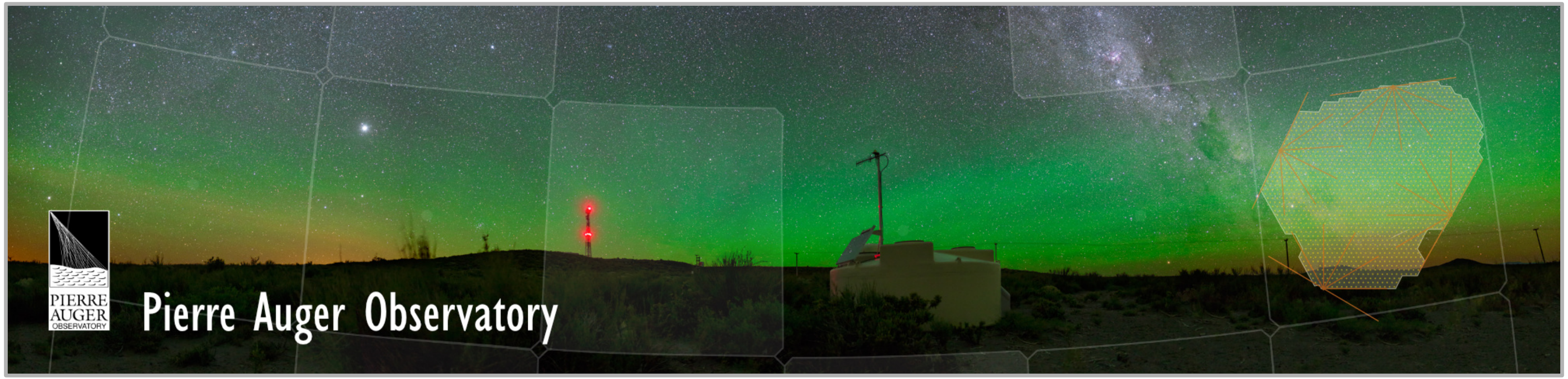
Lateral Distribution Function and Energy Spectrum for the 750 m Array of the Pierre Auger Observatory

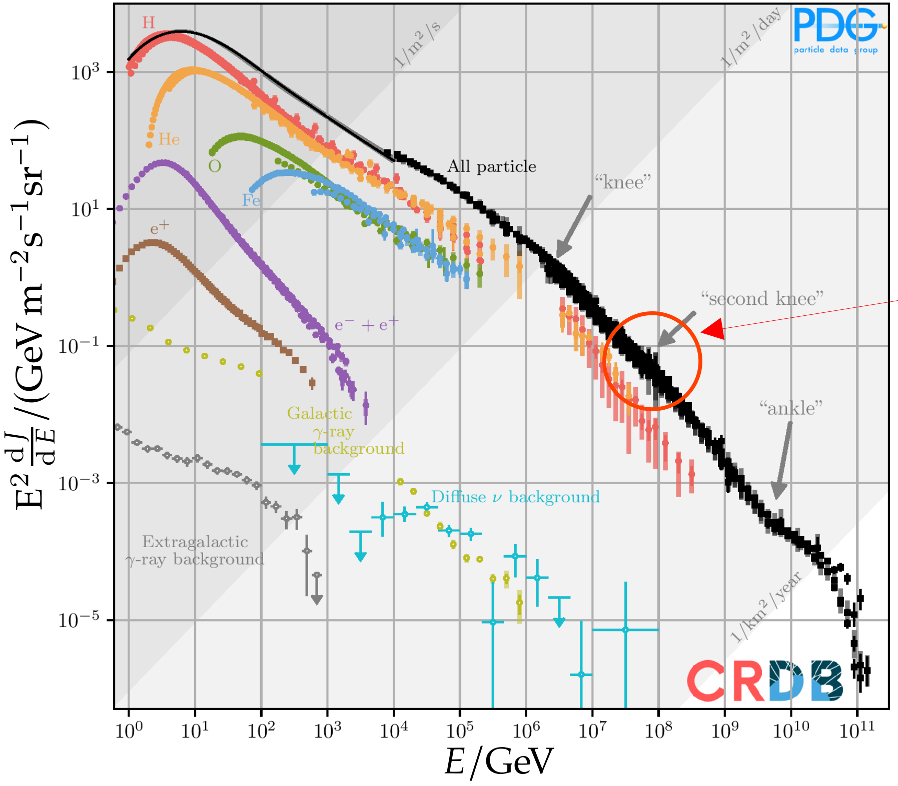
Philipp Meder

Master's thesis

IAP-HEU seminar, 7th November 2024



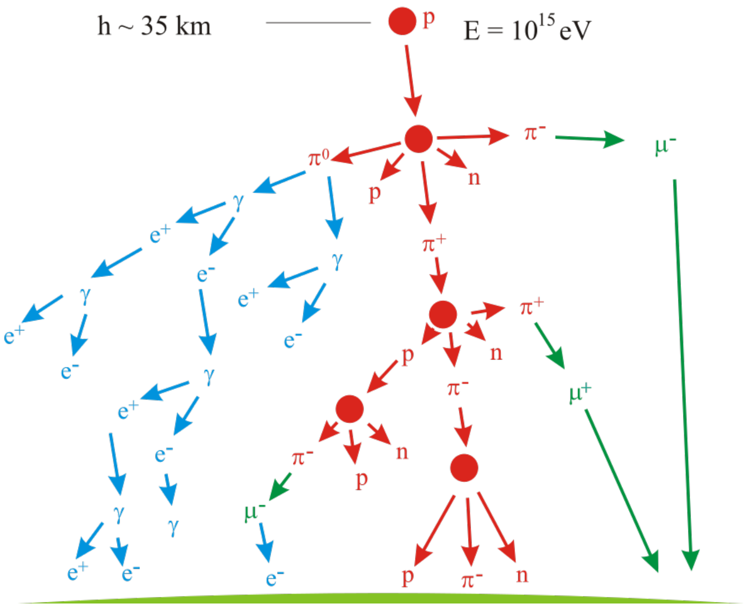
Motivation



- Additional low-signal triggers lower the over-all efficiency threshold
- Low-energy events can be reconstructed

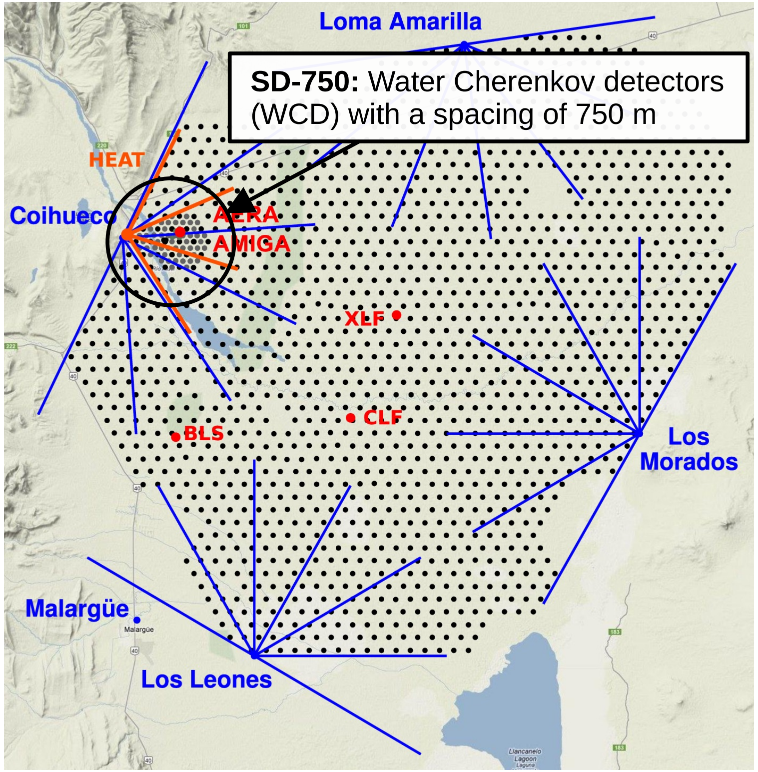
Goal: measure in the range of the second knee with the **SD-750**

