Combined analysis pipeline for joint observations with MAGIC and CTA LST-1

IMPRS Young Scientist Workshop, Ringberg 2022

Giorgio Pirola*,

for the CTA LST project and MAGIC Collaboration

*Ph.D candidate at Max Planck Institut für Physik, Munich

10/05/20222

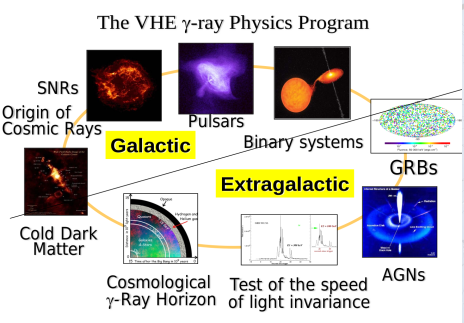
MAX-PLANCK-INSTITUT FÜR PHYSIK

High-Energy gamma-ray astronomy

Multi-wavelength approach: cover the whole electromagnetic spectrum, from radio waves to gamma rays, to study sources at different wavelengths.

- **Why gammas?**
  - Gammas are neutral $\rightarrow$ we can study their production sites through the spectra we measure.
  - Can give information on cosmic rays origins and mechanism of acceleration.
  - Probes for fundamental physics problems: Dark Matter, LIV

- **Energy Classification**
  - High Energy (HE): 0.5 MeV – 100 GeV
  - **Very High Energy (VHE): 100 GeV - 100 TeV**
  - Ultra High Energy (UHE): 100 TeV – 100 PeV
  - Extreme High Energy (EHE): > 100 PeV
- **Sources**
  - Galactic sources: Pulsars, Supernova Remnants, Binary Systems.
  - Extragalactic sources: Starburst, AGN, GRB.



Detection Techniques

Detection of VHE (and more energetic) gammas is not possible with satellites $\rightarrow$ limited effective areas $\sim m^2$

Ground detection

- Relies on Extensive Air Shower (EAS) development
- Effective areas $\rightarrow \sim 10^5 m^2$

2 complementary techniques

- IACT (Imaging Air Cherenkov Technique):
  - sensitive to point-like sources
  - collection area $\sim 10^4 m^2$ (light-pool)
  - energy range $\rightarrow$ VHE
- EAS arrays:
  - nearly 100% duty cycle;
  - higher energy threshold;
  - energy range $\rightarrow$ UHE

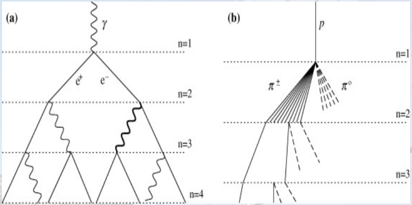
Imaging Air Cherenkov Telescopes

Extensive Air Showers are the result of the interaction of cosmic primary particle with the atmosphere.

There are two kinds of EAS:

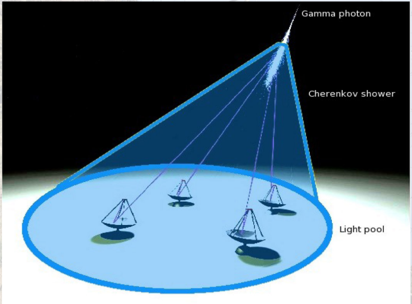
(a) Electromagnetic showers $\rightarrow$ signal for IACTs

(b) Hadronic showers: $\rightarrow$ background ($\sim 10^3$ greater than gamma signal)



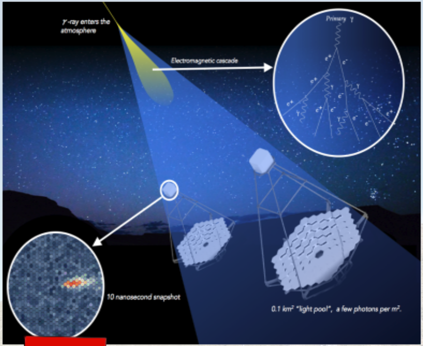
Many of these secondaries travel faster than light in the atmospheric medium $\rightarrow$
flashes of Cherenkov light detected by IACTs:

- Large mirrors collect Cherenkov light and focused it on a camera of PMTs.
- PMTs:
  - Sensitive in the wavelength range of Cherenkov light.
  - Very fast response $\rightarrow$ duration of Cherenkov pulses $\sim$few ns

