Version of Record: https://www.sciencedirect.com/science/article/pii/S0168900221010445
Manuscript_817a8d06102c85cf095283d701578ca4

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 f–“”u″r Physik F–“”o″hringer Ring 6 80805 M–“”u″nchen 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*

Abstract

We introduce a method to reconstruct the kinematics of neutral-current deep inelastic scattering (DIS) using a deep neural network (DNN). Unlike traditional methods, it exploits the full kinematic information of both the scattered electron and the hadronic-final state, and it accounts for QED radiation by identifying events with radiated photons and event-level momentum imbalance. The method is studied with simulated events at HERA and the future Electron-Ion Collider (EIC). We show that the DNN method outperforms all the traditional methods over the full phase space, improving resolution and reducing bias. Our method has the potential to extend the kinematic reach of future experiments at the EIC, and thus their discovery potential in polarized and nuclear DIS.

1. Introduction

The process of deep-inelastic scattering (DIS) is governed by the four-momentum transfer squared of the exchanged boson Q^2, the inelasticity y, and the Bjorken scaling variable x [1, 2, 3]. These kinematic variables are related via the relation $Q^2 = sxy$, where s is the square of the center-of-mass energy.

*Corresponding author

© 2021 published by Elsevier. This manuscript is made available under the Elsevier user license
https://www.elsevier.com/open-access/userlicense/1.0/

Conservation of momentum and energy over constrain the DIS kinematics and leads to a freedom to calculate Q^2, y and x from measured quantities [4, 5, 6, 7, 8, 9, 10, 11]. Each of these methods has advantages and disadvantages and no single approach is optimal over the entire phase space. In addition, each method exhibits different sensitivity to QED radiative effects, which further complicates the choice of an optimal approach. It is a critical time to re-examine reconstruction techniques given ongoing analyses of data from HERA and the future electron-ion colliders in the USA (EIC) [12, 13] and China (EicC) [14], as well as the proposed Large Hadron electron Collider (LHeC) at CERN [15, 16].

Machine learning is a promising tool for kinematic reconstruction in DIS because of its potential to automatically synthesize many dimensions at once. Deep learning has been proposed for a variety of tasks in hadronic final-state (HFS, h) reconstruction, including particle identification [17, 18, 19, 20, 21], jet-energy reconstruction [22, 23, 24, 25, 26, 27], jet tagging [28, 29, 30, 31], unfolding [32, 33, 34, 35, 36, 37, 38, 39, 40], and more [41, 42, 43, 44, 45, 46].

We develop a method to reconstruct DIS kinematic variables and account for QED radiation that relies on deep-neural networks (DNNs). This paper is organized as follows. Section 2 briefly reviews the current methods for kinematic reconstruction in DIS. A DNN trained using fast simulations of the proposed ATHENA detector at the future EIC is studied in Section 3 and the same approach is demonstrated in full simulations of the H1 detector at HERA in Section 4. Section 5 explores the impact of additional acceptance and resolution effects on a fast simulation in comparison with a full detector simulation. The paper ends with conclusions and outlook in Section 6. Concurrently to our proposal, M. Diefenthaler et al. [47] studied the application of DNNs for the combination of the input and output variables of three reconstruction methods for Q^2 and x in NC DIS.

2. Basic kinematic reconstruction in DIS

Figure 1 illustrates the process of lepton-proton scattering[1], which is defined by the incoming electron four-vector k, the outgoing electron four-vector k', the incoming proton four-vector P, and the four-vector of the HFS defined as the sum of all four-vectors that originate from the hadron vertex.

[1] In the following, the lepton can be an e^- or e^+ but will generically just be referred to as an electron.

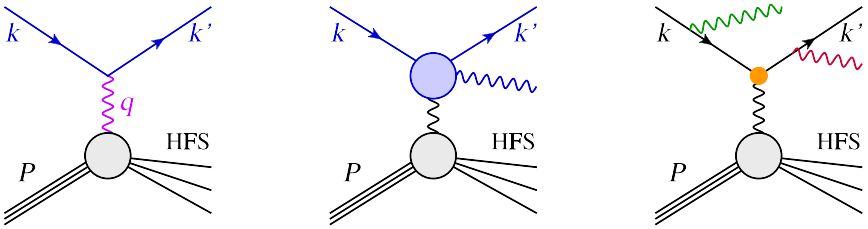

Figure 1: Illustration of *ep* scattering. Left: (QED) Born-level diagram. Middle: generic illustration of a radiative leptonic tensor including higher-order QED corrections at the lepton vertex (the box-diagrams are additive and not shown). Right: practical implementation of QED higher-order corrections in MC event generators in terms of initial state radiation (ISR, green), final state radiation (FSR, red), and an effective coupling (orange).

Using the photon four-vector $q = k - k'$, the QED Born-level kinematics are described by:

$$s = (k + P)^2\,, \quad Q^2 = -q^2\,, \quad y = \frac{q \cdot P}{k \cdot P}\,, \quad \text{and} \quad x = Q^2/(sy)\,. \tag{1}$$

Due to azimuthal symmetry and ignoring mass effects, only two variables of the outgoing four-vectors are of relevance. The usual choice is

- the scattered electron energy E and its polar angle θ, and
- the energy-momentum balance of the HFS, $\Sigma = E_h - p_{z,h}$, and the inclusive angle of the HFS γ.

The HFS quantity Σ can be calculated as the sum of all HFS particles $\Sigma = \sum_i^{\text{HFS}}(E_i - p_{z,i})$, and γ is defined using the transverse momentum of the HFS, T, through $\tan\frac{\gamma}{2} = \frac{\Sigma}{T}$. Together with the electron-beam energy E_0, five observables are known, while three of them suffice to define a *basic reconstruction method* for Q^2, y and x. Only for x one further needs the proton beam energy E_p.

Table 1 summarizes some of the most common reconstruction methods, which use derived quantities from the scattered electron and HFS including $p_{T,e}^2 = E^2\sin^2\theta$, $p_{z,e} = E\cos\theta$, $\Sigma_e = E - p_{z,e} = E(1-\cos\theta)$, $\tan\frac{\theta}{2} = \frac{\Sigma_e}{p_{T,e}}$ and $T^2 = p_{x,h}^2 + p_{y,h}^2 = (\sum_i^{\text{HFS}} p_{x,i})^2 + (\sum_i^{\text{HFS}} p_{y,i})^2 = E_h^2\sin^2\gamma = \Sigma^2\cot^2\frac{\gamma}{2} = \Sigma(2E_h - \Sigma)$.

Each of these methods has pros and cons, and yield good performance in limited kinematic ranges [7, 9, 11]. For example, the methods that mostly rely on the scattered electron yield the best resolution in events with large y, but their resolution on x quickly diverges at low y. In contrast, the

Method name	Observables	y	Q^2	$x \cdot E_p$
Electron (e)	$[E_0,E,\theta]$	$1-\frac{\Sigma_e}{2E_0}$	$\frac{E^2\sin^2\theta}{1-y}$	$\frac{E(1+\cos\theta)}{2y}$
Double angle (DA) [6, 7]	$[E_0,\theta,\gamma]$	$\frac{\tan\frac{\gamma}{2}}{\tan\frac{\gamma}{2}+\tan\frac{\theta}{2}}$	$4E_0^2\cot^2\frac{\theta}{2}(1-y)$	$\frac{Q^2}{4E_0y}$
Hadron (h, JB) [4]	$[E_0,\Sigma,\gamma]$	$\frac{\Sigma}{2E_0}$	$\frac{T^2}{1-y}$	$\frac{Q^2}{2\Sigma}$
ISigma (IΣ) [9]	$[E,\theta,\Sigma]$	$\frac{\Sigma}{\Sigma+\Sigma_e}$	$\frac{E^2\sin^2\theta}{1-y}$	$\frac{E(1+\cos\theta)}{2y}$
IDA [7]	$[E,\theta,\gamma]$	y_{DA}	$\frac{E^2\sin^2\theta}{1-y}$	$\frac{E(1+\cos\theta)}{2y}$
$E_0E\Sigma$	$[E_0,E,\Sigma]$	y_h	$4E_0E-4E_0^2(1-y)$	$\frac{Q^2}{2\Sigma}$
$E_0\theta\Sigma$	$[E_0,\theta,\Sigma]$	y_h	$4E_0^2\cot^2\frac{\theta}{2}(1-y)$	$\frac{Q^2}{2\Sigma}$
$\theta\Sigma\gamma$ [8]	$[\theta,\Sigma,\gamma]$	y_{DA}	$\frac{T^2}{1-y}$	$\frac{Q^2}{2\Sigma}$
Double energy (A4) [7]	$[E_0,E,E_h]$	$\frac{E-E_0}{(xE_p)-E_0}$	$4E_0y(xE_p)$	$E+E_h-E_0$
$E\Sigma T$	$[E,\Sigma,T]$	$\frac{\Sigma}{\Sigma+E\pm\sqrt{E^2+T^2}}$	$\frac{T^2}{1-y}$	$\frac{Q^2}{2\Sigma}$
E_0ET	$[E_0,E,T]$	$\frac{2E_0-E\mp\sqrt{E^2-T^2}}{2E_0}$	$\frac{T^2}{1-y}$	$\frac{Q^2}{4E_0y}$
Sigma (Σ) [9]	$[E_0,E,\Sigma,\theta]$	$y_{\mathrm{I}\Sigma}$	$Q^2_{\mathrm{I}\Sigma}$	$\frac{Q^2}{4E_0y}$
eSigma ($e\Sigma$) [9]	$[E_0,E,\Sigma,\theta]$	$\frac{2E_0\Sigma}{(\Sigma+\Sigma_e)^2}$	$2E_0E(1+\cos\theta)$	$\frac{E(1+\cos\theta)(\Sigma+\Sigma_e)}{2\Sigma}$

Table 1: Summary of basic reconstruction methods that employ only three out of five quantities: E_0 (electron-beam energy), E and θ (scattered electron energy and polar angle), Σ and γ (longitudinal energy-momentum balance, $\Sigma=\sum_{\mathrm{HFS}}(E_i-p_{z,i})$, and the inclusive angle of the HFS). Alternatively, the A4 method makes use of the HFS total energy E_h. Shorthand notations are used for the longitudinal energy-momentum balance of the electron, Σ_e, and for the transverse momentum of the HFS, T. The $E\Sigma T$ and E_0ET methods are under-constrained and have two solutions, referring to two possible electron polar angles, and several more are existent when using E_h (see Ref. [7] for two examples). The two bottom rows provide the equations of the Σ and $e\Sigma$-methods, which combine quantities of different basic reconstruction methods, while further methods (like the PT (rDΣ), DΣ, r$e\Sigma$ or mixed method) are found, e.g., in Ref. [9, 11].

methods that rely mostly on the HFS variables yield better performance at low y, but are rather limited at high y. Consequently, the H1 and ZEUS collaborations have used different methods in different kinematic ranges (see Refs. [48, 49] and references therein). For example, in Refs. [50, 51, 52], the electron method is used for $y \gtrsim 0.19$, while the Σ or $e\Sigma$ method are used at lower y, and DA method is employed for calibration.

In a massless, Born-level calculation, all methods yield equivalent results because of momentum and energy conservation ($2E_0 = \Sigma + \Sigma_e$, $p_{T,e} = T$ and $Q^2 = sxy$). However, once (real) higher-order QED effects are considered, the various methods yield different results and the calculated quantities for Q^2, y and x are not representative for the $\gamma/Z + p$ scattering process at the hadronic vertex.

Higher-order QED effects at the lepton vertex are generically represented as a correction to the leptonic tensor, as displayed in the middle diagram of Figure 1. Such radiative corrections include QED bremsstrahlung off the lepton, photonic lepton-vertex corrections, self-energy contributions at the external lepton lines, and fermionic contributions to the running of the fine-structure constant, and additional box-diagrams representing multi-boson exchange. The complete first-order corrections are calculable semi-analytically [53, 54, 55]. For an implementation in Monte-Carlo (MC) event generators, these calculations are split by partial-fraction decomposition into initial-state and final-state photon radiation (ISR and FSR) and using effective couplings, as displayed in the right diagram of Figure 1.

