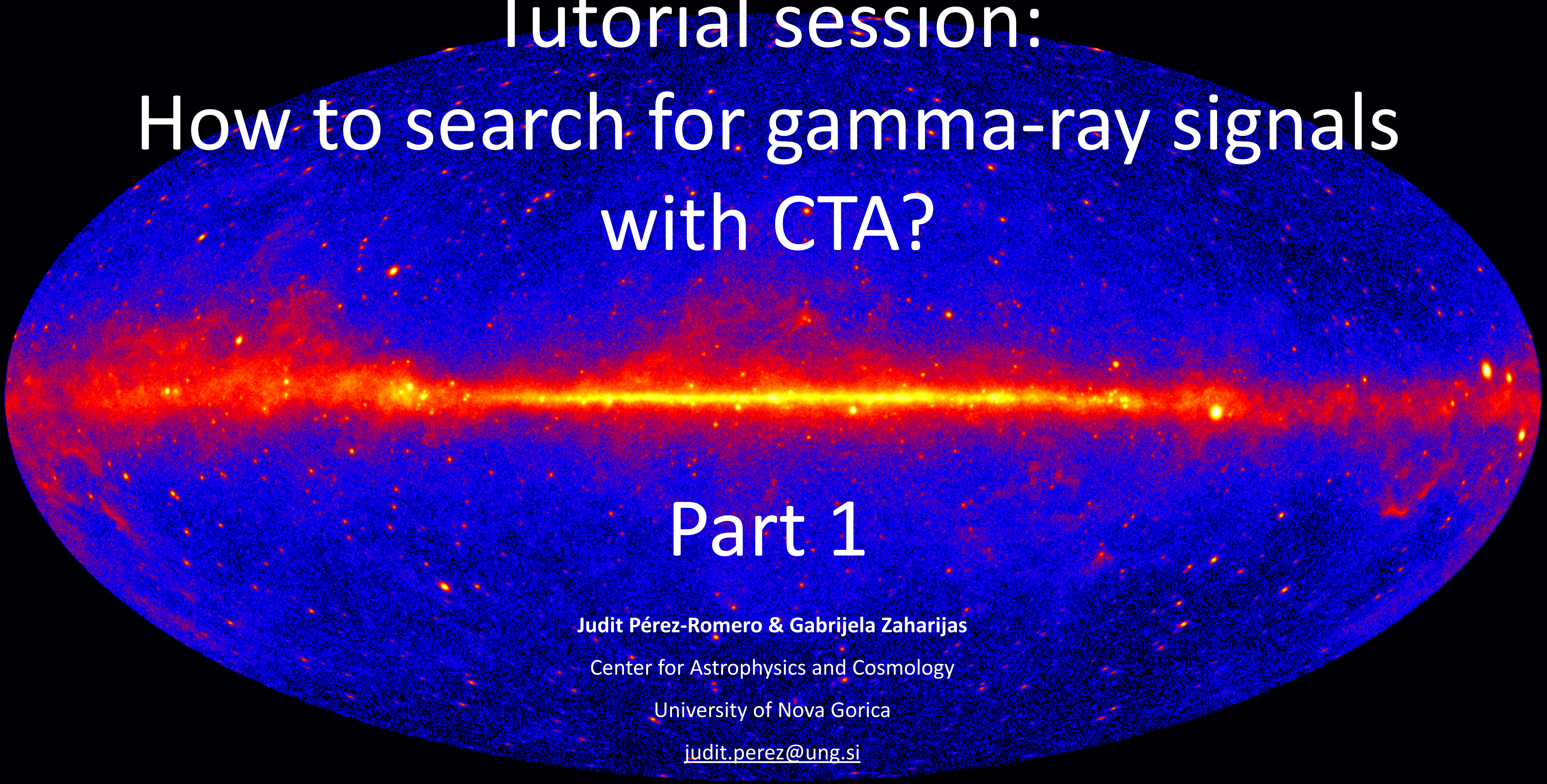
Tutorial session: How to search for gamma-ray signals with CTA?

Part 1



Judit Pérez-Romero & Gabrijela Zaharijas

Center for Astrophysics and Cosmology

University of Nova Gorica

judit.perez@ung.si

Image Credits: Fermi-LAT Collaboration

HOW TO SEARCH FOR GAMMA-RAY SIGNALS WITH CTA

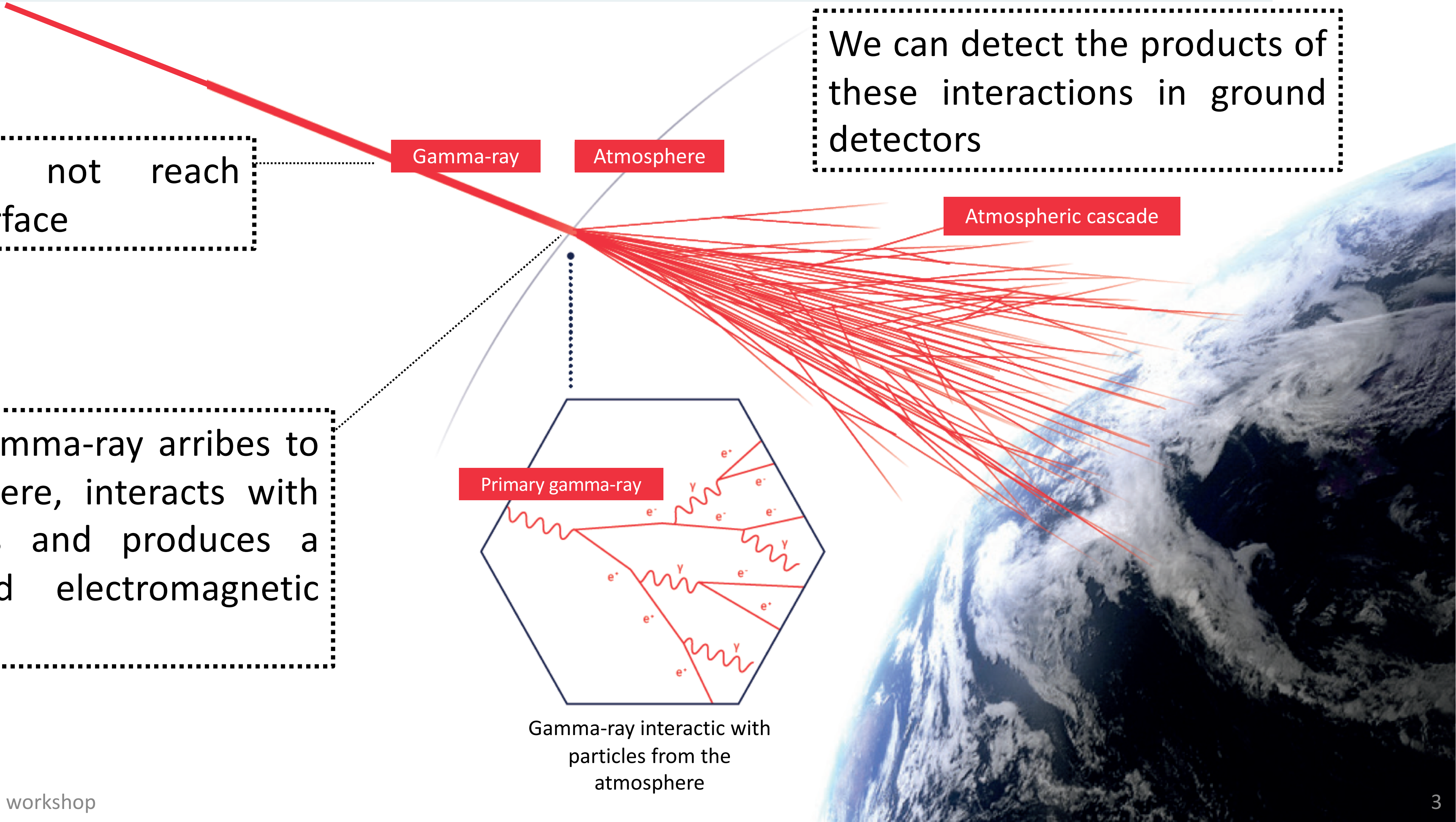
- We will work in Google collab environment, for this you need:
  - Google account
  - Preferably Google Chrome
  - Download "cta-prod5-zenodo-fitsonly-v0.1.zip" from https://zenodo.org/record/5499840#.YUya5WYzbUI
  - Upload the file CTA-Performance-prod5-v0.1-South-20deg.FITS

