Reconstructing the Kinematics of Deep Inelastic Scattering with Deep Learning

M. Arratia, D. Britzger, O. Long, B. Nachman

Terascale workshop 2021
DESY, Hamburg (virtual)
23.11.2021

https://arxiv.org/abs/2110.05505

MAX-PLANCK-INSTITUT
FÜR PHYSIK

Deep-inelastic electron-proton scattering

DIS Kinematic variables in (neutral-current) deep-inelastic scattering

Virtuality
of exchanged boson

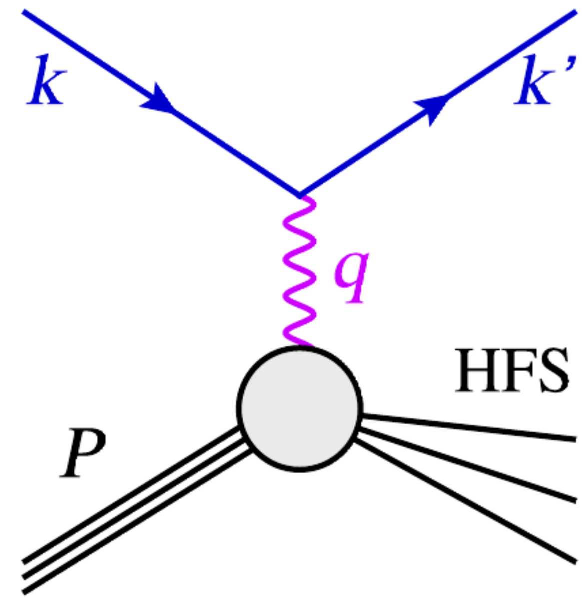
$$\boxed{Q^2 = -q^2}$$

Inelasticity

$$\boxed{y = \frac{q \cdot P}{k \cdot P}}$$

Bjorken-x

$$\boxed{x = Q^2/(sy)}$$



Electron-proton colliders

- HERA 1992–2007 $\sqrt{s}$ = 319GeV
- EIC 2030s+ $\sqrt{s}$ = 141GeV
- LHeC 2030s+ (?) $\sqrt{s}$ = 1300GeV

The ATHENA experiment (w/ DELPHES fast simulation)

Electron Ion Collider (EIC)

- beams: 275 GeV (p), 18 GeV (e)

ATHENA experiment

- 3 T solenoid
- All silicon tracker
- Very good particle ID
- Large acceptance (-4 < η < 4)

DELPHES fast simulation

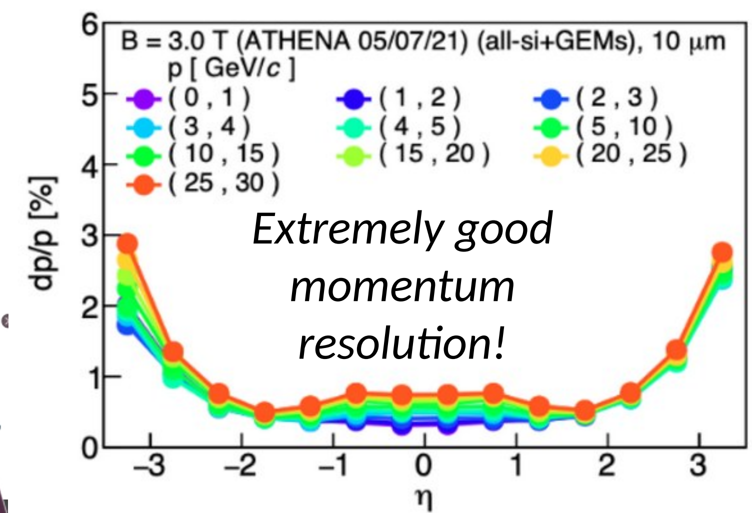
- Detailed momentum smearing of generated particles

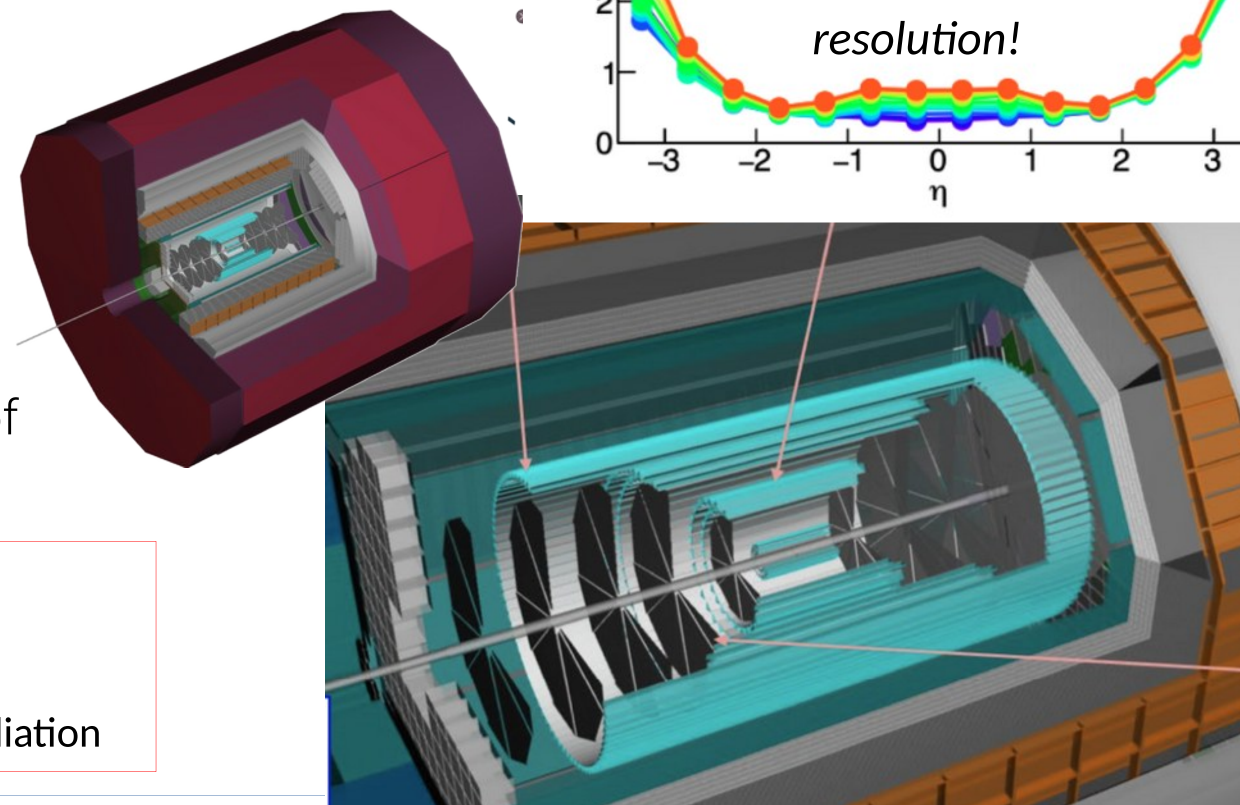
Event selection for this work

Generated $Q^2 > 200\ \mathrm{GeV}^2$

32 GeV < event (E-pz) < 40 GeV ,

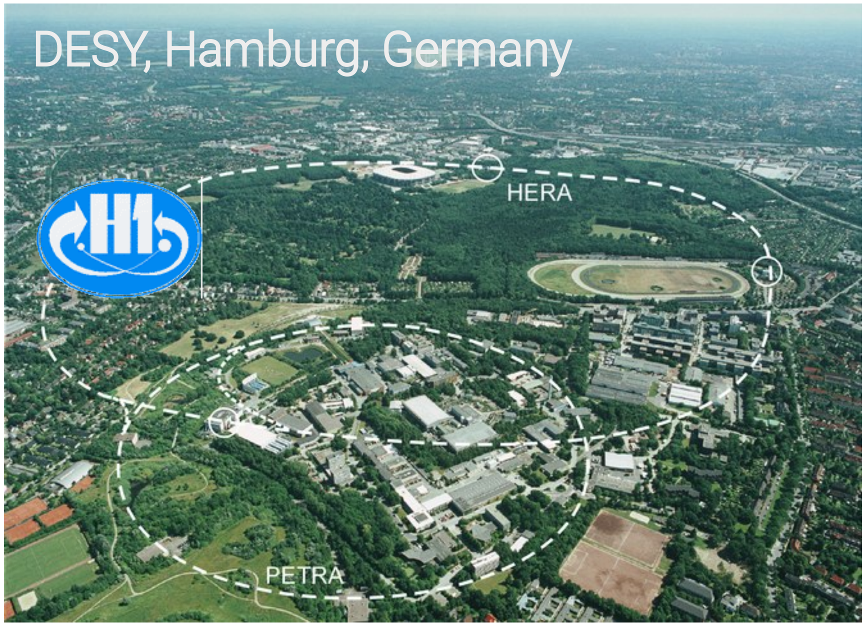
(±4 GeV around $2E_e$) reduces QED radiation





The H1 experiment at HERA

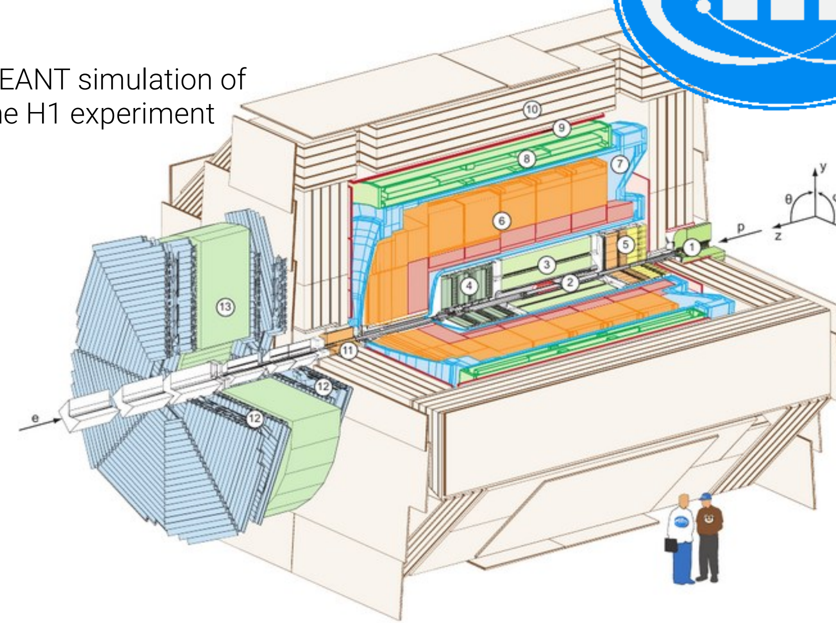
HERA electron-proton collider



- HERA I: 1994 – 2000
  HERA II: 2003 – 2007
- E_e=27.6 GeV, E_p=920GeV
  $\sqrt{s}$ = 300 or 319 GeV

