High-energy event reconstruction at IWCD: Machine Learning

Annalisa De Lorenzis - PhD student

QUANTUM · TECH

14 June 2023

Introduction

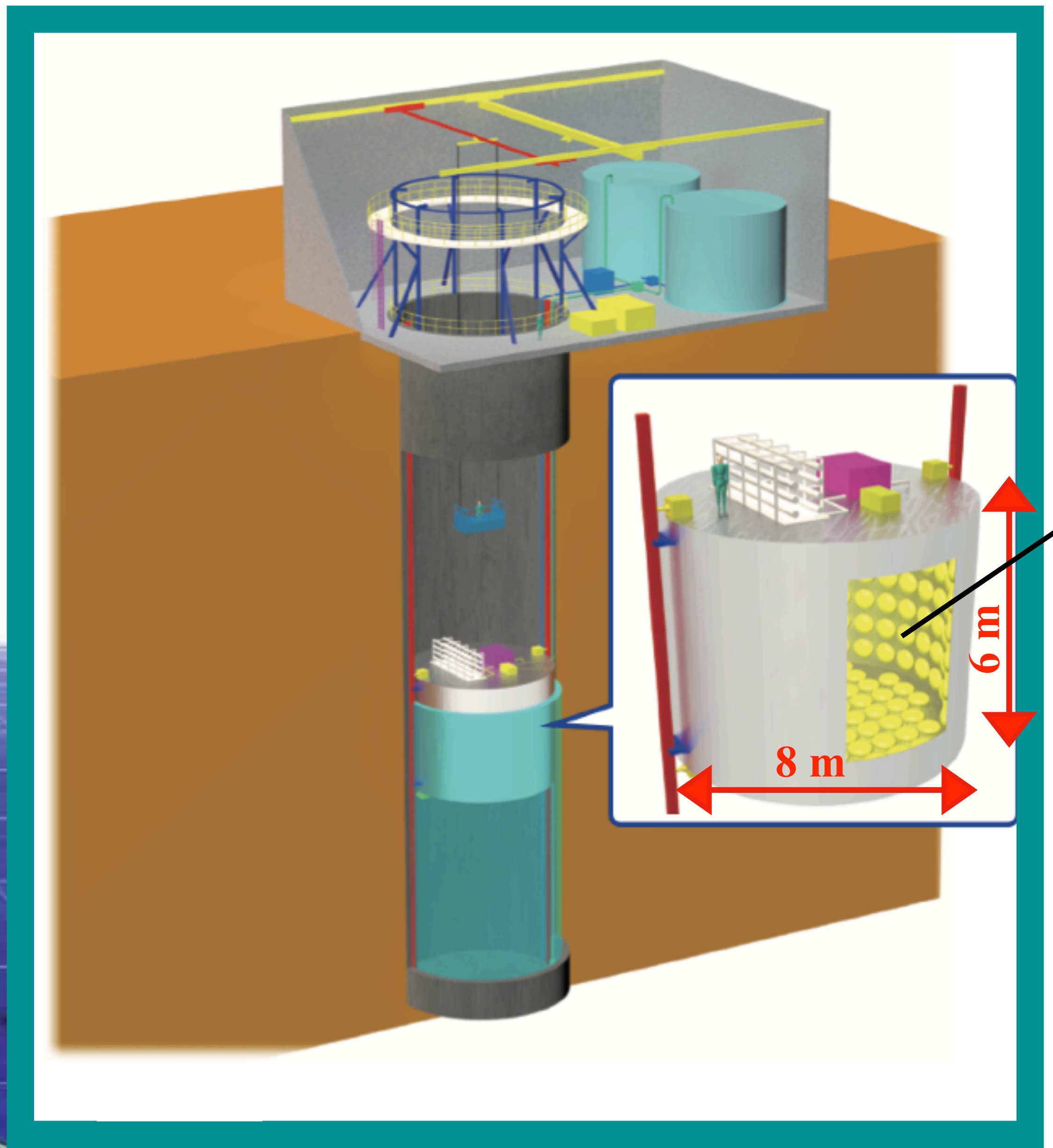
- The Intermediate Water Cherenkov Detector (IWCD)
- Motivations for Machine Learning
- Machine Learning tools and architectures
- Applications in IWCD

The Intermediate Water Cherenkov Detector

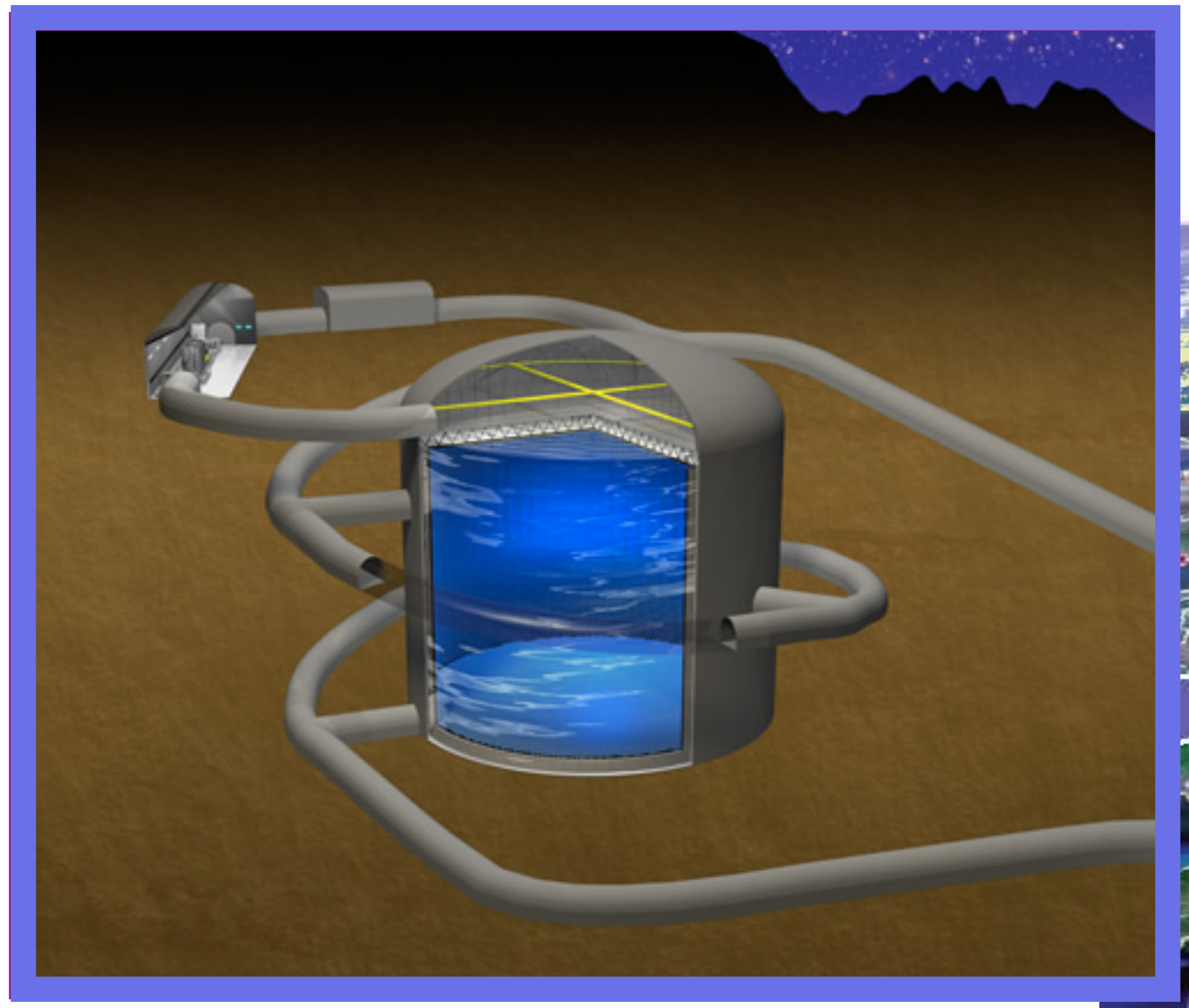
- DATA collected by IWCD:

~ 500 mPMT × 19 PMT × 2 (charge and time)

IWCD

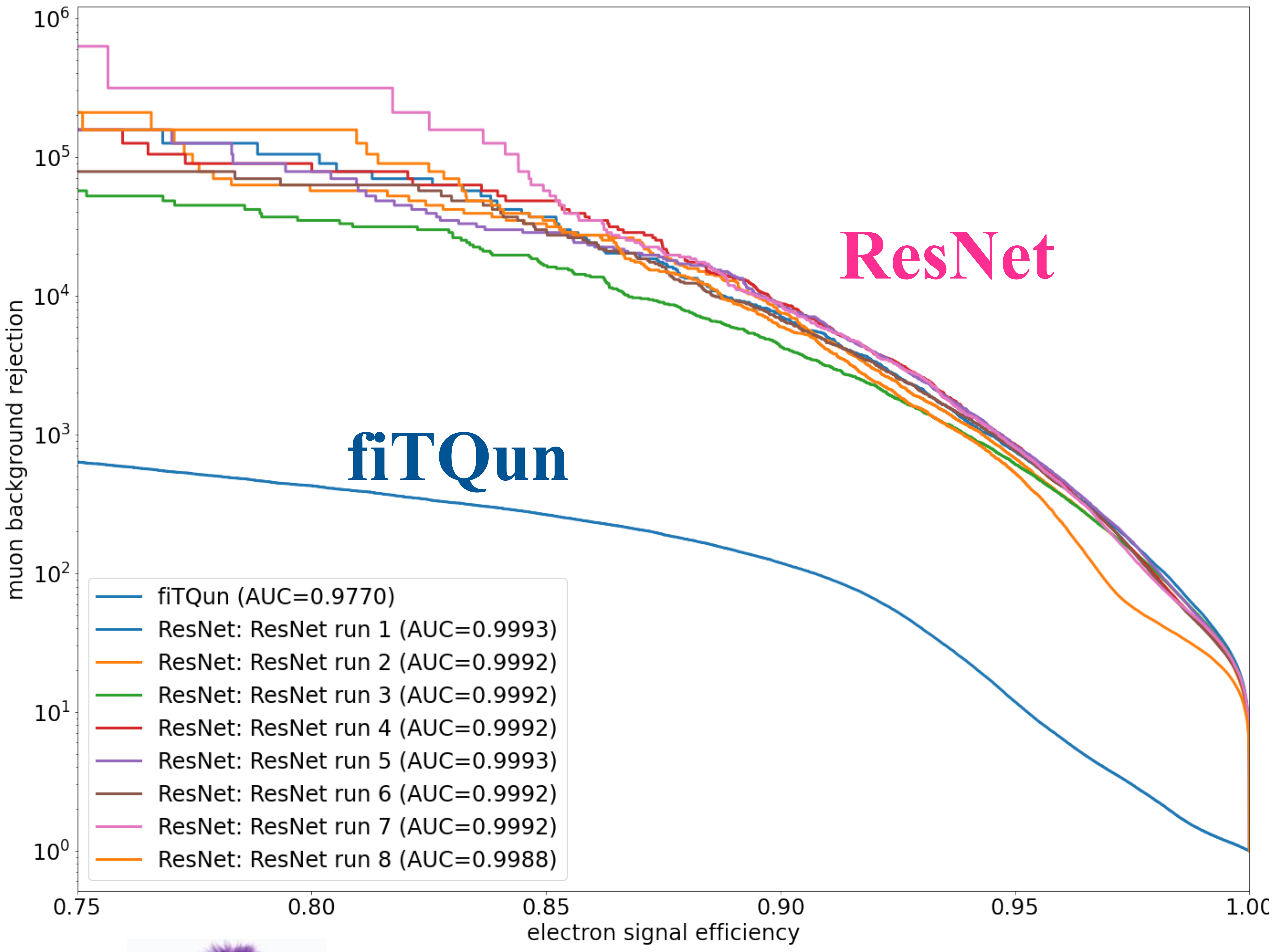


1 mPMT (= 19 PMT)



HYPER-KAMIOKANDE

Motivations for Machine Learning

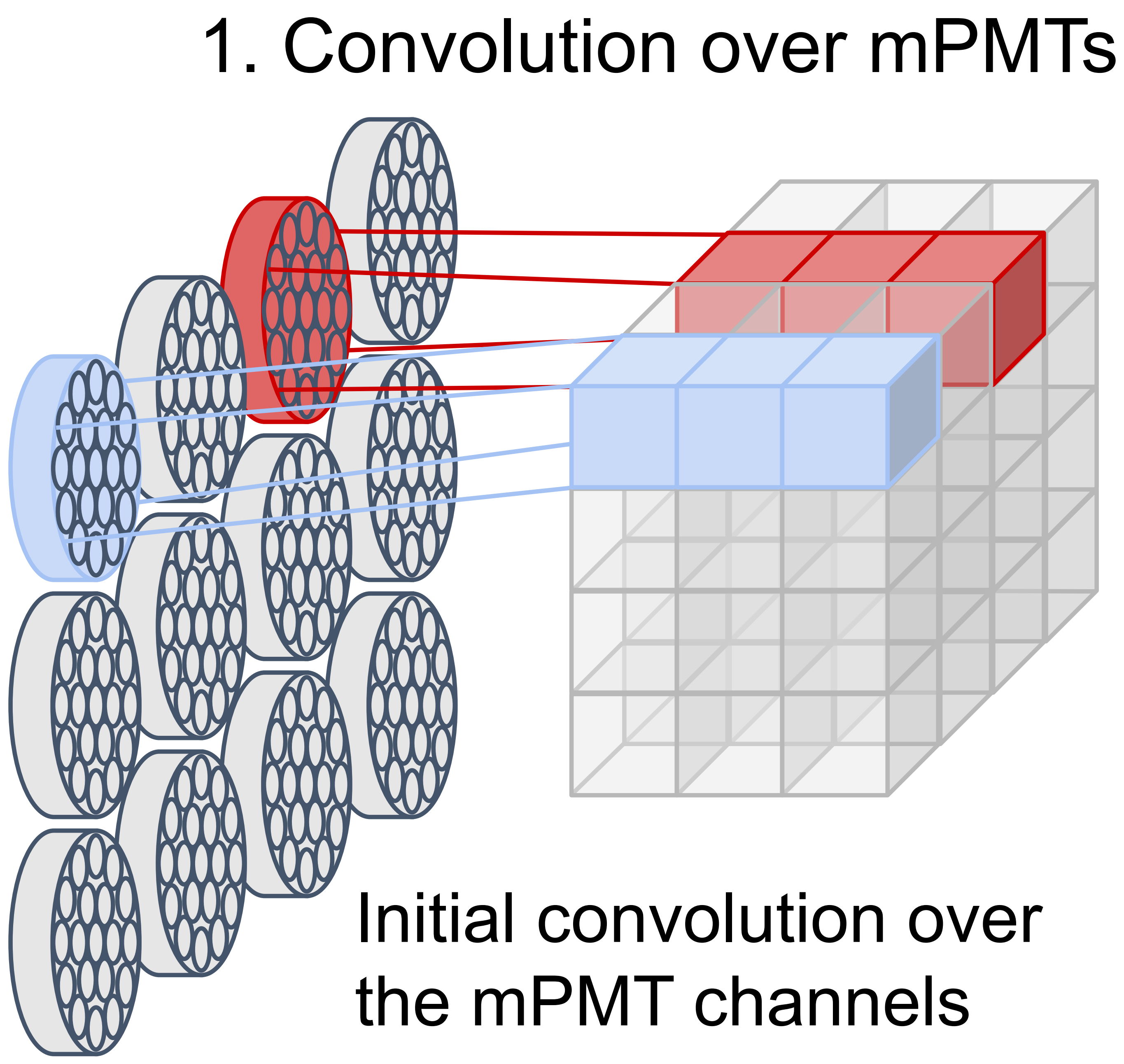


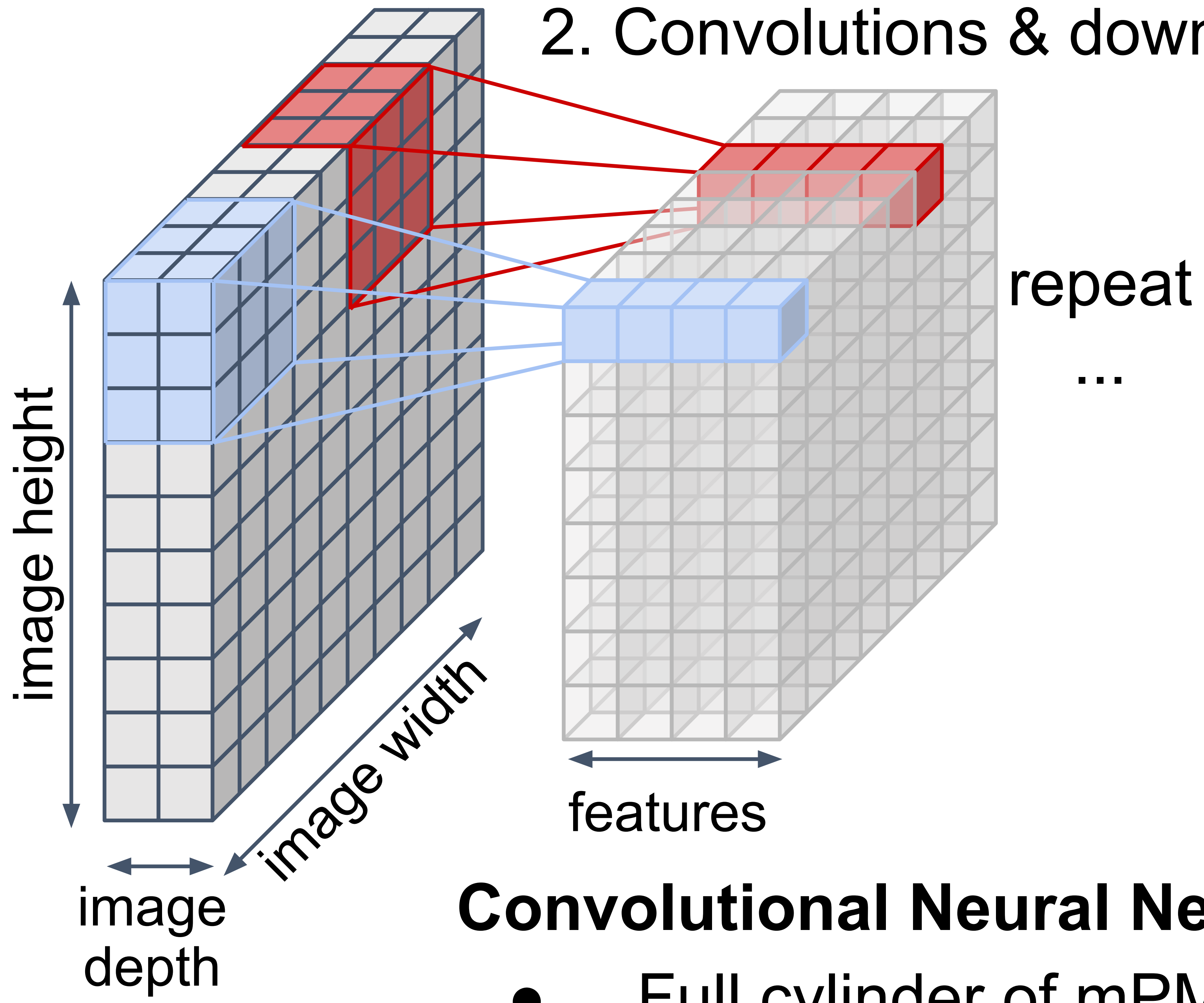
- Very successful in areas of computer vision and image processing
- Improves *particle identification* and *kinematic resolution*
- Remarkable improvements in *speed (very fast to run once neural networks have been trained)*:

