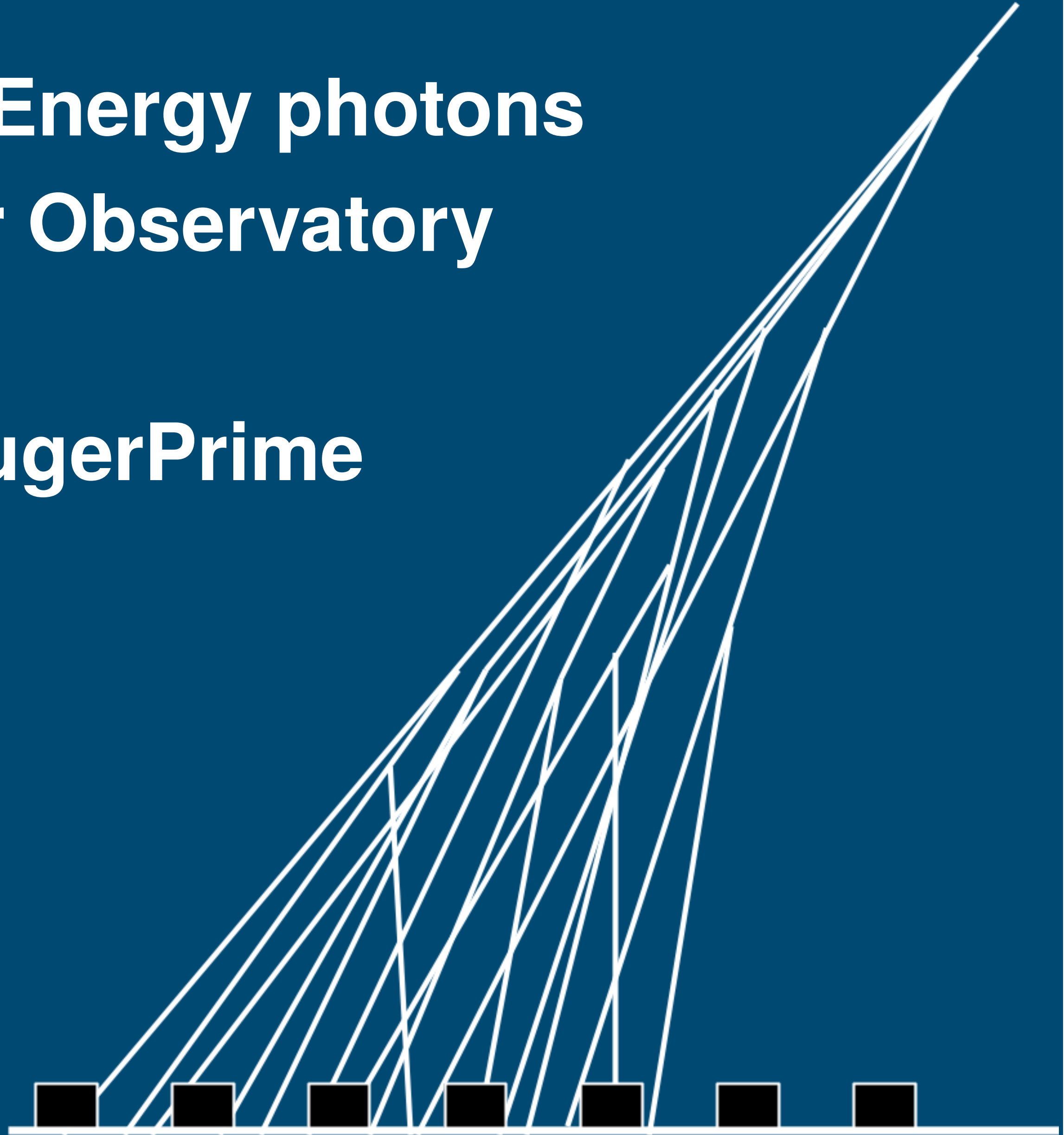
Search for **Ultra High Energy photons** at the **Pierre Auger Observatory** & contribution to **AugerPrime**



In the Auger group.

Thesis Supervisor : Corinne Berat

LPSC Grenoble

PIERRE AUGER OBSERVATORY

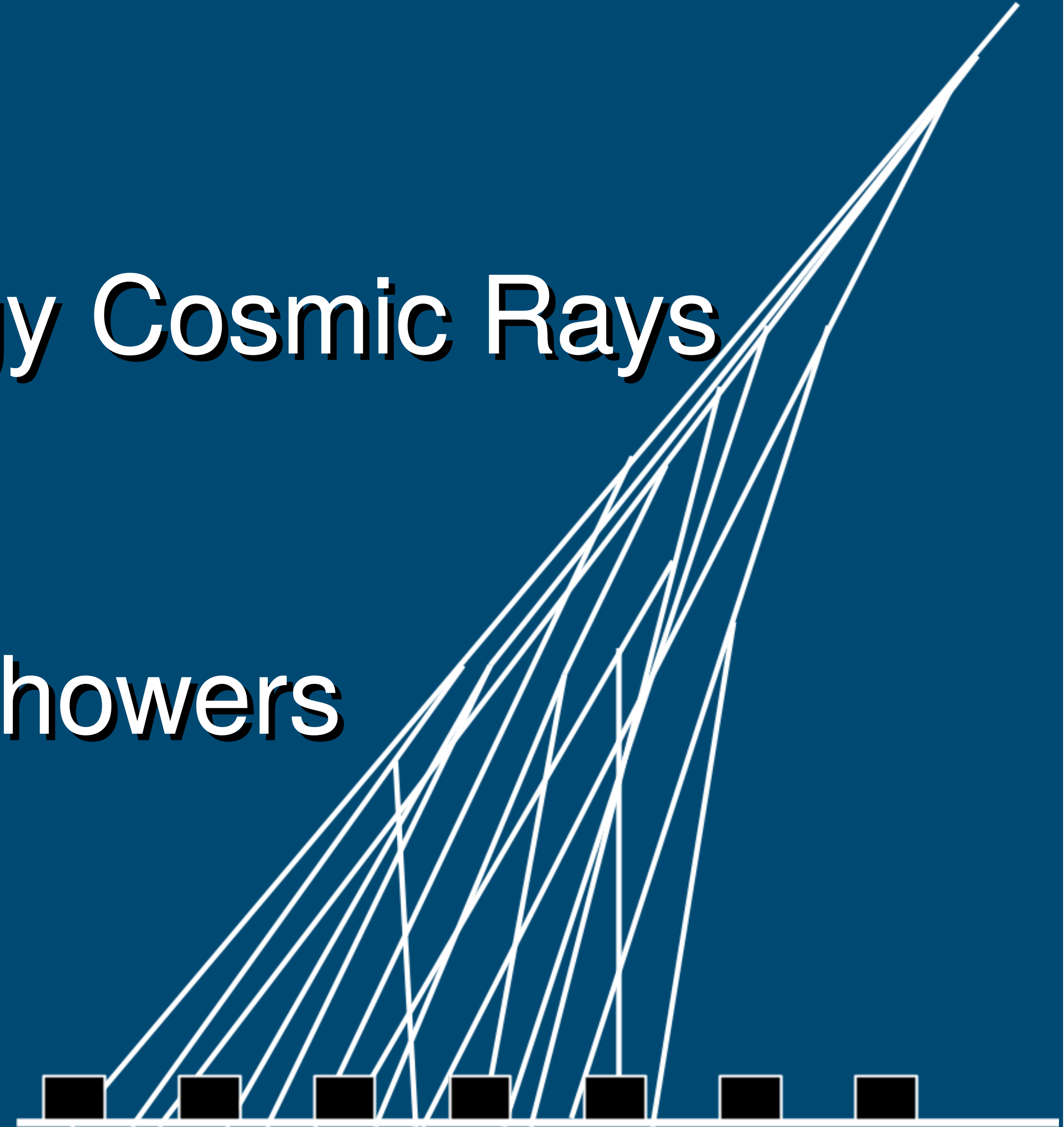
Outline

- Ultra High Energy Cosmic Rays & Extensive Air Showers
- The Pierre Auger Observatory & AugerPrime
- Scintillator Surface Detectors – Construction & Validation
- Search for Ultra-High Energy photon primaries – Motivations & Multi-Variate Analysis

Ultra-High Energy Cosmic Rays

&

Extensive Air Showers



Ultra High Energy Cosmic Rays - Spectrum

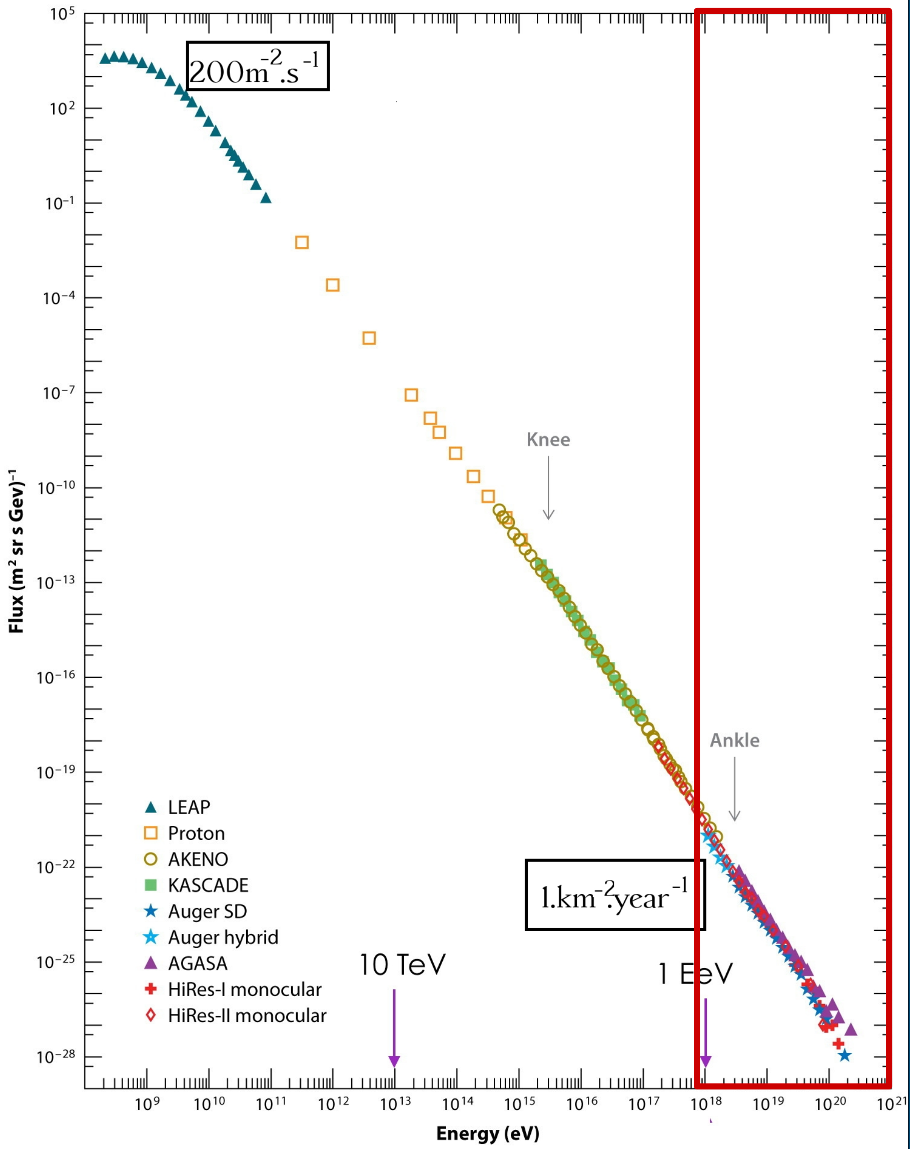
Ultra **H**igh **E**nergy **C**osmic **R**ays (**UHECRs**) :

- ultra high energies : E > 10^{18} eV

- very limited flux : < 1.km^{-2}.year^{-1}

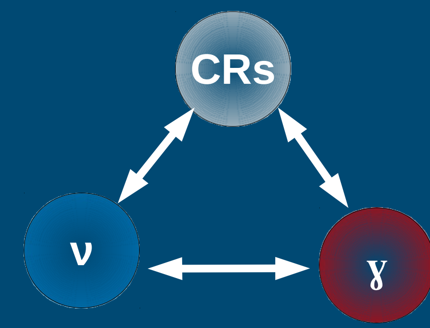
- features in the spectrum

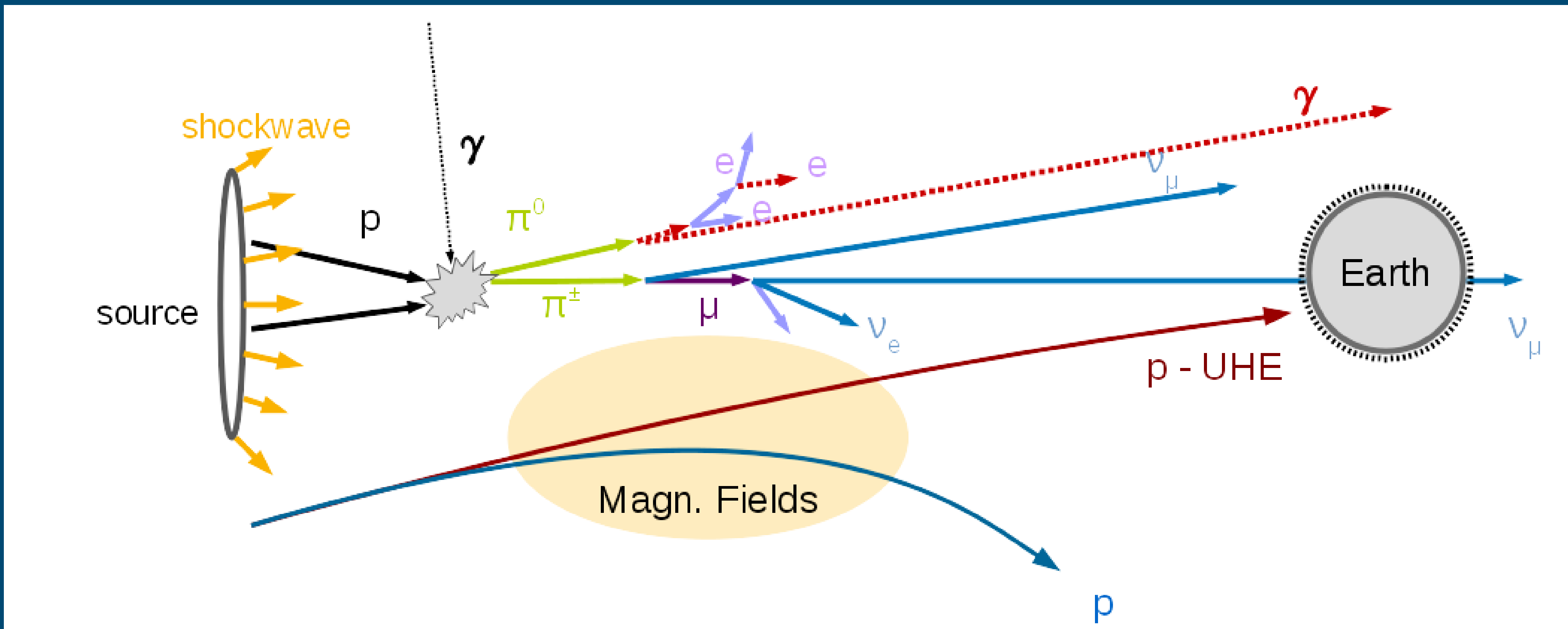
- nucleus from H to Fe & neutrals (n/γ/ν)



Beatty JJ, Westerhoff S. 2009.
Annu. Rev. Nucl. Part. Sci. 59:319–45

Acceleration & Propagation

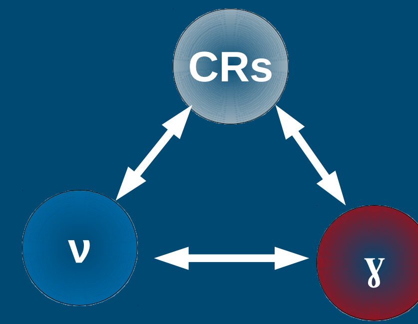




Multi-Messenger Astrophysics

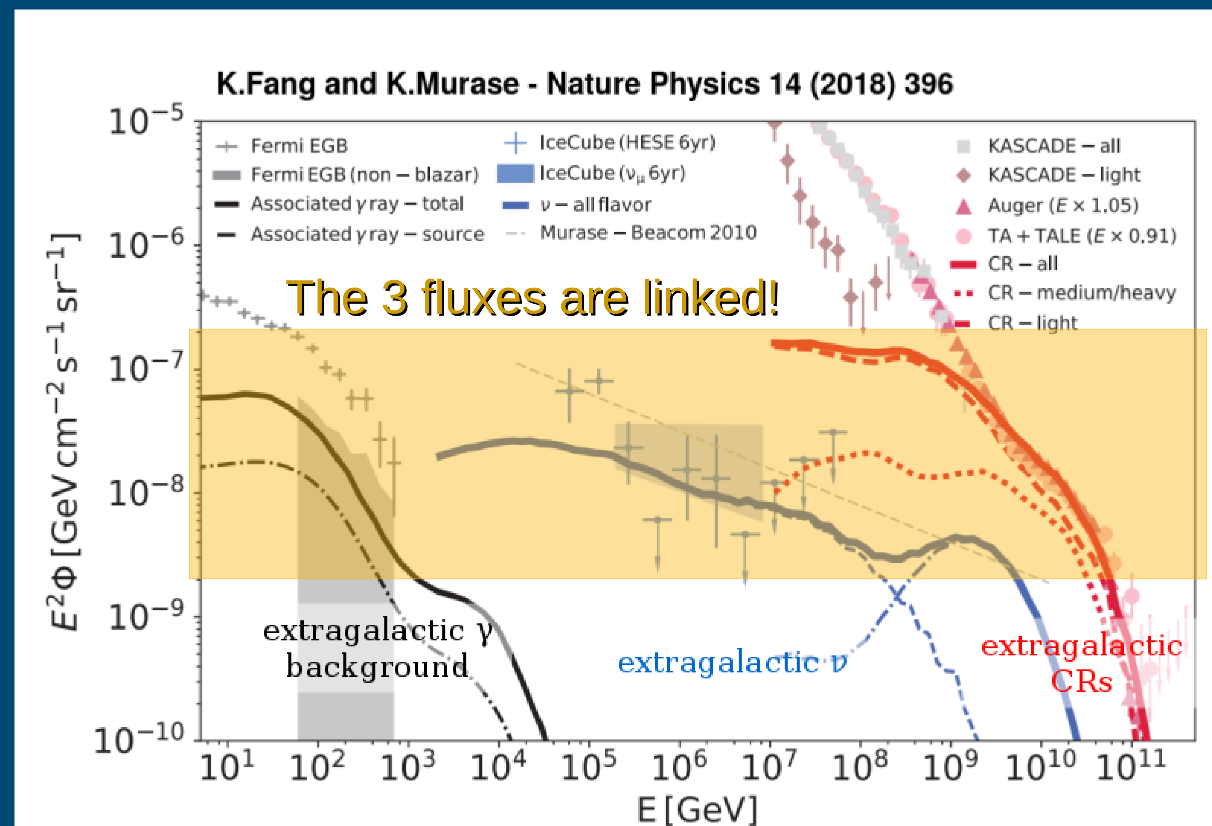
Multi-Messenger era of astrophysics :

