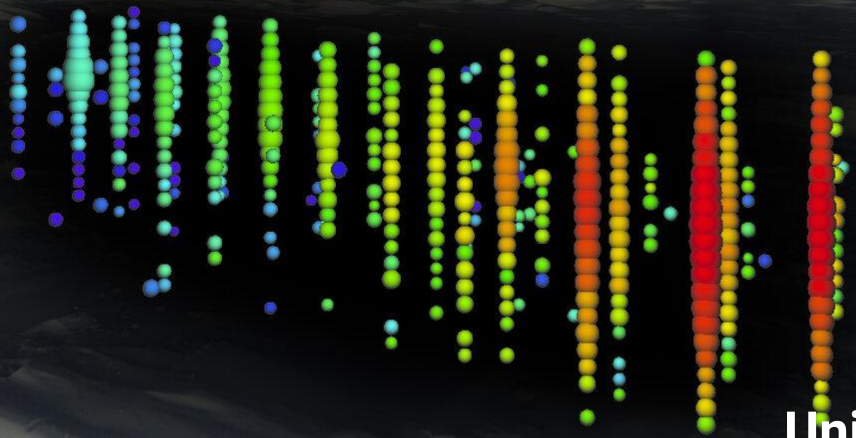
Tiny Particles in The Giant Detector

Reconstruct 10-GeV Scale Neutrinos Using Convolutional Neural Networks in IceCube



April 25th, 2024

Seminar @ iHEP, Beijing

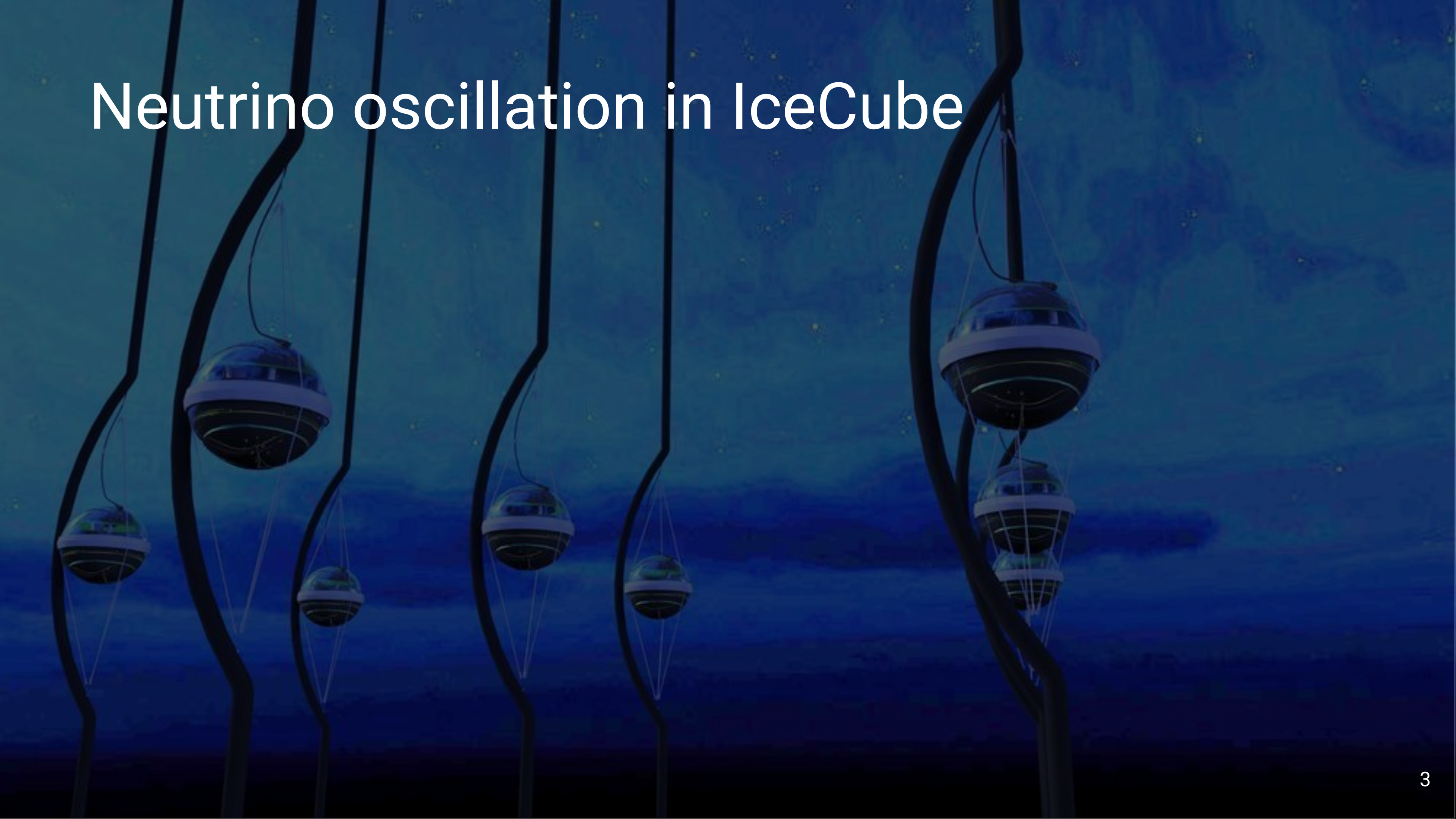
Shiqi Yu

University of Utah

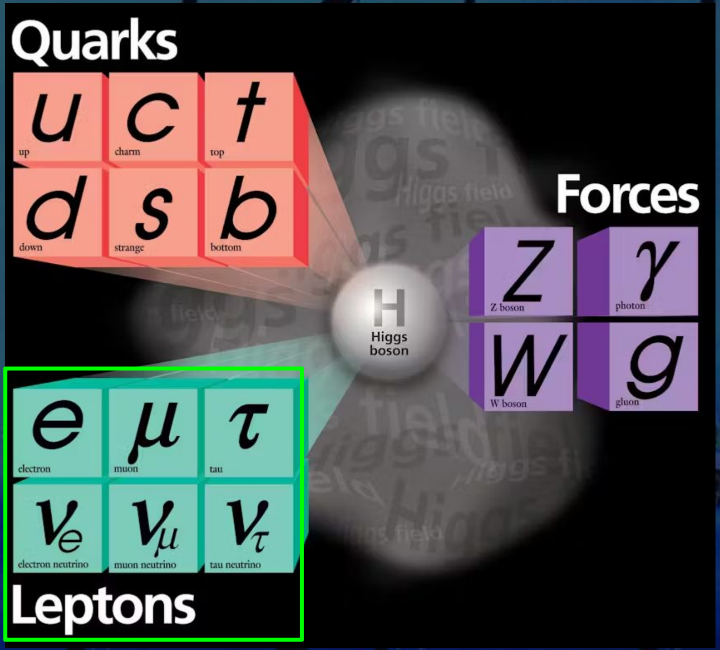
Outlines

- Neutrino oscillation in IceCube
- 10-GeV scale neutrino reconstruction in IceCube
- Performance of CNNs on low-energy events
- Oscillation result and future outlook

Neutrino oscillation in IceCube



Neutrinos



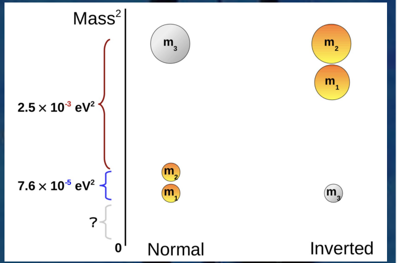
- Second most abundant particles in the Universe
- Have 3 flavors: electron, muon, and tau
- Interact via weak interactions

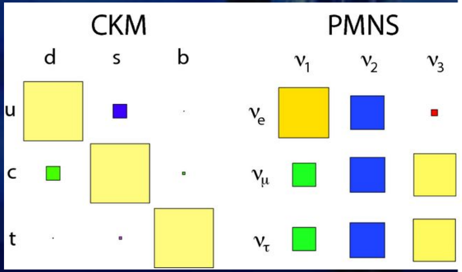
Many interesting questions……

Why study neutrino physics?

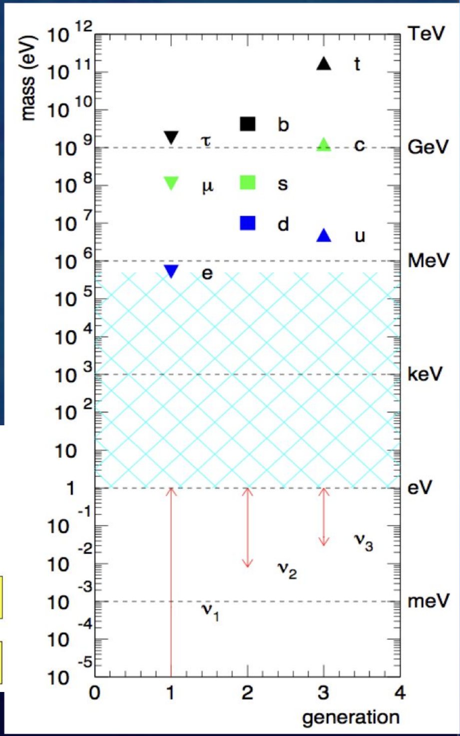
Neutrinos raise many interesting questions:

- Why are the neutrino masses so small?
  - Ordering of neutrino masses?
- Why does neutrino mixing look very different from CKM?
- Is there CP violation in lepton sector?
  - Why is the Universe dominated by matter?

