Analysis methods in gamma-ray astronomy

ORP School on Multi-messenger Astrophysics

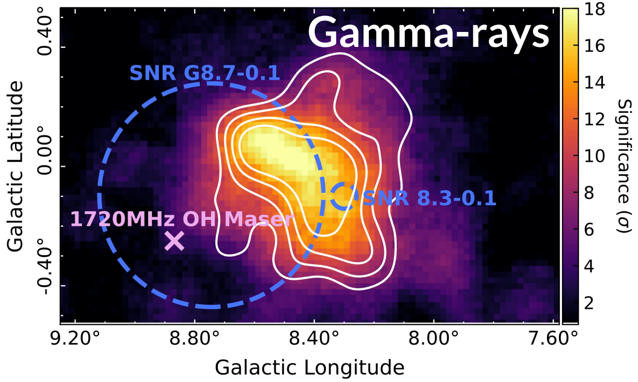
Durham University, 2nd - 6th September 2024

Kirsty Feijen (APC)

https://github.com/Astro-Kirsty/gammapy-ORP

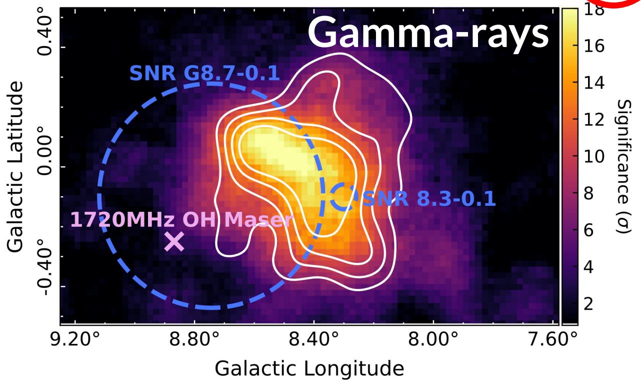
Who am I?

🐨 **Previously:** PhD from the University of Adelaide with Sabrina Einecke and Gavin Rowell



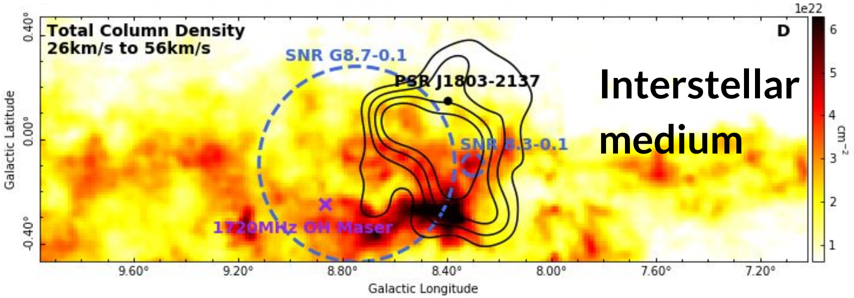
Who am I?

- 🐨 **Previously:** PhD from the University of Adelaide with Sabrina Einecke and Gavin Rowell



Who am I?

🐨 **Previously:** PhD from the University of Adelaide with Sabrina Einecke and Gavin Rowell



Who am I?

- **Previously:** PhD from the University of Adelaide with Sabrina Einecke and Gavin Rowell
  Currently: Post-doc at APC (Astroparticule et Cosmologie) lab in Paris, France

A **Python** package for **gamma-ray** astronomy

CTAO

H.E.S.S.

Research: private!

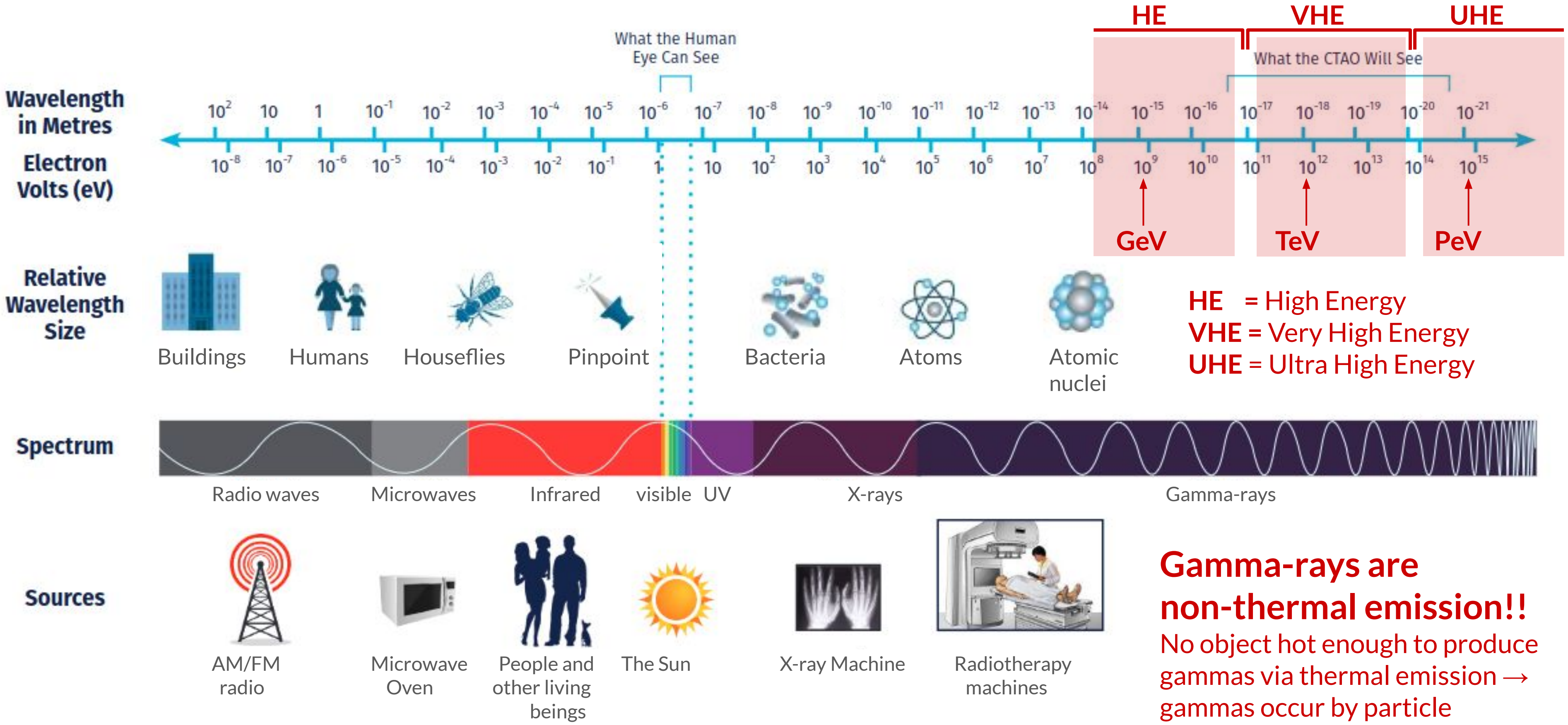
Who am I?

🐨 **Previously:** PhD from the University of Adelaide with Sabrina Einecke and Gavin Rowell
Currently: Post-doc at APC (Astroparticule et Cosmologie) lab in Paris, France

- Gamma-ray astronomer
- Member of both HESS and CTAO with a research focus on galactic TeV gamma-ray sources
- Maintainer of the open source scientific python package gammapy
- **Thank you to the organisers for this invitation!**

A Short Introduction to
Gamma-ray Astronomy

The Electromagnetic Spectrum



Gamma-rays are non-thermal emission!!

No object hot enough to produce gammas via thermal emission → gammas occur by particle acceleration processes

Image credits: Vecteezy.com, Dragonartz.net, NAOJ, NCI, CERN, NASA

Multi-wavelength Milky Way

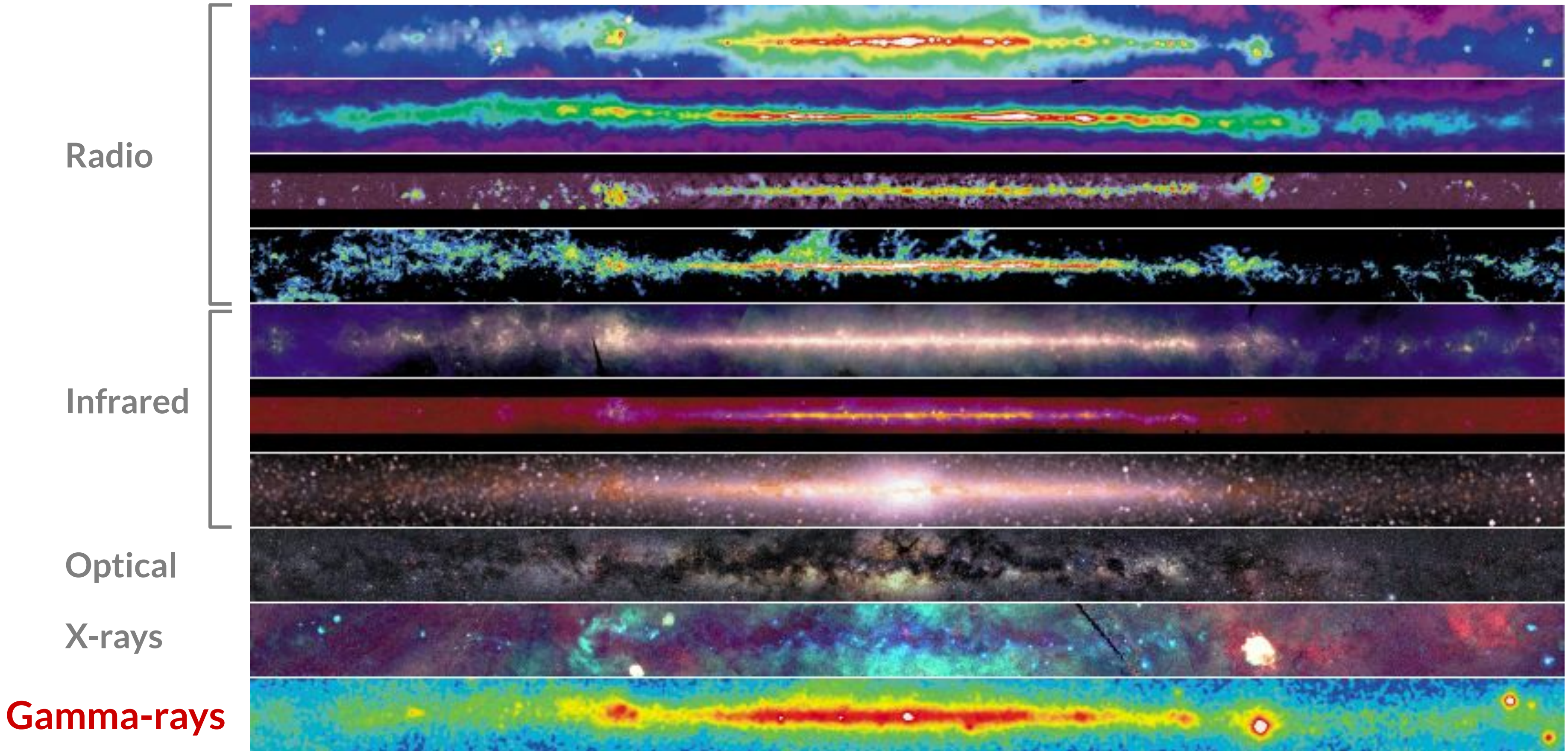
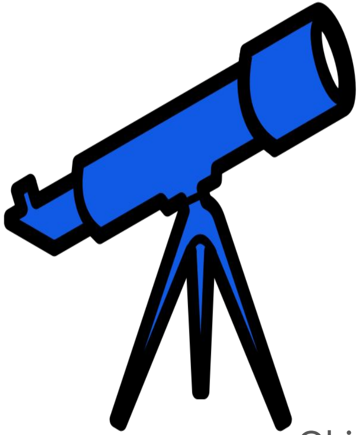


Image credit: https://asd.gsfc.nasa.gov/archive/mwmw/mmw_images.html

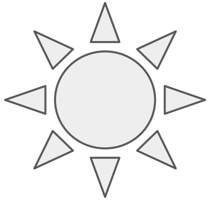
'Typical' telescope:

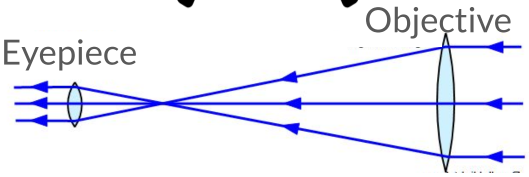
Observer

Telescope



Source



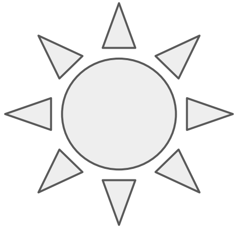


Collect light through reflection or refraction and record the image

Galaxies, stars, planets...

But a gamma-ray telescope.....

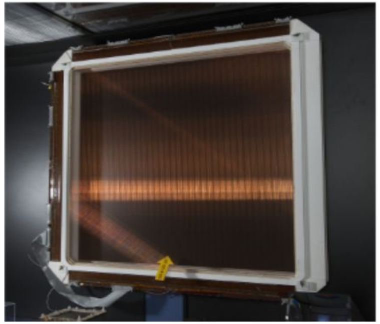
Source



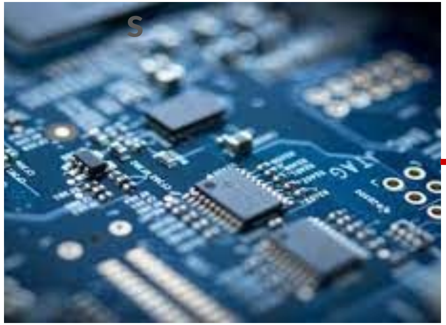
Matter

Radiation
- Electrons
- Electron + positrons
- Photons

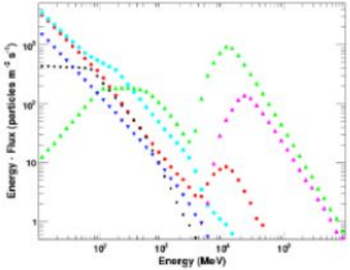
Sensors/electronics



Computer



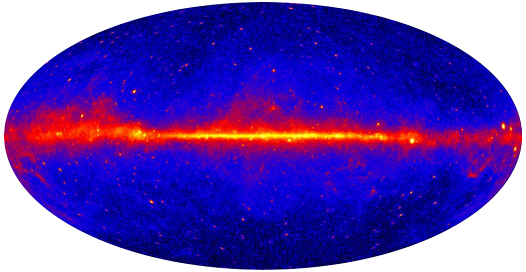
Instrument simulations, background models



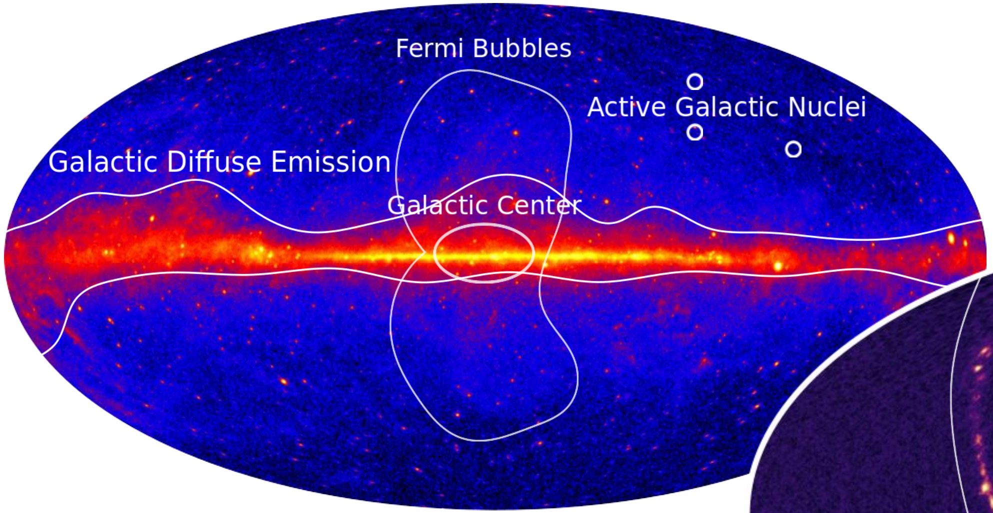
More computers

Observer

Active Galactic nuclei, Supernova Remnants, Pulsar Wind Nebulae ++



The GeV-TeV Gamma-ray Sky



Fermi-LAT all-sky counts map > 1 GeV

Lots of interesting science questions:

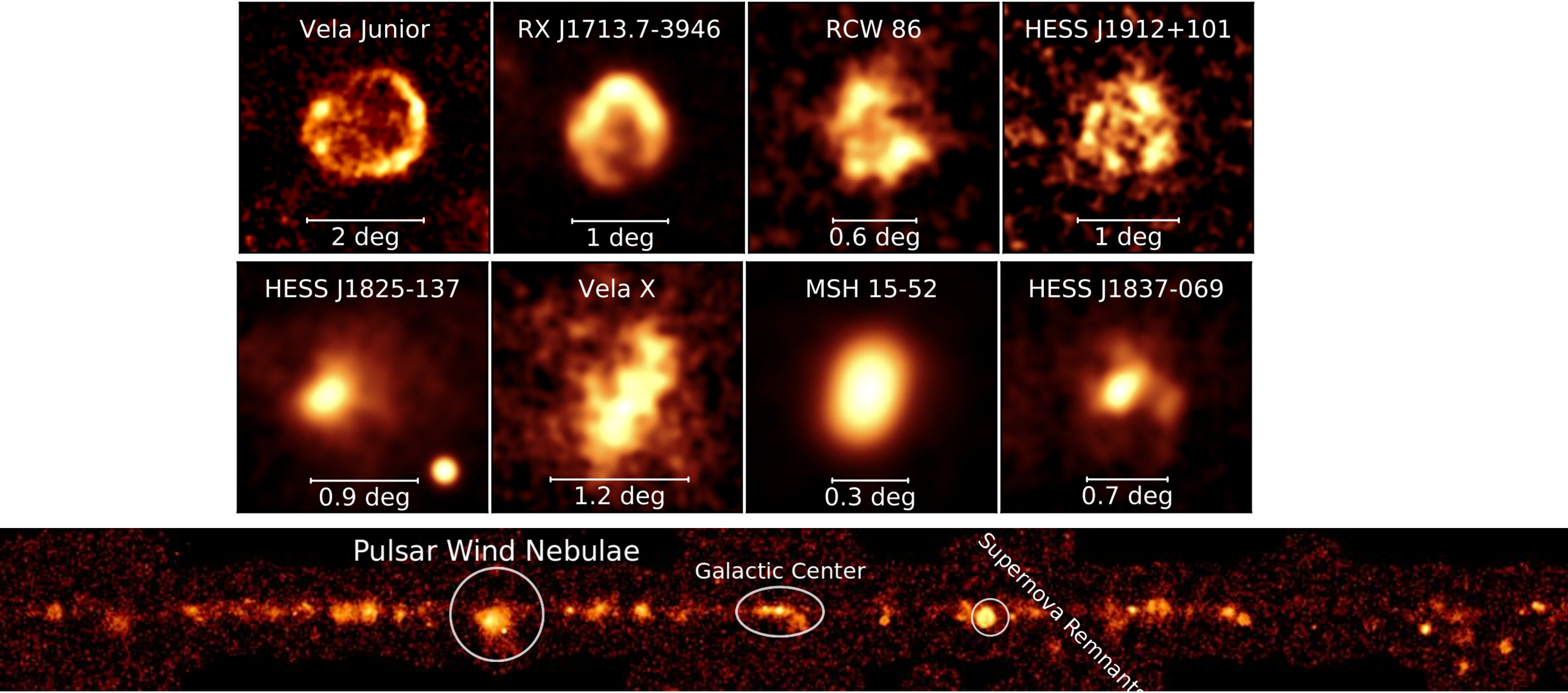
- Where are the acceleration sites of cosmic rays?
- Can particles reach PeV energies?
- Search for signatures of Dark Matter

HAWC all-sky significance map > ~1 TeV

H.E.S.S. Galactic Plane Survey significance map >1 TeV

Slide credit: Axel Donath

The GeV-TeV Gamma-ray Sky



H.E.S.S. Galactic Plane Survey significance map >1 TeV

Slide credit: Axel Donath

Gamma-ray instruments

- Ground based **Imaging Atmospheric Cherenkov Telescopes**, H.E.S.S., VERITAS, MAGIC, FACT, CTA*
  - Pointing instruments with good angular and energy resolution. Short duty cycle, can only operate by night. Large effective area, suited for VHE range, above a few tens of GeV
- **Water Cherenkov Observatories** HAWC, LHAASO and SWGO*
  - All sky coverage, long duty cycle, poorer angular and energy resolution. Large effective area, suited for VHE and UHE range above 1 TeV
- **Satellite based instruments**, currently only Fermi-LAT.
  - Good angular and energy resolution, but limited effective area. Thus limited to GeV energy range long duty cycle. Below 500 GeV

All these observatories have complementary properties, previously one could not make advantage of this BUT with gammapy we can

Slide credit: Axel Donath

Cherenkov Telescope Array Observatory

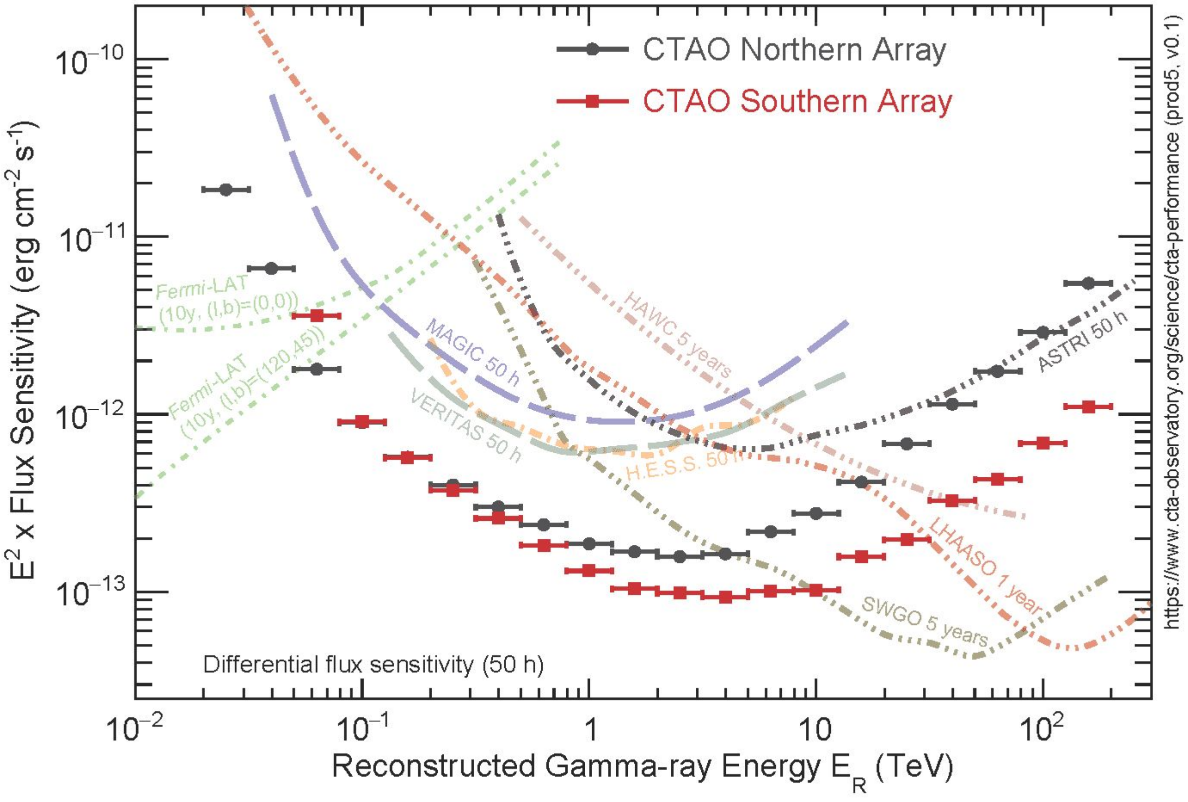
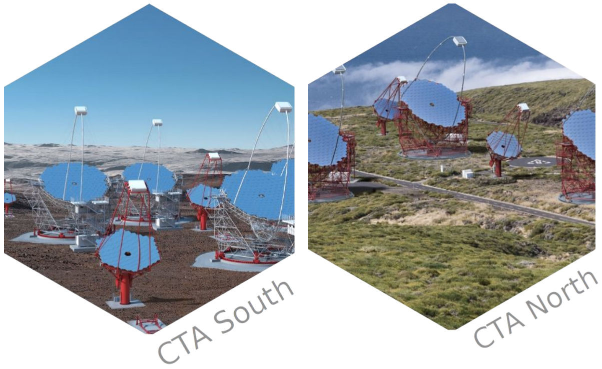


Image credits: https://www.ctao.org/ *Slide credit: Axel Donath*

The sites have access to the entire sky!



