Deutsches Elektronen-Synchrotron DESY

Machine learning-assisted measurement of multi-differential lepton-jet correlations in deep-inelastic scattering with the H1 detector

H1 Collaboration

Abstract

The lepton-jet momentum imbalance in deep inelastic scattering events offers a useful set of observables for unifying collinear and transverse-momentum-dependent frameworks for describing high energy Quantum Chromodynamics (QCD) interactions. We recently performed a measurement of this imbalance in the laboratory frame using positron-proton collisions from HERA Run II [1]. With a new machine learning method, the measurement was performed simultaneously and unbinned in eight dimensions. The results in Ref. [1] were presented projected onto four key observables. This paper extends those results by showing the multi-differential nature of the unfolded result. In particular, we present lepton-jet correlation observables deferentially in kinematic properties of the scattering process, Q^2 and y. We compare these results with parton shower Monte Carlo predictions as well as calculations from perturbative QCD and from a Transverse Momentum Dependent (TMD) factorization framework.

1 Introduction

Jet measurements in deep-inelastic scattering (DIS) at HERA have provided a powerful tool to explore quantum-chromodynamics (QCD), including studies to constrain gluon parton-density functions (PDFs) and to extract the strong coupling constant [2]. While powerful, such measurements have mostly been limited to the domain of collinear QCD. Theoretical and experimental progress in recent years motivate studying aspects of QCD at high momentum transfer (Q) that go beyond the collinear framework. In particular, transverse momentum dependent (TMD) PDFs have a more complex evolution than their collinear counterparts and observables sensitive to TMD PDFs and related quantities will aid the exploration of nucleon structure in multiple dimensions of position and momentum space, as well as quantum correlation amongst nuclear constituents.

Lepton-jet correlations in DIS have recently been suggested as interesting probes of TMD PDFs and TMD evolution [3,4]. These jet-based studies complement traditional measurements using single hadrons (semi-inclusive DIS or SIDIS) by providing observables that can be described theoretically without involving TMD fragmentation functions (FFs). Such decoupling between TMD PDFs and TMD FFs is currently one of the main challenges for accurate extractions of TMD PDFs with global fits.
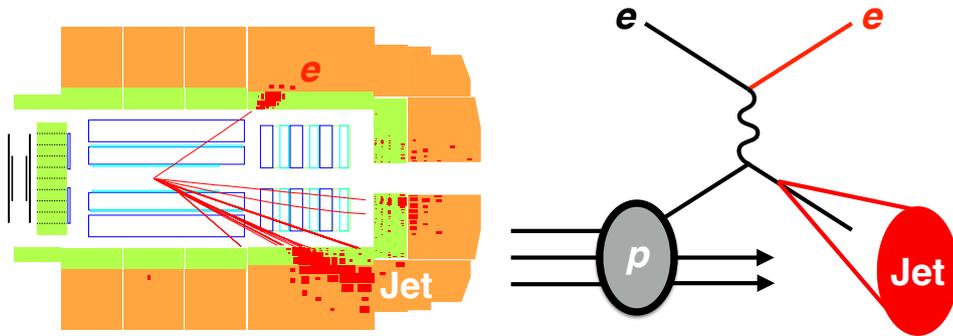

Fig. 1: Left: A display of the H1 tracker (open rectangles) and calorimeter detectors (filled rectangles), showing a neutral current DIS candidate event involving a single jet in the laboratory frame. Right: A leading order Feynman diagram illustrating a contribution to the process in the top display.

Previously, we presented measurements of lepton-jet correlations in the laboratory frame for events $Q^2 > 150$ GeV^2 [1] using H1 data from HERA Run II. In particular, the result included differential cross section measurements of the jet transverse[1] momentum, psseudorapdiity η, relative transverse electron-jet momentum imbalance q_T/Q, and angular separation in the transverse plane, $\Delta\phi$. Analytical calculations within the TMD framework agree well with the data from low to medium values of q_T/Q or $\Delta\phi$, whereas collinear perturbative QCD calculations at next-to-next-to-leading order (NNLO) describe the data well at medium to large values. At intermediate values with $q_T/Q \sim 0.5$, both frameworks agree with the data and with each other. Such observations represent a long-sought matching between the two frameworks [5], which is not observed at lower-energy DIS from fixed target experiments. The data are also well described by Monte Carlo (MC) event generators, including CASCADE [6] that includes TMD effects through the parton-branching method.

Multi-differential measurements of lepton-jet observables in DIS over a wide range of Q^2 are needed to fully constrain TMD evolution effects. The measurement in Ref. [1] was performed simultaneously and unbinned in eight dimensions including the three each for the jet and lepton kinematic properties, q_T/Q, and $\Delta\phi$ (redundant with the first six). This measurement was enabled by a new machine learning technique called MULTIFOLD [7,8]. Even though the data were unfolded simultaneously in eight dimen-

---

[1]This measurement uses a right handed coordinate system defined such that the positive z direction points in the direction of the proton beam and the nominal interaction point is located at $z = 0$. The polar angle θ, is defined with respect to this axis. The pseudorapidity is defined as $\eta_{\text{lab}} = -\ln\tan(\theta/2)$.

sions, the final results of Ref. [1] were presented as four binned differential cross section measurements. The goal of this paper is to explore the unbinned and multidimensional nature of the previous result.

This note is organized as follows. Section 2 briefly introduces the H1 detector and Sec. 3 lists the experimental and synthetic datasets used for the analysis. The event reconstruction is described in Sec. 4. Corrections for detector effects (unfolding) using the MULTIFOLD algorihtm are detailed in Sec. 5. A discussion of uncertainties is in Sec. 6 and theoretical predictions are described in Sec. 7. Results are presented in Sec. 8 and the note ends with conclusions and outlook in Sec. 9.

2   H1 Detector

The H1 detector [9–13] is a general purpose particle detector with cylindrical geometry. The main sub-detectors used in this analysis are the inner tracking detectors and the Liquid Argon (LAr) calorimeter, which are both immersed in a magnetic field of 1.16 T provided by a superconducting solenoid. The central tracking system, which covers $15° < \theta < 165°$ and the full azimuthal angle, consists of drift and proportional chambers that are complemented with a silicon vertex detector in the range $30° < \theta < 150°$ [14]. The tracker yields a transverse momentum resolution for charged particles of $\sigma_{p_T}/p_T = 0.2\% \ p_T/\text{GeV} \oplus 1.5\%$. The LAr calorimeter, which covers $4° < \theta < 154°$ and full azimuthal angle, consists of an electromagnetic section made of lead absorbers and a hadronic section with steel absorbers; both are highly segmented in the transverse and longitudinal directions. The calorimeter energy resolution is $\sigma_E/E = 11\%/\sqrt{E/\text{GeV}} \oplus 1\%$ for leptons [15] and $\sigma_E/E \approx 50\%/\sqrt{E/\text{GeV}} \oplus 3\%$ for charged pions [16]. In the backward region ($153° < \theta < 177.5°$), energies are measured with a lead-scintillating fiber calorimeter [17].

3   Data and Simulated Samples

Data collected for this analysis were collected with the H1 detector in the years 2006 and 2007 when positrons and protons were collided at energies of 27.6 GeV and 920 GeV, respectively, for a center-of-mass energy of about $\sqrt{s} = 320$ GeV. The total integrated luminosity of this data sample corresponds to 136 pb^{-1} [18].

Events in data were selected online using a trigger selecting events with a high energy cluster in the electromagnetic part of the LAr calorimeter. The scattered lepton is identified with the highest transverse momentum LAr cluster matched to a track, and is required to pass the isolation criteria described in Ref. [19]. After fiducial cuts, the trigger efficiency is higher than 99.5% [20, 21] for scattered lepton candidates with energy $E_{e'} > 11$ GeV. A series of fiducial and quality cuts based on simulations [21, 22] suppress backgrounds to a negligible level.

Monte Carlo (MC) simulations are used to correct the data for detector acceptance and resolution effects. Two generators are used for this purpose: DJANGOH [23] 1.4 and RAPGAP [24] 3.1. Both generators implement Born level matrix elements for the neutral current DIS, boson–gluon fusion, and QCD Compton processes and are interfaced with HERACLES [25–27] for QED radiation. The CTEQ6L PDF set [28] and the Lund hadronization model [29] with parameters fitted by the ALEPH Collaboration [30] are used for the non-perturbative components. DJANGOH uses the Colour Dipole Model as implemented in ARIADNE [31] for higher order emissions, and RAPGAP uses parton showers in the leading logarithmic approximation.

Each of these generators is combined with a detailed simulation of the H1 detector response based on the GEANT3 simulation program [32] and reconstructed in the same way as data.

4 Event Reconstruction

DIS reactions are governed by Q^2 and the inelasticity y, or equivalently, the longitudinal momentum fraction $x = Q^2/(sy)$. We use the Σ method [33] to reconstruct Q^2 and y as:

$$Q^2 = \frac{E_{e'}^2 \sin^2 \theta_{e'}}{1 - y} \tag{1}$$

$$y = \frac{\sum_{i \in \text{had}}(E_i - p_{i,z})}{\sum_{i \in \text{had}}(E_i - p_{i,z}) + E_{e'}(1 - \cos \theta_{e'})}, \tag{2}$$

where $\theta_{e'}$ is the polar angle of the scattered lepton and $\sum(E_i - p_{i,z})$ is the total difference between the energy and longitudinal momentum of the entire hadronic final state. Compared to other methods, the Σ reconstruction reduces sensitivity to collinear initial state Quantum Electrodynamic (QED) radiation, $e \to e\gamma$, since the beam energies are not included in the calculation.

The FASTJET 3.3.2 package [34, 35] is used to cluster jets in the laboratory frame with the inclusive k_T algorithm [36, 37] and distance parameter $R = 1$. The inputs for the jet clustering are hadronic final state (HFS) objects with $-1.5 < \eta_{\text{lab}} < 2.75$. These objects are built from calorimeter-cell clusters and reconstructed tracks, after removing those associated with the scattered lepton, using an energy flow algorithm [38–40]. Jets with transverse momentum $p_T^{\text{jet}} > 5$ GeV are selected for further analysis.

The input for the jet clustering at the generator level ("particle level") are final-state particles with proper lifetime $c\tau > 10$ mm generated with RAPGAP or DJANGOH, excluding the scattered lepton. Reconstructed jets (reco) are matched to the generated jets (gen) with an angular distance selection of $\Delta R^2 = (\phi_{\text{gen}}^{\text{jet}} - \phi_{\text{reco}}^{\text{jet}})^2 + (\eta_{\text{gen}}^{\text{jet}} - \eta_{\text{reco}}^{\text{jet}})^2 < 0.9^2$.

