Reconstructing the Kinematics of Deep Inelastic Scattering with Deep Learning

Miguel Arratia,[a,b] Daniel Britzger,[c] Owen Long,[*a] and Benjamin Nachman[d,e]

[a] *Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA*

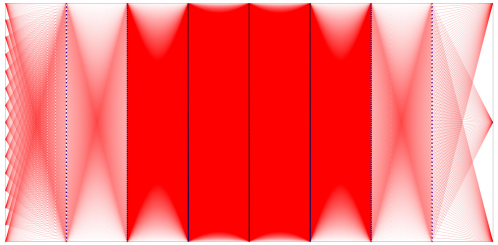
[b] *Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA*

[c] *Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany*

[d] *Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[e] *Berkeley Institute for Data Science, University of California, Berkeley, CA 94720, USA*

https://arxiv.org/abs/2110.05505



UNIVERSITY OF CALIFORNIA UC RIVERSIDE

Overview

- Many methods for reconstructing neutral current DIS variables (Q^2, y, x) using scattered electron and Hadronic Final State (HFS) are complementary. Which method is best depends on Q^2, y, x and the strength of QED radiation.

- QED radiation can produce tails in Q^2, y, x distributions if it is not identified and handled in the event.

- Machine learning can address both of these issues to provide a single reconstruction method that optimally uses all available information in the event.

QED Radiation

We use the following *practical definitions* of QED radiation

Initial State Radiation (ISR): the radiated photon is closer to the electron beam direction.

Final State Radiation (FSR): the radiated photon is closer to the scattered electron direction.

We use the RAPGAP MC generator for our studies.

RAPGAP implements HERACLES for higher-order QED radiation.

$$p_z^{\mathrm{bal}} = 1 - \frac{\Sigma_e + \Sigma}{2\ E_0}$$

Zero if no ISR, >0 with ISR

$$p_T^{\mathrm{bal}} = 1 - \frac{p_{T,e}}{T}$$

Zero if no FSR, >0 with FSR

Definition of DNN regression learning targets or the "true" Q^2, y, x.

Standard definitions for events with no QED radiation in terms of beam electron ℓ , scattered electron ℓ', and beam proton p 4-vectors.

For events with QED radiation

Use post-ISR beam electron 4-vector

Use pre-FSR scattered electron 4-vector

$$q \equiv \ell - \ell'$$

$$Q^2 = -q \cdot q$$

$$y = \frac{p \cdot q}{p \cdot \ell}$$

$$x = -\frac{q \cdot q}{2\ p \cdot q}$$

Machine Learning strategy

Provide complete reconstruction information on the scattered electron and the Hadronic Final State as inputs.

Provide additional event observables that can indicate the presence of a QED radiation photon (ISR or FSR) and help quantify the radiation effects.

Feed all useful information into a regression **Deep Neural Network** (DNN) that predicts Q^2, y, and x.

DNN inputs

QED radiation:

ISR radiation (4):

E, η, and $\Delta\phi$ of photon in event closest to electron beam direction, where $\Delta\phi$ is w.r.t. the scattered electron.

$$p_z^{\text{bal}} = 1 - \frac{\Sigma_e + \Sigma}{2\ E_0}$$ from event reconstruction

FSR radiation (3):

ECAL energy within $\Delta R<0.4$ around electron divided by the electron track momentum..

Number of ECAL clusters within $\Delta R<0.4$ around electron.

$$p_T^{\text{bal}} = 1 - \frac{p_{T,e}}{T}$$ from event reconstruction

DIS reconstruction (Q^2, y, x):

Scattered electron (3): E, p_T, and p_z.

HFS (3): E, p_T (T), and p_z.

$\Delta\phi$ between the scattered electron and HFS.

The difference $\Sigma_e - \Sigma$.

Training samples: RAPGAP generator and DELPHES fast simulation of ATHENA.
NC DIS with $Q^2_{gen} > 200$ GeV^2.
32 GeV < event $(E-p_z)$ < 40 GeV , (±4 GeV around $2E_e$)

Machine Learning : Deep Neural Network

Sequential network with **8 layers**.

15 inputs from previous slide, transformed to zero mean, unit RMS..

Nodes per layer: 64, 128, 512, 1024, 512, 128, 64, N_{out}.

Activation: relu for first layer, selu for middle.

Adam optimizer, learning rate 10^{-4} or 10^{-5}.

Batch size 128 or 1024.

Samples: 28 million, evenly split for training and validation. $Q^2_{gen} > 200$ GeV^2.

Training converges after ~40 to ~100 **epochs** in around 30 to 60 minutes on GPU machine.

Applied this basic DNN structure to the following tasks:

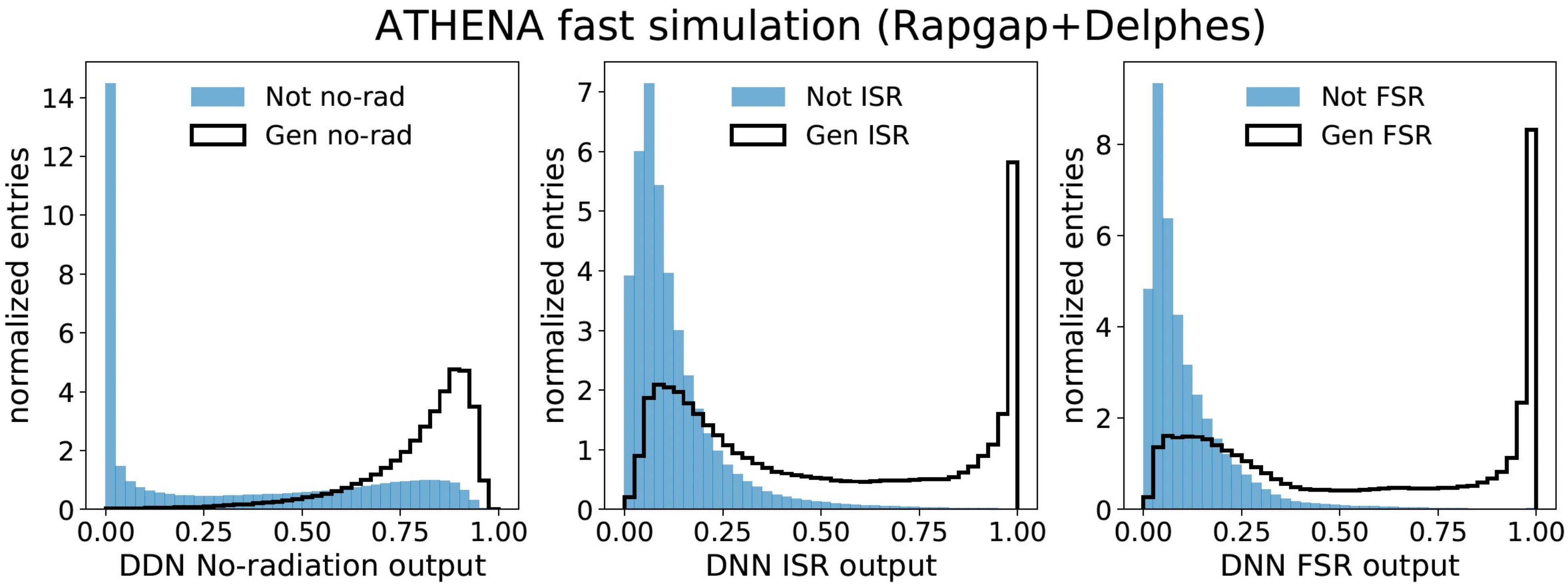
- QED radiation classification : ISR, FSR, NoR
- QED radiation quantification : predict p_z^{bal} and p_T^{bal}.
- DIS reconstruction : Q^2, y, x.

QED radiation : classification DNN

Learning targets: three binary state variables for ISR, FSR, NoR.

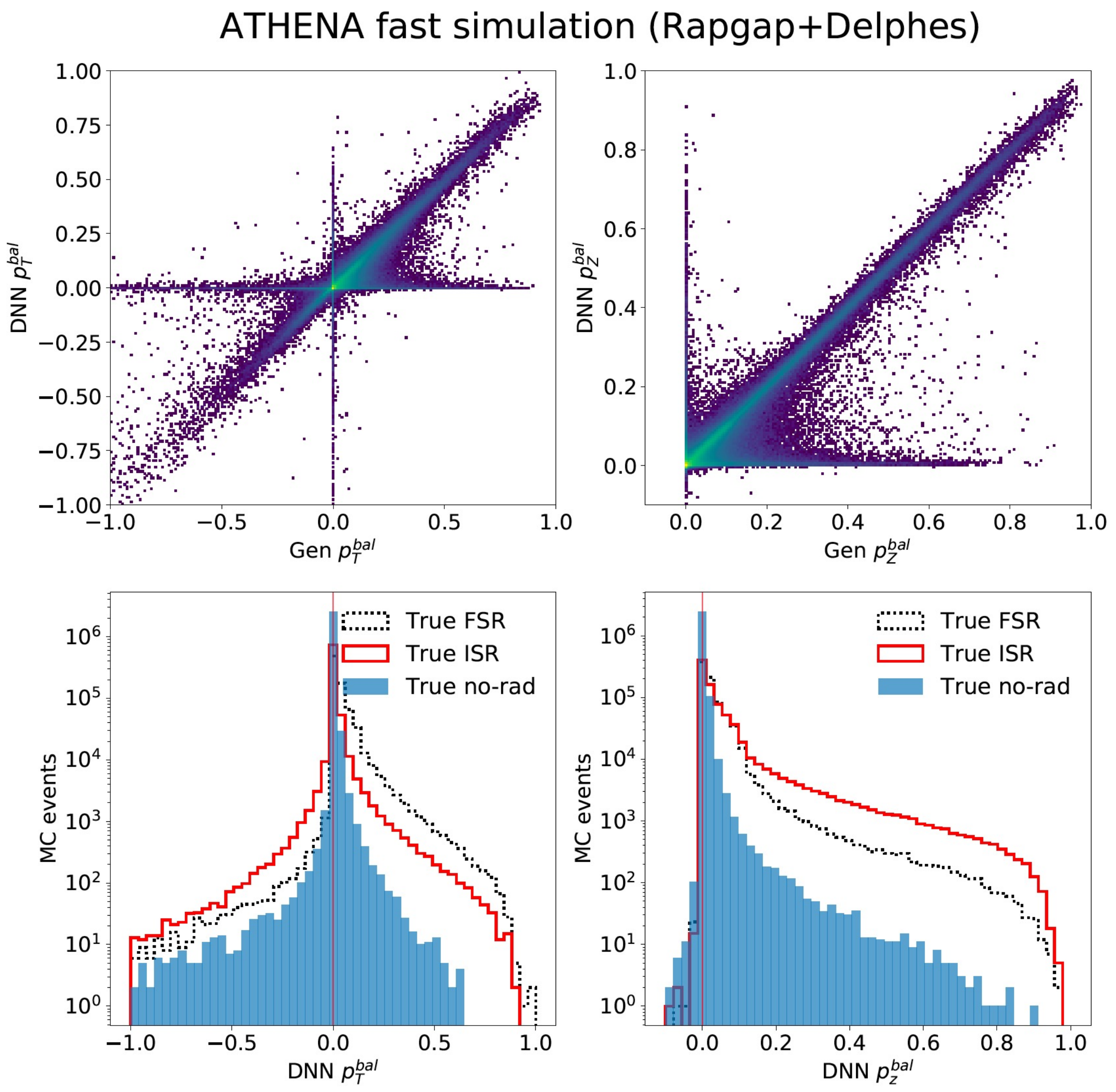
Output layer is 3 nodes for ISR, FSR, NoR with softmax activation (each is 0 to 1, with sum = 1).

