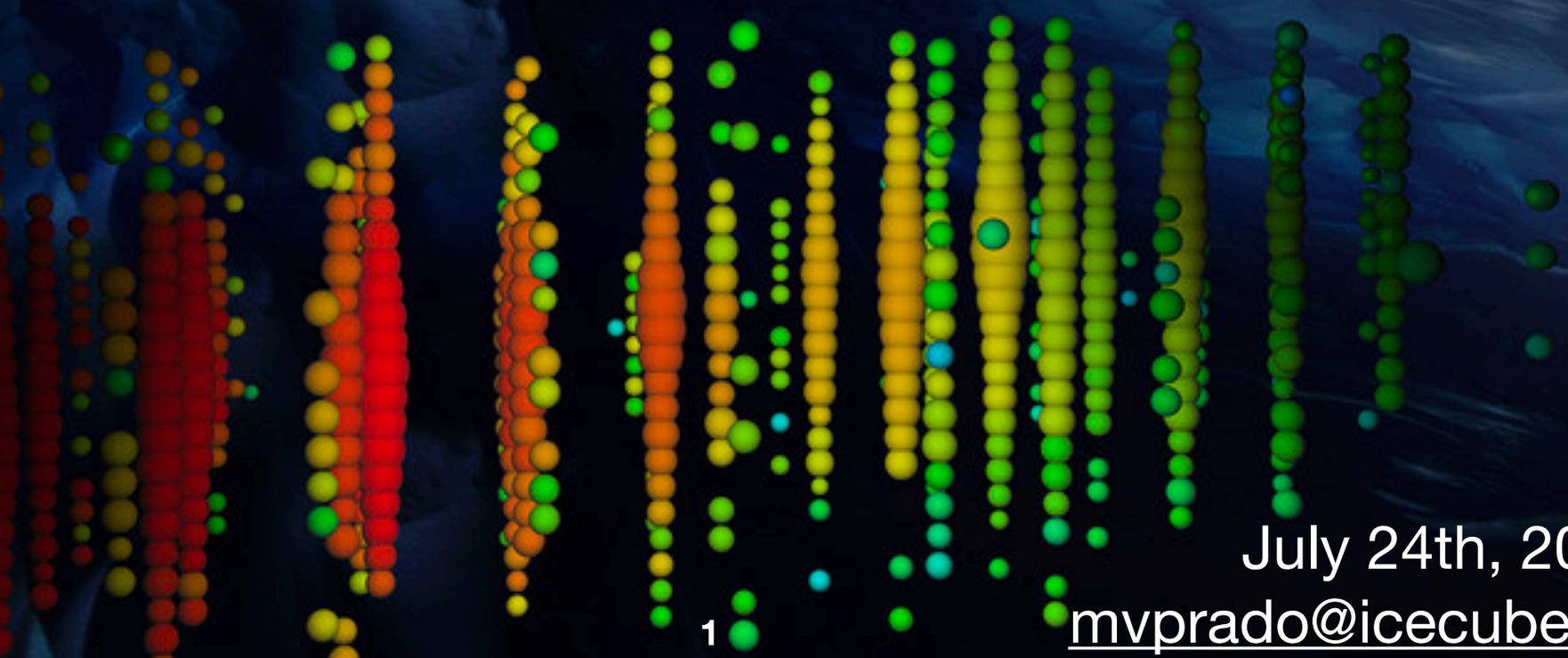
Neutrino Physics and Machine Learning

Deep Learning Classifier for Low-Energy Events in IceCube

Maria Prado Rodriguez
for the IceCube Collaboration



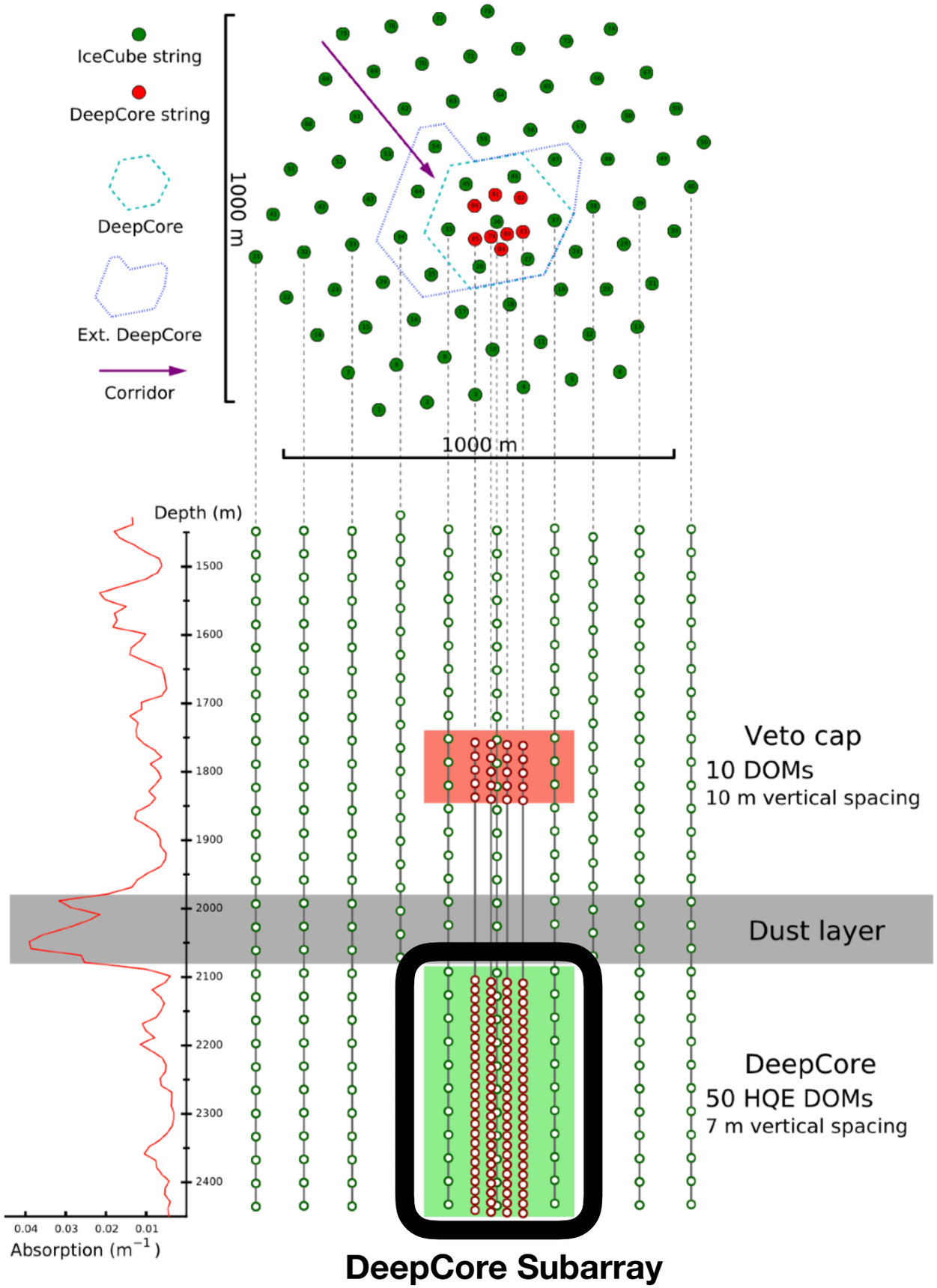
July 24th, 2020

mvprado@icecube.wisc.edu

Overview

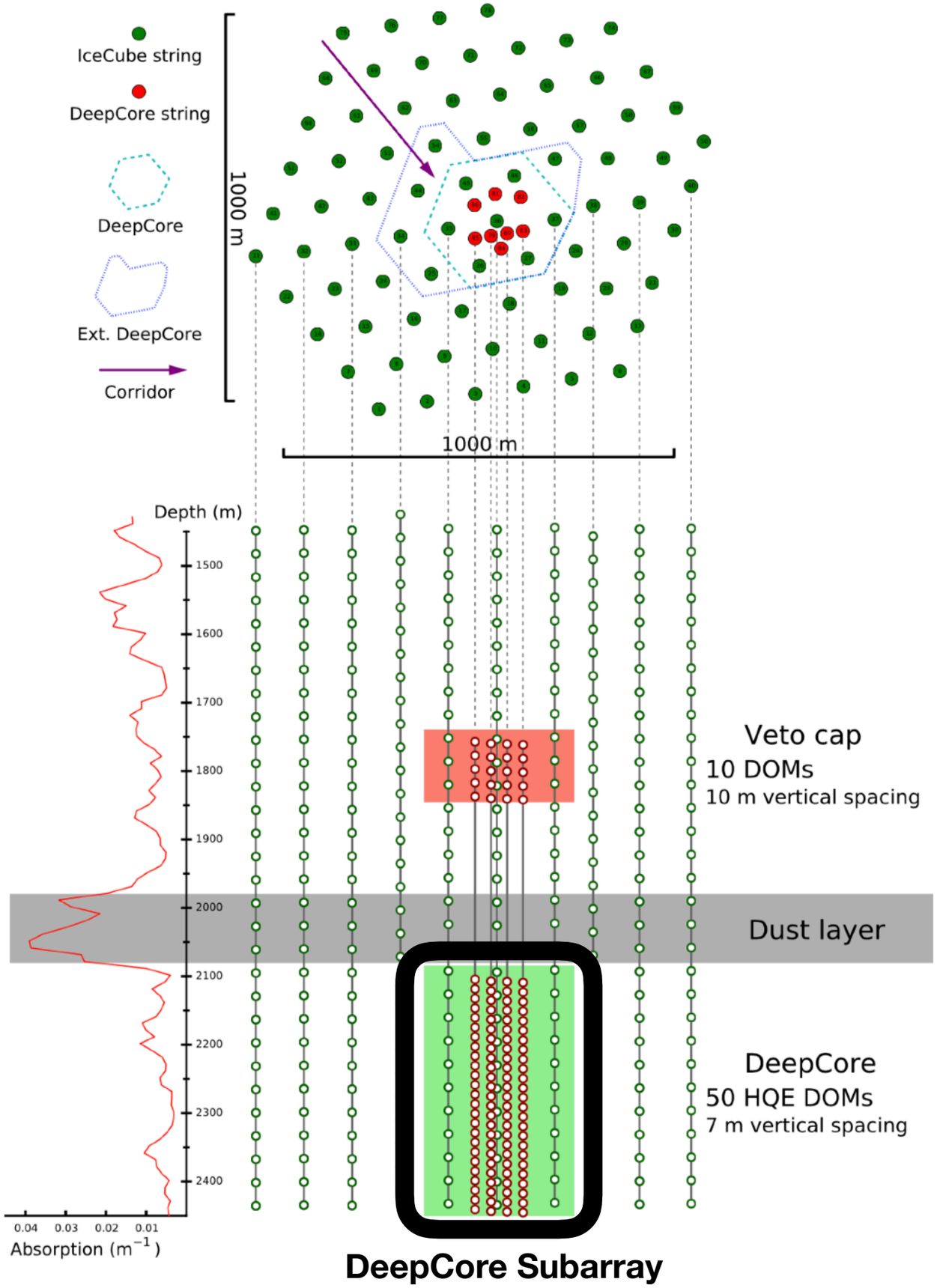
- DeepCore Subarray
- Neutrino Oscillations
- Event Classification in DeepCore
- 2D CNN DeepCore Classifier
- Future Improvements

DeepCore Subarray



	IceCube	DeepCore
Number of strings	78 IceCube strings	8 DeepCore strings
Vertical distance between DOMs	17 meters	7 meters
Horizontal distance between DOMS	110-150 meters	40-90 meters
Total DOMs per string	60	60

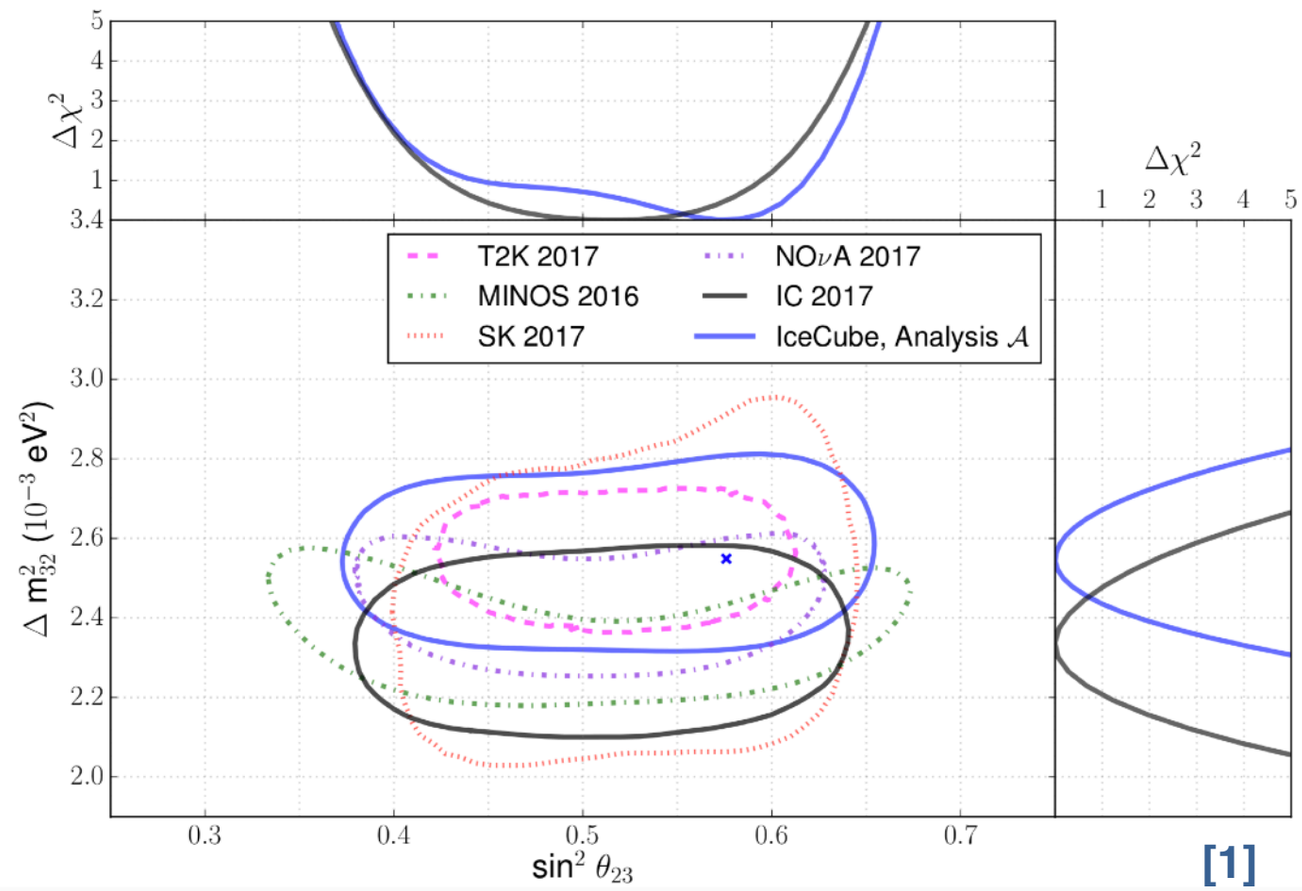
DeepCore Subarray



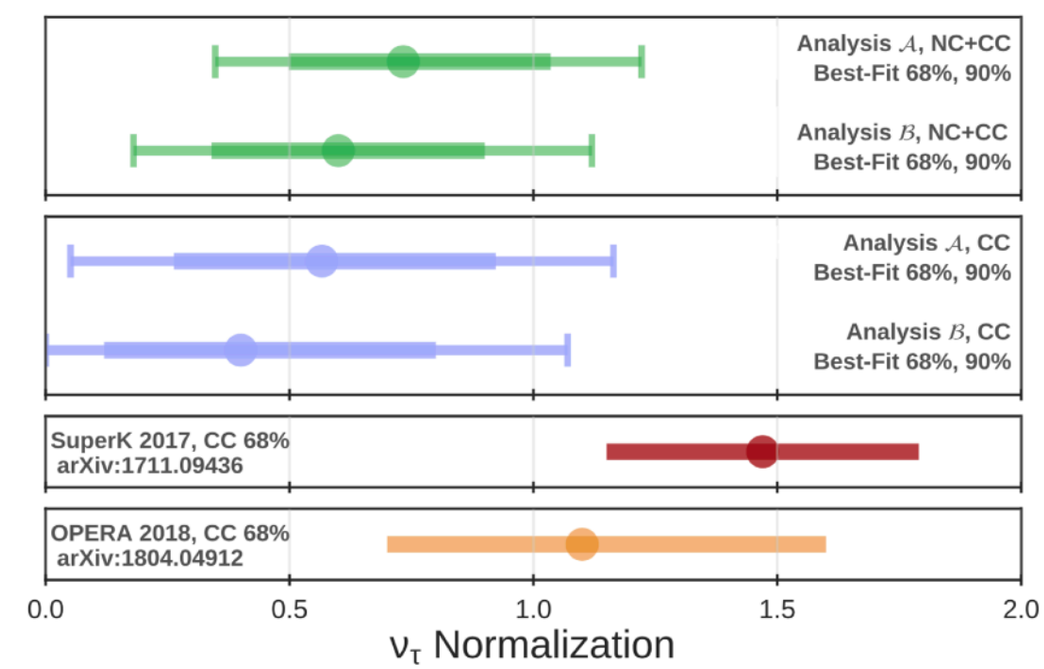
- Tighter vertical spacing provides better event resolution
- Can probe events with energies as low as 5 GeV
- Study “upward-going” neutrinos coming from cosmic rays hitting the Earth’s atmosphere
- “Upward-going” means neutrinos that traverse the Earth

