IMPROVING EVENT RECONSTRUCTION IN HYPER KAMIOKANDE WITH MACHINE LEARNING

OCTOBER 27TH 2025 - ANDREW ATTA

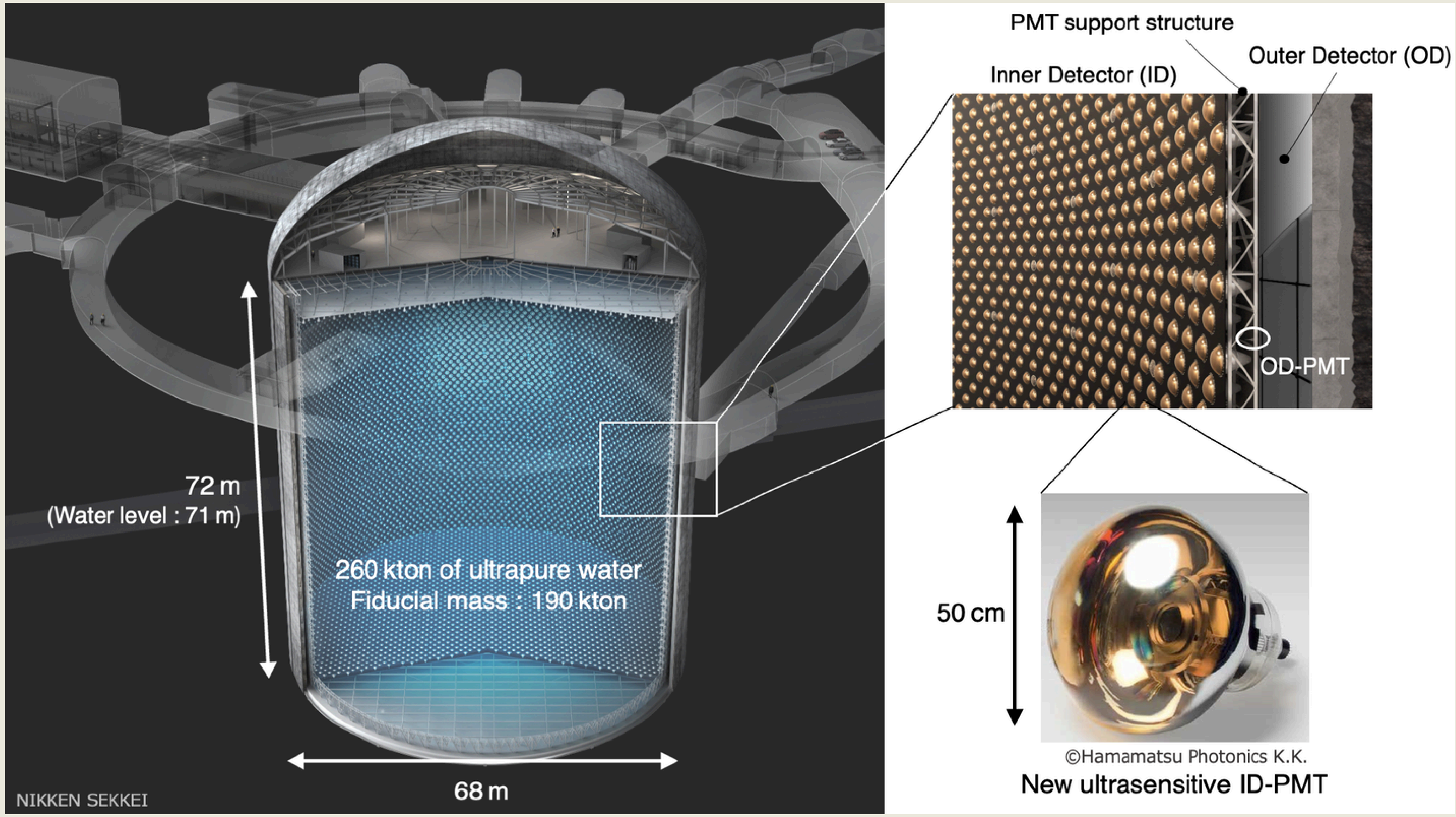
WatChMaL

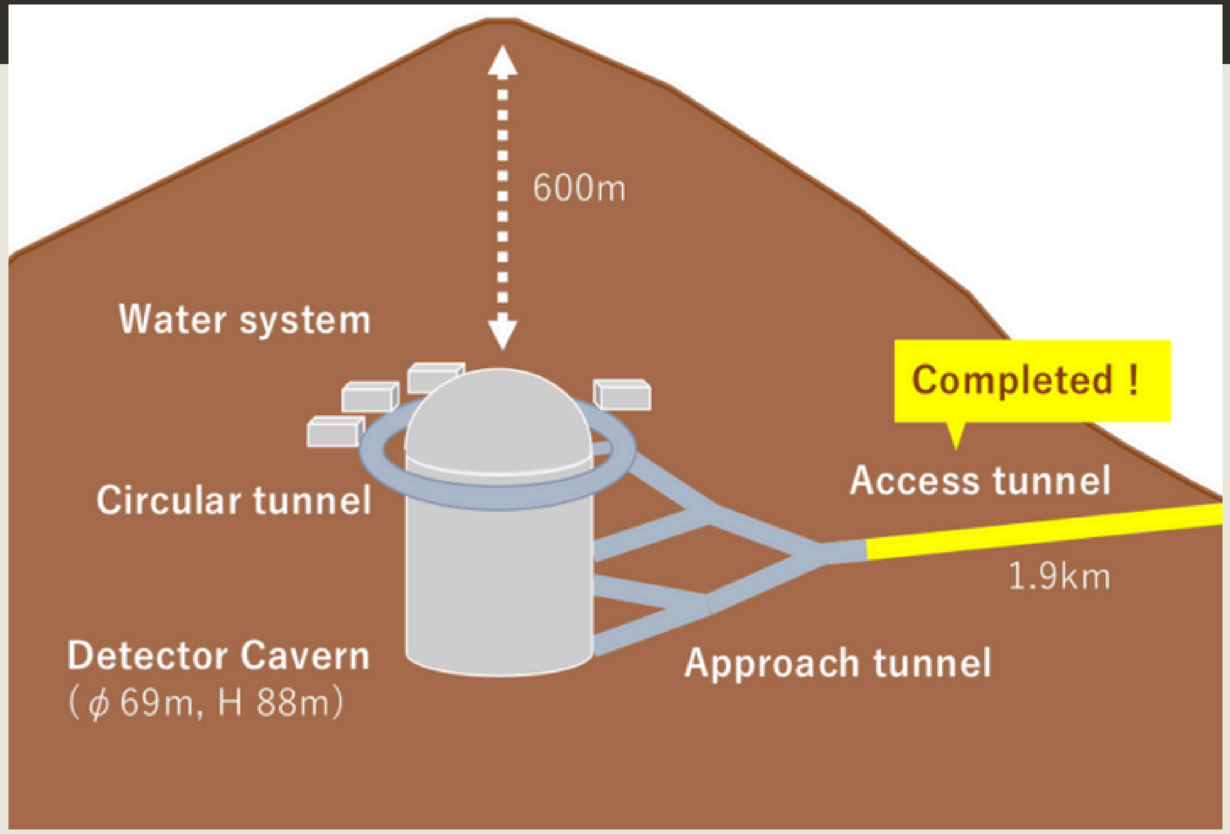
Hyper-K

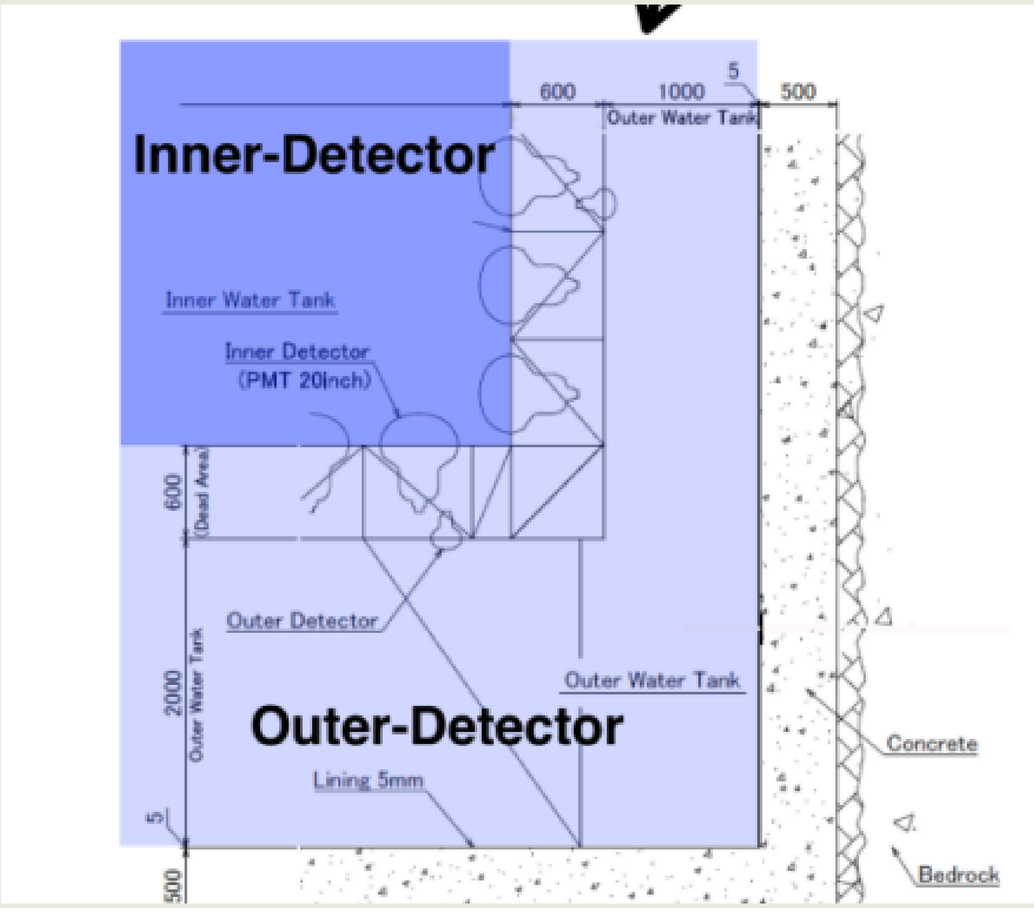
MONASH University

NIKKEN SEKKEI

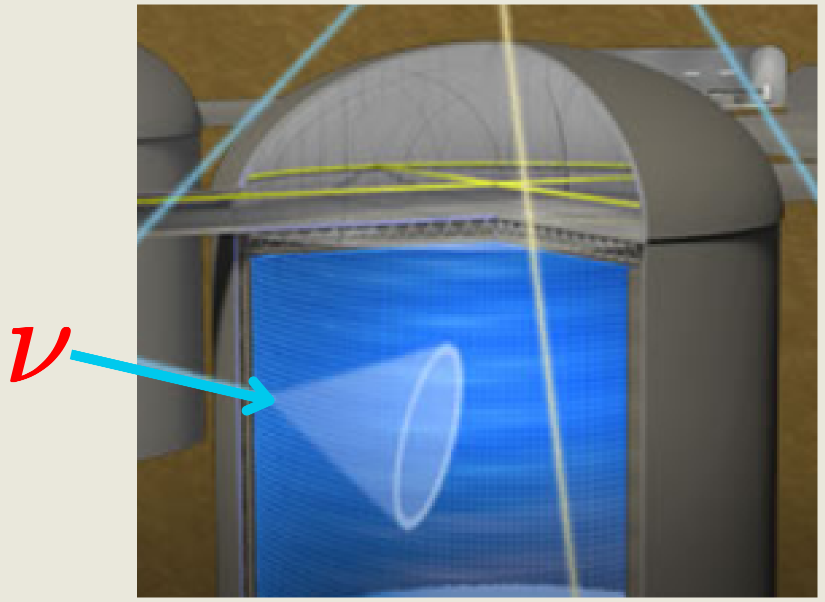
HYPER KAMIOKANDE

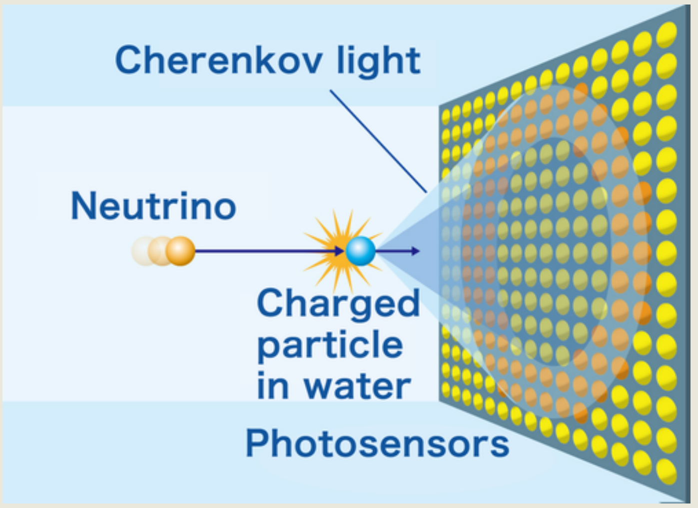


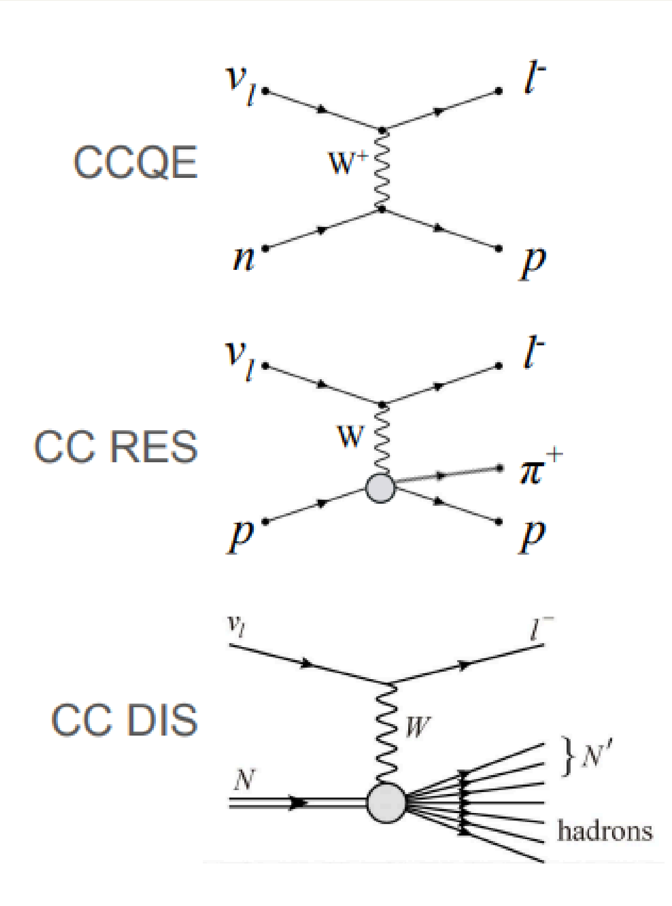




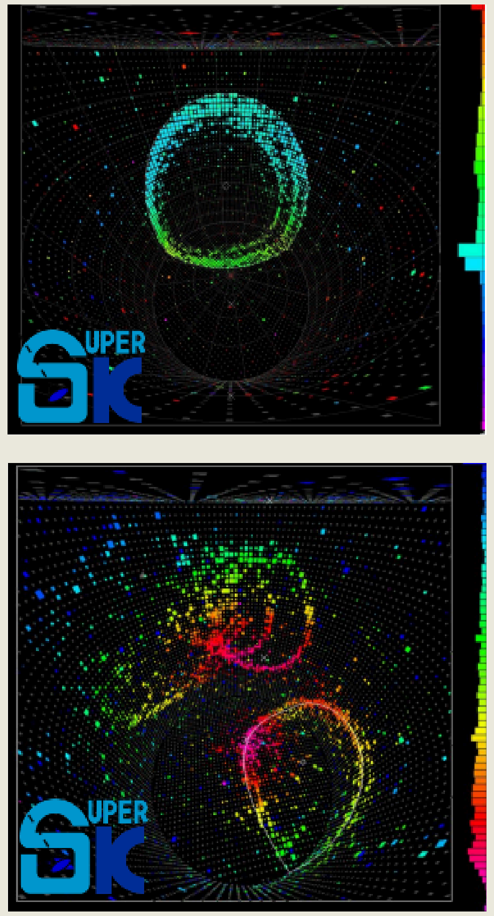
EVENT TYPOLOGY

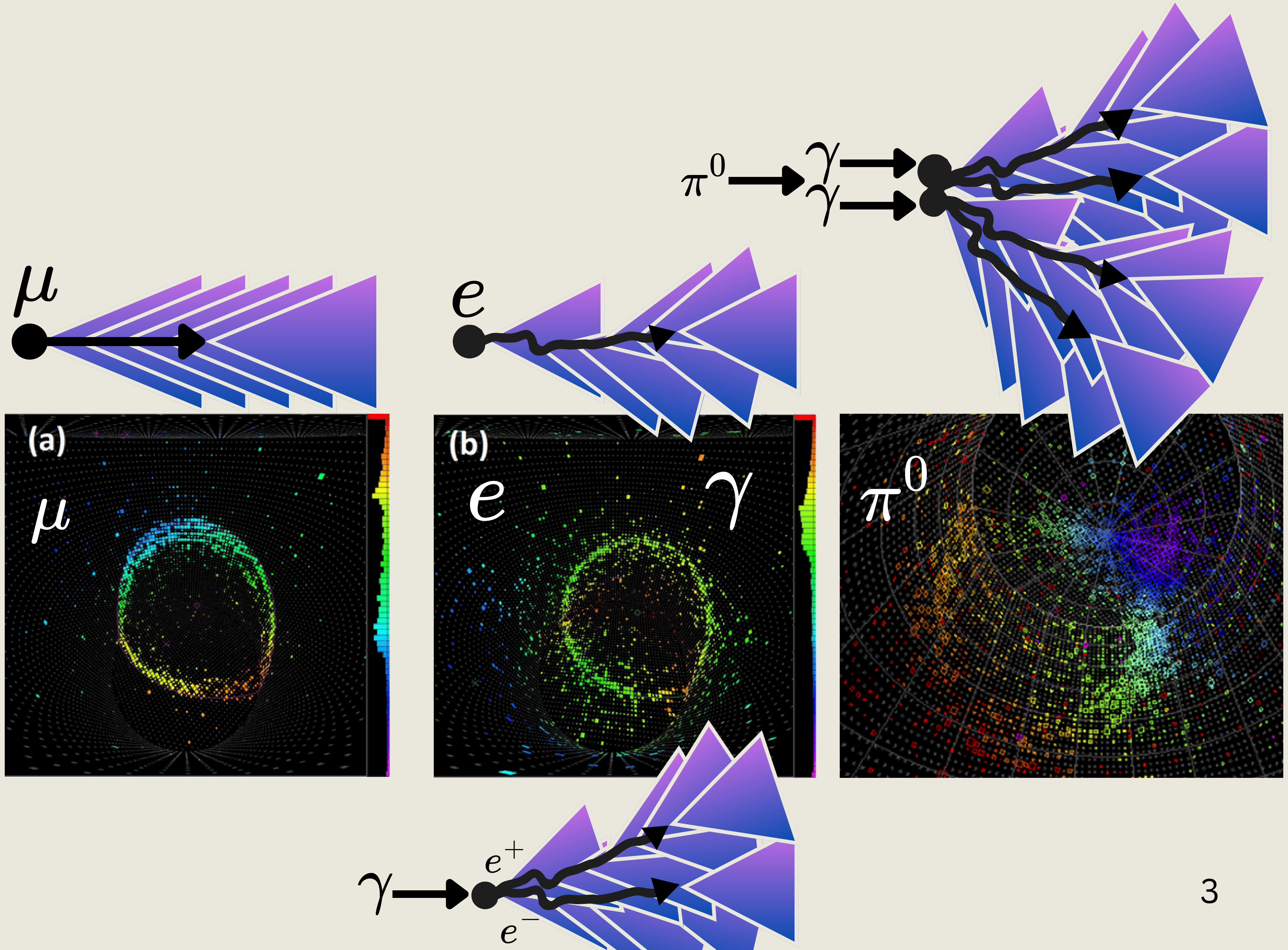




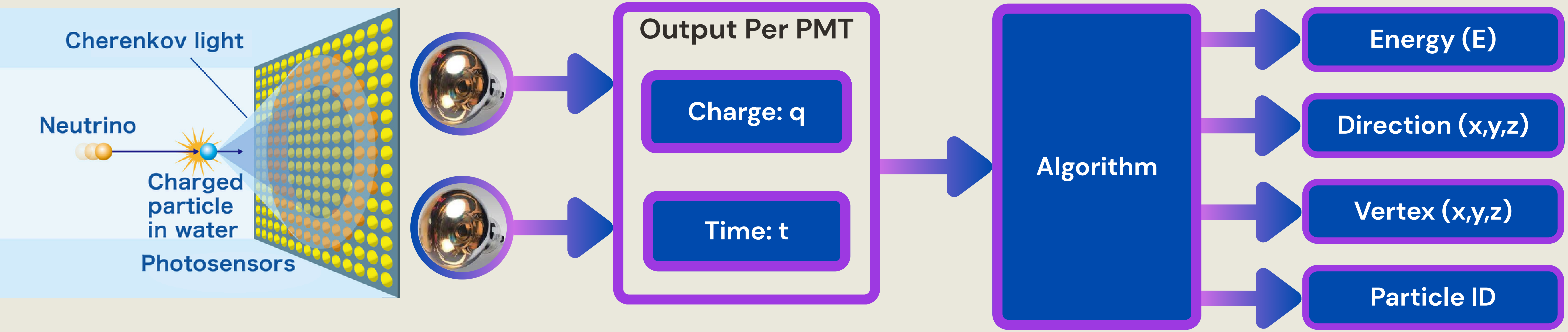


Increasing Energy





EVENT RECONSTRUCTION



TRADITIONAL EVENT RECONSTRUCTION

- In Super-Kamiokande event reconstruction is handled through an algorithm known as **fiTQun**
    - Maximum likelihood estimator:

fit parameters (7D) | Hit Probability | charge of i^{th} PMT | time of i^{th} PMT

$$L(\mathbf{x}) = \prod_{j}^{unhit} P_j(unhit|\mathbf{x}) \prod_{i}^{hit} P_i(hit|\mathbf{x}) f_q(q_i|\mathbf{x}) f_t(t_i|\mathbf{x})$$

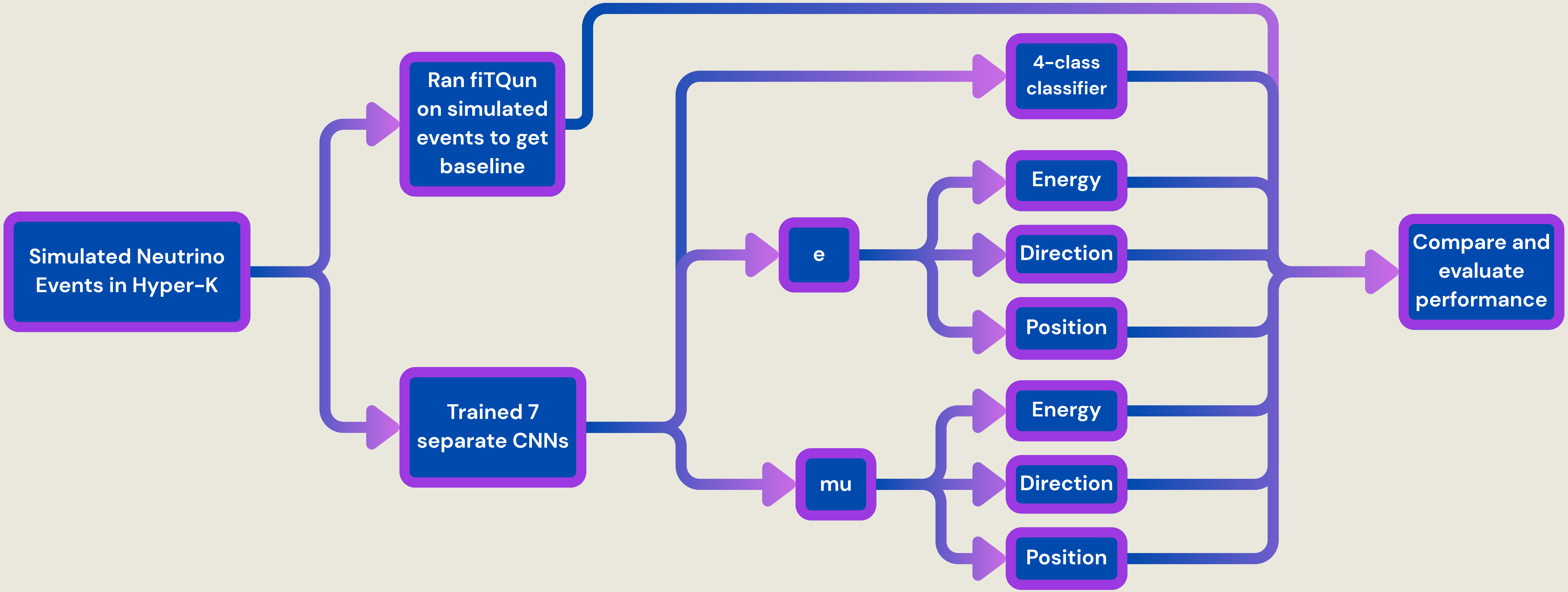
Unhit Probablity | PMT hit charge pdf | PMT hit time pdf

$$\mathbf{x} = (x, y, z, t, p, \theta, \phi)$$

- Does so for multiple different particle hypotheses, then takes a ratio between each to pick the most likely event hypothesis
