Introduction to gammapy

Rubén López-Coto, IAA-CSIC
Enrique Mestre, ICE-CSIC
9th ICE Summer School, 08/07/26

CTAO
Cherenkov Telescope Array Observatory

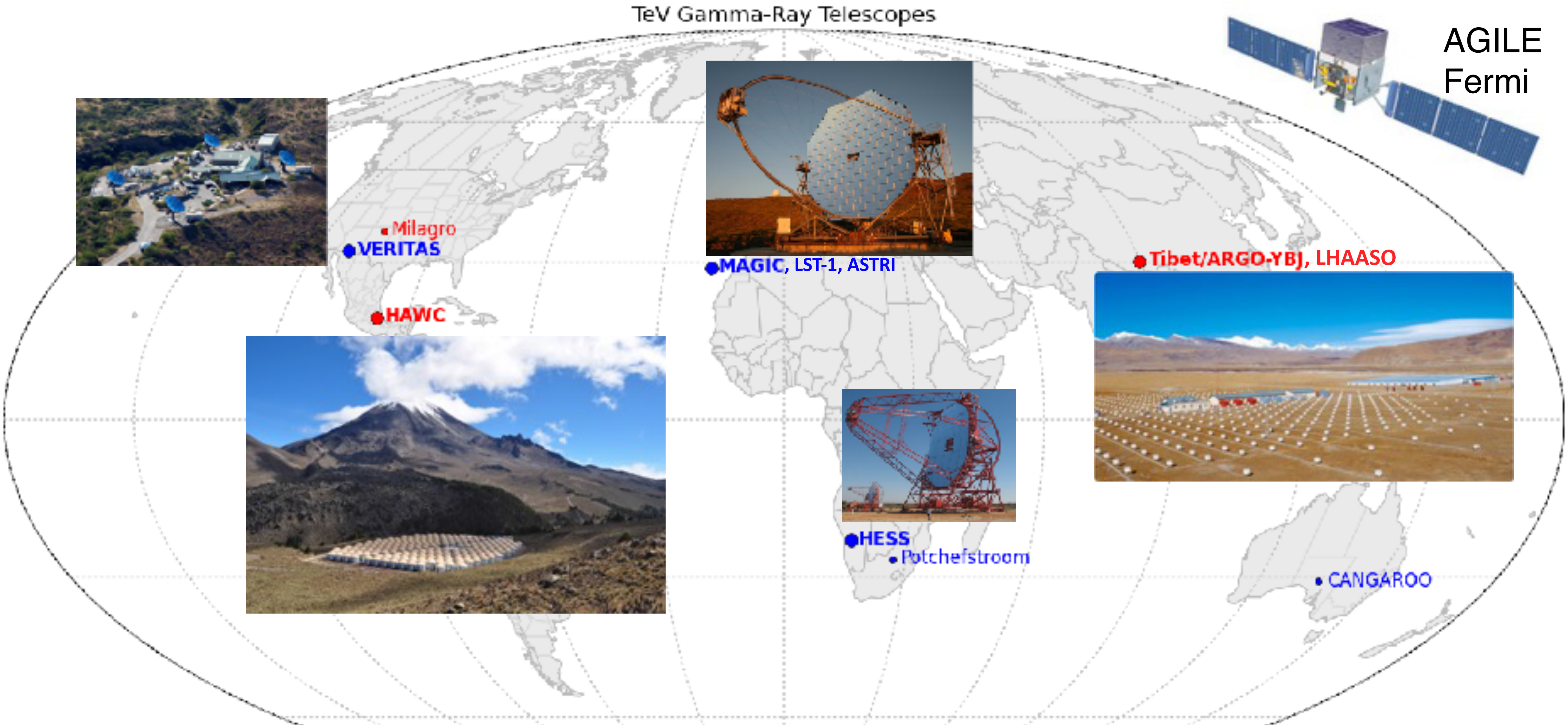
Introduction to gammapy

Special thanks to K.Feijen, F. Acero, A. Sinha, D. Morcuende, B. Khélifi… from whom we ~~stole~~ borrowed some slides

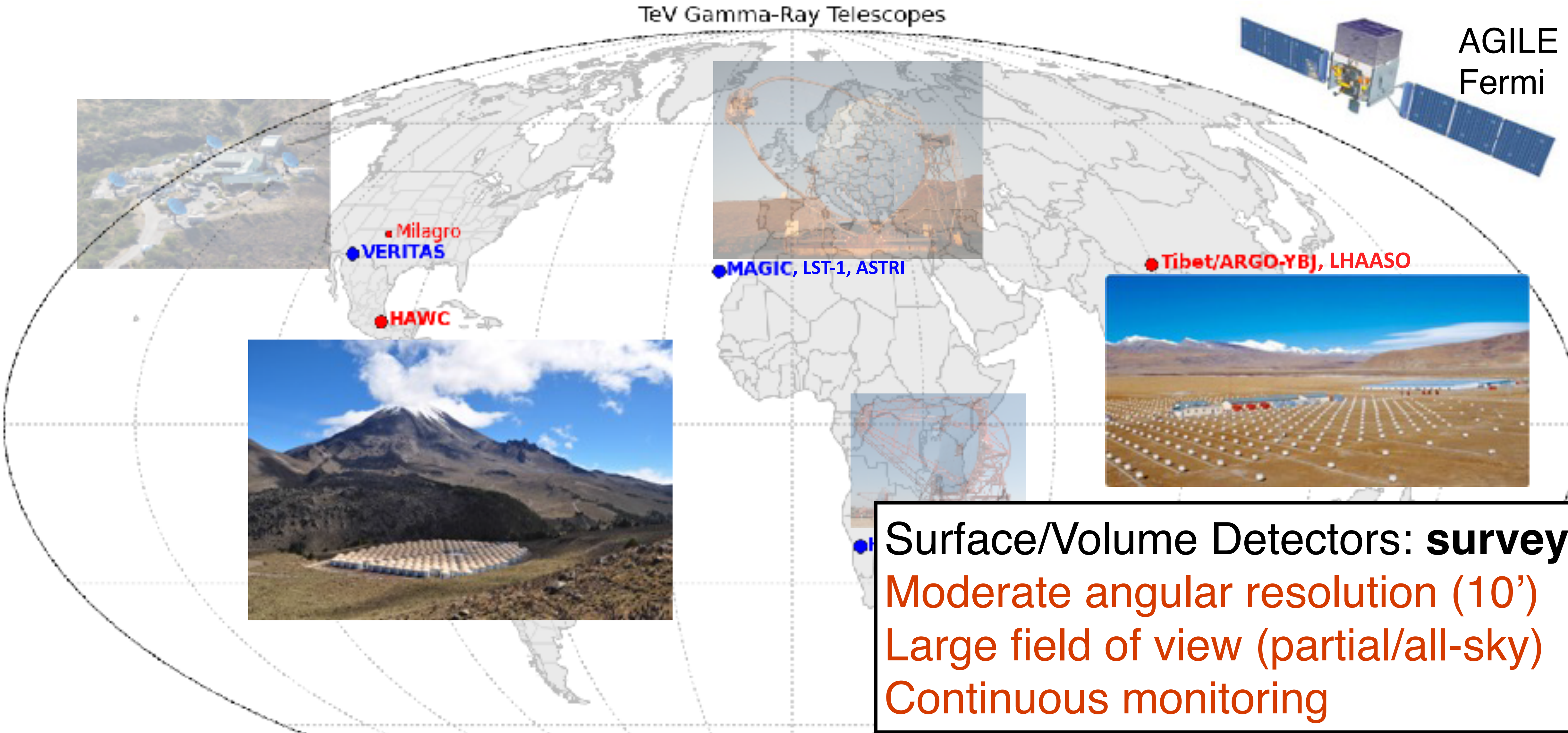
Rubén López-Coto, IAA-CSIC
Enrique Mestre, ICE-CSIC
9th ICE Summer School, 08/07/26

VHEGA
IAA INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA
EXCELENCIA SEVERO OCHOA

Gamma-ray astronomy



Gamma-ray astronomy



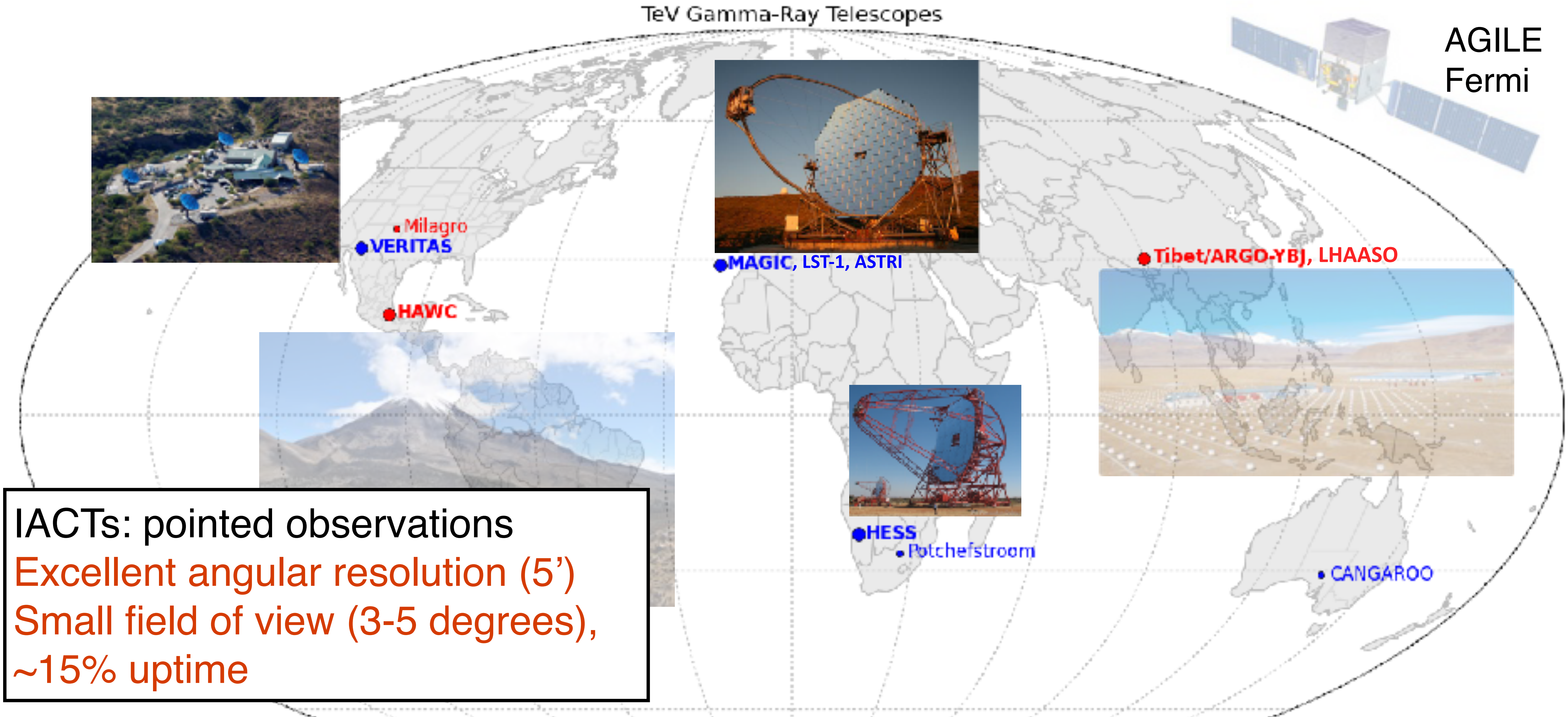
Surface/Volume Detectors: **surveys**

Moderate angular resolution (10’)

Large field of view (partial/all-sky)

