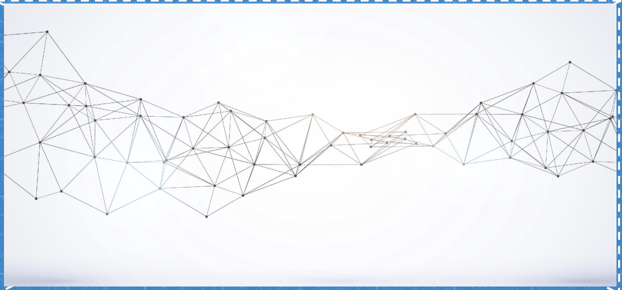
A Graph Neural Network Approach to Low Energy Event Reconstruction in IceCube Neutrino Observatory



Rasmus F. Ørsøe
09-06-2021
pcs557@alumni.ku.dk

Overview

1) The Neutrino and IceCube

The Standard Model, outline of historic experiments, IceCube Detector design, research goals, etc.

2) Graph Neural Networks

Motivation stated through a few fundamental observations, my model.

3) Results

Classification and Regression in MC and in real data

4) Conclusion & Outlook

1

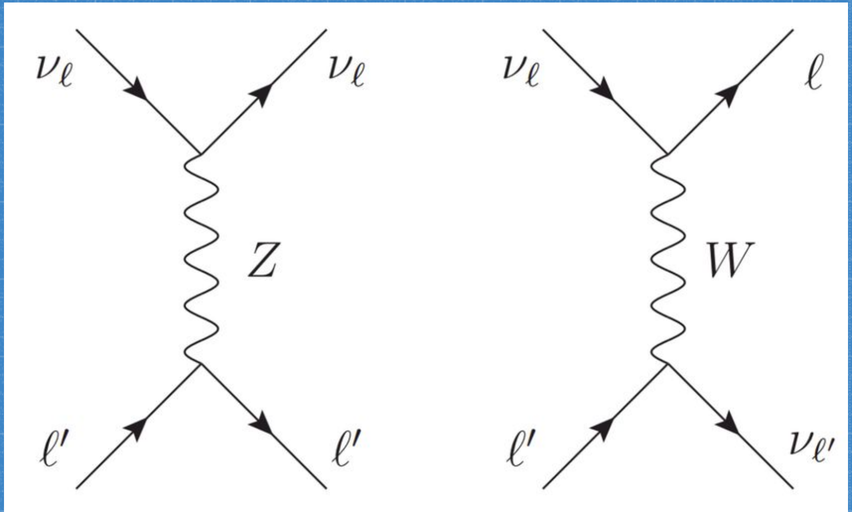
The Neutrino

The Neutrino

- Proposed by Wolfgang Pauli in 1930
- Explained energy spectrum for electrons in beta-decays

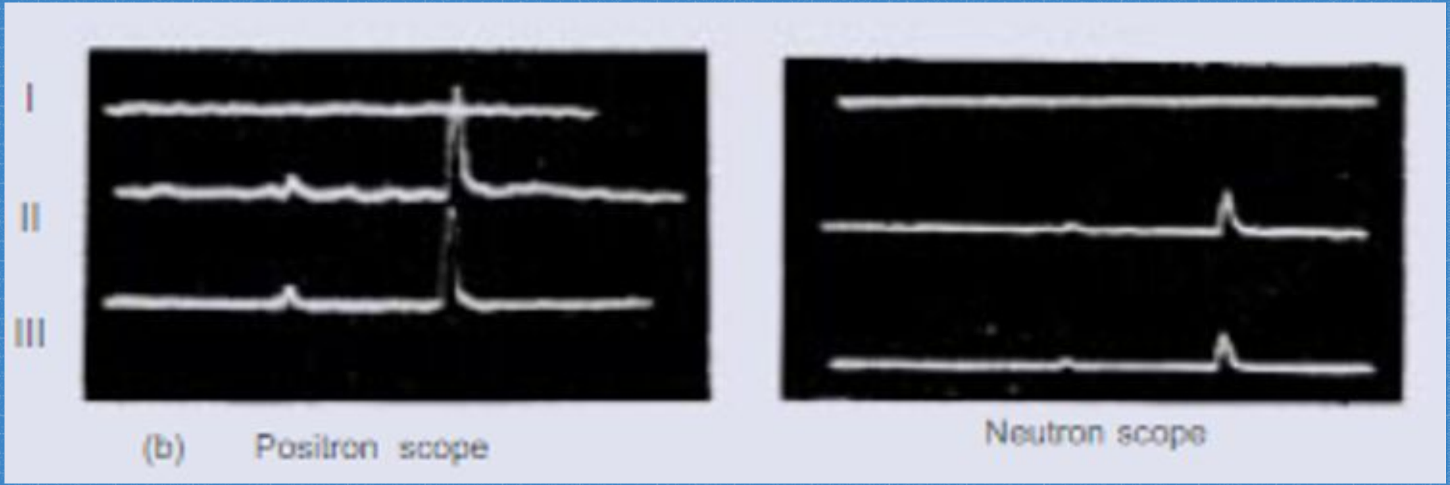
The Neutrino

- Interacts only via weak force (by the standard model)



(Neutral Current) (Charged Current)

Finding The Neutrino



(https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-97-2534-02)

- Cowan-Reines Experiment

  inverse beta decay:

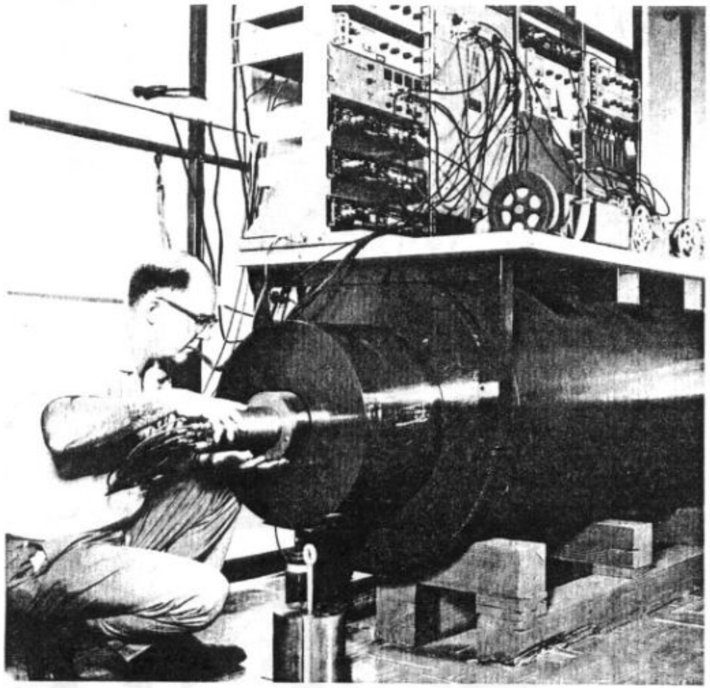
$$\bar{\nu}_e + p \longrightarrow e^+ + n$$

Neutrino Oscillation

- The Homestake Experiment

  Counting solar neutrinos using:

$$\nu_e + Cl_{37} \longrightarrow Ar_{37} + e$$



Dr. Ray Davis of Chemistry is shown placing a low level counter in a cut-down navy gun barrel which acts as a shield from stray cosmic radiation. This equipment is used in the Brookhaven Solar Neutrino Experiment.

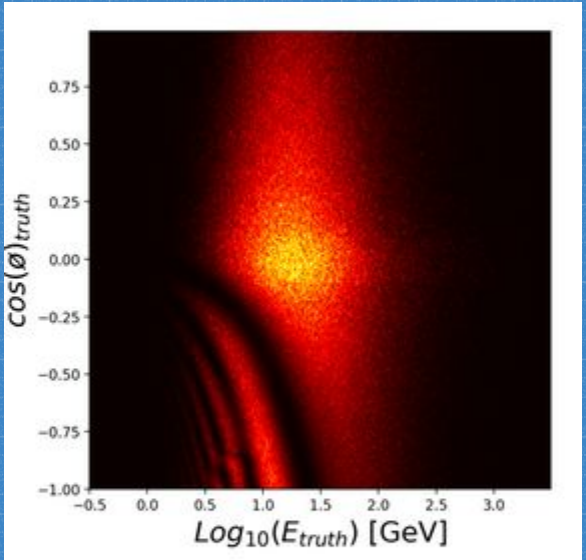
Neutrino Oscillation

$$\begin{pmatrix} v_e \\ v_\mu \\ v_\tau \end{pmatrix} = U_{12}U_{23}U_{13} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$U_{12} = \begin{bmatrix} \cos ø_{12} & \sin ø_{12} & 0 \\ -\sin ø_{12} & \cos ø_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos ø_{23} & \sin ø_{23} \\ 0 & -\sin ø_{23} & \cos ø_{23} \end{bmatrix}$$

