9-12 September 2024
London, UK

PHYSTAT

Statistics meets Machine Learning

IMPERIAL

Identifying Tau Neutrinos in IceCube

PhyStat London 2024

Philipp Eller (TU Munich / IceCube)
philipp.eller@tum.de

*** special thanks to:**
Doug Cowen and team
Matthias Thiesmayer
Philipp Soldin
Philipp Fürst
Oliver Jannik

Part I: Physics Analysis

- Identifying Tau Neutrinos -

Part II: Interpretability

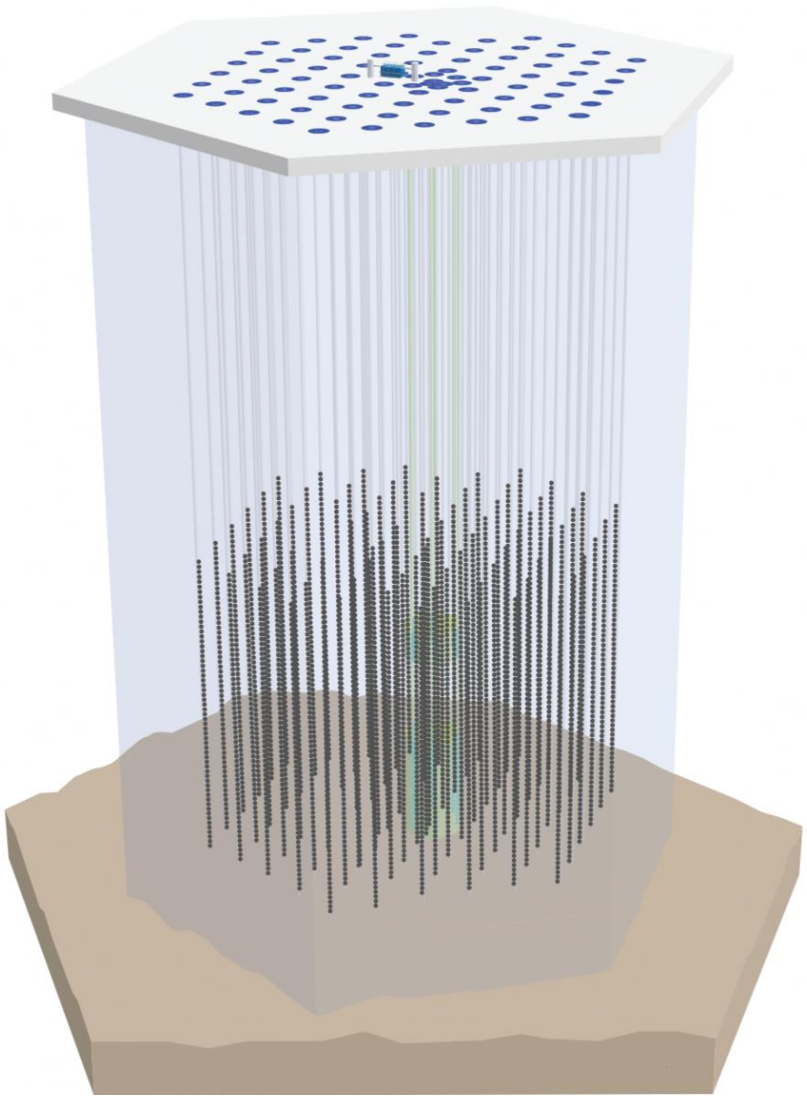
- Dissecting the Neural Net -

IceCube Neutrino Observatory

- Largest Neutrino Detector
  - Built into the South Pole ice
  - Instrumenting a volume of 1 km^3
  - 5160 sensors (DOMs) arranged in 86 strings

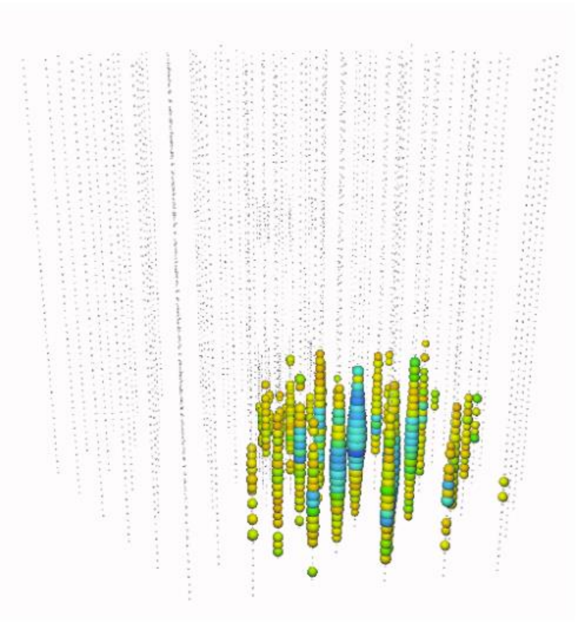
- Capable of detecting neutrino interactions from GeV to PeV energies

- Upgrade of detector underway with 7 new strings
  - advanced sensors
  - new calibration hardware
  - to be deployed 2025/2026



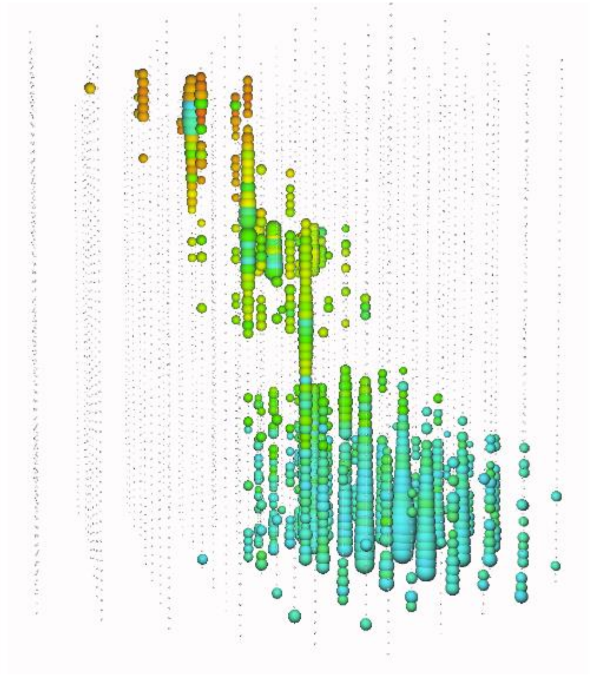
Event Types in IceCube

“Casacdes”
induced by ν_e or NC interactions



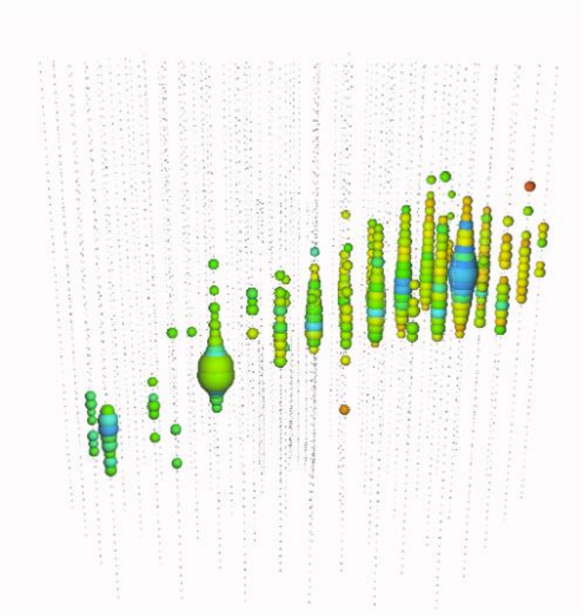
Common events

“Tracks”
induced by Atm. μ (or ν_μ)