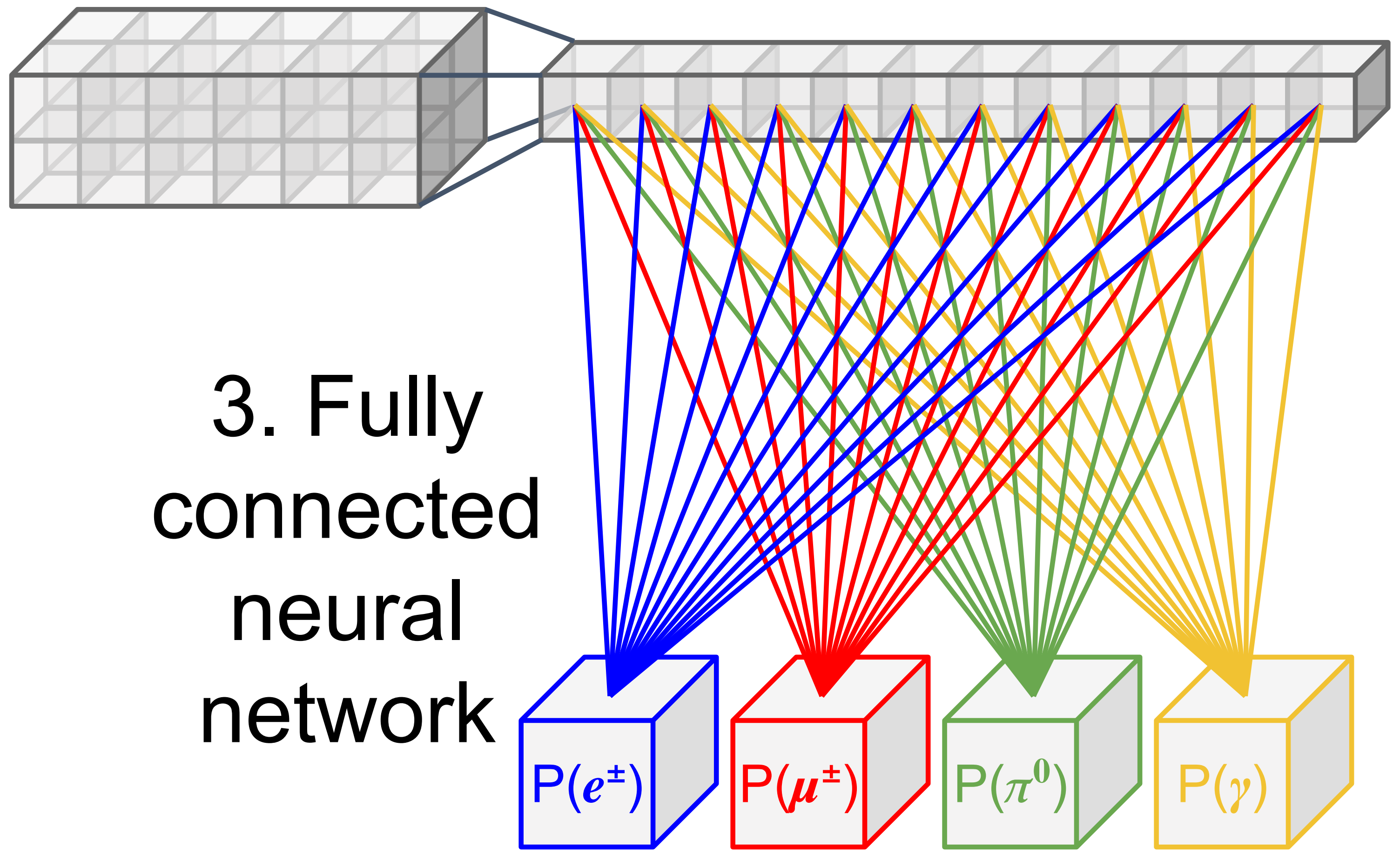
fitQun on CPU: ~ 1 event per minute
ML reconstruction on GPU: 100 000 events per minute

WatChMaL.org : an international working group to develop ML for WC detectors

Deep network architectures for IWCD

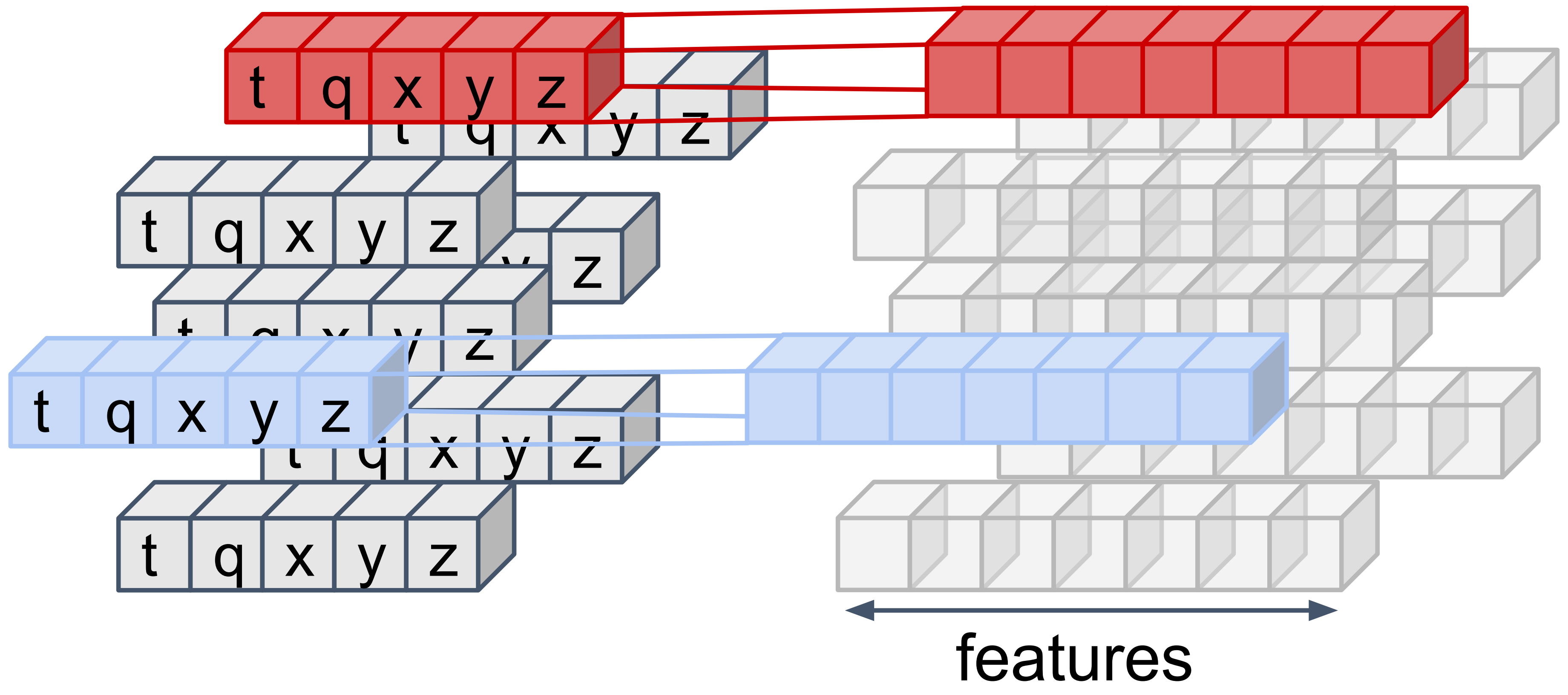






Convolutional Neural Network based on ResNet-18

- Full cylinder of mPMTs is unwrapped onto flat image
- One pixel per multi-PMT



PointNet MLP (convolution over point cloud features)

Point Cloud Neural Network based on PointNet

- Applies to point-cloud of PMT hits in 3D space
- Uses 1x1 convolutions and learns transformations applied to points

**Adapted from N. Prouse's presentation: "NuFact 2022"*

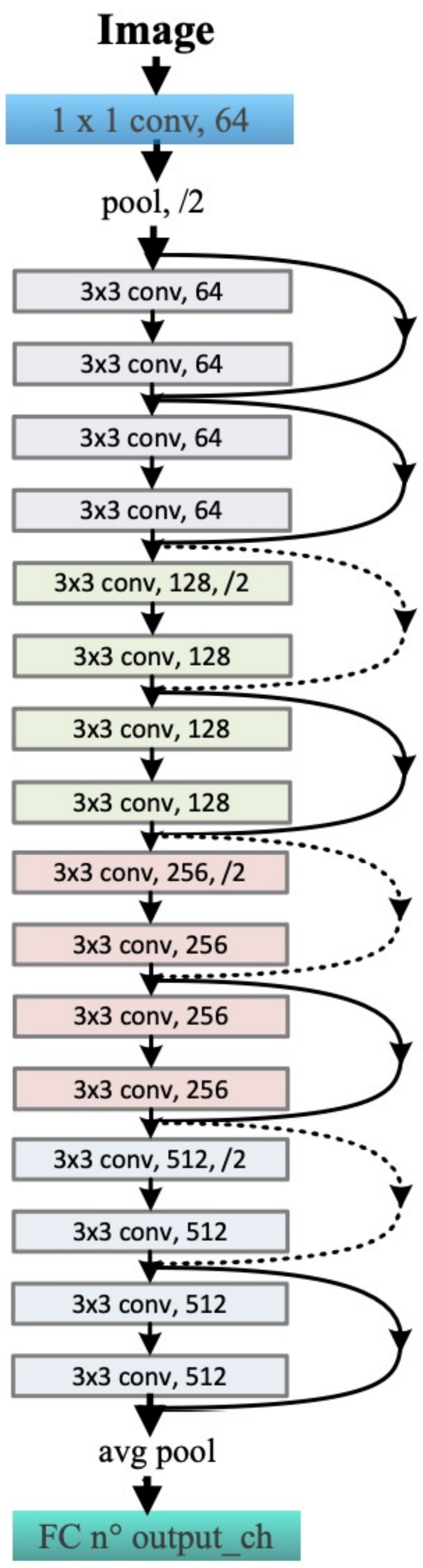
Residual CNN: ResNet18

- Residual CNNs contain Residual blocks, they implement *skip connections*

- Shortcut/skip connections prevent training problems of deep NNs (vanishing and exploding gradient problems)

- For this reason, Residual CNNs have had enormous success on computer vision tasks

