GAMMA-RAY DATA COLLECTION, CALIBRATION AND ANALYSIS (FOR LIV STUDIES)

J. Bolmont - LPNHE - Sorbonne Université
bolmont@in2p3.fr

COST Action CA18108 2nd Training School - Belgrade - September 2022

ABOUT ME

- Gamma-ray astronomy since 2002, LIV studies since 2003
- Associate Professor since 2010 - Sorbonne Université - Paris
- Member of HESS and CTA
- Research activities:
  - Search for LIV-induced energy-dependent time-delays with high-energy gamma-ray sources
  - Study of source-intrinsic effects in flaring blazars
- Past technical activities
  - On-call expert for HESS camera operation
  - Maintenance of HESS cameras

ABOUT THIS COURSE

- Preliminary comment:
  - This course will be entirely focused on ground-based high-energy gamma-ray astronomy
- Four parts:
  - **From image collection to raw data**
  - **Calibration**
  - **Reconstruction**
  - **Analysis** + A brief look at Lorentz Invariance Violation analyses

- Before that: some basics, to get familiar with some of the vocabulary

NOTES/CAVEATS

- This lecture is mostly based on my experience with HESS
- Other IACT Arrays can use different implementation of the same basic methods/ functionalities

- A huge thanks to HESS colleagues who helped me gathering the material for this lecture: M. De Naurois, J.-P. Lenain, F. Toussenel, S. Ohm, a lot of PhD students, and many other colleagues…

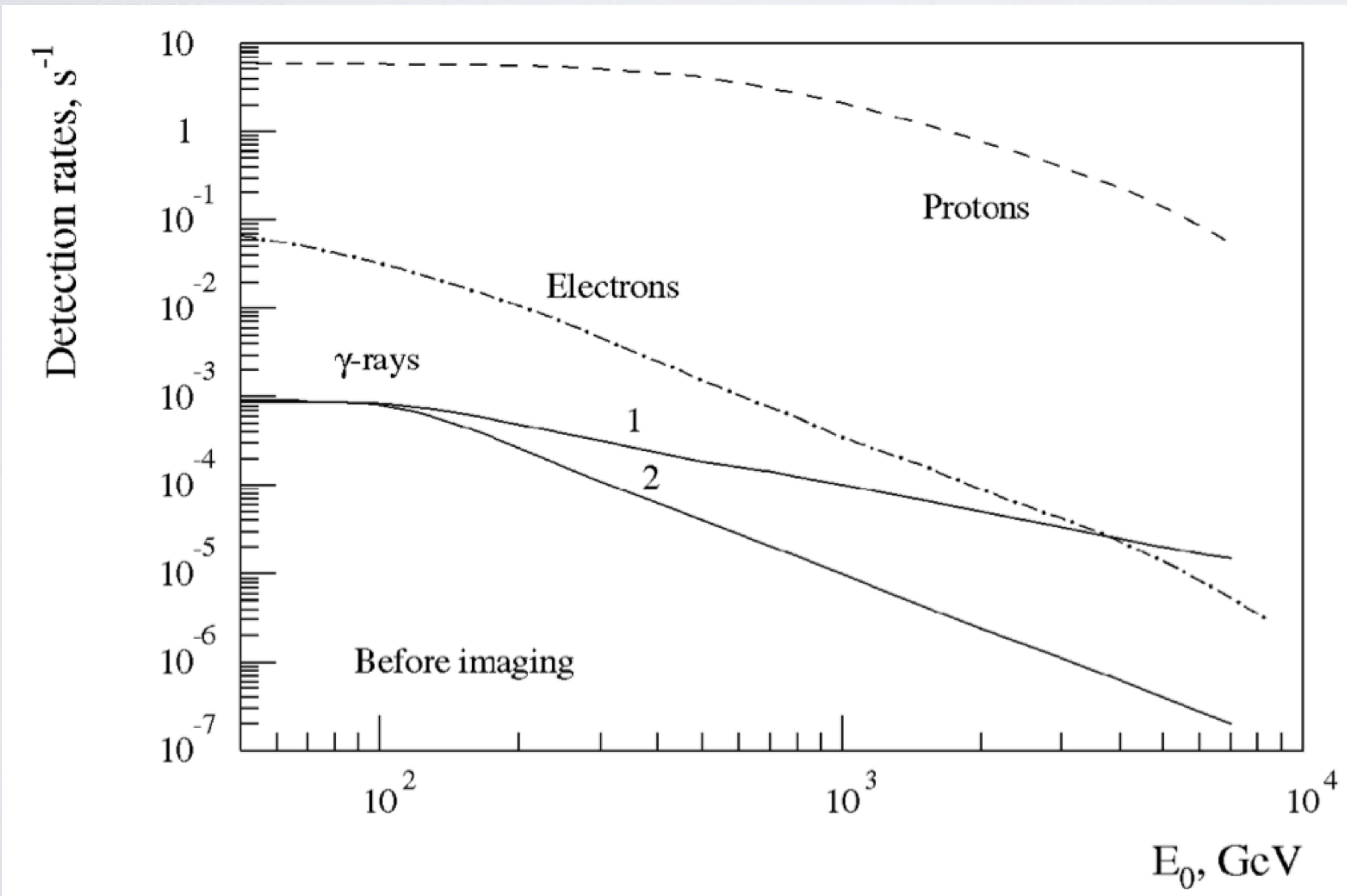
SOME (QUICK) BASICS

SOME QUICK BASICS

- Three major experiments/Cherenkov telescope arrays are currently operating
- The goal is to study the VHE sky
- For this, we **take pictures** of the Cherenkov emission from cascades developing in the atmosphere
    - Atmosphere is a part of the detector
- The signal is **faint** and very **short**
- So, to build a Cherenkov telescope, you need
    - A **large mirror** to collect Cherenkov light
    - A **fast** and **sensitive** camera in the focal plane
- Additional constraints exist!

SOME QUICK BASICS

- We are interested in **photons** produced by an astrophysical source
- Dominant background: **hadrons**



From Aharonian & Akerlof, 1997

γ-ray spectra 1 and 2 correspond to spectral indices of 2 and 3 respectively.

SOME QUICK BASICS

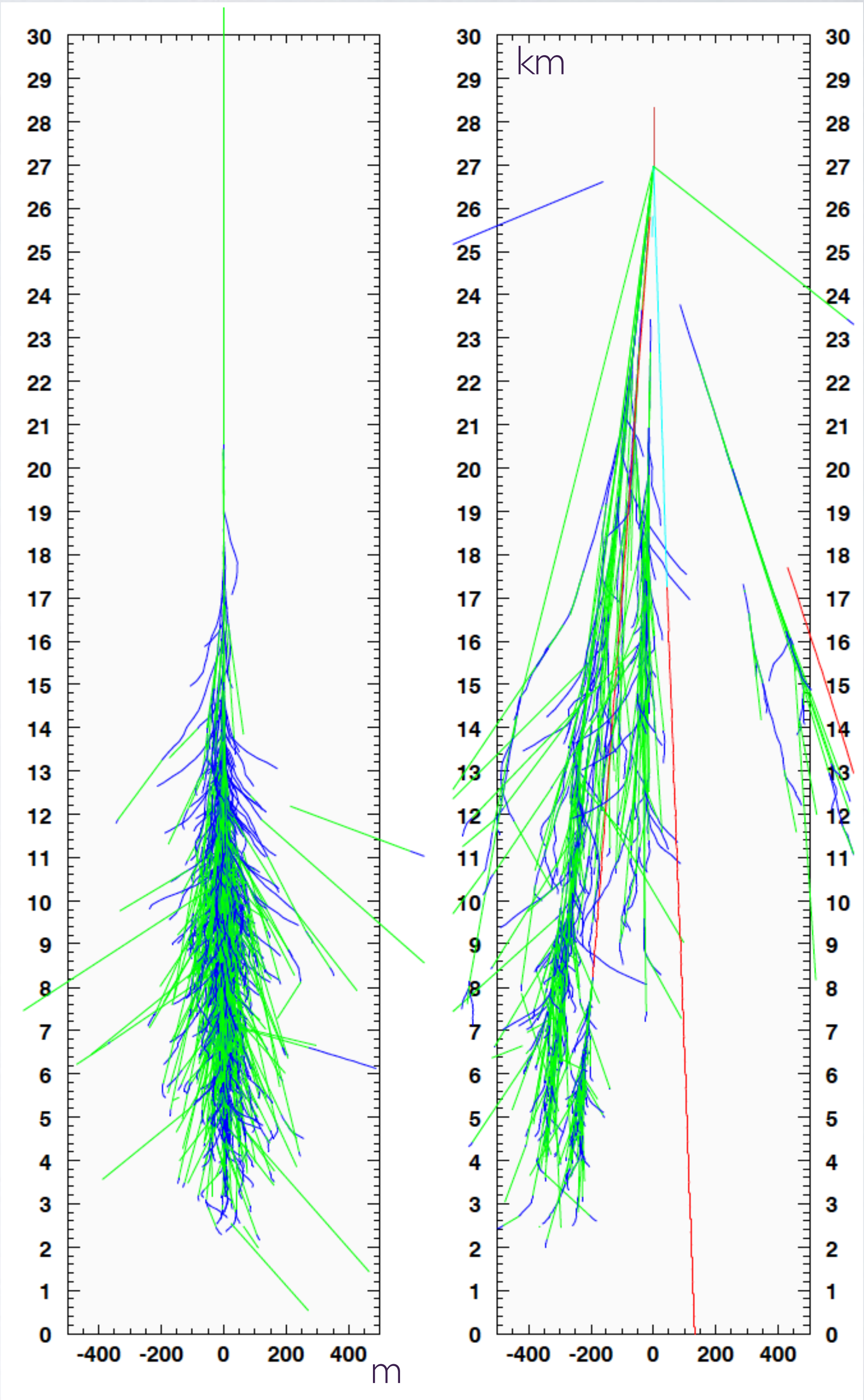
- **Detecting** a particle means
  - **Identifying** it → gamma-ray or hadron?
  - **Characterizing** it → energy, time, direction, impact location
- Notion of **event**
  - A gamma-ray (or a hadron) interacts in the atmosphere, the Cherenkov light is reflected by the mirror and collected by a camera of HESS
  - Necessary condition: the camera is **triggered**
- **Raw data** is recorded only when the camera triggers
  - For each channel (pixel), amplified integrated pulses converted into ADC counts + other parameters
- **Calibration** is necessary to convert ADC counts to **a charge, and assess optical efficiency** → multiple ingredients required!

SOME QUICK BASICS

- **Reconstruction** is used for
  - Discrimination/Identification → **discriminating variables**
    - Selection cuts allows to obtain a sample of **gamma-like events**
    - Different cuts will give different amounts of purity for that sample
  - Characterization
    - Several techniques can be used: Hillas, Model,…
- **Analysis** is used on gamma-like events for
  - Precise **background estimation** and subtraction
  - Evaluation of **statistical significance**
  - **Signal extraction**: energy spectrum, light curve, morphology
    - Sometimes long observations are needed
    - Reconstructed event list can also be provided at this step
- **If the source is suitable, only then can we attempt a LIV analysis**

SOME QUICK BASICS

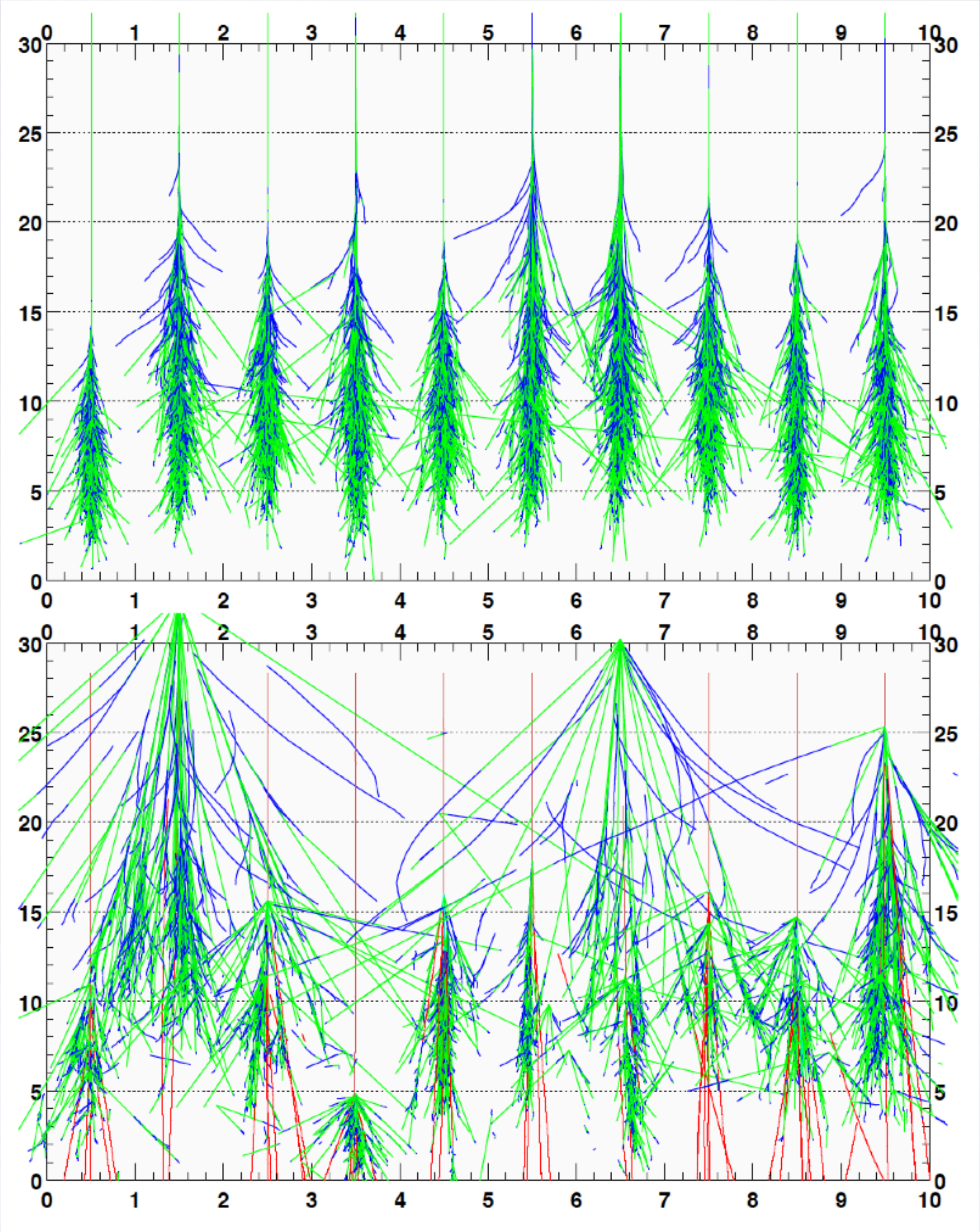
- We are interested in **photons** produced by an astrophysical source
- Dominant background: **hadrons**
- The signal:
  - Primary cosmic rays hit the atmosphere
  - They produce a shower/cascade of secondary particles
  - Charged secondary particles produce Cherenkov light
- A shower develops within O(10 ns)
  - The camera and electronics need to be **fast** enough!



300 GeV photon (left) and hadron (right)

SOME QUICK BASICS

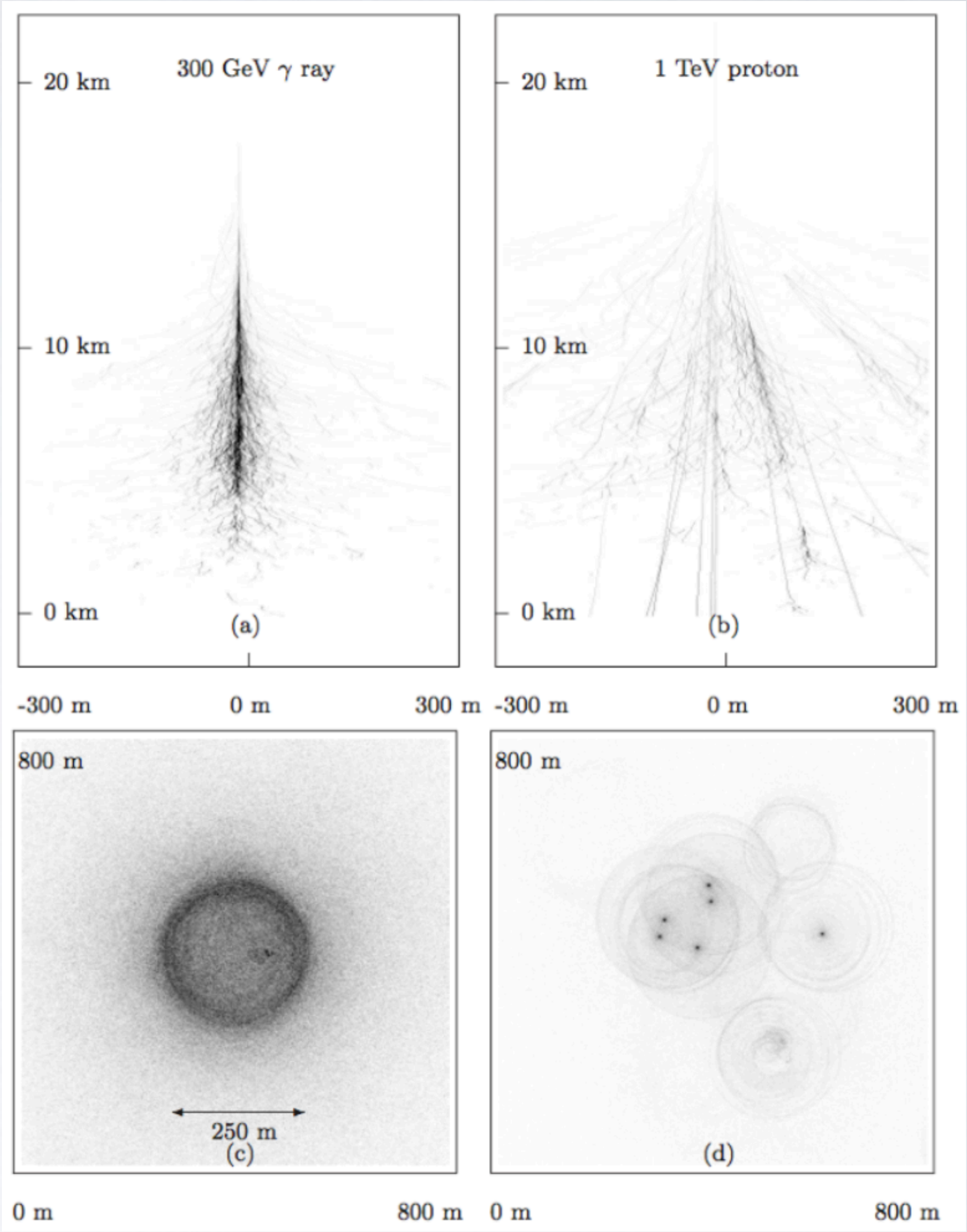
- We are interested in **photons** produced by an astrophysical source
- Dominant background: **hadrons**
- The signal:
  - Primary cosmic rays hit the atmosphere
  - They produce a shower/cascade of secondary particles
  - Charged secondary particles produce Cherenkov light
- A shower develops within O(10 ns)
  - The camera and electronics need to be **fast** enough!



300 GeV photons (top) and hadrons (bottom)

SOME QUICK BASICS

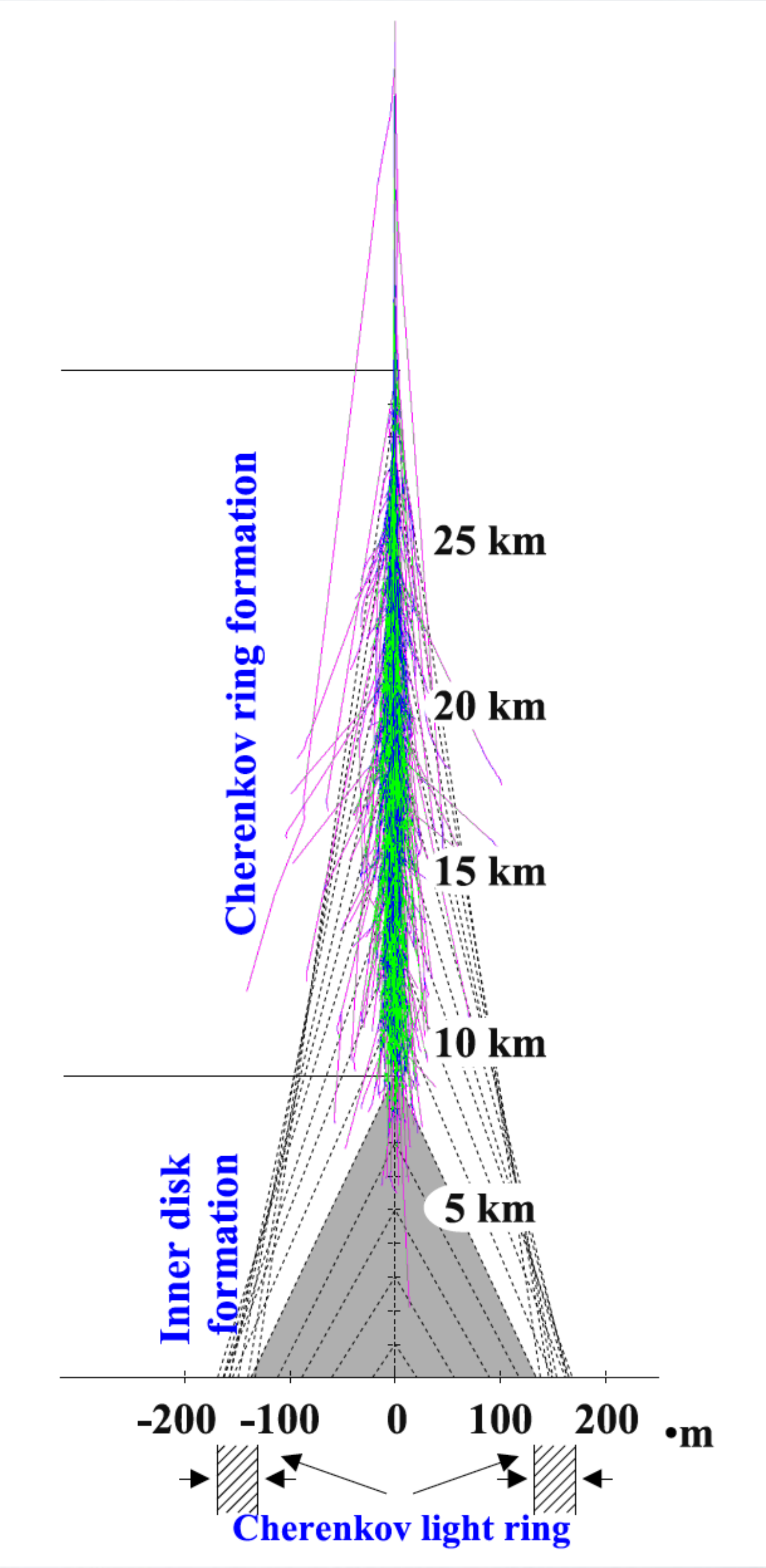
- We are interested in **photons** produced by an astrophysical source
- Dominant background: **hadrons**
- The signal:
  - Primary cosmic rays hit the atmosphere
  - They produce a shower/cascade of secondary particles
  - Charged secondary particles produce Cherenkov light
- A shower develops within O(10 ns)
  - The camera and electronics need to be **fast** enough!



➜ More on cascades during the tutorial!

SOME QUICK BASICS

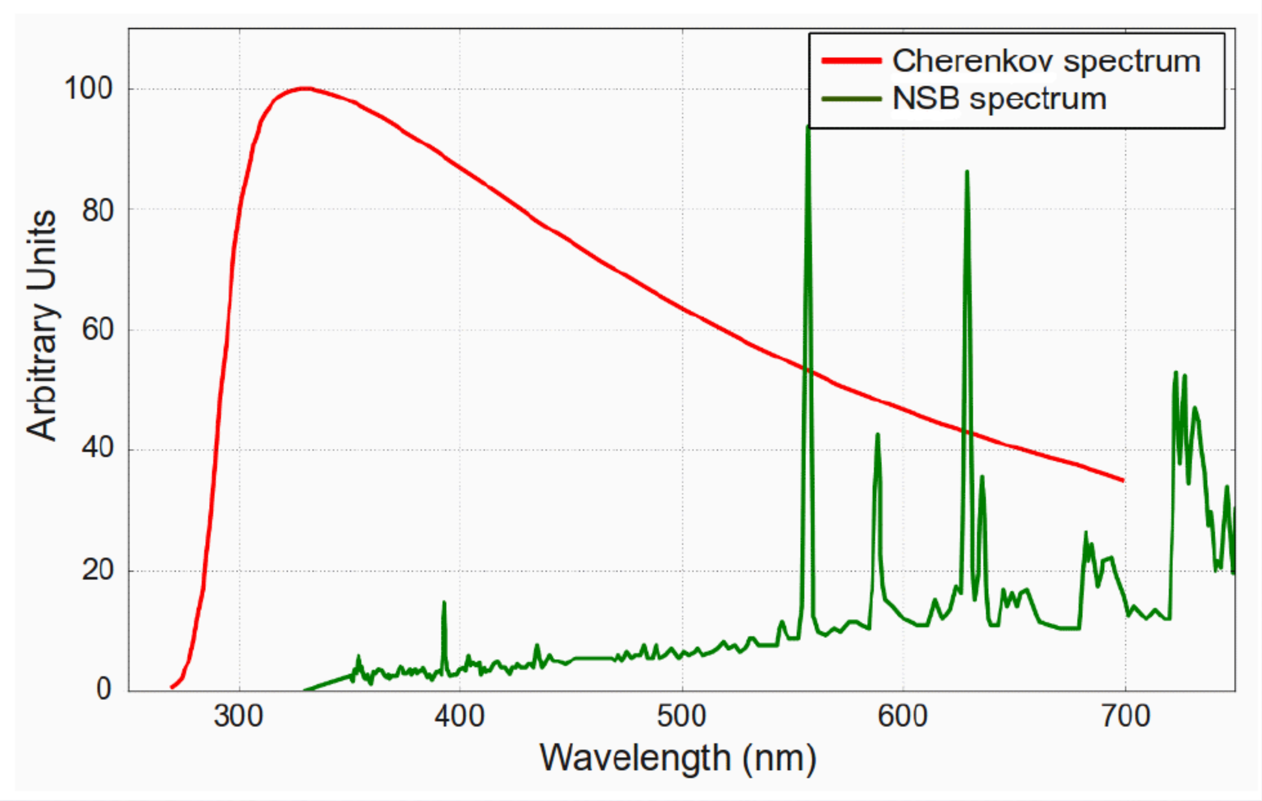
- We are interested in **photons** produced by an astrophysical source
- Dominant background: **hadrons**
- The signal:
  - Primary cosmic rays hit the atmosphere
  - They produce a shower/cascade of secondary particles
  - Charged secondary particles produce Cherenkov light
- A shower develops within O(10 ns)
  - The camera and electronics need to be **fast** enough!



➜ More on cascades during the tutorial!

SOME QUICK BASICS

- Let's be more specific!
- We are interested in **Cherenkov photons** from the cascades which peaks around 350 nm
- But photo-detectors are also sensitive to « normal » light in their range of operation: **Night Sky Background** (NSB)
- Photodetectors have to be carefully chosen

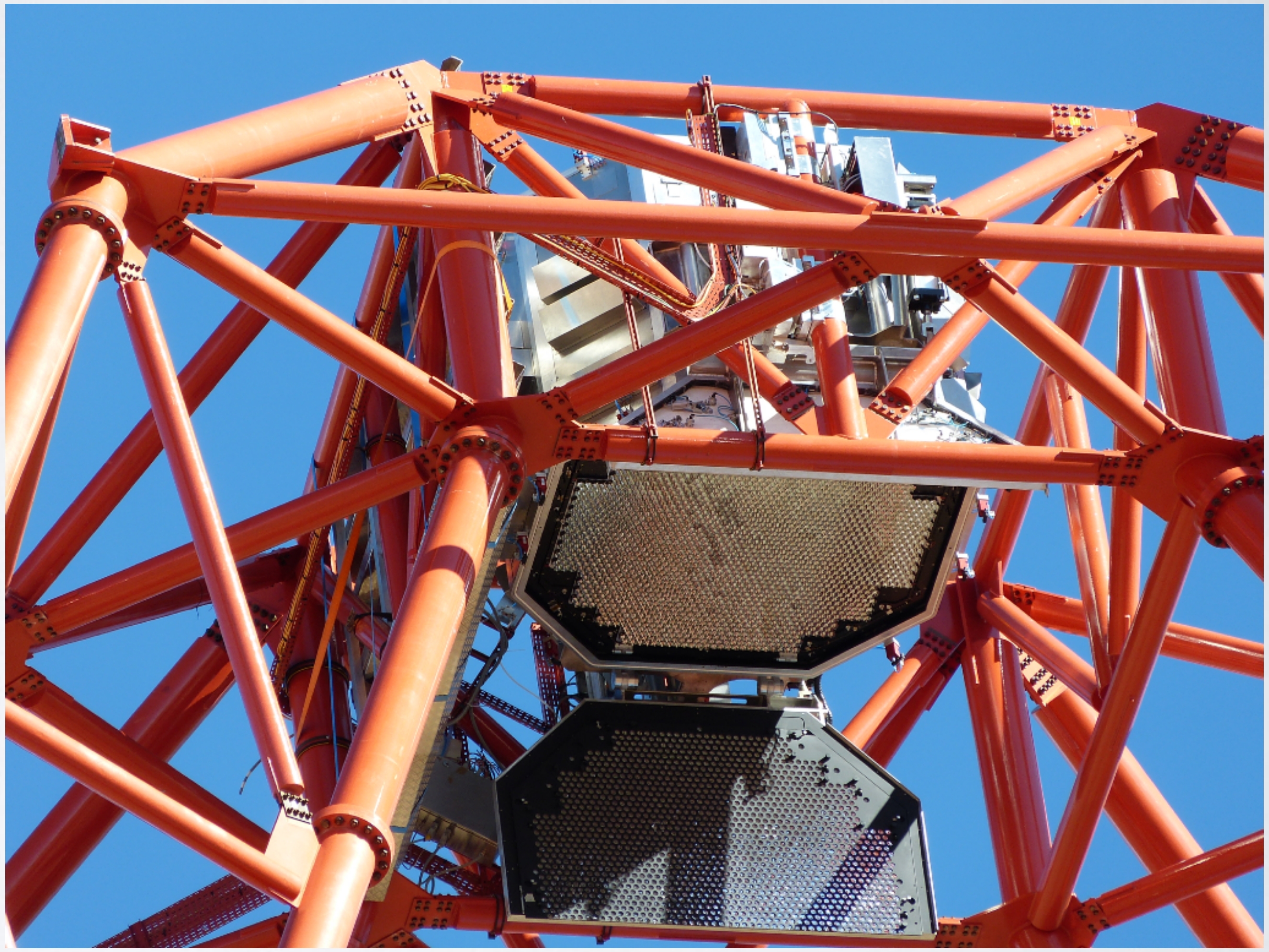


PART 1

FROM IMAGE COLLECTION TO RAW DATA

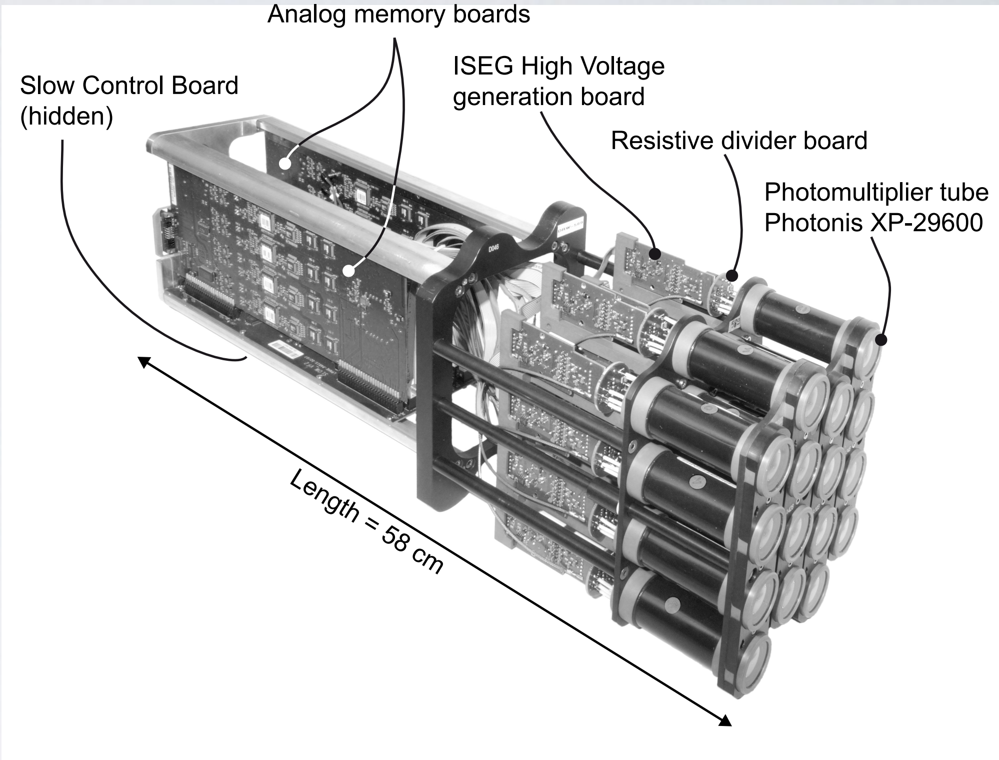
THE CAMERA

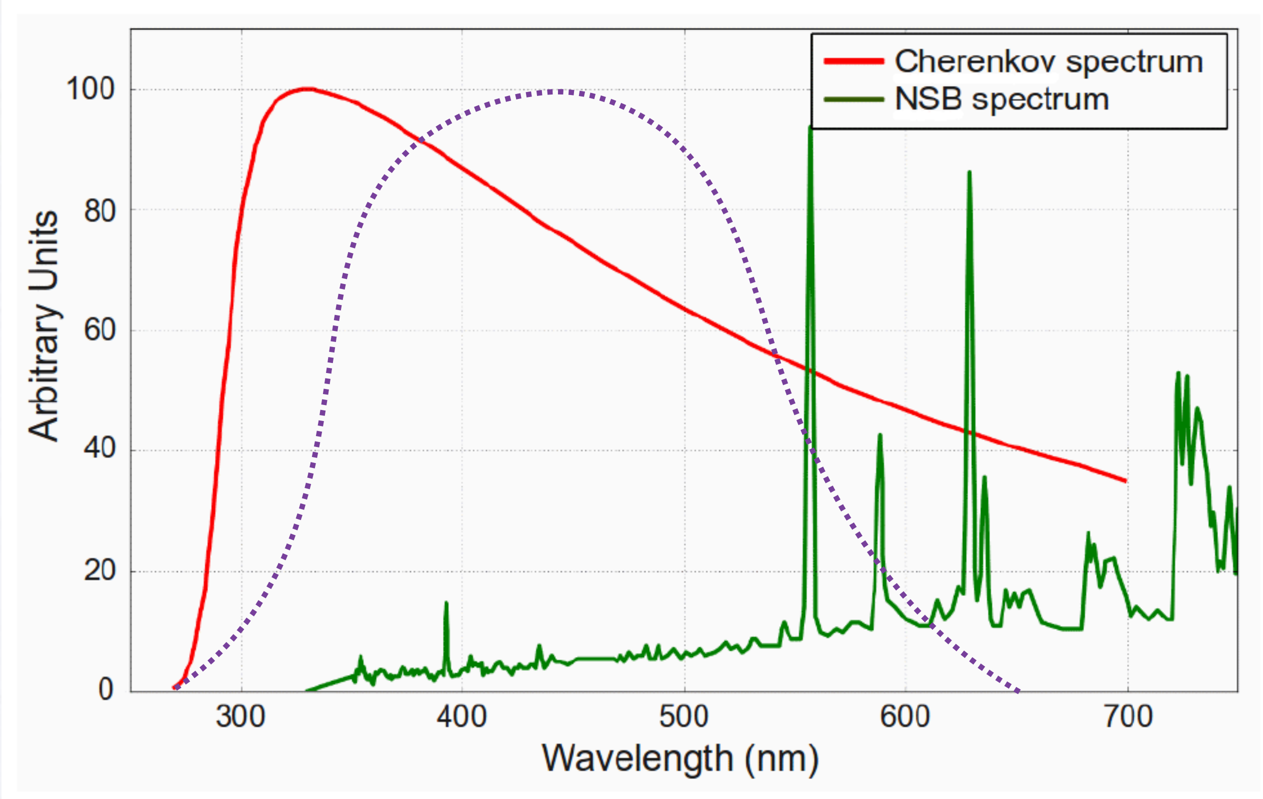
- Characteristics of the pre-upgrade CT5 camera (2012-2019)
  - $2{,}40 \times 2{,}27 \times 1{,}84$ m^3
  - Field of view 3.2°
  - 128 drawers with 16 PMTs each
  - 2048 pixels
  - PMT diameter 29 mm
  - PMT spectral range 270-650 nm
  - Peak QE of 30% at 420 nm
  - Gain 2×10^5 at ~800 V
  - 50 ADC counts for 1 p.e.
  - Integration window width 16 ns
  - Dead time 15 µs
  - Maximum operational trigger rate 5000 Hz



