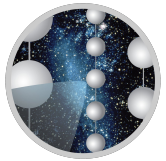



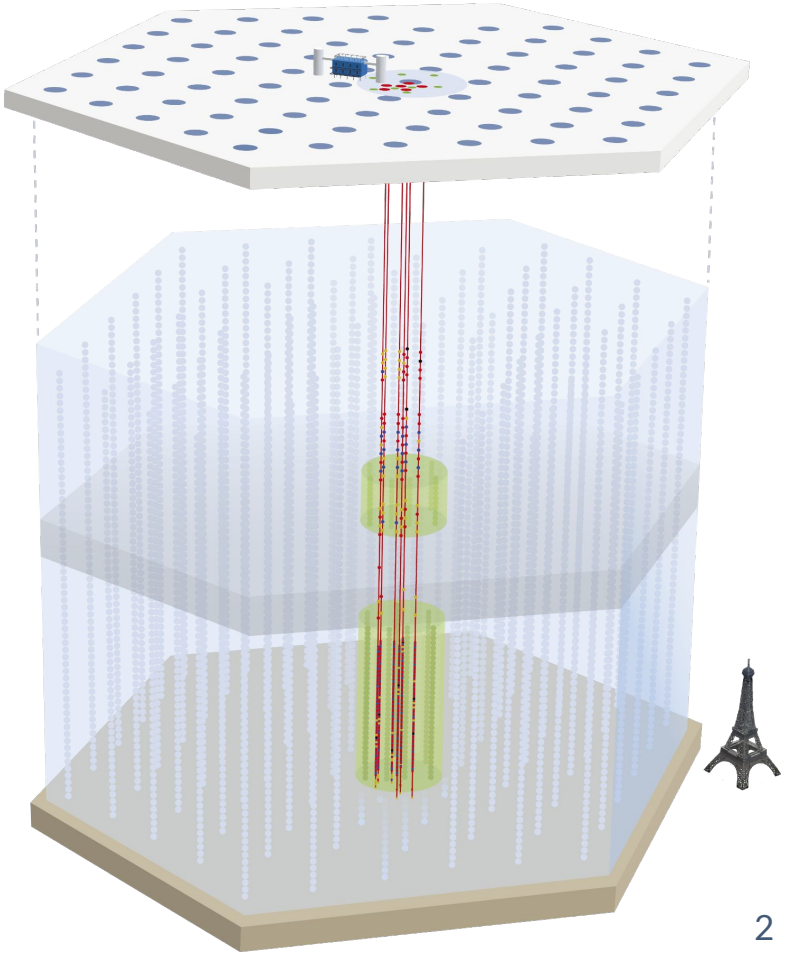
Ongoing Work on ν_τ Classification with IceCube at the TeV scale