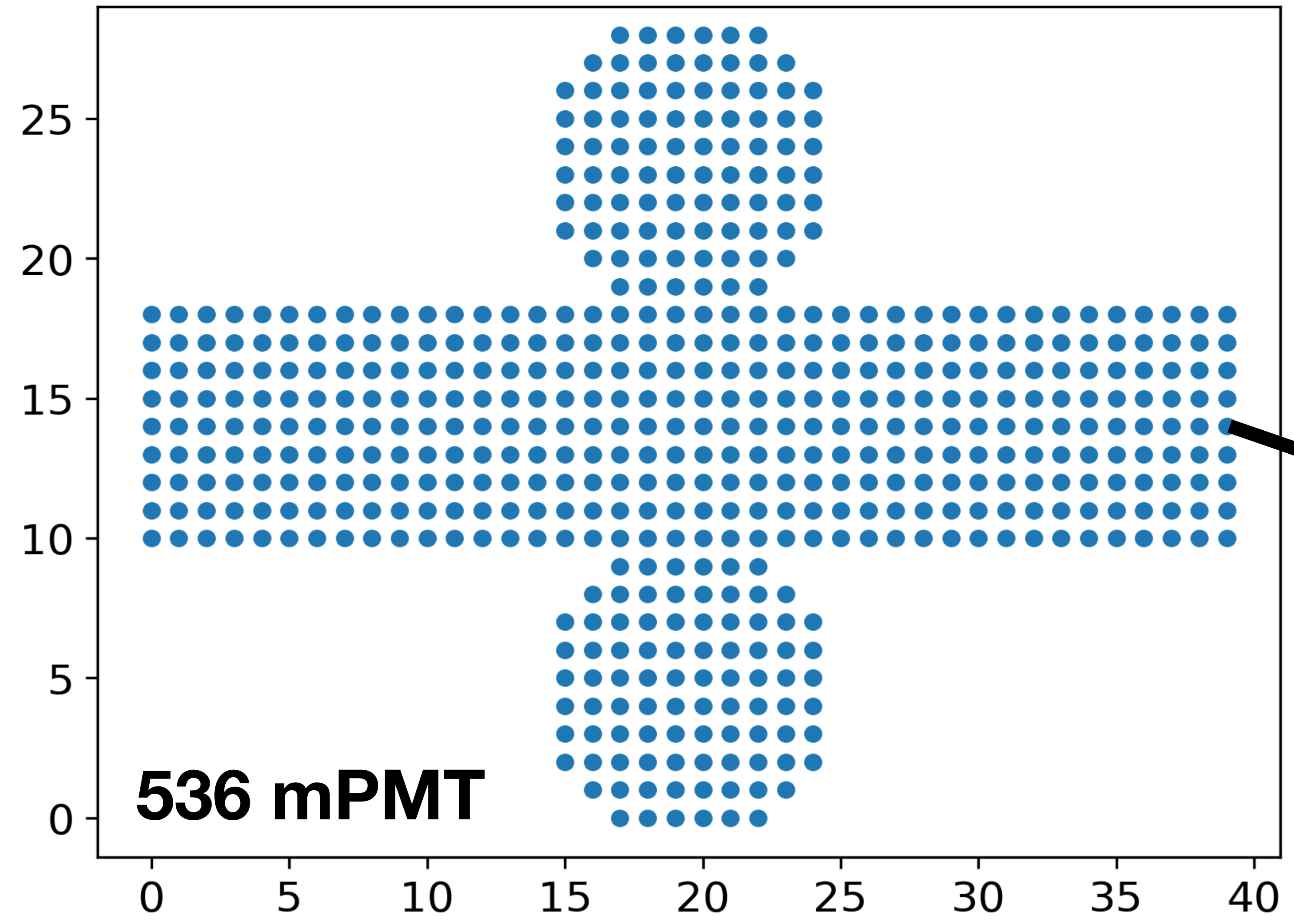
- Used an adapted version of ResNet-18



**Adapted from [arXiv:1512.03385]*

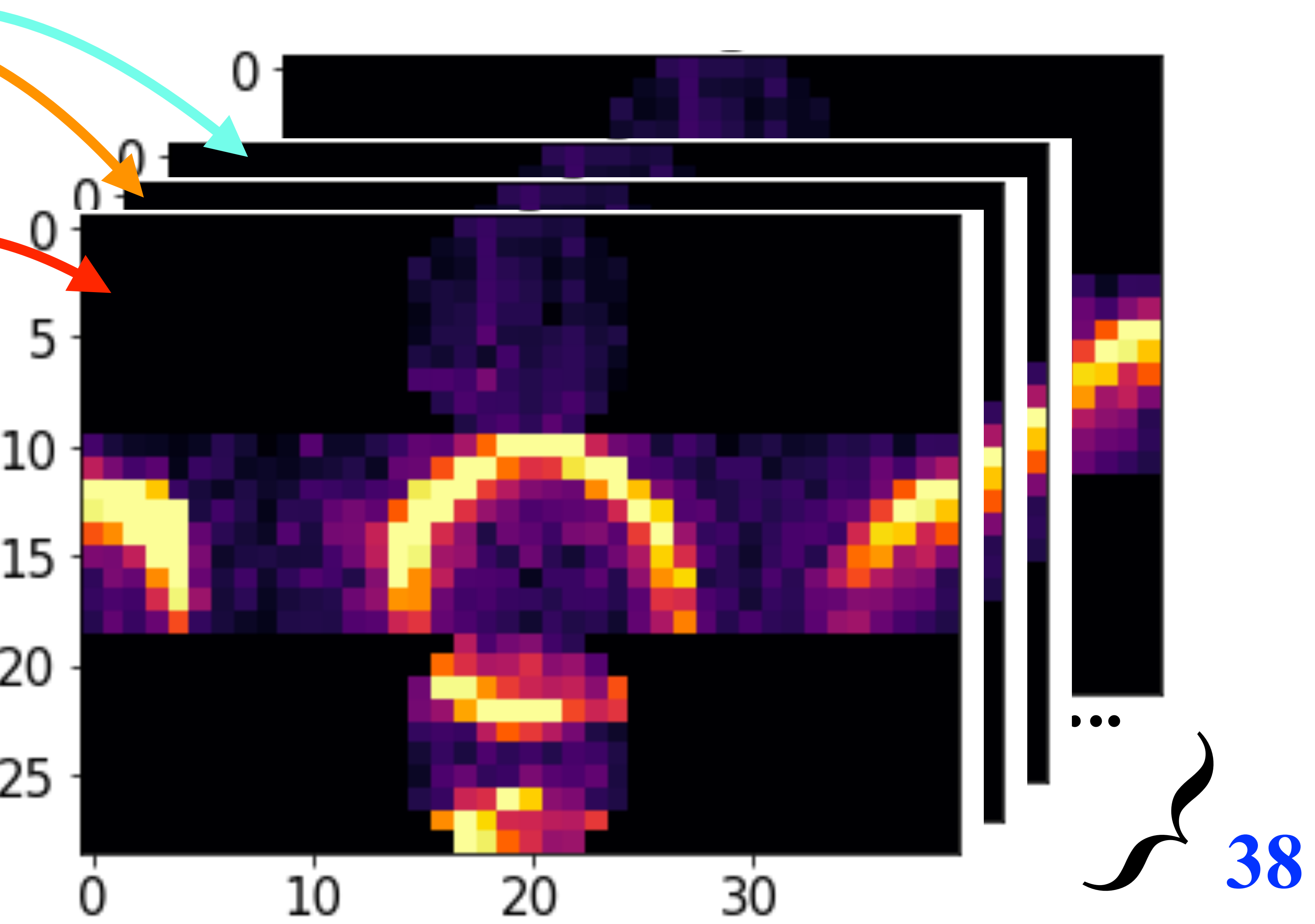
Input image for ResNet-18

“Unrolling” the tank



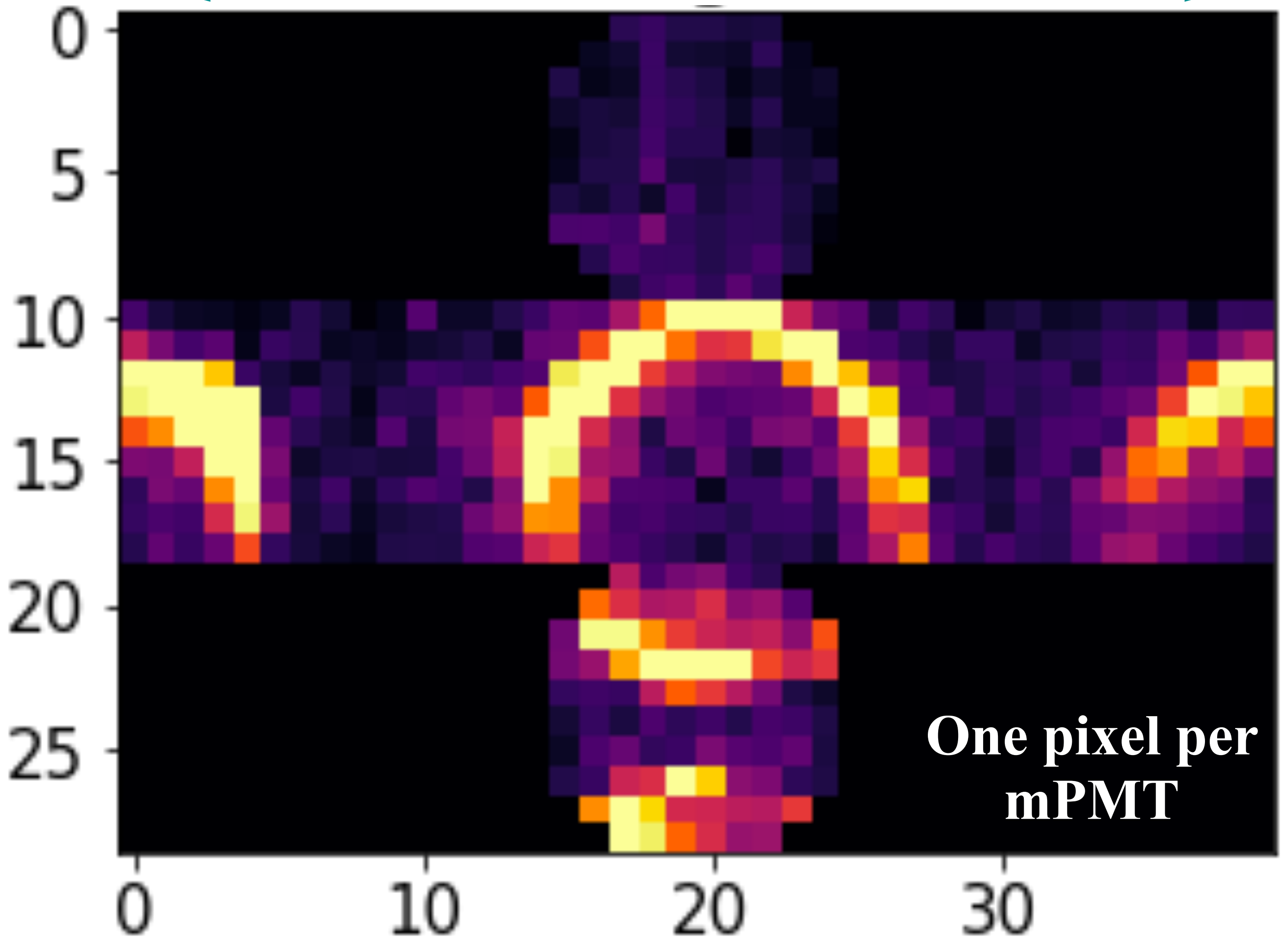
40 pixels

1 mPMT (= 19 PMT)



= 19 channels (charge) 19 Q

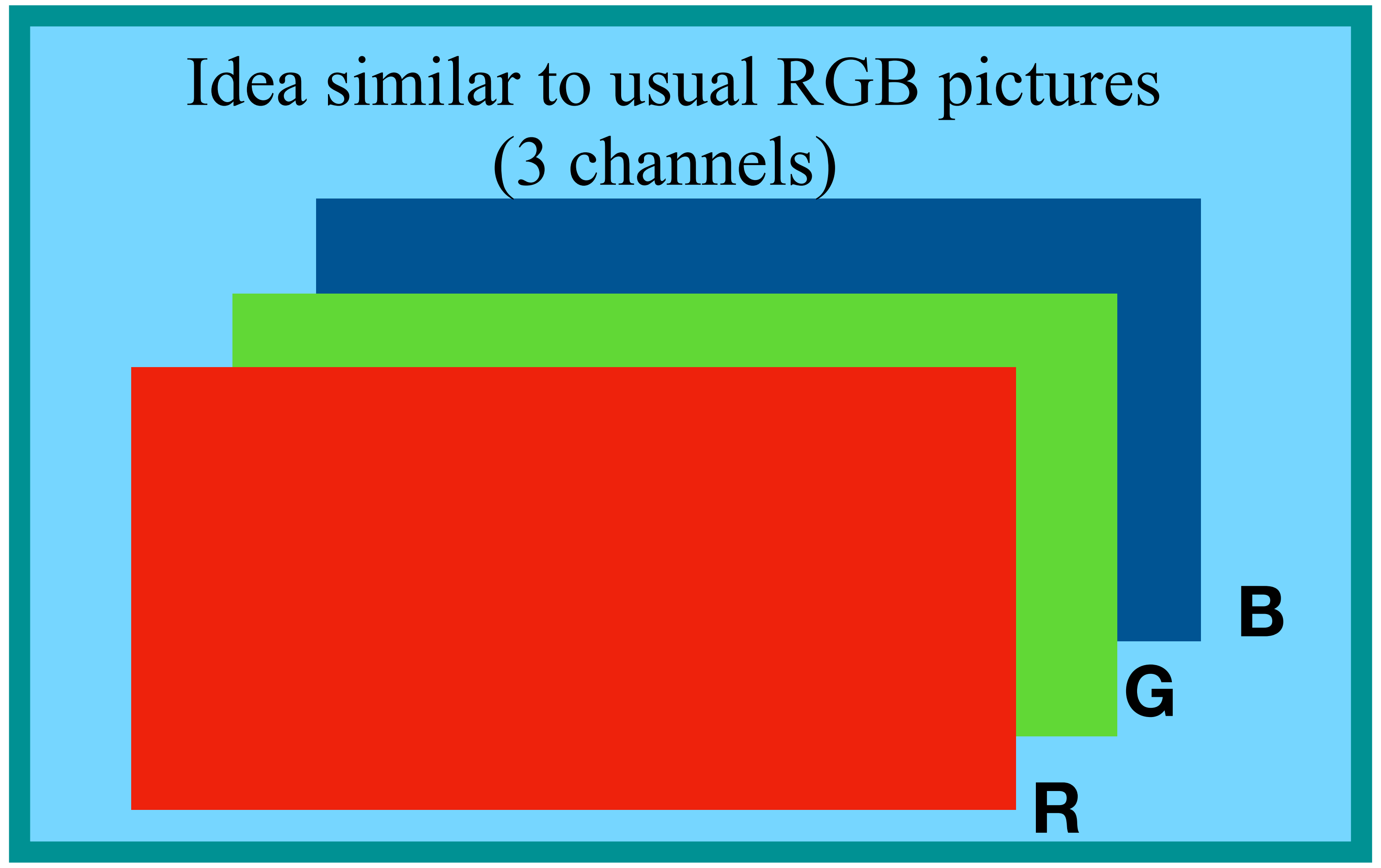
+ 19 channels (time) 19 T



29 pixels

https://github.com/adelorenzis/WatChMaL

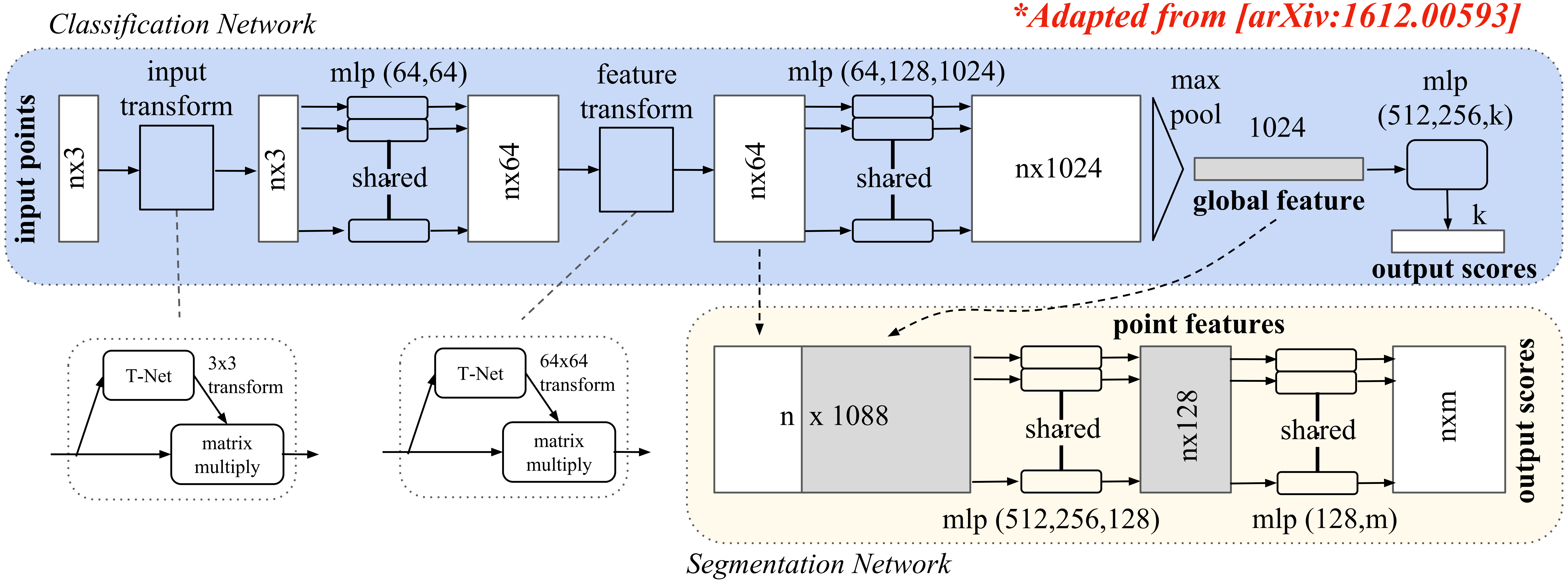
(Added time implementation)



Point Cloud NN: PointNet

- PointNet works on a set of points (*cloud*) rather than on the grid of fixed pixels of an image
- Each point of the cloud represents a hit PMT and has 5 "coordinates": time, charge, 3D position in space (t, q, x, y, z)

**Adapted from [arXiv:1612.00593]*



- The NN exploits the following properties of the input data:
  - points are unordered (not fixed on a grid)
  - there are relations among close points
  - clouds are invariant under transformations (e.g. rotations and translations)

REVIEW: ML Applications in IWCD

- Classification for Particle type Identification (PID):
  - e/μ
  - e/γ
  - e/π^0

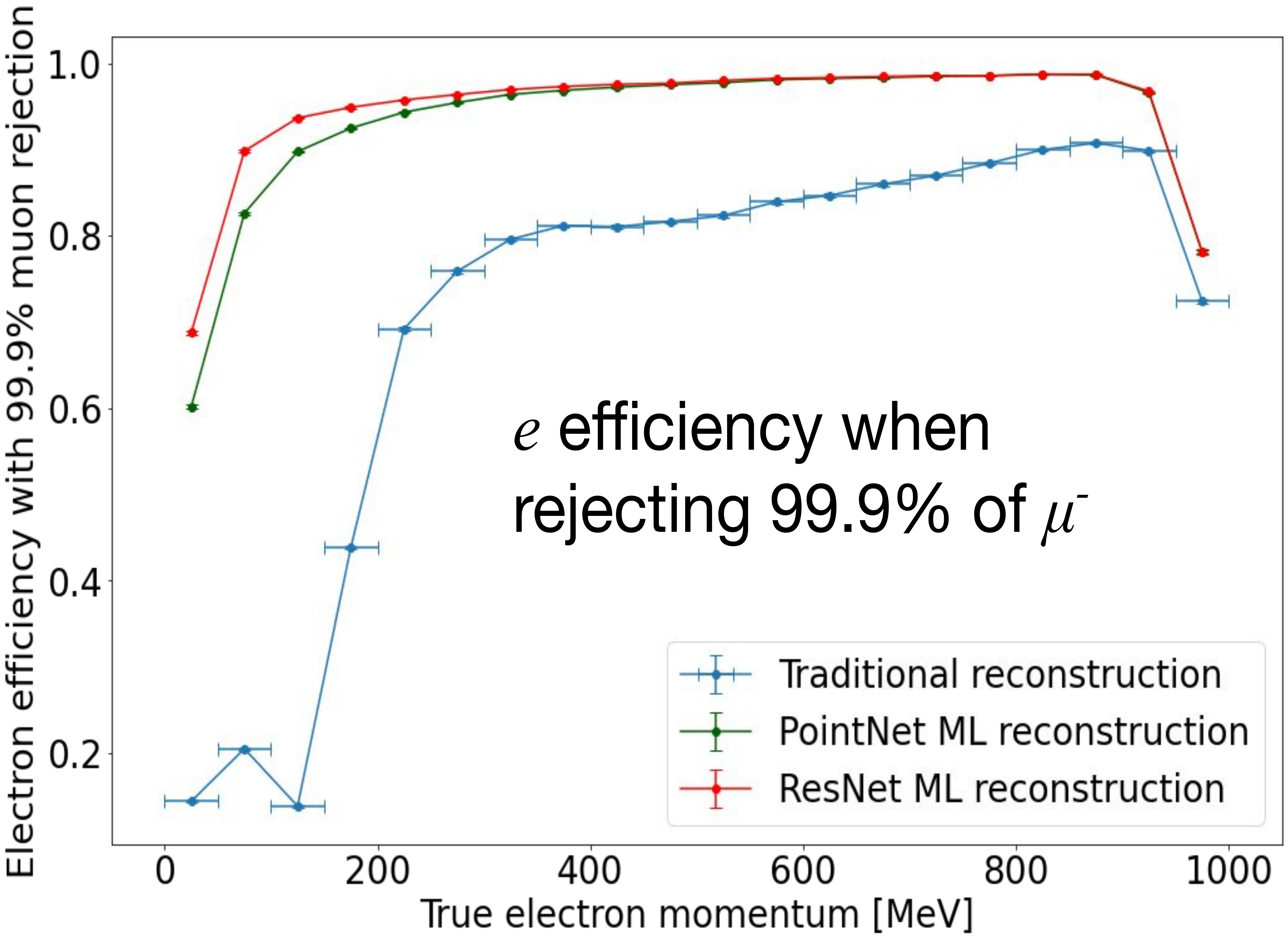
While e/μ are classified reasonably well with traditional methods, there could be large improvements for e/π^0 and e/γ with ML

- Regression: direction, position and energy reconstruction of the particles

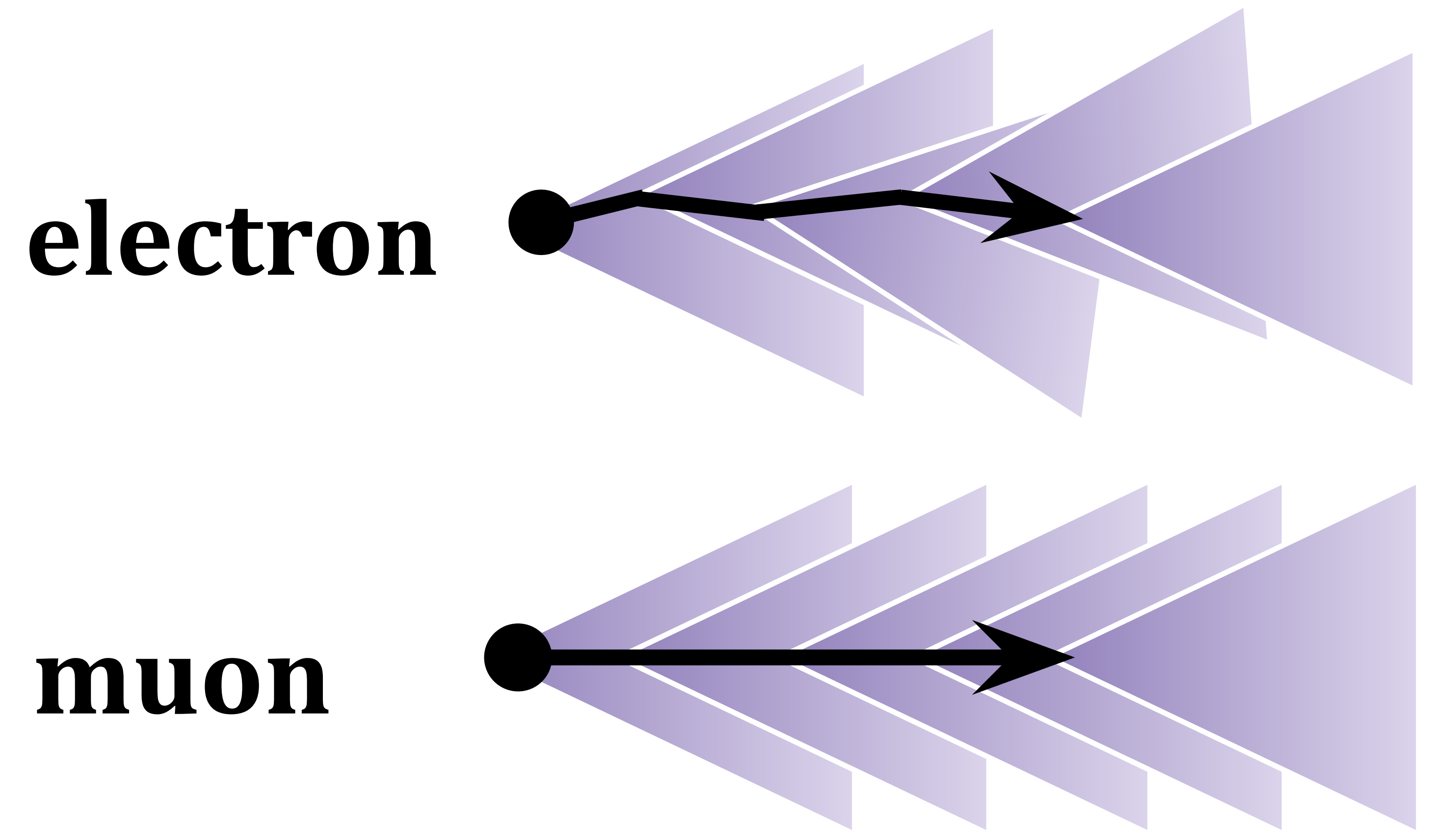
- Classification for Single-Vertex / Multi-Vertex (Pile-up)