H1 experiment at HERA

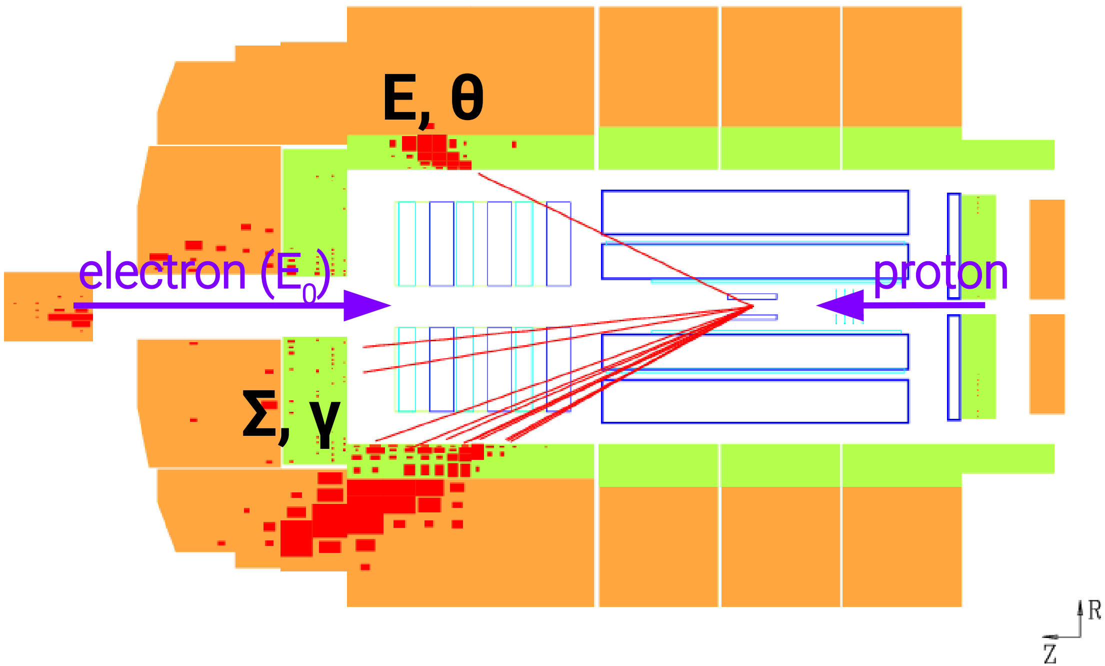
GEANT simulation of the H1 experiment



'multi-purpose' detector

- Asymmetric design with trackers, calorimeter, solenoid, muon-chambers, forward & backward detectors, ...

DIS kinematics in experiment



DIS kinematic variables

- Overconstrained system
  - p_T conservation
  - p_Z conservation
- 3 out of 5 quantities are needed
  → Many formulae with pros and cons in literature

Method name	y	Q^2	$x \cdot E$
Electron (e)	$1 - \frac{\Sigma_e}{2E_0}$	$\frac{E^2 \sin^2 \theta}{1-y}$	$\frac{E(1+\cos\theta)}{2y}$
Double angle (DA) [6, 7]	$\frac{\tan\frac{\gamma}{2}}{\tan\frac{\gamma}{2}+\tan\frac{\theta}{2}}$	$4E_0^2 \cot^2 \frac{\theta}{2}(1-y)$	$\frac{Q^2}{4E_0 y}$
Hadron (h, JB) [4]	$\frac{\Sigma}{2E_0}$	$\frac{T^2}{1-y}$	$\frac{Q^2}{2\Sigma}$
ISigma (IΣ) [9]	$\frac{\Sigma}{\Sigma+\Sigma_e}$	$\frac{E^2 \sin^2 \theta}{1-y}$	$\frac{E(1+\cos\theta)}{2y}$
IDA [7]	y_{DA}	$\frac{E^2 \sin^2 \theta}{1-y}$	$\frac{E(1+\cos\theta)}{2y}$
$E_0 E \Sigma$	y_h	$4E_0E - 4E_0^2(1-y)$	$\frac{Q^2}{2\Sigma}$
$E_0 \theta \Sigma$	y_h	$4E_0^2 \cot^2 \frac{\theta}{2}(1-y)$	$\frac{Q^2}{2\Sigma}$
$\theta\Sigma\gamma$ [8]	y_{DA}	$\frac{T^2}{1-y}$	$\frac{Q^2}{2\Sigma}$
Double energy (A4) [			

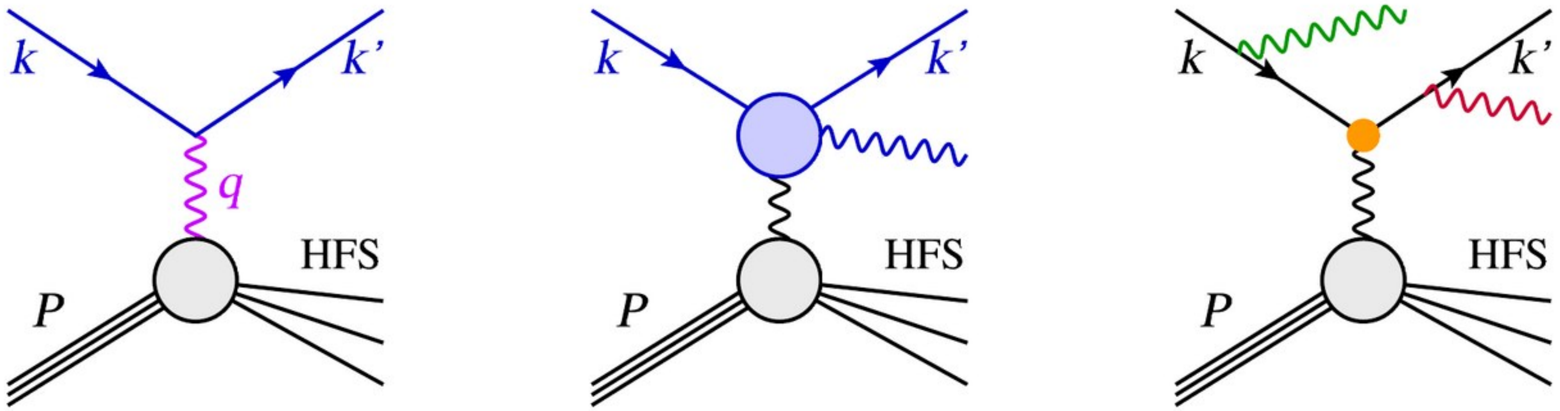
... and many more in literature

Bassler, Bernardi, [NIM A 426 (1999) 583] [NIM A 361 (1995) 197], and several others...

Improvements are not straight forward

- Combination of multiple quantities studied extensively at HERA
- Quantities are calibrated against each other
- QED radiation...

QED radiation



Real photon radiation invalidates conservation laws

→ Different reconstruction methods → different results

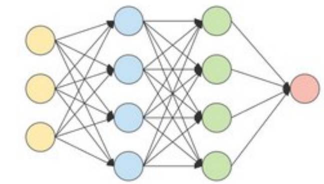
- Careful treatement required for definition of learning targets
- Initial-state photon often collinear with beam → undetected
- final-state photon may become a fake HFS particle

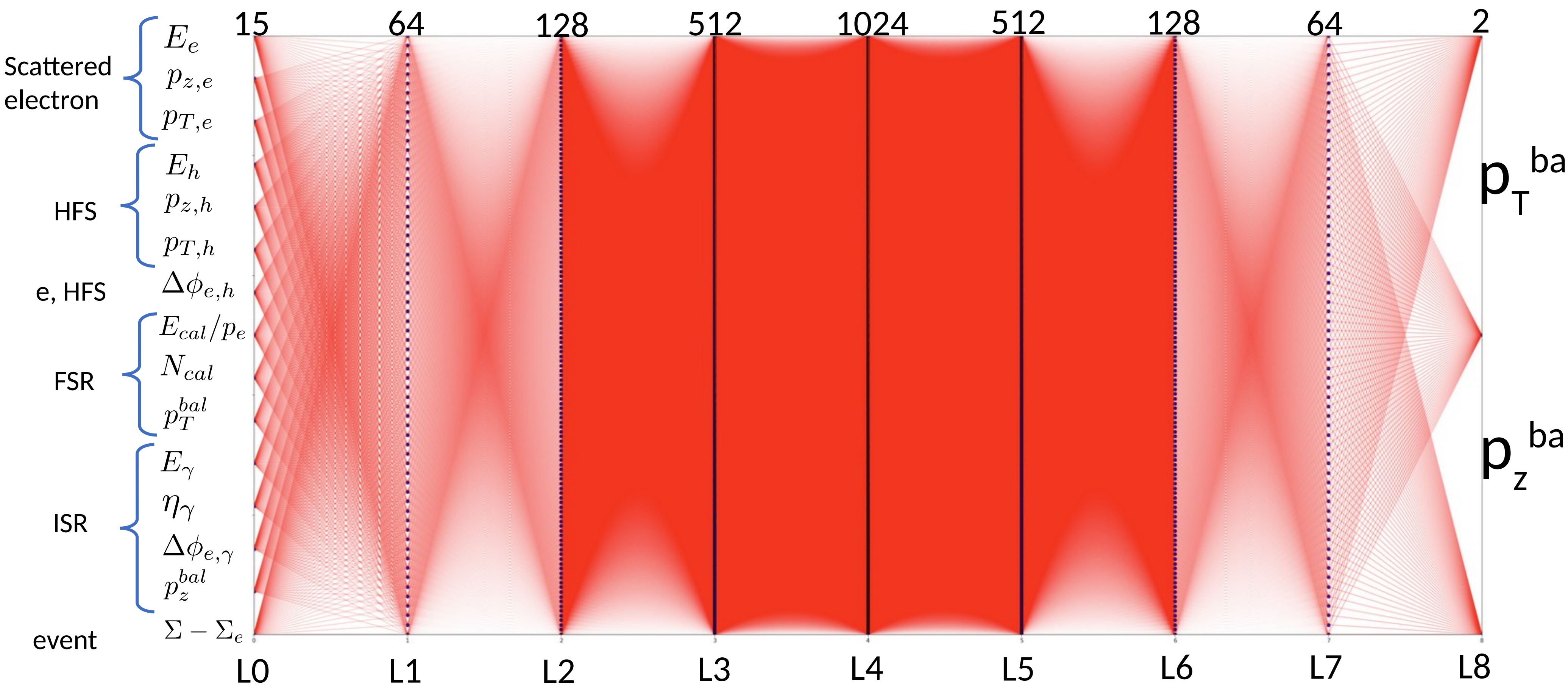
→ Quantify strength of QED effect

- Implementation independent

$$p_T^{\mathrm{bal}} = 1 - \frac{p_{T,e}}{T} \qquad p_z^{\mathrm{bal}} = 1 - \frac{\Sigma_e + \Sigma}{2\ E_0}$$