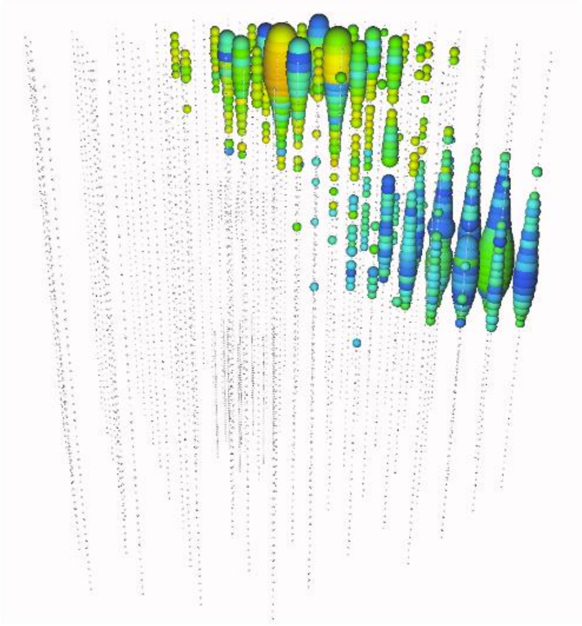
Ubiquitous events

“Starting Tracks”
induced by ν_μ



Common events

“Double Cascades”
induced by ν_τ



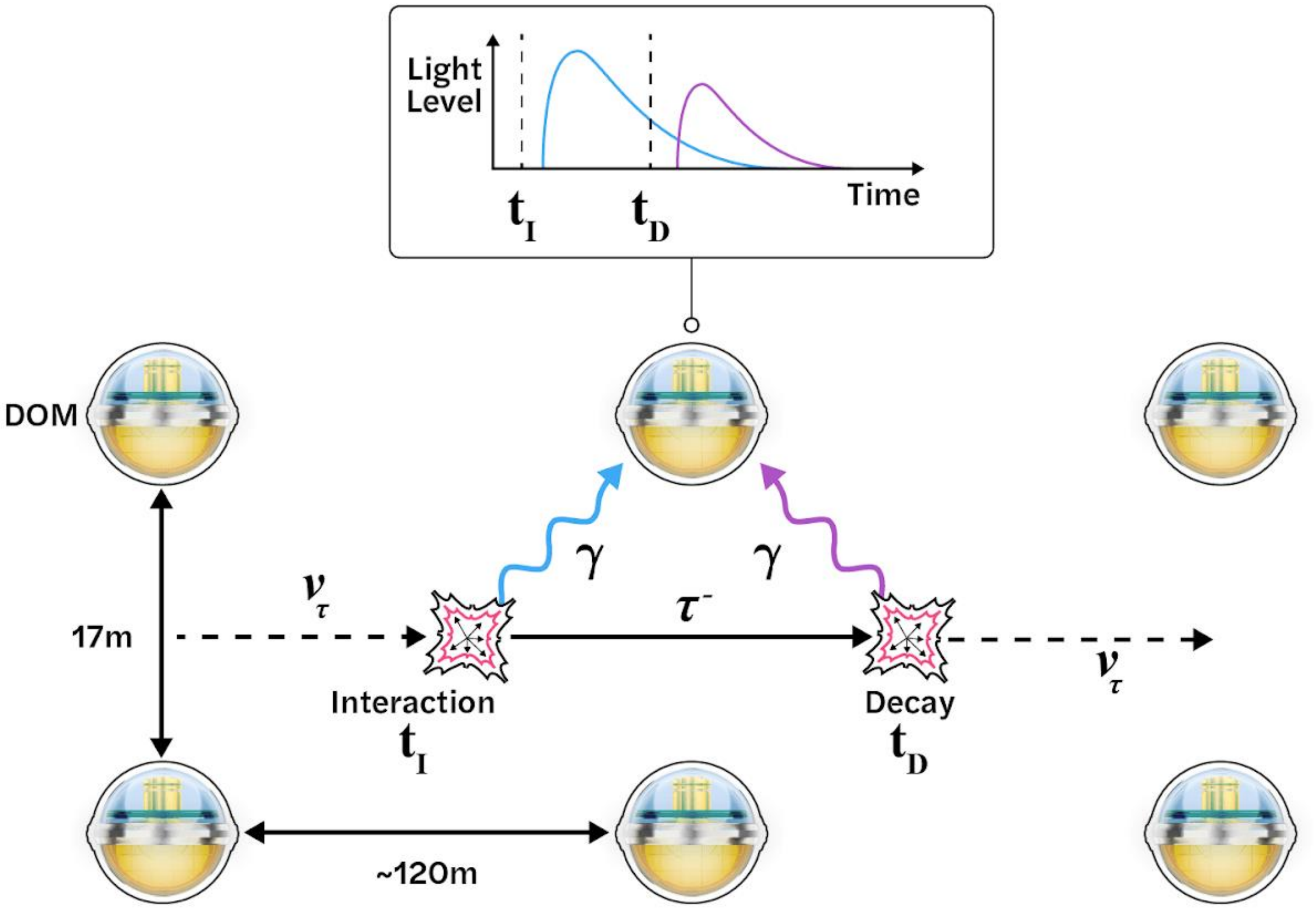
Very rare events

Tau Neutrinos

- A ν_τ **CC** interaction creates a τ lepton
  - Lifetime of 2.9×10^{-13} s
- If sufficiently energetic, it will travel several meters before decay (~50m / PeV)

→ Two separate vertices:
- First: Initial ν_τ **interaction**
- Later: τ **decay**

→ **Double pulse signature**

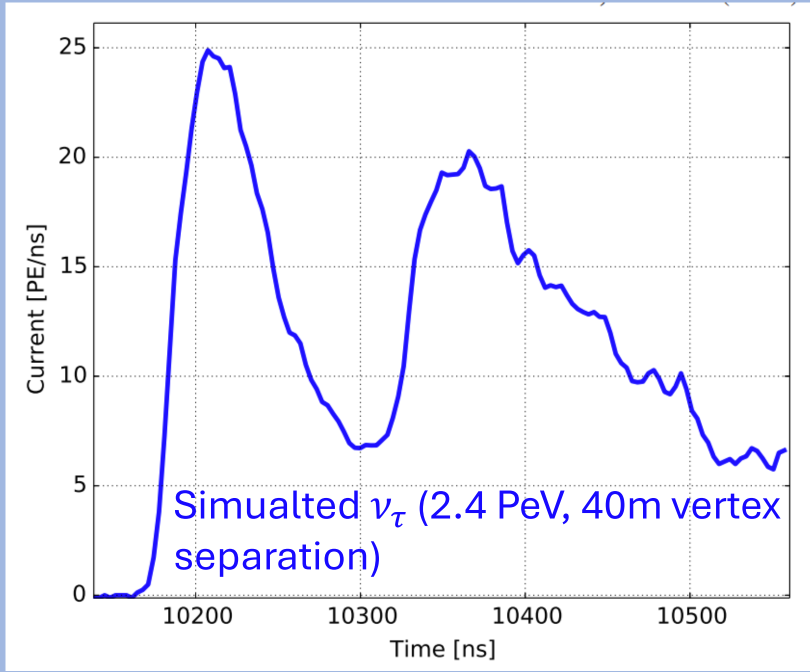


Previous Search Strategies

Double Peak in waveform

- <1 expected in 3 years of data
- 0 events observed

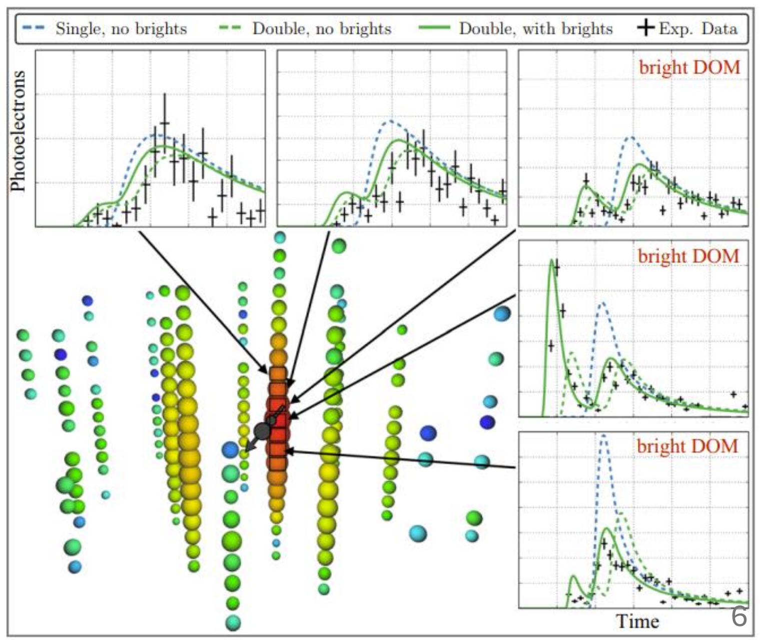
Phys. Rev. D 93, 022001



Double cascade reconstruction

- Using 60 high-energy-starting events (HESE)
- 1.5 signal and 0.8 background expected in 7.5 years of data
- 2 events observed → 2.8 sigma p-value

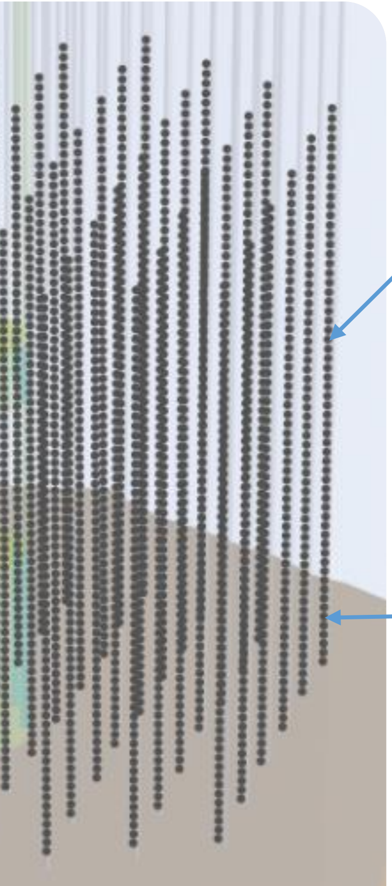
Eur. Phys. J. C 82, 1031 (2022)