Events with $Q^2 > 150$ GeV^2, $0.08 < y < 0.7$, and at least one jet participate in the unfolding (Sec. 5). The final measurement is presented in a fiducial volume defined by $Q^2 > 150\,\text{GeV}^2$, $0.2 < y < 0.7$, $p_T^{\text{jet}} > 10$ GeV, and $-1.0 < \eta_{\text{lab}}^{\text{jet}} < 2.5$; the total inclusive jet cross section in this region is denoted σ_{jet}.

5   Unfolding

The MULTIFOLD method is an iterative two-step procedure to correct for detector effects as illustrated in Fig. 2 illustrates this MULTIFOLD. The goal is to infer the top right box (particle-level data) using detector-level data (top left box) and simulations (lower boxes). The components of MULTIFOLD are explained in more detail below.
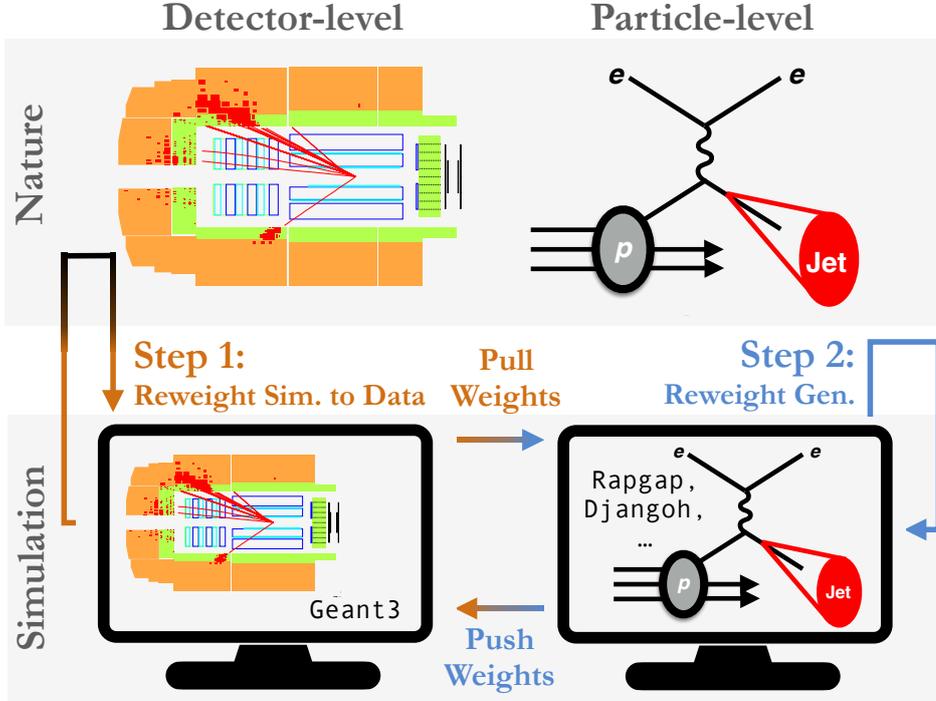

Fig. 2: A schematic diagram of the MULTIFOLD method. The top row represents data ('Nature') while the second row depicts synthetic datasets from MC simulation. The goal is to infer the top right box given the other three boxes. MULTIFOLD is an iterative two-step procedure. The first step uses detector-level inputs (left column) while the second step uses particle-level inputs (right column). Adapted from Ref. [7].

Let $\vec{x} = (p_x^e, p_y^e, p_z^e, p_T^{\text{jet}}, \eta^{\text{jet}}, \phi^{\text{jet}}, q_T^{\text{jet}}/Q, \Delta\phi^{\text{jet}})$. The goal is to make a cross section measurement that is differential in $\vec{x}$. Note that from $\vec{x}$, one can extract Q by using q_T computed from the first six components of $\vec{x}$, $\vec{q}_T = \vec{p}_T^e - \vec{p}_T^{\text{jet}}$. Additionally, y can be computed using Q, the lepton kinematic properties, and Eq. 1. Symbolically:

$$Q^2 = \frac{(q_T^{\text{jet}})^2}{(q_T^{\text{jet}}/Q)^2}$$
$$= \frac{(p_x^e + p_T^{\text{jet}} \cos(\phi^{\text{jet}}))^2 + (p_y^e + p_T^{\text{jet}} \sin(\phi^{\text{jet}}))^2}{(q_T^{\text{jet}}/Q)^2} \tag{3}$$

and

$$y = 1 - \frac{E_e^2 \sin^2 \theta_e}{Q^2}$$
$$= 1 - \frac{((p_x^e)^2 + (p_y^e)^2 + (p_z^e)^2 + m_e^2)\left(\frac{(p_x^e)^2 + (p_y^e)^2}{(p_x^e)^2 + (p_y^e)^2 + (p_z^e)^2}\right)^2}{Q^2}. \tag{4}$$

Even though Eq. 3 and 4 follow directly from the definitions of Q^2 and y, it is useful to express them in terms of $\vec{x}$ to illustrate that they are non-trivial functions of the measured phase space. Deriving measurements of Q^2 and y or other observables as a function of Q^2 and y (if these observables were not directly present in the original phase space) would not be possible with binned unfolding techniques.

The first step of MULTIFOLD uses observables at detector level while the second step operates at particle level. Define $X_{\text{data}} = \{\vec{x}_i\}$ be the set of events in data and $X_{\text{MC,gen}} = \{\vec{x}_{\text{gen},i}\}$ and $X_{\text{MC,reco}} = \{\vec{x}_{\text{reco},i}\}$ be sets of events in simulation with a correspondance between the two sets. In simulation, we have a set of observables at particle-level and detector-level for each event. If an event does not pass the particle-level or detector-level event selection, then the corresponding set of observables are assigned a dummy value $\vec{x} = \emptyset$. Each event i in simulation is also associated with a weight w_i from the MC simulation.

MULTIFOLD achieves an unbinned unfolding by iteratively reweighting the particle-level events. Each event i in simulation is given a weight ν_i and these weights are updated at each iteration. The final result is the simulated events with weights $\nu_i w_i$. From these events, one can compute new observables defined on $\vec{x}$ and can construct histograms or other summary statistics. The MULTIFOLD weights are initialized at $\nu_i = 1$, i.e. the prior is the initial MC simulation.

The first step of MULTIFOLD is to train a classifier f to distinguish the weighted simulation at detector-level from the data. The classifier is trained to maximize the common binary cross entropy:

$$\sum_{\vec{x}_i \in X_{\text{data}}} \log(f(\vec{x}_i)) + \sum_{\vec{x}_i \in X_{\text{MC,reco}}} \nu_i w_i \log(1 - f(\vec{x}_i)), \tag{5}$$

where both sums only include events that pass the detector-level selection. For events that pass the detector-level selection, define $\lambda_i = \nu_i \times f(\vec{x}_i)/(1 - f(\vec{x}_i))$ for $\vec{x}_i \in X_{\text{MC,reco}}$. This manipulation of the classifier output is known (see e.g. Ref. [41, 42]) to produce an estimate of the likelihood ratio between data and simulation. For events that do not pass the detector-level selection, $\lambda_i = \nu_i$.

The second step of MULTIFOLD is a regularization step. The weights λ_i are insufficient because they are not a proper function of the particle-level phase space. In other words, a single phase space point $\vec{x}_{\text{gen}}$ can be mapped to different $\vec{x}_{\text{reco}}$ values under the stochastic detector response. The second step of MULTIFOLD averages the weights λ for a fixed particle-level phase space point. This is accomplished by training a classifier to distinguish the particle-level simulation weighted by ν from the particle-level simulation weighted by λ. The loss function is once again the binary cross entropy:

$$\sum_{\vec{x}_i \in X_{\text{MC,truth}}} \lambda_i w_i \log(f(\vec{x}_i)) + \nu_i w_i \log(1 - f(\vec{x}_i)), \tag{6}$$

where the sum only includes events that pass the particle-level selection. For events that pass the particle-level selection, define $\nu_i = \nu_i \times f(\vec{x}_i)/(1 - f(\vec{x}_i))$ for $\vec{x}_i \in X_{\text{MC,truth}}$. For events that do not pass the particle-level selection, ν_i is left unchanged from its previous value.

The classifiers for Steps 1 and 2 are parameterized as fully connected deep neural networks. These networks are implemented in TENSORFLOW [43] and KERAS [44] and optimized using ADAM [45]. The input layer to the neural networks has 8 nodes, corresponding to the 8 dimensions of $\vec{x}$ used for unfolding. All inputs are standardized so that each dimension of $\vec{x}$ has mean zero and unit standard deviation. Following the input, there are three hidden layers, with 50, 100, and 50 nodes, respectively. Each layer has a rectified linear unit activation function and the network output is a single node with the sigmoid activation function. None of these hyperparameters were optimized and all other hyperparemeters are set to their default values. In particular, the network biases are all initialized to zero and the weights are initialaized using the Glorot uniform distribution [46]. In order to minimize variations from the stochastic nature of the training procedure, we train 10 networks for each configuration and the final result is taken as the median over the 10 values per event.

For training, the inputs are partitioned equally into a training and validation set. This partition is random and redone at each iteration. Training proceeds for 10,000 epochs with an early stopping mechanism that halts training if the validation loss does not decrease for 10 consecutive epochs. Step 1 training uses a batch size of 50,000 events and a learning rate of 2×10^{-6}, while step 2 training uses a batch size of 100,000 events and a learning rate of 5×10^{-6}. The networks are trained using NVIDIA Quadro RTM 6000 Graphical Processing Units (GPUs). These GPUs have sufficient memory (24 GB) to simultaneously fit all of the inputs and the model into memory. The training time for both Step 1 and Step 2 decreases with each iteration since the MC at particle level is reweighted to successively better match the data with each iteration. For instance, the first iteration of Step 1 takes 3350 seconds and Step requires 2500 seconds. In the fifth iteration, Step 1 only takes 660 seconds and Step 2 requires 540 seconds. Unfoldings stops after the fifth iteration after which changes in the weights are negligible.

Learning curves showing the average loss as a function of the epoch number are shown in Fig. 3. The required number of epochs for training decreases with more iterations and Step 2 requires fewer iterations than Step 1. The overall loss values increase with more iterations as the reweighted simulation looks more like the data at the start of each step. Correspondingly, the amount of learning at each step, quantified by the drop in average loss, decreases at each iteration.