THE CAMERA

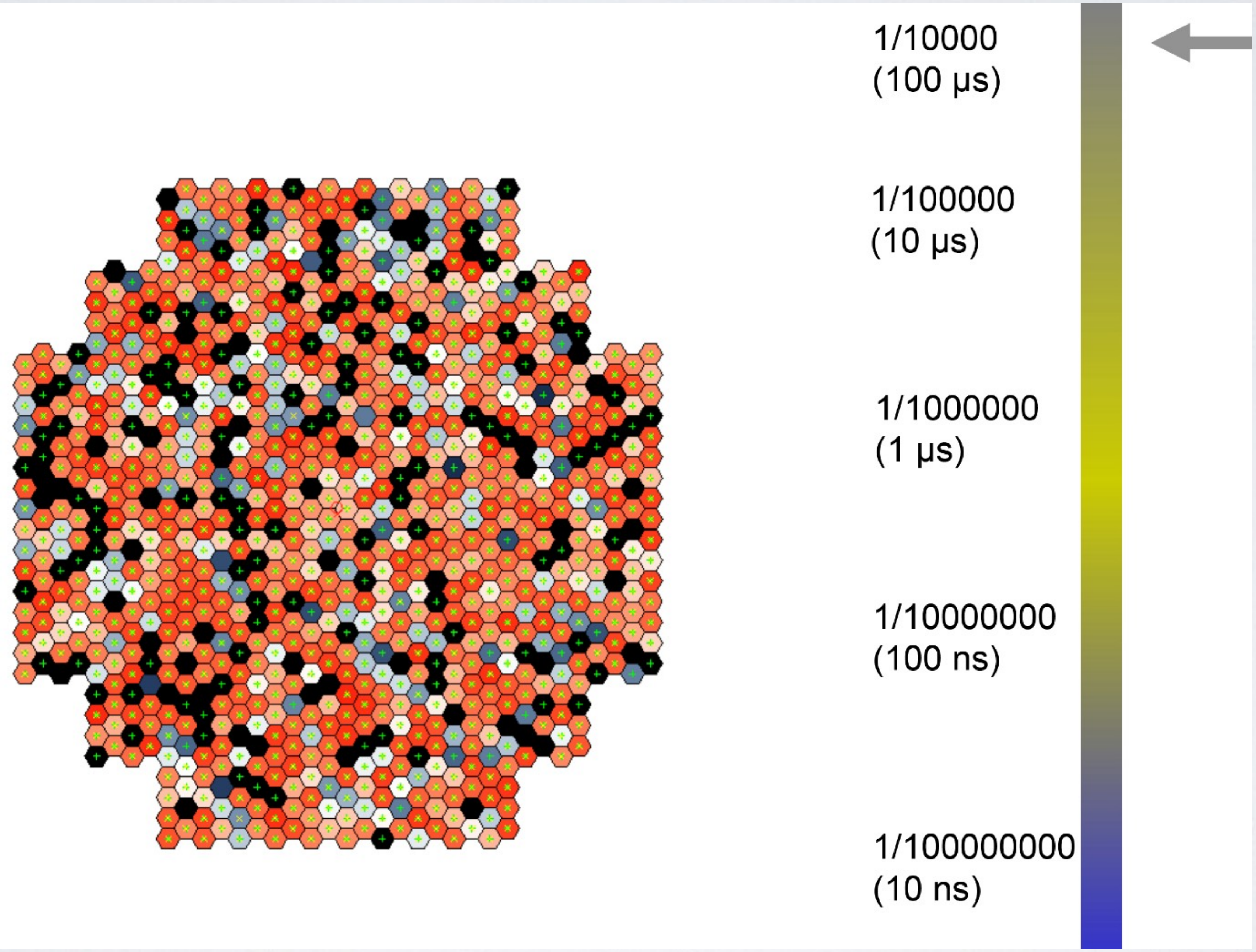
- Characteristics of the pre-upgrade CT5 camera (2012-2019)
  - 2,40 x 2,27 x1,84 m^3
  - Field of view 3.2°
  - 128 drawers with 16 PMTs each
  - 2048 pixels
  - PMT diameter 29 mm
  - PMT spectral range 270-650 nm
  - Peak QE of 30% at 420 nm
  - Gain 2×10^5 at ~800 V
  - 50 ADC counts for 1 p.e.
  - Integration window width 16 ns
  - Dead time 15 μs
  - Maximum operational trigger rate 5000 Hz





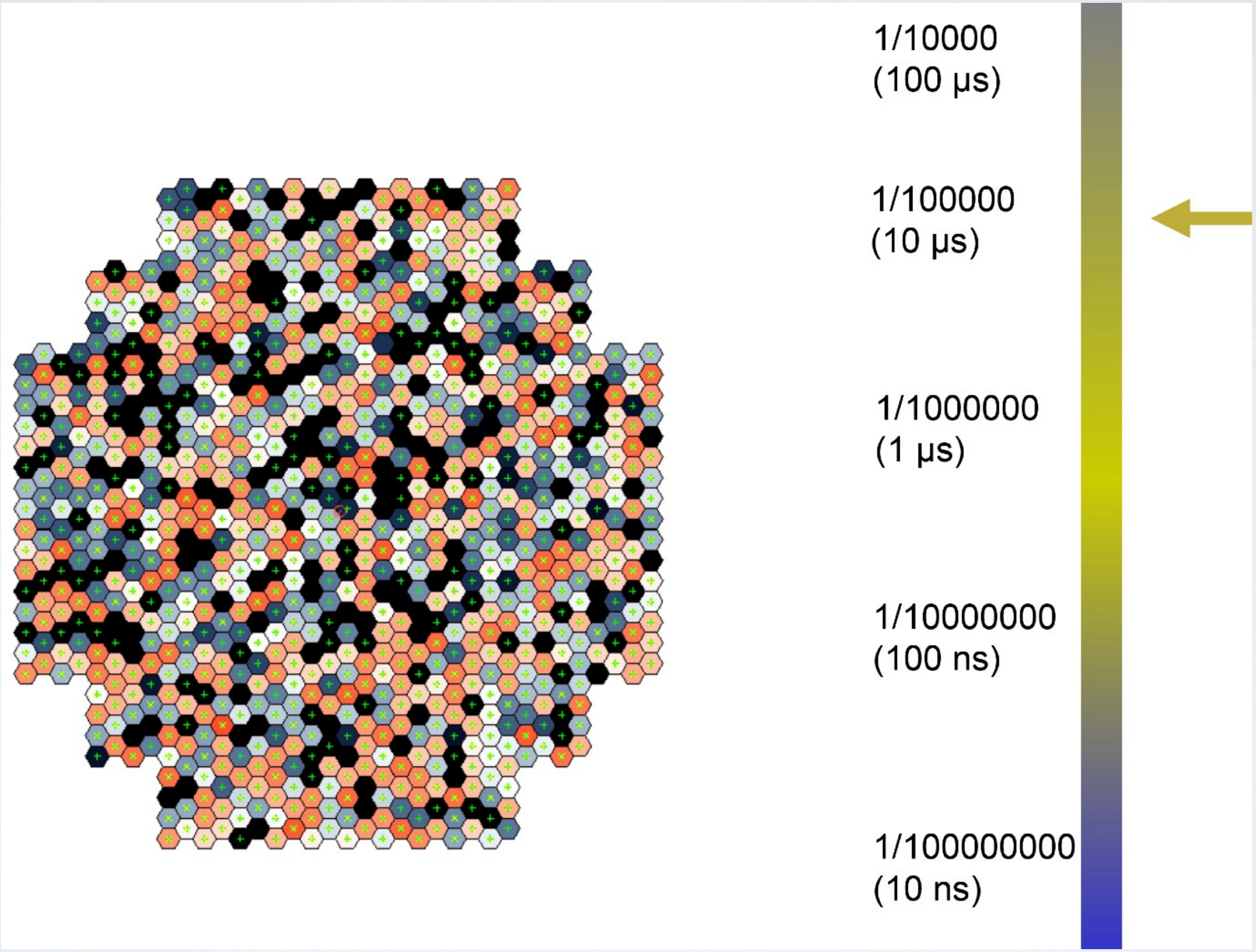
THE CAMERA

- The camera needs to be fast!



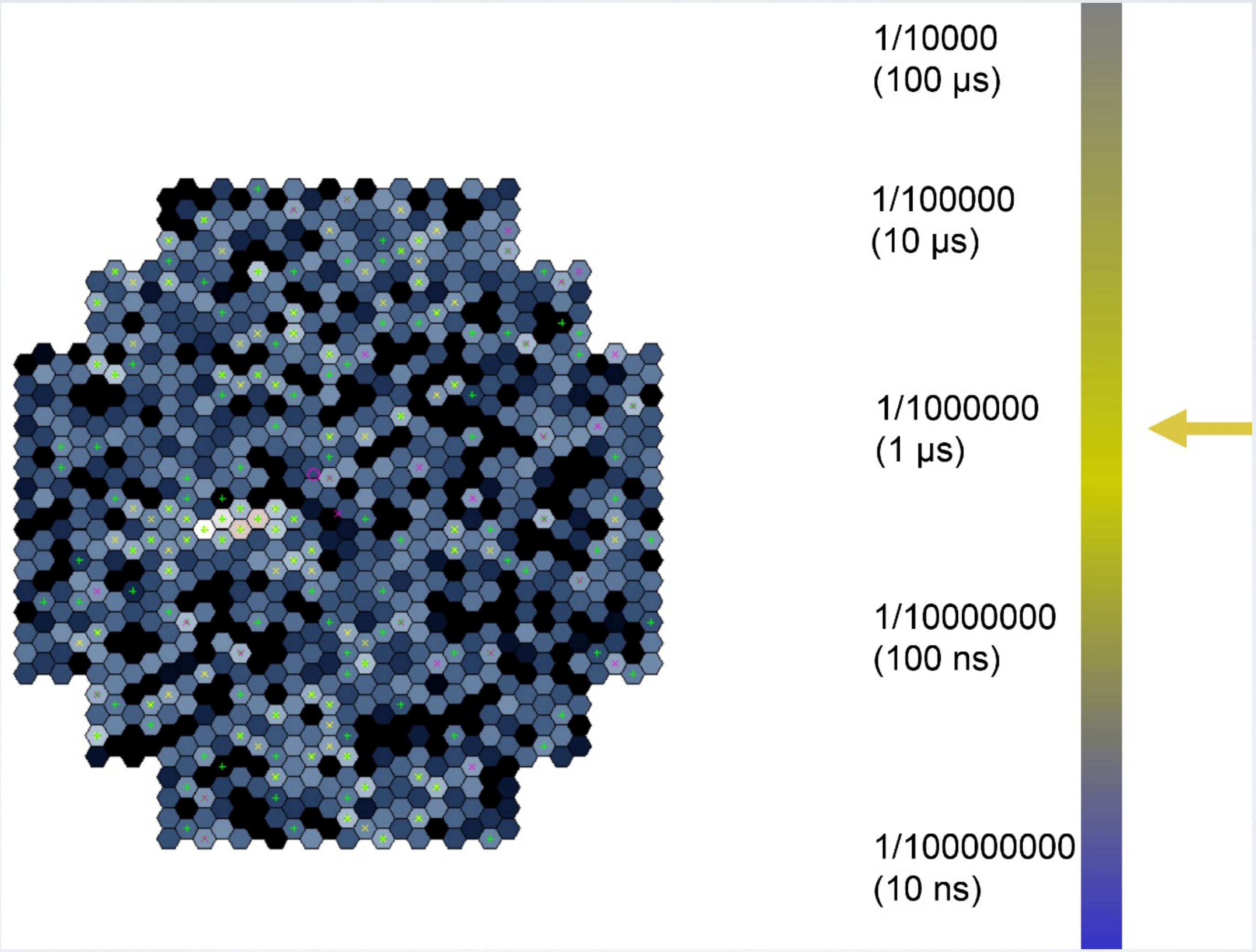
THE CAMERA

- The camera needs to be fast!



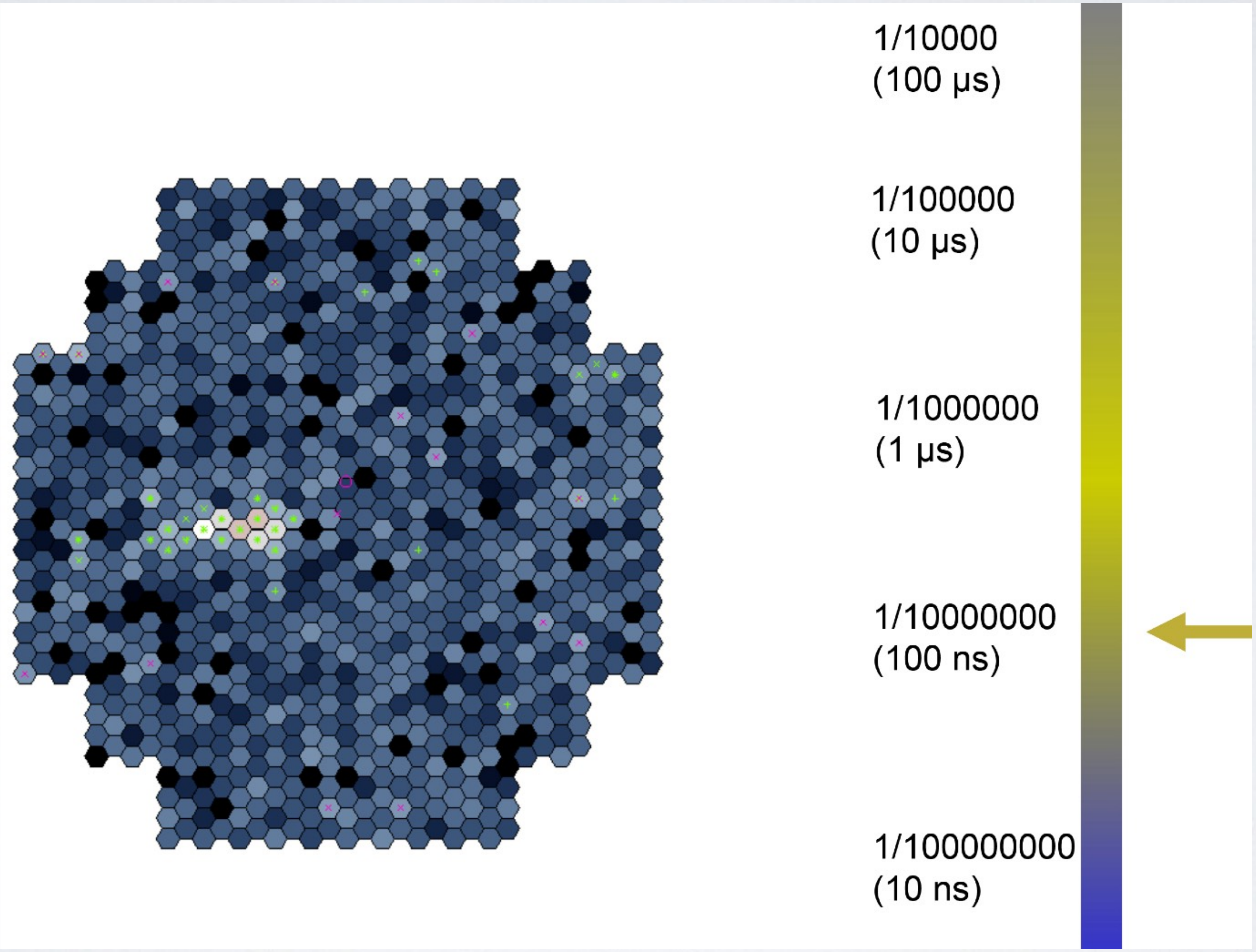
THE CAMERA

- The camera needs to be fast!



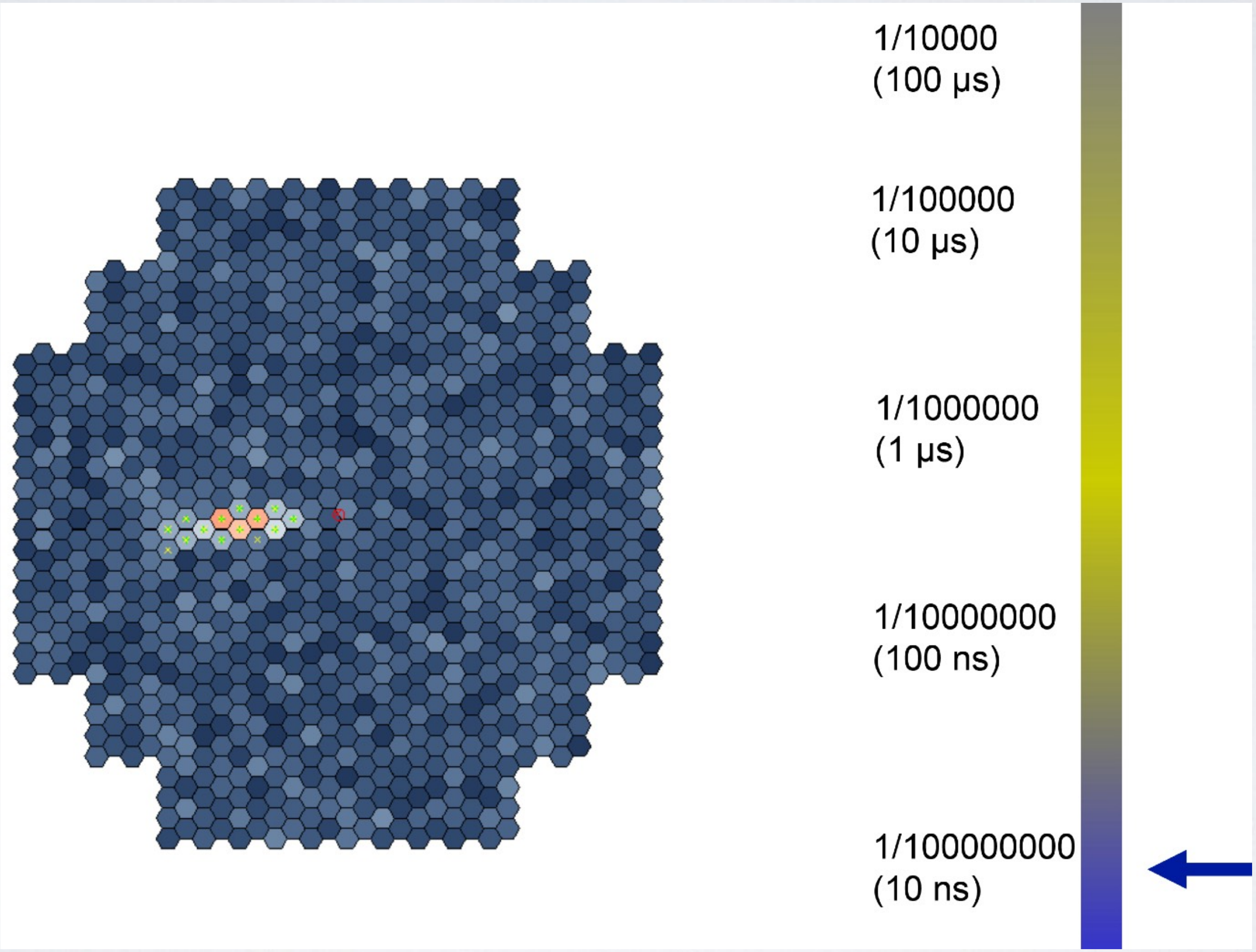
THE CAMERA

- The camera needs to be fast!



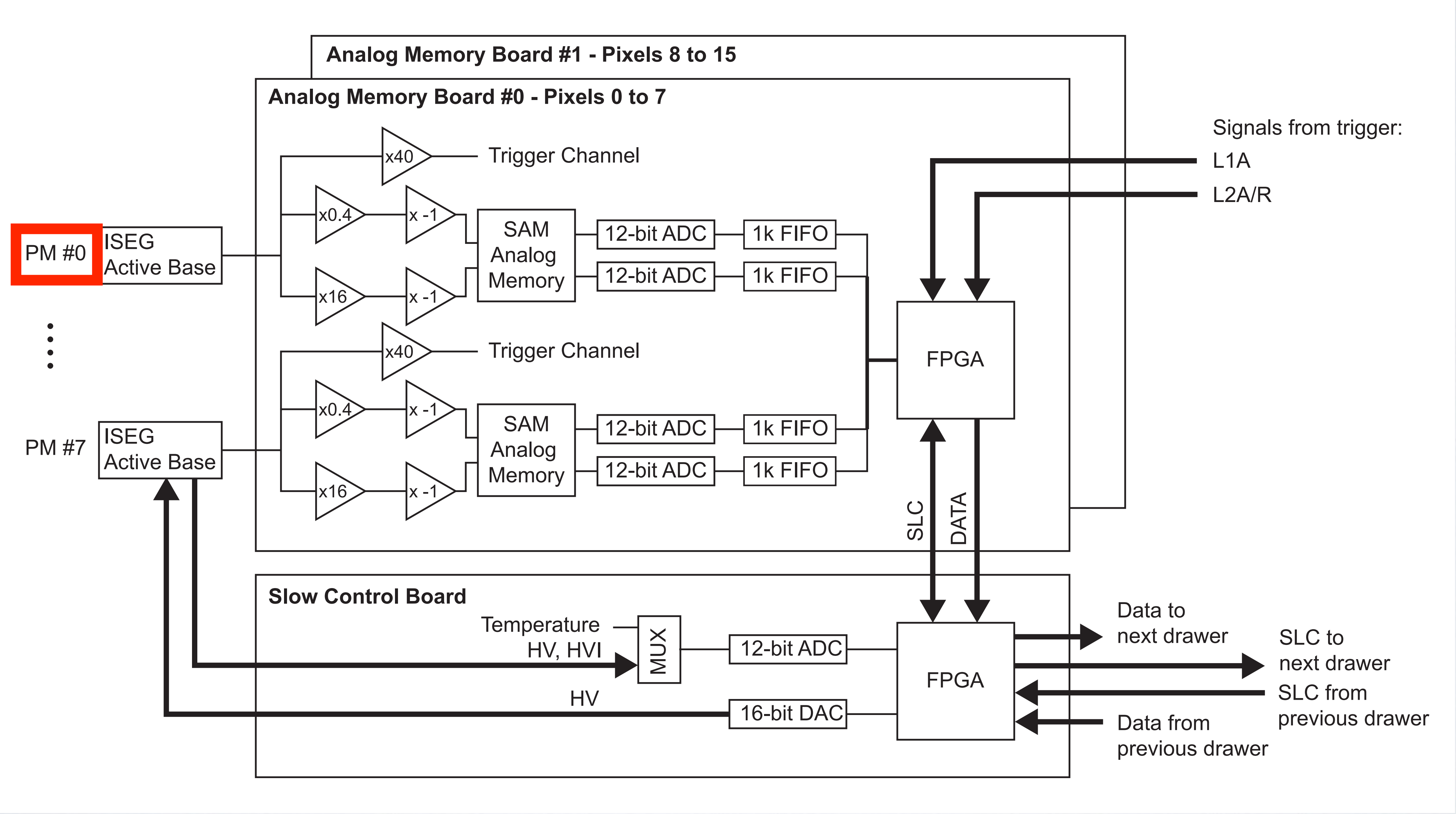
THE CAMERA

- The camera needs to be fast!



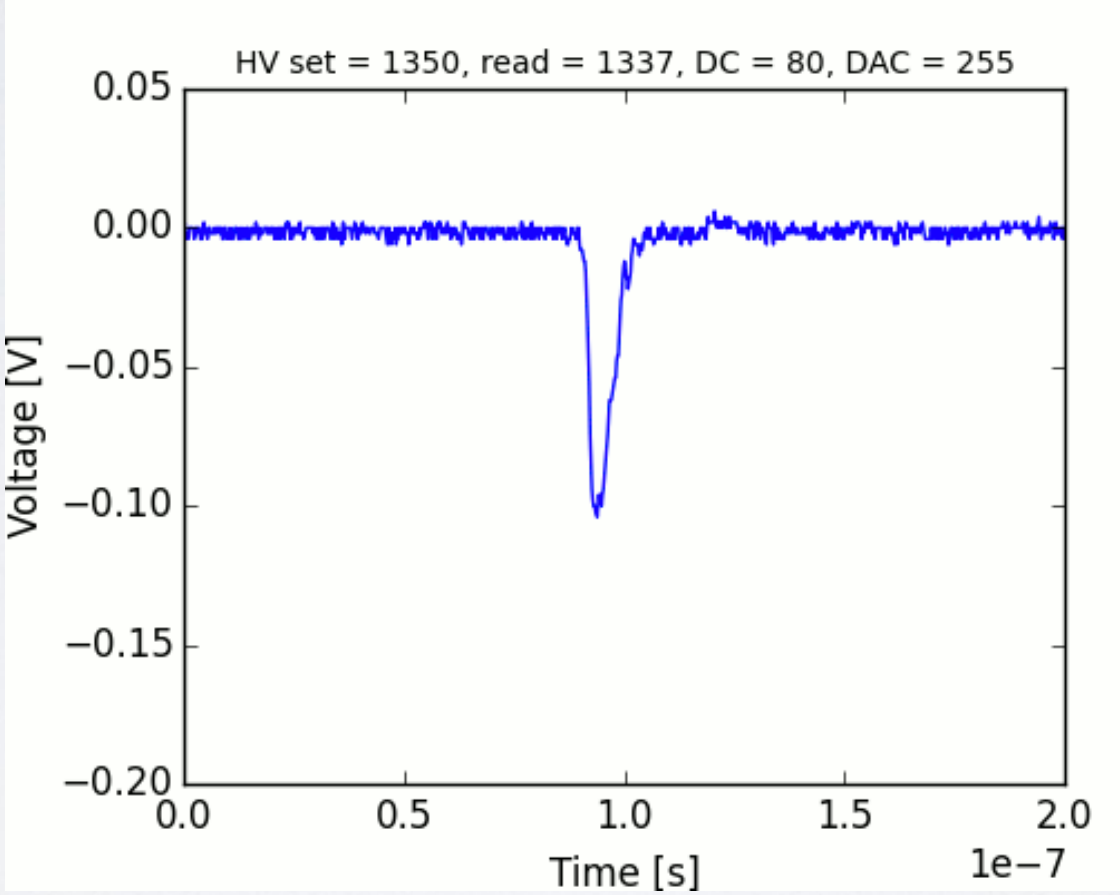
DATA ACQUISITION

- Simplified view of front-end electronics for HESS CT5 camera
  - View of 1 drawer



PMT

- The PMT signal looks like this:

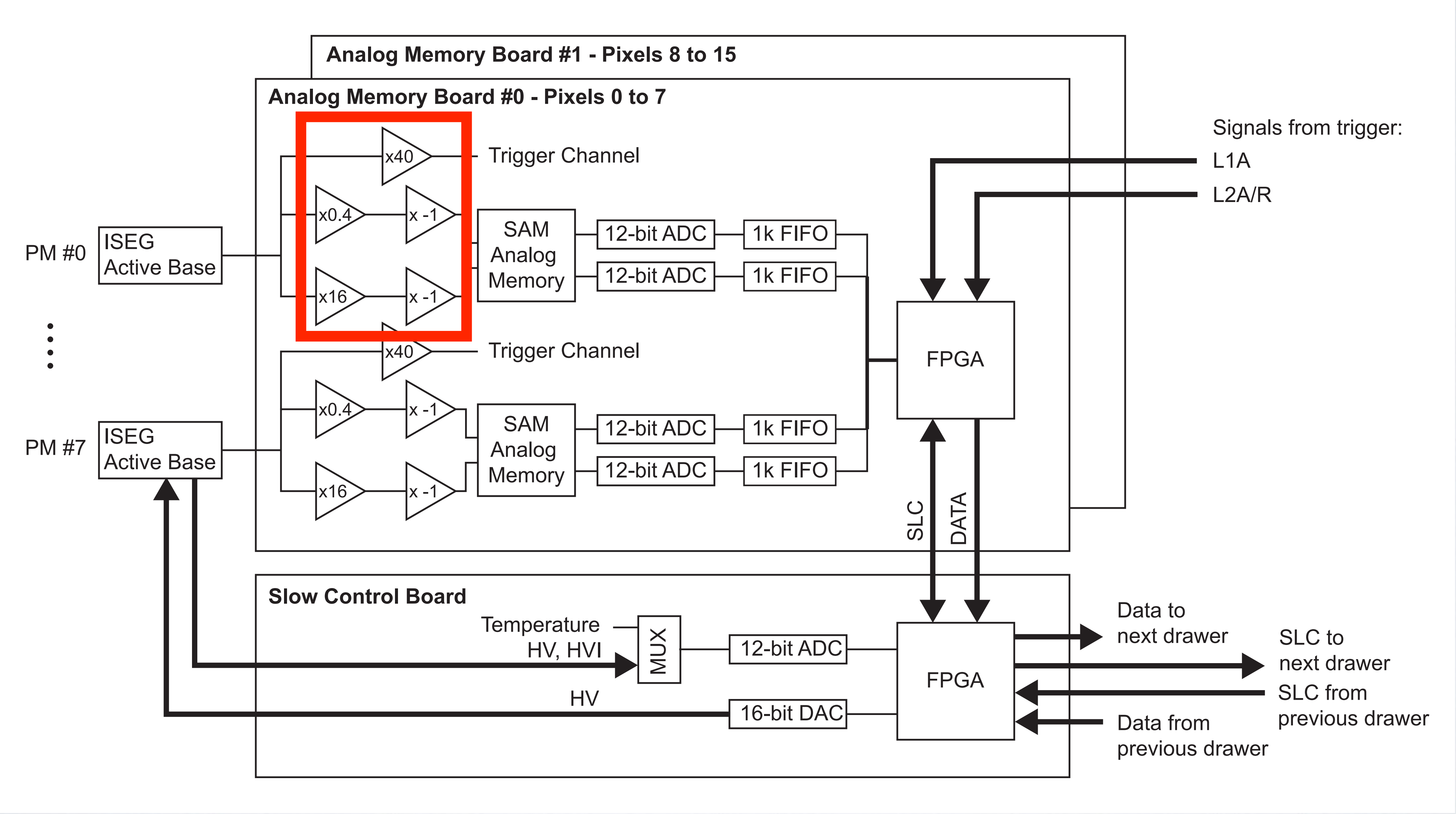


NB: here we see only pulses which triggered the scope, so with an amplitude exceeding a fixed threshold

  - Typical width for HESS PMTs is 3.5 ns
  - The **pulse integral is proportional to the amount of light received**
- PMT bases have a protection system so that HV is switched-off when illumination is too high

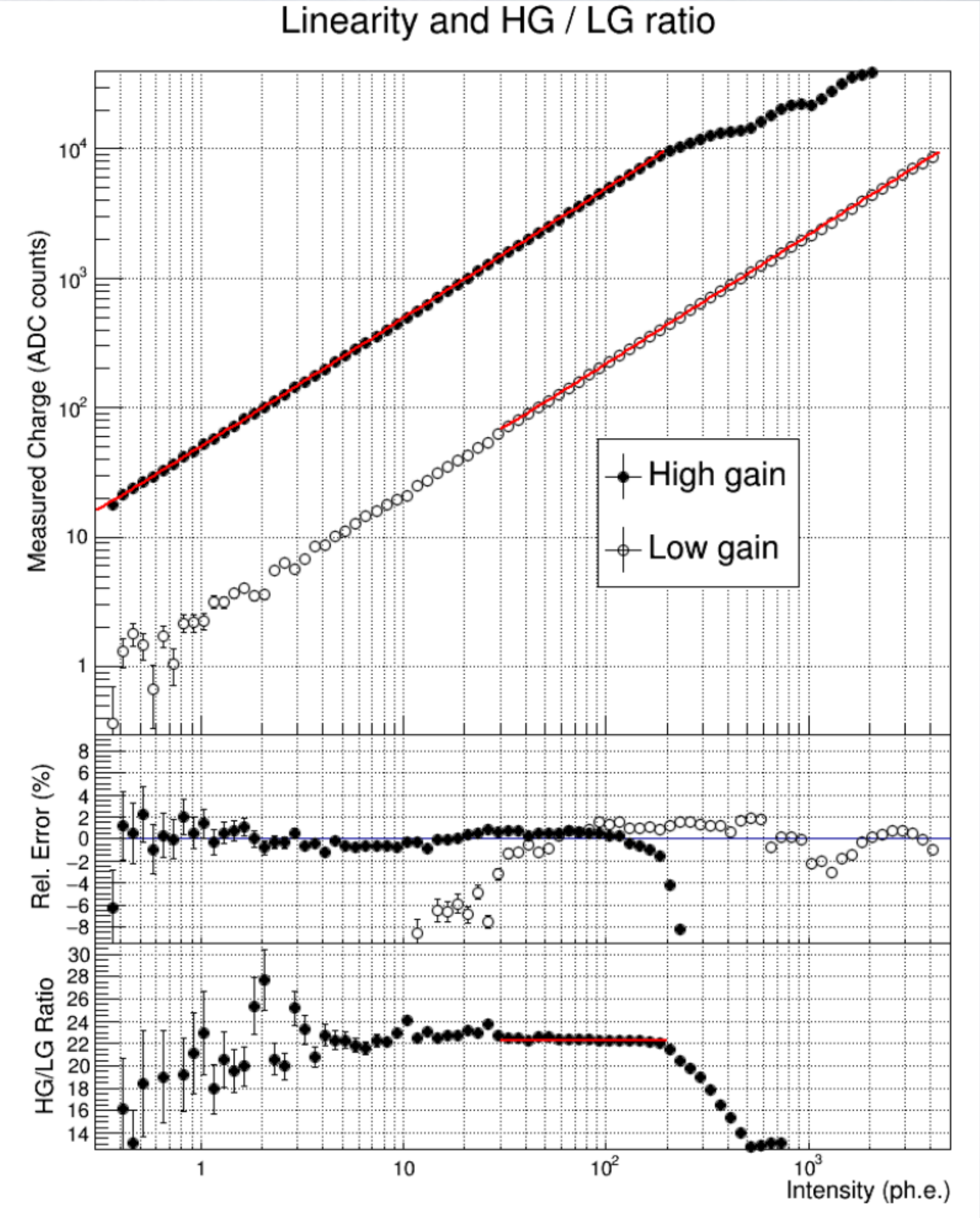
DATA ACQUISITION

- Simplified view of front-end electronics for HESS CT5 camera
  - View of 1 drawer



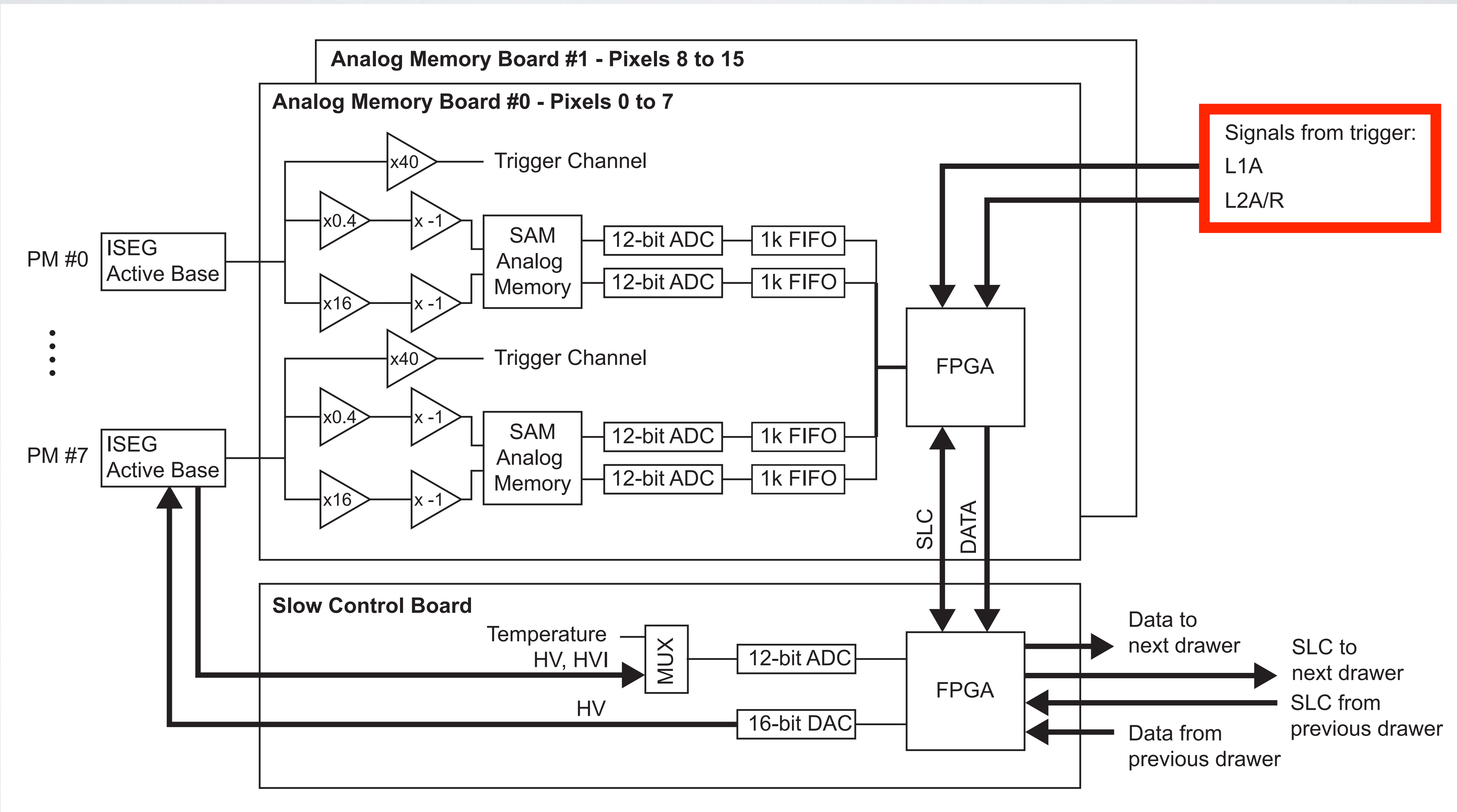
AMPLIFIERS

- **Linear** amplifiers
  - **High gain** (x16) amplification for low amplitude signals (< 200 p.e.)
    - 1 p.e. → 7 mV amplitude
  - **Low gain** (x0.4) amplification for high amplitude signals (< 5000 p.e.)
  - Very high gain (x40) for the **trigger channel**
- Note that
  - Signal is inverted and baseline is shifted towards negative voltages to use the ADC full input range
  - Impedance matching is crucial all along the signal path