Neutrino Oscillations

- Atmospheric muon neutrino oscillations are studied with DeepCore
- Looks into:
  - ❖ Whether θ_{23} is maximal
  - ❖ Unitarity of the PMNS matrix via tau normalization value
- Neutrino energies most relevant for these oscillation studies range from 20 to 35 GeV
- Events at these low energies are sorted as track-like or cascade-like (no double-cascades)
- Tracks: $\nu_\mu CC$, cascades: every other neutrino event
- $\nu_\mu \rightarrow \nu_\tau$ oscillations, for example, appear as an excess in the number of cascades observed
- Therefore, accurate classification is crucial

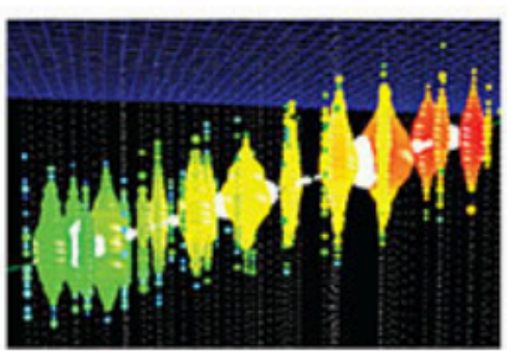


[1]

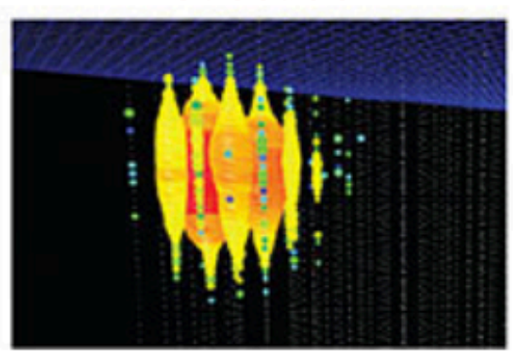


[2]

Track



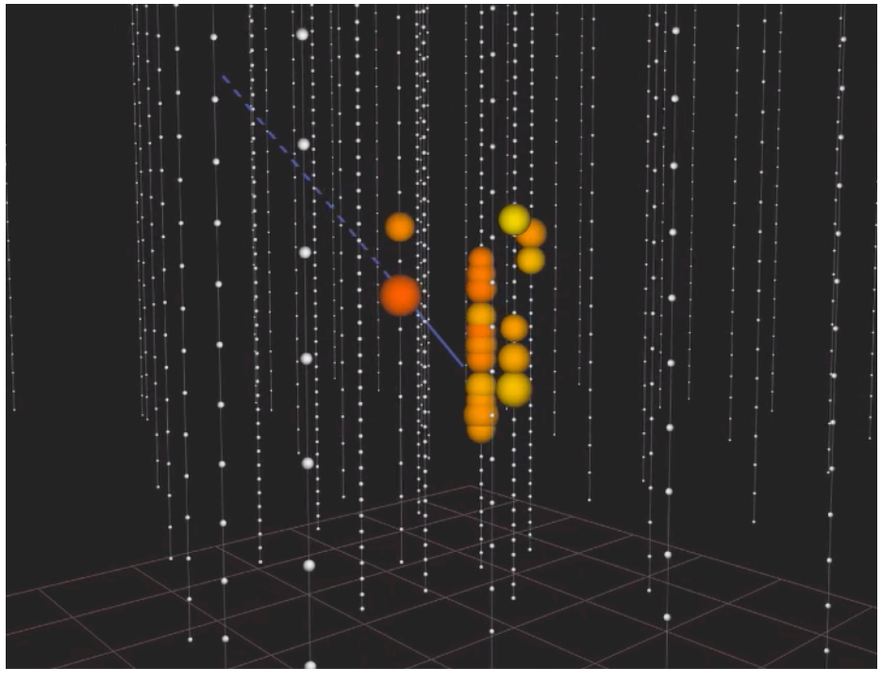
Cascade



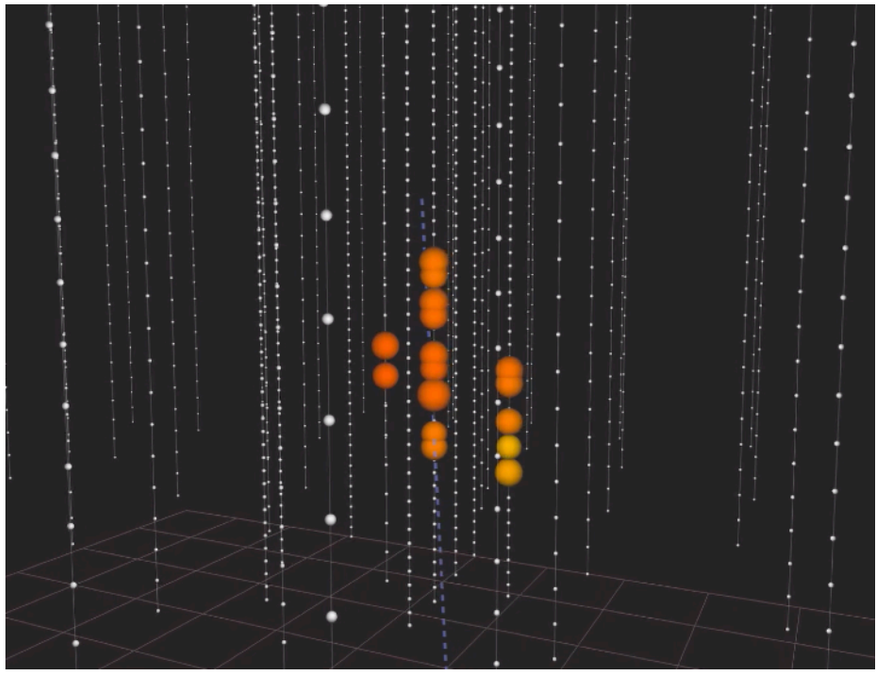
Event Classification in DeepCore

- **Current method:** Feed reconstructed quantities into an XGBoost Boosted Decision Tree for learning
- **Goal:** Use deep learning to exploit symmetries in simulated data for classification
- **Problem:** Mapping DeepCore to an input grid proves more challenging than it sounds, given that it does not have good X-Y-Z spatial symmetry
- **One solution:** Map the DeepCore PE charge hits to a 3D grid (t, z, strings) where the string dimension is a channels dimension

Track with energy of 26 GeV

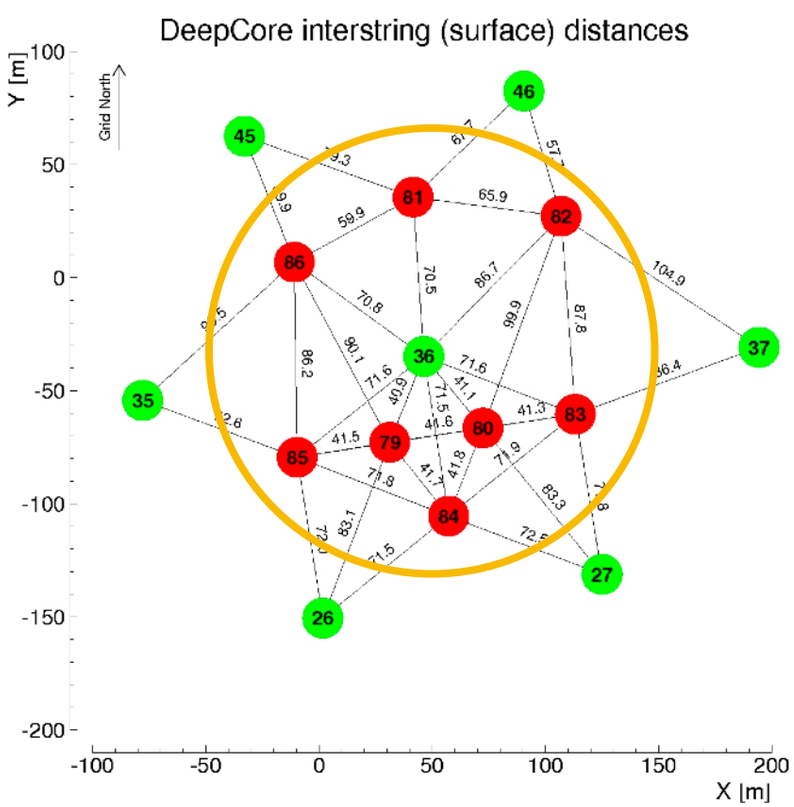


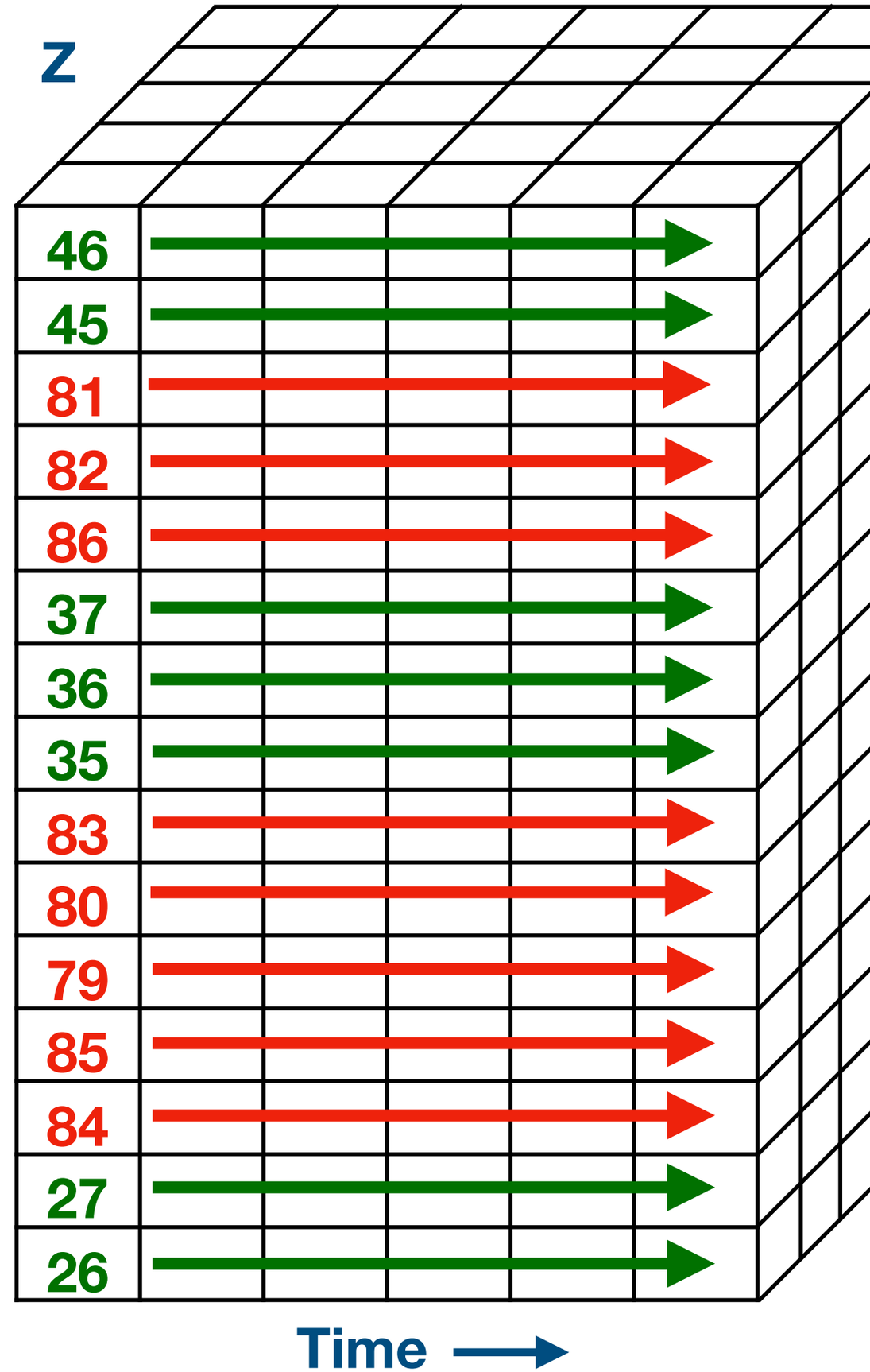
Cascade with energy of 30 GeV



2D Event Classifier

- Event images have dimensions:
  (t, z, strings) = (40, 49, 15)
- Excluded region in ice of increased light-scattering and absorption (Dust Layer)
- Energies > 10 GeV for better resolution
- Events have at least 8 DOM charge hits
- Events start within radius of 100 m from center of DeepCore (orange circle below)