Representative observable distributions corresponding to Fig. 3 are shown in Fig. 4. In particular, histograms of η^{jet} are depicted at various stages of the unfolding for the first and last MUTLIFOLD iterations. The left plots show η^{jet} at detector level while the right plots correspond to η^{jet} at particle level. Due to the asymmetry of the lepton-proton collisions, the jets are more likely to be in the forward ($\eta > 0$) direction compared to the backwards ($\eta < 0$) direction. The dashed solid line in the ratio panel of Fig. 4(a) shows the starting data/MC ratio. This ratio is a few percent high in the backwards direction and up to 20% low in the forward direction. After the Step 1 reweighting (dotted ratio), the ratio is within 1% of unity across the entire spectrum. For Step 2 of the first iteration (Fig. 4(b)), the dashed ratio indicates the difference between the intial MC and the MC with weights from Step 1. By design, this ratio is similar to the dashed ratio in Fig. 4(a). After the Step 2 averaging step, the reweighted MC is within 1% of unity for all bins except the last bin, where the agreement is at the 5% level. At the fifth iteration, the weighted MC at the beginning of Step 1 (dashed ratio) is already quite close to unity across the spectrum (Fig. 4(c)). The 10% disagreement in the last bin is brought to unity with the reweighting. The final weights of Step 2(Fig. 4(d)) are within 1% of unity across the entire spectrum. Note that for Step 2, we reweight the original MC (weights w_i) to the version with weights from Step 1 (weights $\lambda_i w_i$). Another option is to learn an incremental weight with respect to the previous Step 2. We explored both options and found a similar closure. Learning one reweighting is practically easier because it does not require multiplying together many weights.
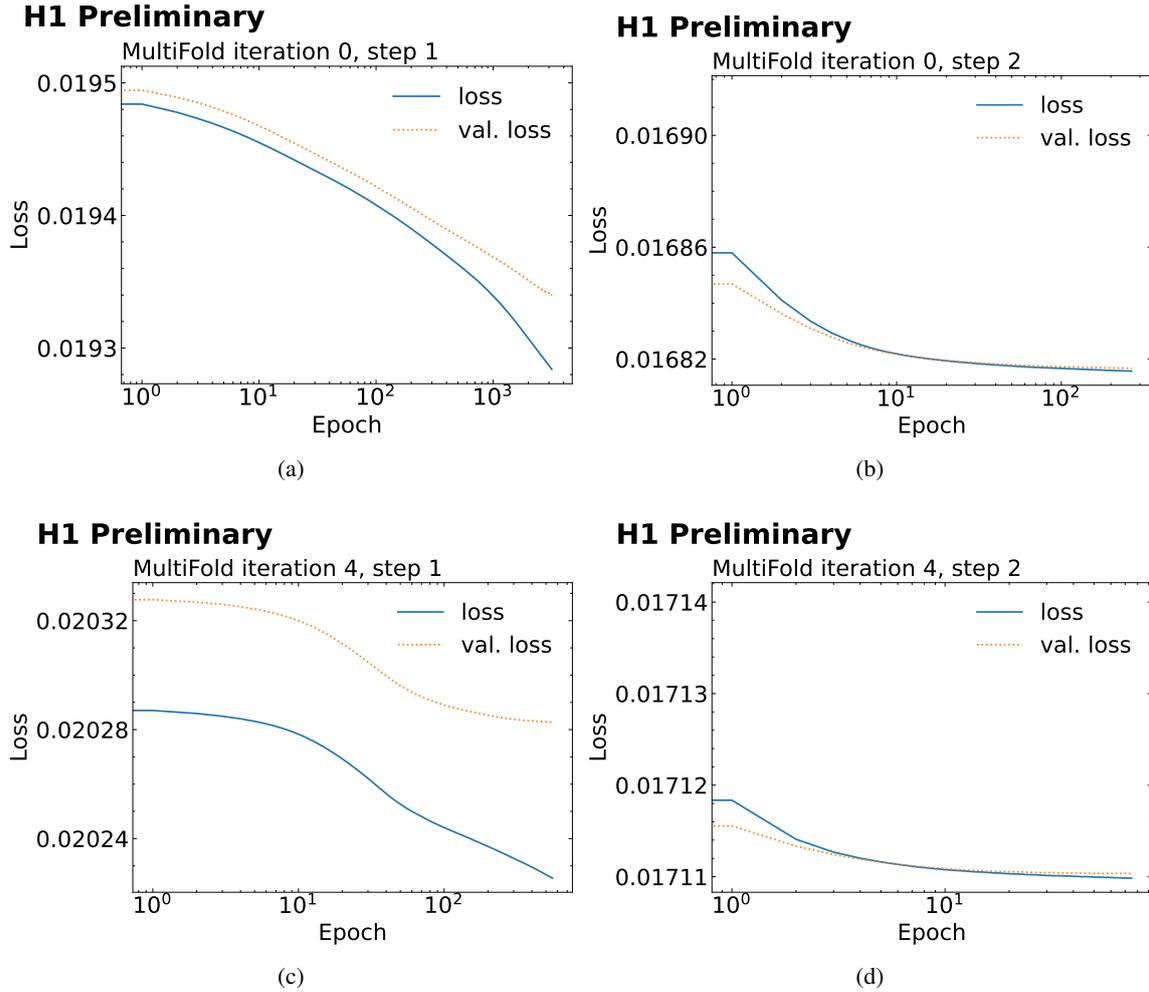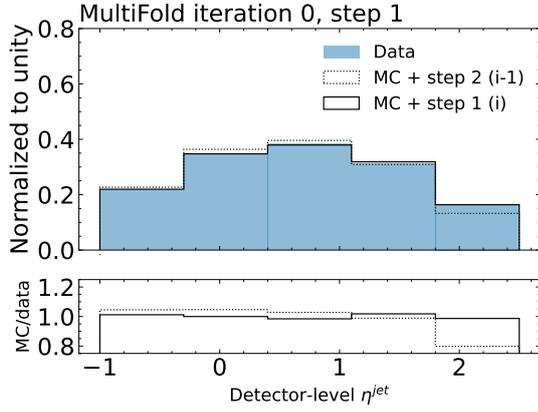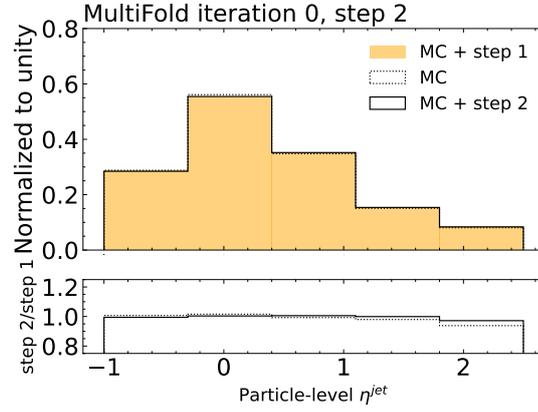
H1 Preliminary

MultiFold iteration 0, step 1

H1 Preliminary

MultiFold iteration 0, step 2

H1 Preliminary

MultiFold iteration 4, step 1

H1 Preliminary

MultiFold iteration 4, step 2

Fig. 3: Training and validation loss as a function of training epoch for the first and last unfolding iterations. (a) and (b) are for the first iteration of MULTIFOLD while (c) and (d) represent the last unfolding iteration. (a) and (c) show the loss for Step 1 and (b) and (d) show the loss for Step 2. The training is allowed to proceed for up to 10,000 epochs, but is truncated early if the validation loss does not decrease for ten consecutive epochs.

(a)

(b)

(c)

(d)
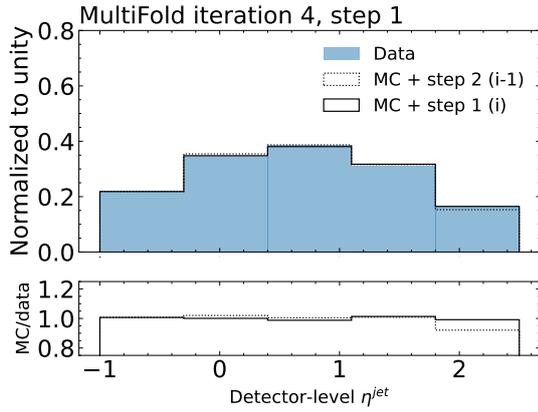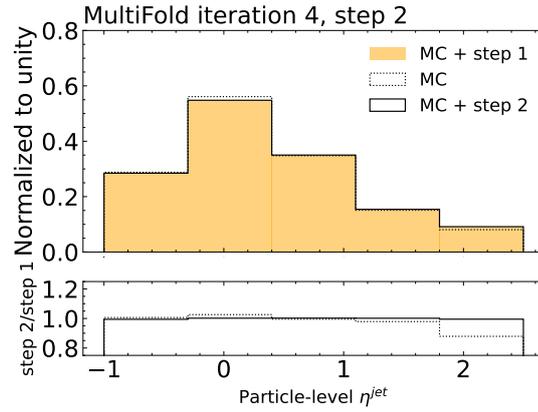
Fig. 4: The binned η^{jet} distribution after the same four iterations and steps as in Fig. 3. In particular, (a) and (b) are for the first iteration of MULTIFOLD while (c) and (d) represent the last unfolding iteration. (a) and (c) show the distribution for Step 1 and (b) and (d) show the distributions for Step 2. The filled histograms in (a) and (c) represent the observed data while the filled histograms in (b) and (d) represent the particle-level MC with weights $\lambda_i w_i$. Solid unfilled histograms correspond to the starting MC for the given iteration and step. For Step 1, these initial weights are $\nu_i w_i$, for ν_i from the previous step; for Step 2, the initial weights are simply w_i. Dashed unfilled histograms correspond to the final weighted MC for the given iteration and step. For Step 1, these final weights are $\lambda_i w_i$; for Step 2, the final weights are $\nu_i w_i$. Ratio panels show the ratio with respect to the target for the given step. In Step 1, the goal is to match the data while in Step 2, the goal is to match the results from Step 1.

6   Uncertainties

Statistical uncertainties are determined using a bootstrapping technique [47]. In particular, the entire unfolding procedure is repeated 100 times. For each time, data events are given a weight that is distributed as a Poission random variable with mean 1. This is analogous to creating 100 pseudo datasets by resampling the original dataset with replacement. The number of MC events exceeds the number of data events by nearly two orders of magnitude and therefore the MC statistical uncertainty is negligible compared to the corresponding data uncertainty. Due to the ensembling procedure described in Sec. 5, variations from the random nature of the network initialization and training is negligible compared to the data statistical uncertainty.

Systematic uncertainties are determined by varying an aspect of the simulation and repeating the unfolding. The variations used in this measurement follow other recent H1 analyses [21, 22] and are briefly summarized in the following.

The HFS-object energy scale uncertainty is decomposed into two parts, one from HFS objects contained in high p_T jets and one from all other HFS objects. In both cases, the energy-scale uncertainty is $\pm 1\%$ [21, 48]. Both uncertainties are estimated separately by varying the corresponding HFS energy by $\pm 1\%$. The uncertainty of the measurement of the azimuthal angle of the HFS objects is ± 20 mrad. The uncertainty of the measurement of the energy of the scattered lepton ranges from $\pm 0.5\%$ at backward and central regions [49] to $\pm 1\%$ at forward regions [21].

