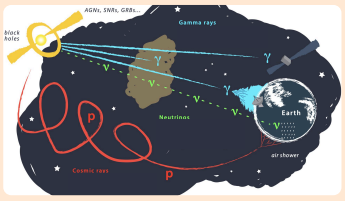
Attention for High Energy Neutrino Reconstruction

Fast multi-task event reconstruction for all IceCube event morphologies

Aske Rosted*[1], Maximilian Meier[1], Shigeru Yoshida[1], Aya Ishihara[1], Nobuhiro Shimizu[1]
[1]Chiba University; International Center For Hadron Physics (ICEHAP)

IceCube and real-time astrophysical neutrinos

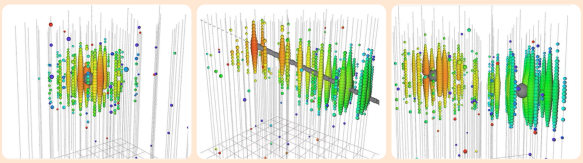
- Real-time reconstruction since 2016
- Computationally efficient algorithms used for *fast* reconstruction of neutrino physics parameters
- Neutrinos of astrophysical origin are rare
- Large sources of background from atmospheric muons & neutrinos
- Different selections/reconstructions targeting different event morphologies



Event Morphologies at IceCube

Current approach
- Low level selection targeting *one* event morphology
- Morphology based reconstruction

This work
- All type reconstruction
- Reconstruction based selections



Cascade
ν_e and neutral current interactions

Track
Charged current interactions

Double-bang
τ production from ν_τ

Model confidence metric

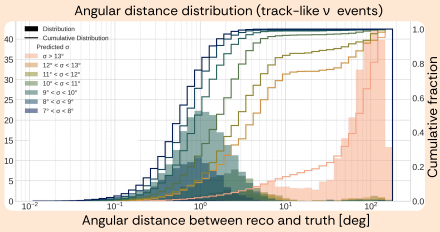
- Von Mises-Fisher loss for the directional reconstruction

$$\kappa = |\bar{r}|$$

- For the 3-dimensional case κ is related to the mean angular spread through

$$\sigma \approx 1/\sqrt{\kappa}$$

- Can be used to remove misreconstructed events



MC target labels for the multitask model

- **Interaction-vertex/closest approach:** (x,y,z,t)
- **Trackness:** Energy deposited from tracks compared to the full energy deposit
- **Visible length:** Length of the track or expansion of the cascade inside the detector volume
- **Deposited energy:** Energy from the primary/daughters of the primary deposited inside the detector volume
- **Energy on entrance:** The energy of the highest energy pseudo-particle as it enters the detector volume
- **Direction:** x,y,z-vector of the direction of the particle along with an confidence metric κ

Multi-task Transformer

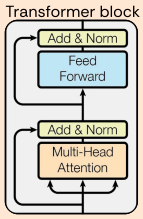
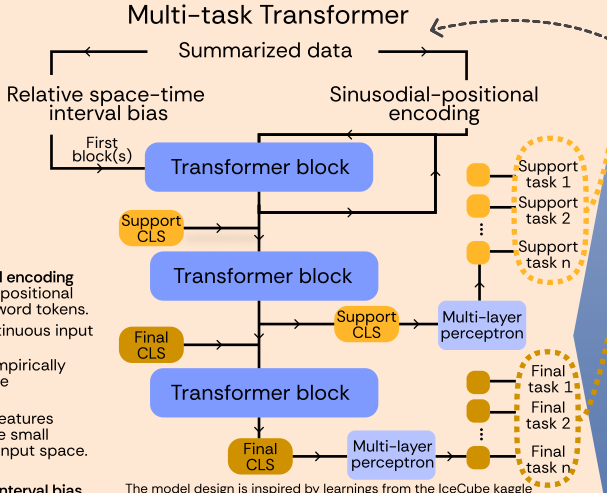


Figure modified from [1]



The model design is inspired by learnings from the IceCube kaggle competition [2]

Sinusoidal-positional encoding
- Used in LLMs for positional encoding of the word tokens.
- Used on the continuous input features.
- Used as it was empirically shown to increase performance.
- Scaling of input features crucial to capture small variations in the input space.