https://colab.research.google.com/drive/185MH2XjbCmMfvXLp5PXkqsoR5-nDkvhj?usp=sharing

GROUND BASED GAMMA-RAY DETECTION

Gamma-rays do not reach directly Earth's surface

We can detect the products of these interactions in ground detectors



Gamma-ray interactic with particles from the atmosphere

When the gamma-ray arribes to the atmosphere, interacts with the particles and produces a chain, called electromagnetic cascades

GROUND BASED GAMMA-RAY DETECTION

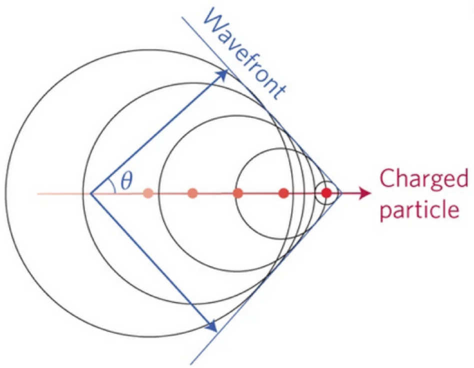
The charged particles of the electomagnetic cascade can travel fastest in air than light!

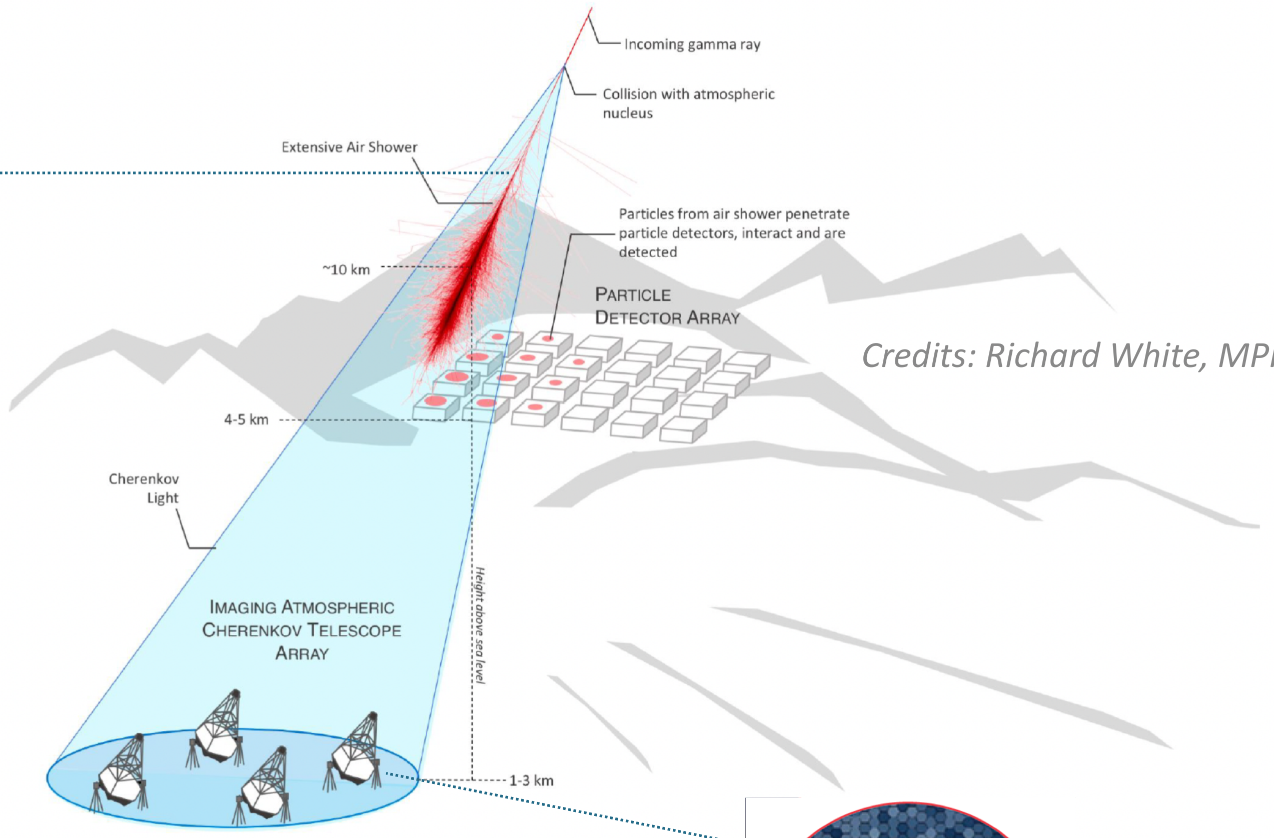
- Particles of the air interact with the charged particles of the electromagnetic cascades producing Cherenkov radiation

$$\cos\theta_c = \frac{1}{\beta n}$$

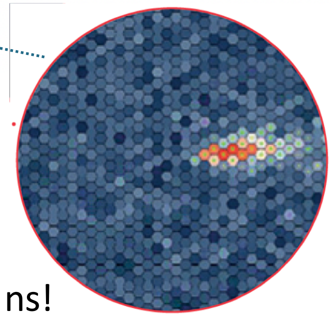
$$\beta = \frac{v_p}{c}$$

Refraction index





Credits: Richard White, MPIK



Picture of 10 ns!

