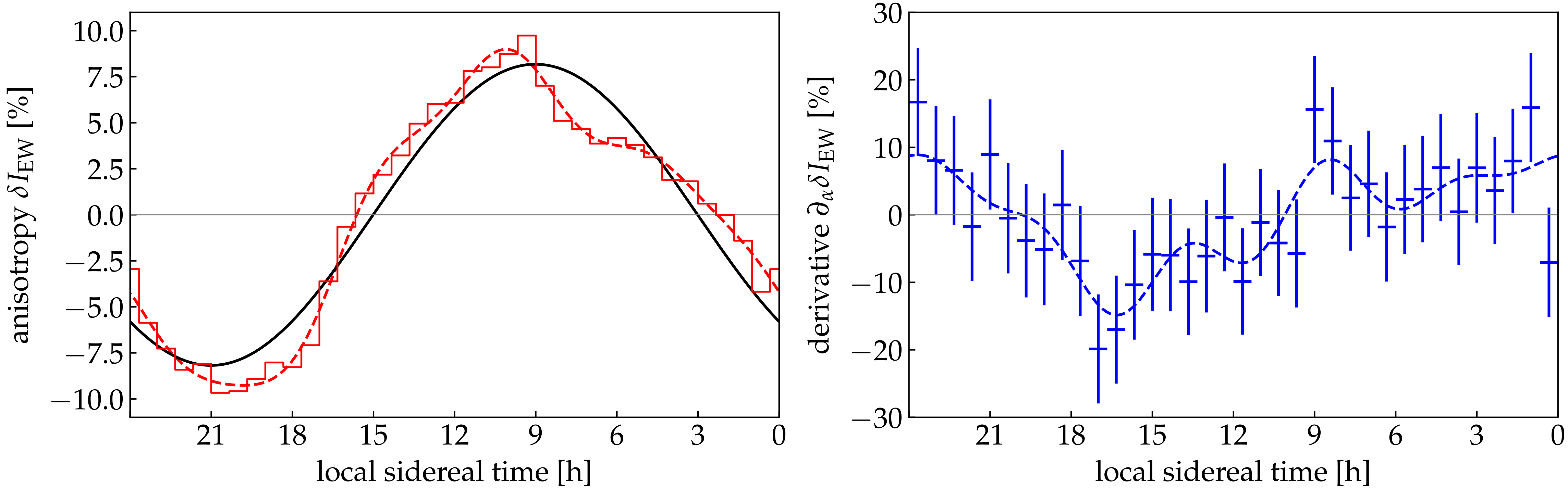
[Pierre Auger Observatory'17; MA'18]

East-West Method

- Strong time variation of CR background level can be compensated by differential methods. [e.g. Bonino *et al.*'11]
- **East-West asymmetry**:

$$A_{\rm EW}(t) \equiv \frac{N_{\rm E}(t) - N_{\rm W}(t)}{N_{\rm E}(t) + N_{\rm W}(t)} \simeq \underbrace{\Delta\alpha \frac{\partial}{\partial\alpha}\delta I(\alpha,0)}_{\text{assuming dipole!}} + \underbrace{\text{const}}_{\text{local asym.}}$$

- For instance, Auger data > 8EeV:



`Auger_EastWest_2017.ipynb`

Likelihood Reconstruction

- East-West method introduces cross-talk between higher multipoles, regardless of the field of view.
- Alternatively, data can be analyzed to simultaneously reconstruct:
  - **relative acceptance** $\mathscr{A}(\varphi, \theta)$ (in local coordinates)
  - **relative intensity** $\mathscr{I}(\alpha, \delta)$ (in equatorial coordinates)
  - **background rate** $\mathscr{N}(t)$ (in sidereal time)
- expected number of CRs in sidereal time bin τ and local "pixel" i:

$$\mu_{\tau i} = \mu(\mathscr{I}_{\tau i}, \mathscr{N}_{\tau}, \mathscr{A}_{i})$$

- reconstruction **likelihood**:

$$\mathscr{L}(\mathbf{n}\,|\,\mathscr{I}, \mathscr{N}, \mathscr{A}) = \prod_{\tau i} \frac{1}{n_{\tau i}!} (\mu_{\tau i})^{n_{\tau i}} e^{-\mu_{\tau i}}$$

- Maximum LH can be reconstructed by iterative methods. [*MA et al.*'15]

Iterative Method

- Expected number of events:

$$\mu_{\tau i} = \mathcal{I}_{\tau i}\mathcal{N}_{\tau}\mathcal{A}_{i}$$

- Maximum LH values can be solved implicitly: [*MA et al.*'15]