- combine cosmic rays, gamma, neutrino and gravitational wave observations
- interplay between all these messengers



Complementarity between the observations

- Gamma-rays :
  +: straight line
  - : UHE horizon < 10 Mpc
- Neutrinos :
  +: straight line, no interaction
  - : isotropic diffuse background
- Cosmic-Rays :
  +: direct accelerator probe
  - : deflected in magn. field



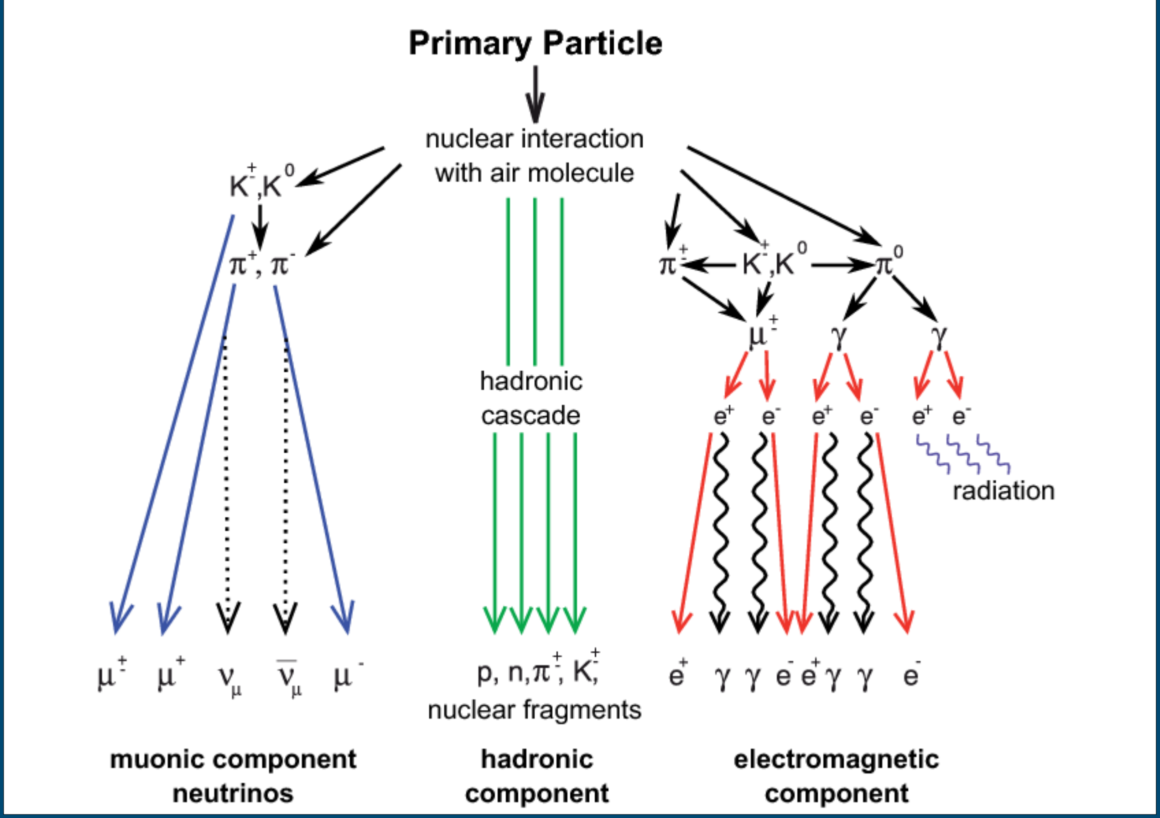
Extensive Air Shower

UHECRs interact with Earth's atmosphere :
⟶ generates an **E**xtensive **A**ir **S**hower (**EAS**)

3 main components :
- muonic component (~4%)
- hadronic component (~1%)
- electromagnetic component (95%)

At ground : ~5.10^{10} particles

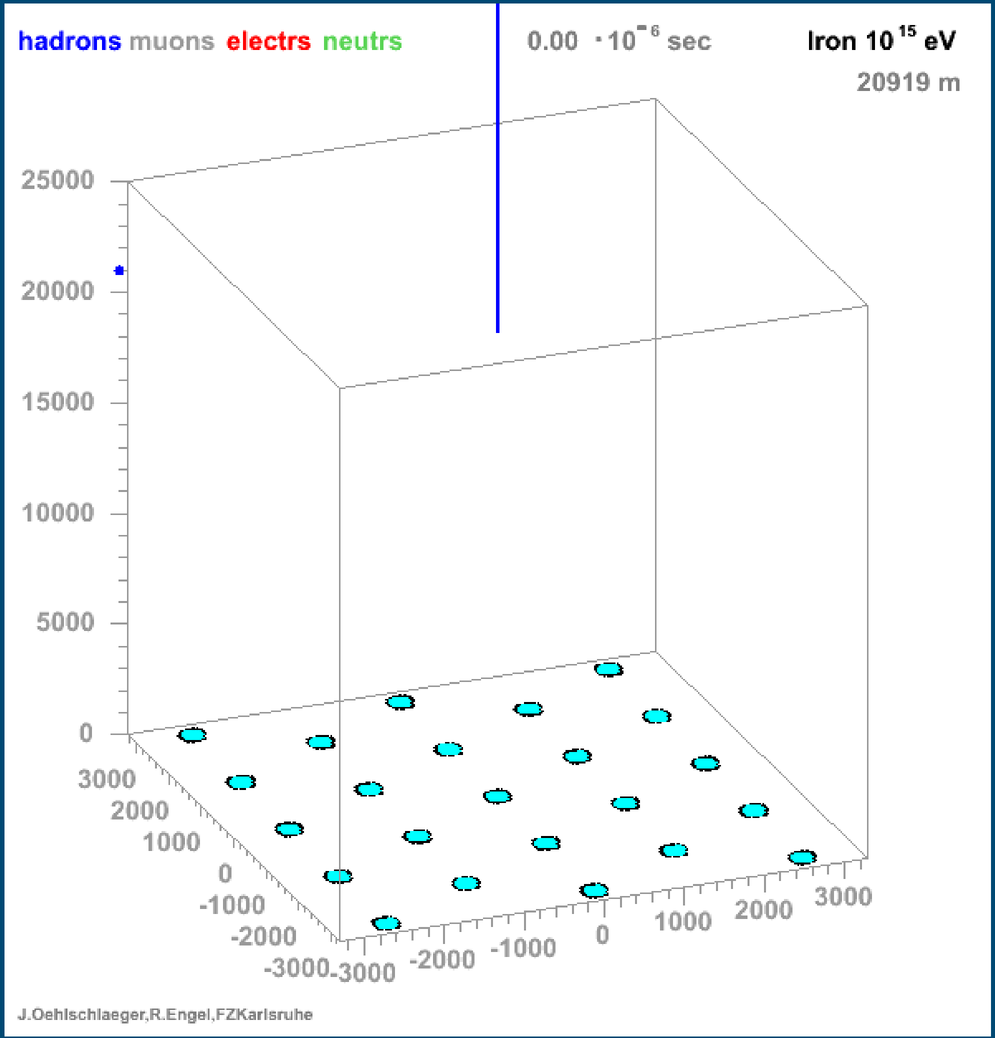
(estimation for a 10^{19}eV p-shower)



Extensive Air Shower

UHECRs interact with Earth's atmosphere :
⟶ generates an **E**xtensive **A**ir **S**hower (**EAS**)

3 main components :
- muonic component
- hadronic component
- electromagnetic component



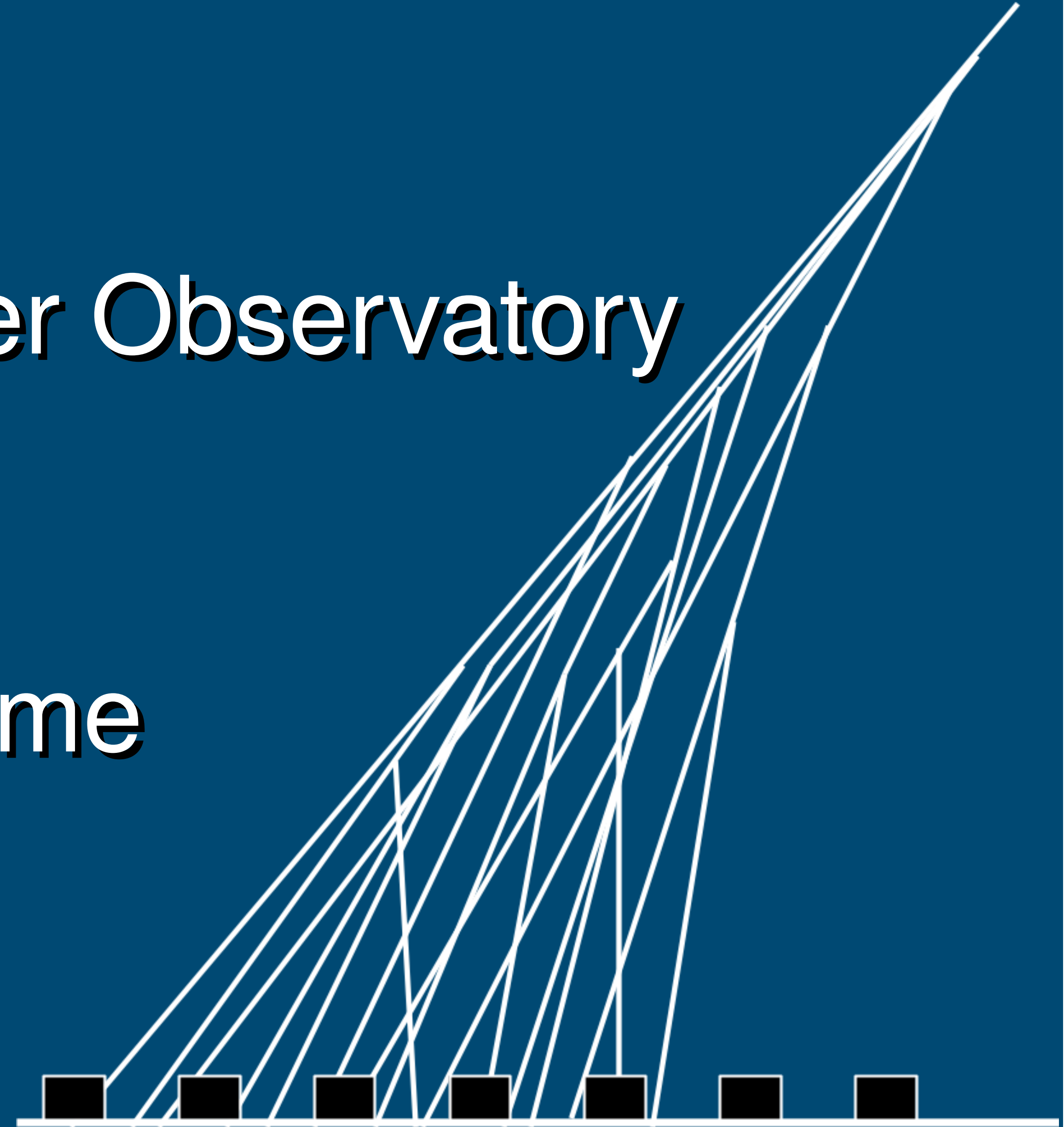
Advantages of the EAS :
- large footprint (up to 15 km)
- multiple observations possibles
- UHE hadronic physics laboratory...

Disadvantages of the EAS :
- indirect information on primary's energy
- inhomogeneous calorimeter
- ...model dependent

The Pierre Auger Observatory

&

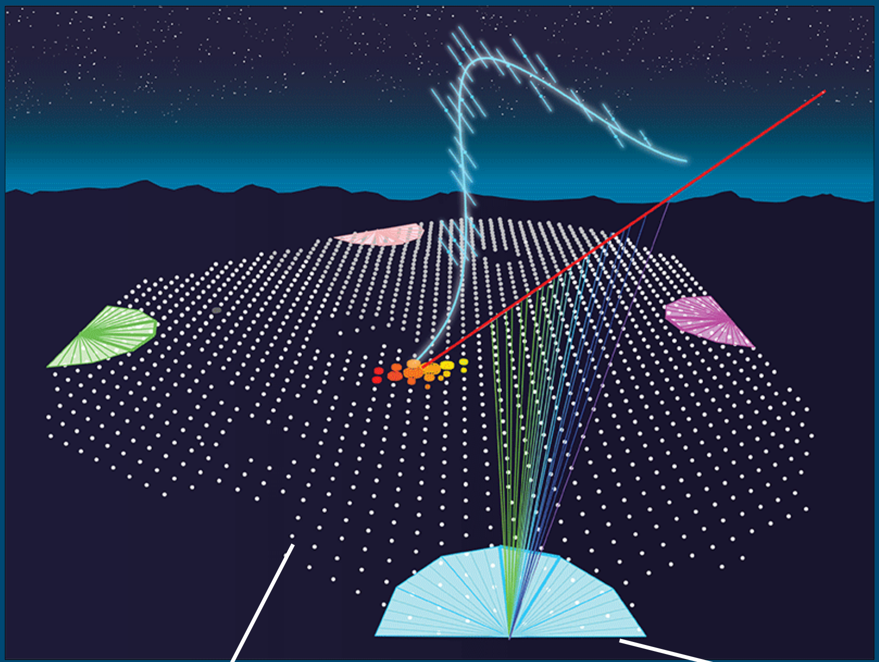
AugerPrime



LPSC Grenoble

PIERRE AUGER OBSERVATORY

The Pierre Auger Observatory



The Pierre Auger Observatory (PAO) :

- Officially completed in 2008
  - started taking data in 2004
- 400 scientists from 18 countries
- Location : pampa near Malargüe, Argentina
- Altitude (mean) : 1400 m above sea-level
- Surface (SD) : 3000 km^2

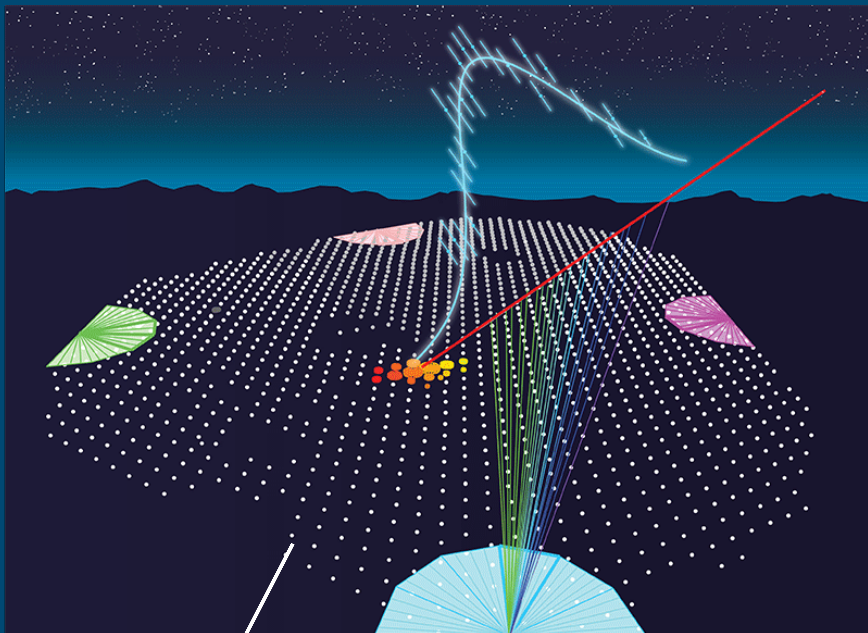
Fluorescence Detector (FD) :

- 24 telescopes located in 4 buildings.
- Overlooking the atmosphere above the array
- Duty-Cycle : 14%

Surface Detector (SD) :

- 1660 Water Cherenkov Detector
- Triangular spacing of 1.5 km
- Duty-Cycle : 100%

The Pierre Auger Observatory



The Pierre Auger Observatory (PAO) :

- Officially completed in 2008
  - started taking data in 2004
- 400 scientists from 18 countries
- Location : pampa near Malargüe, Argentina
- Altitude (mean) : 1400 m above sea-level
- Surface (SD) : 3000 km^2

Fluorescence **D**etector (**FD**) :

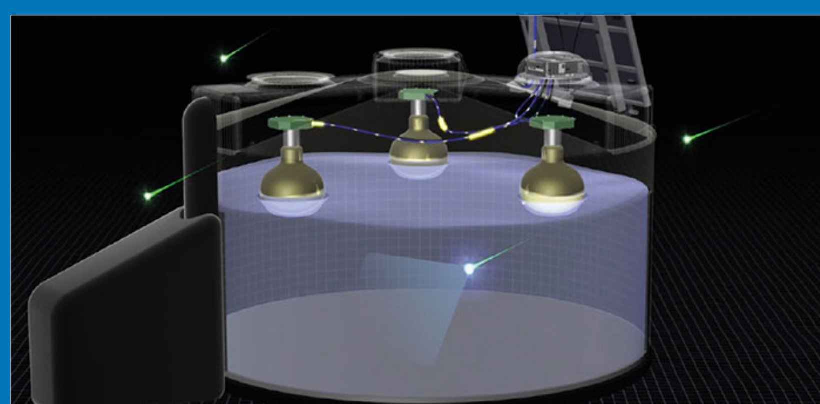
- 24 telescopes located in 4 buildings.
- Overlooking the atmosphere above the array
- Duty-Cycle : 14%

Water **C**herenkov **D**etector (**WCD**) :

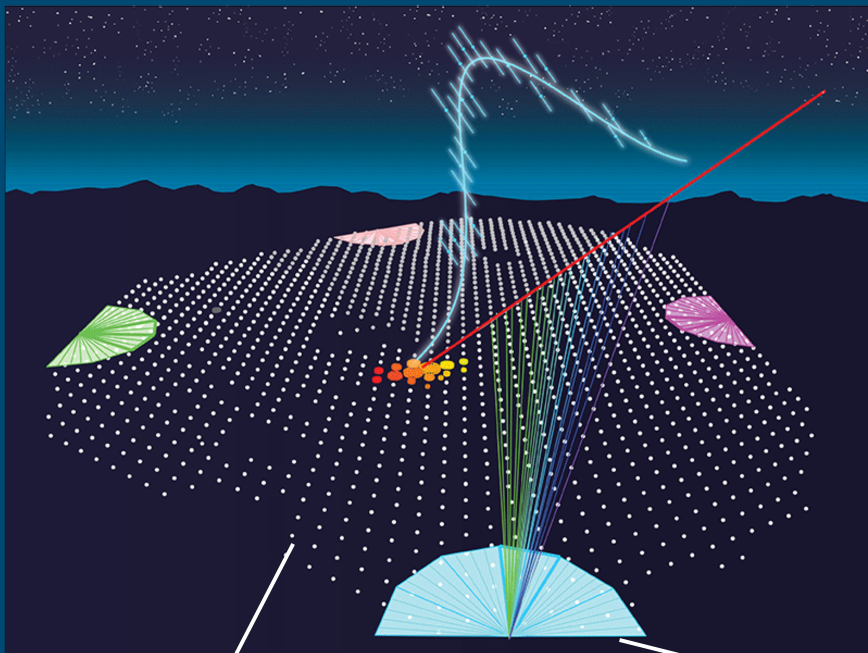
Secondary particles (highly relativistic) going through the detector produce Cherenkov light.