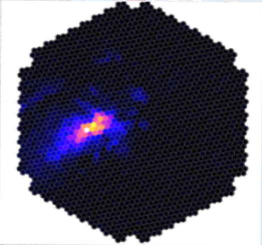


Imaging Air Cherenkov Telescopes

MAX-PLANCK-INSTITUT
FÜR PHYSIK



Hillas Parametrization

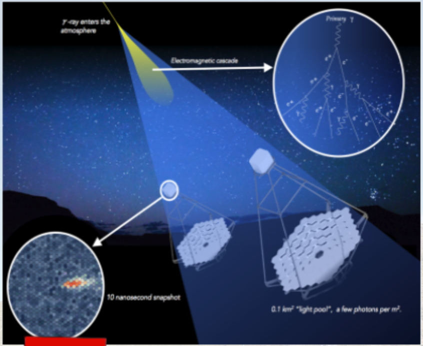


ML → Random Forests (trained with MC data)

Physical properties of the primary particle:

- Energy
- Direction
- "Gammaness"

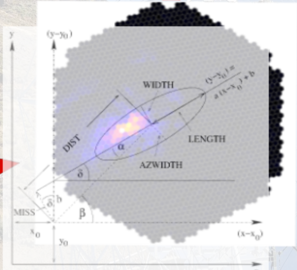
Imaging Air Cherenkov Telescopes



Hillas Parametrization

Physical properties of the primary particle:
- Energy
- Direction
- "Gammaness"

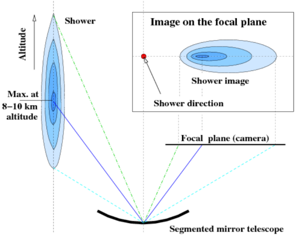
ML → Random Forests (trained with MC data)



Imaging Air Cherenkov Telescopes

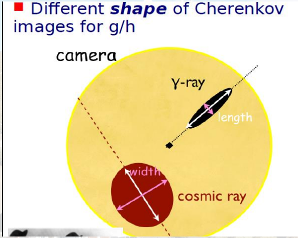
Arrival direction:

- The camera is located in the focal plane of the mirrors: the image has a comet-like shape with the head pointing towards source position.



Energy:

- The higher the energy, the brighter the image.



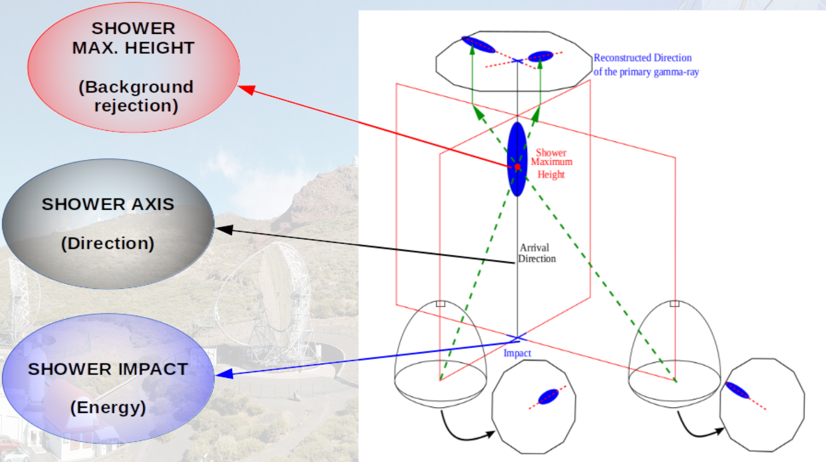
Classification:

- Image produced gamma-initiated EAS ⟶ compact elliptic shape
- Image produced hadron-initiated EAS ⟶ more complex structure due to E.M. sub-showers

The whole parametrization of the images produced MC simulated events, is used to train a ML technique known as **Random Forest** method

Imaging Air Cherenkov Telescopes

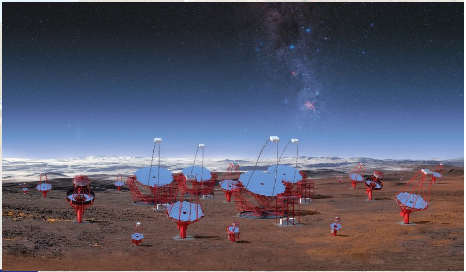
Reconstruction performances are remarkably improved with the **stereoscopic view**