- Robust and has shown great success in Super-K
- Starting to see the upper limit of performance

NEURAL NETWORKS TO THE RESCUE

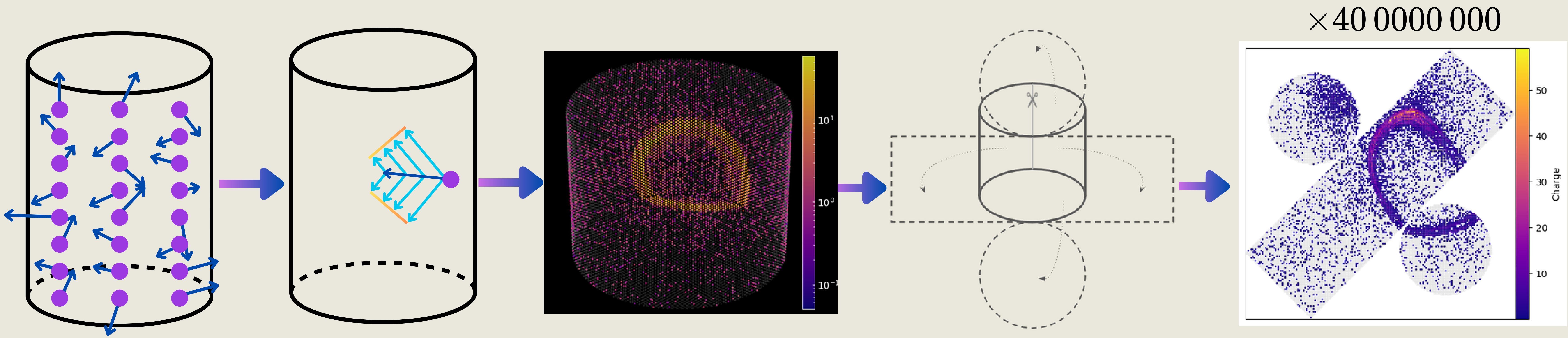
- In recent years there has been a push to adopt machine learning based inference techniques
- 1M events can be processed by fiTQun in 100,000 CPU hours
- Same 1M events can be processed by ML in less than 100 CPU hours or **10 GPU minutes**
- Proven on other water Cherenkov detectors



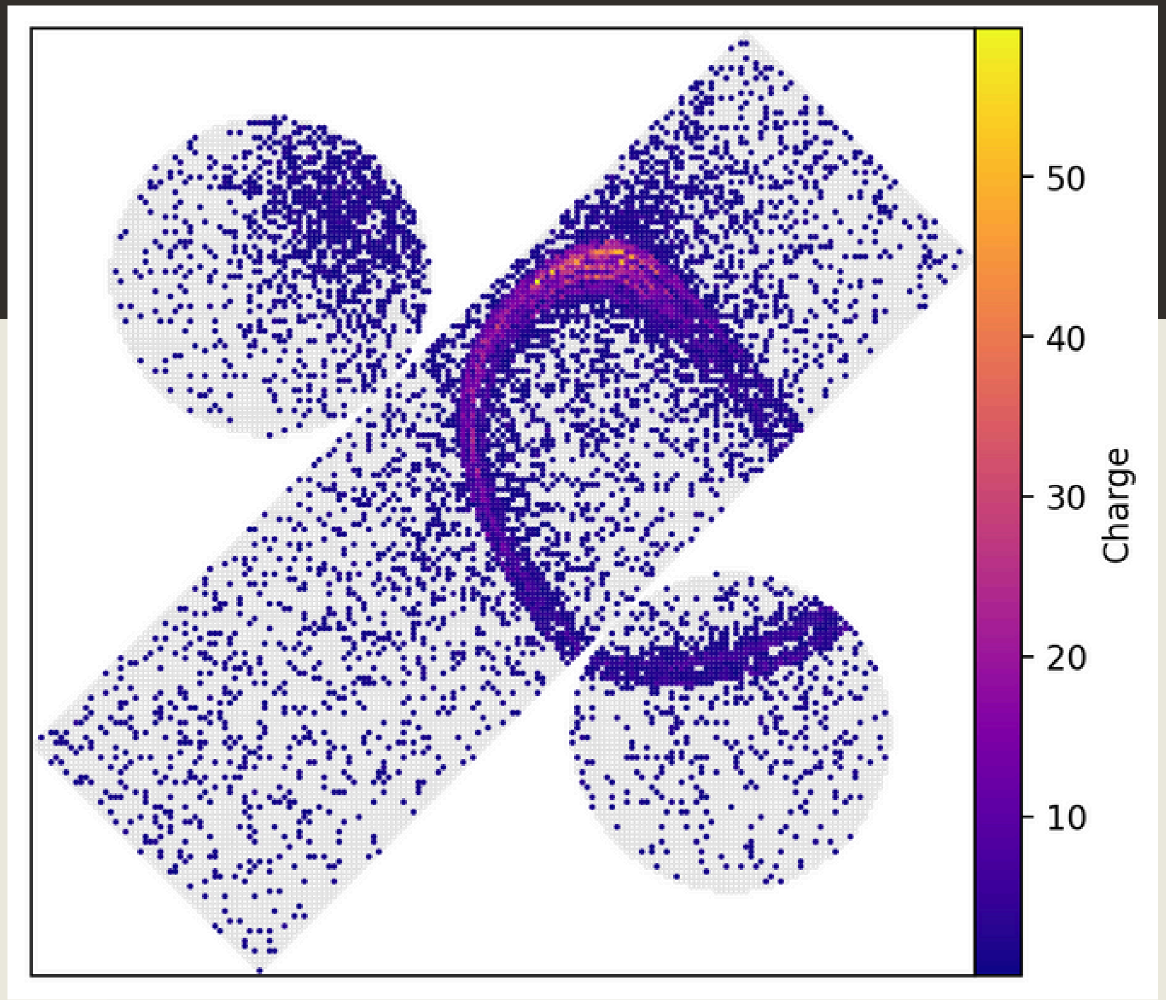
SIMULATING THE DATA

- Simulated the tank and Cherenkov producing process in WCSim - a Geant4 based tool designed for SK and HK
- total of 40 million events - 10 million electrons, muons, gamma rays, and neutral pions
- Uniform energy distribution of 0-2GeV kinetic energy ***above*** Cherenkov threshold for that particular particle
- Uniform initial position in tank
- Isotropic direction distribution
- Only **fully contained** events kept
- Muon decay disabled in muon events

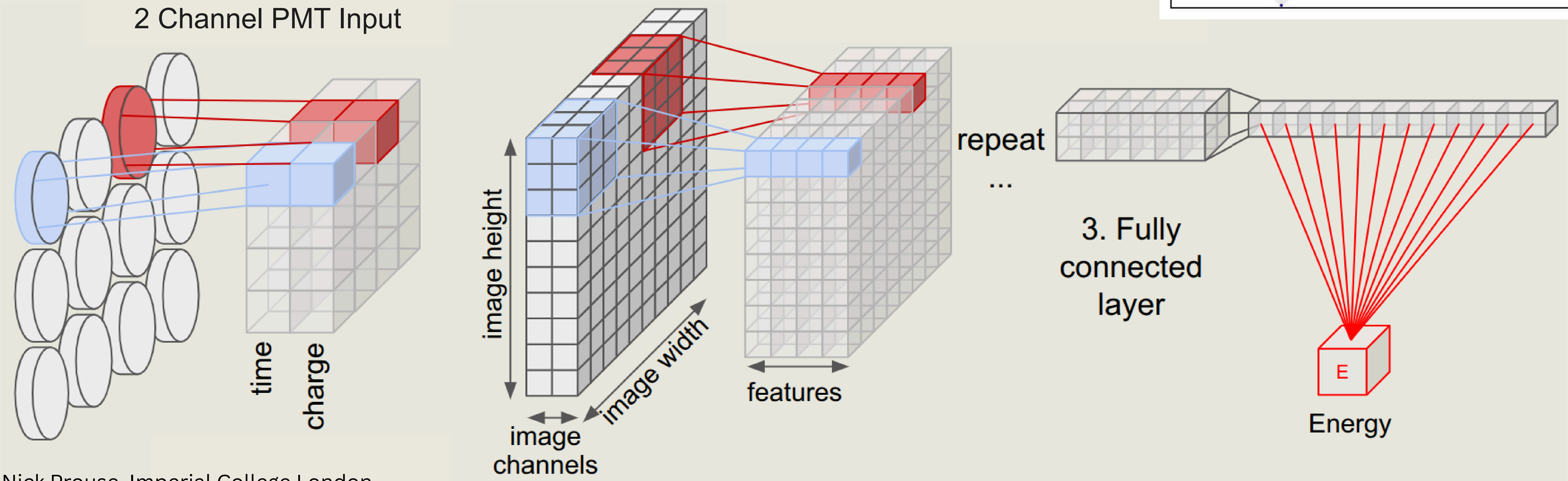
Cuts	FC and nhits>200
Energy	0–2GeV
PID	μ, e, γ, π^0
Muon Decay?	no decay
Training Split	96:2:2
Training Set	~8,569,149
Validation Set	~178,449
Testing Set	~178,595