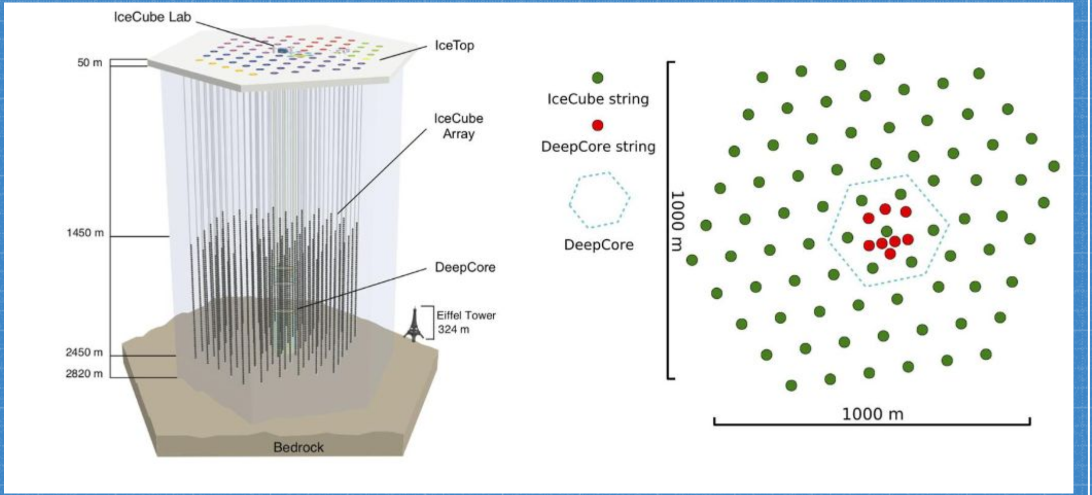
$$U_{13} = \begin{bmatrix} \cos ø_{13} & 0 & \sin (ø_{13}) \cdot e^{-ia} \\ 0 & 1 & 0 \\ -\sin (ø_{23}) \cdot e^{ia} & 0 & \cos ø_{13} \end{bmatrix}$$



2
IceCube

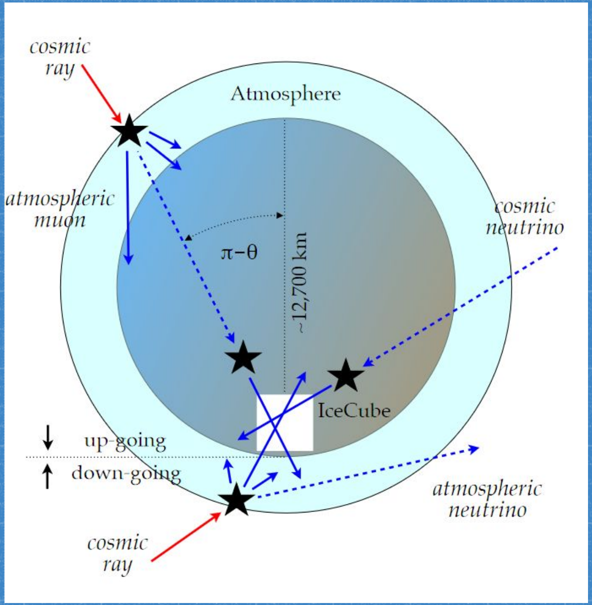
IceCube

- Location: South Pole (In the ice!)
- Largest human-made object by volume
- DeepCore at 2100 - 2450 meters

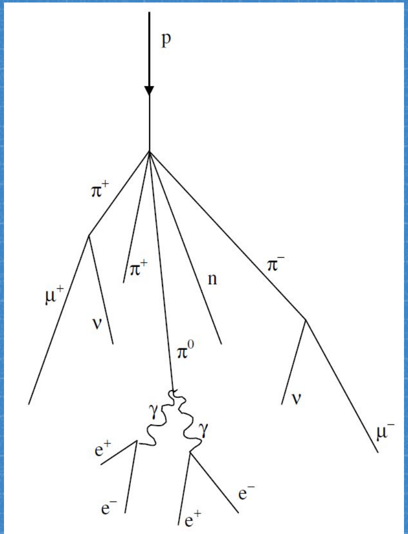


(https://arxiv.org/abs/1806.05696)

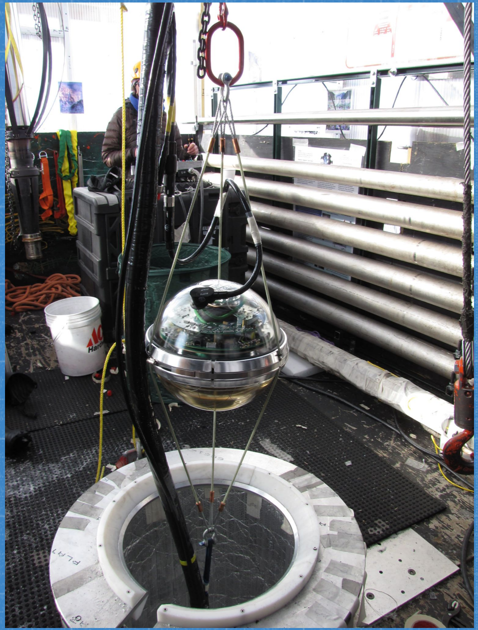
Detecting Neutrinos in IceCube



(https://arxiv.org/pdf/1806.05696.pdf)

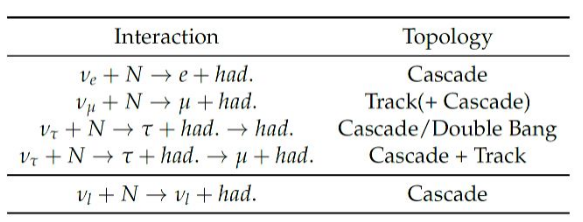


(https://www.researchgate.net/publication/317318317_Near-Space_Muon_Flux_Detection_and_Analysis)



(https://icecube.wisc.edu/gallery/digital-optical-module-dom-development/

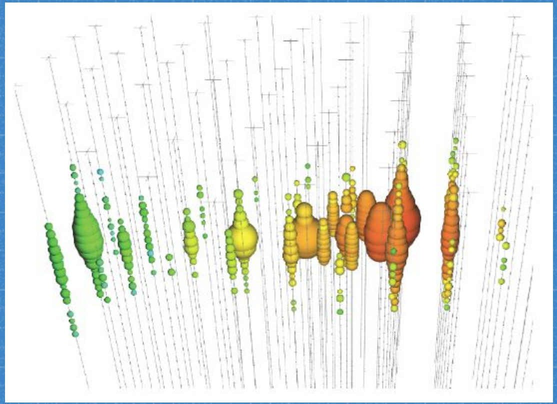
Detecting Neutrinos in IceCube



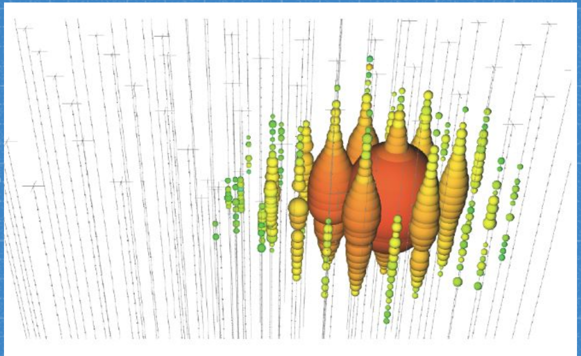
Interaction	Topology
$\nu_e + N \to e + had.$	Cascade
$\nu_\mu + N \to \mu + had.$	Track(+ Cascade)
$\nu_\tau + N \to \tau + had. \to had.$	Cascade/Double Bang
$\nu_\tau + N \to \tau + had. \to \mu + had.$	Cascade + Track
$\nu_l + N \to \nu_l + had.$	Cascade

(https://arxiv.org/pdf/1311.4767.pdf)

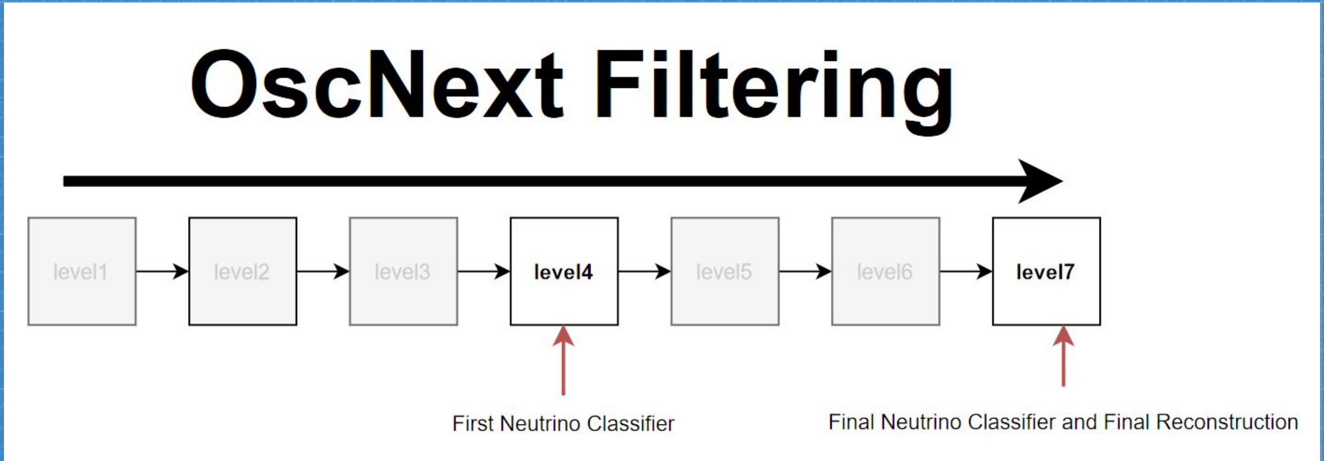
(Track-like)



(Cascade-like)



