Diffuse Astrophysical Neutrinos: Analysis Techniques

Vedant Basu
Postdoctoral Researcher, IceCube

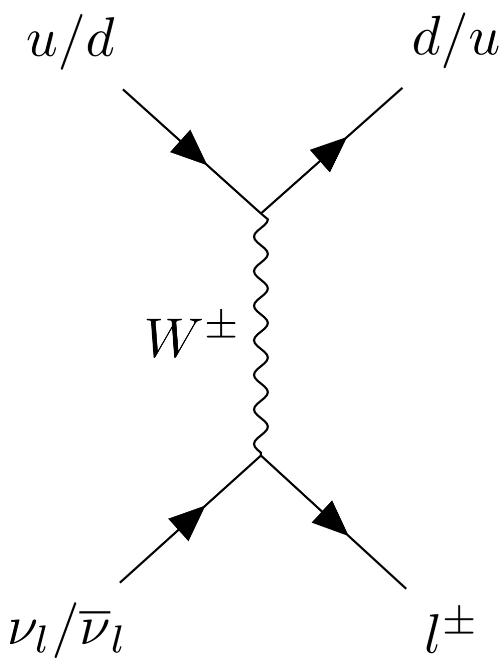
Harvard Neutrino Practicum, 2026

PC: Martin Wolf, IceCube/NSF

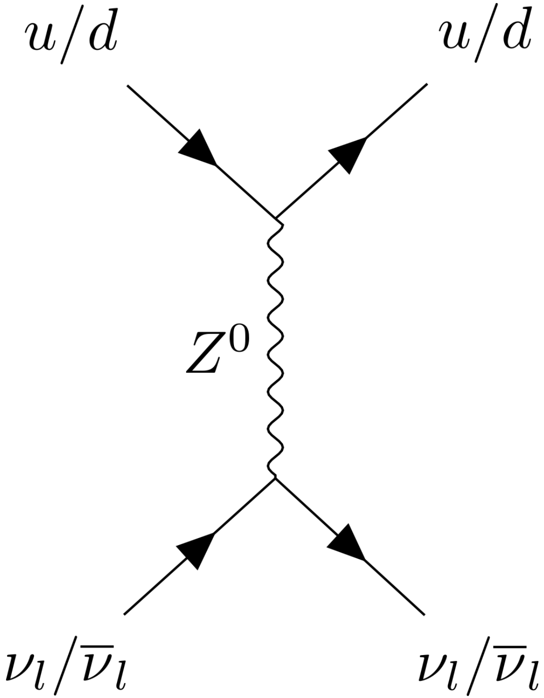
Outline

- **Astrophysical Neutrino Detection at Cherenkov Detectors**
- Event Classification and Selection Techniques
- Neutrino Spectral Measurements
- Fitting a neutrino spectrum!
- Outlook

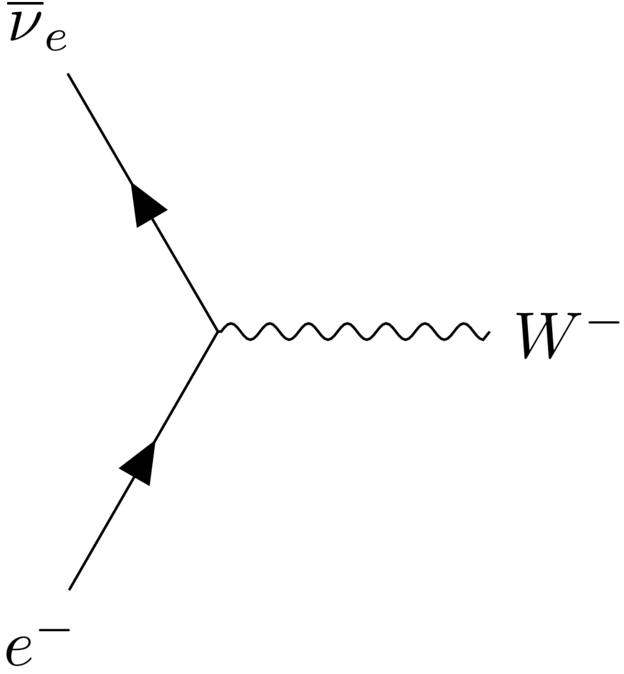
Neutrino Interaction Channels



Charged current interactions yield an outgoing charged lepton and a hadronic shower



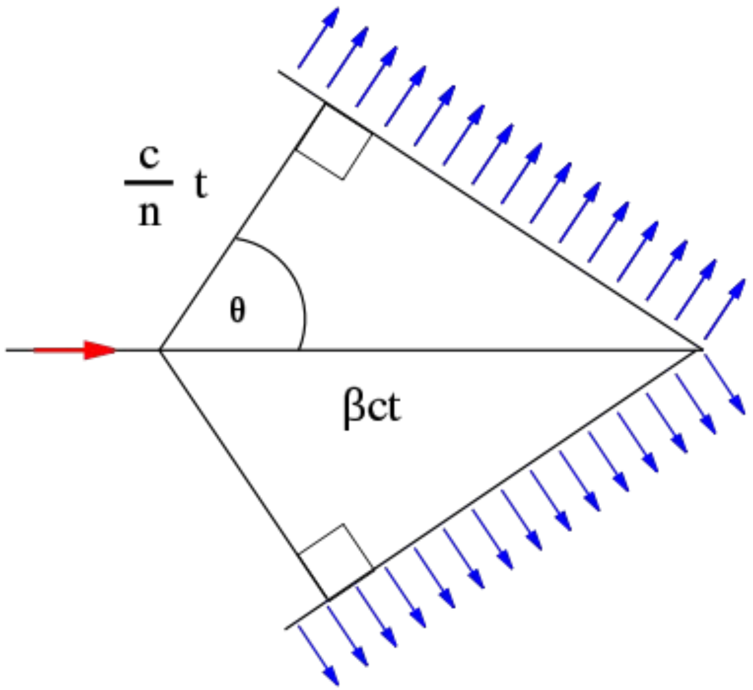
Neutral current interactions have an outgoing scattered hadron and neutrino



Glashow Resonance interactions yield an on-shell W^- boson, which subsequently decays

Observing neutrino interactions via the Cherenkov effect

- **Charged particles deposit energy** as they travel through matter
- When charged particles travel faster than the speed of light in a medium, they emit Cherenkov radiation
- A **cone of light is formed** determined by Cherenkov angle, θ_C
- Cherenkov light collected by DOM photomultiplier tubes (PMTs), which convert the optical signal to an amplified electrical signal

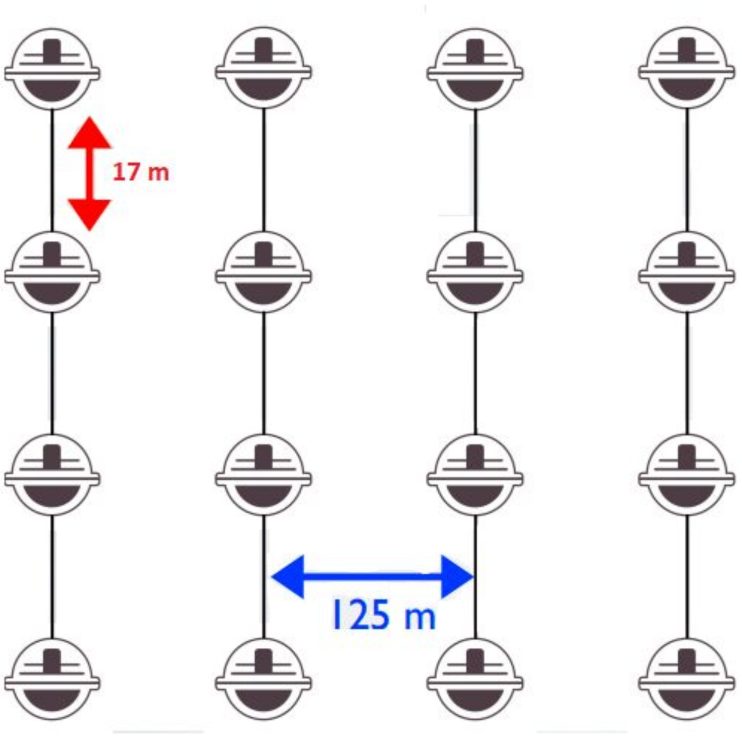


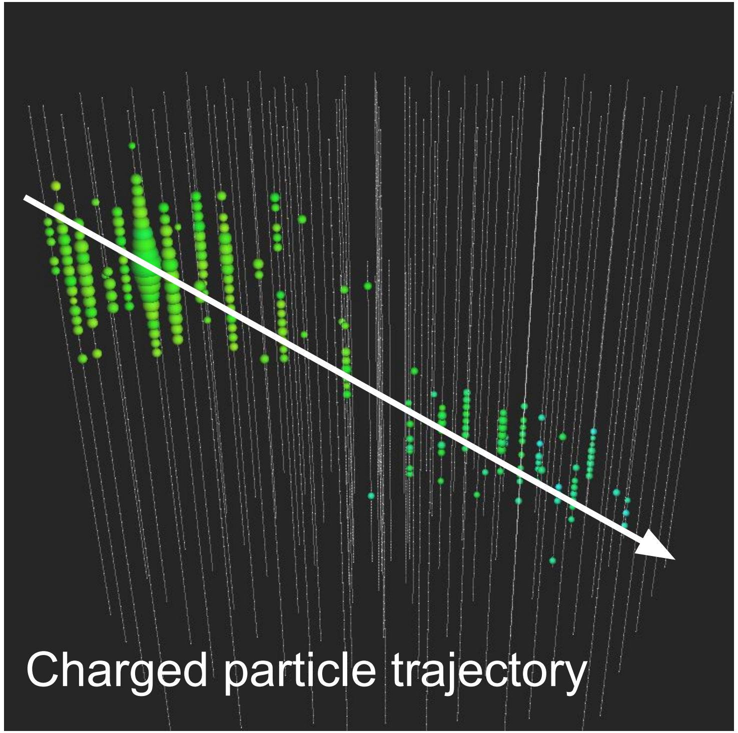
In ice: $\theta_C = 40.2°$

Wikimedia Commons

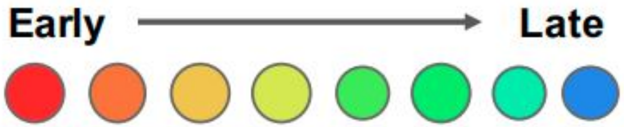
Event Reconstruction

- Number of photons captured ∝ deposited energy. Photon arrival times at DOMs helps reconstruct particle trajectory.

