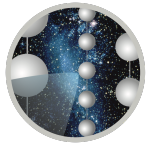
Improving cascade resolution studies

Tianlu Yuan

Calibration workshop

August 4, 2017

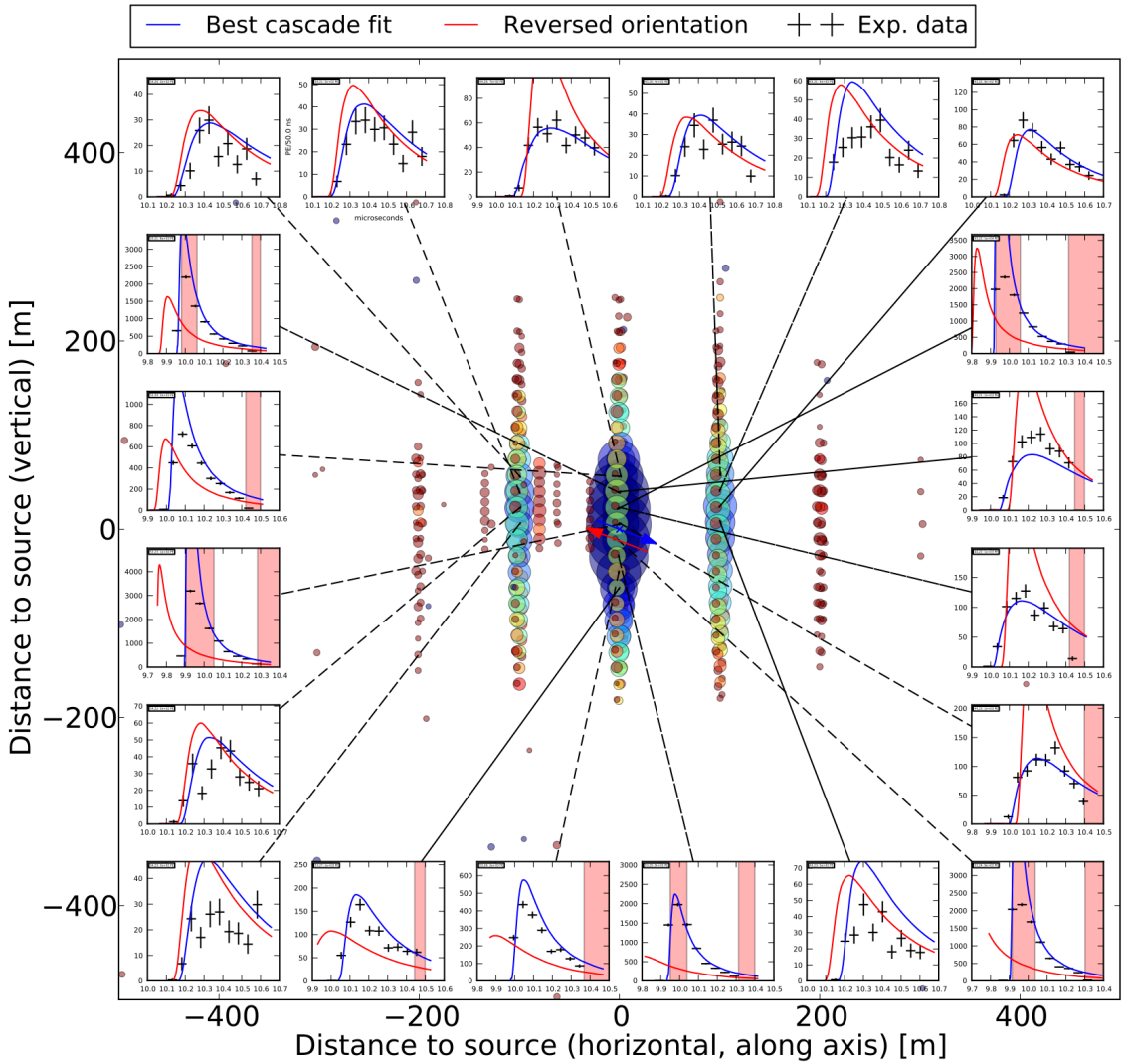


Motivation

- Waveform disagreements between MC and data hinted that there is room to improve reconstruction
- Furthermore, bright-DOM exclusions in HESE mean current resolutions do not take all data into account
- Try to quantitatively evaluate the effect of the ice-model and bright-DOM exclusions on reconstructed resolutions
  - Check what limits are possible for ideal reconstruction

Waveforms and cascade orientation

Bert “Panopticon” plot



Time-windows where PMT saturates or marked as errata are shaded in red

Reconstruction relies on waveform amplitude and timing

Noticeable differences between best-fit and reversed-orientation directions

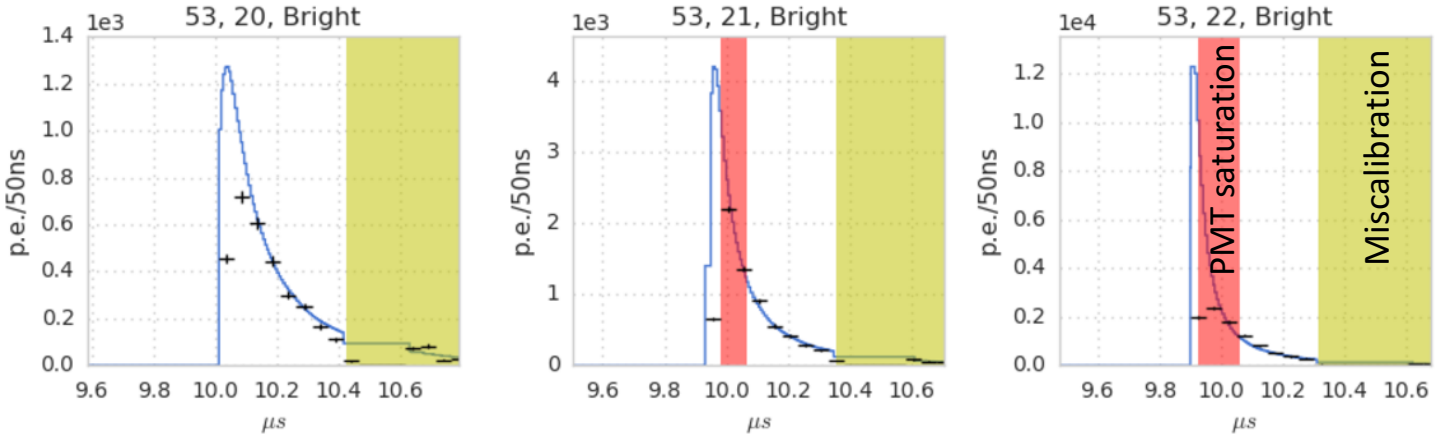
Some disagreement between best-fit and data remain and hint that there is room to improve reconstruction

Two approaches to improved resolutions

1. Include more data

A few examples of unused waveforms

Bert waveforms on closest string

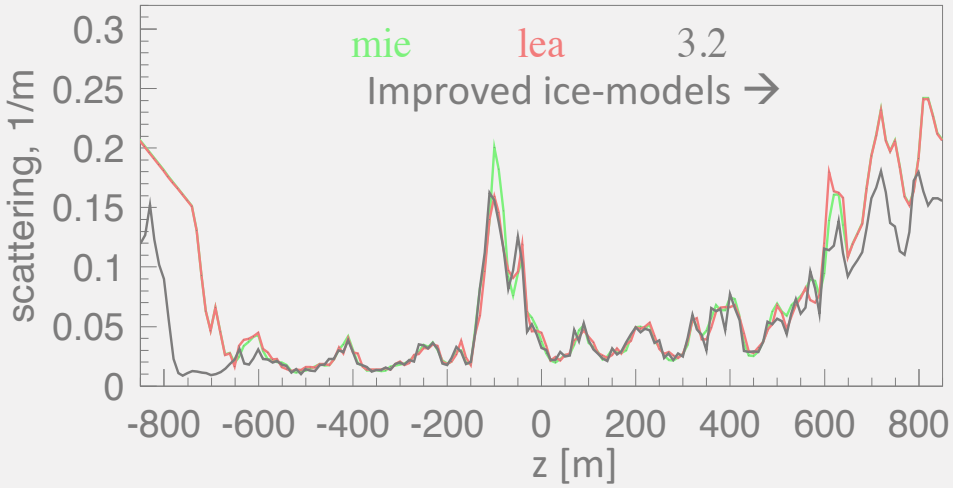


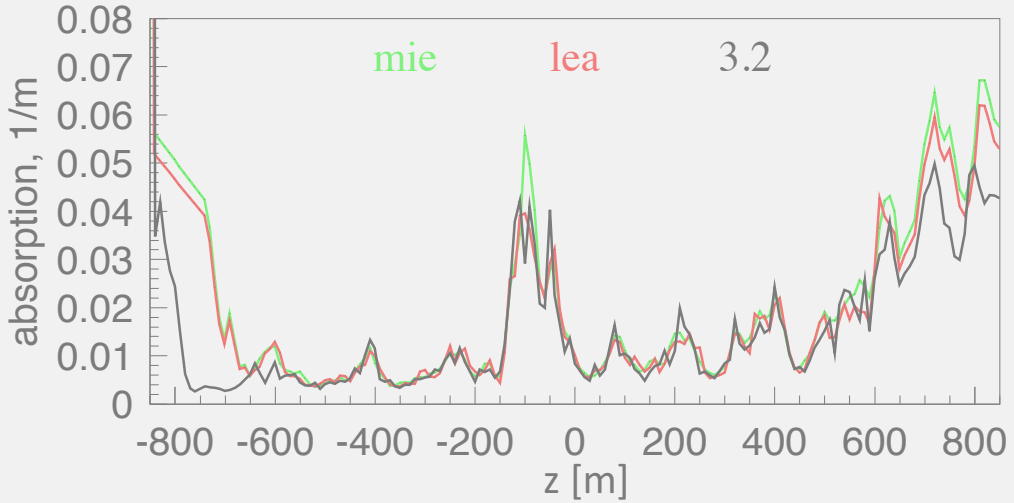
Bright DOMs are ignored completely in reconstruction

