The Muon Puzzle

Analyzed with AGASA and AMIGA data

Flavia Gesualdi
Ralph Engel   Alberto Etchegoyen
Nicolás González   Quentin Luce
Markus Roth   David Schmidt
Daniel Supanitsky   Darko Veberič

UNSAM UNIVERSIDAD NACIONAL DE SAN MARTÍN

HIRSAP Meeting
25-27 November 2020

Double Doctoral degree in Astrophysics DDAp

Karlsruher Institut für Technologie

AGASA muon deficit in simulations

Muon deficit in air shower simulations estimated from AGASA muon measurements

F. Gesualdi, A. D. Supanitsky, and A. Etchegoyen, Phys. Rev. D 101, 083025 – Published 22 April 2020

Observable:

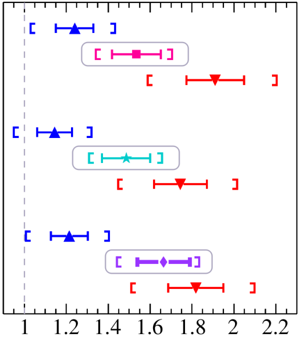
$$\widehat{F} = \frac{\langle \rho_\mu^{\text{data}} / E_R \rangle}{\langle \rho_\mu^{\{\text{p,Fe,mix}\}} / E_R \rangle}$$

With:

- $\log_{10}(E_R/\text{eV}) \in (18.83, 19.46)$
- $\theta \leq 36^\circ$
- $E_{\text{th}}^{\mu} = 0.5\,\text{GeV}$

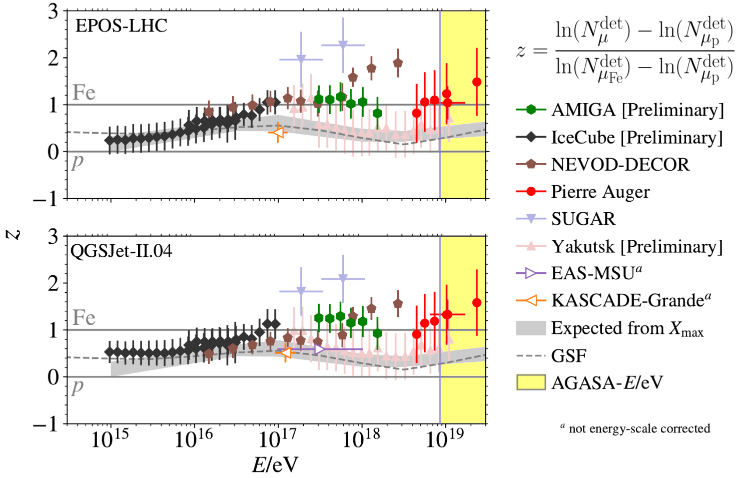
$F \pm (\text{stat}) \pm (\text{syst})$ [68% c.l.]

Sibyll2.3c	Fe	$1.24 \pm 0.09 \pm 0.20$
	Mixed	$\boxed{1.54 \pm 0.12 \pm 0.18}$
	p	$1.91 \pm 0.14 \pm 0.31$
EPOS-LHC	Fe	$1.14 \pm 0.08 \pm 0.19$
	Mixed	$\boxed{1.49 \pm 0.11 \pm 0.17}$
	p	$1.74 \pm 0.13 \pm 0.28$
QGSJetII-04	Fe	$1.21 \pm 0.09 \pm 0.20$
	Mixed	$\boxed{1.66 \pm 0.13 \pm 0.20}$
	p	$1.82 \pm 0.13 \pm 0.29$



Motivation

Muon deficit - H. P. Dembinski *et al.* (2019)



Objective: Computation of the z-value

What can we use to compute z?

Data

$E_{\text{det}}, \ \rho^{\text{det}}_{\mu,\text{data}}$

No detector simulations

We account for energy rec. and binning effects with a convolution.

$\downarrow$

Simulations + mixed composition

$E_{\text{true}}, \ \rho^{\text{true}}_{\mu,\{\text{p,Fe, mix}\}} \ ^*$

Simulations + mixed composition

$E_{\text{det}}, \ \left\langle \rho^{\text{det}}_{\mu,\{\text{p,Fe,mix}\}} \right\rangle$

* A={p, He, N, Fe} Corsika: Fluka + {EPOS-LHC, QGSJetII-04, Sibyll2.3c}, ∼20 per primary per energy (∼30 for p), $\log_{10}(E/\text{eV}) = 18.0, 18.2, ..., 19.8$
Mixed compo.: using mass fractions from the Pierre Auger X_{max} fits

Which z?