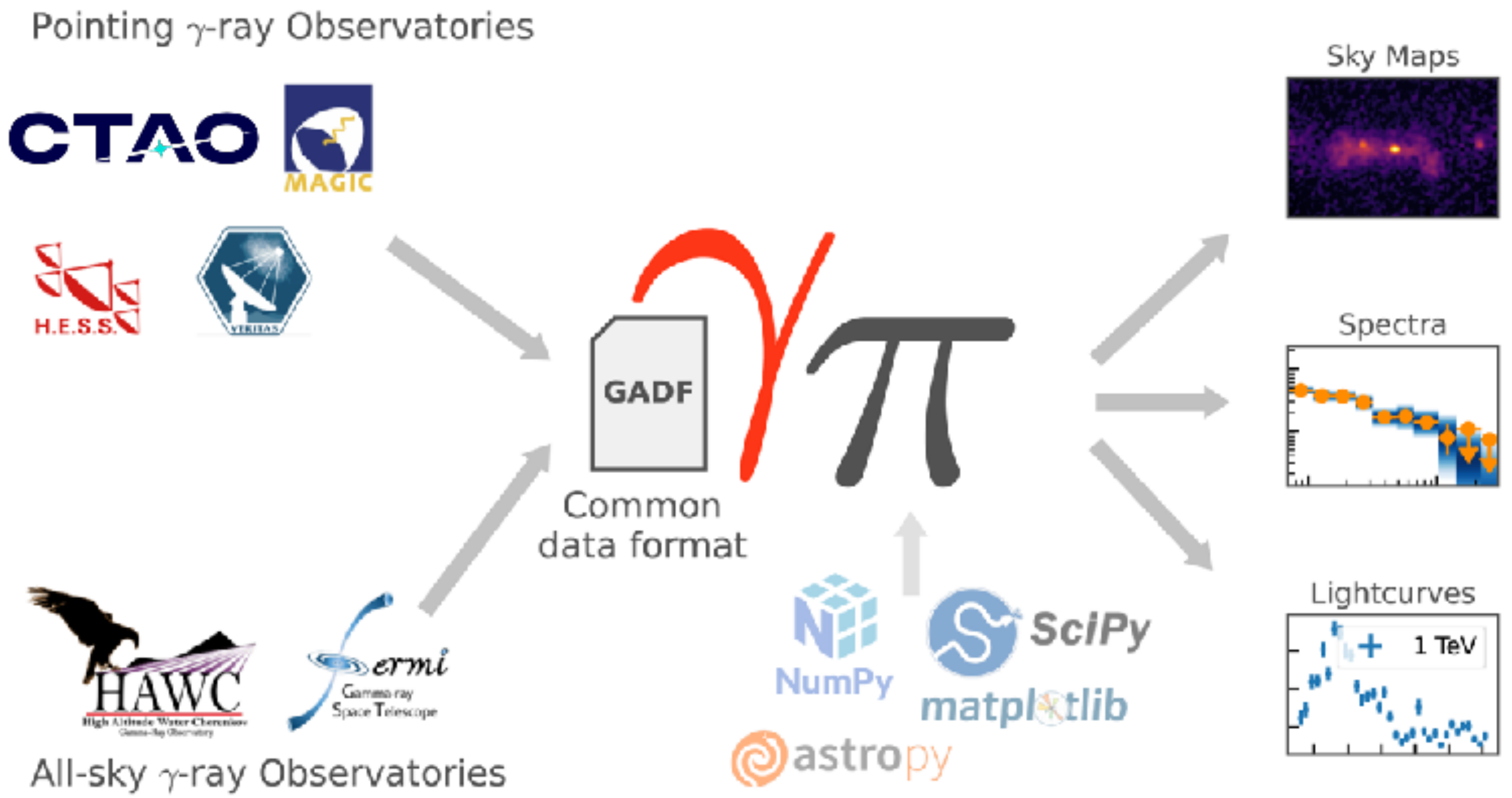
Continuous monitoring

Gamma-ray astronomy

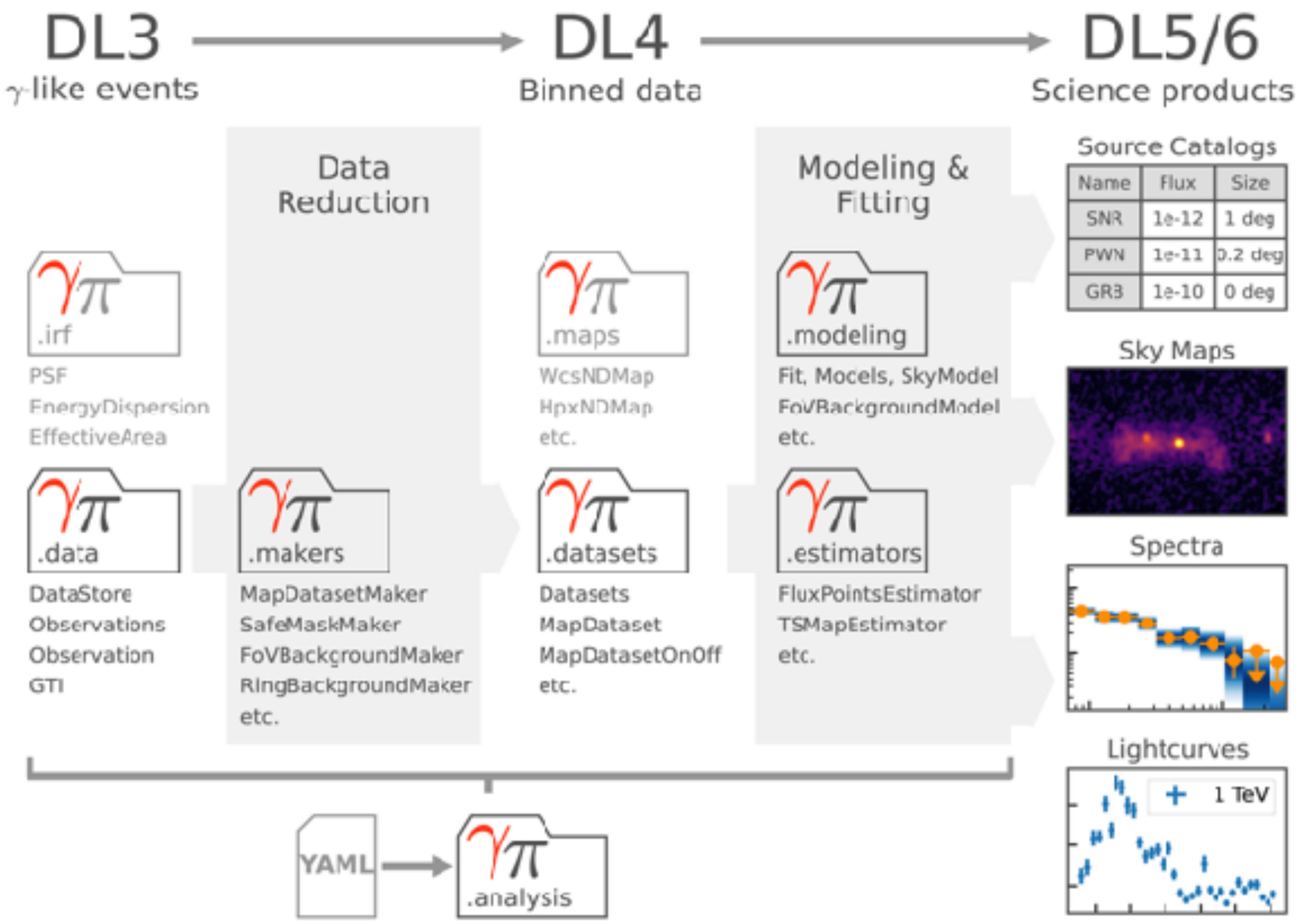


IACTs: pointed observations
Excellent angular resolution (5’)
Small field of view (3-5 degrees),
~15% uptime

Gammapy Overview



Workflow



- Analysis procedure
- Data reduction (DL3 => DL4)
- Data modeling/fitting (DL4 => DL5)

What is DL3-level data?

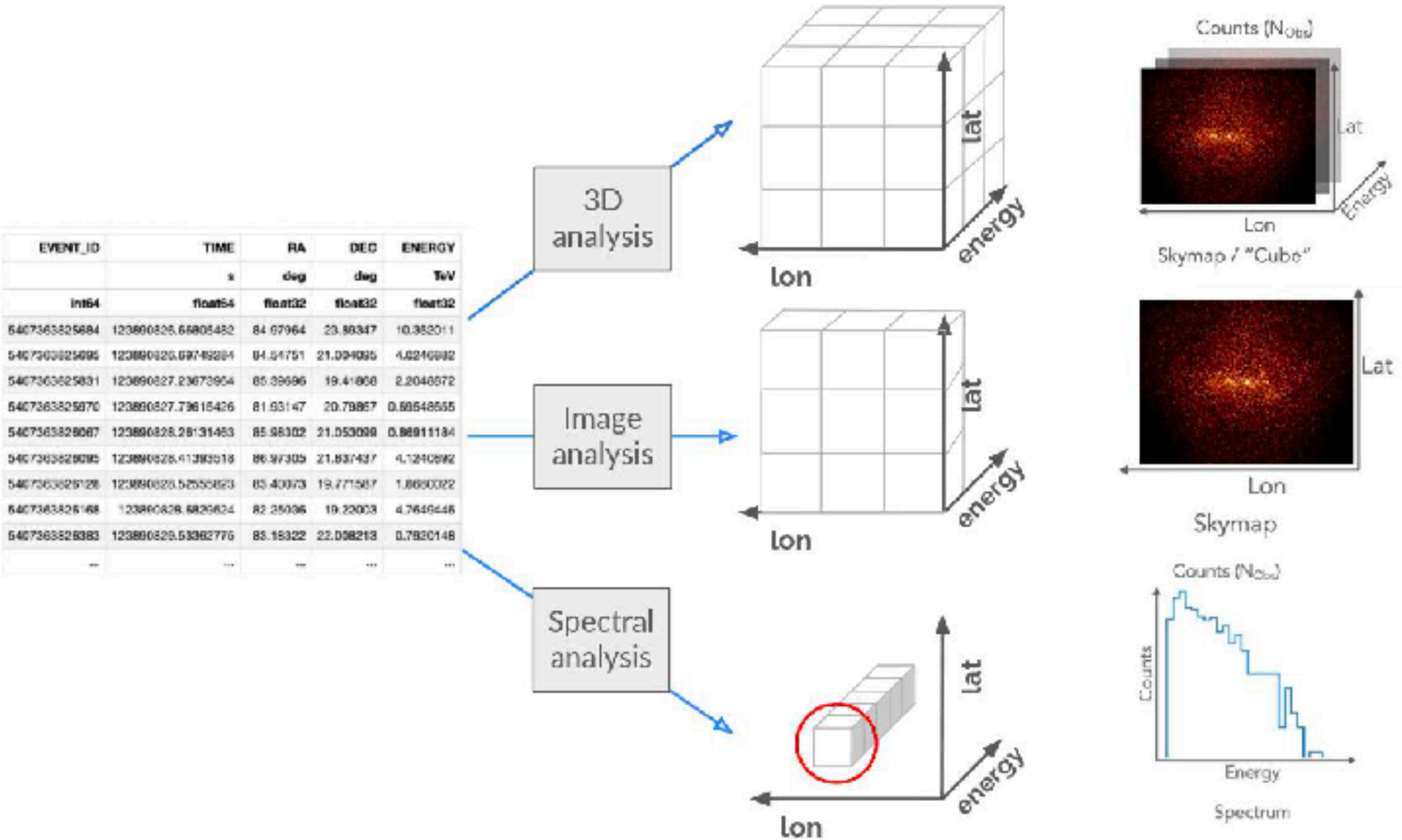
Gamma-like (≠ gamma-rays) event list

TIME	ENERGY	RA	DEC	EVENT_ID	GAMMANESS	MULTIP	ALT
	TeV	deg	deg				deg
Time	float64	float64	float64	int64	float64	int64	float64
59642.900078521834	0.4335927734207155	82.61801476356415	20.09251545574198	15	0.652481984279404 7	1	65.19032192028816
59642.90007852467	0.05659353454897665	82.16785715250593	22.08884029862256 2	16	0.8375578460068874	1	65.42677299793223
59642.90007853735	0.029425252116656653	83.37896570951846	22.900035048568753	25	0.5999593028004220 3	1	66.7589012447007 7
59642.90007853928	0.030256043208540035	82.81581867440378	21.55344982642094	27	0.4757677821120444 7	1	65.8128892116856
59642.90007854366	0.036533710107 76626	84.68662304906163	21.38098156372121 6	28	0.4880631178147 9684	1	67.46672330661772
59642.90007856315	0.043705209573342625	84.04394093457081	21.78631268651758 8	39	0.4763493287114697 3	1	66.93173458120553
59642.90007858677	0.033036188403097 53	81.44006558602877	22.22571651529234 6	56	0.4547665463110159	1	67.43438660081276
59642.90007858725	0.04765515870564858 5	83.73660961310507	20.03401018085537 8	58	0.4330335262597048	1	65.92815536214163
59642.90007860172	0.03483910048666751 6	82.94855834603 02	20.80750585103930 2	69	0.4490918831283351	1	65.80200949434864
...	...	...	...	...	...	...	...
59642.91107596089	0.07561701924159938	81.58348309597052	20.17717553407934 3	7247659	0.4083701923859809 3	1	60.76069537729194
59642.91107596885 6	0.07464419986719237	84.32870838161353	21.06176414791215 7	7247660	0.6123613628321202	1	63.47986991713417 4
59642.9110759871 8	0.11391976707942934	84.69202232534532	21.35384168530805 5	7247663	0.5964388387425119	1	63.90440938341952
59642.91107598558	0.06564583164096852	81.85970087360052	19.96917706984626 8	7247678	0.4563646214098893 3	1	60.92390944812578
59642.91107598093	0.15130850630922363	83.05875475287024	22.46531658397834 8	7247683	0.7682326096401757	1	62.88706525148241
59642.91107601831	0.04722187347629922	85.84709091939387	21.47029354740351 8	7247707	0.4631063843334452	1	64.94356676003194
59642.91107603679	0.05026761444005207	83.08938847079 87	21.96884762468167 6	7247716	0.6108013909166706	1	62.73602679204267 6
59642.91107606607	0.04170633855705597	81.79298794693246	22.76224021308432	7247733	0.4477114797760928	1	61.88056358242251
59642.91107606745	0.05711480516418008	81.65317930593478	23.01953066014323 8	7247743	0.484193931188428	1	61.8494802740055

These lists are prepared for each instrument and represent “science-ready” data or “Data level 3”

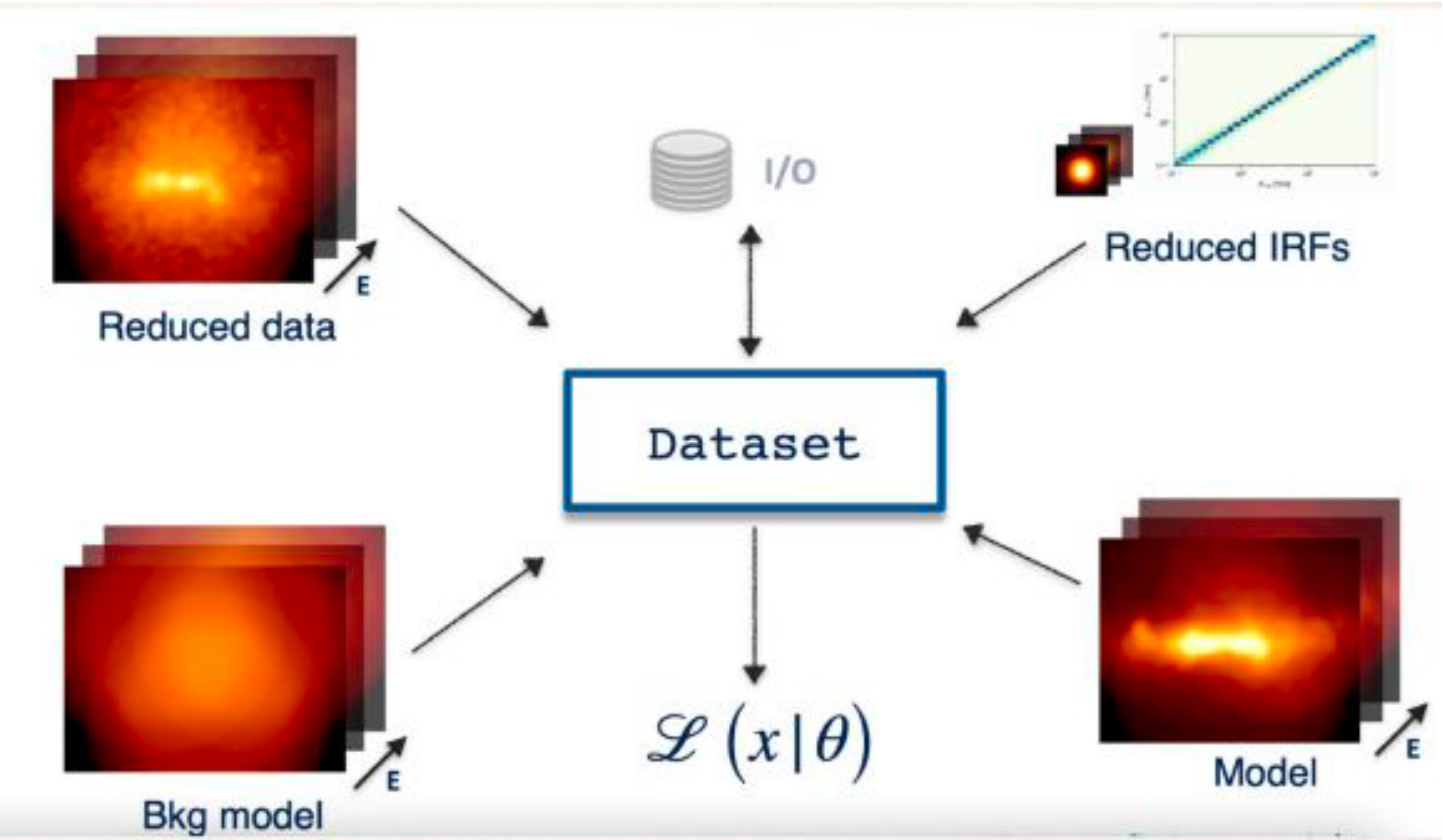
Data reduction DL3→DL4

- Select and retrieve relevant observations from the data store
- Define dataset geometry
- Loop over selected observations
- Bin events (and IRFs) into n-dim sky maps



EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
int64	float64	float32	float32	float32
5407363825684	123890826.66805482	84.67964	23.89347	10.352011
5407363825695	123890826.69749284	84.54751	21.004095	4.0246992
5407363825831	123890827.23673904	85.39696	19.41806	2.2046572
5407363825870	123890827.79615426	81.93147	20.79867	0.68548555
5407363826067	123890828.28131483	85.98302	21.053099	0.86911184
5407363826095	123890828.41393518	86.97305	21.837437	4.1240692
5407363826128	123890828.52555823	83.40073	19.771587	1.8680022
5407363826168	123890828.8829624	82.25036	19.22003	4.7649446
5407363826383	123890829.53362776	83.15322	22.098213	0.7920148
...	...	...	...	...