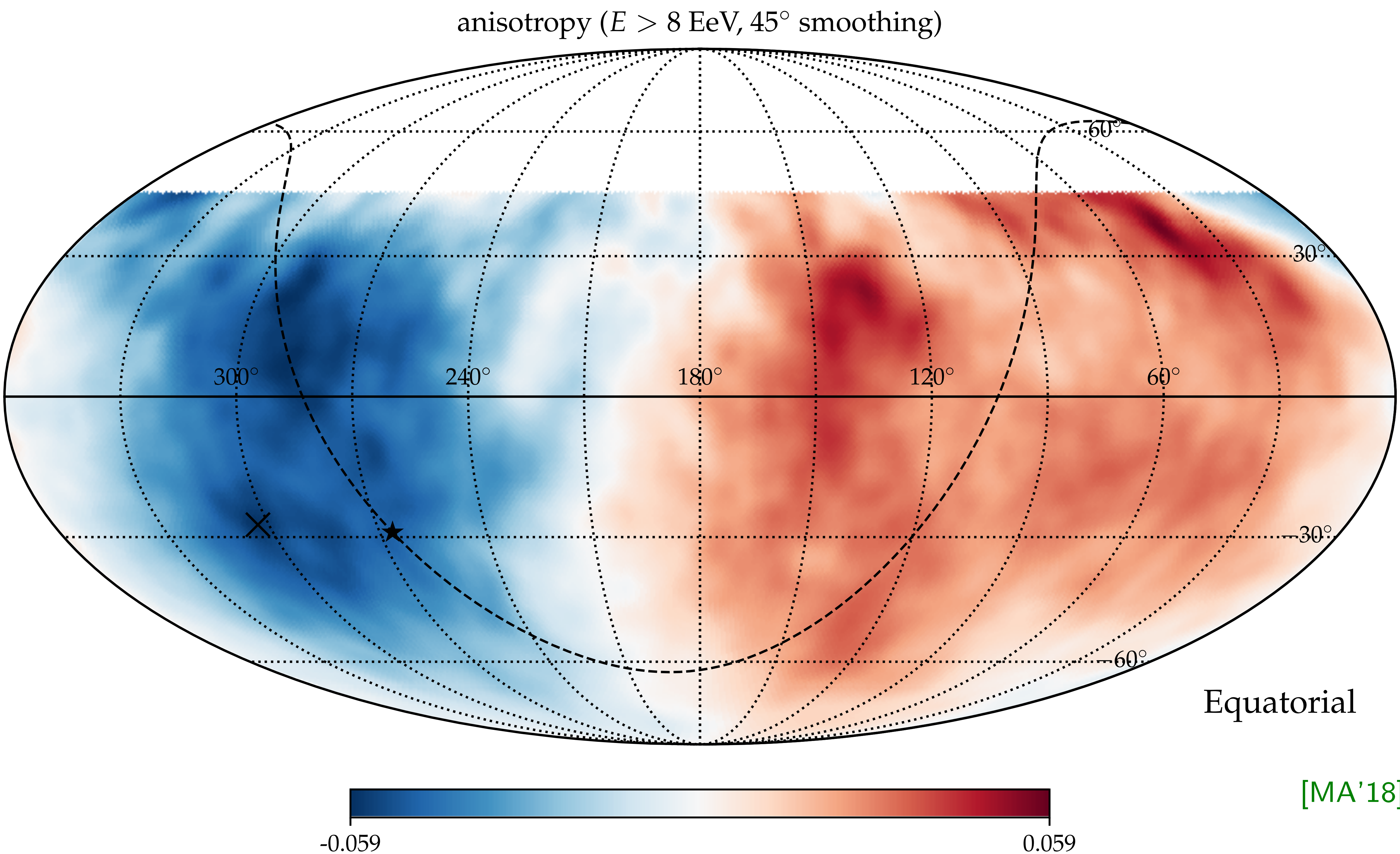
(a) $$\hat{I}_{\mathfrak{a}} = \sum_{\tau} n_{\tau\mathfrak{a}} \Big/ \sum_{\kappa} \hat{\mathcal{A}}_{\kappa\mathfrak{a}}\hat{\mathcal{N}}_{\kappa}$$ ($\mathfrak{a}$: pixel EQ map)

(b) $$\hat{\mathcal{N}}_{\tau} = \sum_{i} n_{\tau i} \Big/ \sum_{j} \hat{\mathcal{A}}_{j}\hat{I}_{\tau j}$$ (τ : sidereal bin)

(c) $$\hat{\mathcal{A}}_{i} = \sum_{\tau} n_{\tau i} \Big/ \sum_{\kappa} \hat{\mathcal{N}}_{\kappa}\hat{I}_{\kappa i}$$ (i : pixel in local map)

- Start from $\hat{I}_{\mathfrak{a}} = \mathrm{const}$ and progressively iterate steps (a), (b) & (c).