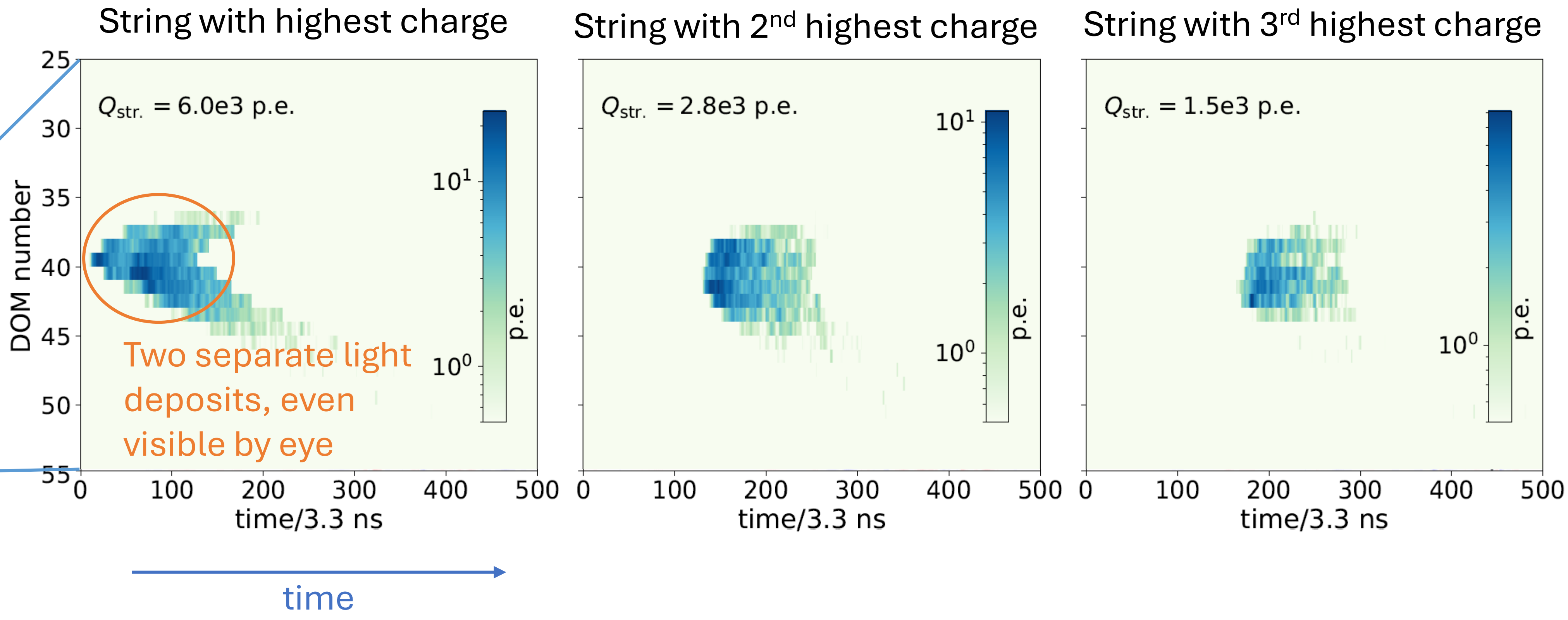


New ML method

Use waveforms of the three detector strings with the highest observed charge

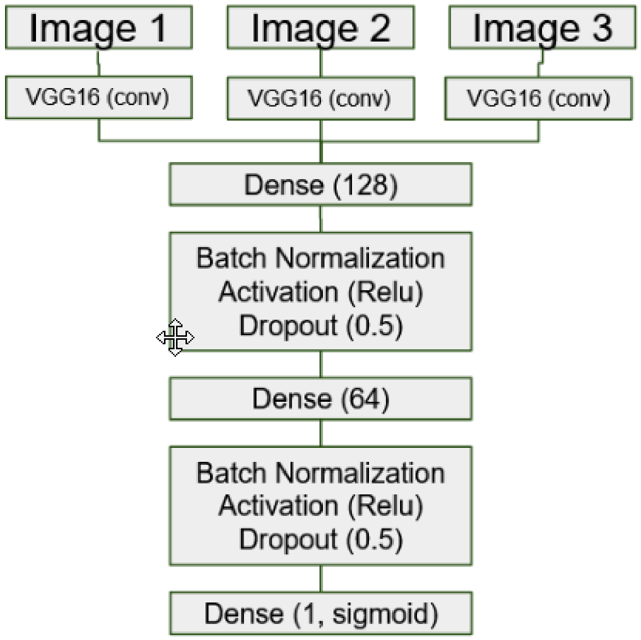
IceCube Array

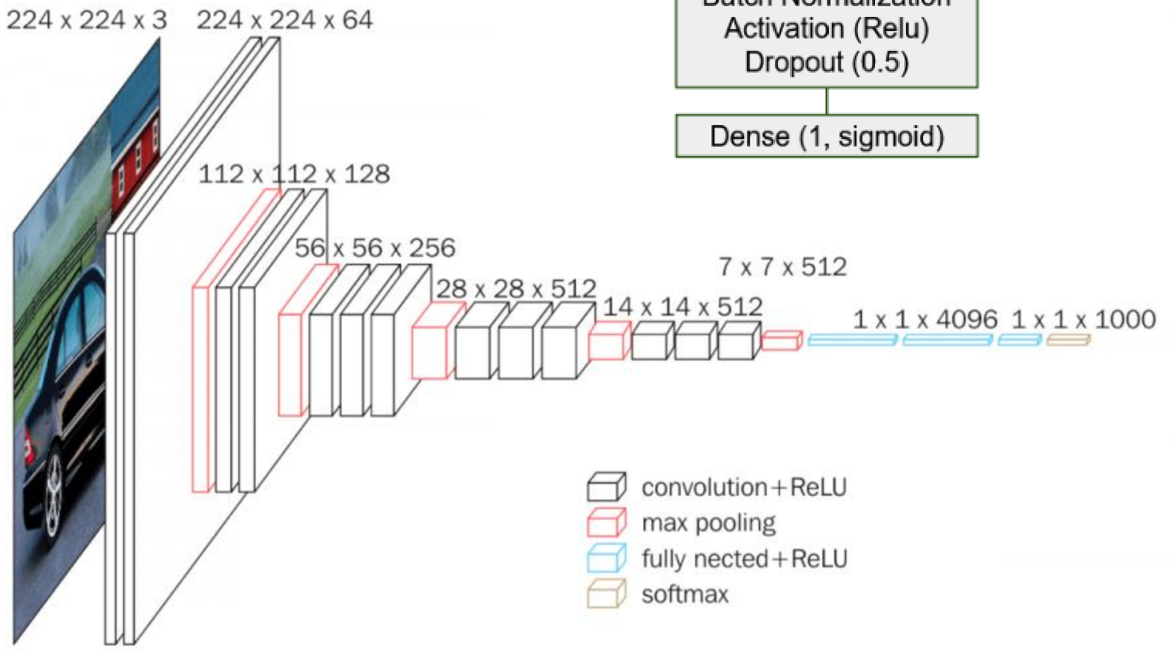




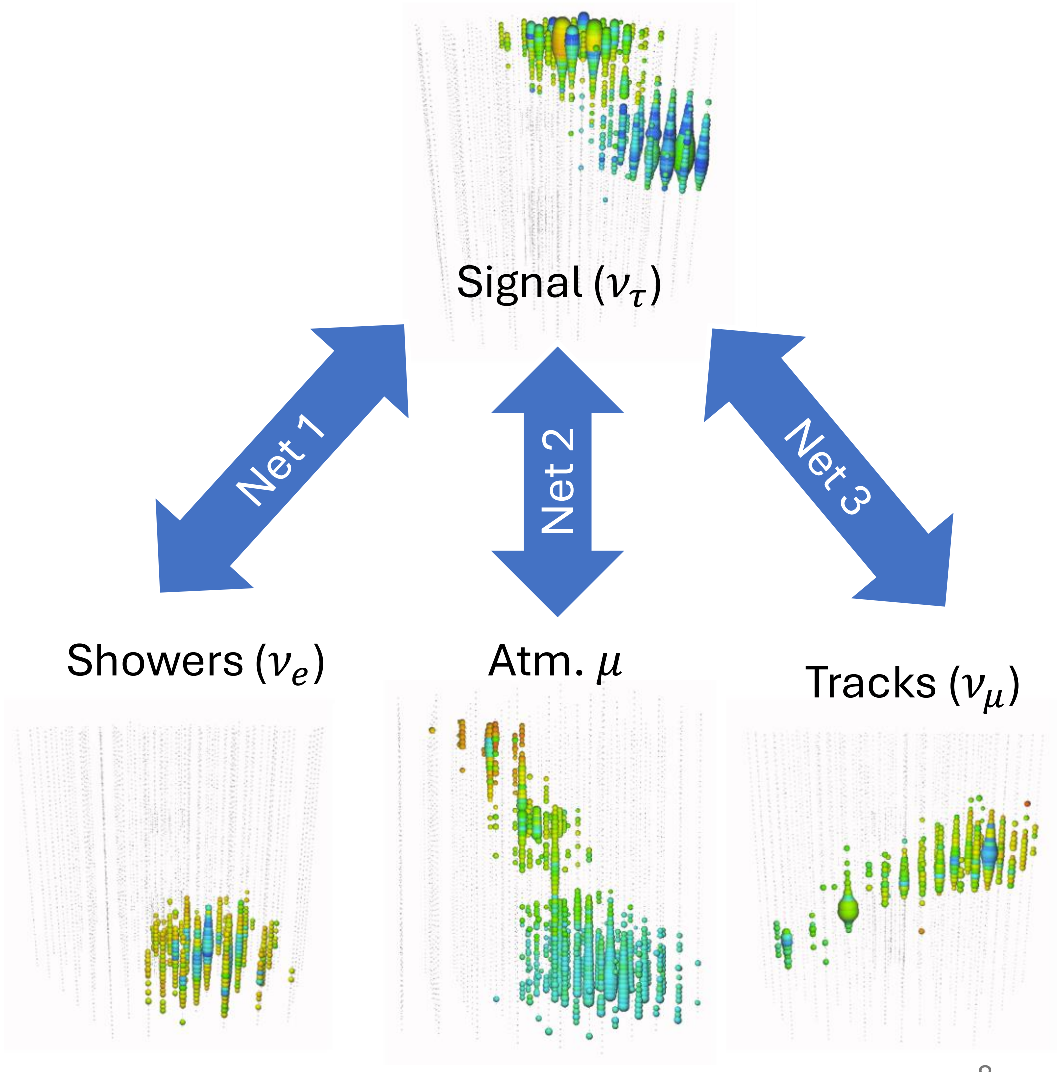
CNN

Based on VGG16





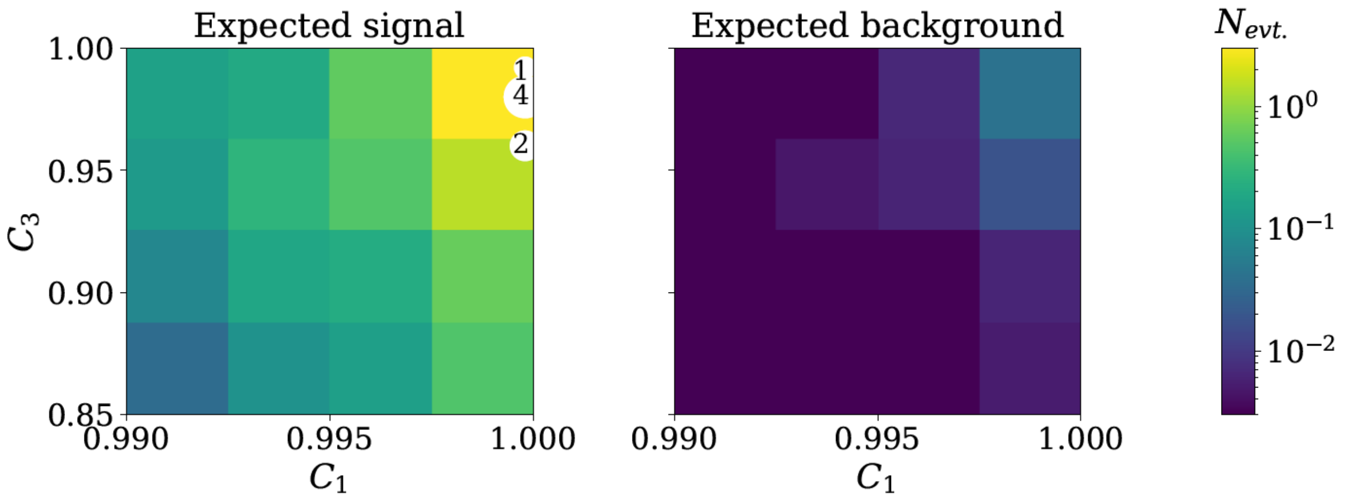
Train thrice with three different classification tasks:

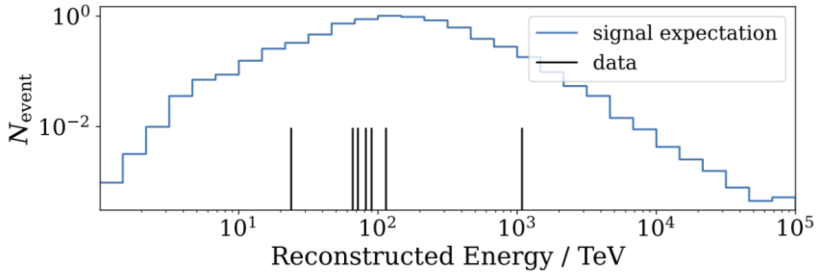


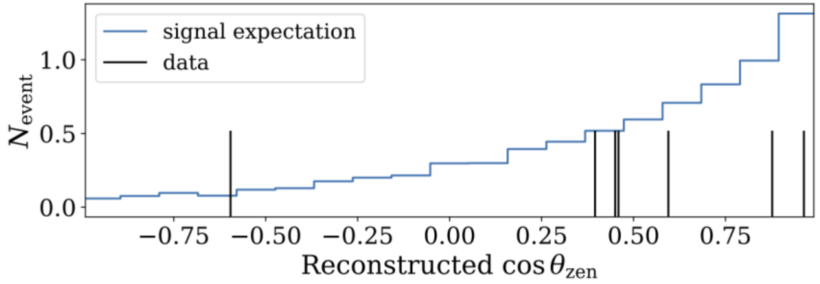
Phys.Rev.Lett. 132 (2024) 15
arXiv: 2403.02516

Results

- While in 9.7 years of data we expected:
  - 6.4 tau neutrinos
  - 0.5 background events
- We find 7 events in the signal region
  - One of these seven was also identified in a previous tau analysis
  - 5 sigma p-value for this being a background fluctuation







Part I: Physics Analysis

- Identifying Tau Neutrinos -

Part II: Interpretability

- Dissecting the Neural Net -

Robustness Tests I

Stress-test the neural networks, to see if they could easily be fooled:

- Uncorrelated variation of the light levels at the DOM level, to mimic mis-modeled uncertainties in DOM efficiencies
- Variation of the light levels for DOMs correlated in bands of depth to mimic unmodeled ice optical property uncertainties
- varying the light levels correlated by entire strings to mimic ice birefringence uncertainty

→ In all cases, the migration probabilities were small and consistent with the expectations from background simulations

Robustness Tests II: Adversarial attacks

Targeted attacks on neural nets aiming to migrate events between classes

→ Adding perturbations to the image in the direction of the gradient

arXiv: 1412.6572

Adversarial Attacks on Tau Analysis

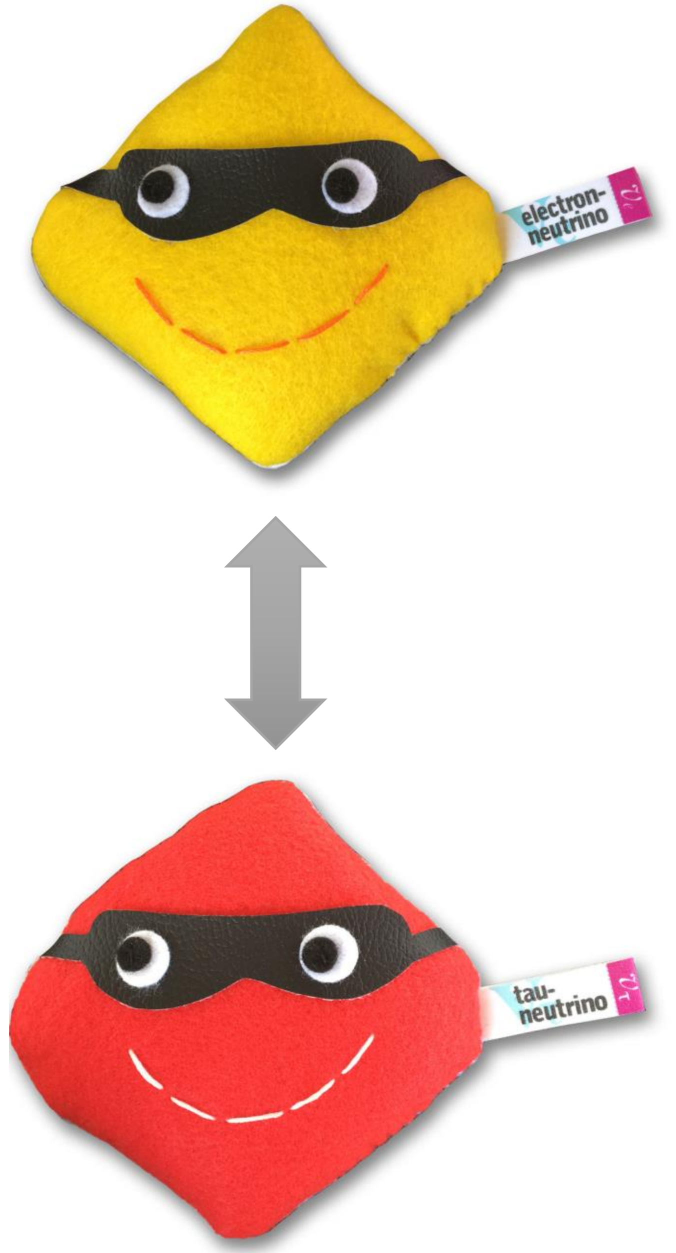
Turn Signal into Background

- For 6 out of the 7 signal events, pixel values would need to be changed so drastically, that it is well outside of what is considered reasonable (>> 10%)
- One event can be turned into background by introducing conceivable changes (~3 %)

  →This is consistent with the expected 0.5 background events

Turn Background into Signal

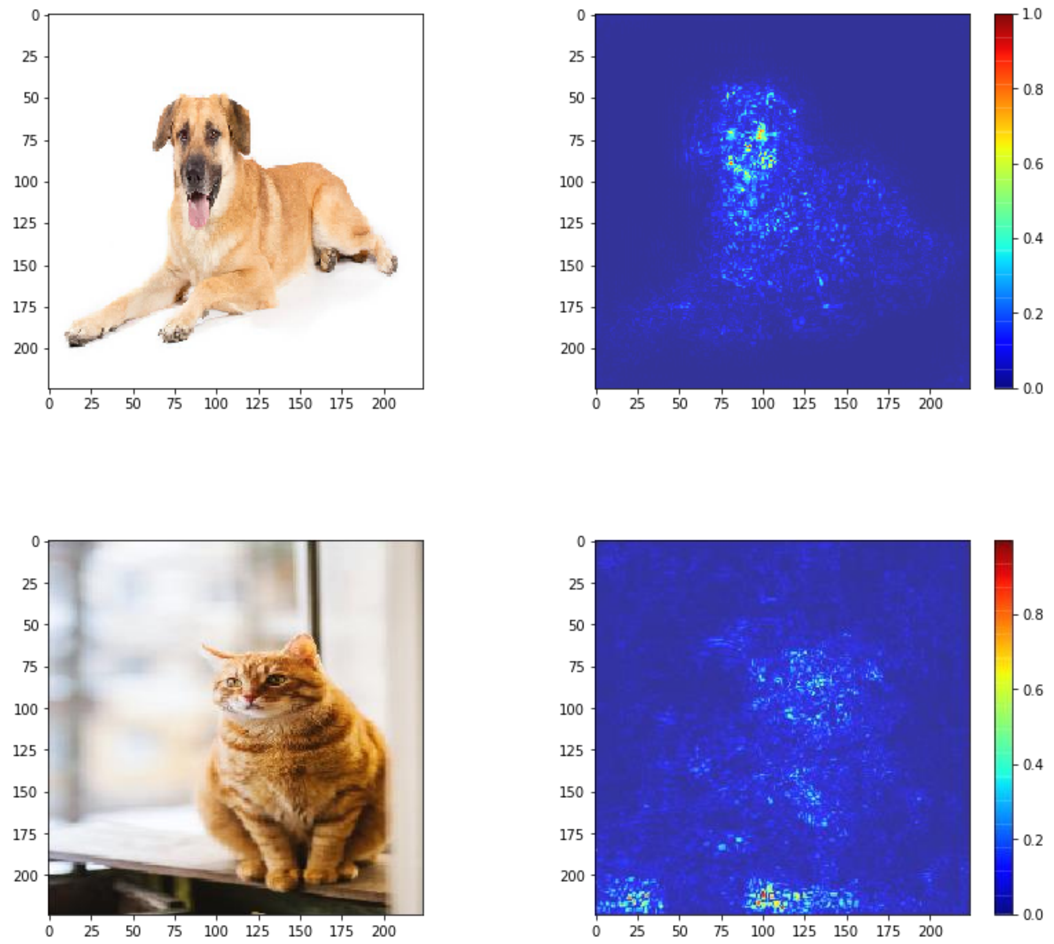
- Only success for one out of 634 attacked background (ν_e) events
- Needed to allow for large deviations (up to 10%)



