Energy and X_{max} reconstruction of Cosmic Rays using the Surface Radio Array of IceCube

Roxanne Turcotte *for the IceCube Collaboration*

At the ARENA Conference 2022

09.06.2022

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 802729).

Prototype station

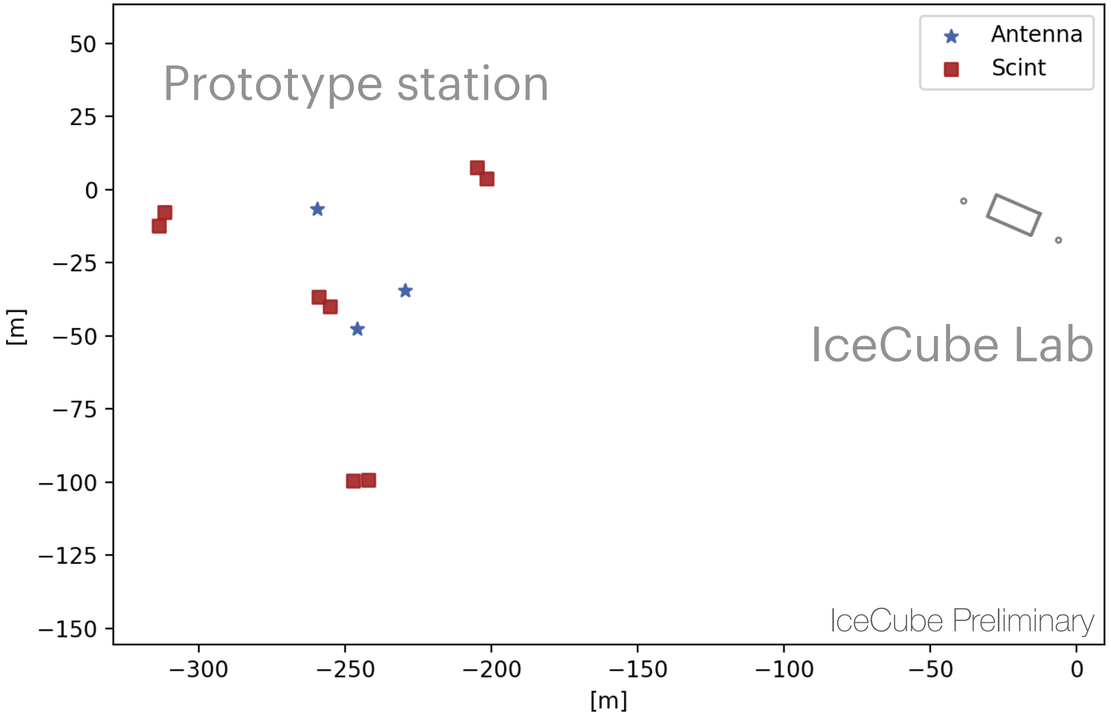
Prototype station deployed in January 2020 over the IceCube detector

Operating ~continuously (background and event data)

Air showers recorded!

*Elevated **scintillator** from the prototype station*

*Elevated **antenna** from the prototype station*



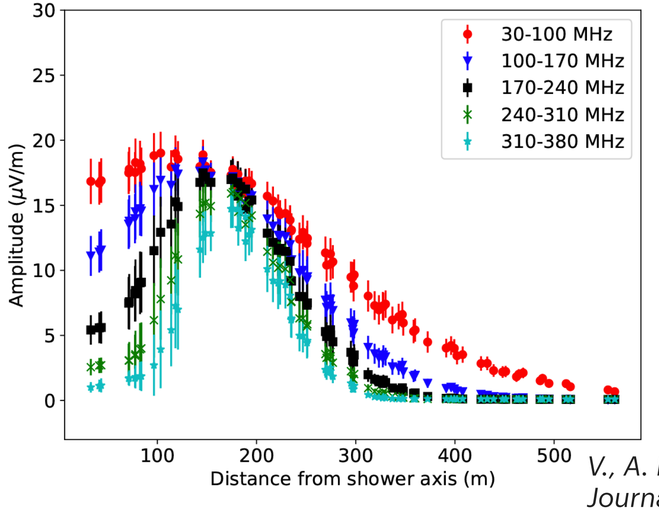
roxanne.turcotte@kit.edu

Outline

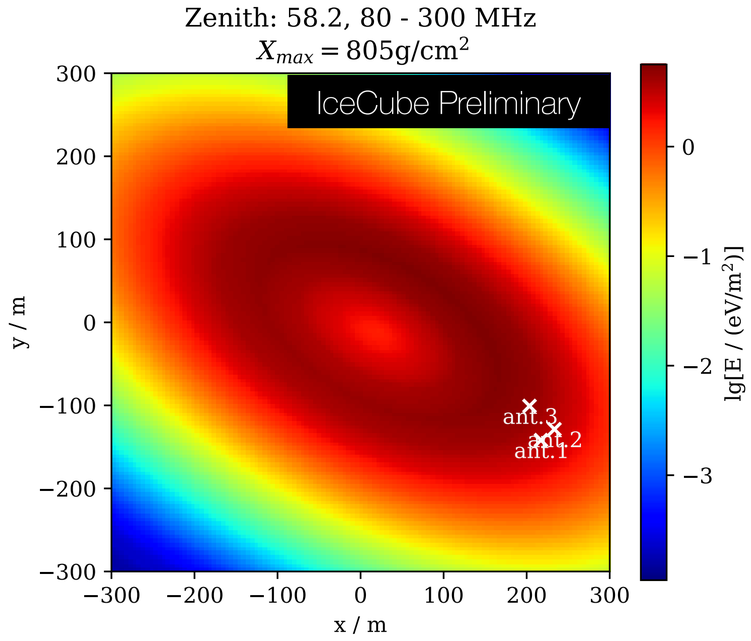
- Implement a reconstruction method inspired by the technique developed by LOFAR *(Buitink, S., et al. Physical Review D 90.8 (2014): 082003.)*
- Try a similar model on our prototype with higher frequency band
- Work towards a log-likelihood
  - Start by a PDF study
- Results & Improvement prospects

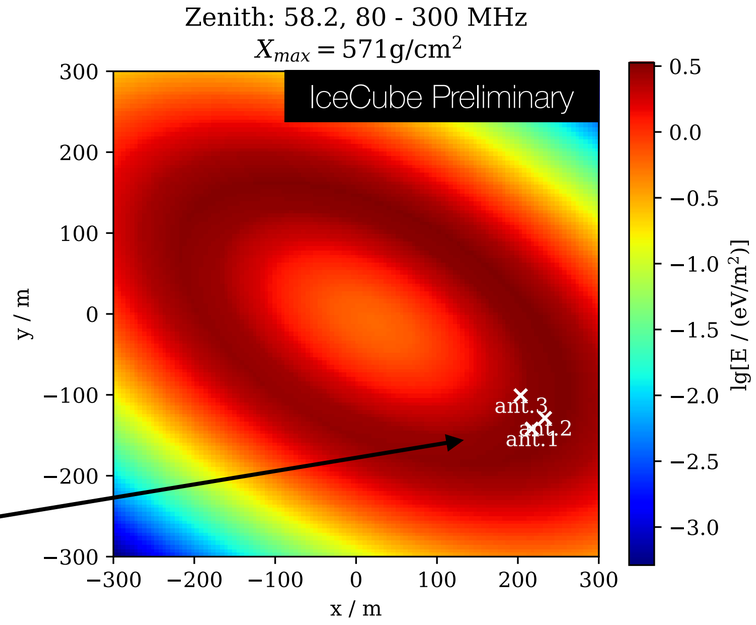
The technique

- Consist of comparing data to a bunch of simulations
- The simulations intrinsically differ because of the shower-to-shower fluctuations which affect the X_{max}
- This method gives us an estimate on:
  - X_{max} via the shape of the footprint
  - the energy via « scaling » the footprint
- Expected better reconstruction —> higher frequency band 70-350MHz (other experiments use ~30-80MHz)
- Test on prototype station -> Only 3 antennas!



V., A. Balagopal et al. The European Physical Journal C vol. 78, no. 2 (Feb. 2018)