Two techniques are commonly applied to reduce sensitivity to QED radiation. The first technique is to merge the FSR photons with the electron, thus providing the four-vector linked to the exchanged boson, and the second technique is to take the ISR radiation to be collinear, which implies that $(\Sigma + \Sigma_e)/2$ provides the incoming electron beam energy that contributes in the interaction. For soft and collinear FSR, the first is done implicitly also at detector level, e.g. when the photon is measured in the same calorimeter cell as the electron. ISR photons often escape undetected through the beam hole of the detector.

In cross-section measurements, the radiative particle-level is commonly just an intermediate step, and additionally requires the application of well-defined QED correction factors. For precision measurements, these factors should be small. Common definitions for cross sections in DIS are:

- Structure-function measurements are made as a function of Q^2 and x and are quoted at the 'Born-level' and thus have to correct for all higher-order QED effects, such that the fine-structure constant factorizes from the calculation of the structure functions.

- At HERA, measurements of the HFS were quoted as non-radiative $\gamma^* p$ cross sections, which are corrected for first-order QED and electroweak effects.

A *radiative cross section* can be defined by merging any radiated photon with the scattered electron that is closer to the scattered electron than to the electron beam. By specifying a single reconstruction method, a well-defined, meaningful and almost model-independent definition of Q^2, y and x is obtained.

In the following, we will deal carefully with radiated photons at the particle level and the detector level to determine kinematic quantities Q^2, y and x, in an optimal way. These observables can then be used for subsequent cross-section measurement with small and well-defined QED corrections.

3. Method

We use TensorFlow [56] to construct and train a DNN to estimate the DIS kinematics using both the scattered electron and the HFS. To demonstrate our methodology, we use a fast simulation of the proposed EIC experiment ATHENA using the Delphes package [57, 58]. After presenting results for ATHENA, we will show results applying the same methodology to a full simulation of the H1 experiment [59, 60, 61]. Both studies use the Rapgap MC generator [62], which employs routines from Refs. [63, 7, 64, 65]. For H1, Rapgap version 3.1 is used, while verion 3.3 is used for the ATHENA studies. In addition, the MC generator Djangoh 1.4 [66] is used to test prior dependence. Both MC generators employ the HERACLES routines [53, 54, 55] for the simulation of higher-order QED effects.

We restrict our study to events with $Q^2 > 200$ GeV^2. This kinematic region is well measured, since the electron is scattered into the central regions of the detector. However, no single reconstruction method gives optimal performance over the full phase space [9].

3.1. Fast simulation of the ATHENA experiment

Neutral-current DIS events are generated with the Rapgap 3.3 event generator for electron-proton scattering with beam energies of $E_0 = 18\,\text{GeV}$ and $E_p = 275\,\text{GeV}$ and processed with the Delphes fast simulation of the ATHENA detector at the EIC. The scattered electron is selected as the highest-p_T track that satisfies the following criteria: correct charge, $p_T > 5$ GeV, electromagnetic fraction in the calorimeter > 0.80, and isolation < 0.20, where isolation is defined as the scalar p_T sum of all other tracks

and neutral hadrons within a cone of $\Delta R < 0.5$ around the electron direction divided by the electron p_T. The HFS is reconstructed from the sum of all Energy-Flow candidates (tracks, photons, and neutral hadrons), excluding the scattered electron and any photon candidates that is within a cone of $\Delta R < 0.4$ around the scattered electron. We require $\Sigma + \Sigma_e$ to be within $\pm 4\,\text{GeV}$ of $2E_0$ to suppress ISR events.

3.2. QED radiation and categorization of events

We introduce a practical categorization of events, which is closely related to our proposed radiative cross-section definition above:

- if the radiated photon is closer to the electron-beam direction, it is an Initial-State Radiation event (ISR);
- if the radiated photon is closer to the scattered-electron direction, it is a Final-State Radiation event (FSR);
- if no photon radiation is emitted by the generator, it is a non-radiative event (NoR).

Within RAPGAP, which implements first-order QED corrections, the radiated photon either branches off from the beam electron before it interacts with the proton, or it branches off of the scattered electron after interacting with the proton. This leads to the natural interpretation of the former as ISR and the latter as FSR, which agrees with our practical definitions in 94% of QED radiation events.

To define the target values of Q^2, y and x for events with QED radiation, we use the generated beam electron after radiation for ISR events and the generated scattered electron prior to radiation for FSR events. While this definition can be considered as the *true* kinematic quantities[2], alternative definitions in terms of particle-level observables are also possible. For example, by applying FSR merging and using the equations of ISR insensitive reconstructions methods for calculating Q^2, y and x. In our studies, these different definitions yield indistinguishable results. The generator-level quantities are generically denoted with a subscript 'gen' in the following.

3.3. DNN inputs

For our main goal of determining Q^2, y and x, the task of the DNN is in fact two-fold. In the absence of QED radiation, the task of the DNN would be to learn to compute Q^2, y and x from the input quantities subject to

[2]Note that learning the true value from detector-level quantities introduces a prior dependence [23].

finite-resolution and acceptance effects. In events with QED radiation, the DNN needs to learn to quantify the extent of QED radiation and account for it while calculating Q^2, y and x. Hence, the regression DNN must learn to treat ISR, FSR, and NoR events separately in order to achieve optimal performance. This is a key feature of our DNN method, which is absent in traditional methods.

We define the following variables to characterize the strength of QED radiation in the event:

$$p_T^{\text{bal}} \;=\; 1 - \frac{p_{T,e}}{T} \;=\; 1 - \frac{\Sigma_e \tan\frac{\gamma}{2}}{\Sigma \tan\frac{\theta}{2}} \qquad \text{and} \qquad p_z^{\text{bal}} \;=\; 1 - \frac{\Sigma_e + \Sigma}{2\ E_0}\,. \tag{2}$$

When calculating them at particle level, both quantities p_T^{bal} and p_z^{bal}, are zero for events with no QED radiation, but positive for events with FSR and ISR, respectively. The p_T^{bal} (p_z^{bal}) value indicates the strength of FSR (ISR).

The following observables that help indicate QED radiation in the event are included as inputs to the regression DNN for Q^2, y and x:

- The values of p_T^{bal} and p_z^{bal}.
- The energy, η, and $\Delta\phi$ of the reconstructed photon in the event that is closest to the electron-beam direction, where $\Delta\phi$ is with respect to the scattered electron.
- The sum ECAL energy within a cone of $\Delta R < 0.4$ around the scattered electron divided by the scattered-electron track momentum.
- The number of ECAL clusters within a cone of $\Delta R < 0.4$ around the scattered electron.

These seven observables are combined with the following eight:

- Scattered-electron quantities $p_{T,e}$, $p_{z,e}$ and E.
- HFS four-vector quantities T, $p_{z,h}$ and E_h.
- $\Delta\phi(e,h)$ between the scattered electron and the HFS momentum vector.
- The difference $\Sigma_e - \Sigma$.

The transverse momenta of the scattered electron and the HFS are highly correlated. We replace the pair of p_T values with the difference and the sum of the p_T values, which removes the correlation and aids the DNN training. The sum $\Sigma_e + \Sigma$ appears in the definition of $p_z^{\rm bal}$ and is sharply peaked at twice the electron-beam energy, making the Σ values anti-correlated; hence, we include the orthogonal combination $\Sigma_e - \Sigma$ in addition.

3.4. DNNs for the classification or quantification of QED radiation

To determine the ability of the DNN to identify and quantify QED radiation, we investigated two approaches: a classification network and a regression network for $p_T^{\rm bal}$ and $p_z^{\rm bal}$. Both DNNs differ only in the activation function for the final layer, the learning targets, and the loss function for the training.

We followed a heuristic approach in designing the DNN, guided by prior experience. We chose to have the number of nodes per hidden layer grow in the first layers, reach a peak size, and then fall at a rate that is symmetric about the peak layer. Several trial configurations were tested, each with a different number of hidden layers and/or number of nodes per layer, though we did not perform a thorough scan. We also tried a few sets of optimizer hyperparameters before finding values that gave good results in a reasonable amount of training time. Various standard activation functions were tested with no appreciable differences in the results. This basic DNN design is effective for a variety of tasks, including classification and regression, as we explain below. The same design works well for both the ATHENA fast simulation and the H1 full simulation. Our design explorations were terminated once we achieved DIS reconstruction performance that exceeded the conventional methods. We have not carried out thorough hyperparameter scans, so further improvements may be possible.

The QED-classification DNN consists of a sequential network with 8 layers. The 15 DNN inputs, defined in the previous section, are transformed prior to training to have zero mean and unit RMS using the `sklearn StandardScaler` [67]. The learning targets are three binary (0 or 1) state variables to tag the events as ISR, FSR, or NoR. The activation function is a rectified linear unit (`relu`) for the first layer, scaled exponential linear unit (`selu`) [68] for the middle layers, and the `softmax` function for the final output layer. The numbers of nodes per layer are 64, 128, 512, 1024, 512, 128, and 64, with 3 outputs in the final layer. The outputs are three numbers, each between 0 and 1, that sum up to 1. The loss function for the training is categorical cross entropy. The training is performed with the `Adam` optimizer [69] with a learning rate of 10^{-4}. The total number of

parameters of the DNN model is 1,199,555. The training and validation are performed using over 28 million simulated events with half for training and half for validation. The batch size for the training is 128. The training terminates after 38 epochs, finding no further improvement in the validation loss function.

Figure 2 shows distributions for the three outputs of the QED-classification DNN. Some events have clear evidence of QED radiation and are strongly identified. Some degree of mis-classification is to be expected, since neither soft ISR/FSR, nor collinear FSR induce a measurable signal in the detector.

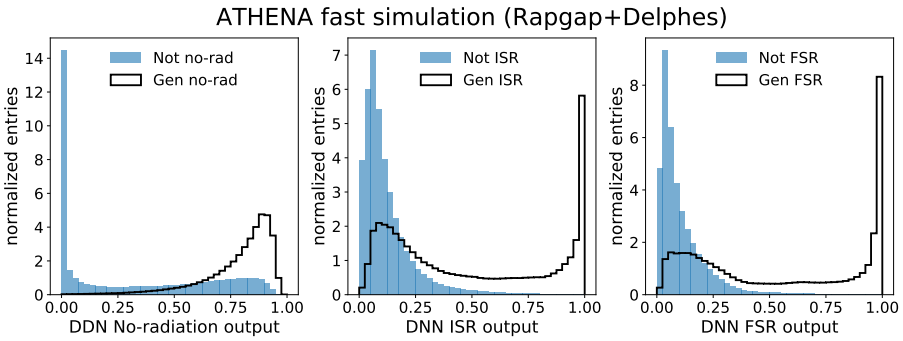

Figure 2: Distributions of the QED-classification DNN predictions for the ATHENA fast simulation. The distributions are normalized to equal area.

In the second study, the QED regression for $p_T^{\rm bal}$ and $p_z^{\rm bal}$, the DNN has the same structure as the QED-classification DNN except for the following differences. The activation function is the `linear` function for the final output layer, which has 2 nodes instead of 3. The learning targets are the particle-level values of $p_T^{\rm bal}$ and $p_z^{\rm bal}$, which are transformed to have zero mean and unit RMS prior to training. The loss function for the training is `Huber` loss [70] with a transition between quadratic and linear loss at ± 0.01. The batch size for the training is 1024. The training and validation are done using the same sample as for the QED-classification DNN.

Figure 3 shows distributions of the DNN predictions for $p_T^{\rm bal}$ and $p_z^{\rm bal}$, separately for ISR, FSR, and NoR events, as well as scatter plots of the predicted values vs the particle-level values. The predicted $p_T^{\rm bal}$ distribution for FSR events is shifted to positive values for FSR events, while the predicted $p_z^{\rm bal}$ distribution is shifted to positive values for ISR events, as expected. For many events, the DNN is able to accurately estimate both $p_T^{\rm bal}$ and $p_z^{\rm bal}$. There are also QED-radiation events where the prediction is zero, which

correspond to cases where the radiated photon is either out of acceptance or not clearly identified.