Saliency Maps

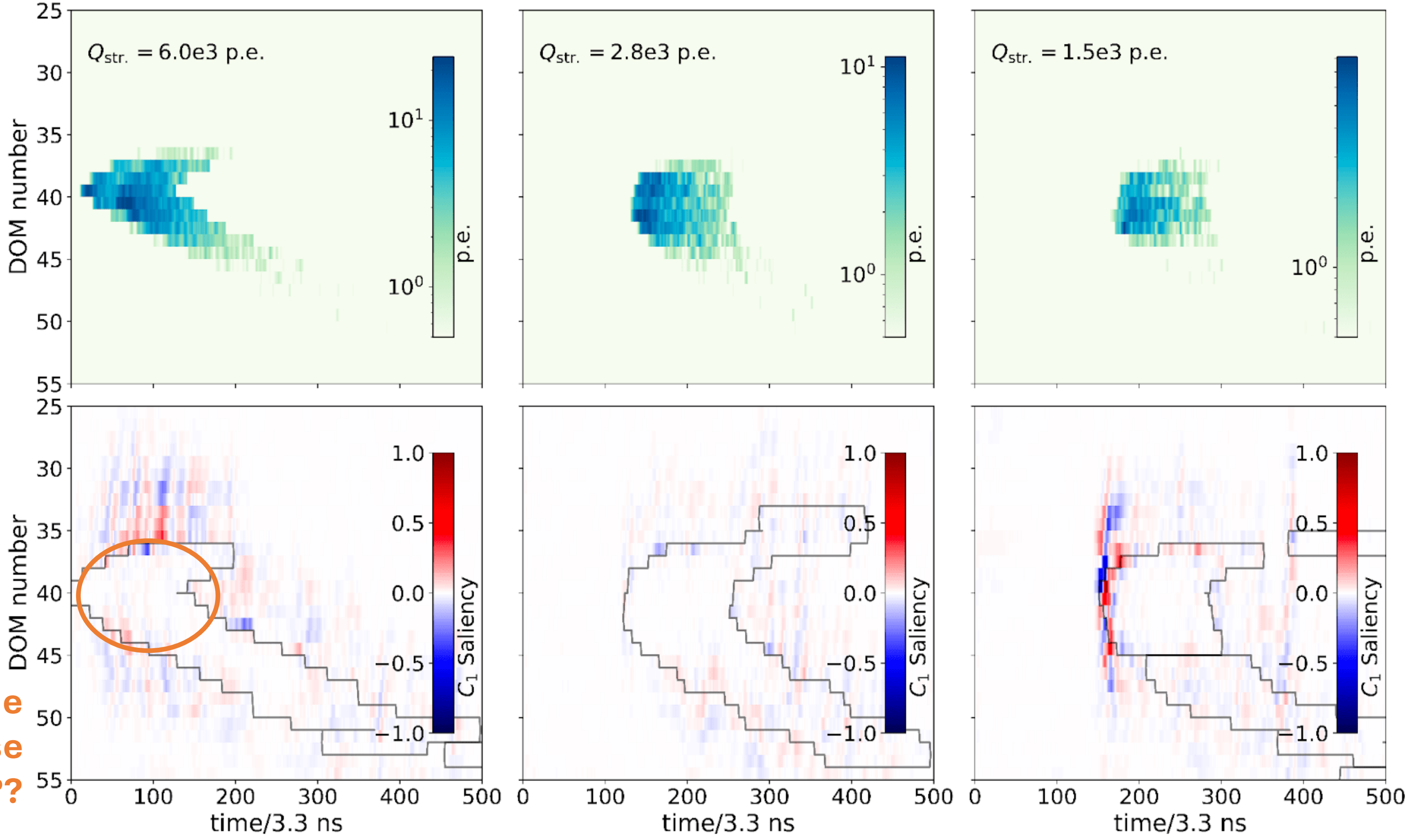
- This shows the importance of the input data on the network output
- Here it is ∂output/∂pixel-value

  (Normalized to 1)
- Large value → change in pixel value will have large impact on network output



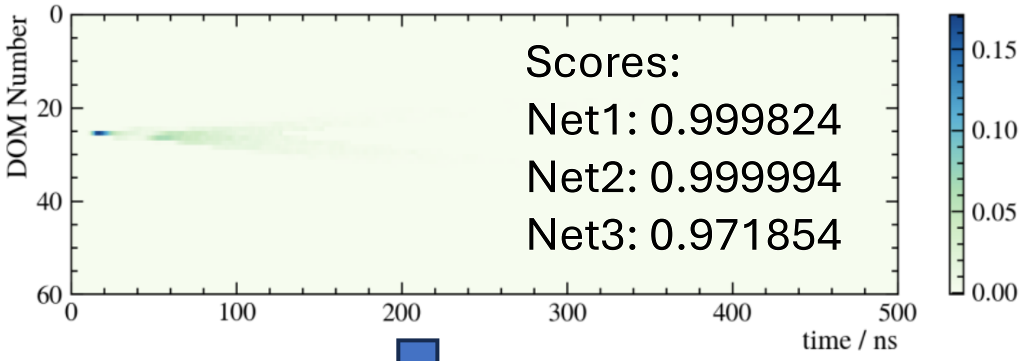
https://usmanr149.github.io/urmlblog/cnn/2020/05/01/Salincy-Maps.html

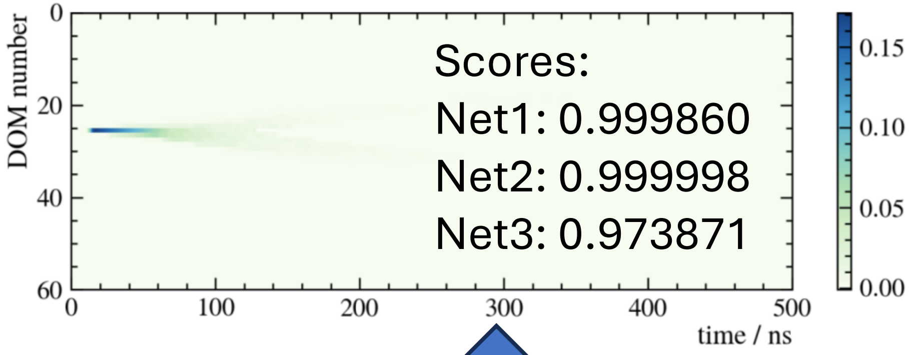
Tau Neutrino Saliency Maps

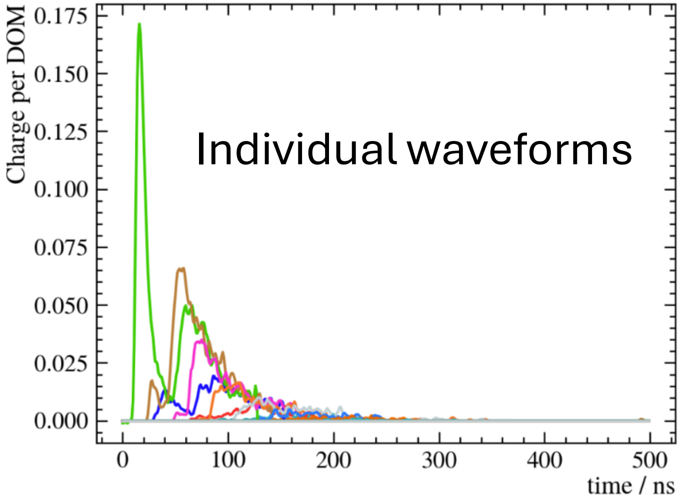


What about the double pulse waveforms??

