# Applications of machine learning to the event reconstruction of Imaging Atmospheric Cherenkov Telescopes

Ilaria Viale, UniPd, ilaria.viale@unipd.it

1st AI INFN Advanced Hackathon, 26 - 28 Nov 2024

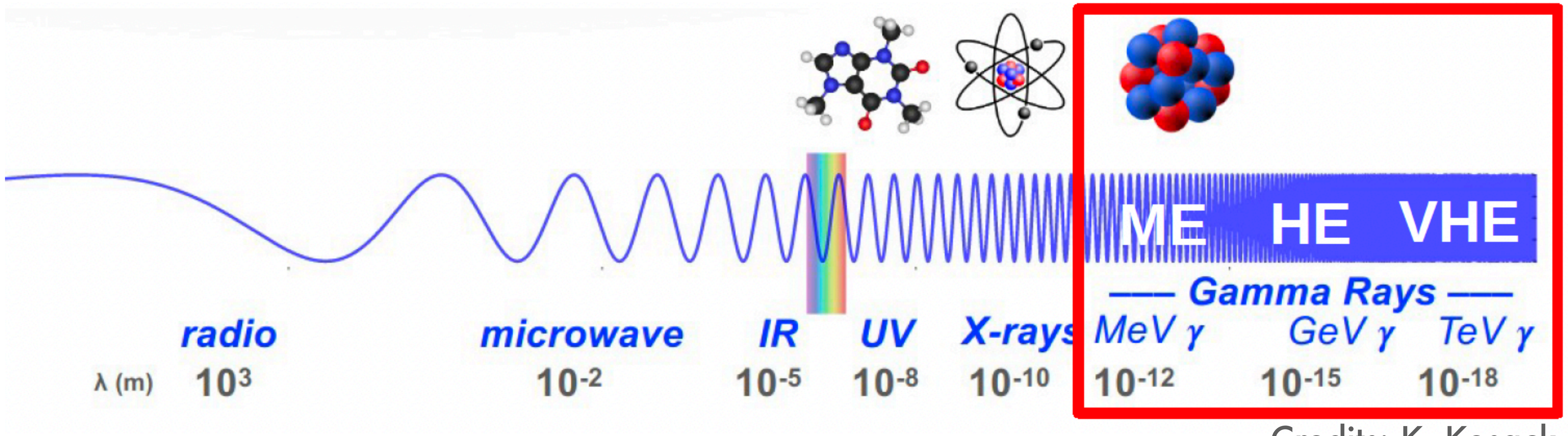
→ Very High Energy astrophysics

&

→ Imaging Atmospheric Cherenkov Telescopes

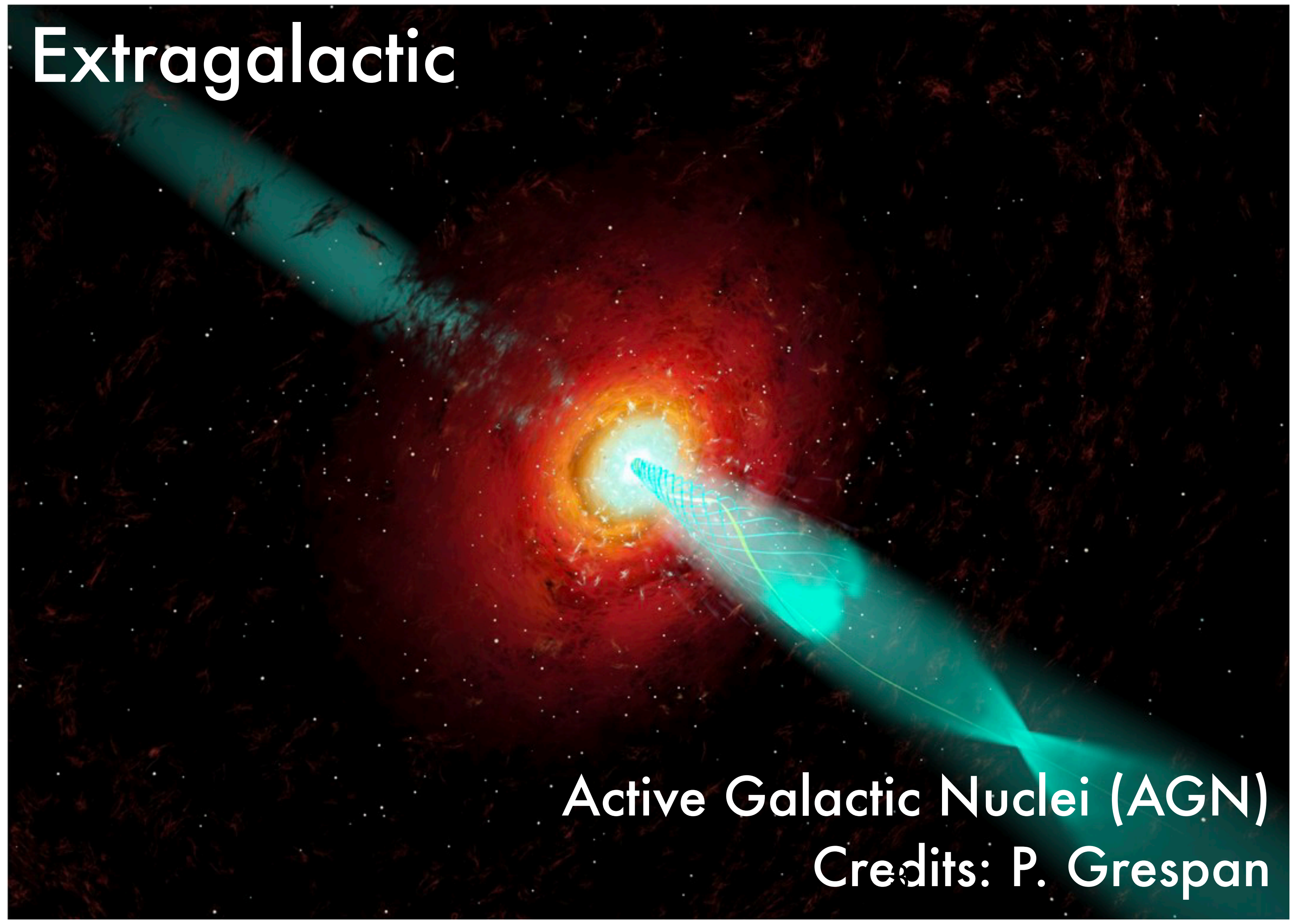
# VHE γ-ray astrophysics

- At the most energetic extreme of the EM spectrum
- γ-rays due to non-thermal emission of accelerated particles
- Study of galactic and extragalactic cosmic accelerators




Credits: K. Kosack


Supernova remnant
Credit: NASA, ESA




Active Galactic Nuclei (AGN)
Credits: P. Grespan


Gamma-ray burst
Credit: ESO/A. Roquette

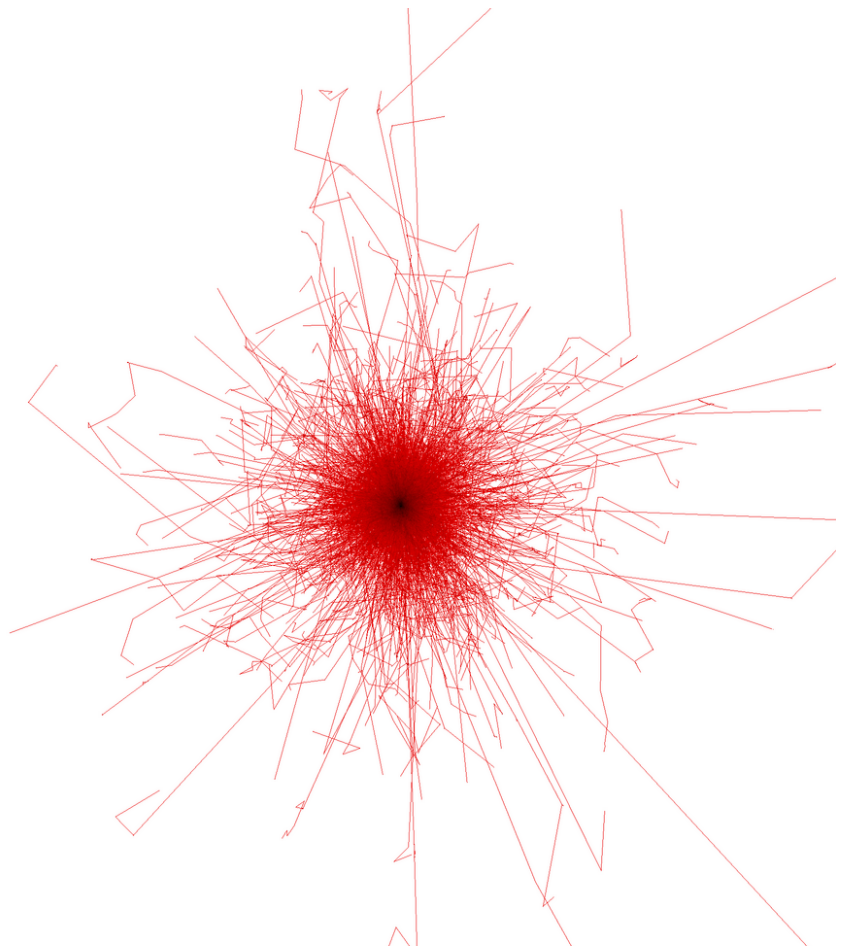
# VHE γ-ray astrophysics

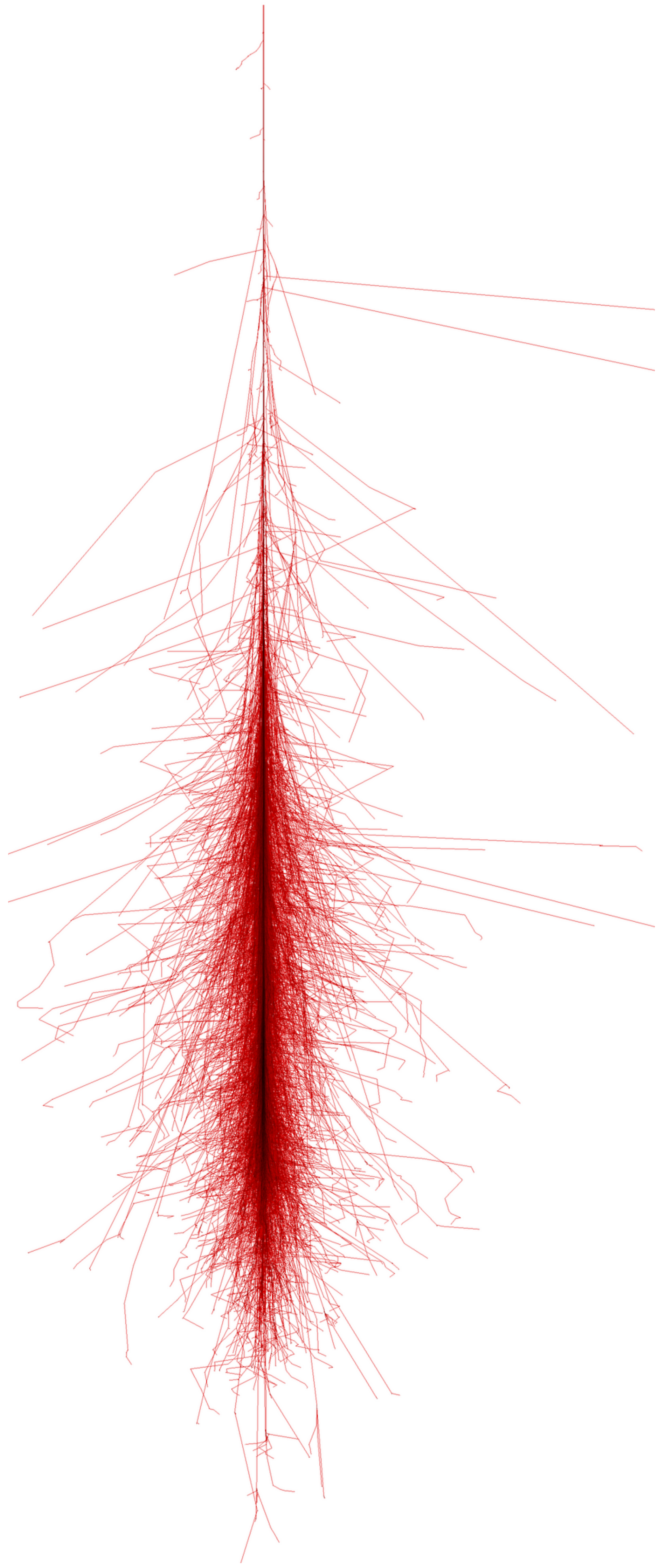
- Effective and high-performing analysis of γ-ray data needed for different science cases:
  - Distant sources
  - Specific AGN types
    - e.g. Flat Spectrum Radio Quasars, Extreme blazars
  - Fundamental physics research
    - e.g. Dark matter, axion like particles, intergalactic magnetic fields

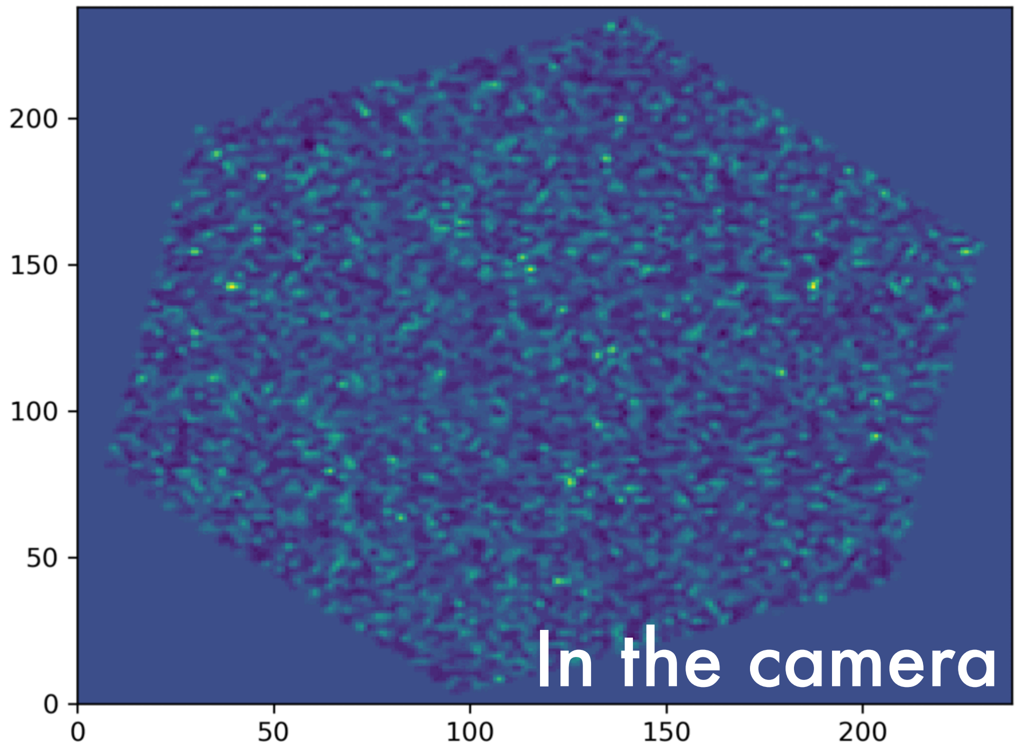

- VHE observations exploit creation of extensive air showers initiated by the incoming gamma-ray

# Extensive Air Showers

## Electromagnetic showers

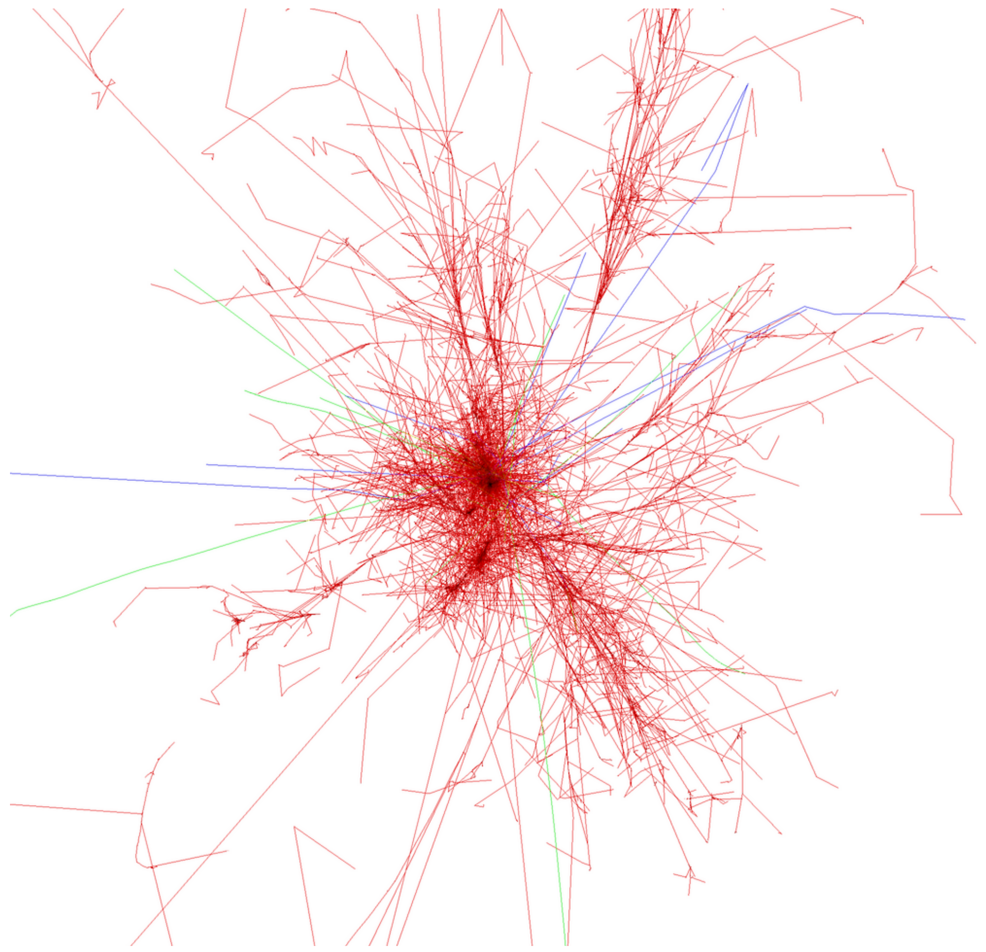


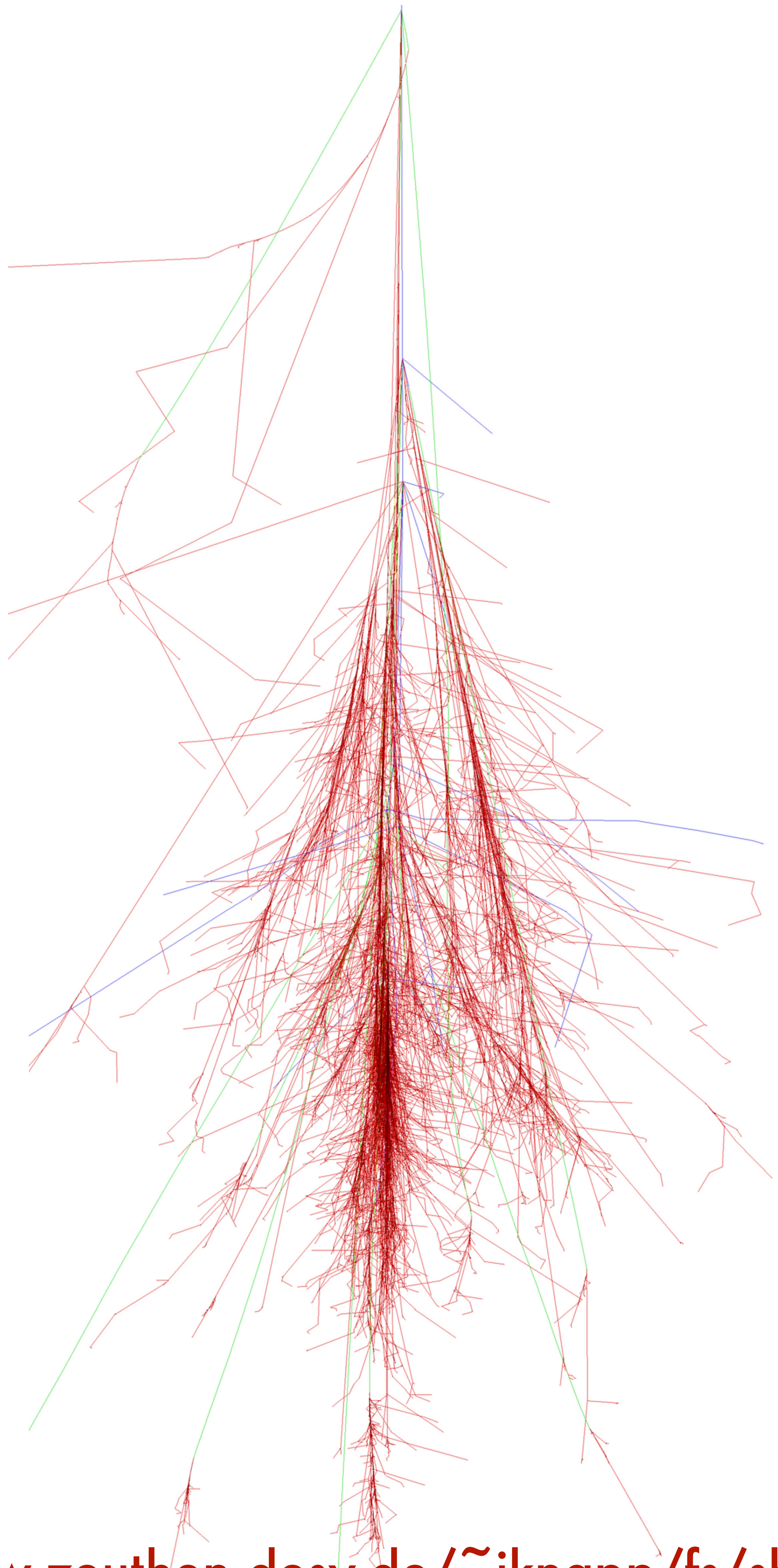


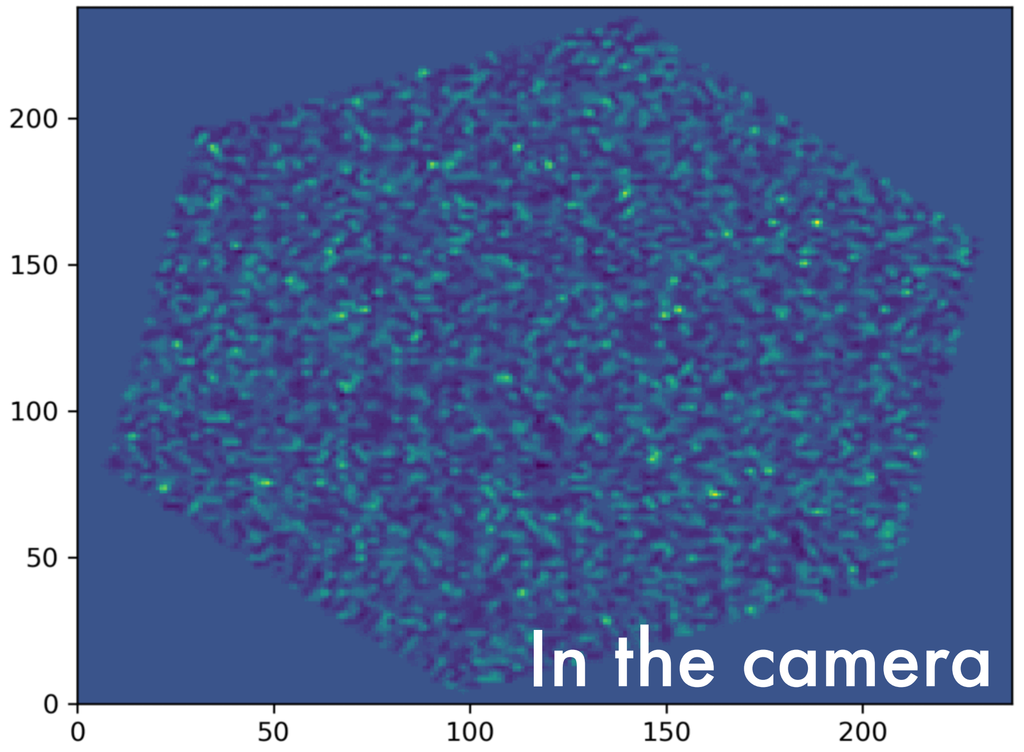



Credits: T. Miener

## Hadronic showers





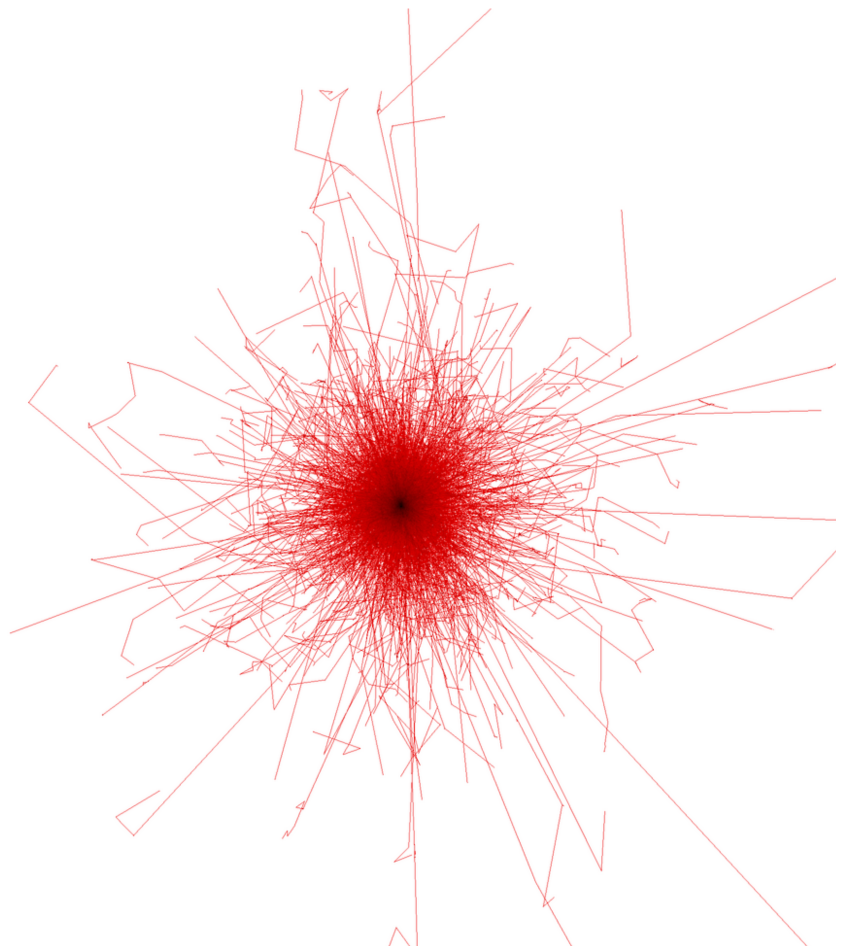


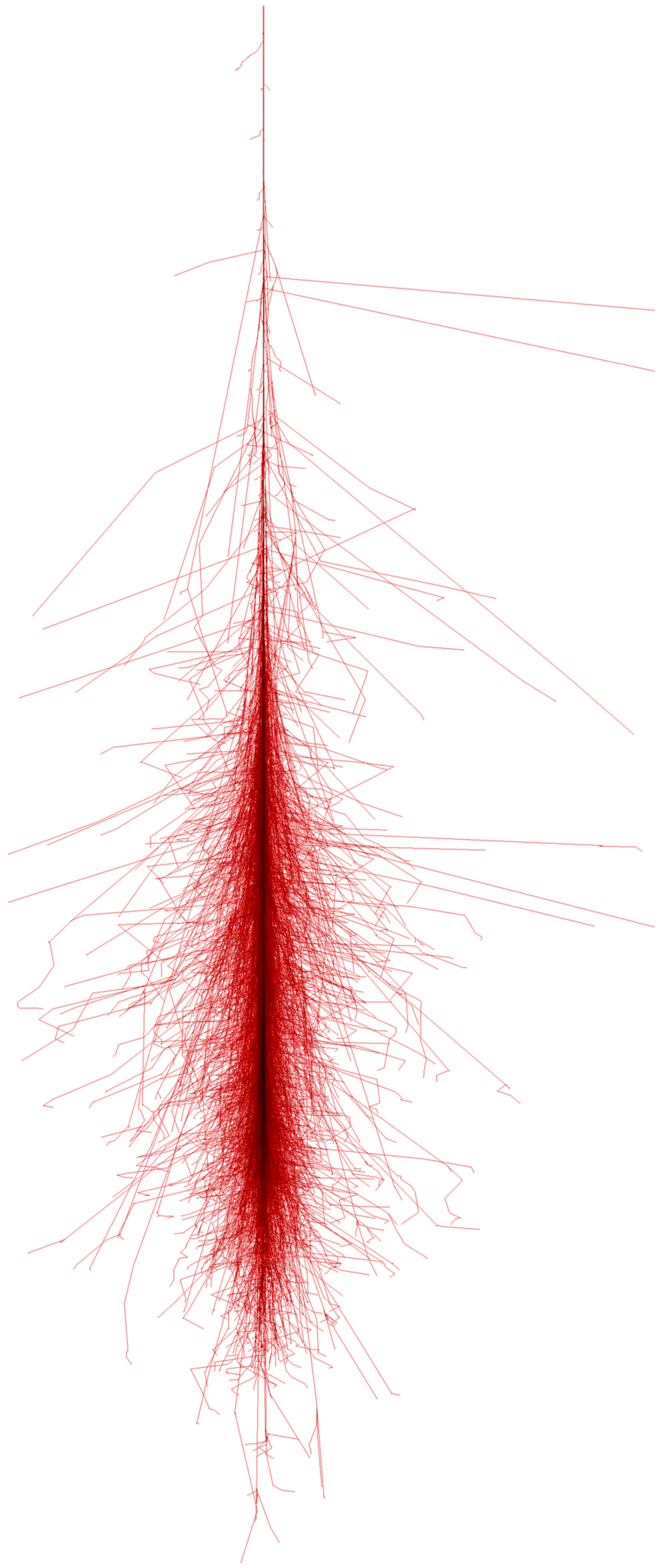

Credits: T. Miener

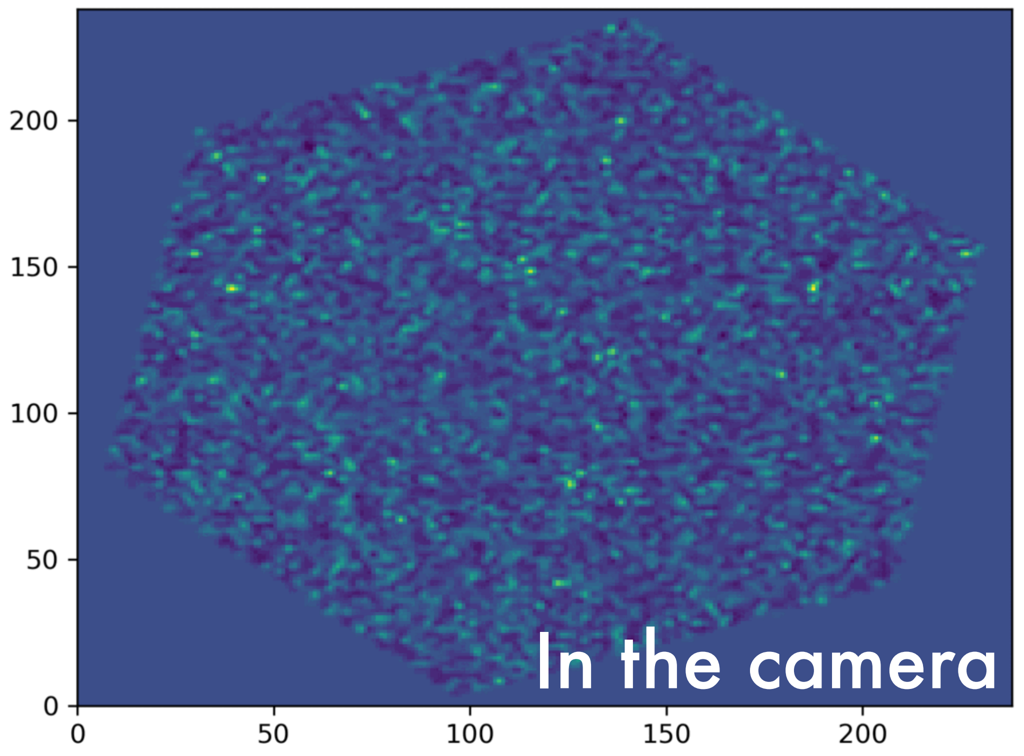
https://www-zeuthen.desy.de/~jknapp/fs/showerimages.html