Collect timing and signal to reconstruct the showers.

Sensitive to $e^{\pm}$, γ, μ



The Pierre Auger Observatory



The Pierre Auger Observatory (PAO) :

- Officially completed in 2008
  - started taking data in 2004
- 400 scientists from 18 countries
- Location : pampa near Malargüe, Argentina
- Altitude (mean) : 1400 m above sea-level
- Surface (SD) : 3000 km^2

Surface **D**etector **(SD)** :

- 1660 Water Cherenkov Detector
- Triangular spacing of 1.5 km
- Duty-Cycle : 100%

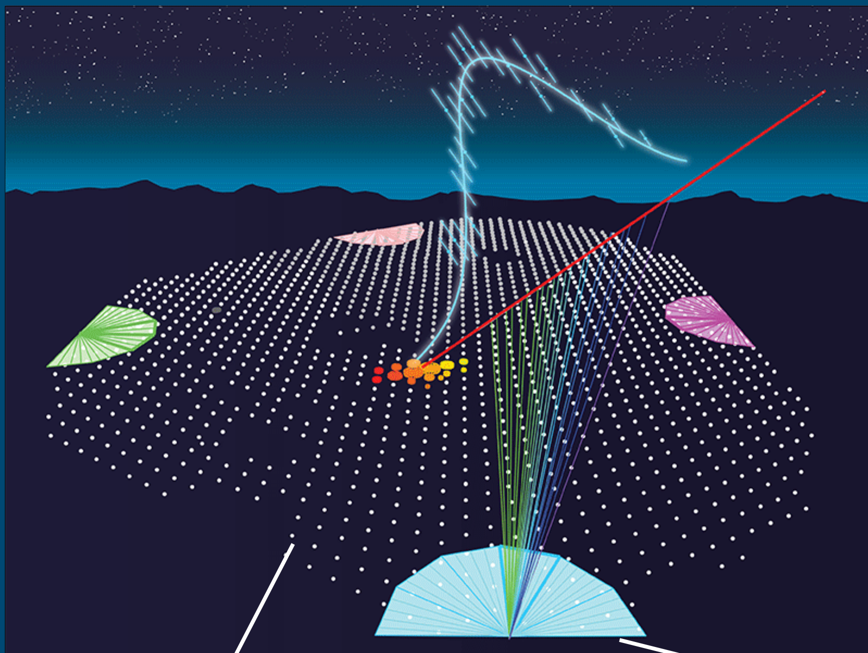
Fluorescence **T**elescope :

An EAS excite the nitrogen molecules in the atmosphere → Fluorescence Emission

Telescopes collect this UV light

Direct measurement of the shower energy

The Pierre Auger Observatory



The Pierre Auger Observatory (PAO) :

- Officially completed in 2008
  - started taking data in 2004
- 400 scientists from 18 countries
- Location : pampa near Malargüe, Argentina
- Altitude (mean) : 1400 m above sea-level
- Surface (SD) : 3000 km^2

Surface **D**etector (**SD**) :

- 1660 Water Cherenkov Detector
- Triangular spacing of 1.5 km
- Duty-Cycle : 100%

Fluorescence **D**etector (**FD**) :

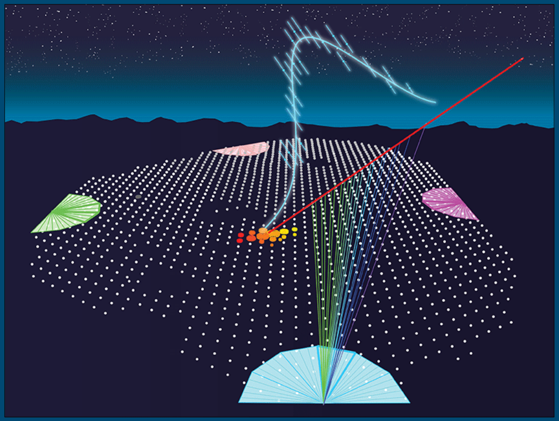
- 24 telescopes located in 4 buildings.
- Overlooking the atmosphere above the array
- Duty-Cycle : 14%

Hybrid Detection & Open questions

Hybrid Detection :

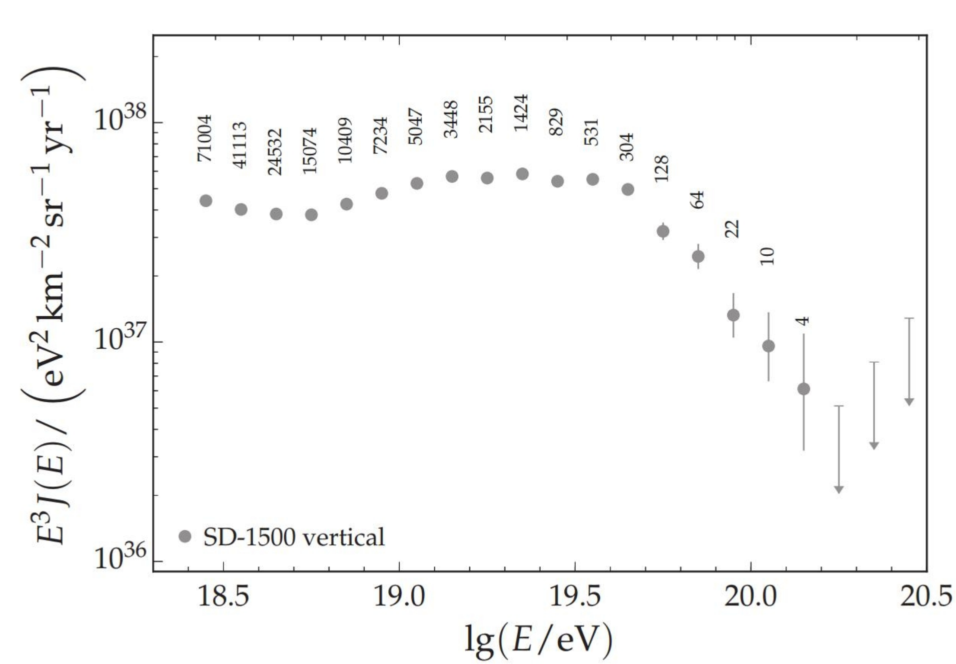
Able to use the FD to calibrate the SD's energy reconstruction

Other complementary detection possible…



Flux suppression around 10^{20}eV :
What is causing it ?

What particles are making up the UHECRs flux ?

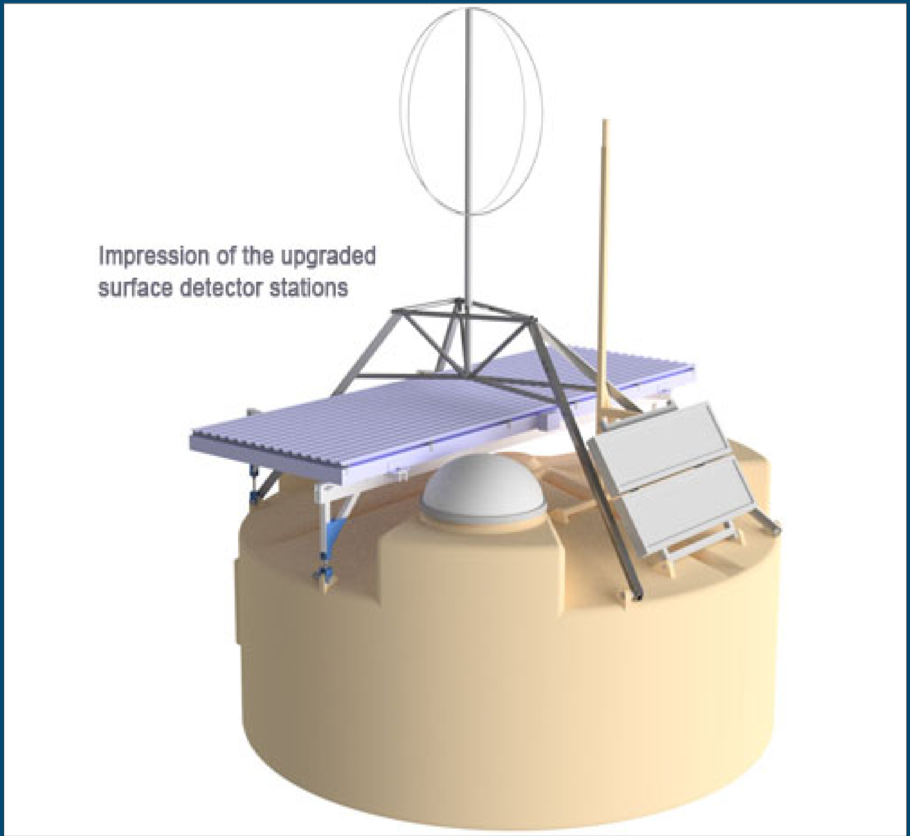


From Francesco Fenu, Contribution to the ICRC2017 : arXiv:1708.06592v2

AugerPrime Upgrade

Science Motivation :

- Probe the flux suppression
- Look at the flux composition at the highest energies
- UHE hadronic physics



AugerPrime :

- Add Scintillator Surface Detectors : on top of the WCD, to disentangle muon/EM components
- Upgraded electronics
- Radio Upgrade : add antennas on top of WCD to detect the showers radio-emissions

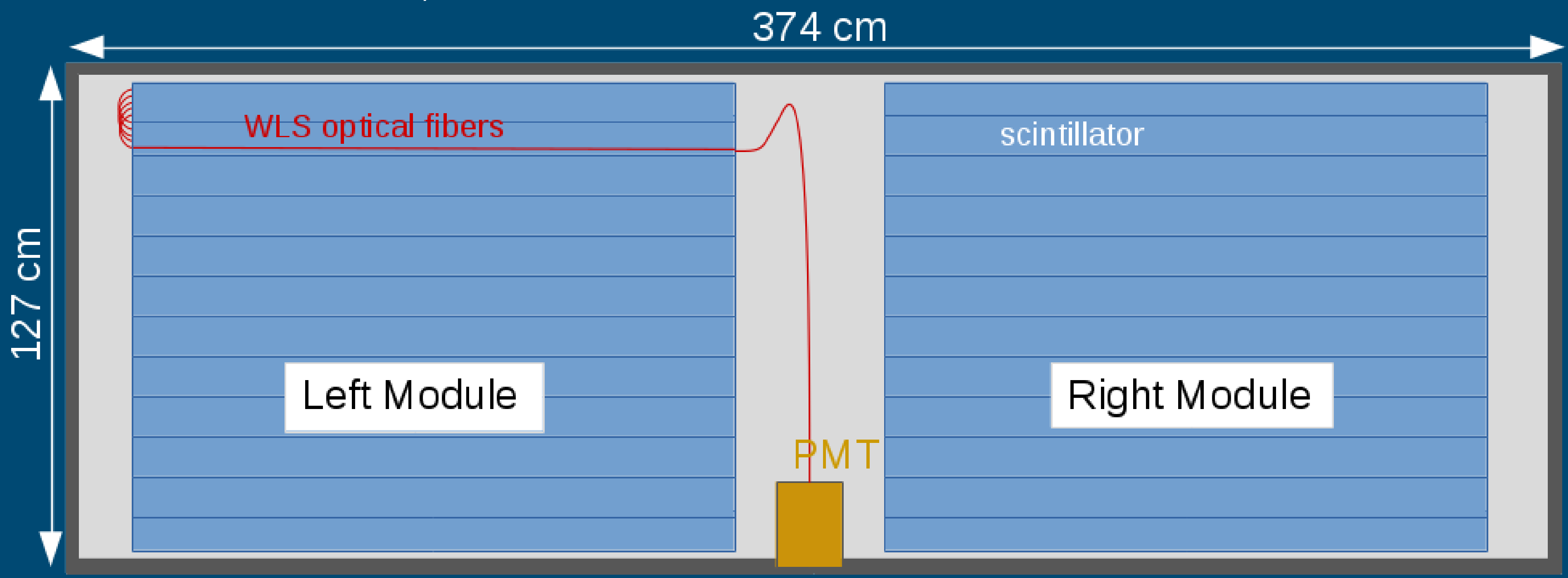
Scintillator Surface Detectors

-

Construction & Validation

Scintillator Surface Detector (SSD)

Objectives :
- add an independent and different measurement of the EAS' components, at the same place as the WCD
- reliable detector, with low maintenance



Detector's signal :
- the SSD's signal will be dominated by the shower's electrons, while the WCD's signal is dominated by the photons and muons.
- particle → scintillator → light → fibers → PMT → signal
- calibrated with atmospheric muons

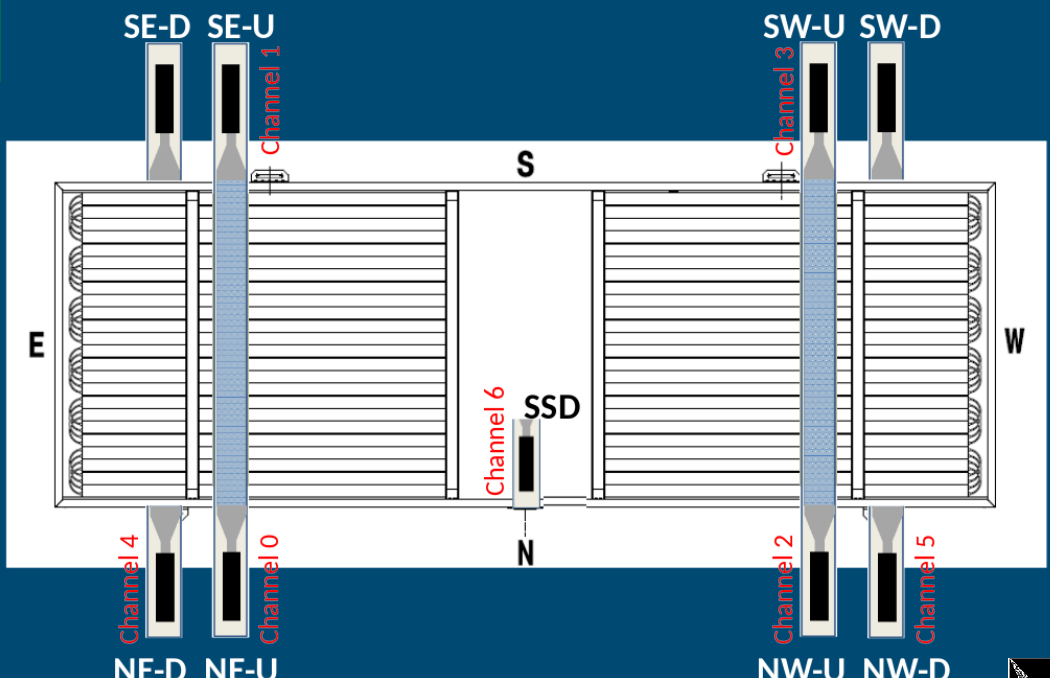
SSD – Construction & Tests Setup

1200 SSDs built in 6 countries : 90 SSDs built at LPSC

Important involvement by the SDI, electronics and administrative departments

The scintillator boards are used as external triggers :

- a particle (muon) goes through both scintillators (up and down)
- read-out electronics : "trigger"
- record the signal from all PMTs inside a user-defined time window *(800 ns before trigger and 300 ns after trigger)*