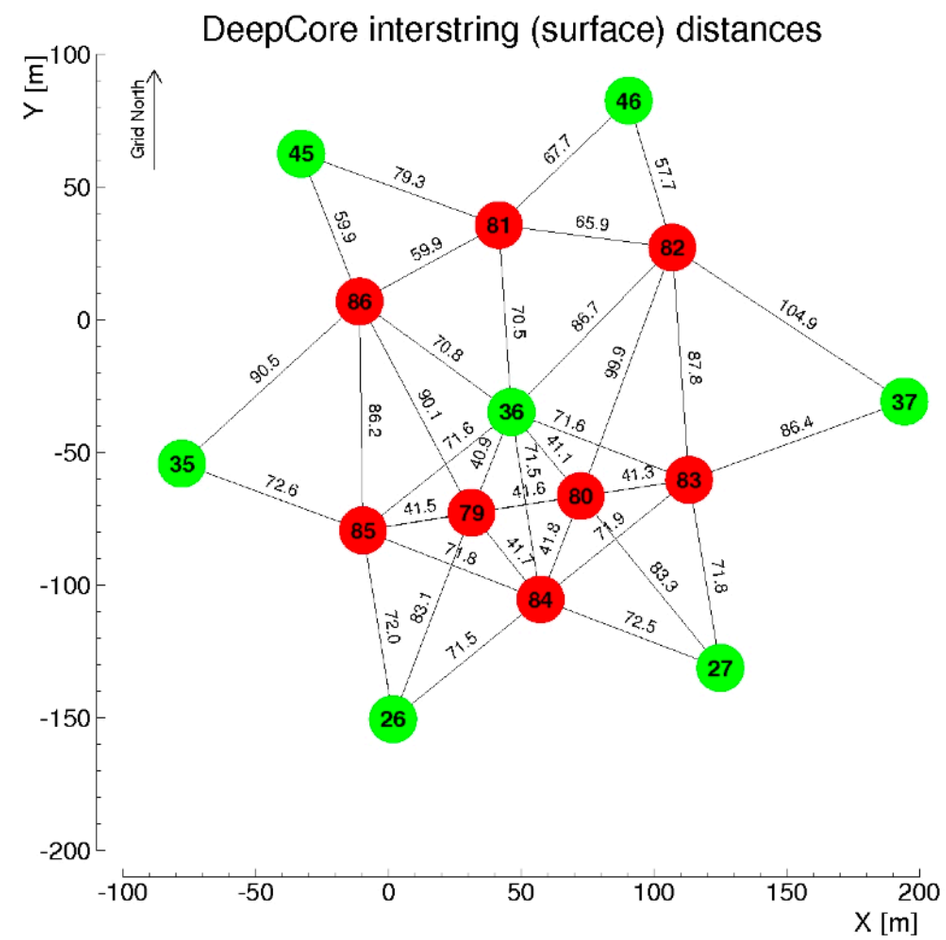


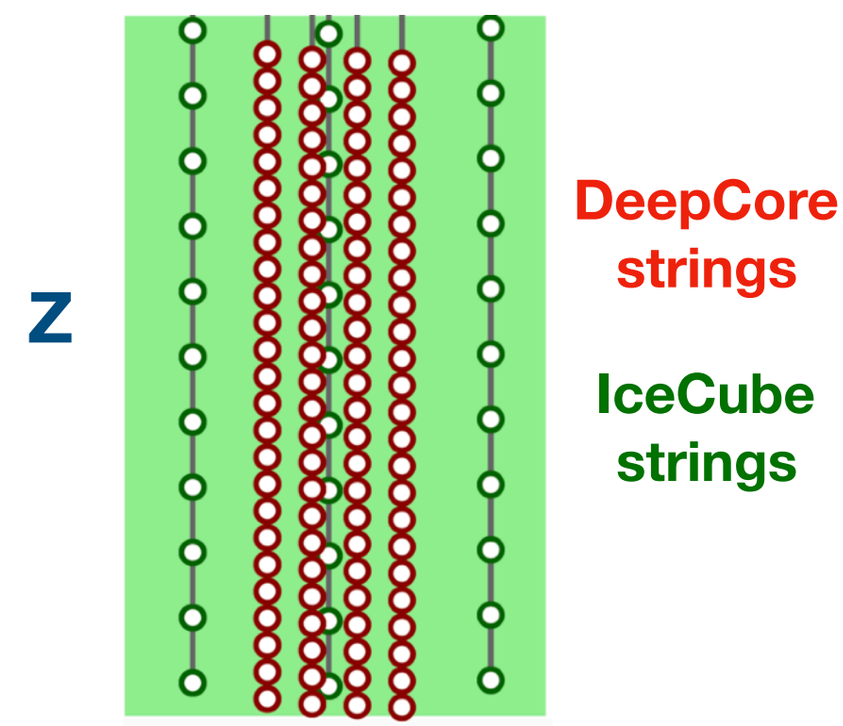
DeepCore strings

IceCube strings

2D Event Classifier

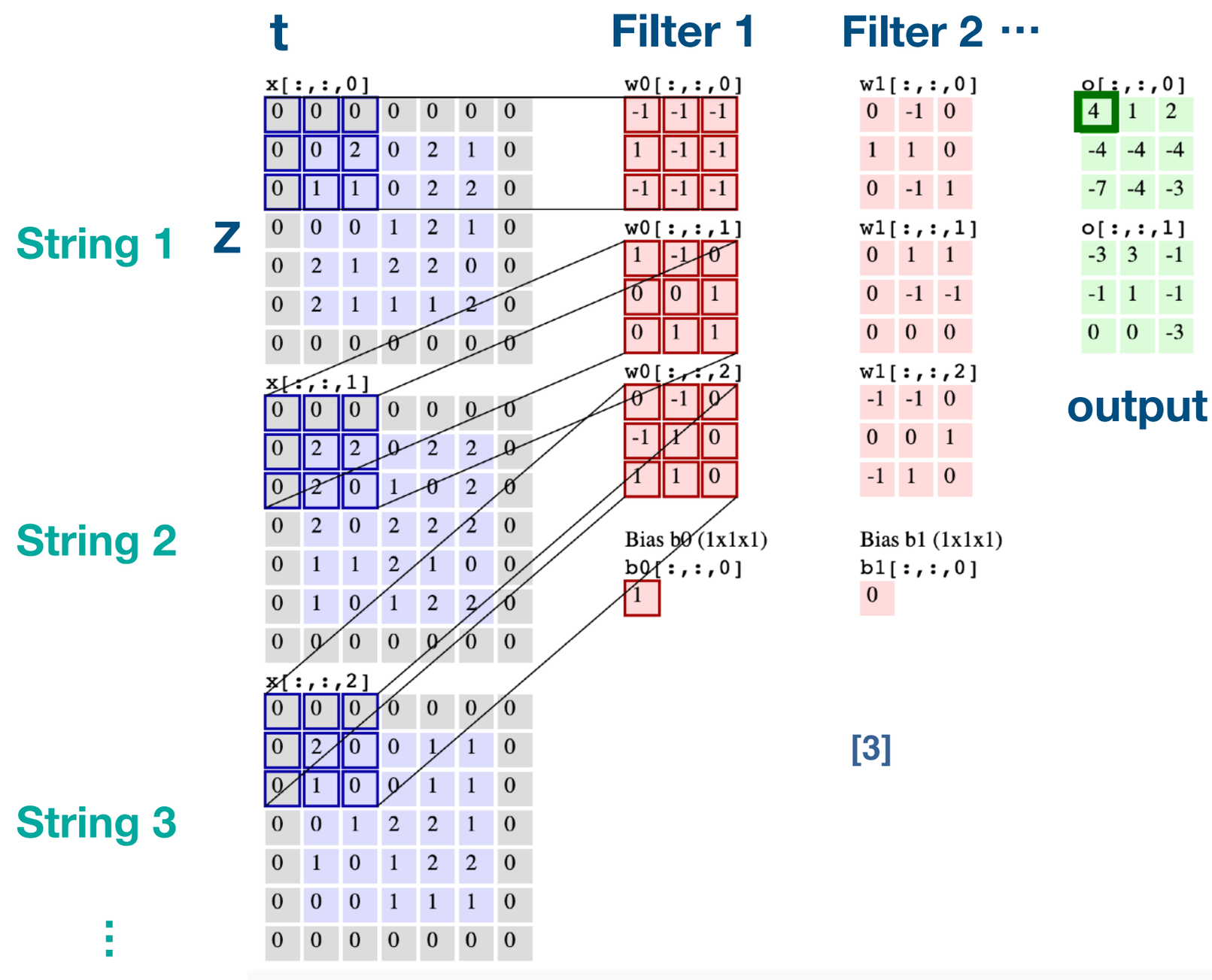
- Accounted for the difference in z-axis spacing for DeepCore and IceCube strings
- Padded IceCube strings with zeros in between for more accurate mapping
- Still not a perfect mapping
- Timing windows are 25 nsec apart
- A 2D CNN over time and z (depth)
- Mapping DeepCore subarray hits is more accurate using time and z (depth) than with x and y
- String dimension is a channels dimension: no specific ordering needed
- Cross-channel patterns





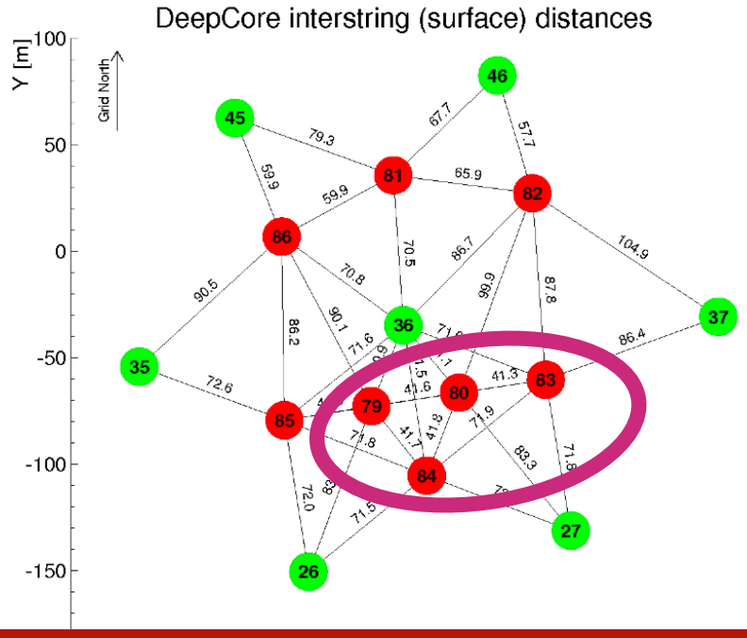
CNN with Strings as Channels

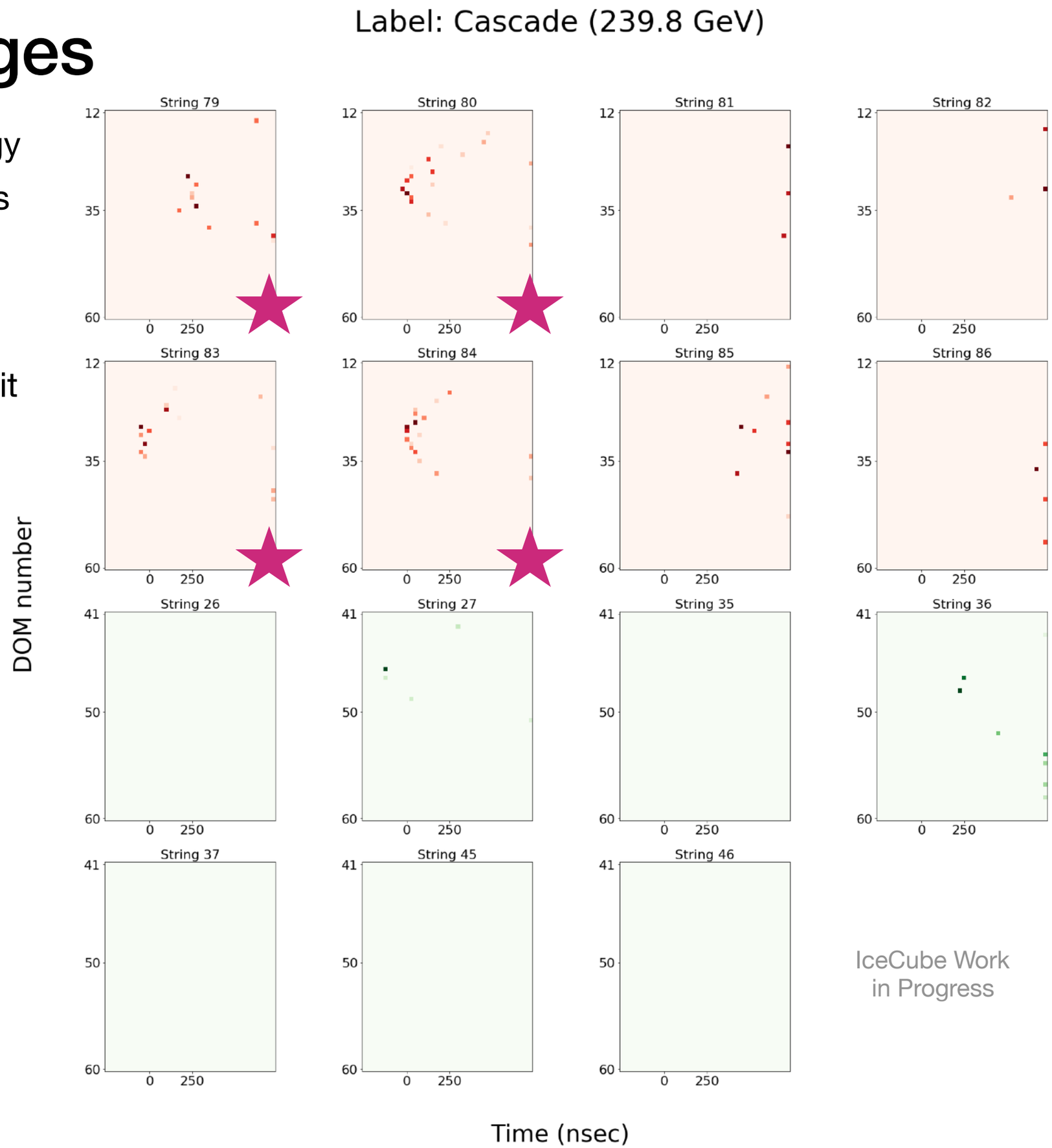
- <u>Cross-channel patterns:</u> Multi-channel CNN highlights and merges together patterns found in each channel
- CNN looks for patterns simultaneously in all channels and finds similarities between channels
- Strings positioned near each other that are hit by an event should exhibit somewhat similar patterns
- By looking at the learned filter values, the CNN can use these similar patterns as a form of linking together the strings that are close to each other, gaining back some information on the x-y axis position



2D DeepCore Images

- Cascade event with a higher energy than that of most DeepCore events
- Highlights patterns better than lower-energy events
- Shows similar patterns in strings hit that are close to each other (magenta stars)
- Sideways “V” is a characteristic pattern of cascades even though some tracks can still exhibit this pattern