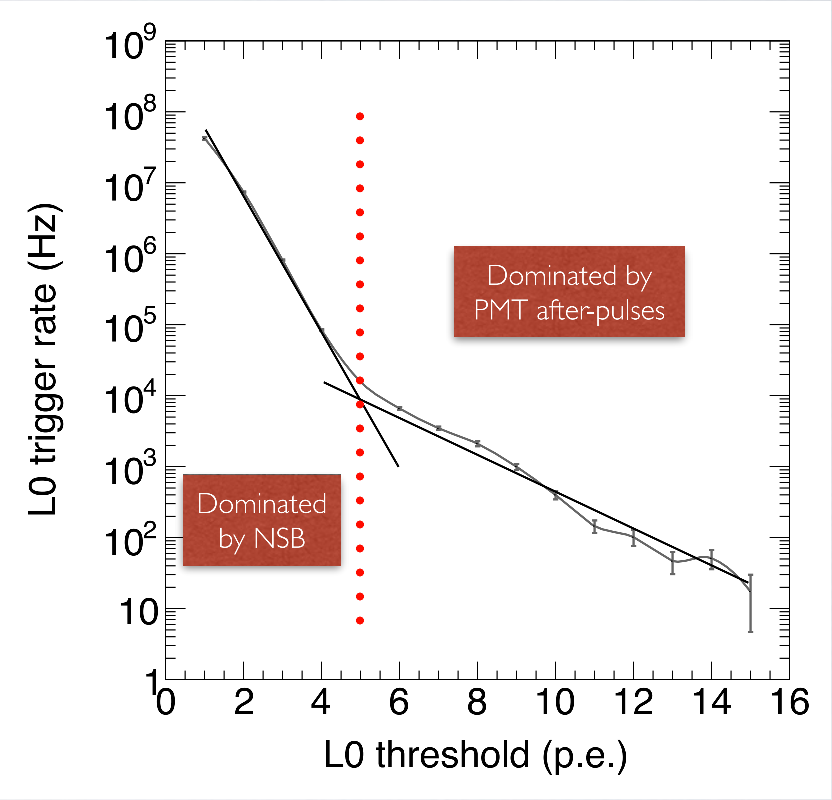
DATA ACQUISITION

- Simplified view of front-end electronics for HESS CT5 camera
  - View of 1 drawer



TRIGGER

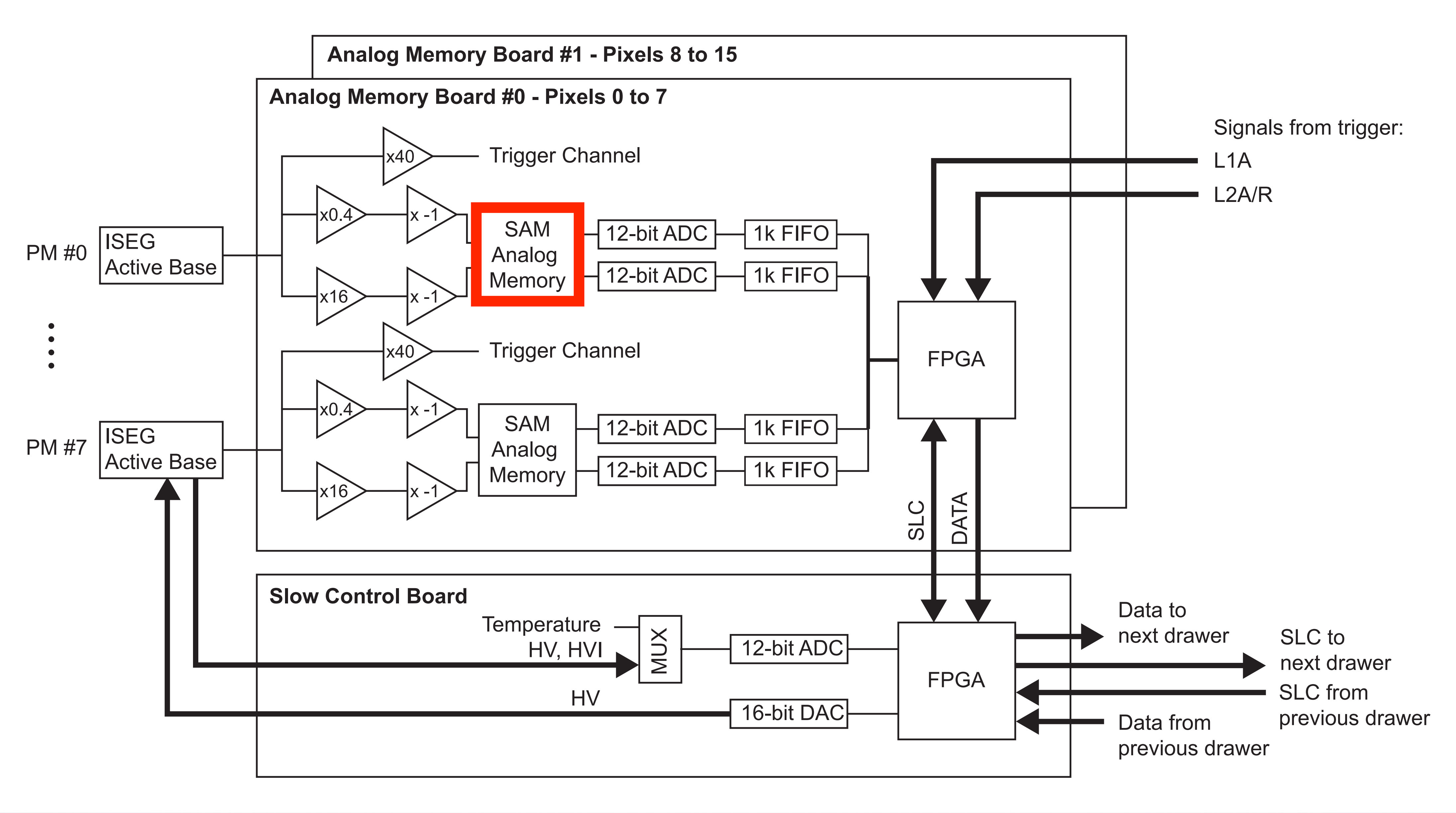
- The array uses **camera triggers** and a **central trigger**
- Different levels of trigger are used
  - **L0**: pixel level → The charge in a pixel exceeds a given threshold (4 p.e.)
  - **L1**: camera level → A number (≥ 3) of close-by pixels fulfill the L0 criteria
  - **L2**: array level
    - L2 trigger makes use of the **central trigger** output
    - Only CT5 has triggered → Monoscopic event
    - Another telescope has triggered within 80 ns → Stereoscopic event



- Different modes are used for different purposes
  - **Internal** trigger for normal data taking
  - **External** trigger for specific calibration runs (Single photo-electron runs)
  - **Software**-generated trigger for testing/warm-up

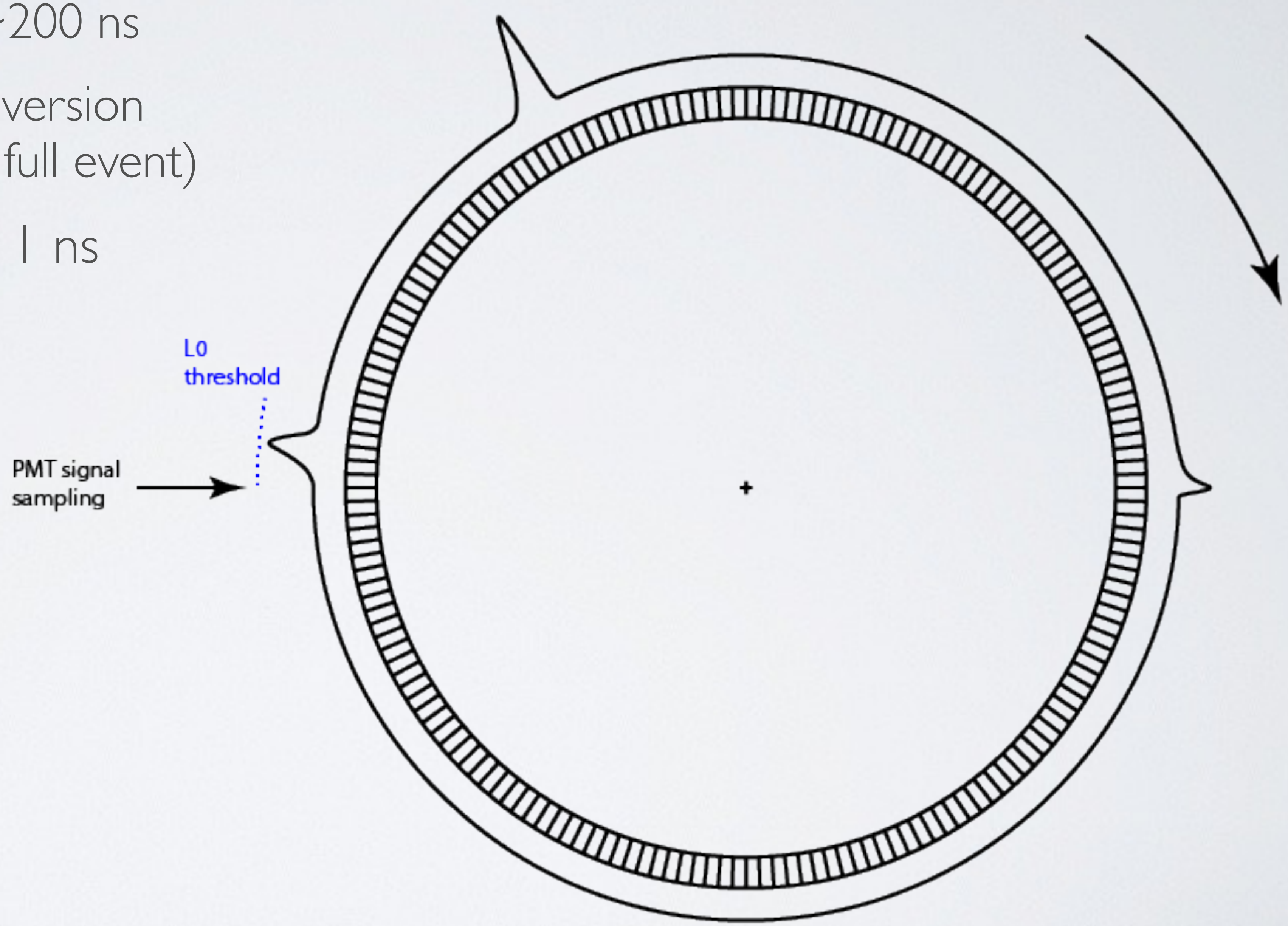
DATA ACQUISITION

- Simplified view of front-end electronics for HESS CT5 camera
  - View of 1 drawer



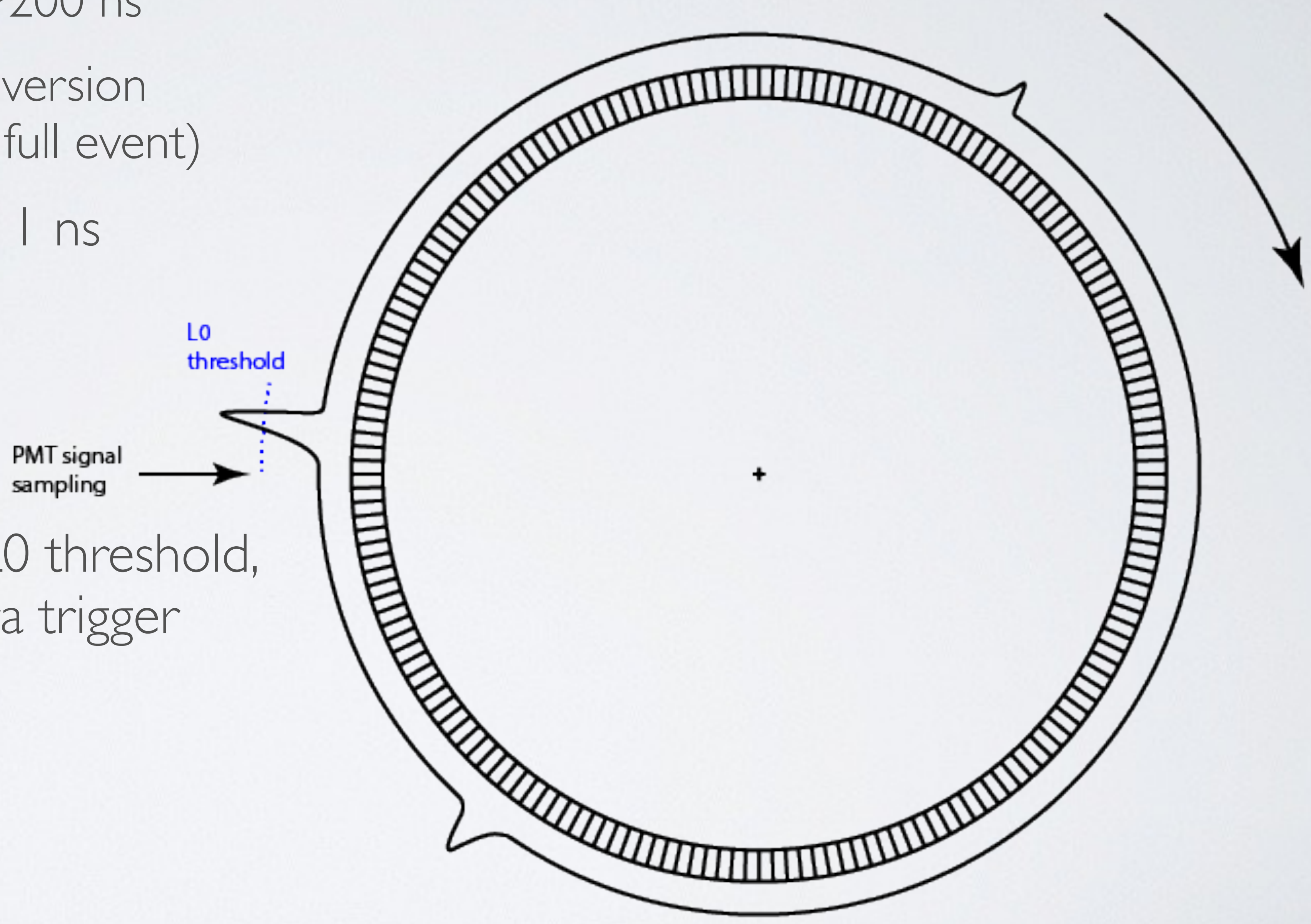
ANALOG MEMORIES

- Analog memory is used to store the signal for some time, until the trigger decision is taken
  - L1 trigger decision takes ~200 ns
  - Analog, because ADC conversion takes time (~250 μs for a full event)
- Signal is sampled at 1 GHz, 1 ns for each sample
- Circular buffer of 256 cells



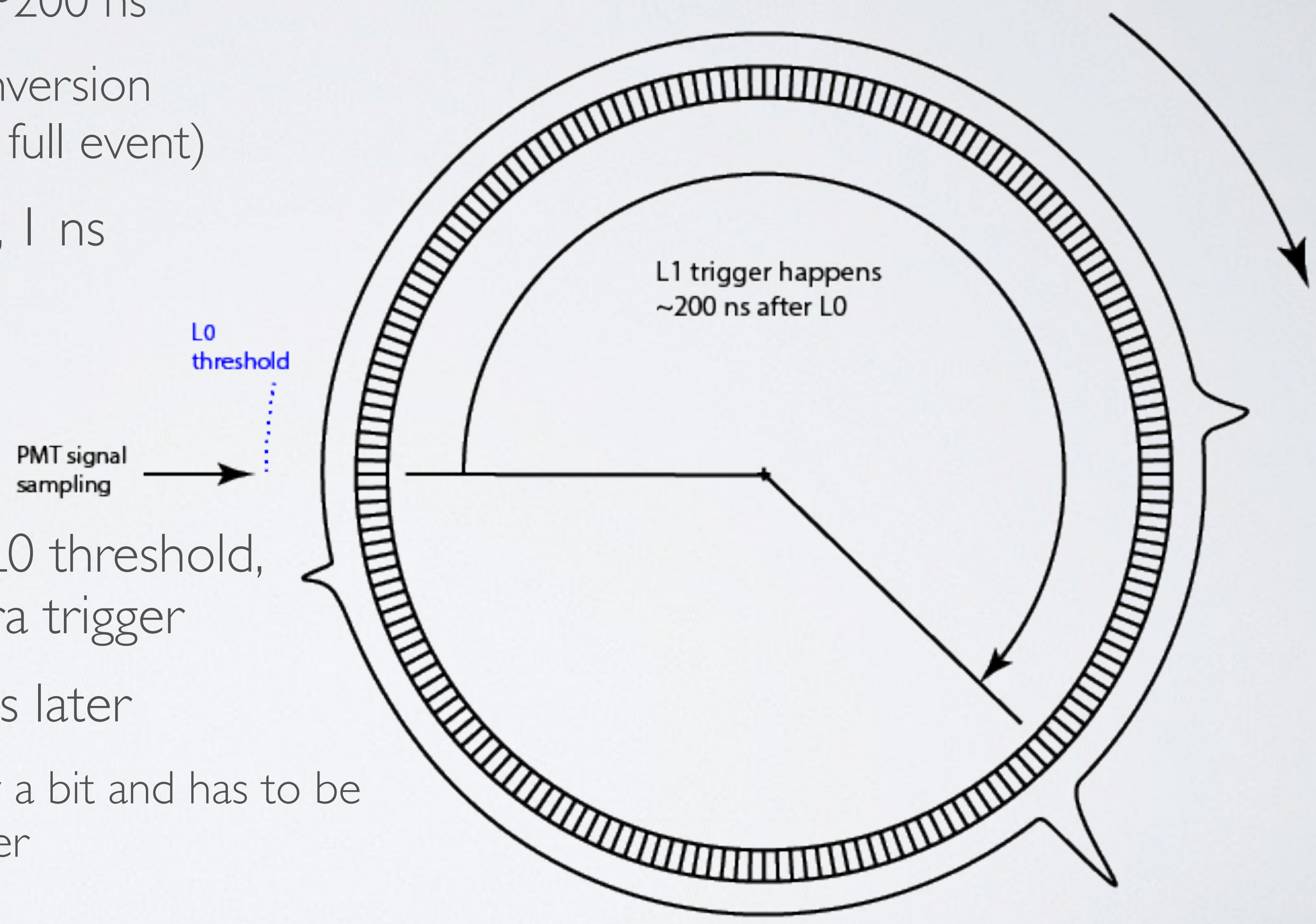
ANALOG MEMORIES

- Analog memory is used to store the signal for some time, until the trigger decision is taken
  - L1 trigger decision takes ~200 ns
  - Analog, because ADC conversion takes time (~250 μs for a full event)
- Signal is sampled at 1 GHz, 1 ns for each sample
- Circular buffer of 256 cells
- If the charge exceeds the L0 threshold, a L0 signal is sent to camera trigger



ANALOG MEMORIES

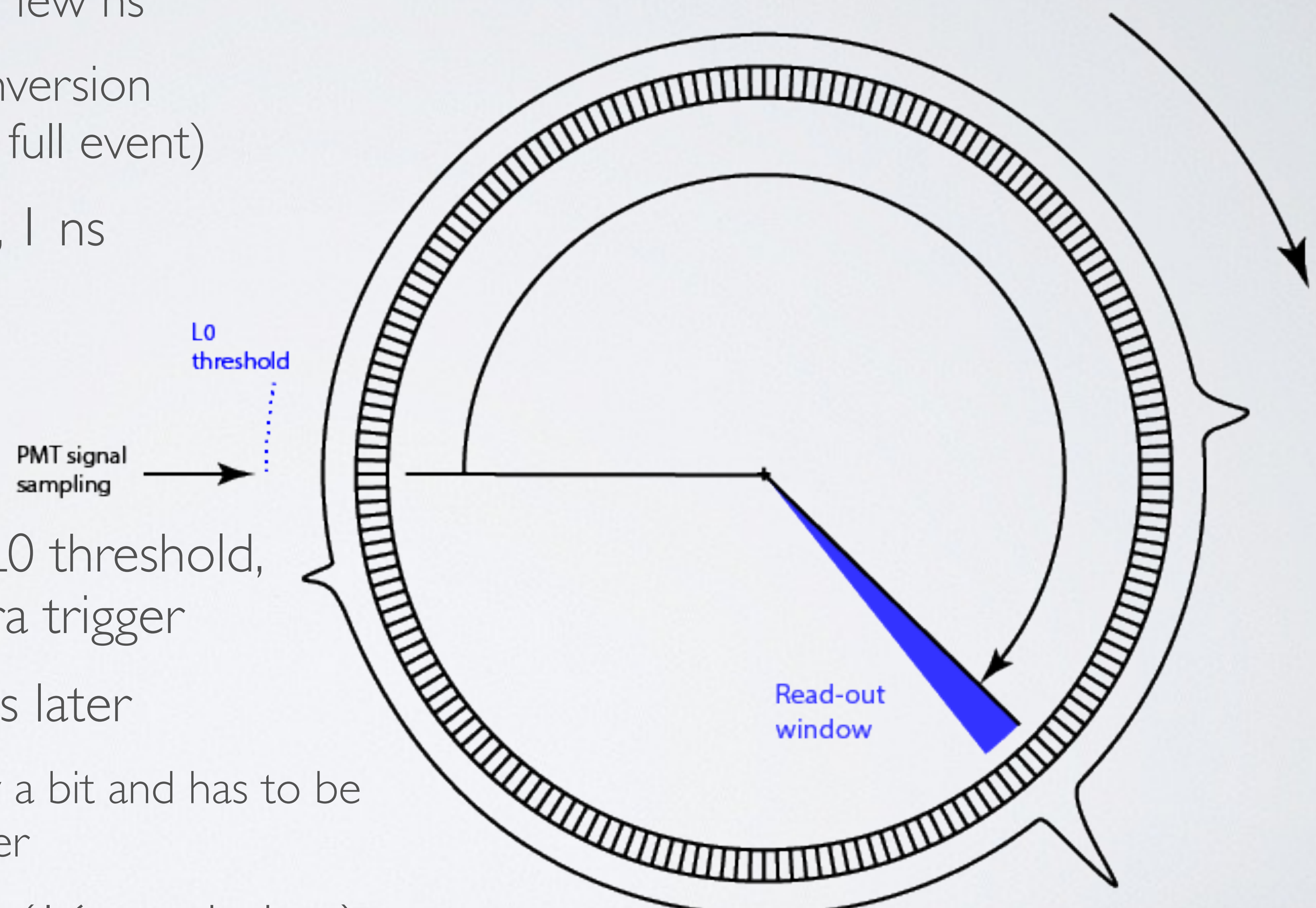
- Analog memory is used to store the signal for some time, until the trigger decision is taken
  - L1 trigger decision takes ~200 ns
  - Analog, because ADC conversion takes time (~250 μs for a full event)
- Signal is sampled at 1 GHz, 1 ns for each sample
- Circular buffer of 256 cells



- If the charge exceeds the L0 threshold, a L0 signal is sent to camera trigger
- L1 trigger happens ~200 ns later
  - The exact delay can vary a bit and has to be calibrated for each drawer

ANALOG MEMORIES

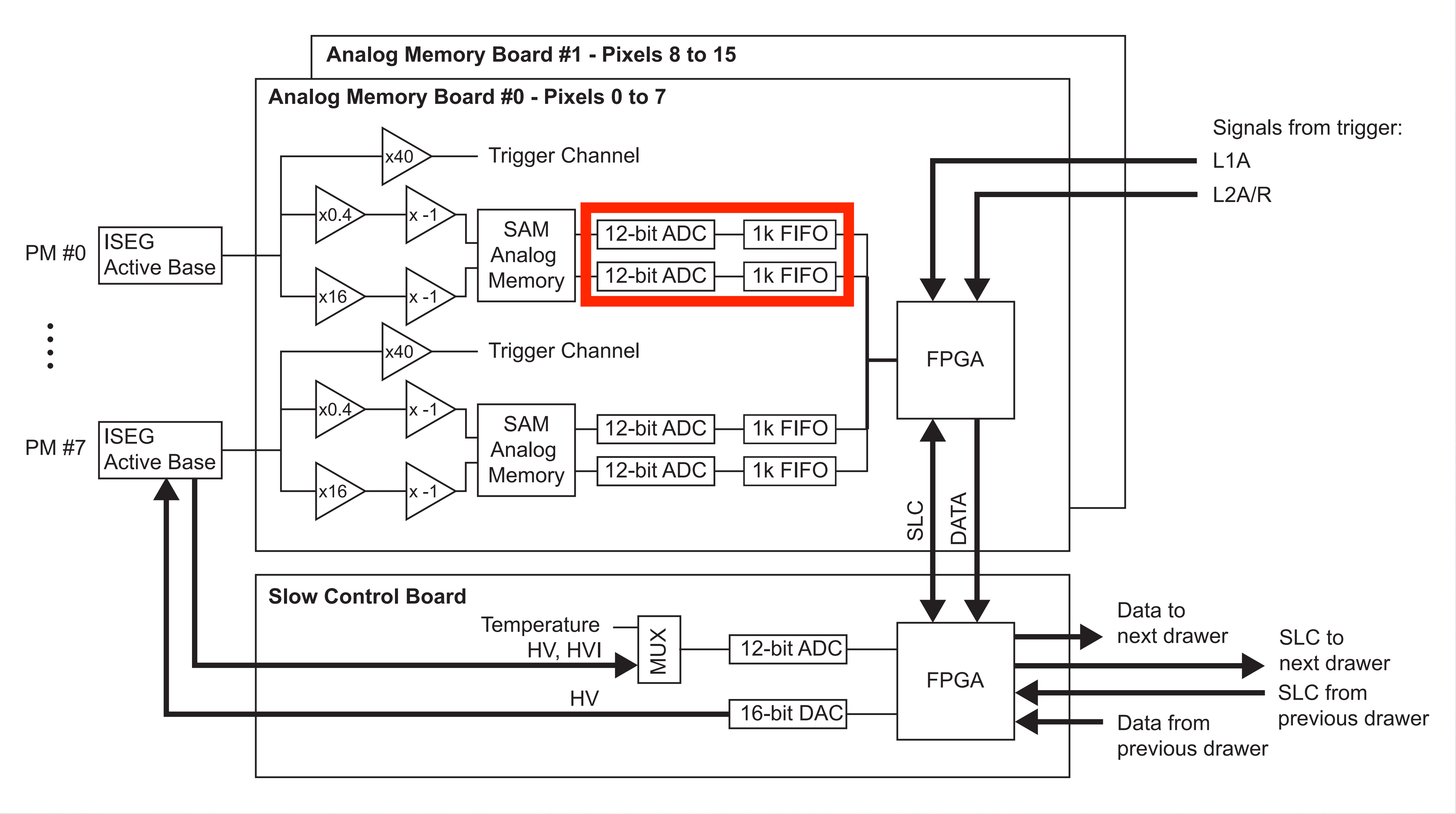
- Analog memory is used to store the signal for some time, until the trigger decision is taken
  - L1 trigger decision takes a few ns
  - Analog, because ADC conversion takes time (~250 μs for a full event)
- Signal is sampled at 1 GHz, 1 ns for each sample
- Circular buffer of 256 cells



- If the charge exceeds the L0 threshold, a L0 signal is sent to camera trigger
- L1 trigger happens ~200 ns later
  - The exact delay can vary a bit and has to be calibrated for each drawer
- The read-out is performed (16 ns window)

DATA ACQUISITION

- Simplified view of front-end electronics for HESS CT5 camera
    - View of 1 drawer



DATA ACQUISITION

- Once the trigger is received, analog memories are stopped, and the signal is digitized to be sent to the DAQ cluster
- FIFO buffers are used all along the signal path to make the tasks of analog memories and of the rest of the DAQ independent
    - It helps reducing the dead time (15 μs) and make it much smaller than the read-out time (250 μs)

- So, **raw data** consists in integrated charge as ADC counts
    - For each pixel
    - For each gain (HG and LG)
- **Slow control** data is intertwined with raw data, at lower rate (1 block for ~1000 events)
    - High voltages, high voltage base currents, temperatures

A SIDE NOTE ON CAMERA DESIGN

- The design presented here is the one used for CT5 « old » « French camera ». It is inherited from CAT, and HESS-I camera designs.
- Other designs exist
- Most recent ones show improved performance
    - Better sensitivity, lower energy threshold, lower dead time
- Some examples:
    - MAGIC: analog signals are sent to the ground through optical fibers
    - HESS-II new camera design (NamCam): signals from the PMTs are digitized directly and sent to an FPGA for trigger, data processing and event transmission
    - NectarCam: evolution from HESS-II old camera readout with improved performance
        - Dead time < 500 ns

PART II

CALIBRATION

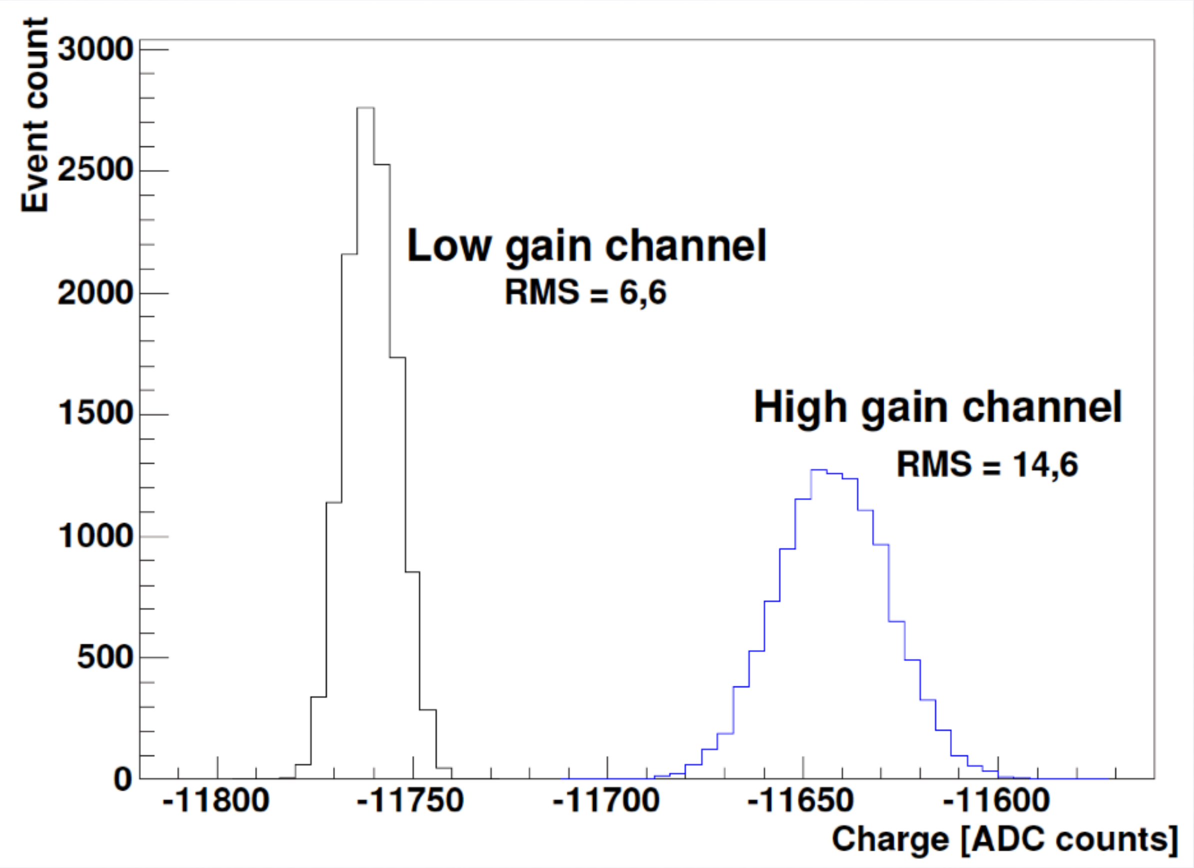
CALIBRATION

- Calibration is the step where **ADC counts are converted into a charge in (photoelectrons) p.e.**
- Cherenkov light intensity is also assessed in order to get the **optical efficiency**
- It is a multi-step process requiring specific types of data
  - What is the electronic signal baseline? → **Pedestal**
  - How to convert ADC counts to a charge at HG? → **Single photo-electron (SPE)**
  - What about LG channel? → **Hi/Lo ratio**
  - How to deal with pixel efficiency non-uniformities? → **Flat-field**
  - How to convert charge to Cherenkov light intensity?
    - What is the optical transmission efficiency? → Use of **muons + atmospheric monitoring**

- From then on, we are going to deal with **charge distributions**
    - Charge is obtained integrating PMT pulses

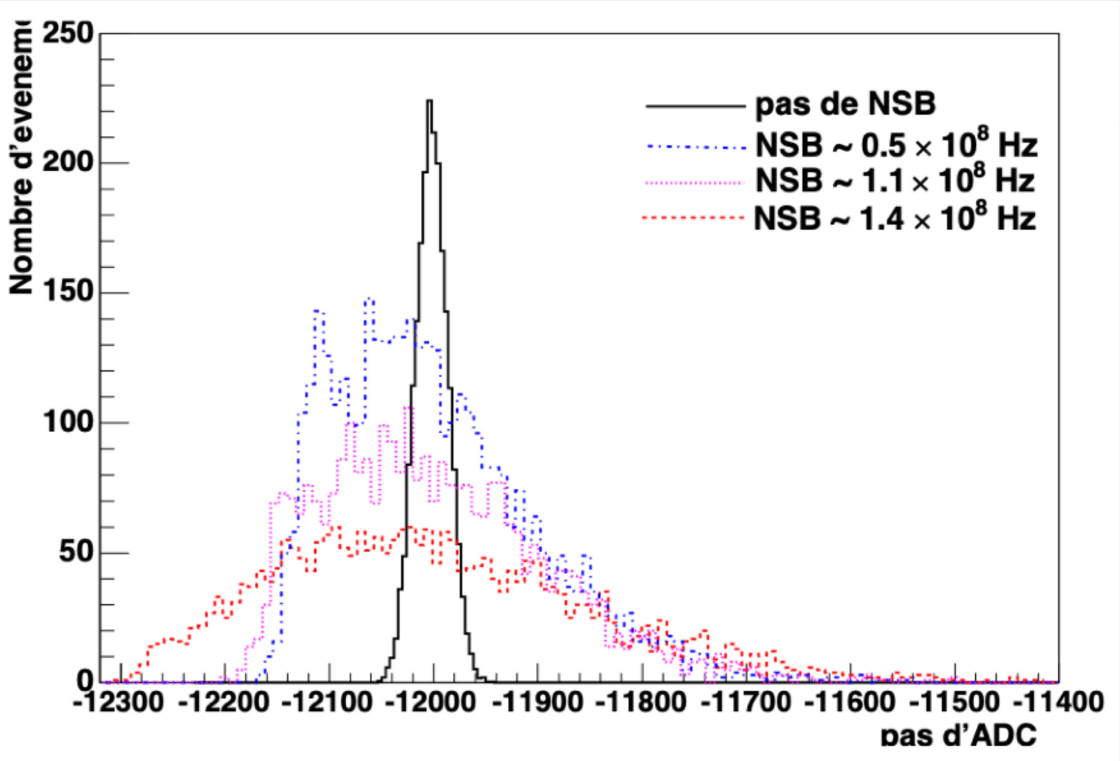
PEDESTAL

- **Pedestal** = baseline of the electronics
- It needs to be calibrated with high accuracy since Cherenkov signals are deduced using this baseline as a reference
- Can be acquired from:
  - Dedicated runs: camera lid closed, random trigger



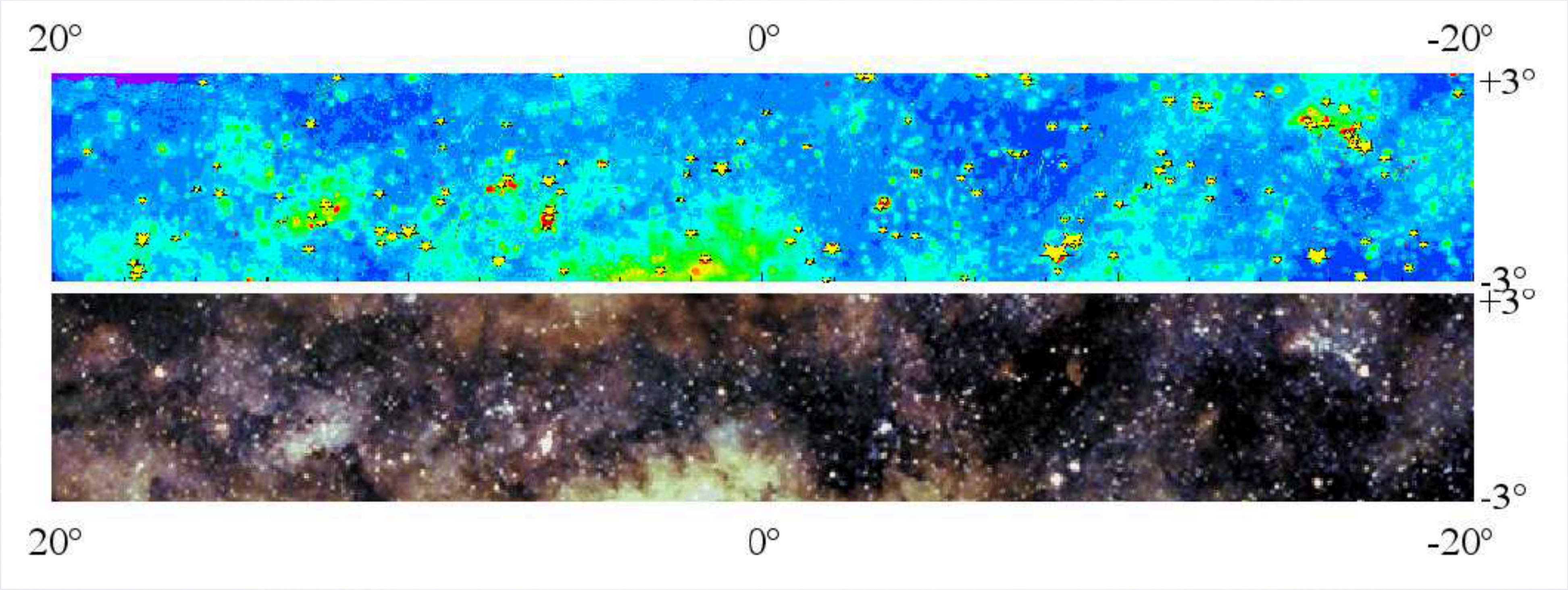
PEDESTAL

- **Pedestal** = baseline of the electronics
- It needs to be calibrated with high accuracy since Cherenkov signals are deduced using this baseline as a reference
- Can be acquired from:
  - Dedicated runs: camera lid closed, random trigger
  - Data taking runs: identifying and filtering out pixels hit by Cherenkov light
    - More NSB → wider pedestal



