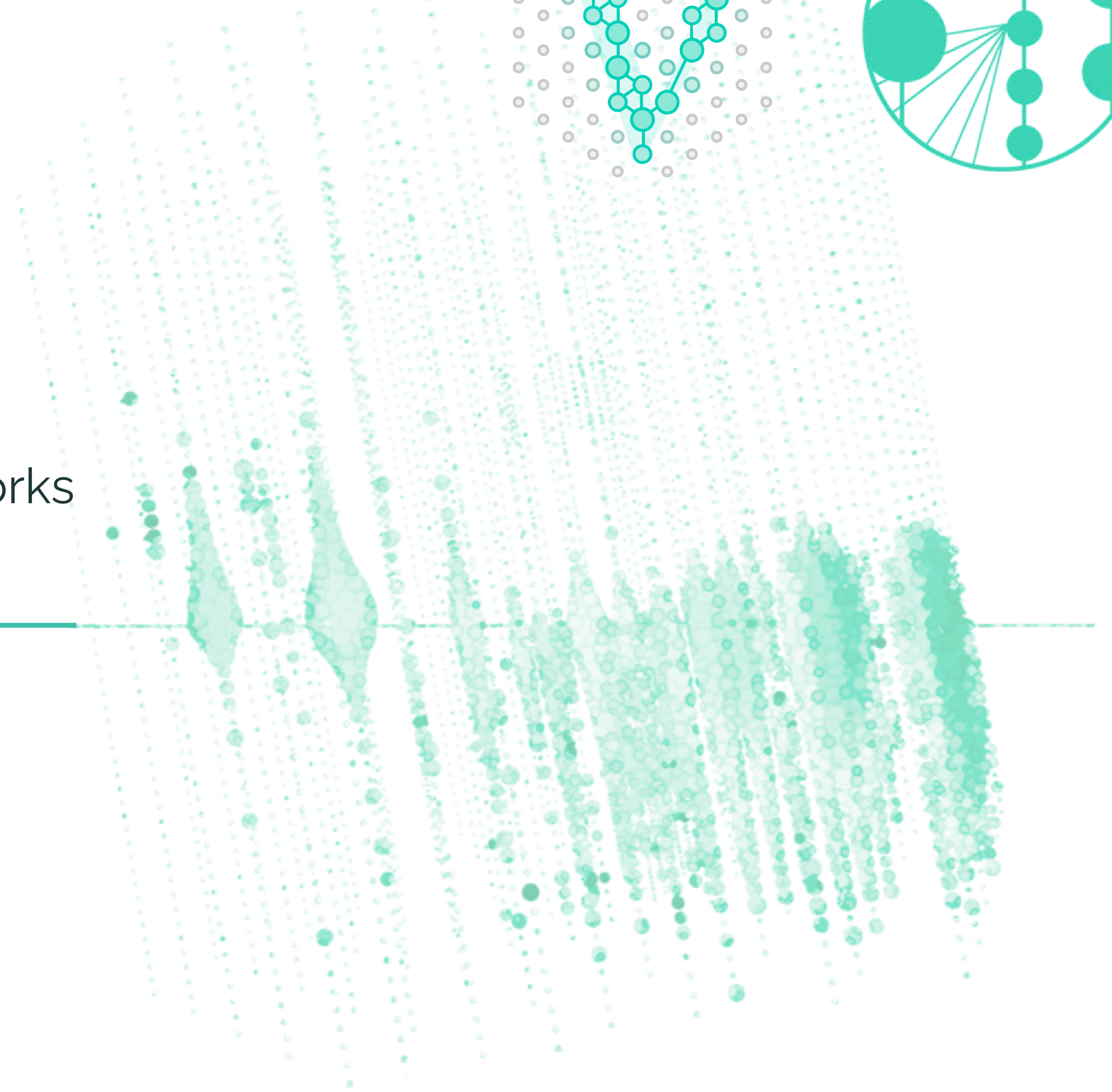
UNIVERSITY OF
COPENHAGEN

The Good, the Bad and the Noisy

Characterisation of IceCube Neutrino Events Using Graph Neural Networks



Master Thesis Defense in Computational Physics

Kaare Endrup Iversen

June 27th 2022

Supervised by Troels Christian Petersen

Introduction

Theory & Motivation

Machine Learning in IceCube

Current State of Affairs

My Work

Time Regression

Event Cleaning

Upgrade Pulse Cleaning

Regression of Cleaned Pulses

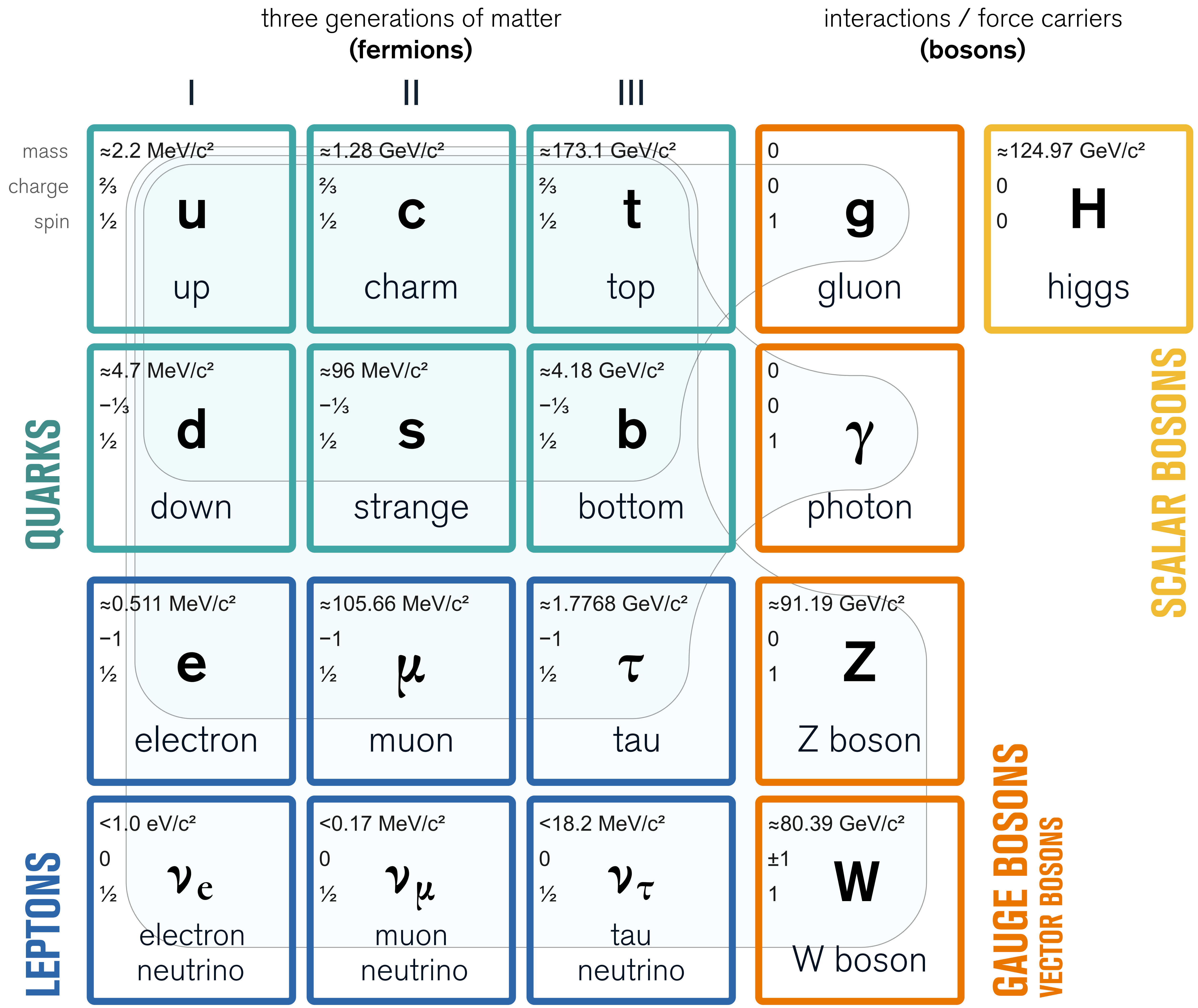
Intermezzo 1: Improvement of Pulse Cleaning

Intermezzo 2: Quality of Cleaned Events

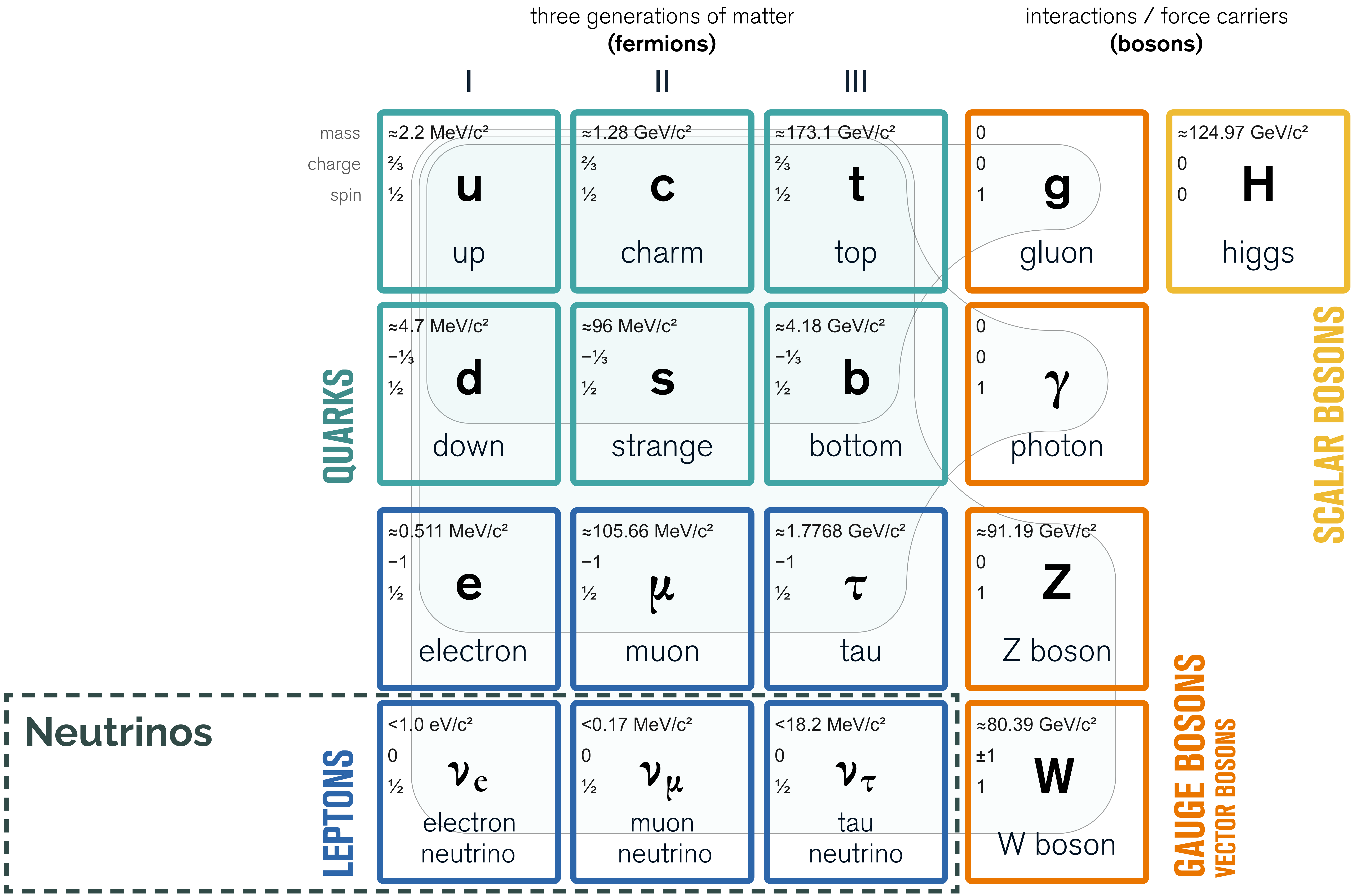
Conclusions & Outlook

The Standard Model

- Coherent description of matter and it's behavior
- Rigorously tested
- Incomplete

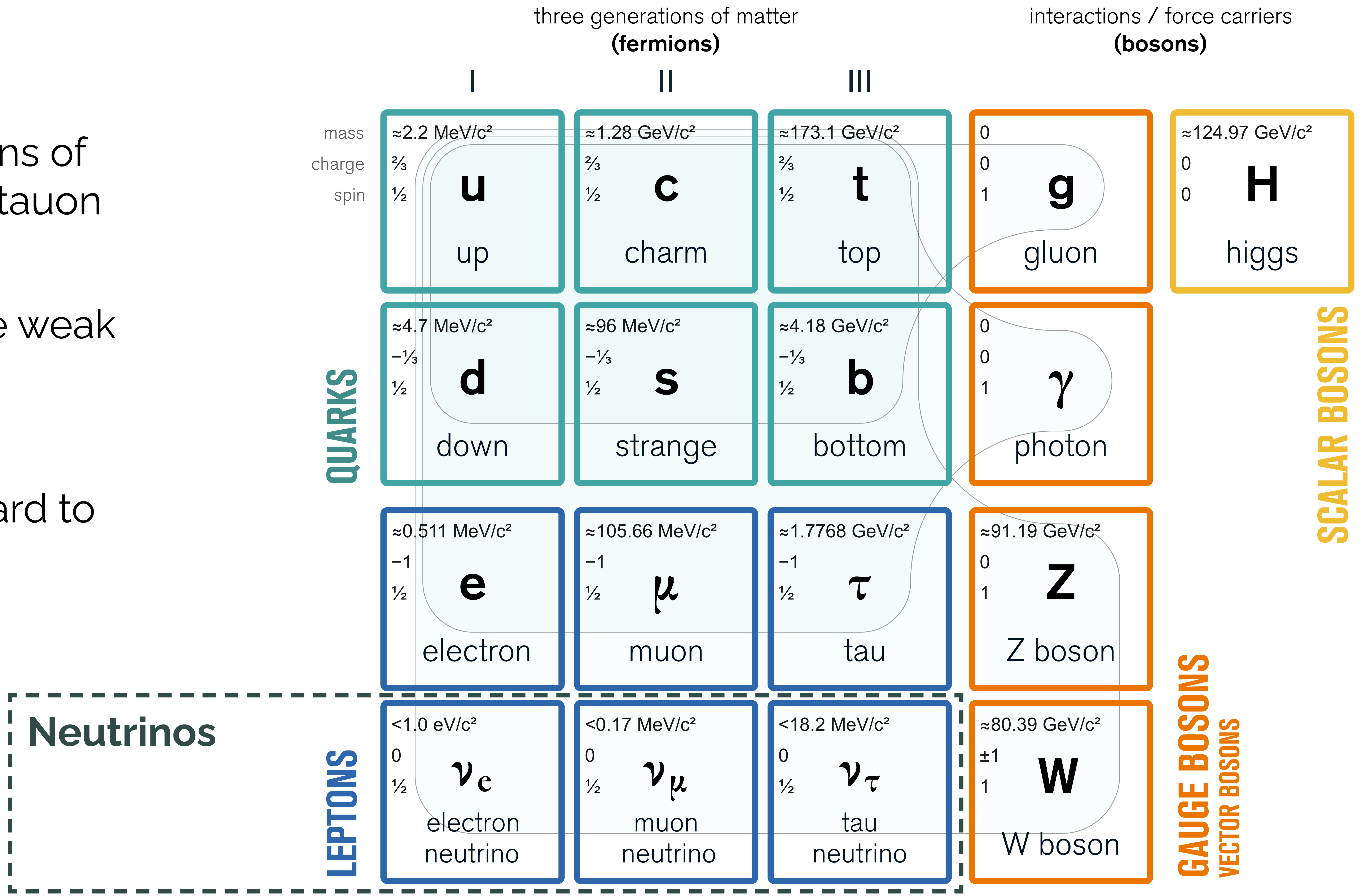


Neutrinos


three generations of matter
(fermions)
interactions / force carriers
(bosons)
I
II
III
mass
charge
spin
≈2.2 MeV/c²
⅔
½
u
up
≈1.28 GeV/c²
⅔
½
c
charm
≈173.1 GeV/c²
⅔
½
t
top
0
0
1
g
gluon
≈124.97 GeV/c²
0
0
H
higgs
≈4.7 MeV/c²
−⅓
½
d
down
≈96 MeV/c²
−⅓
½
s
strange
≈4.18 GeV/c²
−⅓
½
b
bottom
0
0
1
γ
photon
≈0.511 MeV/c²
−1
½
e
electron
≈105.66 MeV/c²
−1
½
μ
muon
≈1.7768 GeV/c²
−1
½
τ
tau
≈91.19 GeV/c²
0
1
Z
Z boson
<1.0 eV/c²
0
½
νe
electron neutrino
<0.17 MeV/c²
0
½
νμ
muon neutrino
<18.2 MeV/c²
0
½
ντ
tau neutrino
≈80.39 GeV/c²
±1
1
W
W boson
QUARKS
LEPTONS
Neutrinos
SCALAR BOSONS
GAUGE BOSONS
VECTOR BOSONS

Neutrinos

- Light and neutral cousins of the eletron, muon and tauon
- Interact only trough the weak force (and gravity)
- Overly abundant but hard to observe



Neutrino Oscillations

- Unexpected ratios of neutrino flavours observed in experiments
- Probability of observing a given flavour changes with distance travelled
- Neutrino oscillations!

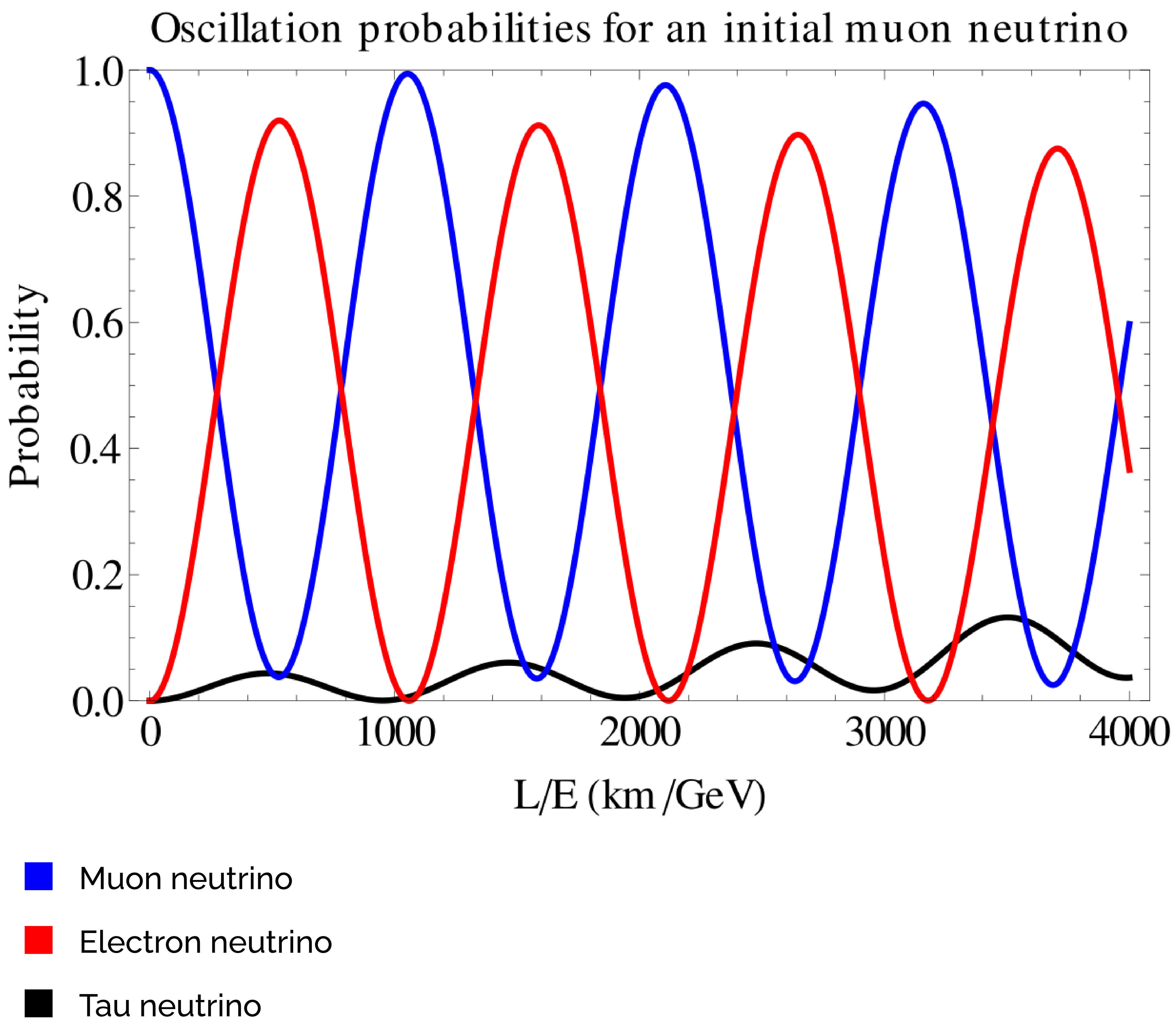


Figure from https://en.wikipedia.org/wiki/Neutrino_oscillation

Neutrino Oscillations

- Eigenstates of weak interactions: ν_e, ν_μ, ν_τ
- Eigenstates of free propagation: ν_1, ν_2, ν_3
- Unitary transformation between the two eigenbases is described by the PMNS-matrix:

$$|\nu_\alpha\rangle = \sum_{k \in 1,2,3} U^*_{\alpha k} \, |\nu_k\rangle$$

$$\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}$$

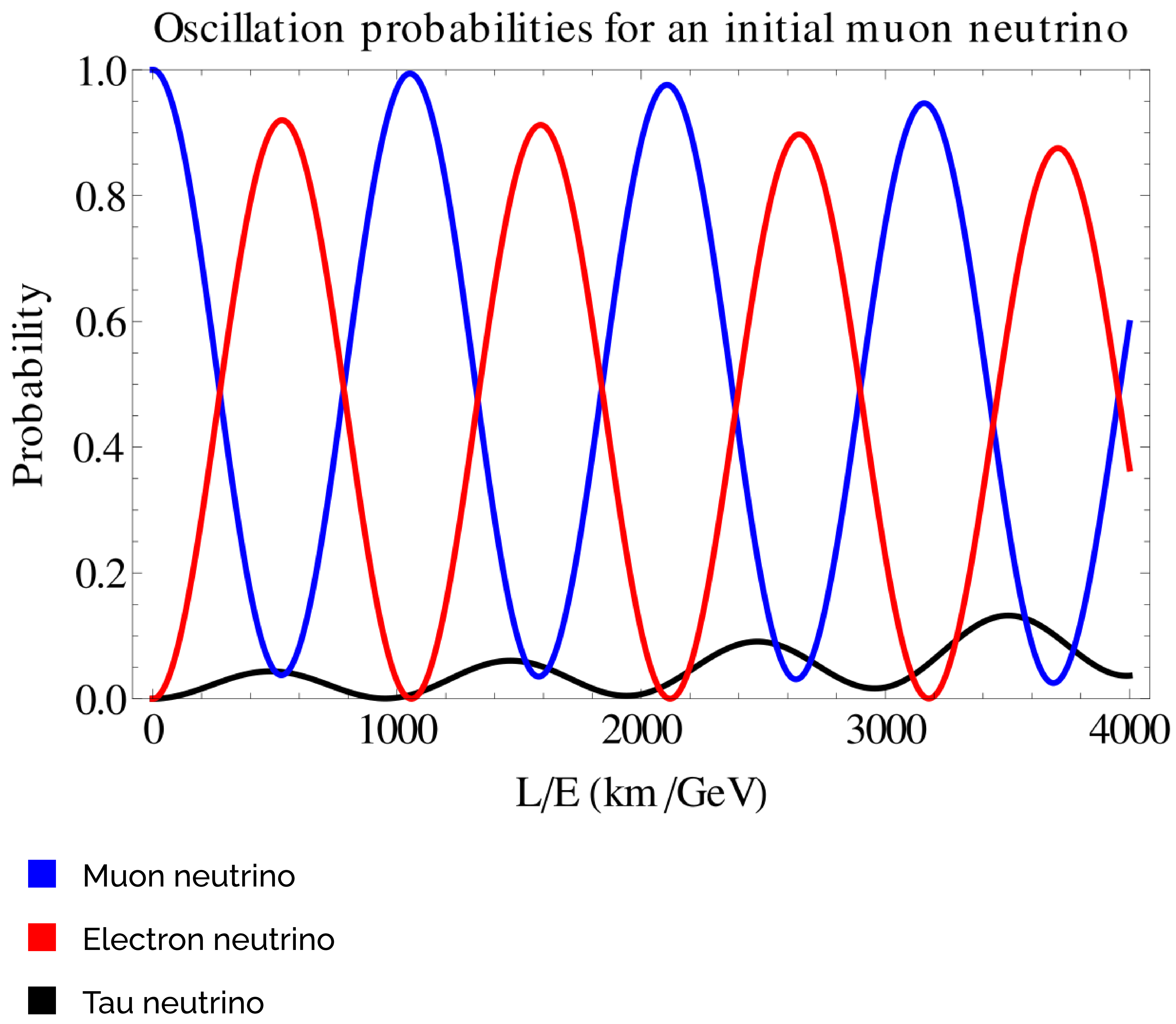


Figure from https://en.wikipedia.org/wiki/Neutrino_oscillation

Neutrino Oscillations

- Eigenstates of weak interactions: ν_e, ν_μ, ν_τ
- Eigenstates of free propagation: ν_1, ν_2, ν_3
- Unitary transformation between the two eigenbases is described by the PMNS-matrix:

$$|\nu_\alpha\rangle = \sum_{k\in 1,2,3} U^*_{\alpha k} |\nu_k\rangle$$

$$c_{ij} = \cos(\theta_{ij}) \text{ and } s_{ij} = \sin(\theta_{ij})$$

$$\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{\mathrm{CP}}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{\mathrm{CP}}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Neutrino Oscillations

- Probability of oscillating from one flavour state to another

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 2 \sum_{k,l \in 1,2,3} \mathrm{Re}\left(U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l}\right) \sin^2\left(\frac{\Delta m^2_{kl} L}{4E}\right) + \sum_{k,l \in 1,2,3} \mathrm{Im}\left(U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l}\right) \sin^2\left(\frac{\Delta m^2_{lk} L}{2E}\right)$$

Neutrino Oscillations

- Probability of oscillating from one flavour state to another

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 2 \sum_{k,l \in 1,2,3} \mathrm{Re}\left(U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l}\right) \sin^2\left(\frac{\Delta m^2_{kl} L}{4E}\right) + \sum_{k,l \in 1,2,3} \mathrm{Im}\left(U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l}\right) \sin^2\left(\frac{\Delta m^2_{lk} L}{2E}\right)$$

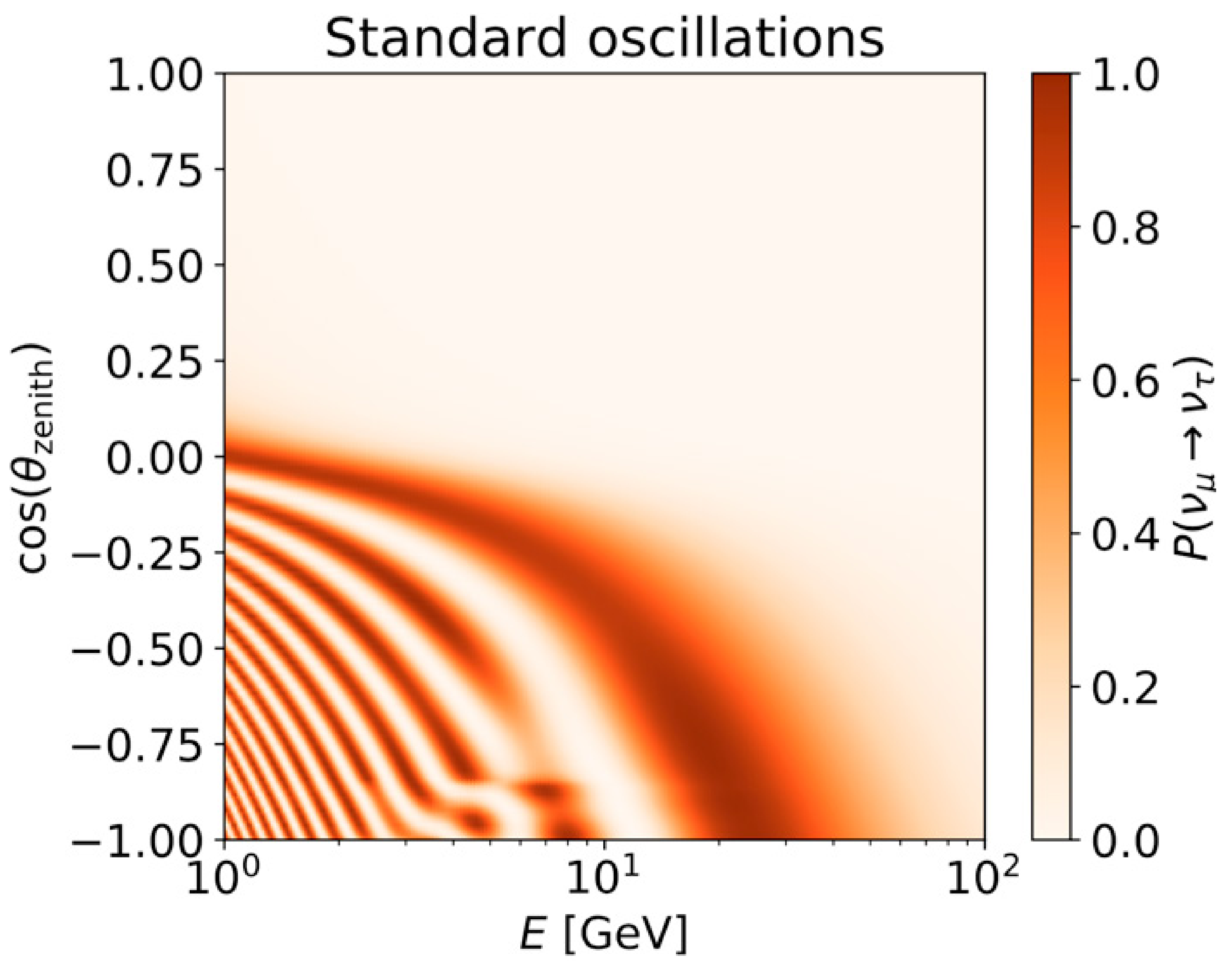
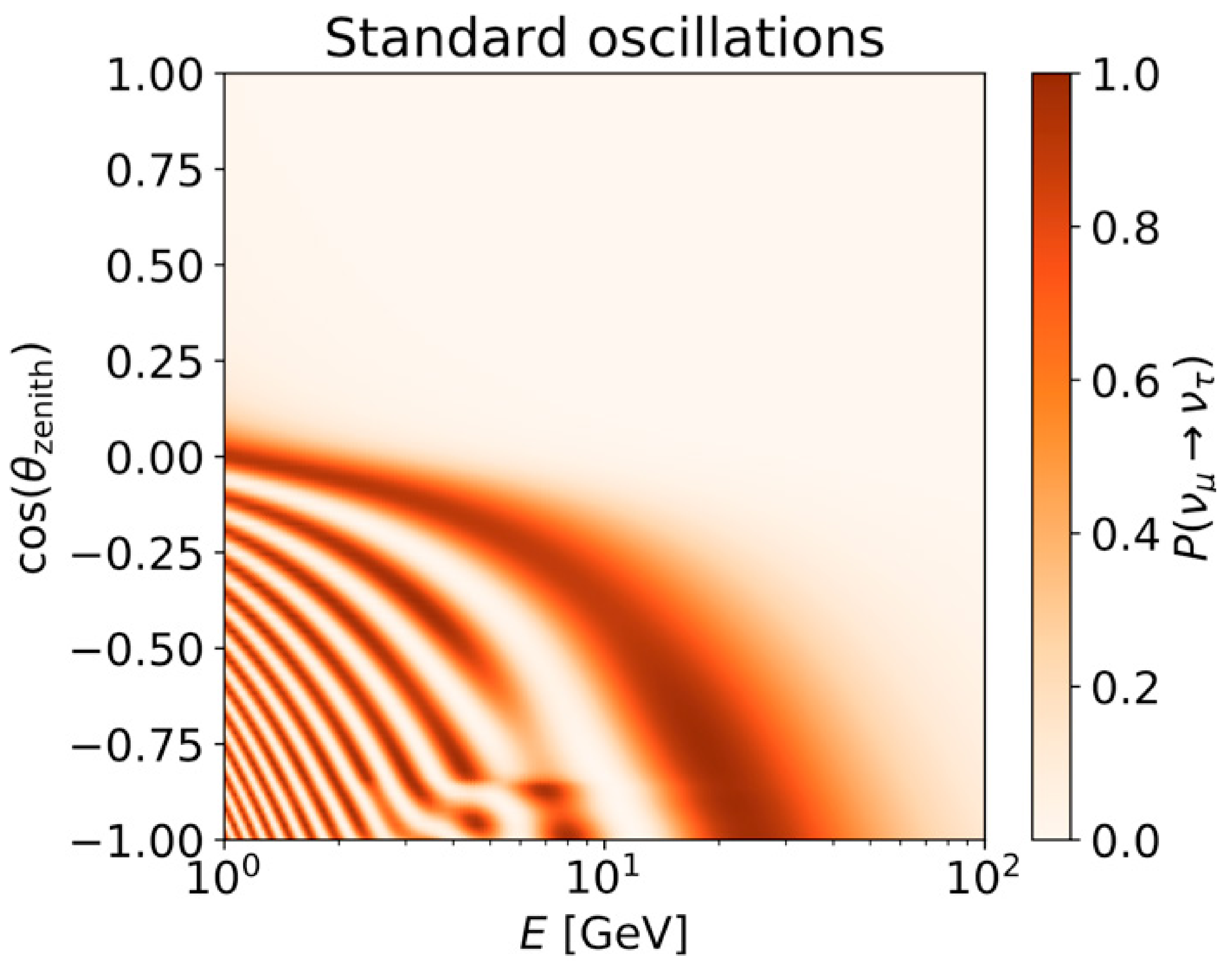


Figure from
https://nbi.ku.dk/english/research/experimental-particle-physics/icecube/neutrino-oscillation/

Partial conclusion: Accurate measurement of neutrino energy, direction and flavour composition will enable accurate calculation of PMNS-matrix elements and mass differences.

