Gran Sasso Science Institute
Astroparticle Physics program
XXXI cycle

GRAN SASSO SCIENCE INSTITUTE
SCHOOL OF ADVANCED STUDIES
Scuola Universitaria Superiore

Reconstruction of events from the surface detector of the Pierre Auger Observatory using air shower Universality

PhD candidate: G. A. Anastasi

Supervisor: Prof. S. Petrera
Co-Supervisor: Dott. F. Salamida

L'Aquila, 8 May 2019

Introduction

Ultra-high energy cosmic rays (UHECR) are particles reaching the Earth from outer space with energies comparable to those of macroscopic objects (namely above $10^{18.5}$ eV).

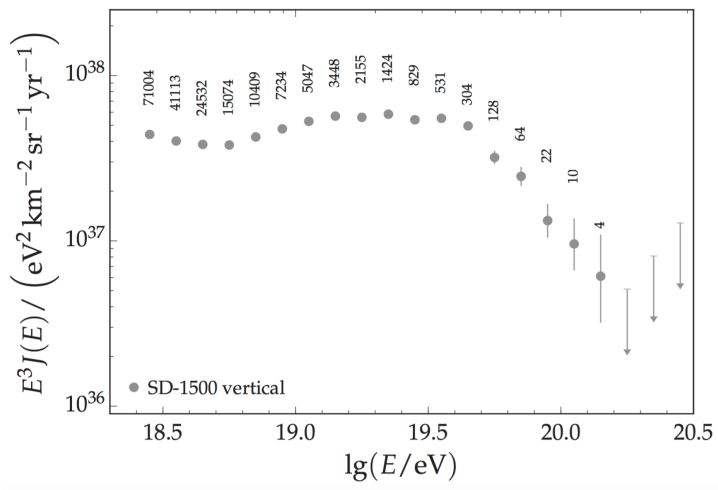
There are still several open questions in the field, such as:

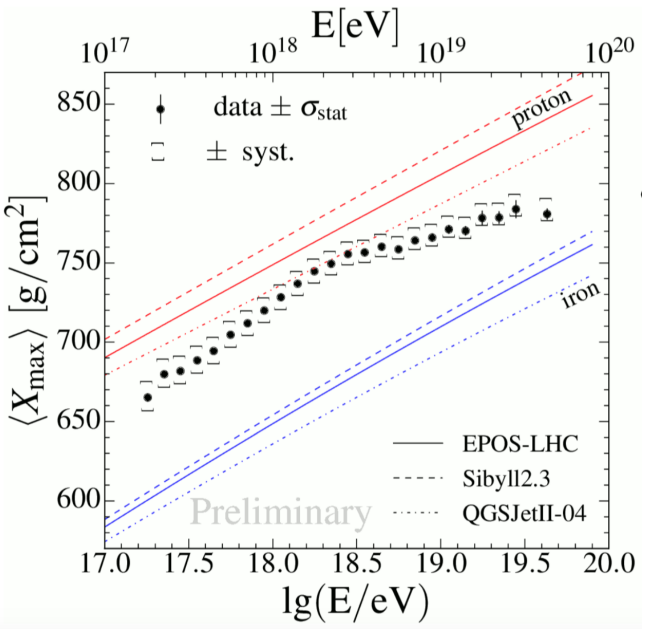
- *What is the nature UHECRs ?*
- *Which (and where) are the sources of UHECRs ?*
- *How are UHECRs accelerated to such extreme energies ?*

and many more.

To address these questions, a critical requirement is the extension of our knowledge about the **mass composition** of UHECR.

Ultra-high energy cosmic rays





Cosmic ray **energy spectrum** known up to $10^{20.2}$ eV.

Cosmic ray **mass composition** only known up to about $10^{19.5}$ eV.

With the aim of obtaining the mass composition at the highest energies, during my PhD I worked in the optimization of a new analysis based on the concept of **air shower universality**, proposed in the framework of the Pierre Auger Collaboration.

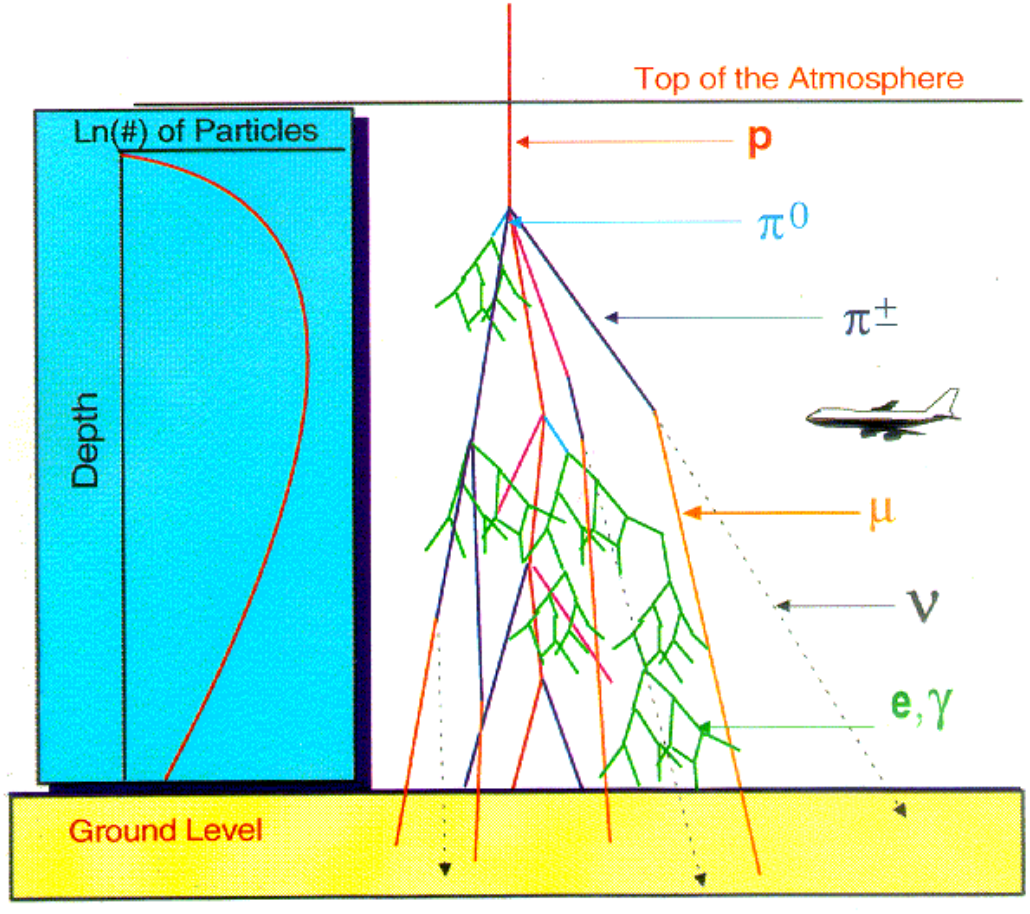
Part I
Detection of UHECR

Extensive Air Showers

The collision of cosmic rays with the atmosphere molecules produces a cascade of particles, called Extensive Air Shower (EAS).

The particles of an EAS initiated by a proton or a nucleus can be roughly divided into three components:

- *hadronic* (mainly pions);
- *electromagnetic* (e^+,e^-,γ);
- *penetrating* (muons and neutrinos).

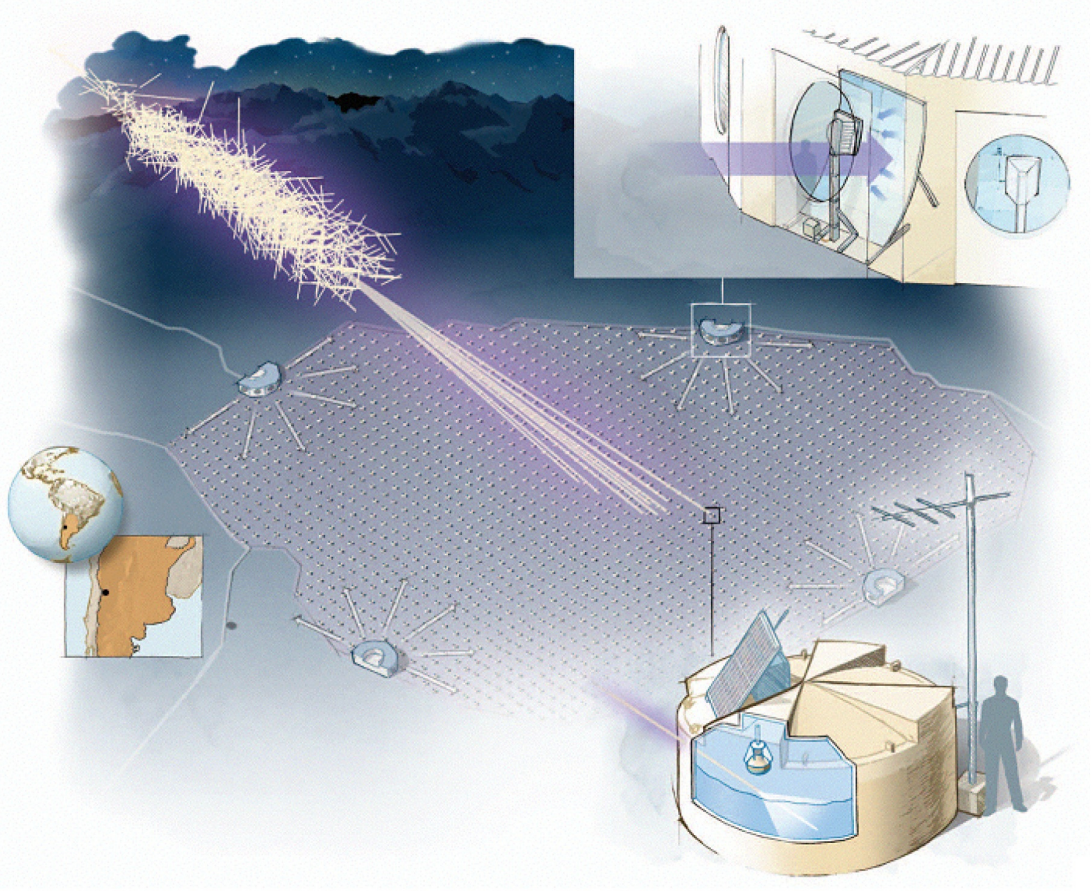


The Pierre Auger Observatory

The Pierre Auger Observatory, placed near the town of Malargüe in Argentina, is the world's largest cosmic ray experiment.

The surface array is composed by more than 1600 water-Cherenkov detectors, distributed over $\sim$3000 km^2.

During clean moonless nights, 24 fluorescence telescopes hosted in four buildings observe the atmosphere above.



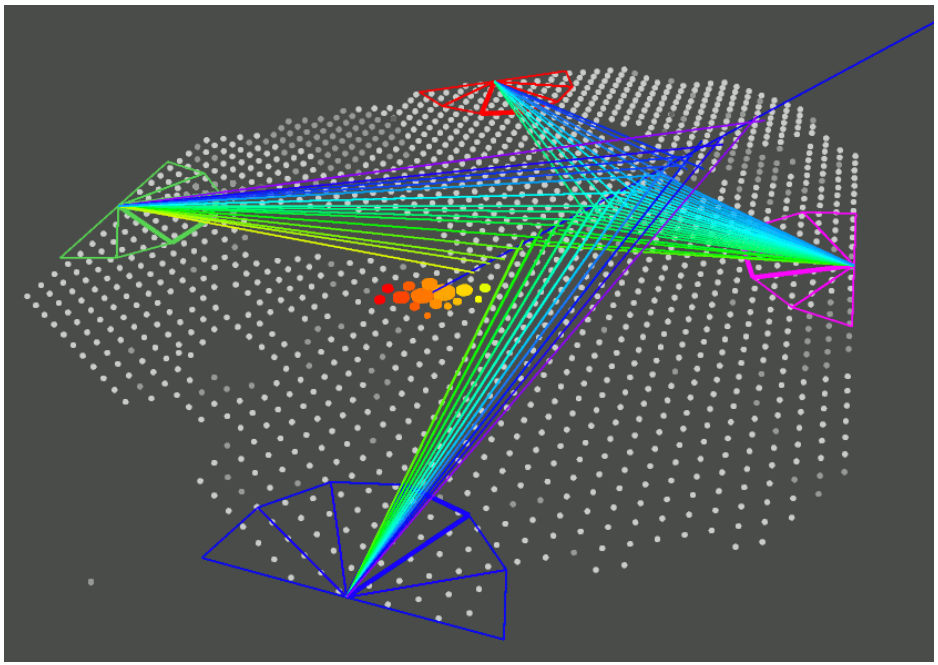
The hybrid detection

A key feature is the hybrid design, that allows the measurement of an EAS with both the surface and fluorescence detectors and thus an accurate reconstruction of the atmospheric shower development.

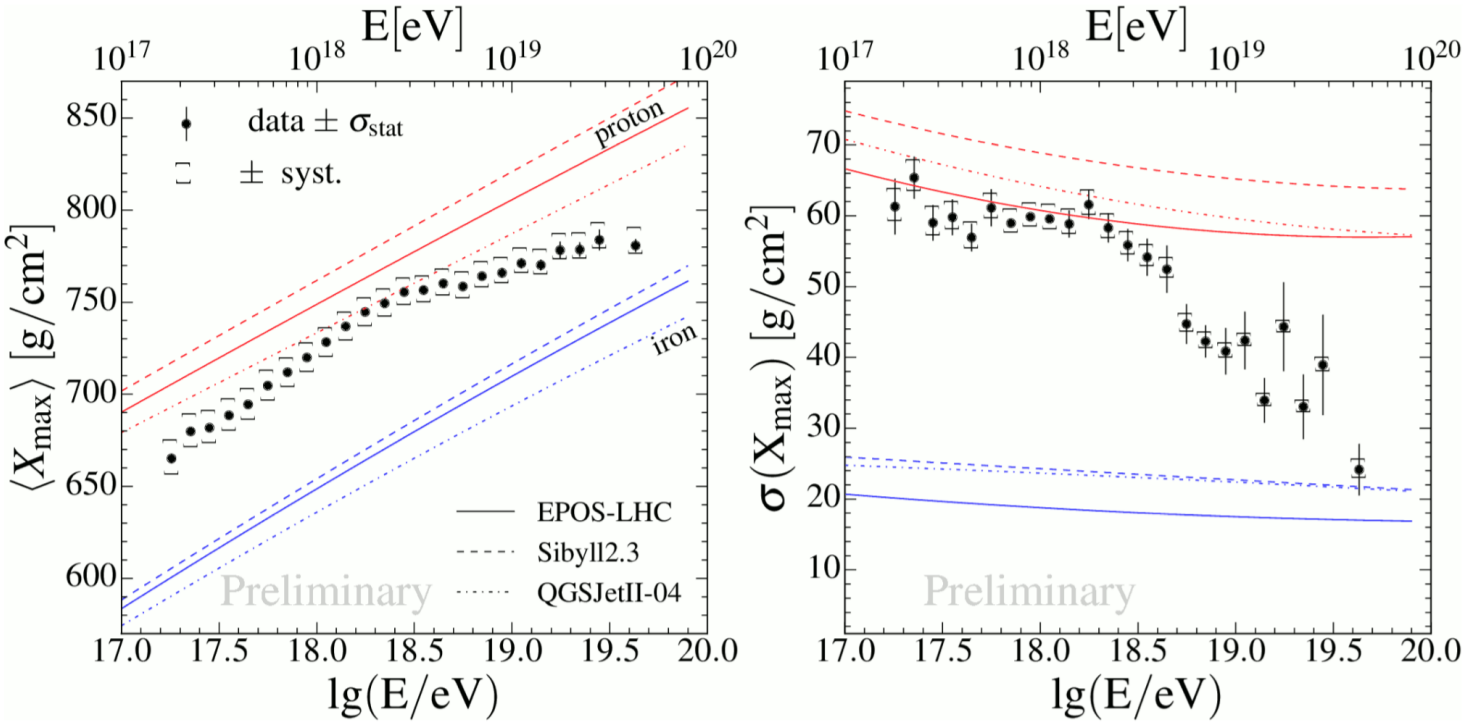
In particular the **depth of the shower maximum** (X_{max}) can be employed in mass composition analysis, since:

$$X_{\mathrm{max}} \propto E/A$$

as obtained for the *superposition* model.



$X_{\max}$ moments



Presented at the 35th ICRC (2017) by J. Bellido.

The first two moments of the $X_{\max}$ distribution reflect different aspects of the underlying distribution of primary masses.

Limitations of the current mass composition studies

The analysis based on the $X_{\max}$ moments presents strong limitations, since the duty cycle of fluorescence telescopes is considerably short (around 10-15%).

As a consequence, the number of events that can be collected above $\sim 10^{19.5}$ eV is not sufficient to drawn conclusive results.

Therefore independent methods which exploit observables available through surface arrays (with their approximately 100% duty cycle) have been studied over the years.

Following this idea, an innovative method based on the paradigm of air shower Universality has been recently proposed in the framework of the Pierre Auger Collaboration.

Part II
Air shower universality

The concept of air shower universality

The basic idea is that extensive air showers are *similar* when observed at the same stage of development.

More specifically the energy, angular and lateral distributions of (low-energy) shower particles can be determined to a remarkable degree of precision only through an **age parameter**[1], such as:

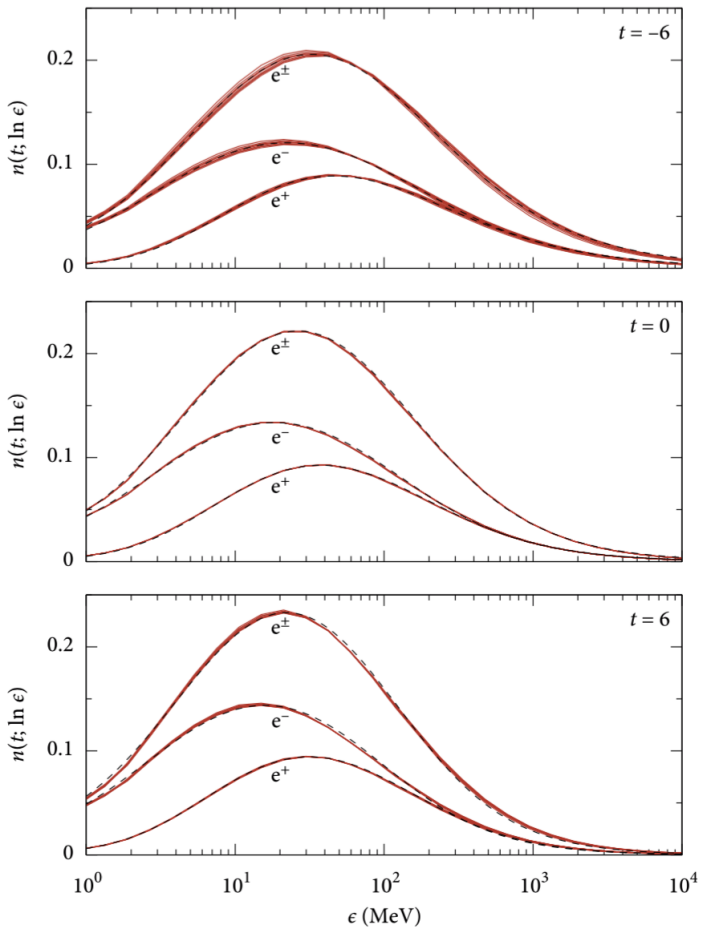
$$s = \frac{3X}{X + 2X_{\max}} \qquad \text{or} \qquad t = \frac{X - X_{\max}}{X_0}$$

where X is the slant depth traversed in the atmosphere, $X_{\max}$ is the depth of the maximum in the development of the e.m. part, $X_0 = 36.7$ g/cm^2 is the radiation length of electrons in air.

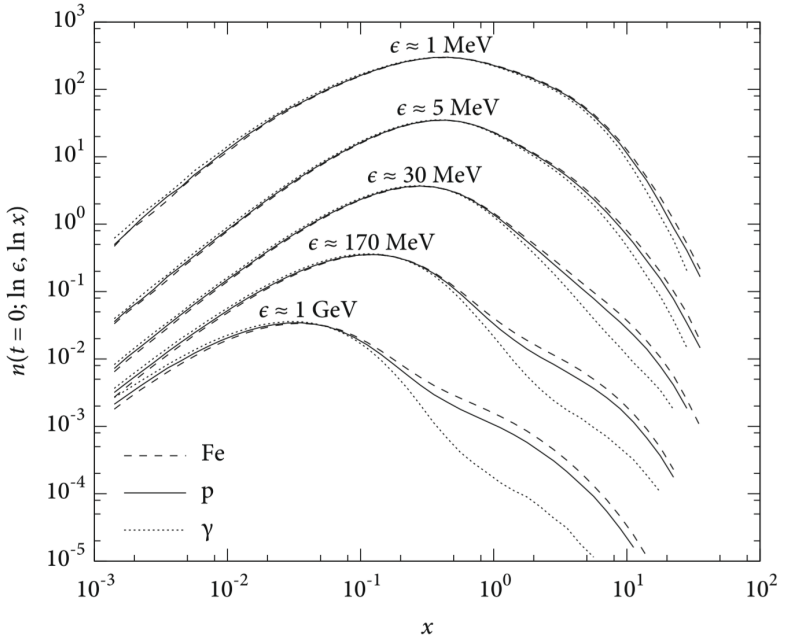
[1] A critical review and a theoretical interpretation of such result is given in *P. Lipari, Phys. Rev. D79 (2009)*.

Universality of the e.m. component

Energy distributions



Lateral spread distributions



Universal behaviours for showers initiated by different species (γ, p, Fe) at different energies.

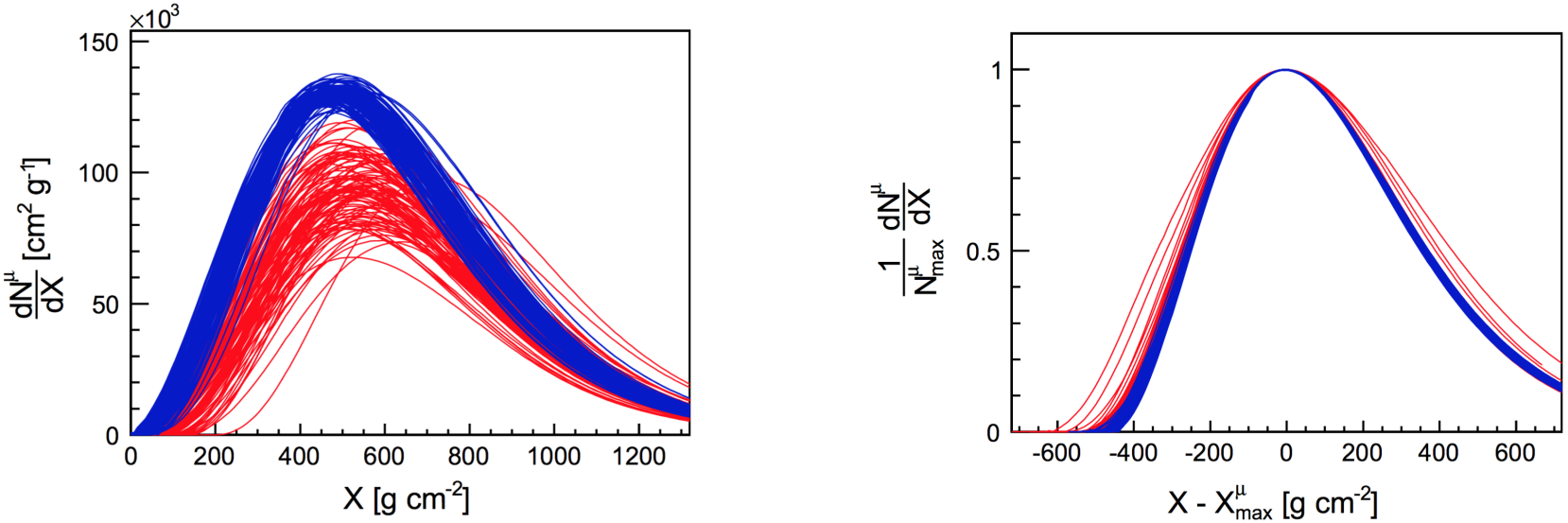
S. Lafebre et al., Astropart. Phys. 31 (2009).

Universality of the muonic component

A rather universal behaviour is found also for the muonic component when the development is described through:

$$X' = X - X^{\mu}_{\max} \qquad N'_{\mu} = N_{\mu}/N^{\max}_{\mu}$$

where X' acts as an age parameter and a renormalization of the total number of muons N_{μ} is accomplished.



Muon production depth profiles for p and Fe-initiated showers.
S. Andringa et al., Astropart. Phys. 35 (2012)

Universality of ground signals

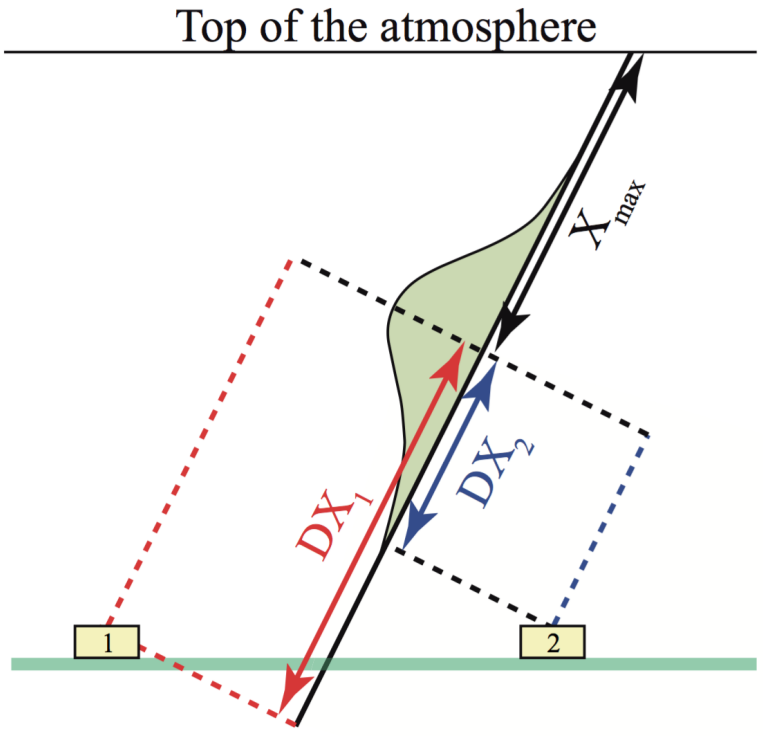
Thus an EAS can be characterised defining only its energy, $X_{\max}$ and a quantity R_μ related with the number of muons.

The next step was to extend the Universality concept to the signal measured by ground-based detectors, using as age parameter

$$DX = X_{\mathrm{gr}} \sec\theta - X_{\max}$$

where X_{gr} is the vertical depth of the detector and θ the zenith angle of the shower.

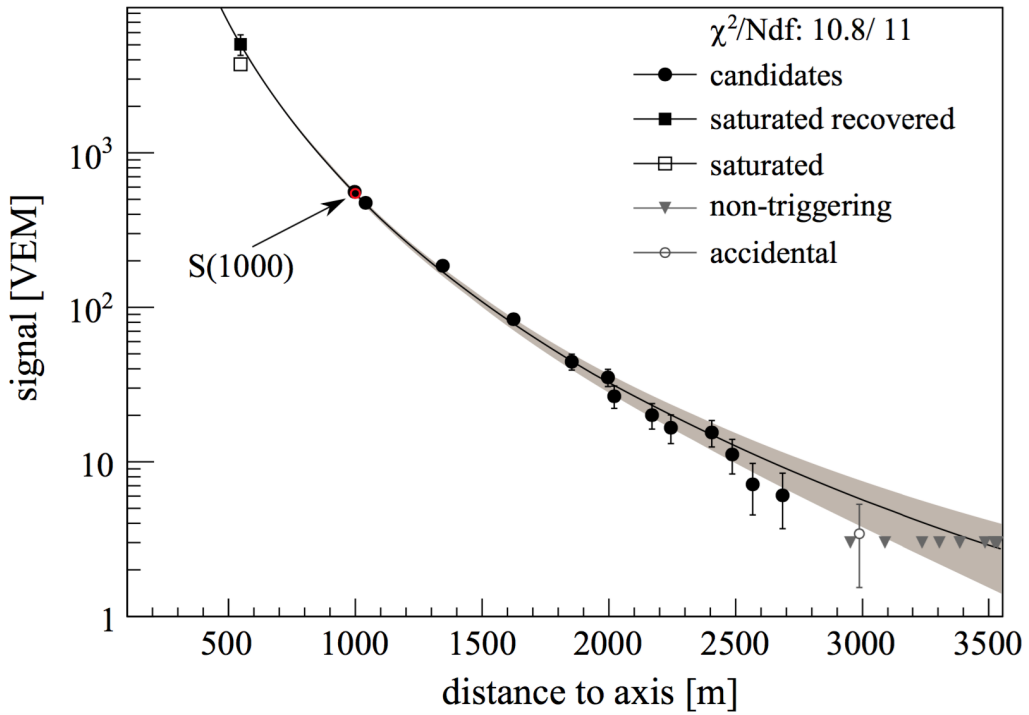
A. S. Chou et al., Contributions to the 29th ICRC (2005)



M. Ave et al., Contributions to the 32th ICRC (2011)

The lateral distribution function

The integrated signals measured by ground-based detectors are analysed as functions of the distance from the shower core to reconstruct the **lateral distribution function** (LDF).

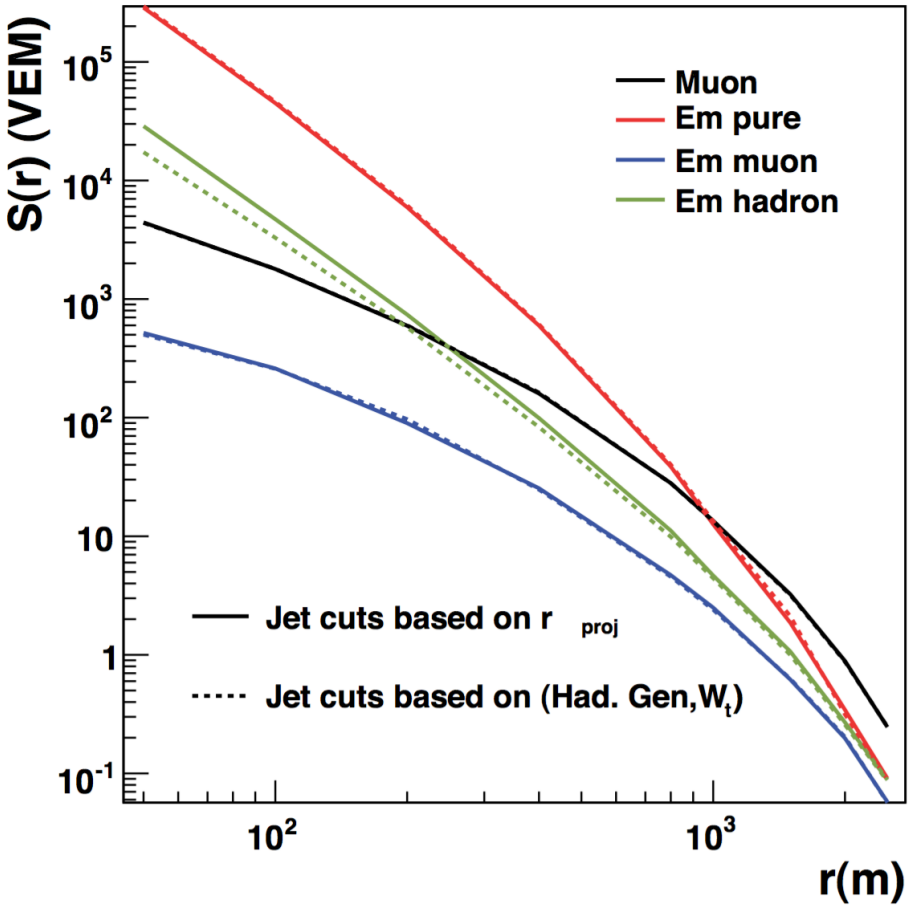


Then the shower **energy** is evaluated using as *estimator* the signal at an optimal distance, which is obtained interpolating from the fitted LDF.

The 4-components shower Universality

A robust parameterisation of the shower development, driven by simulations, requires 4 signal components:

- a purely electromagnetic component;
- a muonic component;
- an electromagnetic component from muon interactions and decay;
- an electromagnetic component from low-energy hadrons (jet component).

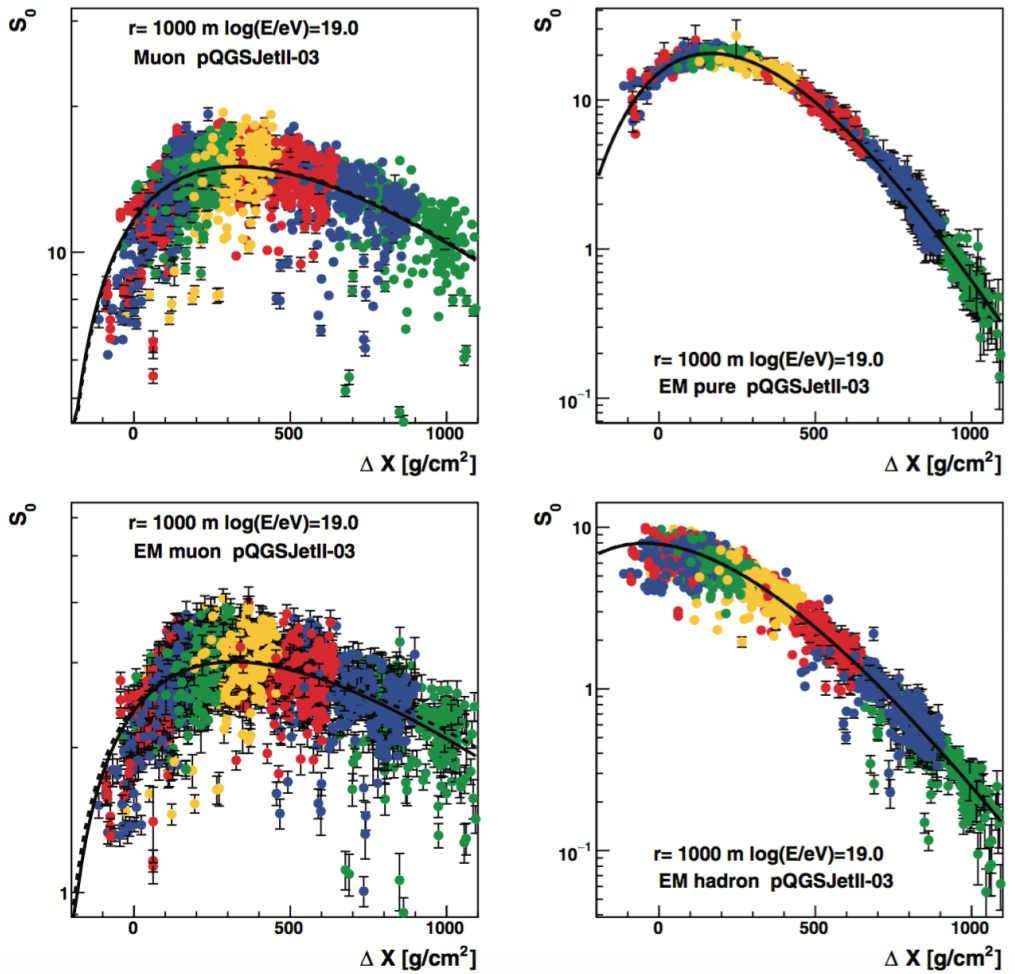


Parameterization of the integrated signal

$$S^i(r,\psi|E,X_{\max},N_\mu,\theta,\rho^{air}) = S_0^i(r,\Delta X|E) \times f^i_{fluct}(r|N_\mu) \times f^i_{\mathrm{mod}}(r,\psi|\theta) \times f^i_{\mathrm{atm}}(r|\rho^{air}) \times f^i_{\mathrm{conv}}(r,\Delta X,\psi|\theta)$$

The signal S_0 measured by an **ideal detector** is parameterized as a Gaisser-Hillas function in the variable ΔX $(= DX)$.

M. Ave et al.,
Astropart. Phys. 87 (2017)

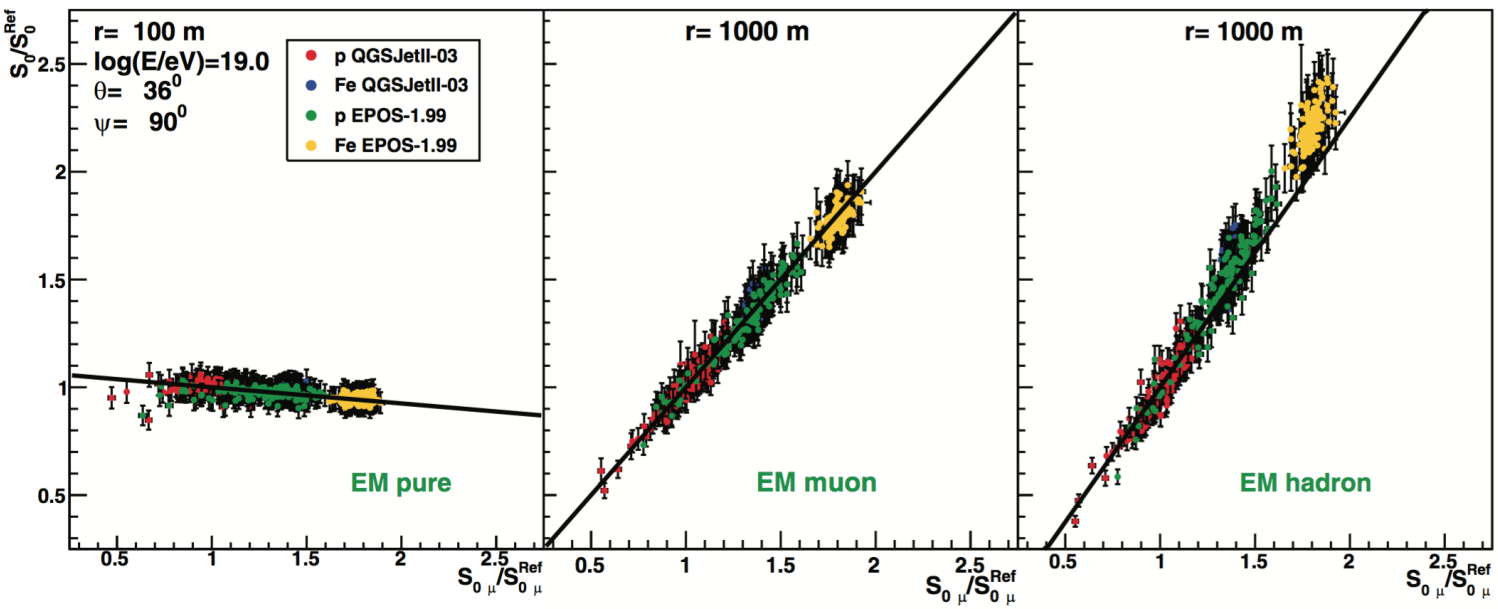


Parameterization of the integrated signal - N_μ

The $f_{\rm fluct}$ factor accounts for primary composition (and hadronic interaction model) effects.

$$N_\mu = \frac{S_{0,\mu}(r = 1000\ m, \psi = 90^\circ | \Delta X, E)}{S^{\rm ref}_{0,\mu}(r = 1000\ m, \psi = 90^\circ | \Delta X, E)}$$

$$f^i_{\rm fluct} = 1 + (N_\mu - 1)\ \alpha_i \qquad \frac{S^i_0}{S^{{\rm ref},i}_0} = 1 + \left(\frac{S^\mu_0}{S^{{\rm ref},\mu}_0}\right) \alpha_i$$



Universality of the time-dependent signal

A natural extension for the Universality of the integrated signals is a parameterization of the signal shape as a function of time (*D. Maurel et al., Contributions to the 33th ICRC (2013)*).