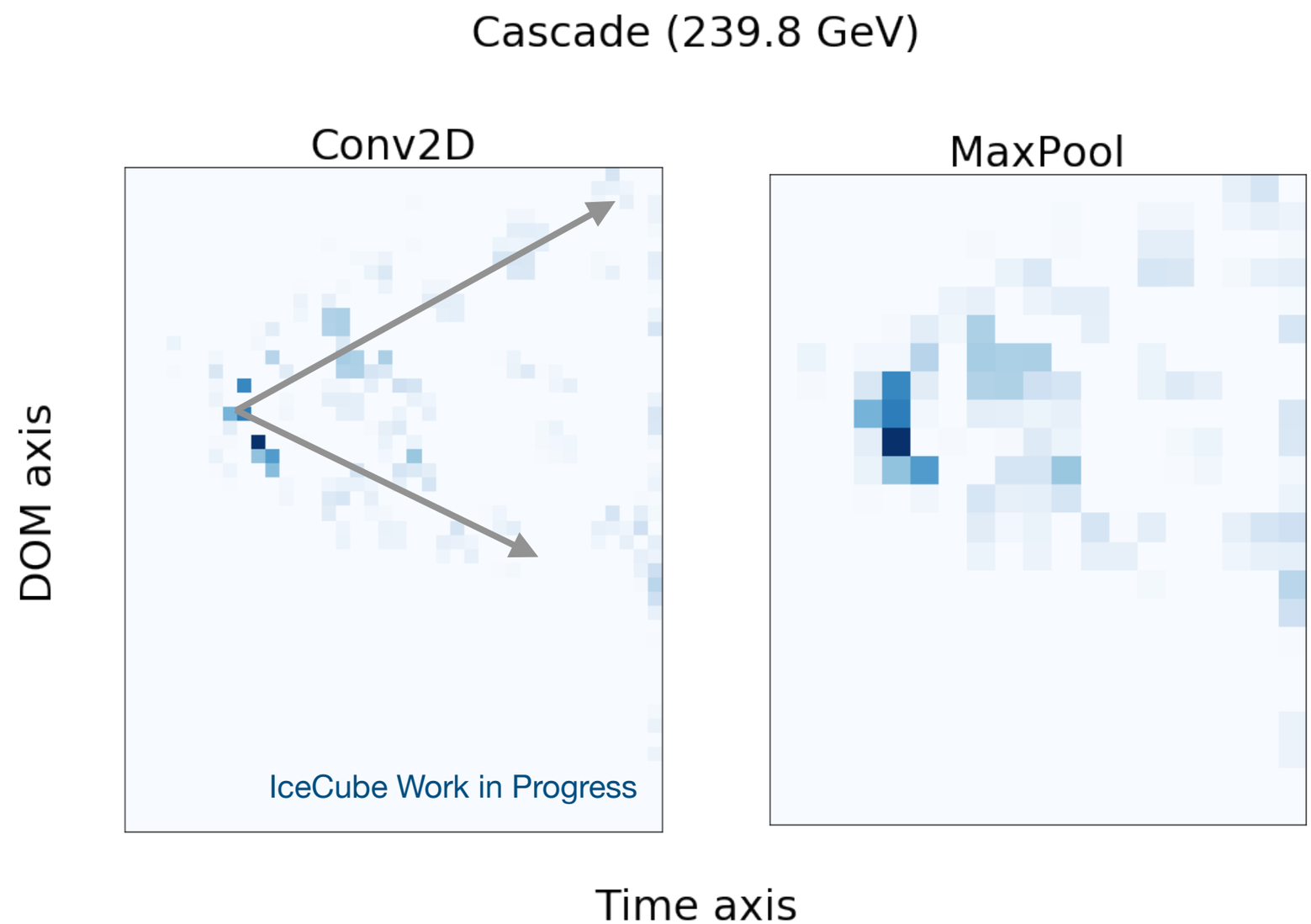




IceCube Work in Progress

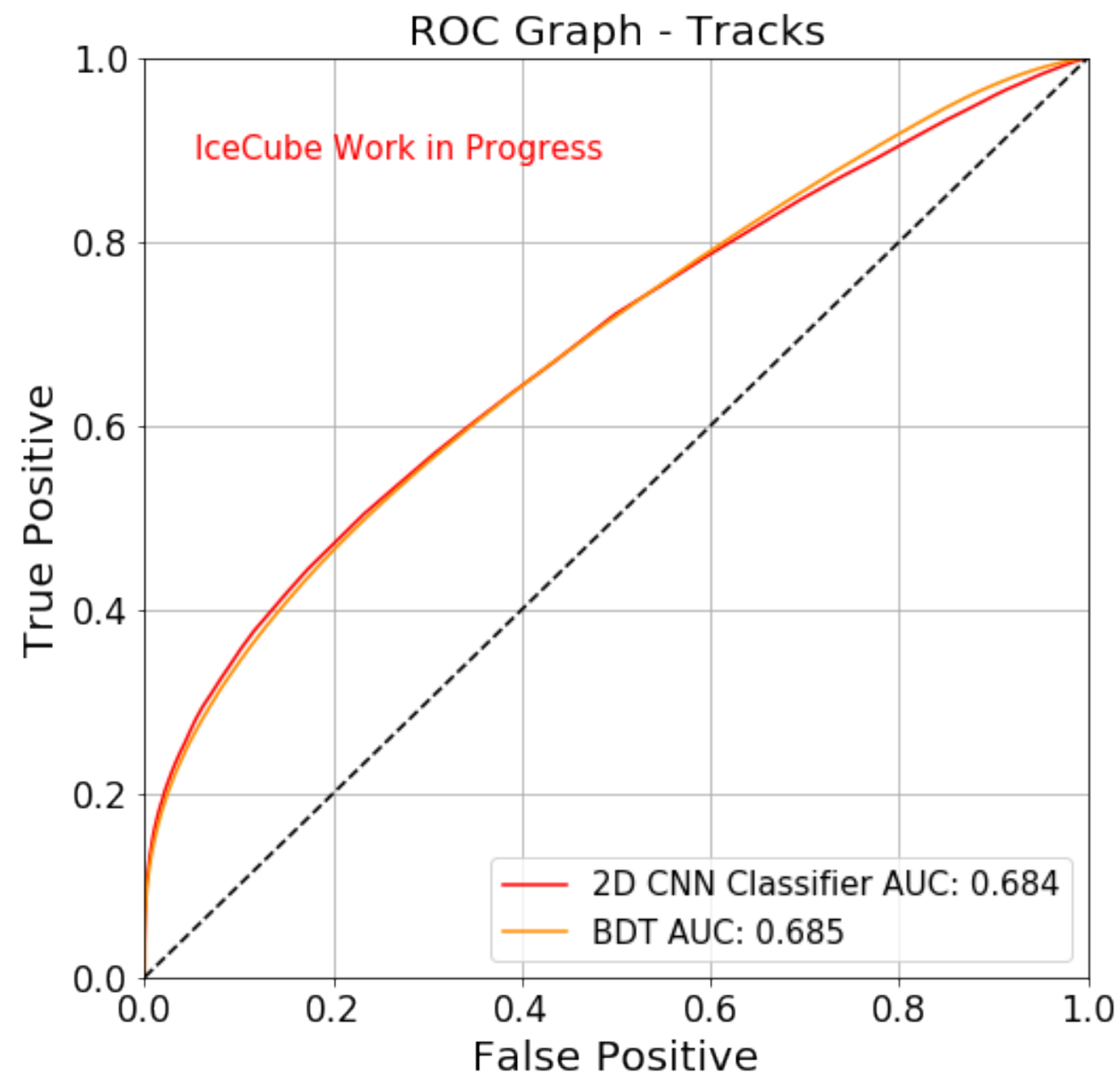
2D DeepCore CNN

- CNN output for 1 filter
- Can see the merged sideways “V” in the top-middle section of the output
- The goal is that patterns like these can still be found in events with energies of about 20 to 35 GeV

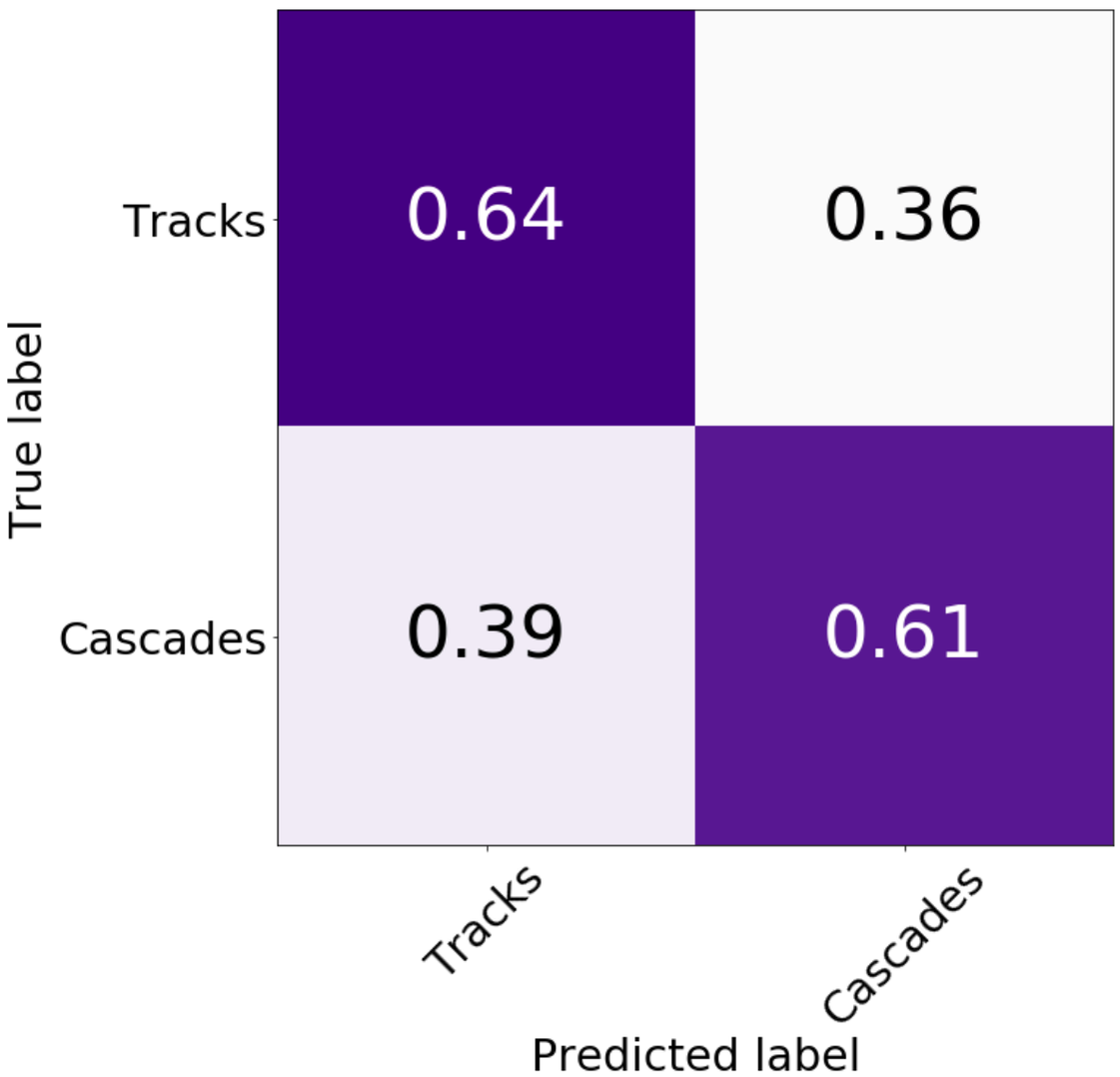


Performance

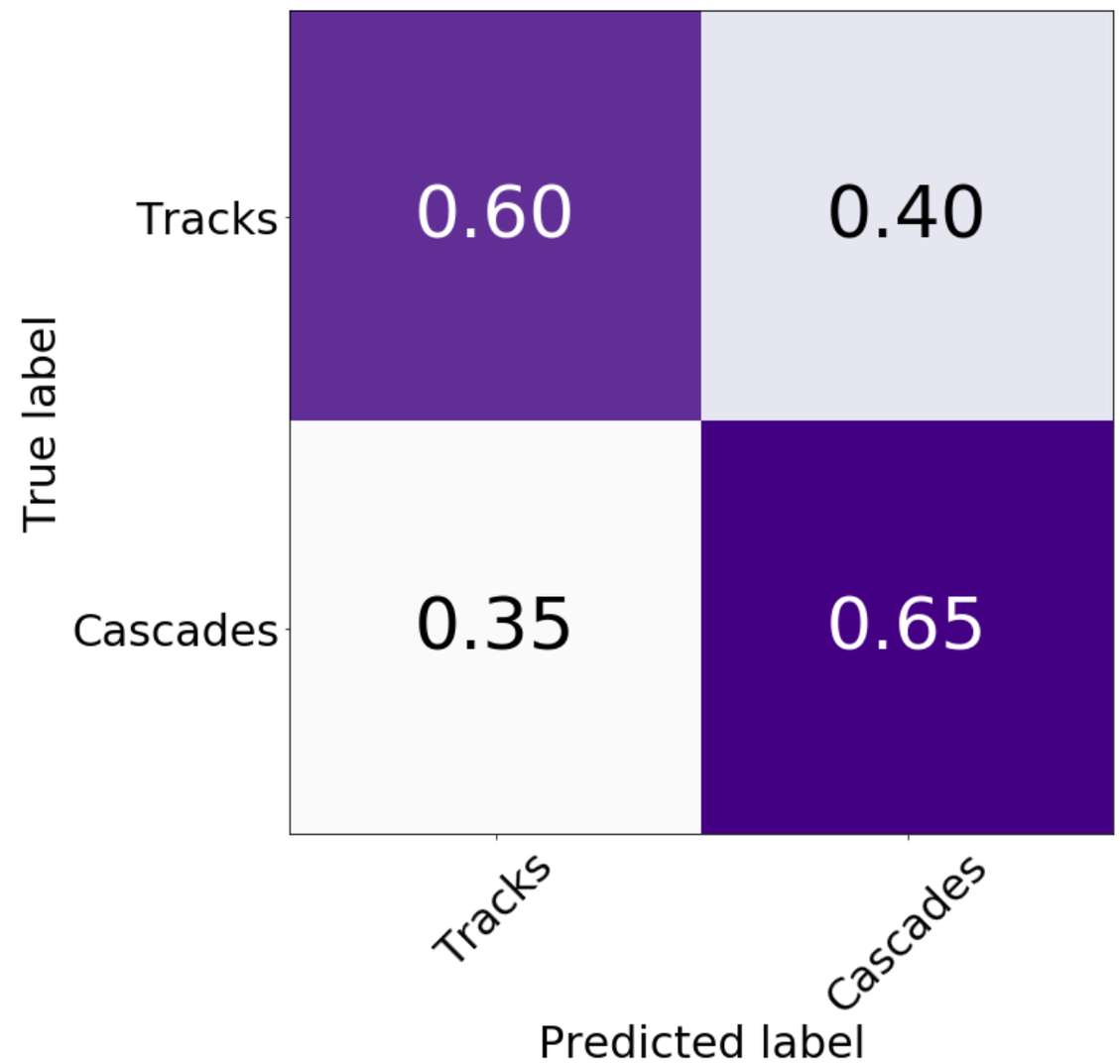
- Used low-level files (not much pre-processing and no reconstruction)
- Training: ~ 5 million events
  Validation: ~ 584k events
  Test: ~ 233k events
- Tracks: ~ 3.5 million
  Cascades: ~ 2.4 million
- Events are weighted according to the corresponding atmospheric weights
- The performance of the 2D CNN classifier proves to be very similar to that of the BDT classifier
- Perhaps the inaccurate z-axis spacing or the sparse nature of the events limits the CNN from finding more useful symmetries
- Perhaps more training data is needed



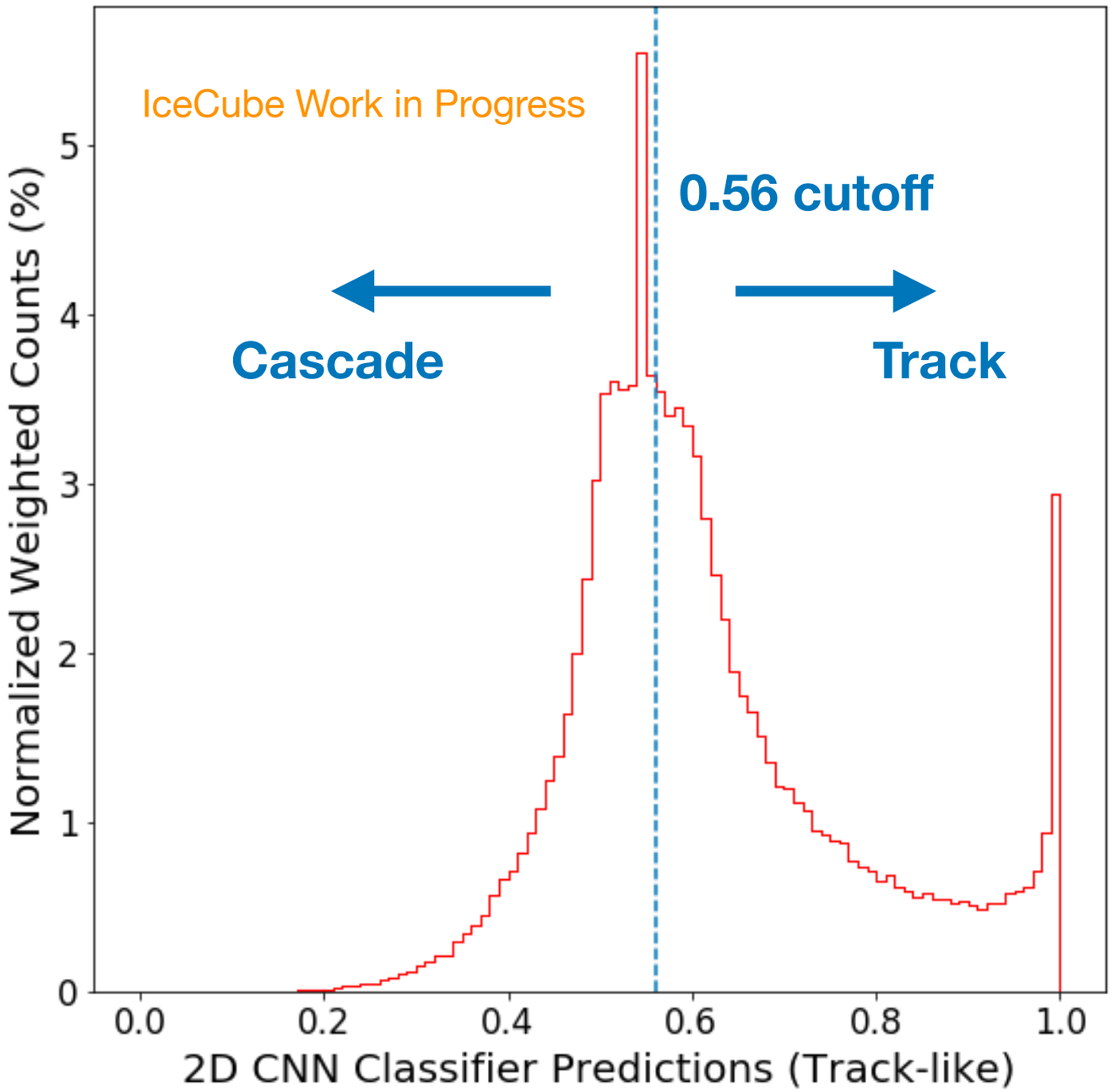
2D CNN Classifier:
track cutoff = 0.56



BDT Classifier:
track cutoff = 0.45



2D CNN Classifier: Prediction Distribution

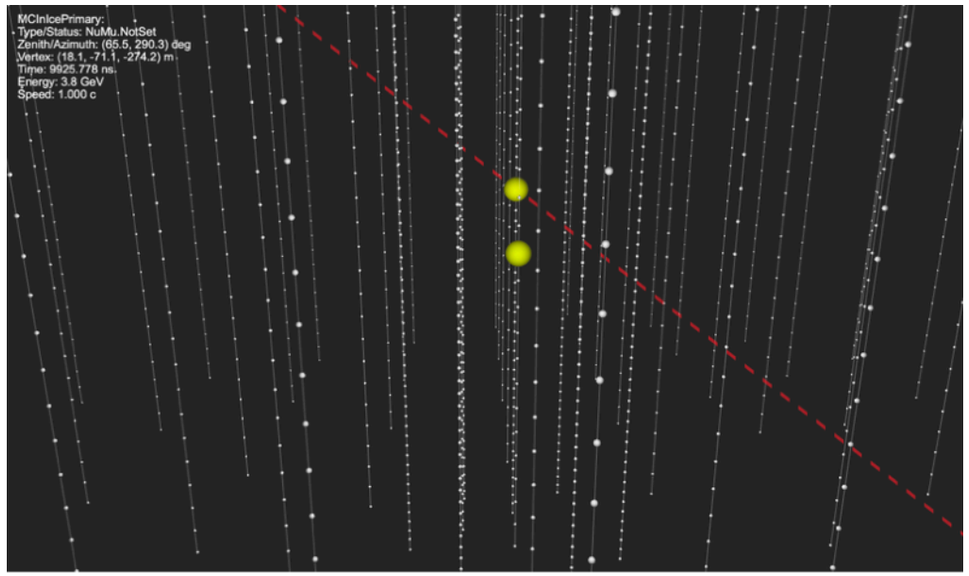


- Prediction distribution for the test sample
- Expected distribution with a peak at 1.0 (track) and no peak at 0 (cascade)
- The network could never be 100% sure of a cascade
- $\nu_\mu CC$ creates a hadronic shower along with a muon
- "Tracks" are technically lollipop-shaped
- A very short track will look just like a cascade
- A very long track will never look like a cascade
- The peak at the center are the events the network could not distinguish at all