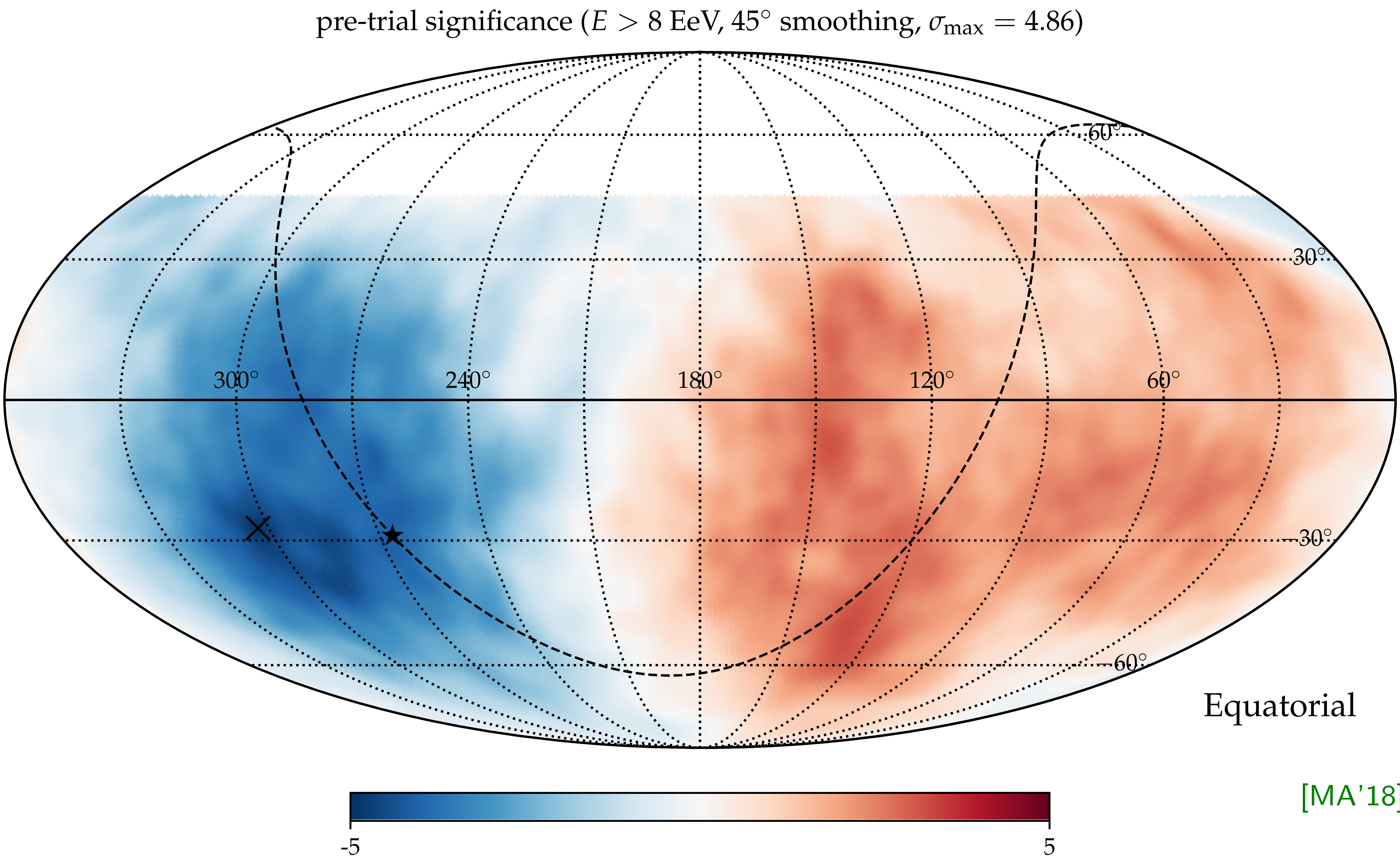
Likelihood Reconstruction



[MA'18]

`Auger_anisotropy_2017.ipynb`

Likelihood Reconstruction

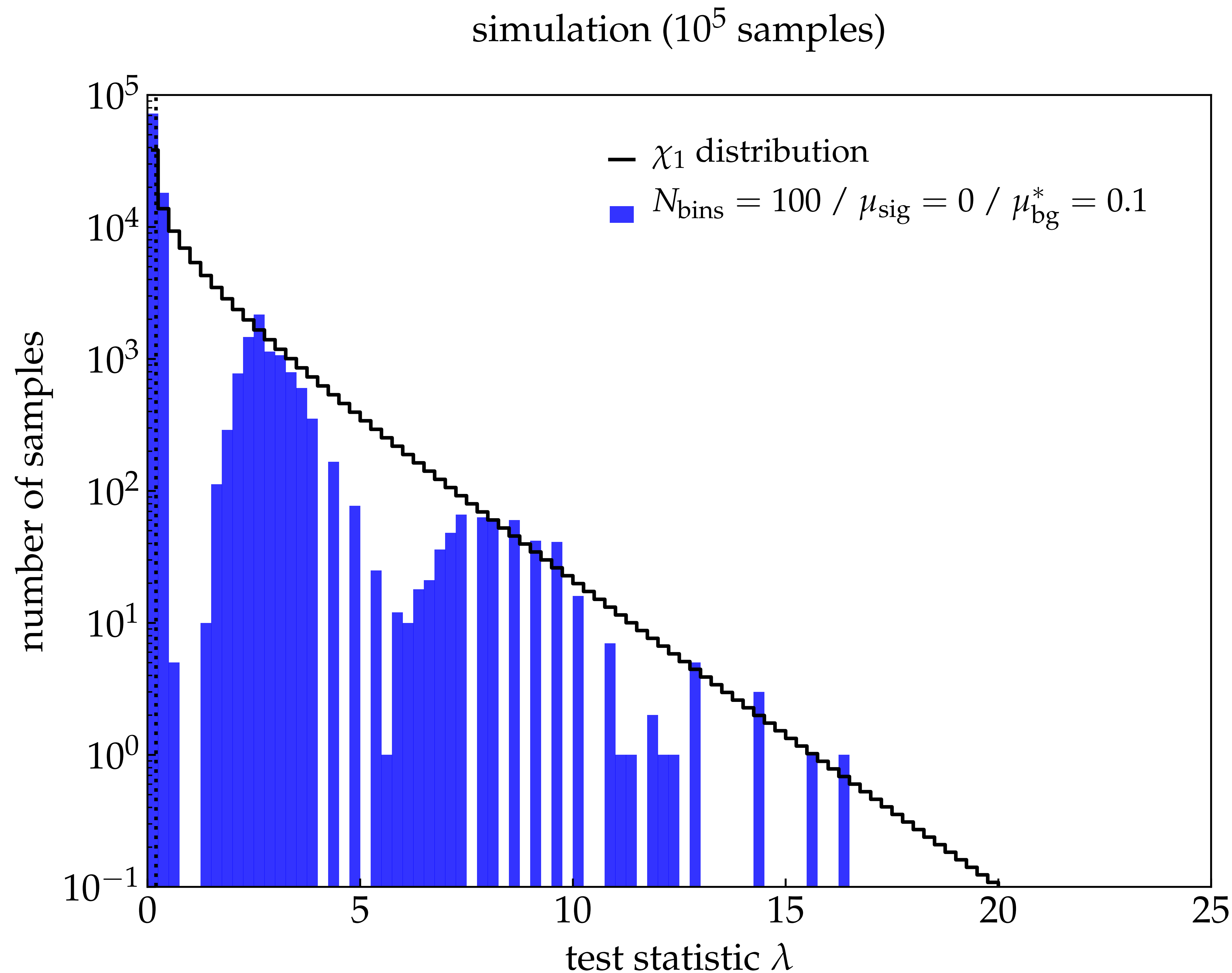


[MA'18]