Data fitting DL4→DL5



- Fitting on pre-computed datasets
  - eg: From HAWC, Fermi-LAT, OGIP files, etc
  - Forward folding with maximum likelihood estimation

End products DL5→DL6

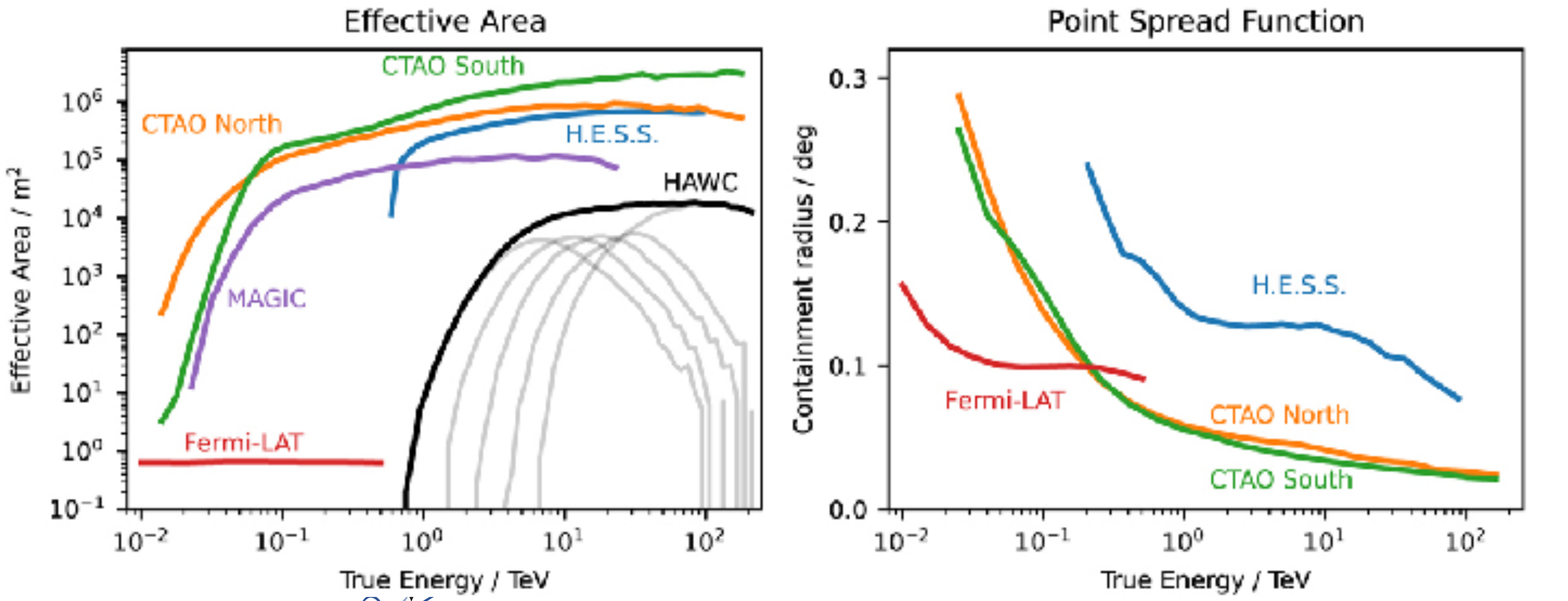
- DL5: Flux points, light curves and flux/TS maps

- Possible to fit DL5 data
  - Eg: published flux points, lightcurves
  - Chi-squared statistics used

- DL6: catalogs
  - Support provide for common catalogs: Fermi 4FGL, H.E.S.S. galactic plane survey, HAWC catalog, etc
  - Create your own catalogs…

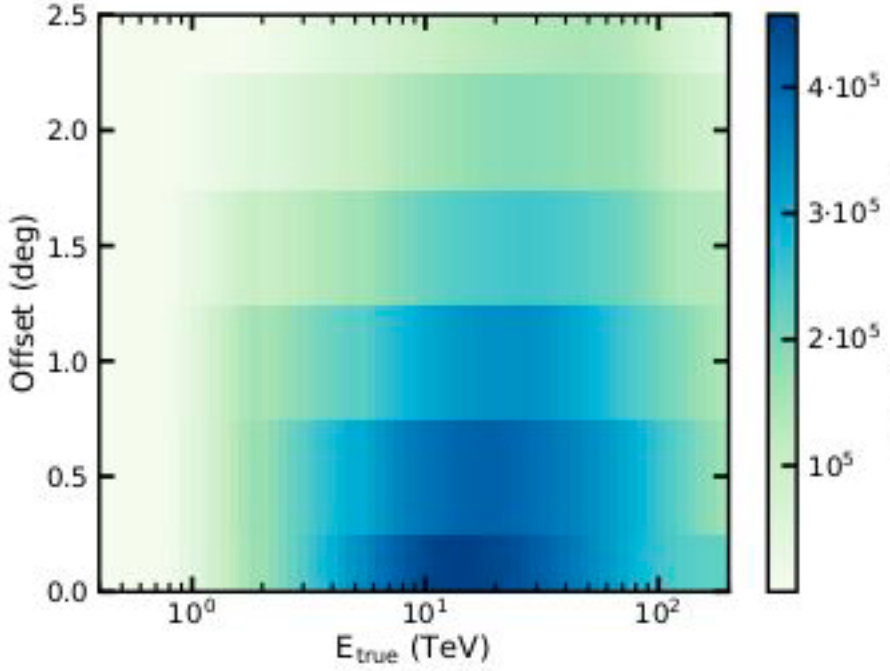
Instrument Response Functions (IRFs)

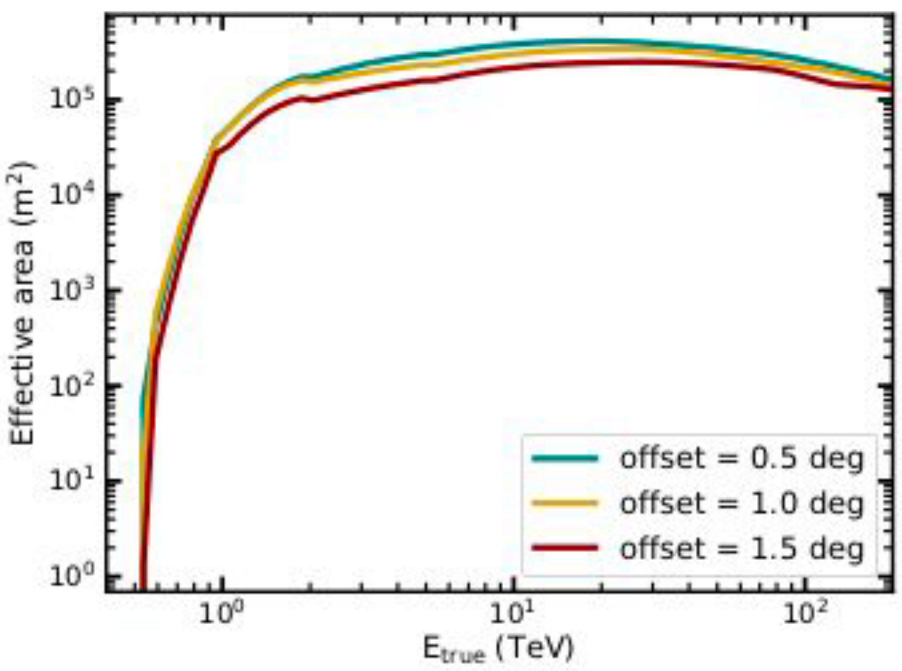
- IRFs are an important concept to understand – specific to each instrument
  - Describe the combined detection abilities and precision of an instrument data-taking and reconstruction procedure
    - Effective area, Point Spread Function (PSF), Energy dispersion, Background



Effective/collection area

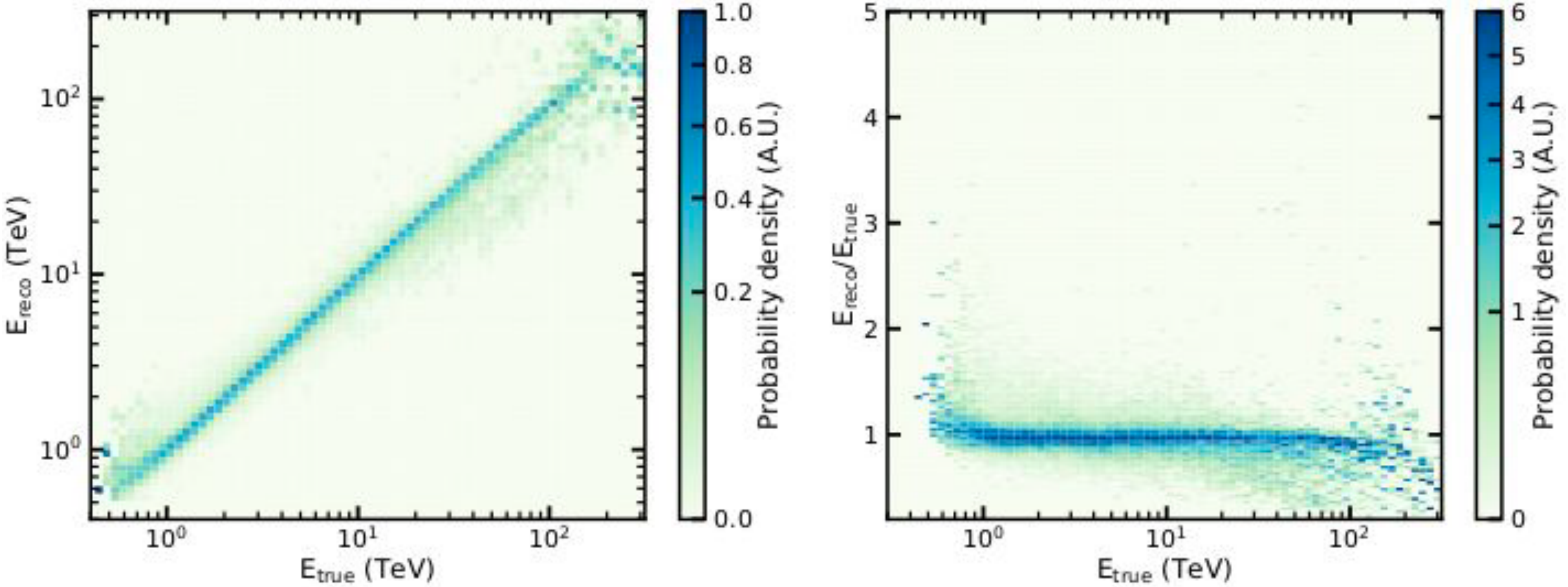
- It gives the number of expected (true) gamma-like events per unit area
- Varies with offset from pointing position and true energy
- Estimated from "Monte Carlo" simulations
- Required to measure flux/brightness of gamma-ray sources
- Typically combined with observation time to derive the effective exposure





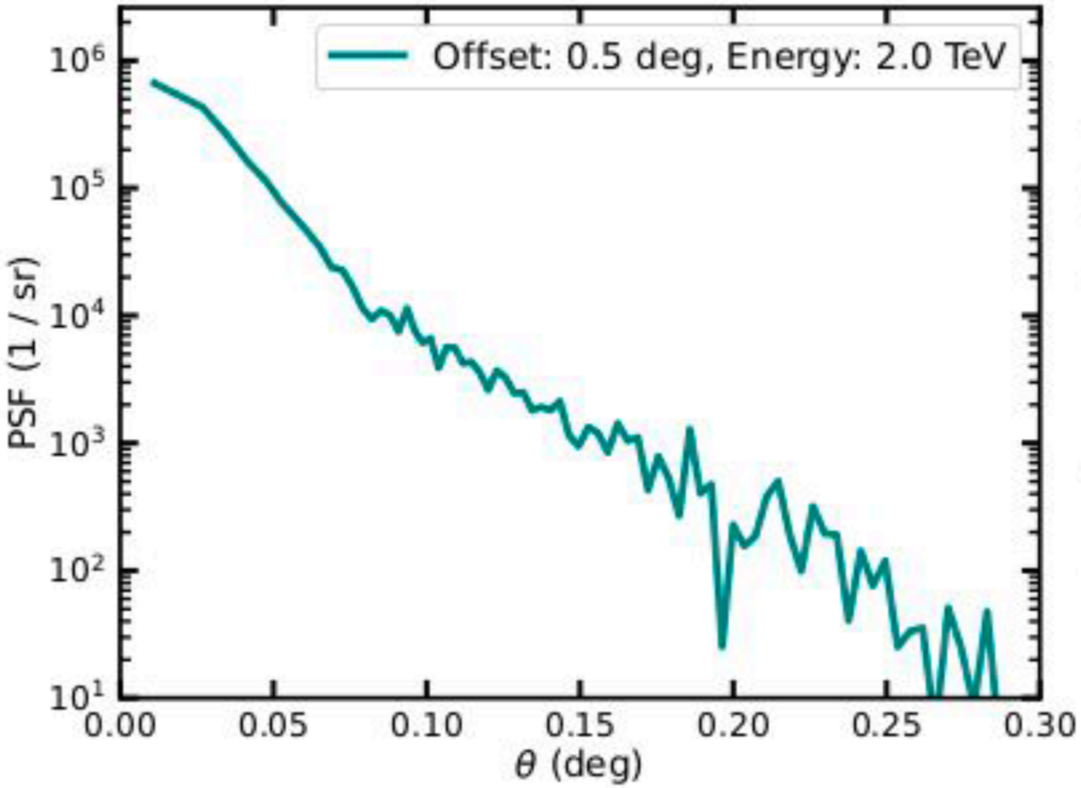
Energy dispersion/migration

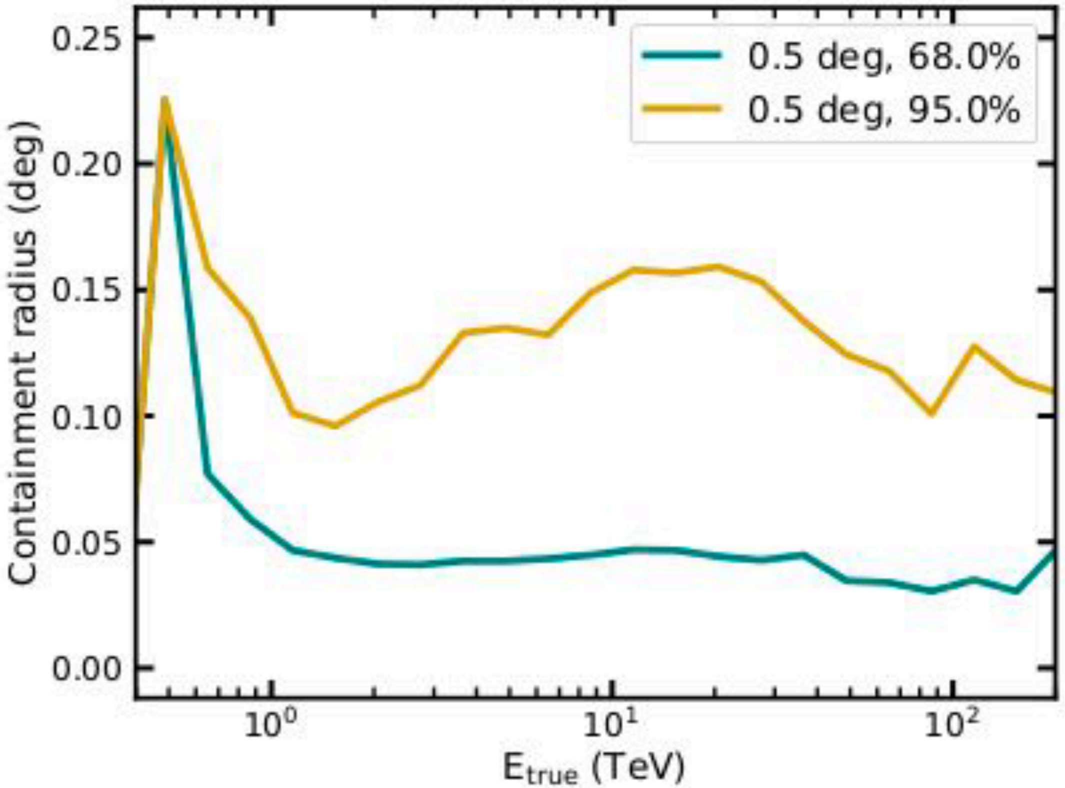
- For each true gamma-ray energy, what is the probability that the event gets assigned a certain reconstructed energy?
    - Accuracy and precision to reconstruct the energy of an event
- Varies with offset from pointing position and true energy