THE CHERENKOV TELESCOPE ARRAY OBSERVATORY

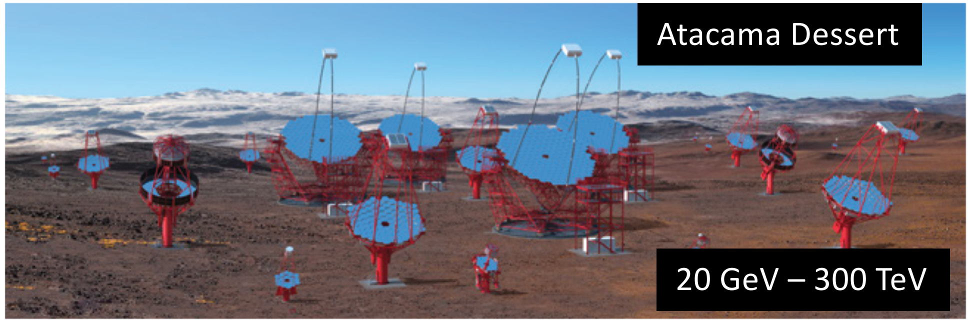
The mean of gamma-rays that reach the Earth is 1 per 1m^2 per year!

↓

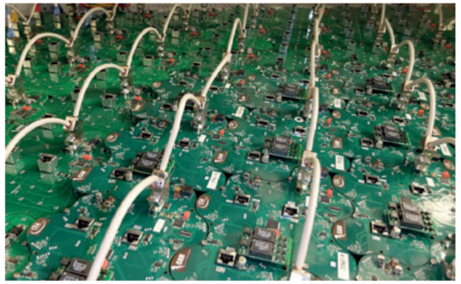
Cherenkov Telescope Array (CTA)

https://www.cta-observatory.org/

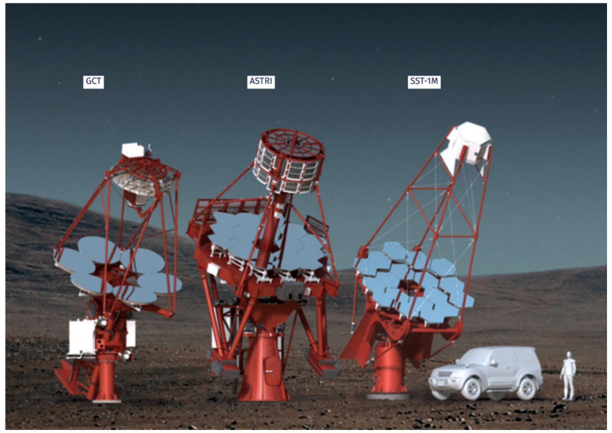
- Two emplacements for the arrays



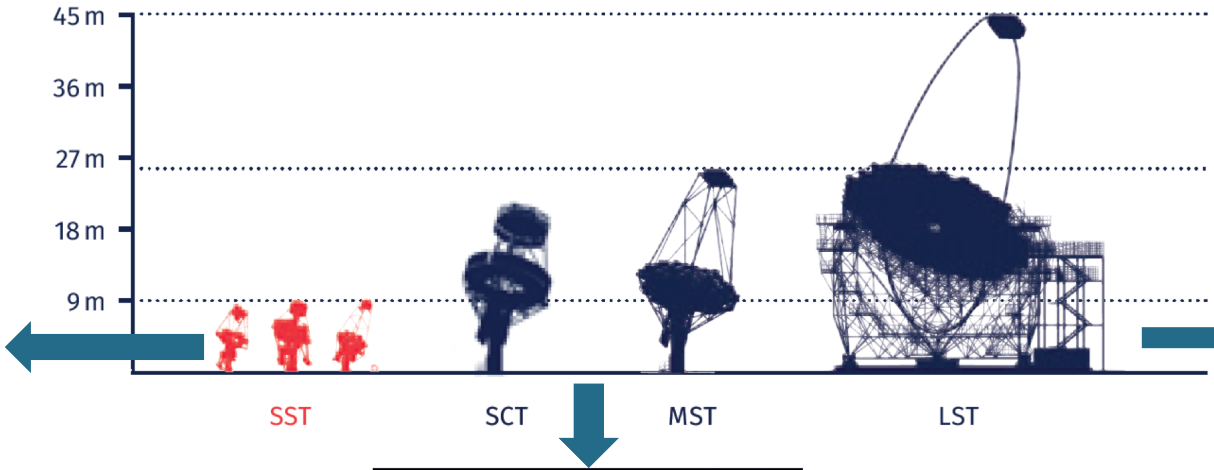
- The telescopes are a complex systems of mirrors that redirects the light to the camera. The signal is then translated to electromagnetic pulses that after data reprocessing, can be used to obtain information about the original gamma-ray: **energy (*E*) & direction (*p*)**

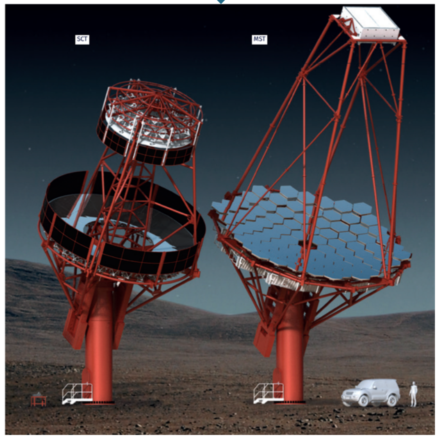


THE CHERENKOV TELESCOPE ARRAY OBSERVATORY

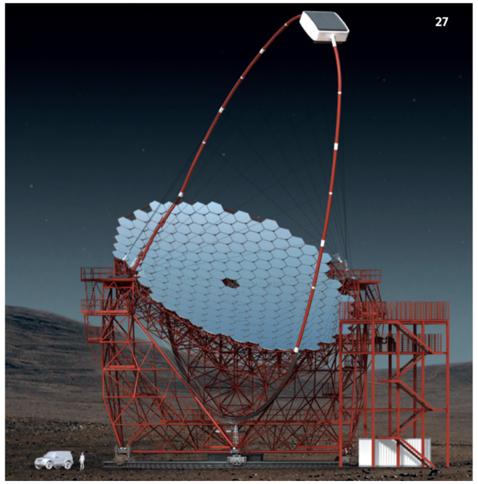


- ~70 telescopes in total
- From 1 – 300 TeV





- ~40 telescopes in total
- From 80 GeV – 50 TeV



- ~8 telescopes in total
- From 20 GeV – 3 TeV

INSTRUMENT RESPONSE FUNCTIONS

Source: https://docs.gammapy.org/1.1/user-guide/irf/index.html?highlight=irf

- To reconstruct the characteristics of the primary gamma-ray, we need to know how the telescopes affects it
- We characterize this effect through the Instrument Response Functions (IRFs)

IRFs

$$N(p, E)\mathrm{d}p\mathrm{d}E = t_{\mathrm{obs}} \int_{E_{\mathrm{true}}} \mathrm{d}E_{\mathrm{true}} \int_{p_{\mathrm{true}}} \mathrm{d}p_{\mathrm{true}} \boxed{R(p, E|p_{\mathrm{true}}, E_{\mathrm{true}})} \times \Phi(p_{\mathrm{true}}, E_{\mathrm{true}})$$

- *N(p, E)* gives the probability to detect a photon emitted from true position p_{true} on the sky and true energy E_{true} at reconstructed position *p* and energy *E*

- We assume that we can characterize the response of the instrument in a factorization of three distributions

$$R(p, E|p_{\mathrm{true}}, E_{\mathrm{true}}) = \boxed{A_{\mathrm{eff}}(p_{\mathrm{true}}, E_{\mathrm{true}})} \times \boxed{PSF(p|p_{\mathrm{true}}, E_{\mathrm{true}})} \times \boxed{E_{\mathrm{disp}}(E|p_{\mathrm{true}}, E_{\mathrm{true}})}$$

INSTRUMENT RESPONSE FUNCTIONS

Source: https://docs.gammapy.org/1.1/user-guide/irf/index.html?highlight=irf

$$R(p, E|p_{\rm true}, E_{\rm true}) = A_{\rm eff}(p_{\rm true}, E_{\rm true}) \times PSF(p|p_{\rm true}, E_{\rm true}) \times E_{\rm disp}(E|p_{\rm true}, E_{\rm true})$$

- **Effective Area *A_{eff}***

Is the effective collection area of the detector. It is the product of the detector collection area times its detection efficiency.