Finn Mayhew, on behalf of the IceCube Collaboration

NPML, June 15, 2026

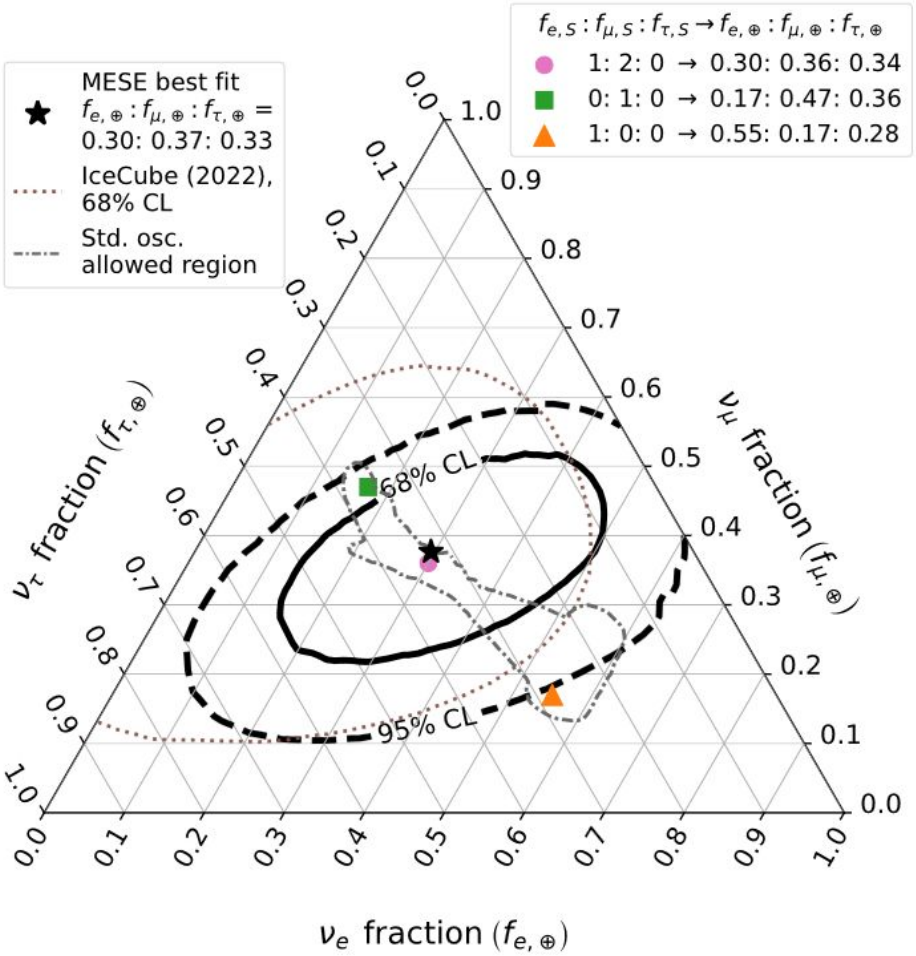
IceCube

- Cubic-kilometer ice Cherenkov detector at the south pole
- Over 5k PMTs housed in digital optical modules (DOMs)
- Taking data in full configuration since 2011
- In 2025-26: new strings installed for the IceCube Upgrade



Astrophysical flavor ratio

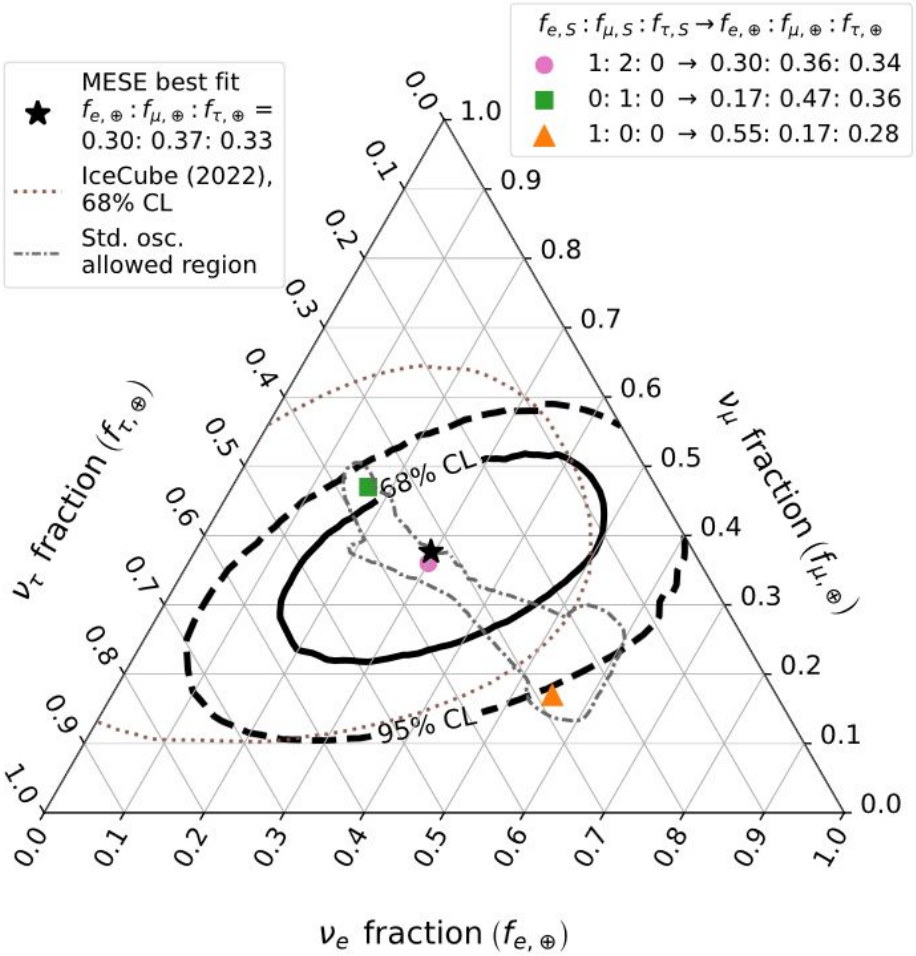
- Measuring astrophysical ν flux constrains ν oscillation and astrophysical production scenarios
  - 1:2:0 = $\pi^{\pm}$ decay
  - 0:1:0 = rapid μ energy loss
  - 1:0:0 = *n* decay
- Big deal if exclude any production scenario!