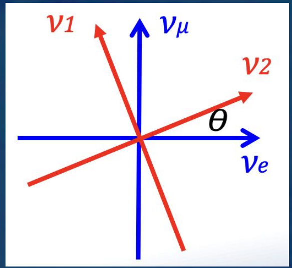




Area of squares represents square of matrix elements.

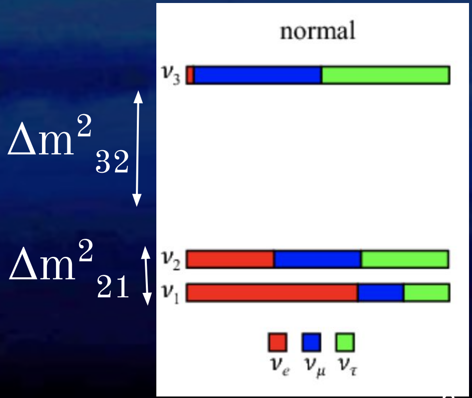


Neutrino Oscillations



- Created in one flavor (ν_μ), but detected in another (ν_e)
- Each flavor (e, μ, τ) is a superposition of masses (1, 2, 3)

- Oscillations among the three neutrino flavors described by:
  - The mixing angles ($\boldsymbol{\theta_{23}}$, θ_{13}, θ_{12}), δ_{CP}
  - The squared mass differences: $\mathbf{\Delta m^2_{32}}$, Δm^2_{21}



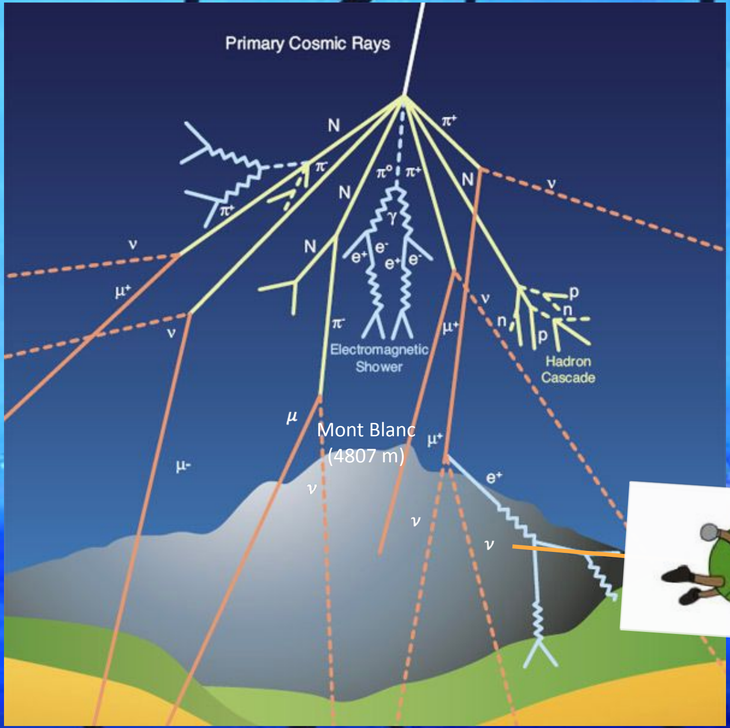
ν_μ disappearance

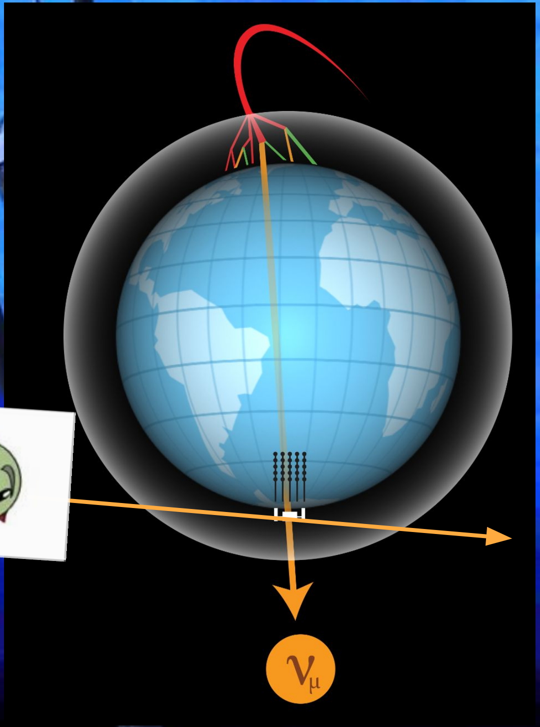
$$P_{\nu_\mu \to \nu_\mu}(L) \approx 1 - \sin^2 2\theta_{23} \cdot \sin^2\left(\frac{1}{4} \cdot \Delta m^2_{32} \cdot \frac{L}{E}\right)$$

* two flavor approximation

- What we need to reconstruct:
  - Neutrino energy (E)
  - Distance of travel (L)
  - Variables to select signal (ν_μ) dominating dataset with reliable reconstruction performance
- What we want to measure:
  - Mixing angle (θ_{23})
  - Mass squared different Δm^2_{32}

IceCube studies atmospheric neutrinos



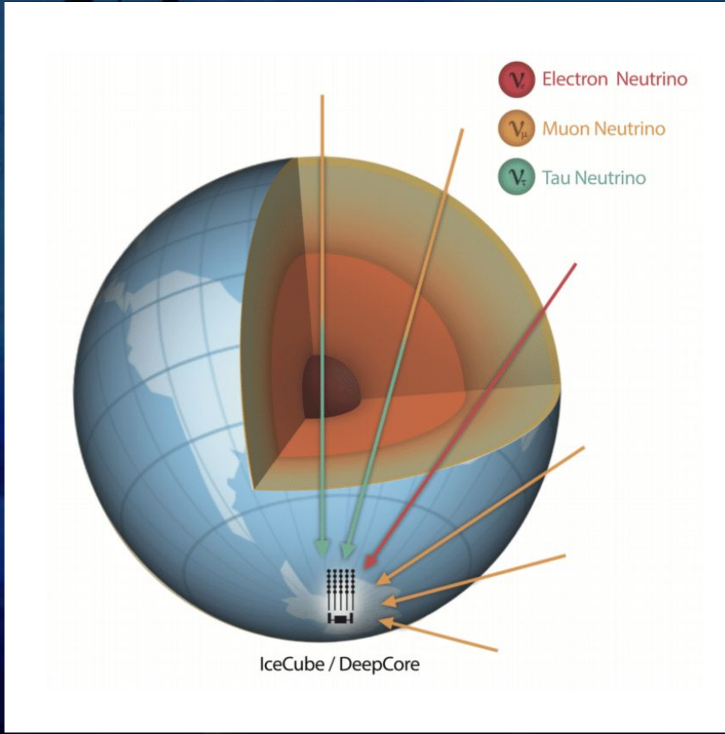


Atmospheric muon neutrinos
from cosmic ray interactions

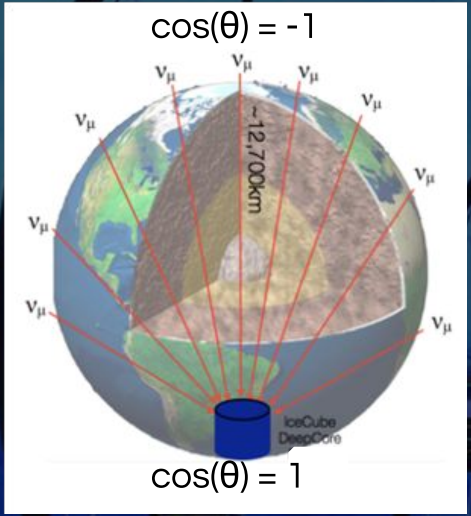
Atmospheric Neutrino Oscillations

- Atmospheric neutrino flux is $\boldsymbol{\nu_\mu}$ dominated good for studying $\boldsymbol{\nu_\mu}$ disappearance
- Wide range of baseline
  - Use zenith angle in detector to determine L