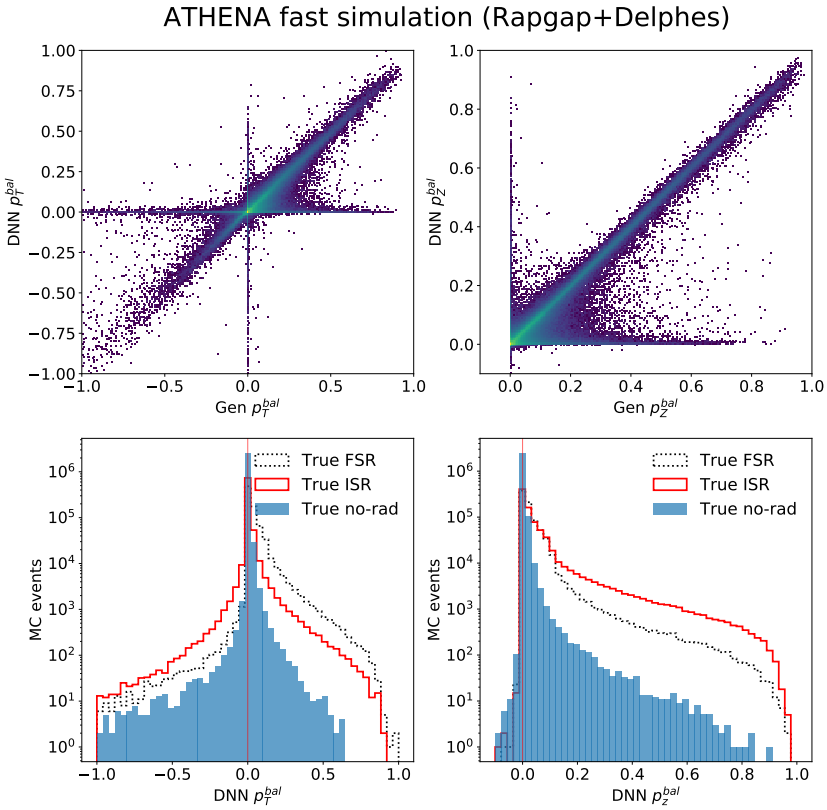

Figure 3: Distributions of the QED-regression DNN predictions and particle-level values of p_T^{bal} and p_z^{bal} for the ATHENA fast simulation.

3.5. Regression DNN for DIS kinematic variables $\boldsymbol{Q^2}$, $\boldsymbol{y}$ *and* $\boldsymbol{x}$

We estimate Q^2, y and x using a regression DNN that has a similar structure as the QED regression DNN described above, except for the final output layer that has three nodes corresponding to the target variables Q^2, y and x. The learning rate is 10^{-5} and the batch size is 1024. Since the distributions of Q^2, y and x are approximately exponential, the DNN is trained to predict the logarithm of each variable. The training and validation are performed using over 28 million events, using half for validation and half for training.

As a pilot study, we consider three different choices for the input variables to the DNN:

- add the three QED-classification DNN outputs (FSR, ISR, NoR) as inputs, in addition to the 15 variables.
- add the QED-regression DNN predictions $p_T^{\rm bal}$ and $p_z^{\rm bal}$ as inputs, in addition to the 15 variables.
- use the same 15 inputs as in the QED-classification and regression DNNs.

We found that the results of these three approaches are essentially the same. This suggests that the QED-classification and the QED-regression DNNs do not provide any additional information beyond what the regression DNN for Q^2, y and x learned. In light of this, we chose the simplest (third) option, which uses only the original 15 variables as inputs. The total number of parameters for the DNN model is 1,199,619. The training terminates after 103 epochs, finding no further improvement in the validation loss function.

3.6. Benchmark of the DNN reconstruction vs. standard methods

Figure 4 shows the results for the DNN and traditional methods using events with $Q^2 > 200$ GeV^2 and two y regions, obtained with the ATHENA fast simulation. The resolutions from the DNN exhibit a peak at unity and mostly Gaussian-like tails; in contrast, classical methods yield a peak but larger tails caused by their limited use of the reconstructed quantities and the presence of ISR or FSR.

The properties of the resolutions, their mean and RMS, are displayed in Figure 5 for small intervals of y [3]. The DNN reconstruction has the smallest RMS among all methods, for all three kinematic variables and all y intervals. Also, the mean distributions are unbiased for Q^2, y and x for all y intervals, while the classical methods exhibit large biases.

We examine more closely the resolution and bias for events with and without QED radiation in Figures 6 and 7, where we use the definition of NoR, ISR and FSR events from Section 3.2. The RMS for events with no QED radiation gives a measure of the core resolution, free from the tails that are visible in the distributions of Figure 4 for the conventional methods. All methods, except the hadron method, show no bias in NoR events. For x and y, the electron method has a better core resolution than the DNN for $y > 0.15$; however, it suffers from poorer resolution and a

[3] More detailed representations of the resolutions for the different methods are shown in Appendix Appendix A.

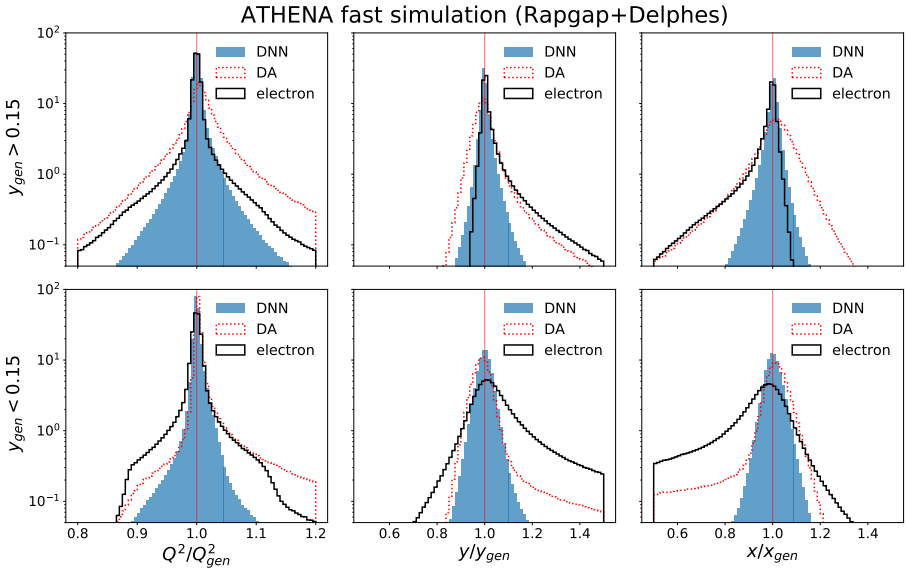

Figure 4: Resolution for Q^2 (left), y (middle), and x (right) for the DNN, electron, and double-angle (DA) methods for the fast simulation of the ATHENA experiment. The top (bottom) row is for events with $y_{\text{gen}} > 0.15$ ($y_{\text{gen}} < 0.15$). All distributions are normalized to the same area.

strong bias in events with QED radiation. The DNN reconstruction has some loss in performance in QED-radiation events, but it is about a factor of two better than the electron method in these events, and shows no bias. We conclude that the DNN has successfully learned to mitigate the effects of QED radiation that spoil the resolution and bias the calculations of the conventional methods.

4. Demonstration using the full simulation of the H1 experiment

We apply our DNN methodology to simulated events of the H1 experiment at HERA. The events were simulated by the H1 Collaboration using the RAPGAP 3.1 [62] and DJANGOH 1.4 [66] generators for the beam energies $E_0 = 27.6\,\text{GeV}$ and $E_p = 920\,\text{GeV}$. The generators employ the HERACLES routines [53, 54, 55] for QED radiation, the CTEQ6L PDF set [71], and the Lund hadronization model [72] with parameters fitted by the ALEPH Collaboration [73]. The simulation of the H1 experiment [60, 61] employs the GEANT 3 package [59] and includes real calorimeter noise and fast shower simulations [74, 75, 76, 77, 78, 79]. The simulation includes time-dependent

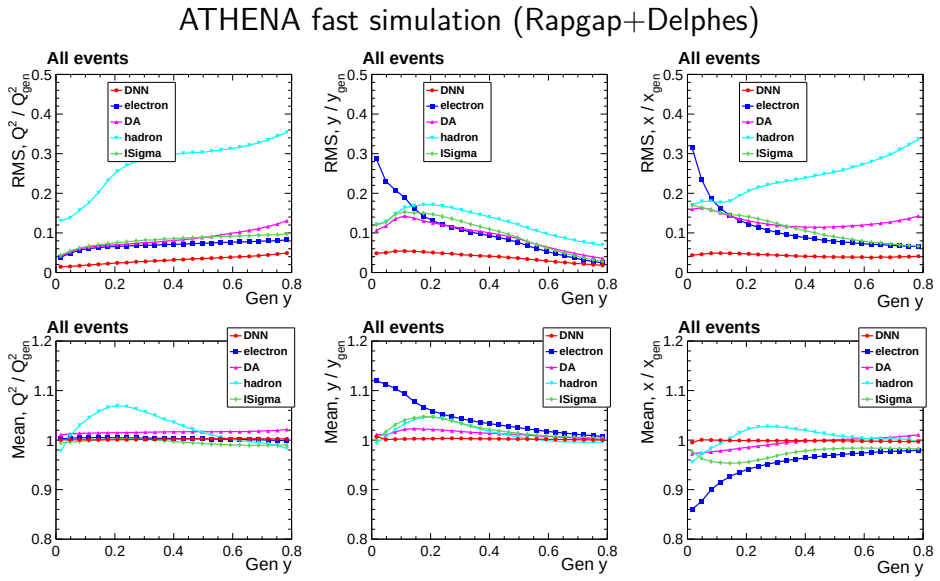

Figure 5: Resolution on the reconstruction of Q^2 (left), y (middle), and x (right) as a function of the generated y for the fast simulation of ATHENA. The top (bottom) row shows the RMS (mean) of the measured-over-generated distribution as a function of generated y_{gen}. The RMS and mean are calculated using events with the measured-over-generated ratio within the interval 0 to 2.

properties ('run-specific'), where the detector state and beam properties correspond altogether to the HERA-II data taking periods.

The simulated events are reconstructed just like data, in particular, an energy-flow algorithm [80, 81, 82] is used to define objects whose sum yields the HFS four-vector, and the scattered electron candidate are defined using the same approach as Refs. [51, 83, 40]. The simulated events also undergo the same (in-situ) calibration procedure as real data, using the latest calibration by the H1 Collaboration [51, 84, 83]. Some technical selections and fiducial cuts are applied as it would be done similarly to real data. In particular, events are required to have $45 < \Sigma + \Sigma_e < 65\,\text{GeV}$ to suppress ISR events; a veto on QED Compton events is imposed; and since a trigger simulation is included, our study is limited to $E \gtrsim 11\,\text{GeV}$ [83]. The simulated events are processed within H1's computing environment [85] and altogether several 10^8 events were simulated, and after 'run'-selection, acceptance effects and our selection of $Q^2 > 200\,\text{GeV}^2$ yield about $7 \cdot 10^7$ simulated and reconstructed events for our DNN studies.

We use the same regression DNN structure, input variables, and training

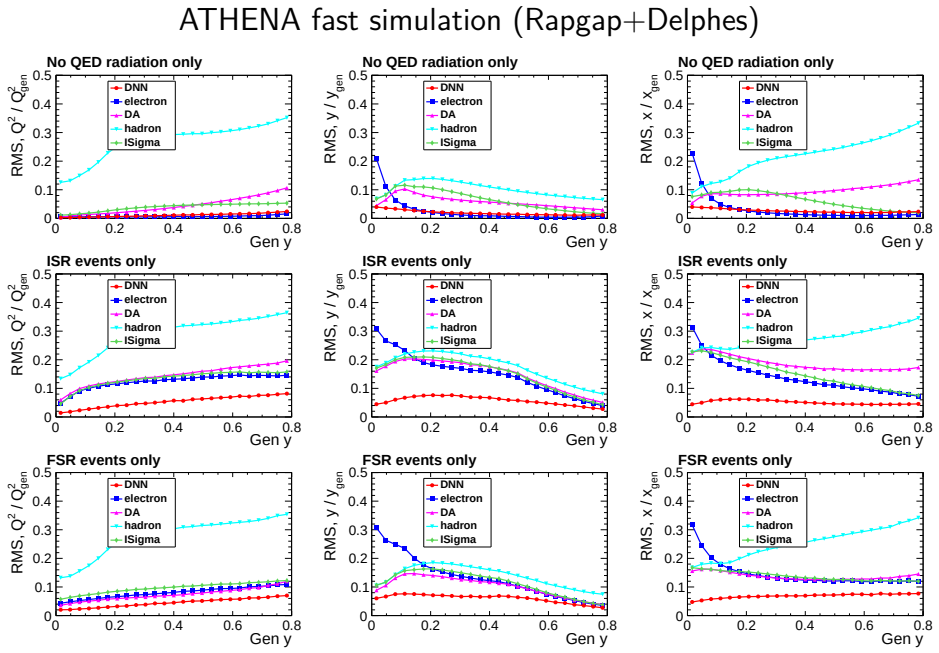

Figure 6: Comparison of the RMS of the resolution of Q^2, y and x (from left to right) for NoR (top), ISR (middle), and FSR (bottom) events. The RMS is calculated using events with the measured-over-generated ratio within the interval 0 to 2.

methods for the simulated H1 events as previously for the ATHENA events. We take a sample of over 12 million Rapgap events and use half of the events for training and half for validation. The training terminates after 125 epochs, finding no further improvement in the validation loss function.