THE MODELS IN QUESTION



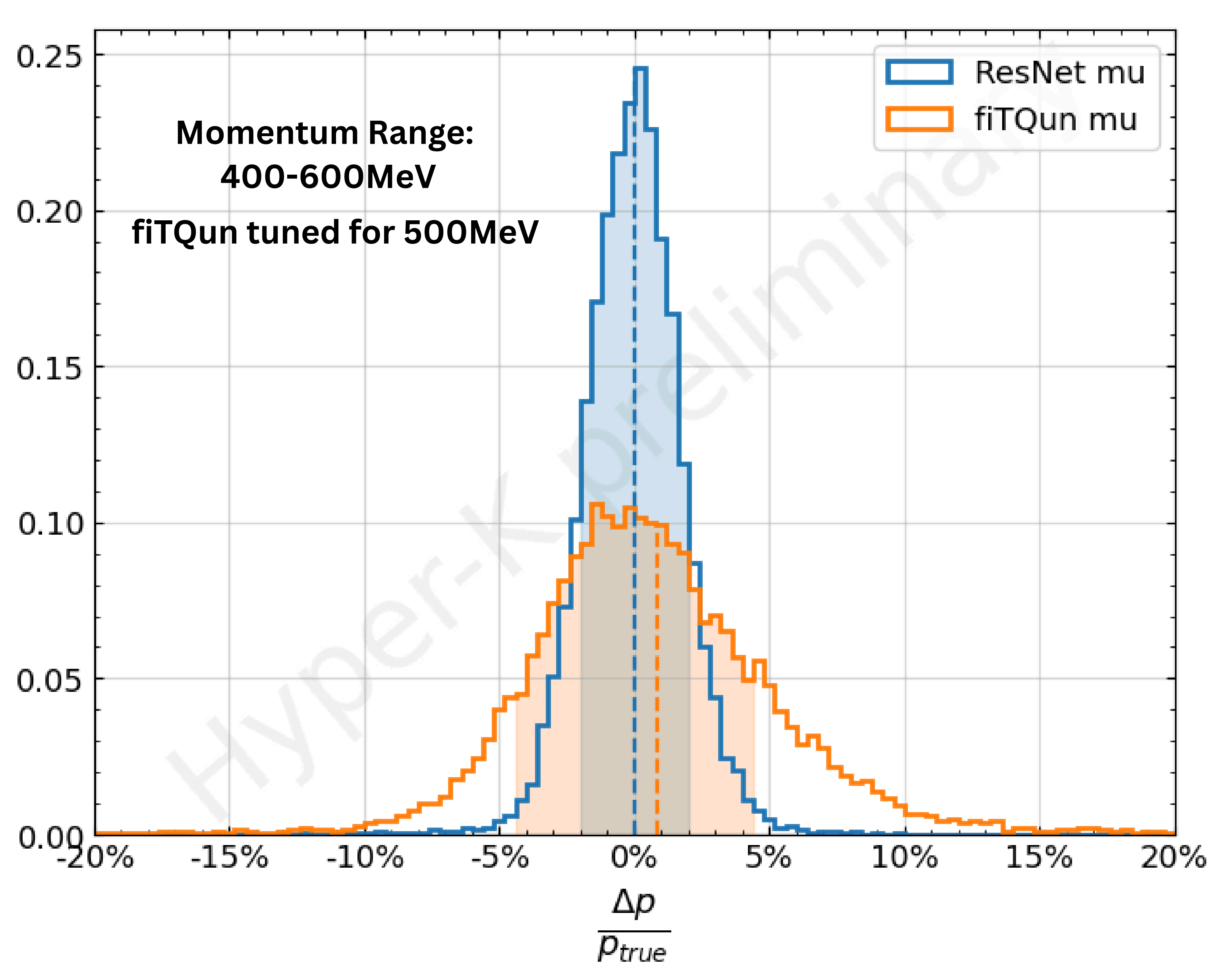
- Each used model used **ResNet-152** as a base and used the WatChMaL (**Wat**er **Ch**erenkov **Ma**chine **L**earning) framework
- Depending on parameter being reconstructed, the input/outputs look different and the loss function is different but the core is still the same
- **7 models in total**
  - 4-class classifier
  - muon regressors - direction, position, energy
  - electron regressors - direction, position, energy

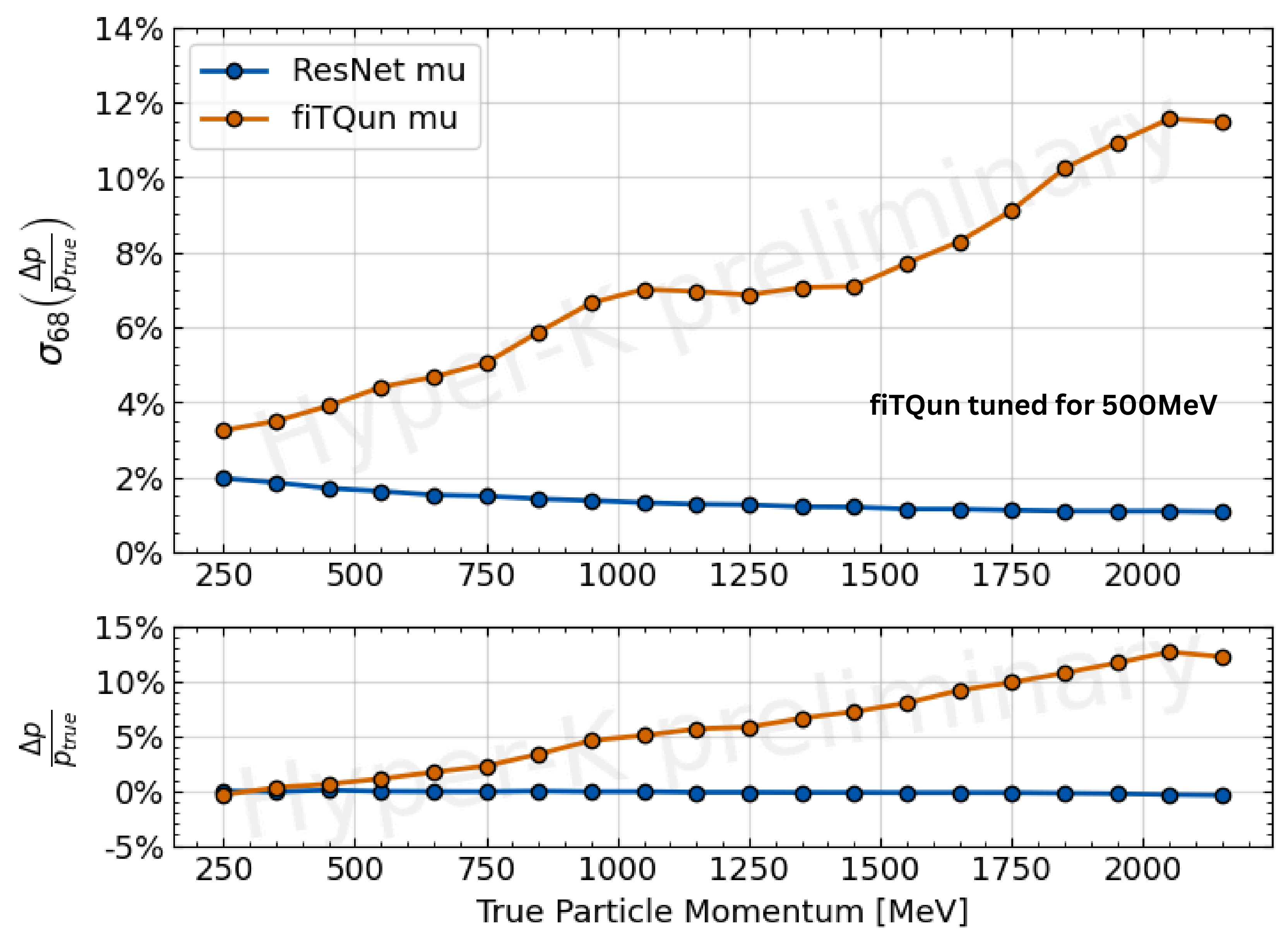


Nick Prouse, Imperial College London

MUON REGRESSION

MOMENTUM RECONSTRUCTION





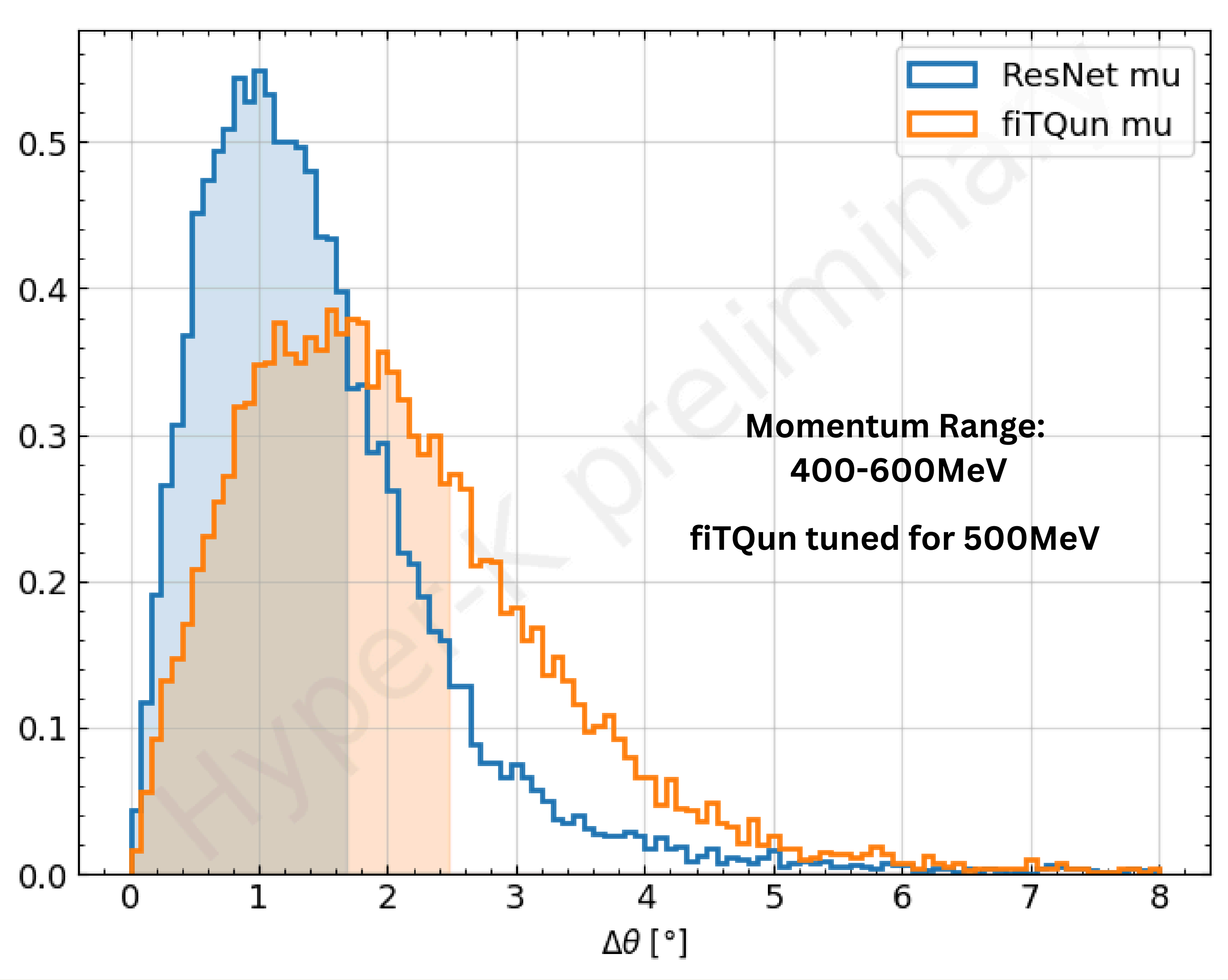
$$\frac{\Delta p}{p_{true}} = \frac{\text{True p} - \text{Predicted p}}{\text{True p}} \times 100$$

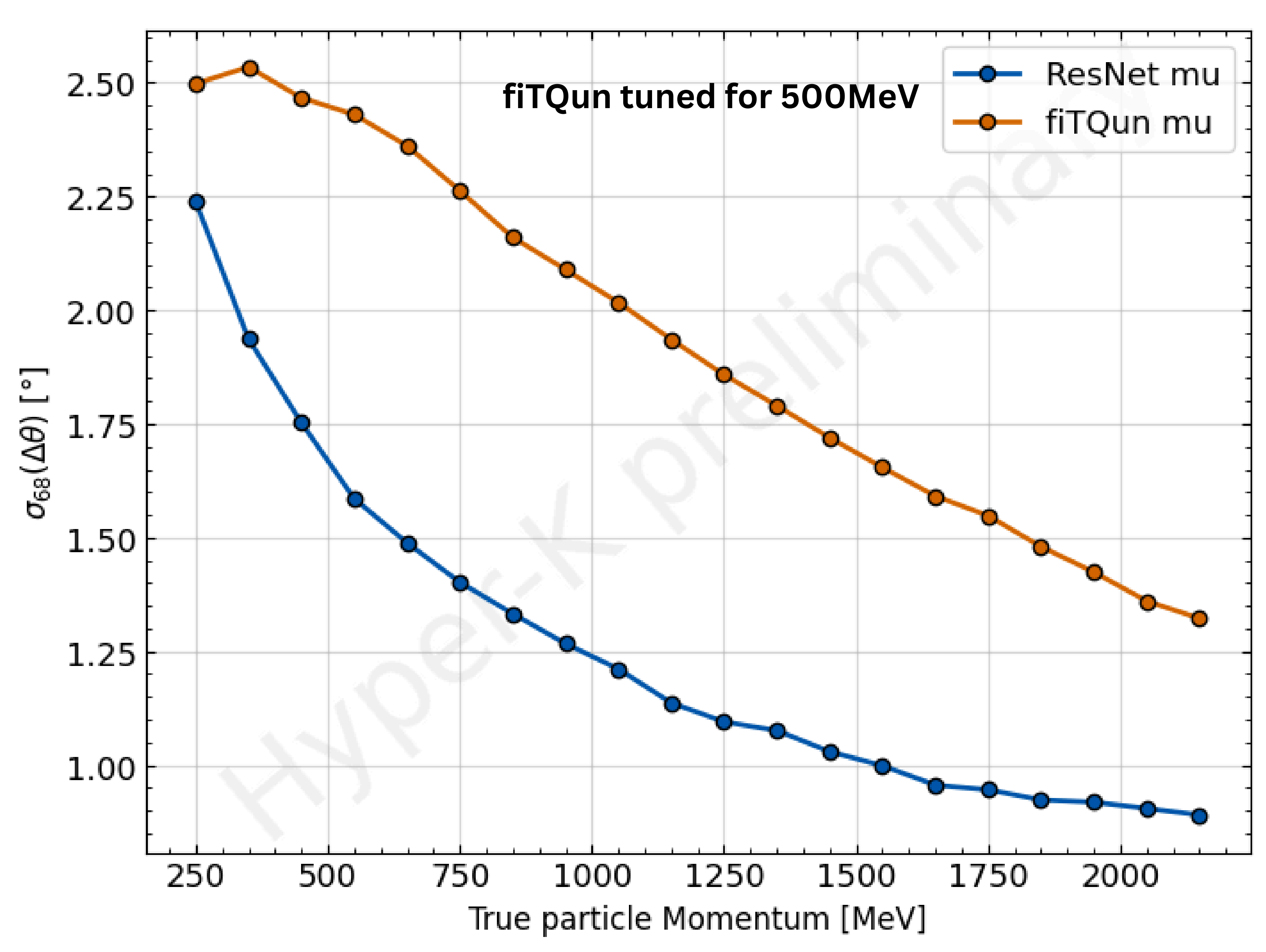
$$\sigma_{68}\left(\frac{\Delta p}{p_{true}}\right) =$$ *absolute* fractional momentum residual below which 68% of events fall under. Analogous to 1 standard deviation