2. Improve ice model, reduce ice uncertainties

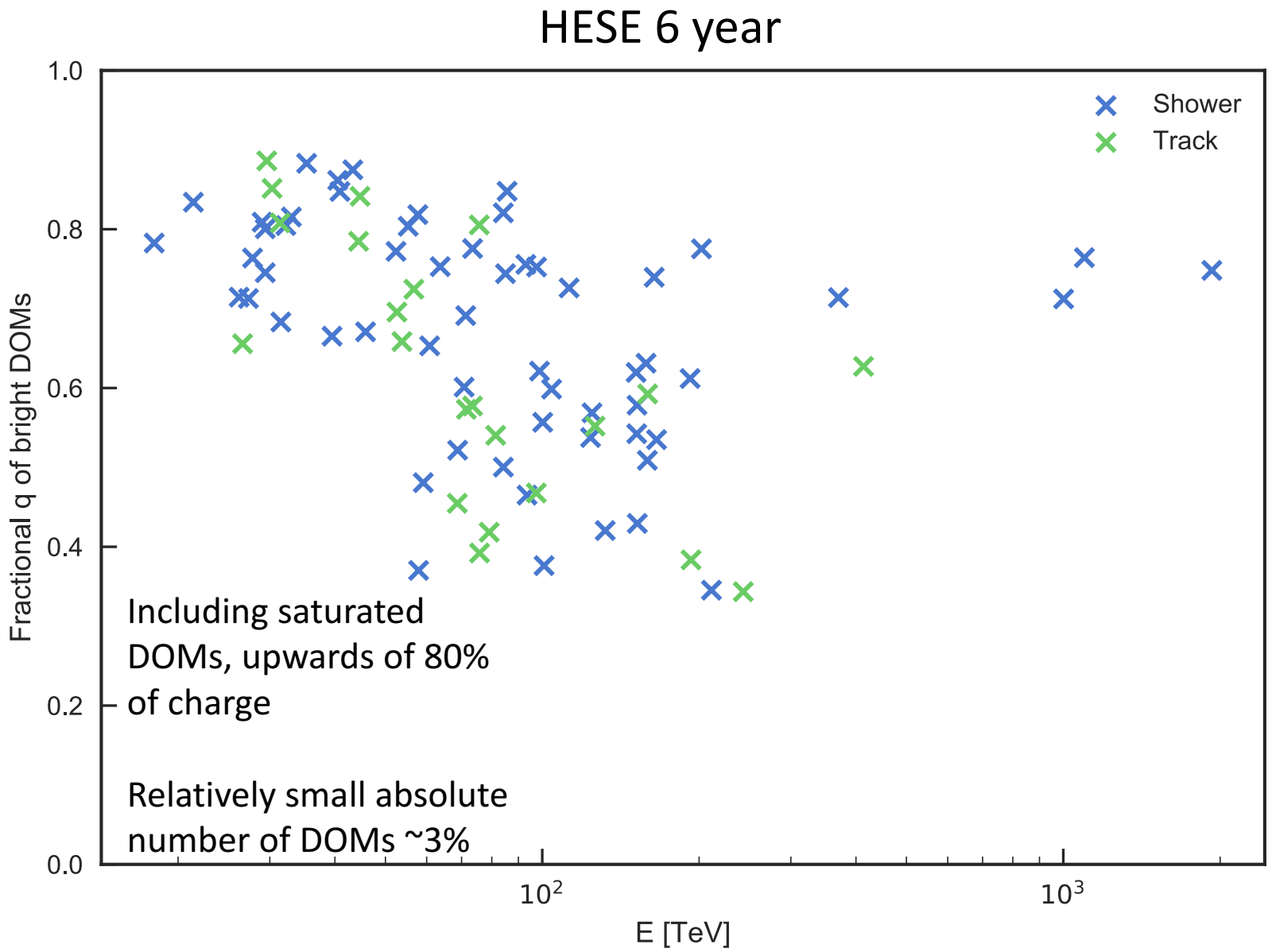
Currently an effective ice-model error of 10%

D. Chirkin





Bright DOMs in high energy events



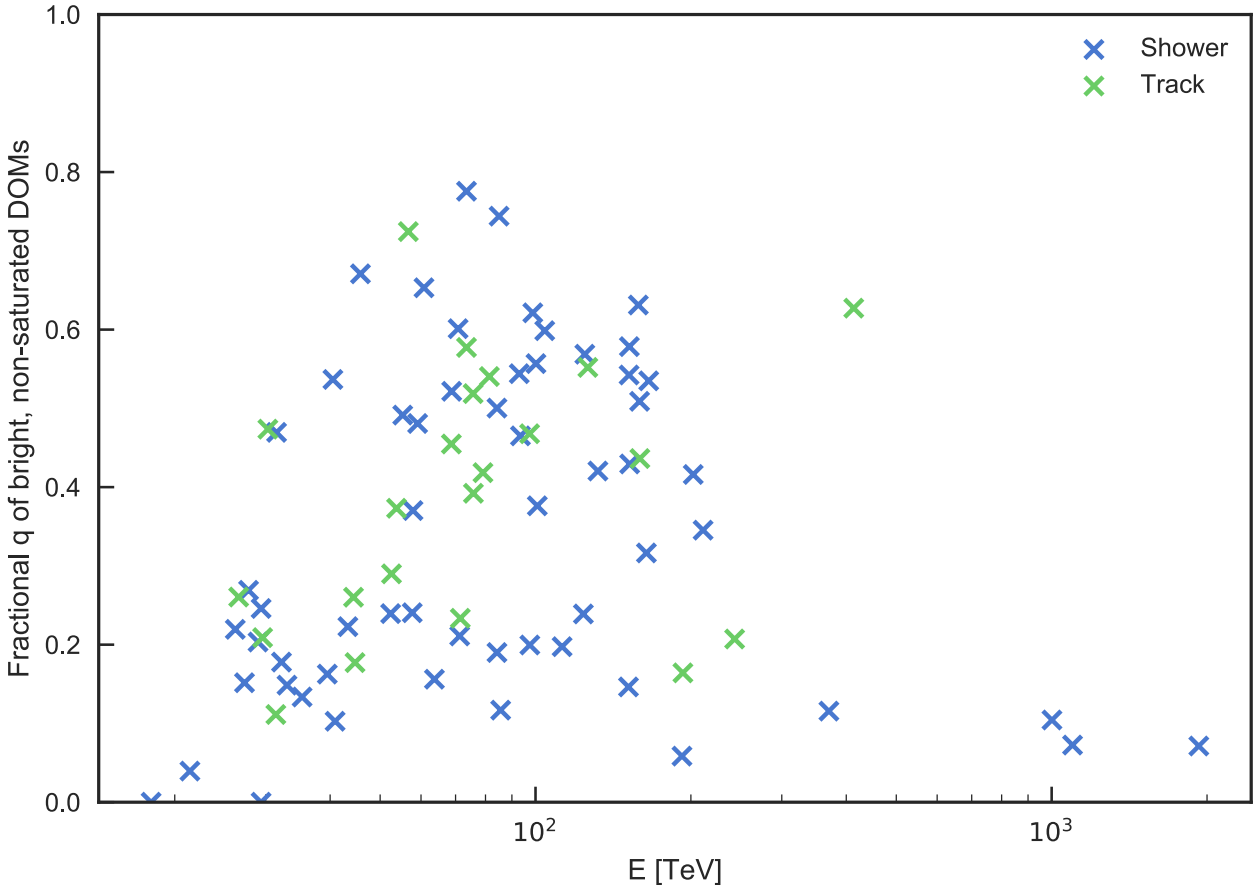
Define Q_{avg} as the mean total charge of all hit DOMs

DOMs with $Q_{bright} > \alpha Q_{avg}$ are classified as "Bright".

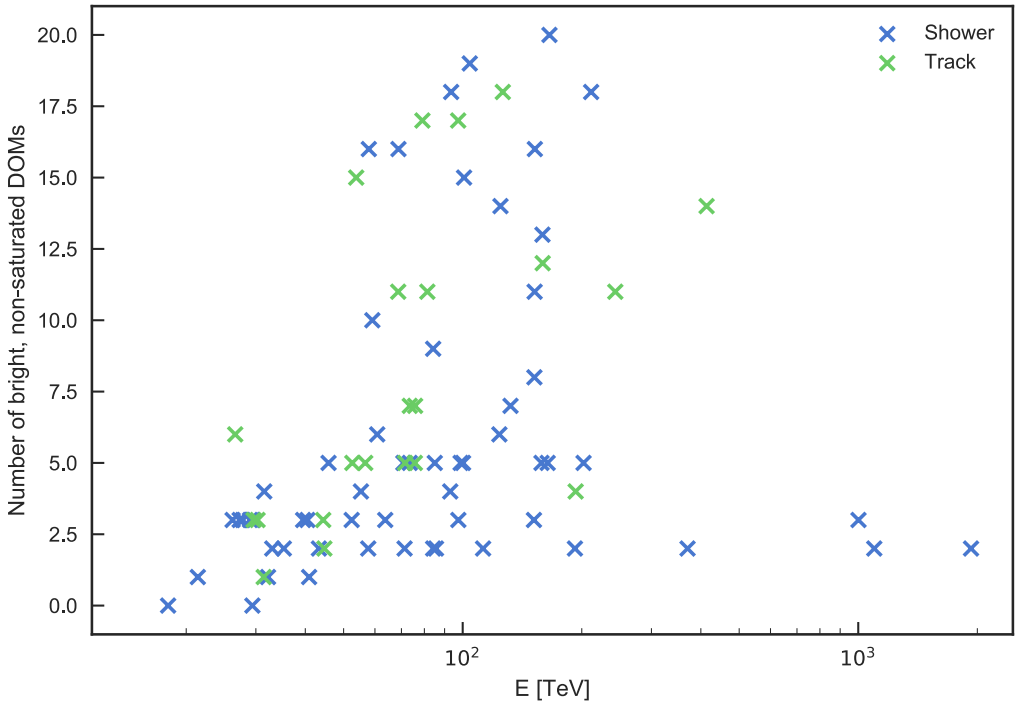
- Default α=10

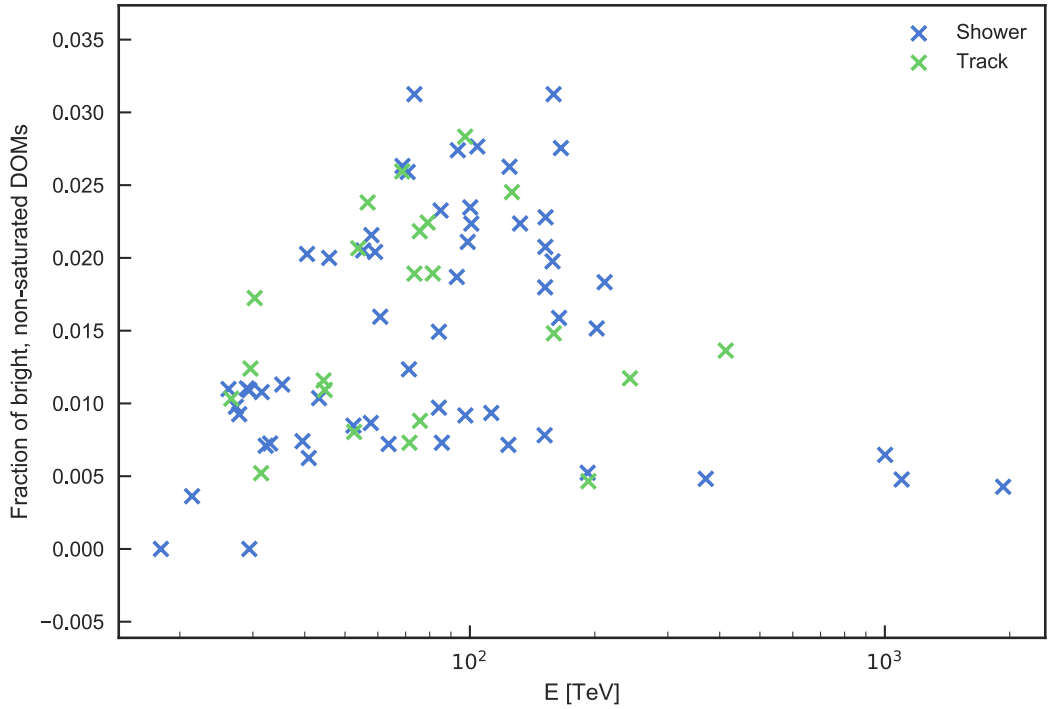
PMT is not necessarily saturated, but excluded because systematic uncertainties start to dominate over statistical errors in fitting the waveforms

Bright but not saturated



Saturated DOMs are typically a subset of bright DOMs.