The uncertainty of the measurement of the azimuthal angle of the scattered lepton is ± 1 mrad [20]. The uncertainty associated with the modeling of the hadronic final state in the event generator used for unfolding and acceptance corrections is estimated by the difference between the results obtained using DJANGOH and RAPGAP. Given that the differential cross sections are reported normalized to the inclusive jet cross section, normalization uncertainties such as luminosity scale or trigger efficiency cancel in the ratio.

The bias of the unfolding procedure is determined by taking the difference in the result when unfolding with RAPGAP and with DJANGOH. This procedure gives a consistent result to unfolding detector-level RAPGAP with DJANGOH (and vice versa). We found that unfolding RAPGAP with itself using statistically independent samples is unbiased within MC statistical uncertainties. The RAPGAP and DJANGOH distributions bracket the data and have rather different underlying models. Therefore, comparing the results with both generators provides a realistic evaluation of the procedure bias.

The statistical uncertainty is subdominant in all bins except the highest Q^2 bin where it reaches 20%. The lepton energy scale is the largest experimental uncertainty, which is comparable to the statistical uncertainty at high Q^2. The HFS in jet scale and model uncertainties are about 1% below $Q^2 = 10^4$ GeV^2. There is no region of the y range that is as statistically limited as for the Q^2 spectrum. The model and lepton energy scale uncertainties reach about 2% and are larger than the other uncertainties in all bins except for the highest y values.

For p_T^{jet}, the HFS in jet scale is mostly the largest experimental uncertainty, although the lepton energy scale also plays a role at lower p_T^{jet} and moderate/high Q^2. The model uncertainty is largest in the highest p_T^{jet} bins. The model and lepton energy scale uncertainties are most important for η^{jet} and $\Delta\phi^{\text{jet}}$, which are mostly between 2-5%. The model uncertainty is most important for q_T^{jet}/Q, where it is mostly less than about 5% except in the highest bin.

7   Theory Predictions

The unfolded data are compared with Parton Shower MC, fixed-order calculations within perturbative QCD (pQCD), and calculations within the TMD factorization framework.

In addition to RAPGAP and DJANGOH MC predictions described in Sec. 3, the data are also compared with predictions from PYTHIA 8.3 [50, 51] and Cascade [6]. Unless stated otherwise, the default parameters for each simulation are used to generate events. The PYTHIA 8.3 predictions are produced with three flavors. The first one (default) uses a p_T-ordered parton pweron shower, the Lund string hadronization, and the NNPDF 3.1 PDF set [52]. Another PYTHIA variation uses instead the VINCIA parton shower [53], with p_T-ordered $2 \rightarrow 3$ branchings. The third PYTHIA flavor uses the DIRE parton shower [54], which implements a p_T-ordered parton shower based on dipoles. CASCADE is a TMD-based MC program that uses matrix elements from KATIE [55] and parton branching TMD PDFs [56–58]. CASCADE is also produced with two flavors. The first setup uses the HERAPDF2.0 PDF set [59] and the second variation uses angular ordering and p_T as the renormalization scale [60, 61].

Perturbative QCD calculations at next-to-next-to-leading order (NNLO) accuracy in QCD are obtained with the POLDIS code [62, 63], which is based on the Projection to Born Method [64]. These calculations are multiplied by hadronization corrections obtained with the default PYTHIA 8.3 setup. Uncertainties on the hadronization corrections are estimated by taking the envelope of the corrections obtained with the default, VINCIA and DIRE simulations. These corrections are smaller than 10% for most kinematic intervals. The uncertainty of the calculations is given by the variation the factorization and renormalization scale Q^2 by a factor of two [62, 63] as well as NLOPDF4LHC15 variations [65].

The TMD calculation employs the framework developed in Refs. [3, 66] using the same jet radius and algorithm used in this work. This differs from the original Refs. [3] paper used the anti-k_T algorithm. The difference is power suppressed at the accuracy of the calculation. The inputs to the TMD calculation are are TMD PDFs and soft functions derived in Ref. [67], which were extracted from an analysis of SIDIS and Drell-Yan data. The calculation is performed at the next-to-leading logarithmic accuracy. This calculation is performed within TMD factorization and no matching to the high q_T region is included, where the TMD approach is expected to be inaccurate. In contrast to pQCD calculations, the TMD calculations do not require non-perturbative corrections, because such effects are already included. Uncertainties are not yet available for the TMD predictions.

8   Results

The unfolded inclusive spectra of Q^2 and y are shown in Fig. 5. Both Q^2 and y exhibit steeply falling distributions that are well-described by RAPGGAP and DJANGOH. A slope in the ratio of the y distribution for DJANGOH at the 10% per unit rigidity level is visible in the ratio plot of Fig. 5(b). The statistical uncertainty is as large or larger than the systematic uncertainties in the last two Q^2 bins and is smaller than the systematic uncertainties below $Q^2 \approx 10^4$ GeV^2. The last bin of the y distribution inside the fiducial volume has a sufficient number of events so that the statistical uncertainty is not large.

Distributions of p_T^{jet}, η^{jet}, q_T^{jet}/Q, and $\Delta\phi^{\text{jet}}$ in bins of Q^2 and y are presented in Figs. 6-8 and compared with various predictions. The jet p_T, q_T/Q, and $\Delta\phi$ spectra are all steeply falling while the jet η peaks near 0 and is asymmetric due to the asymmetry of the colliding beams. As in the inclusive measurement from Ref. [1], no one prediction describes the data everywhere, although a number of calculations are able to model the data well over a wide kinematic range. The disagreements observed in inclusively in Ref. [1] become more pronounced as we examine the phase space differentially. In Figs. 9-11 the data are confronted with three variants of the PYTHIA predictions.

Direct comparisons between Q^2 and y bins are enabled by overlaying the unfolded spectra in Figs. 12 and 13, respectively.
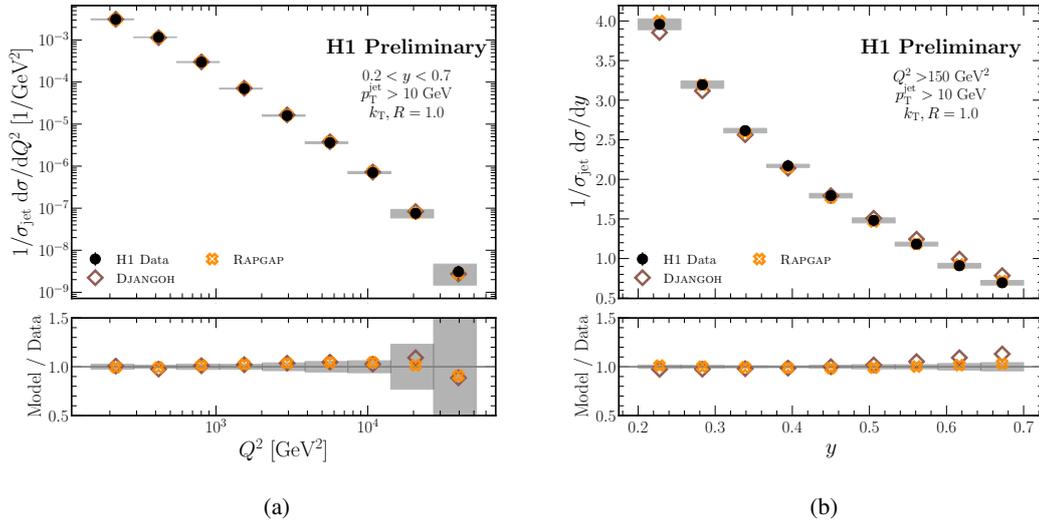
Fig. 5: The unfolded (a) Q^2 and (b) y distributions, compared with various predictions.
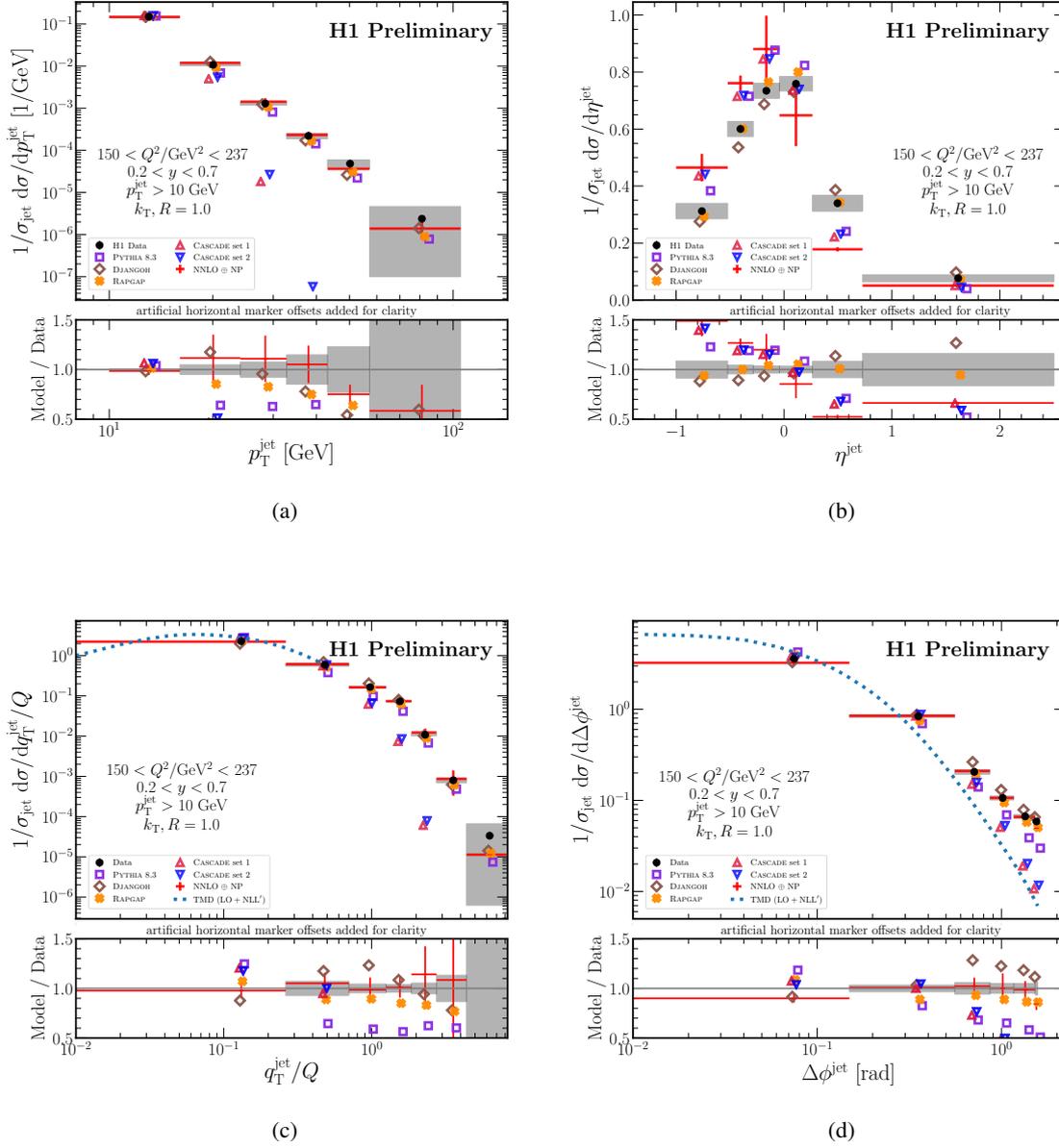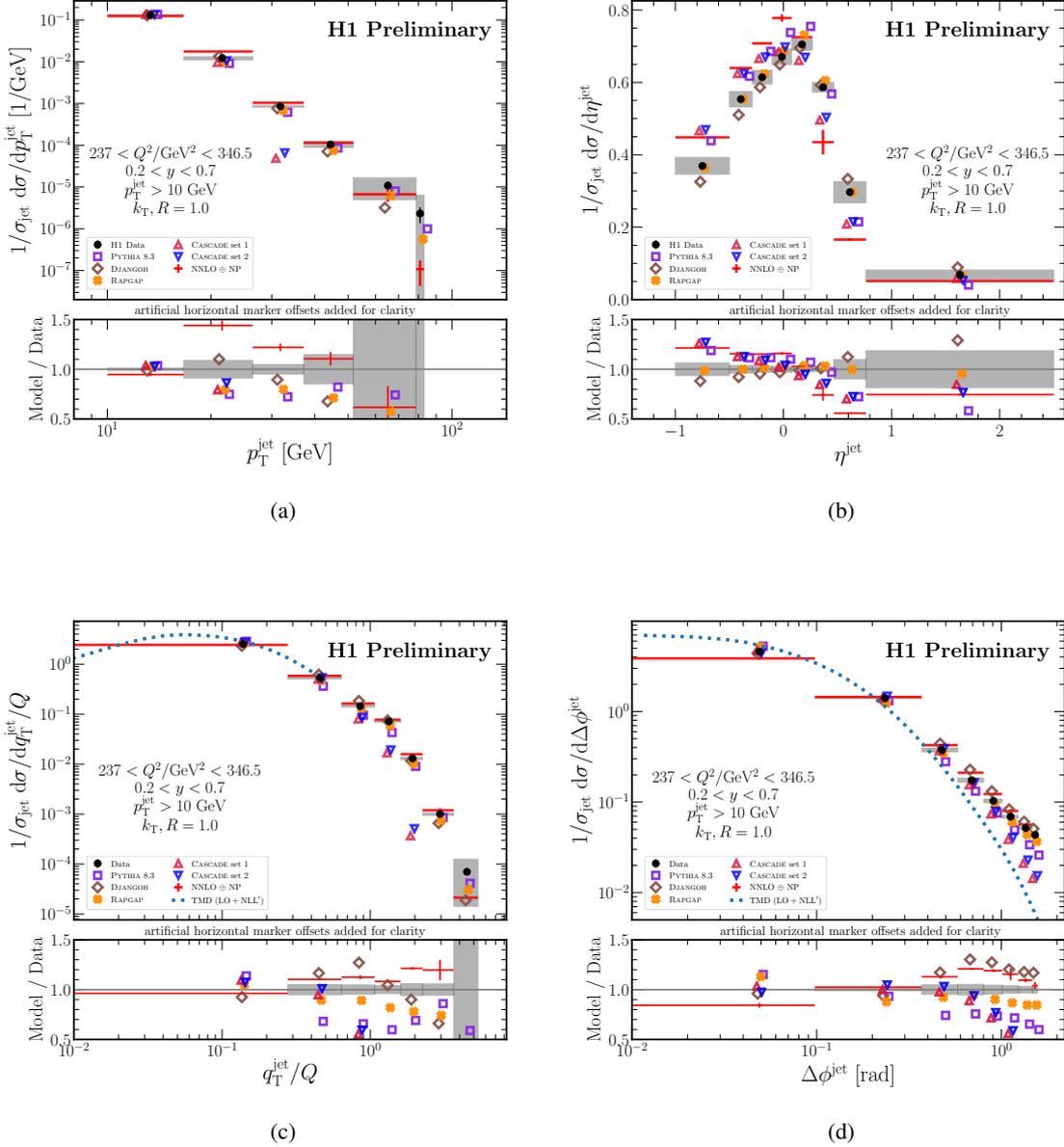
(a)

