Semi-inclusive DIS kinematic reconstruction with ML

Connor Pecar

EIC Early Career Workshop

July 25, 2022, Stony Brook University

Research supported by:

U.S. DEPARTMENT OF ENERGY | Office of Science

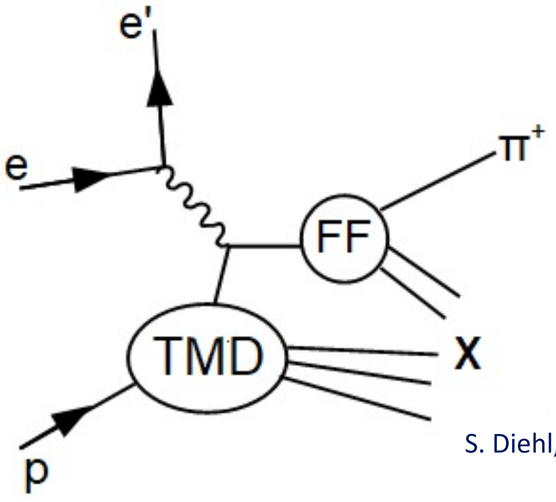
ATHENA

Duke UNIVERSITY

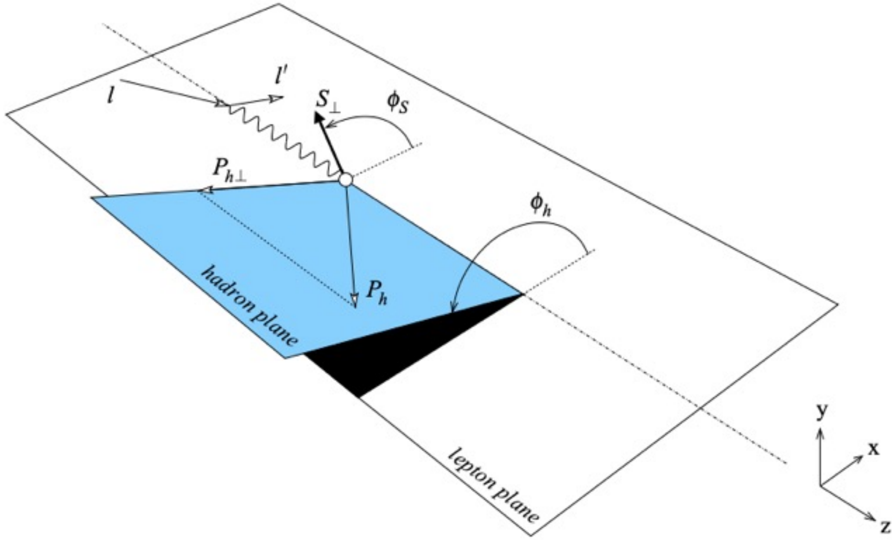
Duke

Semi-inclusive deep inelastic scattering at the EIC

- 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, TMD-FFs
    - Azimuthal angle and transverse momentum defined around virtual photon axis in target COM frame
    - Cross-section a function of $(x, y, z, p_T, \varphi_h, \varphi_s)$
- Many SIDIS observable projections made for ATHENA and ECCE proposals:
    - A_{LL} with kaons, gluon saturation with dihadrons, Sivers