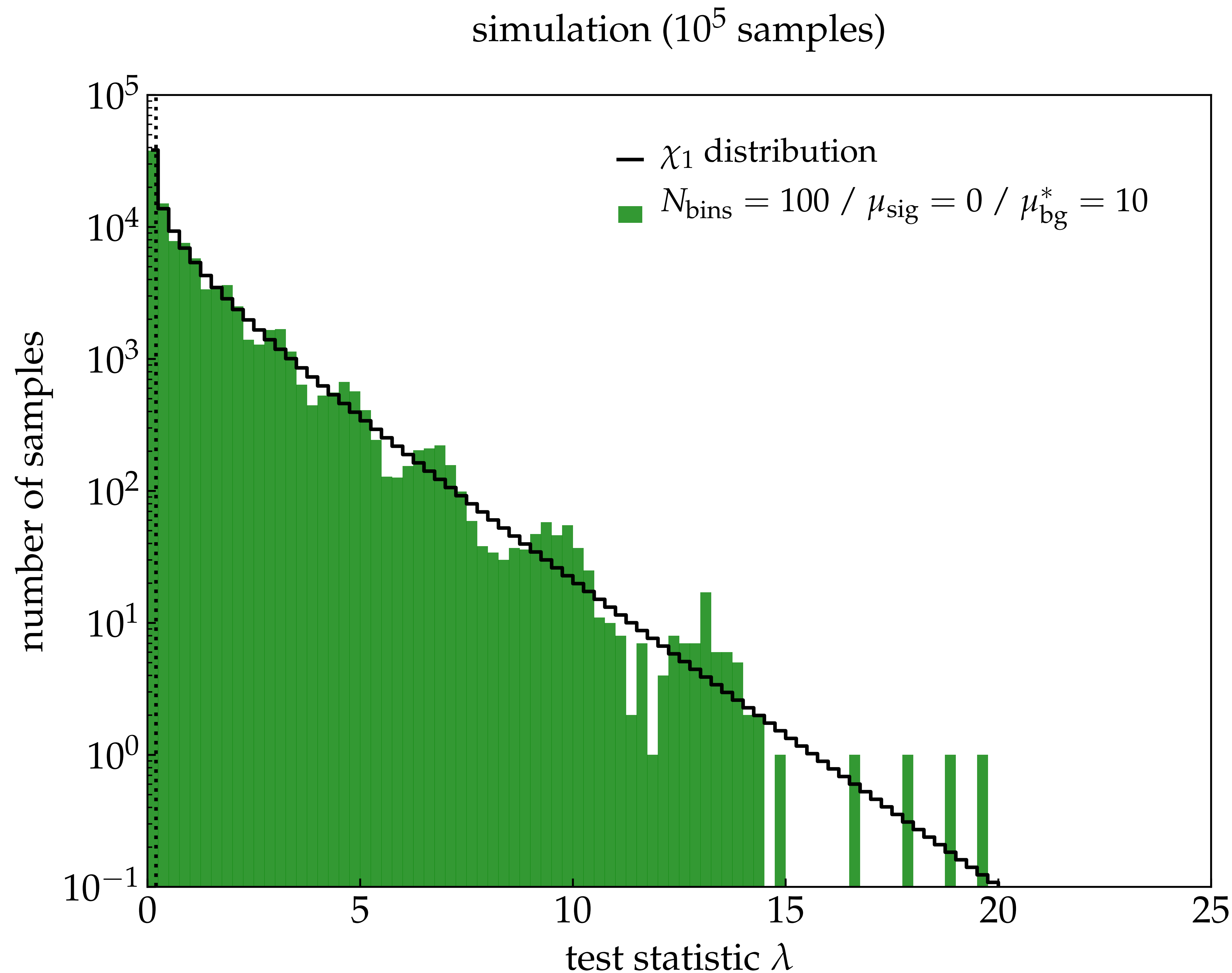
`Auger_anisotropy_2017.ipynb`

Appendix

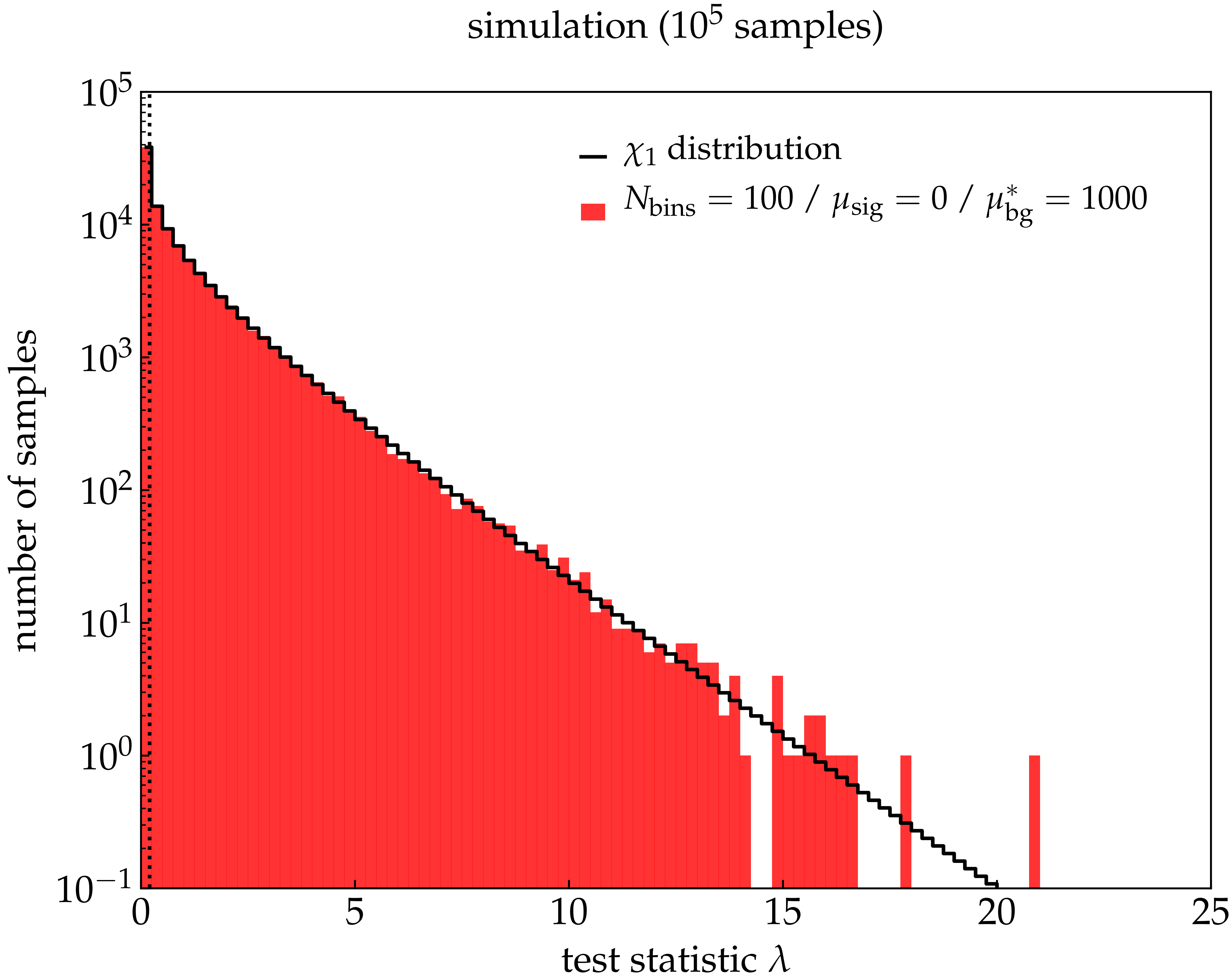
Maximum-Log-Likelihood Ratio



Maximum-Log-Likelihood Ratio



Maximum-Log-Likelihood Ratio



Wilks' Theorem

THE LARGE-SAMPLE DISTRIBUTION OF THE LIKELIHOOD RATIO FOR TESTING COMPOSITE HYPOTHESES[1]

BY S. S. WILKS

(. . .)

Theorem: If a population with a variate x is distributed according to the probability function $f(x, \theta_1, \theta_2 \cdots \theta_h)$, such that optimum estimates $\bar{\theta}_i$ of the θ_i exist which are distributed in large samples according to (3), then when the hypothesis H is true that $\theta_i = \theta_{0i}$, $i = m + 1, m + 2, \cdots h$, the distribution of $-2 \log \lambda$, where λ is given by (2) is, except for terms of order $1/\sqrt{n}$, distributed like χ^2 with $h - m$ degrees of freedom.

[Wilks, Annals Math.Statist. 9 (1938) 1]

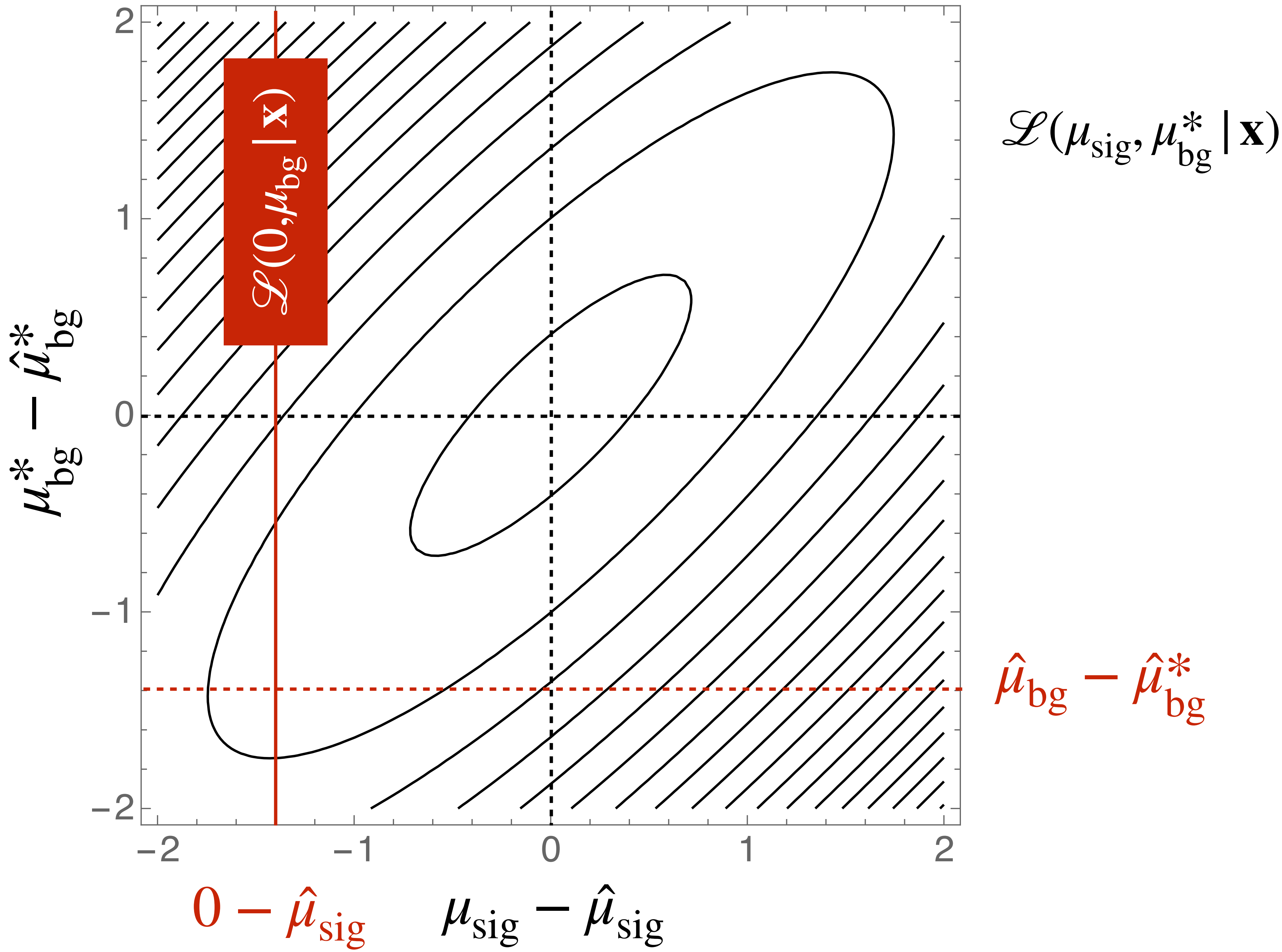