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 2 \sum_{k,l \in 1,2,3} \mathrm{Re}\left(U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l}\right) \sin^2\left(\frac{\Delta m^2_{kl} L}{4E}\right) + \sum_{k,l \in 1,2,3} \mathrm{Im}\left(U^*_{\alpha k} U_{\beta k} U_{\alpha l} U^*_{\beta l}\right) \sin^2\left(\frac{\Delta m^2_{lk} L}{2E}\right)$$



Introduction

Theory and motivation

Machine Learning in IceCube

- IceCube
- Events & Event Signatures
- Machine Learning
- Graphs & GNNs
- GraphNeT & DynEdge

Current State of Affairs

My Work

Conclusions & Outlook

The IceCube Neutrino Observatory

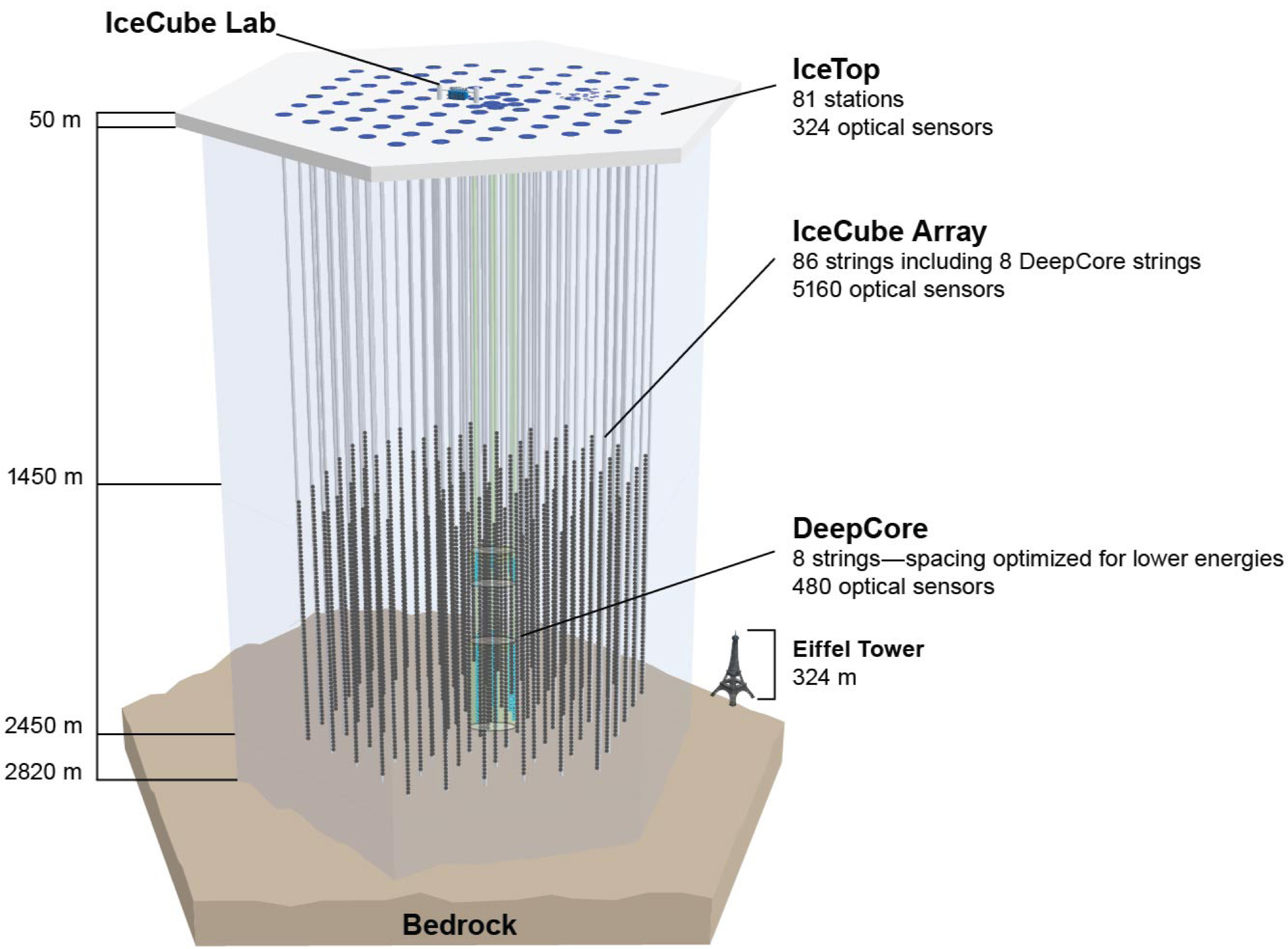
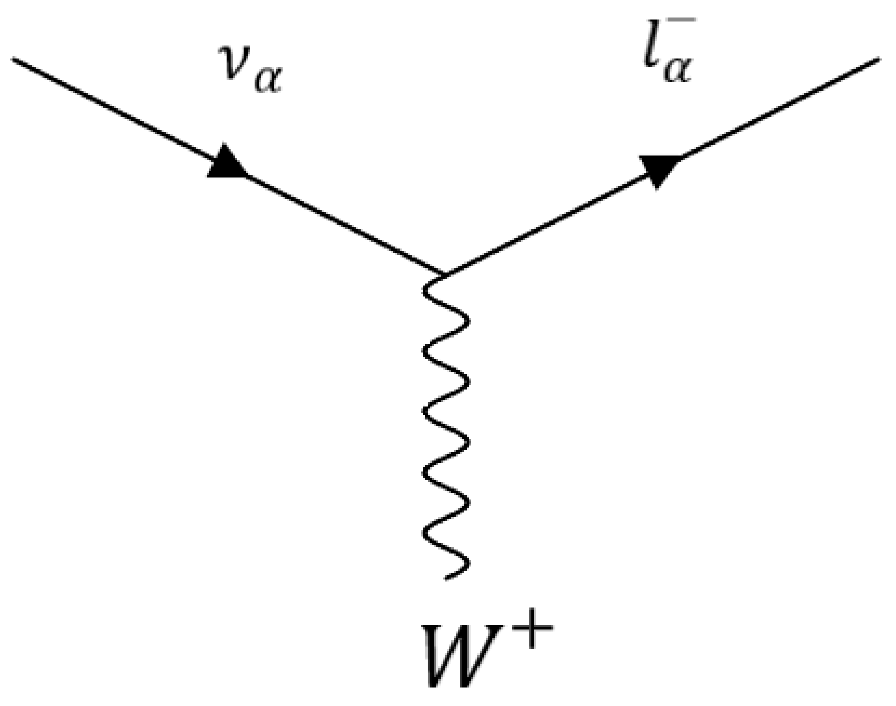


Figure from https://nbi.ku.dk/english/news/news19/upgrade-of-a-research-icecube/

IceCube Events

Neutrino interaction



Cherenkov Radiation

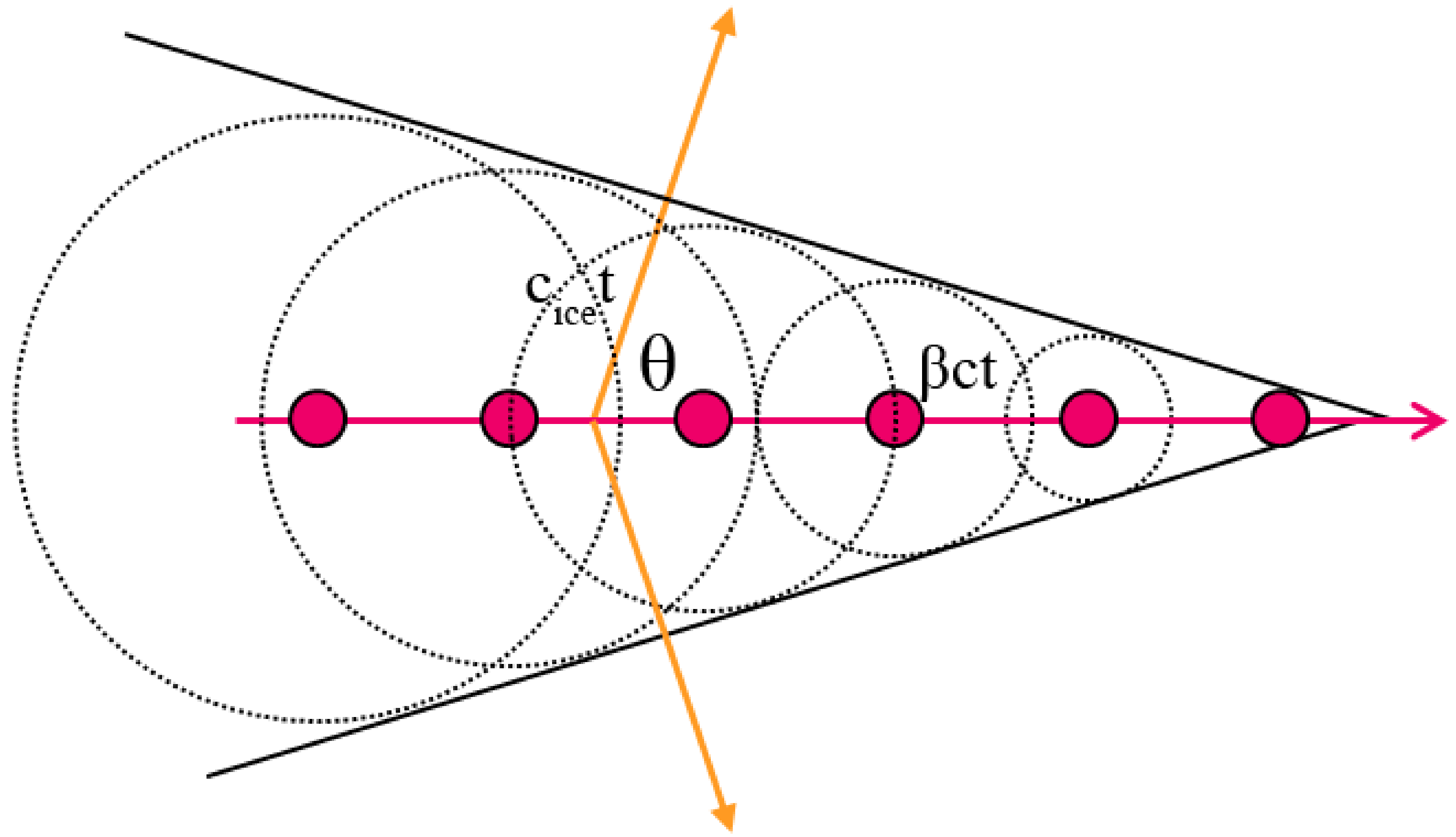


Figure from
http://large.stanford.edu/courses/2014/ph241/alaeian2/

DOMs

Figure from
https://commons.wikimedia.org/wiki/File:ICECUBE_dom_taklampa.jpg

Event signatures

- Different neutrino interactions lead to different event signatures

- Muons have sufficient mass and lifetime to leave a distinct track

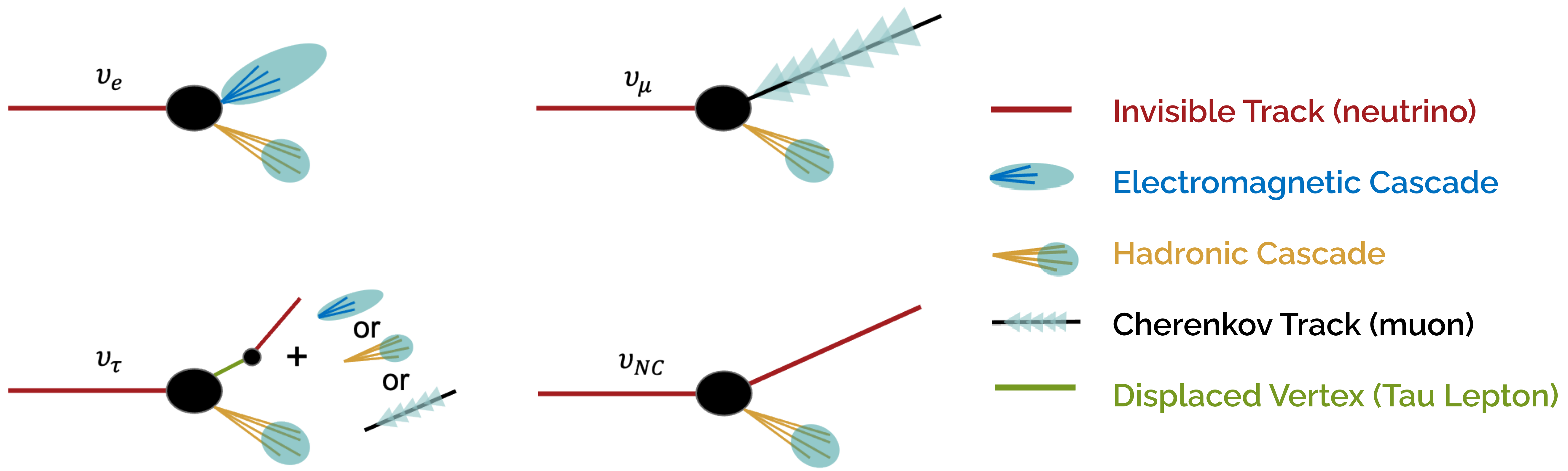


Figure from
Etienne Bourbeau, *Measurement of Tau Neutrino Appearance in 8 Years of IceCube Data*

Event signatures

- Different neutrino interactions lead to different event signatures
- Muons have sufficient mass and lifetime to leave a distinct track

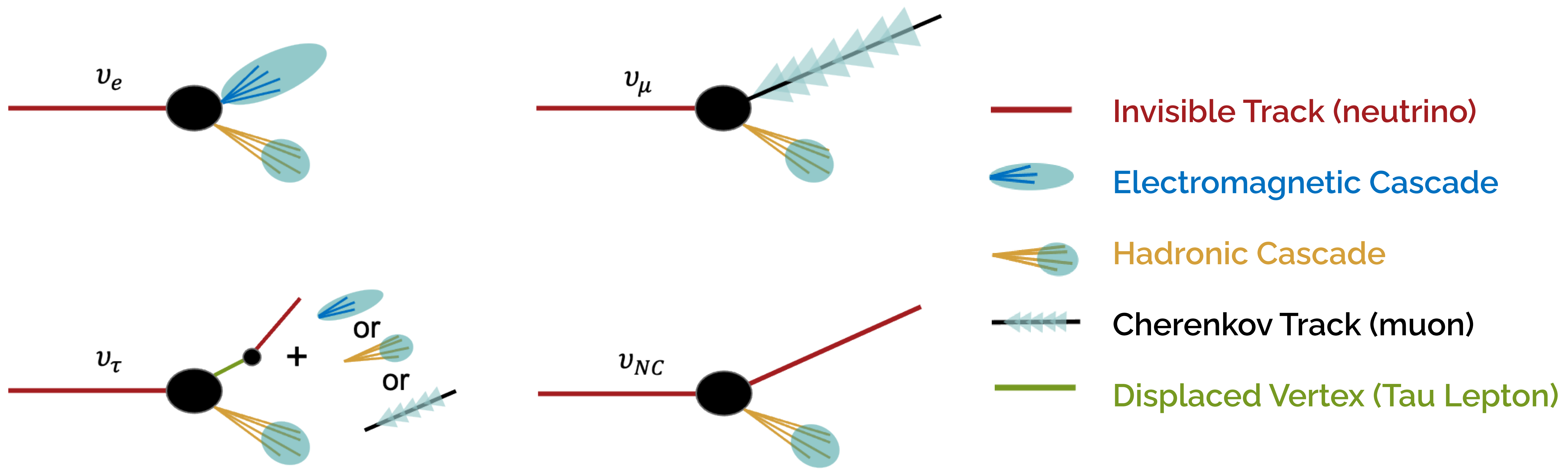


Figure from
Etienne Bourbeau, *Measurement of Tau Neutrino Appearance in 8 Years of IceCube Data*

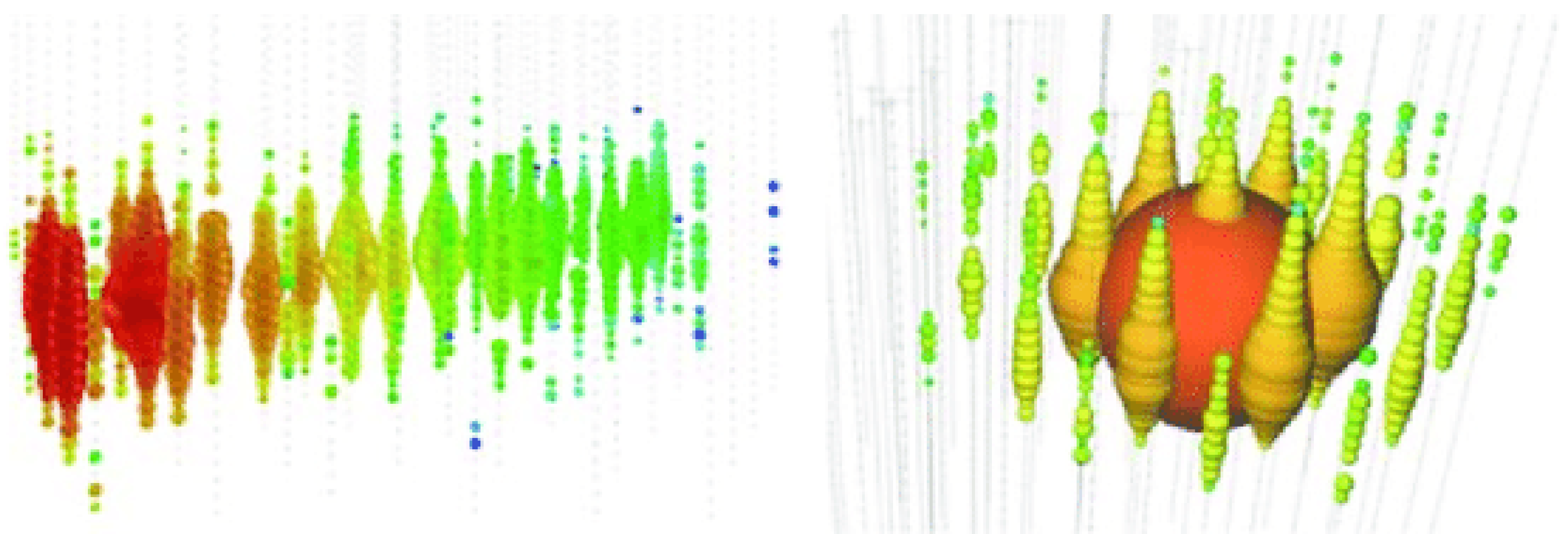


Figure from
Marek Kowalski et al., *Neutrino astronomy with IceCube and beyond*

Atmospheric Neutrinos

Charged particles from the Cosmos interact with atomic nuclei in the atmosphere

The products are mostly light unstable mesons, and predominantly **pions**

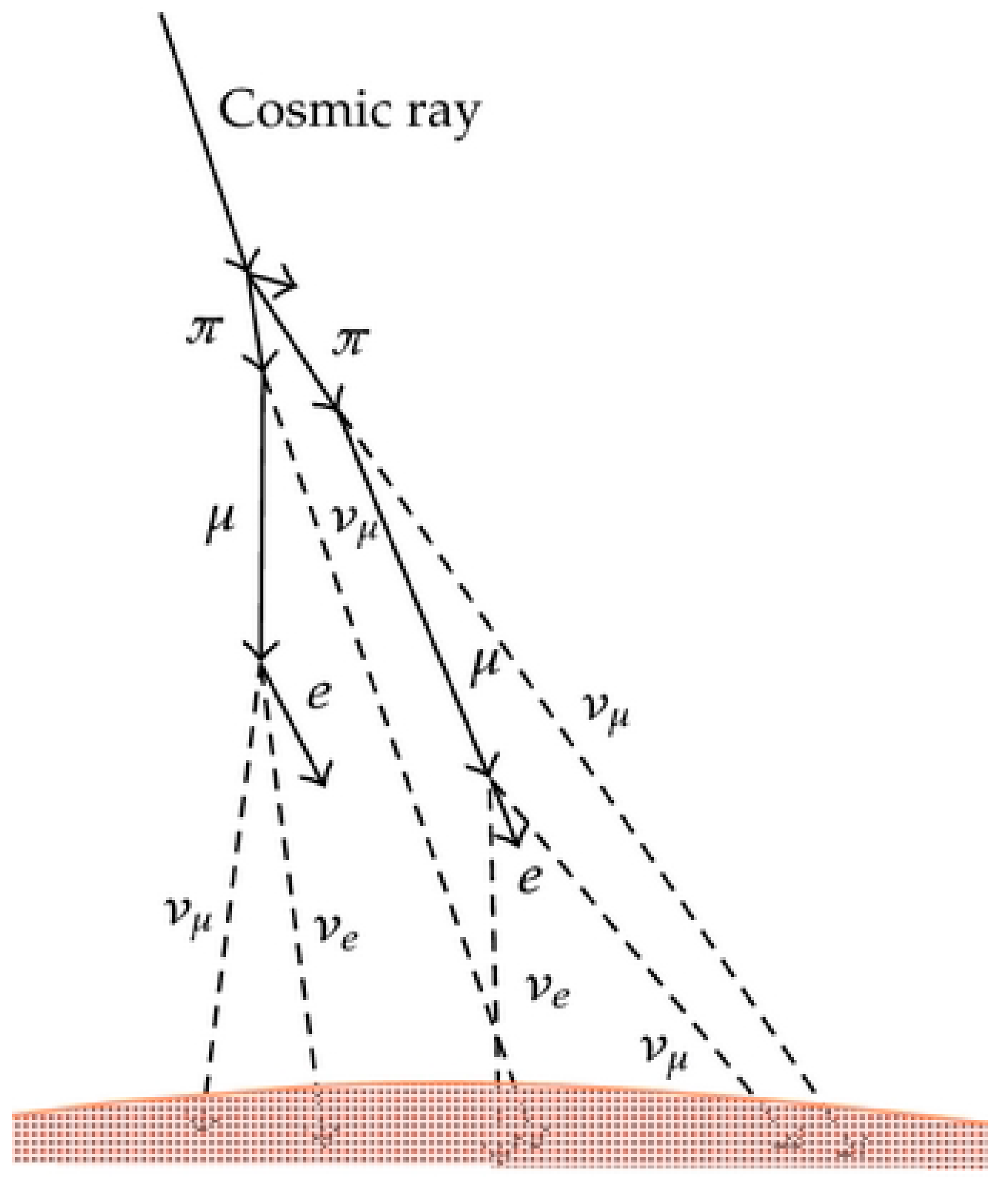


Figure from Takaaki Kajita, *Atmospheric Neutrinos*

Atmospheric Neutrinos

Charged particles from the Cosmos interact with atomic nuclei in the atmosphere

The products are mostly light unstable mesons, and predominantly **pions**

Charged pions decay to **muons** and **muon antineutrinos** or **antimuons** and **muon neutrinos**

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad \text{(99.99\% of the time)}$$
$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \quad \text{(99.99\% of the time)}$$

Muons decays produce **electron antineutrinos**

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \quad \text{(100\% of the time)}$$

Noise & Muons

- Muons decays produce **electron antineutrinos** $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu \quad \text{(100\% of the time)}$
- Many muons **do not decay** but make it to the detector

Noise & Muons

Most of the recorded events in IceCube are caused by muons or noise from the detectors

These need to be filtered out to perform an analysis of neutrino oscillations

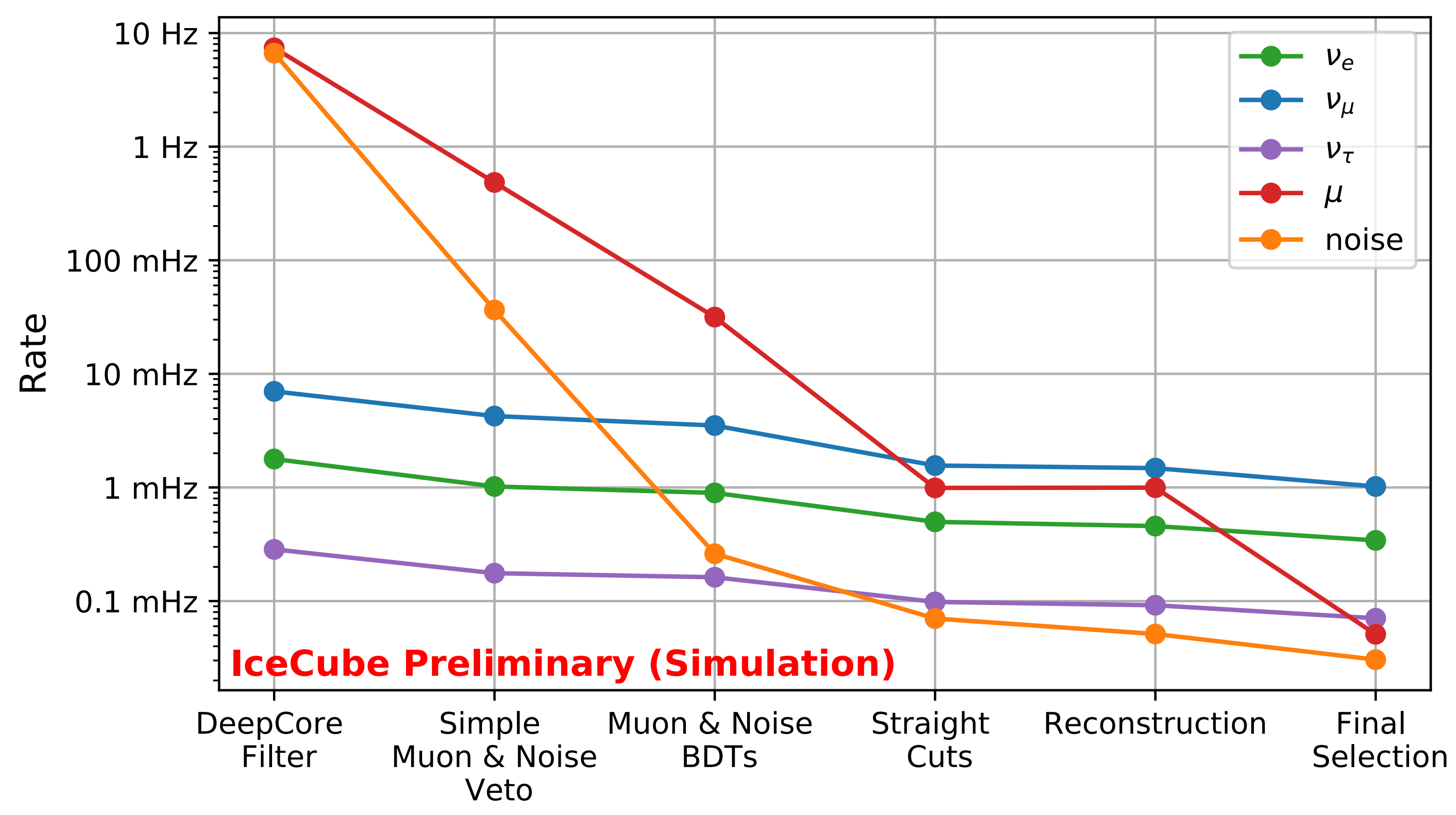
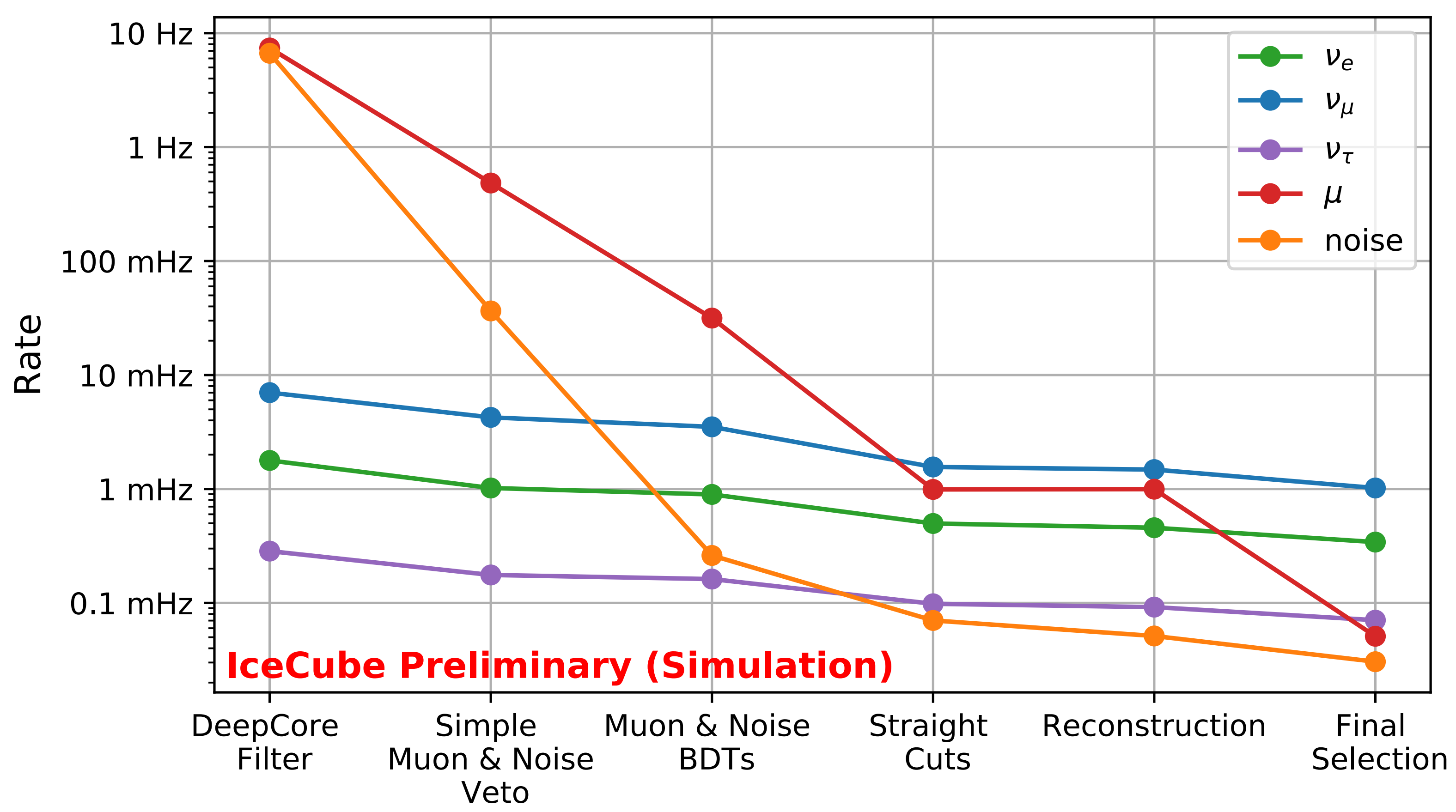


Figure from
Etienne Bourbeau, *Measurement of Tau Neutrino Appearance in 8 Years of IceCube Data*

Partial conclusion: Accurate reconstruction of neutrino events depends on good events selection and reliable reconstruction algorithms. Current methods are slow and inflexible.



Machine Learning

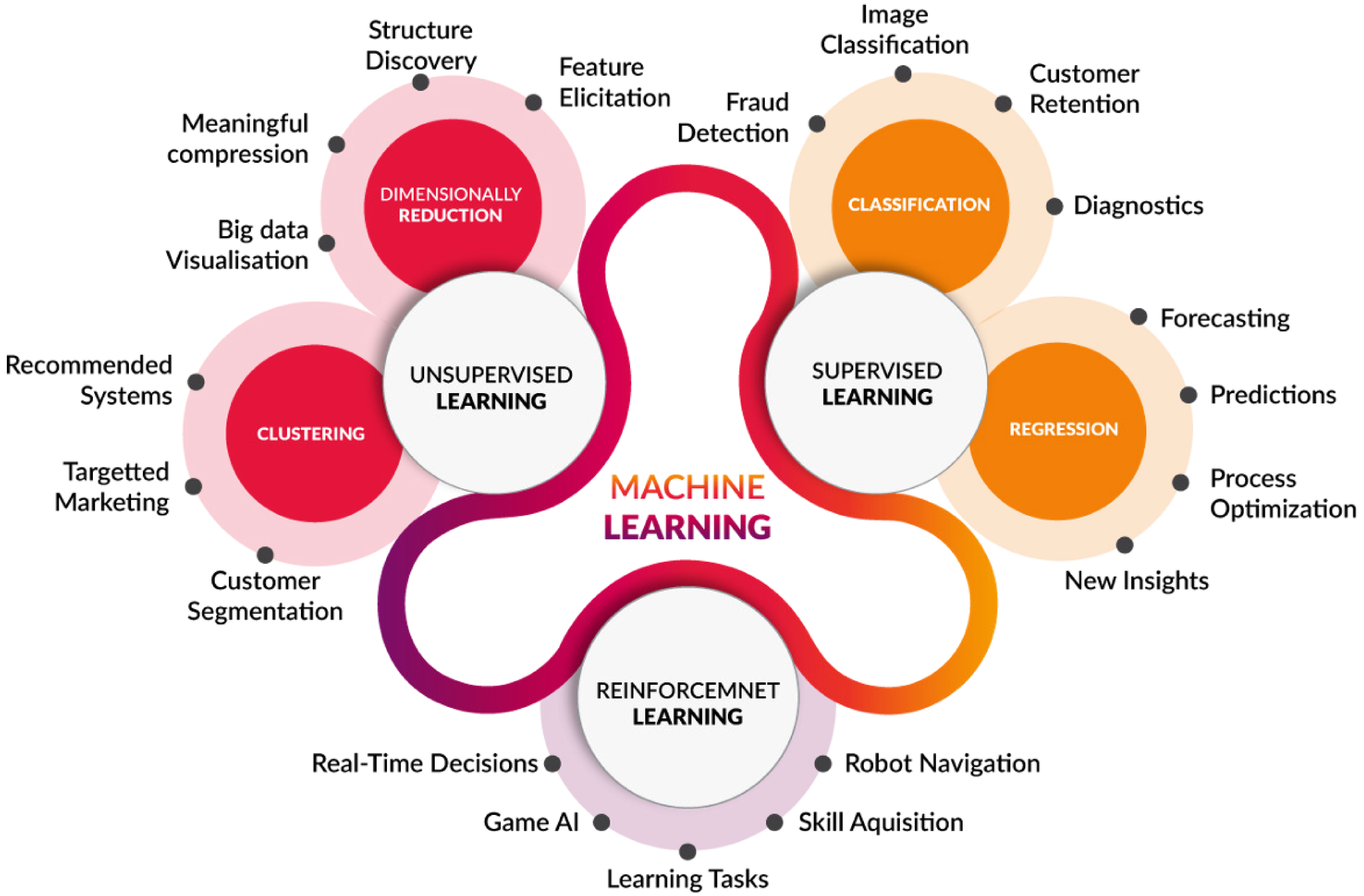


Figure from http://www.cognub.com/index.php/cognitive-platform/

Machine Learning

Machine Learning: Algorithms designed to let computers "learn" from examples without being explixitly programmed

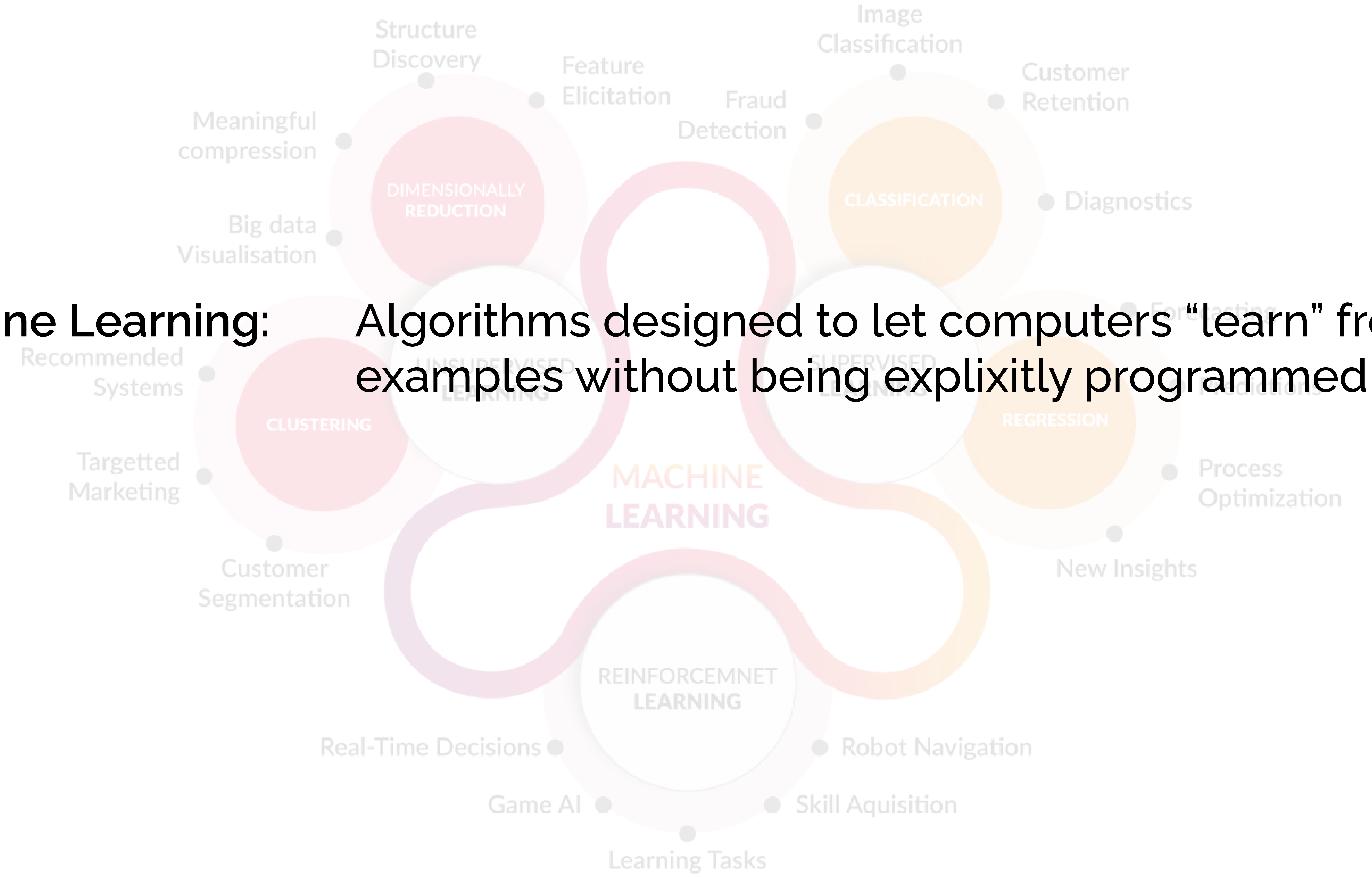
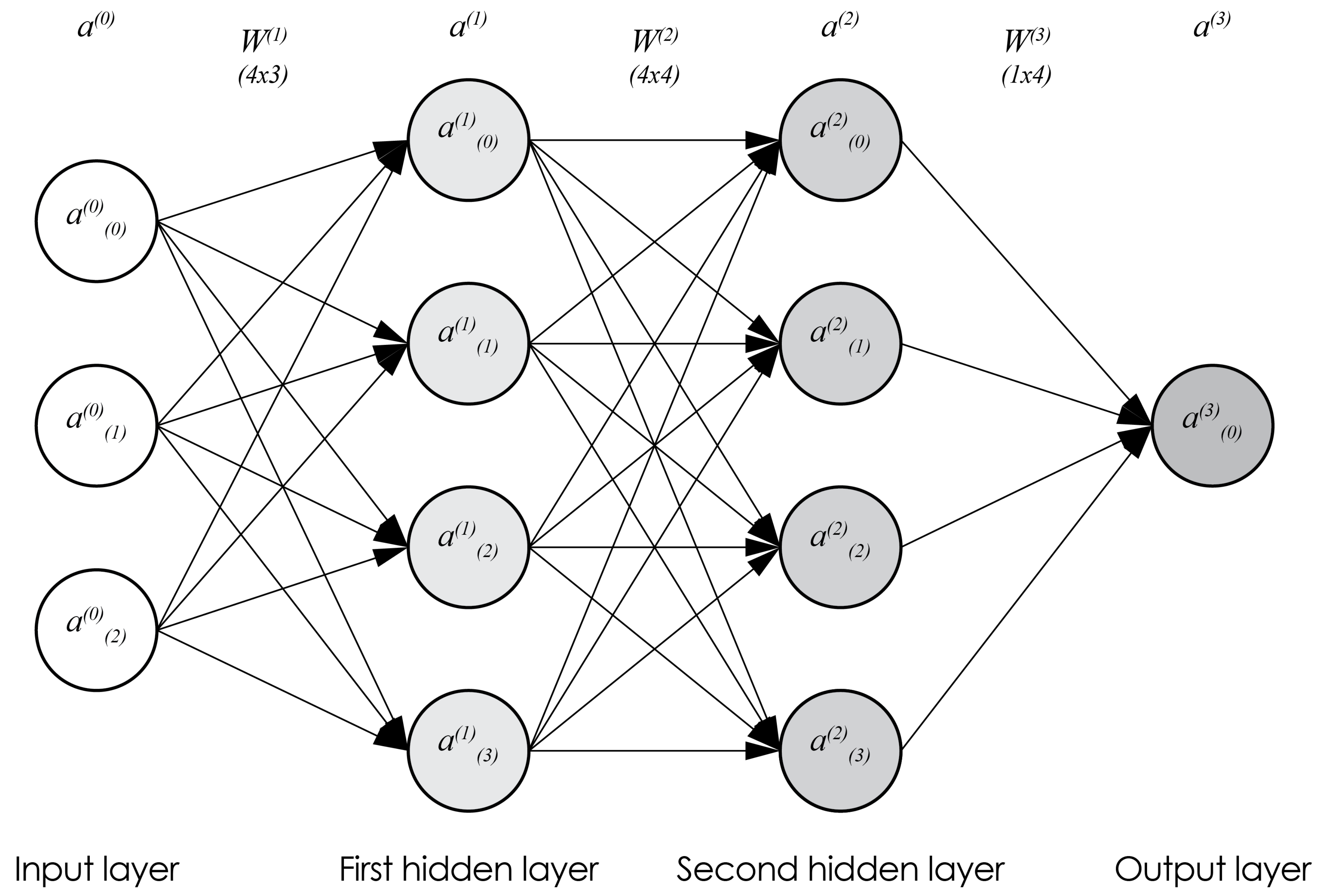


Figure from http://www.cognub.com/index.php/cognitive-platform/

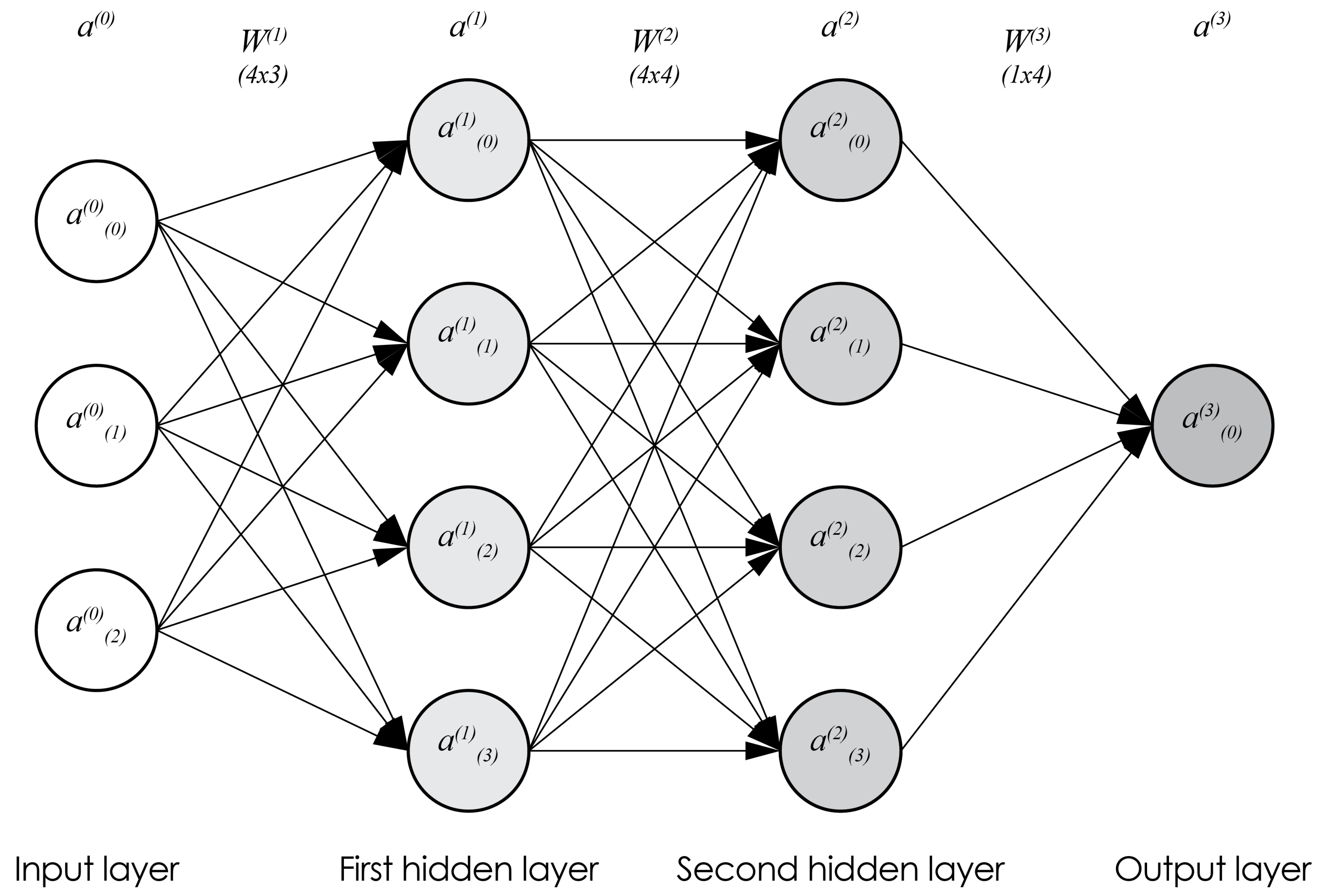
Neural Networks

- Neural networks: Machine learning models made up of **layers** of **nodes** connected by **weights**



