



University of  
Zurich<sup>UZH</sup>

Zurich Open Repository and  
Archive

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2023

---

## Unbinned deep learning jet substructure measurement in high $Q^2$ ep collisions at HERA

Andreev, V ; Arratia, M ; Baghdasaryan, A ; Baty, A ; Begzsuren, K ; Bolz, A ; Boudry, V ; Brandt, G ; Britzger, D ; Buniatyan, A ; Bystritskaya, L ; Campbell, A J ; Cantun Avila, K B ; Cerny, K ; Chekelian, V ; Chen, Z ; Contreras, J G ; Cvach, J ; Dainton, J B ; Daum, K ; Deshpande, A ; Diaconu, C ; Drees, A ; Eckerlin, G ; Egli, S ; Elsen, E ; Favart, L ; Fedotov, A ; Feltesse, J ; Fleischer, M ; et al ; Robmann, Peter ; Müller, Katharina ; Straumann, Ulrich

DOI: <https://doi.org/10.1016/j.physletb.2023.138101>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-240114>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Andreev, V ; Arratia, M ; Baghdasaryan, A ; Baty, A ; Begzsuren, K ; Bolz, A ; Boudry, V ; Brandt, G ; Britzger, D ; Buniatyan, A ; Bystritskaya, L ; Campbell, A J ; Cantun Avila, K B ; Cerny, K ; Chekelian, V ; Chen, Z ; Contreras, J G ; Cvach, J ; Dainton, J B ; Daum, K ; Deshpande, A ; Diaconu, C ; Drees, A ; Eckerlin, G ; Egli, S ; Elsen, E ; Favart, L ; Fedotov, A ; Feltesse, J ; Fleischer, M ; et al ; Robmann, Peter ; Müller, Katharina ; Straumann, Ulrich (2023). Unbinned deep learning jet substructure measurement in high  $Q^2$  ep collisions at HERA. *Physics Letters B*, 844:138101.

DOI: <https://doi.org/10.1016/j.physletb.2023.138101>



ELSEVIER

Contents lists available at ScienceDirect

## Physics Letters B

journal homepage: [www.elsevier.com/locate/physletb](http://www.elsevier.com/locate/physletb)

# Unbinned deep learning jet substructure measurement in high $Q^2$ $ep$ collisions at HERA

V. Andreev<sup>ar</sup>, M. Arratia<sup>ac</sup>, A. Baghdasaryan<sup>an</sup>, A. Baty<sup>p</sup>, K. Begzsuren<sup>ah</sup>, A. Bolz<sup>n</sup>, V. Boudry<sup>y</sup>, G. Brandt<sup>m</sup>, D. Britzger<sup>v</sup>, A. Buniatyan<sup>g</sup>, L. Bystritskaya<sup>ar</sup>, A.J. Campbell<sup>n</sup>, K.B. Cantun Avila<sup>ao</sup>, K. Cerny<sup>w</sup>, V. Chekelian<sup>v</sup>, Z. Chen<sup>ae</sup>, J.G. Contreras<sup>ao</sup>, J. Cvach<sup>aa</sup>, J.B. Dainton<sup>s</sup>, K. Daum<sup>am</sup>, A. Deshpande<sup>ag,aj</sup>, C. Diaconu<sup>u</sup>, A. Drees<sup>ag</sup>, G. Eckerlin<sup>n</sup>, S. Egli<sup>ak</sup>, E. Elsen<sup>n</sup>, L. Favart<sup>d</sup>, A. Fedotov<sup>ar</sup>, J. Feltesse<sup>l</sup>, M. Fleischer<sup>n</sup>, A. Fomenko<sup>ar</sup>, C. Gal<sup>ag</sup>, J. Gayler<sup>n</sup>, L. Goerlich<sup>q</sup>, N. Gogitidze<sup>n</sup>, M. Gouzevitch<sup>ar</sup>, C. Grab<sup>ap</sup>, T. Greenshaw<sup>s</sup>, G. Grindhammer<sup>v</sup>, D. Haidt<sup>n</sup>, R.C.W. Henderson<sup>r</sup>, J. Hessler<sup>v</sup>, J. Hladký<sup>aa</sup>, D. Hoffmann<sup>u</sup>, R. Horisberger<sup>ak</sup>, T. Hreus<sup>aq</sup>, F. Huber<sup>o</sup>, P.M. Jacobs<sup>e</sup>, M. Jacquet<sup>x</sup>, T. Janssen<sup>d</sup>, A.W. Jung<sup>al</sup>, J. Katzy<sup>n</sup>, C. Kiesling<sup>v</sup>, M. Klein<sup>s</sup>, C. Kleinwort<sup>n</sup>, H.T. Klest<sup>ag</sup>, R. Kogler<sup>n</sup>, P. Kostka<sup>s</sup>, J. Kretzschmar<sup>s</sup>, D. Krücker<sup>n</sup>, K. Krüger<sup>n</sup>, M.P.J. Landon<sup>t</sup>, W. Lange<sup>n</sup>, P. Laycock<sup>aj</sup>, S.H. Lee<sup>b</sup>, S. Levonian<sup>n</sup>, W. Li<sup>p</sup>, J. Lin<sup>p</sup>, K. Lipka<sup>n</sup>, B. List<sup>n</sup>, J. List<sup>n</sup>, B. Lobodzinski<sup>v</sup>, O.R. Long<sup>ac</sup>, E. Malinowski<sup>ar</sup>, H.-U. Martyn<sup>a</sup>, S.J. Maxfield<sup>s</sup>, A. Mehta<sup>s</sup>, A.B. Meyer<sup>n</sup>, J. Meyer<sup>n</sup>, S. Mikocki<sup>q</sup>, V.M. Mikuni<sup>e</sup>, M.M. Mondal<sup>ag</sup>, K. Müller<sup>aq</sup>, B. Nachman<sup>e</sup>, Th. Naumann<sup>n</sup>, P.R. Newman<sup>g</sup>, C. Niebuhr<sup>n</sup>, G. Nowak<sup>q</sup>, J.E. Olsson<sup>n</sup>, D. Ozerov<sup>ar</sup>, S. Park<sup>ag</sup>, C. Pascaud<sup>x</sup>, G.D. Patel<sup>s</sup>, E. Perez<sup>k</sup>, A. Petrukhin<sup>af</sup>, I. Picuric<sup>z</sup>, D. Pitzl<sup>n</sup>, R. Polifka<sup>ab</sup>, S. Preins<sup>ac</sup>, V. Radescu<sup>o</sup>, N. Raicevic<sup>z</sup>, T. Ravidandorj<sup>ah</sup>, P. Reimer<sup>aa</sup>, E. Rizvi<sup>t</sup>, P. Robmann<sup>aq</sup>, R. Roosen<sup>d</sup>, A. Rostovtsev<sup>ar</sup>, M. Rotaru<sup>h</sup>, D.P.C. Sankey<sup>i</sup>, M. Sauter<sup>o</sup>, E. Sauvan<sup>u,c</sup>, S. Schmitt<sup>n,\*</sup>, B.A. Schmookler<sup>ag</sup>, G. Schnell<sup>f</sup>, L. Schoeffel<sup>l</sup>, A. Schöning<sup>o</sup>, F. Sefkow<sup>n</sup>, S. Shushkevich<sup>v</sup>, Y. Soloviev<sup>n</sup>, P. Sopicki<sup>q</sup>, D. South<sup>n</sup>, A. Specka<sup>y</sup>, M. Steder<sup>n</sup>, B. Stella<sup>ad</sup>, U. Straumann<sup>aq</sup>, C. Sun<sup>ag</sup>, T. Sykora<sup>ab</sup>, P.D. Thompson<sup>g</sup>, F. Torales Acosta<sup>e</sup>, D. Traynor<sup>t</sup>, B. Tsepeldorj<sup>ah,ai</sup>, Z. Tu<sup>aj</sup>, G. Tustin<sup>ag</sup>, A. Valkárová<sup>ab</sup>, C. Vallée<sup>u</sup>, P. Van Mechelen<sup>d</sup>, D. Wegener<sup>j</sup>, E. Wünsch<sup>n</sup>, J. Žáček<sup>ab</sup>, J. Zhang<sup>ae</sup>, Z. Zhang<sup>x</sup>, R. Žlebčík<sup>ab</sup>, H. Zohrabyan<sup>an</sup>, F. Zomer<sup>x</sup>

<sup>a</sup> I. Physikalisches Institut der RWTH, Aachen, Germany

<sup>b</sup> University of Michigan, Ann Arbor, MI 48109, USA<sup>1</sup>

<sup>c</sup> LAPP, Université de Savoie, CNRS/IN2P3, Annecy-le-Vieux, France

<sup>d</sup> Inter-University Institute for High Energies ULB-VUB, Brussels and Universiteit Antwerpen, Antwerp, Belgium<sup>2</sup>

<sup>e</sup> Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA<sup>1</sup>

<sup>f</sup> Department of Physics, University of the Basque Country UPV/EHU, 48080 Bilbao, Spain

<sup>g</sup> School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom<sup>3</sup>

<sup>h</sup> Horia Hulubei National Institute for R&D in Physics and Nuclear Engineering (IFIN-HH), Bucharest, Romania<sup>4</sup>

\* Corresponding author.

E-mail address: [sschmitt@mail.desy.de](mailto:sschmitt@mail.desy.de) (S. Schmitt).

<sup>1</sup> Supported by the U.S. DOE Office of Science.

<sup>2</sup> Supported by FNRS-FWO-Vlaanderen, IISN-IIKW and IWT and by Interuniversity Attraction Poles Programme, Belgian Science Policy.

<sup>3</sup> Supported by the UK Science and Technology Facilities Council, and formerly by the UK Particle Physics and Astronomy Research Council.

<sup>4</sup> Supported by the Romanian National Authority for Scientific Research under the contract PN 09370101.

<sup>5</sup> Supported by the Bundesministerium für Bildung und Forschung, FRG, under contract numbers 05H09GUF, 05H09VHC, 05H09VHF, 05H16PEA.

<sup>6</sup> Partially supported by Polish Ministry of Science and Higher Education, grant DPN/N168/DESY/2009.

<sup>7</sup> Partially supported by Ministry of Science of Montenegro, no. 05-1/3-3352.

<sup>8</sup> Supported by the Ministry of Education of the Czech Republic under the project INGO-LG14033.

<sup>9</sup> Supported by CONACYT, México, grant 48778-F.

<sup>10</sup> Supported by the Swiss National Science Foundation.

- <sup>i</sup> STFC, Rutherford Appleton Laboratory, Didcot, Oxfordshire, United Kingdom <sup>3</sup>  
<sup>j</sup> Institut für Physik, TU Dortmund, Dortmund, Germany <sup>5</sup>  
<sup>k</sup> CERN, Geneva, Switzerland  
<sup>l</sup> IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France  
<sup>m</sup> II. Physikalisches Institut, Universität Göttingen, Göttingen, Germany  
<sup>n</sup> Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen, Germany  
<sup>o</sup> Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany <sup>5</sup>  
<sup>p</sup> Rice University, Houston, TX 77005-1827, USA  
<sup>q</sup> Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland <sup>6</sup>  
<sup>r</sup> Department of Physics, University of Lancaster, Lancaster, United Kingdom <sup>3</sup>  
<sup>s</sup> Department of Physics, University of Liverpool, Liverpool, United Kingdom <sup>3</sup>  
<sup>t</sup> School of Physics and Astronomy, Queen Mary, University of London, London, United Kingdom <sup>3</sup>  
<sup>u</sup> Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France  
<sup>v</sup> Max-Planck-Institut für Physik, München, Germany  
<sup>w</sup> Joint Laboratory of Optics, Palacký University, Olomouc, Czech Republic  
<sup>x</sup> IJCLab, Université Paris-Saclay, CNRS/IN2P3, Orsay, France  
<sup>y</sup> LLR, Ecole Polytechnique, CNRS/IN2P3, Palaiseau, France  
<sup>z</sup> Faculty of Science, University of Montenegro, Podgorica, Montenegro <sup>7</sup>  
<sup>aa</sup> Institute of Physics, Academy of Sciences of the Czech Republic, Praha, Czech Republic <sup>8</sup>  
<sup>ab</sup> Faculty of Mathematics and Physics, Charles University, Praha, Czech Republic <sup>8</sup>  
<sup>ac</sup> University of California, Riverside, CA 92521, USA  
<sup>ad</sup> Dipartimento di Fisica Università di Roma Tre and INFN Roma 3, Roma, Italy  
<sup>ae</sup> Shandong University, Shandong, PR China  
<sup>af</sup> Fakultät IV – Department für Physik, Universität Siegen, Siegen, Germany  
<sup>ag</sup> Stony Brook University, Stony Brook, NY 11794, USA <sup>1</sup>  
<sup>ah</sup> Institute of Physics and Technology of the Mongolian Academy of Sciences, Ulaanbaatar, Mongolia  
<sup>ai</sup> Ulaanbaatar University, Ulaanbaatar, Mongolia  
<sup>aj</sup> Brookhaven National Laboratory, Upton, NY 11973, USA  
<sup>ak</sup> Paul Scherrer Institut, Villigen, Switzerland  
<sup>al</sup> Department of Physics and Astronomy, Purdue University, West Lafayette, IN 47907, USA  
<sup>am</sup> Fachbereich C, Universität Wuppertal, Wuppertal, Germany  
<sup>an</sup> Yerevan Physics Institute, Yerevan, Armenia  
<sup>ao</sup> Departamento de Física Aplicada, CINVESTAV, Mérida, Yucatán, Mexico <sup>9</sup>  
<sup>ap</sup> Institut für Teilchenphysik, ETH, Zürich, Switzerland <sup>10</sup>  
<sup>aq</sup> Physik-Institut der Universität Zürich, Zürich, Switzerland <sup>10</sup>  
<sup>ar</sup> Affiliated with an institute covered by a current or former collaboration agreement with DESY

