Incorporating Advances in Machine Learning for Reconstruction in T2K and Super-Kamiokande Experiments

Félix Cormier

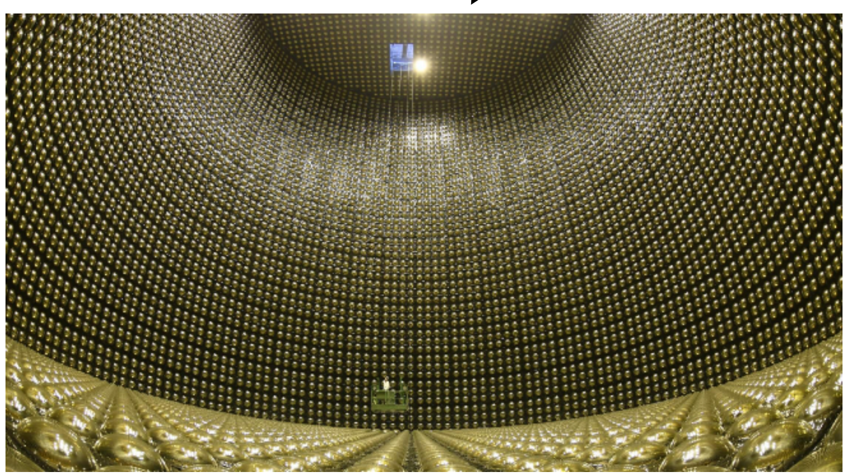
28/05/2024

CAP Congress - PPD Session

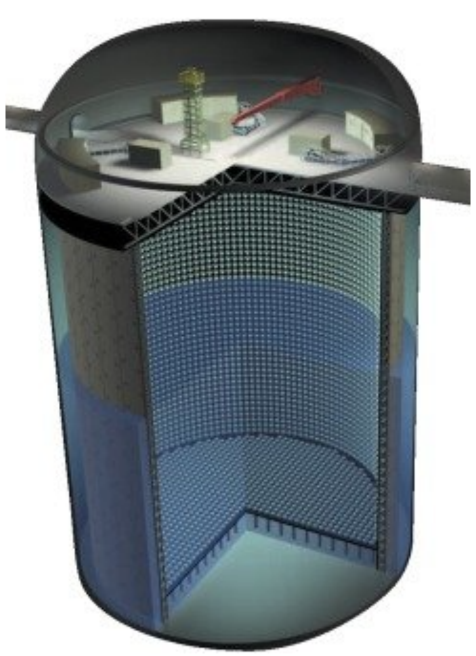
Super-Kamiokande

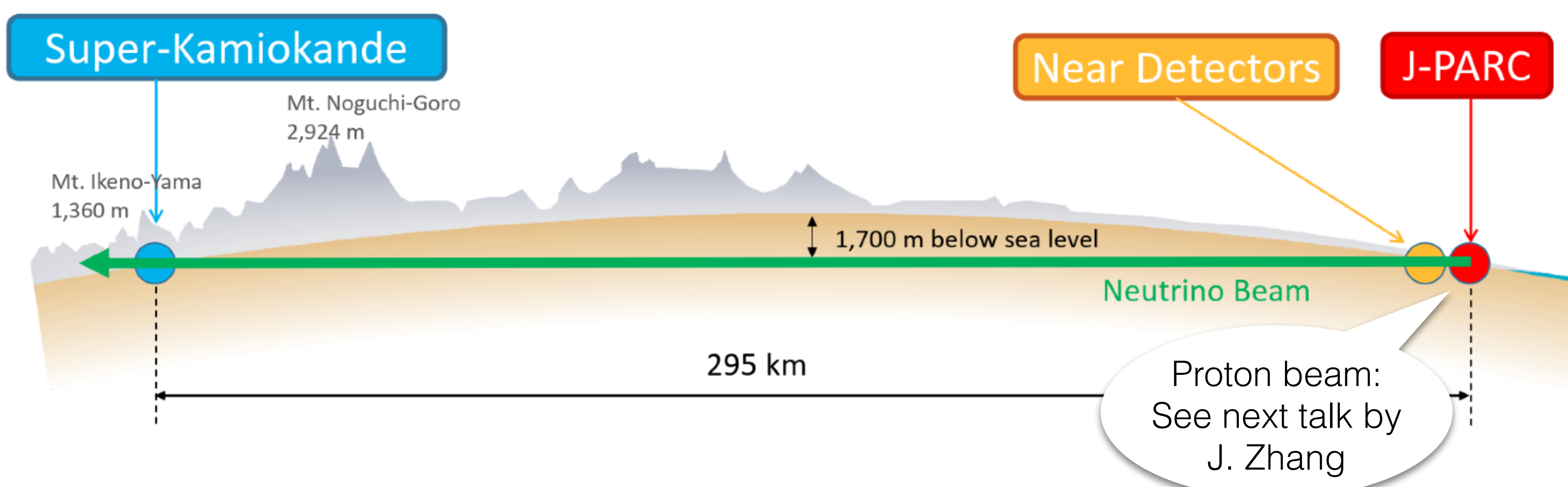
Solar Neutrinos

Atmospheric Neutrinos



- SuperK is a Water Cherenkov detector
  - Large 50 kton water tank lined with PMTs optimal for detecting neutrino interactions
  - Acts as far detector for T2K long-baseline experiment
  - Can also detect solar, atmospheric or supernova neutrinos
  - Cherenkov light from neutrino interaction products can lead to particle identification, reconstruction





The WatChMaL Collaboration

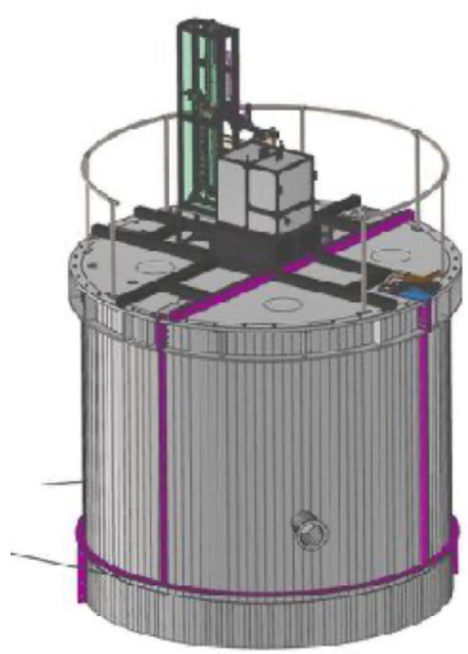
- The **Wat**er **Ch**erenkov **Ma**chine **L**earning Collaboration is a cross-experiment group
- Common data generation, pre-processing and training frameworks are shared by members
- Currently have members in experiments
  - T2K/SuperK
  - Water Cherenkov Test Experiment (WCTE)
  - Intermediate Water Cherenkov Detector (IWCD)
  - Hyper-Kamiokande

Talk Monday by S. Yousefnejad

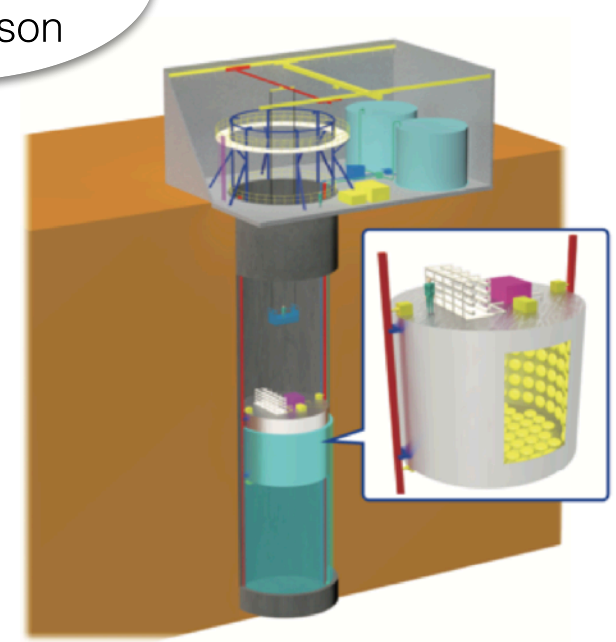
WatChMaL.org

See talk Friday by B. Jamieson

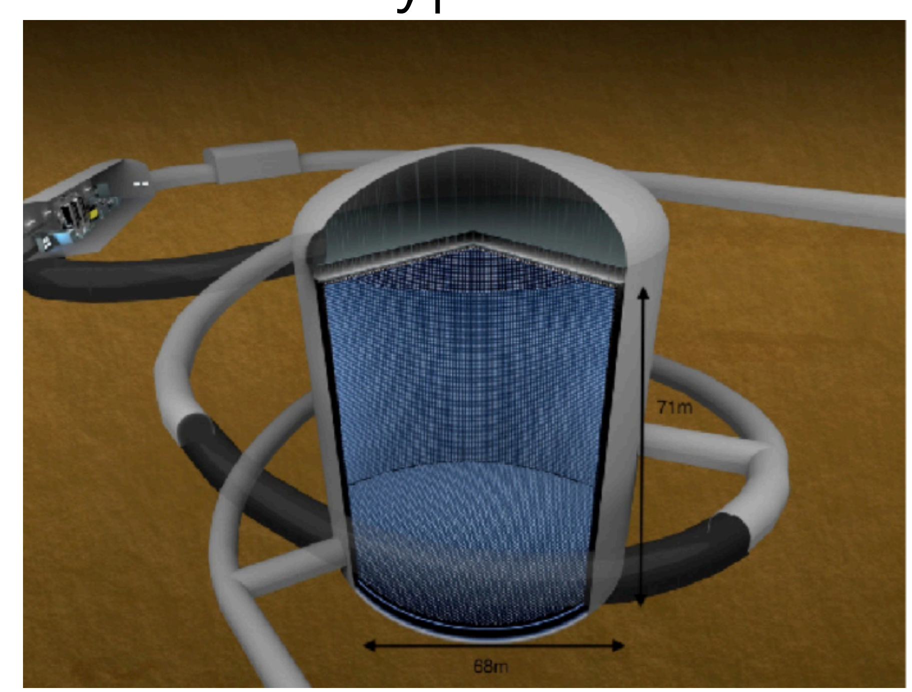
WCTE



IWCD

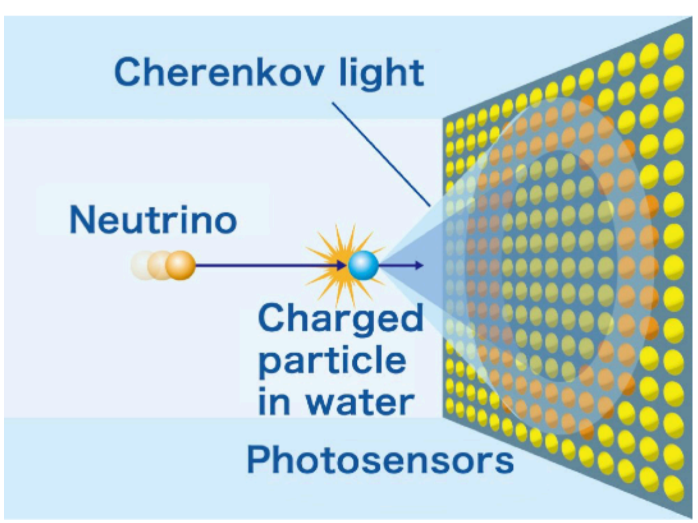


HyperK

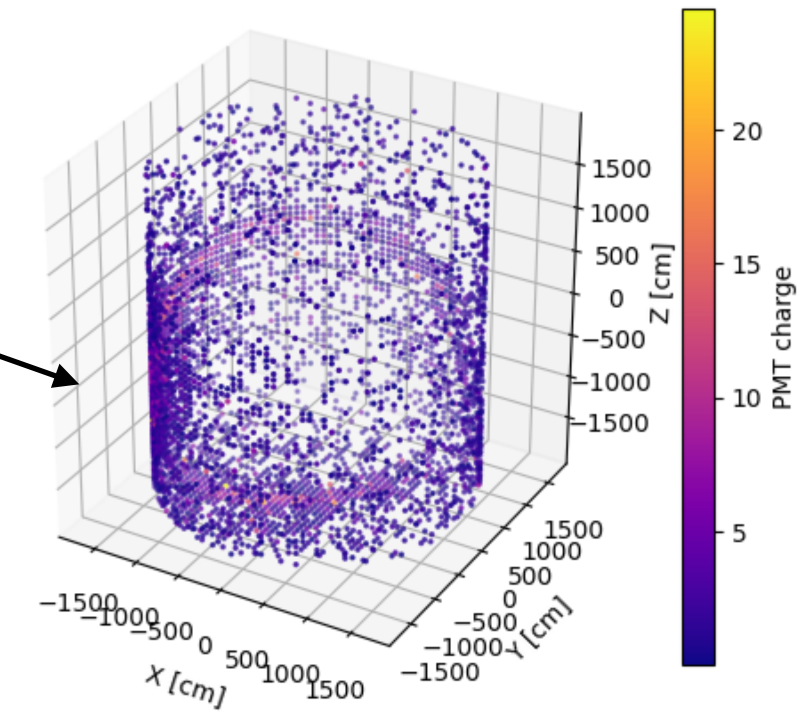


Water Cherenkov data

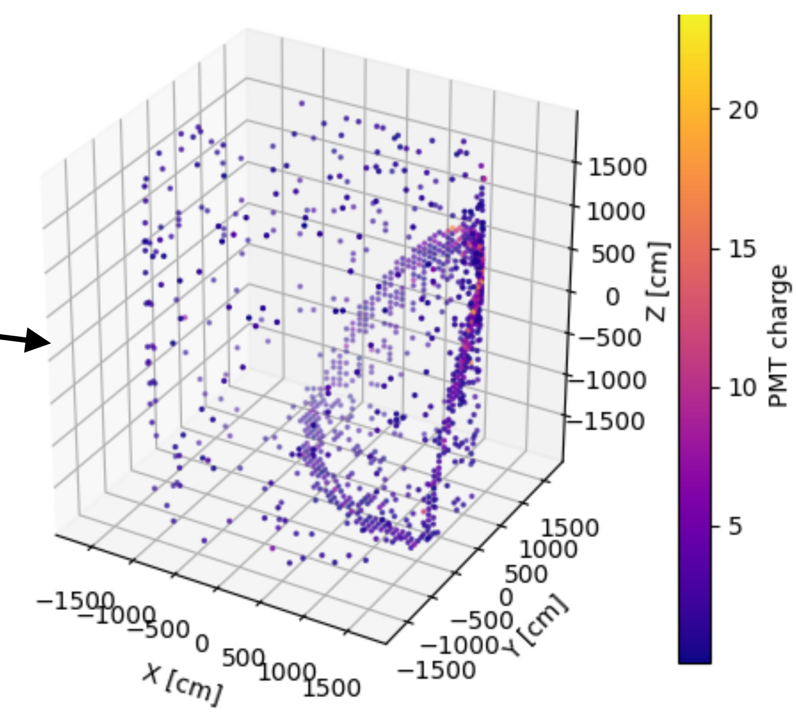
- As products of neutrino interactions travel in water, they produce Cherenkov rings
    - These are imaged by the PMTs
- Products are often electrons (muons) from ν_e (ν_μ) interactions
    - Electrons will produce **larger rings** due to multiple interactions and showering
    - Muons (and π^+) will often have **thinner rings**
- Data is in the form of integrated charge, time of individual PMTs



Electron

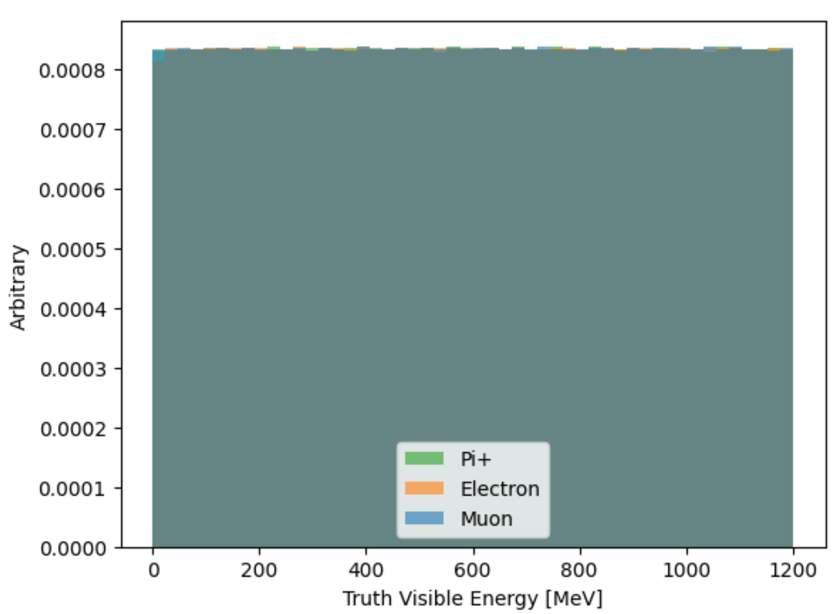


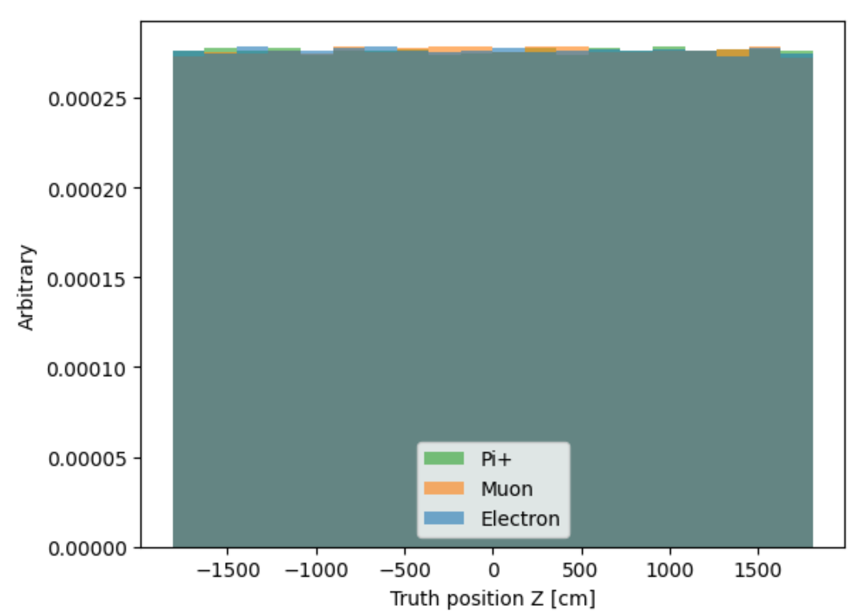
Muon

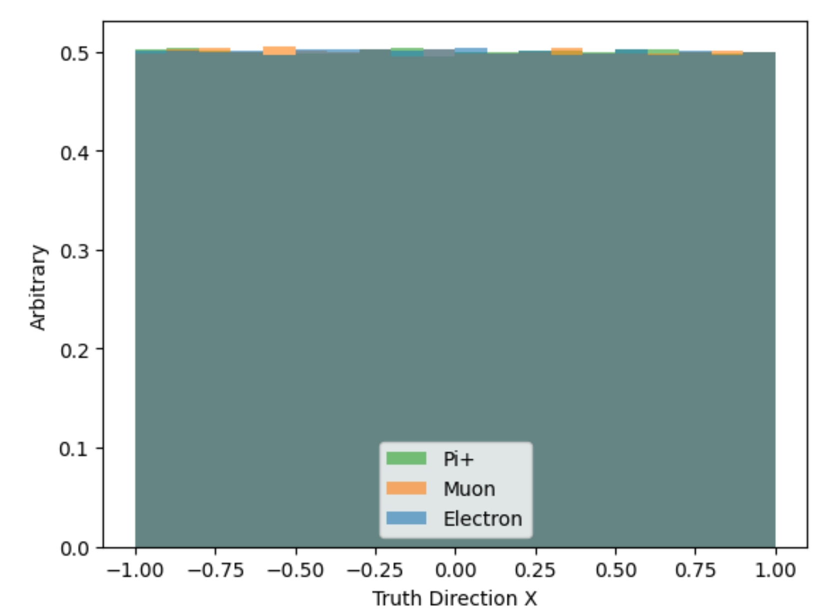


Data generation

- Data generation is important for training deep learning networks in order to avoid biases
- We generate data using official SuperK simulation software, SKDETSIM
  - Generate samples of electrons, muons and π^+
    - **Energy is uniformly sampled** from 0 to 1 GeV above Cherenkov threshold
    - Uniform in vertex position & direction





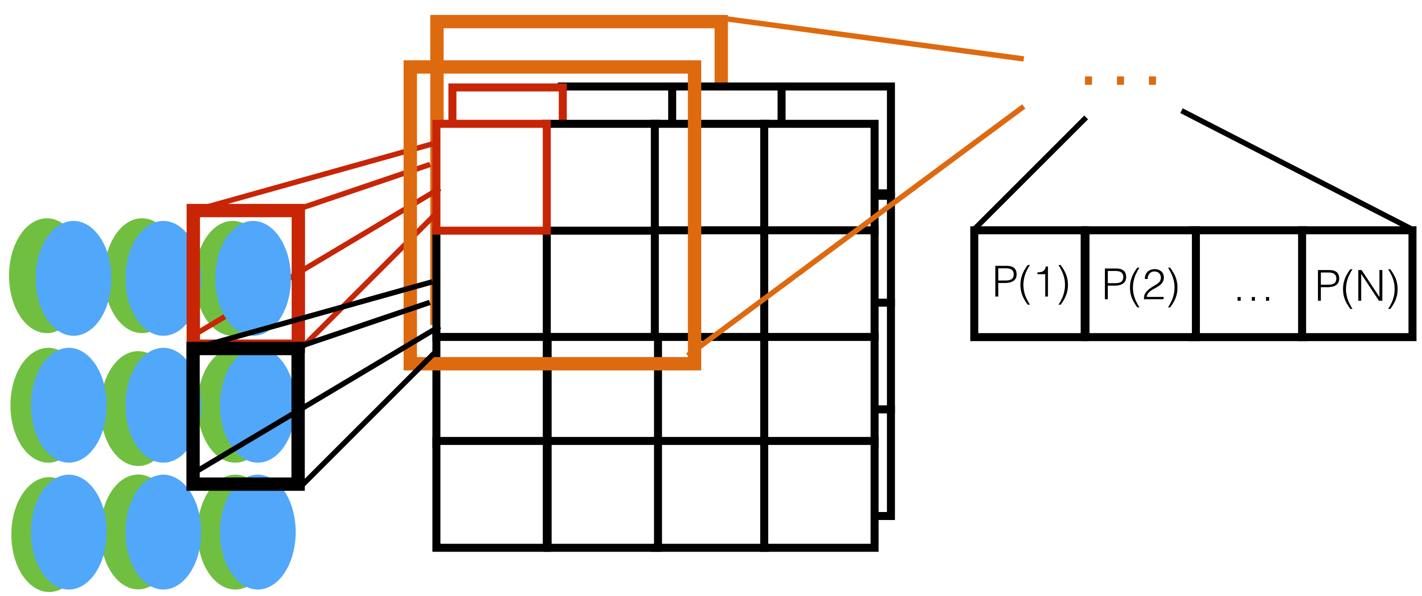


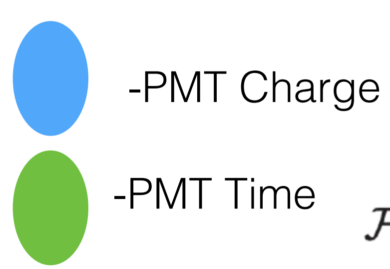
Machine Learning for event reconstruction

- Currently high energy event reconstruction in SuperK is done using the **fiTQun** algorithm
  - This algorithm depends on likelihood minimization, using PMT charges and time to construct the likelihood function
  - Makes assumptions about what the data will look like
  - Calculation complexities make this algorithm difficult to extend to e.g. different particle hypotheses
  - Is very slow to compute - hard to scale
- By contrast, machine learning algorithms can learn this complexity by training over simulated data
  - Can e.g. speed up calculations by over order of magnitude

ResNet Architecture

- **ResNet**
  - Convolutional layers over 2D image-like input PMT data & subsequent layers build features
  - Residual connections between layers help with vanishing gradients





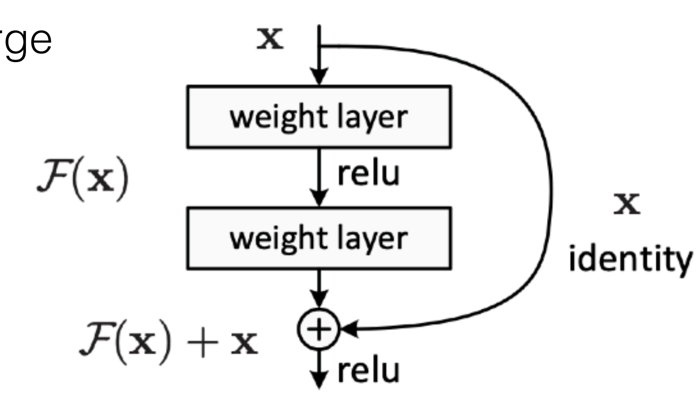
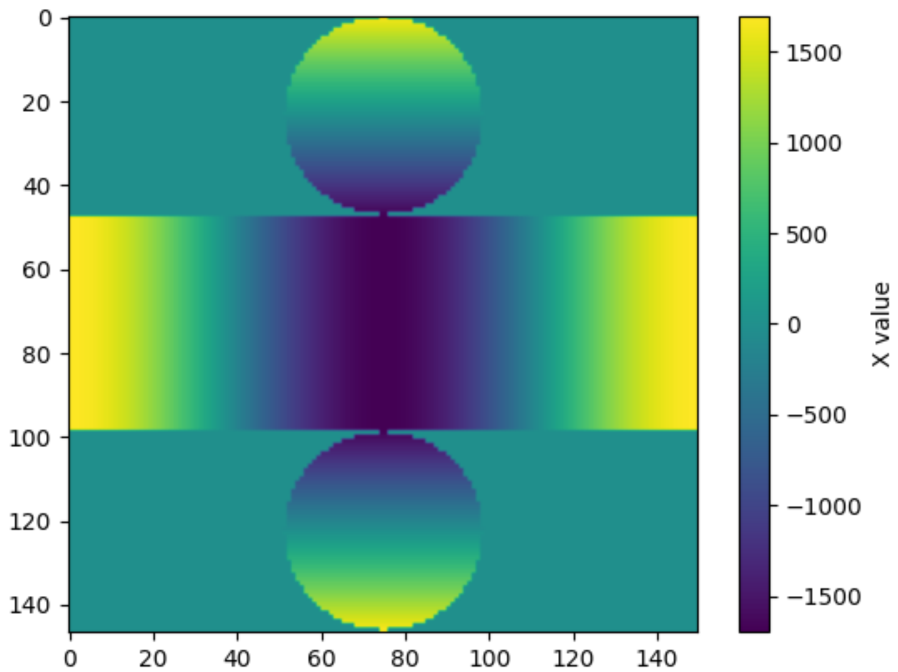


Figure 2. Residual learning: a building block.