- Exploration of alternative methodologies:
  - e/γ using Quantum Machine Learning

Classification: e/μ

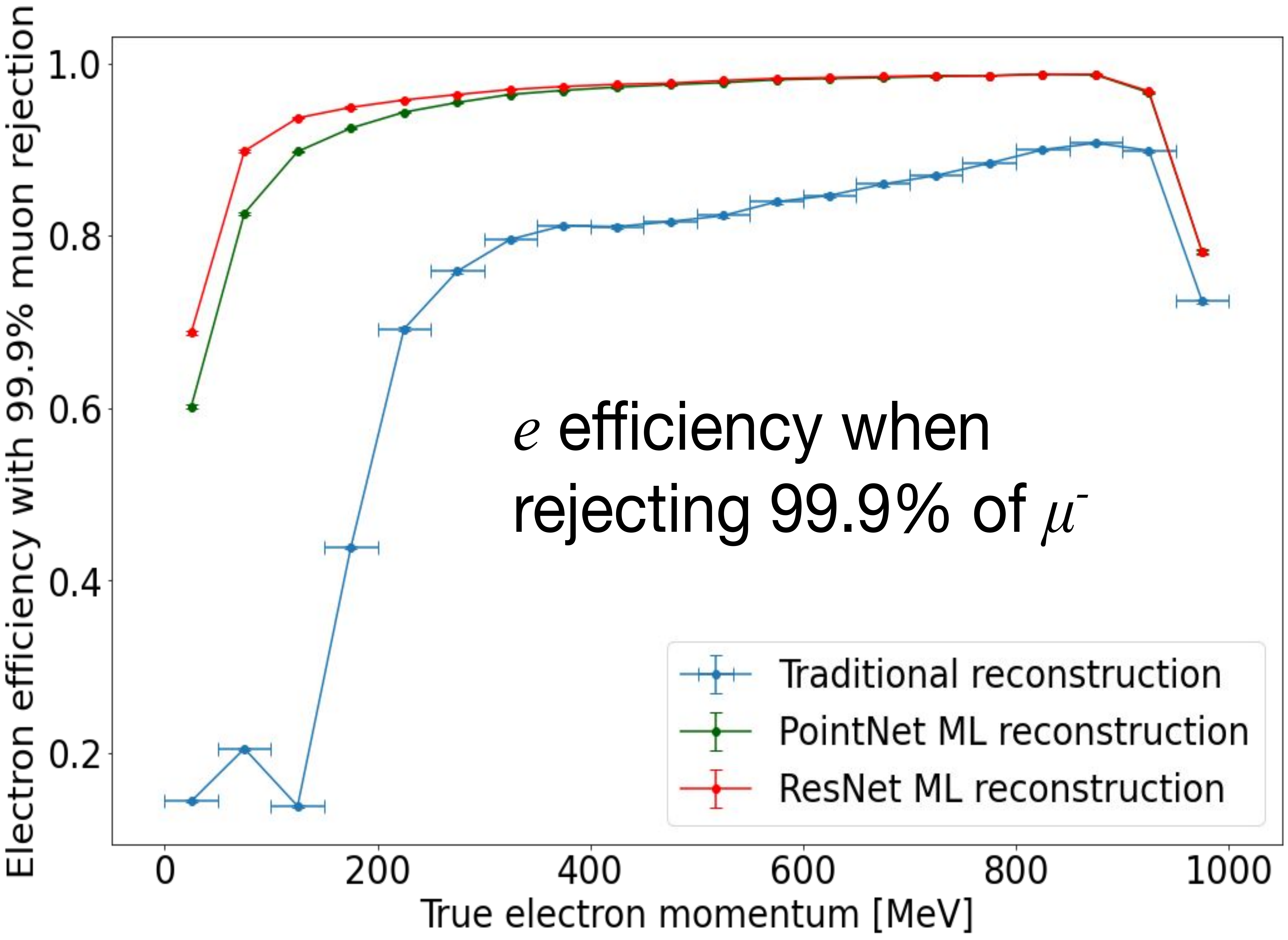


- ML techniques outperform traditional method over all energy range
- ResNet and PointNet give similar results for e/μ, ResNet slightly better for small energies

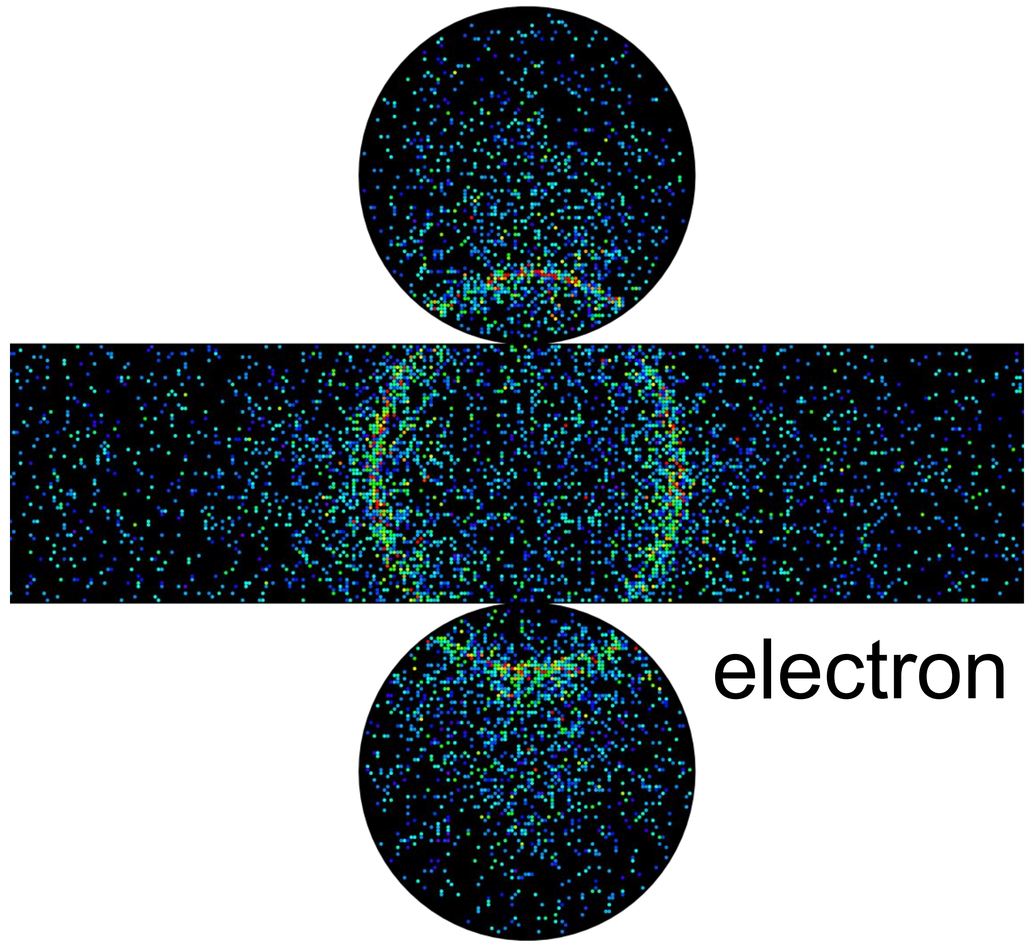


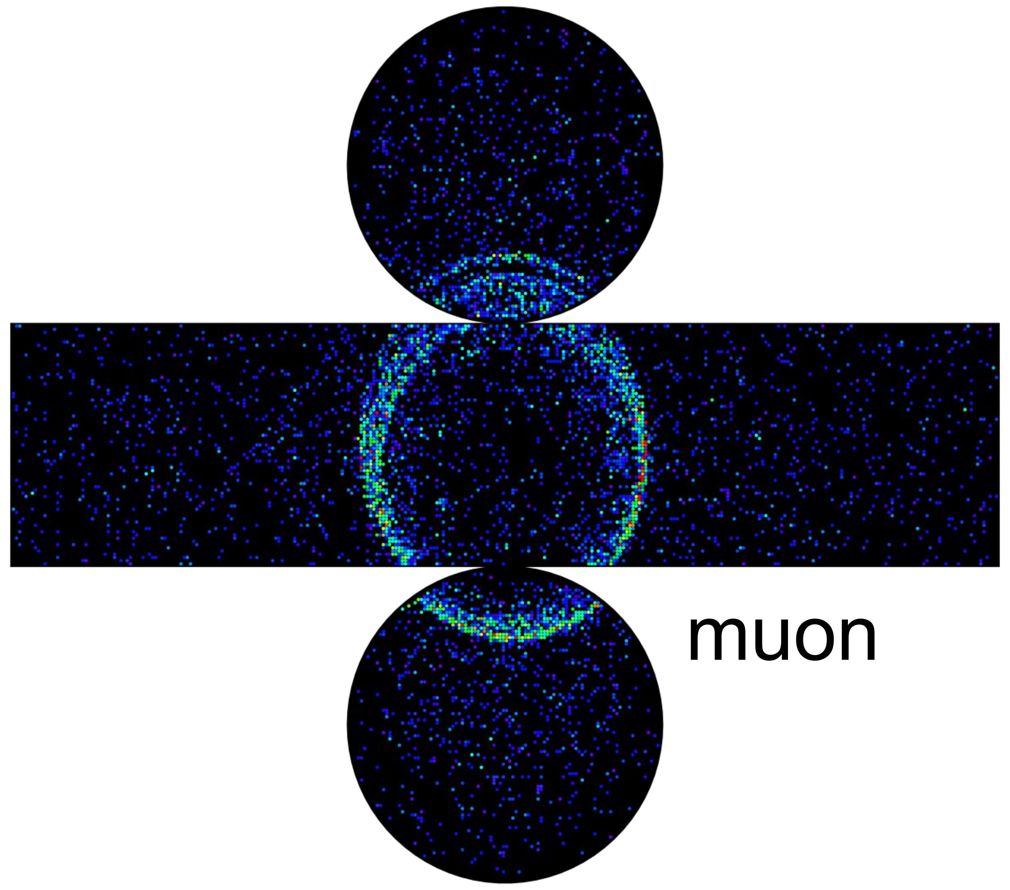
Plot from N. Prouse's presentation: "NuFact 2022"

Classification: e/μ



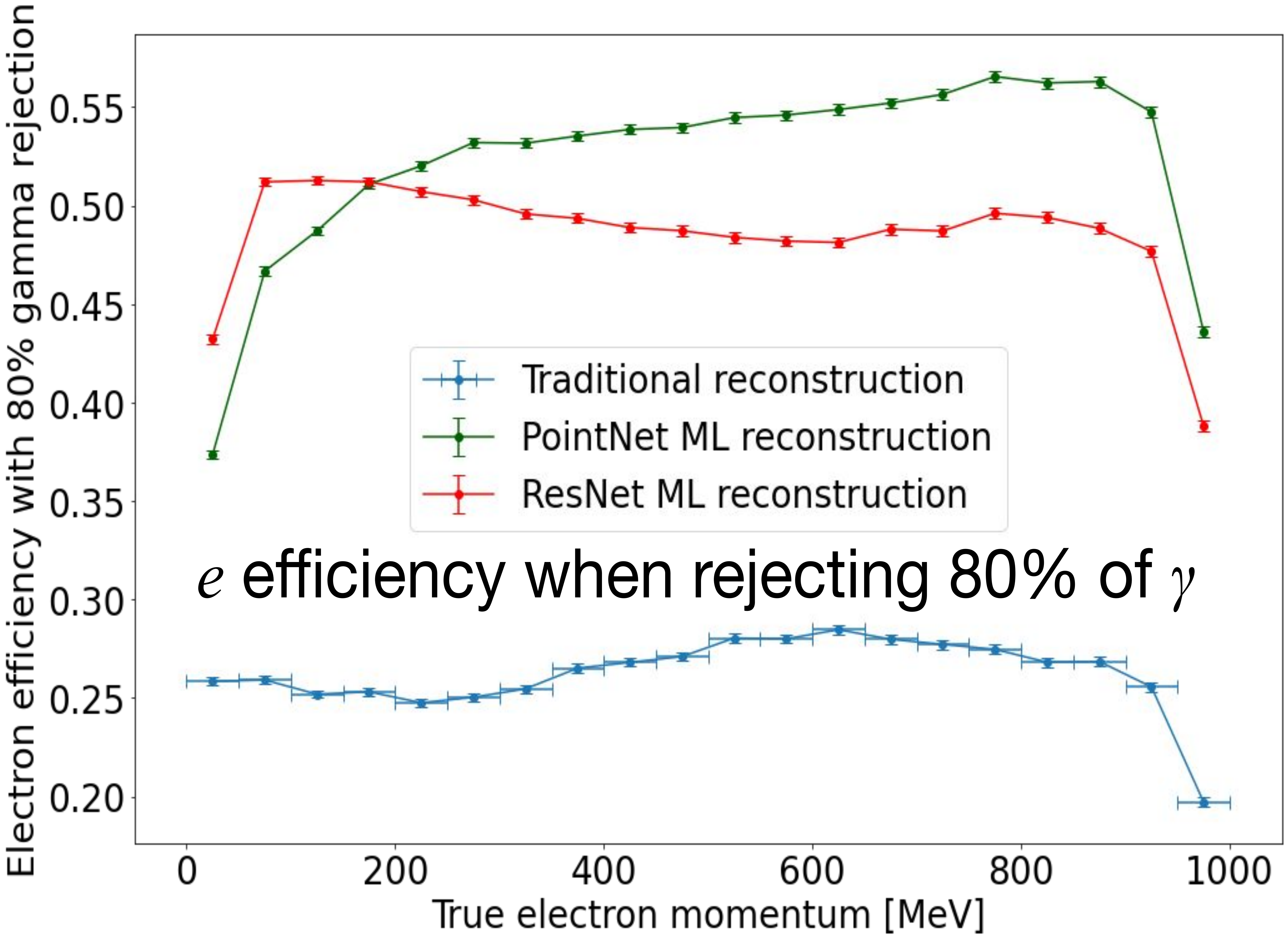
- ML techniques outperform traditional method over all energy range
- ResNet and PointNet give similar results for e/μ, ResNet slightly better for small energies