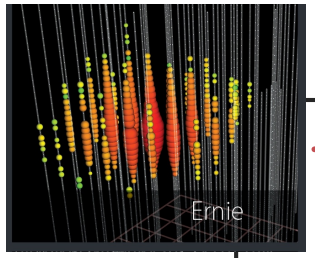
Distance to vertex

Strong dependence on distance from the vertex

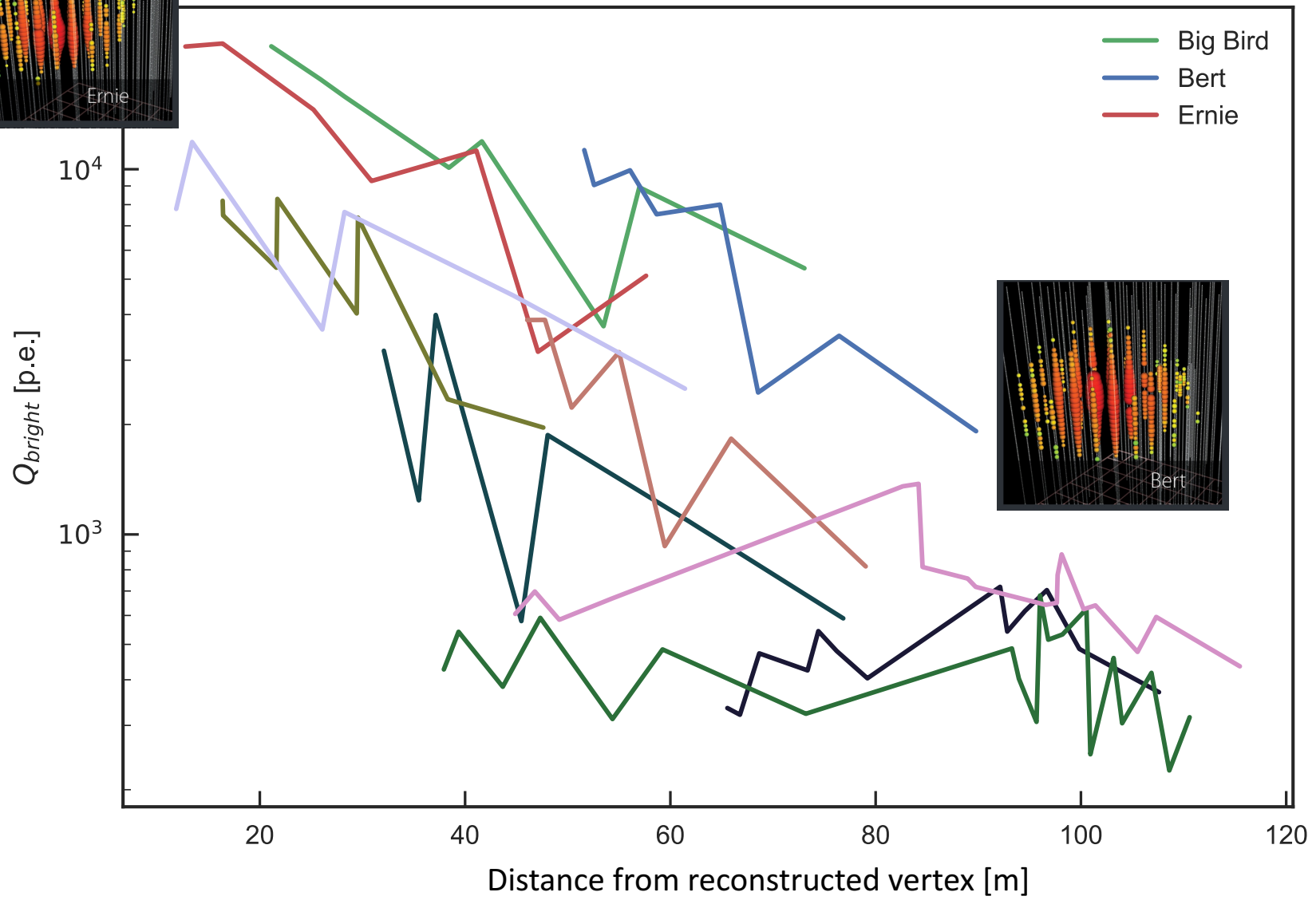
Typically are set of DOMs closest to vertex

Mostly within single string spacing (125m)

Should help with directional reconstruction!



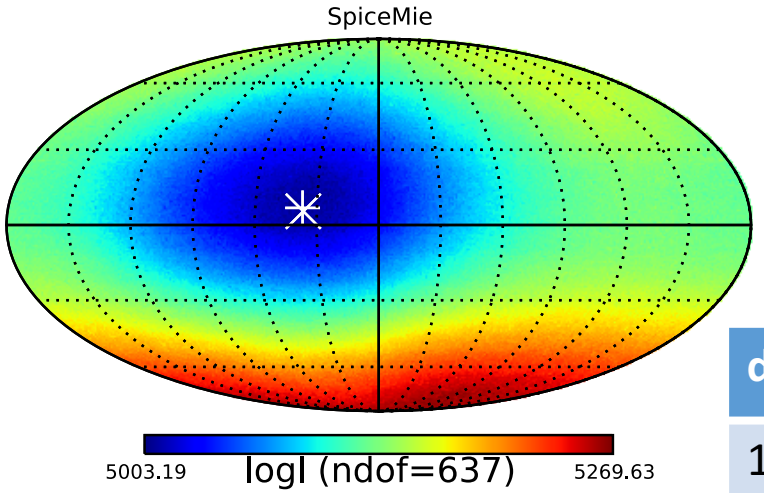
Top 10 energetic cascades in HESE 6 year

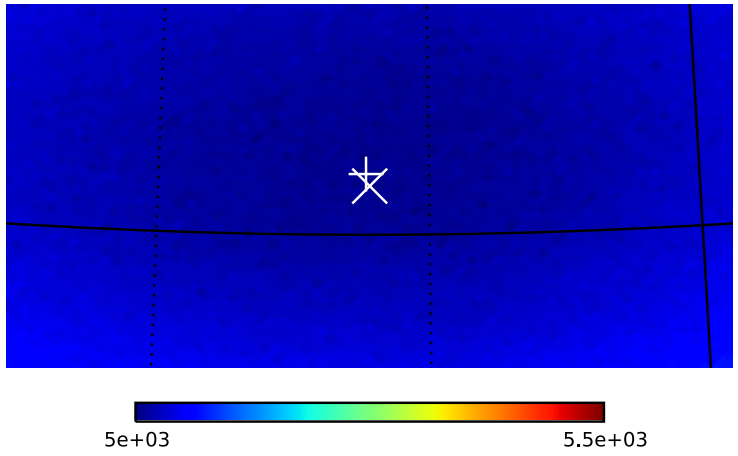


Rowlf: A particularly bad case

x = full sky scan
\+ = iterative monopod

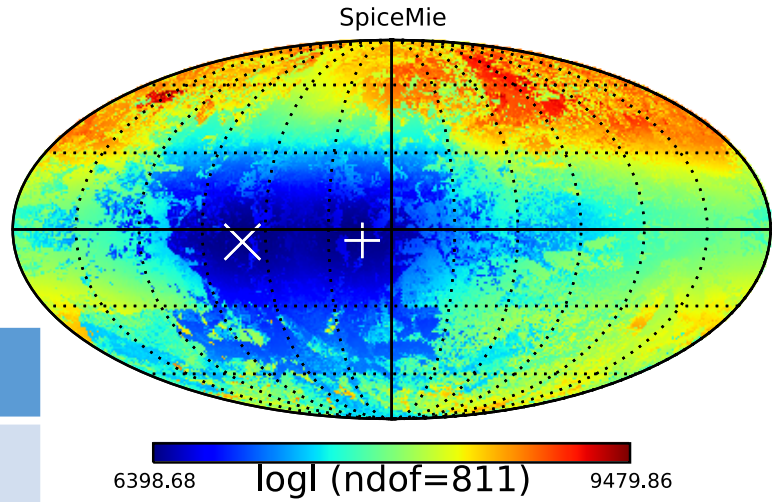
Without Brights

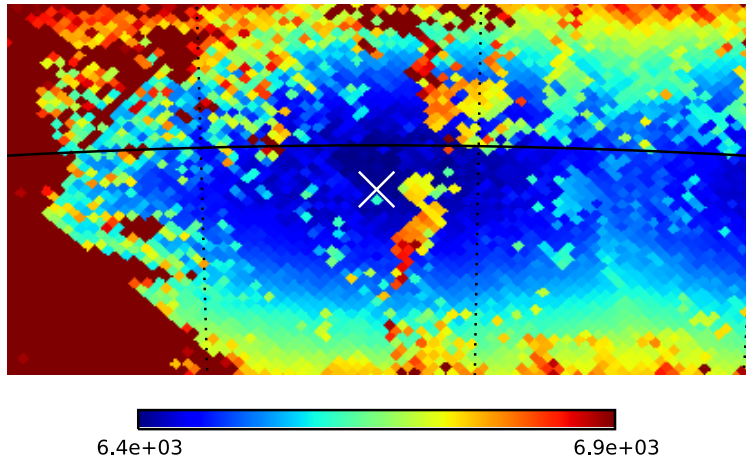




dE/E [%]	dθ [deg]
16	36

With Brights





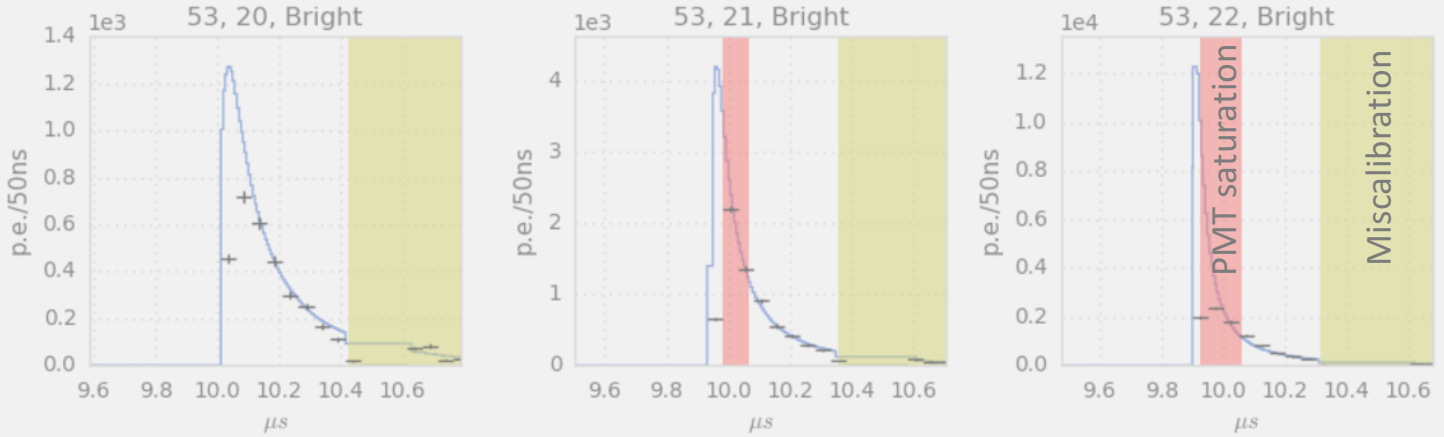
Monopod likelihood map destabilizes with inclusion of bright DOMs