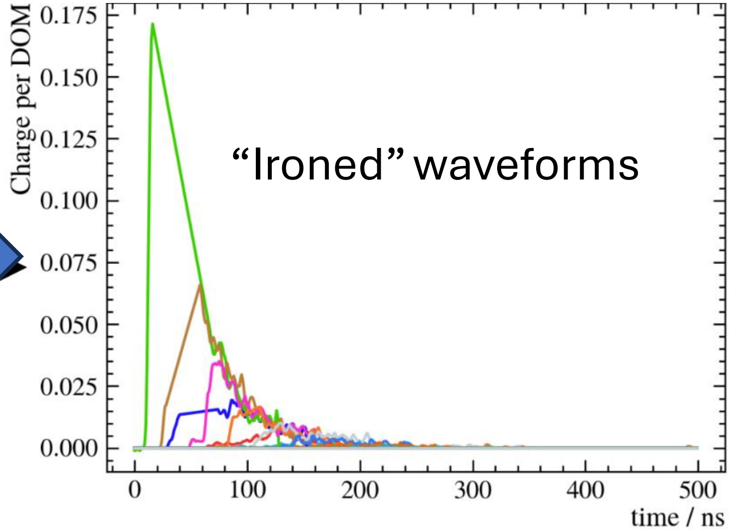
Ironing them out…



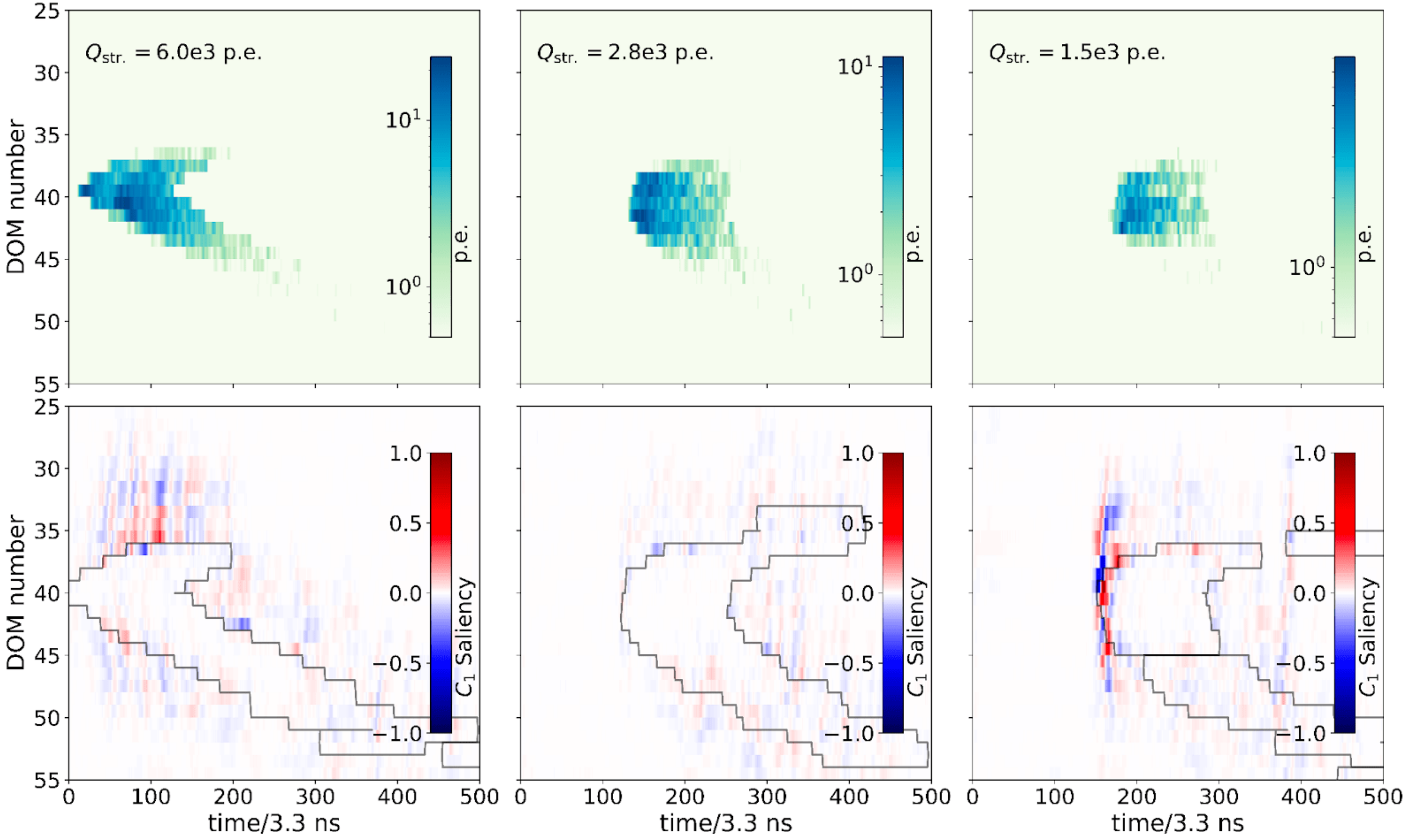




Linear interpolation between separate peaks



Tau Neutrino Saliency Maps



Summary / Conclusions

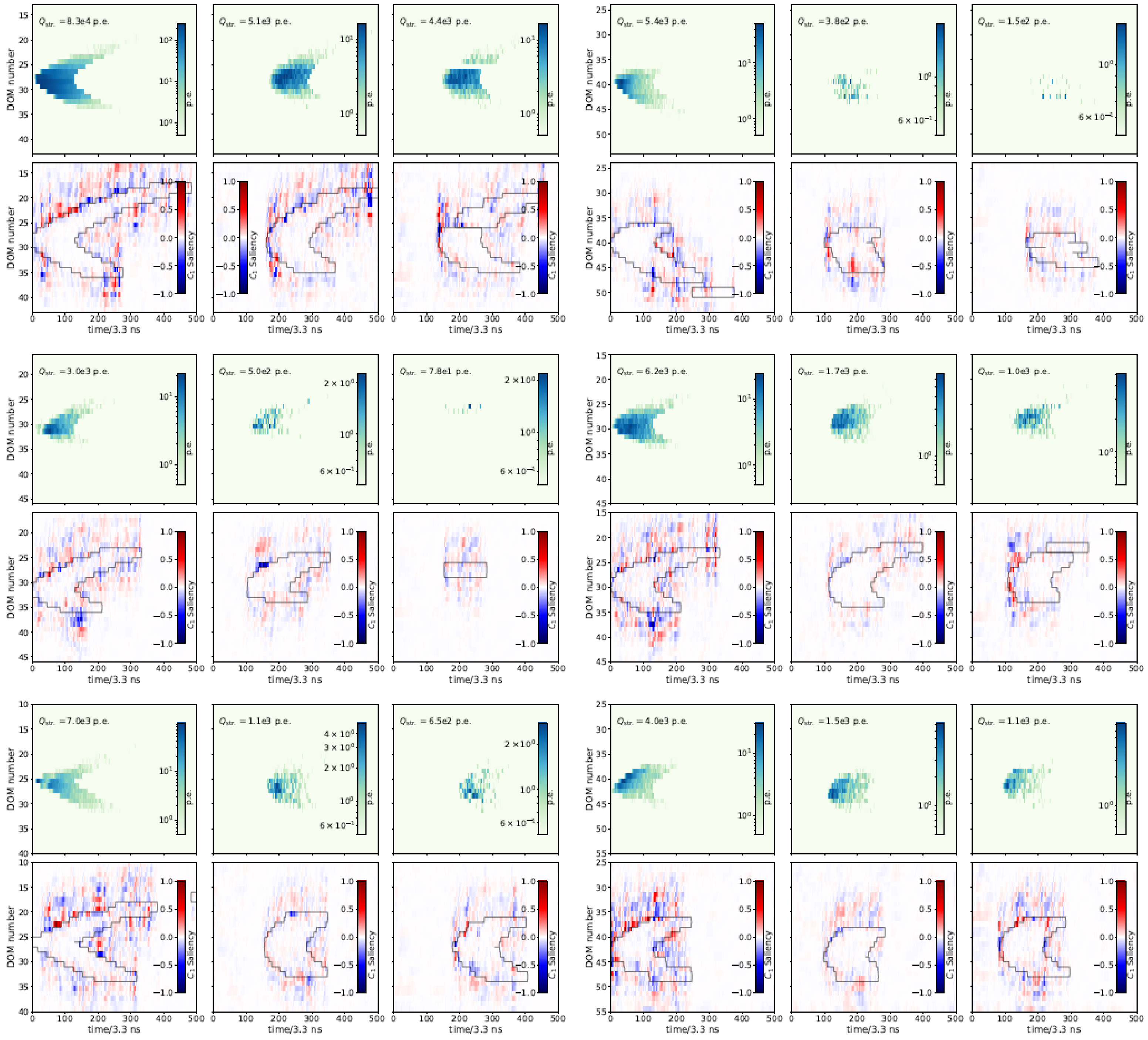
ML methods allowed us to **identify astrophysical tau neutrinos in IceCube for the first time**!

- Applying CNNs to 2d arrays of waveforms from sensors on strings
- 7 signal events observed, while expecting 6.4 tau neutrinos over 0.5 background

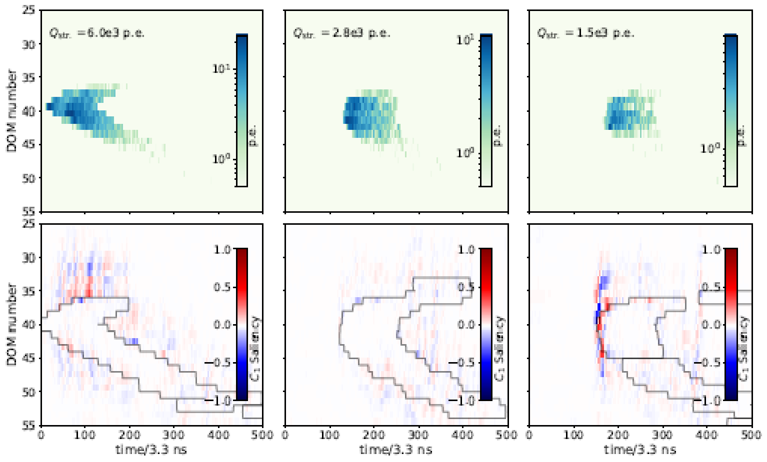
Large efforts were undertaken to test and understand the neural nets:

- **Robustness tests** varying input features according to uncertainties
- Targeted **adversarial attacks** to migrate events between classes
  → NNs proved robust against such attacks
- **Saliency maps** to visualize what features the network uses to form decisions
  - Revealed that double pulse waveform signature is not that useful for signal vs. background classification
  - Rather, the overall shape of the light registered in strings and the timing between strings is important, which gives us **new and better intuition** on how to improve future searches!

Additional Material



All 7 signal candidates (observed events)

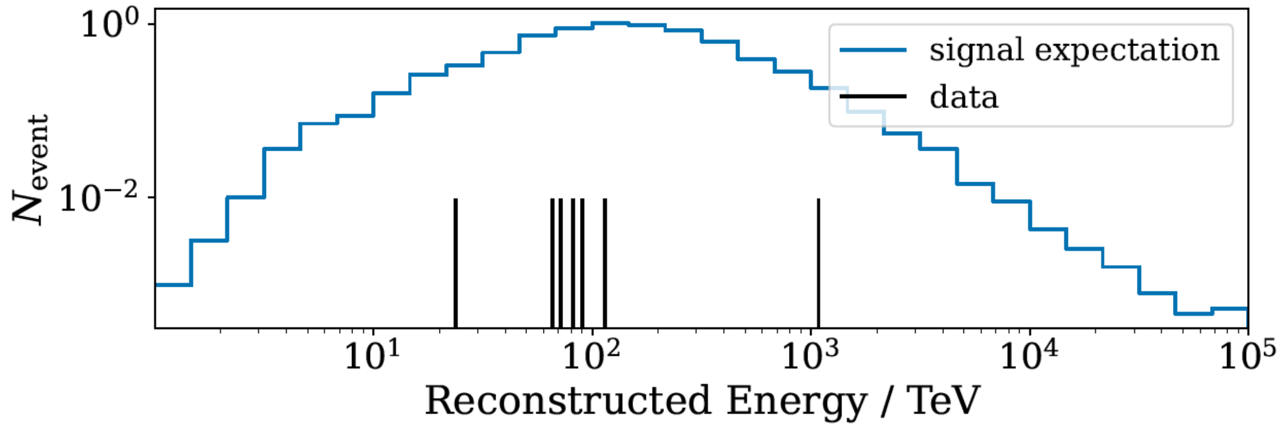


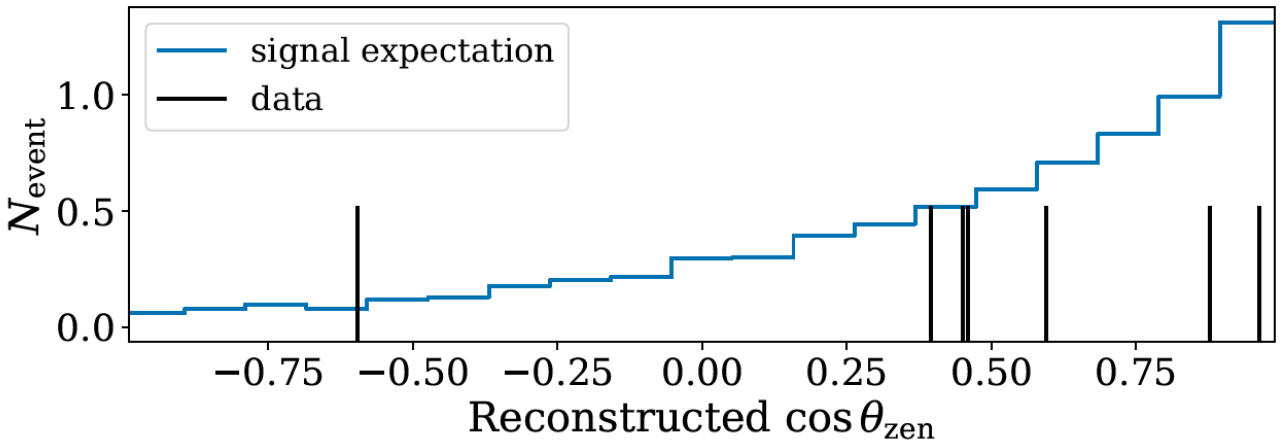
Number of Events

	$\nu^{\text{astro}}_{\tau,CC}$ [59]	$\nu^{\text{astro}}_{\text{other}}$ [59]	$\nu^{\text{atm}}_{\text{conv.}}$ [60–63]	$\nu^{\text{atm}}_{\text{prompt}}$ [56, 64–66]	$\mu^{\text{atm}}_{\text{conv.}}$ [67–70]	all background
initial	160 ± 0.2 (190 ± 0.3)	400 ± 0.7 (490 ± 0.8)	580 ± 7	72 ± 0.1	8400 ± 110	9450 ± 110 (9540 ± 110)
final	6.4 ± 0.02 (4.0 ± 0.02)	0.3 ± 0.02 (0.2 ± 0.01)	0.1 ± 0.008	0.1 ± 0.001	0.01 ± 0.008	0.5 ± 0.02 (0.4 ± 0.02)