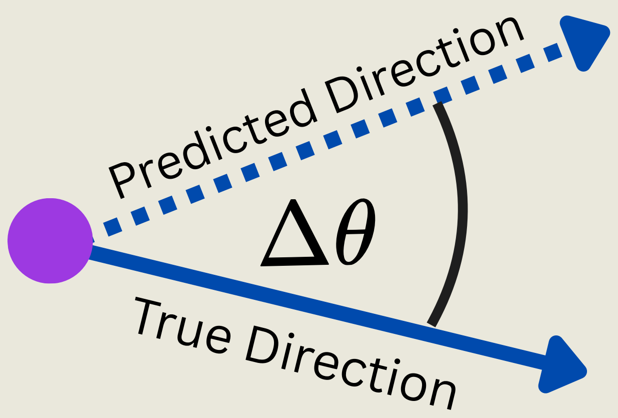
The version of fiTQun I compare results to does not include scattered light despite MC including it

400-600MeV	ResNet	fiTQun
Average $\sigma_{68}\left(\frac{\Delta p}{p_{true}}\right)$	1.67%	4.15%
Average $\frac{\Delta p}{p_{true}}$	0.02%	0.87%

DIRECTION RECONSTRUCTION



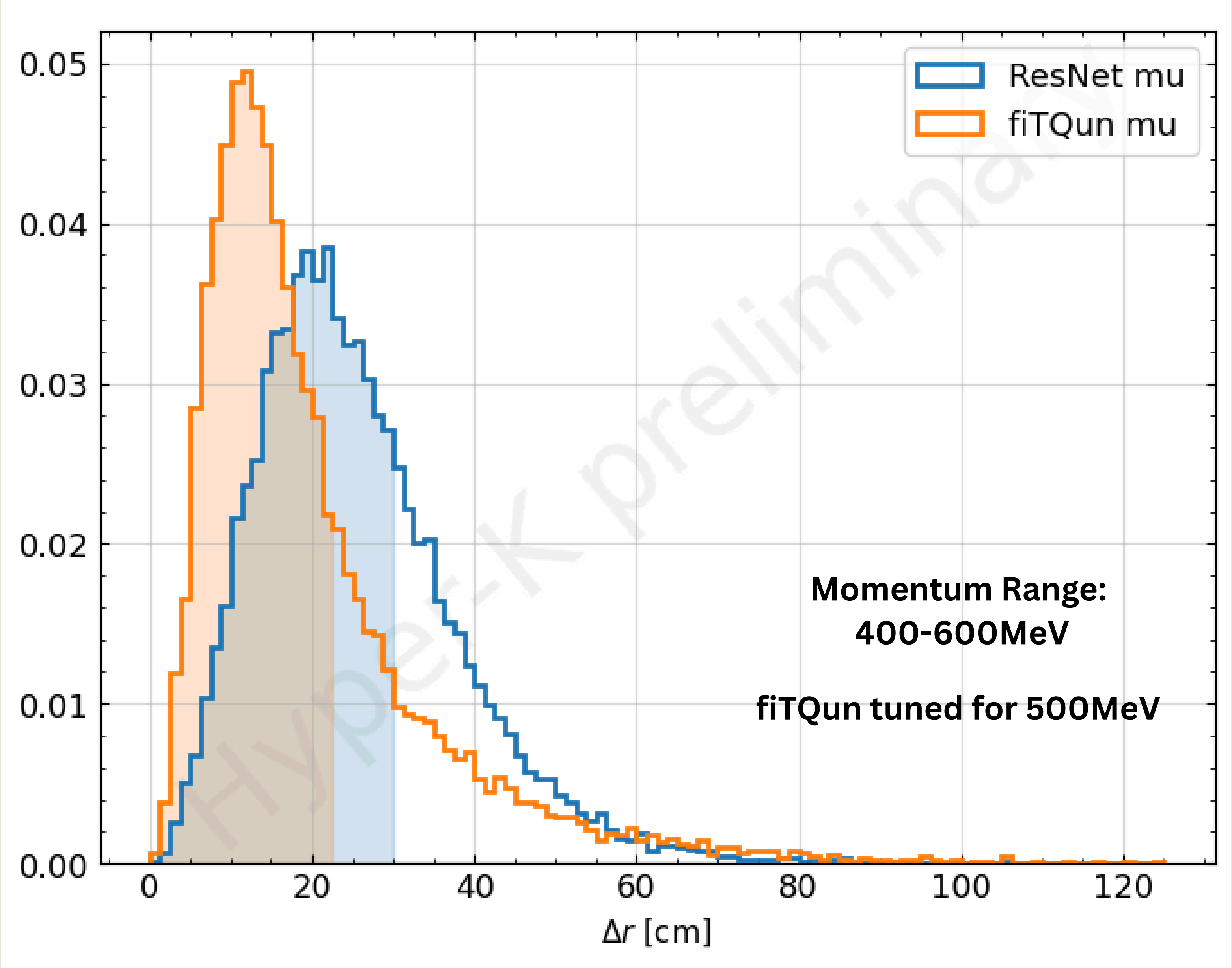


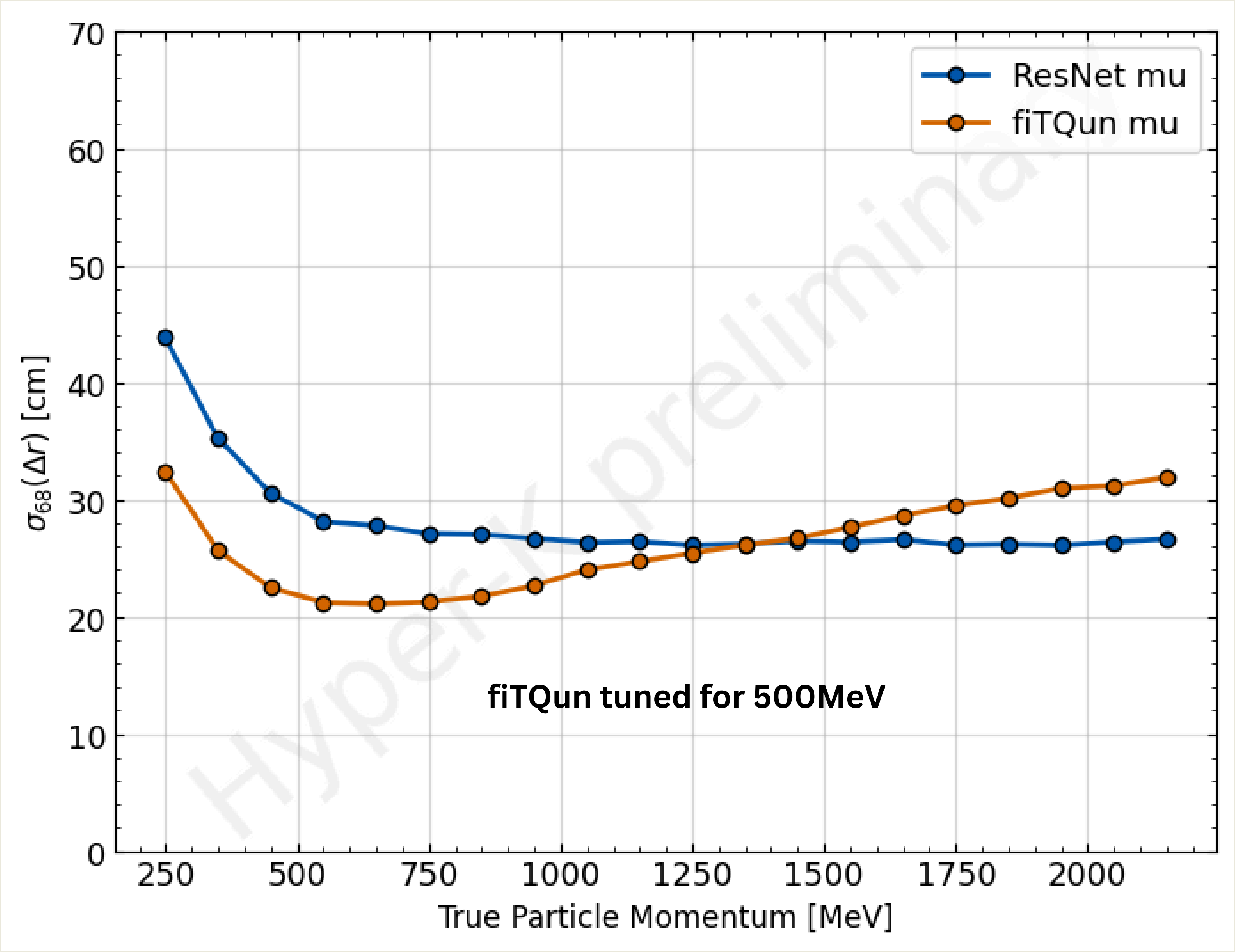


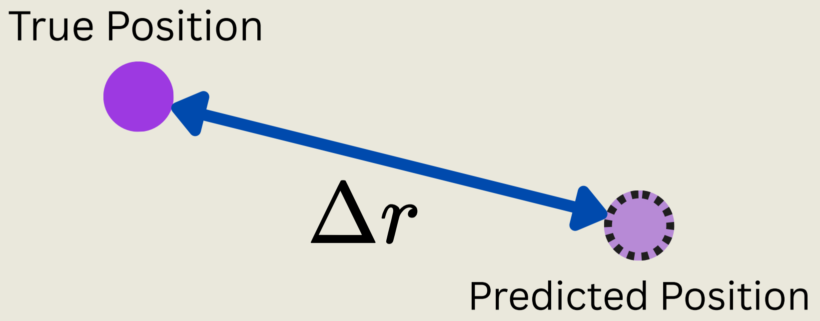
	ResNet	fiTQun
Average $\sigma_{68}\left(\Delta\theta\right)$	1.66°	2.45°

$\sigma_{68}\left(\Delta\theta\right) =$ opening angle residual below which 68% of events fall under

POSITION RECONSTRUCTION







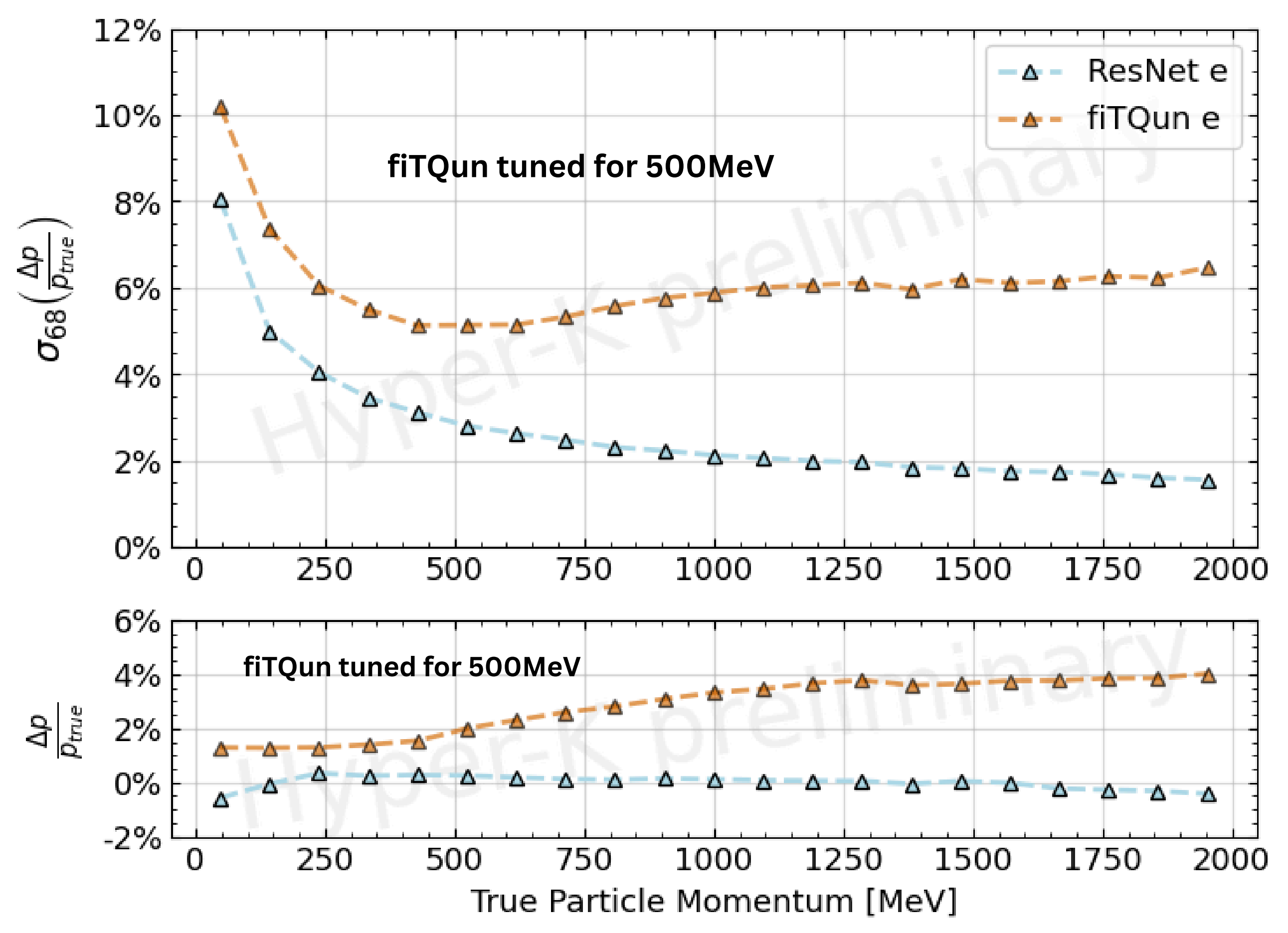
	ResNet	fiTQun
Average $\sigma_{68}\left(\Delta r\right)$	29.4 cm	21.7 cm

$\sigma_{68}\left(\Delta r\right) =$ 3D distance residual below which 68% of events fall under