Figure 8 shows the resolutions for the DNN and two classical methods in two intervals of y. Similarly to the results obtained with the ATHENA simulation, the DNN yields a peak at unity and no asymmetric tails for all three quantities, in contrast with classical methods. The mean and RMS of the resolutions are displayed as a function of y in Figure 9. The DNN yields the smallest RMS. The mean of the DNN reconstruction is closest to unity over a wide range of y. Only for Q^2, or at highest y, the electron method achieves comparably good RMS and mean. In contrast, at lower y the DNN provides a significant improvement and results in a bias-free reconstruction of y and x with superior resolution[4].

[4]More detailed representations of the resolutions for the different methods are collected in Appendix Appendix B, where also further reconstruction methods from Section 2 are

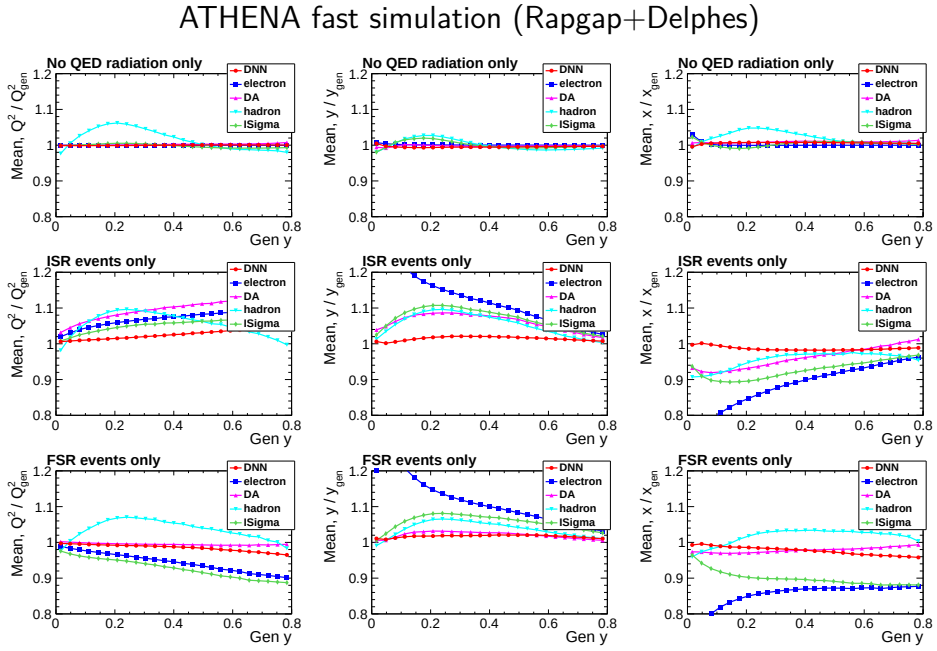

Figure 7: Comparison of the mean of the resolution distributions of Q^2, y and x (from left to right) for NoR (top), ISR (middle), and FSR (bottom) events. The mean is calculated using events with the measured-over-generated ratio within the interval 0 to 2.

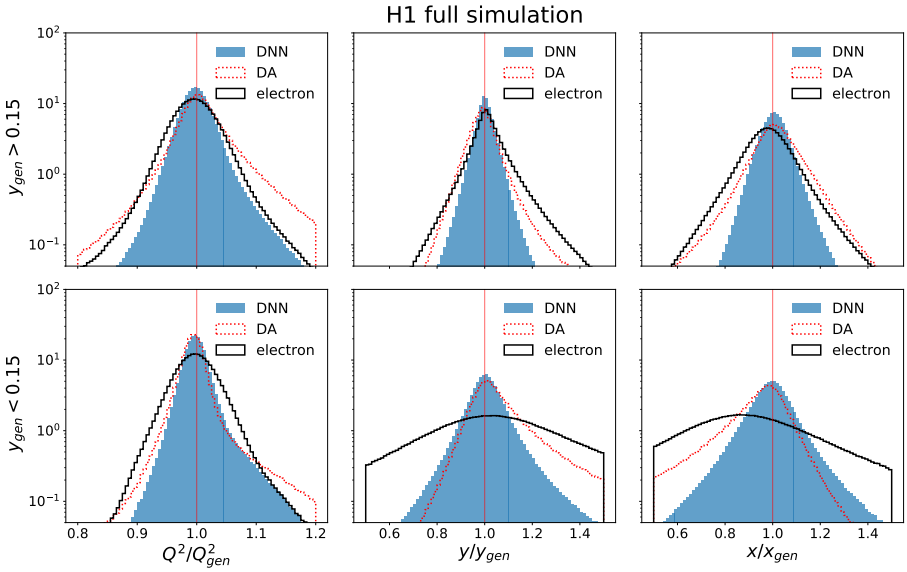

Figure 8: Resolution for Q^2 (left), y (middle), and x (right) for the DNN, electron, and double-angle (DA) reconstruction methods for the full simulation of the H1 experiment. The top (bottom) row is for events with $y_{\text{gen}} > 0.15$ ($y_{\text{gen}} < 0.15$). All distributions are normalized to the same area.

To assess a possible bias in our DNN methodology that may arise from the details of the MC event generator that is used to train the DNN, we study the performance of the DNN reconstruction using two different MC event generators, RAPGAP and DJANGOH. The two event generators differ in the modelling of higher-order QCD radiation that results in significant differences in the prediction of the HFS. DJANGOH employs the color-dipole model, while Rapgap employs a matrix-element plus parton shower model, where the parton shower is in the leading-logarithmic approximation. In this study, we use H1's full simulation and train the DNN with a RAPGAP event sample. Subsequently this DNN is applied to a statistically independent RAPGAP event sample and to a simulated DJANGOH sample.

Figure 10 shows the Q^2, y and x resolutions as a function of the generated y for both the RAPGAP and DJANGOH samples when using the same DNN, where the measured-over-generated ratio is calculated on an event-by-event basis with the respective generated values. The results from the DJANGOH event sample are nearly indistinguishable from the RAPGAP sample. In

studied with a real detector simulation.

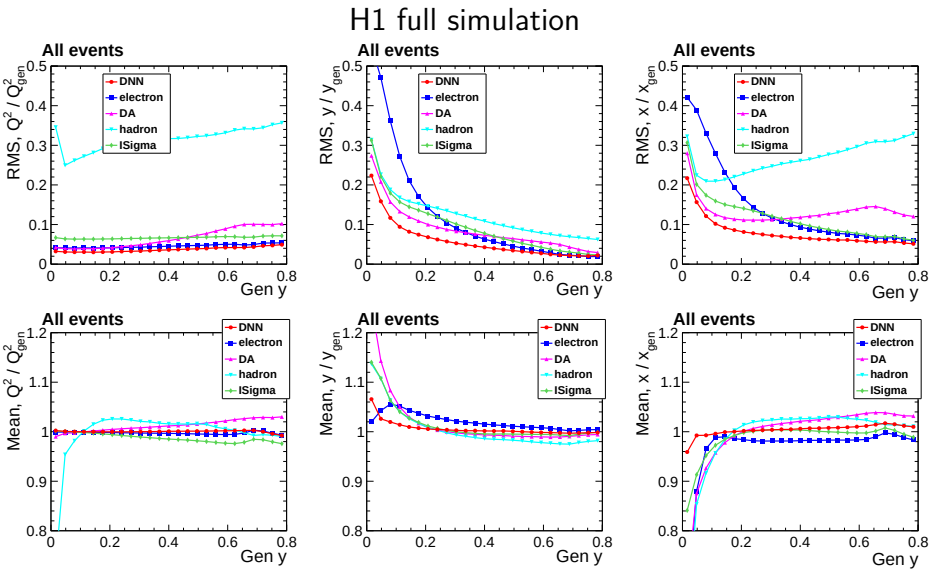

Figure 9: Resolution on the reconstruction of Q^2 (left), y (middle), and x (right) as a function of the generated y for the full simulation of H1. The top (bottom) row shows the RMS (mean) of the measured-over-generated distribution as a function of the generated y. The RMS and mean are calculated using events with the measured-over-generated ratio within the interval 0 to 2.

particular, the mean of the distribution is unbiased. This result suggests that any generator-specific systematic errors in the DNN predictions is negligible.

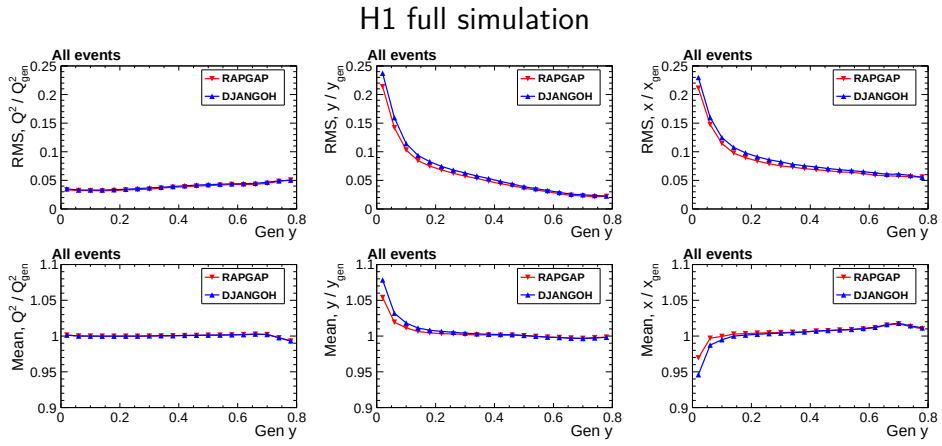

Figure 10: Resolution comparison of the RAPGAP and DJANGOH generators on the DNN reconstruction of Q^2 (left), y (middle), and x (right) as a function of the generated y for the full simulation of H1. The top (bottom) row shows the RMS (mean) of the measured-over-generated distribution as a function of the generated y. The red (blue) curves show the results for the RAPGAP (DJANGOH) event sample, while the DNN was trained with the RAPGAP sample in both cases.

The resolution (RMS) for y and x increase at lower y, even for the DNN reconstruction. Since this pattern is not present in the the ATHENA fast simulation results and may be attributed to further acceptance, noise, or resolution effects that deteriorates the measurement of the HFS [86]. A dedicated study using a DELPHES fast simulation is presented in the following section.

5. Impact of further acceptance and resolution effects at low y

At low y, the HFS-based methods perform better than the electron method. The reason is, that the ratio $E(1-\cos\theta)/2E_0$ gets close to one and cannot be measured accurately because of large values of E. Likewise, however, the HFS momentum balance Σ goes to zero as y goes to zero. Although for kinematic reconstruction at low y the usage of Σ is preferred over Σ_e, the quantity Σ is particularly sensitive to resolution and acceptance effects. In particular HFS components that are more in the central region of the calorimeter contribute more, such making Σ at low y especially sensitive resolution effects or efficiency losses in the central part of the detector or fake components from calorimetric noise.