→ 2D measurement in L vs. E

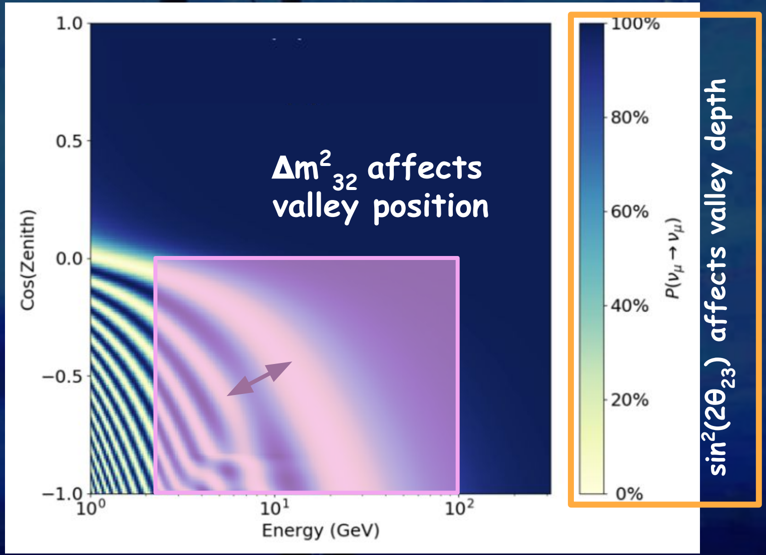


Oscillation Measurement: Baselines vs. Energies



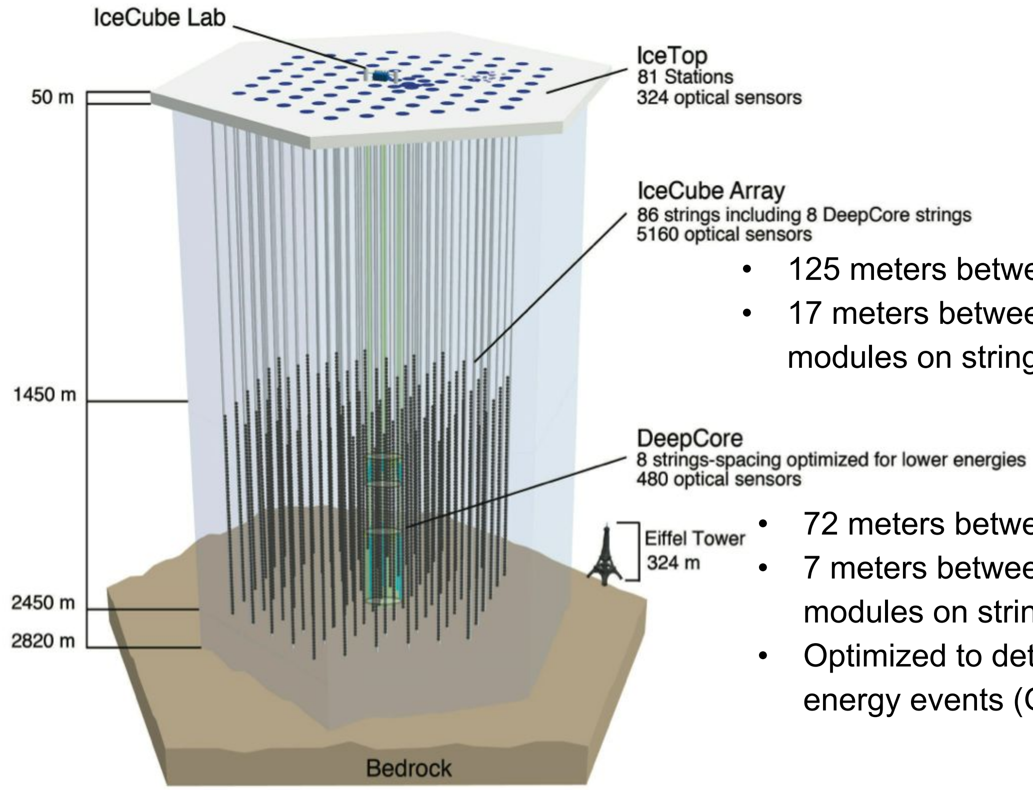
$$P(\nu_\mu \rightarrow \nu_\mu) \approx 1 - \sin^2(2\theta_{23}) \sin^2\left(\frac{1.27\Delta m^2_{32} L}{E}\right)$$

* two flavor approximation



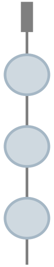
Low-energy (< 100 GeV) reconstruction is critical to oscillation analysis

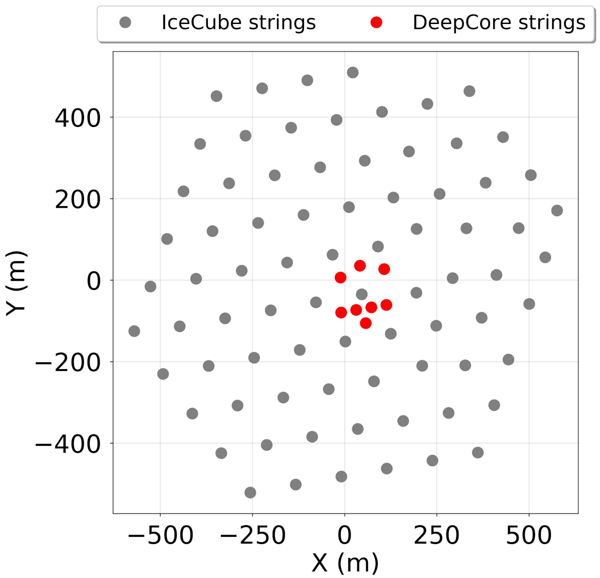
IceCube Detector



- 125 meters between strings
- 17 meters between optical modules on string

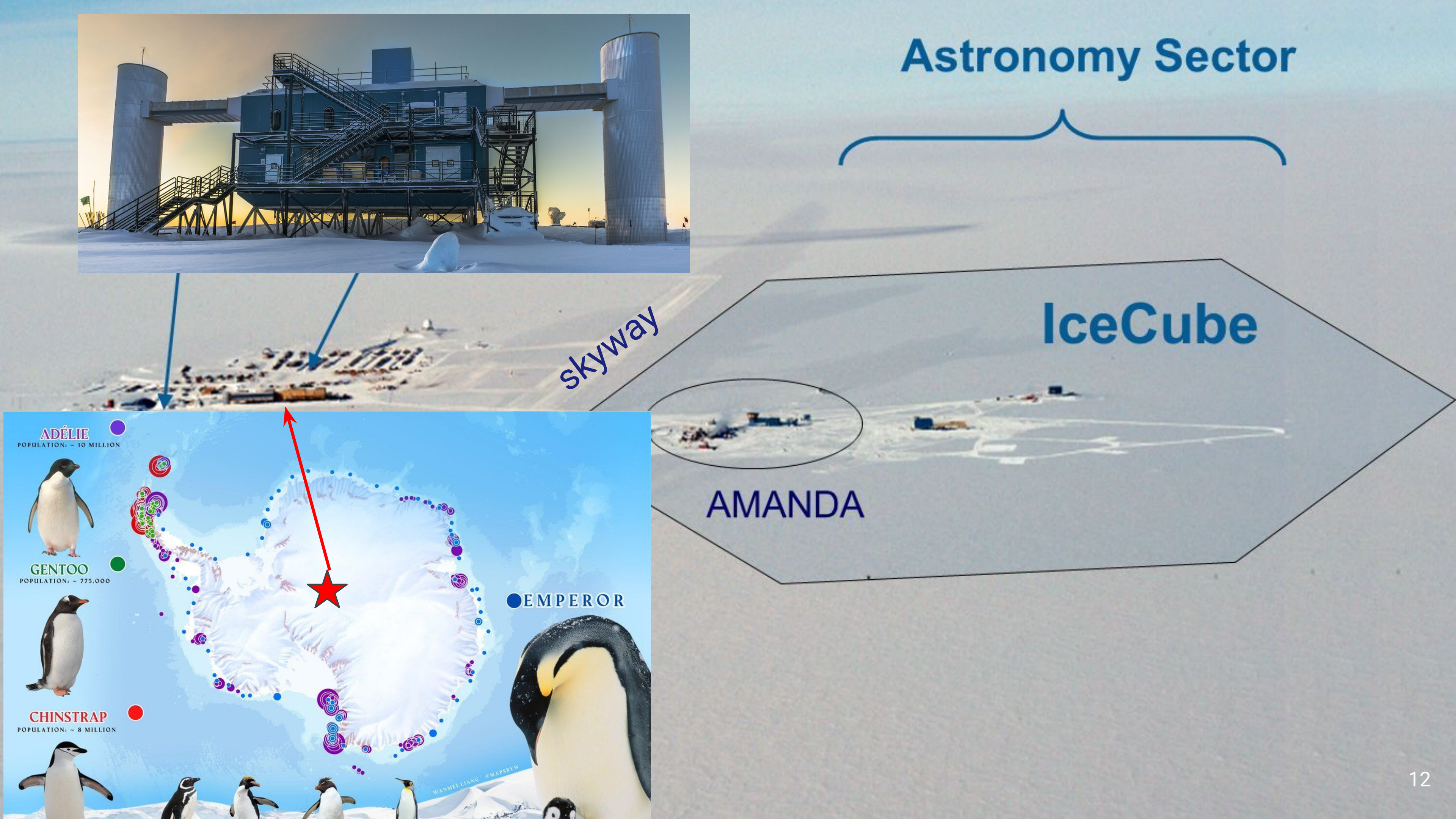
- 72 meters between strings
- 7 meters between optical modules on string
- Optimized to detect lower energy events (GeV-scale)



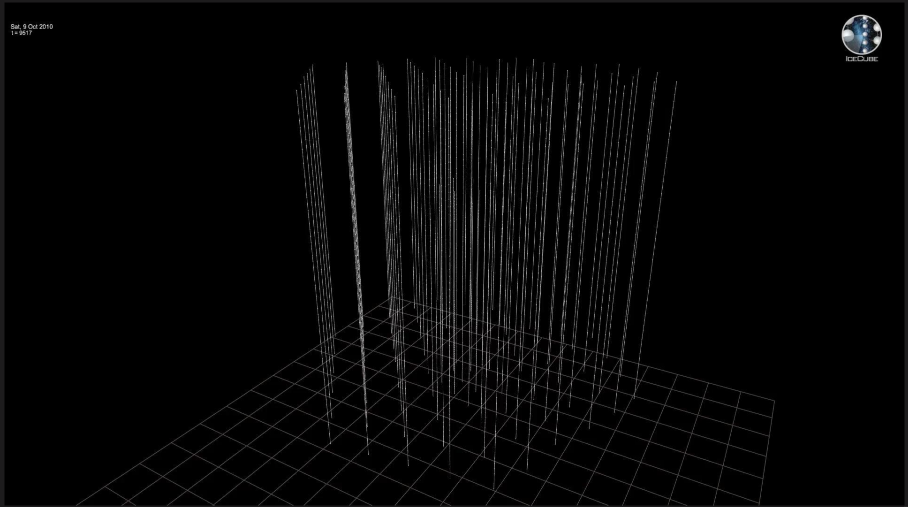


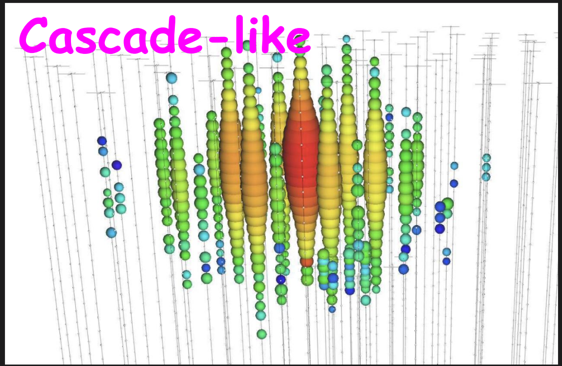
Charged particles travel faster than the speed of light in ice → emitting Cherenkov radiation