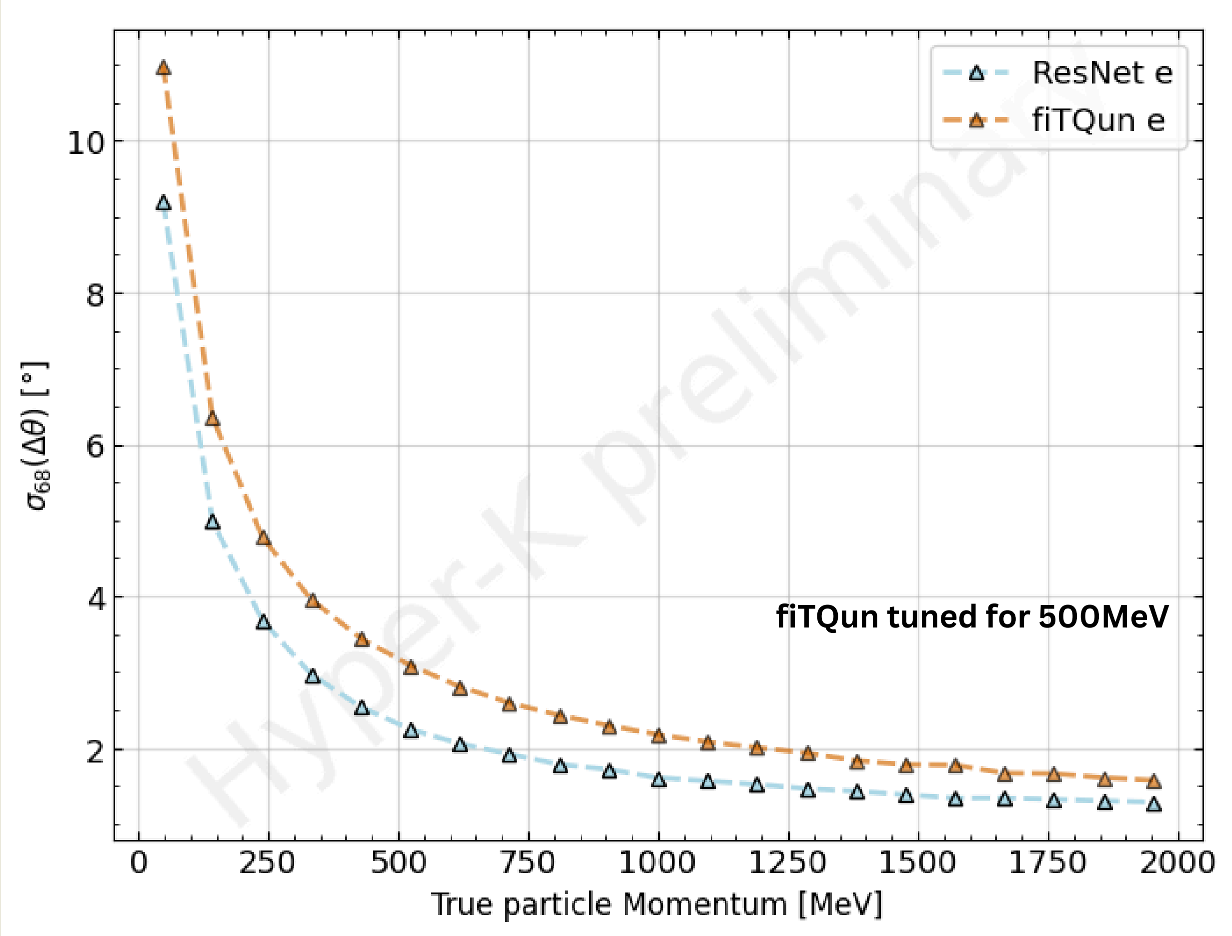
ELECTRON REGRESSION

ELECTRON RECONSTRUCTION

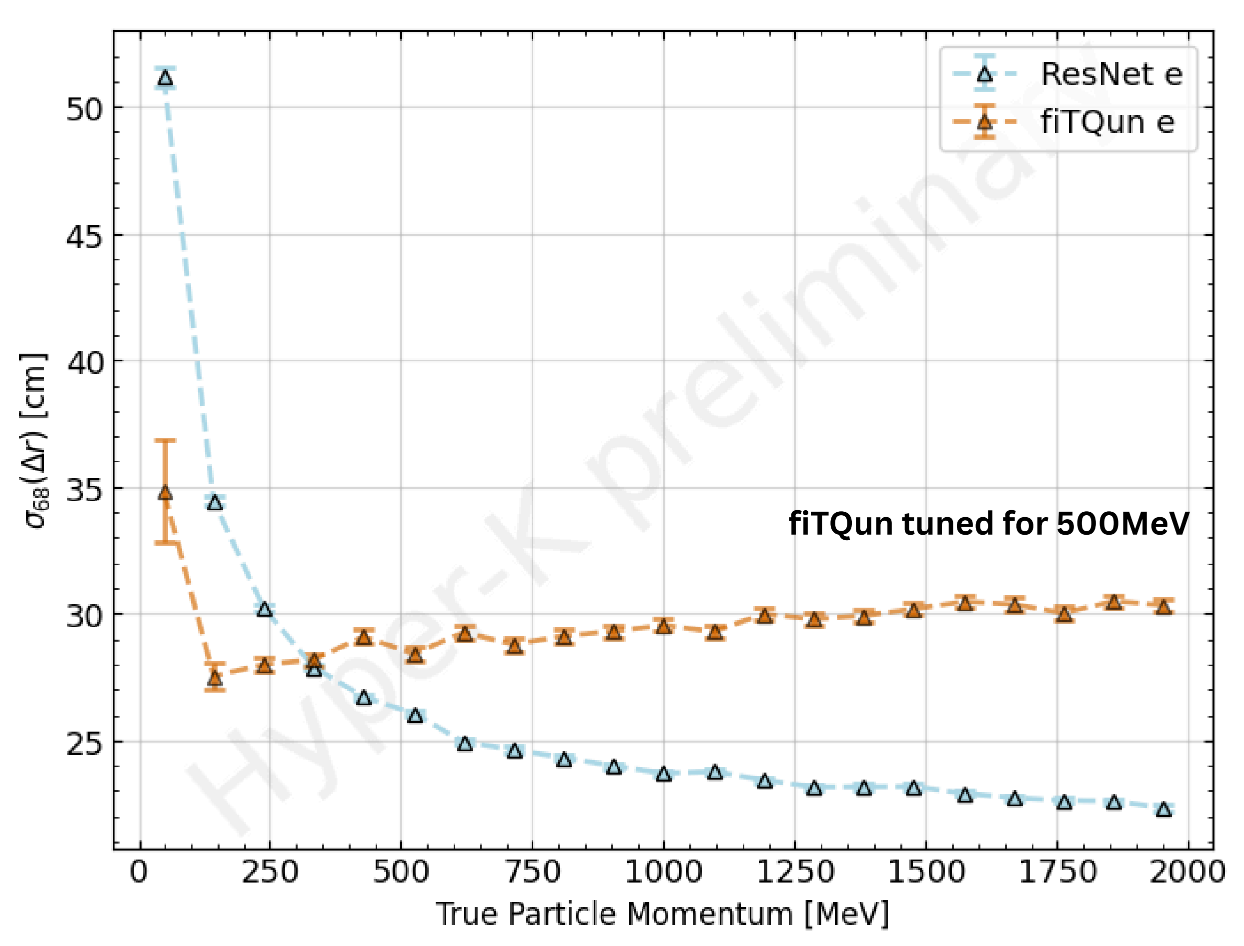
Momentum



Direction



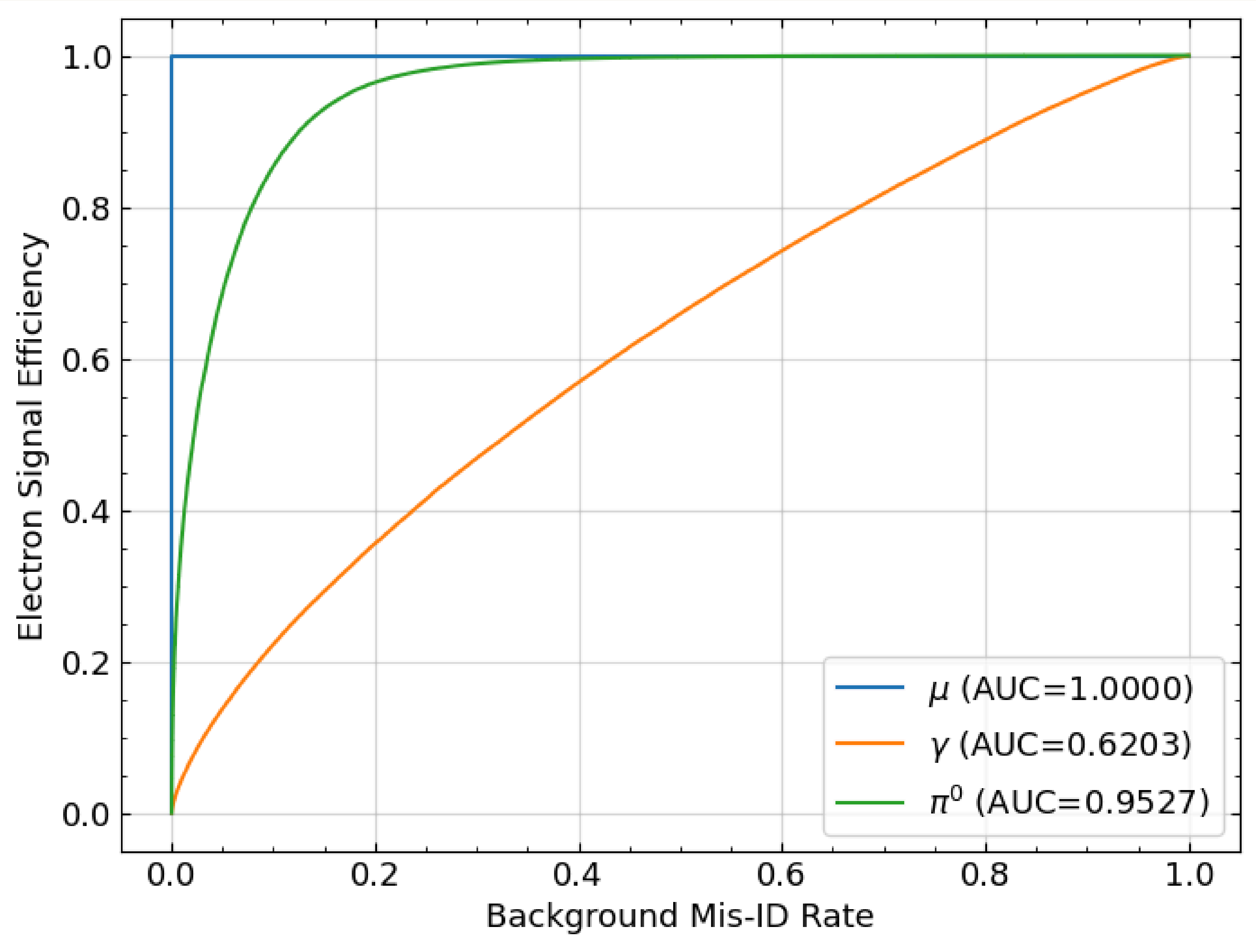
Position

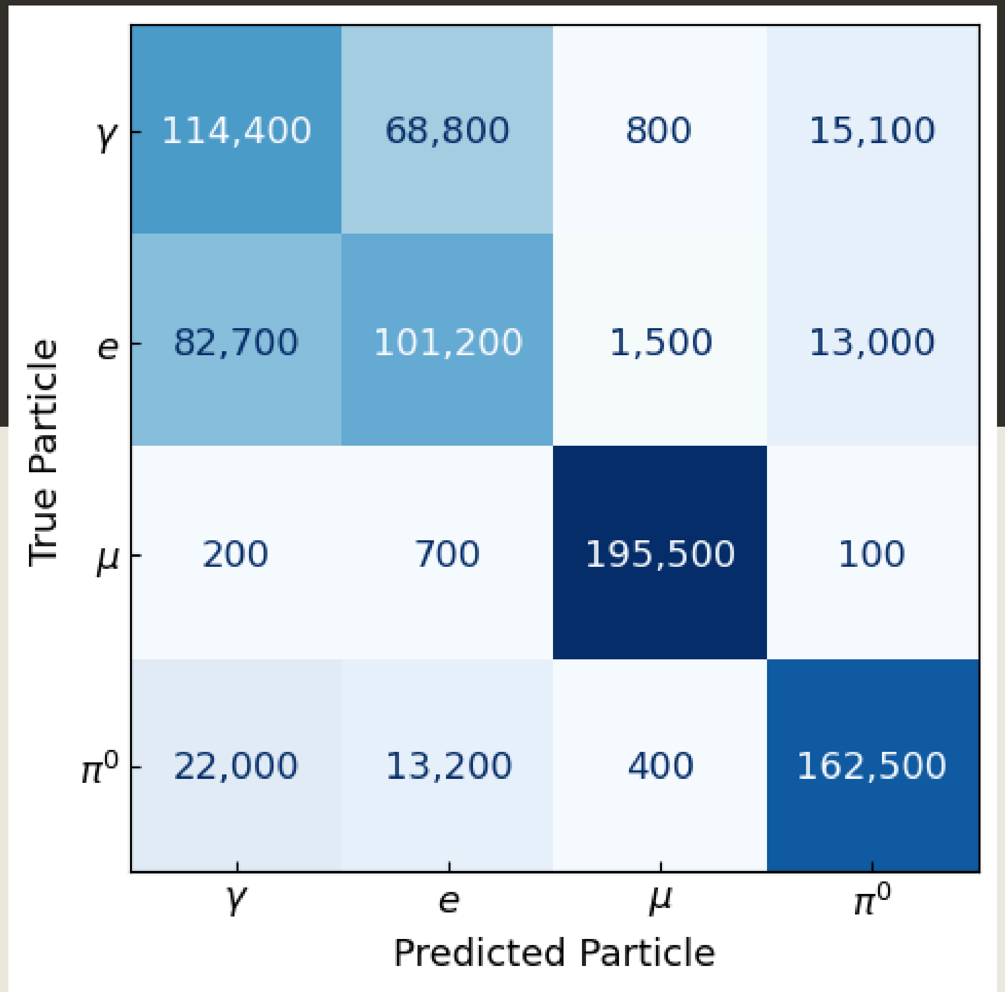


400-600MeV	ResNet	fiTQun
Average $\sigma_{68}\left(\frac{\Delta p}{p_{true}}\right)$	2.91%	5.11%
Average $\frac{\Delta p}{p_{true}}$	0.26%	1.84%
Average $\sigma_{68}\left(\Delta r\right)$	26.1 cm	28.71
Average $\sigma_{68}\left(\Delta\theta\right)$	2.33°	3.20°

CLASSIFICATION

CLASSIFICATION





Discriminators defined as:

$$D_{e/\mu} = P_e + P_\gamma + P_{\pi^0} = 1 - P_\mu$$

$$D_{e/\pi^0} = \frac{P_e + P_\gamma}{P_e + P_\gamma + P_{\pi^0}} = 1 - \frac{P_{\pi^0}}{P_e + P_\gamma + P_{\pi^0}}$$

$$D_{e/\gamma} = \frac{P_e}{P_e + P_\gamma} = 1 - \frac{P_\gamma}{P_e + P_\gamma}$$

CONCLUSIONS

- HK is a next generation neutrino detector that expects to see massively increased statistics
- Traditional reconstruction techniques are no longer fast enough, machine learning techniques are arising to fill the gap
- We've shown that CNNs like ResNet can achieve as good, if not better, reconstruction than traditional techniques while being orders of magnitude faster to compute

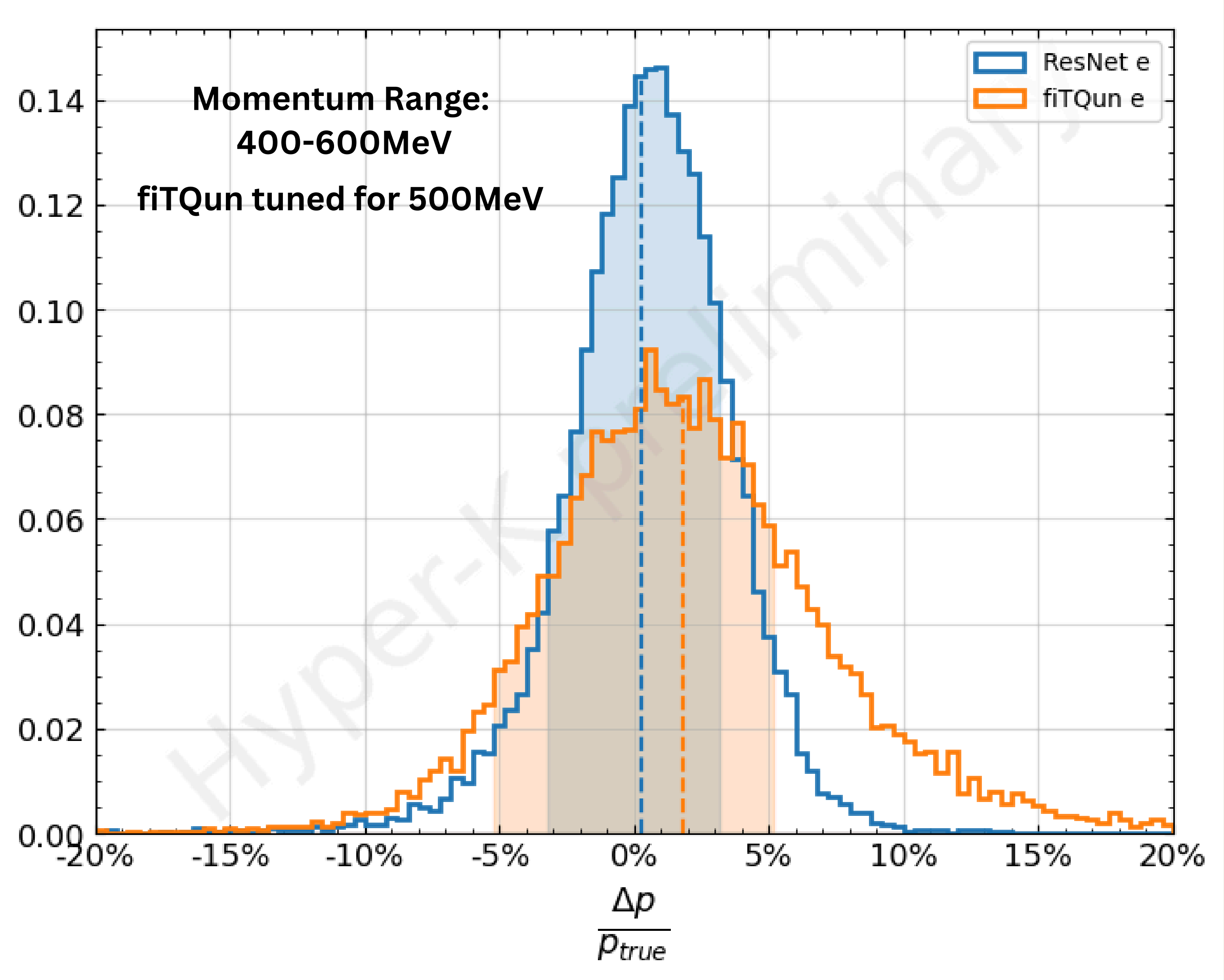
If you'd like to discuss anything with me further, I will be showing off a poster
Or feel free to email me at andrew.atta1@monash.edu