Processing Data for ResNet

- Tests over many Water Cherenkov detectors have shown **ResNet to have better performance** at both classification and regression
- One of the challenges of ResNet is to project 3D cylindrical data from SuperK into a 2D image
- Each PMT is linked to its 3D position in space using a dictionary
  - Each 3D position is then unrolled into a 2D image, with each PMT becoming a pixel

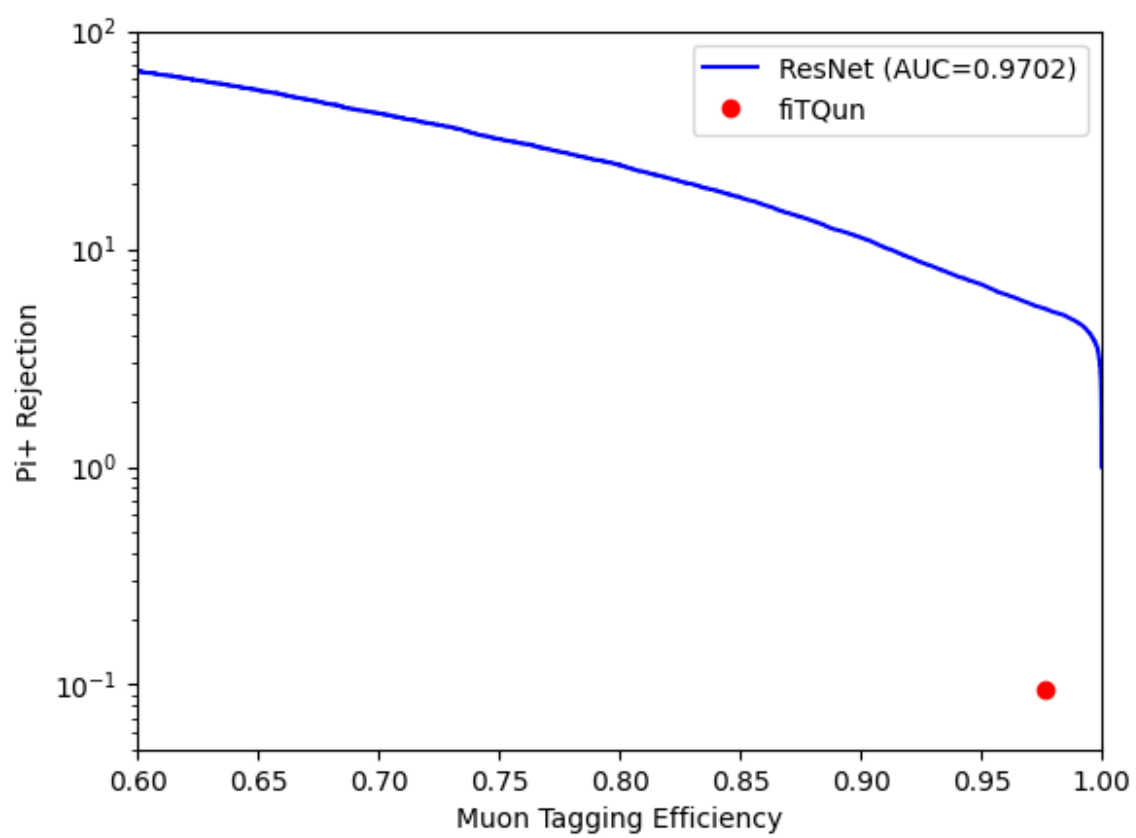


Classification Networks

- Electrons and muons can be easily distinguished in most cases
  - Classical algorithms **already have very high accuracy** in distinguishing these events
- Muons and π^+ are more difficult to distinguish
  - Cherenkov rings from the initial particles virtually indistinguishable
  - Algorithm must use subtle pion hadronic interactions to distinguish between the events
  - Increased classification performance could improve T2K muon disappearance results
    - Potential to impact mass hierarchy measurements

Classification Results

- Run a 3-class classification network using ResNet to classify between e, μ, π^+
- Show results for μ (signal) vs. π^+ (background)
- Can scan across all the class outputs to see how background rejection and signal efficiency vary
  - 100x **better background rejection** at same signal efficiency for μ vs. π^+
  - Currently studies ongoing to understand w**hat the network is learning** for μ vs. π^+ that fiTQun did not learn - hadronic interactions of the charged pion?

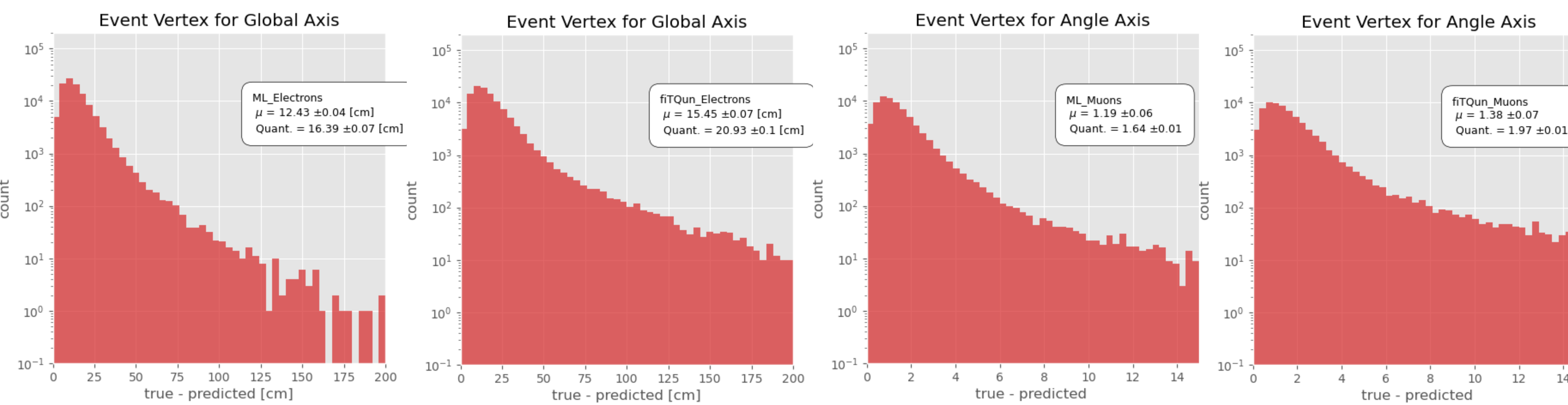


Regression Networks

- An important part of event reconstruction is finding the particle's initial **position**, **direction** and **momentum**
  - Position and direction reconstruction performance can increase efficiency of cuts based on detector location
  - Momentum reconstruction can help reconstruction of initial neutrino energy
- Better resolution in these kinematic variables can have large effects on efficiency during analysis
- Furthermore, better resolution and smaller bias of the reconstruction algorithms can reduce systematic uncertainties

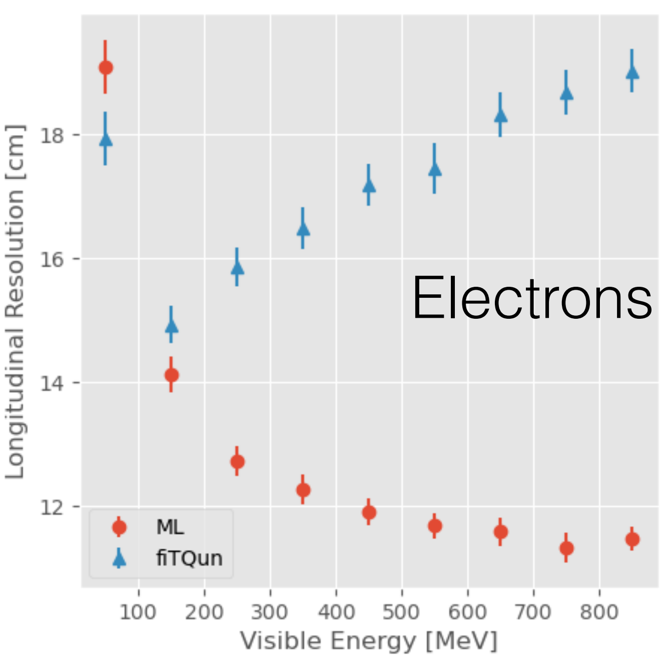
Regression Analysis

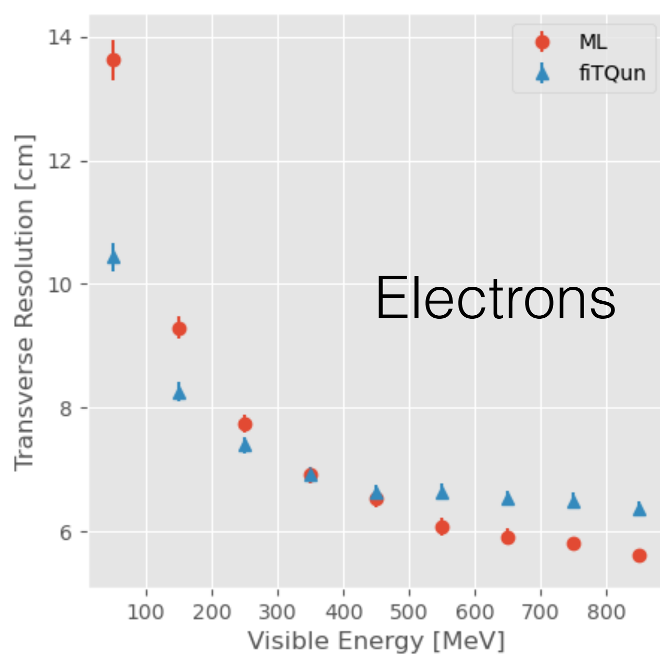
- We construct & train 6 individual networks to reconstruct
  - Electron: position, direction, momentum
  - Muon: position, direction, momentum
- To gauge performance of each network, calculate the residuals of the reconstructed against true value
  - For position: look at 3D distance between true and reconstructed
  - For direction: look at angle between true and reconstructed unit vectors
  - For momentum: look at residual percentage: $(p_{true} - p_{reco})/p_{true}$
- Calculate
  - **Resolution: 68.3% quantile of the residual**
  - Bias: median of the residual

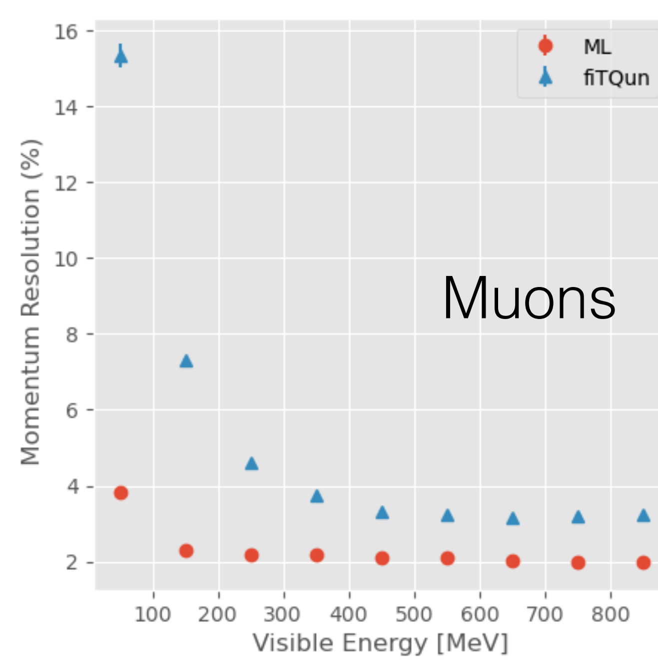


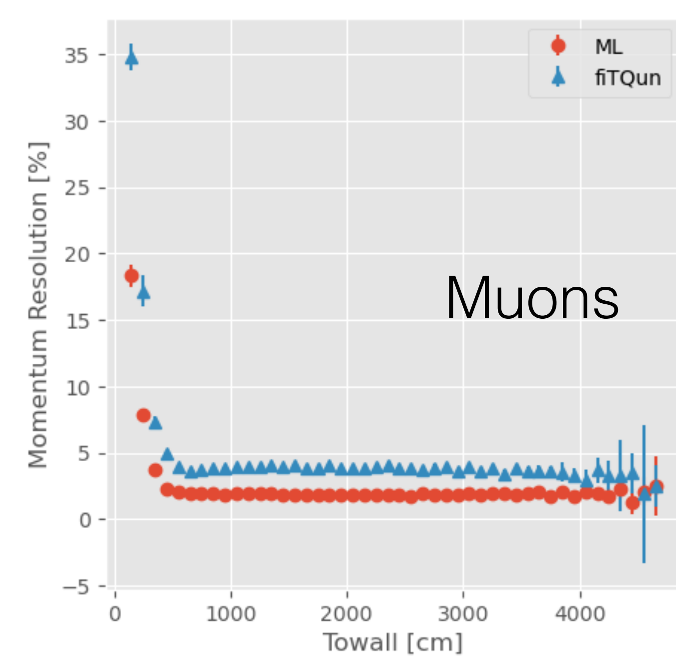
Regression Results

- Can look at position along longitudinal, tranverse direction with respect to true particle direction
    - See that getting better transverse resolution very hard
- Can also analyze regression results as function of underlying variables
    - Visible energy - energy above Cherenkov threshold
    - Towall - distance from initial vertex to detector wall, along true particle direction
- Overall: **10-100% improvement** in resolution depending on particle/ variable

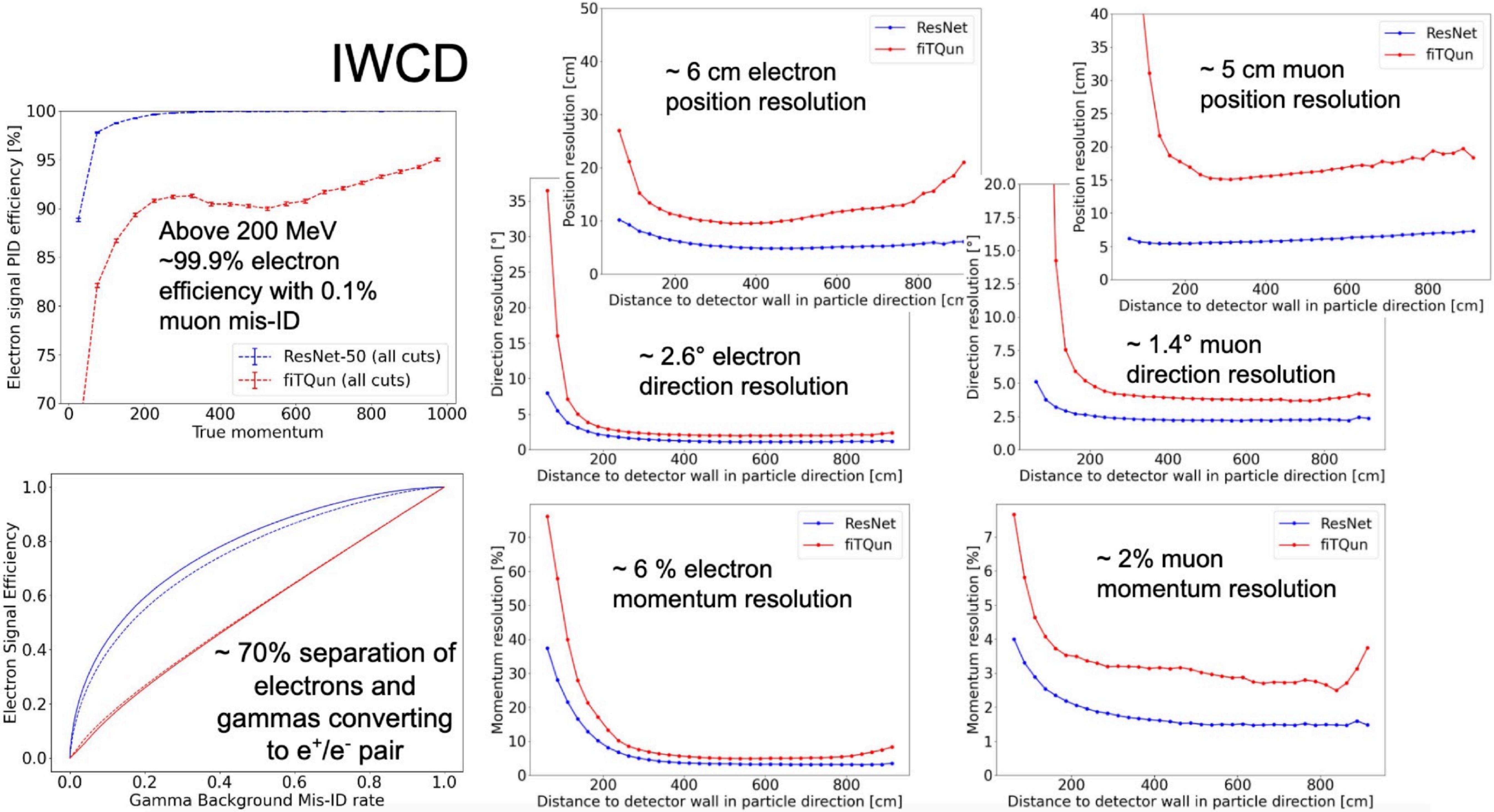








Other WatChMaL Results



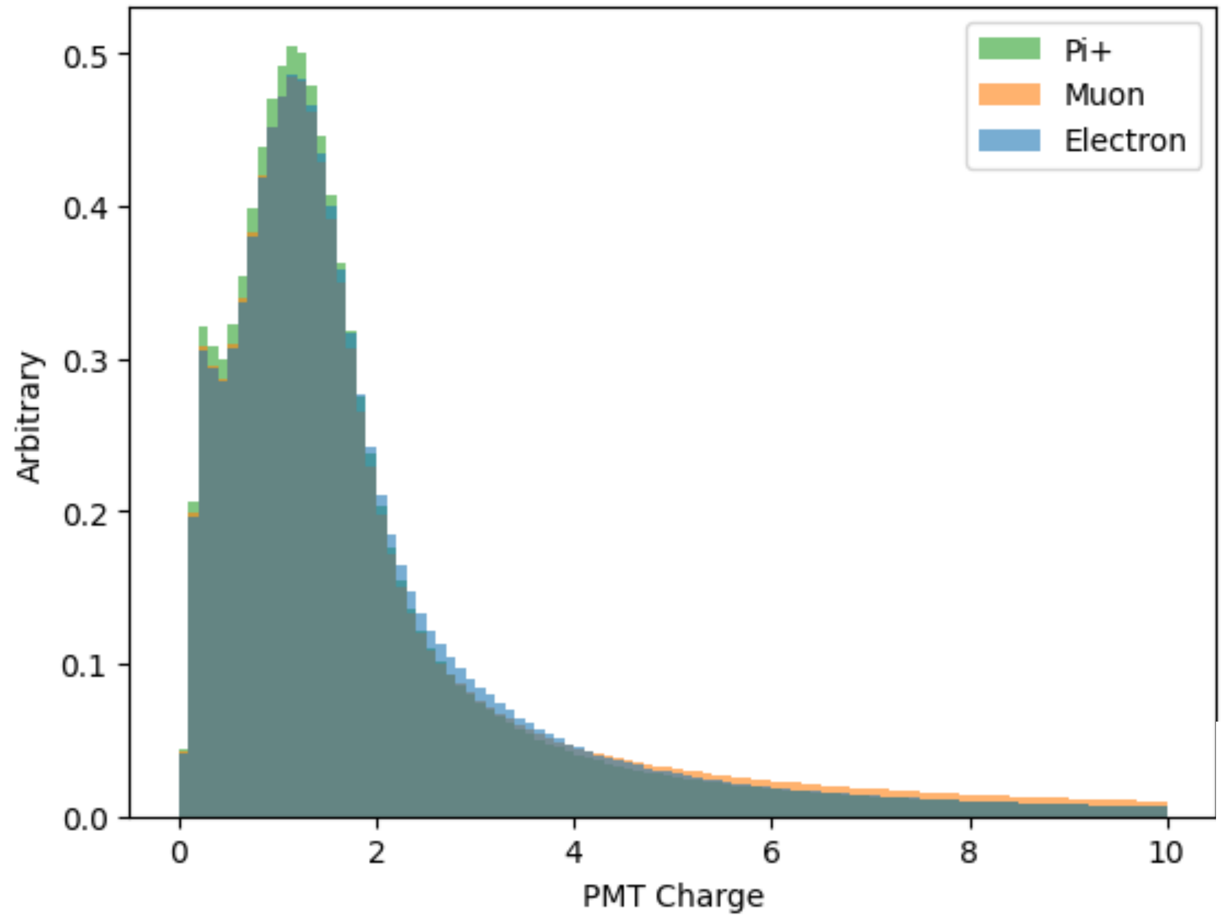
Conclusions & Next Steps

- As part of the WatChMaL Collaboration, applied Deep Learning techniques for Super-Kamiokande event reconstruction
- Use the patterns of PMT charge and time to learn underlying phase space for both classification and regression
  - Classification between electrons, muons and pions shows great improvement over classical reconstruction methods
  - Regression for vertex position, direction and momentum show **10-100% improvement** in resolution, and large reductions in bias for most cases
- WatChMaL analyses on other Water Cherenkov detectors ongoing, also showing good performance
- Next steps will include studying the role of adversarial training on systematic uncertainty reduction

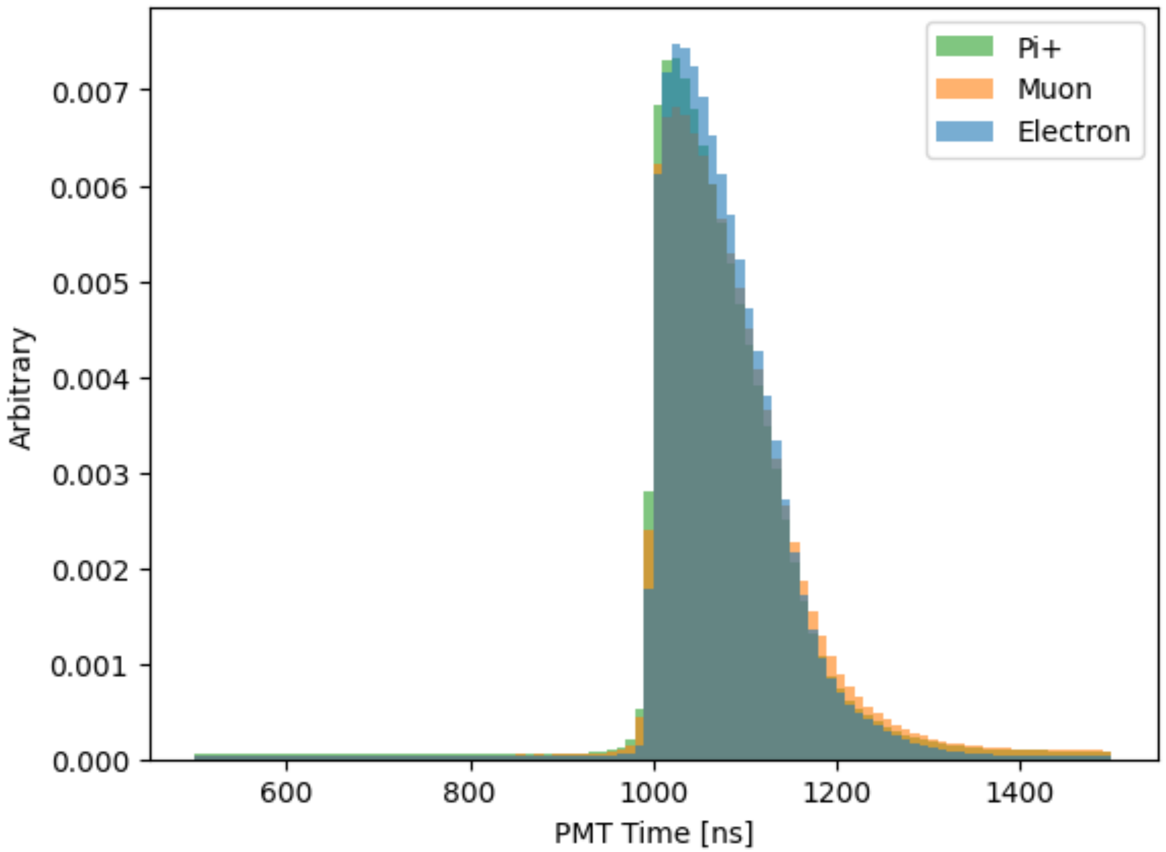
Thank you!