# Extensive Air Showers

## Electromagnetic showers

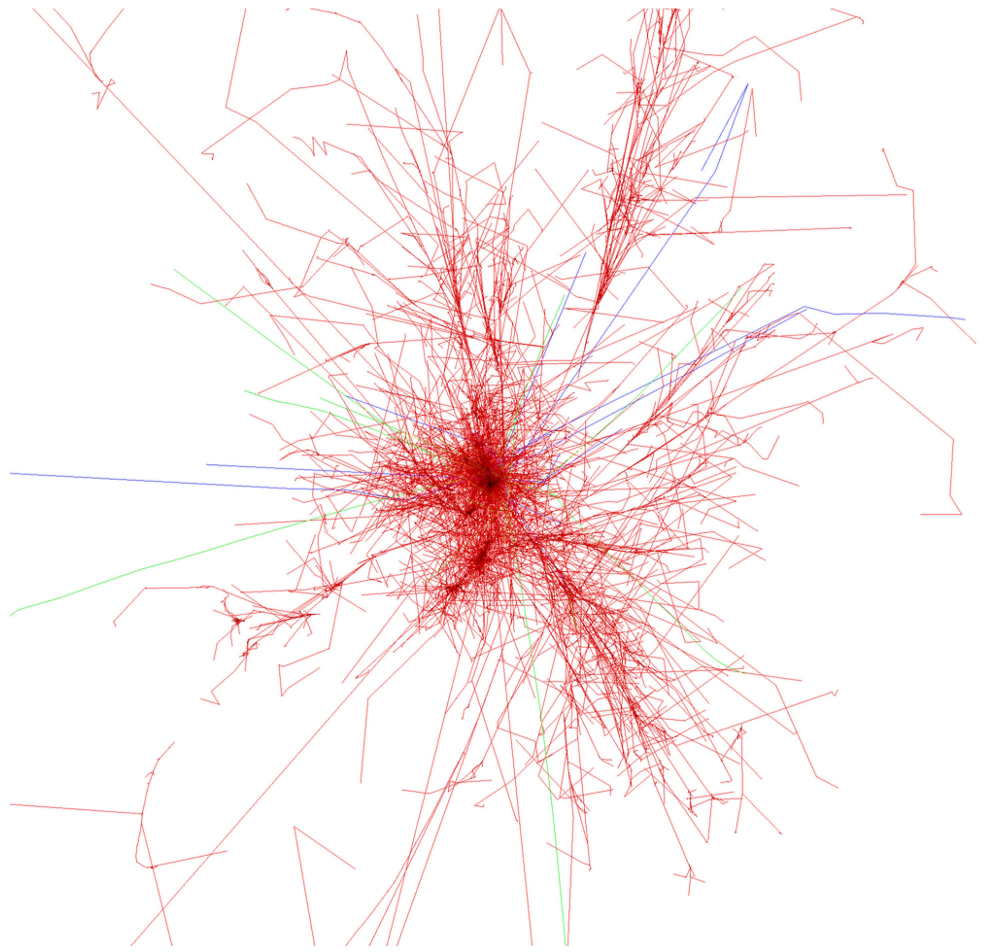


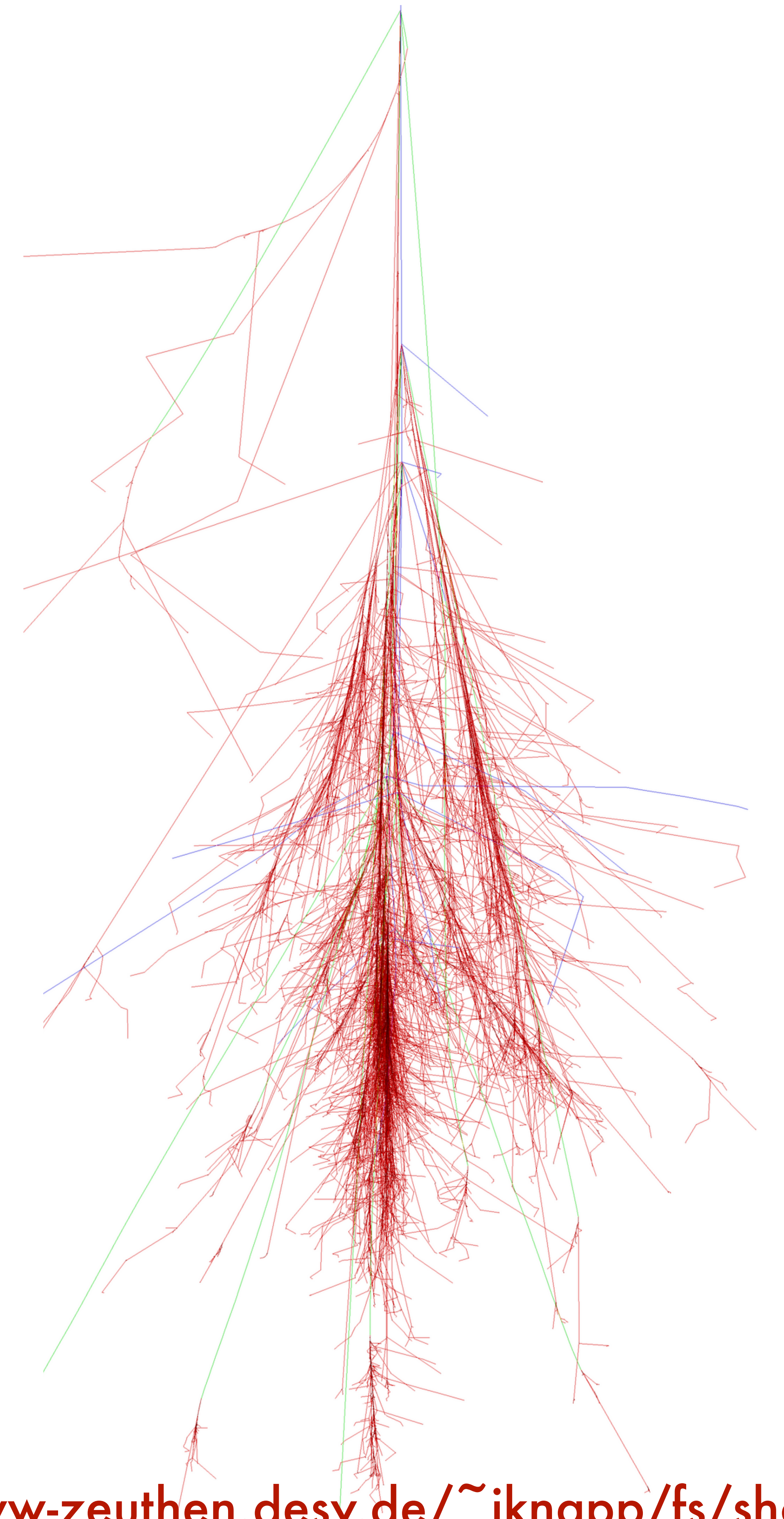


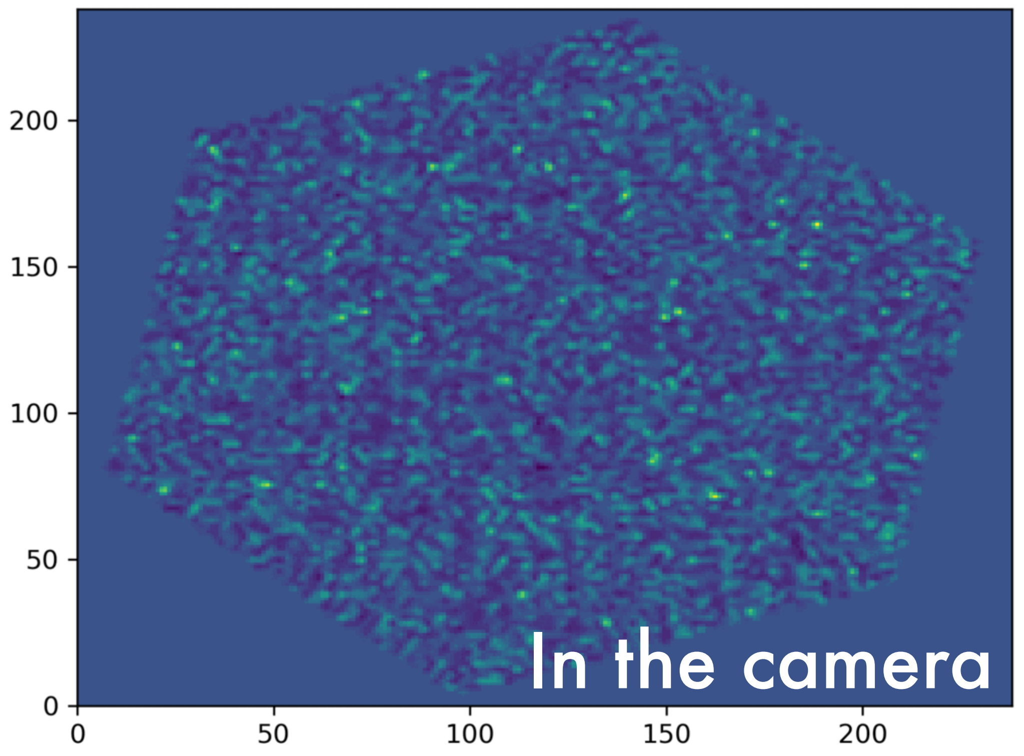



Credits: T. Miener

## Hadronic showers







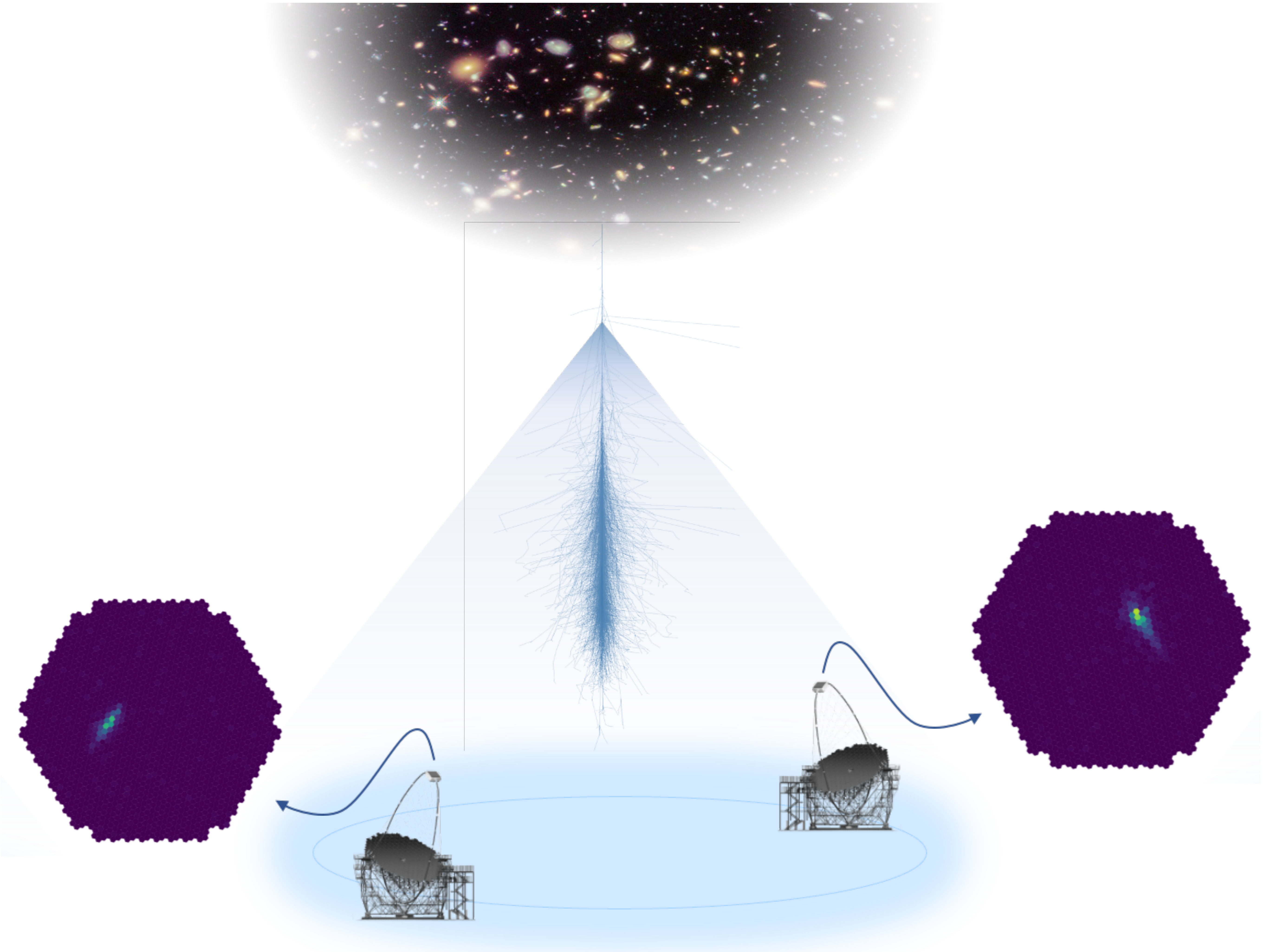

Credits: T. Miener

https://www-zeuthen.desy.de/~jknapp/fs/showerimages.html

# Imaging atmospheric Cherenkov technique

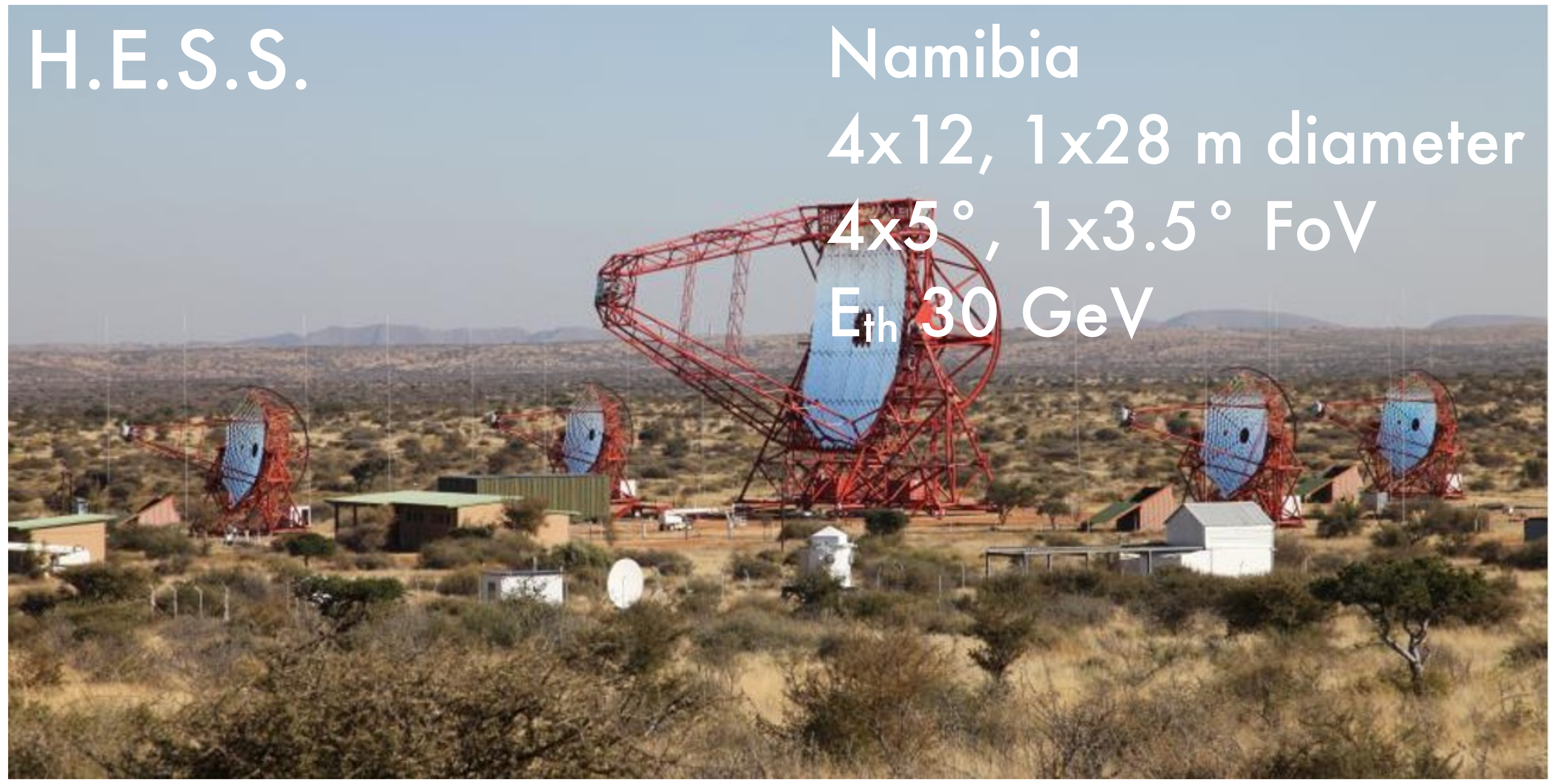
- Air showers initiated by VHE γ-rays
- Detection of Cherenkov light produced by secondary particles
- Image recorded by PMT camera

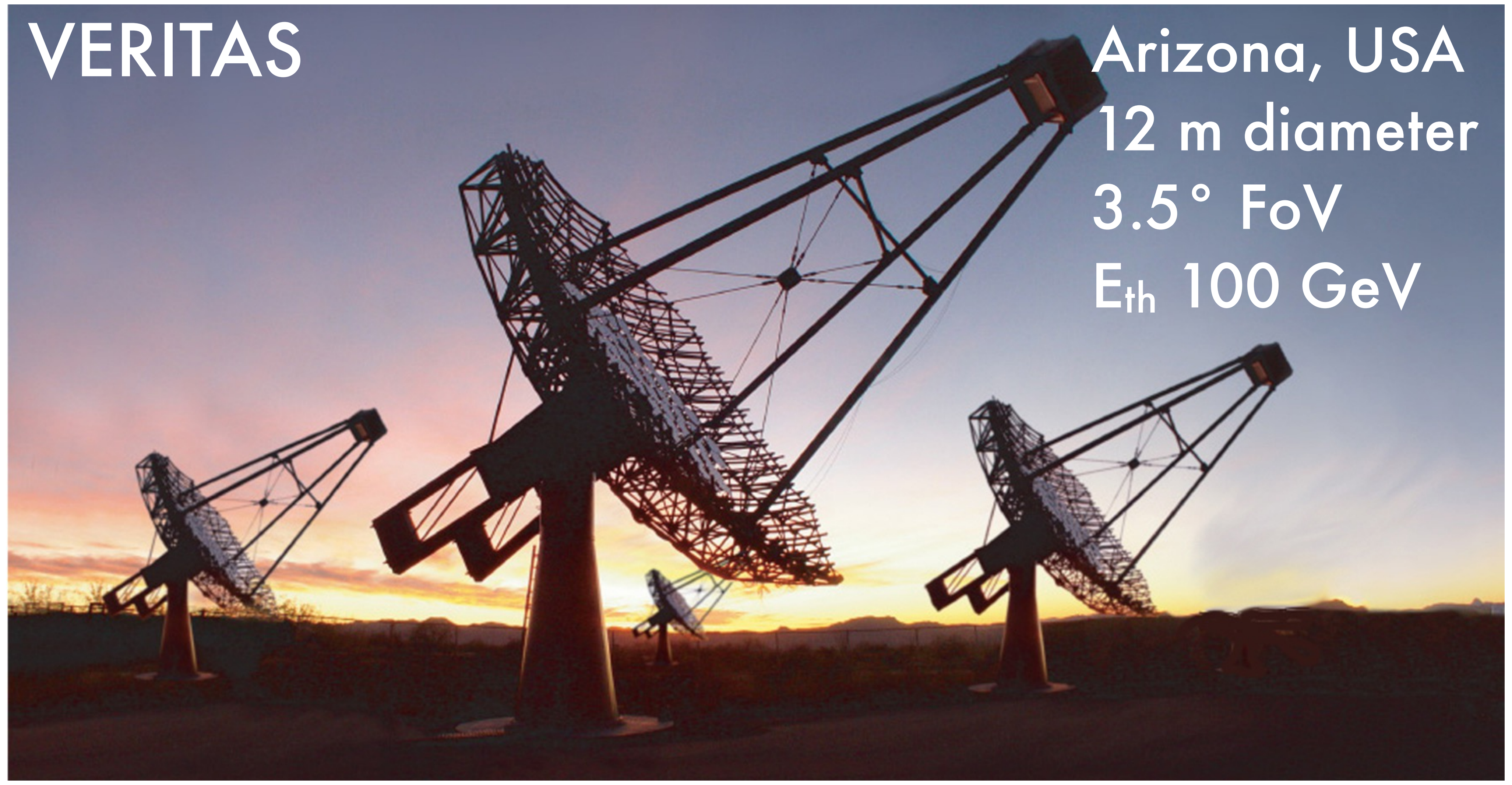
- Use of atmosphere as a calorimeter
  - Energy of primary particle deposited in the form of cascades of secondary particles
- More telescopes improve the reconstruction