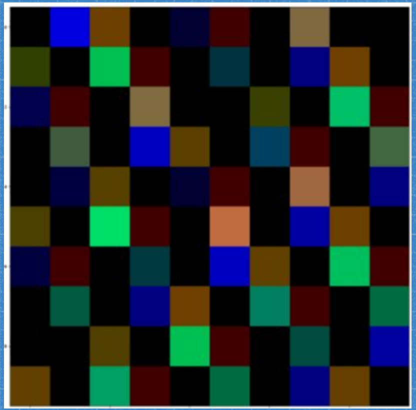
OscNext Filtering Levels



3

Graph Neural Networks

Images at face value



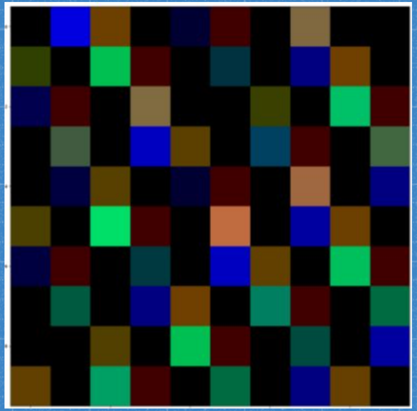
(Image)

img - DataFrame

Index	0	0	0	0	0	0	0	0	0	0
0	[247, 19, 196]	[16, 71, 18]	[252, 62, 228]	[25, 198, 109]	[111, 138, 186]	[93, 19, 197]	[57, 207, 230]	[221, 32, 26]	[150, 119, 104]	[252, 132, 216]
1	[68, 79, 54]	[141, 173, 216]	[180, 112, 97]	[157, 39, 138]	[211, 163, 192]	[70, 171, 211]	[54, 203, 212]	[139, 72, 146]	[18, 5, 162]	[154, 48, 140]
2	[180, 123, 220]	[32, 131, 214]	[67, 83, 145]	[160, 225, 86]	[57, 248, 112]	[171, 157, 126]	[51, 115, 196]	[211, 253, 12]	[183, 115, 67]	[50, 151, 140]
3	[71, 68, 136]	[0, 132, 44]	[253, 86, 25]	[72, 0, 196]	[147, 69, 192]	[163, 44, 171]	[163, 153, 177]	[202, 189, 6]	[167, 252, 168]	[36, 195, 85]
4	[74, 120, 95]	[198, 76, 104]	[234, 62, 158]	[250, 95, 80]	[37, 236, 83]	[88, 137, 88]	[185, 165, 121]	[69, 223, 249]	[174, 86, 95]	[143, 47, 242]
5	[179, 149, 163]	[243, 117, 148]	[70, 186, 169]	[27, 15, 19]	[189, 249, 93]	[164, 253, 220]	[151, 33, 16]	[188, 39, 51]	[21, 27, 52]	[88, 36, 247]
6	[3, 70, 62]	[221, 207, 85]	[101, 65, 132]	[220, 57, 93]	[81, 79, 172]	[156, 242, 64]	[254, 101, 204]	[217, 22, 242]	[46, 242, 77]	[84, 171, 211]
7	[209, 204, 203]	[207, 158, 158]	[212, 248, 11]	[221, 112, 140]	[100, 107, 158]	[118, 122, 252]	[50, 203, 104]	[200, 198, 40]	[254, 227, 104]	[65, 219, 94]
8	[22, 48, 213]	[52, 111, 59]	[215, 155, 231]	[236, 211, 160]	[80, 44, 150]	[22, 202, 149]	[93, 242, 65]	[65, 183, 243]	[43, 91, 98]	[210, 153, 236]
9	[29, 189, 13]	[217, 78, 223]	[27, 104, 151]	[124, 21, 39]	[14, 123, 89]	[18, 33, 23]	[218, 208, 211]	[115, 29, 223]	[26, 171, 241]	[69, 216, 19]

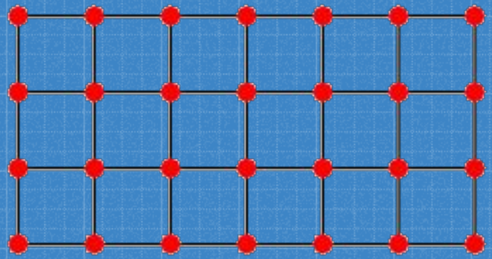
(data)

Images in an abstract sense:



(image)

- Images are grid-like structures, where the distance between neighbouring points in the grid is constant



(image as grid)

- At every point in the grid we associate values ‘red’, ‘green’, ‘blue’

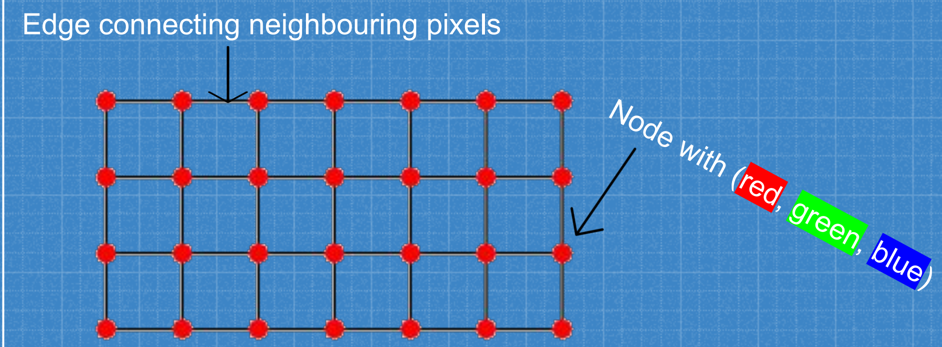
- Images are really a special case of graphs

Graphs

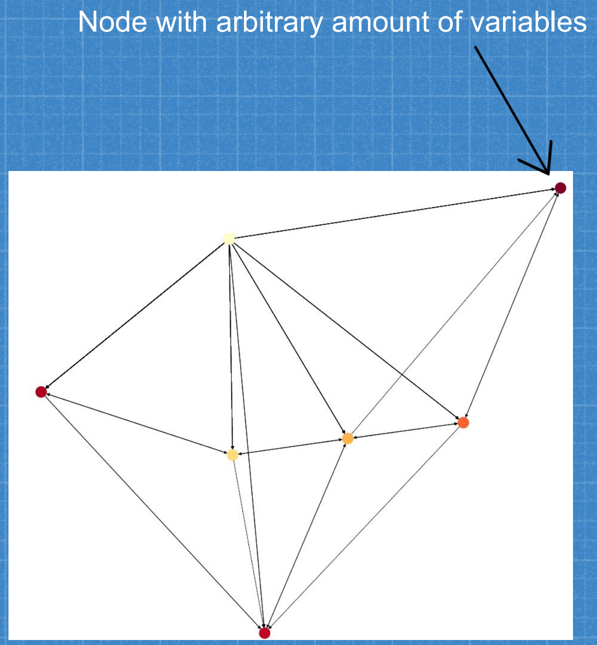
A graph is a generalized image!

Graphs is a collection of

- 1) nodes ("pixels")
- 2) edges (connections between nodes)



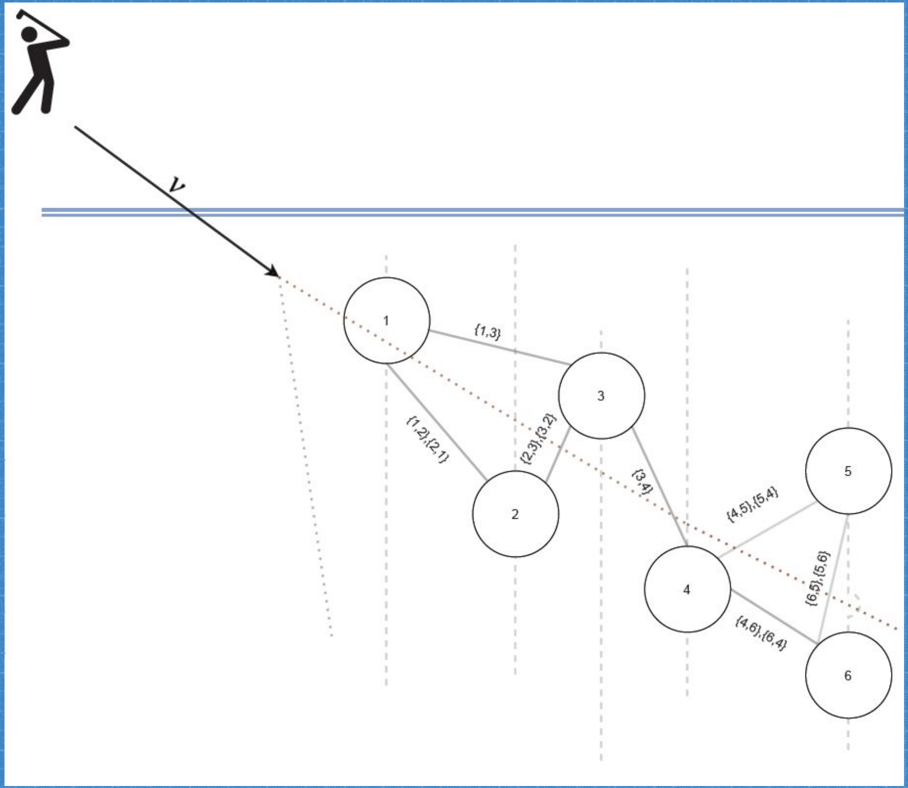
(image as graph)



(IceCube Event as Graph in xy-plane)

Could a graph neural network produce reasonable classification/regression results in IceCube, and if so, how well does it compare to RetroReco?