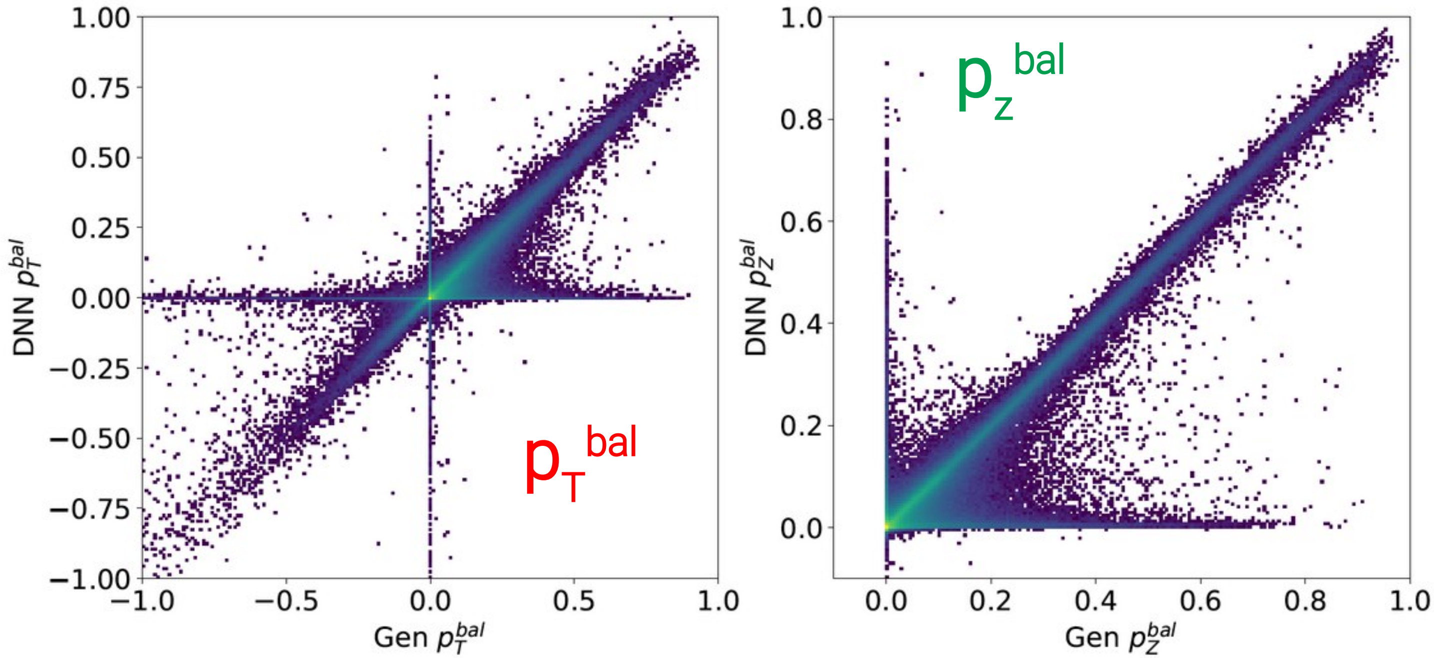
Network diagram

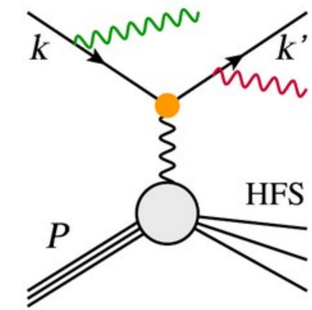




There are 1,197,184 connections in the network.

QED radiation





Regression deep-neural network for p_z^{bal} and p_T^{bal}

- 15 input quantities
- Learning targets are generated values of p_z^{bal} and p_T^{bal}

DNN accurately estimates p_z^{bal} and p_T^{bal} in most events

DIS reconstruction : regression DNN for Q^2, y, x

Learning targets are logarithm of gen values of Q^2, y, and x.

Three output nodes for log of Q^2, y, and x with linear activation.

Loss function is Huber.

We tried three approaches:

1. Add 2 QED regression outputs (p_z^{bal} and p_T^{bal}) to 15 other inputs.
2. Add 3 QED classification outputs (ISR, FSR, NoR) to 15 other inputs.
3. Use the same 15 inputs as in QED DNNs.

DIS reconstruction : regression DNN for Q^2, y, x

Learning targets are logarithm of gen values of Q^2, y, and x.

Three output nodes for log of Q^2, y, and x with linear activation.

Loss function is Huber.

We tried three approaches:

1. Add 2 QED regression outputs (p_z^{bal} and p_T^{bal}) to 15 other inputs.

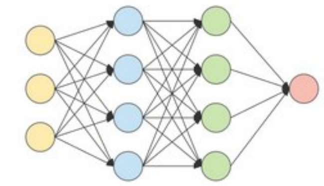
2. Add 3 QED classification outputs (ISR, FSR, NoR) to 15 other inputs.

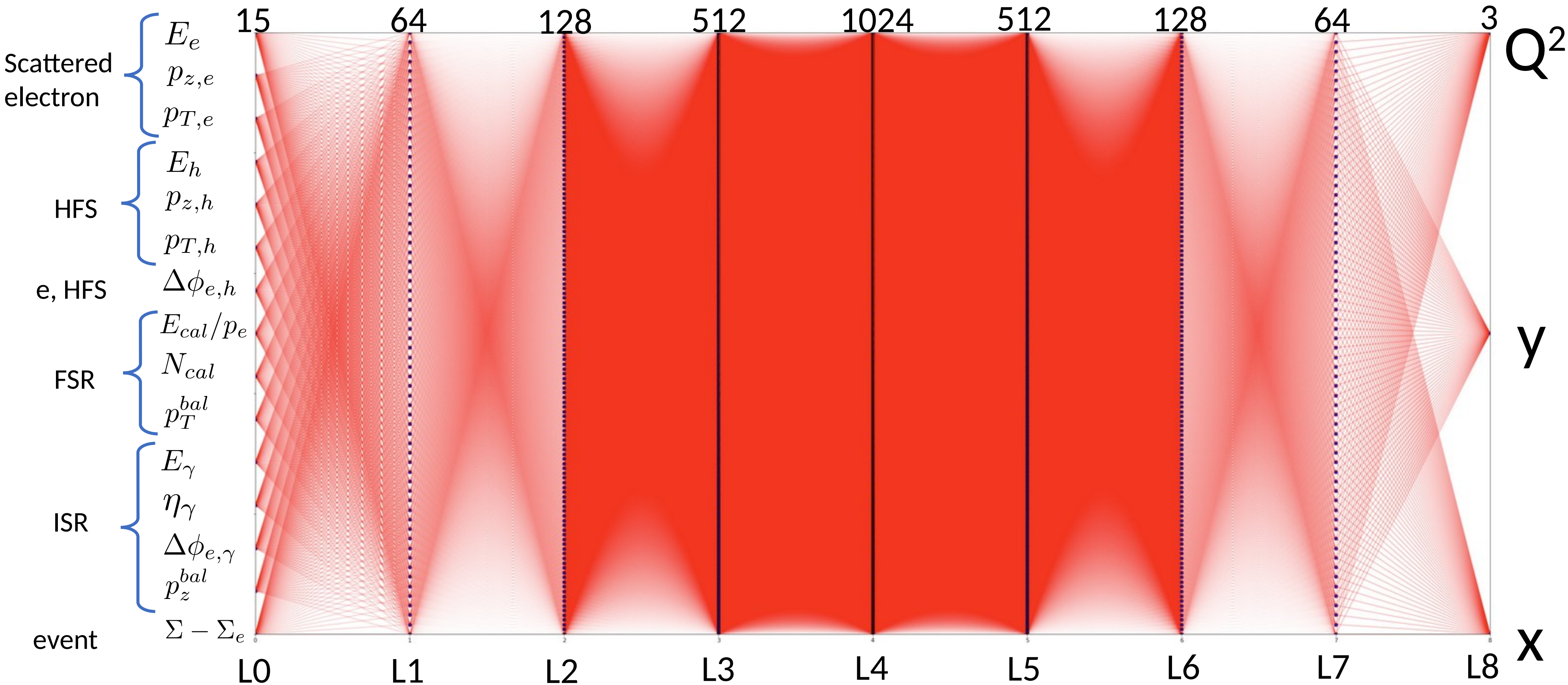
3. Use the same 15 inputs as in QED DNNs.

All 3 give essentially identical results!

We choose the simplest option (3).

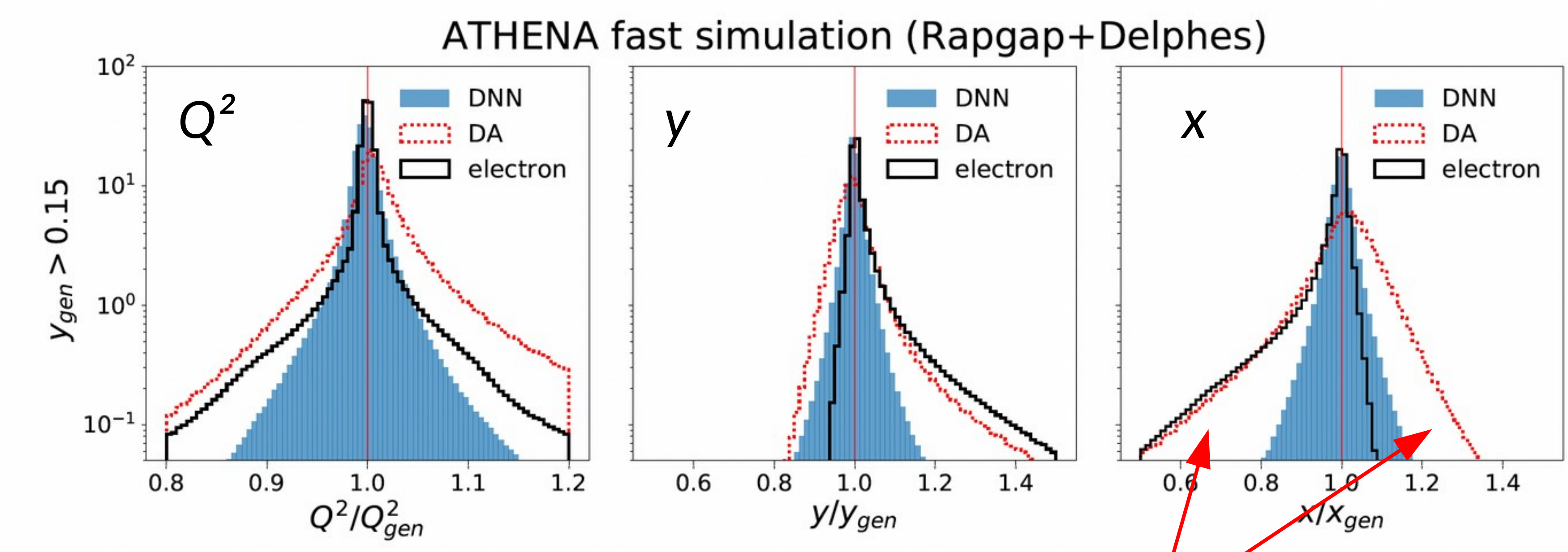
Network diagram





There are 1,197,184 connections in the network.

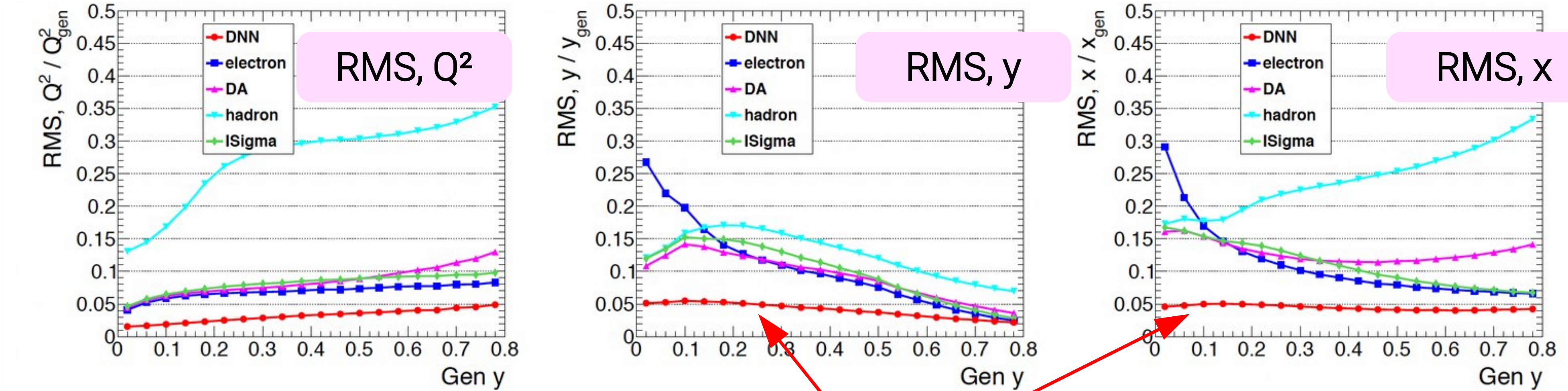
DIS reconstruction – Regression DNN for Q^2, y, x



DNN has similar core resolution to best conventional method
(electron at high y, DA at low y).