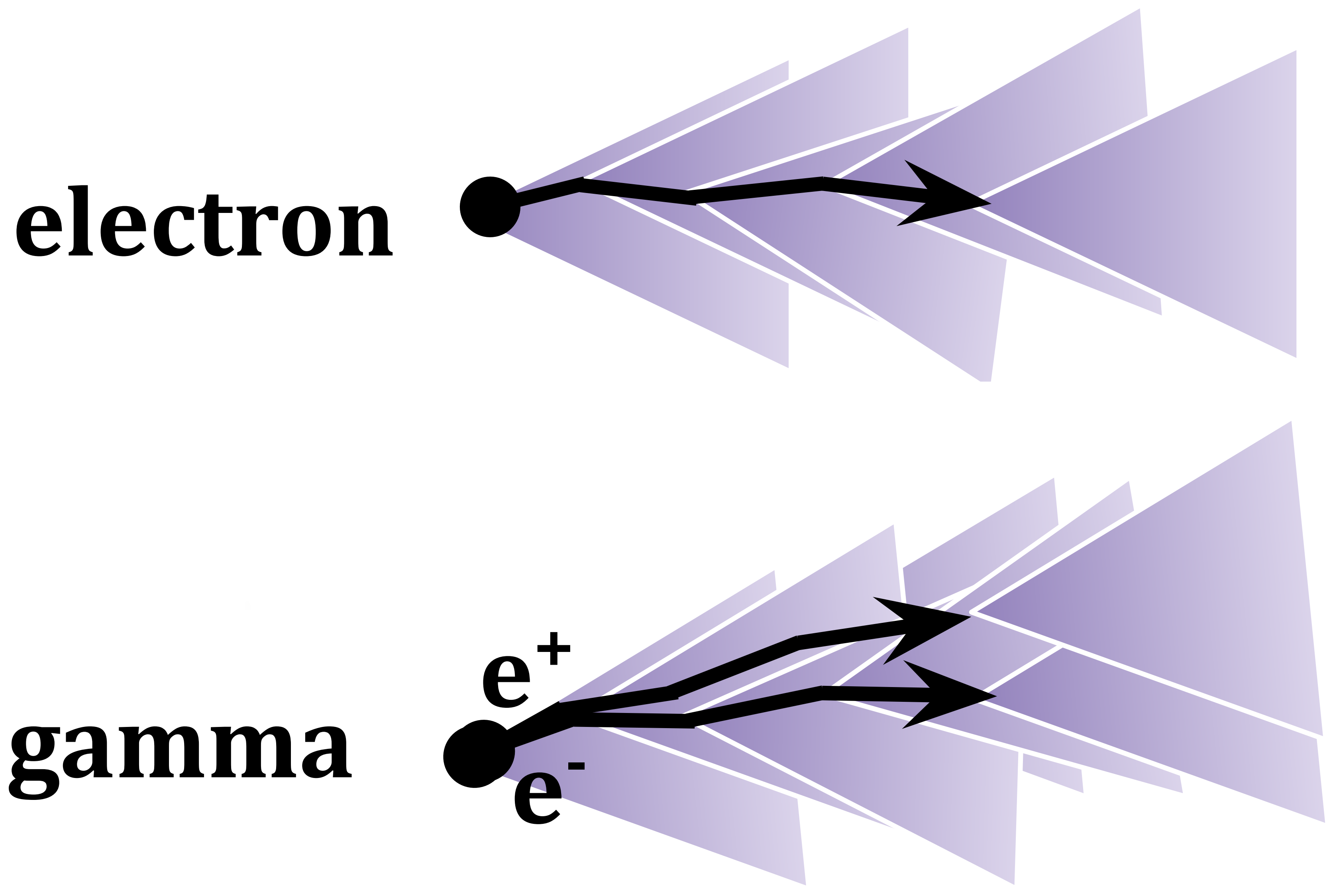


**Plot from N. Prouse's presentation: "NuFact 2022"*

Classification: e/γ

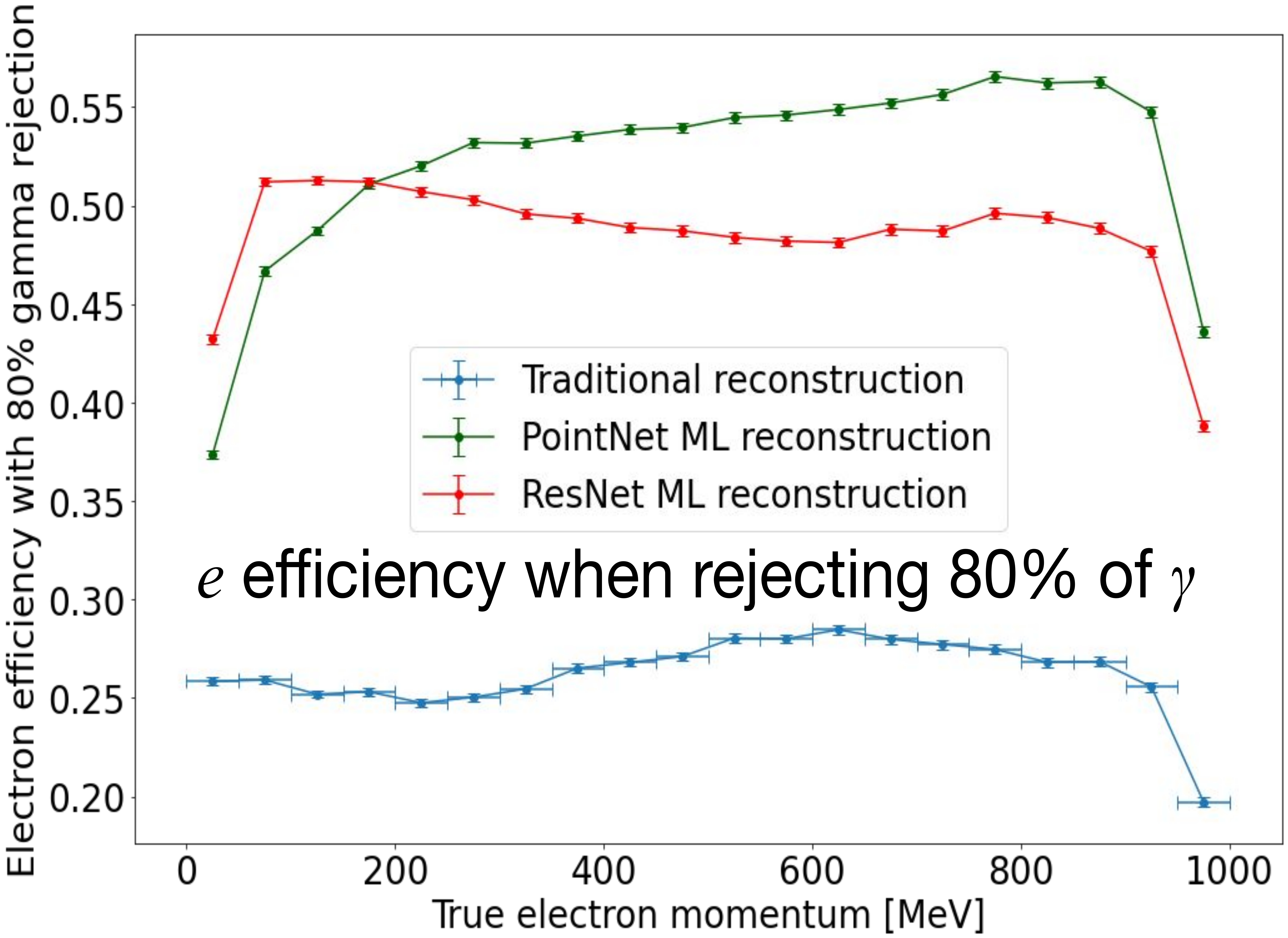


- e and γ are almost indistinguishable in Water Cherenkov detectors
- PointNet performs better than ResNet for e/γ classification, except for small energy



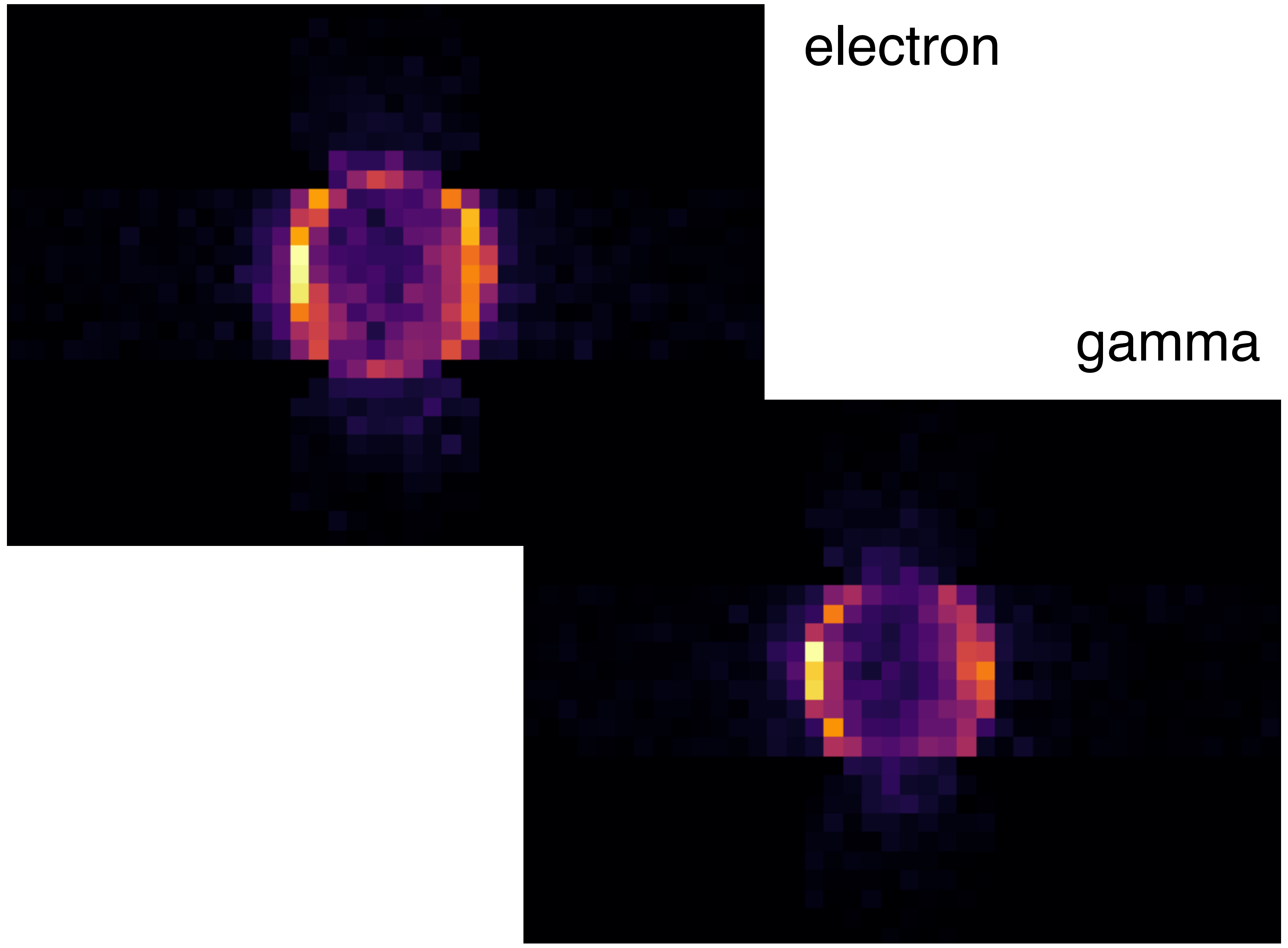
**Plot from N. Prouse's presentation: "NuFact 2022"*

Classification: e/γ



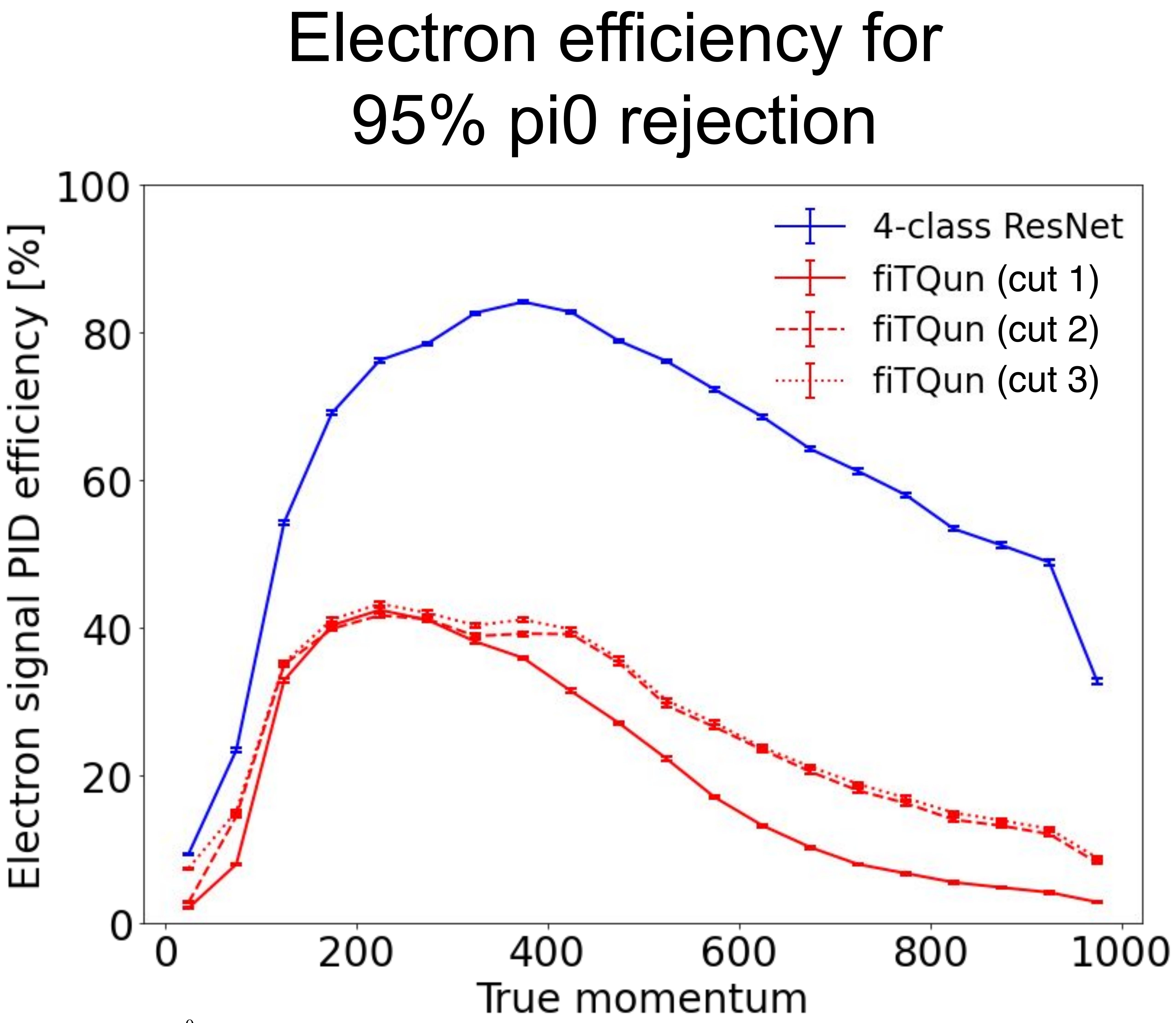
- e and γ are almost indistinguishable in Water Cherenkov detectors

- PointNet performs better than ResNet for e/γ classification, except for small energy

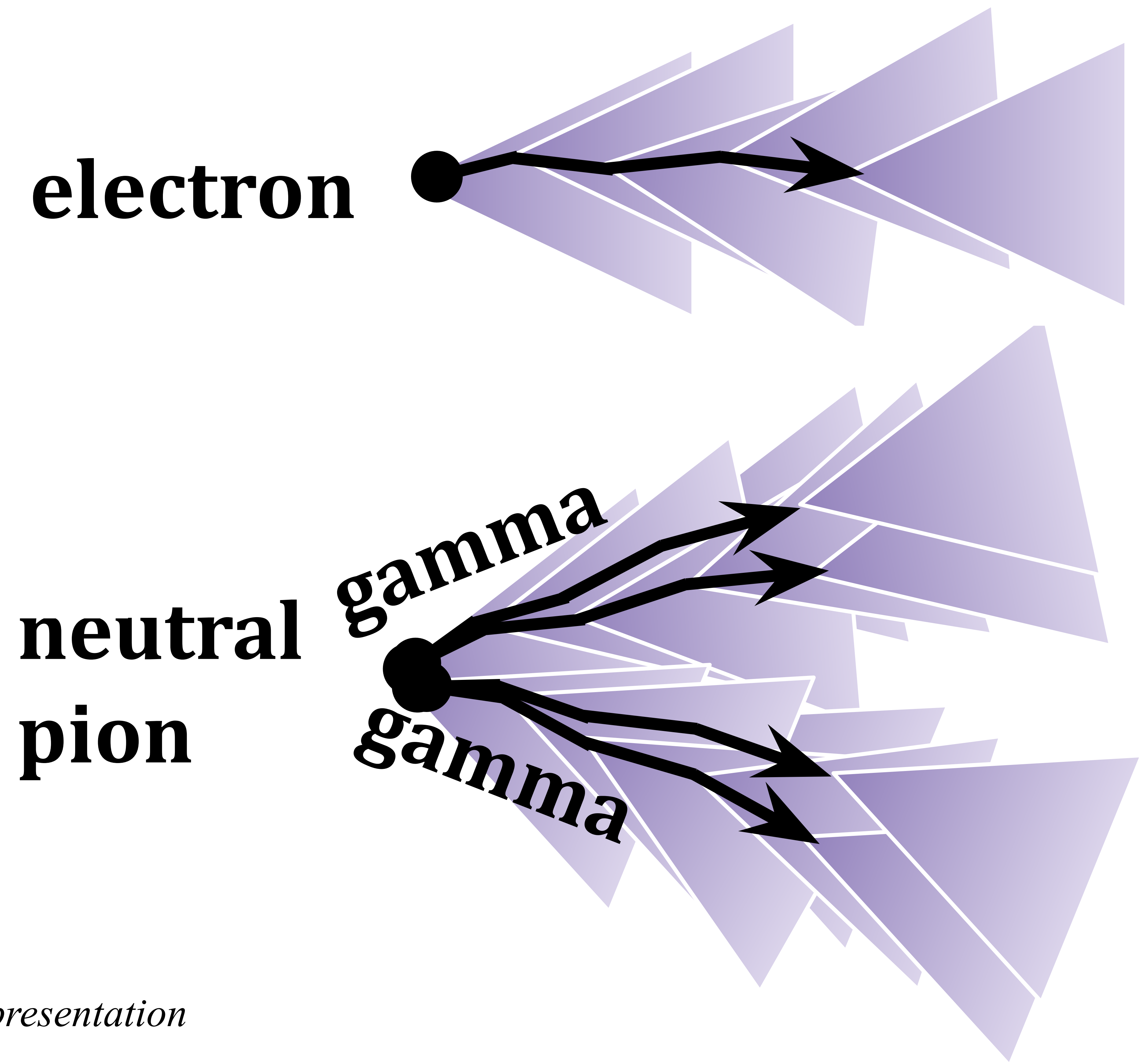


Plot from N. Prouse's presentation: "NuFact 2022"

Classification: e/π^0



- Efficiency significantly drops for higher energies since photons from π^0 become more collimated