- **Point Spread Function *PSF***

Has units of sr^{-1}. It gives the probability of measuring a direction p when the true direction is p_{true}.

- **Energy dispersion *E_{disp}***

It gives the probability to reconstruct the photon at energy E when the true energy is E_{true}.

INSTRUMENT RESPONSE FUNCTIONS

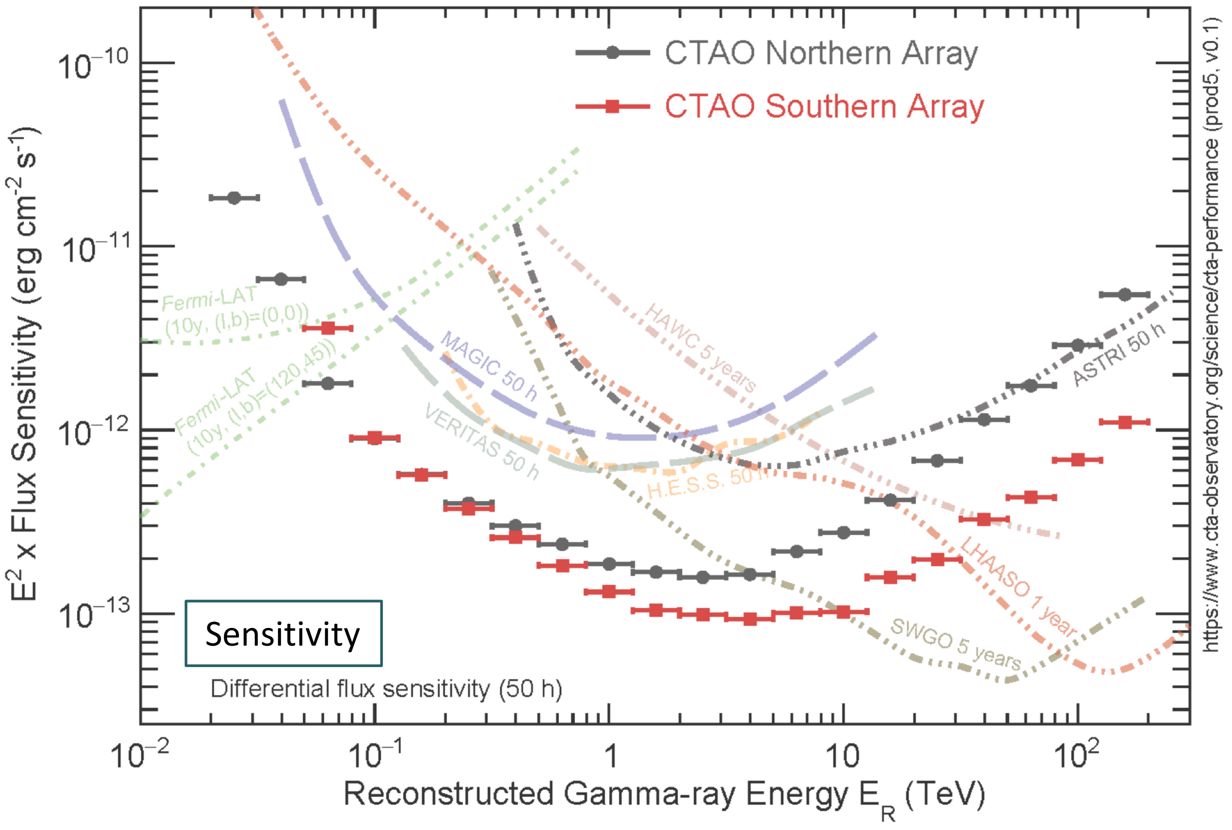
The irreducible background *[Bernlörhr + 12]*

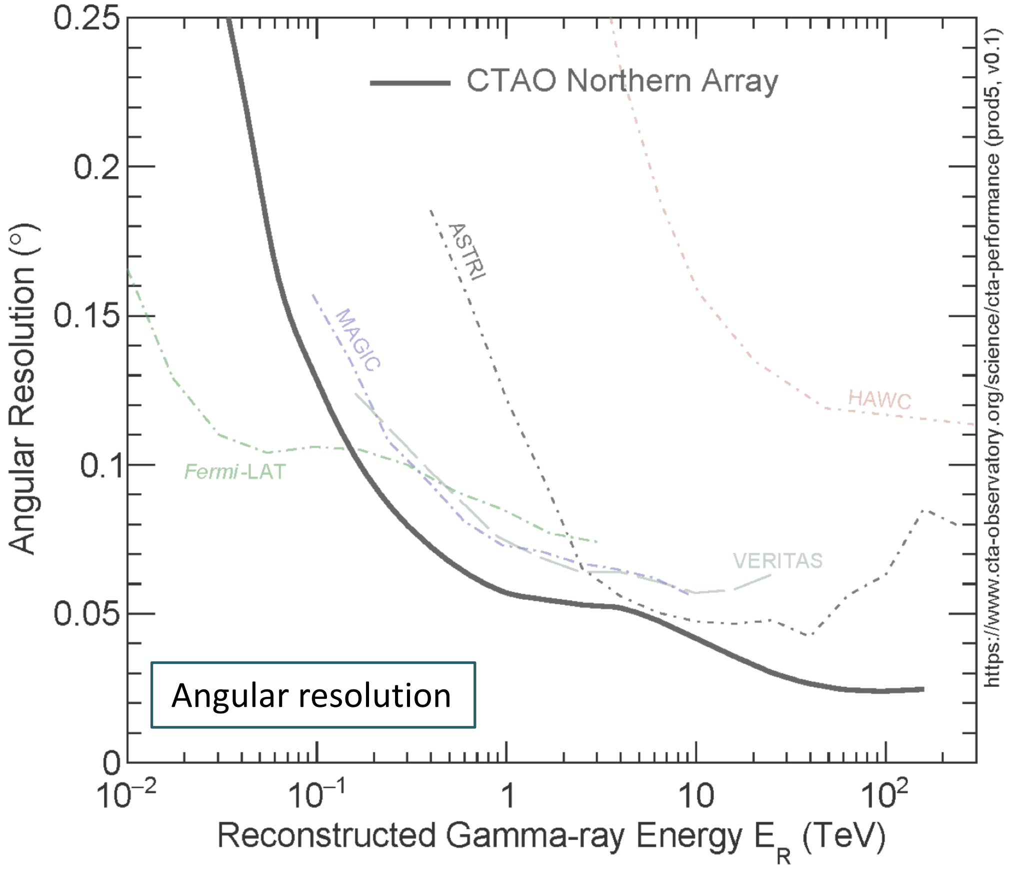
- We need a way to discrimination between gamma-ray initiated showers and hadron showers
- IACTs use image shape (fail if cascade initated by electrons) and shower direction (may fail for extended sources) cuts for that
- Therefore, there will always be a level of confusión that we should quantify as part of the response of the intrument: the background