Two approaches to improved resolutions

1. Include more data

A few examples of unused waveforms

Bert waveforms on closest string

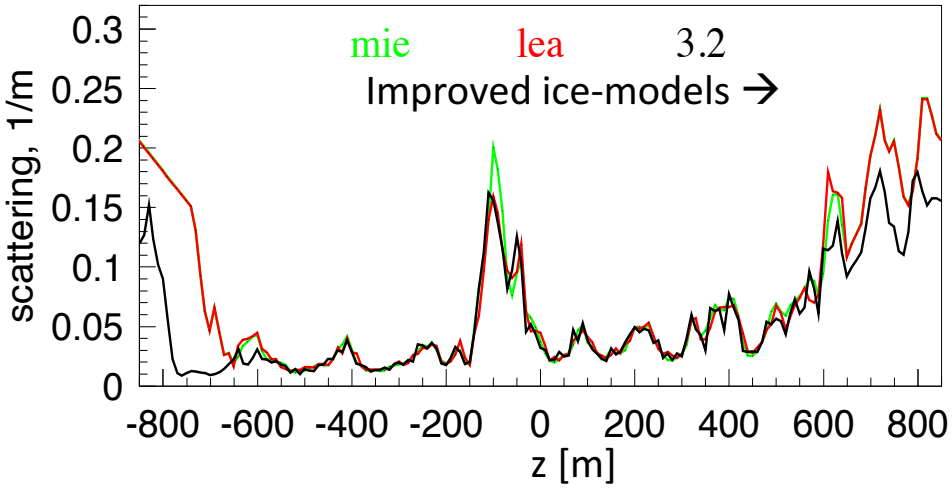


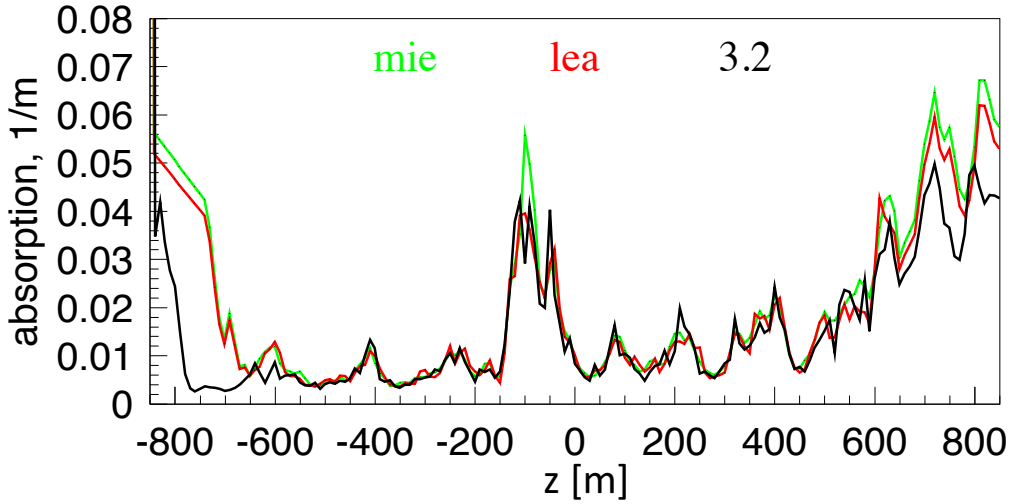
Bright DOMs are ignored completely in reconstruction

2. Improve ice model, reduce ice uncertainties

Currently an effective ice-model error of 10%

D. Chirkin





DirectFit

- DirectFit LLH includes an effective ice-model uncertainty that smears the charge on each DOM +/- 10% (default)
- This ensures that the fit isn't too biased by high statistic DOMs

$$-\ln \mathcal{L} = \sum_i \left[s_i \ln \frac{s_i/n_s}{\mu_s^i} + d_i \ln \frac{d_i/n_d}{\mu_d^i} + \frac{1}{2\sigma^2} \ln^2 \frac{\mu_d^i}{\mu_s} \right].$$

"Likelihood description for comparing data with simulation of limited statistics", D. Chirkin, arXiv:1304.0735

DirectFit

Capable of reconstructing data with direct photon simulation with ppc

Likelihood function different from the mainstream recos as the expectations from simulation is no longer analytic (e.g. Millipede)

Fit routine proceeds through several iterations of a localized random search where many position and direction are tested and the best fit energies at those steps are calculated.

Following fit, approximate Bayesian calculation (ABC) method applied based on fit results to estimate posterior via MCMC.

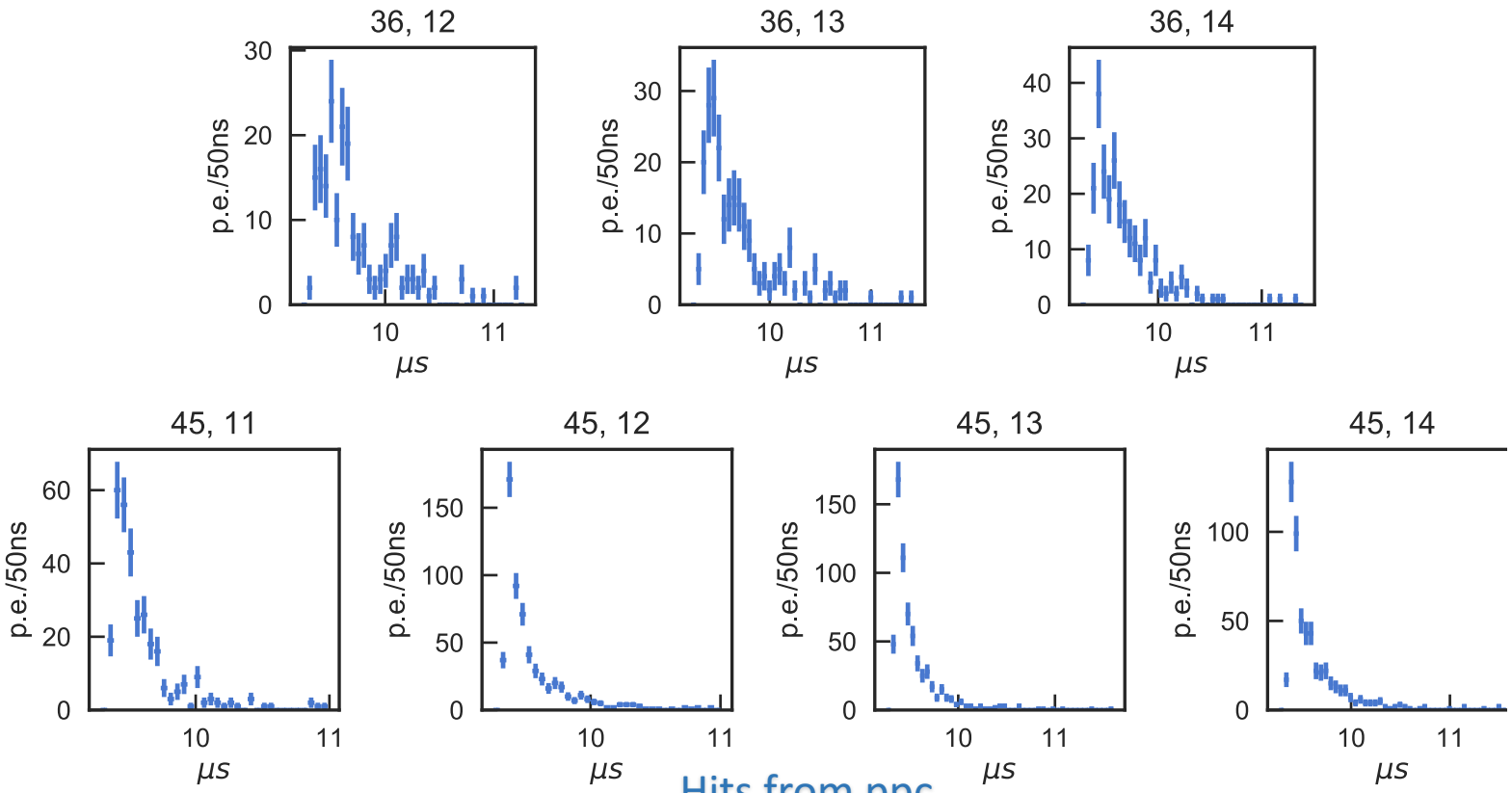
Ref. arXiv:1309.7010

Procedure

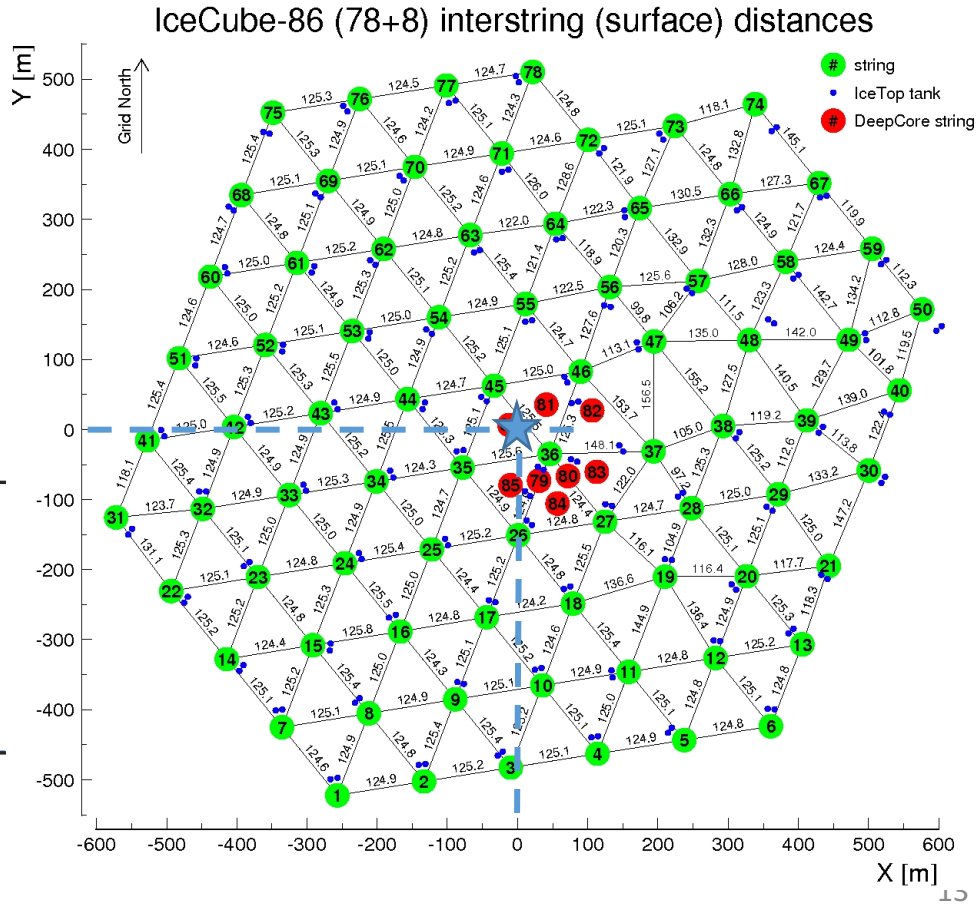
1. Simulate a EM cascade with ppc at
   - r=(0, 0, 300)m → 1648m depth
   - θ=(90 zenith, 0 azimuth)
   - Ice-sim=3.2
   - E=1E[3, 4, … 7] GeV
2. For each simulated cascade, use DirectFit to try and reconstruct the best fit point assuming
   - Ice-rec=(spice-Mie, 3.2)
   - σ=(0.0, 0.05, 0.1) ice model uncertainty
   - Q_{max}=(300, 500, 1000, 3000, 5000, 10000) p.e. cut off such that DOMs with $Q_{DOM}>Q_{max}$ are excluded
3. Once best fit is found, sample from the approximate posterior distribution $P(r, \theta|D)$ for each combination of ice models, energies, and sigmas
   - Std deviation of this sample gives resolution: δr, δθ, δE
   - And pulls: $\frac{E-E_{true}}{\delta E}$ etc.