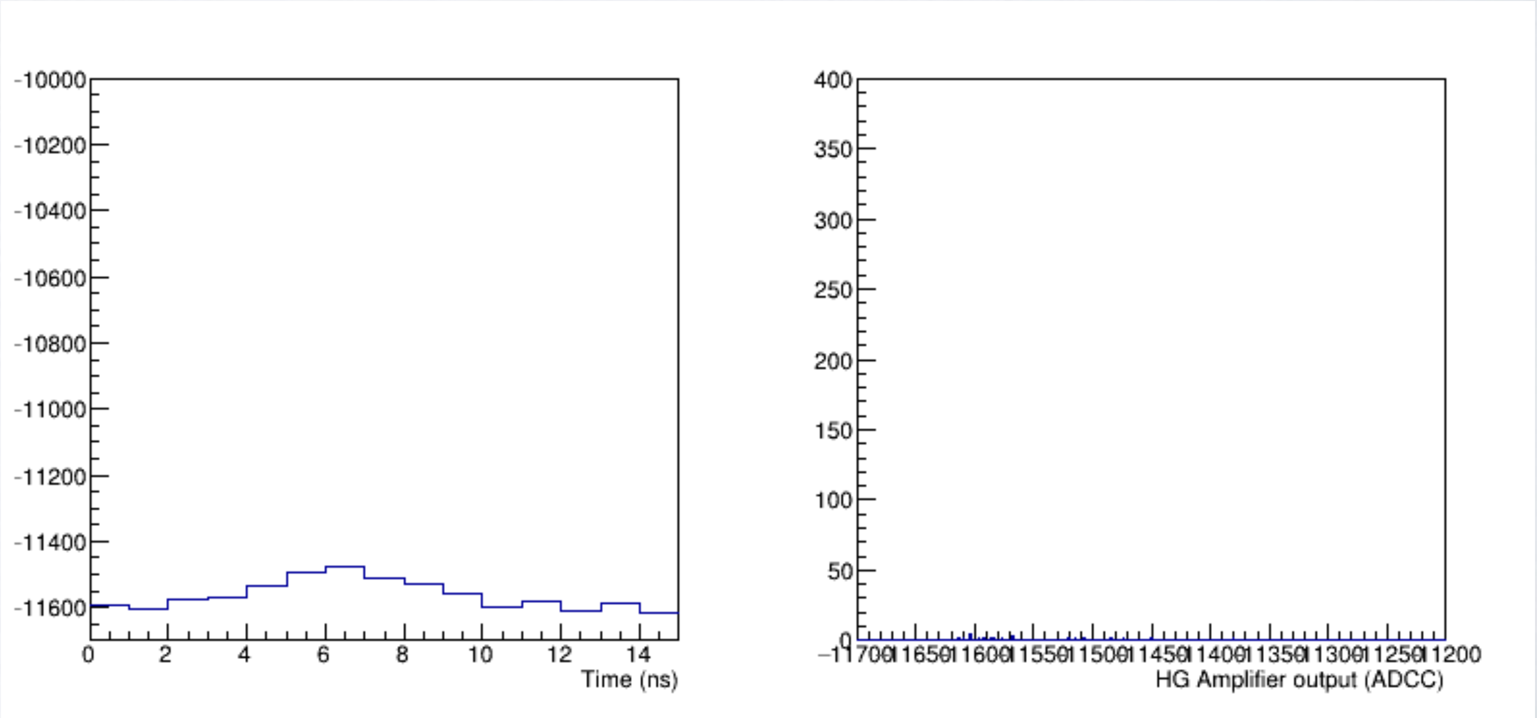
PEDESTAL

- **Pedestal** = baseline of the electronics
- It needs to be calibrated with high accuracy since Cherenkov signals are deduced using this baseline as a reference
- Can be acquired from:
  - Dedicated runs: camera lid closed, random trigger
  - Data taking runs: identifying and filtering out pixels hit by Cherenkov light
    - More NSB → wider pedestal



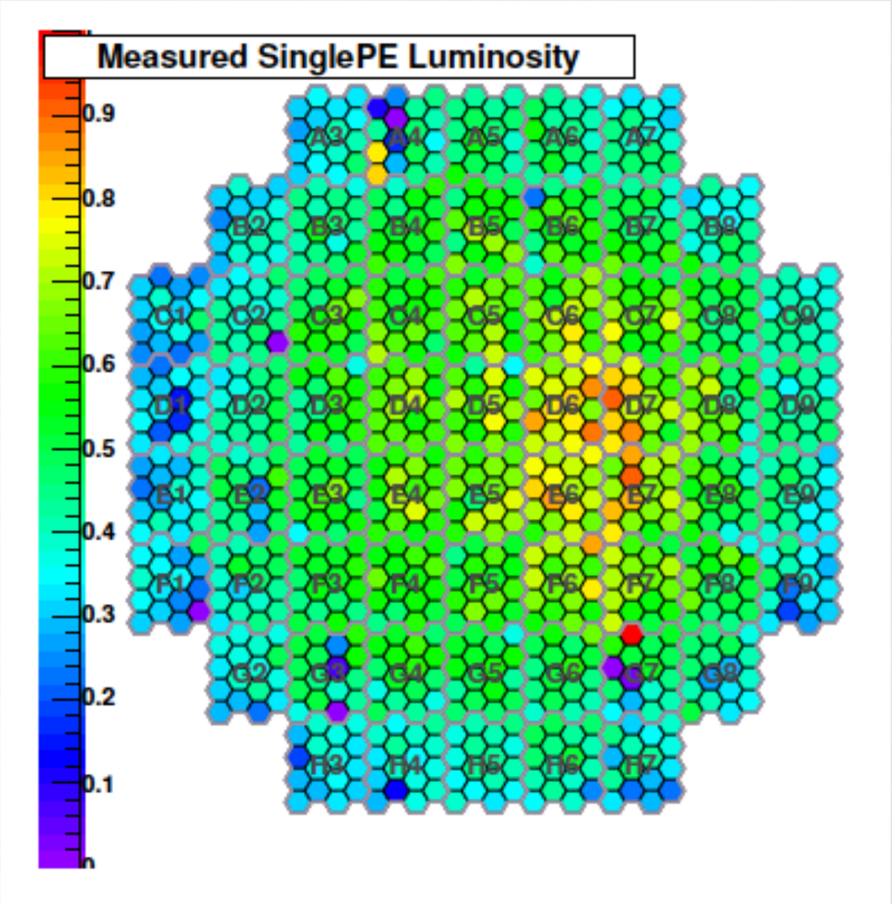
GAIN CALIBRATION

- **Gain** is used to convert ADC counts into a charge expressed in photo-electrons (p.e.)
- Overall gain is from
  - PMT, and
  - Amplifier
    - Obviously, it is different for HG and LG channels
- Dedicated runs are necessary:
  - With camera protected from ambient light (dark or using a filter)
  - Low intensity flashes so that 1 p.e. is produced at the photocathode on average
  - External trigger required
- (Simulated) Signal and charge distribution in the HG channel:



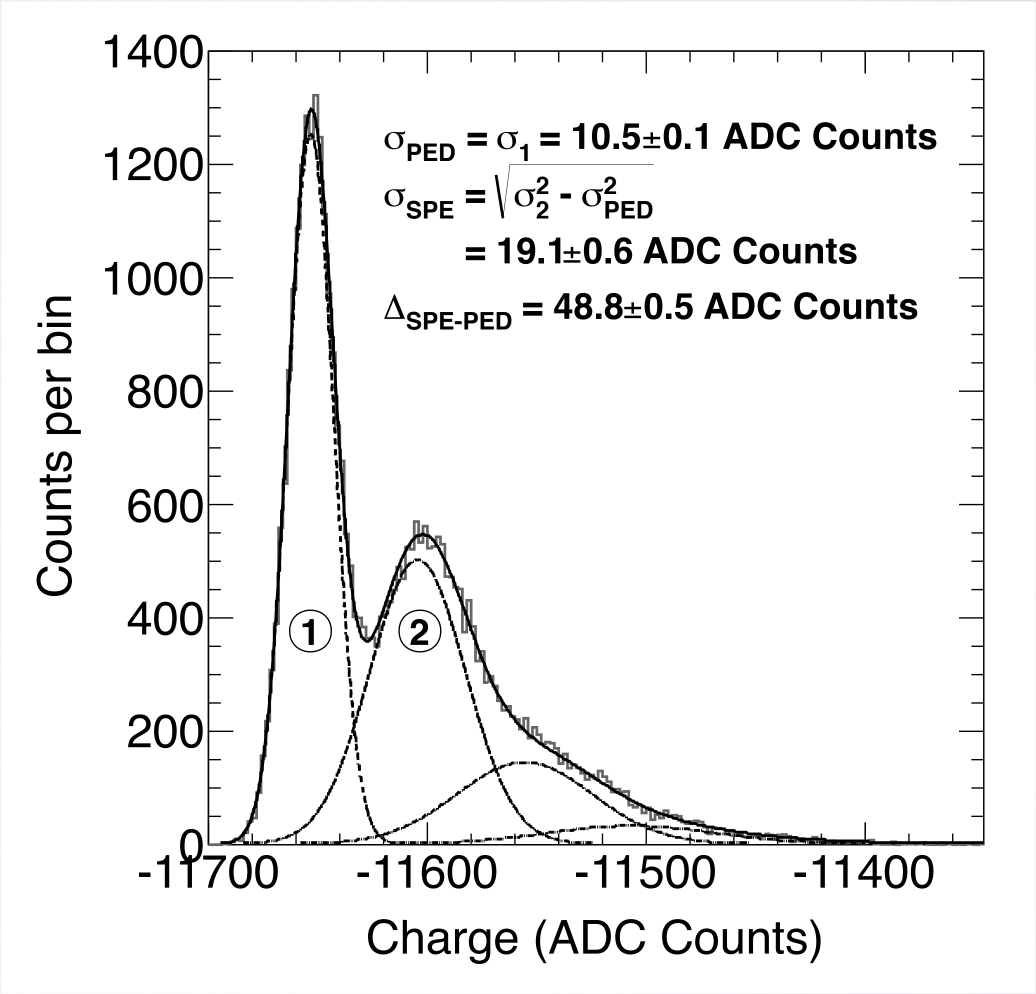
GAIN CALIBRATION

- **Gain** is used to convert ADC counts into a charge expressed in photo-electrons (p.e.)
- Overall gain is from
  - PMT, and
  - Amplifier
    - Obviously, it is different for HG and LG channels
- Dedicated runs are necessary:
  - With camera protected from ambient light (dark or using a filter)
  - Low intensity flashes so that 1 p.e. (SPE, single photo-electron) is produced at the photocathode on average
  - External trigger required
- Luminosity on the camera:



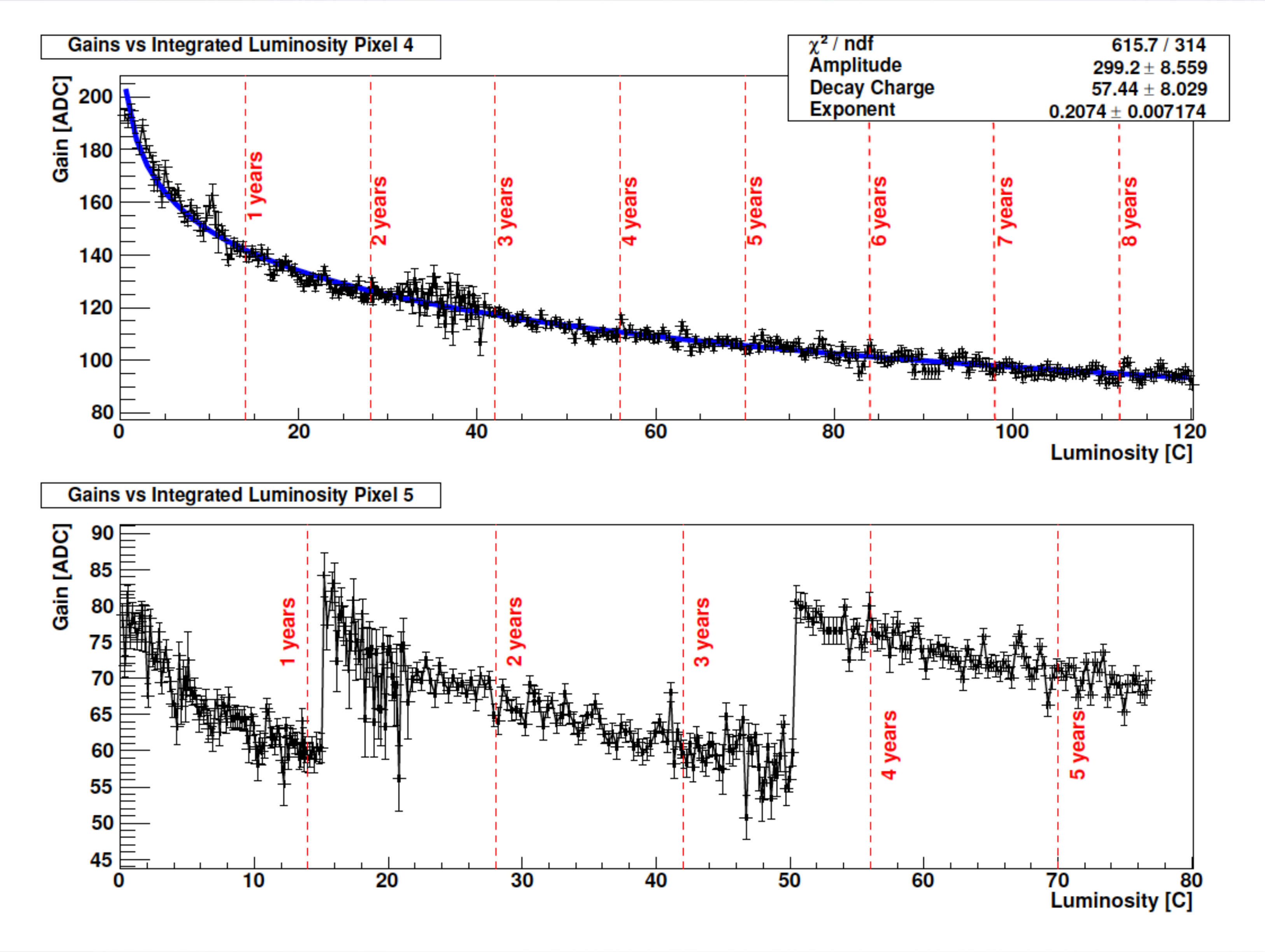
GAIN CALIBRATION

- **Gain** is used to convert ADC counts into a charge expressed in photo-electrons (p.e.)
- Overall gain is from
  - PMT, and
  - Amplifier
    - Obviously, it is different for HG and LG channels
- Dedicated runs are necessary:
  - With camera protected from ambient light (dark or using a filter)
  - Low intensity flashes so that 1 p.e. (SPE, single photo-electron) is produced at the photocathode on average
  - External trigger required
- Charge distribution and fit:
  - Nominal gain is 50 ADC/p.e.
- NB: SPE runs are also used to set PMT HV…



GAIN AND AGING

- PMTs are subject to aging → gain decrease with time → need for regular monitoring
- HV adjustments are made from time to time to re-adjust the gains to 2×10^5

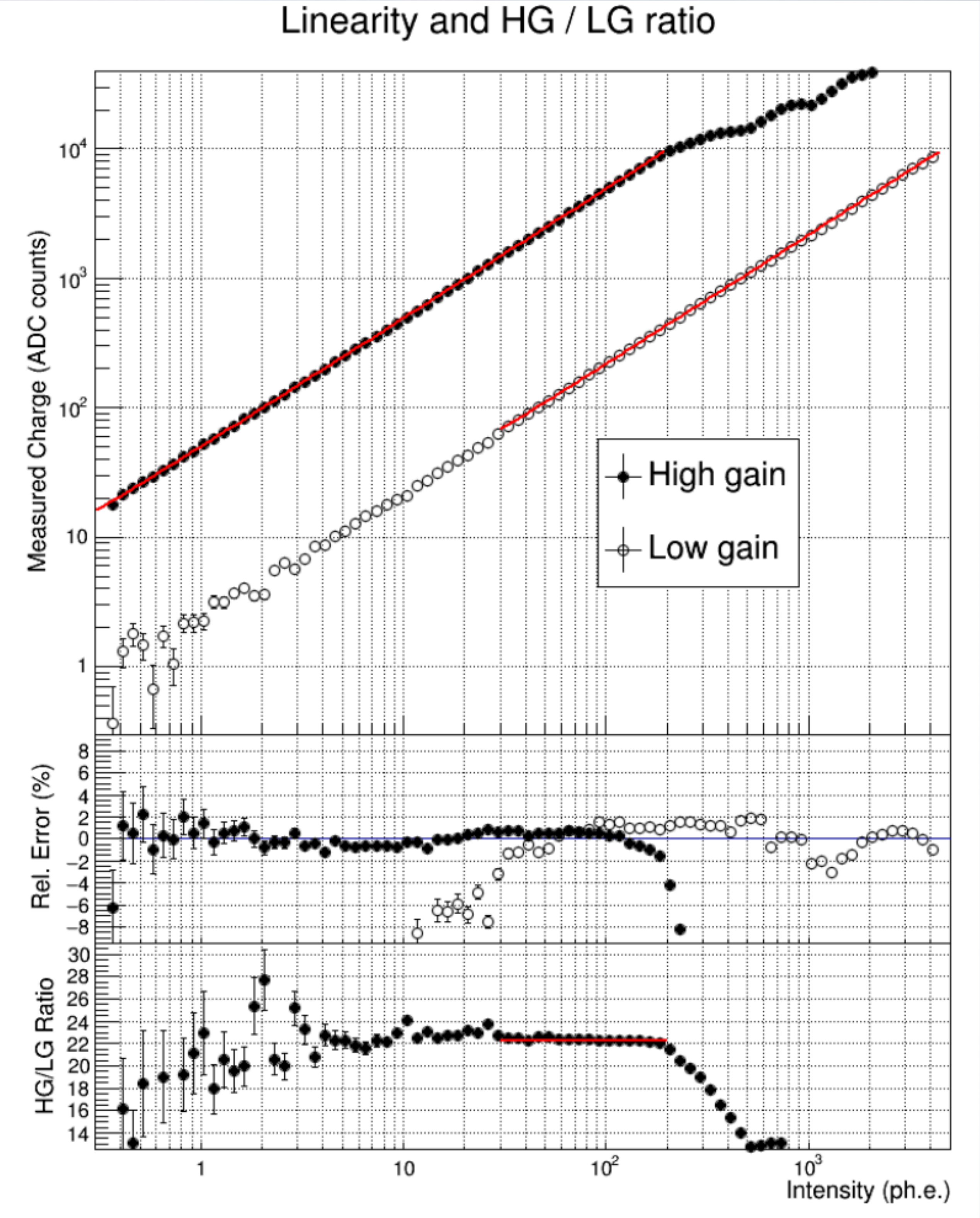


Test bench

On site

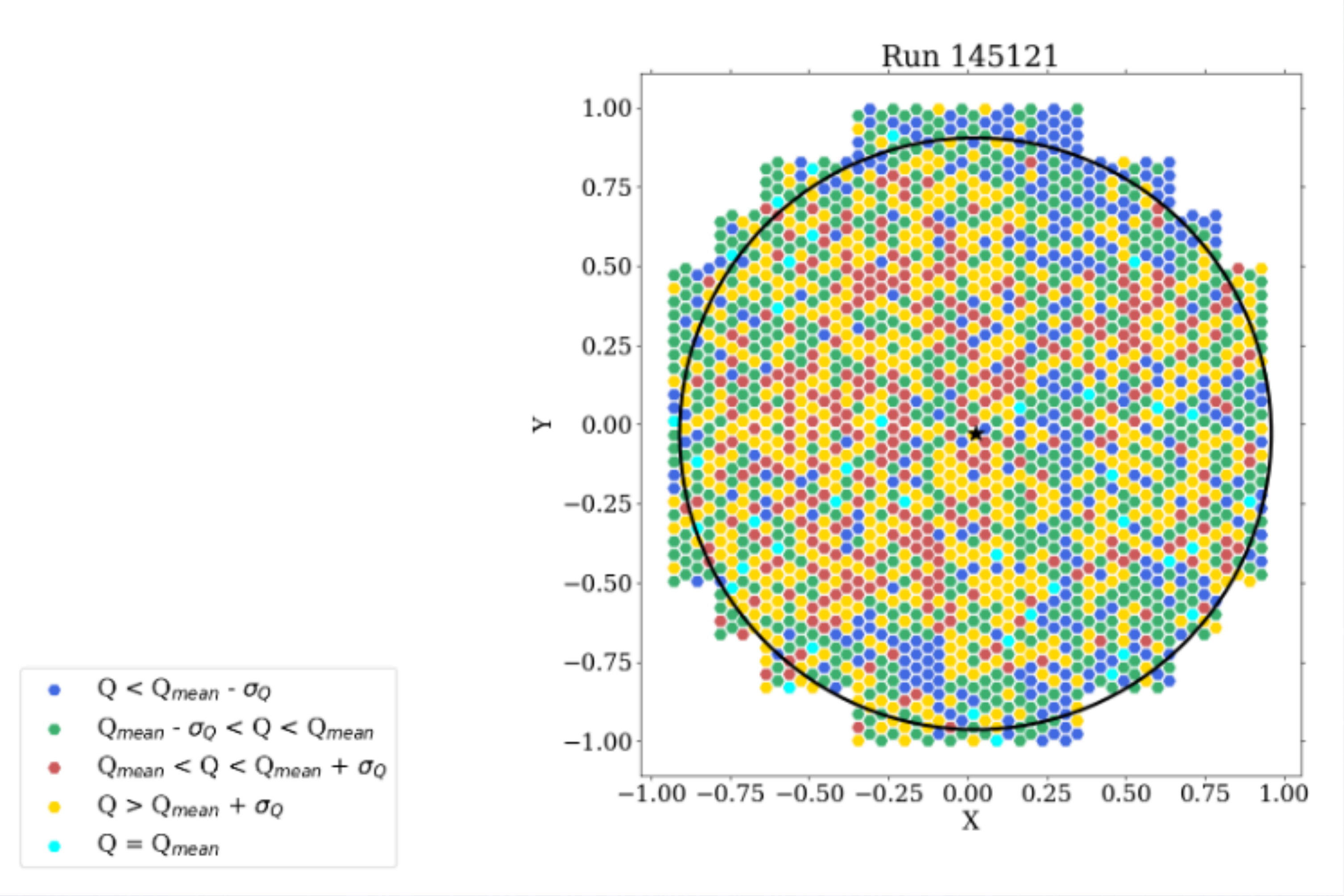
HI/LO GAIN CROSS-CALIBRATION

- SPE cannot be resolved in the LG channel
- Real data is used to **cross-calibrate High and Low gain channels** in the intensity range where they overlap



FLAT-FIELD

- **Flat-fielding** is used to correct for non-uniformities of collection efficiency (cones + PMT) and PMT Quantum Efficiency
  - Camera is illuminated by a very uniform light source located at the center of the dish
  - Cross-calibration coefficients are obtained for each pixel



WHAT WE HAVE SO FAR

- Up to now, we have a conversion from ADC to a number of photo-electrons, averaged over the spatial efficiency of the camera:

$$
\begin{aligned}
S_{HG} &= \frac{ADC_{HG} - P_{HG}}{\gamma_{HG}} \times FF \\
S_{LG} &= \frac{ADC_{LG} - P_{LG}}{\gamma_{HG}} \times (HG/LG) \times FF
\end{aligned}
$$

with:

- **S**: signal in p.e.
- **ADC**: signal in ADCc, **γ**: gain
- **P**: pedestal
- **FF**: flat-filed coefficient
- **HG/LG**: Hi/Lo gain ratio

- Charge in p.e. will be the main ingredient of the next steps of data analysis
- But we still want to evaluate optical efficiency

CHARGE TO CHERENKOV LIGHT

- To get the number of Cherenkov photons hitting the telescope, we need to know:
  - Reflectivity/transmissivity of optical elements as a function of wavelength
  - Telescope collection efficiency
  - PMT QExCE as a function of wavelength.

- All this is difficult to measure…

- Atmospheric muons are used to get the overall optical efficiency

CHARGE TO CHERENKOV LIGHT

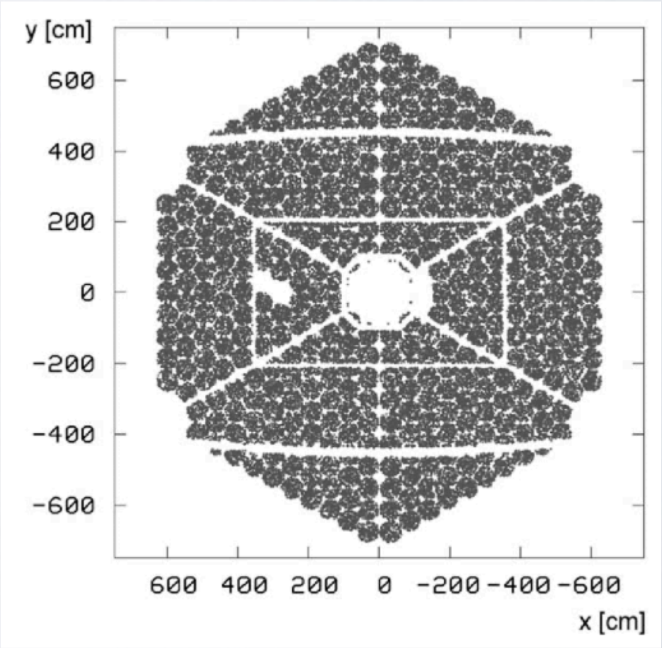
- Muons offer a natural calibration beam:
  - Radiation spectrum is very similar to Cherenkov emission from gamma-ray showers
  - Muons reaching the ground are ultra-relativisitic, making the light yield independent of the muon energy → it depends only on the track length
  - Muon signal is very short in time
- HESS small cameras can see the last ~400 m of the muon track (~700 m for CT5)

- The overall optical efficiency is given by:

$$\text{Eff. opt.} = \text{Re} \times (1\text{-Sh}) \times \text{Co} \times \text{QE}$$

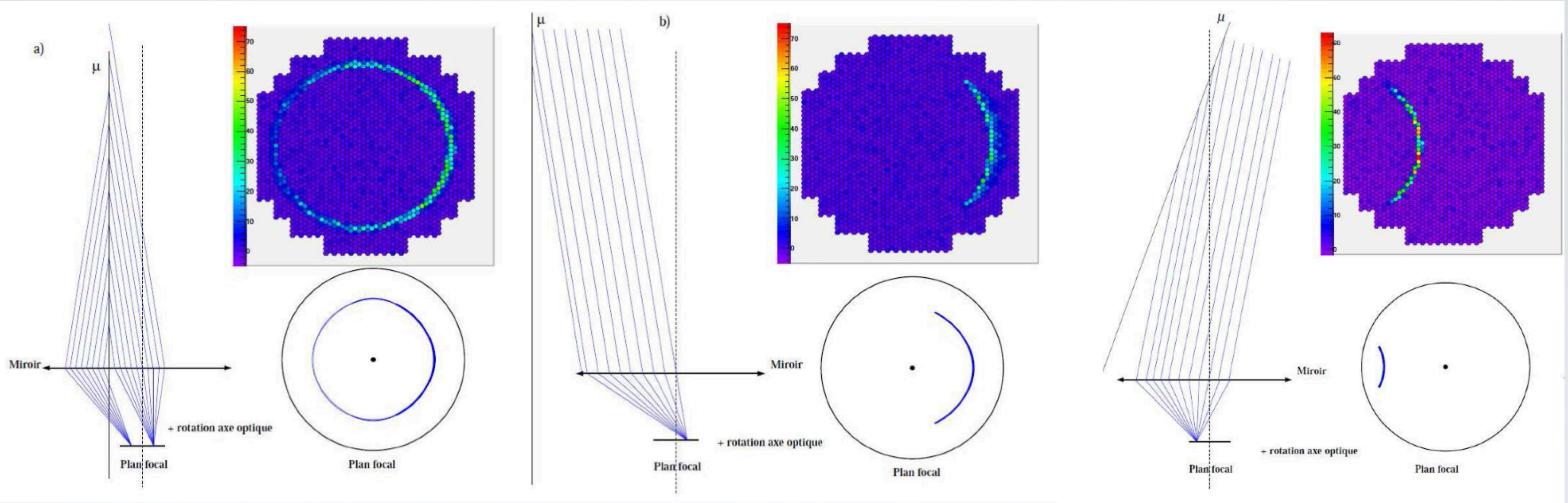
where:

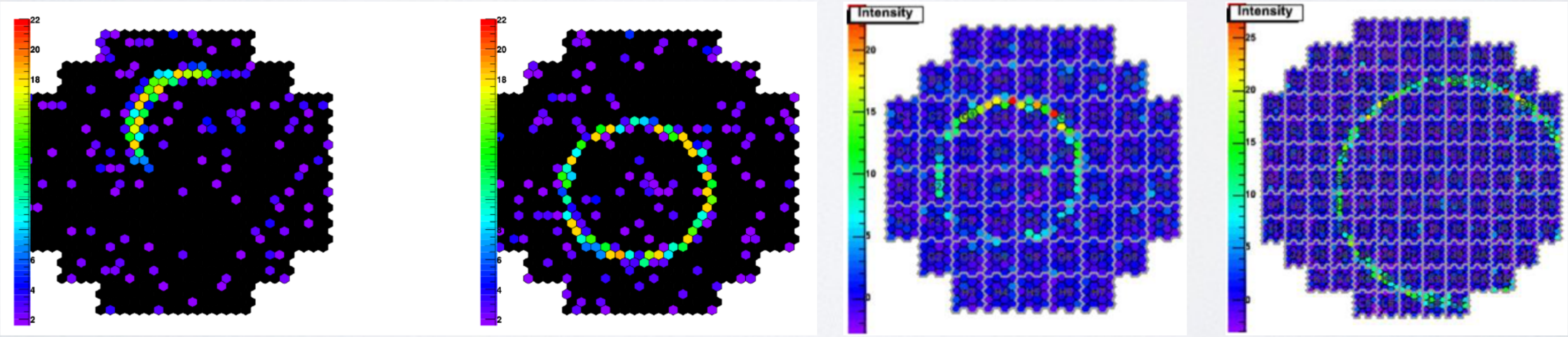
- **Re** is the mirror reflectivity integrated over wavelength (~80%)
- **Sh** is the shadow of camera and masts on the mirror (~10%)
- **Co** is the collection efficiency (~70%)
- **QE** is the PMT quantum efficiency (~30%)



CHARGE TO CHERENKOV LIGHT

- Some simulated and real examples:





CHARGE TO CHERENKOV LIGHT

- The number of photons reaching the camera per unit wavelength, azimuth angle, track length is given by:

$$\frac{d^3N}{dl\,d\phi\,d\lambda} = \frac{\alpha}{2}\boxed{\sin(2\theta_c)}\frac{\boxed{\psi(\lambda)}}{\lambda^2}\boxed{D(\phi,\rho)}\boxed{a(l,\lambda)}$$

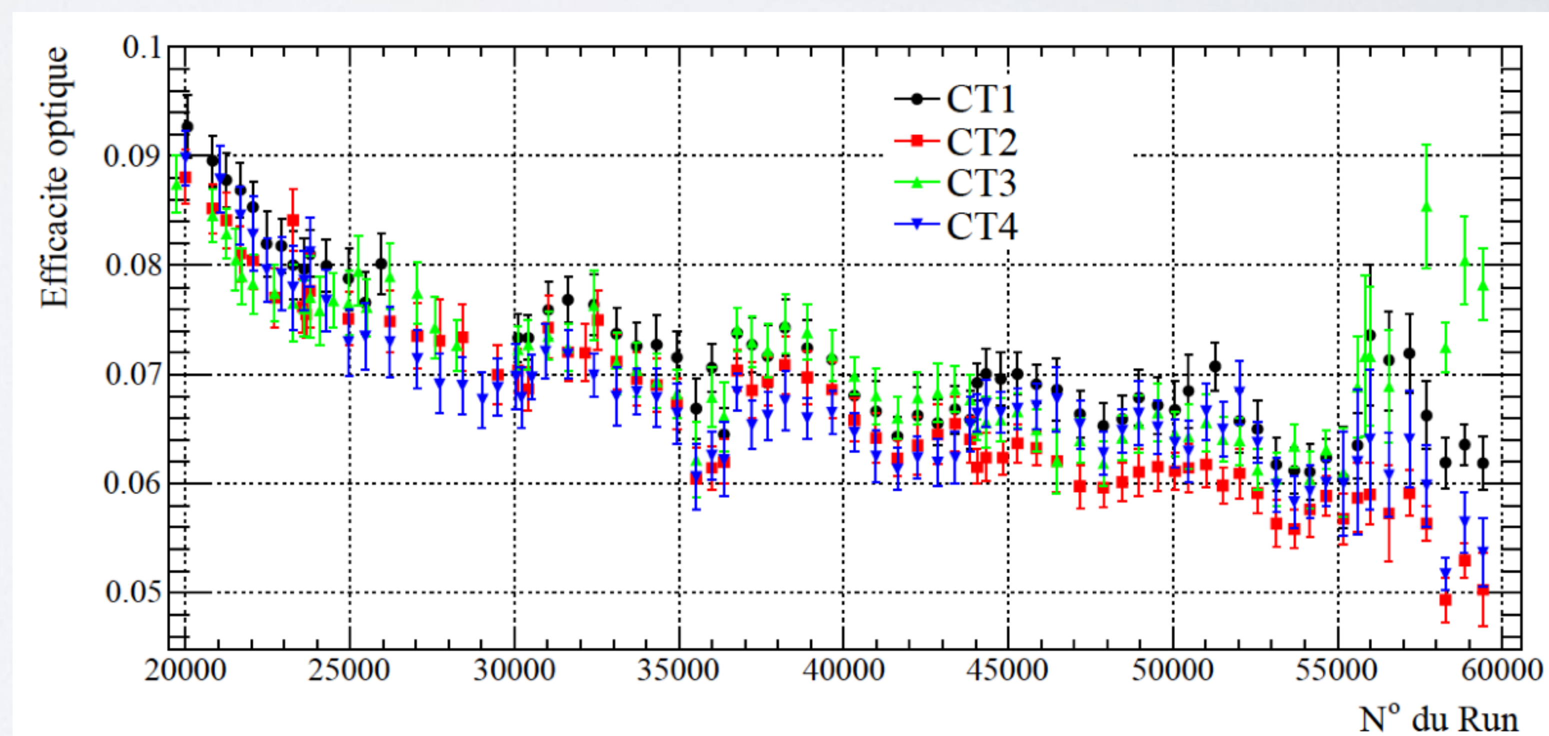
Related to the geometry

Atmospheric absorption

Optical efficiency

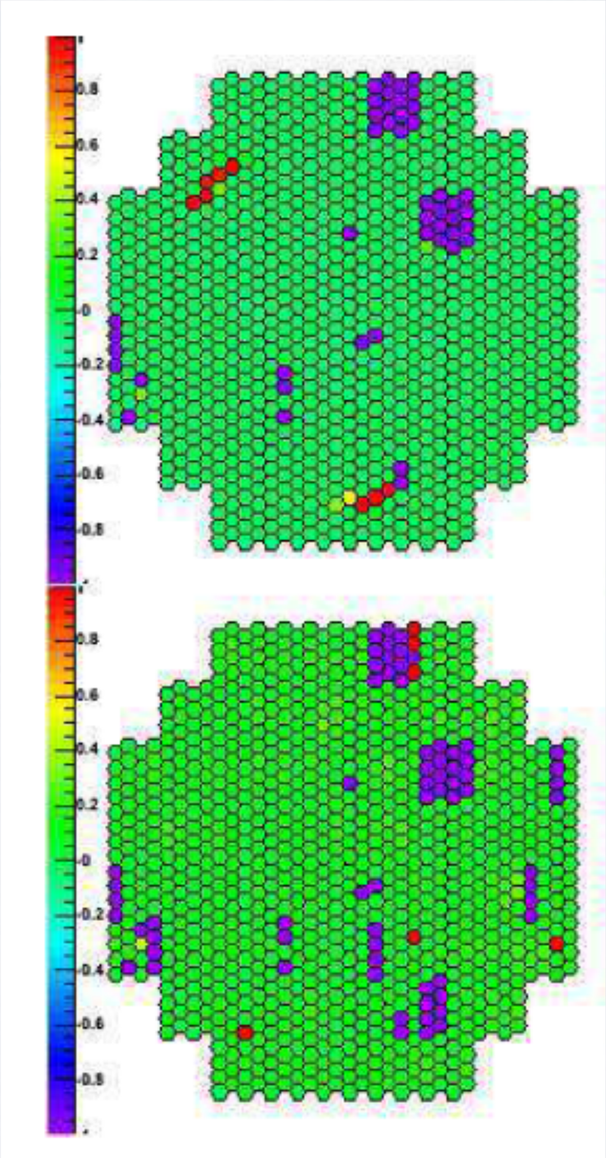
Related to the radius

- Atmospheric absorption can be obtained from models or deduced from measurements (LIDAR)
  - Transparency is reduced by dust, aerosols…
- The technique must be proven with Monte Carlo simulations!
- Long-term monitoring of optical efficiency is important
  - PMT cleaning, Winston cone replacement, mirror refurbishment…

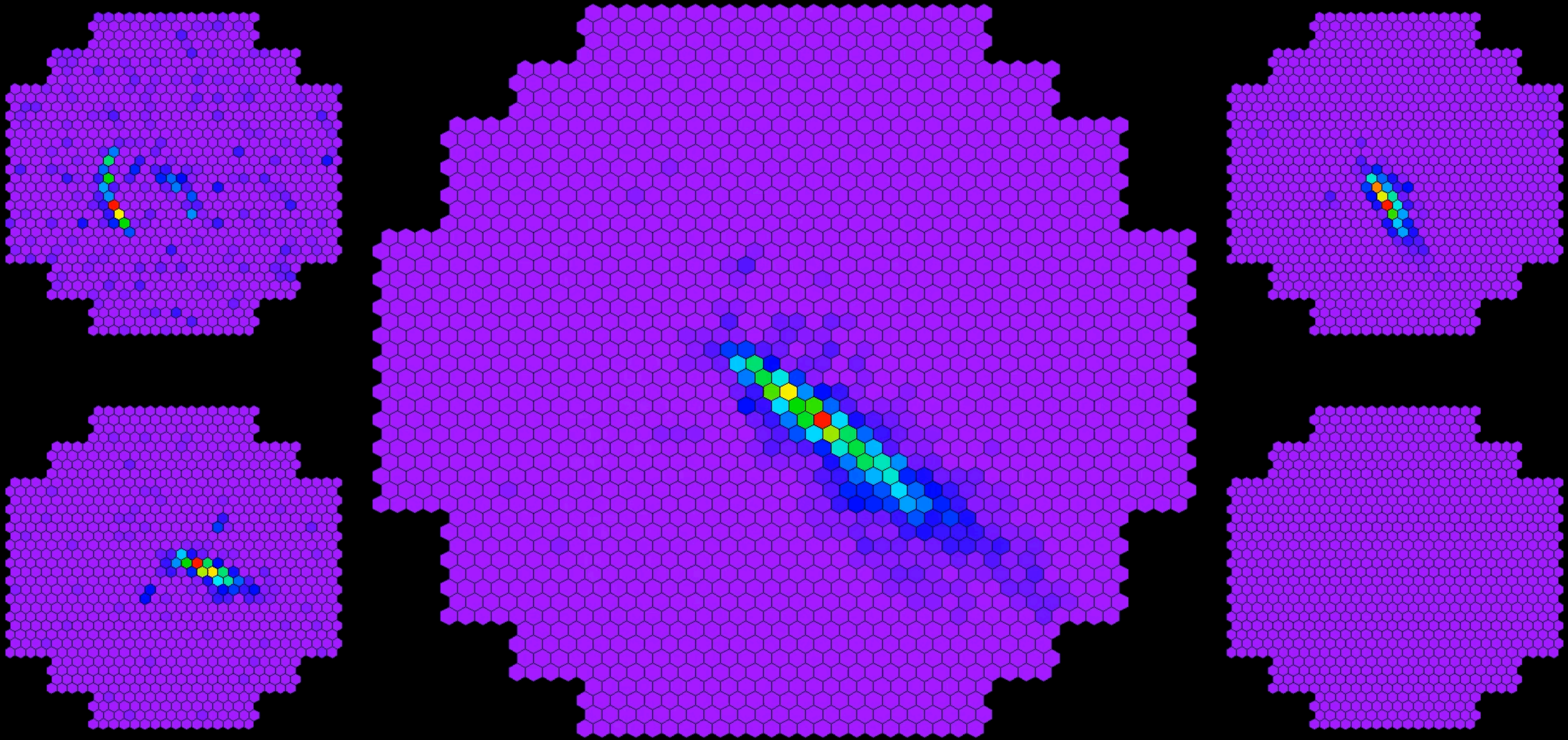


WHERE ARE WE NOW?