$$z = \frac{\ln \left\langle \rho^{\text{det}}_{\mu,\text{data}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle}{\ln \left\langle \rho^{\text{det}}_{\mu,\text{Fe}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle} \tag{1}$$

✓ Easier to compute: experiments report $\ln \left\langle \rho^{\text{det}}_{\mu,\text{data}} \right\rangle$ (comparable exp. to exp.)

$$z^{\langle \ln A \rangle} = \frac{\left\langle \ln(\rho^{\text{det}}_{\mu,\text{data}}) \right\rangle - \left\langle \ln(\rho^{\text{det}}_{\mu,\text{p}}) \right\rangle}{\left\langle \ln(\rho^{\text{det}}_{\mu,\text{Fe}}) \right\rangle - \left\langle \ln(\rho^{\text{det}}_{\mu,\text{p}}) \right\rangle} \tag{2}$$

✓ Directly comparable to $\langle \ln A \rangle$: From the Heitler-Matthews model $z^{\langle \ln A \rangle} = \langle \ln A \rangle / \ln 56$.
$\rightarrow$ Physically relevant

× We can't compute it for AGASA $\rightarrow$ analytic approx.

Results

We compare Eq. (1):

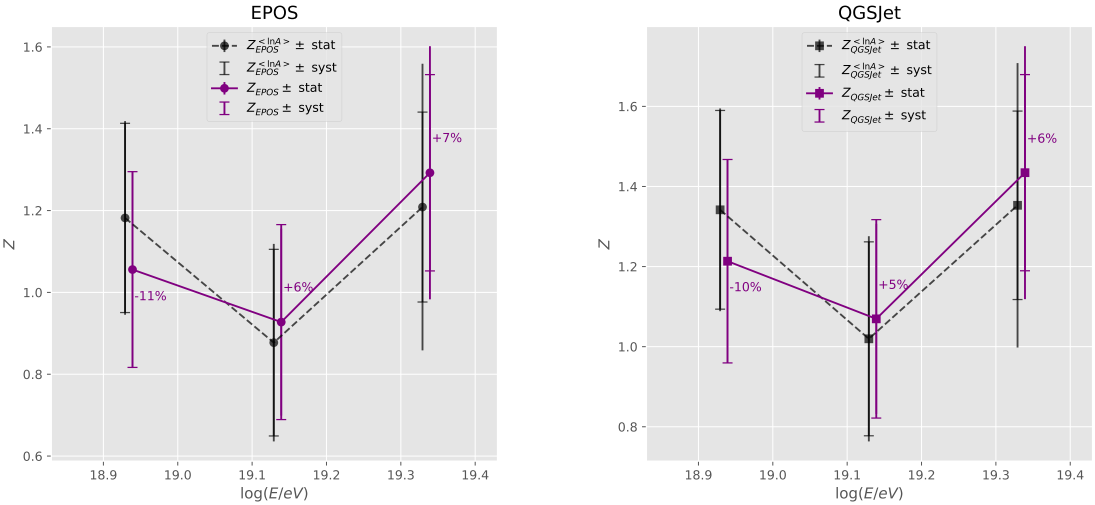
$$z = \frac{\ln \left\langle \rho^{\text{det}}_{\mu,\text{data}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle}{\ln \left\langle \rho^{\text{det}}_{\mu,\text{Fe}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle}$$

to an analytic approx. of Eq. (2):

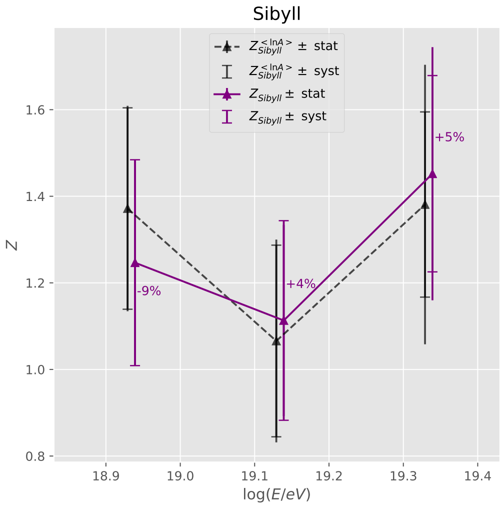
$$z^{\langle \ln A \rangle} \approx \frac{\left\langle \ln(\rho^{\text{det}}_{\mu,\text{data}}) \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle + \frac{1}{2} \overbrace{(\text{Var}[\epsilon^{\text{sh-sh}}_{p}] + \text{Var}[\epsilon^{\text{all but sh-sh}}])}^{\text{Var}[\epsilon_p]}}{\ln \left\langle \rho^{\text{det}}_{\mu,\text{Fe}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle + \frac{1}{2}(\text{Var}[\epsilon^{\text{sh-sh}}_{p}] - \text{Var}[\epsilon^{\text{sh-sh}}_{Fe}])},$$

where $\epsilon = \frac{\rho^{\text{det}}_{\mu} - \langle \rho^{\text{det}}_{\mu} \rangle}{\langle \rho^{\text{det}}_{\mu} \rangle}$ is the **event-wise relative error**, and its variance $\text{Var}[\epsilon] = \langle \epsilon^2 \rangle = \left(\frac{\sigma(\rho_\mu)}{\langle \rho_\mu \rangle} \right)^2$ is the **relative variance** of ρ_μ.