- CTAO will improve sensitivity ~10x compared to existing instruments
- Improved angular resolution by ~3x
- Operate as an open gamma-ray observatory for the first time, with making data public
- Gammapy selected in 2021 as library for the CTA science tools

Example of CTAO performance

- CTAO will explore further and deeper, covering the entire LMC
- Permits the probing the population of VHE emitters in this galaxy and the connection to global galaxy properties

Simulated comparison of CTA's survey of the LMC with current optical and H.E.S.S. images

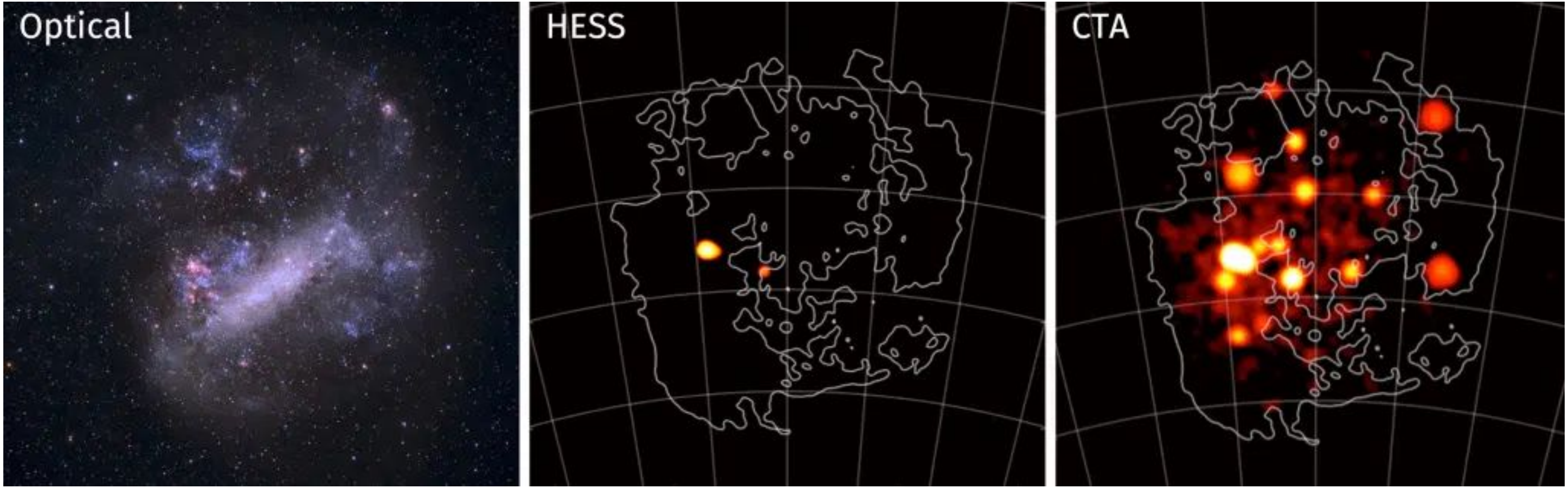


Image credit: https://www.ctao.org/

An Introduction to
Data Analysis Methods

IACT Working Principle

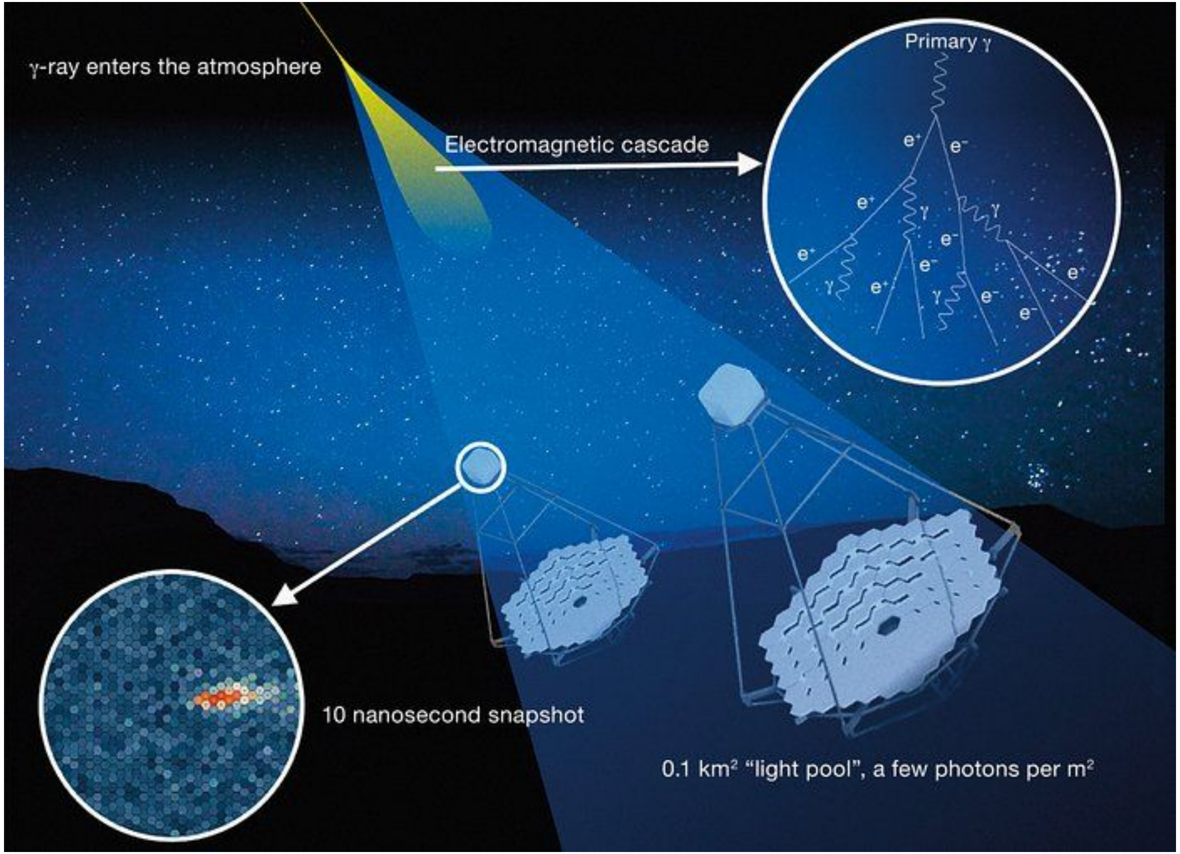


Image credit: https://www.eso.org/public/images/eso1841x/

- Gamma-ray enters the upper atmosphere and triggers a "particle shower": electromagnetic cascade of secondary particles
- Secondary particles move faster than local speed → Cherenkov radiation
- Nanosecond flashes are detected by optical telescopes on the ground with photomultiplier cameras, within the light pool
- With multiple telescopes: can reconstruct energy and arrival direction of incident gamma-ray

IACT Working Principle

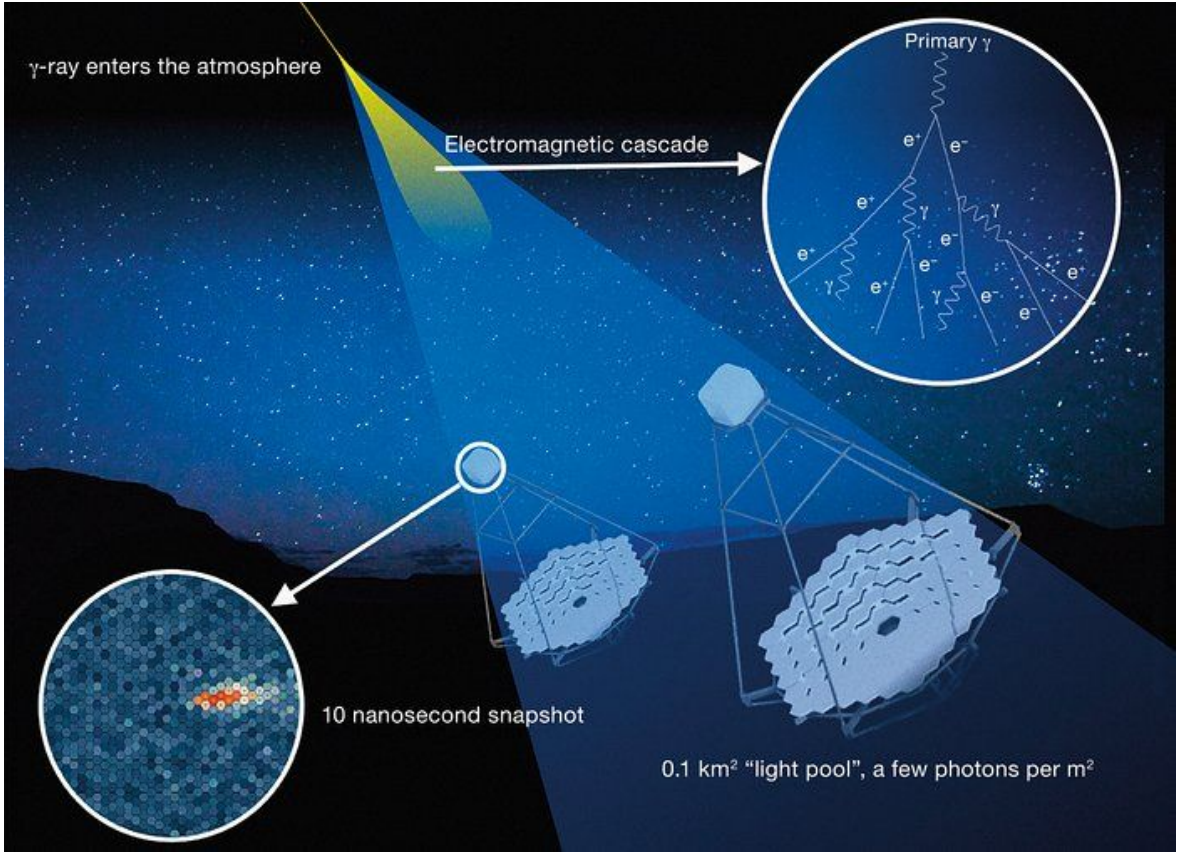
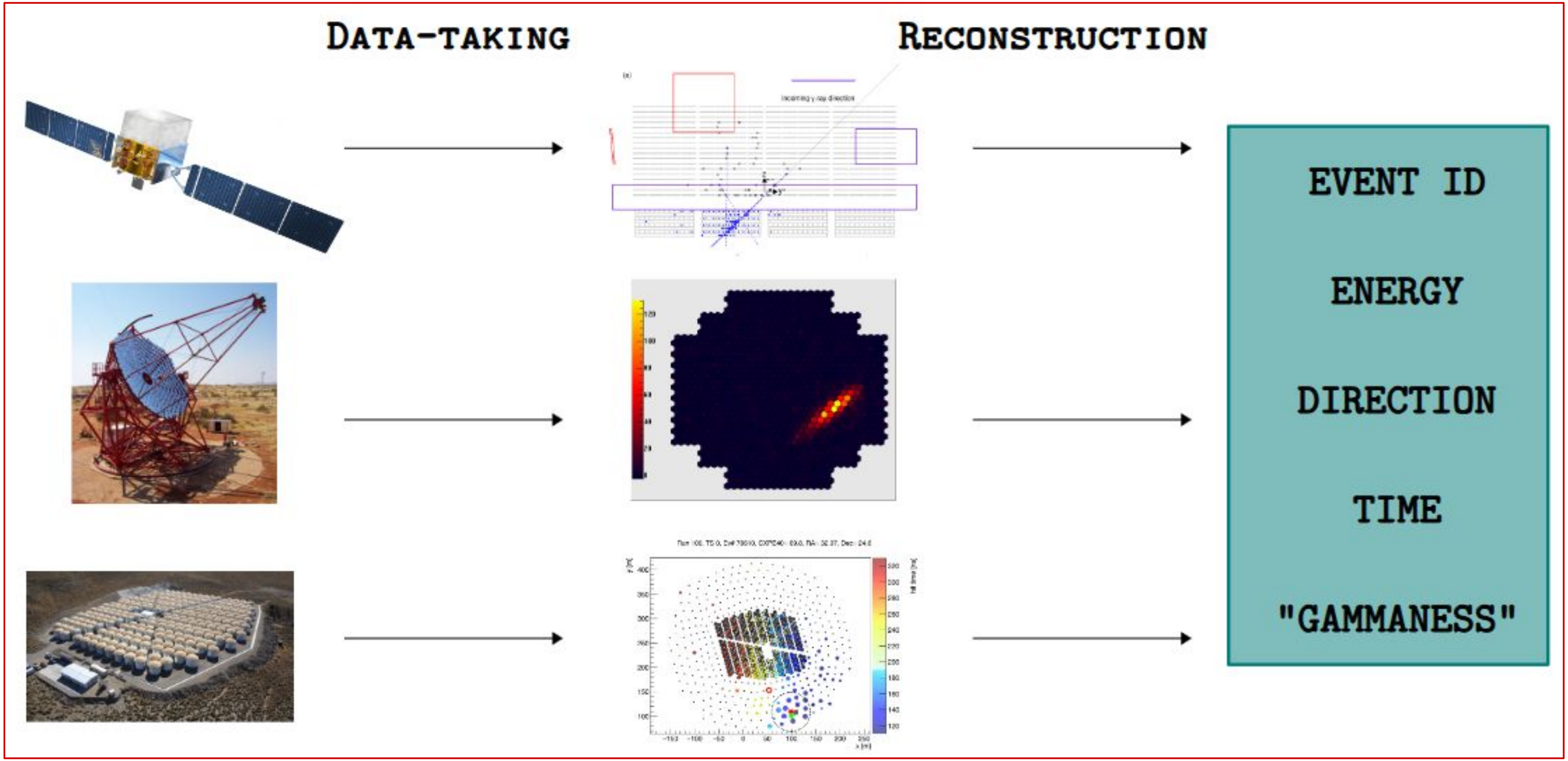


Image credit: https://www.eso.org/public/images/eso1841x/

- Instrument is pointed for a given period of time at position on the sky and measures events
- This is called an “**observation**” or “**run**”
- The instrument characteristics (“response”) is assumed to be stable during the whole observation run
- This is the fundamental “chunk” of data, from which the data is combined (“stacked”) or split, depending on the analysis case.
- For data bookkeeping each observation is assigned a unique identifier

Gamma-ray instrument data



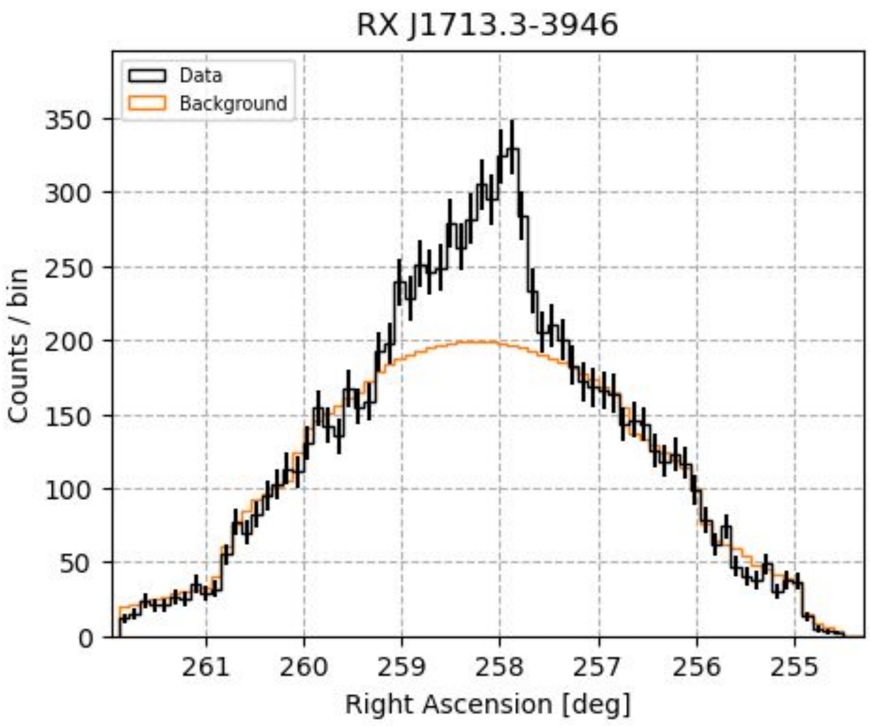
Slide credit: Laura Olivera-Nieto

Aside

- Many analysis methods/data models in gamma-ray astronomy are **inspired by X-ray astronomy**
- In general the structure of the data is very similar, with limited counts, Poissonian nature and requirement to handle the instrumental response
  - Such as limited angular and energy resolution as well as non-uniform exposure and effective detection area
- In detail the structure is of course different, with the largest difference being the complexity of spectral and spatial features at high resolution and the hadronic background domination for ground based gamma-ray observatories
- The standard statistical analysis method in gamma-ray astronomy is binned maximum likelihood fitting of spatial/spectral models, taking the instrument response into account

Gamma-ray analysis

- Reconstruction of gamma-ray events is a complex procedure
- For IACTs we are background dominated
  - Slices along Right Ascension and encompassing the RX J1713.7-3946 (source with significant gamma-ray signal)
  - Figure shows the 'bump' in gammas above the background

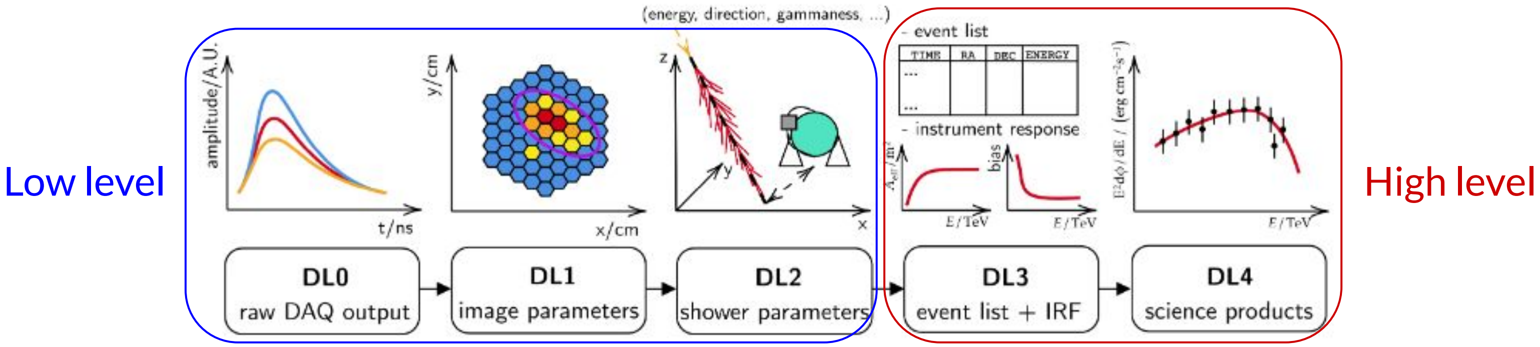


Format for the data?

- In the ~10 years there has been a big effort to define a common format for storing gamma-ray data
- We learn from existing standards for x-ray astronomy and Fermi
- Result: Gamma-Astro-Data-Format (GADF)
  - Based on FITS standards, follow FITS conventions for time and coordinates
  - Information stored in binary tables in specific Header Data Unit (HDU
- If the data looks the same, we can easily share it in a tool
- This is where gammapy comes in!
  - See later

Gamma-ray data analysis

- Low level analysis involves:
  - Calibrating the data
  - Applying cuts to reject hadrons
  - Reconstructing showers
  - Final product are event lists
- High level analysis utilises the event lists to produce a number of interesting products including, but not limited to
- We typically call these products data level (DL) 3-5



https://ui.adsabs.harvard.edu/abs/2021Univ....7..374N/abstract

What does the DL3 data look like?