IceCube Events as Graphs



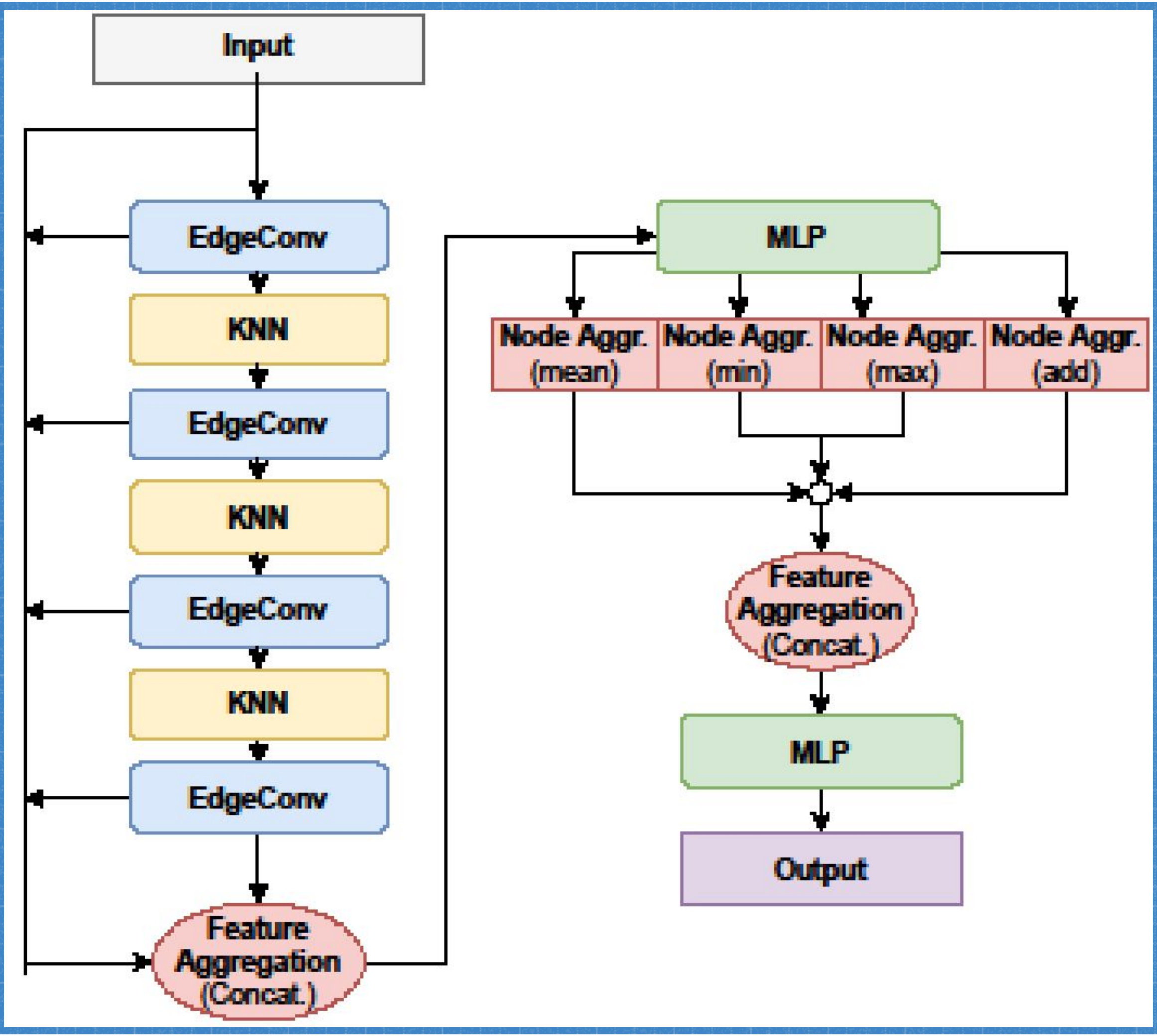
- Pulse + DOM information -> Node
- 8 nearest pulses -> edges

* Multiple pulses from same DOM is included

2

The Model: dynedge

dynedge



Input
EdgeConv
KNN
EdgeConv
KNN
EdgeConv
KNN
EdgeConv
Feature Aggregation (Concat.)
MLP
Node Aggr. (mean)
Node Aggr. (min)
Node Aggr. (max)
Node Aggr. (add)
Feature Aggregation (Concat.)
MLP
Output

EdgeConv

(https://arxiv.org/pdf/1801.07829.pdf)

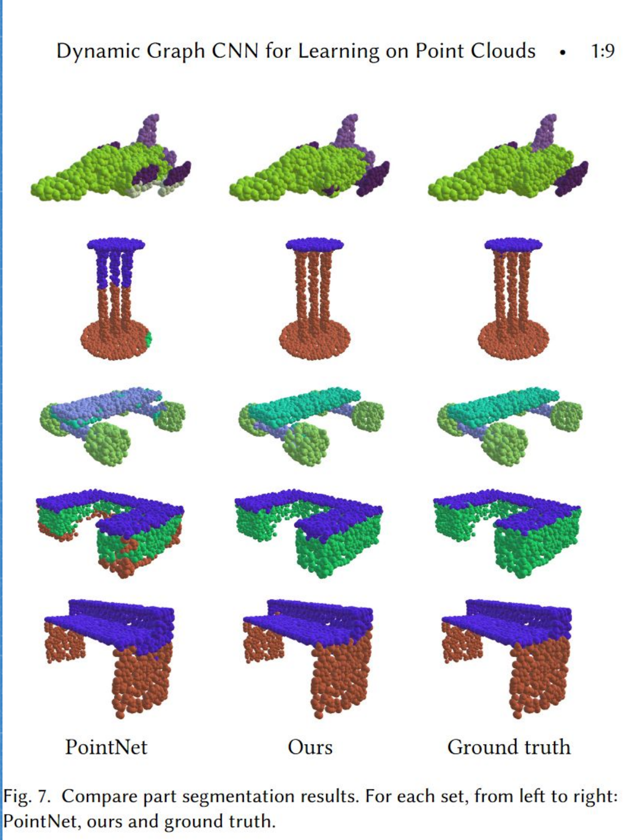
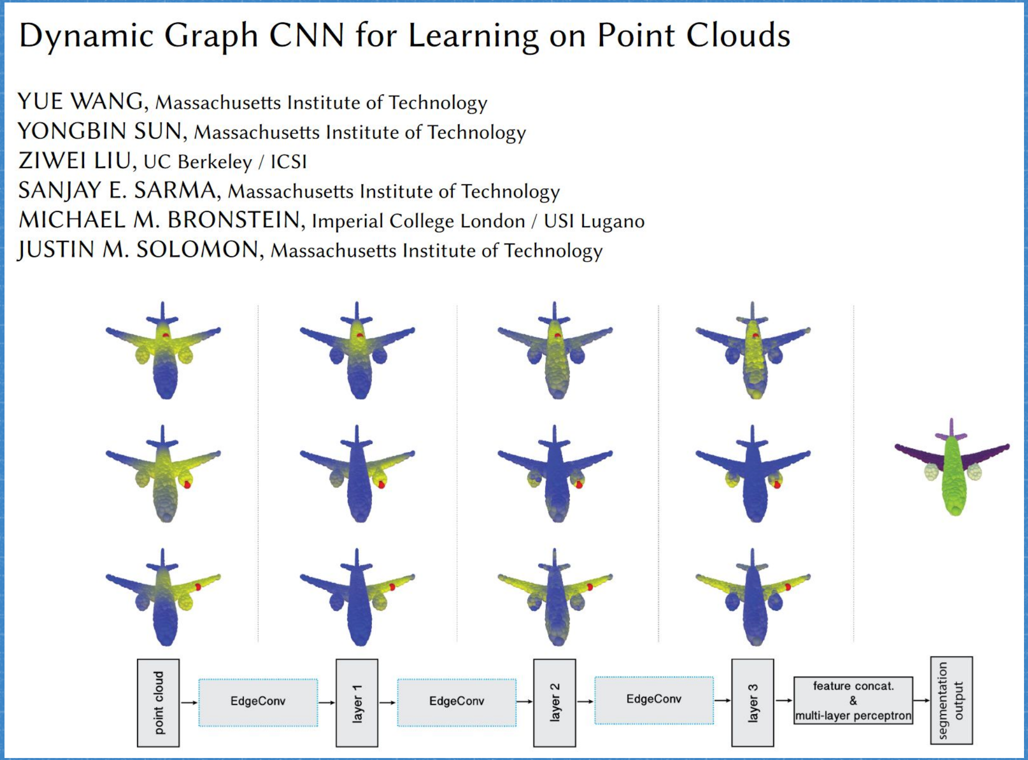

Dynamic Graph CNN for Learning on Point Clouds • 1:9

Fig. 7. Compare part segmentation results. For each set, from left to right: PointNet, ours and ground truth.


Dynamic Graph CNN for Learning on Point Clouds

YUE WANG, Massachusetts Institute of Technology
YONGBIN SUN, Massachusetts Institute of Technology
ZIWEI LIU, UC Berkeley / ICSI
SANJAY E. SARMA, Massachusetts Institute of Technology
MICHAEL M. BRONSTEIN, Imperial College London / USI Lugano
JUSTIN M. SOLOMON, Massachusetts Institute of Technology

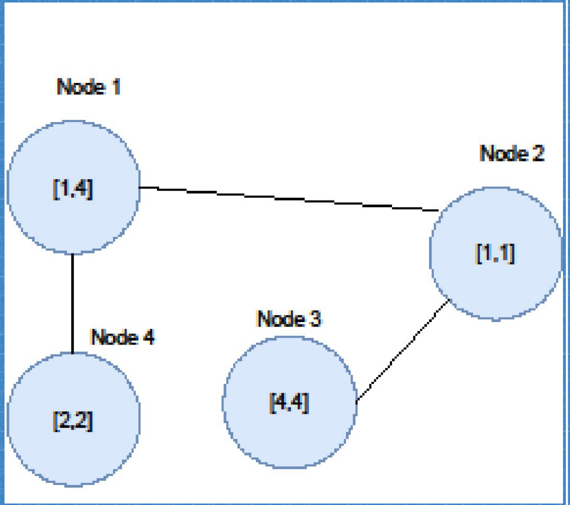
EdgeConv

- The update of values of the j'th node is done via $\hat{x}_j = \sum_{k=1}^{n} f(x_j, x_j - x_k)$

Where f is a learned function (e.g a neural net)