Neural Networks

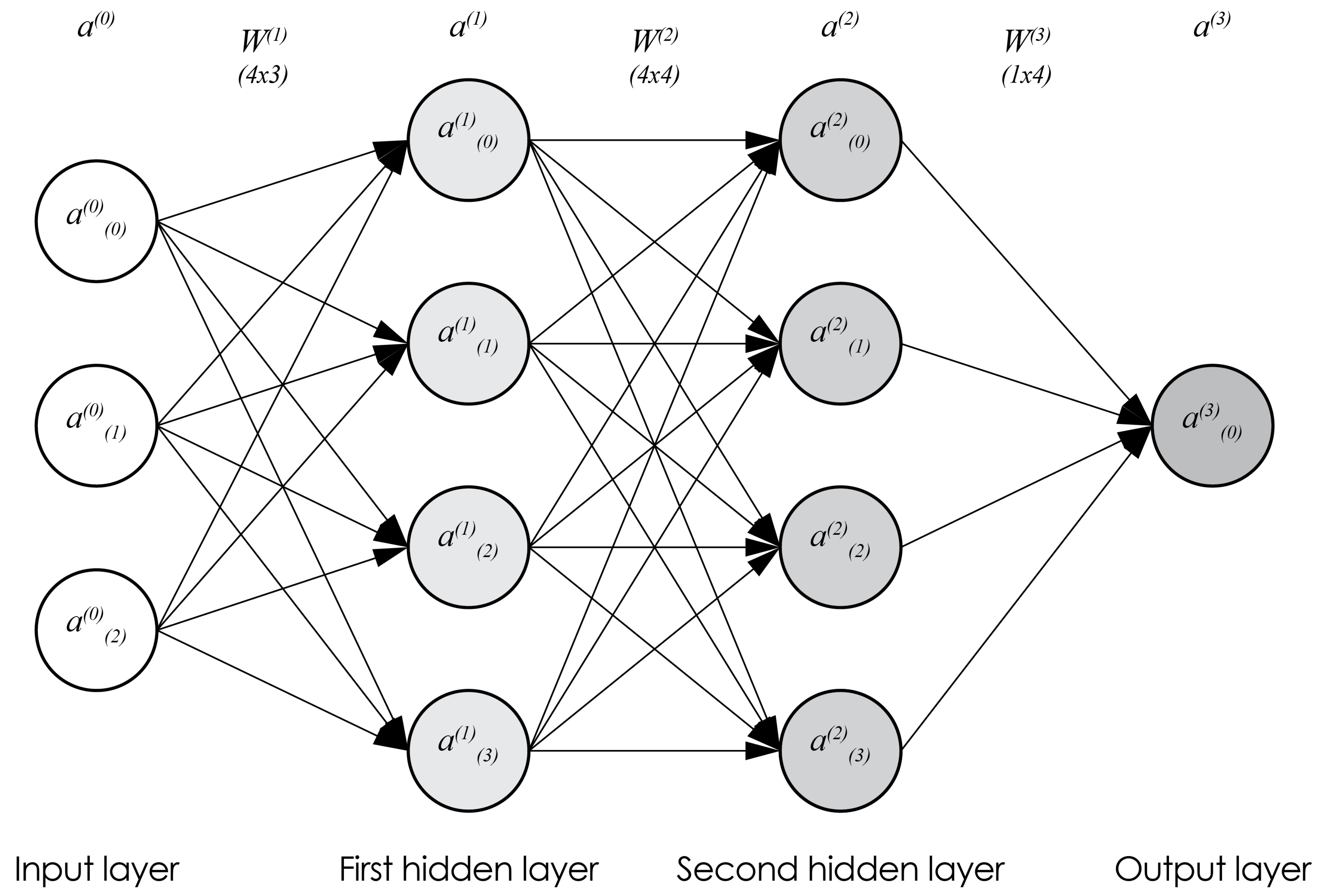
- Neural networks: Machine learning models made up of **layers** of **nodes** connected by **weights**



$$a^{(k)}_{(i)} = f\left(\sum_{j=0}^{N-1}\left(w^{(k)}_{(i,j)} \cdot a^{(k-1)}_{(j)}\right) + b^{(k)}_{(i)}\right)$$

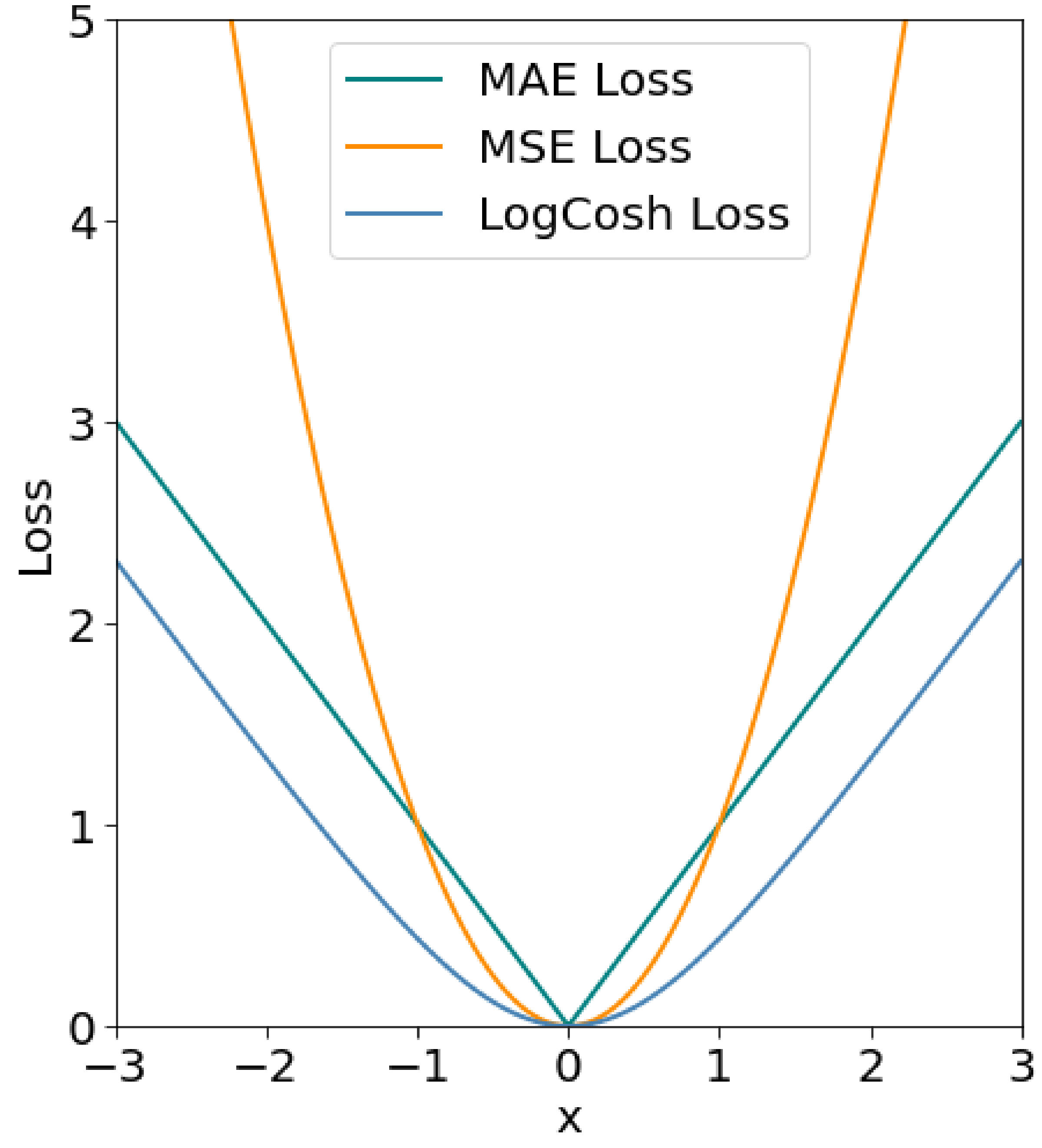
- Weights (and biases) are the learnable parameters of the neural network

Loss



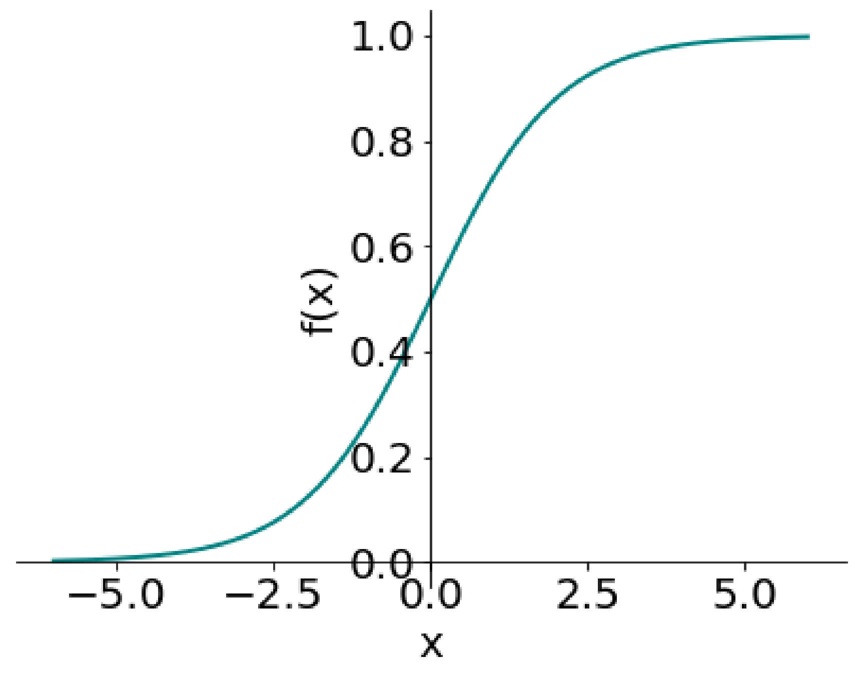
Learning is carried out by optimising the weights and biases to minimise the **loss**

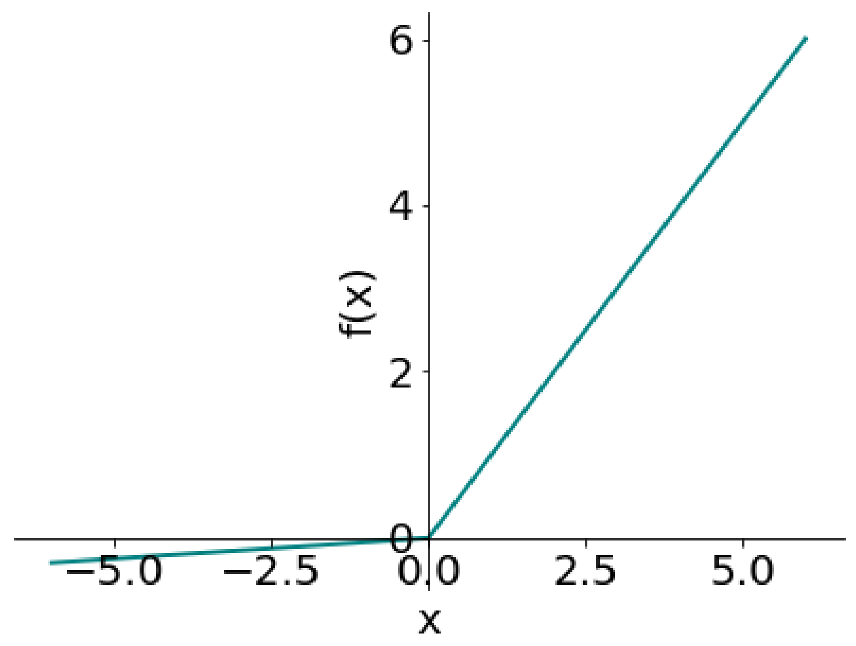
Loss

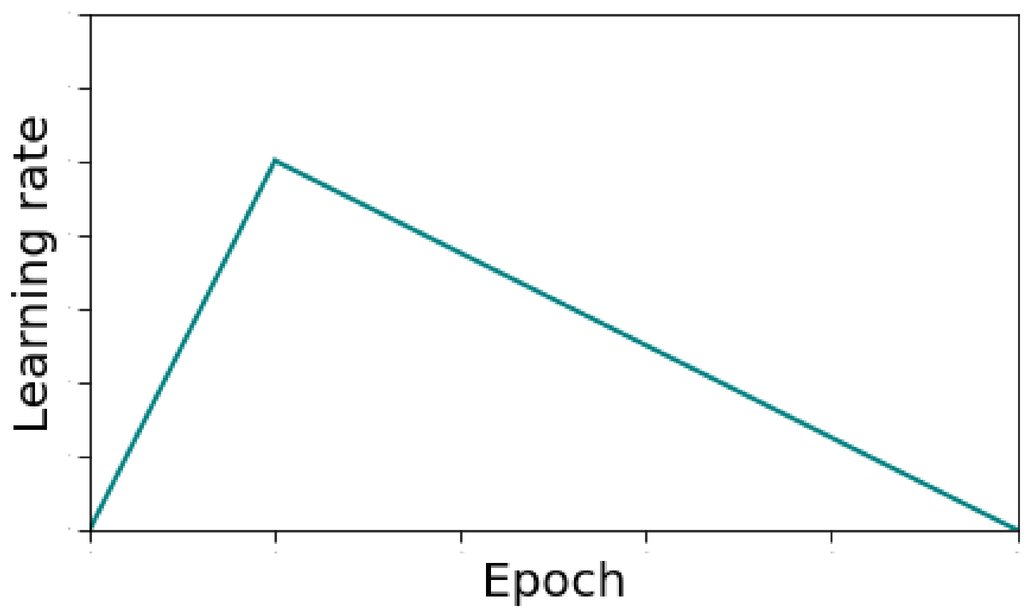


- Learning is carried out by optimising the weights and biases to minimise the **loss**
- The loss function differs between tasks

- Activation functions, optimisers, train/validation split, learning rates, early stopping

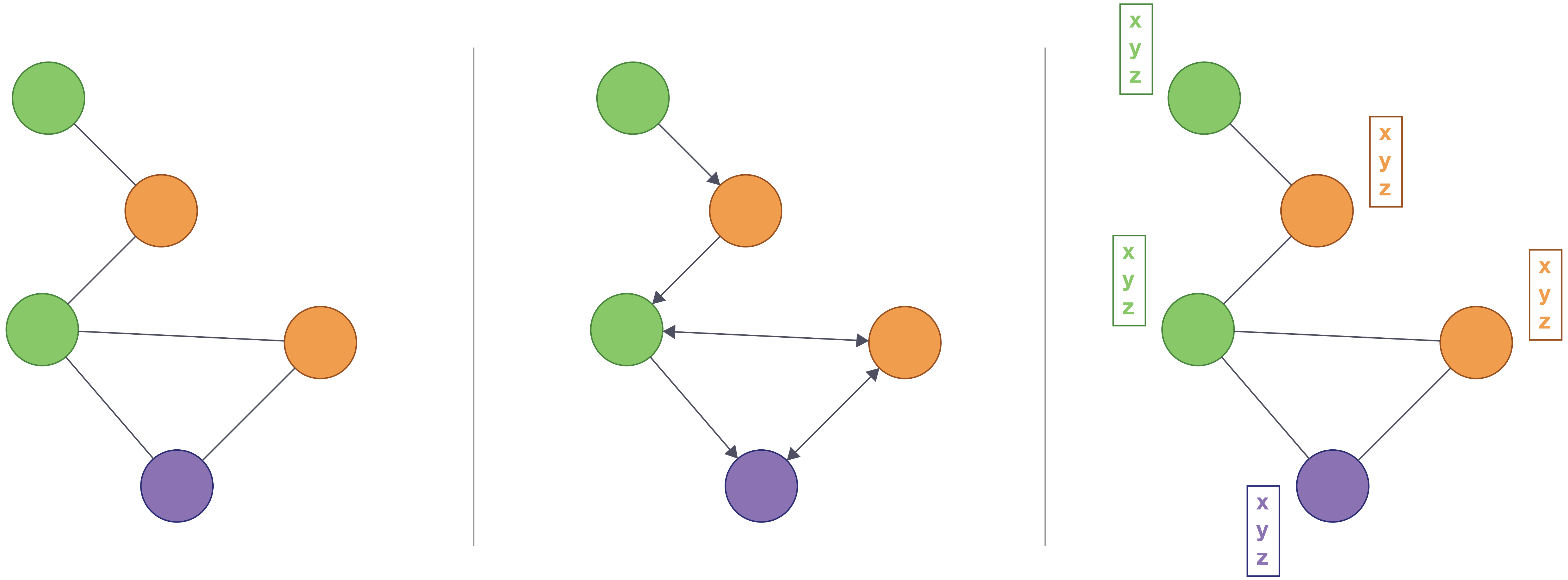






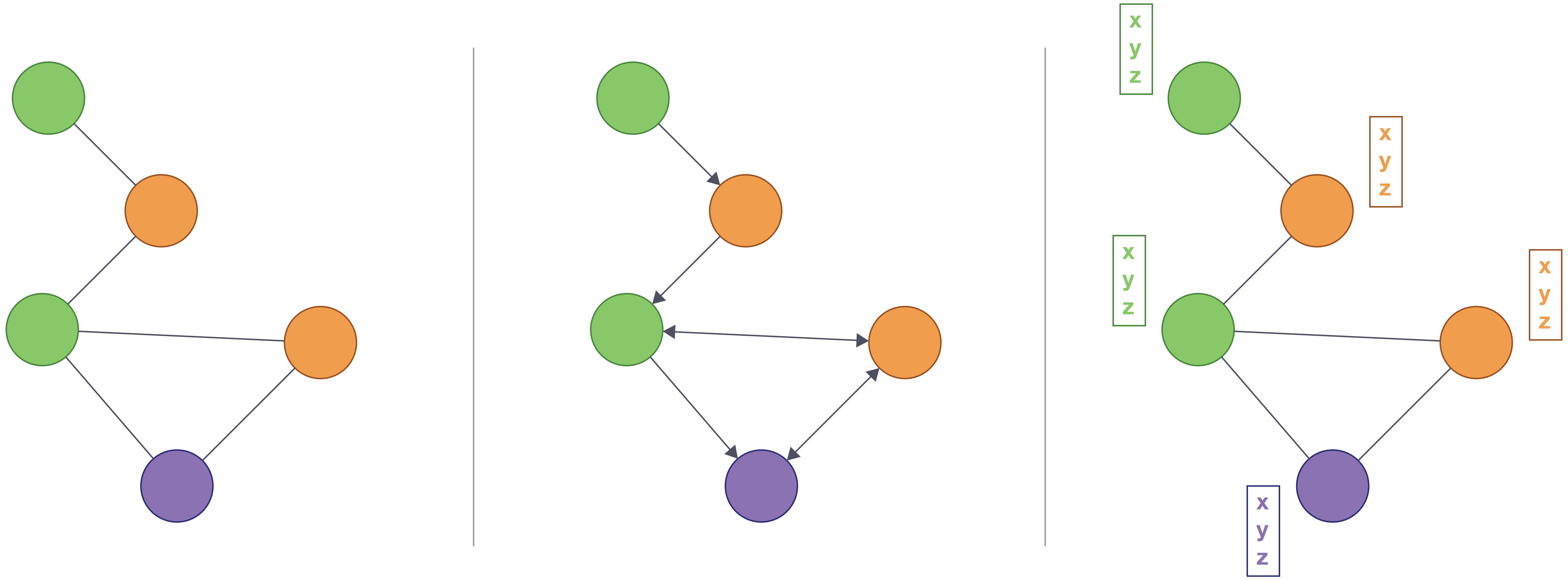
Graph Neural Networks

- Based on mathematical theory of graphs, consisting of nodes and edges
- Well suited for data the lacks a rigid grid-like structure, but still contains a hierarchy



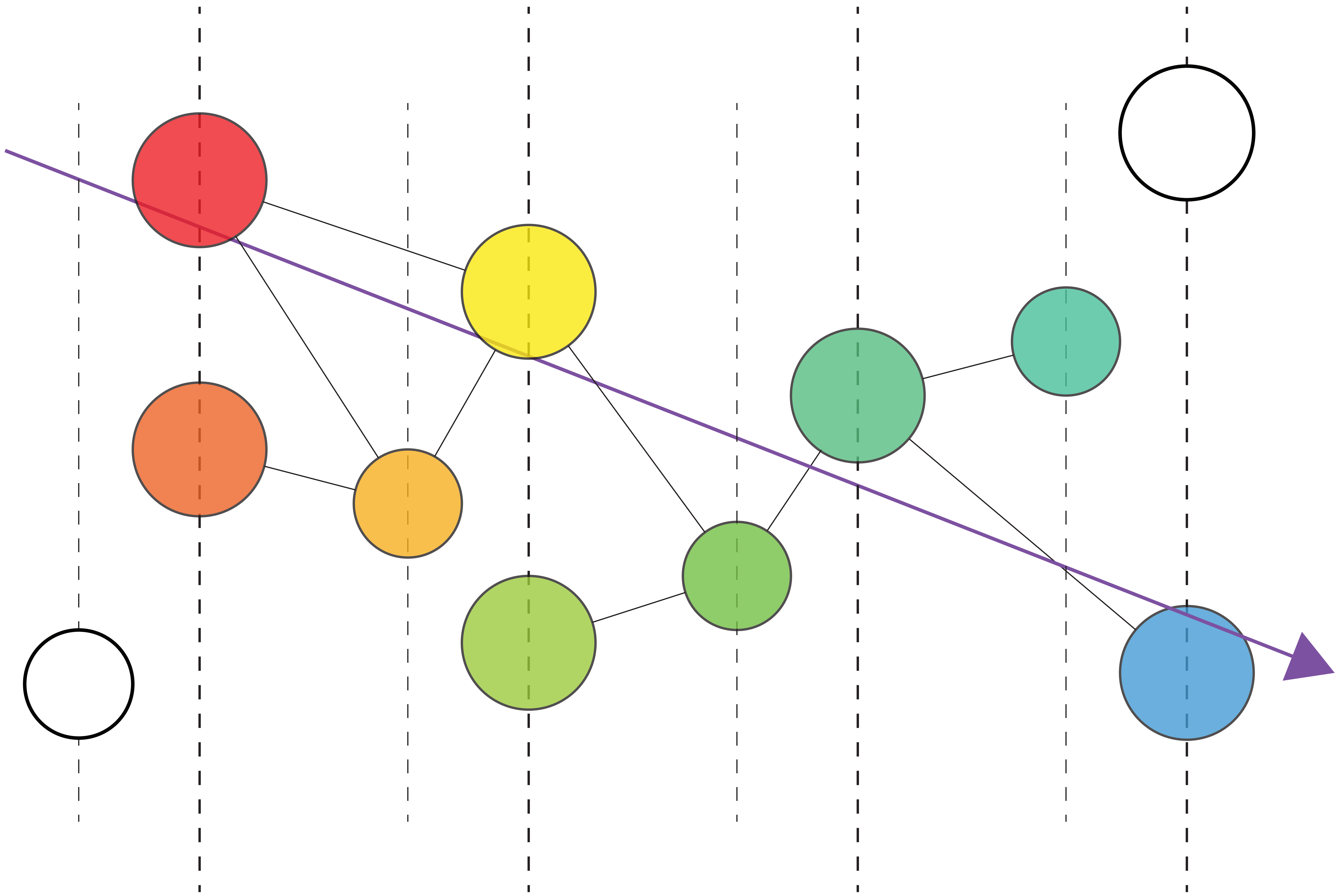
Graph Neural Networks

- Uses message passing schemes between each layer
- The number of node- and edge-features may change but the number of nodes stay the same



IceCube Events as Graphs

- KNN method is used to create graph from pulses in an event



GraphNeT & DynEdge

GraphNeT
Graph Neural Networks for
Neutrino Telescope Event Reconstruction

- GraphNeT is the combined GNN efforts within IceCube
- DynEdge if the common GNN architecture used within GraphNeT and in this work

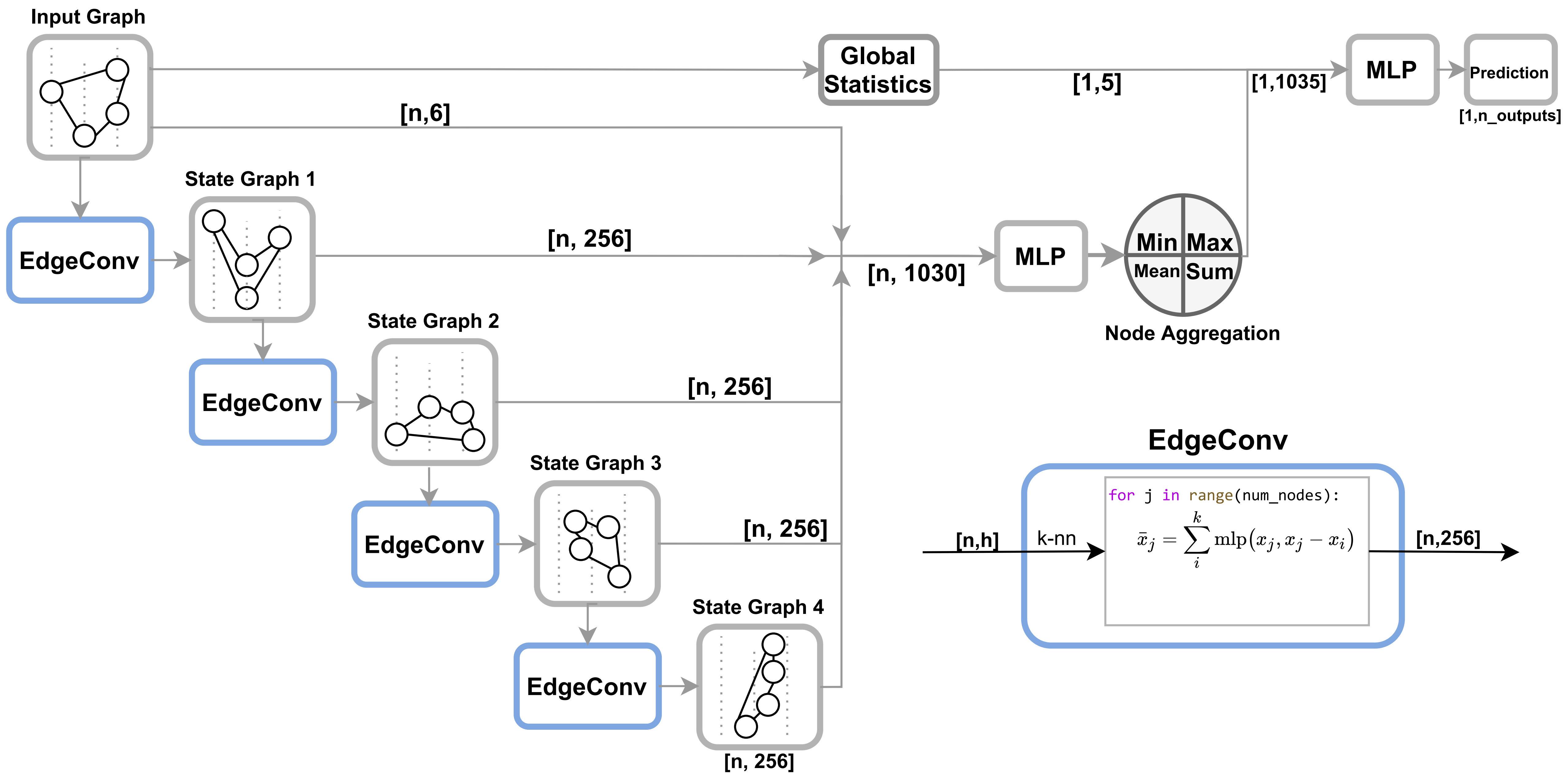


Figure by Rasmus Ørsøe

Partial conclusion: GNNs have already been shown to work well on IceCube regression and classification tasks. But what range of problems are they suited for and what do we stand to gain from it?

Introduction

My Work

Time Regression

Event Cleaning

Upgrade Pulse Cleaning

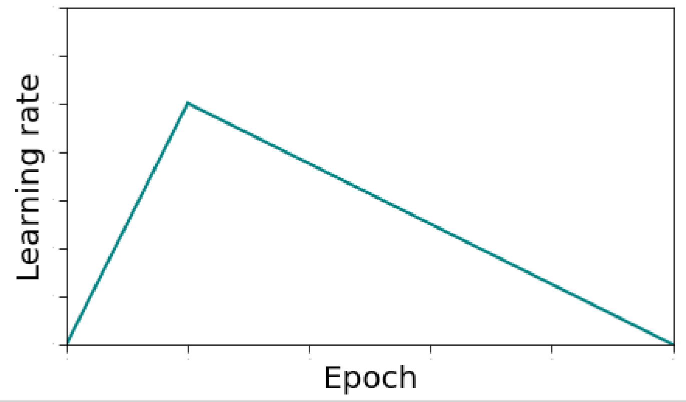
Regression of Cleaned Pulses

Intermezzo 1: Improvement of Pulse Cleaning

Intermezzo 2: Quality of Cleaned Events

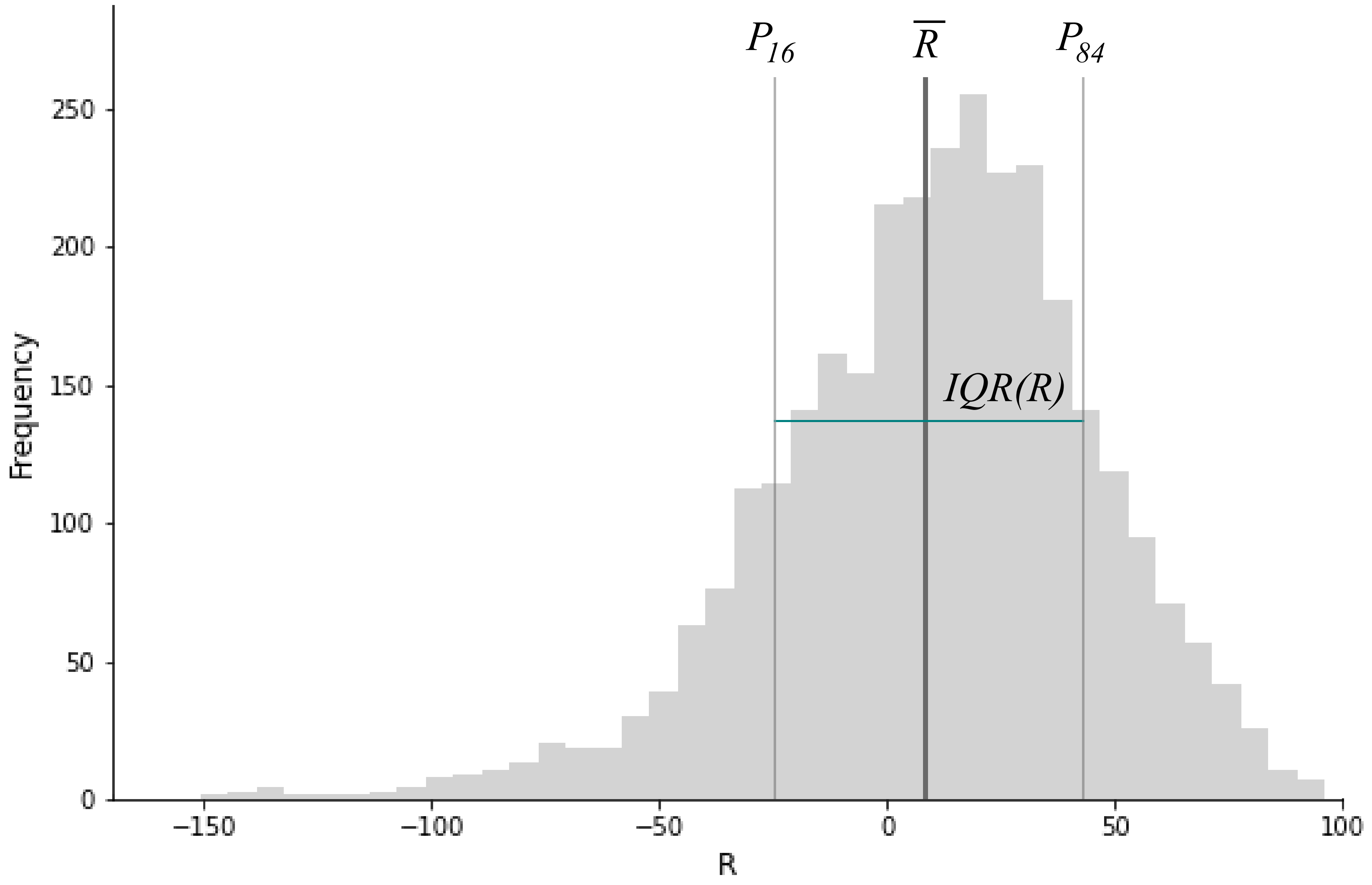
Conclusions & Outlook

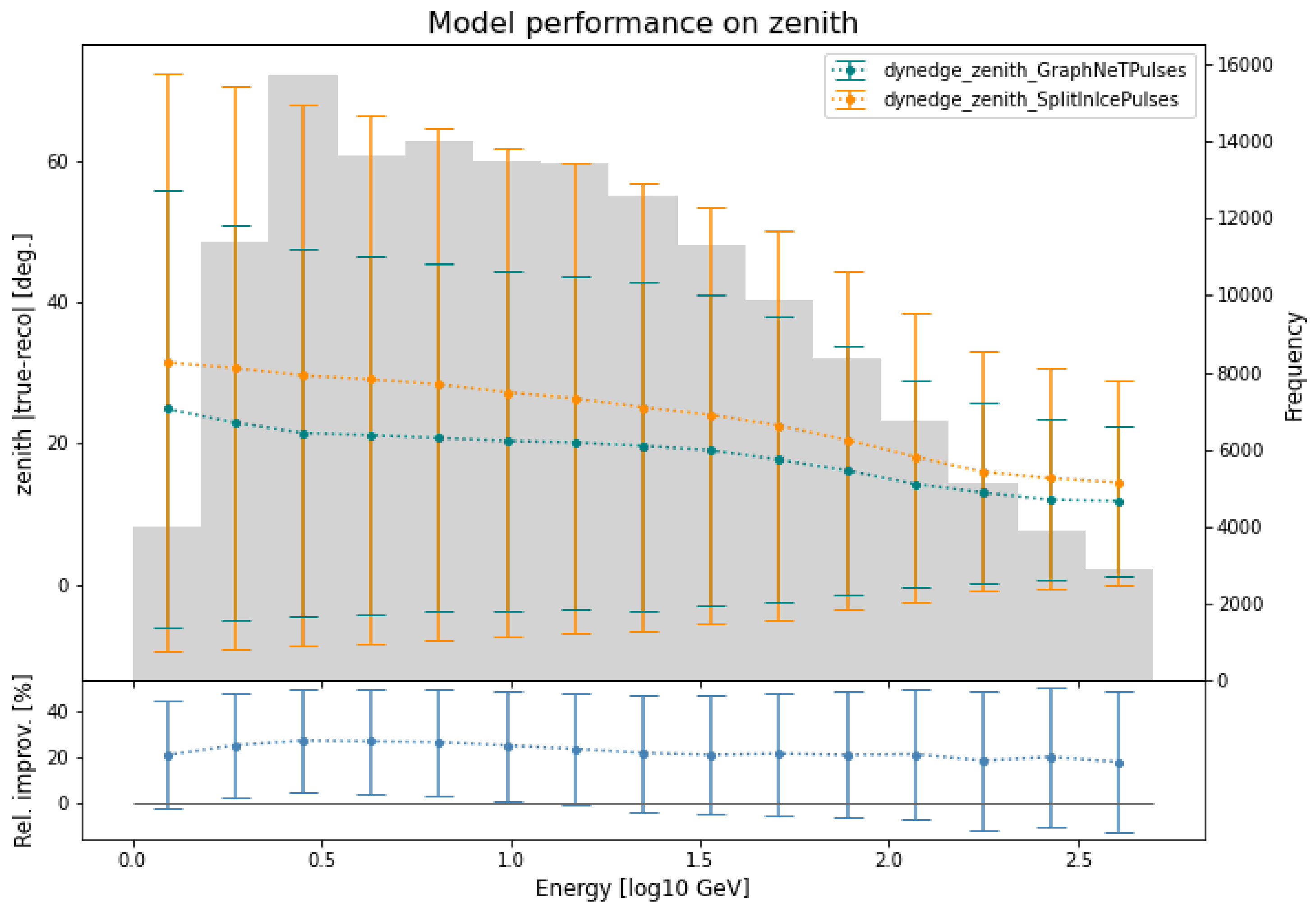
Preface

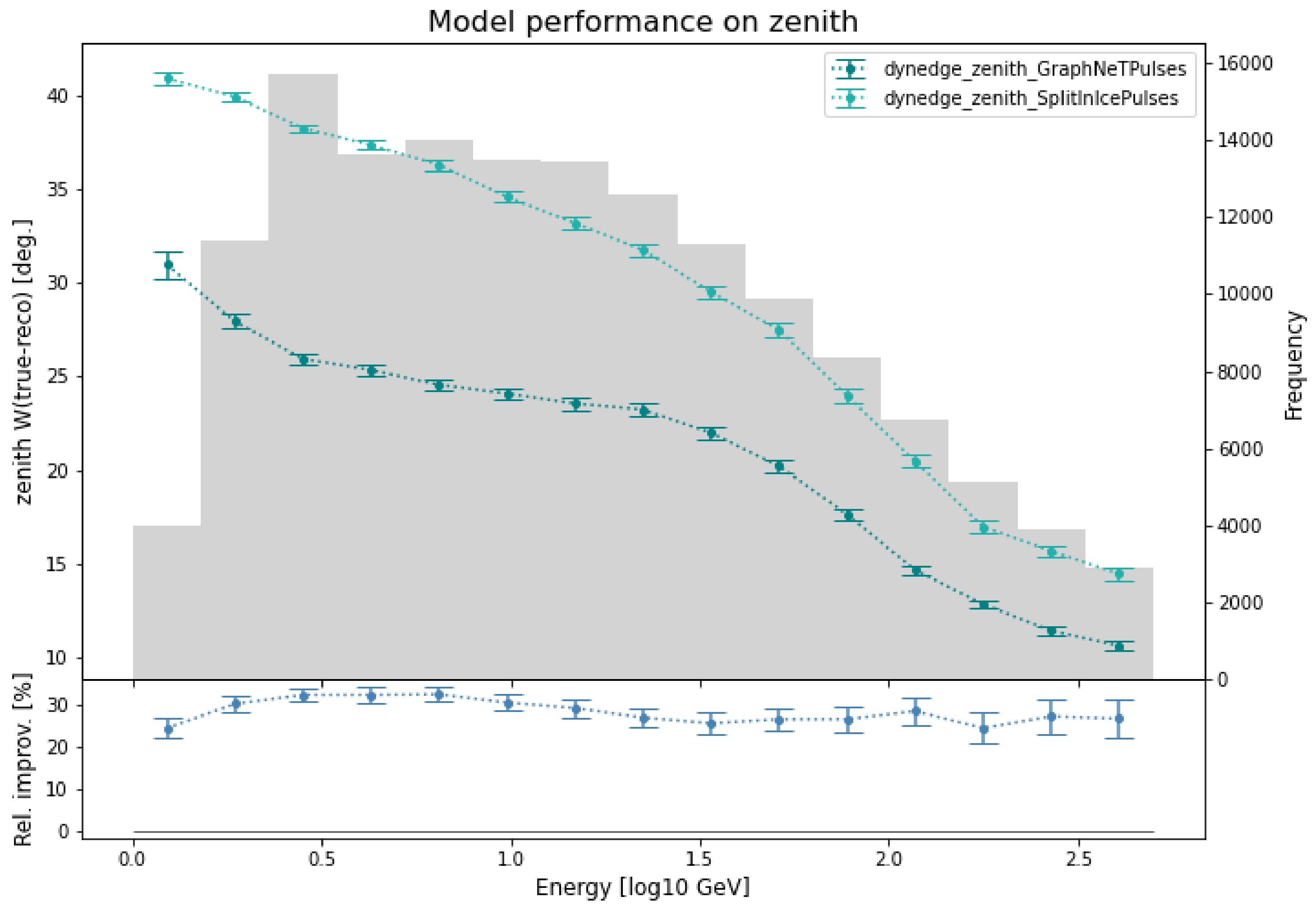
Loss functions:	Classification: Binary Cross Entropy Angular regression: 2D Von Mises-Fisher	Other regressions: LogCosh
Optimiser:	ADAM	
Learning rate:	Slanted triangular LR schedule with early stopping	
Sample distribution:	Even when possible	
Comparison:	RetroReco	
Training data:	IceCube MC simulations	