DOI: 10.48550/arXiv.2510.24957

Astrophysical flavor ratio

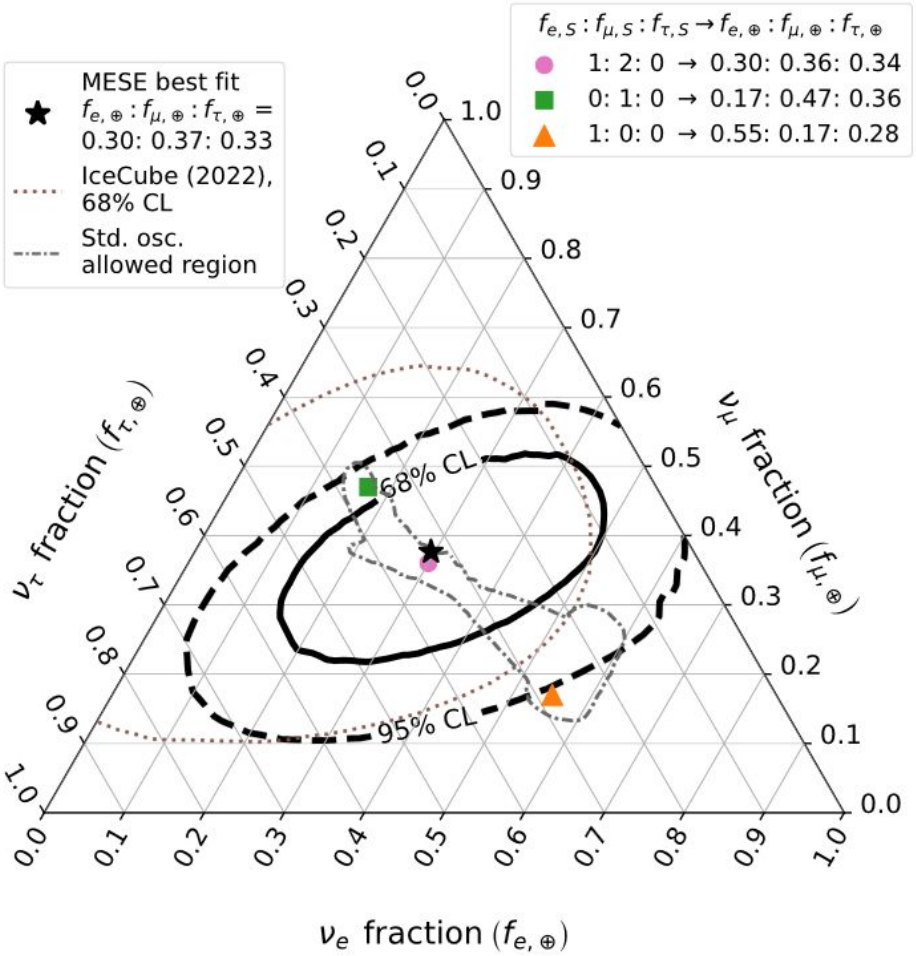
- Standard oscillations imply a narrow region of ratios at Earth
- If measure outside of standard oscillations allowed region, evidence for BSM!