Point Spread Function

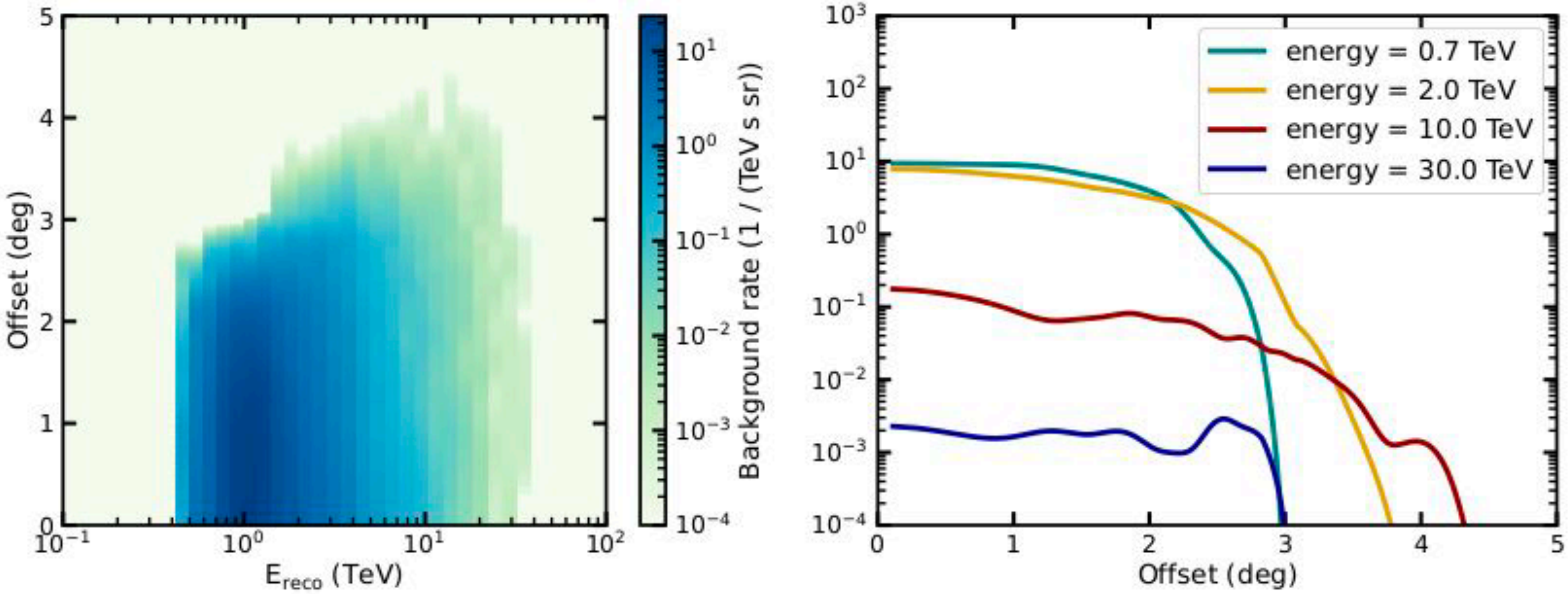
- PSF: angular resolution of the instrument, precision to reconstruct the arrival direction of an event
- Estimated from "Monte Carlo" simulations and by binning events into offset and true energy





Background

- Represents the expected remaining hadronic background after gamma-hadron separation due to misclassified events
- Depends on the reconstructed energy and the offset from the pointing position



Data analysis

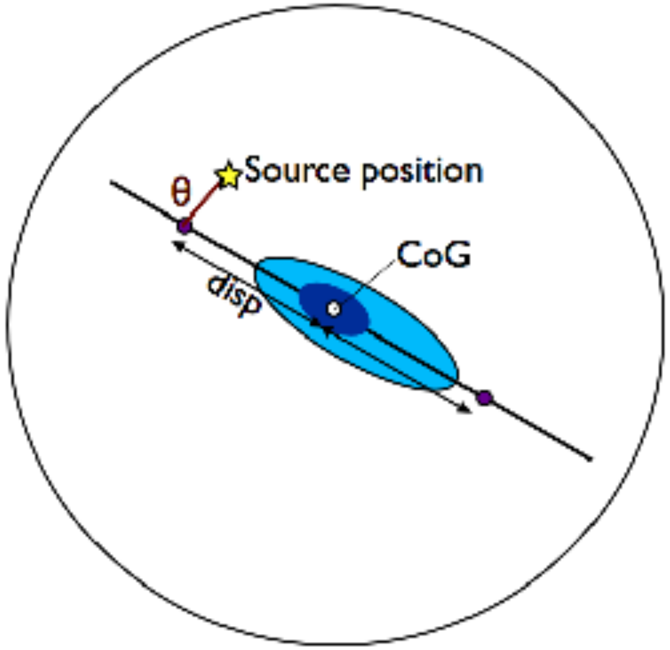
Source detection

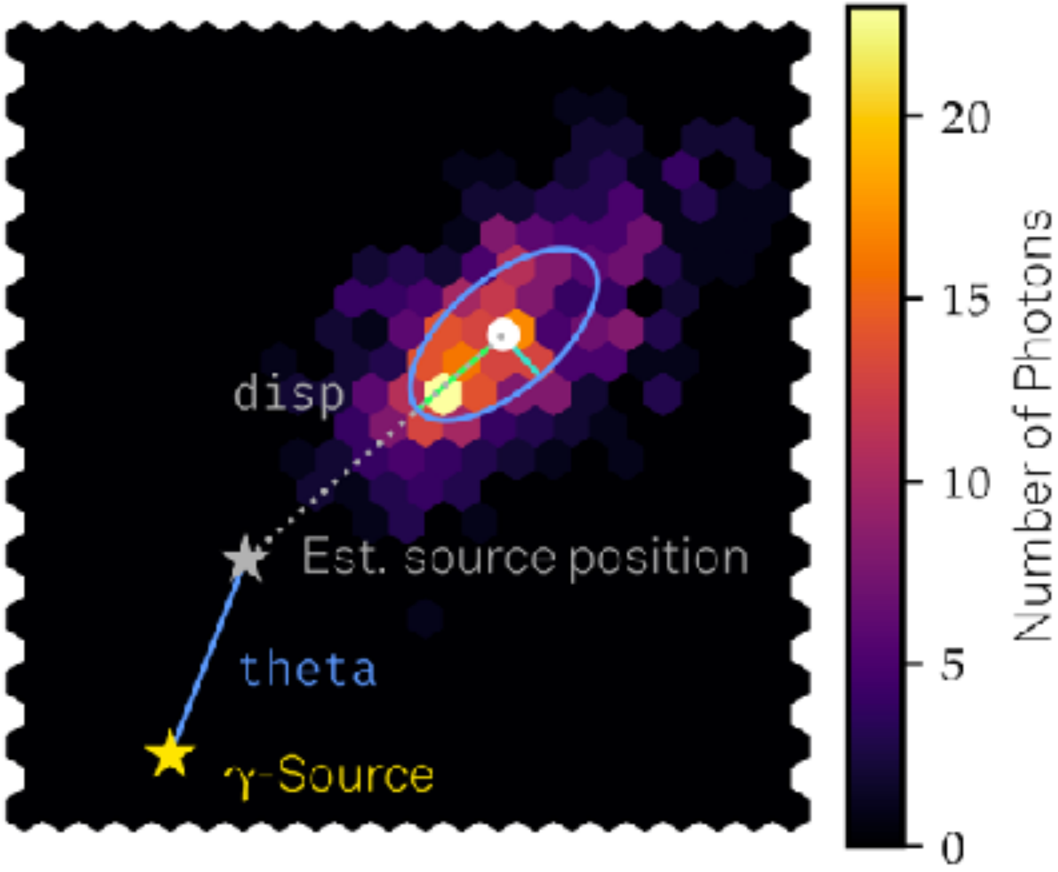
- Likelihood ratio test: compare the likelihood of two hypotheses to see which one is supported by the data
- The two unknown parameters:
  - Expected number of source photons: $< N_S >$
  - Expected number of background events: $< N_B >$
- A model or excess of counts (H1) is tested against the null hypothesis (H0) where no source is present
  - Null hypothesis: $< N_S >= 0 \rightarrow \mathcal{L}_0 = \mathcal{L}(H_0)$
  - Alternative hypothesis: $< N_S > \neq 0 \rightarrow \mathcal{L}_1 = \mathcal{L}(H_1)$
- Use difference in test statistic *TS) through the likelihood ratio:

$$TS = -2 \log \left(\frac{\mathcal{L}_0}{\mathcal{L}_1} \right)$$

Source detection: Theta^2

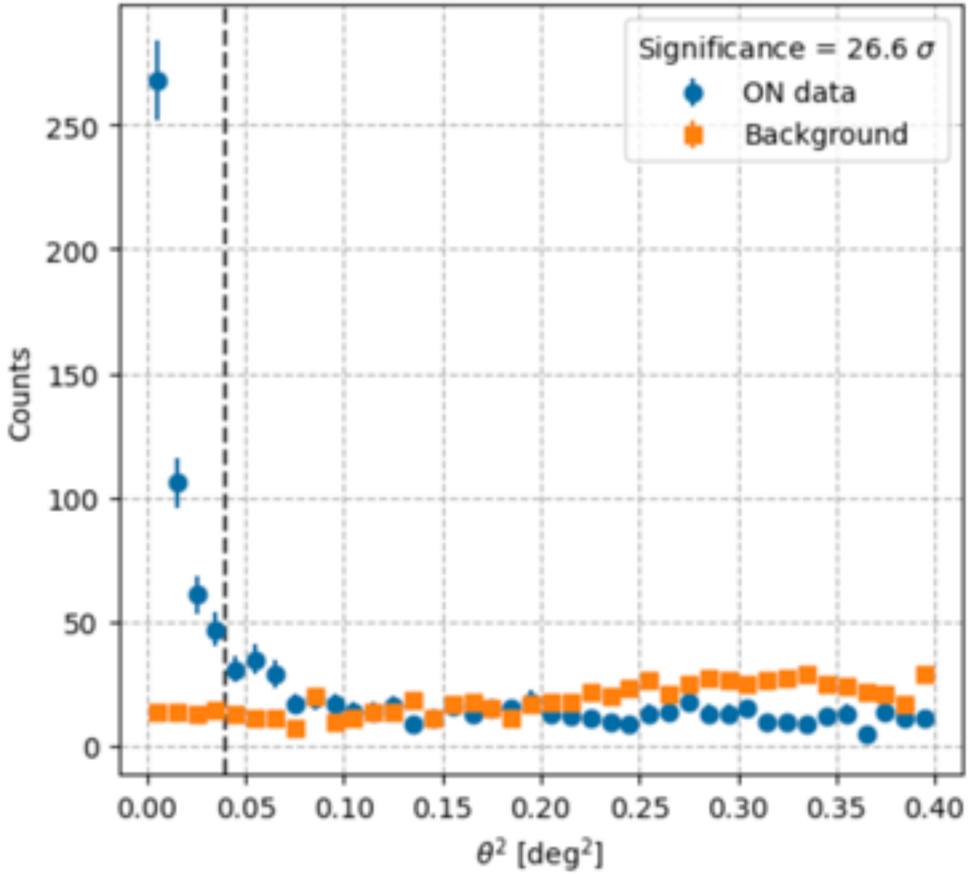
- Squared angular distance between the reconstructed event position and the source position

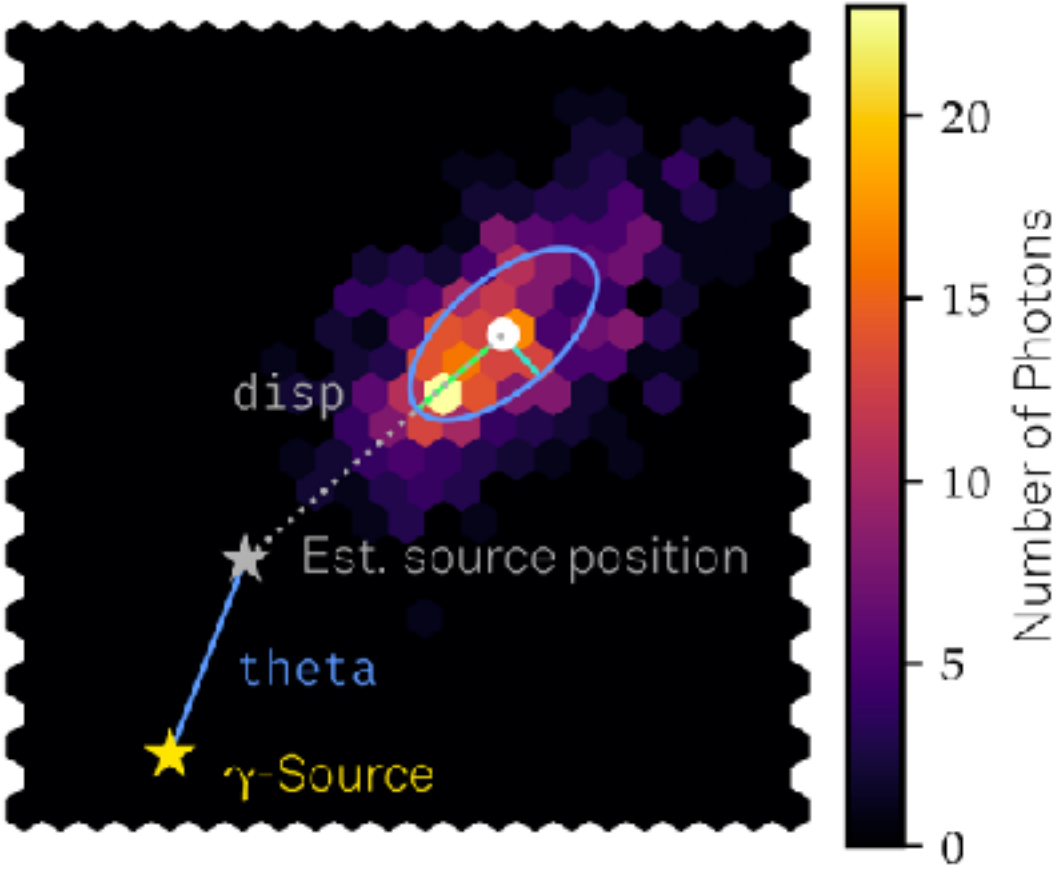




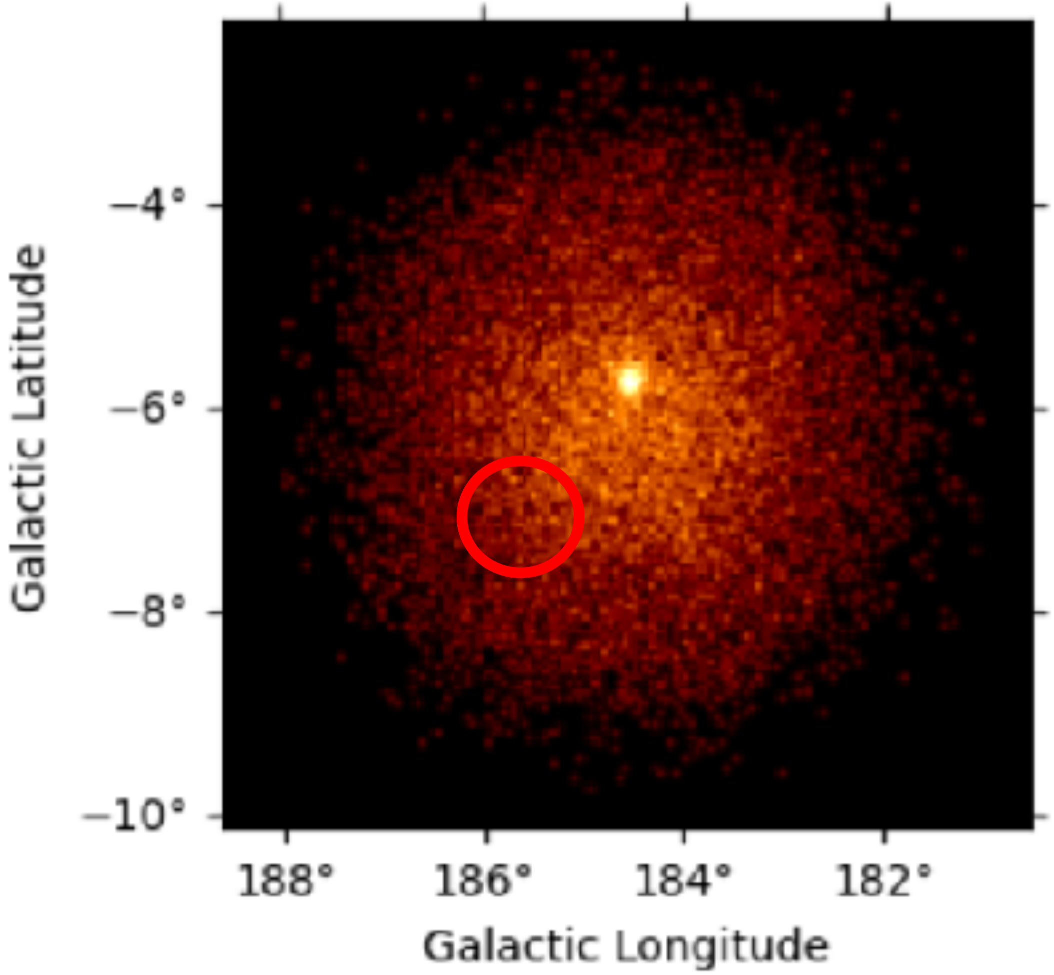
Source detection: Theta^2

- Squared angular distance between the reconstructed event position and the source position