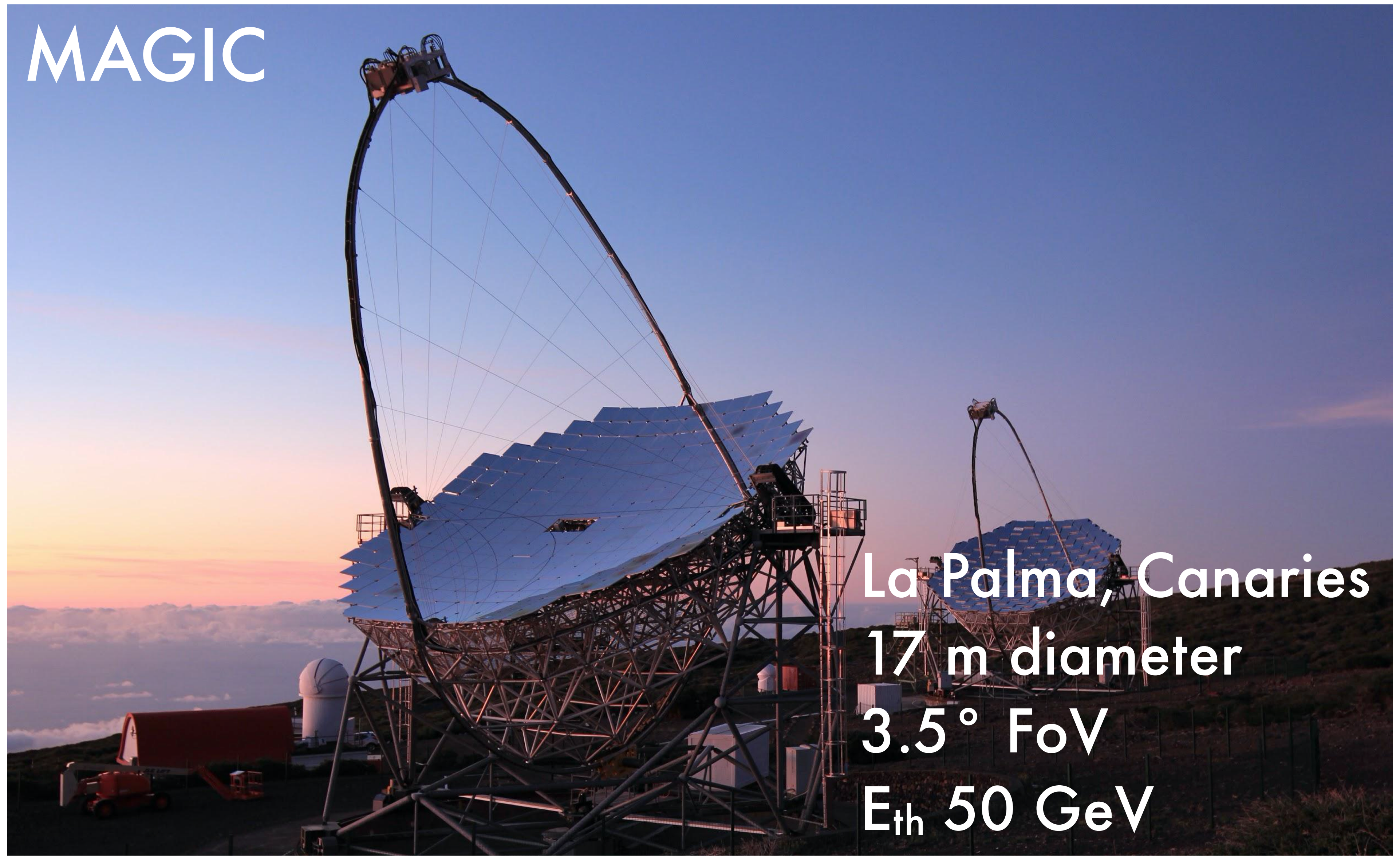


Credits: L. Romanato

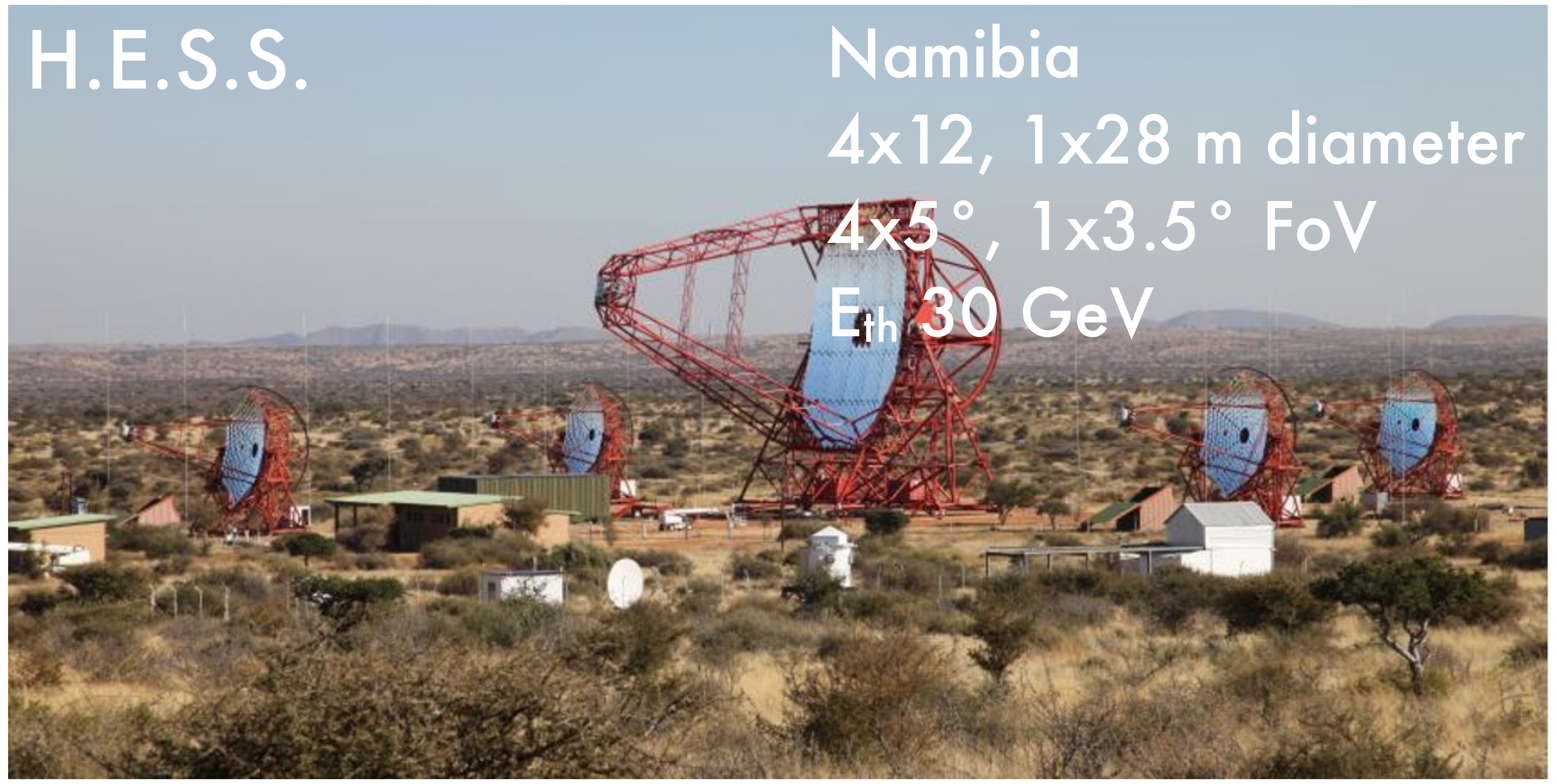
# Current generation of IACTs

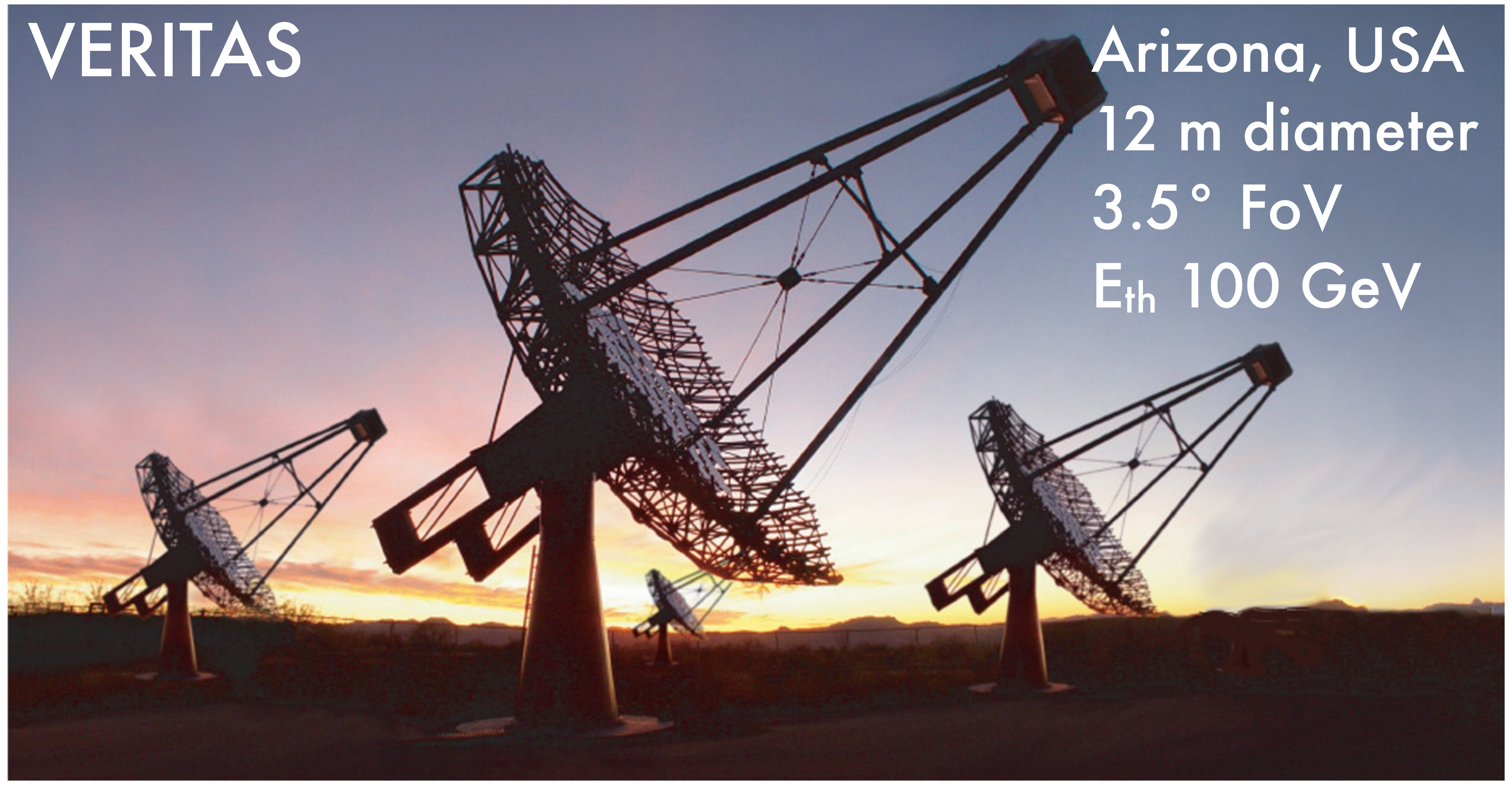


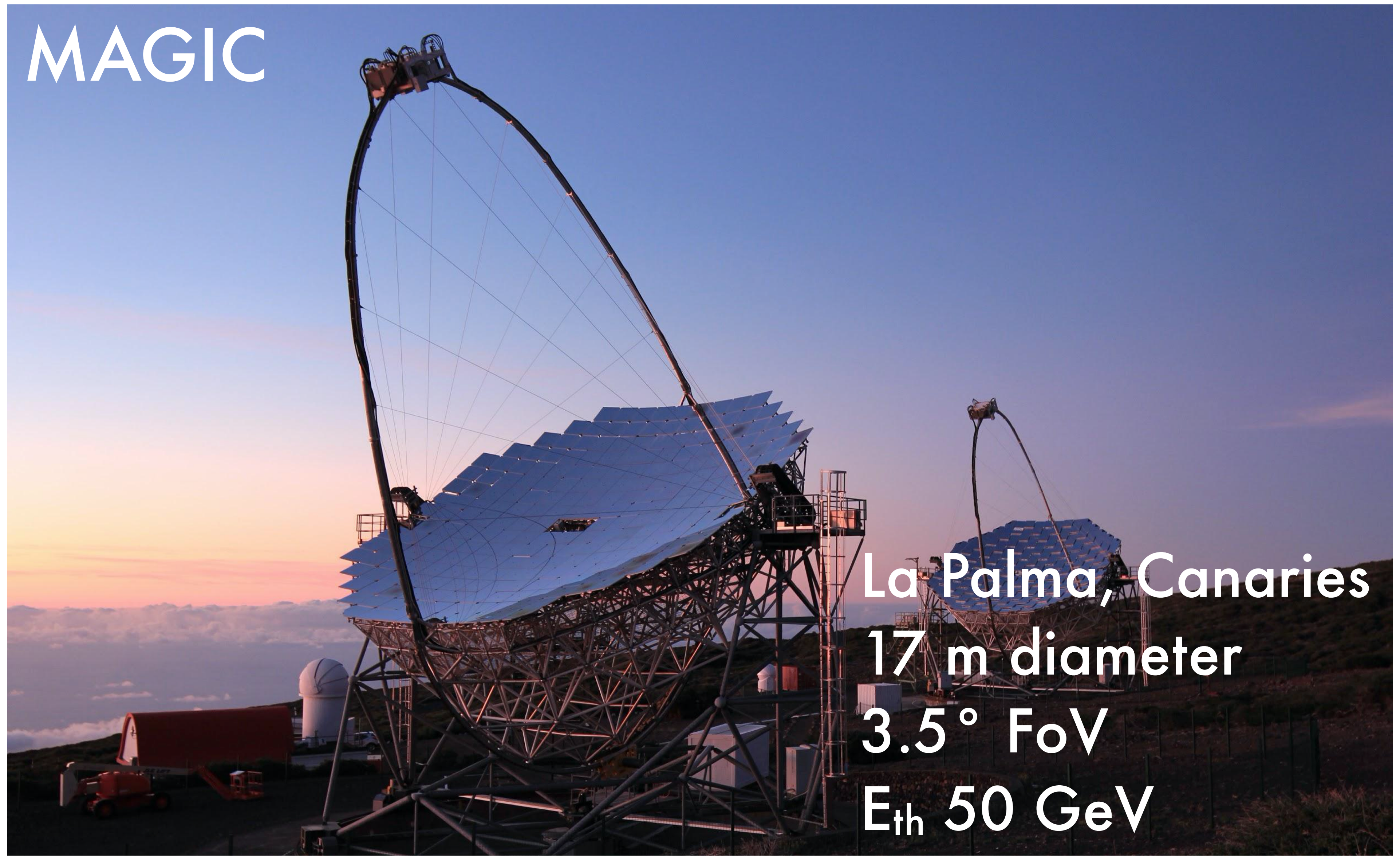



# Current generation of IACTs

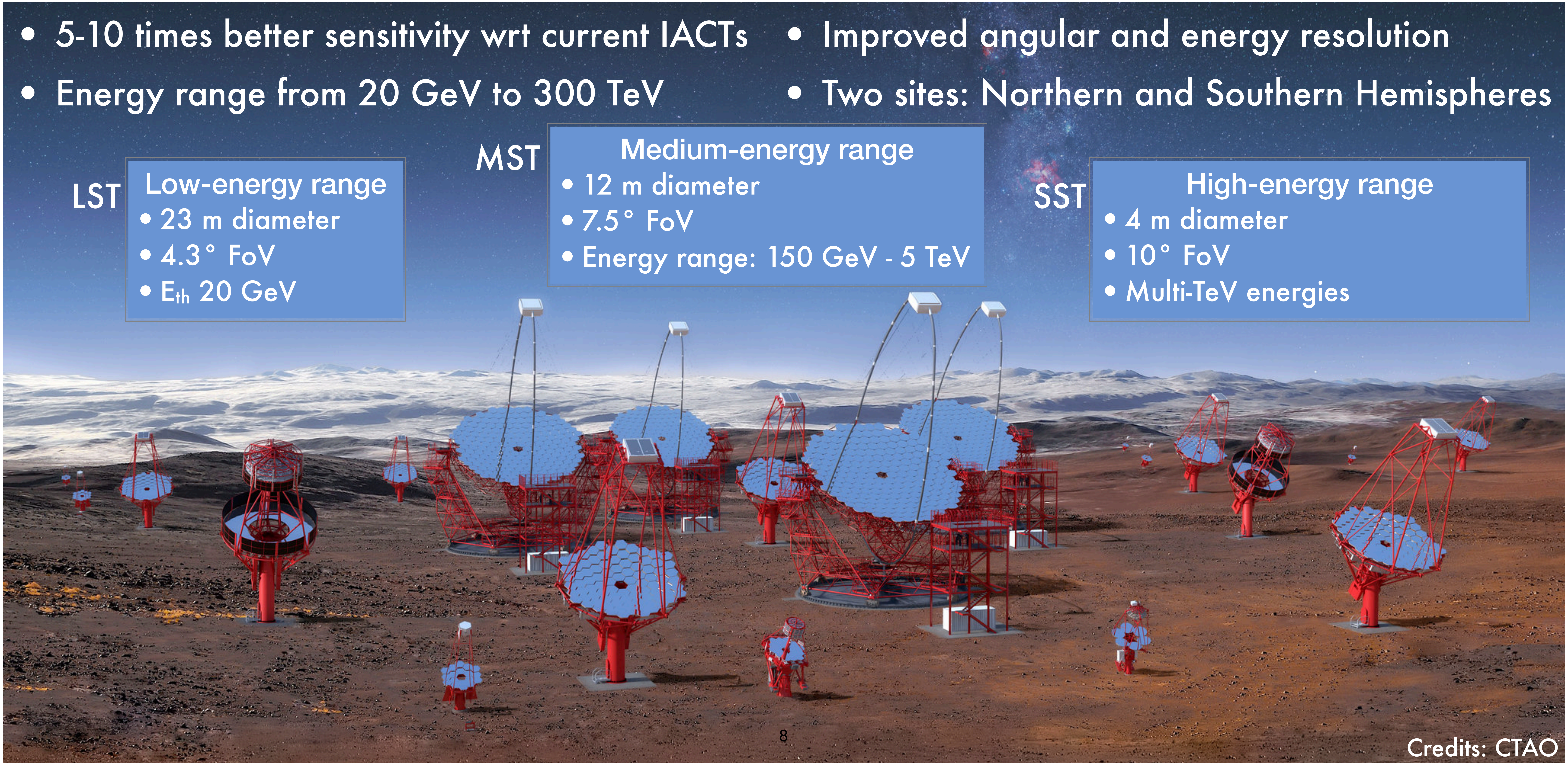



# Next generation of IACTs: CTAO

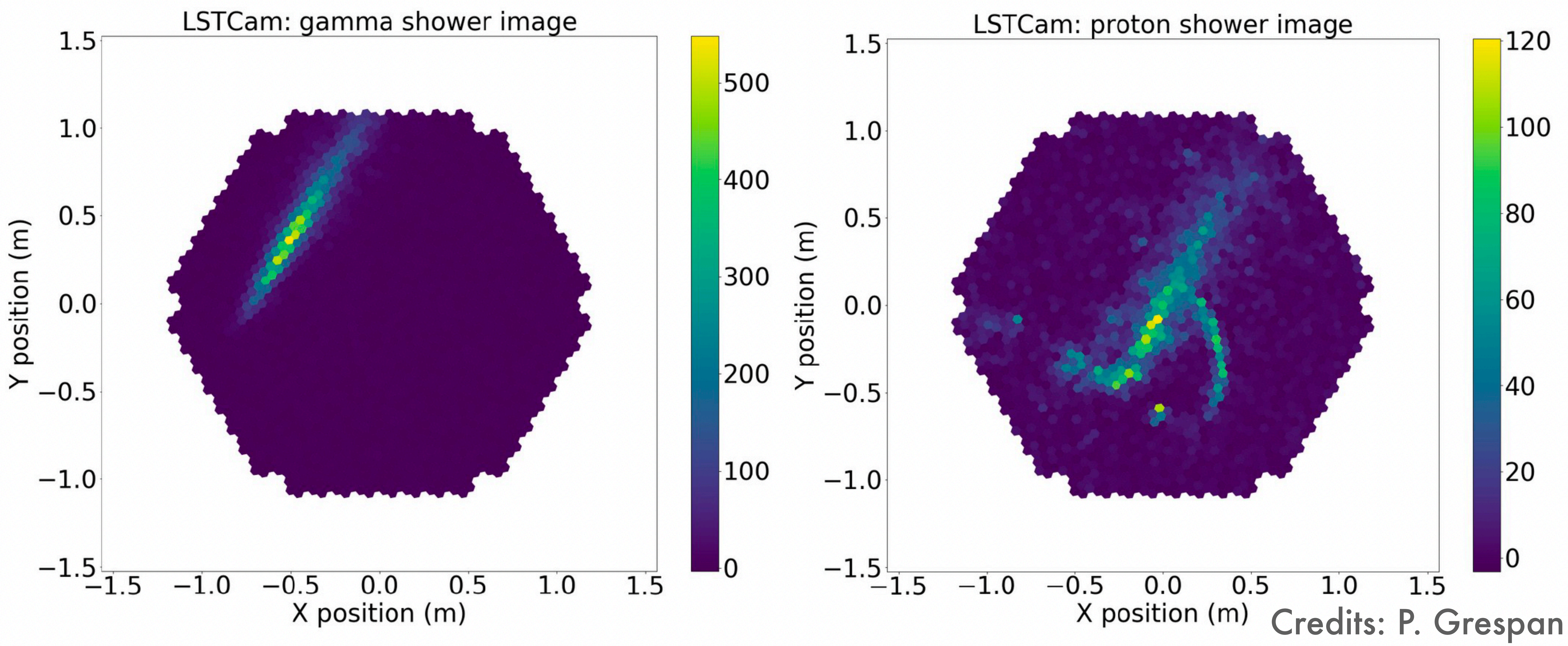
- 5-10 times better sensitivity wrt current IACTs
- Energy range from 20 GeV to 300 TeV
- Improved angular and energy resolution
- Two sites: Northern and Southern Hemispheres




Credits: CTAO

→ IACT event reconstruction

# Aims and issues




Credits: P. Grespan

Aims of event reconstruction:

- Particle type
- Energy
- Incoming direction

Main issue:

Large background from charged cosmic-rays (hadronic showers)

Ratio: 1:1000

# Standard data analysis method

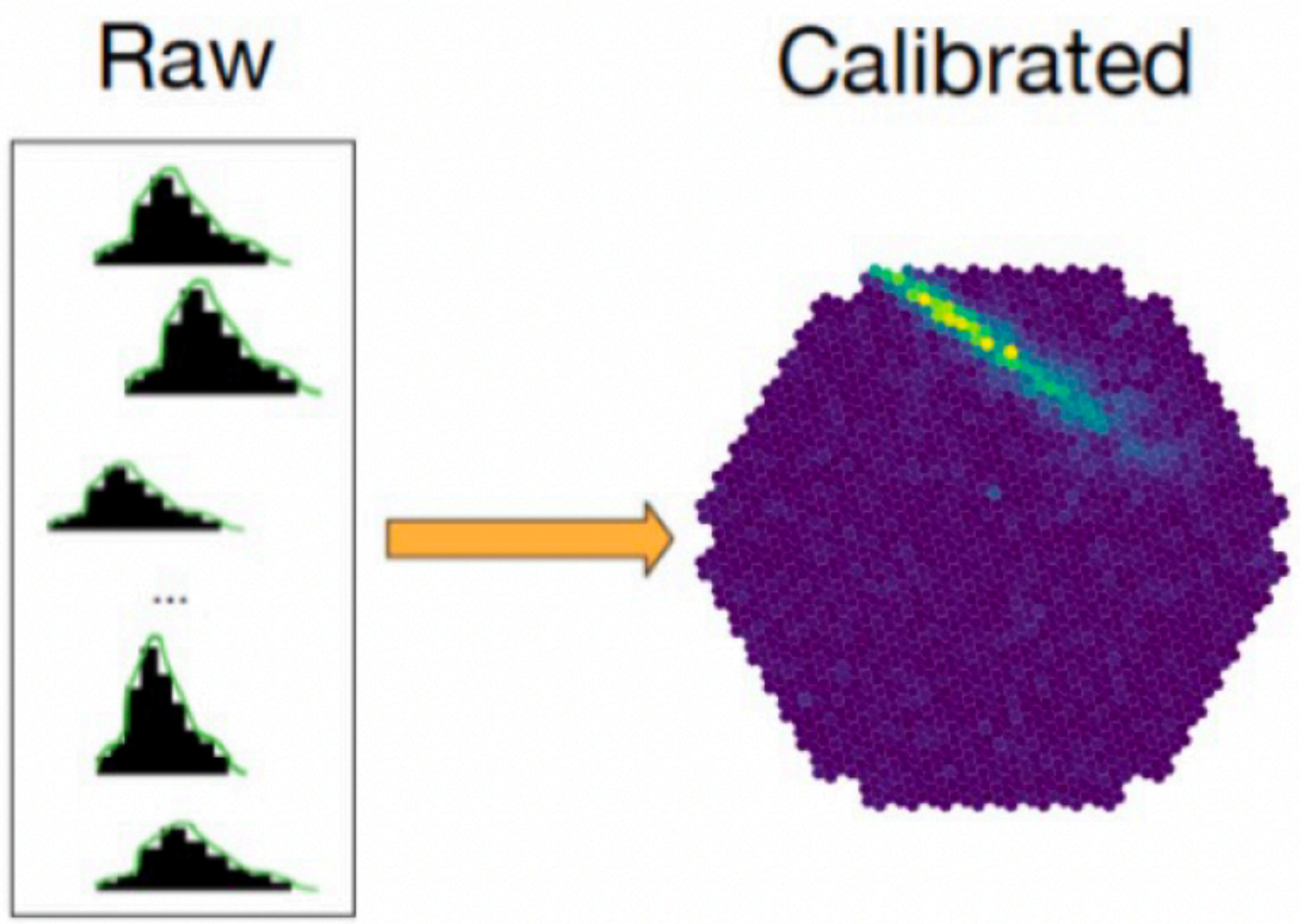





Image credits: P. Grespan

# Standard data analysis method

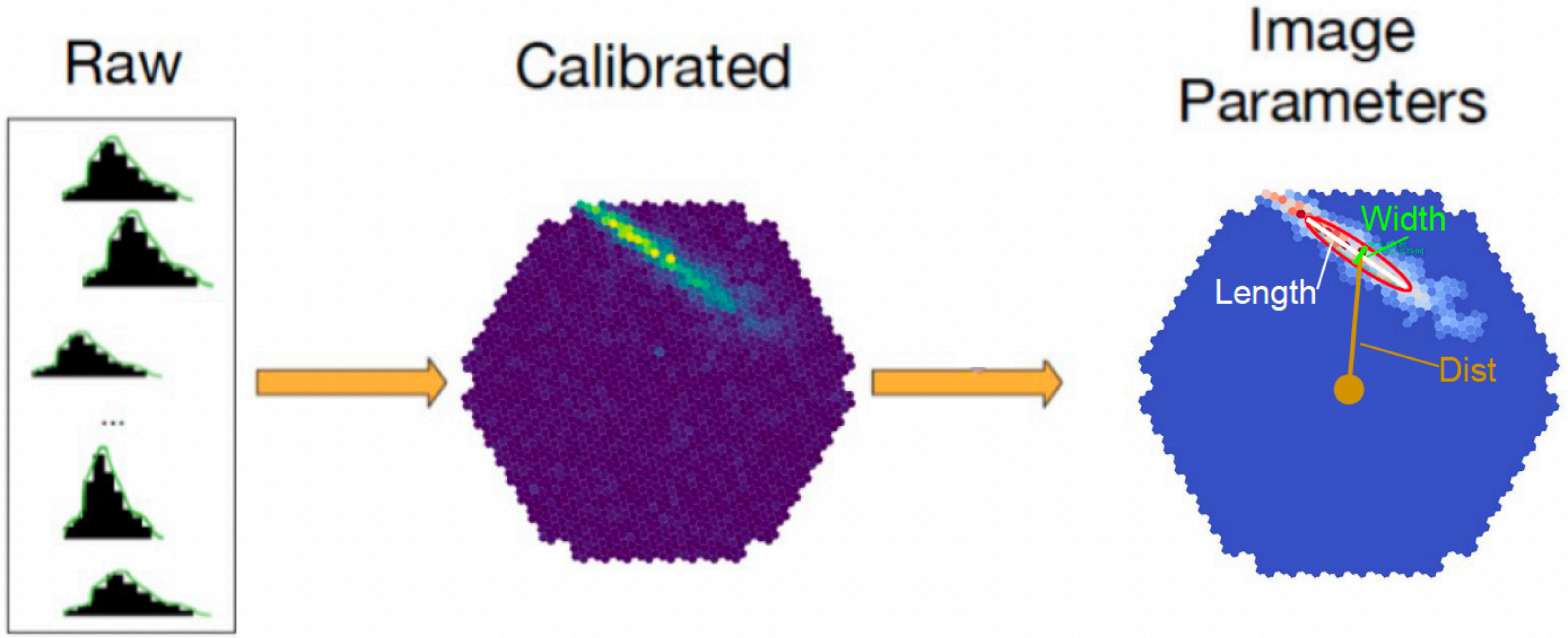


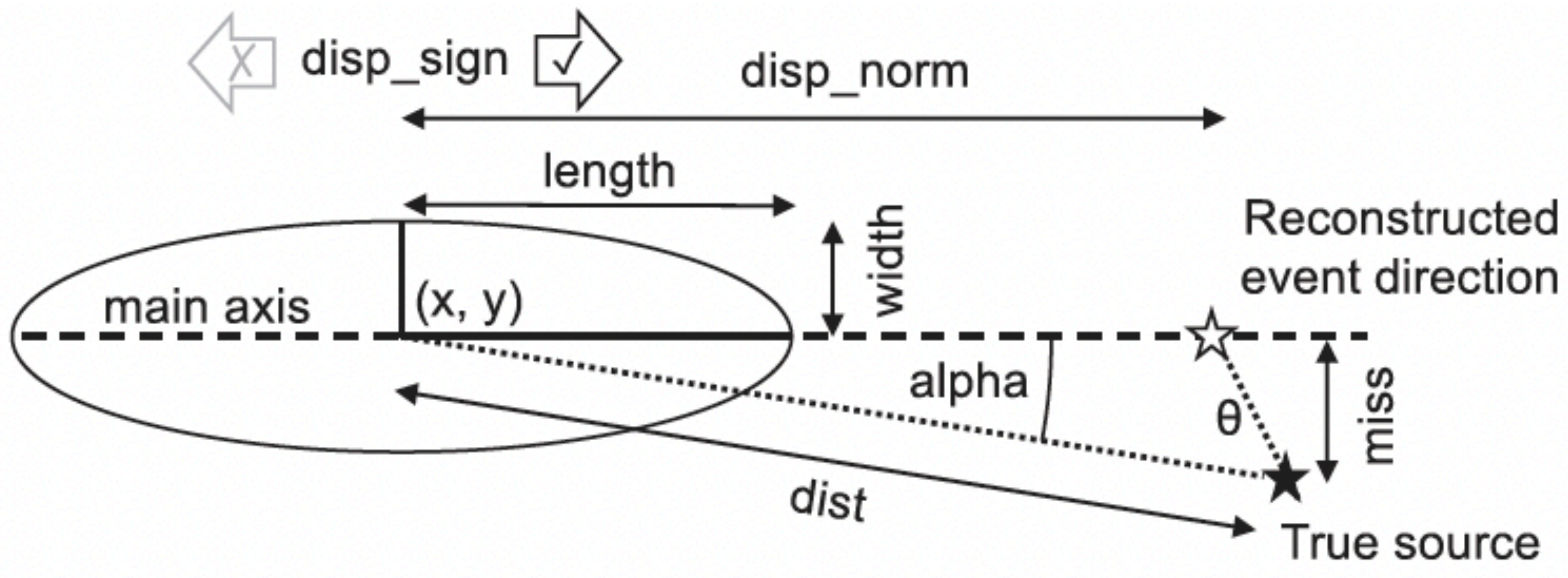



Image credits: P. Grespan

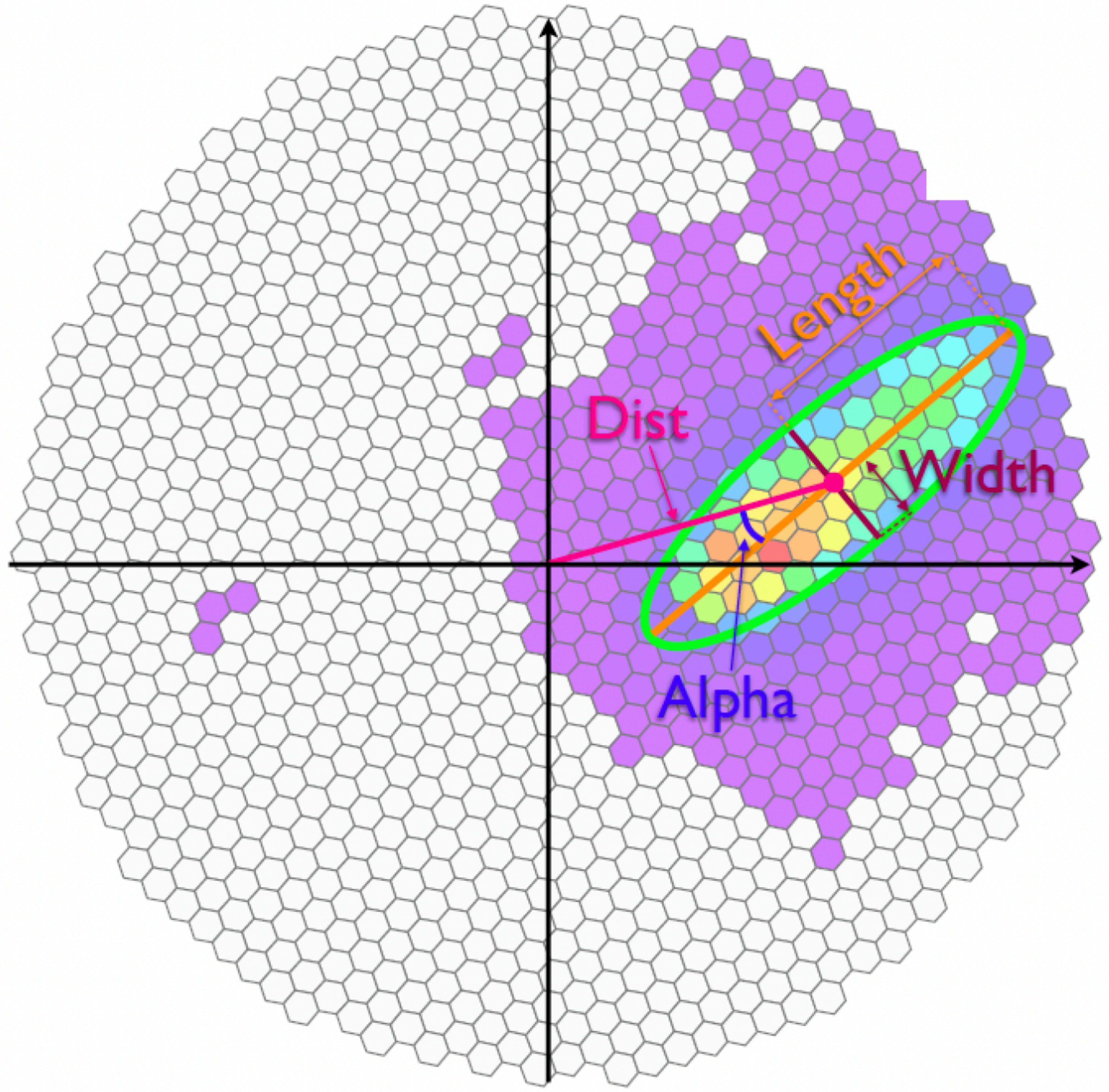
# Image cleaning and parametrization

- Images are cleaned to remove pixels without significant signal or related to spurious signals fro electronics and identify the shower

- Cleaned images are parametrized as elliptical shapes with so-called Hillas parameters




Abe et al., ApJ 956:80, 2023




Credits: A. Fernández Barral

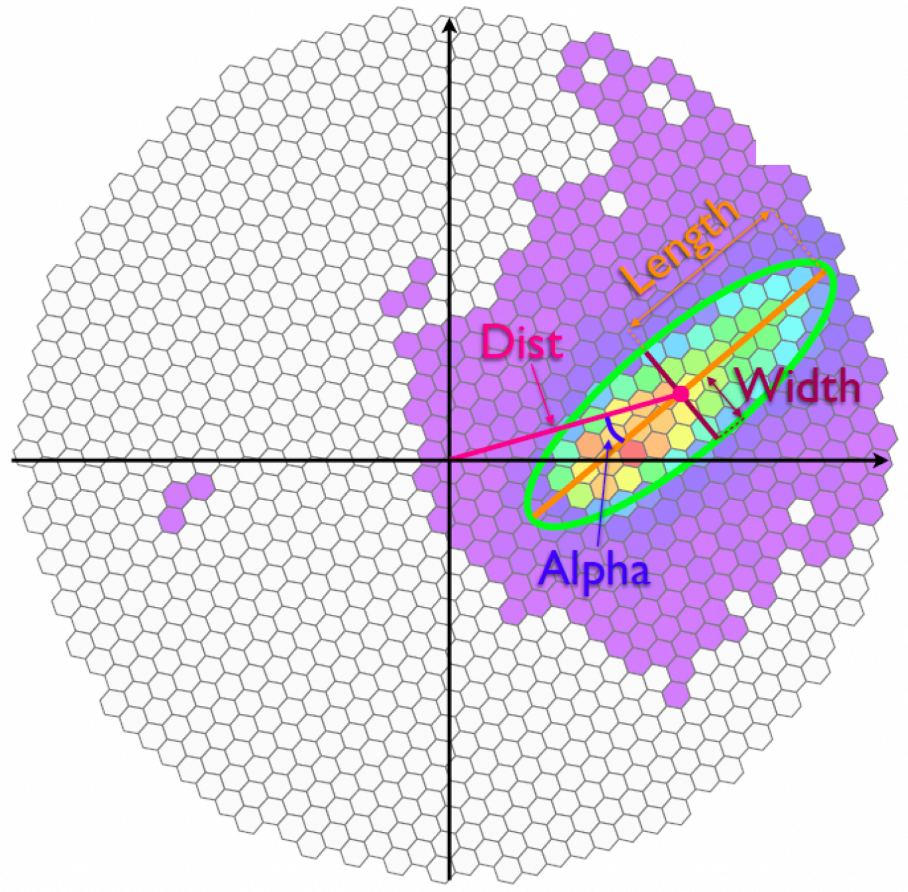
# Image cleaning and parametrization

*Source-independent parameters*

| | |
|---|---|
| `Size` | Total number of phe in the image |
| `Width` | Length of the semi-minor axis of the ellipse |
| `Length` | Length of the semi-major axis of the ellipse |
| `Center of Gravity (CoG)` | Coordinates of the weighted average signal in the camera plane |