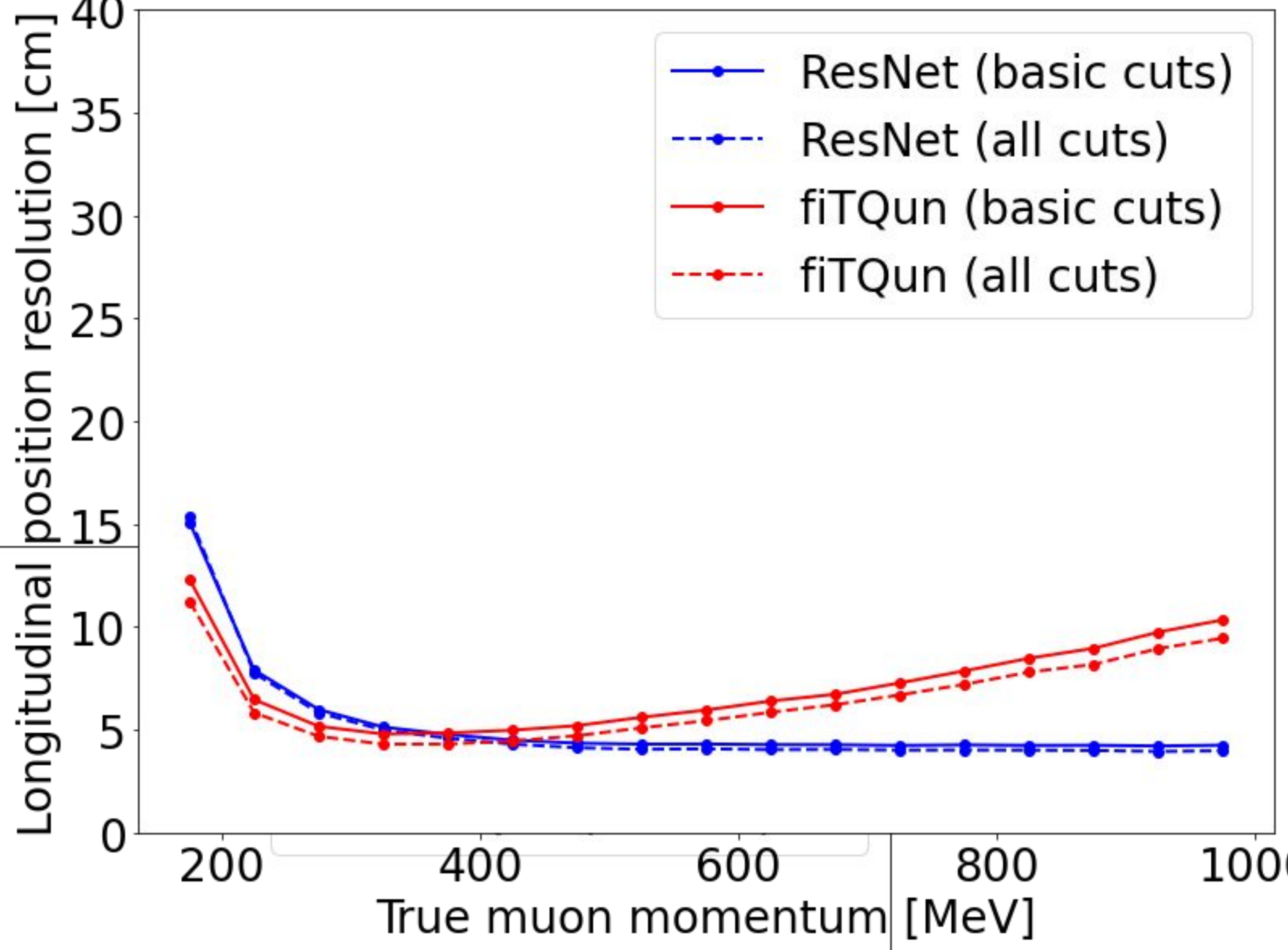


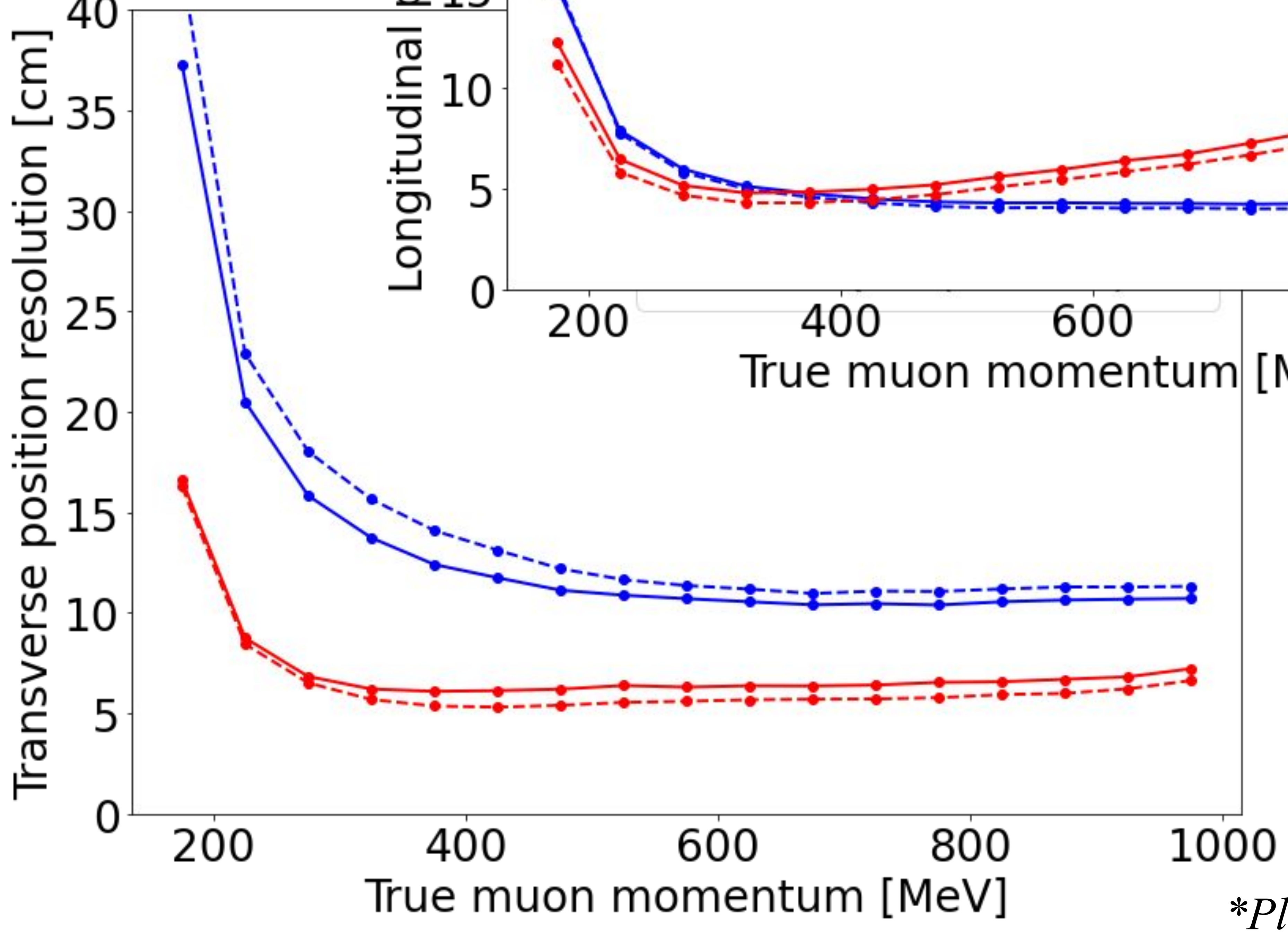
cut 1: 1D cut on e - π^0 log-likelihood difference
cut 2: 2D cut on likelihood difference and reconstructed π^0 mass
cut 3: 2D cut as above but with different cut in each reconstructed momentum bin

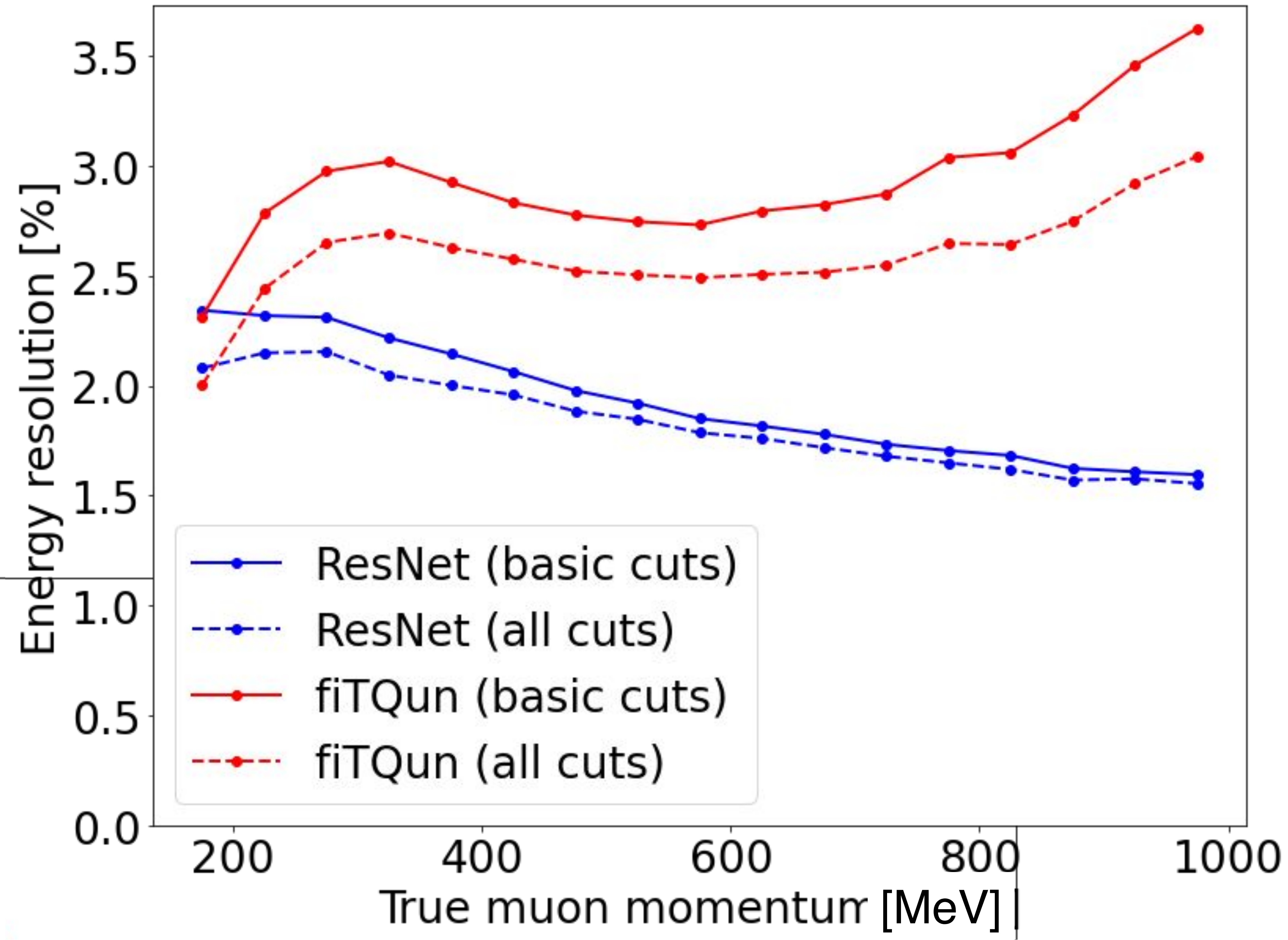
**Plot from WatChMaL meeting's presentation*

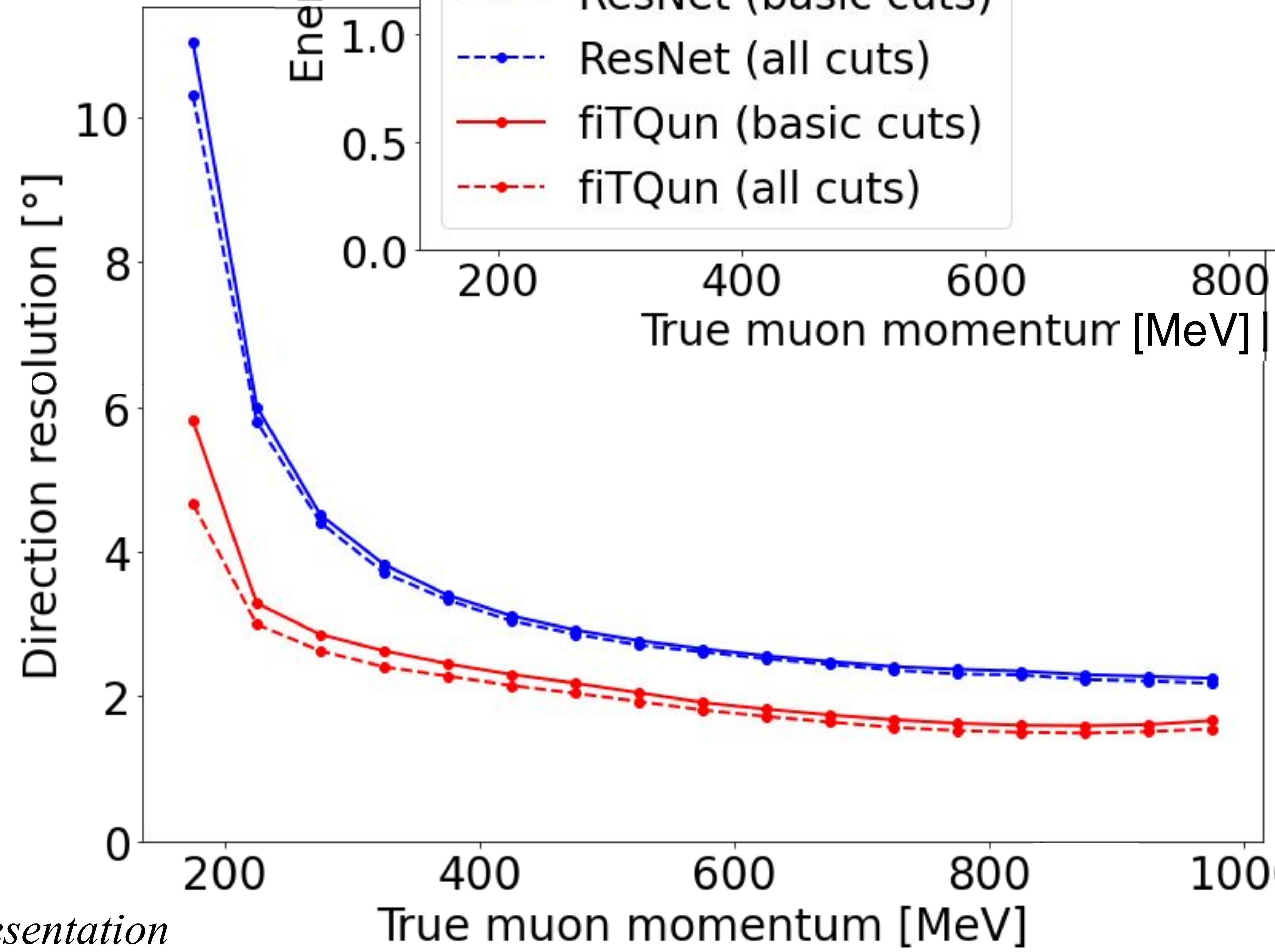
Regression: position, energy and direction reconstruction

- In addition to PID classification, the same networks can be used to reconstruct kinematic quantities of the particles generating the rings









Consider muon rings as an example

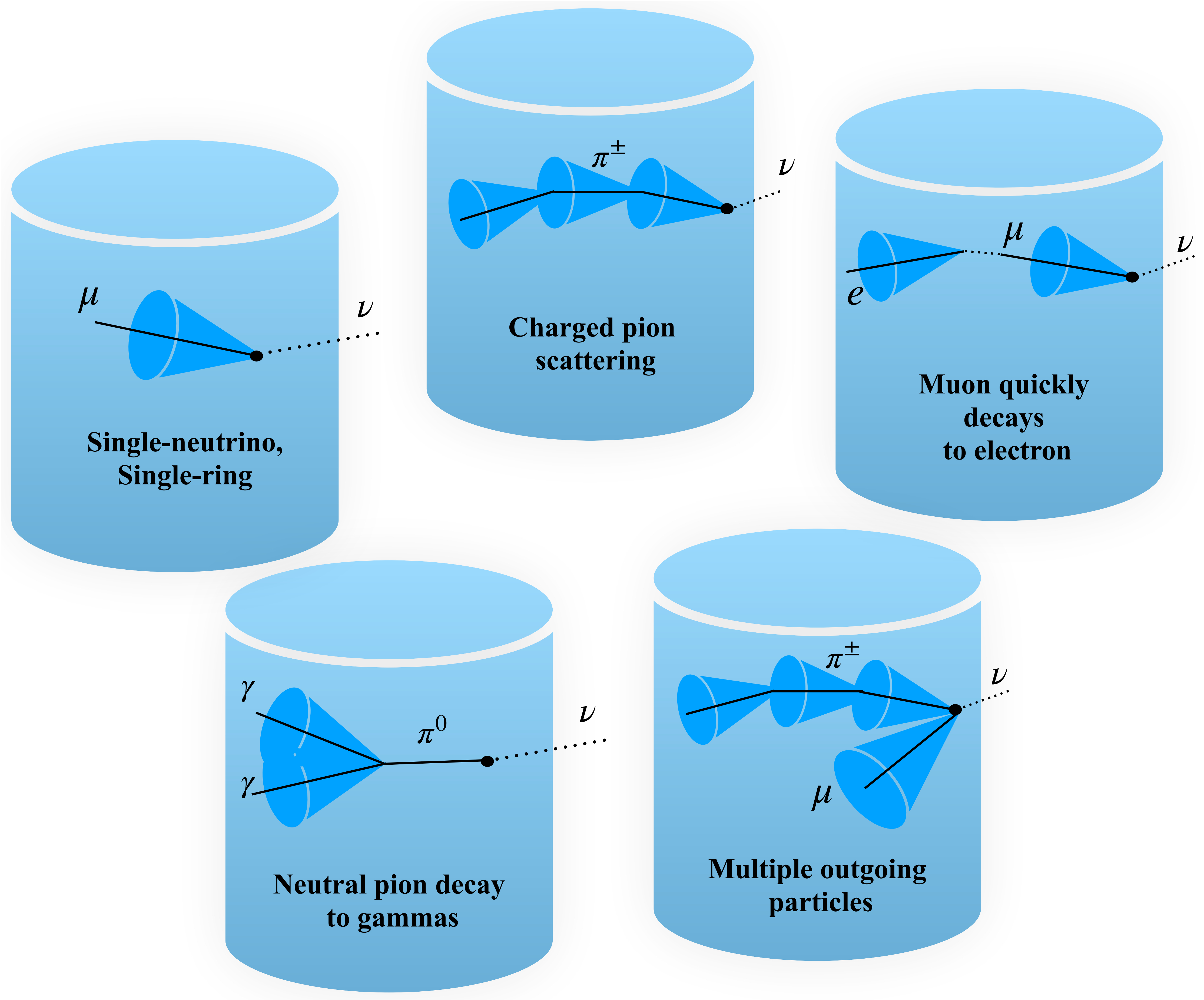
**Plots from WatChMaL meeting's presentation*

Classification: Single-vertex VS Multi-Vertex (Pile-up)