Wilks' Theorem

- *Prerequisites*:
  - Let $\mathbf{x}$ be data that follows a probability function $f(\mathbf{x}\,|\,\theta_1,\ldots,\theta_n)$.
  - **Unconstrained likelihood** $\mathscr{L}(\theta_1,\ldots,\theta_n\,|\,\mathbf{x})$ has maximum at $\hat\theta_1,\ldots,\hat\theta_n$.
  - True hypothesis is $\theta_1^{(0)},\ldots,\theta_m^{(0)}$ with $m < n$.
  - **Constrained likelihood** $\mathscr{L}(\theta_1^{(0)},\ldots,\theta_m^{(0)},\theta_{m+1},\ldots,\theta_n\,|\,\mathbf{x})$ has maximum at $\hat\theta'_{m+1},\ldots,\hat\theta'_n$.
- For a large number of samples $\mathbf{x}$, the distribution of the test statistic:

$$\lambda(\mathbf{x}) \equiv -2\ln\frac{\mathscr{L}(\theta_1^{(0)},\ldots,\theta_m^{(0)},\hat\theta'_{m+1},\ldots,\hat\theta'_n\,|\,\mathbf{x})}{\mathscr{L}(\hat\theta_1,\ldots,\hat\theta_n\,|\,\mathbf{x})}$$

approaches a χ^2_k **distribution** with $k = n - m$ in the limit of large $N_{\rm tot}$.