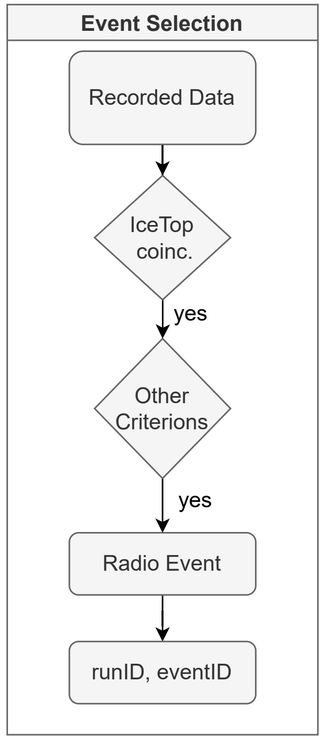




We cover only a very small area of the footprint

roxanne.turcotte@kit.edu

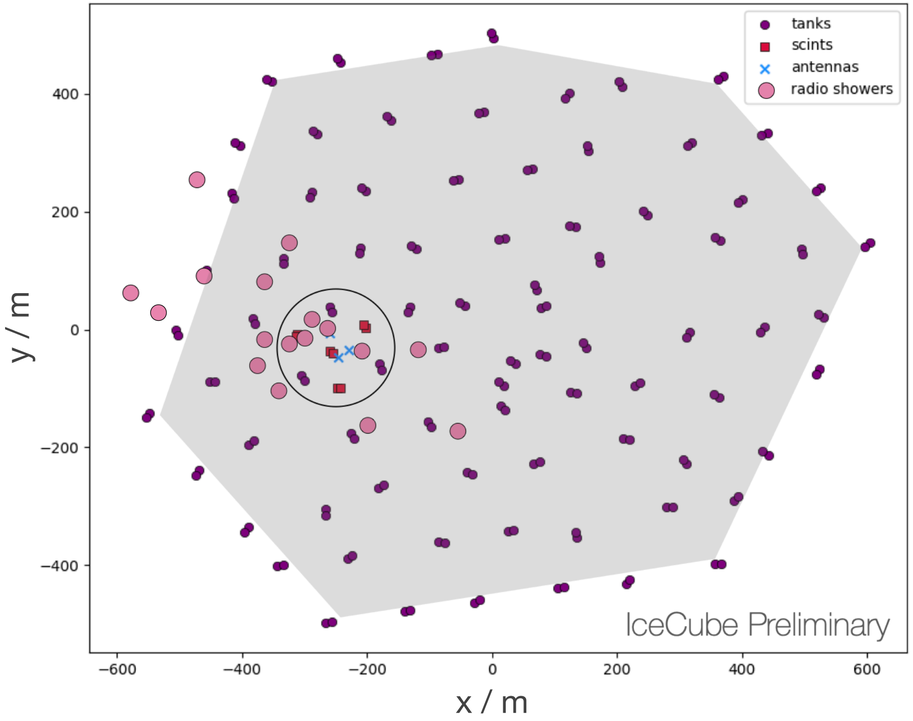
Event selection

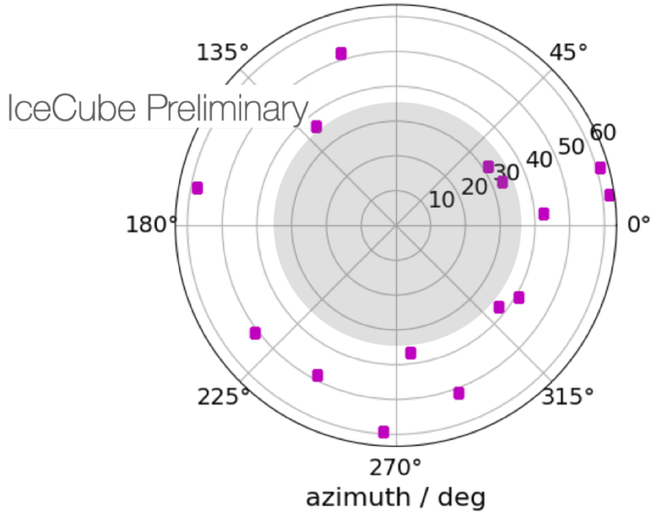


Subset of events

- 16 selected events
- From ~100 PeV to ~350 PeV
- Recorded between
  - 24.10.2020 - 26.01.2021
- With associated scintillator & IceTop data
- No IceTop quality cut and most of the events are outside the calibrated space

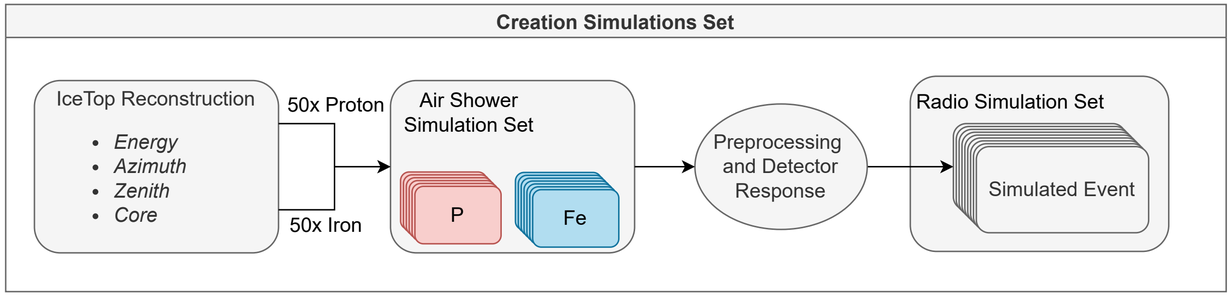
IceTop (*Laputop*) reconstruction





roxanne.turcotte@kit.edu

Creation of a simulation set per event

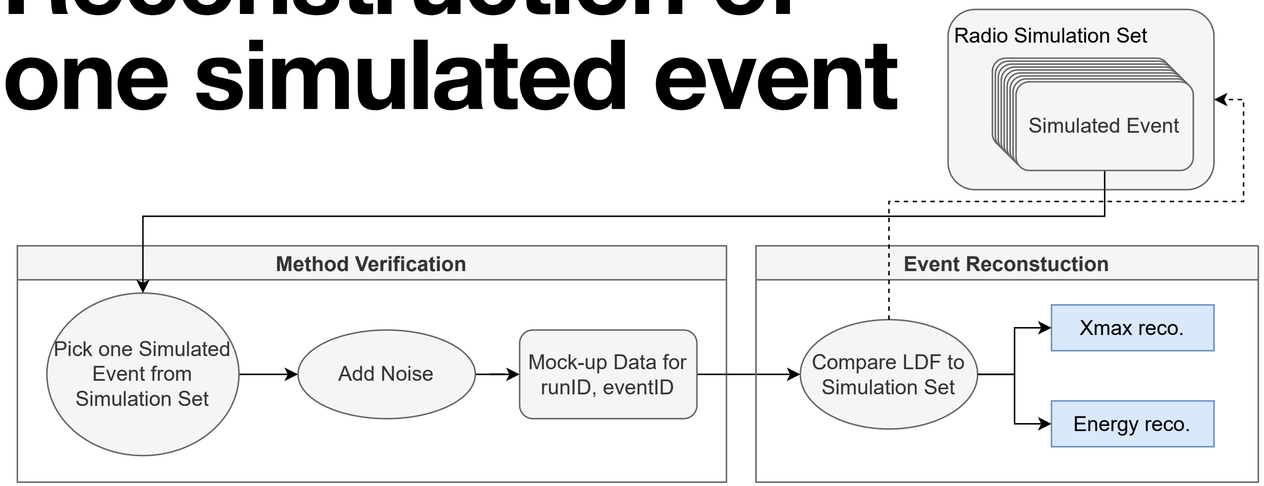


- Create a set of air shower Monte-Carlo simulations for each event
- Use IceTop reconstruction as input for CORSIKA/CoREAS
- IceTop accuracy for our events:
  - Energy -> not characterised
  - Azimuth ~ 0.3°
  - Zenith ~ 0.2°-0.4°
  - Core ~ 2m-8m

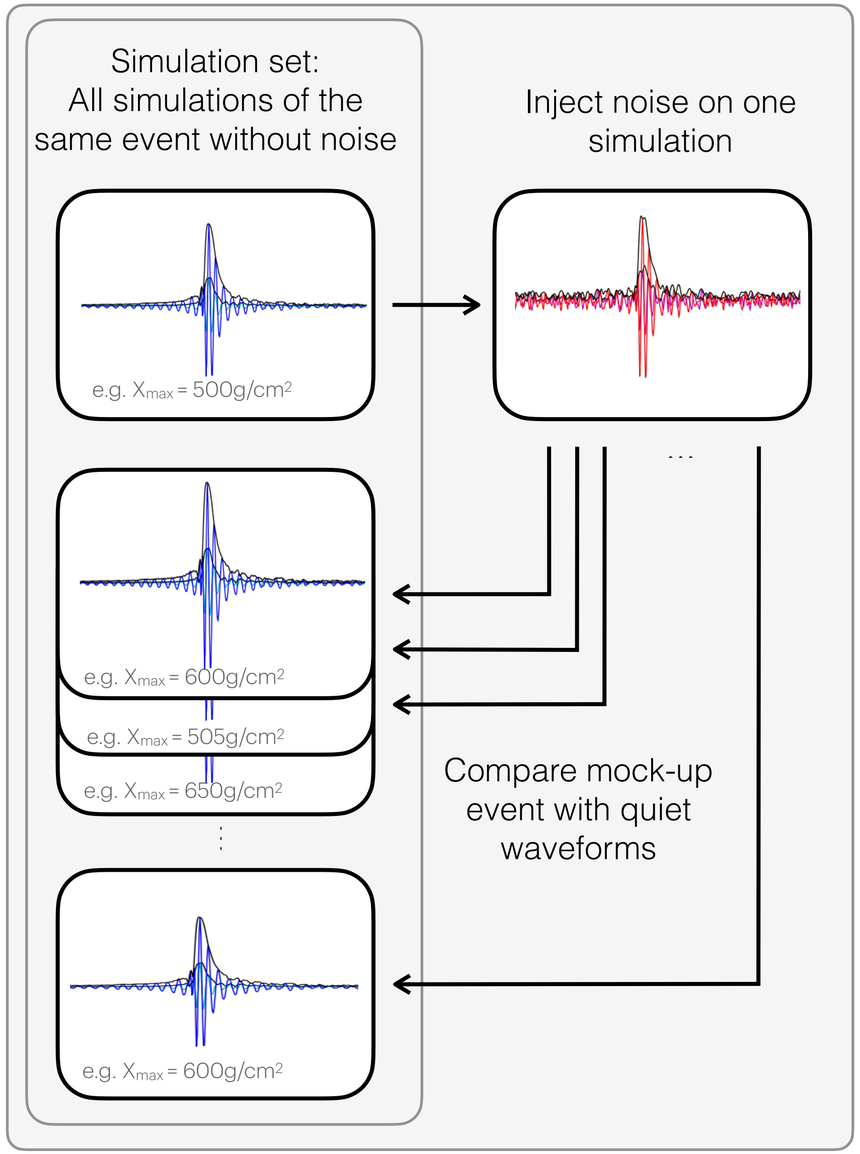
- Each set has 100 simulations (50 H/50 Fe)
- The electronic response is added to the simulations —> Forward folding
- Use the star shape pattern (Coleman, A. arXiv: 2205.02258) and interpolate for the 3 antennas of the prototype station
- Waveforms filtered between 80-300MHz

roxanne.turcotte@kit.edu

Reconstruction of one simulated event



- Pick one simulation in the specific simulation set of one event
- Inject noise (real noise from prototype station) for the creation of the mock-up waveforms
  - *Noisy waveforms*
- Compare it to all the other simulations (with different X_{max}) to find the best fit
  - *Quiet waveforms*