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC).
Nous remercions le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) de son soutien.

Backup

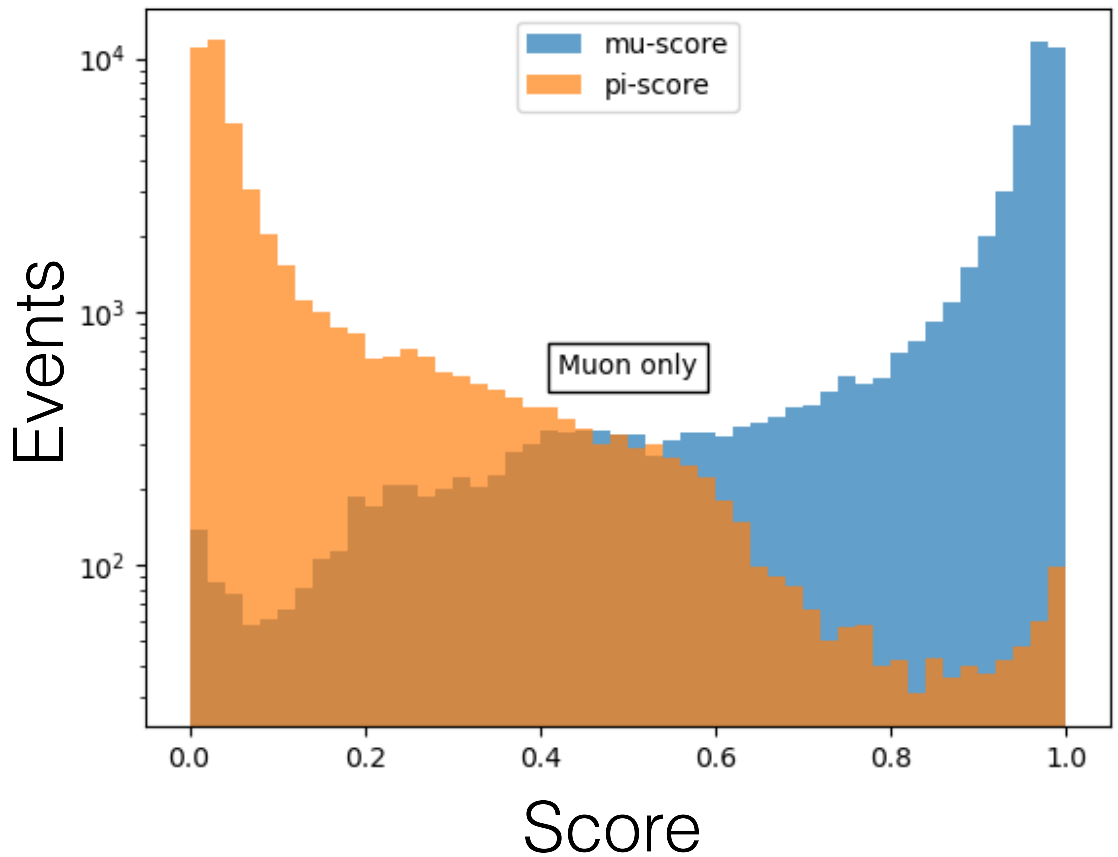


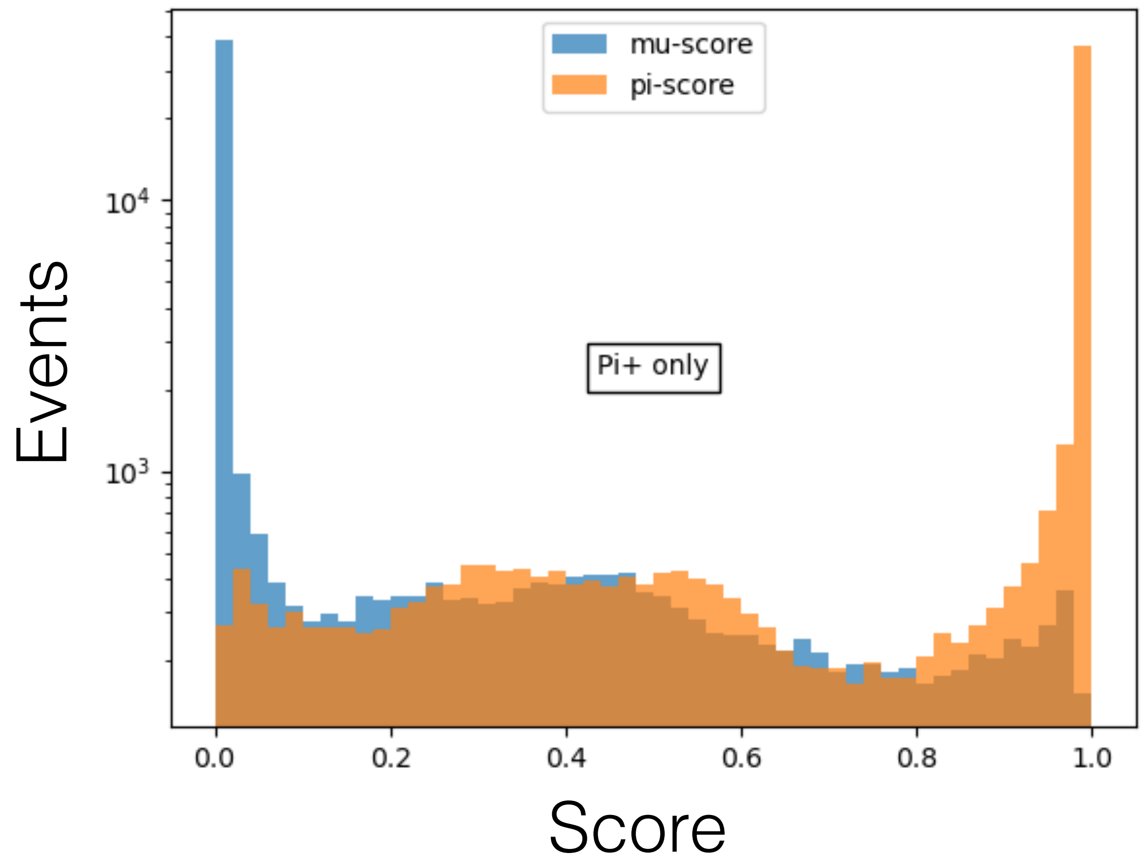
- Today will show current results for classification and regression
  - Classification: 3-class, e/mu/pi+
  - Regression: position, direction and momentum for e/mu
- Data sets are simulated such that
  - Isotropic position, direction in SuperK
  - Uniform visible energy distribution 0-1 GeV
- Training uses the ResNet34 architecture
  - Inputs are all PMT charges, times flattened onto 2D ‘image’

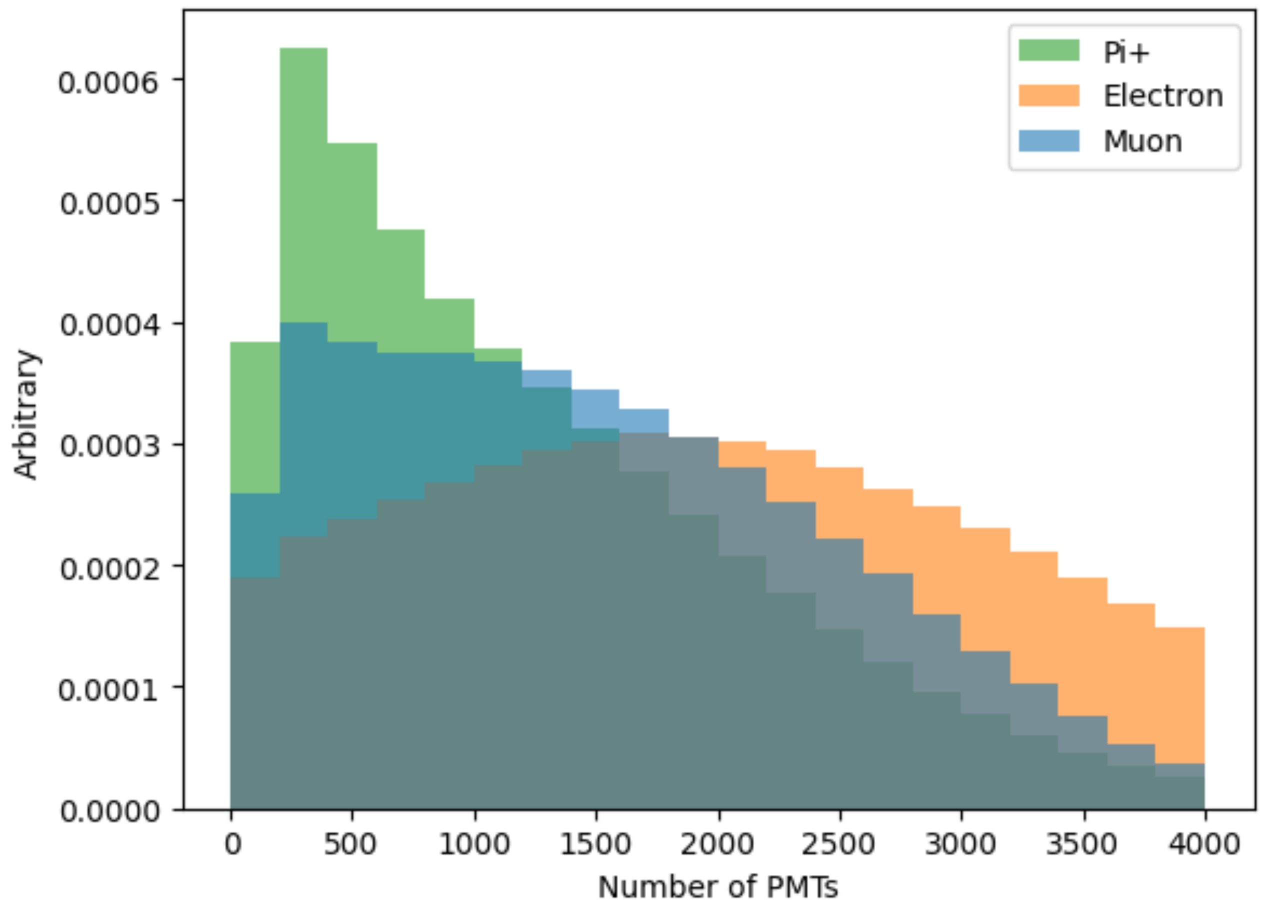


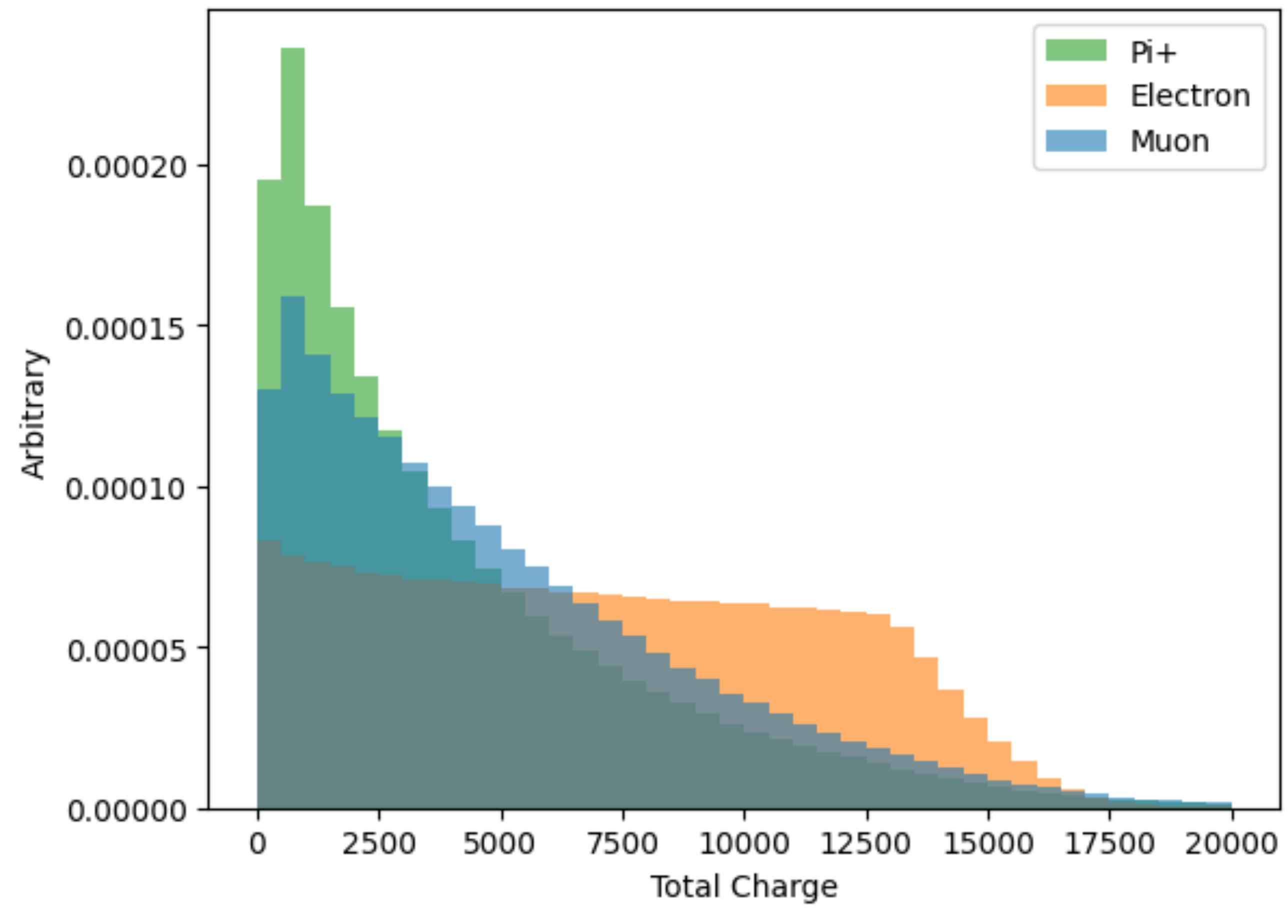
Classification training

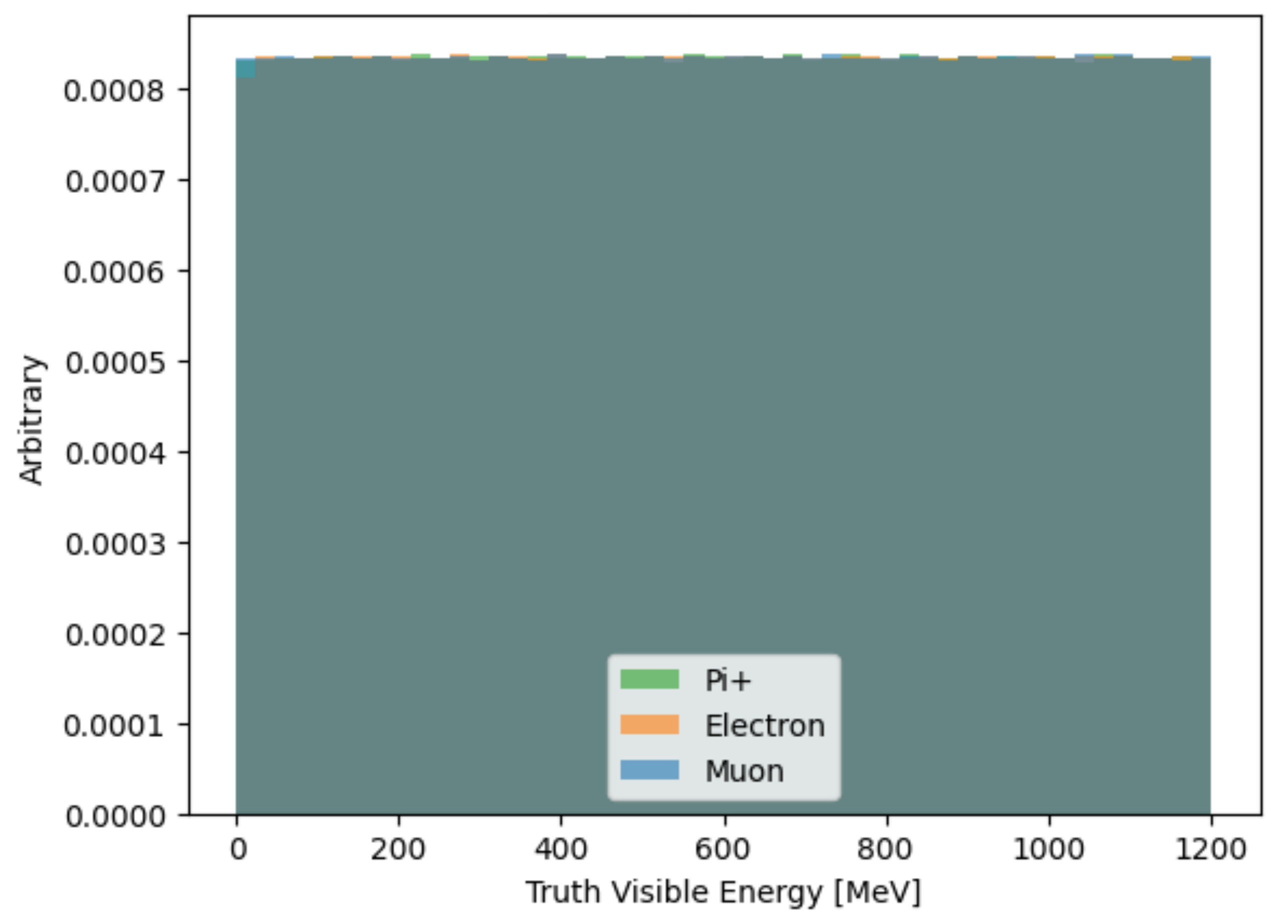
- Classification returns a score between 0 and 1 for each class (Softmax)
    - Electron/muon/pi+
    - 3 scores per event which add up to 1
- Here can see muon/pi+ separation

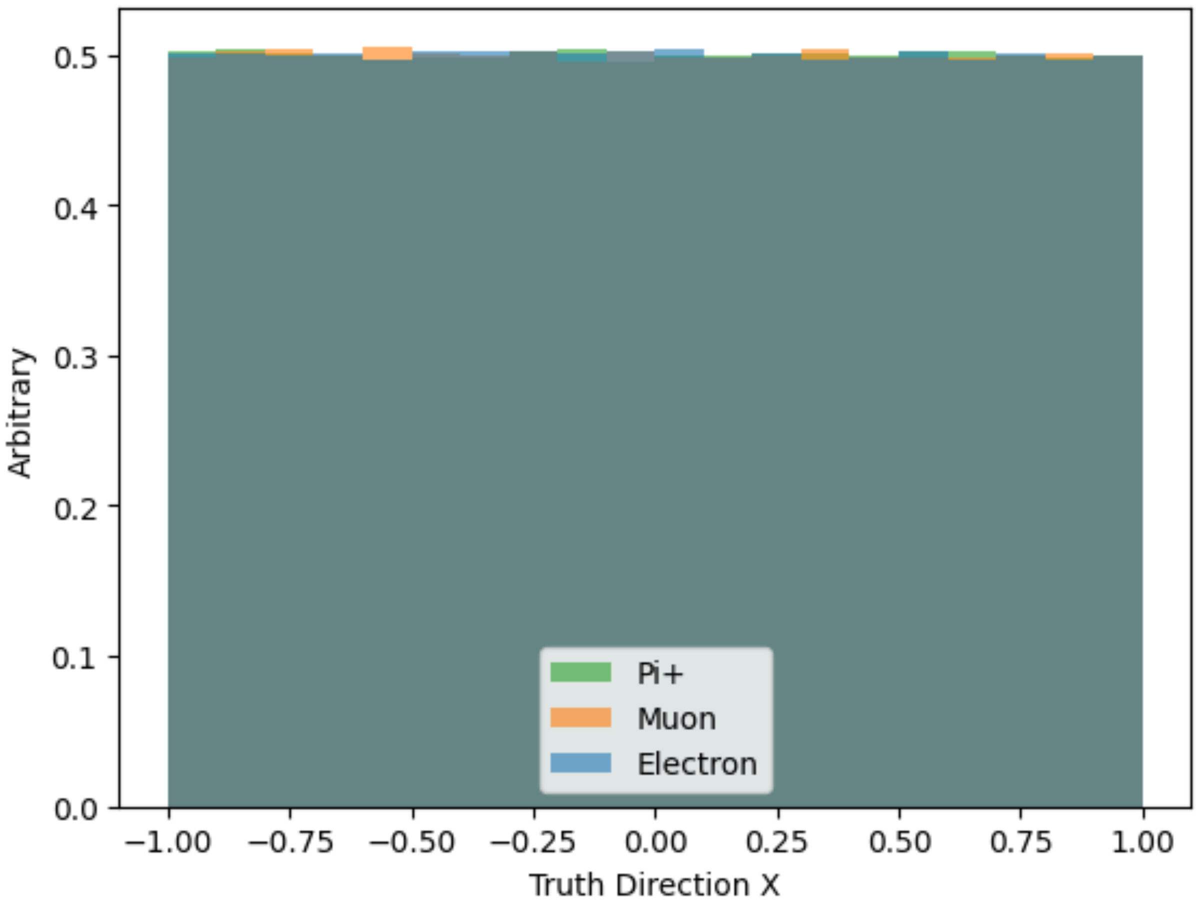


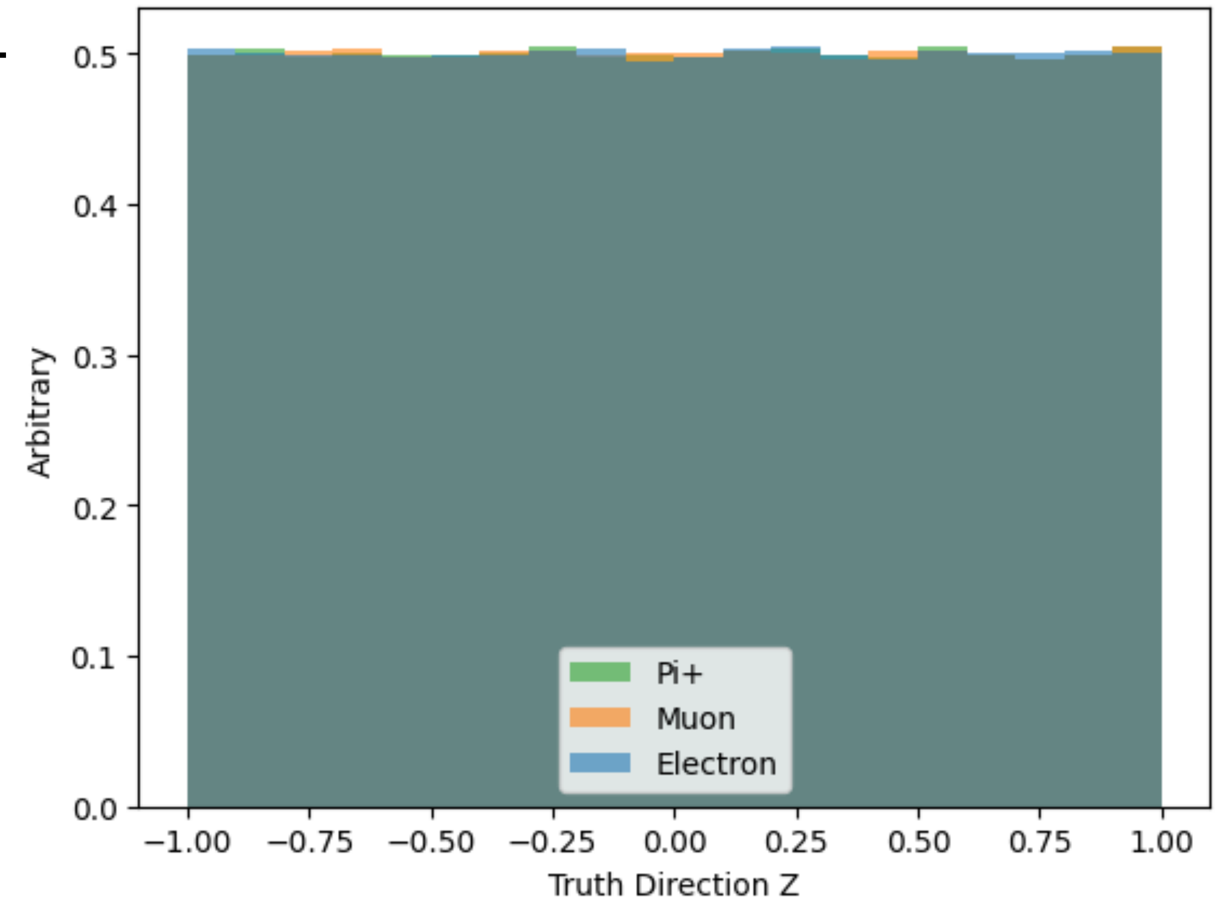


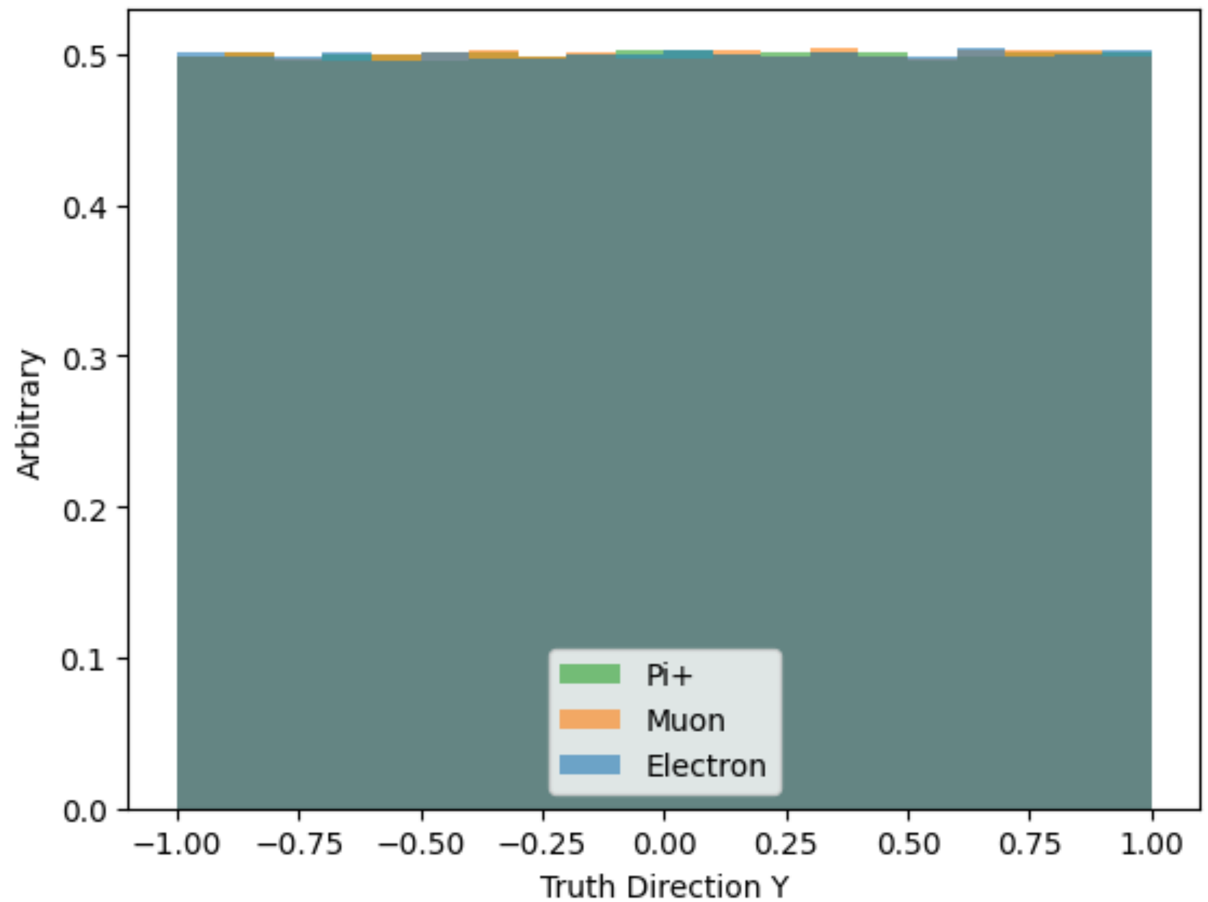

Pi+
Electron
Muon
Arbitrary
0.0000
0.0001
0.0002
0.0003
0.0004
0.0005
0.0006
0
500
1000
1500
2000
2500
3000
3500
4000
Number of PMTs