The model was developed making an ansatz directly for the signal time structure observed in data or simulations:

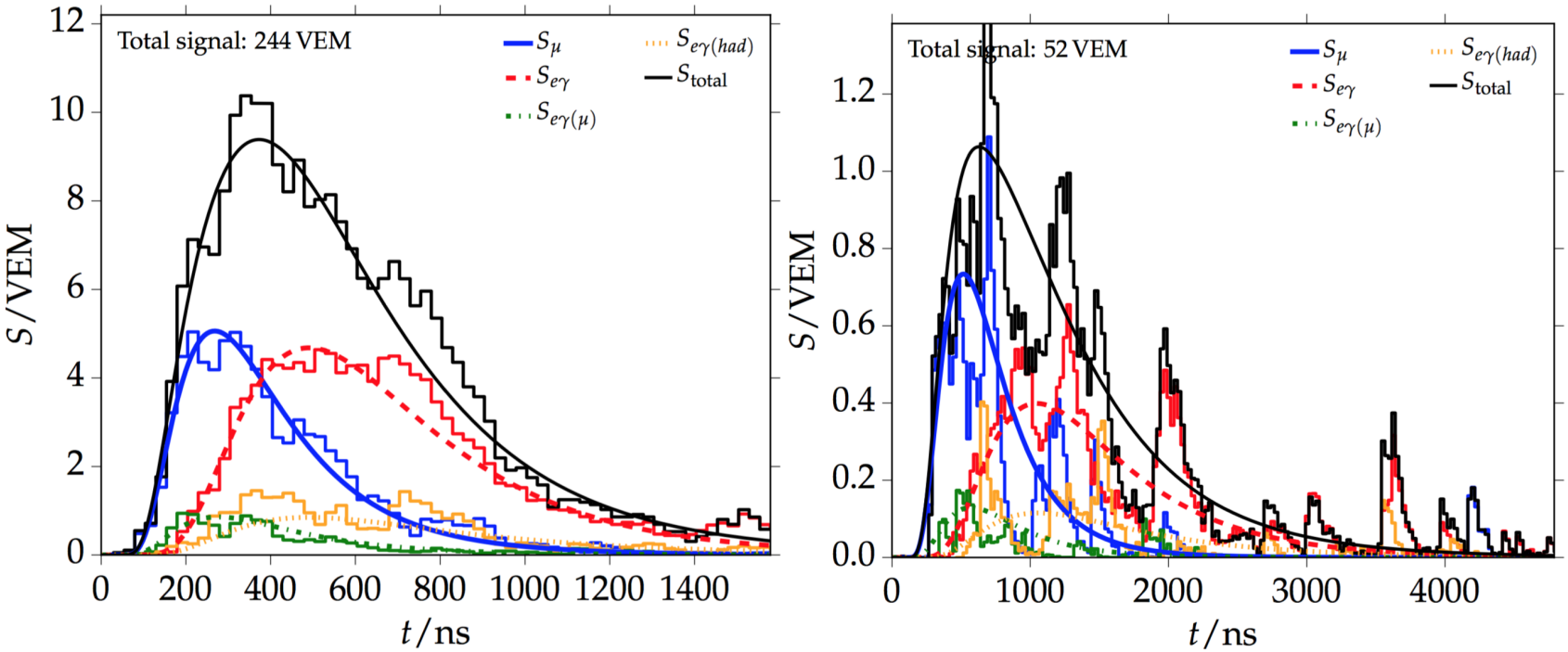
$$\frac{dS_i}{dt}(t) = \frac{1}{\sqrt{2\pi}(t - t_0)s_i} \exp\left[-\frac{(\ln(t - t_0) - m_i)^2}{2s_i^2}\right]$$

that is a **log-normal distribution** (with parameters m and s) where t_0 is the time below which no particles are expected.

Updated studies: *M. Ave et al., Astropart. Phys. 88 (2017)*.

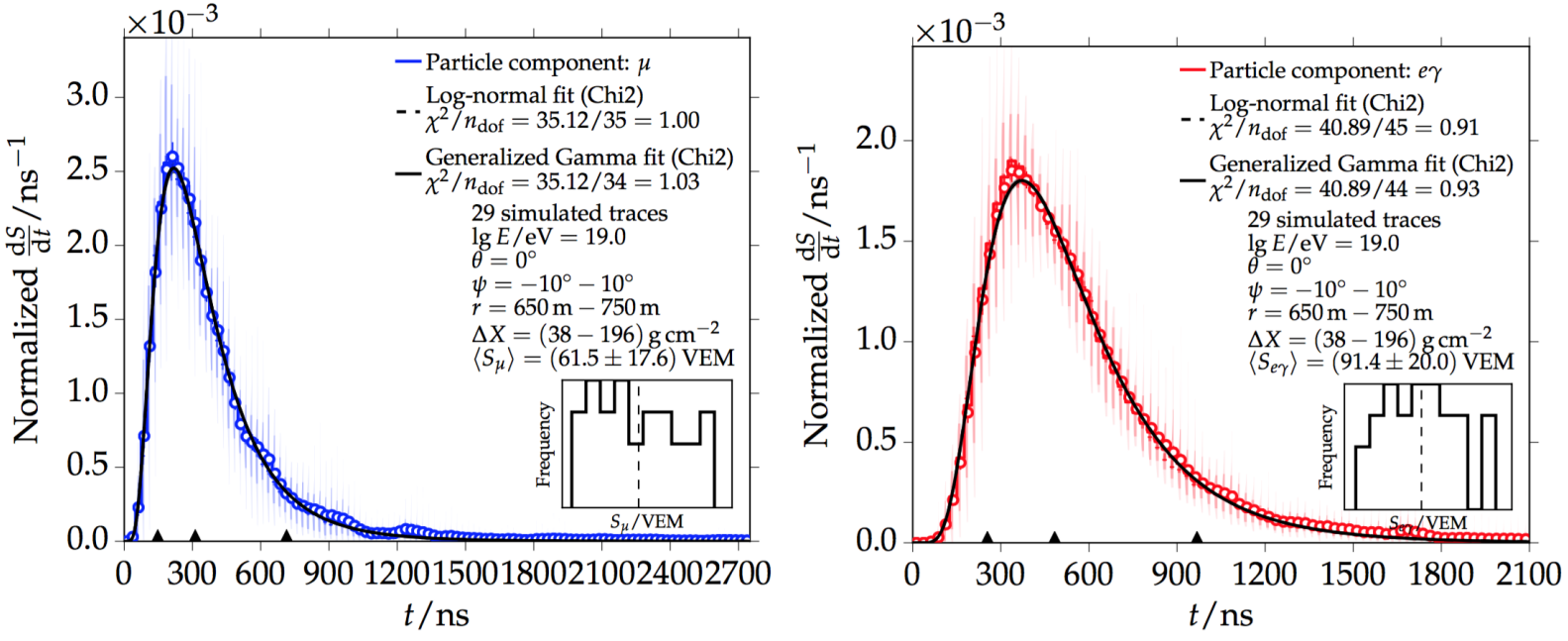
Simulated traces in Universality components

The traces are simulated in bins of 25 ns width (equal to the WCDs integration time) using the un-thinning procedure.
To built the parameterizations, signals averaged in suitable $(r, \theta, \psi, E, \Delta X)$ bins have been employed.



Parameterization of the time-dependent signals

1. Model for the origin of times $D_{T0} = D_{T0}(E, X_{\max}, \theta)$.
2. Using the t_0 parameterization, study of the log-normal parameters (m, s) for the average traces.
3. Parameterizations of (m, s) as functions of $(r, \theta, \psi, E, \Delta X)$.



A. Schulz, PhD thesis, 2016 (GAP 2016-021)

Part III
Universality reconstruction

The Universality reconstruction

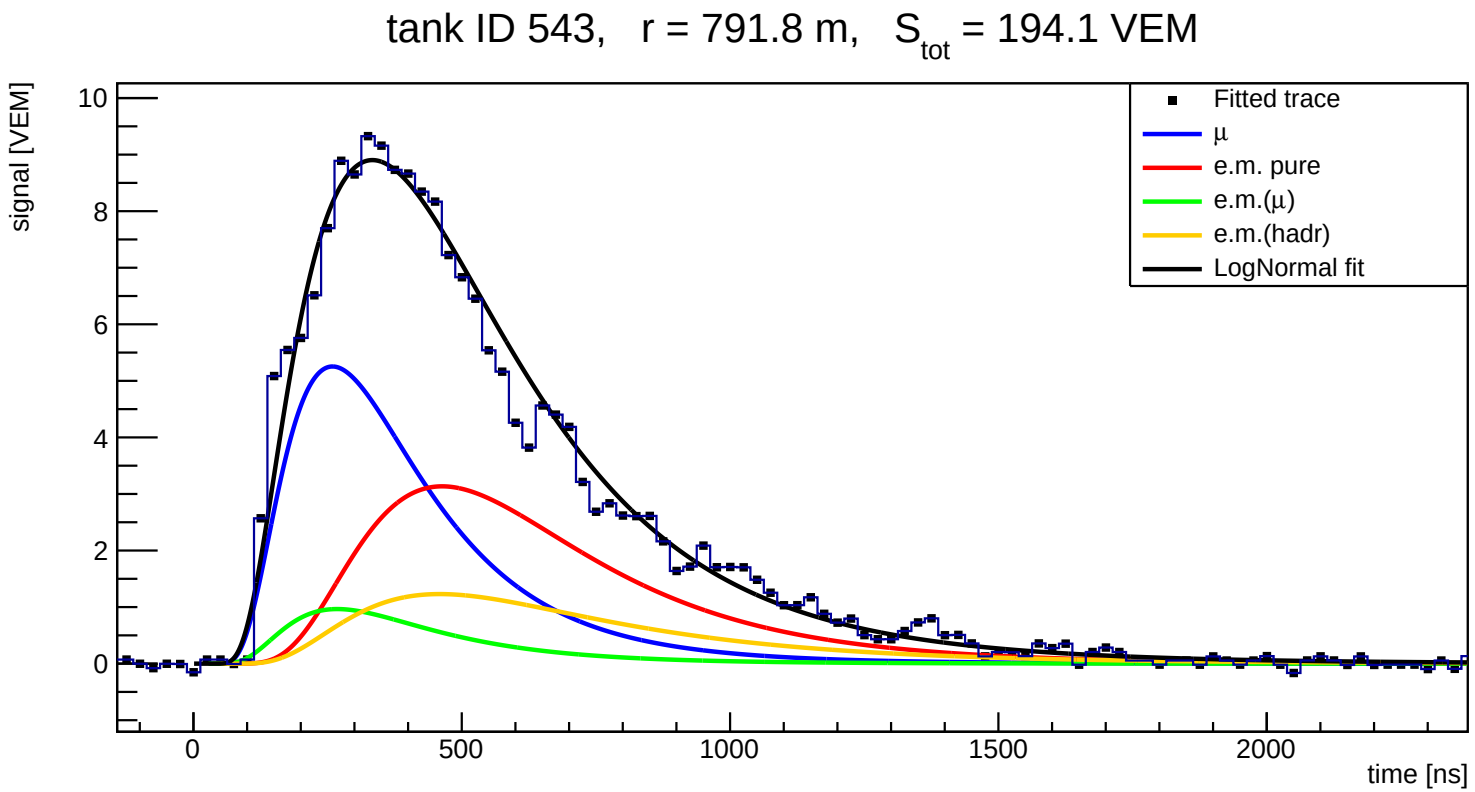
The Universality reconstruction consist in a search for the set of macroscopic shower parameters $(E, X_{\max}, N_\mu, \theta, \psi, \vec{x}_{\text{core}}, t_{\text{core}})$ which better reproduce the signals gathered by the WCDs.

In the *Karlsruhe* **global** version of the reconstruction, $(X_{\max}, N_\mu, \theta, \psi, \vec{x}_{\text{core}}, t_{\text{core}})$ are fitted simultaneously while the **energy E is fixed** to the value from the standard SD fit[2].

The true innovation of this reconstruction is the **shape fit**, that is a bin-to-bin fit of the traces performed using the sum of 4 Log-Normal functions, one for each shower component.

[2] In fact, a complete and successful standard SD reconstruction is required before performing the Universality fit.

Shape fit example



Fit result for a single trace from a real event.
The coloured lines are the 4 log-normal components obtained from the Universality reconstruction; in black, the sum.

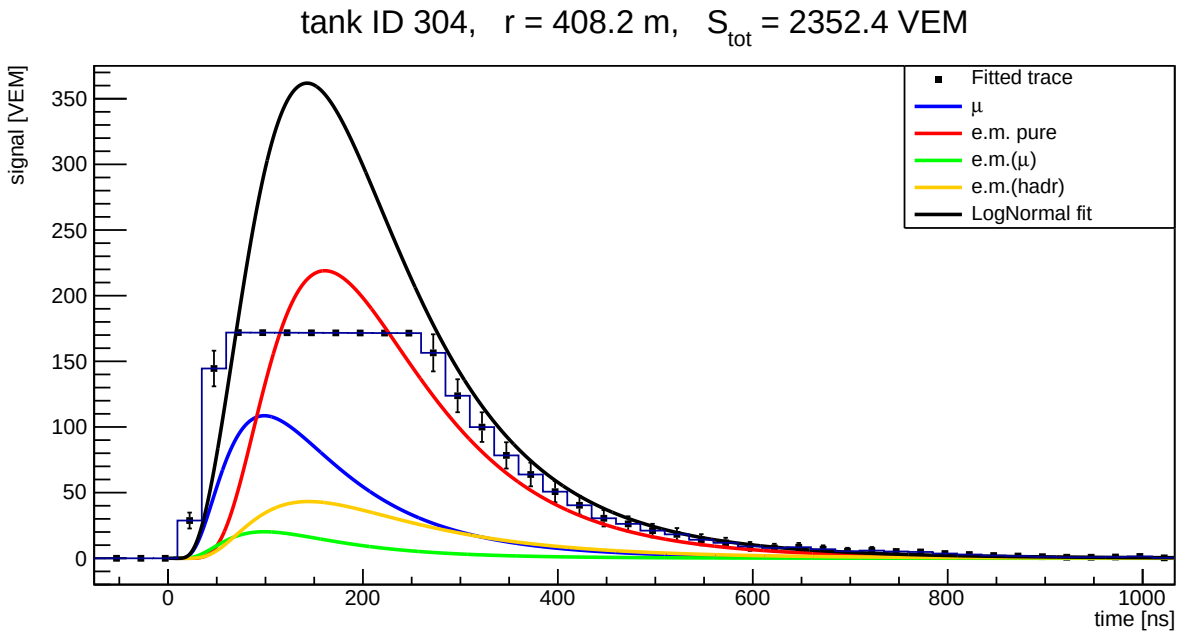
Stages of the analysis

1. The Universality fit is firstly applied on a selected set of **golden** hybrid events, used as benchmark (event-by-event comparison) to improve the reconstruction accuracy.
2. Then the set of (vertical) events with energy larger than 10^{19} eV measured by the **SD-1500** up until the end of 2016 is reconstructed according to an optimized Universality fit.
3. Finally the **mass composition** is obtained from the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distribution analysis over a dataset larger than the hybrid one (between 10 and 3 times, as a function of the energy).

Only **unsaturated** events have been used so far, since the reconstruction is less accurate when saturated signals are included.

Saturated signal

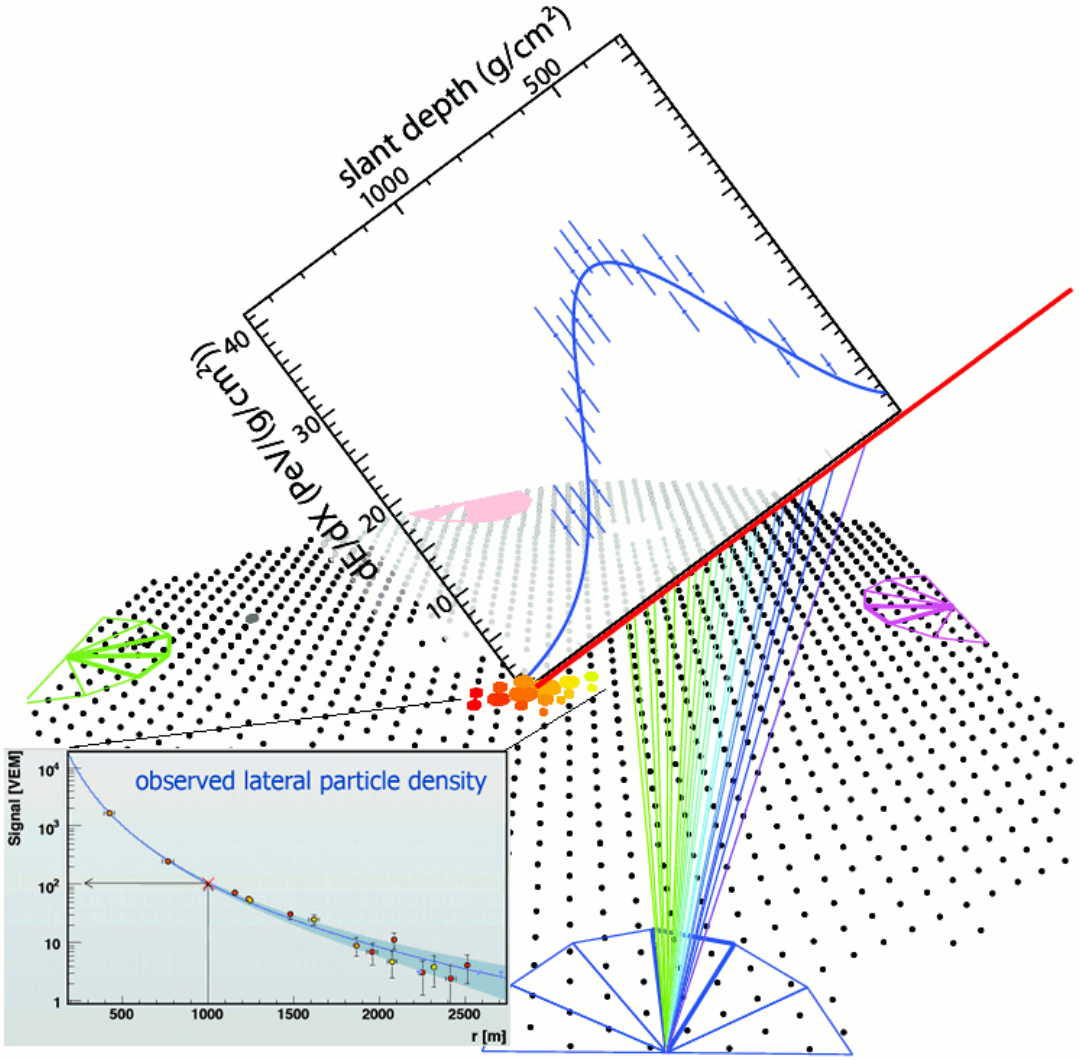
A saturated trace provides much less information, since the time bins with saturated signals cannot be used in the shape fit.



The saturation verifies when a large number of particles reach the detector in a short amount of time (usually near the shower core), causing an overflow of the FADC read-out electronics.

Part III.1
Reconstruction of golden hybrid events

The selected set of golden hybrids



Golden hybrid events present both standard SD and hybrid successful reconstructions.

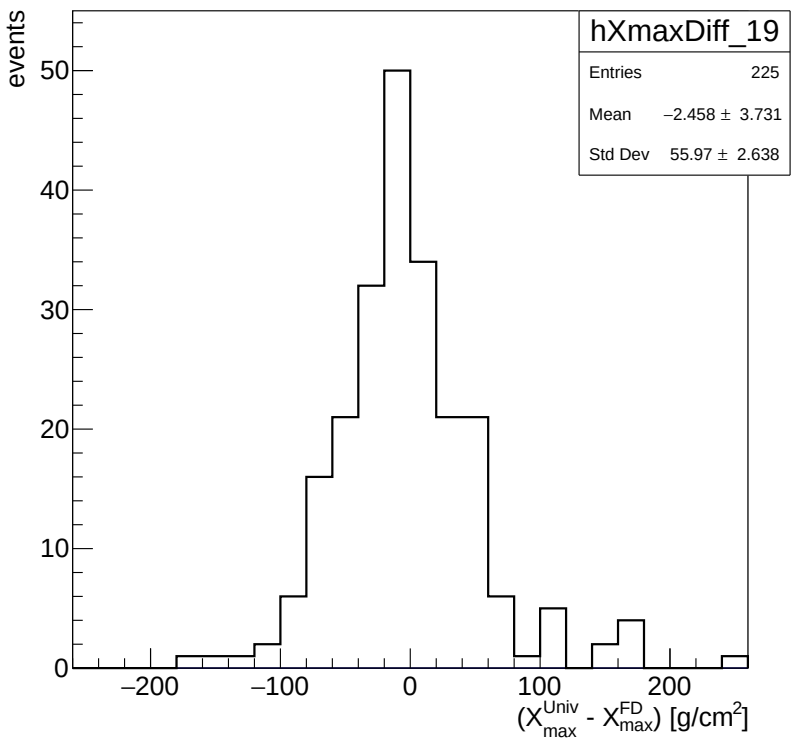
The golden hybrids for the **ICRC-2015** energy calibration of the vertical ($\theta < 60°$) events have been used to evaluate the performances of the Universality fit.

Outcome of the original reconstruction (*UnivKG*)

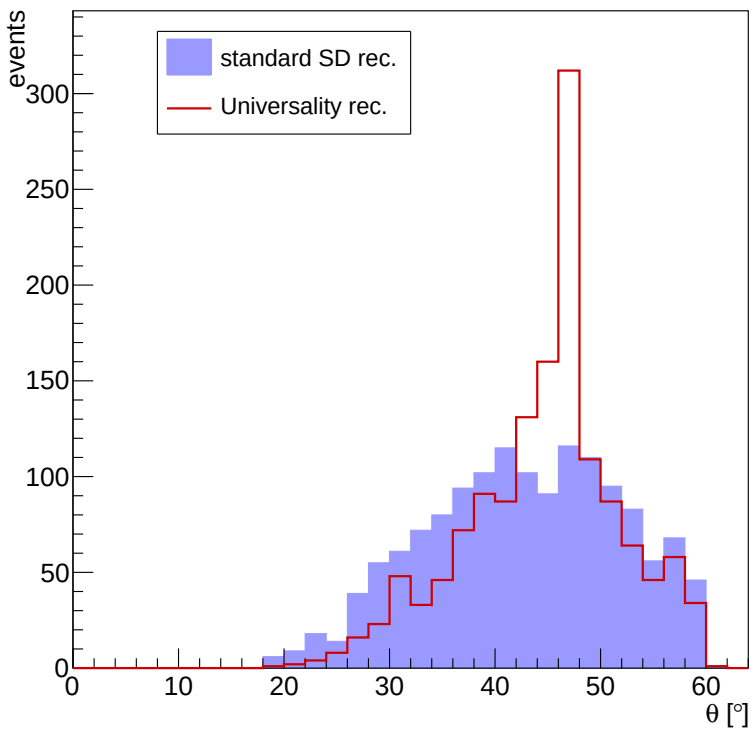
225/228 unsaturated events with $E_{SD} > 10^{19}$ eV reconstructed.

$\left\langle X_{\rm max}^{\rm Univ} - X_{\rm max}^{\rm FD} \right\rangle$ compatible with zero.

$\mathrm{RMS}(X_{\rm max}^{\rm Univ} - X_{\rm max}^{\rm FD}) = 56.0$ g/cm^2.



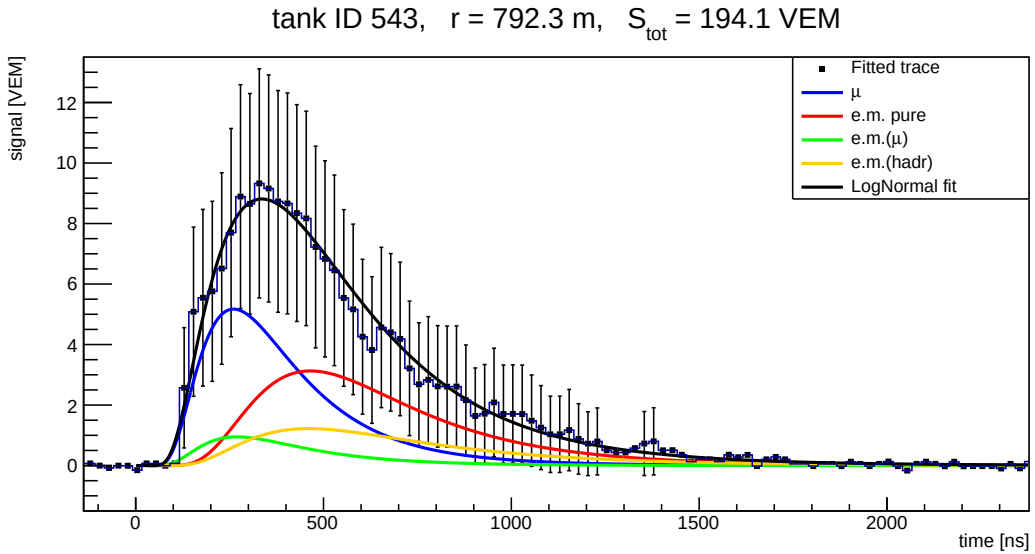
Zenith angle distribution

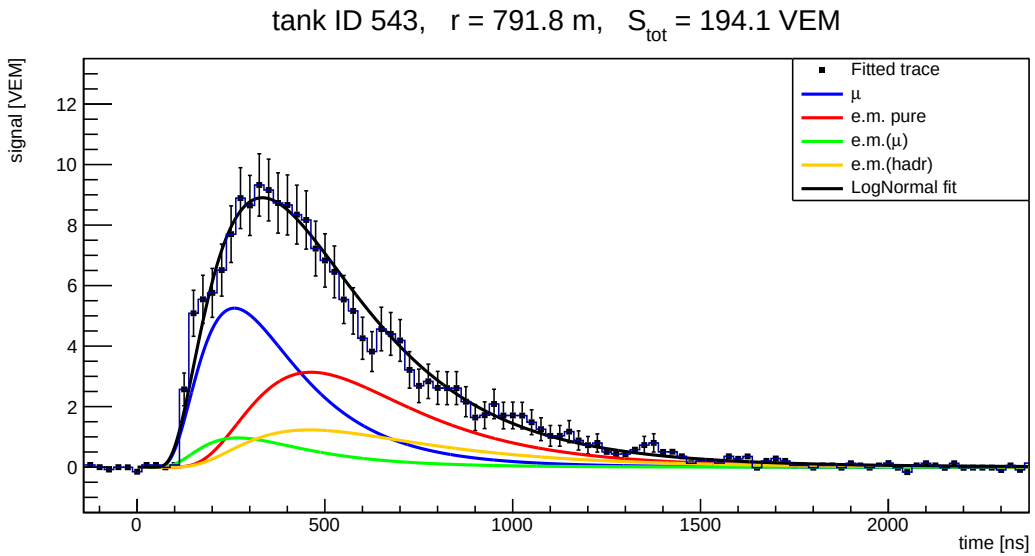


Adjusting the fitting procedure

1. The minimization function is changed from a negative Log-Likelihood into a **deviance**, a quantity equivalent to a generalized χ^2, to be used as a goodness-of-fit estimator.
2. From the analysis of the deviances (D/NDoF), an **overestimation** of the errors assigned for the shape fit to the signal in each time-bin of the traces is evident.
3. A model for the generation of mock traces (**cluster model**) is realized to evaluate the bin content fluctuations.
4. The **bin signal uncertainties** are reduced using a constant factor obtained from the study of a large set of mock traces.

Redefinition of the bin content uncertainty





"Quasi-poissonian"

$$\epsilon(s_i) = f_S \times \sqrt{s_i}$$

Reduced (after cluster model studies)

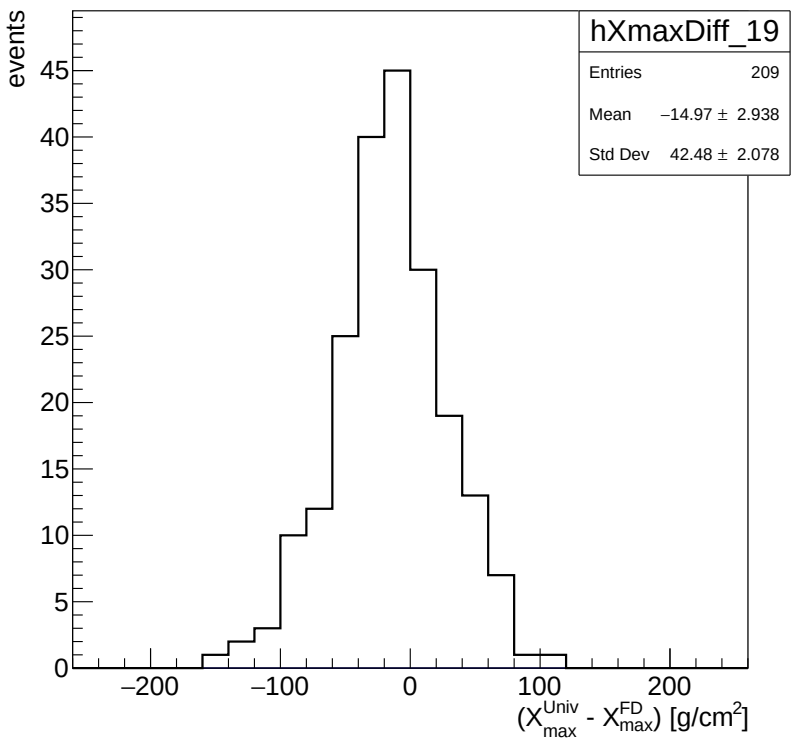
$$\epsilon(s_i) = 0.42\ f_S \times \sqrt{s_i}$$

Reconstruction with deviance and reduced errors

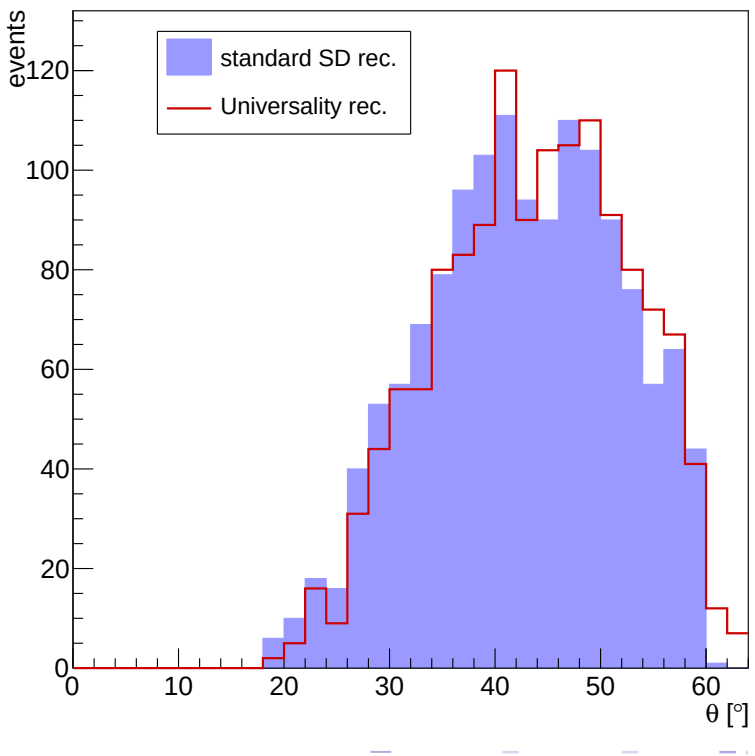
209/228 unsaturated events with $E_{SD} > 10^{19}$ eV reconstructed.

$\left\langle X_{\max}^{\mathrm{Univ}} - X_{\max}^{\mathrm{FD}} \right\rangle = -15.0$ g/cm^2 (negative bias).

$\mathrm{RMS}(X_{\max}^{\mathrm{Univ}} - X_{\max}^{\mathrm{FD}}) = 42.5$ g/cm^2 (smaller dispersion).



Zenith angle distribution



A different paradigm for the Universality reconstruction

Pre-existent assumptions in the fitting procedure :

- **energy fixed** to the SD value, while the other physical quantities are fitted inside large intervals;
- uncertainties on the geometrical parameters (obtained in the standard SD reconstruction) never taken into account.

Then a new fitting procedure (named **constrained Universality reconstruction**) is attempted, where:

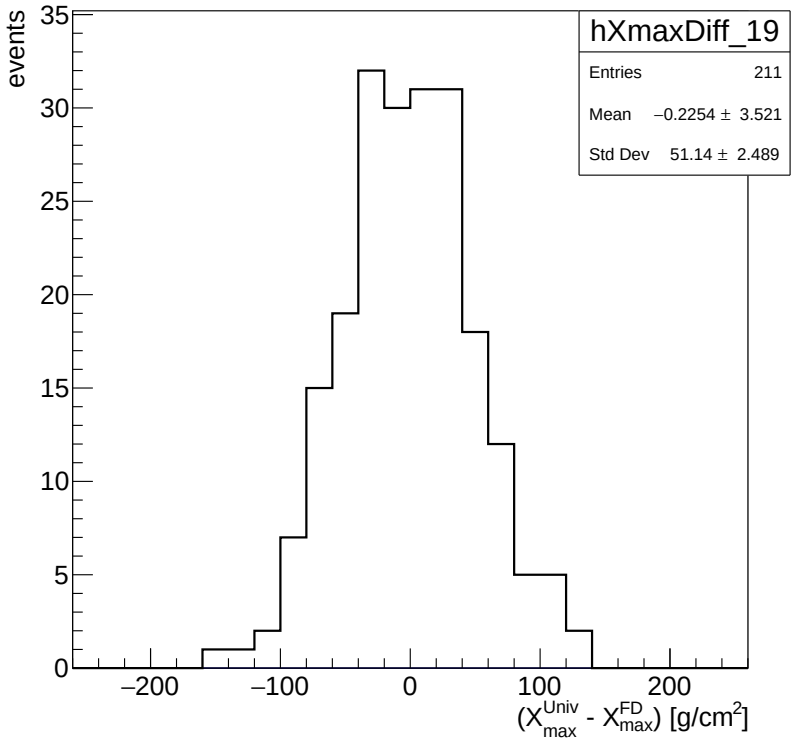
- all the parameters already determined in the standard SD reconstruction (i.e. axis, core and **energy**) are **fitted**;
- constraints around the standard SD outcomes are introduced, exploiting the known reconstruction resolutions.

Outcome of the constrained reconstruction

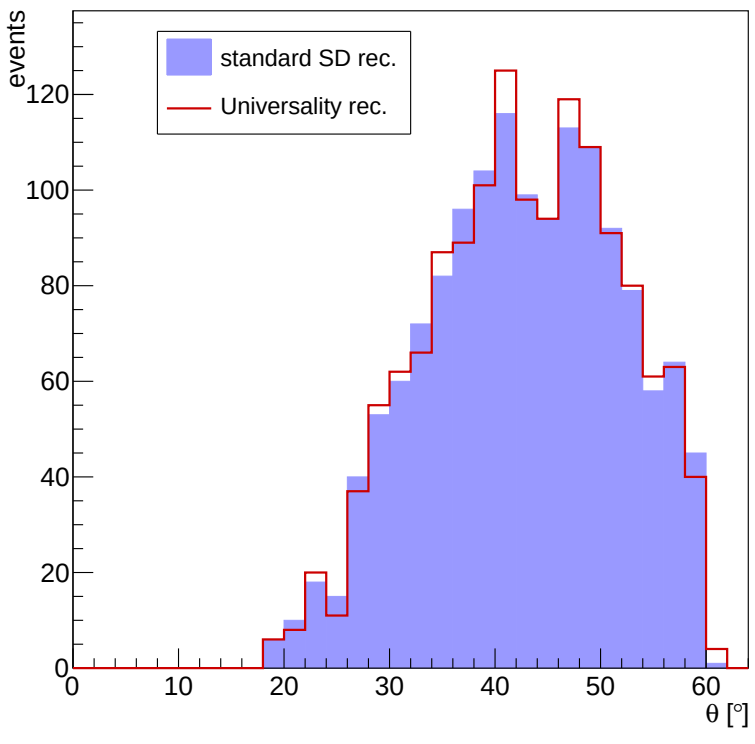
211/228 unsaturated events with $E_{SD} > 10^{19}$ eV reconstructed.

$\left\langle X_{\rm max}^{\rm Univ} - X_{\rm max}^{\rm FD} \right\rangle$ again compatible with zero.

$\mathrm{RMS}(X_{\rm max}^{\rm Univ} - X_{\rm max}^{\rm FD}) = 51.1$ g/cm^2 (again large).