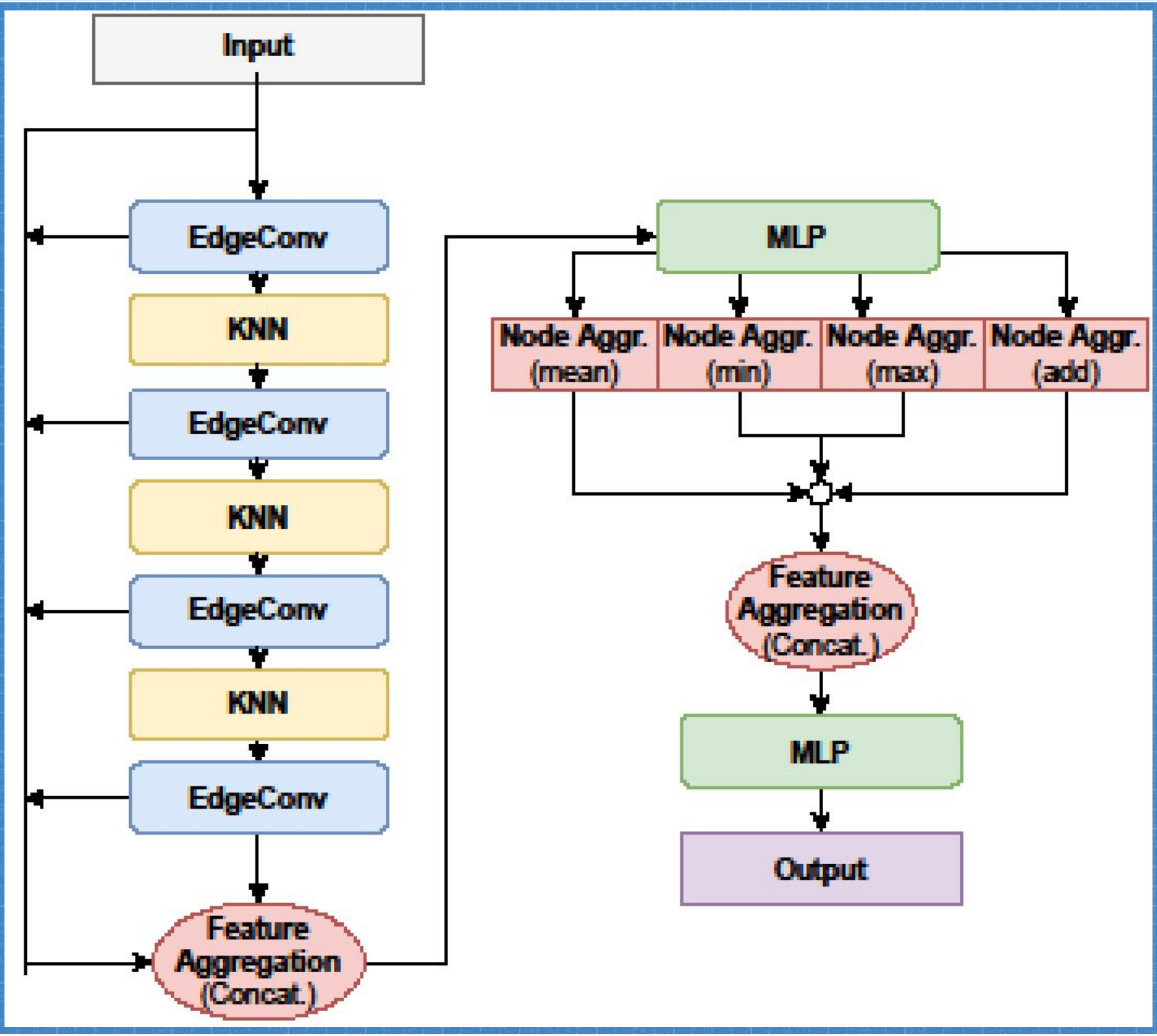
Suppose f = 1 * x + 0. Then the updated values for Node 1 would be:

$$\begin{aligned}
\hat{x}_1 &= f(x_1, x_1 - x_2) + f(x_1, x_1 - x_4) \\
&= f([1,4],[1,4]-[1,1]) + f([1,4],[1,4]-[2,2]) \\
&= f([1,4],[0,3]) + f([1,4],[-1,2]) \\
&= f([1,4,0,3]) + f([1,4,-1,2]) \qquad \text{(by concat.)} \\
&= 1 \cdot [1,4,0,3] + 1 \cdot [1,4,-1,2] \\
&= [2,8,-1,5]
\end{aligned}$$



In a full forward pass, this would be iterated for every node in the graph!

dynedge



Input
EdgeConv
KNN
EdgeConv
KNN
EdgeConv
KNN
EdgeConv
Feature Aggregation (Concat.)
MLP
Node Aggr. (mean)
Node Aggr. (min)
Node Aggr. (max)
Node Aggr. (add)
Feature Aggregation (Concat.)
MLP
Output

(Graph state illustration)

3

Results!

Classification

MC RESULTS (lvl7)

(classification)

The data shown here originates from:

"/data/ana/LE/oscNext/pass2/genie/level7_v02.00/120000”

"/data/ana/LE/oscNext/pass2/genie/level7_v02.00/140000"”

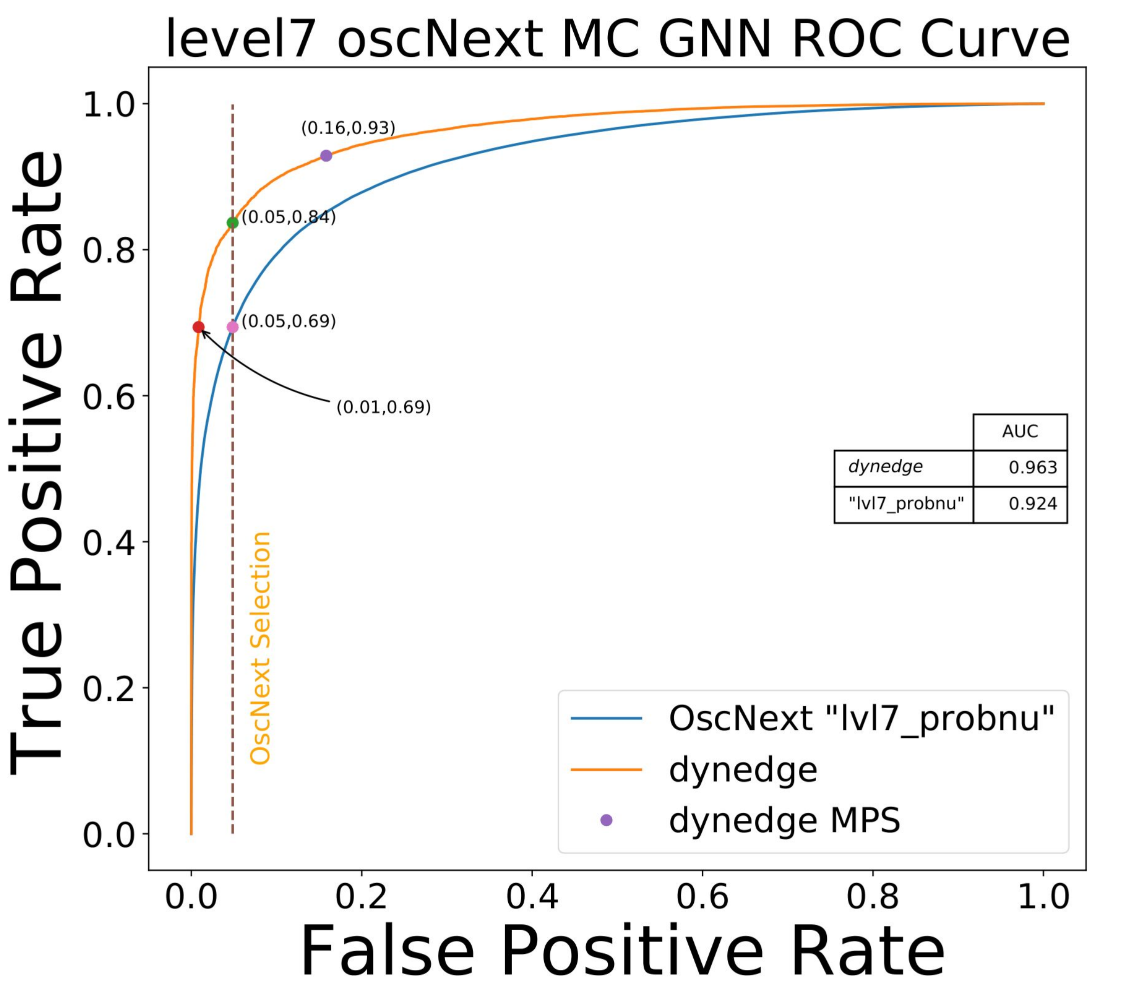
/data/ana/LE/oscNext/pass2/genie/level7_v02.00/160000”

"/data/ana/LE/oscNext/pass2/muongun/level7/130000”

A total of 8.1 million events.

level 7 data selection

- ~642.000 events (out of 8.1 mill)
- 50% muon, 50% neutrino
- The neutrino sample has even distribution of neutrino types
- No noise used!


level7 oscNext MC GNN ROC Curve
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
(0.16,0.93)
(0.05,0.84)
(0.05,0.69)
(0.01,0.69)
OscNext Selection
AUC
dynedge 0.963
"lvl7_probnu" 0.924
OscNext "lvl7_probnu"
dynedge
dynedge MPS

MC RESULTS (lvl4)

(classification)

The data shown here originates from:

"/data/ana/LE/oscNext/pass2/genie/level4_v02.00/120000”

"/data/ana/LE/oscNext/pass2/genie/level4_v02.00/140000"”

/data/ana/LE/oscNext/pass2/genie/level4_v02.00/160000”