Preface

Plot types



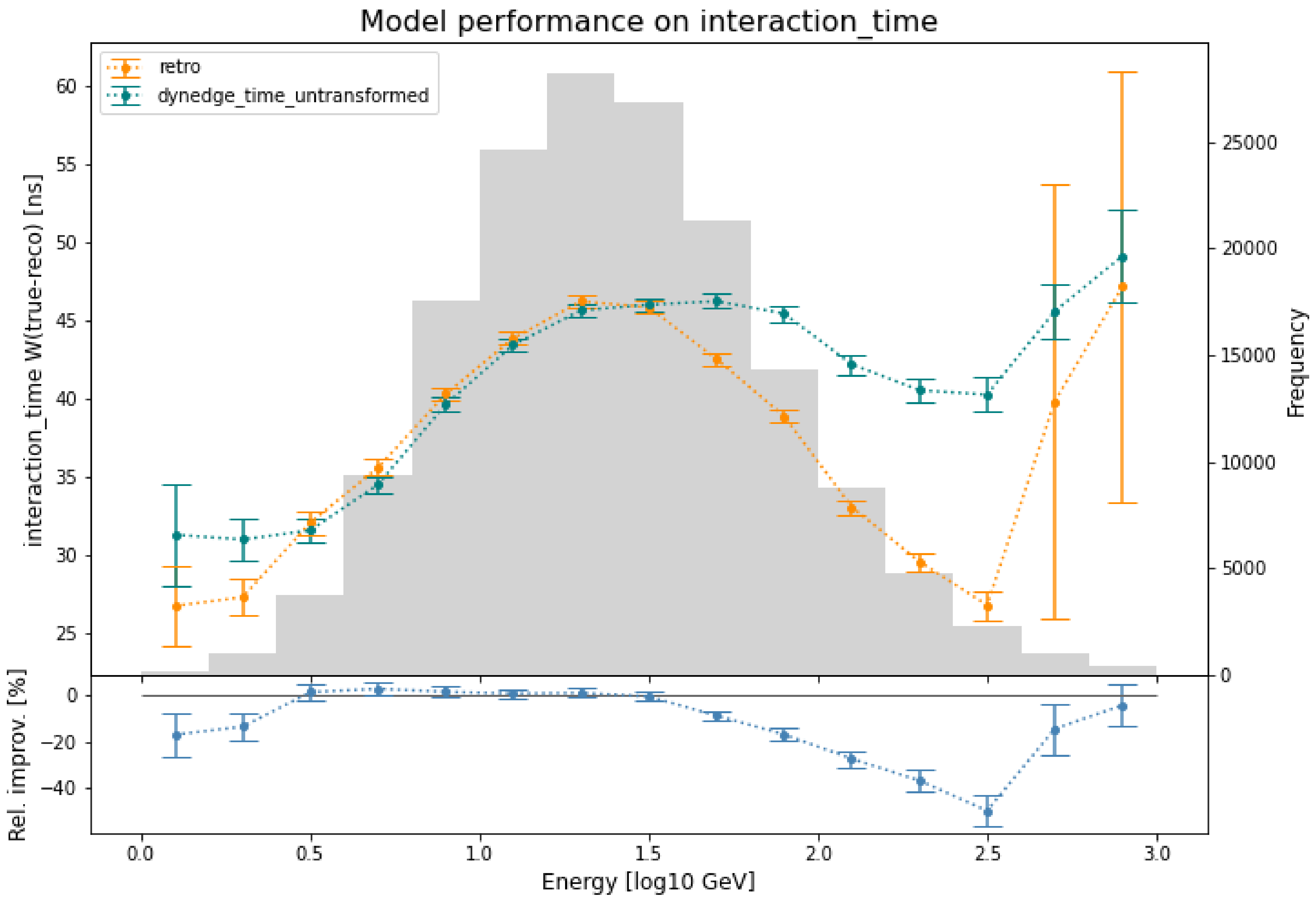


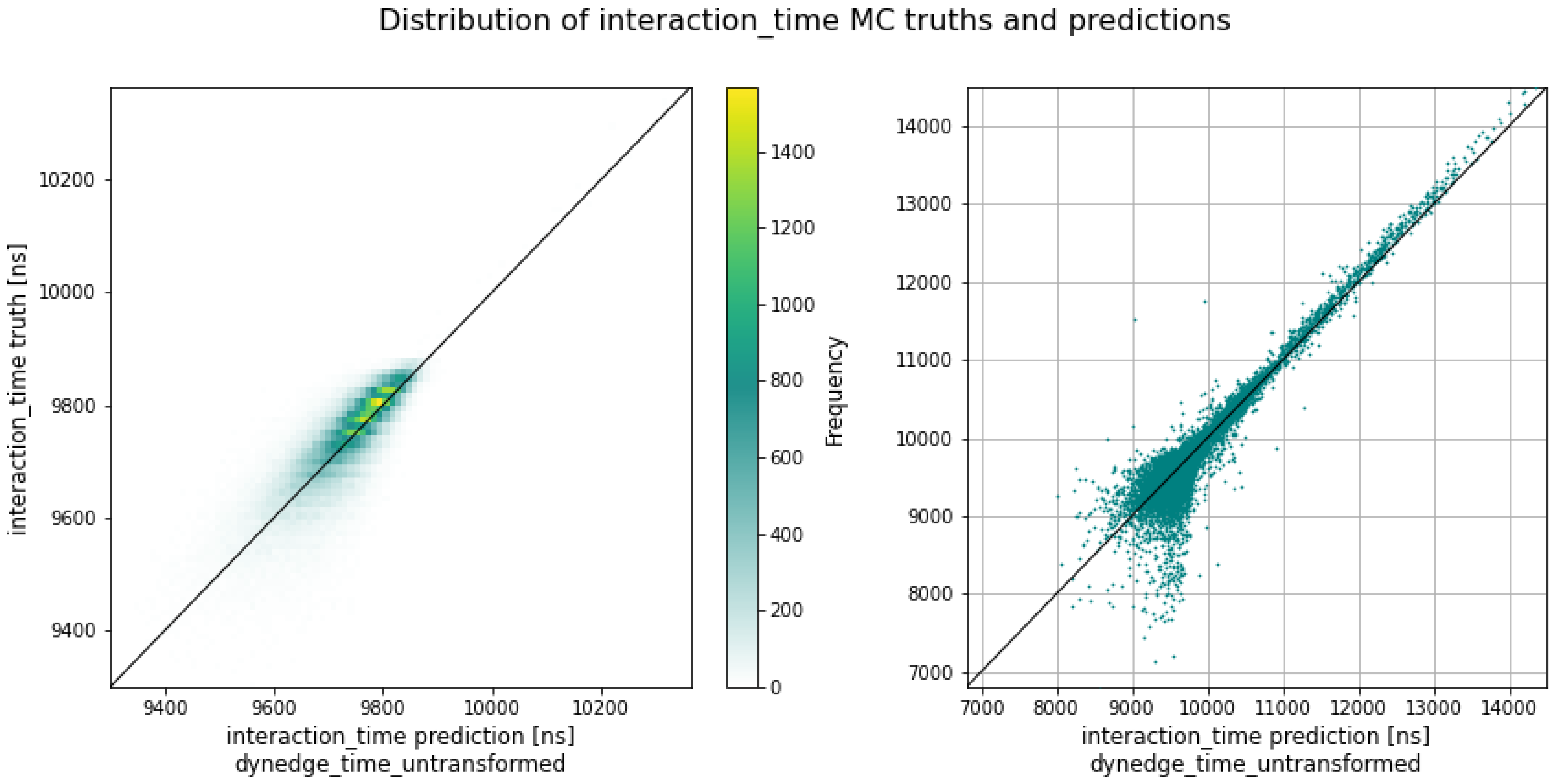


- R = truth - reco (most often)
- W(R) = IQR(R)/1.349

Interaction Time Regression

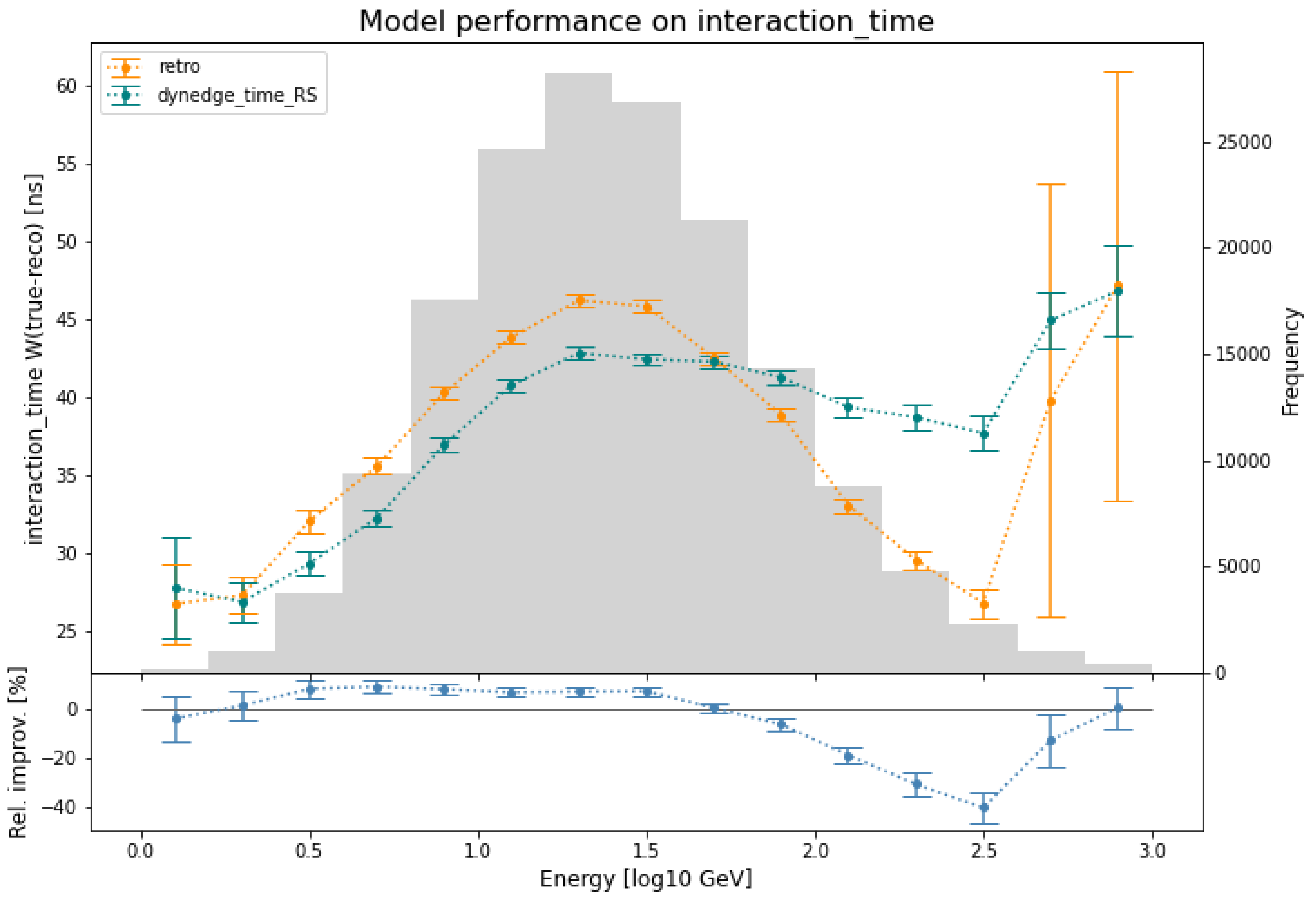
Setup: Regression of the time of neutrino interaction can be useful for some studies, but the GNN reconstruction struggles with this task.
The problem is likely the uneven distribution of the target variable.

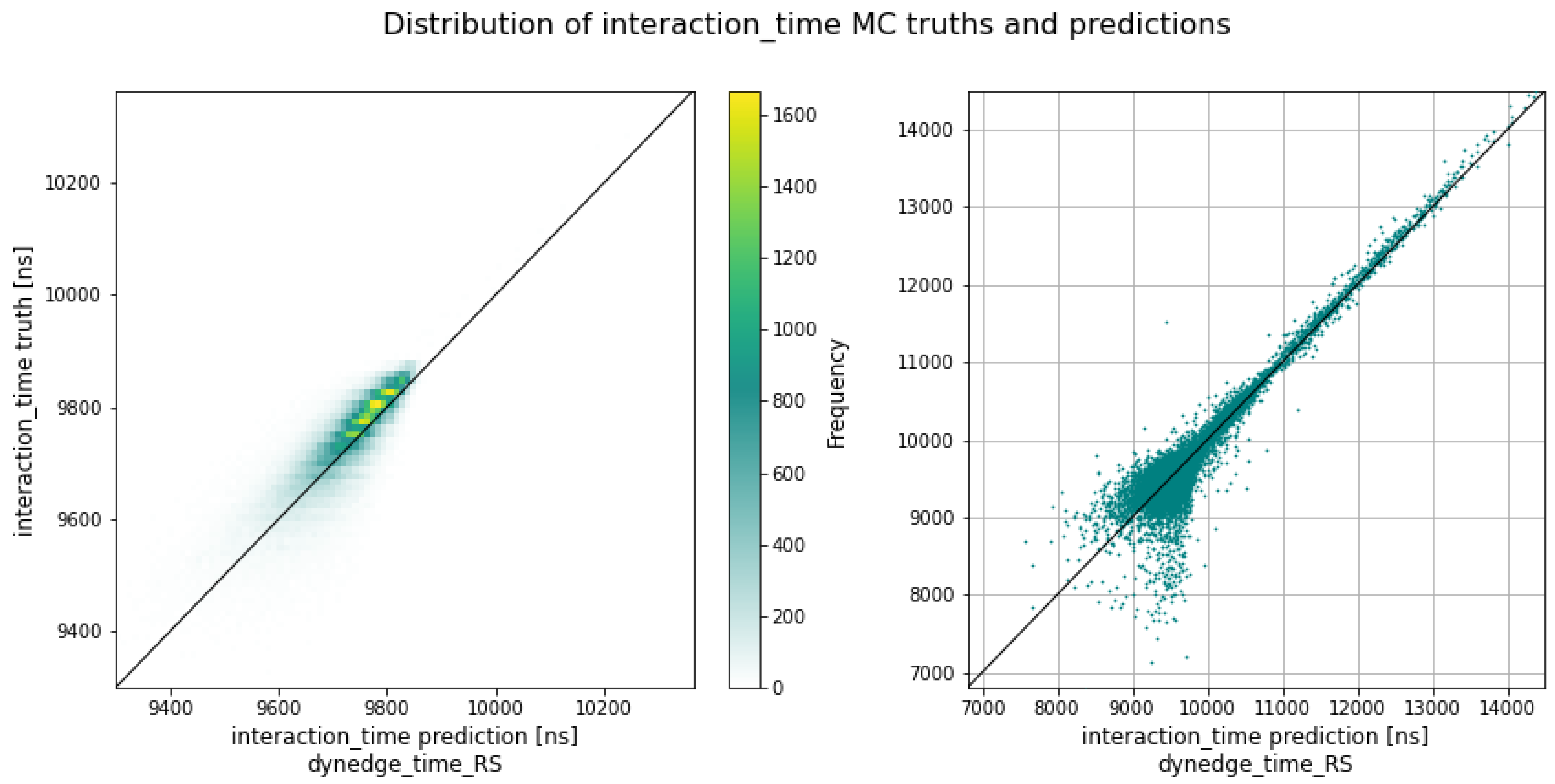




Interaction Time Regression

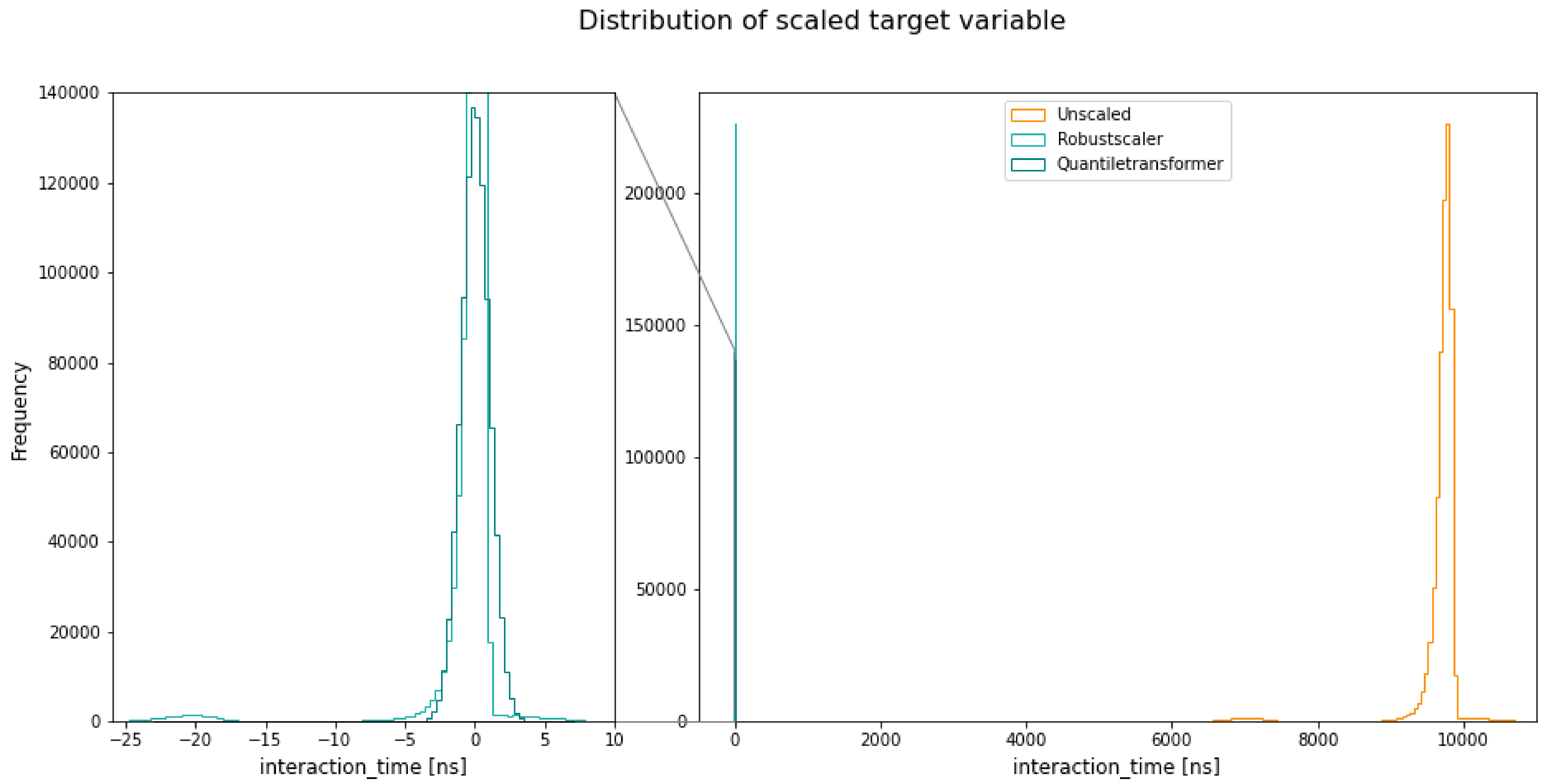
Approach: Apply scaling to the target variable. Attempt 1: **RobustScaler**





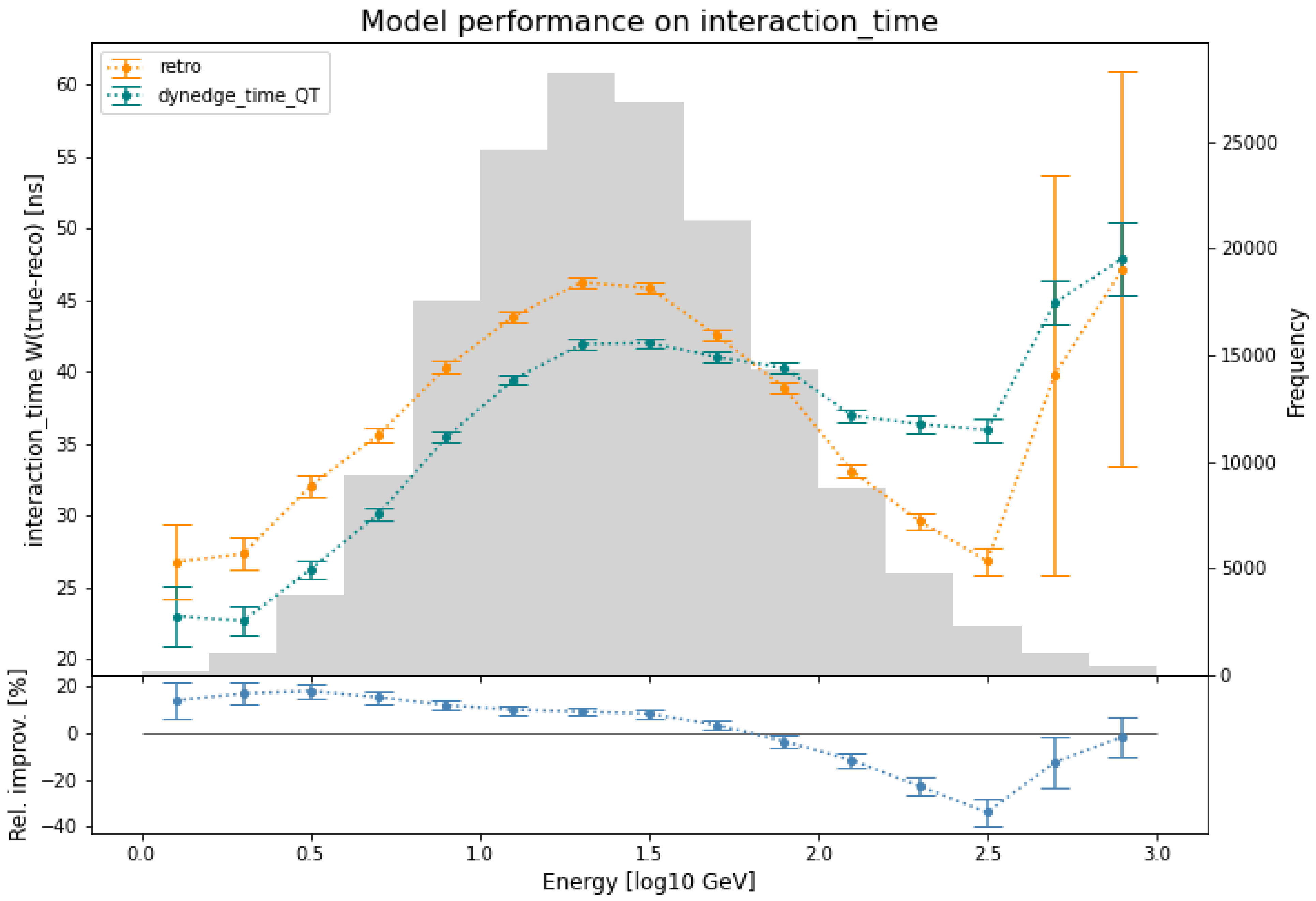
Interaction Time Regression

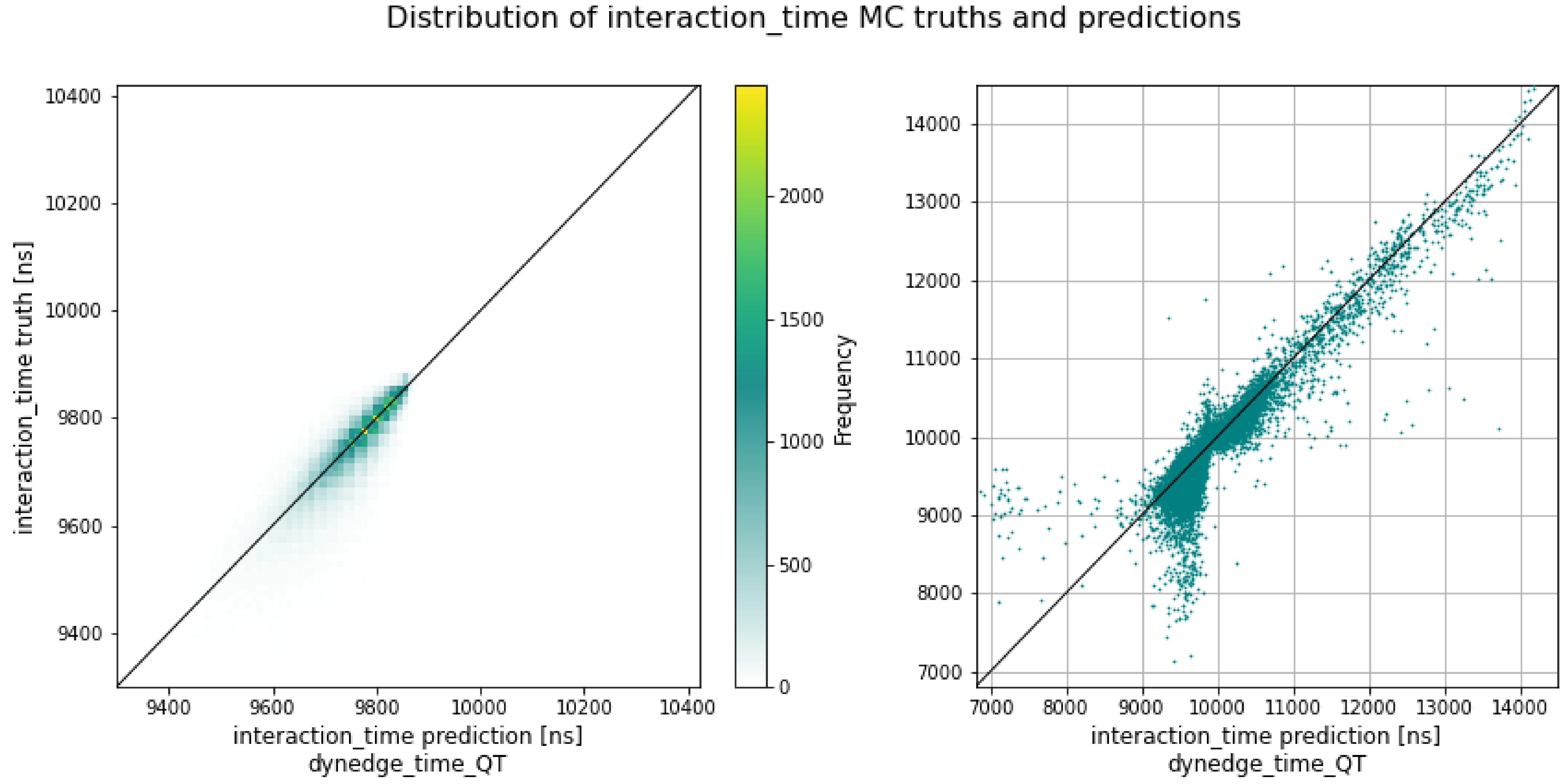
Approach: Apply scaling to the target variable. Attempt 2: **QuantileTransformer**



Interaction Time Regression

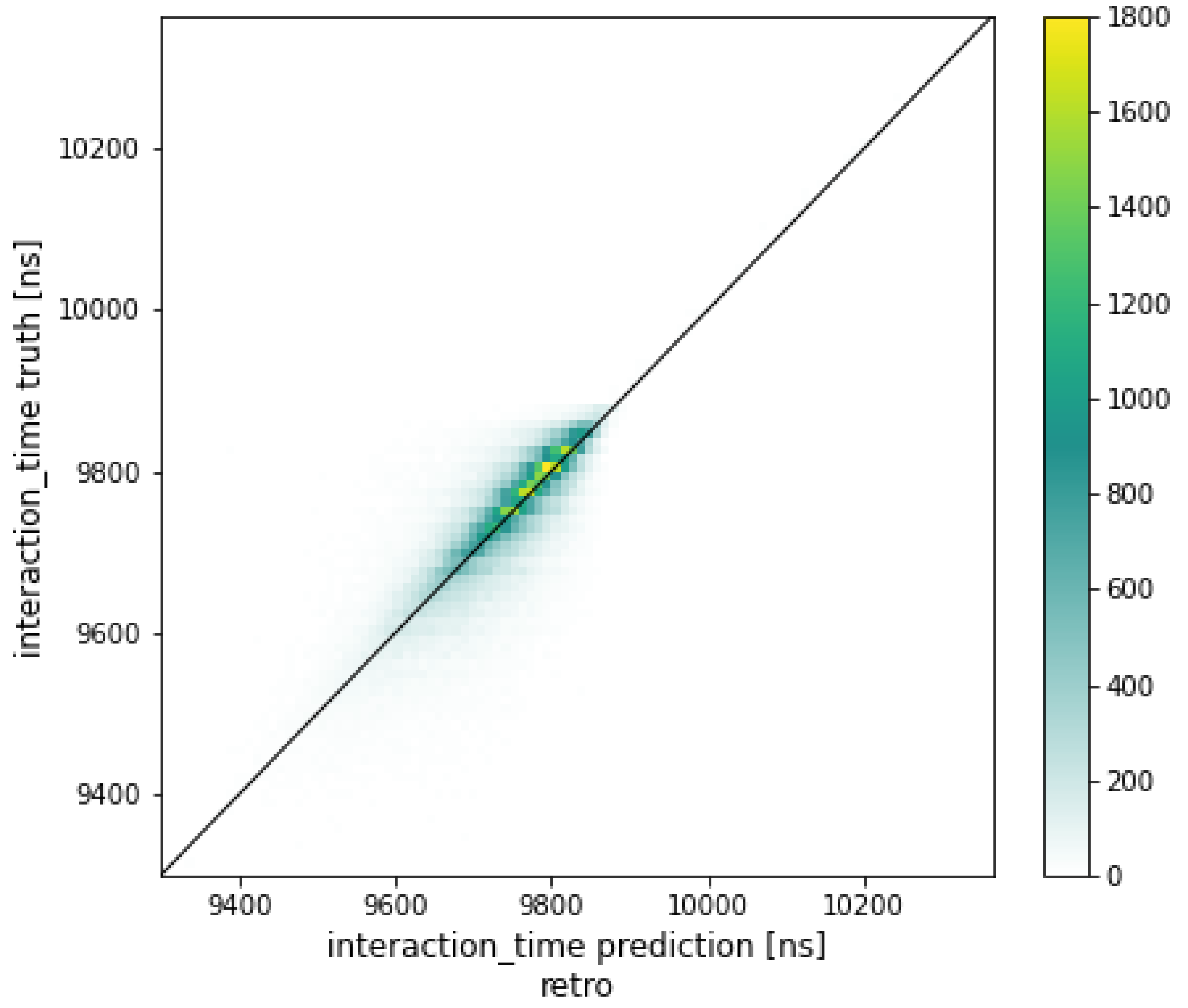
Approach: Apply scaling to the target variable. Attempt 2: **QuantileTransformer**



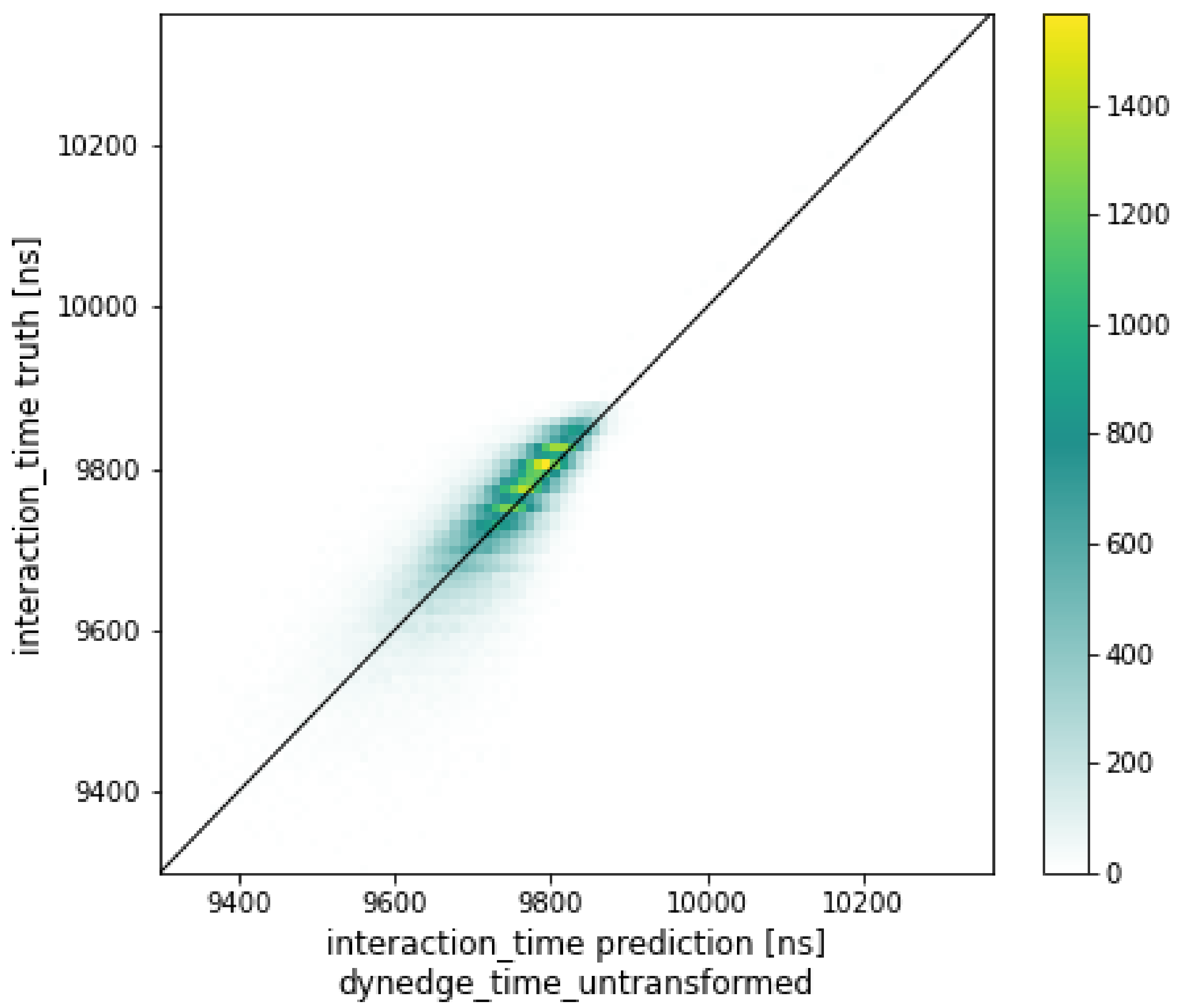


Interaction Time Regression

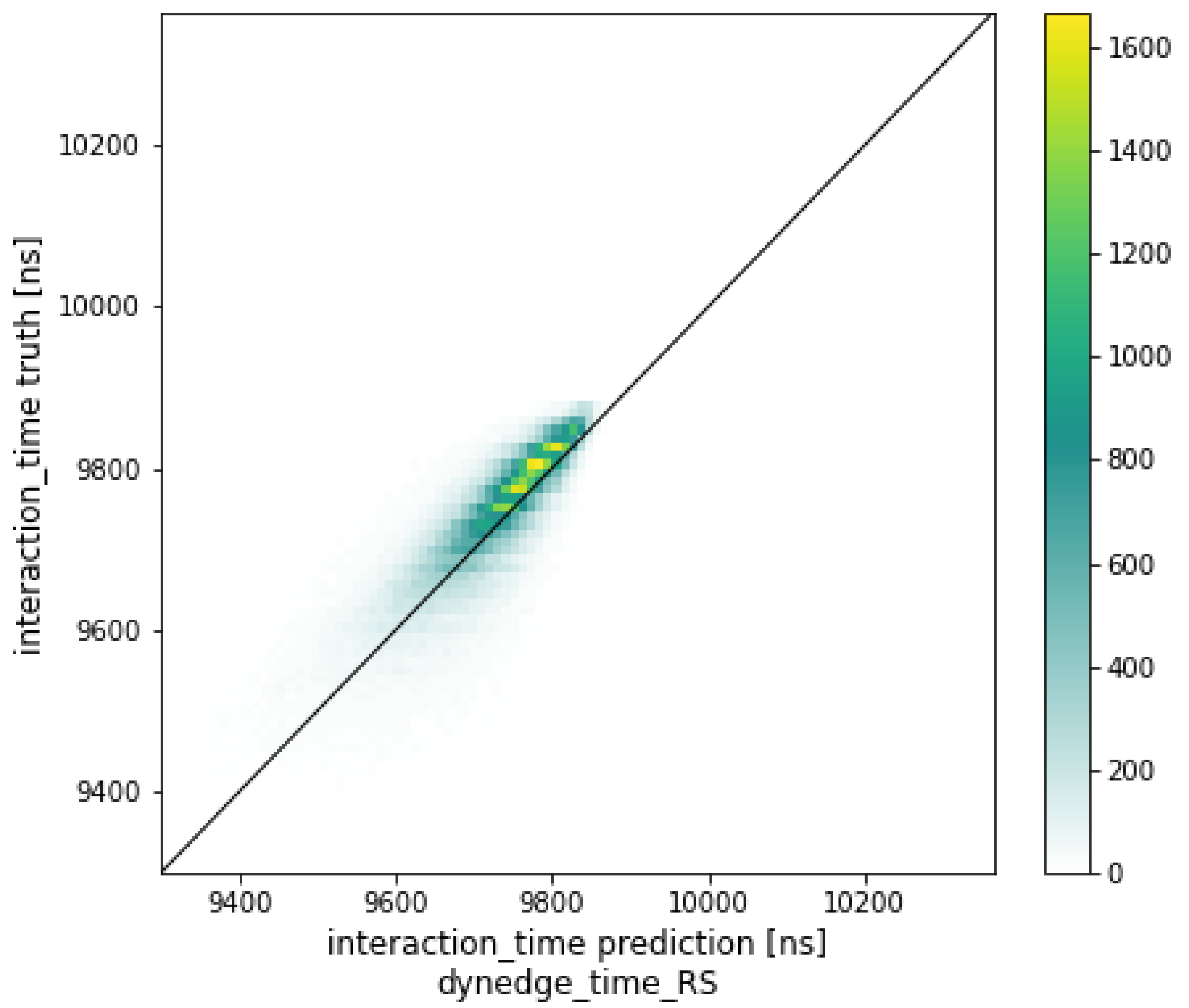
RetroReco



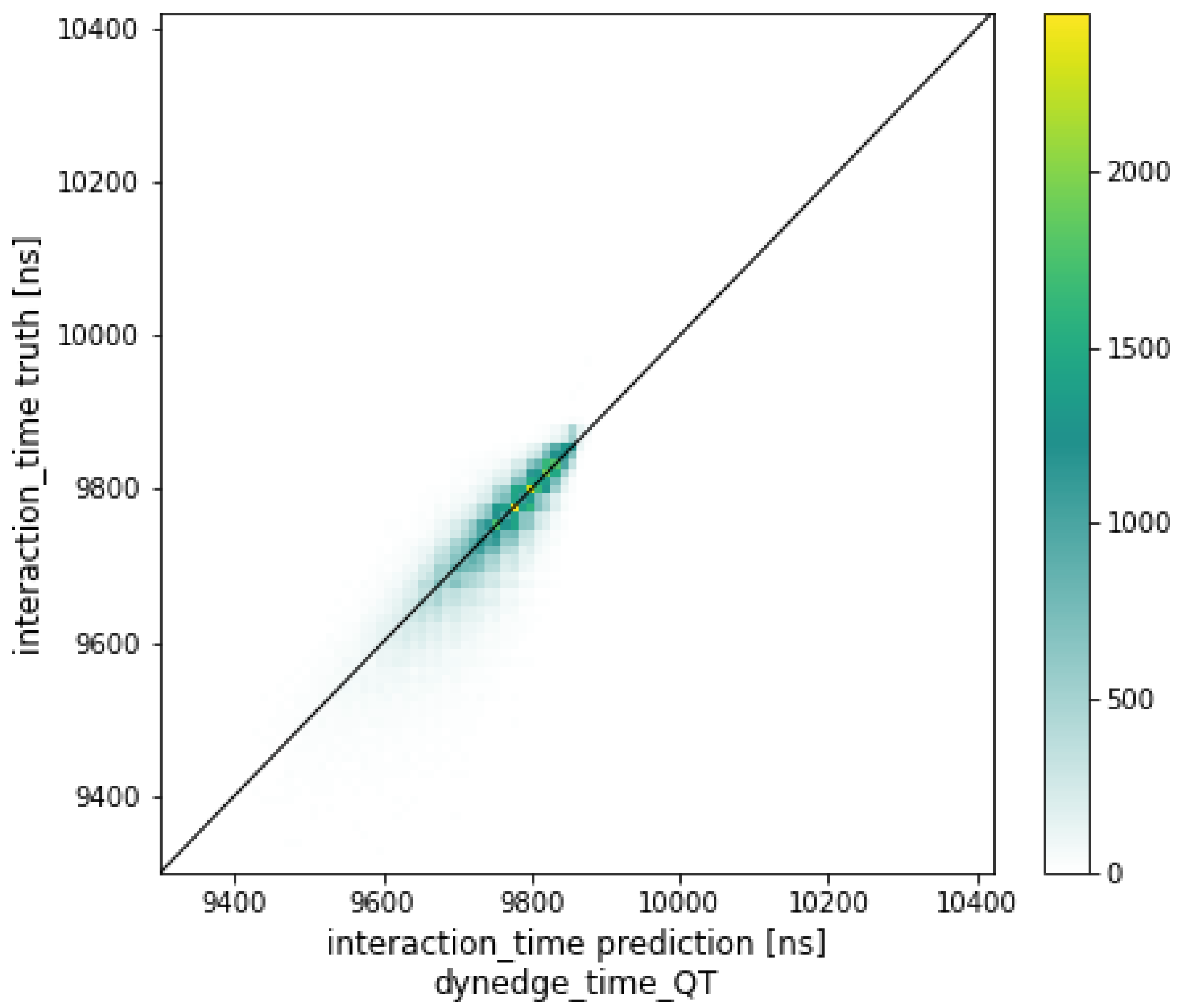
Untransformed



RobustScaler



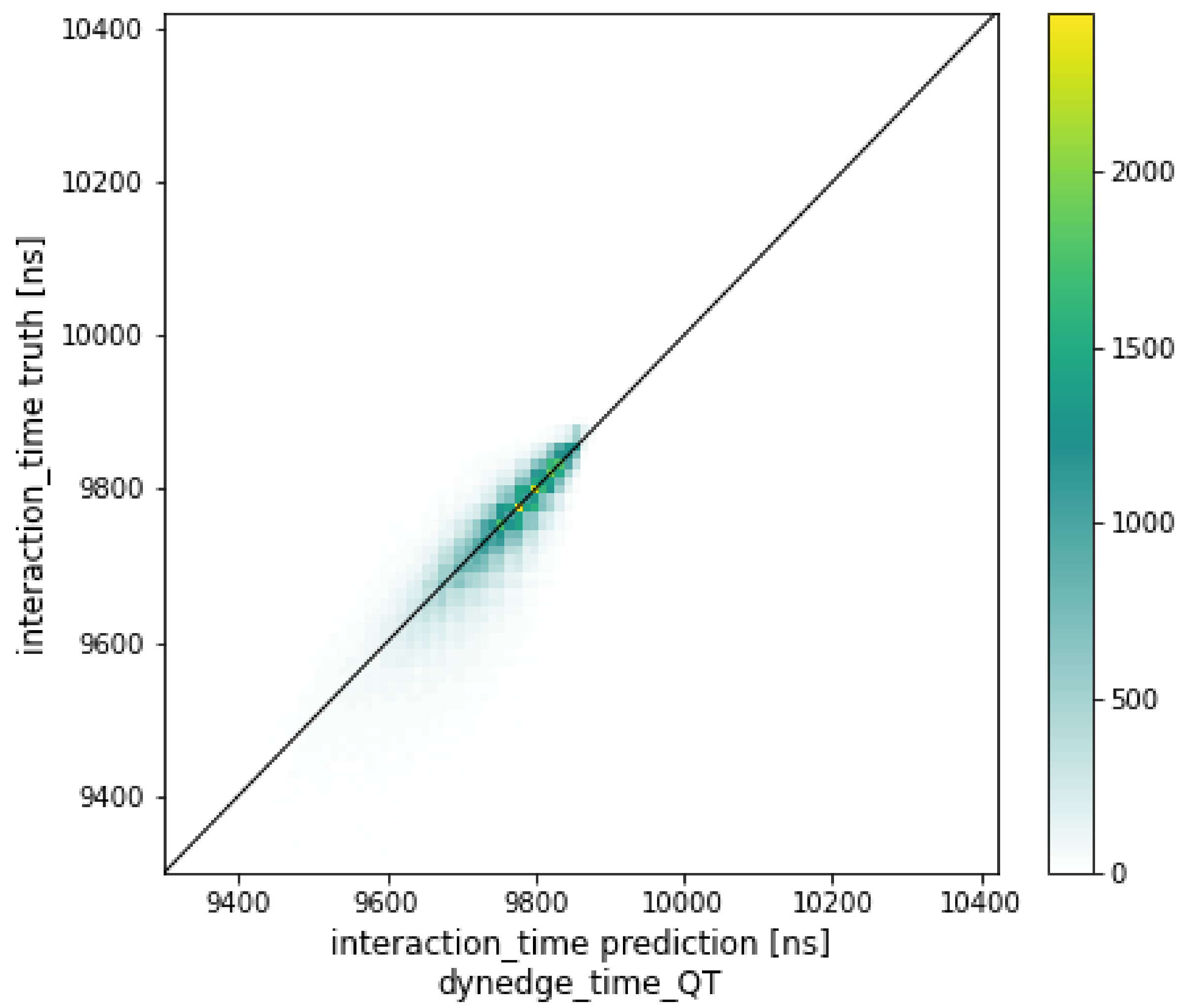
QuantileTransformer



Conclusion:

While non-linear transformations do not work well during ML optimisation, it can be necessary for some regression tasks.