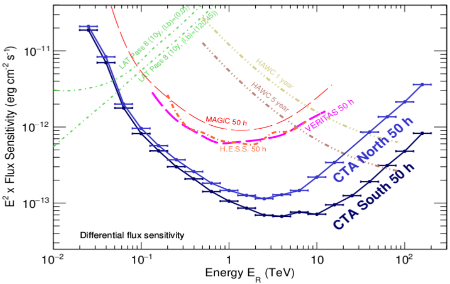


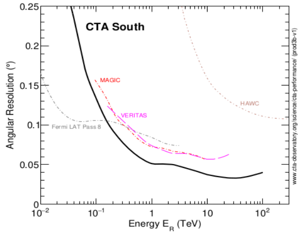
Cherenkov Telescope Array (CTA)

CTA will cover a wide energy range between 20 GeV and 300 TeV

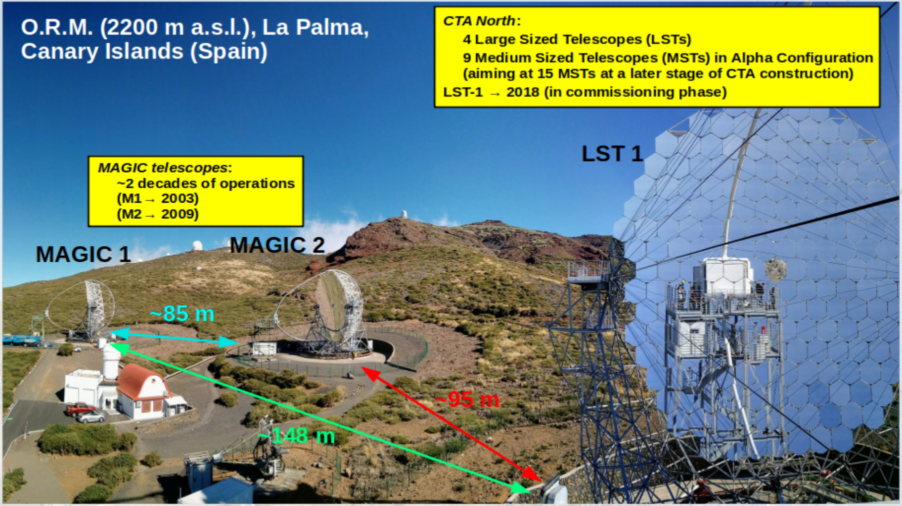
- Sensitivity and performance far superior to current experiments
- More than 100 telescopes in two array sites:
  - **Northern hemisphere array** (La Palma, Canary Islands)
  - **Southern hemisphere array** (Chile, near Paranal)







MAGIC and LST-1



Joint Observations

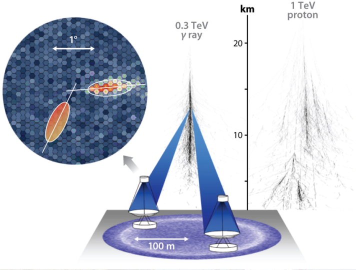
Being at ∼100m from each others, means that can be triggered by the same event: Cherenkov light-pool → ∼100m radium

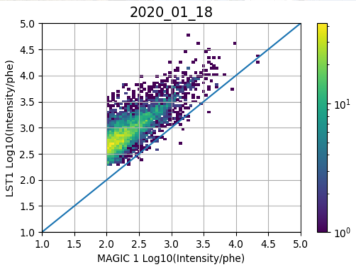
We already have several hours of simultaneous observations:

- first goal → **cross-calibration** of the 2 systems.
- final aim → perform the event-reconstruction with a 3-telescope stereoscopic view.

How?

