

IceCube event reconstruction

Tianlu Yuan

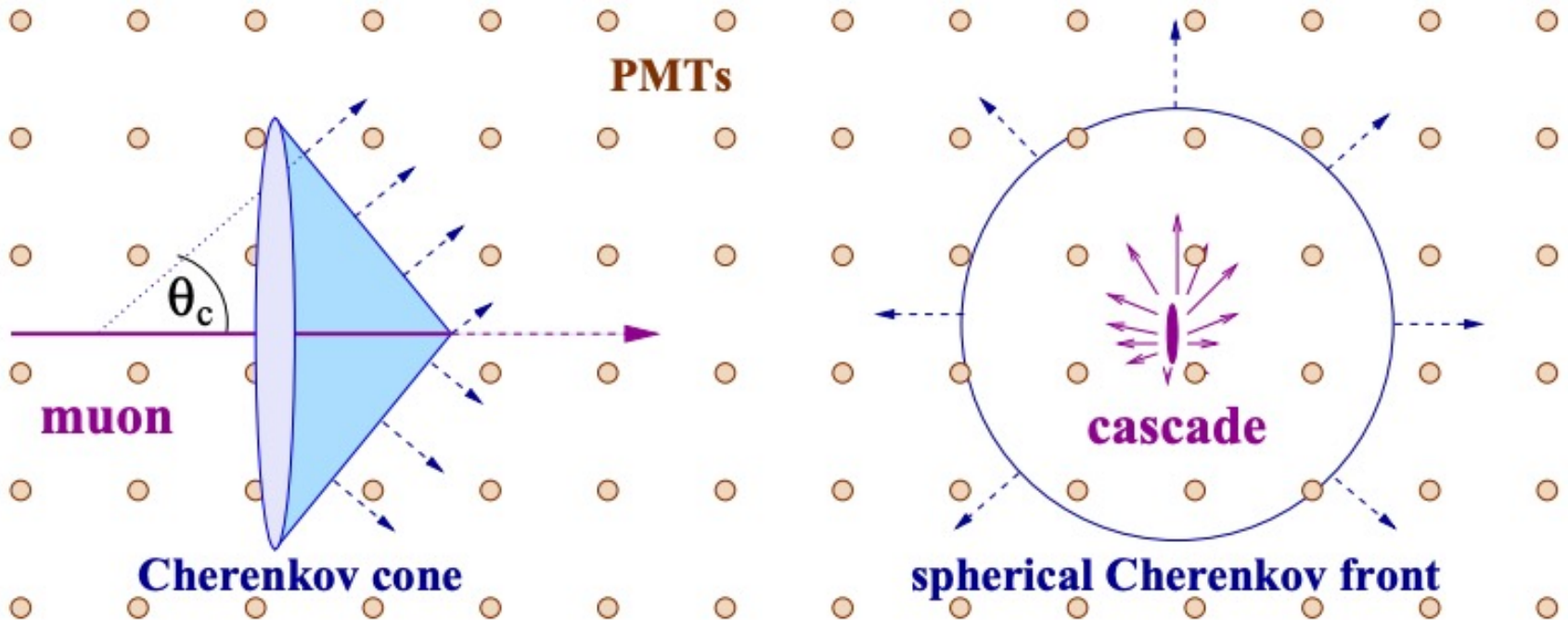
IceCube Summer School 2025



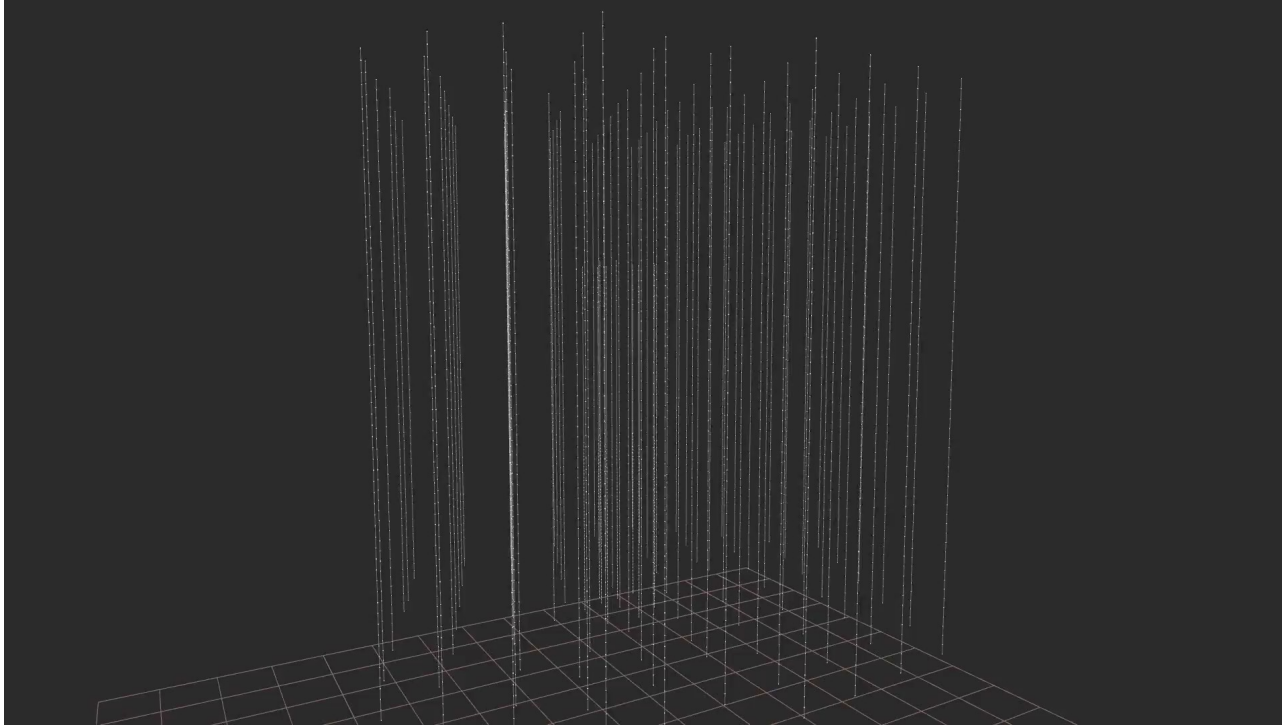
Tracks vs cascades

Tracks can travel large distance $\sim$ first photons on Cherenkov cone

Cascades travel relatively short distance $\sim$ diffuse photons w. spherical front

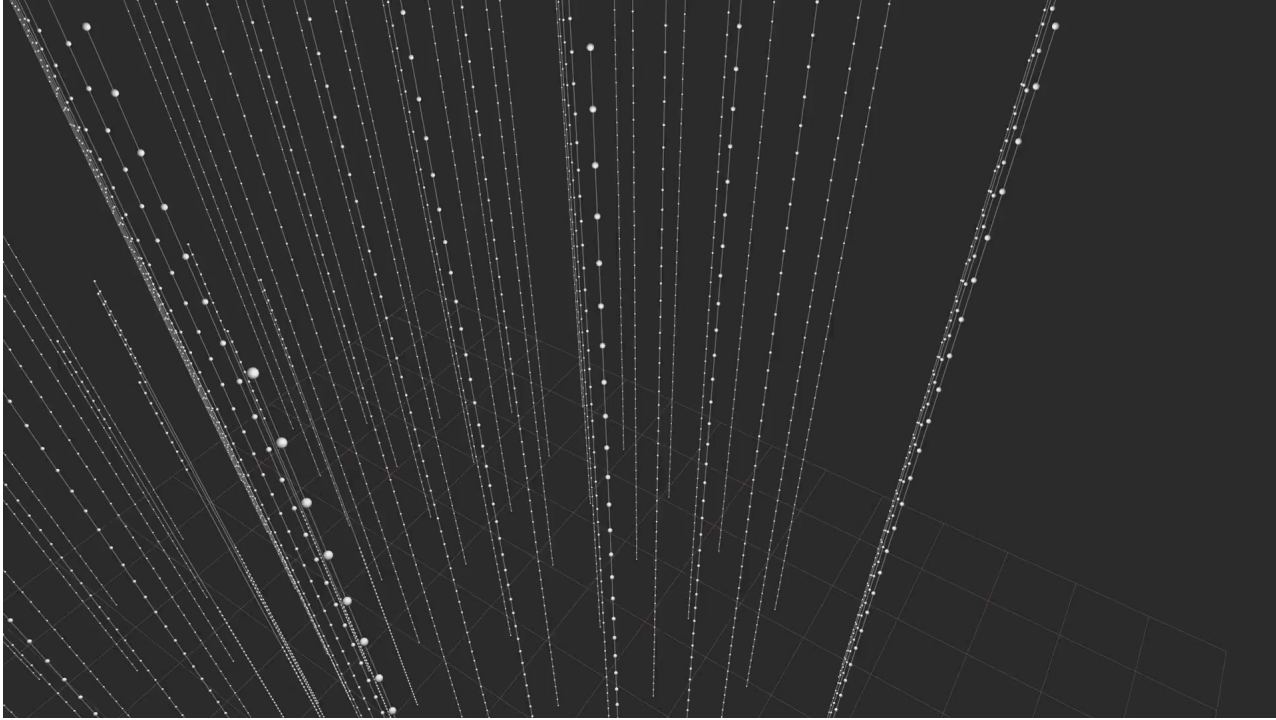


What do neutrinos look like in IceCube?



muons: long paths in the detector → **track**

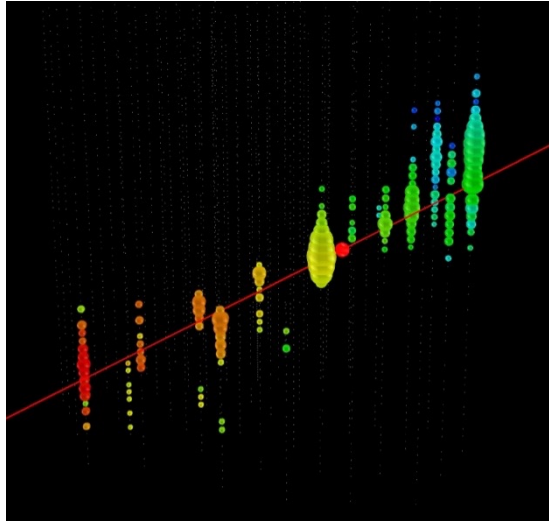
What do neutrinos look like in IceCube?



electrons/hadrons: shower of light → **cascade**

What IceCube actually sees (high-energy)

CC muon neutrino

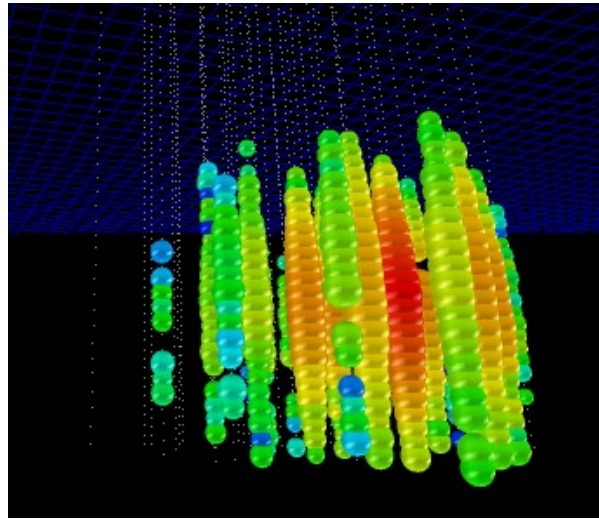


$$\nu_{\mu} + N \rightarrow \mu + X$$

track (data)

angular resolution $\sim 0.5^{\circ}$
energy resolution $\sim \times 2$

NC or CC electron neutrino

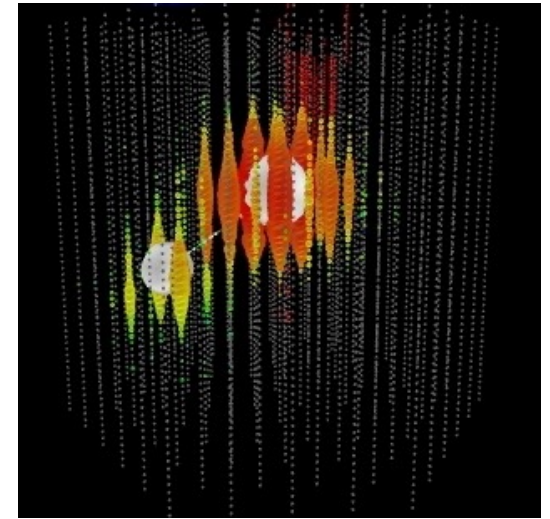


$$\begin{aligned}\nu_e + N &\rightarrow e + X \\ \nu_x + N &\rightarrow \nu_x + X\end{aligned}$$

shower (data)

angular resolution $\sim 10^{\circ}$
energy resolution $\sim 15\%$

CC tau neutrino

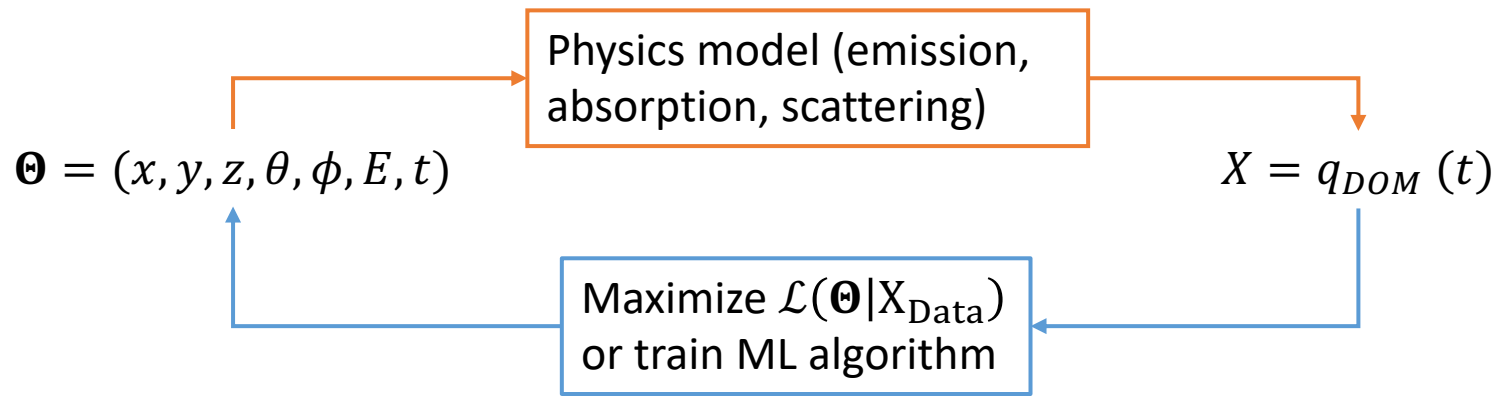
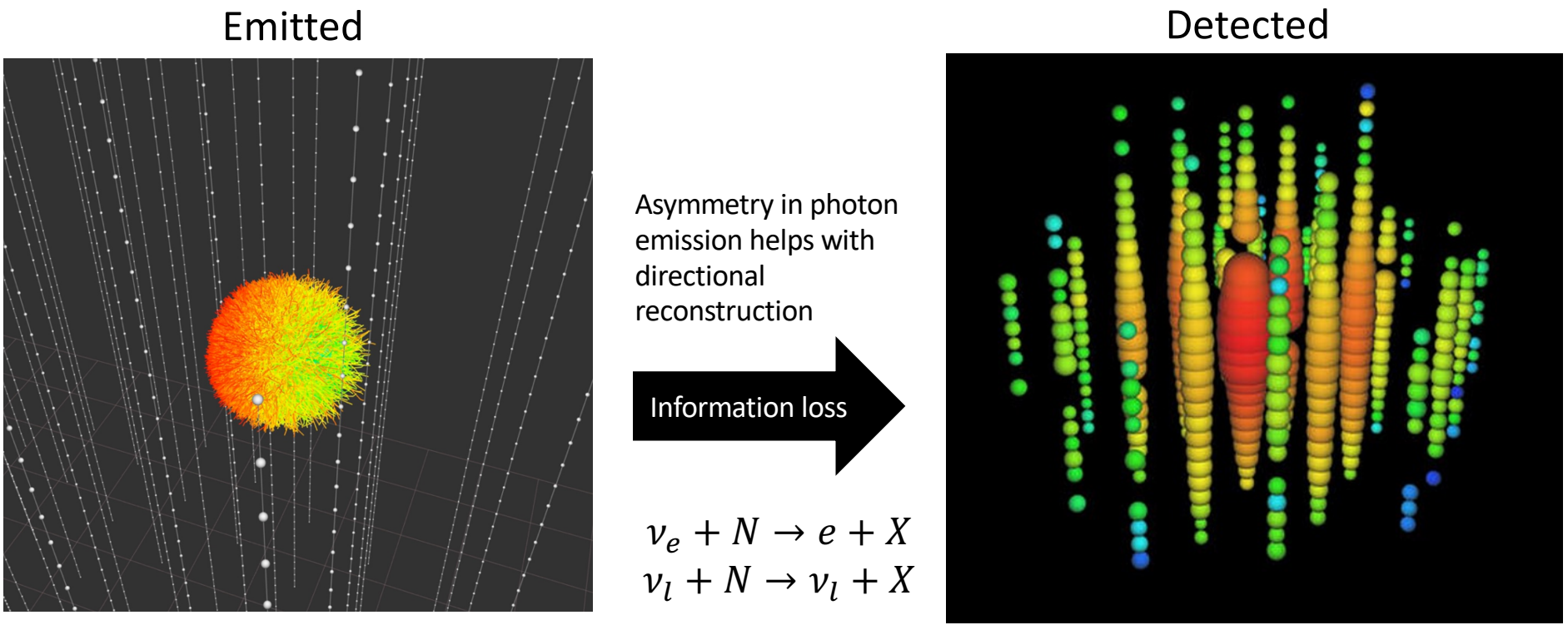