Related to the primary particle energy

Related to the lateral and longitudinal development of the shower




*Source-dependent parameters*

| | |
|---|---|
| `Dist` | Distance between expected source position and CoG position |
| `Alpha` | Angle between ellipse major axis and expected source position - CoG line |

*Timing parameters*

| | |
|---|---|
| `TimeRMS` | Arrival times RMS of pixels surviving image cleaning |
| `Time gradient` | Slope of linear function used to fit the arrival time distribution of the pixels |

Smaller in γ-ray showers

*Directional parameters*

| | |
|---|---|
| `Asymmetry` | Sign of the difference between position of brightest pixel and CoG |

Differentiates between shower head-tail

*Image quality parameters*

| | |
|---|---|
| `LeakageN` | Fraction of image size contained in N outermost pixel rings of the camera |
| `Number of islands` | Number of non-connected pixel groups surviving image cleaning |

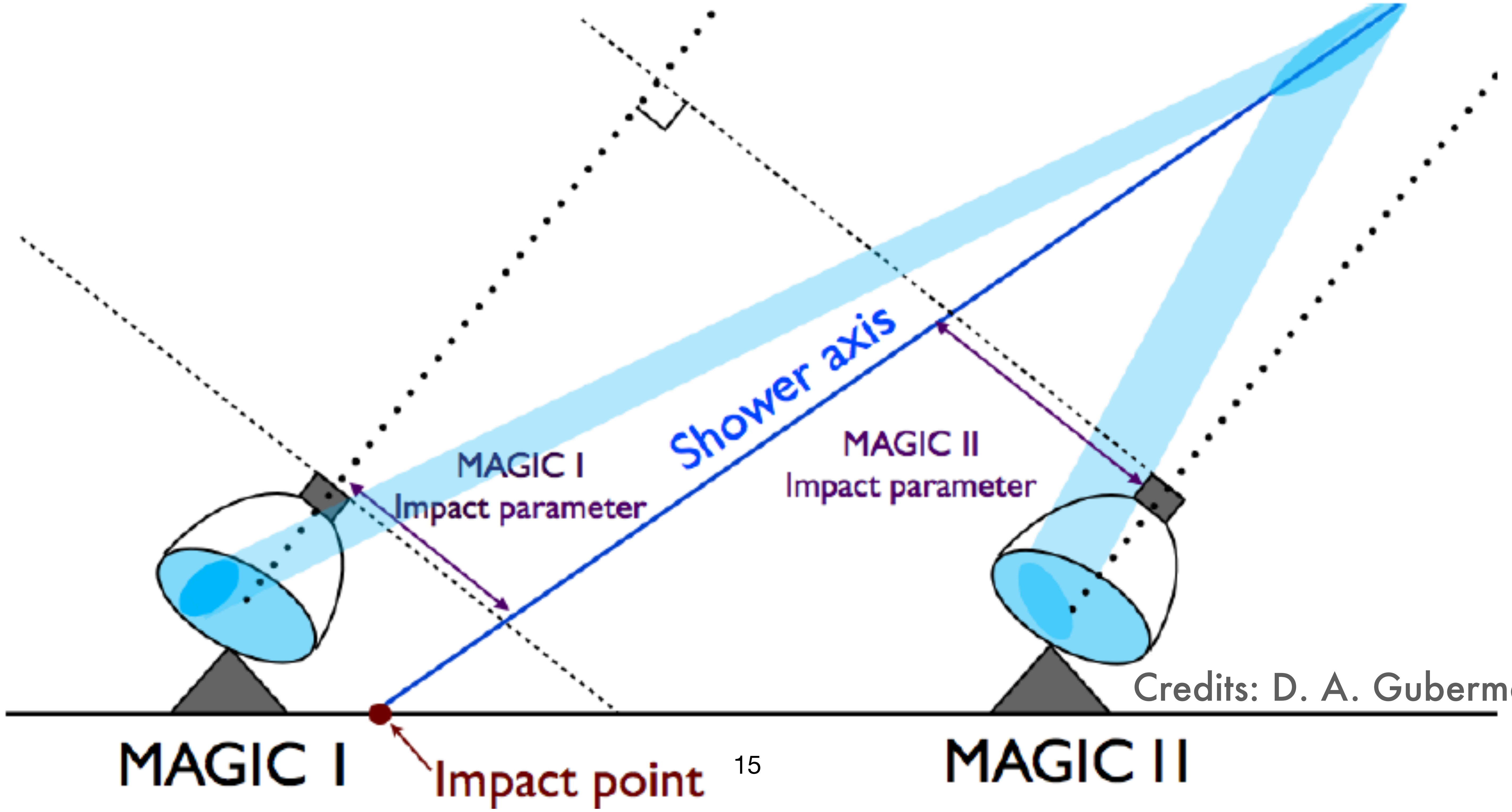
Estimates the fraction of signal loss

Larger for hadronic showers (usually are more fragmentated)

# Image cleaning and parametrization

*Stereo parameters*

| | |
|---|---|
| Shower axis | Direction of the shower |
| Impact parameter | Distance between the shower axis and the pointing direction of the telescope |
| Impact point | Impact position of the shower on the ground |
| Height of shower max | Height at which the number of particles in the EAS is maximum. It depends on the energy of the primary particle. |




Credits: D. A. Guberman

# Standard data analysis method

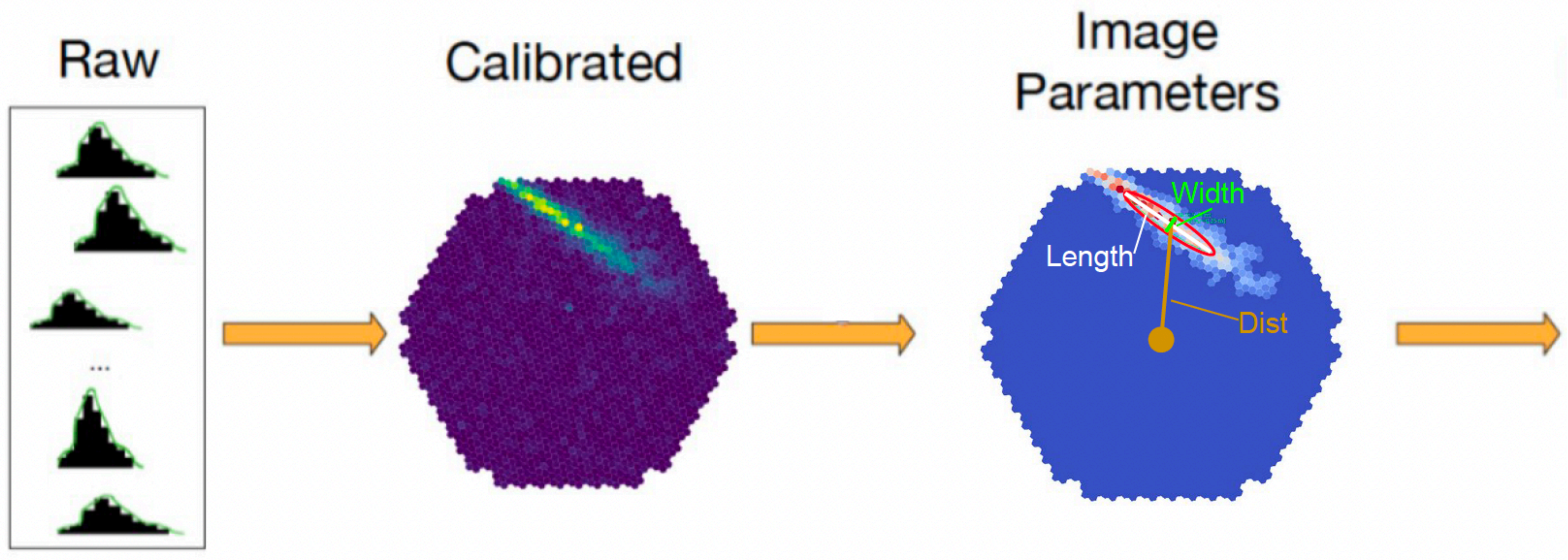





Image credits: P. Grespan

# Standard data analysis method

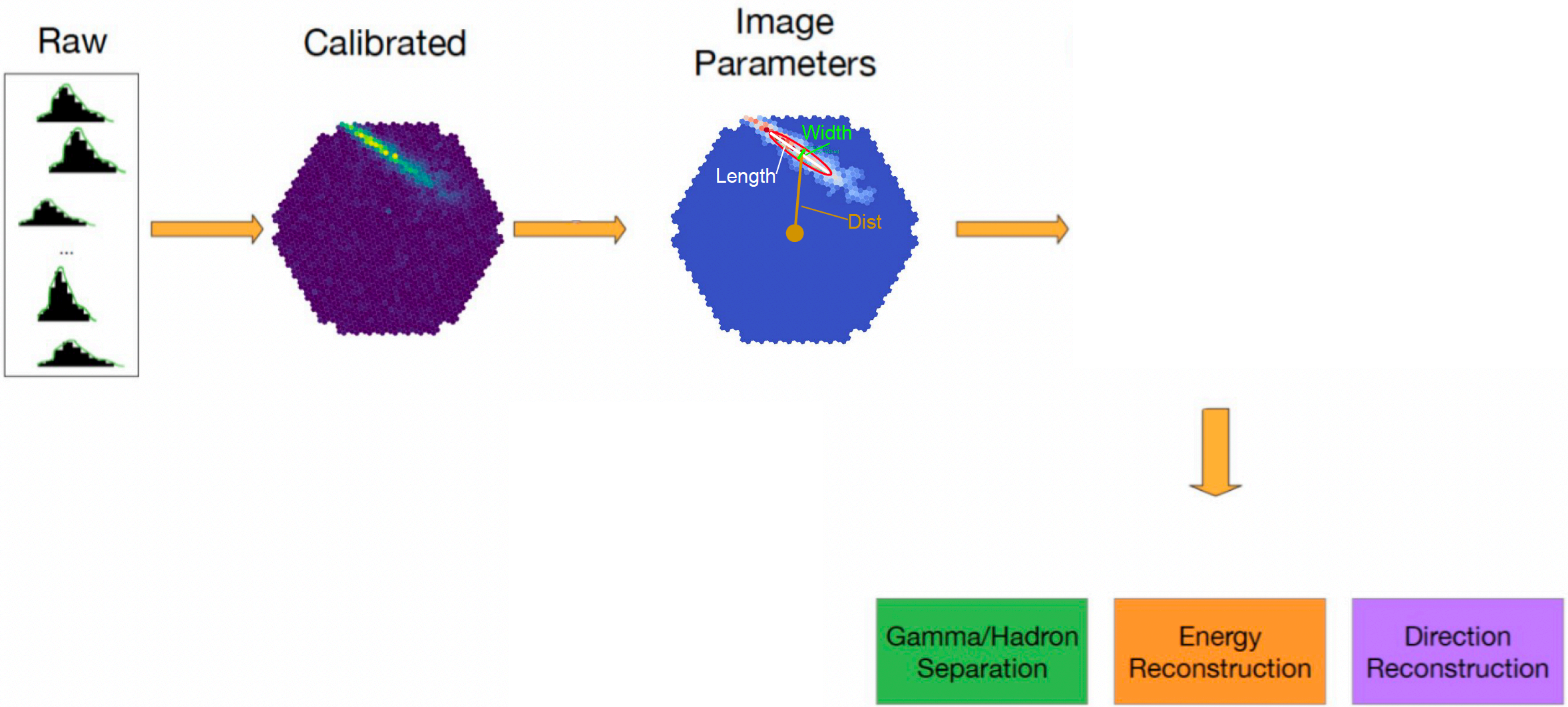





Image credits: P. Grespan

# Standard data analysis method

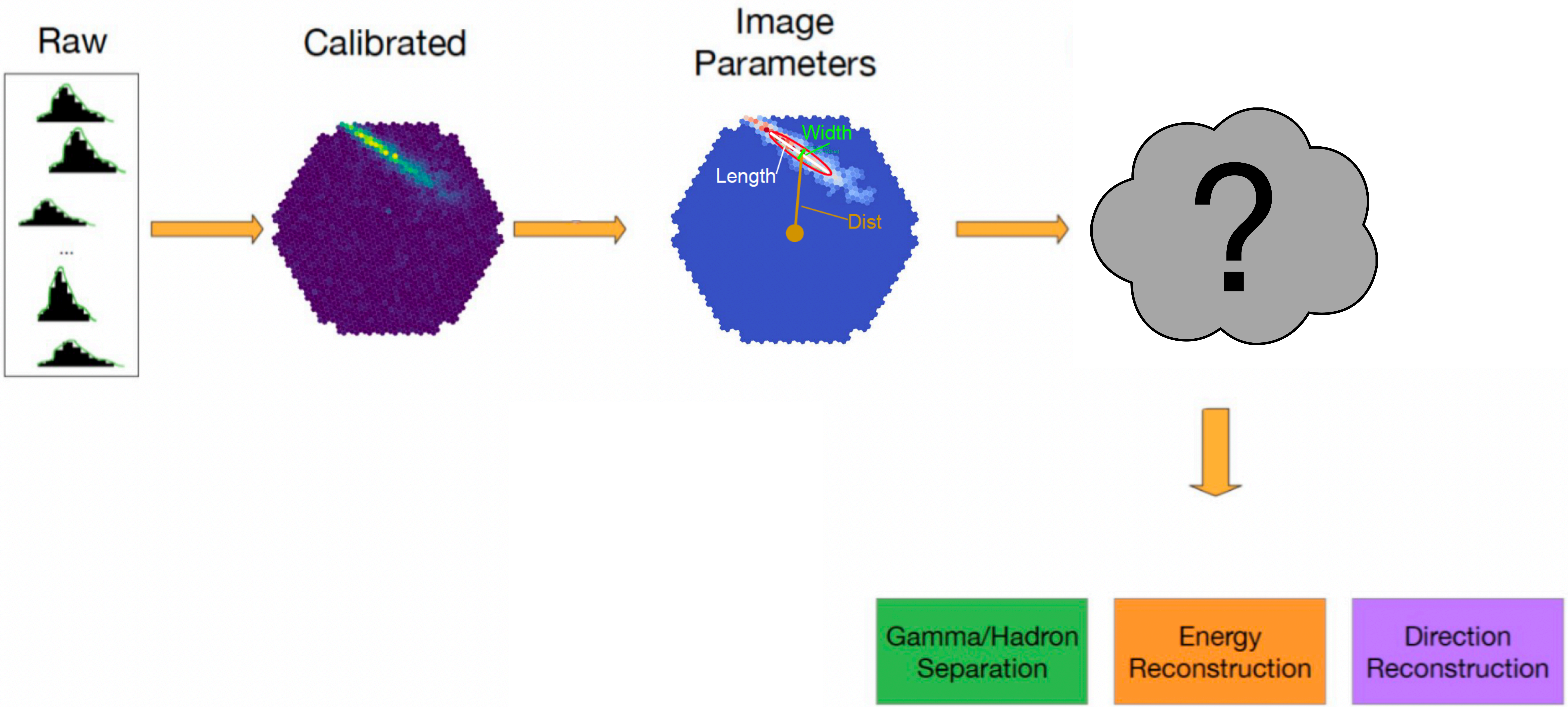




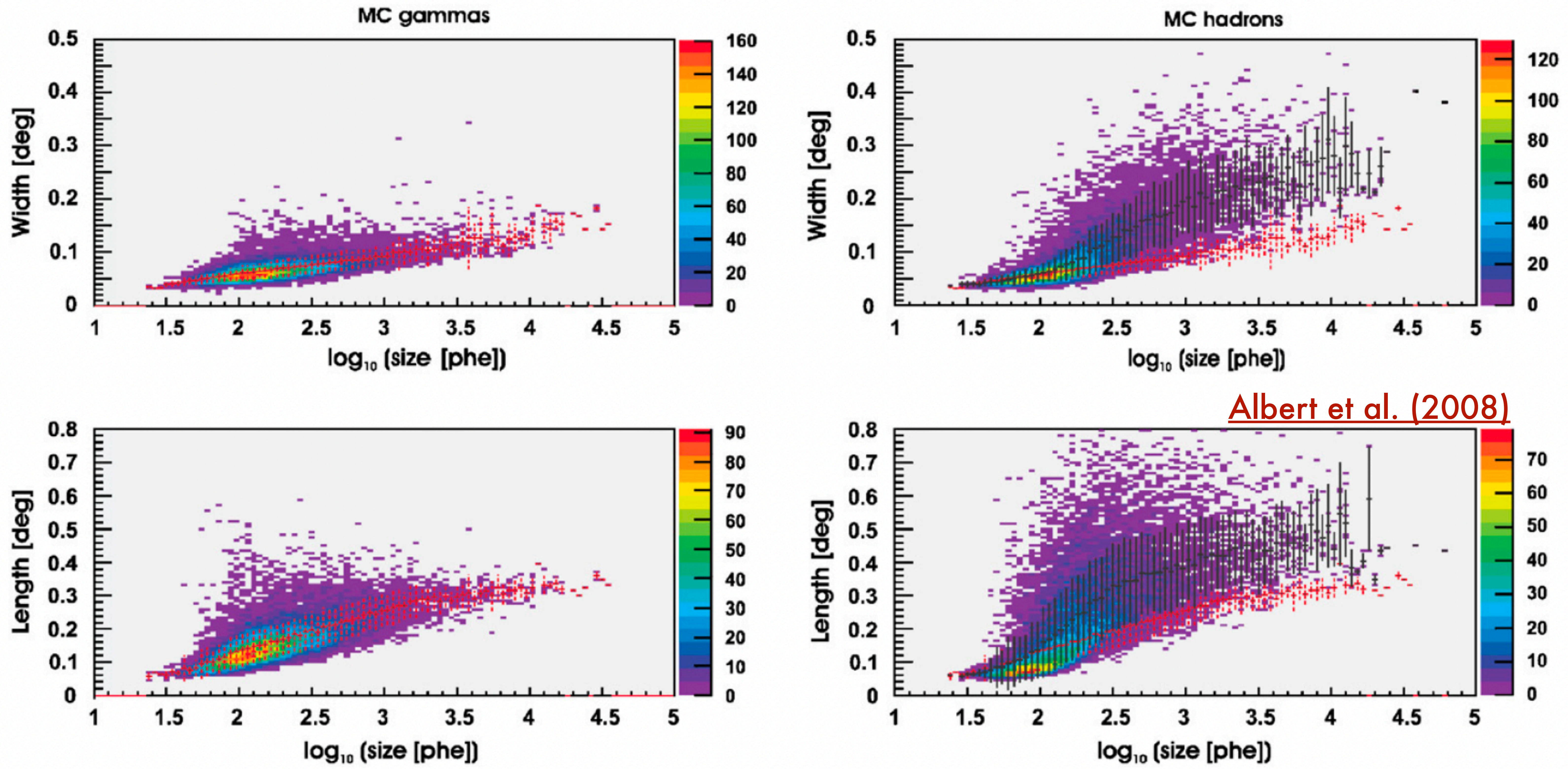

Image credits: P. Grespan

# Analysis technique before machine learning

Different techniques, mostly based on Hillas parametrization or on semi-analytical models

Hillas-parameter based analysis

- Relies on MC simulations
- Defines static cuts to discriminate between γ and hadrons
- Values of image parameters are compared with expectation values from MCs




Albert et al. (2008)

width and length as good separation parameters, at least for size > 200 phe (i.e. E > 100 GeV)

# Standard data analysis method

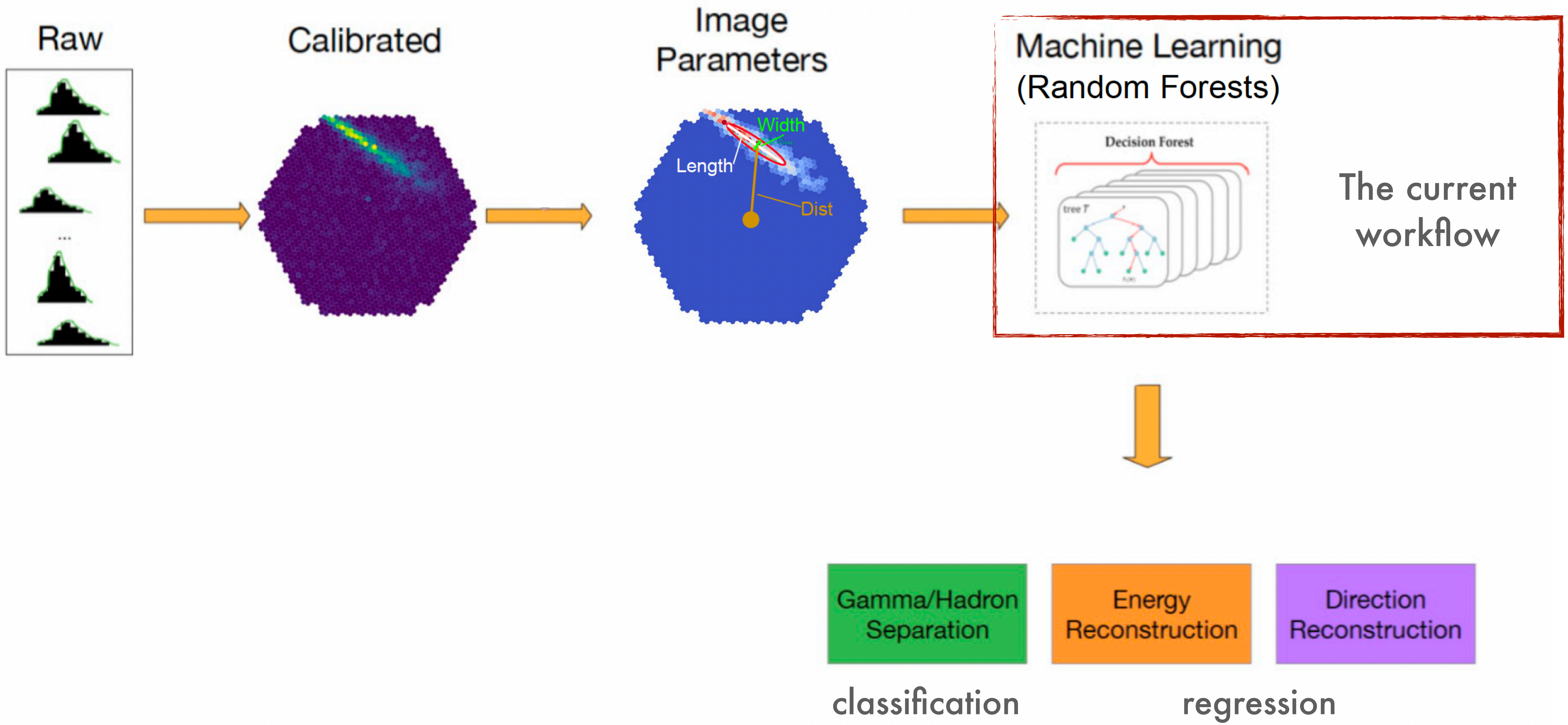
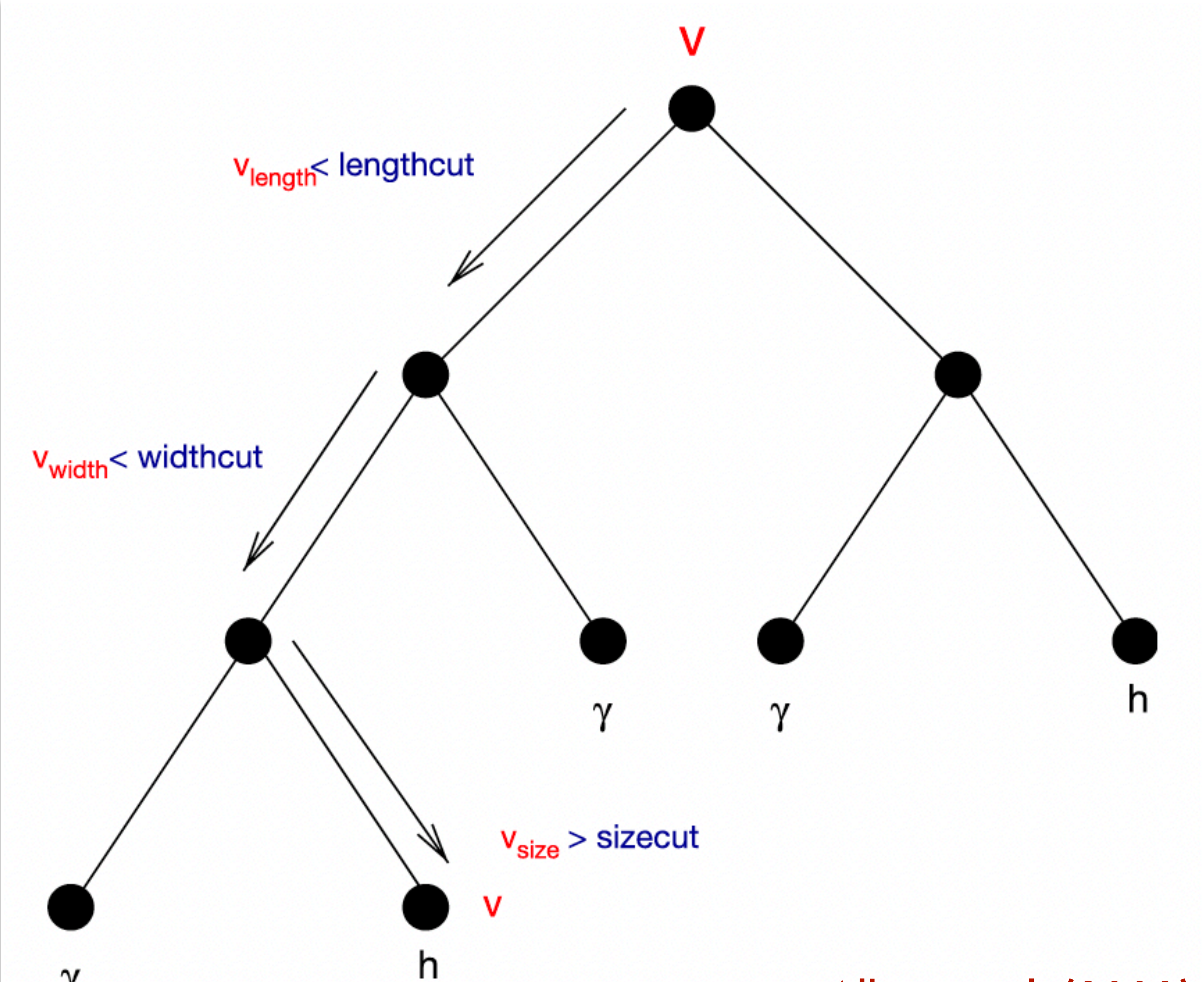





Image credits: P. Grespan

# Current analysis method: Random Forest

- Based on *Random Forests* (RF), one for each task
  - Collection of uncorrelated decision trees, combining their individual results to make the prediction
- Event characterized by vector of image parameters
- Training on MC γ and real background
- MC and bkg data have to match as much as possible the observational conditions of the source data
  - e.g. zenith angle, dark/moon nights, extragalactic/galactic obs.