S. Diehl, PRL

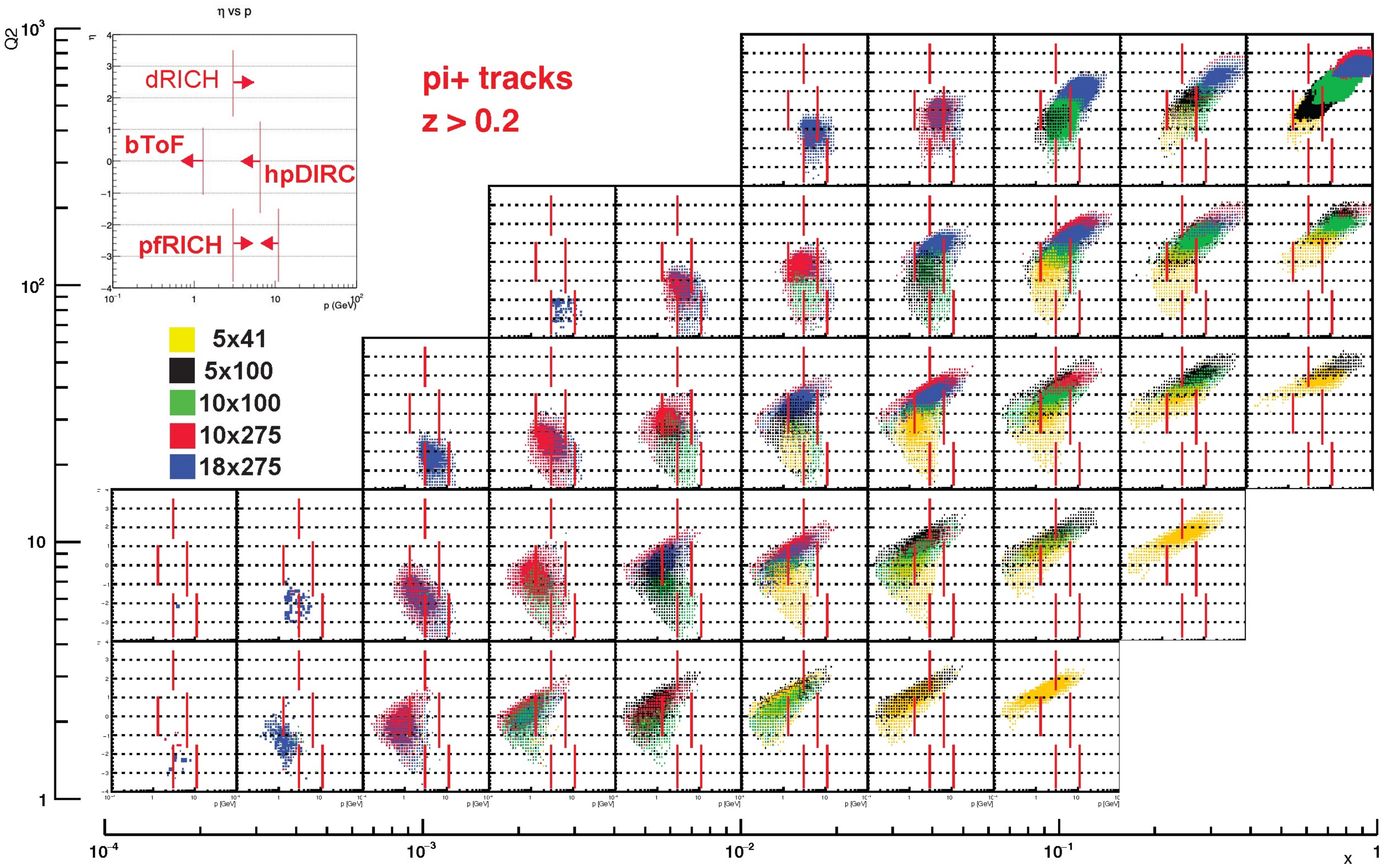


Phys.Rev. D70 (2004) 117504

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
- In addition to inclusive DIS reconstruction, rely on wide PID coverage (flavor separation), precise hadron tracking, virtual photon reconstruction



y > 0.05

Duke

SIDIS kinematic reconstruction

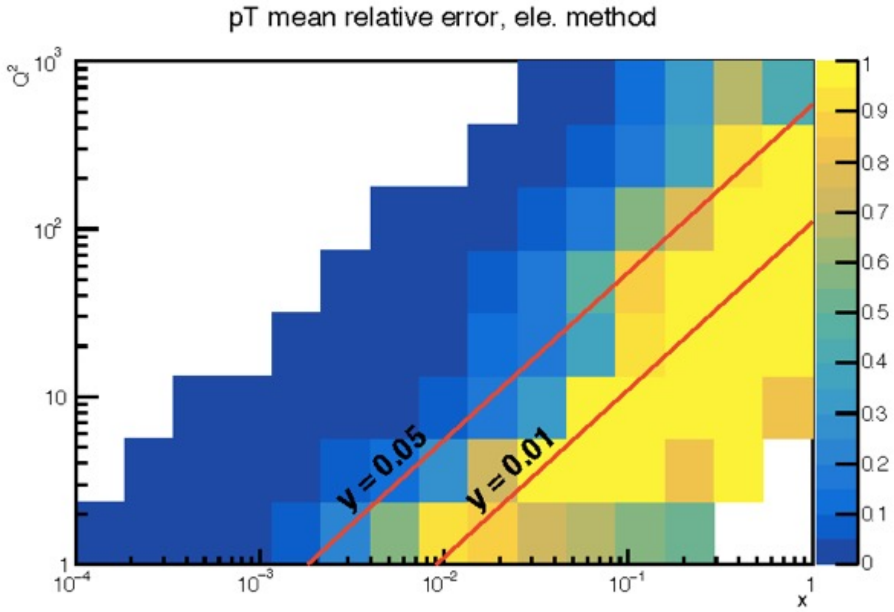
- SIDIS variables: reliant on reconstruction of virtual photon four-momentum, typically determined using

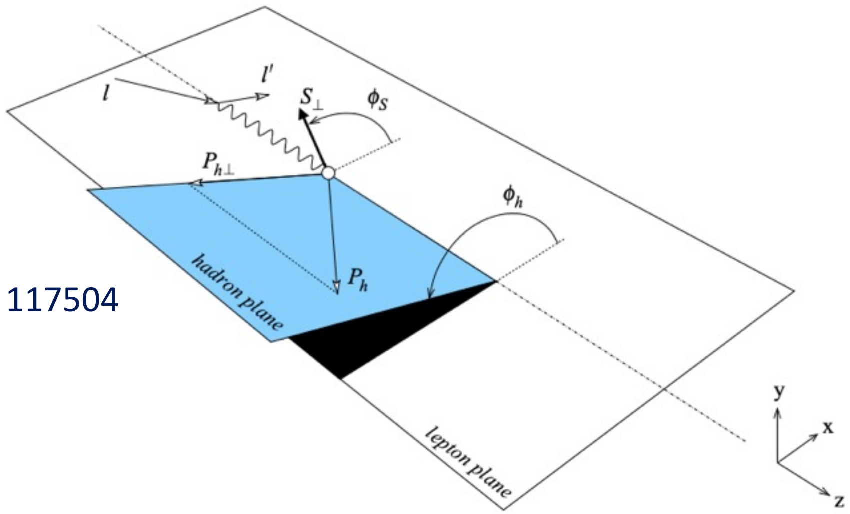
$$q = l - l'$$

- Reliable for larger y, but begins to fail for $y <$ ~0.05

  Low-y: region of interest for TMDs, and important for evolution studies
- **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:



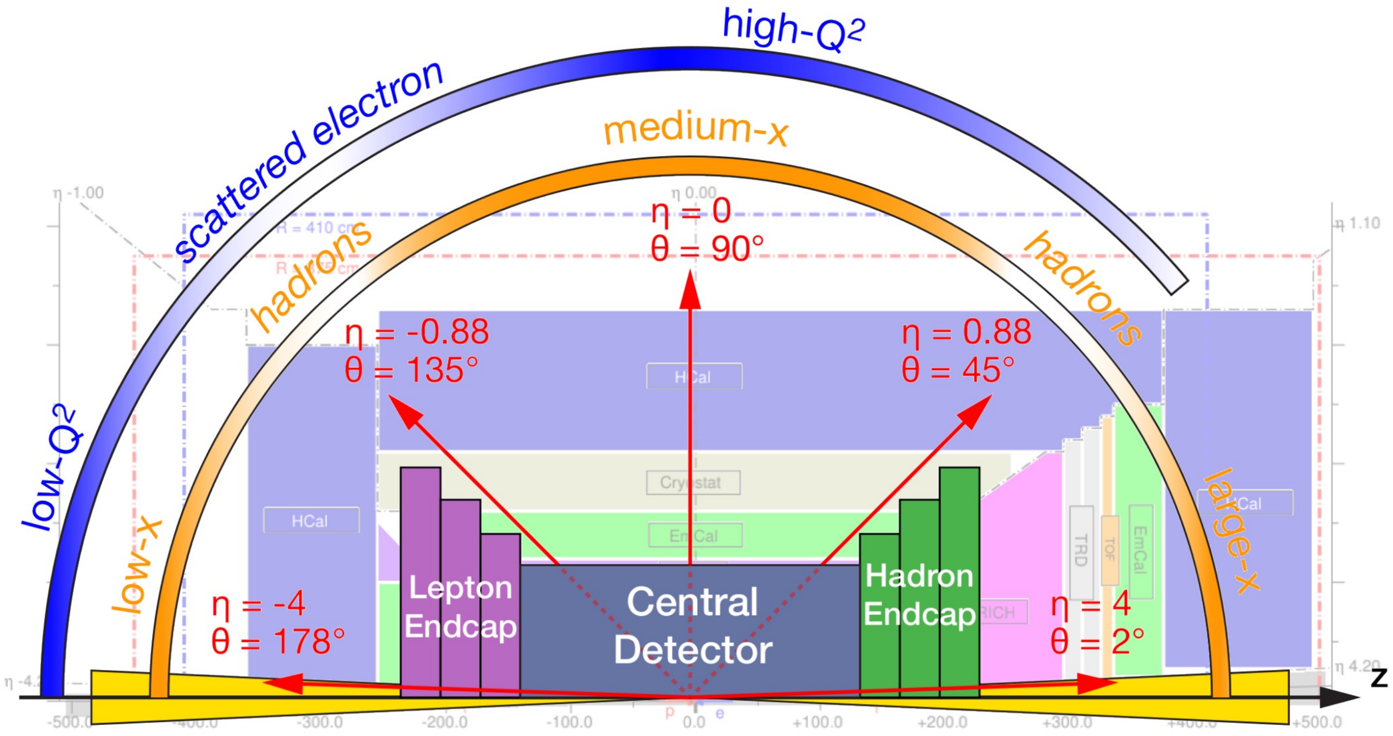


Phys.Rev. D70 (2004) 117504

Duke

Reconstruction with hadronic final state

- Through conservation of momentum and energy, **hadronic final state (HFS) should also contain enough information to constrain q**
- HFS methods developed at HERA for inclusive DIS variables for regions in which electron method less reliable
- Methods utilizing hadronic final state should be more robust with respect to first-order radiative corrections
  - Impact of radiative effects expected to be large for SIDIS observables at EIC



Duke

Reconstruction with hadronic final state

- Method used in EIC YR and ATHENA proposal 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

i)	*Leptonic variables*	$q \equiv q_l = k_2 - k_1, \quad y_l = p_1.(k_1 - k_2)/p_1.k_1$
ii)	*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)	*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)	*Mixed variables* [81]	$q = q_l, y_m = y_{JB}$
v)	*Double angle method* [83]	$Q^2_{DA} = \frac{4E(k_2)^2 \cos^2(\theta(k_2)/2)}{\sin^2(\theta(k_2)/2) + \sin(\theta(k_2)/2)\cos(\theta(k_2)/2)\tan(\theta(p_2)/2)},$ $y_{DA} = 1 - \frac{\sin(\theta(k_2)/2)}{\sin(\theta(k_2)/2) + \cos(\theta(k_2)/2)\tan(\theta(p_2)/2)},$

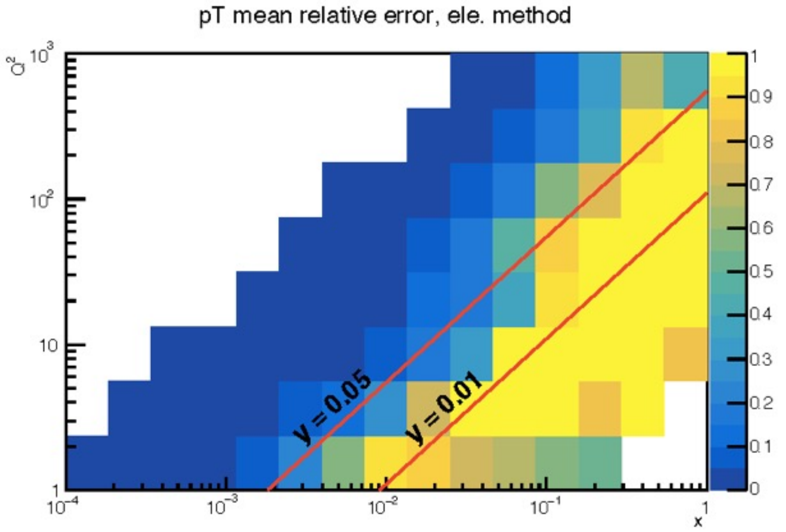
Prog. Part. Nucl. Phys. 2013, Blümlein

- \+ solution in quadratic equation found to always be closer to MC truth
- Resolution improved if this is carried out in head-on frame, then transformed to lab frame
  - Large crossing angle needed for EIC