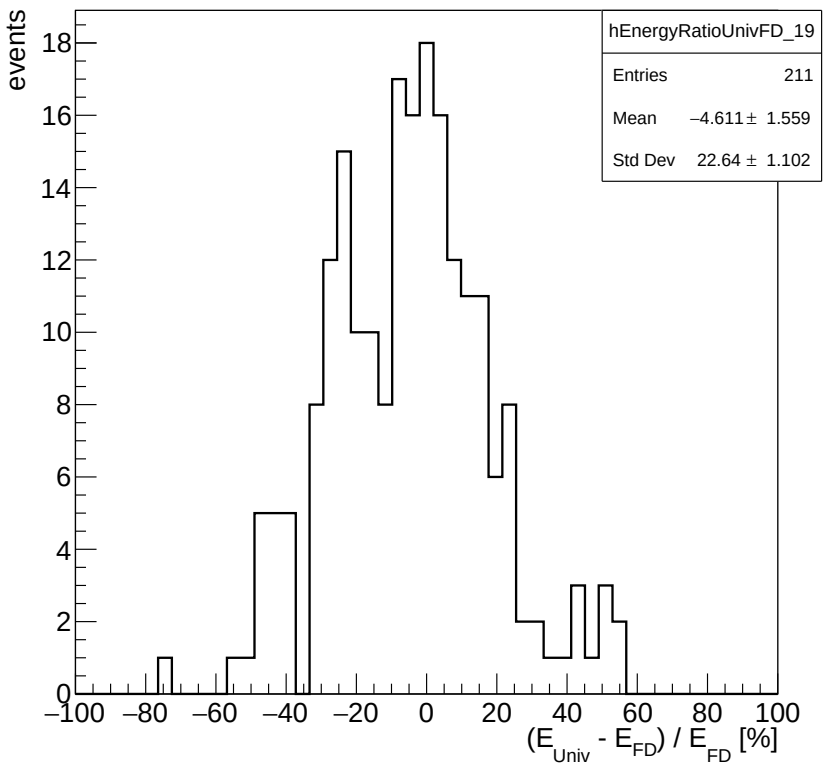
Zenith angle distribution



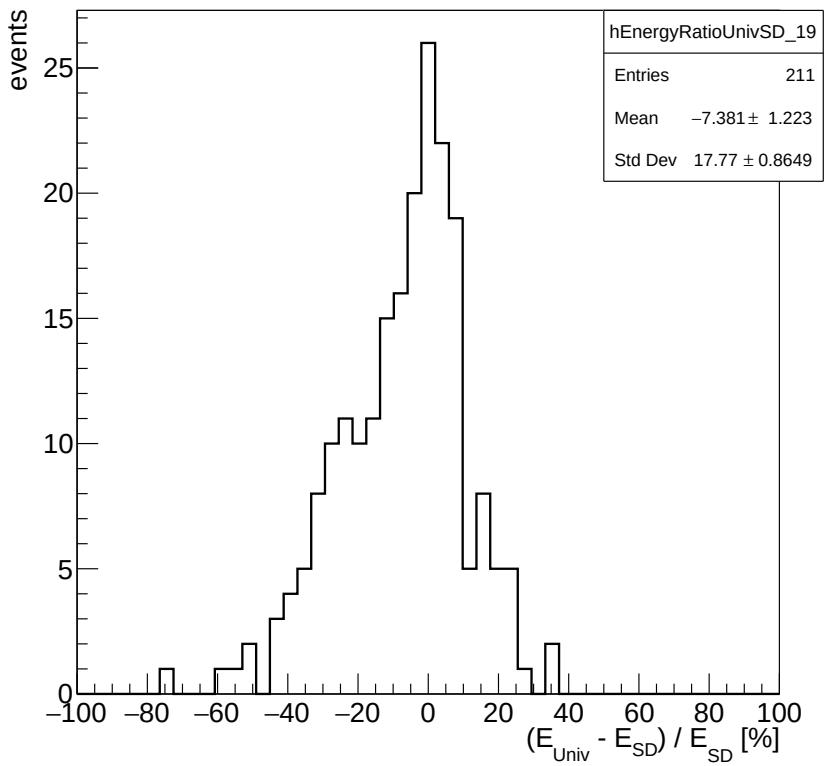
Constrained reconstruction - energy differences

Negative biases are apparent in the distribution of the ratios between E_{Univ} and the correspondent SD and FD determinations.

Universality vs FD

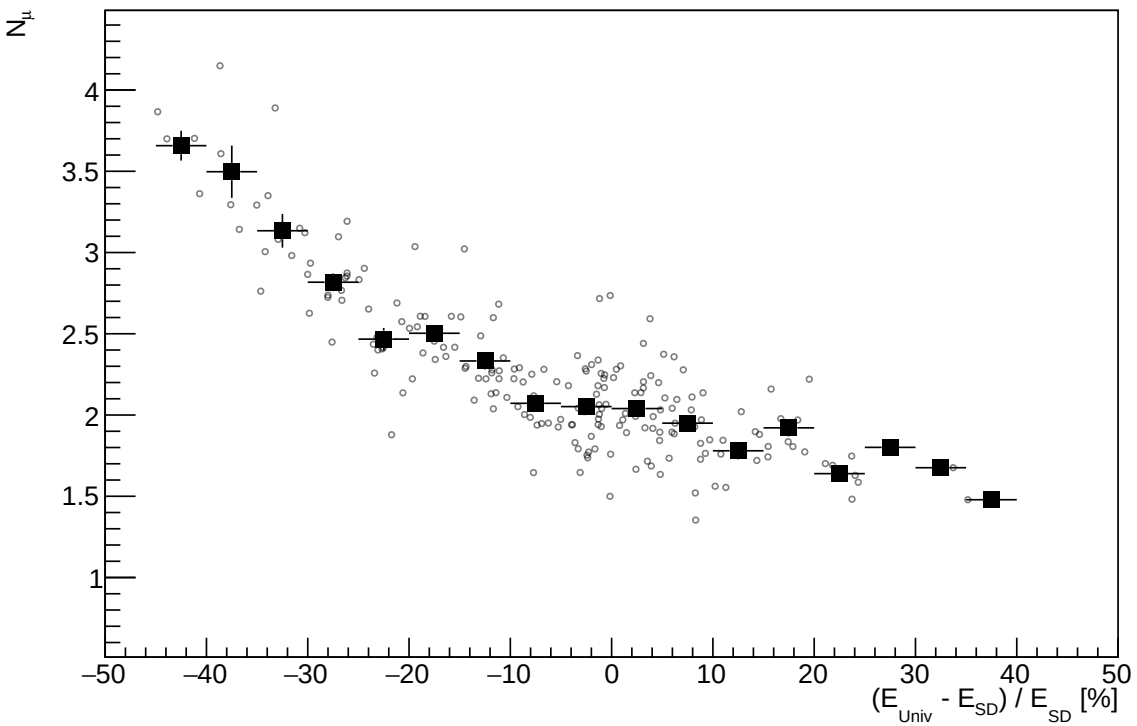


Universality vs SD



Degeneracy of the Universality reconstruction

Possible explanation: degeneracy of the Universality fit due to the **(anti-)correlation between energy and muon content**.



Events with larger fitted values of N_μ have a final energy E_{Univ} smaller than E_{SD}, and viceversa.

A simplified Universality reconstruction

Conclusion (at least for the current Universality model):
the information from only the WCDs surface array in not sufficient to determine at the same moment the energy and the other shower parameters (in particular N_μ).

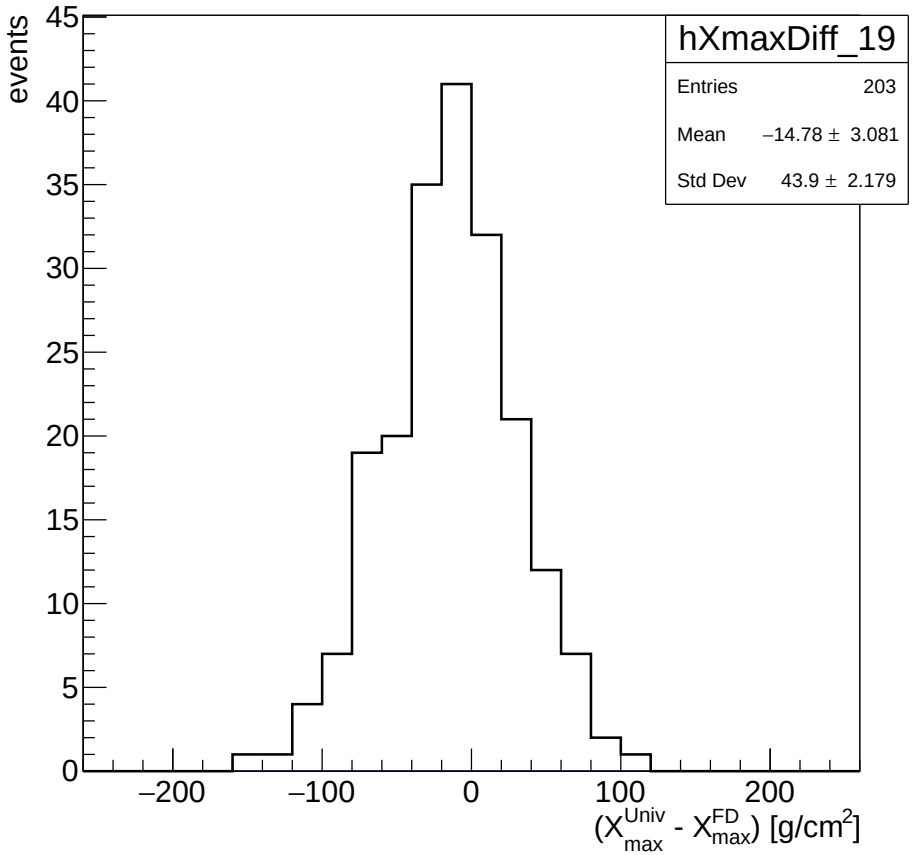
Considering such issue, a **simplified** but faster **fitting procedure** is finally introduced, where:

- the **energy** is again **fixed** to E_{SD};
- the other previously introduced changes (deviance, reduced bin content uncertainties, core constraints) are maintained;
- the **axis** is **fixed** to the standard SD determination.

Outcome of the simplified reconstruction

203/228 unsaturated events with $E_{SD} > 10^{19}$ eV reconstructed.

Not negligible **bias** in the $(X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}})$ distribution.



Internal tensions in the current Universality model (for instance limits in the validity of the parameterizations as well as inaccuracies in simulations/resampling) are the most likely underlying cause.

Part III.2
Reconstruction of the SD-1500 dataset

The SD-1500 dataset

Events with $\theta < 60^\circ$ and energies above 10^{19} eV measured by the SD-1500 between 1st January 2004 and 31st December 2016. (*F. Fenu, Contribution to the 35th ICRC (2017)*)

To assure a good quality in the final outcomes of the Universality reconstruction, additional event selections and quality cuts (applied after the fit, to get rid of *outliers*) are necessary.

cuts	SD-1500 $E_{\mathrm{SD}} > 10^{19}$ eV
ICRC-2017	13966 (100%)
unsaturated	9767 (69.9%)
$\theta < 55^\circ$	8691 (89.0%)
$\theta > \theta_{\mathrm{cut}}(E)$	6854 (78.8%)
fit converged	6115 (89.2%)
$\hat{S}^{\mathrm{hottest}}_{\mathrm{e.m.pure}} > 1$ VEM	6064 (99.2%)
$D_{\mathrm{tot}}/N_{\mathrm{tot}} < 4$	6050 (99.8%)
surviving	6050 (43.3%)

Low zenith angle cut

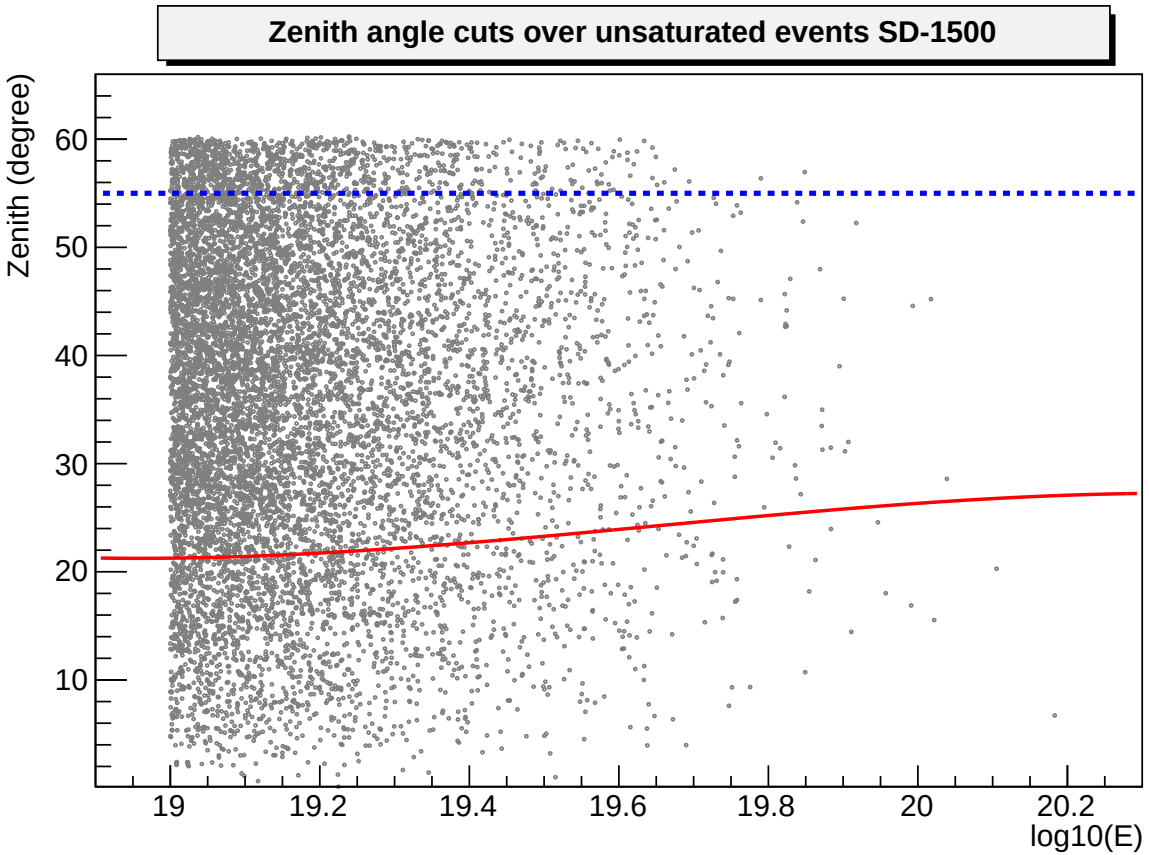
A known issue of the present Universality model is that only simulations where the **maximum** is reached **above the ground level** ($\sim$ 875 g/cm^2) have been used to build the parameterization.

As a consequence, the reconstruction of events with **deeper** $X_{\rm max}$ (that is for a lighter primary and/or a lower zenith angle, at a fixed energy) is less reliable and gives rise to a bias[3].

To un-bias the $X_{\rm max}^{Univ}$ result for the following analysis of the SD-1500 dataset is necessary to introduce a selection over the zenith angle, similarly to the **fiducial field of view selection** performed in the standard hybrid $X_{\rm max}$ analysis.

[3] Such bias is different from what observed in the $(X_{\rm max}^{Univ} - X_{\rm max}^{FD})$ distribution for golden hybrids.

Field of view selection



$\Downarrow \quad \theta < 55^\circ$

Events employed in this analysis

$\Uparrow \quad \theta > \theta_{\mathrm{cut}}(E)$

Events (black points) above the **low zenith cut** (red line) have more than 90% probability to develop with X_{max} above ground.

Limits of the current reconstruction

The event selection (i.e. field of view cuts and unsaturated signals requirement) greatly reduces the available number of events.

These selection cuts are even more severe for increasing energy[4], limiting the extension of mass composition studies.

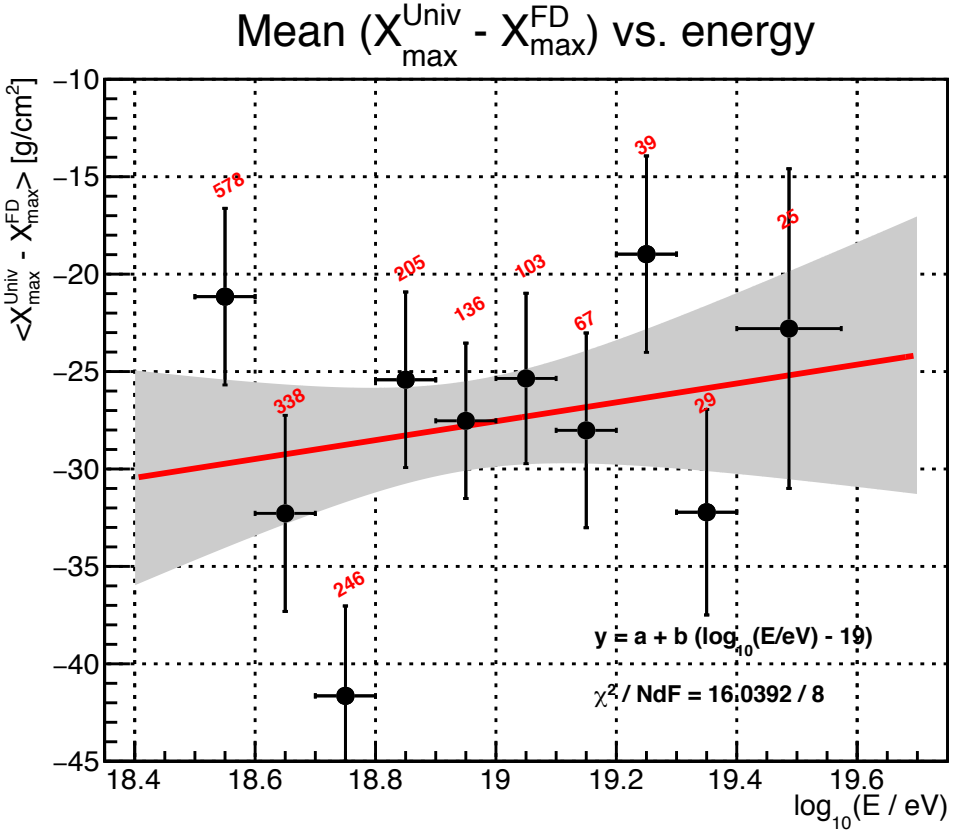
Changes in the fitting procedure cannot solve these issues, being instead required:

- a revision of the Universality parameterizations, to include showers with $X_{\max}$ below ground;
- a measurement of signals without saturation (hardware upgrade of the WCDs ?).

[4] For instance, around 30% of the events above 10^{19} eV present a saturated signal, growing to 50% above $10^{19.5}$ eV.

$X_{\mathrm{max}}^{\mathrm{Univ}}$ bias parameterization

The bias in X_{max}, observed in the reconstructed golden hybrids, must be subtracted from the $X_{\mathrm{max}}^{\mathrm{Univ}}$ of the SD-1500 events.



To estimate the bias as a function of the energy, the $(X_{\mathrm{max}}^{Univ} - X_{\mathrm{max}}^{FD})$ distributions are made separately in energy bins wide 0.1 in $\log_{10}(E/\mathrm{eV})$.

The parameterization is linear in the energy logarithm, even if with large uncertainties.

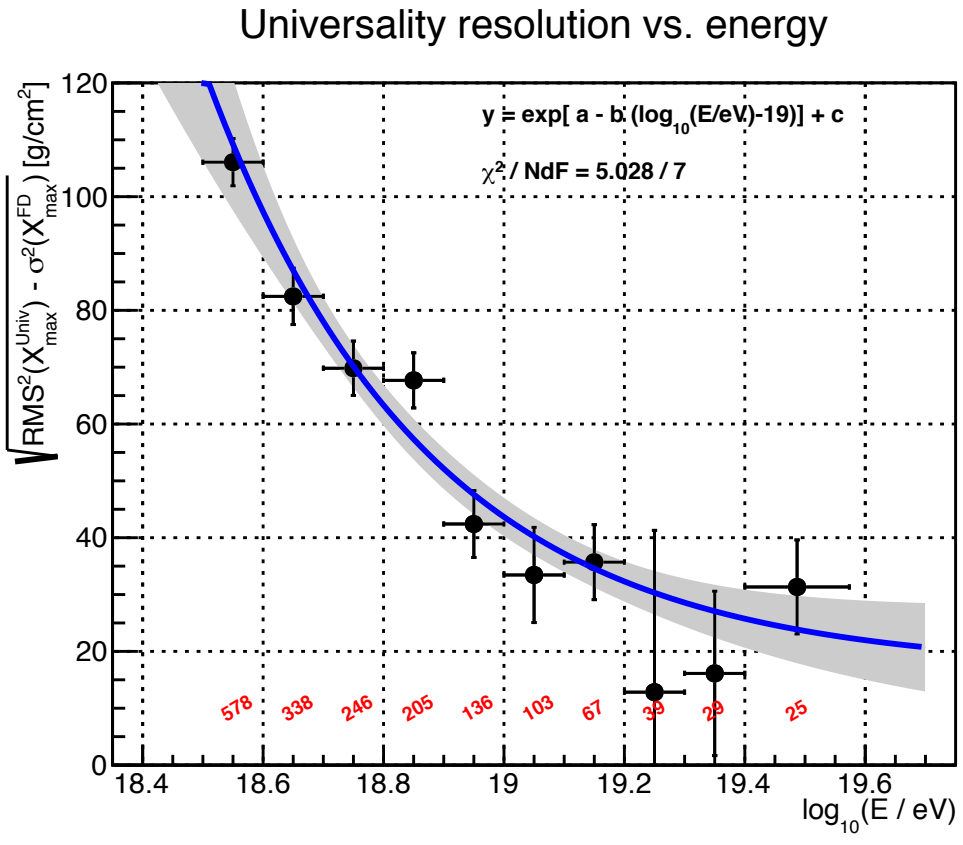
Resolution of the Universality fit

The resolution ($\sigma_{\rm Univ}$) quantifies the broadening of the "true" $X_{\rm max}$ distribution induced by the Universality procedure.

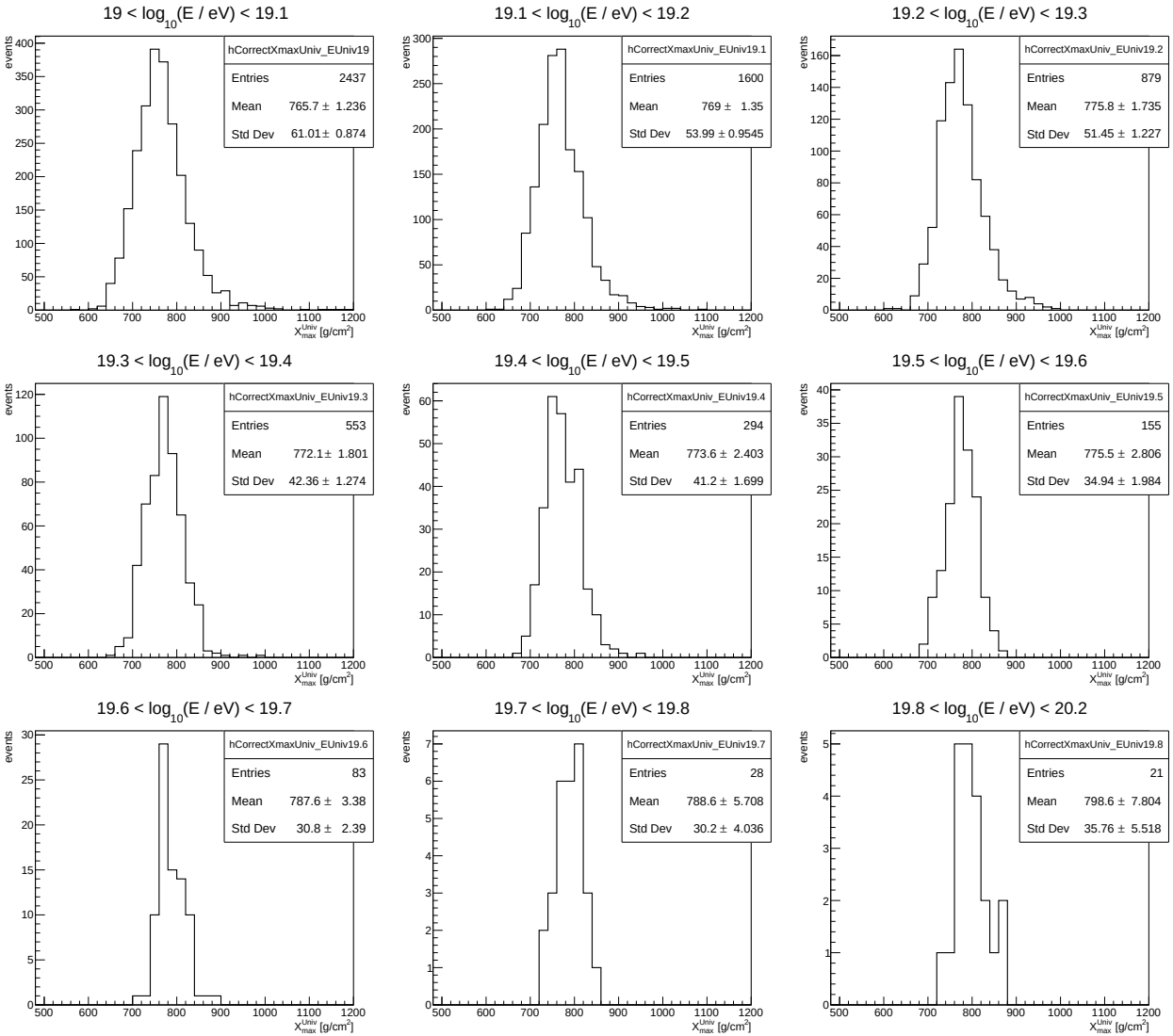
$$\sigma^2_{\rm Univ} = \\ = \sigma^2(X^{\rm Univ}_{\rm max}) - \sigma^2(X_{\rm max})$$

$\sigma(X^{\rm Univ}_{\rm max})$ = standard deviation of the $X^{\rm Univ}_{\rm max}$ distribution

$\sigma(X_{\rm max})$ = value obtained from the hybrid analysis (ICRC-2017) considered as the "true" $X_{\rm max}$ dispersion.

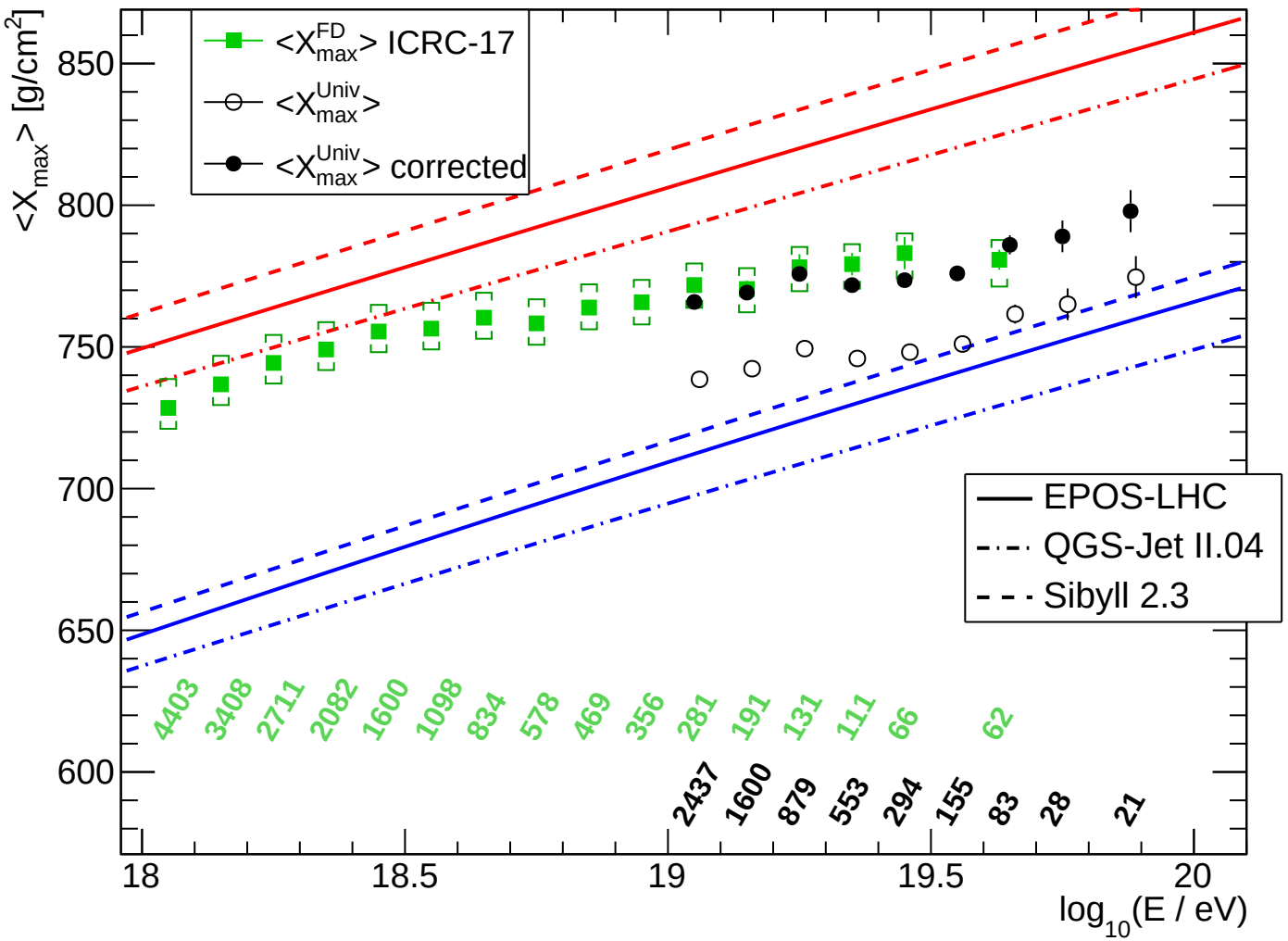


The $X_{\mathrm{max}}^{\mathrm{Univ}}$ distributions for the SD-1500 dataset

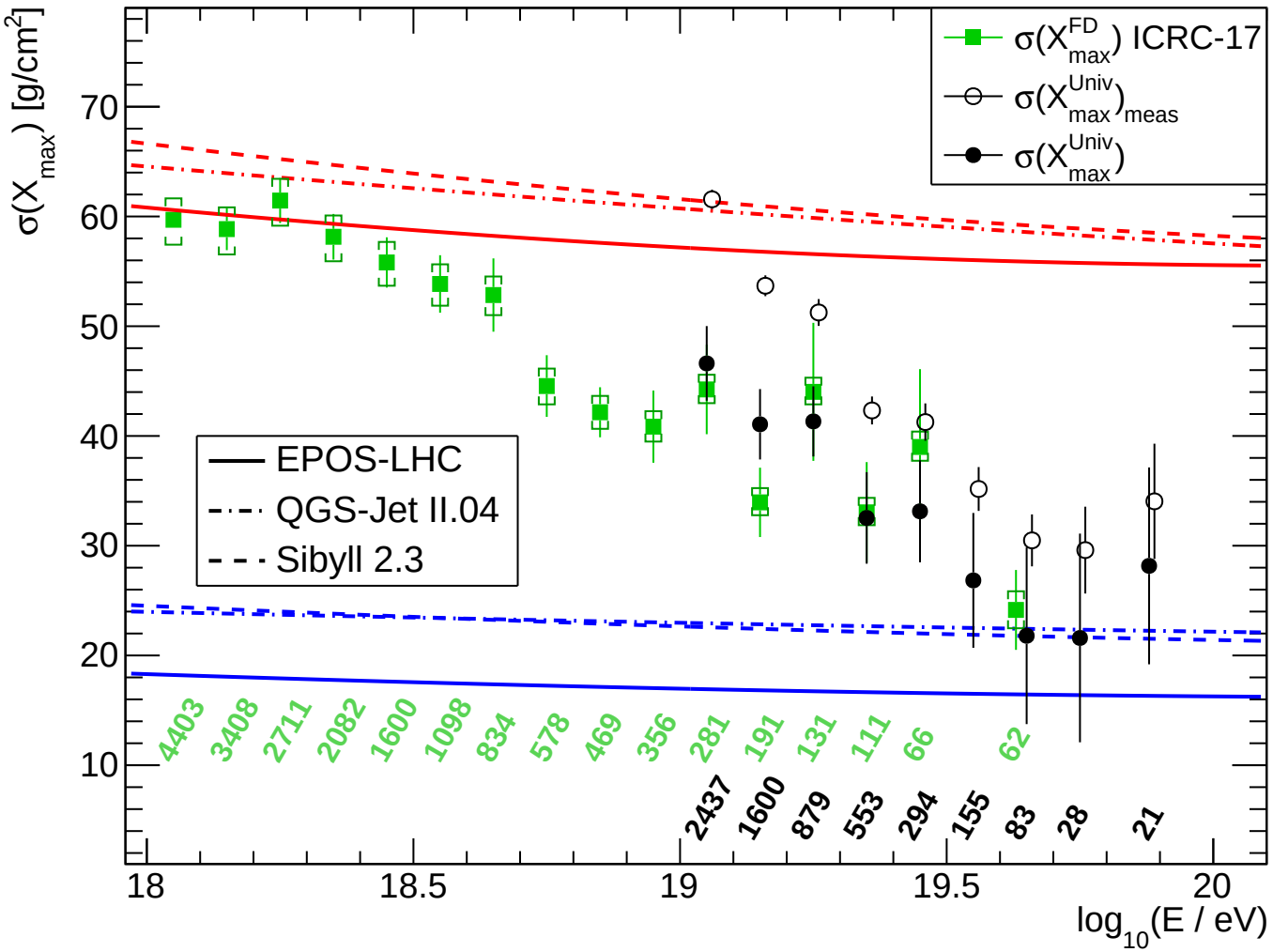


Part IV
Mass composition results

$\langle X_{\mathrm{max}}^{\mathrm{Univ}} \rangle$ for the SD-1500 dataset

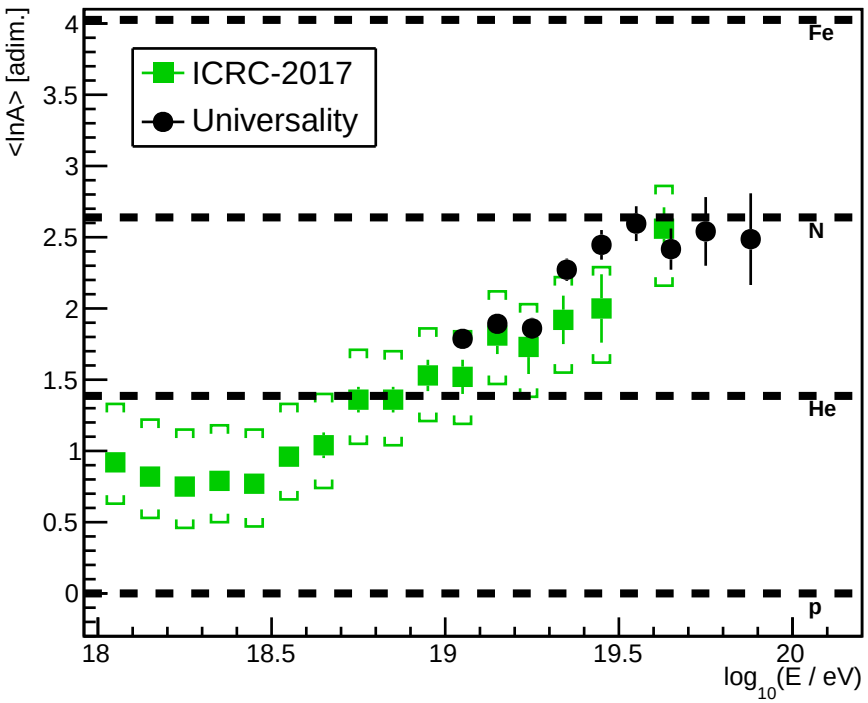


$\sigma(X_{\mathrm{max}}^{\mathrm{Univ}})$ for the SD-1500 dataset

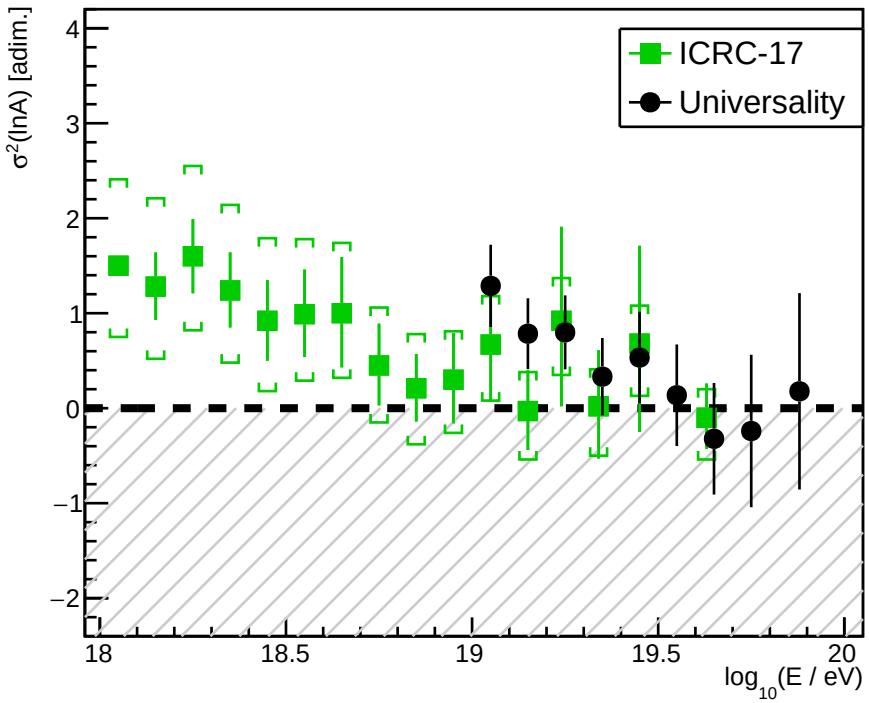


ln *A* moments for the SD-1500 dataset (EPOS-LHC)

Mean lnA vs. energy (EPOS-LHC)



lnA variance vs. energy (EPOS-LHC)



Conversion procedure in *JCAP 1302 (2013) 026*.
Updated values of the parameters given in *GAP 2018-021*.

Results of the moments analysis

- The agreement between the Universality and hybrid results is quite good in the common energy range.
- The value of $\langle \ln A \rangle$ from Universality remains approximately constant above $10^{19.5}$ eV (slow-down of the trend toward heavier composition).
- The last three bins of $\sigma^2(\ln A)$ from Universality do not shown an evident continuation of the previous decrease (slow-down of the trend toward purer composition).

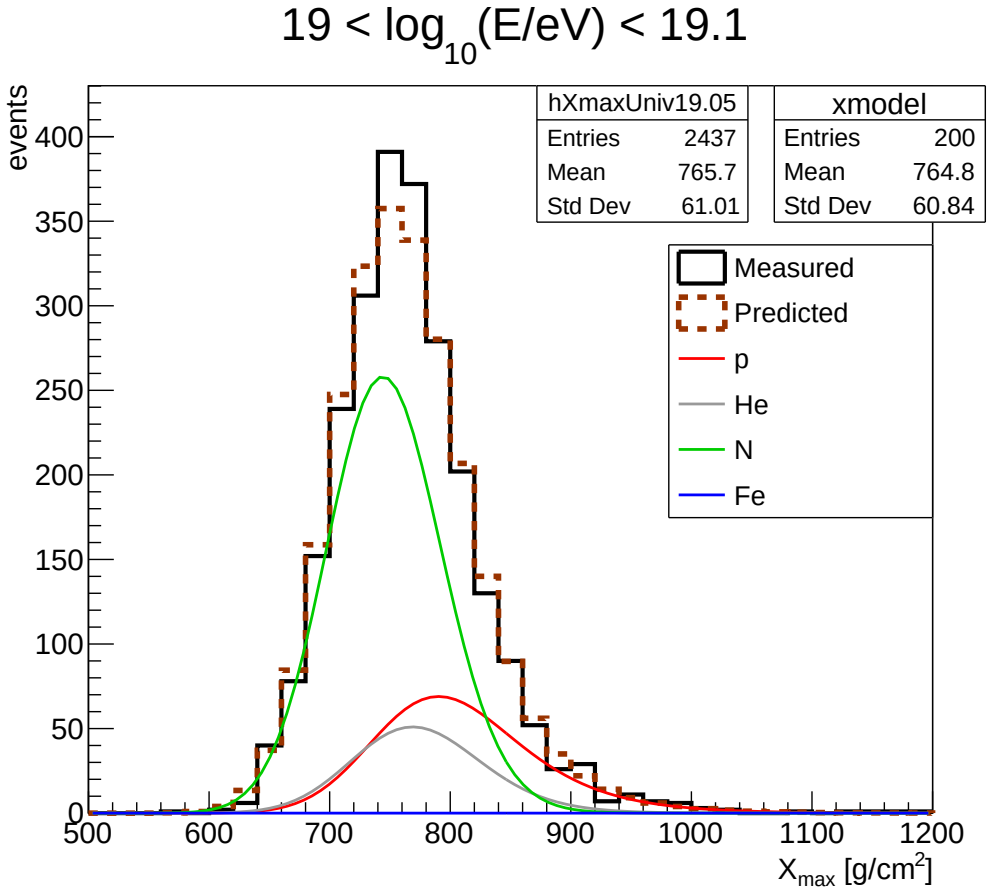
These results are valid also for the other hadronic models. However, the uncertainties are still too large to draw more precise conclusions (also because of the reduced statistics).