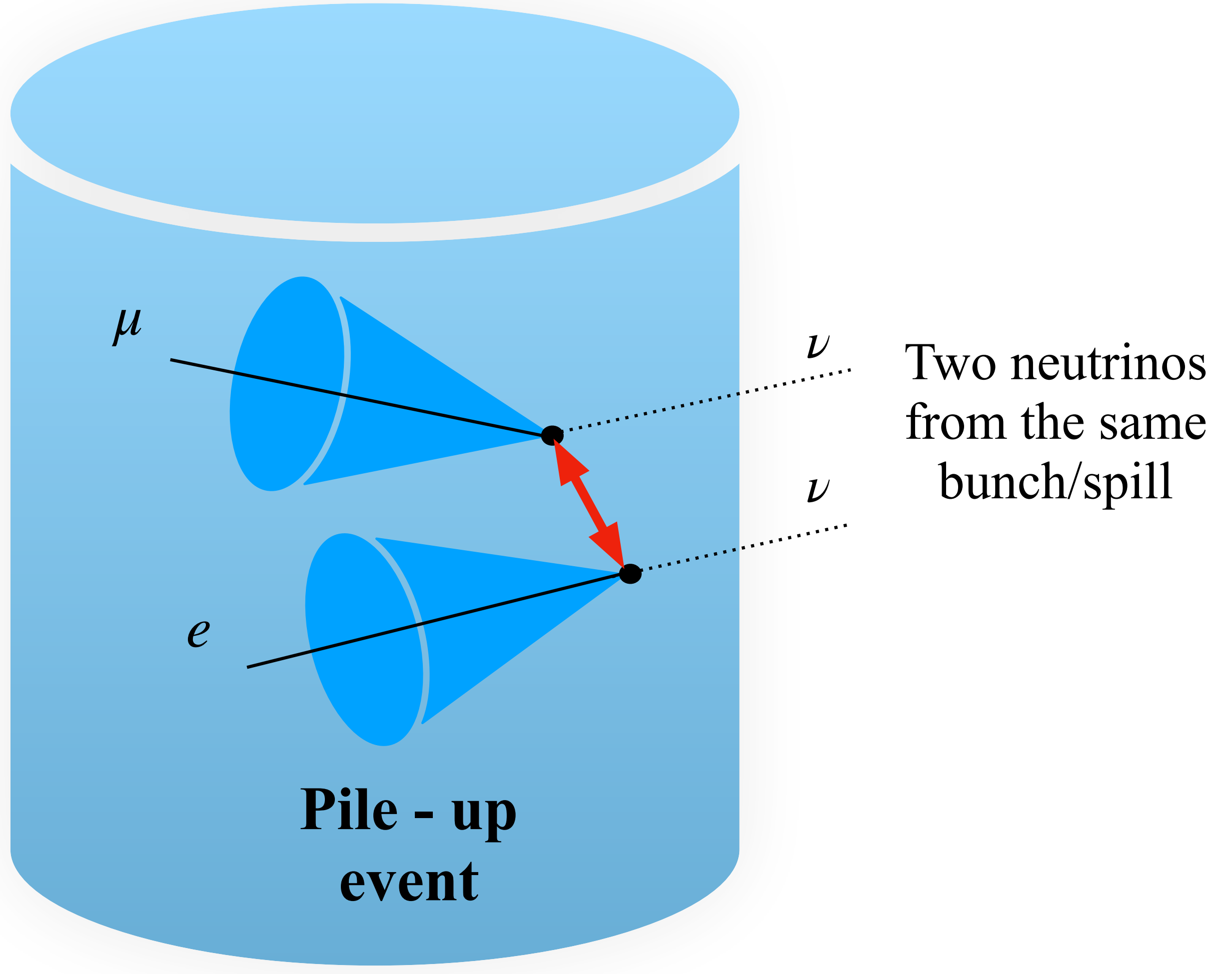
- Due to the high intensity of the beam flux (*d(IWCD,J-PARC) ~ 1km*), **pile-up** events will be significant (~ 20 - 25 %) in **IWCD** detector

- Need to identify pile-up events, separate out complex multi-ring events from multi-vertex pile-up events

Single-vertex

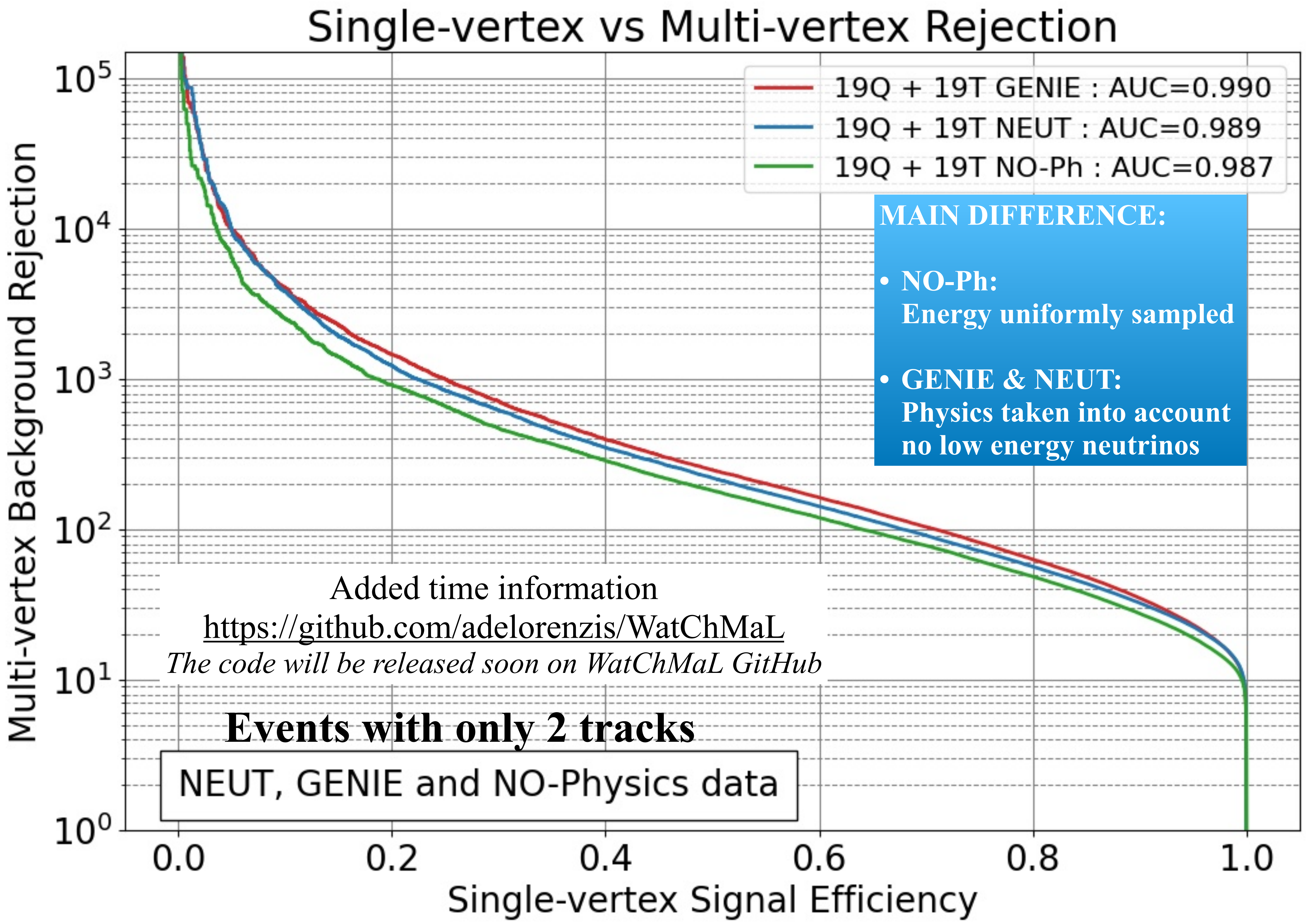


Multi-vertex



- Two *vertices unrelated* to each other and separated in ***time*** and ***space***.

Classification: Single-vertex VS Multi-Vertex (Pile-up)



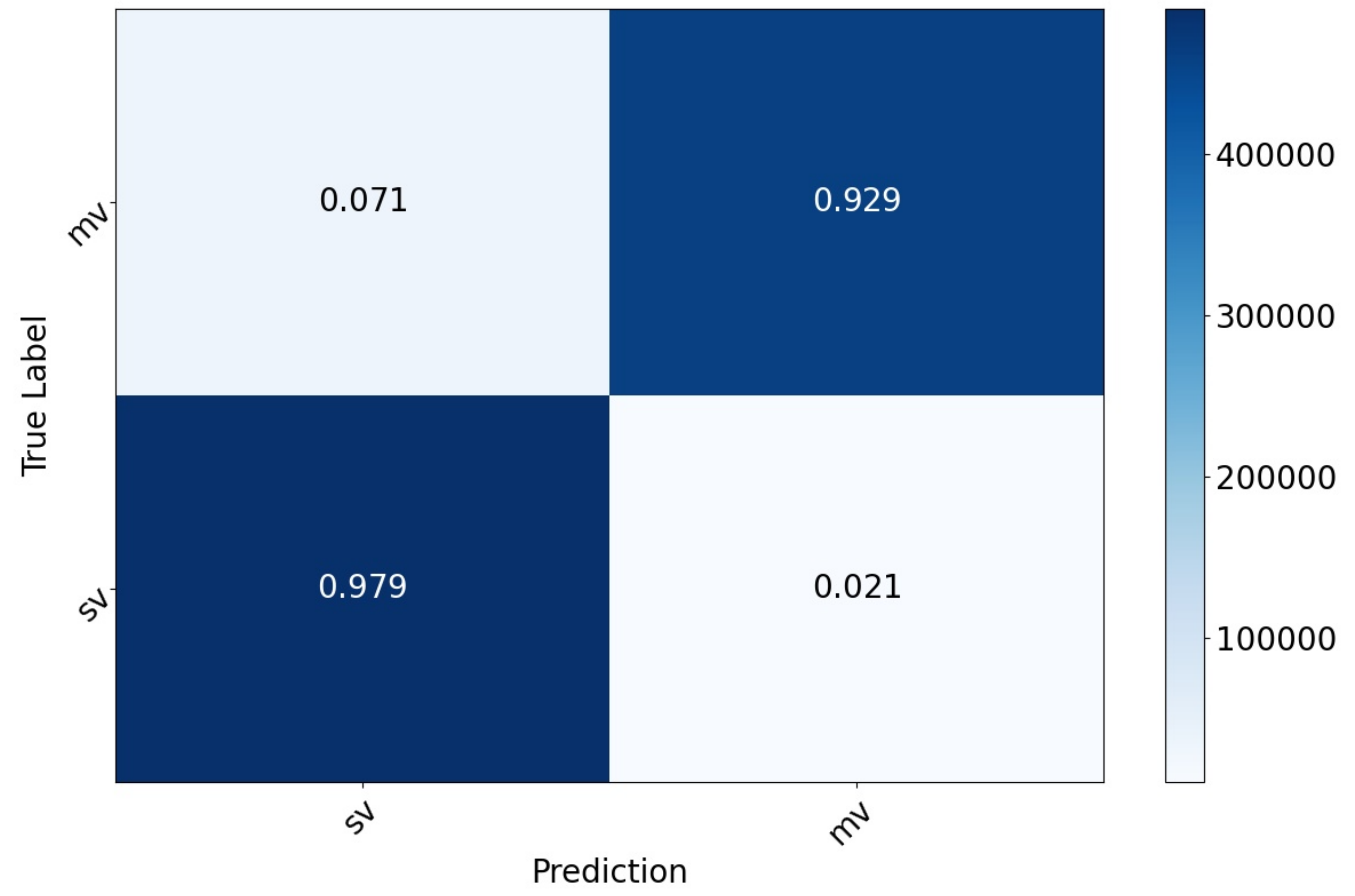
MAIN DIFFERENCE:

- **NO-Ph:**
  Energy uniformly sampled
- **GENIE & NEUT:**
  Physics taken into account
  no low energy neutrinos

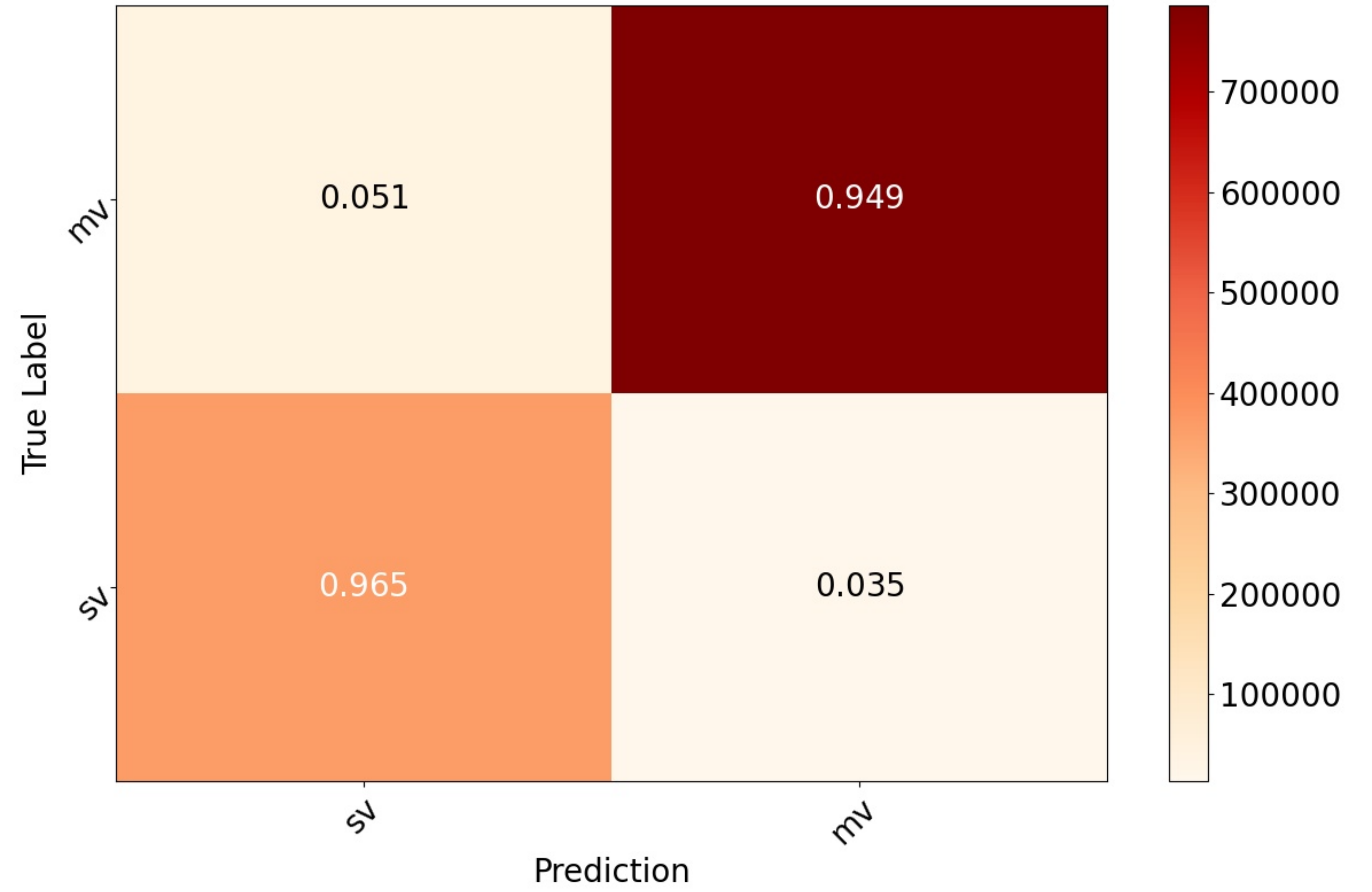
Added time information
https://github.com/adelorenzis/WatChMaL
The code will be released soon on WatChMaL GitHub