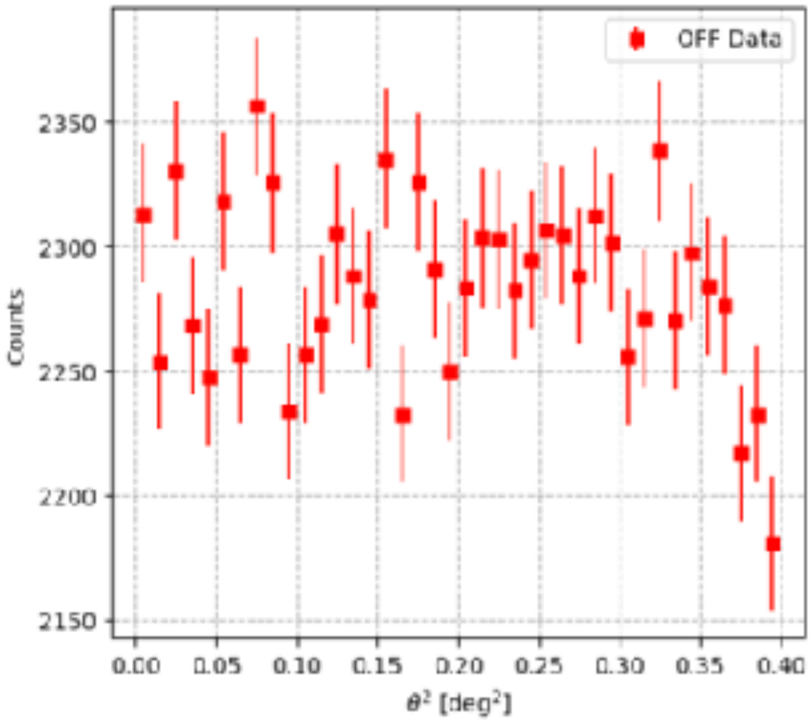




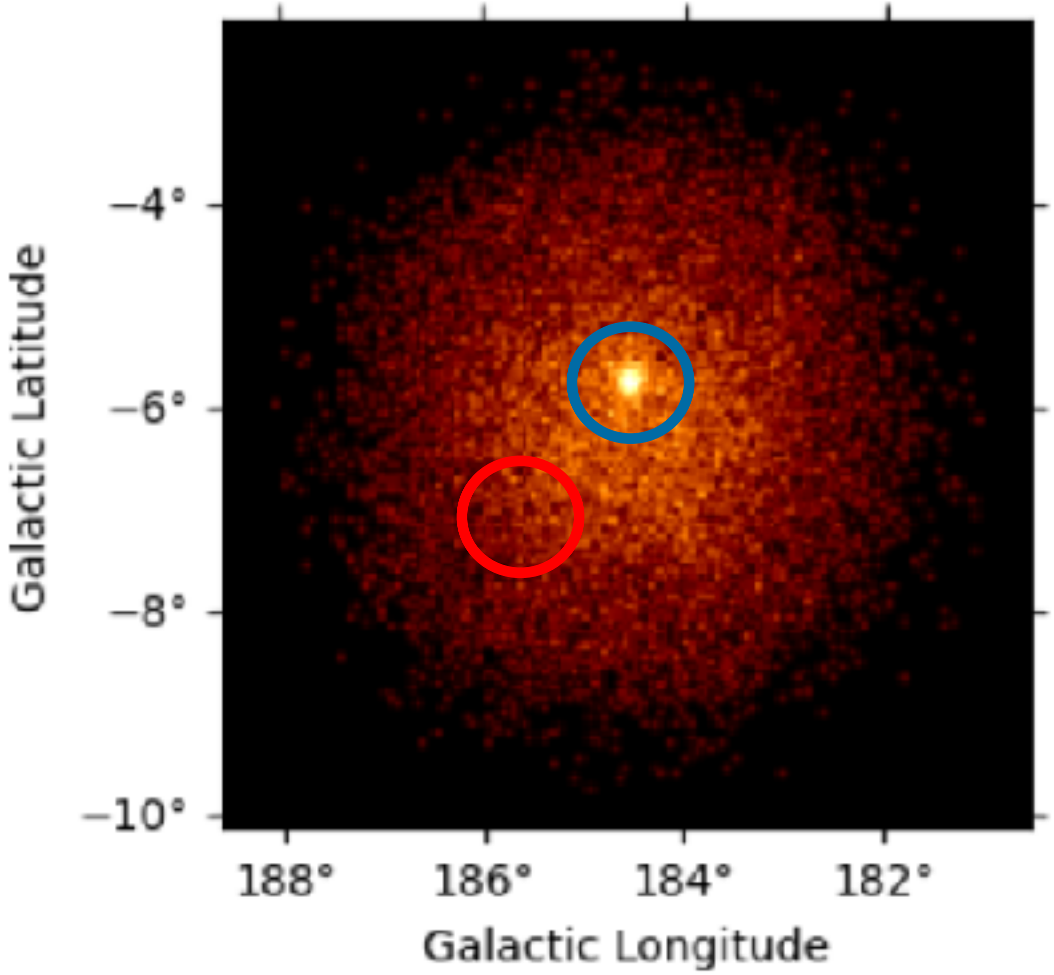
Source detection



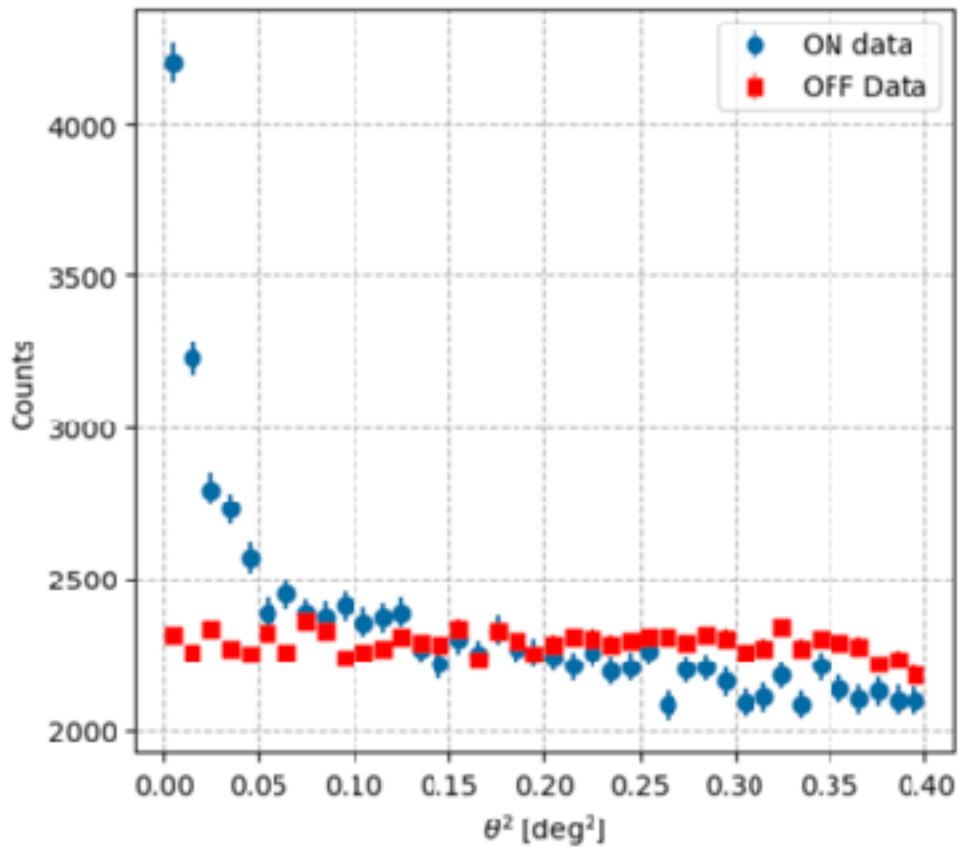
Background region



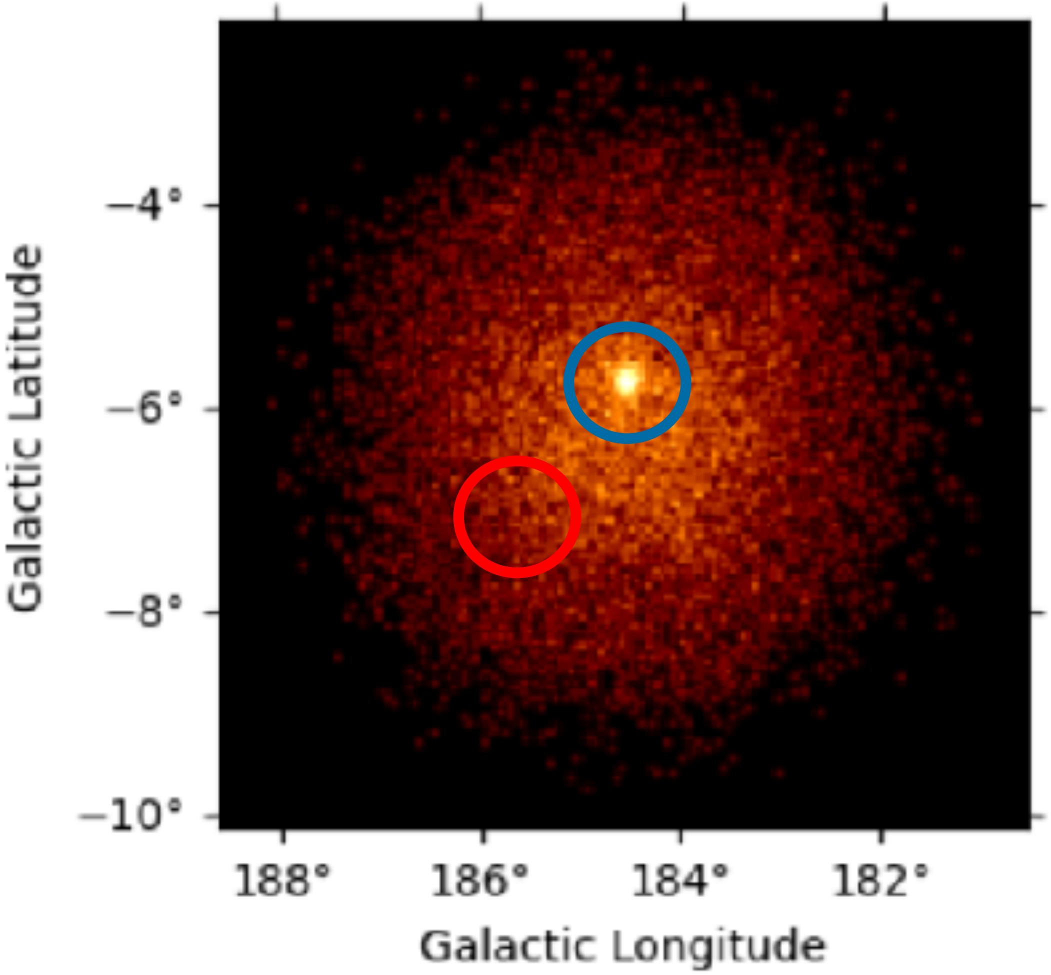
Source detection



Background region
Signal region



Source detection



Background region

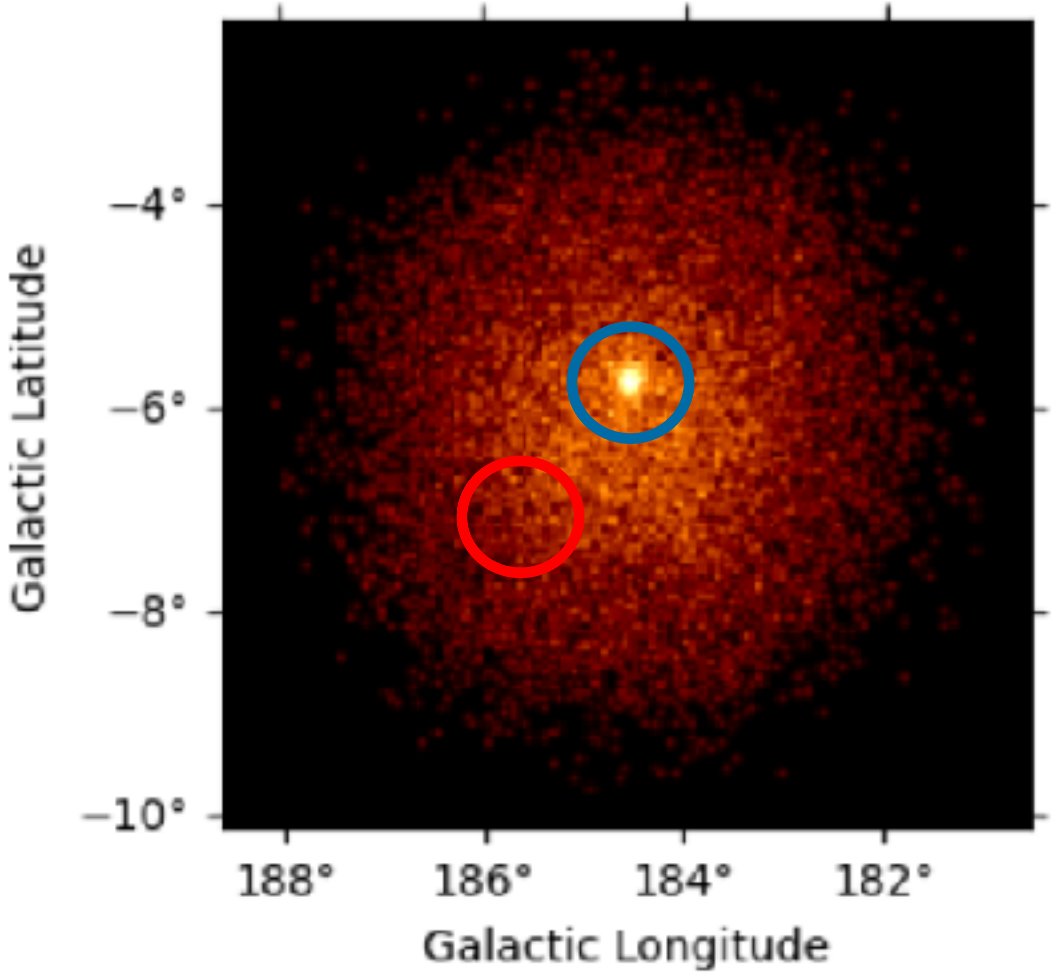
Signal region

Significance:
Likelihood ratio method Li&Ma (1983)