Extensive air showers

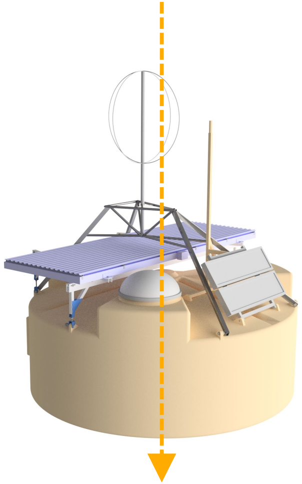


Taken from https://en.wikipedia.org/wiki/File:AirShower.svg

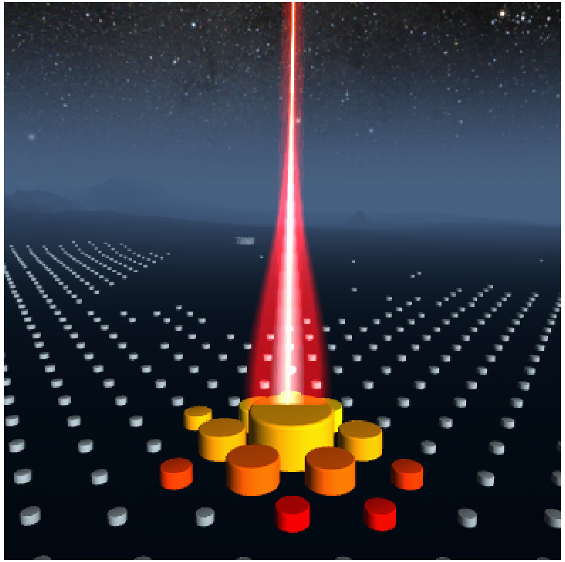
Surface Detector SD-750



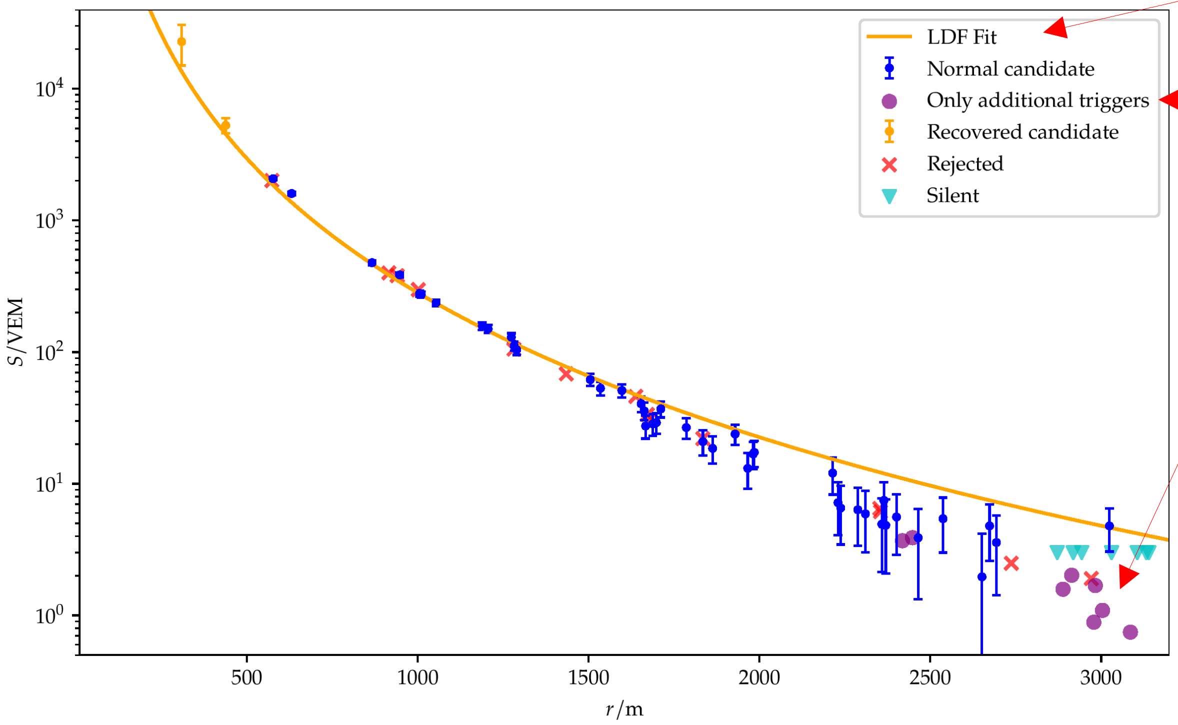
Vertical, central, through-going muon (VEM)



Shower footprint in the surface detector



Surface Detector: lateral shower profile



Lateral distribution function (LDF) fitted with maximum likelihood method

Additional triggers sensitive to low signals. Stations triggered only by those are **currently not used** in the LDF fit!

Parameters to fit:

Geometry (station timing):

θ (zenith angle)

LDF (station signals):

$$S_{450}, x_{\text{core}}, y_{\text{core}}, (\beta, \gamma)$$

signal at $r = 450\,\text{m}$

impact point

LDF slopes (parameterised)

Current likelihood

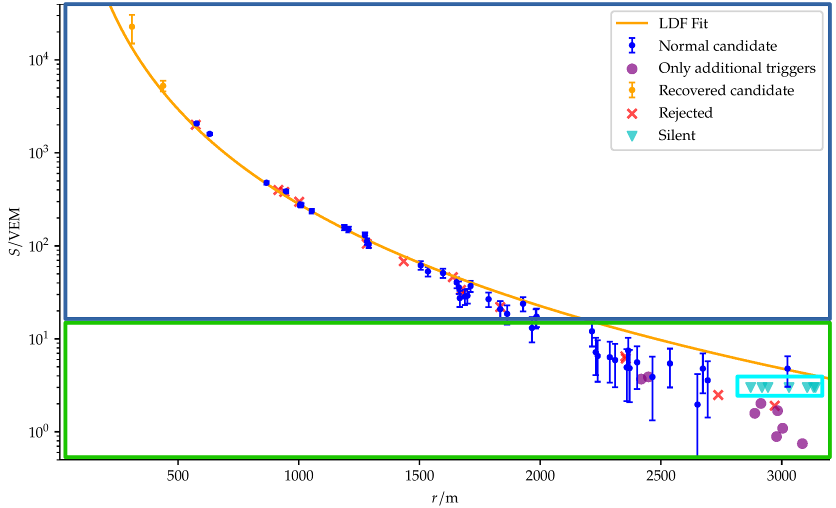
Karlsruhe Institute of Technology

- Signal to particle conversion using Poisson factor $p = \max(1, f_S^{-2}(\theta)) \quad \Rightarrow \quad n = p\,S$
- Different contributions

Poisson needs integers!

$$\mathcal{L} = \prod_i f_{\mathrm{Gauss}}(n_i, \mu_i) \prod_i f_{\mathrm{Poi}}(n_i, \mu_i) \prod_i f_{\mathrm{zero}}(n_i, \mu_i)$$

predicted particle number

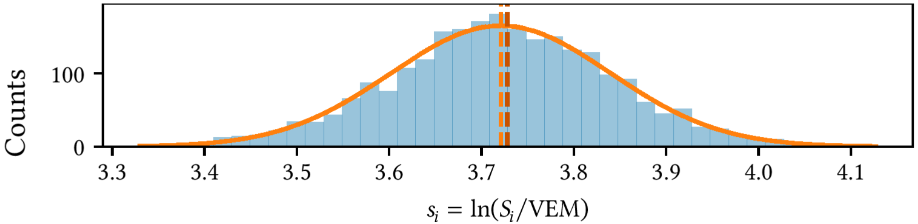


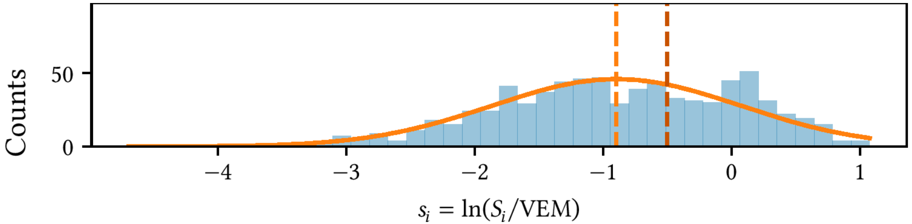
Goal:

- Unify distinct treatment of high and low signals
- Get rid off particle conversion
- Use correct trigger efficiency and all triggers

Simulations (artificially low trigger threshold)

Signals are log-normal distributed!





Vertical lines
Dark orange: $\bar{S}_i$
Light orange: $\bar{s}_i$

Envelopes
Continuous: $\sigma_x = \mathrm{std}\, x_i$
Dashed: $\sigma_x = \mathrm{model}(x_{\mathrm{pred}}, \theta)$

$$s = \ln(S/\mathrm{VEM}), \quad \mu = \ln(S_{\mathrm{pred}}/\mathrm{VEM}),$$

$$\sigma = \sigma(\mu, \theta) = \frac{\mathrm{d}\,\mu}{\mathrm{d}\,S_{\mathrm{pred}}}\,\tilde{\sigma}(S_{\mathrm{pred}}, \theta) = \frac{\tilde{\sigma}(S_{\mathrm{pred}}, \theta)}{S_{\mathrm{pred}}}$$

New log-likelihood with trigger probability:

$$\mathcal{L}_{\ln} = \sum_i^{\mathrm{trig}} [\ln p_{\mathrm{trig}}(\mu_i, \theta) + \ln \mathcal{N}(s_i, \mu_i, \sigma_i)] + \sum_i^{\neg\mathrm{trig}} \ln[1 - p_{\mathrm{trig}}(\mu_i, \theta)]$$

Triggered stations

Silent stations

More information in the thesis

Functional form of the LDF

Karlsruhe Institute of Technology

- Current LDF: **N**ishimura-**K**amata-**G**reisen

$$f_{\mathrm{NKG}} = \left(\frac{r}{r_{\mathrm{ref}}}\right)^{\beta_{\mathrm{NKG}}} \left(\frac{r + r_{\mathrm{scale}}}{r_{\mathrm{ref}} + r_{\mathrm{scale}}}\right)^{\beta_{\mathrm{NKG}}+\gamma_{\mathrm{NKG}}}$$

- New LDF: **E**xponentially **S**uppressed **P**ower **L**aw

$$f_{\mathrm{ESPL}} = \left(\frac{r}{r_{\mathrm{ref}}}\right)^{\beta_{\mathrm{ESPL}}} \exp\left(\frac{r_{\mathrm{ref}} - r}{r_{\mathrm{scale}}}\right)$$

$$\beta_{\mathrm{ESPL}} = -\exp\beta \quad \text{with} \quad \beta \in \mathbb{R}$$

Ensures physical LDFs (only falling) while leaving the fit parameter unbound

For SD-750 array:

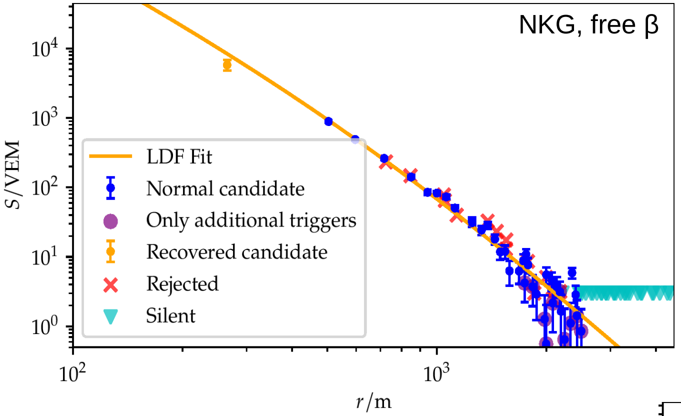
$$r_{\mathrm{ref}} = 450\,\mathrm{m}$$
$$r_{\mathrm{scale}} = 700\,\mathrm{m}$$