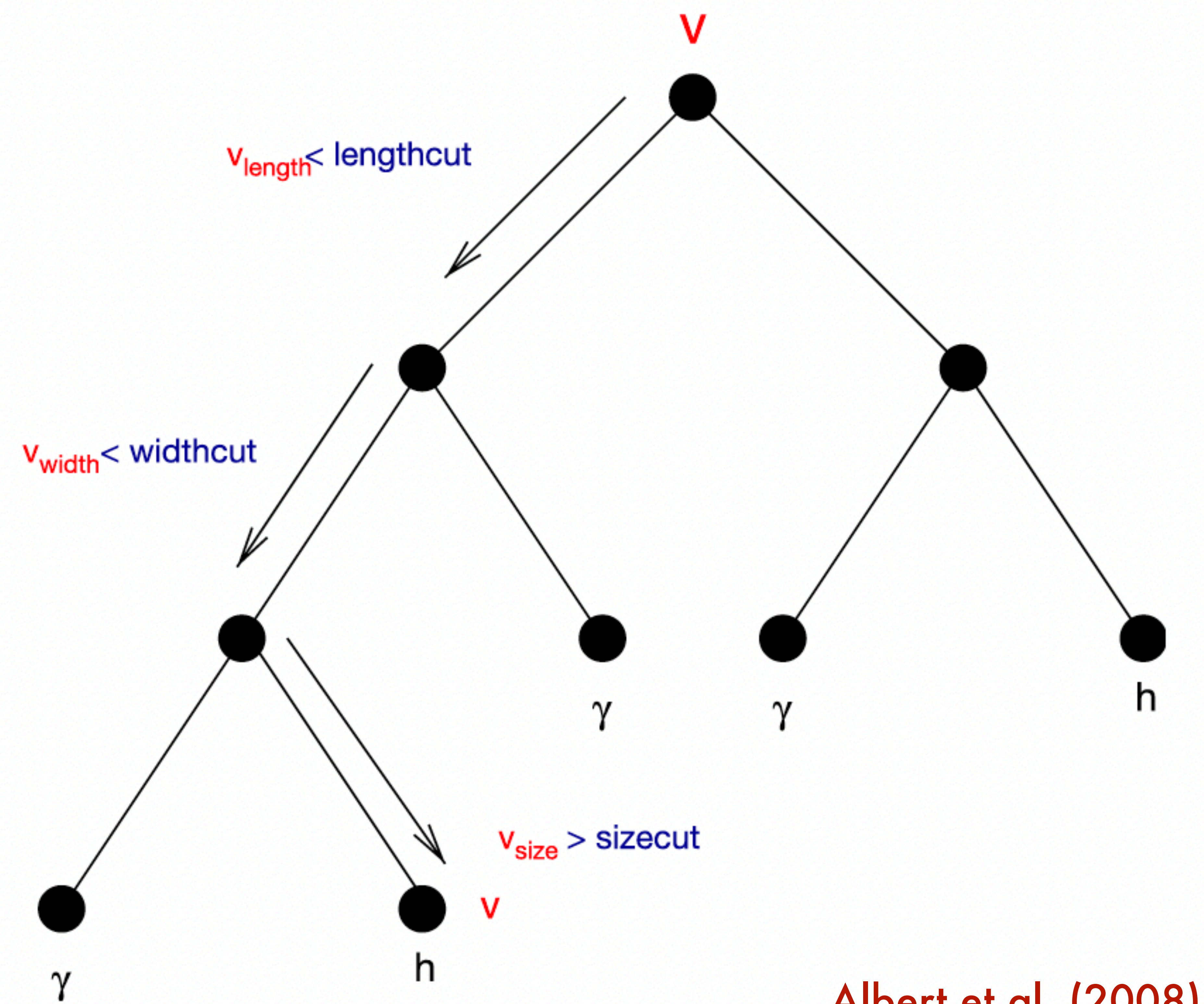



Albert et al. (2008)

# Current analysis method: Random Forest

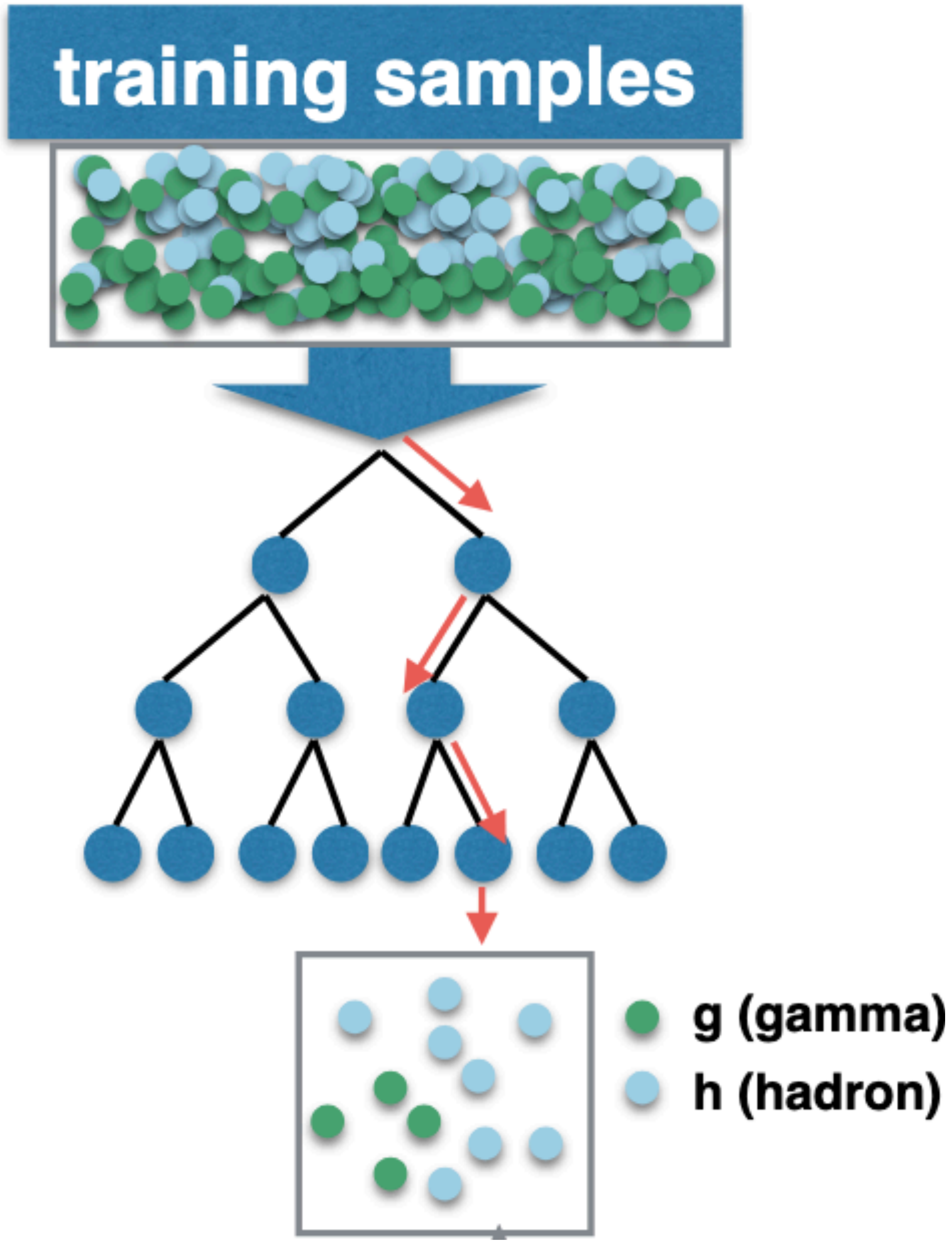
How does the splitting work?

- Full sample (i.e. full parameter space) in root node
- Splitting of each node using on parameter at a time and an optimized cut value
- Splitting process stops if
  - Events per node below defined limit
  - Only events of one class in the node




Albert et al. (2008)

# Identification of particle type

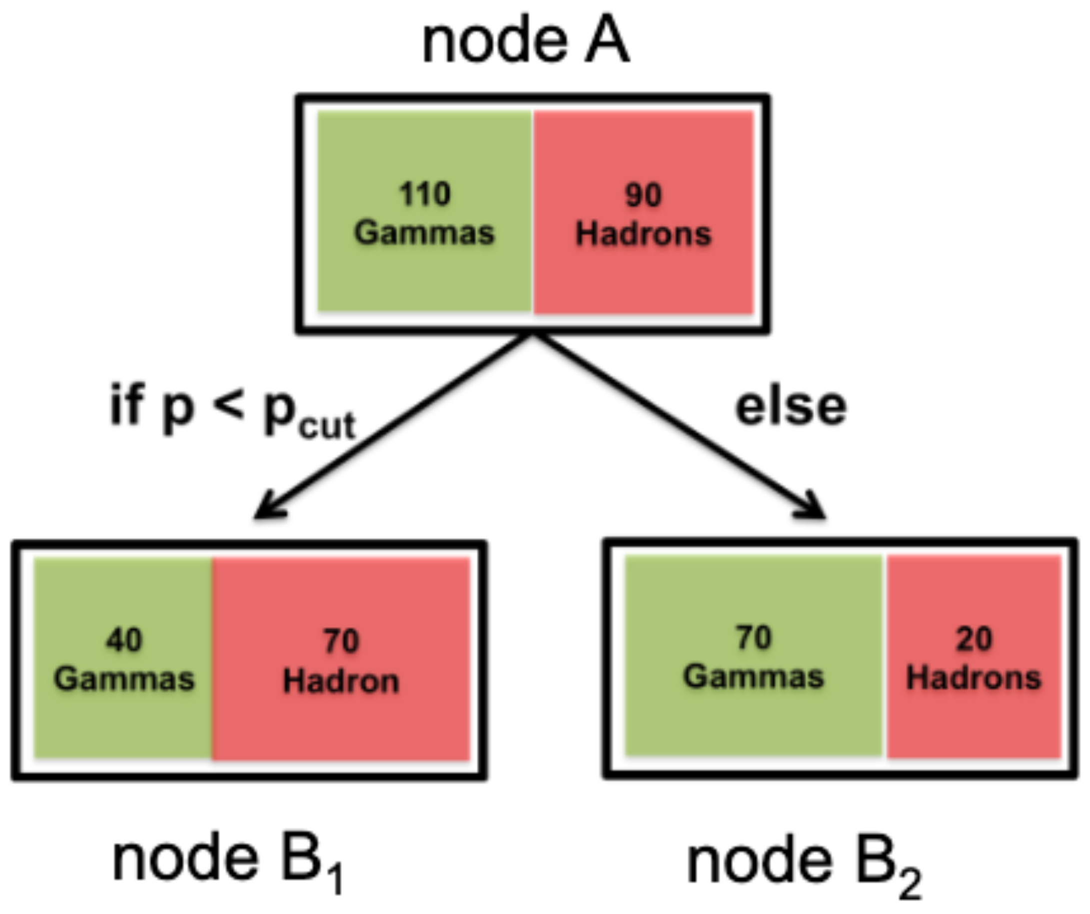



- Best cut value obtained through minimization of the *Gini index*
  - Measure of the inequality in two distributions as a function of a cut in a variable

$$Gini = \frac{4N_\gamma N_h}{(N_\gamma + N_h)^2}$$

After split: weighed average of Gini in each node

Nodes at same depth

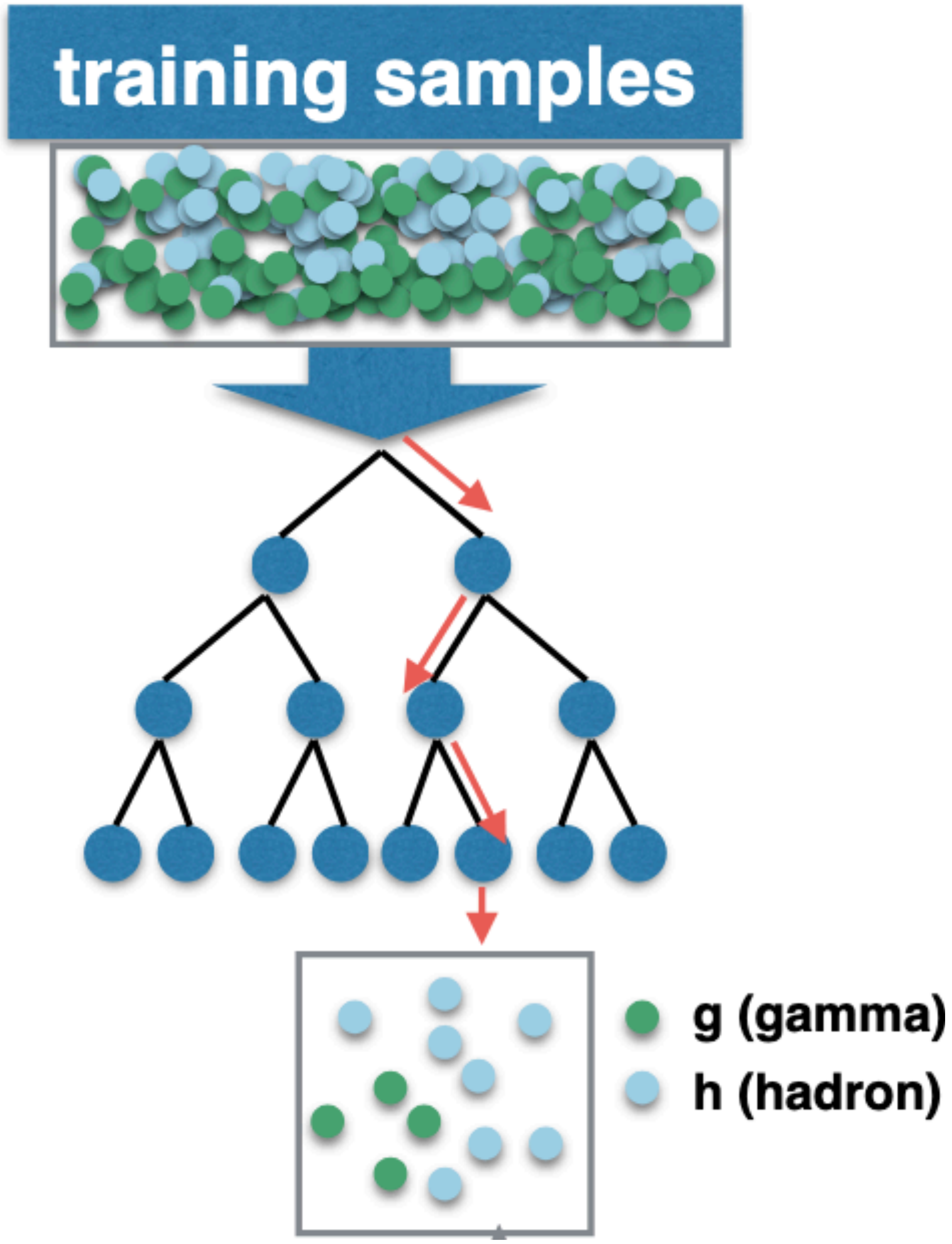



**The smaller the Gini, the better the separation**



Image credits: K. Ishio

# Identification of particle type




- At each terminal node a *hadronness* value is assigned based on its population (gammas=0, hadrons=1)

$$h_i = \frac{N_\gamma \cdot 0 + N_h \cdot 1}{N_\gamma + N_h} \Rightarrow hadronness = \frac{1}{n}\sum_{i=0}^{n} h_i$$

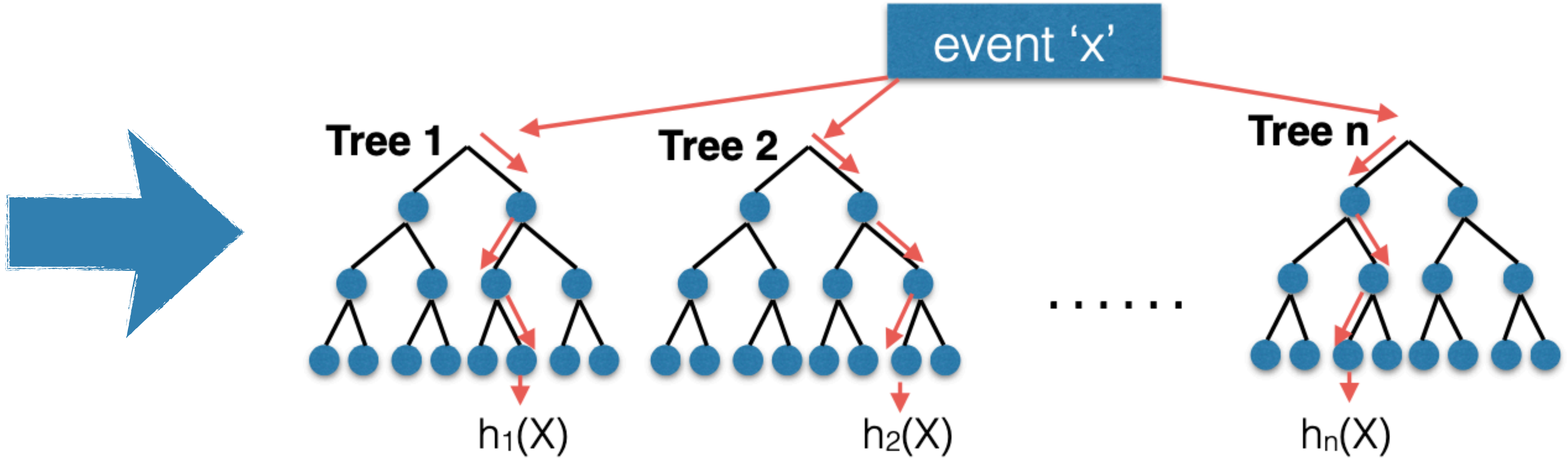
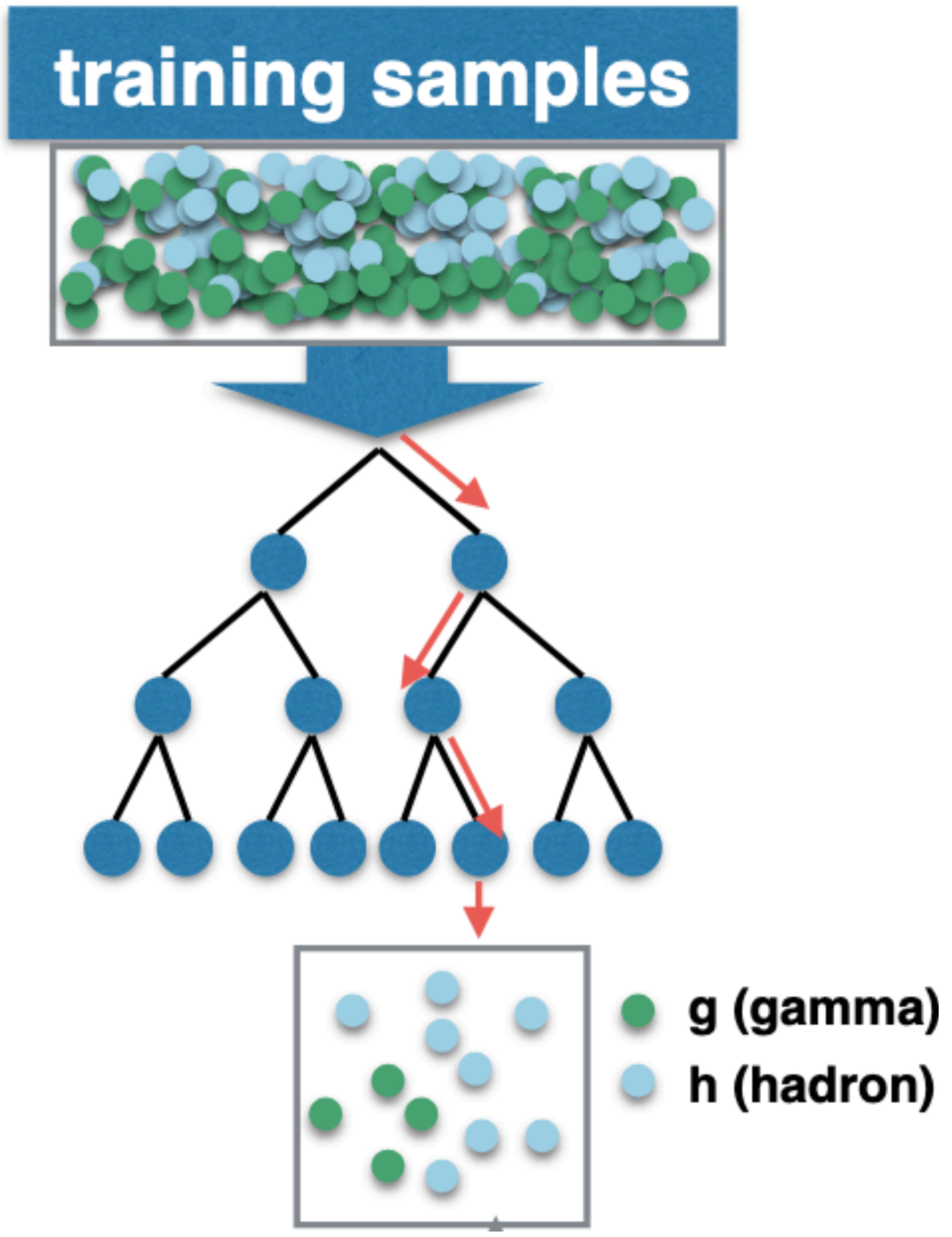





Image credits: K. Ishio

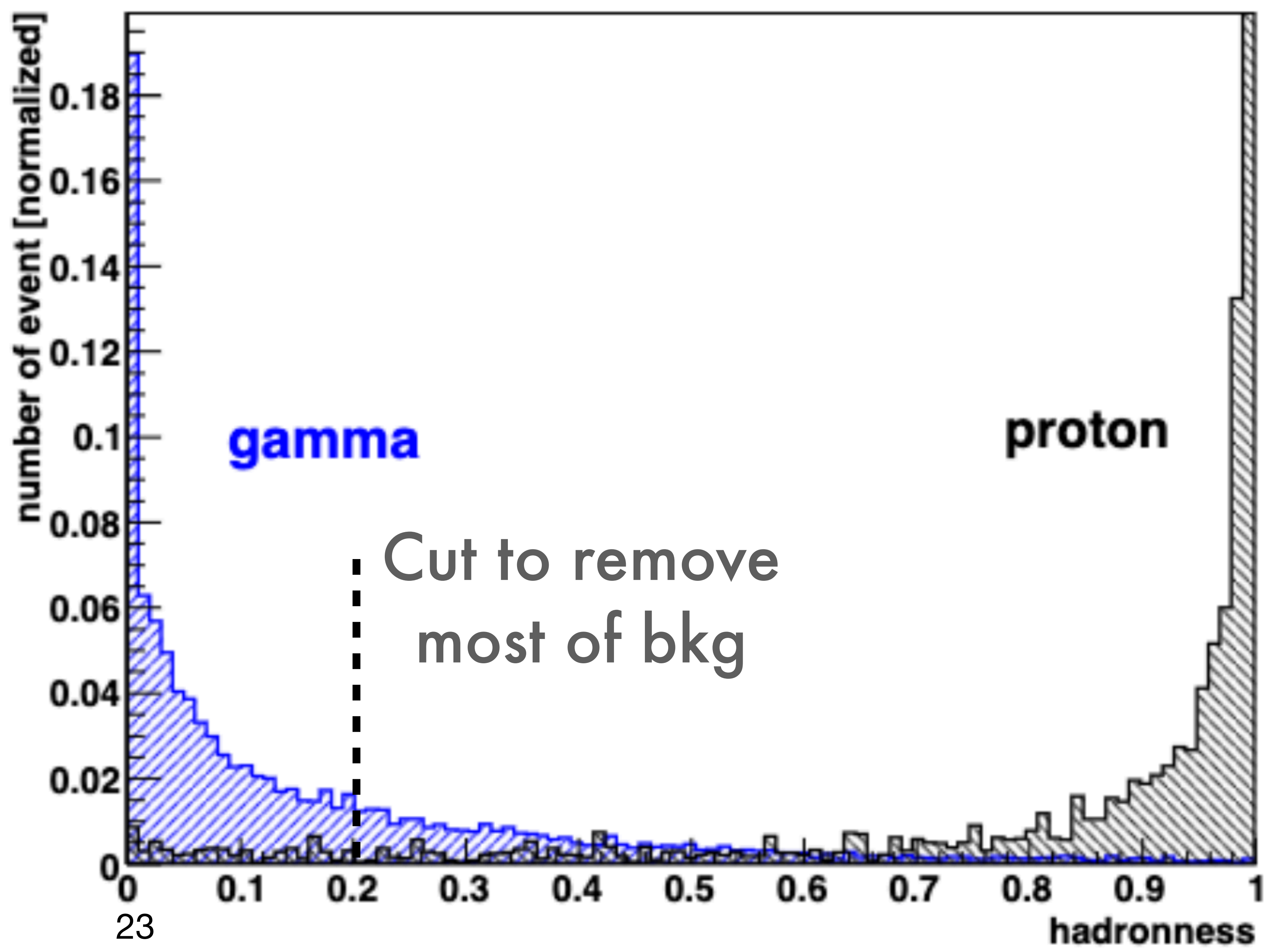
# Identification of particle type




Credits: K. Ishio

- At each terminal node a *hadronness* value is assigned based on its population (gammas=0, hadrons=1)

$$h_i = \frac{N_\gamma \cdot 0 + N_h \cdot 1}{N_\gamma + N_h} \Rightarrow hadronness = \frac{1}{n}\sum_{i=0}^{n} h_i$$



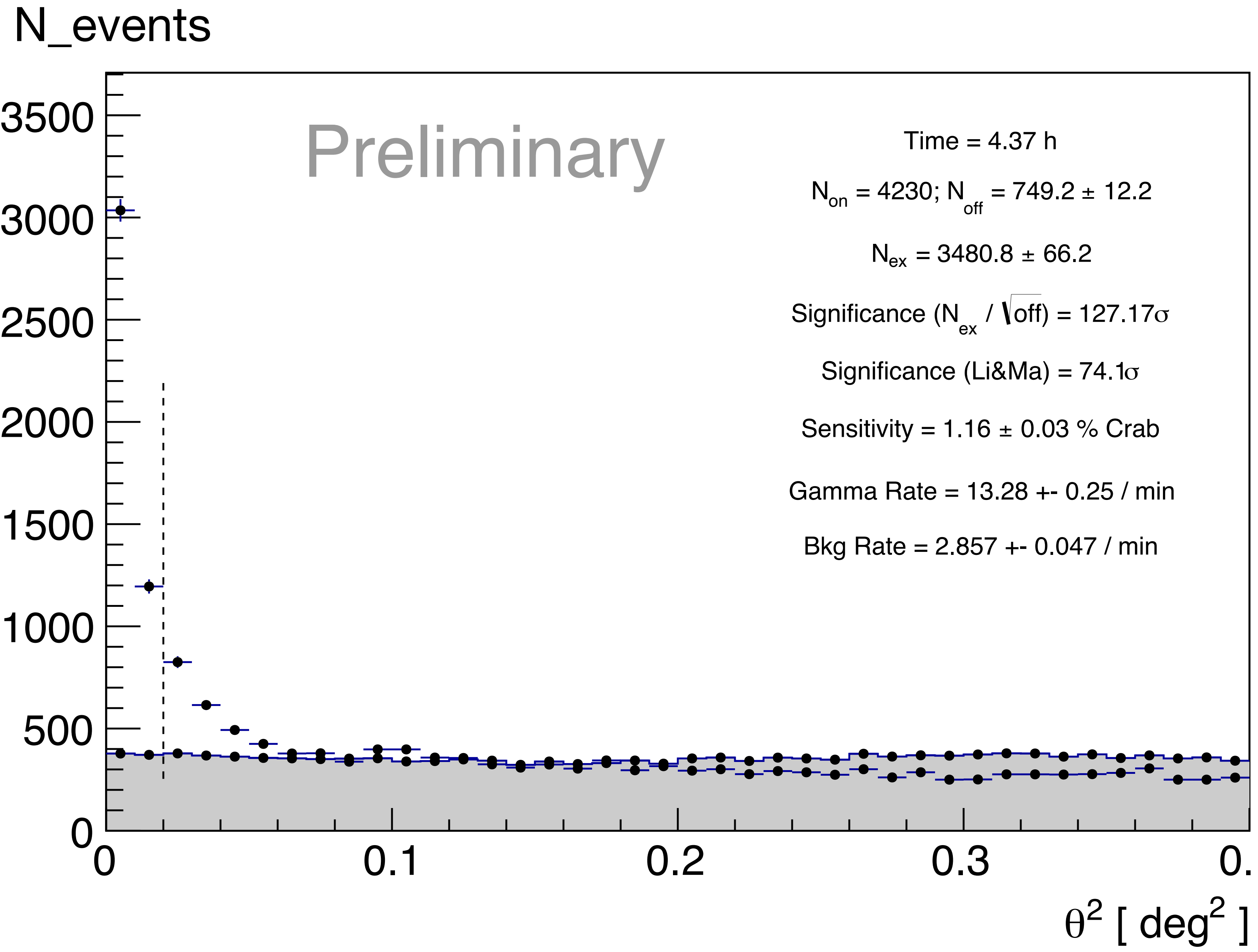

Colin et al., ICRC (2009)