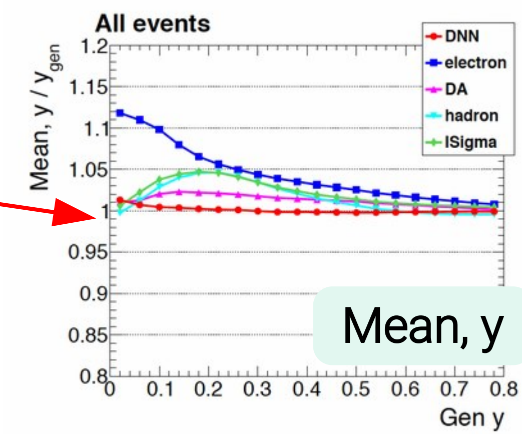
Large tails from QED radiation in conventional reconstruction methods absent in DNN.

DIS reconstruction – Regression DNN for Q^2, y, x



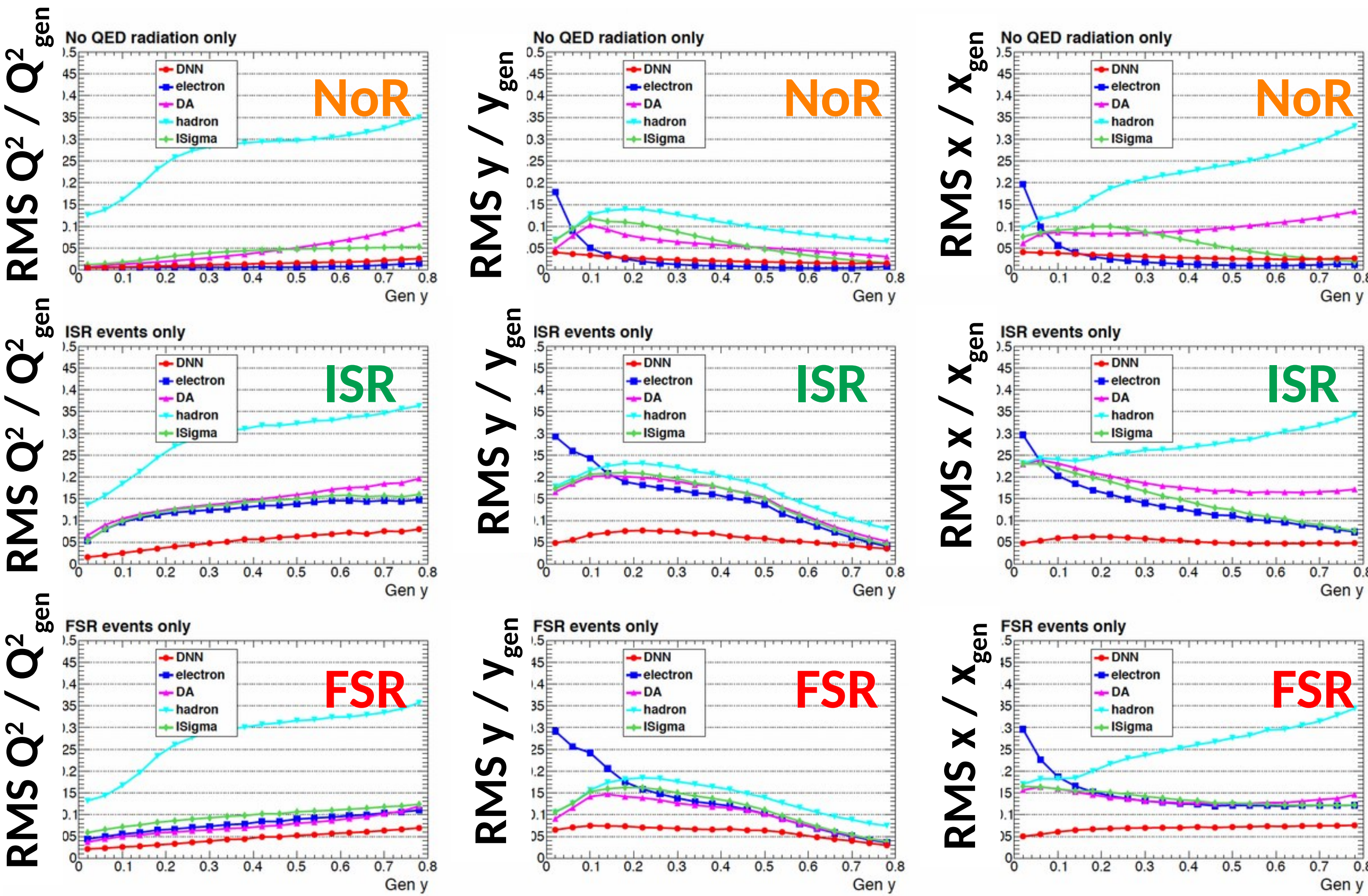
DNN outperforms all conventional reconstruction methods.

DNN has smallest RMS and essentially no bias.



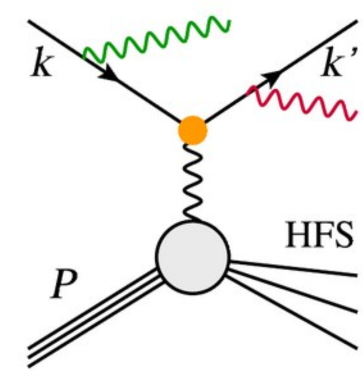
ATHENA

Resolution (RMS) vs. y_{gen}



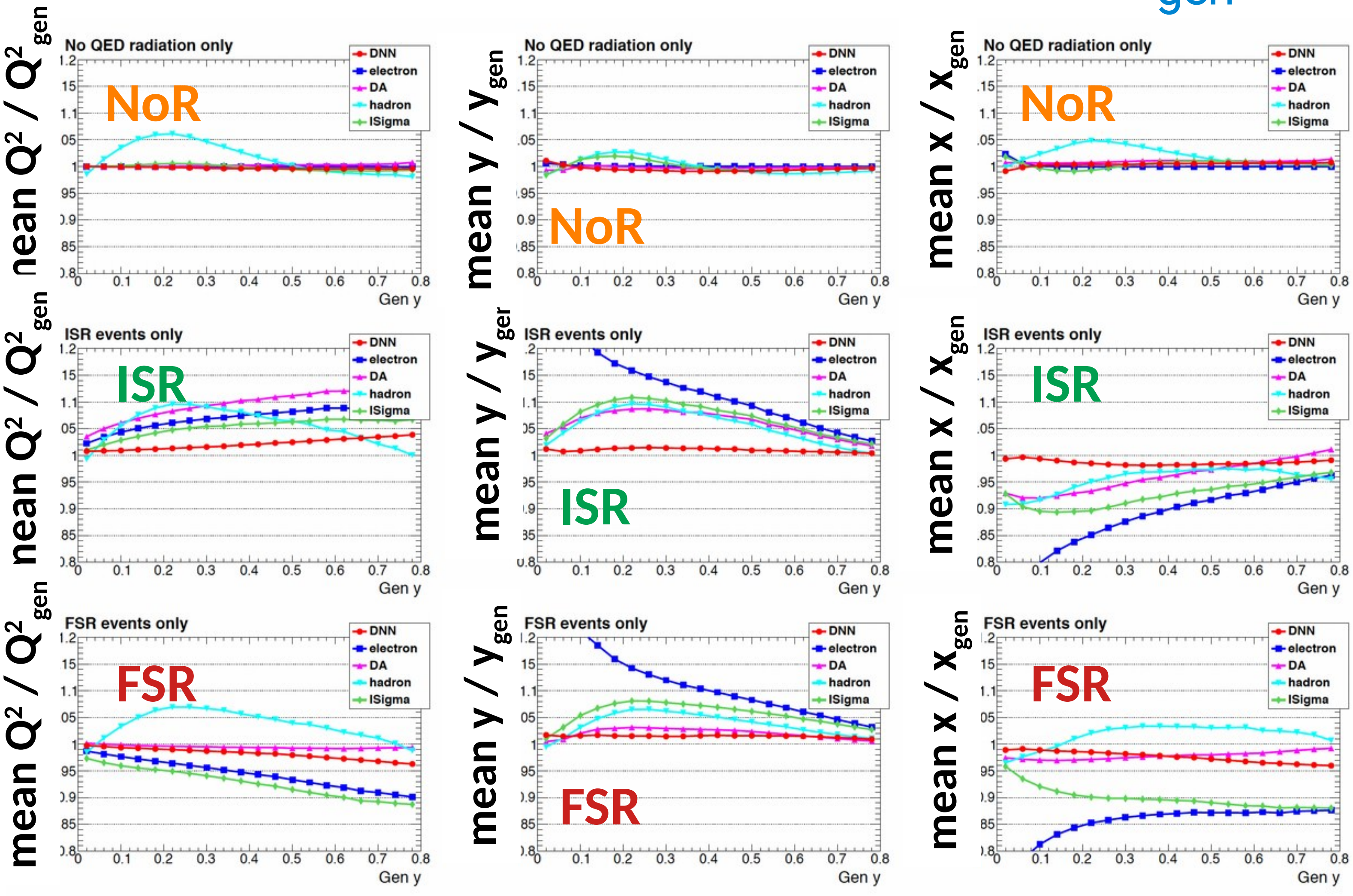
Electron method has better core resolution than DNN for y>0.15 in NoR events (no tails).

DNN resolution much less affected by QED radiation



RMS and mean calculated for events with measured / gen ratio between 0 and 2.

Bias (mean) vs. y_{gen}



All methods (except hadron) are unbiased in events with no QED radiation

DNN remains unbiased in events with QED radiation, while other methods have large bias

DNN has successfully learned how to mitigate QED radiation effects.

RMS and mean calculated for events with measured / gen ratio between 0 and 2.

ATHENA

Demonstration of DNN with H1 full simulation

HERA beam energies: E_e = 27.6 GeV, E_p = 920 GeV.

NC DIS with Q^2_{gen} > 200 GeV^2.

RAPGAP 3.1 generator (includes HERACLES).

GEANT detector simulation with fast showers

Employ standard reconstruction methods for electron and HFS.