Results


EPOS
$Z_{EPOS}^{<\ln A>}$ ± stat
$Z_{EPOS}^{<\ln A>}$ ± syst
Z_{EPOS} ± stat
Z_{EPOS} ± syst
-11%
+6%
+7%
Z
log(E/eV)
QGSJet
$Z_{QGSJet}^{<\ln A>}$ ± stat
$Z_{QGSJet}^{<\ln A>}$ ± syst
Z_{QGSJet} ± stat
Z_{QGSJet} ± syst
-10%
+5%
+6%
Z
log(E/eV)

Results

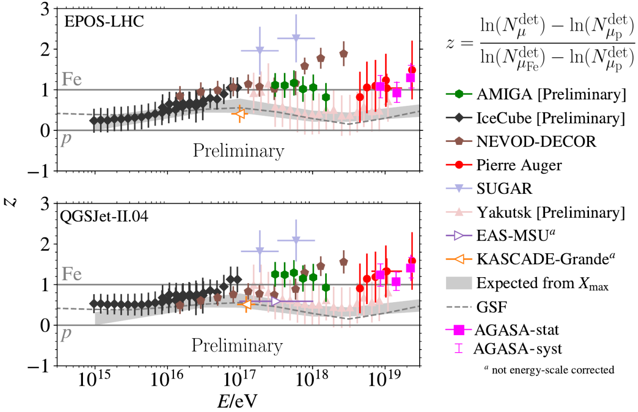


Sibyll
$Z_{Sibyll}^{<\ln A>}$ ± stat
$Z_{Sibyll}^{<\ln A>}$ ± syst
Z_{Sibyll} ± stat
Z_{Sibyll} ± syst
-9%
+4%
+5%
1.6
1.4
1.2
1.0
0.8
Z
18.9
19.0
19.1
19.2
19.3
19.4
log(E/eV)

AGASA: z-values for WHISP

- Two possible definitions:
  $z^{\langle \ln A \rangle}$ (avg. of the log)
  $\neq z$ (log of the avg.)
- We found an analytic approx. to $z^{\langle \ln A \rangle}$ (no det. resolution)
- We computed $z^{\langle \ln A \rangle}$ and z and studied the differences

The z-values calculated from AGASA data are in agreement with those of Pierre Auger and Yakutsk. Evidence of a **muon deficit** in air shower simulations.



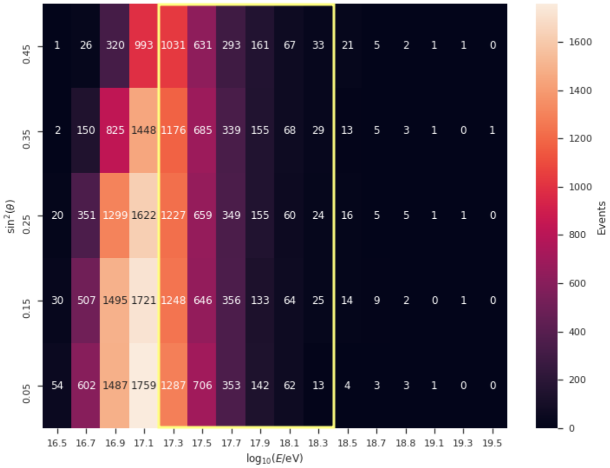
AMIGA Muon LDF

Objective

Study data, models, systematics, (and compare with simulations)

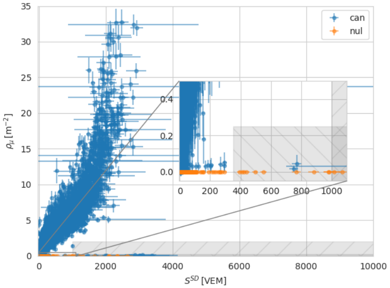
Data selection

1. AMIGA SiPM data 2018+2019
2. Request T4 efficiency > 99 % [1]
3. Request “enough” stats.
4. Request single-station lateral trigger probability > 90 % $\rightarrow r_{\max}(E_{\text{rec}}, \theta_{\text{rec}})$[2]