INSTRUMENT RESPONSE FUNCTIONS

Preliminary Performance Capabilities

https://www.cta-observatory.org/





GAMMAPY

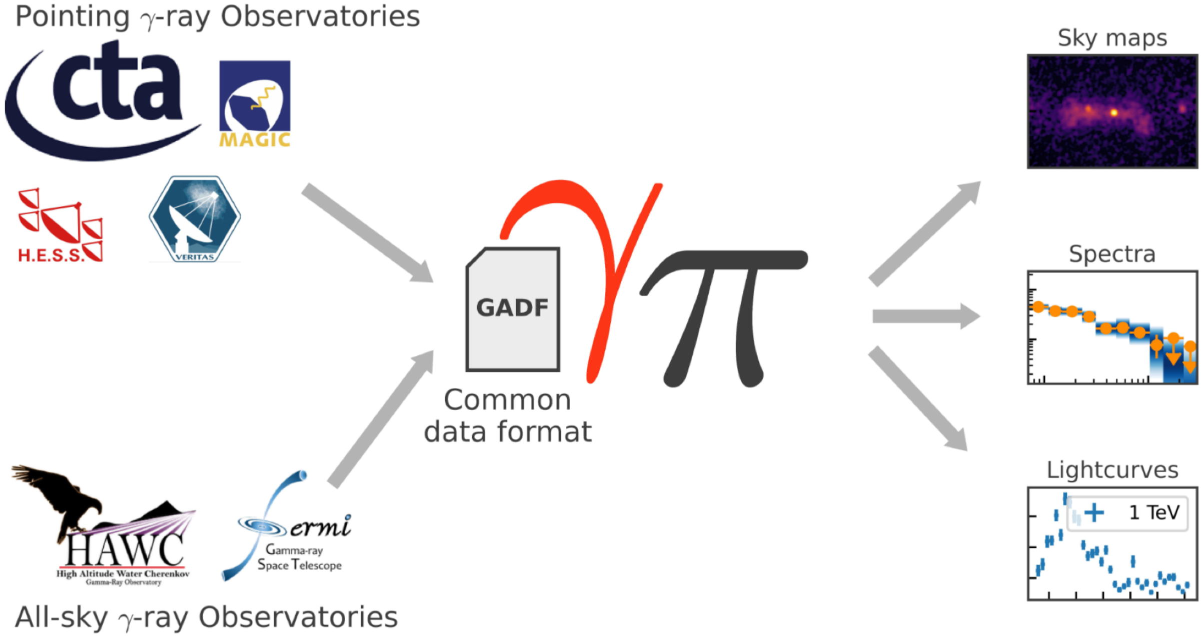
https://docs.gammapy.org/1.1/index.html

- What is Gammapy?
  - Community-developed, **open-source** in Python
  - "Easy" to extend according to community needs
  - Computationally fast

- Stable version sin June v-1.1

Common set of tools

- Unified definitions, methodology
- Avoids repetition of coding
- Easy comparison of results
- Everyone can potentially contribute



Credits: Bruno Khélifi

SEARCHING FOR GAMMA-RAY EMISSION WITH CTA

1. Create a simulated observation of the galactic centre with a gamma-ray extended source

2. Blind analysis of provided dataset

SEARCHING FOR GAMMA-RAY EMISSION WITH CTA

1. **Create a simulated observation of the galactic centre with a gamma-ray extended source**

2. Blind analysis of provided dataset

SIMULATING CTA GAMMA-RAY OBSERVATION

Basic input information to create **ANY** simulated data:

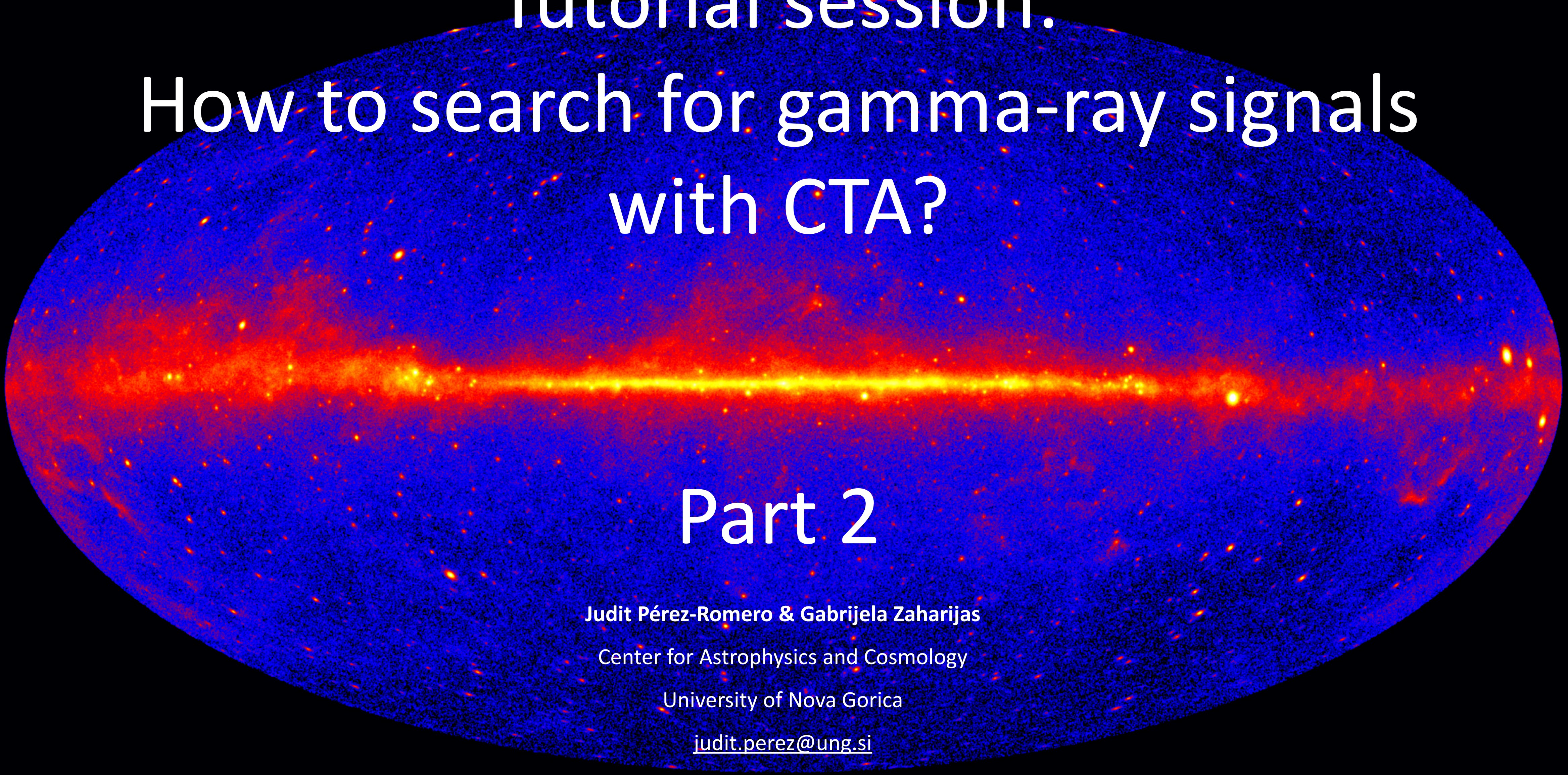
- **Livetime**: duration of the observation
- **Pointing**: position of the sky to look at
- **Region of Interest (ROI)**: area of interest around the pointing
- **Energy range** and/or **binning**
- **Instrument Response Functions (IRFs)**
- **Model**

