# Semi-inclusive DIS kinematic reconstruction at the EIC with machine learning

Connor Pecar

HUGS 2022

Research supported by:
U.S. DEPARTMENT OF ENERGY | Office of Science

ATHENA

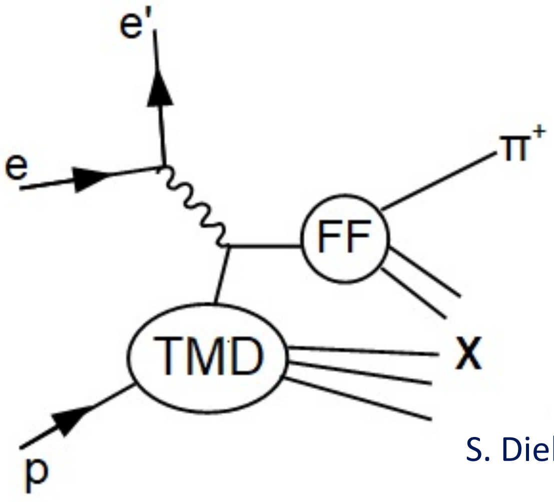
Duke UNIVERSITY

Duke

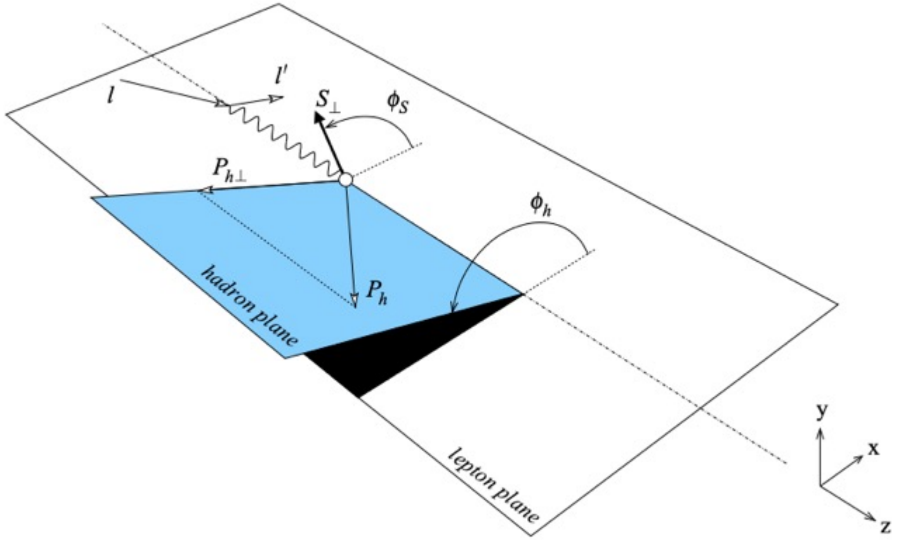
# Semi-inclusive deep inelastic scattering

- SIDIS: virtual photon exchanged with parton, measure scattered lepton and single/di-hadrons
- SIDIS cross-section gives access to parton distribution functions and fragmentation functions
  - Extra degree of freedom from hadron vital for studying TMD-PDFs
  - Also measure TMD-FFs giving information on flavor of struck quark
  - Azimuthal angle and transverse momentum defined around virtual photon axis in target COM frame
  - Cross-section a function of ($Q^2$, x, y, $p_T$, z, $M_h$)