(b)

(c)

(d)

Fig. 6: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance (q_T^{jet}/Q), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{jet}$) for $150 < Q^2/\text{GeV}^2 < 237$. At the bottom, the ratio between predictions and the data are shown. The gray bands represent the total systematic uncertainty of the data; the bars represent the statistical uncertainty of the data, which is typically smaller than the marker size.
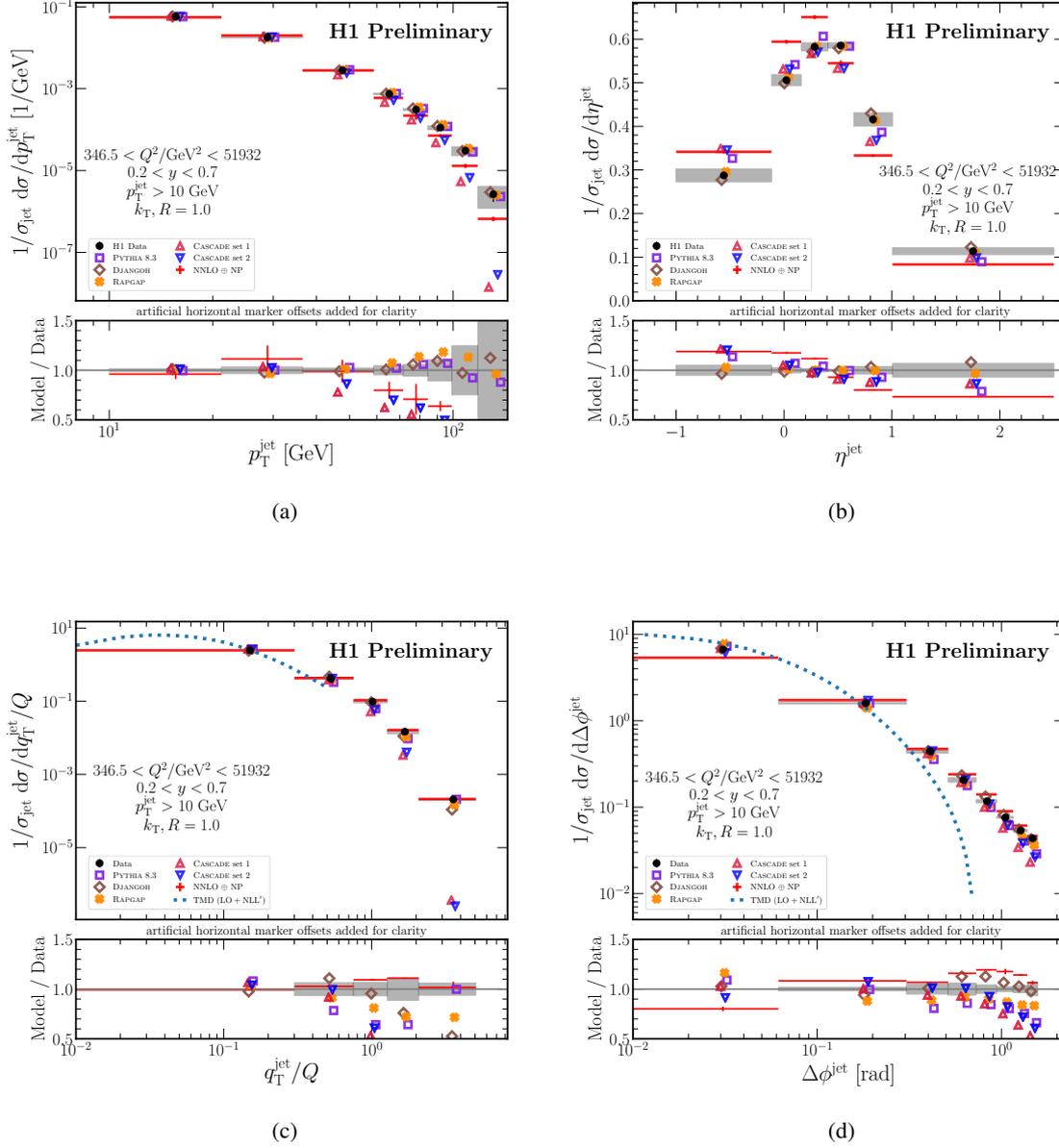
Fig. 7: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance (q_T^{jet}/Q), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{jet}$) for $237 < Q^2/\text{GeV}^2 < 346.5$. At the bottom, the ratio between predictions and the data are shown. The gray bands represent the total systematic uncertainty of the data; the bars represent the statistical uncertainty of the data, which is typically smaller than the marker size.
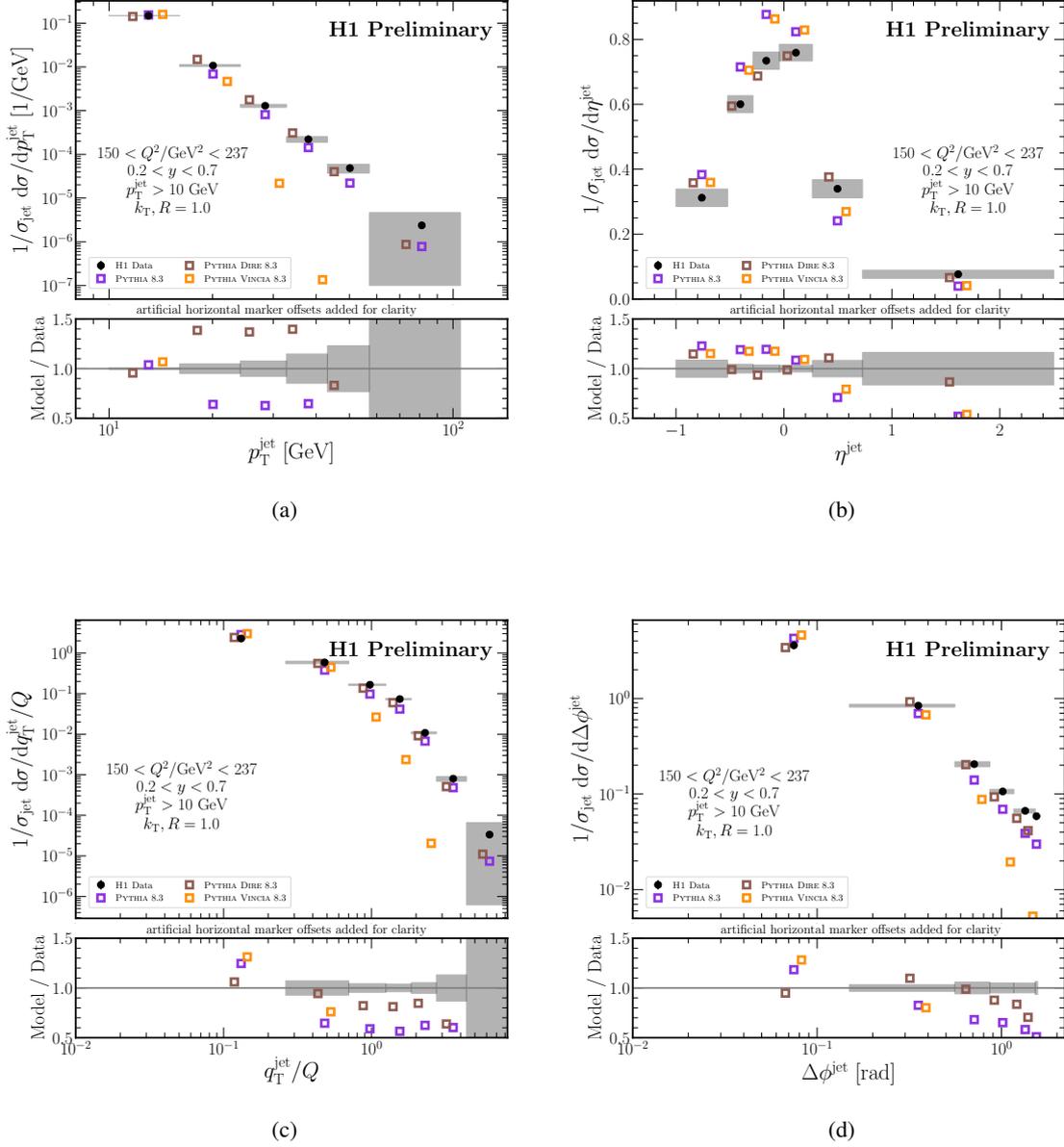
(a)