- By comparing MAGIC and LST-1 event timestamps and performing an offline search of the coincident events

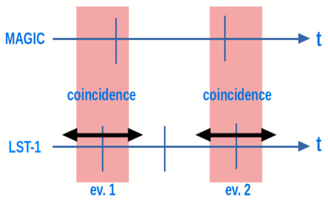


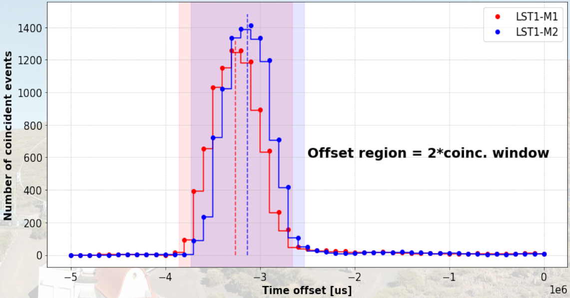


Joint Observations

The coincidence search is performed comparing timestamps:

- a given offset is applied to LST-1 event timestamp
- the algorithm searches for an event in MAGIC included within a certain coincidence window

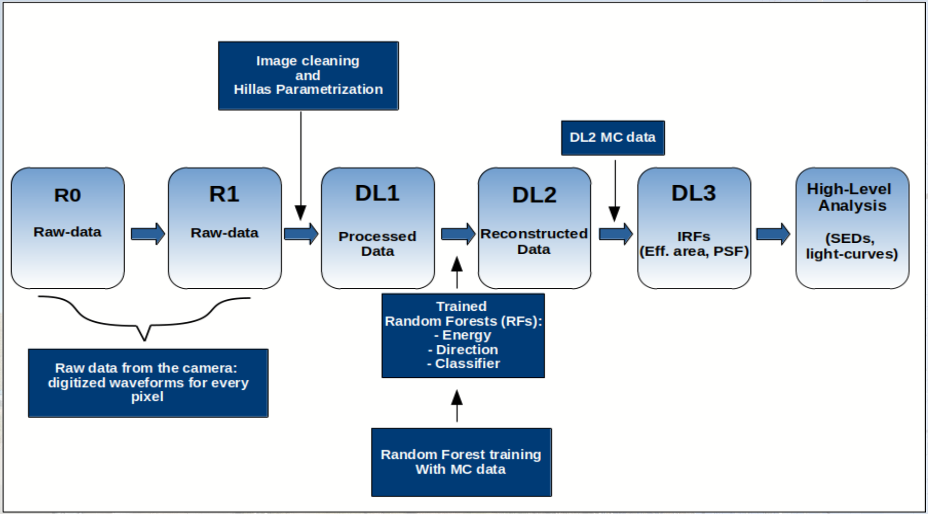




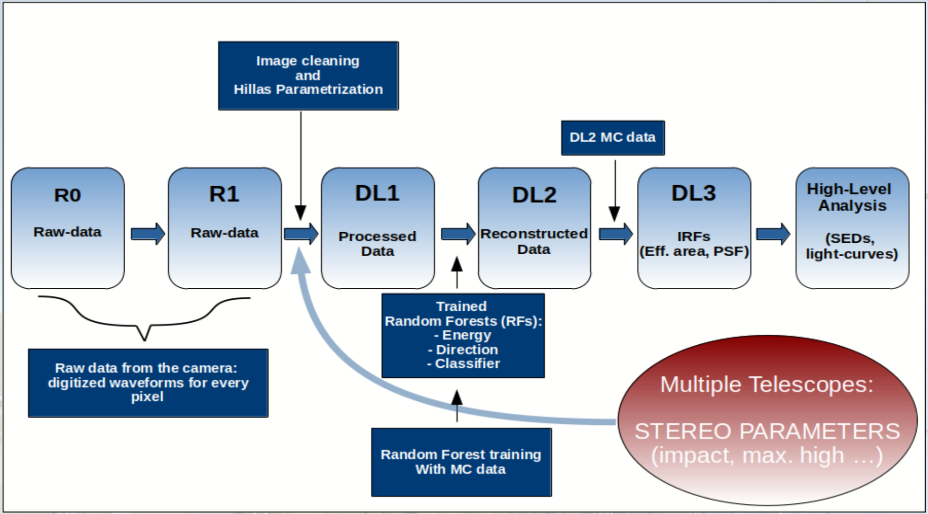
- the process is iterated over different offset values
- the final coincident events are taken with the average offset