$$\nu_{\tau} + N \rightarrow \tau + X$$

“double-bang”
(simulation)

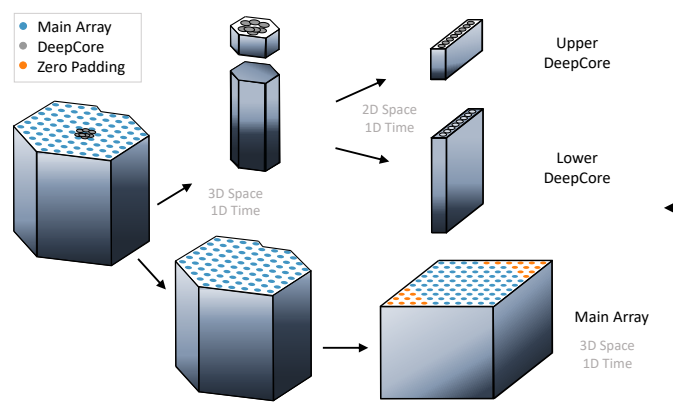
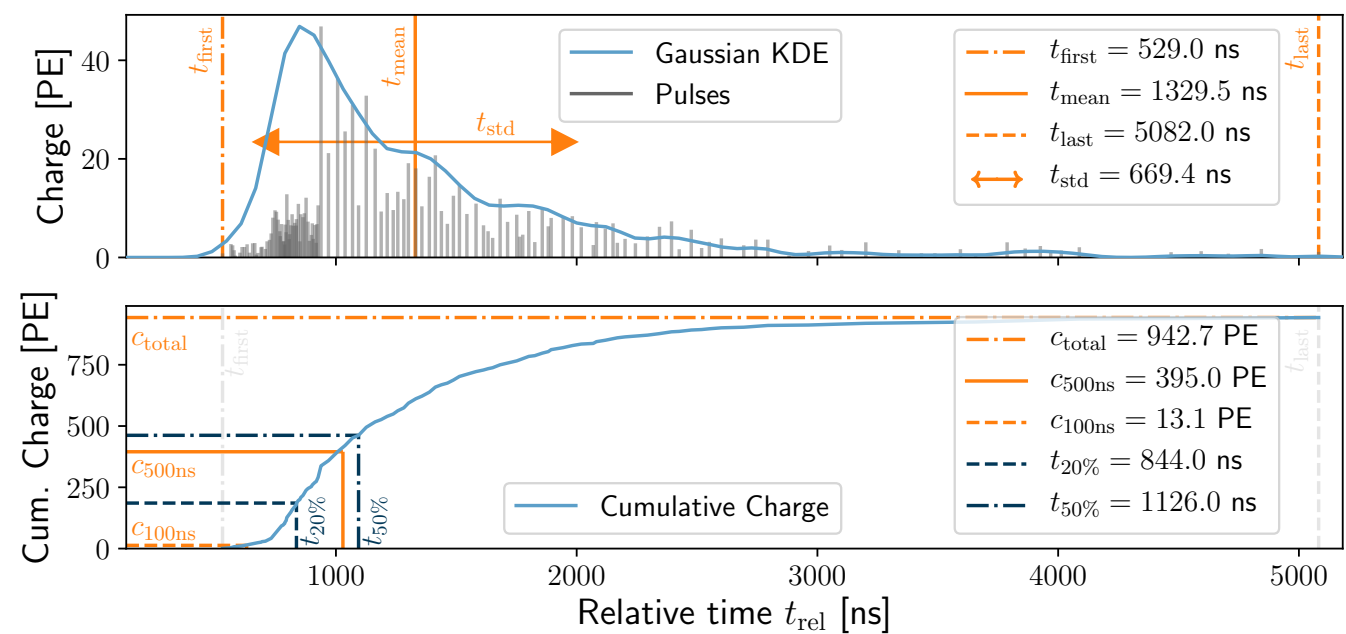
~ 2 expected in 6 years

Event reconstruction



One of the early DNN approaches

Input pulseseries features into CNN



JINST 16, 2021

Physics parameters and IceCube coordinates

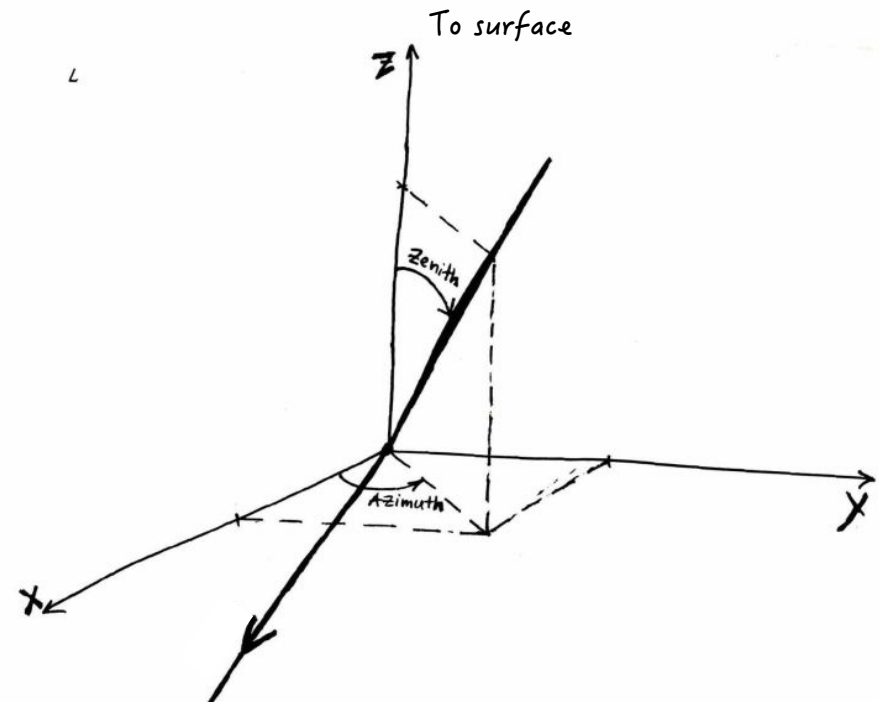
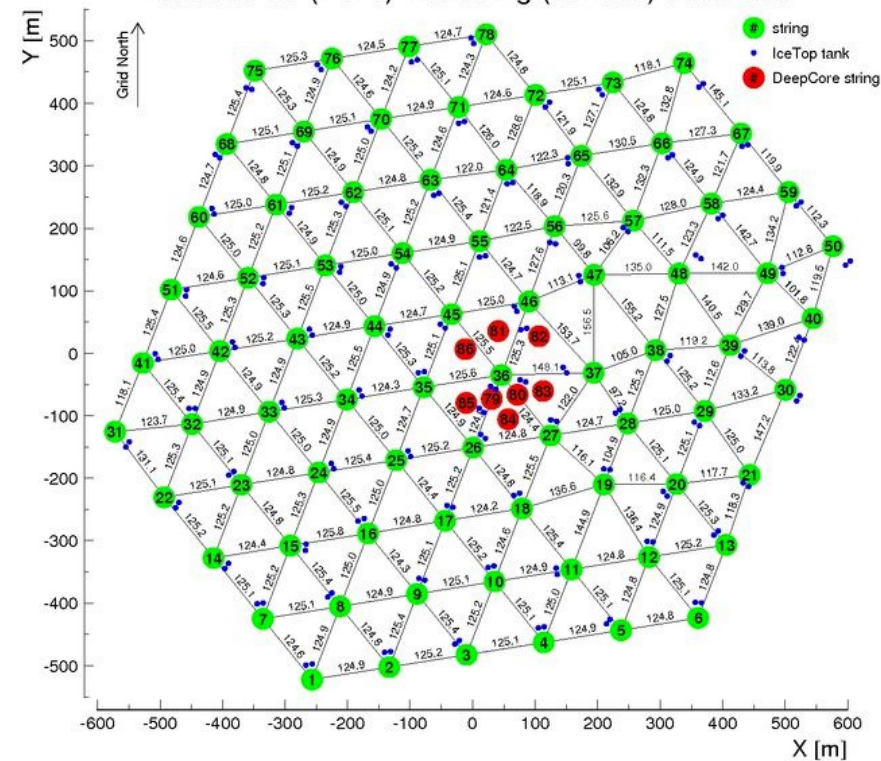
$$\Theta = (x, y, z, \theta, \phi, E, t)$$

Detector coordinate system centered in middle of detector

$(\theta, \phi) = (\text{zenith}, \text{azimuth})$ corresponds to *arrival* direction

Usually, (θ, ϕ, E) are the physics parameters we're most interested in

IceCube-86 (78+8) interstring (surface) distances



Approaches for reconstruction

Tracks

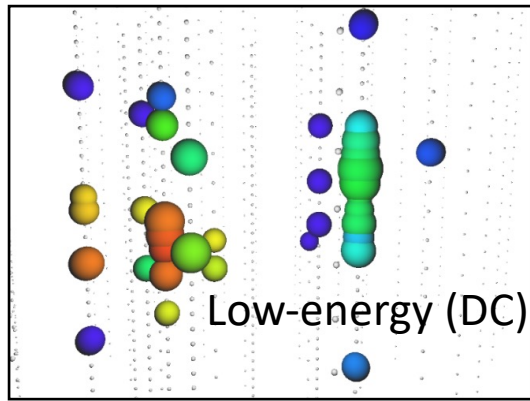
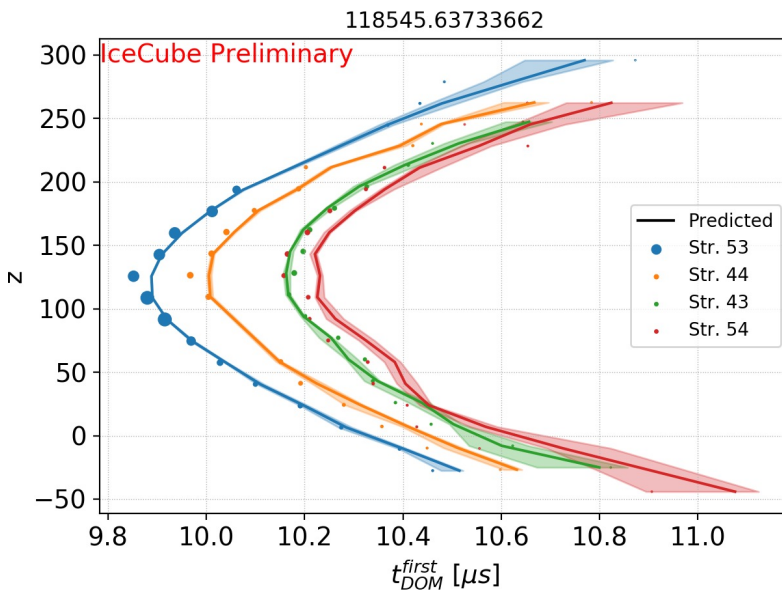
- Use **first-hit times** for *directional* reconstruction (**SANTA**, **SplineReco**)
- Use **full-waveform** information by fitting predicted light yields to what is actually seen (**RetroReco**, **DirectReco**, **MuEx**, **Millipede**, **DirectFit**)
 - Break **high-energy tracks** into multiple cascades along the track due to muon stochastic energy losses
- Likelihood-free inference (**FreeDOM**)
- Energy reco (**TruncatedEnergy**)
- ML (**FLERCNN**, **TNF**)

Approaches for reconstruction

Low energy
High energy

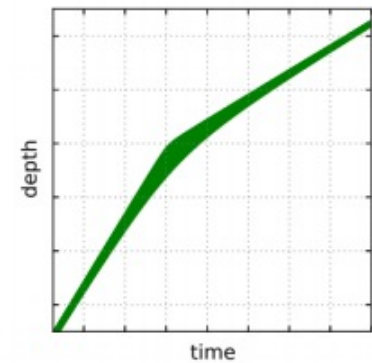
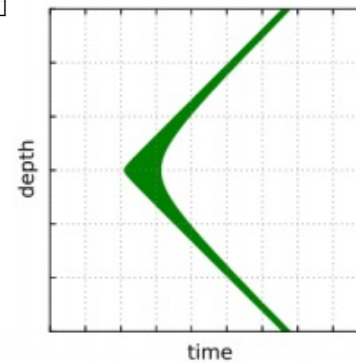
Tracks

- Use **first-hit times** for *directional* reconstruction (**SANTA**, **SplineReco**)



Intersection of cone
with string is hyperbola

Ideally **unscattered**
light



Juan Pablo Yanez

Arrival time pdfs

PMT jitter (Transit time spread) due to spread in initial energies/momenta of photoelectrons

arxiv:0407044

Additional effects due to:

- noise
- additional cascades along track
- scattering

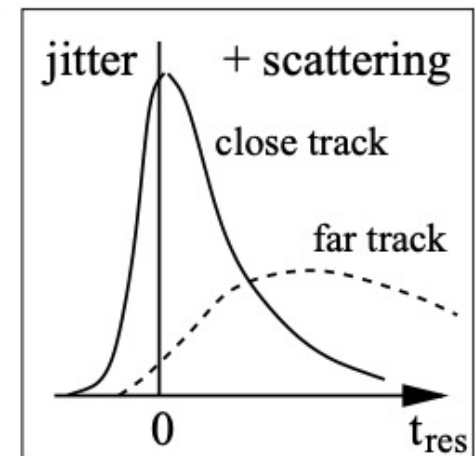
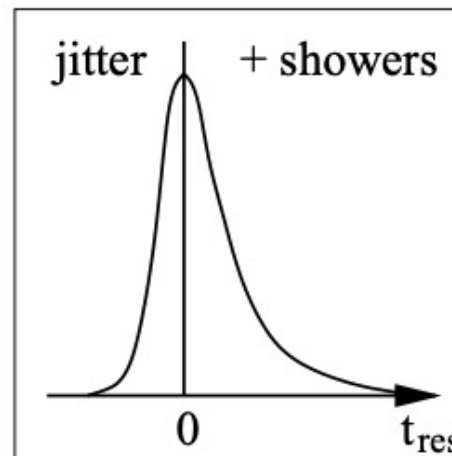
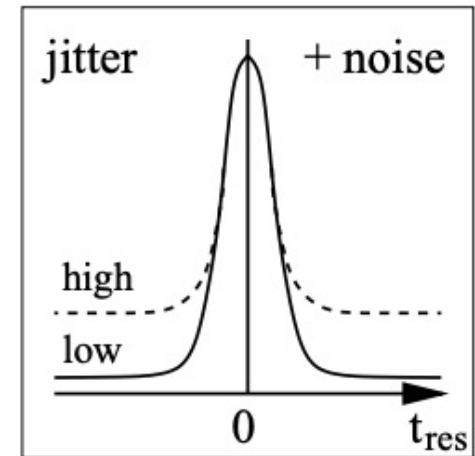
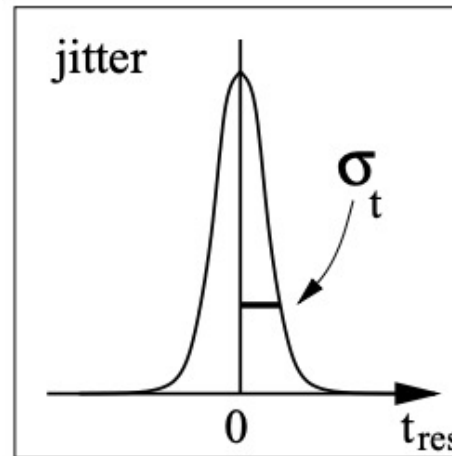
Original analytic parameterization

- “Pandel function” (gamma dist.)

$$p(t_{\text{res}}) \equiv \frac{1}{N(d)} \frac{\tau^{-(d/\lambda)} \cdot t_{\text{res}}^{(d/\lambda-1)}}{\Gamma(d/\lambda)} \cdot e^{-\left(t_{\text{res}} \cdot \left(\frac{1}{\tau} + \frac{c_{\text{medium}}}{\lambda_a}\right) + \frac{d}{\lambda_a}\right)},$$

$$N(d) = e^{-d/\lambda_a} \cdot \left(1 + \frac{\tau \cdot c_{\text{medium}}}{\lambda_a}\right)^{-d/\lambda},$$

- Now based on splines

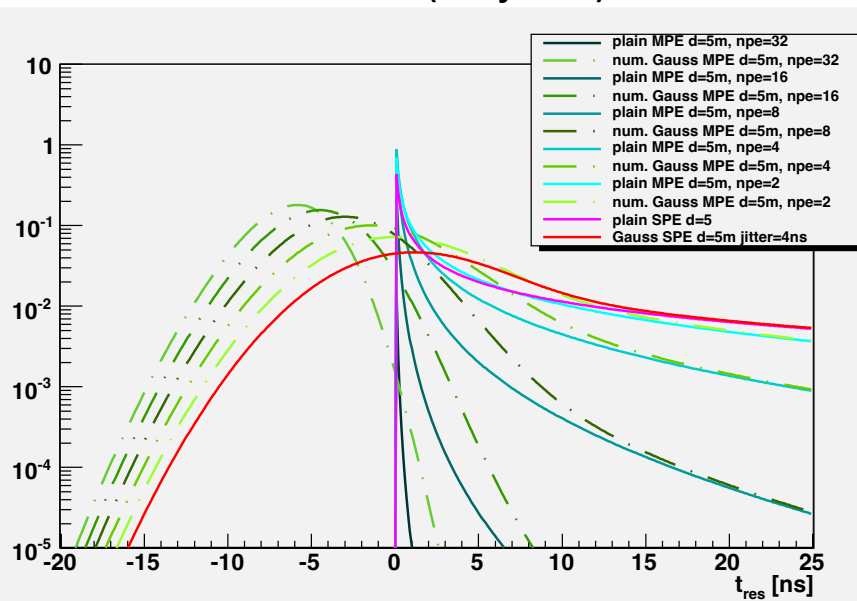


MPE Pandel likelihood

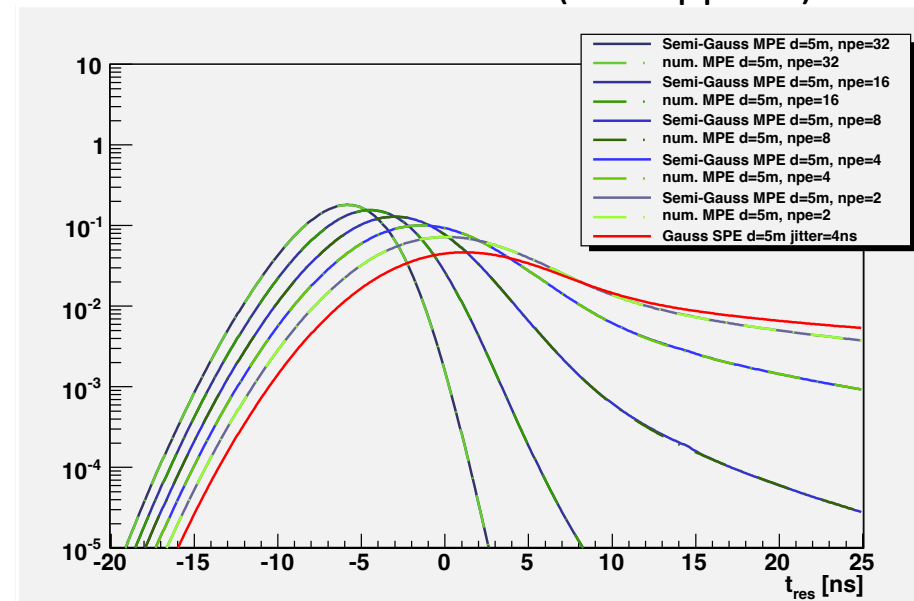
Pandel function cannot cope with negative time residuals so need to convolute with Gaussian

- <https://user-web.icecube.wisc.edu/~boersma/PandelUpdates/MPEplots/>

Plain MPE (no jitter)

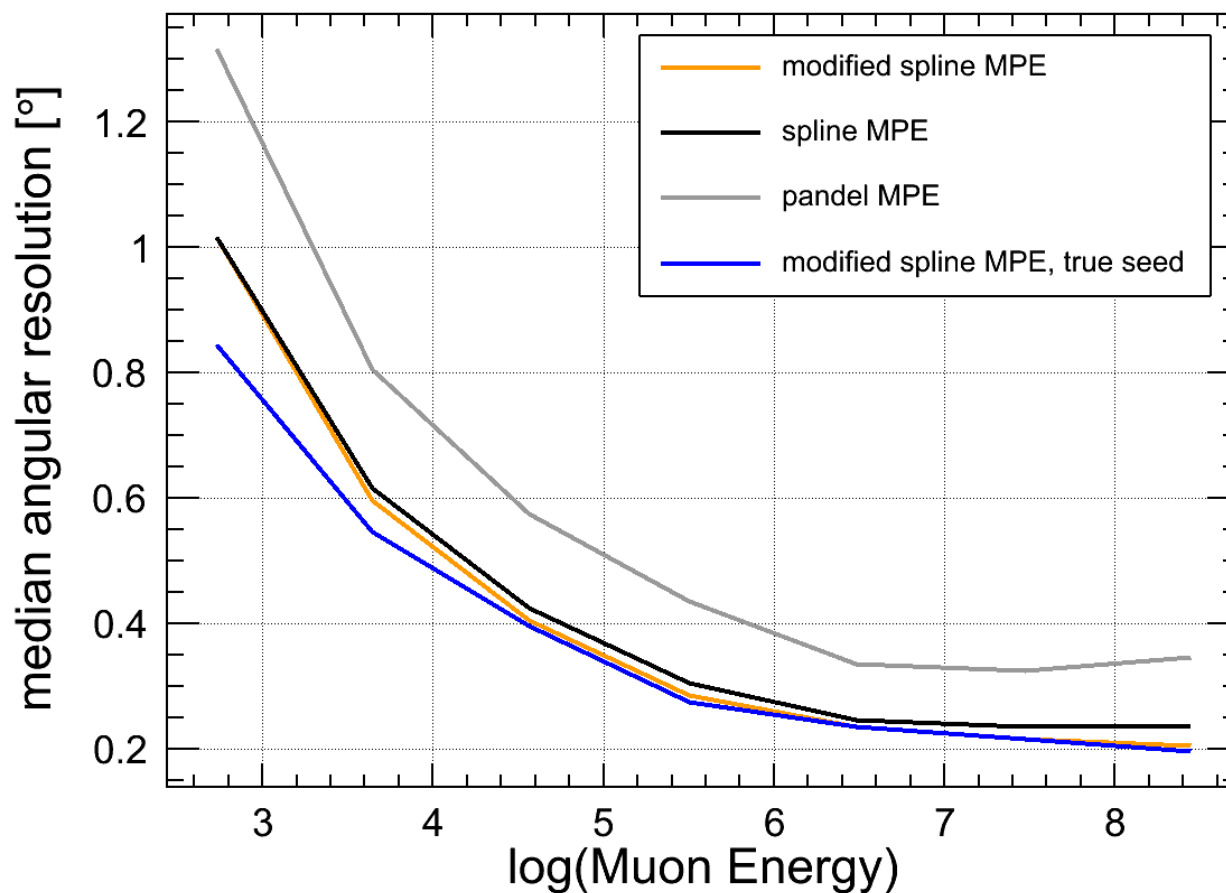


Gauss convoluted (fast-approx.)

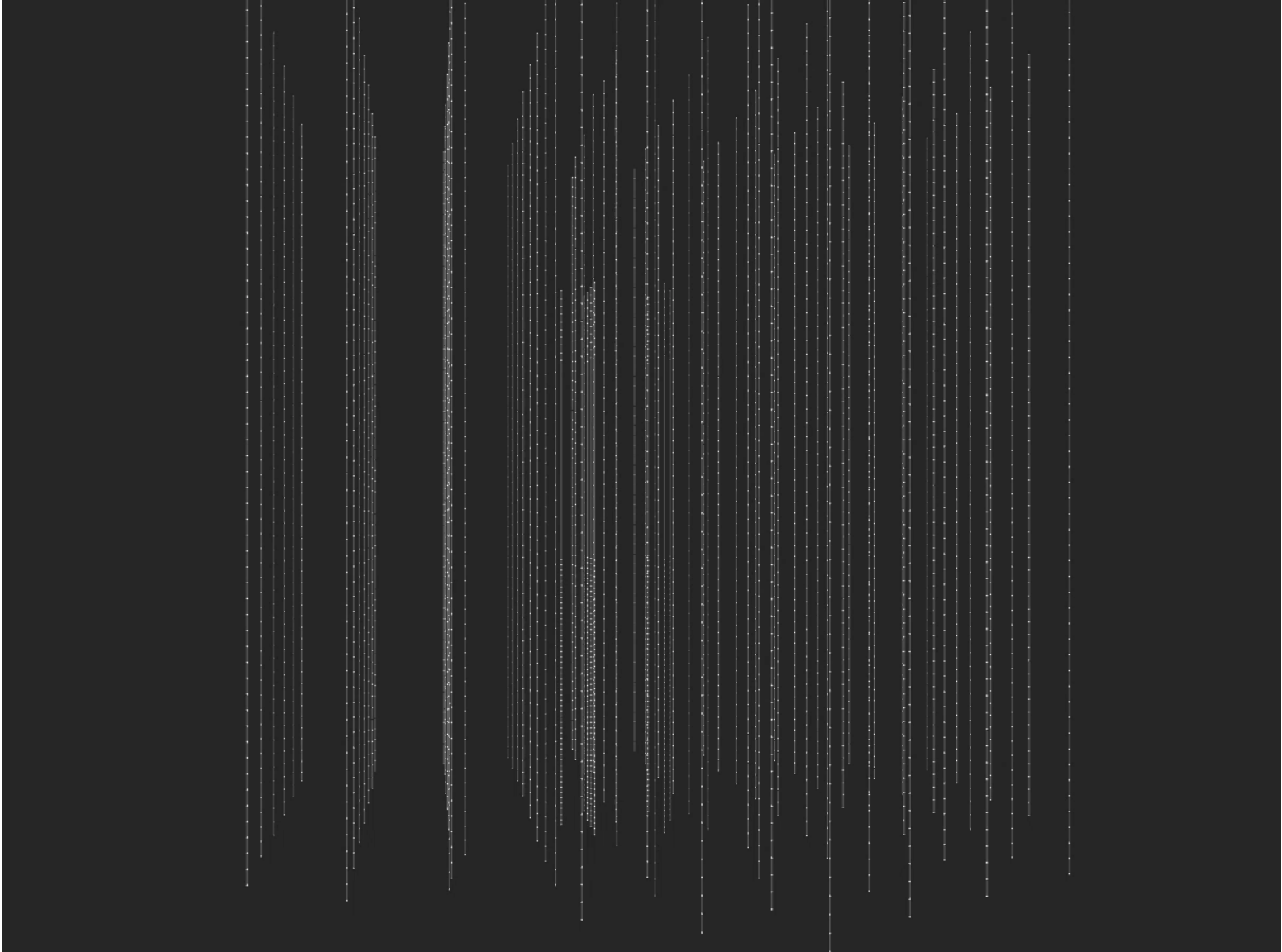


SplineReco Resolutions

Improvements were made by moving to (photo)spline tables based on simulation (c.f. [K. Schatto thesis](#))