Loss function is categorical cross entropy.



Some QED radiation events are strongly identified.

QED radiation : regression DNN for p_z^{bal} and p_T^{bal}



Learning targets are gen values of p_z^{bal} and p_T^{bal}.

Output layer is two nodes for p_z^{bal} and p_T^{bal} with linear activation.

Loss function is Huber.

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

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

Learning targets are log 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 3 QED classification outputs (ISR, FSR, NoR) to 15 other inputs.

2. Add 2 QED regression outputs (p_z^{bal} and p_T^{bal}) 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 log 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.

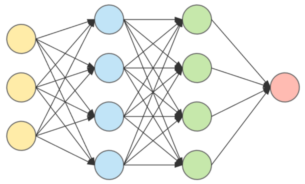
We tried three approaches:

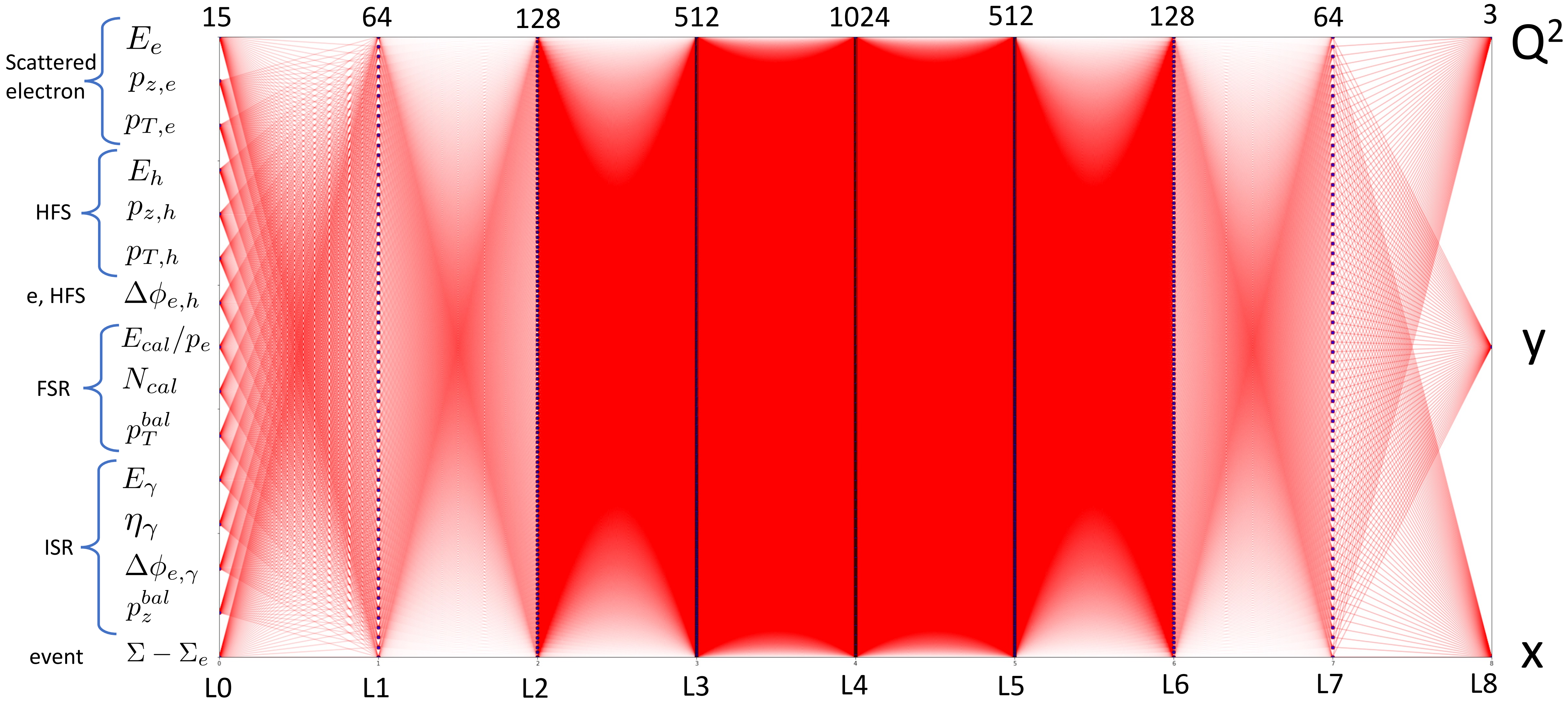
1. Add 3 QED classification outputs (ISR, FSR, NoR) to 15 other inputs.
2. Add 2 QED regression outputs (p_z^{bal} and p_T^{bal}) to 15 other inputs.
3. Use the same 15 inputs as in QED DNNs.

All 3 give essentially identical results!

We choose the simplest option (3).