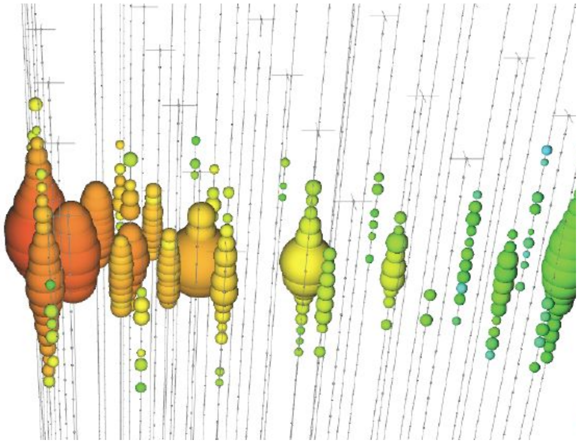




IceCube Event Morphologies



Tracks (Data)

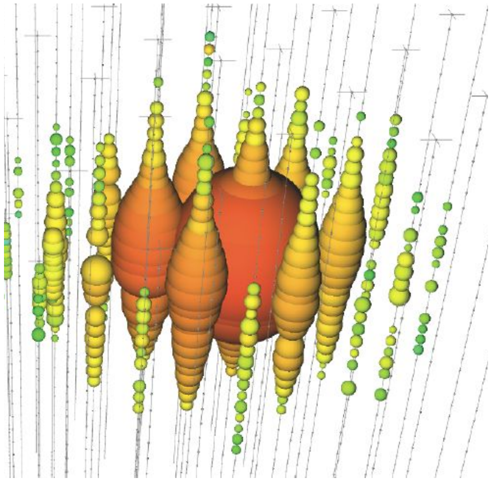


$$\nu_\mu + \mathrm{X} \rightarrow \mu^- + \mathrm{Hadrons}$$

$$\mathrm{Energy\ Res.} : 0.3 \mathrm{\ in\ Log} \frac{\mathrm{E}}{\mathrm{TeV}}$$

$$\mathrm{Angular\ Res.} : 0.3^\circ$$

Cascades (Data)



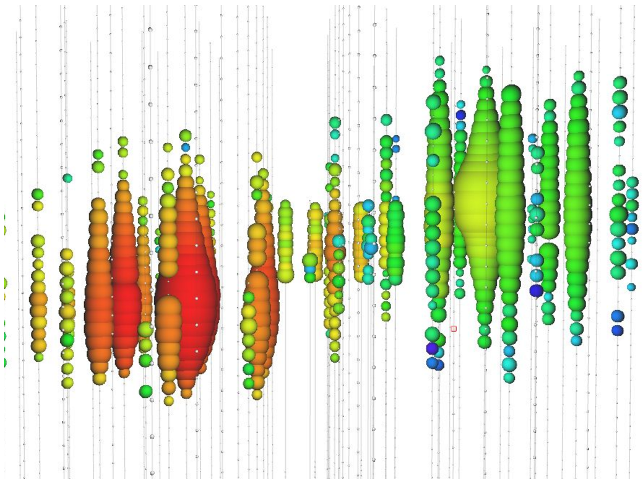
$$\nu_e + \mathrm{X} \rightarrow e^- + \mathrm{Hadrons}$$

$$\nu_l + \mathrm{X} \rightarrow \nu_l + \mathrm{Hadrons}$$

$$\mathrm{Energy\ Res.} : 8\%@100\mathrm{TeV}$$

$$\mathrm{Angular\ Res.} : 4^\circ$$

Double Cascade (Simulated)



$$\nu_\tau + \mathrm{X} \rightarrow \tau^- + \mathrm{Hadrons}$$

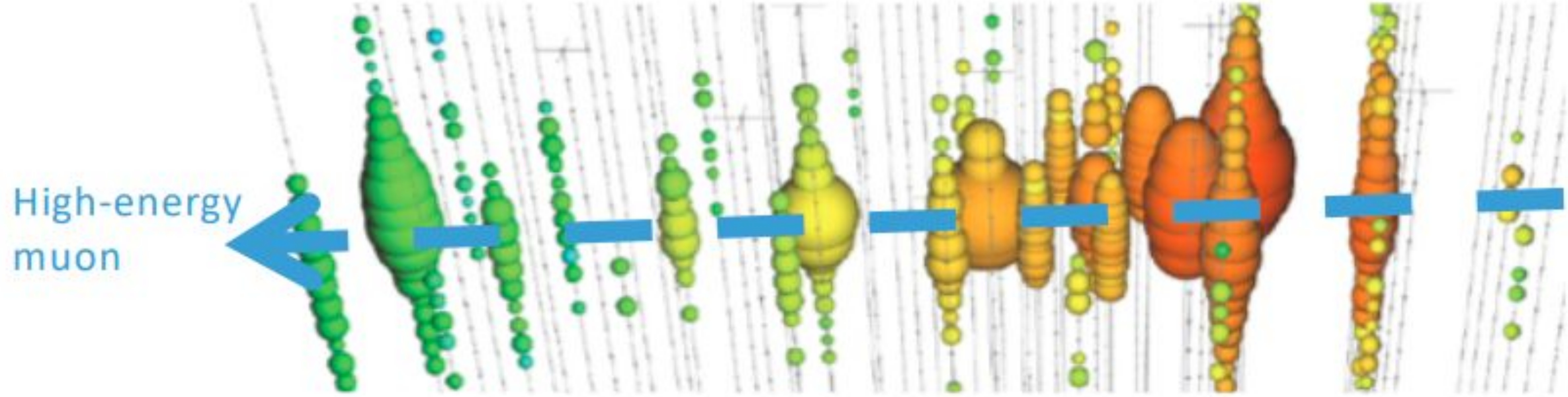
$$\tau^- \rightarrow \nu_\tau + \ldots$$

$$\mathrm{Decay\ Length} : 50\mathrm{m/PeV}$$

Ref.

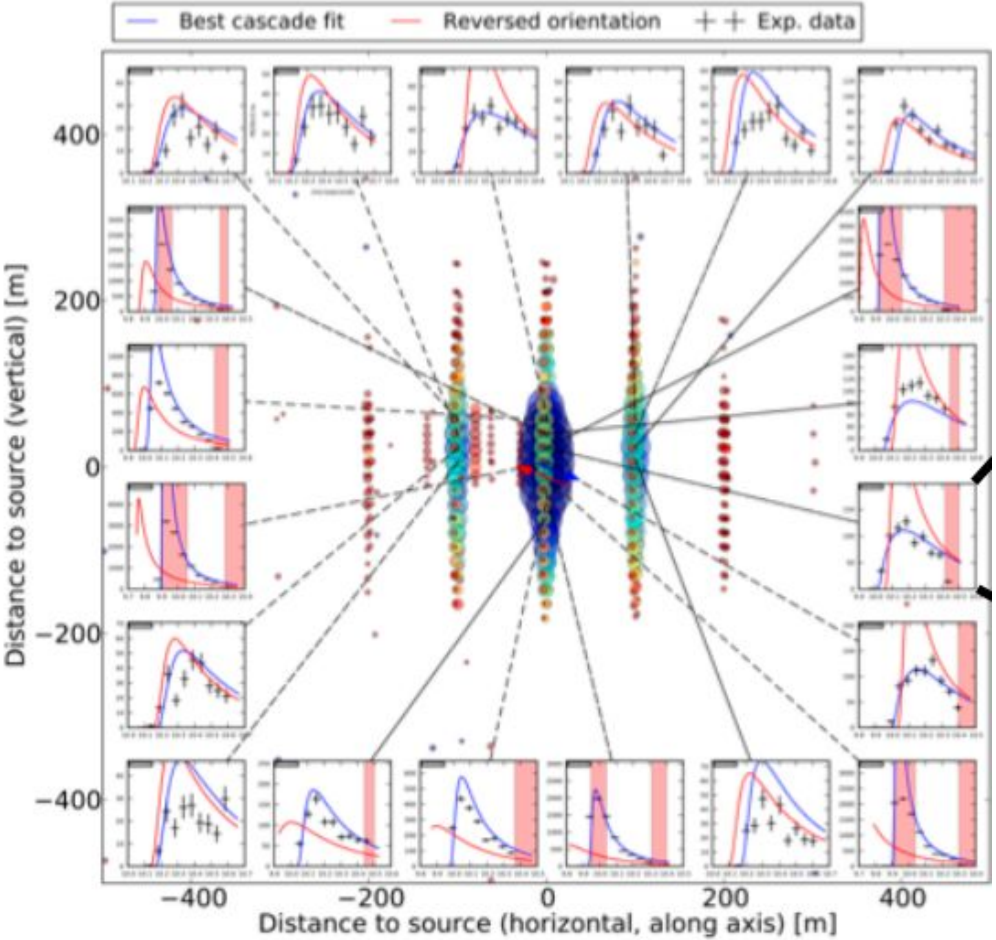
Track Reconstructions

- Use full-waveform information by fitting predicted light yields to what is actually seen
- IceCube's Millipede works for high-energy tracks by breaking it up into multiple cascades along the track due to muon stochastic energy losses



Cascade Reconstructions

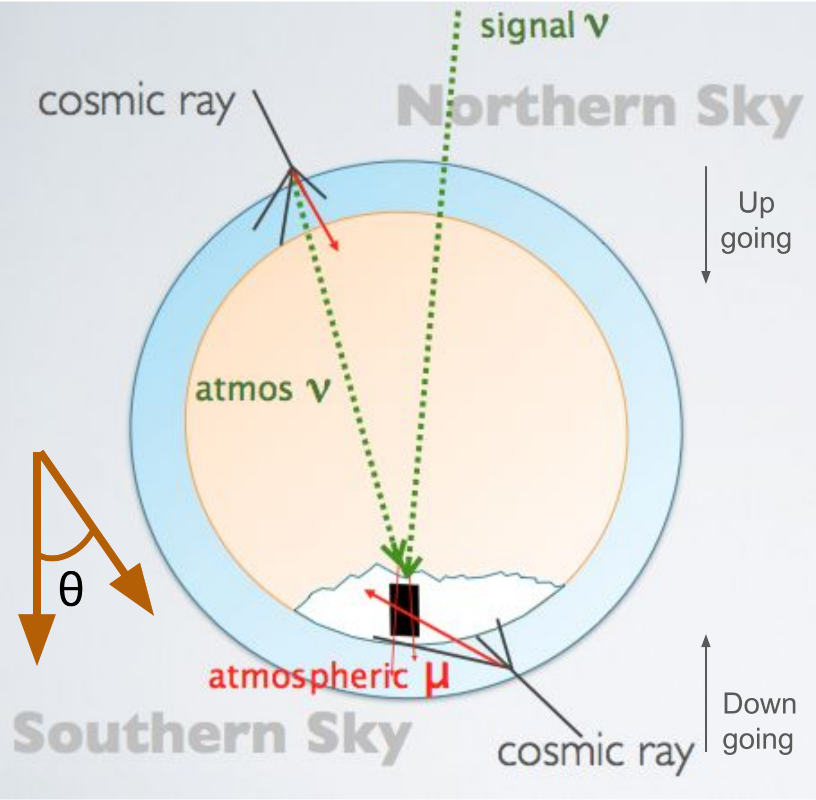
- Large distances between DOMs mean few detected photons
- Sensitivity to asymmetry means high dependence on ice modeling
- IceCube's Monopod algorithm tabulates photon yields for a single ice model