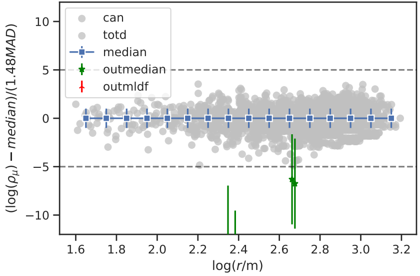
Data cleaning

- ✓ Bad periods: in agreement with previous analyses [3]
- ✓ Saturated counters ok
- ✓ ToTd/MoPS ok
- ✓ Null-signal counters: +17 outliers



- ✓ Other: +126 outliers, filtered using

$$\text{Modif. z-score} = \frac{\log(\rho_\mu) - \text{median}(\log(\rho_\mu))}{1.48\,\text{MAD}(\log(\rho_\mu))} \in [-5, 5]$$



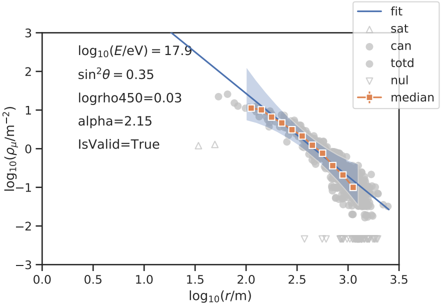
Data fit method

Method

For every $(\log(E), \sin^2\theta)$ bin, we fit the median $\log(\rho_\mu)$ as a function of the $\log(r)$ bins, via a weighted least-squares.

We achieve convergence by:

- using a clean data set
- fitting the median
- expressing the models as linear in the parameters (where possible)
- setting accurate initial values (E, θ) (from a power-law fit)



Example of a fit using a power-law

$$\log_{10}\rho_\mu = \log_{10}\rho_{450} - \alpha\log_{10} r$$

Data fit using modified-NKG

Using a Kascade-Grande/Modified-NKG:

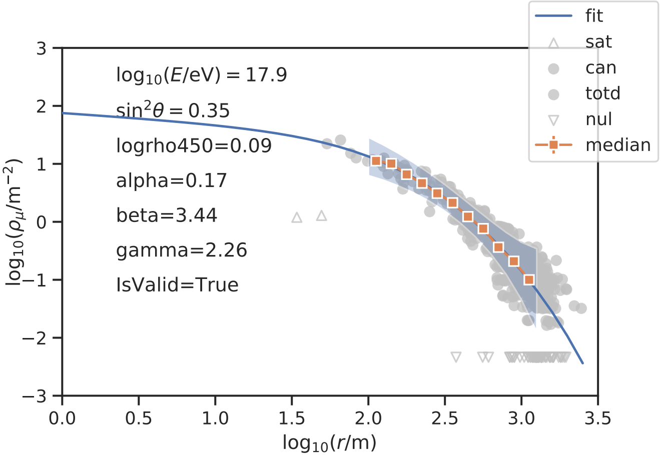
$$\rho_\mu(r) = \rho_{450}\ \frac{f\,(r)}{f\,(450)}$$

$$f\,(r) = \left(\frac{r}{r_0}\right)^{-\alpha}\left(1+\frac{r}{r_0}\right)^{-\beta}\left(1+\left(\frac{r}{10r_0}\right)^2\right)^{-\gamma}$$

The modified-NKG fits data well and **is physically motivated**.

Problems:

- Fitting both α and β is not possible: corr≥ 0.98 + low stats. close to core $\rightarrow$ fit depends on initial val.
- Not enough sensitivity to fit γ.



Toy model

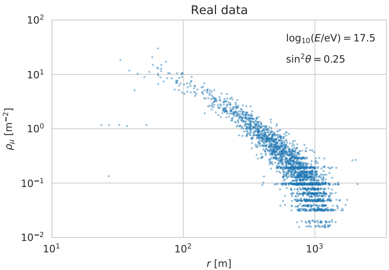
Quantify the systematics introduced by

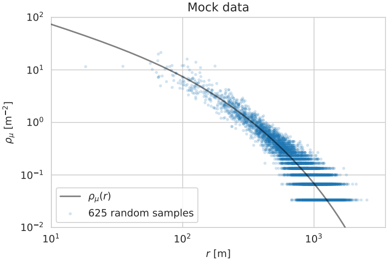
- **fixing a parameter**,
- cutting in distance,
- binning in $\log(r)$,
- fitting the median instead of all points,

in a controlled experiment $\rightarrow$ Toy model

Which effects are taken into account?