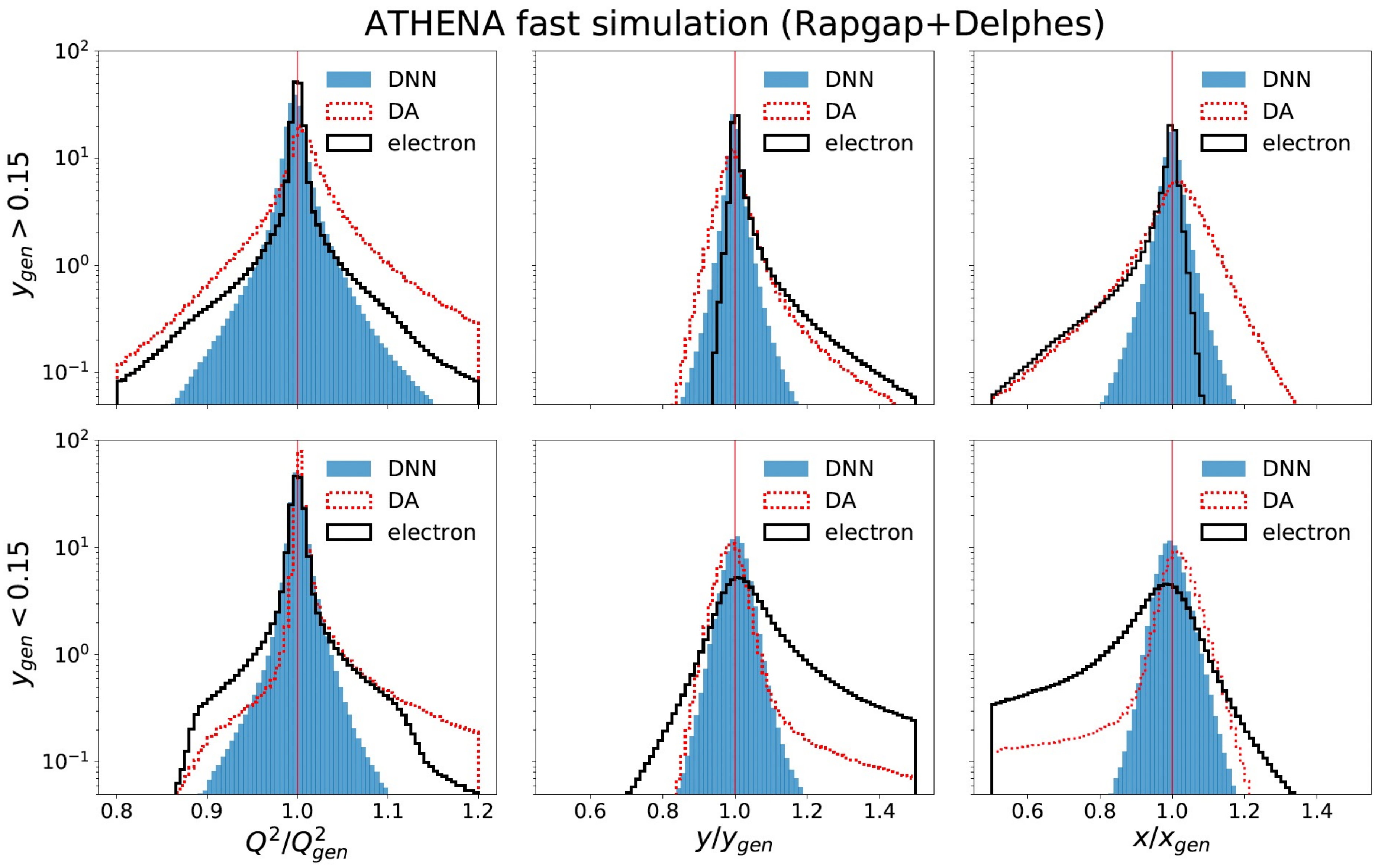
Network diagram





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

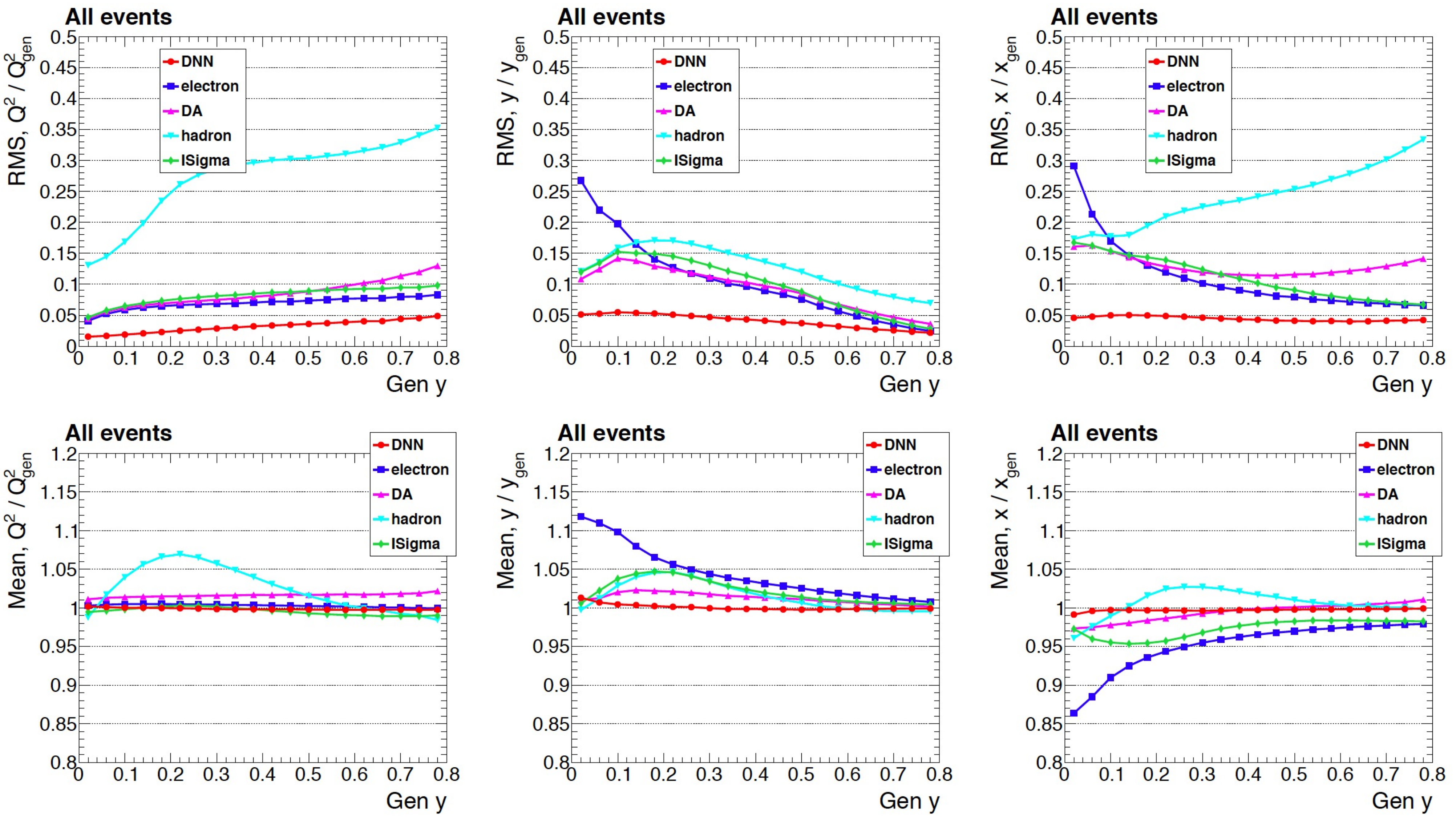
DIS reconstruction : regression DNN for Q^2, y, x



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.

ATHENA fast simulation (Rapgap+Delphes)



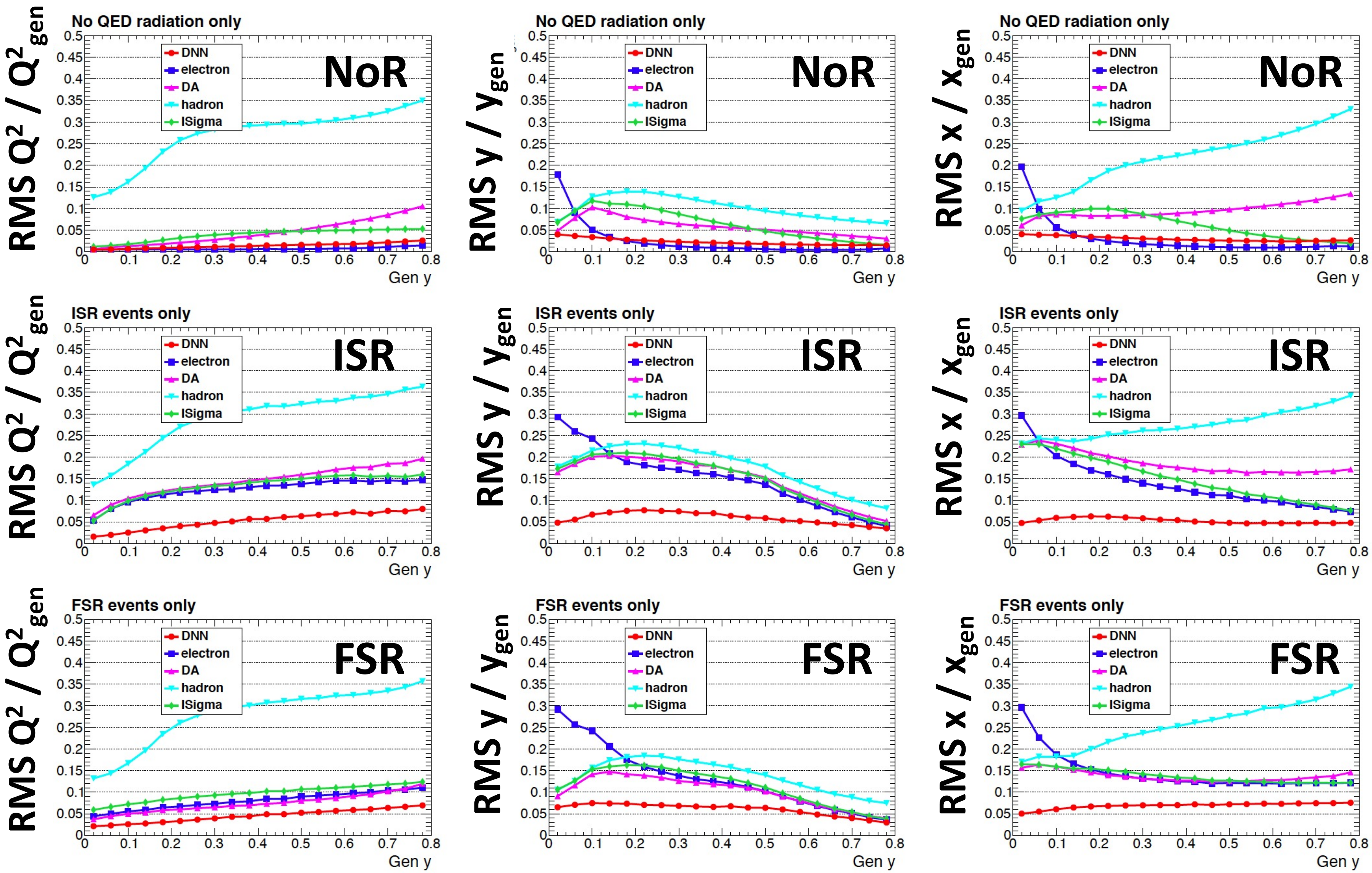
DNN outperforms all conventional reconstruction methods.

DNN has smallest RMS and essentially no bias.

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

ATHENA fast simulation (Rapgap+Delphes)

Resolution (RMS) vs y_{gen}



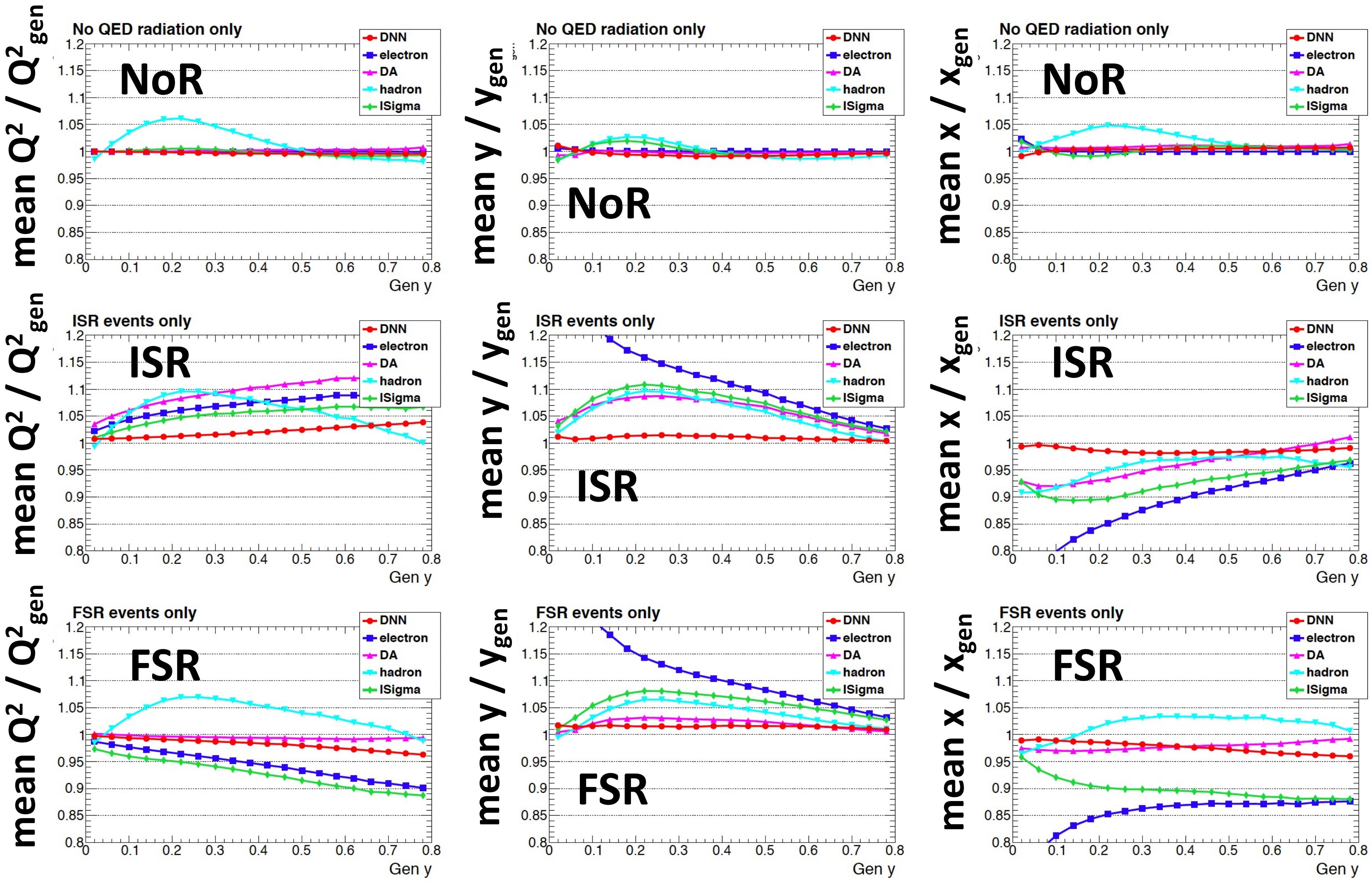
Electron method has better core resolution than DNN for y>0.15 in NoR events (no tails).

DNN resolution much less affected by QED radiation

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

ATHENA fast simulation (Rapgap+Delphes)

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.

Demonstration of DNN with H1 full simulation

HERA beam energies: Ee = 27.6 GeV, Ep = 920 GeV.