SIMULATING CTA GAMMA-RAY OBSERVATION

Basic input information to create **ANY** simulated data:

- Livetime
- Pointing
- Region of Interest (ROI)
- Energy range and/or binning
- Instrument Response Functions (IRFs)
- Model

- Background model: from IRFs
- Source model: can comprehend several sources, each of them could have
  - Spectral model
  - Spatial model
  - Time model



Tutorial session: How to search for gamma-ray signals with CTA?

Part 2

Judit Pérez-Romero & Gabrijela Zaharijas

Center for Astrophysics and Cosmology

University of Nova Gorica

judit.perez@ung.si

Image Credits: Fermi-LAT Collaboration

HOW TO SEARCH FOR GAMMA-RAY SIGNALS WITH CTA

- We will work in Google collab environment, for this you need:
  - Google account
  - Preferably Google Chrome
  - Download the dataset for the analysis https://drive.google.com/file/d/1w8V1znwdh38_-vHfGMS68G2BUdpY-PpF/view?usp=sharing

https://colab.research.google.com/drive/185MH2XjbCmMfvXLp5PXkqsoR5-nDkvhj?usp=sharing

SEARCHING FOR GAMMA-RAY EMISSION WITH CTA

1. Create a simulated observation of the galactic centre with a gamma-ray extended source

2. **Blind analysis of provided dataset**

Maximum Likelihood Method

- Given a set of observed data:
  - produce a model that describes it, including parameters that we wish to estimate
  - derive the probability for the data given the model
  - treat this as a function of the model parameters
  - maximize the likelihood with respect to the parameters

$$\ln\mathcal{L}(\vec{\theta}|D)$$

Observed data

Parameters of the model

- Advantages of the method:
  - “cookbook” through which different kind of problems can be solved
  - for other methods ad-hoc choices may have to be made
  - provides unbiased, minimum variance estimate as sample size increases
  - evaluation of confidence bounds and hypothesis testing
  - well studied in the literatura: particle physics, astroparticles, etc.

GAMMA-RAY ANALYSIS FOR CTA DATA

- Use maximum likelihood approach with a 3D fitting: 2 dimensions in space and 1 in energy

$$\ln\mathcal{L}(\vec{\theta}|D) = \sum_i M_i(\vec{\theta}) - d_i \ln(M_i(\vec{\theta}))$$

Cash statistics *[Cash 79]*

→ Counts predicted by the model to fit, or combination of models, following Poisson distribution

- In the fit, we obtain as **best values** the ones **maximizing the likelihood function**

- To test if a model is better than other to fit a dataset we use the **likelihood ratio test** (*TS*):

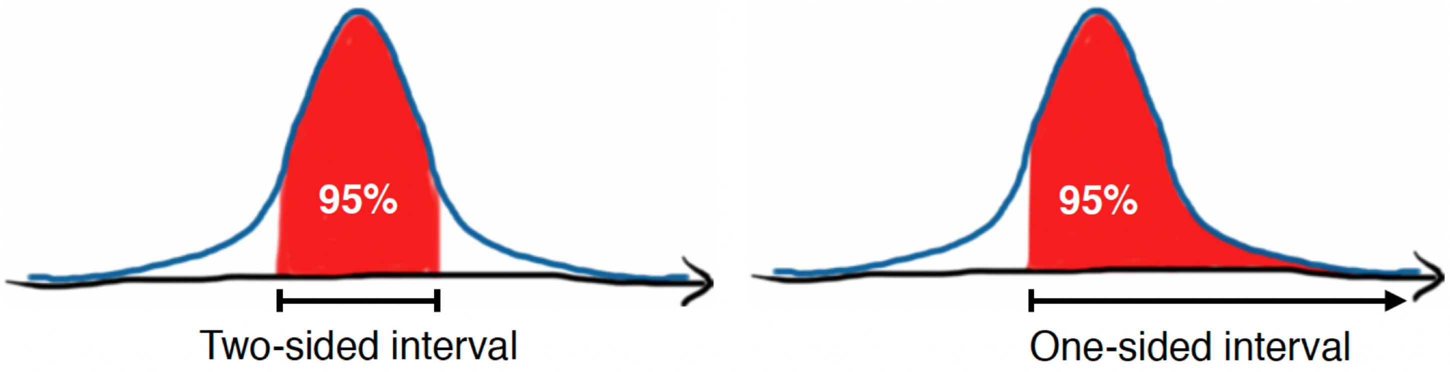
$$TS = 2 \times \frac{\ln \mathcal{L}(H_1)}{\ln \mathcal{L}(H_0)}$$

- If H_0 is the null-hypothesis (only background), we determine a detection when

$$\boxed{TS \geq 25}$$ → $\sim 5\sigma$ detection *[Li&Ma 83]*

GAMMA-RAY ANALYSIS FOR CTA DATA

- If we do not find a signal… Then let's put constraints on the parameter space!
- The likelihood has several dependencies, we need to project over the parameter of interest: Likelihood profile
- The limits can be one-sided or two-sided:



[credits: S. Fegan for Fermi-LAT school]