- ✓ Lateral trigger probability [2]
- ✓ Poisson fluctuations
- ✓ Fluctuations in the rec. core [4] [5]
- ✓ Saturation

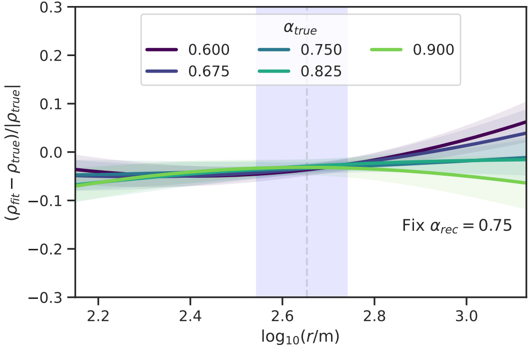


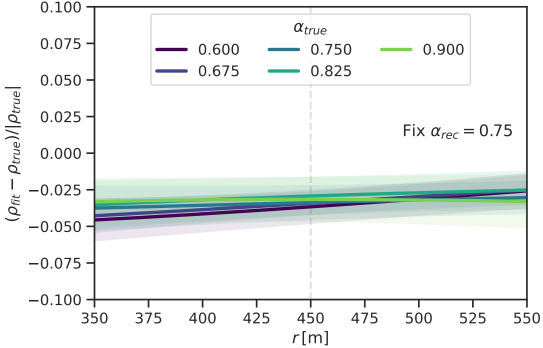


Toy model: Bias from fixing $\alpha_{\text{rec}} = 0.75$

How much bias does fixing α_{rec} to a wrong value introduce?

Idea: Sample from a muon LDF with α_{true}, reconstruct with $\alpha_{\text{rec}} = 0.75$





$20\,\%$ bias in $\alpha \to 10\,\%$ bias in ρ_μ

The systematics around $450\,\mathrm{m}$ are almost minimal

Summary and outlook

Summary:

- We reconstructed AMIGA data, performed quality checks, and obtained a clean data set.
- We compared **how different models fit data**. We found that the **modified NKG** fits well.
- We developed a **toy model** to study the systematics:
  - **450 m is still an optimal distance considering systematics** in the reconstruction

Next step:

- Compare against simulations: **study the muon deficit**

References I

[1] Parameterization of T4 efficiency extracted from GAP2018-045.

[2] Parameterization of lateral trigger probability extracted from GAP2013-114.

[3] GAP2020-021 (N. González *et. al*)

[4] Propagation of the core resolution into the LDF, private comm., Tobias Schultz.

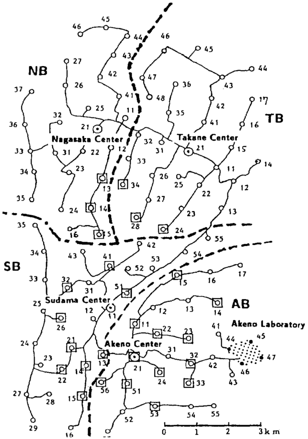
[5] Parameterization of the infill resolution, private comm., Quentin Luce.

Backup slides

AGASA: Akeno Giant Air Shower Array

N. Hayashida *et al.* (1995)

- Akeno, Japan (667 m a.s.l.)
- Hybrid, with **direct μ detection**:
  - 111 plastic scintillation counters spread across $\sim 100\,\mathrm{km}^2$ (~ 1 km separation)
  - 27 proportional counters, $\sim 30\,\mathrm{km}^2$ coverage, shielded with 30 cm of iron or 1 m of concrete ($E^{\mu}_{\mathrm{th.}} = 0.5\,\mathrm{GeV}$)
- $E > 3 \times 10^{16}\,\mathrm{eV}$, $\theta \leq 45°$
- Decomissioned in 2004



S. Yoshida *et al.* (1995)

$\ln\langle\rho_\mu\rangle \neq \langle\ln(\rho_\mu)\rangle$

Two experiments with different detector resolutions would report the same $\ln\left\langle\rho_{\mu,\text{data}}^{\text{det}}\right\rangle$ but different $\left\langle\ln(\rho_{\mu,\text{data}}^{\text{det}})\right\rangle$ [H. Dembinski (2018)]:

$$\left\langle\ln(\rho_\mu^{\text{det}})\right\rangle = \ln\left\langle\rho_\mu^{\text{det}}\right\rangle - \frac{1}{2}\text{Var}[\epsilon] + \mathcal{O}\left(\langle\epsilon^3\rangle\right), \tag{3}$$

where $\epsilon = \frac{\rho_\mu^{\text{det}} - \langle\rho_\mu^{\text{det}}\rangle}{\langle\rho_\mu^{\text{det}}\rangle}$ is the **event-wise relative error**, and its variance $\text{Var}[\epsilon] = \langle\epsilon^2\rangle = \left(\frac{\sigma(\rho_\mu)}{\langle\rho_\mu\rangle}\right)^2$ is the **relative variance** of ρ_μ.