The GraphNeT group has since implemented this in the framework.



Introduction

My Work

Time Regression

Event Cleaning

Upgrade Pulse Cleaning

Regression of Cleaned Pulses

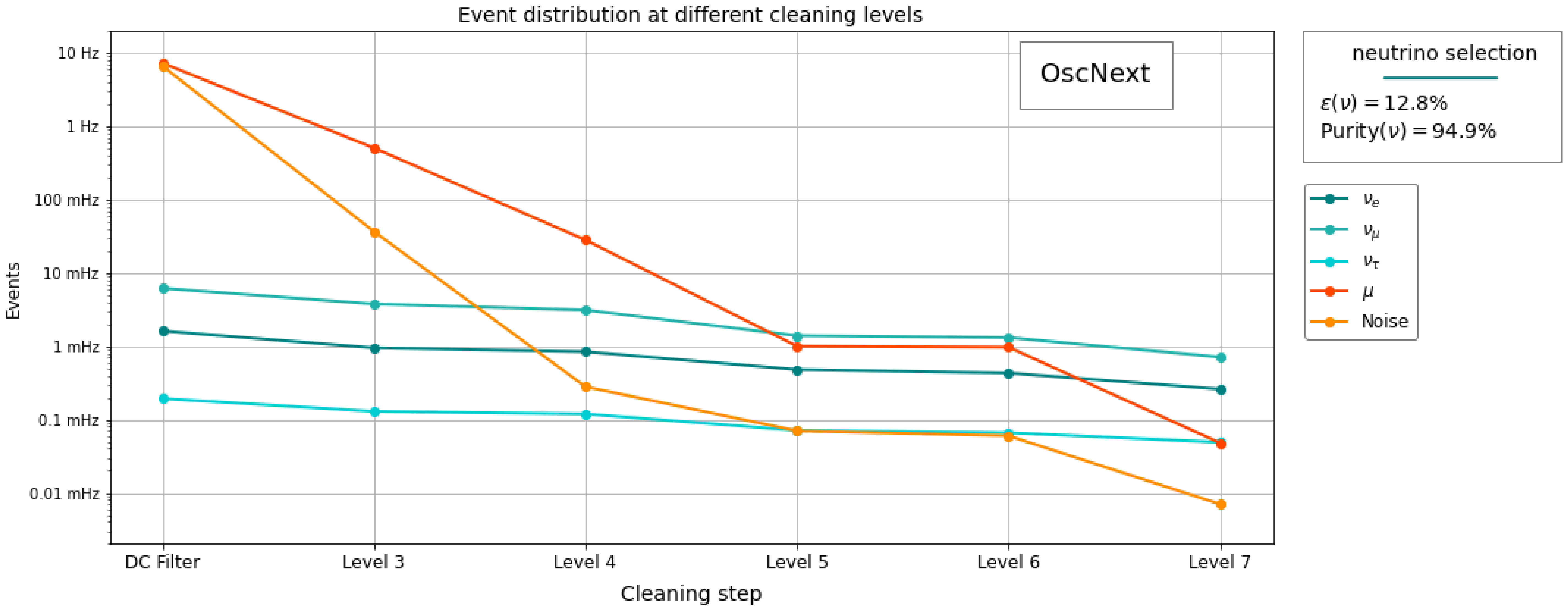
Intermezzo 1: Improvement of Pulse Cleaning

Intermezzo 2: Quality of Cleaned Events

Conclusions & Outlook

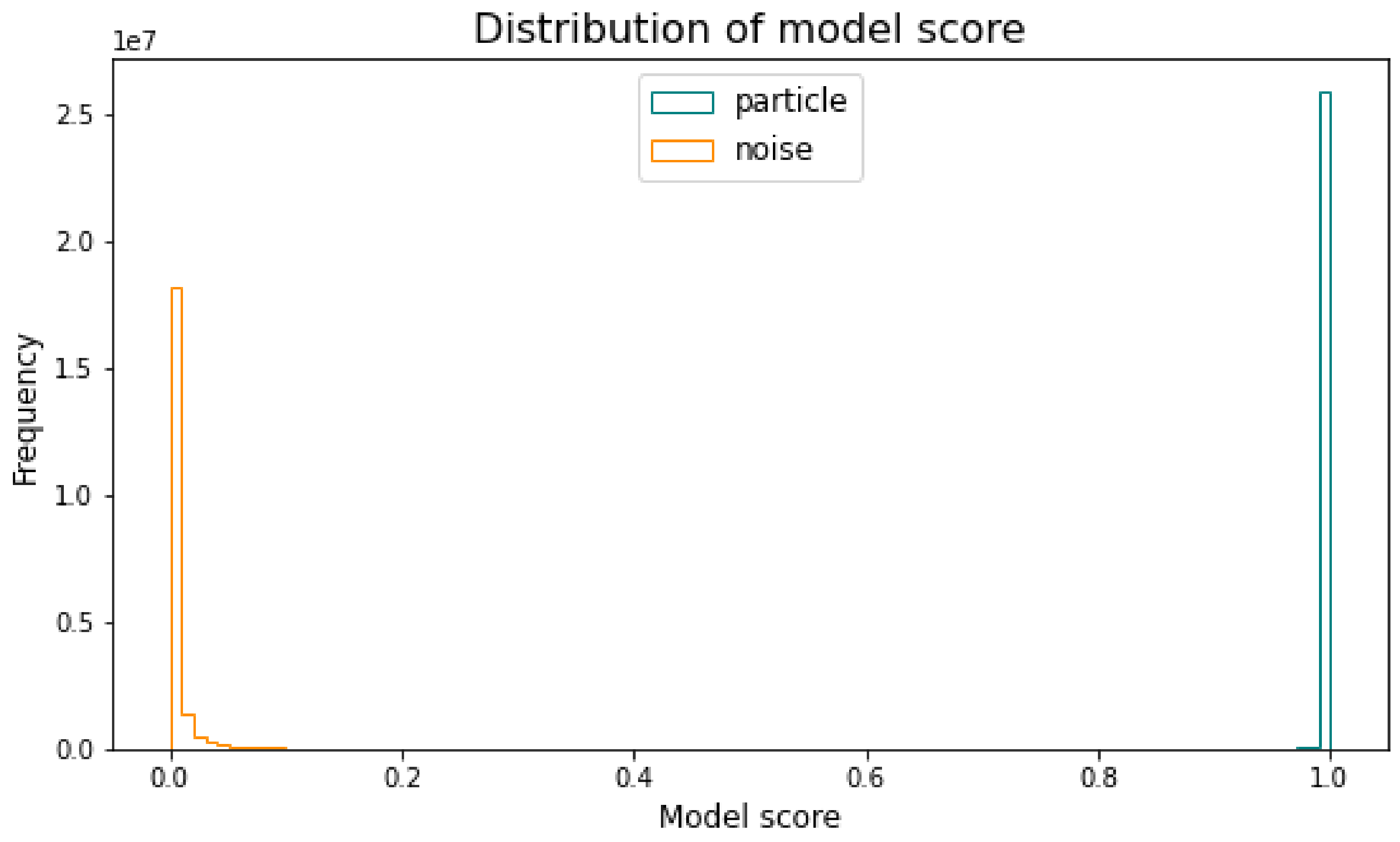
Event Selection Pipeline

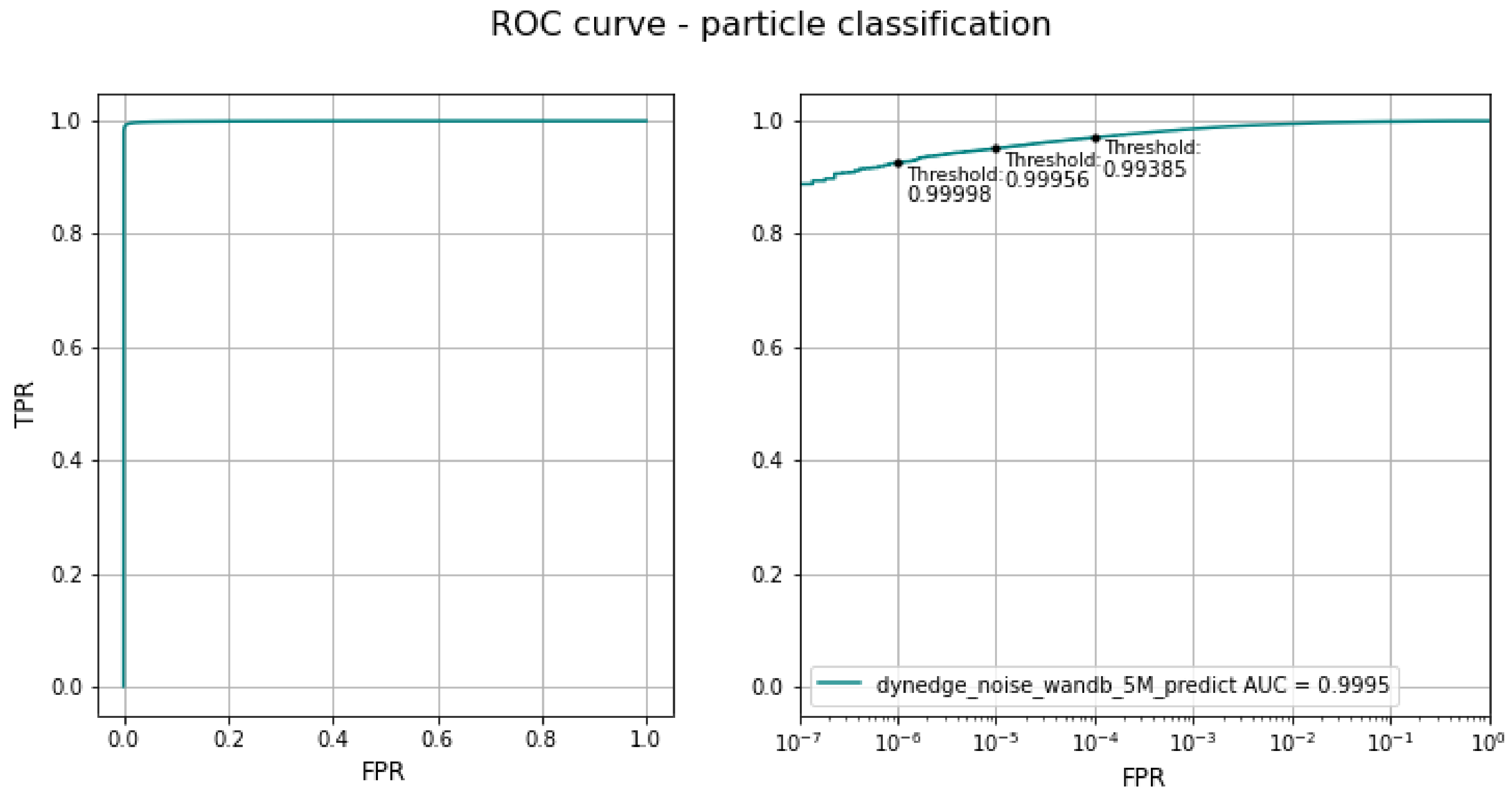
Setup: GNNs have been shown to work well at separating neutrino events from noise and muons. But will it actually improve the OscNext cleaning pipeline, and if so by how much?



Event Selection Pipeline

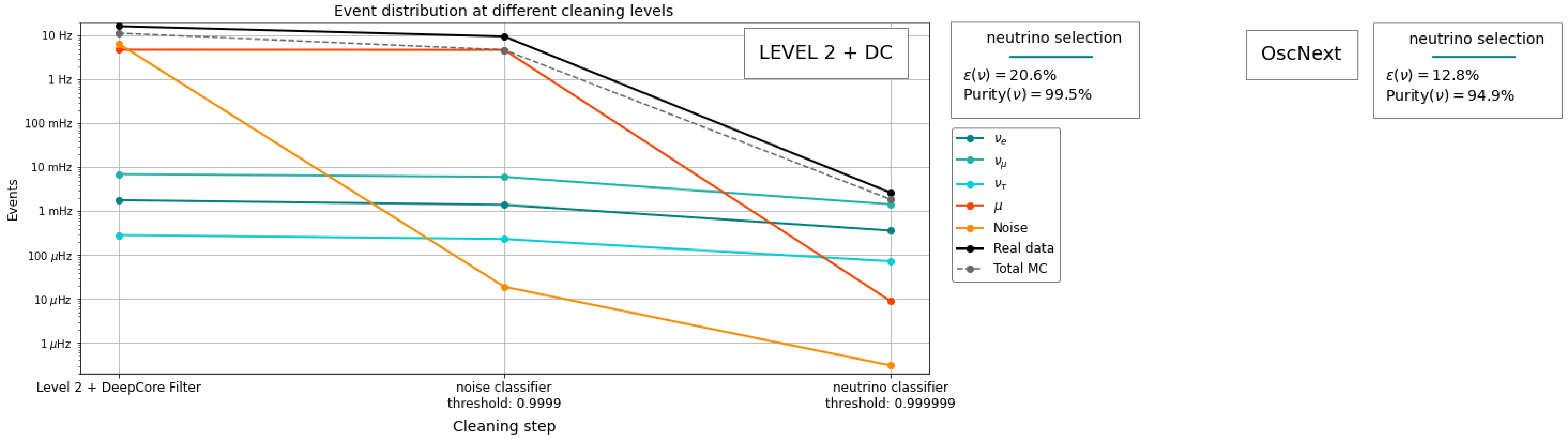
Approach: Train particle/noise classification model





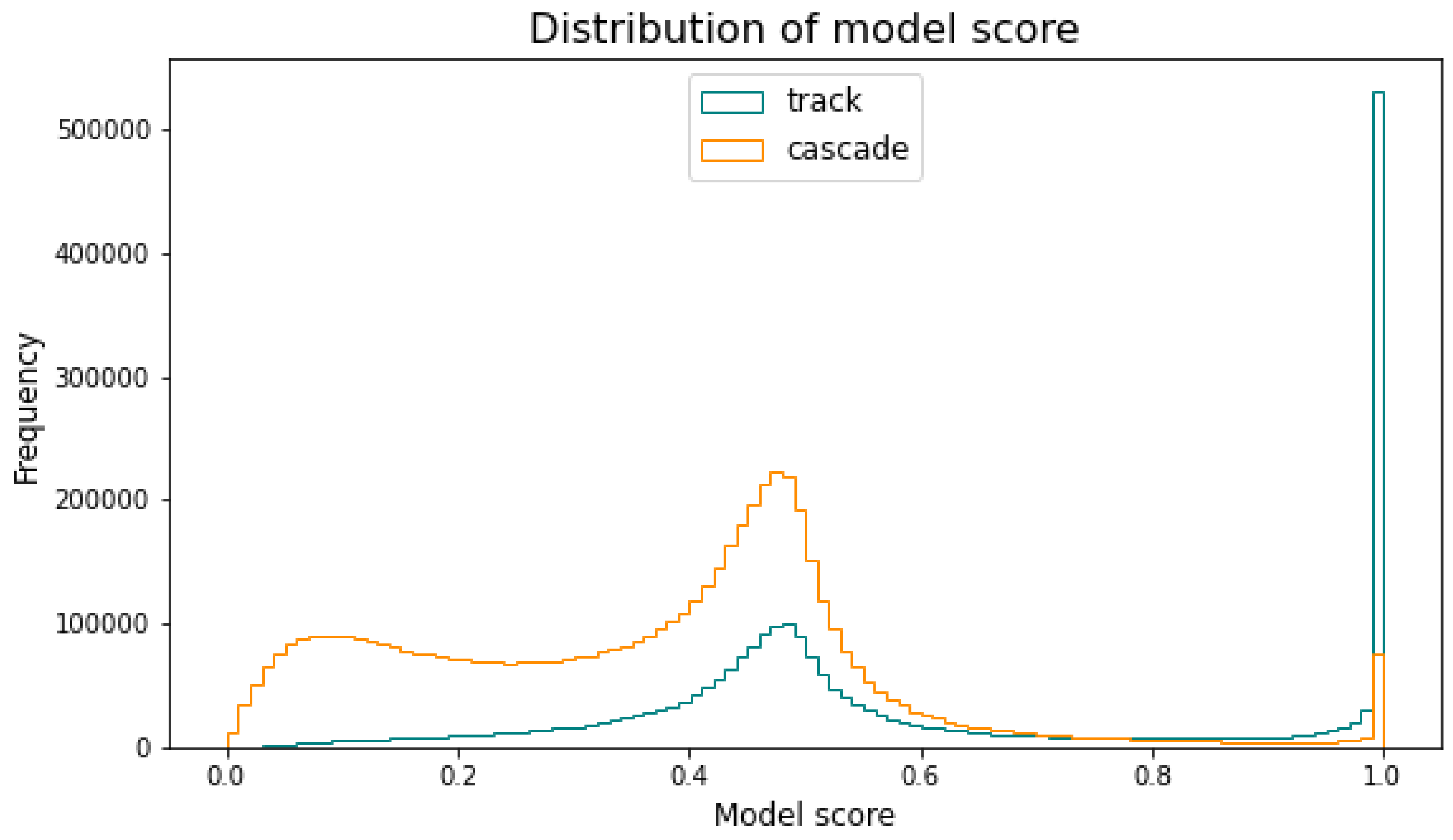
Event Selection Pipeline

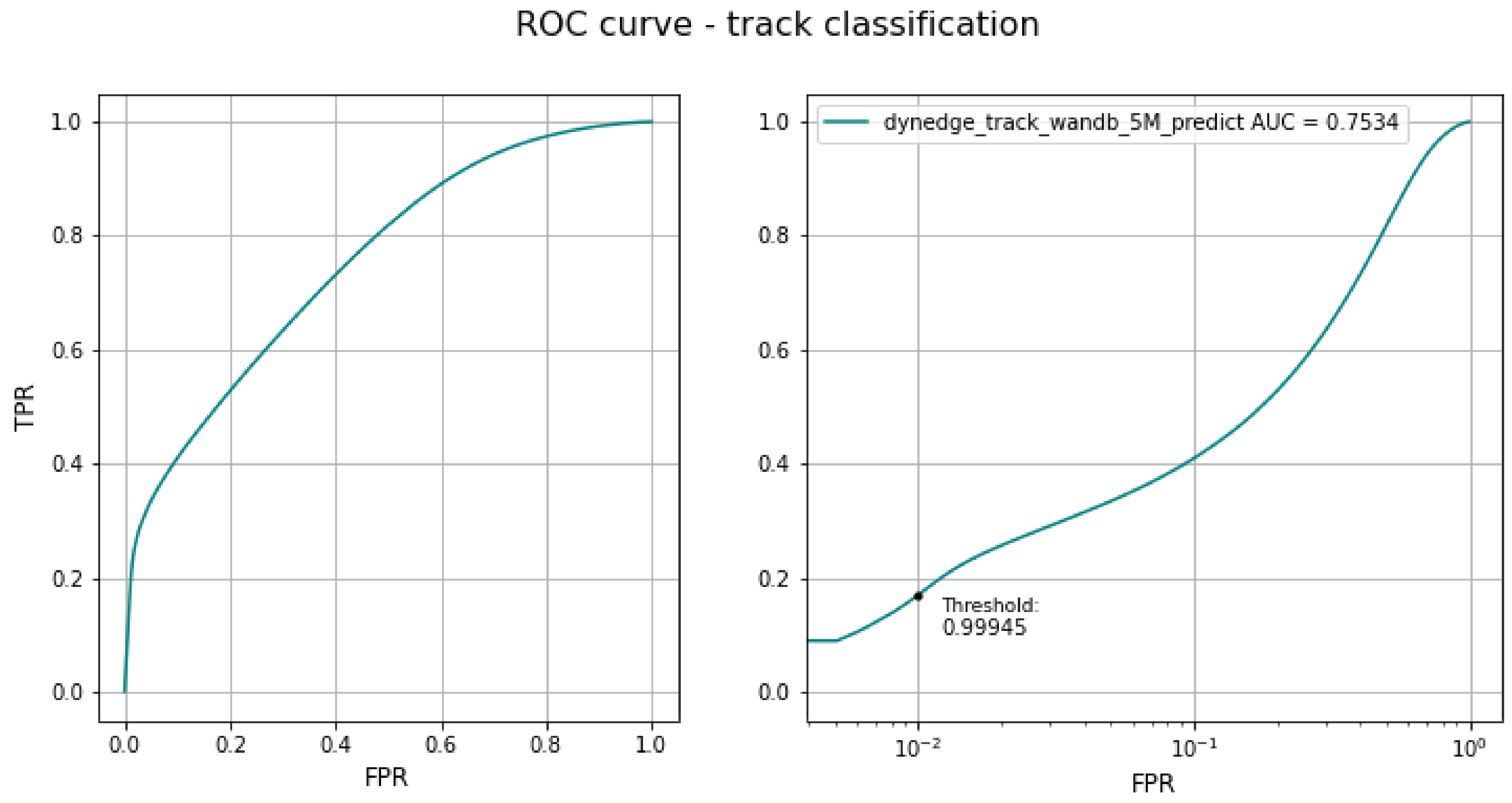
Approach: Train particle/noise classification model (and add neutrino/muon classifier)



Event Selection Pipeline

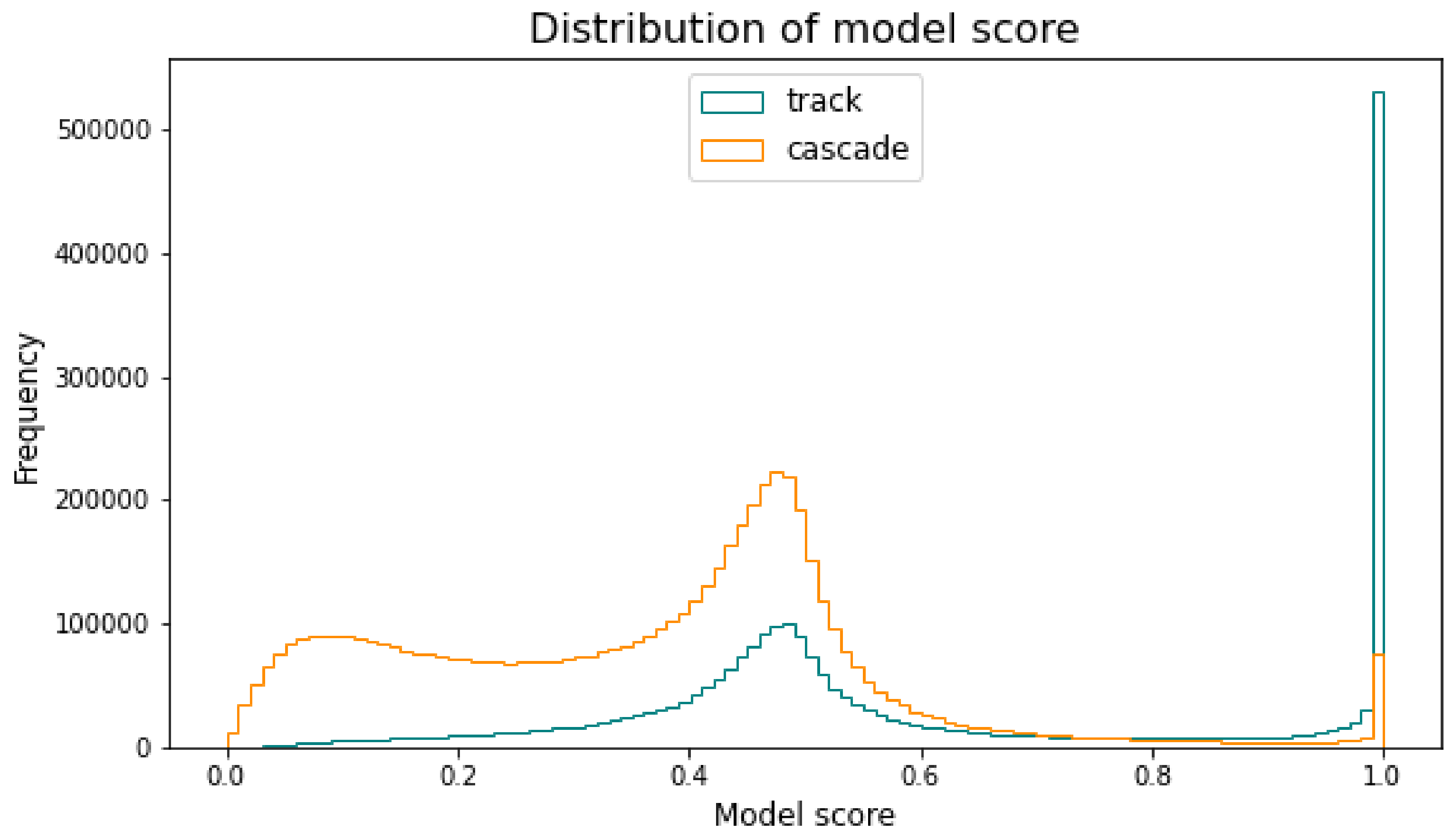
Approach: Add track/cascade classification model

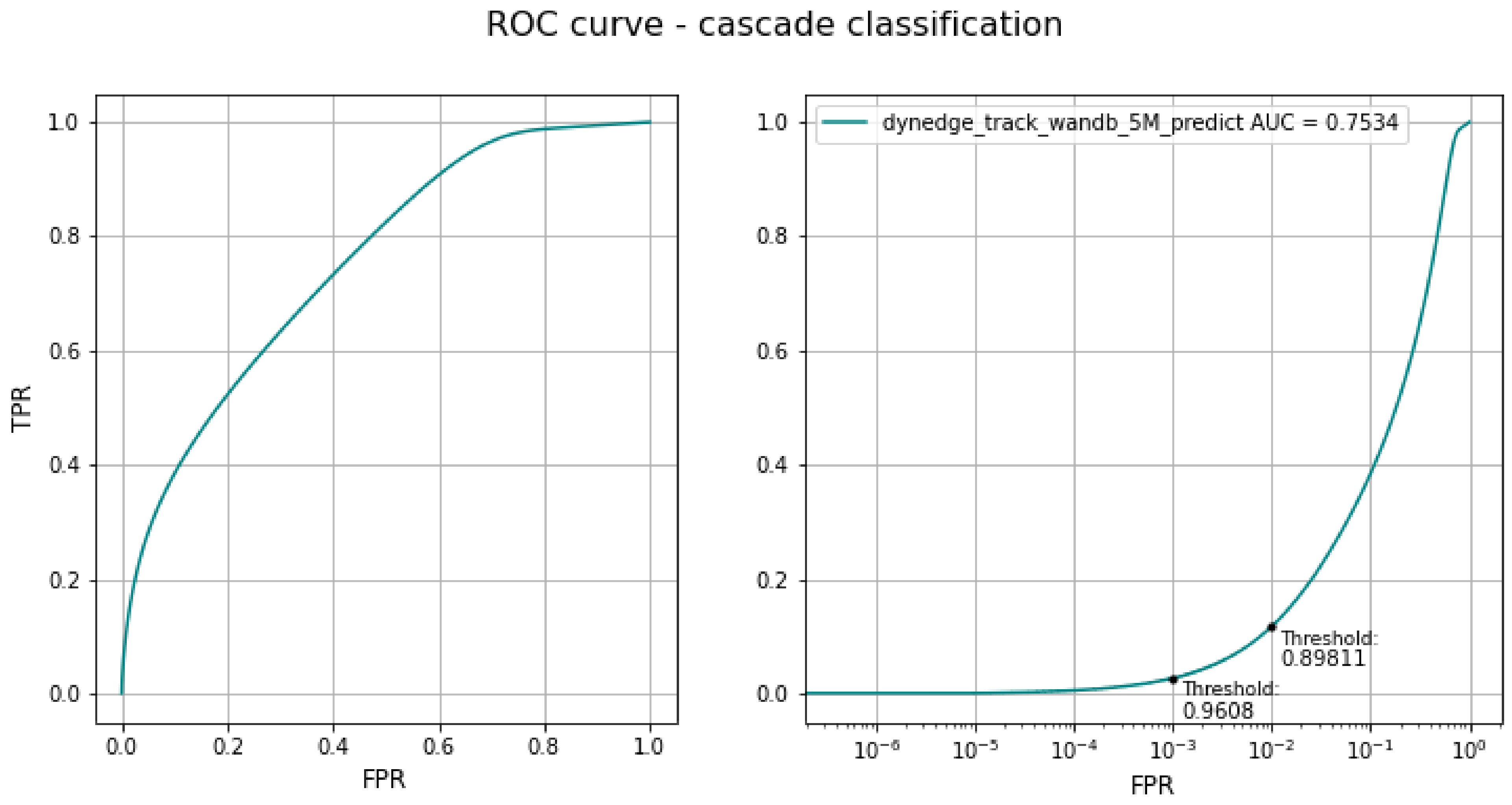




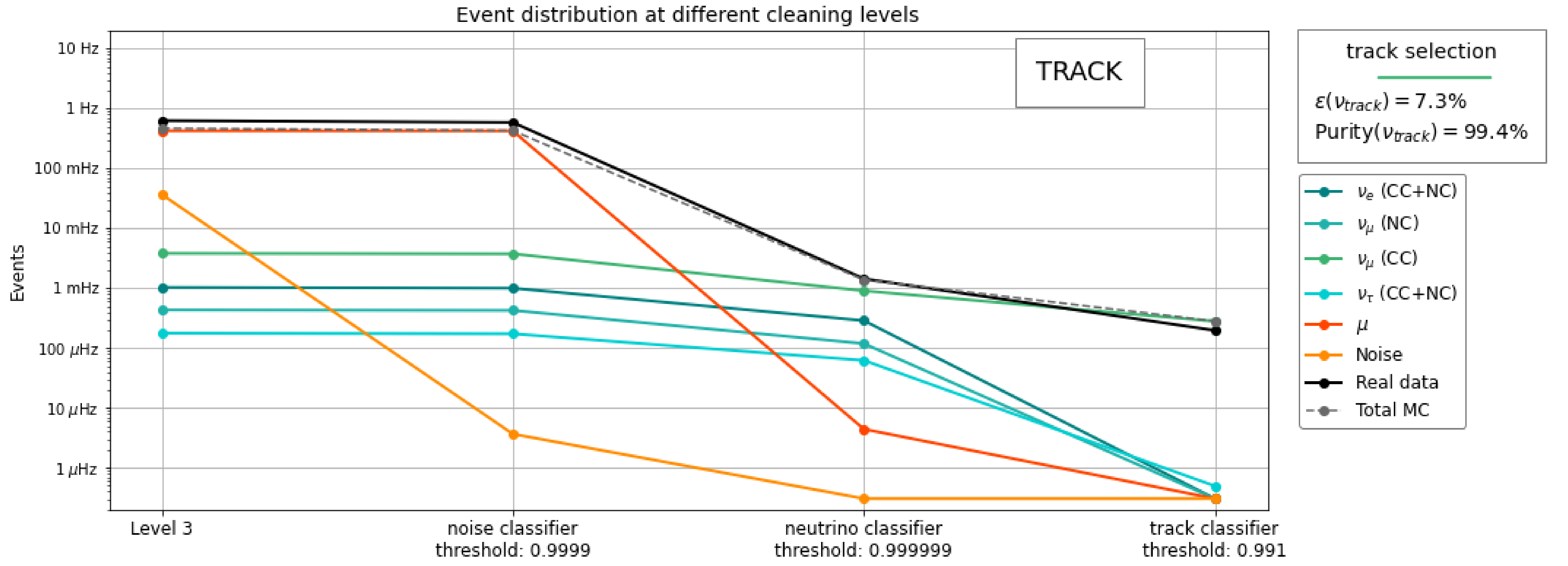
Event Selection Pipeline

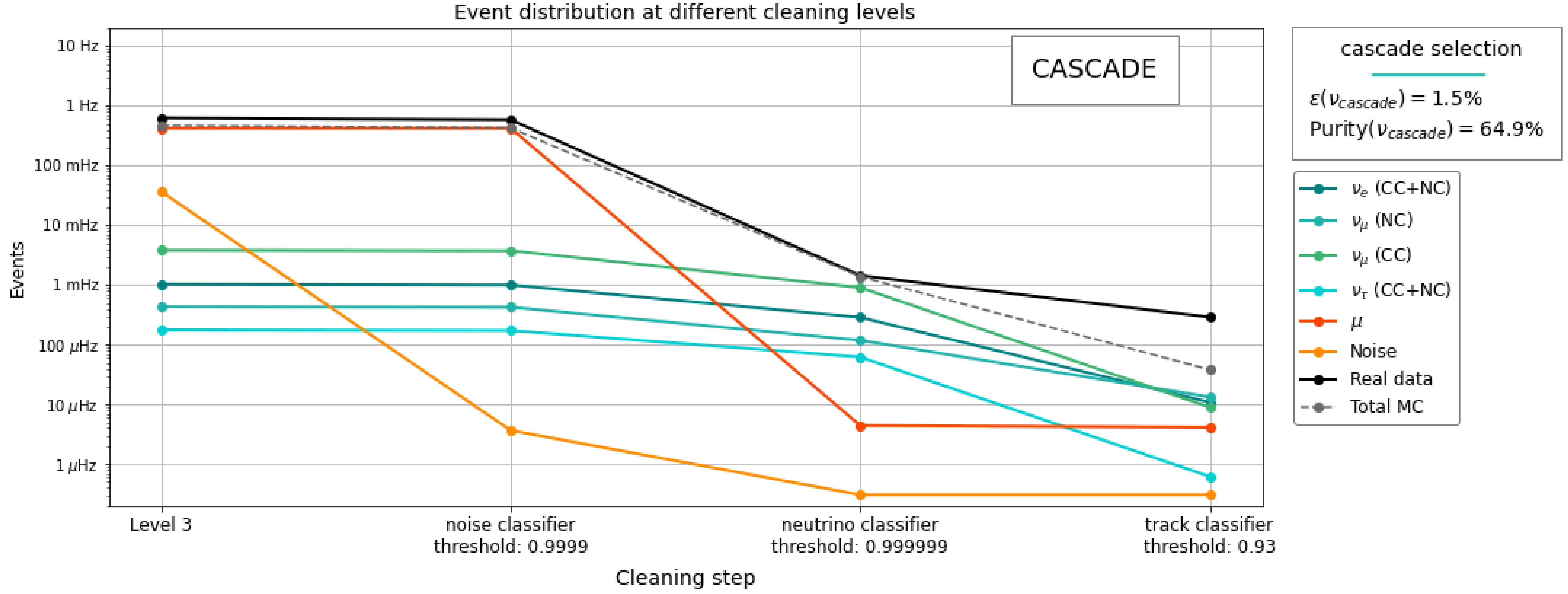
Approach: Add track/cascade classification model





Event Selection Pipeline





Conclusions:

Using GNNs for the OscNext events selection could improve the effeciency from 12.8% to 20.6% and the purity from 94.9 % to 99.5 %.

Real data comparison looks promising.

Track/cascade separation is hard but possible.