Gamma-like event list

EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
uint32	float64	float32	float32	float32
1	664502403.0454683	-92.63541	-30.514854	0.03902182
2	664502405.2579999	-92.64103	-28.262728	0.030796371
3	664502408.8205513	-93.20372	-28.599625	0.04009629
4	664502409.0143764	-94.03383	-29.269627	0.039580025
5	664502414.8090746	-93.330505	-30.319725	0.03035851
6	664502415.5855484	-93.23232	-28.587324	0.034782063
...	...	...	...	...
106212	664504199.9285337	-94.80082	-28.978634	0.6445309
106213	664504199.9315486	-89.076965	-29.361399	0.101013534
106214	664504199.9442778	-93.28163	-30.011133	0.030123588
106215	664504199.9560002	-92.4204	-29.911001	0.086367
106216	664504199.9636843	-93.404785	-28.791838	0.076767094
106217	664504199.9682832	-93.45982	-30.112665	0.037619844

Good time interval (GTI)

START	STOP
Time	Time
53343.92234009259	53343.94186555556

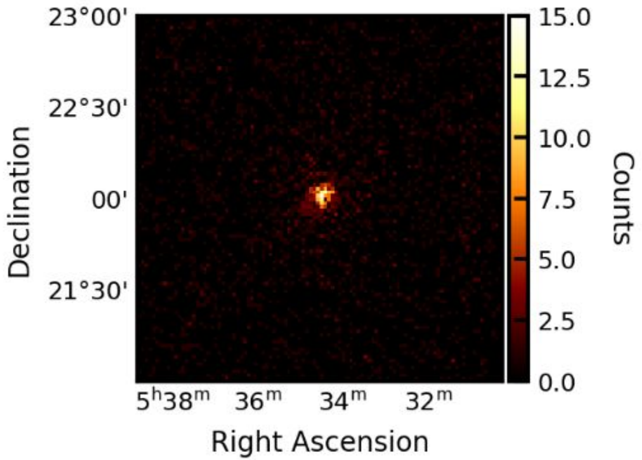
Time interval in MJD that the observations were taken

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

Event lists to physics

This is data reduction

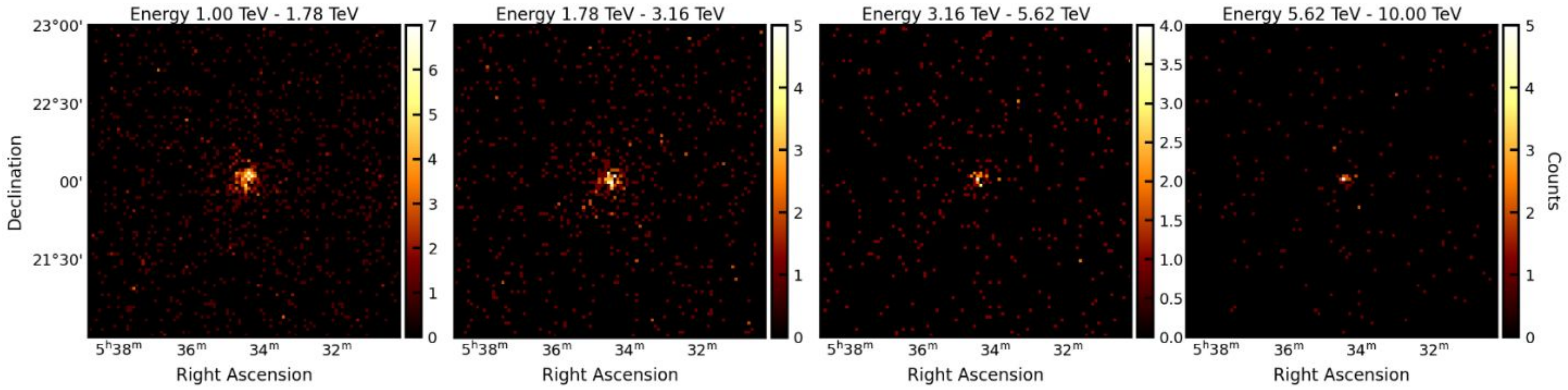
- Let's say you have a list of events taken by a gamma-ray instrument when observing a source, and you want to study that source
- You can make a map! Which is just a 2D histogram of the sky coordinates



Event lists to physics

This is data reduction

- Let's say you have a list of events taken by a gamma-ray instrument when observing a source, and you want to study that source
- You can make a map! Which is just a **3D** histogram of the sky coordinates and **energy***



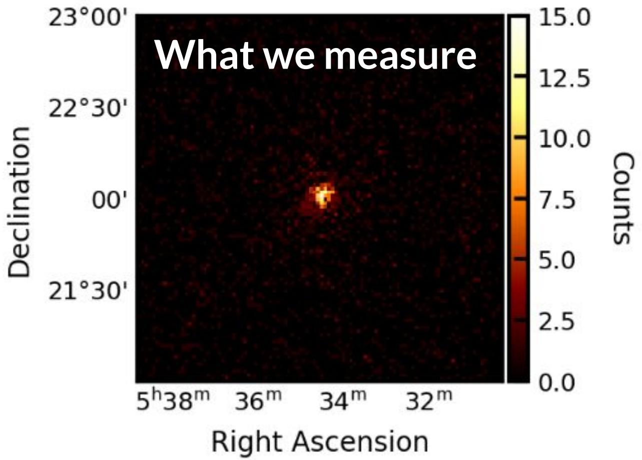
*** or time, or some other quantity...**

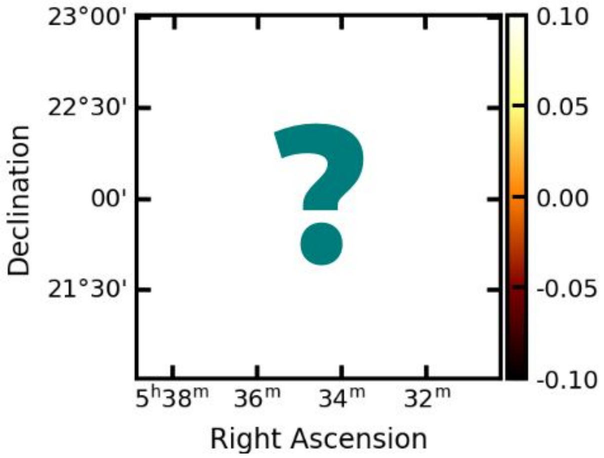
Slide credit: Laura Olivera-Nieto

Event lists to physics

This is data reduction

- Let's say you have a list of events taken by a gamma-ray instrument when observing a source, and you want to study that source
- You can make a map! Which is just a 3D histogram of the sky coordinates and energy





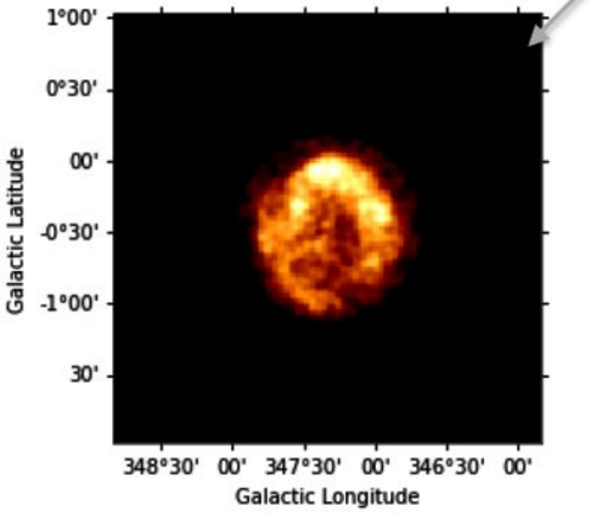
First: lets understand this from another view

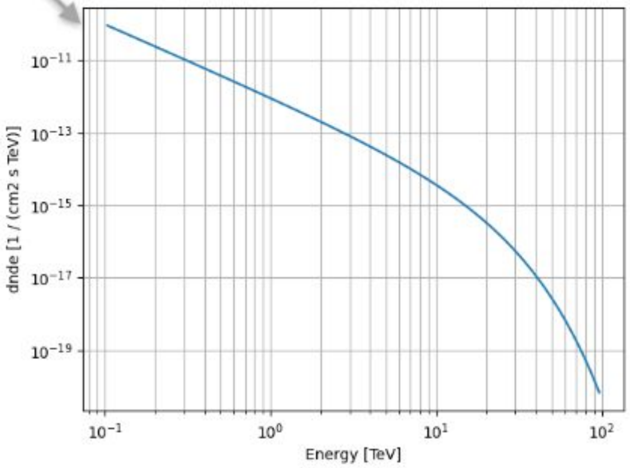
Slide credit: Laura Olivera-Nieto

Event lists to physics

- Assume the source emits *S* gamma-ray photons
- The emission represented by a sky model $\Phi(p_t, E_t)$ with the position of the photon on the sky (p_t) and its energy (E_t)
- Want to measure the model parameters that best reproduce the measured data

$\Phi(p_t, E_t)$





Event lists to physics

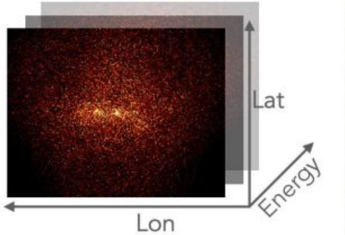
Gamma-like events

EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
uint32	float64	float32	float32	float32
1	664502403.0454683	-92.63541	-30.514854	0.03902182
2	664502405.2579999	-92.64103	-28.262728	0.030796371
3	664502408.8205513	-93.20372	-28.599625	0.04009629

...binned into...

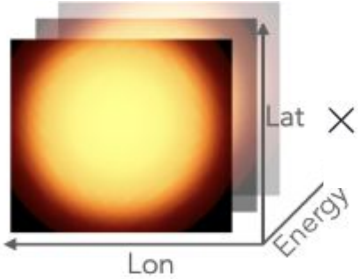
Counts

$N_{obs} =$



Background template

$N_{bkg} =$



$\times\, n_0 \left(\frac{E}{E_0} \right)^{-\Gamma}$

Binned poisson log-likelihood

"Cash statistics": summed over all "bins"

$$\mathscr{C} = 2 \sum_i N^i_{Pred} - N^i_{Obs} \cdot \log N^i_{Pred}$$

$$N_{Pred} = N_{Bkg} + \sum_{Src} N_{Pred,Src}$$

- Predicted counts are computed per model component ("source / object") and summed
- A "global" background model template with "correction parameters" is added

Slide credit: Axel Donath

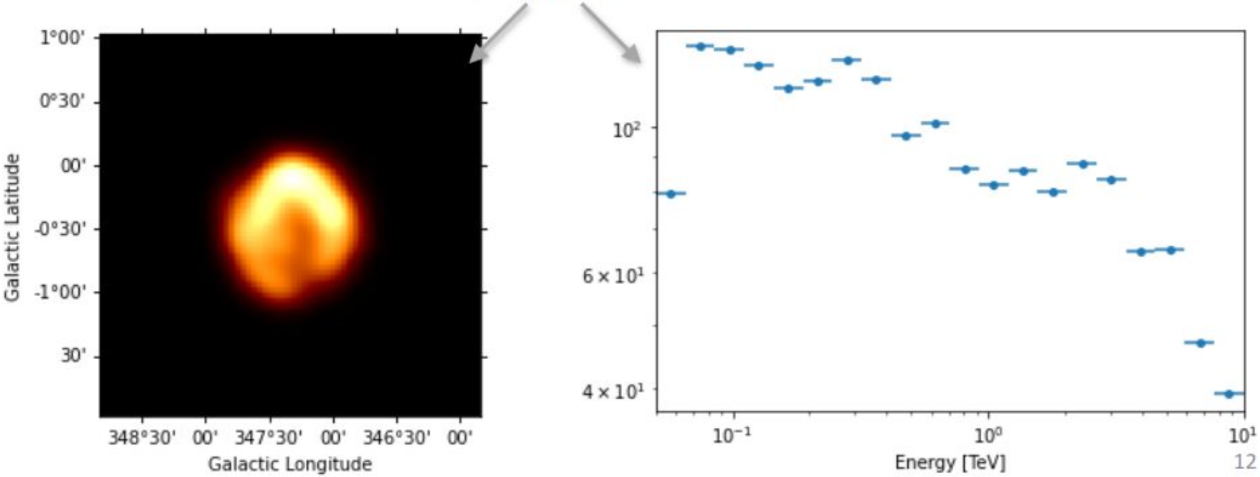
Event lists to physics

- The event list alone is not enough for this, we need **IRFs!**
- Number of observed photons from source *S* is:

$$N(p,E)dpdE = t_{obs}\int_{E_t} dE_t \int_{p_t} dp_t \, R(p,E|p_t,E_t) \times \Phi(p_t,E_t)$$

IRF

$N(p,E)$



Instrument Response Functions (IRFs)

- We assume that the IRF can be simplified to the product of:

$$R(p, E \mid p_t, E_t) = A_{\text{eff}}(p_t, E_t) \times PSF(p \mid p_t, E_t) \times E_{\text{disp}}(E \mid p_t, E_t)$$

Effective area

Point spread function

Energy dispersion

- Measured events also contain background events

$$N(p, E) = \sum_{S} N_S(p, E) + N_{bkg}(p, E)$$

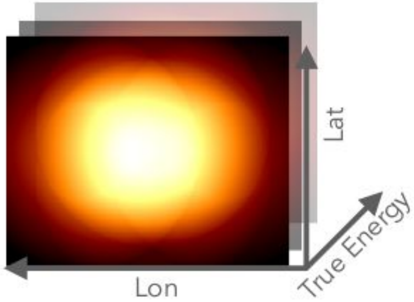
Number of predicted photons from source S

Number of background events

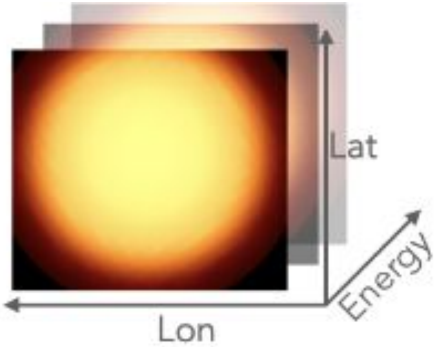
Instrument Response Functions (IRFs)

- There are a number of different IRFs
  - Effective area, PSF, Energy dispersion, Background
- DL3 IRFs are reprojected onto the target geometry

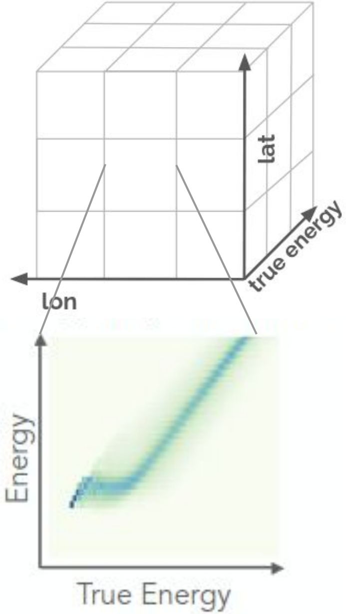
Exposure



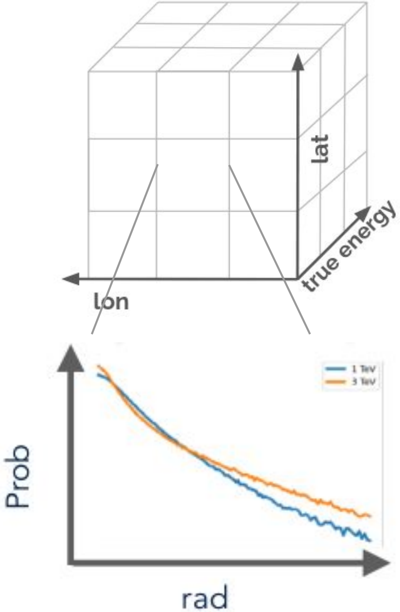
Background template



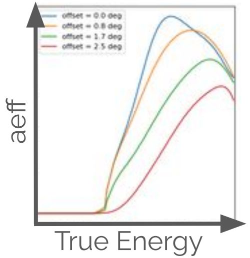
Energy Dispersion



PSF

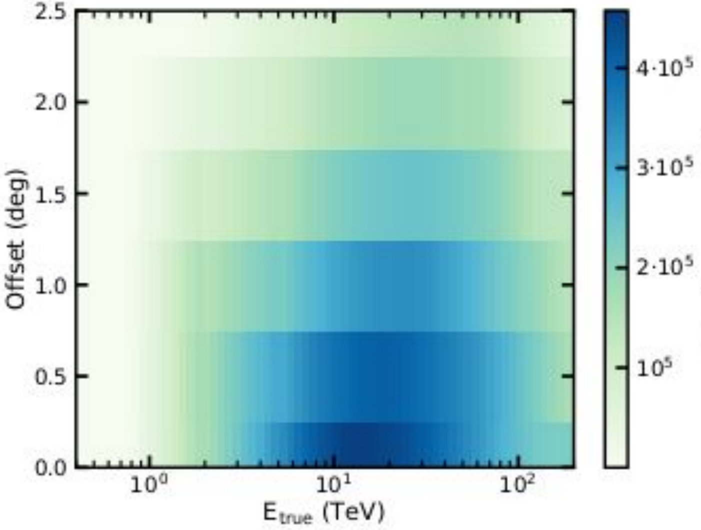


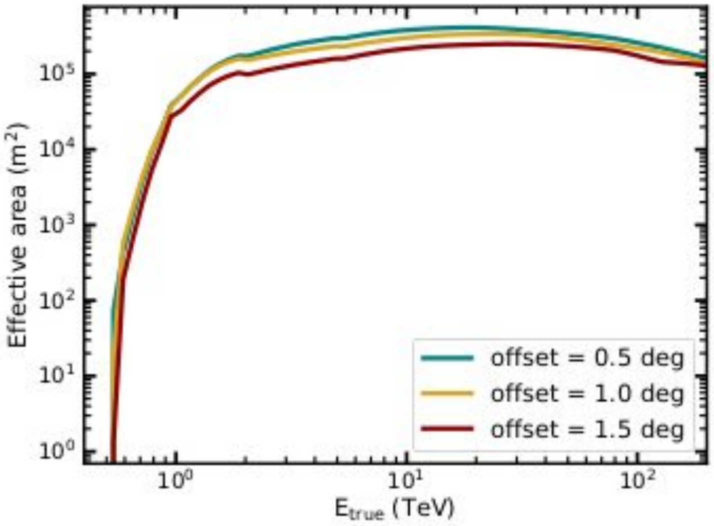
Effective Area



Effective area

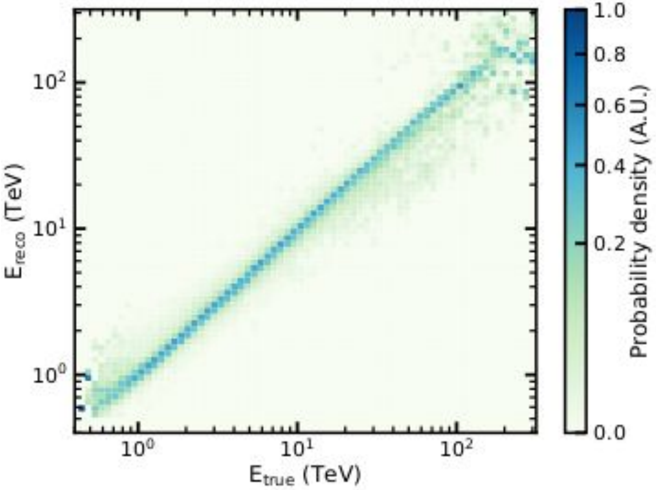
- Effective area: gives the number of expected (true) gamma-like events per area
- Varies with offset from pointing position and true energy
- Estimated from "Monte Carlo" simulations
- Required to measure flux/brightness of gamma-ray sources
- Typically combined with observation time to derive the effective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

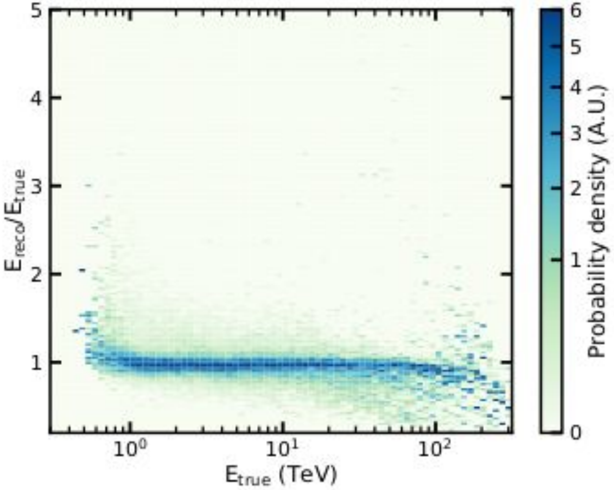




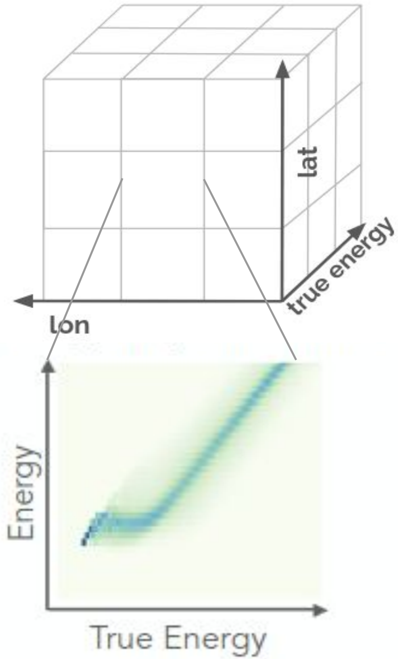
Energy dispersion

- 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
- Required to measure precise spectra of source, especially at low energies (for IACTs)