## ARTICLE INFO

## Article history:

Received 8 May 2023

Received in revised form 17 July 2023

Accepted 25 July 2023

Available online 31 July 2023

Editor: D.F. Geesaman

## ABSTRACT

The radiation pattern within high energy quark- and gluon-initiated jets (jet substructure) is used extensively as a precision probe of the strong force as well as an environment for optimizing event generators with numerous applications in high energy particle and nuclear physics. Looking at electron-proton collisions is of particular interest as many of the complications present at hadron colliders are absent. A detailed study of modern jet substructure observables, jet angularities, in electron-proton collisions is presented using data recorded using the H1 detector at HERA. The measurement is unbinned and multi-dimensional, using machine learning to correct for detector effects. All of the available reconstructed object information of the respective jets is interpreted by a graph neural network, achieving superior precision on a selected set of jet angularities. Training these networks was enabled by the use of a large number of GPUs in the Perlmutter supercomputer at Berkeley Lab. The particle jets are reconstructed in the laboratory frame, using the  $k_T$  jet clustering algorithm. Results are reported at high transverse momentum transfer  $Q^2 > 150 \text{ GeV}^2$ , and inelasticity  $0.2 < y < 0.7$ . The analysis is also performed in sub-regions of  $Q^2$ , thus probing scale dependencies of the substructure variables. The data are compared with a variety of predictions and point towards possible improvements of such models.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Funded by SCOAP<sup>3</sup>.

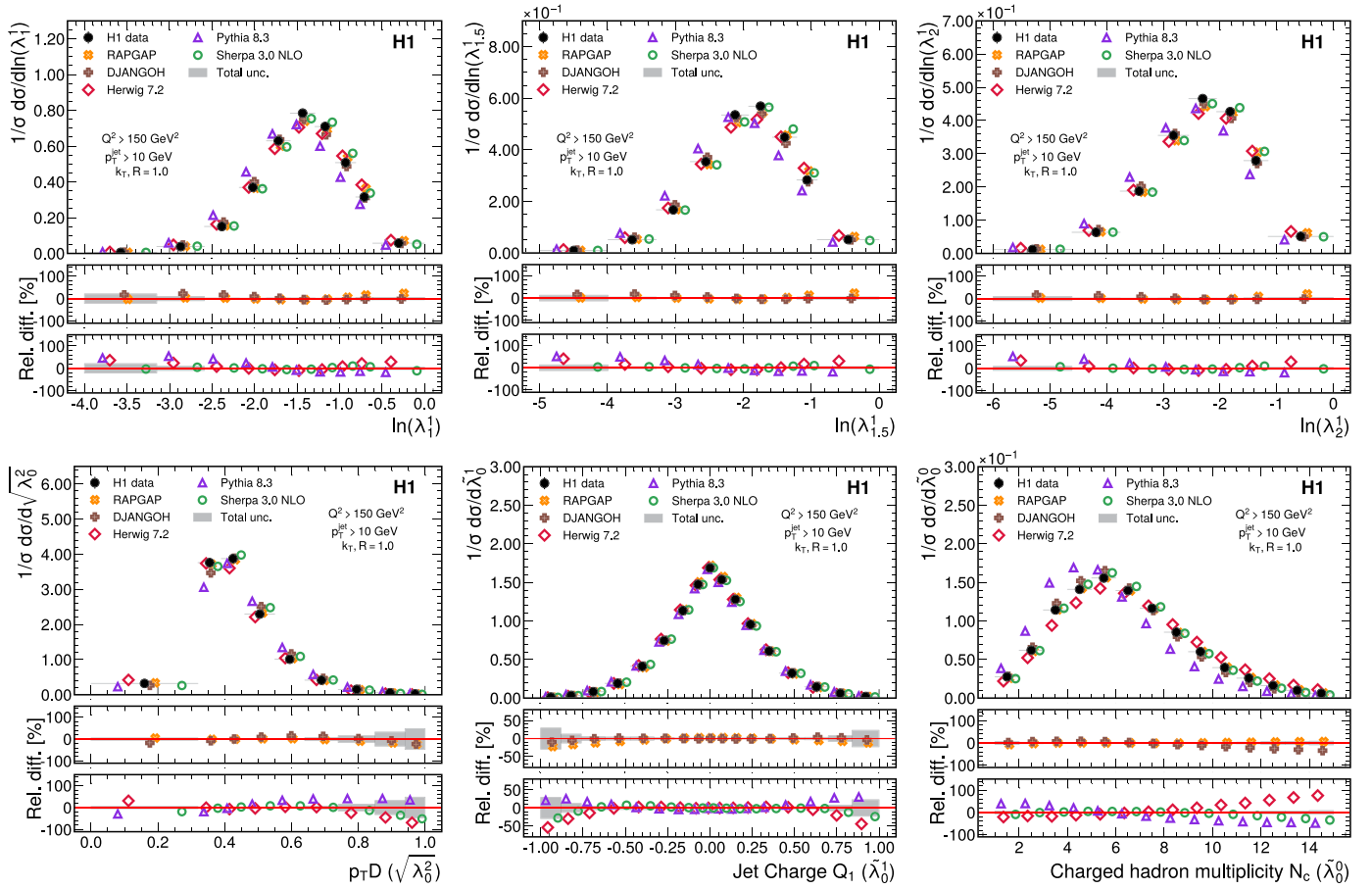
Interactions between quarks and gluons (partons) are described by the theory of Quantum Chromodynamics (QCD) [1]. At high energy particle colliders, outgoing partons produce collimated sprays of particles known as jets. The radiation pattern inside jets (jet substructure) provides insight into the emergent properties of QCD at high energies. Electron-proton collisions are ideal for studies of strong interaction processes. They are induced by an intermediate electroweak gauge particle and provide clean initial conditions. This is in contrast to studies at hadron colliders where the sub-processes are to be studied in the complex context of strong interactions including underlying event effects. The HERA accelerator facility was operated in the years 1992 to 2007 colliding electrons and positrons with protons. Jet properties were extensively studied at HERA during data taking [2–5], but all of these studies pre-date modern jet substructure [6–10]. With new theoretical and

methodological advances, novel insight can be extracted from the preserved HERA data [11,12] for precision QCD studies as well as for event generator improvements.

A canonical set of observables used to explore different aspects of the jet radiation pattern are the generalized angularities [13]:

$$\lambda_{\beta}^{\kappa} = \sum_{i \in \text{jet}} z_i^{\kappa} \left( \frac{R_i}{R_0} \right)^{\beta}, \quad (1)$$

with  $z_i = p_{T,i}/p_T^{\text{jet}}$  for a particle with momentum  $p_{T,i}$  transverse to the incoming beams and clustered inside a jet with distance parameter  $R_0$  and transverse momentum  $p_T^{\text{jet}}$ . The variable  $R_i$  is the Euclidean distance between particle  $i$  and the jet axis in the pseudorapidity-azimuthal angle plane. These observables can be



**Fig. 1.** Measured cross sections, normalized to the inclusive jet production cross section as a function of the jet angularities measured in this work. Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

**Table 1**

Description of the jet substructure observables measured in this work.

Name/Symbol	Observable definition	Charge used
Logarithm of jet broadening	$\ln(\lambda_1^1)$	No
Intermediate observable	$\ln(\lambda_{1,5}^1)$	
Logarithm of jet thrust	$\ln(\lambda_2^1)$	
Momentum dispersion $p_{TD}$	$\sqrt{\lambda_0^2}$	Yes
Jet charge $Q_1$	$\tilde{\lambda}_0^1$	
Charged particle multiplicity $N_c$	$\tilde{\lambda}_0^0$	

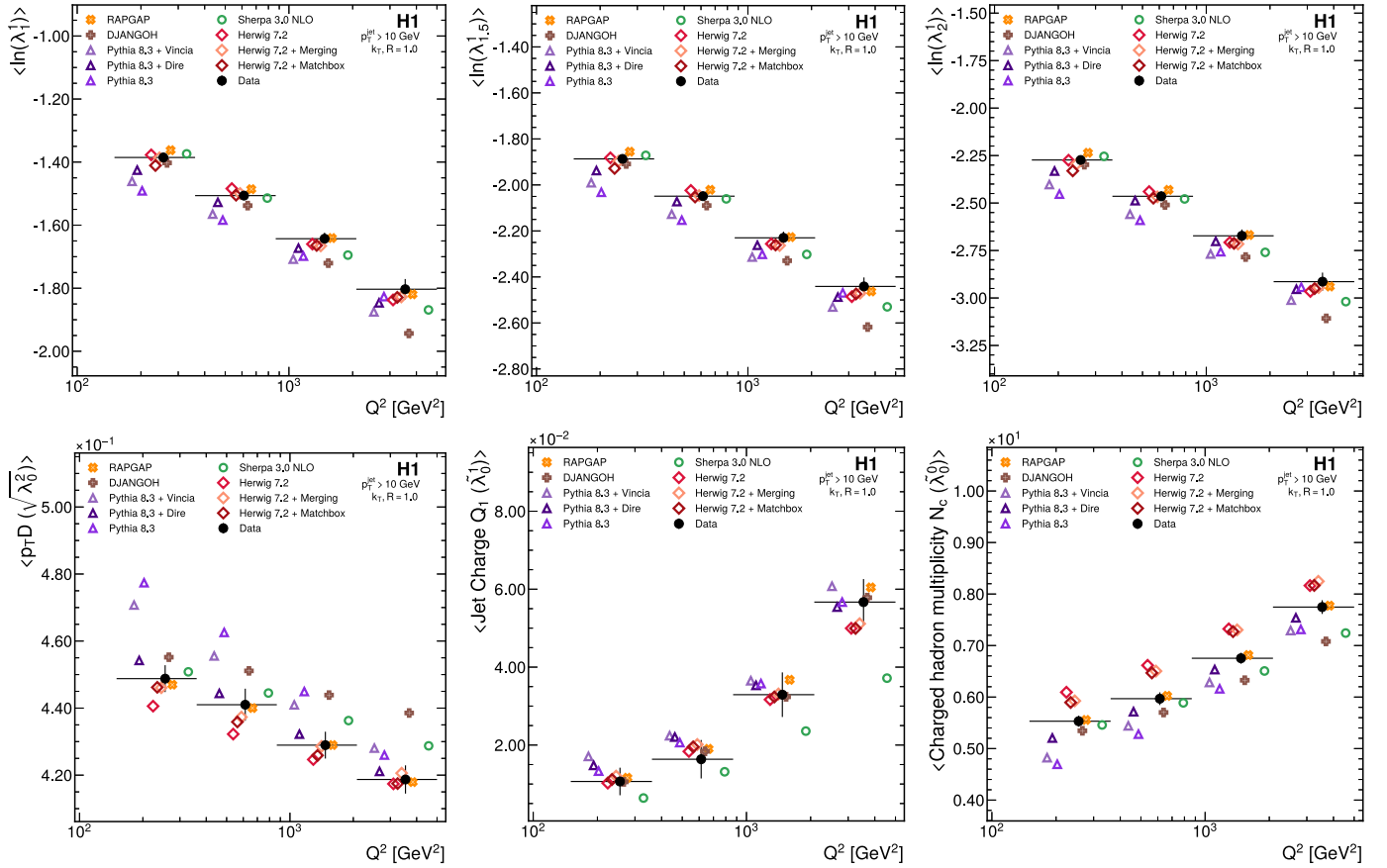
further augmented by multiplying the summand in Eq. (1) with the constituent electric charge  $q_i$  to form the charged-weighted angularities  $\tilde{\lambda}_{\beta}^{\kappa}$ .