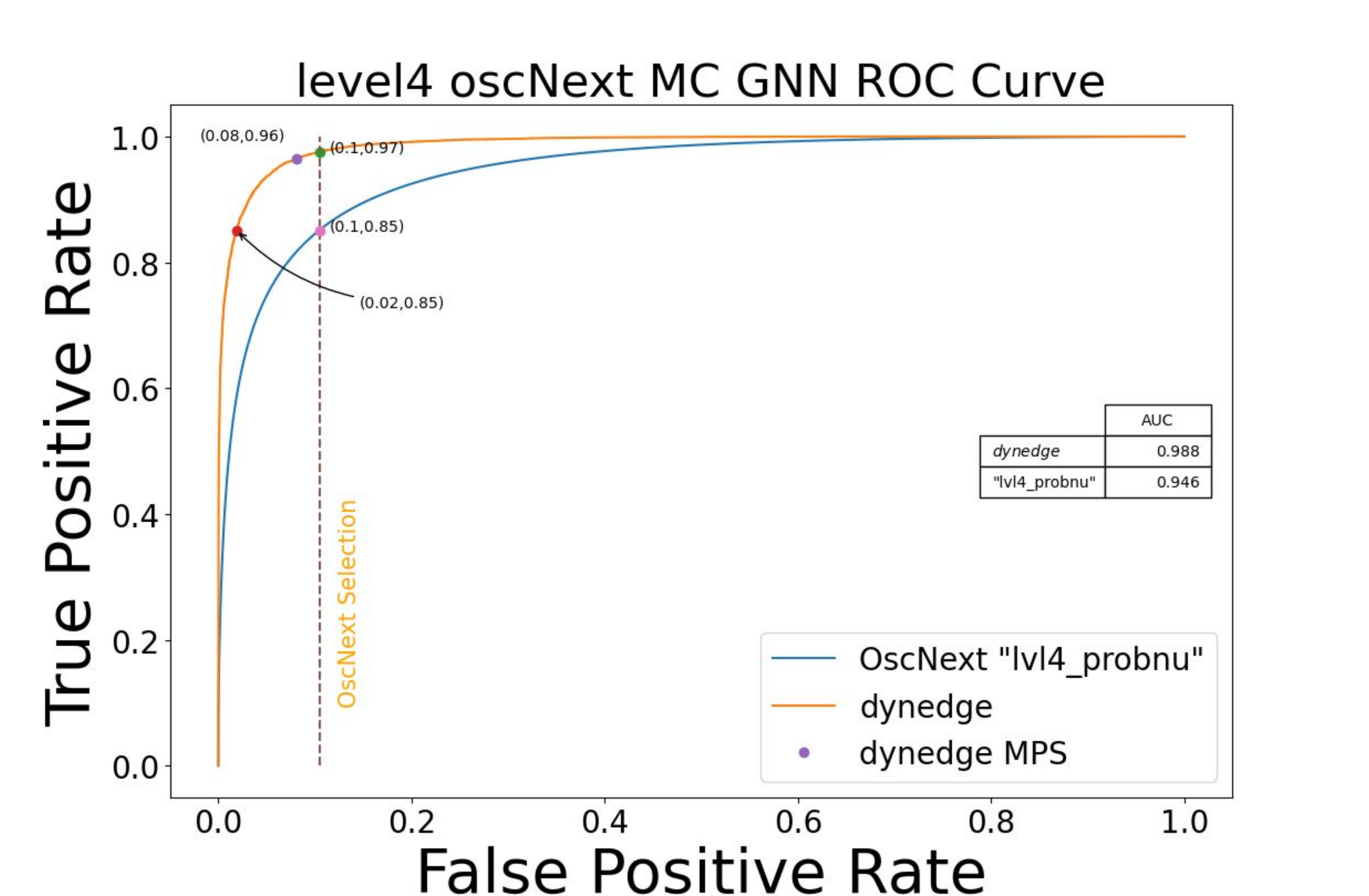
"/data/ana/LE/oscNext/pass2/muongun/level4/130000”

“/data/ana/LE/oscNext/pass2/noise/level4_v02.00”

A total of 15.9 million events.

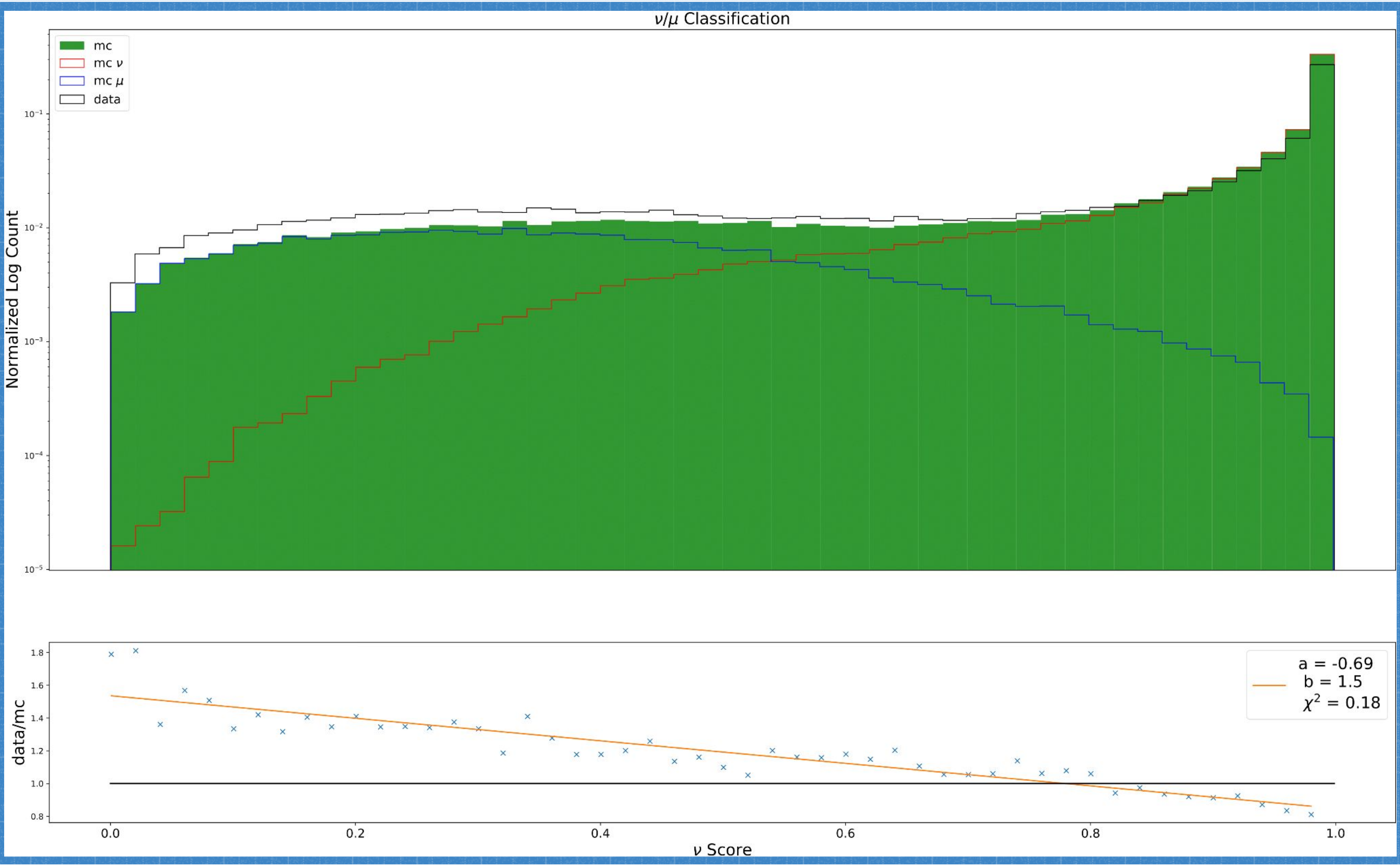
level 4 data selection

- ~600.000 events (out of 15.9 mill)
- Even amount of muons, neutrinos and noise
- The neutrino sample has even distribution of neutrino types


level4 oscNext MC GNN ROC Curve
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
(0.08,0.96)
(0.1,0.97)
(0.1,0.85)
(0.02,0.85)
OscNext Selection
AUC
dynedge 0.988
"lvl4_probnu" 0.946
OscNext "lvl4_probnu"
dynedge
dynedge MPS

REAL DATA RESULTS

Real data is IC86.11 sample



ν/μ Classification
mc
mc ν
mc μ
data
Normalized Log Count
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
data/mc
a = -0.69
b = 1.5
χ^2 = 0.18
1.8
1.6
1.4
1.2
1.0
0.8
0.0
0.2
0.4
0.6
0.8
1.0
ν Score

Regression

(energy_log10, zenith)

Regression Metrics

- Energy : error $= \frac{Pred.-True}{True}$ (otherwise numerical difference)

- Error Width : w $= \frac{IQR(error)}{1.349}$ (interquantile range. Factor applied such that w is approximately 1 sigma if the errors are Gaussian)

- Rel. Imp. = 1 - w_dyn/w_retro

MC RESULTS (lvl7)

(regression)

The data shown here originates from:

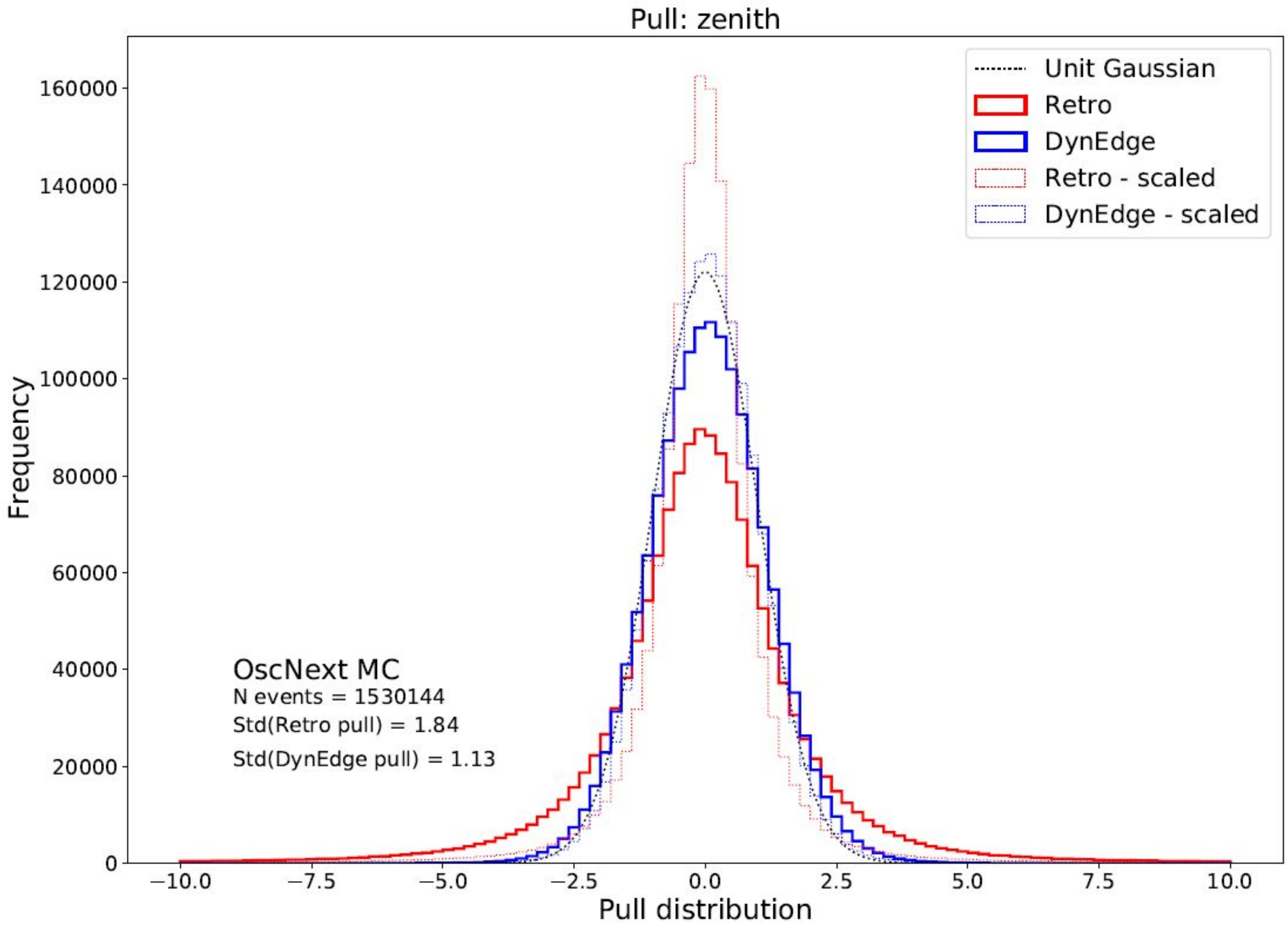
"/data/ana/LE/oscNext/pass2/genie/level7_v02.00/120000”

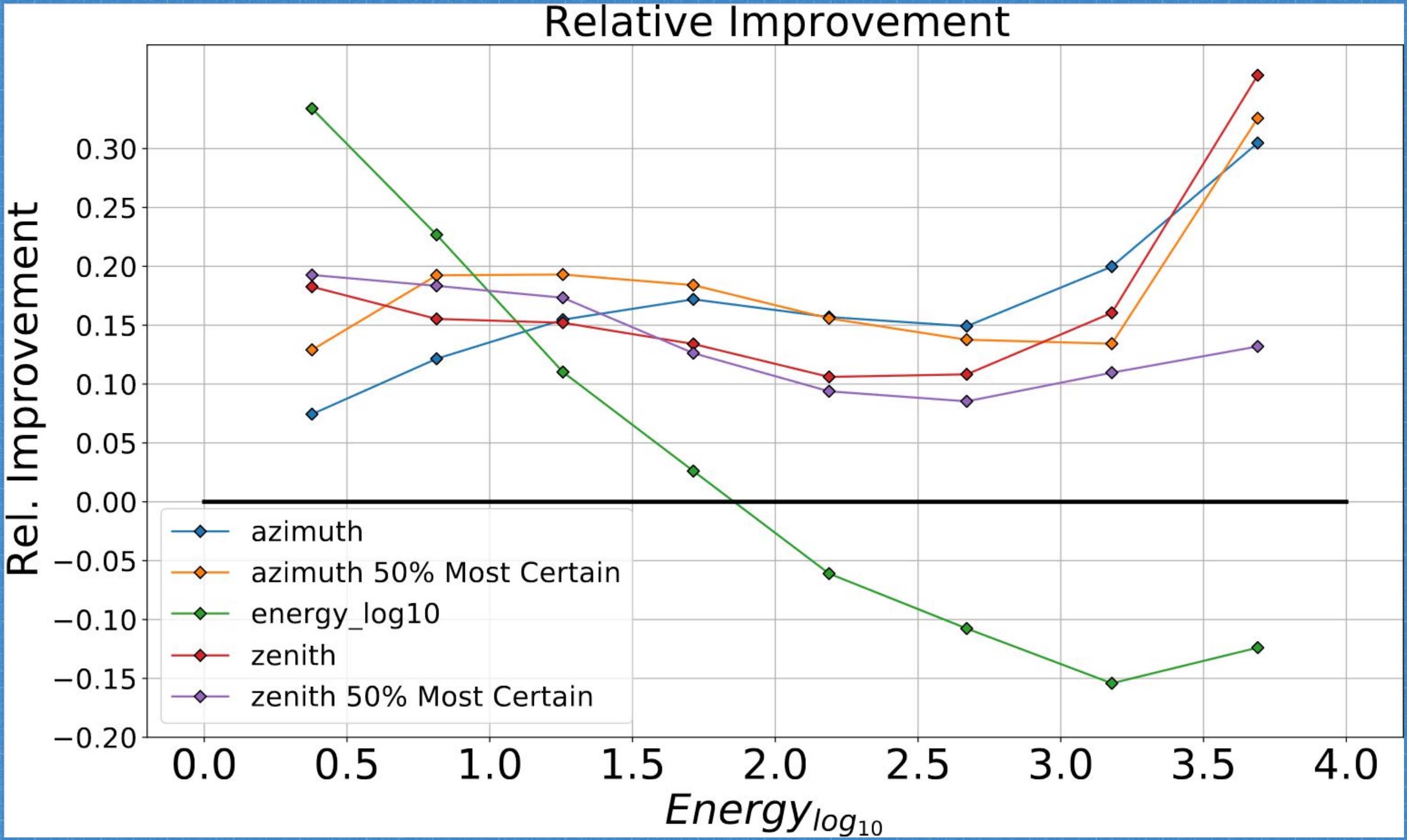
"/data/ana/LE/oscNext/pass2/genie/level7_v02.00/140000"”

/data/ana/LE/oscNext/pass2/genie/level7_v02.00/160000”

NOTE:

This version of RetroReco used a minimizer that made it faster but produces inferior uncertainties as a result

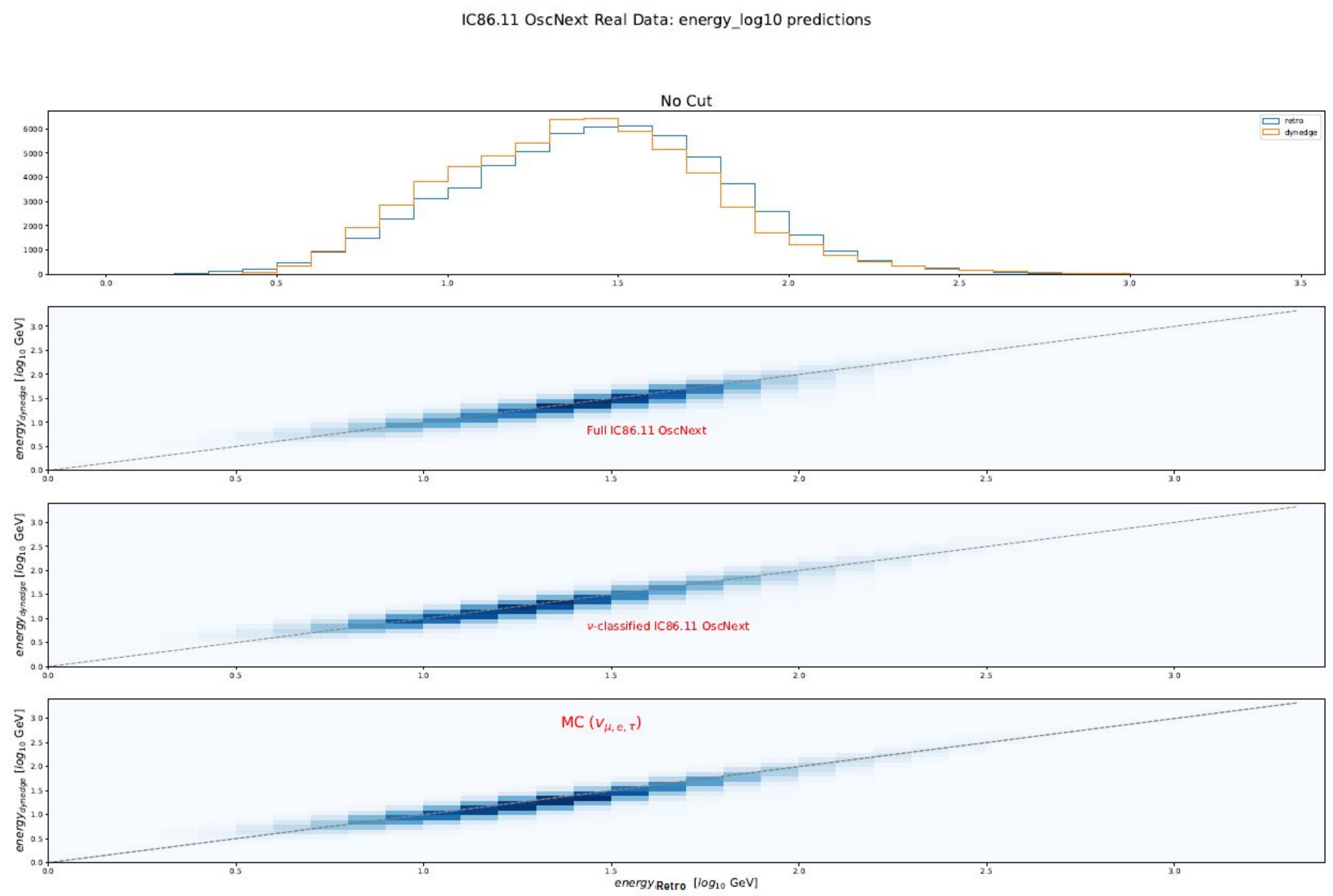



Relative Improvement
Rel. Improvement
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.20
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
$Energy_{log_{10}}$
azimuth
azimuth 50% Most Certain
energy_log10
zenith
zenith 50% Most Certain