Optical modules record pulse charges & times

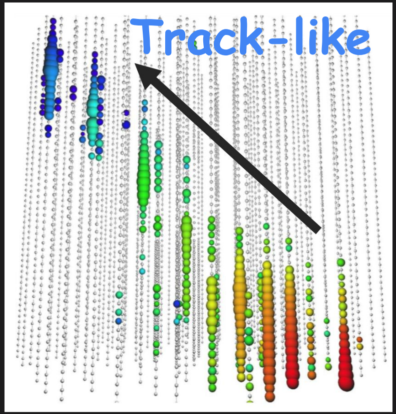


Astronomy Sector
IceCube
AMANDA
skyway
ADÉLIE
POPULATION: ~ 10 MILLION
GENTOO
POPULATION: ~ 775,000
CHINSTRAP
POPULATION: ~ 8 MILLION
EMPEROR

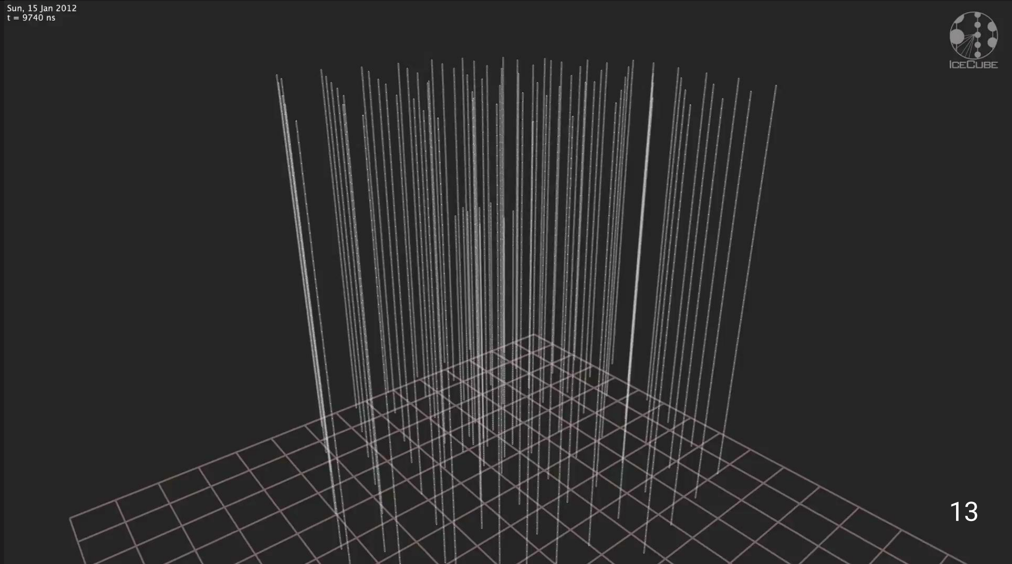




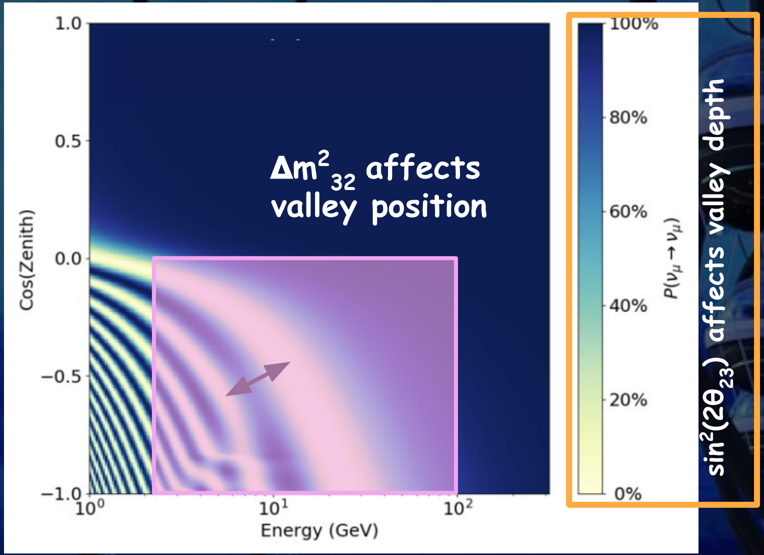
Source: ν_e CC, ν_τ CC, all NC



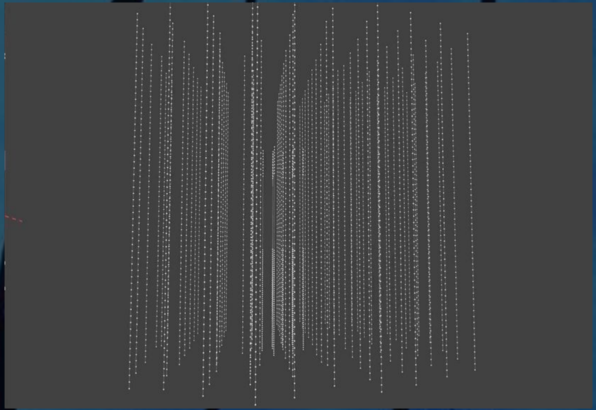
Source: ν_μ CC



10-GeV scale neutrino reconstruction in IceCube



Typical 10-GeV Scale Events

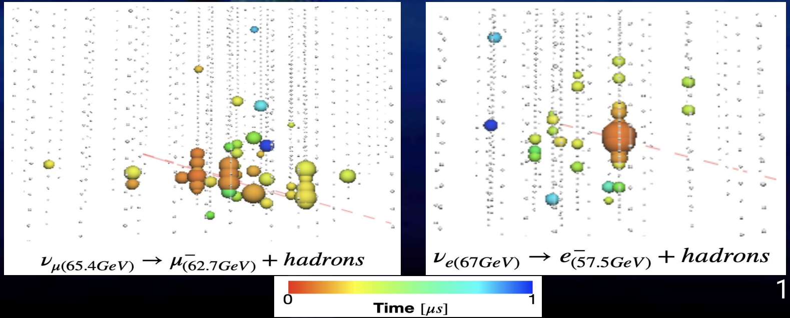


Challenges to reconstruct

- Direction
- Energy
- Track vs. cascade classification
- ……

Track-like events: ν_μ CC, 17% ν_τ CC

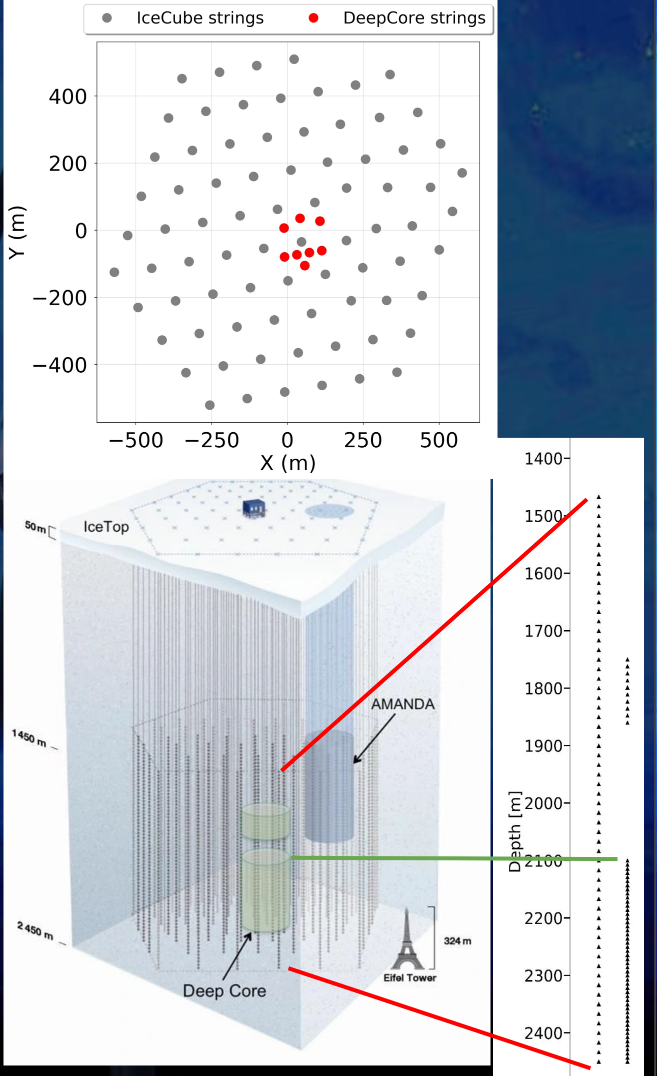
Cascade-like events: ν_e CC, NC, ν_τ CC



- Less light produced in low-E events: fewer pulses recorded by PMTs
- Need to consider IceCube/DeepCore array separately for spatial density and optical differences
- Need to optimize reconstructions for different variables