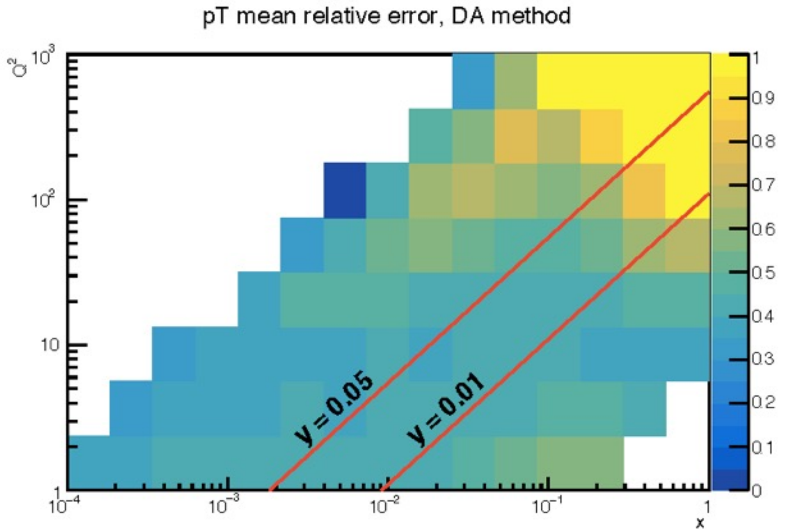
Duke

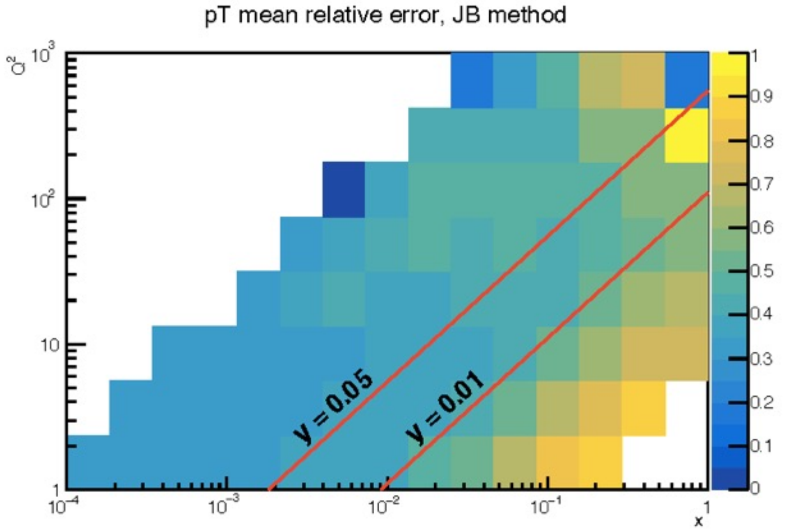
ATHENA full simulation SIDIS resolution, p_T

Transverse momentum (w.r.t. q), 10x275, pi+, z > 0.2:



DA, hybrid method -> pT resolution more acceptable at low y



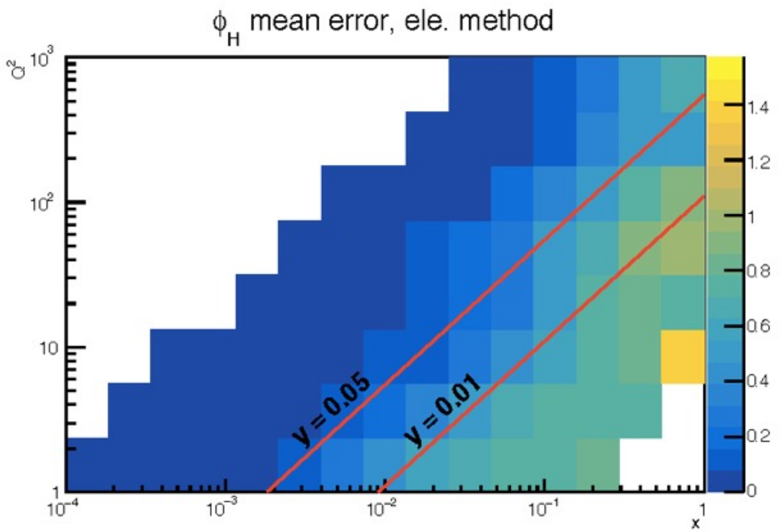


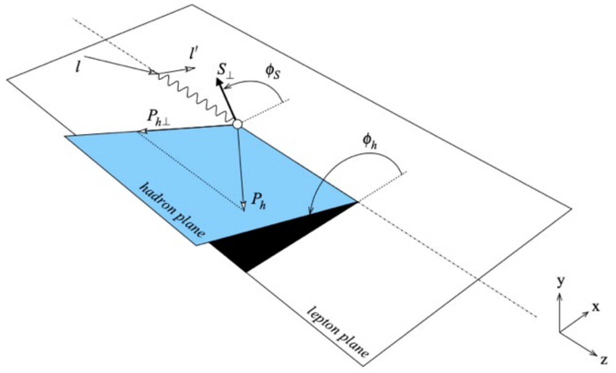
<- JB, HFS only Potential for CC

Duke

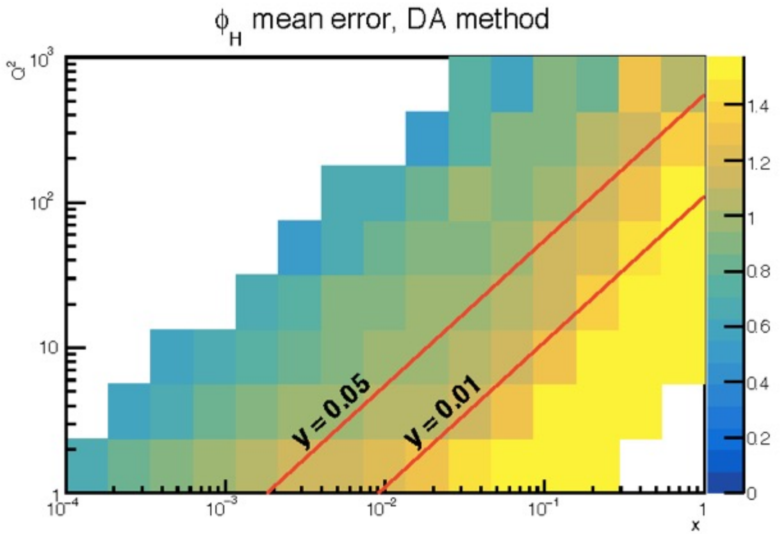
ATHENA full simulation SIDIS resolution, ϕ_h

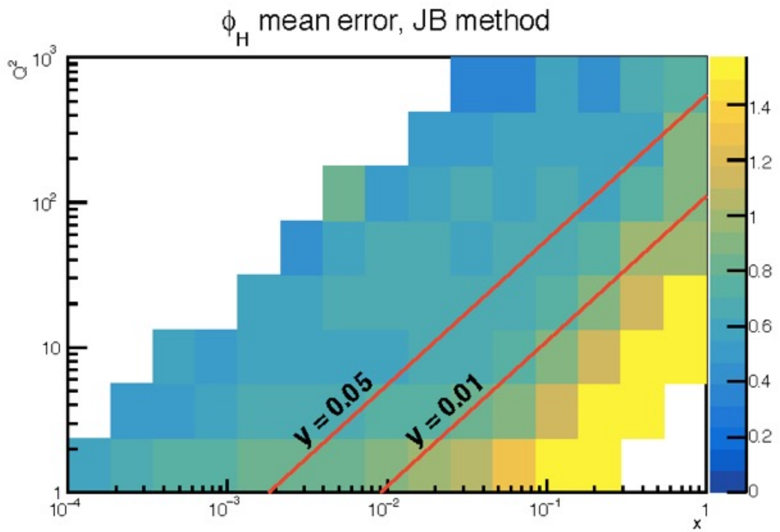
ATHENA full simulation, 10x275, pi+, z > 0.2





Angular resolution still poor at low-y with all methods





Duke

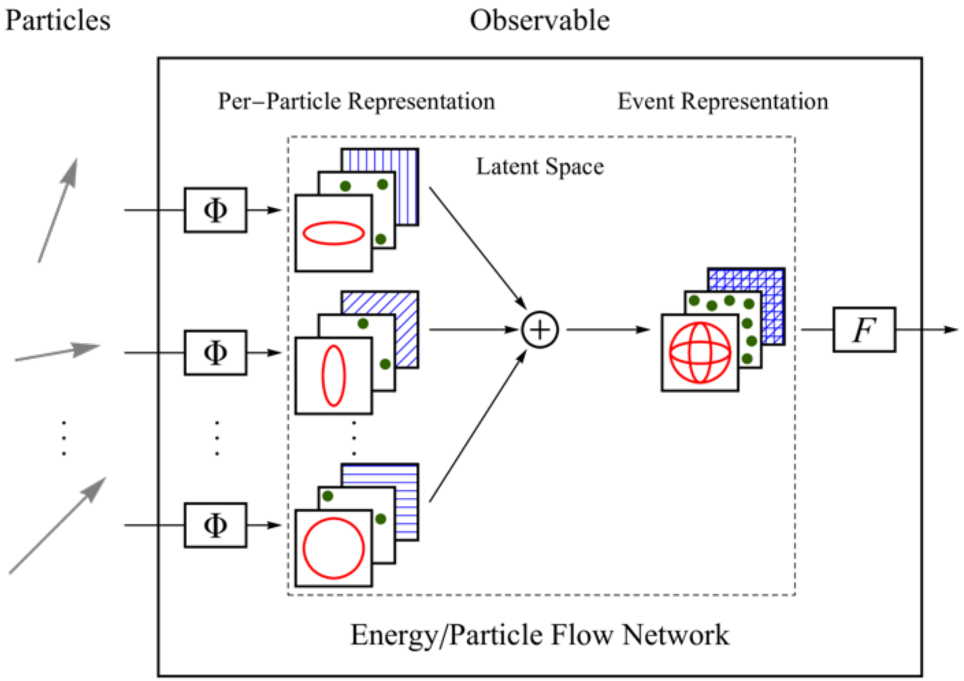
Machine learning reconstruction

- Based on hybrid HFS-electron SIDIS reconstruction, using ML to combine information from both to reconstruct q
  - Potential to correct overall HFS and electron momentum and to more reliably reconstruct z and t components than exact formula
  - 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, as well as NIM-A 1025 (2022) 166164, Arratia, Britzger, Long, Nachman)
- Currently utilizing graph-like neural network architectures designed for jet reconstruction

Duke

Particle flow networks

- Particle flow networks (PFN) developed by Komiske et al., (JHEP 01 (2019) 121, Komiske, Metodiev, Thaler)
  - Accepts unordered set of particles
  - Particles -> 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
- Designed to be general purpose, universal approximator on sets of particles