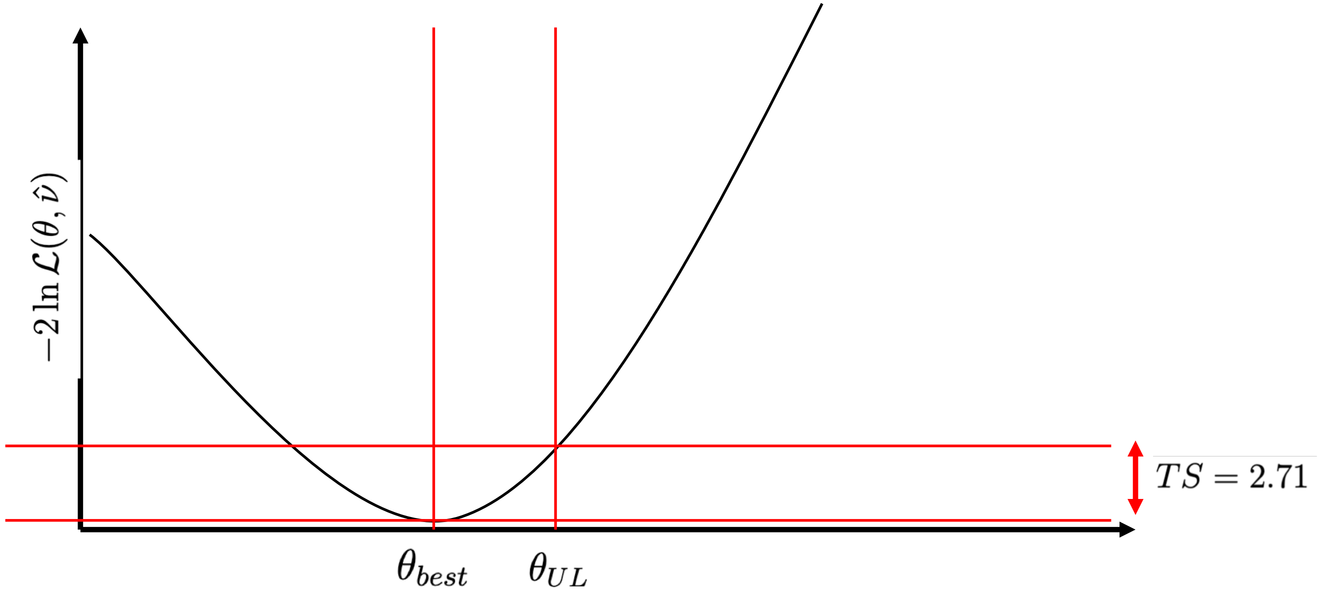
- This limits read as: the upper/lower value most probable to get by 95% of the times (if *TS* is distributed following a χ^2)
  - For the one-sided distribution → $95\,\%\mathrm{C.L} \rightarrow TS = 2.71$
  - For the two-sided distribution → $95\,\%\mathrm{C.L} \rightarrow TS = 3.84$

[Rolke+05]

GAMMA-RAY ANALYSIS FOR CTA DATA

- We have not found a signal… Then let's put constraints!
- The likelihood has several dependencies, we need to project over the parameter of interest: Likelihood profile

Example: 1-sided upper limits



- We need to solve: $-2\ln\frac{\mathcal{L}(\theta_{UL},\hat{\nu})}{\mathcal{L}(\theta_{best},\hat{\nu})} - 2.71 = 0$ *[Rolke+05]*

GUESSING THE DATASET

Spectral Model fitted				Spatial Model fitted				TS	$\sqrt{TS}$
Power Law	Γ = 2.0	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 1 TeV	Point Source	l = 0.0 deg	b = 0.0 deg		3.17x10^{10}	1.78x10^{5}
Power Law	Γ = 2.0	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 1 TeV	Gaussian	l = 0.0 deg	b = 0.0 deg	σ = 0.2	6.65x10^{11}	8.16x10^{5}
LogParabola	α = 2.0 β = 0.1	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 10 TeV	Point Source	l = 0.0 deg	b = 0.0 deg		3.38x10^{10}	1.84x10^{5}
LogParabola	α = 2.0 β = 0.1	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 10 TeV	Gaussian	l = 0.0 deg	b = 0.0 deg	σ = 0.2	7.35x10^{11}	8.58x10^{5}

Power law

$$\phi(E) = \phi_0 \cdot \left(\frac{E}{E_0}\right)^{-\Gamma}$$

Log Parabola

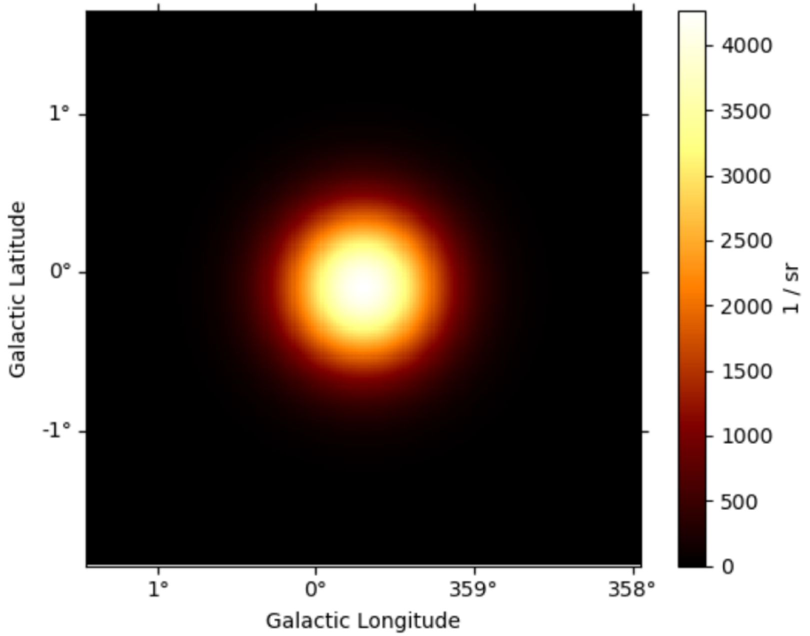
$$\phi(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\log\left(\frac{E}{E_0}\right)}$$

ORIGINAL VALUES OF THE PROVIDED DATASET

Location	Spectral shape		Spatial shape
l = -0.3 deg	Log-Parabola	β= 0.27	Gaussian
b = -0.1 deg	ϕ_0= 11.14^{-11} cm^{-2} s^{-1} MeV^{-1}	E_0 = 1 TeV	σ = 0.35 deg
	α = 2.0		

Inspired by: https://arxiv.org/pdf/2101.04694.pdf

$$\phi(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\log\left(\frac{E}{E_0}\right)}$$



BONUS: DARK MATTER (DM) WITH CTA

Basic input information for DM modelling:

- Model: Annihilation of Weakly Interactive Massive Particles (WIMPs)

$$\frac{d\Phi_\gamma}{dE}(E, l.o.s, \Delta\Omega) = scale \times \boxed{J(l.o.s, \Delta\Omega)} \times \frac{\langle\sigma v\rangle}{4\pi\, m_{\mathrm{DM}}^2} \boxed{\sum_i \mathrm{BR}_i \frac{dN_\gamma^{\mathrm{i}}}{dE}(E)}$$

Spatial model

- Assumes ΛDCM model
- Encodes how the DM is distributed in the object
- We can use different ρ_{DM} parametrizations
- Ends acting as a multiplicative factor to the overall flux

Spectral model

- Encodes the spectrum of the emission
- We can use the tables computed by *[Cirelli+11]*