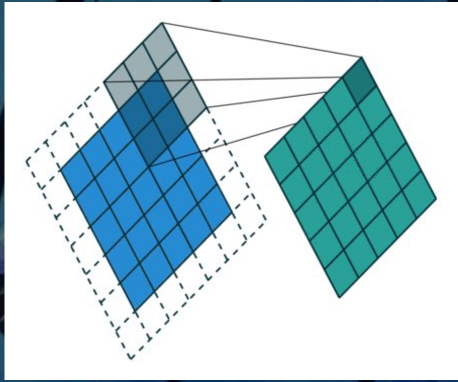
DeepCore Detector

- More densely instrumented in lower center, about 5x higher effective photocathode density
- Detects lower energy events
- Instruments total of 15 MTon
- Use the IceCube array to veto cosmic ray muon background



Convolutional Neural Network

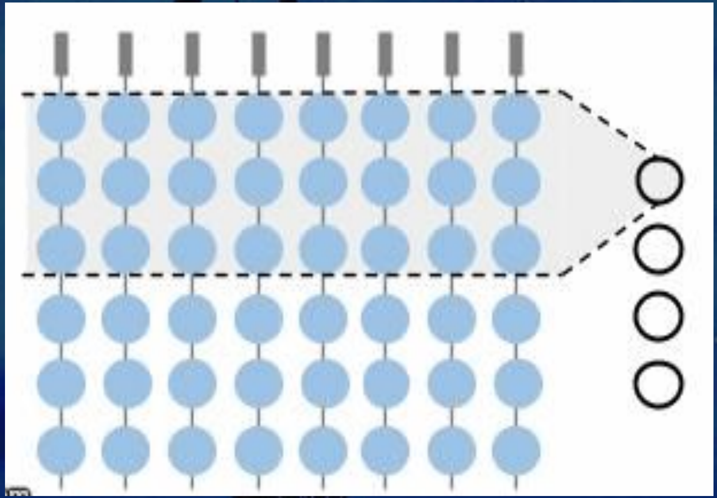
Convolutional Neural Network:



Training Layer 1 → Training Layer 2

CNN kernel (shadow) moving across the input map (blue) to extract features as output map (green).

CNN in IceCube:



CNN kernel in depth going down optical modules

CNN Gif Credit: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

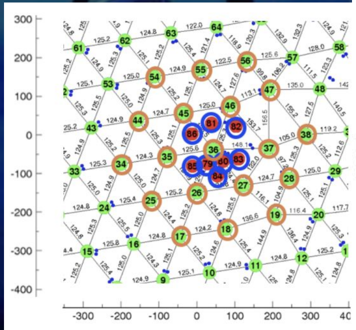
Convolutional Neural Networks

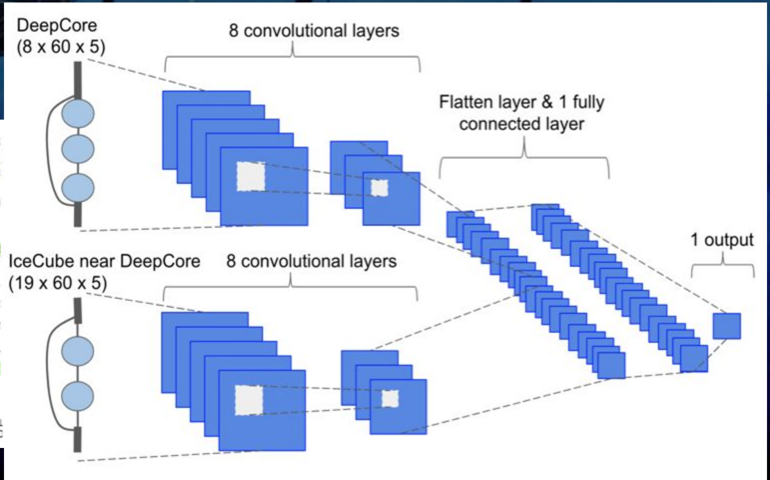
- Only use DeepCore & nearby IceCube strings;
- Noise cleaning applied & hit time within [-500, 4000] ns
- **Five CNNs** trained: optimized individually;
- Comparable resolution to the current likelihood-based* method;

Eur. Phys. J. C 82 (2022) 9, 807

Inputs: 5 summarized variables from all pulses hitting PMT

- sum of charges
- time of first (last) hit
- charge weighted mean (std.) of times of hits





Five reconstructed variables:

- PID: ν_μ CC vs. others
- Neutrino Energy
- Direction (L)
- Interaction vertex
- Atm. muon classifier

PID, Neutrino Energy, Direction (L): Analysis binning

Neutrino Energy, Direction (L), Interaction vertex, Atm. muon classifier: Selections

Convolutional Neural Networks

- Only use DeepCore & nearby IceCube strings;
- Noise cleaning applied & hit time within [-500, 4000] ns
- **Five CNNs** trained: optim
- Comparable resolution to ...iho... method; ...hys.

Inputs: 5 summarized variables from all pulses hitting PMT

- sum of charges
- time of first (last) hit
- charge weighted mean (std.) of times of hits

...tructed variables:

- PID: ν CC vs. others
- Neutrino Energy
- Direction (L)
- Interaction vertex
- Atm. muon classifier

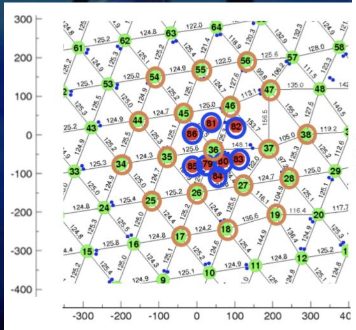
Analysis binning

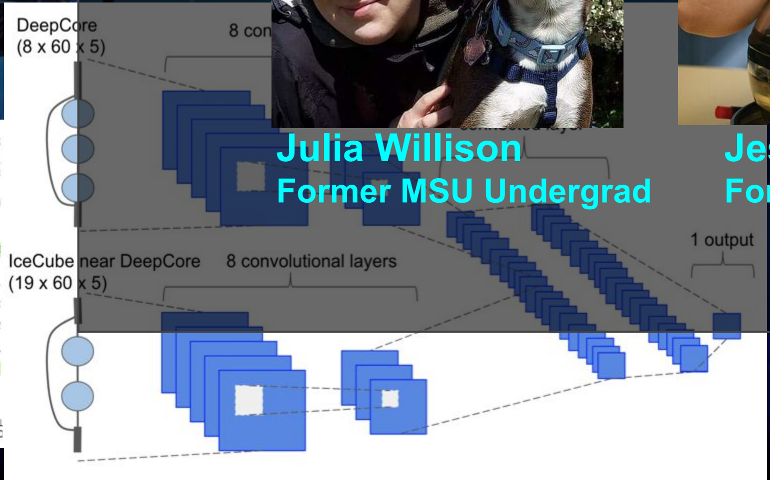
Selections

Students participated in CNN training/development

Julia Willison
Former MSU Undergrad

Jessie Micallef
Former MSU Grad

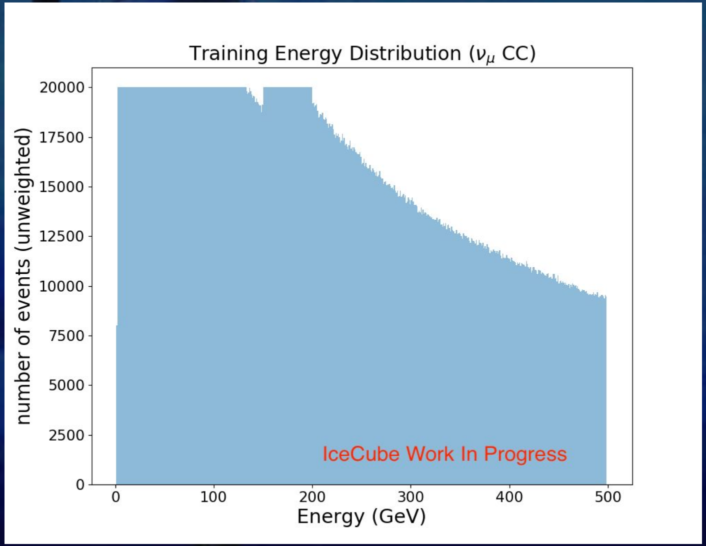




Training Samples: Energy & Interaction Vertex

~ 7 million ν_μ **CC** events

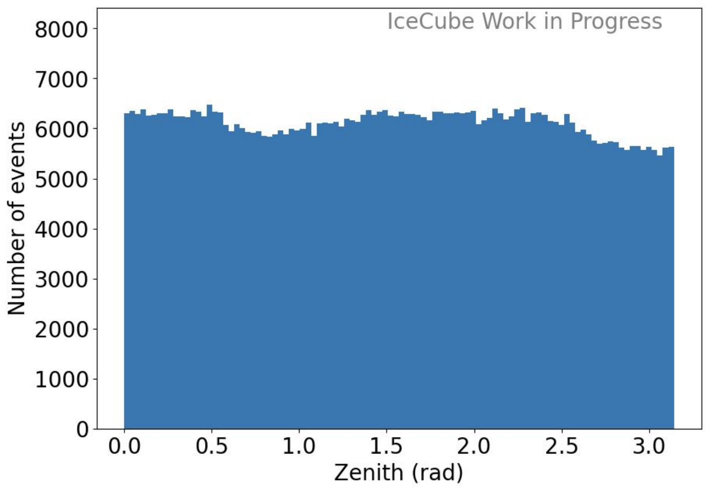
- Flat energy sample in region of interest
  - 1-200 GeV target
  - Extended to 500 GeV
  - nDOM >= 7
- Loss: Mean Absolute Percentage Error