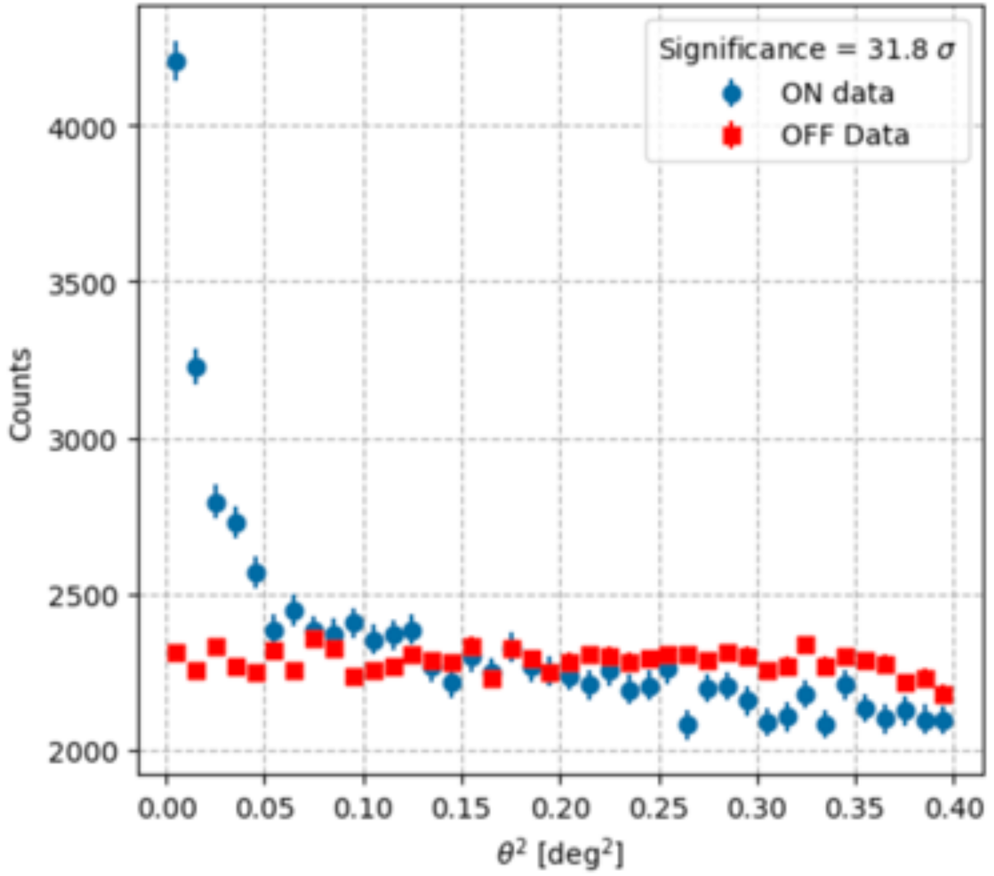
$$S = \sqrt{-2\ln\lambda}$$

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

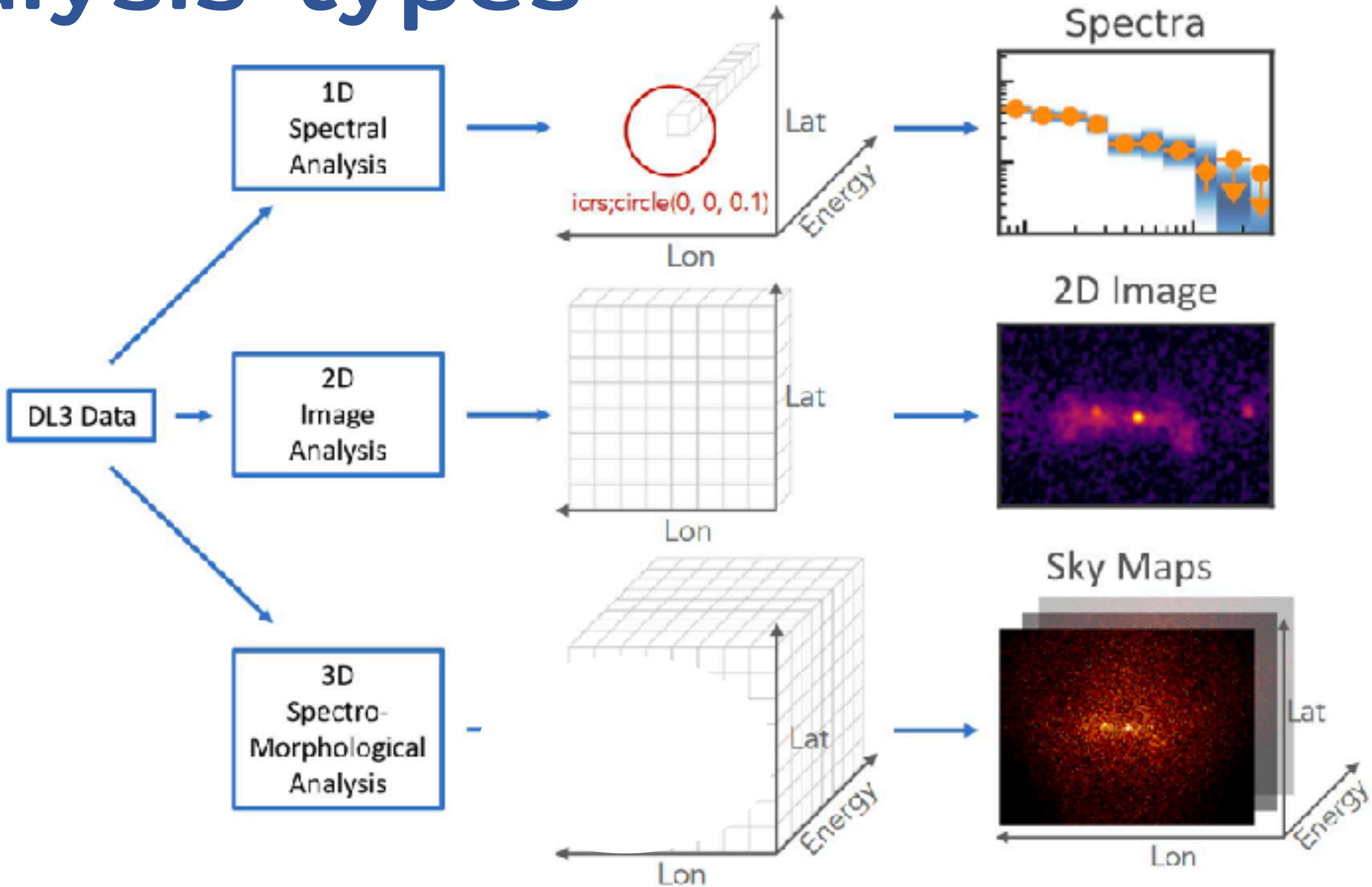
Source detection



Background region
Signal region

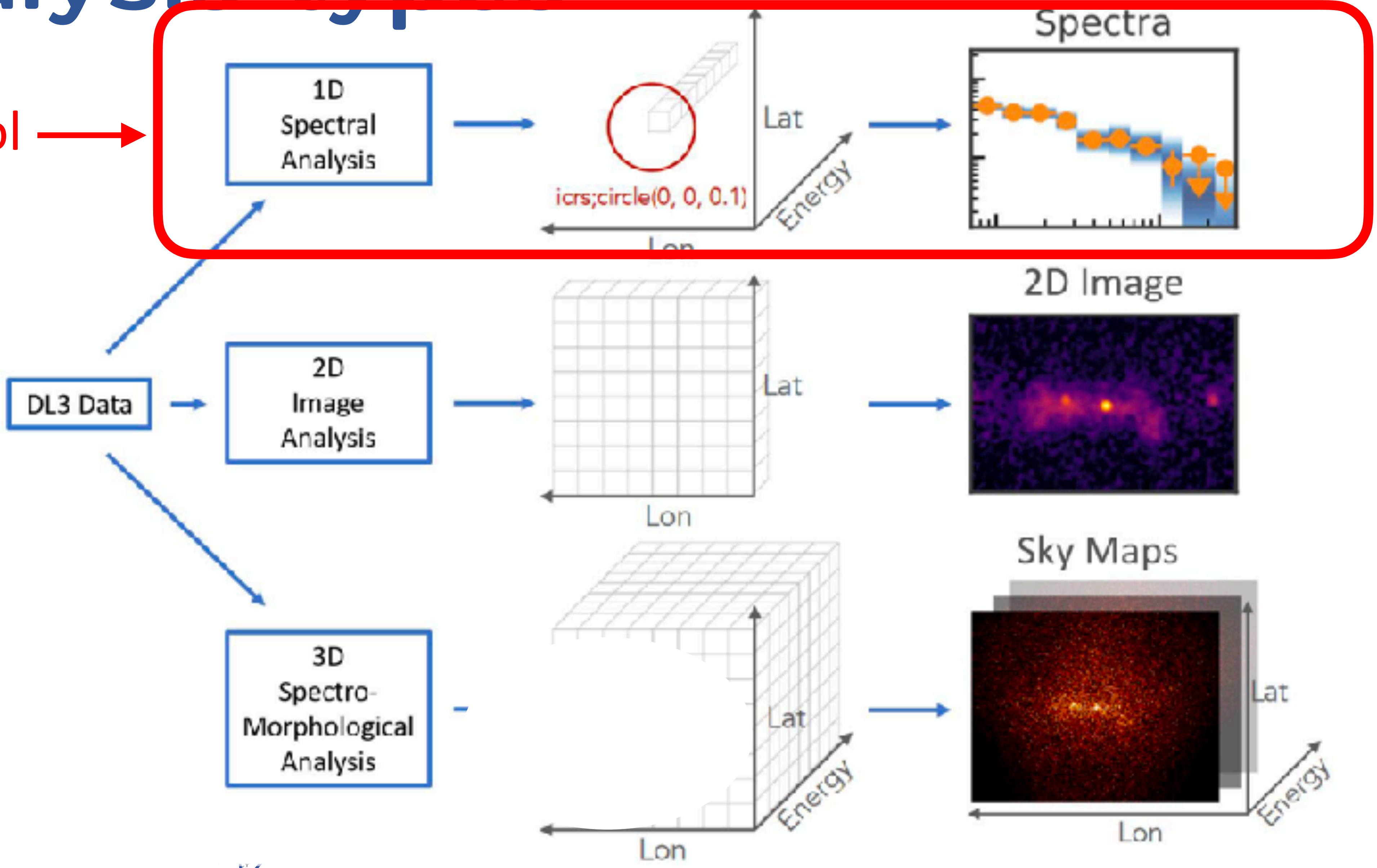


Analysis types



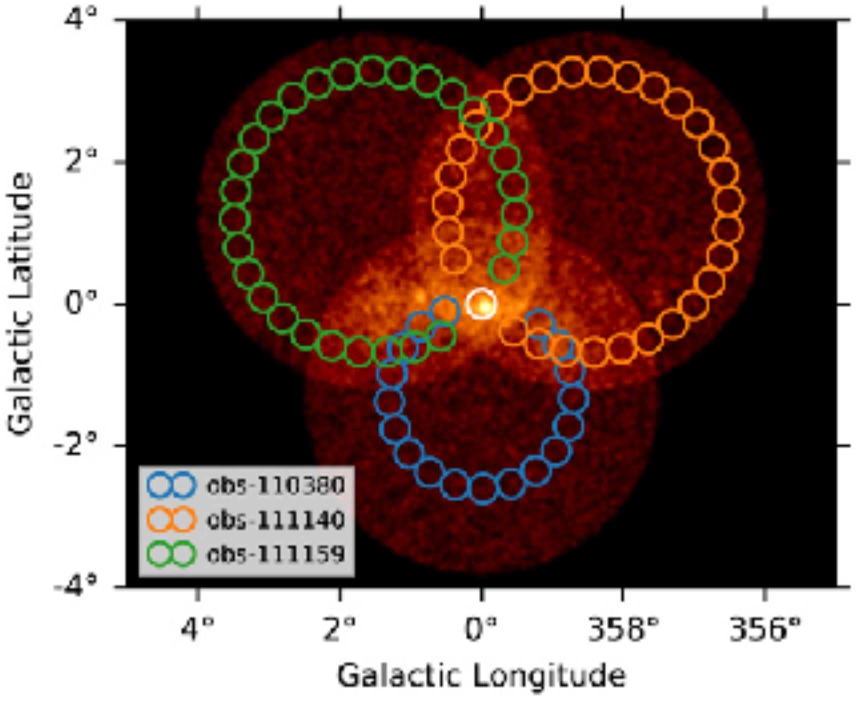
Analysis types

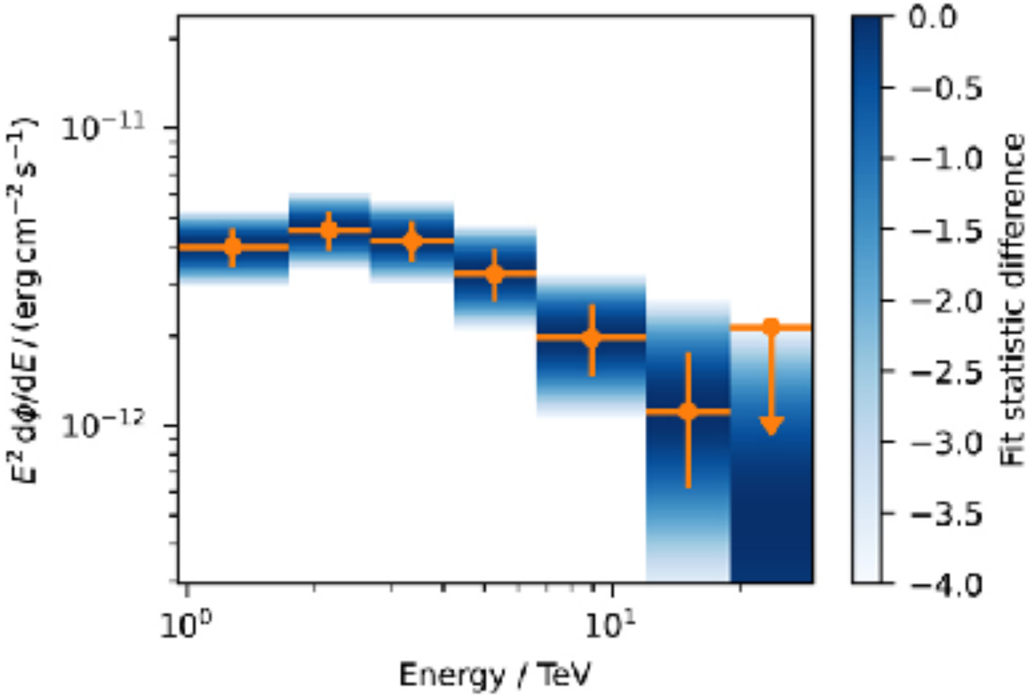
In this school



1D Spectral Analysis

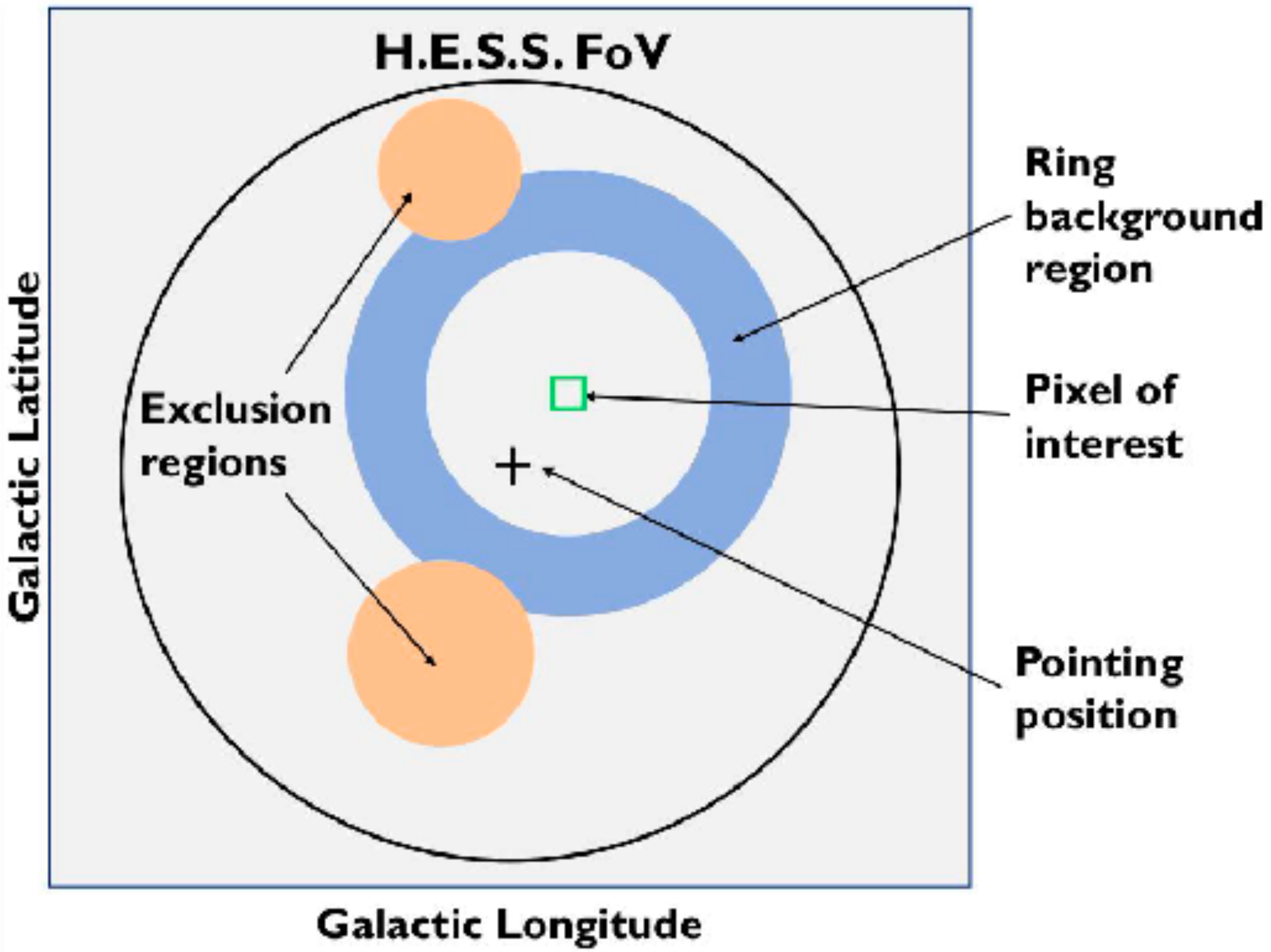
- Use the reflected regions method
- Along with the forward folding method → maximum likelihood
- Use simulated CTAO data – likelihood per energy bin is shown by the blue colored band