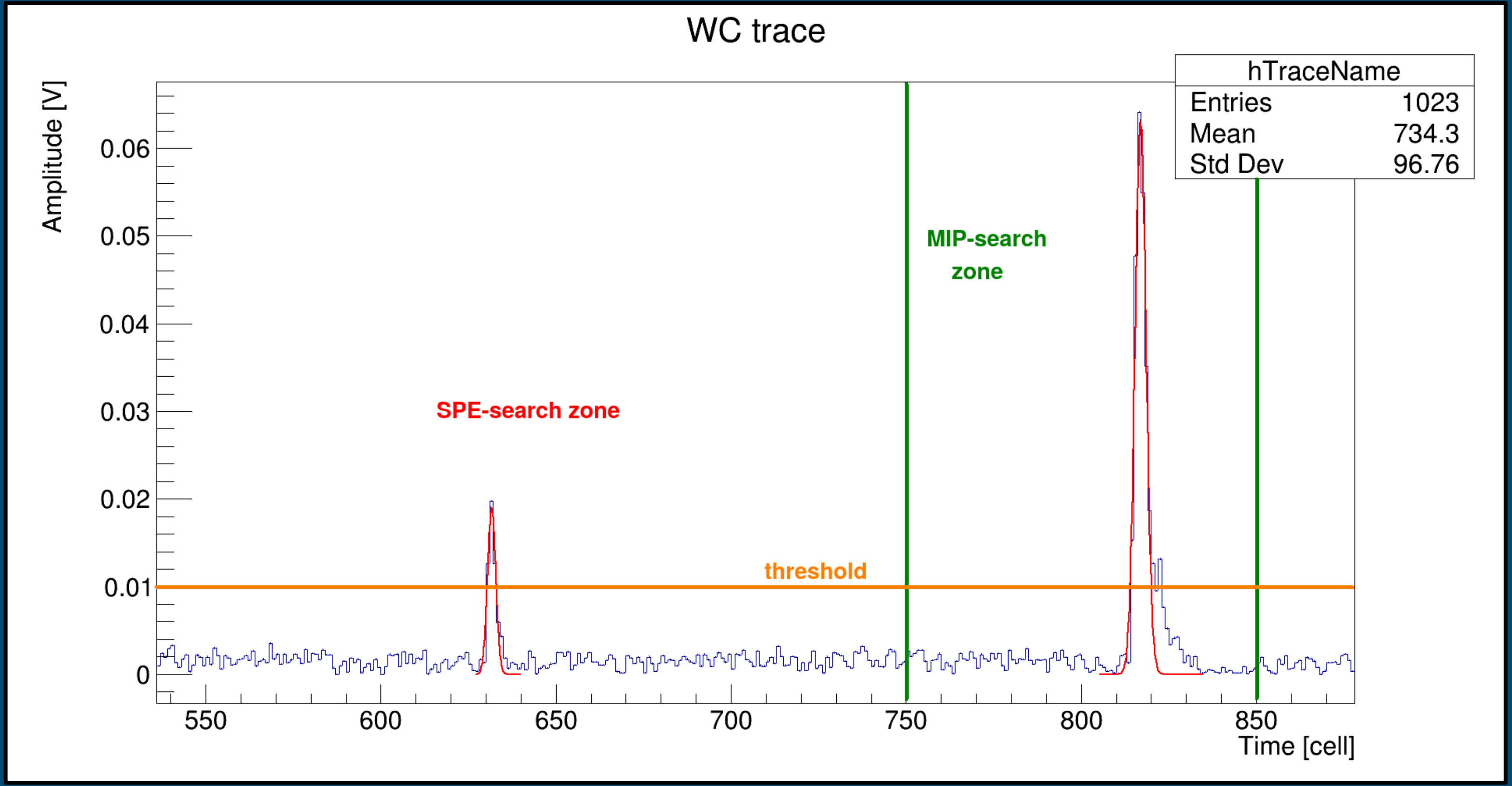


SSD Tests – Data Analysis

Minimum Ionising Particle (MIP) :
Minimum energy deposited by a through-going relativistic particle (i.e muon)

Single Photo-Electron (SPE) peak :
signal picked-up by the PMT for a single photoelectron inside the detector

For each triggered event (50k per run) :
- scan the SSD trace
- look for a MIP and/or SPE peak
- draw the distribution and fit
- use timing information on the external trigger to select vertical events → VMIP

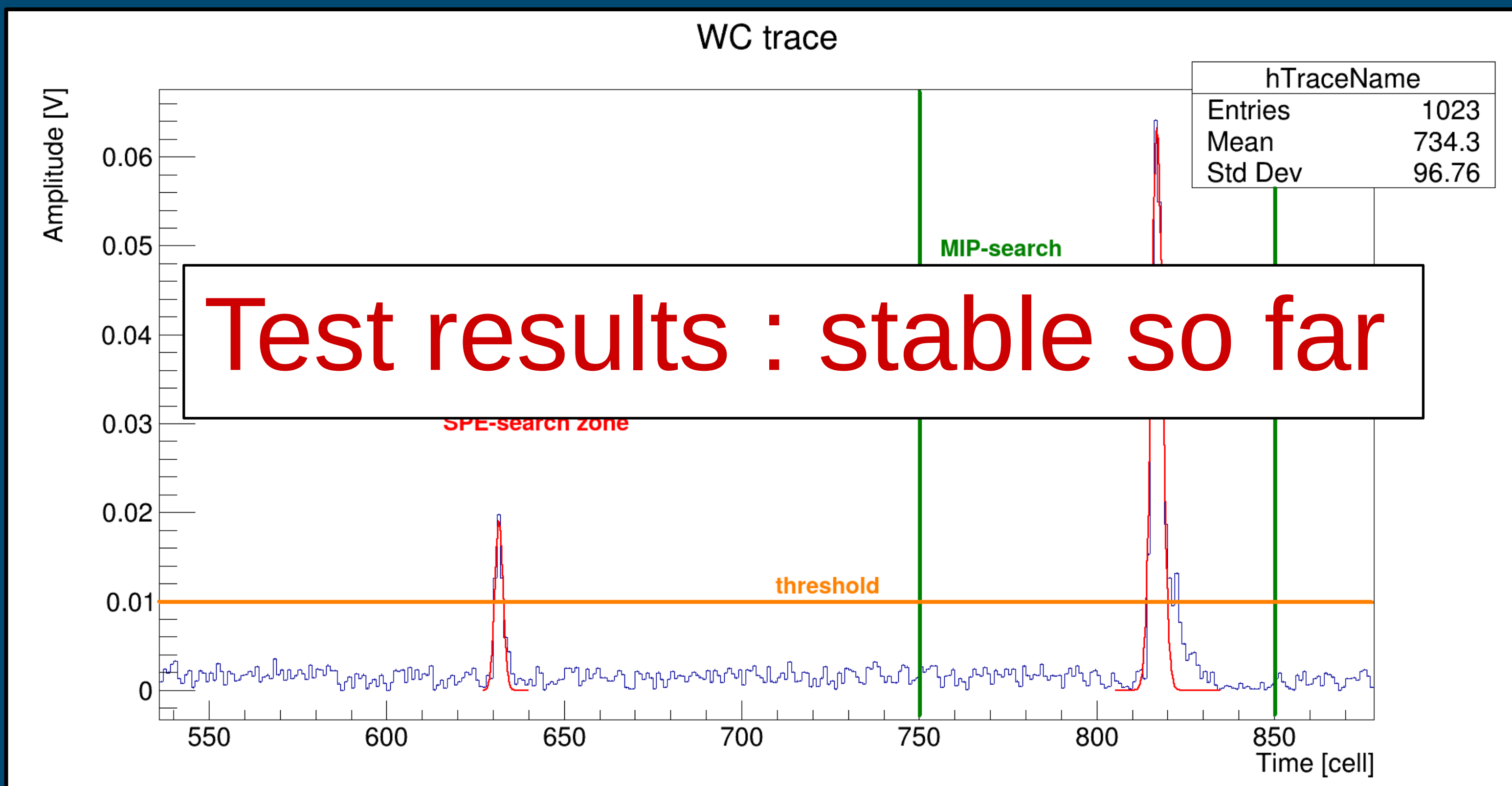


SSD Tests – Data Analysis

<u>Minimum Ionising Particle (MIP) :</u>
Minimum energy deposited by a through-going relativistic particle (i.e muon)

<u>Single Photo-Electron (SPE) peak :</u>
signal picked-up by the PMT for a single photoelectron inside the detector

For each triggered event (50k per run) :
- scan the SSD trace
- look for a MIP and/or SPE peak
- draw the distribution and fit
- use timing information on the external trigger to select vertical events → VMIP



Test results : stable so far

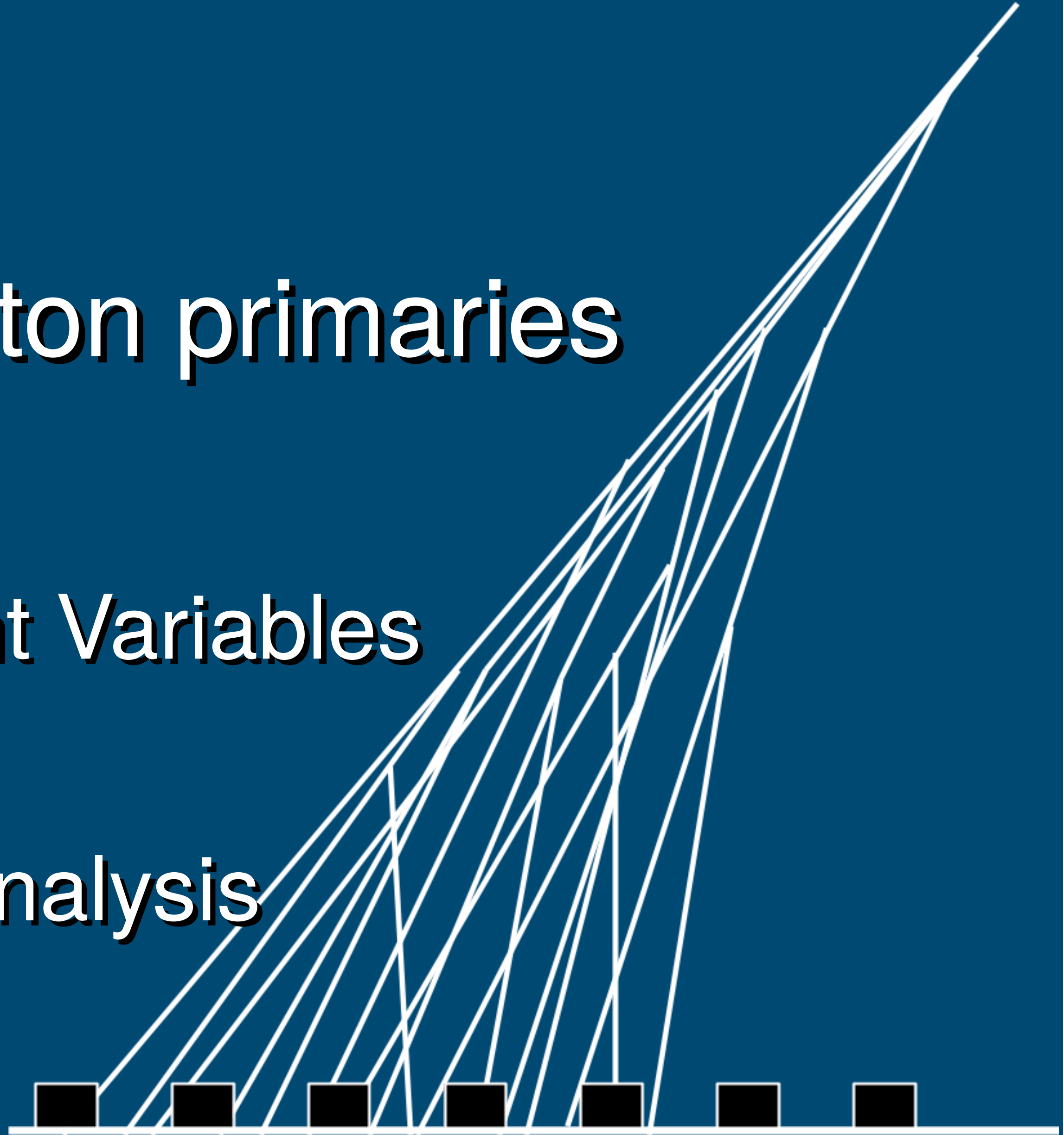
Search for photon primaries

-

Primary Dependent Variables

&

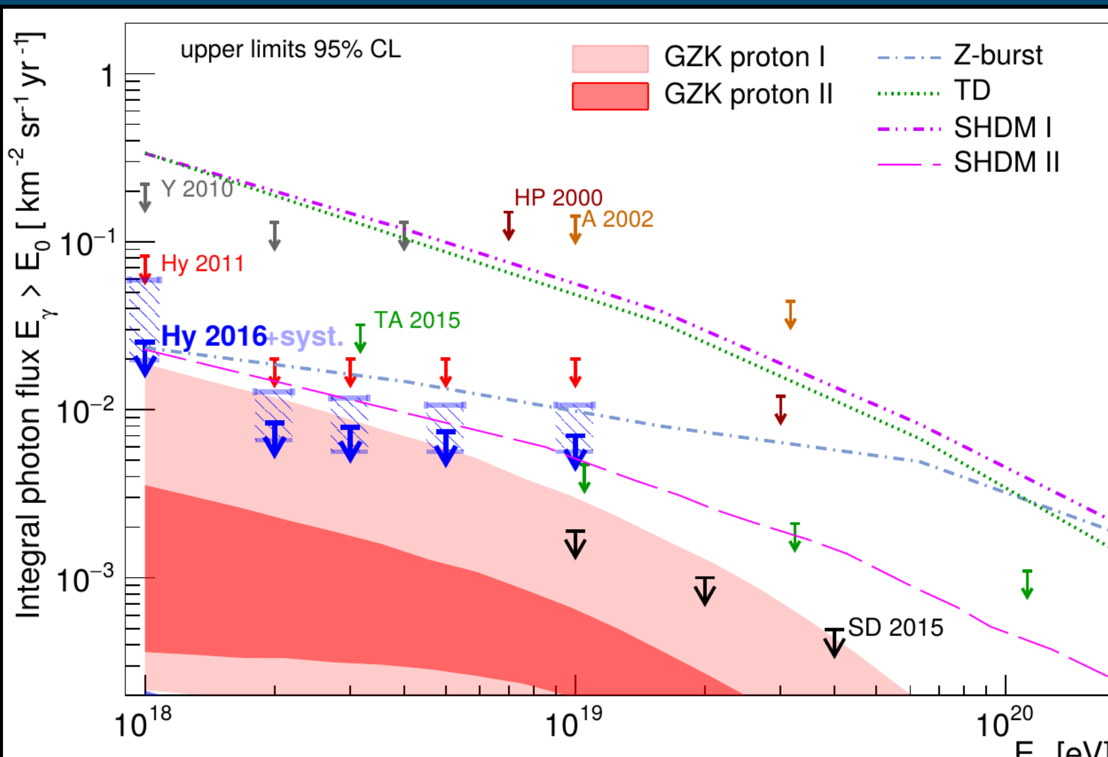
Multi-Variate Analysis



Why look for photon showers?

Why are photon primaries interesting? :

- point back to source directly
- UHE photons follow-up searches
- an excess of photons could independently confirm the GZK effect
- constrain astrophysical scenarios



From Auger Hybrid Photon-search paper, JCAP 1704 (2017) no.04, 009

Greisen **Z**atsepin **K**uzmin effect :

Interaction between UHECRs and Cosmic Microwave Background photons.

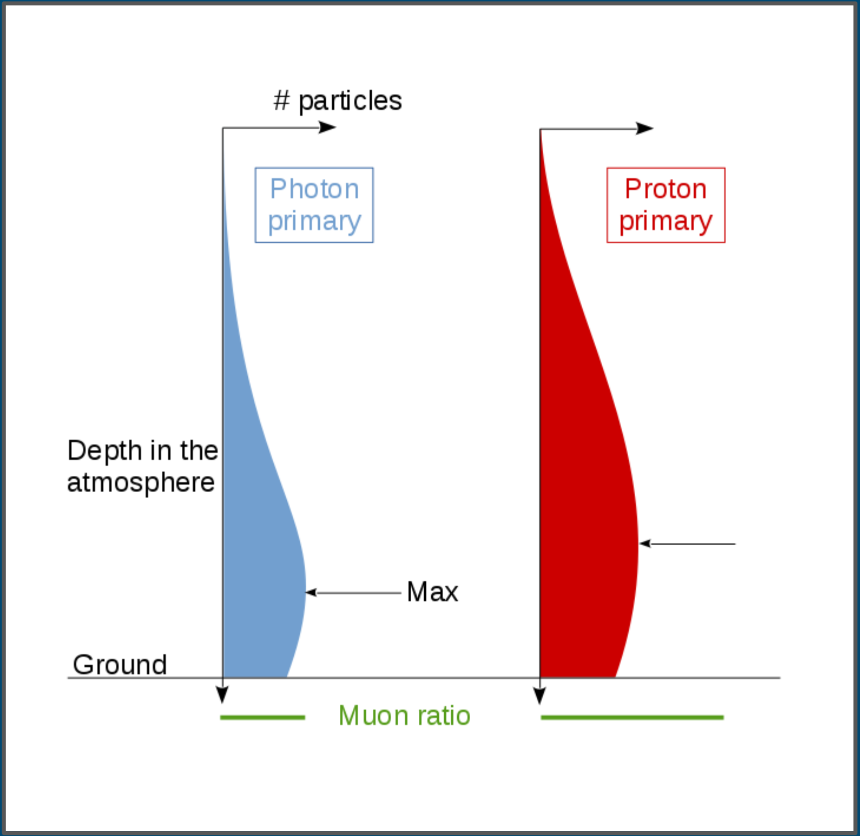
Above an energy threshold :
$E_{min} \sim 10^{19}$eV

$$p + \gamma_{\mathrm{CMB}} \rightarrow \Delta^+(1232) \rightarrow p + \pi^0$$
$$p + \gamma_{\mathrm{CMB}} \rightarrow \Delta^+(1232) \rightarrow n + \pi^+$$

$$\pi^0 \rightarrow \gamma + \gamma$$
$$\pi^0 \rightarrow \gamma + e^+ + e^-$$

Photon-induced showers

- Develop deeper into the atmosphere
- On average, more fluctuations of the shower maximum
- Less muons on ground



Photon Analysis in Auger

Two official analyses in Auger collaboration :
- Hybrid (SD+FD) analysis : more observables
- SD only analysis : *more data*

Aim of my analysis :

Use features inside the WCD signal to estimate the muonic content of the showers and thus differentiate between photon/proton primaries