$$z = \frac{p.p_h}{p.q}$$




S. Diehl, PRL

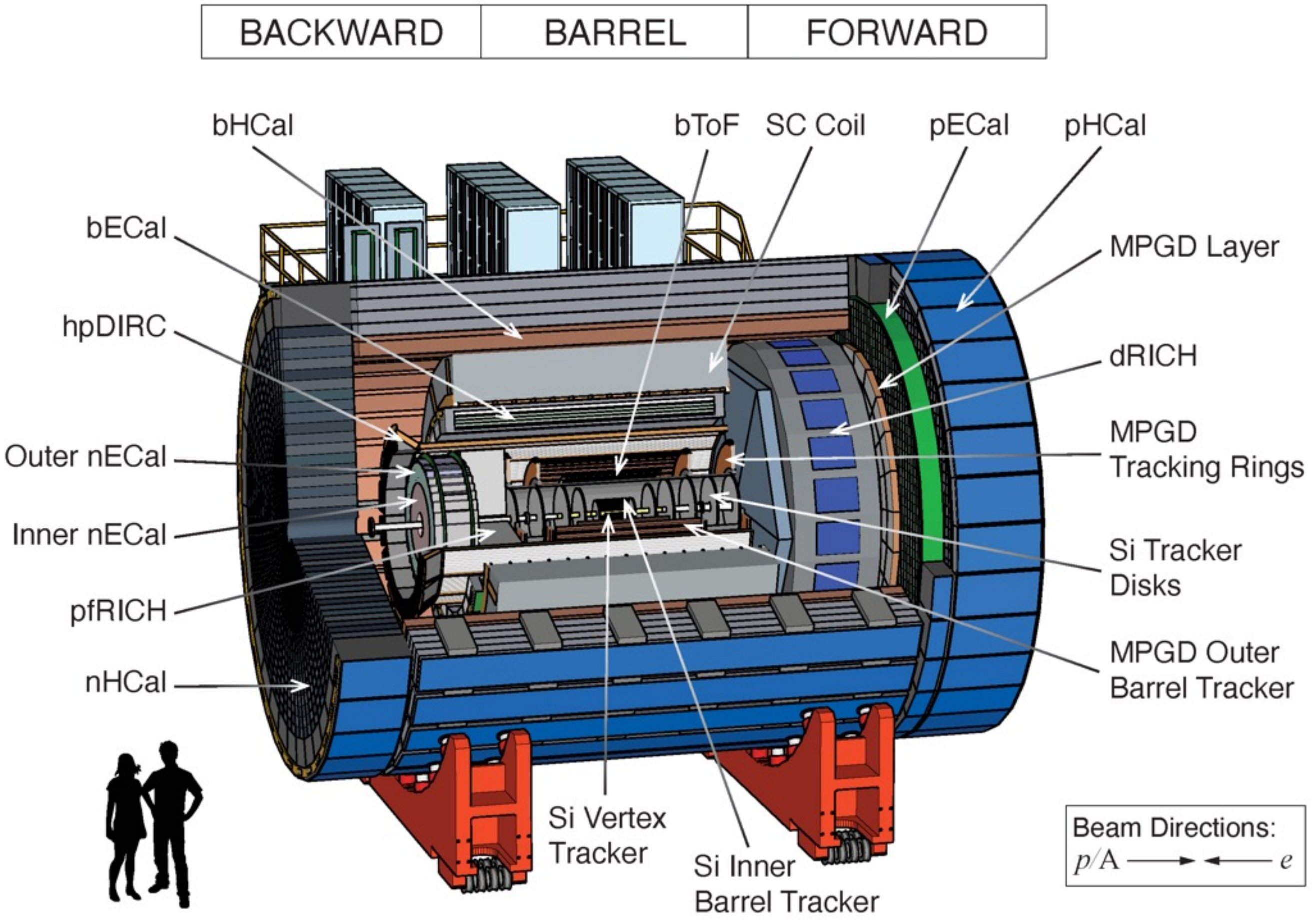



Phys.Rev. D70 (2004) 117504

Duke

# ATHENA detector

- ATHENA (A Totally Hermetic Electron Nucleon Apparatus)
- Precise tracking and PID coverage in wide momentum range – great potential for SIDIS
- Studies shown in this presentation use ATHENA full simulation as "Detector 1" full simulation is still in development




Duke

# SIDIS kinematic reconstruction

- SIDIS variables: reliant on reconstruction of virtual photon four-momentum, typically determined using