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

RAPGAP 3.1 generator (includes HERACLES).

GEANT 3 detector simulation.

Employ standard reconstruction methods for electron and HFS.

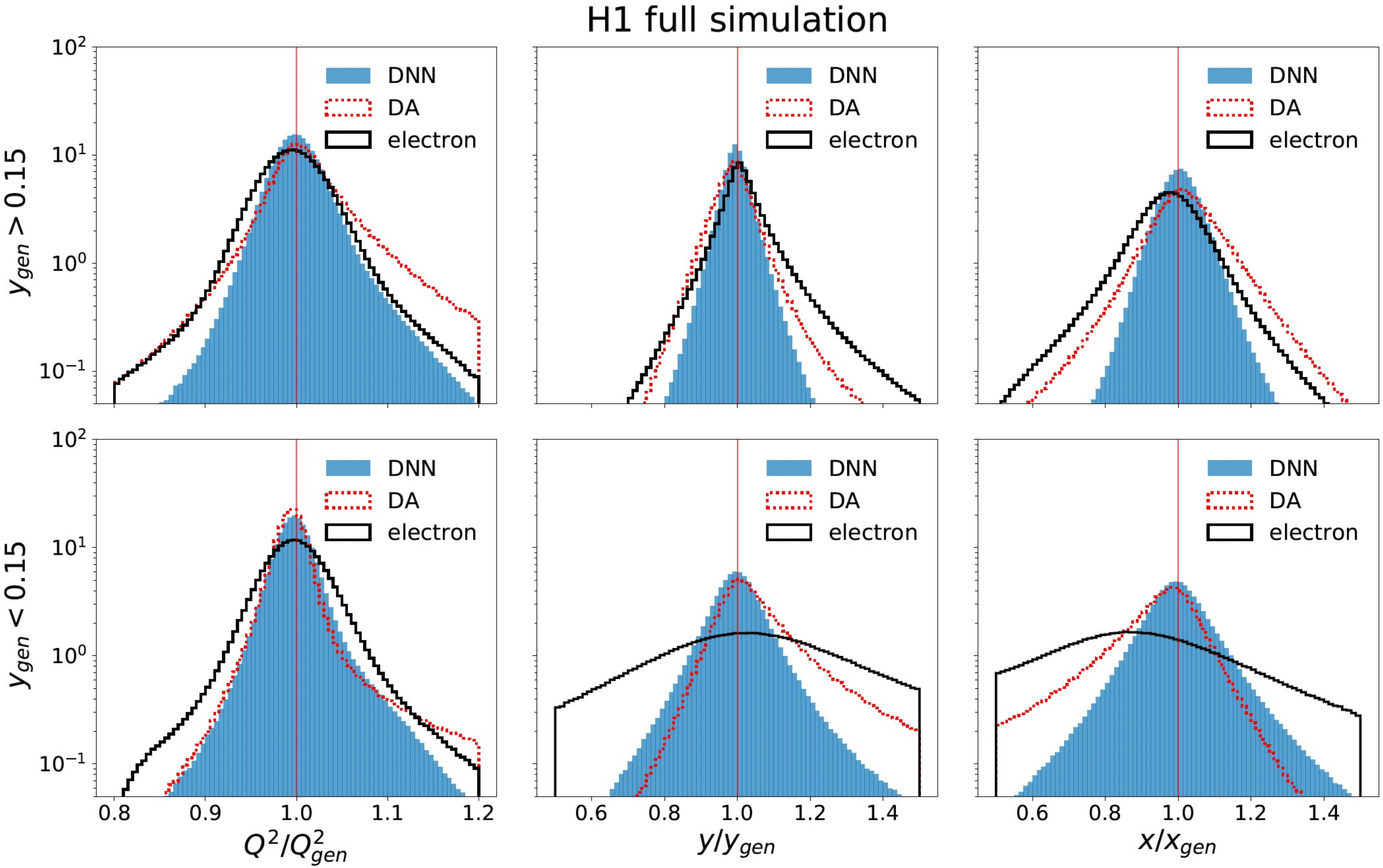
Includes real calorimeter noise.

Includes run-specific conditions.

Event selection:

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

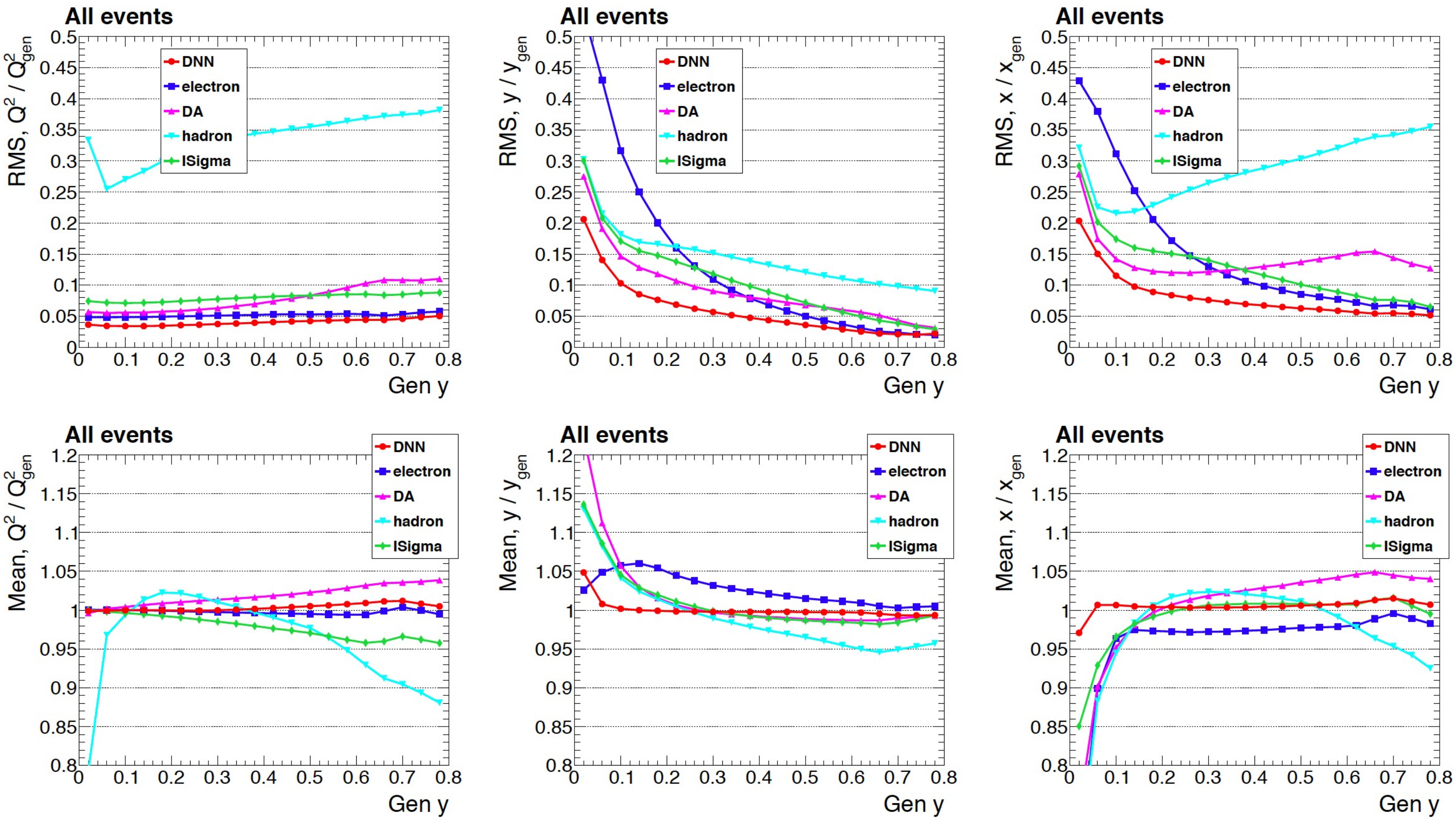
Around 11 million events, evenly split for training and validation.



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

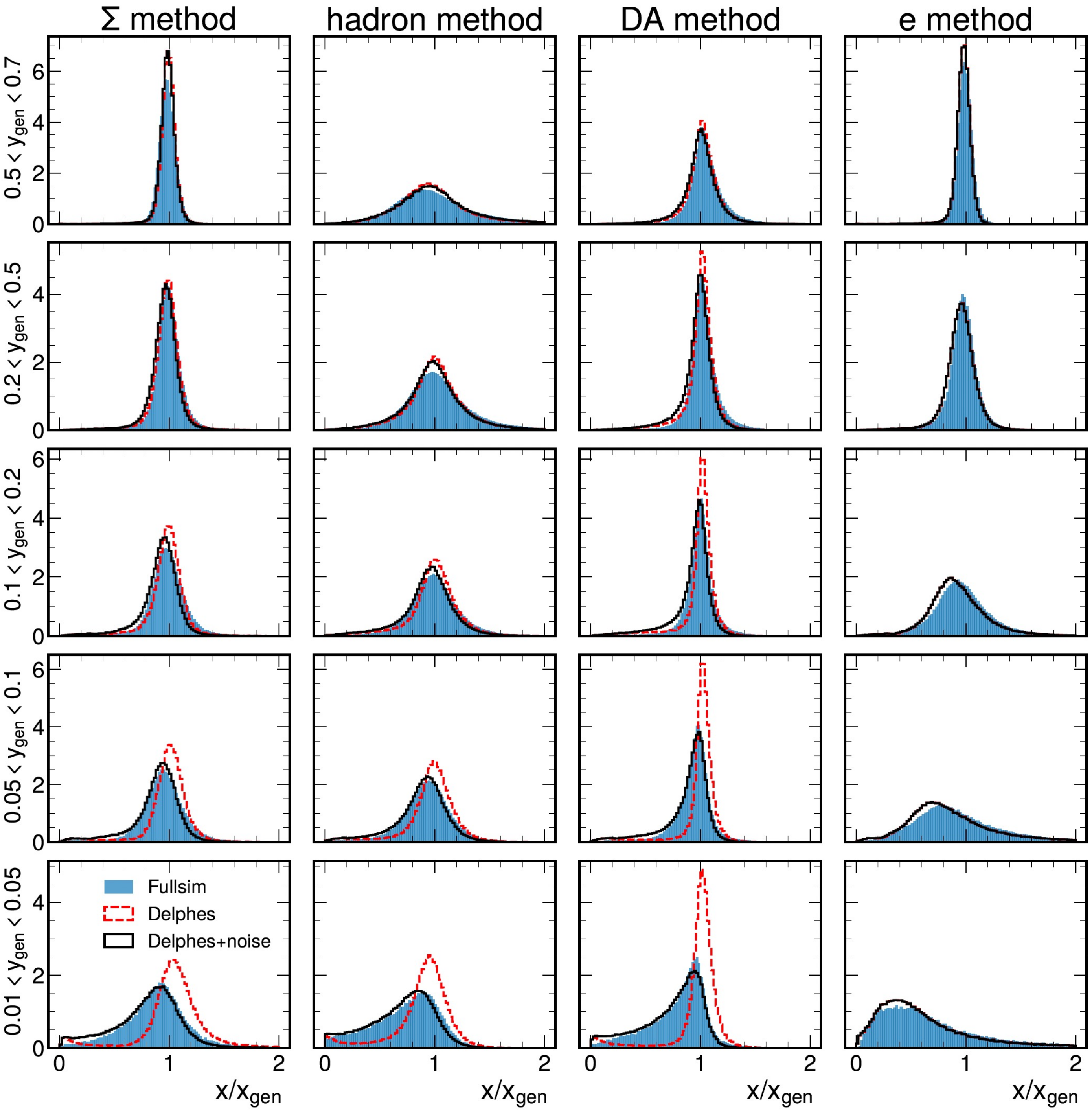
DNN distributions much more symmetric, free of large QED radiation tails.