The fractions fit

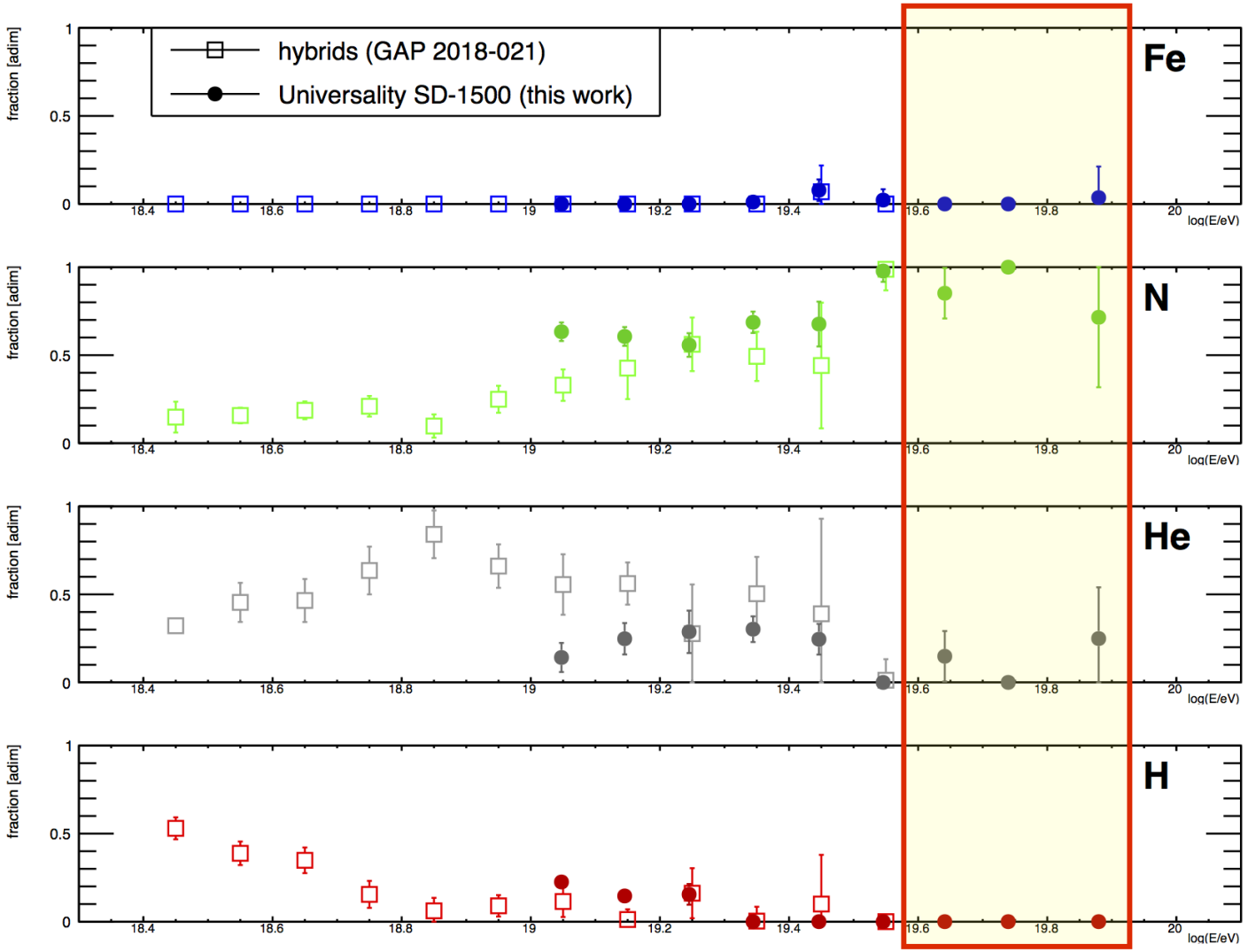
The fractions fit prescribes a comparison of the measured X_{max} distributions with the expectations for a chosen set of primary masses at the Earth.

Details of the procedure in *Phys. Rev. D90 (2014) 122006*.

Also the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distributions can be directly used to infer the fractions of different primary species.



Fractions of representative primaries (EPOS-LHC)

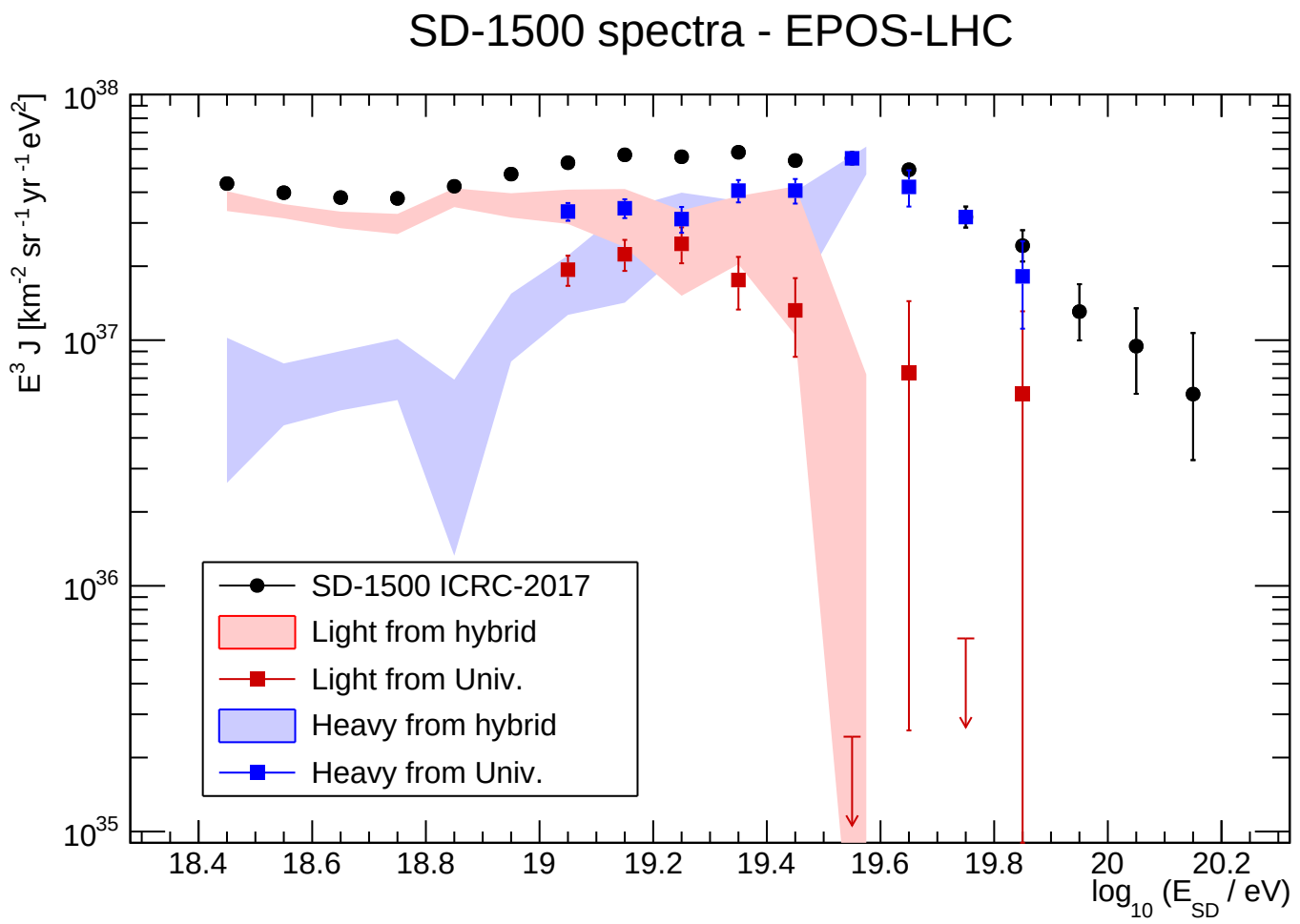


Results of the fractions fit (EPOS-LHC)

- Both in hybrid and Universality results, the fraction of heavy elements (represented by iron) is negligible and a prevalence of nitrogen is observed as the energy increases.
- The helium fractions are smaller in the Universality result than in hybrids (in particular near 10^{19} eV) with correspondingly larger fractions of nitrogen, showing the importance of **correlations** among the 4 fractions.

The uncertainties in the previous plots are calculated simply from the propagation of the errors over the fitted fractions, while the systematics have not been evaluated (more details in the back-up).

Light/Heavy spectra from fractions fit (EPOS-LHC)



Light component = sum of proton and helium components
Heavy component = sum of nitrogen and iron components

Results light/heavy spectra (EPOS-LHC)

The two spectra are obtained multiplying the flux in each bin for the light/heavy fractions (sum of the $H + He$ or $N + Fe$ fractions previously shown).

- The transition from a light-dominated toward and intermediate/heavy-dominated scenario is evident.
- The information added through the fit of the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distributions seems to disfavour a complete disappearance of the light component above $10^{19.5}$ eV.

The upgrade of the Pierre Auger Observatory: *AugerPrime*

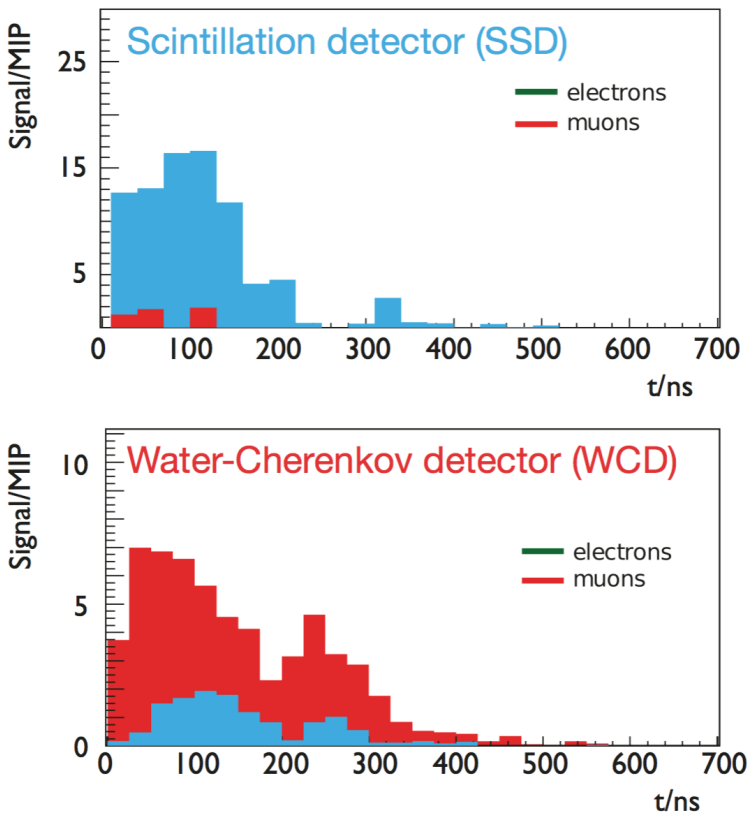
The Observatory has started a major upgrade, called *AugerPrime* (*Preliminary Design Report (2016) - arXiv:1604.03637*).

The key element of the upgrade is a new plastic scintillator, to be installed on top of each WCD, called **Surface Scintillator Detector** (SSD).

At the same time, an additional **small photomultiplier** (SPMT), designed to extend the dynamic range from about 600 VEM to more than 30,000 VEM, will be installed in each WCD.

AugerPrime & Universality reconstruction

The **small-PMTs** will allow to measure signals without saturation in almost all events.



The data from **SSDs** will offer a new measurement of the shower, complementary to the WCDs one, possibly leading to the development of an enhanced fitting procedure.

The **comparison between WCD and SSD signals** could be used to disentangle the electromagnetic and muonic contributions, allowing a constraint of N_μ in the fit.

Conclusions

The feasibility of a mass composition analysis of UHECRs based on the data collected only by the surface detector is proven.

An optimized global Universality reconstruction has been established and then applied to the unsaturated SD-1500 events collected between Jan. 2004 and Dec. 2016.

The $X_{\mathrm{max}}^{\mathrm{Univ}}$ and $\ln A$ moments show a good agreement with the hybrid results and hint a slow-down of the trend toward a heavier and purer composition in the newly added higher energy bins.

Moreover, the fit of four primary (H, He, N, Fe) fractions suggests the presence of a light contribution also at the highest energies.

Backup

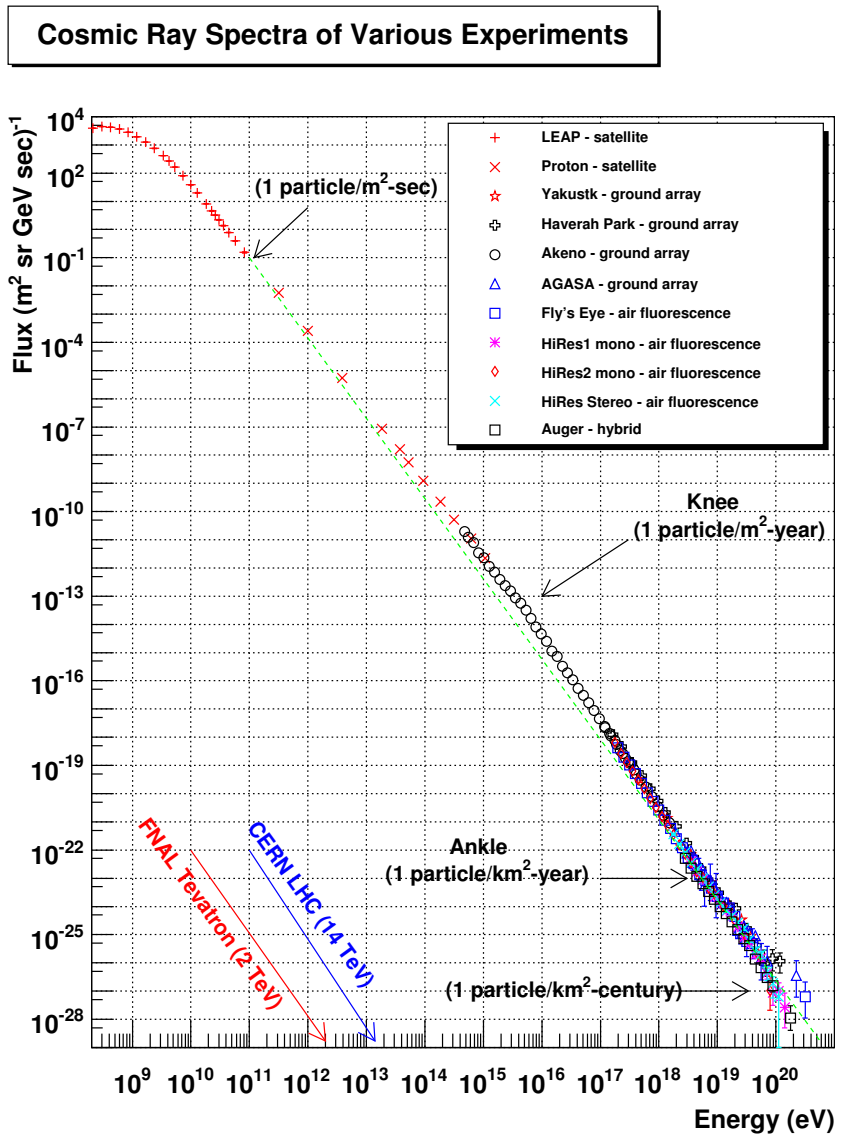
The cosmic ray energy spectrum

The (all-particle) energy spectrum of cosmic rays spans over 12 orders of magnitude in energy and above 30 orders of magnitude in flux.

The differential flux as function of the energy E is a power law:

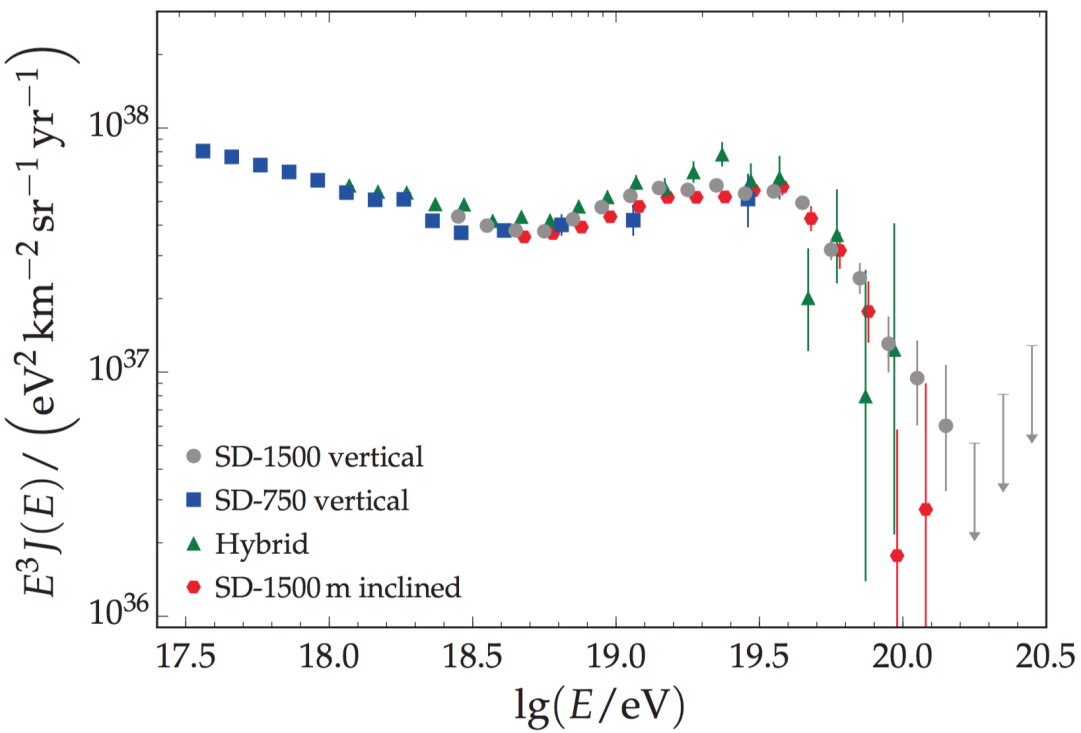
$$\frac{\mathrm{d}N}{\mathrm{d}E} \propto E^{-\gamma} \quad [\mathrm{m}^{-2}\mathrm{s}^{-1}\mathrm{sr}^{-1}\mathrm{eV}^{-1}]$$

where γ is the so-called spectral index.



The ultra-high energy cosmic rays

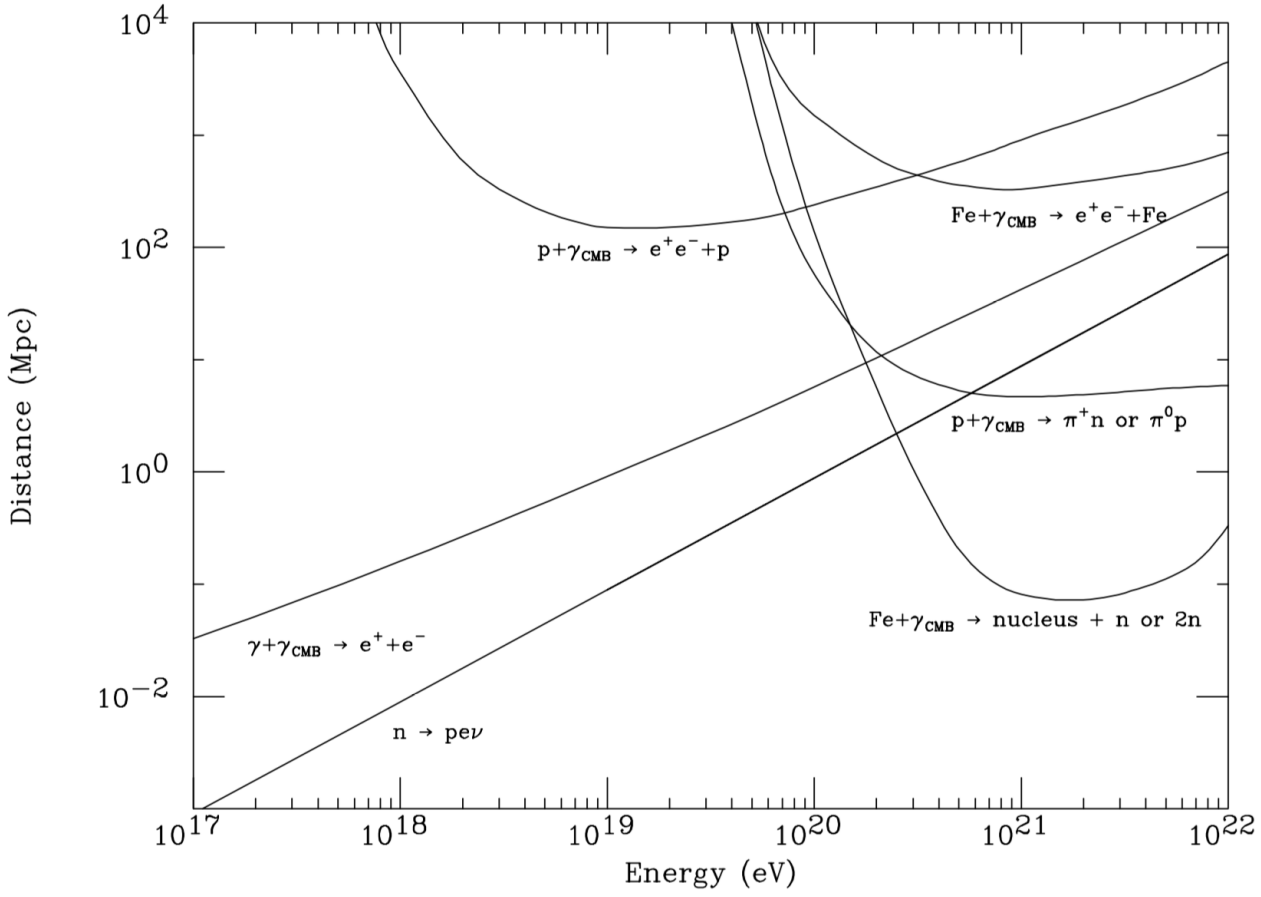
Primary particles with energies above the *ankle* ($E \approx 5 \times 10^{18}$ eV) are usually referred to as ultra-high energy cosmic rays.
A suppression of the flux is anyway observed above $10^{19.5}$ eV.



Energy spectra obtained by Pierre Auger Observatory with different detection techniques.

The differential flux is multiplied by E^3 to emphasise the spectral features.

Energy losses during CR propagation



Summary of interactions lengths for the main reactions of cosmic primaries with the CMB photons.

From $X_{\max}$ to $\ln A$ moments

$$\langle X_{\max} \rangle = \langle X_{\max} \rangle_p + f_E \langle \ln A \rangle \qquad \sigma^2(X_{\max}) = \langle \sigma^2_{sh} \rangle + f_E^2 \, \sigma^2(\ln A)$$

Functional forms of the parameterizations for $\langle X_{\max} \rangle_p$, f_E, σ^2_{sh} are given in *JCAP 1302 (2013) 026*. In particular:

$$\langle \sigma^2_{sh} \rangle = \sigma^2_p [1 + a \langle \ln A \rangle + b \langle \ln A \rangle^2 + b \, \sigma^2(\ln A)]$$

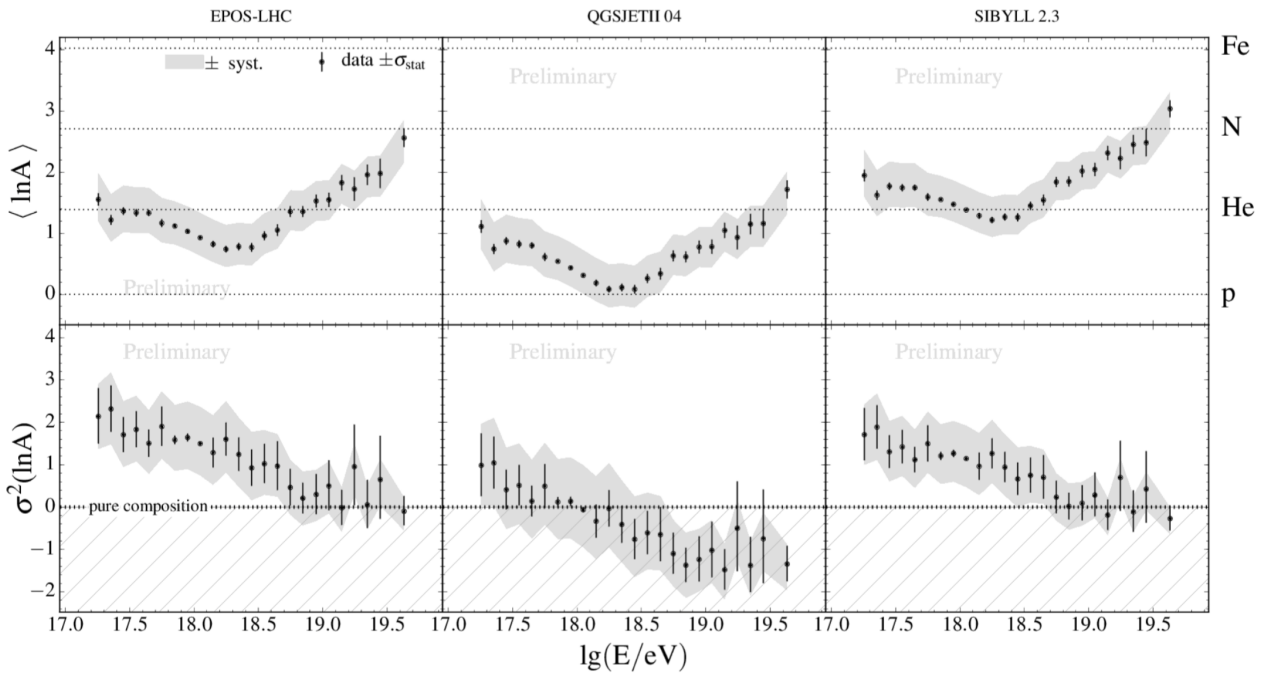
Then:

$$\langle \ln A \rangle = \frac{\langle X_{\max} \rangle - \langle X_{\max} \rangle_p}{f_E} \qquad \sigma^2(\ln A) = \frac{\sigma^2(X_{\max}) - \sigma^2_{sh}(\langle \ln A \rangle)}{b \, \sigma^2_p + f_E^2}$$

Updated values of the parameters given in *GAP 2018-021*.

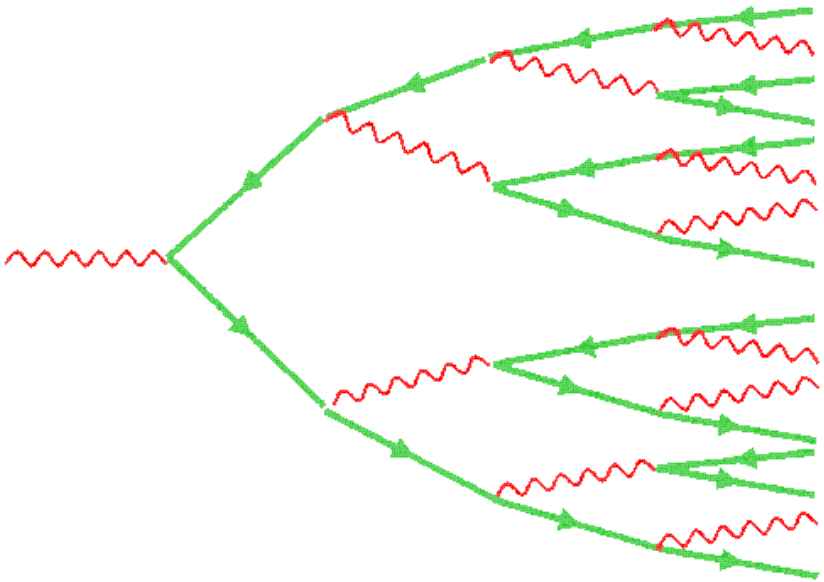
Mass composition

Using the procedure described in *JCAP 1302 (2013) 026*, the first two moments of the distribution of (the logarithm of) the primaries mass A are directly calculated from $\langle X_{\max} \rangle$ and $\sigma(X_{\max})$.



Presented at the 35th ICRC (2017) by J. Bellido.

Heitler model for electromagnetic showers



Sketch of the shower development according to the Heitler model. Each step corresponds to a radiation length $\lambda_r \times \ln 2$; the wave lines are γs, the green lines oriented to the right are electrons, if to the left positrons.

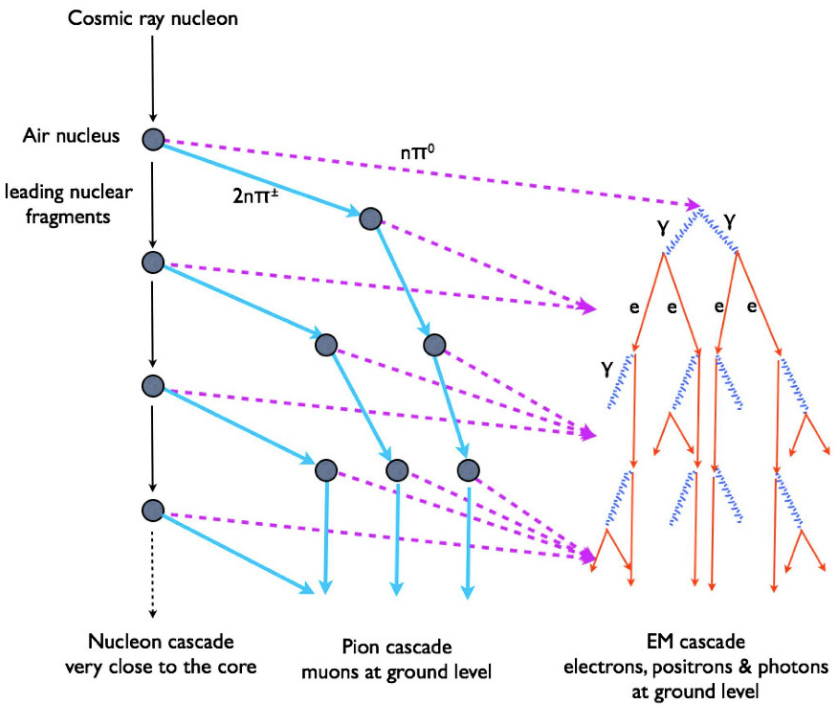
Mean energy at a step n given the initial energy E_0

$$\langle E \rangle = \frac{E_0}{2^n}$$

The shower reaches its maximum when the average energy per particles is not enough to generate new secondaries, that is when $\langle E \rangle$ approaches the critical energy ε_c. Then:

$$X_{\max} = X_0 + \lambda_r \ \ln(E_0/\varepsilon_c)$$

Generalized Heitler model for hadronic showers



Sketch of the model for the development of a hadronic cascade. At each step about 1/3 of the energy is transferred to the e.m. component.

After n interaction lengths, the energy is divided between the hadronic and e.m. parts:

$$E_{had} = (2/3)^n \; E_0$$

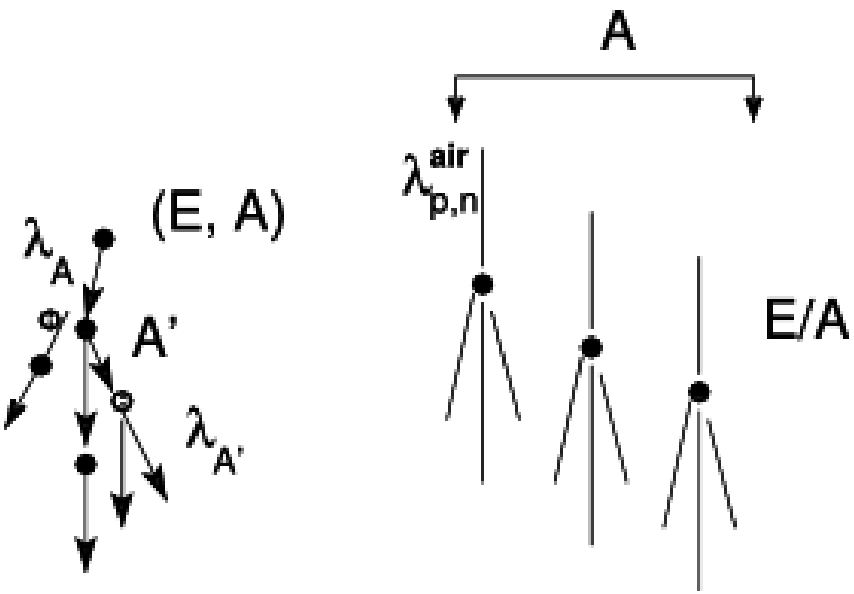
$$E_{e.m.} = [1 - (2/3)^n] \; E_0$$

The cascade is numerically dominated by e.m. particles, thus $X_{\max}$ is defined as the maximum of the electromagnetic part.

$$X^{p}_{\max}(E_0) \approx X_0 + \lambda_I \ln 2 + X^{e.m.}_{\max}\left(\frac{(1/3)E_0}{2N_\pi}\right)$$

Generalized Heitler model for hadronic showers

Basic assumption of the superposition model: a primary nucleus with mass number A and energy E behaves as A independent primary nucleons with energy E/A.



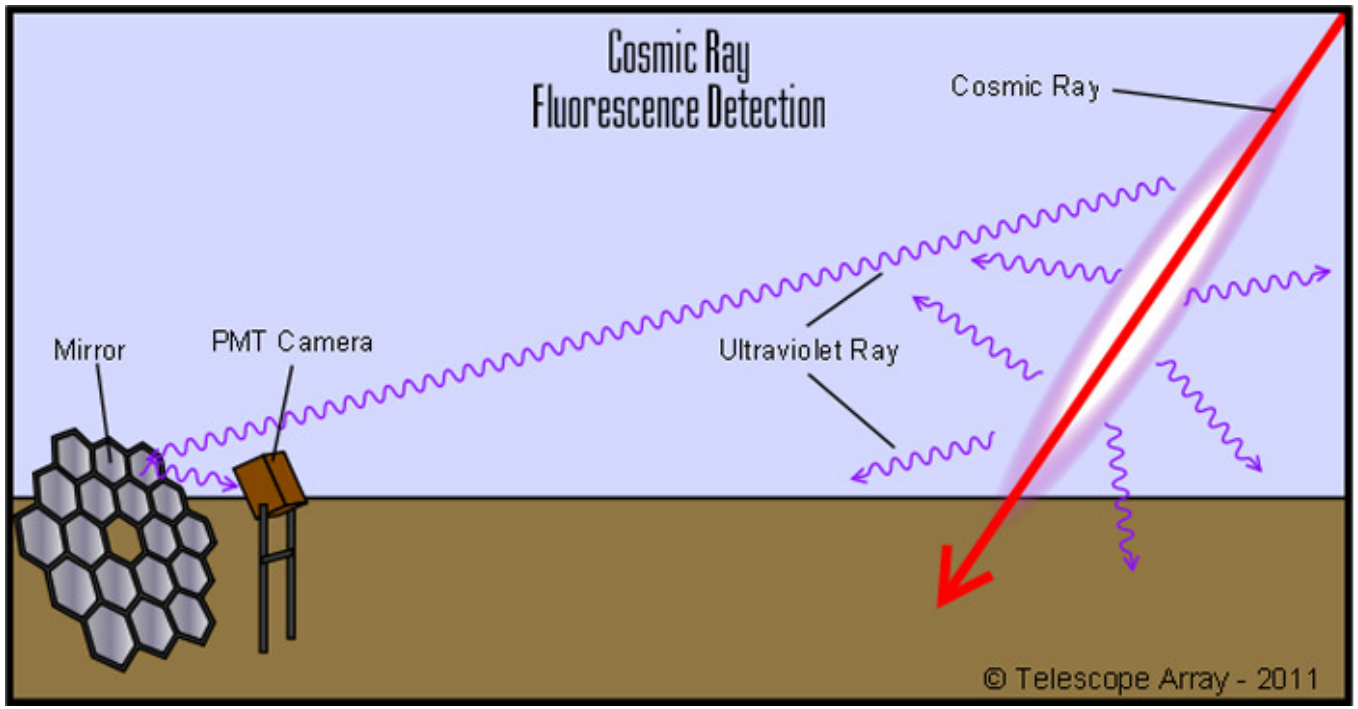
$$X^A_{\max}(E_0) = X^p_{\max}(E_0/A) =$$
$$= X^p_{\max}(E_0) - \lambda_r \ln A$$

The superposition model also predicts a faster development of the shower, meaning that the pions reach their critical energy sooner. Therefore:

$$N^A_\mu(E_0) = \left(\frac{E_0}{AE^\pi_c}\right)^\alpha = A^{1-\alpha}\, N^p_\mu(E_0)$$

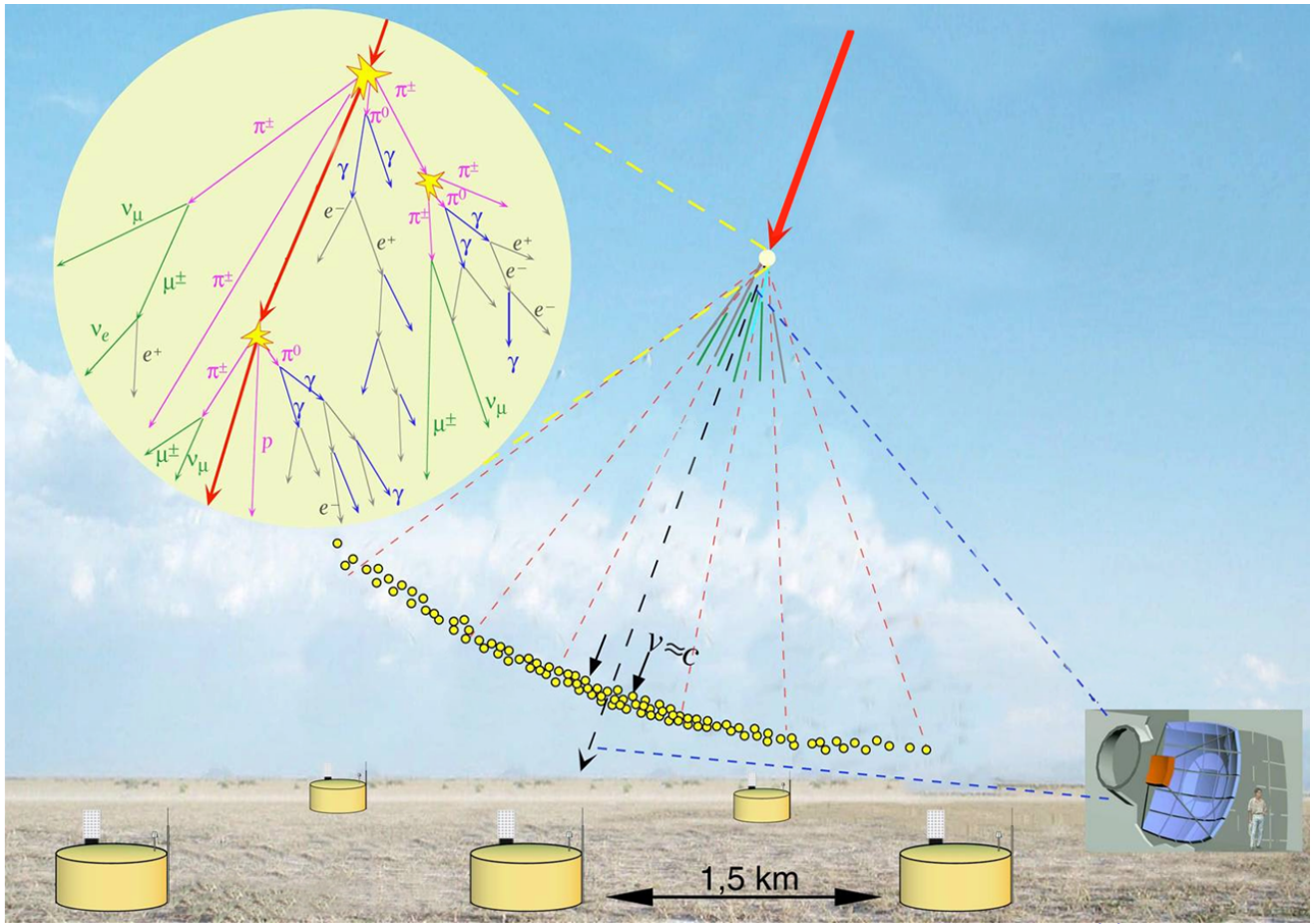
The fluorescence technique

This experimental technique takes advantage of the radiation, called **fluorescence light**, emitted by the air nitrogen molecules excited by the passage of the (ionizing component of the) cascade.