- **For each pixel, we end up with a charge in p.e., and a way to assess Cherenkov light intensity**
- **This will be the main ingredient for the next step: reconstruction and analysis**.



- The calibration step is also very important as a first assessment of instrument performance
- Hardware problems detected during calibration must be taken into account in the reconstruction and analysis steps
  - An inoperative pixel can be removed from the analysis only for a part of a run…
- Calibration software requires great care: a bug can introduce biaises in the reconstruction and analysis steps
  - Each time the calibration software is updated, all data sets need to be reprocessed…
  - Fortunately, effects of these bugs are usually very small!



To be continued…

PART III

RECONSTRUCTION AND ANALYSIS

WHAT DID WE LEARN SO FAR?

- Yesterday, we've seen
  - How images are collected
  - How we can convert from ADC to a charge and to the Cherenkov light intensity
    - Multiple step process
- In practice, reconstruction and analysis use charge in p.e. as main input
- Optical efficiency must be taken into account in Monte Carlo simulations, which provide a charge in p.e. for each pixel
- The next steps are
  - **Reconstruction**: shower images are processed to discriminate gamma-like and hadron-like events and to reconstruct shower parameters
  - **Analysis**:
    - if the signal is strong enough (significant), spectral analysis, light curve extraction…
    - If the signal is not significant, limits on the flux can still be derived
  - Additional criteria apply for LIV analysis
- Before that: run selection, data quality checks, etc. → not covered…

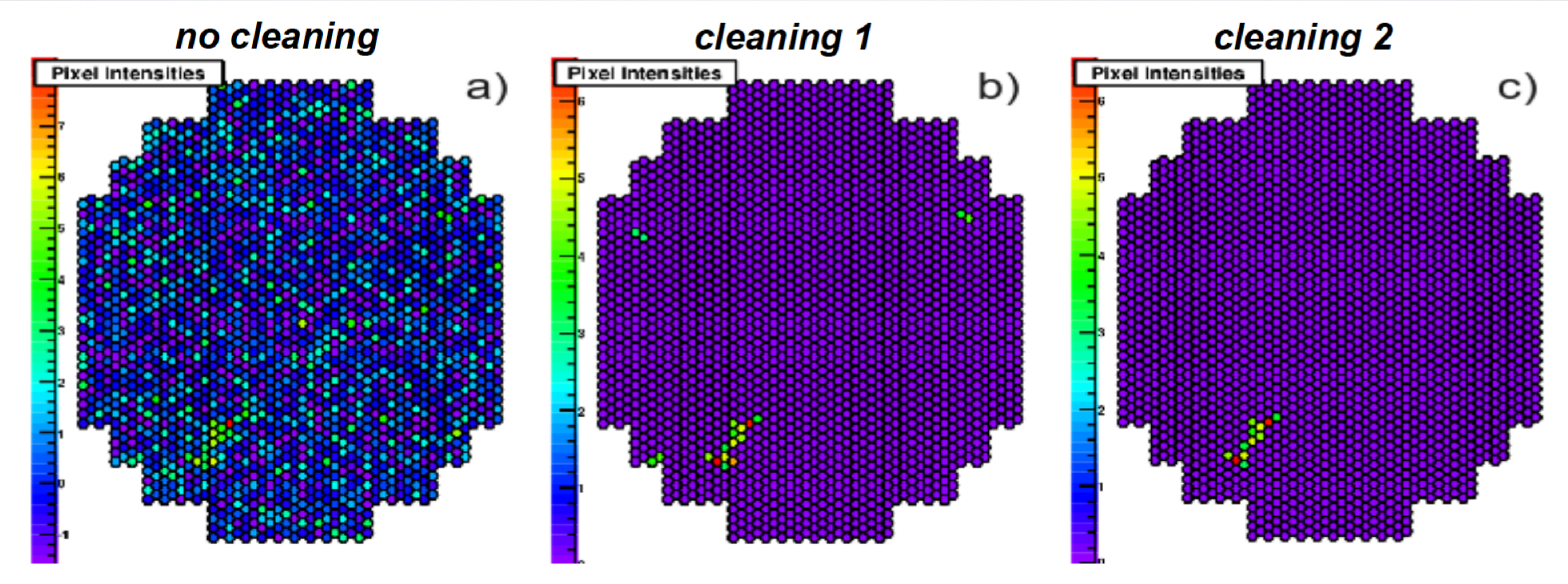
RECONSTRUCTION

- **Goals of reconstruction**:
  - Reconstruct the parameters of the primary particles: direction, impact location and energy
  - Provide discriminating variables to reduce as much as possible the amount of background events. Rejection factor is typically $O(10^4)$

- Two methods will be covered in this session
  - **Hillas** reconstruction
  - **Model** reconstruction
- They are quite simple to implement, and widely used
- They are also complementary and can be cross-checked or even combined
- Many more exist! Some are based on the use of neural networks, boosted decision trees, etc.
  - More complexity, longer computing times… but better overall performance

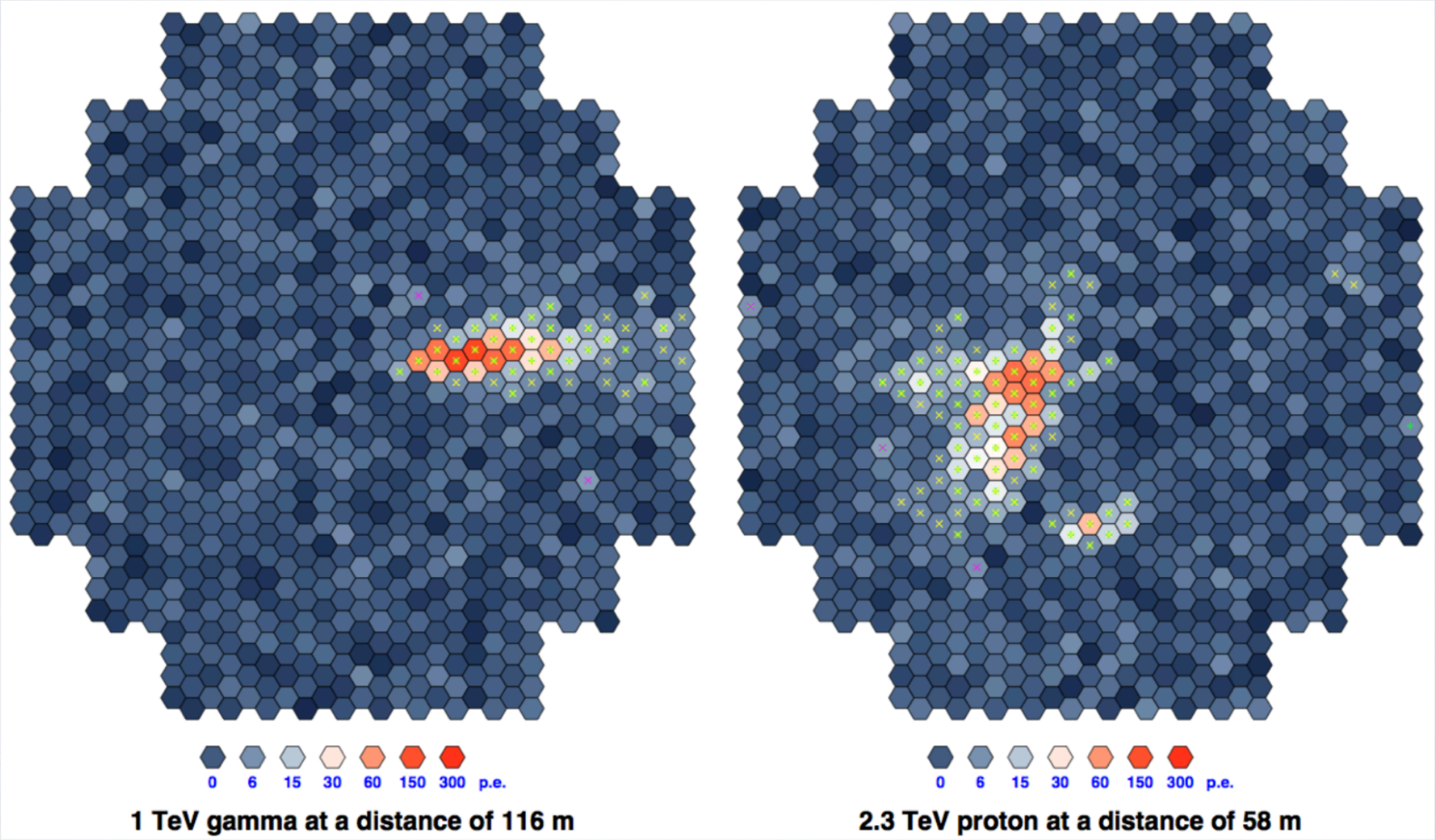
HILLAS RECONSTRUCTION

- First step: **image cleaning**
  - Ideally, keeps only the pixels with light from the shower
  - Two thresholds:
    - A pixel with a signal $>s_1$ is kept only if it has at least a neighbor with intensity $>s_2$
  - Different values can be used depending on the amount of NSB (4, 7), (7, 10)…



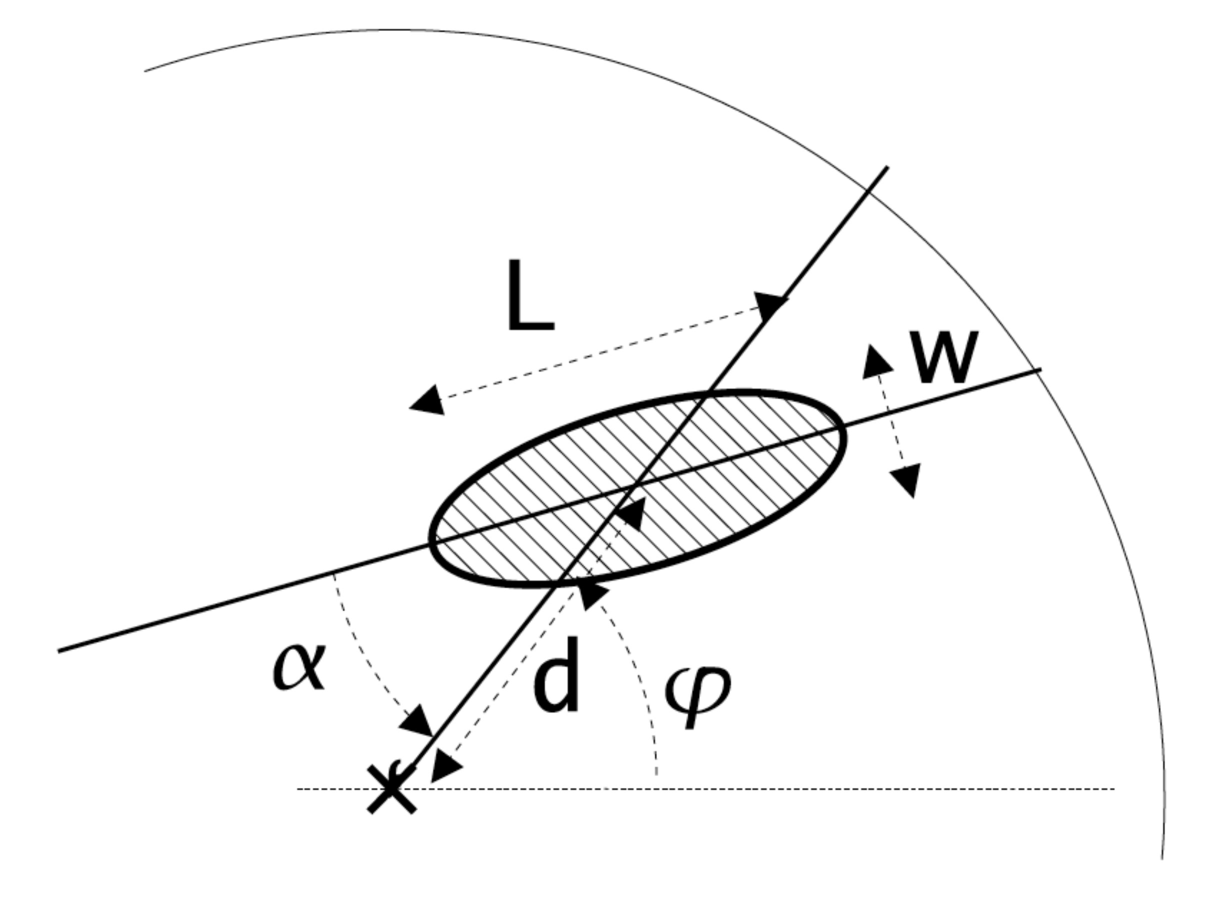
HILLAS RECONSTRUCTION

- Images look like this, for a photon, and for a hadron
- The gamma shower is quite elliptical in shape…



HILLAS RECONSTRUCTION

- In 1985, Hillas proposed to model the images by an ellipse
- **Hillas parameters**:
  - Length **L** and width **w** of the ellipse
  - **Size** (total image amplitude in p.e.)
  - Nominal distance **d**
  - Azimuthal angle of the image main axis **φ**
  - Orientation angle **α**



- This parametrization can be expressed from the intensity q_i (in p.e.) and position (x_i, y_i) of each pixel
- From these, several quantities are computed:

$$\langle x\rangle = \frac{\sum_i x_i q_i}{\sum_i q_i} \quad \langle xy\rangle = \frac{\sum_i x_i y_i q_i}{\sum_i q_i}$$

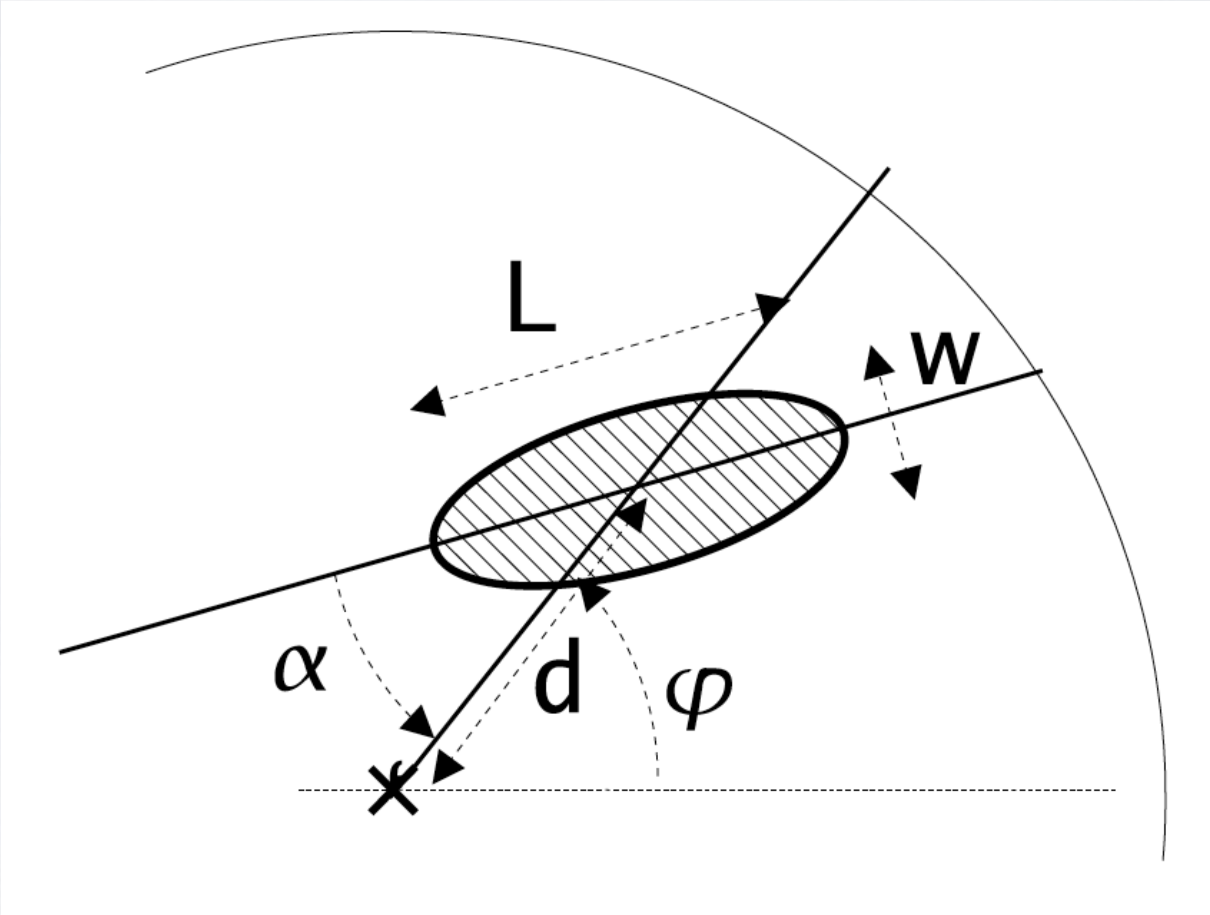
$$\langle y\rangle = \frac{\sum_i y_i q_i}{\sum_i q_i} \quad \langle x^2\rangle = \frac{\sum_i x_i^2 q_i}{\sum_i q_i} \quad \langle y^2\rangle = \frac{\sum_i y_i^2 q_i}{\sum_i q_i}$$

$$\sigma_{y^2} = \langle y^2\rangle - \langle y\rangle^2$$

$$\sigma_{x^2} = \langle x^2\rangle - \langle x\rangle^2$$

$$\sigma_{xy} = \langle xy\rangle - \langle x\rangle\langle y\rangle$$

HILLAS RECONSTRUCTION

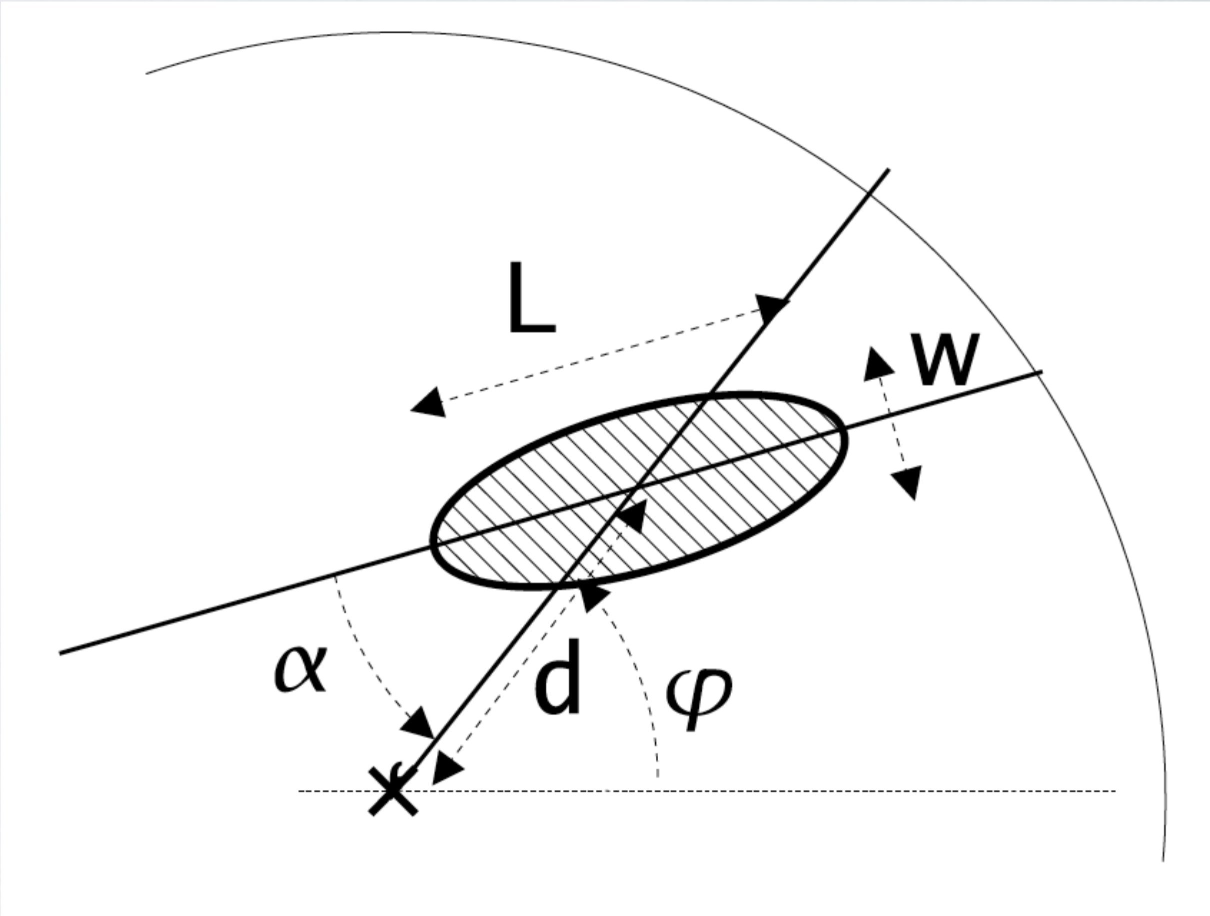


- In 1985, Hillas proposed to model the images by an ellipse
- **Hillas parameters**:
  - Length **L** and width **w** of the ellipse
  - **Size** (total image amplitude in p.e.)
  - Nominal distance **d**
  - Azimuthal angle of the image main axis **φ**
  - Orientation angle **α**
- This parametrization can be expressed from the intensity q_i (in p.e.) and position (x_i, y_i) of each pixel
- From these, several quantities are computed:

$$\begin{aligned}
\chi &= \sigma_{x^2} - \sigma_{y^2} \\
z &= \sqrt{\chi^2 + 4\sigma_{xy}} \\
b &= \sqrt{\frac{(1+\chi/z)\langle x\rangle^2 + (1-\chi/z)\langle y\rangle^2 - 2\sigma_{xy}\langle x\rangle\langle y\rangle}{2}}
\end{aligned}$$

HILLAS RECONSTRUCTION

- In 1985, Hillas proposed to model the images by an ellipse
- **Hillas parameters**:
  - Length **L** and width **w** of the ellipse
  - **Size** (total image amplitude in p.e.)
  - Nominal distance **d**
  - Azimuthal angle of the image main axis **φ**
  - Orientation angle **α**

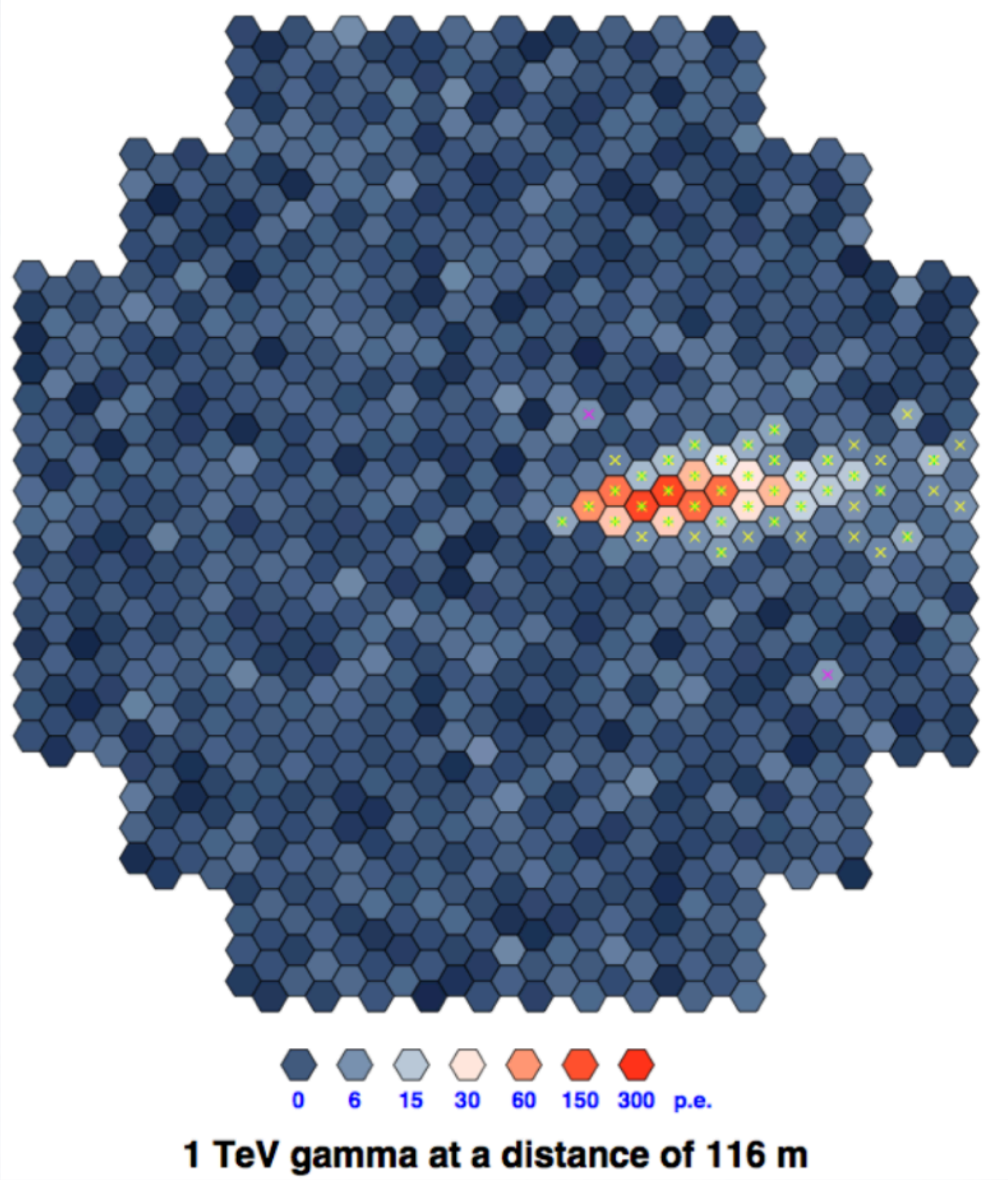


- This parametrization can be expressed from the intensity q_i (in p.e.) and position (x_i, y_i) of each pixel
- From these, several quantities are computed
- And then we get Hillas parameters from image quantities:

$$
\begin{aligned}
d &= \sqrt{\langle x \rangle^2 + \langle y \rangle^2} \\
L &= \sigma_{x^2} + \sigma_{y^2} + z \\
w &= \sigma_{x^2} + \sigma_{y^2} - z \\
\alpha &= \arcsin\left(\frac{b}{d}\right)
\end{aligned}
$$

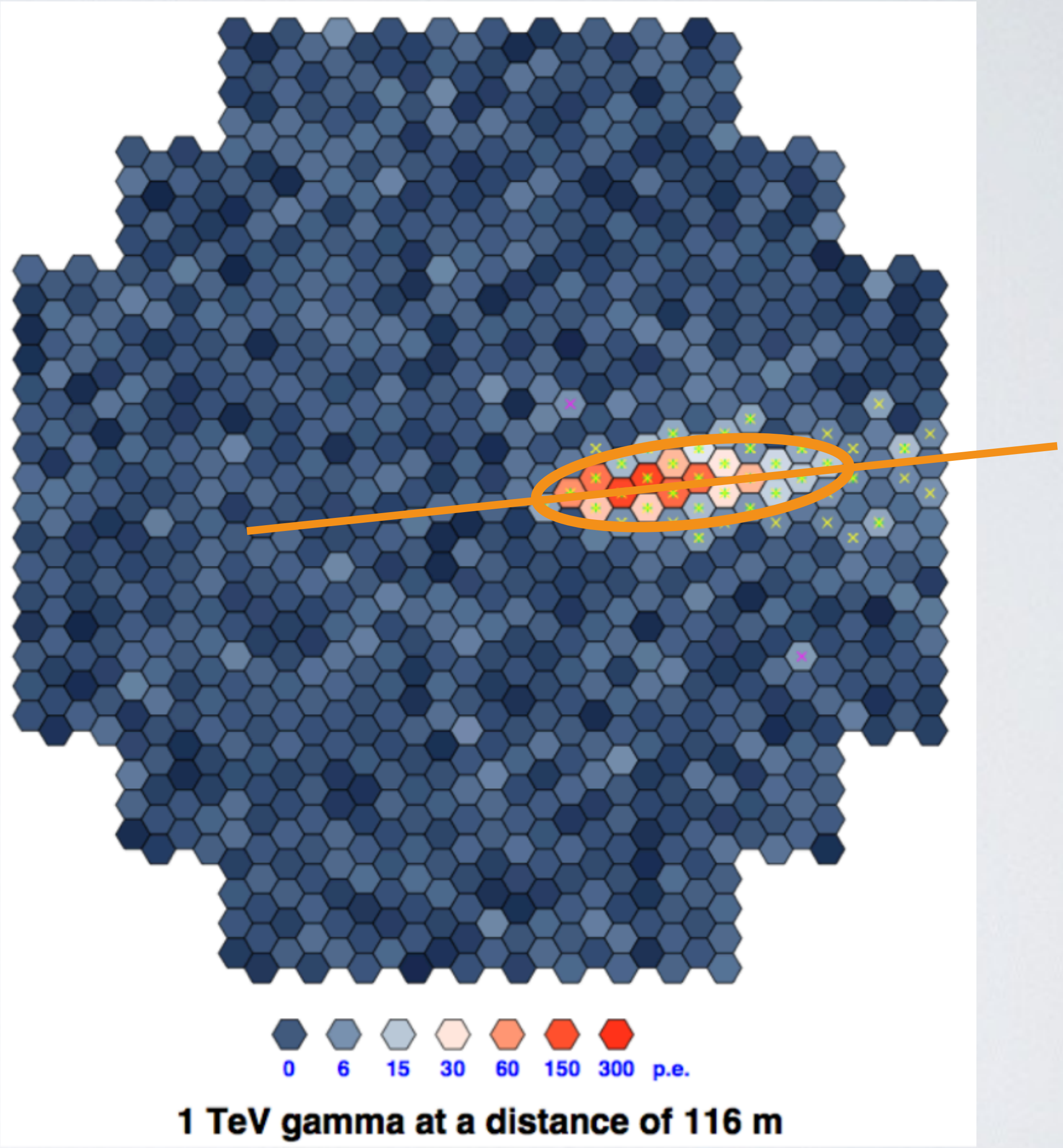
HILLAS RECONSTRUCTION

- Single telescope reconstruction
  - Degeneracy for source direction…



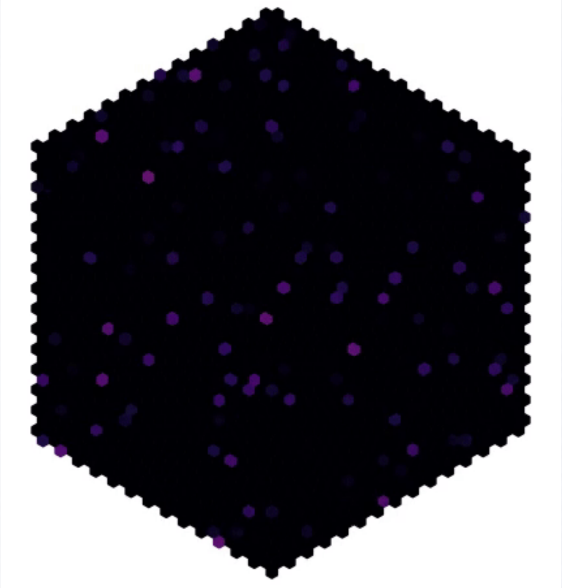
HILLAS RECONSTRUCTION

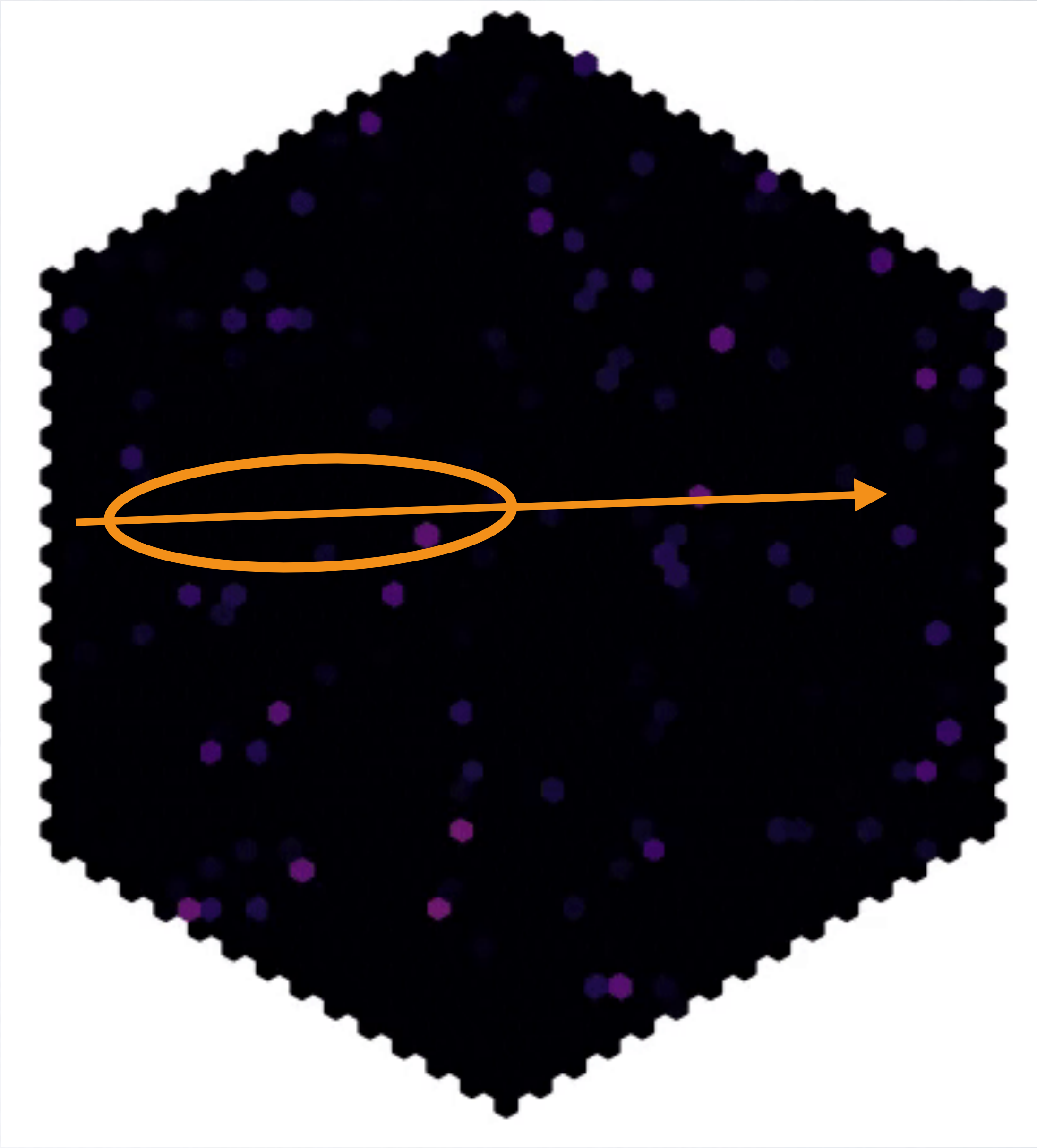
- Single telescope reconstruction
  - Degeneracy for source direction…
  - Additional parameters are added to remove the degeneracy
  - Recent hardware developments can help: addition of timing in image analysis