H1 full simulation



DNN resolution outperforms all conventional methods.

DNN resolution degrades for y<0.15, not seen in ATHENA fastsim study.

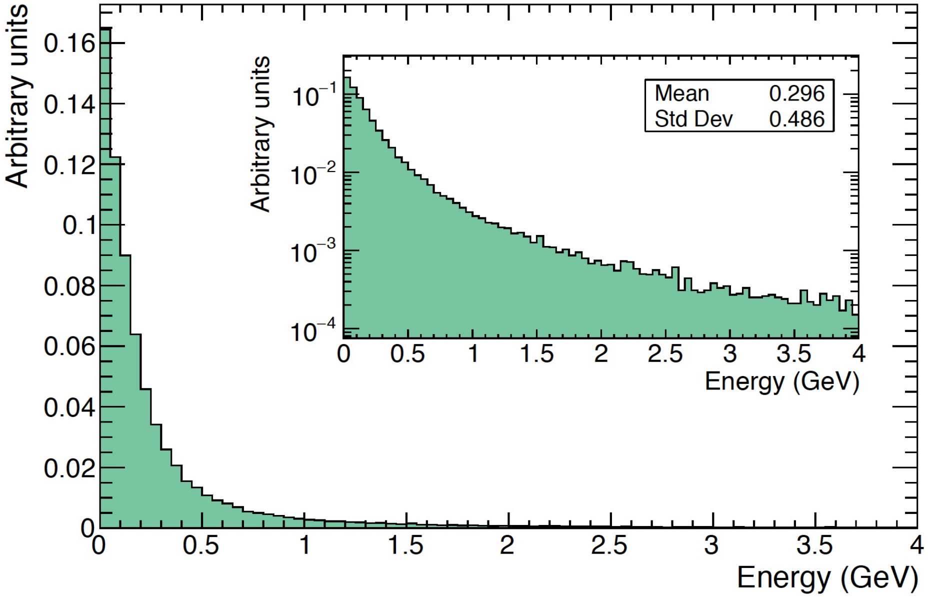


H1 fastsim vs fullsim

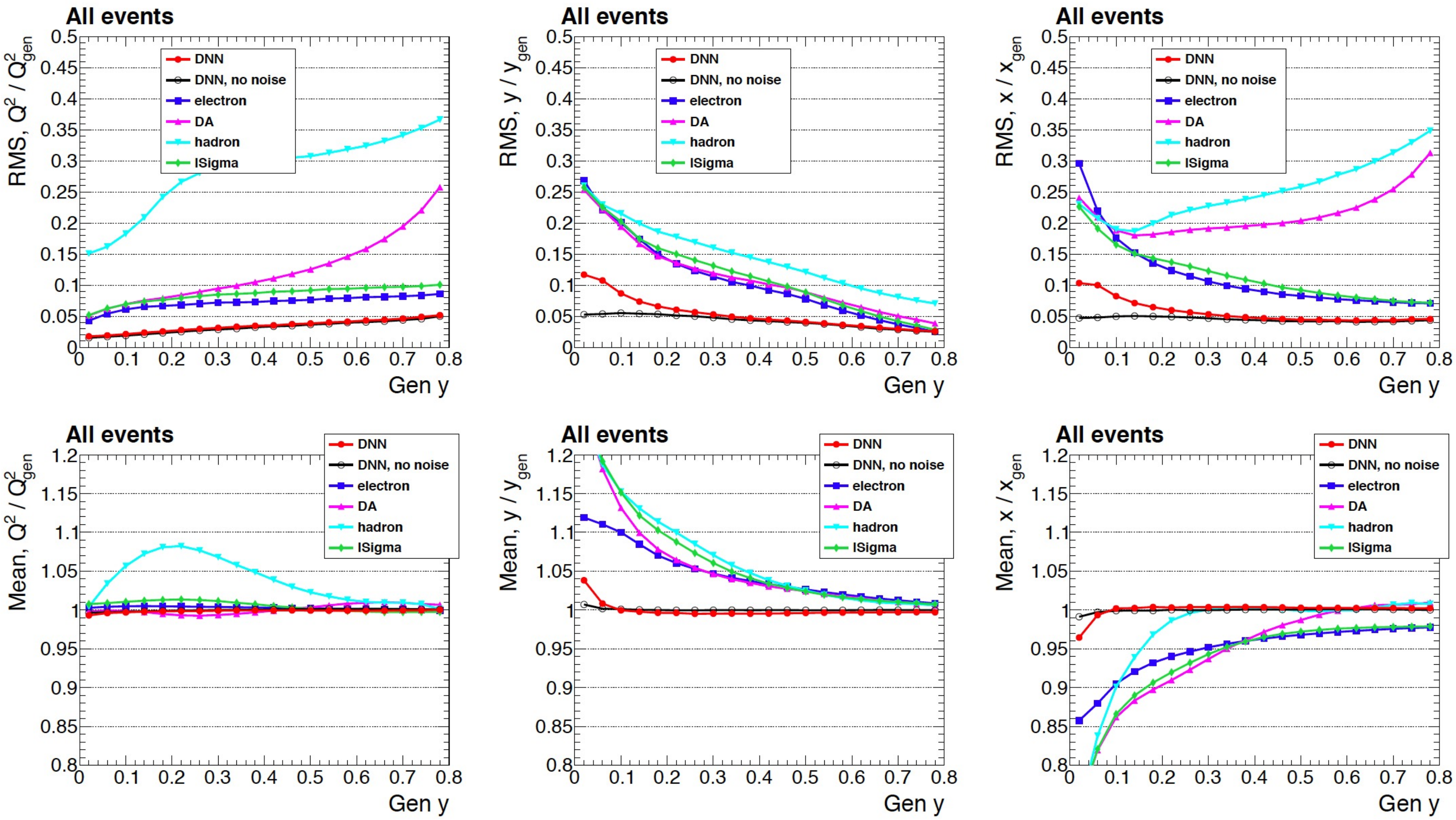
DELPHES fastsim does not include **noise hits in the calorimeter.**

Adding a calorimeter noise component (simple ad-hoc model) to the fastsim gives much better agreement with fullsim.

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



ATHENA fast simulation with additive resolution effect (Rapgap+Delphes)



Test of adding same ad-hoc noise model (conservative) to ATHENA fastsim.

Reran entire procedure, including DNN training, with noise-added fastsim sample.

DNN resolution does degrade for y<0.15.

DNN still significantly outperforms conventional methods.

Conclusions

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

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

The DNN approach outperforms all conventional reconstruction methods in the full range of y for $Q^2 > 200$ GeV^2.

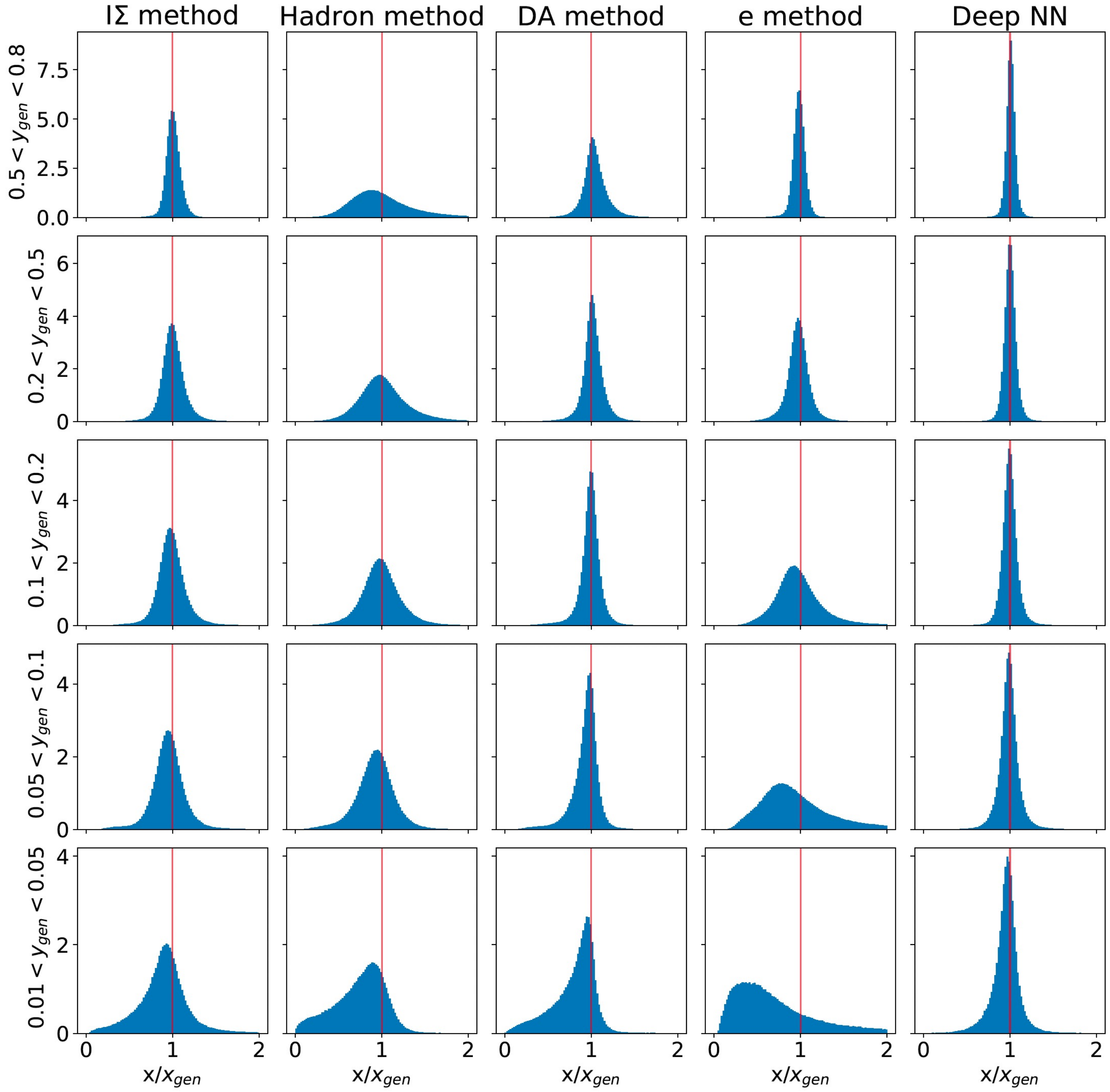
We have demonstrated our method in the full simulation of H1.