Training Samples: Zenith

~5 million ν_μ **CC** events

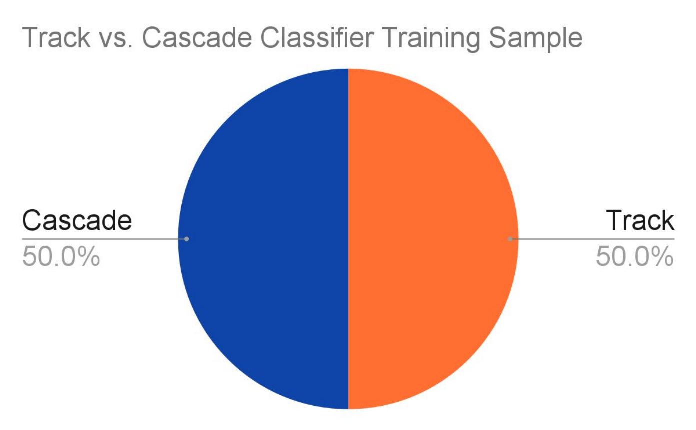
- Flat zenith sample across all zenith angles
  - 5-300 GeV
  - Containment cut on true starting & ending points
- Loss : MSE



Training Samples: Neutrino Identification (PID)

~6 million neutrino events in 5–200 GeV

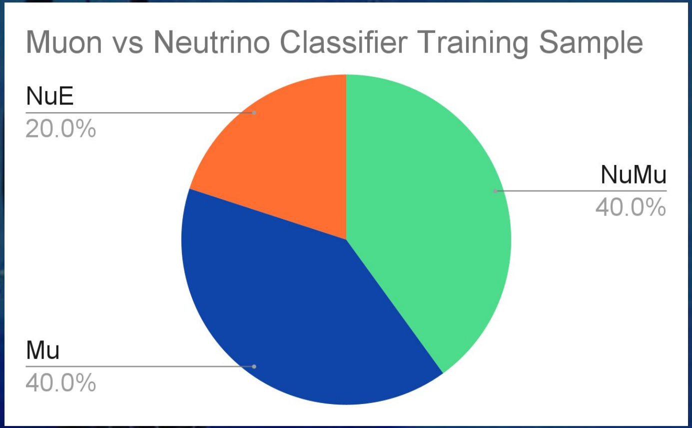
- Track: ν_μ **CC**
- Cascade: ν_μ NC and ν_e
- Loss: Binary Cross Entropy



Training Samples: Muon Classification

~7 million events

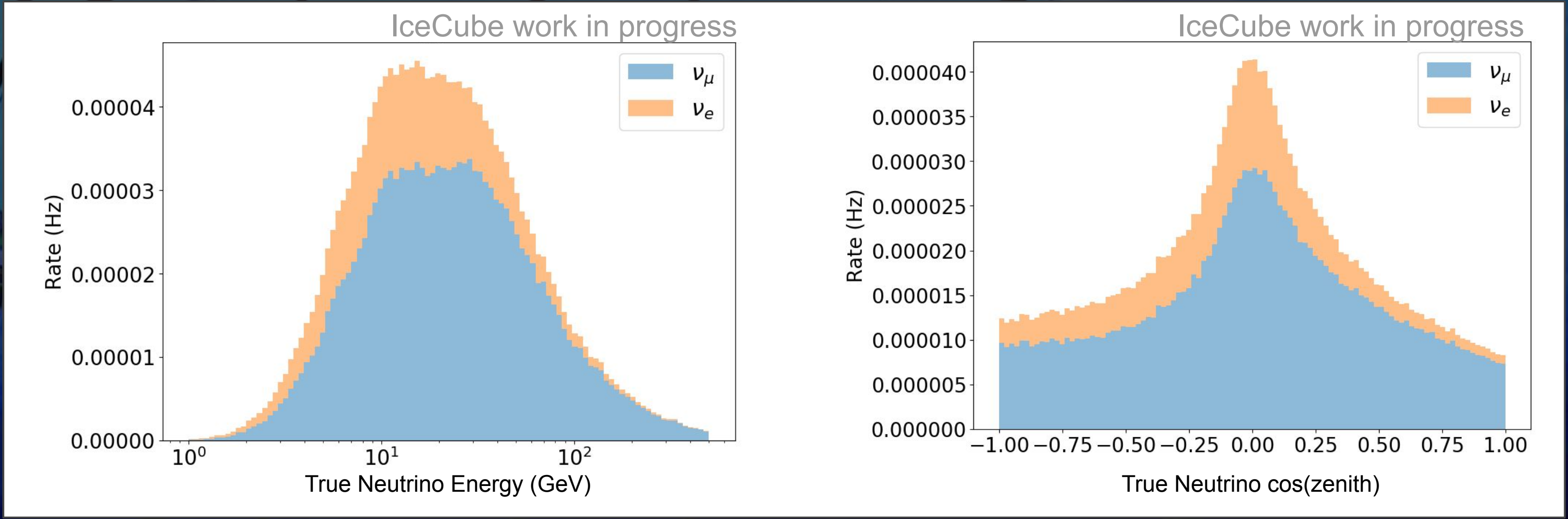
- Neutrinos are in 5–200 GeV
- Muon events: lower-level MuonGun (L4)
  - Before low-level muon BDT filter for more statistics
- nDOM >= 4
- Loss: Binary Cross Entropy
- Used as an input to a final muon BDT for a better performance on final level muon rejection



Performance of CNN Reconstruction

Testing Samples

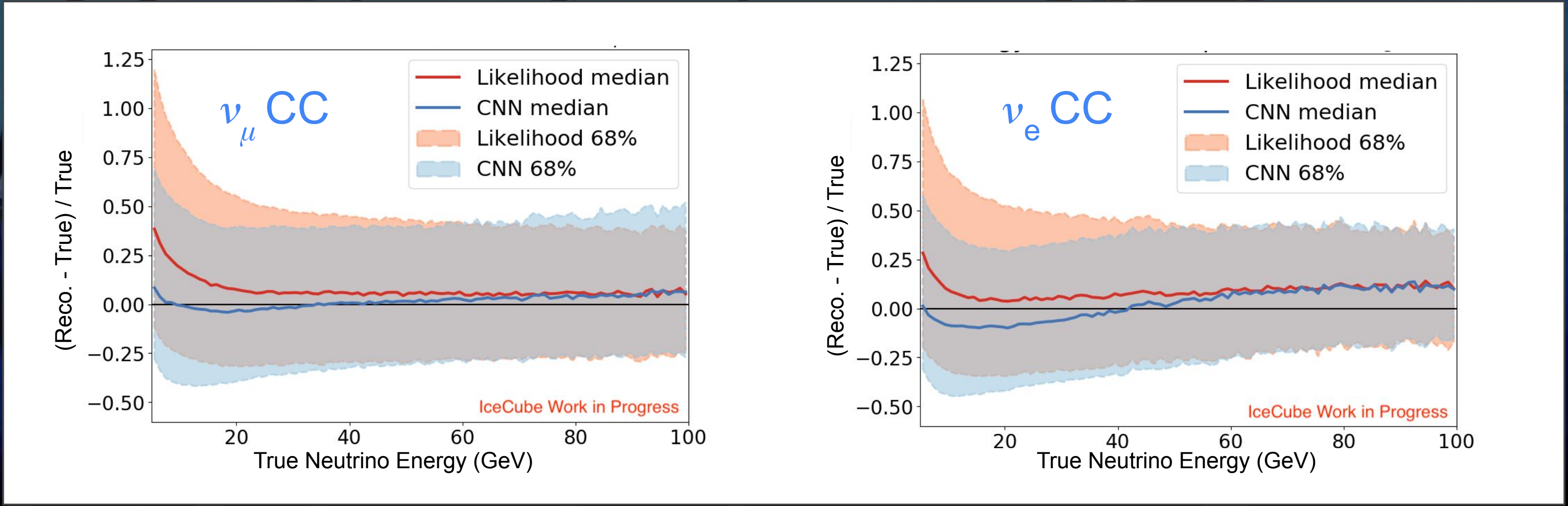
- Nominal MC sample with flux, xsec, and oscillation weights applied;
- Testing on signal (ν_μ CC) and major background (ν_e CC);
- Baseline: likelihood-based method