Utilizing full-waveform information



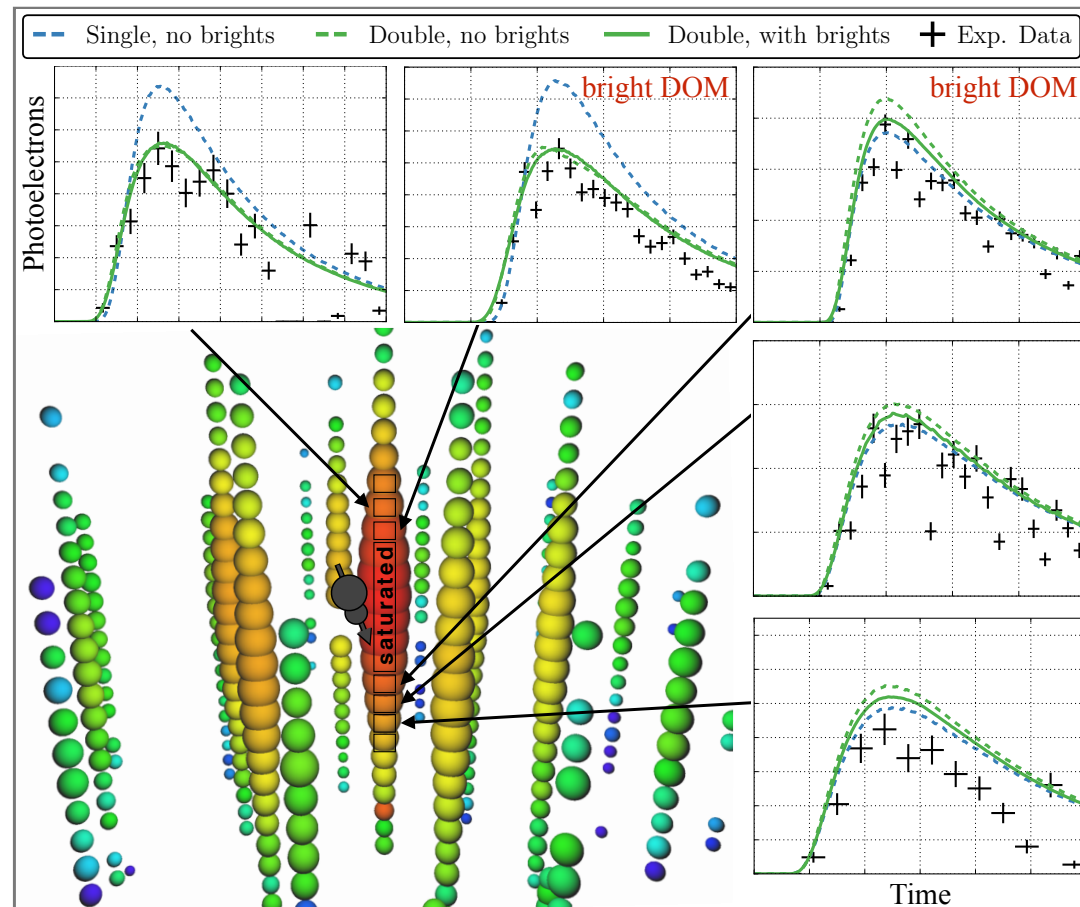
Full-waveform expectations

Tabulated photon yields

- Pros: Fast runtime; gradients
- Cons: Approximate ice model

Direct photon propagation

- Pros: Any ice-model can be used exactly
- Cons: Statistical errors from both data and MC; slow



IC collaboration, 1311.4767
D. Chirkin, arXiv:1304.0735

NN modeling (EventGenerator)

Models photoelectron yields with NN using basis functions (e.g. asymmetric gaussian)

Generator NN learns mapping:

$$f: \vec{\theta} \rightarrow \vec{\lambda}, \vec{p}(t_i | \vec{\theta})$$

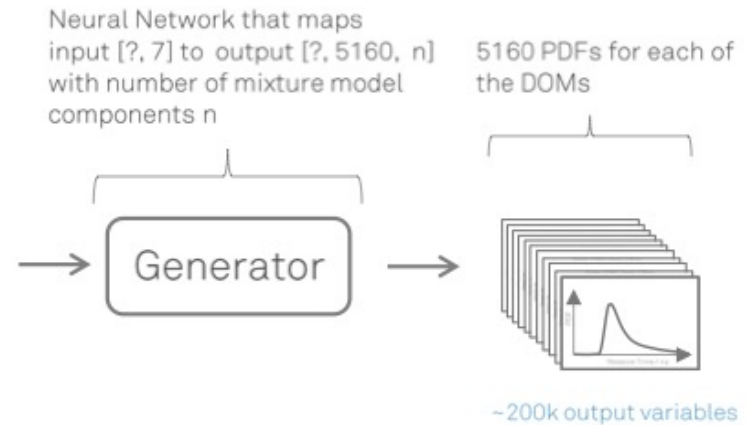
$\vec{\theta}$: source parameters

$\vec{\lambda}$: expected DOM charge

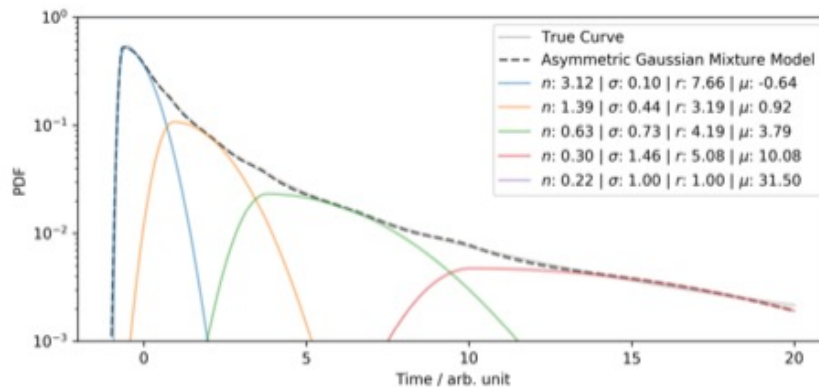
$\vec{p}(t_i | \vec{\theta})$: pulse arrival PDFs

Example: Cascade Hypothesis

$$\vec{\theta} = (x, y, z, \varphi, \theta, E, t)$$



Parameterization of pulse arrival time PDF:



$$g(x|\mu, \sigma, r) = \begin{cases} N \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), & x \leq \mu \\ N \cdot \exp\left(-\frac{(x-\mu)^2}{2(\sigma r)^2}\right), & \text{otherwise} \end{cases}$$

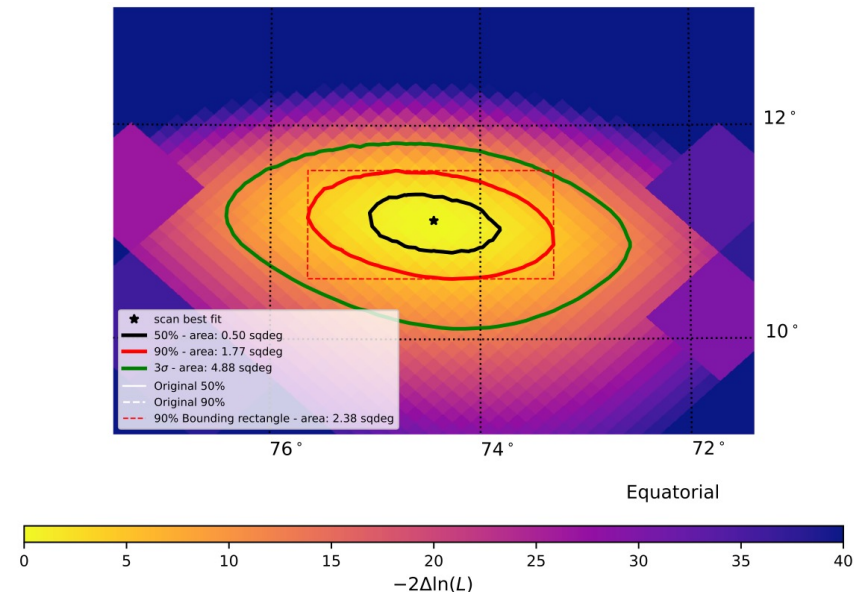
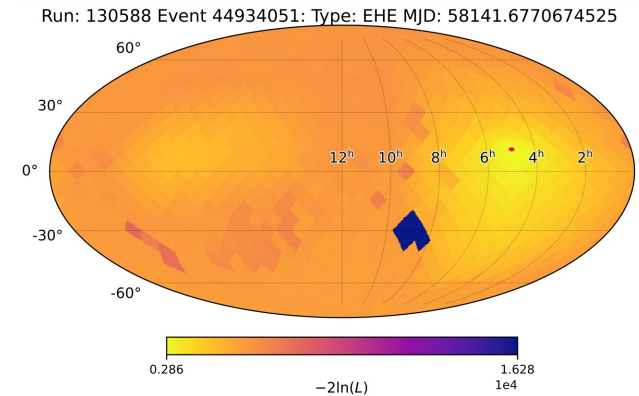
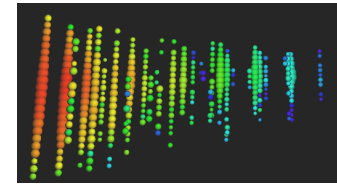
$$N = \frac{2}{\sqrt{(2\pi) \cdot \sigma(r+1)}}$$

DOI: 10.1007/3-540-70659-3_42

Minimization approaches

Given a likelihood $\mathcal{L}(\Theta|X_{\text{Data}})$ as a function of $\Theta = (x, y, z, \theta, \phi, E, t)$, want to find Θ_0 that minimizes the negative likelihood

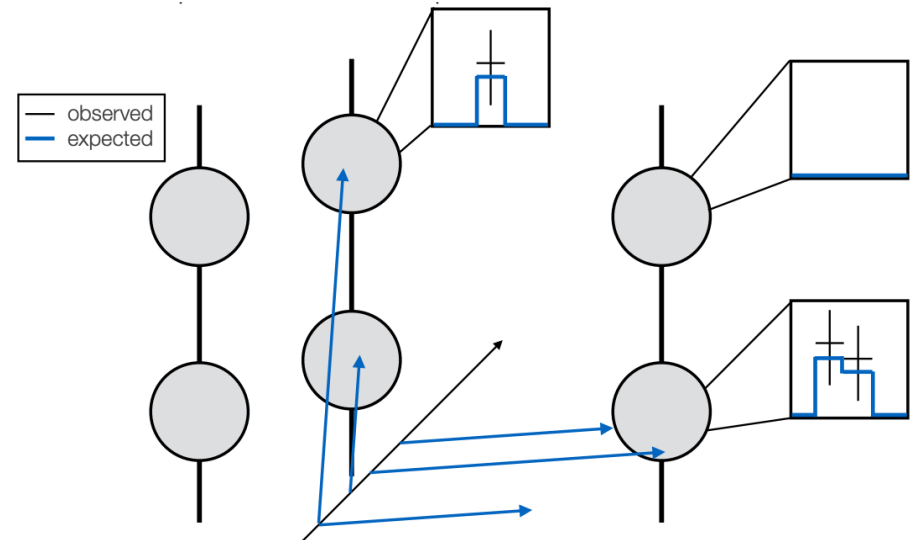
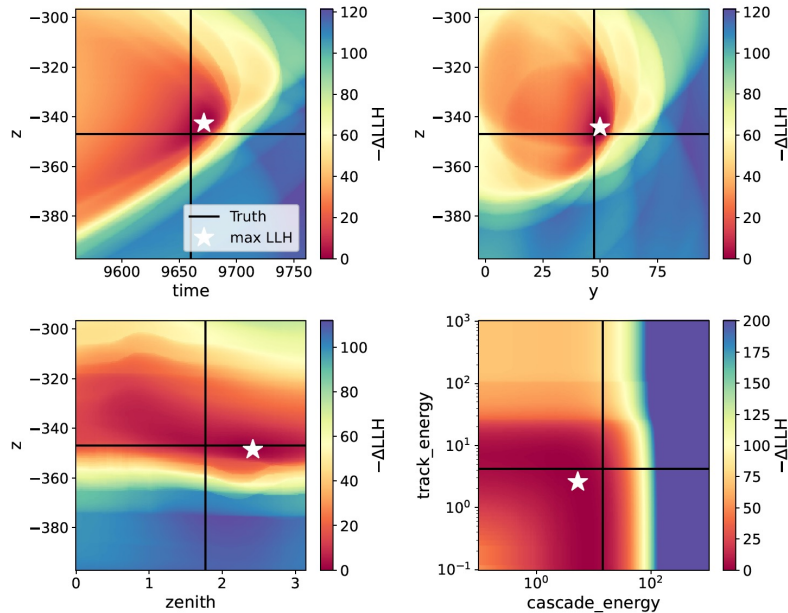
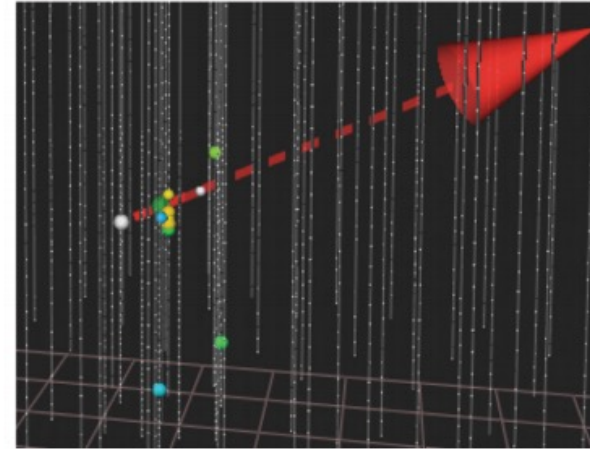
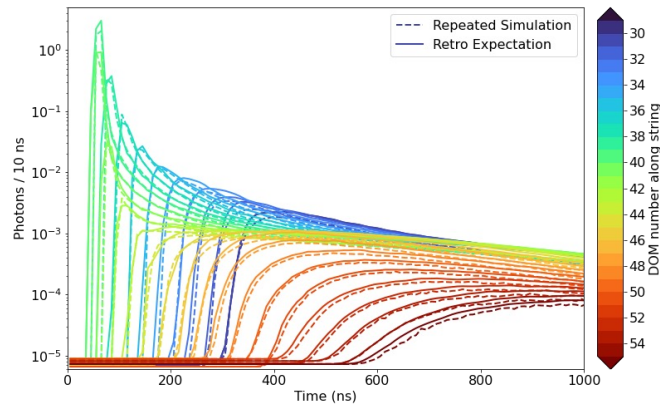
- Millipede uses **photon tables** which allows for iterative gradient descent
- Can also **brute force** all possible directions (θ, ϕ) to reduce the minimization to only 5 dimensions (realtime alerts)



Low-energy reco: RetroReco and DirectReco

RetroReco: emit photons from DOM and track $\rightarrow$ then build retro tables

DirectReco: like DirectFit but for lower energies



Approaches for reconstruction

Tracks

- Use **first-hit times** for *directional* reconstruction (**SANTA**, **SplineReco**)
- Use **full-waveform** information by fitting predicted light yields to what is actually seen (**RetroReco**, **DirectReco**, **MuEx**, **Millipede**, **DirectFit**)
 - Millipede works for **high-energy tracks** by breaking it up into multiple cascades along the track due to muon stochastic energy losses
- Likelihood-free inference (**FreeDOM**)
- Energy reco (**TruncatedEnergy**)
- ML (**FLERCNN**, **TNF**)

Cascades

- Use **full-waveform** information by fitting predicted light yields to what is actually seen (**RetroReco**, **DirectReco**, **Monopod**, **DirectFit**)
- ML (**FLERCNN**, **DNN**, **TNF**)
- ML+LLH approaches (**EventGenerator**)
- Likelihood-free inference (**FreeDOM**)