ESPL LDF: catch the correct LDF behaviour far from the core

First term: overall fall-off
Second term: suppression at large distances

Depends in principal on another slope parameter, for SD-750 this simplified form looks reasonable when checked on data

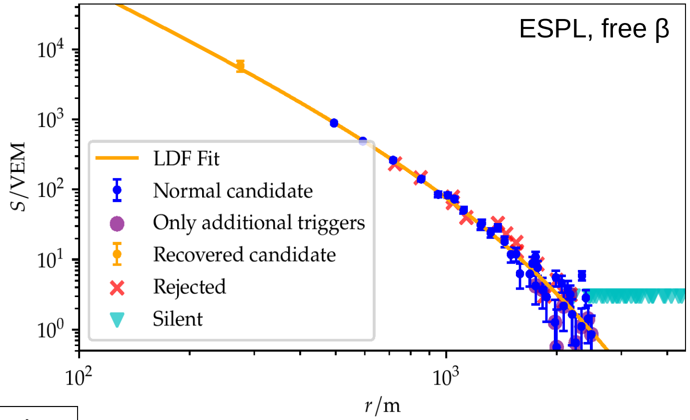
Fitting events with the new setup

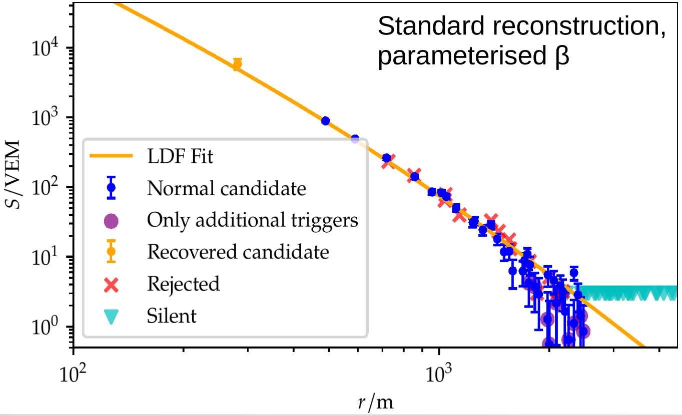


Event: 34352067

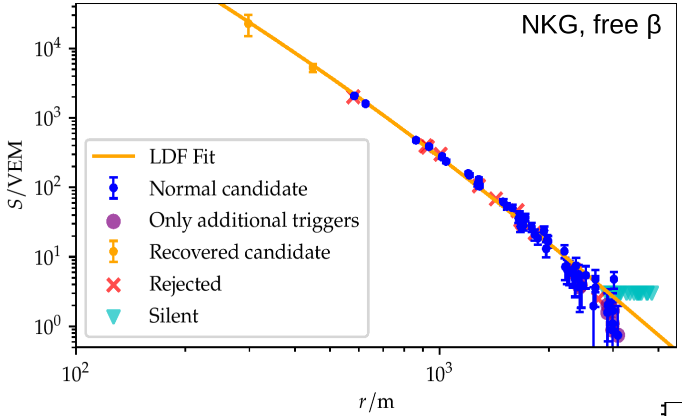
$$\ell_i = -2\,\mathcal{L}_{\ln,i}$$
$$\ell_{\mathrm{NKG}} = 41.41$$
$$\ell_{\mathrm{ESPL}} = 24.61$$

$$\Delta\ell = \ell_{\mathrm{ESPL}} - \ell_{\mathrm{NKG}} = -16.79$$





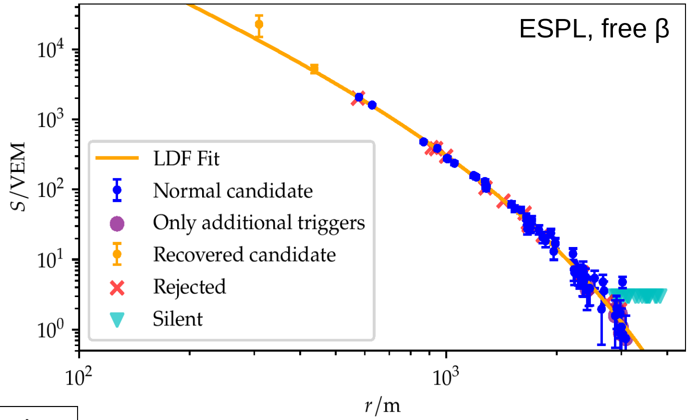
Fitting events with the new setup

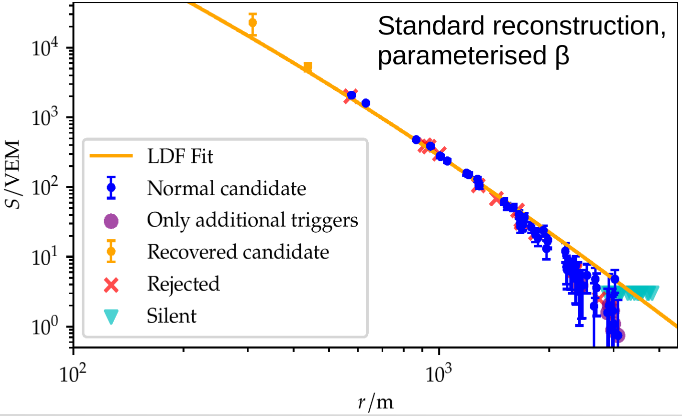


Event: 27057820

$$\ell_i = -2\,\mathcal{L}_{\ln,i}$$
$$\ell_{\mathrm{NKG}} = 25.69$$
$$\ell_{\mathrm{ESPL}} = -4.23$$

$$\Delta\ell = \ell_{\mathrm{ESPL}} - \ell_{\mathrm{NKG}} = -29.92$$





Parameterising the ESPL slope parameter

- Simple parameterisation:

$$x = \lg(S_{450}/\mathrm{VEM})$$

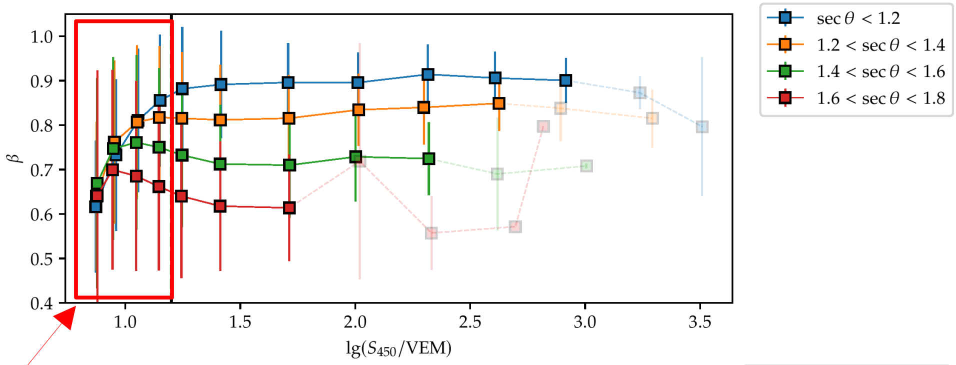
$$\beta = a + b\,x + c\,\sec\theta$$

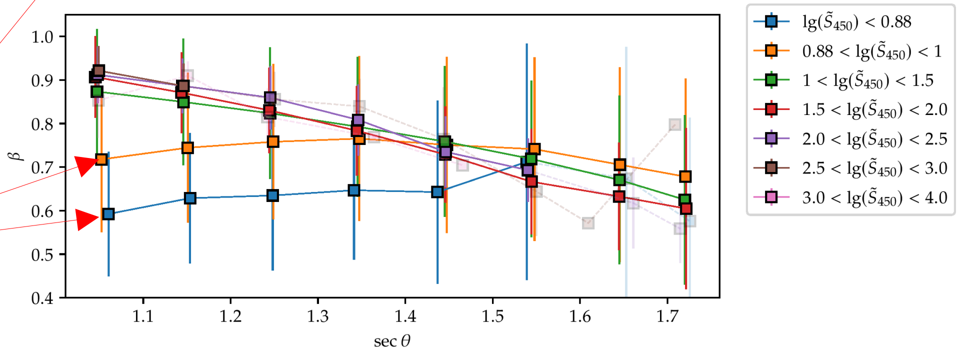
$$\sigma_\beta^2 = \exp\left(s_1 + s_2\,x + s_3\,\sec\theta\right)$$

Add term to likelihood to allow a variation around the parameterised mean

$\sim \mathcal{N}(\beta_{\mathrm{recon}}, \beta, \sigma_\beta)$

Physical or statistical effect?
Huge errorbars!
Extrapolate into this region!