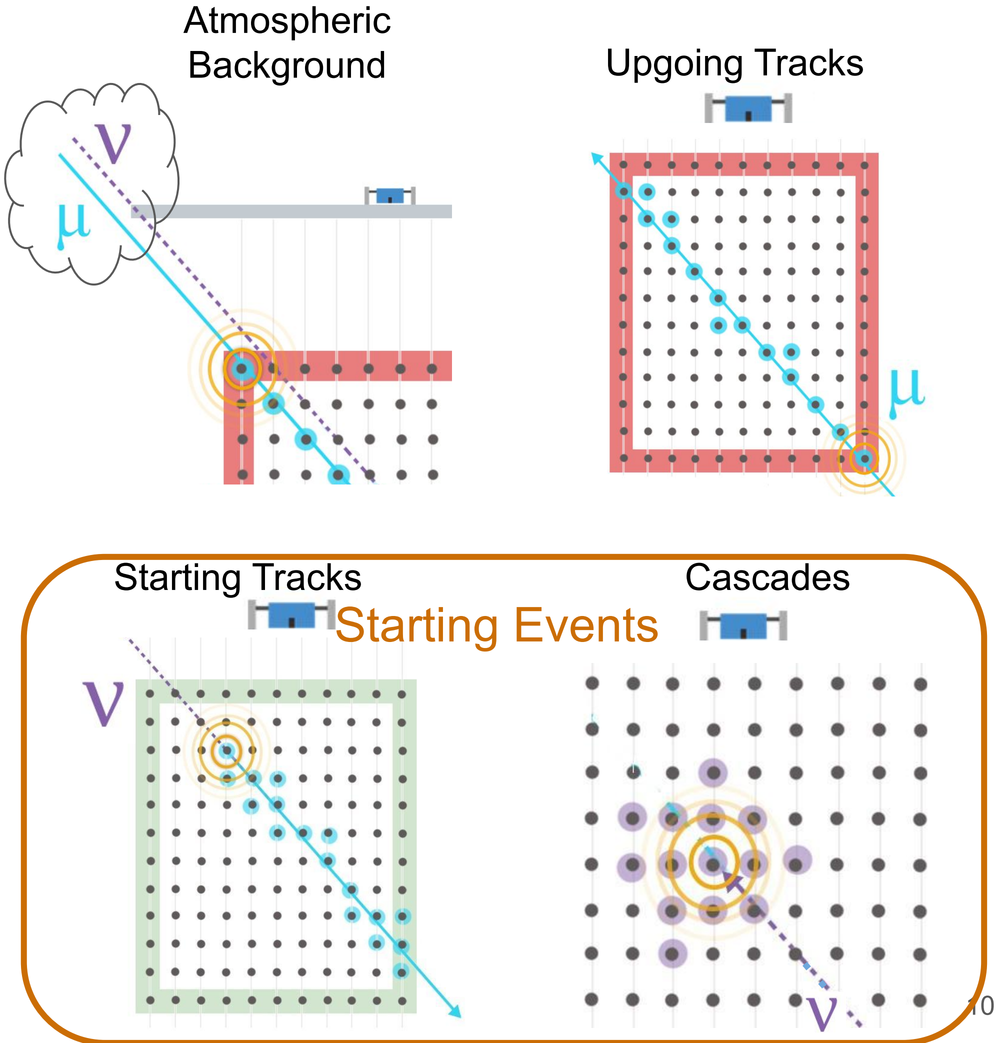


Outline

- Astrophysical Neutrino Detection at Cherenkov Detectors
- **Event Classification and Selection Techniques**
- Neutrino Spectral Measurements
- Fitting a neutrino spectrum!
- Outlook

Event Selections

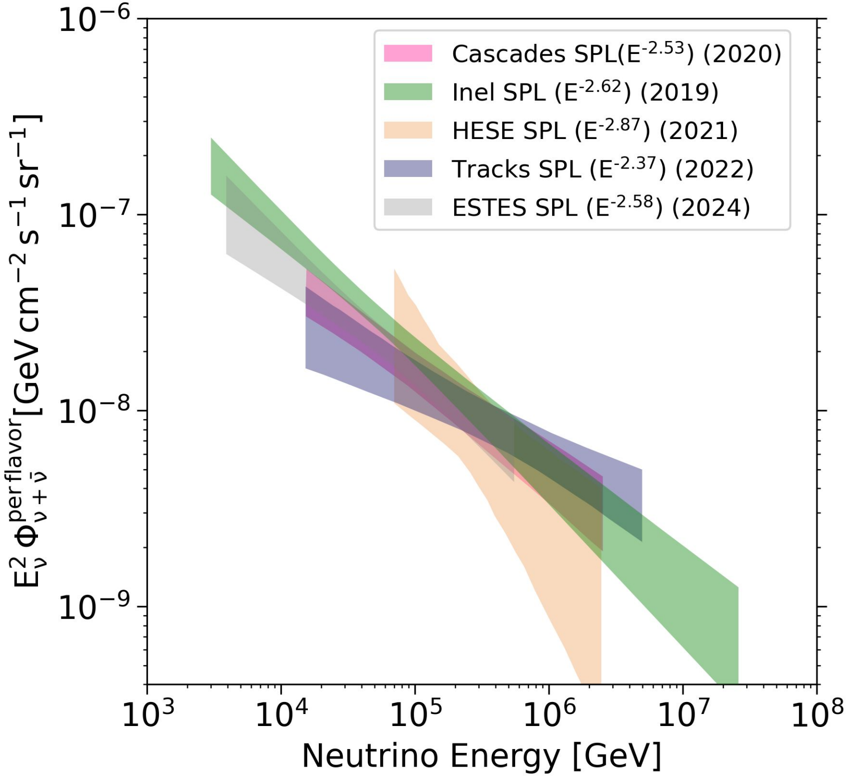




Diffuse Cosmic Neutrino Flux

IceCube measurements of the diffuse spectrum prior to 2026 found a single power law ($E^{-\gamma}$) to be the best fit

- High Energy Starting Event (HESE): all flavours & all sky, above 60 TeV
- Through-going tracks: muon neutrinos from the Northern sky
- Cascades: All sky electron and tau neutrinos, and muon neutrino NC interactions
- Enhanced Starting Track Event Selection (ESTES): All sky muon neutrinos

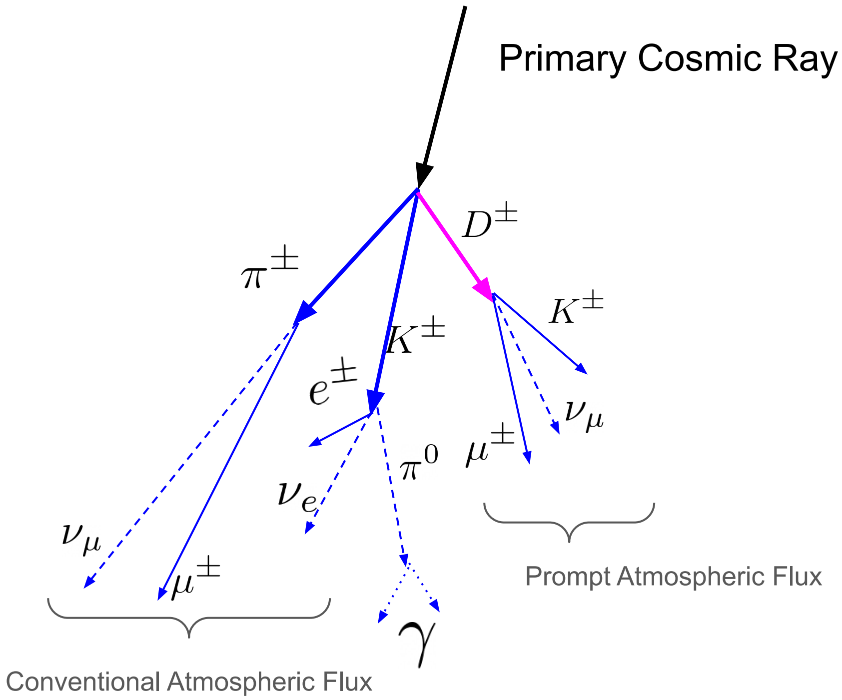


Outline

- Astrophysical Neutrino Detection at Cherenkov Detectors
- Event Classification and Selection Techniques
- **Neutrino Spectral Measurements**
- Beyond Astronomy
- Outlook

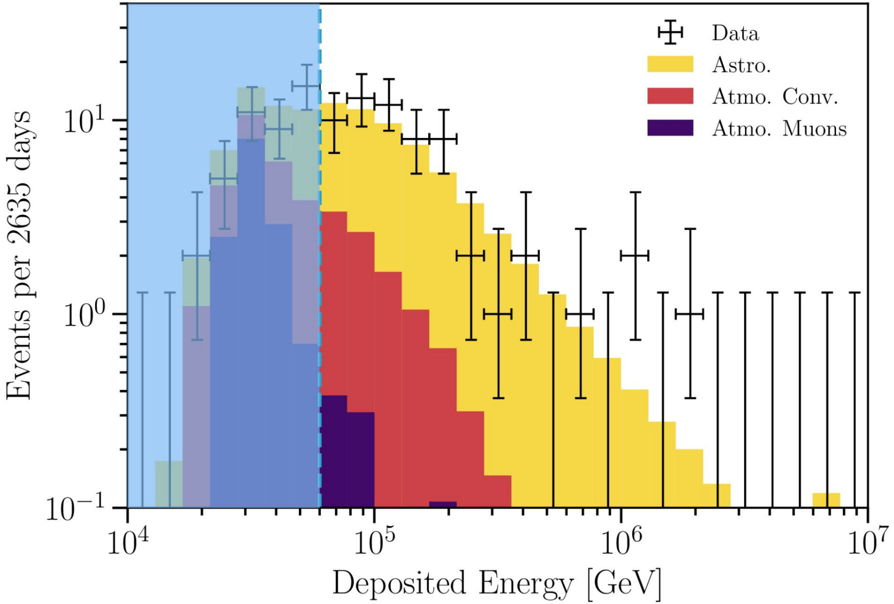
Atmospheric Background

- Main background from atmospheric muons
- Neutrinos also produced in atmospheric cosmic rays interactions.
- Conventional flux from pion/kaon decay follows ~$E^{-3.7}$
- Prompt contribution from charmed meson decay- follows ~$E^{-2.7}$
- Q: Why do the conventional and prompt fluxes have different spectra?