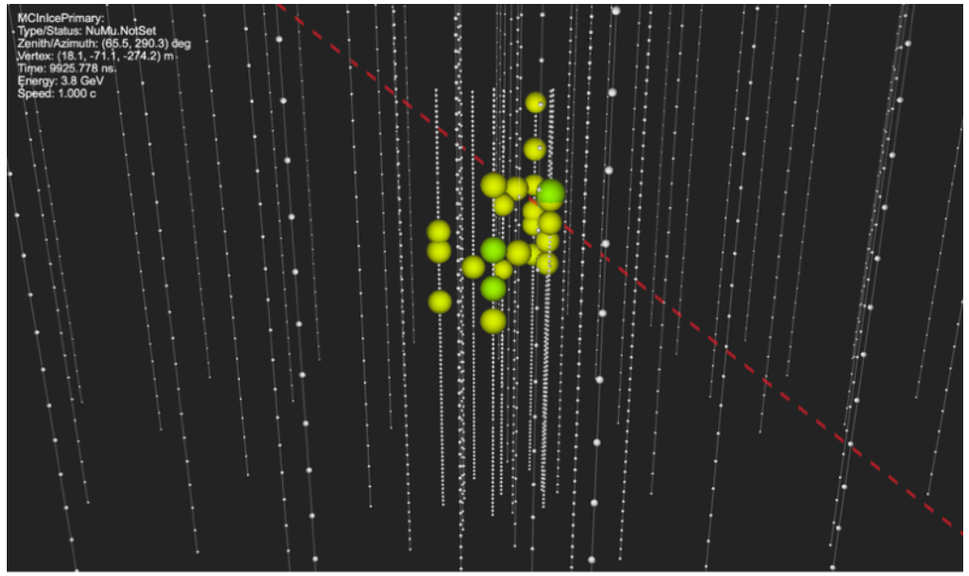
Future Improvements

- More accurate z-axis spacing of DeepCore vs IceCube strings
- Perhaps Cumulative charge
- More training data
- Sparse convolutions
- Graph Neural Networks (already in the works)
- IceCube Upgrade: seven more strings with even shorter DOM spacing
- Will help us probe lower energy events (1-10 GeV) by providing more information channels
- Could provide better separation power between tracks and cascades

DeepCore 3.8 GeV event



Upgrade 3.8 GeV event



References

[1] & [2] IceCube Collaboration, Phys. Rev. D 99, 032007 (2019), arXiv:1901.05366

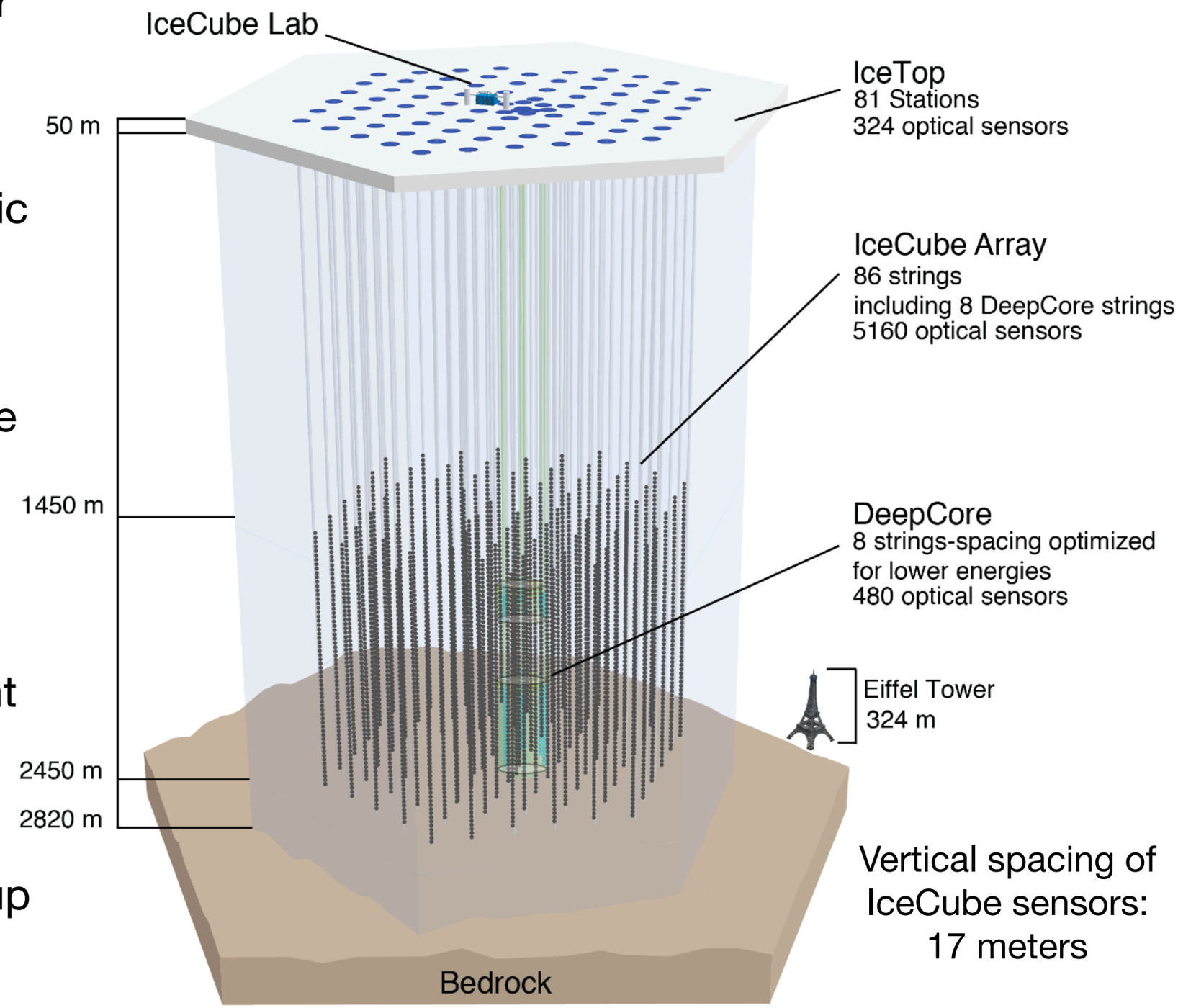
[3] https://cs231n.github.io/convolutional-networks/

Thank you!

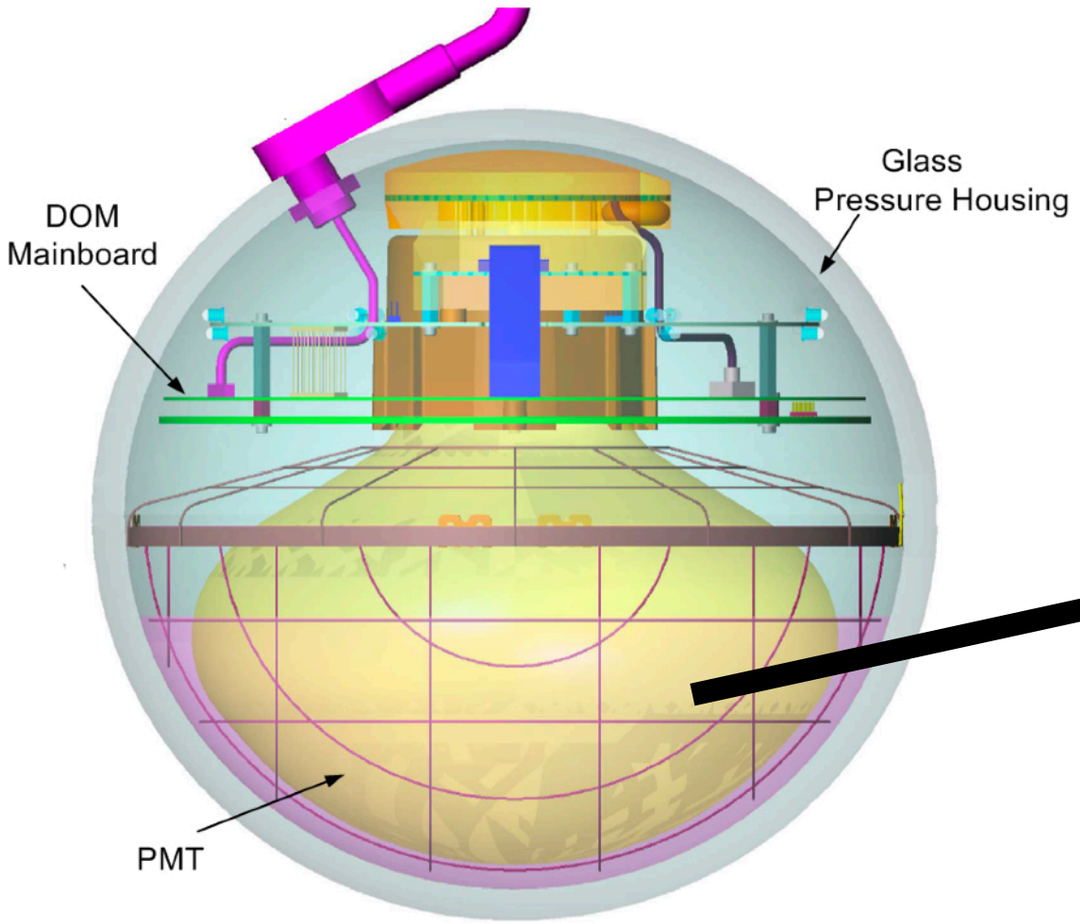
Back-up Slides

The IceCube Neutrino Observatory

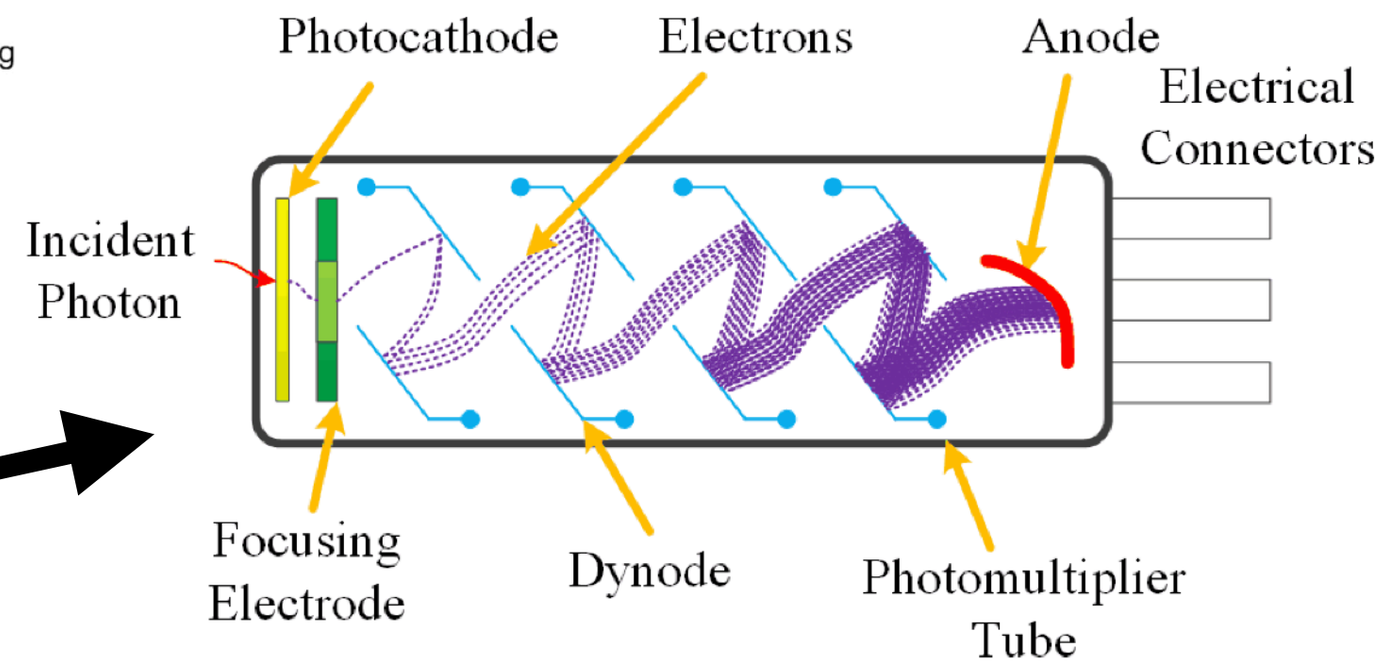
- IceCube is a neutrino detector located at the South Pole
- It consists of a lattice of light sensors embedded in Antarctic ice ~1.5 kilometers below the surface
- Neutrinos can interact with the atoms in the ice and produce detectable charged particles that leave behind a light trace
- Our digital optical module light sensors pick up this light, convert it into a current and transfer the digitized current up the cable to the lab



Digital Optical Modules (DOMs)



Digital Optical Module (DOM)

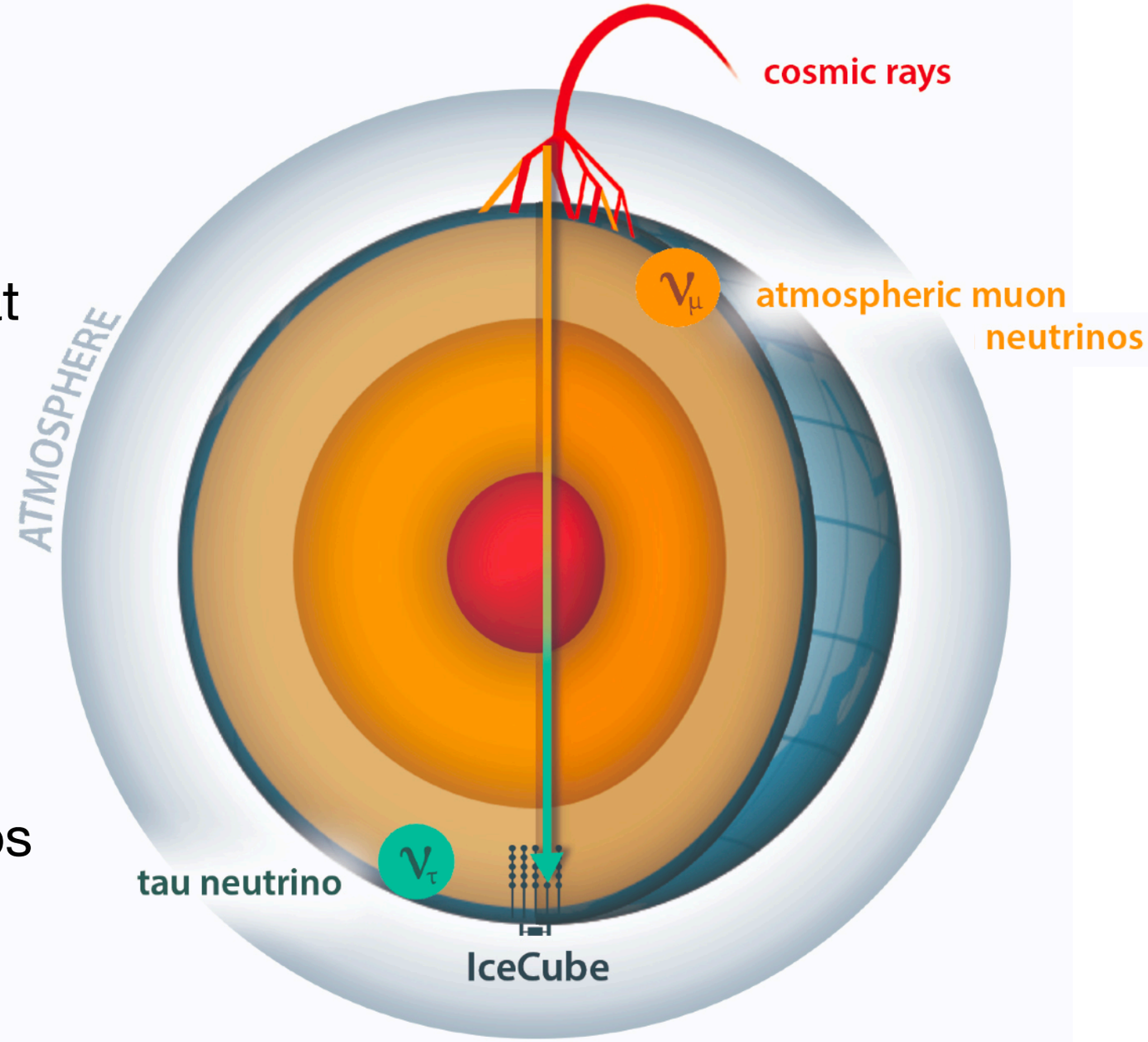


Photomultiplier Tube (PMT)