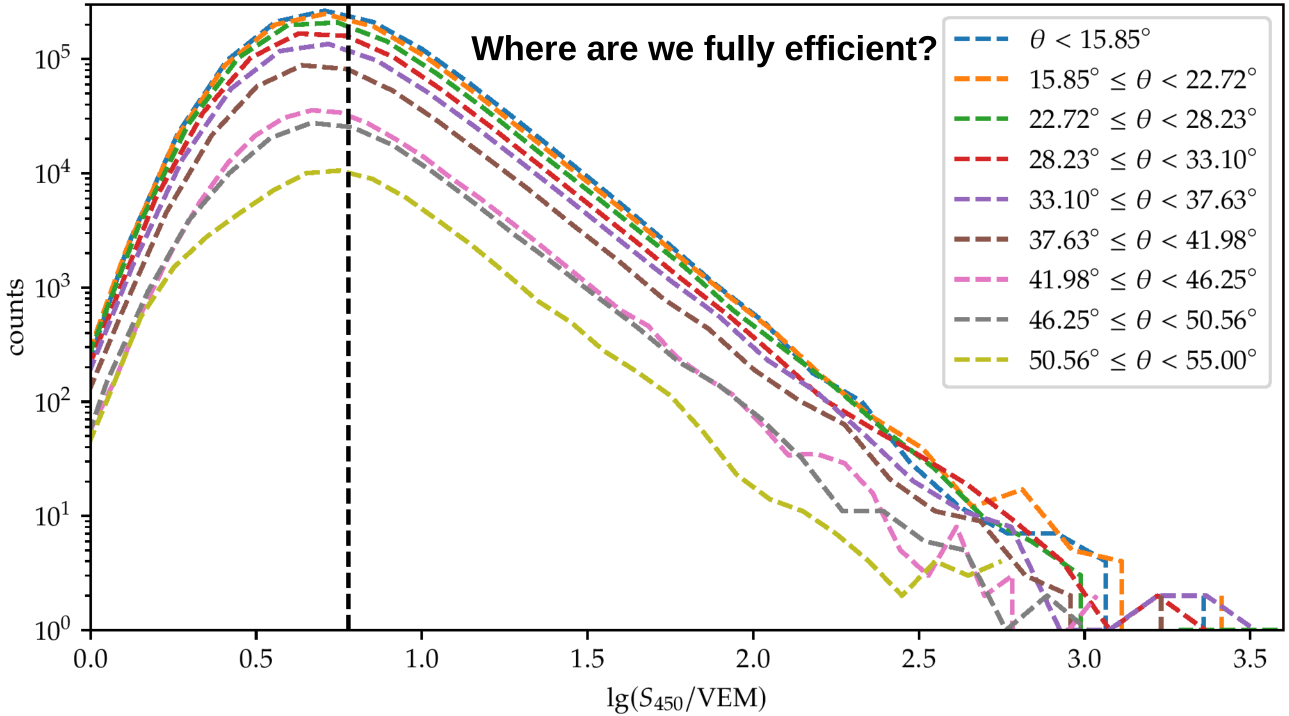


Transparent: bins with less than 30 events

Raw spectrum – two issues

Obviously there is a violation to isotropy due to the attenuation of the shower in the atmosphere

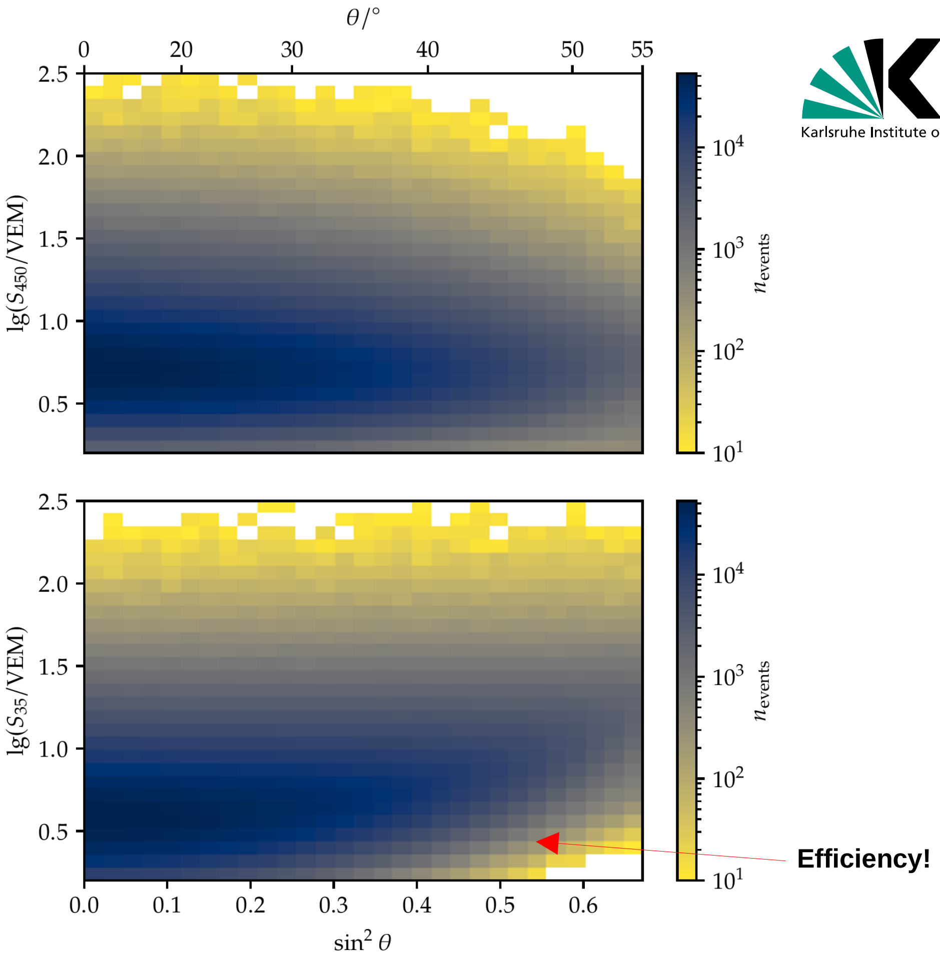
CIC needed



Constant Intensity Cut

Before: event number not isotropic due to attenuation effects

After: event number corrected for attenuation effects so constant intensity is achived



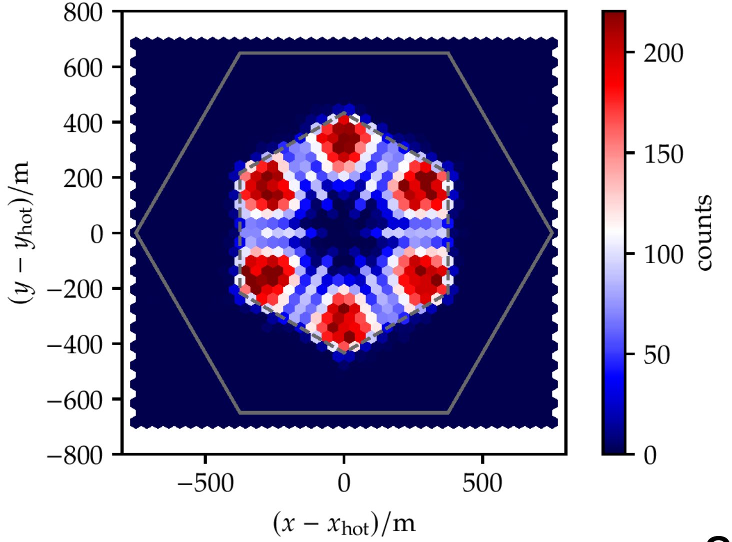
Data driven efficiency check: core distribution

Karlsruhe Institute of Technology

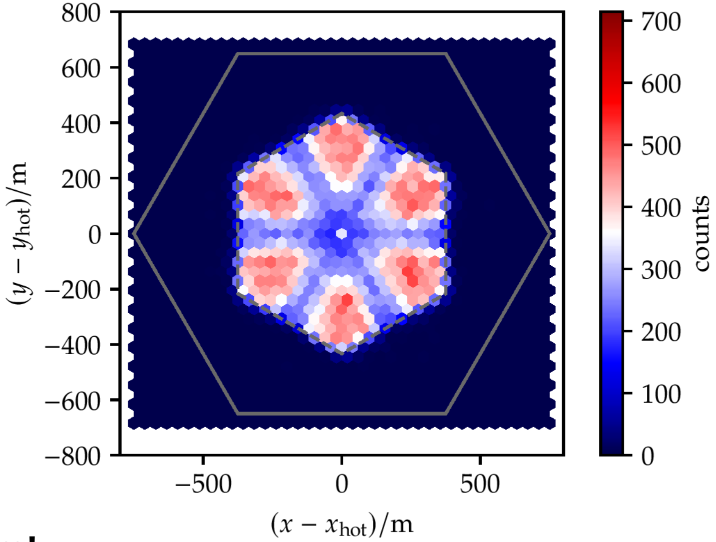
attenuation-corrected shower size

Example: data up to 30° $S_{35} \in [6, 6.5]\,\mathrm{VEM}$

Only old triggers



Old and new triggers



Clearly not uniform!

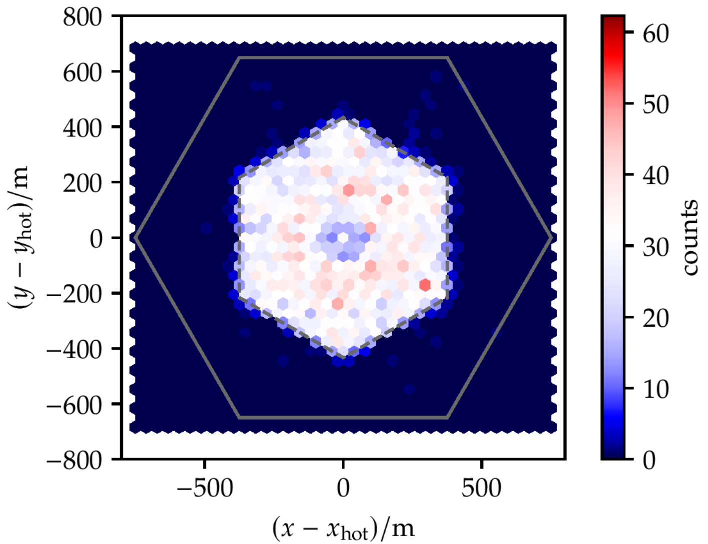
Data driven efficiency check: core distribution