Introduction

My Work

Time Regression

Event Cleaning

Upgrade Pulse Cleaning

Regression of Cleaned Pulses

Intermezzo 1: Improvement of Pulse Cleaning

Intermezzo 2: Quality of Cleaned Events

Conclusions & Outlook

Upgrade Pulse Cleaning

- IceCube Upgrade is the envisioned low-energy-focused extension of the IceCube detector array
- Upgrade features higher detector density and new DOM types

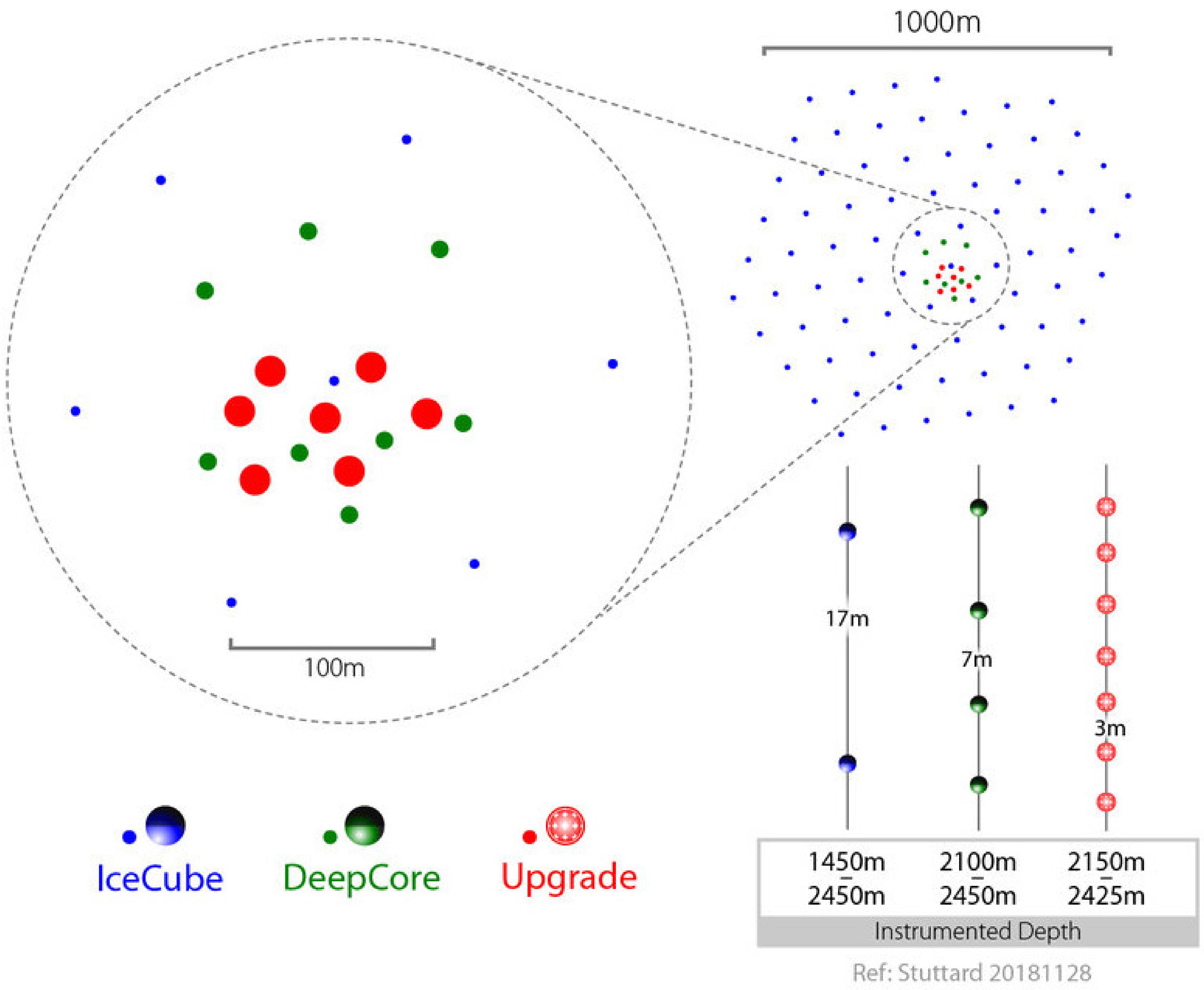


Figure from Marek L. Kowalski, *Next generation neutrino detectors at the South Pole*

Upgrade Pulse Cleaning

Setup: The new Upgrade DOMs are very noisy due to contamination with radioactive isotopes.

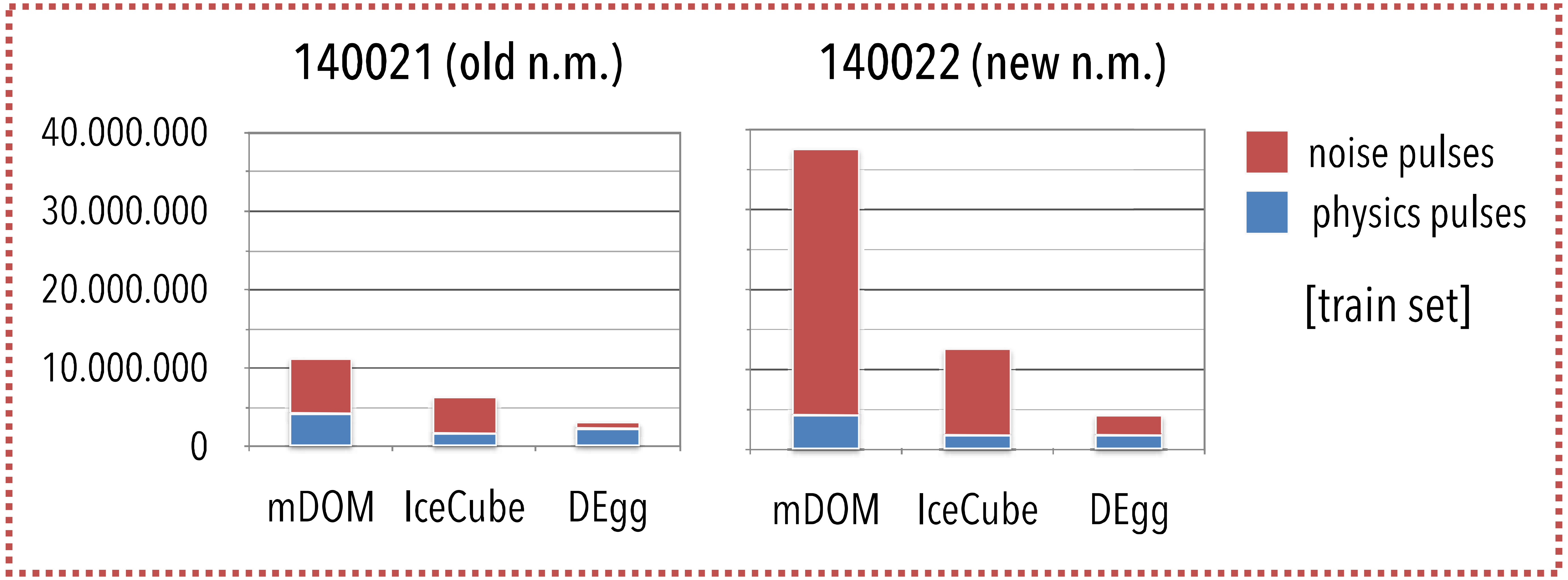
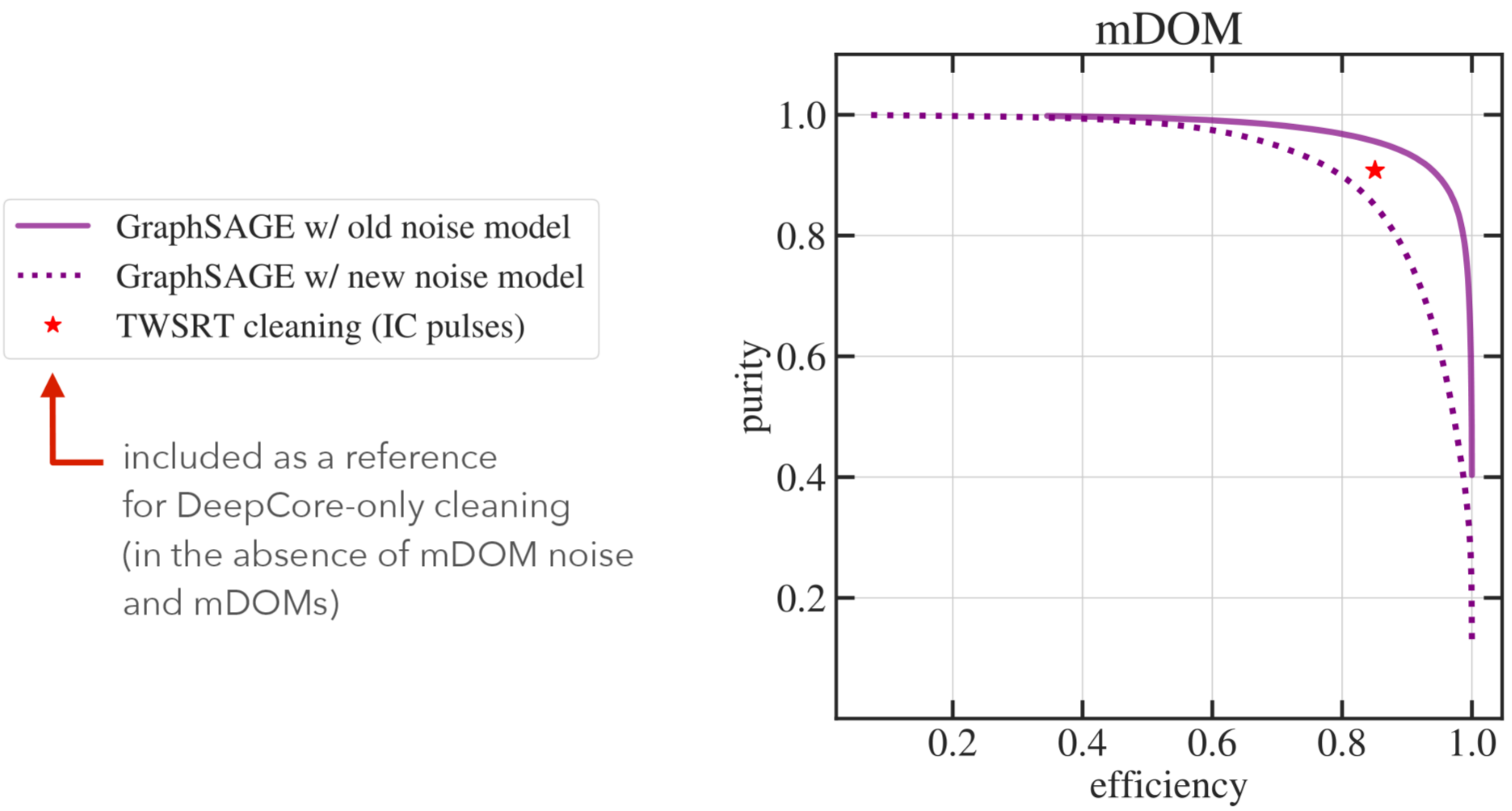


Figure from Tania Kozynets & Tom Stuttard, *GNN-assisted pulse cleaning for the IceCube-Upgrade*

Upgrade Pulse Cleaning

Setup: The new Upgrade DOMs are very noisy due to contamination with radioactive isotopes.

Promising pulse cleaning has already been demonstrated using a simple GNN.

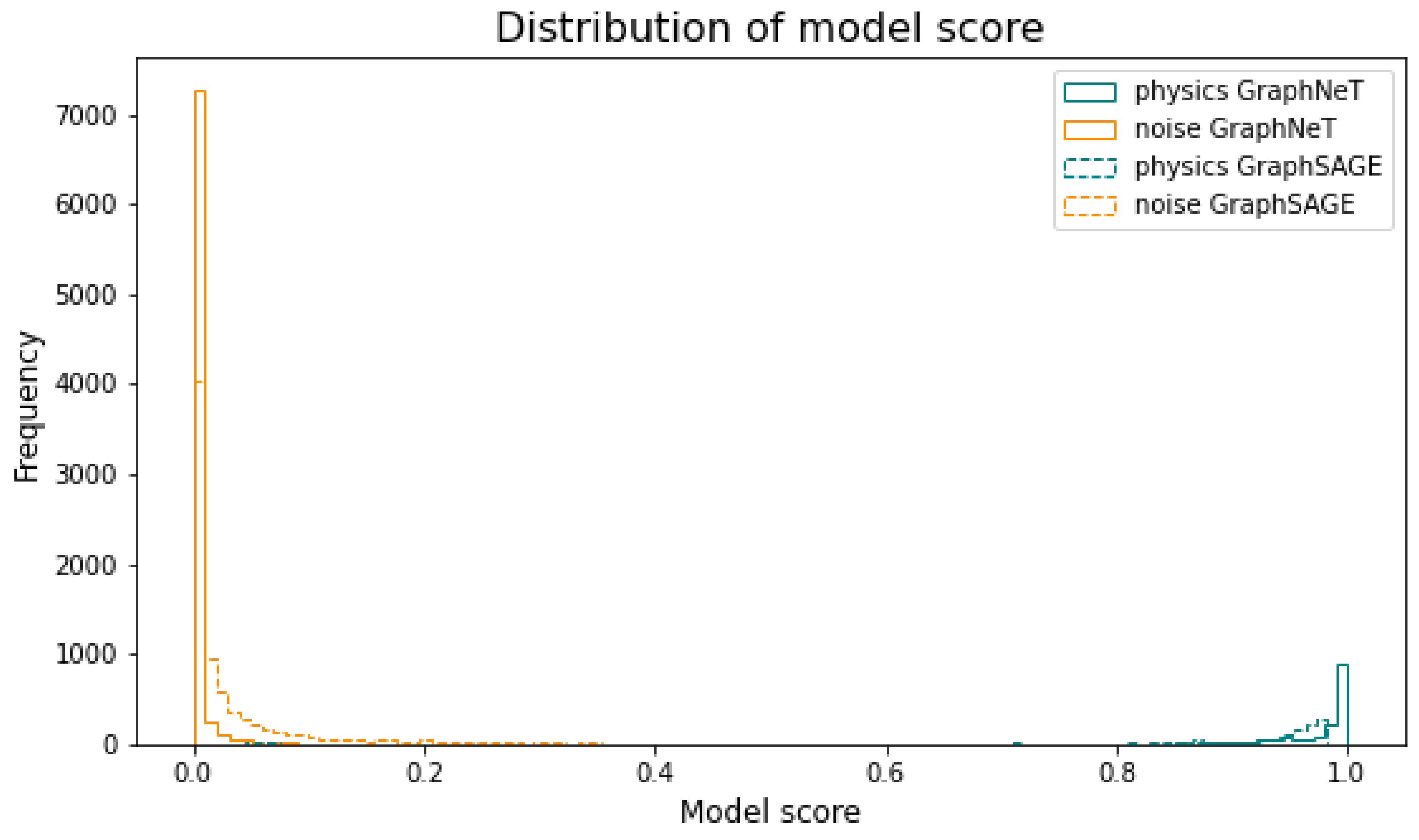


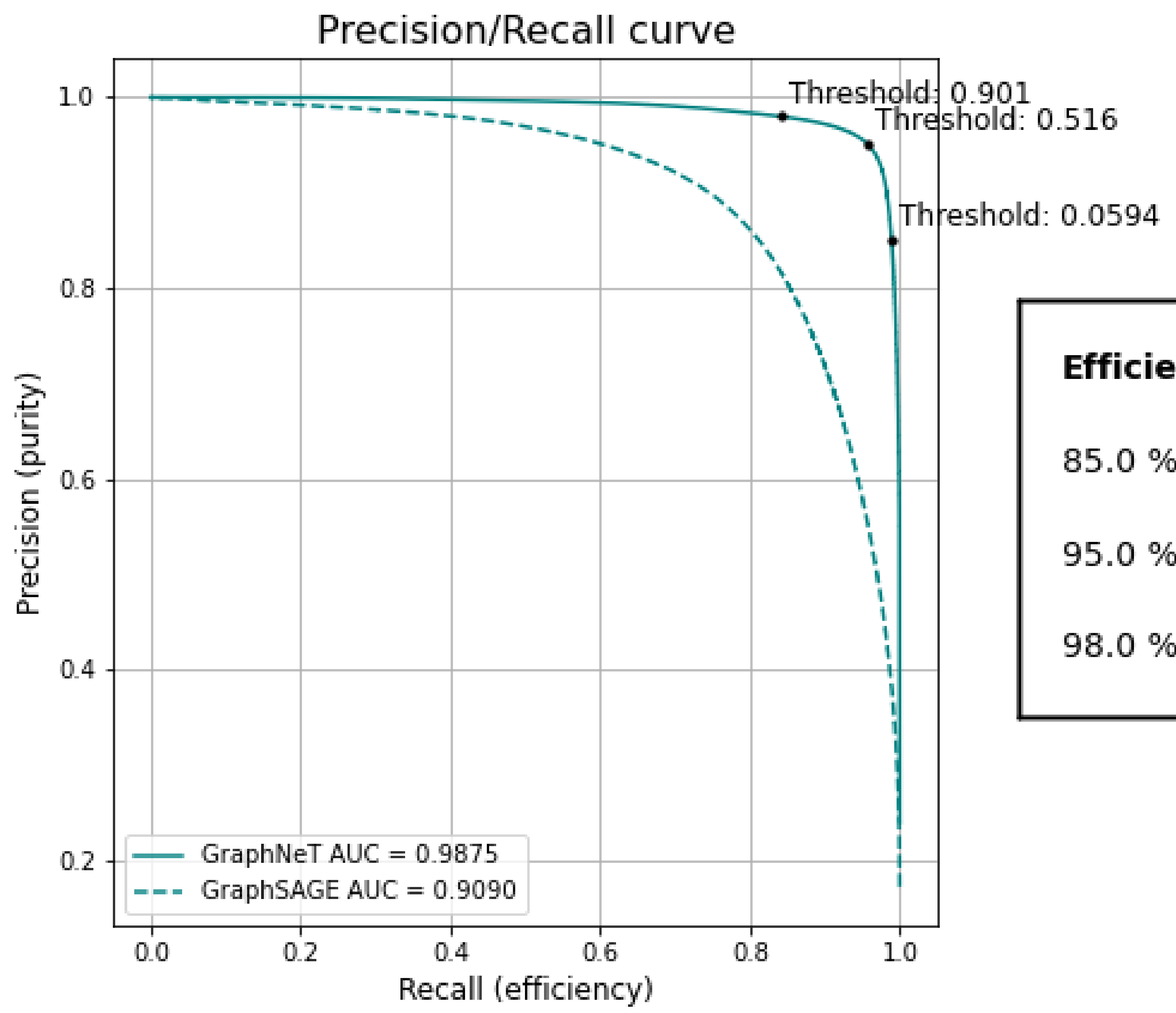
	old (new) model
90% efficiency →	94% (77%) purity
90% purity →	95% (80%) efficiency

Figure from Tania Kozynets & Tom Stuttard, *GNN-assisted pulse cleaning for the IceCube-Upgrade*

Upgrade Pulse Cleaning

Approach: Train a model for pulse-level noise cleaning using the DynEdge architecture



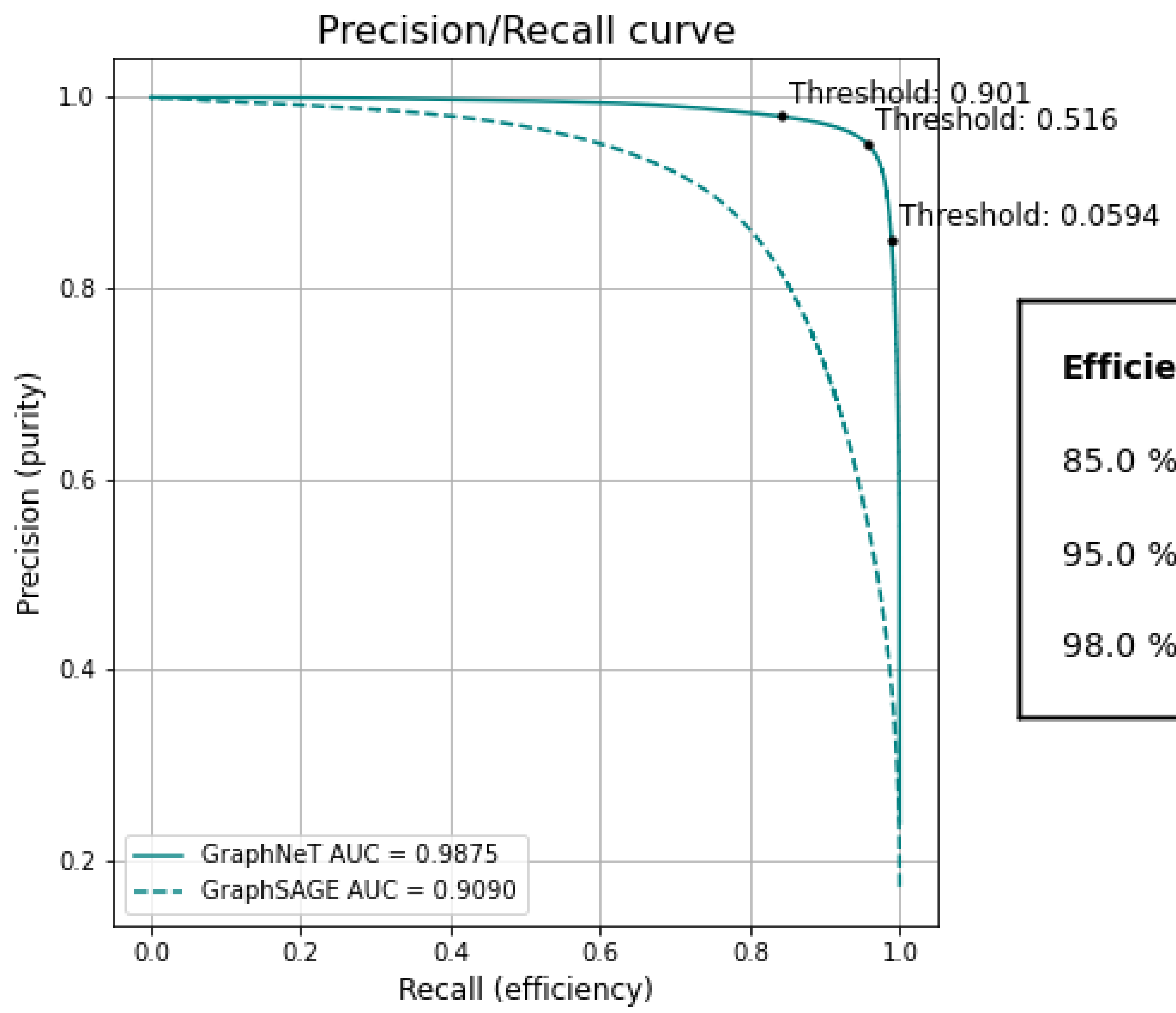


Efficiency	Purity GraphNeT	Purity GraphSAGE
85.0 % ⟶	98.9 %	81.2 %
95.0 % ⟶	95.8 %	60.6 %
98.0 % ⟶	84.2 %	40.7 %

Conclusions:

Pulse cleaning using DynEdge works well compared to the previous model.

The real test should to see if the additional noise can be nullified by the pulse cleaning.



Efficiency	Purity GraphNeT	Purity GraphSAGE
85.0 % →	98.9 %	81.2 %
95.0 % →	95.8 %	60.6 %
98.0 % →	84.2 %	40.7 %

Introduction

My Work

Time Regression

Event Cleaning

Upgrade Pulse Cleaning

Regression of Cleaned Pulses

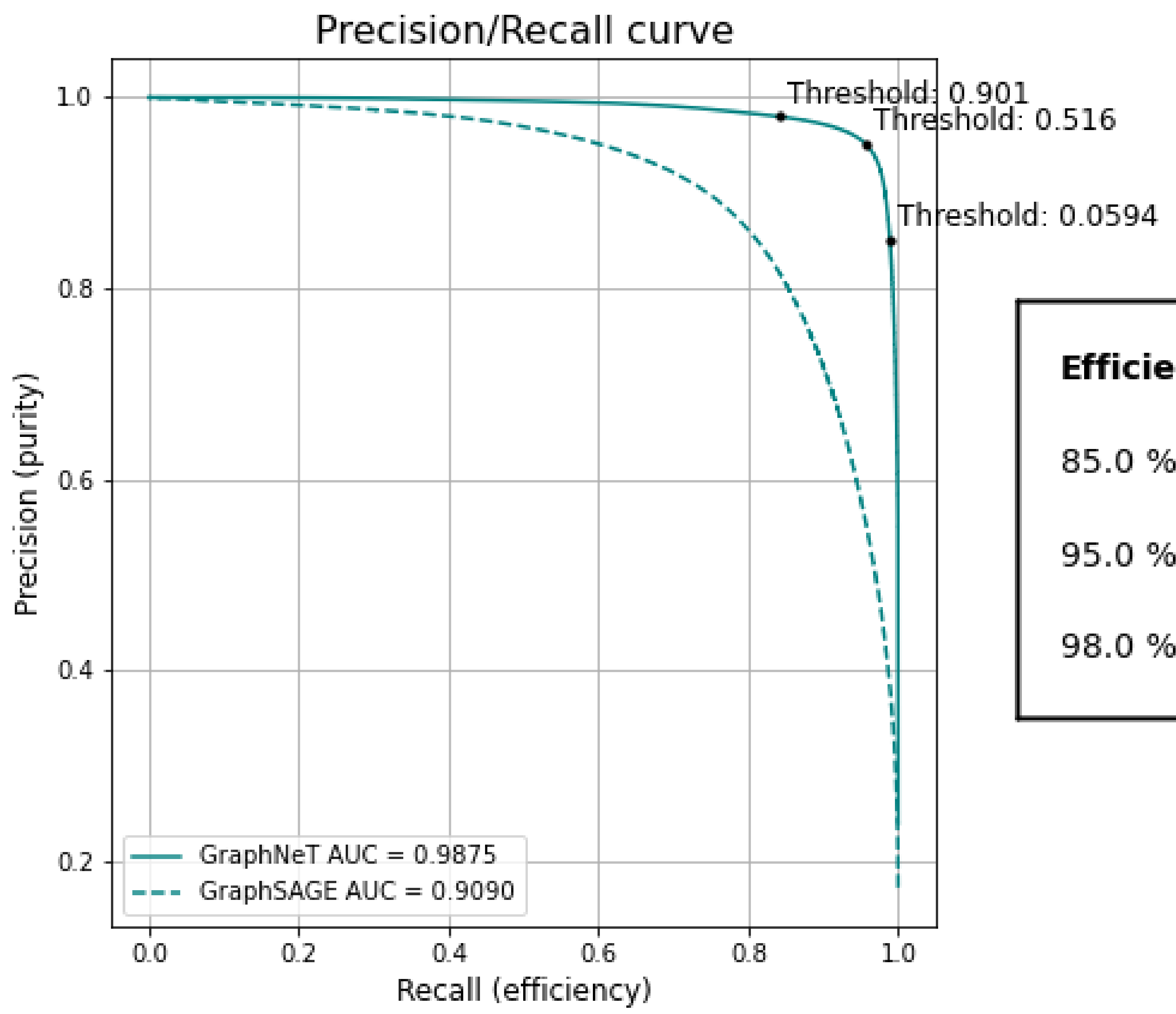
Intermezzo 1: Improvement of Pulse Cleaning

Intermezzo 2: Quality of Cleaned Events

Conclusions & Outlook

Regression of Cleaned Pulses

Setup: With no real comparison for the pulse cleaning, we want to get a measure of its effect on the upgrade data.

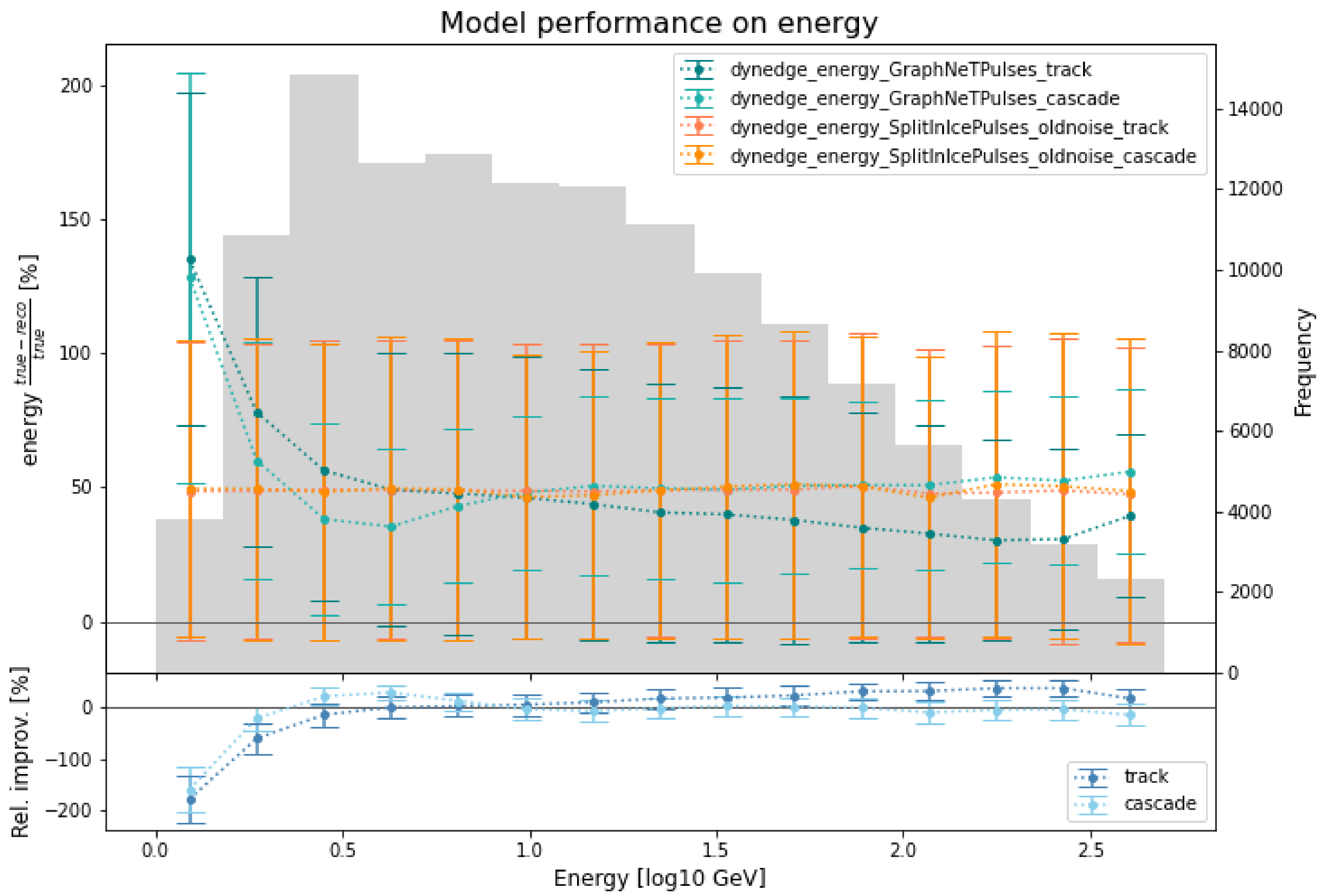


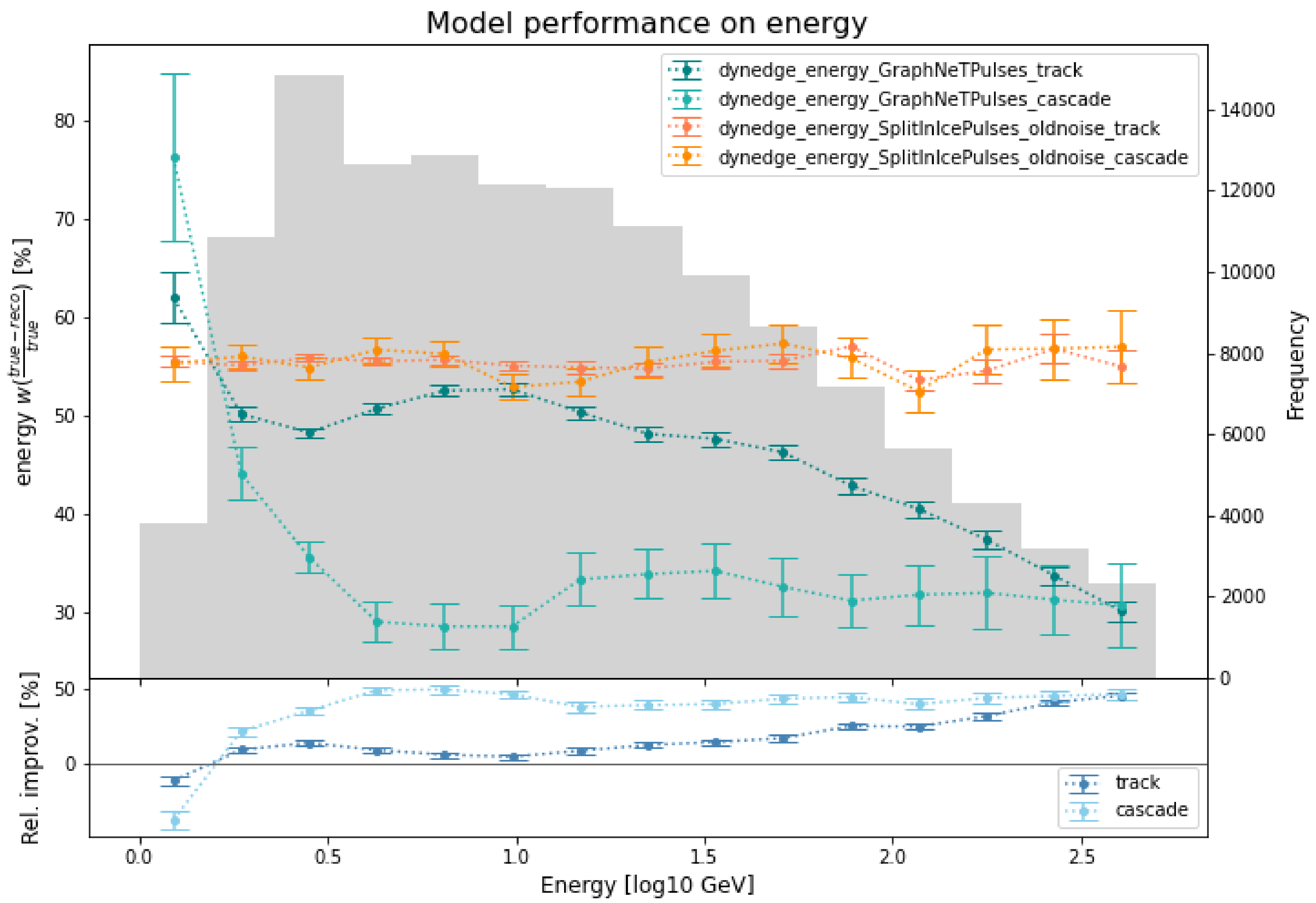
Efficiency	Purity GraphNeT	Purity GraphSAGE
85.0 % →	98.9 %	81.2 %
95.0 % →	95.8 %	60.6 %
98.0 % →	84.2 %	40.7 %

Regression of Cleaned Pulses

Approach: Train regression models on cleaned data and uncleaned data with the old and new noise models.

Energy

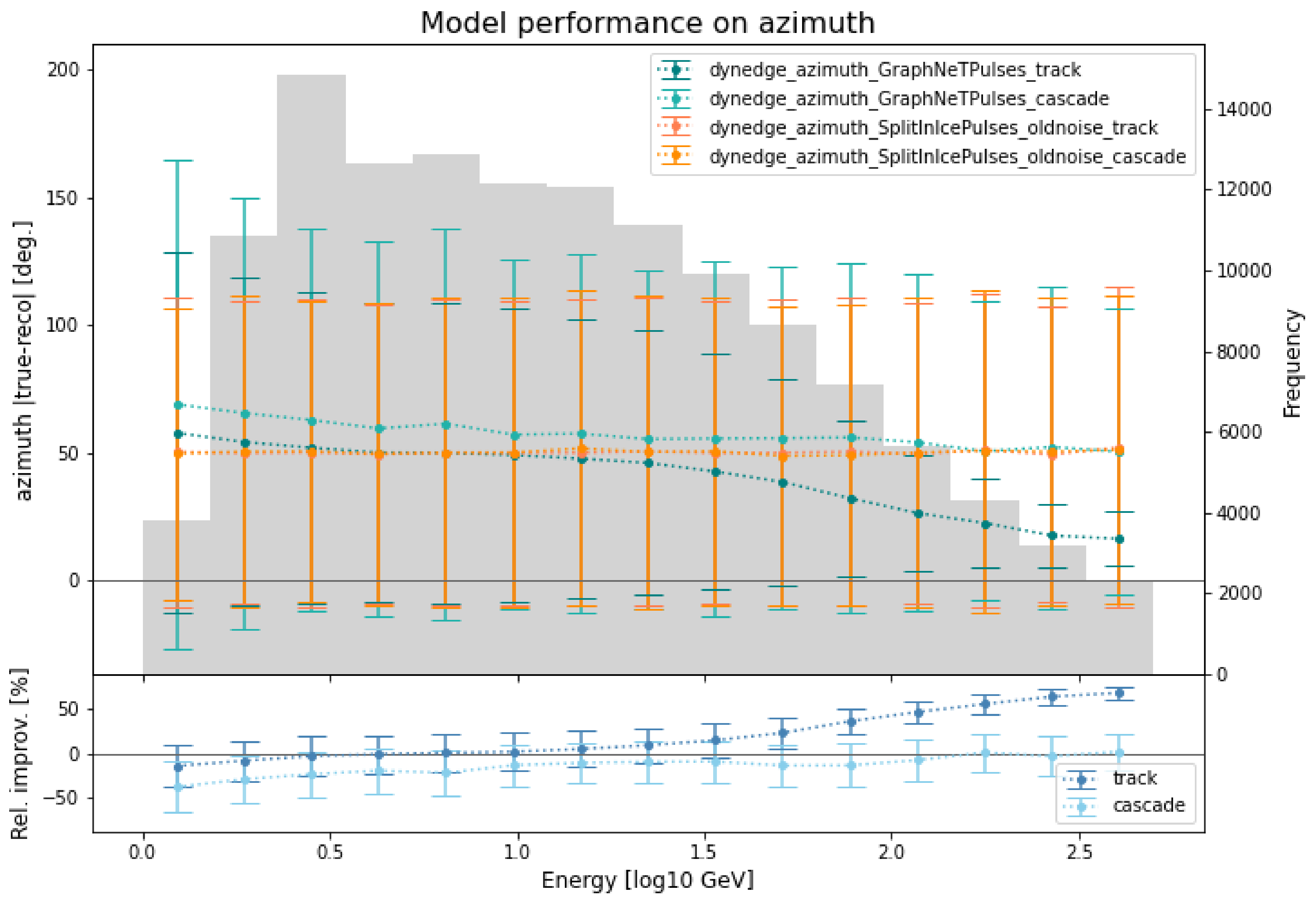


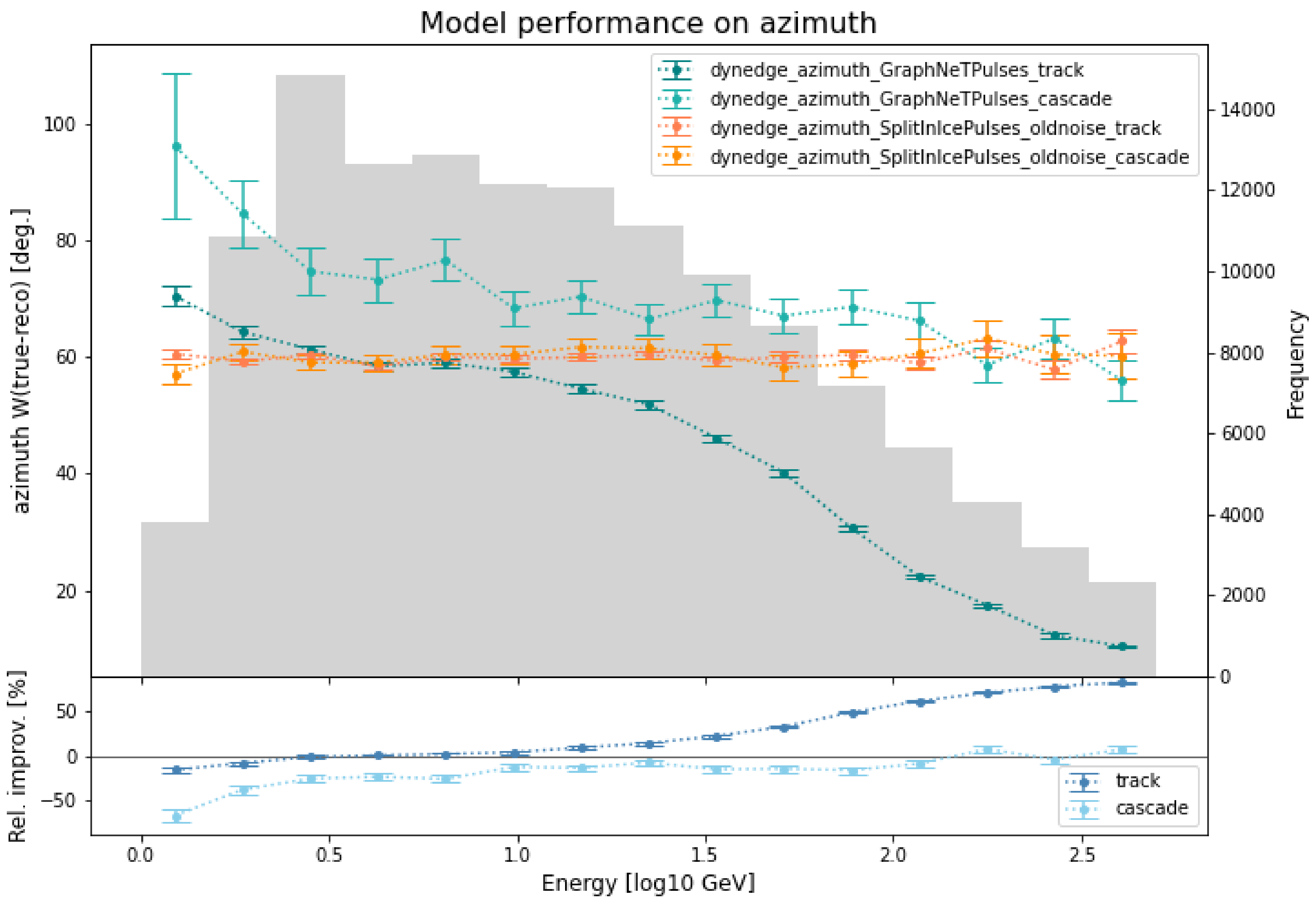


Regression of Cleaned Pulses

Approach: Train regression models on cleaned data and uncleaned data with the old and new noise models.