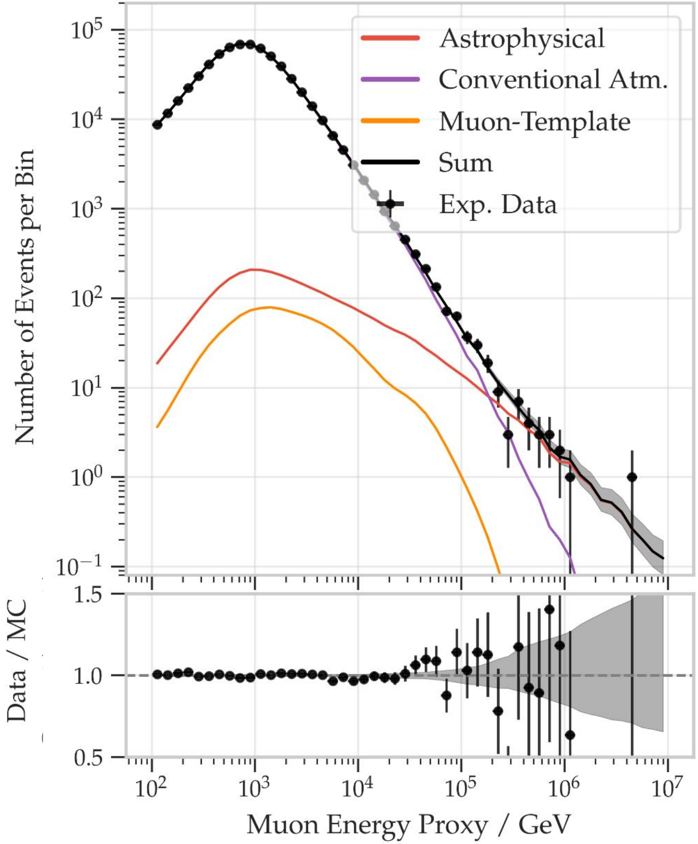


Low energy sensitivity

- The search for structure in the astrophysical neutrino spectrum necessitates pushing the energy threshold lower.
- Tradeoff: Better background rejection/modelling



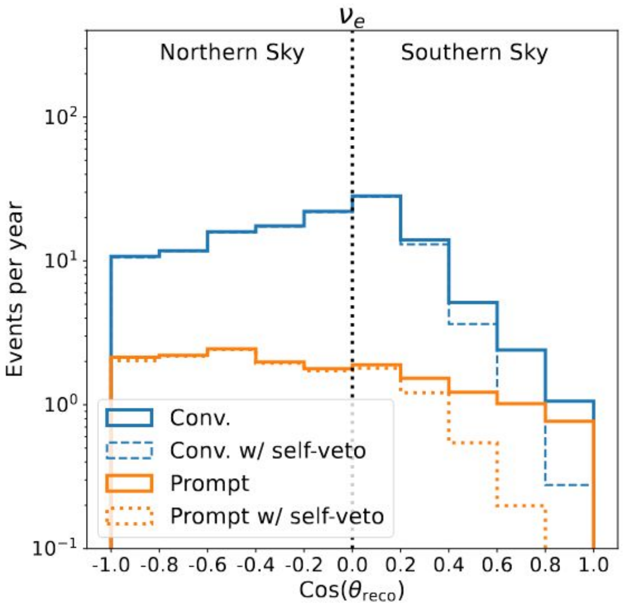
HESE flux characterization, 7.5 years of data
(events below 60 TeV (shaded) ignored in fit)
IceCube HESE: PRD 104.2 (2021): 022002.

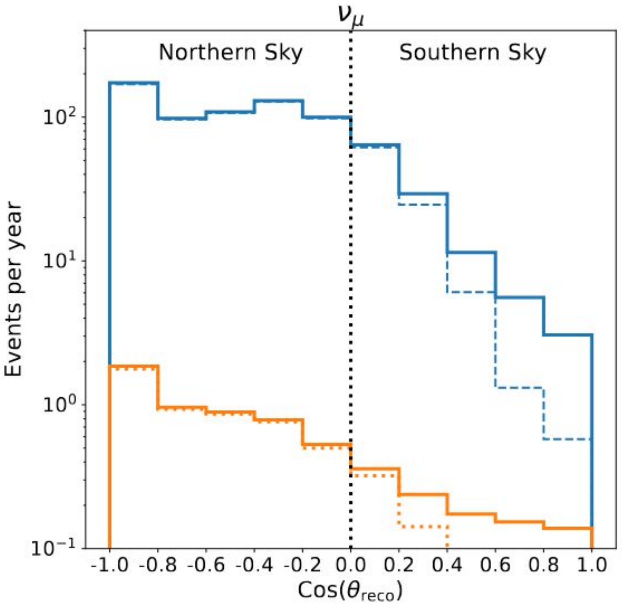


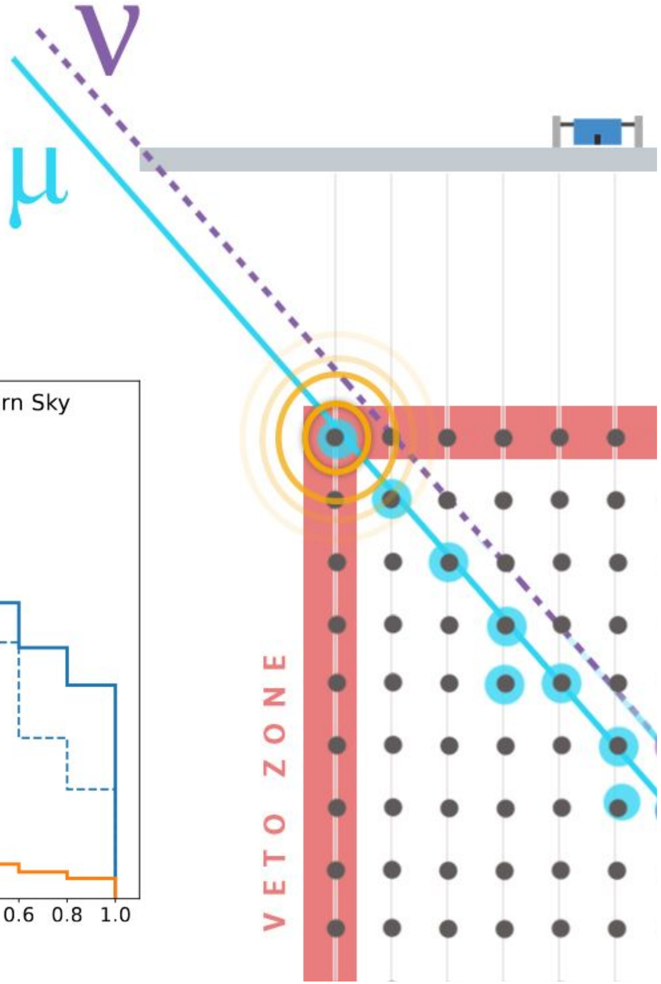
The energy distribution of upgoing tracks for the period 2010-18
The Astrophysical Journal 928.1 (2022): 50

Neutrino self-veto

- Neutrinos from CR showers often accompanied by muons.
- Vetoing these muons suppresses atm. neutrino background
- Accurate modeling of the self-veto suppression via muon bundle injection.

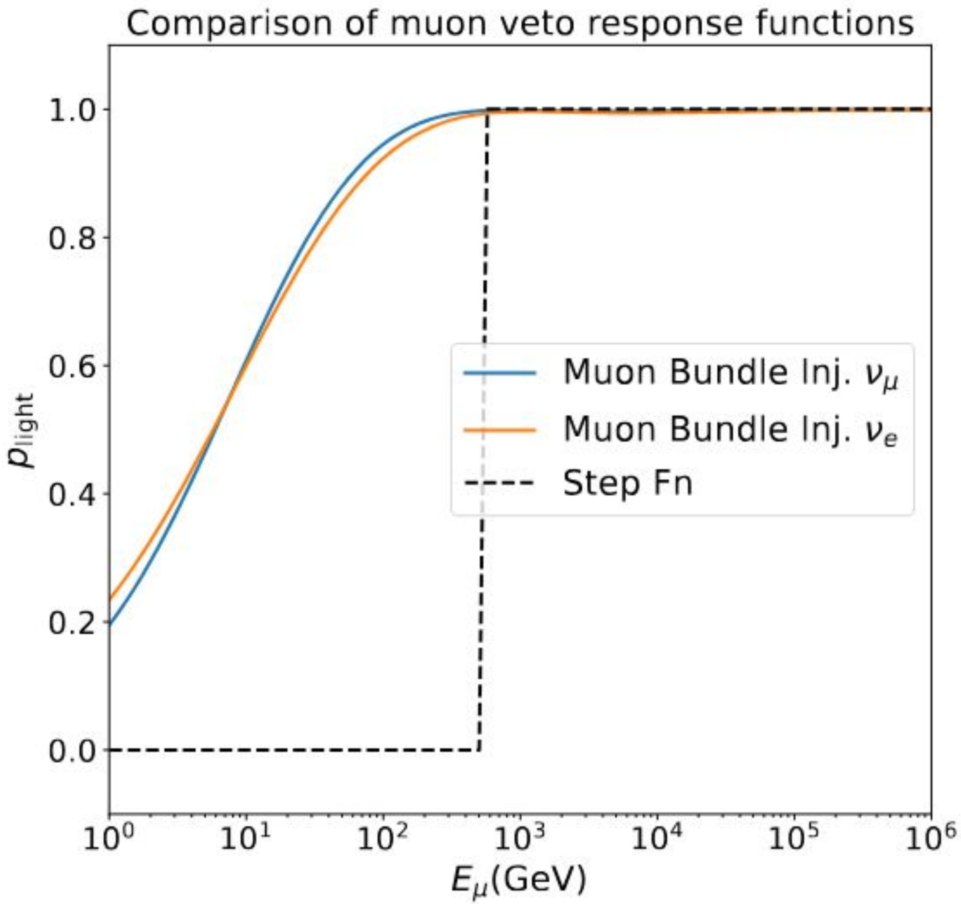






How do we model it?

- Need to model detector response to air shower muons
- Generate PDFs of Muon Bundle multiplicity and energy from Corsika MC
- Inject these muons into neutrino events which survive to final level
- Parametrise detector response by reprocessing through event selection, using the P_{light} function

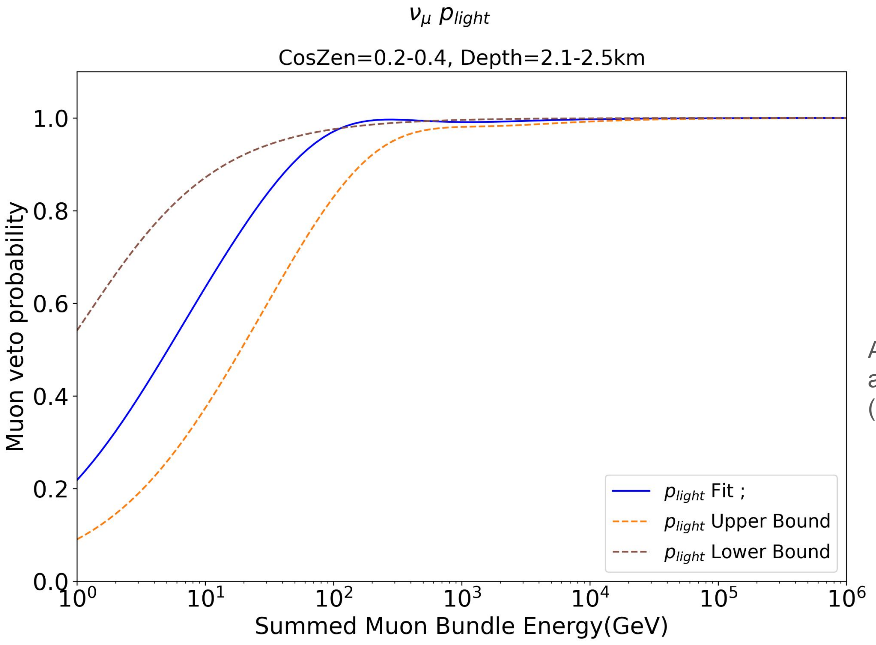


$P_{light}(E_\mu)$: fraction of events rejected by addition of accompanying muons