Hardware trigger under development → should be ready by this year

Data analysis chain

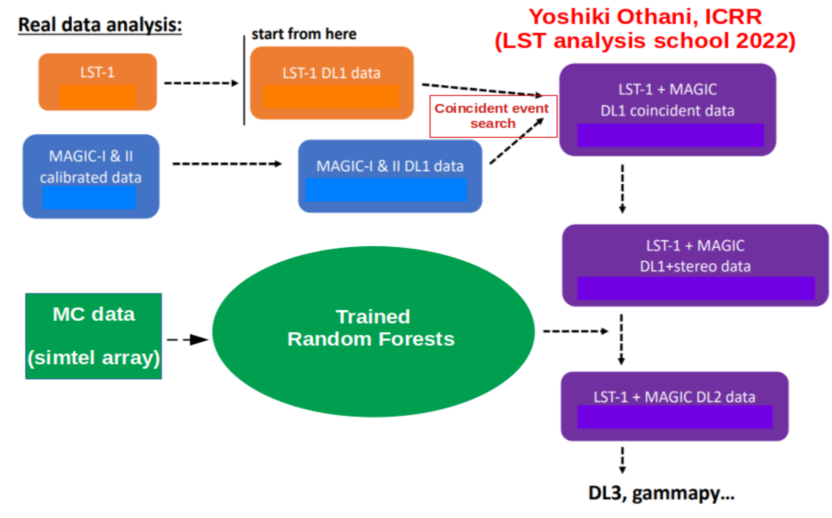


Data analysis chain



→ **Stereo parameters added to the Random Forests training features**

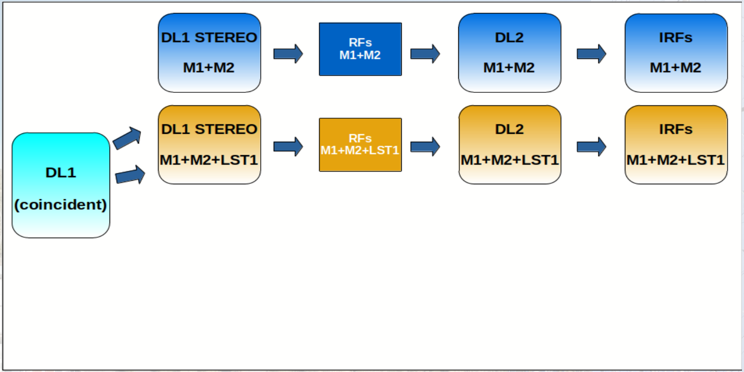
Combined analysis pipeline



Performance for different telescope combinations

MAX-PLANCK-INSTITUT FÜR PHYSIK

- MC data:
  - point-like gamma with 0.4° wobble
  - 5 GeV < E < 50 TeV
  - Zd=40° Az=90°
- Crab data:
  - 4 runs of joint observations (2021/12/15)
  - 35°<Zd<50°

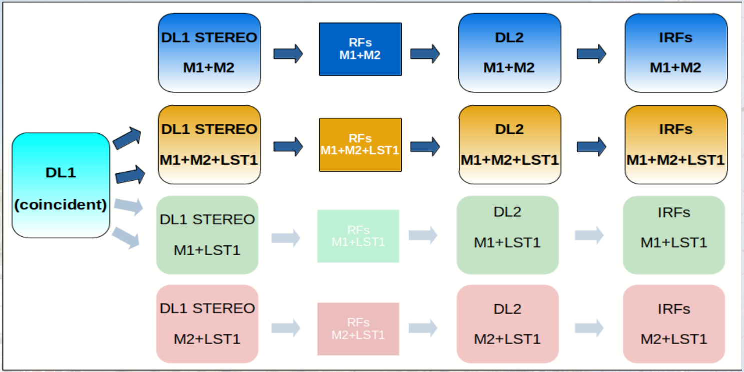


MAGIC

cta

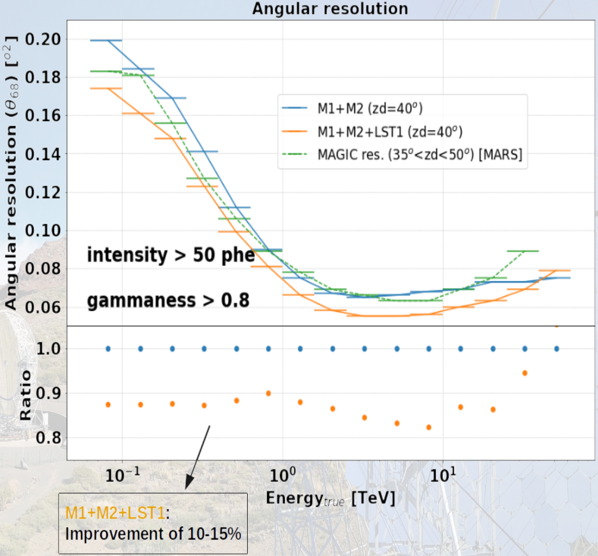
Performance for different telescope combinations