J. Micallef, et al. ICRC 2021 proceeding

Performance: Energy

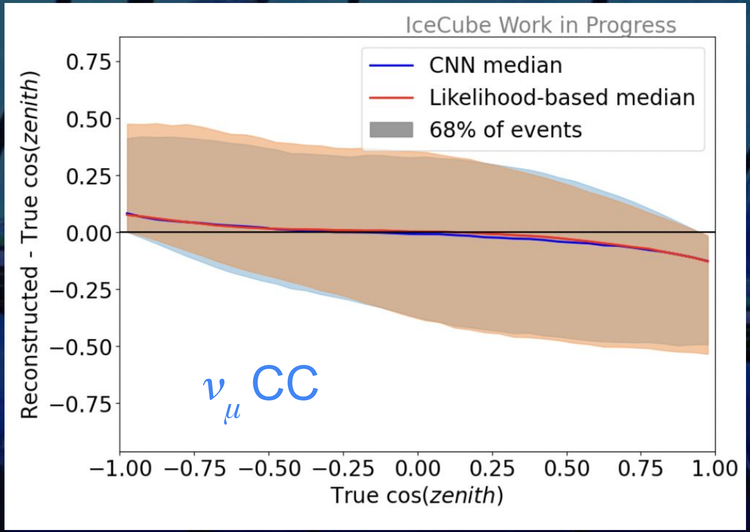
- Flat median against true neutrino energy;
  - CNN has better resolution at low energy (majority of sample)
- Comparable performance to likelihood method at higher energy and in background;

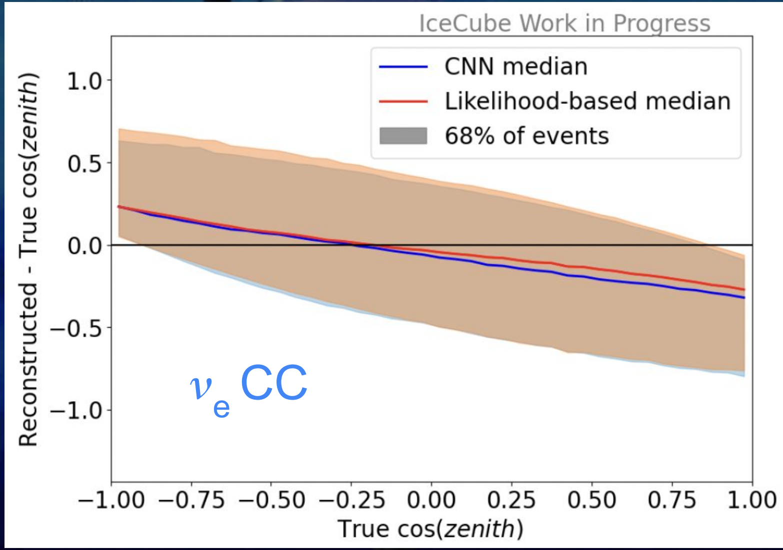


S. Yu, et al. ICRC 2021 proceeding.

Performance: Zenith

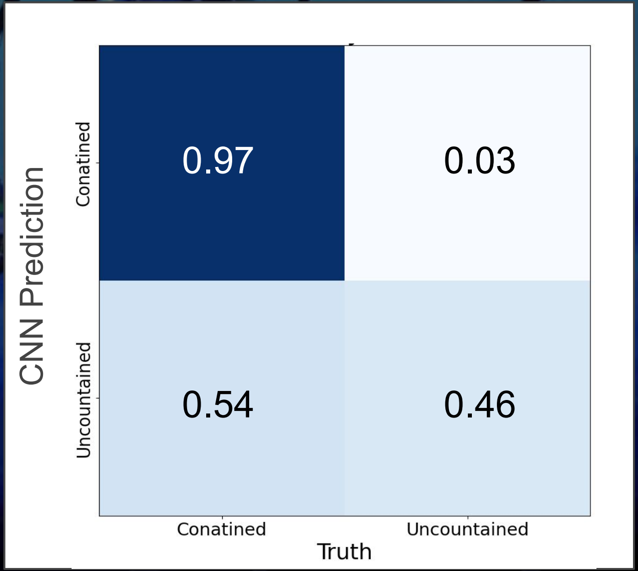
- Flat median against true direction;
- Comparable to current method in both signal and background.

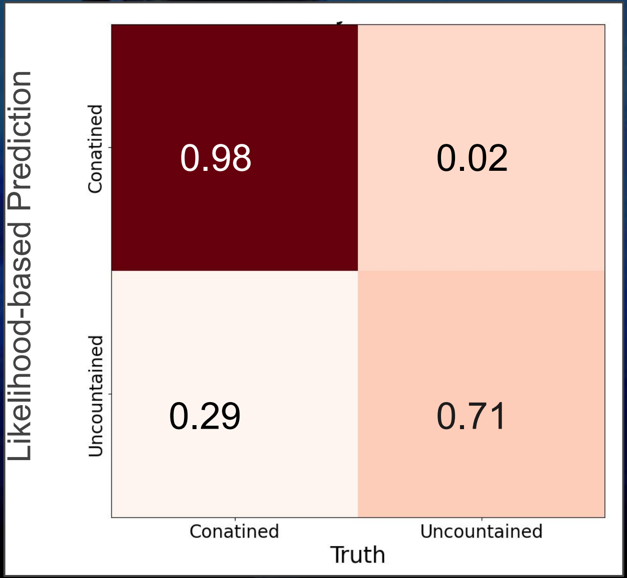




Performance: Vertex

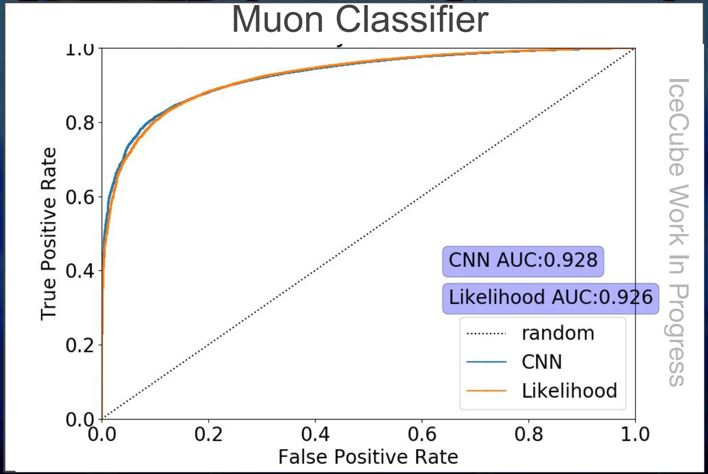
- Selecting events starting near DeepCore;
- Comparable purities* in selected ν_μ CC samples.
    - *Selected events are truly contained.

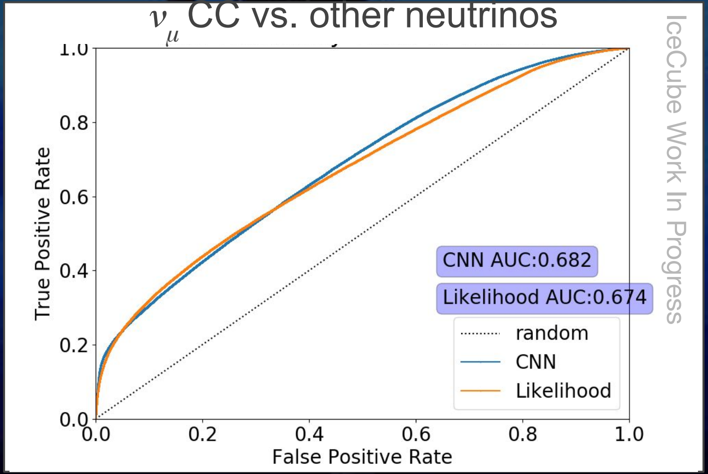




Performance: Muon and PID Classifiers

- Comparable performance to the current methods:
    - Similar AUC values.
- Hard to identify track from cascades at low energy → less DOMs see photons.





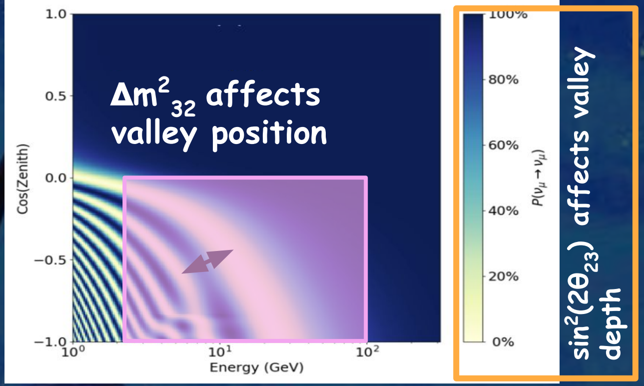
Performance: Speed

	Second per file (~3k events)	Time for full sample assuming 1000 cores
CNN on GPU	21	~ 13 minutes
CNN on CPU	45	~ 7.5 hours
Current Likelihood-based method (CPU only)	120,000	~ 46 days

- CNN runtime improvement: **about 5,000 times faster;**
  - CNNs are able to process in parallelize with clusters → can be even faster!
- Big advantage: large production of full Monte Carlo simulations ~$O(10^8)$.