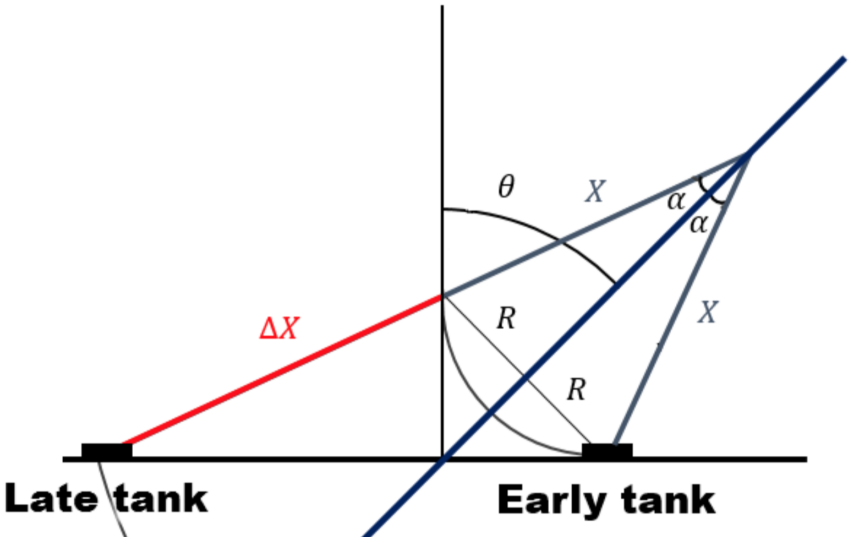


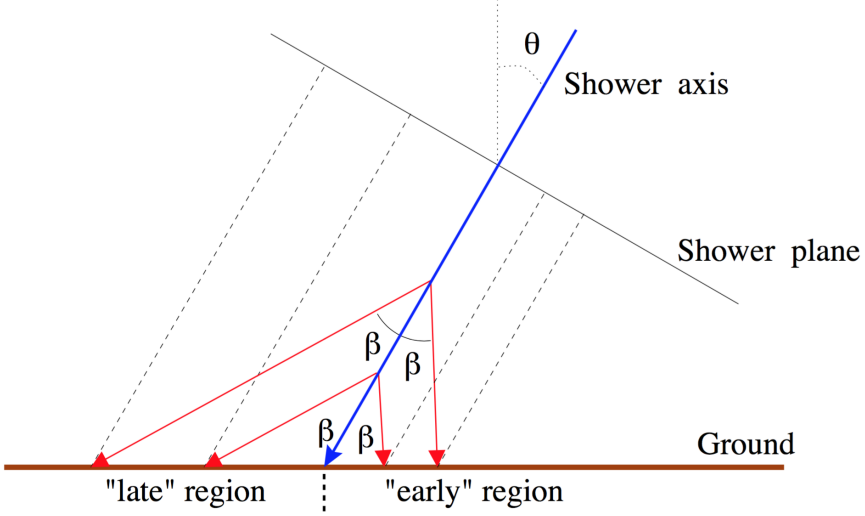
The surface array detection

Another experimental solution is a network of ground-based particle detectors which **sample** the shower particles at ground level.



Azimuthal asymmetry of signals



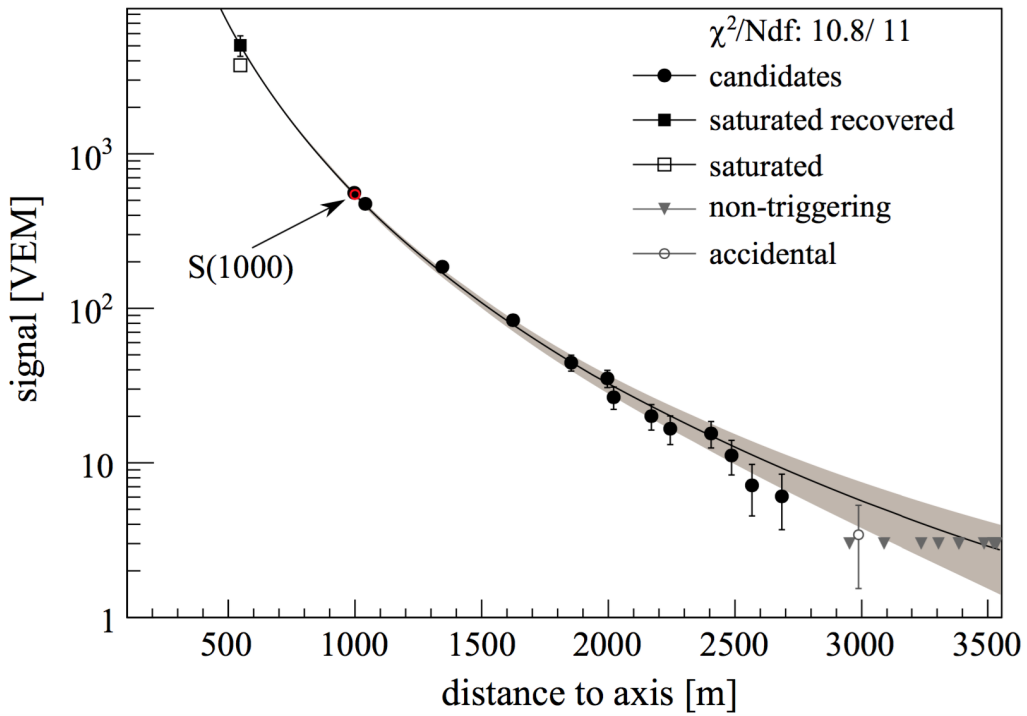


An asymmetry affects the signals collected by the SD and it is due to the combination of the longitudinal evolution of the shower (left) and of geometrical effects related to the incidence angles of the particles on the WCDs (right).

SD standard reconstruction - LDF

Fit of the lateral distribution function (LDF) of the WCD signals with a modified Nishimura-Kamata-Greisen function:

$$S(r) = S(r_{\mathrm{opt}}) \left(\frac{r}{r_{\mathrm{opt}}} \right)^{\beta} \left(\frac{r + r_1}{r_{\mathrm{opt}} + r_1} \right)^{\beta+\gamma}$$

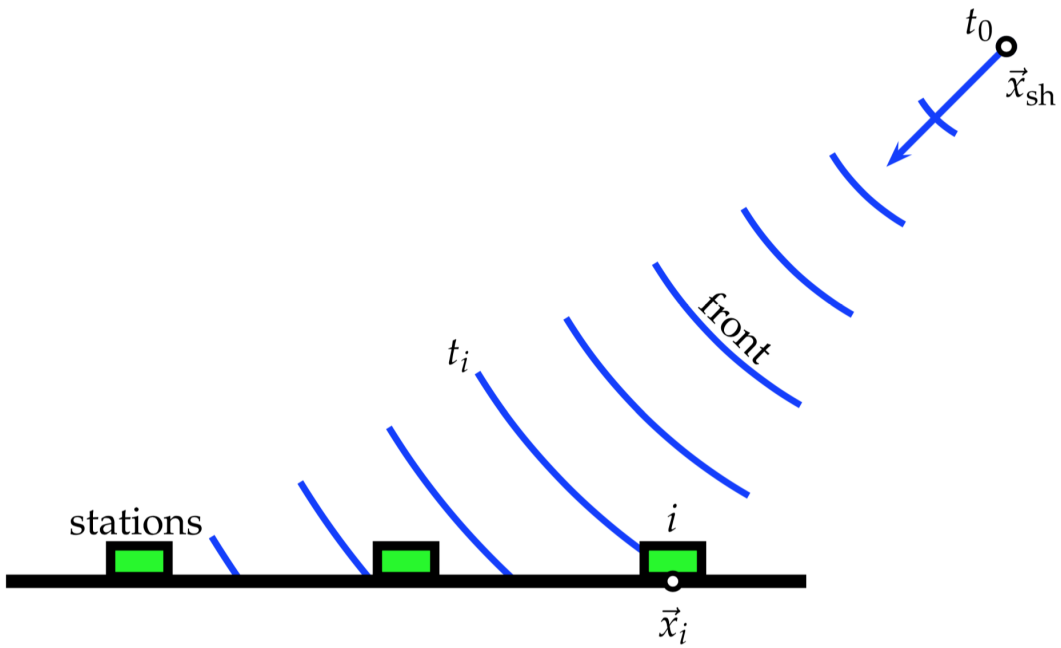


where $r_1 = 700$ m, β and γ are slope parameters functions of the zenith angle and $r_{\mathrm{opt}} = 1000$ m is the optimal distance (*Astropart. Phys. 26 (2007), p. 414*).

SD standard reconstruction - geometry

The shower geometry is reconstructed by fitting the start times of the signals (t_i) of each WCD with a model describing the evolution of the shower front as a sphere inflating at the speed of light:

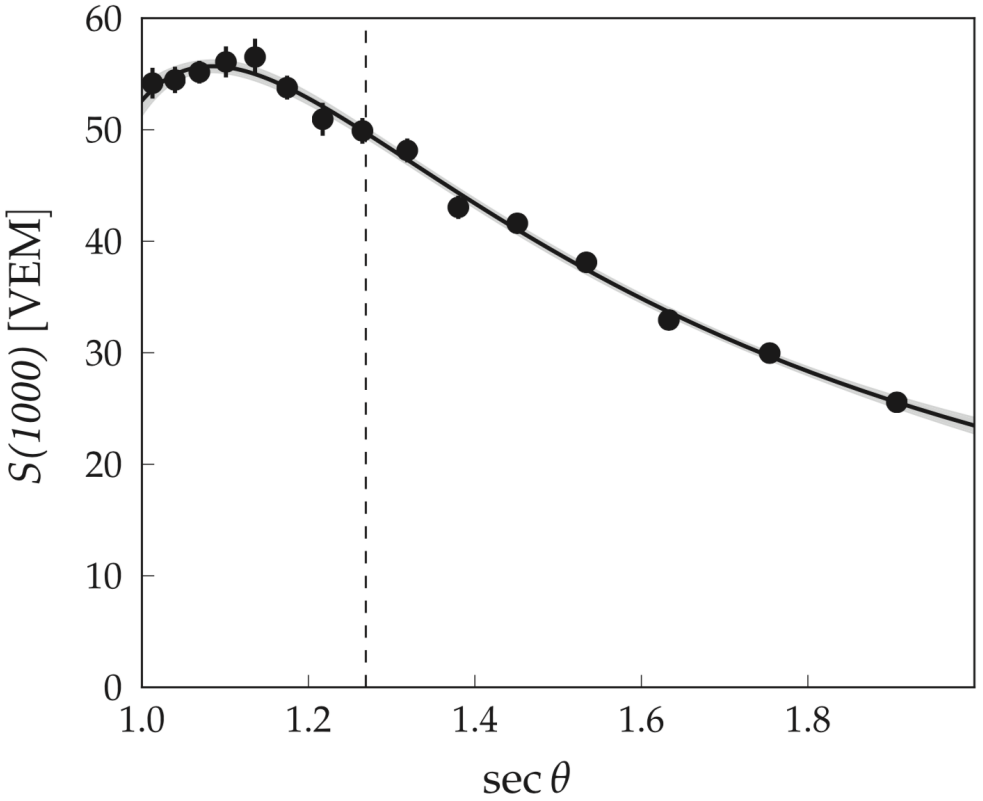
$$c(t_i - t_0) = |\vec{x}_{sh} - \vec{x}_i|$$



where t_0 and $\vec{x}_{sh}$ identify a virtual origin (in space and time) of the shower.

SD standard reconstruction - energy

The value of $S(1000)$ is converted in an energy estimator using the Constant Intensity Cut method, that is taking into account the shower attenuation with the zenith angle and other geometrical effects.

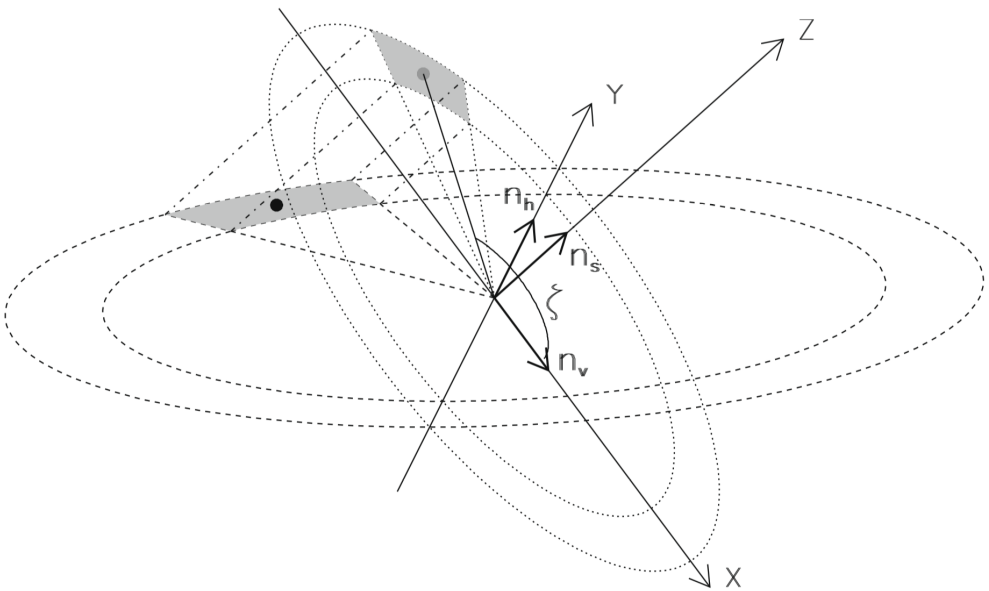


$$S_{38} \equiv \frac{S(1000)}{f_{\mathrm{CIC}}(\theta)}$$

which can be considered as the signal that a shower with a certain size $S(1000)$ would have produced arriving with an inclination of 38°.

Parameterization of the integrated signal

The 4-components description allowed a parameterization of integrated signals measured by the WCDs of the Pierre Auger Observatory as a function of the shower macroscopic characteristic and of the detector position (r, ϕ).

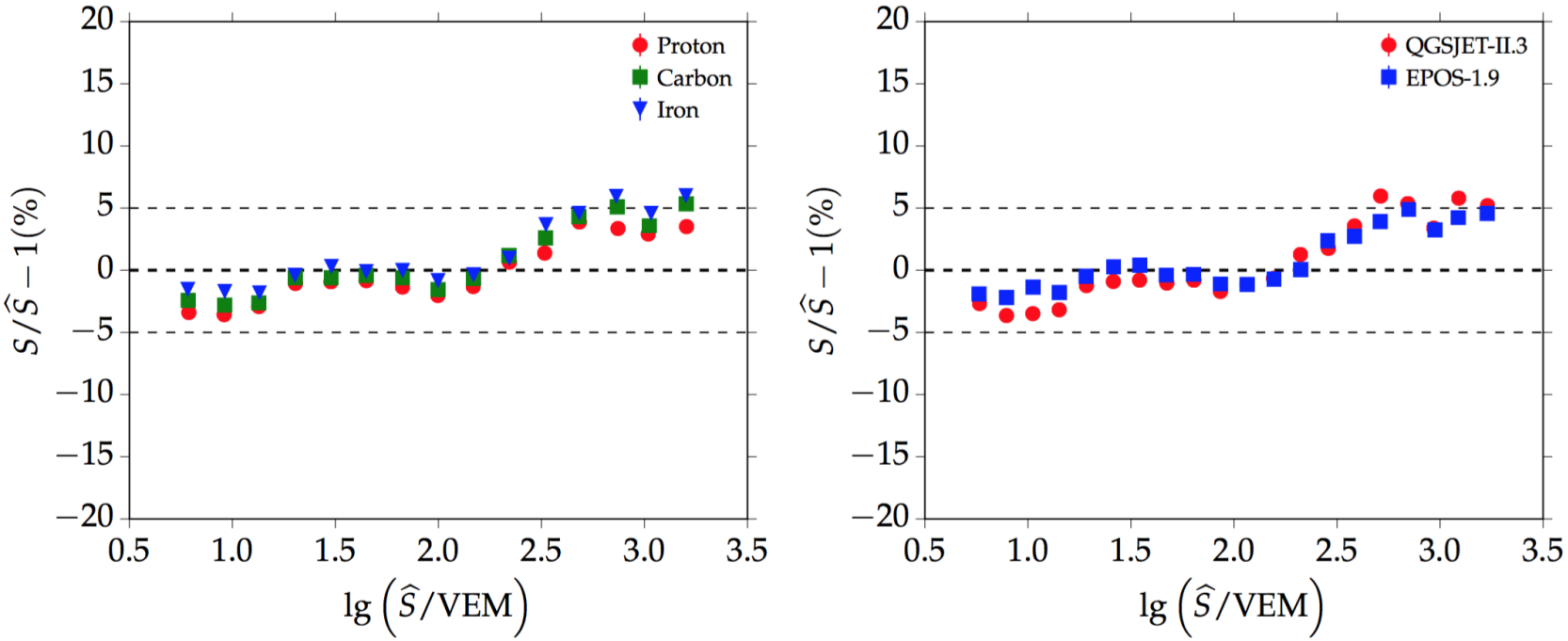


P. Billoir, Astropart. Phys. 30 (2008)

The (simulated!) signals employed to build the parameterizations are obtained **averaging over a large area** around each detector.

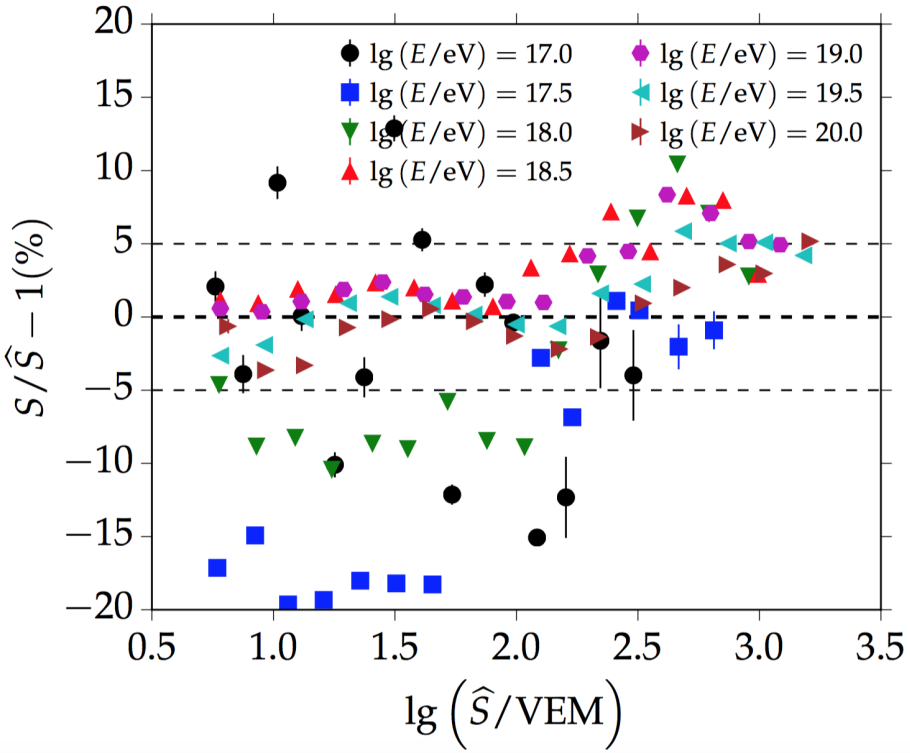
M. Ave et al., GAP 2011-087

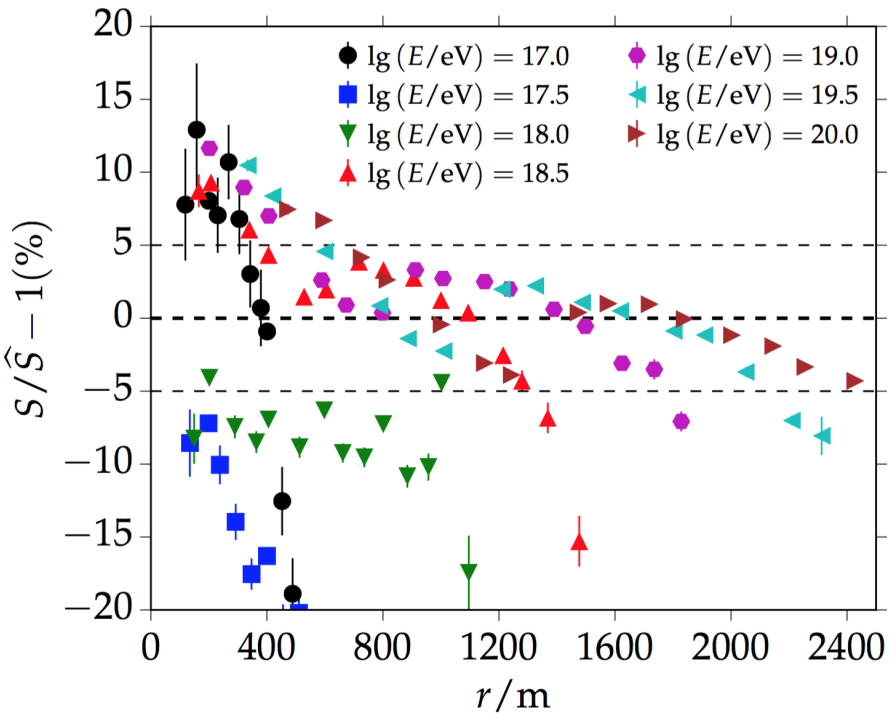
Parameterization of the integrated signal - residuals



Relative signal residuals as a function of the predicted signal $\hat{S}$ for different shower primaries and hadronic interaction models.

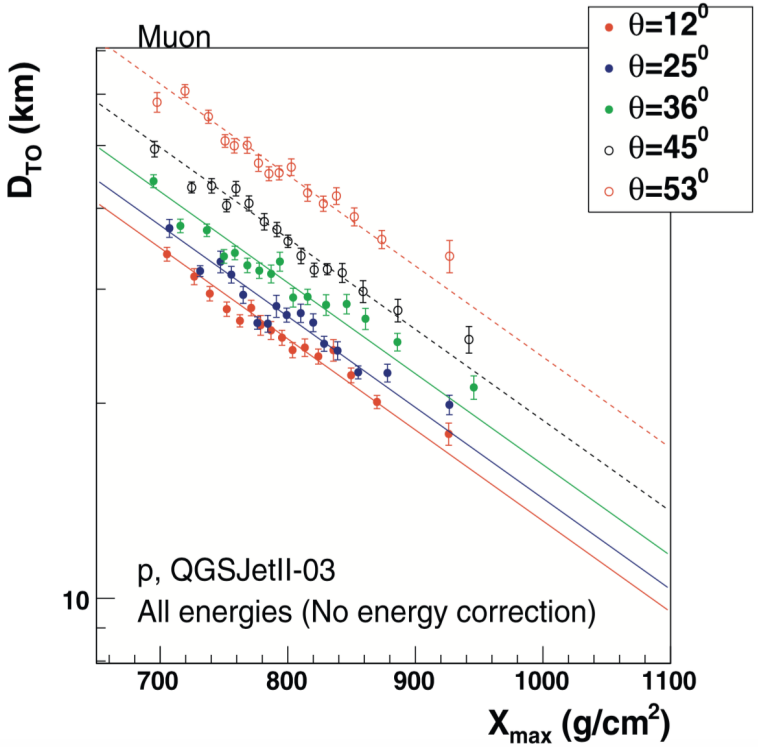
Parameterization of the integrated signal - residuals

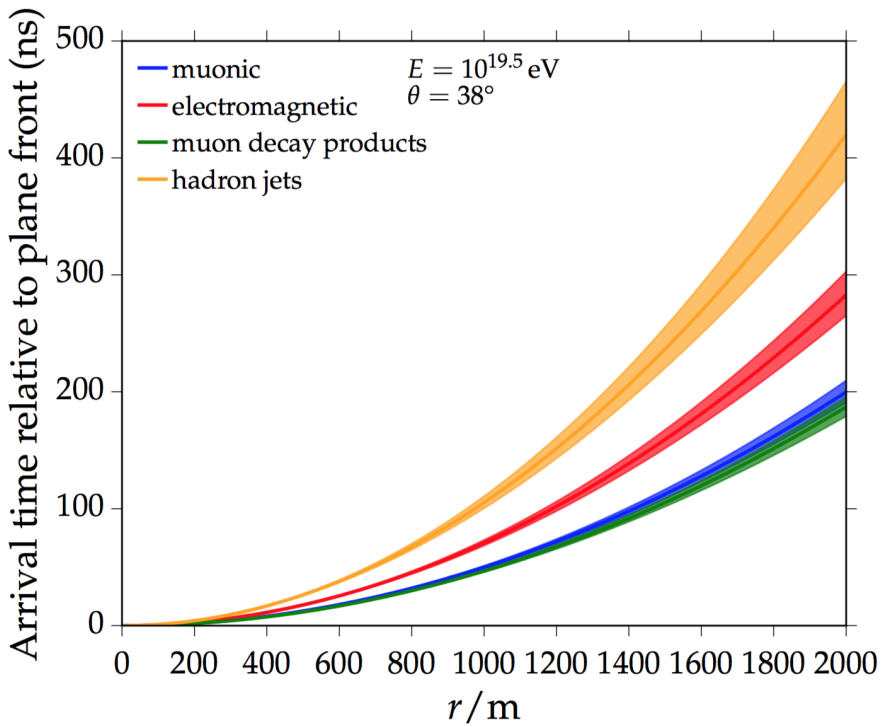




Relative signal residuals as a function of the predicted signal $\hat{S}$ and of the core distance r for different primary energies.

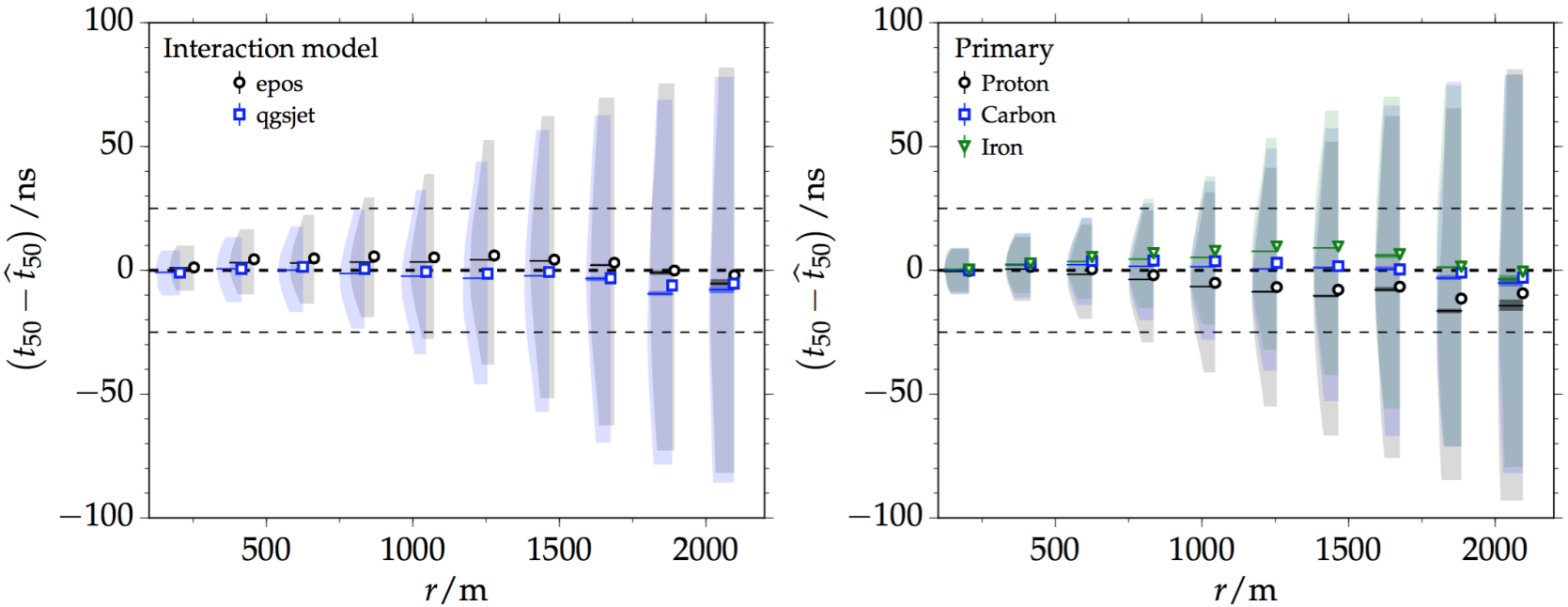
Parameterization of the origin of times





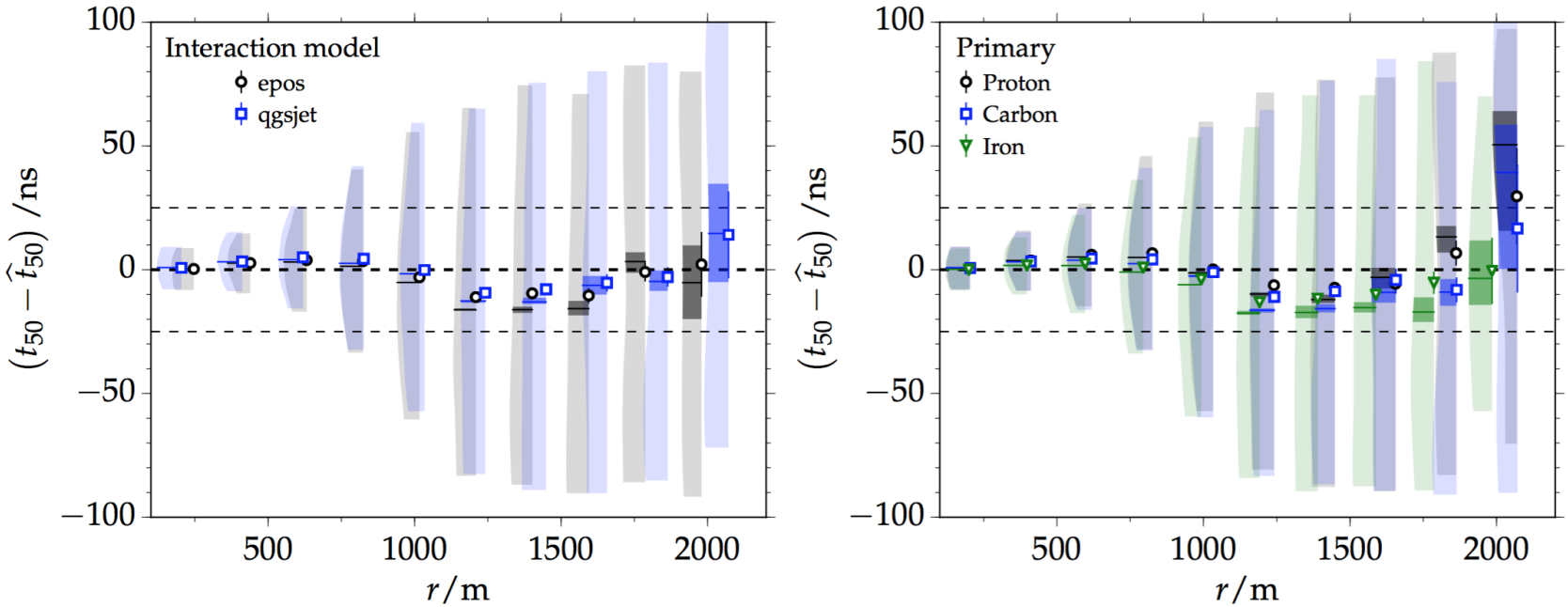
(Left) Mean D_{TO} for the muonic component as a function of $X_{\max}$ for different zenith angles. (Right) Start time t_0 as a function of the core distance, for a fixed energy and zenith angle; the 4 shower components are shown with different colours.

Parameterization of traces - residuals muonic



Residuals for the 50% time quantile of the muonic component, as functions of the core distance r, for different hadronic interaction models and primary species.

Parameterization of traces - residuals e.m. pure



Residuals for the 50% time quantile of the pure e.m. component, as functions of the core distance r, for different hadronic interaction models and primary species.

The minimization procedure

The original fitting procedure is a negative log-likelihood minimisation. The function to be minimized is given by:

$$-\ell = -\sum_{k}^{N_{\text{stations}}} \left(\ell^{k}_{\text{LDF}} + \sum_{i}^{n^{k}_{\text{bins}}} \ell^{k,i}_{\text{shape}} \right)$$

Symmetrically, the total deviance to be minimized is:

$$D = \sum_{k}^{N_{\text{stations}}} \left(D^{k}_{\text{LDF}} + \sum_{i}^{n^{k}_{bins}} D^{k,i}_{\text{shape}} \right)$$

It must be stressed that the traces of each tank are not independently fitted. This means that the minimum does not give the best fit function (sum of 4 parameterized log-normal distributions) for each single trace.

The shape likelihood

A trace is used for the **shape fit** if at least 5 bins have a signal above 0.7 VEM and the station it belongs to has a core distance $r \in [100, 2200]$ m. Then for each bin **above 0.7 VEM** it is calculated a **gaussian** log-likelihood:

$$\ell^i_{shape} = \ln \mathcal{L}^i_{shape} = \ln[\mathcal{N}(s_i \mid \hat{s}_i, \epsilon(s_i))]$$

where $\mathcal{N}$ indicates the normal distribution of the variable s_i (the measured signal in the bin i) with mean $\hat{s}_i$ (the signal predicted by the sum of the 4 log-normal parameterisations) and dispersion equal to $\epsilon(s_i)$ (the assigned bin content uncertainty).

Since the traces are described as gaussian distributed binned data, the corresponding deviance is equal to:

$$D^k_{\text{shape}} = \sum_{i=1}^{N^k_{\text{bins}}} \left(\frac{s_i - \hat{s}_i}{\epsilon(s_i)} \right)^2$$

The LDF likelihood

A station is used for the **LDF fit** if $r \in [100, 2500]$ m and if the associated trace is not used for the shape fit. For each selected station a **poissonian** log-likelihood is calculated:

$$\ell^k_{LDF} = \ln \mathcal{L}^k_{LDF} = \ln \left[\mathcal{P} \left(\frac{S^k_{\mathrm{tot}}}{f_S^2}, \frac{\hat{S}^k_{tot}}{f_S^2} \right) \right]$$

where $\mathcal{P}$ indicates the Poisson distribution of the variable S^k_{tot} (the integrated signal measured in the station k) with mean equal to $\hat{S}^k_{tot}$ (the expected signal from the parameterisation), both divided by the factor $1/f_S^2$, which transforms the signal into a poissonian variable (representing an *effective number of particles*).

In the case of poisson distributed binned data the deviance is equal to:

$$D^k_{\mathrm{LDF}} = 2 \left(\frac{1}{f_S^2} \cdot S^k_{\mathrm{tot}} \ln \frac{S^k_{\mathrm{tot}}}{\hat{S}^k_{\mathrm{tot}}} + \frac{1}{f_S^2} \cdot \hat{S}^k_{\mathrm{tot}} - \frac{1}{f_S^2} \cdot S^k_{\mathrm{tot}} \right)$$

Uncertainty over the bin signals

The bin signals s_i (measured in intervals of 25 ns) were initially assumed to have the same fluctuations as for the total signal (see for example GAP-2014-035), which means:

$$\epsilon(s_i) = f_S \times \sqrt{s_i}$$

where the uncertainty factor f_S is parameterized as a function of the shower zenith angle θ:

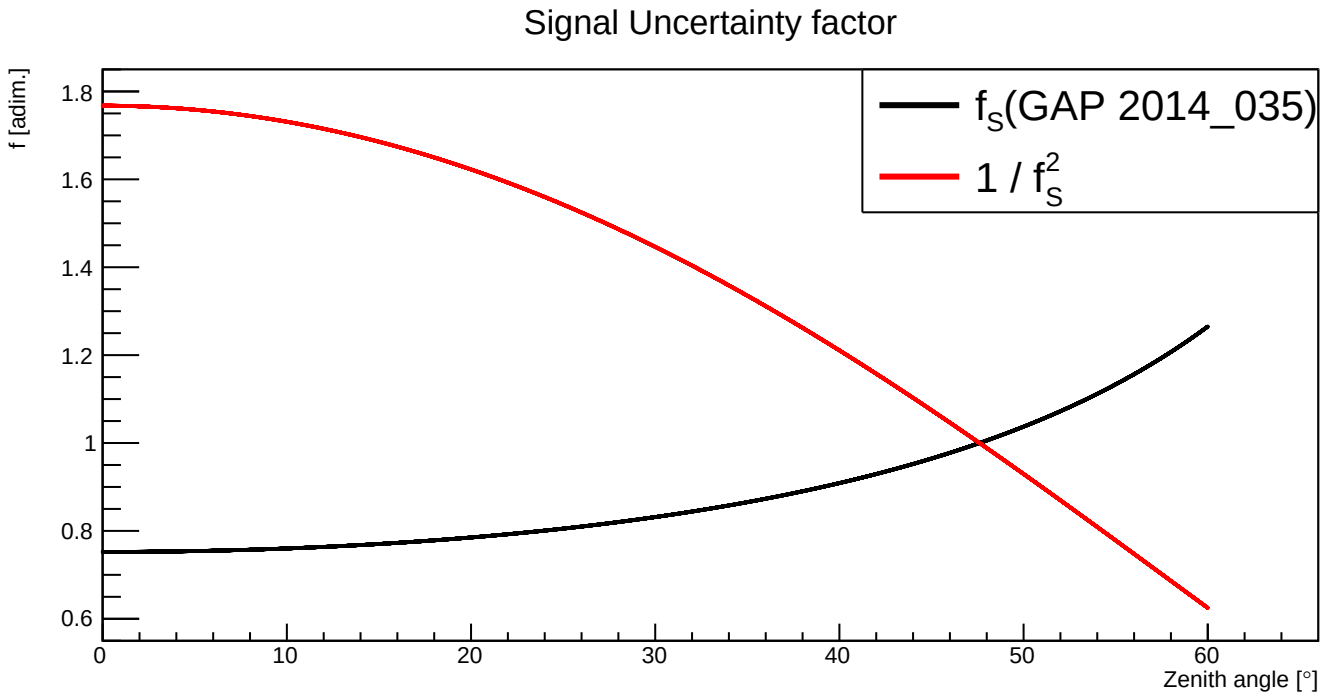
$$f_S(\theta) = 0.865 \cdot [1 + 0.593 \cdot (\sec\theta - \sec 35^\circ)]$$

Such starting assumption implies that the bin-to-bin correlations are not considered. Then it can be demonstrated to be wrong and, using the **cluster model**, a reduction of the uncertainties is obtained:

$$\epsilon(s_i) = 0.42\ f_S \times \sqrt{s_i}$$

Such value is employed in the optimized Universality reconstruction.