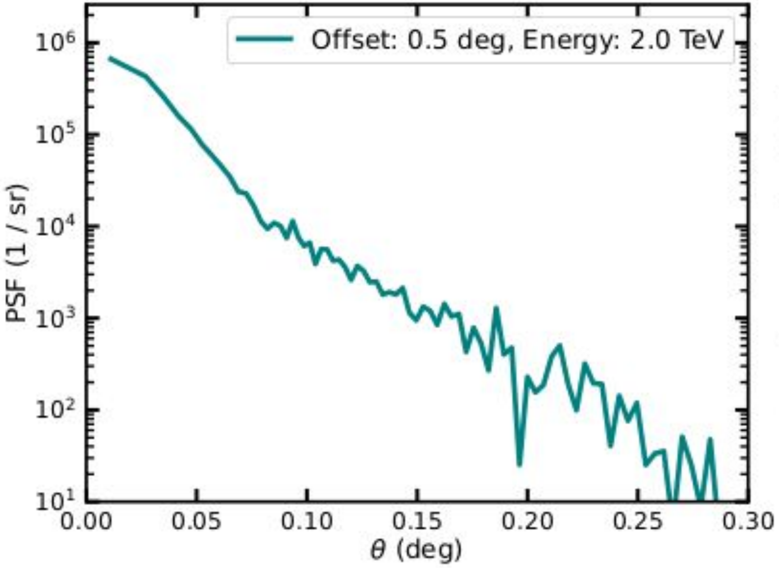


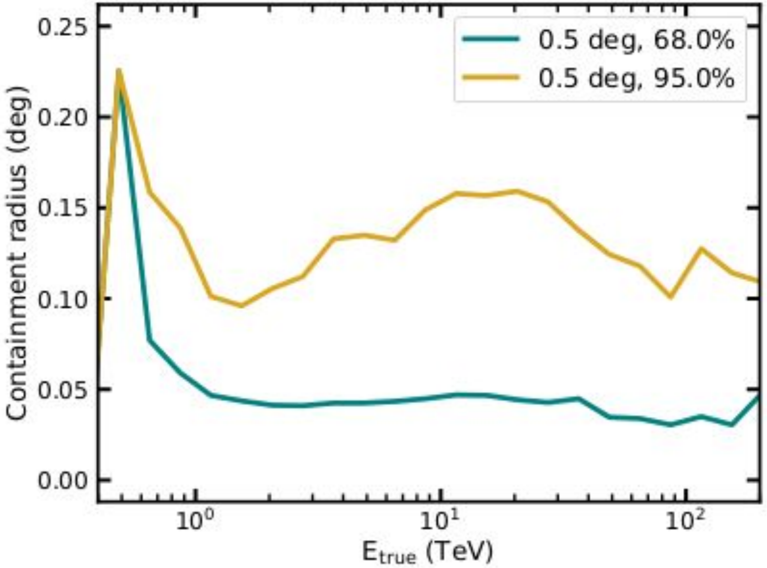
Energy Dispersion



Point Spread Function

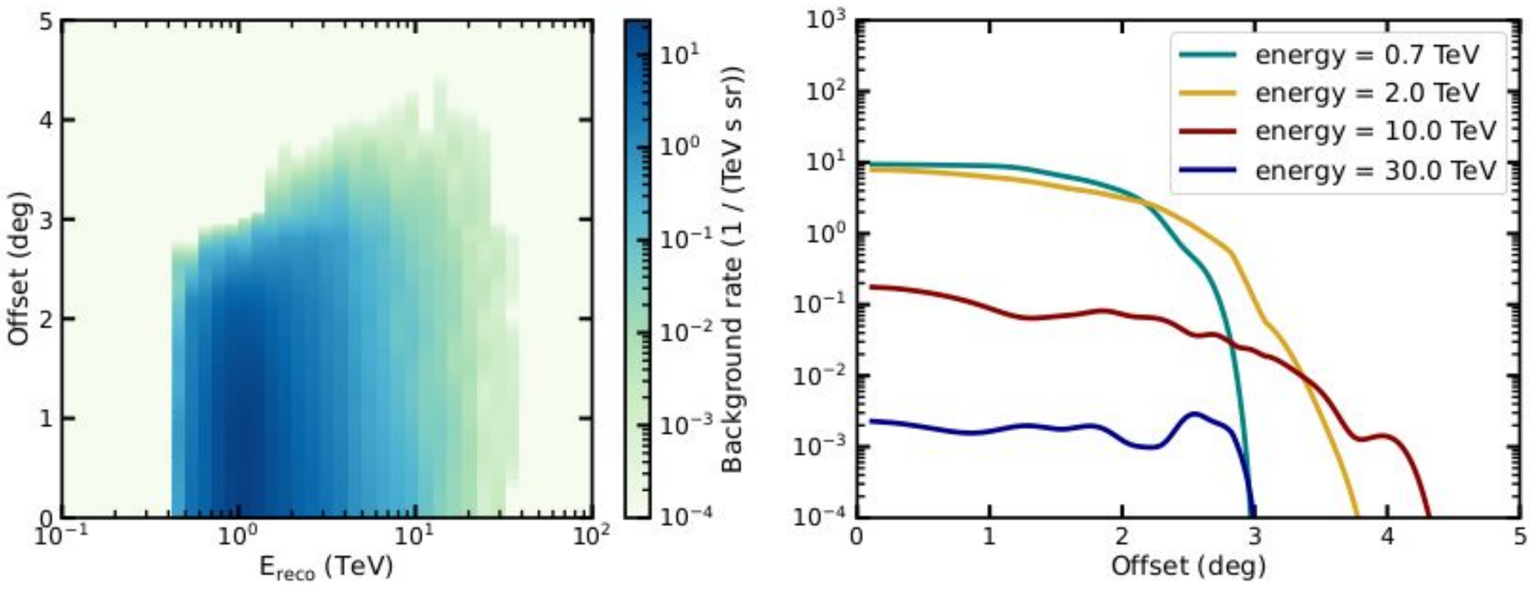
- PSF: angular resolution of the instrument, precision to reconstruct the arrival direction of an event
- Estimated from "Monte Carlo" simulations and by binning events into offset and true energy
- Typically stored as a "radial profile"
- Required to measure extension of galactic sources and precise flux of point sources





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



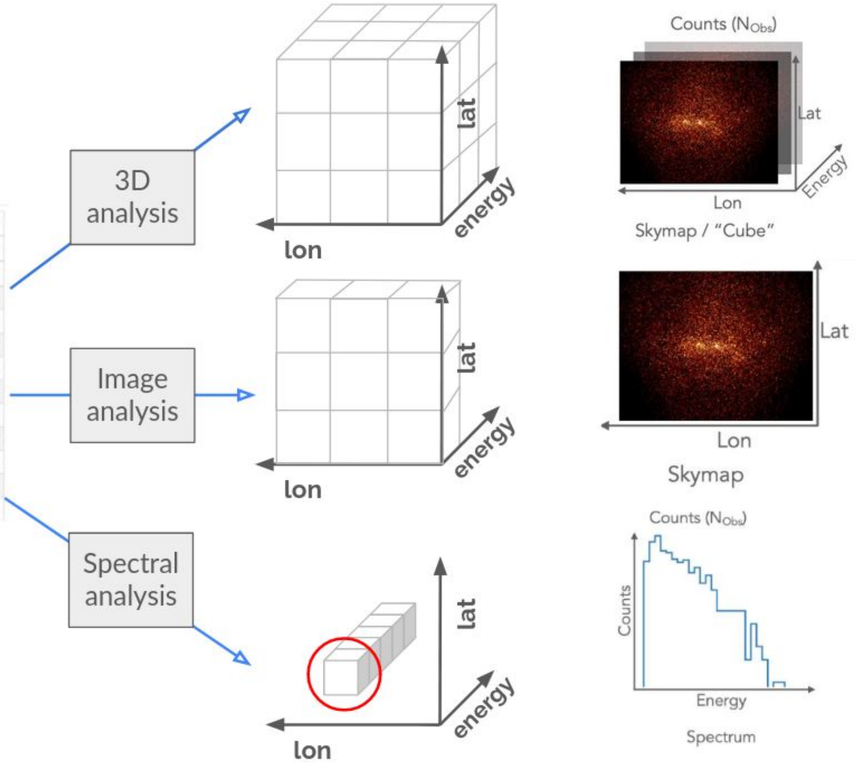
DL3 - DL4 : data reduction

1. Select and retrieve relevant observations
2. Define the reduced dataset geometry
   a. Is the analysis 1D (spectral only) or 3D?
   b. Define target binning and projection
3. Initialize the data reduction methods (makers)
   a. Data and IRF projection
   b. Background estimation
   c. Safe Mask determination
4. Loop over selected observations
   a. Apply makers to produce reduced datasets
   b. Combine them for stacked or joint analysis

Data reduction summary DL3 → DL4

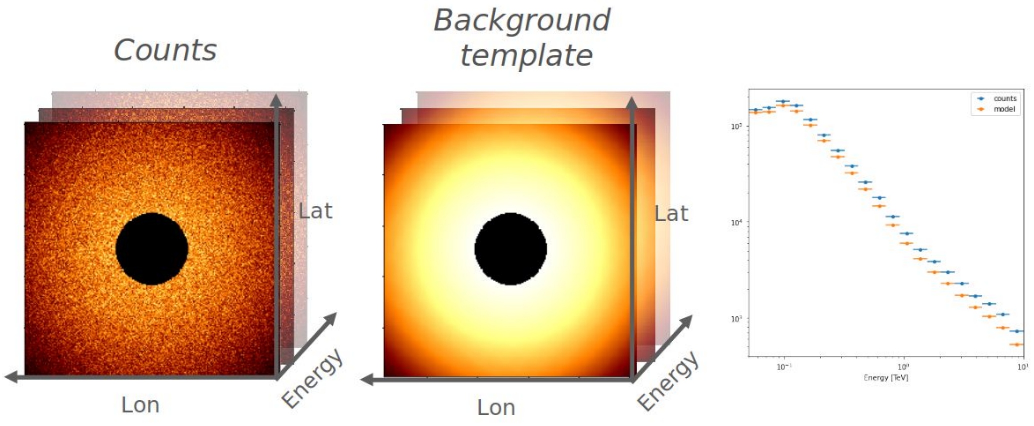
- Bin events (and IRFs) into n-dim sky maps
  - Apply event selections (time, offset, etc)
  - Spatial and energy binning
- Generalised case: 3D maps
  - Image analysis: cube with one energy bin
  - Spectral analysis: Cube with one spatial bin

EVENT_ID	TIME	RA	DEC	ENERGY
	s	deg	deg	TeV
int64	float64	float32	float32	float32
5407363825684	123890826.66805482	84.97964	23.89347	10.352011
5407363825695	123890826.69749284	84.54751	21.004095	4.0246882
5407363825831	123890827.23673964	85.39696	19.41868	2.2048872
5407363825970	123890827.79615426	81.93147	20.79867	0.69548655
5407363826067	123890828.26131463	85.98302	21.053099	0.86911184
5407363826095	123890828.41393518	86.97305	21.837437	4.1240892
5407363826128	123890828.52555823	83.40073	19.771587	1.6680022
5407363826168	123890828.6829524	82.25036	19.22003	4.7649446
5407363826383	123890829.53362775	83.18322	22.008213	0.7920148
...	...	...	...	...

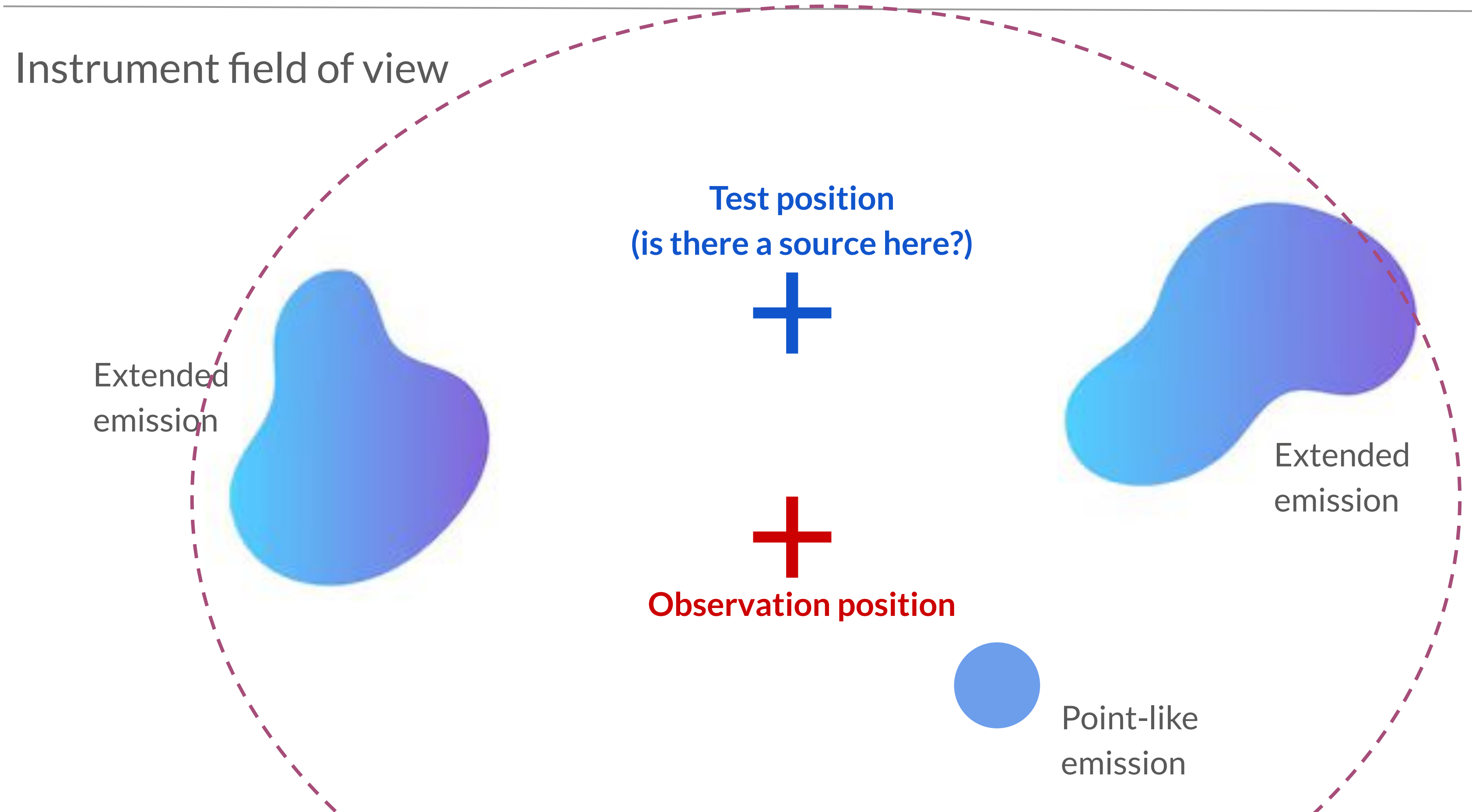


Correcting the background

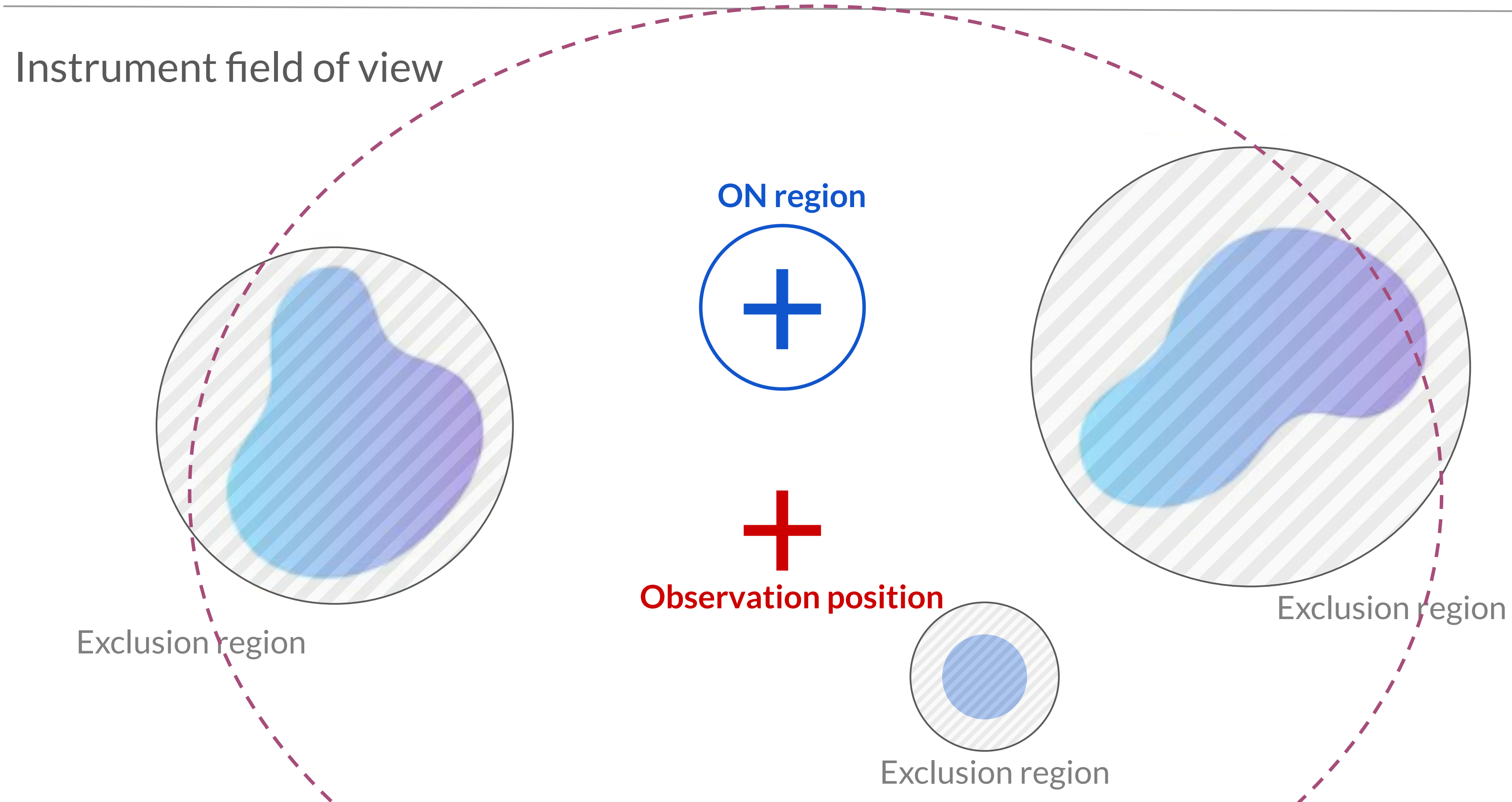
- Background model provides the shape of the expected hadronic background and average rate
- Run-to-run variation in trigger rate, means the background needs to be corrected to the conditions of each observation run
- To further reduce systematics, the background can be measured directly in the data
- In general, we ensure obtain the number of counts where no sources are expected to adjust the background level



A piece of VHE sky...

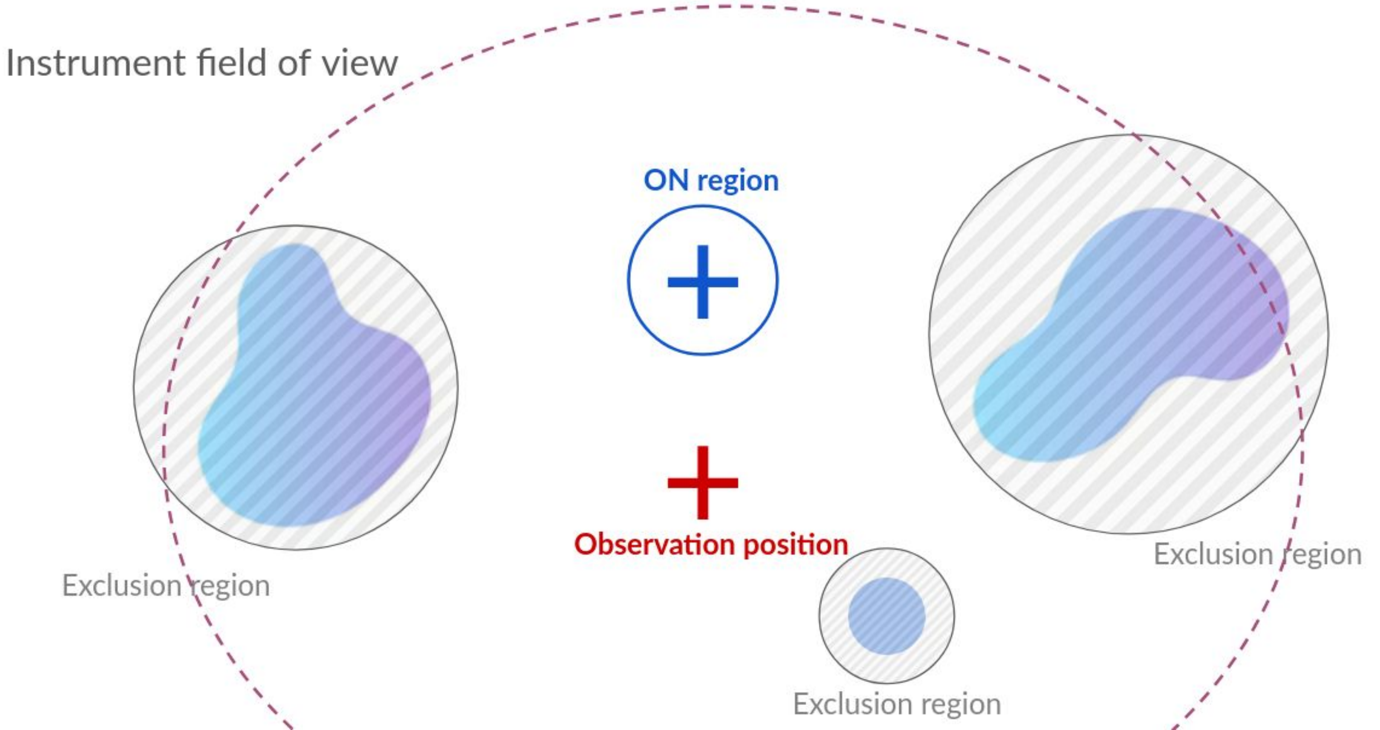


ON region


Instrument field of view
ON region
Observation position
Exclusion region
Exclusion region
Exclusion region

OFF region

- For each observation, the 'on' region is where gamma-ray sources are expected and the 'off' region contains no gamma-ray source emission

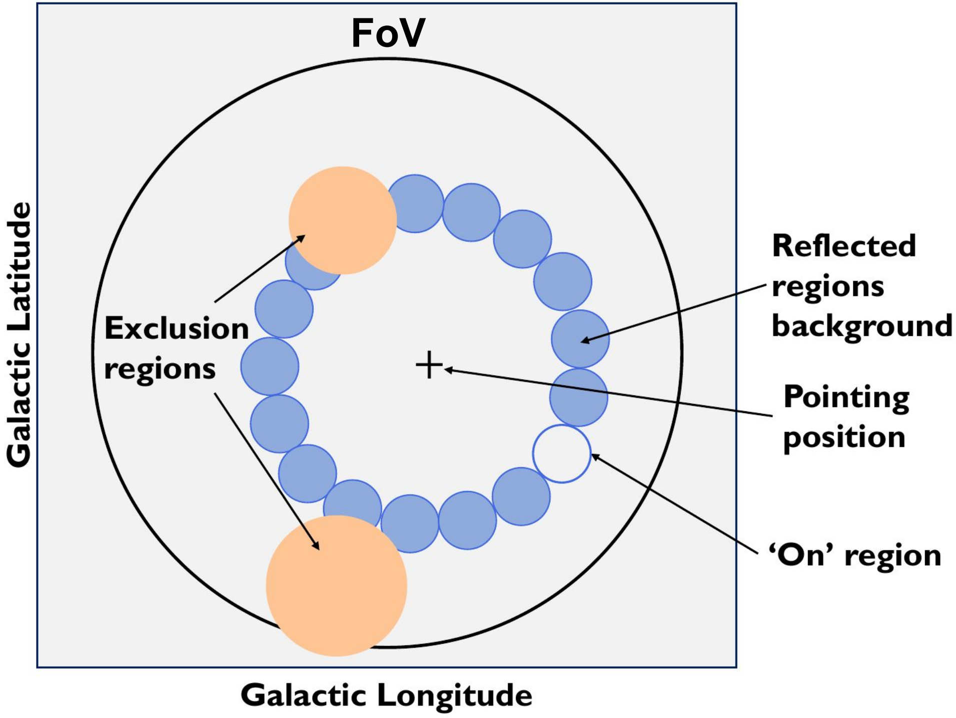


- Need to define an OFF region to estimate the background → few techniques to do this...