In this work, normalized multi-differential cross sections are measured as a function of six jet angularities. The jets are selected at high transverse momenta in neutral current deep inelastic scattering (DIS) at high photon virtualities  $Q^2$ , resulting for the majority of the collisions in single jet events.

The angularities include three infrared and collinear (IRC) safe observables: the jet broadening ( $\lambda_1^1$ ) [14–16], an intermediate observable  $\lambda_{1,5}^1$ , and the jet thrust ( $\lambda_2^1$ ) [17]. Three IRC unsafe variables are also studied: the momentum dispersion  $p_{TD} = \sqrt{\lambda_0^2}$  [18–20], the jet charge  $Q_1 = \tilde{\lambda}_0^1$  [21,22], and the number of charged constituents  $N_c = \tilde{\lambda}_0^0$ . All observables measured in this work are summarized in Table 1.

These observables are chosen because they have been extensively studied in  $e^+e^-$  and  $pp$  collisions and they cover a diverse set of physics effects. The momentum dispersion is known to be one of the best probes for separating jets originating from quarks versus jets originating from gluons [23]. The charged particle multiplicity is also an excellent quark-versus-gluon jet discriminant [24,25] and its evolution with energy scale has been measured at nearly all colliders that have studied jets, including the SPS [26–28], PETRA [29,30], PEP [31–34], TRISTAN [35], CESR [36], LEP [37–48], Tevatron [49], and LHC [50–54]. Precise theory predictions are provided despite of the quantity not being IRC safe [55,56]. The charged constituent multiplicity also is important input to event generator tuning [57]. The jet charge can be used to disentangle quark jets of different flavors [58,59]. Even though it is not collinear-safe, its energy scale dependence can be predicted in perturbation theory [22,60] and can be used to search for scaling violation [58]. The IRC safe observables can be computed in perturbation theory both at a fixed energy scale and as a function of energy scale. The thrust is formally equivalent to the jet mass, which is the most precisely known observable in  $e^+e^-$  and  $pp$  collisions [61–66] and may be used to extract the strong coupling constant [67]. All IRC safe observables are presented here after a logarithmic transformation, highlighting the region dominated by non-perturbative effects. Results are also reported as a function of the energy scale set by the DIS photon virtuality, thus probing the evolution of jet substructure.

New machine learning methods are used to simultaneously correct (unfold) all observables for detector effects, where Graph Neu-



**Fig. 2.** Measured mean of the unfolded jet angularities for multiple  $Q^2$  intervals. Data are shown as solid dots, horizontal bars indicate the bin ranges. Vertical bars represent the total data uncertainties. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity.

ral Networks (GNN) are used for the first time to process all of the reconstructed particles inside jets. Making use of the unbinned nature of the data unfolding, both mean and standard deviation of the measured distributions are provided at multiple  $Q^2$  intervals. They are reported after unfolding and are free of binning effects.

These first measurements, derived from the HERA preserved data, are going to set the foundation for future studies at the future electron-ion collider or possible other electron-proton collider experiments [68–71], which all will operate in kinematic regimes complementary to the HERA machine.

## 1. Experimental method

A full description of the H1 detector can be found elsewhere [72–76]. The detector components that are most relevant for this measurement are described below. Results are reported using a right handed coordinate system with positive  $z$  direction pointing towards the outgoing proton beam and positive  $x$ -axis pointing to the center of the HERA ring. The nominal interaction point is located at  $z = 0$ . The polar angle  $\theta$  is defined with respect to the  $z$  axis, the azimuthal angle  $\phi$  is measured in the  $xy$  plane. The pseudorapidity is defined as  $\eta_{\text{lab}} = -\ln \tan(\theta/2)$ . 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  $15^\circ < \theta < 165^\circ$  and the full azimuthal angle, consists of drift and proportional chambers that are complemented with a silicon vertex detector in the range  $30^\circ < \theta < 150^\circ$  [77]. It 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  $4^\circ < \theta < 154^\circ$  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. Its energy resolution is  $\sigma_E/E = 11\%/\sqrt{E/\text{GeV}} \oplus 1\%$  for leptons [78] and  $\sigma_E/E \approx 50\%/\sqrt{E/\text{GeV}} \oplus 3\%$  for charged pions [79]. In the backward region ( $153^\circ < \theta < 177.5^\circ$ ), energies are measured with a lead-scintillating fiber calorimeter [76]. Results are reported using data recorded by the H1 detector in the years 2006 and 2007 when protons and electrons/positrons (henceforth referred to as ‘electrons’) were collided at energies of 920 GeV and 27.6 GeV, respectively. The total integrated luminosity of this data sample corresponds to  $228 \text{ pb}^{-1}$  [80].

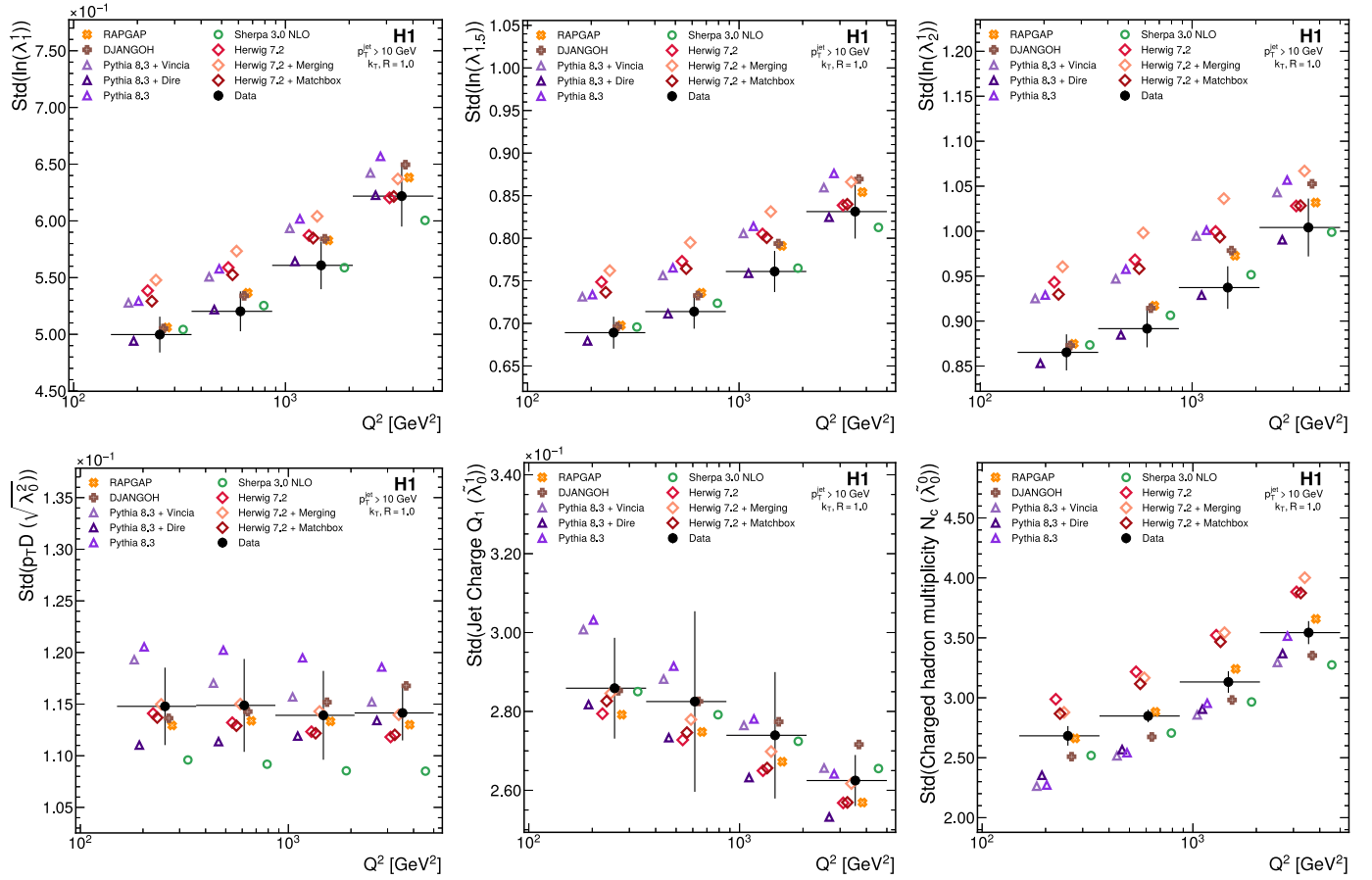
Events are triggered by requiring a high energy cluster in the electromagnetic part of the LAR calorimeter. The scattered electron is identified as the highest transverse momentum LAR cluster matched to a track passing an isolation criteria [81]. Events containing scattered electrons with energy  $E_{e'} > 11 \text{ GeV}$  are kept for further analysis, resulting in a trigger efficiency higher than 99.5% [82,83]. Backgrounds from additional processes such as cosmic rays, beam-gas interactions, photoproduction, charged-current DIS and QED Compton processes are rejected after dedicated selection [83,84], resulting in negligible background contamination.

The inelasticity and photon virtuality are reconstructed using the  $\Sigma$  method [85] defined as

$$y = \frac{\Sigma_{\text{had}}}{\Sigma_{\text{had}} + \Sigma_{e'}}, \quad Q^2 = \frac{E_{e'}^2 \sin^2 \theta_{e'}}{1 - y}, \quad (2)$$

where  $\Sigma_{e'} = E_{e'}(1 - \cos \theta_{e'})$  with  $\theta_{e'}$  the polar angle of the scattered electron and  $\Sigma_{\text{had}} = \sum (E_i - p_{i,z})$  is the difference be-





**Fig. 3.** Measured standard deviation of the unfolded jet angularities for multiple  $Q^2$  intervals. Data are shown as solid dots, horizontal bars indicate the bin ranges. Vertical bars represent the total data uncertainties. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity.

tween the energy and longitudinal momentum sums of the entire hadronic final state (HFS). HFS objects are reconstructed using an energy flow algorithm [86–88] after removing energy clusters and tracks associated to the electron. Additionally, events are required to have  $45 < \Sigma_{\text{had}} + \Sigma_{e'} < 65$  GeV to suppress initial-state QED radiation and contributions from photoproduction.

Jets are defined in the laboratory frame by clustering HFS objects satisfying  $-1.5 < \eta_{\text{lab}} < 2.75$ . The FASTJET 3.3.2 package [89,90] is used with longitudinally invariant  $k_T$  clustering algorithm [91,92] using the default  $E$ -scheme and distance parameter  $R_0 = 1$ . All jets in the event with  $p_T > 5$  GeV are kept for further analysis.

The unfolding targets ('particle-level') are calculated in the simulation using final-state particles with proper lifetime  $c\tau > 10$  mm and excluding the scattered lepton and photons radiated from one of the lepton lines. Reconstructed and generator level jets are matched by requiring the distance  $\Delta R = \sqrt{(\phi_{\text{gen}}^{\text{jet}} - \phi_{\text{reco}}^{\text{jet}})^2 + (\eta_{\text{gen}}^{\text{jet}} - \eta_{\text{reco}}^{\text{jet}})^2} < 0.9$ .

Results are presented after unfolding the data to particle level for events in the kinematic region defined by  $Q^2 > 150$  GeV<sup>2</sup>,  $0.2 < y < 0.7$ ,  $p_T^{\text{jet}} > 10$  GeV, and  $-1.0 < \eta^{\text{jet}} < 2.5$ .