Uncertainty factors



$$\sigma(S/\mathrm{VEM}) = f_S \times \sqrt{S/\mathrm{VEM}}$$

Effective number of particles (poissonian variable)

$$n_{\mathrm{eff}} = f_S^{-2} \times S/\mathrm{VEM}$$

Constraints on the shower parameters

Axis constraint ←-- angular resolution of the surface detector.
C. Bonifazi, Nucl. Phys. Proc. Suppl. 190 (2009)

$$D_{\vec{a}} = \left(\frac{\widehat{\vec{a}_{\mathrm{Univ}}, \vec{a}_{\mathrm{SD}}}}{\sigma_{\vec{a}}} \right)^2 - 2 \ln \left(\frac{\widehat{\vec{a}_{\mathrm{Univ}}, \vec{a}_{\mathrm{SD}}}}{\sigma_{\vec{a}}} \right) - 1$$

Energy constraint ←-- resolution of the energy calibration.
I. Valiño, Contribution to the 34th ICRC (2015)

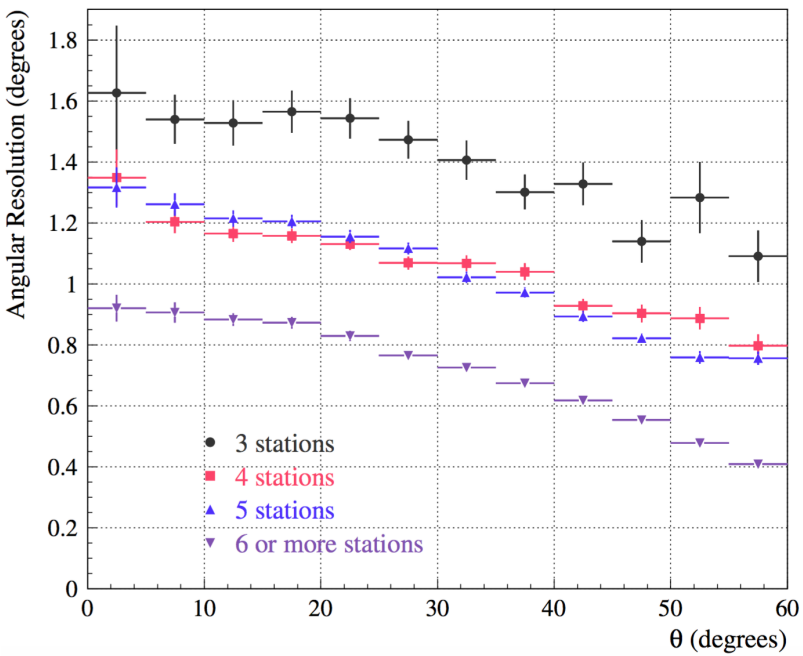
$$D_E = \left(\frac{E_{\mathrm{Univ}} - E_{\mathrm{SD}}}{0.15\ E_{\mathrm{SD}}} \right)^2$$

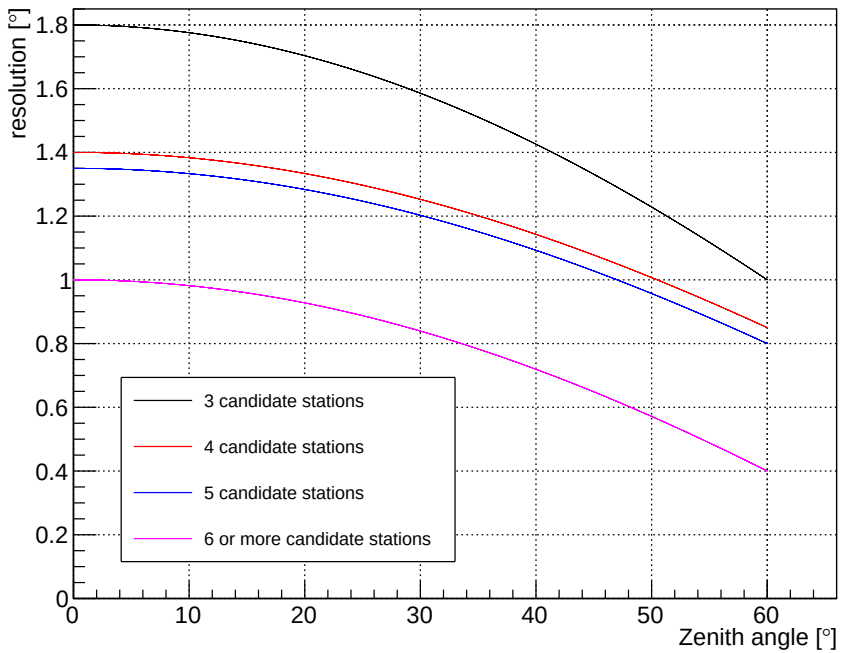
Core constraints ←-- comparison of SD and hybrid reconstructions.
Biases x_0 and t_0 due to **azimuthal asymmetry**.

$$D_{core} = \left(\frac{x_{\mathrm{Univ}} - (x_{\mathrm{SD}} + x_0)}{\sigma_x} \right)^2 + \left(\frac{y_{\mathrm{Univ}} - y_{\mathrm{SD}}}{\sigma_y} \right)^2 + \left(\frac{t_{\mathrm{Univ}} - (t_{\mathrm{SD}} + t_0)}{\sigma_t} \right)^2$$

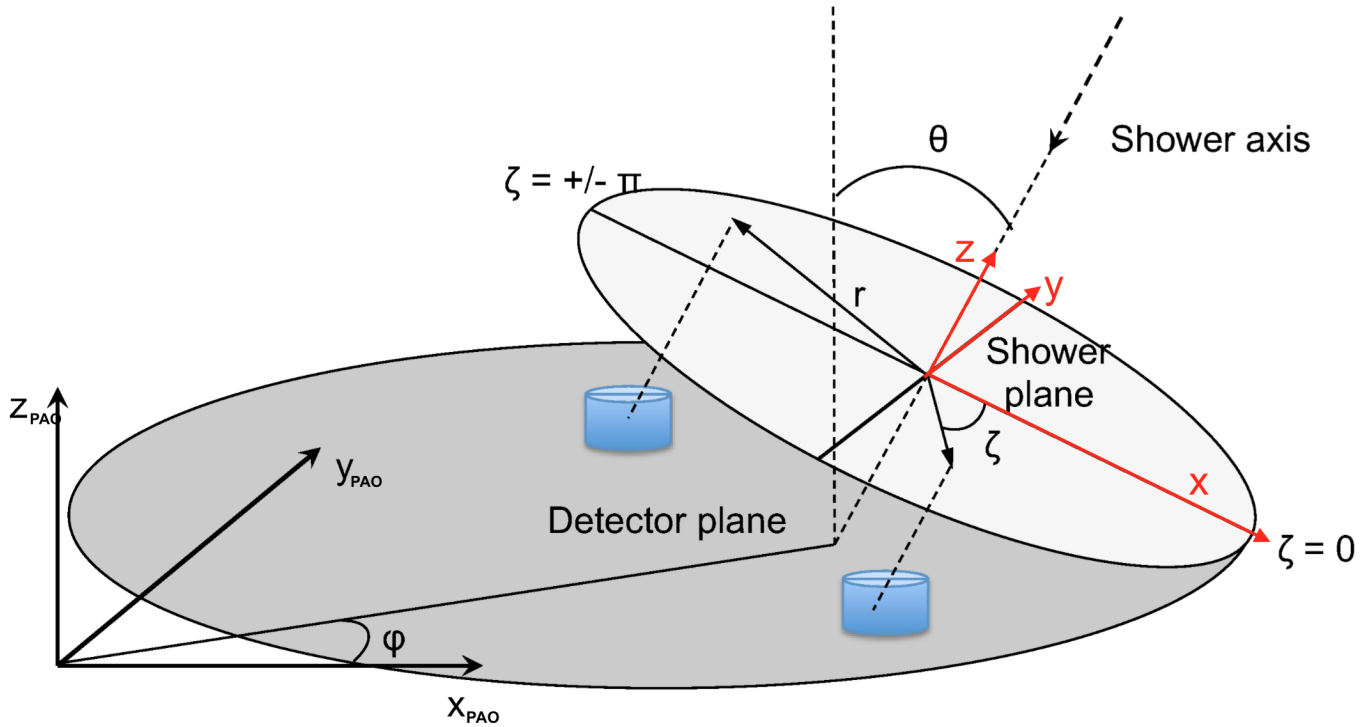
SD angular resolution

Angular resolution as a function of the zenith angle for events with an energy above 3 EeV (*C. Bonifazi, Nucl. Phys. Proc. Suppl. 190 (2009)*) and functions used in my work as approximations of the experimental behaviours.





Azimuthal asymmetry - Shower coordinate system

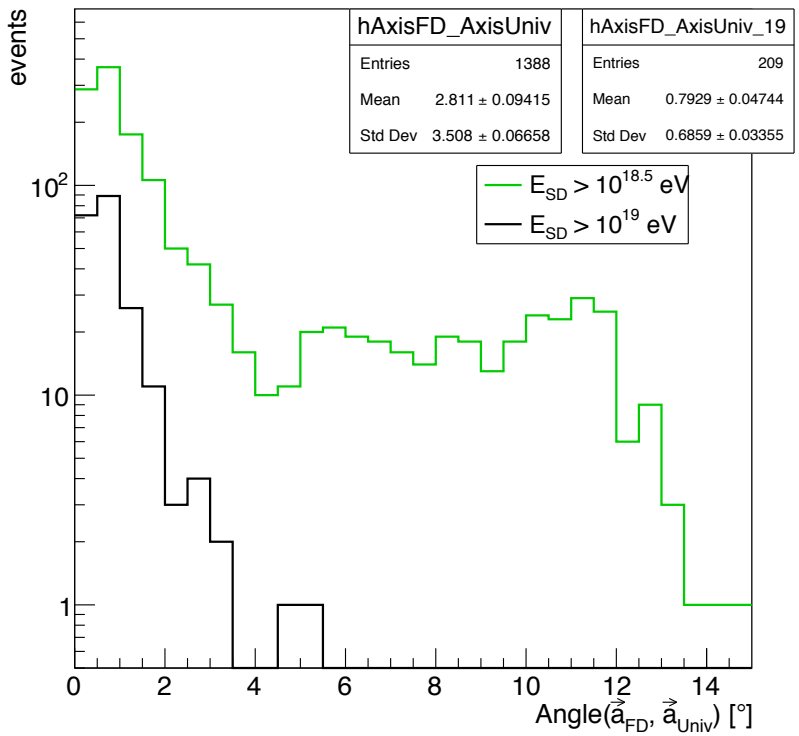


On the left, the fixed coordinate system of the Pierre Auger Observatory, where the azimuth (ϕ) and zenith (θ) of an event are defined.In red, the coordinate system relative to an event (*shower coordinate system*), where the shower plane is perpendicular to the shower axis.

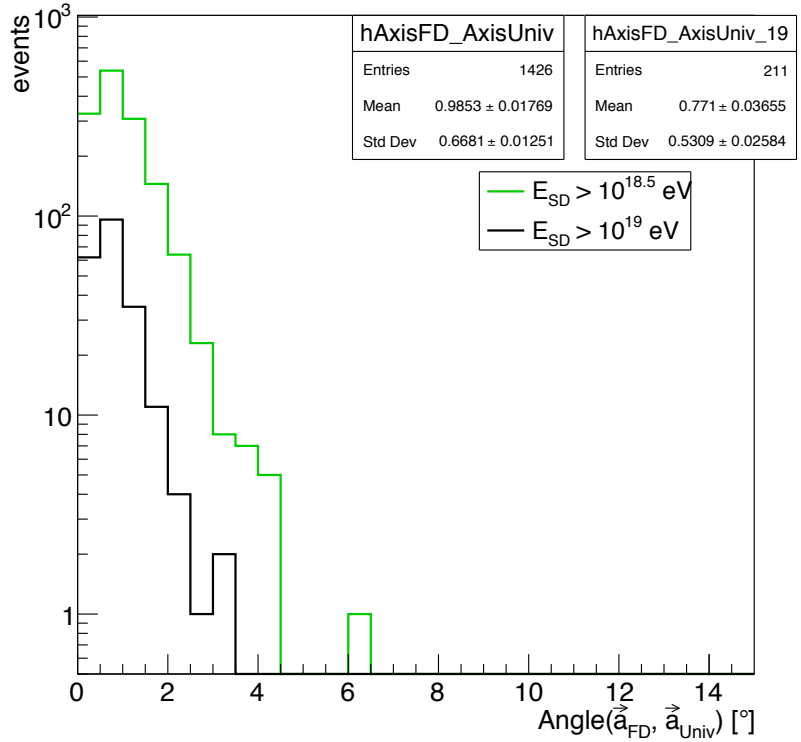
Geometrical reconstruction always accurate

The shower geometry (and in particular the axis) for events with $E > 10^{19}$ eV is determined with good accuracy in all the "adjusted" versions of the Universality reconstructions.

Energy fixed, deviance & bin error reduced



Energy released, deviance & bin error reduced, axis constrained



Details about the field of view cut

The reconstruction is considered reliable for zenith angles where a proton-initiated shower has at least 90% probability to develop with a maximum above ground (for a fixed energy value), evaluated using **generalized Gumbel distributions**.

The **low zenith cut** $\theta_{\rm cut}(E)$ is obtained from an average among the results for 3 different hadronic interaction models:

$$\theta_{\rm cut} = a + b\ lgE + c\ (lgE)^2 + d\ (lgE)^3$$

$lgE = (\log_{10}(E/eV) - 19)$, $a = 21.3°$, $b = 0.8°$, $c = 8.8°$, $d = -4.5°$.

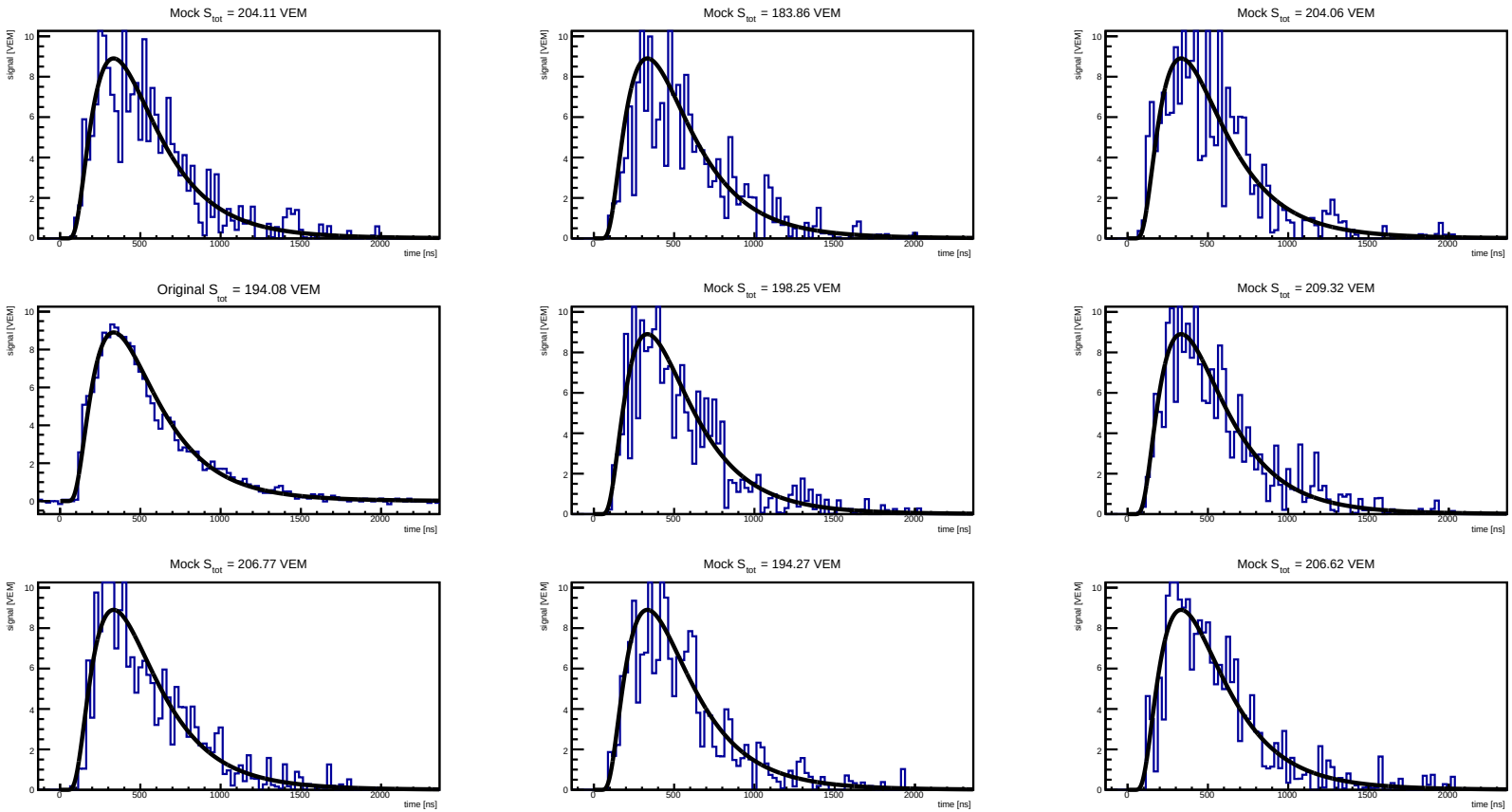
The cut $\theta < 55°$ allows to remove many events with poor reconstruction accuracy and with a larger $X_{\rm max}^{\rm Univ}$ bias w .r. t. the golden dataset as a whole.

Additional quality cuts

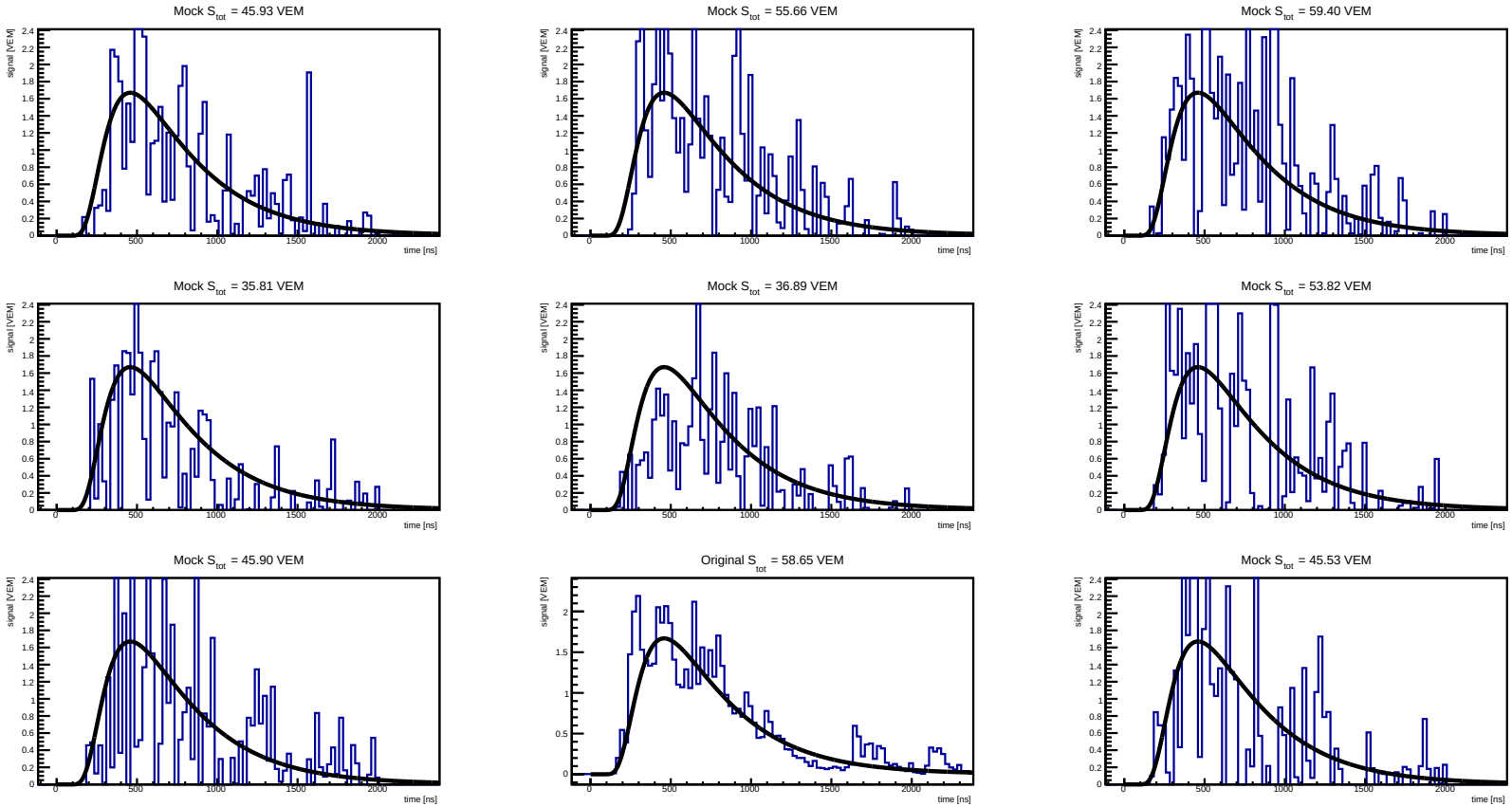
The optimized Universality minimization converges around 90% of the times over the previously selected set of events. Then two quality cuts are performed to reject outliers:

- The predicted pure electromagnetic signal in the hottest tank ($\hat{S}^{\mathrm{hottest}}_{\mathrm{e.m.pure}}$) is required to be larger than 1 VEM, rejecting events with small $X^{\mathrm{Univ}}_{\mathrm{max}}$ values (i.e. near the low boundary of 200 g/cm^2) or unrealistic couples of $X^{\mathrm{Univ}}_{\mathrm{max}}$ and N_μ.
- The ratio between the sum of all the deviance terms D_{tot} and the total number of "degrees of freedom" N_{tot} (given by the number of bins employed in the shape fit plus the number of stations in the LDF fit) is required to be $D_{\mathrm{tot}}/N_{\mathrm{tot}} > 4$.

"Quasi-poissonian" assumption - high signal example

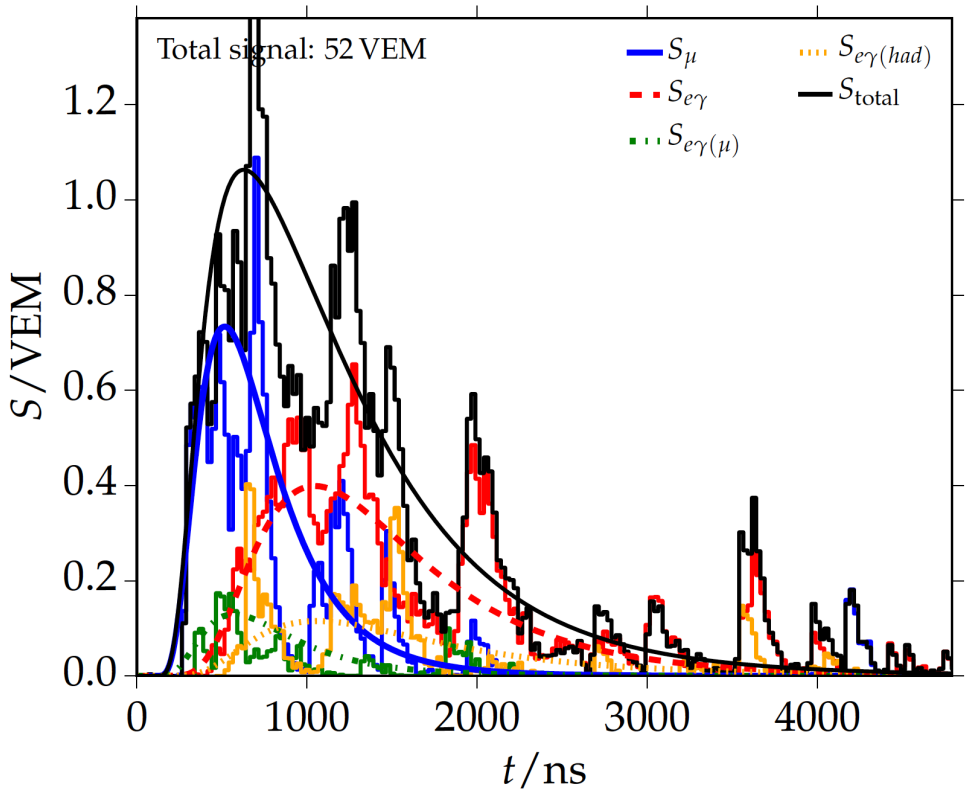


"Quasi-poissonian" assumption - low signal example



The clusters of signal

Isolated peaks are clearly visible in the measured traces.
Given appropriate simulations, **clusters** of correlated bins are observed also in each of the 4 Universality components.

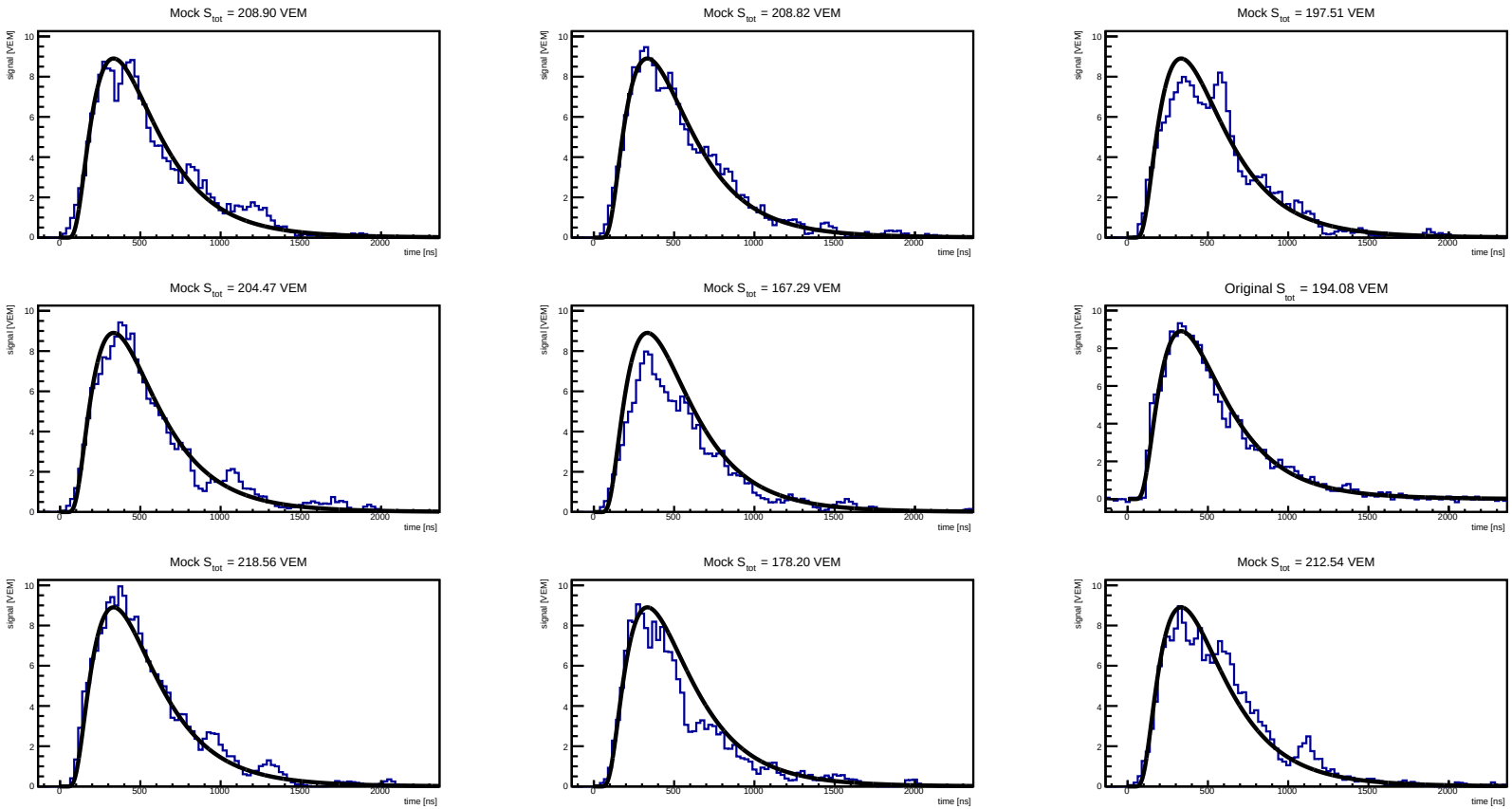


This suggests a simple model for generating random traces from a basic assumption: traces are the sum of clusters with (approximately!) the same basic shape.

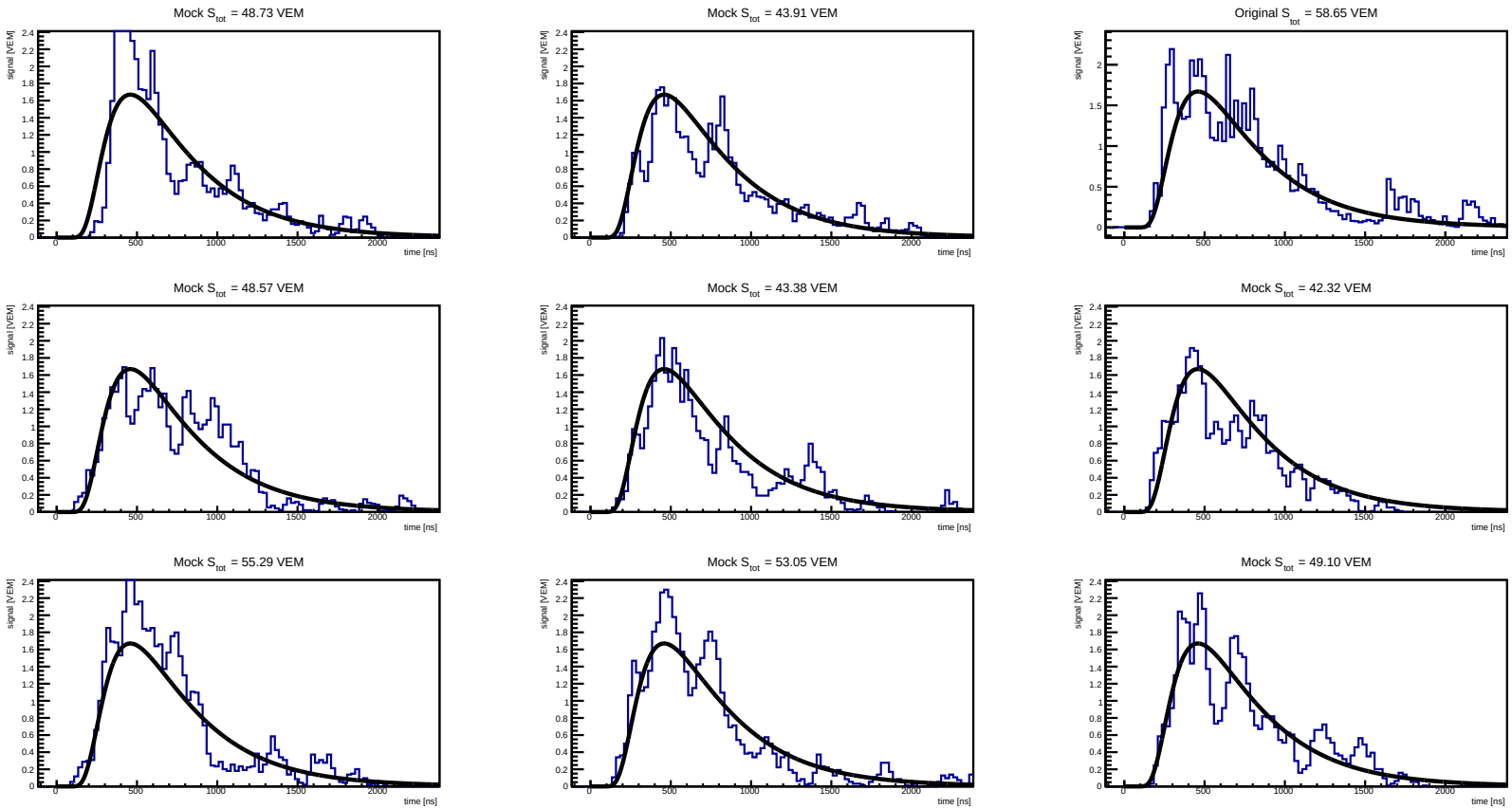
The cluster model assumptions

1. the basic block of each trace is a cluster with mean integrated signal $\langle S_{\rm clu} \rangle = (1 \times f_S^2)$ VEM, ideally correspondent to the signal produced by one *effective particle*; $f_S = f_S(\theta)$ is the uncertainty factor.
2. the cluster's time distribution is assumed to be a log-normal with scale parameter $m = 100$ ns and shape parameter $s = 0.4$ (the location parameter is put equal to 0).
3. each cluster has a mean content of $\langle n_{p.e.} \rangle = 94 \times f_S^2$ photoelectrons (number of p.e. per VEM ~ 94). The actual number of p.e. in a cluster is extracted from a Poisson distribution with mean equal to $\langle n_{p.e.} \rangle$.
4. the mean number of clusters for each Universality component i is $\langle N_i^{\rm clu} \rangle = S_i / S_{\rm clu}$, where S_i is the predicted signals from the fit. The number of clusters $N_i^{\rm clu}$ for the i-th component of an actual mock trace is extracted from a Poisson distribution with mean equal to $\langle N_i^{\rm clu} \rangle$.
5. the distribution of the clusters temporal positions in the trace corresponds, separately for each shower component, with the related log-normal function obtained from the shape fit.

Cluster model mock traces - high signal example

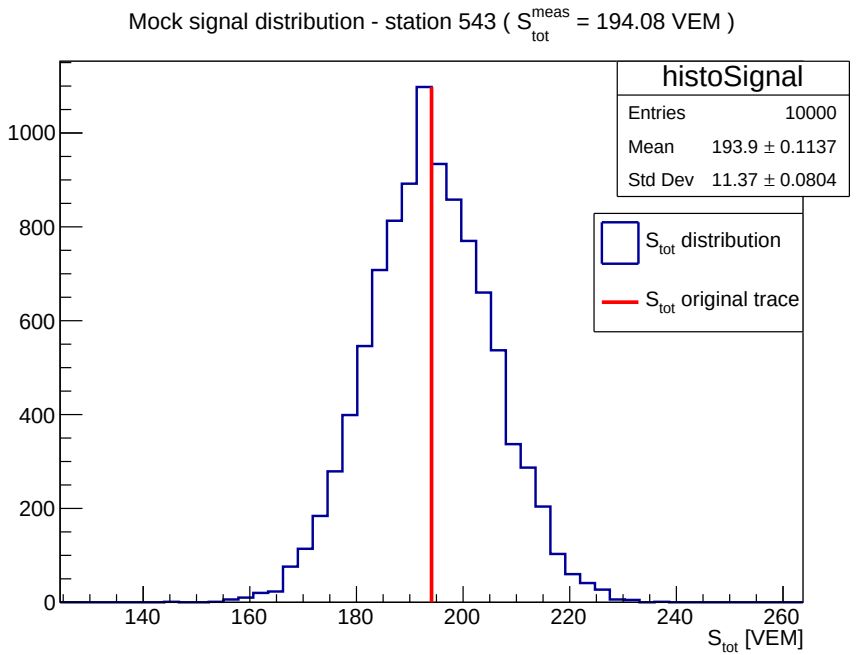


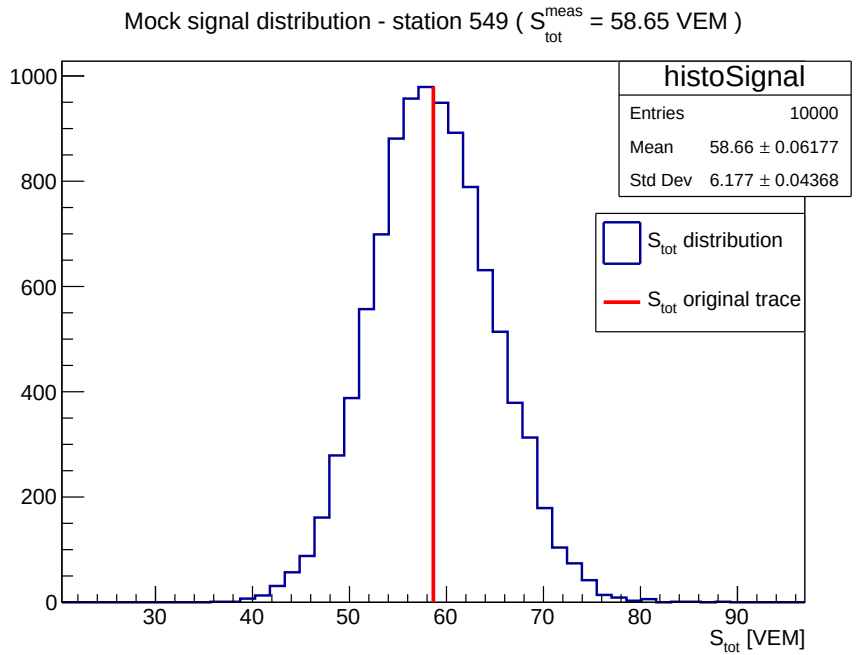
Cluster model mock traces - low signal example



Cluster model mock traces - distributions of total signals

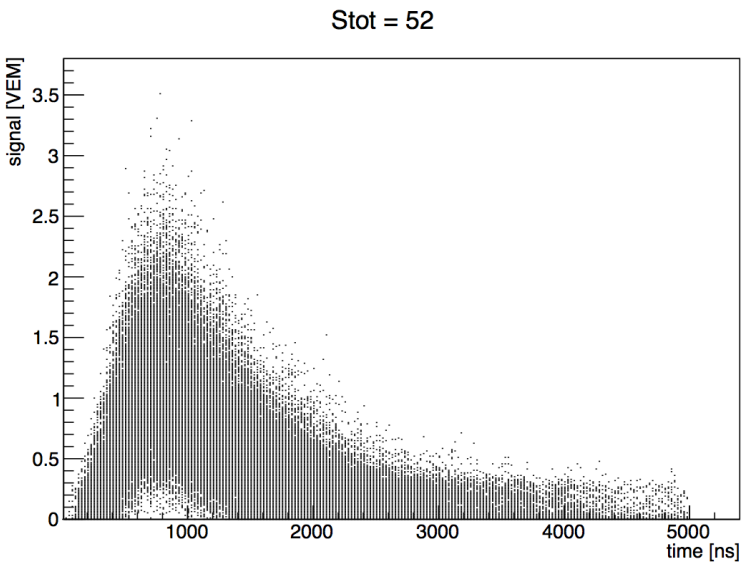
The integrated signal of a set of mock traces generated from the same original one through the cluster model results distributed as in the experimental observations, in particular with a dispersion equal to the integrated signal fluctuations $f_S \times S_{\mathrm{tot}}$.