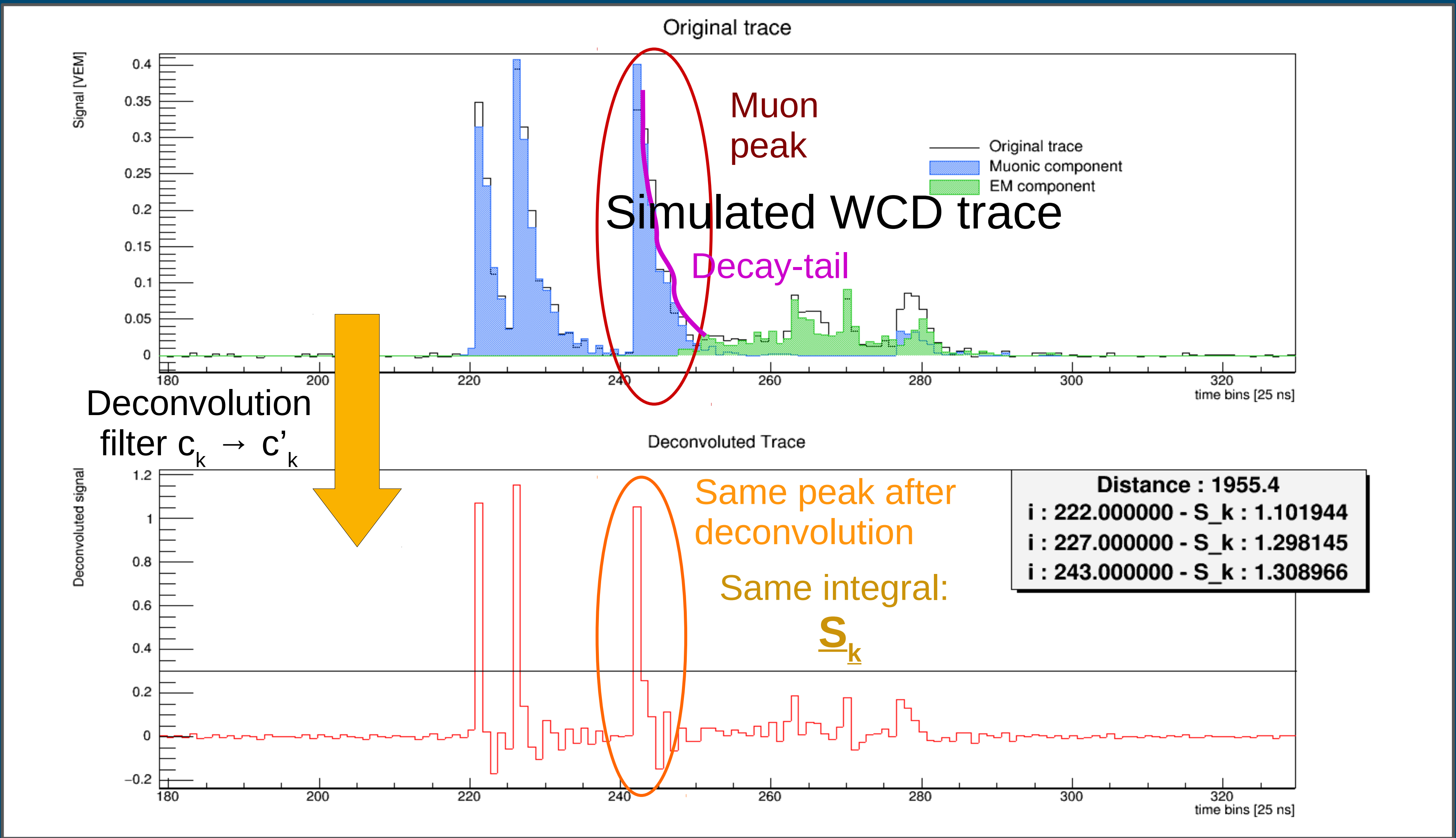
Improve the discrimination power of the analysis

SD only analysis = 100% duty cycle

Trace = WCD signal

Muonicity Method – st-level variable

Core of the Muonicity Method : perform a linear deconvolution to remove the exponential decay-tail of a muon signal while integrating the total muon-signal.



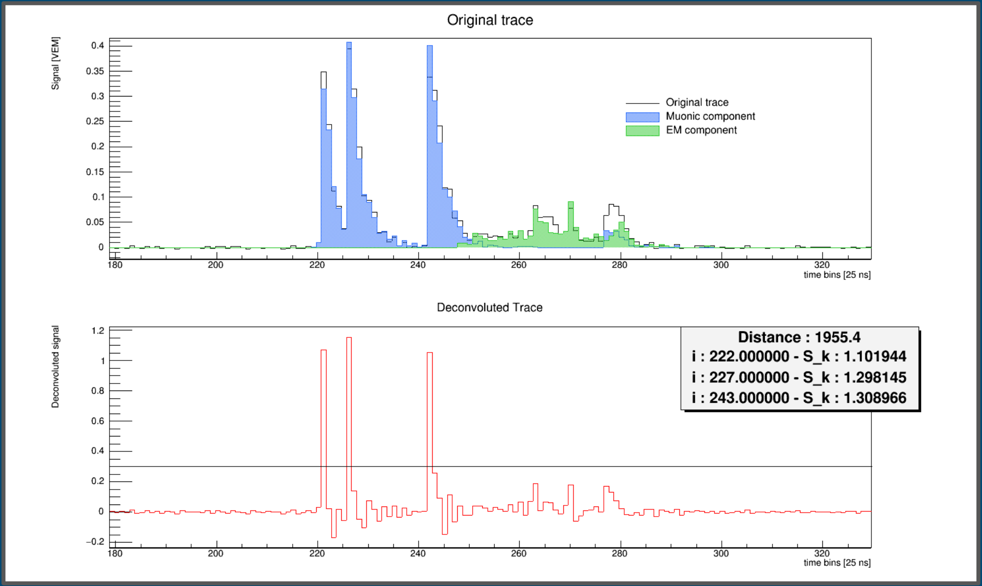
Muonicity Method – st-level variable

Core of the Muonicity Method : perform a linear deconvolution to remove the exponential decay-tail of a muon signal while integrating the total muon-signal.

Deconvolution Filter :

$$c_k' = \frac{c_k - \alpha . c_{k-1}}{1 - \alpha}$$

with : $\alpha = exp(-\Delta t / \tau)$



For each station *j* deconvolute the VEM trace, identify the peaks :

- Calculate S_k for each peak
- Sum S_k for each station : $\mathbf{S_{peaks}}$

$$S_{peaks,j} = \sum_{k=0; S_k > 0.7 VEM}^{N_{peaks}} S_k$$

Muonicity Method – evt-level variable

Value calculated on trace, for proton and photon showers : S^{calc}_{peaks}

Training with proton simulations

Machine Learning Model :

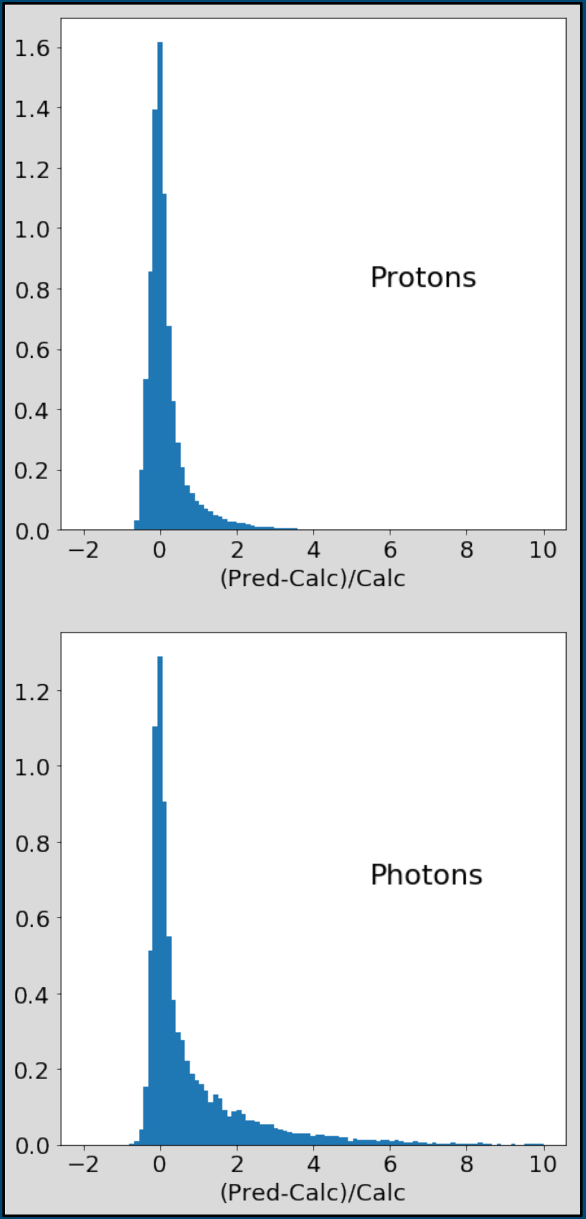
- Parameters : Distance, Zenith, SD_Energy
- Predict S^{pred}_{peaks} values from parameters

Predict a value from parameters

Value predicted by proton-trained Machine Learning model : S^{pred}_{peaks}

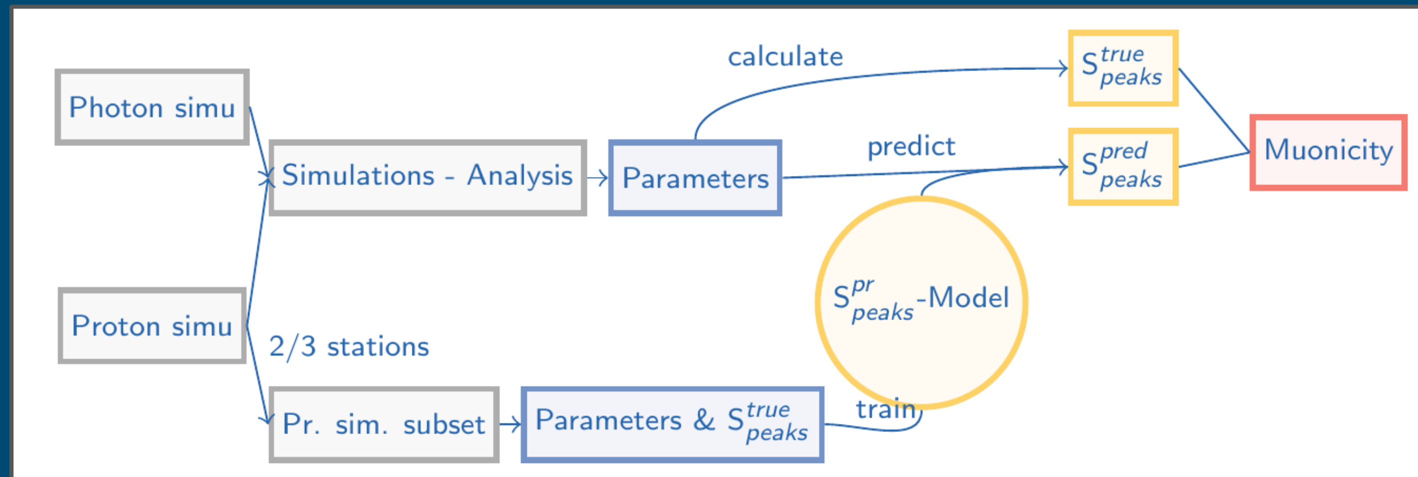
→ *the model is supposed to better reconstruct proton-* S^{pred}_{peaks} *(closer to* S^{calc}_{peaks} *)*

$$\mathcal{M} = \frac{1}{N}\sum_{j=0}^{N}\frac{S^{calc}_{peaks,j}}{S^{pred}_{peaks,j}}$$



Muonicity Method – Preliminary Results

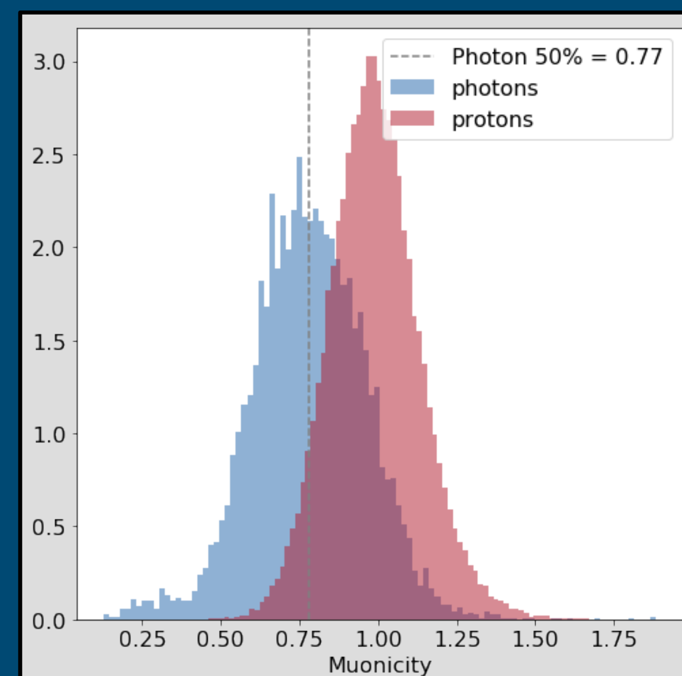
Simplified pipeline of the Muonicity Method →



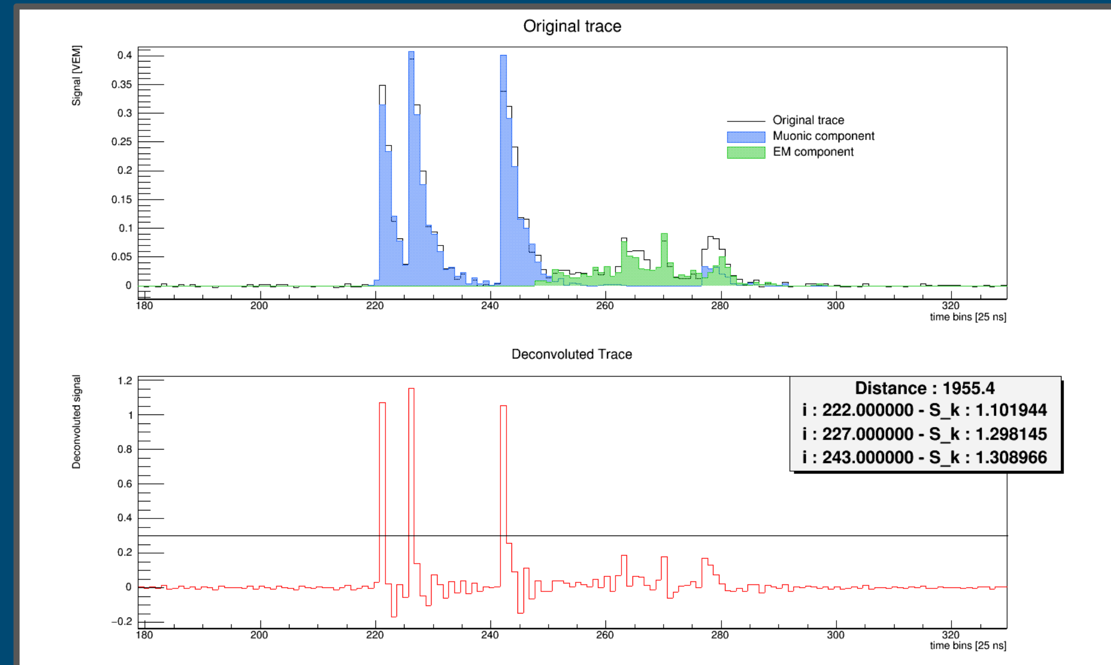
On its own and without further parameter tuning, the Muonicity variable has some separation power

Merit Factor η=0.91

→ Can we extract more discrimination from the muonic component of the SD traces?