DOI: 10.48550/arXiv.2510.24957

Astrophysical flavor ratio

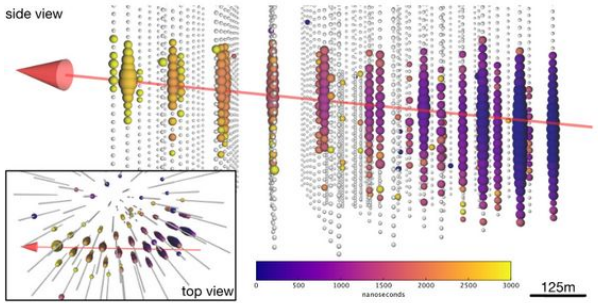
- Room for improvement in ν_τ: difficult to distinguish from ν_e's, especially below ~100 TeV
- **Goal: refine ν_τ flux measurement at low energies, shrink flavor triangle contour**



DOI: 10.48550/arXiv.2510.24957

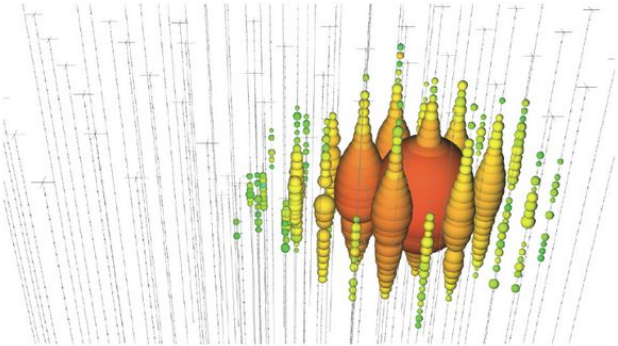
ν signatures in IceCube

Track



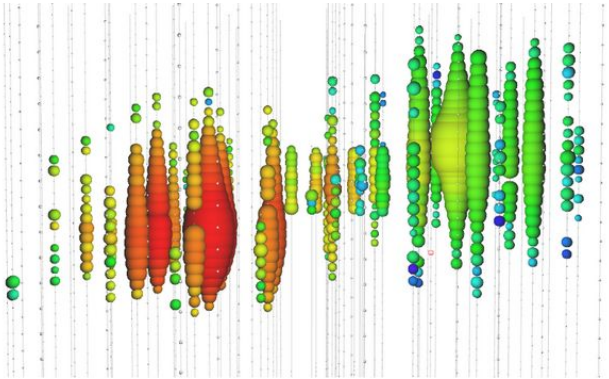
ν_μ CC

Cascade



All ν NC
ν_e CC
ν_τ CC $\lesssim$ several PeV

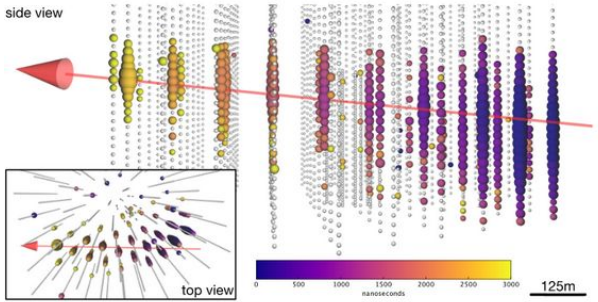
Double Cascade



ν_τ CC $\gtrsim$ several PeV

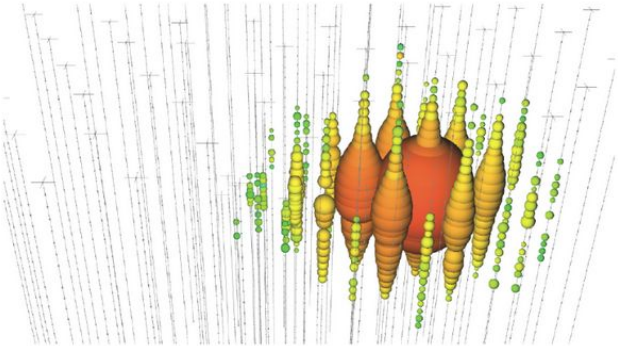