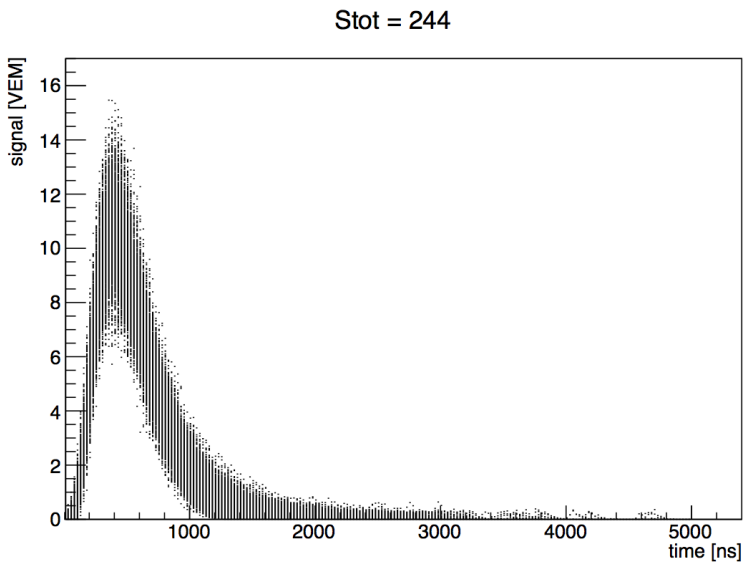




Bin signal dispersion with cluster model

The needed reduction of the bin content uncertainties can be quantified requiring the compatibility between such uncertainties and the bin signal fluctuations.
Using the toy model, one can generate a large set of mock traces from a single log-normal and then calculate the mean and dispersion of the signal in each bin.

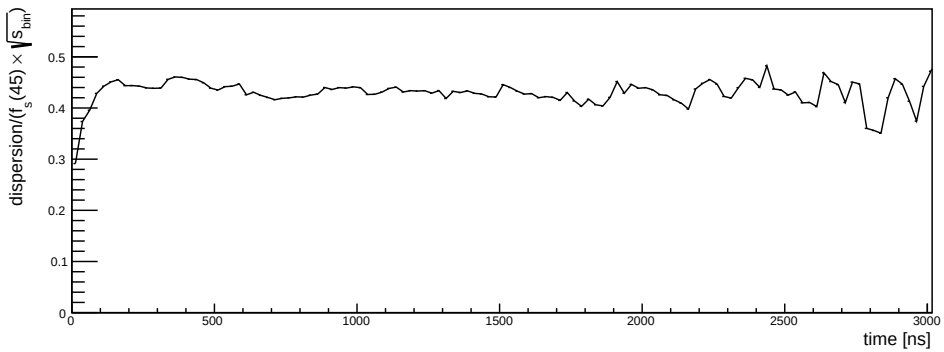


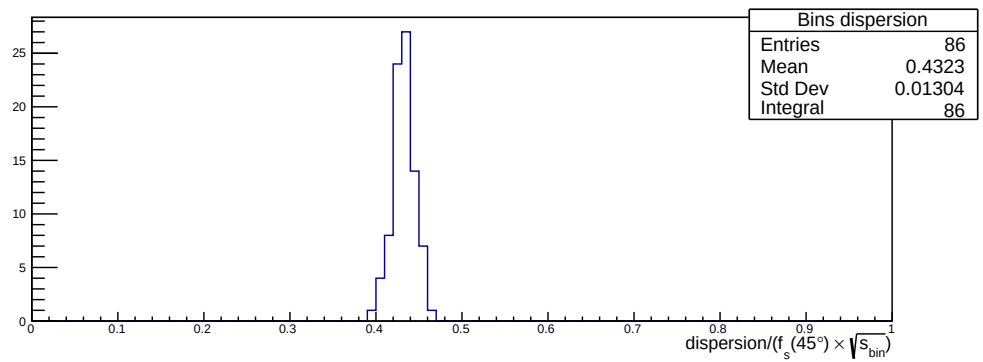


Bin signal dispersion with cluster model

The idea is to compare the bin content dispersion $\sigma[s_i]$ with the "quasi poissonian" bin signal errors $\epsilon(s_i) = f_S \times \sqrt{s_i}$.
In particular, the ratio $r_i = \sigma[s_i]/(f_S \times \sqrt{s_i})$ is studied.

The dispersion $\sigma[s_i]$ in a single Log-Normal is found to be on average less than half of $\epsilon(s_i) = f_S \times \sqrt{s_i}$, confirming that a reduction of such errors is necessary, also to accomplish more correctly the shape fit.





Reducing the "quasi poissonian" bin errors

The values of $r_i = \sigma[s_i]/(f_S \times \sqrt{s_i})$ are studied for different integrated signals, zenith angles, shapes of the traces (i.e. different values of the Log-Normal parameters m and s).

It arises that the reducing factor does not substantially change inside the chosen parameter space, assuming only values in the (quite) small range [0.41, 0.44].

Therefore the uncertainty over the bin content can be safely resized (as previously reported):

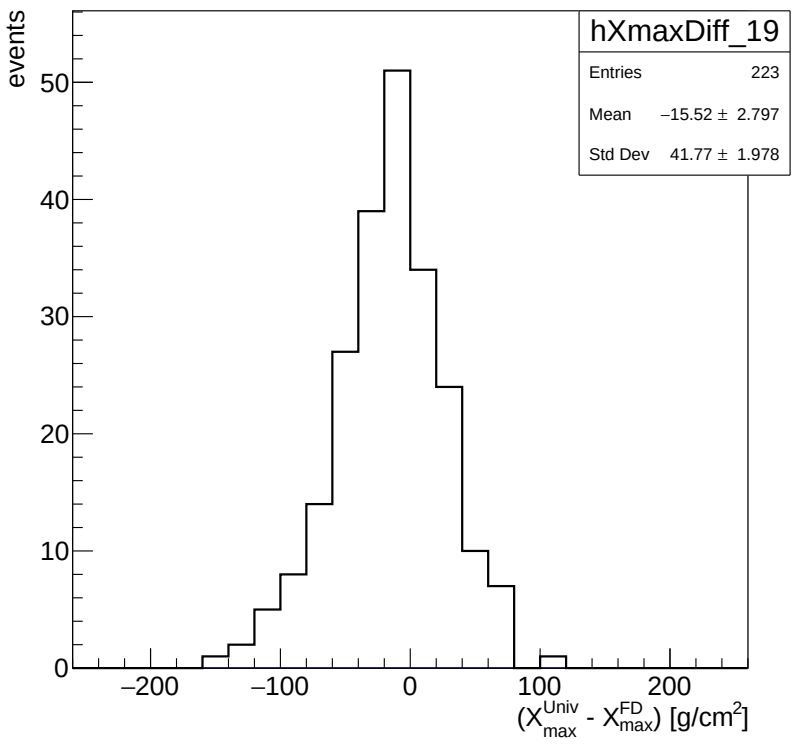
$$\epsilon(s_i) = 0.42\ f_S \times \sqrt{s_i}$$

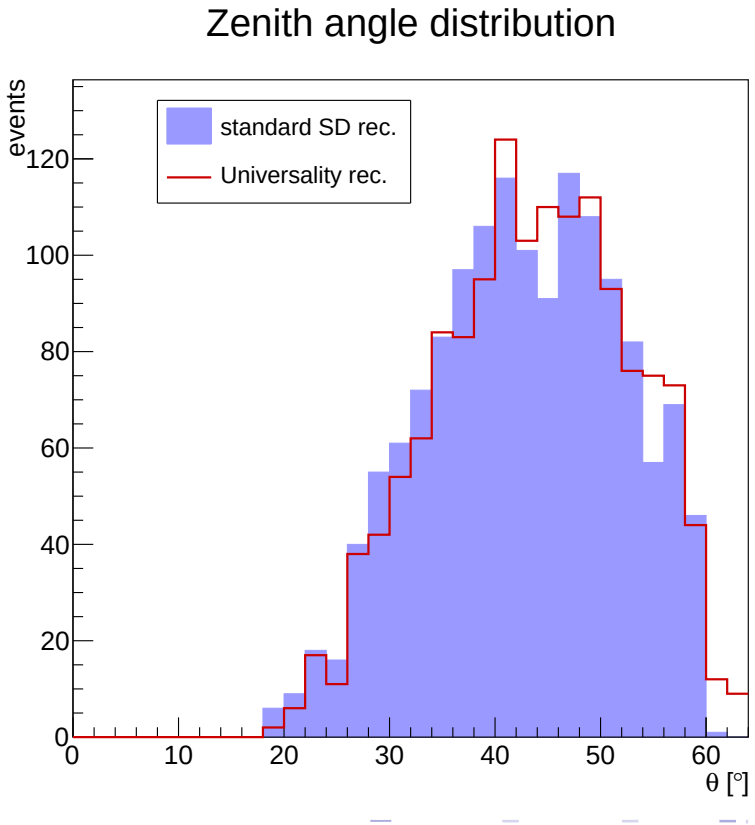
Outcome after deviance introduction

223/228 unsaturated events with $E_{SD} > 10^{19}$ eV reconstructed.

$\left\langle X_{\max}^{\mathrm{Univ}} - X_{\max}^{\mathrm{FD}} \right\rangle = -15.5$ g/cm^2 (bias appearance).

$\mathrm{RMS}(X_{\max}^{\mathrm{Univ}} - X_{\max}^{\mathrm{FD}}) = 41.8$ g/cm^2 (smaller than original fit).



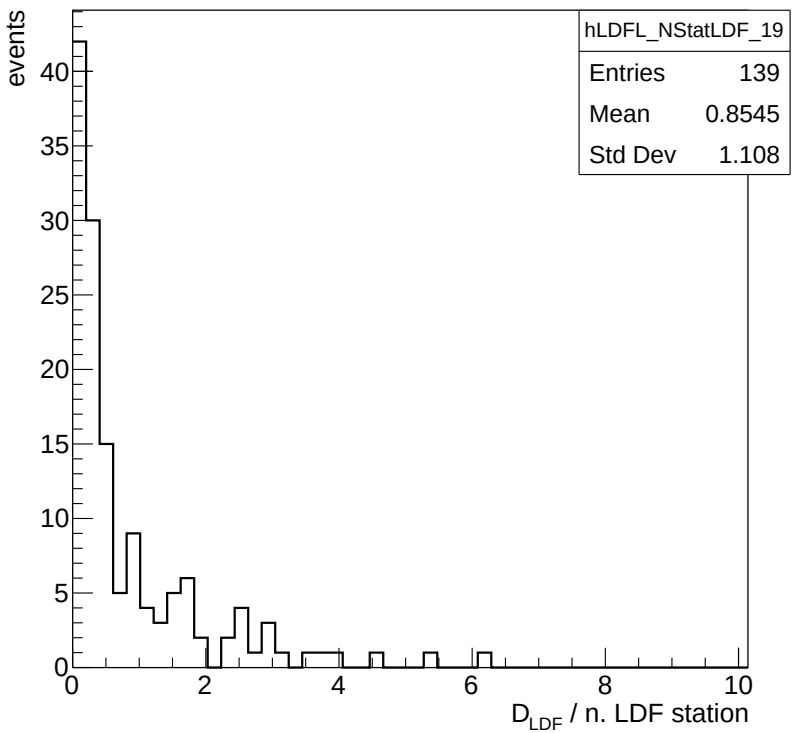


Outcome after deviance introduction - Deviances

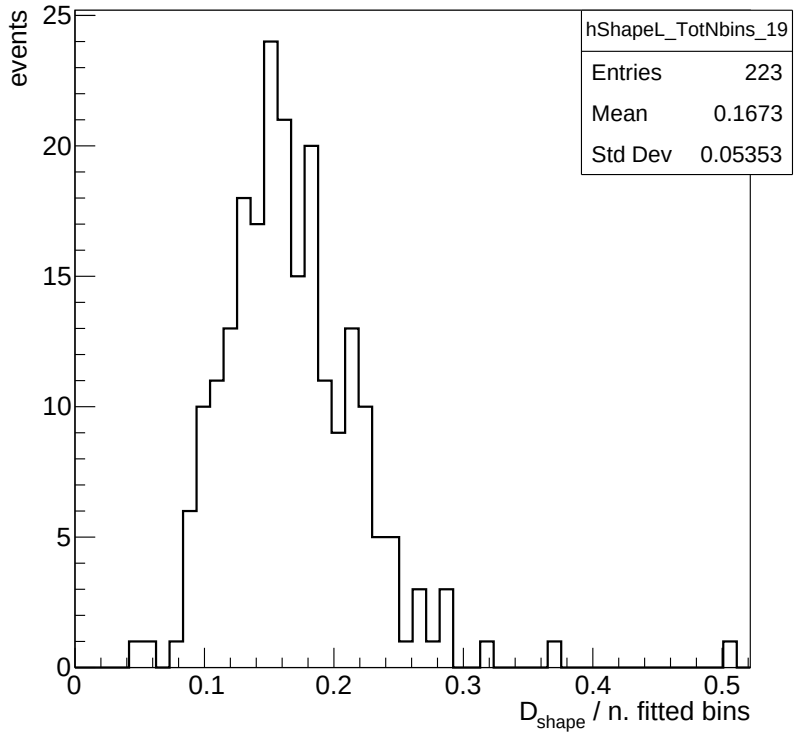
D_{LDF}/n. stations for LDF likelihood similar to a χ^2 distribution with few degrees of freedom.

Values of D_{shape}/total number of fitted bins noticeably small.

Tot. LDF Deviance per station



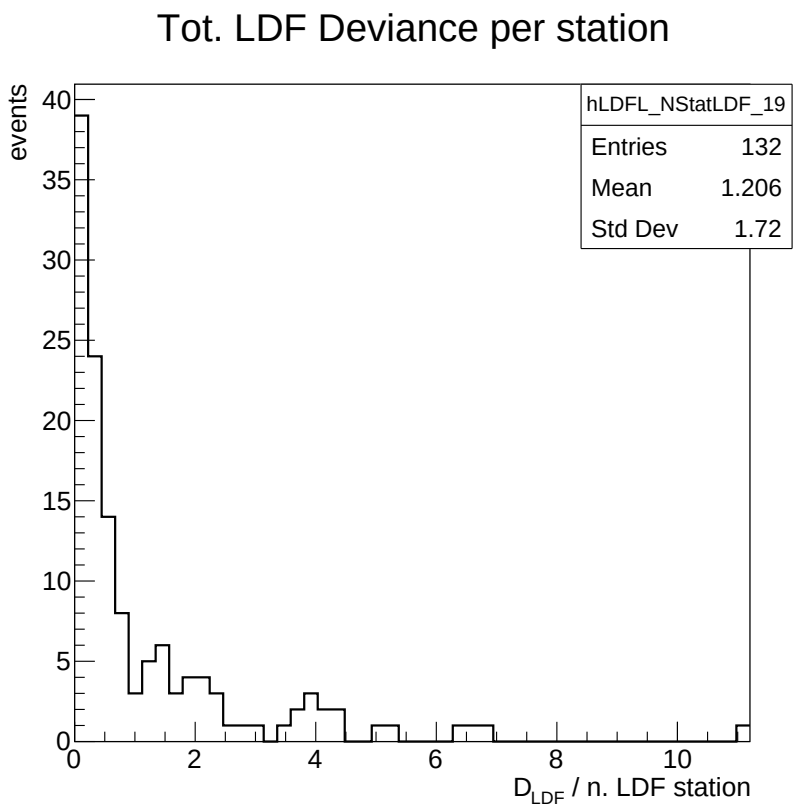
Tot. shape Deviance per fitted bin

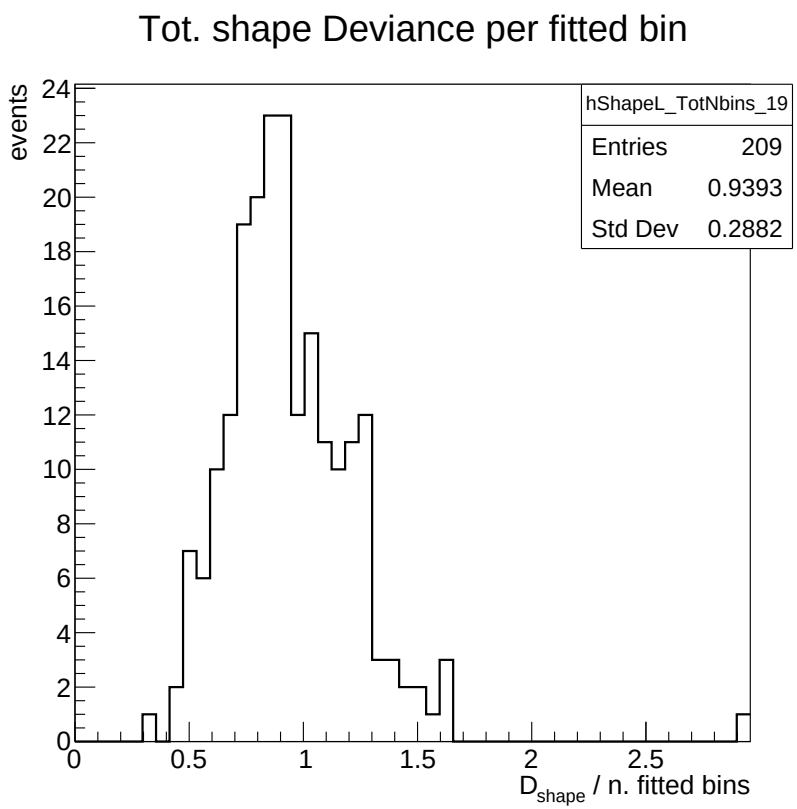


Outcome after errors reduction - Deviances

D_{LDF}/n. stations for LDF likelihood as good as before.

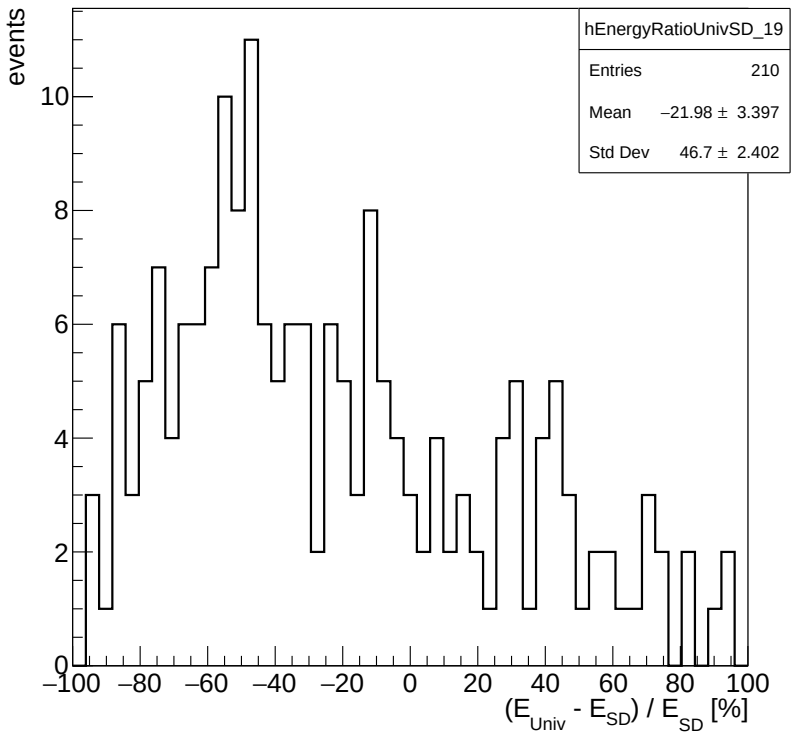
Values of D_{shape}/total number of fitted bins noticeably increased, showing now a distribution centered approximately around 1.

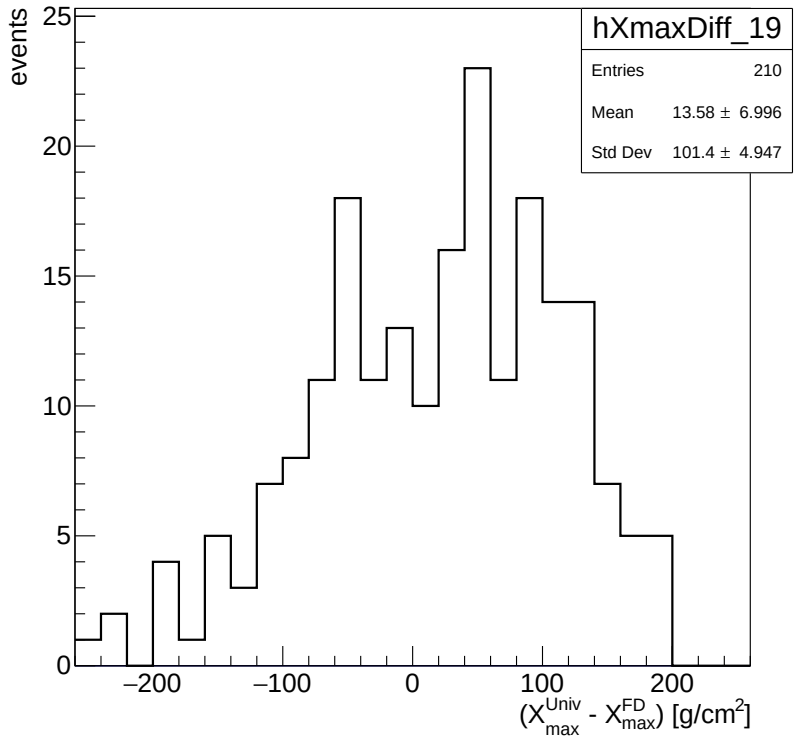




Constrained reconstruction - no energy constraint

The reliability of E_{Univ} rapidly deteriorates for increasingly wider energy constraint and, as a consequence, the reconstruction outcomes (in particular the $(X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}})$ distribution) get worse.

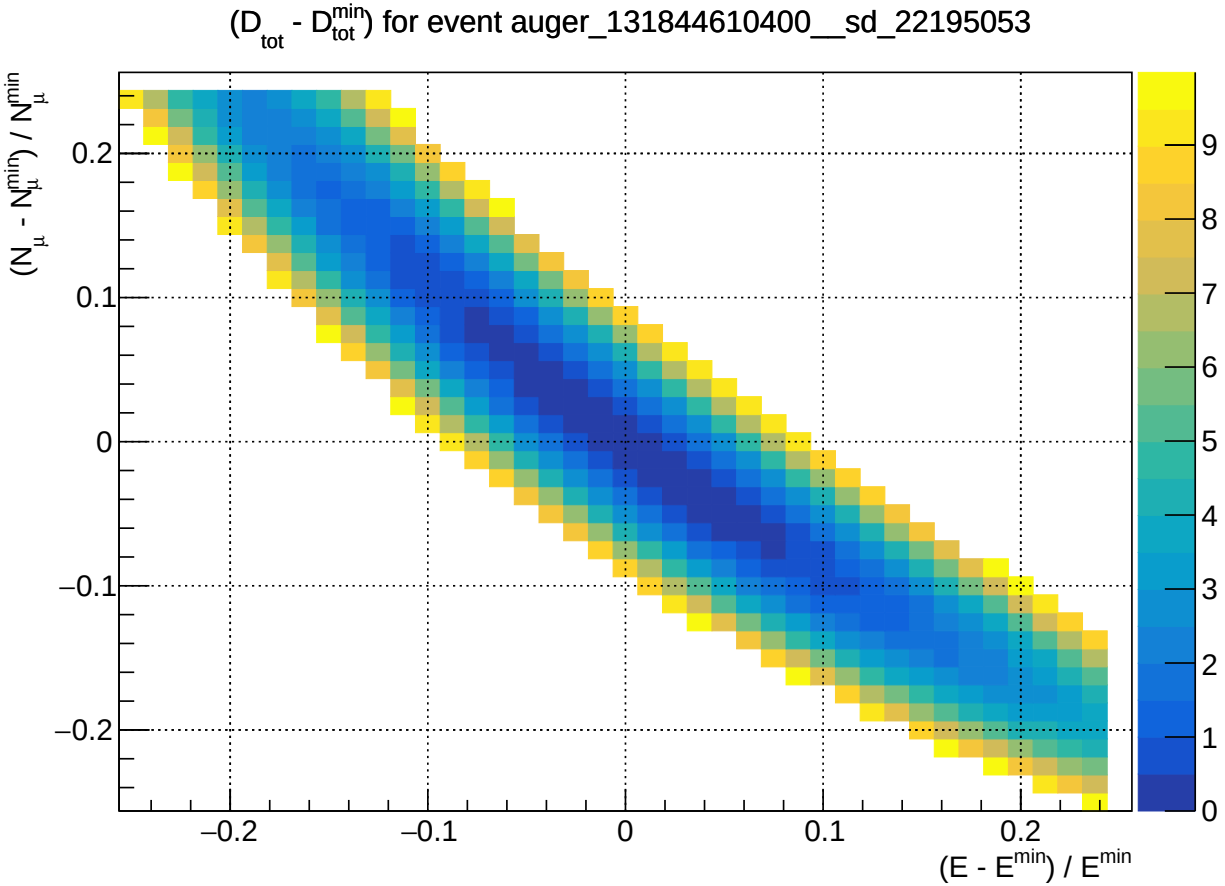




Reconstruction outcomes when the energy constraint is removed.

Degeneracy of minimum

Degeneracy of the Universality fit due to the **(anti-)correlation between energy and muon content** ...



... meaning that, in each event, different $(E_{\mathrm{Univ}}, N_\mu)$ couples can produce nearly the same prediction for the signals at ground level.

$E - N_\mu$ anti-correlation in the signal parameterization

In first approximation:

$$S_{em} \propto K_{em}(\Delta X, r)\ (0.925 - 0.075 N_\mu) \left(\frac{E}{10^{19}\ eV}\right)$$

$$S_{\mu} \propto K_{\mu}(\Delta X, r)\ N_\mu \left(\frac{E}{10^{19}\ eV}\right)^{0.93}$$

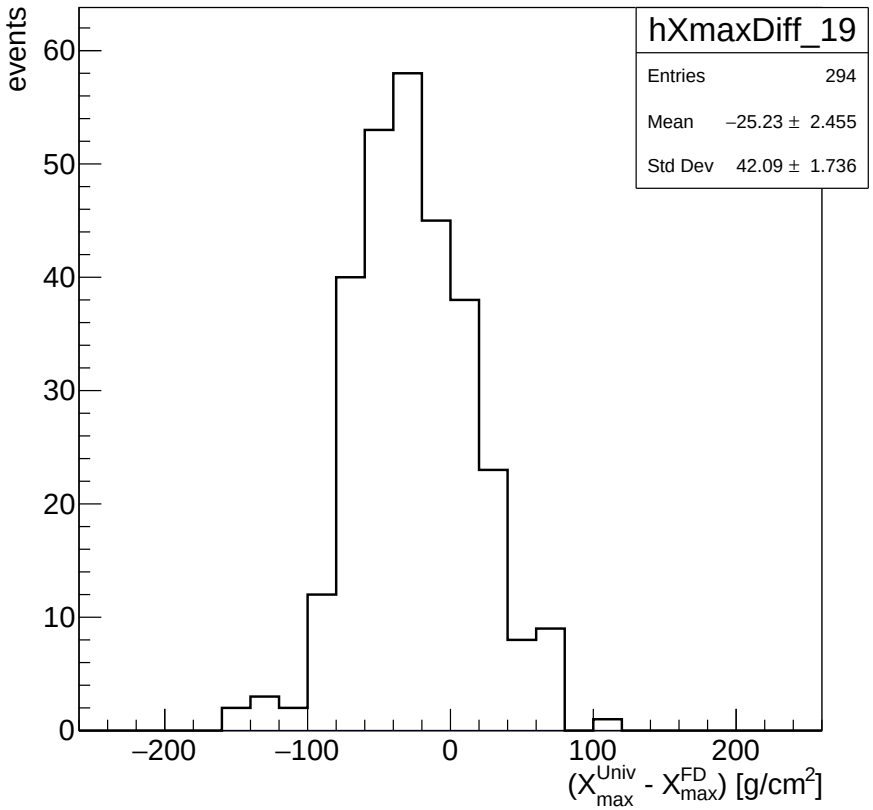
$$S_{em(\mu)} \propto K_{em(\mu)}(\Delta X, r)\ N_\mu \left(\frac{E}{10^{19}\ eV}\right)^{0.93}$$

$$S_{em(had)} \propto K_{em(had)}(\Delta X, r)\ (1.2 N_\mu - 0.2) \left(\frac{E}{10^{19}\ eV}\right)^{0.93}$$

Then three over four components (i.e. the muonic, e.m. from muon and e.m. from hadrons) are roughly functions of the product of energy and N_μ, apart from dependencies on ΔX and r.

Outcome of the simplified reconstruction - ICRC-2017

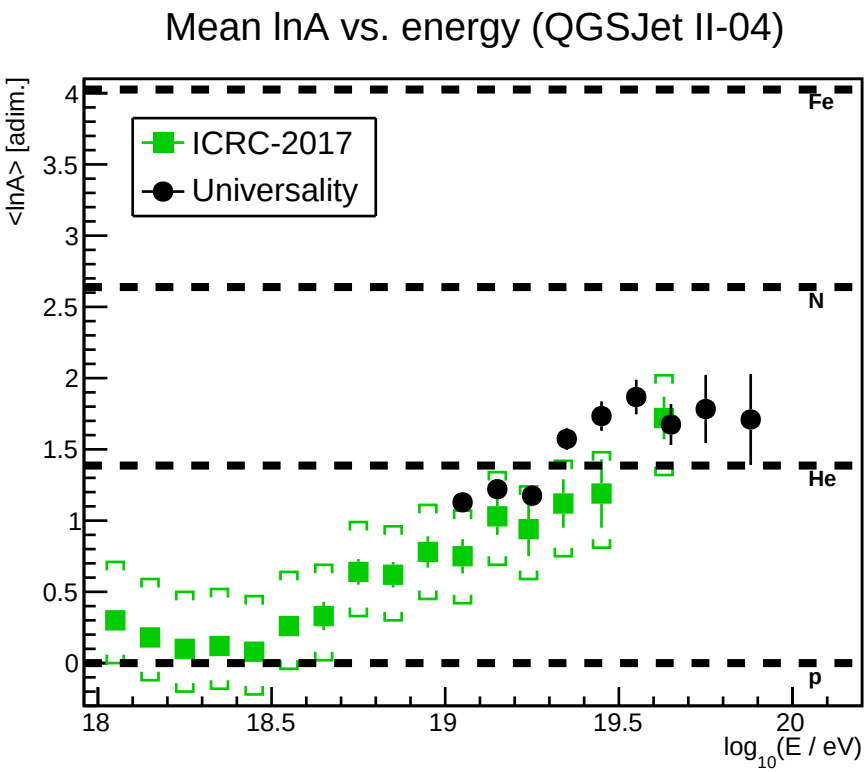
294/334 unsaturated golden hybrids with $E_{\mathrm{SD}} > 10^{19}$ eV and $\theta < 60^{\circ}$ used for the **ICRC-2017** energy calibration are successfully fitted through the simplified reconstruction.

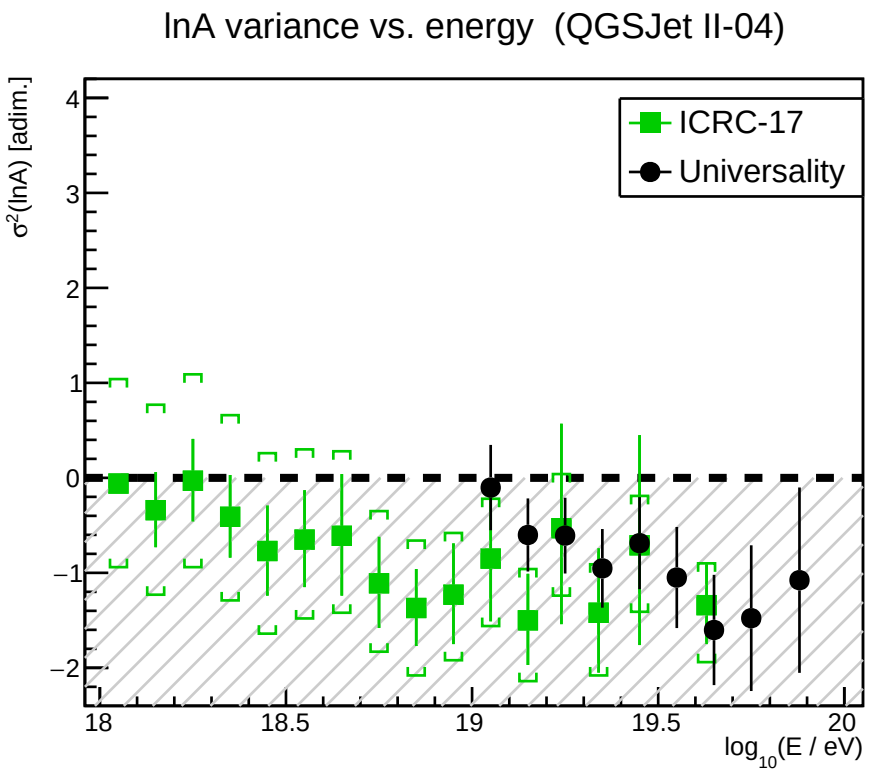


The $(X_{\mathrm{max}}^{Univ} - X_{\mathrm{max}}^{FD})$ bias is substantially larger than the ICRC-2015 one.

This shift is due to changes in the golden event selection, in the energy scale and in the energy calibration.

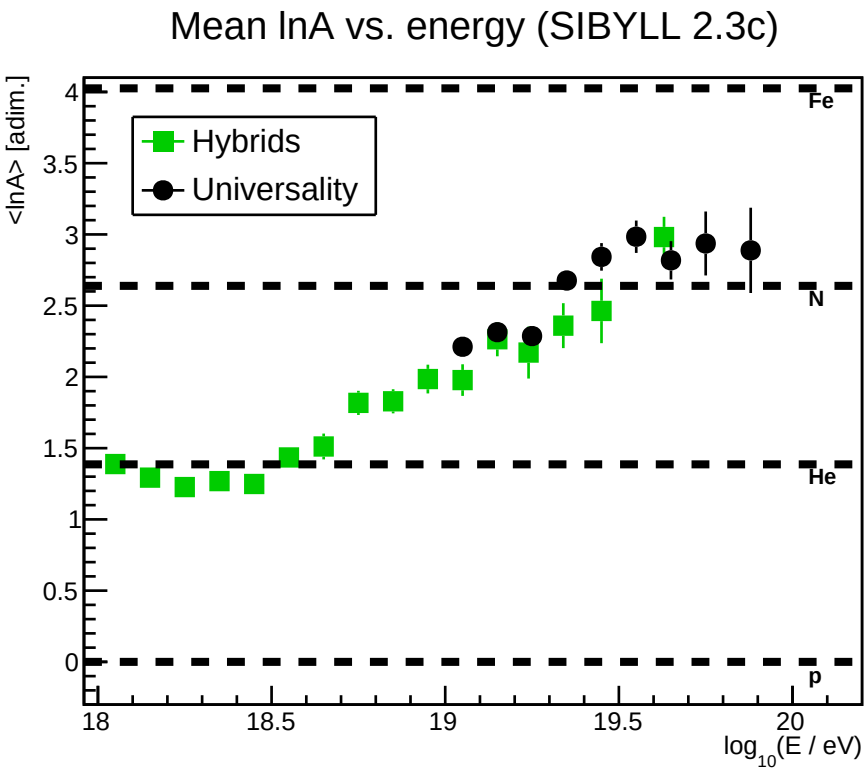
ln A moments for the SD-1500 dataset (QGS-Jet II-04)

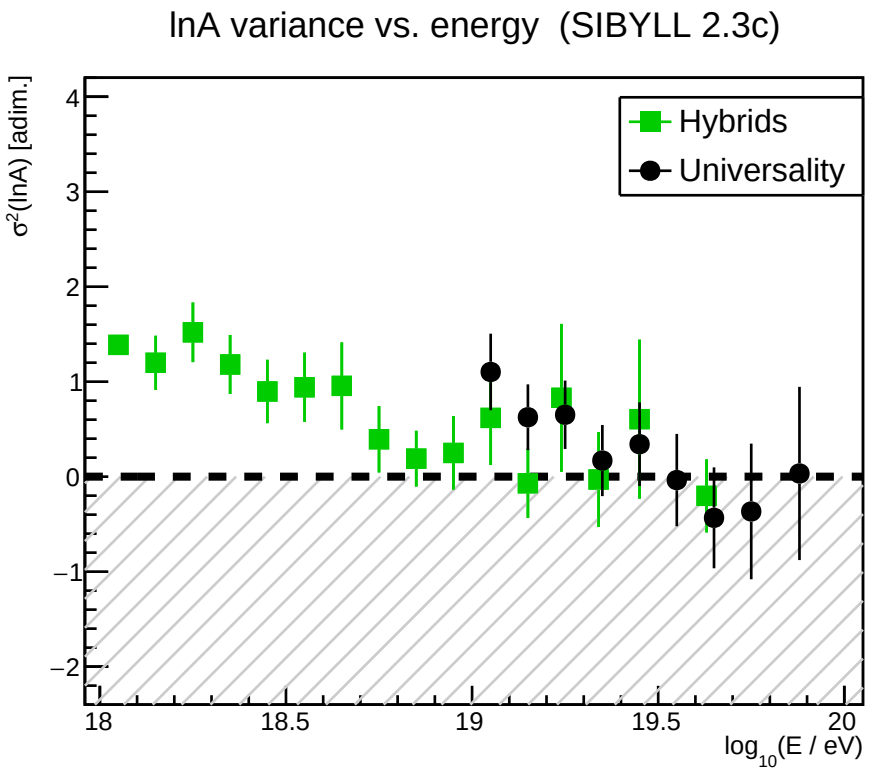




QGS-Jet II-04 fails to provide consistent interpretation of data, considering the negative values in the ln A dispersion plots.

ln *A* moments for the SD-1500 dataset (Sibyll 2.3c)



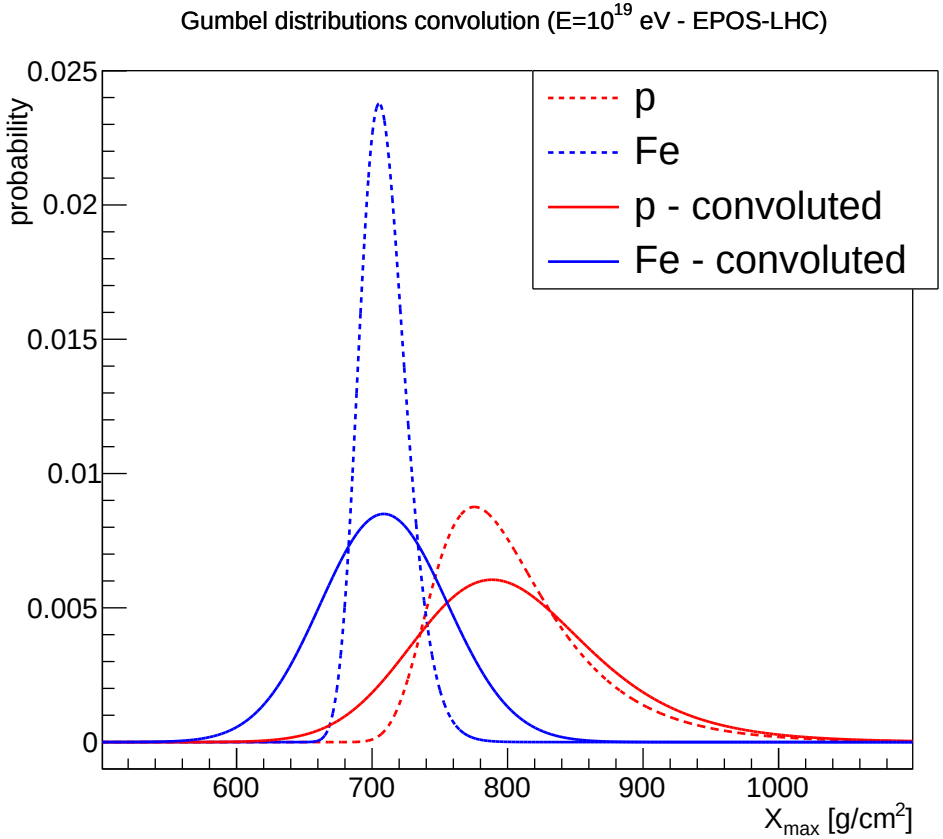


The Sibyll model was updated to the version 2.3c after the ICRC-2107. Therefore, the reference values for the hybrid analysis are taken from GAP 2018-021.