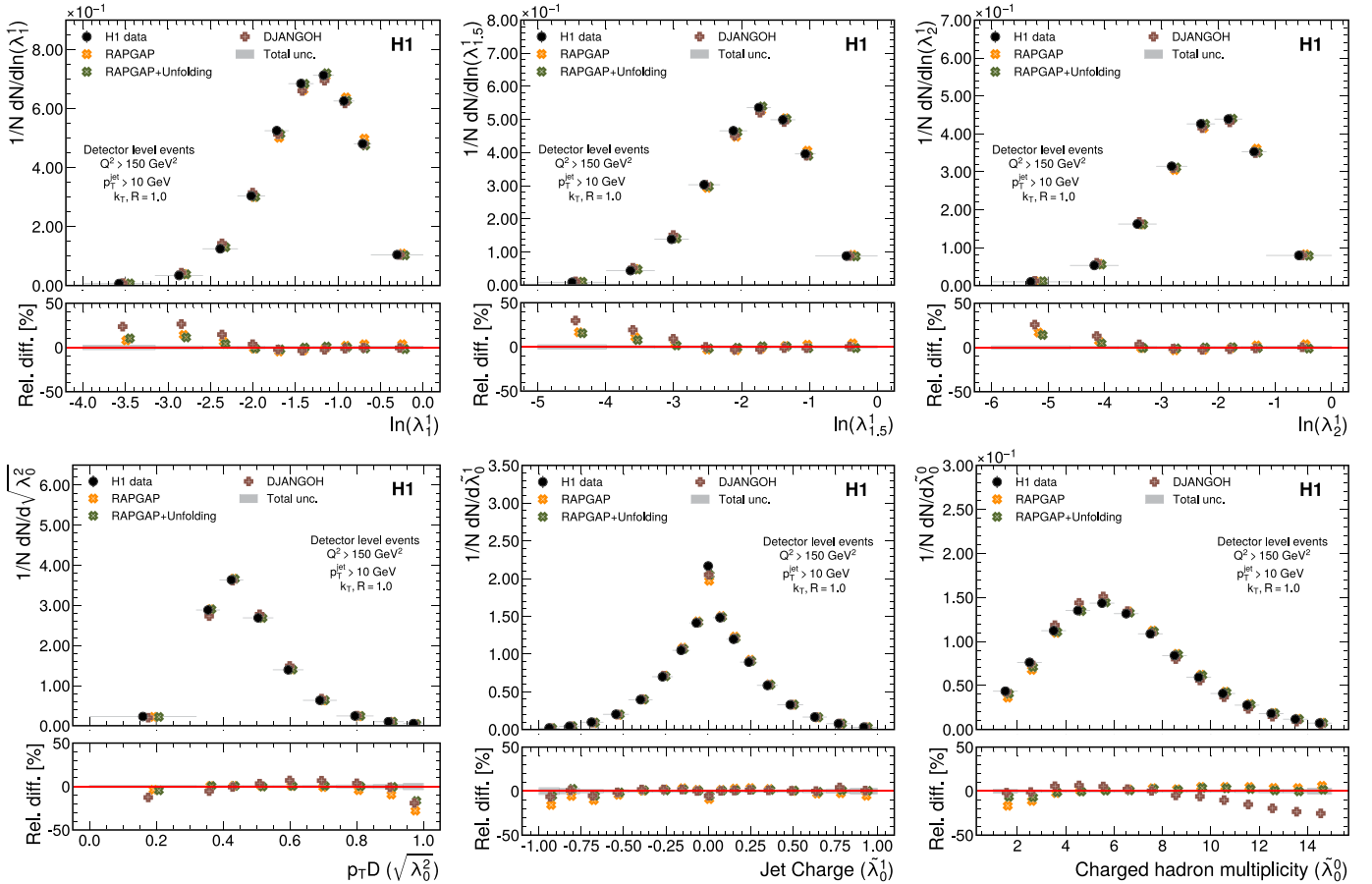
## 2. Monte Carlo simulations

Monte Carlo (MC) simulations are used to correct the data for detector acceptance and resolution effects as well as to compare theoretical predictions with unfolded results.

Detector acceptance and resolution effects are estimated using large samples of simulated events, generated with the DJAN-

GOH 1.4 [93] and RAPGAP 3.1 [94] event generators. Both implement Born level matrix elements and are interfaced with HERACLES [95–97] for QED radiation. The CTEQ6L Parton distribution function (PDF) set [98] and the Lund hadronization model [99] with parameters determined by the ALEPH Collaboration [100] are used for the non-perturbative components. DJANGO uses the Colour Dipole Model as implemented in ARIADNE [101] 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 [102] and reconstructed in the same way as data.

Additional predictions, yet without a detailed simulation of the H1 detector, are made using a set of state-of-the-art generators developed mostly for  $pp$  collisions. Predictions from PYTHIA 8.3 [103, 104] are used for comparison using the default implementation and two additional parton shower implementations: VINCIA [105, 106] and DIRE [107]. VINCIA uses a  $p_T$ -ordered model for QCD + QED showers based on the antenna formalism while DIRE implements a  $p_T$  ordered dipole shower similar to ARIADNE. The NNPDF3.1 PDF set [108] is used for both default and VINCIA implementation and the MMHT14nlo68cl PDF set [109] is used for the DIRE implementation. Predictions from HERWIG 7.2 [110,111] are calculated using the cluster hadronisation model [112] with default implementation parameters and alternative MATCHBox [113] matching and merging [112] schemes. Predictions from a pre-release version of SHERPA 3.0 [114] are provided by the SHERPA authors featuring a new cluster hadronisation model [115] and matrix element calculation at next-to-leading order (NLO) obtained



**Fig. A.4.** Measured cross sections at reconstruction level, normalized to the inclusive jet production cross section, as a function of the jet angularities. Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from DJANGO and RAPGAP simulations are shown for comparison, as well as RAPGAP predictions with unfolding weights. The predictions are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels. Gray bands represent the total systematic uncertainties from detector effects. Data statistical uncertainties are shown as vertical bars, while the statistical uncertainties on the simulations are smaller than the marker sizes.

from OpenLoops with the SHERPA Dipole Shower [116] based on the truncated shower method [117,118].

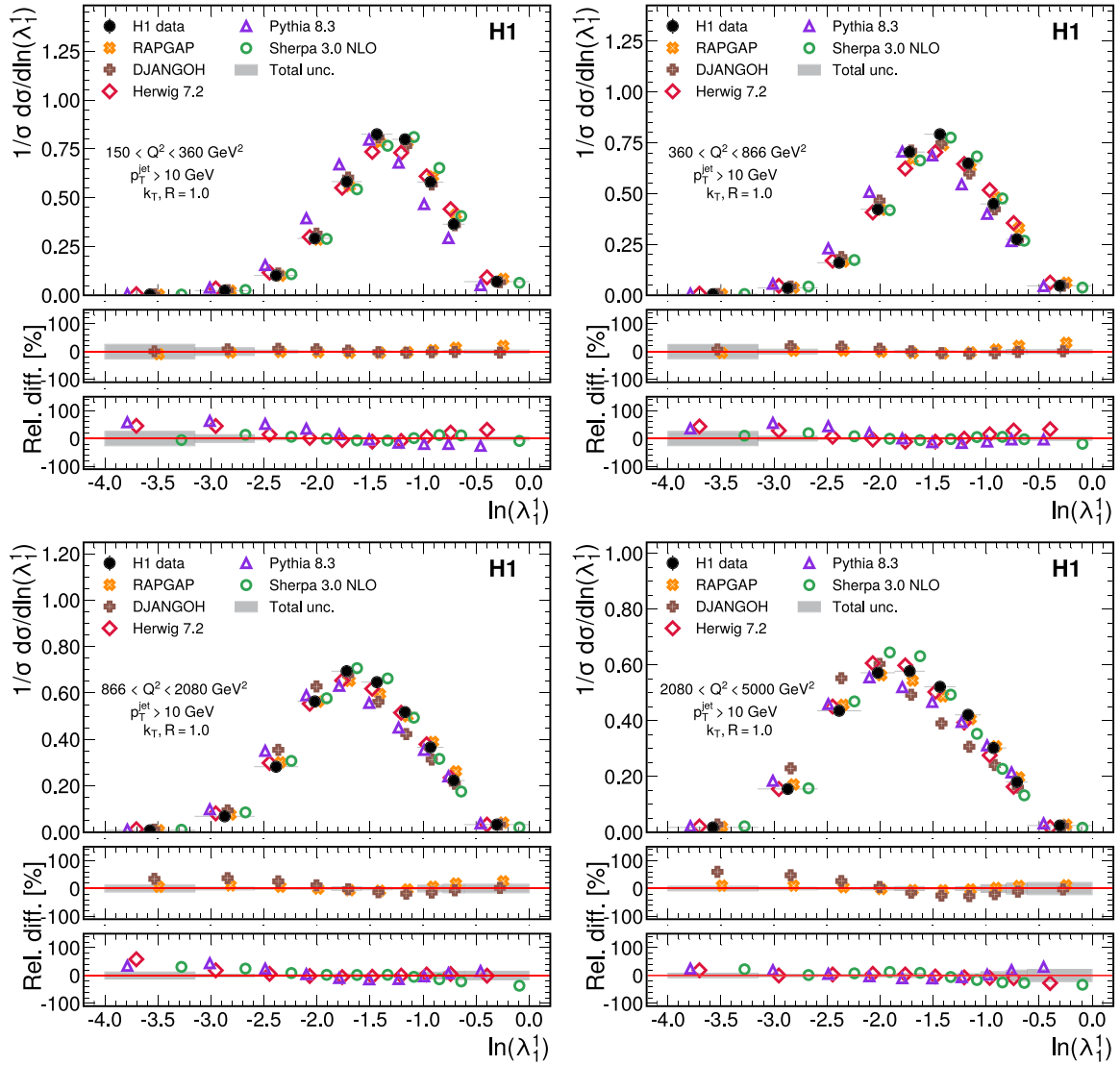
### 3. Unfolding methodology

The unfolding procedure is carried out by simultaneously unfolding the six jet angularities, jet momentum ( $p_T, \eta, \phi$ ), and photon virtuality  $Q^2$  using the OMNIFOLD method [119,120]. OMNIFOLD corrects the data for detector effects using machine learning in an iterative reweighting process that is unbinned and naturally incorporates high dimensional inputs. This method generalizes the Lucy-Richardson deconvolution technique [121–123] designed for binned spectra.

At each iteration, OMNIFOLD employs a set of classifiers trained using neural networks to estimate two reweighting functions used during the unfolding procedure. Each iteration has two steps, with one classifier each. In the first step, a classifier is trained to distinguish data from simulated events, resulting in event weights based on the classifier output. The second step updates the simulation by training a classifier to distinguish between the original simulation and the simulation weighted after the first step. This second step only uses the particle-level information of simulated events as inputs, producing a weight map that is a proper function of the particle level inputs. The updated simulation is then used as the starting point for the following OMNIFOLD iteration. The process is repeated a total of 13 times; six iterations are chosen for the final results as that choice yields the lowest overall uncertainty.

The first step uses up to 30 HFS objects clustered inside jets as inputs to a GNN, where HFS objects are represented as nodes of the graph. This novel hybrid approach has been found to reduce uncertainties by accounting for all possible covariates of the detector response. For each HFS object with momentum ( $p_T, \eta, \phi$ ) and electric charge  $q$ , the kinematic information used for the first step classifier is the set  $(z, \Delta\eta, \Delta\phi, q)$ , with  $z = p_T/p_T^{\text{jet}}$ ,  $\Delta\eta = \eta - \eta^{\text{jet}}$ , and  $\Delta\phi = \phi - \phi^{\text{jet}}$ . Jet momentum ( $p_T^{\text{jet}}, \eta^{\text{jet}}, \phi^{\text{jet}}$ ) and  $Q^2$  information are also included in the aggregation function of the graph implementation. In total,  $30 \times 4 + 4 = 124$  features are considered during this first step, requiring the network implementation to learn the differences between data and simulation in a high-dimensional setting. Leveraging the success of GNNs for jet classification [124–135] based on particle sets as inputs, the Point Cloud Transformer (PCT) [132,136] GNN architecture is chosen. This GNN architecture is optimized to learn the relationship between nearby particles through the use of attention layers, resulting in state-of-the-art classification performance for different types of jets [132]. The local neighborhood in PCT is defined by selecting the five closest neighbors of each particle with distances calculated in the  $\eta - \phi$  plane. The impact of adding more than 30 particles was found to be negligible.