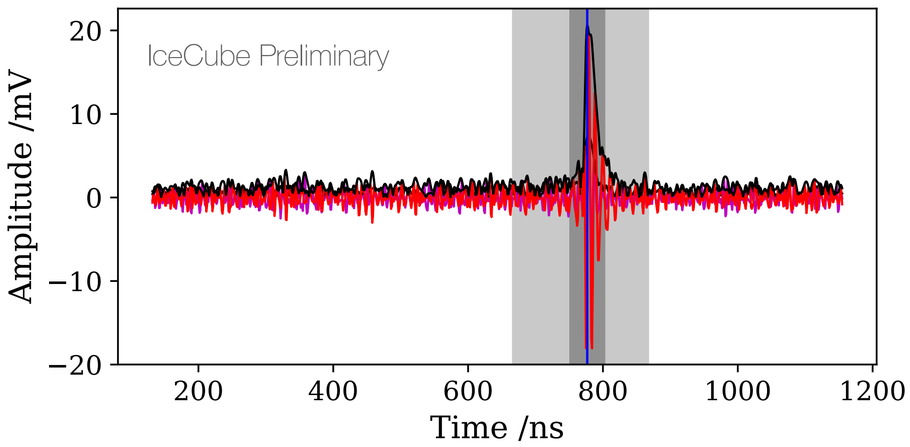
roxanne.turcotte@kit.edu

Finding the pulse and defining the observable

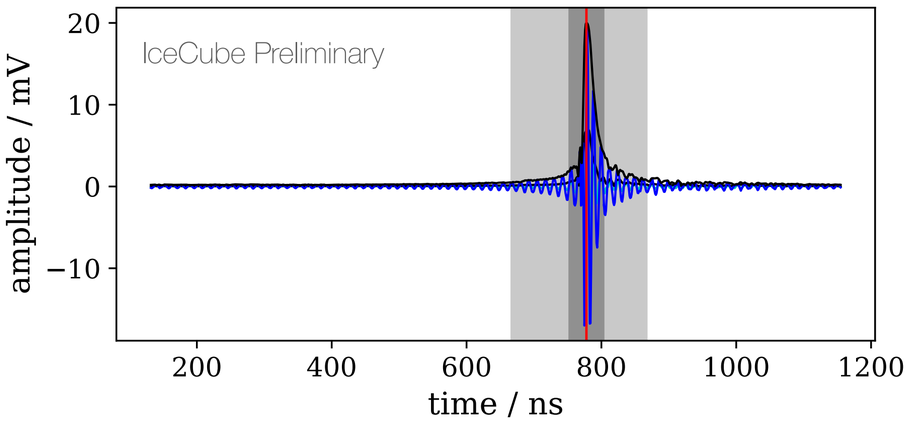
- Find pulses - same for quiet/noisy :
  - Define a 200ns window around where the pulse is expected
  - For each antenna, add the two channels together
  - Find maximum
  - Make a symmetrical window around it
  - Calculate <u>average power of the Hilbert envelope</u> of each channels individually
  - —> reconstruction in the « data-space"!

$$P = \frac{1}{\Omega \cdot t_{\text{tot}}} \int_{-\Pi}^{\Pi} H^2(t)dt$$

CoREAS simulation — noisy waveform

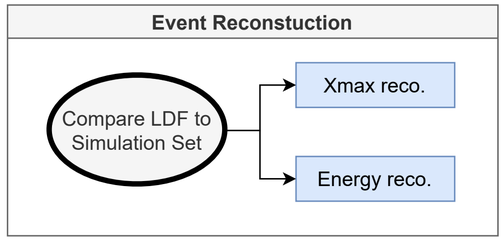


CoREAS simulation — quiet waveform



roxanne.turcotte@kit.edu

Compare LDF - Minimisation



- Treat the two channels of the three antennas individually —> 6 - 1 d.o.f
- Minimise using χ^2
- Find optimal f - scaling factor for energy

χ^2 - Power Hilbert envelope

$$\chi^2_{\text{power}} = \sum_{i}^{\text{nCh}=6} \left(\frac{\left(P_i - \sigma_i\right) - f^2 \cdot P_{\text{sim}_i}}{\sigma_i} \right)^2$$