Azimuth

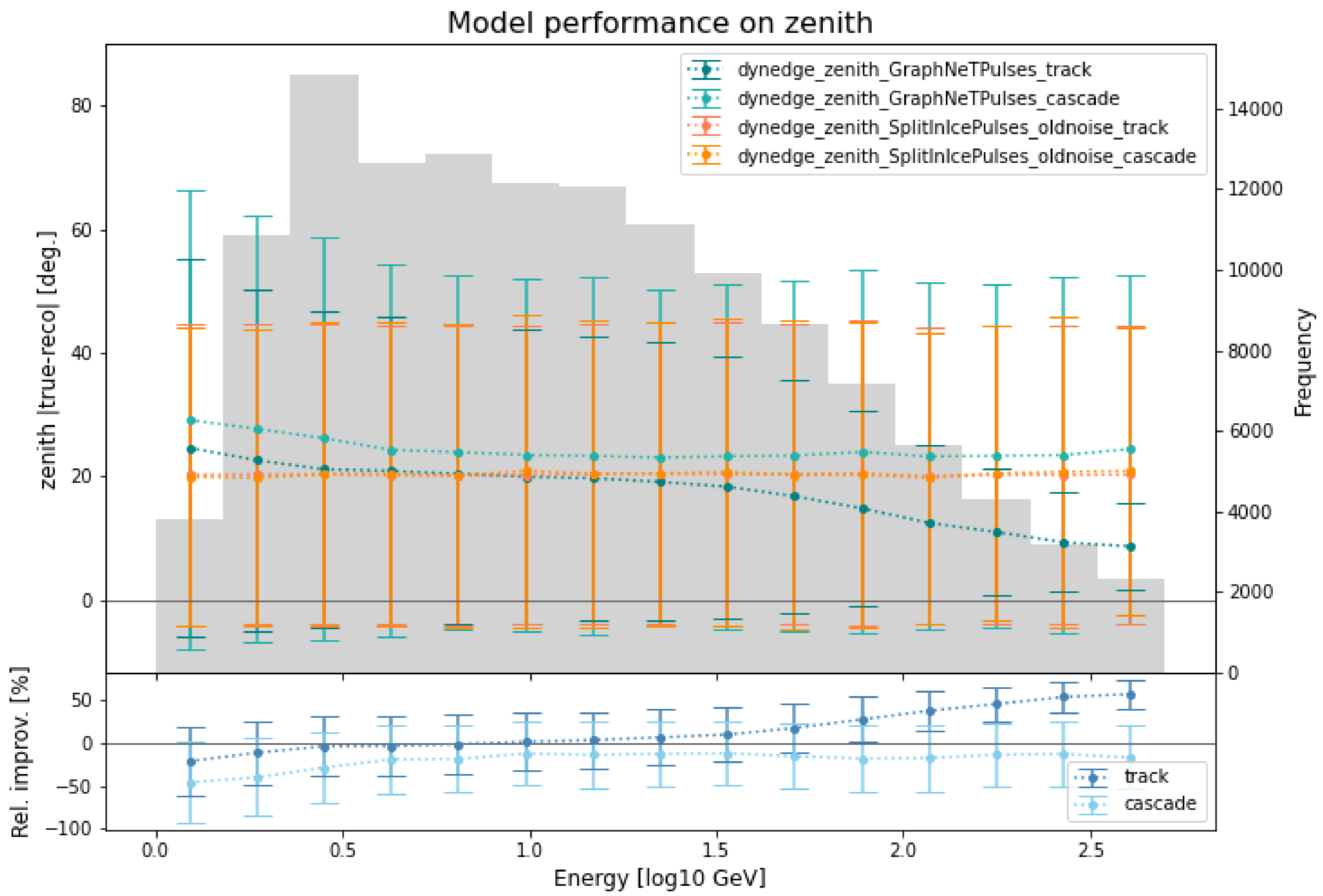


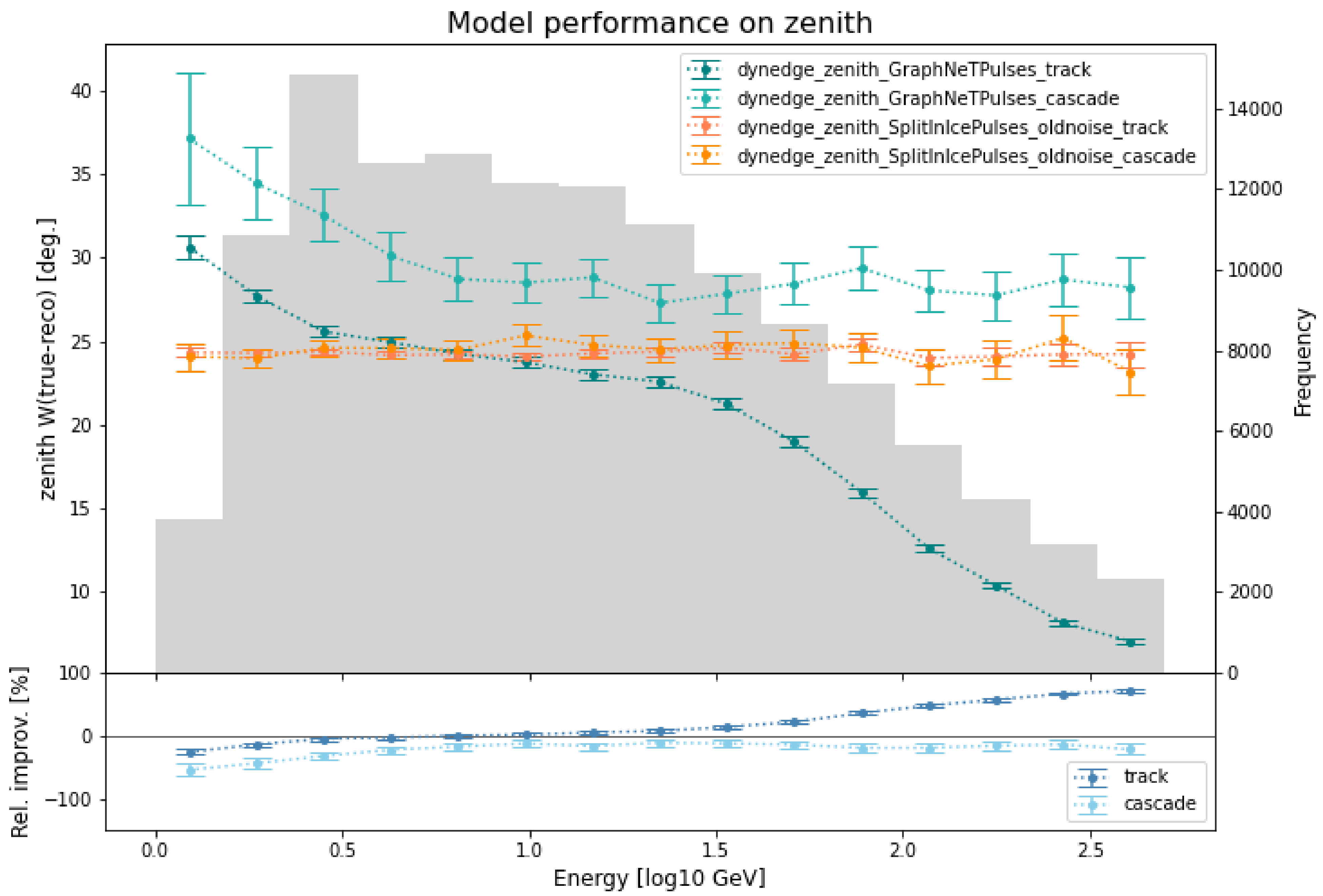


Regression of Cleaned Pulses

Approach: Train regression models on cleaned data and uncleaned data with the old and new noise models.

Zenith





Conclusions:

Pulse cleaning improves reconstruction performance so that it is comparable to the old noise model.

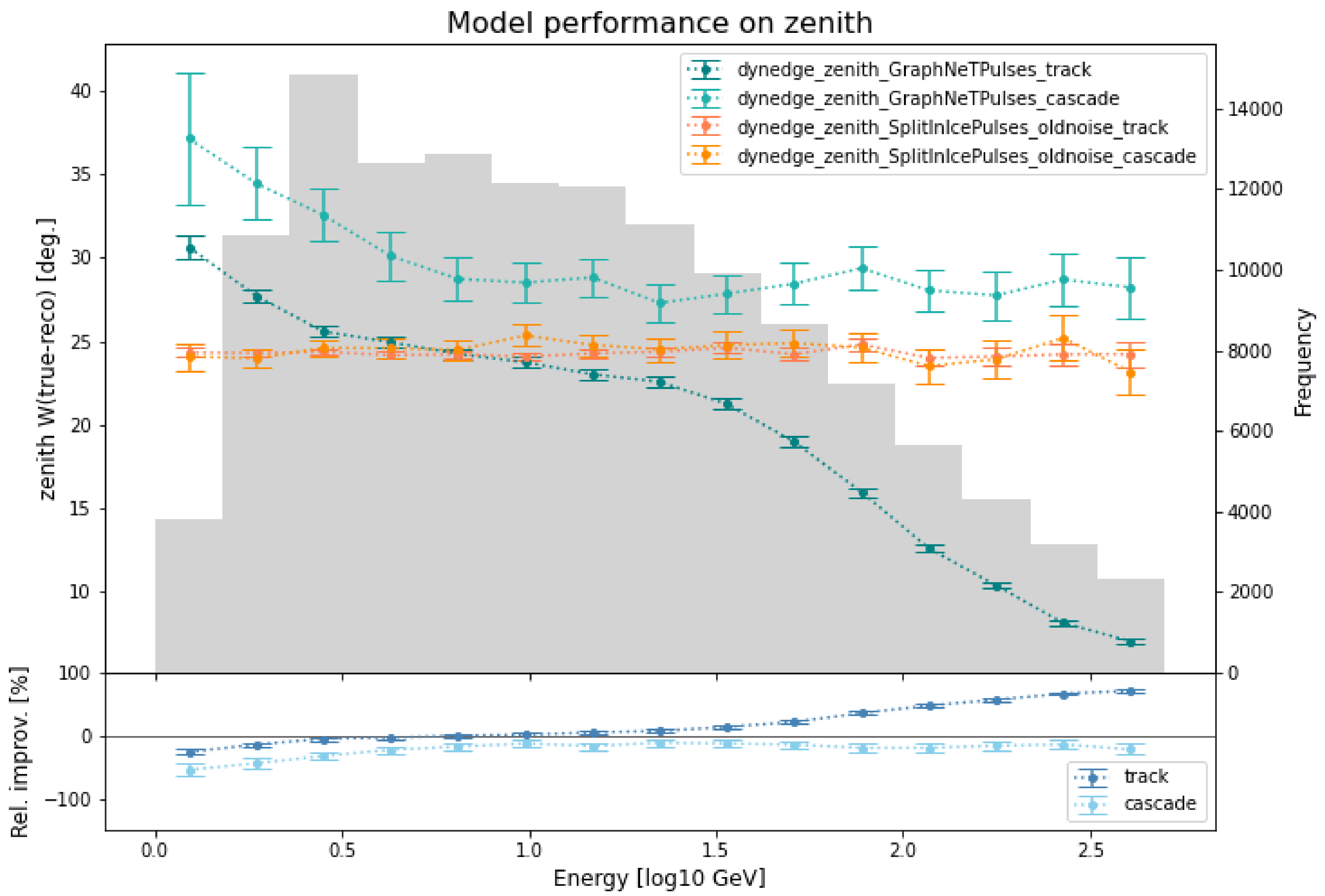
Reconstruction is still lacking in certain ranges, namely the low-energy regime and the cascade-like events.

Intermezzo 1

Revisiting Pulse Cleaning

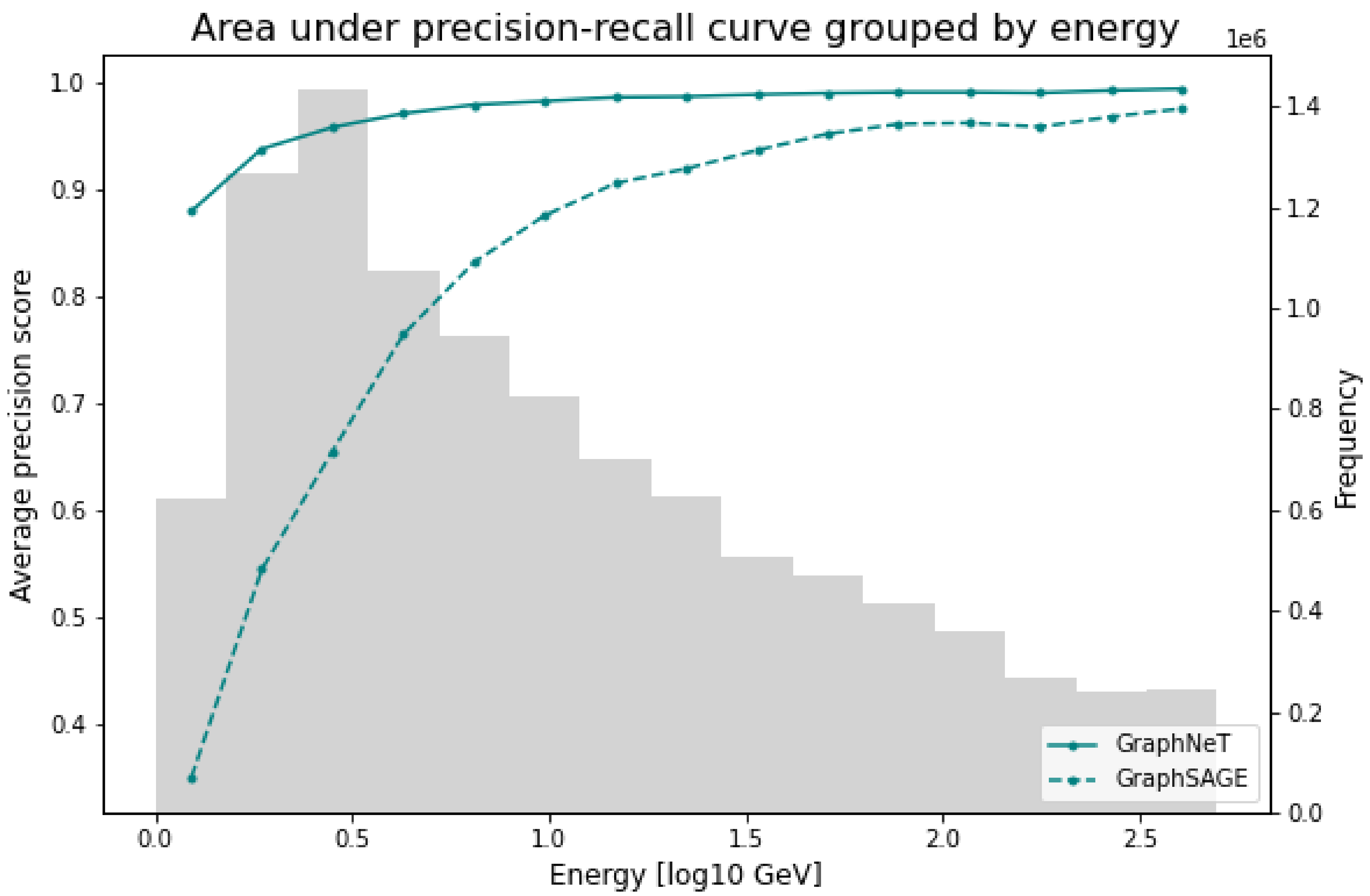
Revisiting Pulse Cleaning

Setup: With reconstruction performance lacking in distinct ranges, improvements to the pulse cleaning or event selection could improve reconstruction.



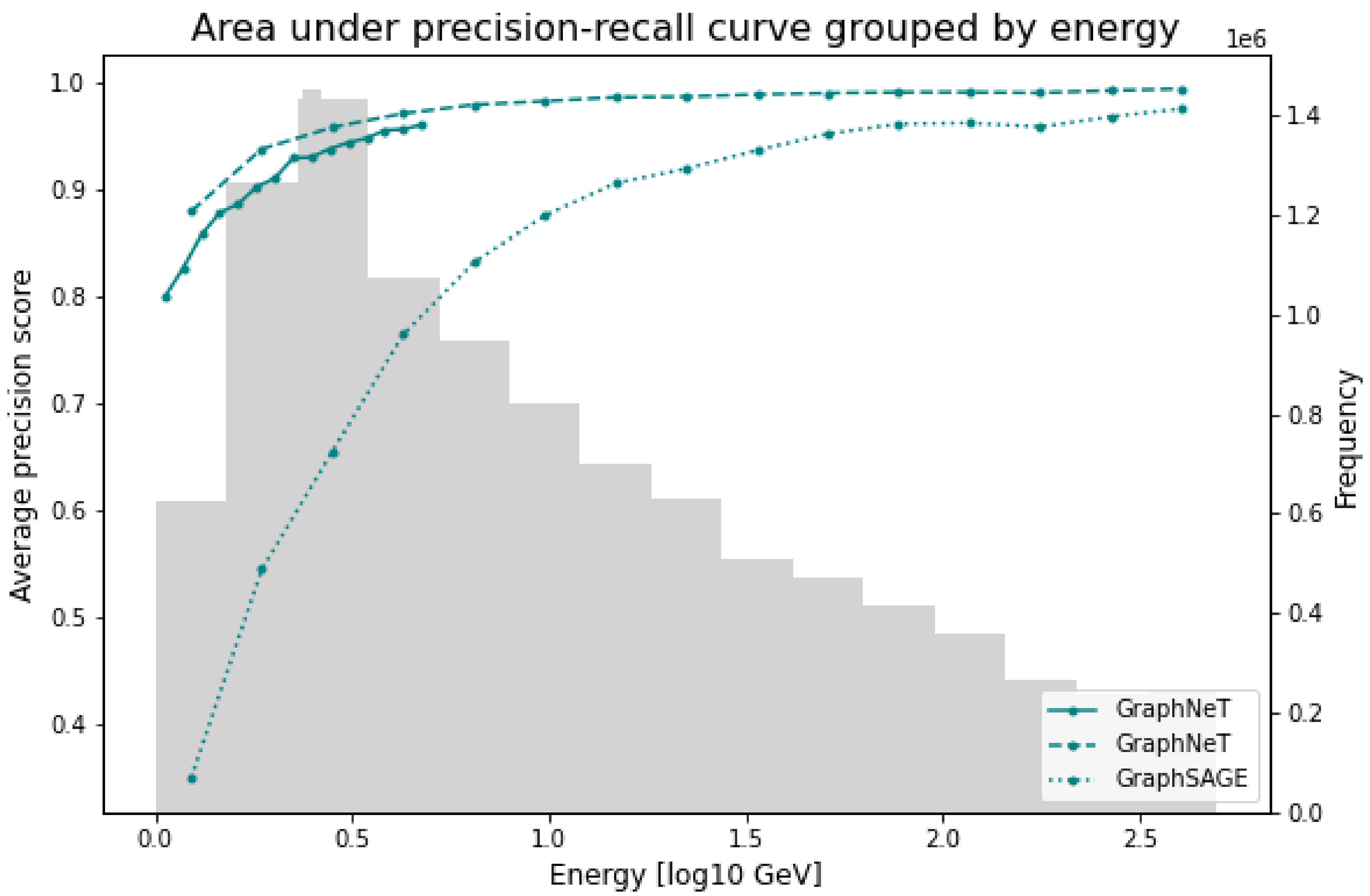
Revisiting Pulse Cleaning

Approach 1: Train a model for low energy events separately



Revisiting Pulse Cleaning

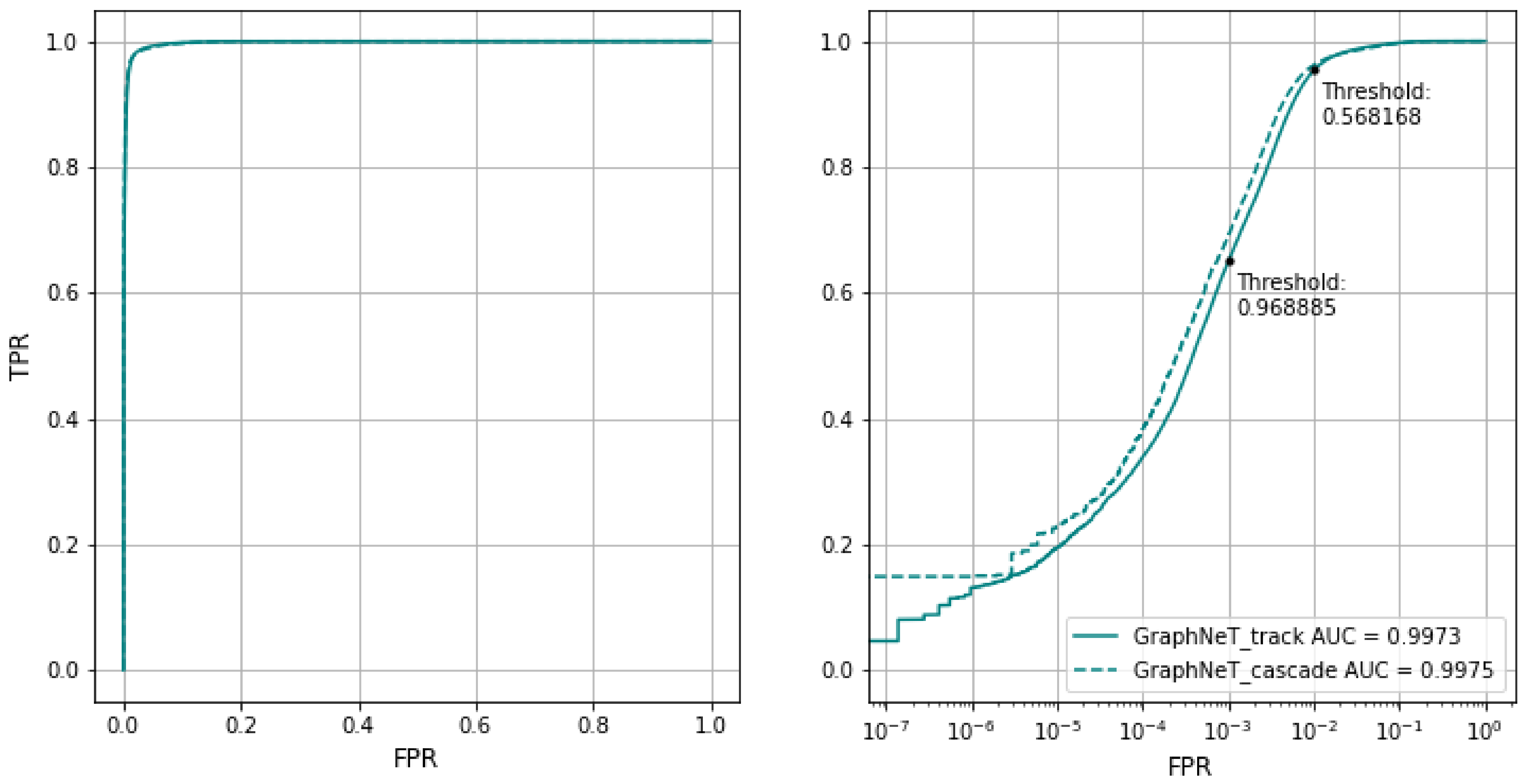
Approach 1: Train a model for low energy events separately



Revisiting Pulse Cleaning

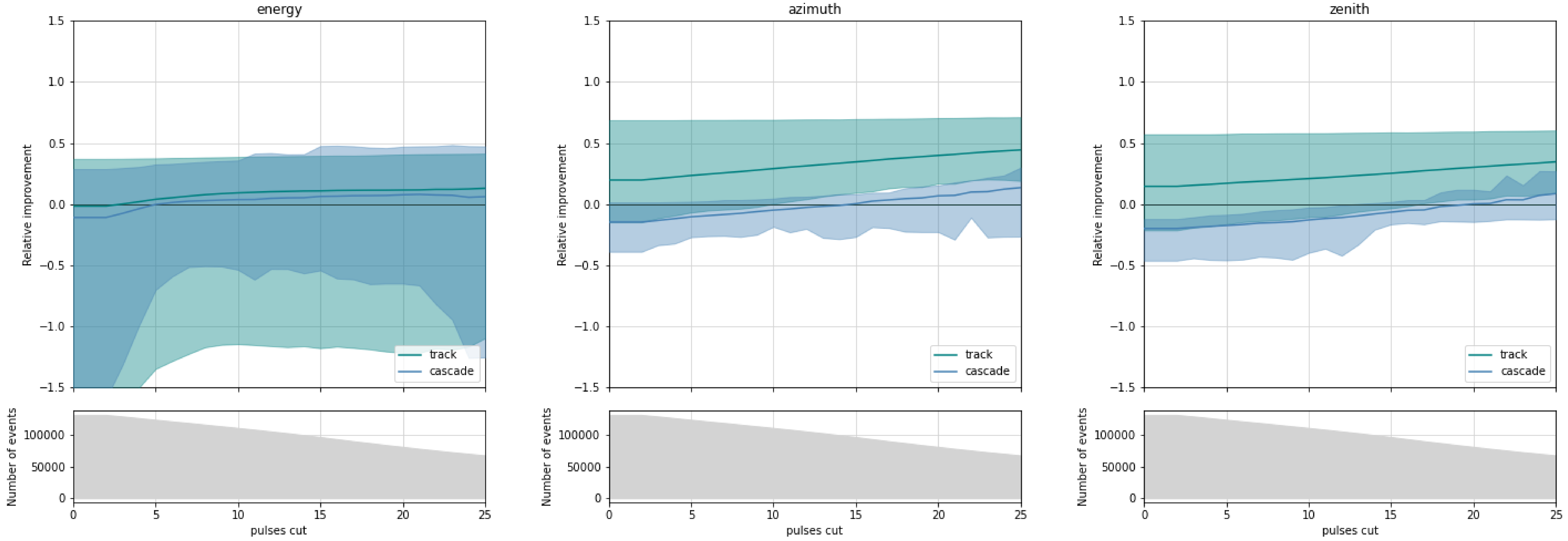
Approach 2: Train a model for track/cascade-like events separately

ROC curve



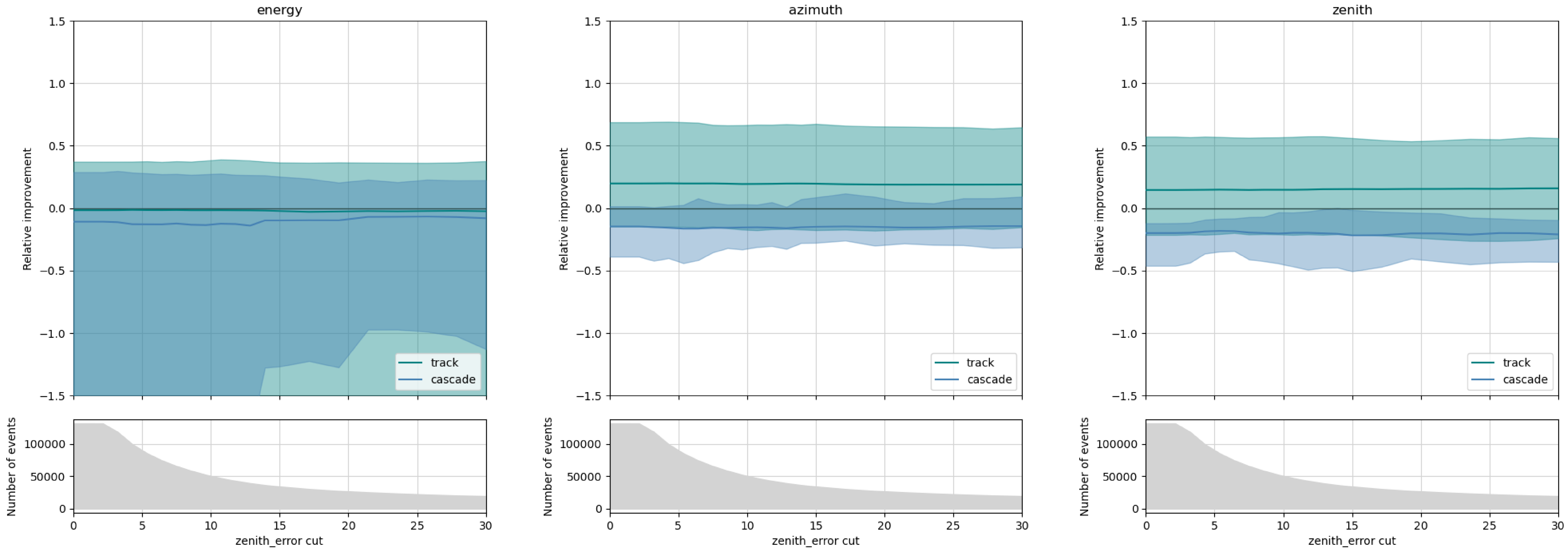
Revisiting Pulse Cleaning

Approach 3: Use number of pulses as a parameter for event selection



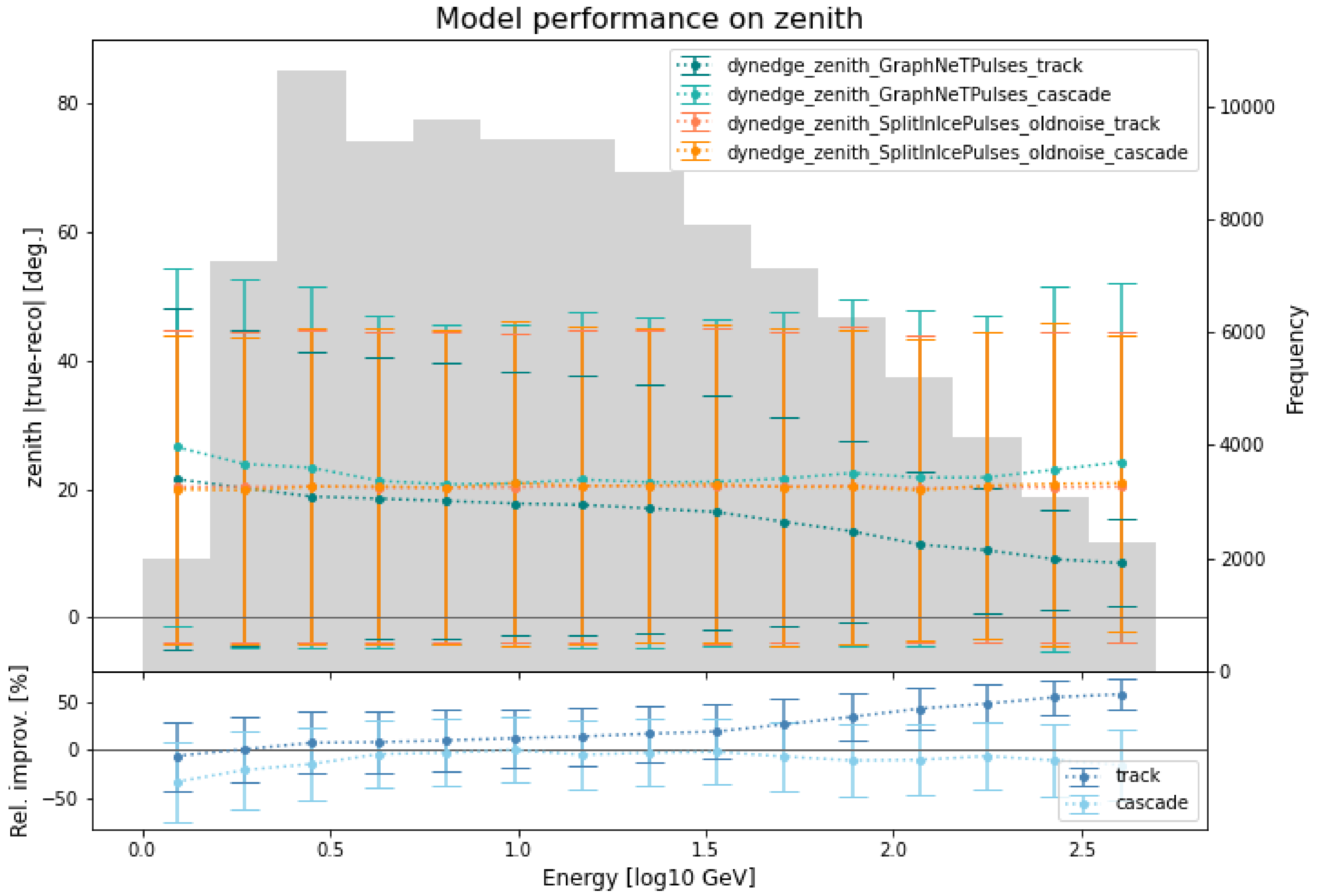
Revisiting Pulse Cleaning

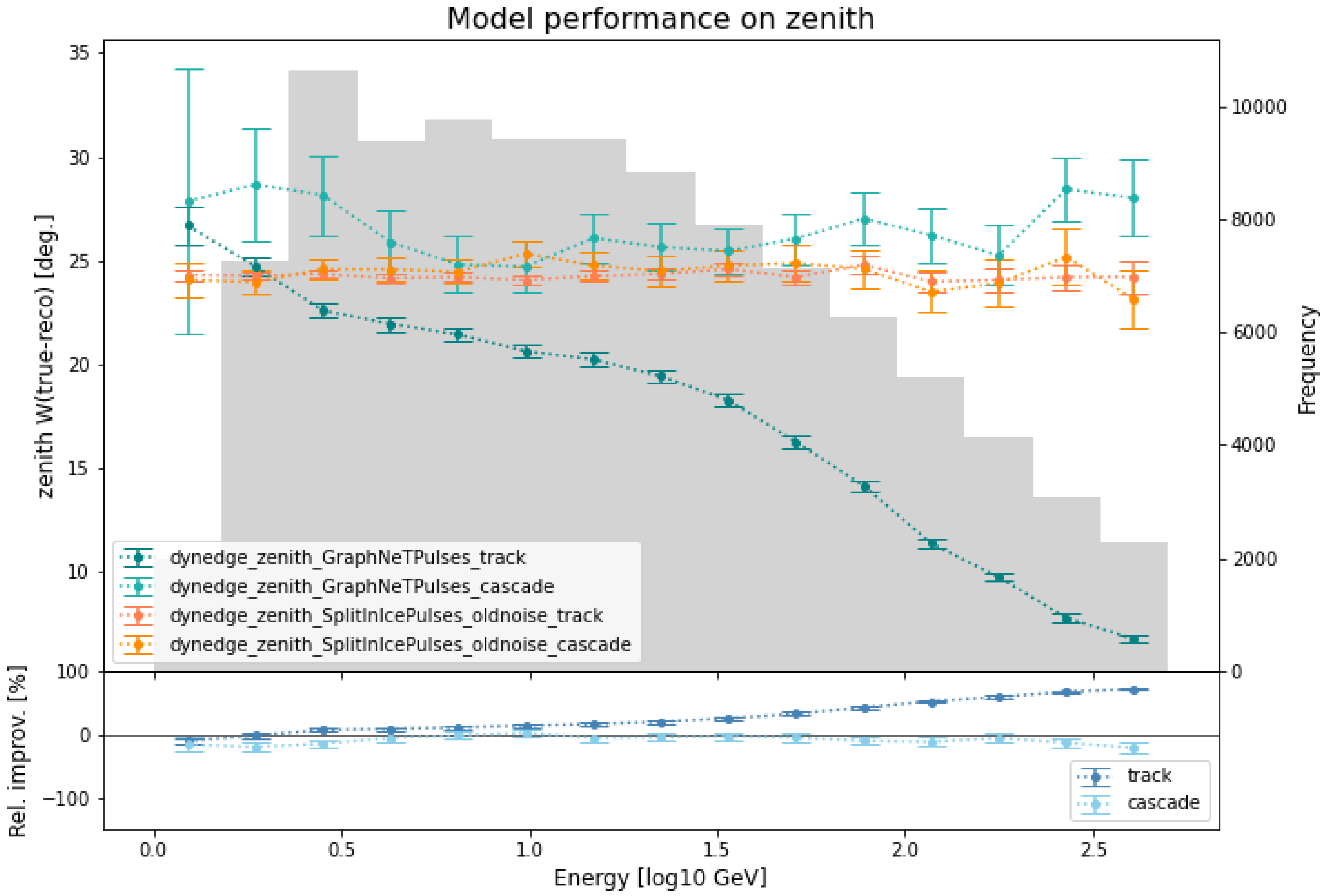
Approach 4: Use zenith kappa estimation as a parameter for event selection