JHEP 01 (2019) 121

Duke

ML SIDIS model and training

- Model combining electron and HFS:
  - Particle features for PFN: momentum, energy, η, φ 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

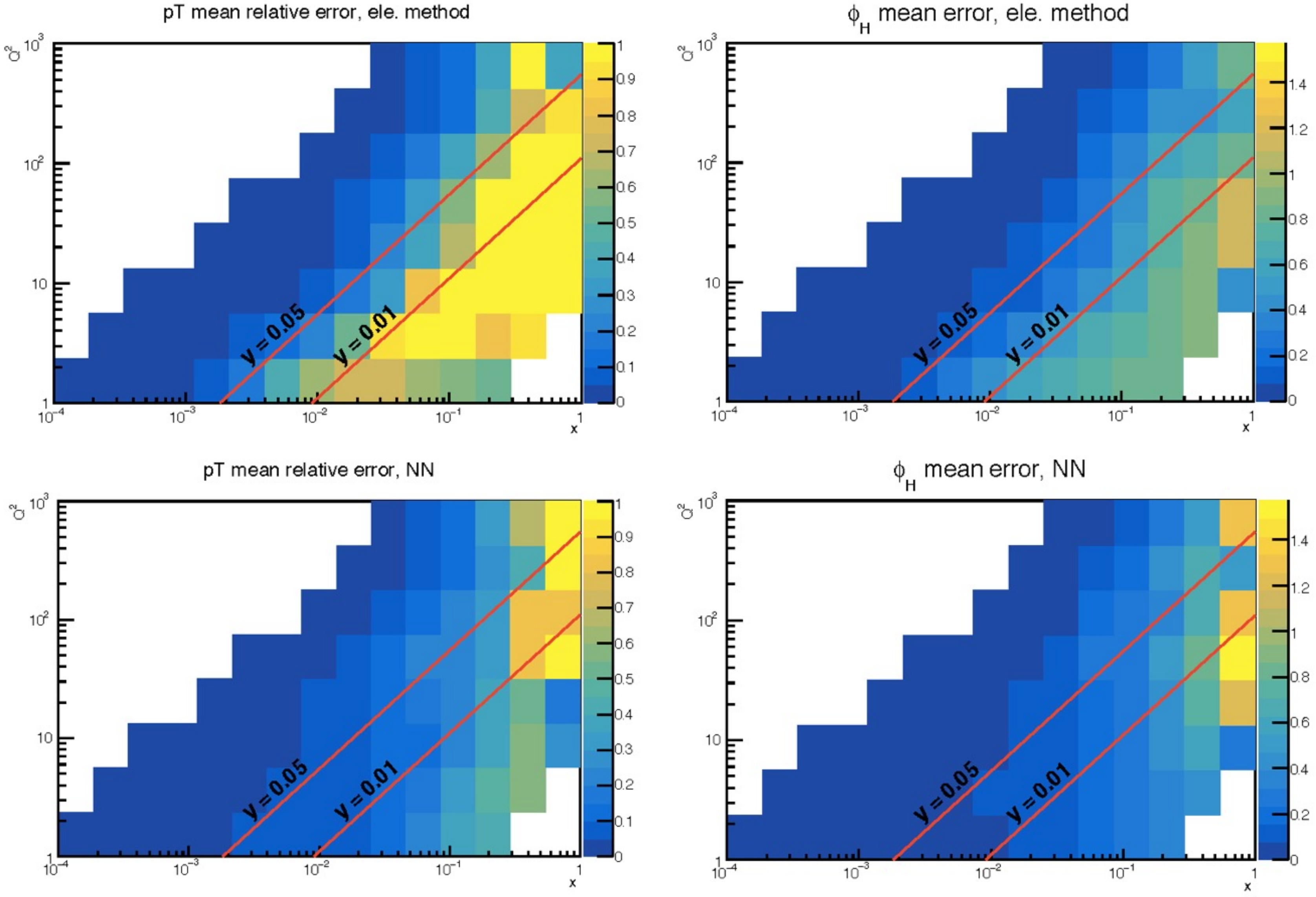
Duke

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

$$\frac{p_T - p_{T,true}}{p_{T,true}} \qquad \phi_H - \phi_{H,true}$$