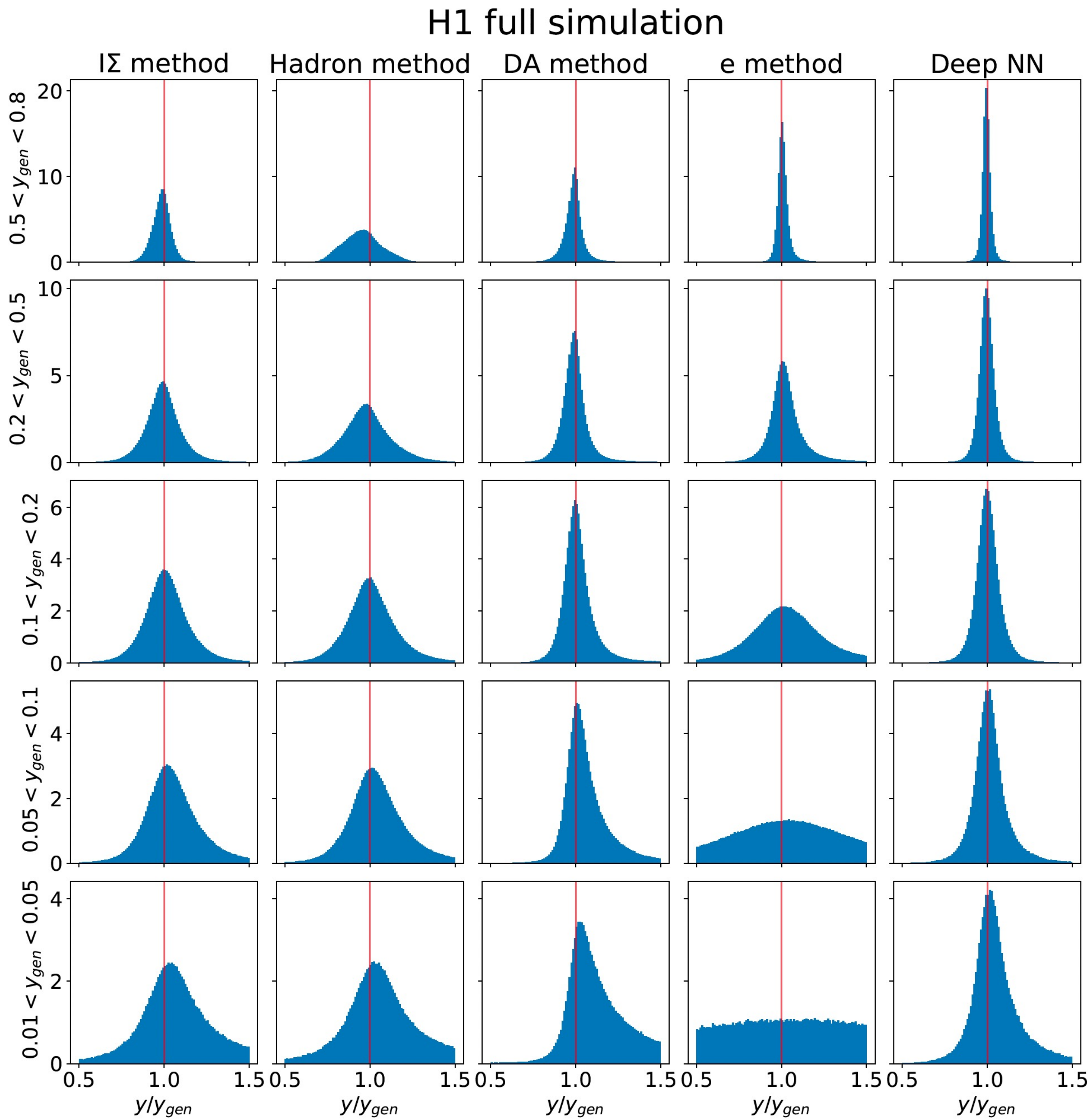
Calorimeter noise is important at low y and drives the resolution on x and y.