- DOMs are made up mainly of a PMT and the readout electronics
- PMTs detect photons through the photoelectric effect, amplifying the photoelectrons knocked off of the photocathode through a dynode chain
- By the time the photoelectrons reach the anode, the amplification is big enough to create a readable current

Cosmic Rays

- High-energy protons and nuclei that collide with the Earth's atmosphere
- Produce a shower of particles
- Charged pions and kaons from this shower decay into muons and neutrinos
- "Upward-going" direction: Neutrinos will travel through the Earth
- Some will interact with the ice in DeepCore

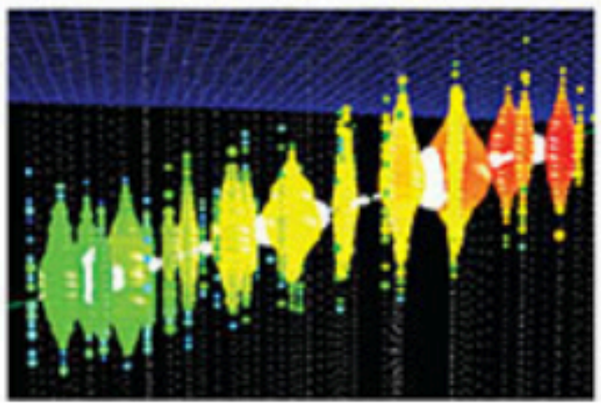


$$p^+ + N \rightarrow \pi^+ + X$$
$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$
$$\mu^+ \rightarrow e^+ + \bar{\nu_\mu} + \nu_e$$

In 2014, we saw about **44,465** neutrino events in DeepCore with an expected ratio of about **5 : 15 : 1** for $\nu_e : \nu_\mu : \nu_\tau$.

IceCube Events

Tracks

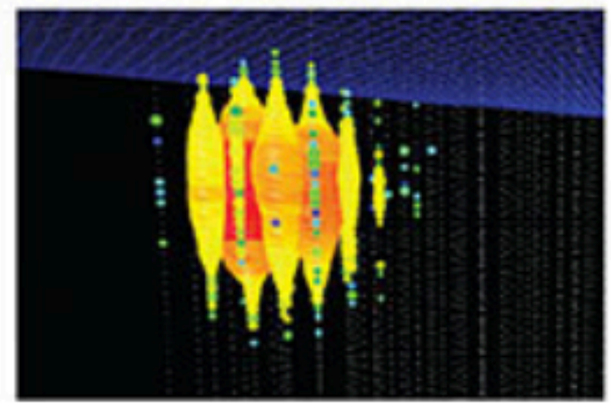


$$\nu_\mu + N \rightarrow \boxed{\mu^- + X}$$

Hadronic shower plus Muon track

Muons travel long distances before they decay

Cascades



$$\nu_x + N \rightarrow \nu_x + \boxed{X}$$

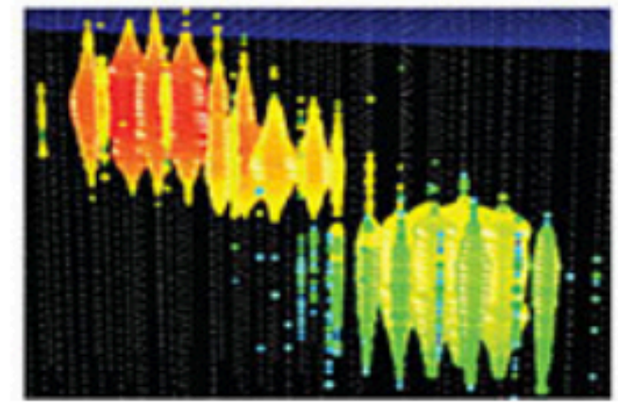
Hadronic shower

$$\nu_e + N \rightarrow \boxed{e^- + X}$$

Hadronic shower plus Electromagnetic shower

Electrons scatter

Double-Bangs

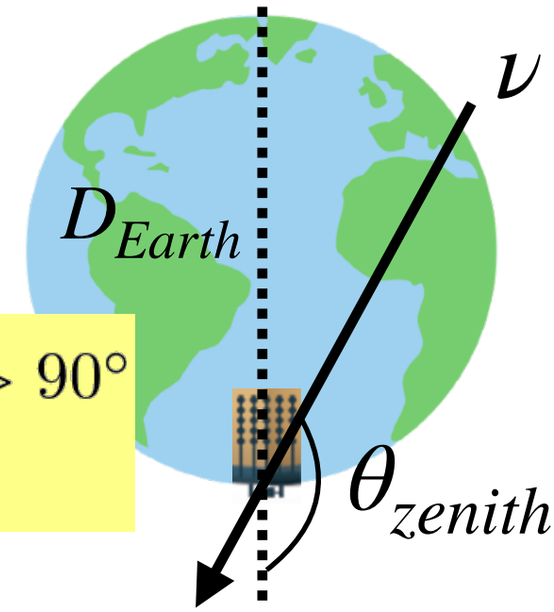


$$\nu_\tau + N \rightarrow \tau^- + \boxed{X}$$
$$\tau^- \rightarrow \nu_\tau + \boxed{X'}$$

Two Hadronic showers

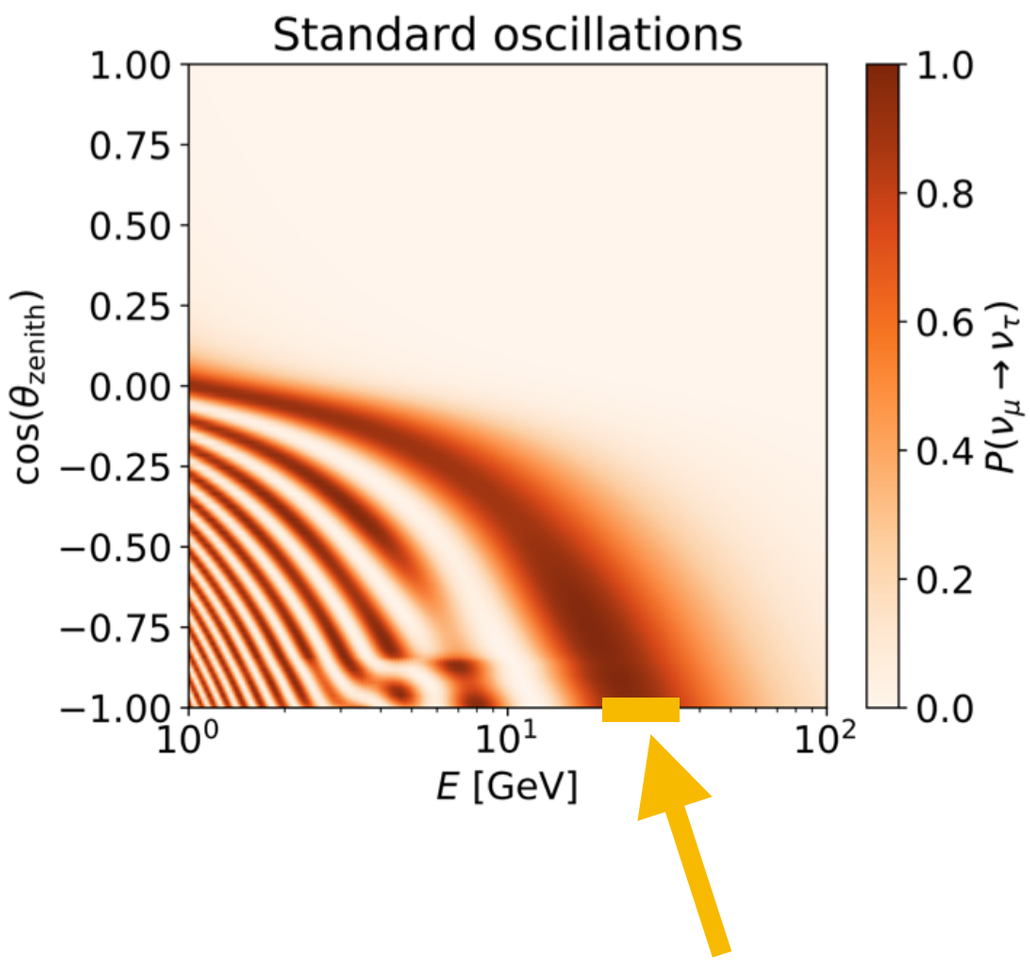
Taus decay very quickly

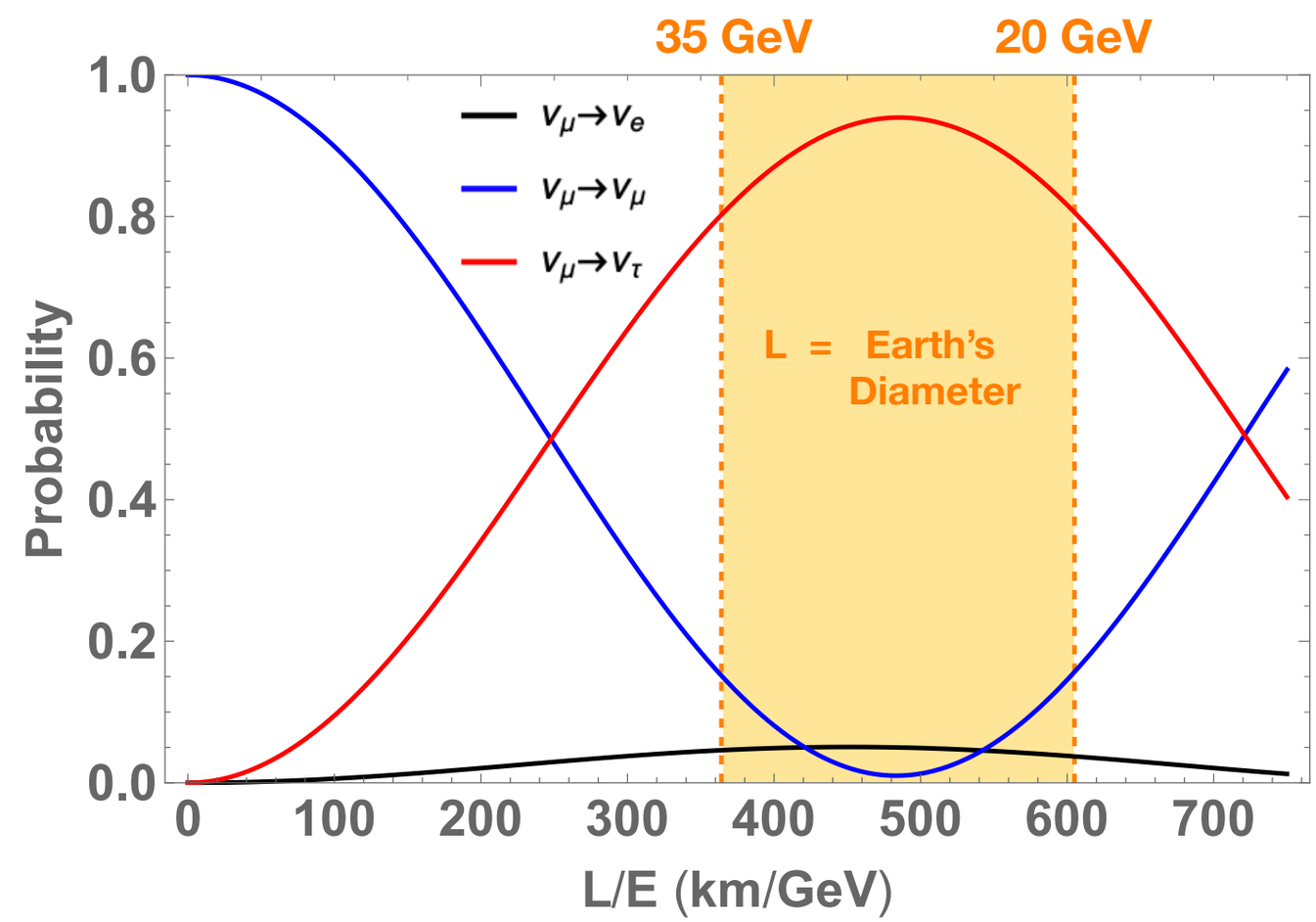
Tau Neutrino Appearance in DeepCore



$$P_{\nu_\mu \to \nu_\tau} \approx \cos^4(\theta_{13}) \sin^2(2\theta_{23}) \sin^2\left(\frac{\Delta m^2_{32} L}{4E_\nu}\right)$$

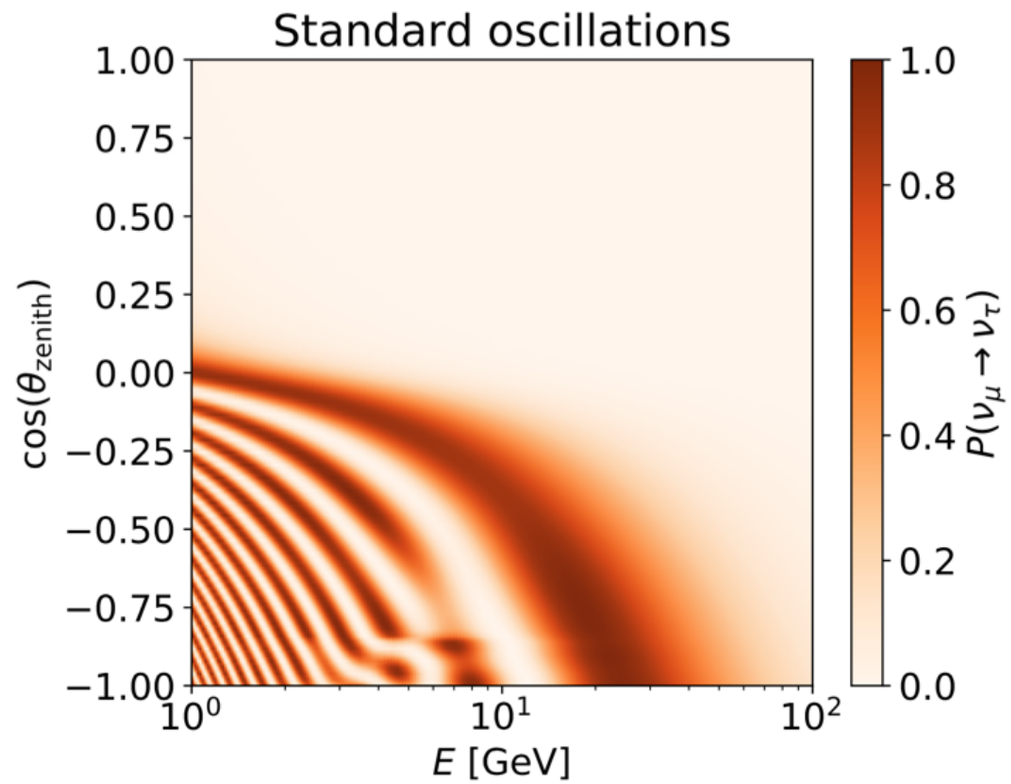
$$L \approx -D_{Earth}\cos(\theta), \quad \theta > 90^\circ$$
$$L \approx 0, \quad \text{otherwise}$$