- MC data:
  - point-like gamma with 0.4^o wobble
  - 5 GeV < E < 50 TeV
  - Zd=40^o Az=90^o
- Crab data:
  - 4 runs of joint observations (2021/12/15)
  - 35^o<Zd<50^o



Instrument Response Functions (IRFs)

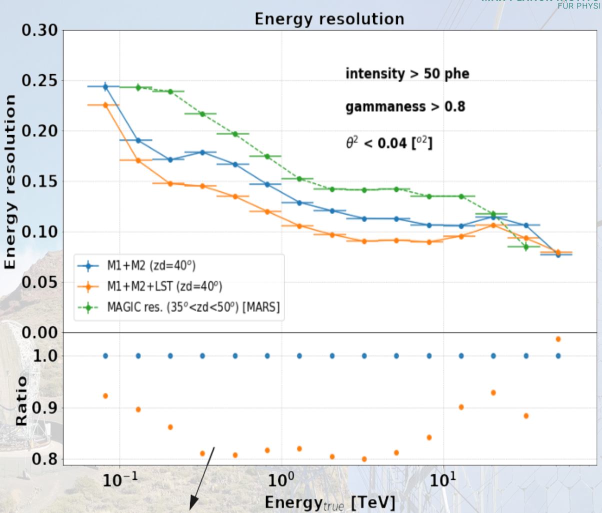
- M1+M2
  - **Only by this combo: 0.2%**
  - **Total: 56.1%**
- M1+M2+LST1
  - **Total : 55 .9%**
- M1+LST1
  - **Only by this combo: 16.0%**
  - **Total: 71.8%**
- M2+LST1
  - **Only by this combo: 27.9%**
  - **Total: 83.8%**



Instrument Response Functions (IRFs)

MAX-PLANCK-INSTITUT FÜR PHYSIK

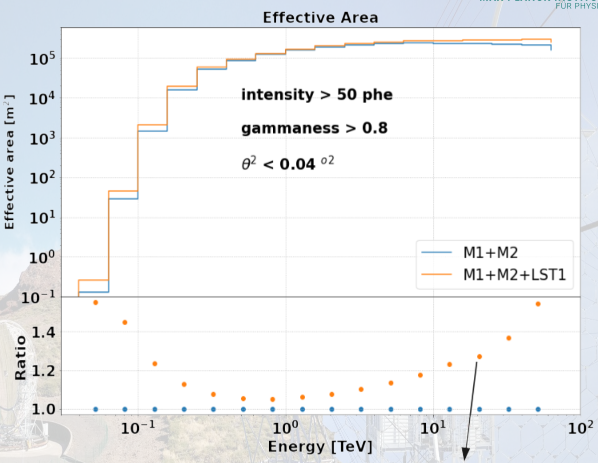
- M1+M2
  - Only by this combo: 0.2%
  - Total: 56.1%
- M1+M2+LST1
  - Total : 55 .9%
- M1+LST1
  - Only by this combo: 16.0%
  - Total: 71.8%
- M2+LST1
  - Only by this combo: 27.9%
  - Total: 83.8%



M1+M2+LST1:
Improvement of 10-20%

Instrument Response Functions (IRFs)

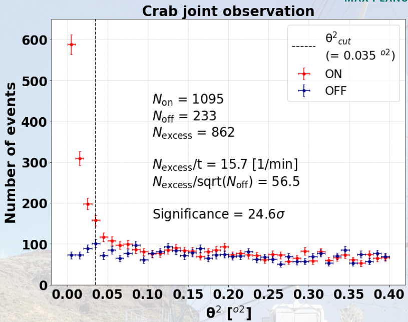
- M1+M2
  - **Only by this combo: 0.2%**
  - **Total: 56.1%**
- M1+M2+LST1
  - **Total : 55 .9%**
- M1+LST1
  - **Only by this combo: 16.0%**
  - **Total: 71.8%**
- M2+LST1
  - **Only by this combo: 27.9%**
  - **Total: 83.8%**