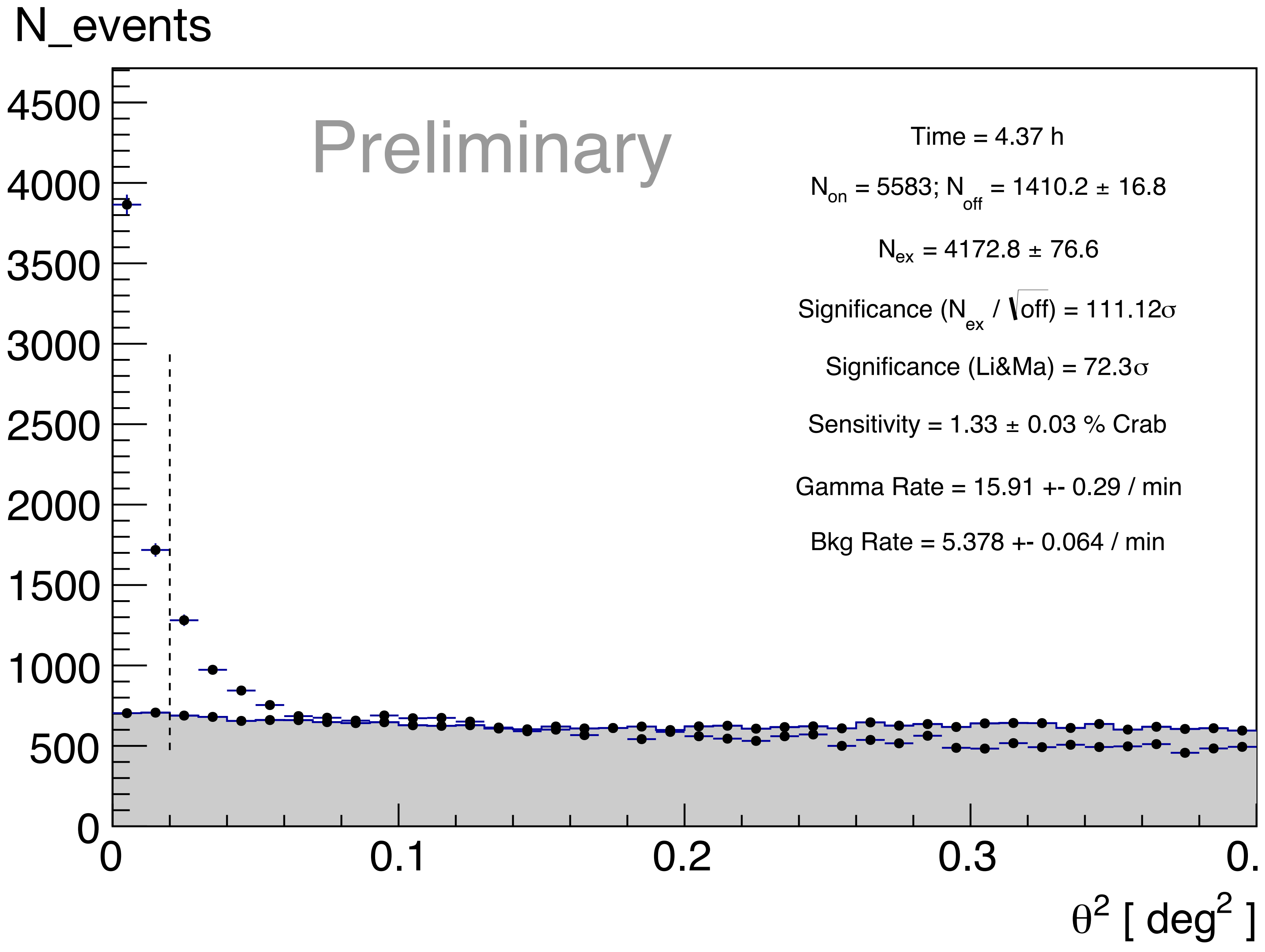
# Effect of gamma/hadron cut

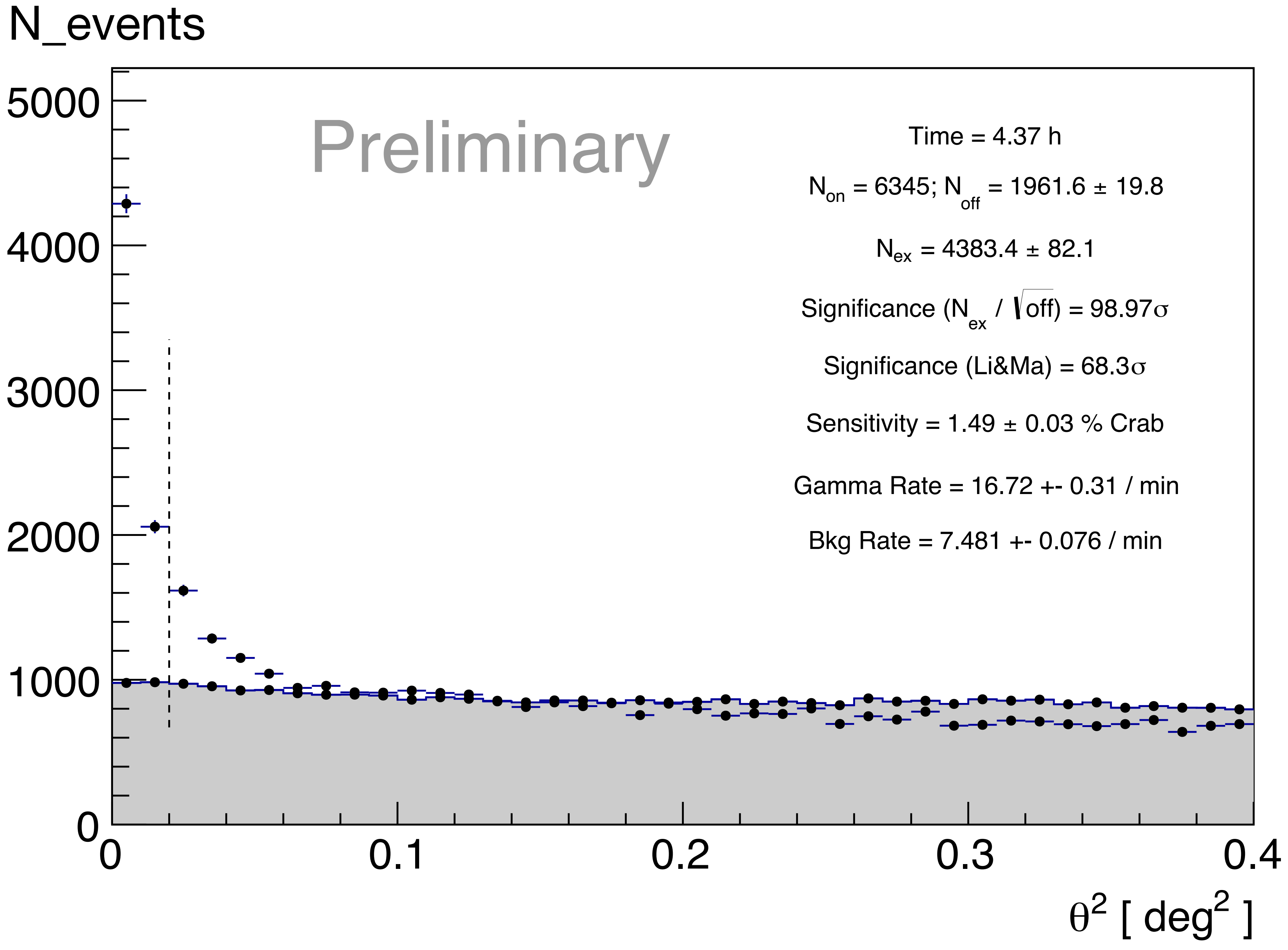
Hadronness < 0.2

Hadronness < 0.4

Hadronness < 0.6



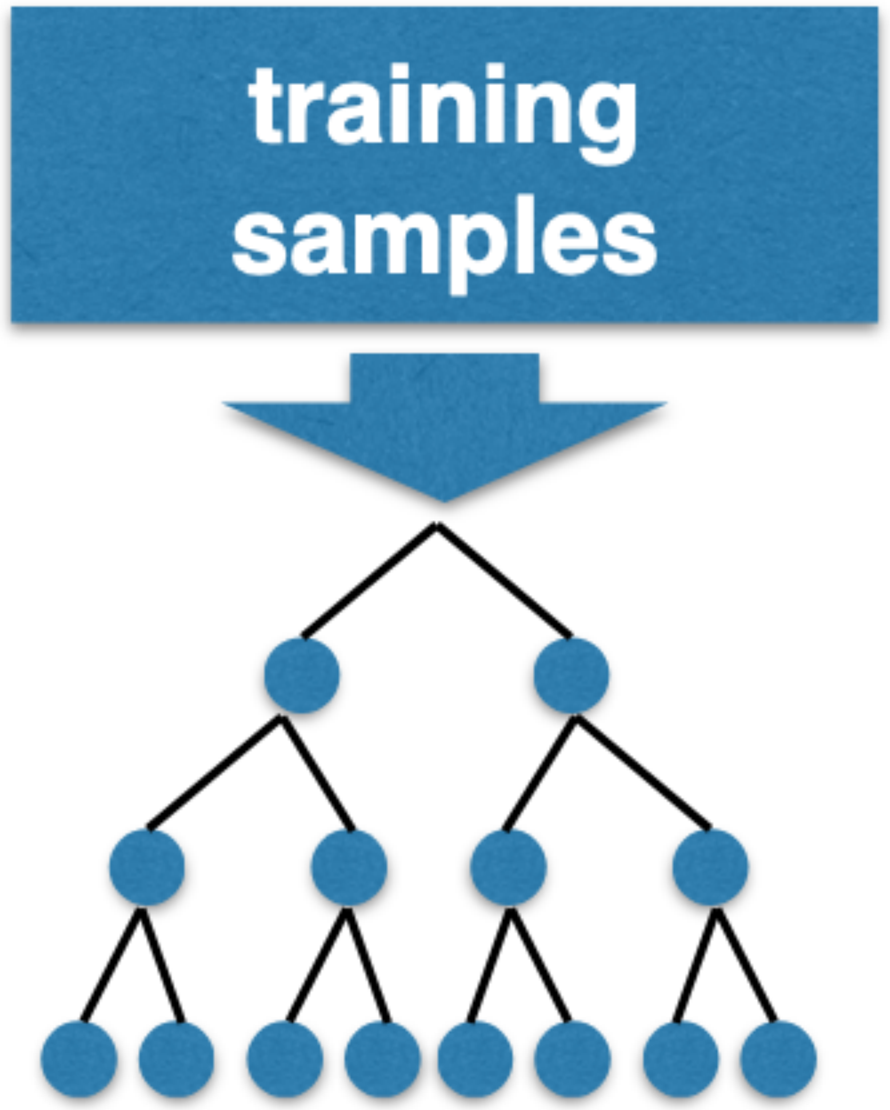



# Energy reconstruction

- Splitting rule not relying on class population
- To purify the node population wrt the energy distribution the variance is used
- In analogy to the Gini index, the weighted average of the variance in minimized to find the best cut:

$$\sigma^2(E) = \frac{1}{N_L + N_R}(N_L\sigma_L^2 + N_R\sigma_R^2)$$



Credits: K. Ishio

- The energy at the node is given by the average of the population at the node
- The final energy is given by averaging the results in each tree: $E = \frac{1}{n}\sum_{i=0}^{n} E_i$

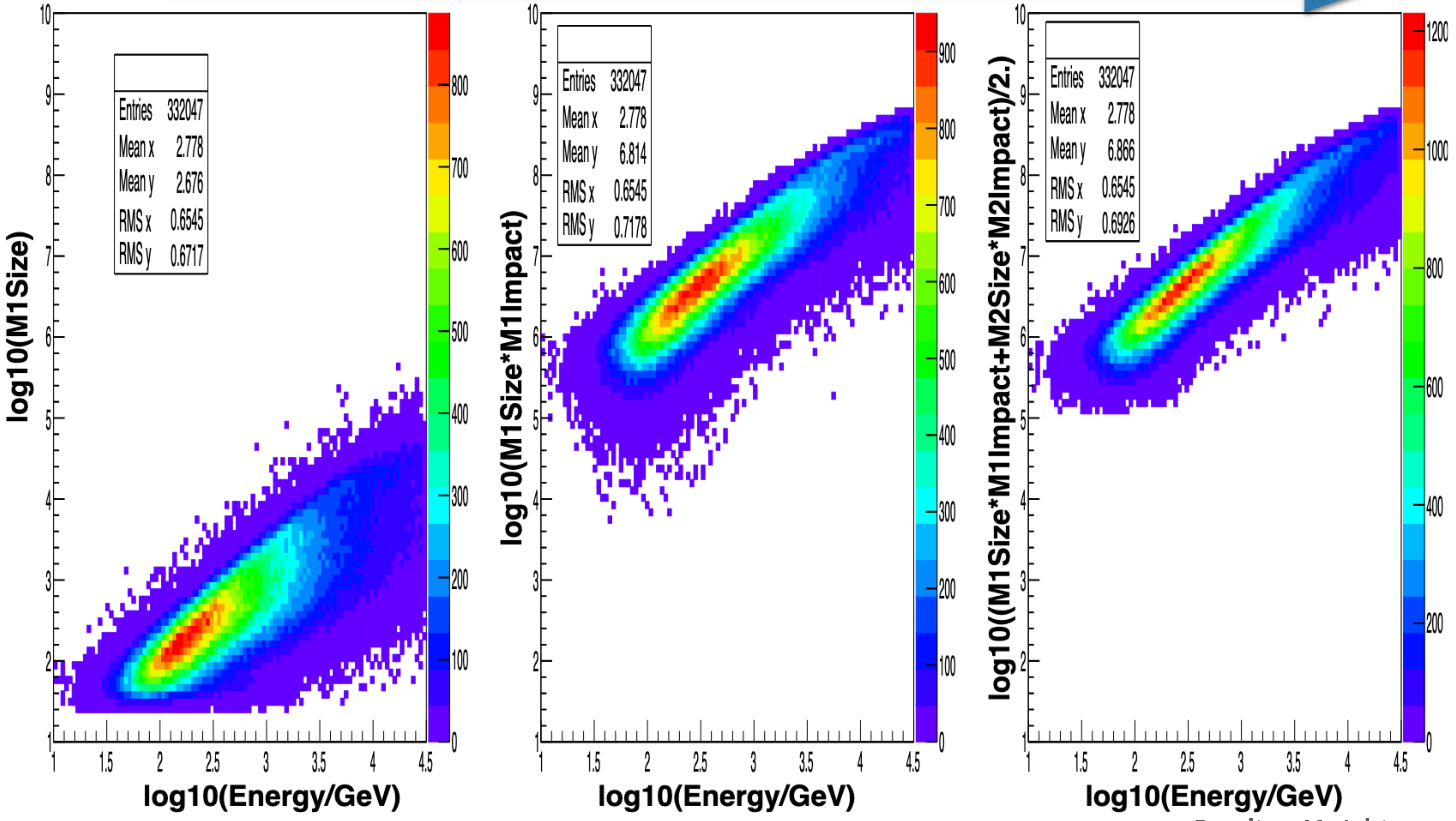
# Energy reconstruction

Main parameters related to energy are:

- Size
  - number of Cherenkov photons
  - ∝ Energy
- Impact
  - distance of the telescope axis to the shower

    ⇒ smaller size

Correcting for distance improves size-energy correlation




Credits: K. Ishio

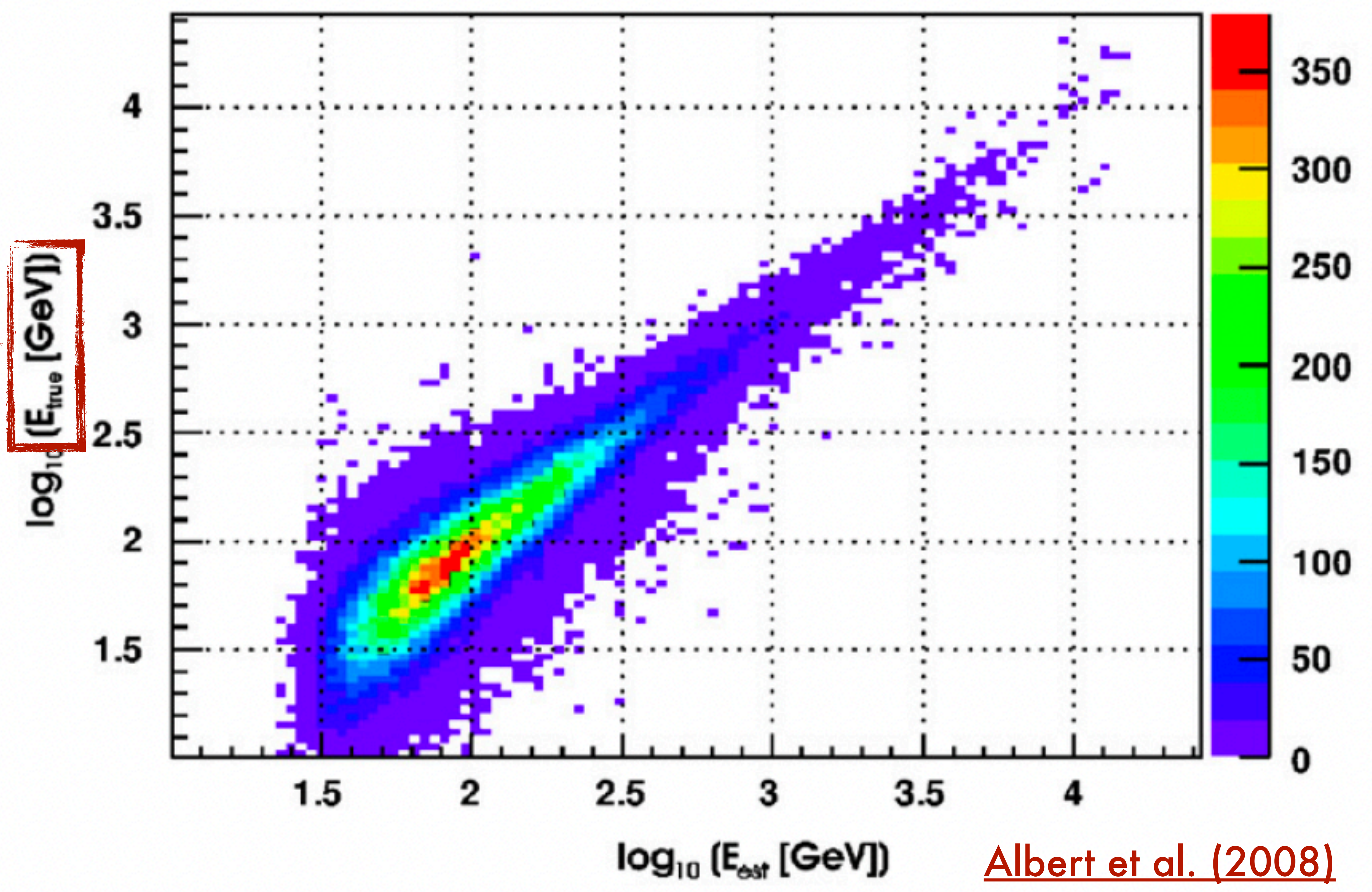
# Energy reconstruction

Main parameters related to energy are:

- Size
    - number of Cherenkov photons
    - $\propto$ Energy
- Impact
    - distance of the telescope axis to the shower

      $\Rightarrow$ smaller size

Reconstructed energy

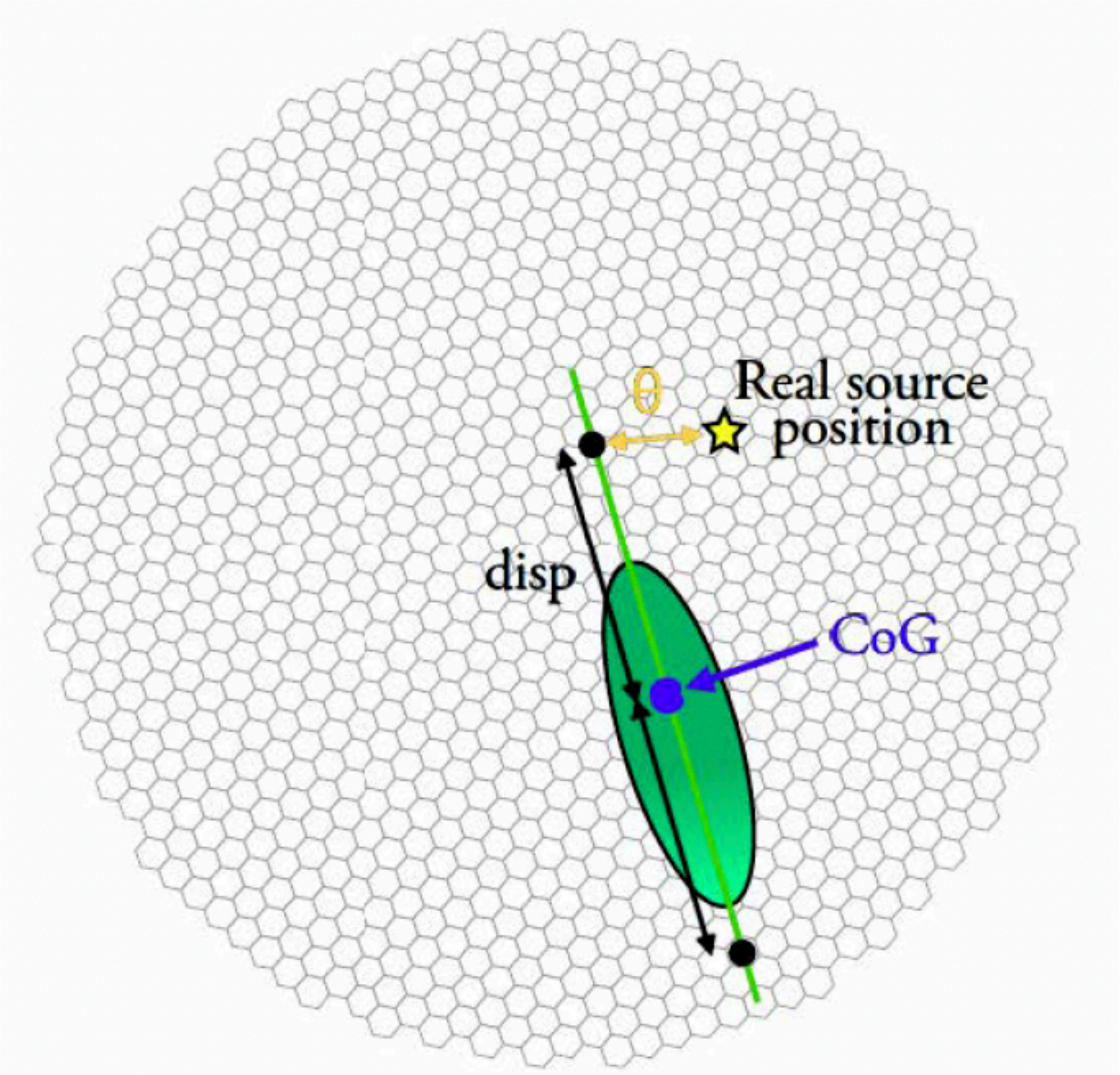



Albert et al. (2008)

True energy of the shower known from MCs

# Direction reconstruction

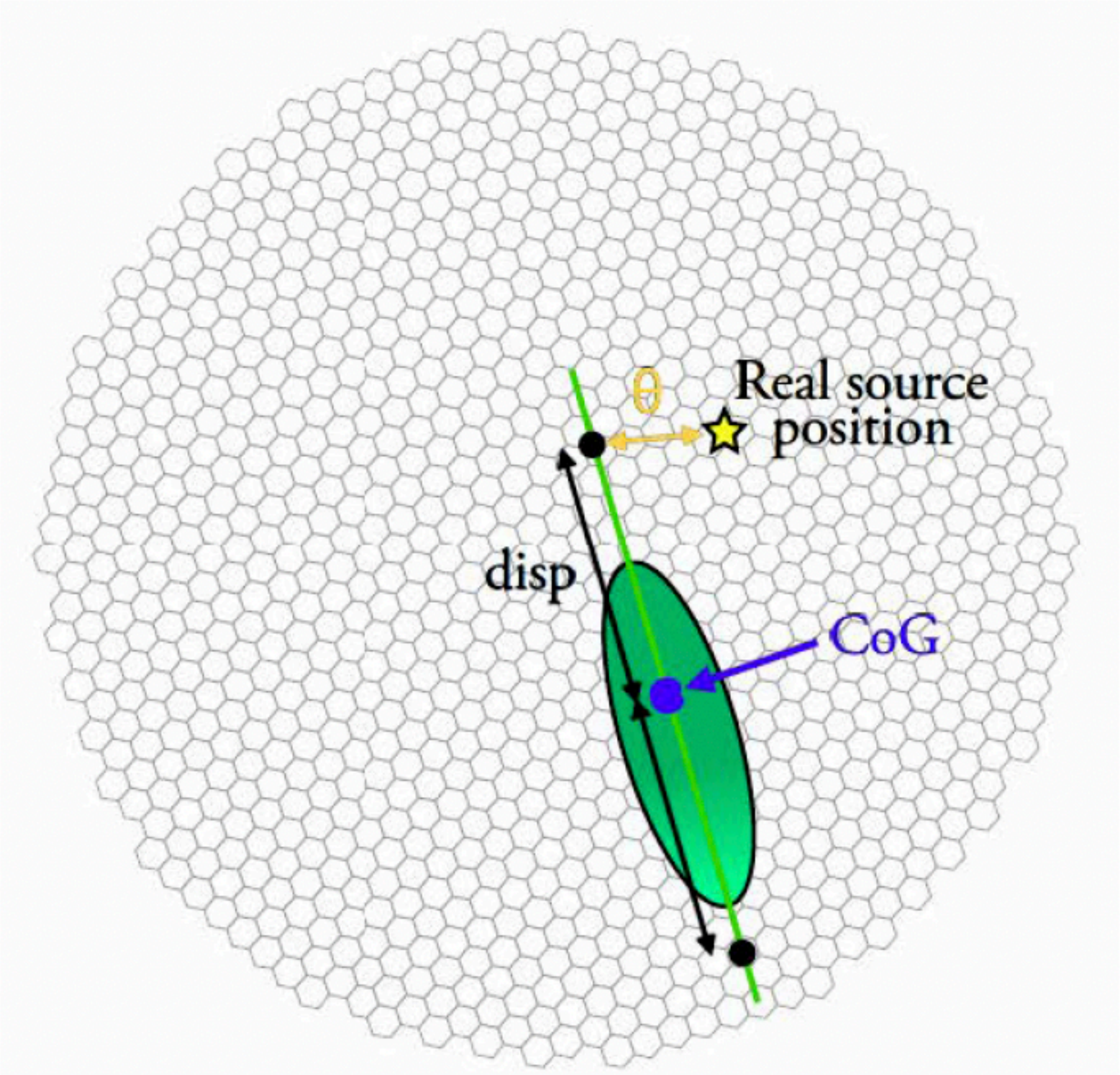
- Based on *DISP* parameter
- Source position assumed to be on the ellipse major axis
- Source position related to image shape and photons arrival time
  - Old *DISP* method: image shape only
  - New *DISP-RF*: also timing information
- Two possible positions are found for each image
- Degeneracy is broken thanks to asymmetry in charge distribution → "head-tail discrimination"




Credits: A. Fernández Barral

# Direction reconstruction

- Based on *DISP* parameter
- Source position assumed to be on the ellipse major axis
- Source position related to image shape and photons arrival time
  - Old *DISP* method: image shape only
  - New *DISP-RF*: also timing information
- Two possible positions are found for each image
- Degeneracy is broken thanks to asymmetry in charge distribution → "head-tail discrimination"

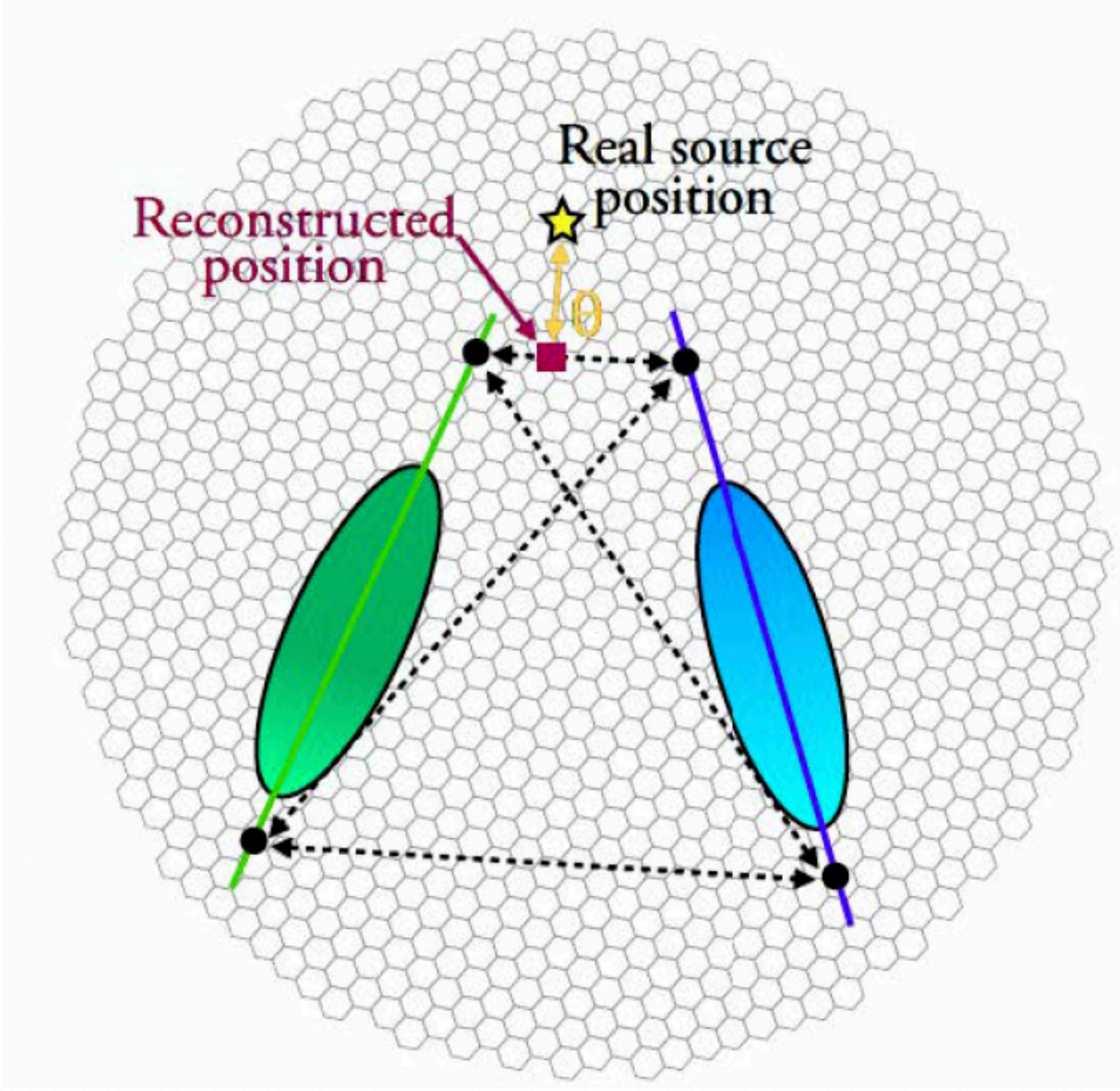



Credits: A. Fernández Barral

Reconstruction of DISP and ASYMMETRY parameters

# Direction reconstruction

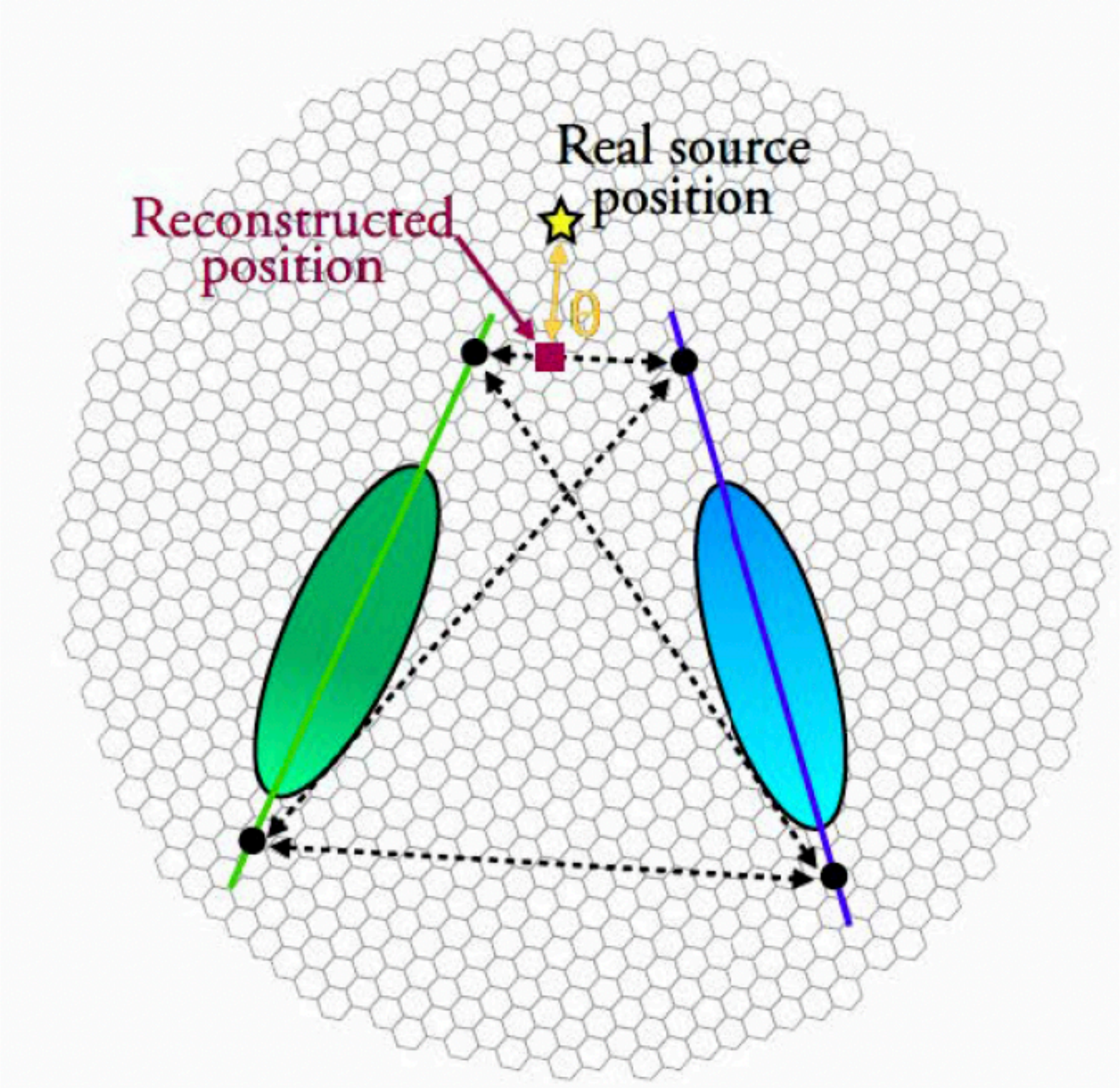
- For stereo observations:
  - All possible position combinations are considered
  - The combination giving the smaller distance is selected
  - Final source position is estimated as average of computed positions weighted with number of pixels in images




Credits: A. Fernández Barral

# Direction reconstruction

- For stereo observations:
  - All possible position combinations are considered
  - The combination giving the smaller distance is selected
  - Final source position is estimated as average of computed positions weighted with number of pixels in images