HILLAS RECONSTRUCTION

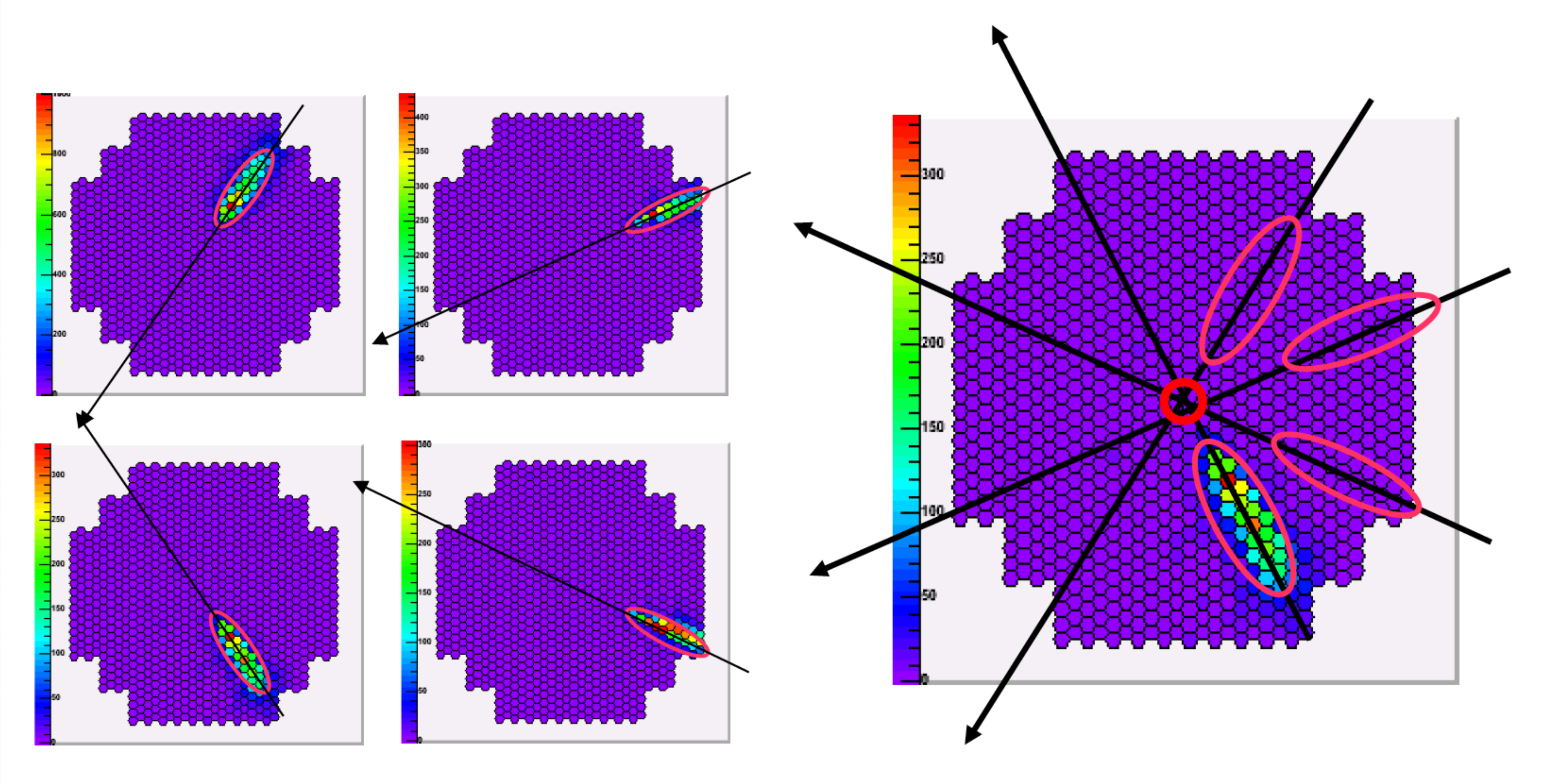
- Single telescope reconstruction
  - Degeneracy for source direction…
  - Additional parameters are added to remove the degeneracy
  - Recent hardware developments can help: addition of timing in image analysis
    - Possibility to resolve cascade development as a function of time
  - Example with a hadron:





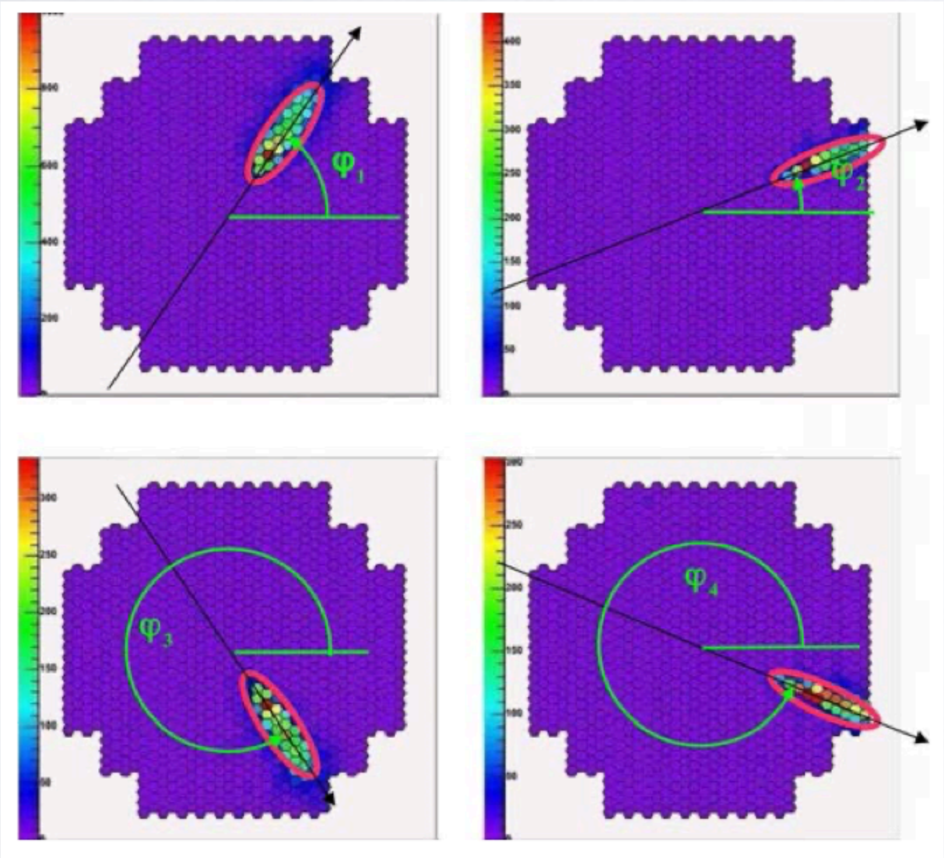
HILLAS RECONSTRUCTION

- Stereoscopy is used by all major experiments
  - Intersection of image axes in the camera frame gives the position of the source

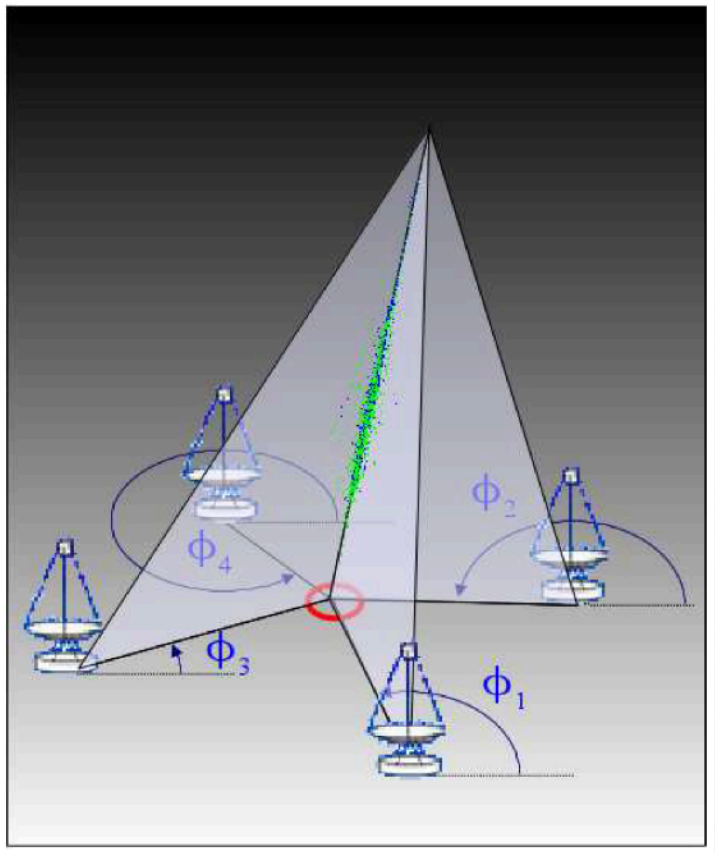


HILLAS RECONSTRUCTION

- Stereoscopy is used by all major experiments
  - Intersection of image axes in the camera frame gives the position of the source
  - Intersection of planes containing the telescope position and the shower axis gives the position of shower impact on the ground (also called shower core)
- NB: pointing accuracy, camera position in the focal plane have to be carefully controlled!



⚠️ These two figures do not depict the same situation!



HILLAS RECONSTRUCTION

- Gamma-hadron separation
  - Hillas parameters can be used to discriminate between photons and hadrons
  - Several methods exist. An example is the so-called **Scaled Cuts** method
  - The idea is to compare image width and length with the expectation values and variance obtained from simulations as a function of image charge **q** and impact distance **ρ**
    - **Scaled Width** (SW) and **Scaled Length** (SL)

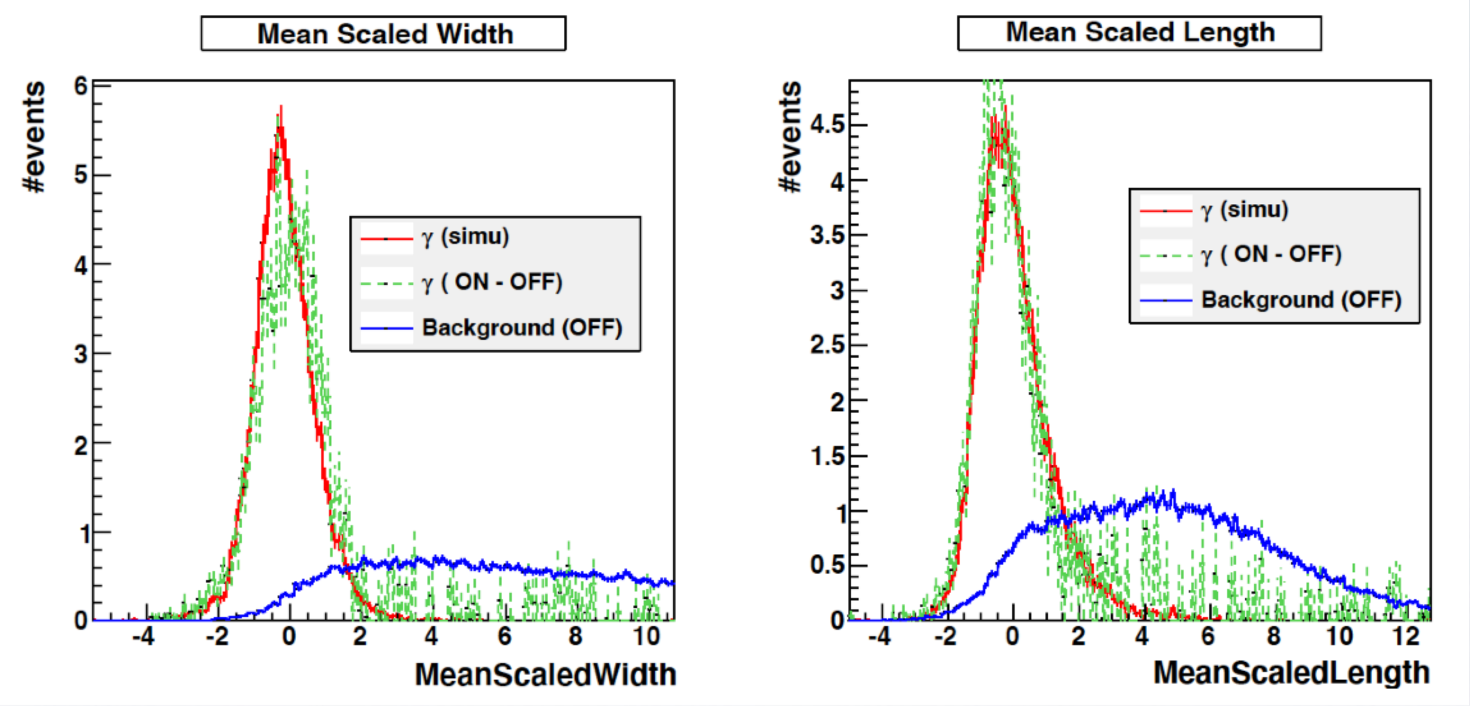
$$SW = \frac{w(q,\rho) - \langle w(q,\rho)\rangle}{\sigma_w(q,\rho)} \qquad SL = \frac{l(q,\rho) - \langle l(q,\rho)\rangle}{\sigma_l(q,\rho)}$$

  - For stereoscopic observations with **ntels** telescopes:
    - **Mean Scaled Width** (SW) and **Mean Scaled Length** (SL)

$$MSW = \frac{\sum_{tels} SW}{\sqrt{ntels}} \qquad MSL = \frac{\sum_{tels} SL}{\sqrt{ntels}}$$

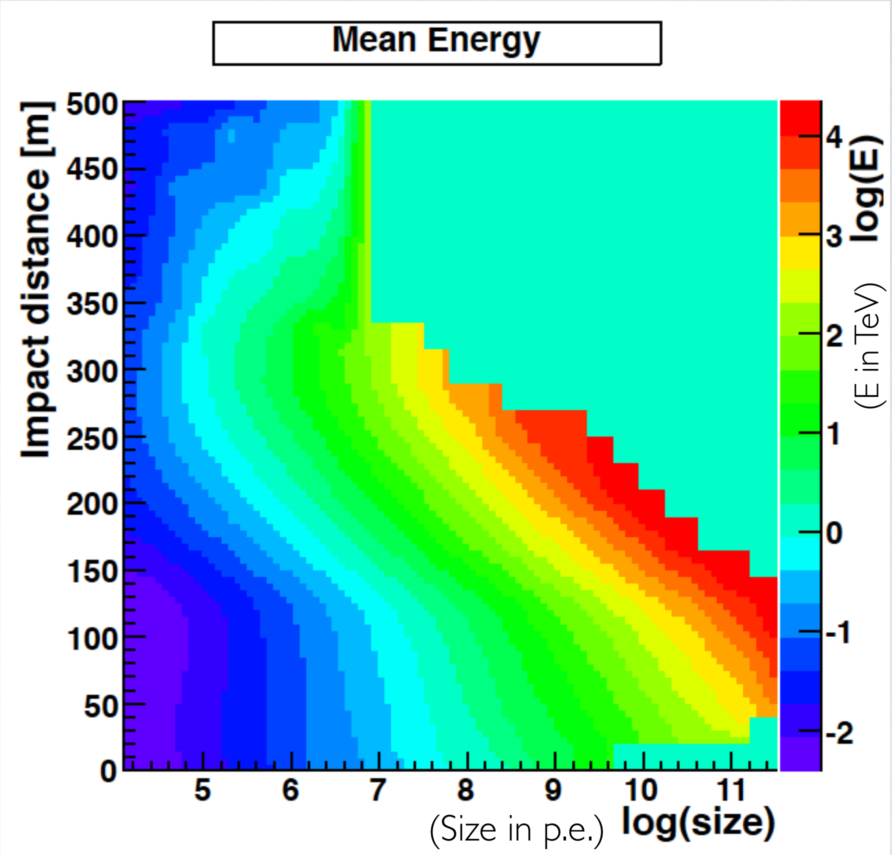
HILLAS RECONSTRUCTION

- Example of MSW and MSL distributions for the Crab Nebula
  Simulated gamma-rays in red, Gamma-like events from the Crab in green, Background events in blue



- Main interest of Scaled Cuts:
  - They are scaled by the variance of the parameters, e.g.
    - Small energy → big variance → small parameters
    - Big energy → small variance → big parameters
  - This will turn into the fact selection efficiency will not depend on energy
    → No bias for energy spectrum determination

HILLAS RECONSTRUCTION



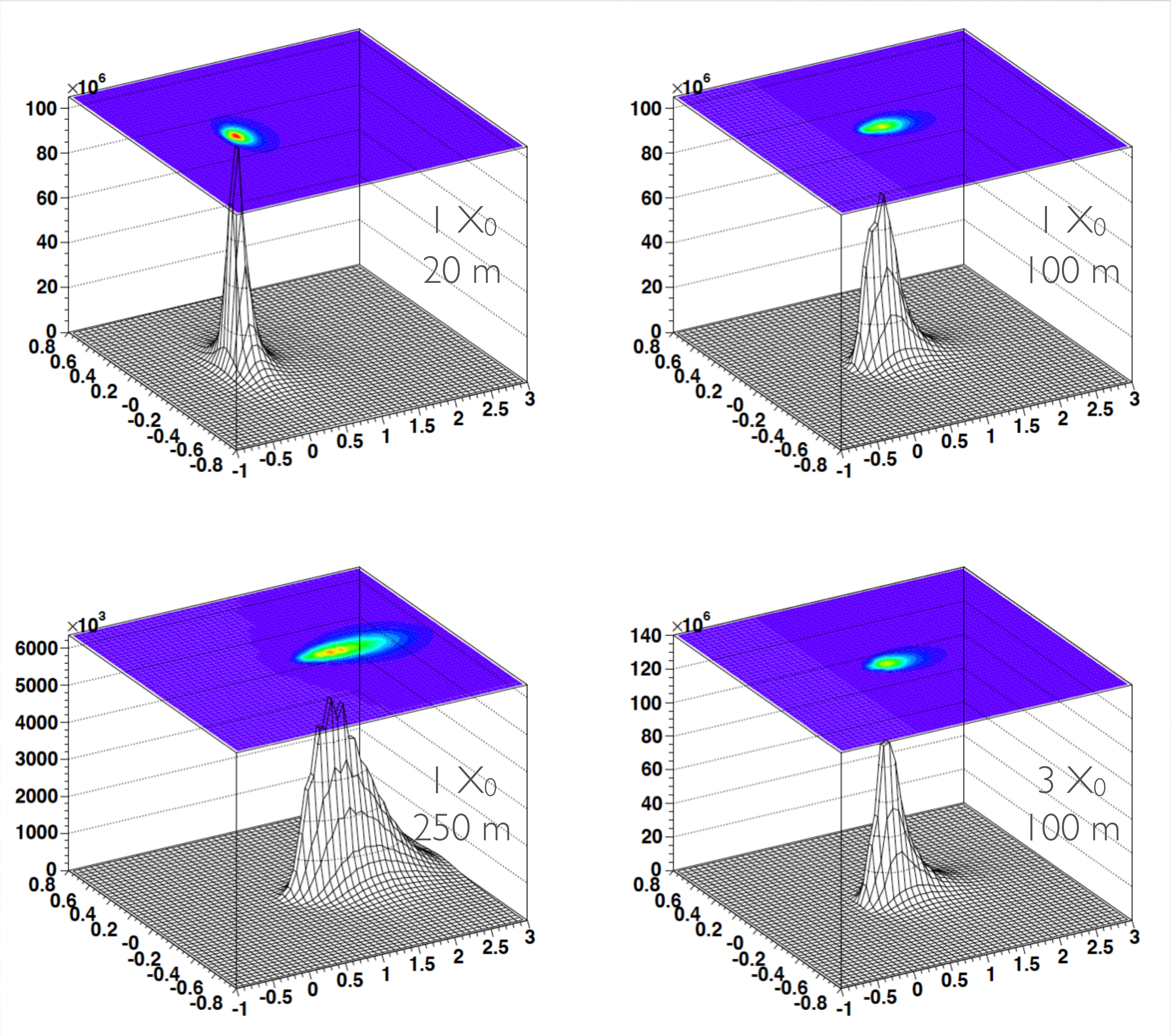
- Energy reconstruction
  - Use of simulated events to build look-up tables
  - Primary energy is deduced from impact parameter and image size (intensity)
- Tables
  - For each shower:
    - Reconstruct the impact point (the core)
    - Compute the intensity in each telescope
  - Events are binned by reconstructed distance to the telescope and image intensity
  - For each of these bins, average and variance of *true* energy is computed
  - This is repeated for 20 bins for zenith angle, 7 bins for optical efficiency, 6 bins for off-axis angle. Intermediate values are obtained through interpolation
- This interpolation allows to get the reconstructed energy for each telescope from the size, using the impact point reconstructed with the full array
- Final energy estimation is obtained from a weighted average. Energy resolution ~20%, with biases at low and high energies

MODEL RECONSTRUCTION

- **Model reconstruction** was developed by De Naurois & Rolland (2009) from early works done in CAT in the 90s (Le Bohec, 1998)
- NB: Cherenkov light distribution of a shower is given by the longitudinal, lateral and angular distributions of charged particles
- The method:
  - Produce these distributions from Monte Carlo simulations
  - Parameterize the distributions → shower model
    - The first interaction depth parameter is taken into account
  - Produce the images in the cameras, adding the noise due to NSB in each pixel
  - Reconstruction and discrimination are achieved through the comparison between raw images and modeled raw images
    - Done with a log-likelihood maximization technique
- In HESS, this technique allows a x2 improvement of sensitivity compared to Hillas reconstruction

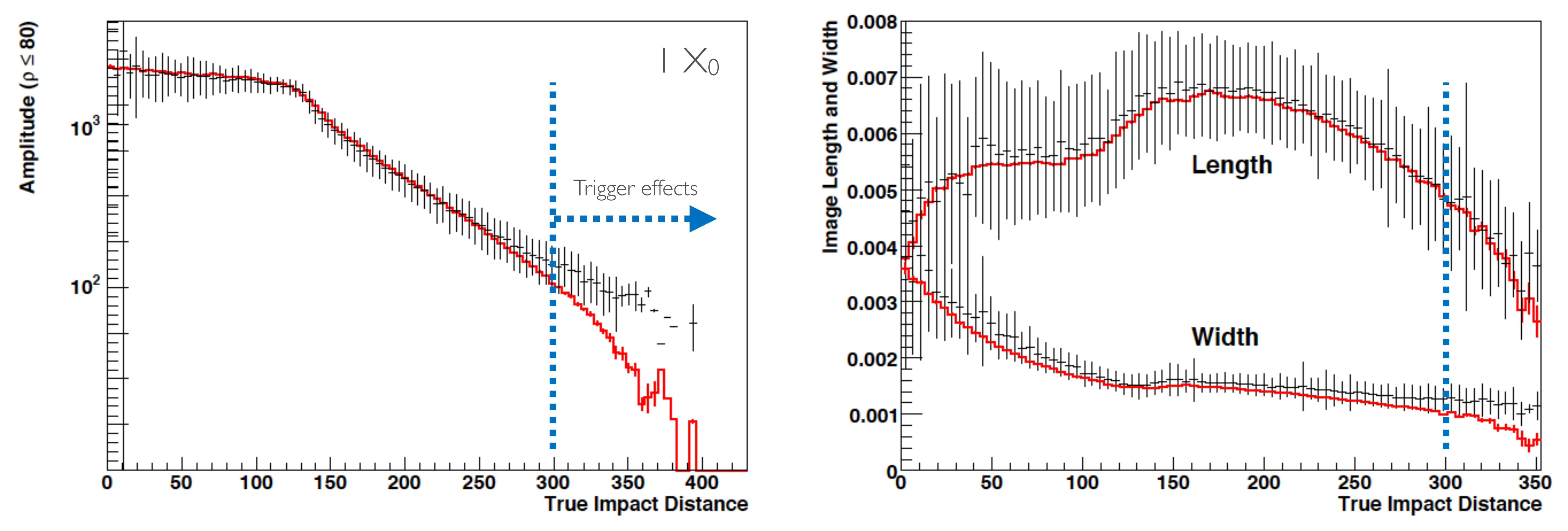
MODEL RECONSTRUCTION

- Examples of modeled shower images
- 1 TeV vertical showers
- X,Y units: degrees in the camera frame
- Z scale: intensity in p.e.
- 0.01° pixel size



MODEL RECONSTRUCTION

- It is important that the model is accurate
- Need to compare it with simulations
- Example:



- Average and shower-to-shower fluctuations from simulations is shown in black (1 TeV, zenith)
- Model prediction in red, pixel size 0.16° (H.E.S.S.)

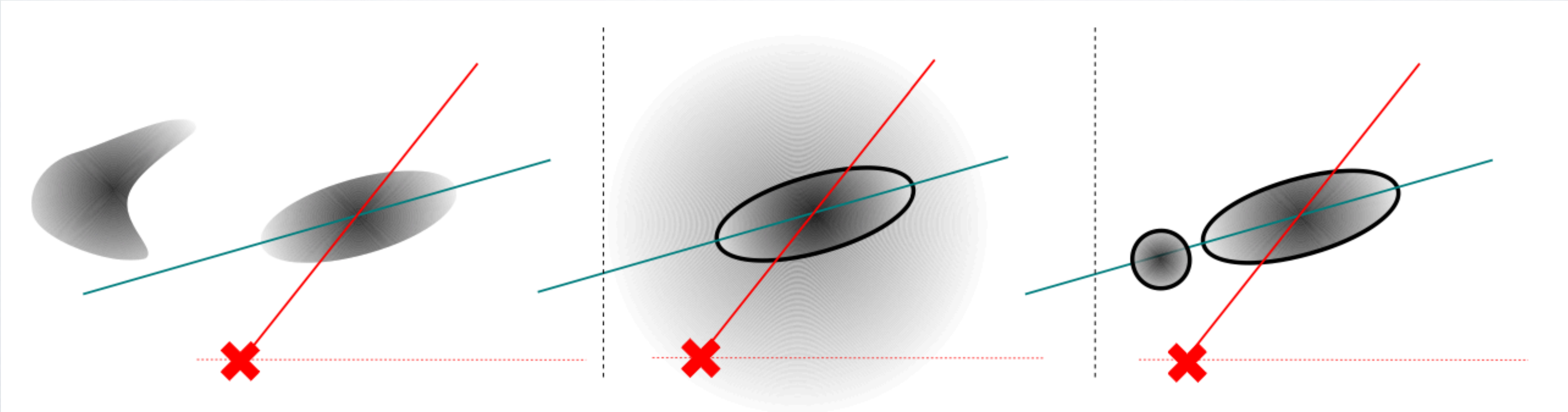
MODEL RECONSTRUCTION

- A likelihood fit is used to reconstruct energy, direction, impact, first interaction depth by comparing real raw images and models
    - This comparison is done at the pixel and p.e. level!
    - The main ingredient is the charge distribution in each pixel
    - Pixels are assumed to be independent
- The output of the fit
    - Best guess for the 6 shower coefficients: direction (2), impact (2), depth of first interaction, energy
    - Uncertainties
    - Final log-likelihood value
- Compatibility between recorded events and pure gamma-ray hypothesis is assessed from Goodness-of-fit:

$$G = \frac{\sum_{\text{pixel } i} \left[\overbrace{\boxed{\ln L(s_i|\mu_i)}}^{\text{Pixel log-likelihood}} - \overbrace{\boxed{\langle \ln L \rangle |_{\mu_i}}}^{\text{Pixel expectation value}} \right]}{\underbrace{\boxed{\sqrt{2 \times \text{NDoF}}}}_{\text{Normalization}}}$$

MODEL RECONSTRUCTION

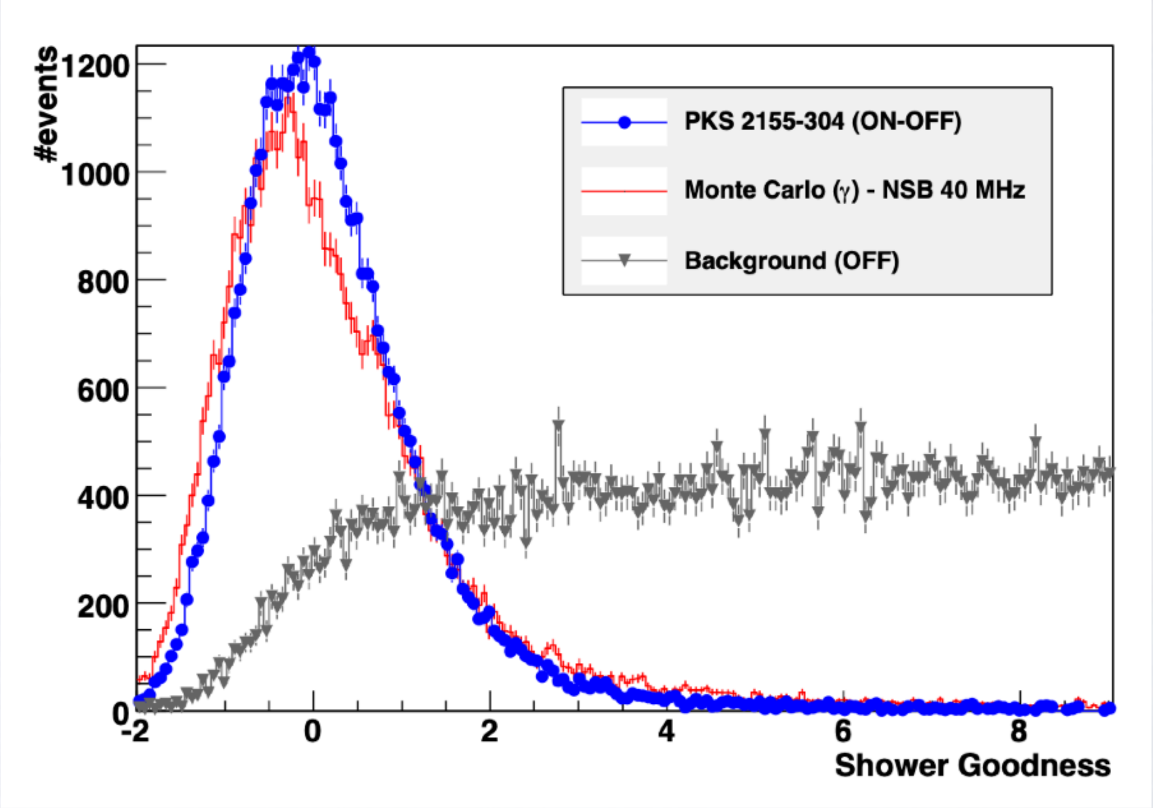
- Discrimination between gamma-ray and hadron induced showers relies on several facts:
    - Hadron induced showers are irregular → image usually show clusters
    - Hadronic part of the shower produce low intensity Cherenkov light which is spread over the camera → « hadronic rain »
    - Primary charged particle can produce Cherenkov light



MODEL RECONSTRUCTION

- From this, we can build discrimination variables
- **Shower Goodness** (SG)
  - Goodness-of-fit, using pixels above 0.01 p.e. grouped together with three rows of neighbors
- Background Goodness (BG)
  - Goodness-of-fit, using all the other pixels

- As already done before, we can define another parameter using images from several telescopes: **Mean Scaled Shower Goodness** (MSSG)



RECONSTRUCTION

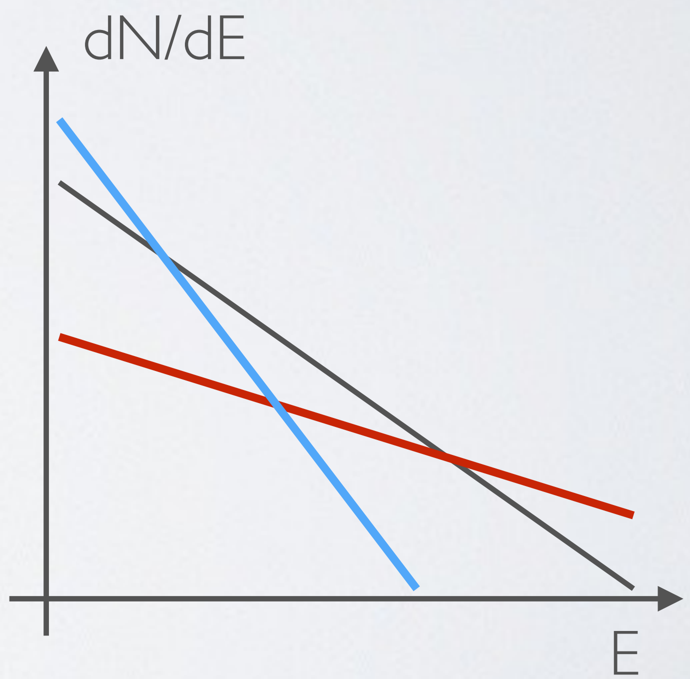
- Where we are now
  - Two methods for reconstructing event parameters
  - These methods can be used also as discrimination tools
- Before reconstruction, we need to chose a number of initial selections such as
  - Minimum charge (p.e.) in the camera
  - Maximum nominal distance (deg.)
  - Minimum number of telescopes required
  - Maximum Shower Goodness
  - Primary depth range
  - Max θ^2 (θ is the angle between the expected position of the source and the reconstructed position) → post-reconstruction criteria
- Each set of cuts will lead to
  - Different performance!
  - Different list of reconstructed events!

RECONSTRUCTION

- Examples:

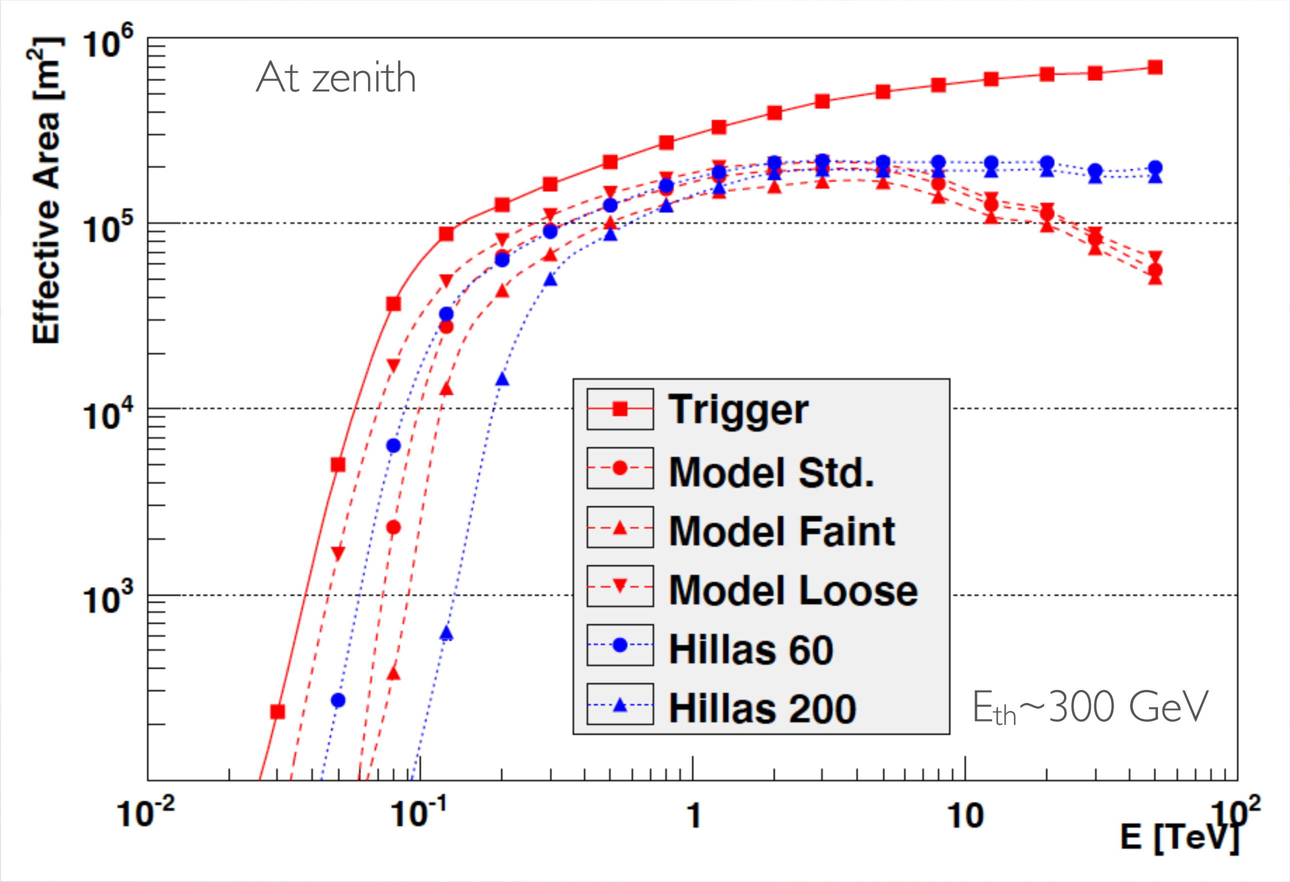
Name	Min. Charge (p.e.)	Max. Nom. Distance (deg.)	#Tels	$SG_{\max}$	t_0 (X_0)	$\theta^2_{\max}$ (deg^2)
Standard	60	2	2	0.6	$[-1, 4]$	0.01
Faint Source	120	2	2	0.4	$[-1, 4]$	0.005
Loose Cuts	40	2	2	0.9	N/A	0.0125

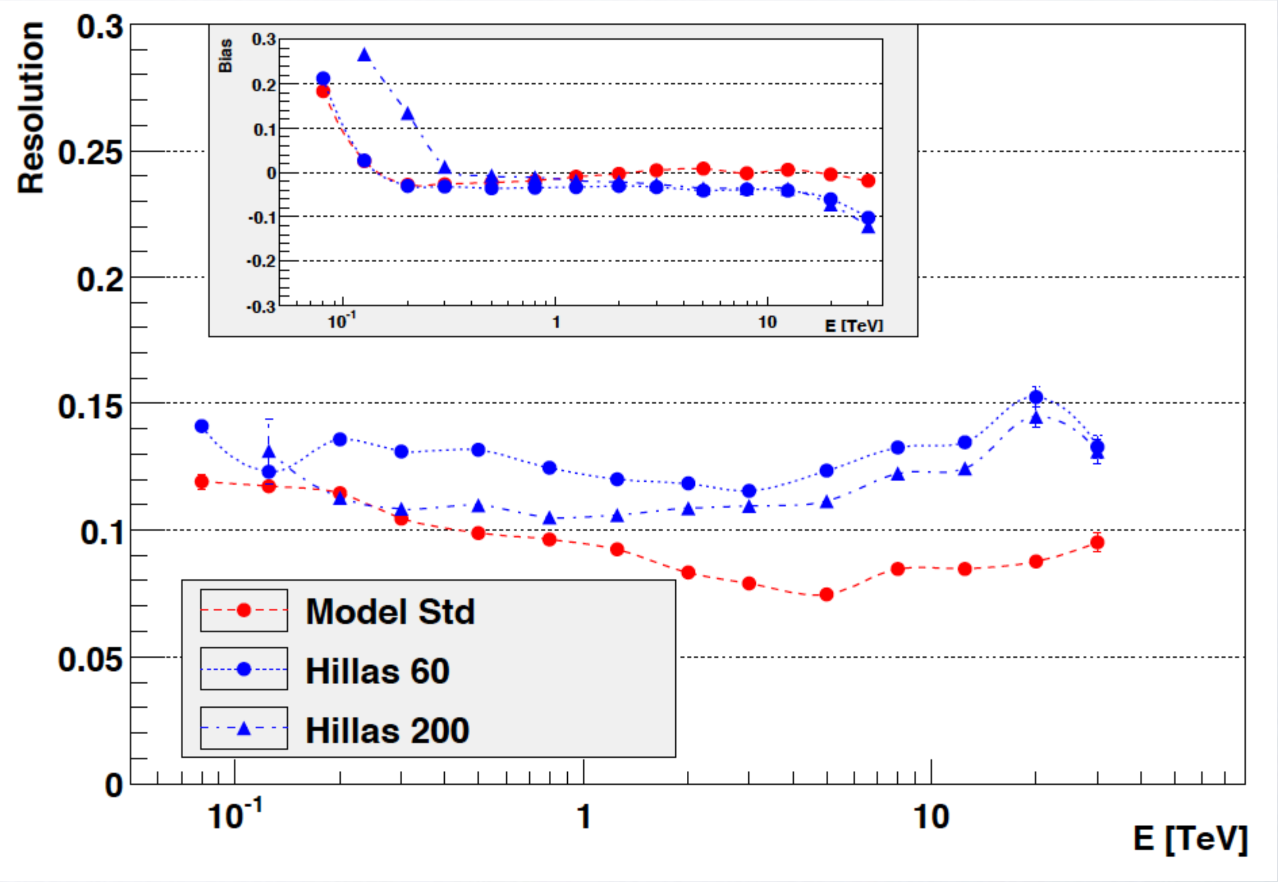
- Min. Charge has a direct influence on energy threshold
- **Faint source**: for low flux, hard spectrum
  - Good angular resolution, so smaller cut on θ^2
- **Loose Cuts**: for high flux, soft spectrum