ν signatures in IceCube

Track



ν_μ CC

Cascade

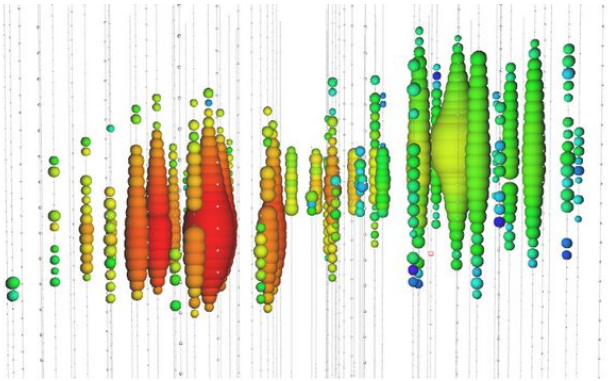


All ν NC

ν_e CC

ν_τ CC $\lesssim$ several PeV

Double Cascade

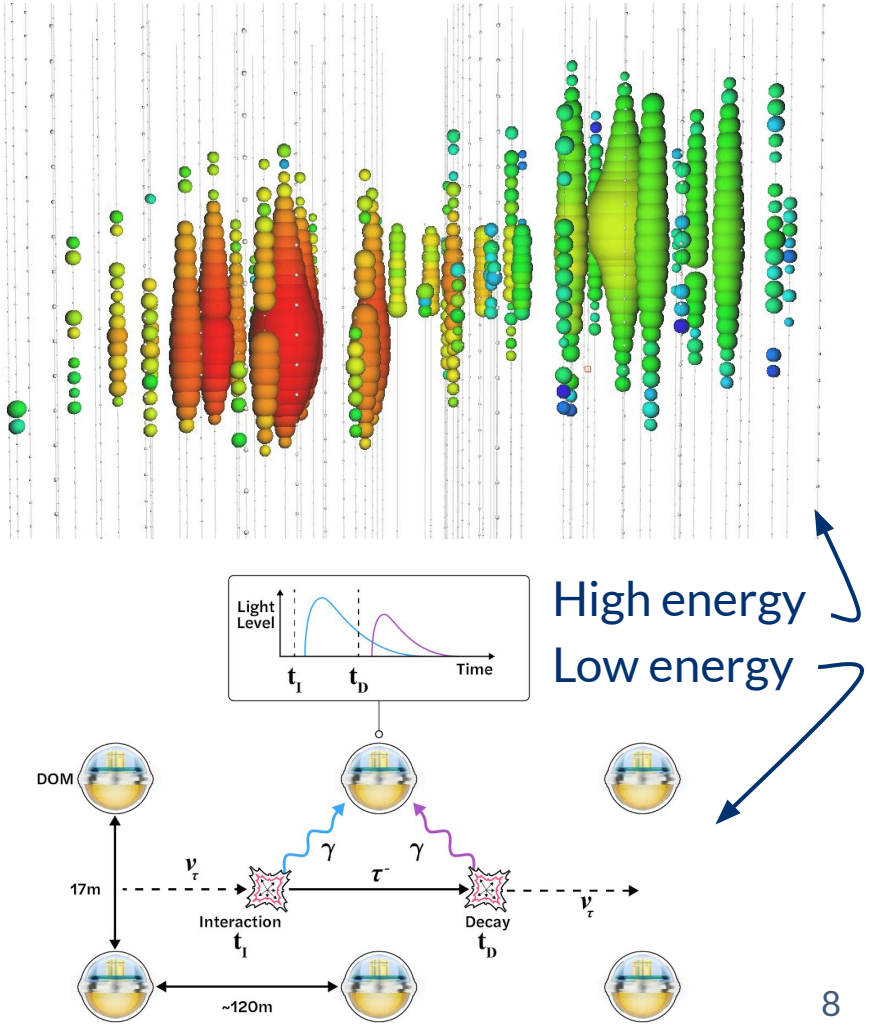


ν_τ CC $\gtrsim$ several PeV

Looking to distinguish these

ν_τ signature

- ν_τ's deposit light as **two showers**:
  1. τ lepton production
  2. τ lepton decay
- τ decay length $\propto \nu_\tau$ energy
- $\gtrsim$ several PeV, showers are visible as two cascades in the detector
- $\lesssim$ several PeV, showers overlap, harder to distinguish from ν_e single showers
- **=> Look for double pulse pattern in individual DOMs**