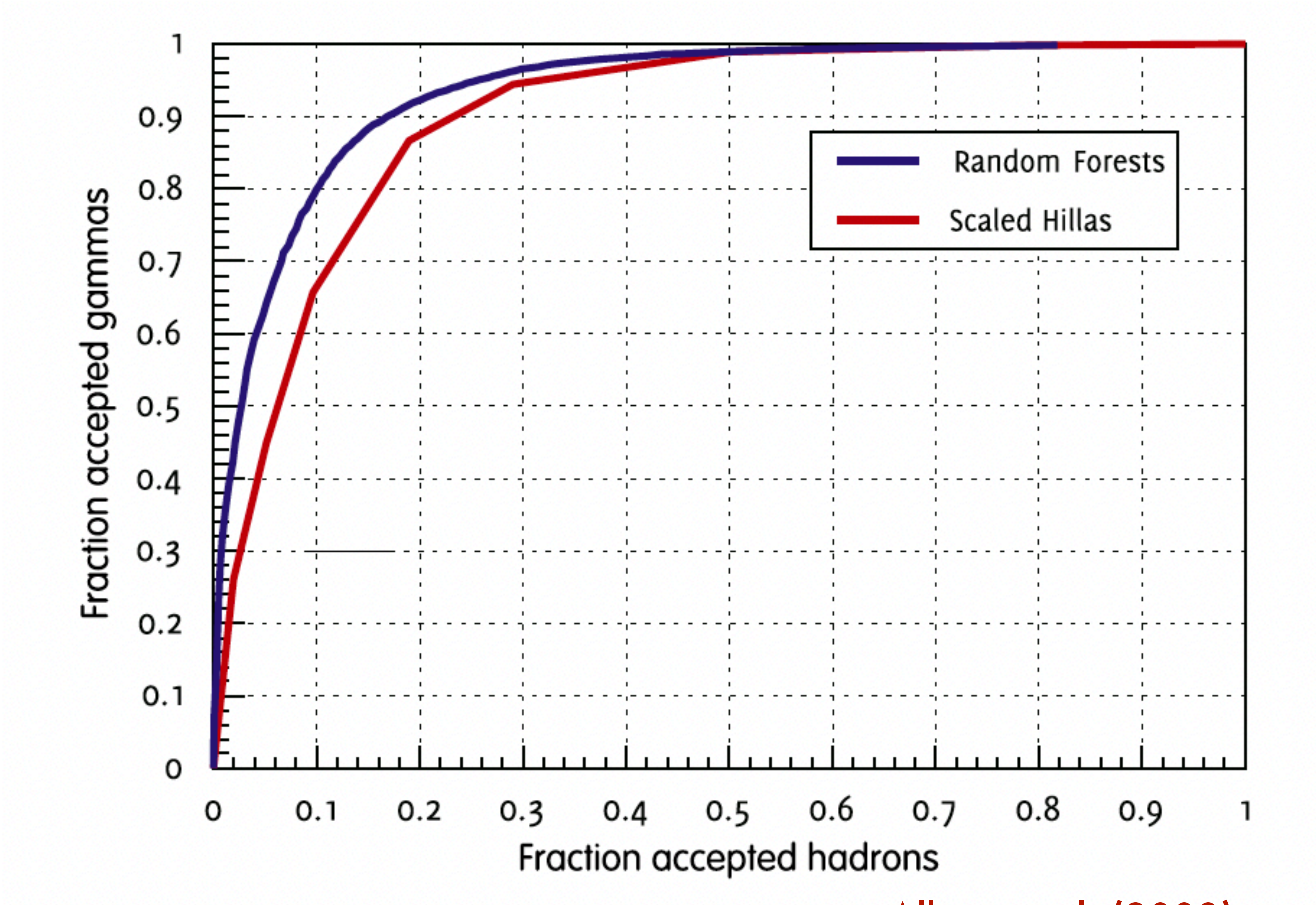



Credits: A. Fernández Barral

In both mono/stereo cases: training aims at finding a relation between the disp (known for MCs) and a defined set of parameters

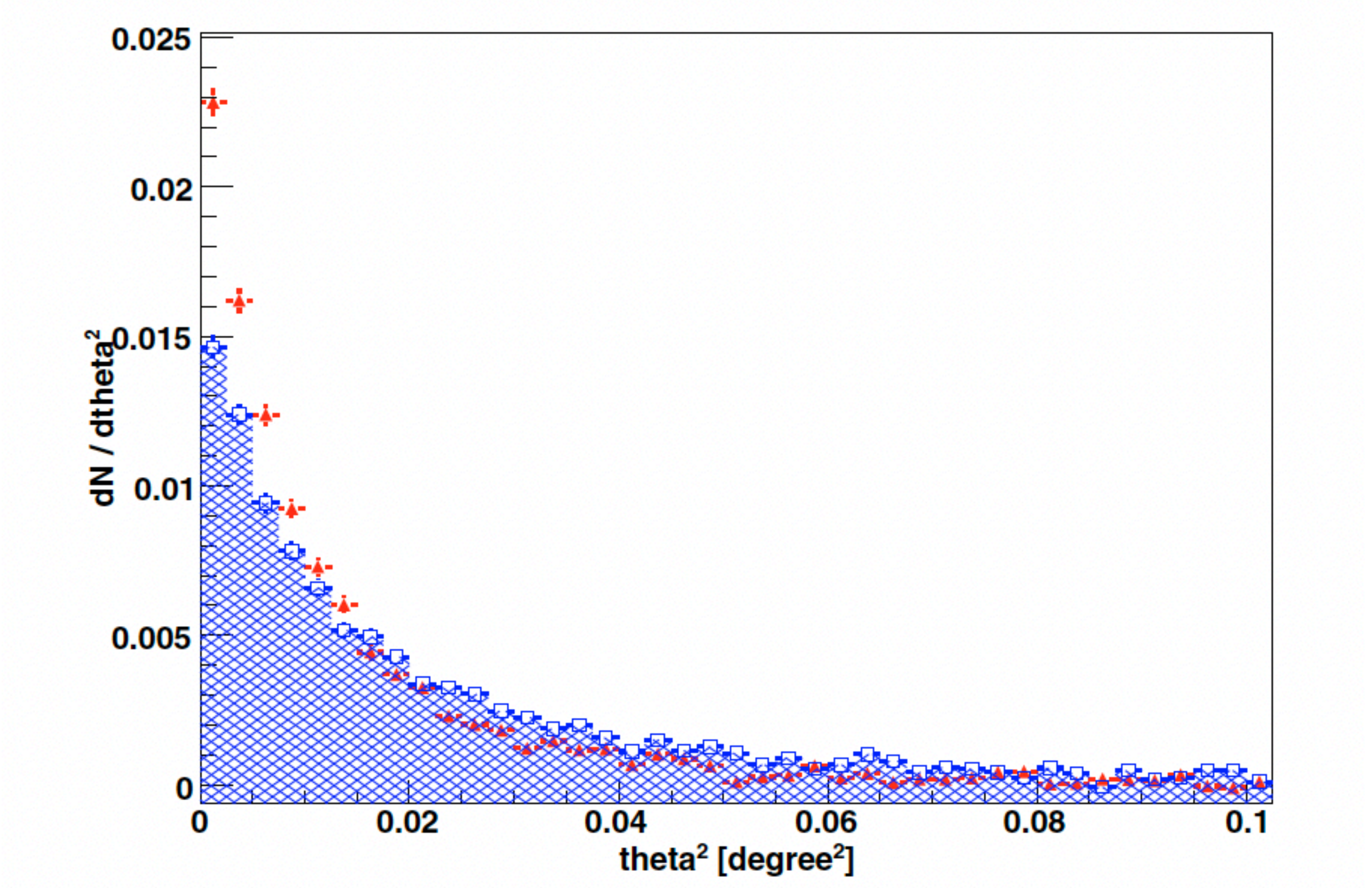
# Model comparison

## Hillas-based analysis vs RF



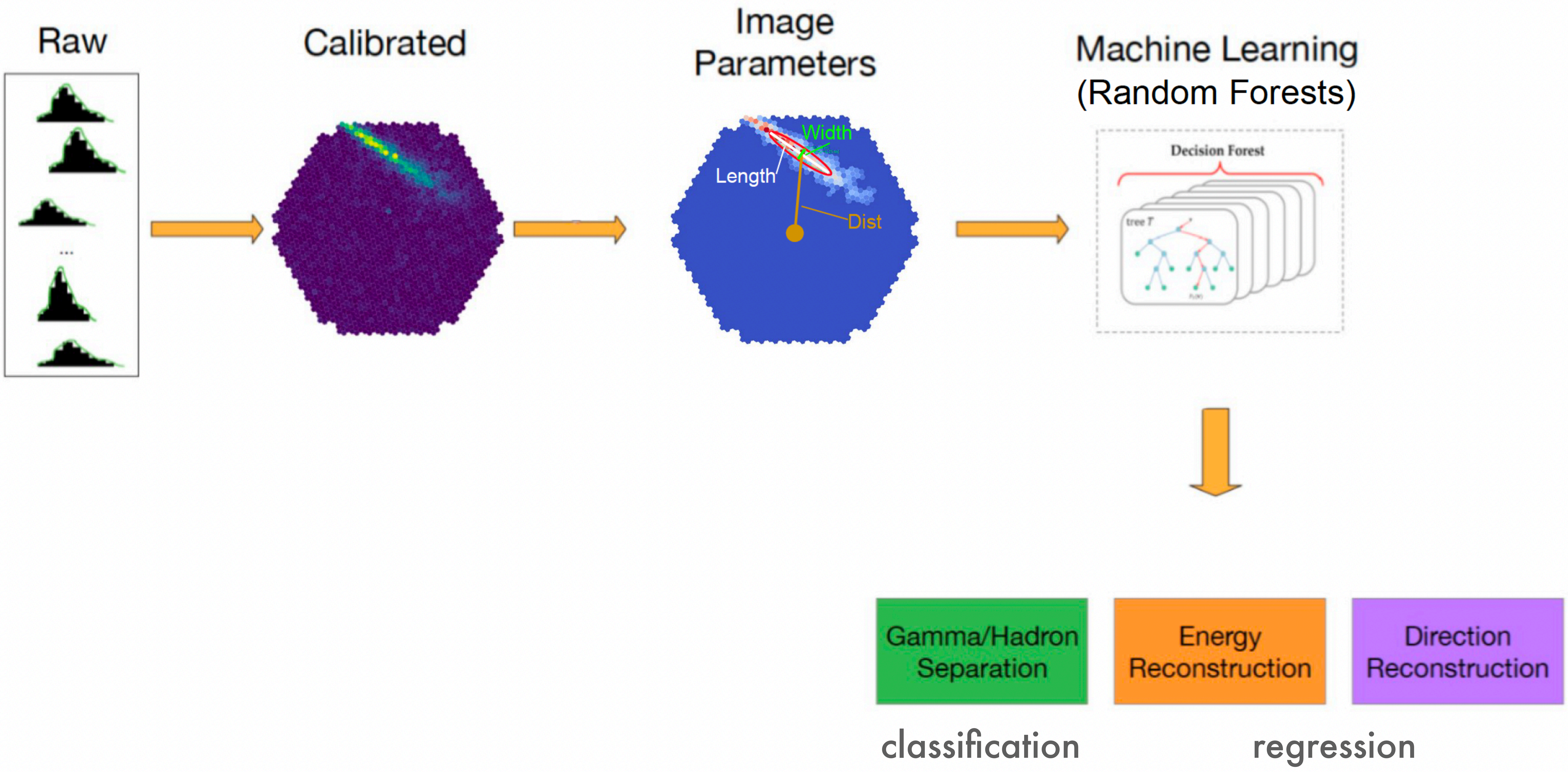

[Albert et al. (2008)]

## DISP vs DISP-RF



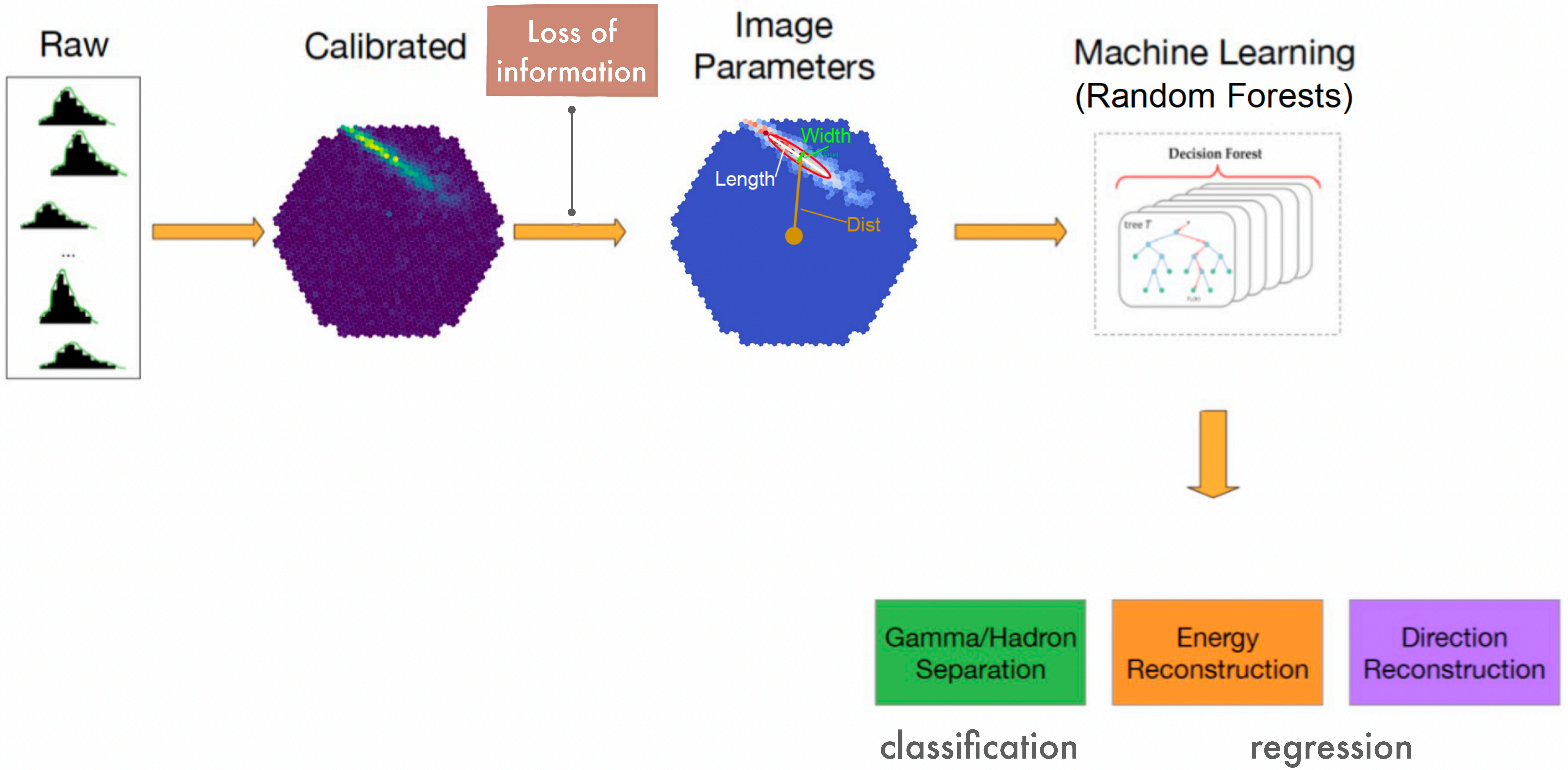

[Aleksić et al. (2010)]

# Standard data analysis chain




Credits: P. Grespan

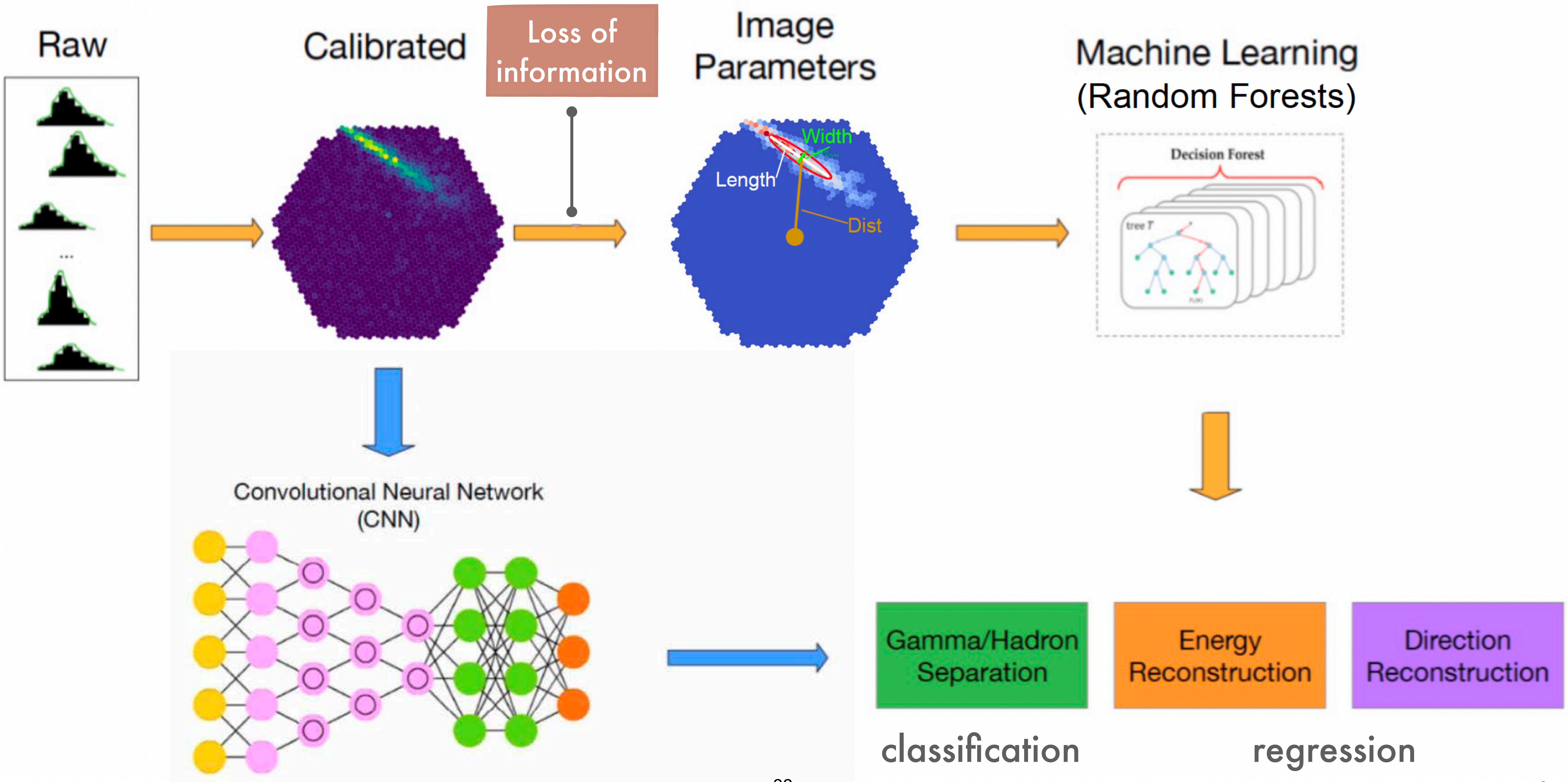
# Standard data analysis chain






Credits: P. Grespan

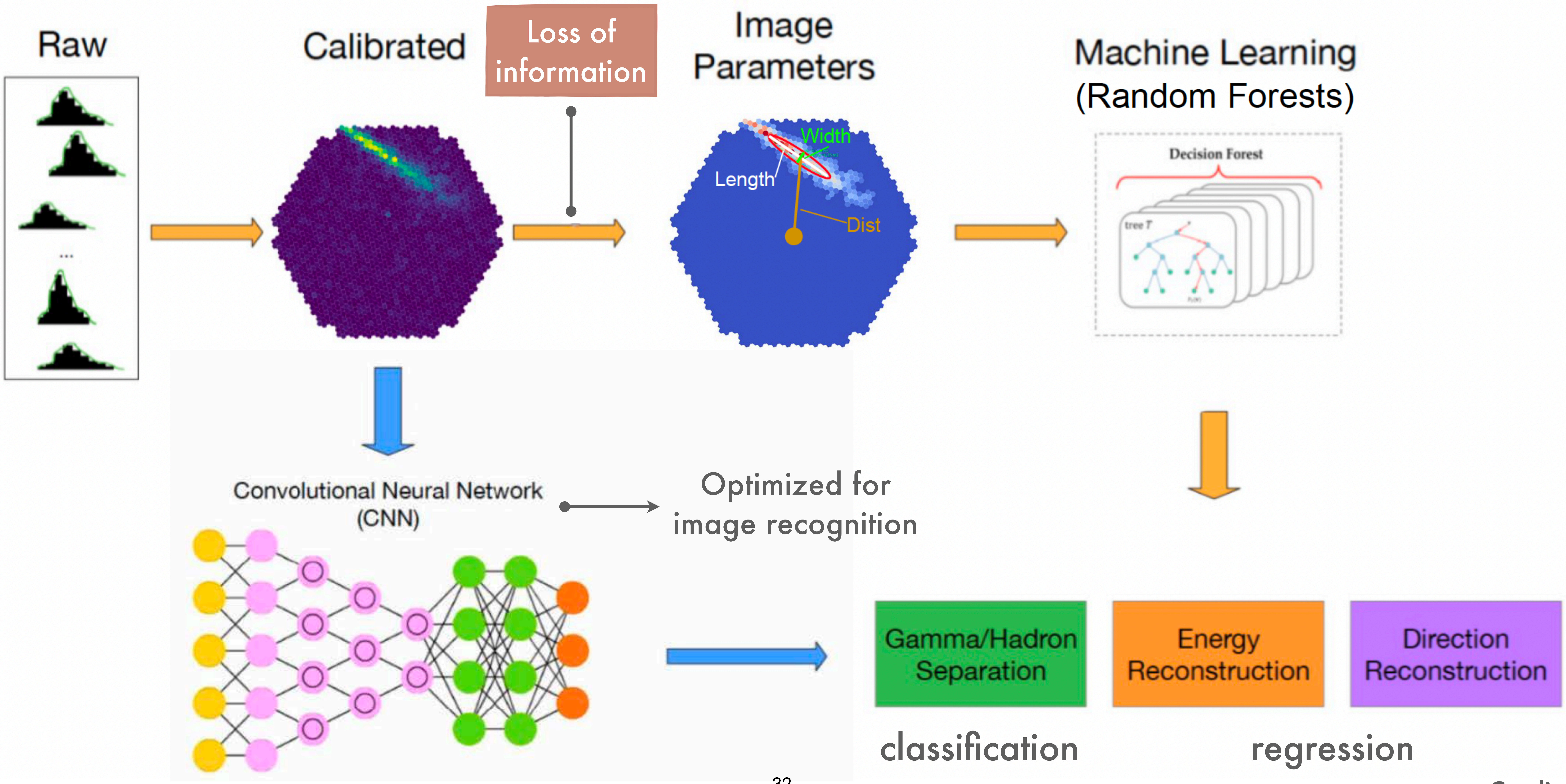
# Towards a Deep Learning approach






Credits: P. Grespan

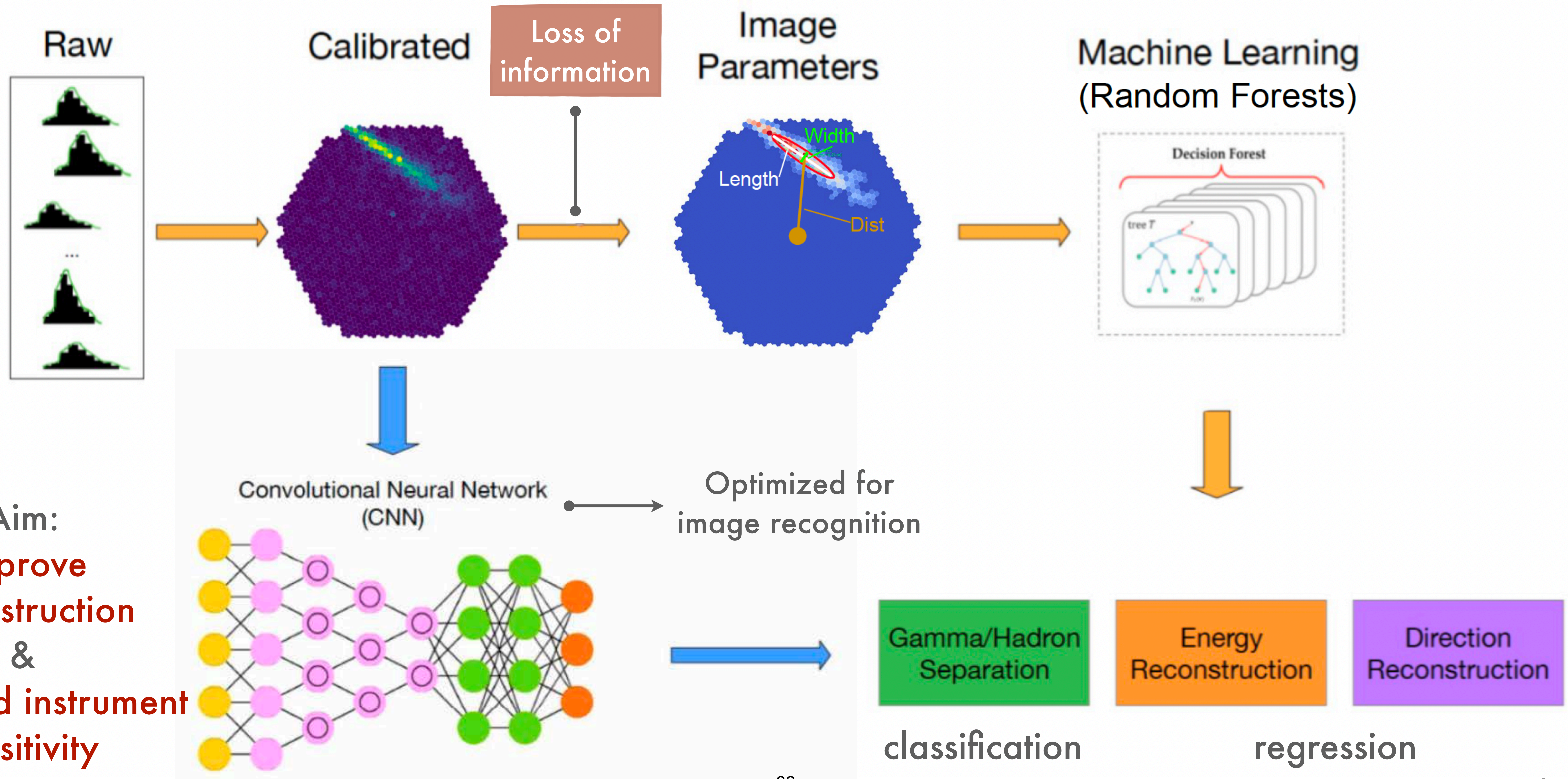
# Towards a Deep Learning approach






Credits: P. Grespan

# Towards a Deep Learning approach

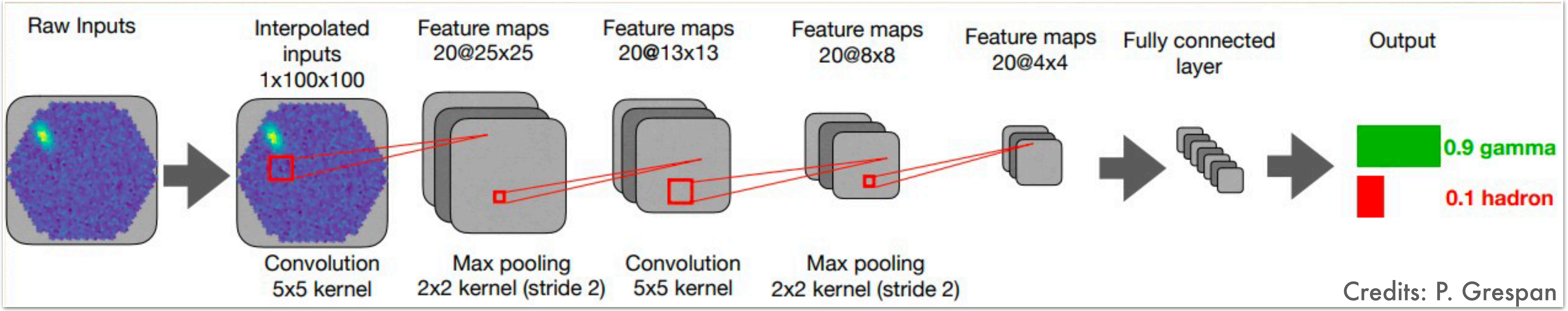



Aim:
improve reconstruction
&
increased instrument sensitivity



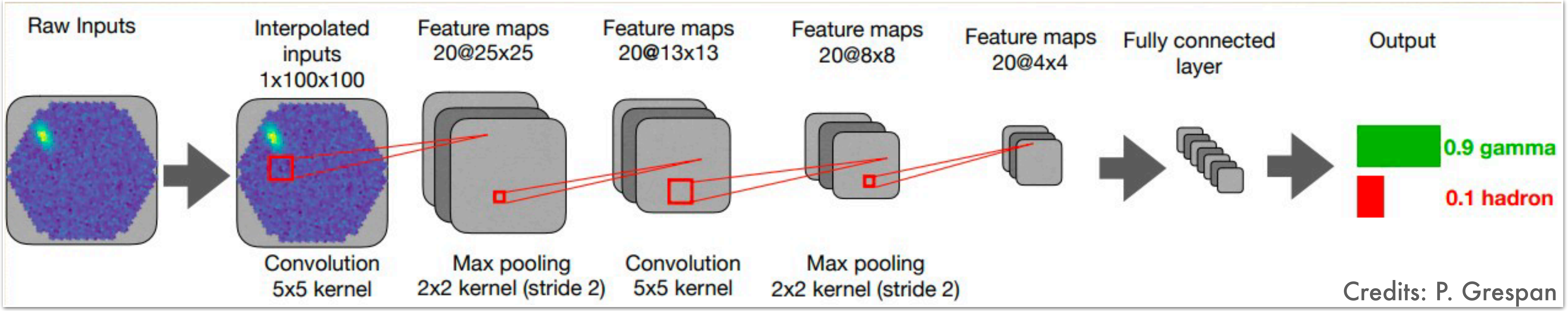
Credits: P. Grespan

# Convolutional Neural Networks




- Able to access spatial and temporal image information
- Able to identify relevant image features with unprecedented accuracy through the use of convolutions
- Thanks to the extracted features, it can make a prediction of the quantity of interest
- Performance checked by a loss function