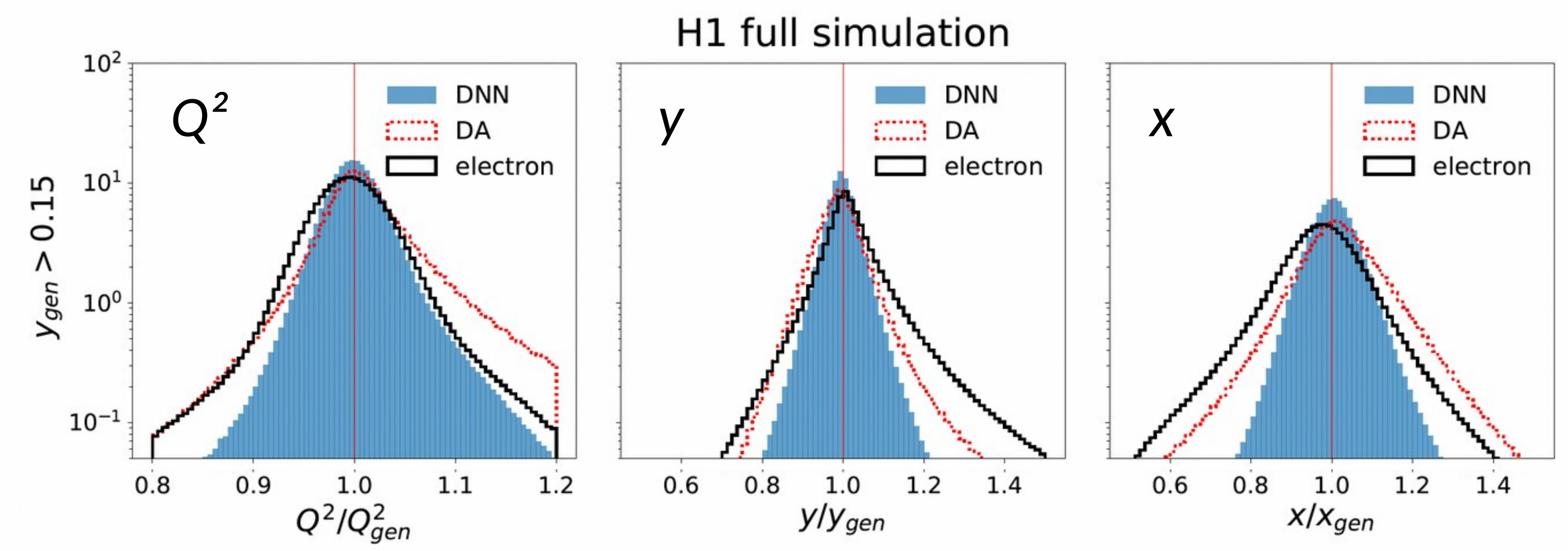
Includes real calorimeter noise.

Includes run-specific conditions.

Event selection:

45 GeV < event ($E-p_z$) < 65 GeV (±10 GeV around 2 E_e)

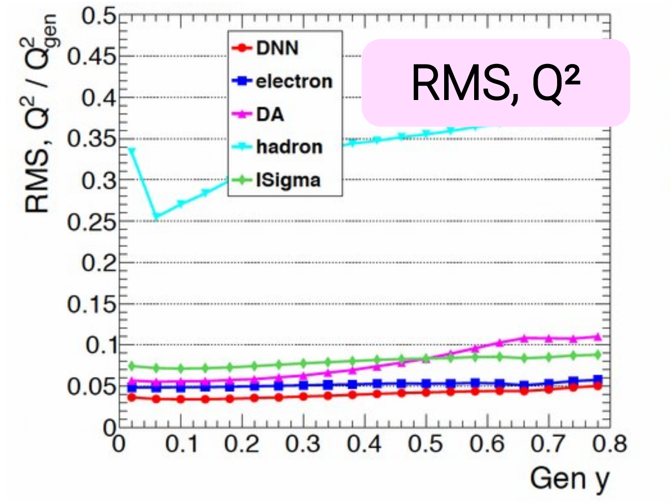
H1 @ HERA – Regression DNN for Q^2, y, x

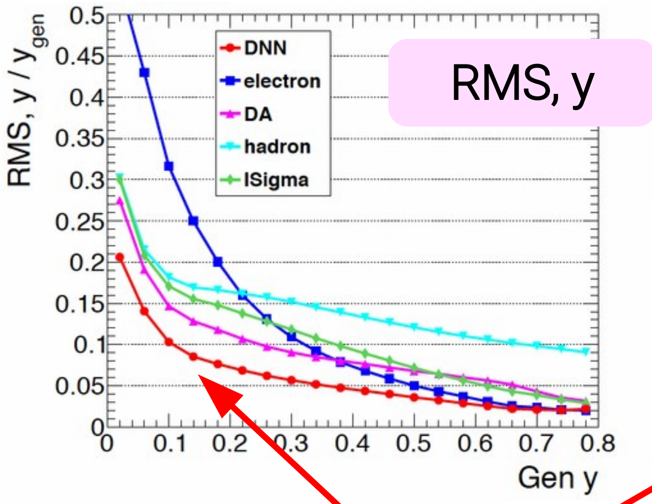


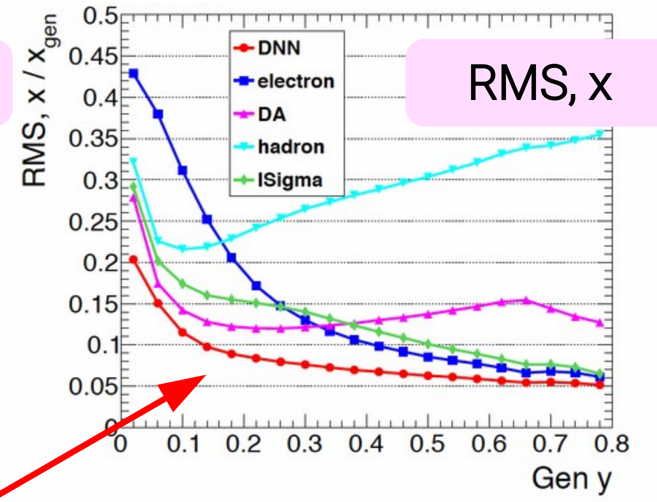
*DNN has **better** core resolution than best conventional method (electron at high y, DA at low y).*

DNN distributions much more symmetric, free of large QED radiation tails.

H1 @ HERA – RMS and mean





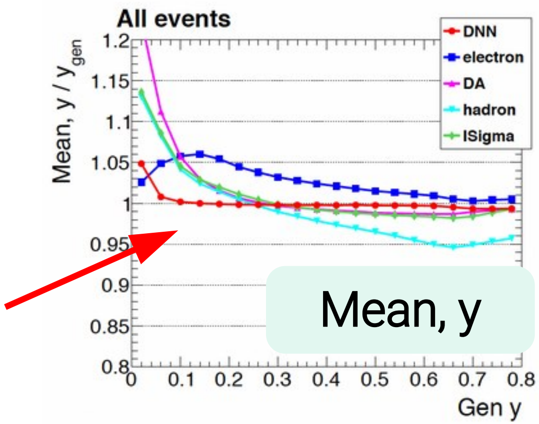


DNN resolution outperforms all conventional methods.

Small improvement in Q^2 (because of calibration)

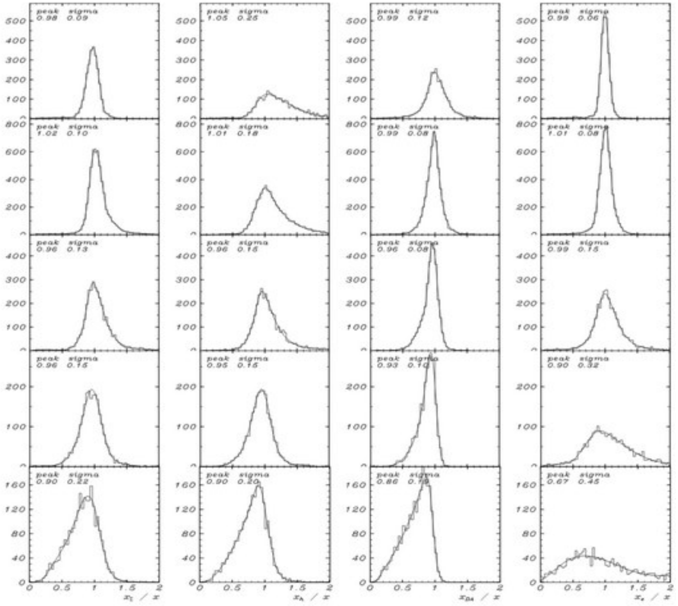
Good improvement at low-y (high-x)

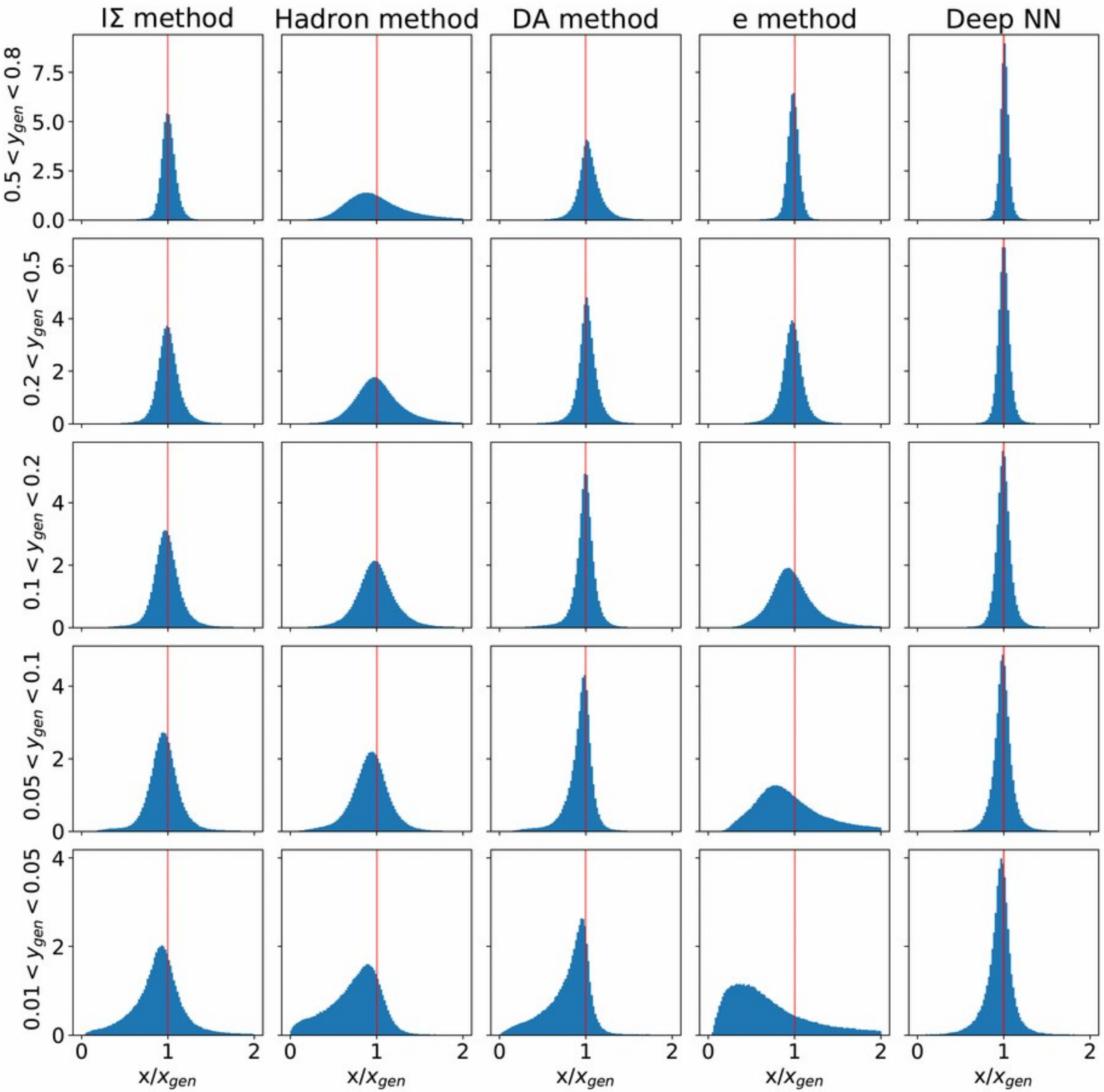
DNN learns to calibrate and to account for acceptance effects



Resolution on x in bins of y.