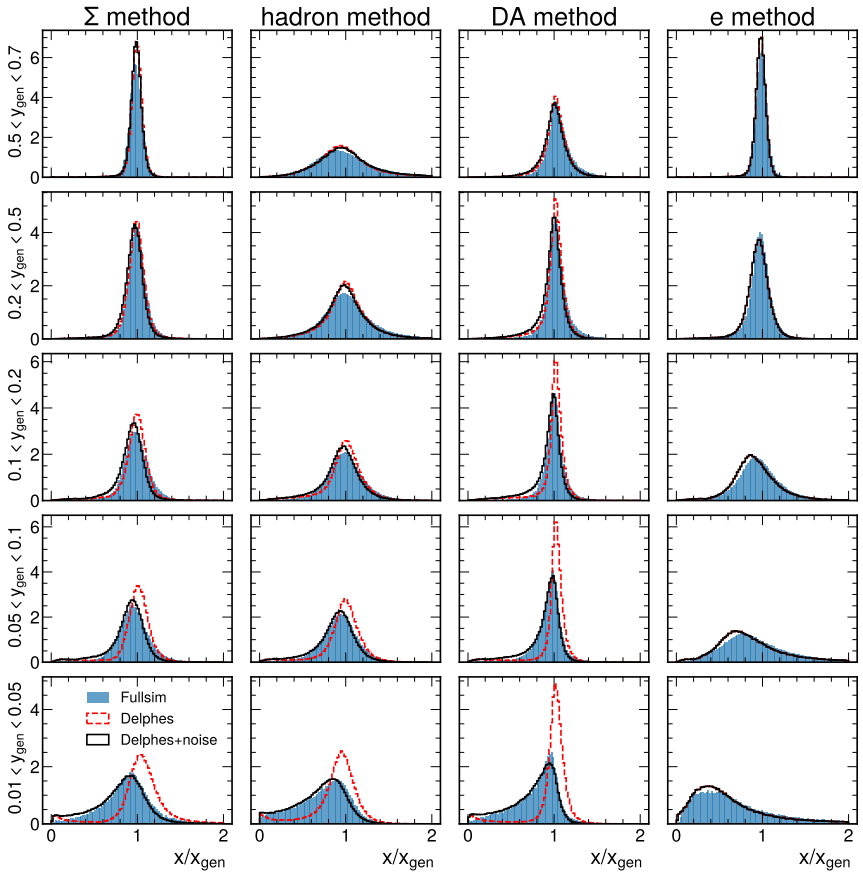

Figure 11: Resolutions for x in various y ranges and reconstruction methods obtained with a full GEANT simulation of H1 (blue histogram) and a fast simulation based on DELPHES with (black) and without (red) an added ad-hoc noise contribution.

The DELPHES fast simulation does not include calorimeter noise hits, nor does it account in a full-fledged manner for single-particle acceptance effects and efficiency losses as they can be present at the boundaries of calorimeter stacks or because of insensitive material. To test the hypothesis that such detector effects can be responsible for the resolution decrease in x and y for hadronic reconstruction methods at low y ($y \lesssim 0.15$), we have implemented the H1 experiment in DELPHES. Figure 11 shows the x resolution for the standard reconstruction methods for our fast simulation of H1 compared to the full simulation. The agreement between the fast and full simulation at high y is fairly good. At low y, however, there is a low-side tail for the Σ, hadron and DA method for events processed with the full H1 simulation, while that tail is absent in the fast simulation, and also the mean value is shifted.

We apply an additional additive component with random sign to the HFS to the fast simulated events, which mimics further detector effects, like acceptance or efficiency losses, reduced resolution or artificial components from electronic noise in the calorimeter. The model we use to simulate these effects is to generate random numbers from `TRandom::Landau` in ROOT [87] with `mu=0` and `sigma=0.05` in units of GeV, and add it with a random sign to the p_x, p_y, and p_z components of the reconstructed HFS four-vector.

The results are also displayed in Figure 11. We find that adding such an additional detector effect to the HFS in the fast simulation brings the fast simulation into good agreement with the full detector simulation. Adding these ad-hoc detector effects produces a low-side tail in the x resolution at low y but does not affect the x reconstruction at high y. The electron remains naturally unaltered in that procedure.

This study suggests that further detector effects that are otherwise not included in the DELPHES fast simulation impact the precise measurement of the HFS and reduce the x and y resolution at low y. We do not currently have an estimate for how large that effect in ATHENA will be, what is the actual correspondence in the full simulation (an acceptance, efficiency, resolution or noise effect), or to which extent it is impacted by calibration or noise-suppression algorithms. Though, due to a larger coverage of the calorimeter, less dead material and newer detector technologies of ATHENA as compared to H1, we expect this additive component to be significantly smaller for ATHENA than what our ad-hoc model adds to H1 to bring the fast and full simulation into agreement.

To place an upper bound on the impact of these effects in the ATHENA results, we investigated adding our ad-hoc component to the ATHENA fast simulation. The full analysis is then repeated, including the DNN training.

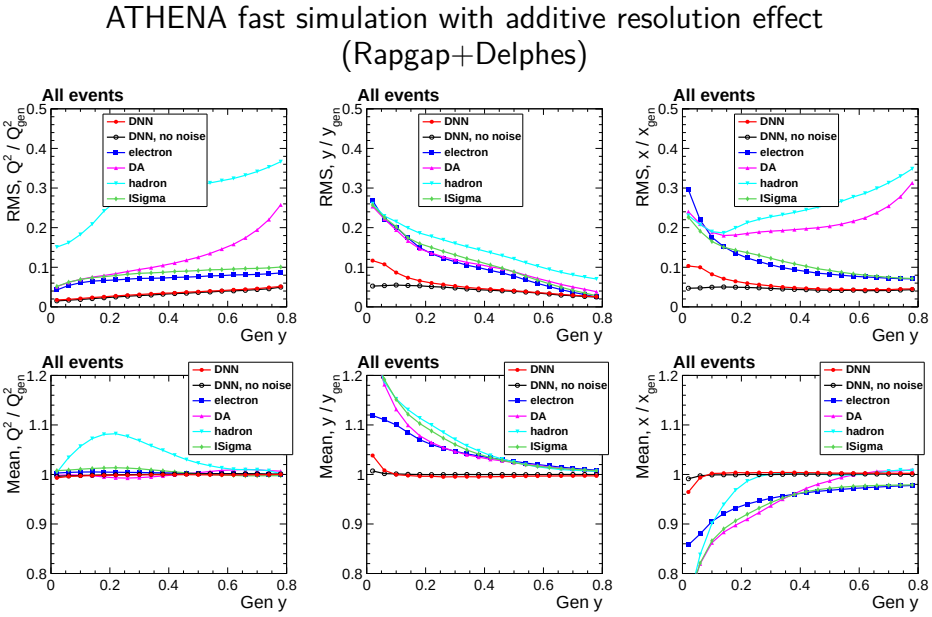

Figure 12: Resolution on the reconstruction of Q^2 (left), y (middle), and x (right) as a function of the generated y for the fast simulation of ATHENA with the ad-hoc additive resolution effects model added. The top (bottom) row shows the RMS (mean) of the measured-over-generated distribution as a function of the generated y. For comparison, the DNN curves for that unaltered sample are shown in black for comparison.

The results are shown in Figure 12 for both the conventional reconstruction methods and the DNN reconstruction, where the DNN reconstruction for the unaltered ATHENA sample is included for comparison. The DNN resolution does get worse below y of around 0.2 and there is a small bias at very low y, as we expected based on the H1 study. The Q^2 reconstruction is insensitive to that, since it is dominated by the electron reconstruction. The results in Figure 12 show that also with a very conservative ad-hoc model, the DNN outperforms all standard reconstruction methods.

6. Summary and outlook

We have presented a novel method to reconstruct the DIS kinematic variables (Q^2, y and x) using a deep neural network (DNN) that takes as input the electron and hadronic-final-state measurements as well as observables that can indicate the presence of QED radiation.

We have introduced our methodology using a DELPHES-based fast simulation of the ATHENA experiment for the EIC and validated our methods

using the well-understood full simulation of the H1 experiment at HERA.

Our method outperforms traditional methods over a wide kinematic range, improving the resolution, and decreasing the bias. We validated that our method is independent of the Monte Carlo event generator used to train the DNN. We further performed a study to validate the fast simulation approach by comparing a DELPHES model of the H1 detector with the H1 full simulation; these comparisons allowed us to identify key effects for low y events.

Our DNN-based method shows promise for improving the resolution and extending the kinematic reach of flagship EIC measurements. Given that EIC experiments are being designed, our method offers a way to benchmark detector performance against DIS using an optimal combination of electron and HFS measurements.

Code availability

The code in this work can be found in: `https://github.com/owen234/DIS-reco-paper/`

Acknowledgments

We thank our colleagues from the H1 Collaboration for allowing us to use the simulated MC event samples and for providing valuable feedback to the analysis an on the manusript, in particular Günter Grindhammer, Sergey Levonian, Stefan Schmitt and Zhiqing Zhang. We also thank members of the ATHENA Collaboration, and in particular Steven Sekula, for help with the ATHENA DELPHES model. We thank Hubert Spiesberger for valuable discussions about QED radiative effects and insights into the HERACLES routines. We thank Hannes Jung for providing the RAPGAP event generator. Thanks to DESY-IT and MPI für Physik for providing some computing infrastructure and supporting the data preservation project of the HERA experiments. M.A was supported through DOE Contract No. DE-AC05-06OR23177 under which JSA operates the Thomas Jefferson National Accelerator Facility, and by the University of California, Office of the President award number 00010100. B.N. was supported by the Department of Energy, Office of Science under contract number DE-AC02-05CH11231.

Appendix A. Detailed resolution plots for the ATHENA fast simulation

In this section detailed resolution plots for the variables Q^2 (Figure A.13), y (Figure A.14) and x (Figure A.15) for the ATHENA fast simulation at the EIC with $\sqrt{s} = 141\,\text{GeV}$ are shown. Our DNN-reconstruction method is compared to four widely used basic reconstruction methods (IΣ, hadron, double-angle and electron-method) for $Q^2 > 200\,\text{GeV}^2$ in five kinematic ranges in y_{gen}.

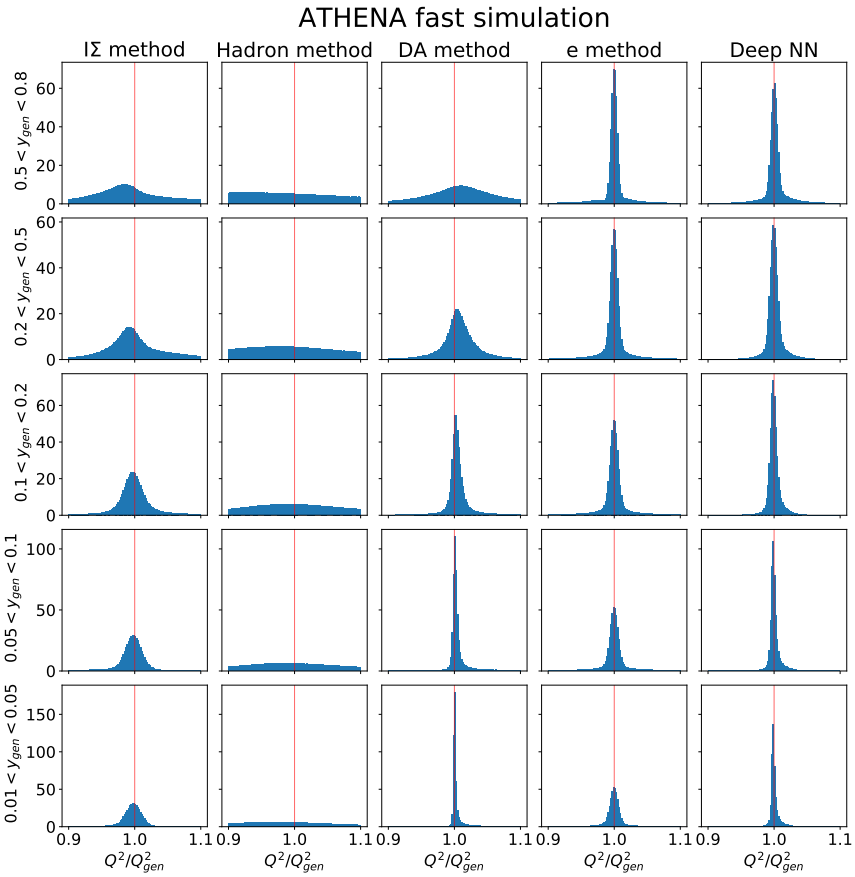

Figure A.13: Resolutions for Q^2 in various y ranges and reconstruction methods from the DELPHES fast simulation of ATHENA for events with $Q^2 > 200$ GeV^2. The tails in the distributions are from events with ISR and FSR radiation.