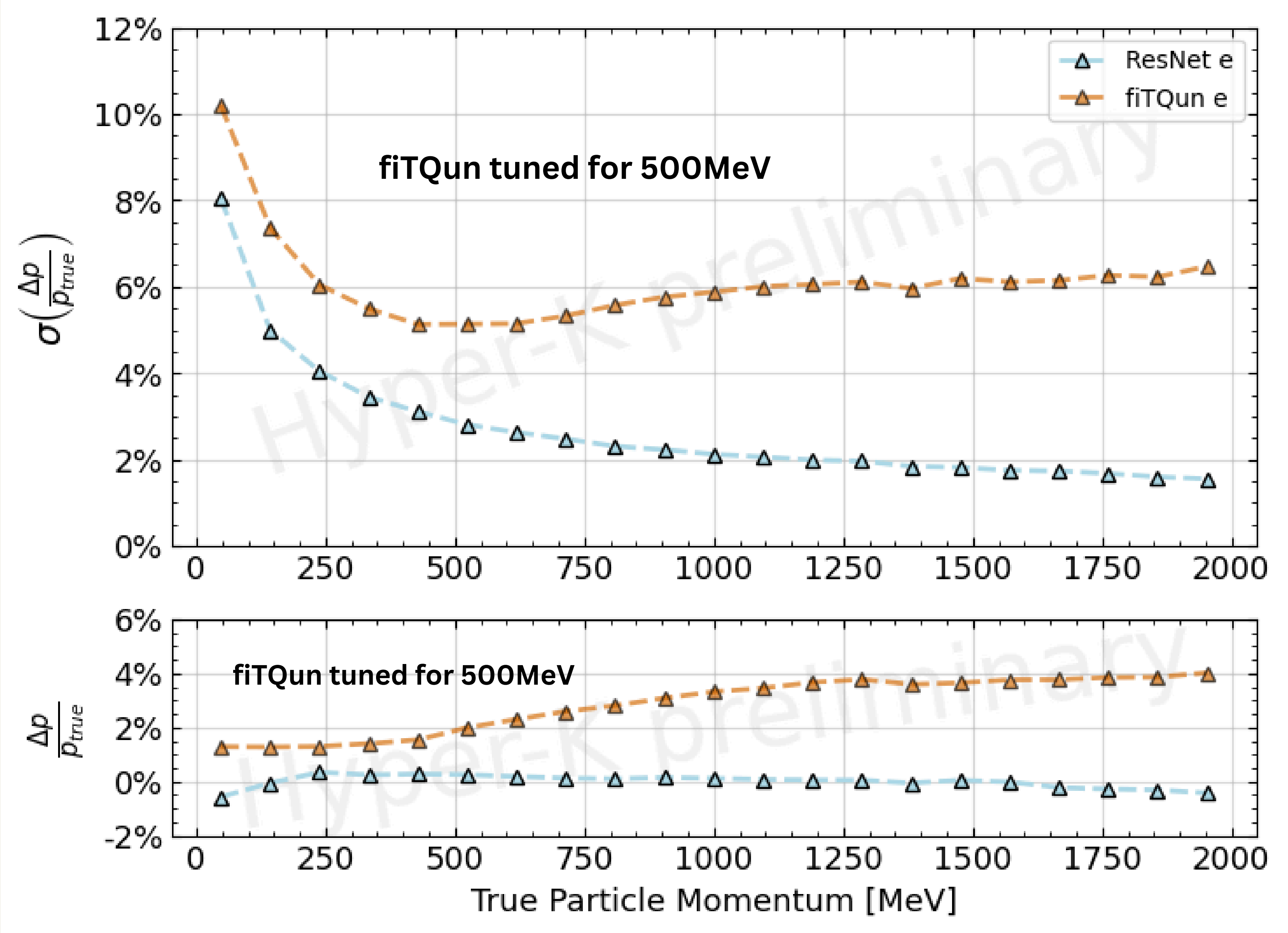
THANKS :)

APPENDIX

[redacted]

ELECTRON MOMENTUM RECONSTRUCTION



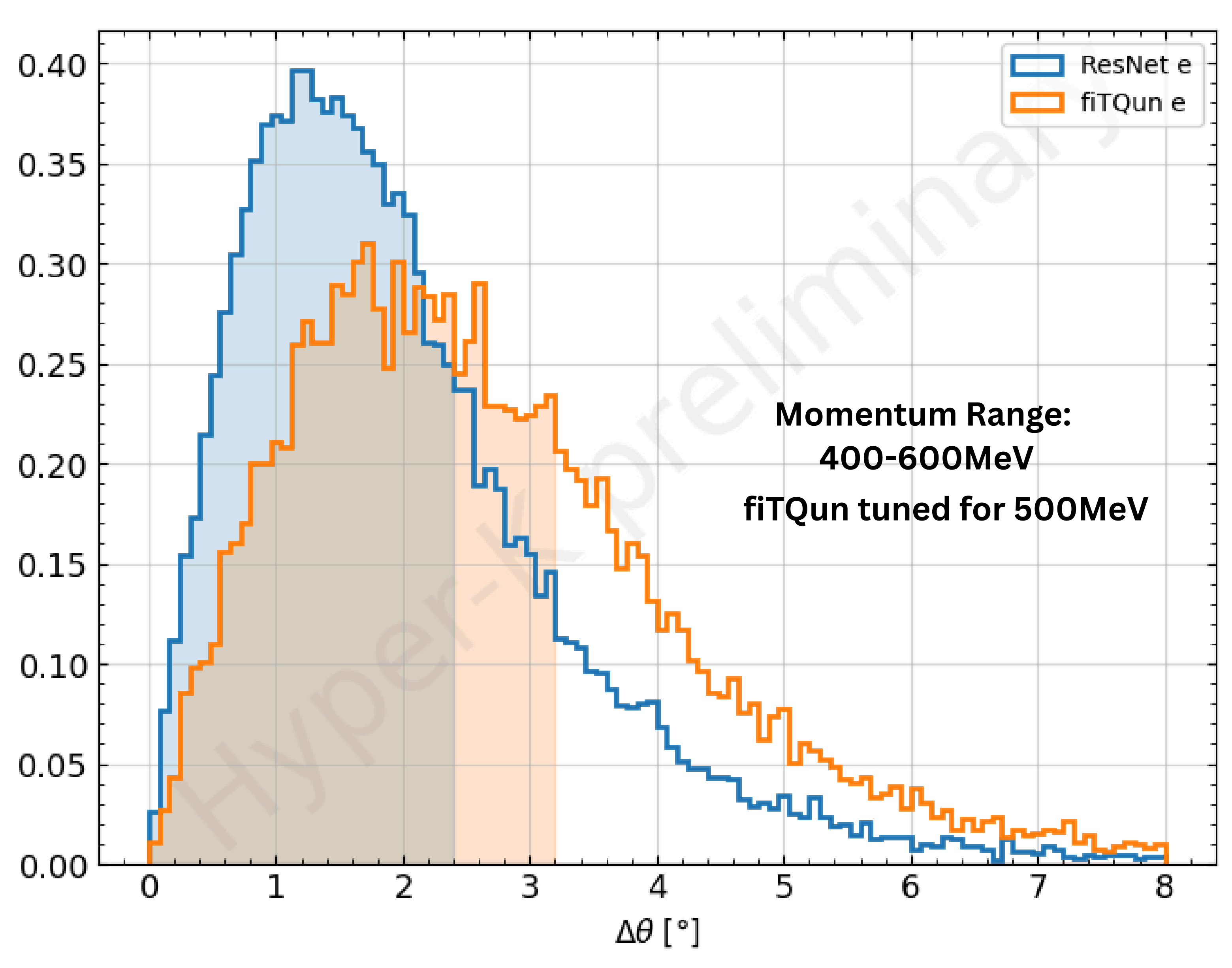


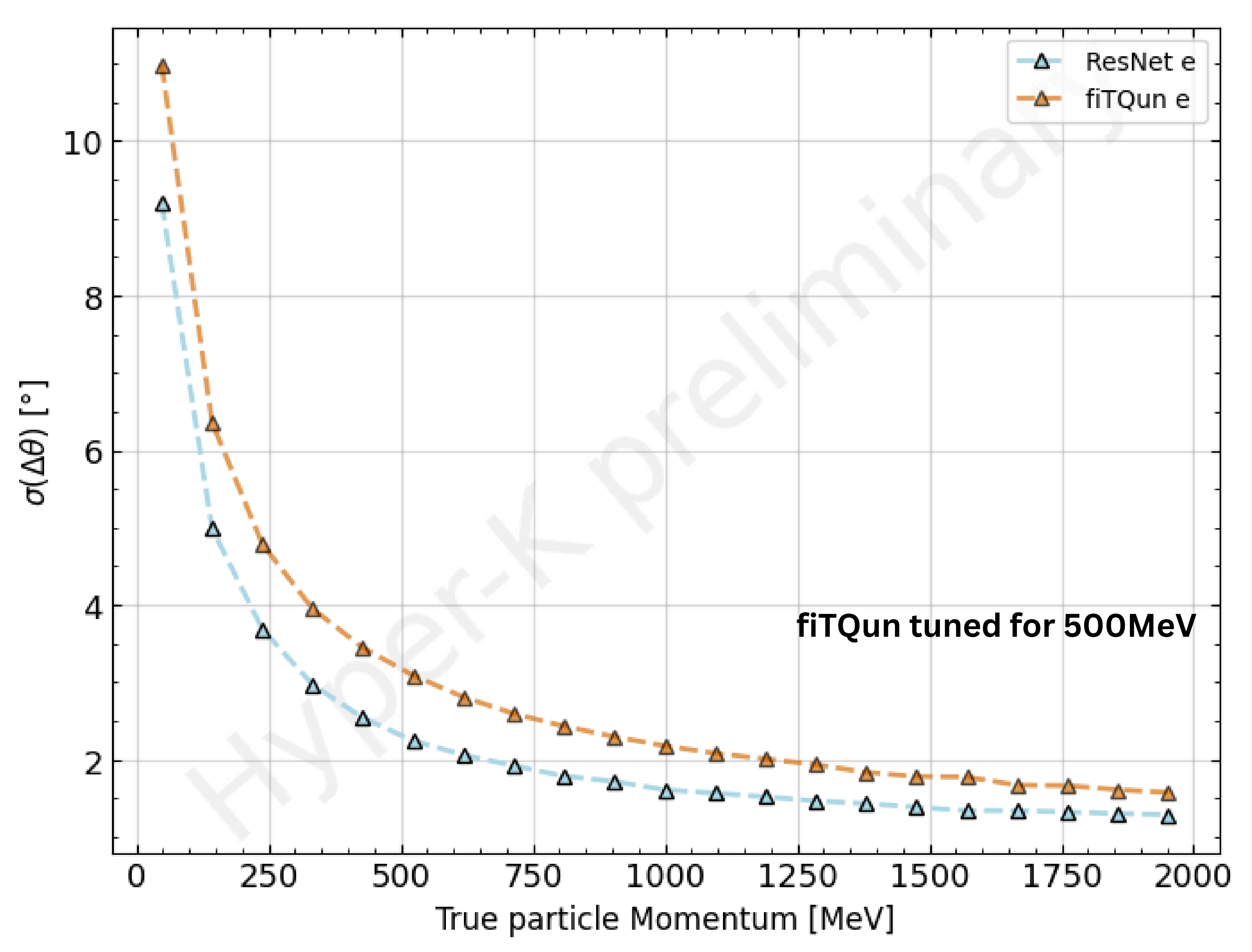
$$\frac{\Delta p}{p_{true}} = \frac{\text{True p} - \text{Predicted p}}{\text{True p}} \times 100$$

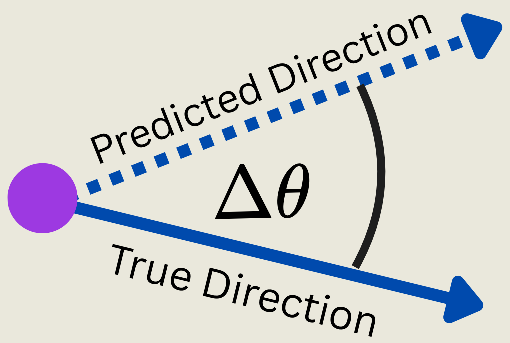
$$\sigma\left(\frac{\Delta p}{p_{true}}\right) =$$ *absolute* fractional momentum residual below which 68% of events fall under. Analogous to 1 standard deviation

	ResNet	fiTQun
Average $\sigma\left(\frac{\Delta p}{p_{true}}\right)$	2.91%	5.11%
Average $\frac{\Delta p}{p_{true}}$	0.26%	1.84%