Example: data up to 30°

attenuation-corrected shower size

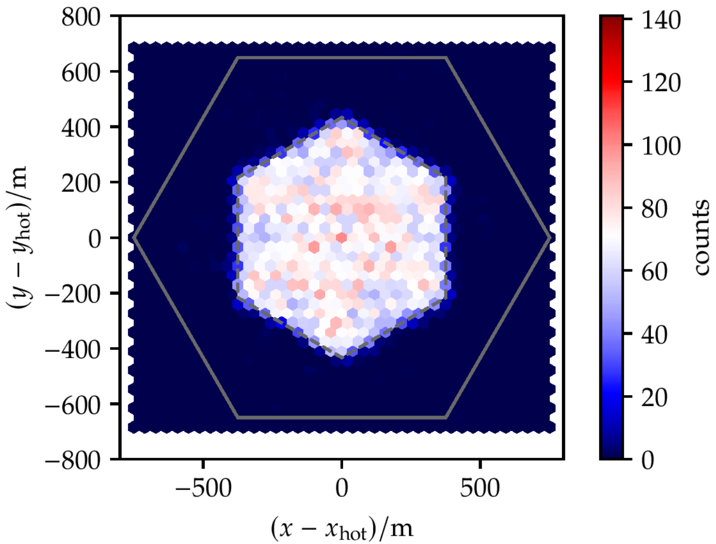
$S_{35} \in [10, 10.5]\,\mathrm{VEM}$

Only old triggers



does not look uniform

Old and new triggers

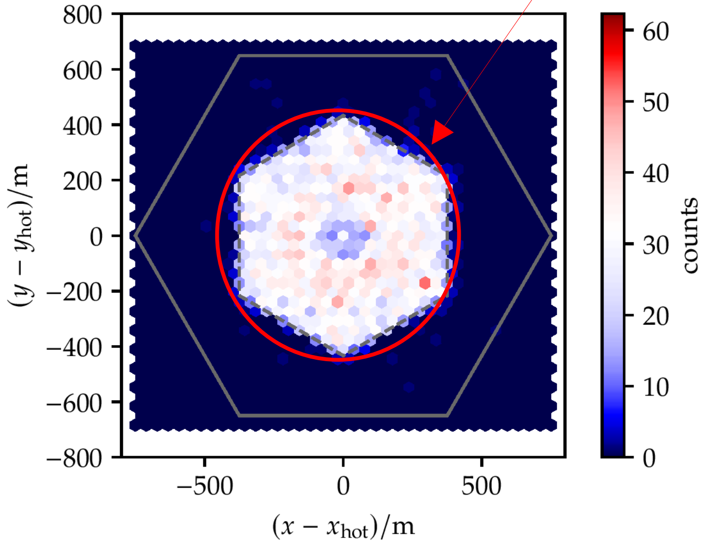


looks uniform

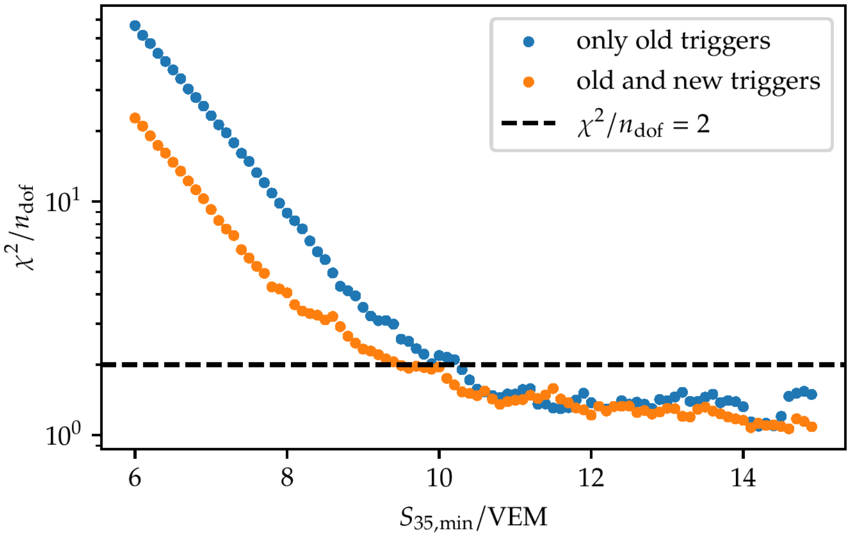
Data driven efficiency check: core distribution

Example: data up to 30°

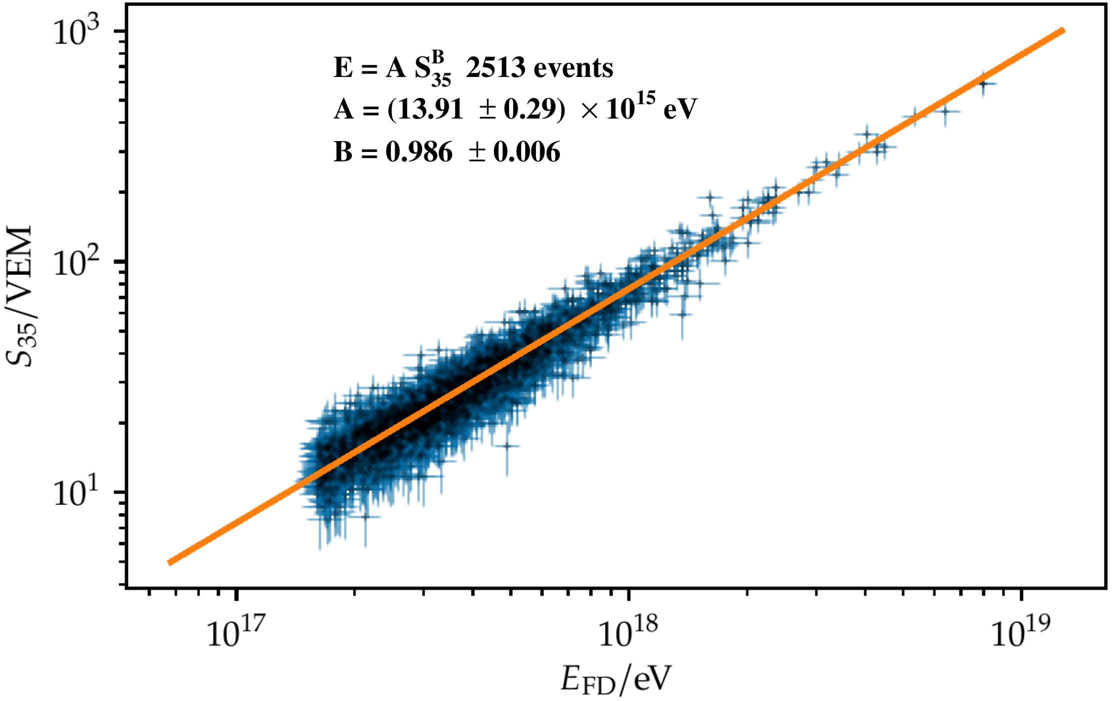
Reduced chi-square test for uniformity in this hexagon



Maybe we can estimate full efficiency with this test?



Energy calibration with SD-FD hybrids



Analysing the spectrum

Karlsruhe Institute of Technology

- Measured flux $J_{\text{meas}}(E) = \frac{n_{\text{events}}(E, \theta_{\max})}{\Delta E\ \varepsilon(\theta_{\max})}$

Array exposure (in praxis: hexagon counting)

$$\varepsilon(\theta_{\max}) = A_{\text{hex}}\, \pi\, \sin^2 \theta_{\max} \int \mathrm{d}t\ n_{\text{hex}}(t)$$

- Take detector effects into account

$J(E_i) = c_i\, J_{\text{meas}}(E_i)$ with correction factors $c_i = \frac{\mu_i(\hat{\alpha})}{\nu_i(\hat{\alpha})}$

Expected true number of events

$$\mu_j(\vec{\alpha}) = \varepsilon \int_{E_j}^{E_j + \Delta E_j} \mathrm{d}E'\ J_{\text{model}}(E', \vec{\alpha})$$

Expected measured number of events

$$\nu_i(\vec{\alpha}) = \sum_j R_{ij}\, \mu_j(\vec{\alpha})$$

Detector response (not fully investigated due to thesis deadline)

The fit model depends on a set of parameters $\vec{\alpha}$

- Find the optimal parameters $\hat{\alpha}$ by minimising a log-likelihood based on $f_{\text{Poi}}(n_{\text{events}}(E_i), \nu_i(\vec{\alpha}))$

Analysing the spectrum

Fit model $$J_{\text{model}}(E,\vec{\alpha}) = J_0 \left(\frac{E}{100\,\text{PeV}}\right)^{-\gamma_0} \prod_{i=0}^{1} \left(1 + \left(\frac{E}{E_{ij}}\right)^{\frac{1}{\omega_{ij}}}\right)^{(\gamma_i-\gamma_j)\omega_{ij}} \quad , \quad j = i+1$$

- J_0 is the over-all normalisation constant
- E_{01} is the position of the second knee with transition width ω_{01}
- E_{12} is the position of the ankle with transition width ω_{12}
- γ_0 is the spectral index before the second knee
- γ_1 is the spectral index between the second knee and the ankle
- γ_2 is the spectral index after the ankle

Analysing the spectrum

- We get consistent features when varying the full efficiency threshold!

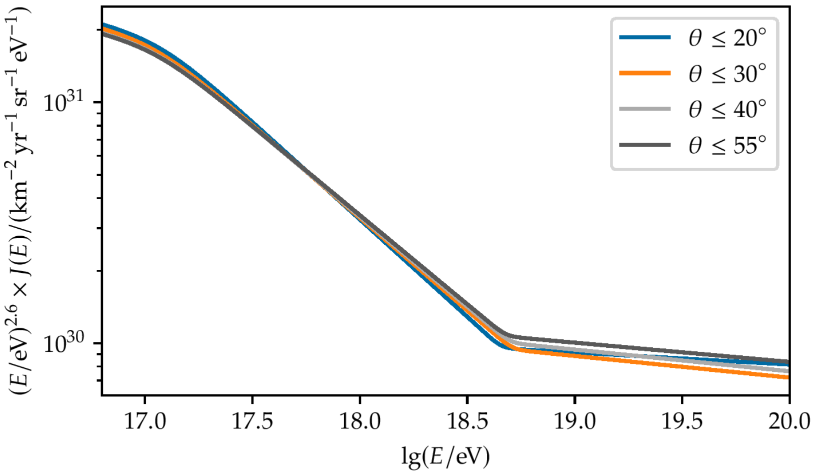
$\theta_{max} = 20°$

parameter	value $\pm\sigma_{stat} \pm \sigma_{bias}$		
$\lg(E_{full}/eV)$	16.9	17.0	17.1
$\frac{J_0 \times 10^{13}}{km^{-2}\,yr^{-1}\,sr^{-1}\,eV^{-1}}$	$1.21 \pm 0.04 \pm 0.14$	$1.20 \pm 0.03 \pm 0.17$	$1.20 \pm 0.01 \pm 0.16$
$\lg(E_{01}/eV)$	$17.07 \pm 0.03 \pm 0.02$	$17.08 \pm 0.08 \pm 0.10$	$17.08 \pm 0.04 \pm 0.33$
γ_0	$2.83 \pm 0.07 \pm 0.17$	$2.87 \pm 0.13 \pm 0.02$	$2.85 \pm 0.03 \pm 0.28$
γ_1	$3.41 \pm 0.01 \pm 0.03$	$3.41 \pm 0.02 \pm 0.01$	$3.41 \pm 0.02 \pm 0.03$
ω_{01} (fixed)	0.25	0.25	0.25
$\lg(E_{12}/eV)$	$18.66 \pm 0.08 \pm 0.02$	$18.65 \pm 0.08 \pm 0.01$	$18.65 \pm 0.08 \pm 0.05$
γ_2	$2.65 \pm 0.27 \pm 0.01$	$2.65 \pm 0.27 \pm 0.00$	$2.65 \pm 0.29 \pm 0.08$
ω_{12} (fixed)	0.05	0.05	0.05