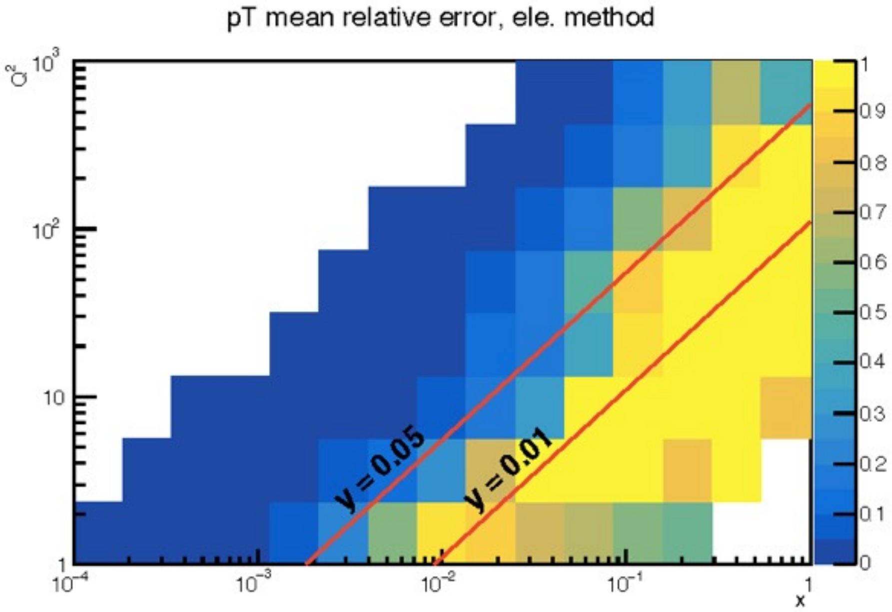
$$q = l - l'$$

- Reliable for larger y (“inelasticity”), but begins to fail for y < ~0.05

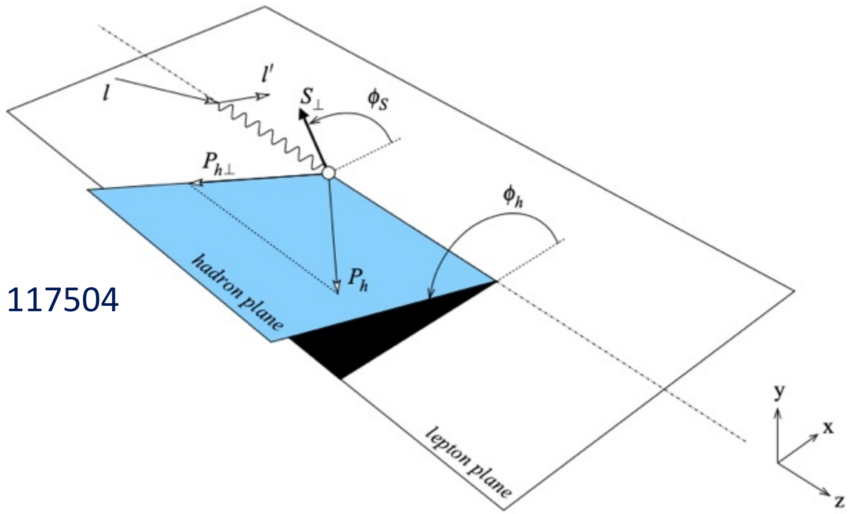
Low-y: region of interest for TMDs, valence quarks, evolution studies, overlap with Jlab coverage

- **To utilize full EIC kinematic reach for SIDIS studies, need improved methods to determine SIDIS variables**
- CC – would require first method without electron

ATHENA full simulation:




$$y = \frac{Q^2}{xs}$$




Phys.Rev. D70 (2004) 117504

Duke

# Reconstruction with hadronic final state

- Through conservation of momentum and energy, **hadronic final state (HFS) should also contain enough information to constrain virtual photon four-momentum**
- Method used in EIC YR and detector proposals to reconstruct virtual photon using hadronic final state (HFS)
  - x and y components - summed HFS momentum
  - z and t components - solved for algebraically using

$$y = \frac{p.q}{p.l} \qquad Q^2 = -q^2$$

and DIS variables from any DIS reconstruction method

$$
\begin{array}{rll}
i) & \textit{Leptonic variables} & q \equiv q_l = k_2 - k_1, \quad y_l = p_1.(k_1 - k_2)/p_1.k_1 \\
ii) & \textit{Hadronic variables } [81] & q \equiv q_h = p_2 - p_1, \quad y_l = p_1.(p_2 - p_1)/p_1.k_1 \\
iii) & \textit{Jacquet-Blondel variables } [82] & Q^2_{JB} = (\vec{p}_{2,\perp})^2/(1 - y_{JB}), \quad y_{JB} = \Sigma/(2E(k_1)) \\
 & & \Sigma = \sum_h (E_h - p_{h,z}) \\
iv) & \textit{Mixed variables } [81] & q = q_l, y_m = y_{JB} \\
v) & \textit{Double angle method } [83] & Q^2_{DA} = \dfrac{4E(k_2)^2 \cos^2(\theta(k_2)/2)}{\sin^2(\theta(k_2)/2) + \sin(\theta(k_2)/2)\cos(\theta(k_2)/2)\tan(\theta(p_2)/2)}, \\
 & & y_{DA} = 1 - \dfrac{\sin(\theta(k_2)/2)}{\sin(\theta(k_2)/2) + \cos(\theta(k_2)/2)\tan(\theta(p_2)/2)},
\end{array}
$$

Prog. Part. Nucl. Phys. 2013, Blümlein

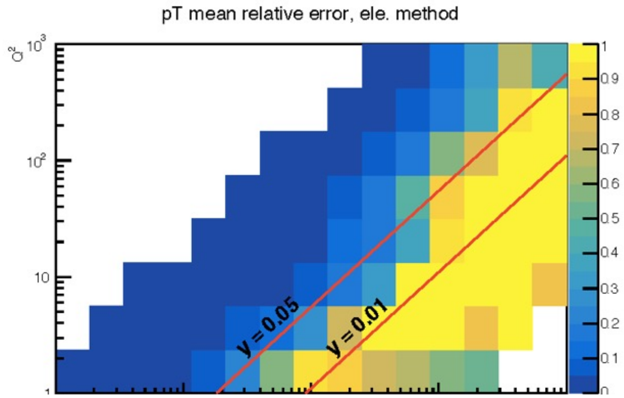
Duke

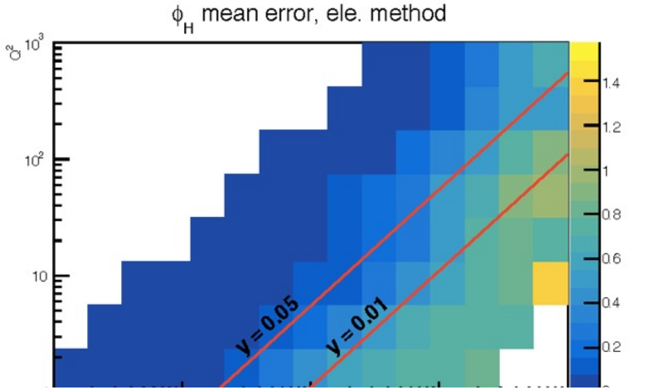
# Resolutions:

ATHENA full simulation, 10x275, pi+, z > 0.2

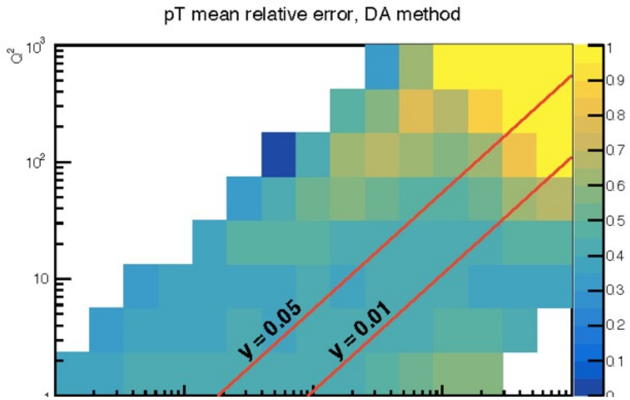
$p_T$ | $\varphi_h$

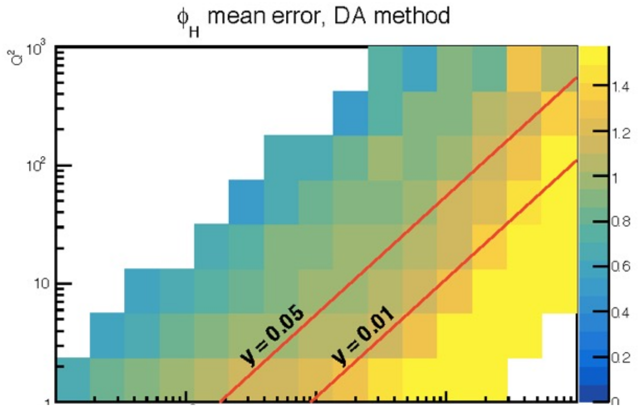
Electron method



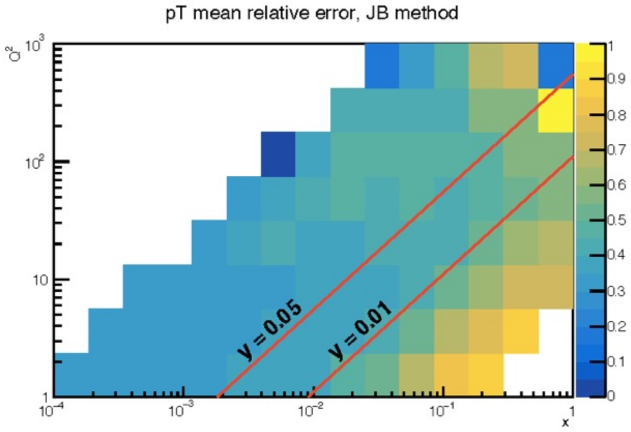



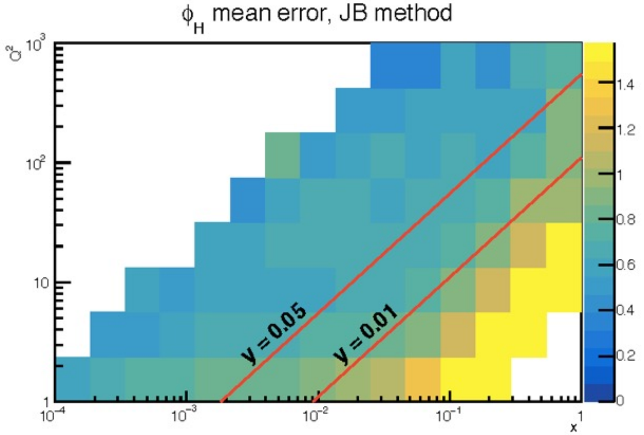

DA method (Electron and HFS)




JB method (Only HFS)