ELECTRON DIRECTION RECONSTRUCTION



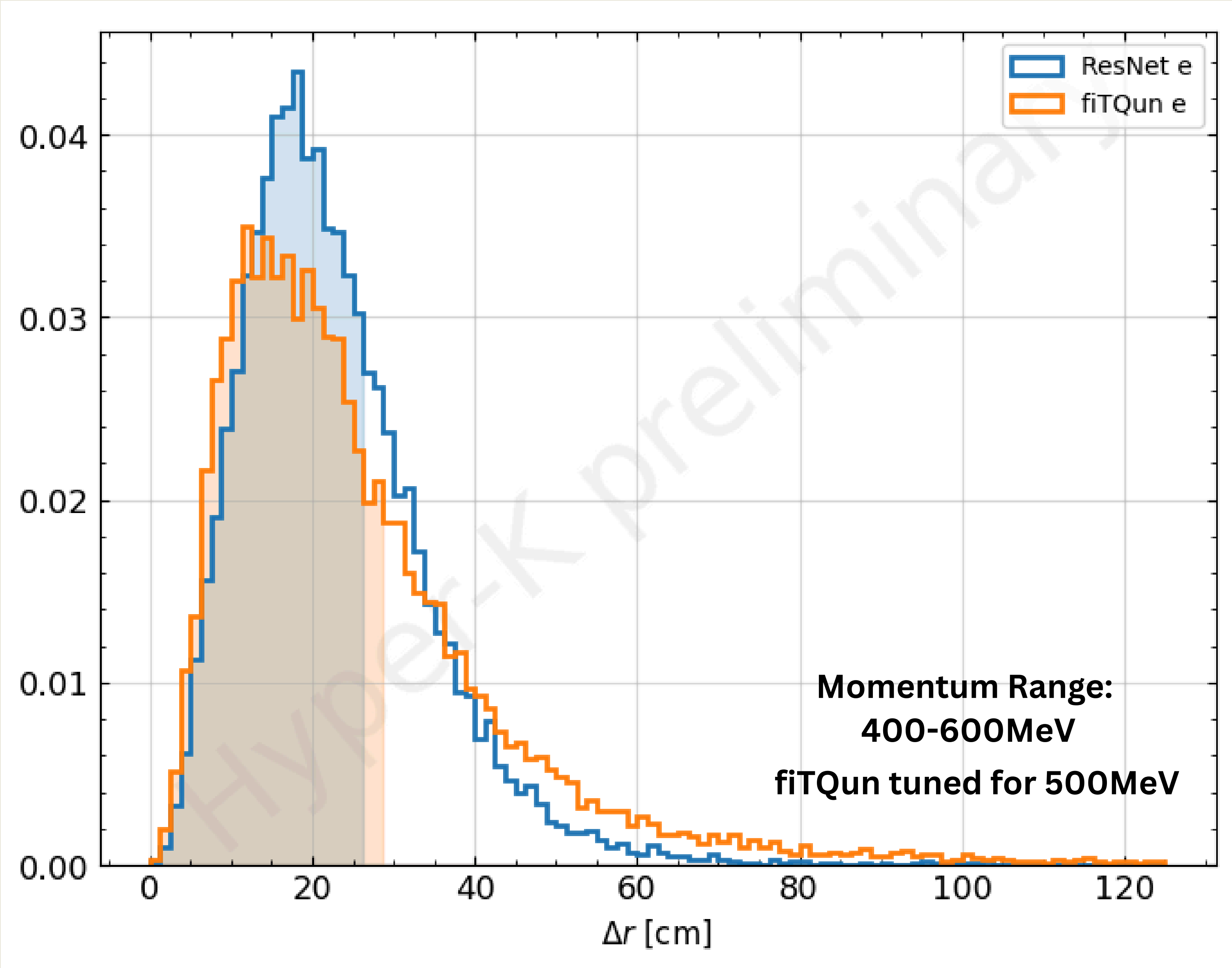


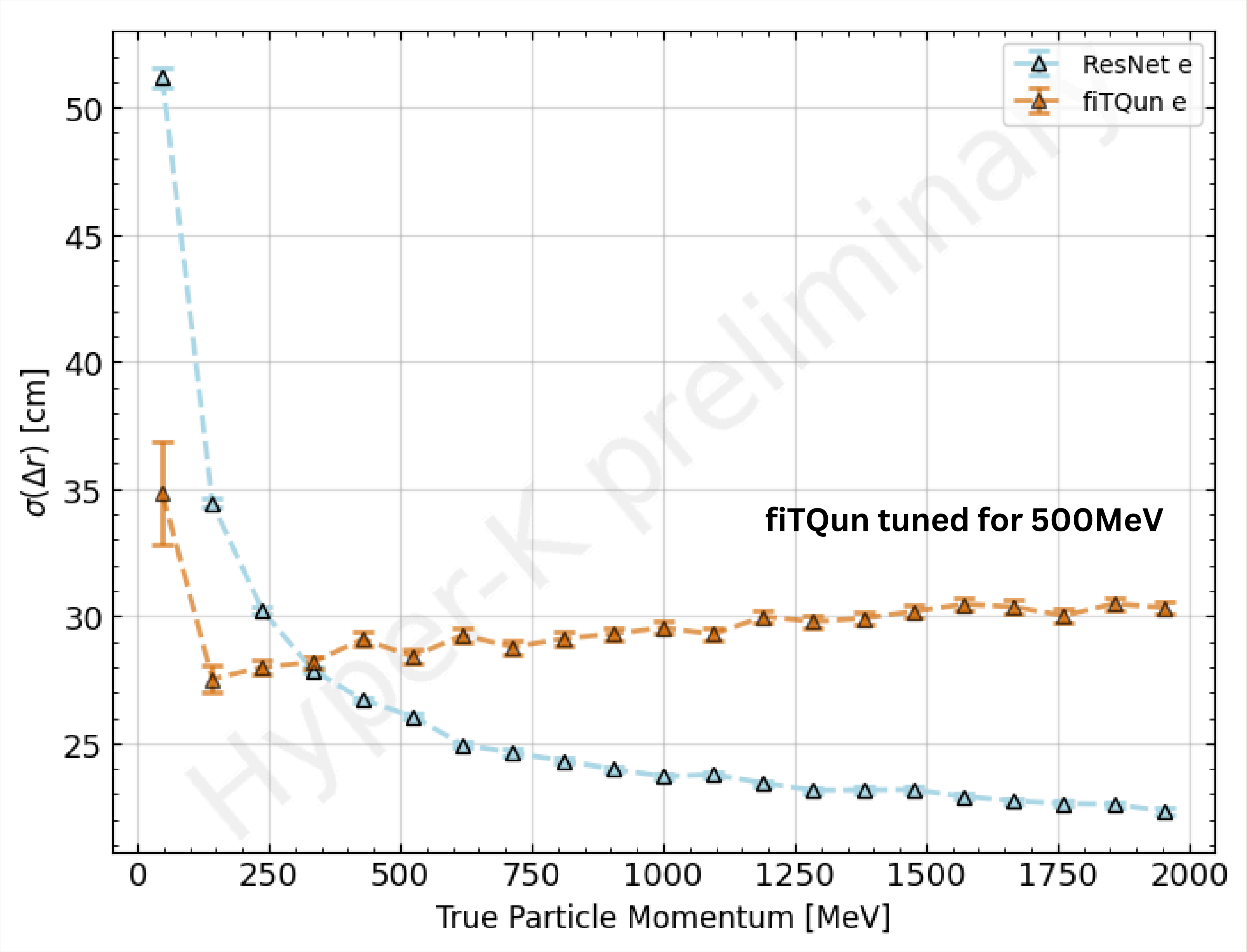


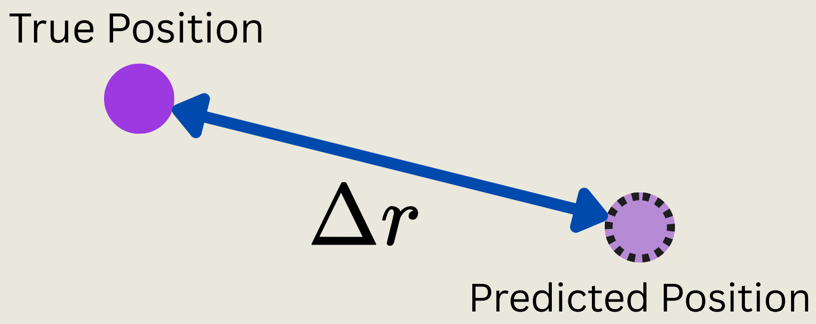
	ResNet	fiTQun
Average $\sigma\left(\Delta\theta\right)$	2.33°	3.20°

$\sigma\left(\Delta\theta\right) =$ opening angle residual below which 68% of events fall under

ELECTRON POSITION RECONSTRUCTION



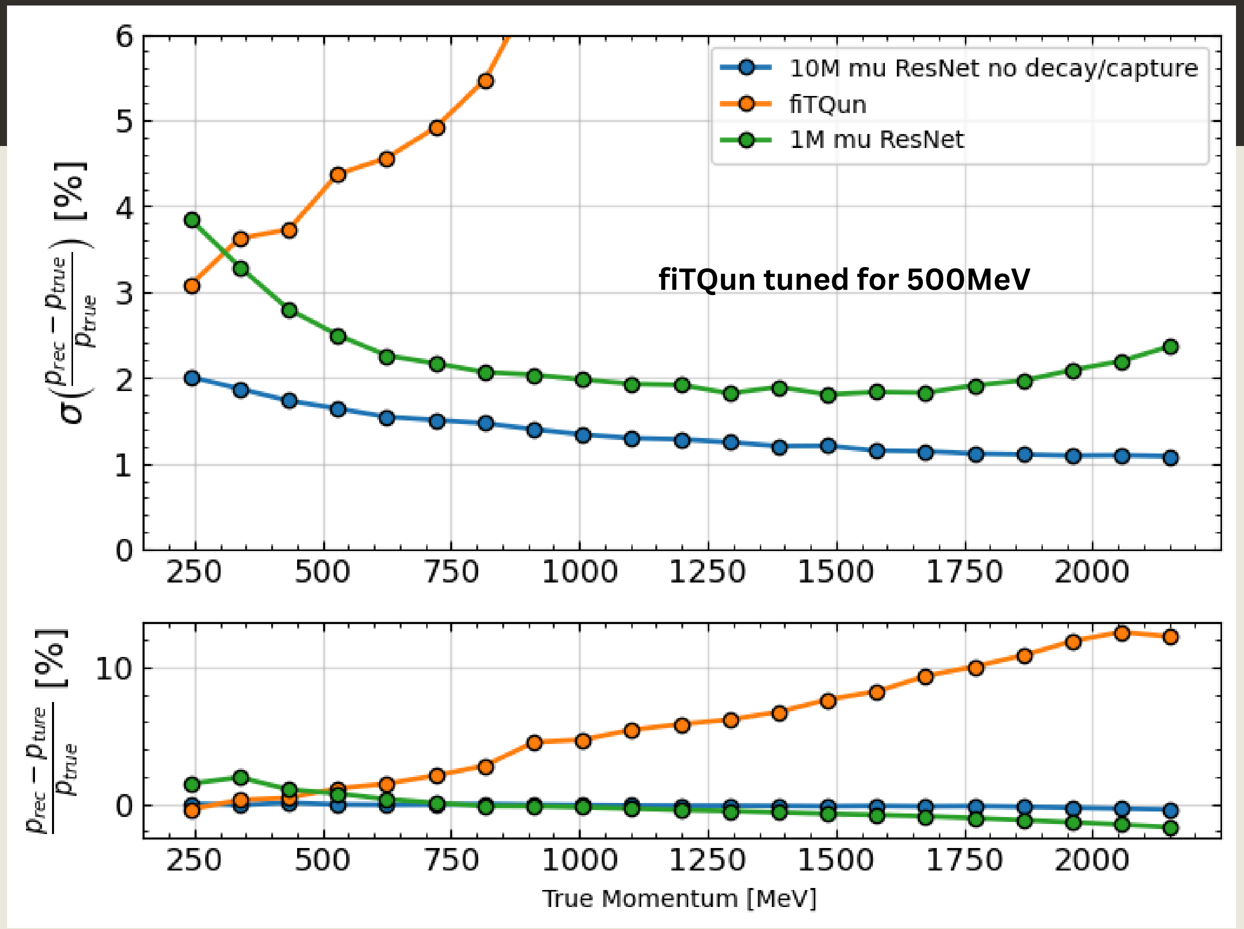


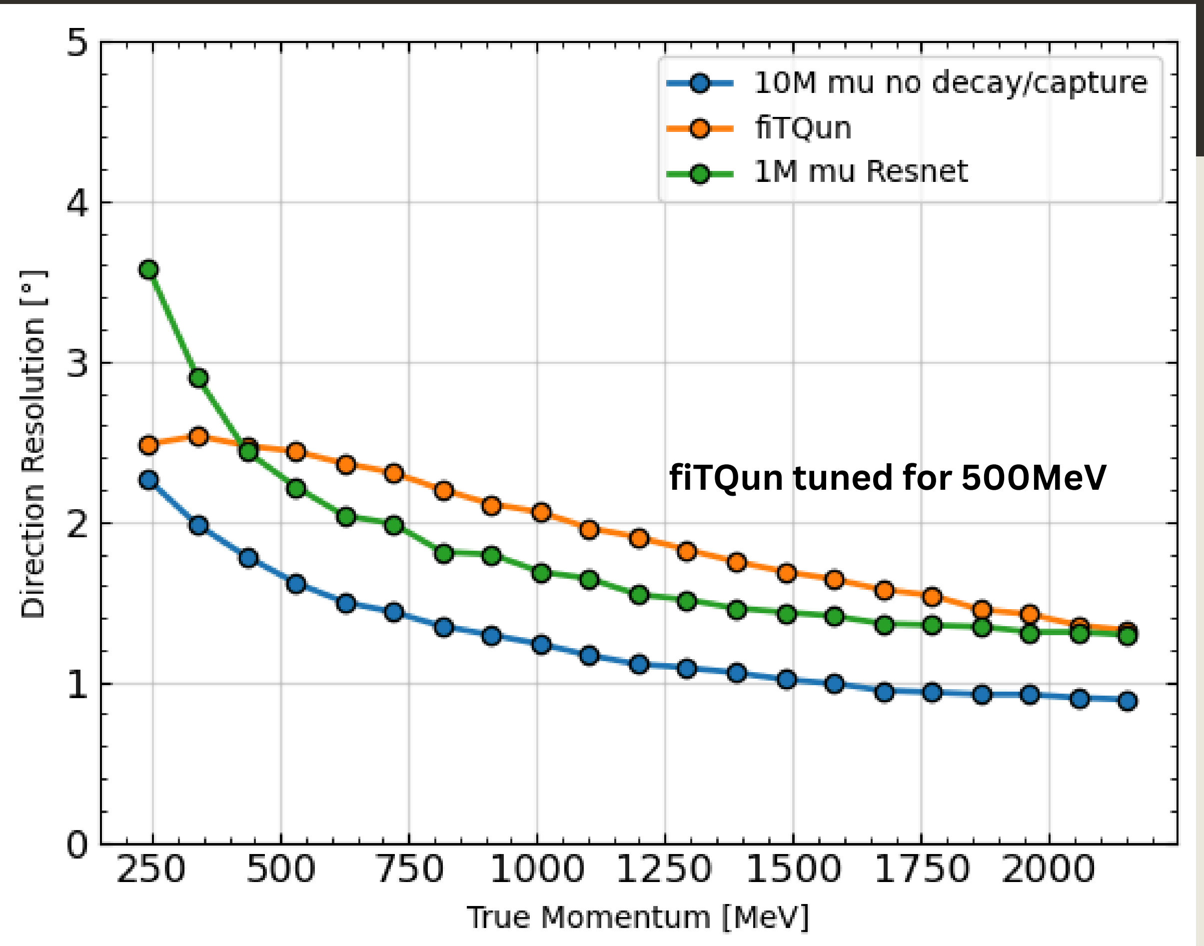


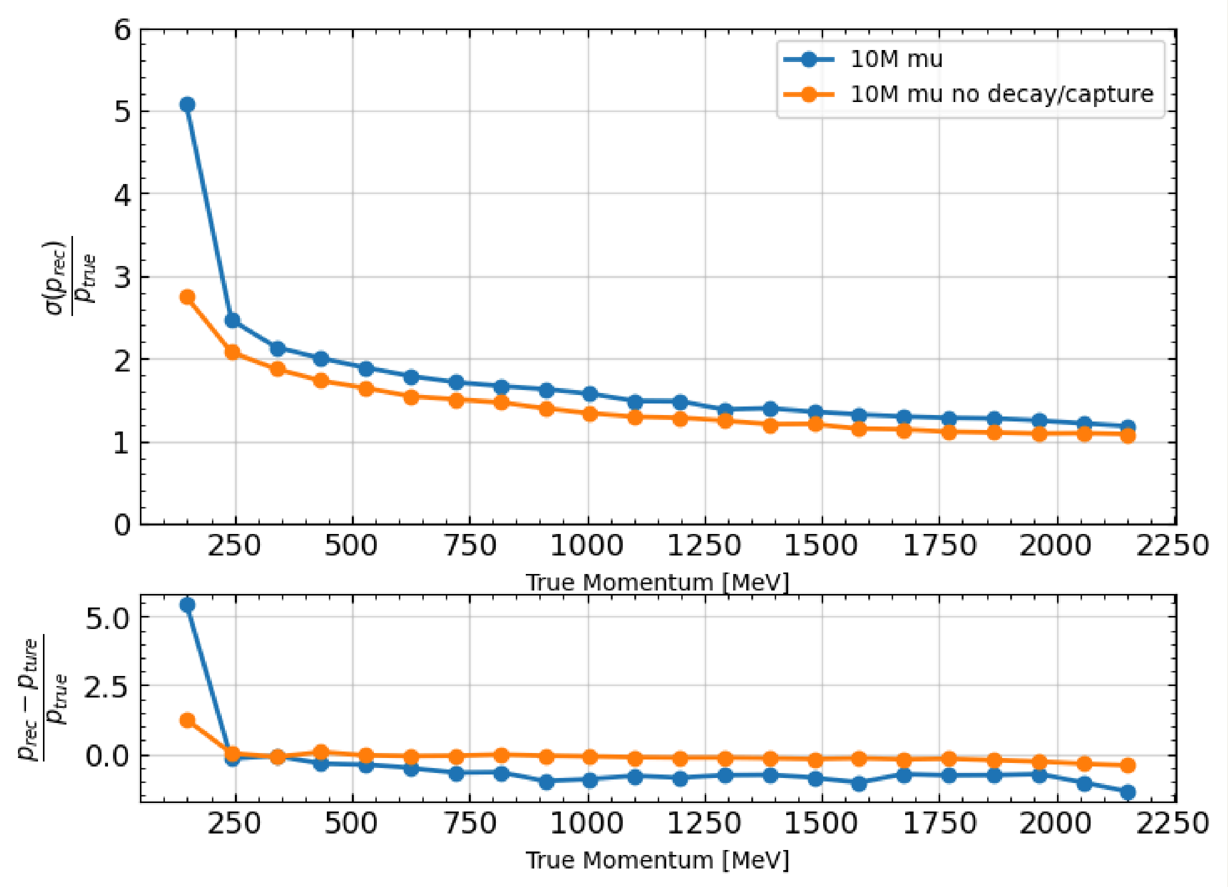
	ResNet	fiTQun
Average $\sigma\left(\Delta r\right)$	26.1 cm	28.71