H1, Bassler, Bernardi, 1994
[NIM A 361 (1995) 197]

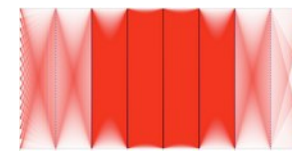


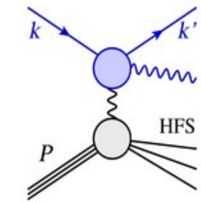


Summary (arXiv:2110.05505)

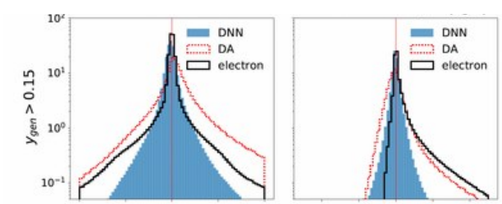
We have applied modern machine learning techniques to reconstruct the kinematics of Deep Inelastic Scattering

Our method includes observables that allow QED radiation effects to be significantly reduced in the reconstruction





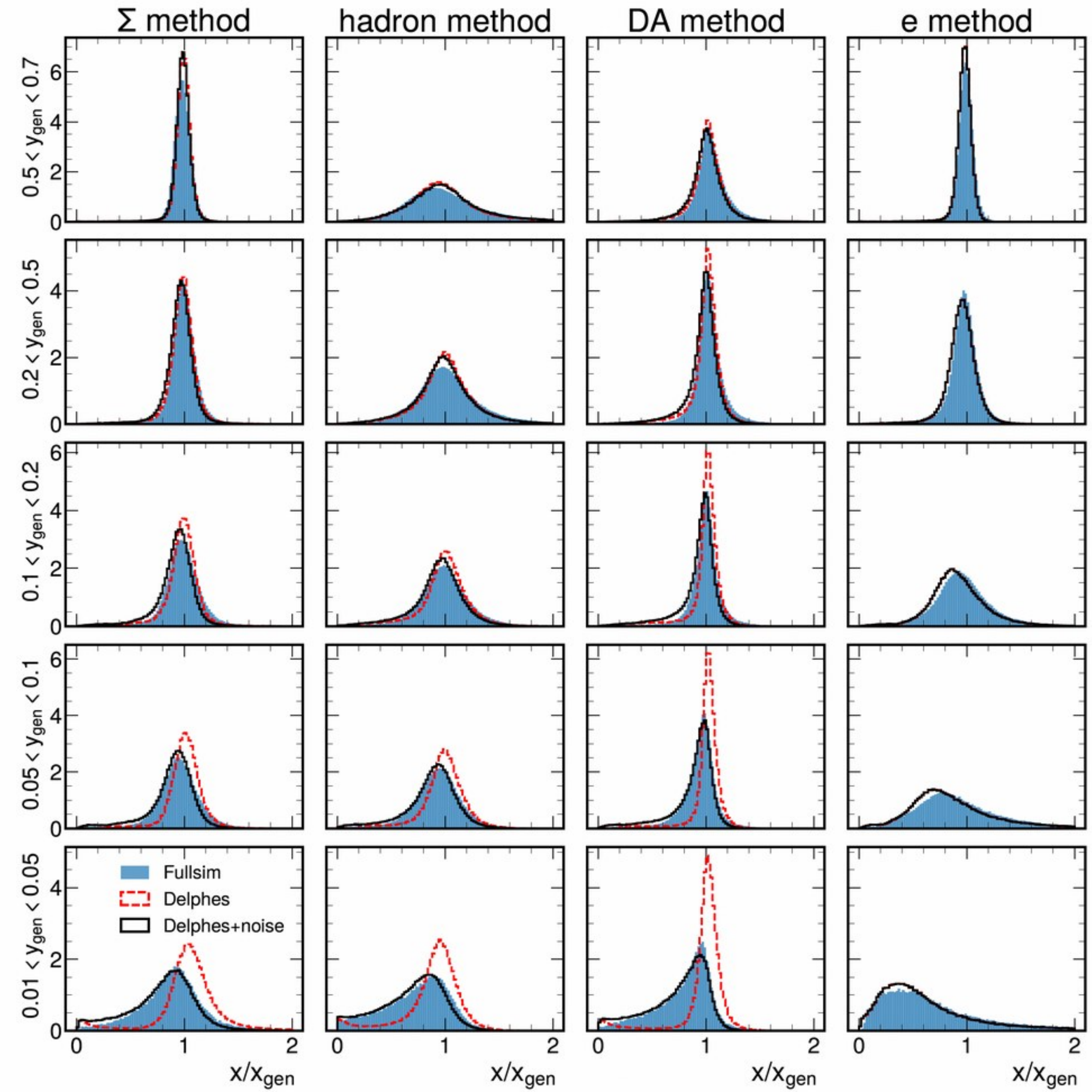
The DNN approach outperforms conventional reconstruction methods in the full range of y for $Q^2 > 200$ GeV², and is most successful for y and x at high-x



An EIC fast-simulation was benchmarked against a real experiment at HERA (H1)

Subtle details of a full simulation may be important for conclusions (calibration, noise, acceptance, trigger etc…)
→ Ask HERA experiments, they are readily available thanks to long-term data and software preservation efforts

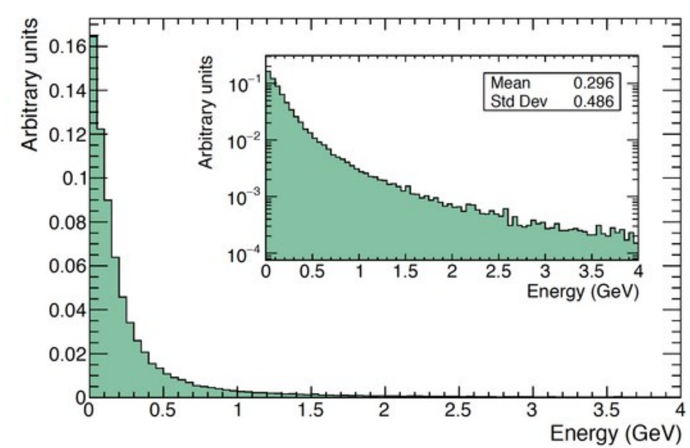
H1 fastsim vs fullsim



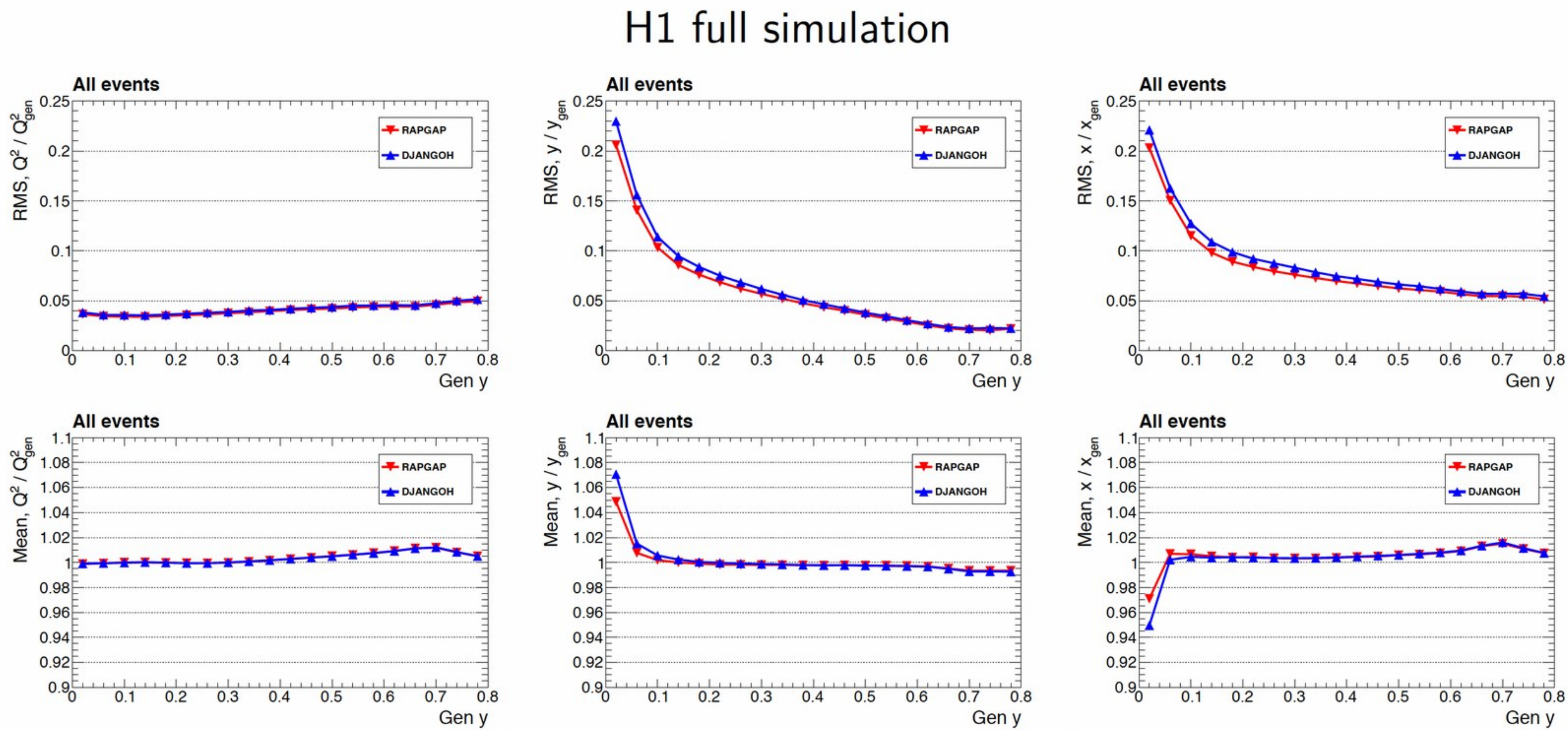
DELPHES fastsim does not include some detector effects (inefficient ('dead') detector components, dead material, noise hits in the calorimeter,e tc…

Adding an additional noise-resolution-acceptance component (simple ad-hoc model) to the fastsim gives much better agreement with fullsim.