where,

σ_i = noise power in waveform i

P_i = average power in the pulse of noisy waveform i

P_{sim_i} = average power in the pulse of quiet waveform i

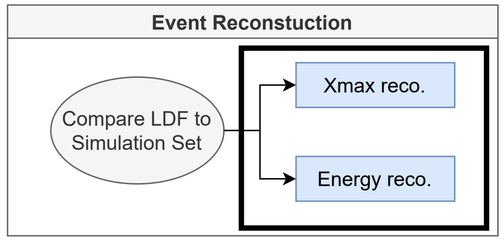
f = free variable to fit

$$E_{\text{CR}_{\text{reco}}} \approx f \cdot E_{\text{sim}}$$

based on : arXiv:2005.13441v2 [astro-ph.HE] 29 Sep 2020

roxanne.turcotte@kit.edu

Example with mock-up event

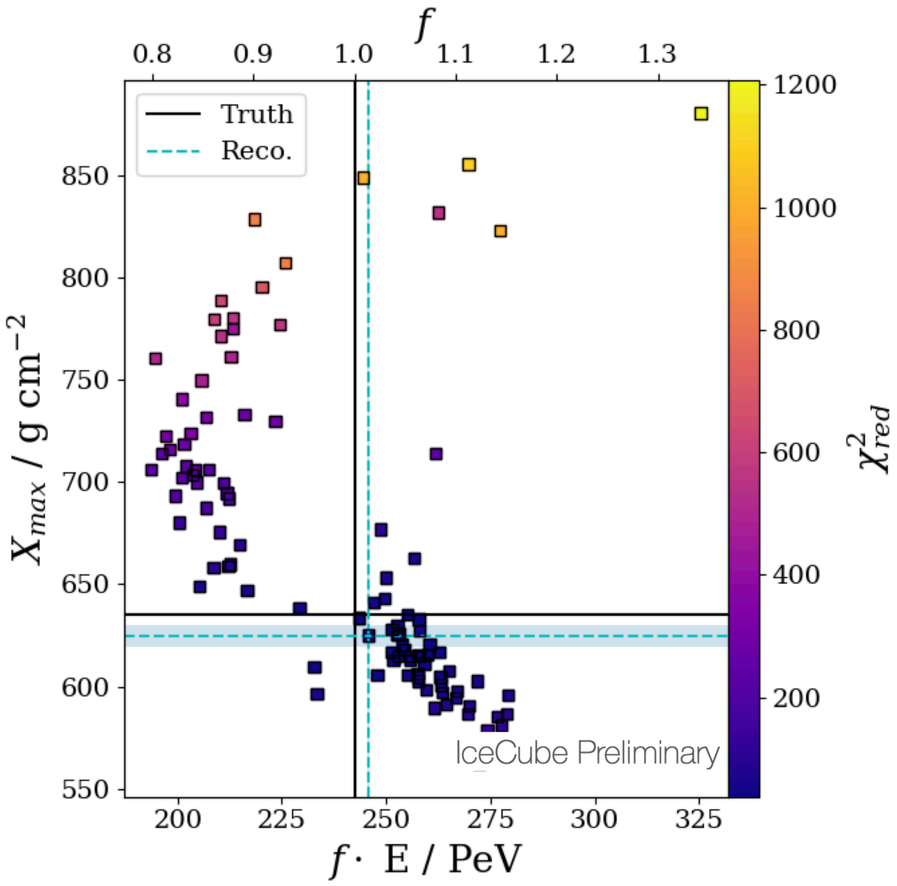


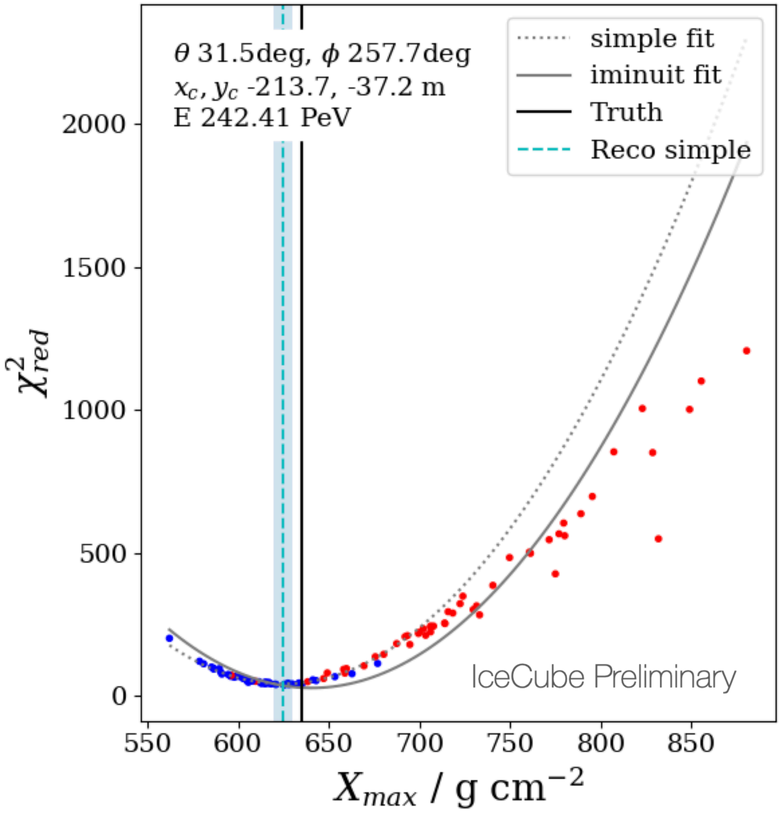
Xmax reconstructed value

Fit a parabola on the 50 smallest χ^2 values

Energy reconstructed value

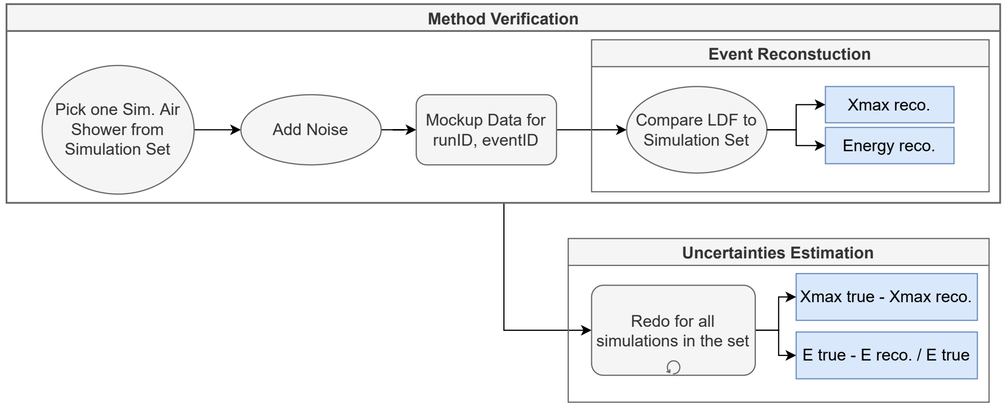
Take f from the lowest χ^2 value



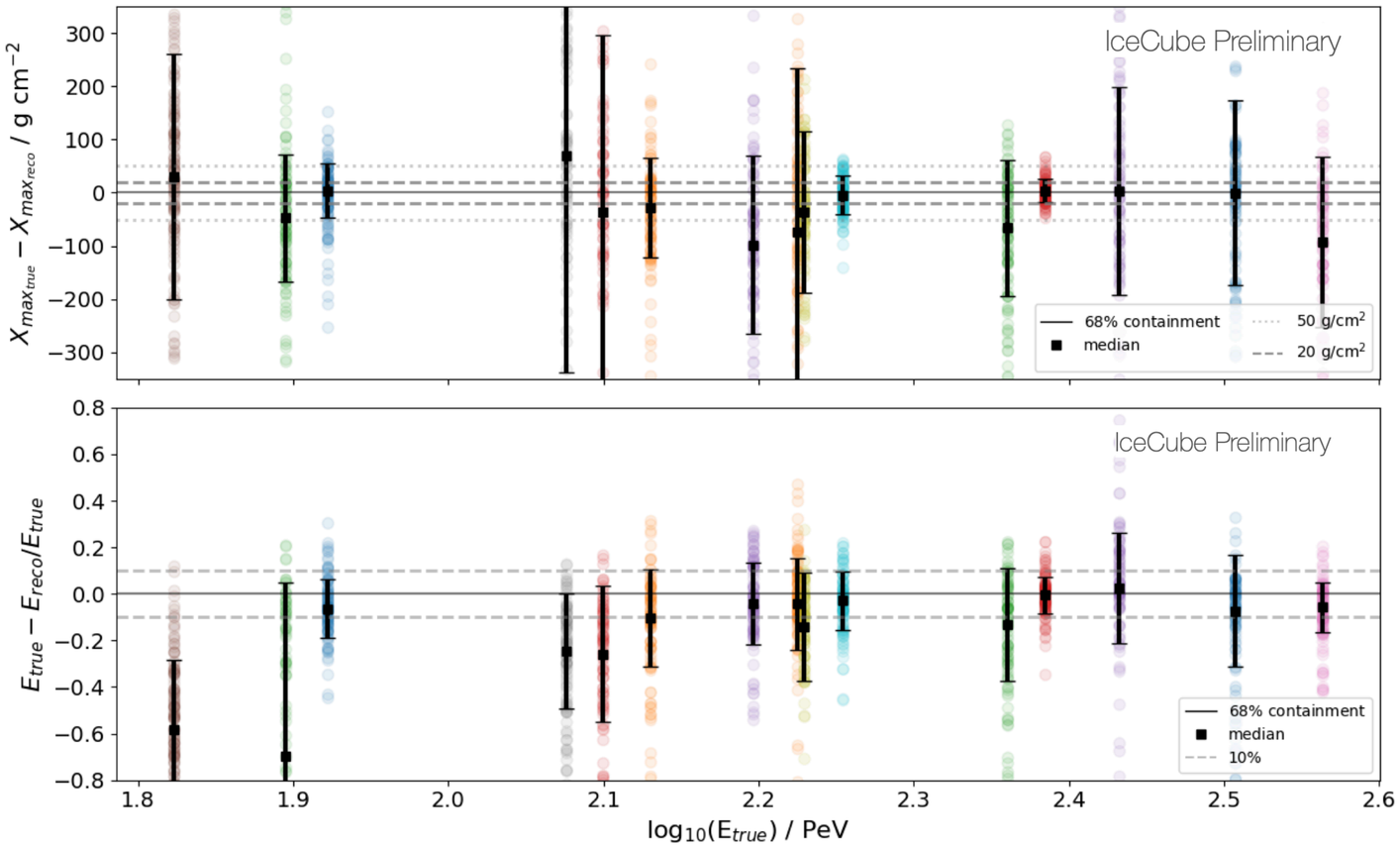


roxanne.turcotte@kit.edu

Estimation of the accuracy of reconstruction per event



- The previous step is redone for all the simulations in the set -> estimated resolution per event
- We do it for all the events of the subset
- Bias in lower energy
- Characteristics of event well reconstructed under investigation
  - No clear dependance on zenith/azimuth/energy
  - Probably lying on the Cherenkov-like ring
  - Those having a core outside the IceTop array are in general badly reconstructed
  - Reminder that most of the events are outside the optimal IceTop reconstruction
  - —> Better quality cut needed in the future!



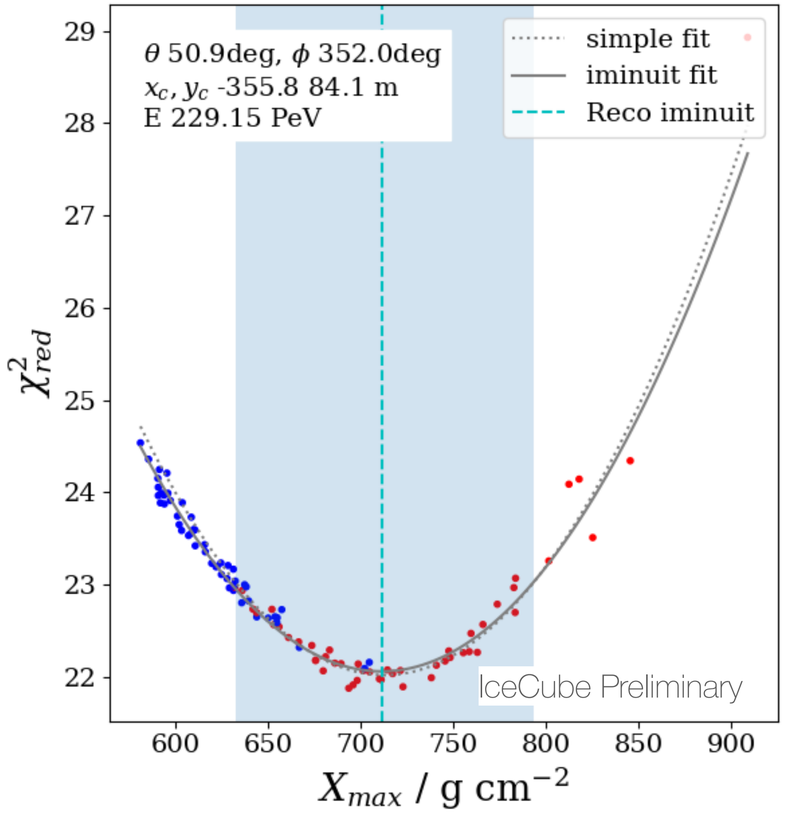
roxanne.turcotte@kit.edu

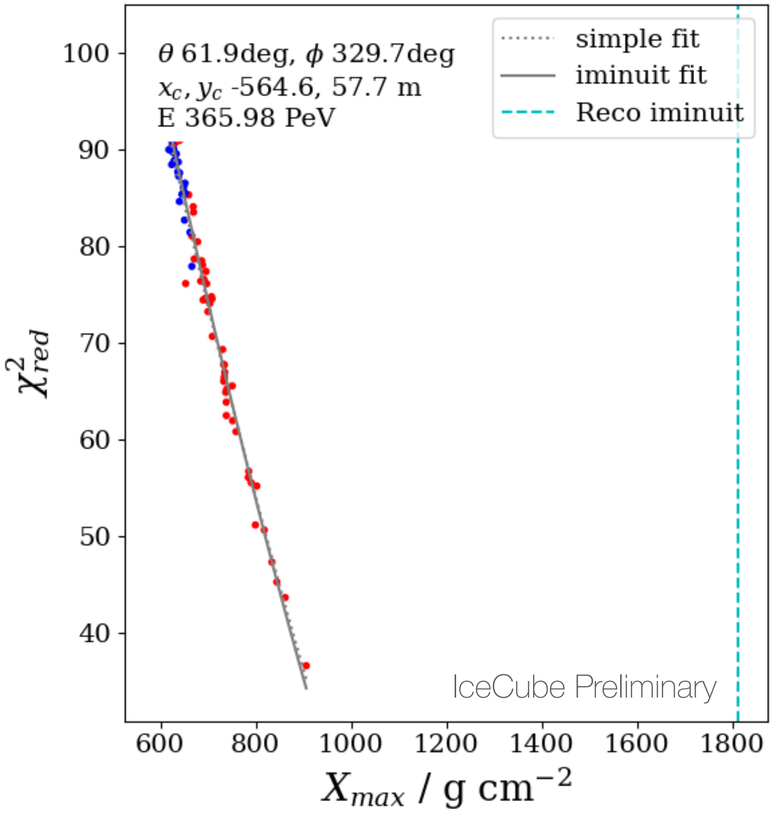
Real measurements - 50ns window χ^2

We see a similar trend in the real data.