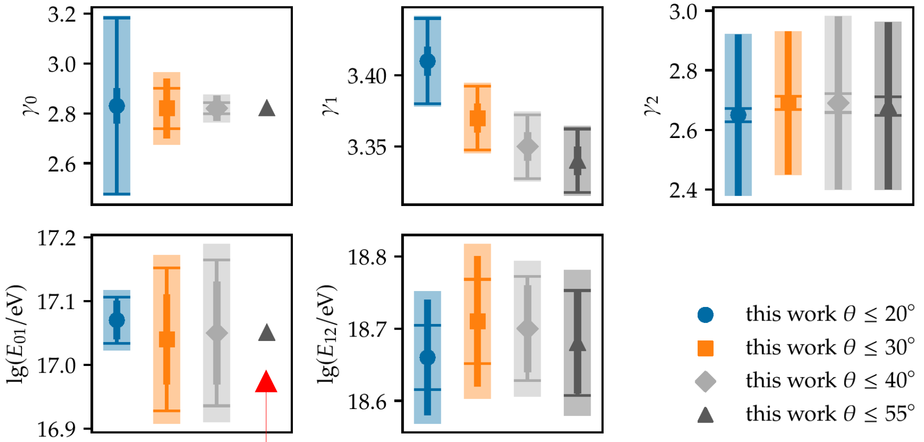
Uncertainty because of not fully investigated detector effects

Analysing the spectrum

- In general, the features are consistent for different θ_{max}

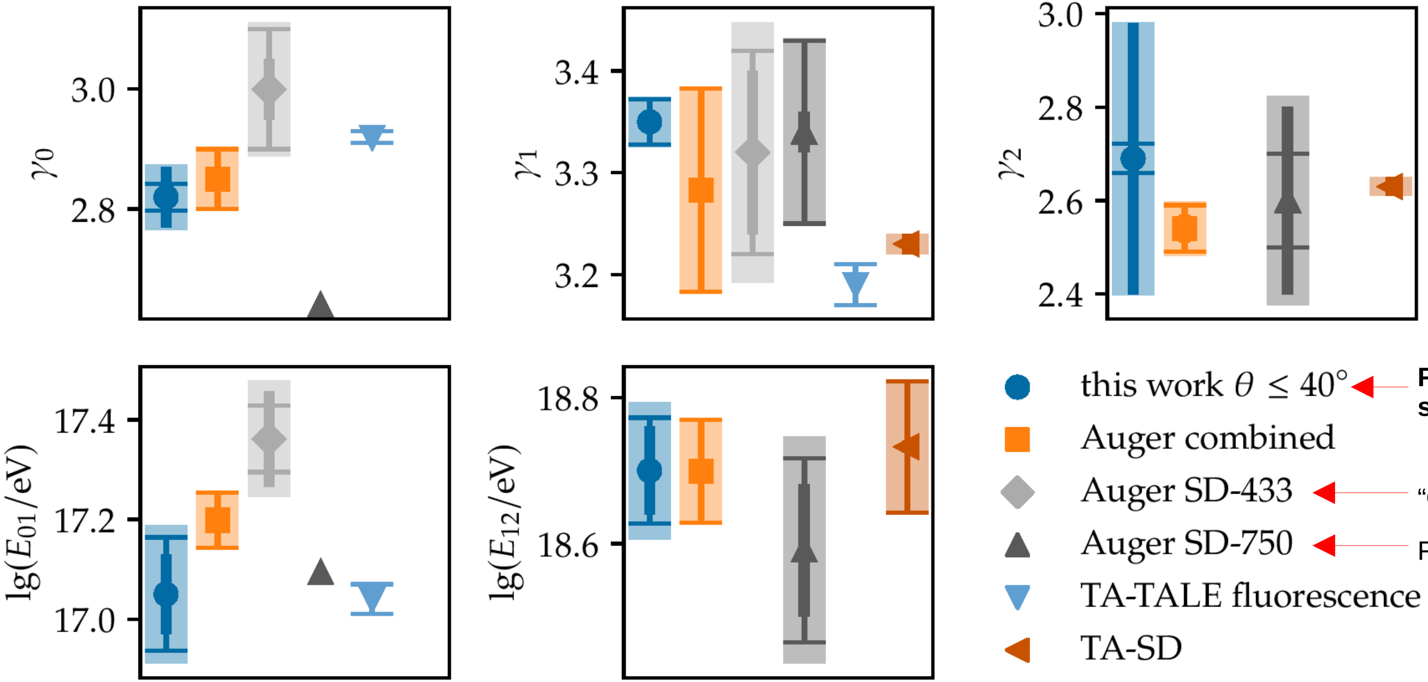


Not understood



Second knee fixed to mean of the other measurements

Comparing the measurements to other publications



First measurement of the second knee with the SD-750

“Old” cross calibration with SD-750

Fixed second knee

What to do next?

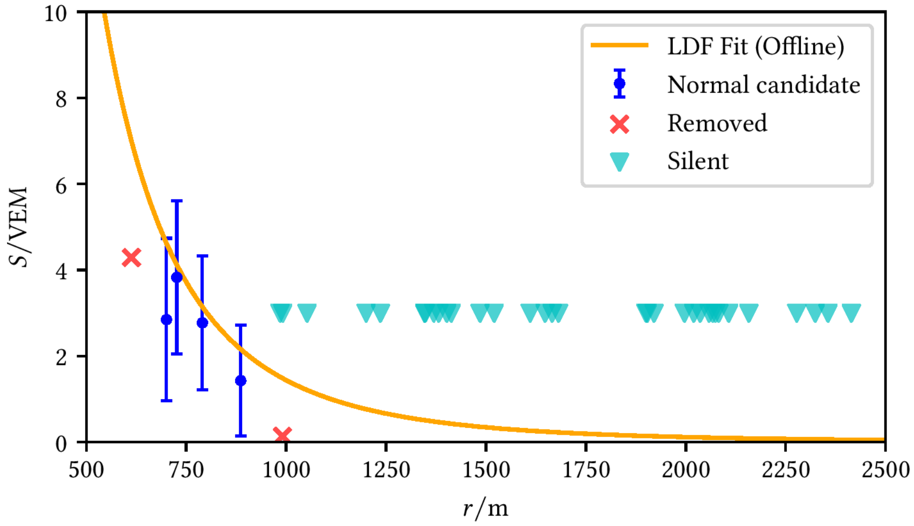
- Include the new reconstruction into the standard software
- Rerun the parameterisation-calibration-chain
- Investigate the necessary properties for the detector response
- Refit the spectral features
- Do the same for the other surface detector arrays

Backup

Selecting events for the parameterisation (lever-arm criteria)

- Many events require parameterised LDF slopes due to low station multiplicity
- Not all events with enough stations constrain the LDF slopes good enough
- We want to prevent biases in the slope parameterisation!

Find high-quality events where the slope parameters can be fitted freely and use the results for the slope parameterisation!



Lever-arm criteria for a free LDF slope fit (β)

- New criteria: LDF in log-log space is approximately a linear function

$$s_i = \ln(S_i/\mathrm{VEM}), \qquad \rho_i = \ln(r_i/\mathrm{m})$$

$$\tilde{s}_i = s_i - \langle s \rangle\,, \quad \sigma(\tilde{s}_i) \approx \sigma(s_i), \qquad \tilde{\rho}_i = \rho_i - \langle \rho \rangle$$

Fit $\tilde{s}(\tilde{\rho}) = m\,\tilde{\rho} + c$

Correlation coefficient

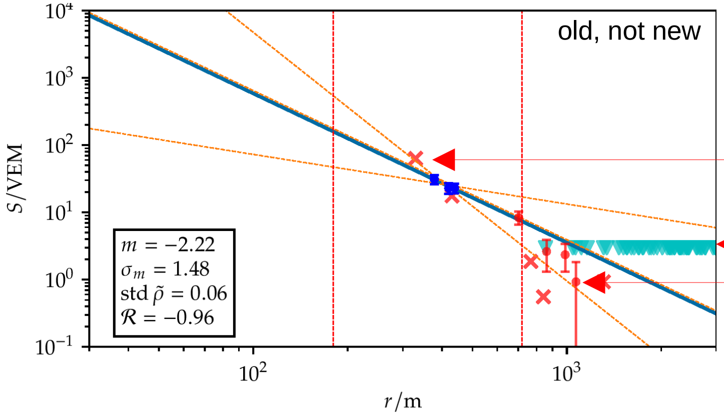
array	$\sigma_m^{\max}$	$[m_{\min}, m_{\max}]$	$\mathrm{std}(\tilde{\rho})_{\min}$	$\mathcal{R}(\tilde{s},\tilde{\rho})_{\max}$
SD-1500 (Not verified with data)	0.5	[−4, −1]	0.2	−0.7
SD-750	0.5	[−4, −1]	0.2	−0.7
SD-433 (Not verified with data)	0.5	[−4, −1] (?)	0.2 (?)	−0.7

Fit accuracy — Physical behaviour — Not dominated by clustered stations — Physical behaviour

Lever-arm criteria for a free LDF slope fit (β)

- Event quality conditions:
  - Minimum of 5 stations
  - 6T5 and $\theta \leq 55°$
  - For comparison with old criteria:
    - T4 trigger efficiency $p_{\mathrm{T4}} \geq 0.99$
    - Fitted curvature, no saturation

More information in GAP-2024-033

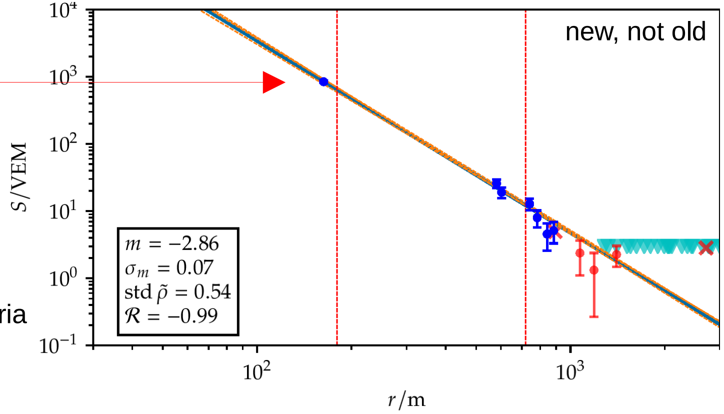


Stations used in the fit

Rejected stations

Silent stations

Stations with $p_{\mathrm{trig}} < 0.9$
Not used for the criteria (criteria should not be dominated by low-signal stations)



Is the shower size correctly estimated?