Trandom::Landau , mu=0, sigma = 0.05 GeV

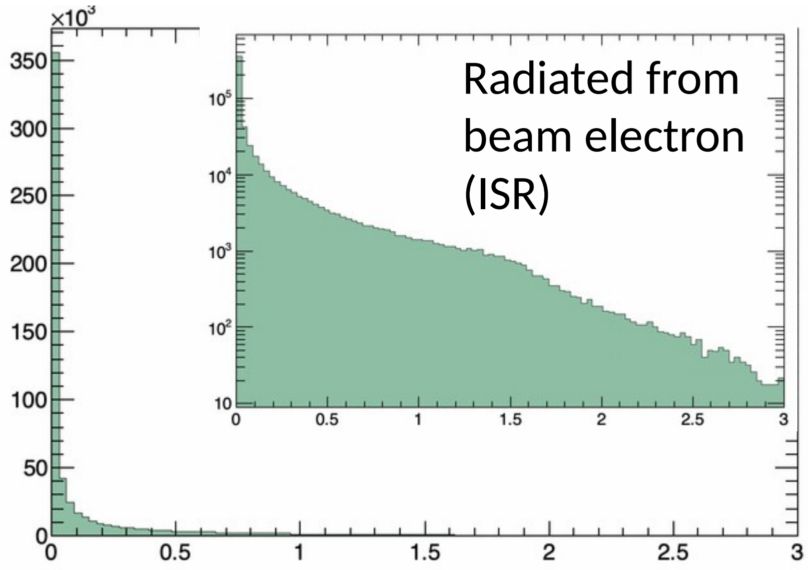


Tow independent MC generators: Djangoh & Rapgap

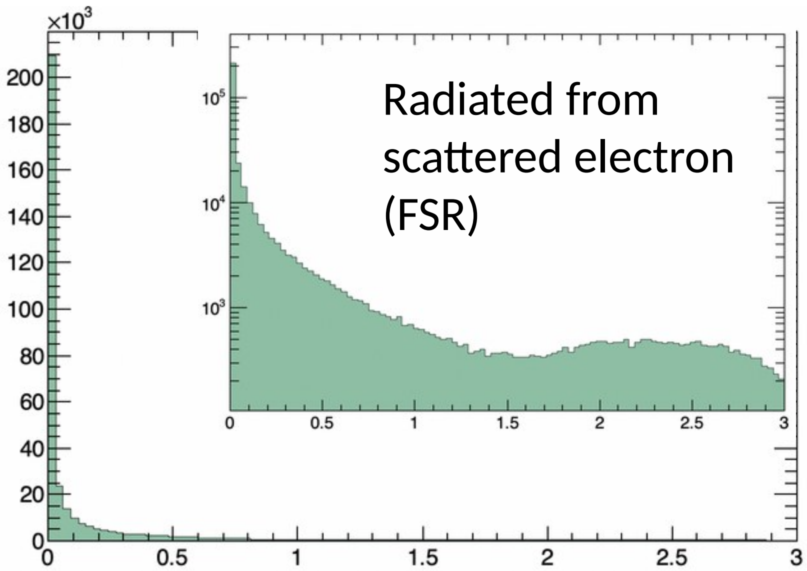


Test of using DNN trained in RAPGAP sample to make predictions in DHANGOH sample.

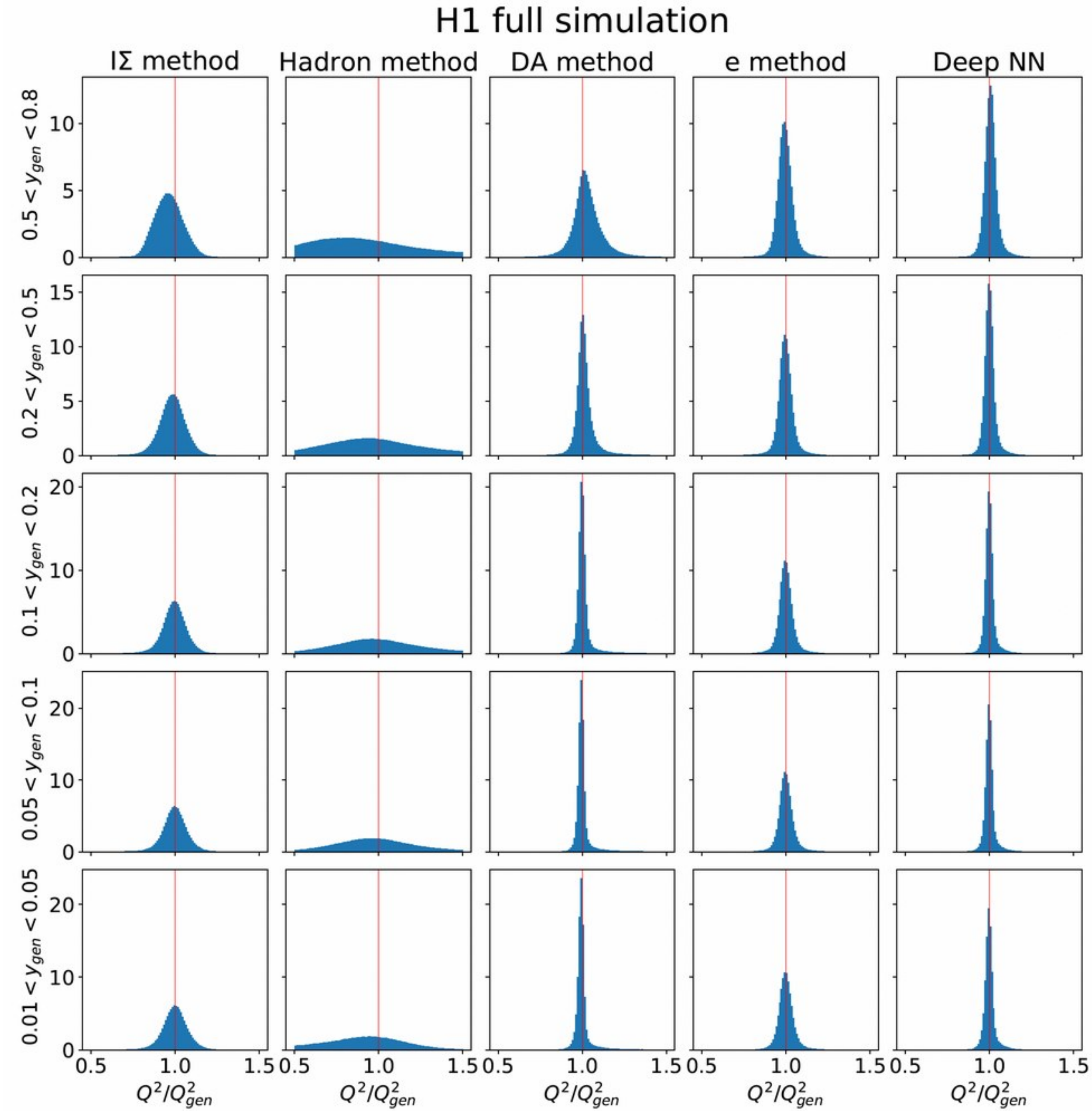
QED radiation



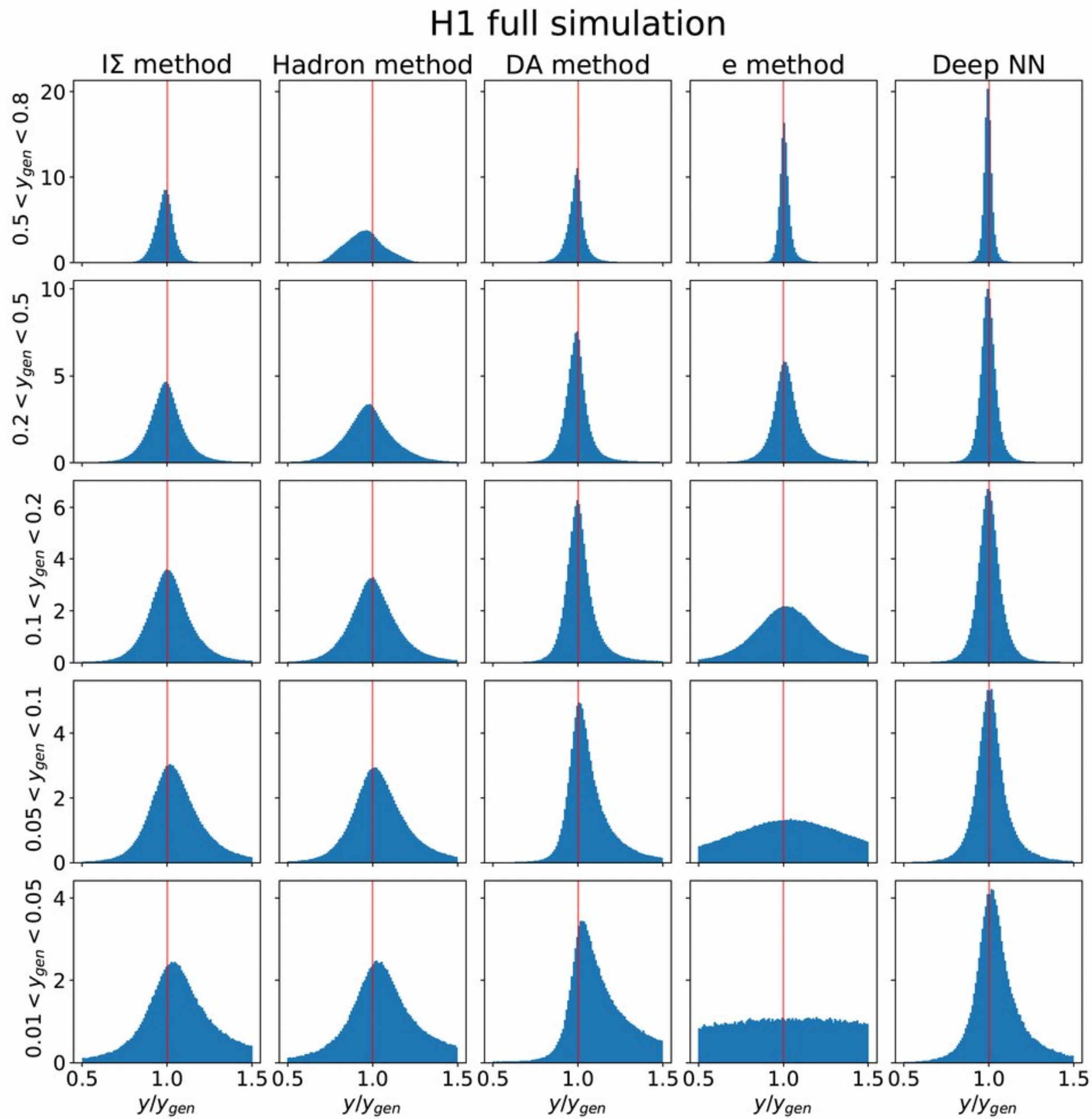
Angle between electron and photon (radians)