1D analysis technique

Reflected regions method:

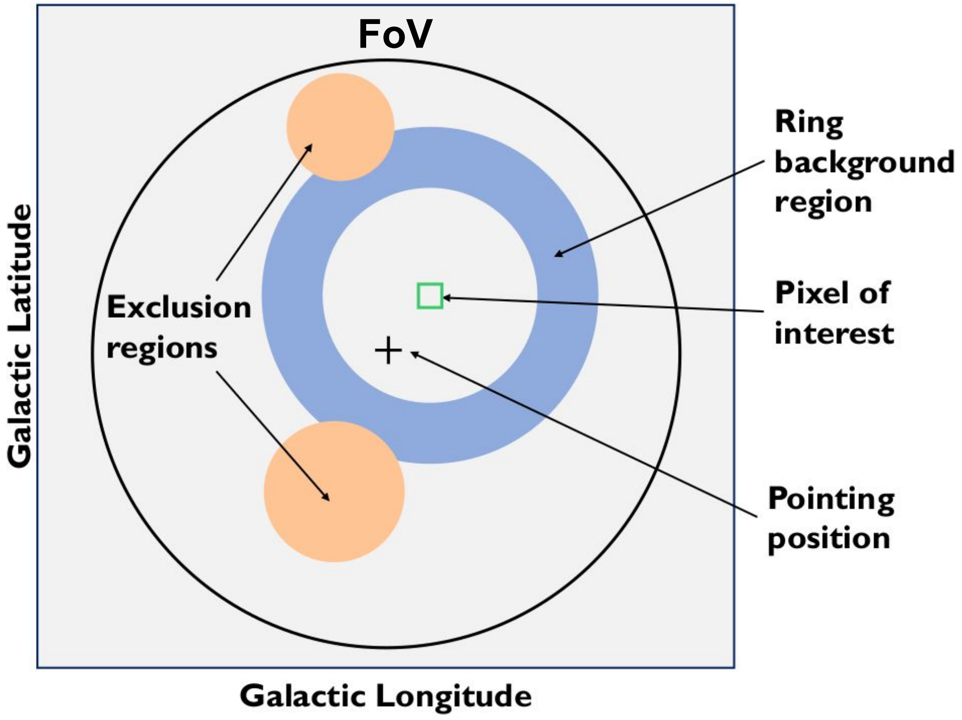
- The 'off' region has the same shape, size and offset from the pointing position as the 'on' region
- This technique is used for spectral analysis, as the spatial dimension is lost when grouping the pixels inside the 'on' region



2D analysis technique

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



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
- Below shows normalisation and tilt varies in the spectral shape for the model

$$\mathcal{B}(E) \longrightarrow \mathcal{B}(E) \times \texttt{norm} \times \left(\frac{E}{1\,\mathrm{TeV}}\right)^{-\texttt{tilt}}$$

Modeling and fitting DL4 → DL5

- For modeling and fitting, gamma-ray analysis techniques rely on forward-folding:
  - Measured counts N is compared to predicted counts Npred

$$N_{\text{pred}}(p, E) = \sum_{S} E_{disp} \left[PSF \star \left(expo \times \Phi_S(p_t, E_t) \right) \right] + N_{\text{bkg}}(p, E)$$

Maximum likelihood technique

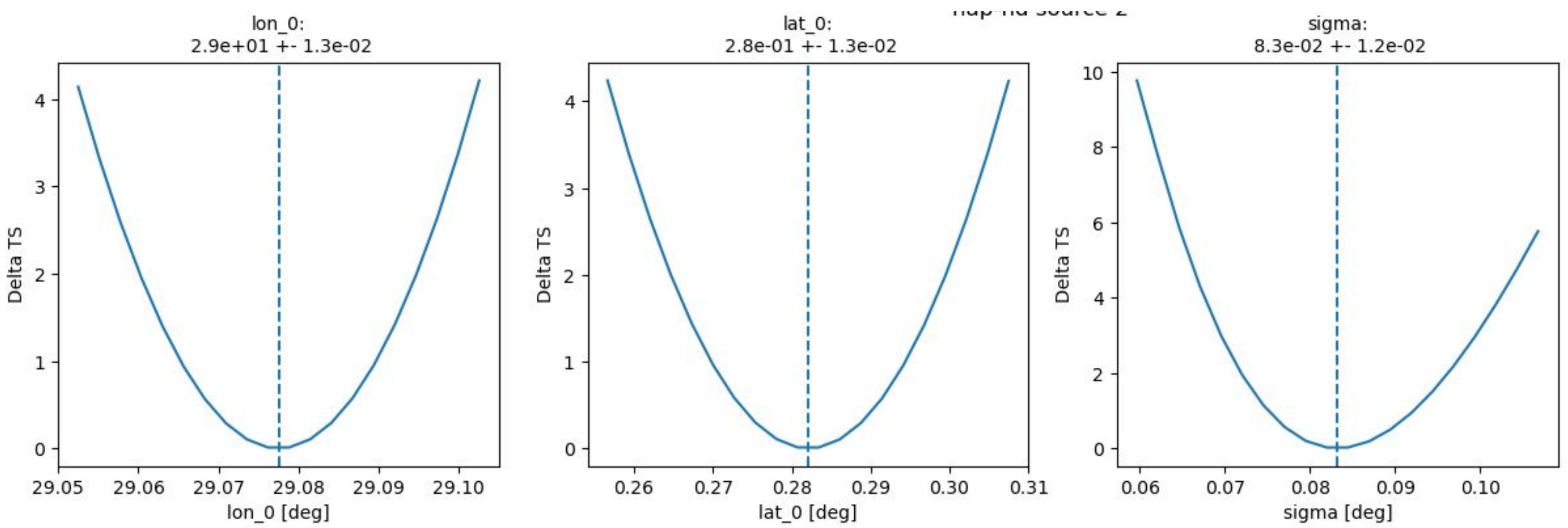
- We define the likelihood function for Poisson probability density functions

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

- Assume a model of family of models and compute the predicted counts for that set of models with their associated parameters
- Utilise the maximum likelihood technique to determine the best fit parameters and their associated uncertainties
  - The likelihood is different for whether you know the background or measure it on the background i.e you utilise two different statistics but they are just two different ways of writing the likelihood

Maximum likelihood technique

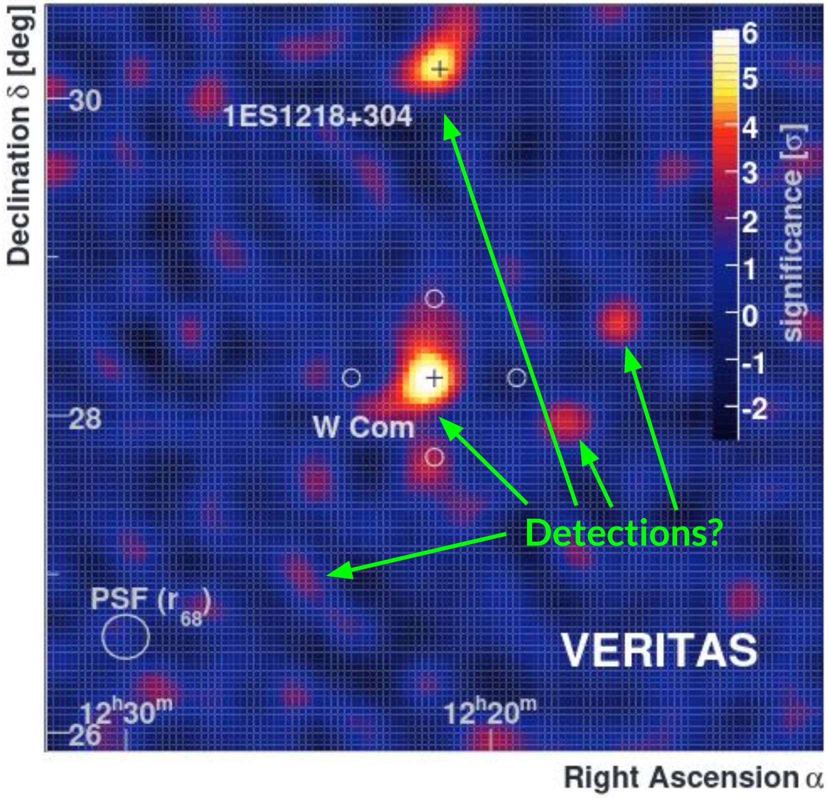
- Another way of saying that is, we find the minimum of the log-likelihood function
- For example here the source was fit with a Gaussian model with longitude, latitude and sigma free
- By plotting the scan over the parameters we can see that it is well minimised



Analysis

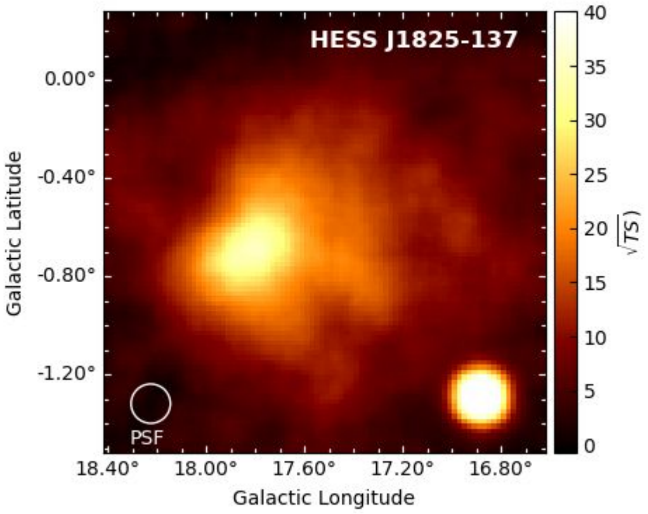
Analysis questions

- Source detection probability?
    - Test statistics (TS), significance



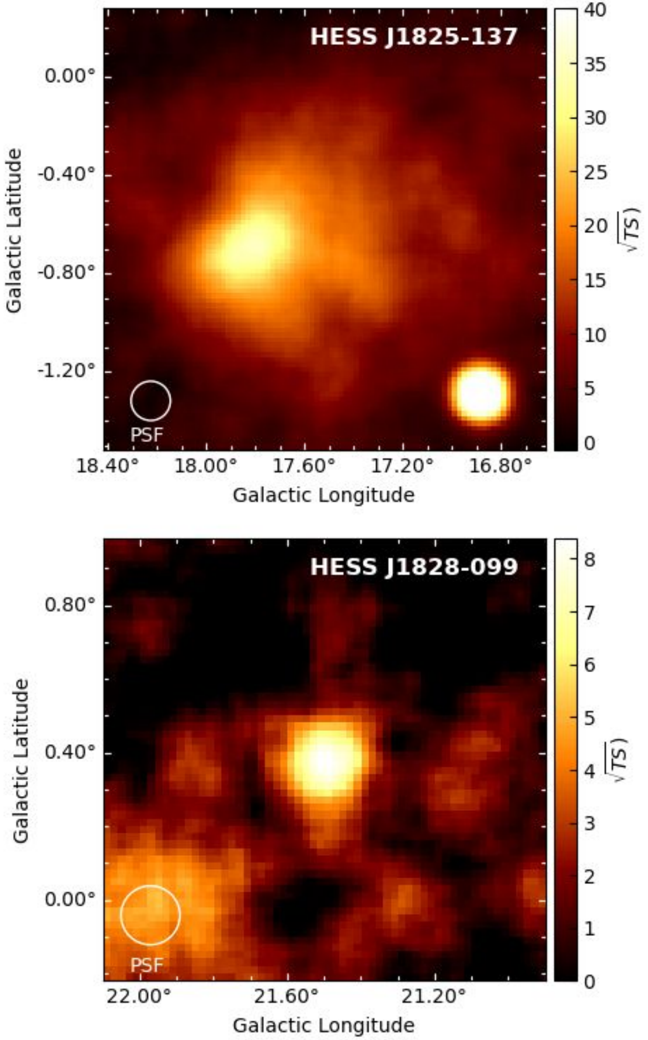
Analysis questions

- Source detection probability?
  - Test statistics (TS), significance
- What shape does the source have?
  - Extended



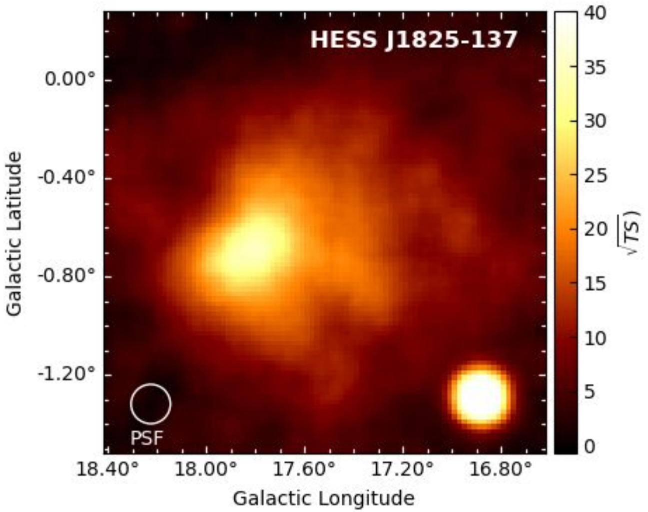
Analysis questions

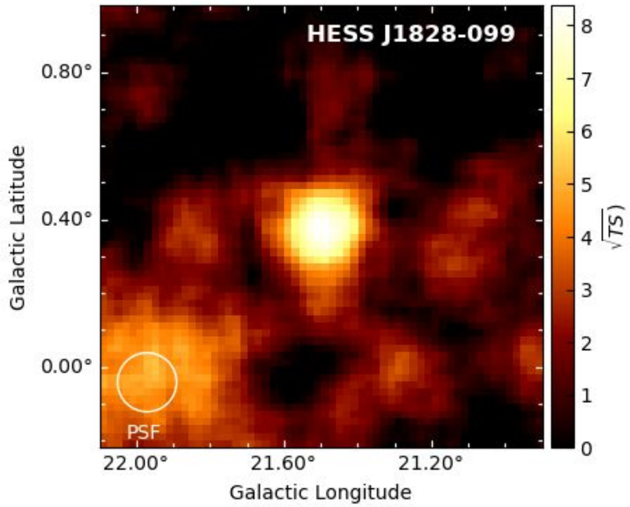
- Source detection probability?
  - Test statistics (TS), significance
- What shape does the source have?
  - Extended, point-like, …



Analysis questions

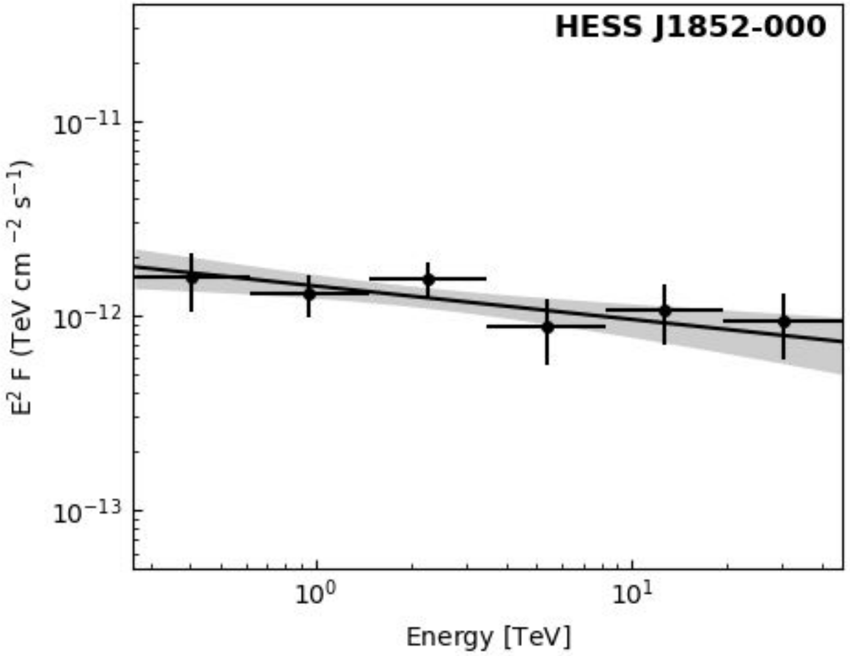
- Source detection probability?
  - Test statistics (TS), significance
- What shape does the source have?
  - Extended, point-like, …
- What is the source position?
  - Source confusion,
- How bright is the source?
  - Flux, light curves, etc





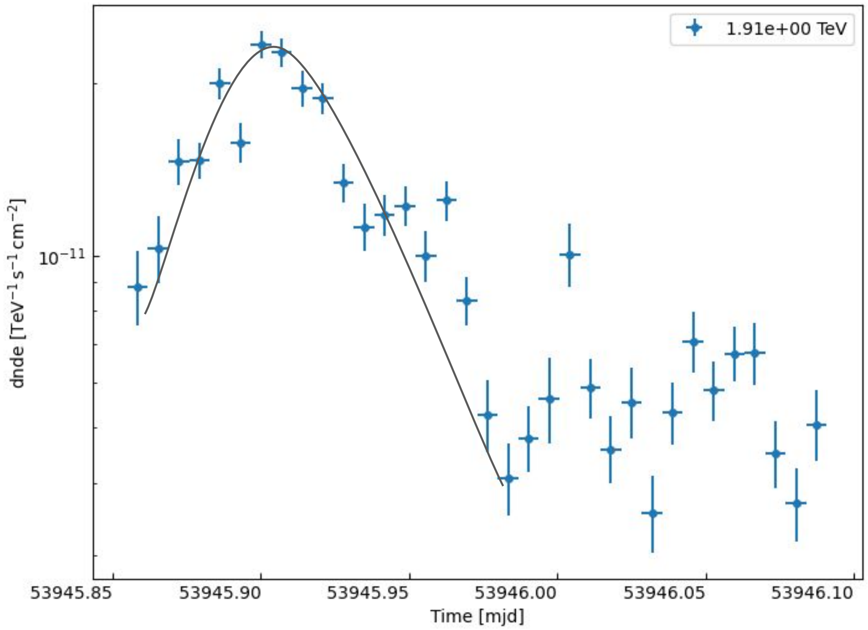
Analysis questions

- Source detection probability?
  - Test statistics (TS), significance
- What shape does the source have?
  - Extended, point-like, …
- What is the source position?
  - Source confusion,
- How bright is the source?
  - Flux, light curves, etc
- What does the energy spectrum look like?
  - Power law, log-parabola, …



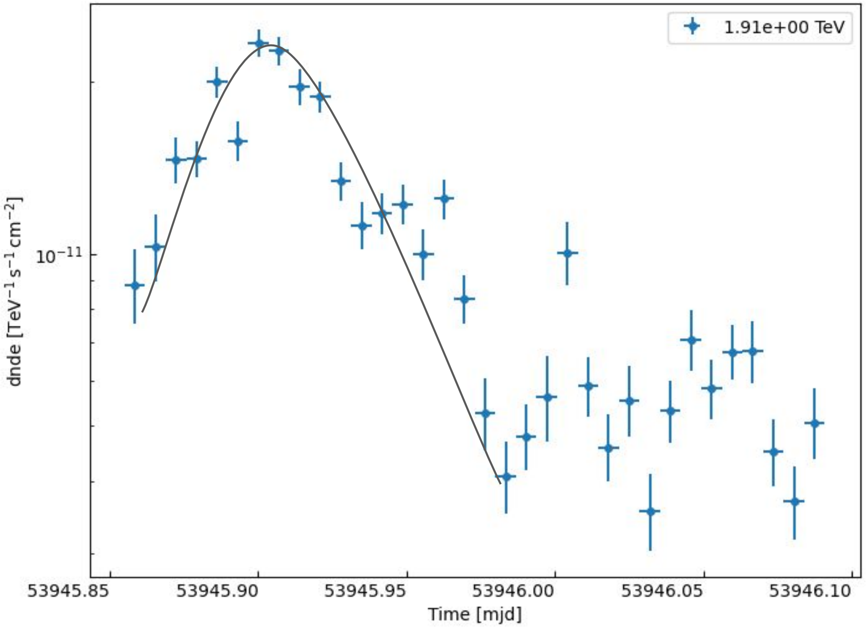
Analysis questions

- Source detection probability?
  - Test statistics (TS), significance
- What shape does the source have?
  - Extended, point-like, …
- What is the source position?
  - Source confusion,
- How bright is the source?
  - Flux, light curves, etc
- What does the energy spectrum look like?
  - Power law, log-parabola, …
- Is the emission variable?
  - Constant emission, flares, periodic, …



Analysis questions

- Source detection probability?
  - Test statistics (TS), significance
- What shape does the source have?
  - Extended, point-like, …
- What is the source position?
  - Source confusion,
- How bright is the source?
  - Flux, light curves, etc
- What does the energy spectrum look like?
  - Power law, log-parabola, …
- Is the emission variable?
  - Constant emission, flares, periodic, …



What are the statistical and systematic uncertainties on all of that?