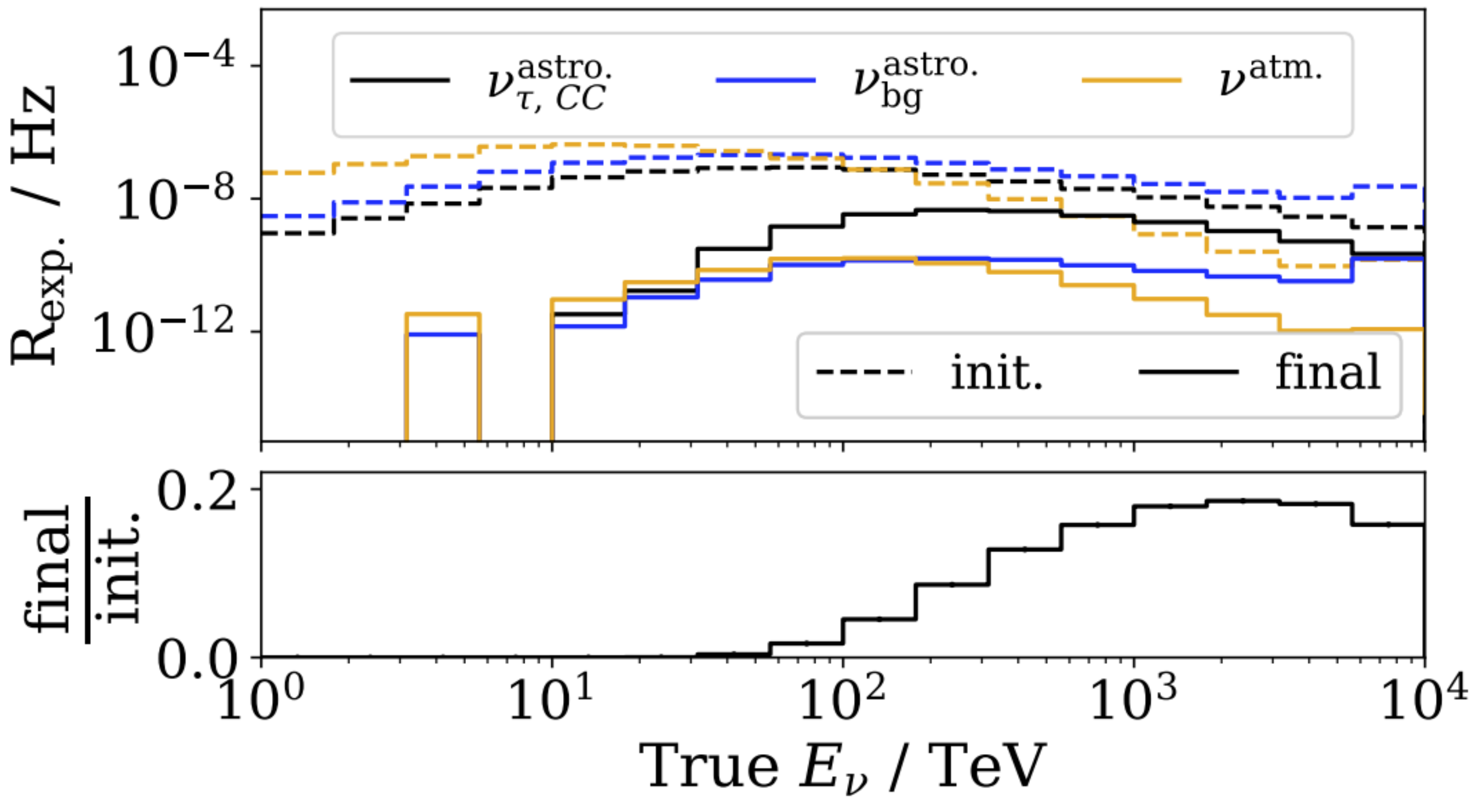
TABLE I. Expected number of events after initial and final set of selection criteria (including all corrections described in the text) for signal ($\nu^{\text{astro}}_{\tau,\,\text{CC}}$) and backgrounds, assuming IceCube's flux from Refs. [53] and (in parentheses) [56]. About 85% of the estimated contribution from $\nu^{\text{atm}}_{\text{prompt}}$ is from ν_τ. Signal and astrophysical background levels vary with the flux. The simulation did not include the self-veto effect [71] that would reduce the conventional (conv.) and prompt ν^{atm} backgrounds. References to associated simulation packages are given; see text for details. Errors are statistical only, arising from finite simulation samples.