Electron method

Neural network



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

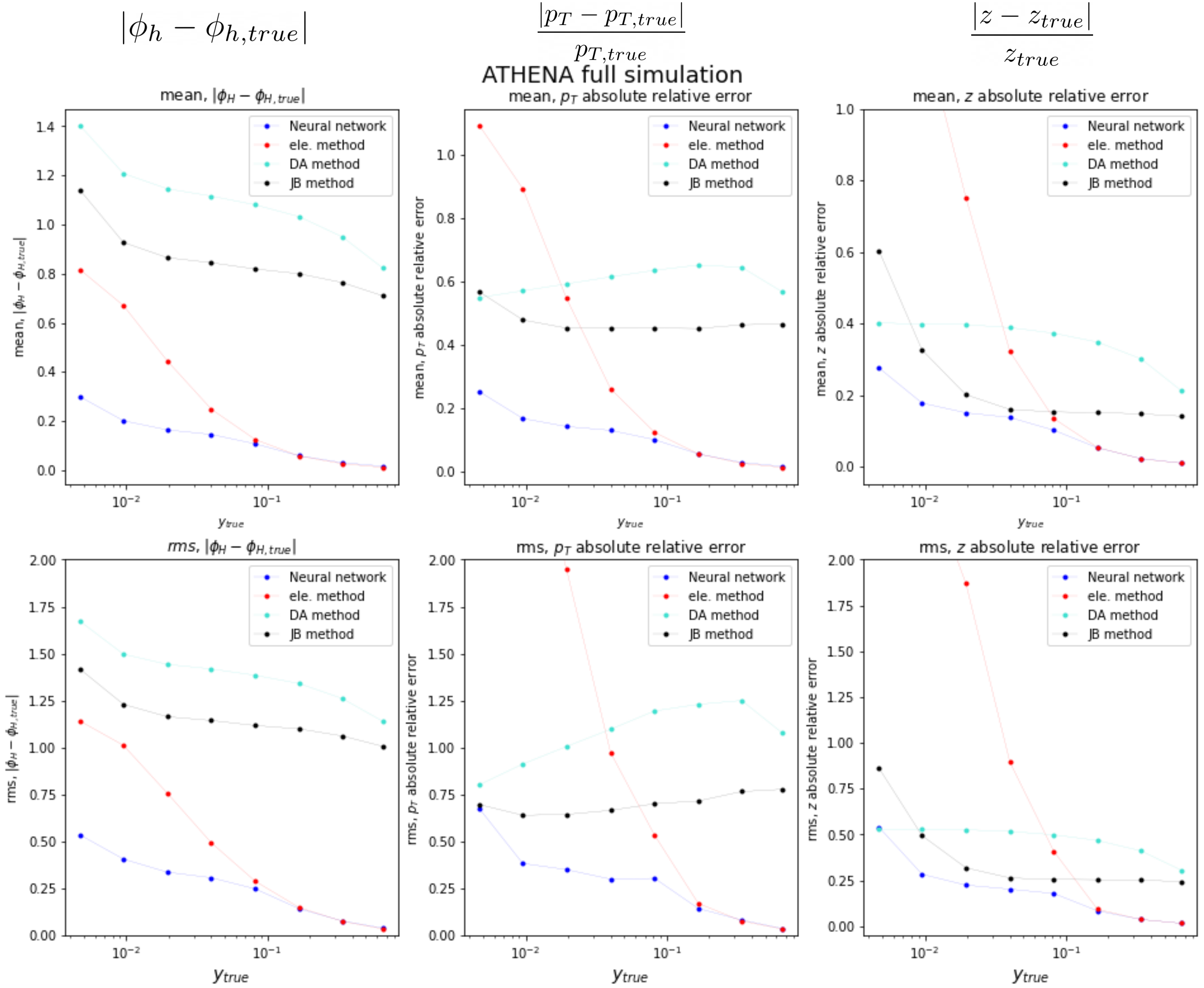
Duke

ATHENA full simulation, 10x275, pi+, z > 0.2
Reconstruction as a function of y

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:



Duke

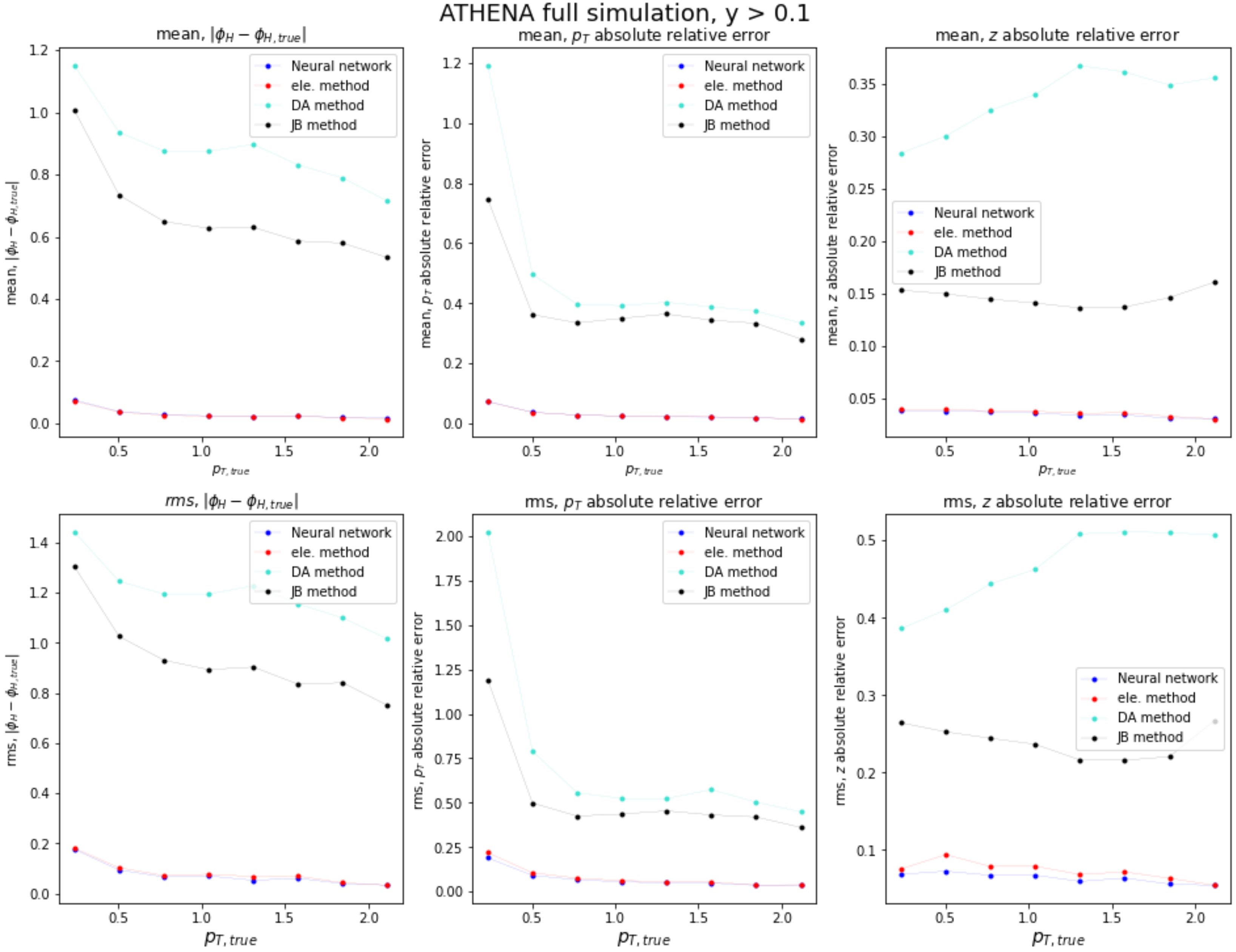
ATHENA full simulation, 10x275, pi+, z > 0.2