An example: step 1, simulation

1. E=100 TeV, ice=spice-3.2 (latest), r=(0, 0, 300), θ=(90z, 0a)

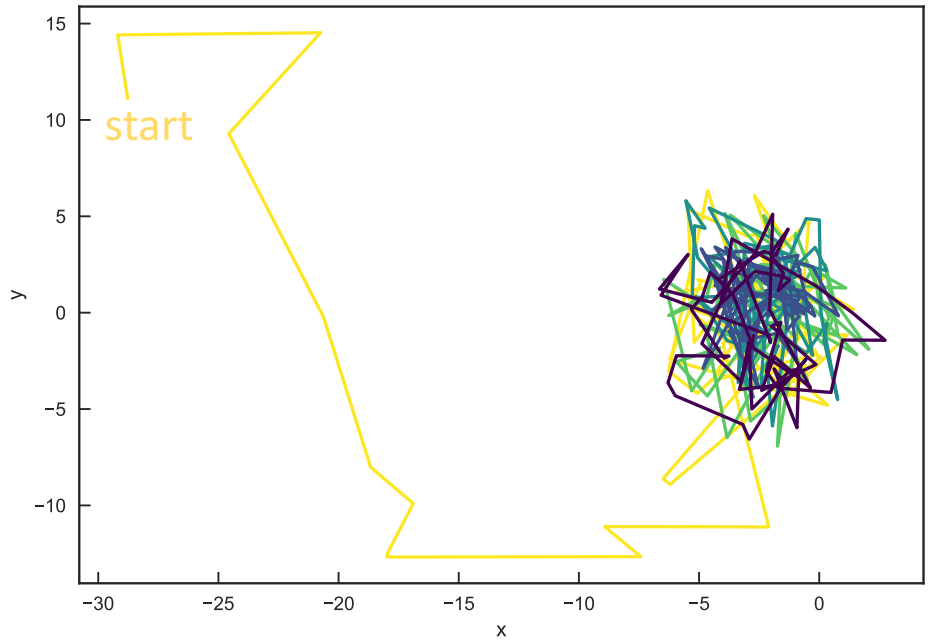


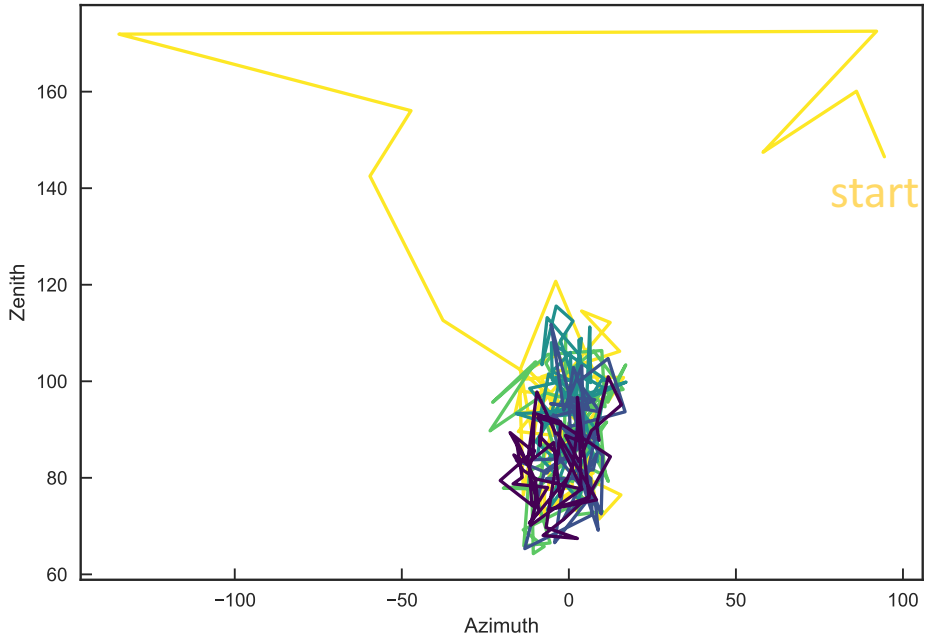
Hits from ppc



An example: step 2, reconstruction

1. E=100 TeV, ice=spice-3.2 (latest), r=(0, 0, 300), θ=(90z, 0a), σ=0.0
2. DirectFit steps to the minimum

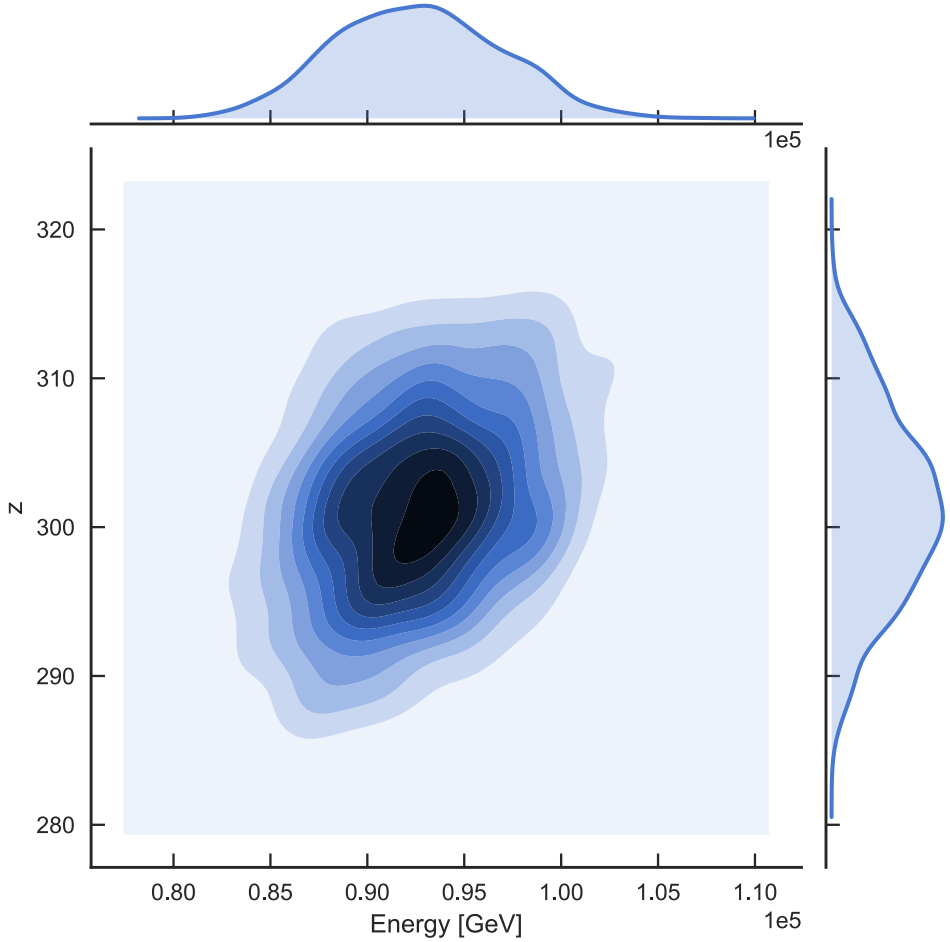


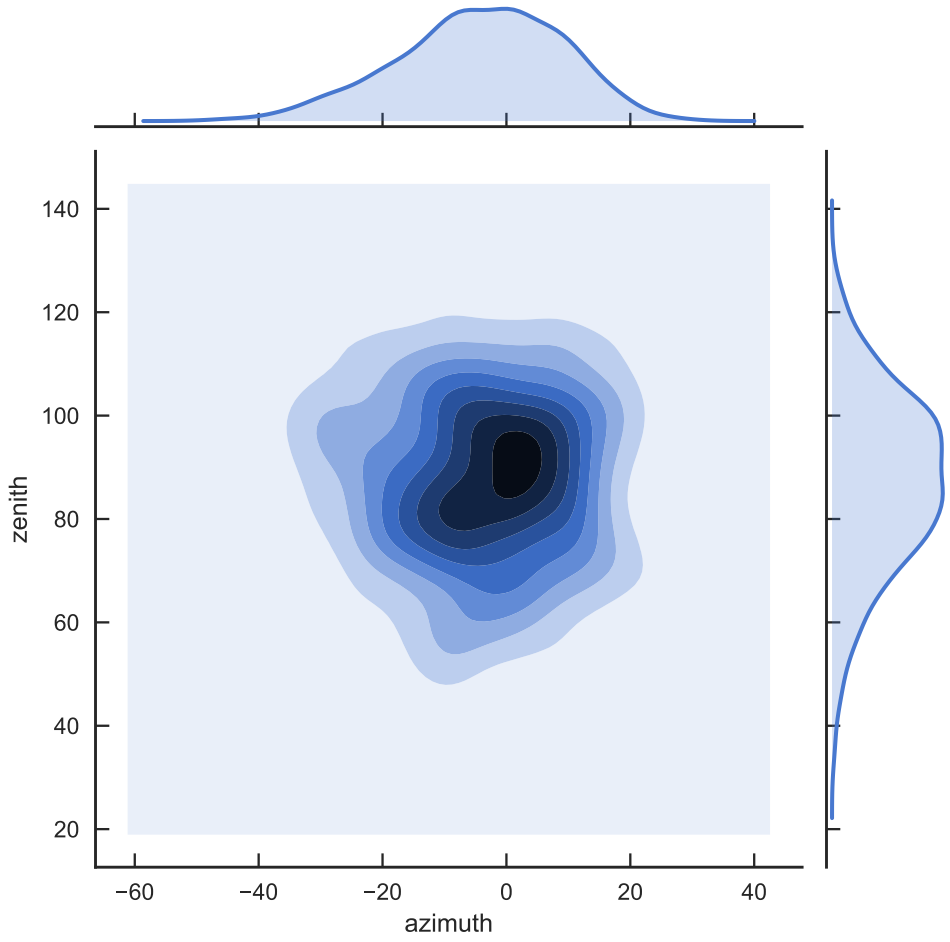


Spread and mean of last 5% of steps
used to initialize step 3

An example: step 3, error calculation

1. E=100 TeV, ice=spice-3.2 (latest), r=(0, 0, 300), θ=(90z, 0a), σ=0.0
2. DirectFit steps to the minimum
3. Generate probabilities across the parameter space