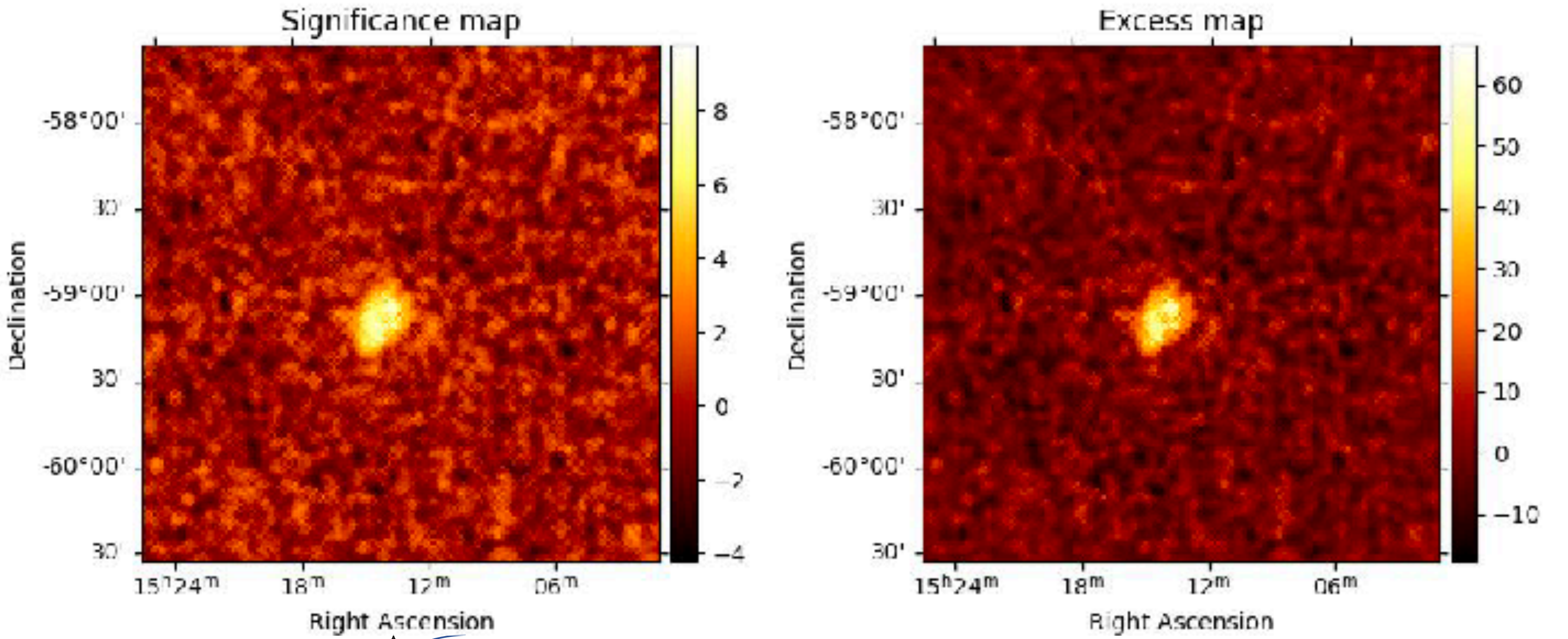
2D Spectral Analysis

- Ring background method:
- Take a ring region for each pixel in the FoV where counts inside the ring estimate the background
- As this is performed for every point within the FoV it can be utilised to create a map
- Note: this relies on a background model



2D Spectral Analysis

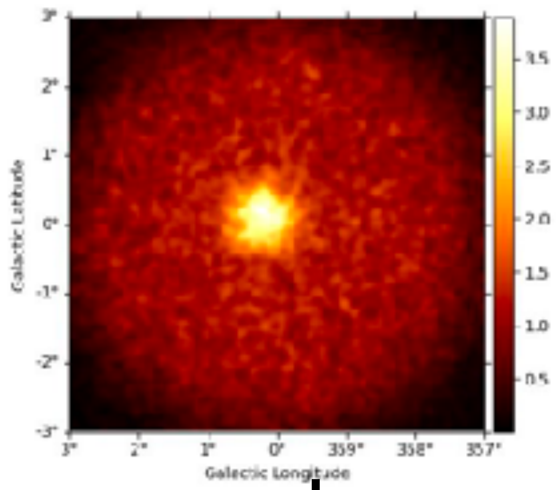
- “Ring background” method
- Select OFF events from an annulus about the ON region
- Requires good knowledge of 2D acceptance!



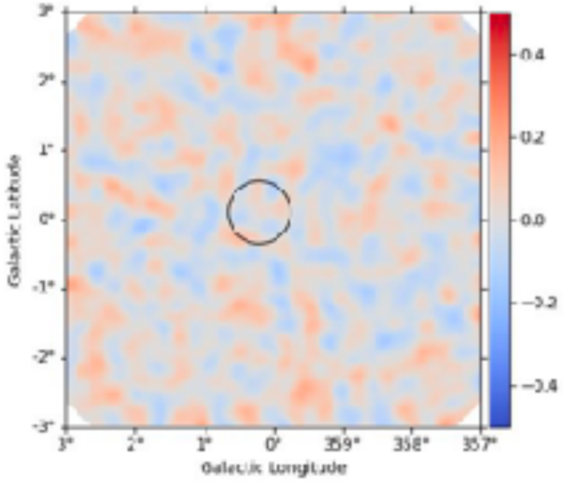
3D analysis technique

- FoV background method:
  - Advanced method to estimate the background is through a 3D model
- Construct a model of acceptance, and then predict the background over the field of view (FoV) as a function of energy
- Runwise FoV background is adjusted to the counts measured outside the exclusion regions, implement a specific spectra model
  - Corrects for effects that are not taken into account during the FoV background model construction
- Use the FoV background method which allows for a 3D analysis
- In one analysis chain we can fit both the spectra and the morphology of a source

Initial model guess

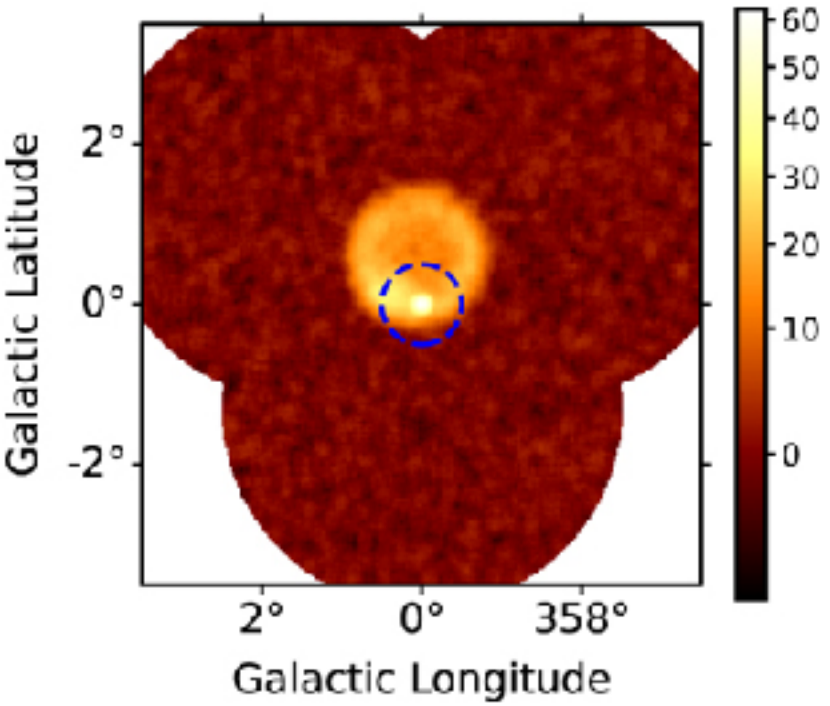


Residuals

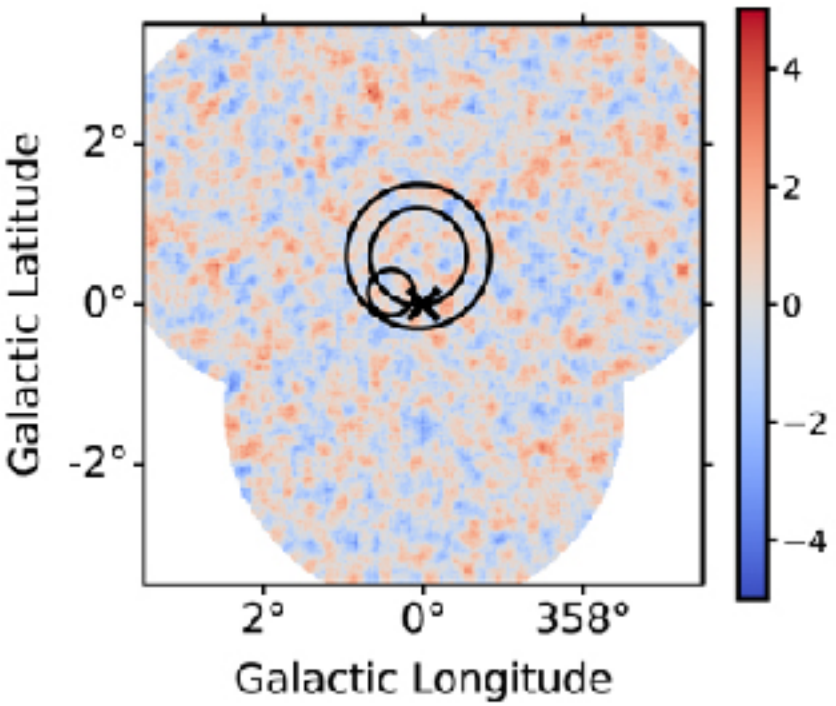


3D analysis technique

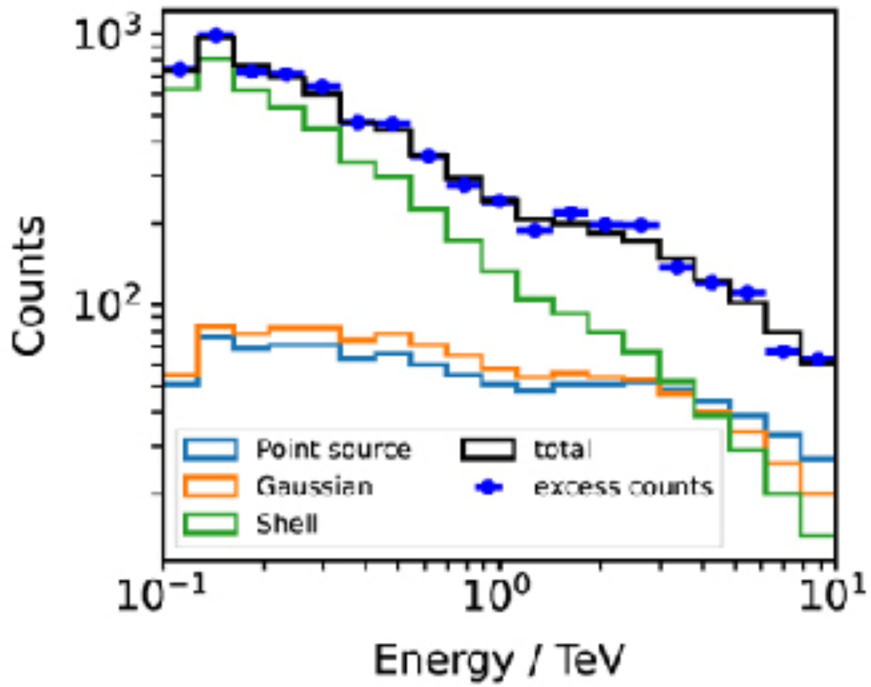
- Disentangling the contribution from overlapping sources



Significance map



Residuals



Contribution of each source model to the signal region

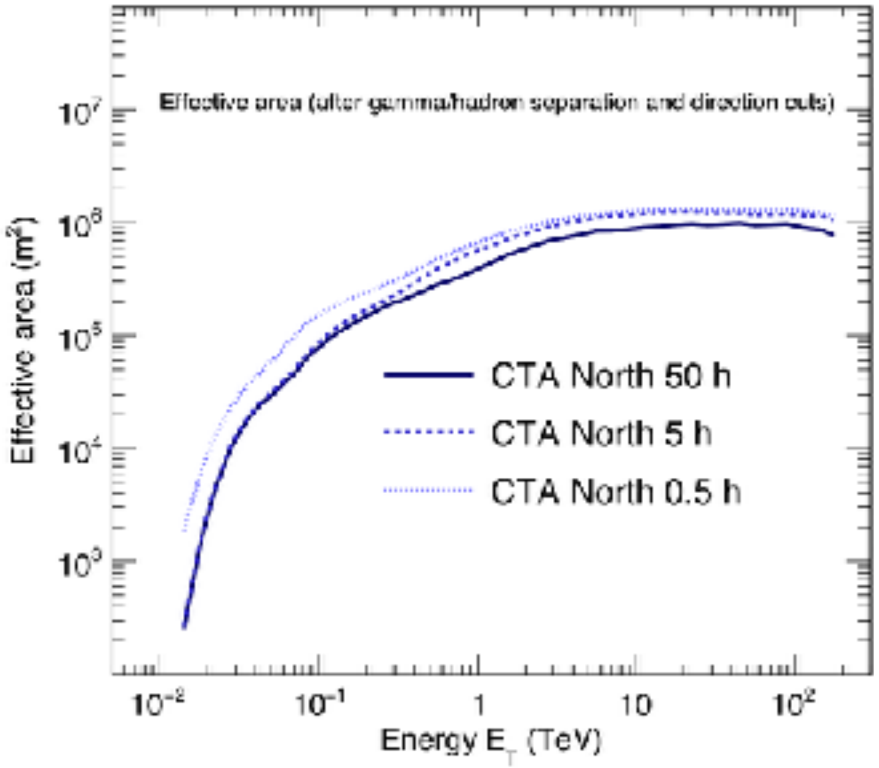
HANDS ON!