Wilks' Theorem



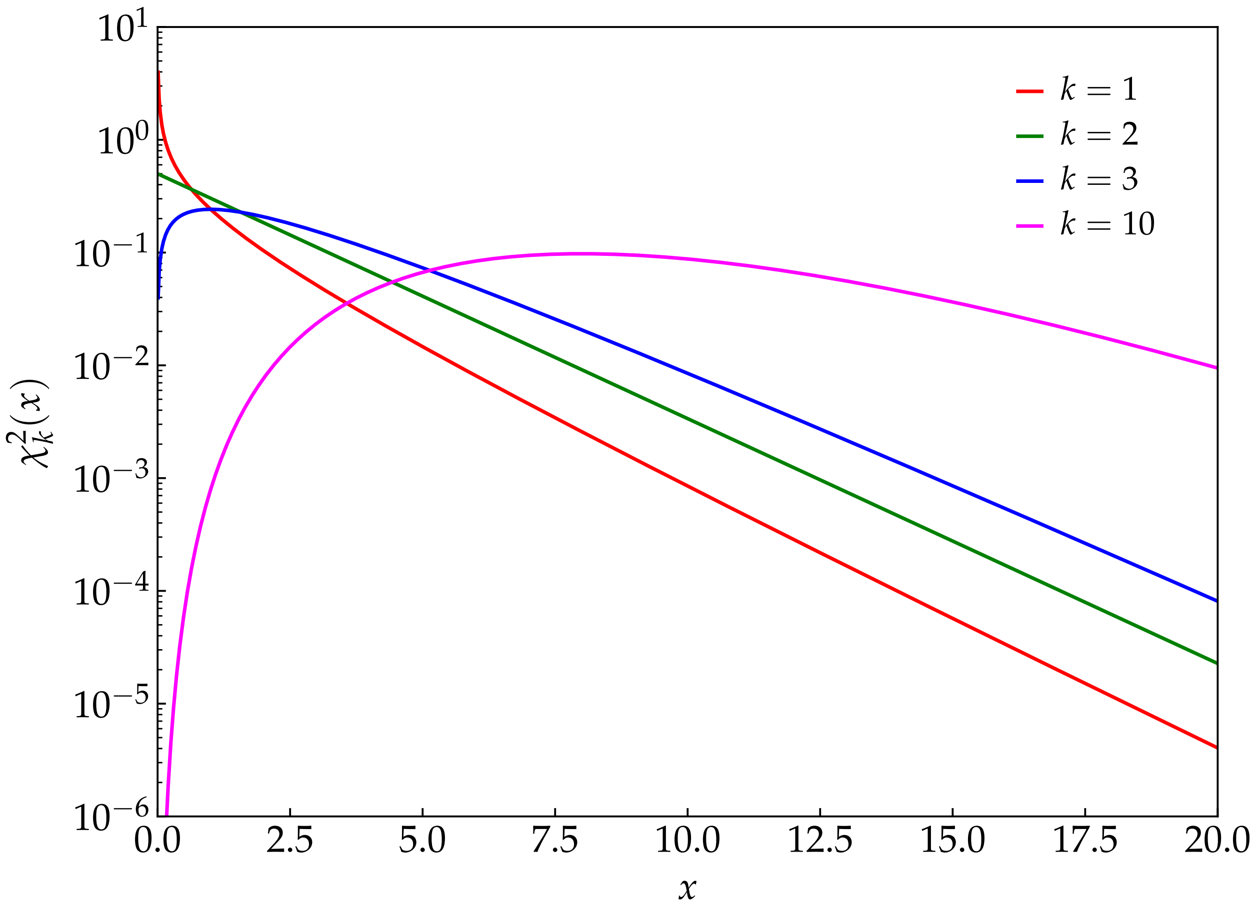
Chi-Square Distributions

- Definition of the χ^2_k **distribution:**

$$\chi^2_k = \frac{x^{k/2-1}e^{-x/2}}{2^{k/2}\Gamma(k/2)}$$

- **degrees of freedom** in our example:

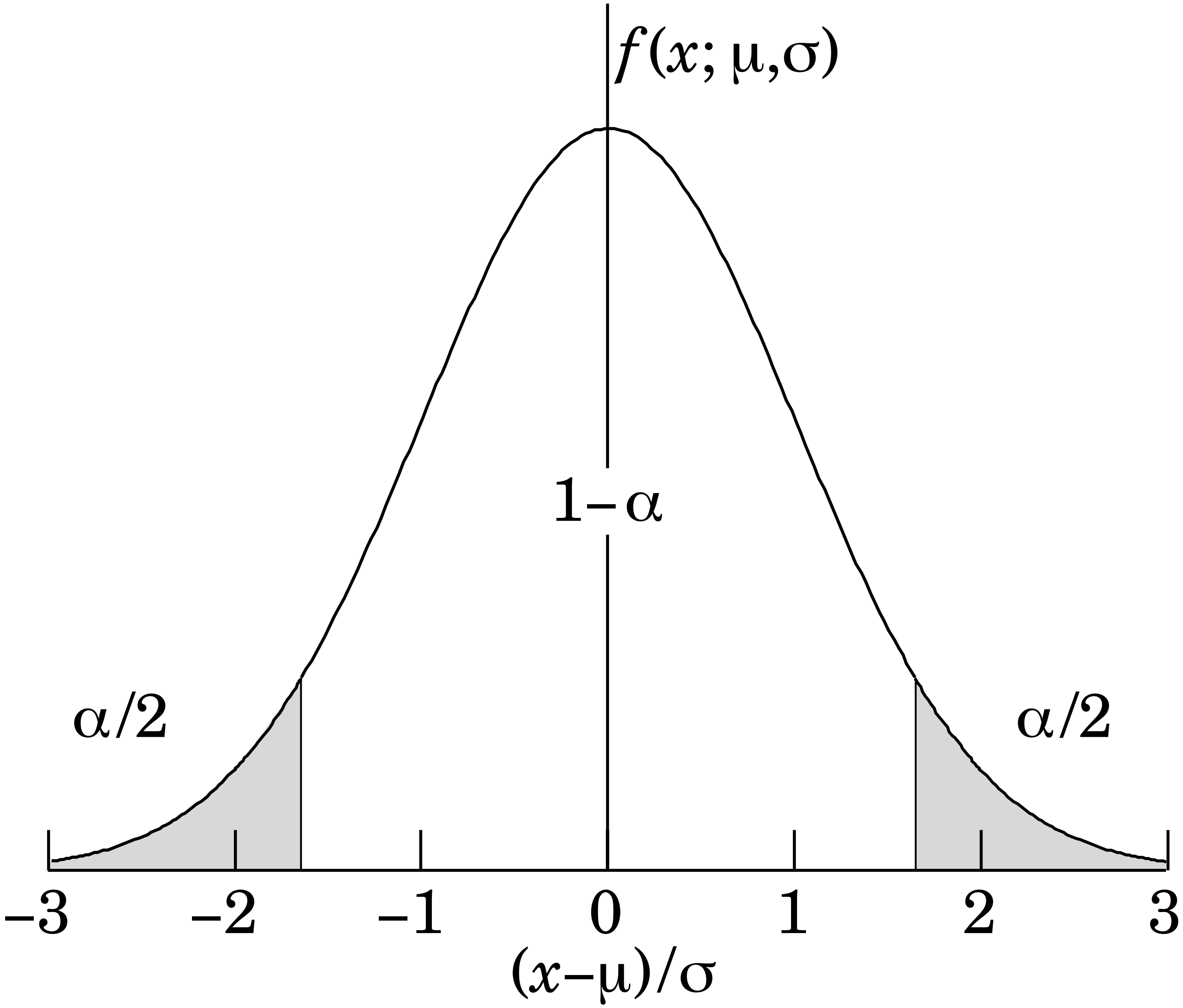
$$k = 2 - 1 = 1$$



- integral over χ^2_k distribution is related to the integrated probability of a k-variate normal distribution:

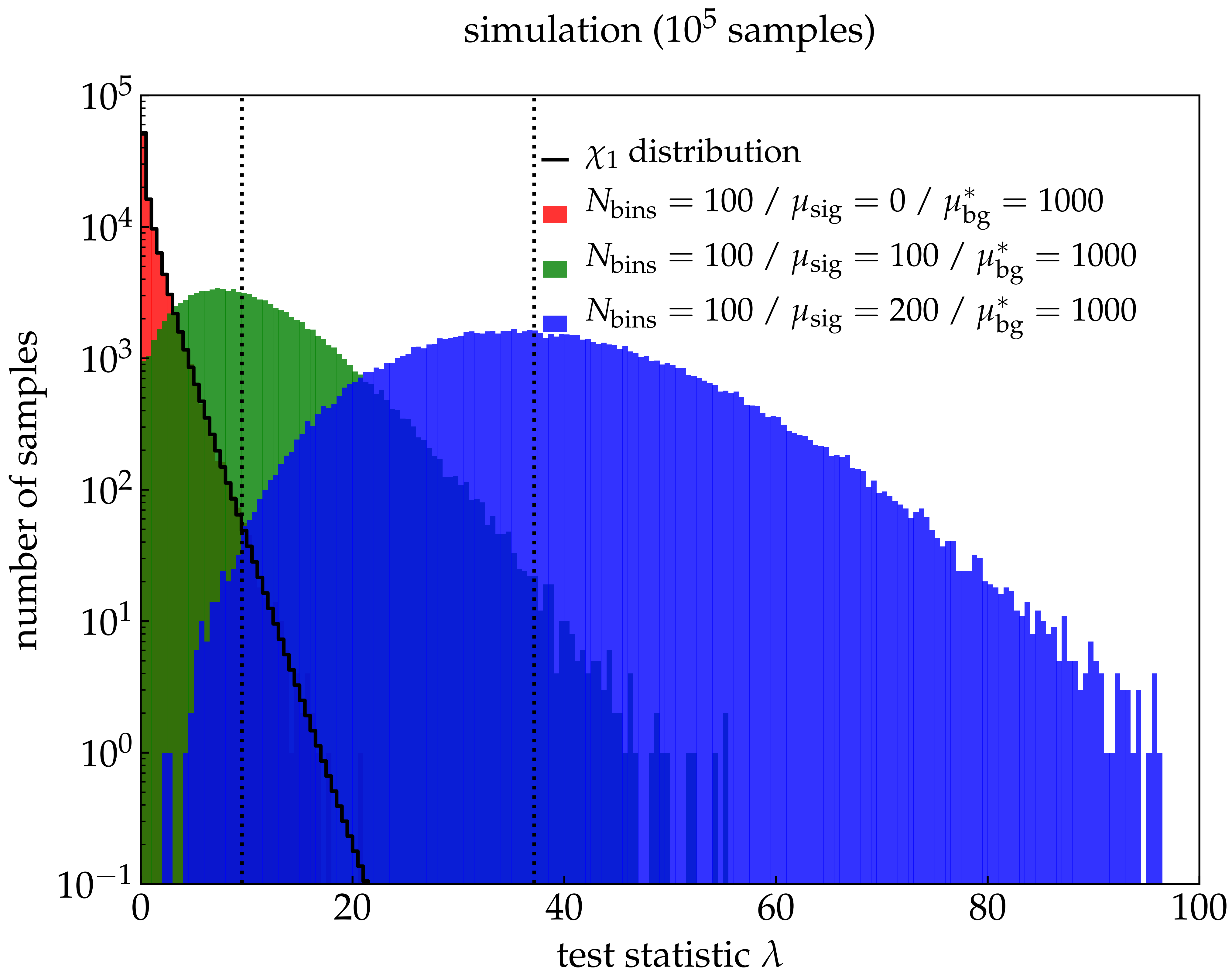
$$\int\limits_{\lambda_{\rm obs}} {\rm d}\lambda\, \chi^2_k(\lambda) = \int\limits_{\mathbf{r}^T\boldsymbol{\Sigma}^{-1}\mathbf{r}>\lambda_{\rm obs}} {\rm d}r_1\ldots{\rm d}r_k \frac{1}{\sqrt{(2\pi)^k \det\boldsymbol{\Sigma}}} \exp(-\mathbf{r}^T\boldsymbol{\Sigma}^{-1}\mathbf{r}/2)$$