In the second step, a simplified classifier architecture is used. The inputs of the classifier are the same jet kinematic information and  $Q^2$  used in the previous step, but HFS objects are replaced by the jet observables that are reported in the final results. This choice decreases the number of inputs by an order of magnitude



**Fig. B.5.** Measured cross sections, normalized to the inclusive jet production cross section, for multiple  $Q^2$  intervals as a function of  $\ln(\lambda_1^j)$ . Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

and reduces the training time of the classifier by a factor five. Since the first step is tasked to incorporate as much information as possible from the data, the simplified classifier adopted for the second step does not degrade the unfolding results while greatly reducing the computational cost of the method. This simplified classifier is implemented using a fully-connected neural network with three hidden layers with 64, 128, and 64 nodes per layer. Different hyperparameter choices of the neural networks were also tested, but no significant difference in the final results was observed. The robustness of the classifiers is improved by training an ensemble of 10 network models for each OMNIFOLD step. The average response of the ensemble is then used to derive the reweighting functions.

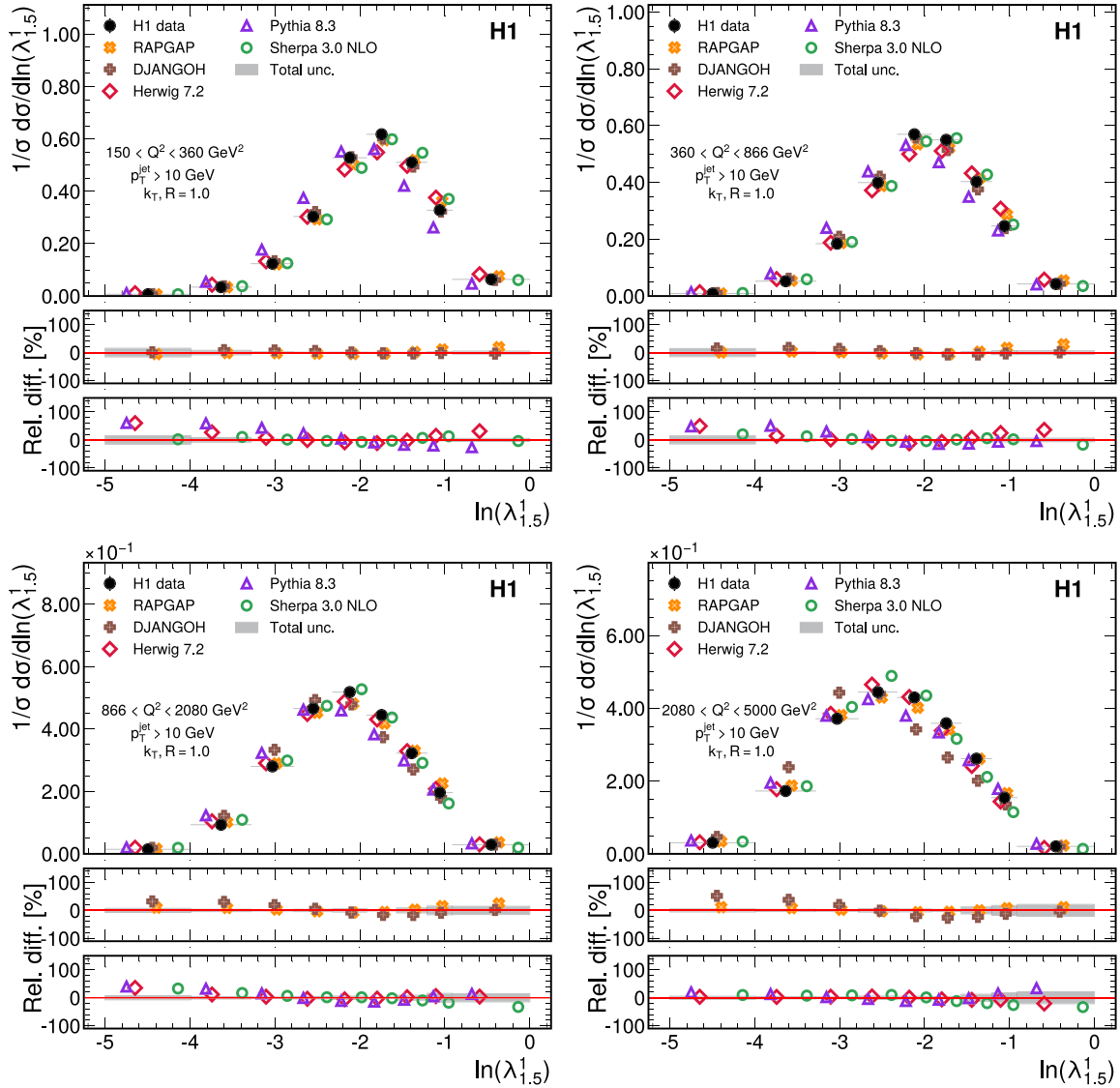
The output of all classifiers is passed through a sigmoid activation function and trained using the binary cross-entropy loss function. The training proceeds until the validation loss, estimated with a statistically independent data set, does not improve for 10 consecutive epochs. All machine learning methods are implemented using the KERAS [137] and TENSORFLOW [138] libraries, using the ADAM [139] optimization algorithm. In total around  $50 \times 10^6$  jets are used during the training of the method split between data and simulation. To derive the unfolded results and uncertainties,

2800 neural networks are trained independently using 128 graphic processing units (GPUs) simultaneously with the Perlmutter supercomputer [140] and HOROVOD [141] library used for parallel distributed training.

#### 4. Uncertainties

Systematic uncertainties on the reconstruction of the data observables are estimated by varying the relevant parameters in the simulation and carrying out the full unfolding procedure for each simulation variant. Uncertainties on HFS objects include the energy scale from two different contributions: HFS objects contained in high  $p_T$  jets and other HFS objects. In both cases, the energy-scale uncertainty is  $\pm 1\%$ . Both uncertainty sources are estimated separately [83,142] by varying the corresponding HFS energy by  $\pm 1\%$ . An uncertainty of  $\pm 20$  mrad is assigned to the azimuthal angle determination of HFS objects. Lepton uncertainties are considered in the  $Q^2$  determination and the uncertainty on the lepton energy scale ranges from  $\pm 0.5\%$  to  $\pm 1\%$  [83,143]. Uncertainties on the azimuthal angle of the scattered lepton are estimated to be  $\pm 1$  mrad [82]. In Appendix A distributions on reconstructed level





**Fig. B.6.** Measured cross sections, normalized to the inclusive jet production cross section, for multiple  $Q^2$  intervals as a function of  $\ln(\lambda_{1,5}^1)$ . Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

and the uncertainties are compared with predictions from DJANGO and RAPGAP generators.

Additional uncertainties from the unfolding procedure are estimated to cover a possible bias from the generator choice used to perform the unfolding procedure. A model uncertainty is estimated by unfolding the data with DJANGO instead of the default RAPGAP. In addition, a non-closure uncertainty is estimated by unfolding DJANGO as pseudo-data using the RAPGAP simulation and comparing to DJANGO at particle level. The combined effects of the unfolding uncertainties typically are below 10% and increase only in the more extreme regions of the jet angularities. QED corrections accounting for virtual and real higher-order QED effects as implemented in HERACLES are small and the full effect is taken as an uncertainty, estimated by comparing predictions with and without first order QED radiative effects at the leptonic vertex. The statistical uncertainty is estimated using the bootstrap technique [144]. The unfolding procedure is repeated on 100 pseudo datasets, each defined by resampling the original dataset with replacement. Normalization uncertainties, such as luminosity scale

and trigger efficiencies, are not considered since they cancel in the ratio of the normalized differential cross section results.

## 5. Results

The results of this analysis are unbinned normalized cross sections as a function of  $Q^2$  and the generalized jet angularities. Unfolded results of the normalized differential particle level cross sections of the jet substructure observables are presented in Fig. 1 in the kinematic region described in Sec. 1. In order to visualize the unfolded distributions bins have been introduced based on the respective detector resolution. All distributions are observed to accumulate around a distinct peak. Mean and standard deviation are calculated from the unbinned, unfolded event sample, and are shown for the six jet angularities in four  $Q^2$  ranges in Figs. 2 and 3, respectively, with numerical results reported in Tables 2 and 3. Distributions of the angularities in the four  $Q^2$  ranges are shown in Appendix B.

Predictions by the DIS MC generators, DJANGO and RAPGAP, and by general purpose simulators, SHERPA, HERWIG, PYTHIA show

**Table 2**

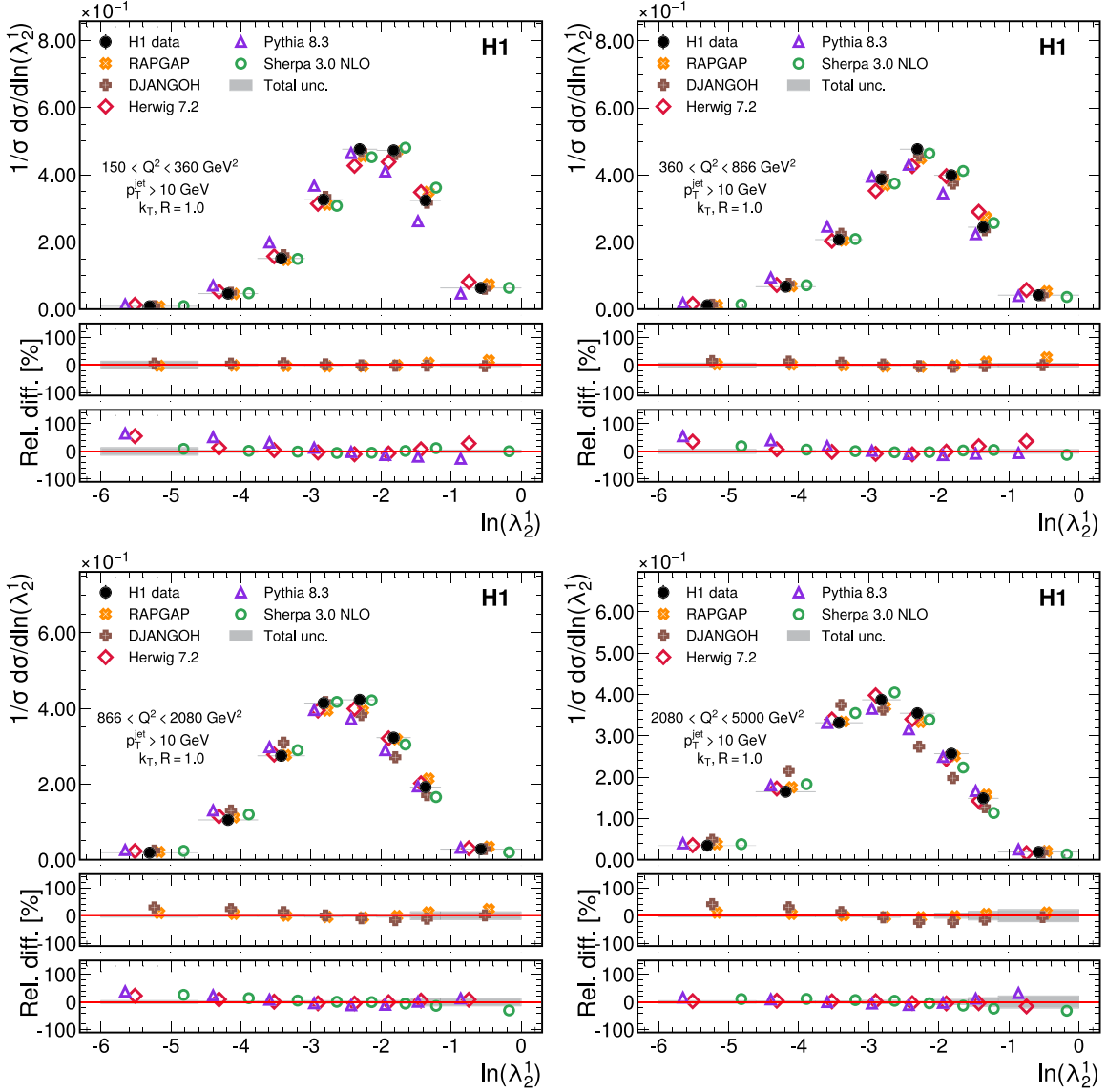
Measured means of the jet angularities in multiple  $Q^2$  intervals. The statistical (Stat.) and total (Tot.) uncertainties are given, as well as the seven systematic uncertainty sources described in the main text.