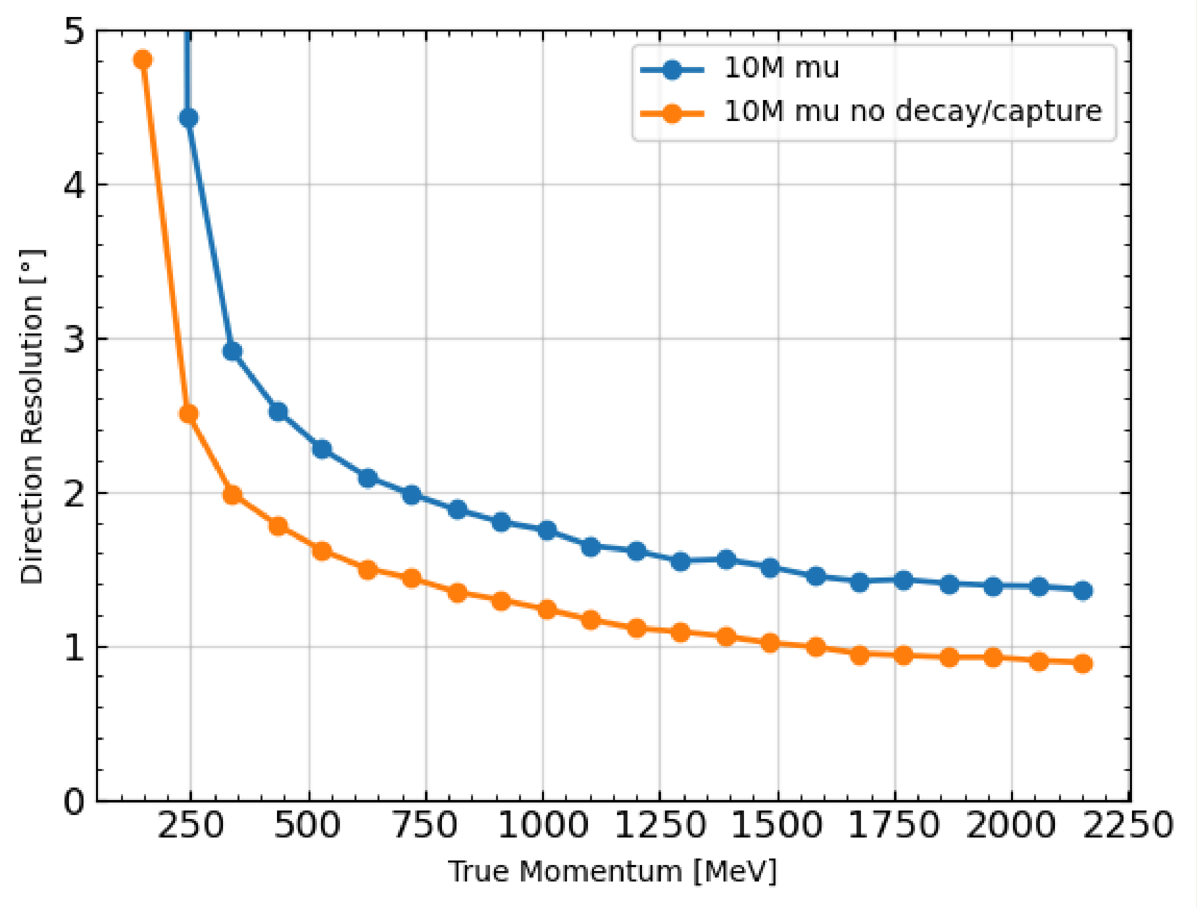
$\sigma\left(\Delta r\right) =$ 3D distance residual below which 68% of events fall under

A NOTE ON SCALABILITY









- Going from 700,000 → 9,600,00 Training events does not improve momentum or direction reconstruction much
- In fact, benefit is likely caused by removing decay muons and junk events arising from incorrect trigger timestamping in muon capture events

TRAINING NETWORKS

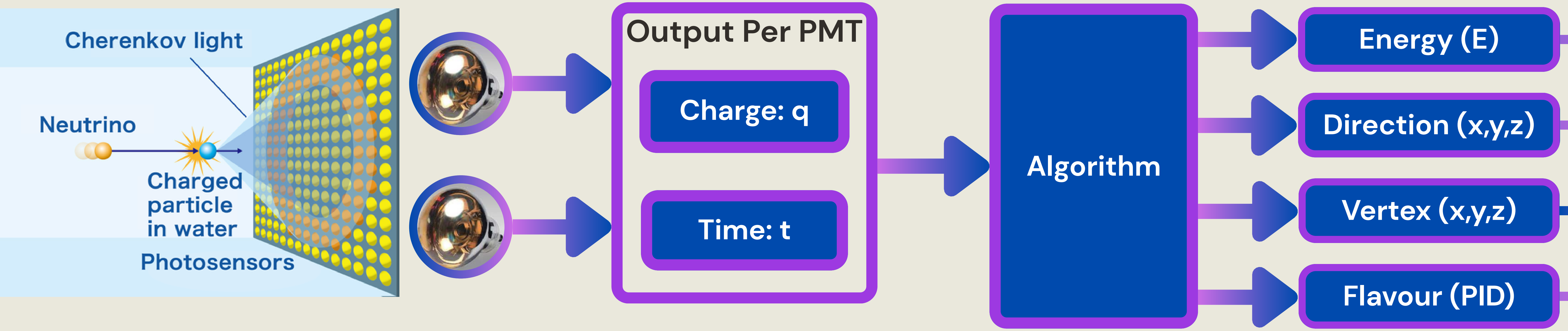
- Trained 7 networks
  - 6 Regression Networks
    - Three Muon Regressors for energy, position, direction - 10 million events
    - Three Electron Regressors for energy, position, direction - 10 million events
    - Trained for 20 epochs each, batch size=512, AdamW optimiser, learning rate=1e-3 - 1e-6 (cosine annealing LR every 2.5 epochs)
    - Trained on fully contained events only
    - Used Huber Loss function. Relative Huber loss for momentum reconstruction
  - One Classification Network
    - 4-class classifier for PID - γ vs. e vs. μ vs. π^0
    - Same config as above (besides loss function)
- All used ResNet-152 architecture using **WatChMaL**
- 190x189 input images with two channels - PMT hit time and charge
  - Input channels scale by dividing hit charges by 10 and hit times 1350 (readout window
- Trained on 4 A100 GPUs for 20 epochs for regression networks and 8 epochs for classification network

Cuts	FC and nhits>200
Energy	0–2GeV
PID	μ, e, γ, π^0
Muon Decay?	no decay
Training Split	96:2:2
Training Set	~8,569,149
Validation Set	~178,449
Testing Set	~178,595

PERFORMANCE SUMMARY

	400-600MeV	ResNet	fiTQun
μ	Average $\sigma\left(\frac{\Delta p}{p_{true}}\right)$	1.67%	4.15%
	Average $\frac{\Delta p}{p_{true}}$	0.02%	0.87%
	Average $\sigma\left(\Delta r\right)$	29.4 cm	21.7 cm
	Average $\sigma\left(\Delta \theta\right)$	1.66°	2.45°
e	Average $\sigma\left(\frac{\Delta p}{p_{true}}\right)$	2.91%	5.11%
	Average $\frac{\Delta p}{p_{true}}$	0.26%	1.84%
	Average $\sigma\left(\Delta r\right)$	26.1 cm	28.71
	Average $\sigma\left(\Delta \theta\right)$	2.33°	3.20°

EVENT RECONSTRUCTION BUT REAL THIS TIME



$$E_\nu^{rec} = \frac{2(m_n - V)E_l + m_p^2 - (m_n - V)^2 - m_l^2}{2(m_n - V - E_l + p_l \cos\theta_l)}$$

$$P_{\alpha\to\beta}(L,E) = \delta_{\alpha\beta} - 4\sum_{i>j}\mathrm{Re}\left(\mathrm{U}^*_{\alpha\mathrm{i}}\mathrm{U}_{\beta\mathrm{i}}\mathrm{U}_{\alpha\mathrm{j}}\mathrm{U}^*_{\beta\mathrm{j}}\right)\sin^2(\Delta\mathrm{m}^2_{\mathrm{ij}}\frac{\mathrm{L}}{2\mathrm{E}}) + 2\sum_{\mathrm{i}>\mathrm{j}}\mathrm{Im}\left(\mathrm{U}^*_{\alpha\mathrm{i}}\mathrm{U}_{\beta\mathrm{i}}\mathrm{U}_{\alpha\mathrm{j}}\mathrm{U}^*_{\beta\mathrm{j}}\right)\sin^2(\Delta\mathrm{m}^2_{\mathrm{ij}}\frac{\mathrm{L}}{2\mathrm{E}})$$

FULL OSCILLATION EXPRESSION

$$
\begin{aligned}
P(\nu_\mu \to \nu_e) = {} & 4c_{13}^2 s_{13}^2 s_{23}^2 \cdot \sin^2 \Delta_{31} \\
& + 8c_{13}^2 s_{12} s_{13} s_{23} (c_{12} c_{23} \cos\delta_{CP} - s_{12} s_{13} s_{23}) \cdot \cos\Delta_{32} \cdot \sin\Delta_{31} \cdot \sin\Delta_{21} \\
& \mp 8c_{13}^2 c_{12} c_{23} s_{12} s_{13} s_{23} \sin\delta_{CP} \cdot \sin\Delta_{32} \cdot \sin\Delta_{31} \cdot \sin\Delta_{21} \\
& + 4s_{12}^2 c_{13}^2 (c_{12}^2 c_{23}^2 + s_{12}^2 s_{23}^2 s_{13}^2 - 2c_{12} c_{23} s_{12} s_{23} s_{13} \cos\delta_{CP}) \cdot \sin^2\Delta_{21} \\
& \mp 8c_{13}^2 s_{13}^2 s_{23}^2 \cdot \frac{aL}{4E_\nu} (1 - 2s_{13}^2) \cdot \cos\Delta_{32} \cdot \sin\Delta_{31} \\
& \pm 8c_{13}^2 s_{13}^2 s_{23}^2 \cdot \frac{a}{\Delta m_{31}^2} (1 - 2s_{13}^2) \cdot \sin^2\Delta_{31},
\end{aligned}
$$

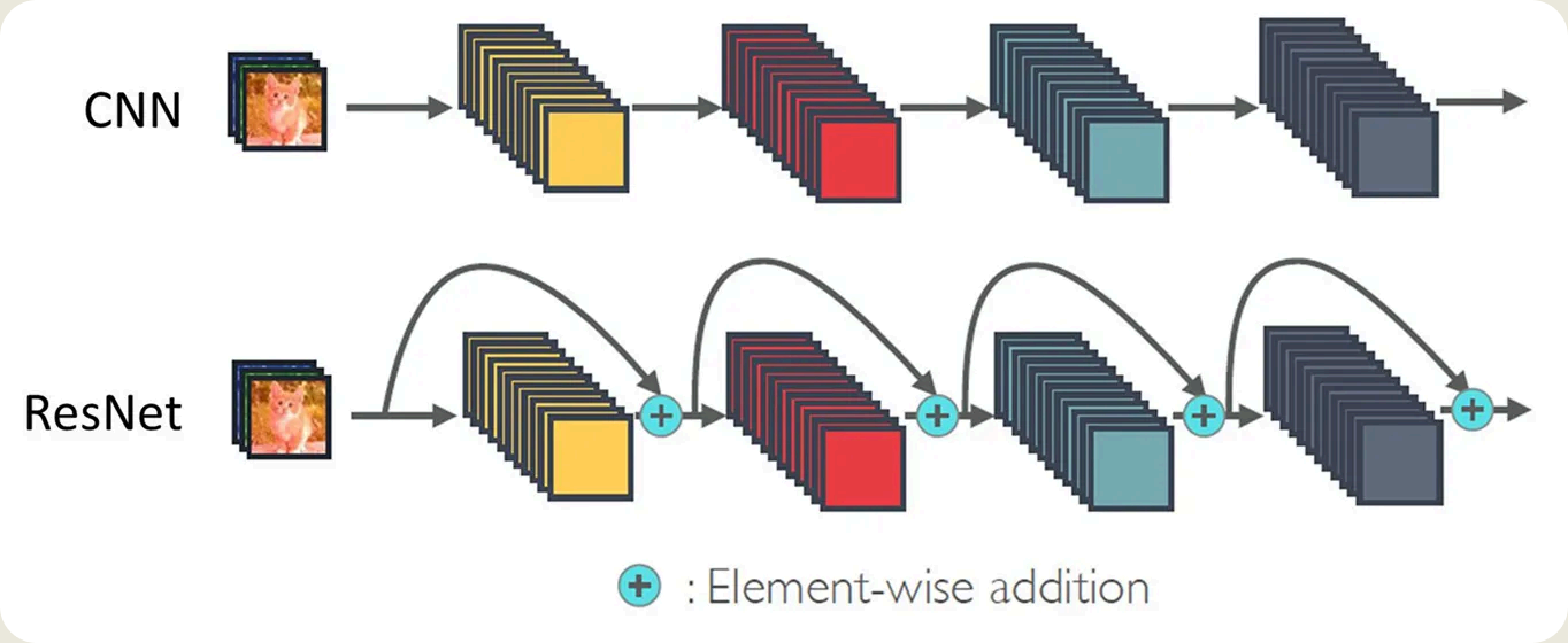
$c_{ij} = \cos\theta_{ij},\ s_{ij} = \sin\theta_{ij}$ $\Delta_{ij} = \dfrac{\Delta m_{ij}^2 L}{4E_\nu}$ $\Delta m_{ij}^2 = m_i^2 - m_j^2$ $a = 2\sqrt{2} G_F n_e E_\nu$

Fermi Coupling Constant

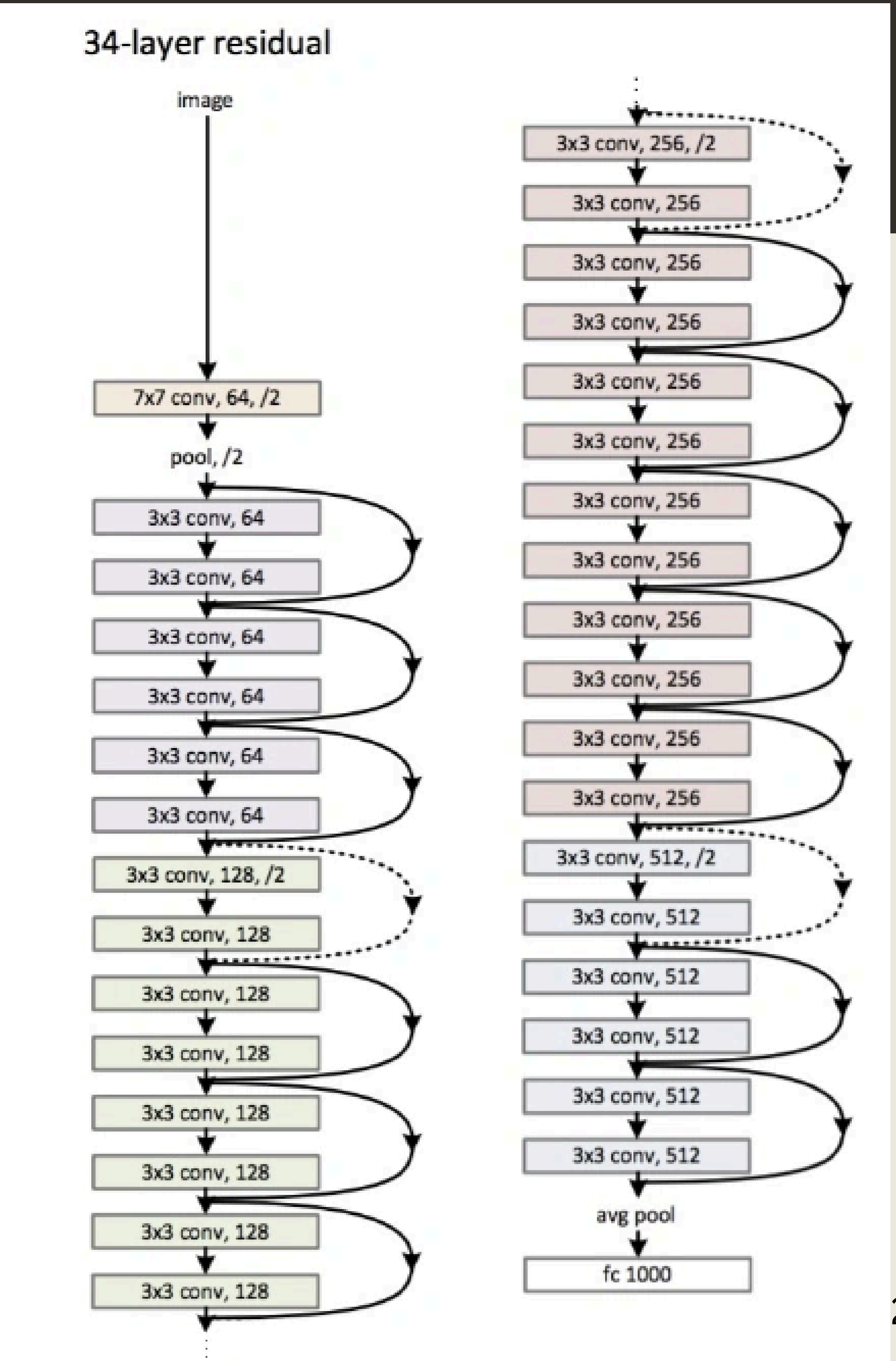
electron density

Neutrino Energy

RESNET



CNN
ResNet
: Element-wise addition



34-layer residual
image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
avg pool
fc 1000

LOSS FUNCTION

$$Huber = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}(y_i - \hat{y}_i)^2 \qquad |y_i - \hat{y}_i| \leq \delta$$

$$Huber = \frac{1}{n}\sum_{i=1}^{n}\delta\left(|y_i - \hat{y}_i| - \frac{1}{2}\delta\right) \qquad |y_i - \hat{y}_i| > \delta$$