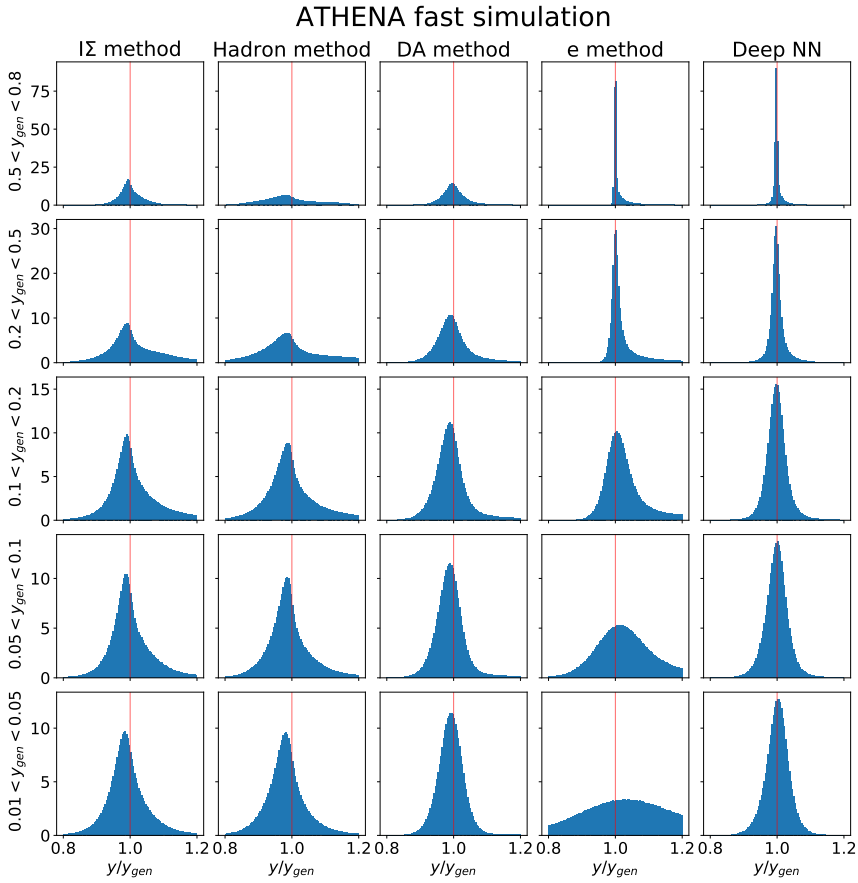

Figure A.14: Resolutions for y in various y ranges and reconstruction methods from the DELPHES fast simulation of ATHENA for events with $Q^2 > 200$ GeV^2. The tails in the distributions are from events with ISR and FSR radiation.

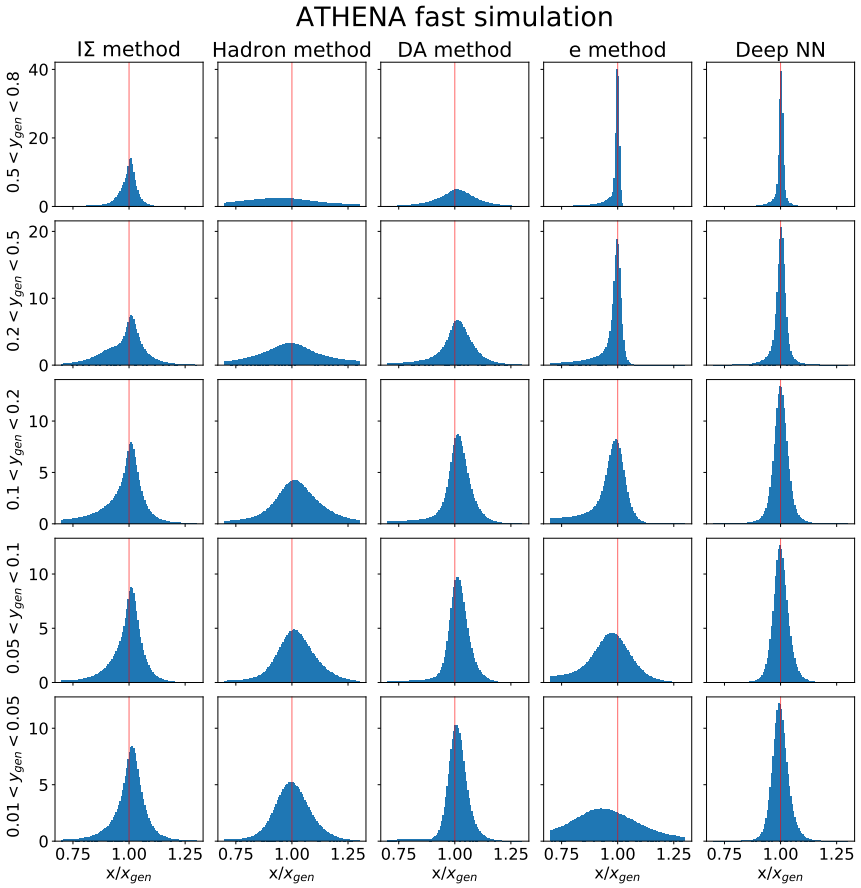

Figure A.15: Resolutions for x in various y ranges and reconstruction methods from the DELPHES fast simulation of ATHENA for events with $Q^2 > 200$ GeV^2. The tails in the distributions are from events with ISR and FSR radiation.

Appendix B. Detailed resolution plots for the H1 full simulation

In this section detailed resolution plots for the variables Q^2 (Figures B.16 and B.17), y (Figure B.18) and x (Figures B.19 and B.20) for H1's full simulation at HERA at $\sqrt{s} = 319\,\text{GeV}$ are shown. Our DNN-reconstruction method is compared to the full set of eight basic reconstruction methods (IΣ, hadron, double-angle (DA), electron (e), IDA, $E_0E\Sigma$, $E_0\theta\Sigma$, and $\theta\Sigma\gamma$-method) for $Q^2 > 200\,\text{GeV}^2$ in five kinematic ranges in y_{gen}. The double-energy method is omitted since E_h is not measureable du to relevant acceptance losses through the forward beam-hole. For y there are only four unambiguous independent basic methods.

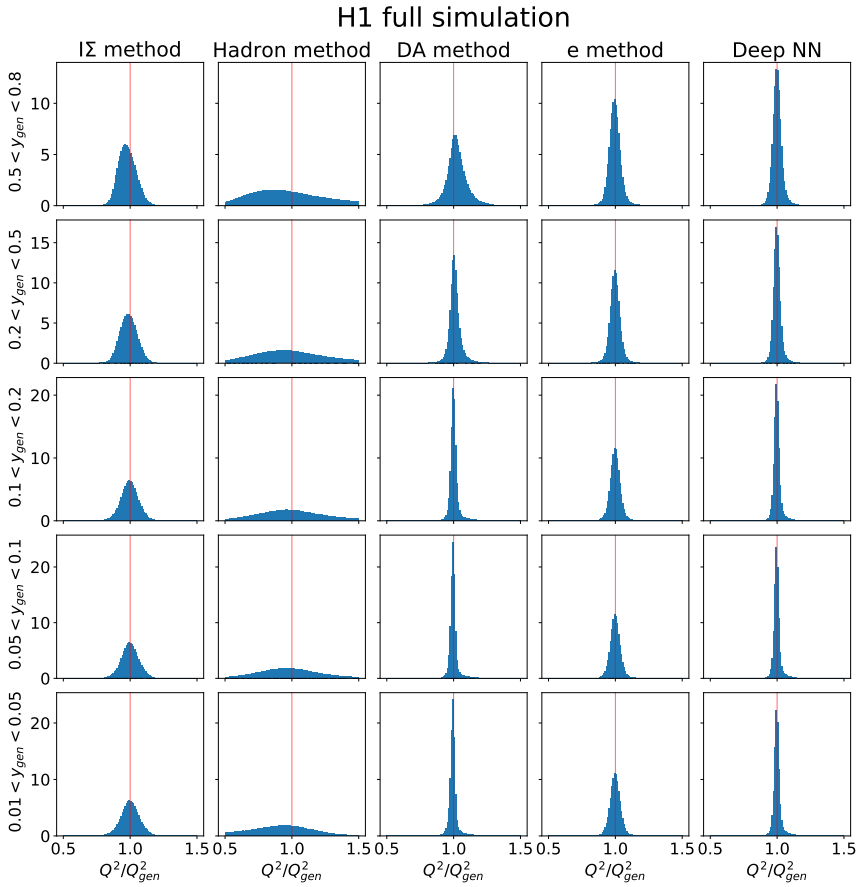

Figure B.16: Resolutions for Q^2 in various y ranges and reconstruction methods from the full simulation of H1 for events with $Q^2 > 200\ \text{GeV}^2$.

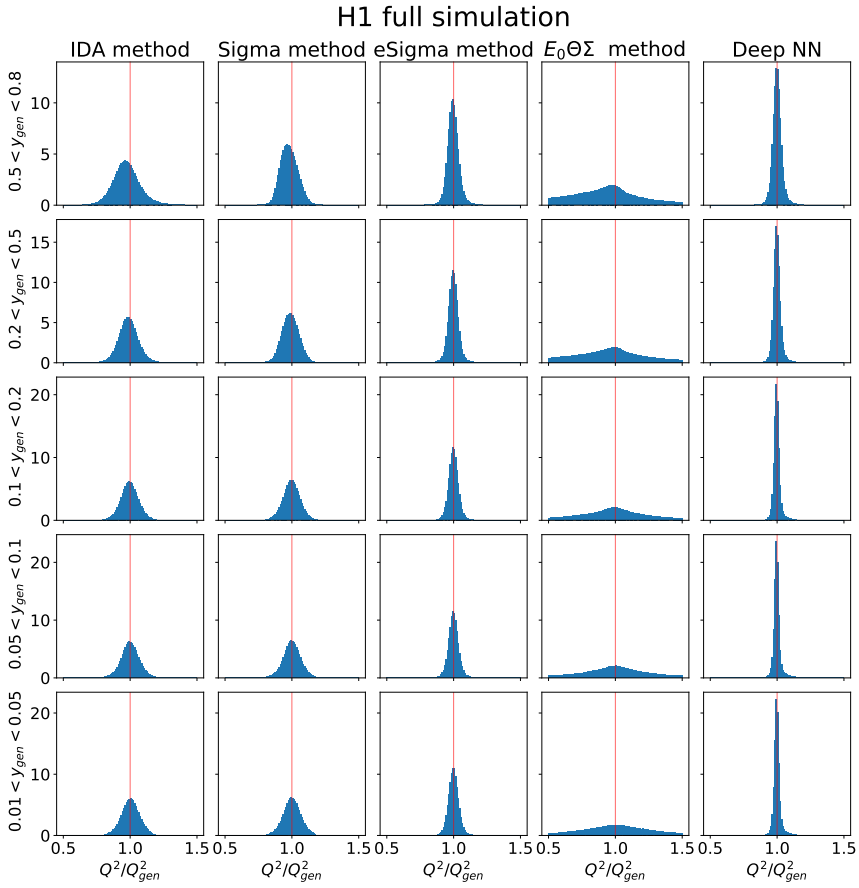

Figure B.17: Resolutions for Q^2 in various y ranges and reconstruction methods from the full simulation of H1 for events with $Q^2 > 200$ GeV^2.

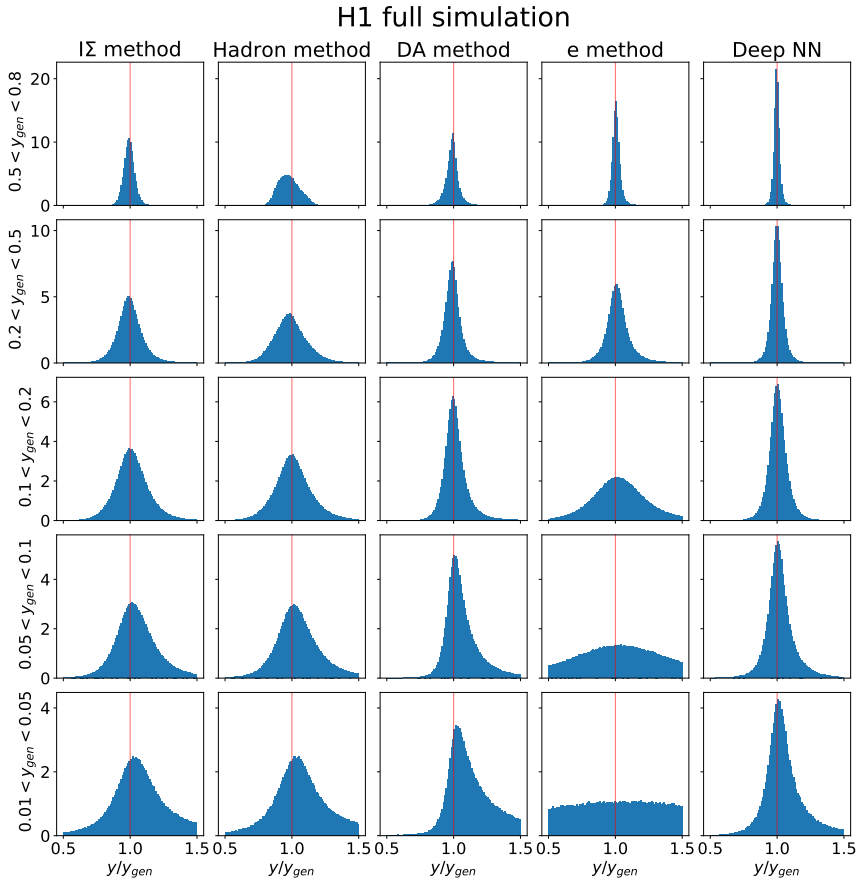

Figure B.18: Resolutions for y in various y ranges and reconstruction methods from the full simulation of H1 for events with $Q^2 > 200$ GeV^2.