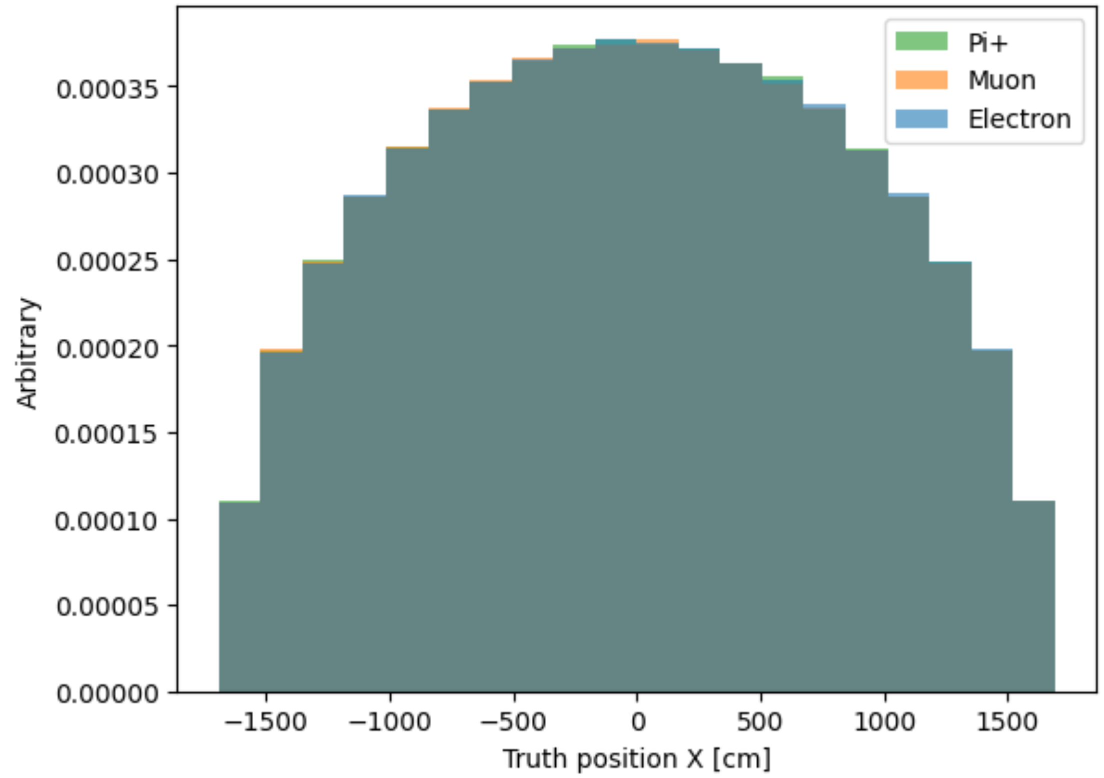

Pi+
Electron
Muon
Arbitrary
0.00000
0.00005
0.00010
0.00015
0.00020
0
2500
5000
7500
10000
12500
15000
17500
20000
Total Charge

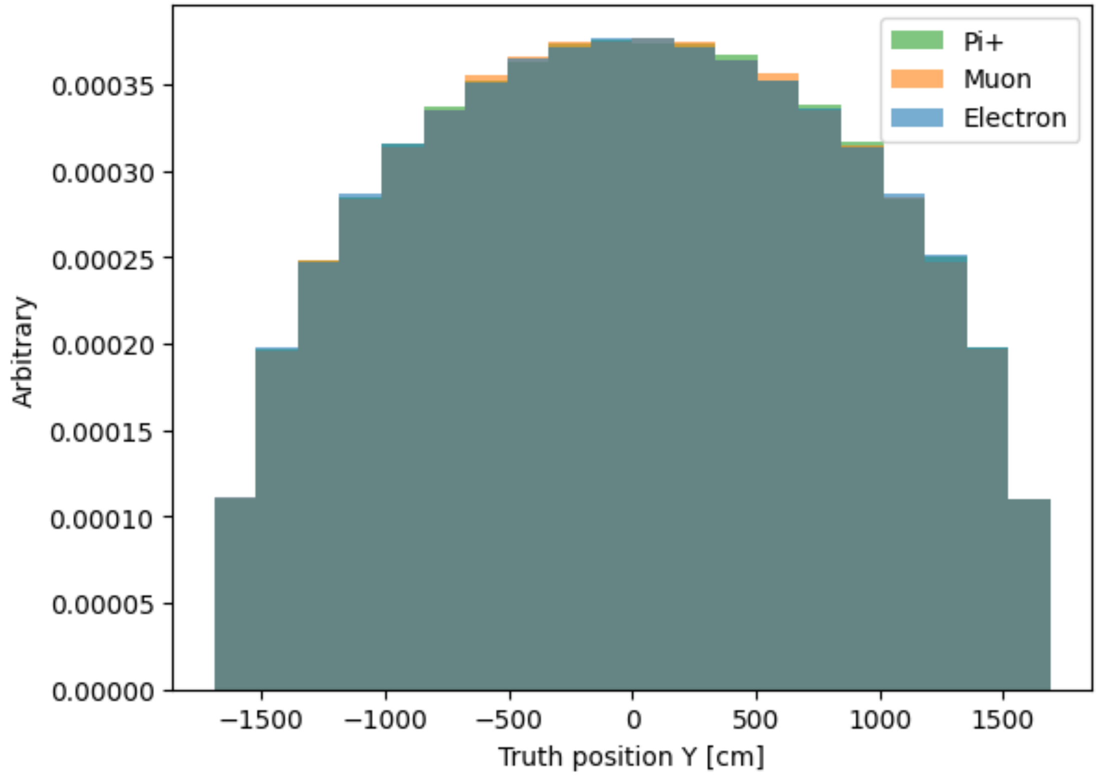

Pi+
Electron
Muon
Arbitrary
0.0000
0.0001
0.0002
0.0003
0.0004
0.0005
0.0006
0.0007
0.0008
0
200
400
600
800
1000
1200
Truth Visible Energy [MeV]

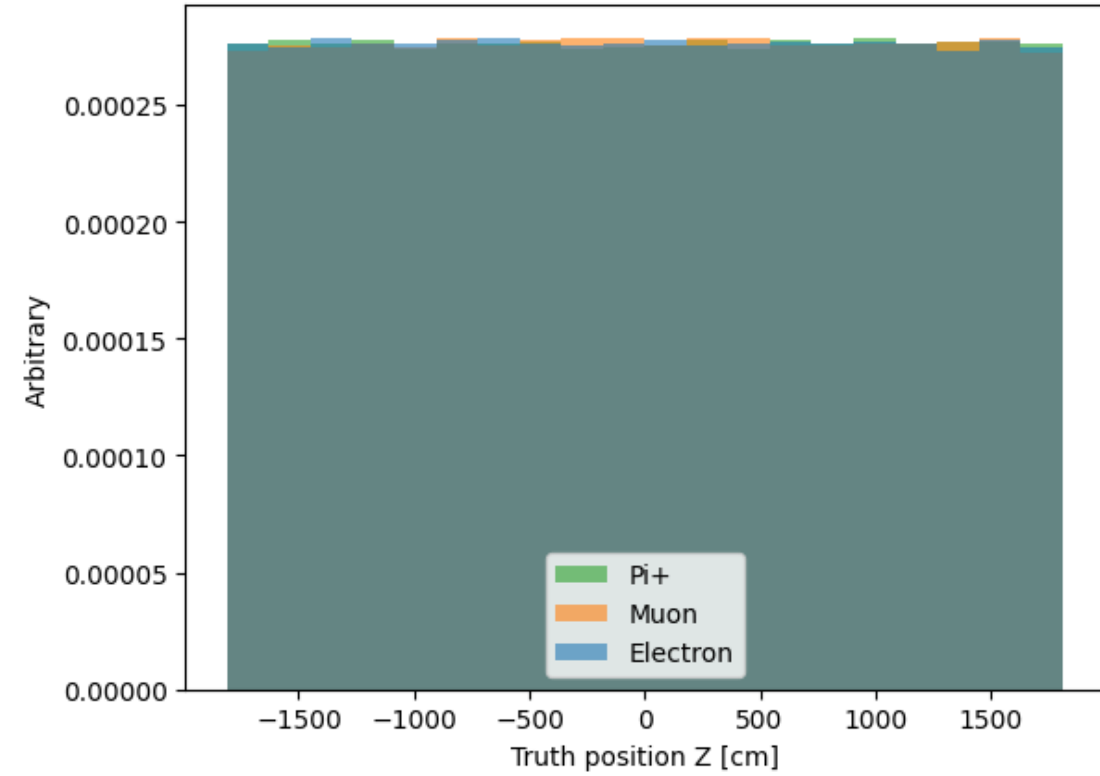

0.5
0.4
0.3
0.2
0.1
0.0
Arbitrary
Pi+
Muon
Electron
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Truth Direction X


0.5
0.4
0.3
0.2
0.1
0.0
Arbitrary
Pi+
Muon
Electron
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Truth Direction Z


0.5
0.4
0.3
0.2
0.1
0.0
Arbitrary
Pi+
Muon
Electron
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
Truth Direction Y

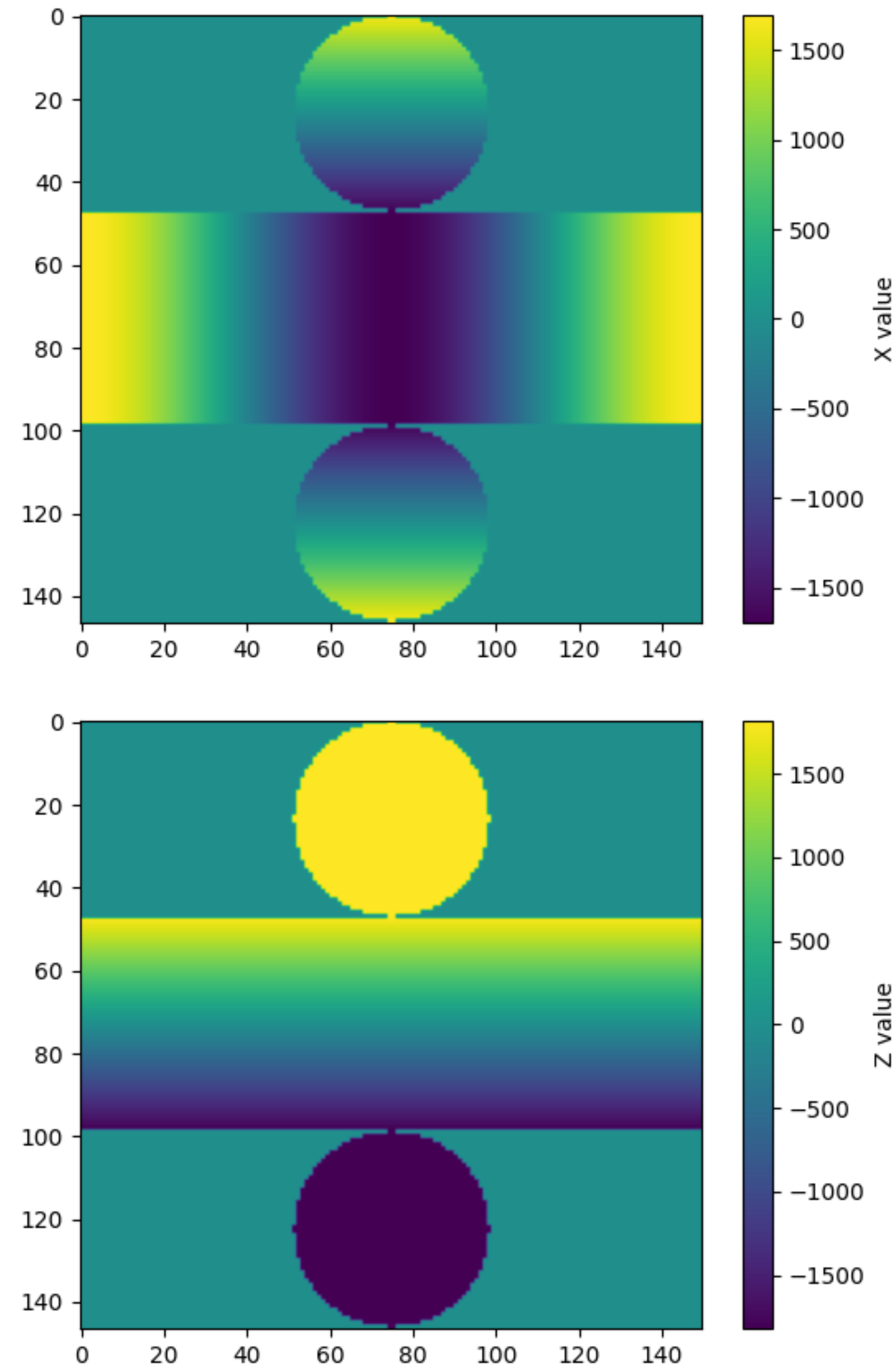

Pi+
Muon
Electron
Arbitrary
Truth position X [cm]


Pi+
Muon
Electron
Arbitrary
Truth position Y [cm]


Pi+
Muon
Electron
Arbitrary
Truth position Z [cm]

Training - ResNet

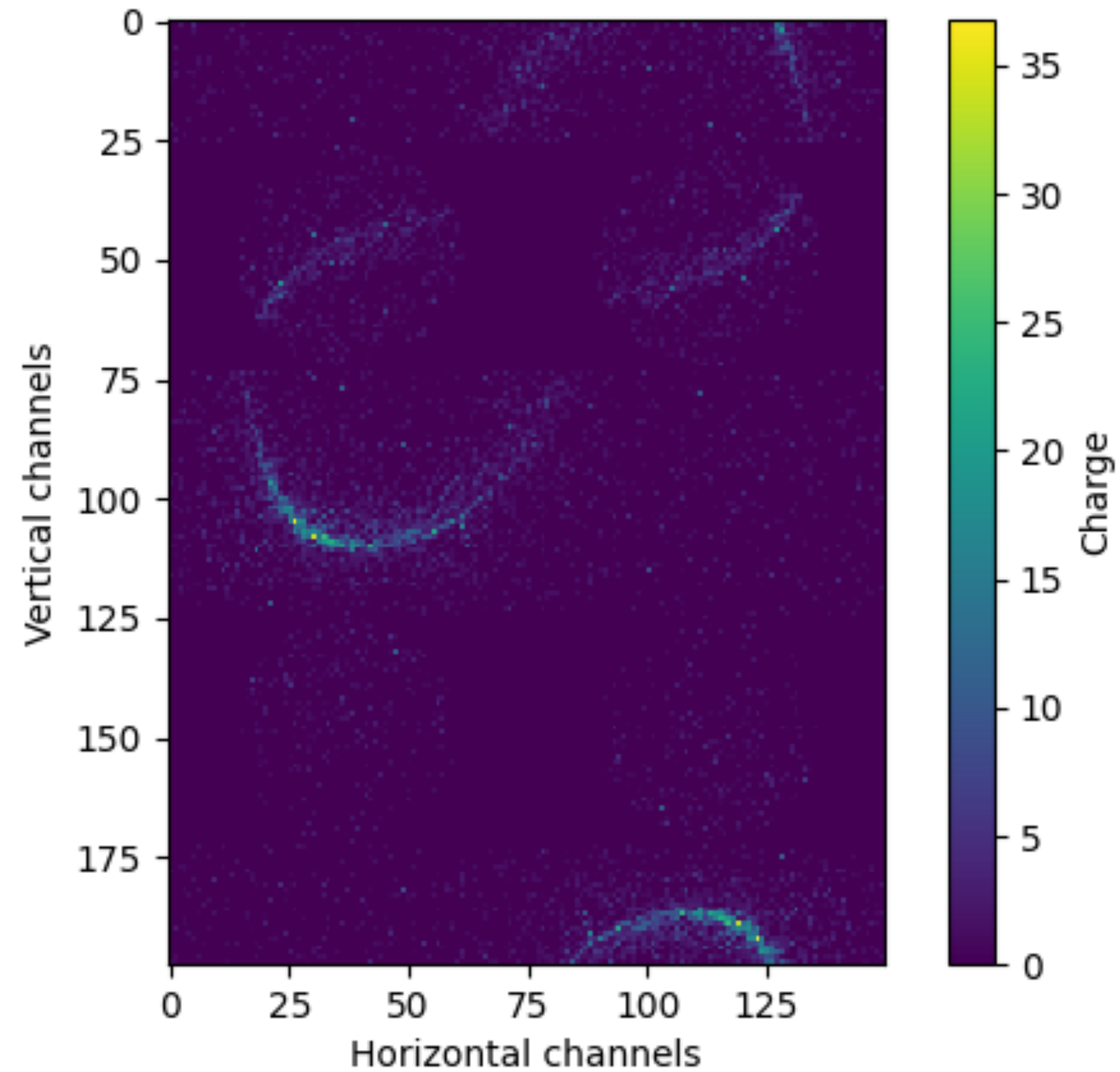
- Based on prior HyperK experience & r WCSIM, the **ResNet** ML architecture p performance
- ResNet is a CNN which requires input 2D map
    - See examples of this projection on the plc
    - Decision on how to unroll cylinder arbitrar with this



Double cover & Transforms

- One way which was found to work well to remove effects due to arbitrary choice of cylinder unrolling is double cover padding

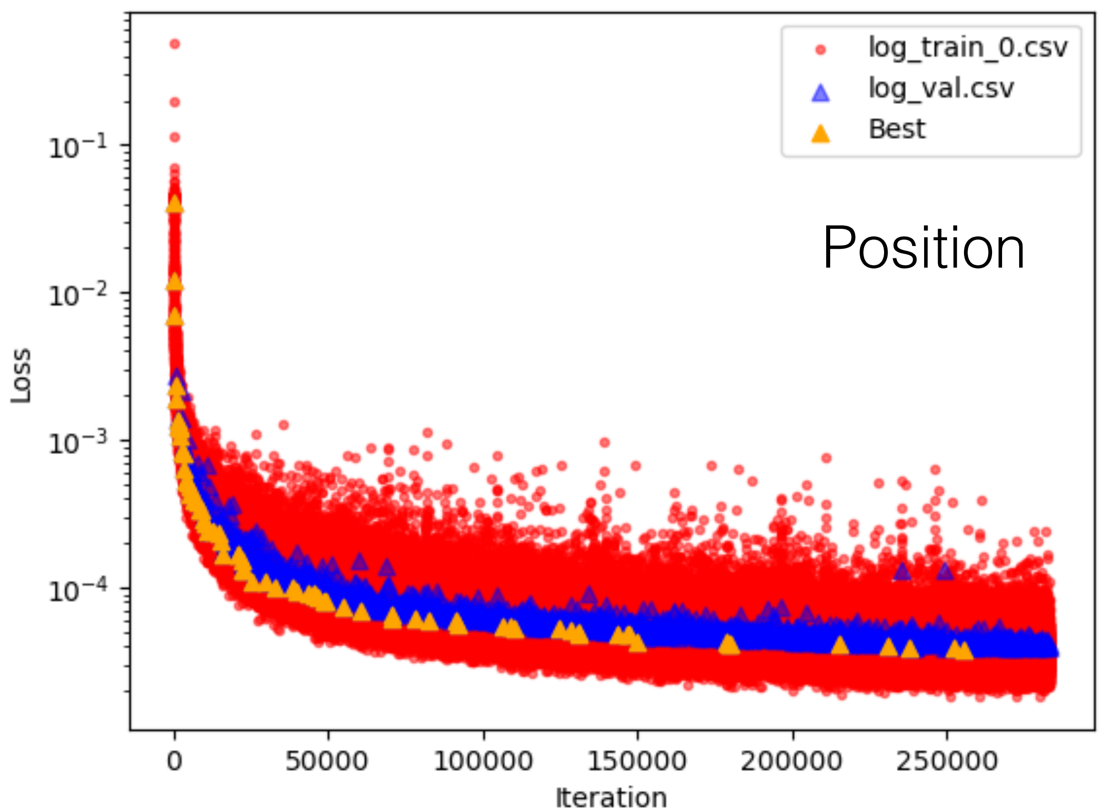
```
                         CBALKJIHGFED
       01                    01    32
       23                    23    10
ABCDEFGHIJKL  -->       DEFGHIJKLABC
MNOPQRSTUVWX            PQRSTUVWXMNO
       45                    45    76
       67                    67    54
                         ONMXWVUSTRQP
```