https://arxiv.org/abs/2110.05505

Extra Slides

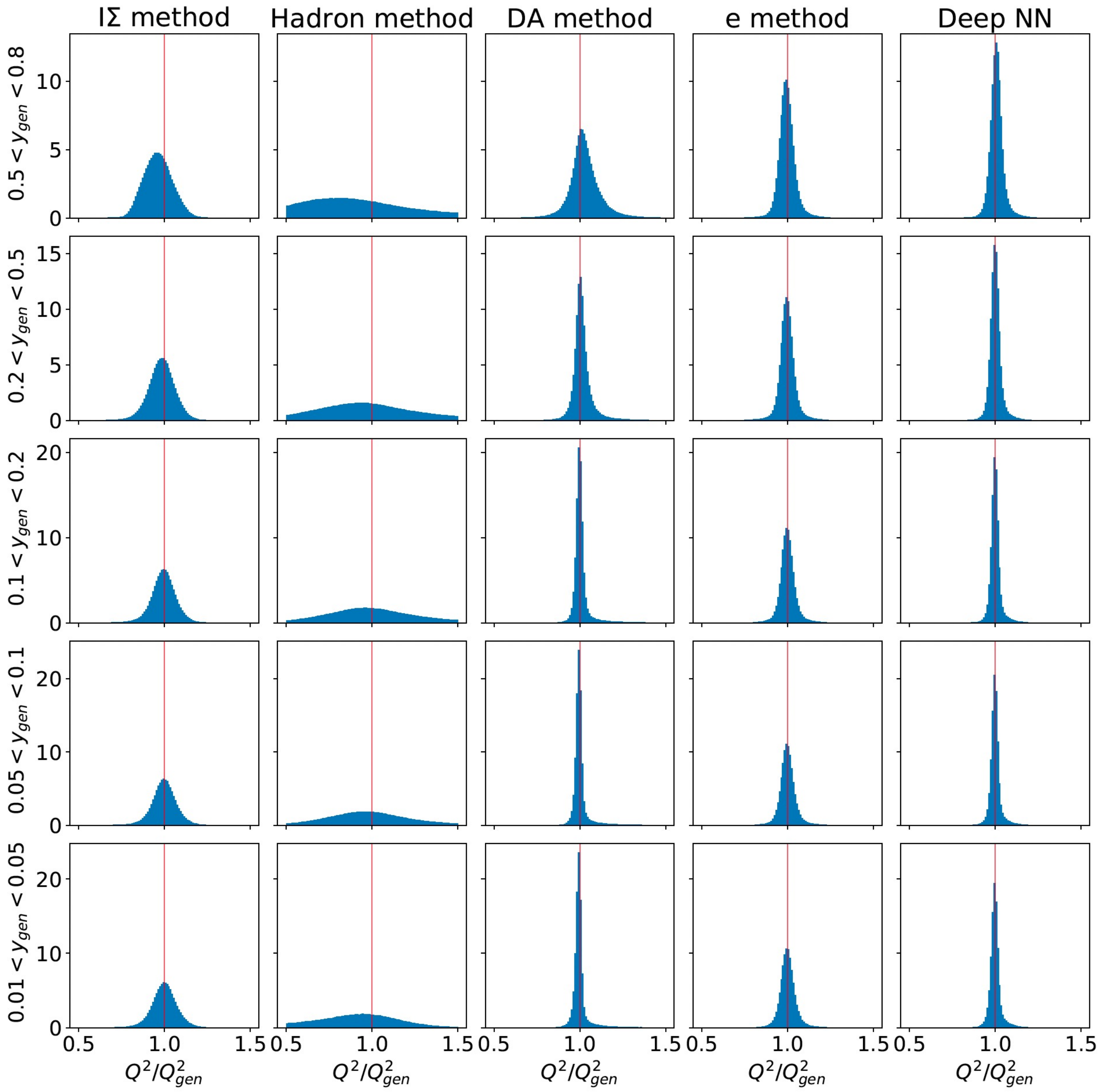
H1 full simulation



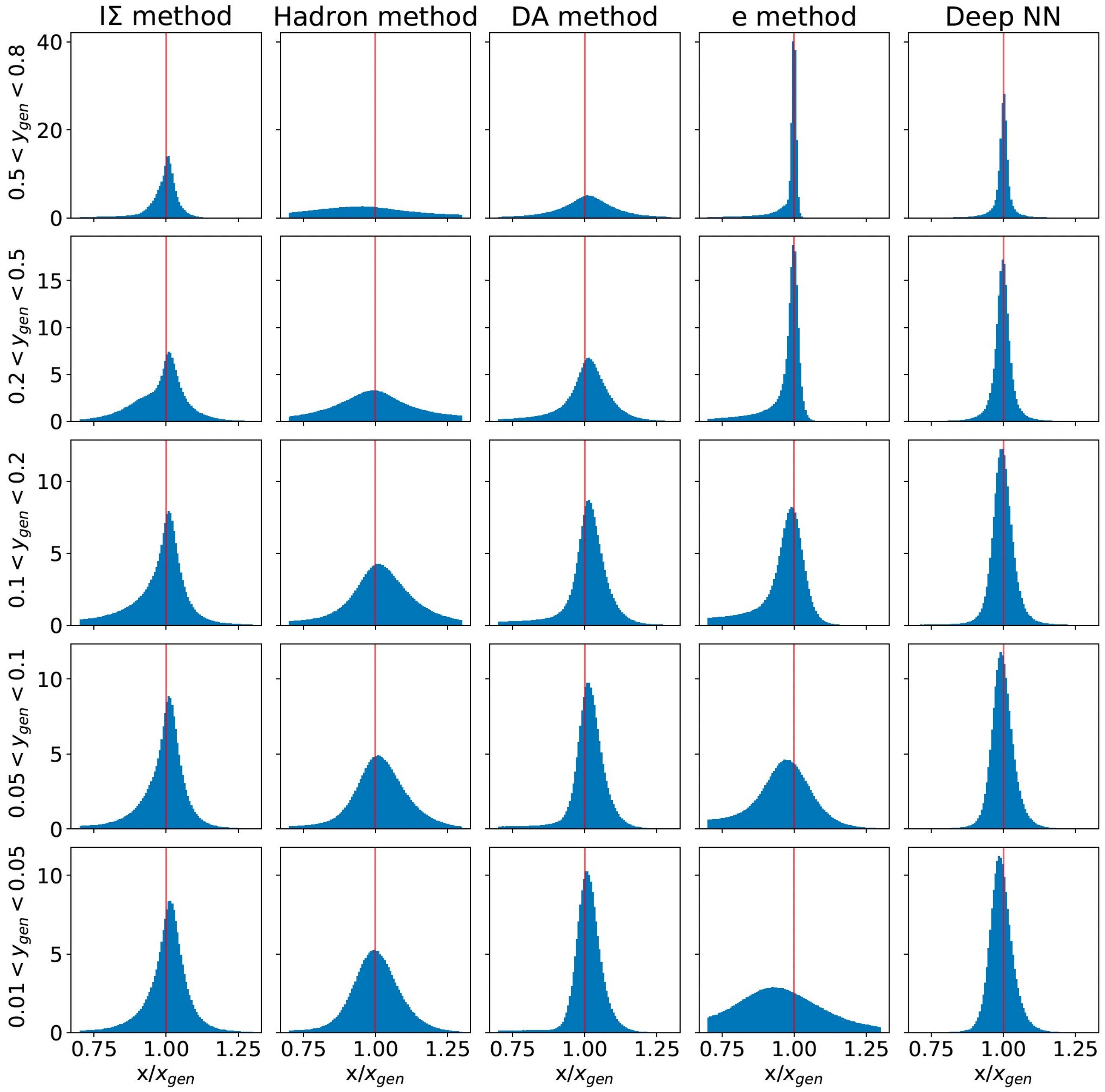


H1 full simulation
IΣ method
Hadron method
DA method
e method
Deep NN
0.5 < y_gen < 0.8
0.2 < y_gen < 0.5
0.1 < y_gen < 0.2
0.05 < y_gen < 0.1
0.01 < y_gen < 0.05
y/y_gen