$Q^2$ [GeV $^2$ ]	$\langle \ln(\lambda_1^1) \rangle$	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	-1.385	0.004	0.007	0.004	0.004	0.004	0.002	0.002	0.002	0.002
[360.42, 866.03]	-1.506	0.005	0.009	0.003	0.000	0.004	0.002	0.004	0.007	0.009
[866.03, 2080.90]	-1.643	0.006	0.012	0.006	0.005	0.000	0.014	0.005	0.004	0.005
[2080.90, 5000.00]	-1.803	0.010	0.018	0.008	0.009	0.010	0.020	0.008	0.009	0.014
$Q^2$ [GeV $^2$ ]	$\langle \ln(\lambda_{1.5}^1) \rangle$	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	-1.887	0.006	0.007	0.006	0.006	0.004	0.003	0.004	0.003	0.002
[360.42, 866.03]	-2.049	0.006	0.010	0.004	0.004	0.006	0.001	0.006	0.009	0.013
[866.03, 2080.90]	-2.230	0.008	0.012	0.009	0.006	0.004	0.020	0.007	0.007	0.006
[2080.90, 5000.00]	-2.441	0.013	0.016	0.010	0.012	0.013	0.026	0.011	0.008	0.008
$Q^2$ [GeV $^2$ ]	$\langle \ln(\lambda_2^1) \rangle$	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	-2.273	0.007	0.007	0.007	0.007	0.004	0.004	0.006	0.004	0.003
[360.42, 866.03]	-2.464	0.008	0.010	0.002	0.006	0.007	0.003	0.008	0.010	0.018
[866.03, 2080.90]	-2.673	0.010	0.012	0.011	0.006	0.007	0.025	0.008	0.011	0.003
[2080.90, 5000.00]	-2.914	0.015	0.016	0.011	0.014	0.016	0.031	0.013	0.014	0.013
$Q^2$ [GeV $^2$ ]	$\langle \sqrt{\lambda_0^2} \rangle$	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.449	0.001	0.009	0.001	0.001	0.002	0.001	0.002	0.001	0.001
[360.42, 866.03]	0.441	0.001	0.011	0.001	0.001	0.002	0.001	0.003	0.001	0.001
[866.03, 2080.90]	0.429	0.001	0.009	0.001	0.001	0.001	0.001	0.003	0.001	0.001
[2080.90, 5000.00]	0.419	0.002	0.010	0.002	0.002	0.002	0.001	0.001	0.002	0.001
$Q^2$ [GeV $^2$ ]	$\langle \tilde{\lambda}_0^1 \rangle$	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.011	0.001	0.332	0.001	0.001	0.001	0.001	0.001	0.001	0.001
[360.42, 866.03]	0.016	0.001	0.302	0.001	0.001	0.002	0.003	0.001	0.001	0.001
[866.03, 2080.90]	0.033	0.002	0.174	0.001	0.001	0.003	0.004	0.001	0.001	0.001
[2080.90, 5000.00]	0.057	0.003	0.104	0.002	0.003	0.002	0.001	0.001	0.002	0.002
$Q^2$ [GeV $^2$ ]	$\langle \tilde{\lambda}_0^0 \rangle$	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	5.531	0.019	0.019	0.010	0.020	0.047	0.006	0.041	0.012	0.081
[360.42, 866.03]	5.969	0.024	0.021	0.021	0.023	0.020	0.020	0.036	0.070	0.081
[866.03, 2080.90]	6.753	0.035	0.016	0.012	0.028	0.032	0.048	0.006	0.037	0.066
[2080.90, 5000.00]	7.747	0.056	0.017	0.053	0.056	0.050	0.050	0.028	0.019	0.050

**Table 3**

Measured standard deviations of the jet angularities in multiple  $Q^2$  intervals. The statistical (Stat.) and total (Tot.) uncertainties are given, as well as the seven systematic uncertainty sources described in the main text.

$Q^2$ [GeV $^2$ ]	std ( $\ln(\lambda_1^1)$ )	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.500	0.002	0.016	0.008	0.005	0.008	0.008	0.004	0.002	0.002
[360.42, 866.03]	0.520	0.003	0.018	0.009	0.005	0.009	0.009	0.006	0.003	0.002
[866.03, 2080.90]	0.561	0.003	0.021	0.012	0.006	0.005	0.011	0.007	0.008	0.003
[2080.90, 5000.00]	0.622	0.004	0.027	0.012	0.006	0.000	0.011	0.007	0.018	0.004
$Q^2$ [GeV $^2$ ]	std ( $\ln(\lambda_{1.5}^1)$ )	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.689	0.003	0.019	0.010	0.005	0.010	0.009	0.004	0.002	0.002
[360.42, 866.03]	0.714	0.004	0.020	0.010	0.005	0.010	0.010	0.006	0.004	0.003
[866.03, 2080.90]	0.761	0.004	0.024	0.014	0.006	0.005	0.012	0.008	0.010	0.004
[2080.90, 5000.00]	0.831	0.005	0.032	0.015	0.009	0.002	0.013	0.008	0.021	0.002
$Q^2$ [GeV $^2$ ]	std ( $\ln(\lambda_2^1)$ )	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.865	0.004	0.020	0.011	0.005	0.012	0.010	0.003	0.002	0.002
[360.42, 866.03]	0.892	0.004	0.021	0.010	0.004	0.011	0.010	0.005	0.004	0.005
[866.03, 2080.90]	0.937	0.005	0.024	0.014	0.003	0.004	0.012	0.008	0.010	0.003
[2080.90, 5000.00]	1.004	0.006	0.032	0.016	0.010	0.003	0.013	0.009	0.019	0.005
$Q^2$ [GeV $^2$ ]	std ( $\sqrt{\lambda_0^2}$ )	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.115	0.001	0.004	0.002	0.001	0.002	0.002	0.002	0.001	0.001
[360.42, 866.03]	0.115	0.001	0.005	0.002	0.001	0.002	0.001	0.002	0.001	0.001
[866.03, 2080.90]	0.114	0.001	0.004	0.002	0.000	0.002	0.000	0.002	0.002	0.001
[2080.90, 5000.00]	0.114	0.001	0.003	0.000	0.001	0.001	0.001	0.001	0.002	0.001
$Q^2$ [GeV $^2$ ]	std ( $\tilde{\lambda}_0^1$ )	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	0.286	0.001	0.013	0.001	0.001	0.001	0.001	0.001	0.001	0.012
[360.42, 866.03]	0.283	0.001	0.023	0.001	0.001	0.001	0.001	0.002	0.000	0.023
[866.03, 2080.90]	0.274	0.001	0.016	0.000	0.001	0.001	0.001	0.002	0.001	0.016
[2080.90, 5000.00]	0.262	0.002	0.006	0.002	0.001	0.001	0.000	0.002	0.000	0.005
$Q^2$ [GeV $^2$ ]	std ( $\tilde{\lambda}_0^0$ )	Stat.	Tot.	HFS (jet)	HFS (other)	HFS ( $\phi$ )	Lepton (E)	Lepton ( $\phi$ )	Model	Closure
[150.00, 360.42]	2.683	0.013	0.082	0.012	0.034	0.023	0.006	0.036	0.043	0.038
[360.42, 866.03]	2.848	0.012	0.052	0.012	0.003	0.009	0.011	0.026	0.011	0.038
[866.03, 2080.90]	3.132	0.020	0.090	0.030	0.015	0.019	0.021	0.018	0.068	0.029
[2080.90, 5000.00]	3.544	0.031	0.096	0.028	0.007	0.029	0.033	0.030	0.066	0.015



**Fig. B.7.** Measured cross sections, normalized to the inclusive jet production cross section, for multiple  $Q^2$  intervals as a function of  $\ln(\lambda_2^1)$ . Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

a good agreement with data for most observables. For HERWIG, the largest differences to data are observed in the charged hadron multiplicity, predicting a higher number of particles than observed. In contrast, PYTHIA underpredicts the charged particle multiplicity. SHERPA shows a good agreement in all IRC-safe observables and is able to describe the data at low observables values, where contributions from soft-radiation are dominant and challenging to describe, as evidenced by a worse agreement by PYTHIA and HERWIG.

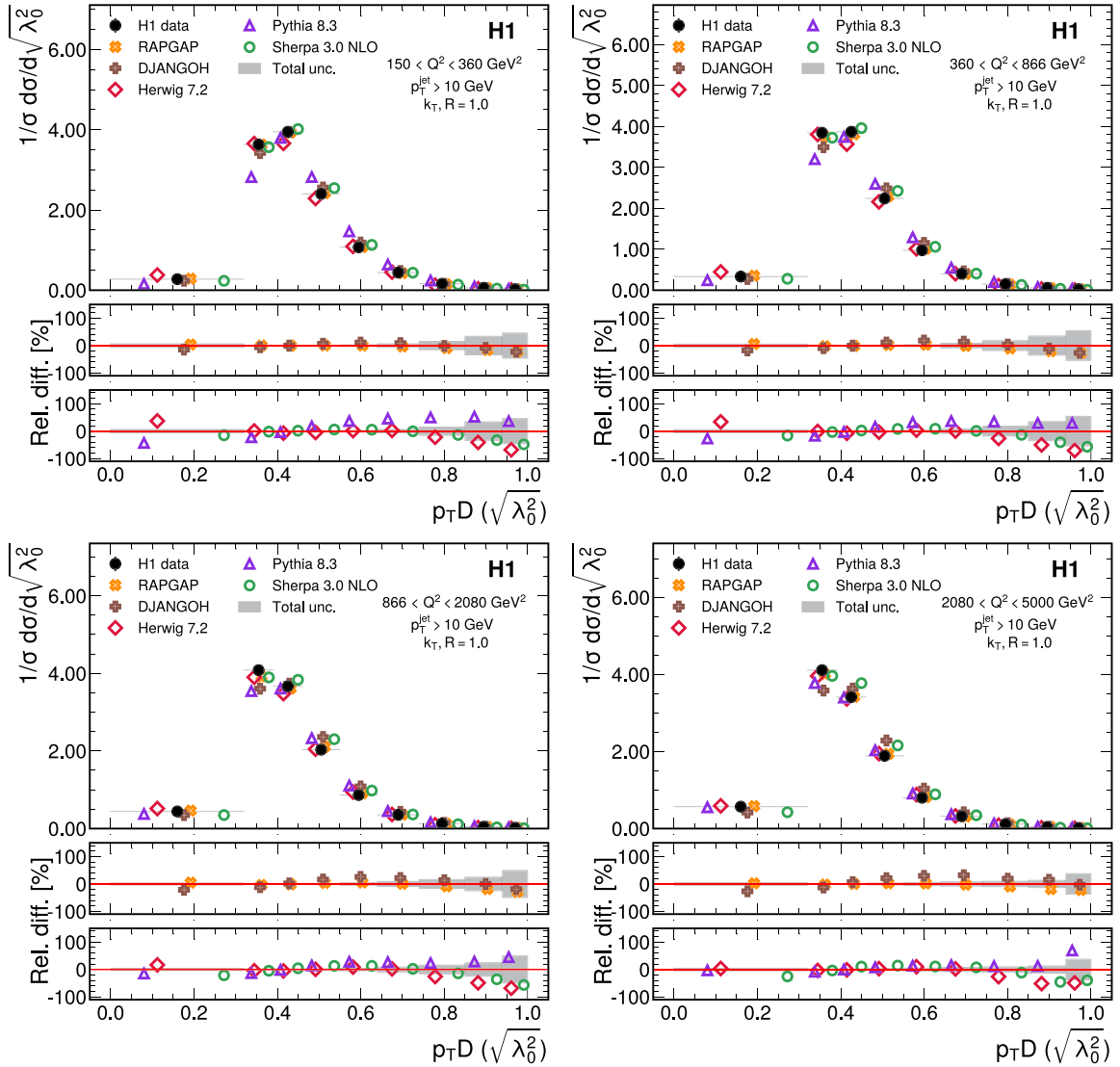
A larger selection of predictions is used to study the mean and standard deviations as a function of  $Q^2$  in Figs. 2 and 3. RAPGAP shows a good agreement with data in all distributions and  $Q^2$  intervals, while DJANGO shows worse agreement with data at high  $Q^2$  values. Of the three PYTHIA prediction variants, where different parton shower models are studied, the DIRE parton shower provides the best overall description of the data. The HERWIG variants, where NLO matrix elements and details of the matching procedure are studied, all are in overall good agreement with data. Large deviations from data are observed in the charged hadron multiplicity,

which has a high sensitivity to the modeling of the hadronization process. Predictions from SHERPA show a better agreement with data at lower  $Q^2$  values and larger discrepancies at the highest  $Q^2$  interval. SHERPA is also able to provide a good description of the standard deviation of the IRC jet angularities together with DIS dedicated generators and the PYTHIA implementation with DIRE parton shower. In contrast, all other implementations predict a larger standard deviation than observed.