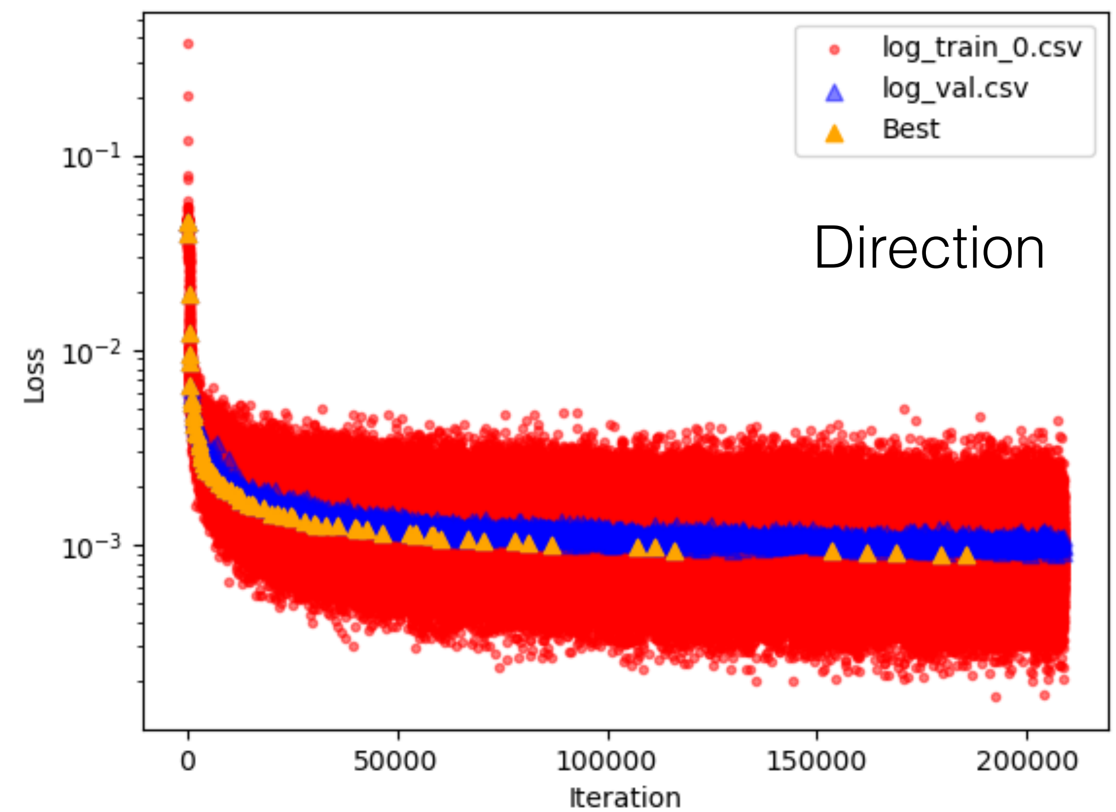


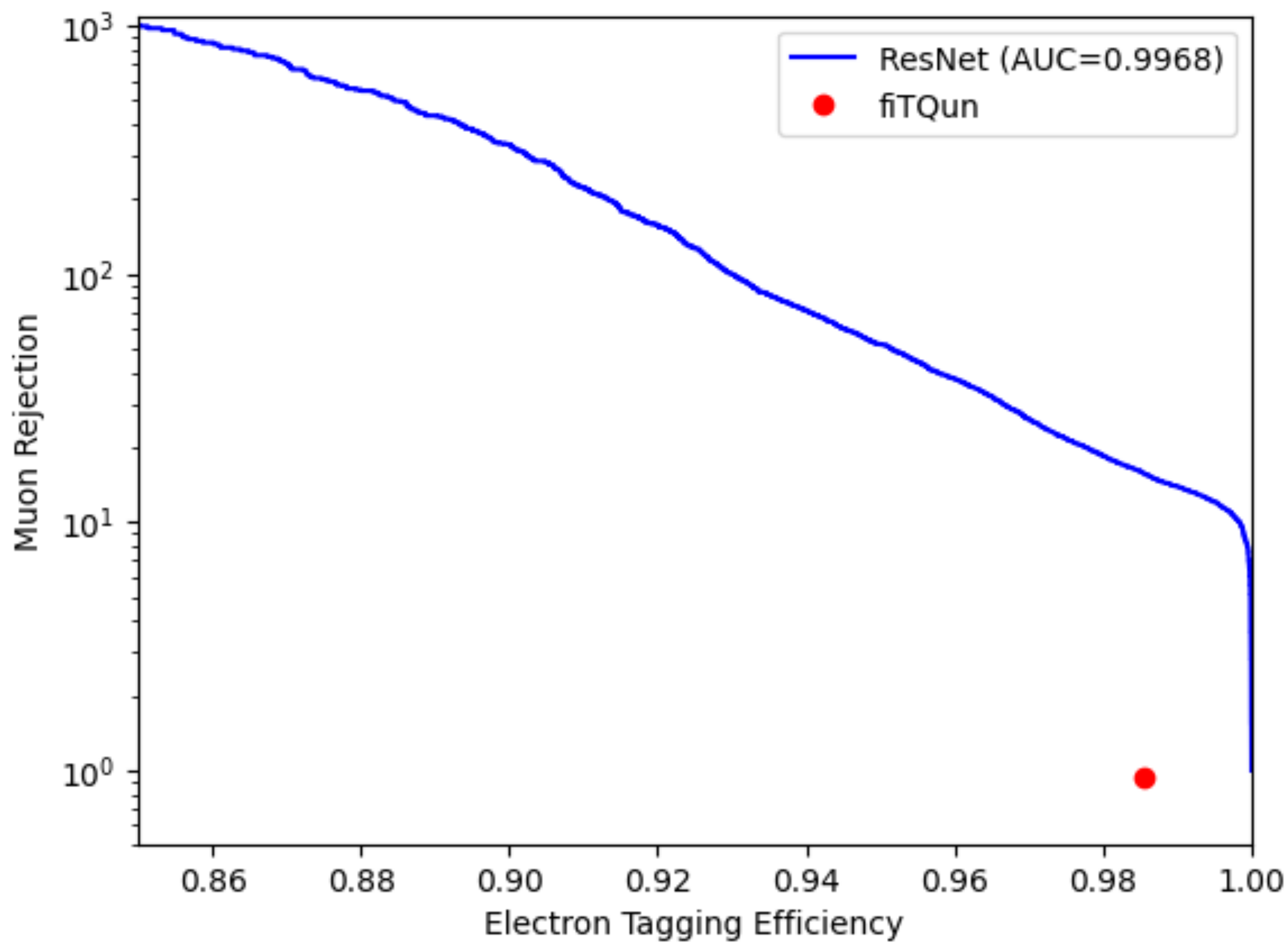
- Another augmentation to the data is transforms, where we flip horizontally, vertically and do a front/back reflection
  - Allows us to artificially inflate data, leading to higher stats

Position & Direction Regression

- Loss per iteration plots for position and direction
  - Each iteration is a batch which is ~100 events
  - Plots corresponds to ~10 epochs
- Shown is training loss and validation loss
  - Best is when validation loss is best yet during training

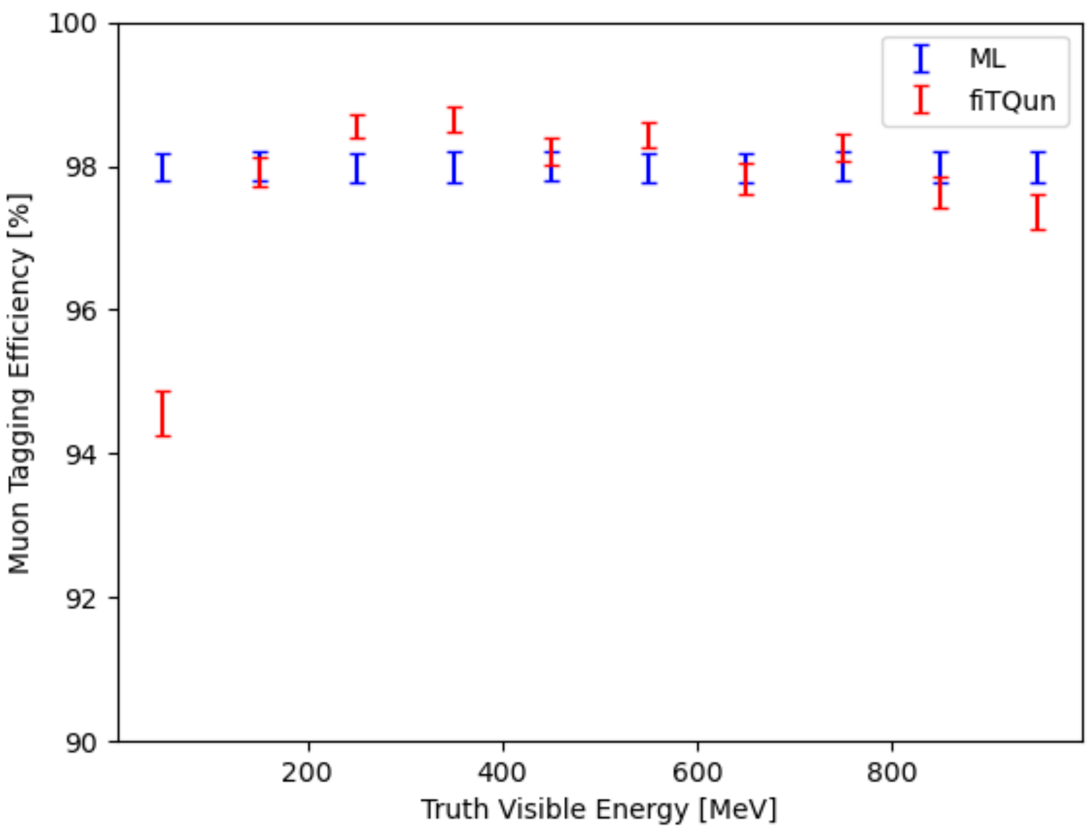


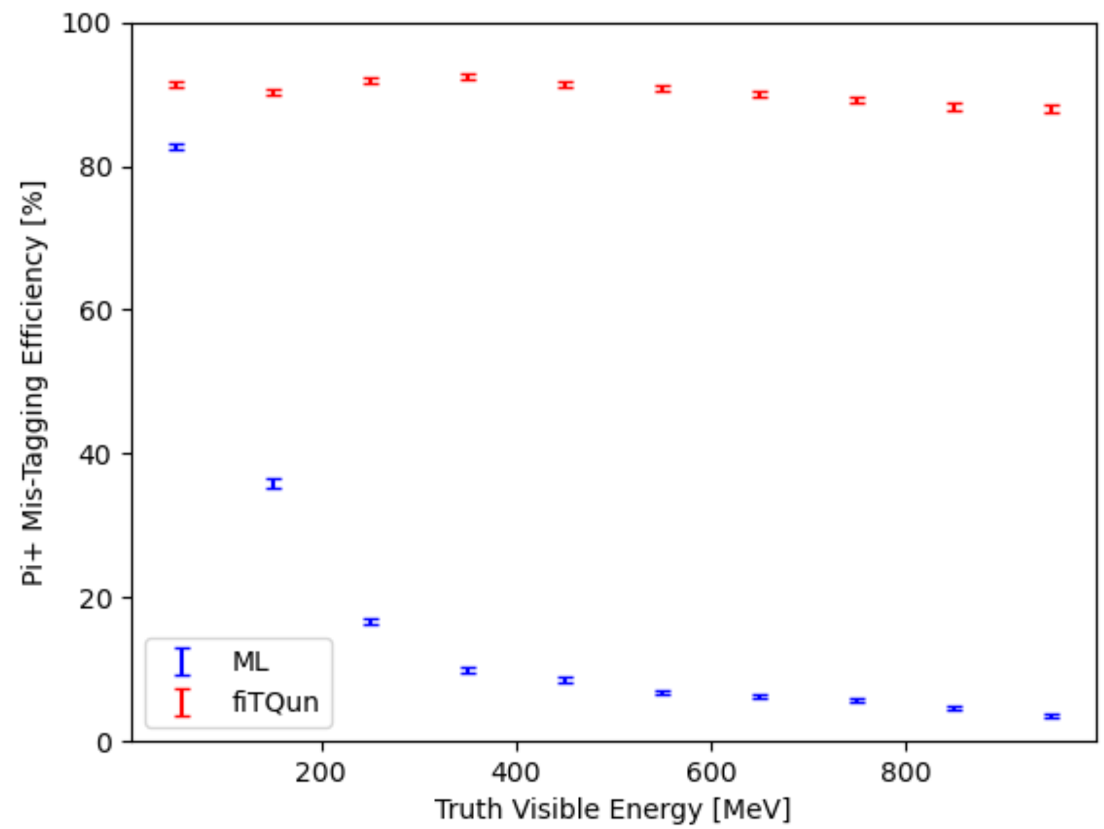


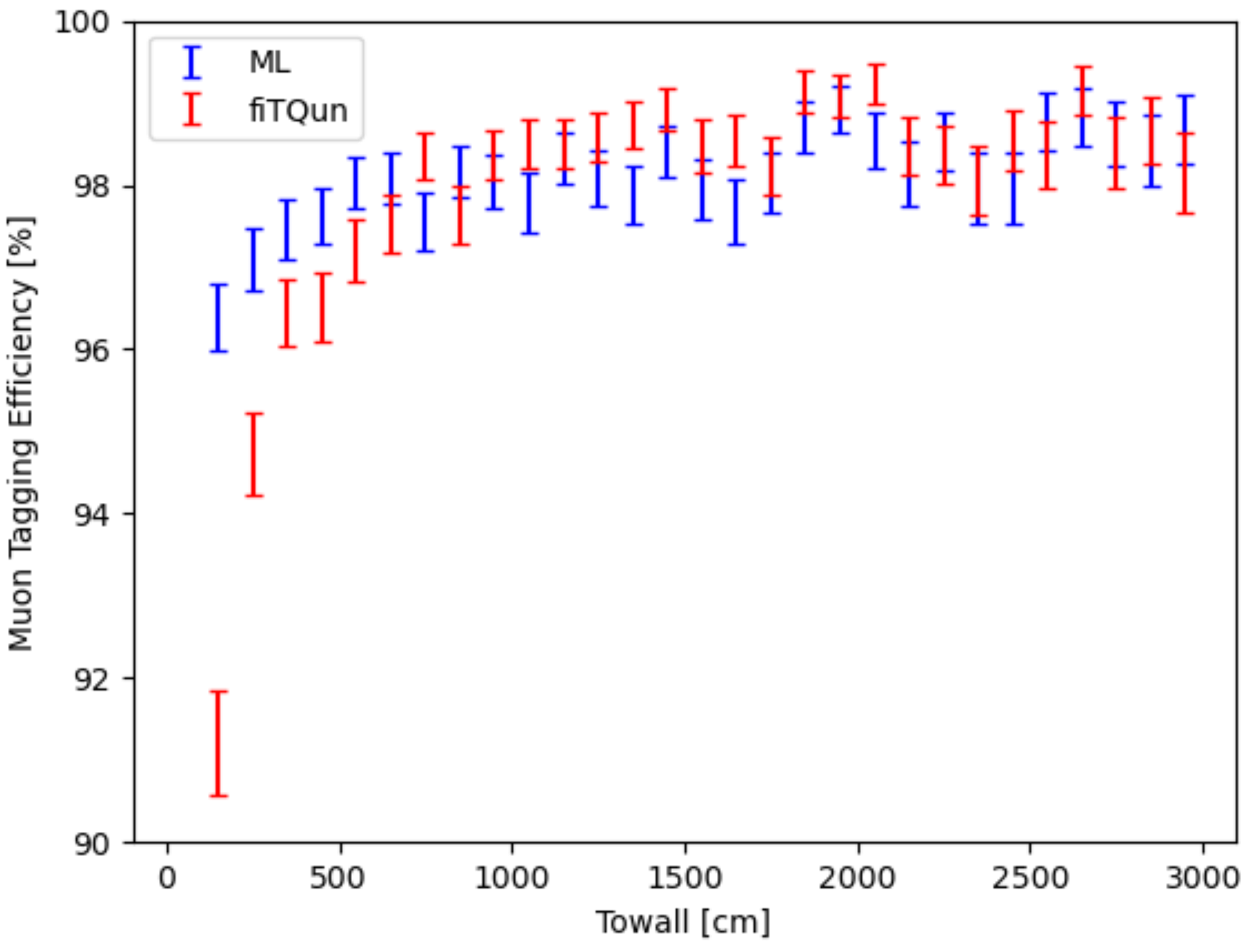

ResNet (AUC=0.9968)
fiTQun
Muon Rejection
10^3
10^2
10^1
10^0
0.86
0.88
0.90
0.92
0.94
0.96
0.98
1.00
Electron Tagging Efficiency

*Added after talk

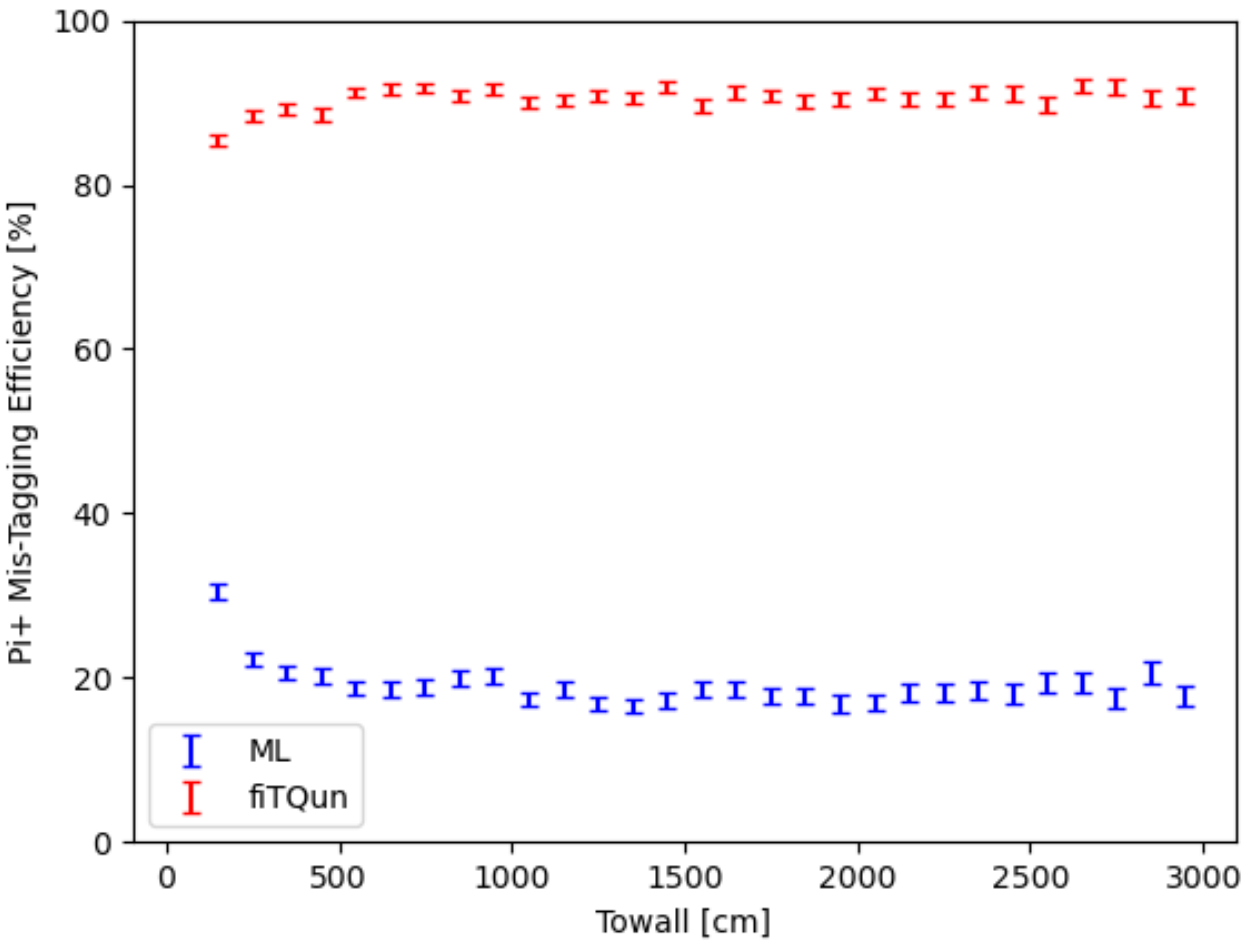
- Muon/Pi+ Classification performance as function of truth visible energy
  - Truth visible energy defined as initial particle energy over Cherenkov threshold



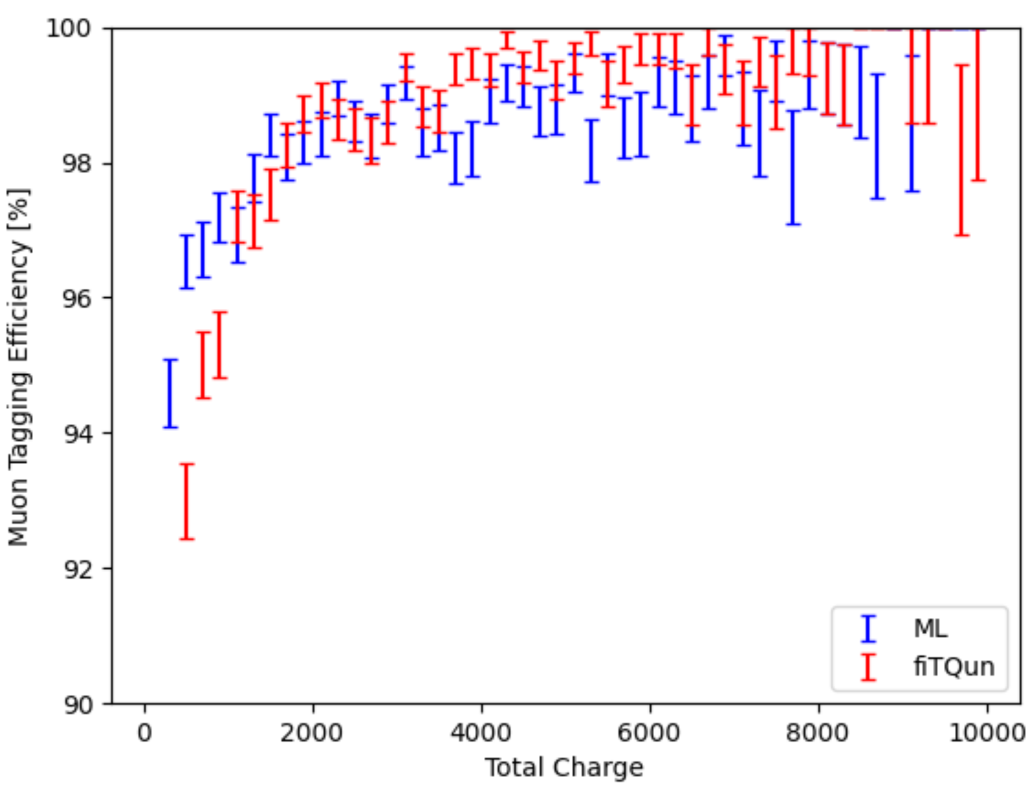


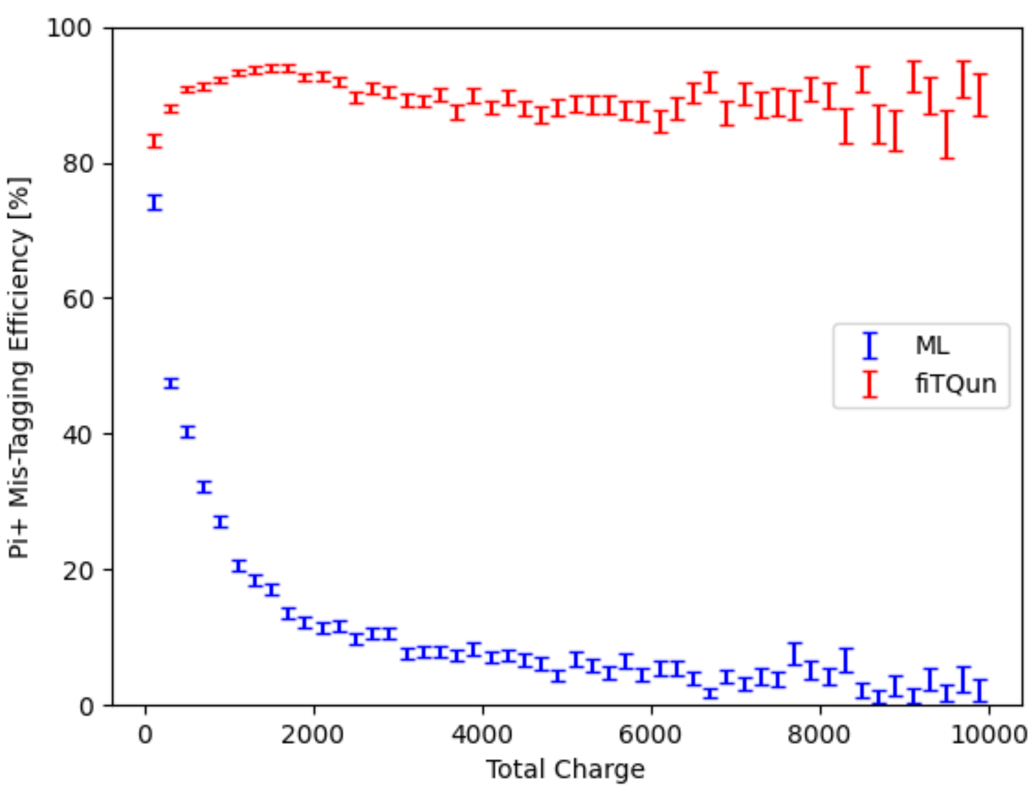


Muon Tagging Efficiency [%]
Towall [cm]
ML
fiTQun

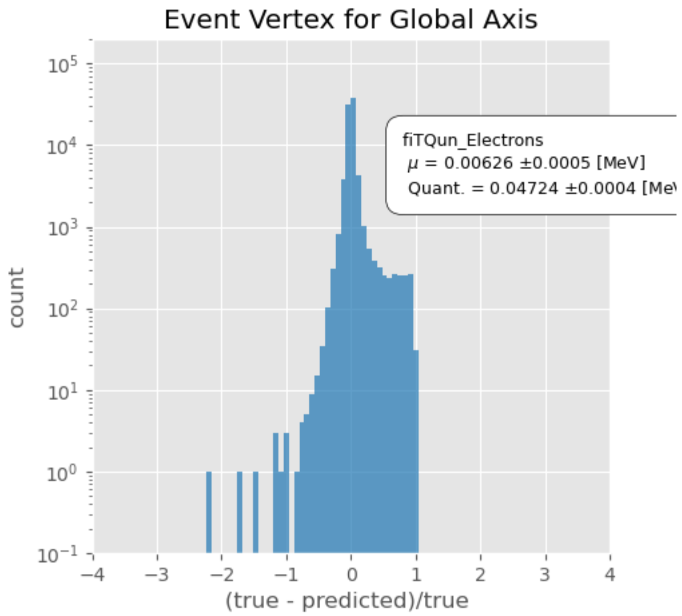


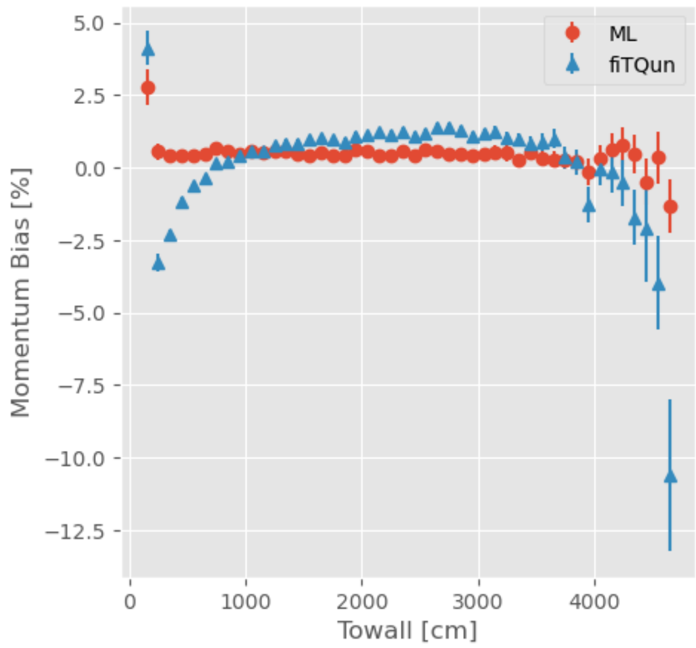
Pi+ Mis-Tagging Efficiency [%]
Towall [cm]
ML
fiTQun