(b)

(c)

(d)

Fig. 8: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance (q_T^{jet}/Q), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{jet}$) for $346.5 < Q^2/\text{GeV}^2 < 51932$. At the bottom, the ratio between predictions and the data are shown. The gray bands represent the total systematic uncertainty of the data; the bars represent the statistical uncertainty of the data, which is typically smaller than the marker size.
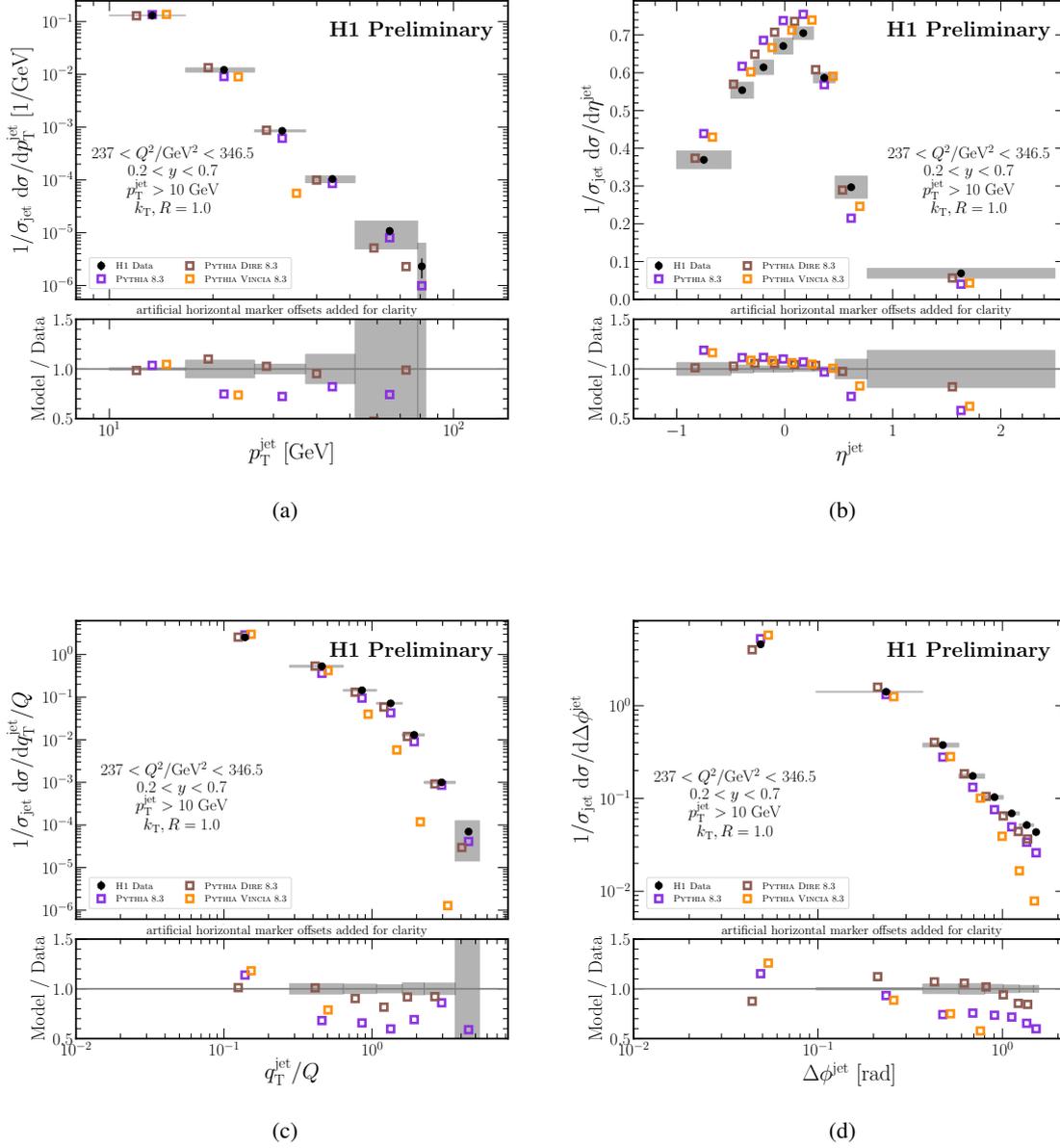
Fig. 9: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance ($q_{\mathrm{T}}^{\mathrm{jet}}/Q$), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{\mathrm{jet}}$) for $150 < Q^2/\mathrm{GeV}^2 < 237$. At the bottom, the ratio between predictions and the data are shown. The gray bands represent the total systematic uncertainty of the data; the bars represent the statistical uncertainty of the data, which is typically smaller than the marker size.
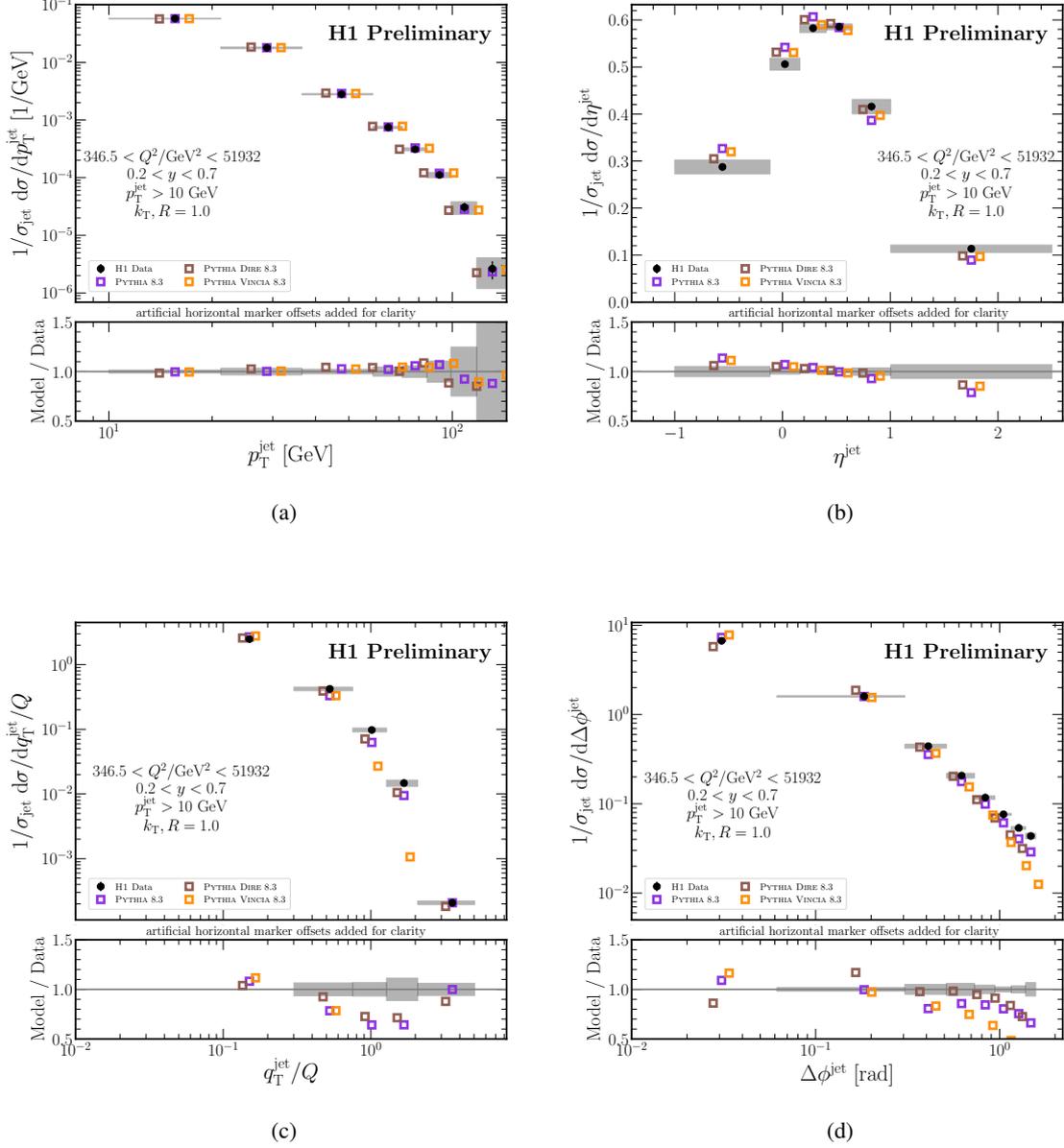
Fig. 10: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance (q_T^{jet}/Q), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{jet}$) for $237 < Q^2/\text{GeV}^2 < 346.5$. At the bottom, the ratio between predictions and the data are shown. The gray bands represent the total systematic uncertainty of the data; the bars represent the statistical uncertainty of the data, which is typically smaller than the marker size.