Analysis

Do we detect a source?
What are its properties?

Source detection Hypothesis testing

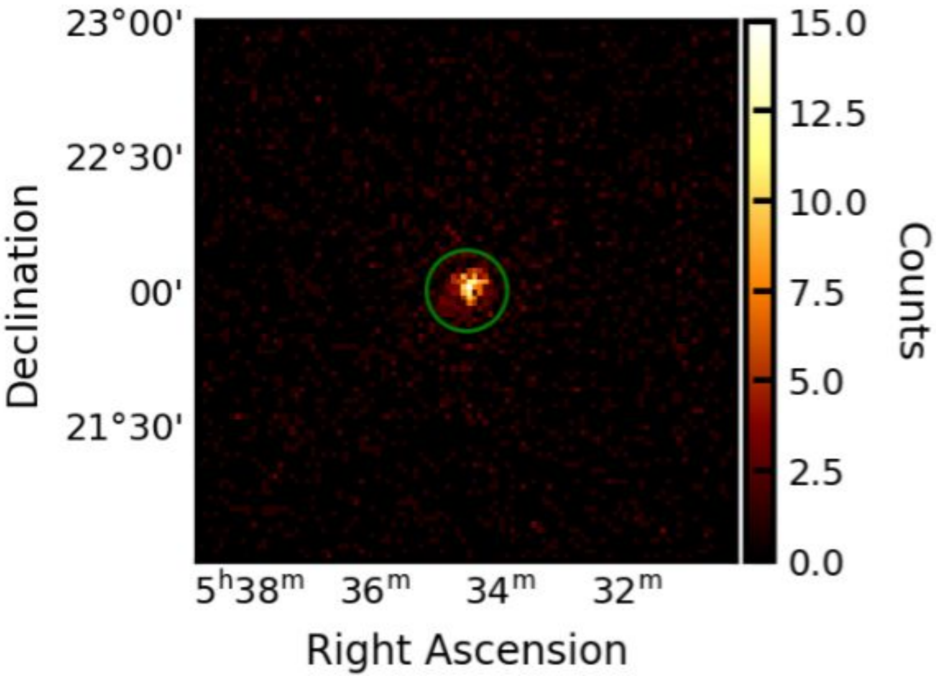
Likelihood ratio test: compare the likelihood of two hypotheses to see which one is supported by the data

- The two unknown parameters:
  - Expected number of source photons: $< N_S >$
  - Expected number of background events: $< N_B >$
- A model or excess of counts (H_1) is tested against the null hypothesis (H_0) where no source is present
  - Null hypothesis: $< N_S >= 0 \rightarrow \mathcal{L}_0 = \mathcal{L}(H_0)$
  - Alternative hypothesis: $< N_S > \neq 0 \rightarrow \mathcal{L}_1 = \mathcal{L}(H_1)$
- Utilise difference in test statistic (TS) through the likelihood ratio:

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

Wilk's Theorem

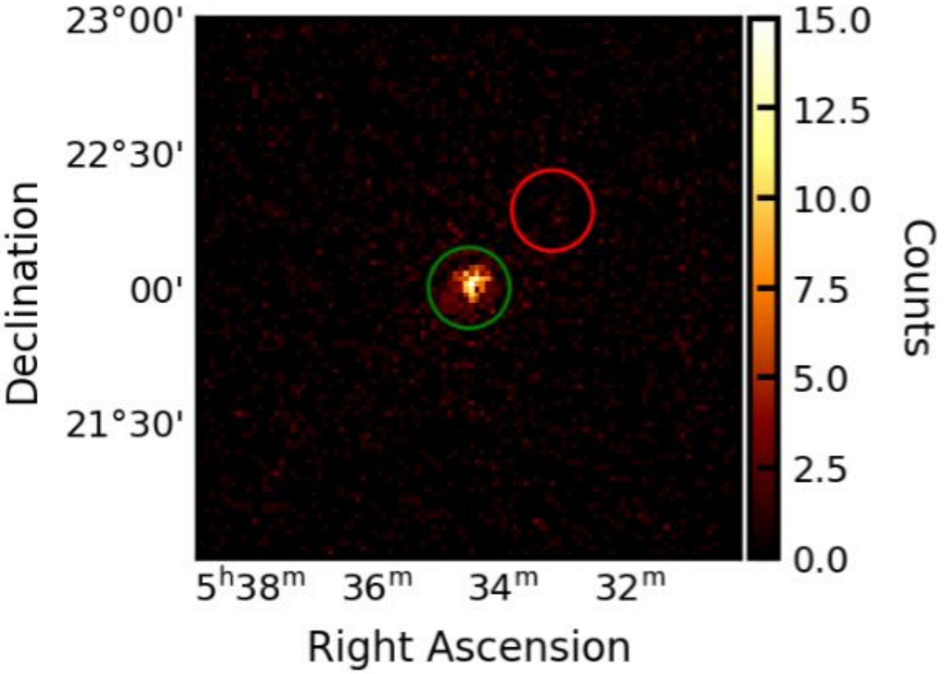
Source detection - simplified

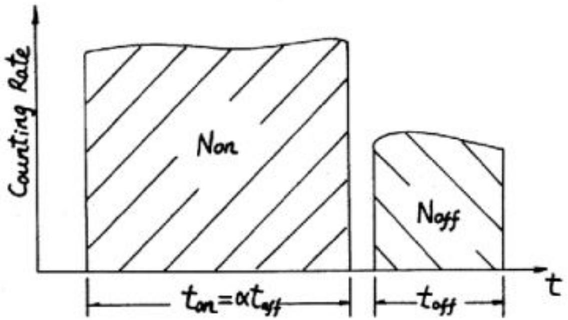


- **Signal region**: 412 counts

Slide credit: Laura Olivera-Nieto

Source detection - simplified





- **Signal region**: 412 counts
- **Background region**: 40 counts

Slide credit: Laura Olivera-Nieto

Source detection - recall hypothesis testing

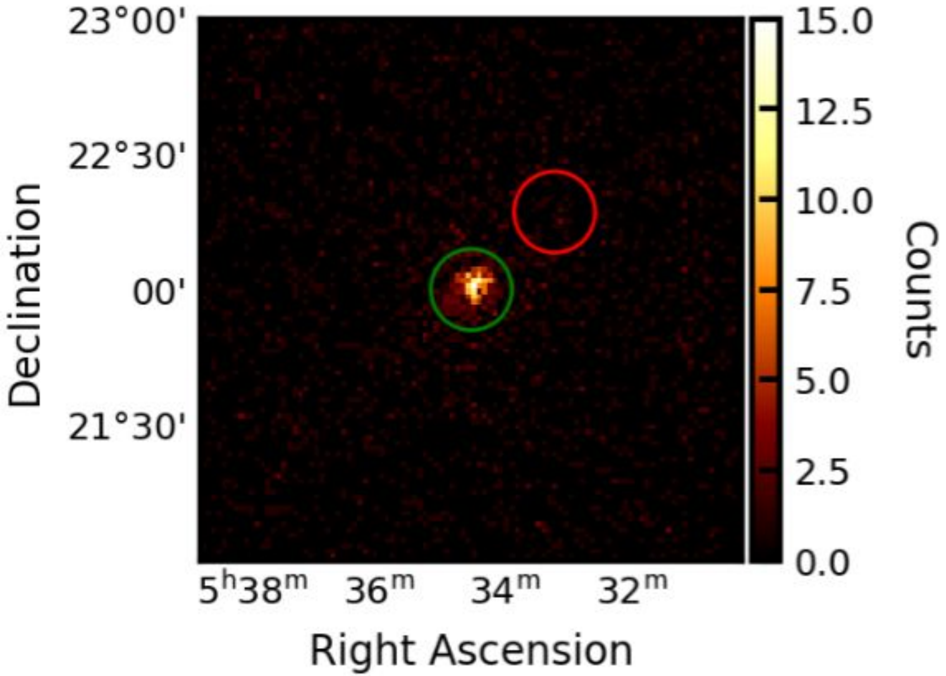
- A model or an excess of counts is tested against a null hypothesis where no source is present
- Recall: the (poisson) likelihood ratio

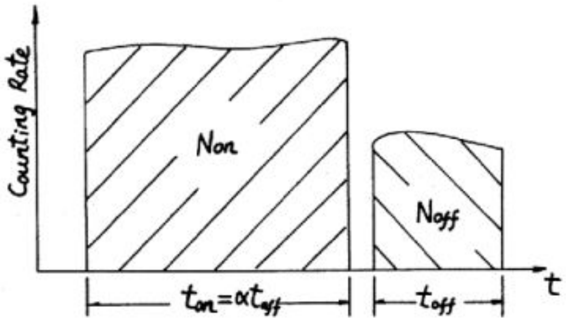
$$TS = -2\log\left(\frac{\mathcal{L}_0}{\mathcal{L}_1}\right)$$

- This test statistic is a measure of how much the source hypothesis is preferred over statistical background fluctuations for a measured excess signal
- With only 1 degree of freedom $\sigma = \sqrt{TS}$, we typically require 5σ to claim a detection

Slide credit: Laura Olivera-Nieto

Source detection - simplified





- **Signal region**: 412 counts
- **Background region**: 40 counts
- Likelihood ratio method Li&Ma (1983):

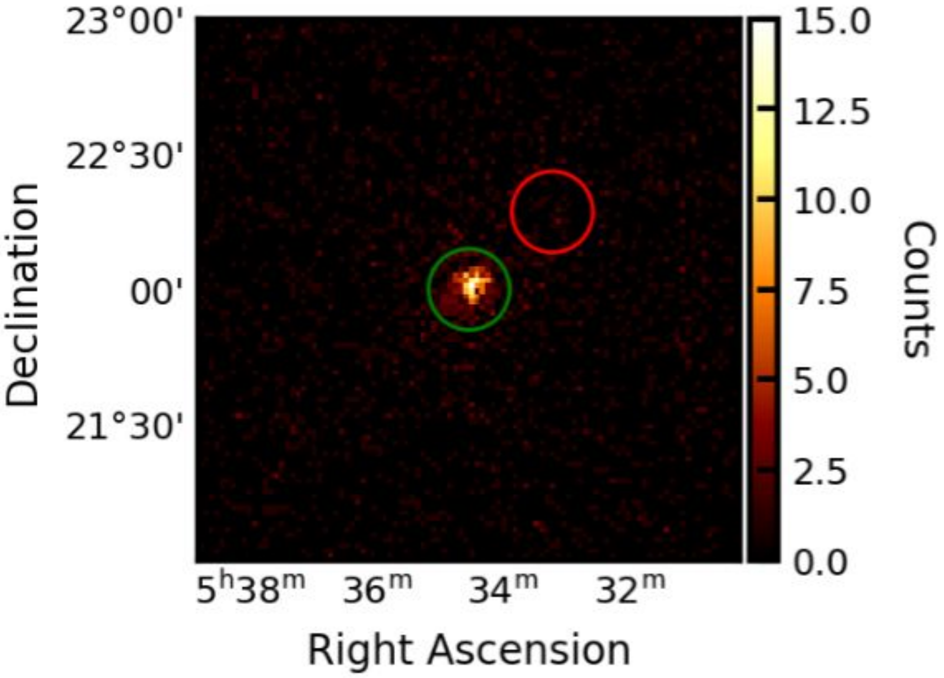
$$S = \sqrt{-2\ln\lambda}$$
$$= \sqrt{2}\left\{N_{\mathrm{on}}\ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{\mathrm{on}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right] + N_{\mathrm{off}}\ln\left[(1+\alpha)\left(\frac{N_{\mathrm{off}}}{N_{\mathrm{on}}+N_{\mathrm{off}}}\right)\right]\right\}^{1/2}$$

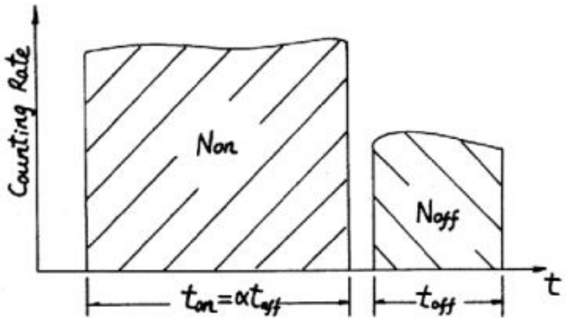
- Significance ~ 35σ

In reality one would use e.g. "reflected regions background", as mentioned previously

Slide credit: Laura Olivera-Nieto

Source detection - simplified





- **Signal region**: 412 counts
- **Background region**: 40 counts
- Likelihood ratio method Li&Ma (1983):

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

- Significance ~ 35σ

Source detection! ✅

In reality one would use e.g. "reflected regions background", as mentioned previously

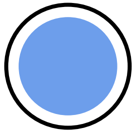
Slide credit: Laura Olivera-Nieto

Analysis

Do we detect a source?
What are its properties?

Spectral analysis

- We detected a source and now we want to calculate the flux
- Define a source region where we want to extract the spectrum from
- Point source is very simple as we can just take the optimal excess
- Extended sources can start to become a little more complex
  - Complex (i.e. non-circular) regions and more importantly, the instrument response is non-uniform over the selected region



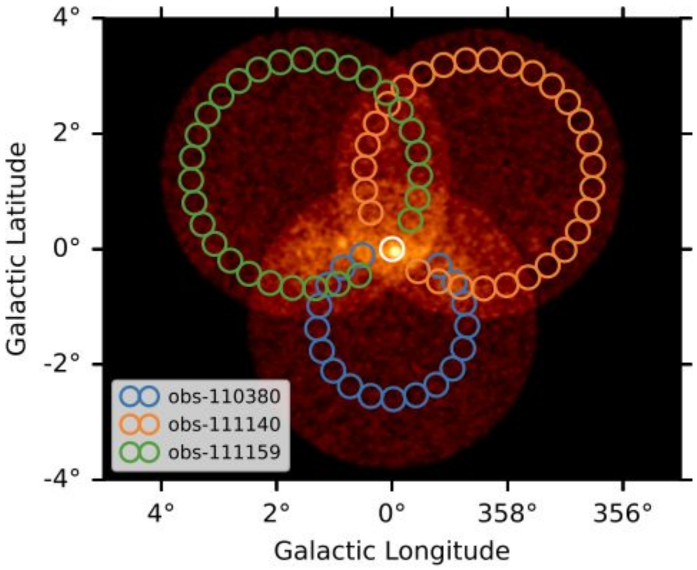
Point source

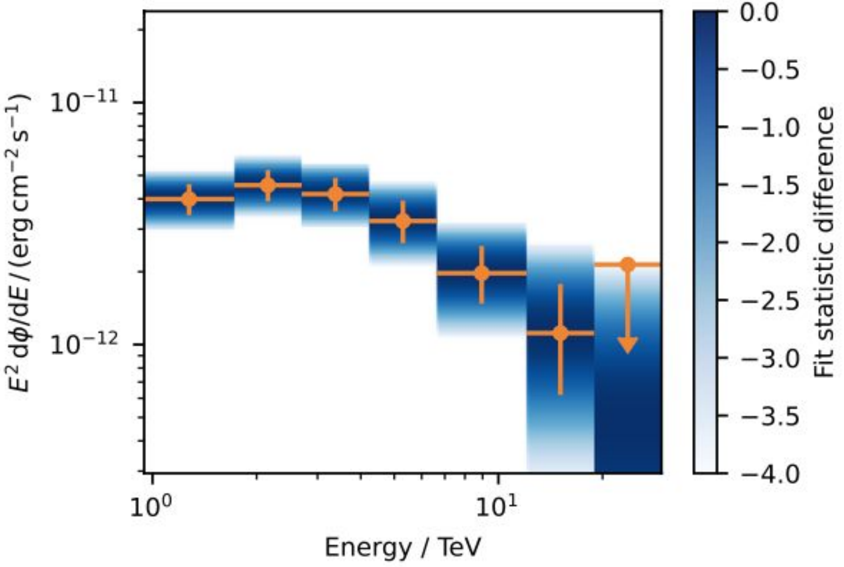
Full enclosure

Internal region

Spectral analysis Classical method

- Utilise the reflected regions method as mentioned previously
- Along with the forward folding method → maximum likelihood
- Utilise simulated CTA data – likelihood per energy bin is shown by the blue colored band

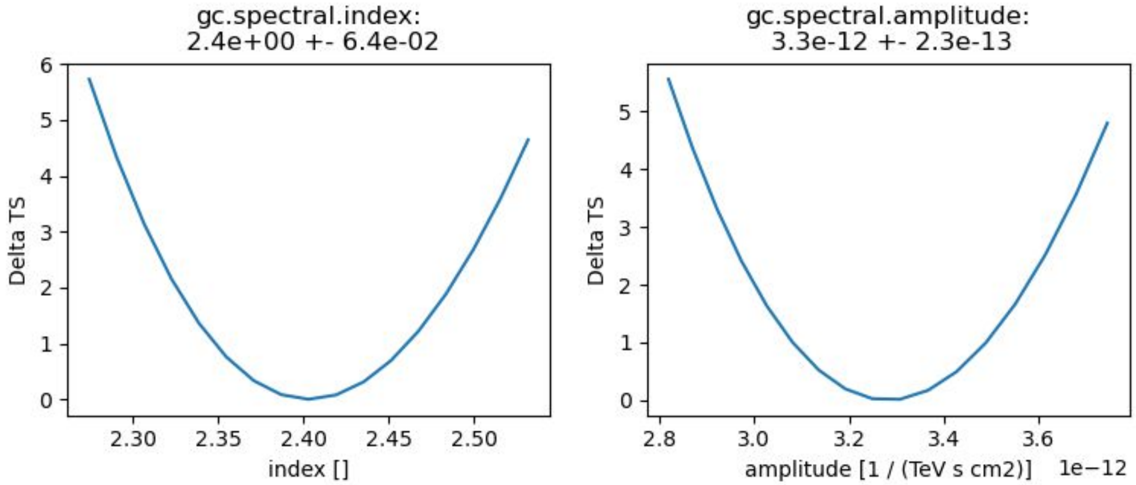




Spectral analysis Classical method

- Maximum likelihood = minimum test statistic
- Is the minimum well defined? **YES!** ✅

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



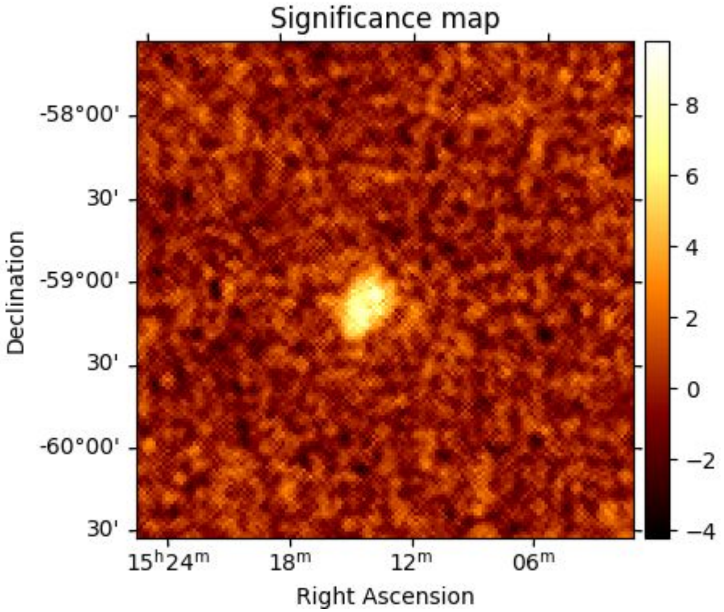
See this tutorial

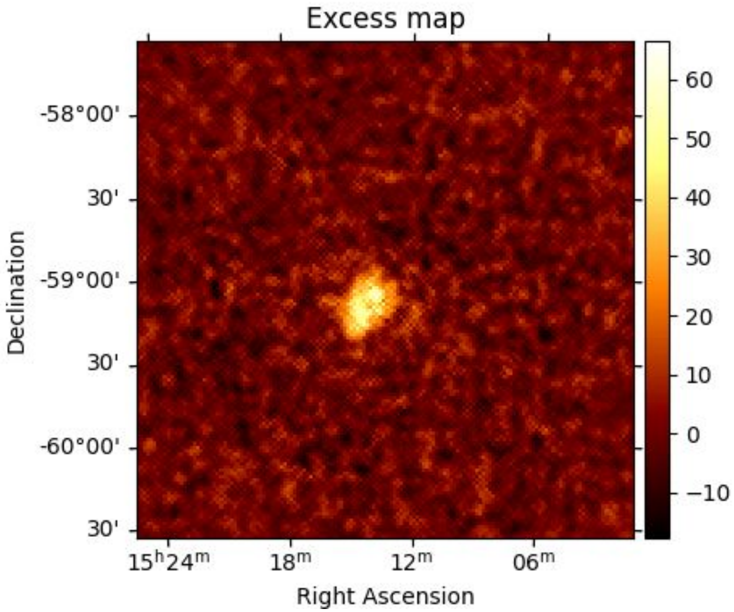
- The same methodology is utilised for an extended source with one key difference:
  - The values of the IRFs are averaged over the entire region
  - See this tutorial

Spatial analysis

Classical 2D method

- Utilise the ring background method as explained previously
- Select OFF events from an annulus about the ON region
- Requires good knowledge of 2D acceptance!