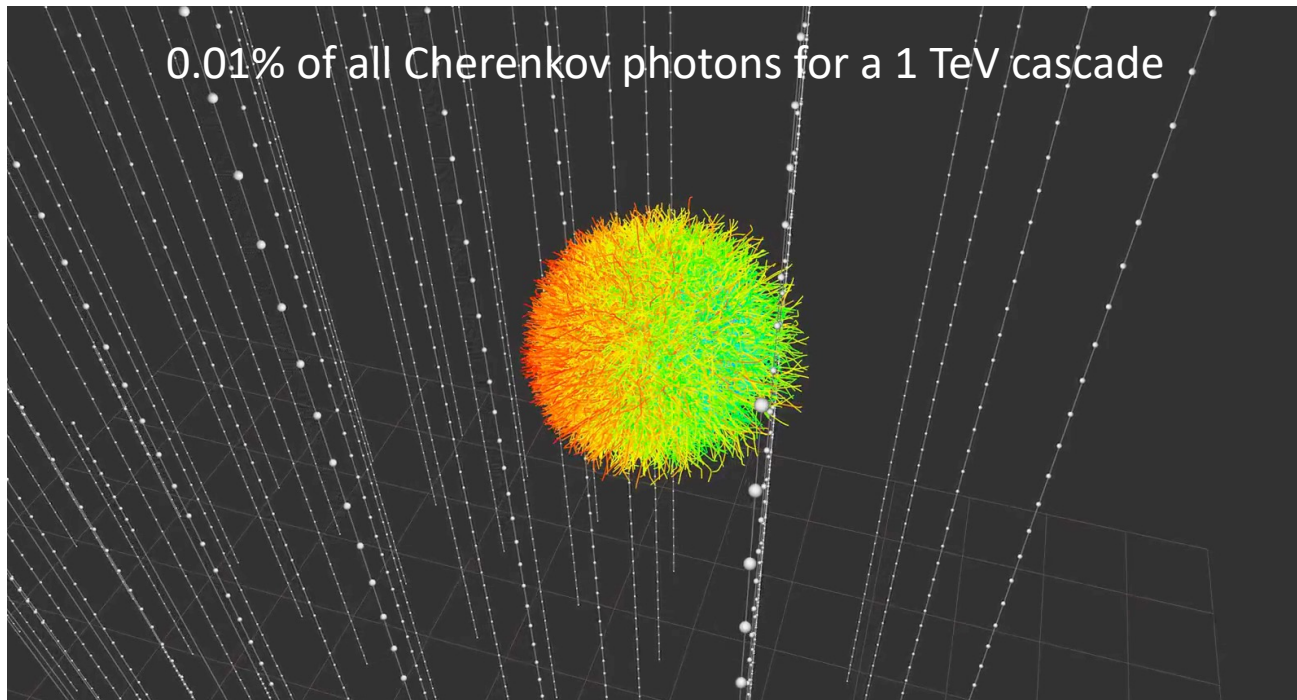
Challenges in cascade reconstruction

Large distances between DOMs means not many detected photons

Small asymmetry means high dependence on ice modeling

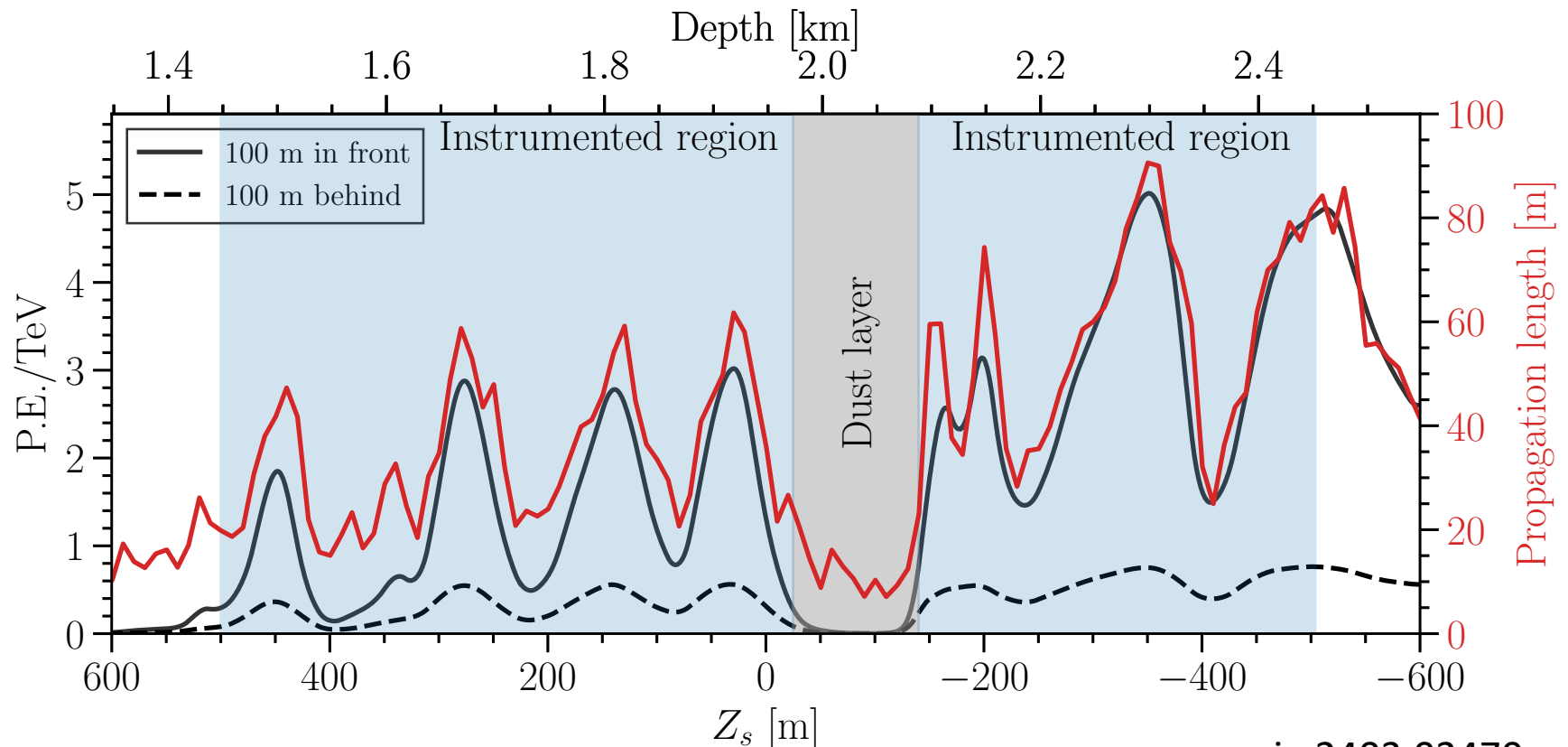
Sheer number of photons difficult to simulate

1. Tabulate photon yields for a single ice model (Millipede/Monopod)
 - Fast, table generation time-consuming
2. Directly propagate all photons for any ice model (DirectFit)
 - Slow but accurate



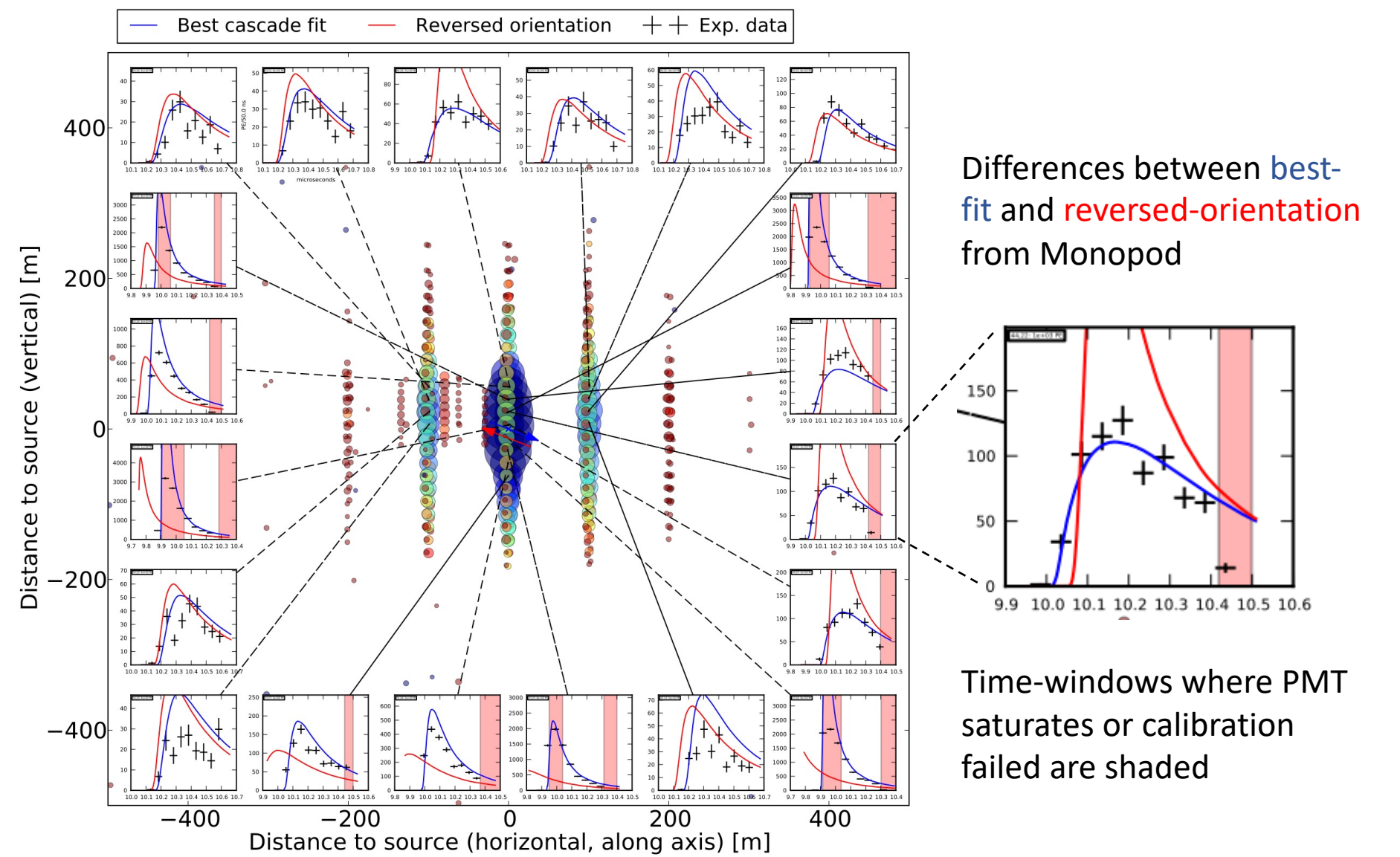
Photon yields vs ice

Number of arrival P.E.s shows clear correlation with ice optical properties
Differences between receiver in front of / behind shower allows for directional reconstruction



arxiv:2403.02470

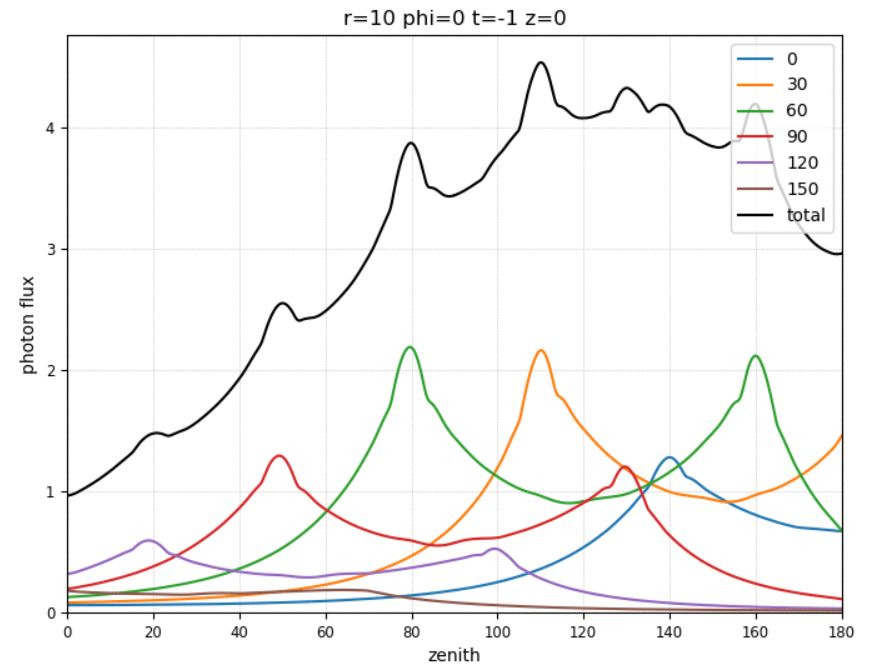
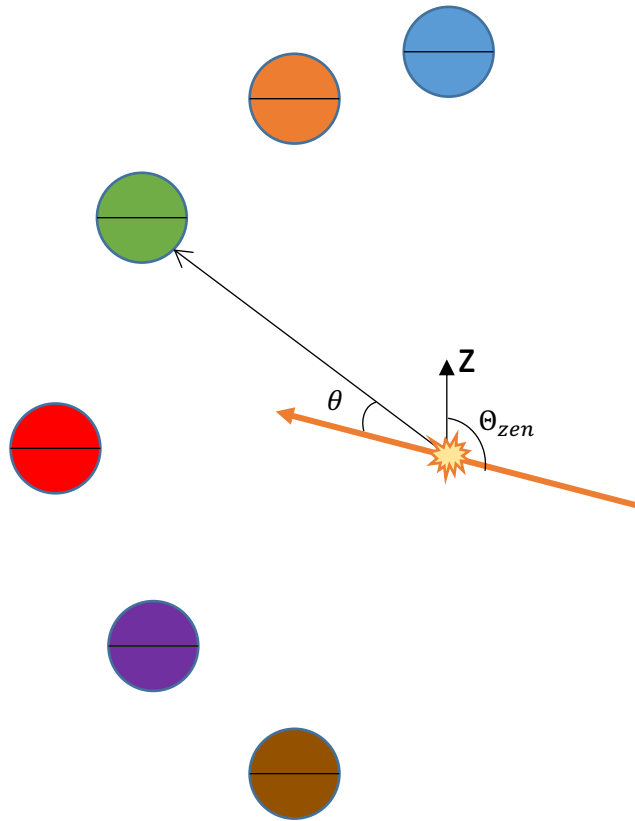
Cascade orientation from full-waveform



Photon amplitudes

Photon flux at different receivers as taken from photospline

Cherenkov peaks visible nearby, falls off rapidly with distance



Minimization approaches

Given a likelihood $\mathcal{L}(\Theta|X_{\text{Data}})$ as a function of $\Theta = (x, y, z, \theta, \phi, E, t)$, want to find Θ_0 that minimizes the negative likelihood

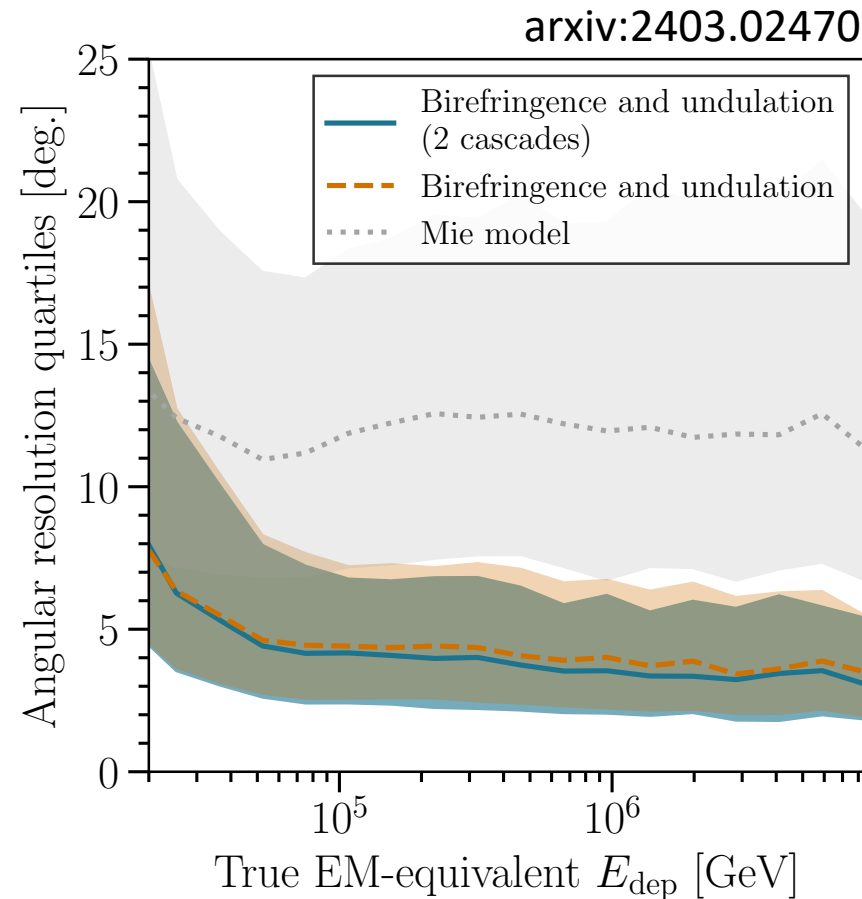
Need to explore 7D space which is challenging