Fig. 11: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance (q_T^{jet}/Q), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{jet}$) for $346.5 < Q^2/\text{GeV}^2 < 51932$. At the bottom, the ratio between predictions and the data are shown. The gray bands represent the total systematic uncertainty of the data; the bars represent the statistical uncertainty of the data, which is typically smaller than the marker size.
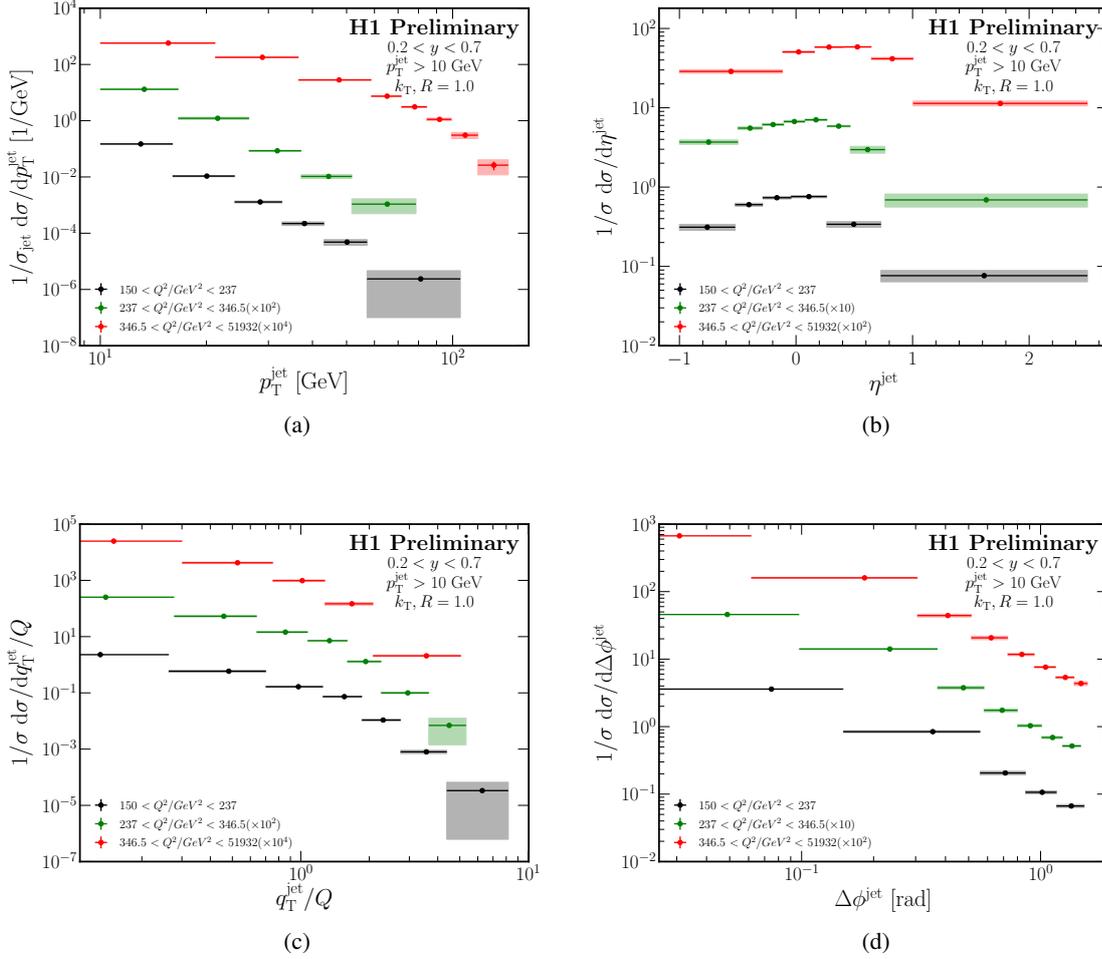
Fig. 12: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance ($q_{\mathrm{T}}^{\mathrm{jet}}/Q$), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{\mathrm{jet}}$) in three bins of Q^2. Data in higher Q^2 regions are vertically offset in the plots to improve clarity. Only unfolded data is presented, while the systematic uncertainty is denoted by the corresponding color-shaded box. The statistical uncertainty is given by the vertical colored bars attached to the data point. However, it is not visible in most plots since it is smaller than the size of the data marker.
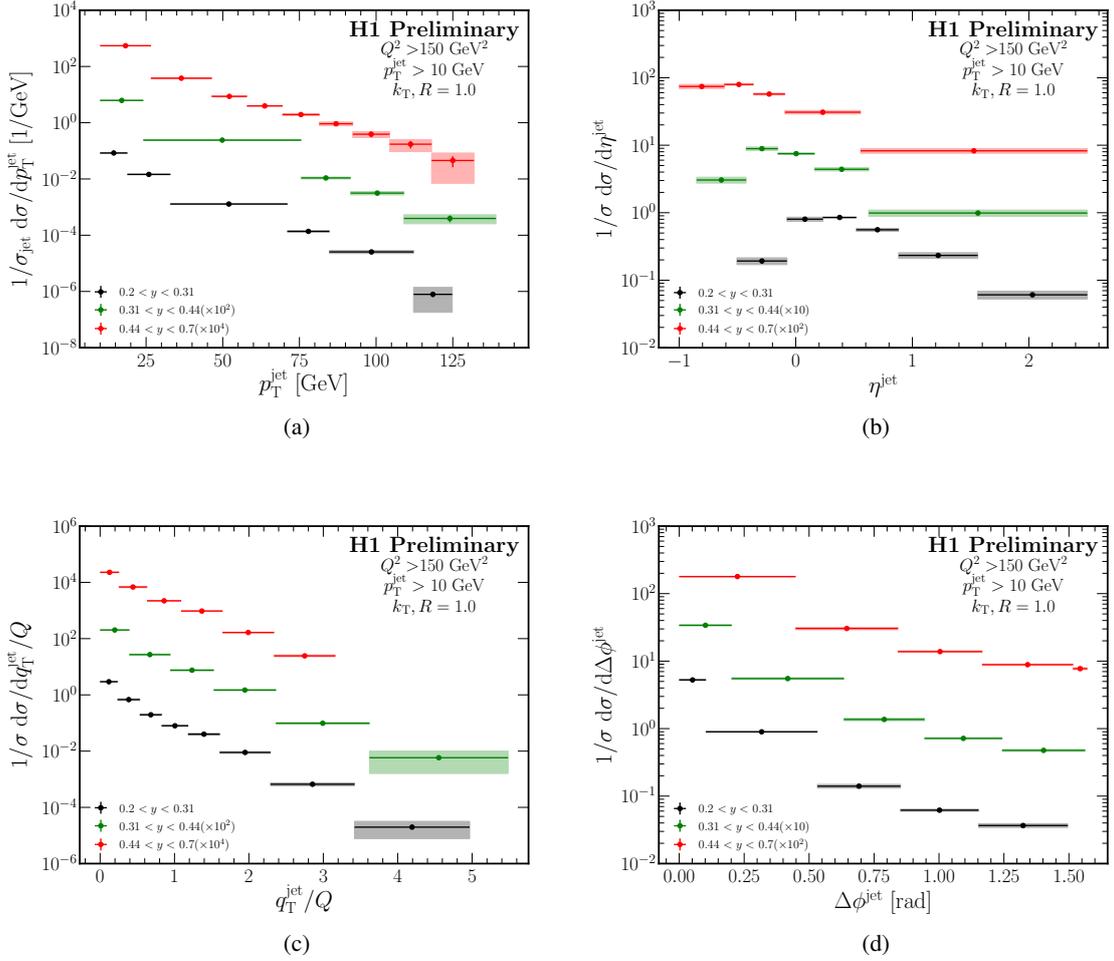
Fig. 13: Measured cross sections, normalized to the inclusive jet production cross section, as a function of (a) the jet transverse momentum, (b) the jet pseudorapidity, (c) the lepton-jet momentum balance (q_T^{jet}/Q), and (d) the lepton-jet azimuthal angle correlation ($\Delta\phi^{jet}$) in three bins of y. Data in higher y regions are vertically offset in the plots to improve clarity. Only unfolded data is presented, while the systematic uncertainty is denoted by the corresponding color-shaded box. The statistical uncertainty is given by the vertical colored bars attached to the data point. However, it is not visible in most plots since it is smaller than the size of the data marker.

9   Conclusions

This paper has reported multi-differential cross section measurements of the laboratory frame jet-lepton imbalance in neutral current DIS events collected with the H1 detector at HERA. Using the same machine learning-based unfolding as in Ref. [1], we show the differential nature of the result by presenting the four observables from Ref. [1] in bins of Q^2 and y. While Q^2 and y were not part of the original phase space, they can be derived from the measurement. As the result is unbinned, there are no artifacts from the indirect Q^2 and y measurements. We compare the data to a number of predictions and while we find that no one calculation can describe the data everywhere, a number of predictions are able to accurately model the data across a wide kinematic range. These data will be provide an important input to connecting perturbative QCD and TMD calculations in a unified framework and they also represent an important methodological step towards publishing unbinned differential cross section measurements [68].

Bibliography

[1] H1 Collaboration. Measurement of lepton-jet correlation in deep-inelastic scattering with the H1 detector using machine learning for unfolding. 2021.

[2] Paul Newman and Matthew Wing. The hadronic final state at HERA. *Rev. Mod. Phys.*, 86(3):1037, 2014.

[3] Xiaohui Liu, Felix Ringer, Werner Vogelsang, and Feng Yuan. Lepton-jet correlations in deep inelastic scattering at the Electron-Ion Collider. *Phys. Rev. Lett.*, 122(19):192003, 2019.

[4] Daniel Gutierrez-Reyes, Ignazio Scimemi, Wouter J. Waalewijn, and Lorenzo Zoppi. Transverse momentum dependent distributions with jets. *Phys. Rev. Lett.*, 121(16):162001, 2018.

[5] J. Collins, L. Gamberg, A. Prokudin, T. C. Rogers, N. Sato, and B. Wang. Relating Transverse Momentum Dependent and Collinear Factorization Theorems in a Generalized Formalism. *Phys. Rev. D*, 94(3):034014, 2016.

[6] S. Baranov et al. CASCADE3 A Monte Carlo event generator based on TMDs. *Eur. Phys. J. C*, 81(5):425, 2021.

[7] Anders Andreassen, Patrick T. Komiske, Eric M. Metodiev, Benjamin Nachman, and Jesse Thaler. OmniFold: A Method to Simultaneously Unfold All Observables. *Phys. Rev. Lett.*, 124(18):182001, 2020.

[8] Anders Andreassen, Patrick T. Komiske, Eric M. Metodiev, Benjamin Nachman, Adi Suresh, and Jesse Thaler. Scaffolding Simulations witiih Deep Learning for High-Dimensional Deconvolution. *ICLR SimDL Workshop*, page 12, 2021.