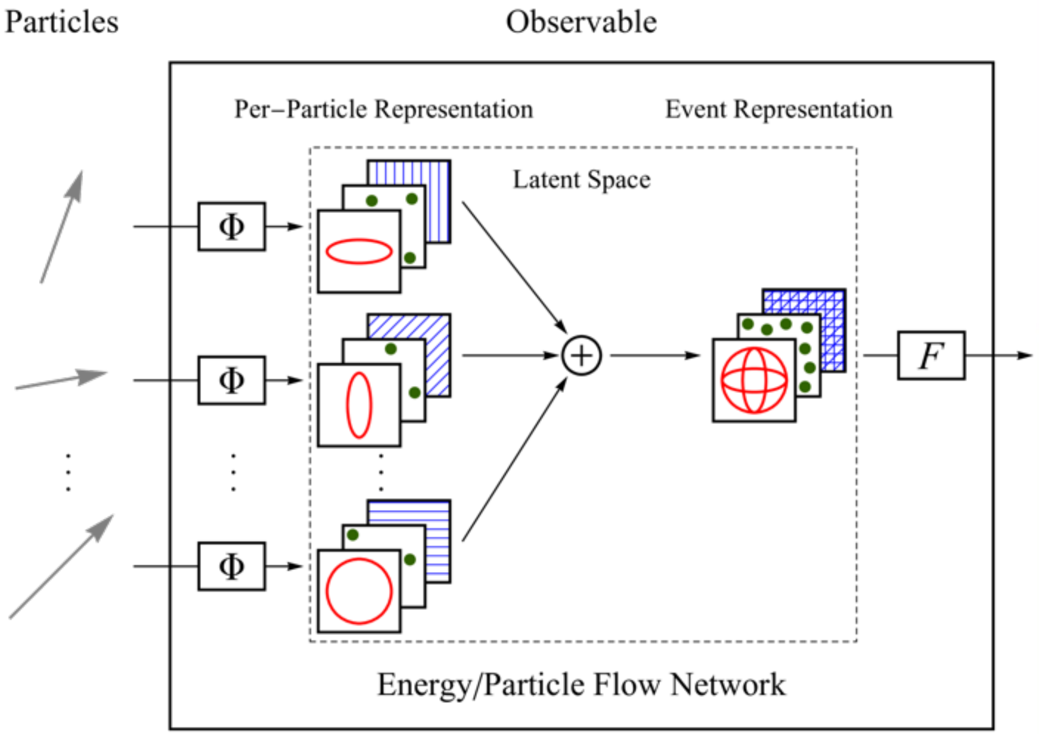

Angular resolution still poor at low-y with all methods

Duke

# Machine learning reconstruction

- Based on hybrid HFS-electron SIDIS reconstruction, using ML to combine information from both to reconstruct q
  - ML models used for DIS reconstruction have been shown to be able to naturally account for radiative effects
    - (arXiv:2108.11638 Diefenthaler, Farhat, Verbytskyi, Xu, and NIM-A 1025 (2022) 166164, Arratia, Britzger, Long, Nachman)
- Currently utilizing graph-like neural network architectures designed for tasks like jet classification
  - Particle flow networks:
    - Accepts unordered set of particles
    - Particle features -> input to layers Φ
      - Summed over to created latent space of ℓ variables
    - Global features of event concatenated with latent space variables
    - Latent space variables and global features fed to layers F, produce final output

Particle flow networks (PFN) developed by Komiske et al., (JHEP 01 (2019) 121, Komiske, Metodiev, Thaler):




Duke

# ML SIDIS model and training

- Model combining electron and HFS:
  - Particle features for PFN: momentum, energy, $\eta$, $\varphi$ in lab frame
  - Event-wide features: electron four momentum, DIS variables from JB, DA, electron methods
    - DIS variables will eventually be replaced with final reconstructed Q2 and x likely using another ML method, but in this study statistics for training were limited
  - Target: MC virtual photon four-momentum in lab frame
- Training sample: ATHENA full simulation
  - Version of dd4hep ATHENA full sim. used for detector proposal
    - Still some features missing, e.g. proper scattered electron ID
  - HFS at the level of reconstructed particles
  - 10 GeV electron beam, 275 GeV proton beam, crossing angle -25 mrad
  - Trained on 3 million events with $Q^2 > 1$ GeV$^2$, 2 million with $Q^2 > 10$ GeV$^2$
  - 1 million $Q^2 > 1$ GeV$^2$ events for validation

Duke

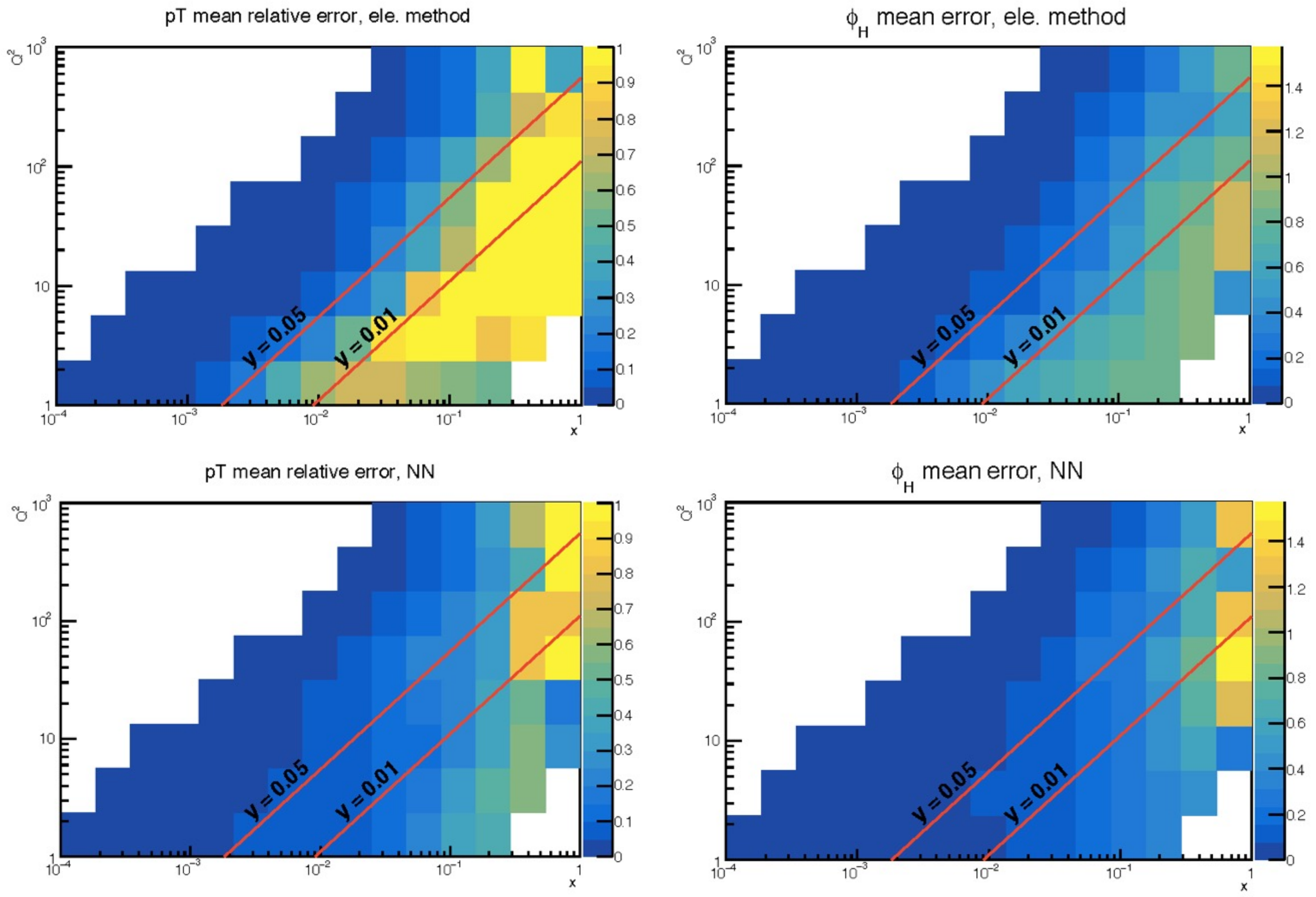
ATHENA full simulation,
10x275, pi+, z > 0.2