Smoothing/Muonicity Method

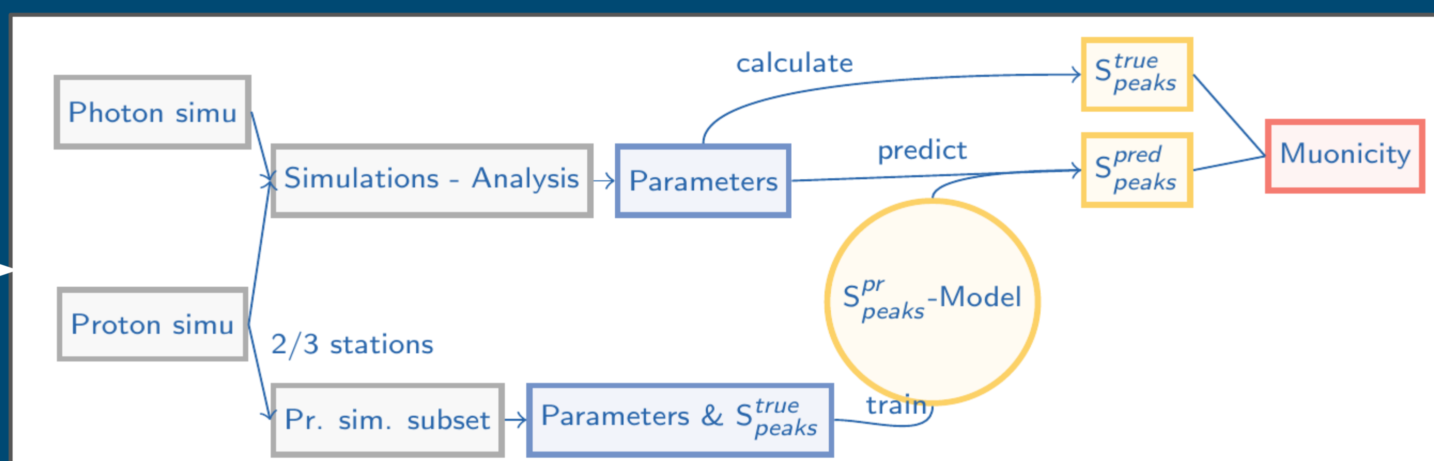


<u>Muonicity :</u>

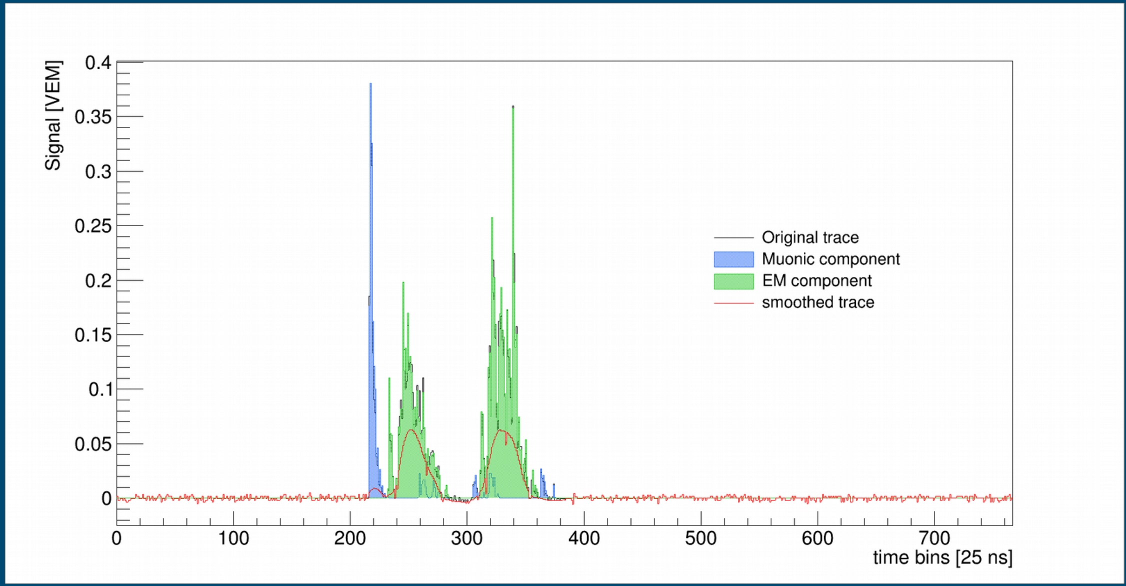
- deconvolute muon peaks
- sum peaks signal

→ S_{peaks}

Simplified pipeline of the Muonicity Method →



Smoothing/Muonicity Method

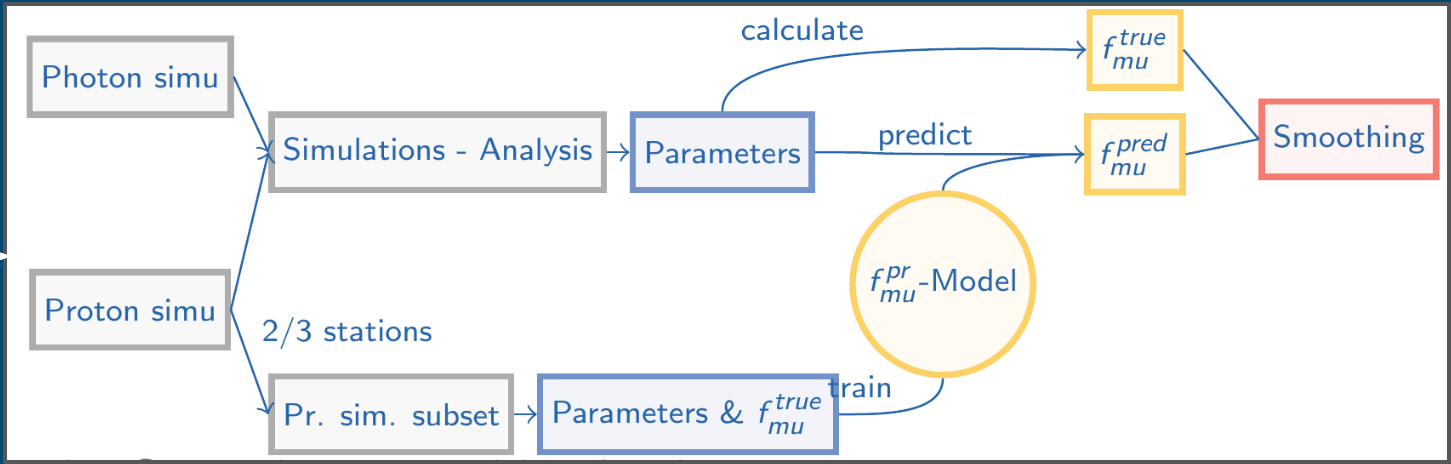


<u>Smoothing :</u>

- smoothen out the muon peaks
- estimate the EM signal

→ f_{mu}

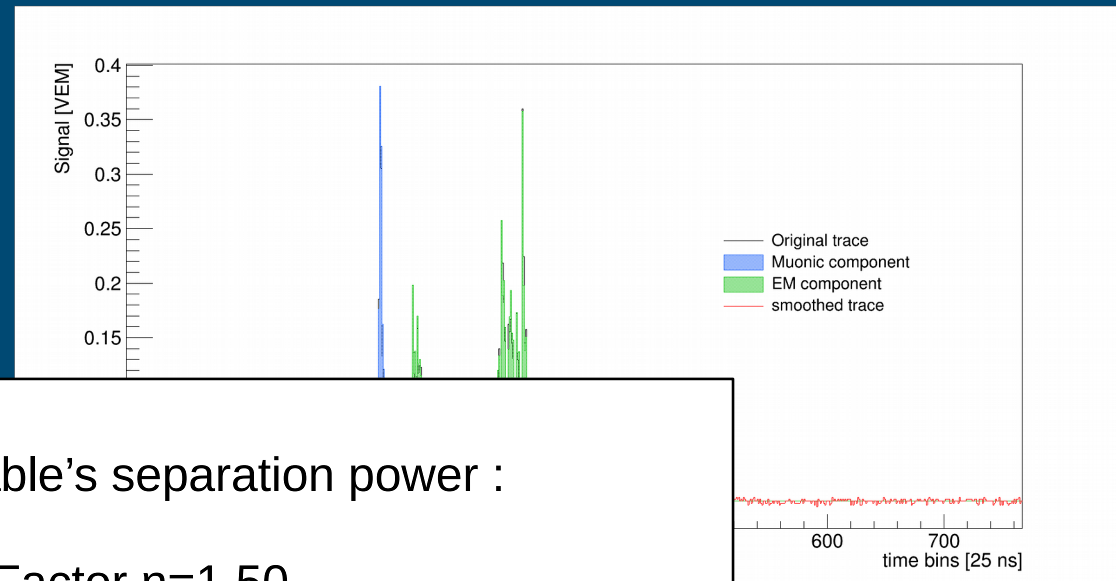
Simplified pipeline of the Smoothing Method →



Smoothing/Muonicity Method

Photon Search

<u>Smoothing :</u>
- smoothen out the muon peaks
- estimate the EM signal

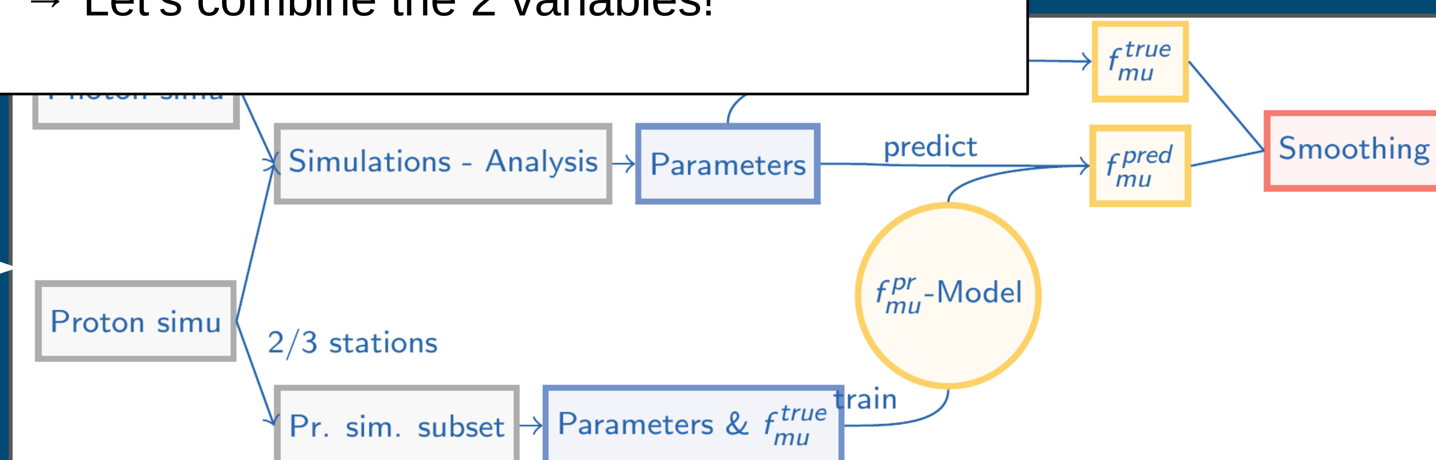


Smoothing variable's separation power :

Merit Factor η=1.50

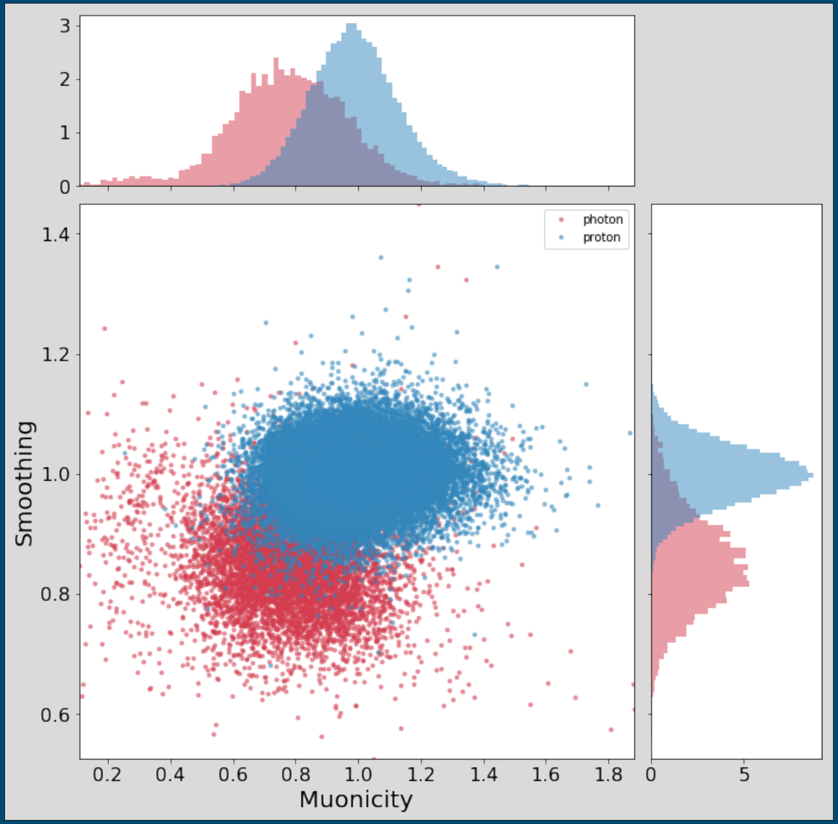
→ Let's combine the 2 variables!

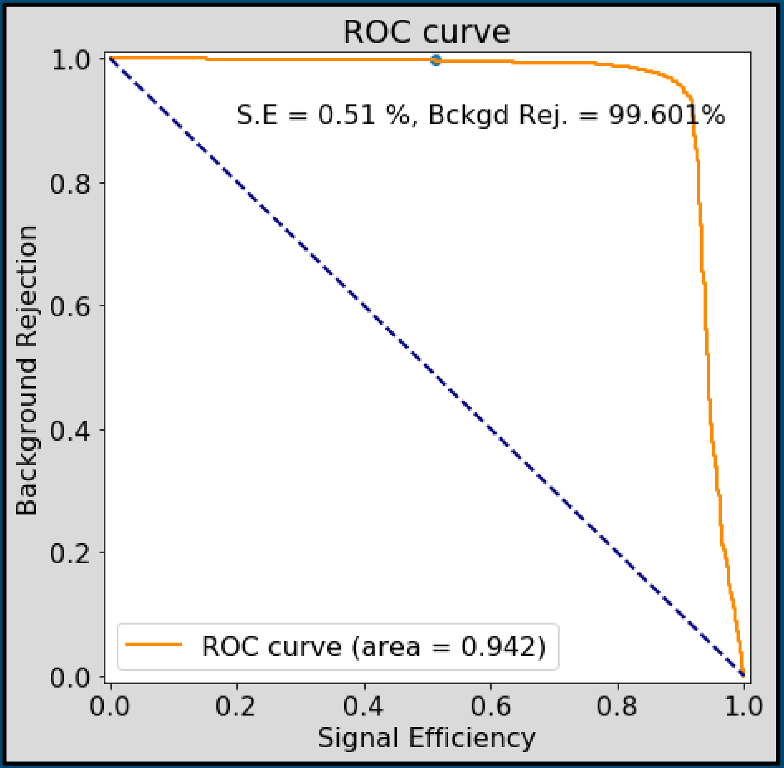
Simplified pipeline of the Smoothing Method



Multi-Variate Analysis

- As expected the 2 variables are correlated
- But the separation power is still better by combining the two





- Using a Support Vector Machine classifier with RBF kernels
- At 50% signal efficiency, 99.6% background rejection

What I did so far

- <u>SSDs assembly :</u>
  - Wrote a procedure to standardize the SSDs construction process
  - Helped design and install the SSDs test setup
  - Performed the SSDs validations
- <u>Photon-search :</u>
  - Implemented discrimination observables calculation inside the Auger framework
  - Designed :
    - a flexible method for photon/proton discrimination
    - a pipeline for MVA based on machine learning
- <u>Shifts, Formations & Outreach :</u>
  - Participated to FD shifts and beta-tested the SD shifts
  - Outreach to scholars
  - Label REI
  - SOS school, ISAPP school

What's next ?

- SSDs tests :
  - Re-run the analysis on the whole dataset once the production is finished
- Muonicity & Smoothing methods :
  - Optimize the observables
  - Re-run the pipeline with tuned parameters, updated simulations and heavier primaries.
  - Try changing the ML input parameters
  - Add other observables?
- Apply the method to data
- Detect UHE photons or improve the limits on the flux

Thank you !

LPSC Grenoble

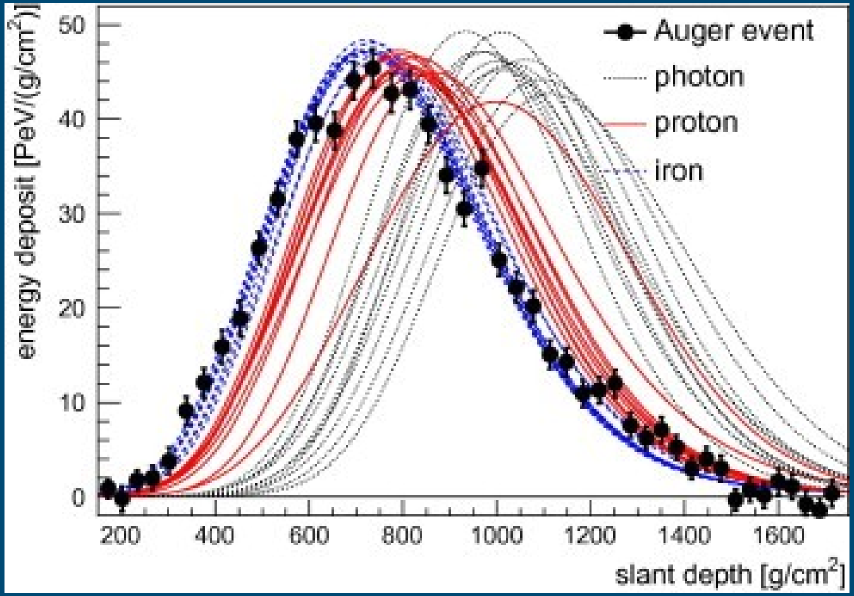
PIERRE AUGER OBSERVATORY

Primary dependence of the showers

Differences in the "shape" of the showers :

- First interaction further down for light CRs

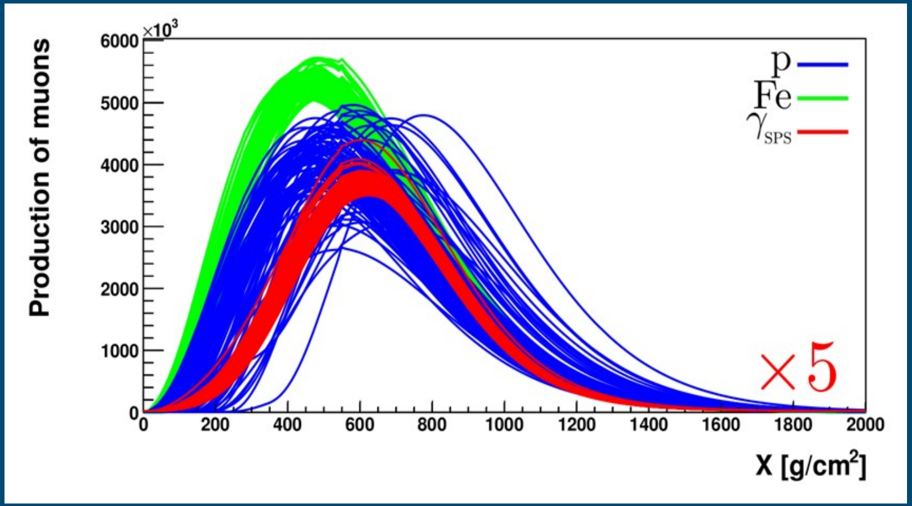
- Spread of the $z(N_{max})$ tighter for heavier CRs



Differences in the composition of the showers :