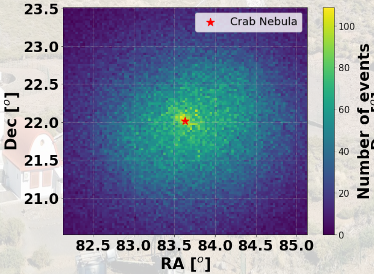


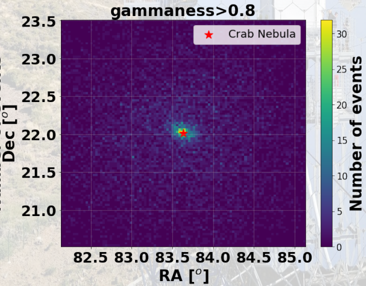
M1+M2+LST1: Almost 100% events seen by M1+M2, but more events survive the gamma selection → better background rejection

Crab Detection

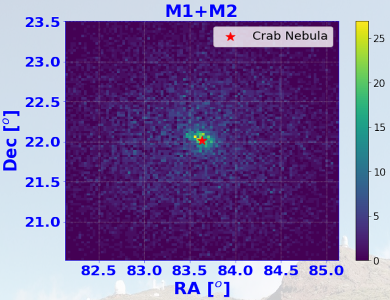
- intensity> 50 phe
- gammaness>0.8
- $\theta^2 < 0.035^{o2}$
- 1 OFF region
- Effective time: 55.1 min

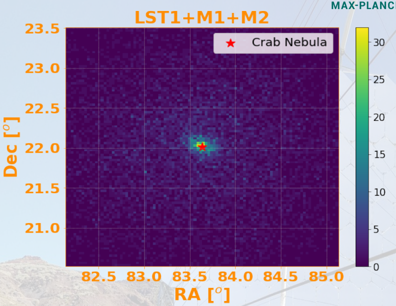


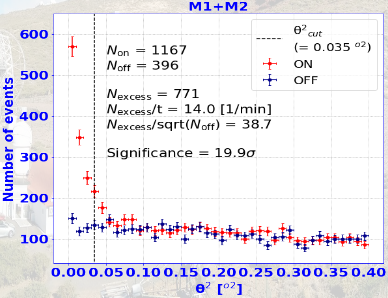


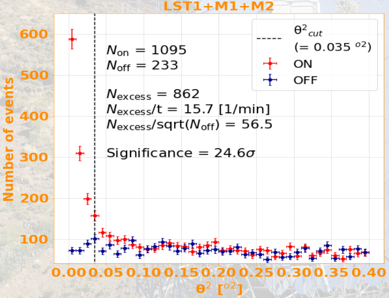


Crab Detection









Outlook

The pipeline for the combined analysis is still under development, but it has already demonstrated its potential, showing improvements in the performance reconstruction.

Next steps:

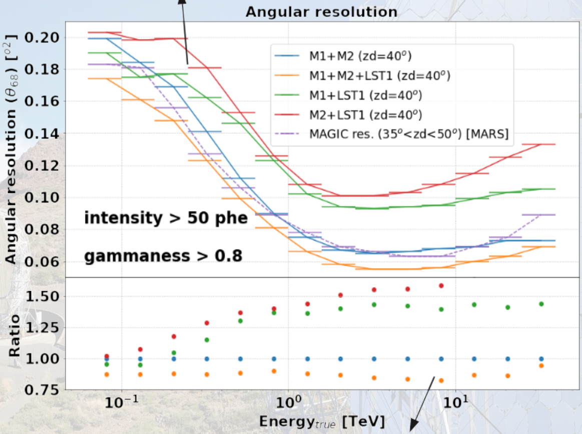
- **Optimization of the reconstruction process:**
  - Random Forests optimization (now 3 RFs for combination)
- **Check of the reliability of the simtel array MC production:**
  - New MC production covering a wider range of zenith angles
  - Detailed study and comparison of the MC data with Crab Observations
- **Development of the last part of the pipeline:**
  - More automatic tools for the IRFs production
  - gammapy tools for the high level analysis
- **Observing new sources and do physics!**

Backup slides

Backup: Angular resolution

- M1+M2+LST1
  - **Total : 55 .9%**
- M1+M2
  - **Only by this combo: 0.2%**
  - **Total: 56.1%**
- M1+LST1
  - **Only by this combo: 16.0%**
  - **Total: 71.8%**
- M2+LST1
  - **Only by this combo: 27.9%**
  - **Total: 83.8%**