# Convolutional Neural Networks




*In IACTs event reconstruction...*

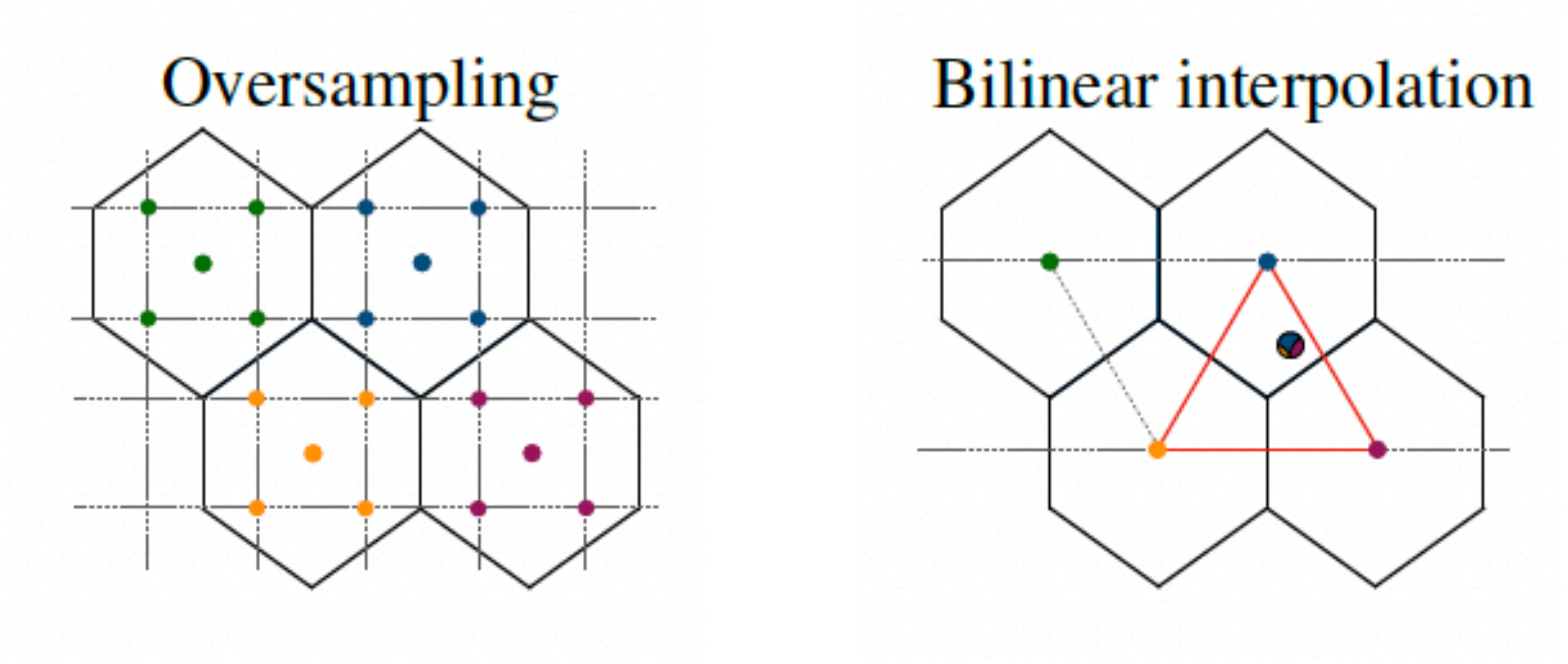
- Application of CNN on the non-parametrised images to enhance telescope sensitivity

*Risks and issues...*

- Need for MC hadrons: less reliable than MC γ in approximating real data
- Developed for squared pixels

# Solving the hexagonal pixel challenge

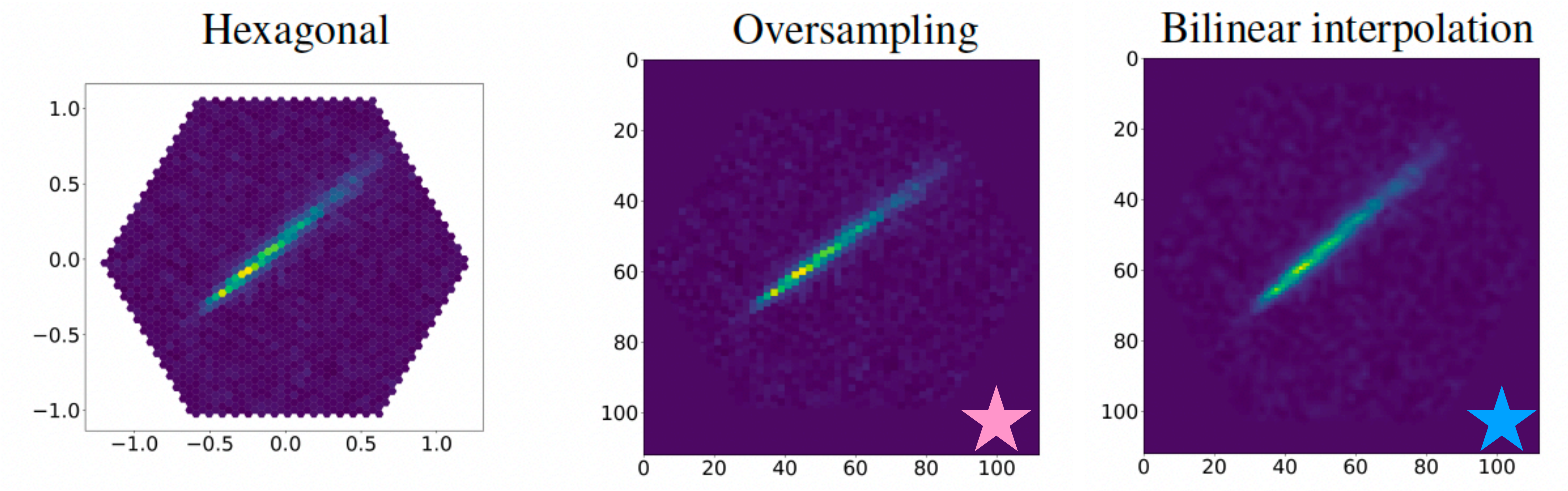
- IACT camera pixels are hexagonal
- Need for:
  - Mapping method turning them into cartesian lattice
  - Dedicated convolution implemented in the CNN to operate on hexagonal pixel organization




Nieto et al. (2019)

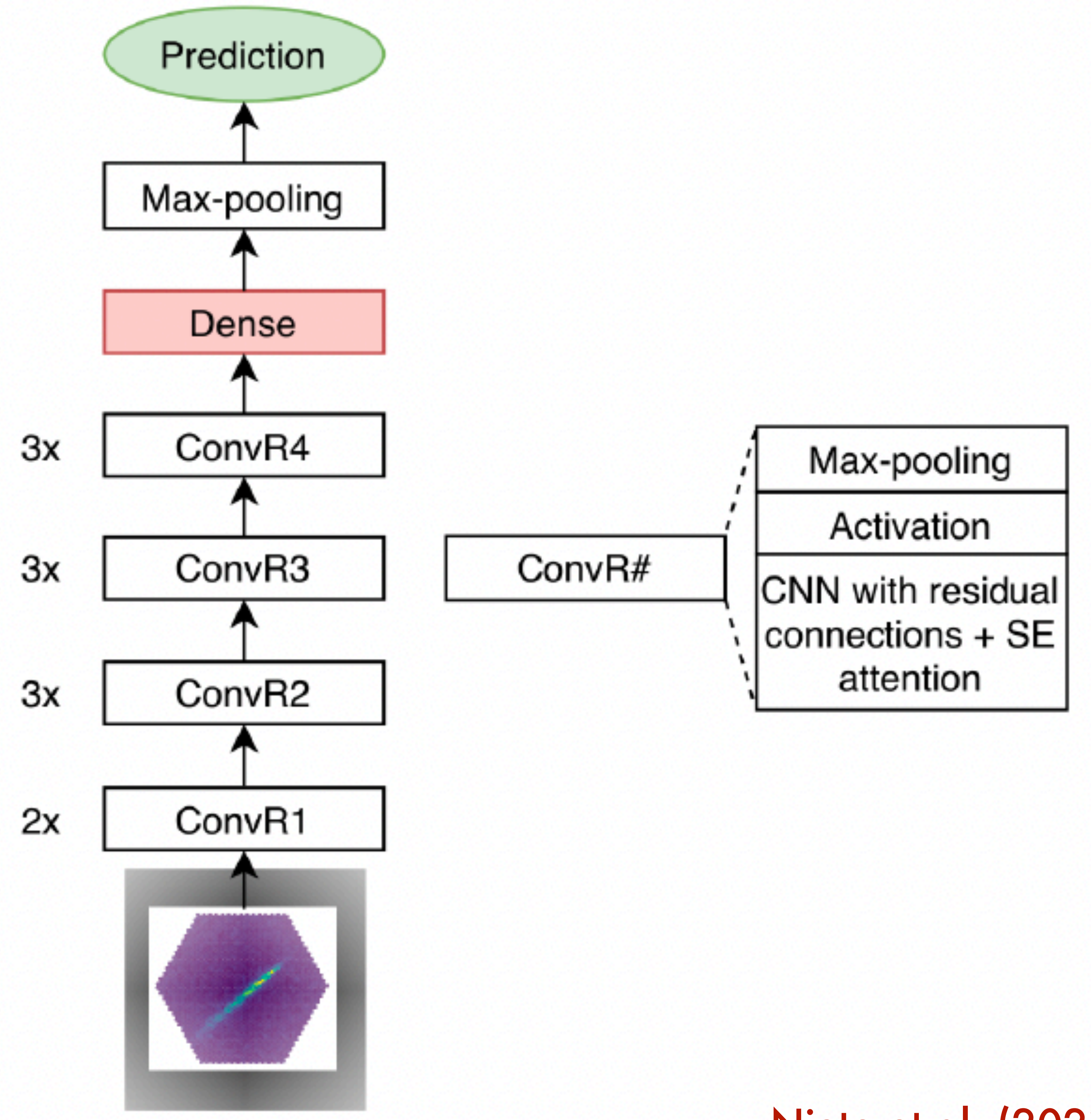
★ Preserves image charge

★ Preserves angles and distances



# Single telescope full event reconstruction

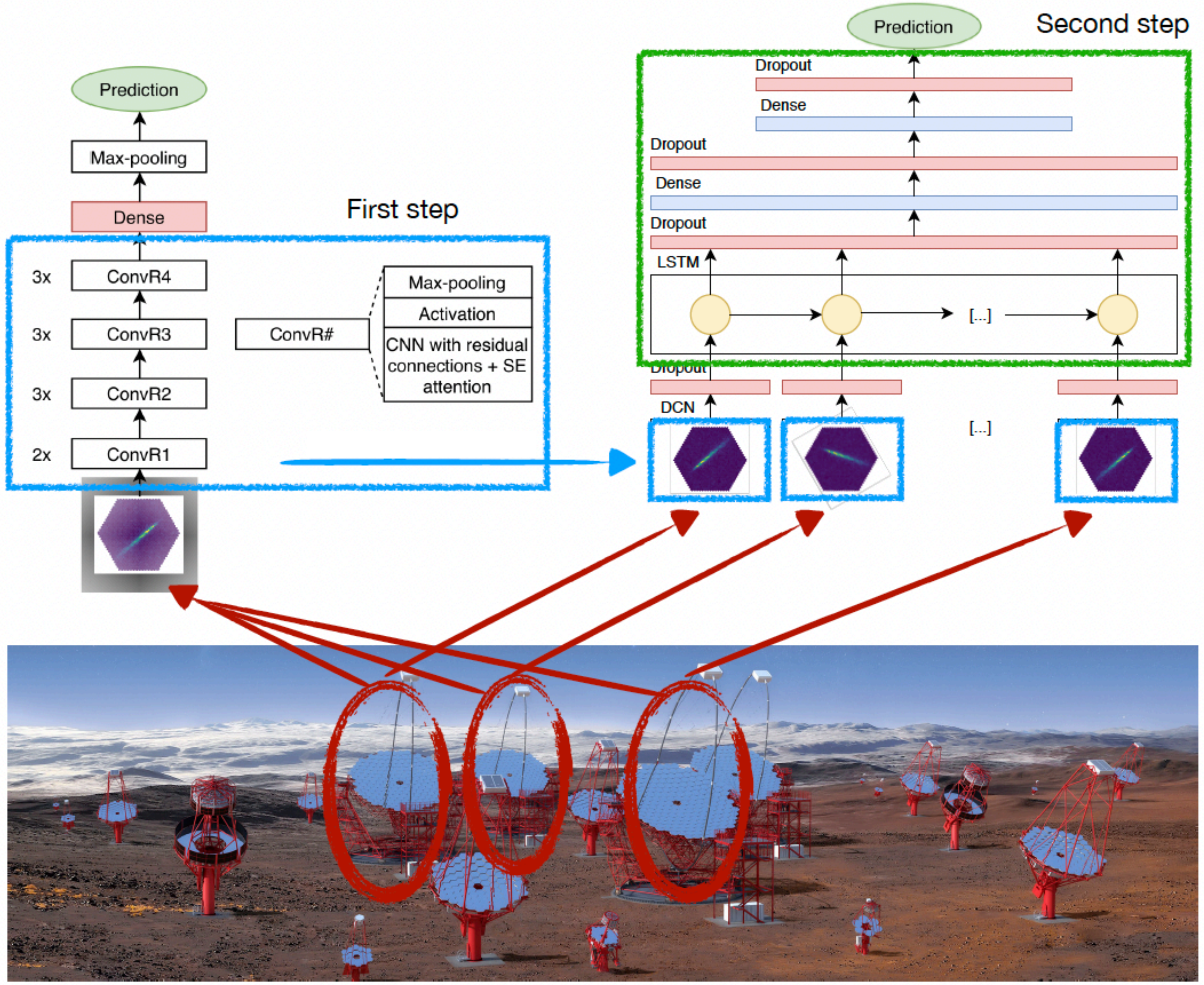
- One of main efforts: CTLearn framework
  - Open source python package for IACT event reconstruction with Deep Learning
  - Pixel mapping into cartesian lattices
  - Model based on a 33-layers CNN with residual connections
  - One model of each reconstruction task
  - Both mono and stereo analyses allowed



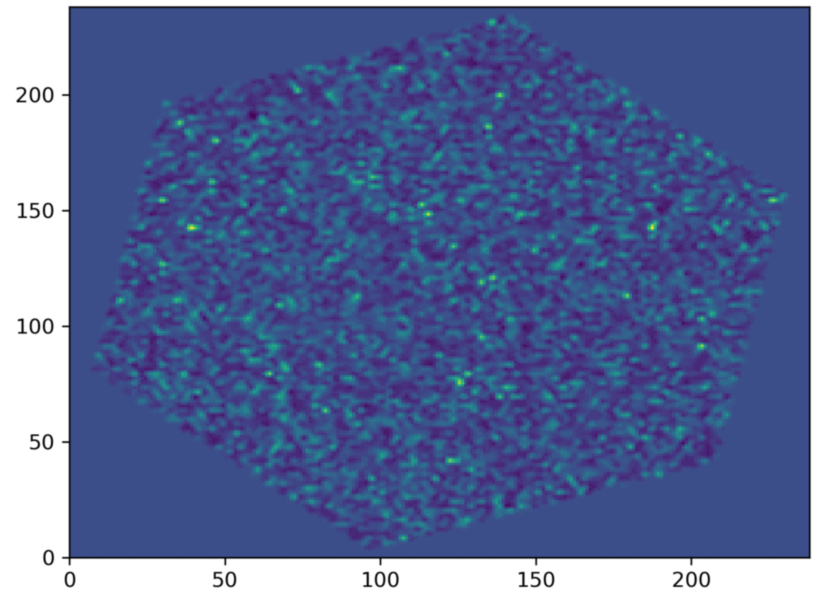

Nieto et al. (2021)

# Stereo full event reconstruction

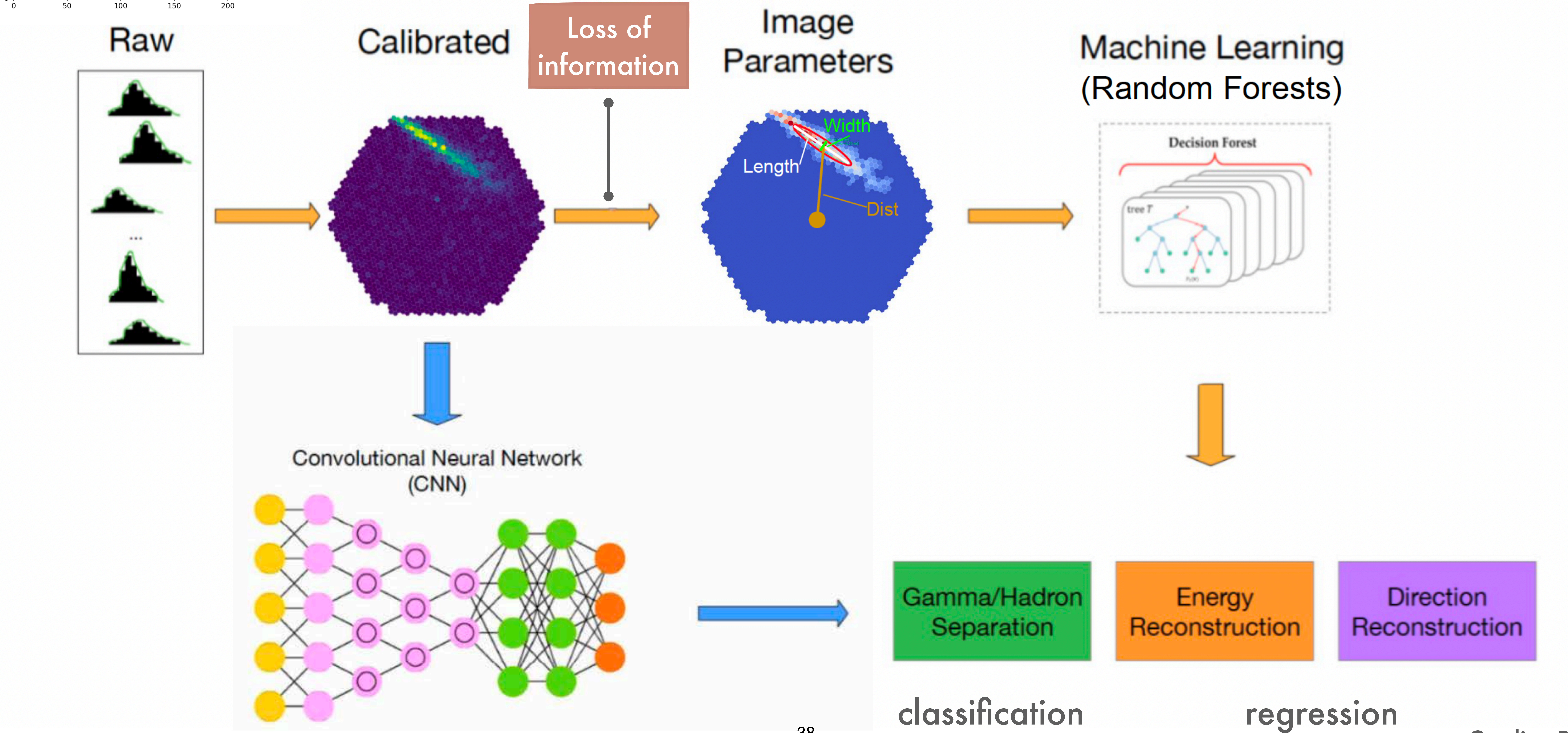
- Output of single-telescope network as input for stereo network processing multiple images in parallel
- Stereo network size adjusted based on number of telescopes triggered by the event




Miener et al. (2021)



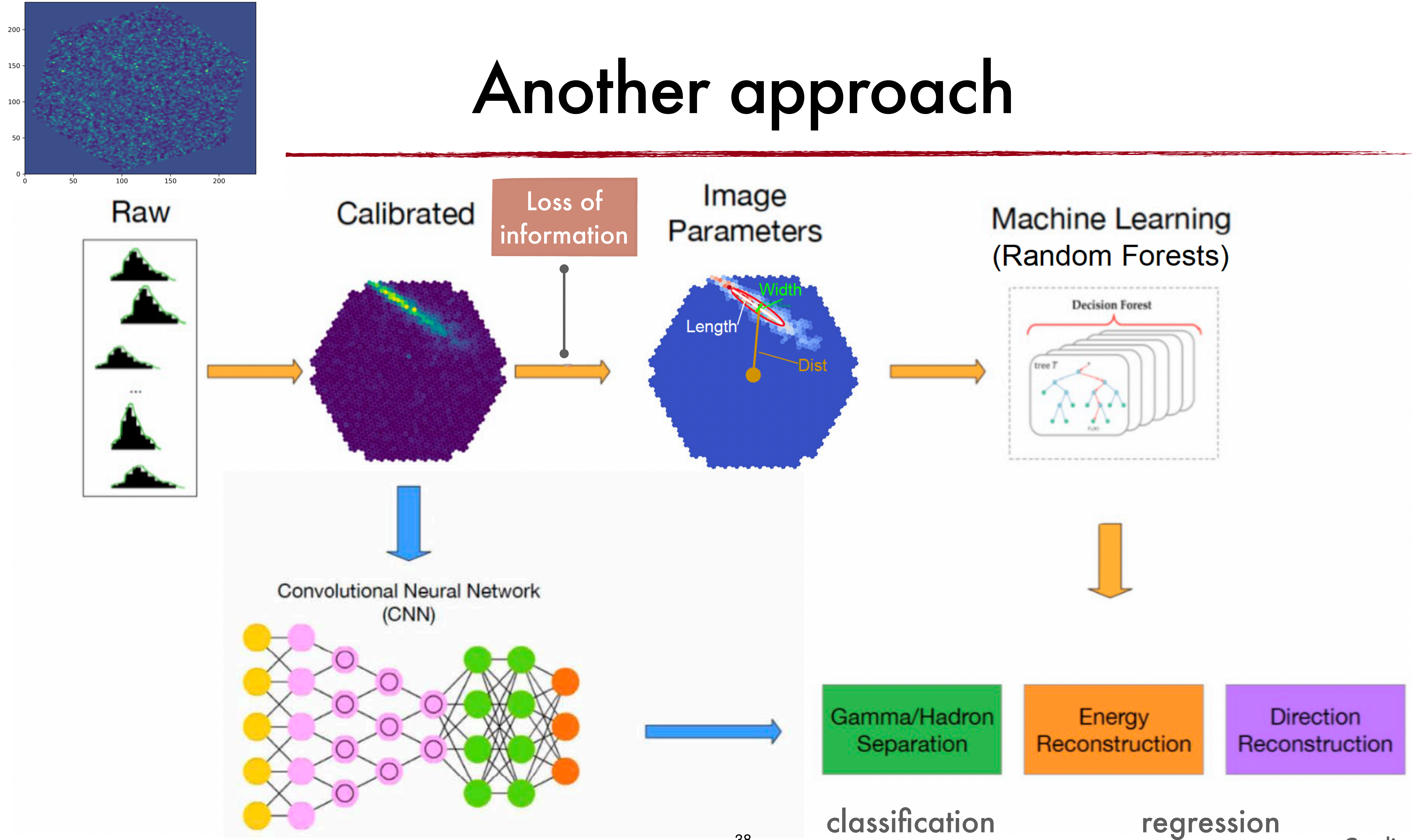
# Another approach






Credits: P. Grespan

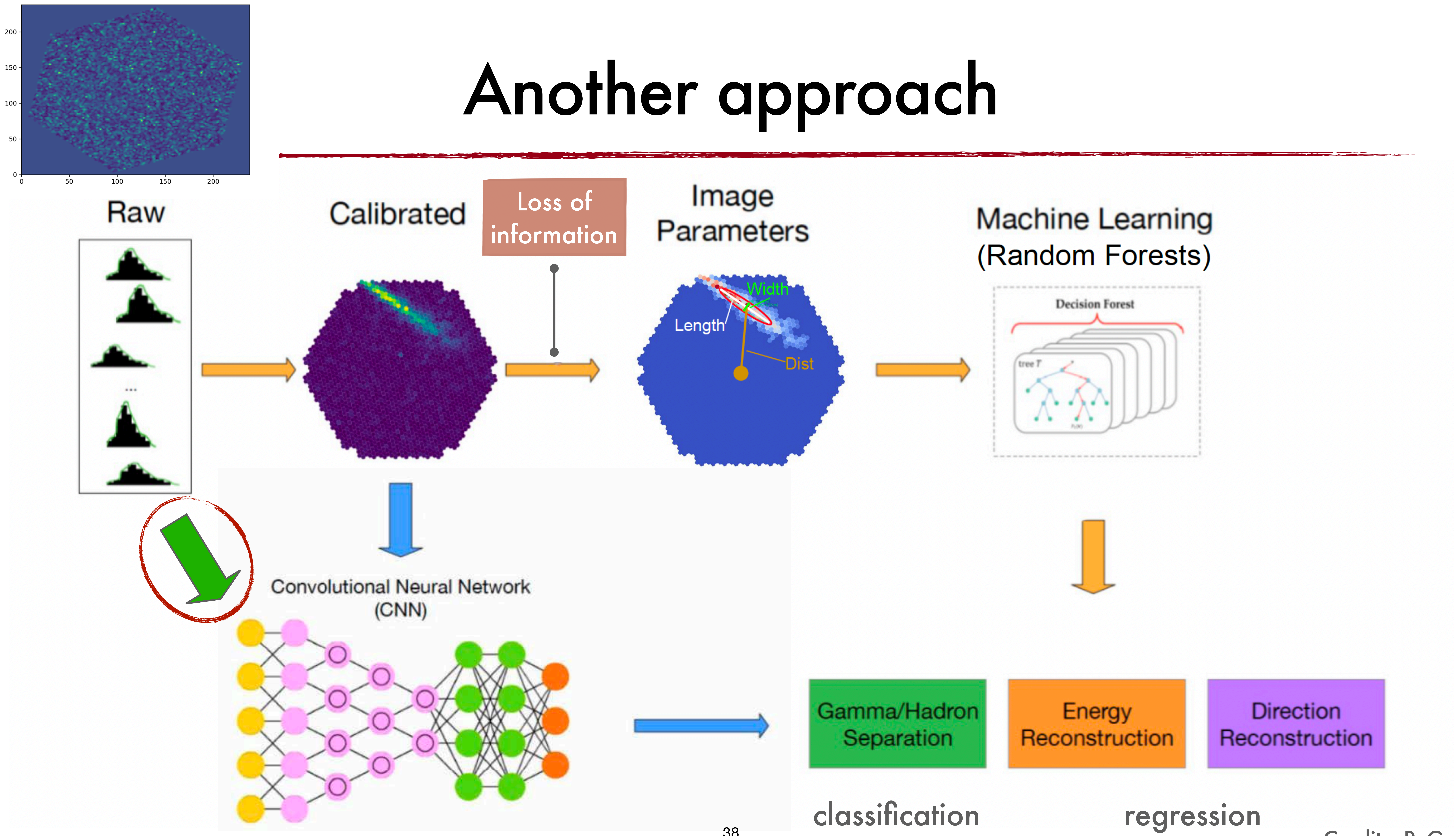
# Another approach






Credits: P. Grespan

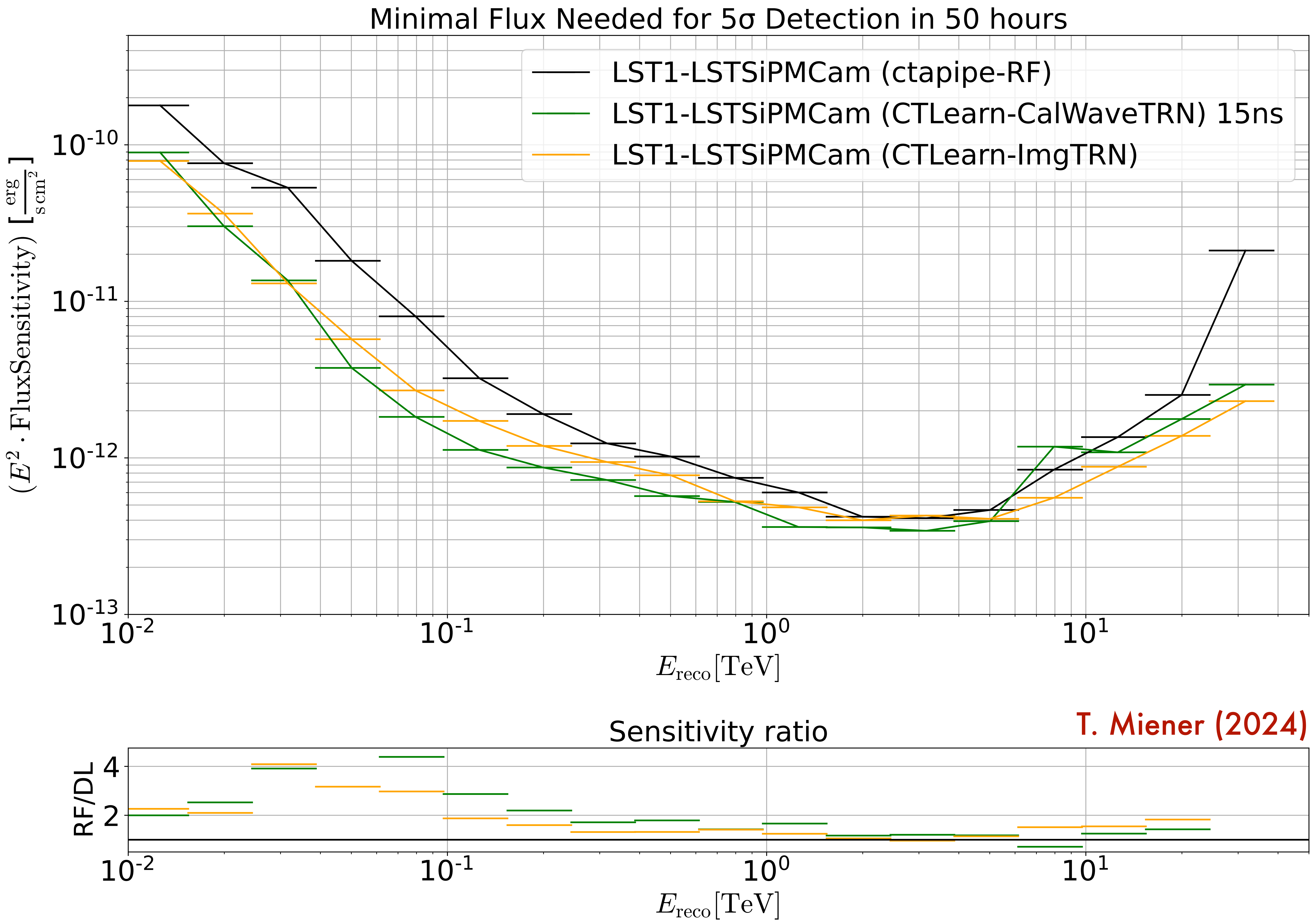
# Another approach






Credits: P. Grespan

# Results



Minimal Flux Needed for 5σ Detection in 50 hours
LST1-LSTSiPMCam (ctapipe-RF)
LST1-LSTSiPMCam (CTLearn-CalWaveTRN) 15ns
LST1-LSTSiPMCam (CTLearn-ImgTRN)
$(E^2 \cdot \mathrm{FluxSensitivity})\ [\frac{\mathrm{erg}}{\mathrm{s\,cm^2}}]$
$10^{-10}$
$10^{-11}$
$10^{-12}$
$10^{-13}$
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$E_{\mathrm{reco}}[\mathrm{TeV}]$
Sensitivity ratio
T. Miener (2024)
RF/DL
4
2
$10^{-2}$
$10^{-1}$
$10^{0}$
$10^{1}$
$E_{\mathrm{reco}}[\mathrm{TeV}]$

# Summary and future prospects

- VHE γ-ray astrophysics has a crucial role in exploring the most energetic phenomena in the universe
- Classical machine learning techniques as Random Forests represent a robust and reliable method for the analysis of VHE γ-ray data
- However, the pre-processing of the γ-ray images needed for the application of this techniques can lead to a loss of information on the original event
- In this context Deep Learning methods can be of help, as they are able to work directly on the raw images
- Stereo analysis with CNNs on both images and waveforms show promising results

# Thank you for your attention