- Heavier primary → faster energy loss in the atmosphere

- Muon ratio α $A_{primary}$



Smoothing Method – st-level variable

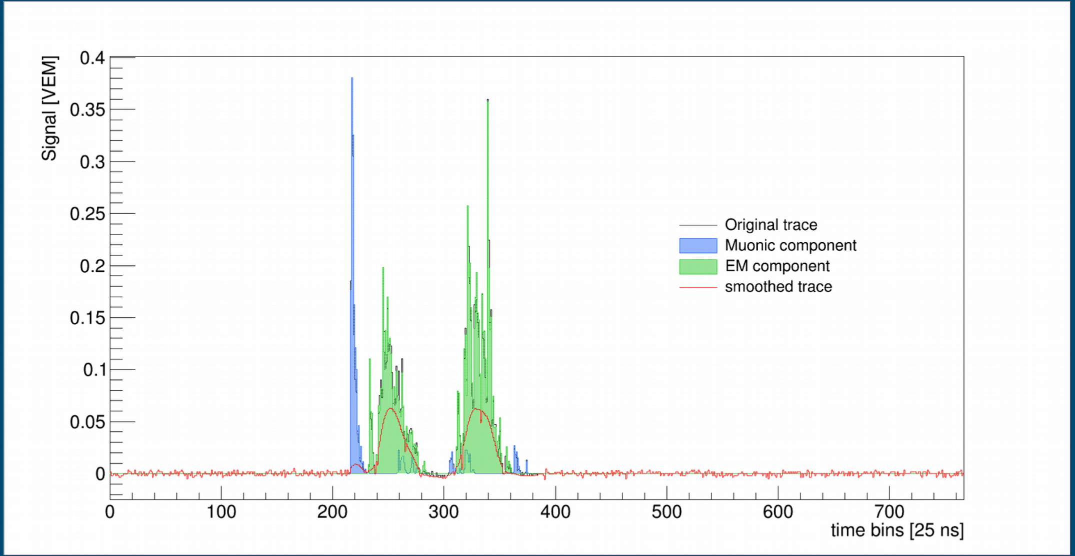
Back-up

Core of the Smoothing Method : extract the muonic component of the trace by performing a sliding-window averaging to remove spikes

- $S^{sm}(t_i) = \sum_k \frac{S(t_k)}{2.N_{bin}+1}$ with $k \in [i - N_{bin}, i + N_{bin}]$
- the muonic content is then : $S^{\mu}(t_i) = [S(t_i) - S^{sm}(t_i)]$
- This muonic content is substracted from the trace and the process is repeated a number of times

The smoothened trace is then compared to the original trace to obtain a st-level variable : f_{mu}

- $S^{\mu}_{tot} = \sum_i S(t_i) - S^{sm}(t_i)$
- $S^{EM}_{tot} = S_{tot} - S^{\mu}_{tot}$
- $f_{mu} = \frac{S^{mu}_{tot}}{S_{tot}}$

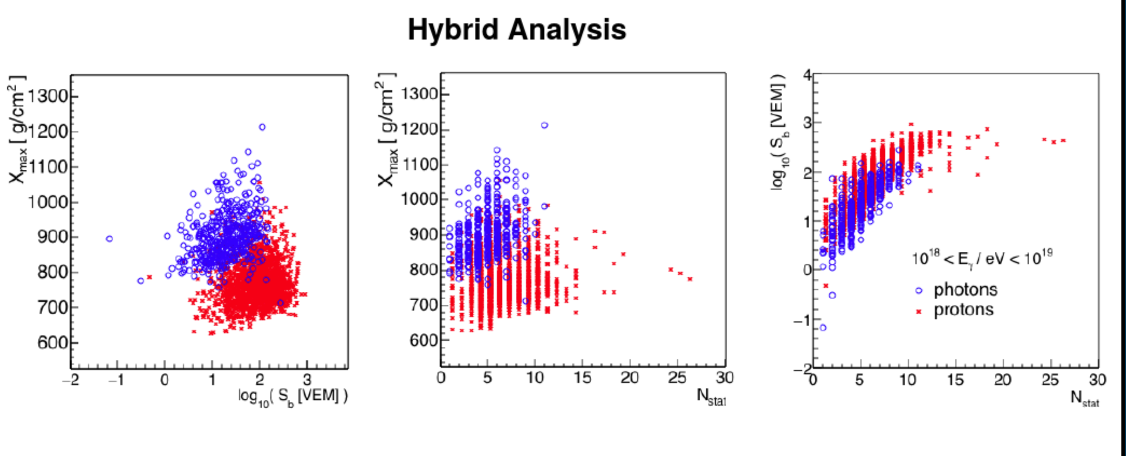


f_{mu} here, has the same role as S_{peaks} in the Muonicity Method

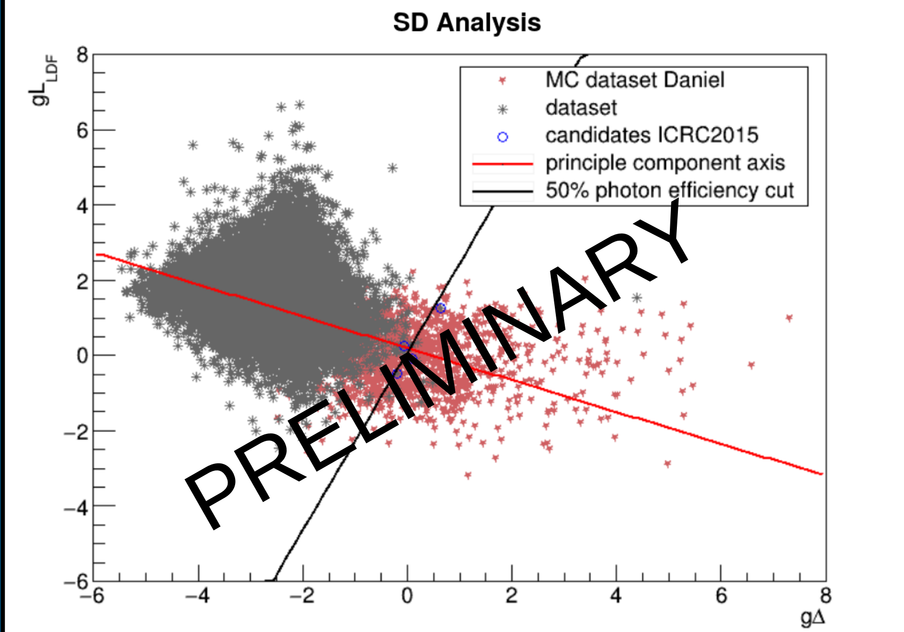
Photon Analysis in Auger

Back-up

- Hybrid (SD+FD) analysis : more observables
- SD only analysis : *more data*



From Auger Hybrid Photon-search paper, JCAP 1704 (2017) no.04, 009



From Auger SD Photon-search paper, to be published

Aim of my analysis :

Use features inside the WCD signal to estimate the muonic content of the showers and thus differentiate between photon/proton primaries

SD only analysis = 100% duty cycle

Trace = WCD signal

Anisotropy maps

Back-up

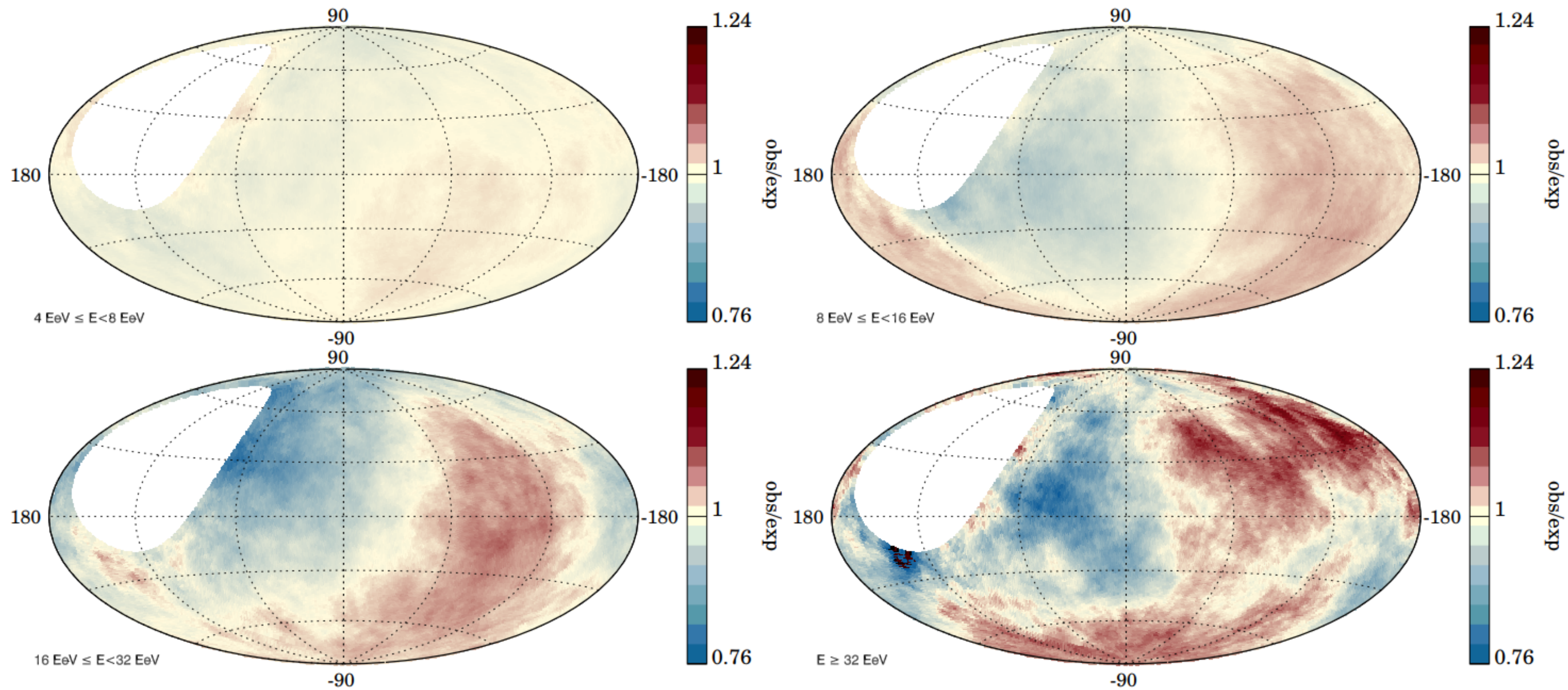
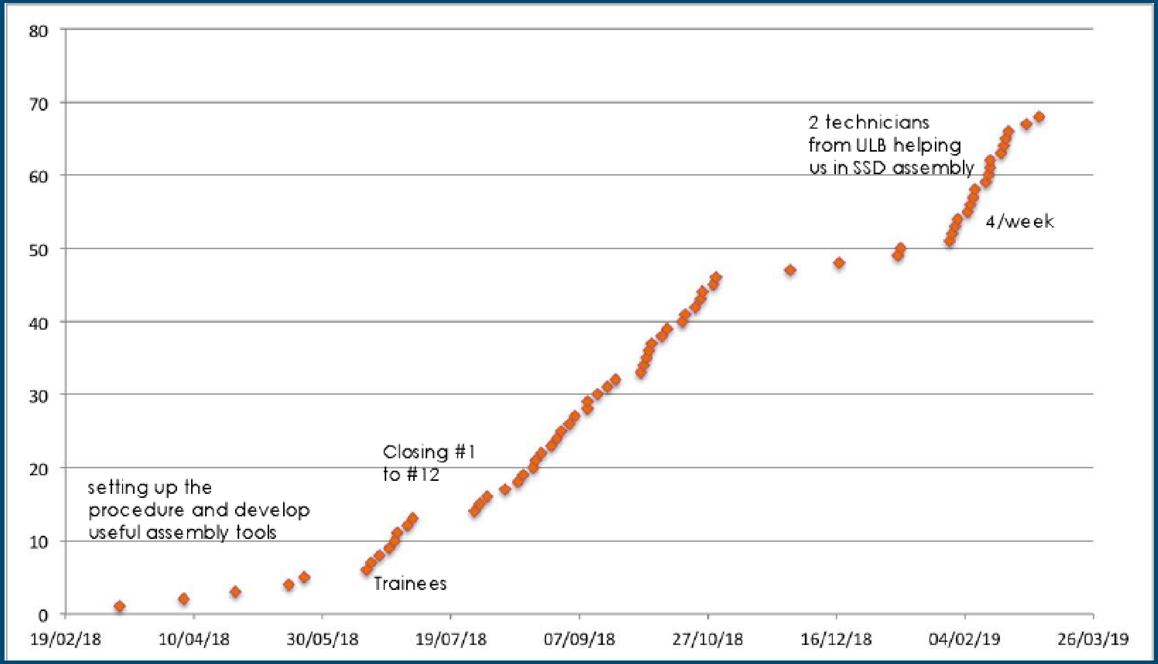
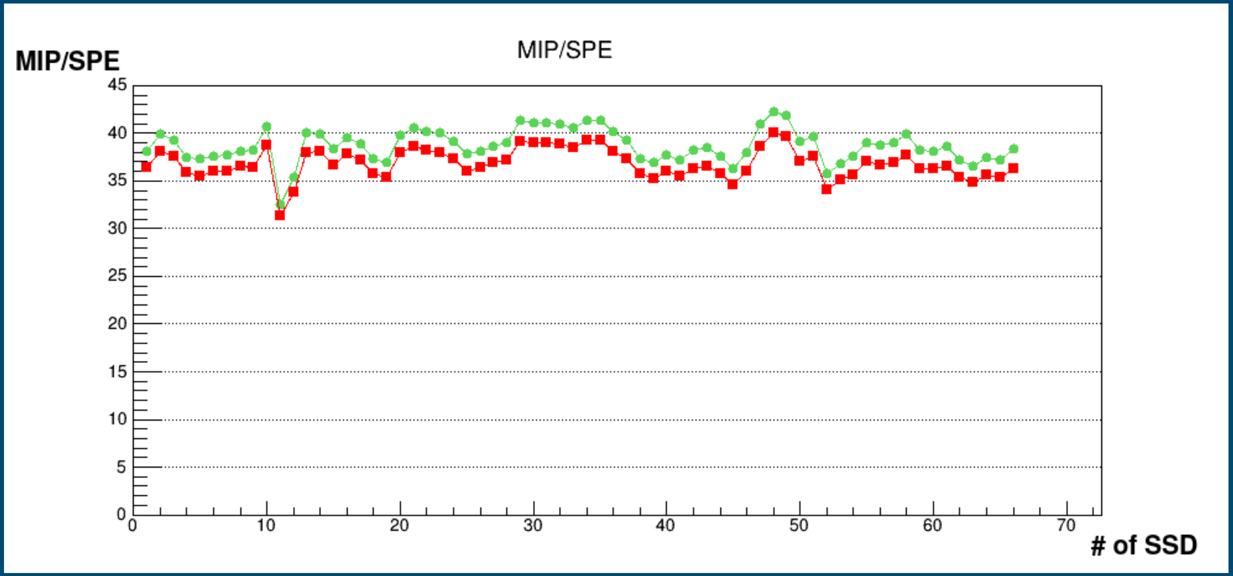


Figure 4. Maps in Galactic coordinates of the ratio between the number of observed events in windows of 45° over those expected for an isotropic distribution of arrival directions, for the four energy bins above 4 EeV.