- Monopod uses **photon tables** which allows for iterative gradient descent
 - May not always find the global minimum

Impact of ice on cascade angular resolution

“Mie model” neglects birefringence and layer undulations

Some slight additional improvement when using two-cascade fit to model shower extension



Minimization approaches

Given a likelihood $\mathcal{L}(\Theta|X_{\text{Data}})$ as a function of $\Theta = (x, y, z, \theta, \phi, E, t)$, want to find Θ_0 that minimizes the negative-likelihood

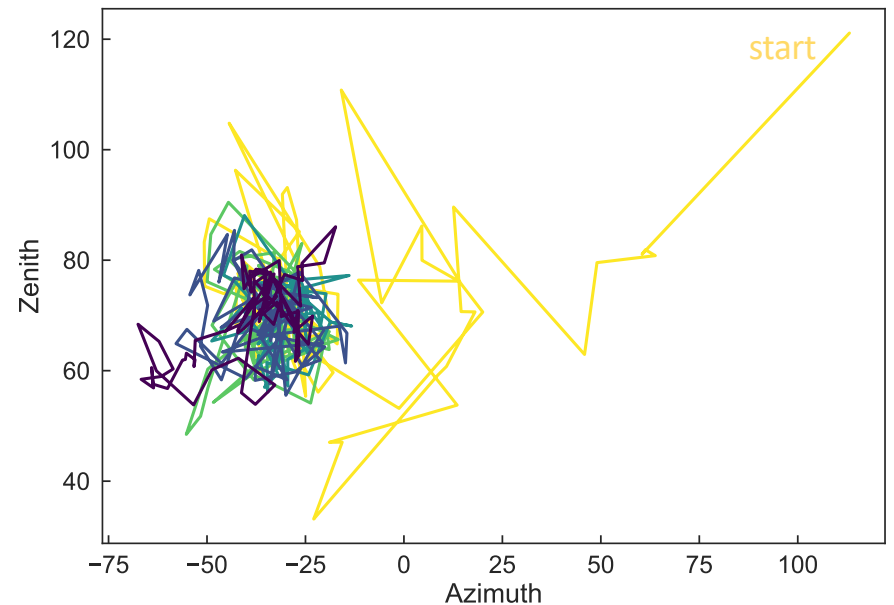
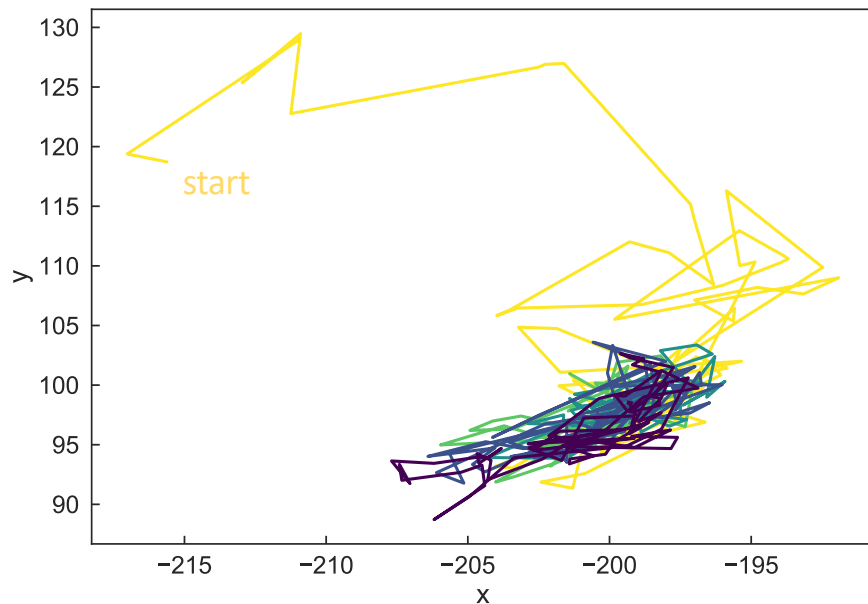
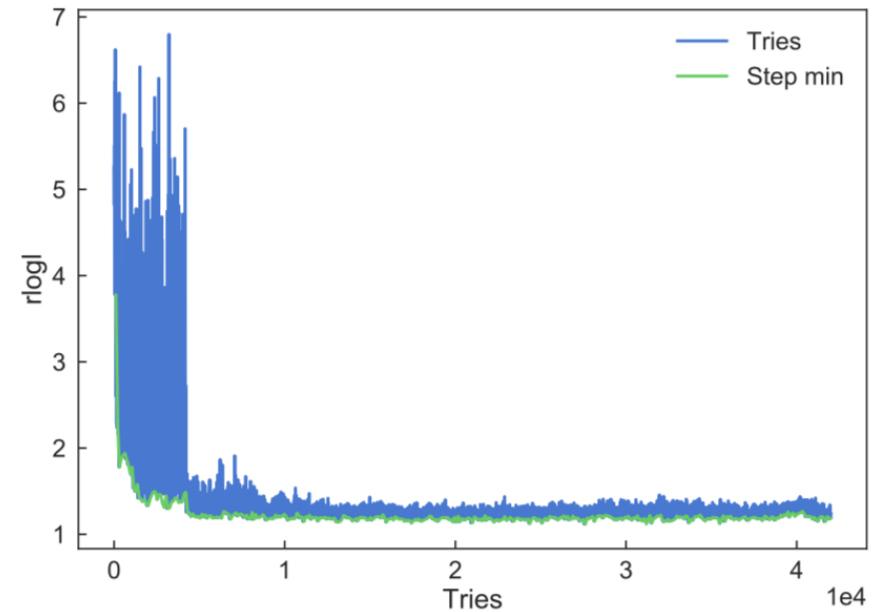
Need to explore 7D space which is challenging

- Millipede/Monopod uses **photon tables** which allows for iterative gradient descent
 - May not always find the global minimum
- DirectFit attempts to find minimum using **localized random search**, randomly sampling points in (x, y, z, θ, ϕ) within a “search radius” that is refined iteratively

DirectFit minimization

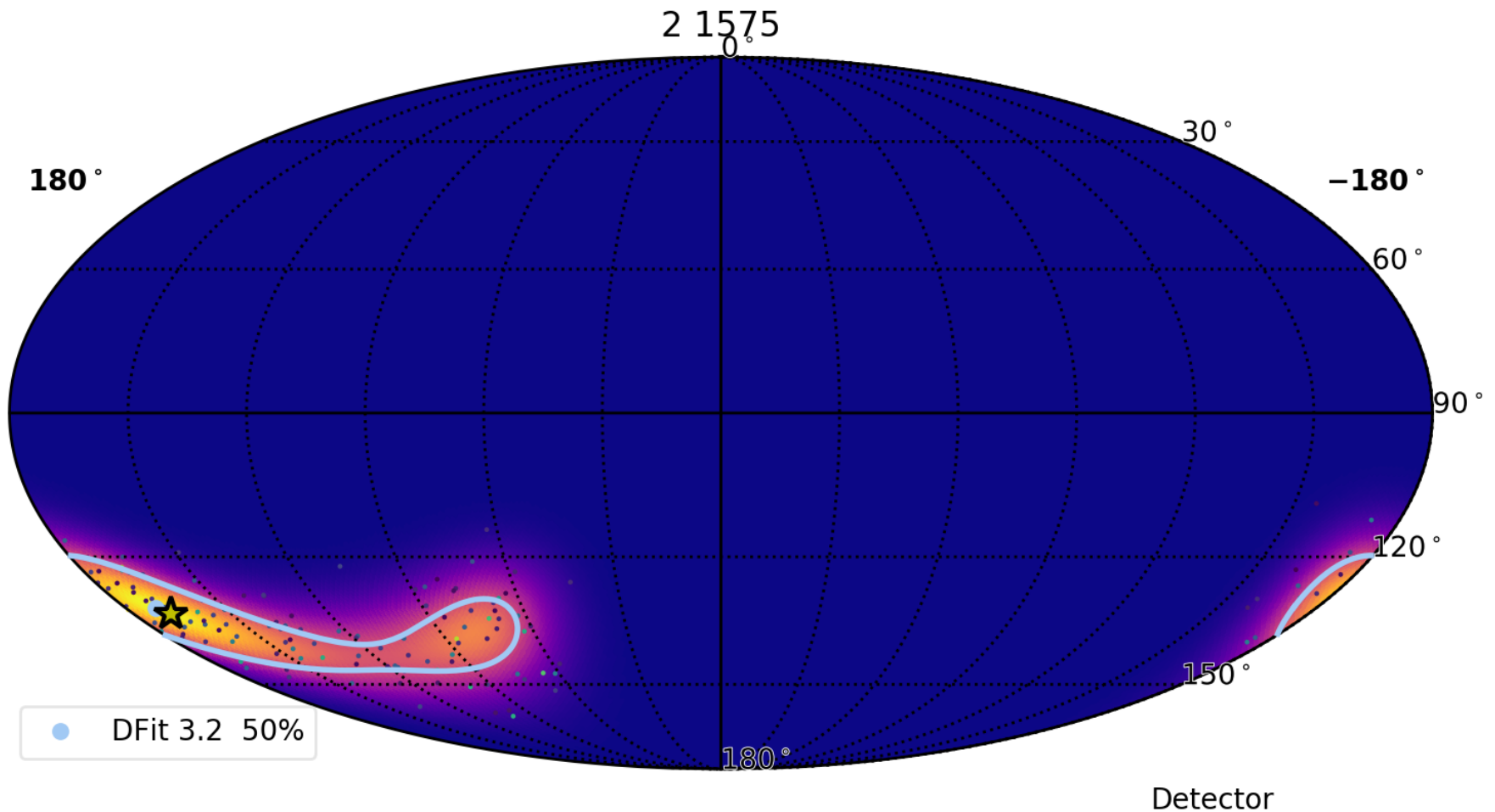
Likelihood gradually improves from start to finish

Following this, MCMC approach to sample from posterior pdf



DirectFit with directional PDFs

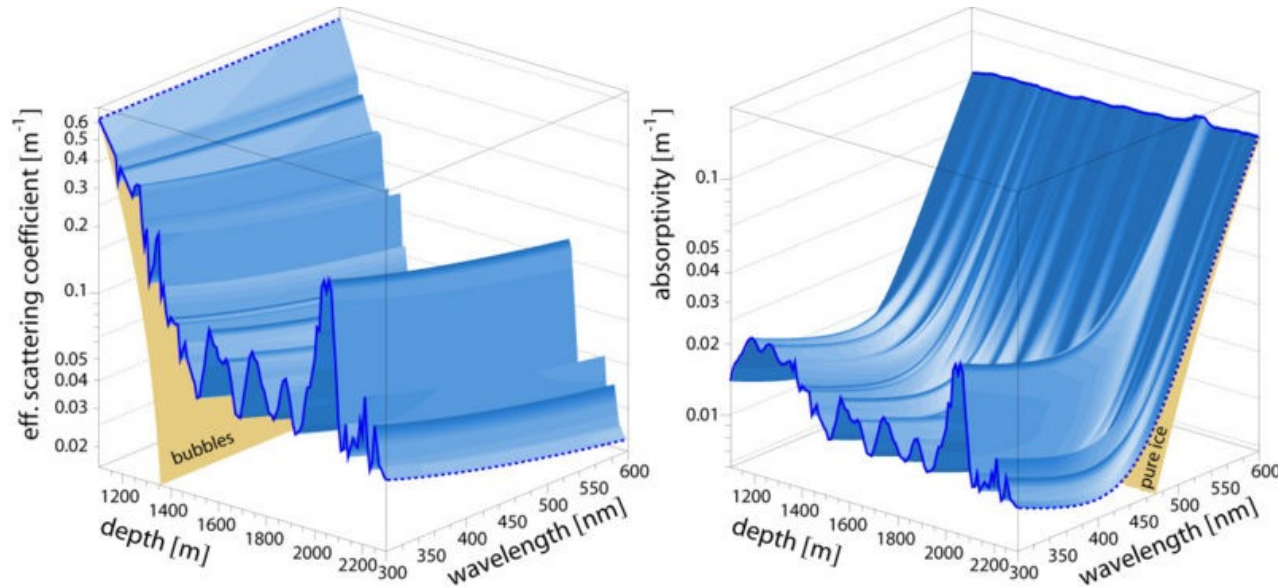
ABC outputs points on unit sphere (simulated event)



Can then fit a PDF on a sphere to those points

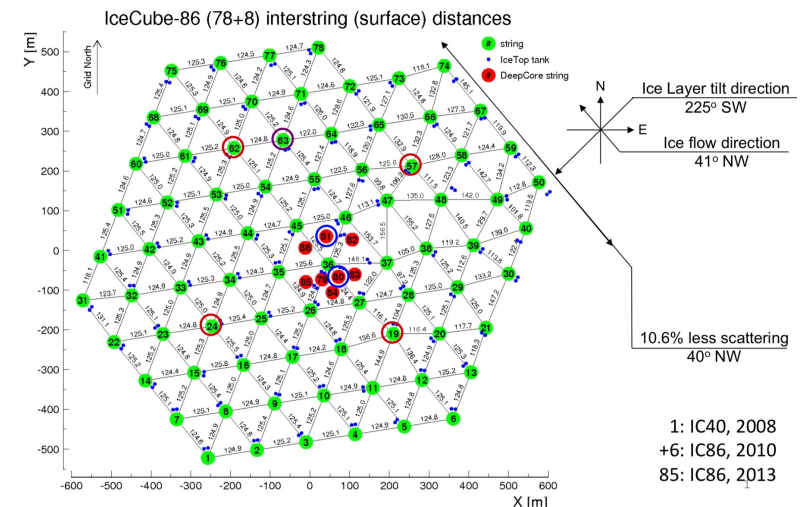
Ice modeling is important!