M1+LST1 & M2+LST1:
Worse performance:
less events surviving the gamma selection
→ better way to weight the 2 contributions needed



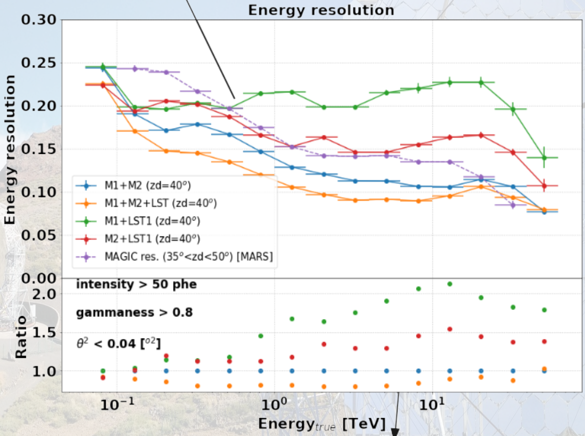
M1+M2+LST1:
Improvement of 10-15%

Backup: Energy resolution

MAX-PLANCK-INSTITUT FÜR PHYSIK

- M1+M2+LST1
  - **Total : 55 .9%**
- M1+M2
  - **Only by this combo: 0.2%**
  - **Total: 56.1%**
- M1+LST1
  - **Only by this combo: 16.0%**
  - **Total: 71.8%**
- M2+LST1
  - **Only by this combo: 27.9%**
  - **Total: 83.8%**

M1+LST1 & M2+LST1:
- Worse performance
- Less events surviving the gamma selection
  → better way to weight the 2 contributions needed



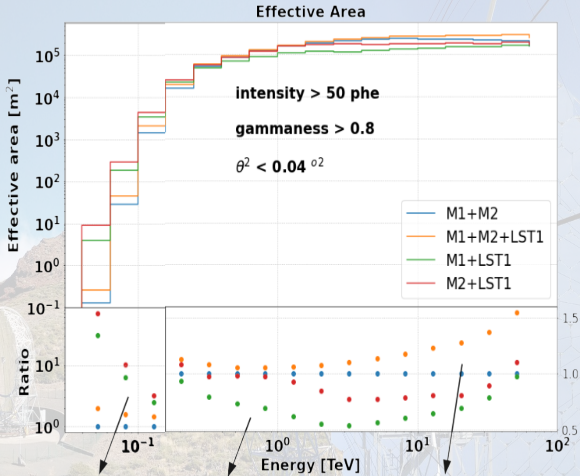
M1+M2+LST1:
Improvement of 10-15%

MAGIC

cta

Backup: Effective Area

- M1+M2+LST1
  - Total : 55 .9%
- M1+M2
  - **Only by this combo: 0.2%**
  - **Total: 56.1%**
- M1+LST1
  - **Only by this combo: 16.0%**
  - **Total: 71.8%**
- M2+LST1
  - **Only by this combo: 27.9%**
  - **Total: 83.8%**



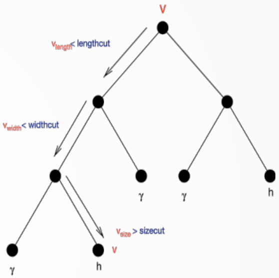
M1+LST1 & M2+LST1:
- Low energies: high statistics
- High energies: events surviving the gamma selection decrease

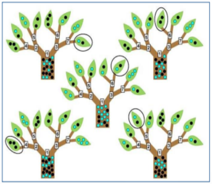
M1+M2+LST1:
Almost 100% events seen by M1+M2, but more events survive the gamma selection
→ better background rejection

Backup: Random Forest

Random Forest Training

- Each events is a point in a multidimensional space constituted by its **Hillas Parameters**;
- 2 MC samples (gammas/hadrons) are created in the space;
- One dimension is chosen at random and the optimal cut ($\mathbf{x_{cut}}$) that better separates the 2 distribution is found (using the so called *Gini* index);
- The sample is divided: $\mathbf{x<x_{cut}}$ and $\mathbf{x>x_{cut}}$;
- The process is iterated over all the dimensions until the final sub-sample contains only same type elements (g/h);
- Once trained the single **decision tree** the operation is then repeated (~100 trees);





Every real event goes through the forest and it is associated to a value ('**gammaness**'), depending on how many times ended to a "**h**" or "**g**" leaf.