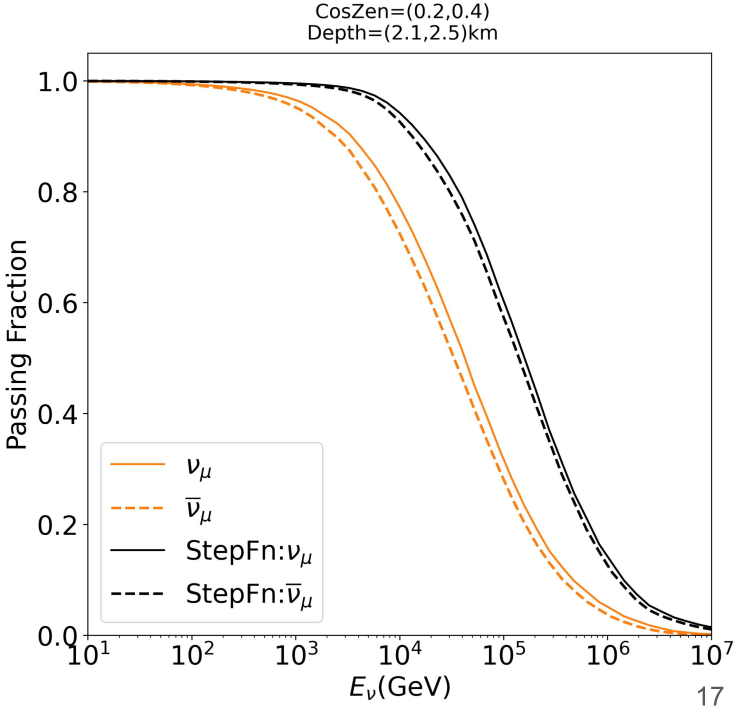
Neutrino passing fractions

- The muon detector response (**$P_{light}(E_\mu)$**) evaluated using fraction of neutrino+muon events which survive reprocessing.

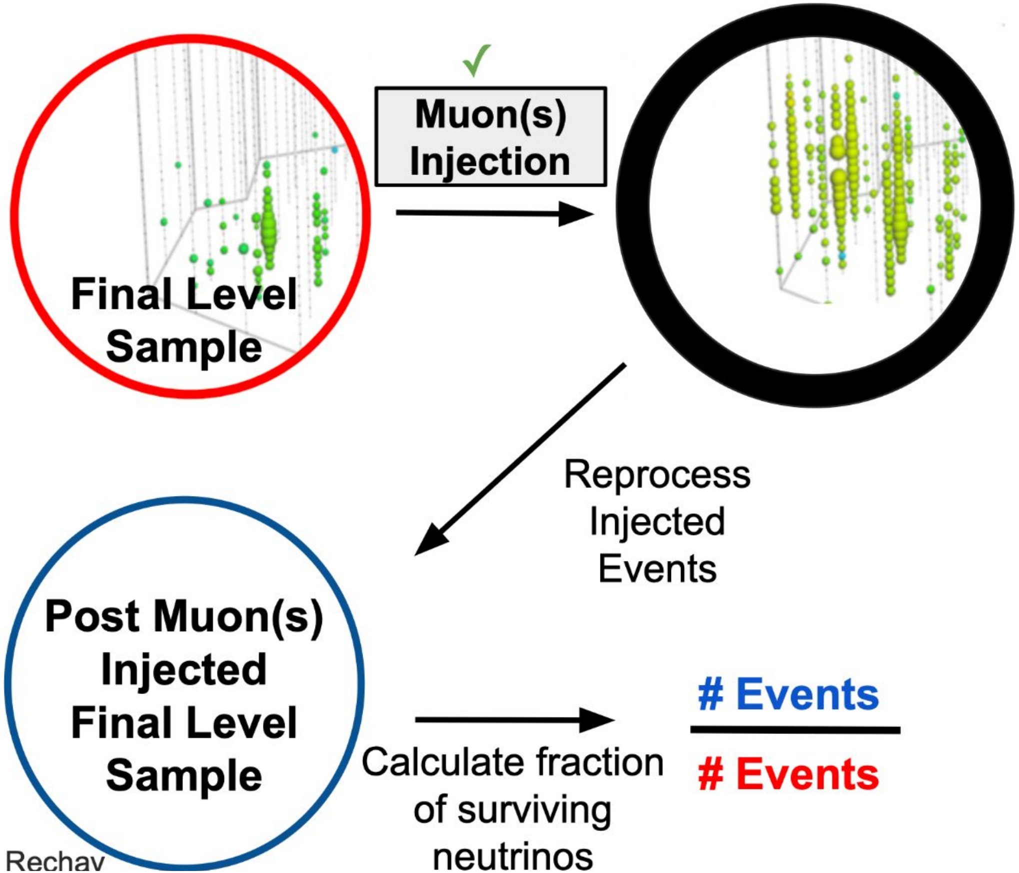
- **P_{light}** is used to calculate **neutrino** passing fractions with the nuVeto package



Argüelles, C. A., et al, JCAP 2018.07 (2018): 047.



Muon Bundle Injection



$P_{survive}$ = fraction of neutrino to survive selection with muon(s)

- # of neutrino events at final level
- # of neutrino events with injected muons at final level

Astrophysical Flux Models

- Single Power Law (SPL):

$$\Phi_{\text{astro}}\left(\frac{E_\nu}{100\text{TeV}}\right)^{-\gamma}$$

- Broken Power Law (BPL):

$$\Phi_{\text{astro}}\left(\frac{E_\nu}{E_{\text{break}}}\right)^{-\gamma_{\text{BPL}}}\left(\frac{E_{\text{break}}}{100\text{TeV}}\right)^{-\gamma_1}$$

$$\gamma_{\text{BPL}} = \begin{cases} \gamma_1\ (\mathrm{E}_\nu < E_{\text{break}}) \\ \gamma_2\ (\mathrm{E}_\nu > E_{\text{break}}) \end{cases}$$

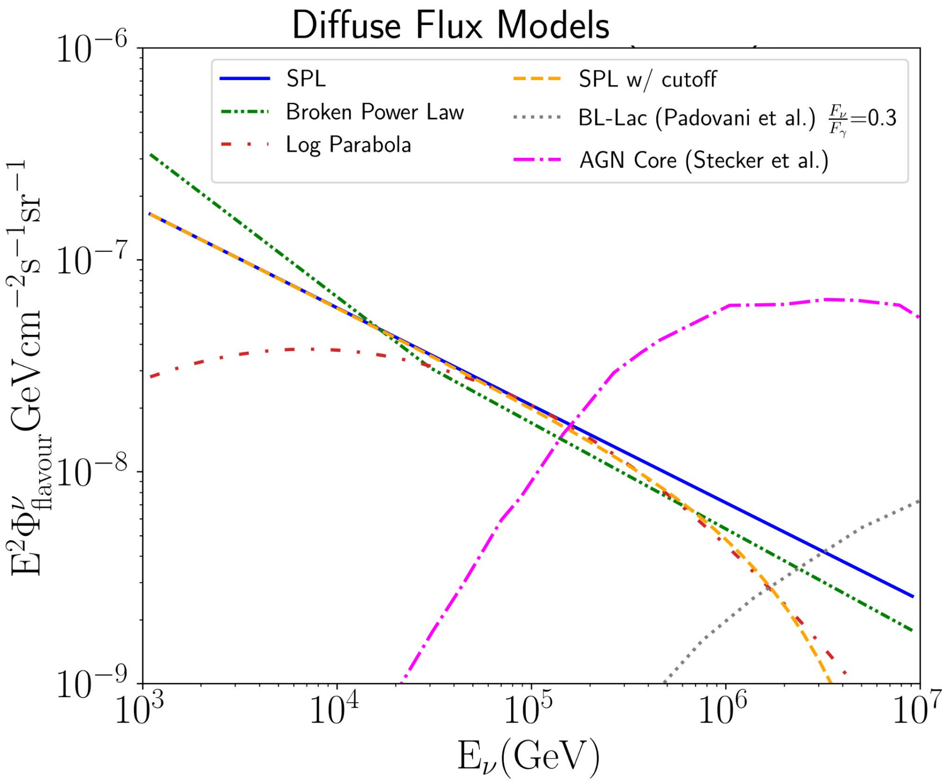
- Log Parabola (LP):

$$\Phi_{\text{astro}}\left(\frac{E_\nu}{100\text{TeV}}\right)^{-\alpha-\beta\log_{10}\left(\frac{E_\nu}{100\text{TeV}}\right)}$$

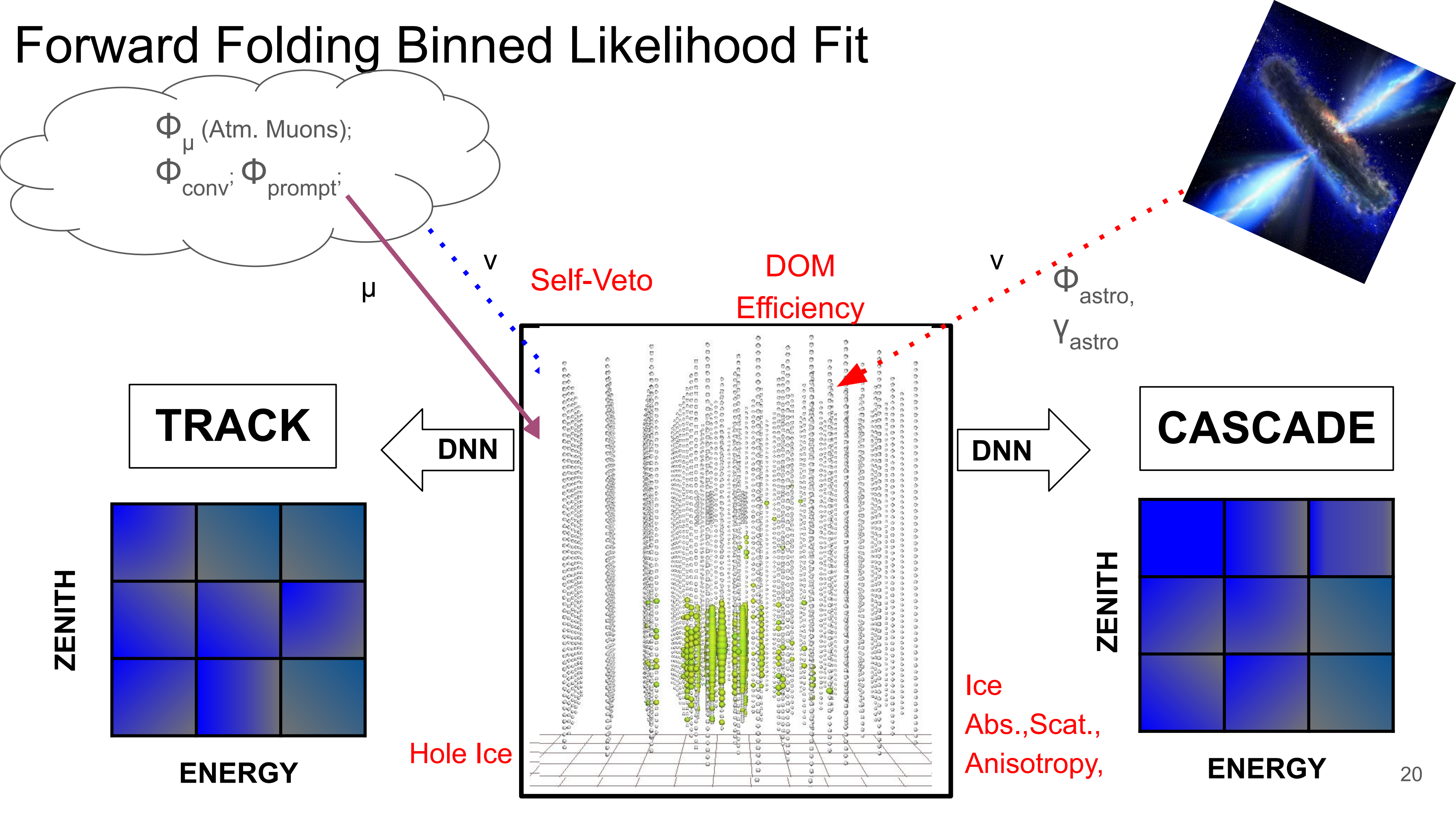
- SPL+Cutoff:

$$\Phi_{\text{astro}}\left(\frac{E_\nu}{100\text{TeV}}\right)^{-\gamma} e^{-\left(\frac{E_\nu}{E_{\text{cutoff}}}\right)}$$