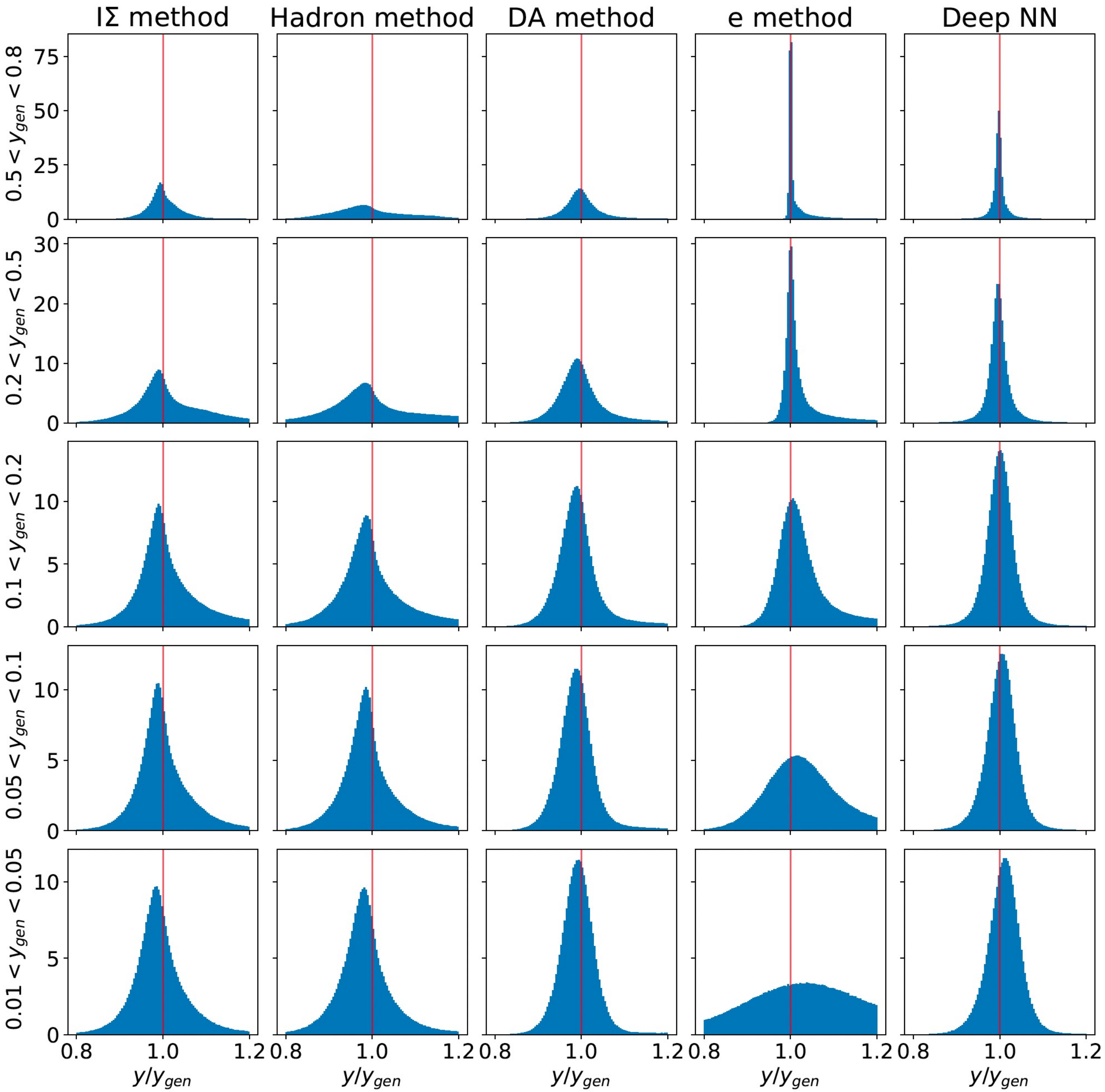
H1 full simulation



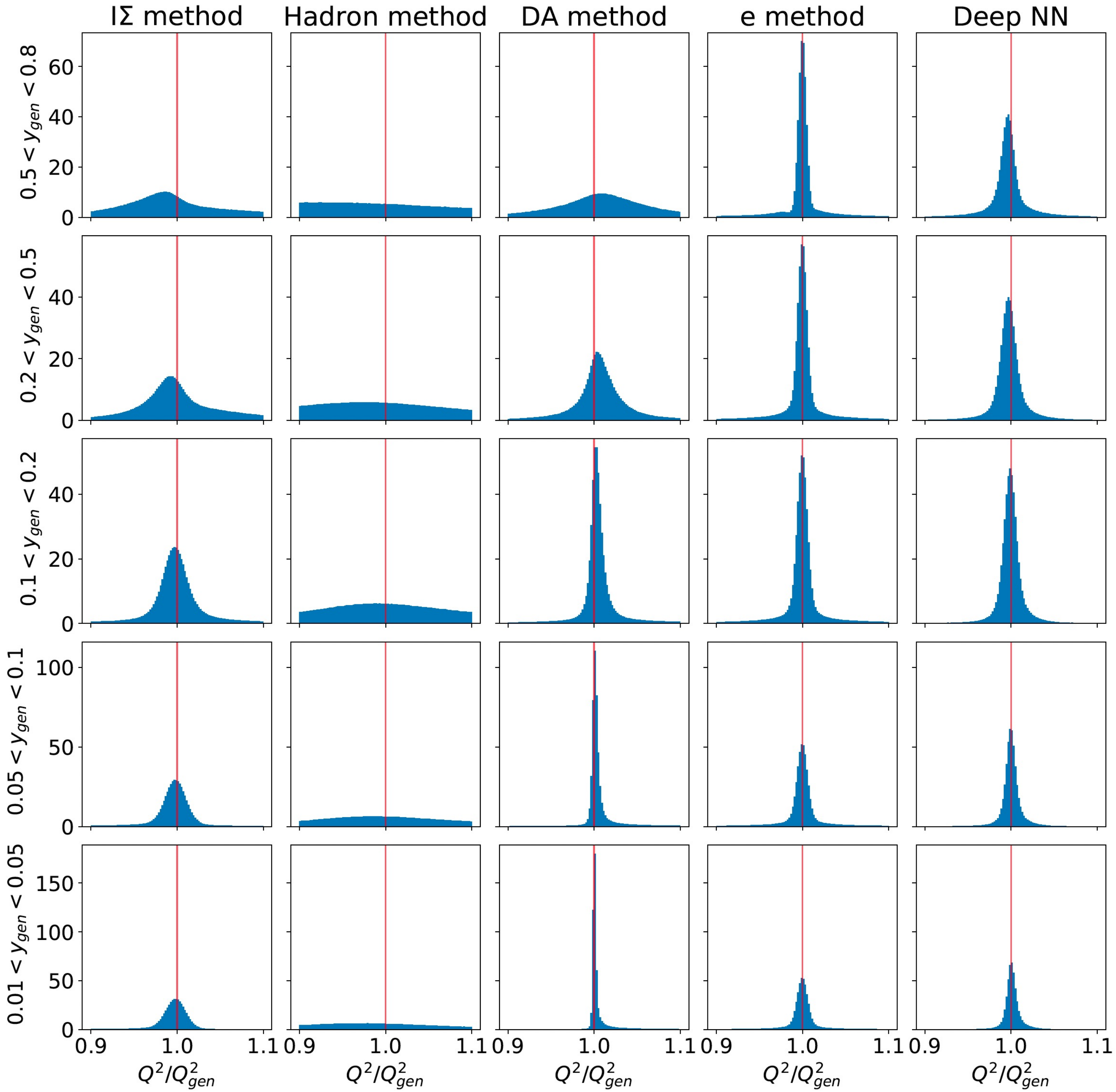
ATHENA fast simulation



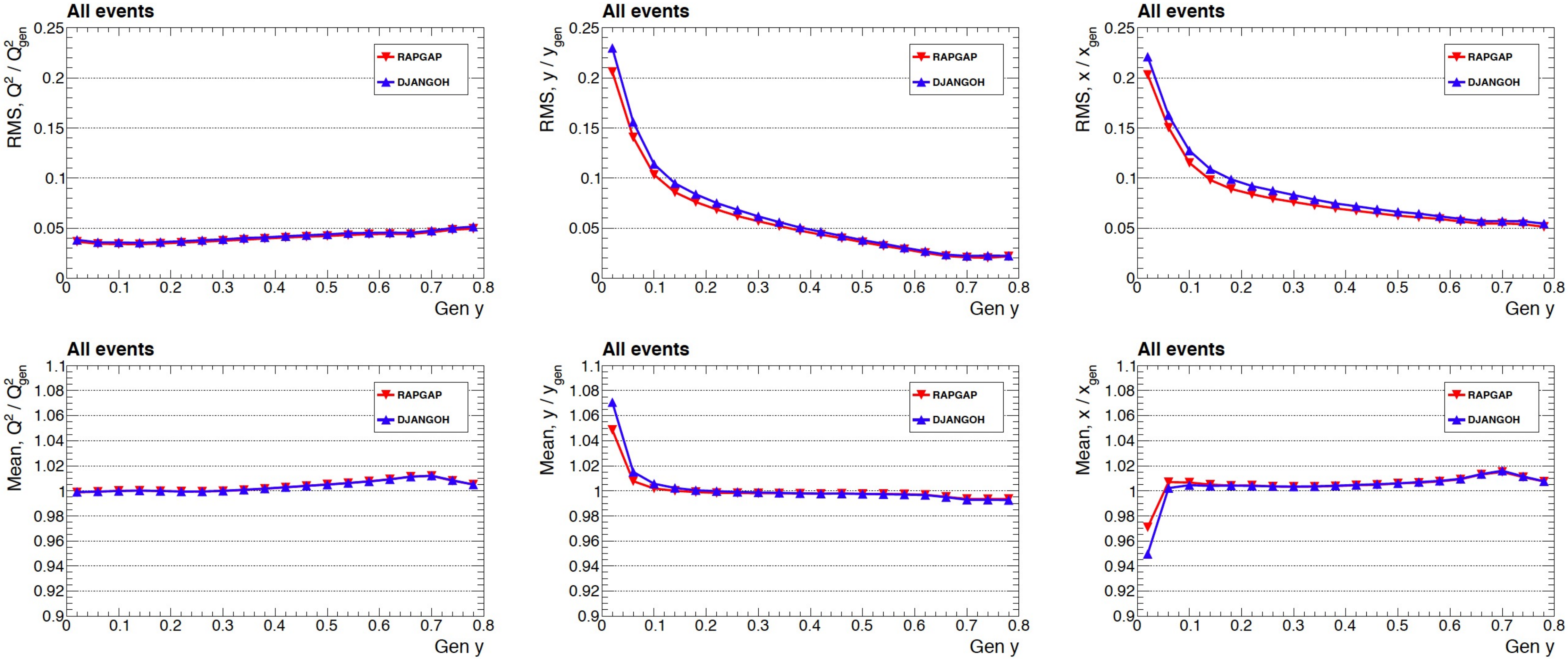
ATHENA fast simulation



ATHENA fast simulation

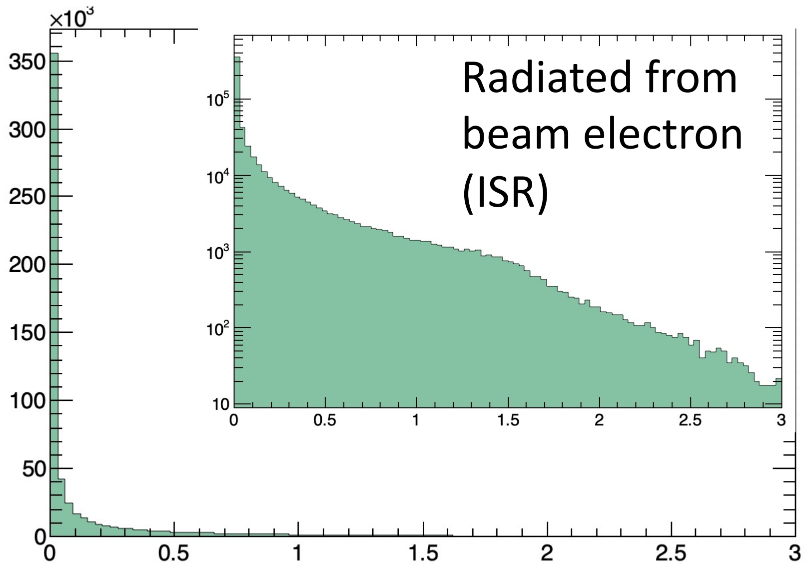


H1 full simulation

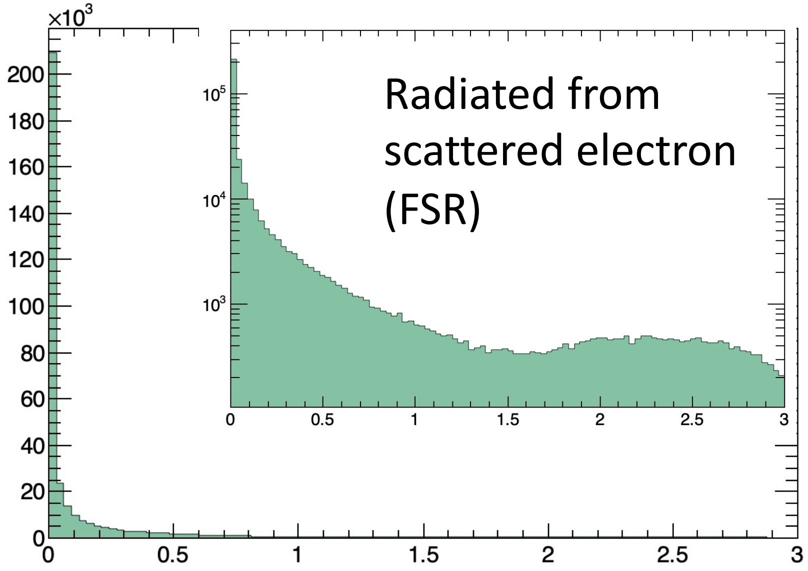


Test of using DNN trained in RAPGAP sample to make predictions in DHANGOH sample.

QED Radiation



Angle between electron and photon (radians)



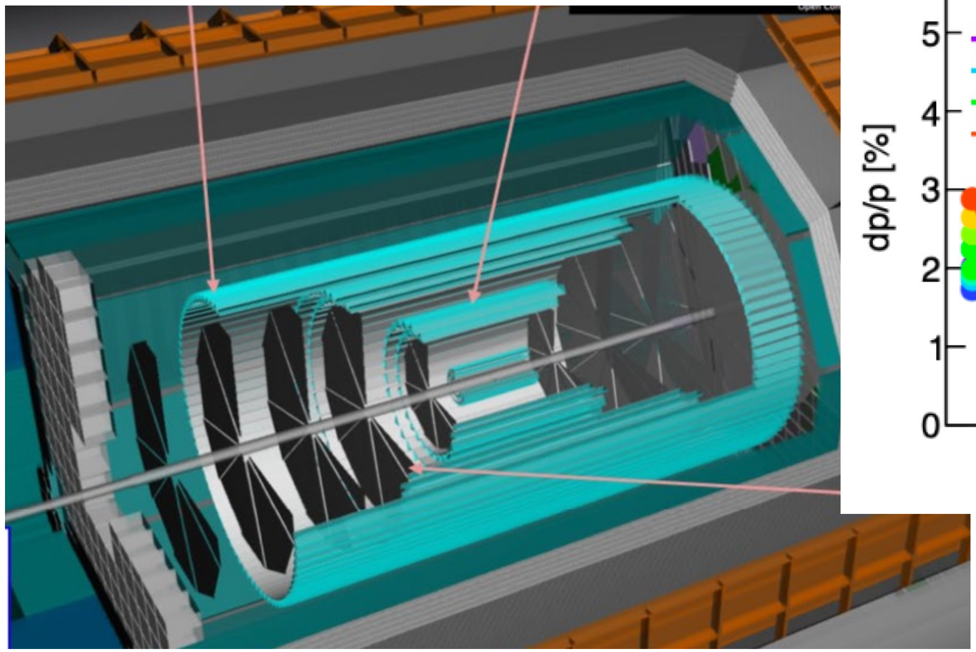
Angle between electron and photon (radians)

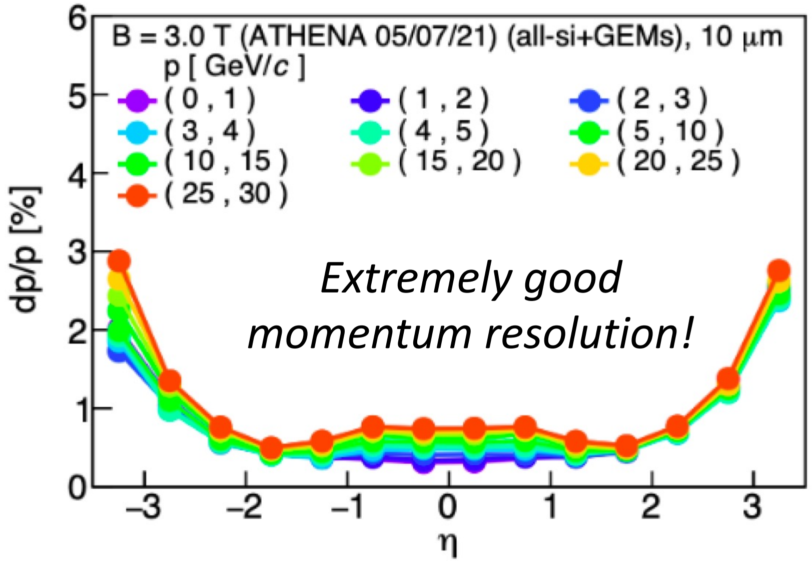
The ATHENA experiment and DELPHES fast simulation

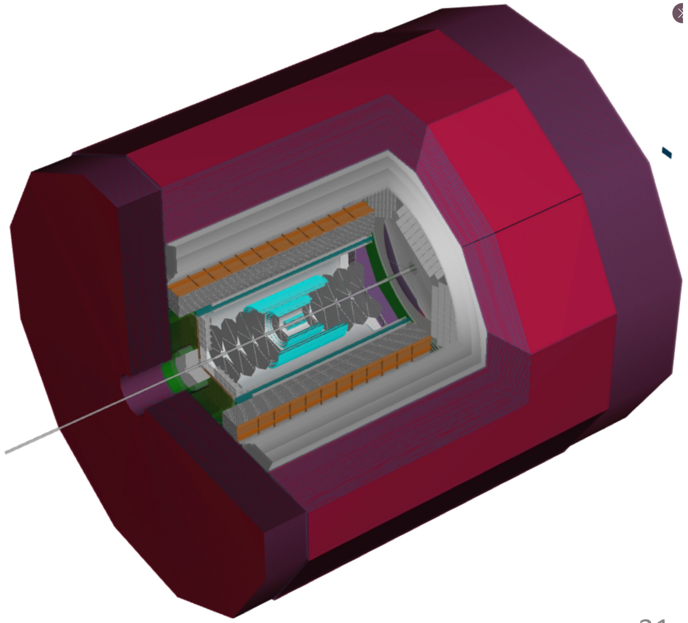
Electron Ion Collider (EIC)
beams: 275 GeV (p), 18 GeV (e)

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

DELPHES fast simulation
Detailed momentum smearing of gen particles







Event selection for this presentation

Generated $Q^2 > 200$ GeV^2

32 GeV < event (E-pz) < 40 GeV , (±4 GeV around $2E_e$) reduces QED radiation

See references in extra slides for more info.