Chi-Square Distributions



$$\alpha = \int_{\lambda_{obs}} \mathrm{d}\lambda\, \chi^2_1(\lambda) = 1 - \mathrm{erf}\left(\sqrt{\lambda_{obs}/2}\right)$$

Application of Wilks' Theorem



Application of Wilks' Theorem

- For large $N_{\rm tot}$ we can apply Wilks' theorem and calculate the **"p-value"** of an observed excess:

$$p = \int_{\lambda_{\rm obs}}^{\infty} {\rm d}\lambda\, \chi^2_k(\lambda) = 1 - {\rm erf}(\sqrt{\lambda_{\rm obs}/2})$$

- From signal simulations ($\mu_{\rm bg} = 1000$, $N_{\rm bins} = 100$), we can determine the median test statistic and corresponding **significance level,** *e.g.*:
  - $\mu_{\rm sig} = 100 \;\rightarrow\; \lambda_{\rm med} \simeq 9.8 \;\rightarrow$ Wilks' theorem: $p_{\rm med} \simeq 0.0017$
  - $\mu_{\rm sig} = 200 \;\rightarrow\; \lambda_{\rm med} \simeq 38.0 \;\rightarrow$ Wilks' theorem: $p_{\rm med} \simeq 7.1 \times 10^{-10}$
- The **5σ discovery threshold** corresponds to $\mu_{\rm sig} \simeq 162$ events.

Application of Wilks' Theorem

- **discovery potential:**

Level of $\mu_{\rm sig}$ such that $50\,\%$ of samples have a chance probability of less than 5.7×10^{-7} to be generated by background only.

- This is a challenge for brute-force background simulation – you need $N_{\rm samples} \gg 10^7$ for accuracy!

- **Wilks' theorem** allows to extrapolate the background distribution much simpler:

Level of $\mu_{\rm sig}$ such that $50\,\%$ of samples have $\lambda \geq \lambda_{\rm threshold} = 5^2 = 25.$

Li-Ma Formula

ANALYSIS METHODS FOR RESULTS IN GAMMA-RAY ASTRONOMY

TI-PEI LI AND YU-QIAN MA
High Energy Astrophysics Group, Institute of High Energy Physics,
Academia Sinica, Beijing, China

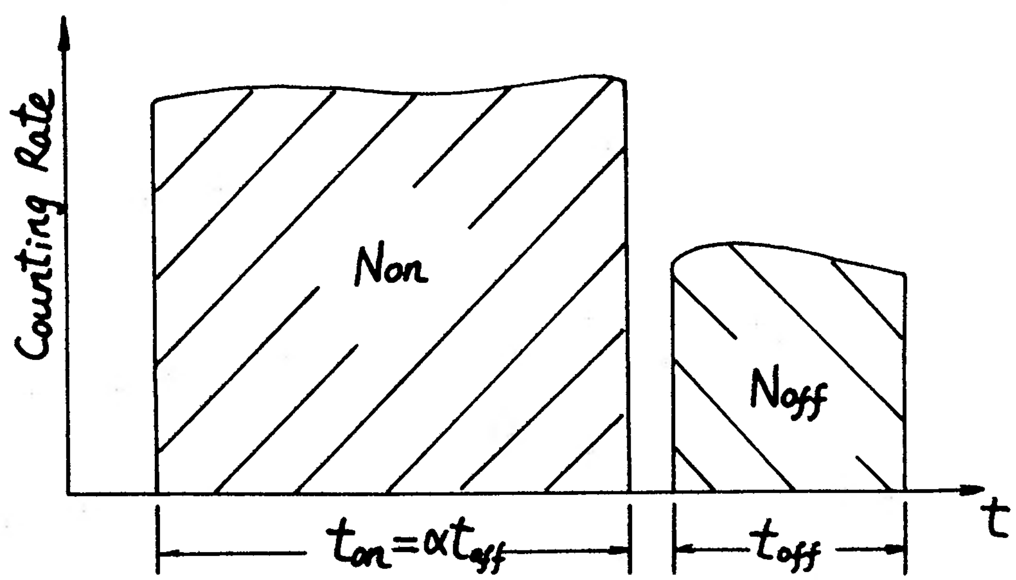
Received 1982 September 20; accepted 1983 February 7

ABSTRACT

The current procedures for analyzing results of γ-ray astronomy experiments are examined critically. We propose two formulae to estimate the significance of positive observations in searching γ-ray sources or lines. The correctness of the formulae are tested by Monte Carlo simulations.

[Li & Ma, ApJ 272 (1983) 317]

$$S = \sqrt{-2\ln\lambda} = \sqrt{2}\left\{N_{\rm on}\ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{\rm on}}{N_{\rm on}+N_{\rm off}}\right)\right] + N_{\rm off}\ln\left[(1+\alpha)\left(\frac{N_{\rm off}}{N_{\rm on}+N_{\rm off}}\right)\right]\right\}^{1/2}$$



$$N_{\rm on} \to x_1$$

$$N_{\rm off} \to N_{\rm tot} - x_1$$

$$\alpha^{-1} \to N_{\rm bins} - 1$$