[9] I. Abt et al. The H1 detector at HERA. *DESY-93-103*, 7 1993.

[10] B. Andrieu et al. The H1 liquid argon calorimeter system. *Nucl. Instrum. Meth. A*, 336:460–498, 1993.

[11] I. Abt et al. The H1 detector at HERA. *Nucl. Instrum. Meth. A*, 386:310–347, 1997.

[12] I. Abt et al. The Tracking, calorimeter and muon detectors of the H1 experiment at HERA. *Nucl. Instrum. Meth. A*, 386:348–396, 1997.

[13] R. D. Appuhn et al. The H1 lead / scintillating fiber calorimeter. *Nucl. Instrum. Meth. A*, 386:397–408, 1997.

[14] D. Pitzl et al. The H1 silicon vertex detector. *Nucl. Instrum. Meth. A*, 454:334–349, 2000.

[15] B. Andrieu et al. Beam tests and calibration of the H1 liquid argon calorimeter with electrons. *Nucl. Instrum. Meth. A*, 350:57–72, 1994.

[16] B. Andrieu et al. Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations. *Nucl. Instrum. Meth. A*, 336:499–509, 1993.

[17] R.-D. Appuhn, C. Arndt, E. Barrelet, R. Barschke, U. Bassler, V. Boudry, R. Buchholz, F. Brasse, D. Bruncko, S. Chechelnitski, B. Claxton, G. Cozzika, J. Cvach, S. Dagoret-Campagne, W.D. Dau, H. Deckers, T. Deckers, F. Descamps, M. Dirkmann, J. Dowdell, V. Efremenko, E. Eisenhandler, A.N. Eliseev, G. Falley, J. Ferencei, B. Fominykh, K. Gadow, U. Goerlach, L.A. Gorbov, I. Gorelov, M. Grewe, L. Hajduk, I. Herynek, J. Hladký, M. Hütte, H. Hutter, W. Janczur, J. Janoth, L. Jönsson, H. Kolanoski, V. Korbel, F. Kriváň, D. Lacour, F. Lamarche, M.P.J. Landon, B. Laforge, J.-F. Laporte, F. Lehner, R. Maraček, K. Meier, A. Meyer, A. Migliori, F. Moreau, G. Müller, P. Murín, V. Nagovizin, T.C. Nicholls, D. Ozerov, E. Perez, J.P. Pharabod, R. Pöschl, K. Rybicki, A. Rostovtsev, C. Royon, S. Schleif, A. Schuhmacher, A. Semenov, V. Shekelyan, Y. Sirois, P.A. Smirnov, V. Solochenko, J. Špalek, S. Spielmann, H. Steiner, J. Stiewe, M. Taševský, V. Tchernyshov, K. Thiele, E. Tzamariudaki, S. Valkár, D. VanDenPlas, G. Villet, K. Wacker, A. Walther, M. Weber, D. Wegener, T. Wenk, J. Žáček, A. Zhokin, and K. Zuber. The h1 lead/scintillating-fibre calorimeter. *Nucl. Instrum. Meth. A*, 386(2):397–408, 1997.

[18] F. D. Aaron et al. Determination of the Integrated Luminosity at HERA using Elastic QED Compton Events. *Eur. Phys. J. C*, 72:2163, 2012. [Erratum: Eur. Phys. J. C **74**, 2733 (2012)].

[19] C. Adloff et al. Measurement and QCD analysis of neutral and charged current cross-sections at HERA. *Eur. Phys. J. C*, 30:1–32, 2003.

[20] F. D. Aaron et al. Inclusive Deep Inelastic Scattering at High Q^2 with Longitudinally Polarised Lepton Beams at HERA. *JHEP*, 09:061, 2012.

[21] 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(2):65, 2015.

[22] Vladimir Andreev et al. Measurement of Jet Production Cross Sections in Deep-inelastic ep Scattering at HERA. *Eur. Phys. J. C*, 77(4):215, 2017.

[23] K. Charchula, G. A. Schuler, and H. Spiesberger. Combined QED and QCD radiative effects in deep inelastic lepton - proton scattering: The Monte Carlo generator DJANGO6. *Comput. Phys. Commun.*, 81:381–402, 1994.

[24] Hannes Jung. Hard diffractive scattering in high-energy e p collisions and the Monte Carlo generator RAPGAP. *Comput. Phys. Commun.*, 86:147–161, 1995.

[25] H Spiesberger, A A Akhundov, H Anlauf, Dimitri Yuri Bardin, Johannes Blümlein, H D Dahmen, Stanislaw Jadach, L V Kalinovskaya, A Kwiatkowski, P Manakos, T Mannel, H J Möhring, G Montagna, O Nicrosini, T Ohl, Wiesiek Placzek, Tord Riemann, and L Viola. Radiative corrections at

HERA. *CERN-TH-6447-92*, page 43, Mar 1992.

[26] A Kwiatkowski, H. Spiesberger, and H. J. Mohring. Characteristics of radiative events in deep inelastic e p scattering at HERA. *Z. Phys. C*, 50:165–178, 1991.

[27] A. Kwiatkowski, H. Spiesberger, and H. J. Mohring. Heracles: An Event Generator for *ep* Interactions at HERA Energies Including Radiative Processes: Version 1.0. *Comput. Phys. Commun.*, 69:155–172, 1992.

[28] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, Pavel M. Nadolsky, and W. K. Tung. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

[29] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Phys. Rept.*, 97:31–145, 1983.

[30] S. Schael et al. Bose-Einstein correlations in W-pair decays with an event-mixing technique. *Phys. Lett. B*, 606:265–275, 2005.

[31] L. Lönnblad. ARIADNE version 4: A Program for simulation of QCD cascades implementing the color dipole model. *Comput. Phys. Commun.*, 71:15–31, 1992.

[32] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini. GEANT3. *CERN-DD-EE-84-01*, 9 1987.

[33] Ursula Bassler and Gregorio Bernardi. On the kinematic reconstruction of deep inelastic scattering at HERA: The Sigma method. *Nucl. Instrum. Meth. A*, 361:197–208, 1995.

[34] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J. C*, 72:1896, 2012.

[35] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett. B*, 641:57–61, 2006.

[36] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys. B*, 406:187–224, 1993.

[37] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev. D*, 48:3160–3166, 1993.

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

[39] S Hellwig. Untersuchung der D^* - π slow Double Tagging Methode in Charmanalysen. (Diploma, Univ. Hamburg), 2005.

[40] 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), 2005.

[41] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

[42] Masashi Sugiyama, Taiji Suzuki, and Takafumi Kanamori. *Density Ratio Estimation in Machine Learning*. Cambridge University Press, 2012.

[43] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. *OSDI*, 16:265, 2016.

[44] Fraciois Chollet. Keras. https://github.com/fchollet/keras, 2017.

[45] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *ICLR*, 3, 2015.

[46] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In Yee Whye Teh and D. Mike Titterington, editors, *AISTATS*, volume 9 of *JMLR Proceedings*, pages 249–256. JMLR.org, 2010.

[47] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1 – 26, 1979.

[48] Roman Kogler. Measurement of jet production in deep-inelastic *ep* scattering at HERA. (Thesis,

Univ. Hamburg), 2011.

[49] F. D. Aaron et al. Measurement of $D^{*\pm}$ Meson Production and Determination of F_2^c at low Q^2 in Deep-Inelastic Scattering at HERA. *Eur. Phys. J. C*, 71:1769, 2011. [Erratum: Eur. Phys. J. C **72**, 2252 (2012)].

[50] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.

[51] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.

[52] Richard D. Ball et al. Parton distributions from high-precision collider data. *Eur. Phys. J. C*, 77(10):663, 2017.

[53] Nadine Fischer, Stefan Prestel, Mathias Ritzmann, and Peter Skands. Vincia for Hadron Colliders. *Eur. Phys. J. C*, 76(11):589, 2016.

[54] Stefan Höche and Stefan Prestel. The midpoint between dipole and parton showers. *Eur. Phys. J. C*, 75(9):461, 2015.

[55] A. van Hameren. KaTie : For parton-level event generation with k_T-dependent initial states. *Comput. Phys. Commun.*, 224:371–380, 2018.

[56] F. Hautmann, H. Jung, A. Lelek, V. Radescu, and R. Zlebcik. Soft-gluon resolution scale in QCD evolution equations. *Phys. Lett. B*, 772:446–451, 2017.

[57] F. Hautmann, H. Jung, A. Lelek, V. Radescu, and R. Zlebcik. Collinear and TMD Quark and Gluon Densities from Parton Branching Solution of QCD Evolution Equations. *JHEP*, 01:070, 2018.

[58] A. Bermudez Martinez, P. Connor, H. Jung, A. Lelek, R. Žlebčík, F. Hautmann, and V. Radescu. Collinear and TMD parton densities from fits to precision DIS measurements in the parton branching method. *Phys. Rev. D*, 99(7):074008, 2019.

[59] 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(12):580, 2015.

[60] A. Bermudez Martinez et al. Production of Z-bosons in the parton branching method. *Phys. Rev. D*, 100(7):074027, 2019.

[61] A. Bermudez Martinez et al. The transverse momentum spectrum of low mass Drell–Yan production at next-to-leading order in the parton branching method. *Eur. Phys. J. C*, 80(7):598, 2020.

[62] Ignacio Borsa, Daniel de Florian, and Iván Pedron. Jet production in polarized deep inelastic scattering at next-to-next-to-leading order. *Phys. Rev. Lett.*, 125(8):082001, 2020.

[63] Ignacio Borsa, Daniel de Florian, and Iván Pedron. Inclusive-jet and dijet production in polarized deep inelastic scattering. *Phys. Rev. D*, 103(1):014008, 2021.

[64] Matteo Cacciari, Frédéric A. Dreyer, Alexander Karlberg, Gavin P. Salam, and Giulia Zanderighi. Fully Differential Vector-Boson-Fusion Higgs Production at Next-to-Next-to-Leading Order. *Phys. Rev. Lett.*, 115(8):082002, 2015. [Erratum: Phys. Rev. Lett. **120**, 139901 (2018)].

[65] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*, 43:023001, 2016.

[66] Xiaohui Liu, Felix Ringer, Werner Vogelsang, and Feng Yuan. Lepton-jet Correlation in Deep Inelastic Scattering. *Phys. Rev. D*, 102(9):094022, 2020.

[67] Peng Sun, Joshua Isaacson, C. P. Yuan, and Feng Yuan. Nonperturbative functions for SIDIS and Drell–Yan processes. *Int. J. Mod. Phys. A*, 33(11):1841006, 2018.

[68] Miguel Arratia et al. Presenting Unbinned Differential Cross Section Results. 9 2021.