RECONSTRUCTION

- **Performance of reconstruction** is assessed through
  - Its ability to reconstruct shower parameters with accuracy
  - Its ability to reject hadron background
- **Performance of the detector** is commonly quantified by **effective area,** E_{true} vs E_{rec} relation i.e. **migration matrix**, which also contains values for **energy resolution** and **bias** → **Instrument Response Functions** (IRF)
  - Depend on zenith angle, offset angle, observation conditions, reconstruction strategy…

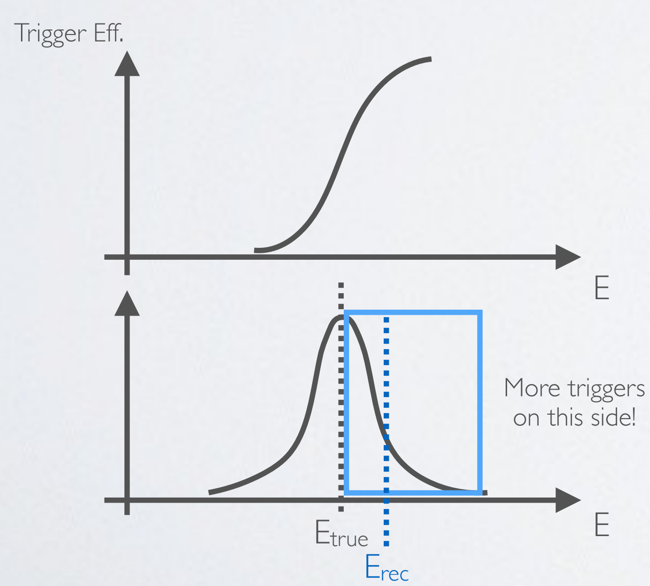


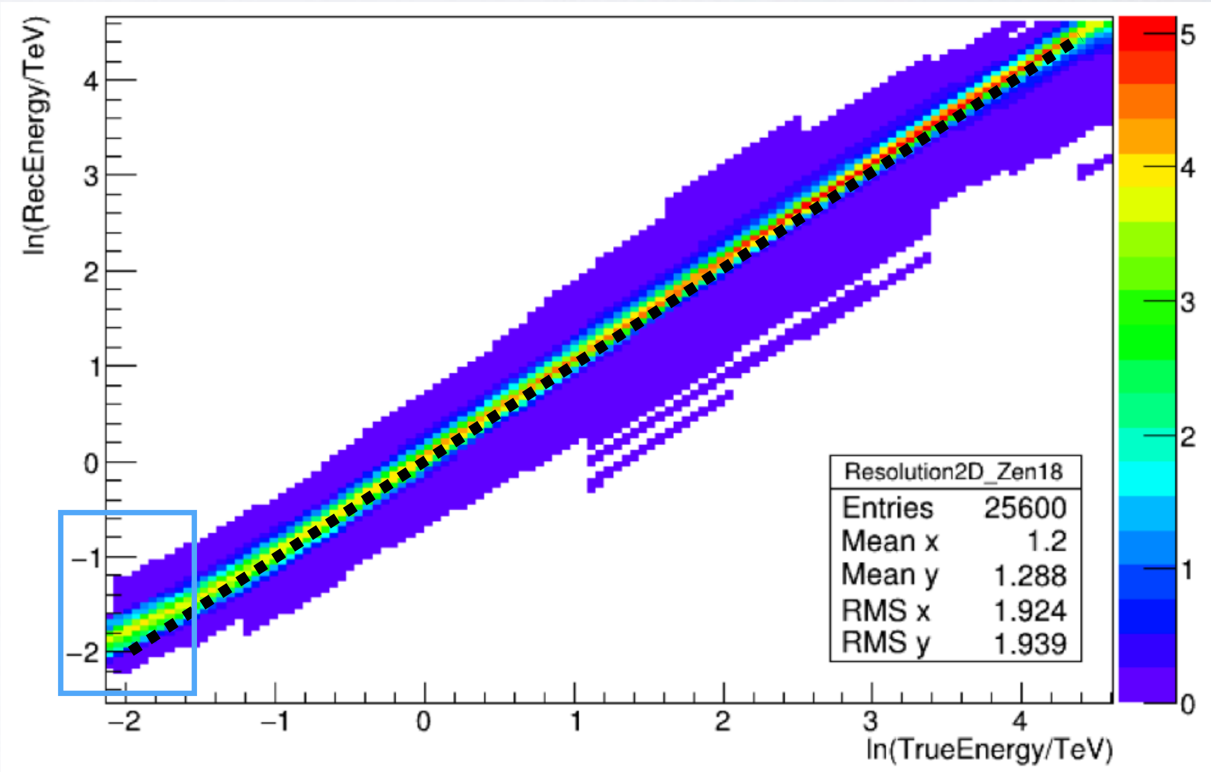


RECONSTRUCTION

- **Performance of reconstruction** is assessed through
  - Its ability to reconstruct shower parameters with accuracy
  - Its ability to reject hadron background
- **Performance of the detector** is commonly quantified by **effective area,** E_{true} vs E_{rec} relation i.e. **migration matrix**, which also contains values for **energy resolution** and **bias** → **Instrument Response Functions** (IRF)
  - Depend on zenith angle, offset angle, observation conditions, reconstruction strategy…

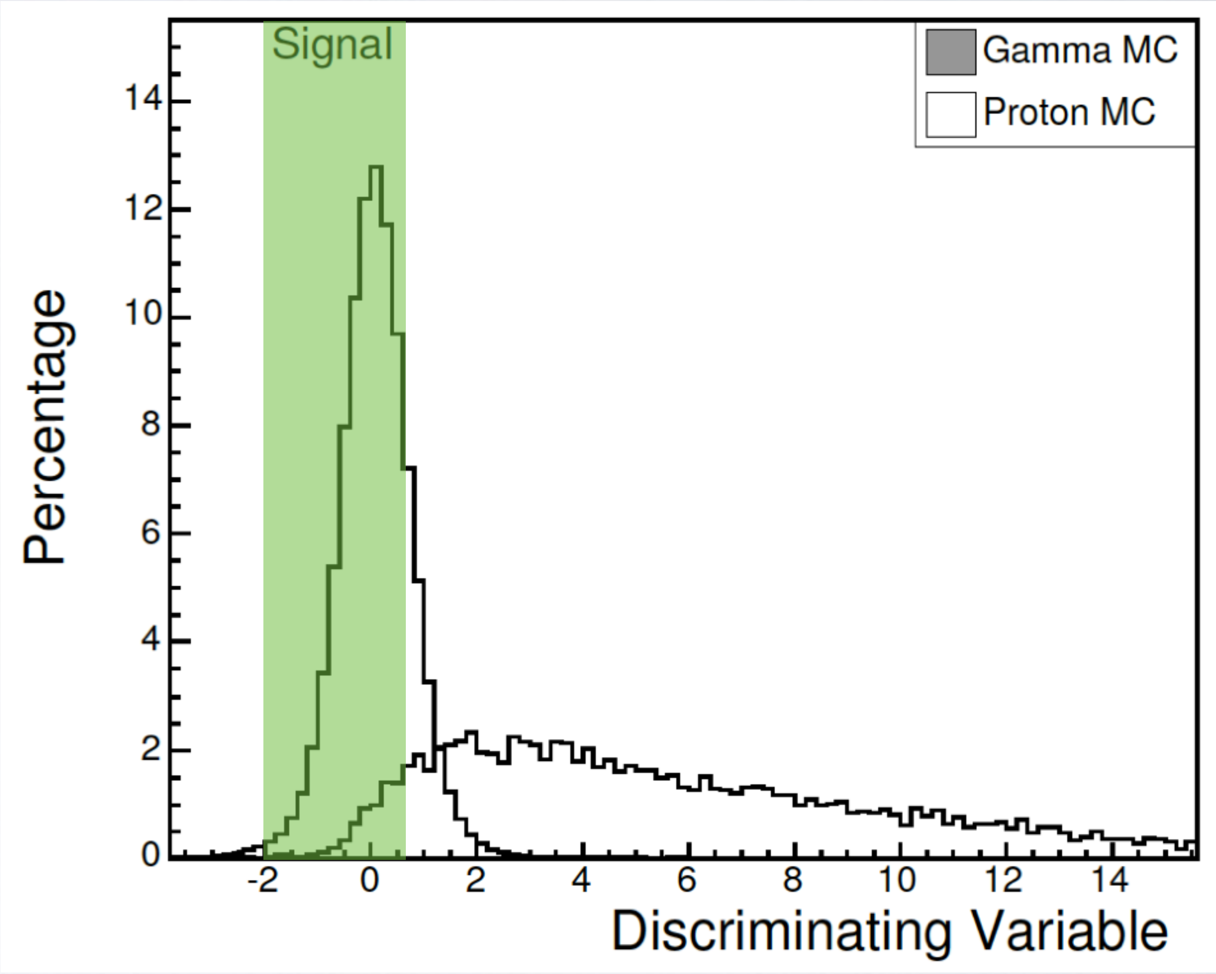
NB: at low energies, trigger efficiency decreases and reconstructed energy is overestimated





FROM RECONSTRUCTION TO ANALYSIS

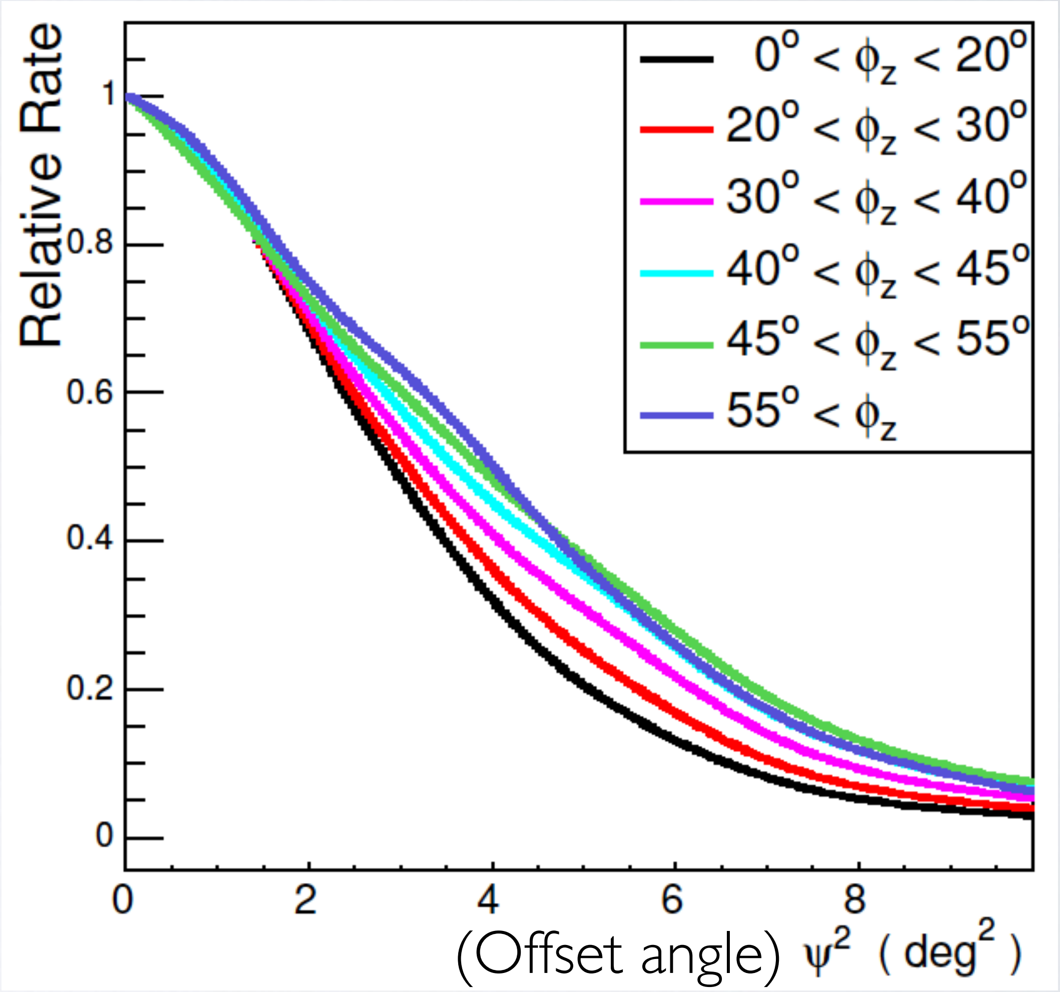
- The reconstruction step gives us a list of **gamma-like** events (« **Signal** »), including
  - Real gamma from the source under study
  - Real gamma from background
  - Real electrons
  - Mis-identified hadrons
- We now want to focus as much as possible on source gamma events to extract source properties:
  - Spectrum
  - Light curve
  - Morphology



- So, we need to estimate the background as precisely as possible…
  - Goal of the analysis

ANALYSIS

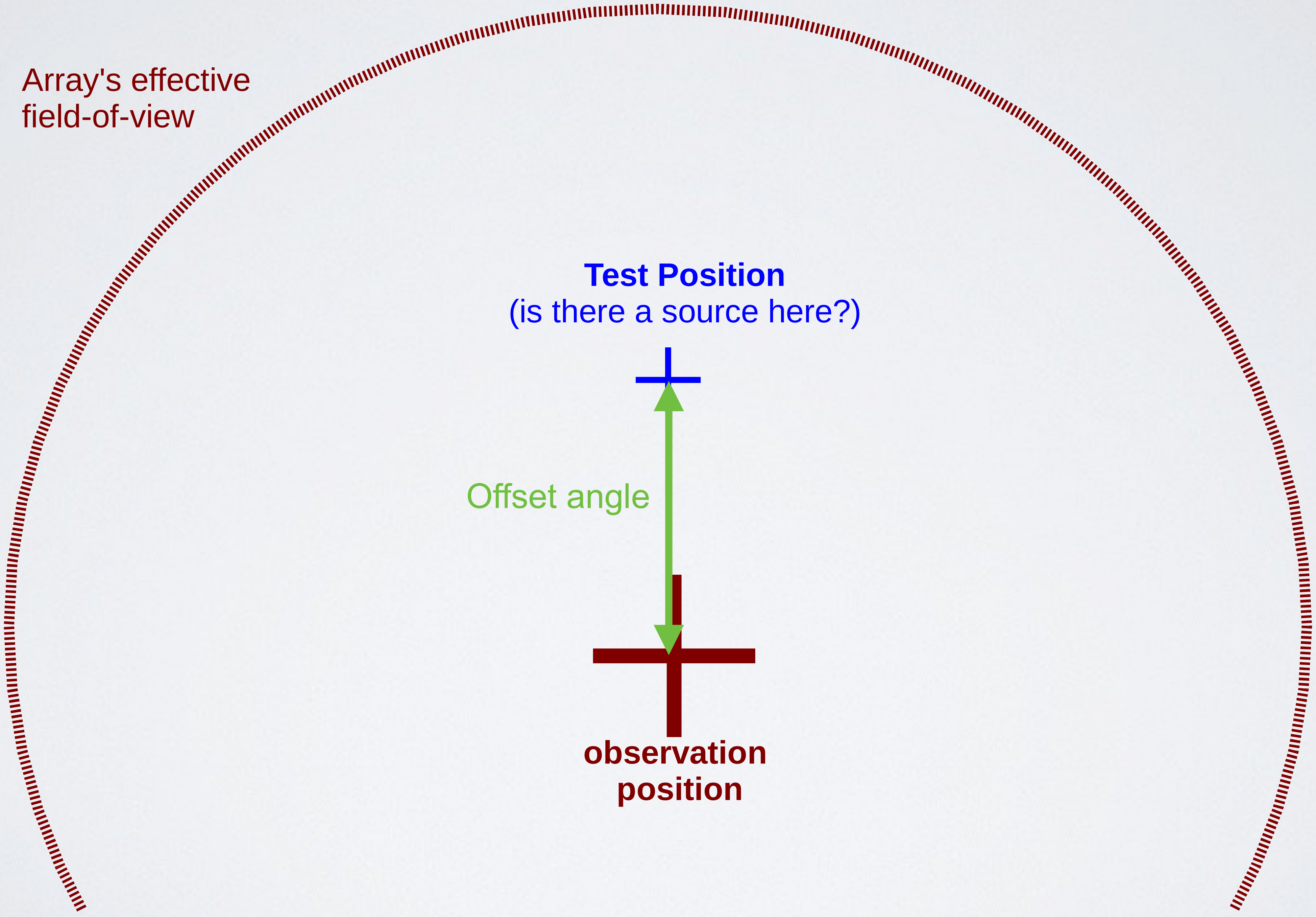
- The detection of a source usually requires multiple runs
- These runs are usually taken in different observation conditions
  - Zenith angle
  - Offset angle
  - Presence of hardware problems
- The goal of the **analysis** is to combine these runs and precisely estimate the background
- Some background estimation techniques require to know the **acceptance** of the detector → heavy use of **simulations**
  - Probability that an event of a given energy and type triggers the system, is reconstructed, and is selected according to the considered event class



- It's not the case for the « **Multiple-OFF** » or « reflected background » technique

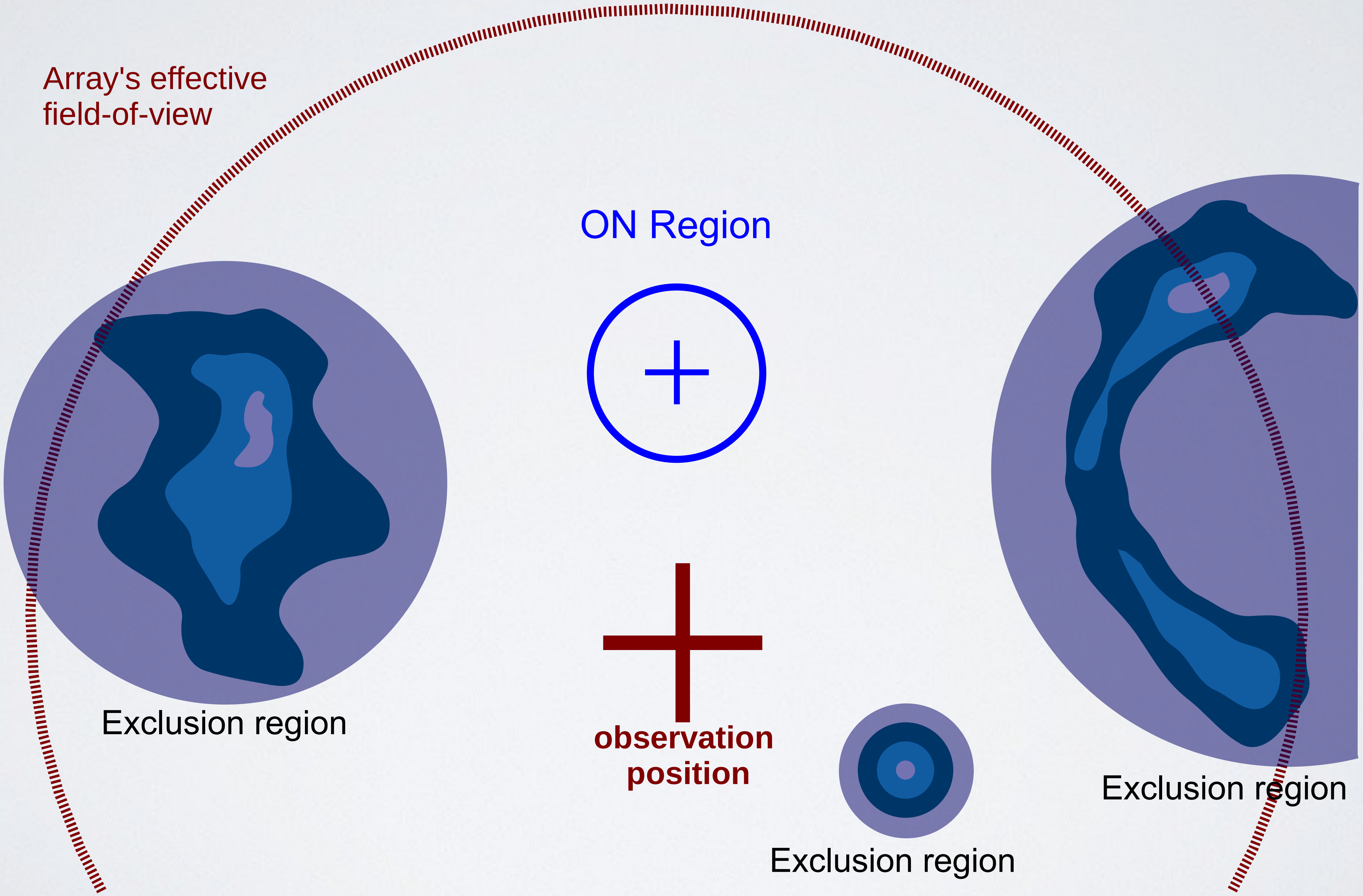
BACKGROUND ESTIMATION

- Say we want to investigate a particular position in the sky



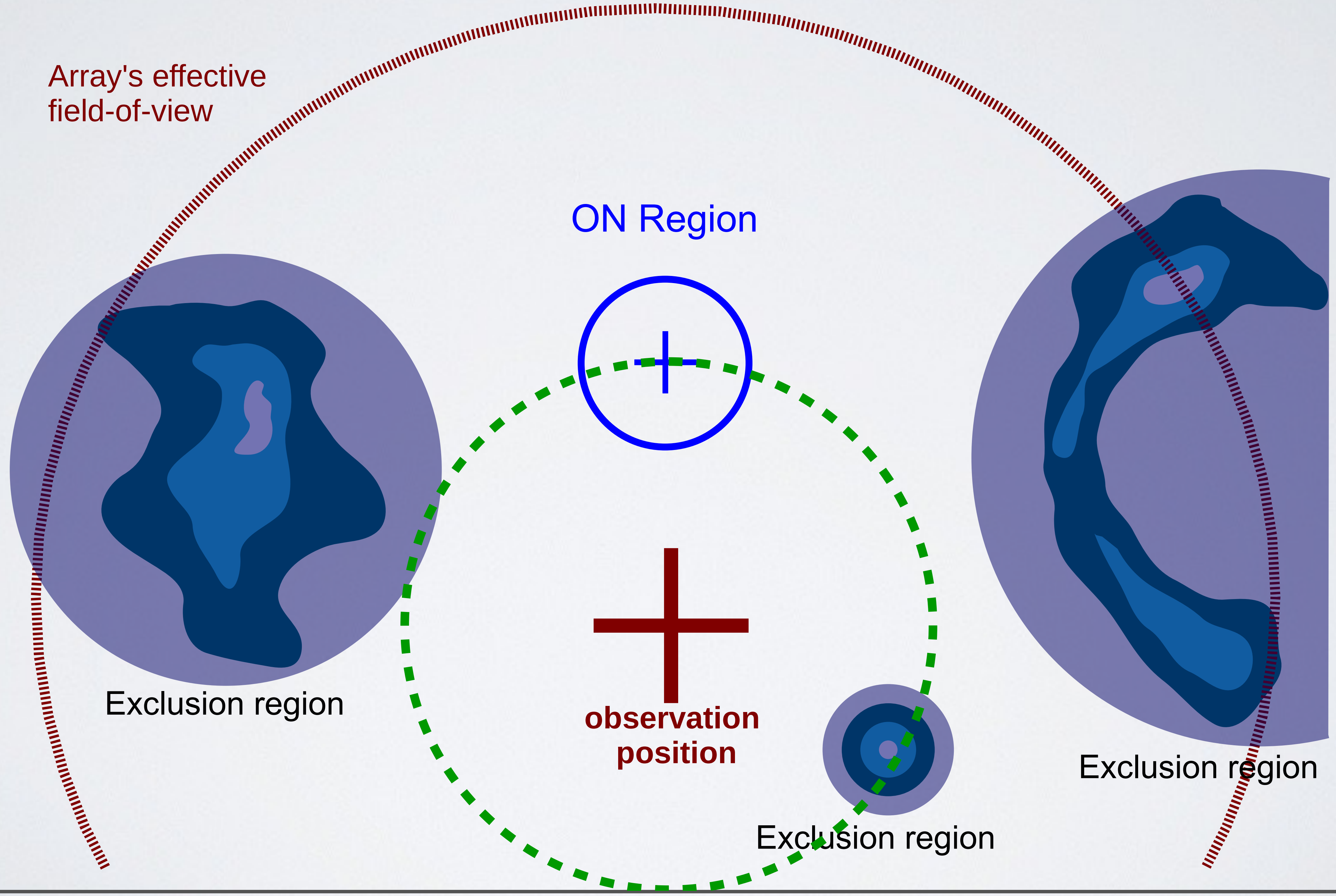
BACKGROUND ESTIMATION

- Problem: there may be other sources in the same FOV! → exclusion regions
- The **ON Region** is centered on the expected position of the source



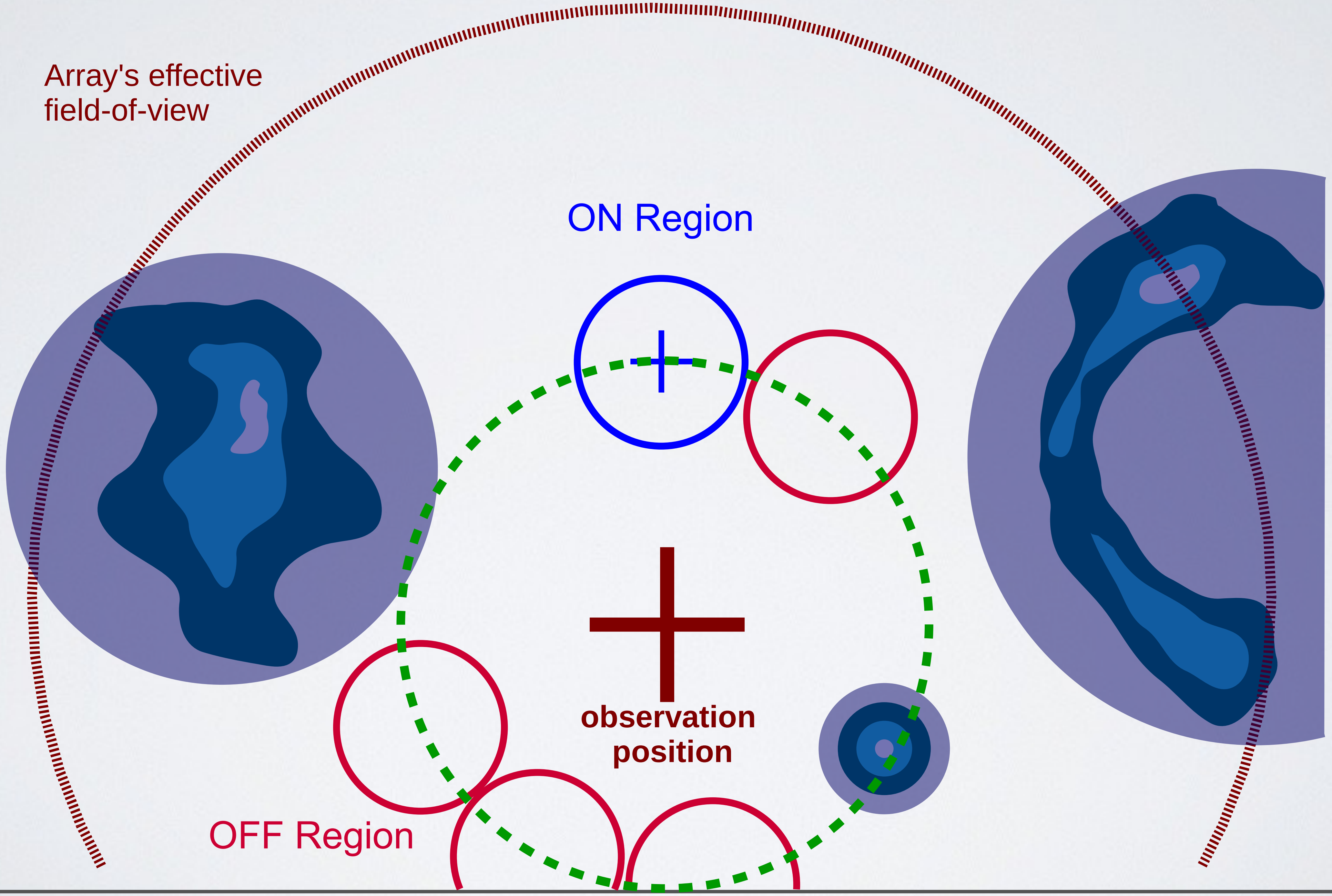
BACKGROUND ESTIMATION

- **Acceptance** is radially symmetric, so we consider a circle around the observation position



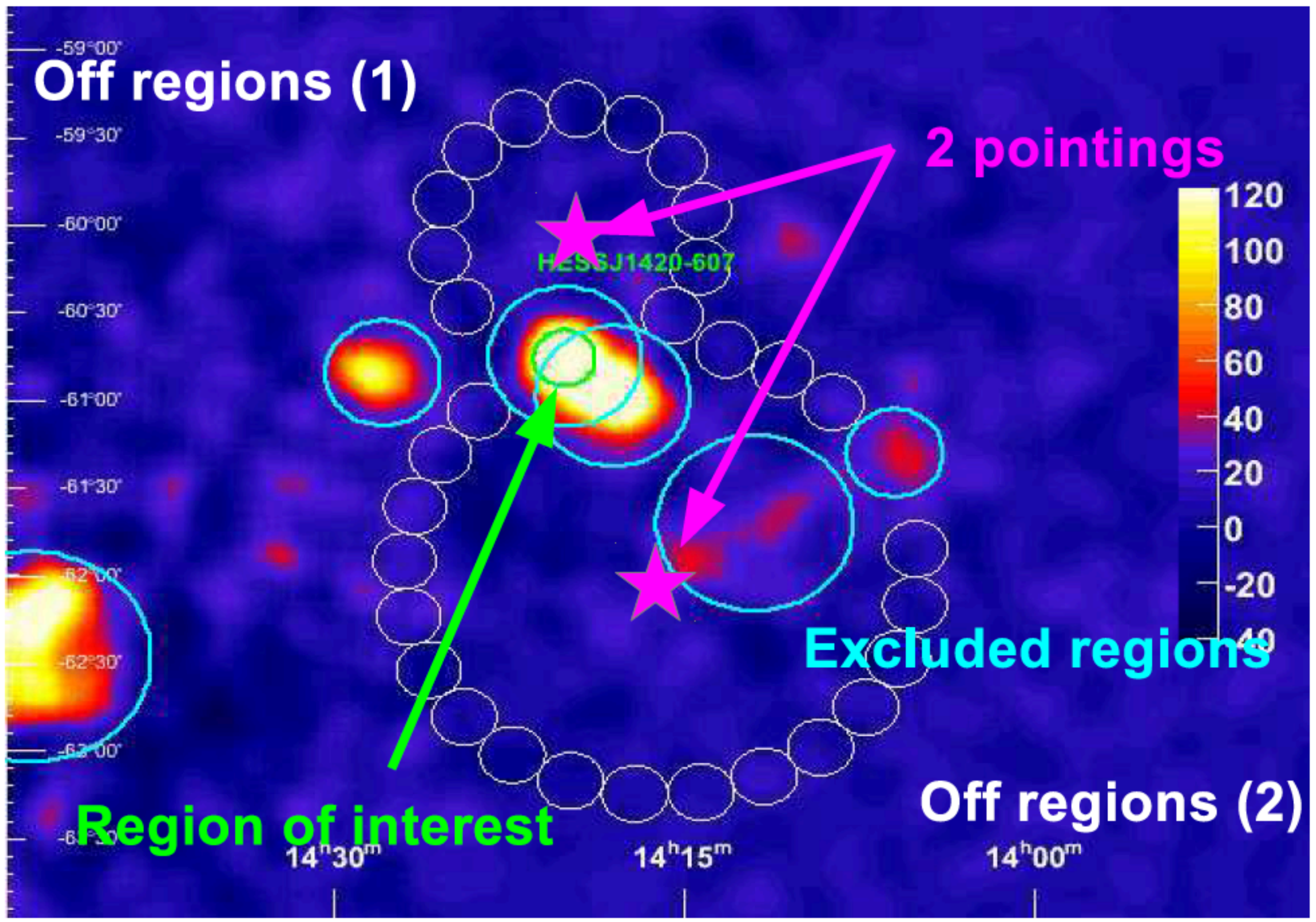
BACKGROUND ESTIMATION

- Background is taken in **OFF Regions** which are centered on that ring
- There can be several OFF Regions → « **Multiple OFF** »



BACKGROUND ESTIMATION

- A real example for two different runs



- When a high number of runs is necessary, the whole process needs to be automated…
- A database with exclusion regions is updated regularly

SIGNIFICANCE

- We define:
  - **N_{ON}** the number of events in the region of interest (ROI)
  - **N_{OFF}** the number of events in control regions (OFF regions) with only background events
  - **α** the ratio of the expected number of background events in the ROI over the actual number of background events accumulated in OFF regions
    - α can be understood as the ratio of exposures between the ON region and the OFF regions
  - The **excess** is given by **$N_{ON} - \alpha N_{OFF}$**
  - **Significance** is obtained from Li & Ma (1983):

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

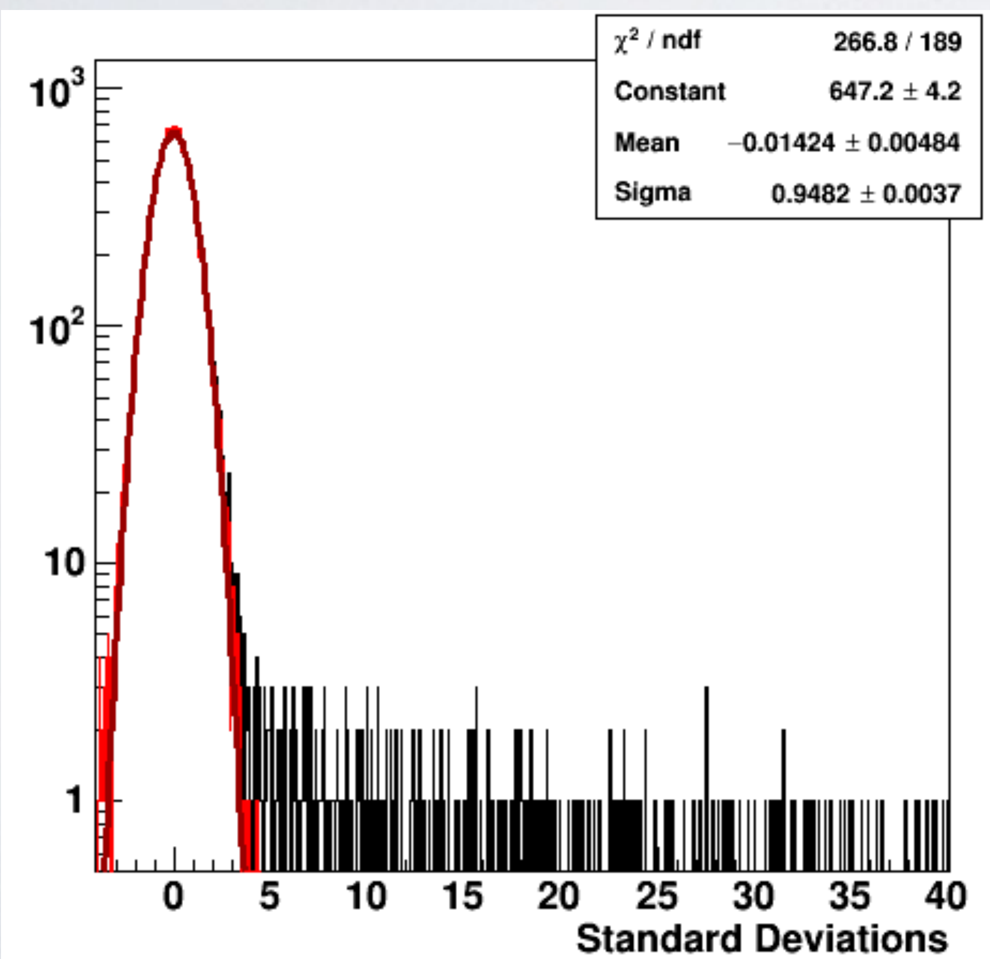
with

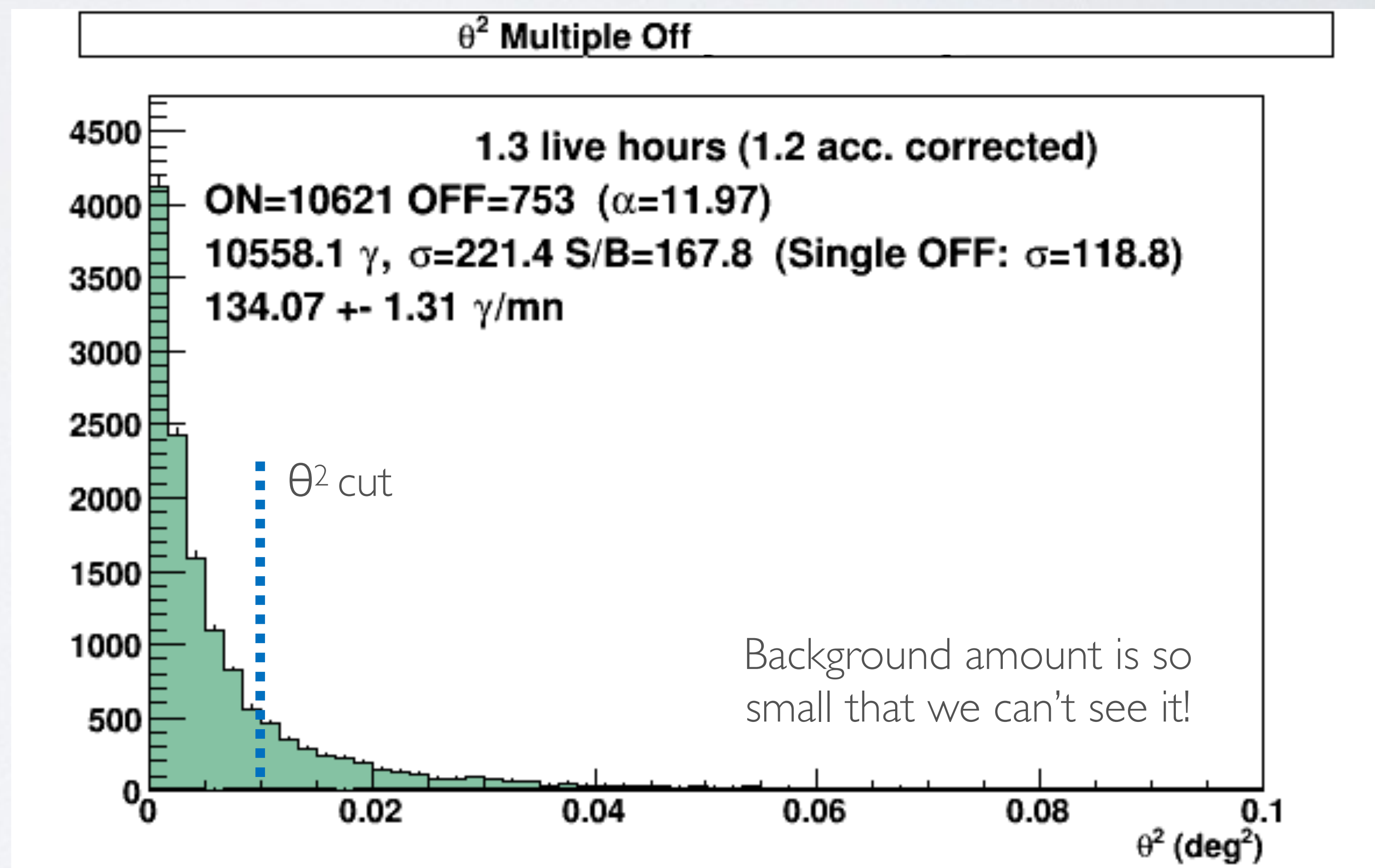
$$\lambda = \left[\frac{\alpha}{1+\alpha}\left(\frac{\mathrm{N_{on}} + \mathrm{N_{off}}}{\mathrm{N_{on}}}\right)\right]^{\mathrm{N_{on}}} \times \left[\frac{1}{1+\alpha}\left(\frac{\mathrm{N_{on}} + \mathrm{N_{off}}}{\mathrm{N_{off}}}\right)\right]^{\mathrm{N_{off}}}$$