- AGN core neutrino emission (Stecker, F. W. et al, 1991)
- Neutrino emission from BLLac objects (Padovani, P. et al, 2015)

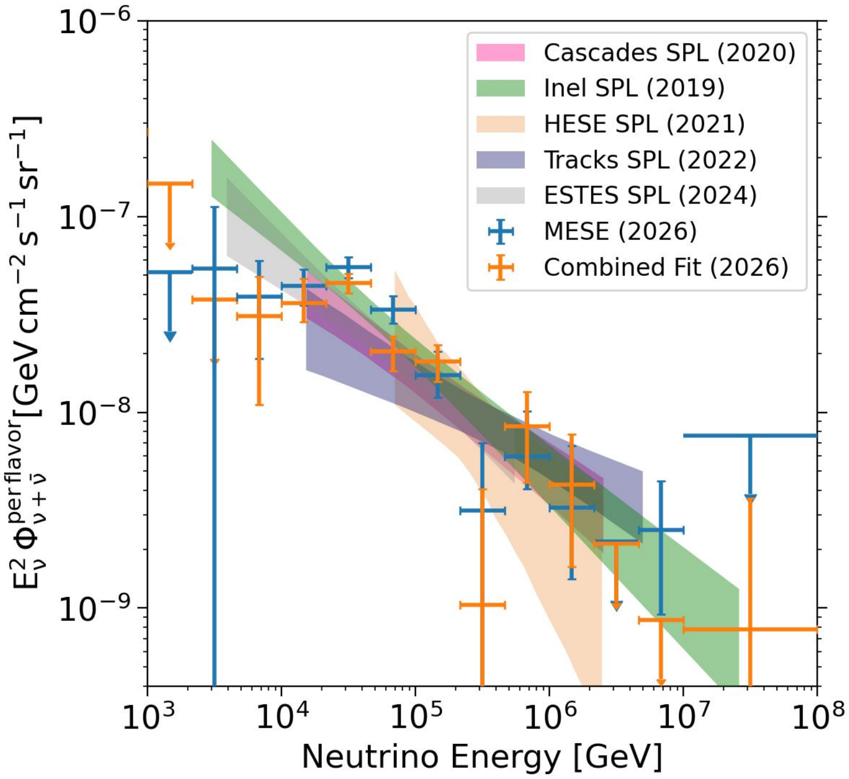


Forward Folding Binned Likelihood Fit



Diffuse Cosmic Neutrino Flux

- In 2026, two analyses were published which observed structure in the Cosmic Neutrino Spectrum
  - MESE: starting events down to 1 TeV
  - Combined Fit: Joint fit of IceCube cascades and upgoing tracks samples
- Reject SPL model for first time by more than 4σ
- Next generation of analyses underway to better resolve features of the neutrino spectrum down to TeV scales

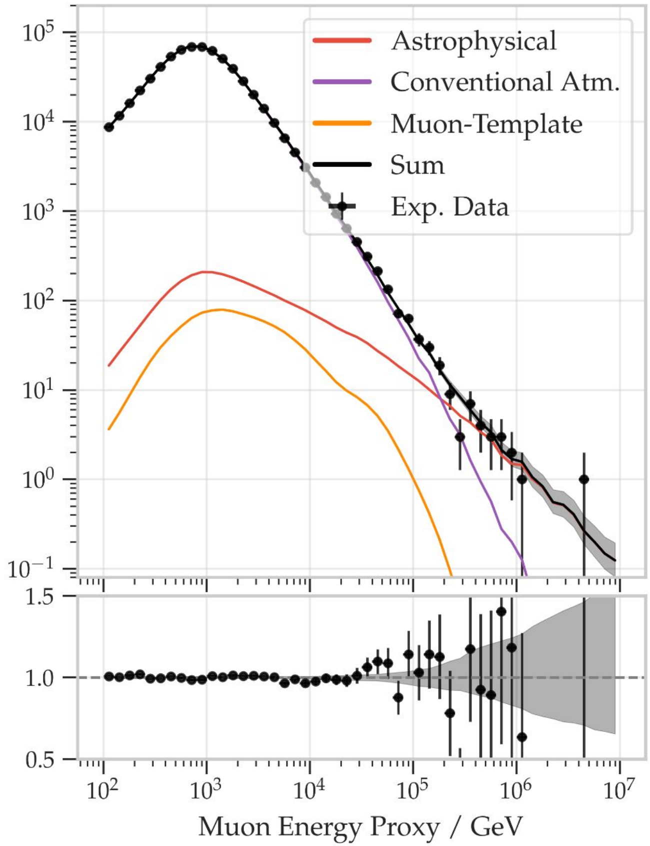


Outline

- Astrophysical Neutrino Detection at Cherenkov Detectors
- Event Classification and Selection Techniques
- Neutrino Spectral Measurements
- **Fitting a neutrino spectrum! (Time to open the Colab Notebooks!)**
- Outlook

Components of the astrophysical fit

- Astrophysical Neutrinos: For the purpose of this talk, this is the signal we are trying to measure.
- Atmospheric Neutrinos: An irreducible background
- Atmospheric Muons: Background which can be discriminated and vetoed, but which we will ignore for this exercise



The energy distribution of upgoing tracks for the period 2010-18
The Astrophysical Journal 928.1 (2022): 50

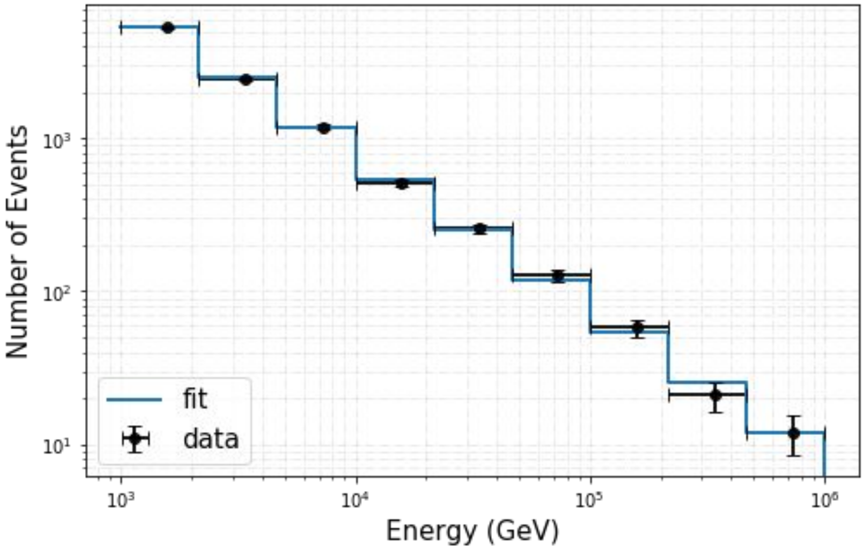
Link to Notebook

- The practical diffuse exercise is hosted here.
- Plan to run scripts on Google Colab, but you can clone the repository and run it in an environment of your choice.

Datasets

- *hese_toy_simulation.json*: A pre-made dataset of simulated neutrino events scaled to $E^{-2.5}$

- *hese_toy_data.json:* The data we will try to fit with our simulated neutrino events



Computing expectation values

- Best way to compute an expectation value is to generate Monte Carlo simulation (MC).
  - Start by generating neutrinos and "inject" them near your detector
- You usually inject neutrinos following chosen energy, zenith, azimuthal distributions, and over a certain area of the detector
- Our MC is very resource-intensive to produce-> need to re-use for different astrophysical models
- Solution: Reweighting

Reweighting MC

- We normalize the simulated events to represent different flux models.
  - We start by computing the gen-weight for each event

$$\Phi_{generated} = N_{generated}[GeV^{-1} * sr^{-1} * cm^{-2}]$$

  - Now we choose a theoretical flux model, either atmospheric or astrophysical. Let us use

$$\Phi_{model} = 2 * 10^{-18} * (\frac{E_\nu}{100TeV})^{-2.6}[s^{-1} * GeV^{-1} * sr^{-1} * cm^{-2}]$$

  - You can now combine the generated flux and theoretical fluxes to compute event rates
  - * Each event will now have it's own expectation rate in units s^{-1}, the weight w

$$w = \frac{\Phi_{model}}{\Phi_{generated}}$$

Binning the data

- We now "bin" the simulation into bins such as energy/zenith/topology, sum all the simulated events within a bin and denote as weight "w_i", how many neutrinos we expect per second in that bin for our event selection.

- For the total number of events, we add up the total weights from each contributing flux- astrophysical signal and atmospheric backgrounds.

- The dataset we have provided is from the HESE analysis, primarily high energy above 60TeV

- We create bins in energy, cosine zenith, and morphology. The morphologies we consider for the toy analysis are cascades and tracks

Binned Likelihood analyses

- We now define a poissonian likelihood, which is a statement about the probability of a given set of model parameters being correct given the data we observe.
- For a given bin, the probability, of detecting **k** counts/s when expecting **λ** counts/s is

$$P(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

- Given that you already see **k** events, we estimate the true expectation to be the value which maximizes the probability of detecting that many events.

$$\mathcal{L}(\lambda|k) = P(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

- Across multiple bins “ *i* ”, we have

$$\mathcal{L}(\overrightarrow{\theta|d}) = \prod_i \frac{(\lambda_i)^{k_i} e^{-\lambda_i}}{k_i!}$$

Performing the fit

- As most optimizers are written as minimizers, we define the log-likelihood which turns the product across bins into a sum

$$\log \mathcal{L}(\theta \overrightarrow{\,\Big|\,} d) = \sum_i k_i \cdot \log\left(\lambda_i\right) - \lambda_i - \log \Gamma\left[k_i + 1\right]$$

- We then take the negative of the log likelihood, turning the maximum into a minimum, to take advantage of computational minimizers

Hypothesis Testing: Likelihood ratio test

- To test whether a given model fits our data better or worse than another, we run a series of pseudo-experiments with Poisson sampled histograms, and try to fit the test histogram with both models.