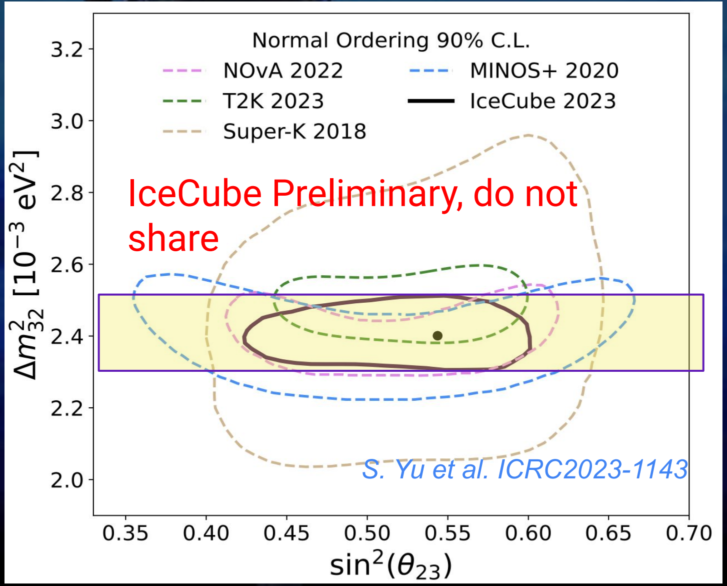
Oscillation result and future outlook

Application in Oscillation Analysis

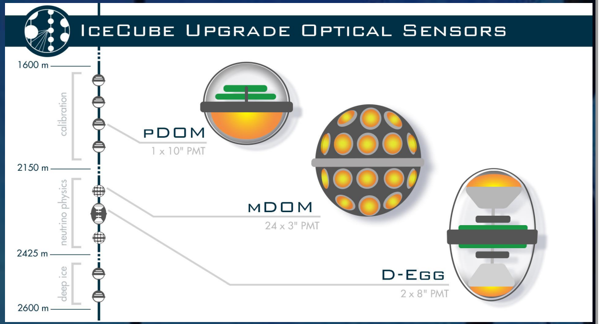


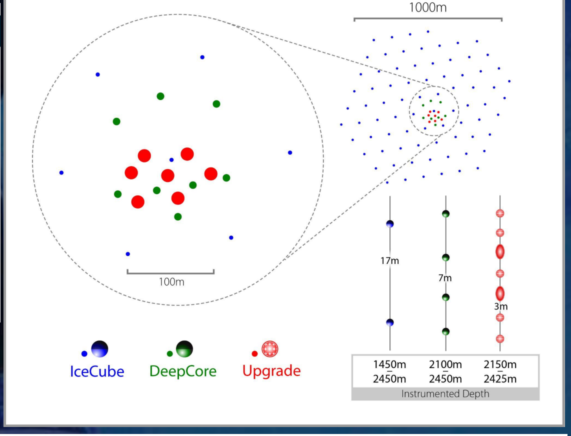
A compatible and complementary result compared with the existing measurements, with room for future improvements (MC models, calibration, reconstruction, etc…)

- Competitive on Δm^2_{32} measurement.



Future





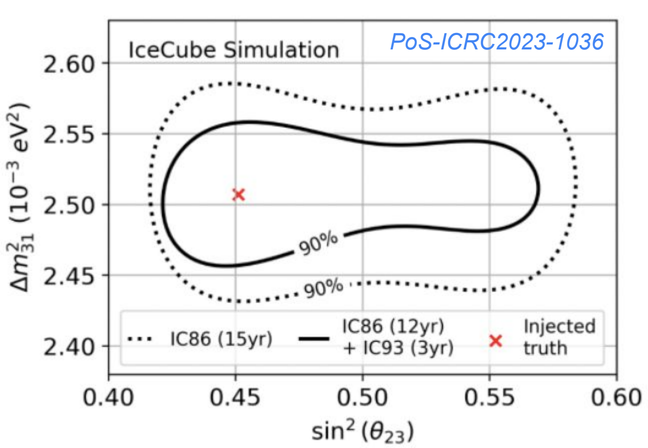
Upcoming DeepCore results of neutrino physics:

- mass ordering, non-standard interaction, etc…

New reconstruction development: GNN, RNN, Transformer…

<u>**The Upgrade detector:**</u>

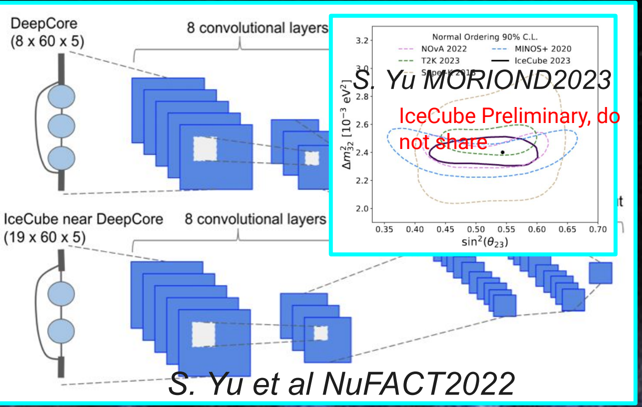
- More densely instrumented strings in the center
  - Better event resolution!
- DOM: multiple PMT designs
  - Great for calibration studies!
- Target deploying 2025/26



After 3 years running, expect 20-30% improvement on the sensitivity to the atmospheric mixing parameters θ_{23} and $\Delta m^2{}_{32}$

Enabling Multimessenger Science and Neutrino Physics with **Machine Learning in IceCube**

Atmospheric neutrino



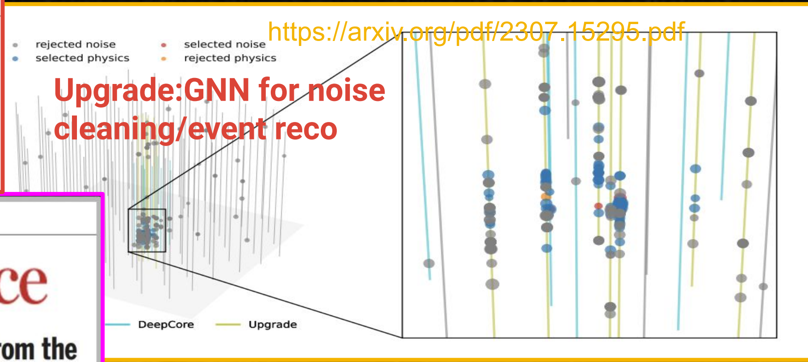
RESEARCH

RESEARCH ARTICLE

NEUTRINO ASTROPHYSICS

Evidence for neutrino emission from the nearby active galaxy NGC 1068

IceCube Collaboration*†



RESEARCH

RESEARCH ARTICLES

Science

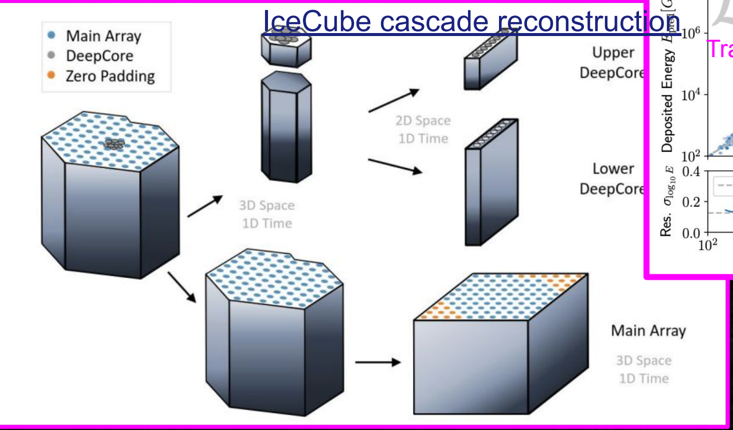
NEUTRINO ASTROPHYSICS

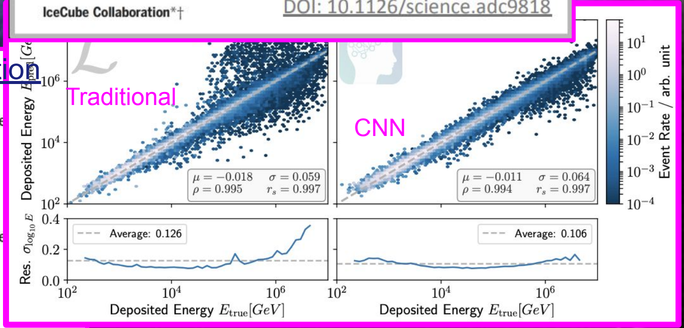
Observation of high-energy neutrinos from the Galactic plane

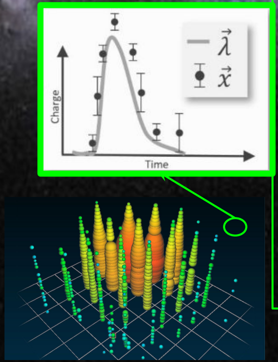
IceCube Collaboration*†

DOI: 10.1126/science.adc9818

Astrophysical neutrino







Hybrid approach (ongoing): NN predict photon distribution feed to likelihood (traditional) method

Use **Neural Networks** to reconstruct high-energy astrophysical and atmospheric low-energy neutrinos (*3k times faster* than traditional methods) allowing for timely reconstruction and great scientific outcomes.

Conclusion

- CNNs are used for **multipurpose reconstructions** for IceCube oscillation analysis:
  - Energy, direction, interaction vertex;
  - PID (numu CC vs. background neutrinos), muon classifier.
- Approximately **5000 times faster** in runtime than the current method;
  - Big advantage for IceCube full production → large atmospheric neutrino sample.
- CNNs have **better or comparable performances** to the traditional likelihood-based reconstruction method;
- The result using CNN-reconstructed neutrino dataset reports compatible and complementary measurement on ν_μ disappearance parameters
- Many more ML developments and applications in IceCube in all energy range for different scientific motivations.


2025/26 winter:
Will drill 7 holes and deploy
Upgrade strings/DOMs!
Chance to visit the Pole!

Thank you!

IceCube Collaboration Meeting
in 2025 Fall will be @ U of Utah!

>400 collaborators
>200 students and postdocs
>60 institutes over 14 countries

IceCube Collaboration Meeting
March 2024, Munster, Germany

Overflow slides