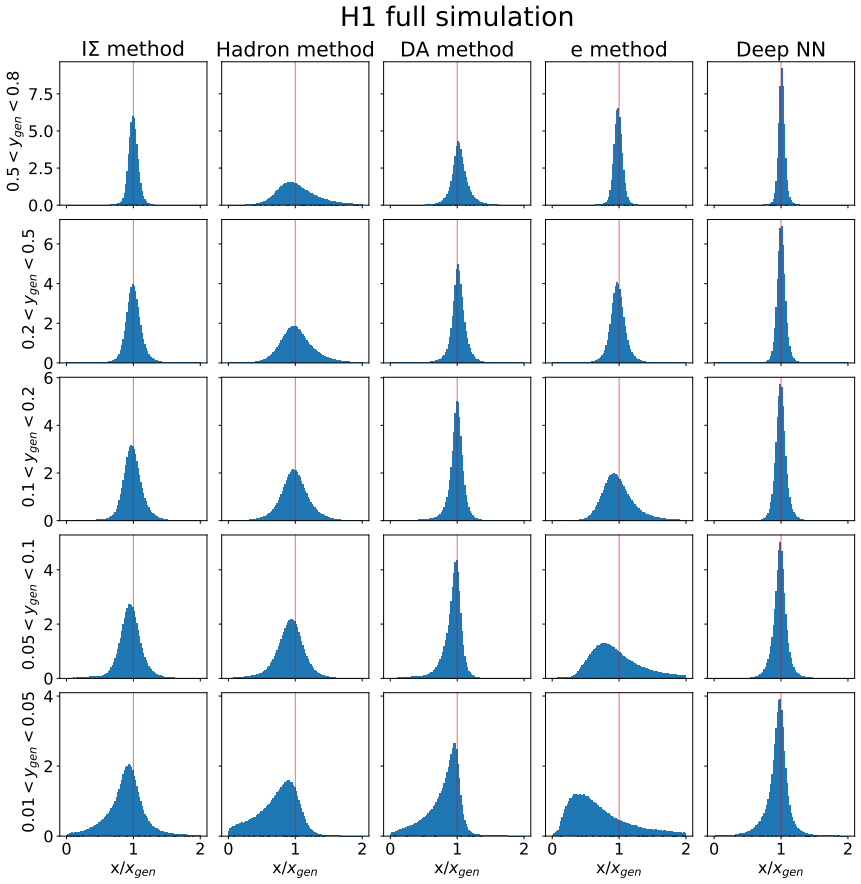

Figure B.19: Resolutions for x in various y ranges and reconstruction methods from the full simulation of H1 for events with $Q^2 > 200$ GeV^2.

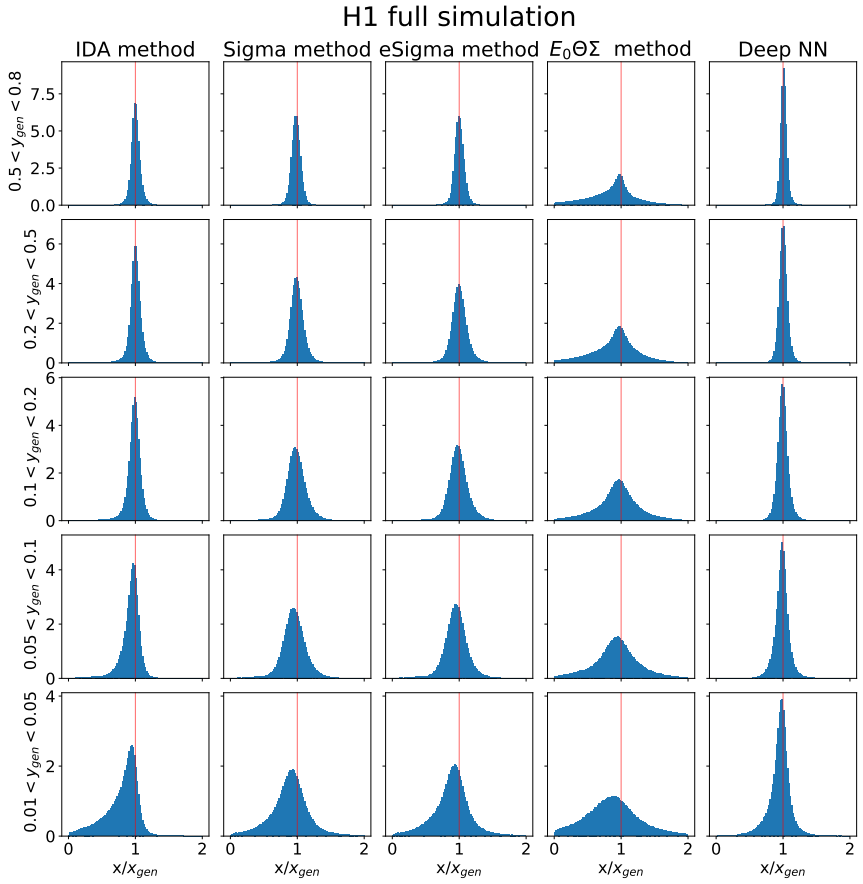

Figure B.20: Resolutions for x in various y ranges and reconstruction methods from the full simulation of H1 for events with $Q^2 > 200$ GeV^2.

References

[1] R. K. Ellis, W. J. Stirling, B. R. Webber, QCD and Collider Physics, Cambridge University Press, 1996. doi:10.1017/CBO9780511628788.

[2] R. Devenish, A. Cooper-Sarkar, Deep inelastic scattering, Oxford University Press, 2004. doi:10.1093/acprof:oso/9780198506713.001.0001.

[3] P. A. Zyla, et al., Review of Particle Physics, PTEP 2020 (2020) 083C01. doi:10.1093/ptep/ptaa104.

[4] A. Blondel, F. Jacquet, DETECTION AND STUDY OF THE CHARGED CURRENT EVENT , in: U. Amaldi, et al. (Eds.), Proceedings of the Study of an *ep* Facility for Europe, 1979, p. 391.

[5] J. Blümlein, M. Klein, Kinematics and resolution at future *ep* colliders, in: 1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90), 1990, pp. 549–551.

[6] K. C. Hoeger, Measurement of x, y, Q^2 in neutral current events, in: Workshop on Physics at HERA 1991, 1991.

[7] S. Bentvelsen, J. Engelen, P. Kooijman, Reconstruction of (x, Q^2) and extraction of structure functions in neutral current scattering at HERA, in: Workshop on Physics at HERA 1991, 1992.

[8] U. Bassler, G. Bernardi, Progress on Kinematical Variables Reconstruction. Consequences for D.I.S. Physics Analysis at Low x, H1 internal note, H1-03/93-274 (Mar 1993).

[9] U. Bassler, G. Bernardi, On the kinematic reconstruction of deep inelastic scattering at HERA: The Sigma method, Nucl. Instrum. Meth. A 361 (1995) 197–208. arXiv:hep-ex/9412004, doi:10.1016/0168-9002(95)00173-5.

[10] M. Derrick, et al., Measurement of the F2 structure function in deep inelastic e^+ p scattering using 1994 data from the ZEUS detector at HERA, Z. Phys. C 72 (1996) 399. arXiv:hep-ex/9607002, doi:10.1007/s002880050260.

[11] U. Bassler, G. Bernardi, Structure function measurements and kinematic reconstruction at HERA, Nucl. Instrum. Meth. A 426 (1999) 583–598. arXiv:hep-ex/9801017, doi:10.1016/S0168-9002(99)00044-3.

[12] A. Accardi, et al., Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all, Eur. Phys. J. A 52 (2016) 268. arXiv:1212.1701, doi:10.1140/epja/i2016-16268-9.

[13] R. Abdul Khalek, et al., Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report (Mar 2021). arXiv:2103.05419.

[14] D. P. Anderle, et al., Electron-ion collider in China, Front. Phys. (Beijing) 16 (2021) 64701. arXiv:2102.09222, doi:10.1007/s11467-021-1062-0.

[15] J. L. Abelleira Fernandez, et al., A Large Hadron Electron Collider at CERN: Report on the Physics and Design Concepts for Machine and Detector, J. Phys. G 39 (2012) 075001. arXiv:1206.2913, doi:10.1088/0954-3899/39/7/075001.

[16] P. Agostini, et al., The Large Hadron-Electron Collider at the HL-LHC (Jul 2020). arXiv:2007.14491.

[17] L. De Oliveira, B. Nachman, M. Paganini, Electromagnetic Showers Beyond Shower Shapes, Nucl. Instrum. Meth. A 951 (2020) 162879. arXiv:1806.05667, doi:10.1016/j.nima.2019.162879.

[18] D. Belayneh, et al., Calorimetry with Deep Learning: Particle Simulation and Reconstruction for Collider Physics (Dec 2019). arXiv:1912.06794, doi:10.1140/epjc/s10052-020-8251-9.

[19] ATLAS Collaboration, Deep Learning for Pion Identification and Energy Calibration with the ATLAS Detector, ATL-PHYS-PUB-2020-018 (2020).
URL `http://cdsweb.cern.ch/record/2724632`

[20] P. Abratenko, et al., Convolutional neural network for multiple particle identification in the MicroBooNE liquid argon time projection chamber, Phys. Rev. D 103 (2021) 092003. arXiv:2010.08653, doi:10.1103/PhysRevD.103.092003.

[21] A. Aurisano, A. Radovic, D. Rocco, A. Himmel, M. D. Messier, E. Niner, G. Pawloski, F. Psihas, A. Sousa, P. Vahle, A Convolutional Neural Network Neutrino Event Classifier, JINST 11 (2016) P09001. arXiv:1604.01444, doi:10.1088/1748-0221/11/09/P09001.

[22] A. M. Sirunyan, et al., A Deep Neural Network for Simultaneous Estimation of b Jet Energy and Resolution, Comput. Softw. Big Sci. 4 (2020) 10. arXiv:1912.06046, doi:10.1007/s41781-020-00041-z.

[23] Generalized Numerical Inversion: A Neural Network Approach to Jet Calibration, Tech. Rep. ATL-PHYS-PUB-2018-013, CERN, Geneva (Jul 2018).
URL http://cds.cern.ch/record/2630972

[24] Simultaneous Jet Energy and Mass Calibrations with Neural Networks, Tech. Rep. ATL-PHYS-PUB-2020-001, CERN, Geneva (Jan 2020).
URL http://cds.cern.ch/record/2706189

[25] P. Baldi, L. Blecher, A. Butter, J. Collado, J. N. Howard, F. Keilbach, T. Plehn, G. Kasieczka, D. Whiteson, How to GAN Higher Jet Resolution (Dec 2020). arXiv:2012.11944.

[26] G. Kasieczka, M. Luchmann, F. Otterpohl, T. Plehn, Per-Object Systematics using Deep-Learned Calibration (Mar 2020). arXiv:2003.11099, doi:10.21468/SciPostPhys.9.6.089.

[27] S. Cheong and A. Cukierman and B. Nachman and M. Safdari and A. Schwartzman, Parametrizing the Detector Response with Neural Networks, JINST 15 (2020) P01030. arXiv:1910.03773, doi:10.1088/1748-0221/15/01/P01030.