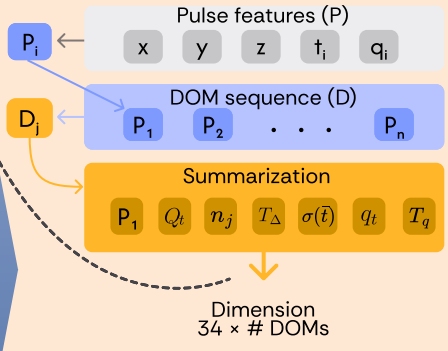
Relative space time interval bias
- Minkowski space-time

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$
$$ds = sign(ds^2) \cdot \sqrt{|ds^2|}$$

- Clipped at (-4,4) and passed through the positional-encoder as well as a learnable affine linear transformation

Multi-task neural networks
- Often used in visual transformers where the final task might be too complex to train directly
- Based on the idea of *inductive transfer* - learning multiple task at the same time can increase stability and reduce parameters by sharing part of the network

Summarization of pulses



n : number of pulses per DOM

$Q_t = \Sigma_{i=0}^{n_j} q_i$

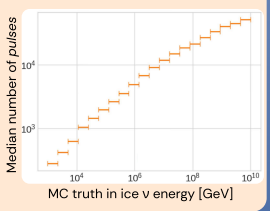
$T_\Delta = t_n - t_1$

T_q = Timing at fraction of total charge

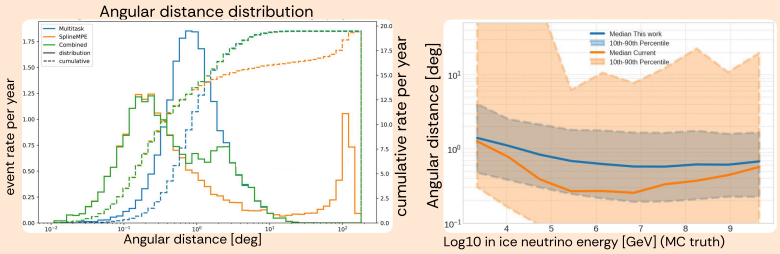
q_t = Charge after t time

$\sigma(\bar{t})$ = time standard deviation

Summarization of pulses
- Necessary due to the large number of pulses (individual charge recordings by the PMTs) for high energy events
- The transformer network scales as $\mathcal{O}(n^2 \cdot d)$ with d being the feature dimension and n being the number of activated DOMs
- These summarization features are inspired by summarization features used by the DNN cascade reconstruction used by IceCube

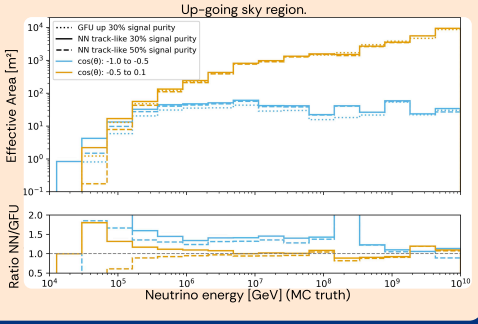


Directional reconstruction performance

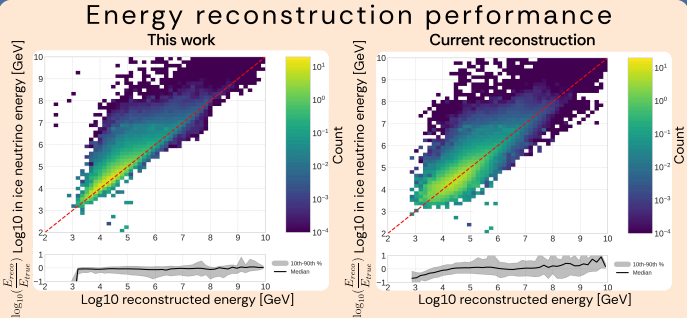


- Performance here shown for a quality selection designed for this analysis.
- These performance metrics are weighted to an astrophysical neutrino flux.
- This selection targets track-like events that are well reconstructed by the multi-task model.

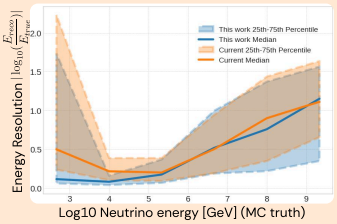
GFU-like selection
- By using the new reconstructed features a similar selection to the current GFU selection can be created.
- Results in **~50% more** astrophysical up-going neutrinos at the same purity (30%)
- Median angular distance of **0.29°** compared to the 0.22° of the GFU selection on the same MC sample
- ~50% of the selection are already present in the current GFU selection
- yearly astrophysical rate **~18.5 (~10.5)** at 30% (50%) purity compared to GFU ~12.3 (~7.5)



Energy reconstruction performance



- Comparison showed for the current **GFU selection** to ensure good performance of Likelihood based method
- The GFU selection targets *track-like* events
- Less biased and smaller uncertainty across the entire energy range.
- Largest difference appears for "low" energy events
- The improved energy resolution can be used to more effectively separate ν_{atmos} from ν_{astro}



References

[1] *Attention is all you need* June 2017
[2] IceCube -- Neutrinos in Deep Ice The Top 3 Solutions from the Public Kaggle Competition October 2023