- We define our null hypothesis H_0 and our test hypothesis H_1. Then our Test Statistic is given by

$$TS = 2(\ln \mathcal{L}_{H1} - \ln \mathcal{L}_{H0})$$

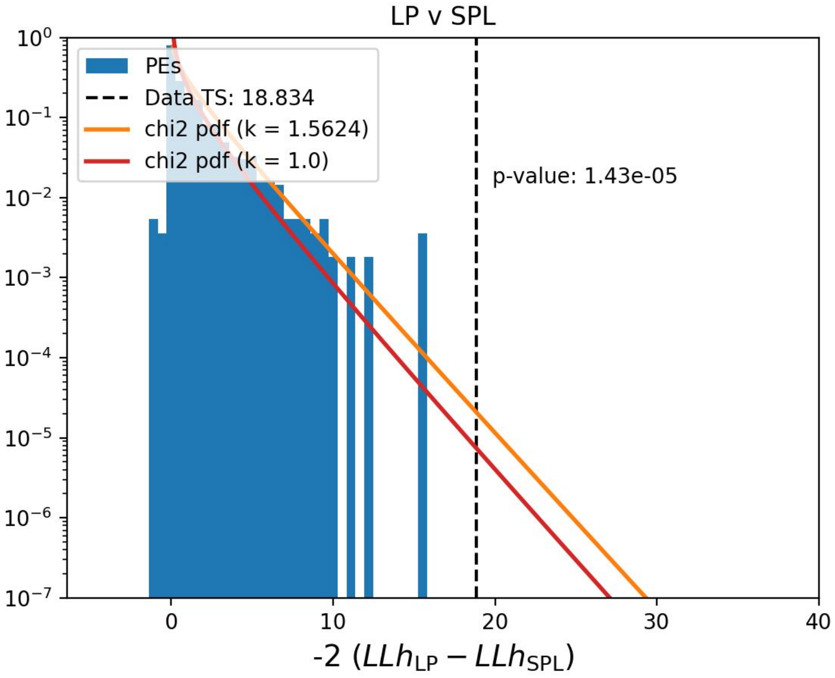
- We obtain TS_{data} by fitting our hypotheses to the data histogram, but to determine the statistical significance of our result, we must also obtain a background distribution obtained from fitting a histogram obtained from the null hypothesis

Hypothesis Testing: Background distribution

- We can then define a p-value for our TS_{data} based on the background distribution.

- It becomes computationally very expensive to run pseudo-experiments to map out the tail of the distribution

- We therefore appeal to **Wilks' Theorem**, which provides a model of the underlying background distribution, which we can use for the asymptotes.

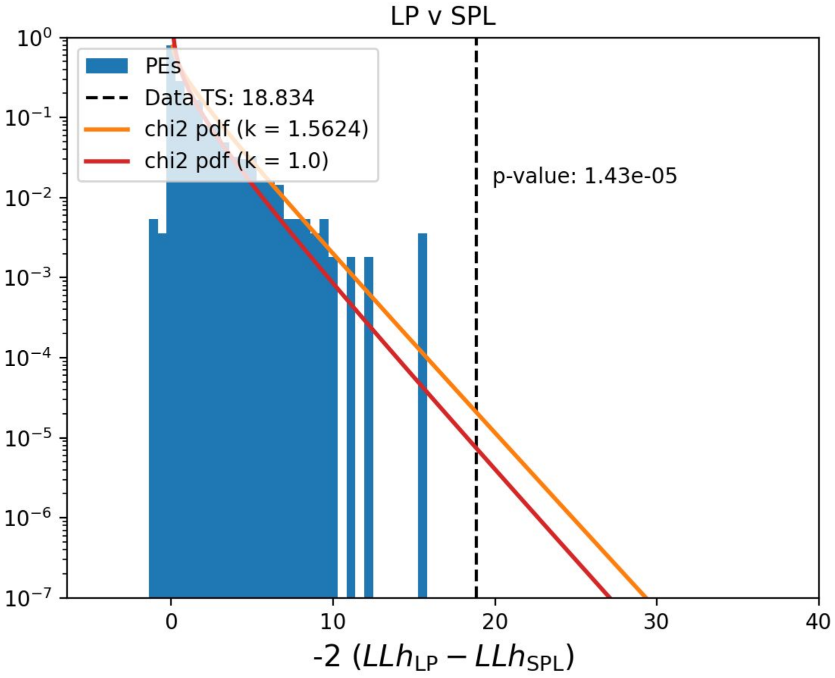


Hypothesis Testing: Wilks' Theorem

- The theorem states that, if H_0 is true, then as the sample size grows, the distribution of the TS converges to a chi-square (χ^2) distribution, with degrees-of-freedom equal to the number of extra parameters in H_1 vs H_0, regardless of the specific details of the tested models.

- Key underlying assumptions:
  - The two models are nested: H_0 must be a special case of the H_1 parameters
  - The true parameter must lie in the interior of the parameter space: Not strictly true for LP vs SPL!

- We use a modification due to **Chernoff:** The distribution is a delta function at zero plus a chi-square.



Outline

- Astrophysical Neutrino Detection at Cherenkov Detectors
- Event Classification and Selection Techniques
- Neutrino Spectral Measurements
- Fitting a neutrino spectrum! (Time to open the Colab Notebooks!)
- **Outlook**

Outlook

- We have covered the mechanics of performing a simple diffuse fit, along with all the systematic uncertainties that must be accounted for

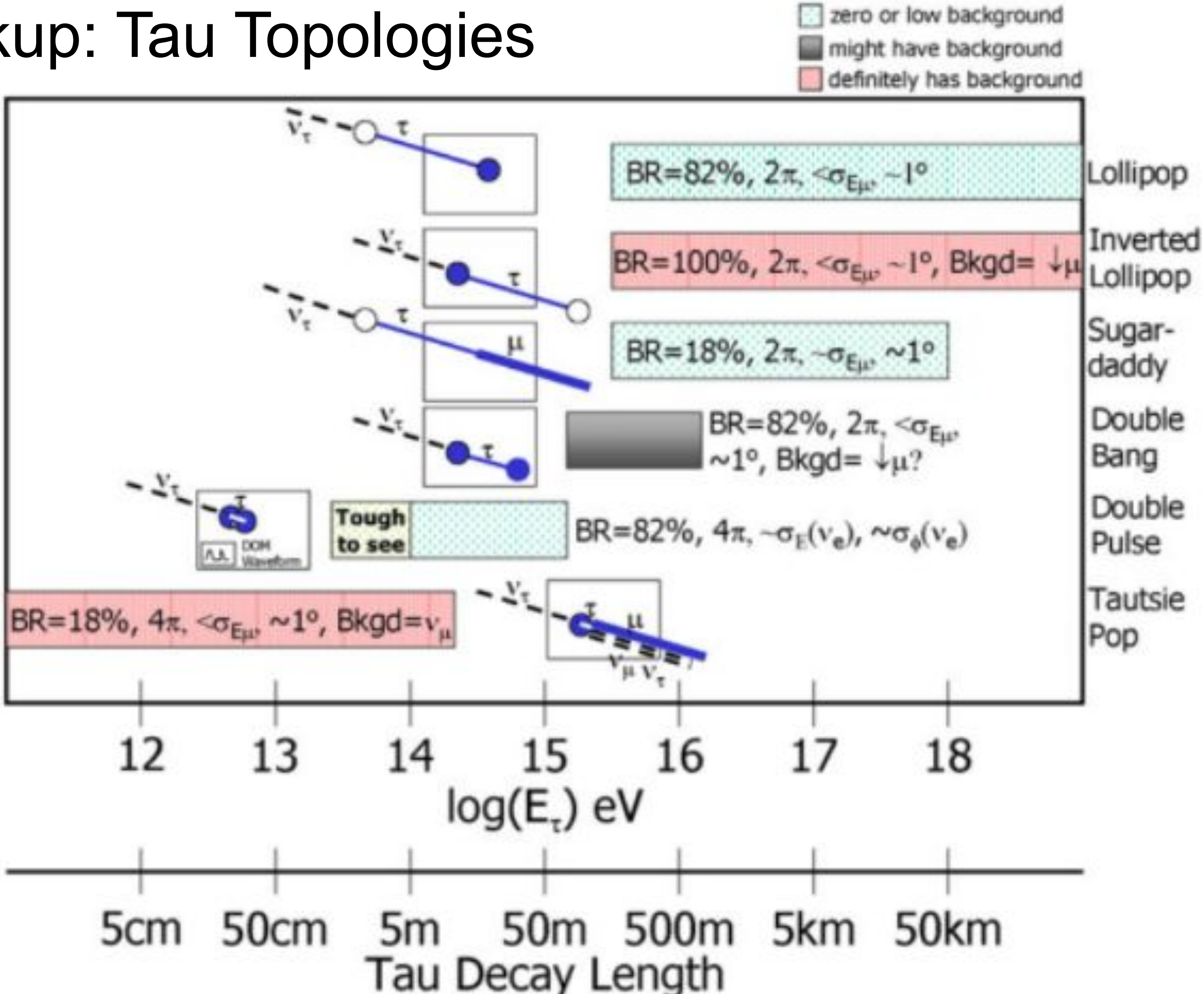
- New results from IceCube indicate the presence of structure in the diffuse astrophysical neutrino spectrum, with a steep spectrum above ~30 TeV and a harder spectrum below that threshold

- New analyses underway to better resolve the low energy spectrum, using machine learning techniques which provide much higher purity neutrino samples!

- Questions?