$$\frac{p_T - p_{T,true}}{p_{T,true}}$$

$$\phi_H - \phi_{H,true}$$

Electron method

Neural network




PFN able to correct electron method in almost all of x-Q2

Duke

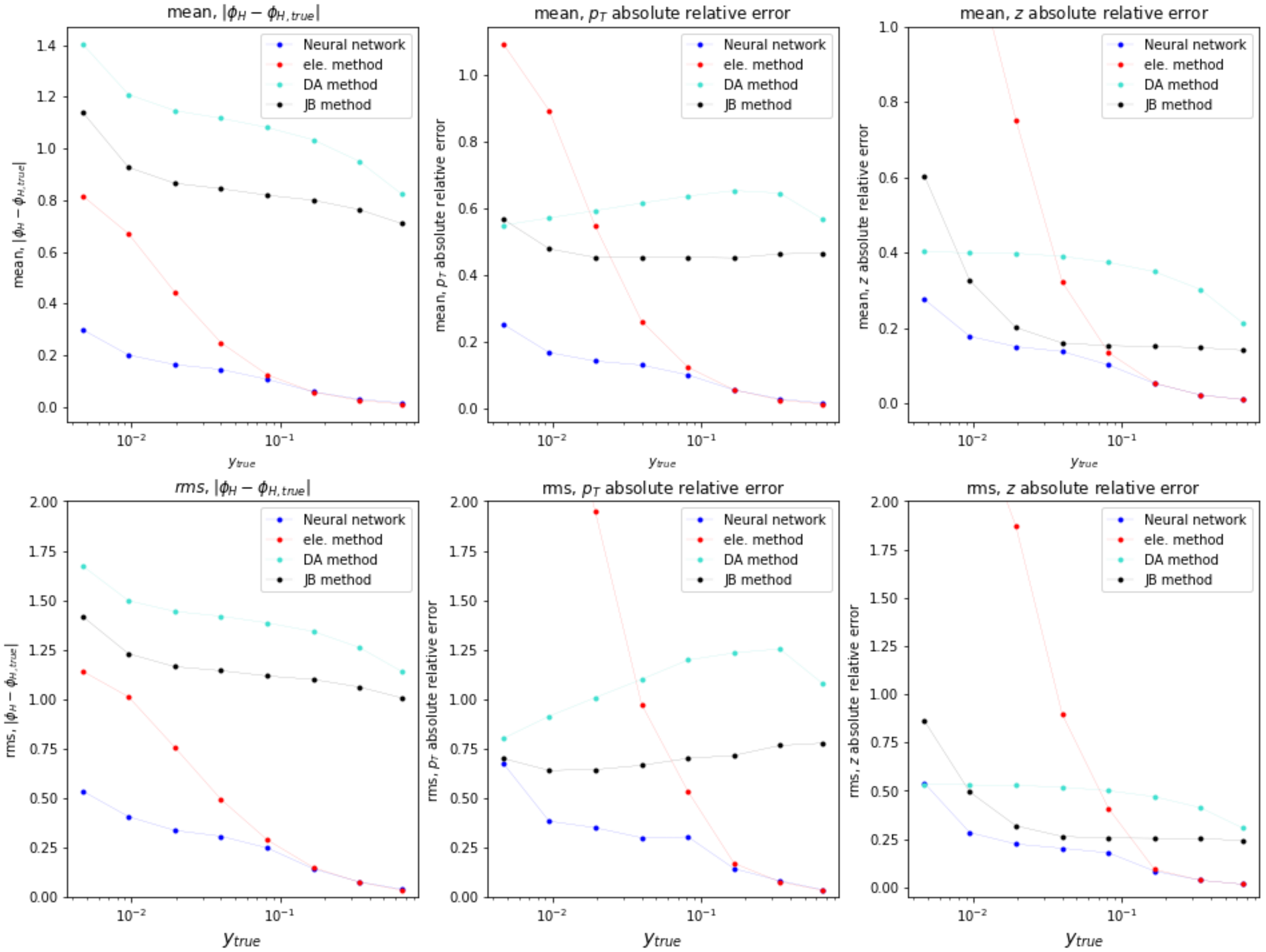
ATHENA full simulation, 10x275, pi+, z > 0.2