Bulk ice described by scattering and absorption coefficients as a function of depth → these have been refined over time



Ice layers were found to be tilted
[arXiv:1301.5361]

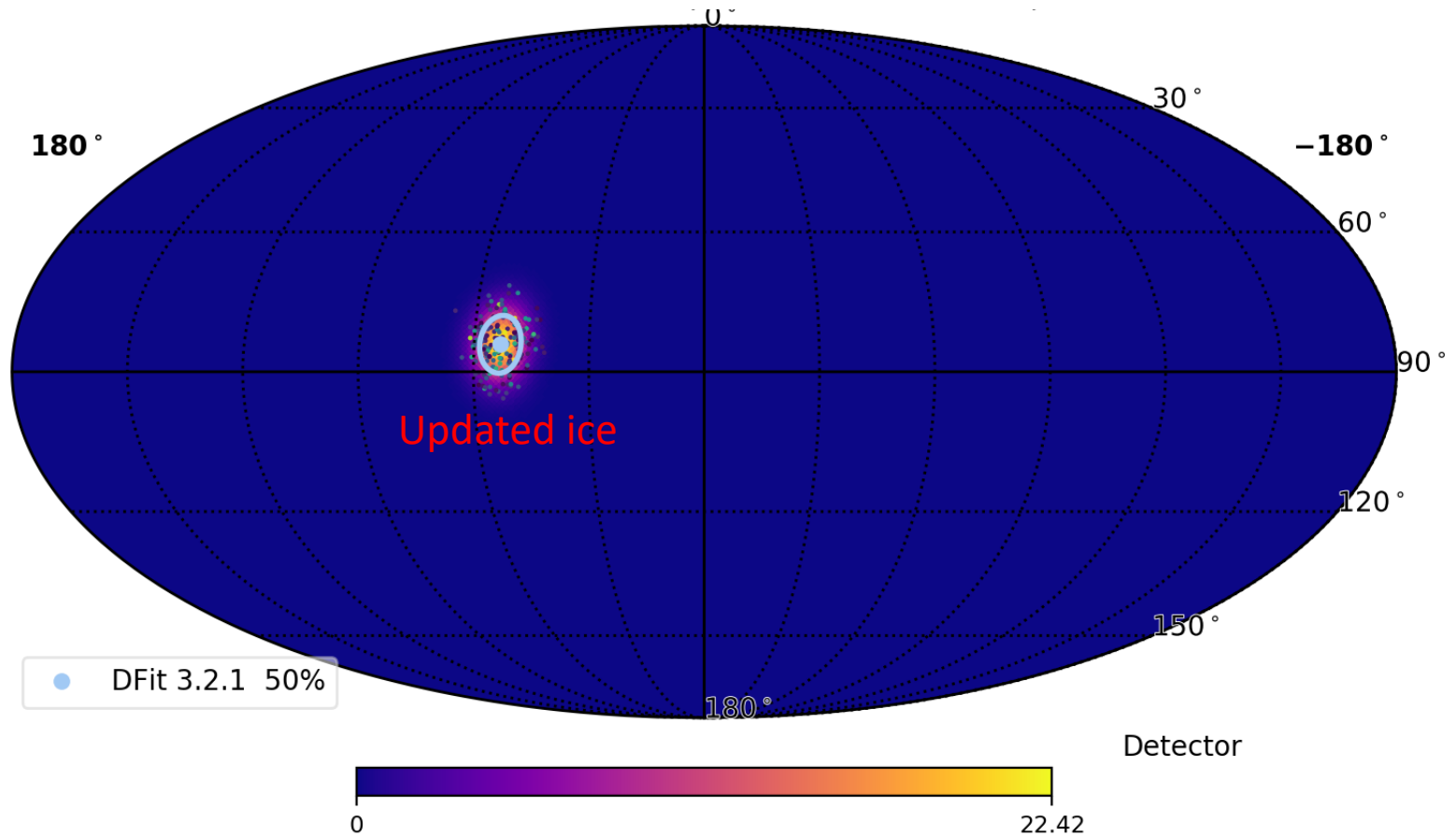
Ice was also discovered to be anisotropic
[ICRC 2013, 0580]



1: IC40, 2008
+6: IC86, 2010
85: IC86, 2013

Directional bias due to different ice models

Ice affects cascade reconstruction



Transformer normalizing flows

Use transformer architecture to model full posterior distributions with normalizing flows

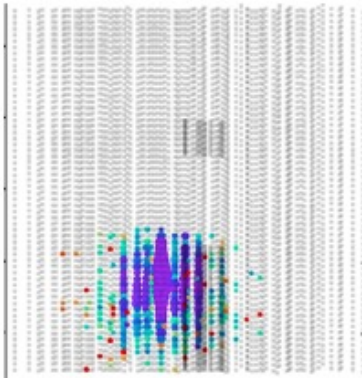
NN essentially learns joint pdf $f(\Theta, X_{\text{Data}})$ which can then be conditioned

(conditional) Normalizing flows

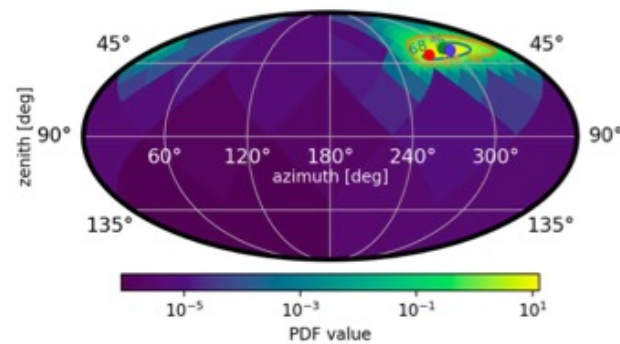
powerful machine-learned conditional PDFs from raw data

In particular **posterior distribution** - Example direction: **p(direction|data)**

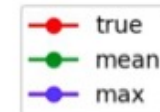
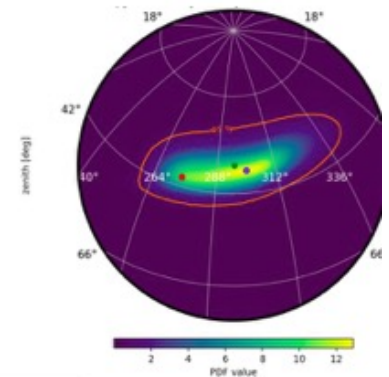
Raw pulse data



Full sky



Zoom



Summary

Reconstruction in IceCube is often a challenge

Many algorithms exist, separable into high-energy/low-energy and track/shower

- Ice modeling is most important for cascades

Traditionally LLH-based approaches; now lots of ML developments (also relevant for future extensions)

Each has pros and cons ~ymmv

New approaches always welcome!

References

SANTA: <https://doi.org/10.1016/j.astropartphys.2011.01.003>

SplineReco: <https://docs.icecube.aq/icetray/main/projects/spline-reco/index.html>

RetroReco: <https://github.com/icecube/retro>

DirectReco:

https://indico.cern.ch/event/593812/contributions/2499791/attachments/1468178/2270620/snowicki_IC_directreco_CAPtalk2017.pdf

MuEx: <https://docs.icecube.aq/icetray/main/projects/mue/muex.html>

TruncatedEnergy: https://docs.icecube.aq/icetray/main/projects/truncated_energy/index.html

Millipede: <https://docs.icecube.aq/icetray/main/projects/millipede/index.html>

DirectFit: http://icecube.wisc.edu/~dima/work/WISC/papers/2013_ICRC/dir/icrc2013-0581.pdf

FLERCNN: <https://github.com/jessimic/LowEnergyNeuralNetwork>

DNN: https://icecube.wisc.edu/~mhuennefeld/docs/dnn_reco/html/pages/about.html

EventGenerator:

https://events.icecube.wisc.edu/event/213/contributions/10356/attachments/7984/10418/2024_09_24_IceCube_Madison.pdf

TNF:

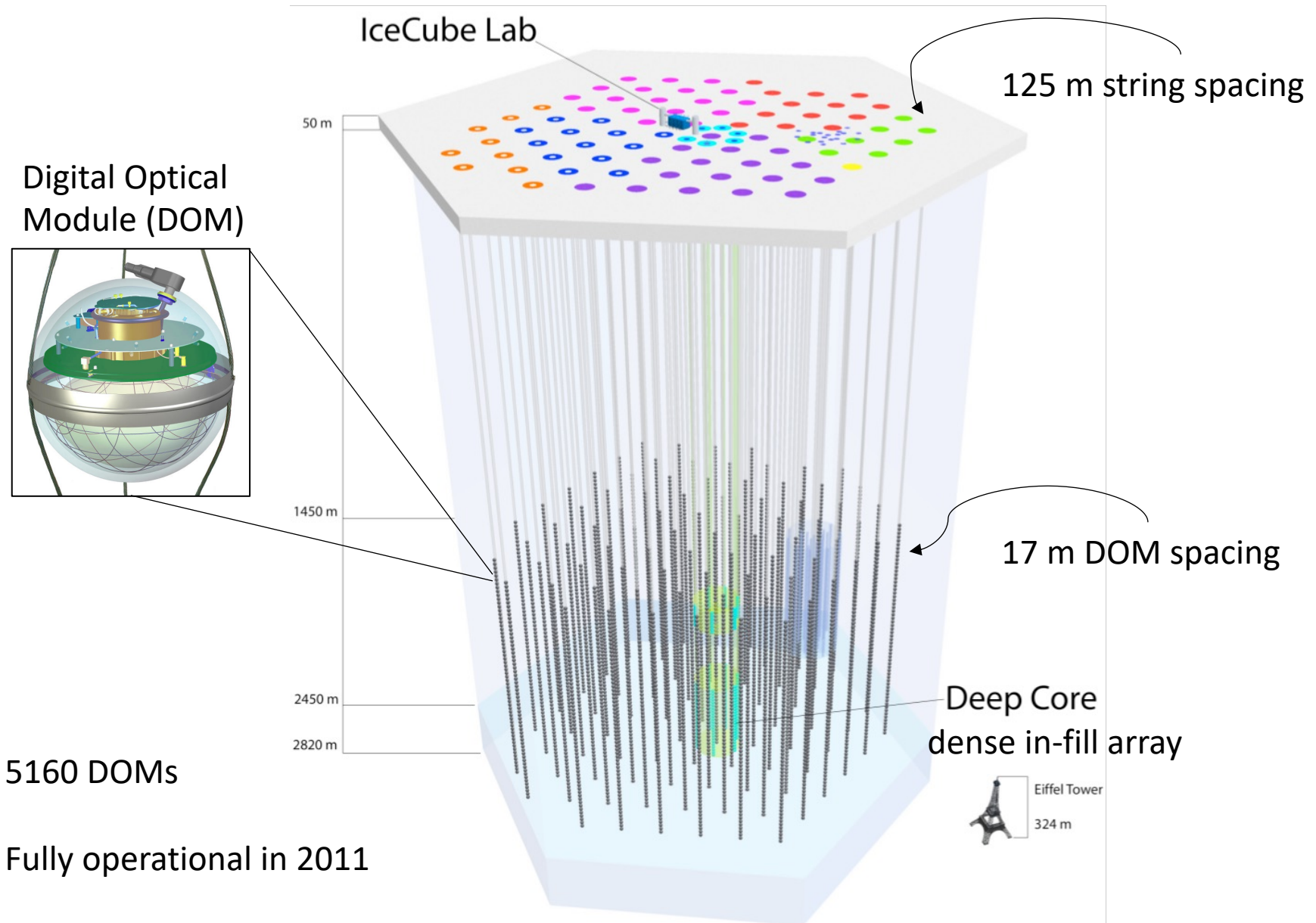
https://events.icecube.wisc.edu/event/180/contributions/10010/attachments/7810/10142/icecube_muenster_talk.pdf

FreeDOM:

https://events.icecube.wisc.edu/event/125/contributions/7228/attachments/5679/6634/fienberg_freeDOM_plenary.pdf

Backups

IceCube



Detection principals

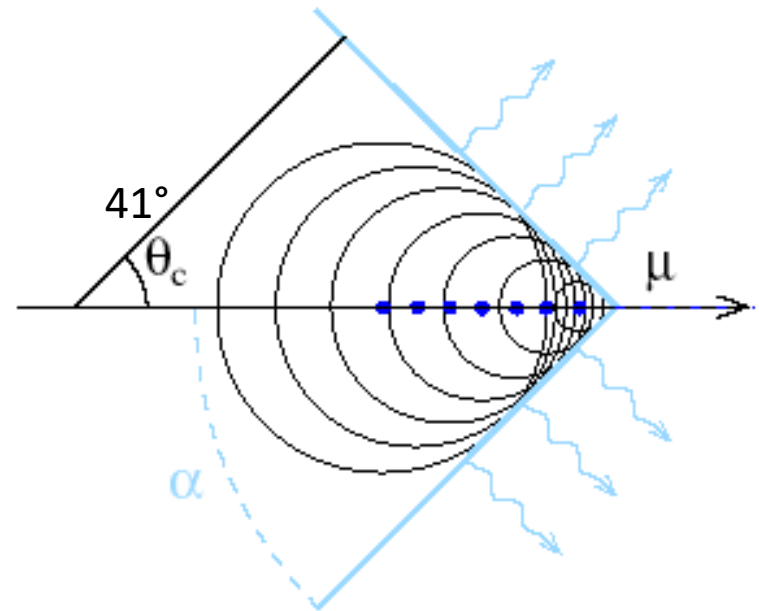
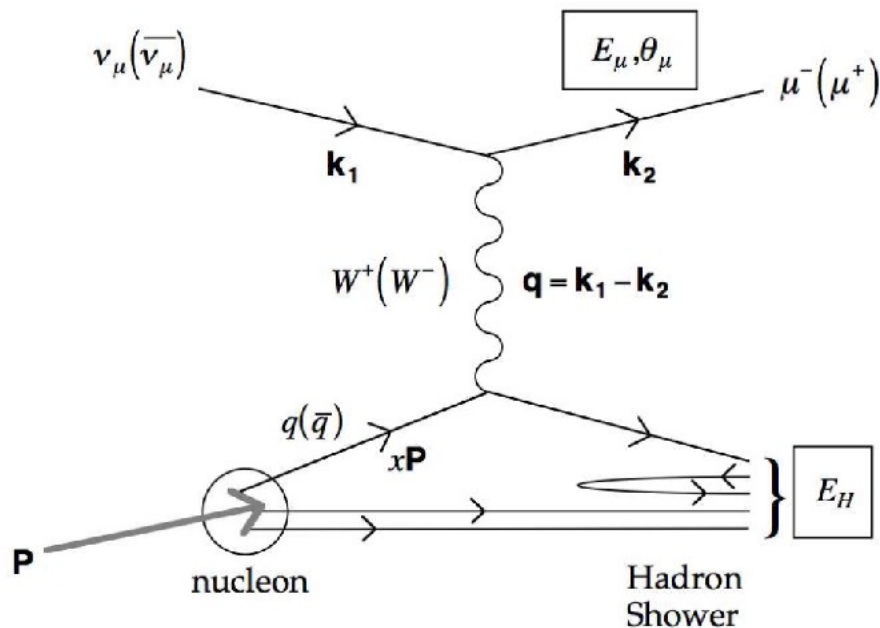
Neutrino interacts via weak force with targets in ice