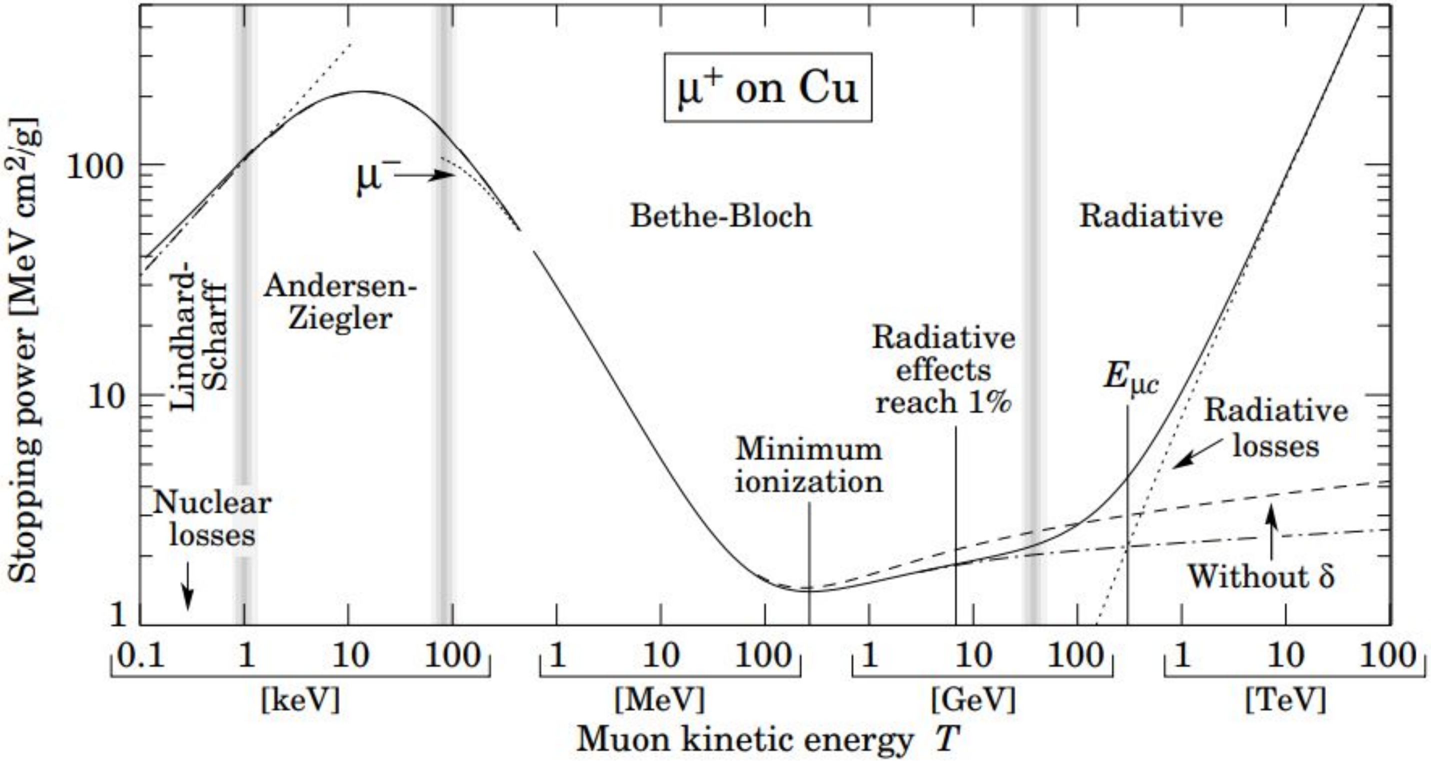
Backup

Backup: Tau Topologies



Cowen, D. F., and IceCube Collaboration. "Tau neutrinos in IceCube." *Journal of Physics: Conference Series*. Vol. 60. No. 1. IOP Publishing, 2007.

Backup: Muon Energy Losses



Groom, Donald E., Nikolai V. Mokhov, and Sergei I. Striganov. "Muon stopping power and range tables 10 MeV–100 TeV." Atomic Data and Nuclear Data Tables 78.2 (2001): 183-356.

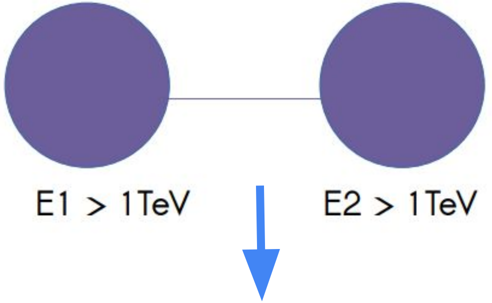
Systematics: Atmospheric Neutrino Self-Veto

We use the P_light as an input to analytically compute the probability of a neutrino with energy E, and zenith angle θ not getting vetoed due to an accompanying sibling muon, this is the passing fraction

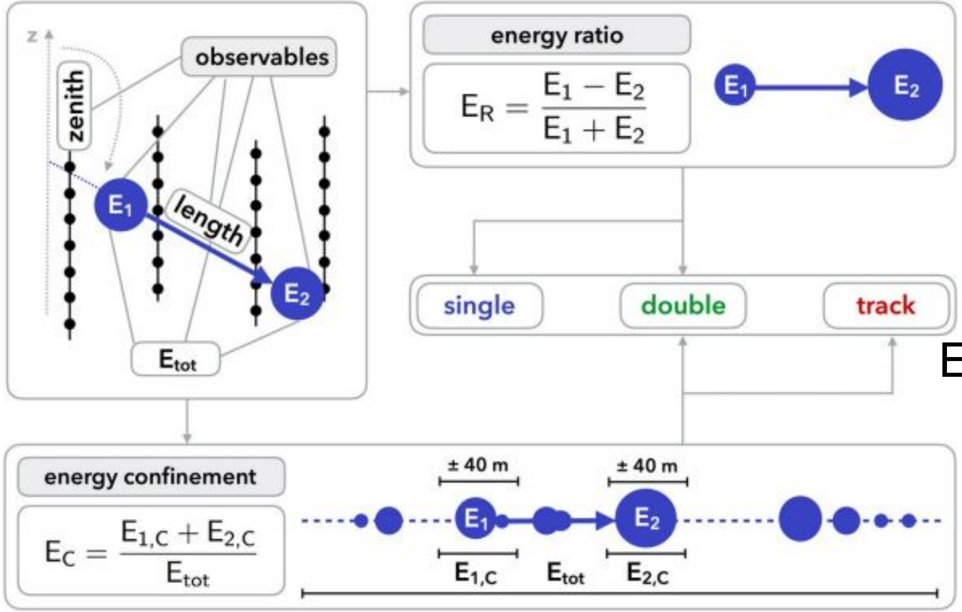
$$\mathcal{P}_{\rm pass}(E_\nu,\theta_z) = \frac{1}{\phi_\nu(E_\nu,\theta_z)} \sum_A \sum_p \int dE_p \int \frac{dX}{\lambda_p(E_p,X)} \int dE_{\rm CR} \frac{dN_{p,\nu}}{dE_\nu}(E_p,E_\nu) \frac{dN_{A,p}}{dE_p}(E_{\rm CR},E_p,X)\, \phi_A(E_{\rm CR})$$

$$\times \left[1 - \int dE_\mu^{\rm i} \frac{dN_{p,\mu}}{dE_\mu^{\rm i}}(E_p,E_\nu,E_\mu^{\rm i}) \underbrace{\int dE_\mu^{\rm f}\, \mathcal{P}_{\rm light}(E_\mu^{\rm f})\, \mathcal{P}_{\rm reach}\left(E_\mu^{\rm f}|E_\mu^{\rm i},\theta_z\right)}_{\mathcal{P}_{\rm det}\left(E_\mu^{\rm i},\, X_\mu(\theta_z)\right)}\right] e^{-N_{A,\mu}(E_{\rm CR}-E_p,\theta_z)}$$

Tau neutrino tagging



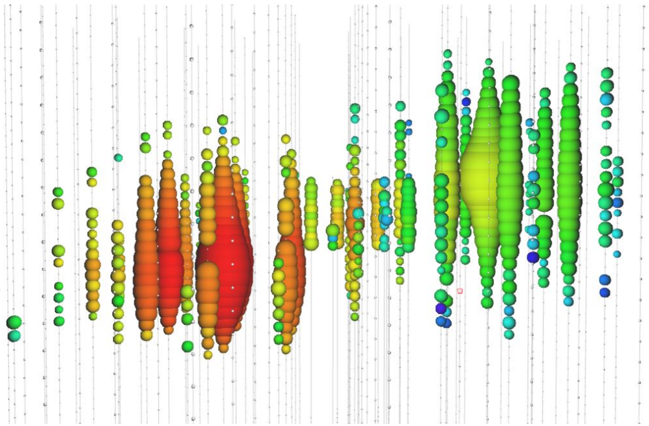
- Using Taupede: likelihood-based fitter for double cascades (DC), previously used in HESE
  (M. Usner DOI:10.18452/19458, J. Stachurska DOI:10.18452/21611)
- As MESE has a DNN-based cascade/track classifier, we use only the DC selection with Taupede



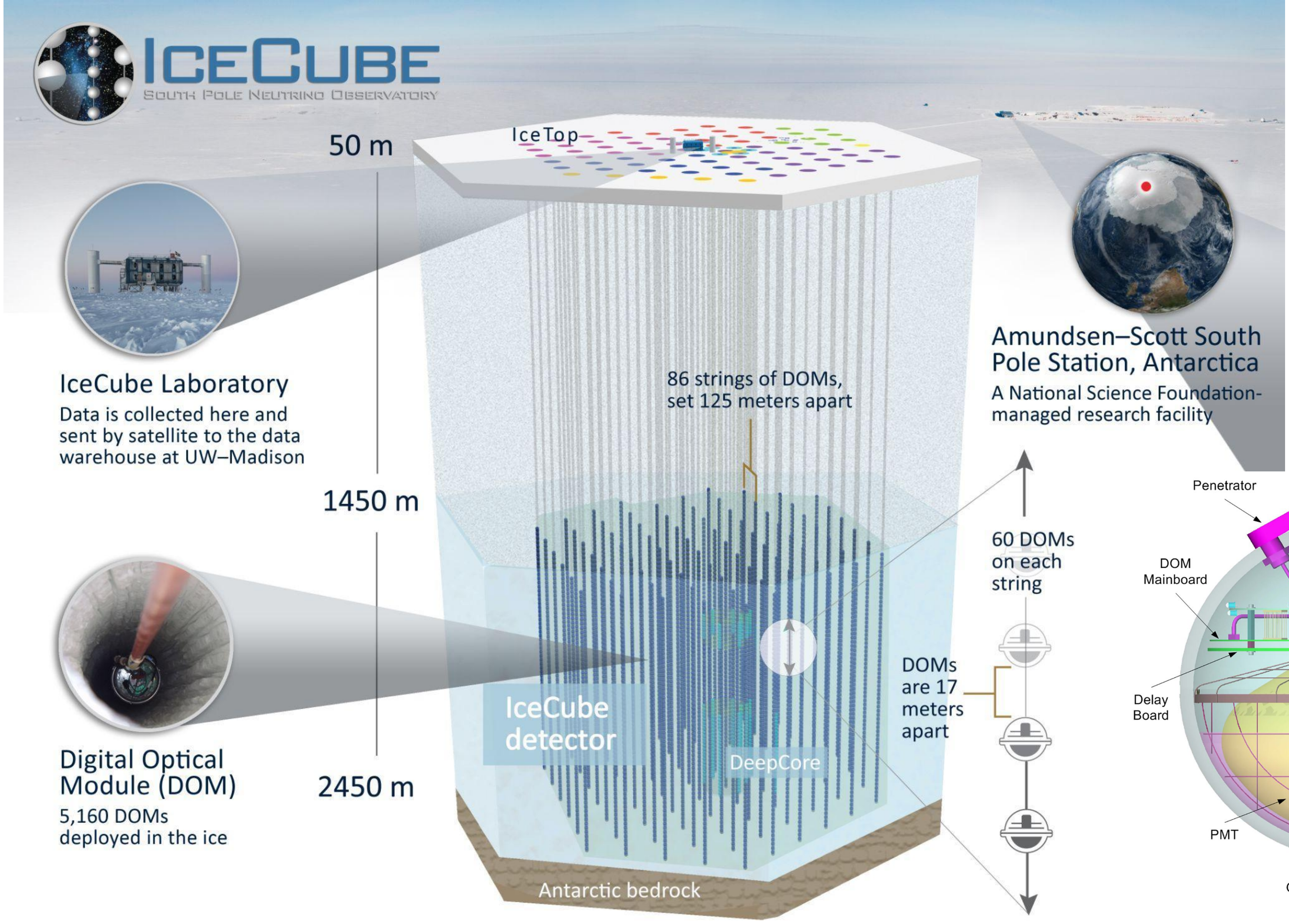
(EC≥0.99)

ER ϵ (0.98,0.3)

(E_{tot}≥ 30TeV)



Simulated Double Cascade Event



Digital **O**ptical **M**odule