Approximating $z^{\langle \ln A \rangle}$

- We **don't** know $\left\langle \ln(\rho_{\mu,\{\mathrm{p,Fe}\}}^{\mathrm{det}}) \right\rangle$.
- We can't numerically compute it without knowing the detector res.
- But we **do** know $\ln \left\langle \rho_{\mu,\{\mathrm{p,Fe}\}}^{\mathrm{det}} \right\rangle$.

Then we use (3) to approximate $z^{\langle \ln A \rangle}$ (Eq. (1)):

$$z^{\langle \ln A \rangle} \approx \frac{\left\langle \ln(\rho_{\mu,\mathrm{data}}^{\mathrm{det}}) \right\rangle - \ln \left\langle \rho_{\mu,\mathrm{p}}^{\mathrm{det}} \right\rangle + \frac{1}{2}\mathrm{Var}[\epsilon_p]}{\ln \left\langle \rho_{\mu,\mathrm{Fe}}^{\mathrm{det}} \right\rangle - \ln \left\langle \rho_{\mu,\mathrm{p}}^{\mathrm{det}} \right\rangle + \frac{1}{2}(\mathrm{Var}[\epsilon_p] - \mathrm{Var}[\epsilon_{Fe}])}. \tag{4}$$

What is Var$[\epsilon]$?

Var$[\epsilon]$ is the scatter of the detected muon densities in one energy bin.

Sources

$$\left.\begin{array}{c} \mathrm{Var}[\epsilon^{\text{sh-sh}}] \\ + \\ \mathrm{Var}[\epsilon^{\text{all but sh-sh}}] \\ = \\ \mathrm{Var}[\epsilon] \end{array}\right.$$

- shower-to-shower fluctuations

- detector resolution
- reconstructed energy resolution
- different energies contributing to one same bin

Computation of $z^{\langle \ln A \rangle}$

- $\mathrm{Var}[\epsilon^{\text{sh-sh}}]$ depends on the primary
- $\mathrm{Var}[\epsilon^{\text{all but sh-sh}}]$ doesn't depend on the primary, and can be calculated from data:

$$\mathrm{Var}[\epsilon^{\text{all but sh-sh}}] = \mathrm{Var}[\epsilon_{\text{data}}] - \mathrm{Var}[\epsilon^{\text{sh-sh}}_{\text{mix}}]. \tag{5}$$

Finally:

$$z^{\langle \ln A \rangle} \approx \frac{\left\langle \ln(\rho^{\text{det}}_{\mu,\text{data}}) \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle + \frac{1}{2} \overbrace{\left(\mathrm{Var}[\epsilon^{\text{sh-sh}}_{p}] + \mathrm{Var}[\epsilon^{\text{all but sh-sh}}] \right)}^{\mathrm{Var}[\epsilon_p]}}{\ln \left\langle \rho^{\text{det}}_{\mu,\text{Fe}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle + \frac{1}{2} \left(\mathrm{Var}[\epsilon^{\text{sh-sh}}_{p}] - \mathrm{Var}[\epsilon^{\text{sh-sh}}_{Fe}] \right)}. \tag{6}$$

Computation of Δz

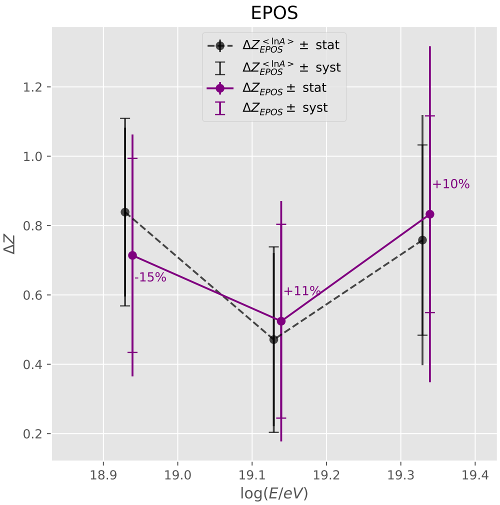
Δz is the difference between z from data and z from the mix. compo.