The fractions fit - Gumbel functions

The **generalized Gumbel functions** $g(X_{\max}|E, A)$ have been employed as expected $X_{\max}$ distributions, using the updated parameterization in *GAP 2018-021*.

Each parameterized Gumbel functions has to be corrected for the Universality resolution through a **convolution with a gaussian** with variance equal to σ_{Univ}.



Fractions fit systematics

A hint about the magnitude of systematic uncertainties in the fractions fit of the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distribution is obtained considering an artificial shift of $+5$ g/cm^2 (well inside the 68% uncertainty region of the parameterized Universality bias), as shown in the next slide.

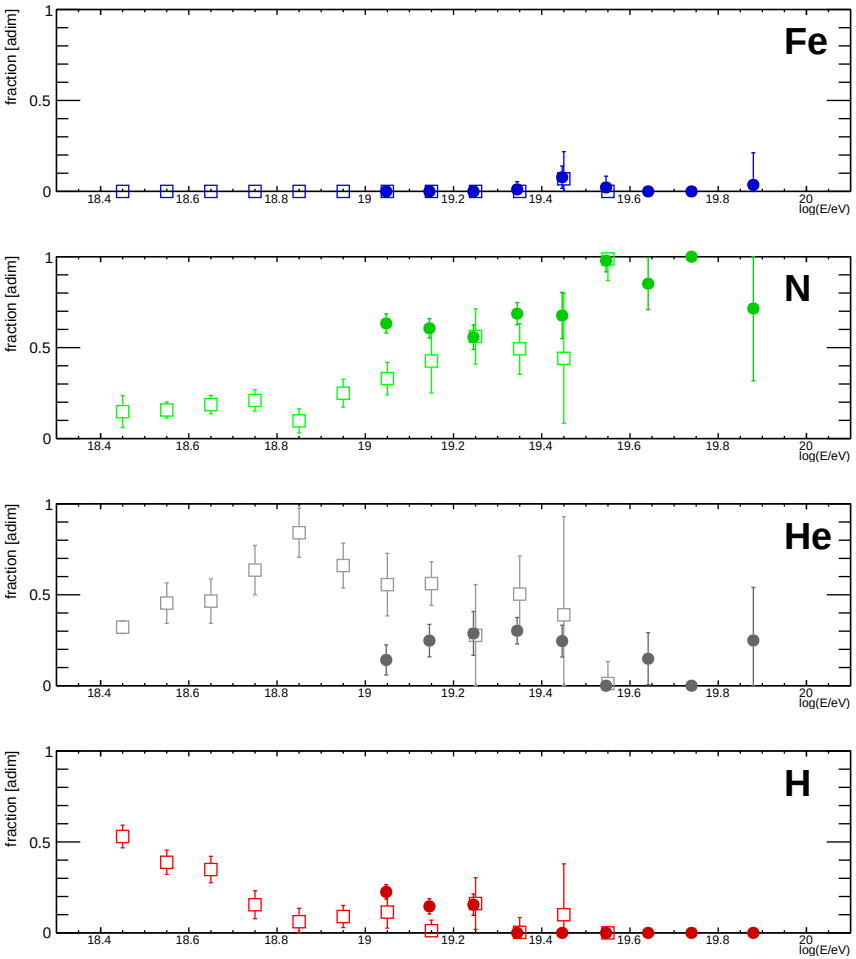
The main sources of systematic uncertainties are the bias and resolution of the Universality method:

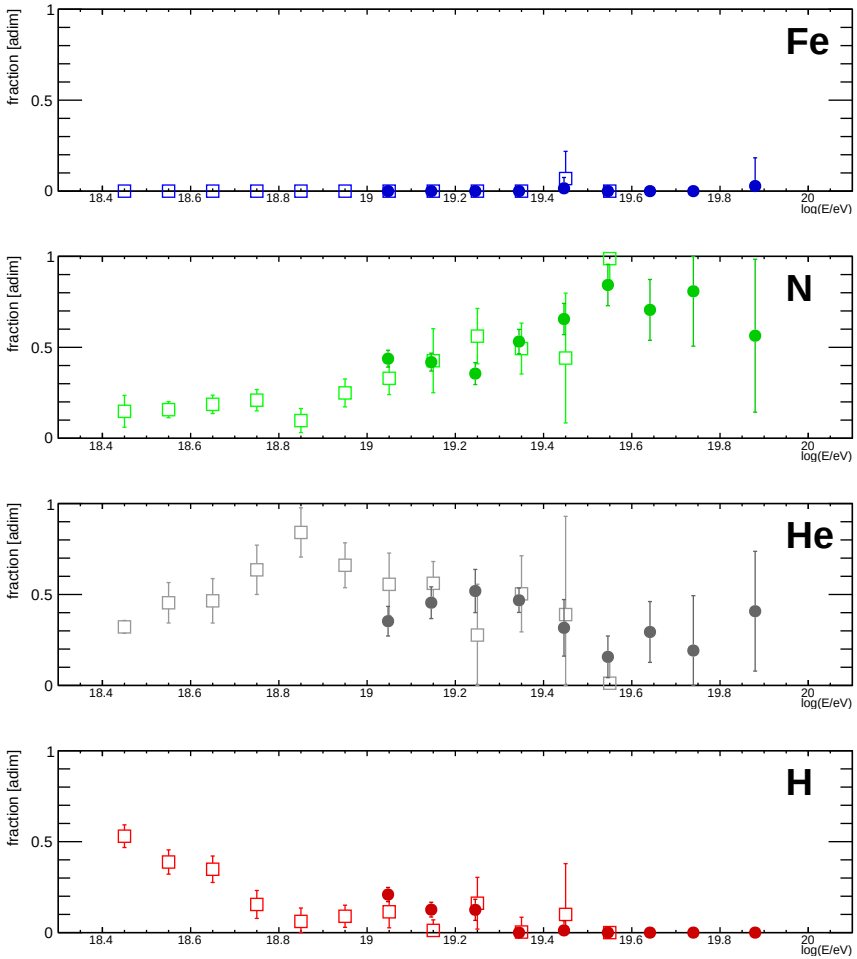
- the assumption of a gaussian representation for the Universality resolution could be simplistic (for instance, it could be asymmetrical instead);
- the choice of a rigid shift of the distributions to correct the Universality bias could be inappropriate: more sophisticated methods may be required to recover the correct mean value.

Fractions fit - rigid shift toward deeper $X_{\mathrm{max}}^{\mathrm{Univ}}$

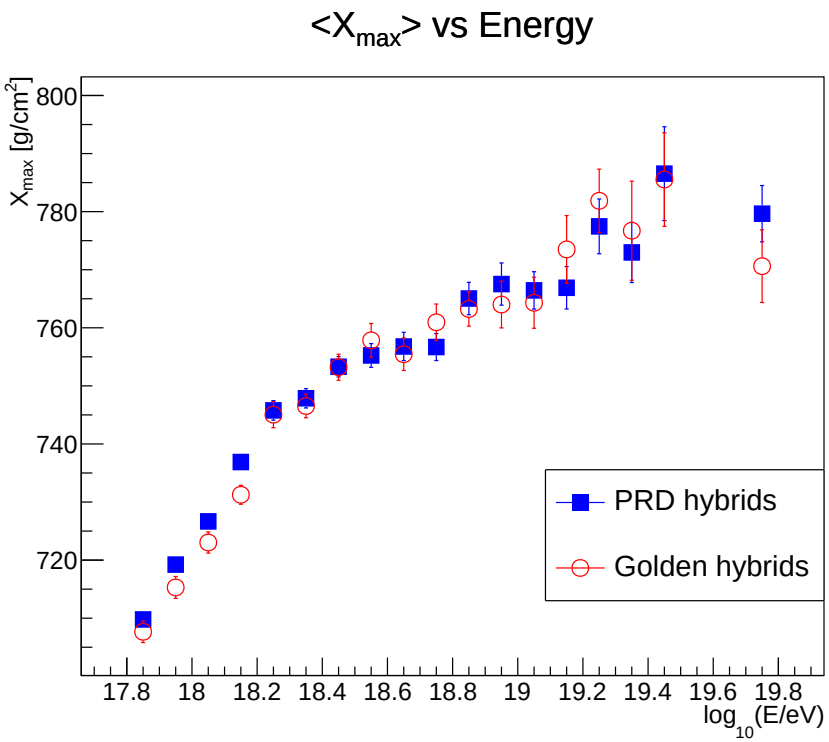
Unchanged

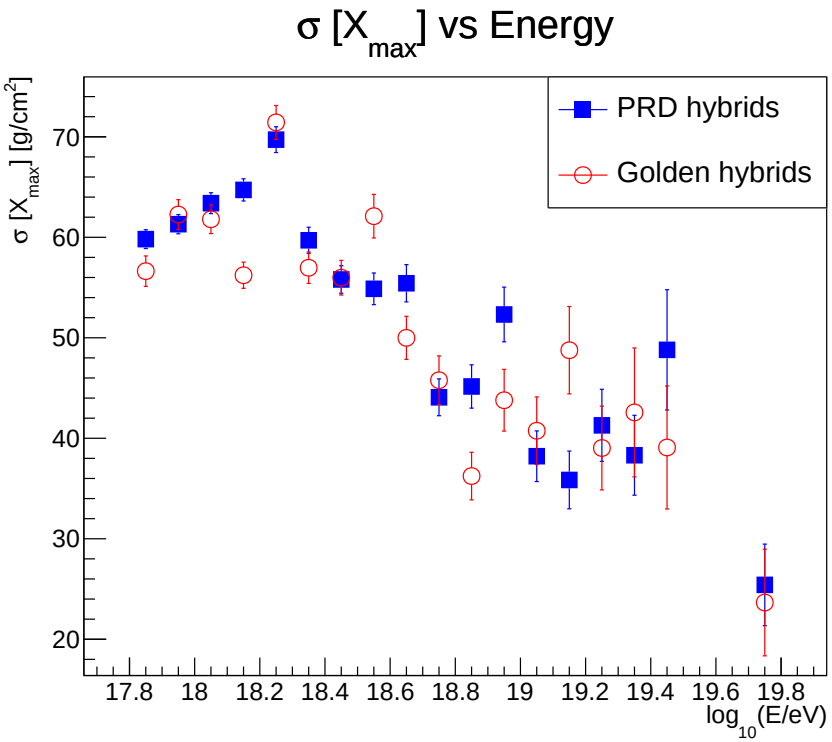
$X_{\mathrm{max}}^{\mathrm{Univ}} \rightarrow X_{\mathrm{max}}^{\mathrm{Univ}} + 5$ g/cm^2





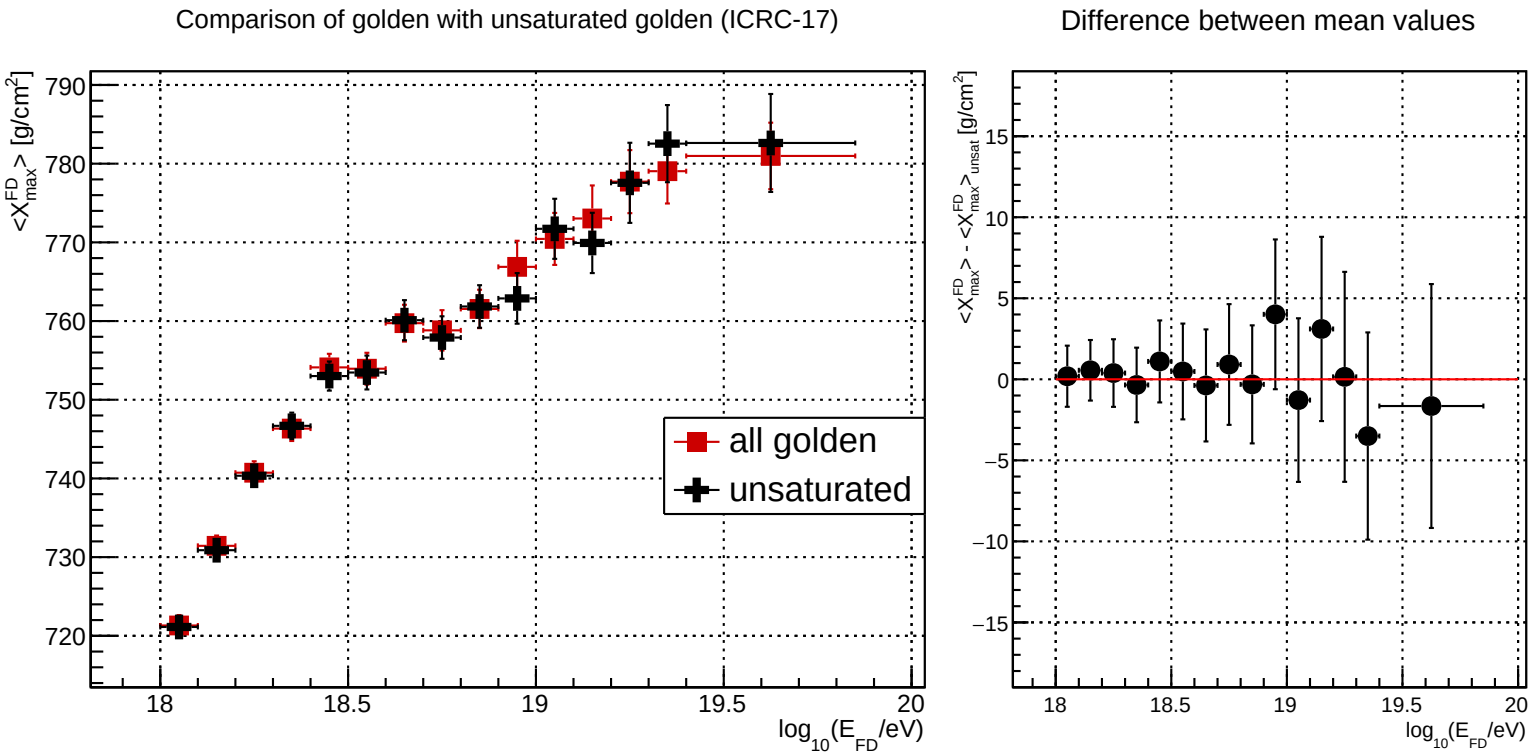
Golden VS hybrid





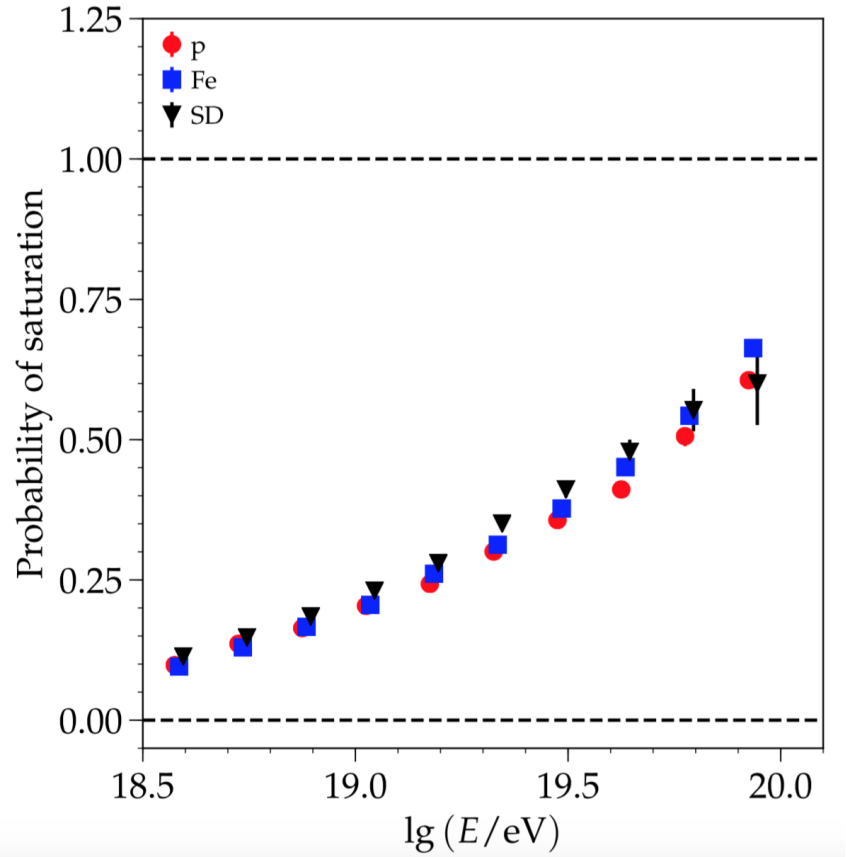
Comparison between the $X_{\rm max}^{\rm FD}$ moments for the selected set of ICRC-2015 golden events and for the set of hybrid events employed in *Phys. Rev. D 90 (2014) 122005*.

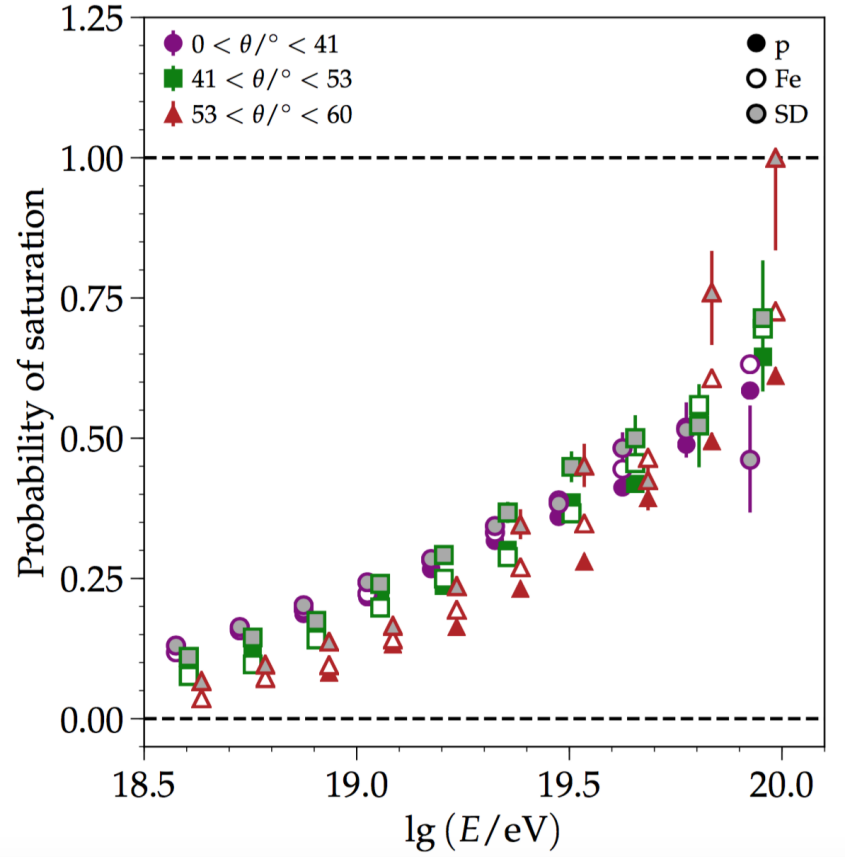
Golden ICRC-2017 - all events VS unsaturated



The differences in $\langle X_{\max}^{\mathrm{FD}} \rangle$ for the entire set of ICRC-2017 golden hybrids and for the unsaturated sub-set are compatible with zero at all energies. The exclusion of saturated events from the dataset from Universality studies should not introduce any bias.

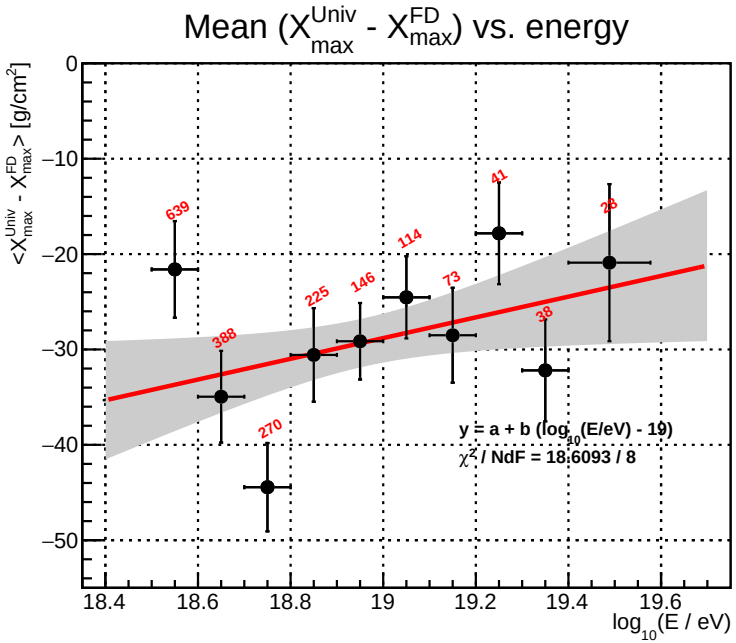
Probability of saturation

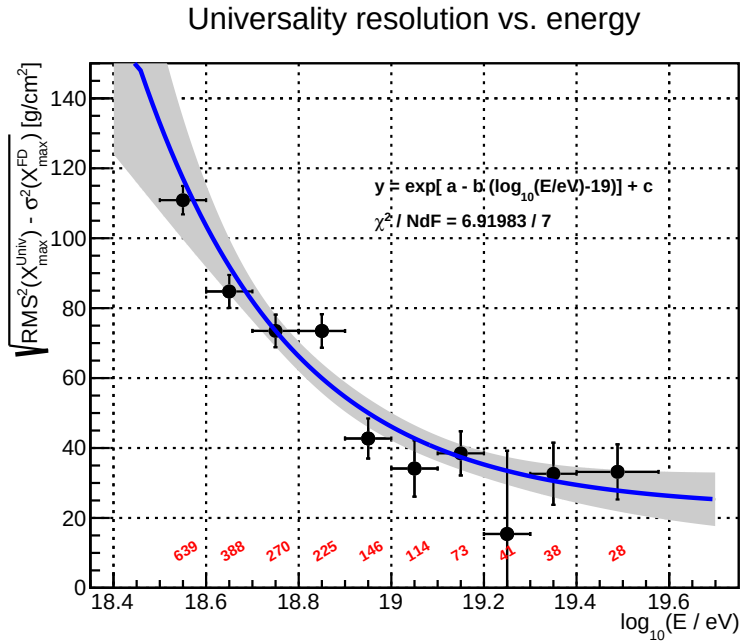




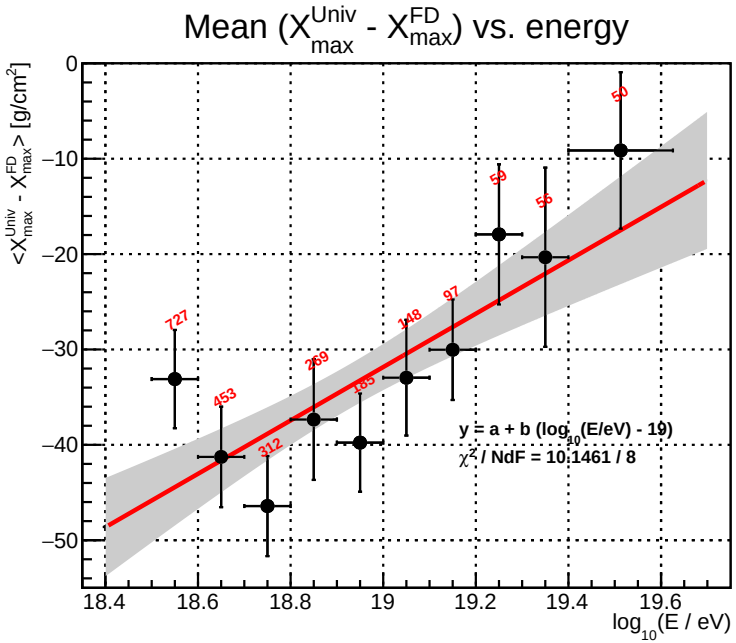
Probability of saturation for air showers simulated with QGS-Jet II.04 proton and iron primaries as well as for SD data.
Taken from Ariel Bridgeman PhD thesis (GAP 2018-039).

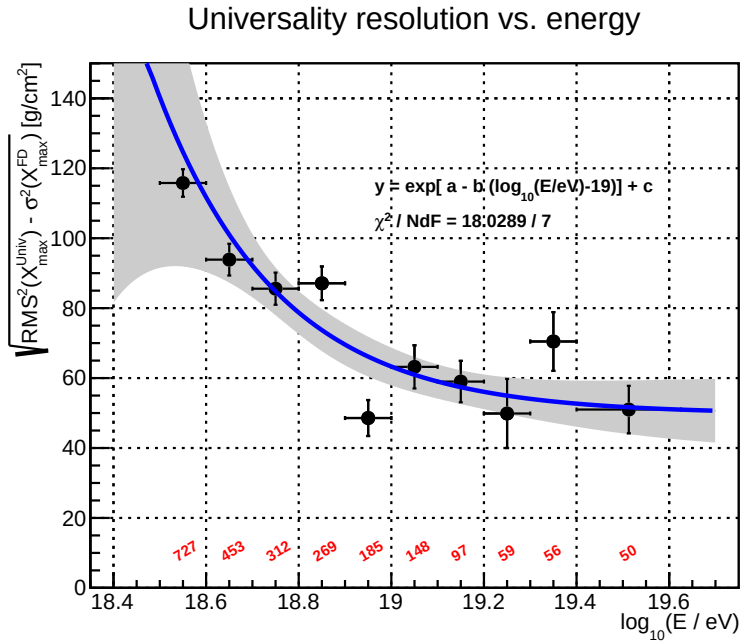
Bias and resolution - unsaturated (no selection cuts)



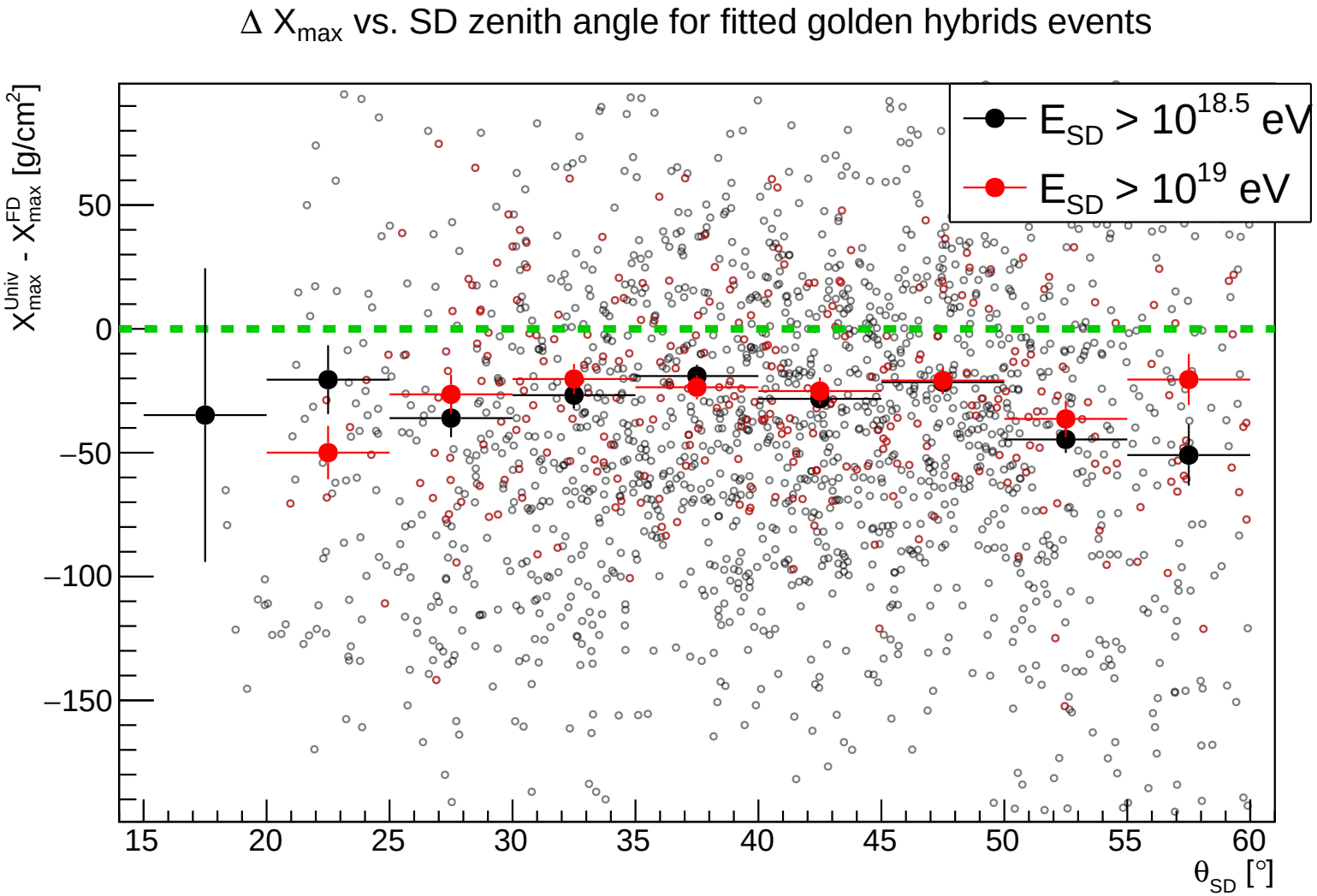


Bias and resolution - all events (no selection cuts)

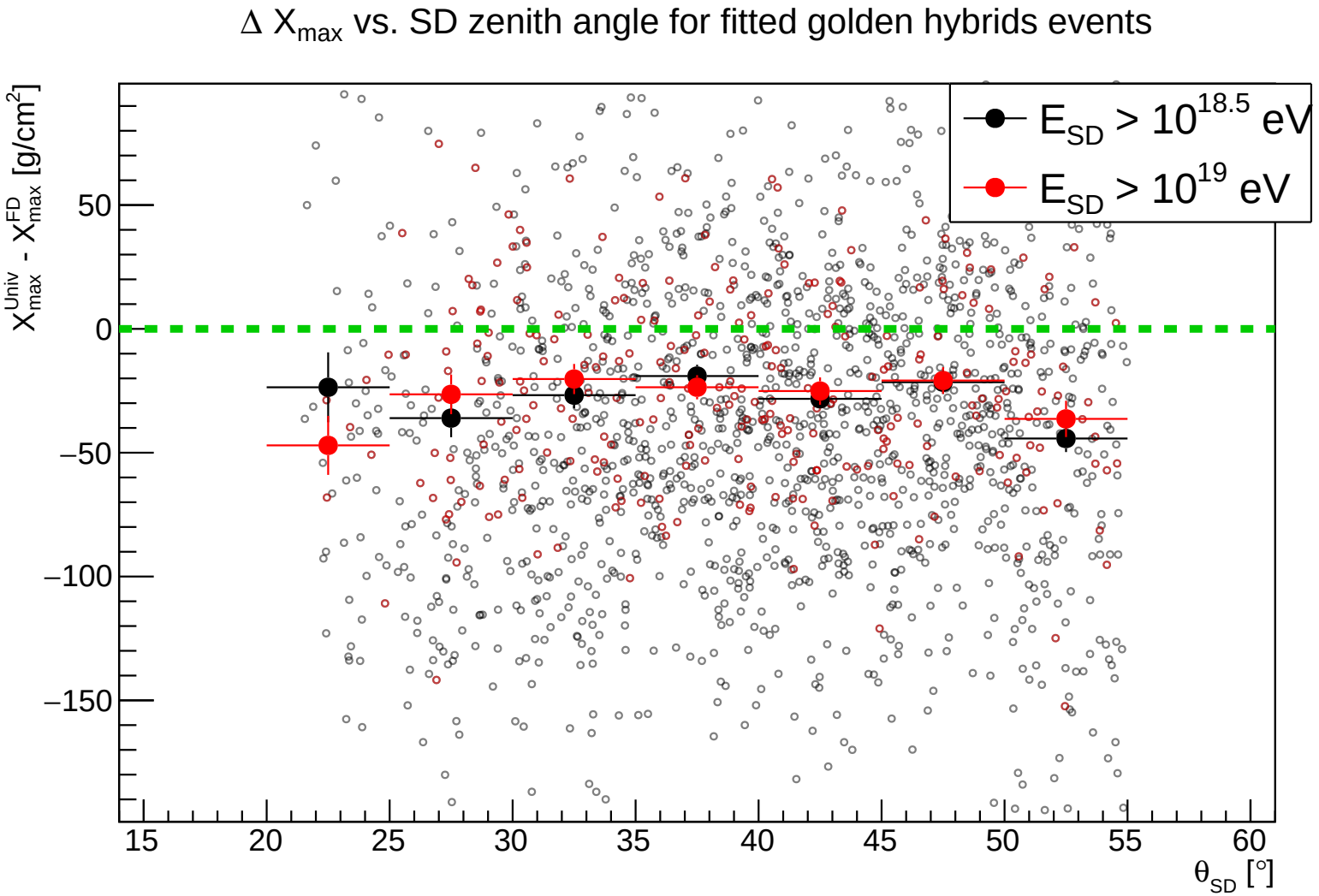




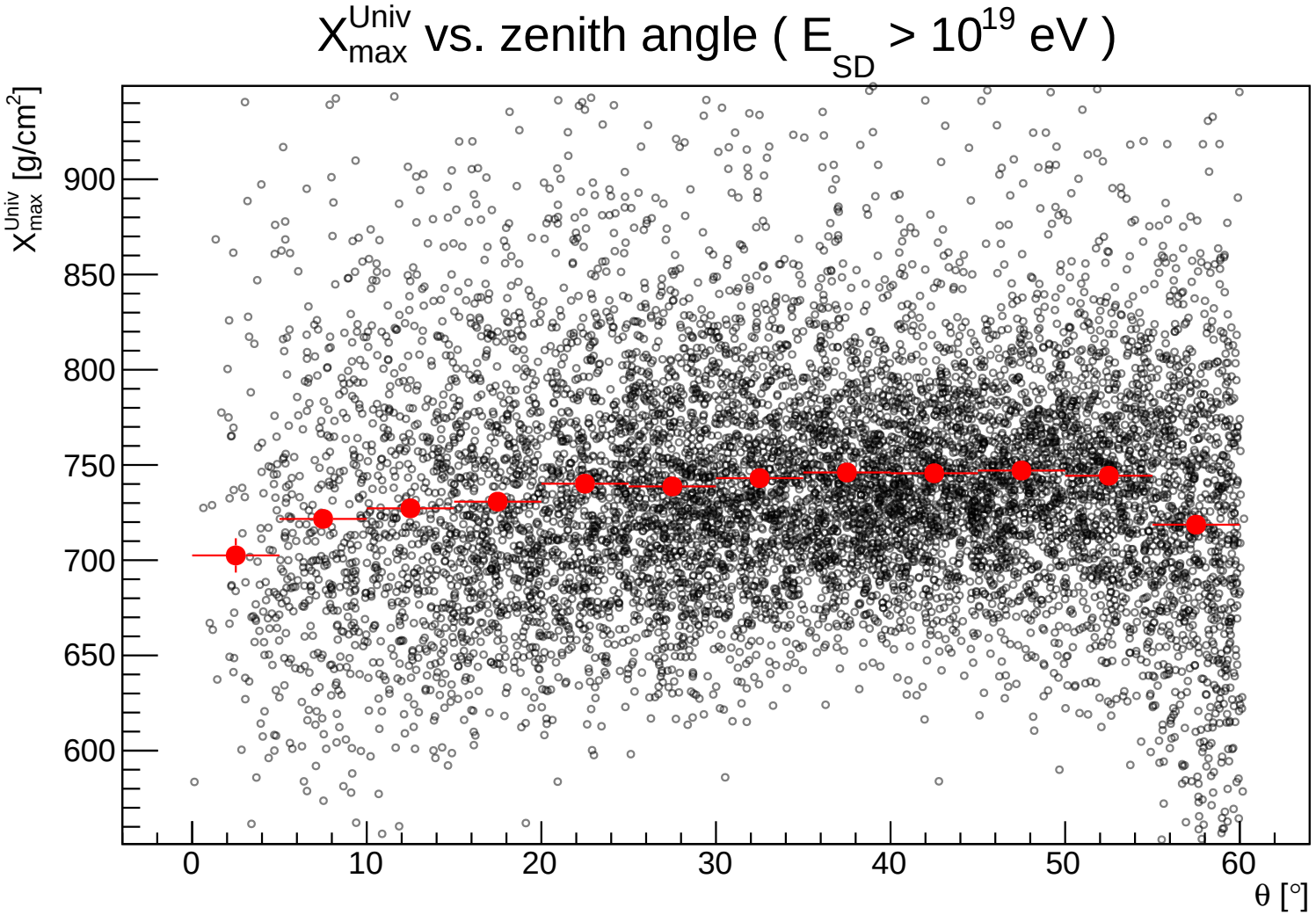
X_{max} bias before FoV cut



X_{max} bias after FoV cut



SD-1500 $\left\langle X_{\mathrm{max}}^{\mathrm{Univ}} \right\rangle$ before FoV cut



SD-1500 $\left\langle X_{\mathrm{max}}^{\mathrm{Univ}} \right\rangle$ after FoV cut