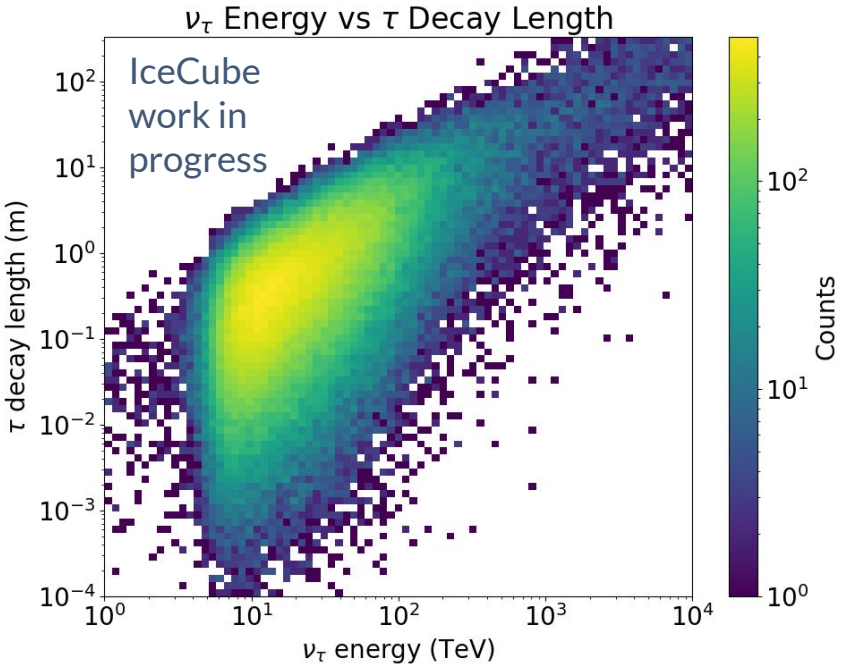


Training sample

- Energy range: 100 GeV to 100 PeV
- ν_e and ν_τ CC only
  - First, distinguish a single background
  - May also be able to distinguish NC (extremely similar signature to ν_e CC)
  - Another network is being trained that we hope will distinguish other background
- Cut ν_τ -> τ -> μ events, which have a different signature
- Pass through pre-existing event selection (through Section II.B.1 here)
  - Will reduce background in real data
  - Included here to better align training sample with real data

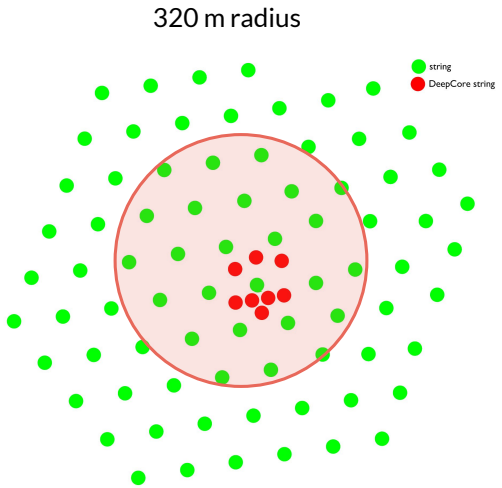
Networks specialized in energy regimes

- Two regimes in reco energy:
  - Low (5-65 TeV): more difficult to distinguish from ν_e
  - High (65+ TeV): less difficult
  - Note: plot shows true quantities
- Shared architecture
- Separate training and potentially different inputs and hyperparameters