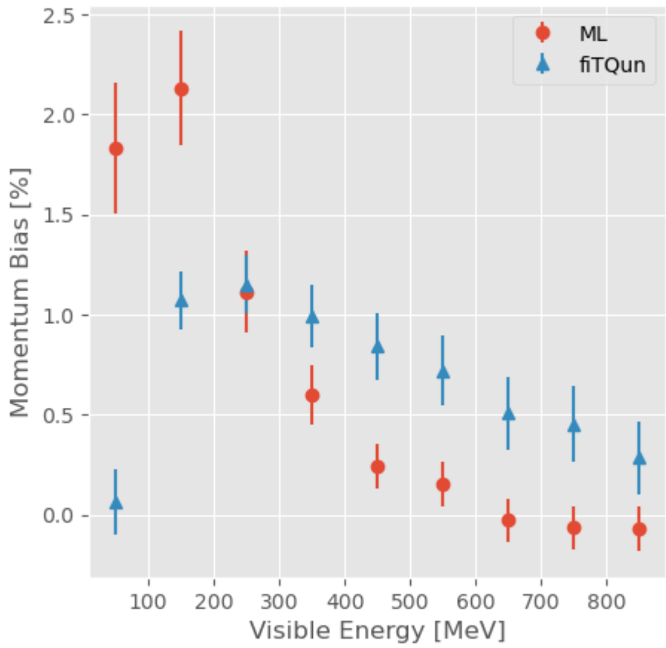

100
98
96
94
92
90
Muon Tagging Efficiency [%]
0
2000
4000
6000
8000
10000
Total Charge
ML
fiTQun

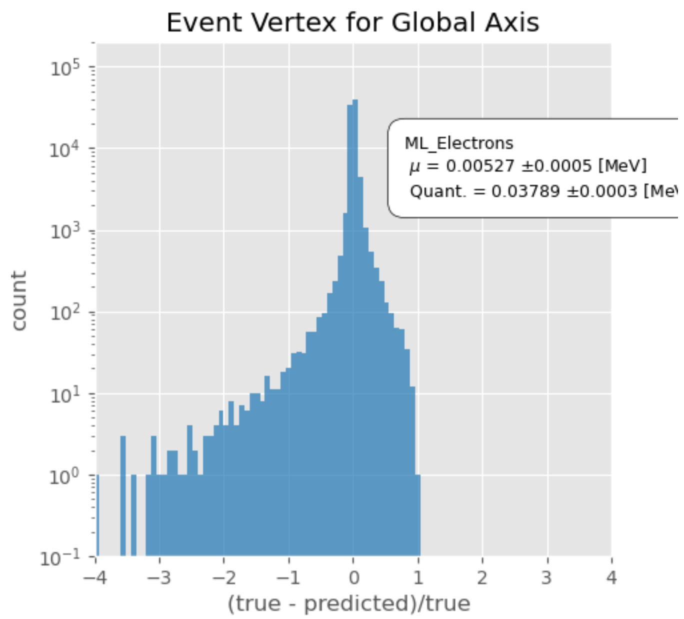

100
80
60
40
20
0
Pi+ Mis-Tagging Efficiency [%]
0
2000
4000
6000
8000
10000
Total Charge
ML
fiTQun