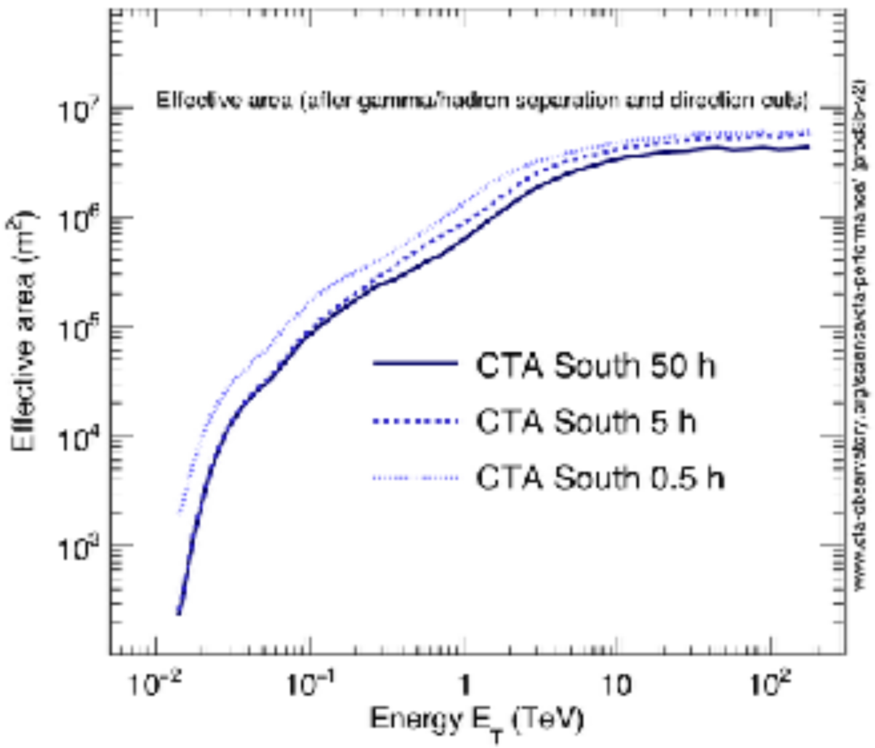
- Perform a 1D spectral analysis of the RX J1713.7-3946
- Get your hands dirty: Real Data!
  - Data analysis of DL3 observations of the Crab Nebula with the LST-1 of CTAO

Backup

Collection area

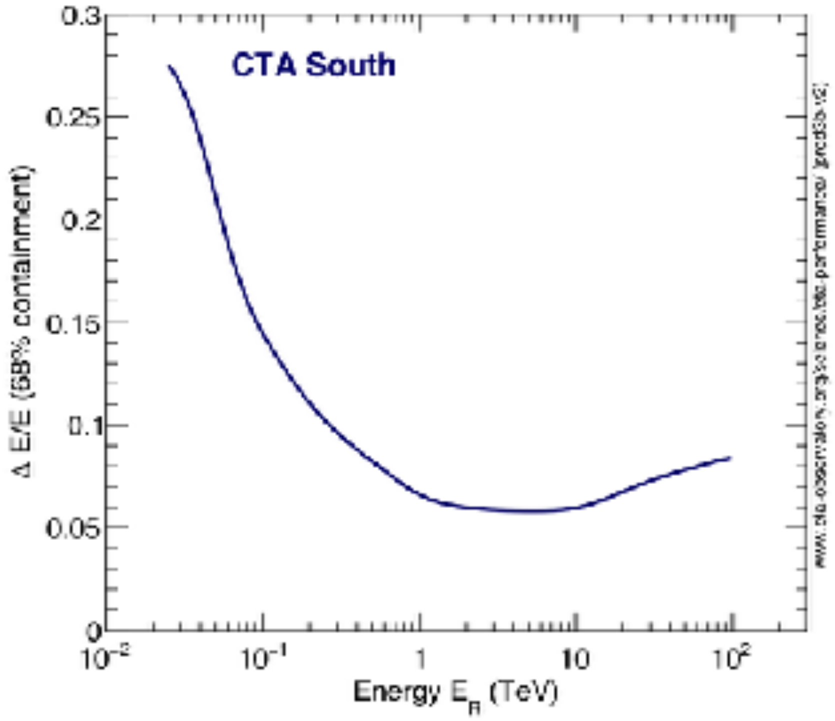
- It indicates up to which distance your array effectively work
- Calculated using the number of simulated events, those that survive the reconstruction cuts and the simulated area

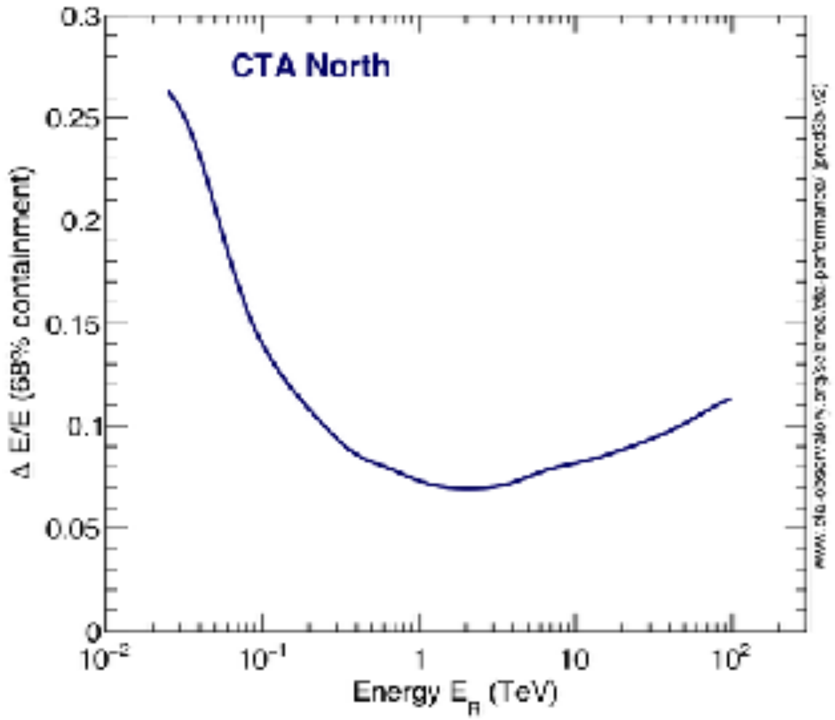




Energy resolution

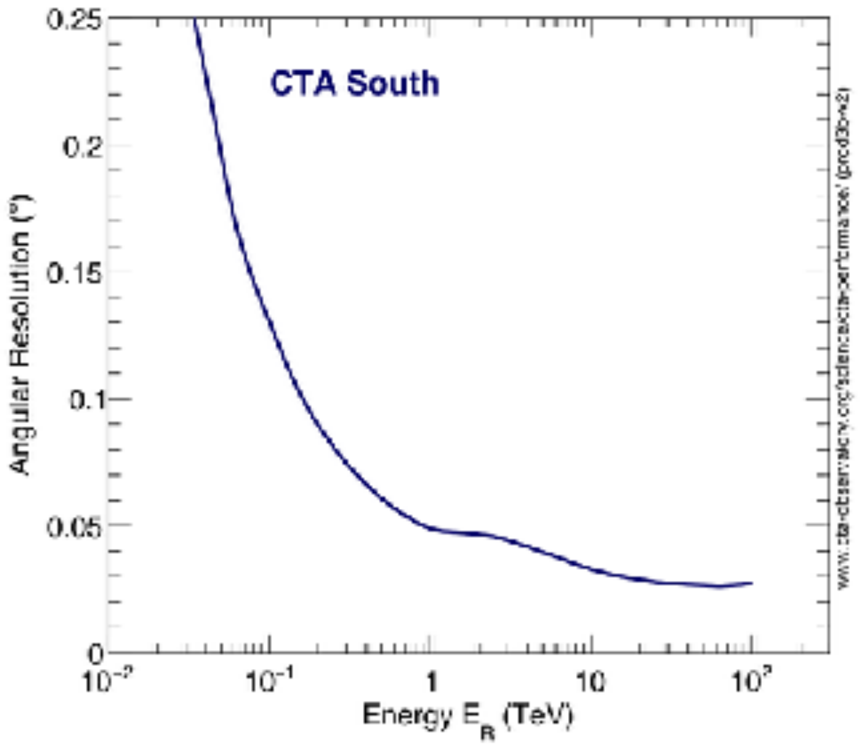
- Defined as the sigma of the gaussian of the relative error between the reconstructed and the true energy
- A lower energy resolution implies a better spectral reconstruction

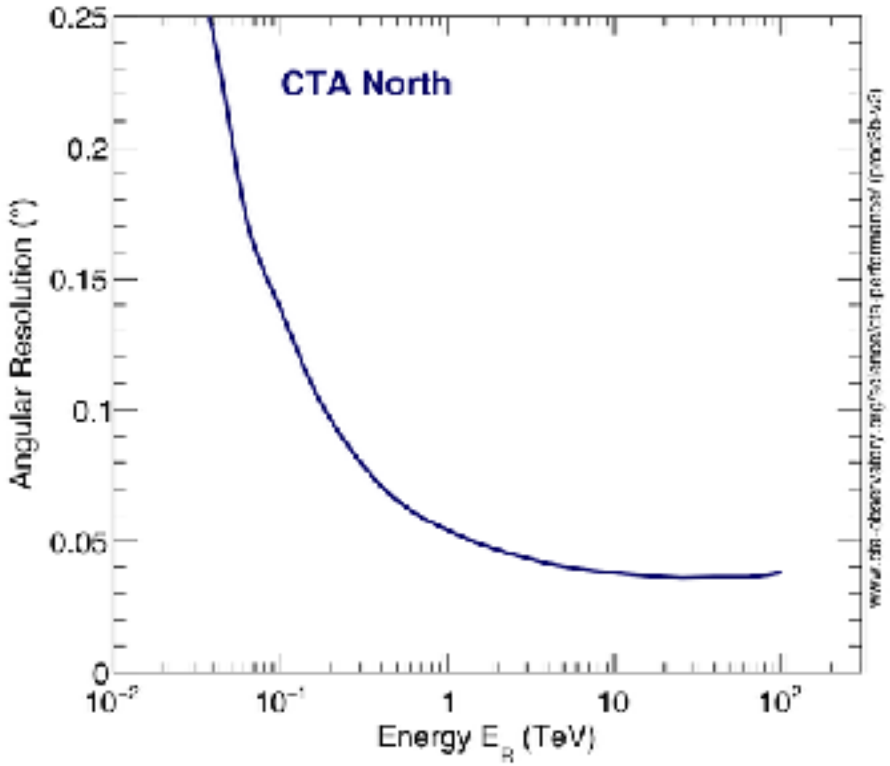




Angular resolution

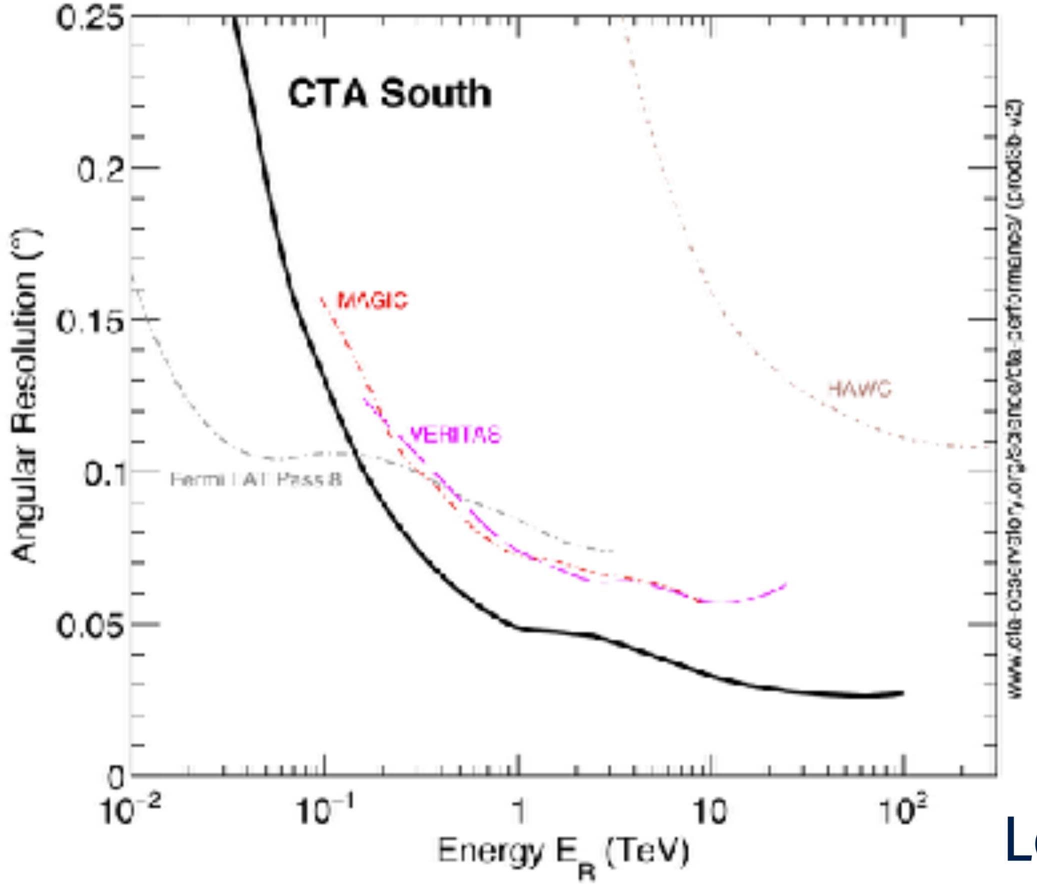
- Usually defined as the 68% containment radius of well reconstructed events in an energy bin
- It evaluates the ability of your instrument to resolve a source





Angular resolution comparison

- Significant improvement with respect to the currently existing facilities
- CTA will be able to resolve most of the known galactic sources



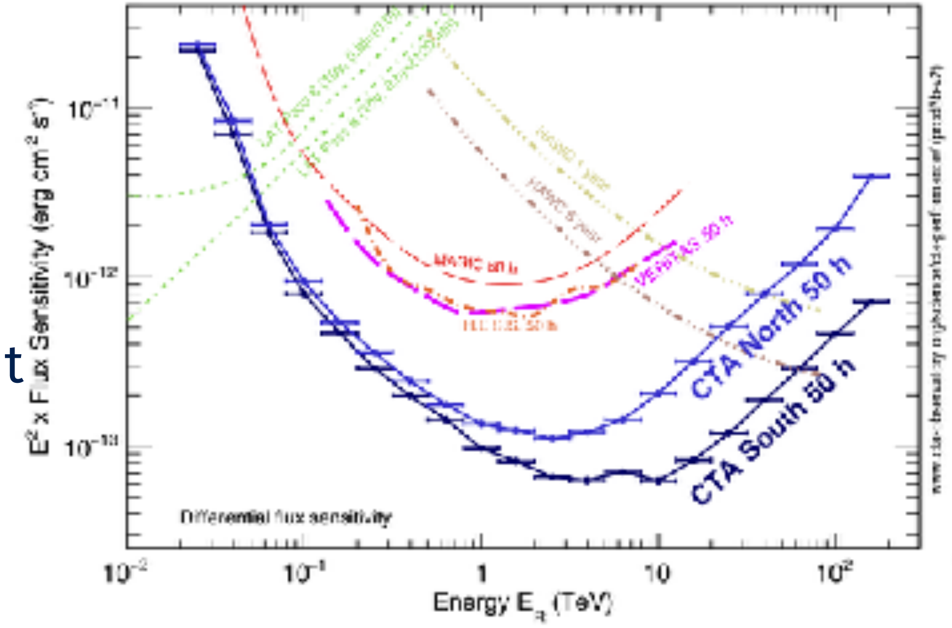
Lower is better

Differential Sensitivity

- One of the most important figures of merit that you will see

- It represents the lower differential flux an instrument is able to measure with 5σ significance

- The improvement of CTA with respect to the currently existing facilities is around one order of magnitude



Lower is better