## 6. Conclusion

A first measurement of jet angularities in neutral current DIS events with  $Q^2 > 150 \text{ GeV}^2$  and  $0.2 < y < 0.7$ , as well as divided into four  $Q^2$  intervals is presented. The distributions as well as mean and standard deviation of these jet substructure observables are sensitive to perturbative QCD effects and to hadronization models used to describe jet formation.

The process of unfolding detector effects makes use of novel machine learning methods. All measured distributions are simul-



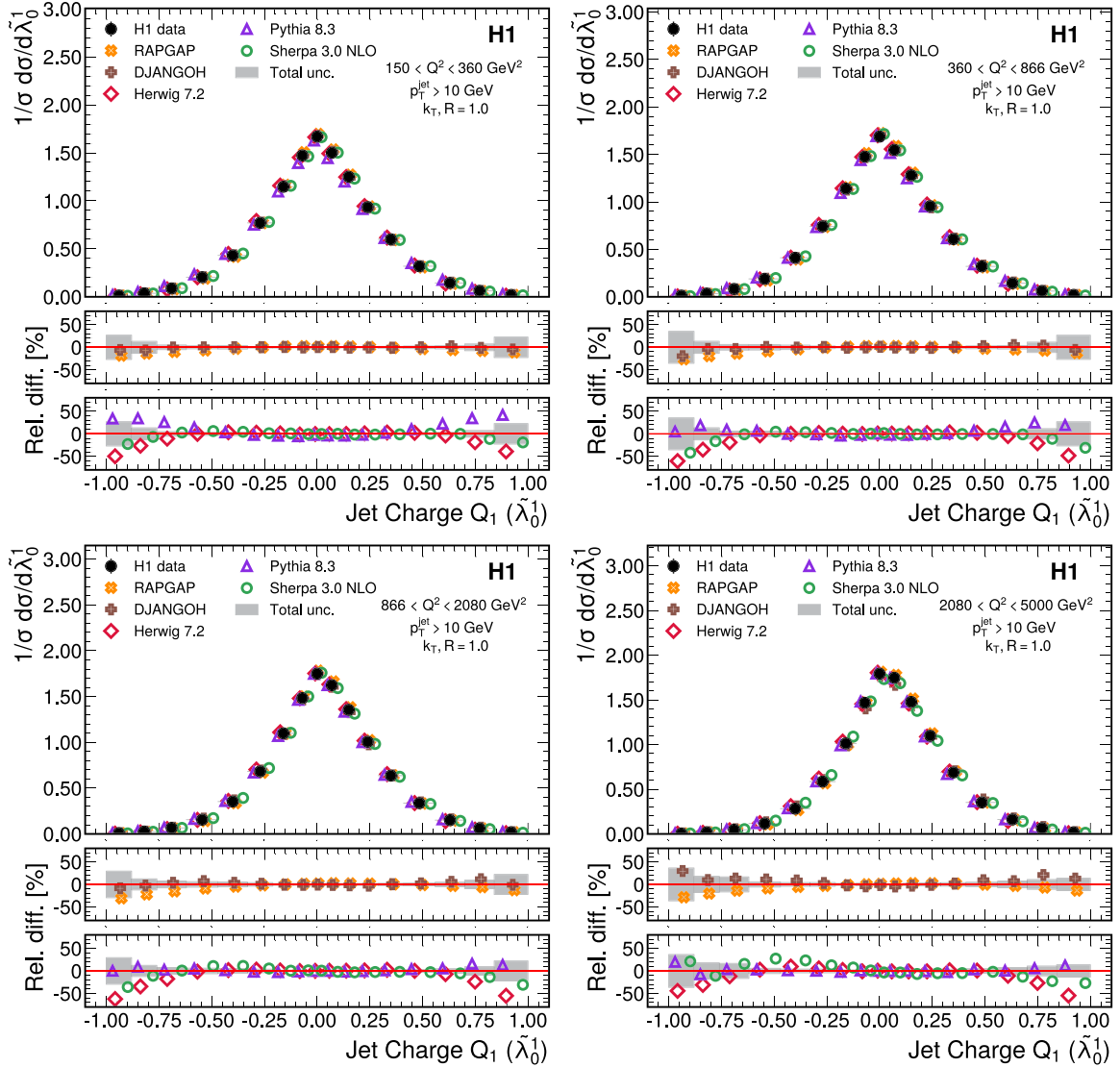
**Fig. B.8.** Measured cross sections, normalized to the inclusive jet production cross section, for multiple  $Q^2$  intervals (columns) as a function of  $p_T D$ . Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

taneously unfolded using the OMNI FOLD approach. Kinematic information from each reconstructed particle clustered inside a jet is input to a dedicated Graph Neural Network implementation that learns the correlation between particles clustered inside a jet. On particle level, however, the jet substructure is represented by the generalized angularities alone. The success of this approach demonstrates that different sets of unbinned observables at detector level and particle level can be used to determine the unfolded distributions. The training of over two thousand neural networks is carried out on the Perlmutter supercomputer using 128 GPUs simultaneously during the unfolding procedure.

While the unfolding procedure is unbinned, results are provided as histograms for both single- and multi-differential cross sections to ease the comparison with different theory predictions. Mean and standard deviation of all observables are calculated after unfolding in multiple  $Q^2$  ranges. Theory predictions from dedicated DIS simulators provide an overall good description of all measured quantities. In particular, a good agreement between data and RAPGAP simulation is observed for all distributions and  $Q^2$  intervals. General purpose simulators are also able to describe the data well. PYTHIA interfaced with DIRE parton shower, HERWIG, and SHERPA

predictions all show a good agreement with most of the jet angularities studied. The largest differences to data are observed for the charged hadron multiplicity, which is expected to be most sensitive to non-perturbative effects. The H1 data can be used to further improve the precision of event generators and analytic calculations. The  $Q^2$  dependence is measured and can be compared to scale evolution predictions.

The measurements can serve as a complementary guide for jet substructure studies at the future Electron-Ion Collider, which is going to record much larger data samples in a similar kinematic region. Results presented in this work make use of all detector objects during the unfolding procedure, learning the relationship between reconstructed objects using neural networks. This points to a possible more general algorithm for data unfolding. In the generalization, all particles or jets in a collisions could be made available after unfolding. This would greatly simplify future comparisons to predictions, in particular for specialized theories valid only in small phase-space regions, or for new observables not known to date.



**Fig. B.9.** Measured cross sections, normalized to the inclusive jet production cross section, for multiple  $Q^2$  intervals as a function of  $Q_1$ . Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

Data will be made available on request.

### Acknowledgements

We are grateful to the HERA machine group whose outstanding efforts have made this experiment possible. We thank the engineers and technicians for their work in constructing and maintaining the H1 detector, our funding agencies for financial support, the DESY technical staff for continual assistance and the DESY directorate for support and for the hospitality which they extend to the non-DESY members of the collaboration.

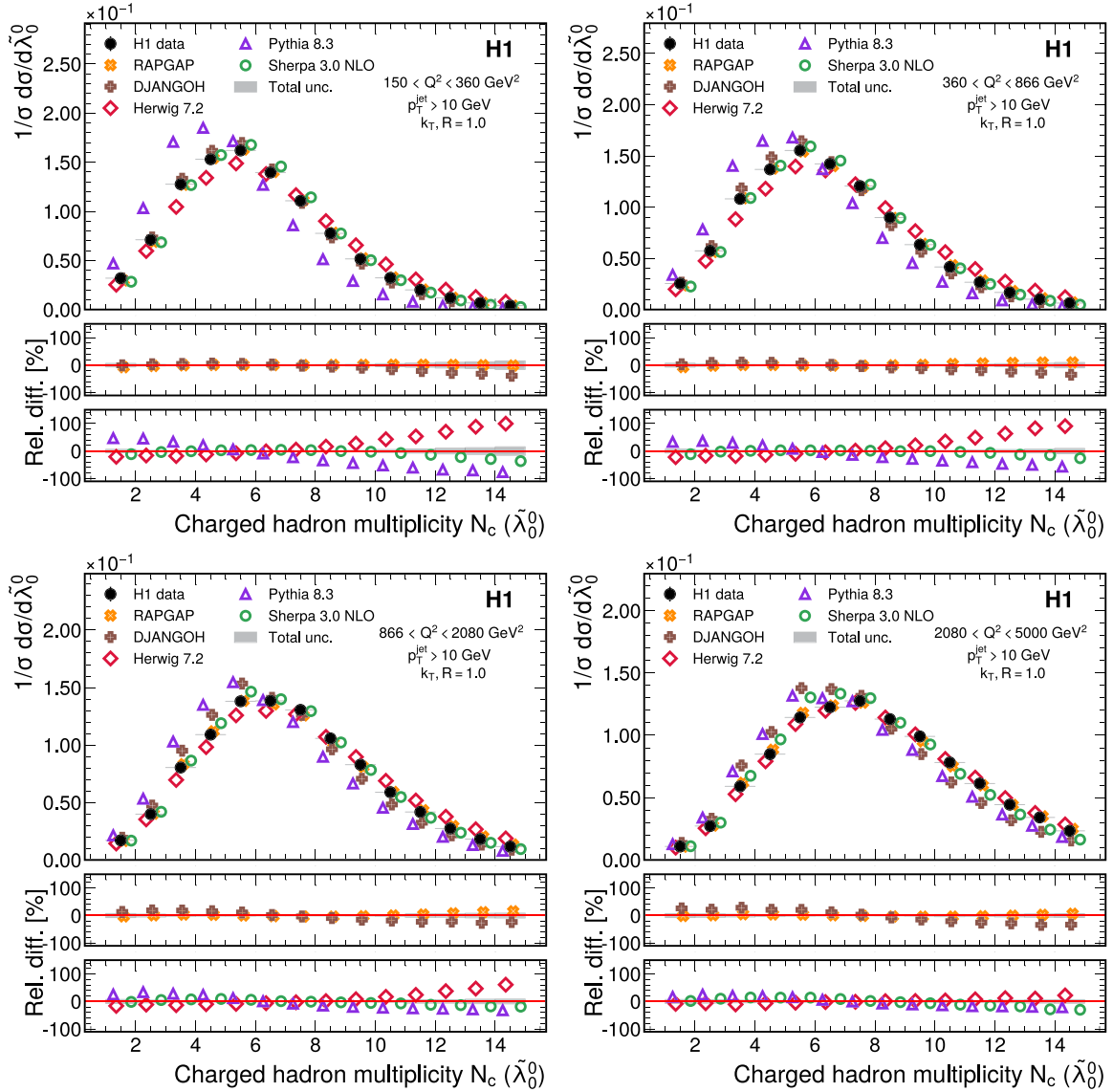
We express our thanks to all those involved in securing not only the H1 data but also the software and working environment

for long term use, allowing the unique H1 data set to continue to be explored. The transfer from experiment specific to central resources with long term support, including both storage and batch systems, has also been crucial to this enterprise. We therefore also acknowledge the role played by DESY-IT and all people involved during this transition and their future role in the years to come.

### Appendix A. Data and simulation comparison at reconstruction level

Prior to unfolding, a comparison between the data and simulation is performed using the selection criteria described in Sec. 1 and the angularities calculated from the reconstructed detector objects. Both RAPGAP and DJANGO simulations are compared to data in Fig. A.4. The RAPGAP simulation with unfolding weights applied is also shown.





**Fig. B.10.** Measured cross sections, normalized to the inclusive jet production cross section, for multiple  $Q^2$  intervals as a function of  $N_c$ . Data are shown as solid dots, horizontal bars indicate the bin ranges. Predictions from multiple simulations are shown for comparison, and are offset horizontally for visual clarity. The relative differences between data and predictions are shown in the bottom panels, split between dedicated DIS simulators (middle) and general purpose simulators (bottom). Gray bands represent the total data uncertainties.