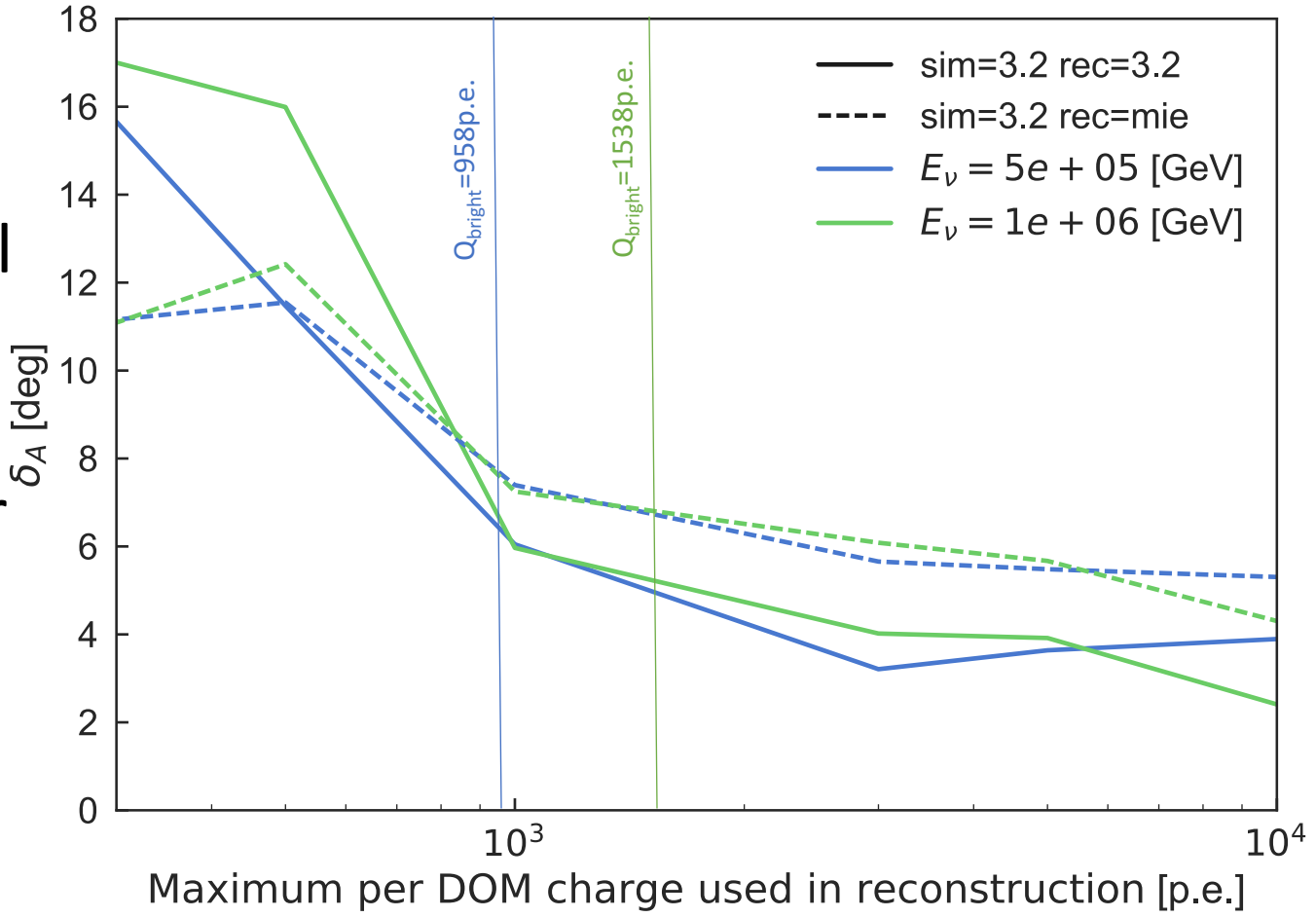


Effect of Q_{max} on angular resolution

Reconstruct simulated 500TeV and 1PeV cascade for a set of Q_{max}

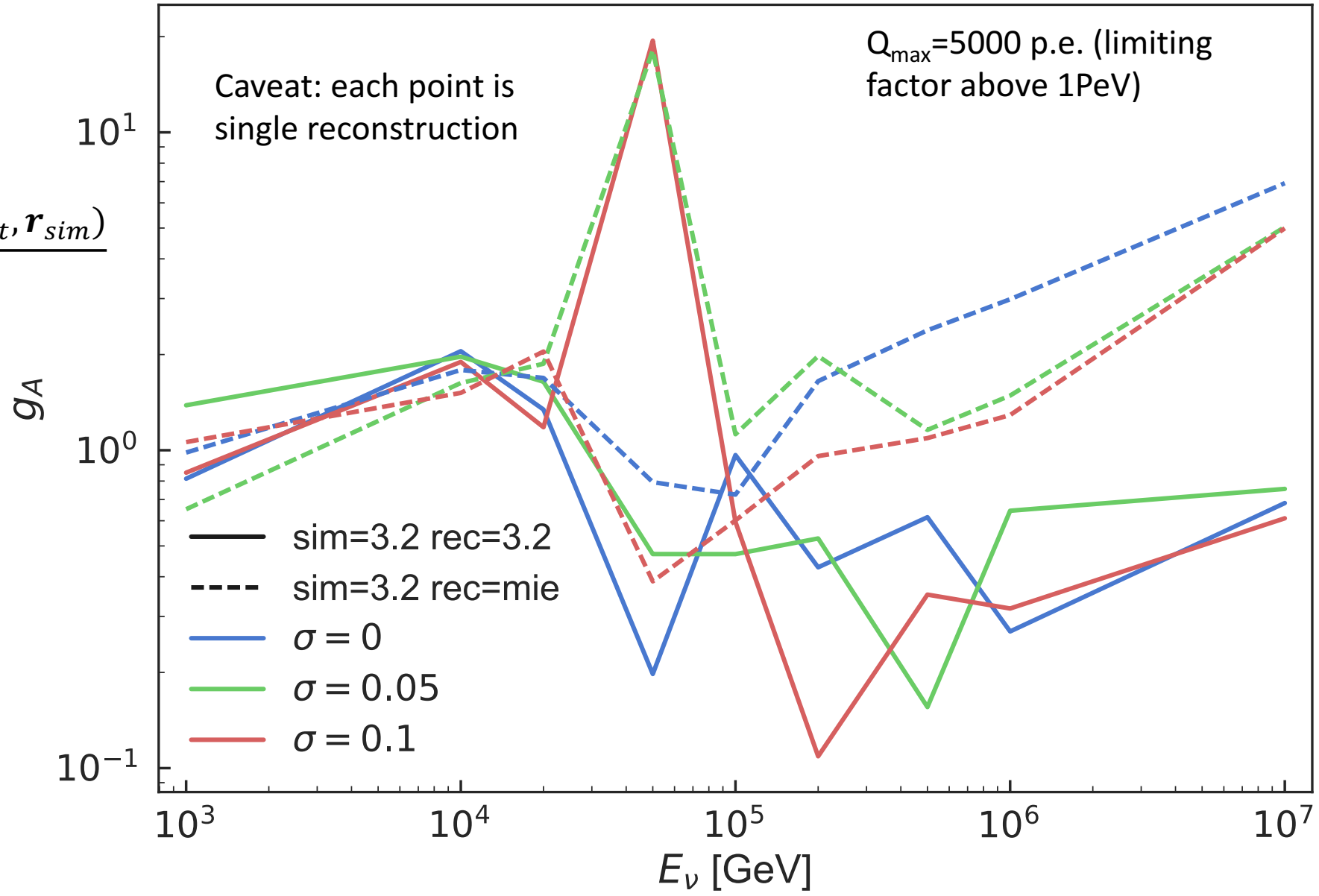
Tested with an identical sim-reco ice-model and a different ice-model (SPICE-mie)

Both show a trend towards a better angular resolution as more DOMs are included (increasing Q_{max})