  - λ is the **likelihood-ratio** between two hypotheses:
    - N_{ON} events result from Signal + α Background
    - N_{ON} events result from background only (*null hypothesis*)

SIGNIFICANCE

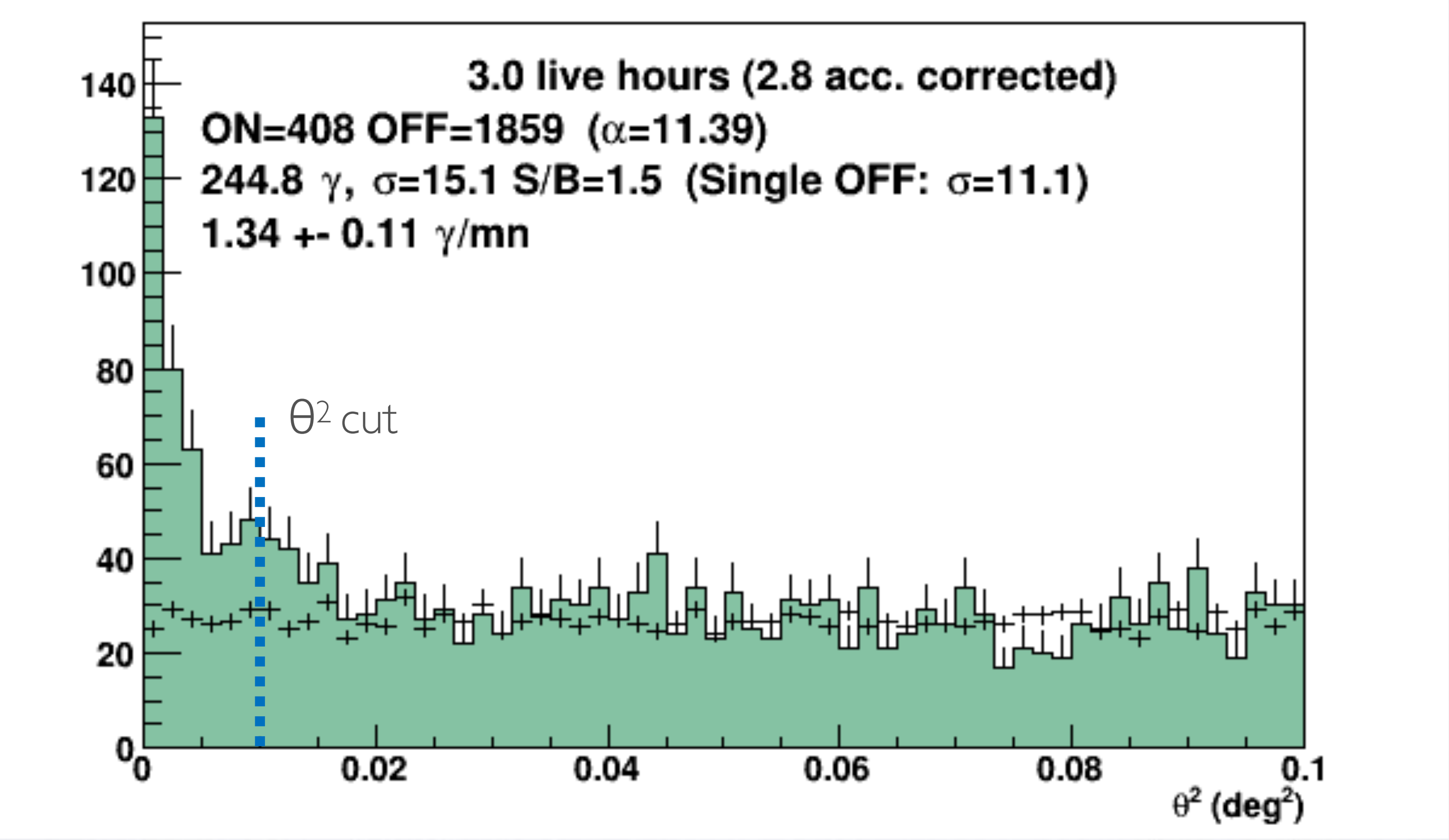
- For pure Poisson fluctuations for both ON and OFF events, S follows a normal distribution of average 0 and width 1
- Excess events show up as a tail on top of this distribution
- Another control plot: θ^2 distribution





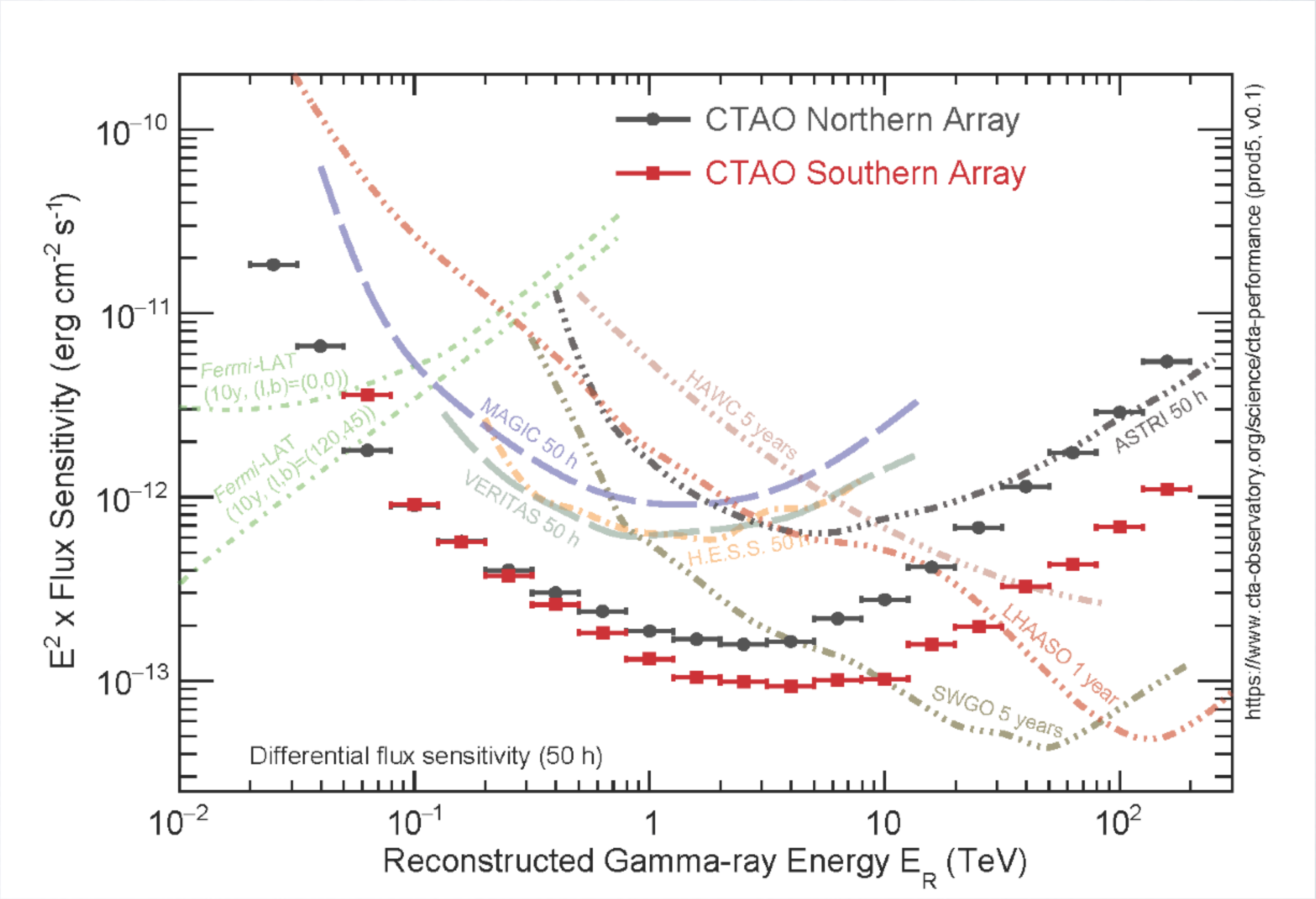
SIGNIFICANCE

- For pure Poisson fluctuations for both ON and OFF events, S follows a normal distribution of average 0 and width 1
- Excess events show up as a tail on top of this distribution
- Another control plot: θ^2 distribution



SENSITIVITY

- For a given analysis chain and a given set of cuts, it is possible to estimate the **sensitivity** of the array
- **Sensitivity**: minimum flux which can be detected with Nσ significance during T hours of observations. Other criteria may be added: 10 detected gamma-rays per energy bin, S/B > 1/20…
- Example for point-like sources:



WHAT'S NEXT?

- From now on, we will focus on a simple case: point source, with negligible background
    - This can happen for e.g. a flaring AGN or a GRB!
- We will also consider that the source is detected with a significant signal
    - Usually a detection is claimed when $S > 5\ \sigma$
- We are usually interested in
  - The energy spectrum of the source
  - Its light curve
    - Ingredients for LIV studies!

- We'll now briefly see how to get the spectrum

SPECTRUM

- A commonly used technique for spectrum extraction is the so-called « **forward folding** » technique
- It requires a good knowledge of Instruments Response Functions (IRFs): migration matrix (E_{true} vs E_{rec}), energy resolution $\mathscr{R}$, effective area $\mathscr{A}$
  - IRFs are different for each reconstruction method and depend on selection cuts!
  - IRFs vary with observation conditions!
- Steps:

- Assume a spectral shape $\boldsymbol{\Phi}$ (power law, broken power law, etc.)
- Compute in each bin of energy the **expected** number of gamma events:

$$n_\gamma = \int_{E_{rec,1}}^{E_{rec,2}} \mathrm{d}E_{rec} \int_0^\infty \mathrm{d}E_{true} \mathcal{R}(E_{rec}, E_{true}|\delta, \Psi, \epsilon) \times \mathcal{A}(E_{true}|\delta, \Psi, \epsilon) \times \Phi(E_{true}|\vec{\alpha})$$

  with $\boldsymbol{\delta}$ the zenith angle, $\boldsymbol{\Psi}$ the off-axis angle, $\boldsymbol{\varepsilon}$ the optical efficiency

- A log-likelihood is built from the number of ON and OFF events, taking into account a proper live time normalization $\boldsymbol{\beta}$

$$P(N_{ON}, N_{OFF}|n_\gamma, n_h) = \frac{n_\gamma + \beta n_h{}^{N_{ON}}}{N_{ON}!} e^{-(n_\gamma+\beta n_h)} \times \frac{n_h{}^{N_{OFF}}}{N_{OFF}!} e^{-n_h}$$

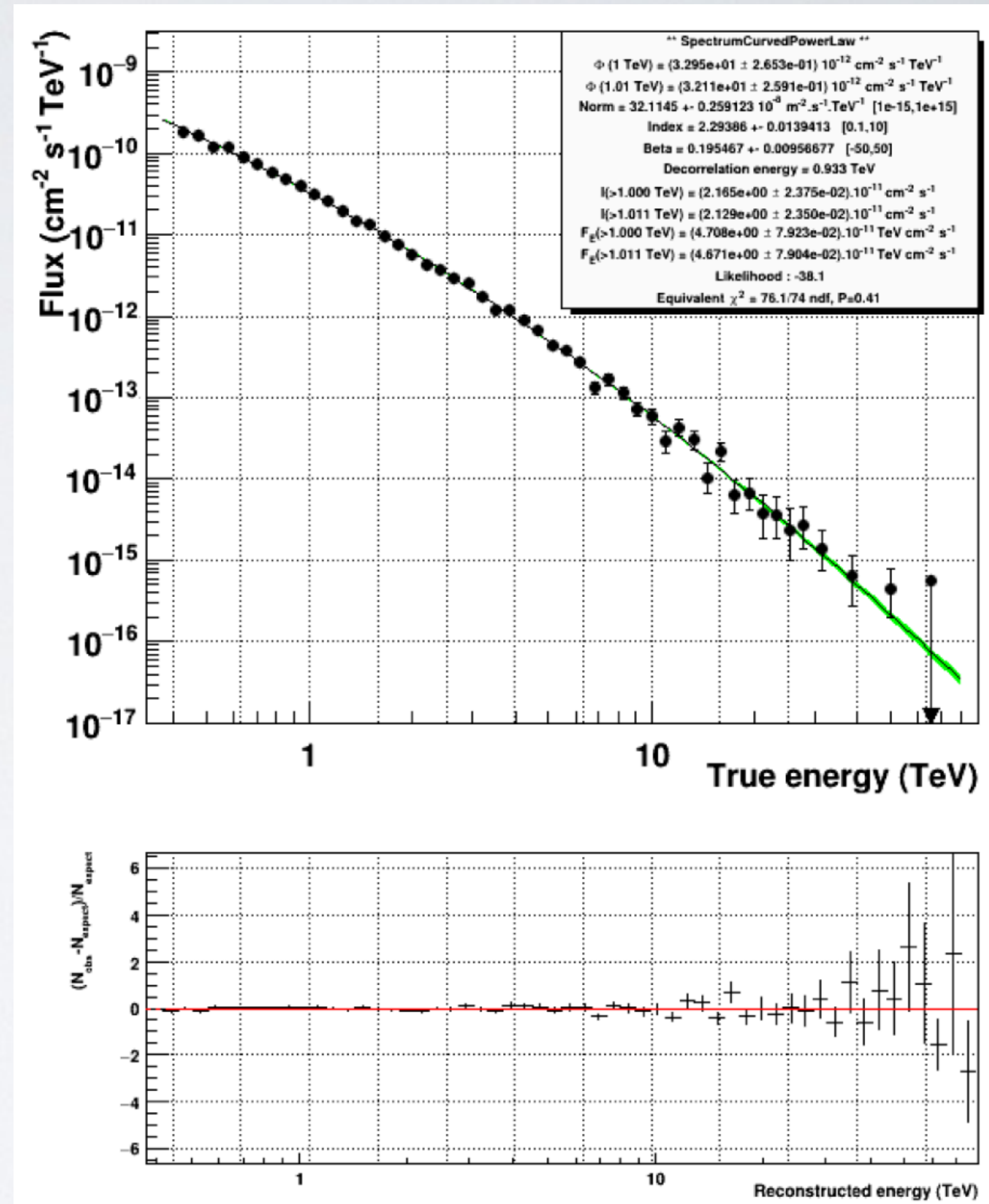
  where n_γ and n_h are the expected numbers of gamma and background events in the energy bin.

- The log-likelihood is then maximized

SPECTRUM

- The results of the maximization are:
  - The best fit parameters α
  - The expected number of events in each bin
  - The covariance matrix → used to derive flux uncertainty in each bin
  - The final value for log-likelihood → ≈ goodness-of-fit
- Flux points are obtained from the model using best fit parameters and residuals obtained in each energy bin

- Main advantage of the forward-folding technique: N_{ON} and N_{OFF} used directly, no need for flux point calculations

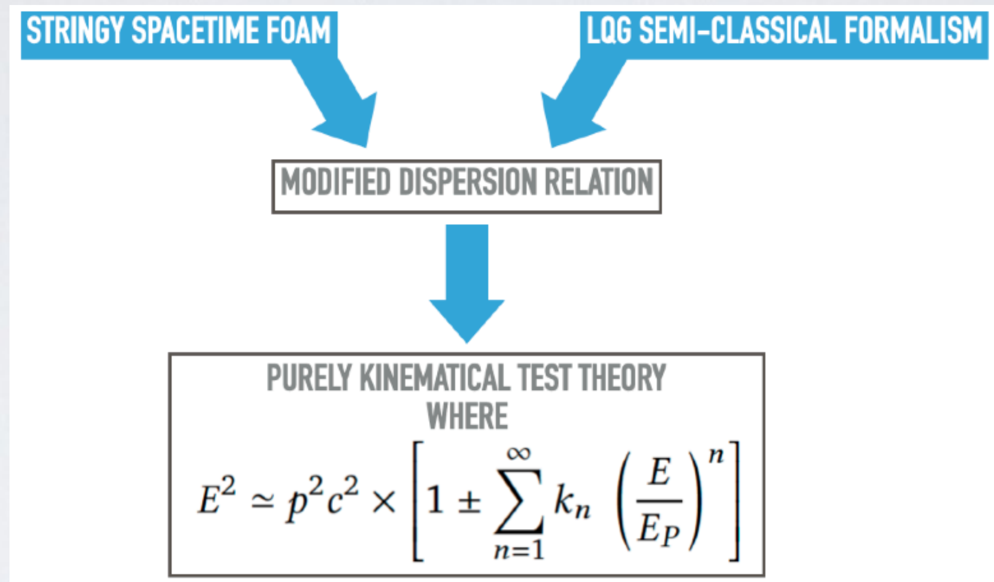


PART IV

LIV ANALYSIS

WHAT WE WANT TO TEST

- Some theoretical approaches predict a modified dispersion relation for photons in vacuum



E_P: Planck scale ~10^{19} GeV

- So, photons of different energies could travel at different speeds in vacuum
- Two photons emitted at the same time at redshift z would arrive with a delay

$$\Delta t_n \simeq s_\pm \frac{n+1}{2} \frac{E_h^n - E_l^n}{E_{QG}^n} \int_0^z \frac{(1+z')^n}{H(z')} dz'$$

where the integral deals with Universe expansion

- This effect is very small but cumulate on large distances
  - Example values for $E_{QG} = E_P$
- When no significant lag is measured, a limit is set on E_{QG}

	z = 0.1	z = 1
ΔE = 100 GeV	0.3 s 3x10^{-18} s	4 s 5x10^{-17} s
ΔE = 1 TeV	3 s 3x10^{-16} s	40 s 5x10^{-15} s

Δt_n for n = 1, n = 2

HOW

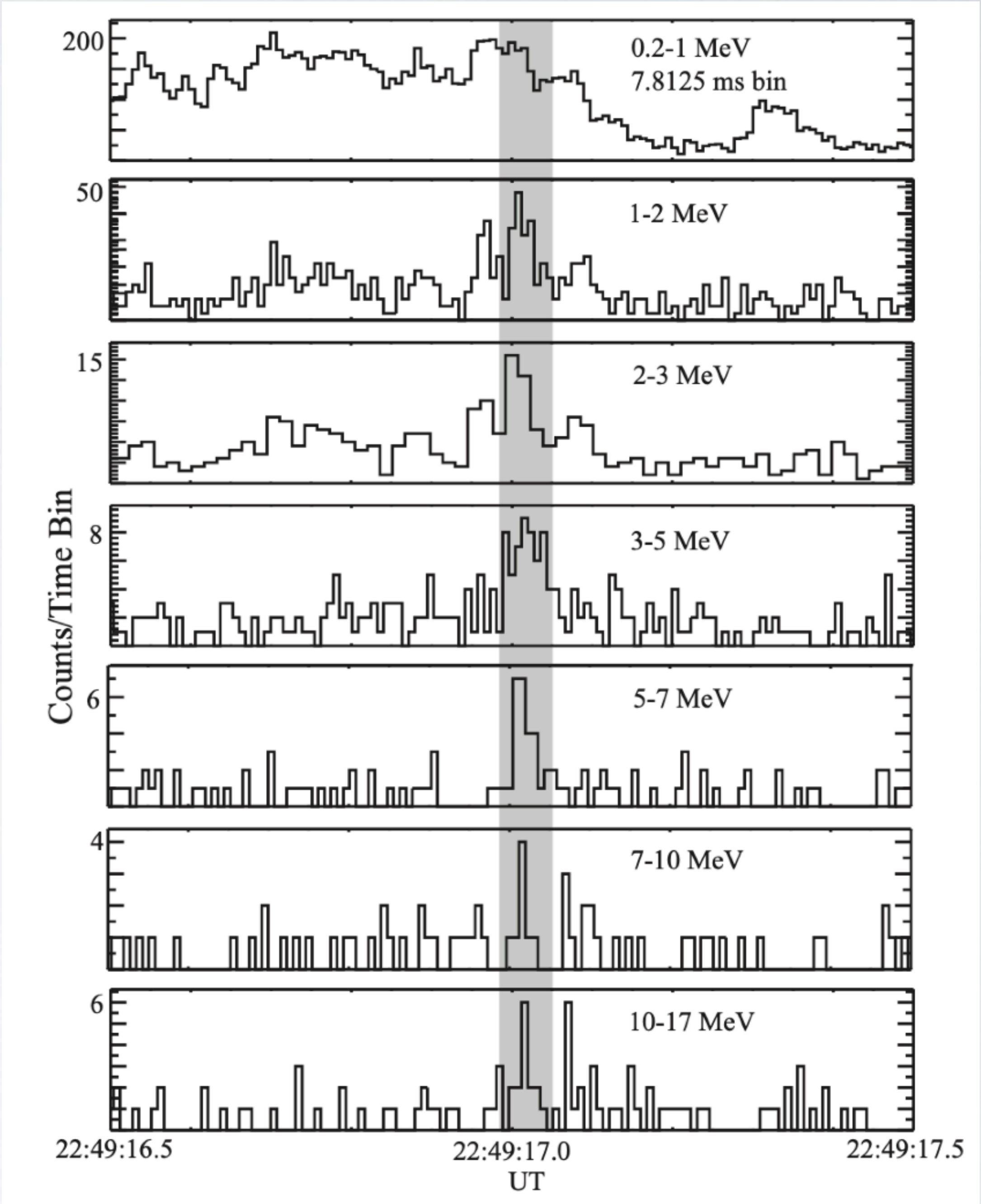
- So, we want to measure a time lag between photons of different energies
- Need for sources that are
  - Distant
  - Variable or transient
  - Energetic
- Candidates:
  - Gamma-Ray Bursts (GRBs)
  - Flaring Active Galactic Nuclei (AGNs)
  - Pulsars (PSRs)
- These sources all have advantages and drawbacks
- The « Variable or transient » criteria is essential
  - The more variability we have, the more sensitive we'll be
  - A galactic millisec. PSR can give more stringent limits than a distant flaring AGN…
- Several techniques have been developed to measure energy-dependent time delays

Source	Observed by	Distance	Variability time scale
AGN flare	IACT	z_{max} ~ 0.6	O(1 min)
GRB	Satellites	z_{max} ~ 9	O(0.1 s)
PSR	Satellites & IACTs	d_{max} ~ 50 kpc	O(1 ms)

Source	Random ?	Intrinsic effects	Emission mechanisms
AGN flare	Yes	Hints	poorly understood
GRB	Yes	Yes	poorly understood
PSR	No	Can be separated	~Ok

BASIC TECHNIQUES

- The main ingredients of LIV analyses are:
  - light curves in different energy bands, or
  - a list of photons (gamma-like) with arrival times, reconstructed true energies, etc. together with IRFs
    - As seen previously, each type of analysis, cuts, etc. will lead to different IRFs and different lists of photons
- A simple technique:
  - Fit the peaks in the light curves in different energy bands
  - Compare the fitted positions of the peaks to get Δt and the corresponding uncertainty
  - From the spectral index and energy ranges, get the value of ΔE
- Other methods were used based on signal processing techniques: CCF, wavelets…



Boggs et al. 2004 - GRB 021206 - RHESSI
7.8 ms binning

A MORE ADVANCED TECHNIQUE

- Likelihood technique proposed by Martinez & Errando (2008)
- The pdf is computed from the time and energy distributions and takes IRFs into account (effective area, energy resolution)

$$P(E_i, t_i|\tau_n) = N \int_0^\infty D(E_i, E_S)\, \Lambda(E_S)\, F(t_i - \tau_n E_S^n)\, dE_S$$

- The likelihood is then obtained from the events with

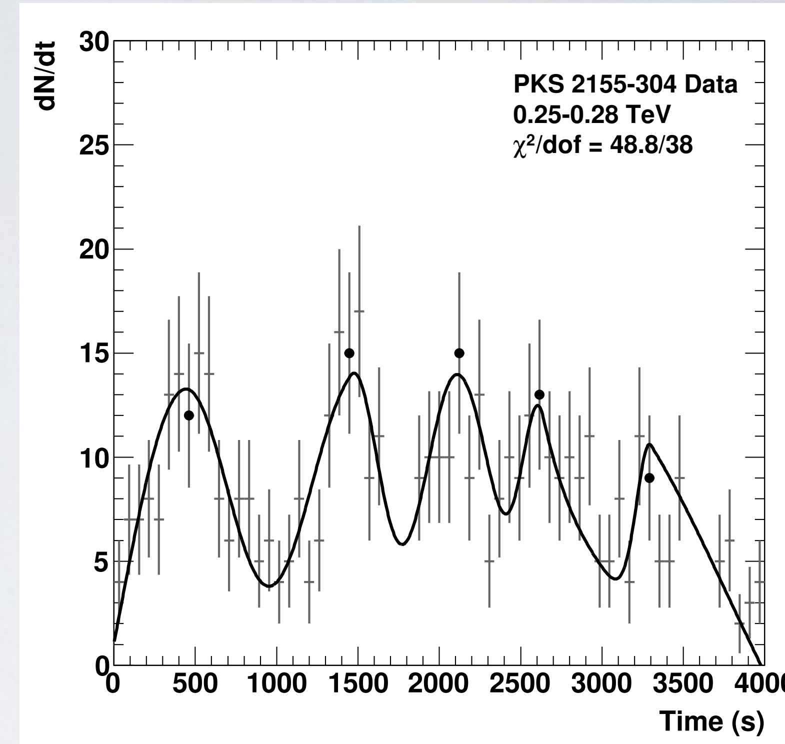
$$\mathcal{L} = \prod_i P(E_i, t_i|\tau_n)$$

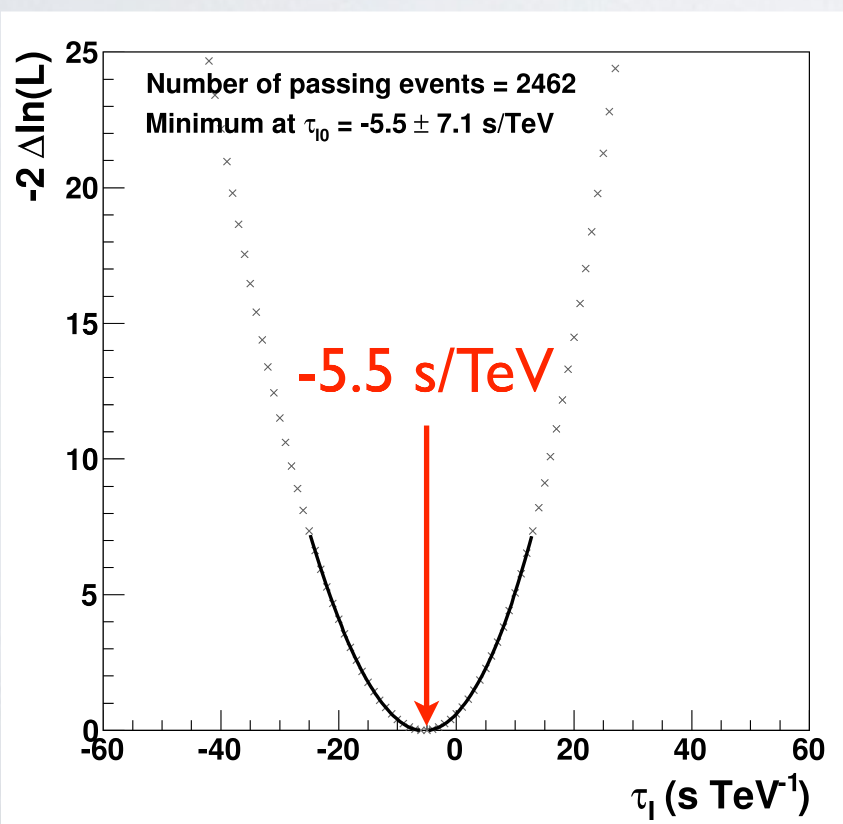
- Maximizing the likelihood gives the best estimate for parameter τ

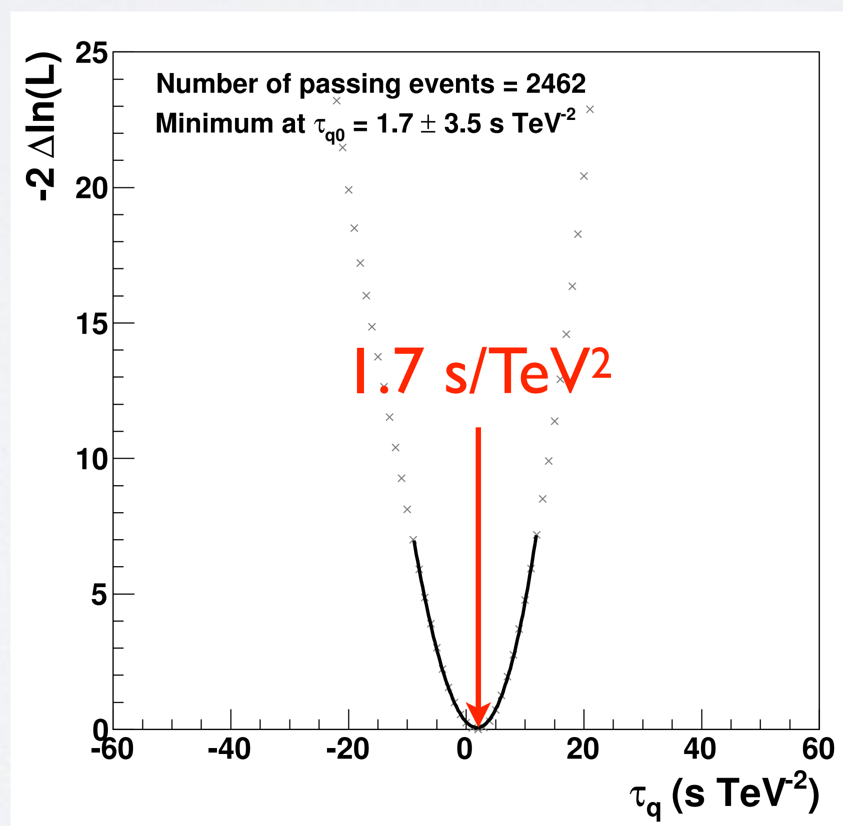
- Main advantages:
  - Good sensitivity
  - Easy to combine multiple sources

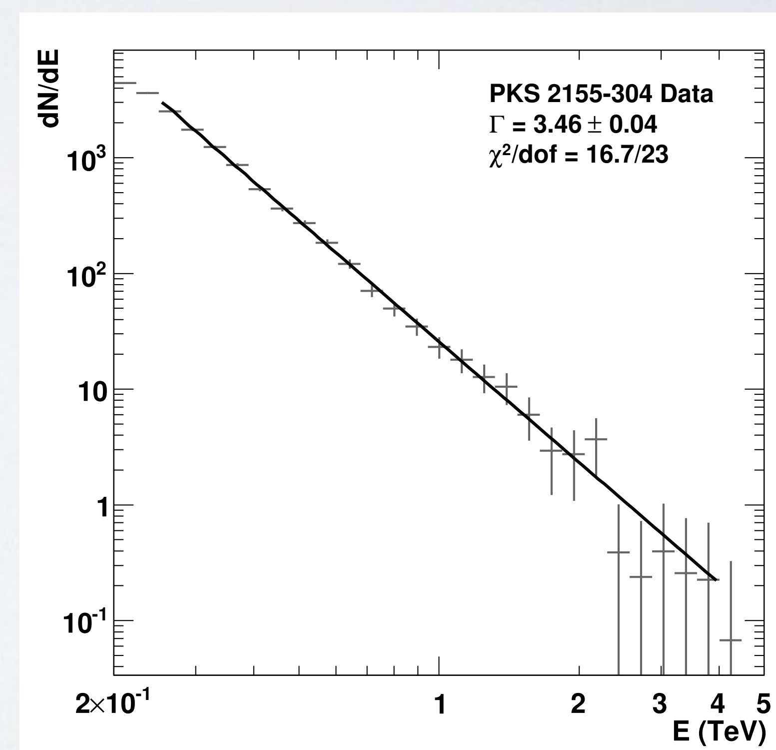
A MORE ADVANCED TECHNIQUE

- An example with PKS2155-304 2006 flare:
  - Parameterization of the light curve at low energies
  - Parameterization of the spectrum
  - Likelihood maximization and obtained best estimates



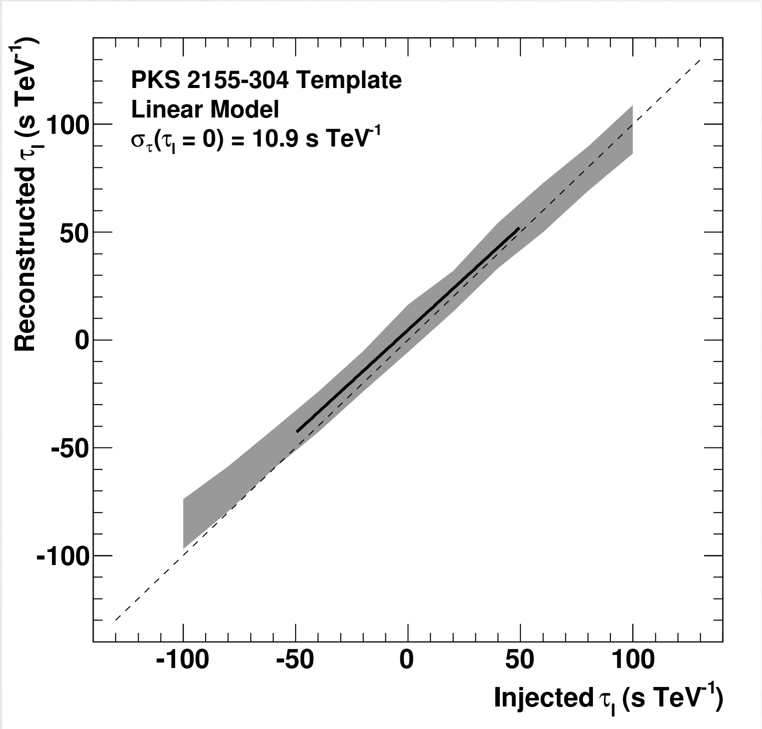






IMPORTANT CONSIDERATIONS

- **Key points of analyses are**
  - **Performance assessment**
  - **Statistical and systematic error assessment**
- Both of these are usually done using dedicated toy-MC simulations
  - Simulate hundreds of light curves with same properties as real data
  - Reconstruct the lag for all of them and study its distribution → statistical error estimate
- Systematic uncertainties have also to be careful assessed.



	Change in estimated τ_l (s/TeV)	Change in estimated τ_q (s/TeV^2)
Selection cuts	< 5	
Background contribution	< 1	
Acceptance factors	< 1	
Energy resolution	< 1	
Energy calibration	< 2	
Spectral index	< 1	
Calibration systematics	< 5	< 1
$F_S(t)$ parameterization	≈ 7	≈3
Total	< 10.3	< 6.6

FINAL REMARKS

FINAL REMARKS

- Hardware, reconstruction, analysis is a huge topic to cover
- With only 4 hours available, a lot had to be ignored
  - Physics of the showers (we'll cover that a bit with the tutorial)
  - New generation instruments (sorry CTA, SiPMs, etc. !)
  - New techniques in reconstruction and analysis (neural networks, etc.)
  - Acceptance determination
  - Other background rejection techniques
  - Details on LIV analyses → will be covered during the 3rd training school!
  - …
- Lecture notes will cover that a just a little bit more, including references, and a glossary of most important terms

WE MADE IT!

THANKS!