About half of the events are reconstructable or very close.

The rest have most likely and initial reconstruction too far from the truth or the technique is not optimal (?) also the chi2 values are too high

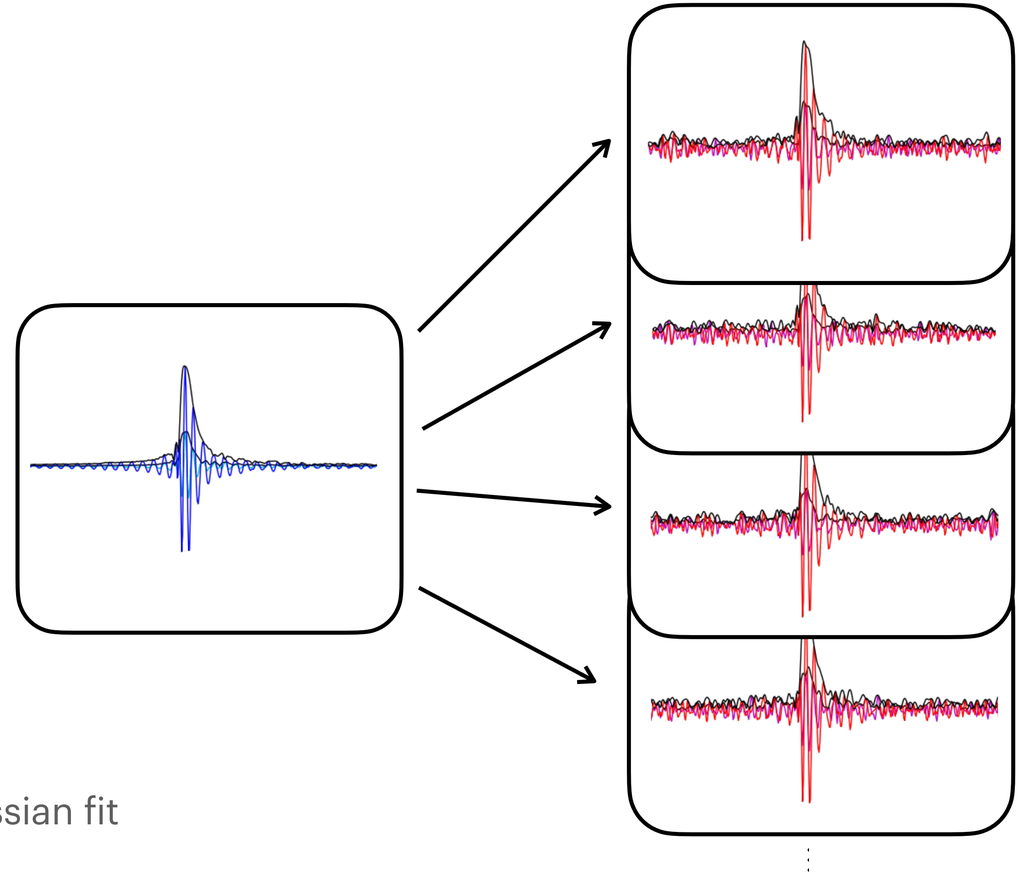


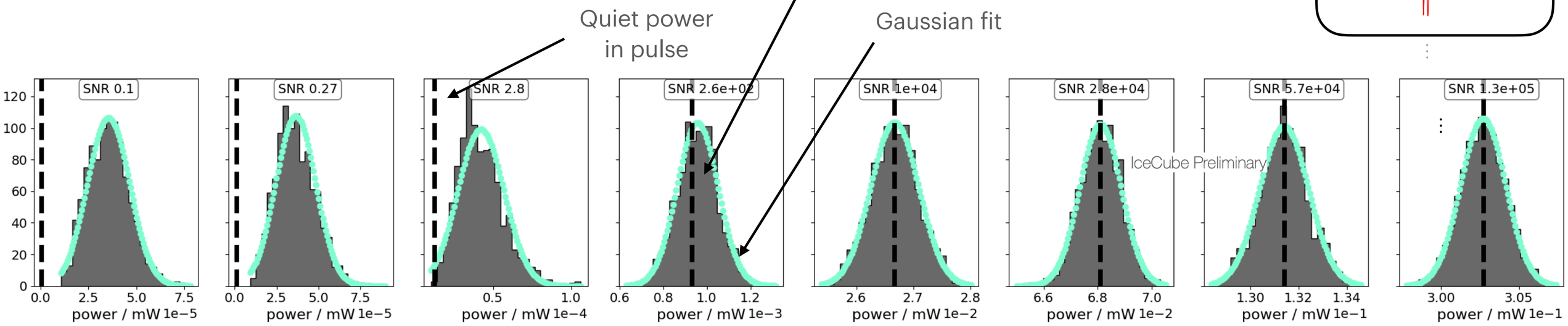


roxanne.turcotte@kit.edu

Investigating the PDF

- Studying the effect of the noise on the power
- Take all the pulses of the antennas in a star shaped layout —> many different pulses strength
- Inject 1000x some ideal noise on each pulse
  - ideal noise = cane expectation + thermal noise
- Each histogram represent one pulse



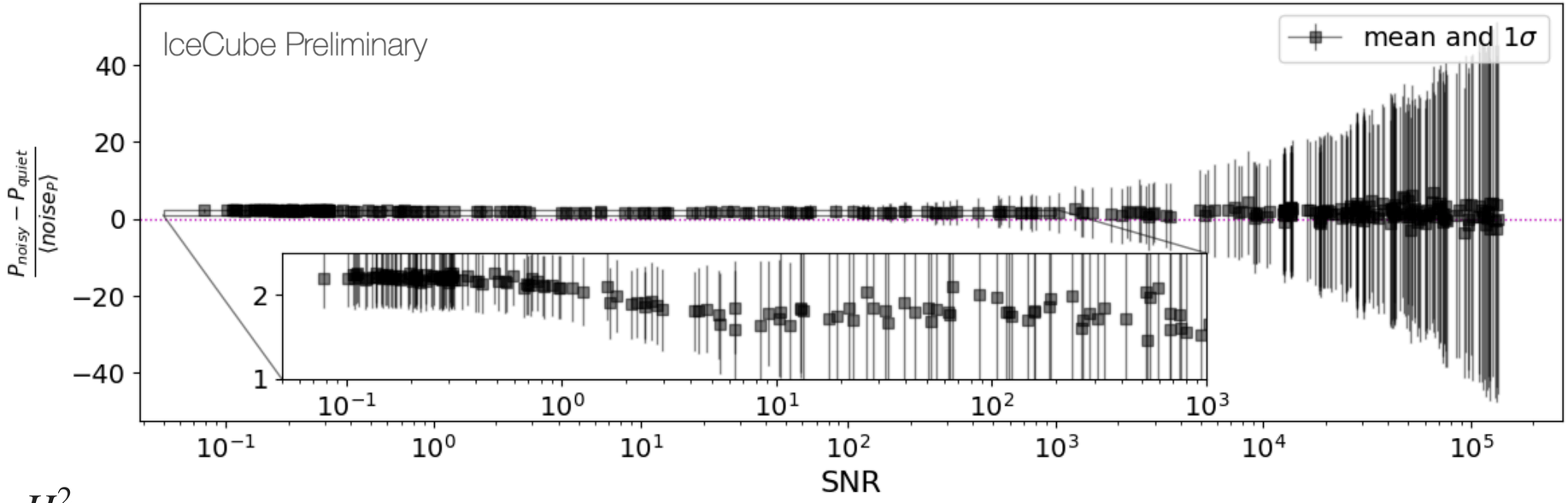


Increasing SNR

roxanne.turcotte@kit.edu

Behaviour of the Gaussian distributions

Gaussian mean with ideal noise (Cane + thermal from LNA)