Reconstruction as a function of pT, y > 0.1

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

Mean:

RMS:



- Comparison with other HFS methods vs true pT
- Machine learning method equals/slightly outperforms electron method for large y (expected)

Duke

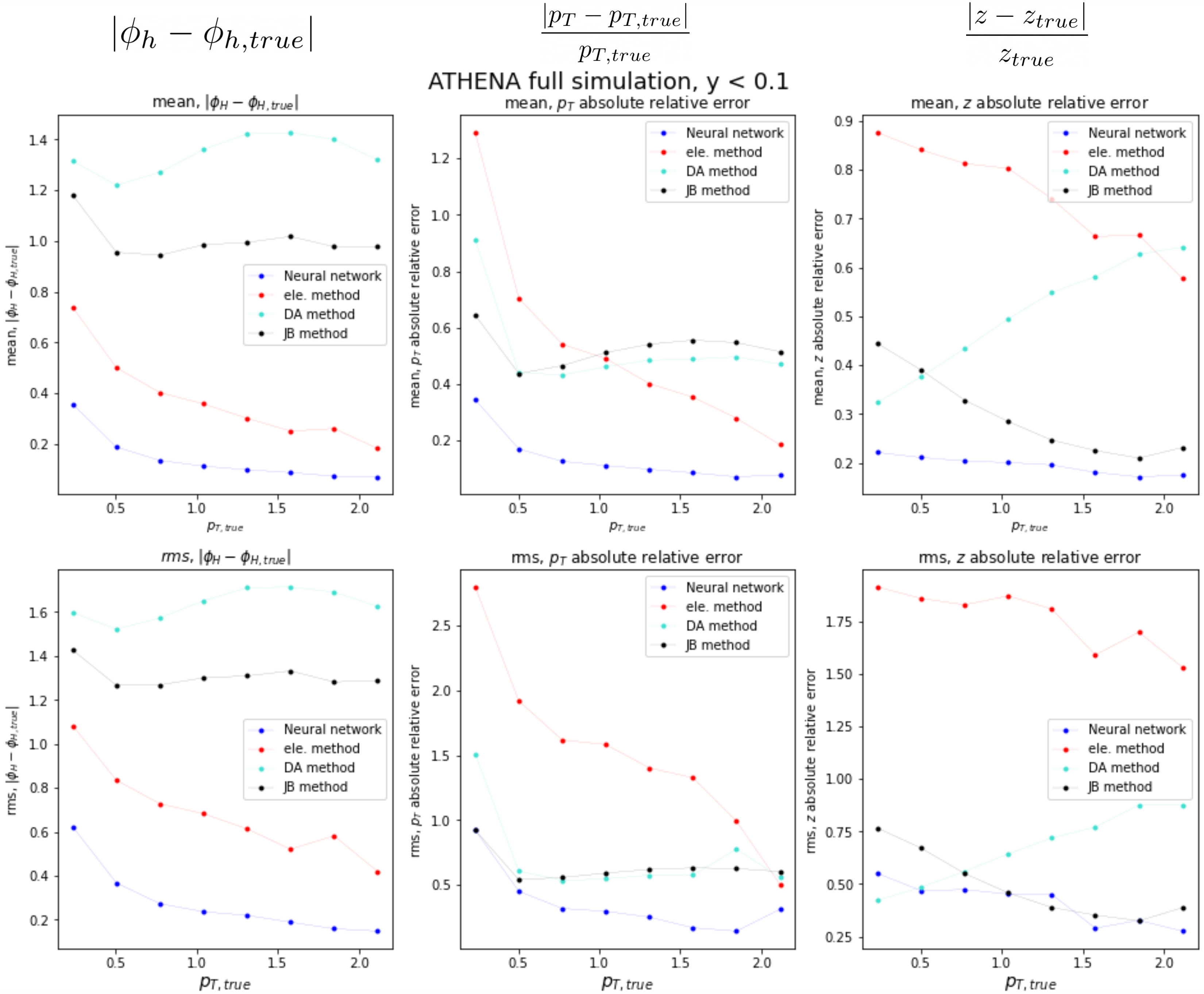
ATHENA full simulation, 10x275, pi+, z > 0.2

Reconstruction as a function of pT, y < 0.1

Mean:

RMS:

- Comparison with ele., other HFS methods vs true pT
- At low-y, network is able to reliably reconstruct kinematics even at very low pT (w.r.t virtual photon)
- Mean pT ~ 0.5 GeV



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
- We demonstrate a machine learning approach combining the hadronic final state and scattered electron which surpasses existing methods for all of x-Q2 and p_T
- Next steps in reconstruction:
  - Currently working on replacing the particle flow network with an architecture which can learn correlations between HFS particles (such as a GNN), as well as exploring other deep learning approaches
  - Method will need to be tested with better implementation of radiative effects

Research supported by:

U.S. DEPARTMENT OF ENERGY | Office of Science

Duke