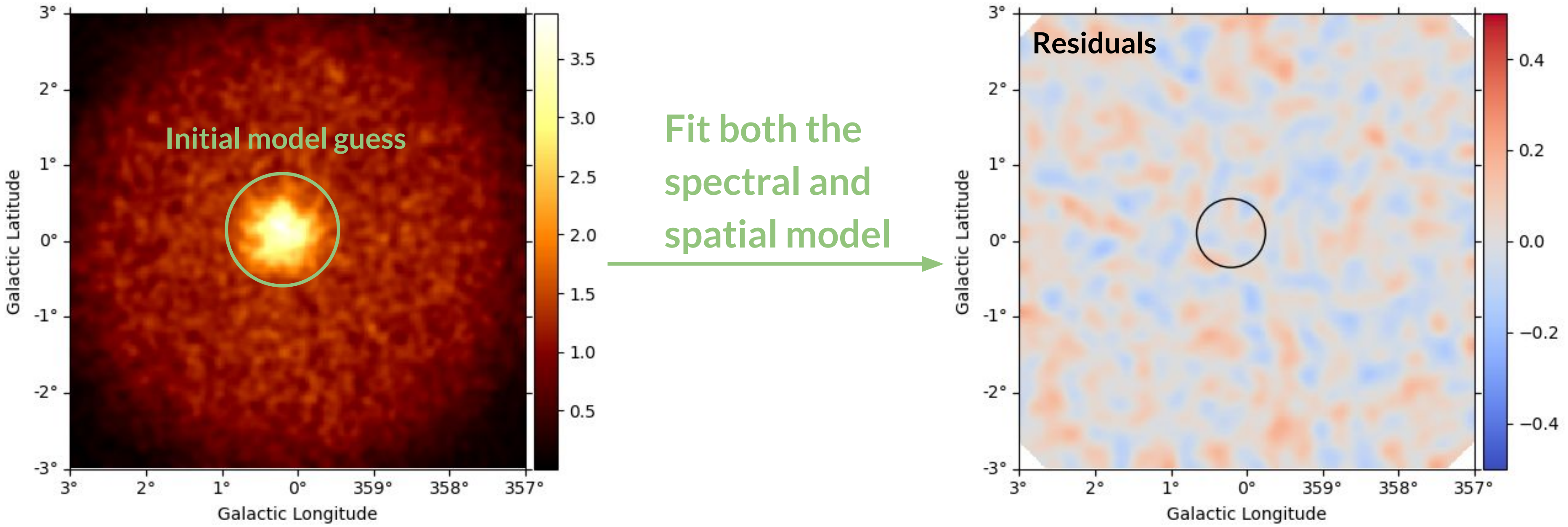




Spectral and Spatial analysis

✨The new 3D method! ✨

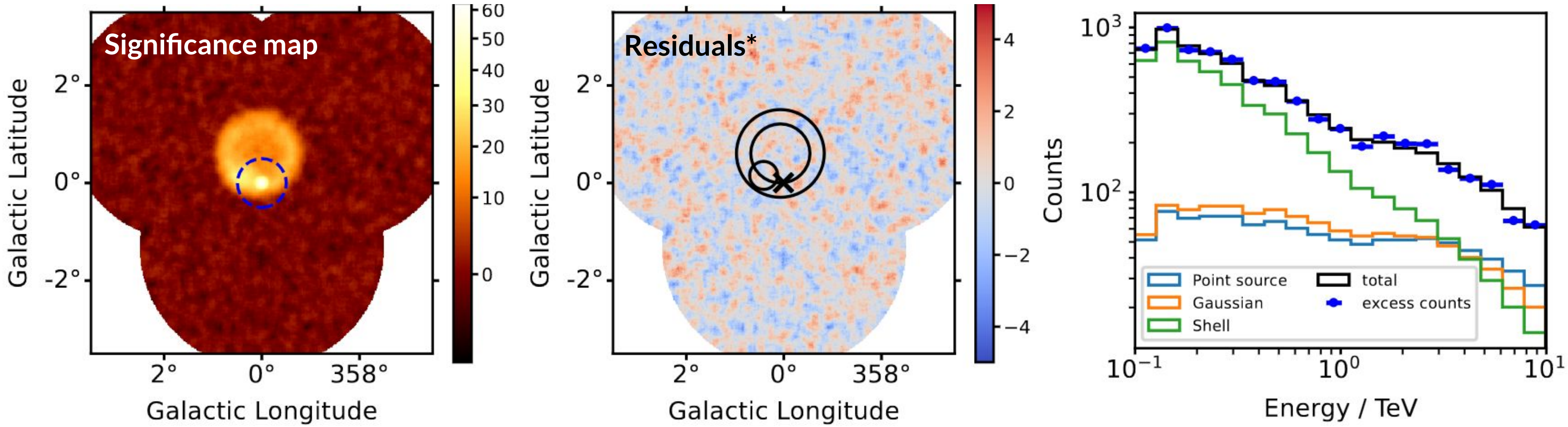
- Utilise the FoV background method which allows for a 3D analysis
- In one analysis chain we can fit both the spectra and the morphology of a source



✨ Power of the 3D method ✨

It is possible to disentangle contributions of overlapping sources!

- For example:
  - Point source with power law spectrum
  - Gaussian source with log-parabola spectrum
  - Shell with power law spectrum



*Significance map after subtracting the best-fit model for each of the sources

Contribution of each source model to the circular region of radius 0.5° drawn in the left image, together with the excess counts inside that region.

An Introduction To Gammapy

Getting started: installation

Quickstart installation with conda/mamba

```
curl -O https://gammapy.org/download/install/gammapy-1.2-environment.yml
conda env create -f gammapy-1.2-environment.yml
conda activate gammapy-1.2
```

See Getting started

Installation with pip

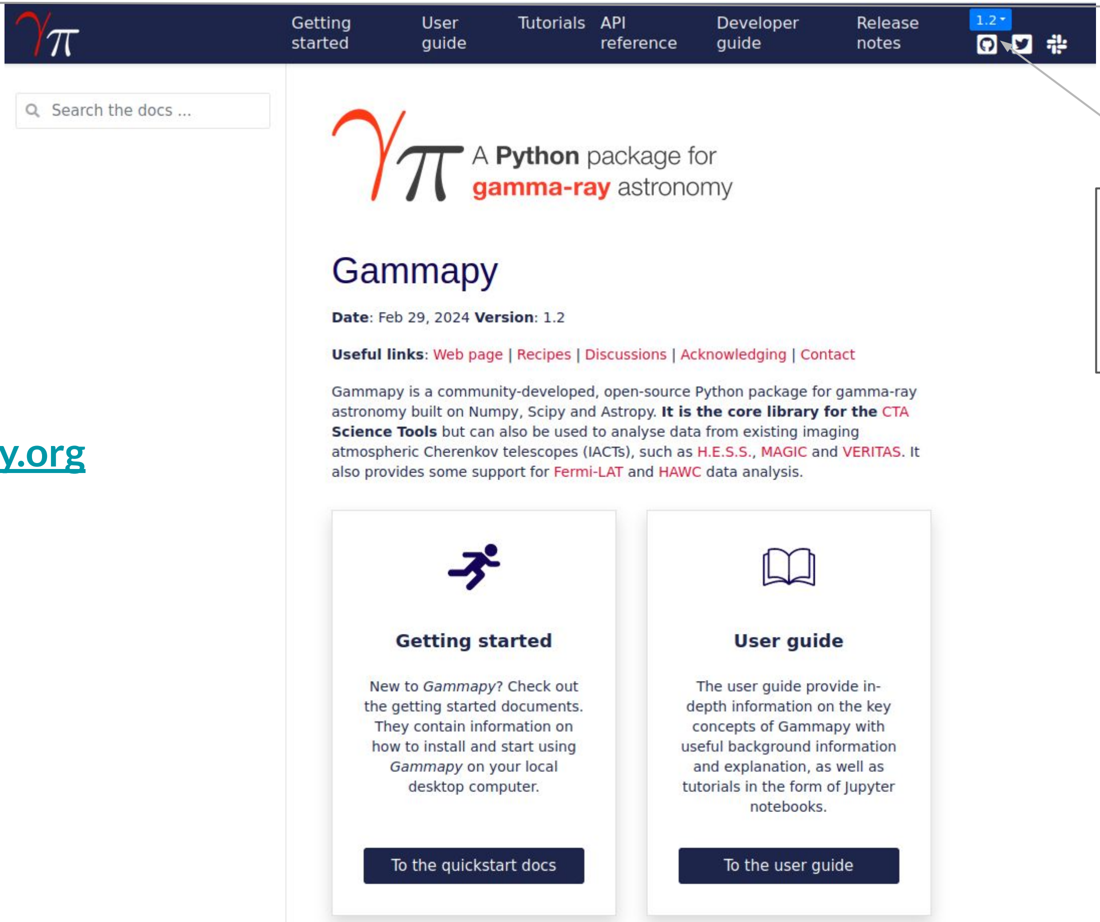
```
pip -m install gammapy[all]
```

install all dependencies

For this workshop, the other instructions are found on the GitHub

Getting started: documentation

See docs.gammapy.org



Switch between versions

Getting helping, reporting issues

How to provide feedback/get help

- #help channel on gammapy.slack
- #gammapy channel on hesschat.slack (if you are a HESS member)
- GitHub discussion, in particular the help category

How to report issues and bugs or request a new feature

- GitHub issues page (requires a GitHub account)

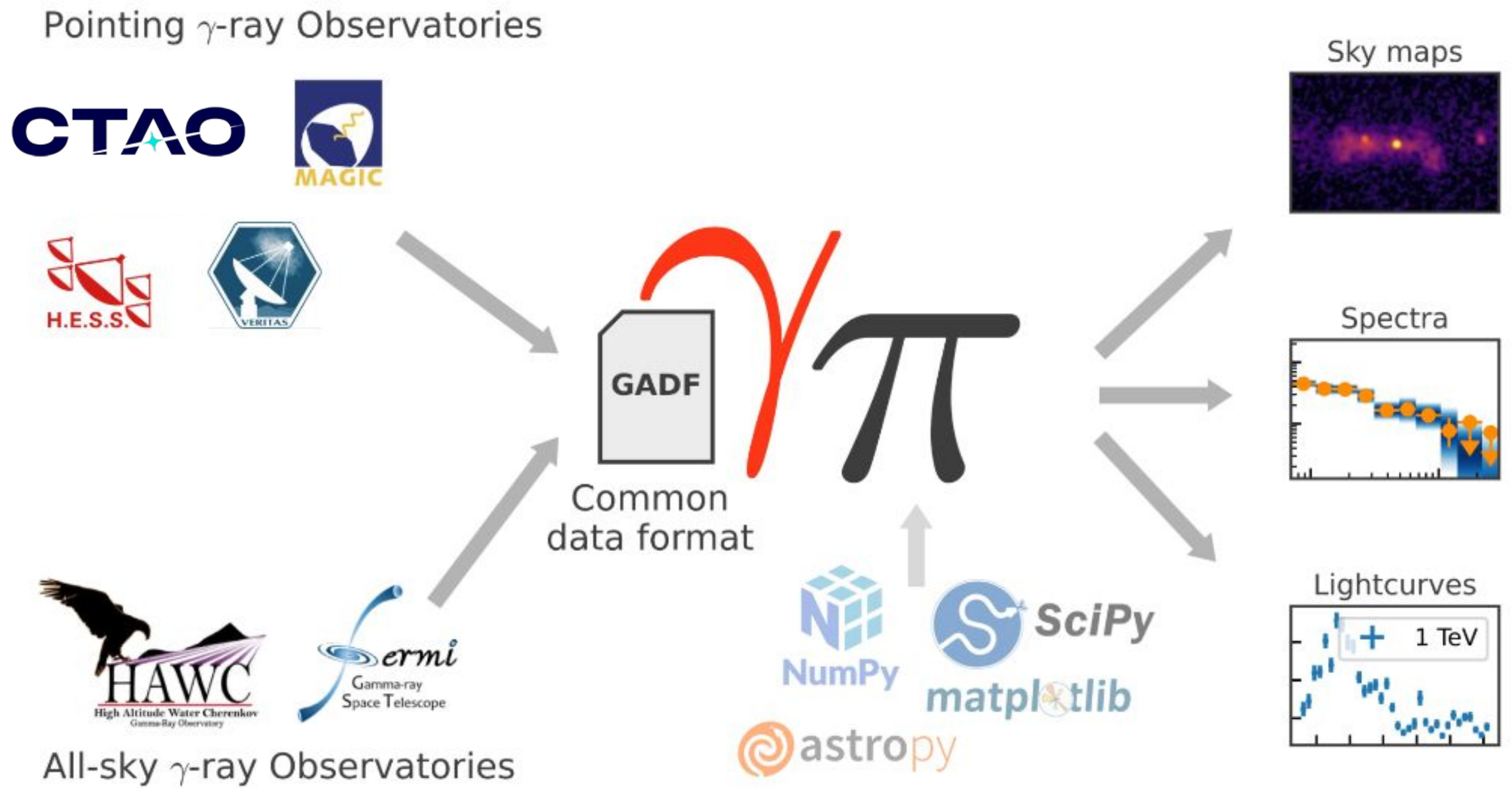
Gammapy & CTAO

- v1.2 was released February 2024

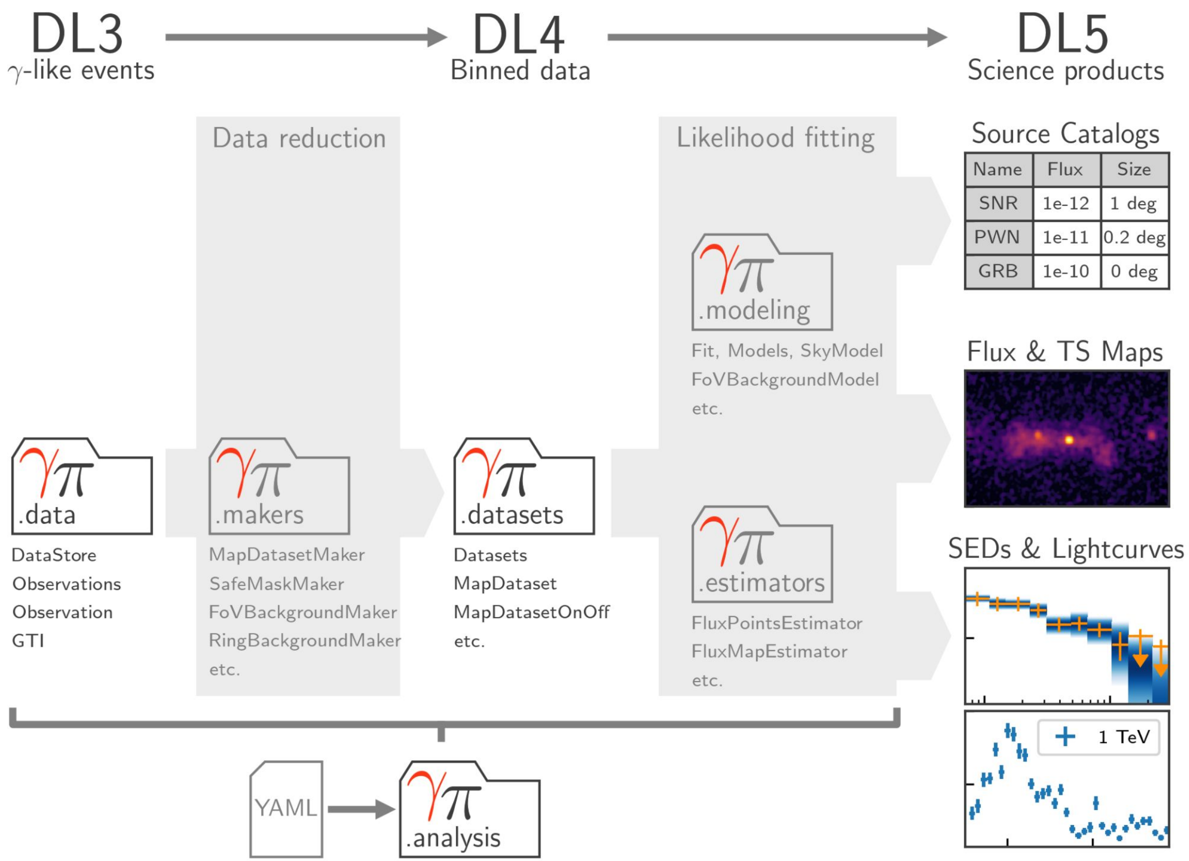
Gammapy team involved in SDC effort

- Event types are supported
  - Distributed as two different data stores
  - Stacked or joint analyses are possible
- Missing features for SDC have been added:
  - Time dependent spectral models in v1.1
  - Metadata containers to support CTAO data model
- SDC will be important to prepare first SAT release

Gammapy overview



Gammapy overview

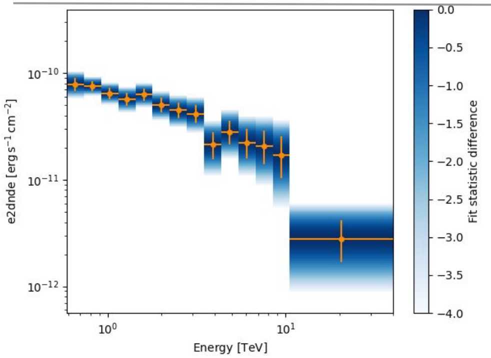


2-step analysis procedure

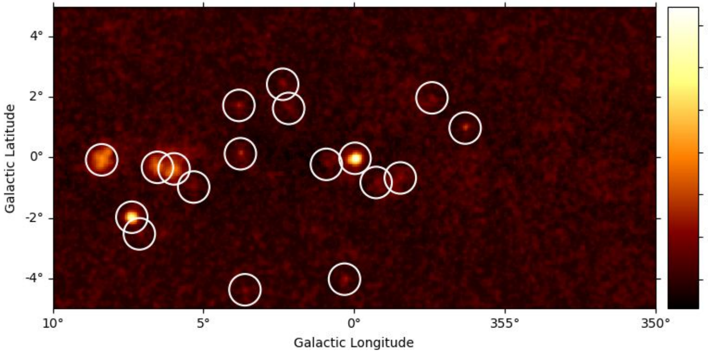
- Data reduction (DL3 to DL4)
- Data modeling/fitting (DL4 to DL5)

Analysis use cases

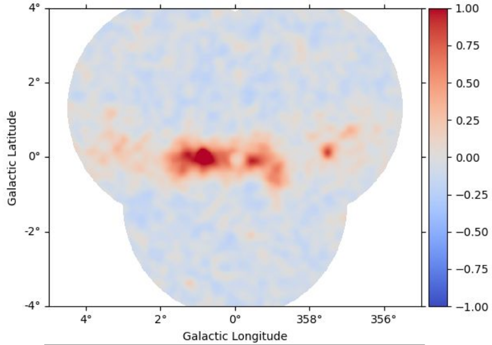
Links



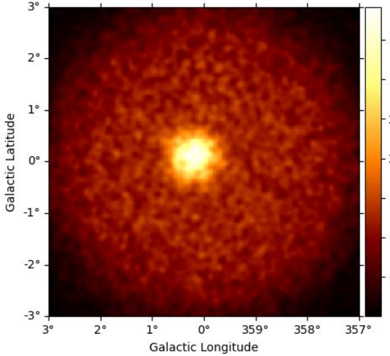
[1D spectral analysis]



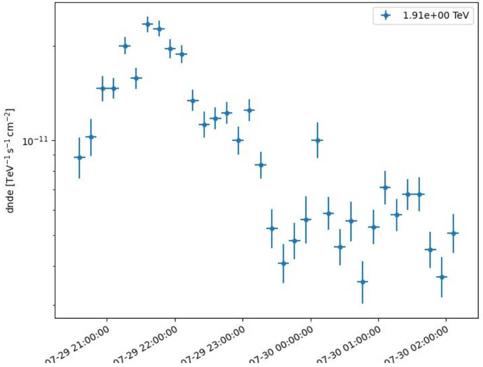
[Source detection]



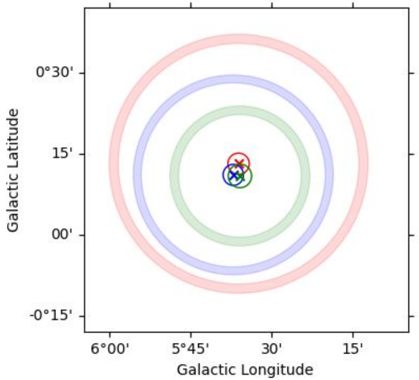
[3D analysis]



[Observation simulation]



[Light curve variability]



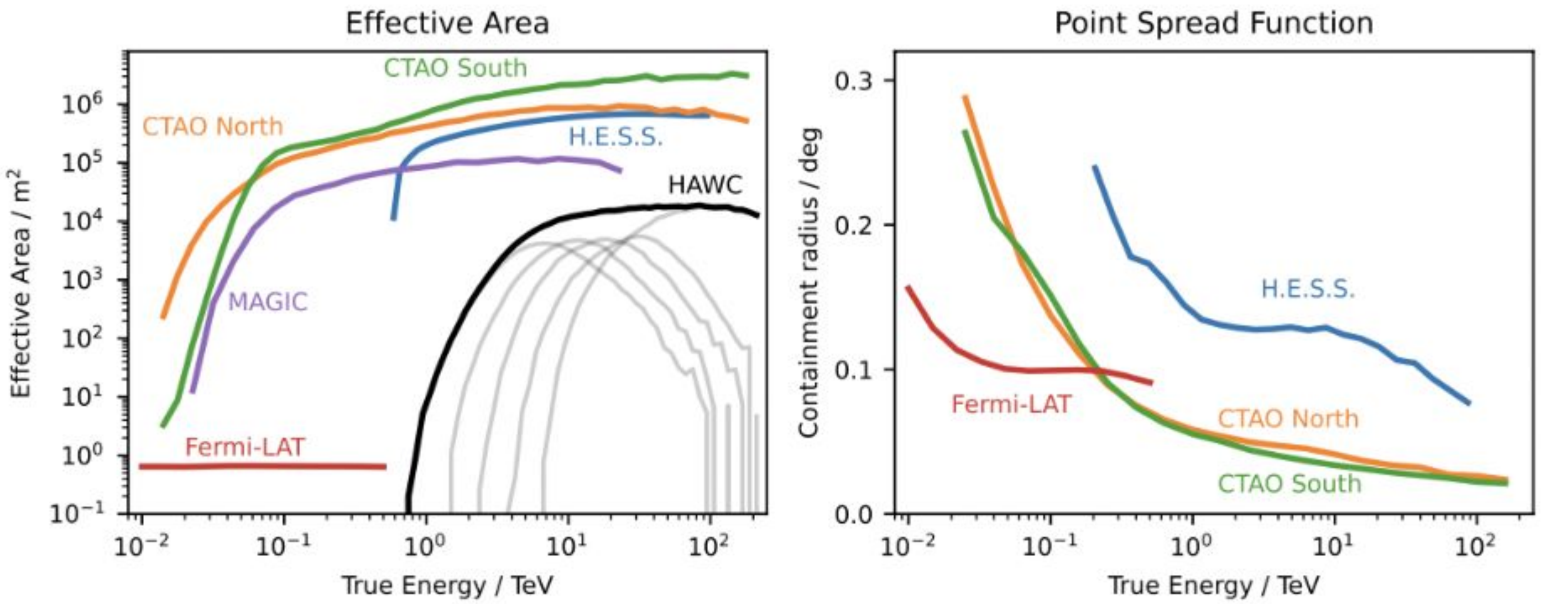
[Energy dependent morphology]