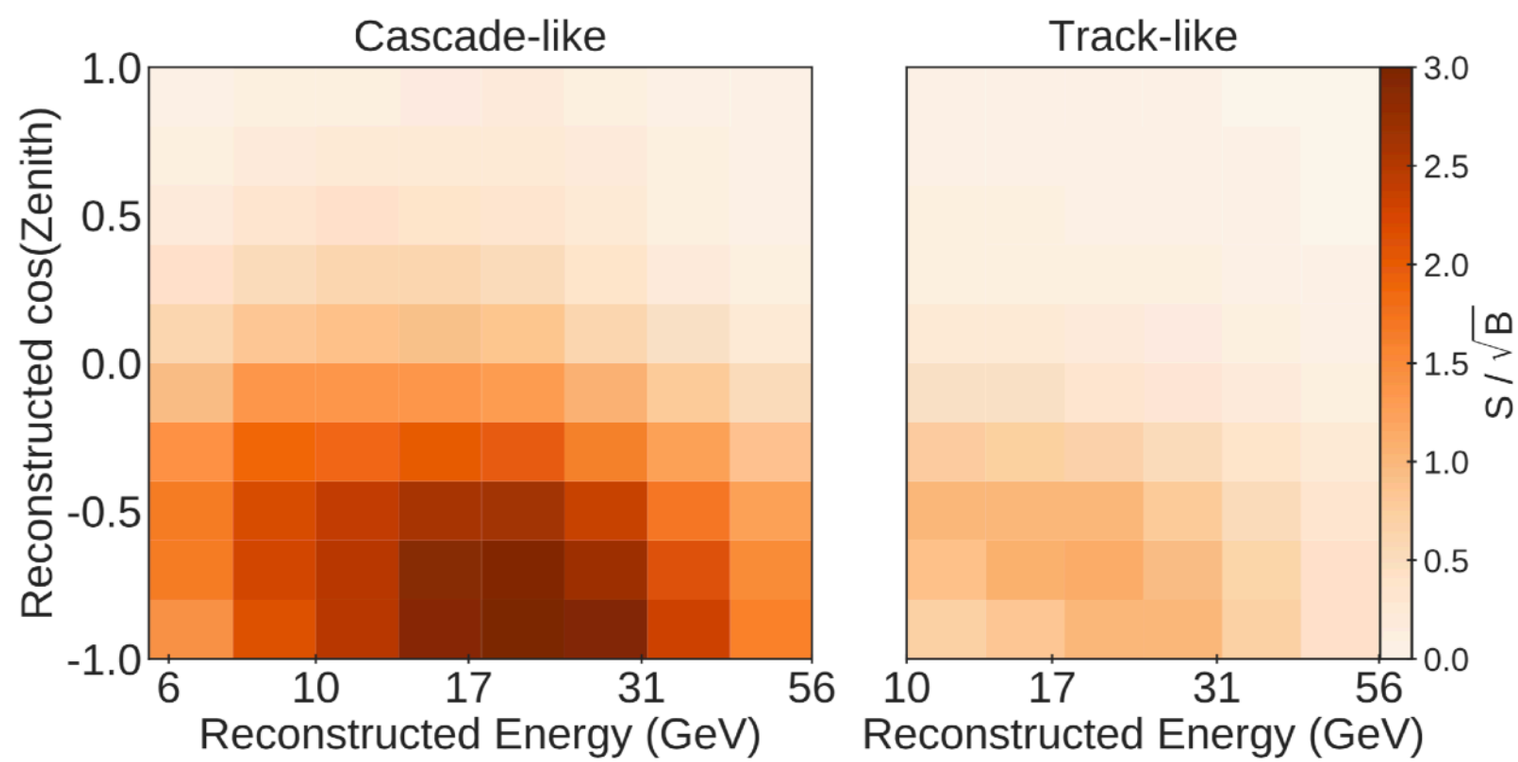




Tau Neutrino Appearance in DeepCore

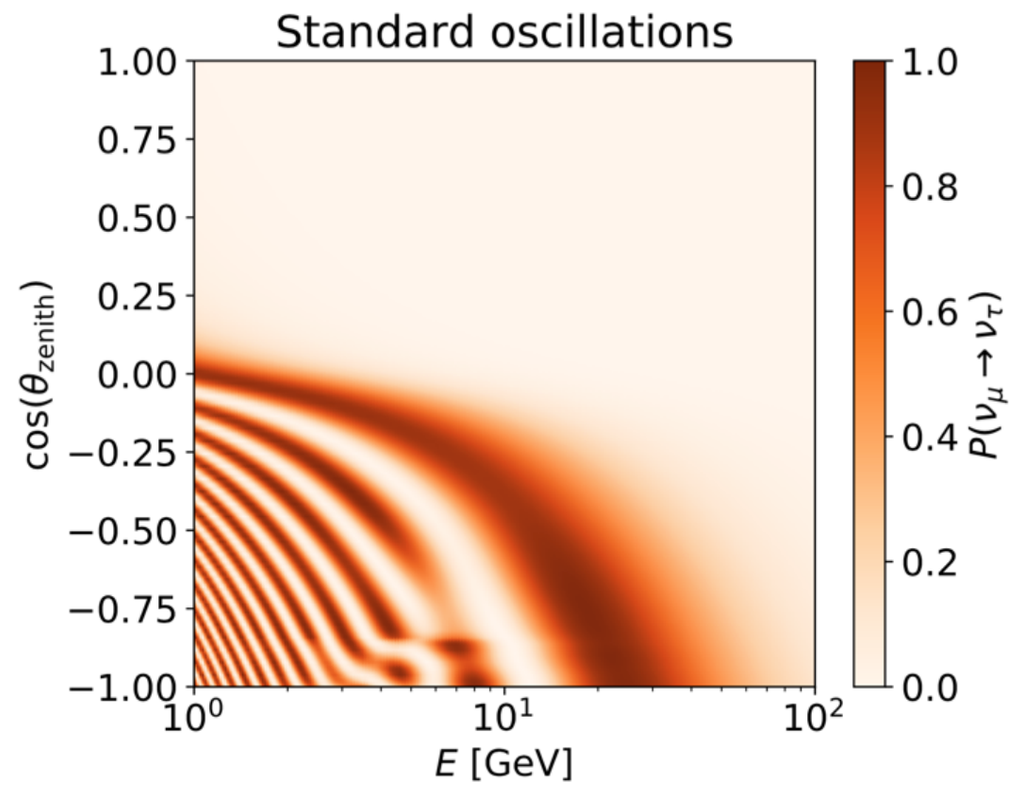
- Optimized to look for an excess of tau neutrinos at energy ranges of 5 - 56 GeV for “upward-going” neutrinos
- Events are sorted into two categories: “Track-like” and “Cascade-like”
- Expected number of tau neutrino events over the uncertainty in the background events shows how likely we are to see a tau neutrino signal per bin

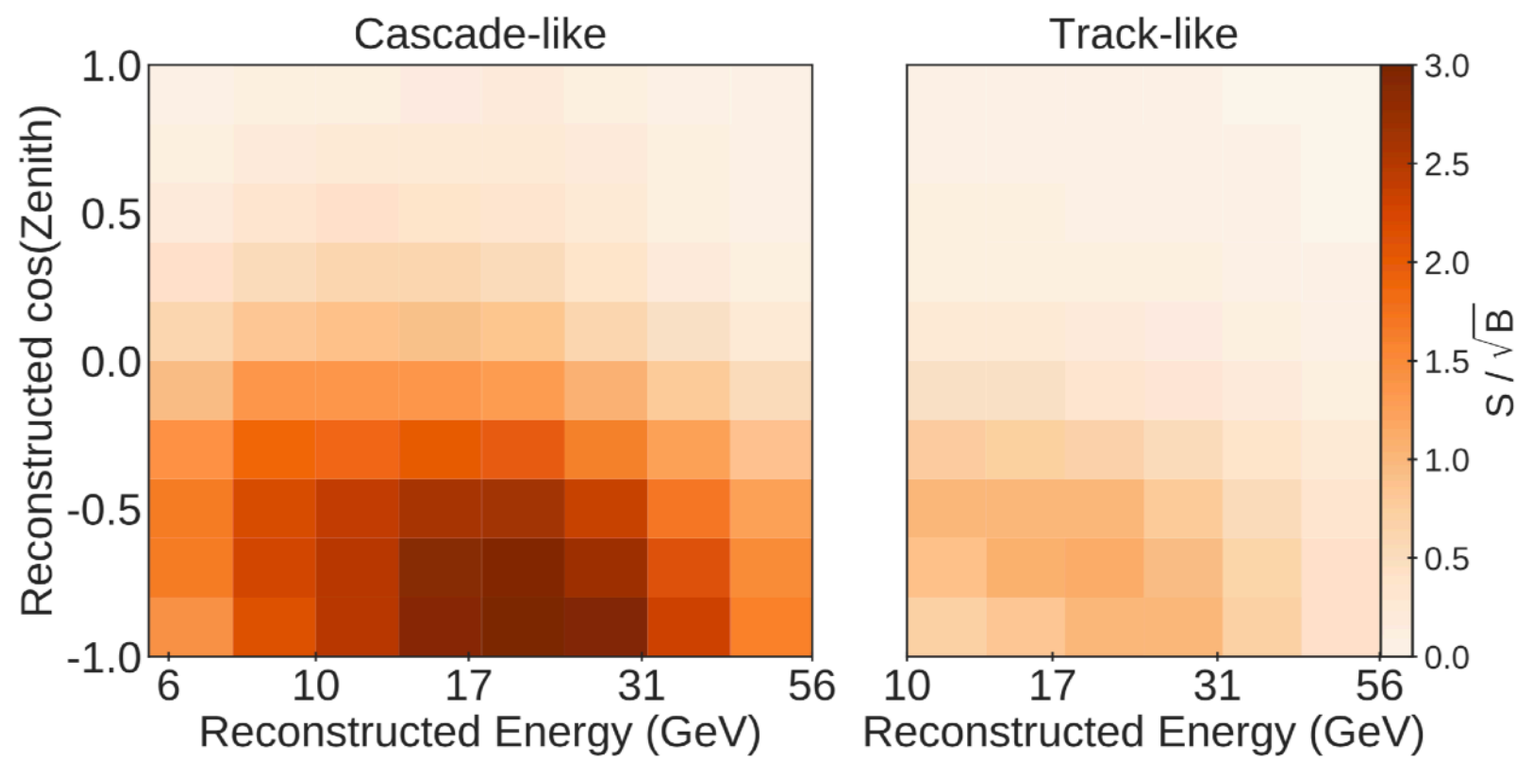




Nu Tau Appearance

- Although most tau neutrino events will have a "Cascade-like" signature, in about 17% of the tau neutrino CC interactions, the tau will decay into a muon and muon neutrino, so this is why some "Track-like" bins appear non-zero

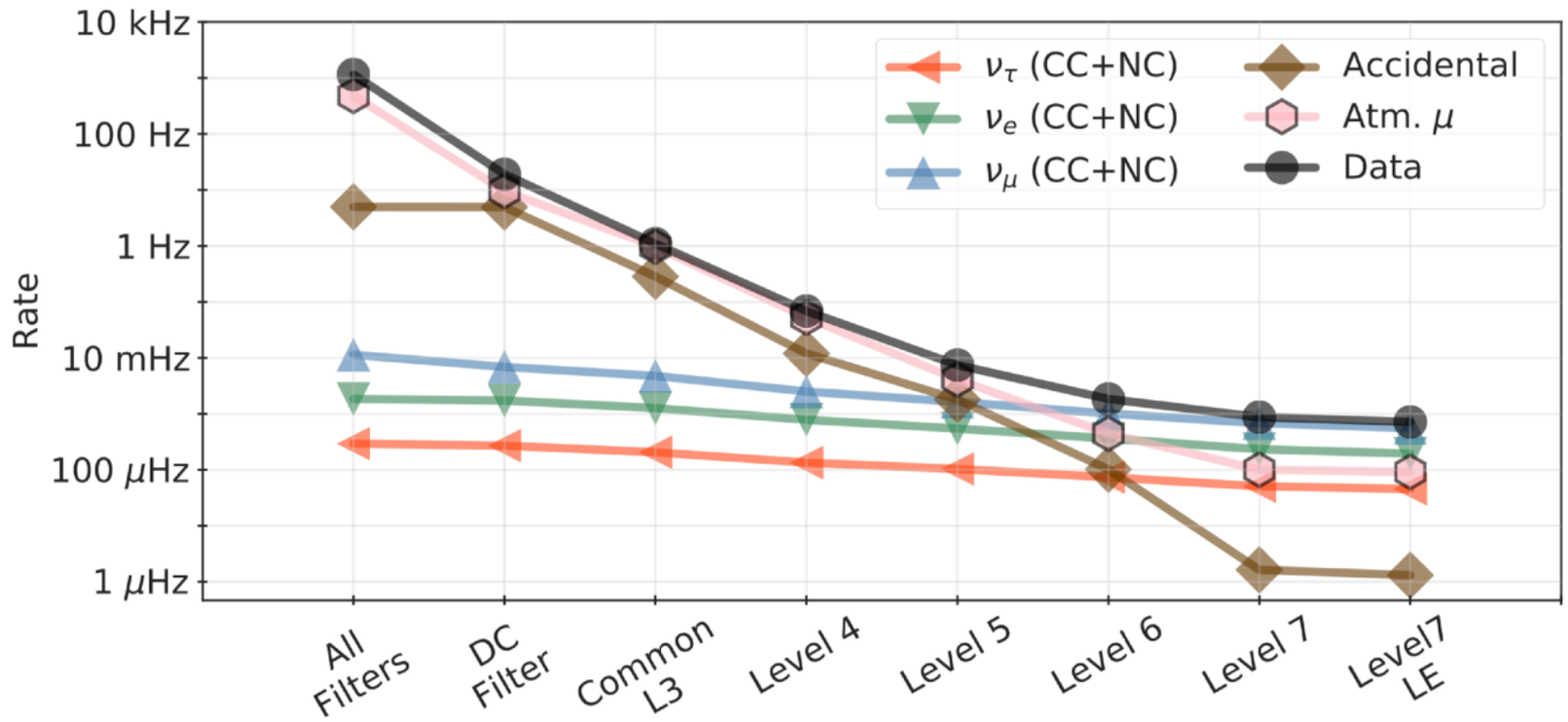




Tau Neutrino Appearance

Parameter	Prior	Analysis $\mathcal{A}$ Best fit (CC+NC)	Analysis $\mathcal{A}$ Best fit (CC)	Analysis $\mathcal{B}$ Best fit (CC+NC)	Analysis $\mathcal{B}$ Best fit (CC)
Neutrino Flux & Cross Section:					
ν_e/ν_μ Ratio	1.0 ± 0.05	1.03	1.03	1.03	1.03
ν_e Up/Hor. Flux Ratio (σ)	0.0 ± 1.0	-0.19	-0.18	-0.25	-0.24
$\nu/\bar{\nu}$ Ratio (σ)	0.0 ± 1.0	-0.42	-0.33	0.01	0.04
$\Delta\gamma_\nu$ (Spectral Index)	0.0 ± 0.1	0.03	0.03	-0.05	-0.04
Effective Livetime (years)	-	2.21	2.24	2.45	2.46
M_A^{CCQE} (Quasi-Elastic) (GeV)	$0.99^{+0.248}_{-0.149}$	1.05	1.05	0.88	0.88
M_A^{res} (Resonance) (GeV)	1.12 ± 0.22	1.00	0.99	0.85	0.85
NC Normalization	1.0 ± 0.2	1.05	1.06	1.25	1.26
Oscillation:					
θ_{13} (°)	8.5 ± 0.21	-	-	8.5	8.5
θ_{23} (°)	-	49.8	50.2	46.1	45.9
Δm^2_{32} (10^{-3}eV^2)	-	2.53	2.56	2.38	2.34
Detector:					
Optical Eff., Overall (%)	100 ± 10	98.4	98.4	105	104
Optical Eff., Lateral (σ)	0.0 ± 1.0	0.49	0.48	-0.25	-0.27
Optical Eff., Head-on (a.u.)	-	-0.63	-0.64	-1.15	-1.22
Local Ice Model	-	-	-	0.02	0.07
Bulk Ice, Scattering (%)	100.0 ± 10	103.0	102.8	97.4	97.3
Bulk Ice, Absorption (%)	100.0 ± 10	101.5	101.7	102.1	101.9
Atmospheric Muons:					
Atm. μ Fraction (%)	-	8.1	8.0	4.6	4.6
$\Delta\gamma_\mu$ (μ Spectral Index, σ)	0.0 ± 1.0	0.15	0.15	-	-
Coincident $\nu + \mu$ Fraction	$0.0 + 0.1$	0.01	0.01	-	-
Measurement:					
ν_τ Normalization	-	0.73	0.57	0.59	0.43