- At IceCube energies, primarily deep-inelastic scattering (DIS) off nucleons

Nucleon breaks apart; outgoing particles may be charged

Charged particles emit **Cherenkov radiation** detectable by PMTs

Rev. Mod. Phys. 84, 1307



Cherenkov radiation

Occurs when a charged particle travels faster than light-in-medium

Constructive interference of EM-field to form a plane wave

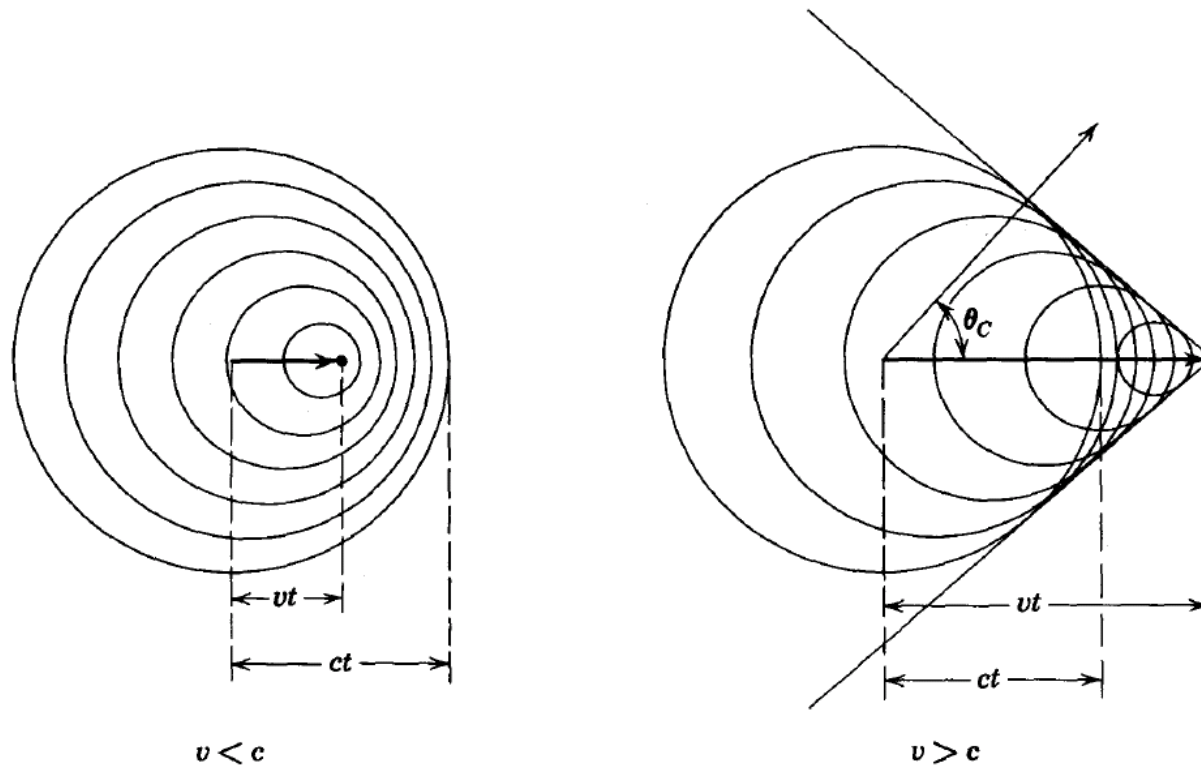


Fig. 14.14 Cherenkov radiation. Spherical wavelets of fields of a particle traveling less than, and greater than, the velocity of light in the medium. For $v > c$, an electromagnetic “shock” wave appears, moving in the direction given by the Cherenkov angle θ_c .

Local effects

Hole-ice

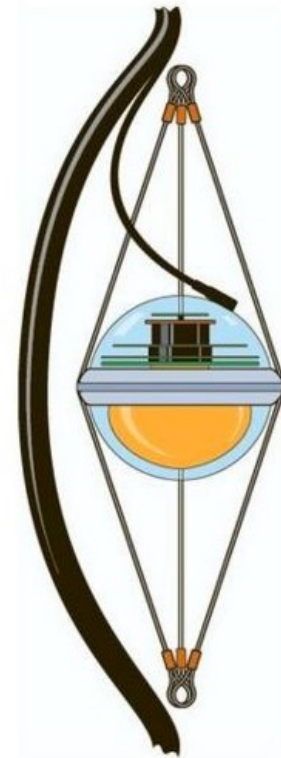
- Refrozen central column with high scattering

Looking up the string



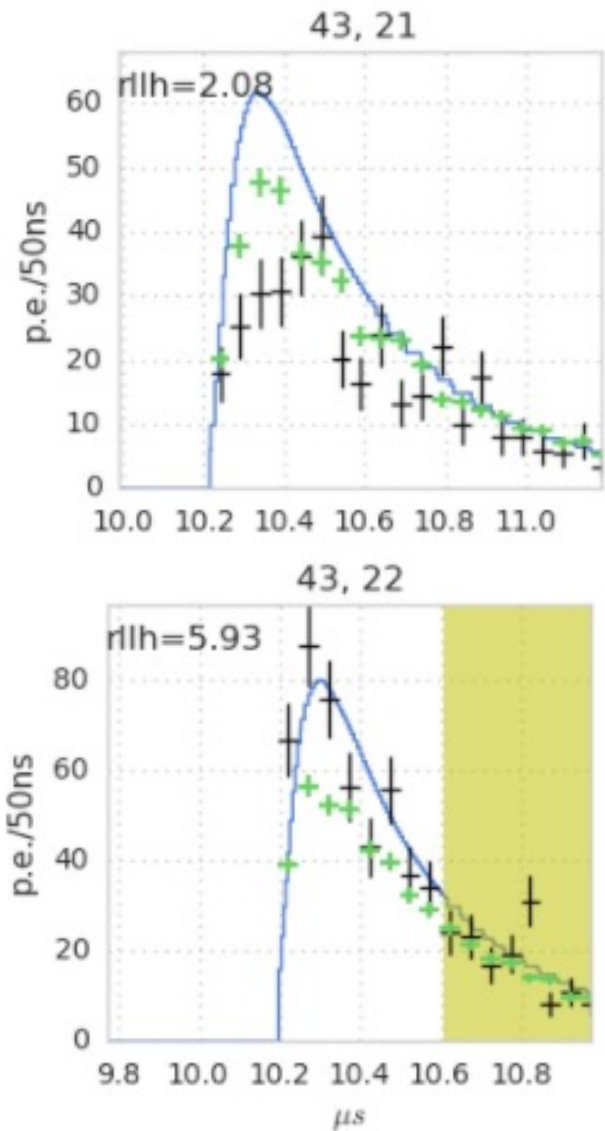
DOM orientation

- Thick, support cable may impede direct photons if vertex is nearby
- A few DOMs may not be perfectly horizontal

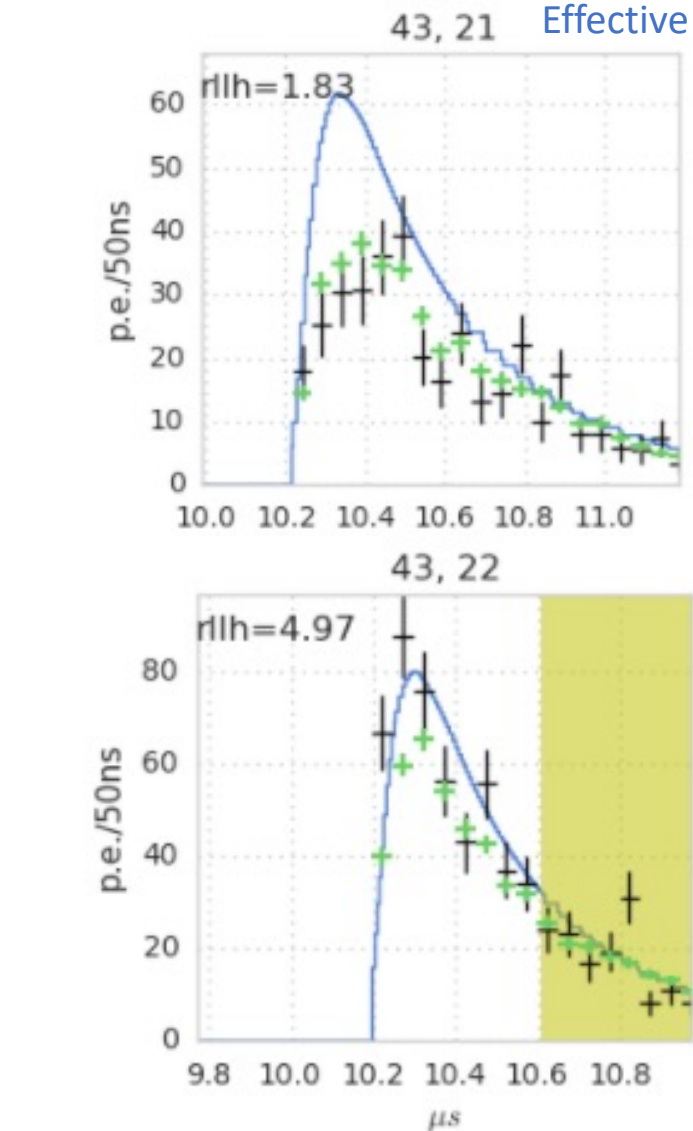


Local effects: DOM orientation and cable position

Without local effects

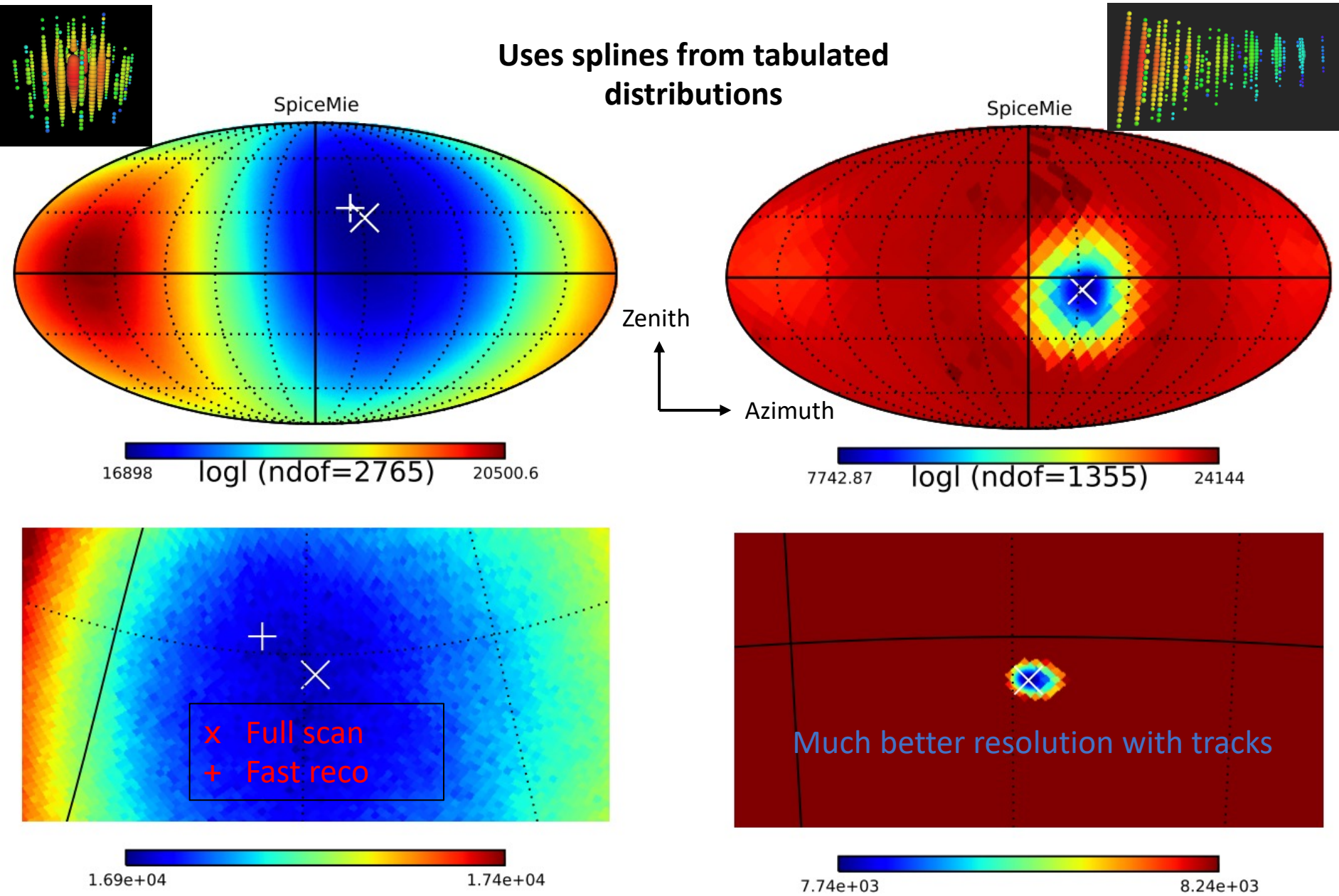


With local effects



Bert data
Direct photon MC
Effective photon MC

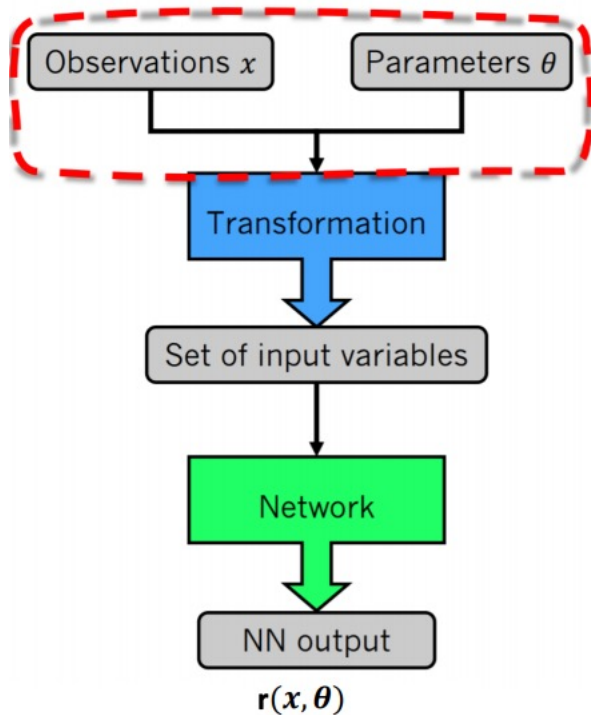
Cascade vs track skymap



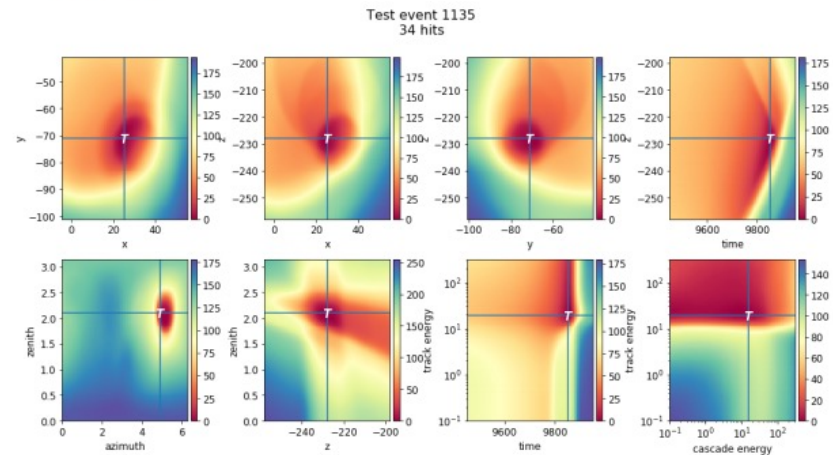
FreeDOM

Likelihood-free inference using NN

Train a binary classifier that can be converted back into a likelihood



- We replace $\frac{\mathcal{L}(\theta|x)}{p(x)}$ with the output of our neural network, $r(x, \theta)$
 - r is a ratio estimator; approximates the likelihood-to-evidence ratio
- $r(x, \theta)$ can be used anywhere you'd typically use a likelihood function
- Evaluating $r(x, \theta)$ is very fast (tens of microseconds)



T: true parameters

eight hypothesis parameters:

x , y , z , t , azimuth, zenith, cascade energy, track energy

parameters not being scanned are set to their truth values