Data reduction

1. Select and retrieve relevant observations from the data store
2. Define the reduced dataset geometry
   a. Is the analysis 1D (spectral only) or 3D?
   b. Define target binning and projection
3. Initialise the data reduction methods (makers)
   a. Data and IRF projection
   b. Background estimation
   c. Safe Mask determination
4. Loop over selected observations
   a. Apply makers to produce reduced datasets
   b. Optionally combine them (stacking)

IRF visualisation

- Users can read and visualise IRFs which have been provided by various instruments and telescopes including CTA
  - **We will see how to extract IRFs in this workshop**

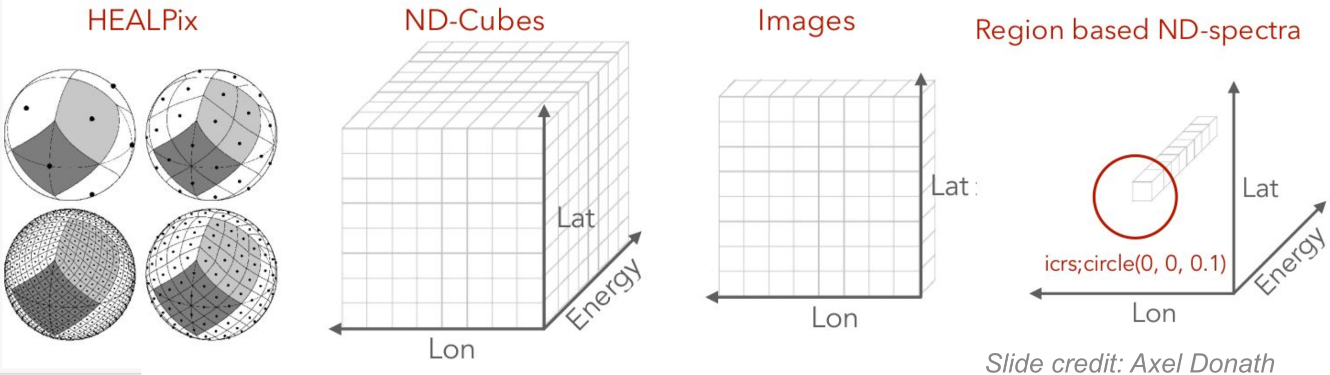


https://github.com/gammapy/gammapy-v1.0-paper/blob/main/src/figures/irfs.py

Gammapy Maps

```python
from gammapy.maps import Map, MapAxis
from astropy.coordinates import SkyCoord
from astropy import units as u

skydir = SkyCoord("0d", "5d", frame="galactic")

energy_axis = MapAxis.from_energy_bounds(
    energy_min="1 TeV", energy_max="10 TeV", nbin=10
)

# Create a WCS Map
m_wcs = Map.create(
    binsz=0.1,
    map_type="wcs",
    skydir=skydir,
    width=[10.0, 8.0] * u.deg,
    axes=[energy_axis])


# Create a HEALPix Map
m_hpx = Map.create(
    binsz=0.1,
    map_type="hpx",
    skydir=skydir,
    axes=[energy_axis]
)

# Create a region map
region = "galactic;circle(0, 5, 1)"
m_region = Map.create(
    region=region,
    map_type="region",
    axes=[energy_axis]
)
```

- **N-dimensional coordinate aware data structures** for storing gamma-ray data, with arbitrary number of non-spatial dimensions, such as energy, time or a label for the data class
- **Uniform API for WCS, HEALPix and region** based pixelization schemes



Slide credit: Axel Donath

Gammapy Makers

```python
import astropy.units as u

from gammapy.data import DataStore
from gammapy.datasets import MapDataset
from gammapy.makers import (
    FoVBackgroundMaker,
    MapDatasetMaker,
    SafeMaskMaker
)
from gammapy.maps import MapAxis, WcsGeom

data_store = DataStore.from_dir(
    base_dir="$GAMMAPY_DATA/hess-dl3-dr1"
)

obs = data_store.obs(23523)

energy_axis = MapAxis.from_energy_bounds(
    energy_min="1 TeV",
    energy_max="10 TeV",
    nbin=6,
)

geom = WcsGeom.create(
    skydir=(83.633, 22.014),
    width=(4, 3) * u.deg,
    axes=[energy_axis],
    binsz=0.02 * u.deg,
)

empty = MapDataset.create(geom=geom)
```

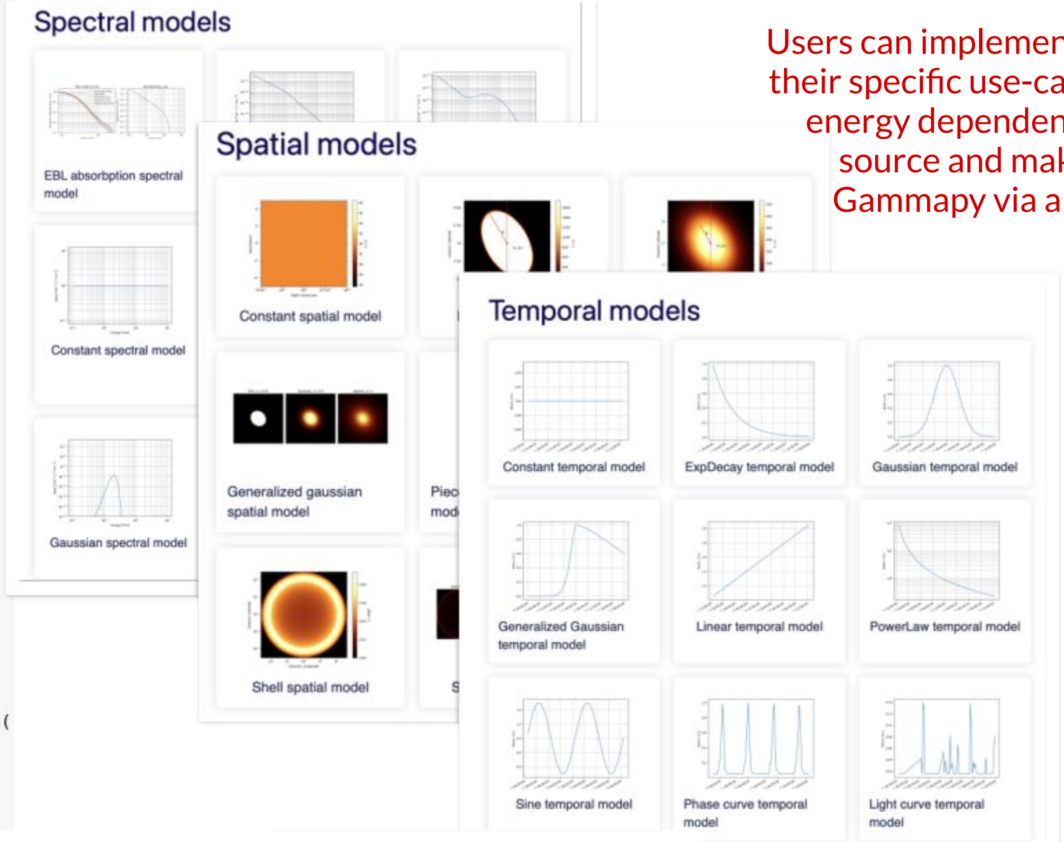
```python
maker = MapDatasetMaker()

mask_maker = SafeMaskMaker(
    methods=["offset-max", "aeff-default"],
    offset_max="2.0 deg",
)

bkg_maker = FoVBackgroundMaker(
    method="scale",
)

dataset = maker.run(empty, observation=obs)
dataset = bkg_maker.run(dataset, observation=obs)
dataset = mask_maker.run(dataset, observation=obs)
dataset.peek()
```

- Maker are configurable, stateless objects the represent an algorithm/data reduction step
- Data represented by a MapDataset is passed between the makers is modified and holds the state
- This allow users to define data reduction chains, without access to data and also implement custom steps
- The whole process can be also defined from YAML based configuration files and executed from a “high level” Analysis class

Slide credit: Axel Donath

Gammapy Models

```
from astropy import units as u
from gammapy.modeling.models import (
    ConstantTemporalModel,
    EBLAbsorptionNormSpectralModel,
    PointSpatialModel,
    PowerLawSpectralModel,
    SkyModel,
)

# define a spectral model
pwl = PowerLawSpectralModel(
    amplitude="1e-12 TeV-1 cm-2 s-1", index=2.3
)

# define a spatial model
point = PointSpatialModel(
    lon_0="45.6 deg",
    lat_0="3.2 deg",
    frame="galactic"
)


# define a temporal model
constant = ConstantTemporalModel()

# combine all components
model = SkyModel(
    spectral_model=pwl,
    spatial_model=point,
    temporal_model=constant,
    name="my-model",
)
print(model)

ebl = EBLAbsorptionNormSpectralModel.read_builtin(
    reference="dominguez", redshift=0.5
)

absorbed = pwl * ebl
absorbed.plot(energy_bounds=(0.1, 100) * u.TeV)
```



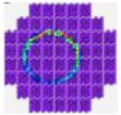
$$f_{Src} = f_{Spectral}(E) \cdot f_{Spatial}(E, l, b) \cdot f_{Temporal}(t)$$

Users can implement **custom models** for their specific use-case, such as modeling energy dependent morphology of a source and make it available to Gammapy via a **registry system**

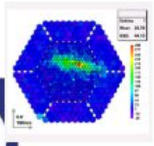
Slide credit: Axel Donath

Joint analysis?

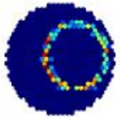
H.E.S.S.



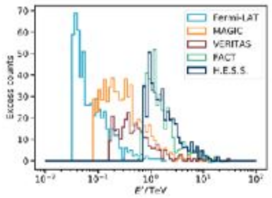
MAGIC

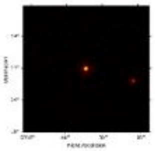


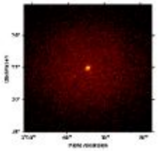
VERITAS

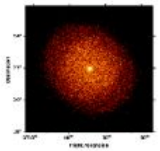


Dataset	T_{obs}	E_{min} (TeV)	E_{max} (TeV)	N_{on}	N_{bkg}	R_{on} (deg)
Fermi-LAT	~7 yr	0.03	2	578	1.2	0.30
MAGIC	0.66 h	0.08	30	784	129.9	0.14
VERITAS	0.67 h	0.16	30	289	13.7	0.10
FACT	10.33 h	0.45	30	691	272.8	0.17
H.E.S.S.	1.87 h	0.71	30	459	27.5	0.11









Raw data

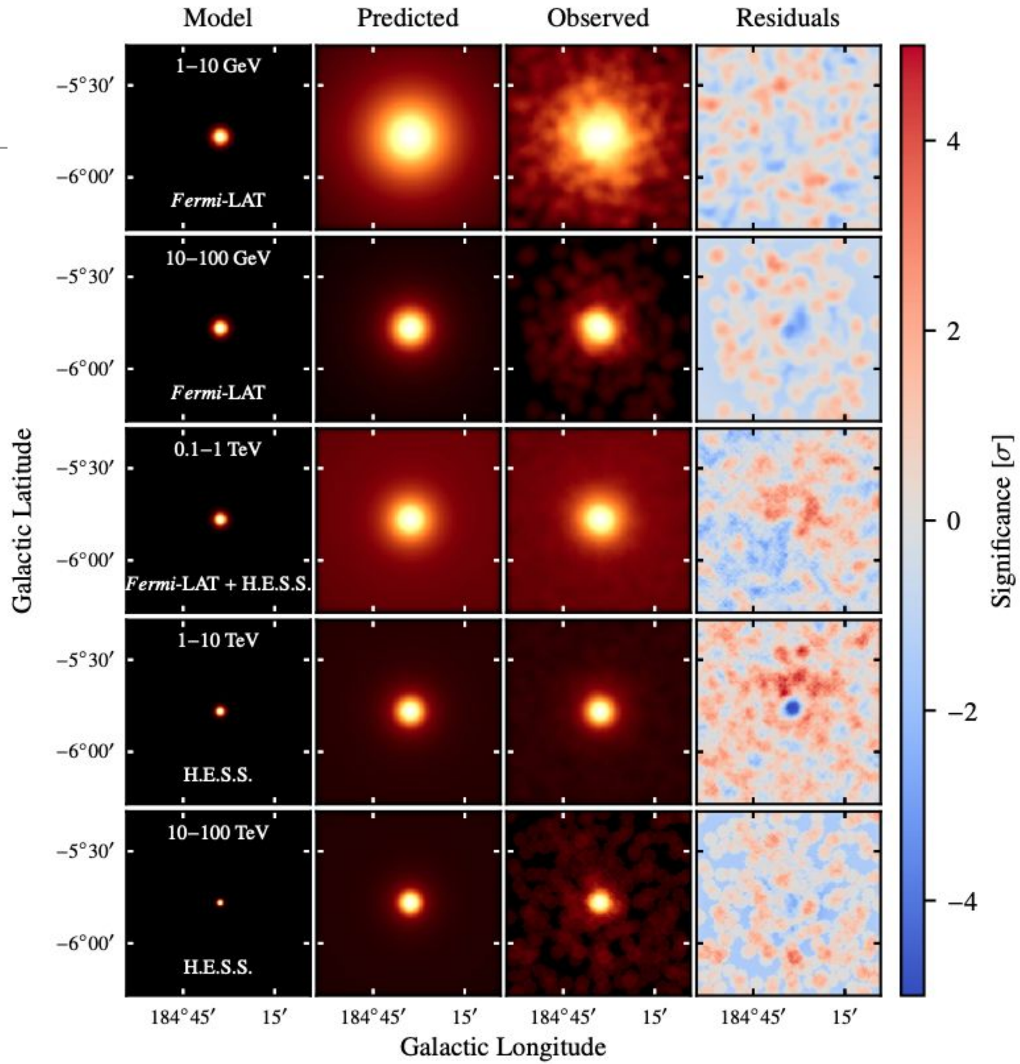
DL3 GADF

High level science products

Joint analysis

Constrain extensions

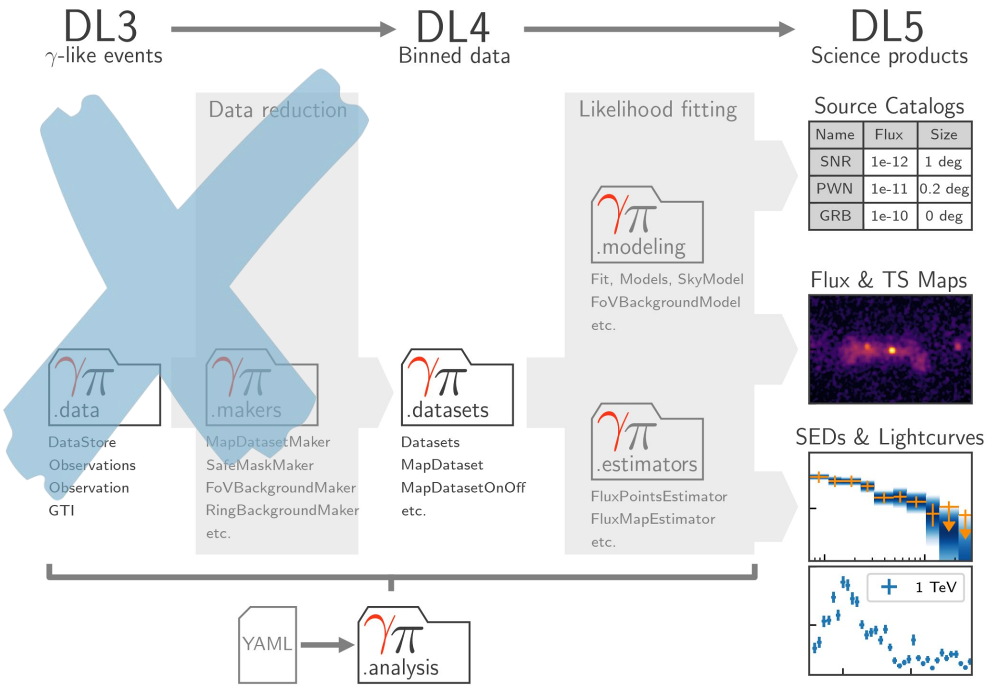
- Joint Fermi-LAT H.E.S.S. analysis used to constrain the extension of the Crab Nebula
- Probe structures to understand the underlying mechanisms



Aharonian et al., 2024

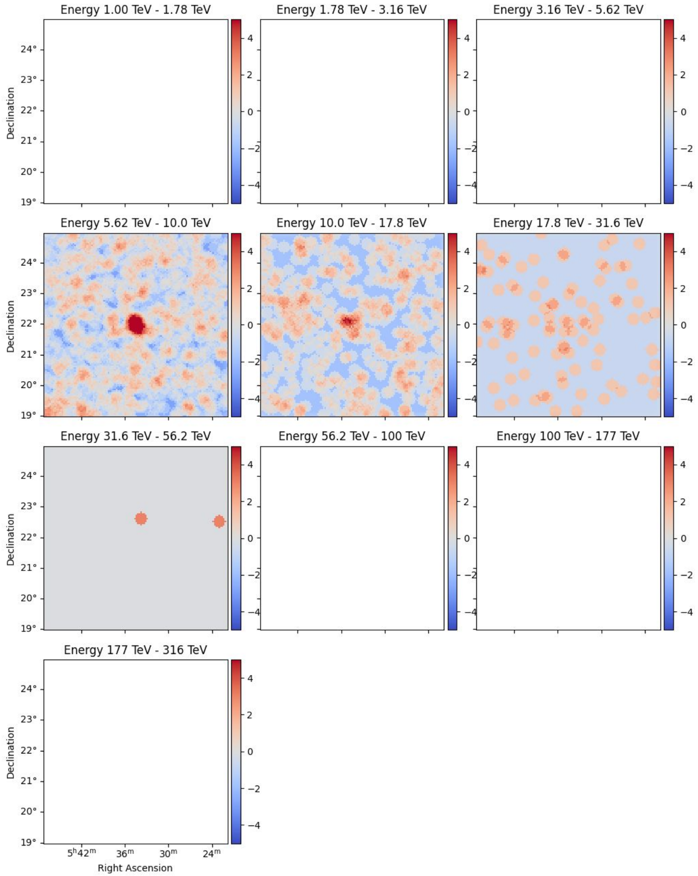
Fermi-LAT with gammapy

- Analysis starts from DL4 data levels:
    - After binning and reproduction
- Once you have a DL4 product "Dataset", modelling and fitting proceeds as before
- Bonus: Simulating datasets
- Note: Fermi-LAT analysis is always 3D



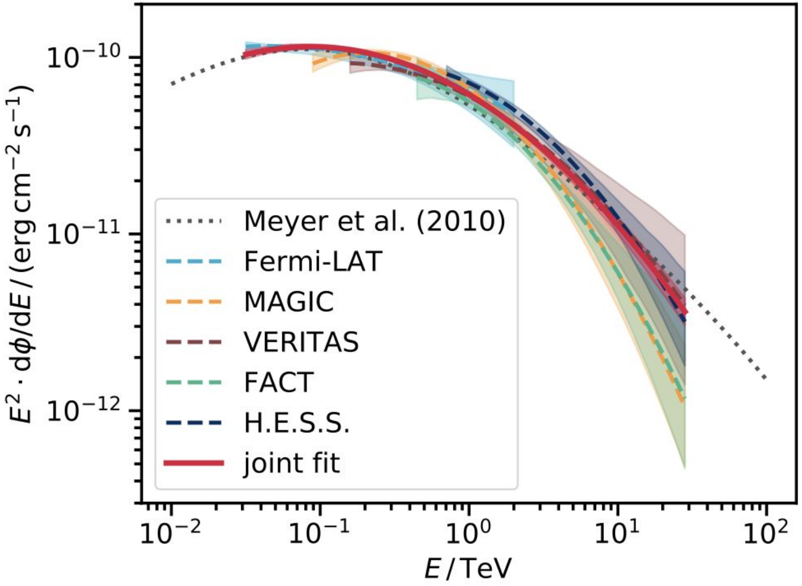
HAWC with gammapy

- Events: DL3 level
- IRFs: DL4 level
- Joint analysis for different fHit bins
- Background and exposure calculated per transit for a source
- Correct by the number of transits per source computed from the GTIs

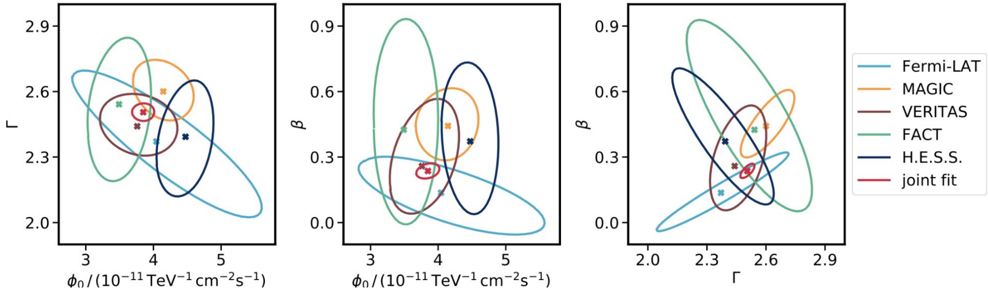


Multi-instrument analysis

- Spectral fit combining different types of data
  - Fermi-LAT DL4 data - full 3D analysis
    7yrs of data
  - MAGIC DL3 data - point like 1D spectral analysis 40 mins of data
  - VERITAS DL3 data - point like 1D spectral analysis
    40 mins of data
  - FACT - point like 1D spectral analysis
    10.3 hrs of data
  - H.E.S.S. - full containment 3D analysis
    2 hrs of data



Better constrained parameters



Nigro et al., 2019

Thanks for your attention

```
from gammapy import song
song(karaoke=True)
```