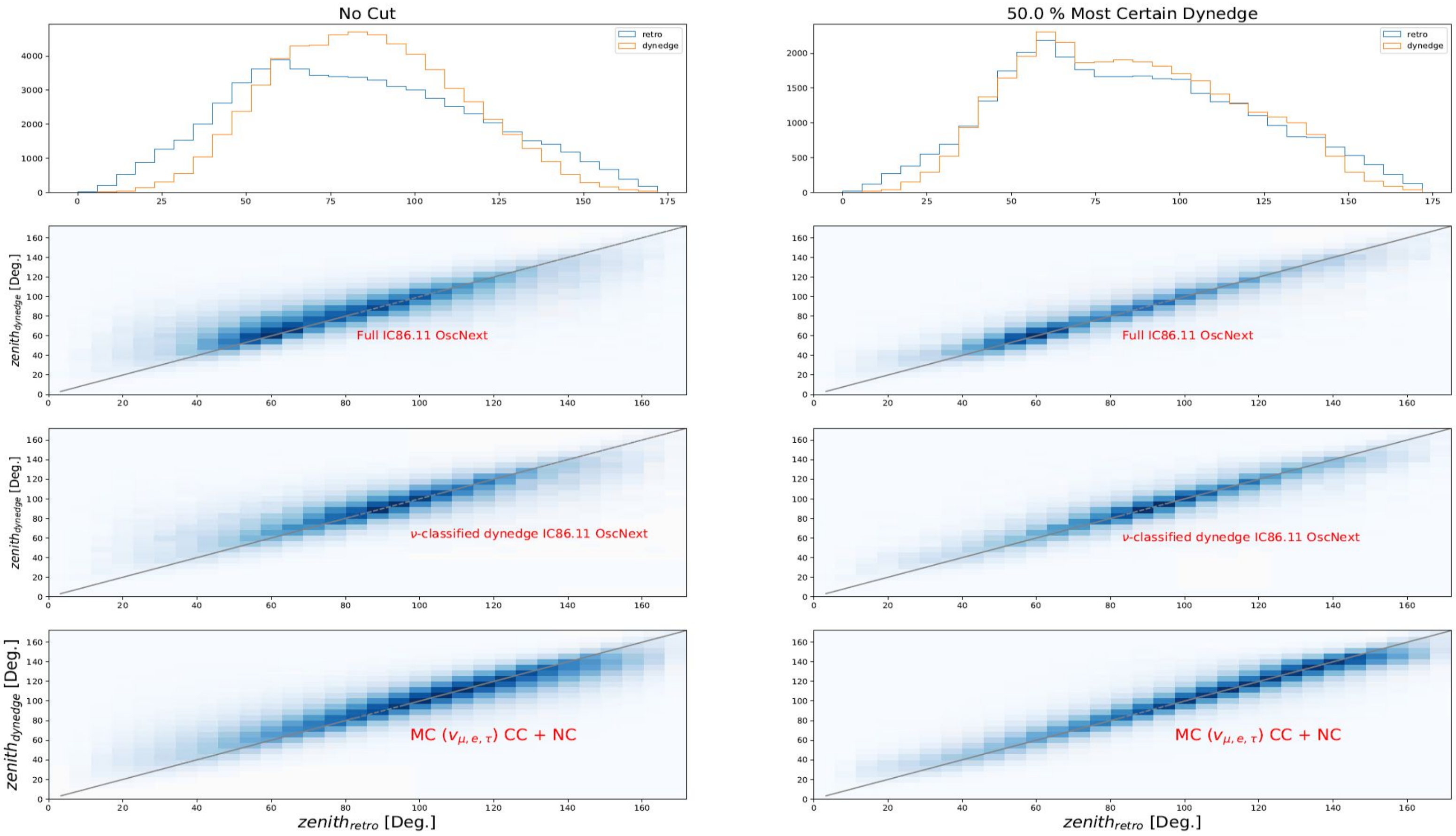
(oscillation plots)

REAL DATA RESULTS

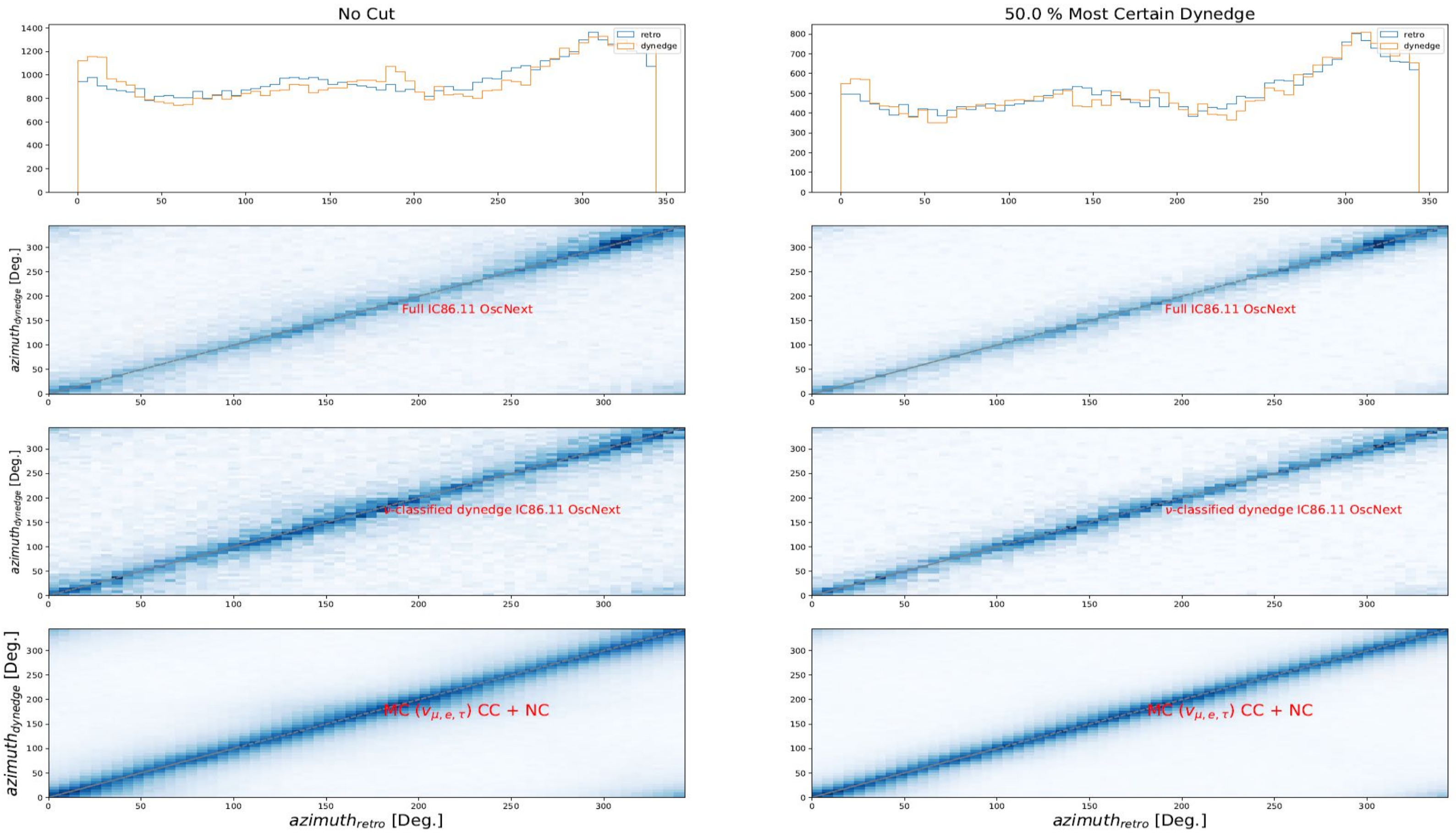
Real data is IC86.11 sample


IC86.11 OscNext Real Data: energy_log10 predictions
No Cut
retro
dynedge
Full IC86.11 OscNext
ν-classified IC86.11 OscNext
MC ($\nu_{\mu,e,\tau}$)
$energy_{dynedge}$ [log_{10} GeV]
$energy_{Retro}$ [log_{10} GeV]

IC86.11 OscNext Real Data: zenith predictions



IC86.11 OscNext Real Data: azimuth predictions



Conclusion

Conclusion

- Reconstruction Speed: ~15.000 events / s
- 11.7%, 22.4% and 16.3% avg. rel. improvement for azimuth, energy_log10 and zenith at 0 - 1.5 log10 GeV.
- 12% and 15% increased signal at level4 and level7, or factor 5 reduction in FPR.

Outlook

- Integrate dynedge into IceCube Framework (Already started!)
- Do PISA-analysis (with help from TUM) (Already started!)
- DoubleBang Classification with PhD Student Leander (DESY) (Starting Monday)
- Re-iterate the angular uncertainties
- So much more..

Thanks!

ANY QUESTIONS?