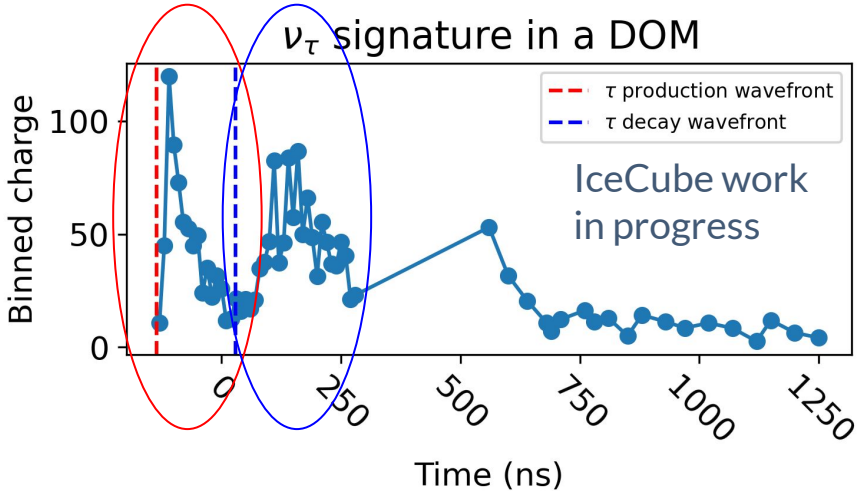
Input

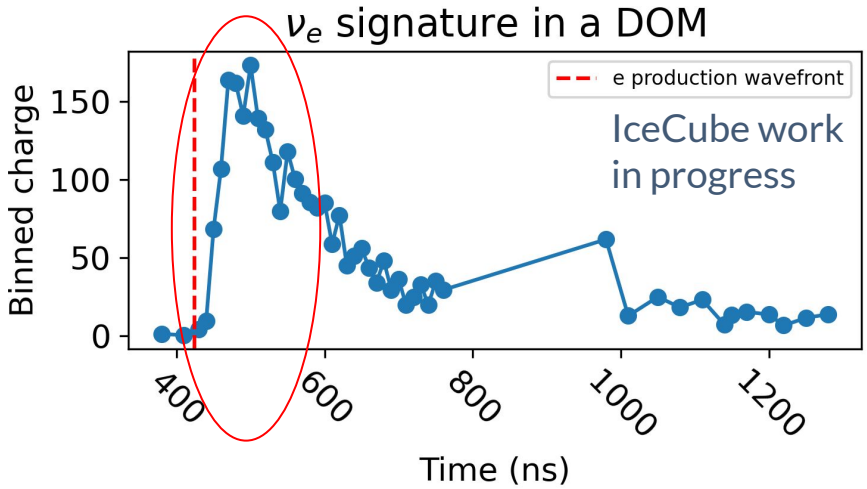
- Memory usage ∝ number of tokens squared
  => must **keep only relevant info**
- These settings may later be optimized and may vary between the two networks



Step	Reason	Current value
Cut info far from τ	Photon scattering + absorption blurs signature	Cut outside of 320 m radius cylinder + after 1.5 μs
Include only a set number of tokens, prioritize by closeness to reco vertex	Directly limit memory usage	1000 tokens
Reduce time resolution	Allow network to see farther with same number of tokens, helpful for high-energy events	Bin charges every 10 ns