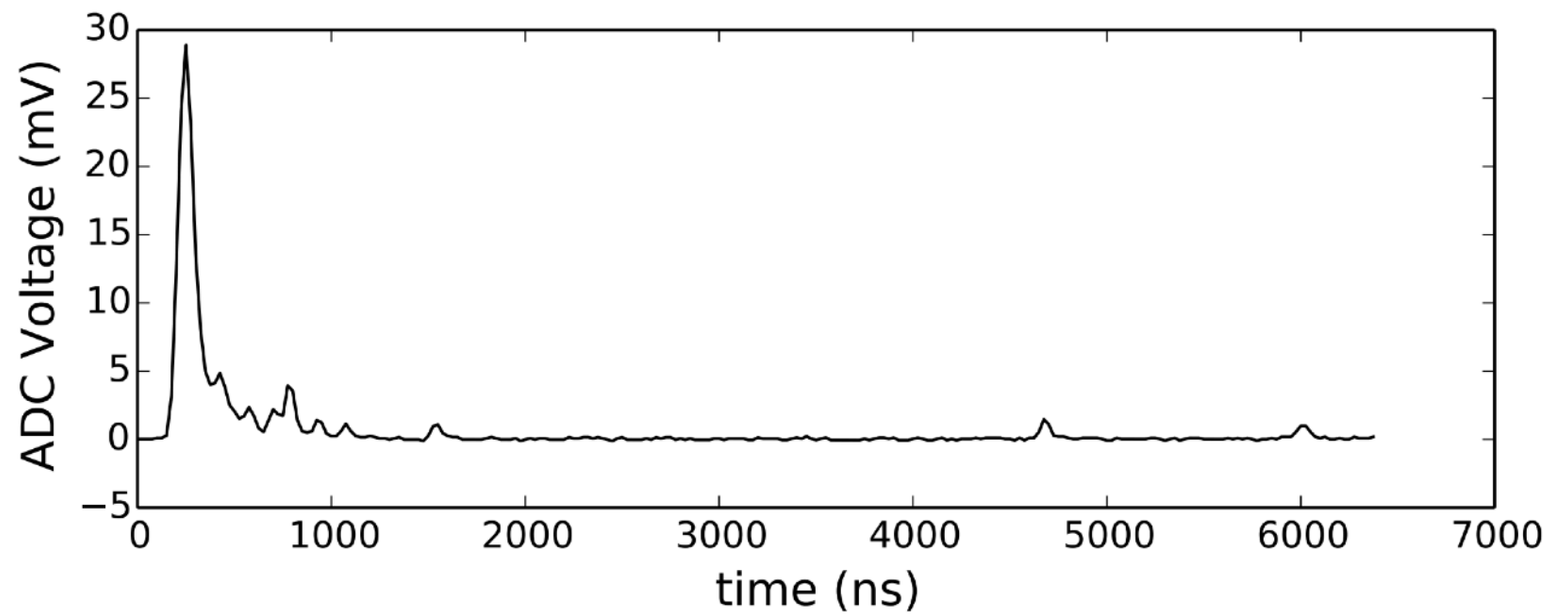
Background Rates



Type	Analysis $\mathcal{A}$ Events	$\pm 1\sigma$	Analysis $\mathcal{B}$ Events	$\pm 1\sigma$
$\nu_e + \bar{\nu}_e$ CC	13462	29	9545	23
$\nu_e + \bar{\nu}_e$ NC	1096	9	923	8
$\nu_\mu + \bar{\nu}_\mu$ CC	35706	48	23852	39
$\nu_\mu + \bar{\nu}_\mu$ NC	4463	19	3368	17
$\nu_\tau + \bar{\nu}_\tau$ CC	1804	9	934	5
$\nu_\tau + \bar{\nu}_\tau$ NC	556	3	445	4
Atmospheric μ	5022	167	1889	45
Noise Triggers	93	27	< 25	< 5
total (best fit)	**62203**	**180**	**40959**	**68**
observed	**62112**	**249**	**40902**	**202**

DOM PE Charge

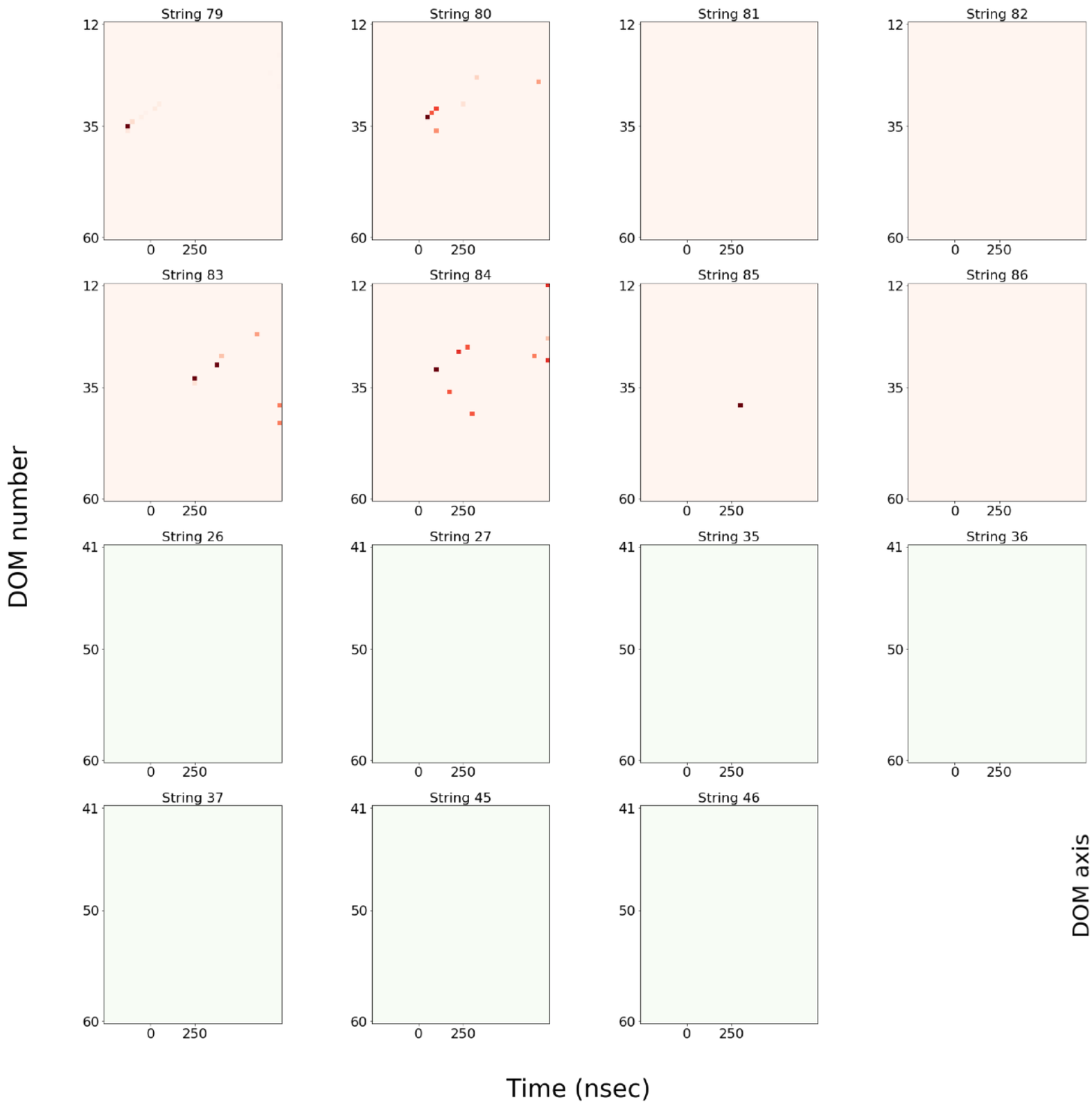
- DOM records a shower of PEs per photoelectron that gets knocked off the photocathode
- Records pulses as a function of time
- A discriminator threshold is set to reject noise and accept signals
- A fit for one PE (SPE) is set and all other pulses are fit compared to that
- So if a pulse was 3/4 of that, the PE charge would read “0.75 PE charge”
- The discriminator threshold is set to accept any pulse above 0.25

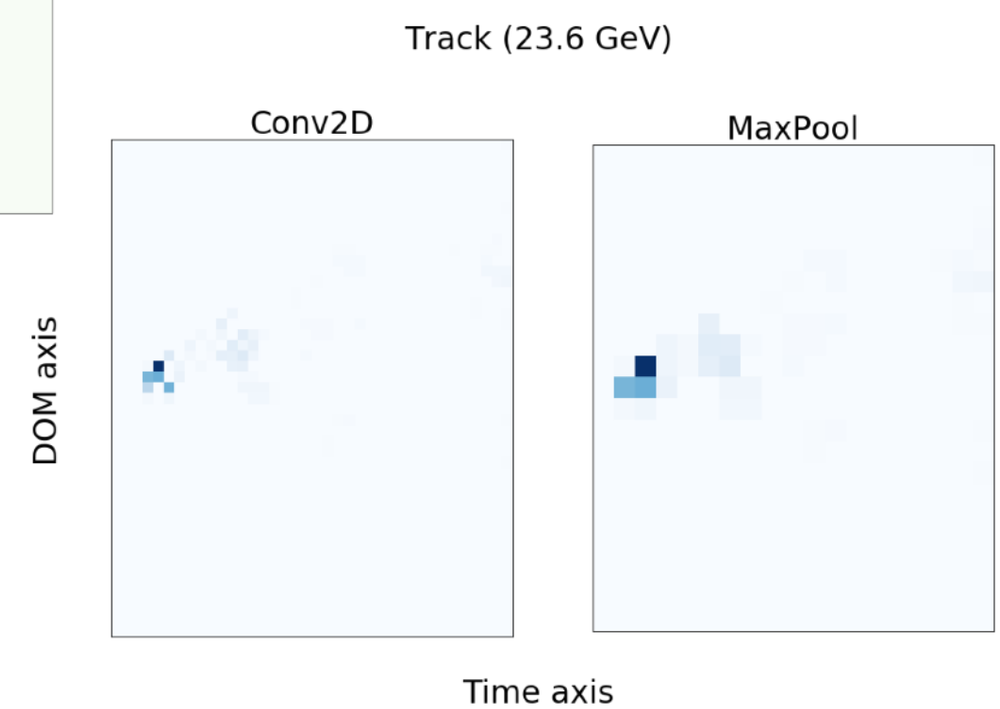


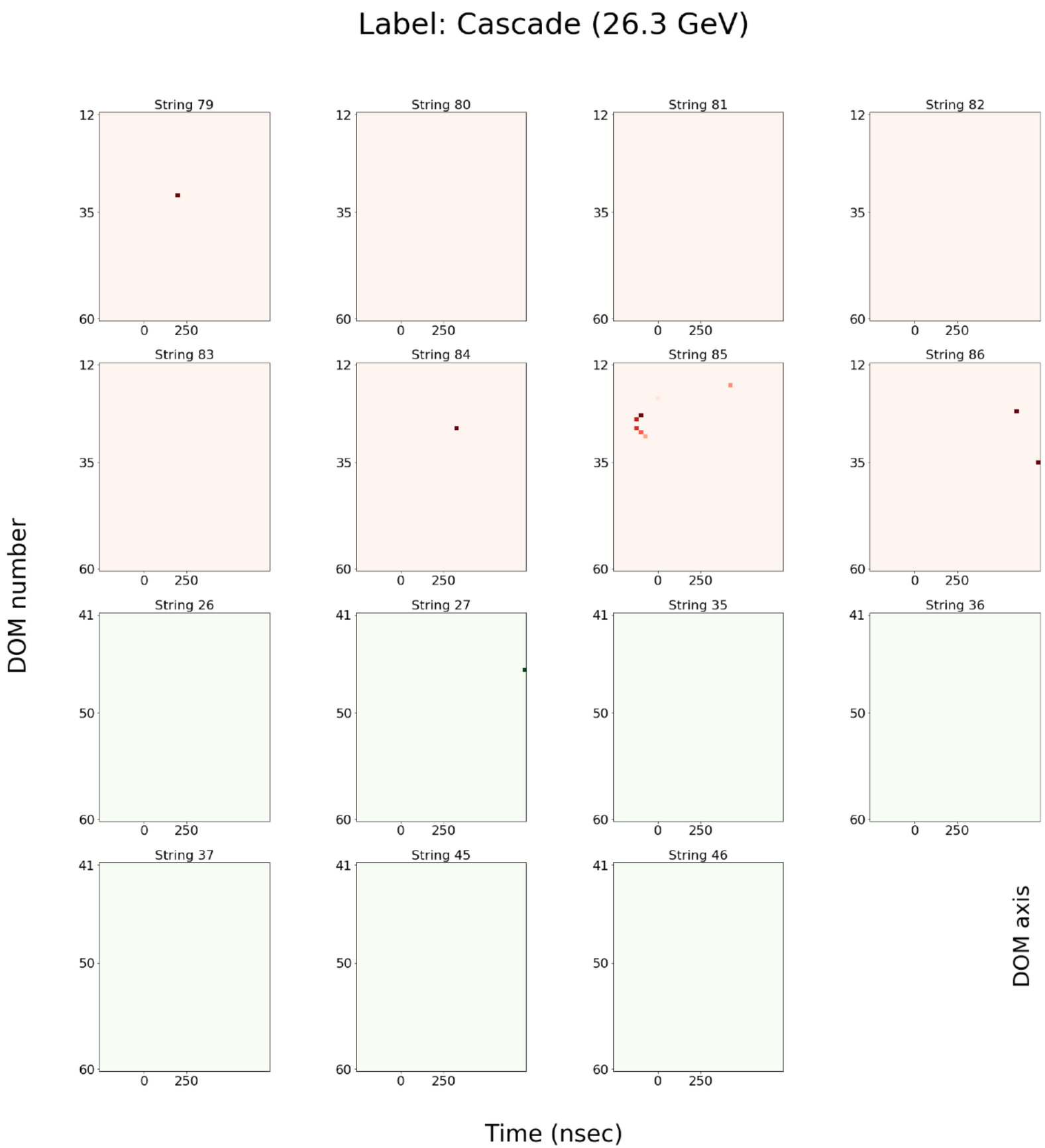
XGBoost BDT Feature Importance

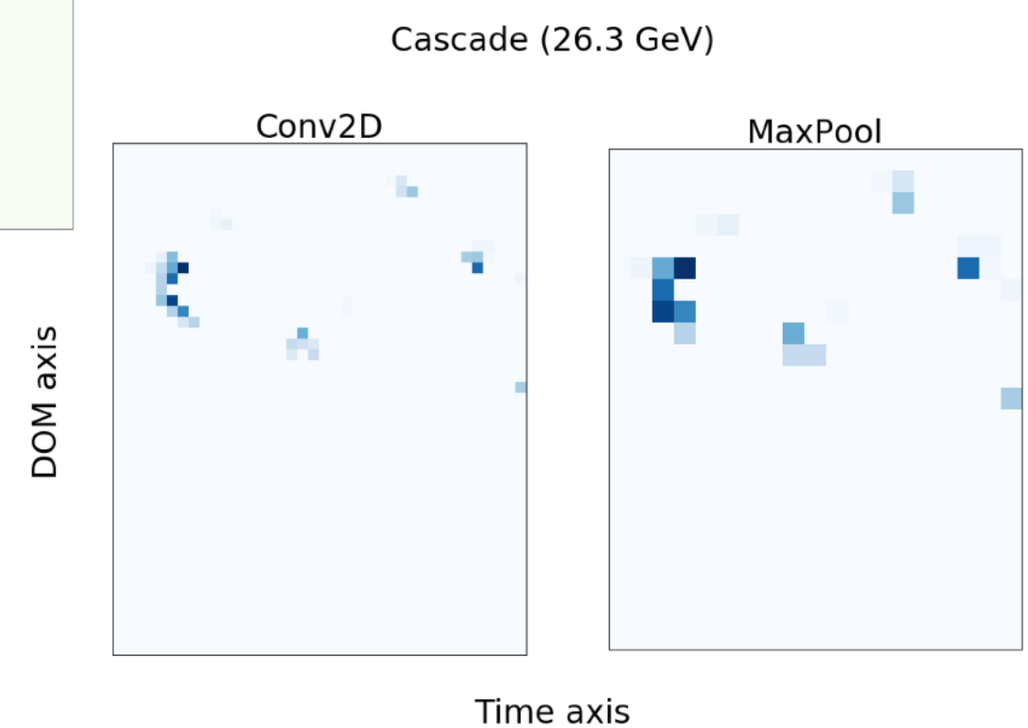
Feature name	Feature importance
reconstructed cascade energy	0.081
reconstructed track length	0.657
reconstructed $\cos(\theta_{zenith})$	0.068
reconstructed $\cos(\theta_{zenith})$ uncertainty	0.099
LLH ratio of cascade+track reco to shower only reco	0.096

Label: Track (23.6 GeV)






Label: Cascade (26.3 GeV)
String 79
String 80
String 81
String 82
String 83
String 84
String 85
String 86
String 26
String 27
String 35
String 36
String 37
String 45
String 46
DOM number
Time (nsec)


Cascade (26.3 GeV)
Conv2D
MaxPool
DOM axis
Time axis

IceCube Upgrade