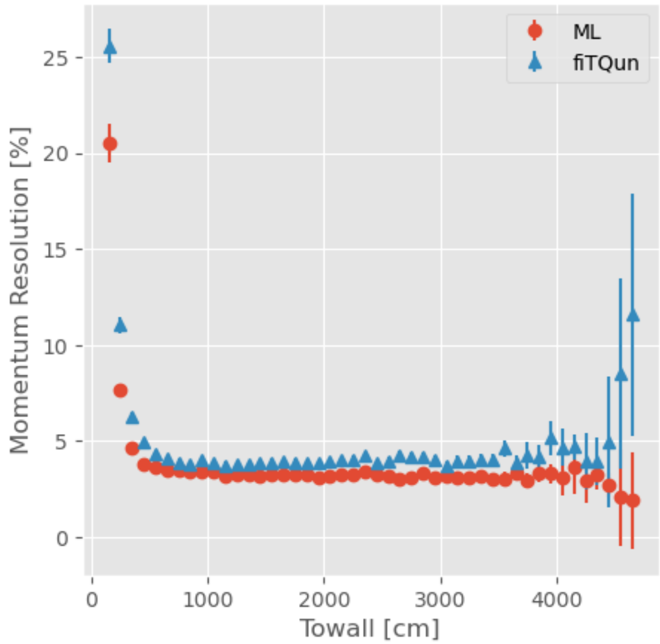
Electron - Momentum

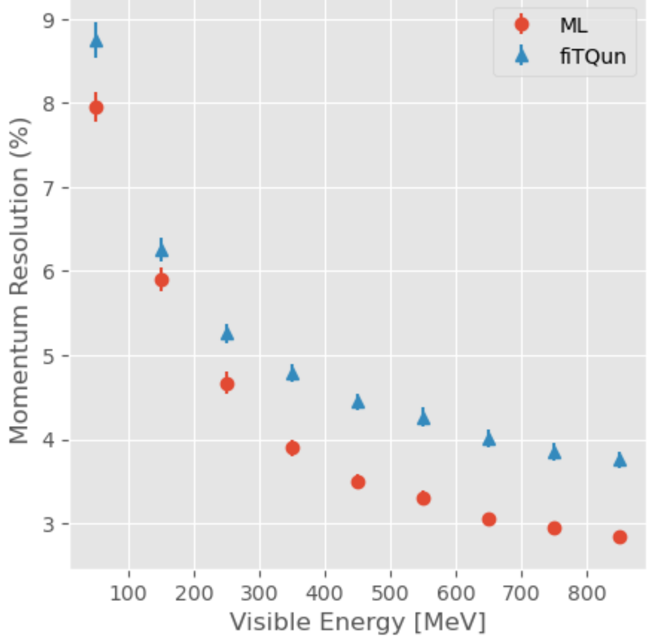




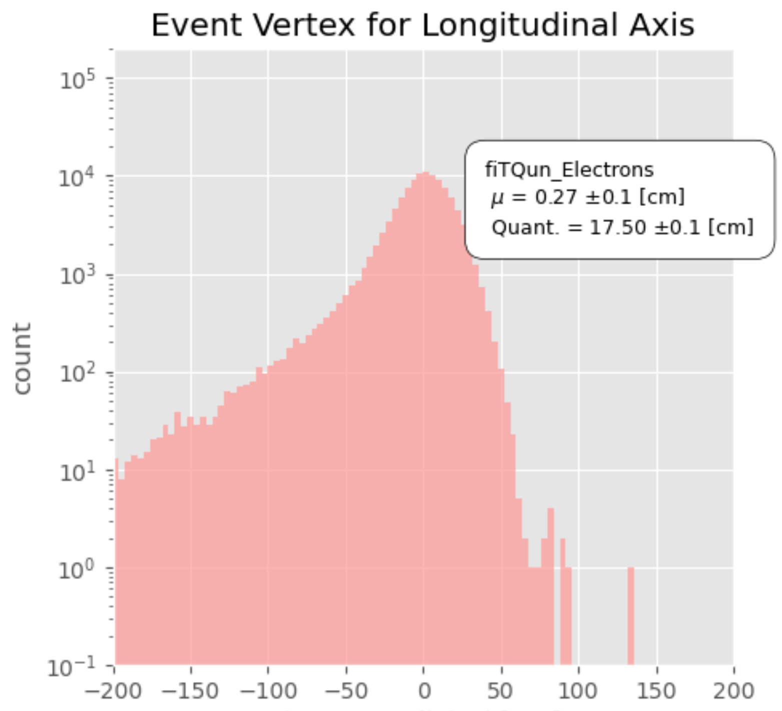


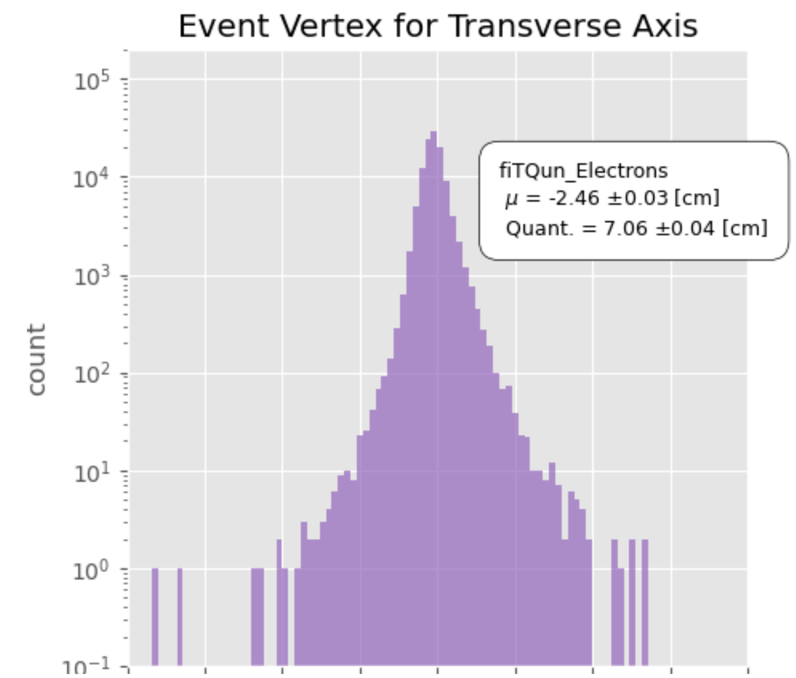


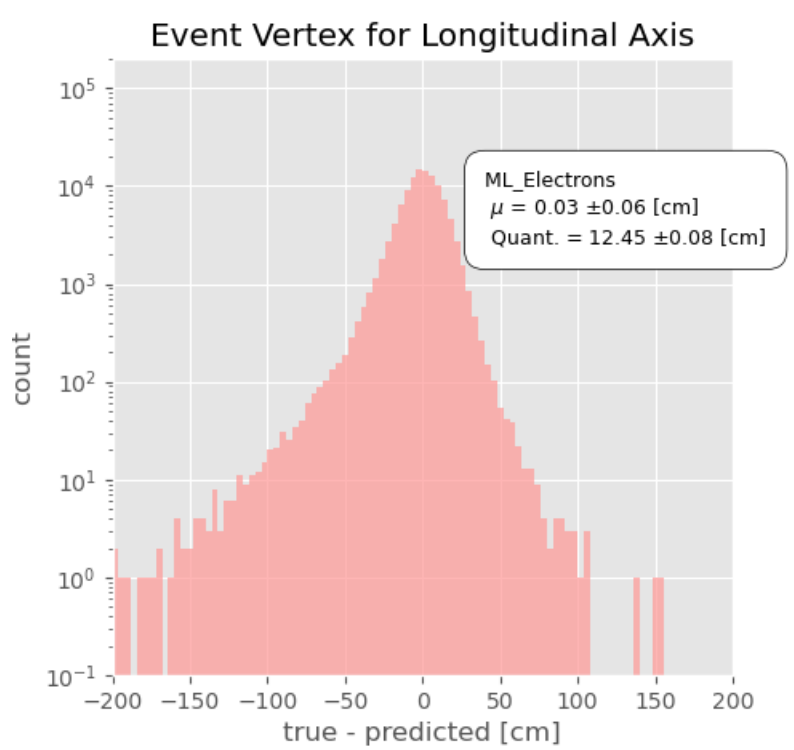


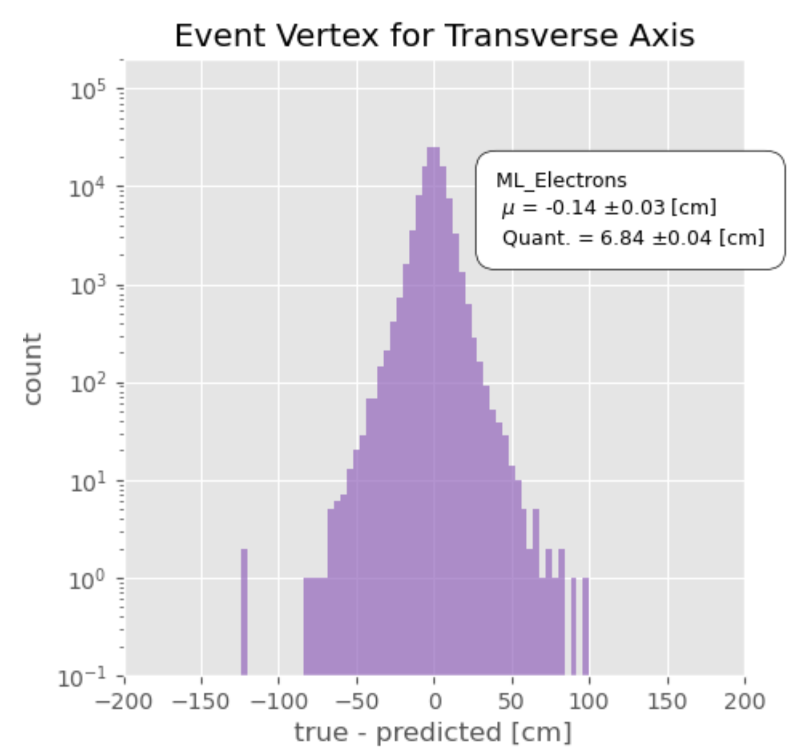


Electron - Position

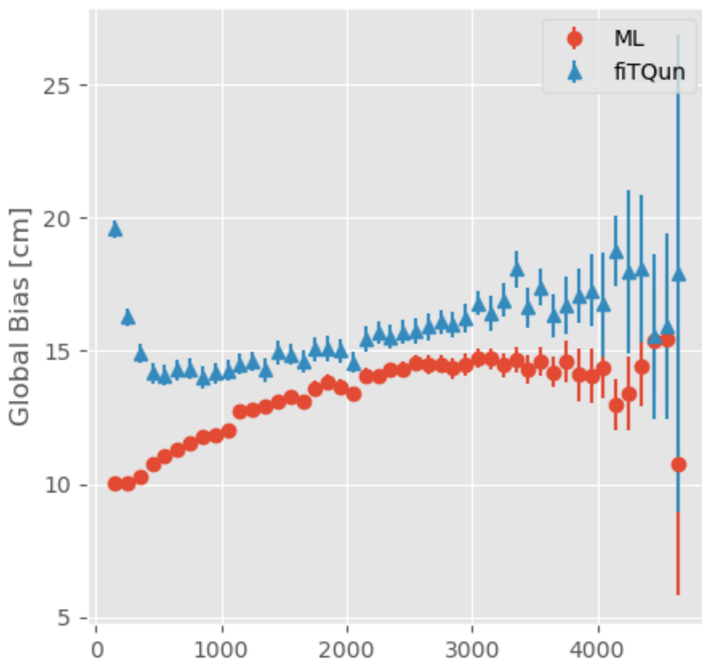


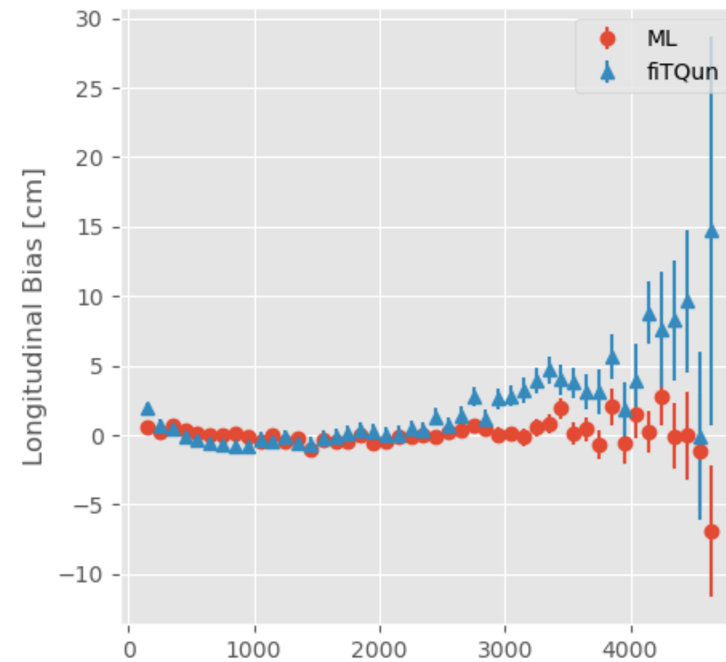


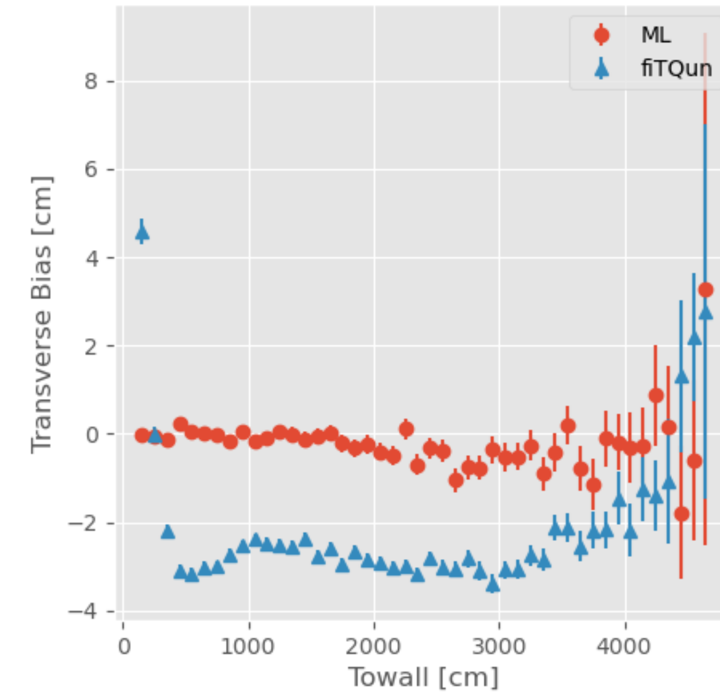


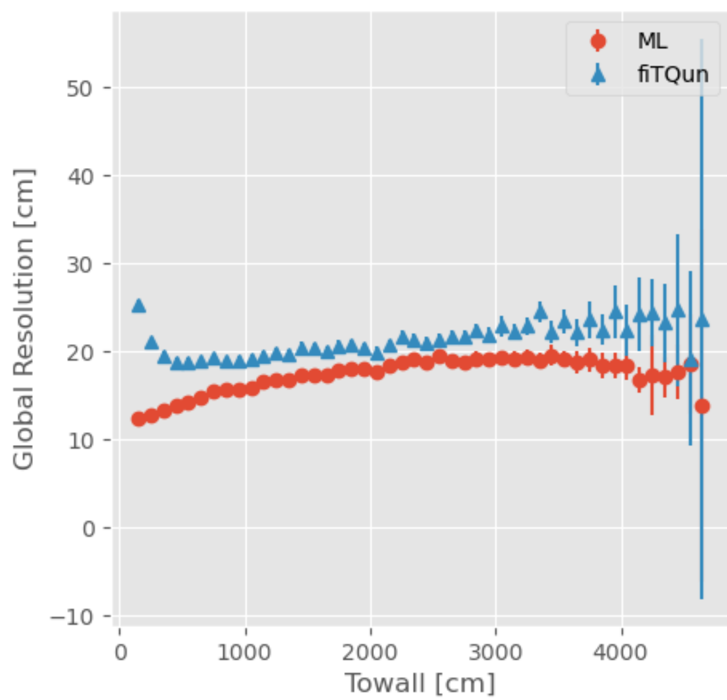


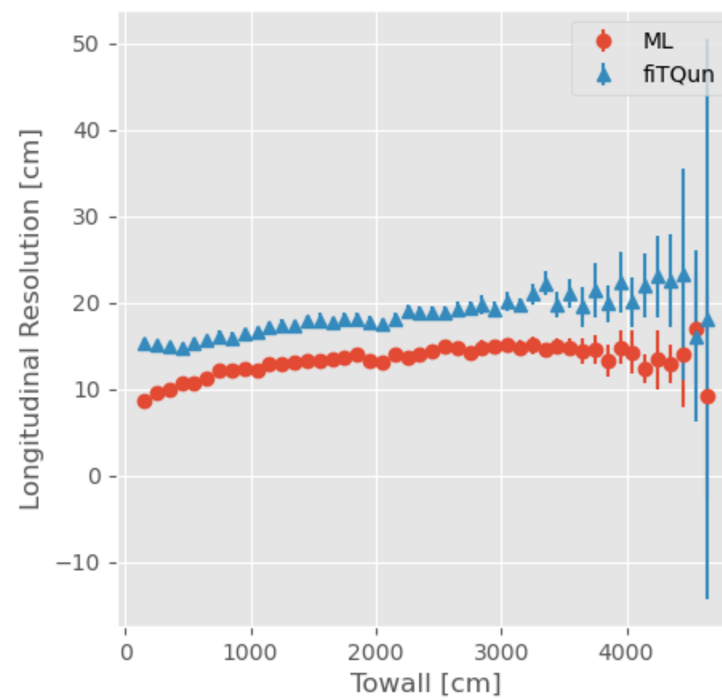
Electron - Position - Towall

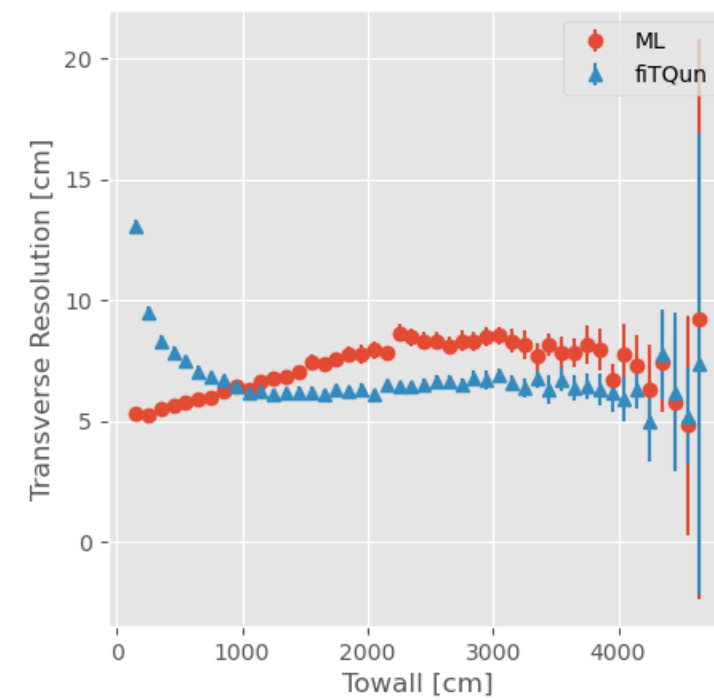


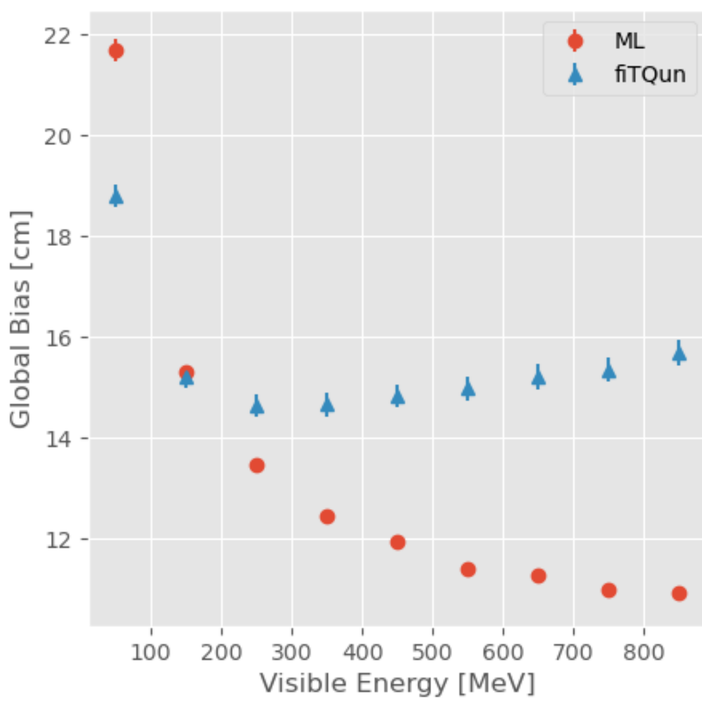




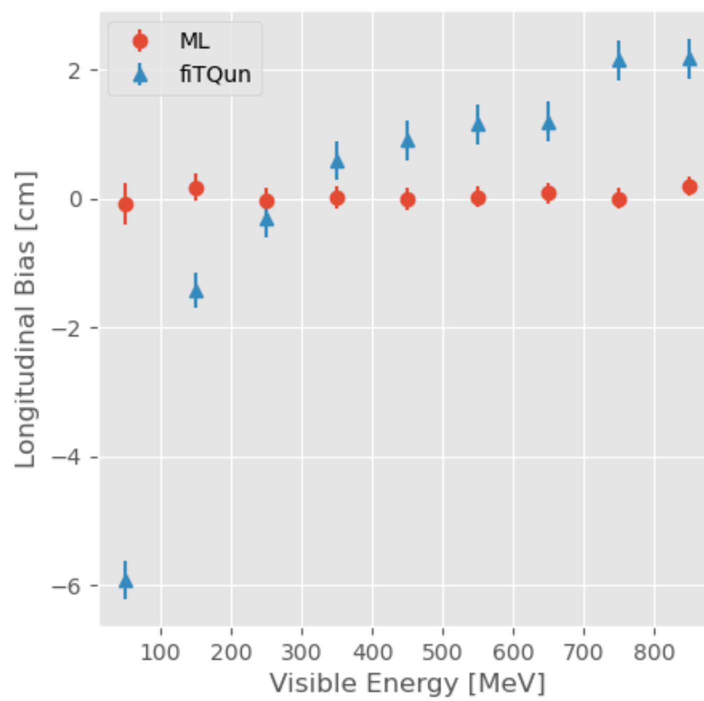




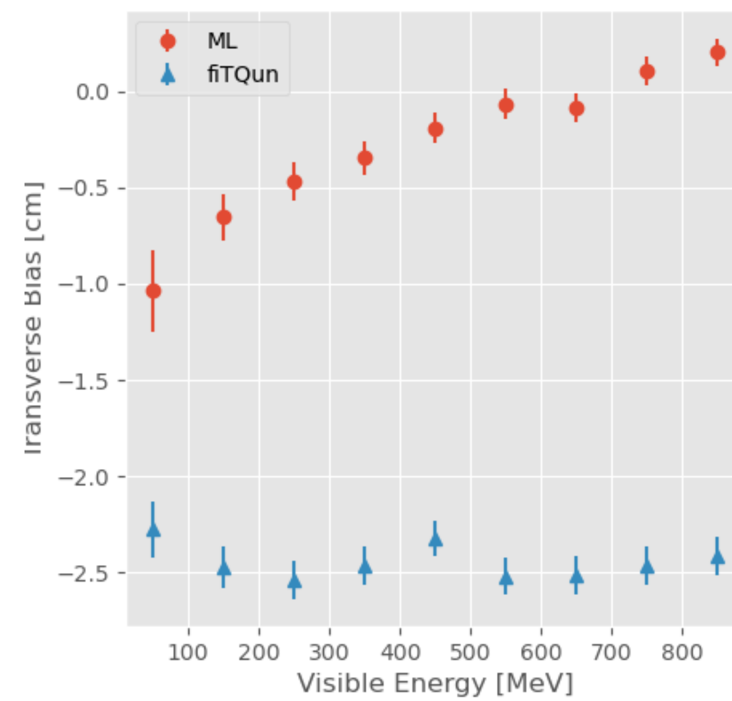




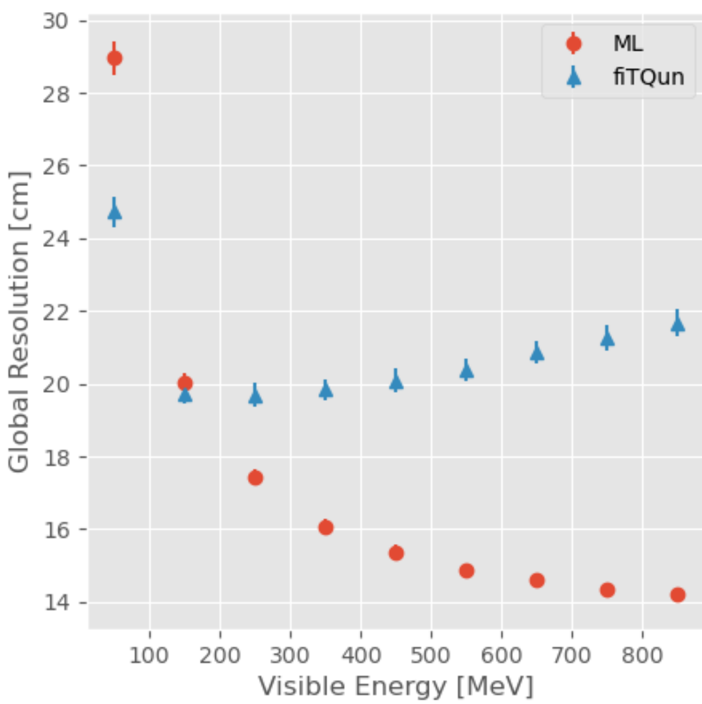
ML
fiTQun
Global Bias [cm]
Visible Energy [MeV]



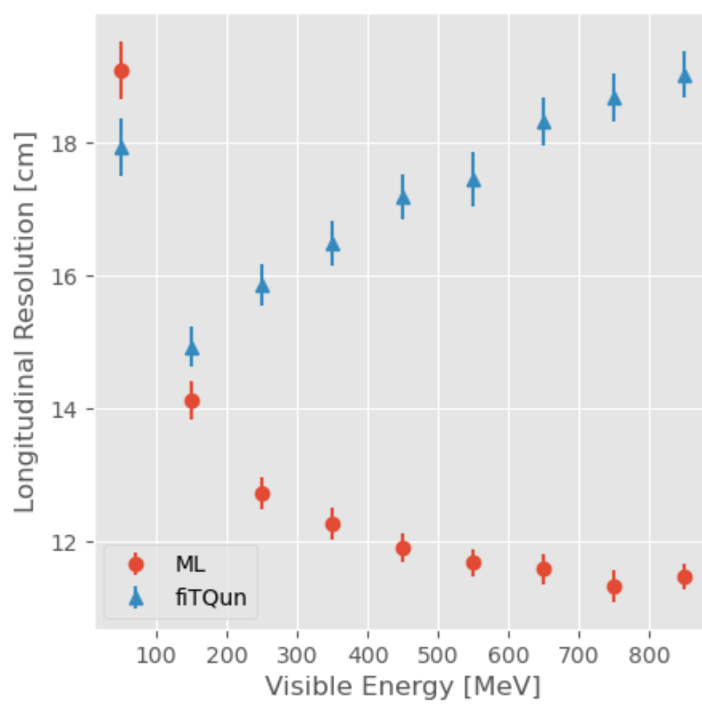
ML
fiTQun
Longitudinal Bias [cm]
Visible Energy [MeV]



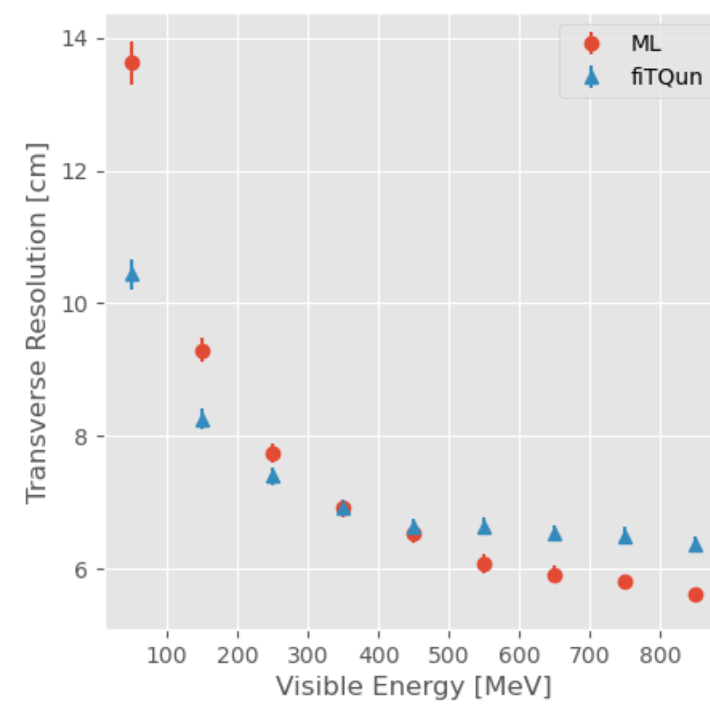
ML
fiTQun
Iransverse Bias [cm]
Visible Energy [MeV]



ML
fiTQun
Global Resolution [cm]
Visible Energy [MeV]

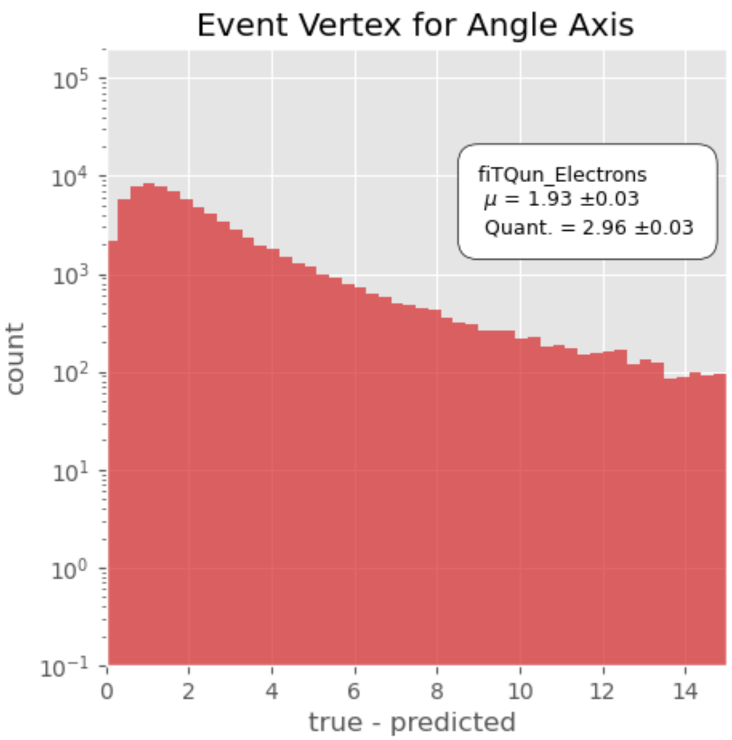


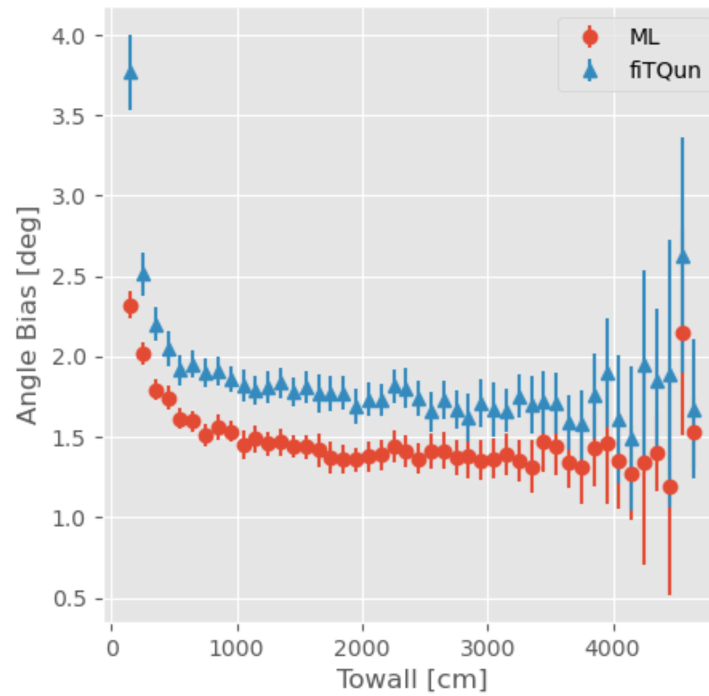
ML
fiTQun
Longitudinal Resolution [cm]
Visible Energy [MeV]

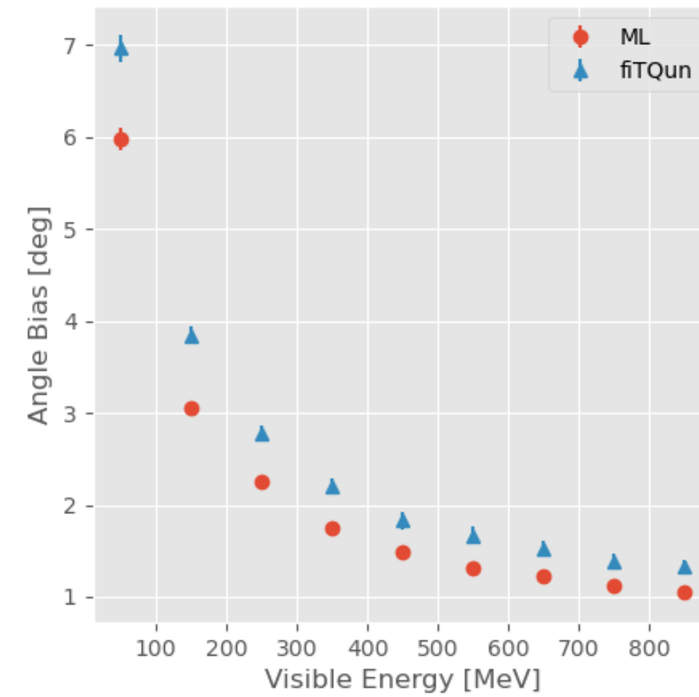


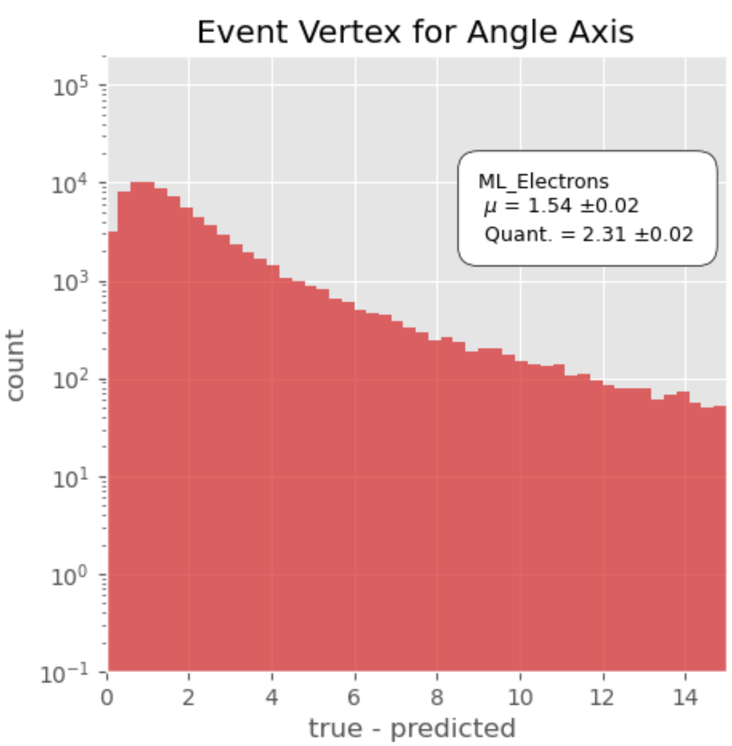
ML
fiTQun
Transverse Resolution [cm]
Visible Energy [MeV]