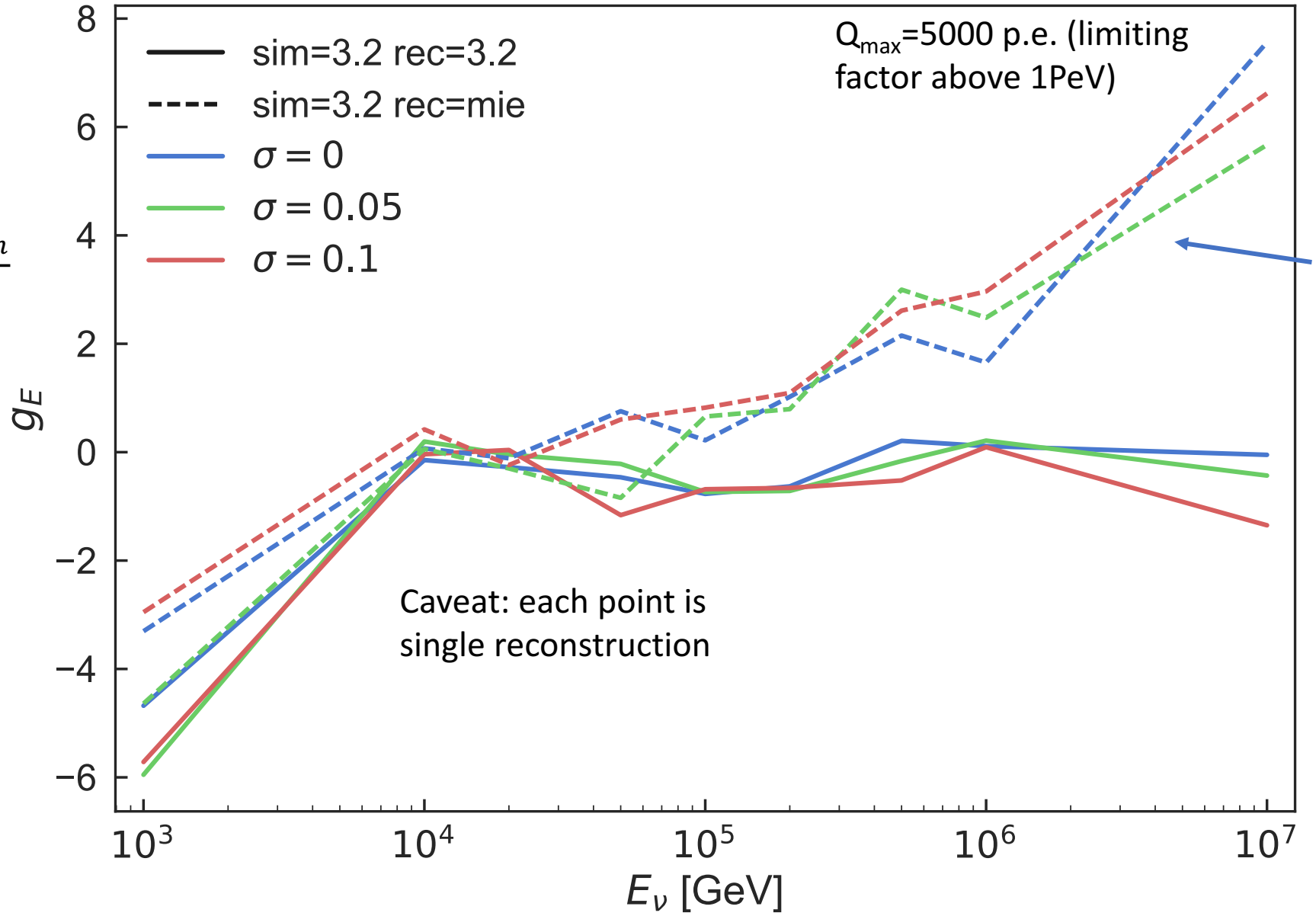
Angular pull vs energy and σ

$$g_A = \frac{inner(\boldsymbol{r}_{fit}, \boldsymbol{r}_{sim})}{\delta_A}$$



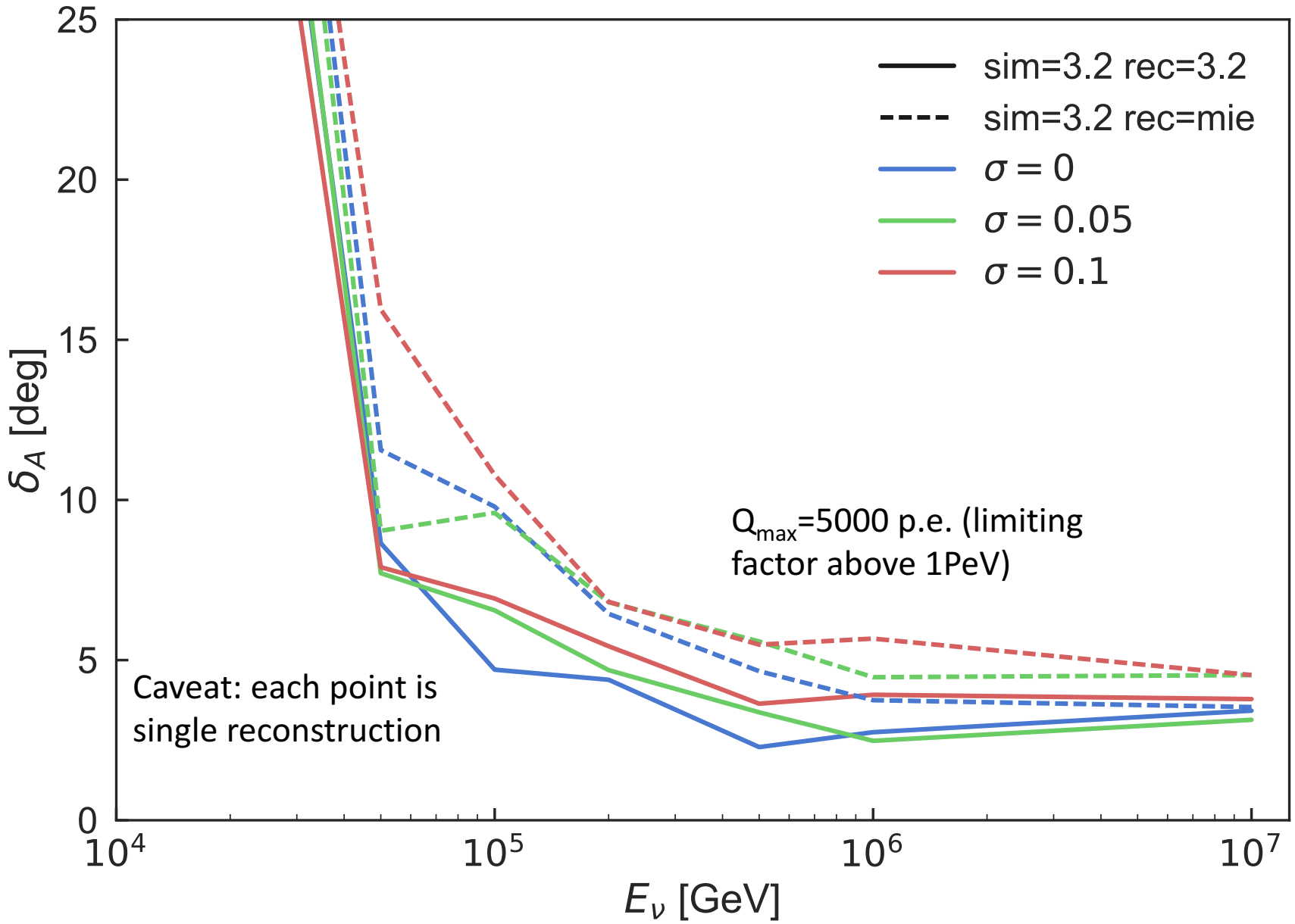
Energy pull vs energy and σ

$$g_E = \frac{E_{fit} - E_{sim}}{\delta_E}$$



Biased energy reconstruction with Mie

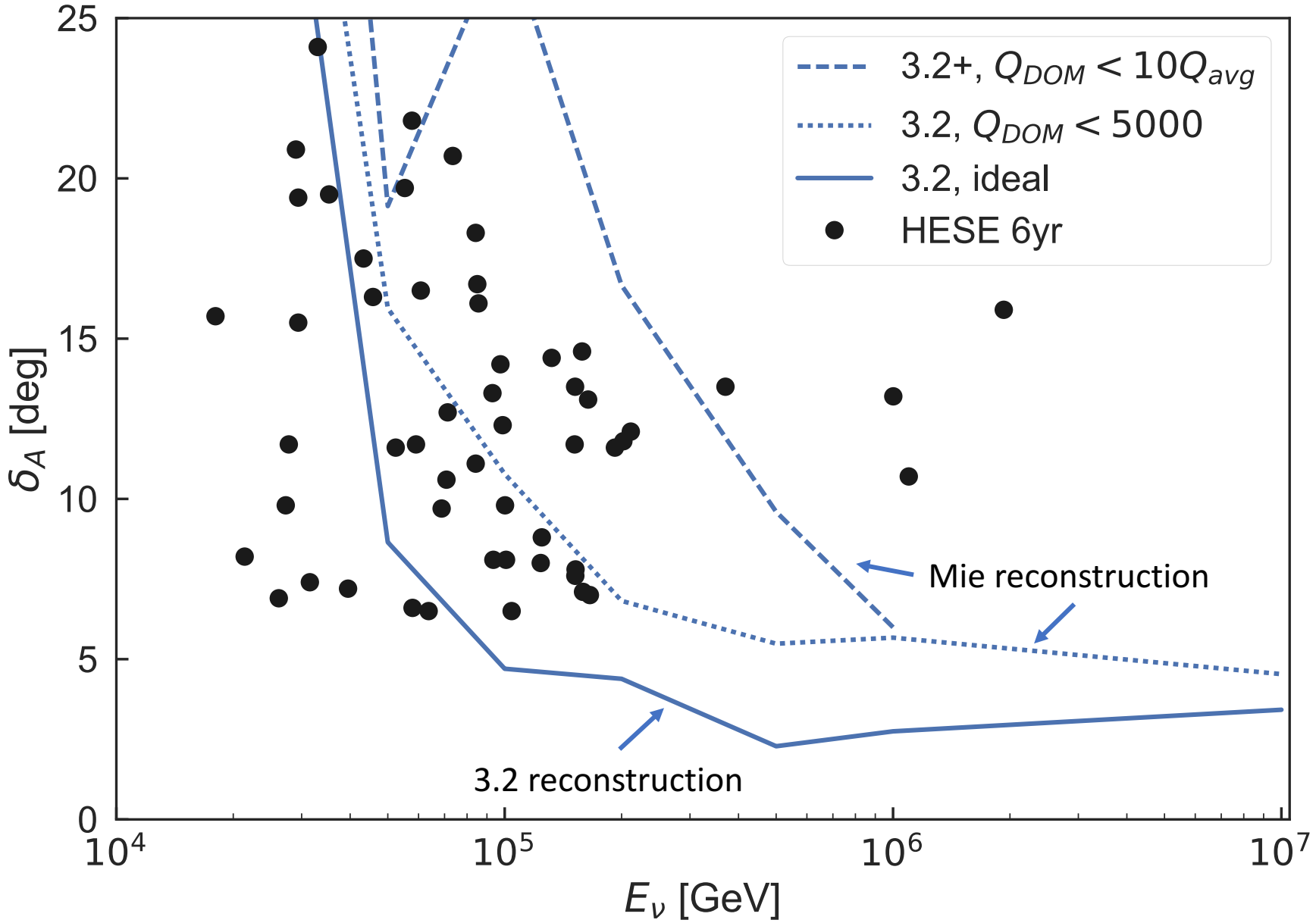
Angular resolution vs energy and σ



Reducing σ generally reduces δ_A (colors)

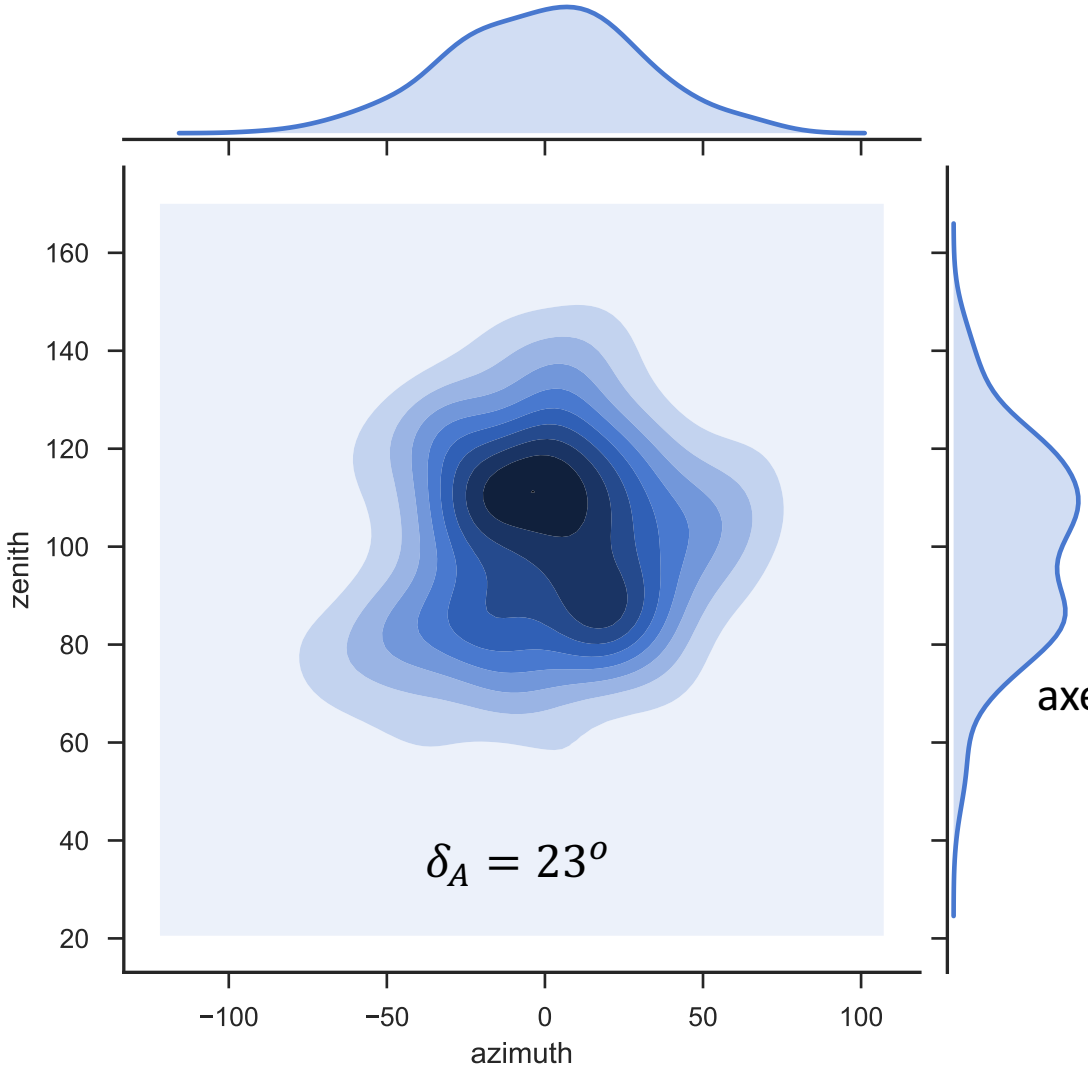
Correct knowledge of ice also reduces δ_A (dashed vs solid)