[28] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, A. Schwartzman, Jet-images — deep learning edition, JHEP 07 (2016) 069. arXiv:1511.05190, doi:10.1007/JHEP07(2016)069.

[29] A. M. Sirunyan, et al., A deep neural network to search for new long-lived particles decaying to jets, Mach. Learn. Sci. Tech. 1 (2020) 035012. arXiv:1912.12238, doi:10.1088/2632-2153/ab9023.

[30] M. Aaboud, et al., Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC, Eur. Phys. J. C 79 (2019) 375. arXiv:1808.07858, doi:10.1140/epjc/s10052-019-6847-8.

[31] A. Butter, et al., The Machine Learning Landscape of Top Taggers, SciPost Phys. 7 (2019) 014. arXiv:1902.09914, doi:10.21468/SciPostPhys.7.1.014.

[32] A. Andreassen, P. T. Komiske, E. M. Metodiev, B. Nachman, J. Thaler, OmniFold: A Method to Simultaneously Unfold All Observables, Phys. Rev. Lett. 124 (2020) 182001. arXiv:1911.09107, doi:10.1103/PhysRevLett.124.182001.

[33] K. Datta, D. Kar, D. Roy, Unfolding with Generative Adversarial Networks (2018). arXiv:1806.00433.

[34] M. Bellagente, A. Butter, G. Kasieczka, T. Plehn, R. Winterhalder, How to GAN away Detector Effects (2019). arXiv:1912.00477, doi:10.21468/SciPostPhys.8.4.070.

[35] A. Glazov, Machine learning as an instrument for data unfolding (Dec 2017). arXiv:1712.01814.

[36] M. Vandegar, M. Kagan, A. Wehenkel, G. Louppe, Neural Empirical Bayes: Source Distribution Estimation and its Applications to Simulation-Based Inference, in: A. Banerjee, K. Fukumizu (Eds.), Proceedings of The 24th International Conference on Artificial Intelligence and Statistics, Vol. 130 of Proceedings of Machine Learning Research, PMLR, 2021, pp. 2107–2115. arXiv:2011.05836.
URL https://proceedings.mlr.press/v130/vandegar21a.html

[37] J. N. Howard, S. Mandt, D. Whiteson, Y. Yang, Foundations of a Fast, Data-Driven, Machine-Learned Simulator (Jan 2021). arXiv:2101.08944.

[38] P. Baroň, Comparison of Machine Learning Approach to other Unfolding Methods (Apr 2021). arXiv:2104.03036.

[39] A. Andreassen, P. T. Komiske, E. M. Metodiev, B. Nachman, A. Suresh, J. Thaler, Scaffolding Simulations with Deep Learning for High-dimensional Deconvolution (May 2021). arXiv:2105.04448.

[40] V. Andreev, et al., Measurement of lepton-jet correlation in deep-inelastic scattering with the H1 detector using machine learning for unfolding (Aug 2021). arXiv:2108.12376.

[41] M. Feickert and B. Nachman, A Living Review of Machine Learning for Particle Physics (2021). arXiv:2102.02770.

[42] A. J. Larkoski, I. Moult, B. Nachman, Jet Substructure at the Large Hadron Collider: A Review of Recent Advances in Theory and Ma-

chine Learning, Phys. Rept. 841 (2020) 1–63. arXiv:1709.04464, doi:10.1016/j.physrep.2019.11.001.

[43] D. Guest, K. Cranmer, D. Whiteson, Deep Learning and its Application to LHC Physics (2018). arXiv:1806.11484, doi:10.1146/annurev-nucl-101917-021019.

[44] K. Albertsson, et al., Machine Learning in High Energy Physics Community White Paper (2018). arXiv:1807.02876, doi:10.1088/1742-6596/1085/2/022008.

[45] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. 91 (2019) 045002. arXiv:1903.10563, doi:10.1103/RevModPhys.91.045002.

[46] D. Bourilkov, Machine and Deep Learning Applications in Particle Physics, Int. J. Mod. Phys. A 34 (2020) 1930019. arXiv:1912.08245, doi:10.1142/S0217751X19300199.

[47] M. Diefenthaler, A. Farhat, A. Verbytskyi, Y. Xu, Deeply Learning Deep Inelastic Scattering Kinematics (Aug 2021). arXiv:2108.11638.

[48] M. Klein, R. Yoshida, Collider Physics at HERA, Prog. Part. Nucl. Phys. 61 (2008) 343. arXiv:0805.3334, doi:10.1016/j.ppnp.2008.05.002.

[49] H. Abramowicz, et al., Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data, Eur. Phys. J. C 75 (2015). arXiv:1506.06042, doi:10.1140/epjc/s10052-015-3710-4.

[50] F. D. Aaron, et al., A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA, Eur. Phys. J. C 64 (2009) 561. arXiv:0904.3513, doi:10.1140/epjc/s10052-009-1169-x.

[51] F. D. Aaron, et al., Inclusive Deep Inelastic Scattering at High Q^2 with Longitudinally Polarised Lepton Beams at HERA, JHEP 09 (2012) 061. arXiv:1206.7007, doi:10.1007/JHEP09(2012)061.

[52] V. Andreev, et al., Measurement of inclusive ep cross sections at high Q^2 at $\sqrt{s} = 225$ and 252 GeV and of the longitudinal proton structure function F_L at HERA, Eur. Phys. J. C 74 (2014) 2814. arXiv:1312.4821, doi:10.1140/epjc/s10052-014-2814-6.

[53] H. Spiesberger, et al., Radiative corrections at HERA (1992) 798–839. URL https://cds.cern.ch/record/237380

[54] A. Kwiatkowski, H. Spiesberger, H. J. Mohring, Characteristics of radiative events in deep inelastic *ep* scattering at HERA, Z. Phys. C 50 (1991) 165–178. doi:10.1007/BF01558572.

[55] A. Kwiatkowski, H. Spiesberger, H. J. Mohring, Heracles: An Event Generator for *ep* Interactions at HERA Energies Including Radiative Processes: Version 1.0, Comput. Phys. Commun. 69 (1992) 155–172. doi:10.1016/0010-4655(92)90136-M.

[56] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al., Tensorflow: A system for large-scale machine learning., in: OSDI, Vol. 16, 2016, pp. 265–283.

[57] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, M. Selvaggi, DELPHES 3, A modular framework for fast simulation of a generic collider experiment, JHEP 02 (2014) 057. arXiv:1307.6346, doi:10.1007/JHEP02(2014)057.

[58] M. Arratia, S. Sekula, A delphes card for the eic yellow-report detector (Mar 2021). doi:10.5281/zenodo.4592887.
URL https://doi.org/10.5281/zenodo.4592887

[59] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, P. Zanarini, GEANT3, CERN-DD-EE-84-01 (Sep 1987).

[60] I. Abt, et al., The H1 detector at HERA, Nucl. Instrum. Meth. A 386 (1997) 310–347. doi:10.1016/S0168-9002(96)00893-5.

[61] I. Abt, et al., The Tracking, calorimeter and muon detectors of the H1 experiment at HERA, Nucl. Instrum. Meth. A 386 (1997) 348–396. doi:10.1016/S0168-9002(96)00894-7.

[62] H. Jung, Hard diffractive scattering in high-energy e p collisions and the Monte Carlo generator RAPGAP, Comput. Phys. Commun. 86 (1995) 147–161. doi:10.1016/0010-4655(94)00150-Z.

[63] H.-U. Bengtsson, T. Sjostrand, The Lund Monte Carlo for Hadronic Processes: Pythia Version 4.8, Comput. Phys. Commun. 46 (1987) 43. doi:10.1016/0010-4655(87)90036-1.

[64] G. Ingelman, A. Edin, J. Rathsman, LEPTO 6.5: A Monte Carlo generator for deep inelastic lepton - nucleon scattering, Comput. Phys. Commun. 101 (1997) 108–134. arXiv:hep-ph/9605286, doi:10.1016/S0010-4655(96)00157-9.

[65] M. Dobbs, J. B. Hansen, The HepMC C++ Monte Carlo event record for High Energy Physics, Comput. Phys. Commun. 134 (2001) 41–46. doi:10.1016/S0010-4655(00)00189-2.

[66] K. Charchula, G. A. Schuler, H. Spiesberger, Combined QED and QCD radiative effects in deep inelastic lepton - proton scattering: The Monte Carlo generator DJANGO6, Comput. Phys. Commun. 81 (1994) 381–402. doi:10.1016/0010-4655(94)90086-8.

[67] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[68] G. Klambauer, T. Unterthiner, A. Mayr, S. Hochreiter, Self-normalizing neural networks, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems, Vol. 30, Curran Associates, Inc., 2017.
URL https://proceedings.neurips.cc/paper/2017/file/5d44ee6f2c3f71b73125876103c8f6c4-Paper.pdf

[69] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization (2017). arXiv:1412.6980.

[70] P. J. Huber, Robust Estimation of a Location Parameter, The Annals of Mathematical Statistics 35 (1964) 73 – 101. doi:10.1214/aoms/1177703732.
URL https://doi.org/10.1214/aoms/1177703732

[71] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, W. K. Tung, New generation of parton distributions with uncertainties from global QCD analysis, JHEP 07 (2002) 012. arXiv:hep-ph/0201195, doi:10.1088/1126-6708/2002/07/012.

[72] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, Parton fragmentation and string dynamics, Phys. Rept. 97 (1983) 31–145. doi:10.1016/0370-1573(83)90080-7.

[73] S. Schael, et al., Bose-Einstein correlations in W-pair decays with an event-mixing technique, Phys. Lett. B 606 (2005) 265–275. doi:10.1016/j.physletb.2004.12.018.

[74] H. Fesefeldt, The Simulation of Hadronic Showers: Physics and Applications, PITHA-85-02 (Dec 1985).

[75] G. Grindhammer, M. Rudowicz, S. Peters, The Fast Simulation of Electromagnetic and Hadronic Showers, Nucl. Instrum. Meth. A 290 (1990) 469. doi:10.1016/0168-9002(90)90566-O.

[76] J. Gayler, Simulation of H1 calorimeter test data with GHEISHA and FLUKA, in: Workshop on Detector and Event Simulation in High-energy Physics (MC '91), 1991, p. 312.

[77] M. Kuhlen, The Fast H1 detector Monte Carlo, in: 26th International Conference on High-energy Physics, 1992, pp. 1787–1790. doi:10.1063/1.43288.

[78] G. Grindhammer, S. Peters, The Parameterized simulation of electromagnetic showers in homogeneous and sampling calorimeters, in: International Conference on Monte Carlo Simulation in High-Energy and Nuclear Physics (MC '93), 1993. arXiv:hep-ex/0001020.

[79] A. Glazov, N. Raicevic, A. Zhokin, Fast simulation of showers in the H1 calorimeter, Comput. Phys. Commun. 181 (2010) 1008–1012. doi:10.1016/j.cpc.2010.02.004.

[80] M. Peez, Search for deviations from the standard model in high transverse energy processes at the electron proton collider HERA. (Thesis, Univ. Lyon), Ph.D. thesis (Jun 2003).

[81] S. Hellwig, Untersuchung der D^* - π slow Double Tagging Methode in Charmanalysen, Diploma thesis, Univ. Hamburg (Jun 2004).

[82] B. Portheault, First measurement of charged and neutral current cross sections with the polarized positron beam at HERA II and QCD-electroweak analyses. (Thesis, Univ. Paris XI), Ph.D. thesis (Mar 2005).

[83] V. Andreev, et al., Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s, Eur. Phys. J. C 75 (2015) 65. arXiv:1406.4709, doi:10.1140/epjc/s10052-014-3223-6.

[84] R. Kogler, Measurement of jet production in deep-inelastic e p scattering at HERA, Phd thesis (Feb 2011). doi:10.3204/DESY-THESIS-2011-003.

[85] D. Britzger, S. Levonian, S. Schmitt, D. South, Preservation through modernisation: The software of the H1 experiment at HERA, EPJ Web Conf. 251 (2021) 03004. arXiv:2106.11058, doi:10.1051/epjconf/202125103004.

[86] M. Peez, et al., An energy flow algorithm for hadronic reconstruction in oo: Hadroo2, H1-Internal Note, vol. H1-01/05-616, DESY, 2005 (2005).

[87] R. Brun, F. Rademakers, Root - an object oriented data analysis framework, in: AIHENP'96 Workshop, Lausane, Vol. 389, 1996, pp. 81–86.