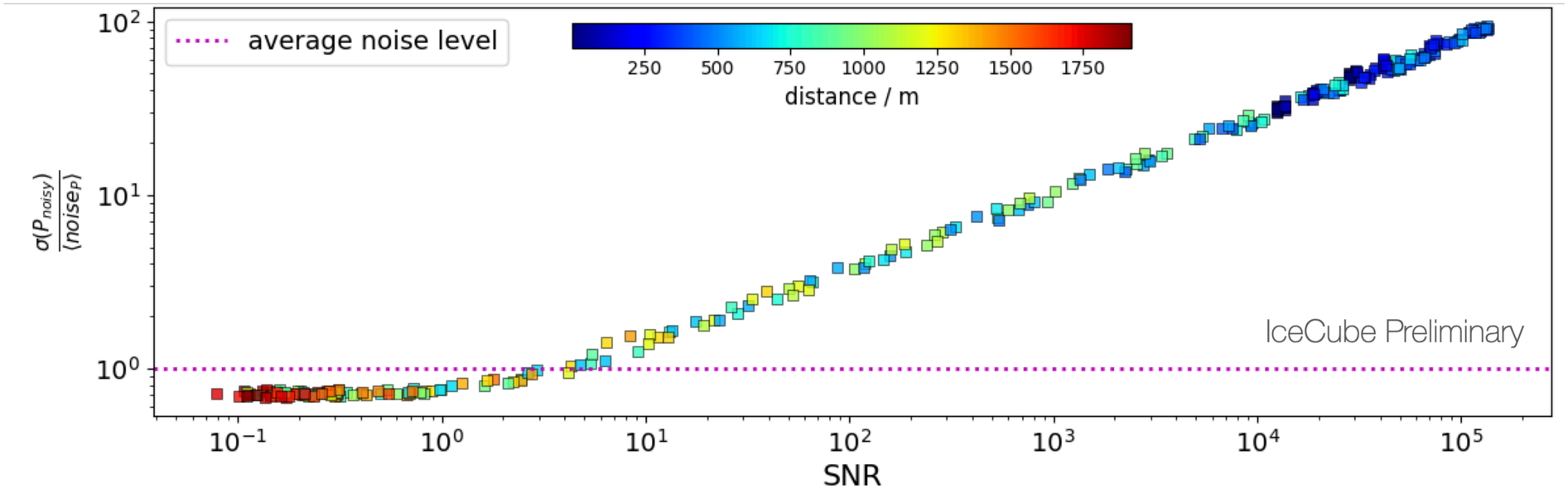
- $SNR = \frac{H_{\text{max}}^2}{\left\langle \text{RMS}_{\text{noise}} \right\rangle^2}$
- Where,
  - $\left\langle \text{noise}_{\text{RMS}} \right\rangle^2$ —> RMS of the noise averaged over all waveforms
  - $\left\langle \text{noise}_{\text{P}} \right\rangle^2$ —> same using power instead of RMS
  - Calculated using a 128ns sub-window method

$$P'_{\text{sim}} = P_{\text{sim}} + c \cdot \left(1 - d \cdot e^{-\beta x^{-\alpha}}\right)$$

roxanne.turcotte@kit.edu

Behaviour of the standard deviations

Gaussian std deviation with ideal noise (Cane + thermal from LNA)

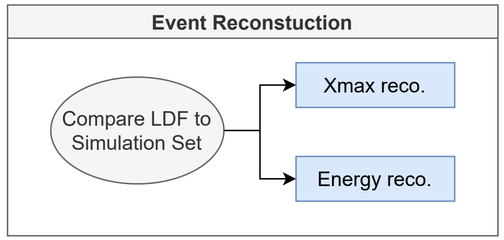


- The standard deviation of the noisy pulses is constant below $SNR \lesssim 1$
- Then increase as a function of the pulse strength
- Parametrised with a broken power law

$$\sigma\left(P_{\text{sim}}\right) = a \cdot \left(1 + \left(\frac{P_{\text{sim}}}{P_0}\right)^k\right)^{\frac{b}{k}}$$

roxanne.turcotte@kit.edu

Migrating from χ^2 to LLH



- Log-likelihood with the average power of Hilbert Envelope
  - Assuming Gaussian distribution everywhere
  - Using the functions $\sigma\left(f, P_{\text{sim}_i}\right)$ and P'_{sim} derived earlier

From

χ^2 - Power Hilbert envelope

$$\chi^2_{\text{power}} = \sum_{i}^{\text{nCh}=6} \left(\frac{\left(P_i - \sigma_i\right) - f^2 \cdot P_{\text{sim}_i}}{\sigma_i} \right)^2$$

To

LLH - Power Hilbert envelope

$$-2ln\mathscr{L} = \sum_{i=0}^{\text{nCh}=6} \left(2\sqrt{2\pi} \ln \sigma\left(f, P_{\text{sim}_i}\right) + \left(\frac{\left(P_i - f^2 P'_{\text{sim}_i}\right)^2}{\sigma\left(f, P_{\text{sim}_i}\right)^2} \right) \right)$$

where,

$$\sigma\left(f, P_{\text{sim}_i}\right) = a \cdot \left(1 + \left(\frac{f^2 P_{\text{sim}_i}}{P_0} \right)^k \right)^{\frac{b}{k}}$$

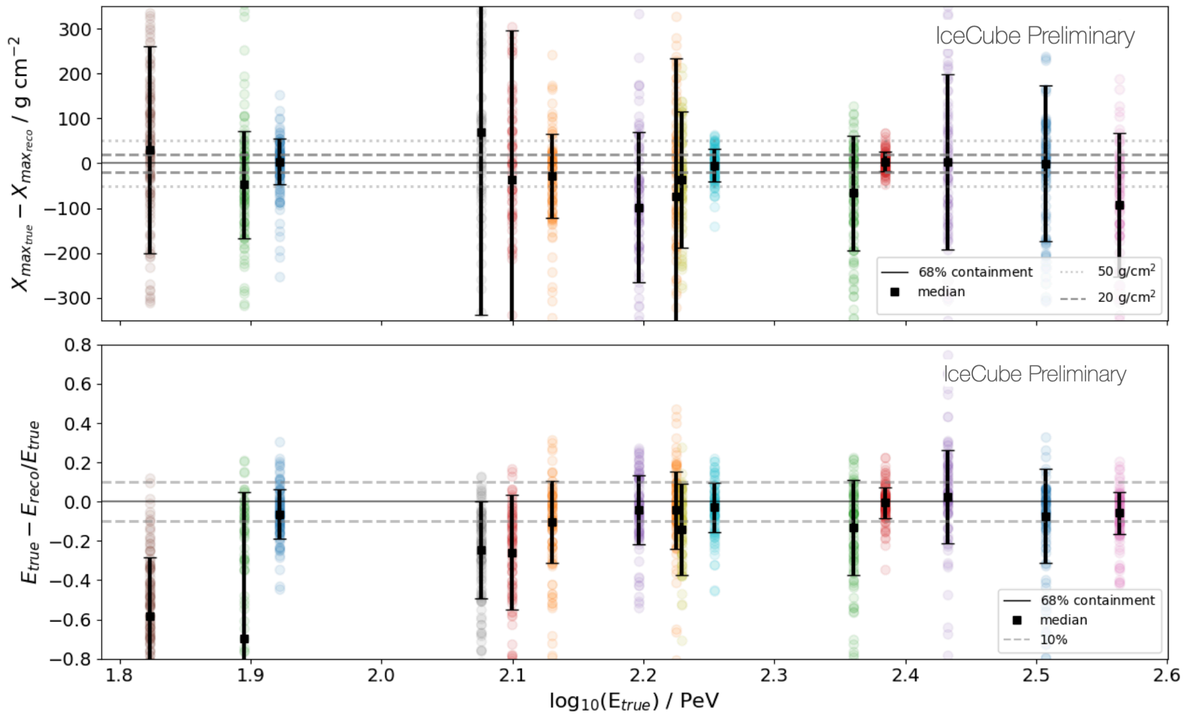
$$P'_{\text{sim}} = P_{\text{sim}} + c \cdot \left(1 - d \cdot e^{-\beta x^{-\alpha}}\right)$$

roxanne.turcotte@kit.edu

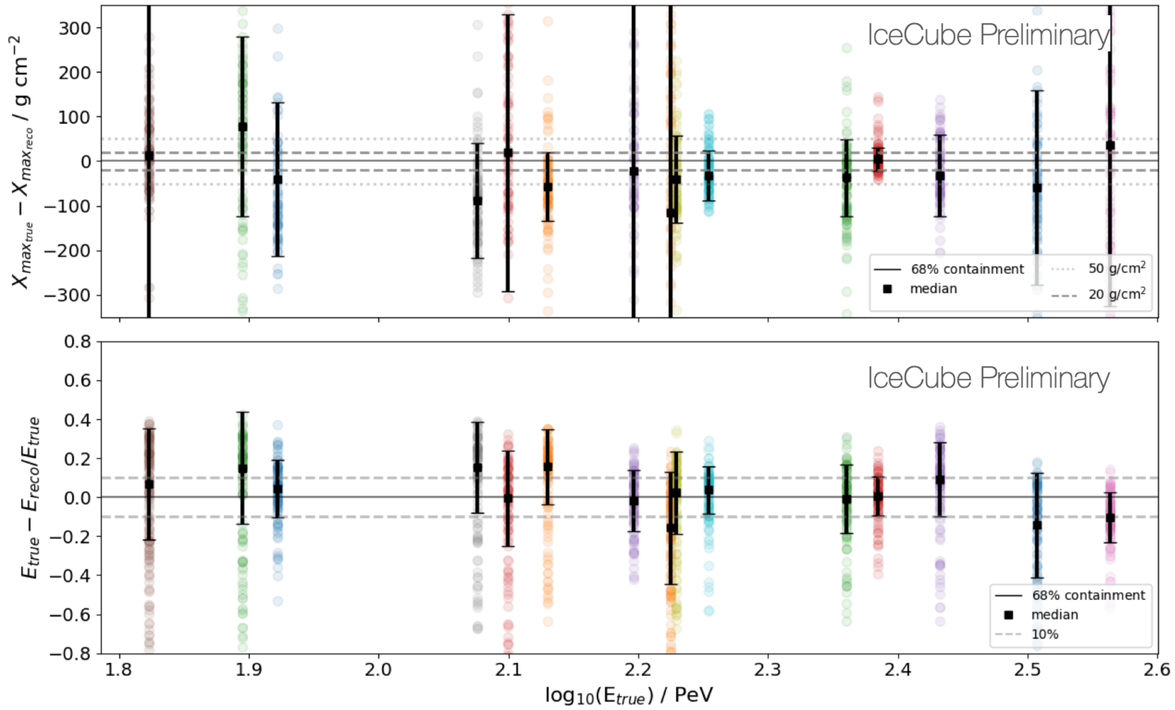
Results

- Hard to compare because...
  - There is still events with badly reconstructed in Xmax.
    - Also points towards better cuts
  - Need more and better events, nevertheless bias seems reduced even overshot