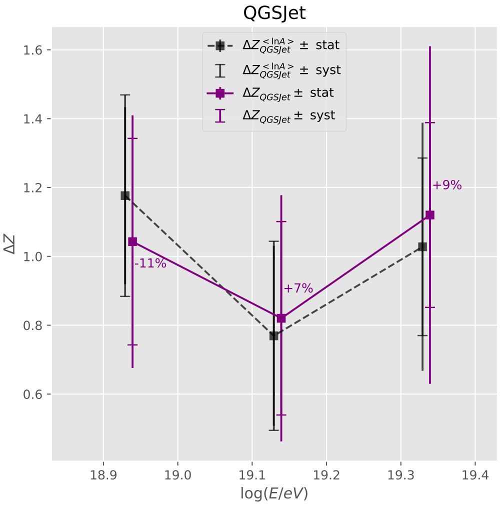
$\Delta z = z - z_{\text{mass}}$
$\Delta z^{\langle \ln A \rangle} = z^{\langle \ln A \rangle} - z^{\langle \ln A \rangle}_{\text{mass}}$, where:

$$z^{\langle \ln A \rangle}_{\text{mass}} = \frac{\left\langle \ln(\rho^{\text{det}}_{\mu,\text{mix}}) \right\rangle - \left\langle \ln(\rho^{\text{det}}_{\mu,\text{p}}) \right\rangle}{\left\langle \ln(\rho^{\text{det}}_{\mu,\text{Fe}}) \right\rangle - \left\langle \ln(\rho^{\text{det}}_{\mu,\text{p}}) \right\rangle} \tag{7}$$

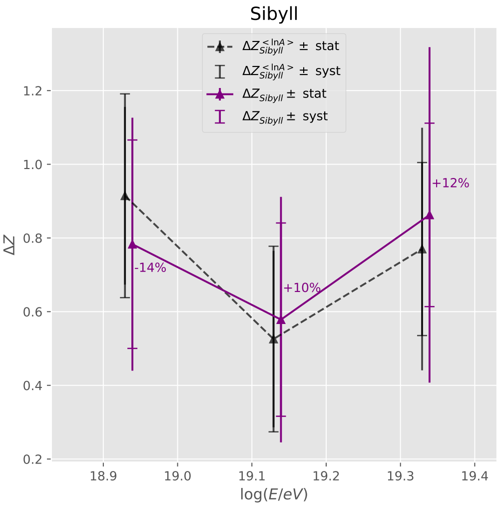
$$\approx \frac{\ln \left\langle \rho^{\text{det}}_{\mu,\text{mix}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle + \frac{1}{2}(\text{Var}[\epsilon^{\text{sh-sh}}_{p}] - \text{Var}[\epsilon^{\text{sh-sh}}_{\text{mix}}])}{\ln \left\langle \rho^{\text{det}}_{\mu,\text{Fe}} \right\rangle - \ln \left\langle \rho^{\text{det}}_{\mu,\text{p}} \right\rangle + \frac{1}{2}(\text{Var}[\epsilon^{\text{sh-sh}}_{p}] - \text{Var}[\epsilon^{\text{sh-sh}}_{Fe}])}. \tag{8}$$

Results


EPOS
$\Delta Z_{EPOS}^{<\ln A>}$ ± stat
$\Delta Z_{EPOS}^{<\ln A>}$ ± syst
ΔZ_{EPOS} ± stat
ΔZ_{EPOS} ± syst
-15%
+11%
+10%
ΔZ
log(E/eV)
1.2
1.0
0.8
0.6
0.4
0.2
18.9
19.0
19.1
19.2
19.3
19.4


QGSJet
$\Delta Z_{QGSJet}^{<\ln A>}$ ± stat
$\Delta Z_{QGSJet}^{<\ln A>}$ ± syst
ΔZ_{QGSJet} ± stat
ΔZ_{QGSJet} ± syst
-11%
+7%
+9%
ΔZ
log(E/eV)
1.6
1.4
1.2
1.0
0.8
0.6
18.9
19.0
19.1
19.2
19.3
19.4

Results



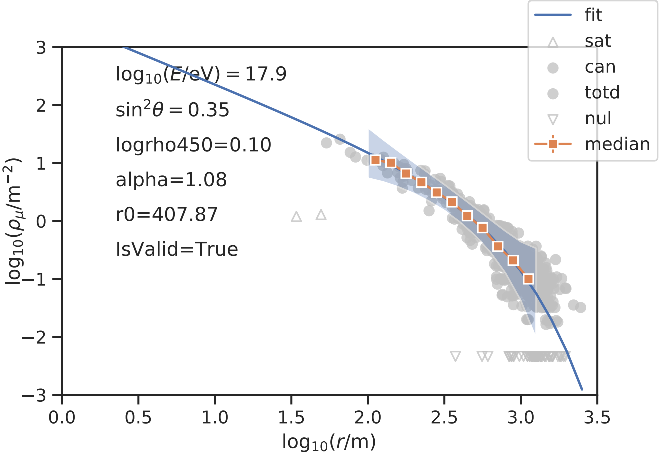
Sibyll
$\Delta Z^{<\ln A>}_{Sibyll}$ ± stat
$\Delta Z^{<\ln A>}_{Sibyll}$ ± syst
ΔZ_{Sibyll} ± stat
ΔZ_{Sibyll} ± syst
-14%
+10%
+12%
ΔZ
1.2
1.0
0.8
0.6
0.4
0.2
18.9
19.0
19.1
19.2
19.3
19.4
log(E/eV)