- Stations with

  $r_i = (450 \pm 20)\,\mathrm{m}$

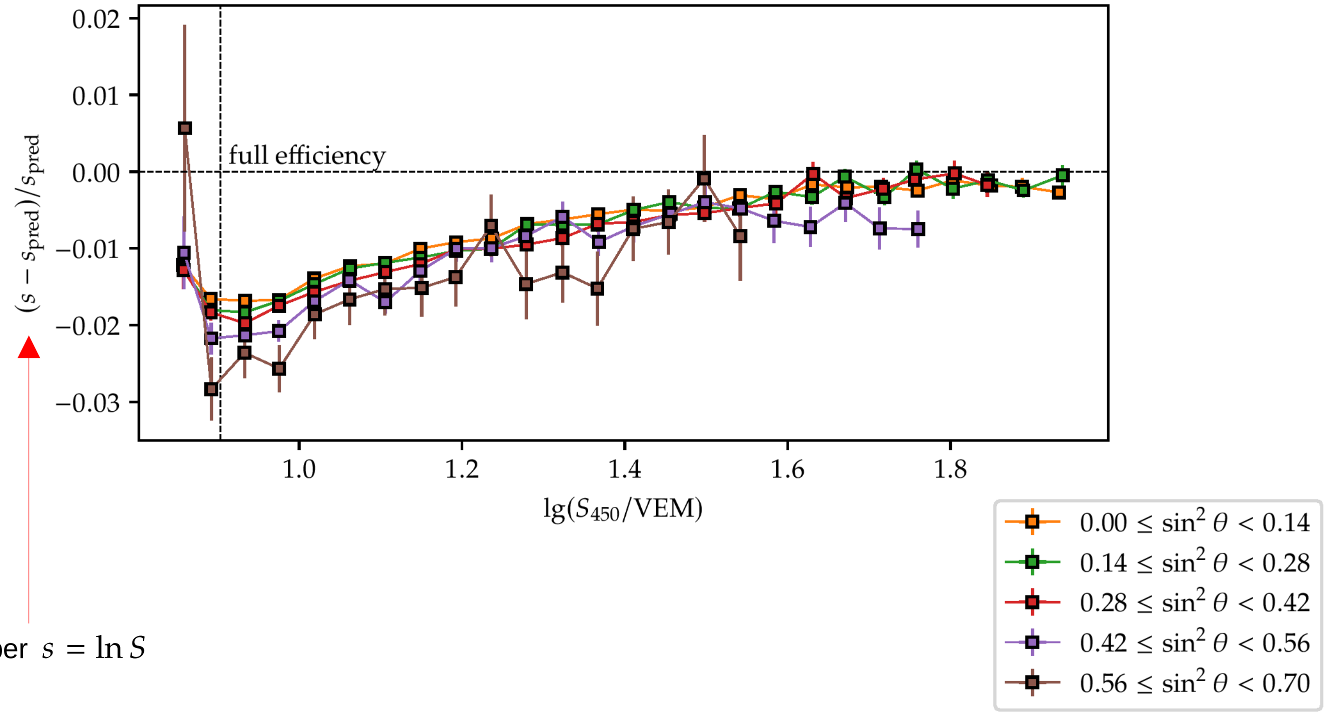
- Residuals:

  Only reasonable in a gaussian limit, **goodness of fit is preferred**

- Only stations with

  $S_{\mathrm{pred}} \geq 8\,\mathrm{VEM}$

  Statistical effect or real bias?



Remember $s = \ln S$

Is the shower size correctly estimated?

- Stations with

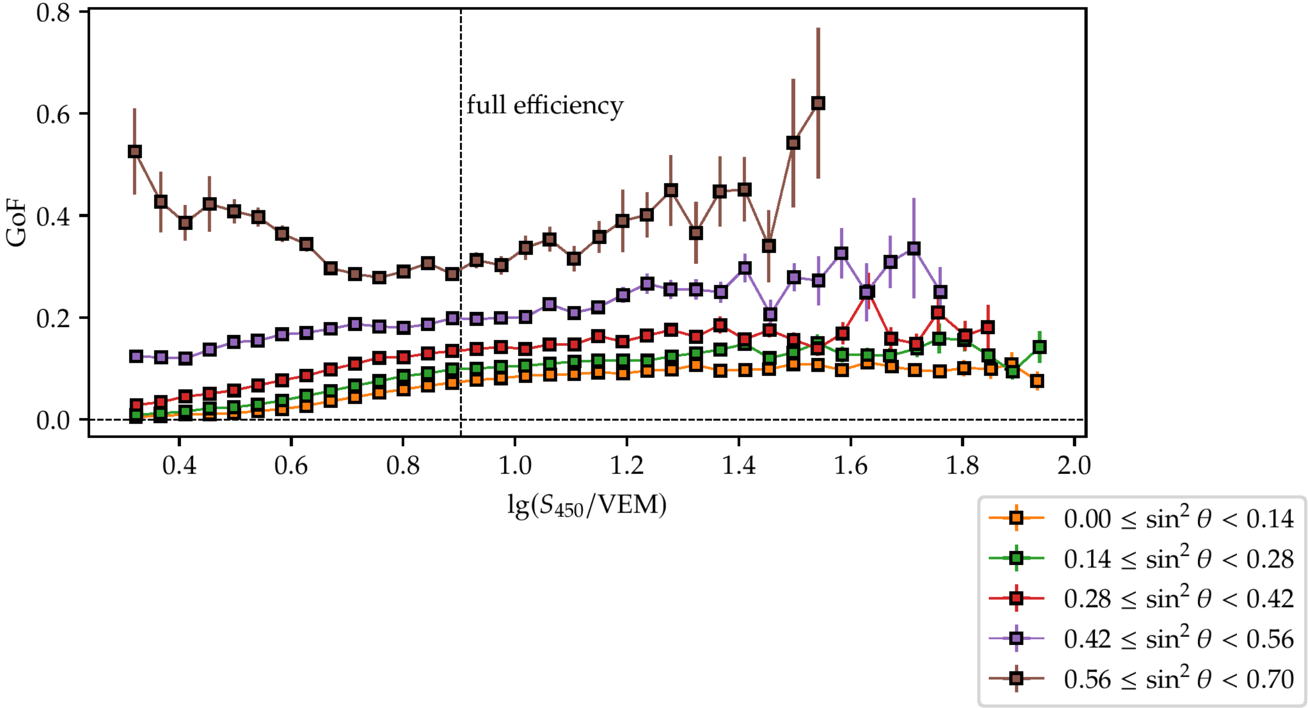
  $r_i = (450 \pm 20)\,\mathrm{m}$

- Goodness of fit:

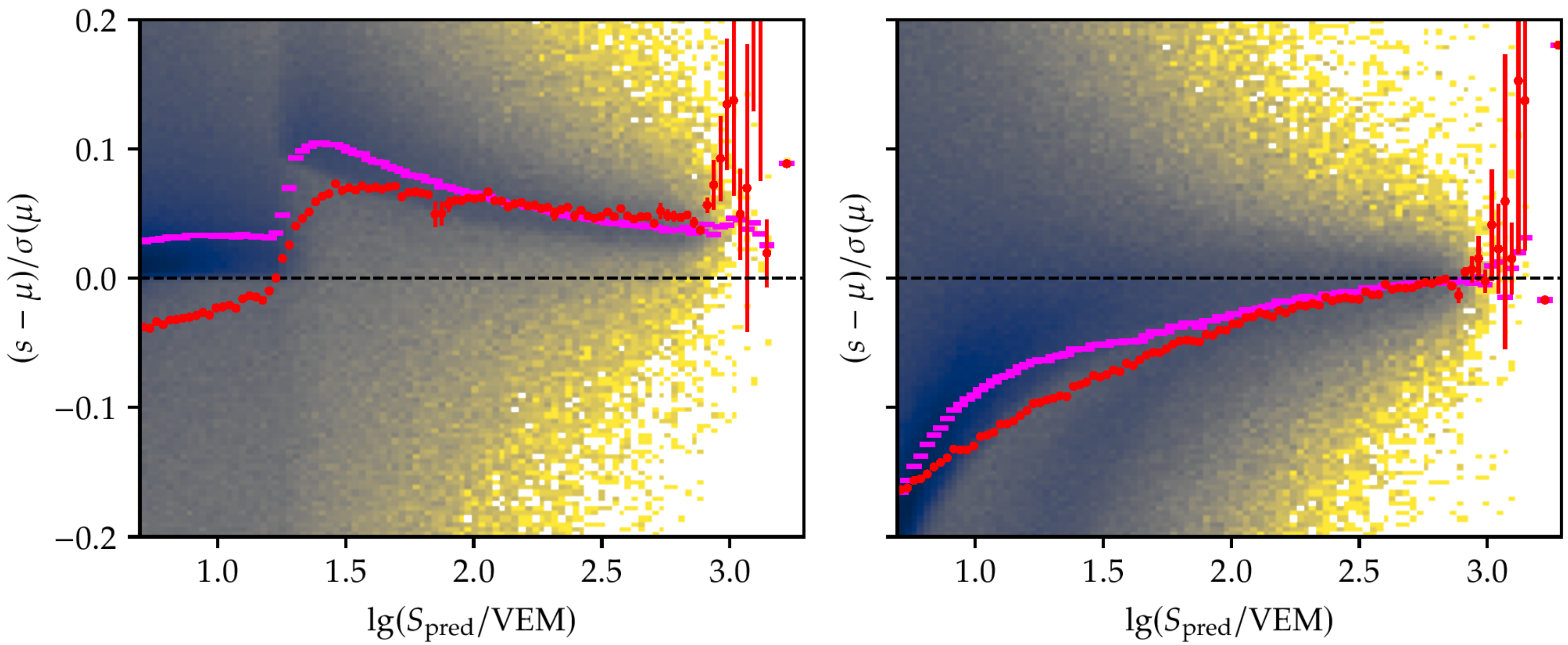
  $-2(\ln \mathcal{L}_i - \ln \mathcal{L}_{\mathrm{sat},i})$