Reconstructed Signal Events



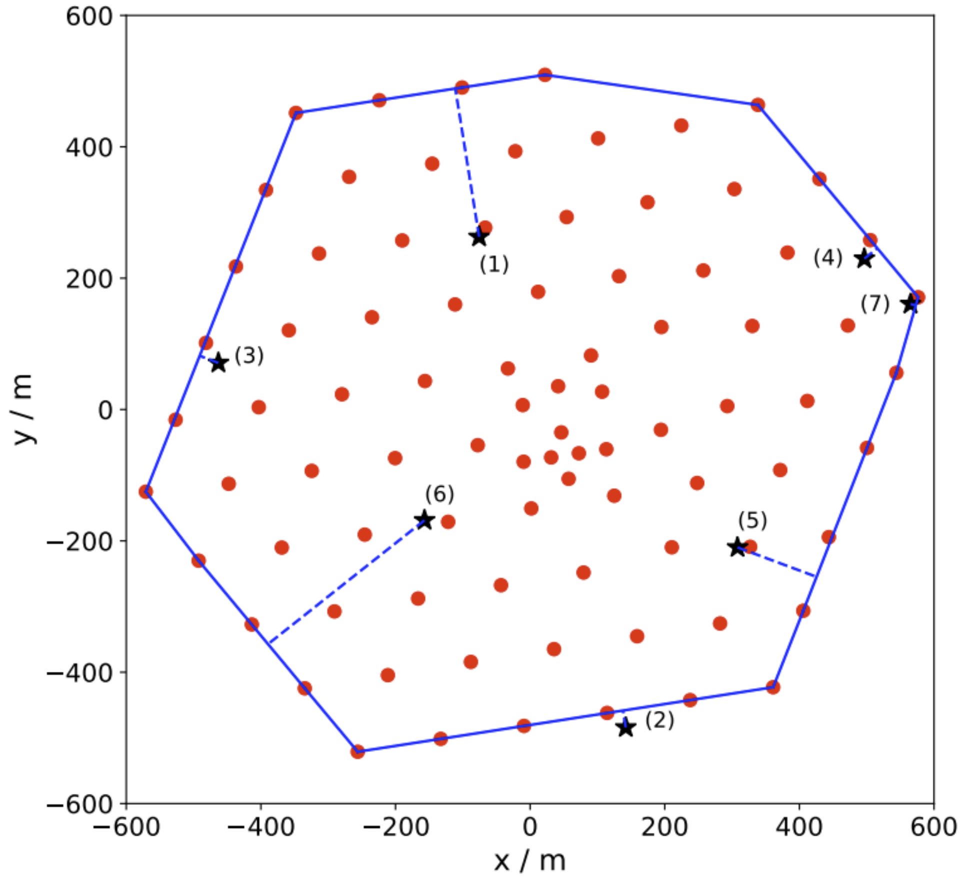


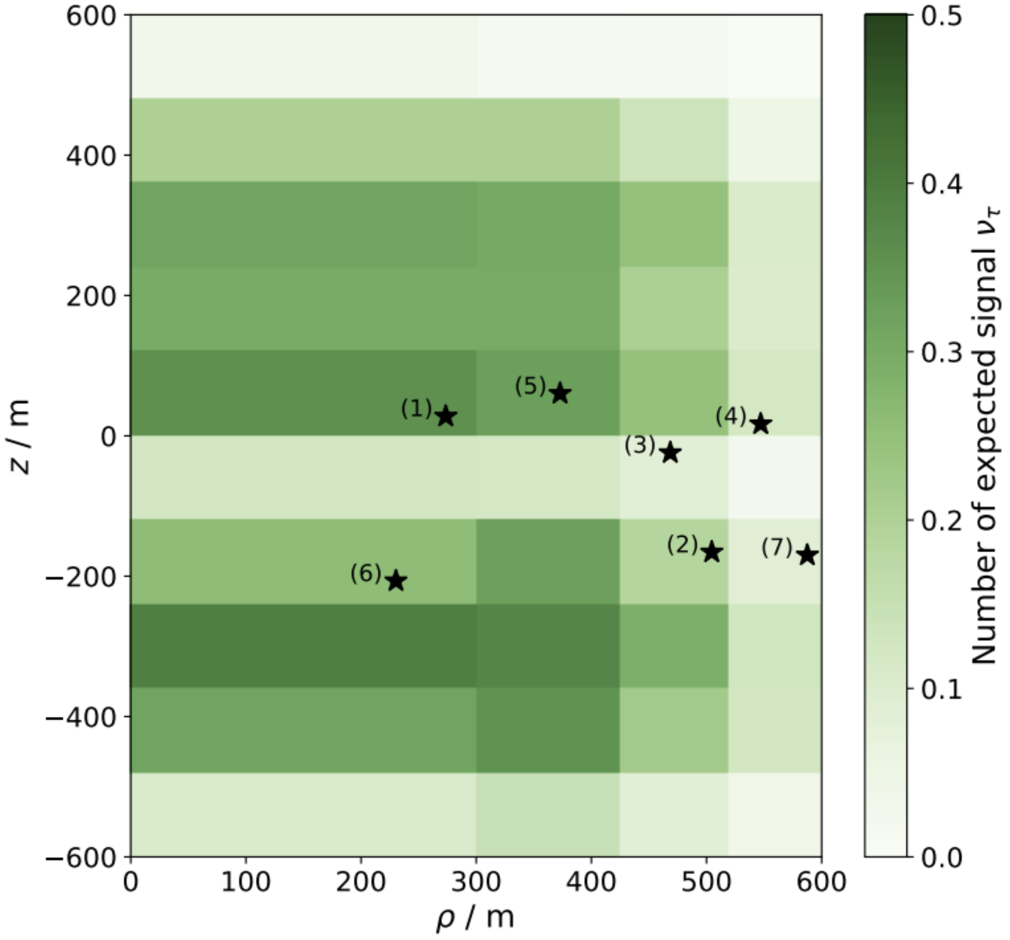
Spectrum



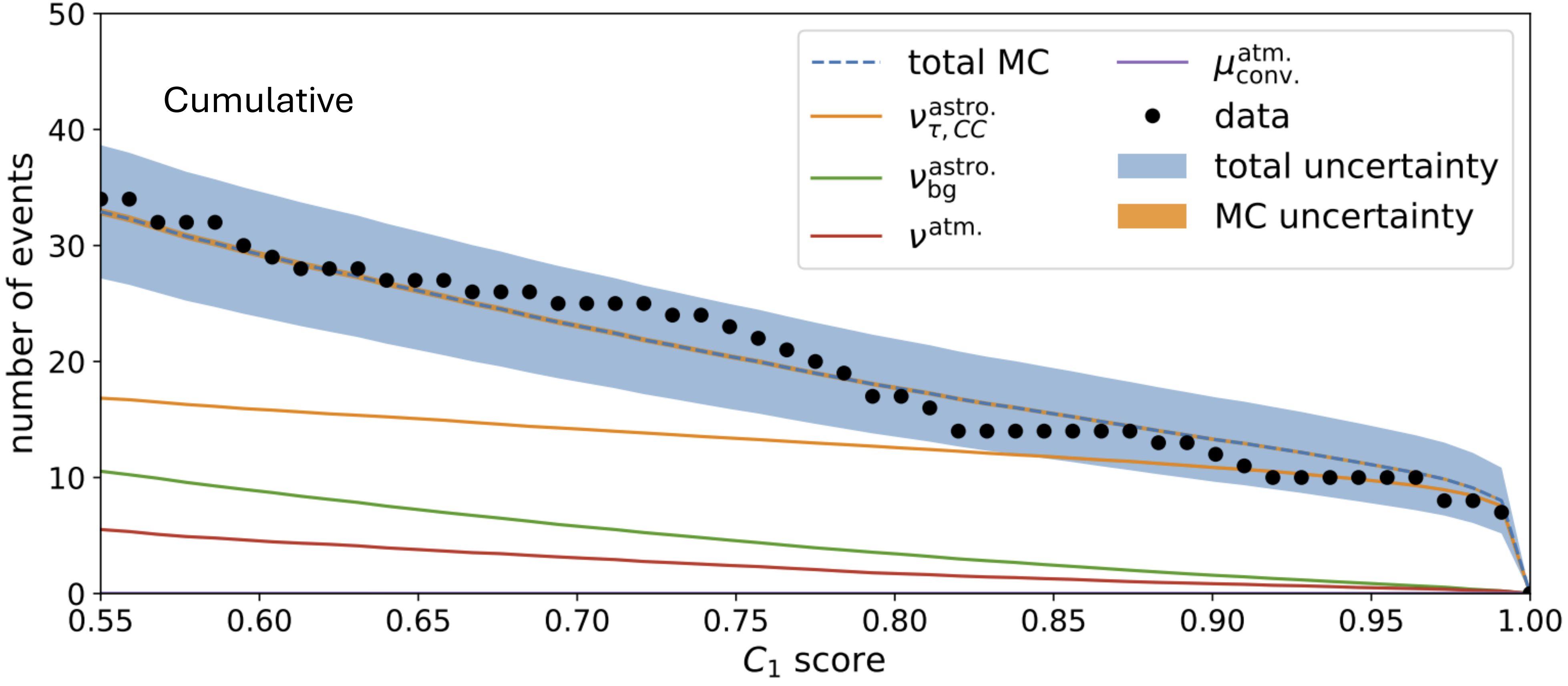
10^{-4}
10^{-8}
10^{-12}
$R_{exp.}$ / Hz
$\nu_{\tau,\ CC}^{astro.}$
$\nu_{bg}^{astro.}$
$\nu^{atm.}$
init.
final
0.2
0.0
$\frac{final}{init.}$
10^0
10^1
10^2
10^3
10^4
True E_ν / TeV

Spatial Distribution



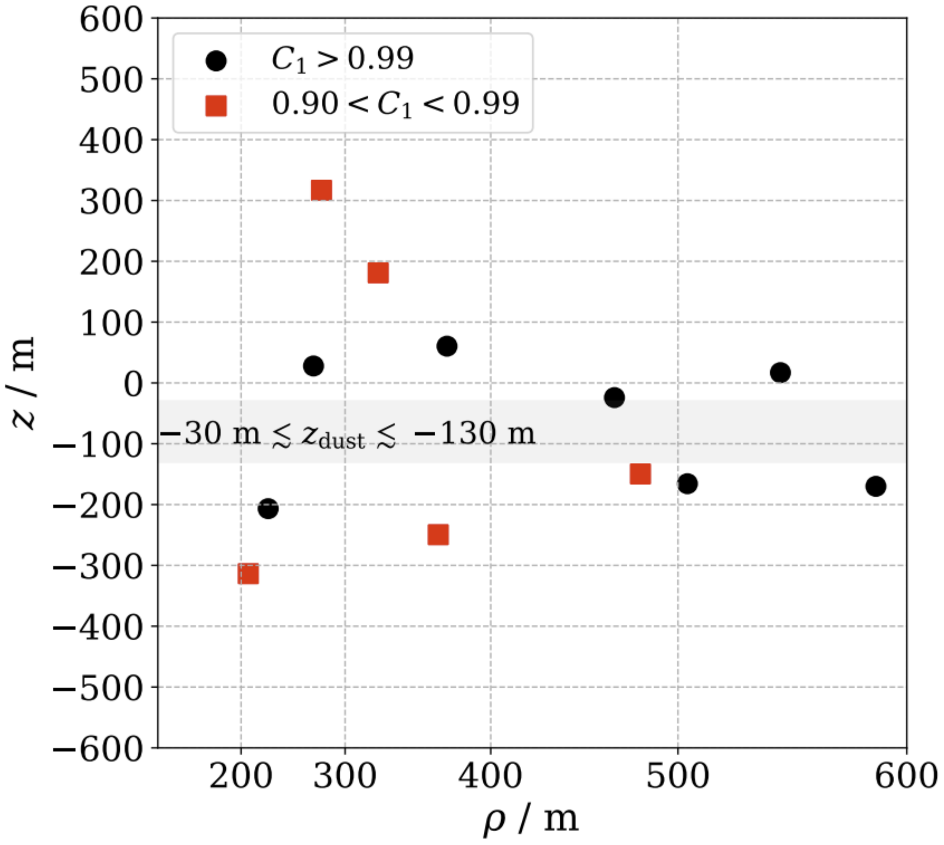


Classifier vs. data

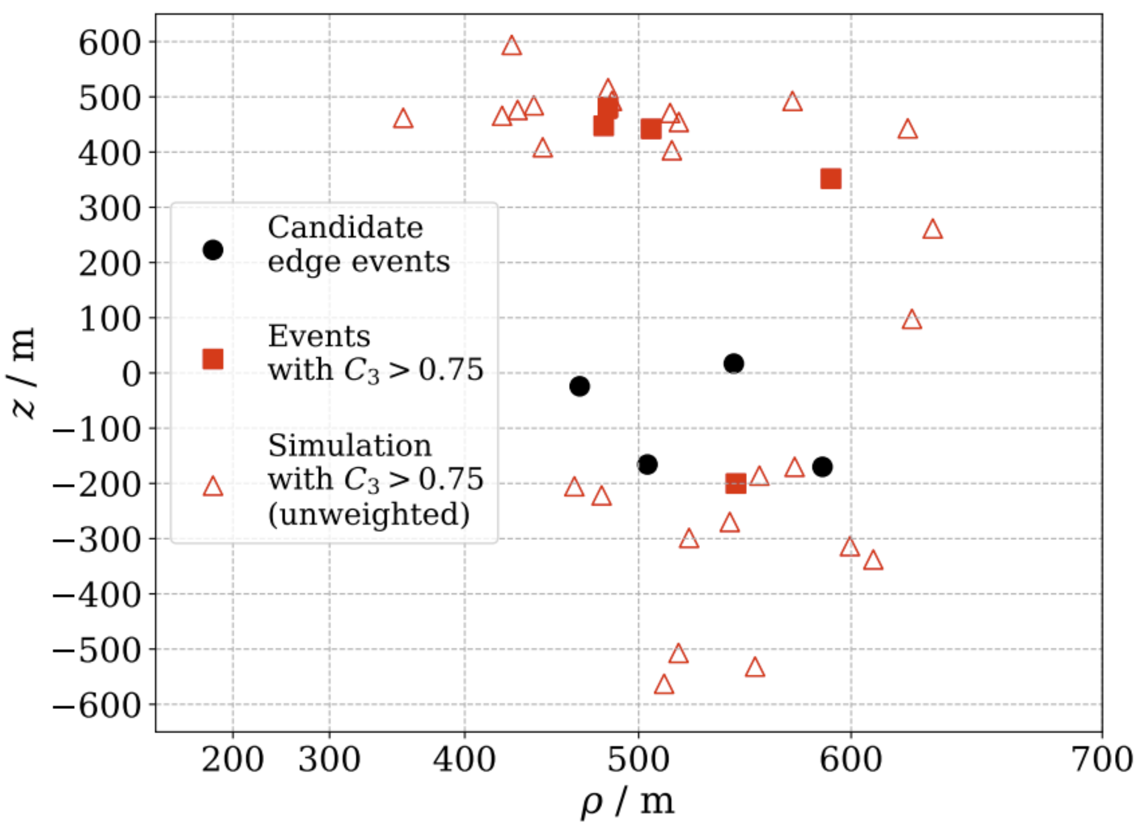


Relaxed Cuts

Net 1



Net 3



Adversarial Attack

Successful conversion of the one signal event to background