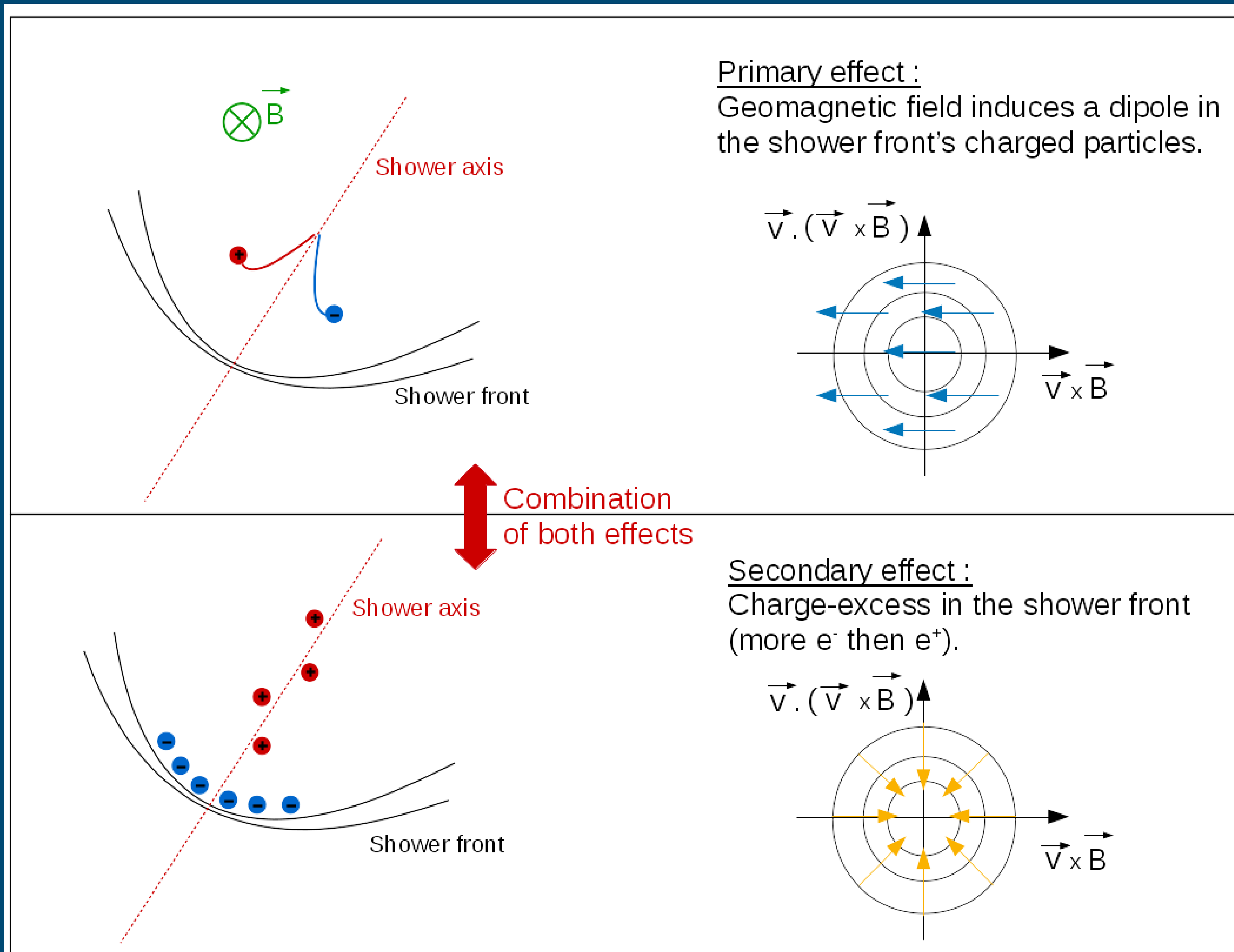
Constructing and validation status

Back-up





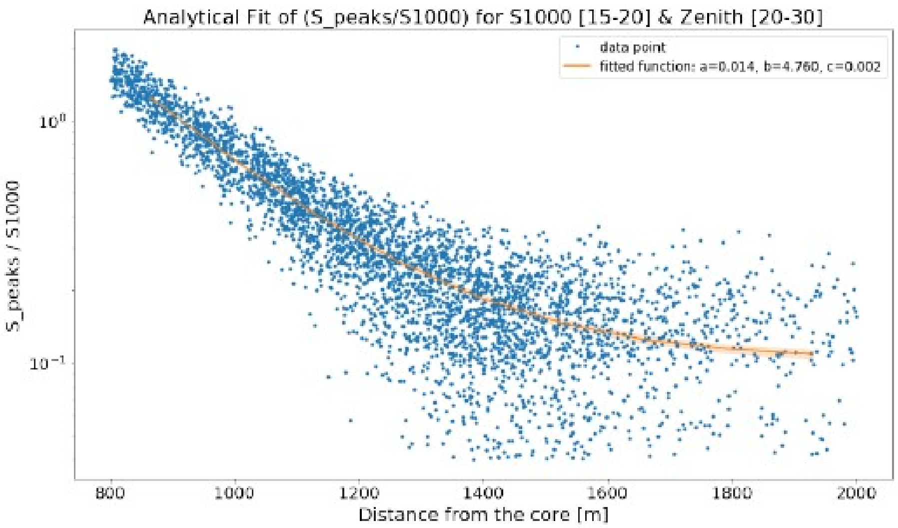
Radio Detection

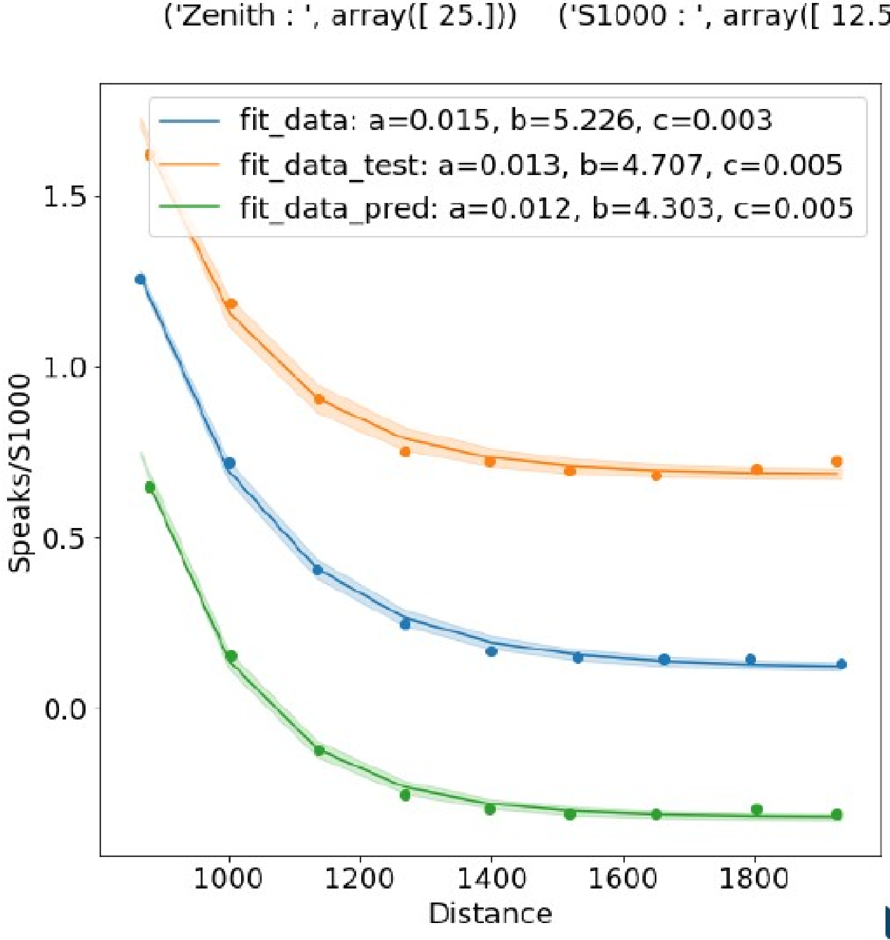


Reproduce analytical function with model

The analytical function fitted to predicted S_{peaks} is the same as the one fitted on data

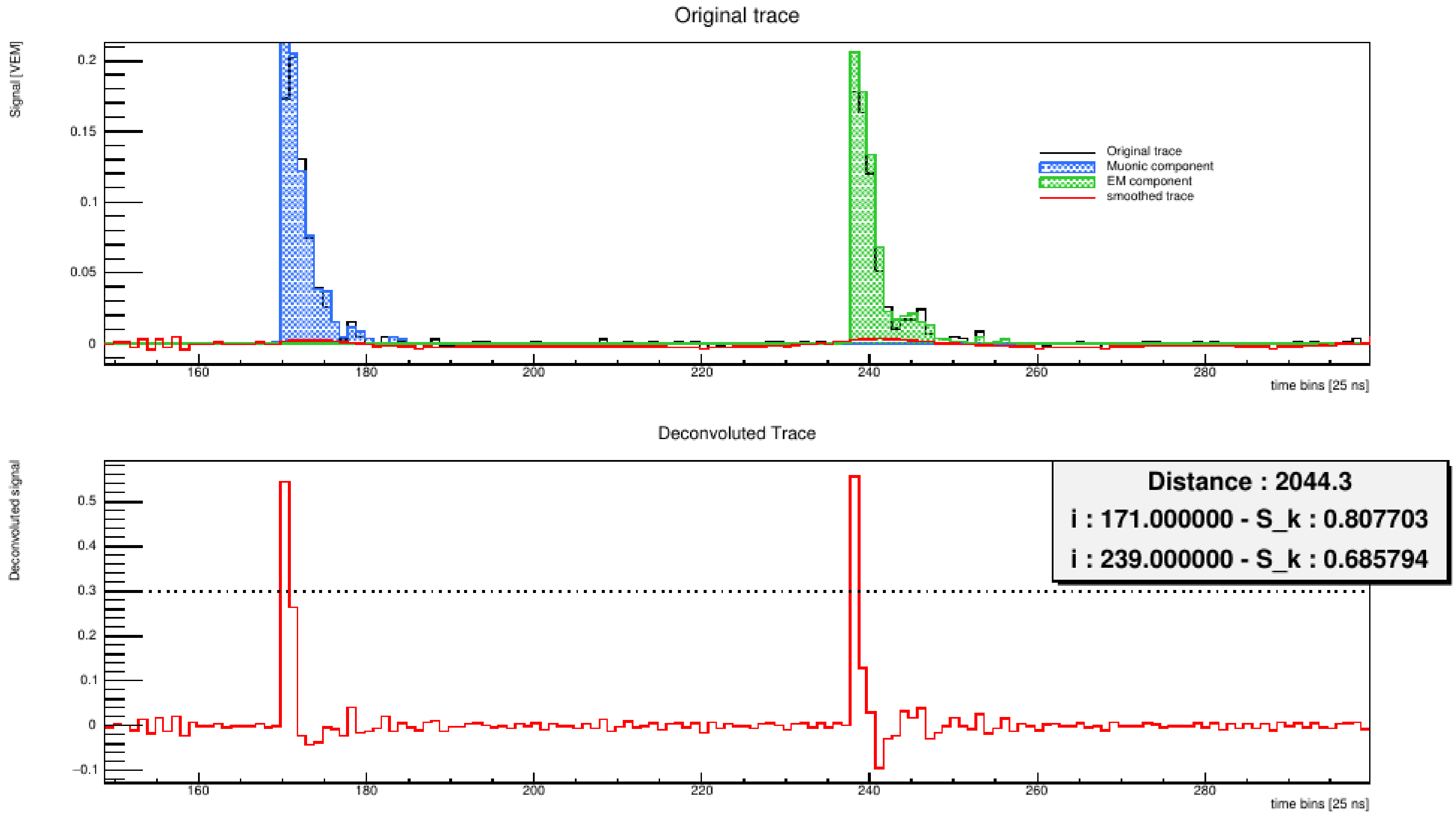
- Fitted function : $S_0 + A.e^{(-(Dist-1000)/B)}$ where : $(A/B/S_0) = (a/b/c).(\theta, S1000)$
- Here *fit data pred* & *fit data test* are shifted for clarity





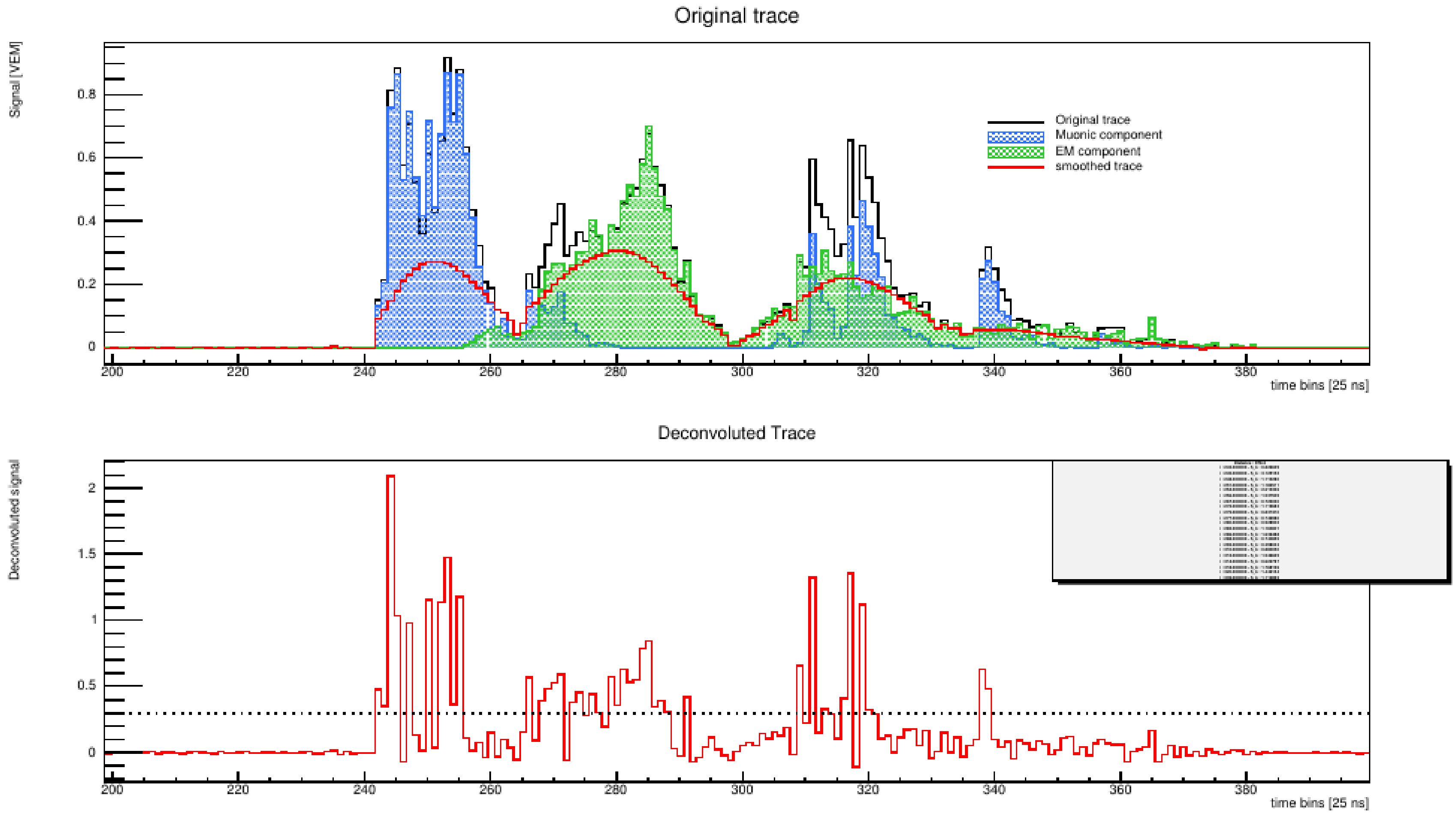
Not so convincing traces

Muon-like γ



Not so convincing traces

Overlapping peaks



RandomForestRegressor

A Random Forest Regressor trains several Decision Tree Regressor on data and parameter subsets. Then it combines those estimators to obtain a better one.
A Decision Tree recursively split the dataset to group identically labeled data.

- training vectors : $\vec{x_i} \in R^n$
  with $i = 1, ..., l$
- target value vector : $\vec{y} \in R^l$
- $y_m = \frac{1}{N_m} \sum_{i \in N_m} y_i$
- Mean Squared Error : $H(X_m) = \frac{1}{N_m} \sum_{i \in N_m} (y_i - y_m)^2$

At each node m, $H(X_m)$ is minimized to split dataset.

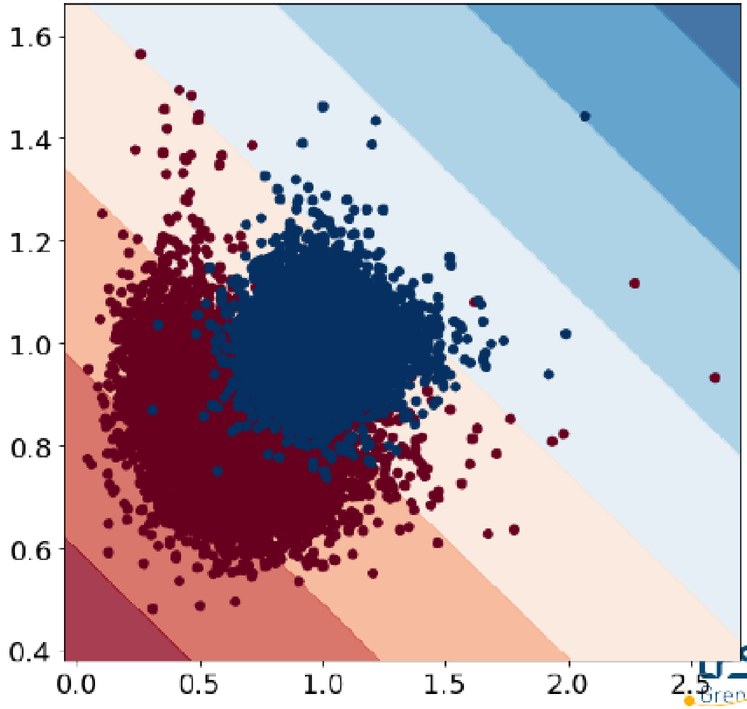
Support Vector Classification - Hyperplane

A **S**upport **V**ector **M**achine will construct a hyper-plane with the highest distance between two points of different classes.

$$[\frac{1}{n}\sum_{i=1}^{n} max(0, 1 - y_i(\vec{w}\cdot\vec{x_i} - b))] + \lambda \parallel \vec{w} \parallel^2$$

The aim is to minimize the above function given that :

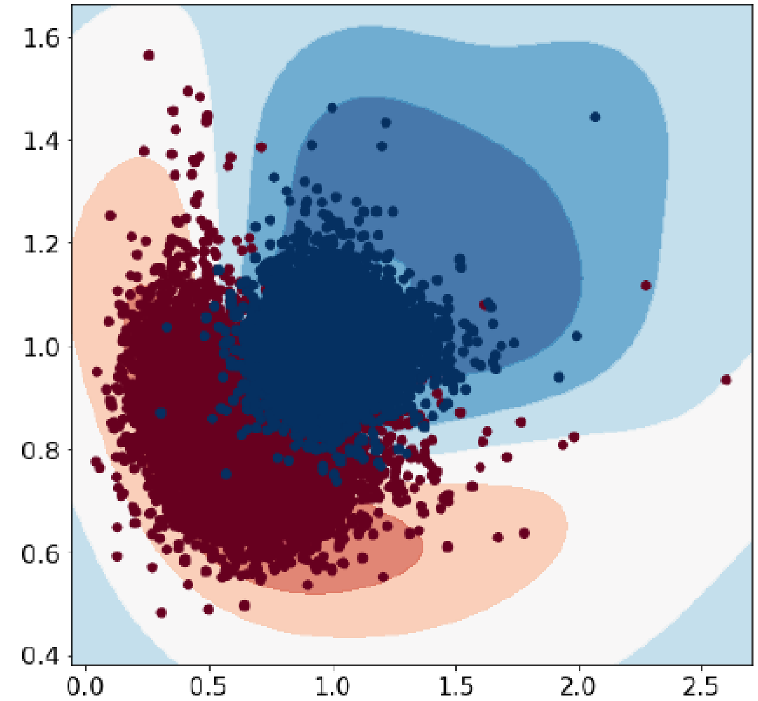
- $(\vec{x_i}, y_i)$ is a training dataset of n point.
  $\vec{x_i}$:parameters, y_i:(1 or -1) label
- hyperplane : $\vec{w}\cdot\vec{x} - b = 0$.
  if $y_i(\vec{w}\cdot\vec{x_i} - b) \geq 1$, then every data point is correctly predicted
- λ : tradeoff between margin-size and precision



Support Vector Classification - Kernels

For non-linearly separated data, it might be useful to project it in higher dimensions → **kernel functions**

Previous frame : linear kernel → no projections
Here : Radial Basis Function (RBF)



Mathematically, a kernel **K** applied to 2 data points can be written as :

$\mathbf{K}(\vec{x_i}, \vec{x_j}) = \vec{x_i} \cdot \vec{x_j}$: linear kernel
$\mathbf{K}(\vec{x_i}, \vec{x_j}) = \exp(-\gamma \|x - x'\|^2)$: RBF kernel