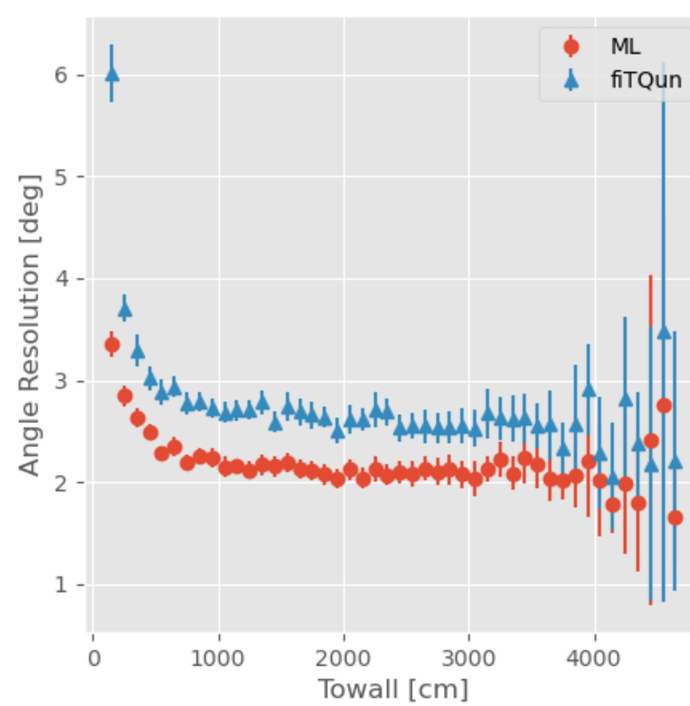
Electron - Direction

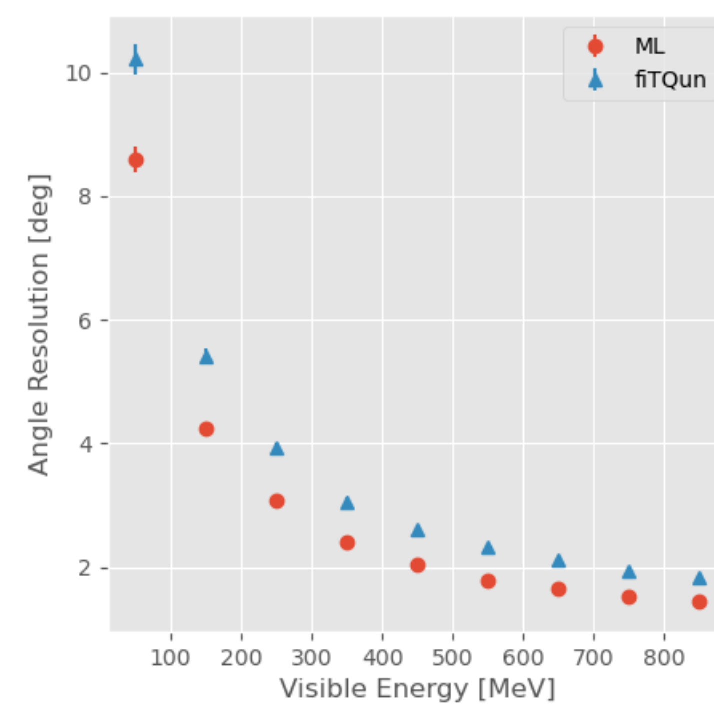




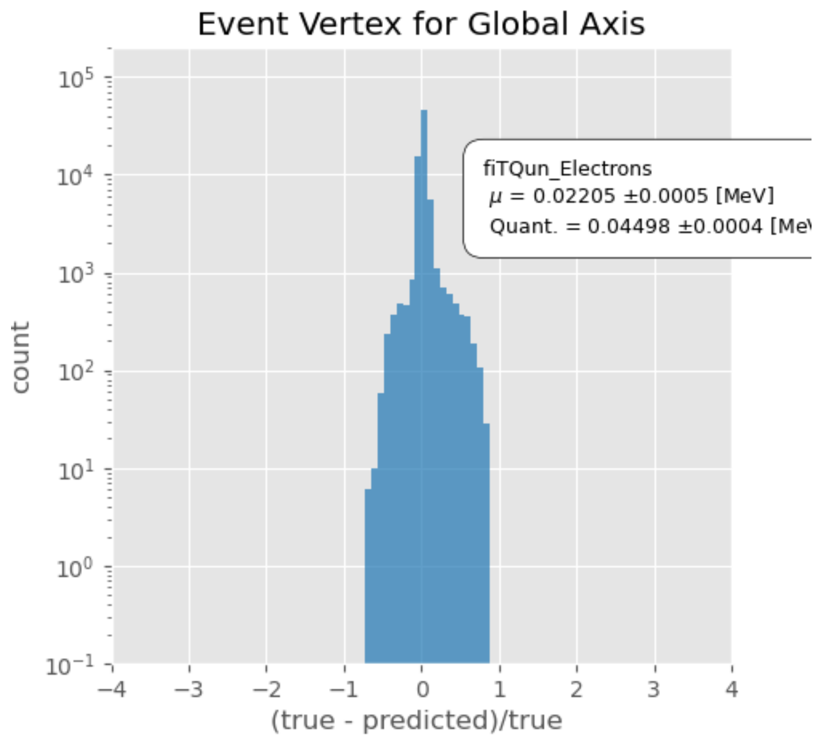


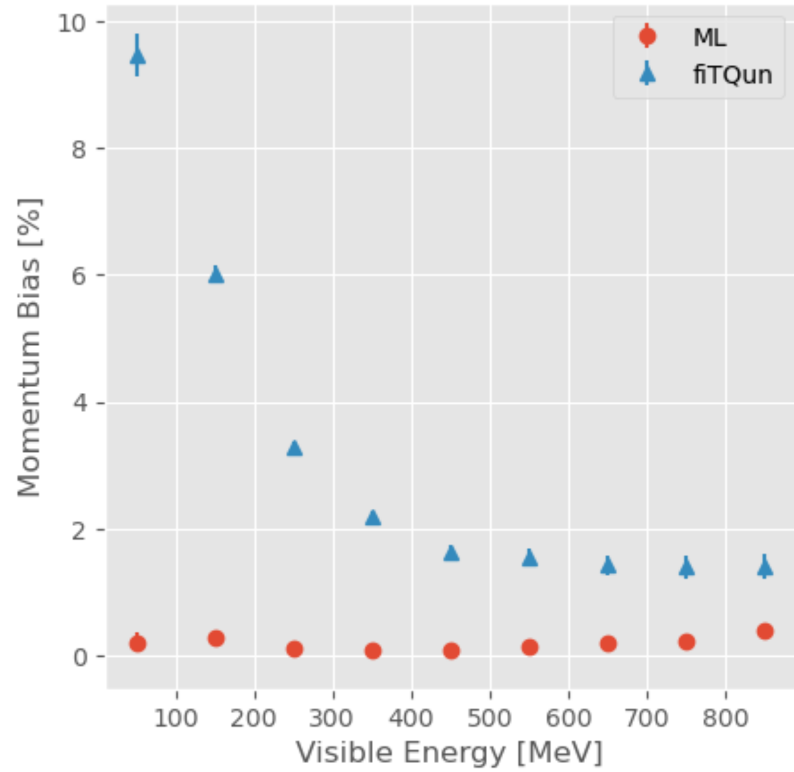


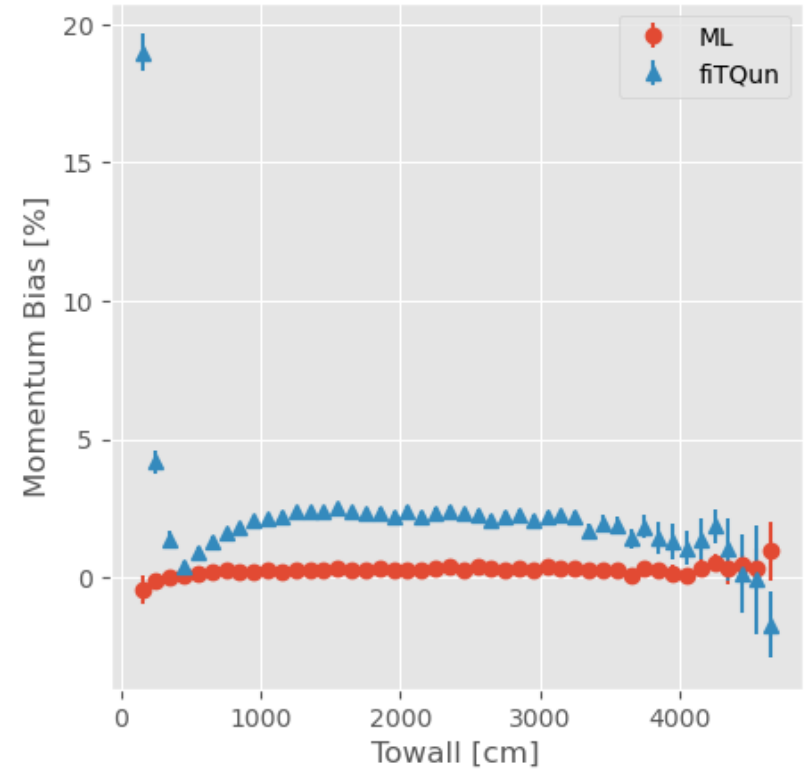


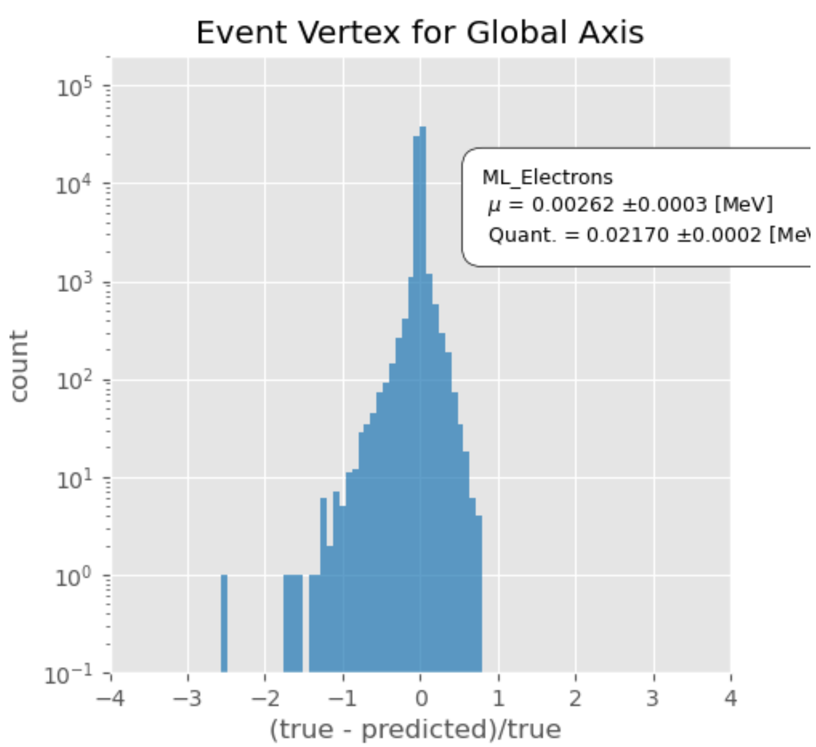


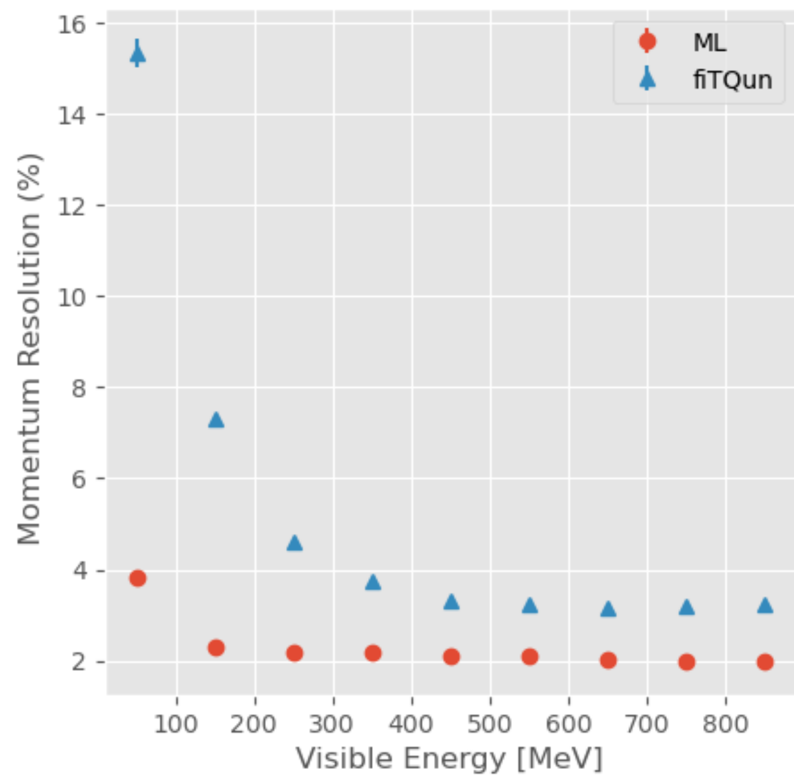
Muon - momentum

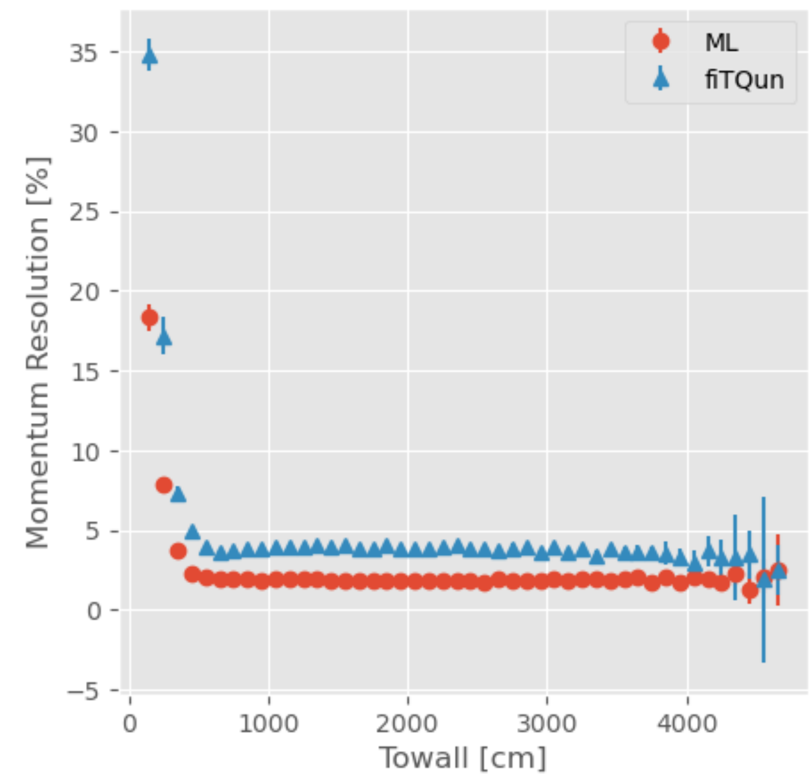




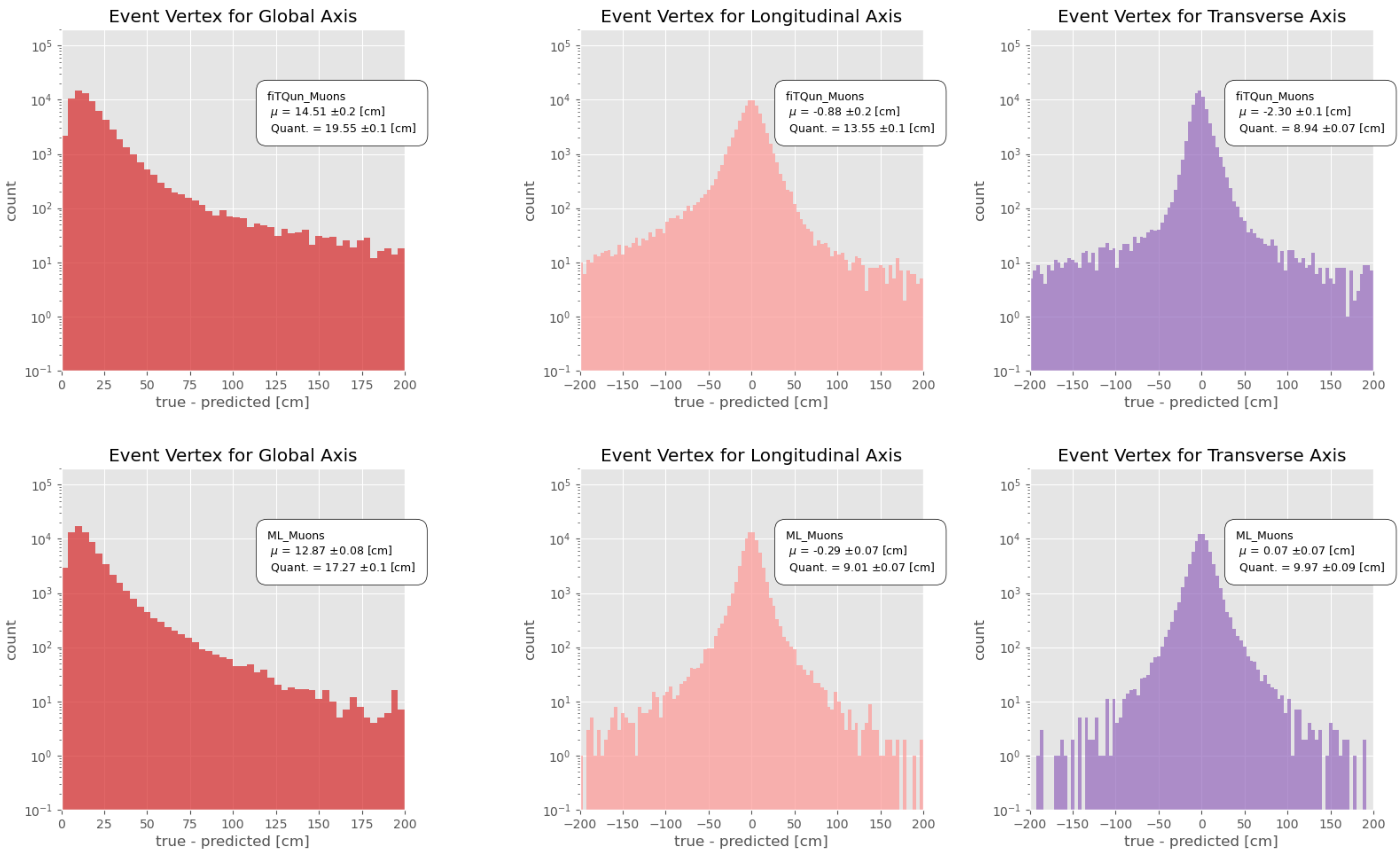




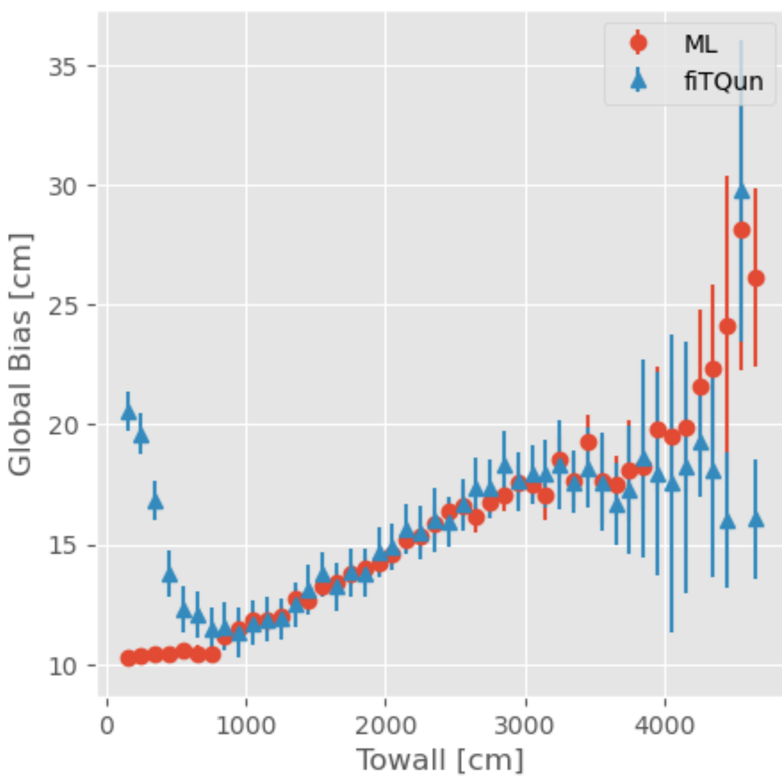


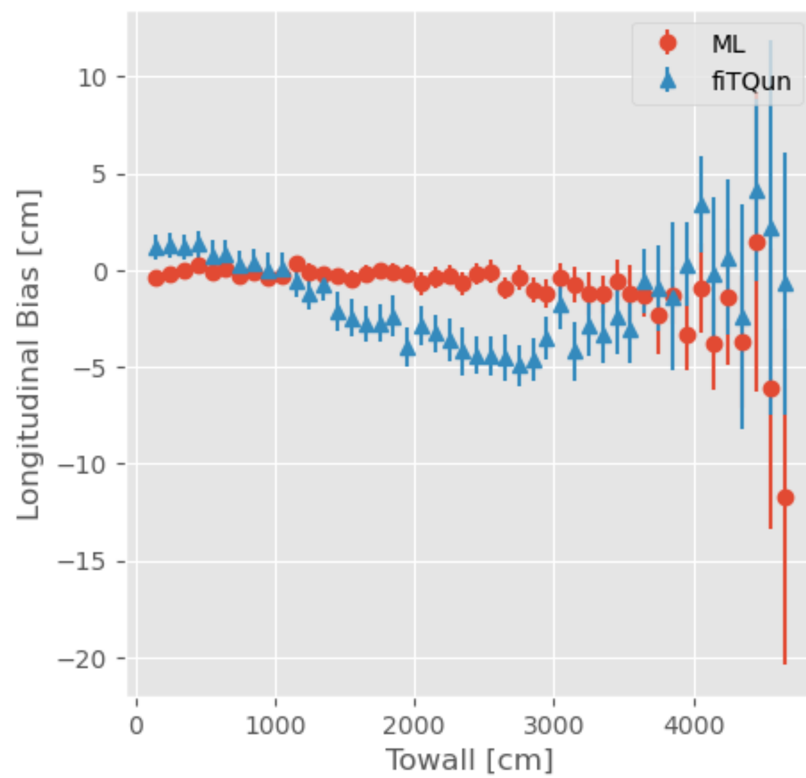


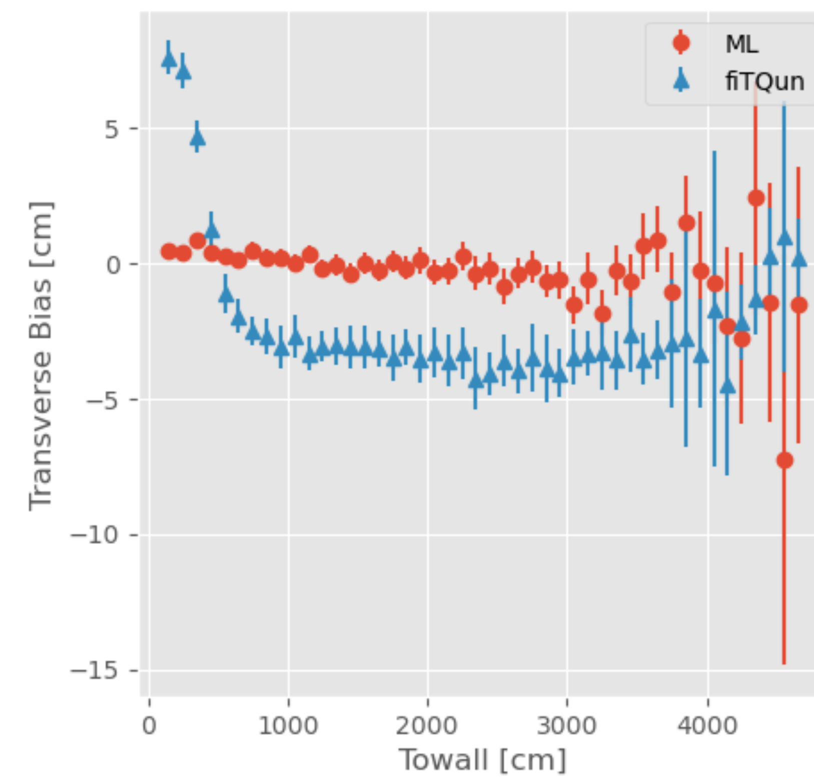
Muon - position

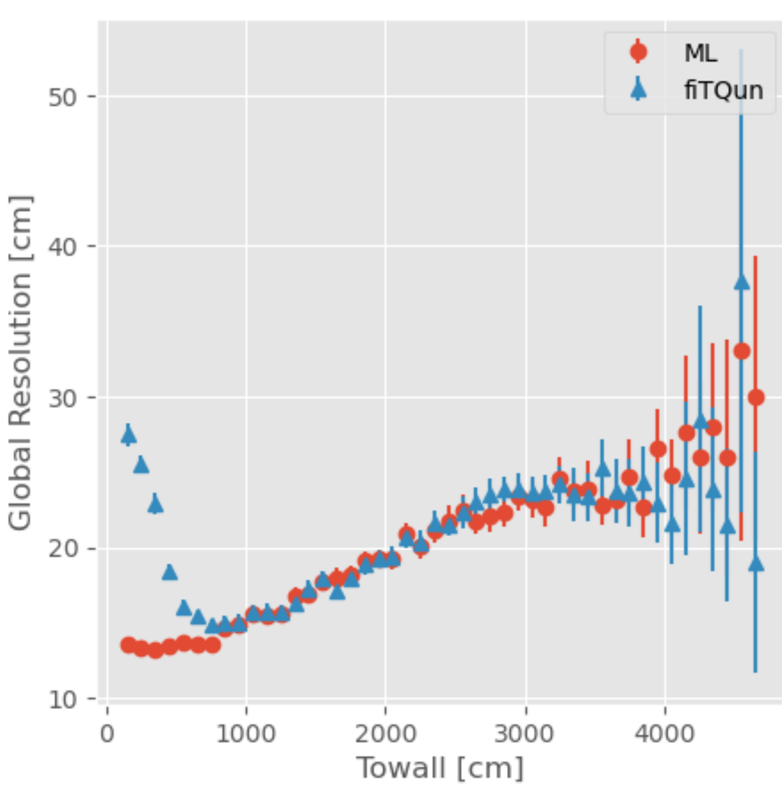


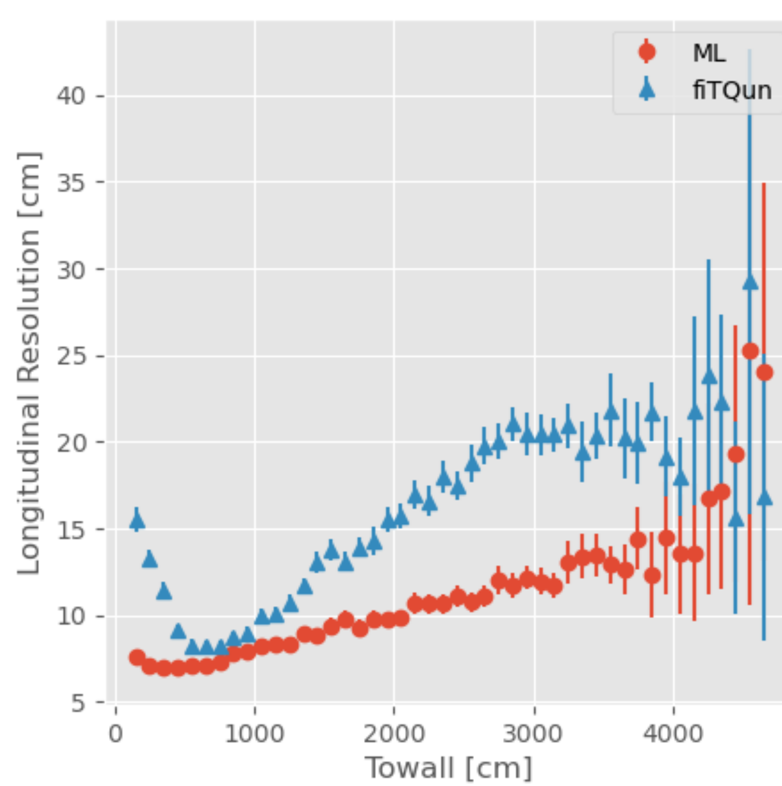
Muon - position - towall

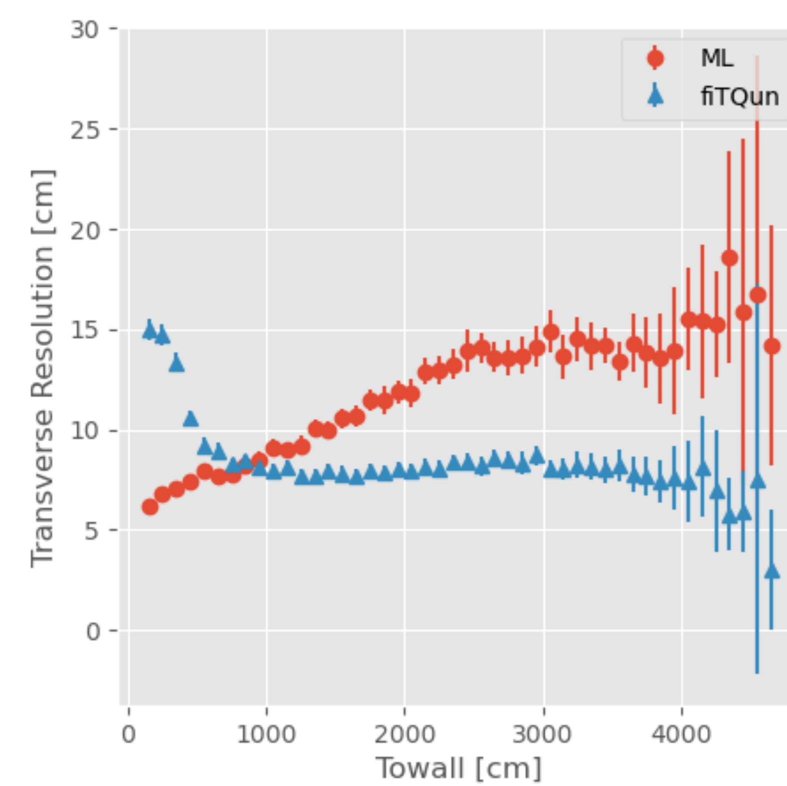




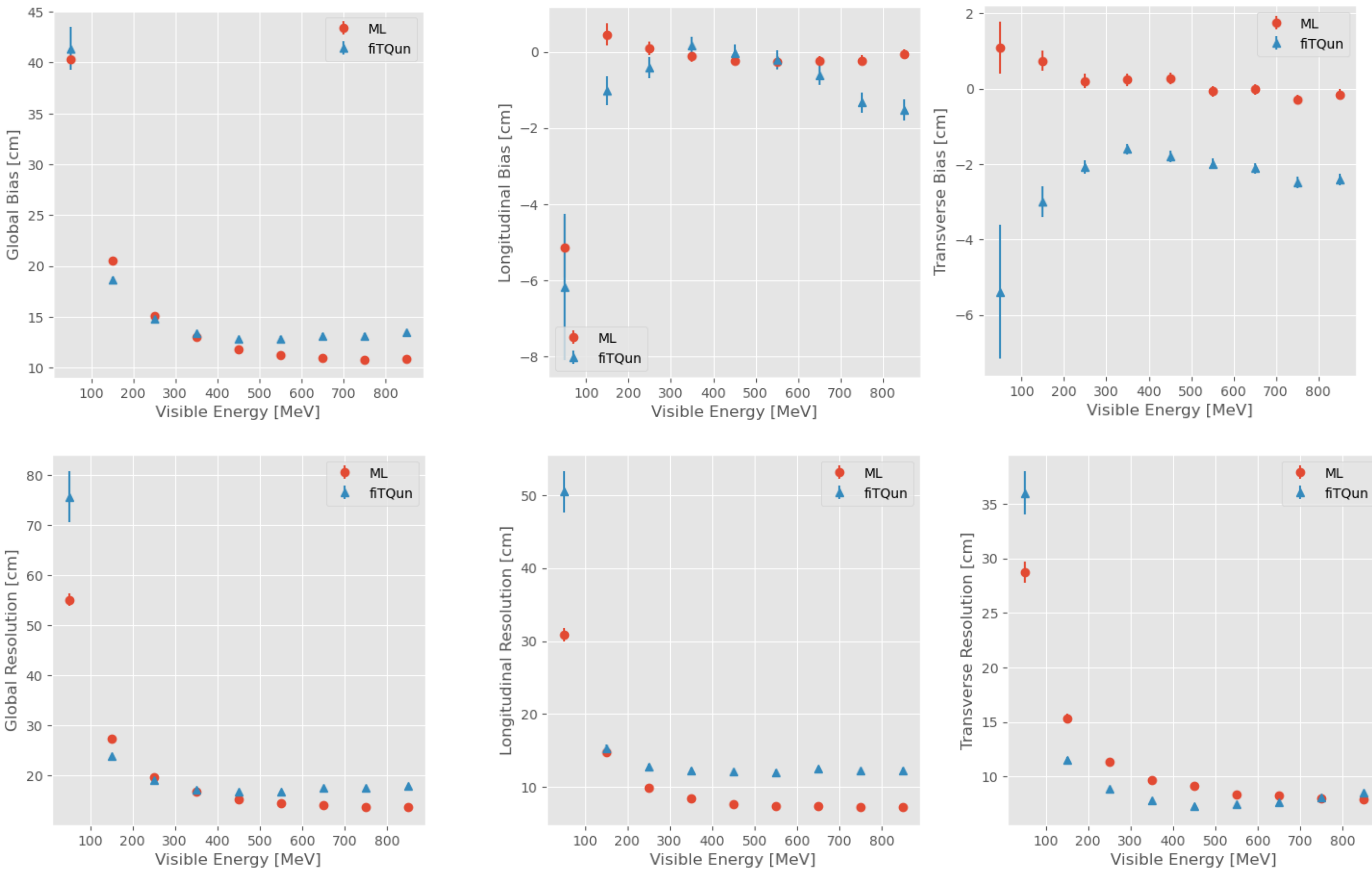








Muon - position - visible energy



Muon - direction