Approach 3: Use number of pulses as a parameter for event selection

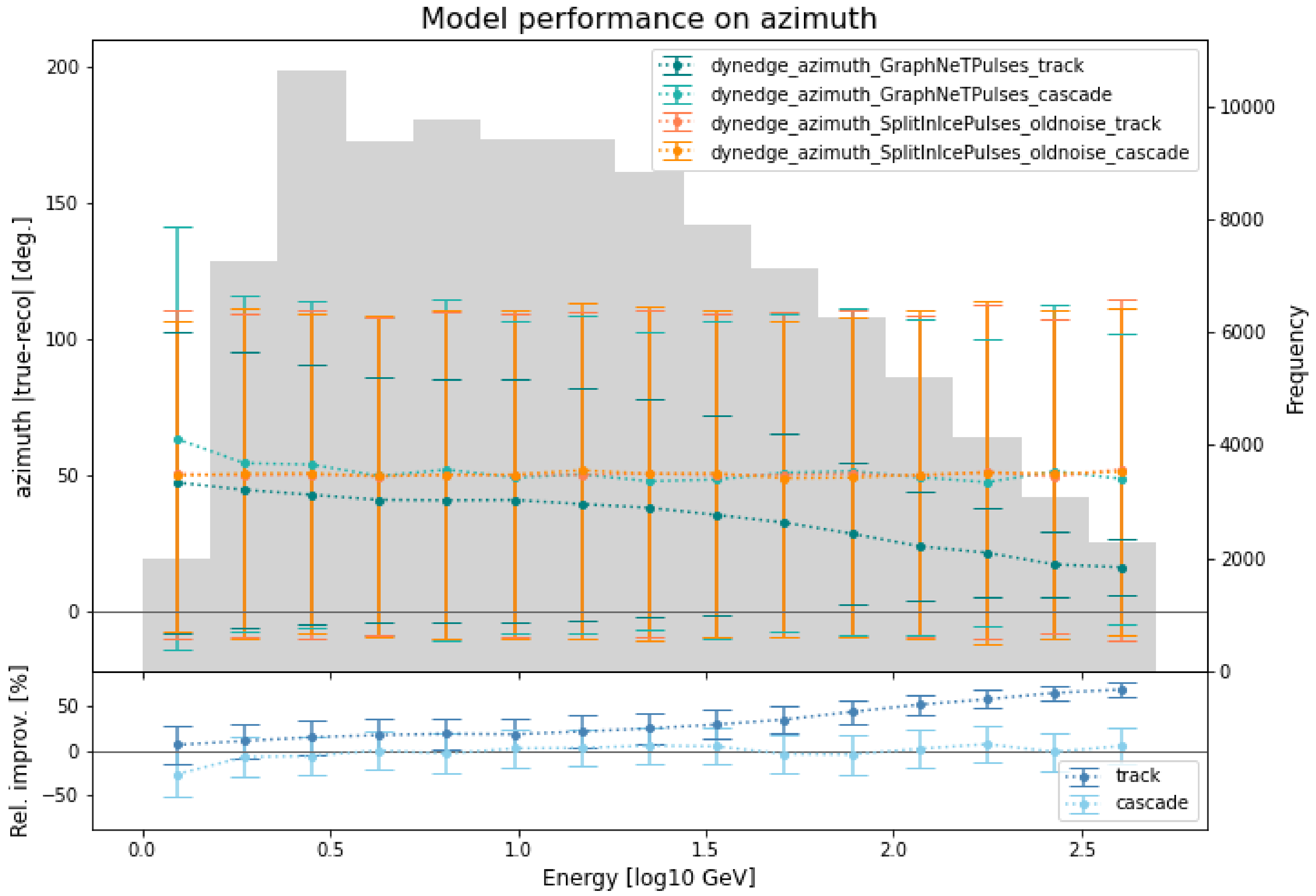
Zenith

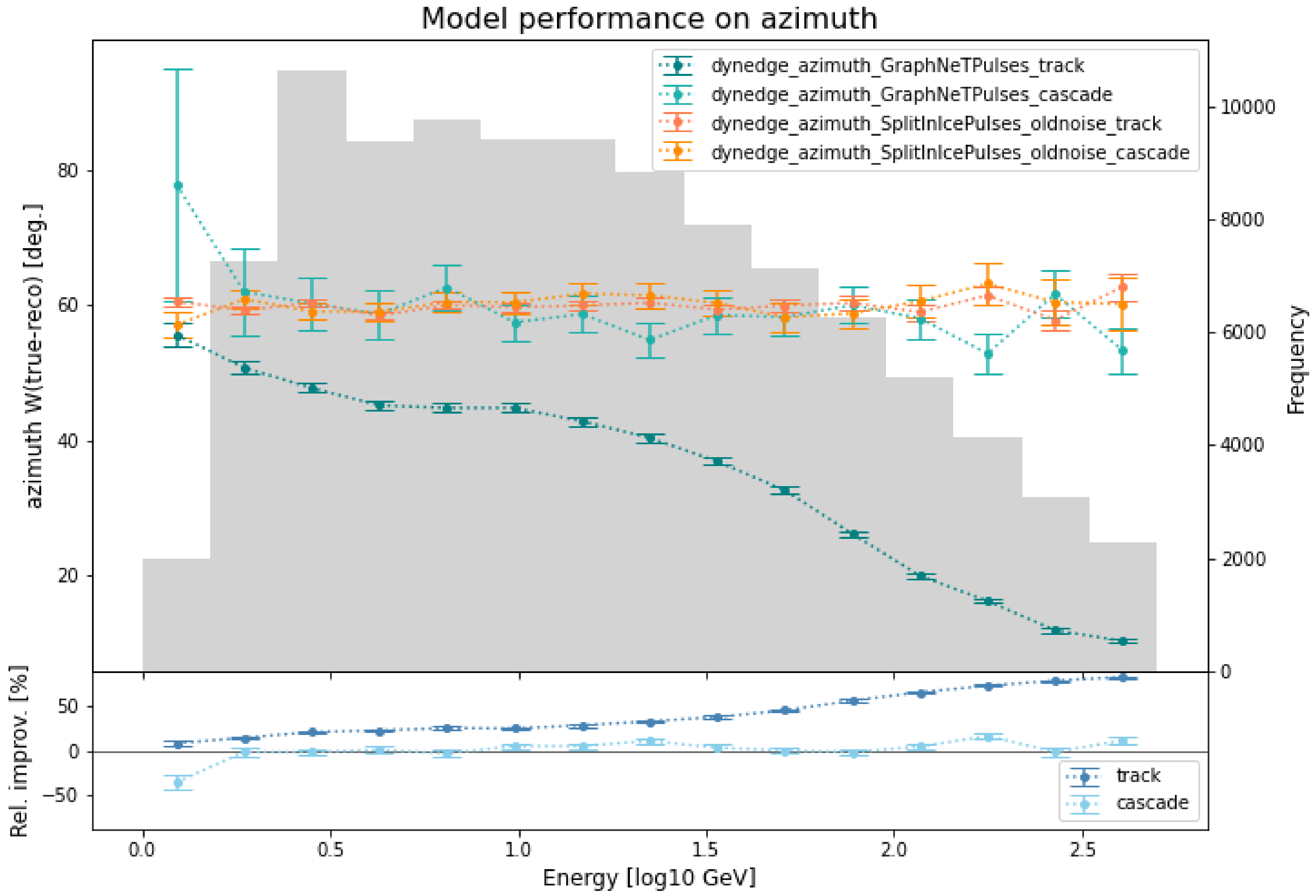




Approach 3: Use number of pulses as a parameter for event selection

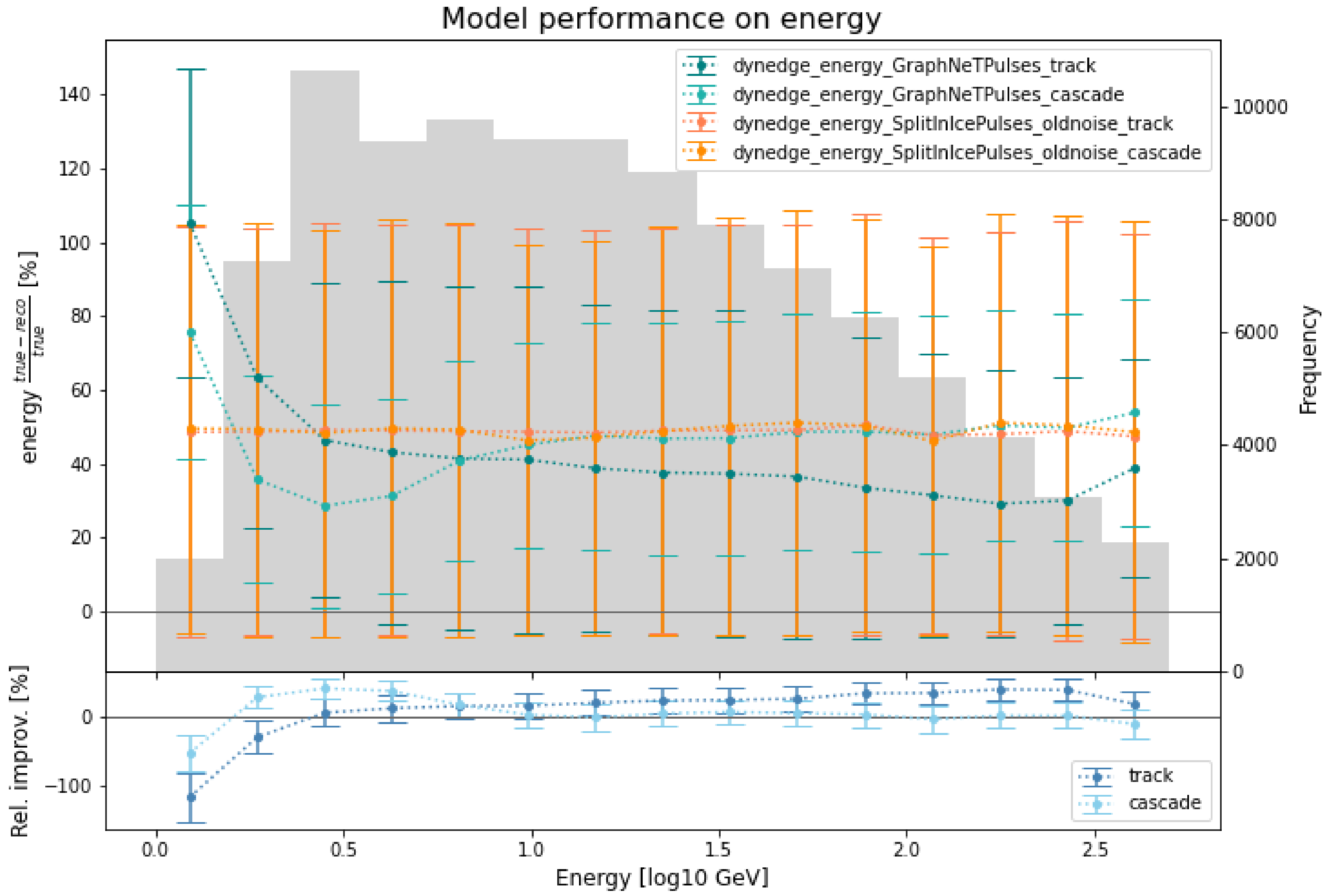
Azimuth

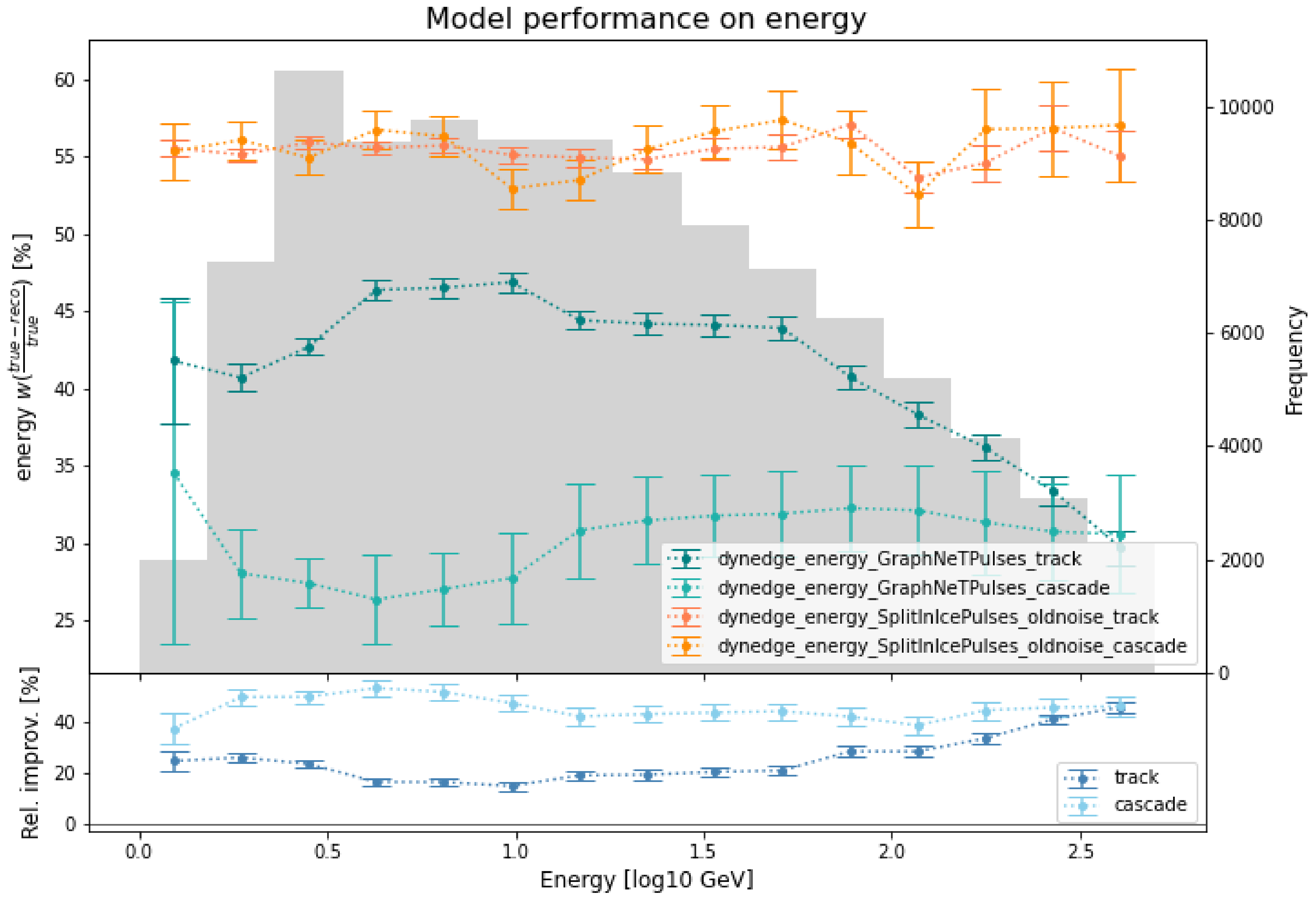




Approach 3: Use number of pulses as a parameter for event selection

Energy





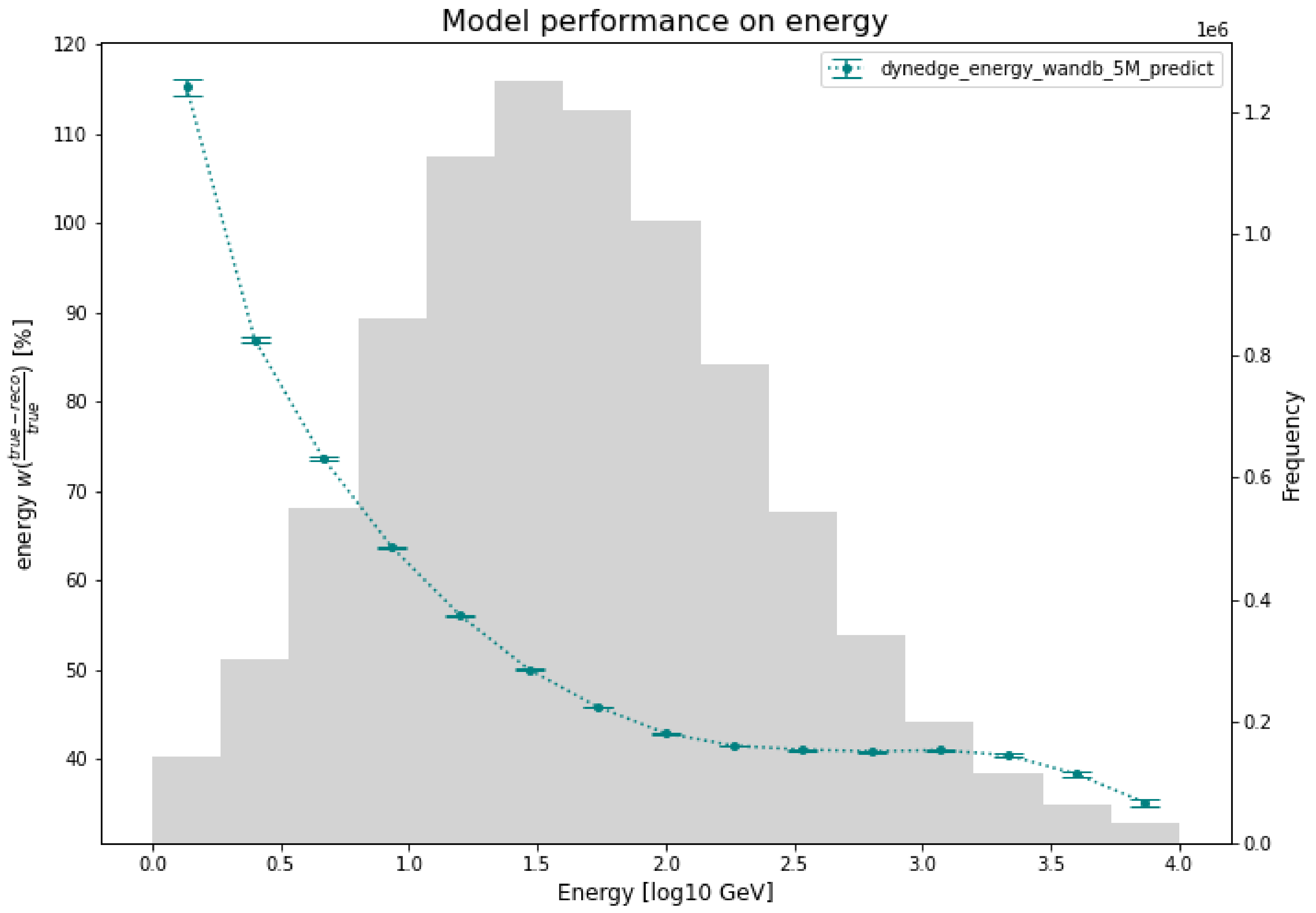
Conclusion: Simple event selection can improve reconstruction performance.

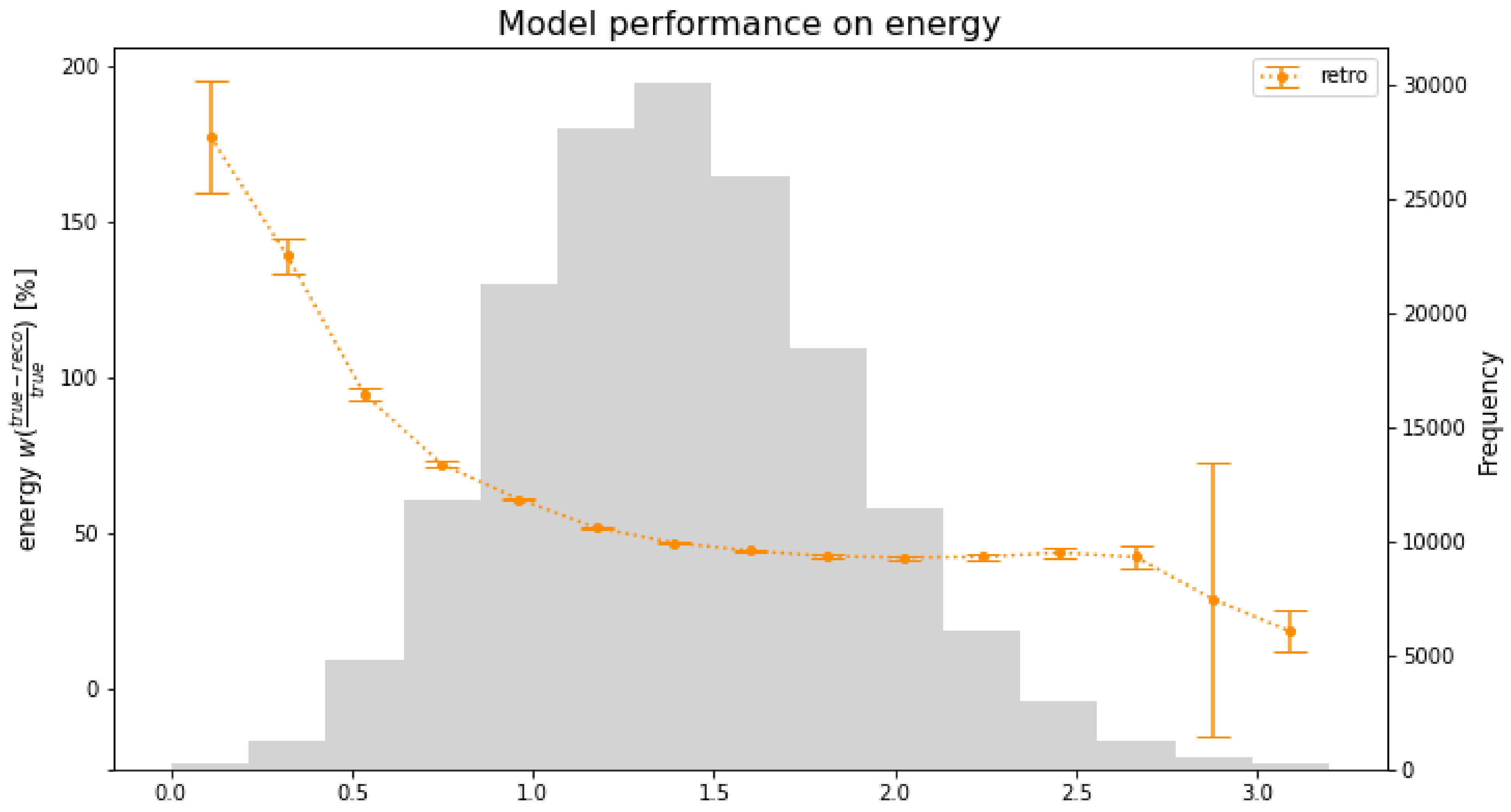
More sophisticated methods like GNNs or current OscNext tools should improve the cleaned events beyond the old noise model.

Intermezzo 2

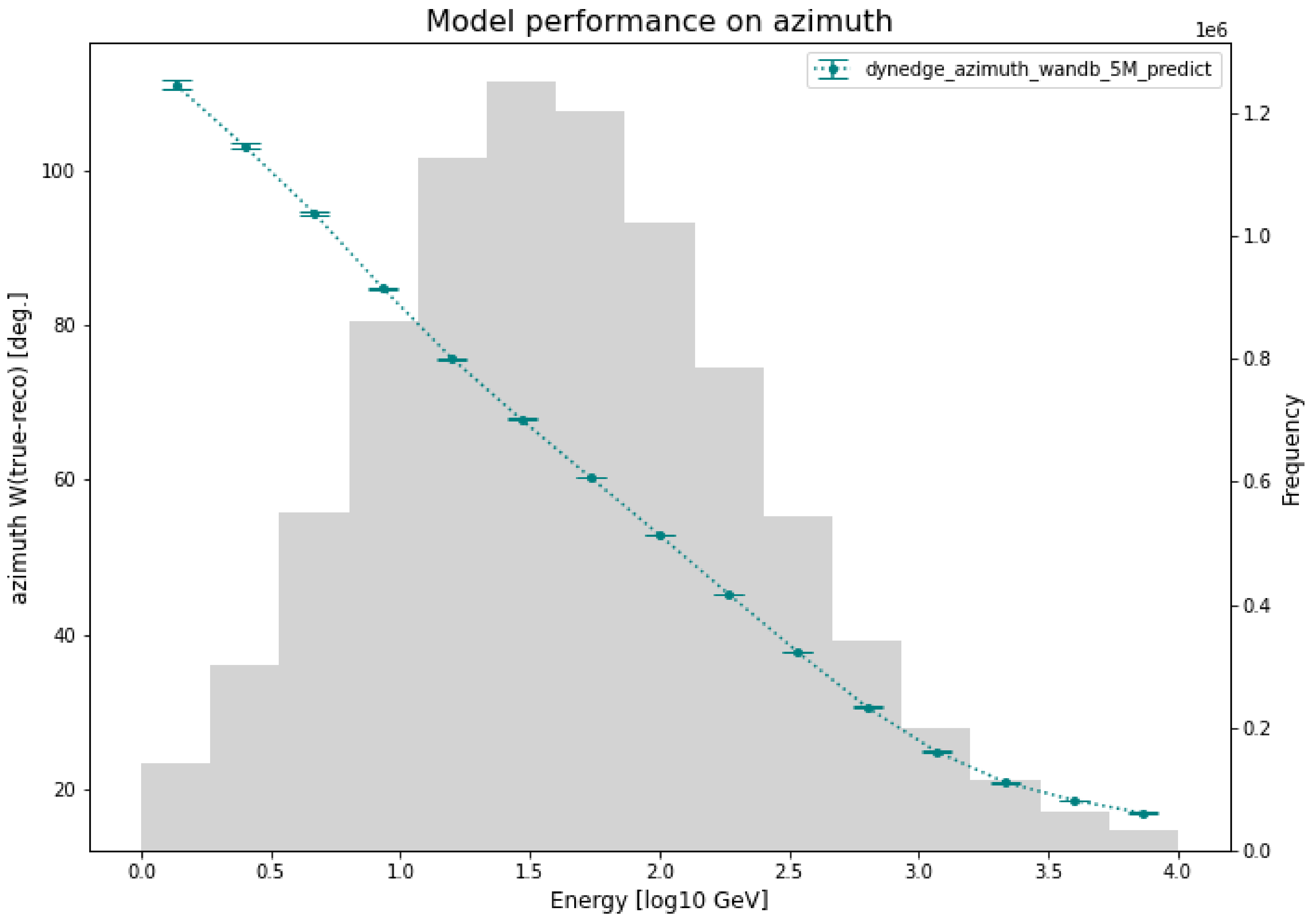
Quality of Cleaned Events

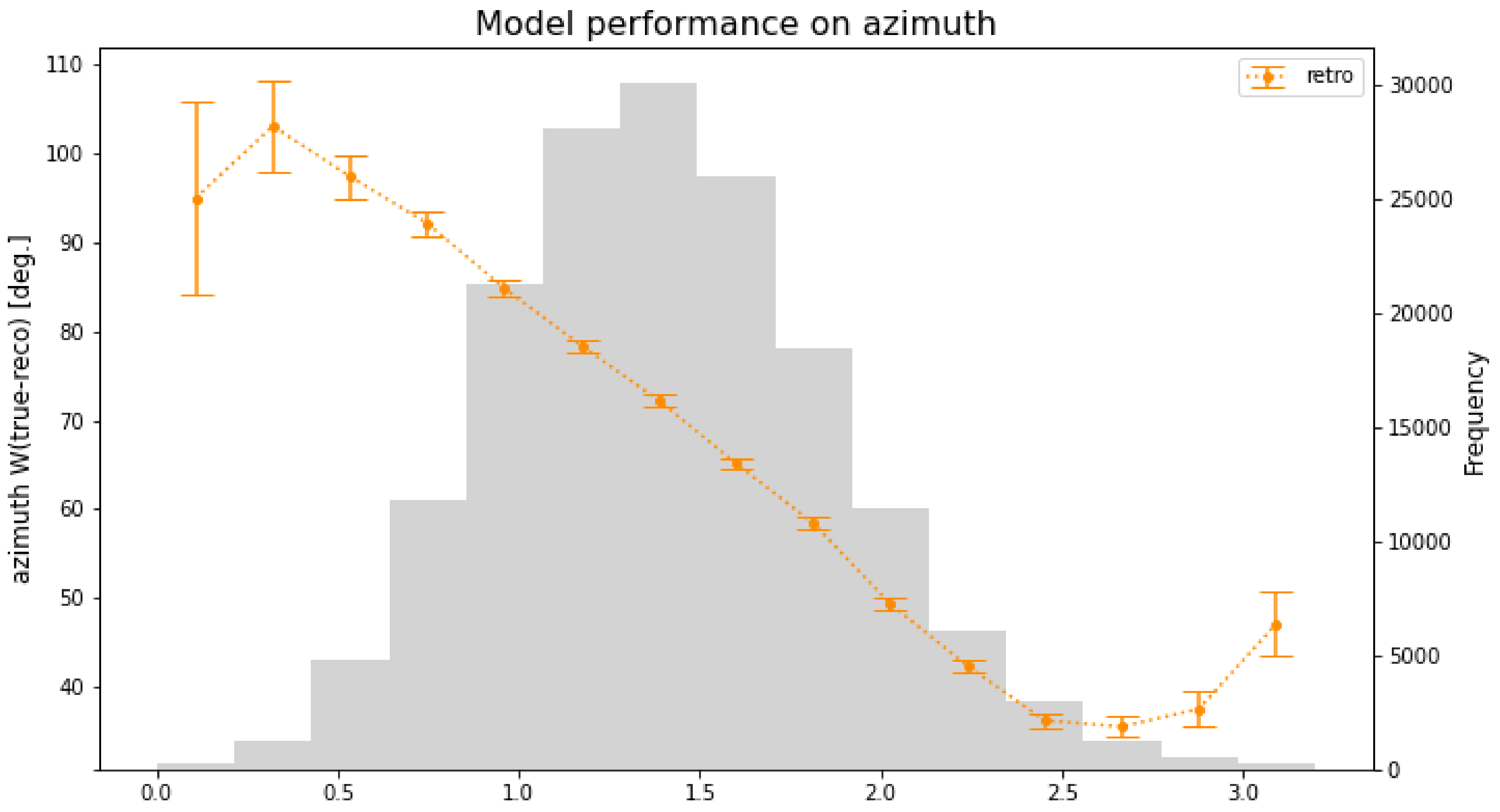
Quality of Cleaned Events



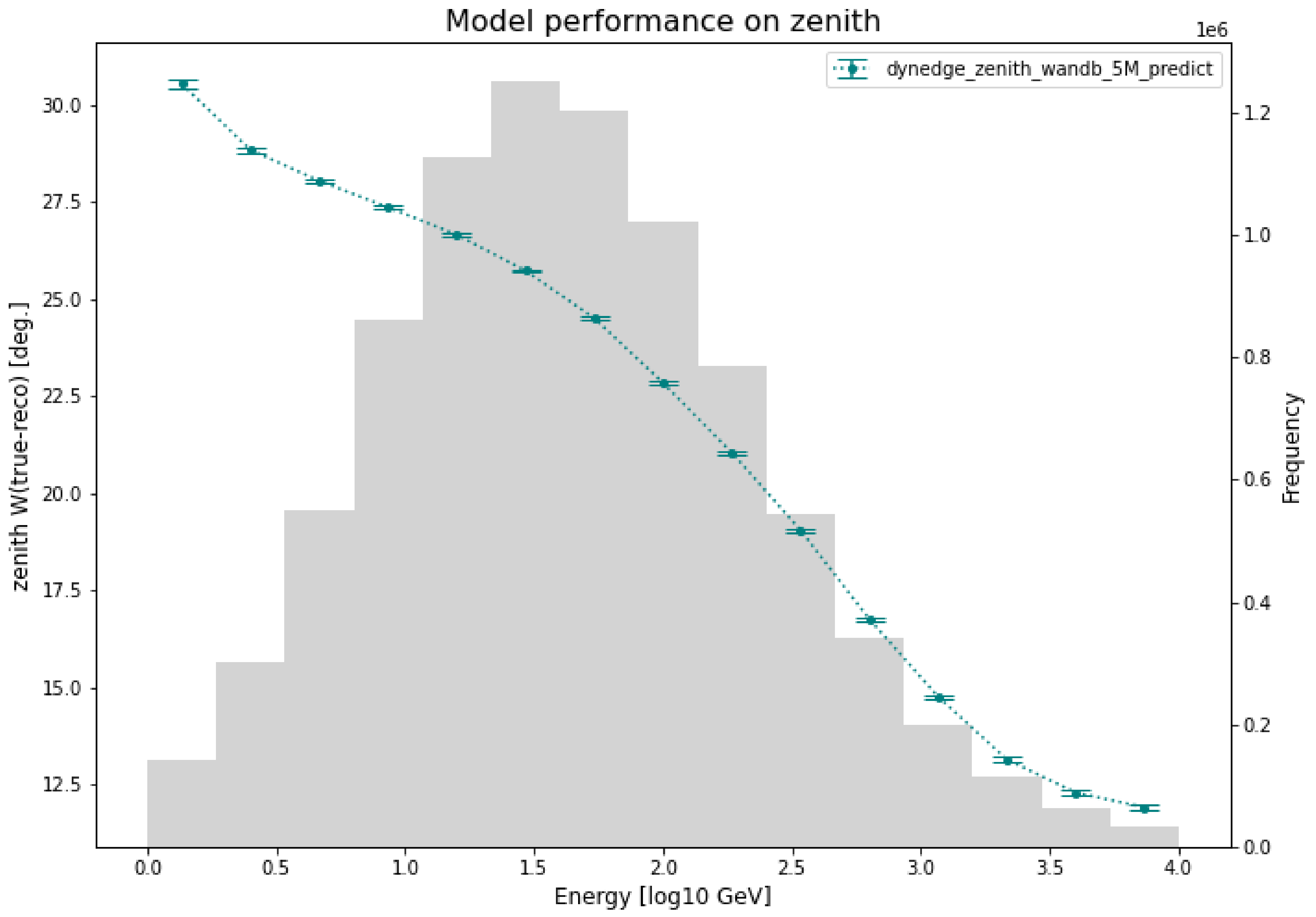


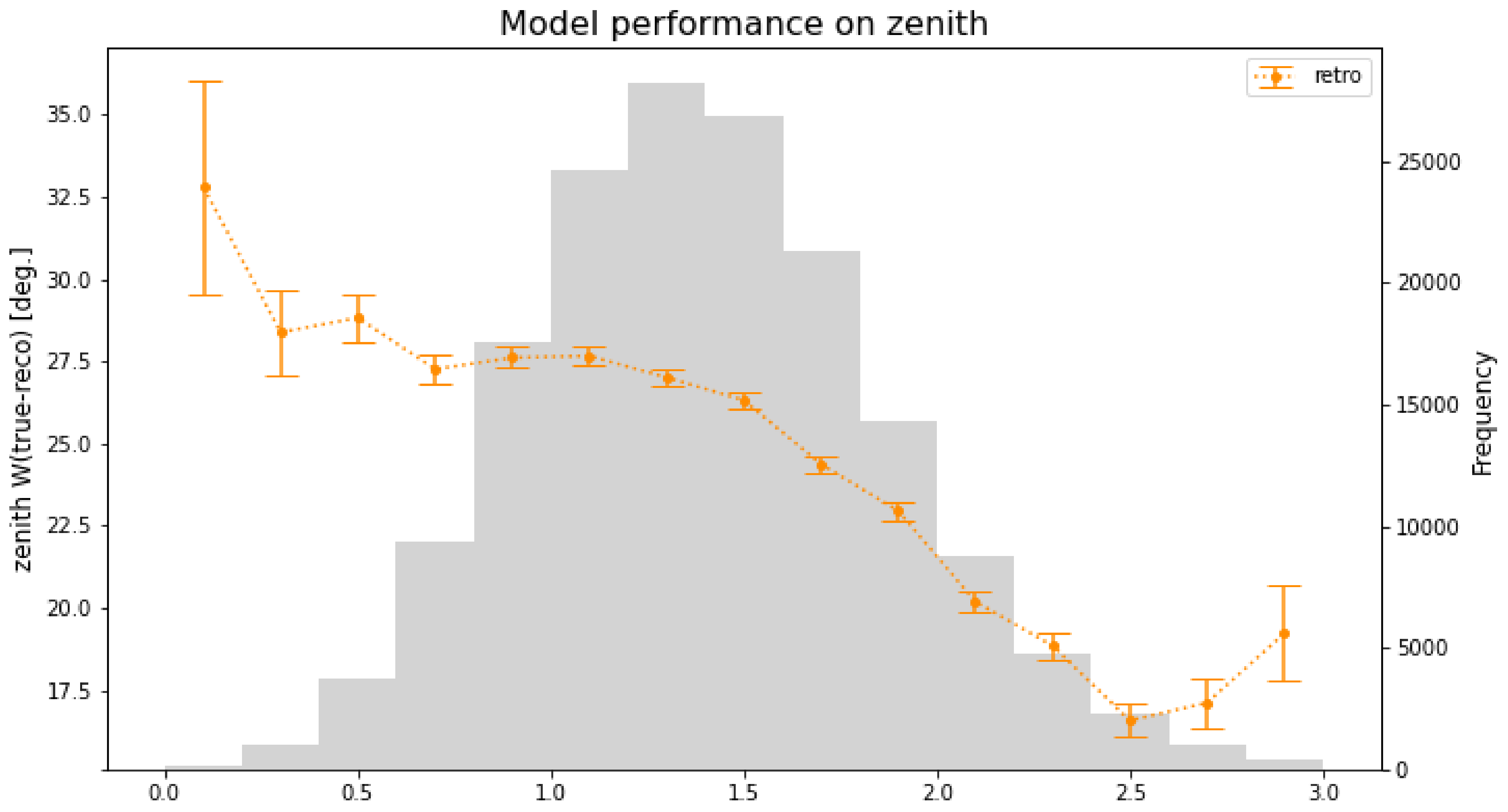
Quality of Cleaned Events





Quality of Cleaned Events





Introduction

My Work

Conclusions & Outlook

Conclusions

Interaction Time Reconstruction

Non-linear transformation can be beneficial for some regression tasks.

It the case of interaction time, it allows the GNN to outperform retro in the desired energy ranges.

Pulse Cleaning

The more sophisticated GNN is better at cleaning the additional Upgrade noise than the simple (but already good) GNN.

Event Cleaning Pipeline

GNNs allow for higher effeciency and purity rates than the current OscNext cleaning pipeline.

+ The new selection of events are well suited for regression with the GNN

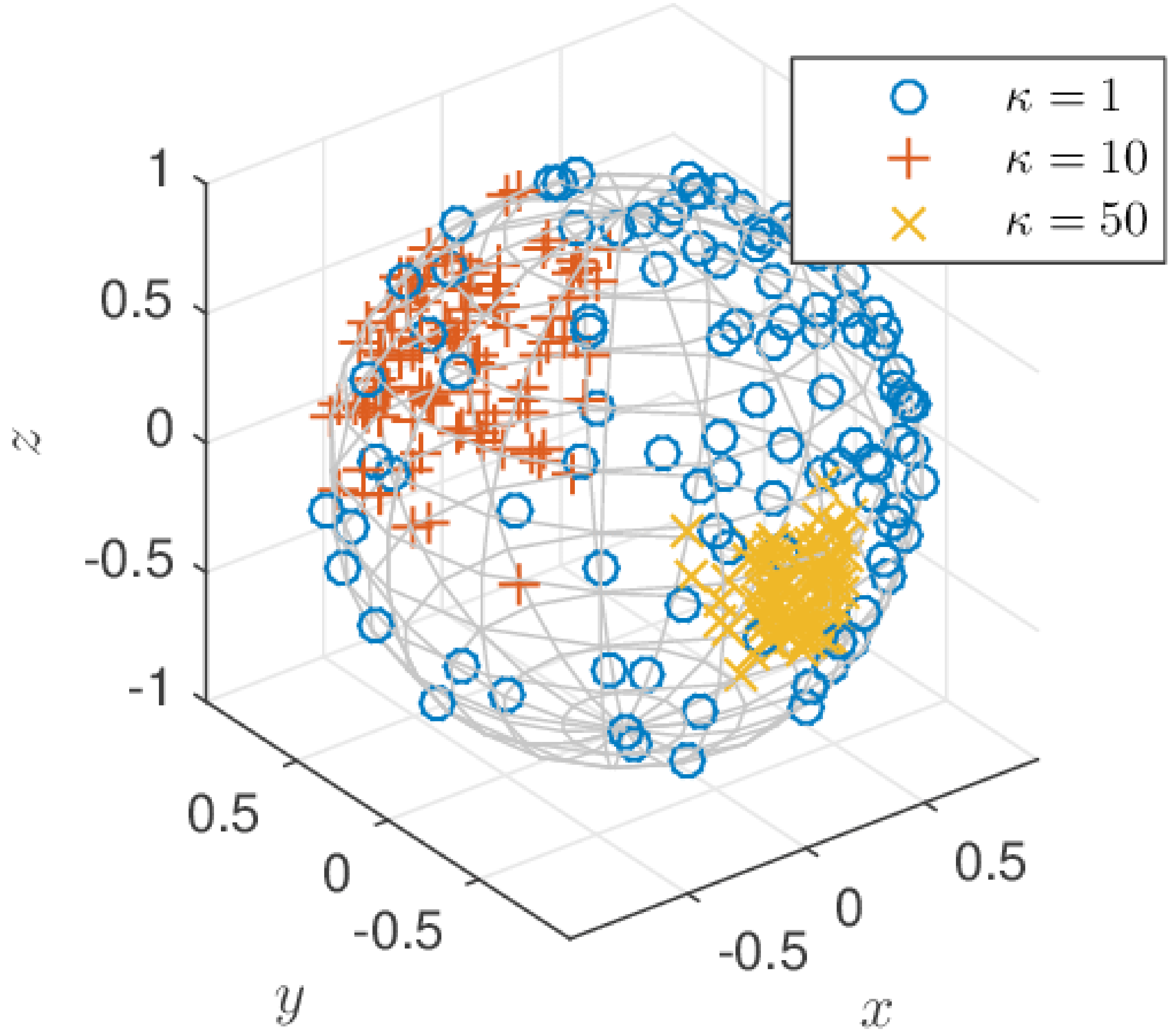
Regression of cleaned pulses

Reconstruction of cleaned pulses nullify the effect of the extra noise in many byt not all cases.

+ Applying a simple event selection criterion makes this true in all but a few ranges.

Thank you!

Von Mises-Fisher Loss



$$p_n(\boldsymbol{x}|\boldsymbol{\mu},\kappa) = C_n(\kappa)\exp\left(\kappa\boldsymbol{\mu}^T\boldsymbol{x}\right)$$

$$\mathcal{L} = -\log p_n(\boldsymbol{x}|\boldsymbol{\mu},\kappa) = -\log(\kappa) + \log(2\pi) + \log(I_0(\kappa))$$