Mean:

- Comparison with other HFS/hybrid methods vs $y_{true}$
- NN clearly best performance for low y, and at least equaling electron method for large y

RMS:

$|\phi_h - \phi_{h,true}|$ $\quad \frac{|p_T - p_{T,true}|}{p_{T,true}}$ $\quad \frac{|z - z_{true}|}{z_{true}}$




Duke

# Summary

- Projections for the ATHENA and ECCE detectors demonstrate the exciting capabilities and kinematic coverage of the EIC for SIDIS measurements
- The electron method fails for $y < 0.05$, but can be improved using the hadronic final state and DIS variables to reconstruct virtual photon axis
- The use of particle flow networks to combine hadronic final state and scattered electron information surpasses existing reconstruction methods for all of x-Q2 and $p_T$
- Next steps in reconstruction:
    - Currently working on virtual photon reconstruction with an architecture which could also learn correlations between HFS particles (graph neural network), as well as exploring other deep learning approaches
    - Method will need to be tested with better implementation of radiative effects and continuously as the Detector 1 simulation is developed

Research supported by:
U.S. DEPARTMENT OF ENERGY | Office of Science

Duke

# Extra

Duke

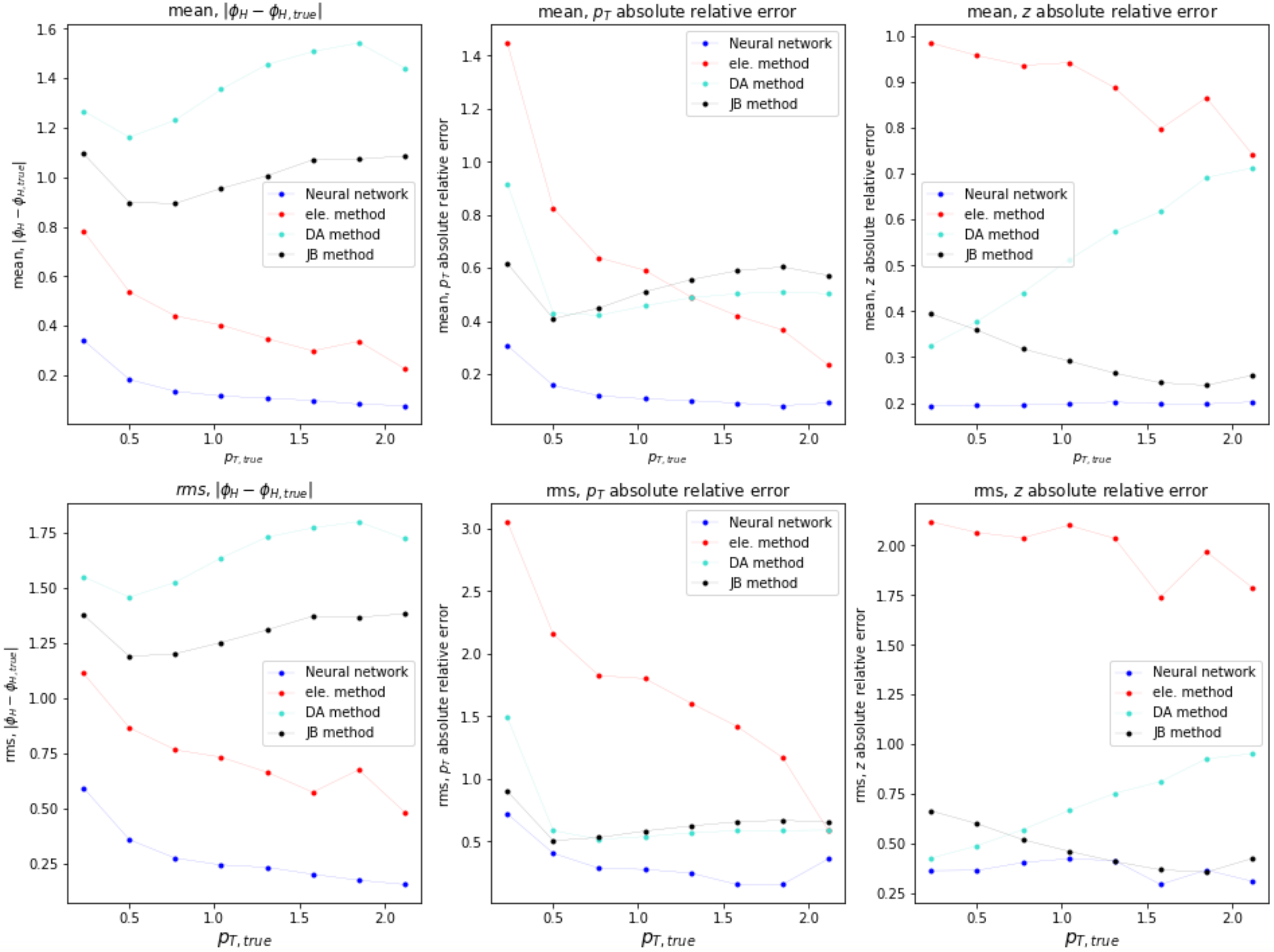
ATHENA full simulation,
10x275, pi+, z > 0.2