χ^2 - real noise



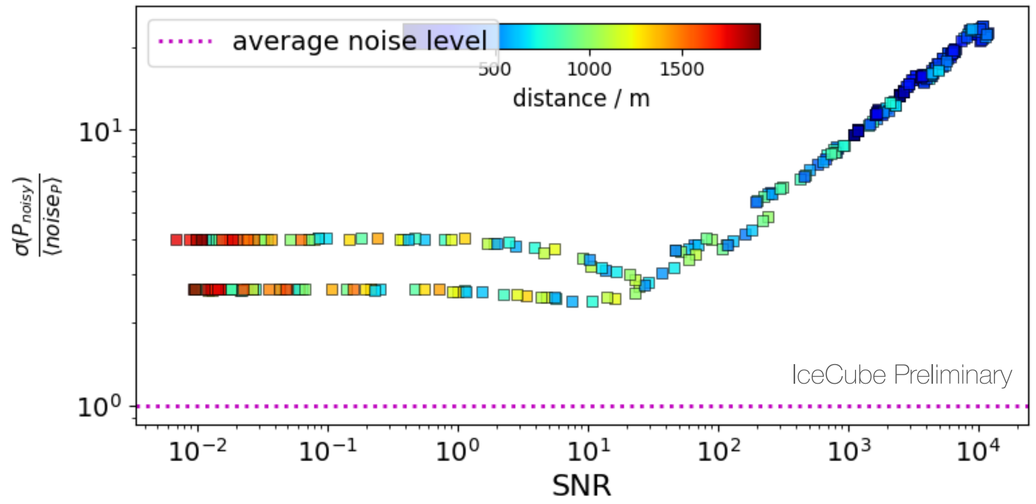
LLH - real noise



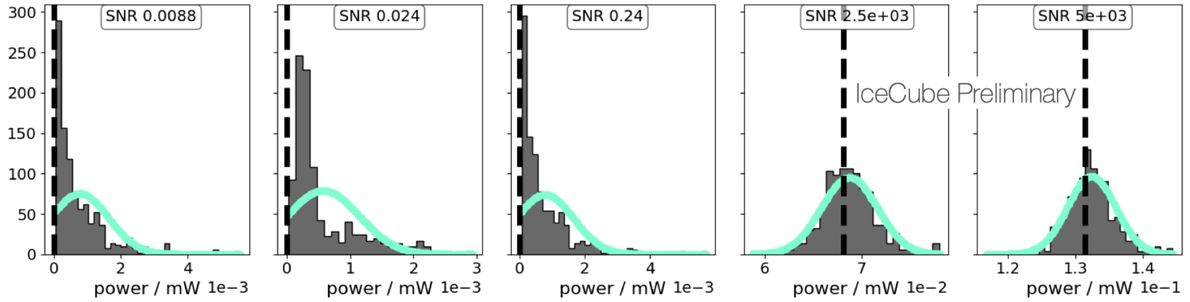
roxanne.turcotte@kit.edu

Why is the parametrisation not optimal ?

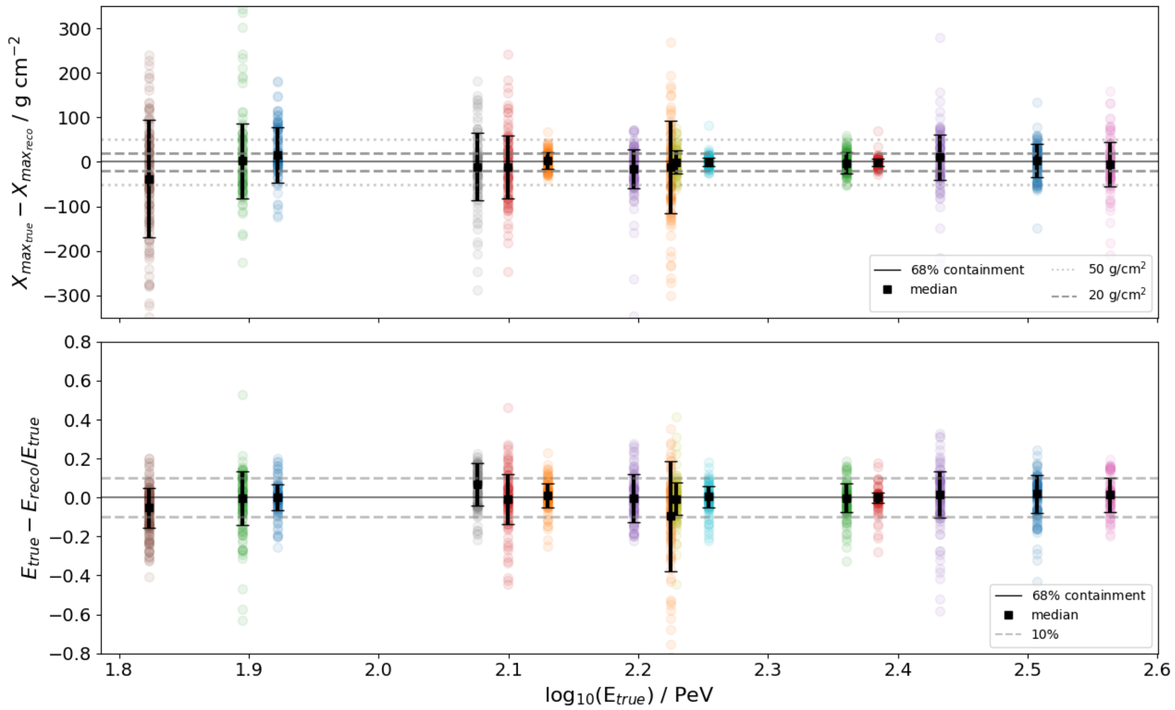
- Real noise
  - Bimodal distribution
  - Non-gaussian for small pulses
  - —> 2d parametrisation for LLH !
- Reducing the noise
  - Better filtering (median filter)
  - Denoiser ML - Abdul
  - Hardware improvements



Real noise distribution —> does not follow Gaussian



LLH - ideal noise (Cane + thermal from LNA)



roxanne.turcotte@kit.edu

Conclusion

- We implemented a technique based on the state-of-the art analysis
  - which works in this higher frequency band and with 3 antennas
- LLH represent better the system with this observable
- It's a prototype station, we are still learning about our data
- More events soon —> better quality cuts

Outlook

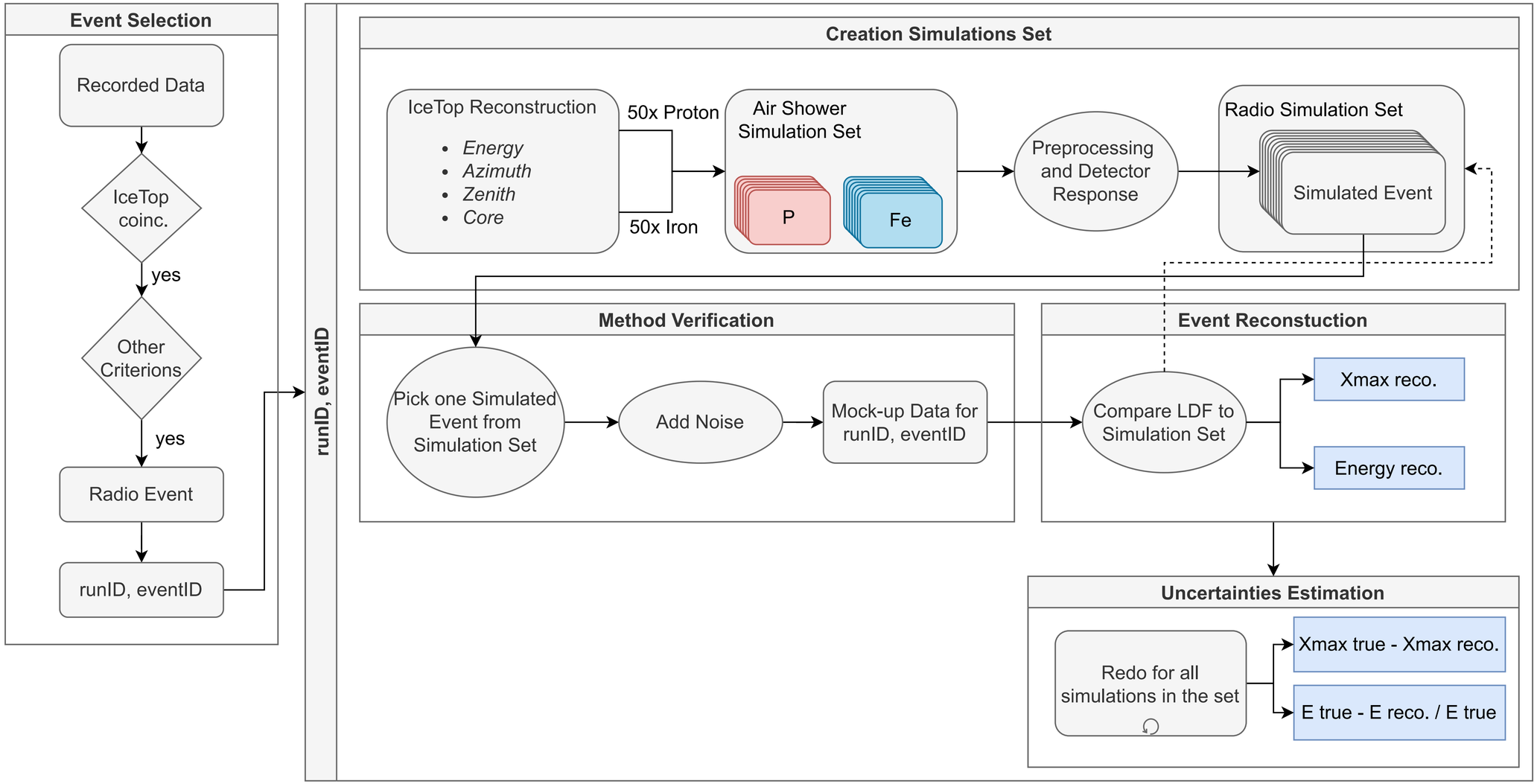
- Going for a 2D PDF parametrisation of noise
- Incorporating information about the spectrum and moving the core position
- Using information from the scintillation detectors as first input
- More stations !!