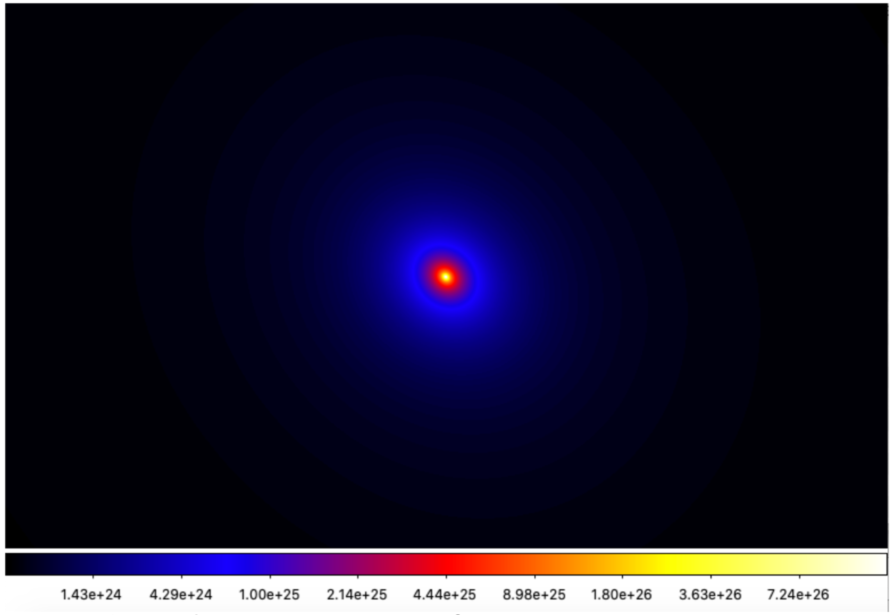
  http://www.marcocirelli.net/PPPC4DMID.html

BONUS: DARK MATTER WITH CTA

Basic input information for DM modelling:

- Model: Annihilation of Weakly Interactive Massive Particles (WIMPs)

Spatial model

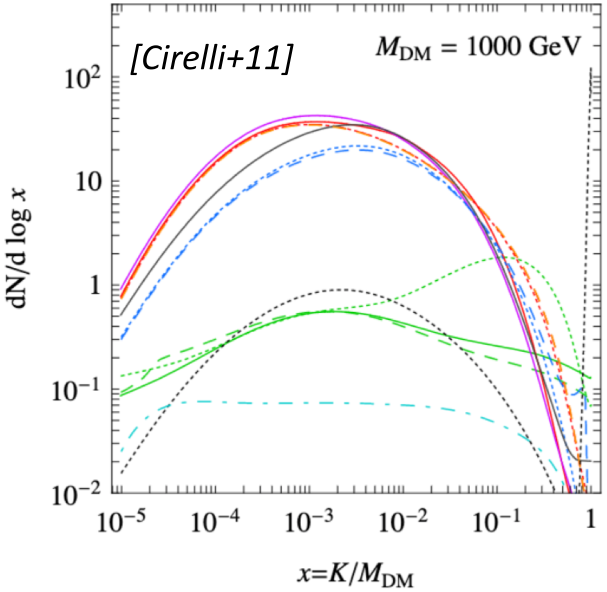


Created with CLUMPY software

[Charbonnier+12, Bonnivard+15, Hütten+18]

https://clumpy.gitlab.io/CLUMPY/

Spectral model



Gamma-ray emission spectrum from:

- WIMP annihilation
- m_{DM} = 5 TeV
- $b\bar{b}$ channel

GUESSING THE DATASET

Spectral Model fitted				Spatial Model fitted				TS	$\sqrt{TS}$
Power Law	Γ = 2.0	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 1 TeV	Point Source	l = 0.0 deg	b = 0.0 deg		3.17x10^{10}	1.78x10^{5}
Power Law	Γ = 2.0	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 1 TeV	Gaussian	l = 0.0 deg	b = 0.0 deg	σ = 0.2	6.65x10^{11}	8.16x10^{5}
LogParabola	α = 2.0 β = 0.1	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 10 TeV	Point Source	l = 0.0 deg	b = 0.0 deg		3.38x10^{10}	1.84x10^{5}
LogParabola	α = 2.0 β = 0.1	ϕ_0 = 1.0x10^{-12} cm^{-2} s^{-1} TeV^{-1}	E_0 = 10 TeV	Gaussian	l = 0.0 deg	b = 0.0 deg	σ = 0.2	7.35x10^{11}	8.58x10^{5}
DM annihilation	m_{DM} = 5 TeV	J = 1x10^{21} GeV^2 cm^{-5}	ch = $b\bar{b}$	Gaussian	l = 0.0 deg	b = 0.0 deg	σ = 0.2	5.84x10^{11}	7.64x10^{5}

Power law

$$\phi(E) = \phi_0 \cdot \left(\frac{E}{E_0}\right)^{-\Gamma}$$

Log Parabola

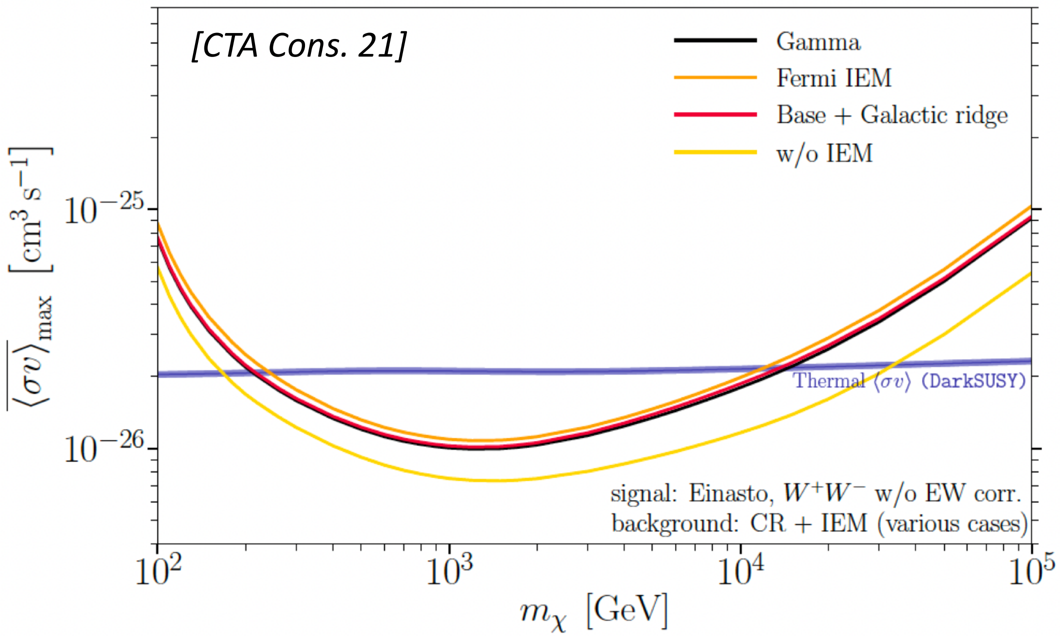
$$\phi(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\log\left(\frac{E}{E_0}\right)}$$

Dark Matter Annihilation

$$\frac{\mathrm{d}\phi}{\mathrm{d}E} = \frac{\langle\sigma\nu\rangle}{4\pi k m_{\mathrm{DM}}^2}\frac{\mathrm{d}N}{\mathrm{d}E} \times J(\Delta\Omega)$$

BONUS: DARK MATTER WITH CTA

- Okey but, these are not the cool plots for constraints!



- We have fitted our dataset to one channel and one mass. We will need to loop over the mass range of interest and the available channels to get these lines.

- On top of that, we have only used one Poisson realization. To have statistically meaningful predictions, we need to average over O(300) realizations.

15 mass values x 3 channels x 300 realizations = 13500 fits!