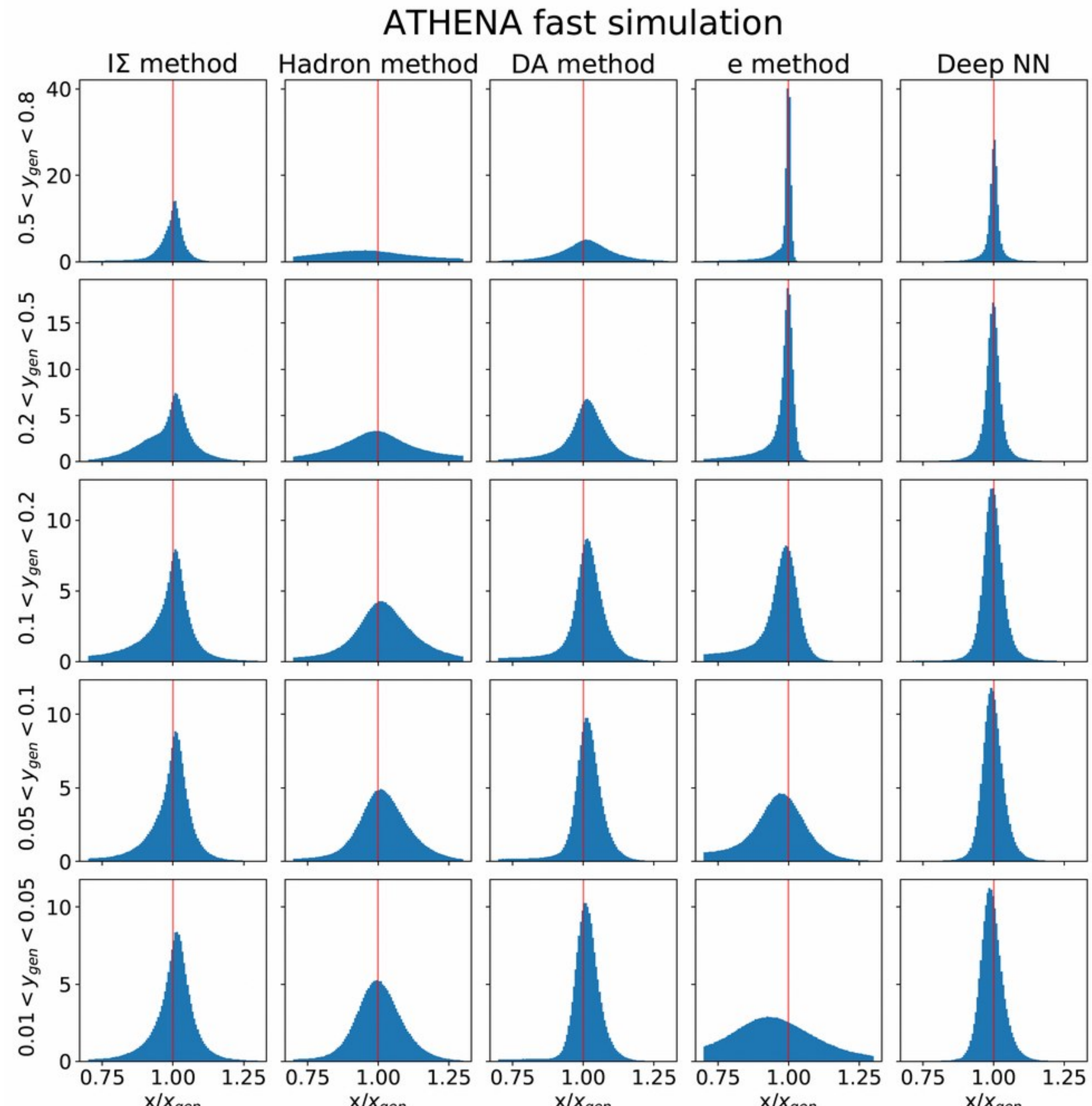
Angle between electron and photon (radians)



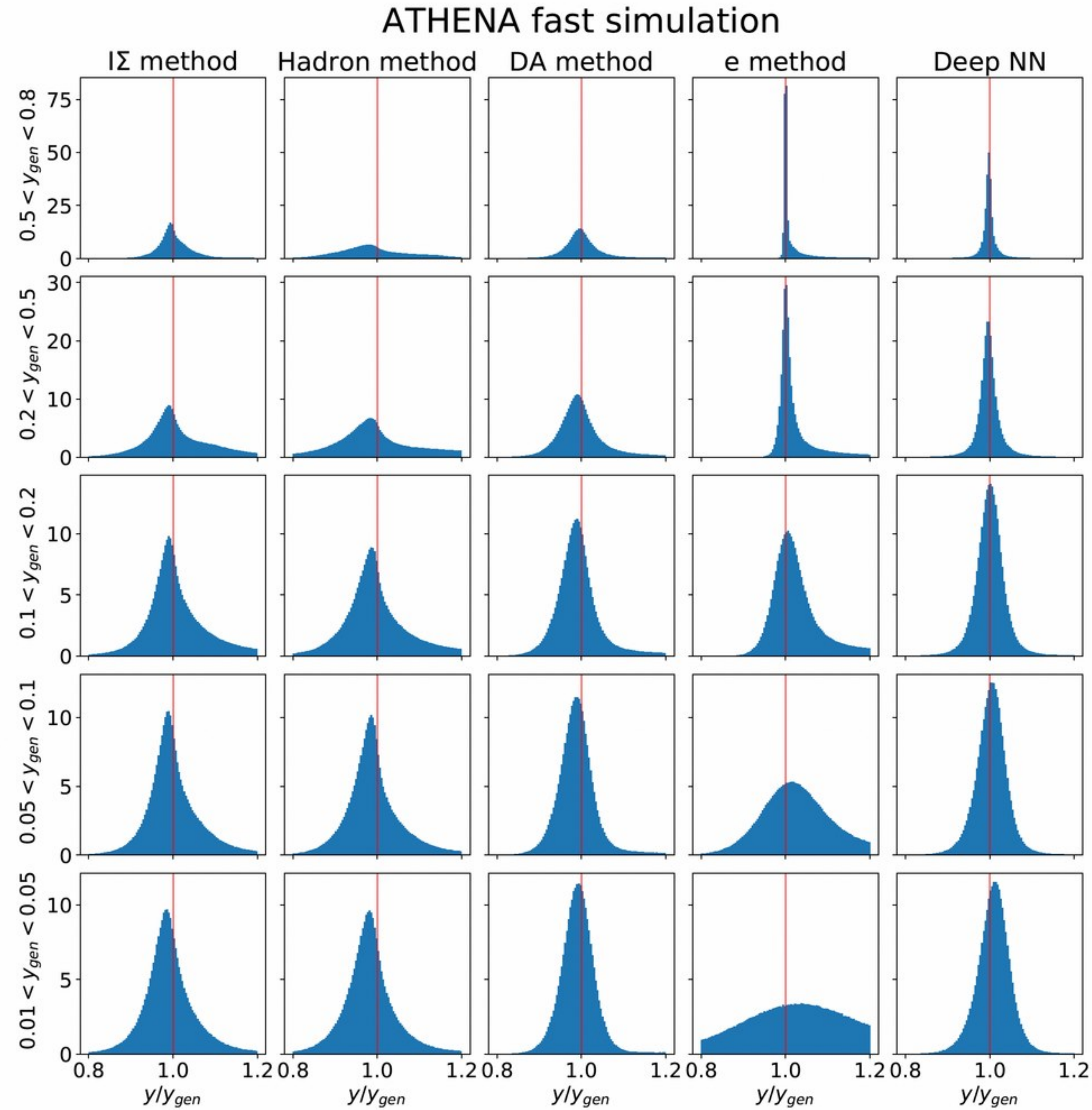
H1 full simulation
IΣ method
Hadron method
DA method
e method
Deep NN
0.5 < y_{gen} < 0.8
0.2 < y_{gen} < 0.5
0.1 < y_{gen} < 0.2
0.05 < y_{gen} < 0.1
0.01 < y_{gen} < 0.05
Q^2/Q^2_{gen}



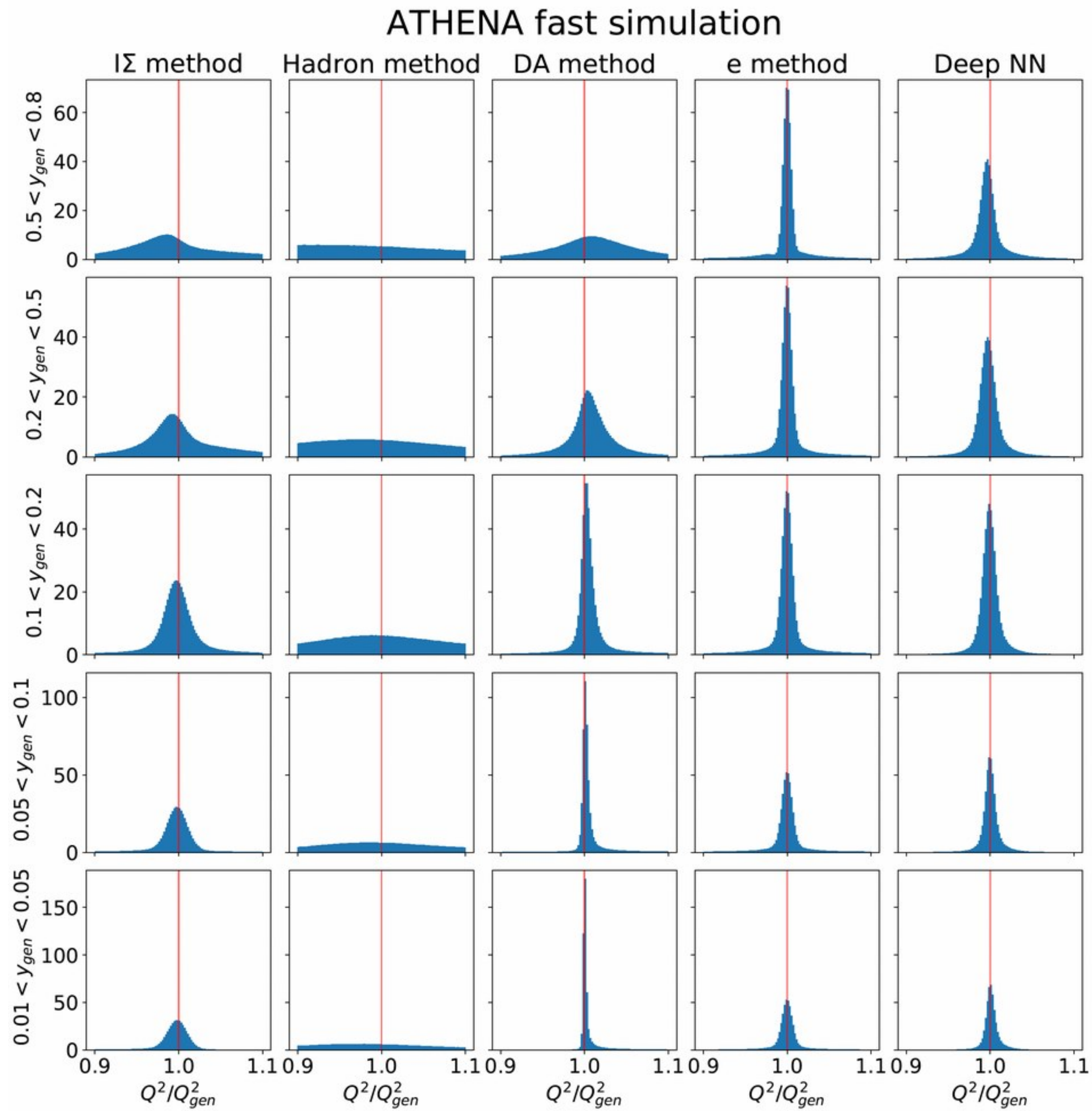
H1 full simulation
IΣ method
Hadron method
DA method
e method
Deep NN
$0.5 < y_{gen} < 0.8$
$0.2 < y_{gen} < 0.5$
$0.1 < y_{gen} < 0.2$
$0.05 < y_{gen} < 0.1$
$0.01 < y_{gen} < 0.05$
y/y_{gen}



ATHENA fast simulation
IΣ method
Hadron method
DA method
e method
Deep NN
0.5 < y_gen < 0.8
0.2 < y_gen < 0.5
0.1 < y_gen < 0.2
0.05 < y_gen < 0.1
0.01 < y_gen < 0.05
x/x_gen



ATHENA fast simulation
IΣ method
Hadron method
DA method
e method
Deep NN
$0.5 < y_{gen} < 0.8$
$0.2 < y_{gen} < 0.5$
$0.1 < y_{gen} < 0.2$
$0.05 < y_{gen} < 0.1$
$0.01 < y_{gen} < 0.05$
y/y_{gen}



ATHENA fast simulation
IΣ method
Hadron method
DA method
e method
Deep NN
$0.5 < y_{gen} < 0.8$
$0.2 < y_{gen} < 0.5$
$0.1 < y_{gen} < 0.2$
$0.05 < y_{gen} < 0.1$
$0.01 < y_{gen} < 0.05$
0.9
1.0
1.1
Q^2/Q^2_{gen}