Thank you for your attention!

Questions?

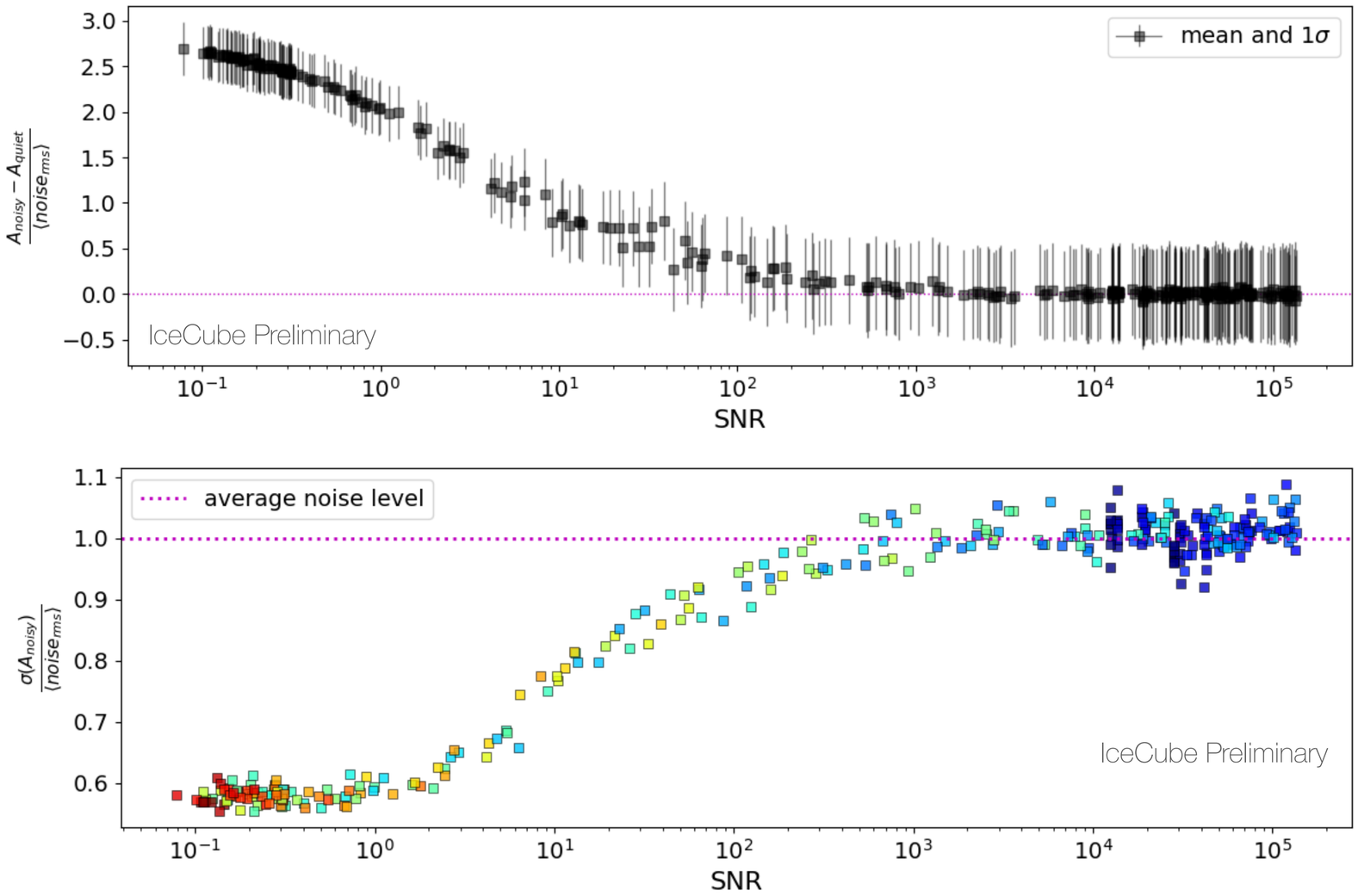
Back up slides

Methodology



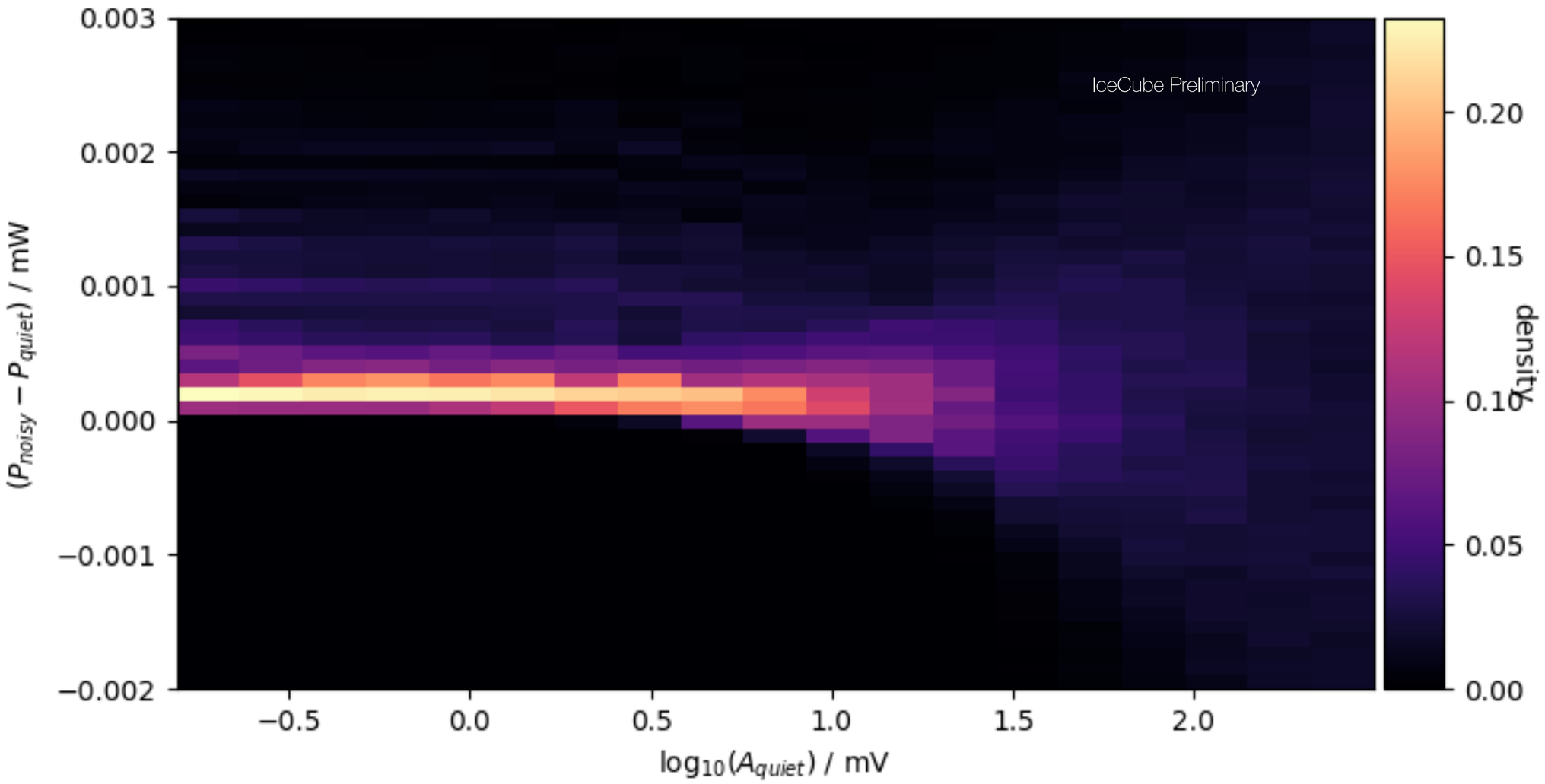
roxanne.turcotte@kit.edu

Maximum of the Hilbert envelope - Bias and standard deviation



roxanne.turcotte@kit.edu

2D PDF of noisy waveforms



roxanne.turcotte@kit.edu