Mean:

RMS:

- Comparison with other HFS methods vs true $p_T$
- NN again clearly outperforming other methods for all $p_T$ (integrated over all y)

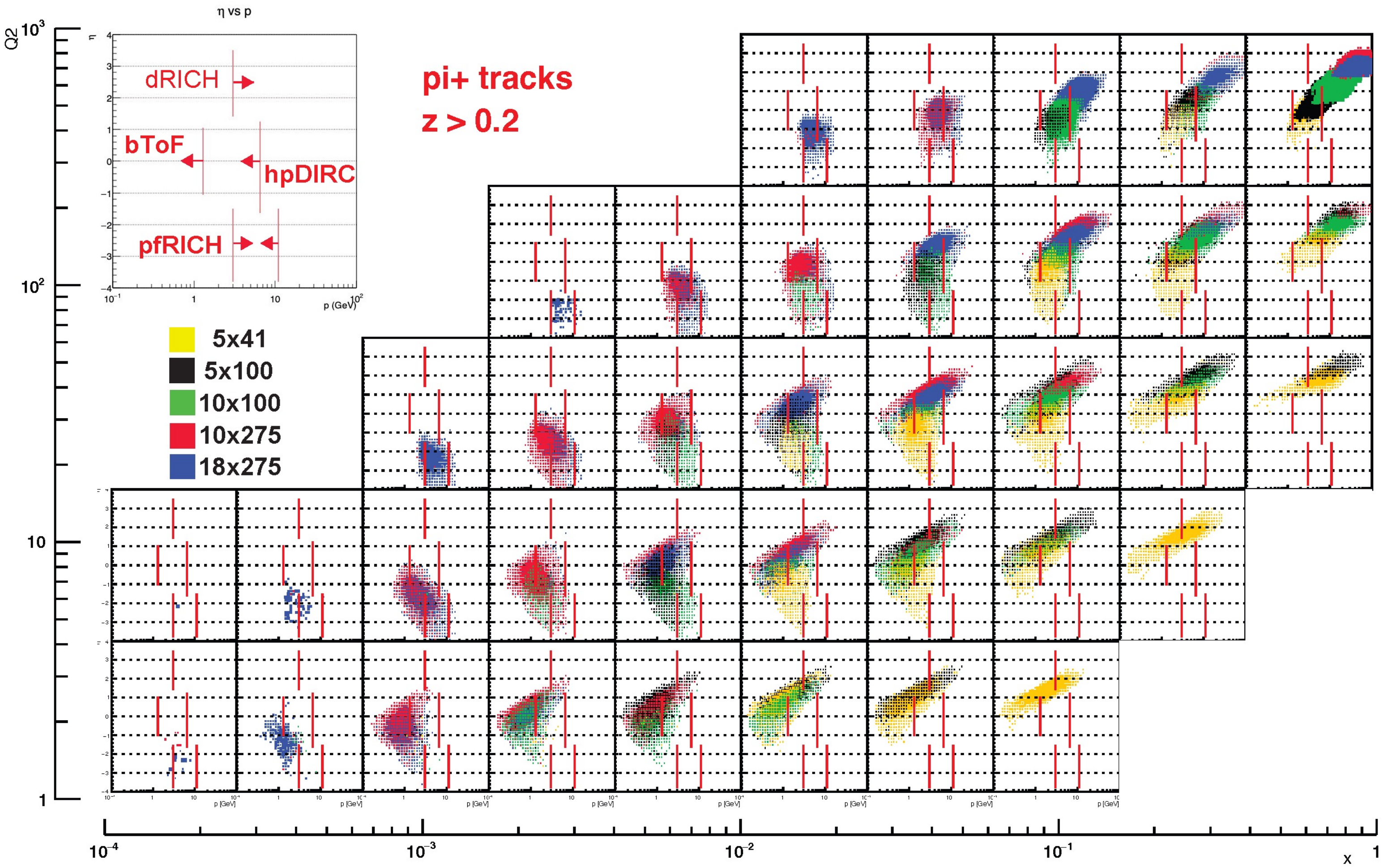
$$|\phi_h - \phi_{h,true}| \qquad \frac{|p_T - p_{T,true}|}{p_{T,true}} \qquad \frac{|z - z_{true}|}{z_{true}}$$




Duke

# SIDIS observables and coverage at the EIC

- Broad kinematics and PID coverage available at EIC/ATHENA
  - large lever arm for SIDIS multiplicities and asymmetries
- Many SIDIS projections made for ATHENA and ECCE proposals:
  - Beam/target/double spin asymmetries, gluon saturation with dihadrons, etc.




$$z = \frac{p.p_h}{p.q}$$

y > 0.05

Duke