  LDF fit   best possible fit

  $\rightarrow \chi^2$ in a gaussian limit

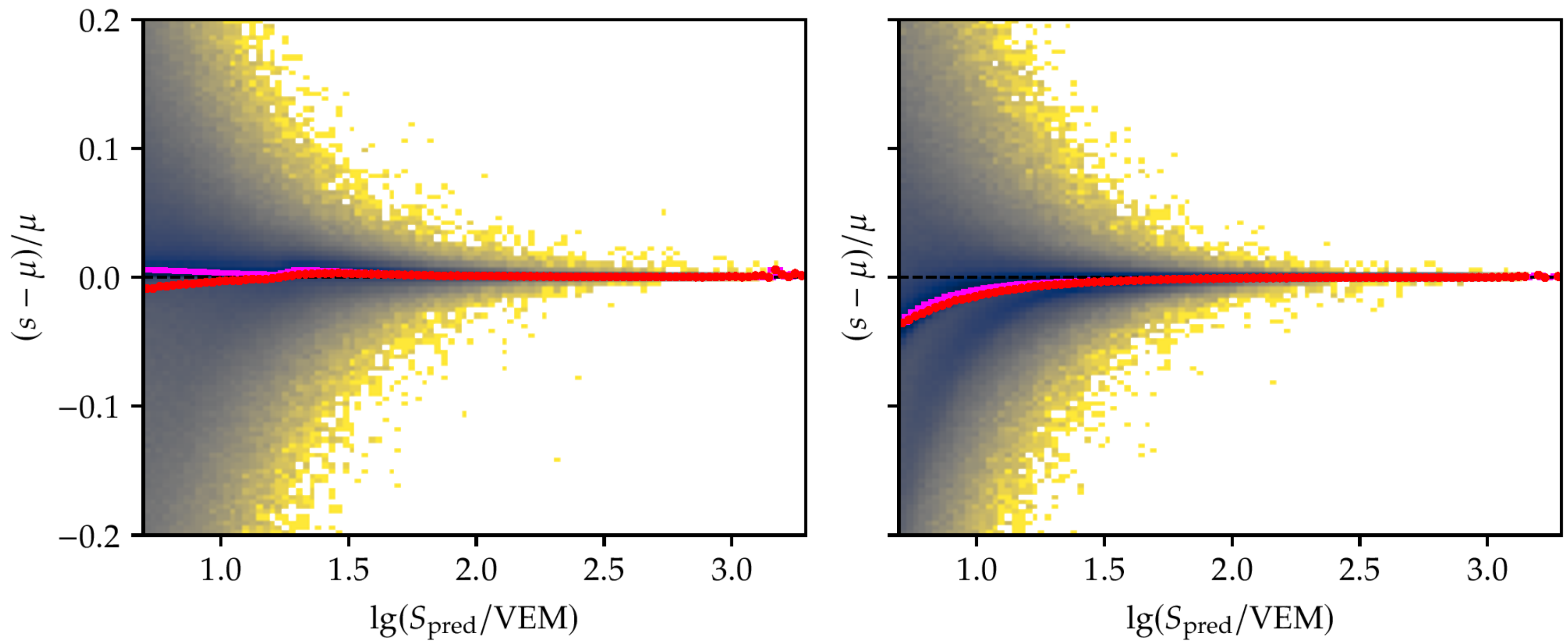


Comparison with old reconstruction



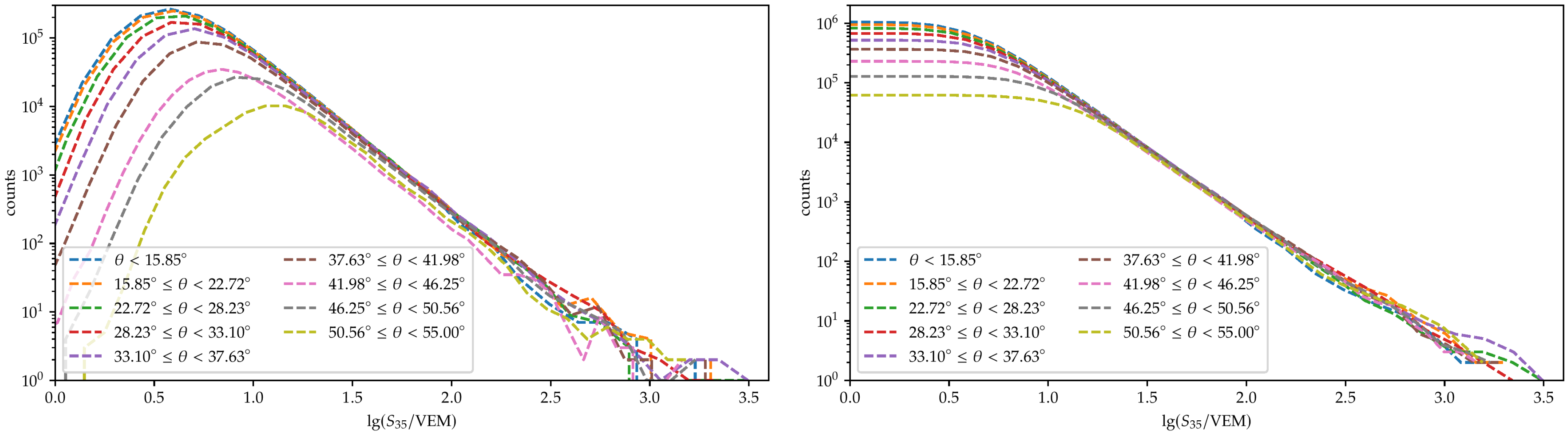
Trigger probability > 90%

Comparison with old reconstruction

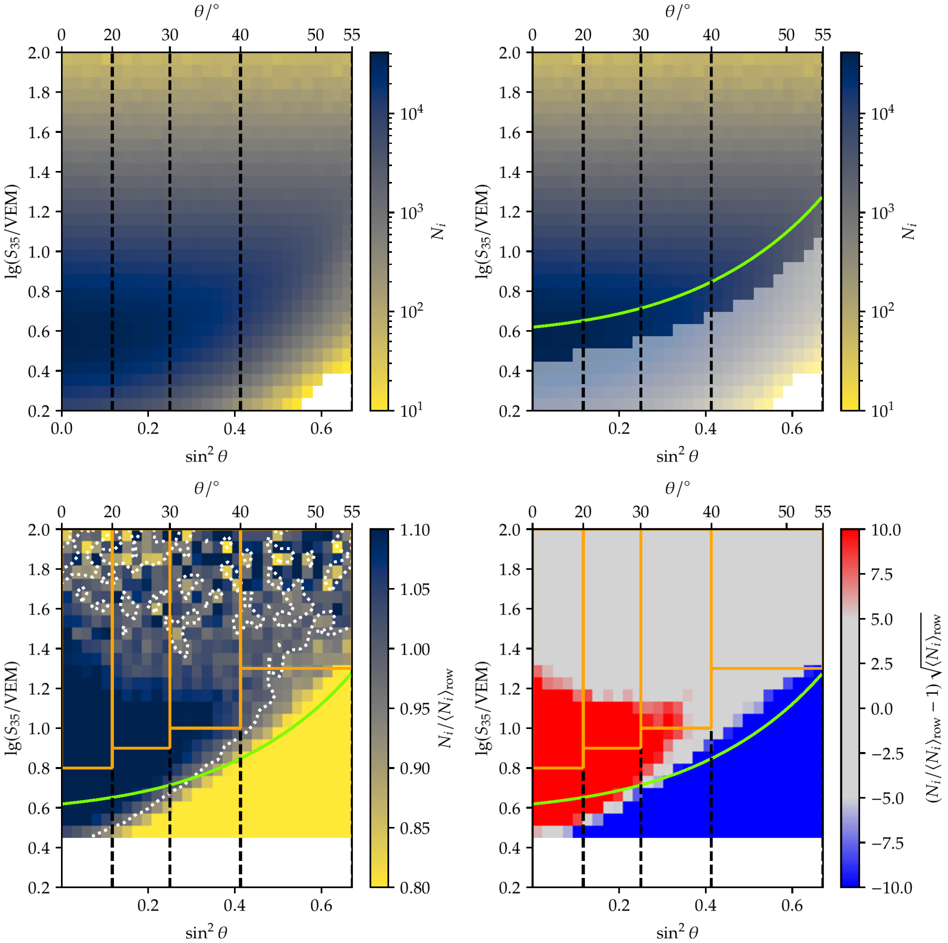


Trigger probability > 90%

Constant Intensity Cut



Efficiency



Raw spectrum (uncorrected for detector effects)