Combined effect on angular resolution



3.2+: Includes cable, DOM orientation

With more simulated photons



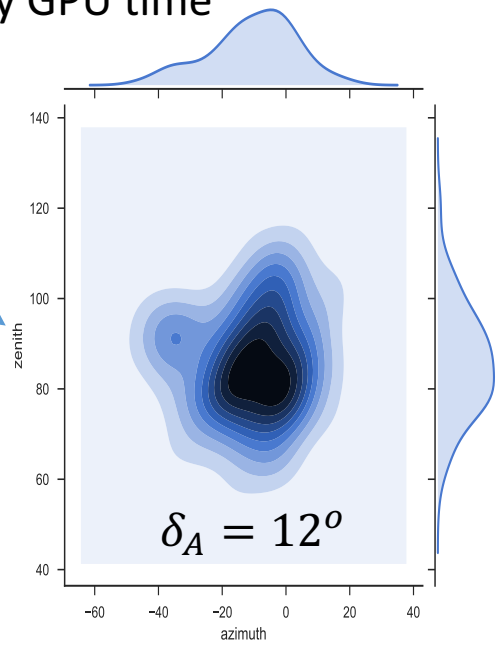
1x data statistics

Direct photon reconstruction mean statistical uncertainties in MC

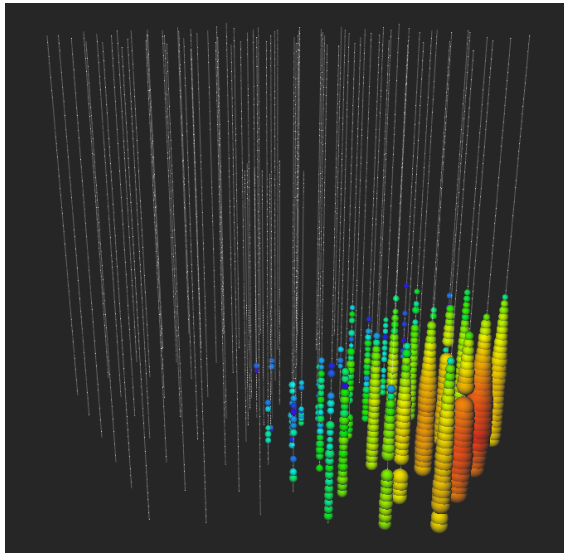
Increased photon statistics improves angular resolution even more!

Limited by GPU time

axes to scale

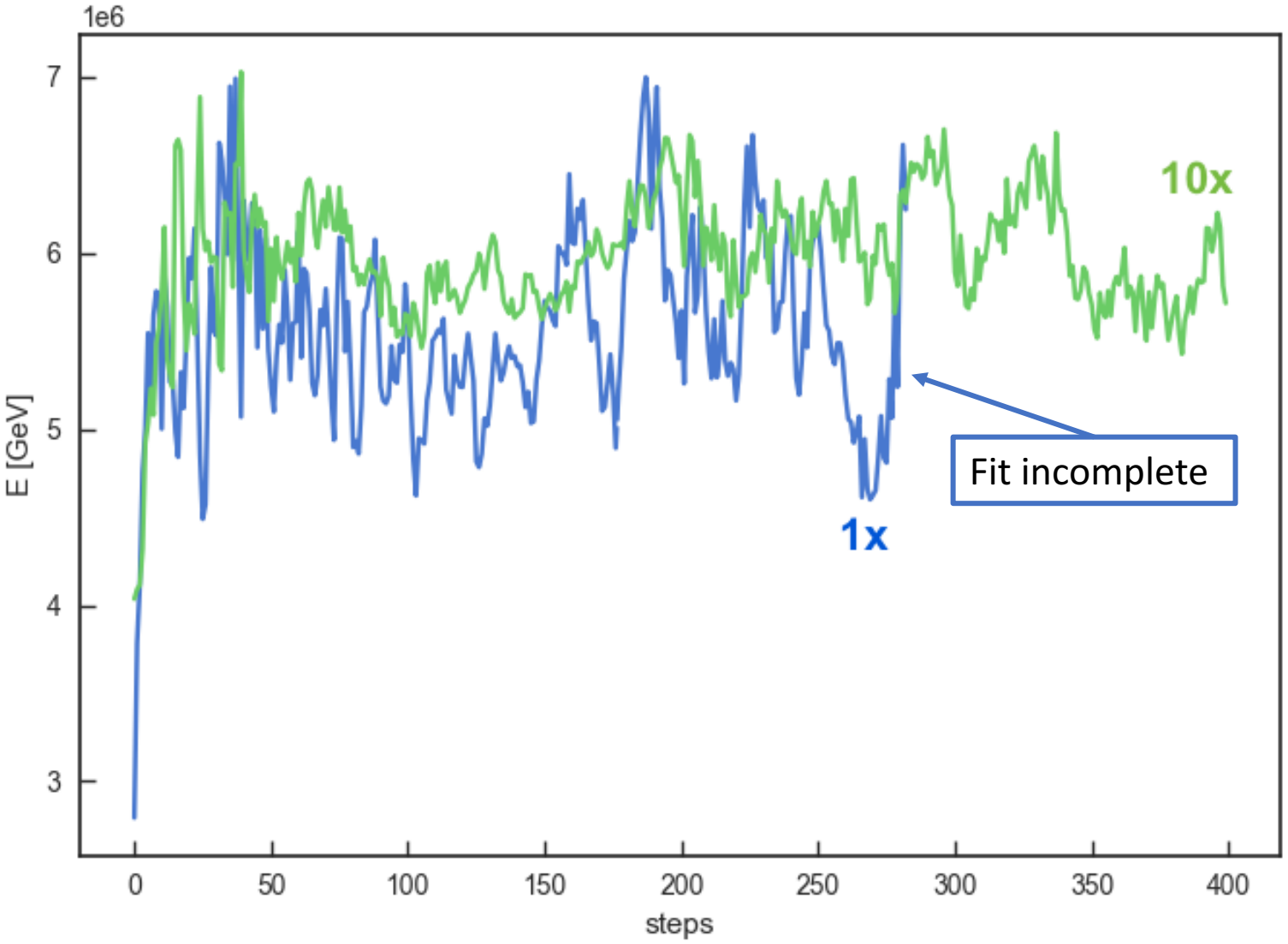


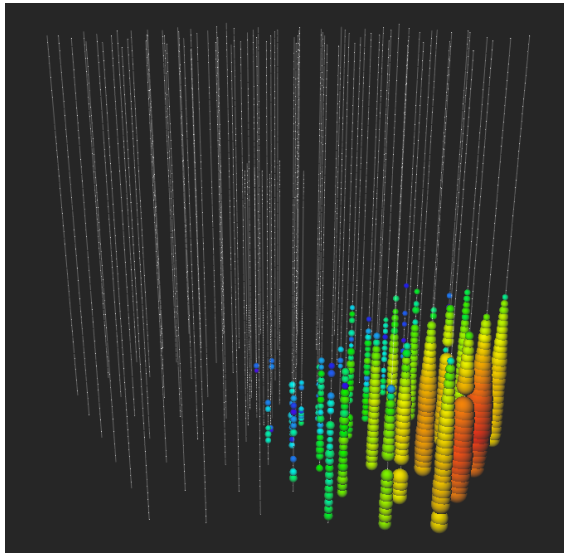
10x data statistics



Performed on new non-contained PeV cascade “Hydrangea”!

Improved energy reconstruction too!





Performed on new non-contained PeV cascade "Hydrangea"!

$$\frac{\delta_E}{E}: 8.3\% \rightarrow 3.6\%$$

Summary

Room to improve cascade reconstruction

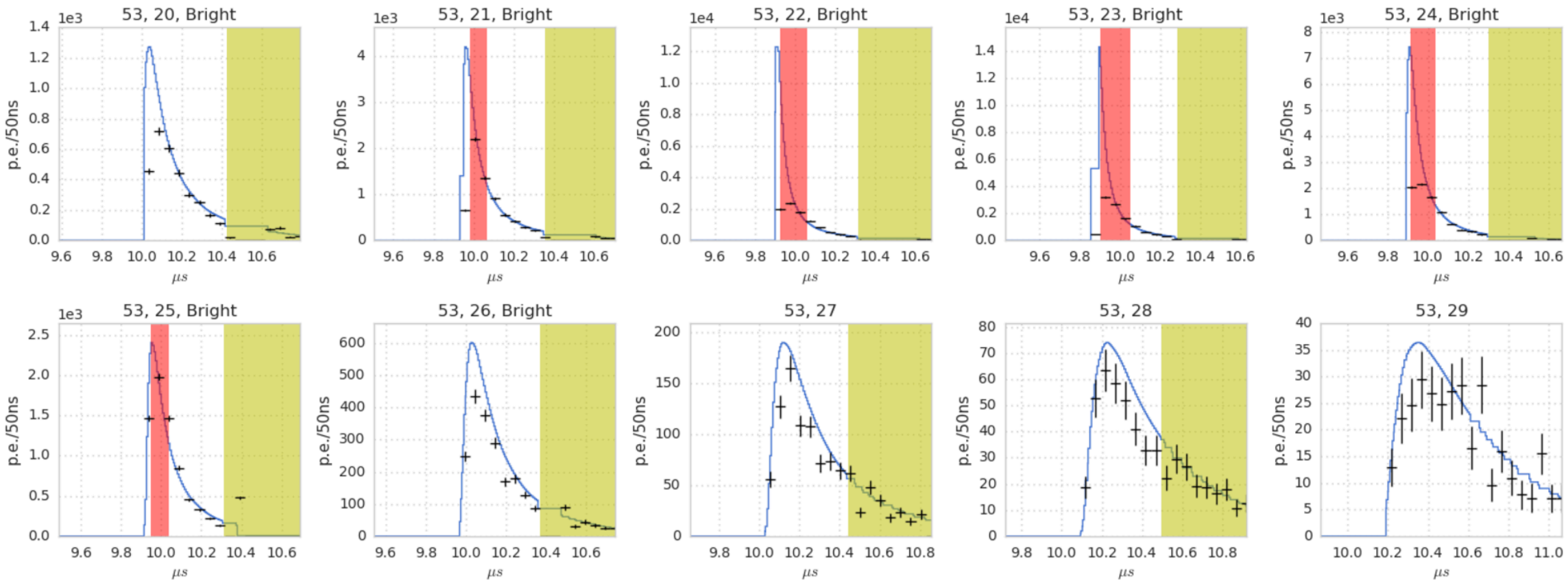
Currently affected by

1. Bright DOM exclusions
2. Ice-model and ice-model uncertainty

Even more improvement with increased direct photon statistics but this may prove to be impractical.

Backups

Bert



http://icecube.wisc.edu/~tyuan/share/fig/waveform/hese_exclusions/mjd55782.52/str53.png