Events with only 2 tracks

NEUT, GENIE and NO-Physics data



No-Physics



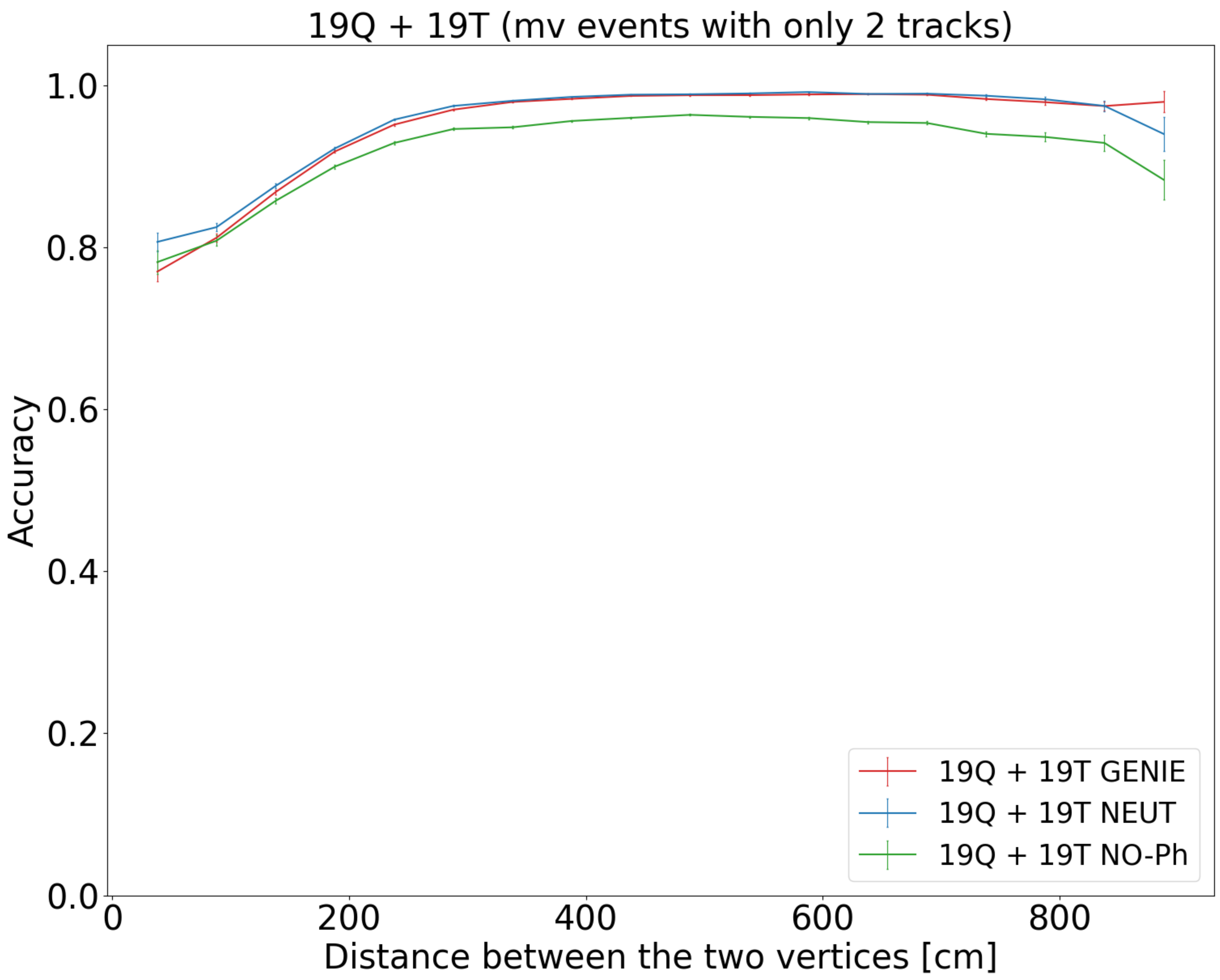
GENIE

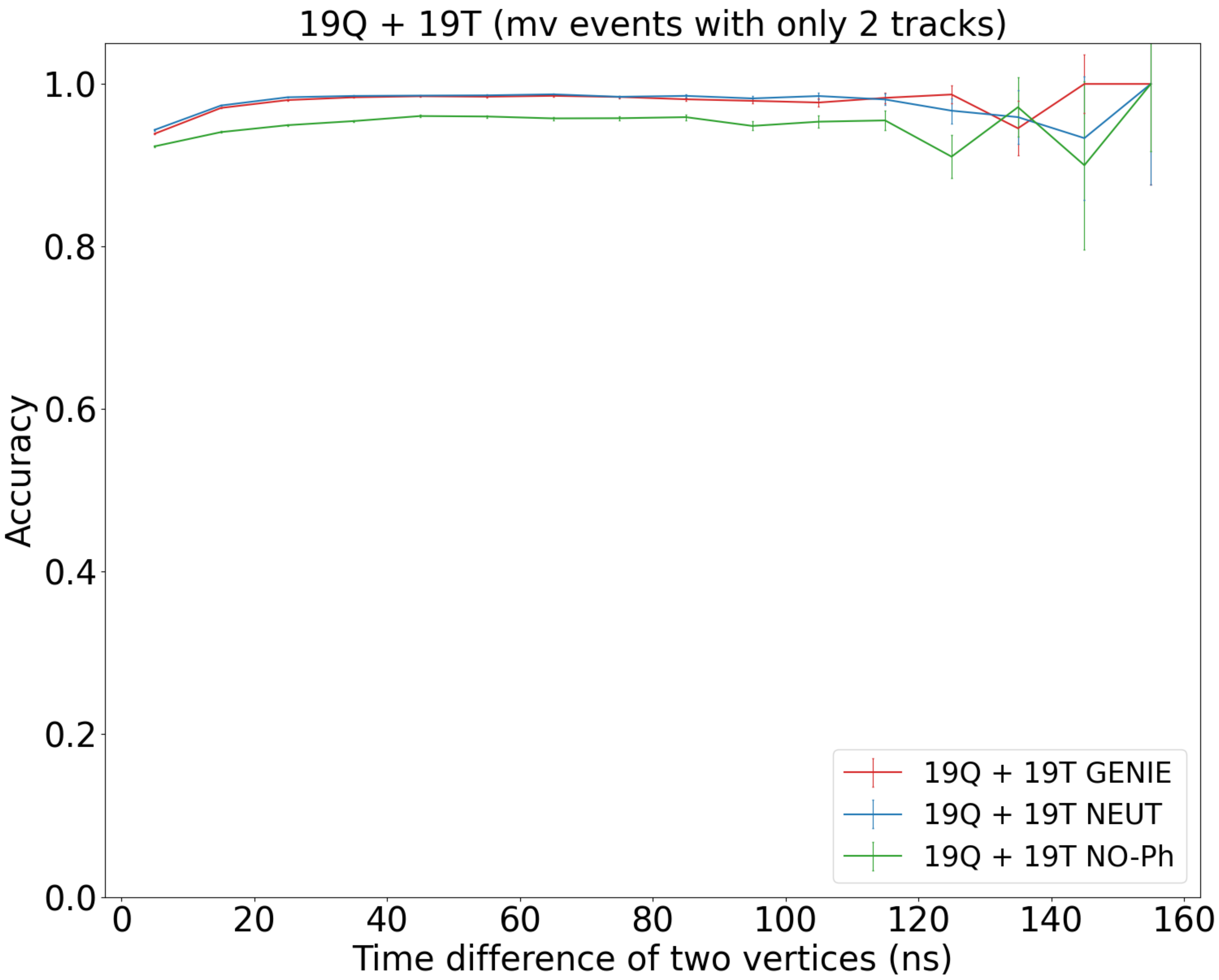
For the training: Dataset generated with 2 to 5 tracks originating from 1 or 2 vertices
For the testing: Dataset generated with 2 tracks originating from 1 or 2 vertices

**ROC : Receiver operating characteristic curve*

*AUC : *Area under the curve*

Classification: Single-vertex VS Pile-up





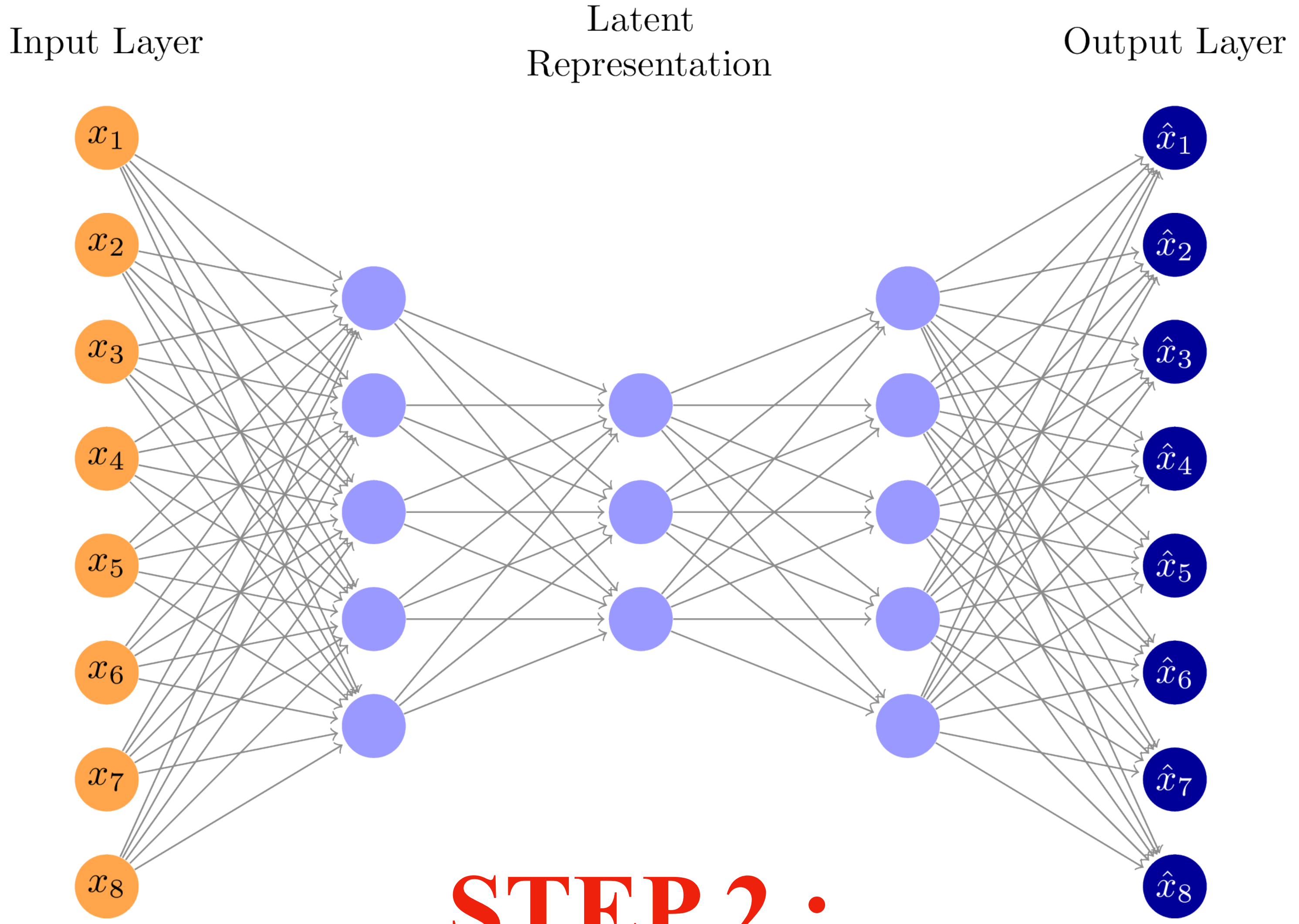
*Accuracy = (# of correct predictions) / (# of total predictions)

*Err_Accuracy= $\sqrt{\frac{(Acc*N+1)(Acc*N+2)}{(N+2)(N+3)} - \frac{(Acc*N+1)^2}{(N+2)^2}}$ arXiv:physics/0701199

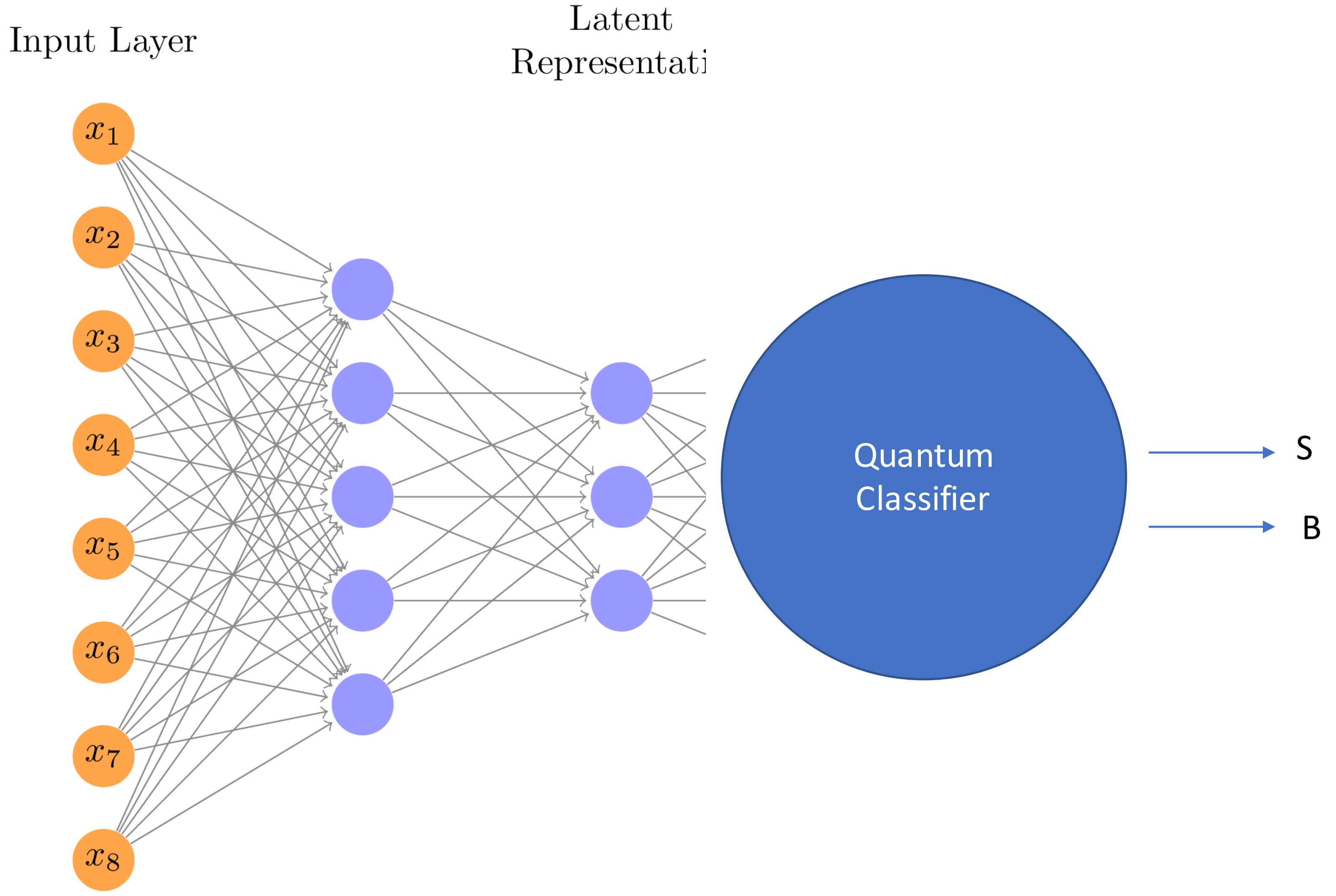
Quantum Machine Learning

- Classification e/γ is a difficult task, need to explore alternative methods
- We are investigating the use of a ***Quantum Classifier***
- Currently Quantum methods are restricted to low dimensional feature space
- We can use the ***auto-encoder*** to construct a dimensionally-reduced “latent” space of new features (compress the input image into a smaller “**latent representation**”)
- The “latent” layer can be subsequently used as a (smaller) input for other analysis (e.g. classification)

STEP 1 : Auto-encoder



STEP 2 : Encoder + Quantum Classifier



Conclusions

- ResNet-18 and PointNet show promising results
- Adding the time information provides significant improvement*
- WCTE will allow to test new ideas with real data
- Exploring new techniques (e.g. quantum classifier)

**Details in the back-up slides*

Back-up slides

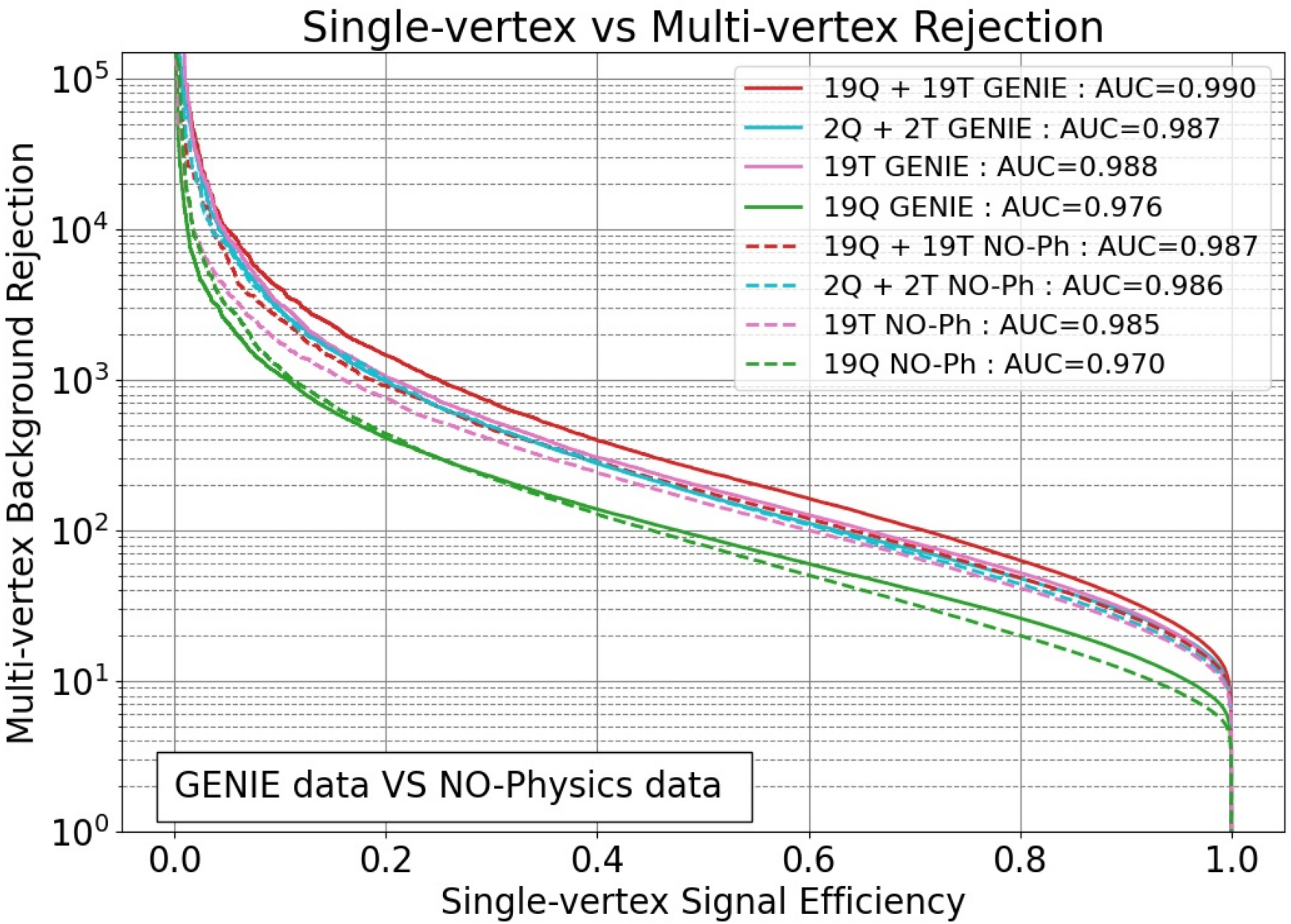
RESULTS (Pile-up study)

GENIE RESULTS	LOSS	ACC	F1	AUC
19 Q + 19 T	0,114	0,954	0,966	0,990
2 Q + 2 T	0,126	0,950	0,963	0,987
19 T	0,126	0,950	0,963	0,988
19 Q	0,180	0,924	0,943	0,976

NO-Physics RESULTS	LOSS	ACC	F1	AUC
19 Q + 19 T	0,128	0,954	0,953	0,987
2 Q + 2 T	0,136	0,951	0,950	0,986
19 T	0,142	0,949	0,948	0,985
19 Q	0,212	0,916	0,913	0,970

The ROC

Events with only 2 tracks



**ROC : Receiver operating characteristic curve*

*AUC : *Area under the curve*

Classification: Single-vertex VS Multi-Vertex (Pile-up)

DETAILS OF DATASETS USED

DATASET NO-PHYSICS (5M for the training + 1M for testing)

Obtained using an adapted *ParticleGunGeneretor* and *WCSim*

Track:
- for each track, energy sampled from a uniform distribution with range [0, remaining_energy]
- tracks isotropically generated

SV (MV)
MaxTotEvisPerEvt=2000 MeV
MaxVtxPerEvt=1 (2)
MaxTrkPerEvt=5
MinTrkPerVtx=2 (1)

e^-, μ^-, π^0, π^-

DATASETS GENIE and NEUT (for testing)

Obtained using *GENIE and NEUT software* and *WCSim*

Physics taken into account

SV (MV)
MaxVtxPerEvt=1 (2)
MinTrkPerVtx=2 (1)

$e^-, e^+ \mu^-, \pi^0, \pi^-, \pi^+, p, \gamma$

For all datasets

Vertex position (= track initial position):
- x, z uniformly sampled in a circle of radius R = 400 cm
- y uniformly sampled in [-300, 300] cm

Vertex time (= track initial time)
Sampled from a Gaussian distribution with μ = 200 ns and σ = 25 ns

Detector information used in the simulation

daq.mac

```
/DAQ/DigitizerOpt/TimingPrecision 0.1
/DAQ/TriggerSaveFailures/Mode 1
/DAQ/TriggerSaveFailures/TriggerTime 100
/DAQ/TriggerSaveFailures/PreTriggerWindow  -400
/DAQ/TriggerSaveFailures/PostTriggerWindow +2950
/DAQ/TriggerNDigits/Threshold 2500000
/DAQ/TriggerNDigits/Window 200
#/DAQ/TriggerNDigits/AdjustForNoise true
```