Data fit using exponential

Using a normalized exponential/Hillas:

$$\rho_\mu(r) = \rho_{r_{\text{ref}}} \left(\frac{r}{r_{\text{ref}}}\right)^{-\alpha} \exp\left(-\frac{r-r_{\text{ref}}}{r_0}\right)$$

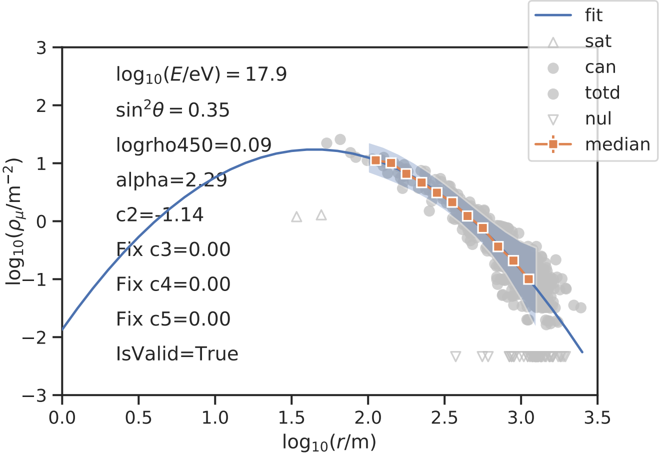
Fits with 3 free parameters are $\sim$ to fits with modif. NKG with 2 free parameters.



Data fit using loglog-polynomials

Using a log-log polynomial degree 2:
$\log_{10}\rho_\mu(\log_{10} r) = \log_{10}\rho_{r_{\mathrm{ref}}} + c_1 \log_{10}(r/r_{\mathrm{ref}}) + c_2 \log^2_{10}(r/r_{\mathrm{ref}})$

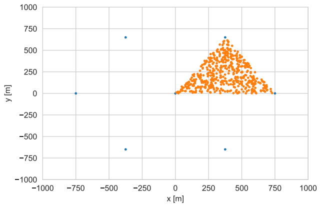
Log-log polynomials yield unphysical results.

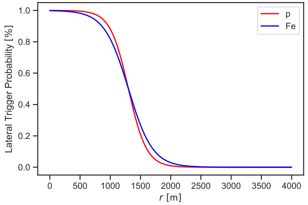


Toy model

We **generate mock-data** by:

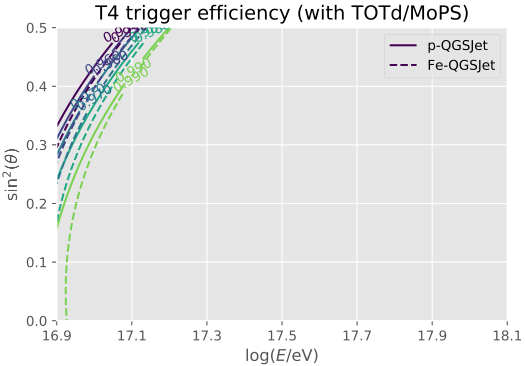
1. Creating an infill-like array
2. Defining a “true” underlying MLDF
3. Throwing it at the array
4. Sampling ρ_μ
5. Keeping the station with a probability following the LTP
6. Fluctuating ρ_μ (Poisson)
7. Fluctuating the reconstructed position of the core $\rightarrow$ the distance to the shower axis
8. Flagging saturation with a certain probability (computed from data)





Efficiency

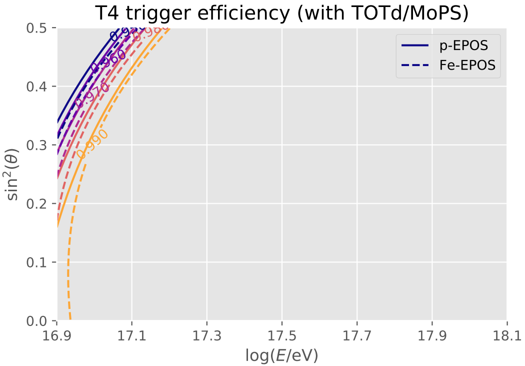
T4 efficiency



Parameterization from Alan Coleman’s PhD thesis (GAP2018 045)

Efficiency

T4 efficiency



Parameterization from Alan Coleman's PhD thesis (GAP2018 045)

Lateral Trigger Probability

Single stations: Distance to the core at which the trigger probability of triggering is 90 %.