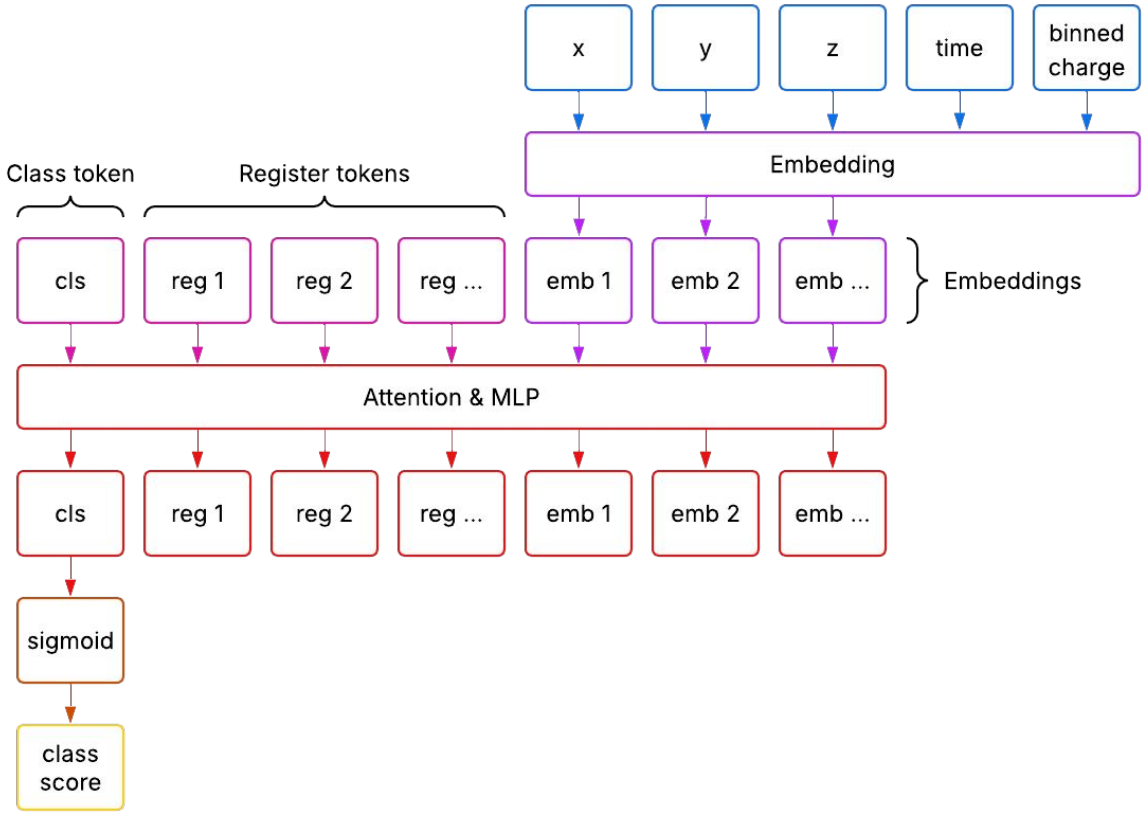
Example signatures in binned charge





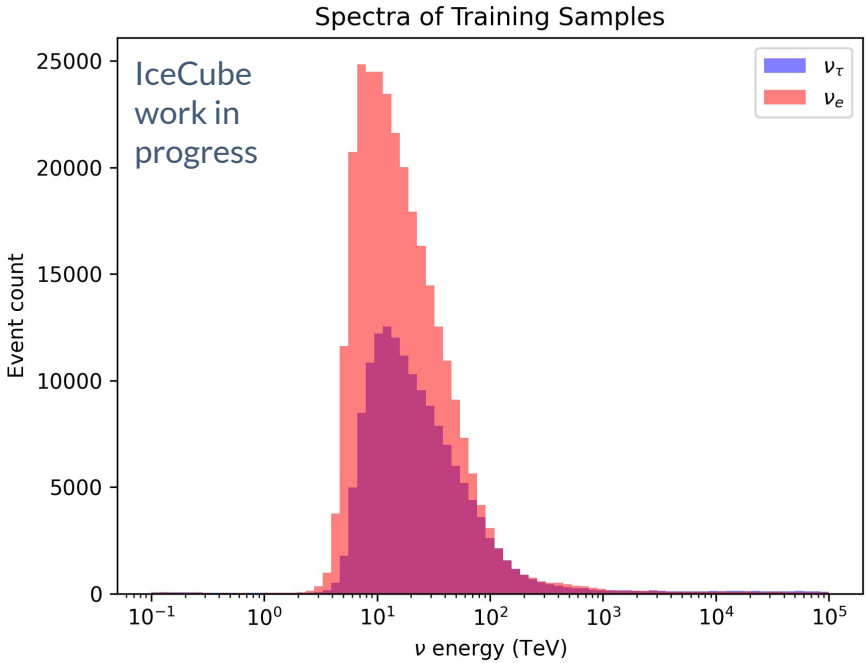
Architecture

- Features:
  - DOM position
  - Binned charge time
  - Binned charge value
- Fourier encode for normalization
- Use register tokens to smooth attention maps



Current status

- Training is ongoing
- Upcoming:
  - Balancing classes
  - Generating more simulation
  - Downsampling to a realistic spectrum
  - Looking into additional features + PCA
  - Hyperparameter optimization



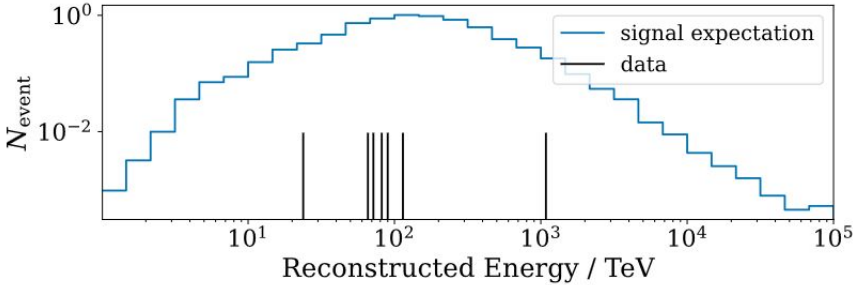
Backup

Current architecture hyperparameter values

Parameter	Value
Embedding space dimension	384
Number of (attention + MLP) layer pairs	16
Number of attention heads	12
MLP layer dimension	1536
Number of relative position buckets	32
Maximum relative position	256
Number of register tokens	3

Most recently published IceCube ν_τ analysis (2024)

- Observed 7 candidate ν_τ events
- 9.7 years of data
- CNN-based approach: 2D images of DOM number (~depth) vs time



DOI: 10.48550/arXiv.2403.02516

Attention in our context

- Attention process allows network to learn relationships between tokens
- Our relationship is the double pulse pattern (our signature)
- Input is general enough that unforeseen relationships can be learned
    - Including patterns of info from multiple DOMs