## Appendix B. Binned distribution of unfolded observables at different energy scales

In this section, the normalized differential cross section of the jet substructure observables after unfolding are presented in Figs. B.9, B.10, B.8, B.5, B.6, and B.7. Histograms showing the results are created after detector unfolding for comparison with different generators. The normalized differential cross section for observable  $\mathcal{O}$  is calculated as:

$$\frac{1}{\sigma} \frac{d\sigma}{d\mathcal{O}} = \frac{1}{N} \frac{\Delta N}{\Delta X}, \quad (\text{B.1})$$

with  $N$  being the total number of unfolded jets in the considered  $Q^2$  range and  $\Delta N$  representing the number of unfolded jets in the respective bin of  $\mathcal{O}$ . The bin width is denoted  $\Delta X$ .

## Appendix C. Data tables for all presented histograms

Measured values of the normalized differential cross sections and contributions from each uncertainty source to the unfolded values are listed in Tables C.4, C.5, C.6, C.7, and C.8.

## References

- [1] F. Gross, et al., 50 years of quantum chromodynamics, arXiv:2212.11107, 2022.
- [2] C. Adloff, et al., H1, Measurement of internal jet structure in dijet production in deep inelastic scattering at HERA, Nucl. Phys. B 545 (1999) 3–20, [https://doi.org/10.1016/S0550-3213\(99\)00118-2](https://doi.org/10.1016/S0550-3213(99)00118-2), arXiv:hep-ex/9901010.
- [3] J. Breitweg, et al., ZEUS, Measurement of jet shapes in high  $Q^2$  deep inelastic scattering at HERA, Eur. Phys. J. C 8 (1999) 367–380, <https://doi.org/10.1007/s100520050471>, arXiv:hep-ex/9804001.
- [4] S. Chekanov, et al., ZEUS, Measurement of subjet multiplicities in neutral current deep inelastic scattering at HERA and determination of  $\alpha(s)$ , Phys. Lett. B 558 (2003) 41–58, [https://doi.org/10.1016/S0370-2693\(03\)00216-8](https://doi.org/10.1016/S0370-2693(03)00216-8), arXiv:hep-ex/0212030.

**Table C.4**

Measured values of the normalized unfolded differential cross sections and uncertainties for each systematic source considered in the range  $Q^2 > 150 \text{ GeV}^2$ .

$\ln(\lambda_1^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_1^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-4.00, -3.15]	0.006	3.122	21.454	11.697	6.089	10.095	9.280	8.787	3.030	1.483
[-3.15, -2.59]	0.038	2.238	10.056	5.187	4.311	3.258	4.283	3.378	2.237	2.295
[-2.59, -2.18]	0.151	1.412	4.036	0.717	0.186	0.904	1.390	1.141	1.385	2.789
[-2.18, -1.86]	0.369	0.841	2.339	0.640	0.520	0.828	0.839	0.081	0.810	1.425
[-1.86, -1.58]	0.629	0.749	3.105	1.670	1.440	0.527	1.369	1.183	0.724	0.375
[-1.58, -1.29]	0.784	0.750	2.799	1.408	1.291	0.399	1.604	0.706	0.379	0.500
[-1.29, -1.05]	0.709	0.607	1.696	0.568	0.517	0.597	0.419	0.960	0.312	0.604
[-1.05, -0.81]	0.508	1.070	3.443	1.140	1.150	1.038	1.143	0.572	0.590	2.242
[-0.81, -0.61]	0.315	1.960	6.917	2.648	2.136	1.616	2.544	1.956	0.757	4.352
[-0.61, 0.00]	0.057	2.871	8.053	1.292	2.468	2.674	1.761	1.152	2.760	5.445
$\ln(\lambda_{1,5}^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_{1,5}^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-5.00, -3.99]	0.010	2.489	14.137	7.035	5.404	5.577	5.227	4.817	1.854	5.467
[-3.99, -3.28]	0.052	1.755	5.929	2.495	1.510	1.007	2.365	0.570	1.516	3.787
[-3.28, -2.78]	0.167	1.100	3.832	1.063	1.073	1.068	0.590	1.042	1.074	2.731
[-2.78, -2.32]	0.354	0.759	2.256	0.941	1.305	0.532	0.312	0.875	0.637	0.609
[-2.32, -1.92]	0.535	0.735	2.873	1.565	1.378	0.209	1.514	0.498	0.655	0.598
[-1.92, -1.57]	0.569	0.624	1.690	0.572	0.516	0.552	0.972	0.293	0.392	0.621
[-1.57, -1.21]	0.449	0.789	2.226	0.499	0.944	0.765	0.206	1.156	0.789	0.781
[-1.21, -0.91]	0.283	1.659	5.811	2.316	1.985	1.358	2.248	1.588	0.256	3.496
[-0.91, 0.00]	0.052	2.678	7.593	1.015	2.240	2.524	1.613	1.072	2.542	5.278
$\ln(\lambda_2^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_2^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-6.00, -4.61]	0.012	2.039	11.029	4.839	4.291	3.877	3.747	3.223	1.913	5.705
[-4.61, -3.76]	0.064	1.384	5.820	0.773	1.058	1.184	0.450	1.376	1.380	4.984
[-3.76, -3.09]	0.188	0.897	2.421	0.724	0.575	0.871	0.885	0.099	0.722	1.460
[-3.09, -2.55]	0.355	0.716	2.745	1.263	1.808	0.303	0.865	0.839	0.643	0.449
[-2.55, -2.06]	0.466	0.622	2.002	1.002	0.340	0.453	1.158	0.622	0.464	0.596
[-2.06, -1.58]	0.427	0.597	1.880	0.565	0.195	0.597	0.587	1.170	0.588	0.638
[-1.58, -1.15]	0.279	1.346	4.385	1.814	1.388	1.135	1.620	0.989	1.003	2.511
[-1.15, 0.00]	0.051	2.485	6.921	0.944	1.929	2.328	1.459	1.289	2.427	4.692
$\sqrt{\lambda_0^2}$	$1/\sigma \, d\sigma/d\sqrt{\lambda_0^2}$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[0.00, 0.32]	0.322	1.336	5.555	1.335	0.810	1.963	1.212	4.160	1.815	0.844
[0.32, 0.39]	3.761	0.599	1.486	0.272	0.183	0.694	0.224	0.859	0.296	0.621
[0.39, 0.46]	3.875	0.353	2.429	0.427	0.295	0.308	0.197	0.329	1.931	1.239
[0.46, 0.55]	2.296	0.553	1.531	0.501	0.552	0.544	0.550	0.168	0.456	0.805
[0.55, 0.64]	1.011	1.062	2.868	1.042	1.042	0.866	0.920	0.685	1.628	0.459
[0.64, 0.74]	0.415	1.563	7.673	2.566	2.178	3.974	1.441	4.361	2.434	1.514
[0.74, 0.85]	0.152	2.551	15.804	6.838	2.984	8.837	3.763	9.109	2.539	2.436
[0.85, 0.94]	0.057	4.153	32.704	14.773	9.364	16.999	13.020	16.045	3.663	4.124
[0.94, 1.00]	0.024	5.688	49.586	21.621	13.182	24.299	22.638	21.964	11.637	8.026
$\bar{\lambda}_0^1$	$1/\sigma \, d\sigma/d\bar{\lambda}_0^1$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-1.00, -0.88]	0.012	4.945	30.010	15.368	6.793	13.754	9.789	16.375	4.940	4.041
[-0.88, -0.76]	0.034	3.307	13.156	6.265	1.598	5.088	2.438	8.461	3.302	2.451
[-0.76, -0.62]	0.082	2.142	5.746	1.401	2.138	2.053	0.339	3.181	2.075	1.771
[-0.62, -0.47]	0.191	1.325	4.767	0.979	1.606	1.440	3.015	1.211	1.323	1.748
[-0.47, -0.33]	0.411	0.786	4.584	1.544	1.585	2.288	2.898	0.990	0.809	0.478
[-0.33, -0.21]	0.745	0.553	2.226	0.796	0.428	0.996	1.438	0.579	0.542	0.375
[-0.21, -0.11]	1.134	0.488	1.242	0.478	0.487	0.488	0.393	0.344	0.368	0.438
[-0.11, -0.03]	1.474	0.433	1.435	0.388	0.086	0.240	0.411	0.427	0.430	1.058
[-0.03, 0.03]	1.688	0.337	1.630	0.585	0.887	0.796	0.460	0.227	0.699	0.167
[0.03, 0.11]	1.537	0.367	1.270	0.430	0.277	0.499	0.628	0.231	0.663	0.285
[0.11, 0.19]	1.280	0.388	1.041	0.131	0.332	0.265	0.530	0.359	0.356	0.445
[0.19, 0.29]	0.955	0.491	1.101	0.491	0.287	0.447	0.240	0.395	0.239	0.421
[0.29, 0.41]	0.610	0.716	1.635	0.344	0.718	0.291	0.694	0.127	0.711	0.661
[0.41, 0.56]	0.323	1.068	2.522	0.865	1.094	1.000	0.780	1.050	0.375	0.651
[0.56, 0.70]	0.145	1.563	4.131	1.508	1.537	1.246	0.858	0.959	1.294	2.259
[0.70, 0.84]	0.062	2.357	8.393	1.441	1.244	4.414	4.344	3.894	2.301	1.567
[0.84, 1.00]	0.020	3.775	23.469	6.677	6.895	13.110	11.793	10.396	3.408	3.717
$\bar{\lambda}_0^0$	$1/\sigma \, d\sigma/d\bar{\lambda}_0^0$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[1.00, 2.00]	0.028	1.839	7.749	1.503	1.802	4.571	1.504	3.082	3.129	2.952
[2.00, 3.00]	0.062	1.173	4.658	1.169	1.167	1.596	0.996	1.064	2.173	2.864
[3.00, 4.00]	0.114	0.798	4.048	0.541	0.761	0.685	0.191	0.598	2.080	3.113
[4.00, 5.00]	0.141	0.575	2.199	0.156	0.571	0.524	0.345	0.402	0.897	1.671
[5.00, 6.00]	0.156	0.384	0.964	0.378	0.362	0.328	0.356	0.208	0.339	0.337
[6.00, 7.00]	0.139	0.341	1.024	0.313	0.343	0.340	0.341	0.228	0.242	0.611
[7.00, 8.00]	0.116	0.459	2.159	0.389	0.214	0.347	0.425	0.445	0.980	1.672
[8.00, 9.00]	0.085	0.645	3.522	0.609	0.482	0.337	0.559	0.322	1.808	2.754
[9.00, 10.00]	0.060	0.877	5.289	0.734	0.850	0.373	0.499	0.588	2.958	4.056
[10.00, 11.00]	0.040	1.137	5.345	1.070	0.333	1.125	0.879	1.928	2.262	3.893
[11.00, 12.00]	0.026	1.430	4.706	1.219	0.203	1.173	0.609	2.796	0.741	2.908
[12.00, 13.00]	0.016	1.738	6.529	1.732	0.799	2.490	1.509	4.521	1.737	2.009
[13.00, 14.00]	0.010	2.075	6.515	1.945	0.713	2.390	1.280	4.835	1.669	0.575
[14.00, 15.00]	0.006	2.390	8.898	2.375	1.335	3.504	1.947	6.307	2.316	2.197

**Table C.5**Measured values of the normalized unfolded differential cross sections and uncertainties for each systematic source considered in the range  $150 < Q^2 < 360 \text{ GeV}^2$ .

$\ln(\lambda_1^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_1^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-4.00, -3.15]	0.004	3.21	26.30	11.99	8.20	14.77	11.96	9.95	2.53	1.91
[-3.15, -2.59]	0.024	2.11	14.28	5.73	6.77	7.64	5.56	4.86	2.00	1.97
[-2.59, -2.18]	0.101	1.41	5.23	1.93	0.84	3.28	2.83	0.86	1.03	0.65
[-2.18, -1.86]	0.292	0.96	3.10	0.96	0.95	0.27	0.50	0.88	2.24	0.85
[-1.86, -1.58]	0.582	0.77	2.51	0.81	0.90	0.70	0.64	0.37	1.72	0.49
[-1.58, -1.29]	0.825	0.69	2.24	1.05	1.08	0.11	1.03	0.68	0.45	0.74
[-1.29, -1.05]	0.799	0.55	2.24	0.15	0.53	0.43	0.67	0.57	1.35	1.27
[-1.05, -0.81]	0.580	0.94	2.44	0.76	0.75	0.72	0.90	0.92	0.94	0.93
[-0.81, -0.61]	0.364	1.83	4.66	1.31	1.34	1.78	1.02	1.78	1.79	2.06
[-0.61, 0.00]	0.070	2.85	6.12	0.62	2.70	2.81	2.78	1.28	1.02	1.86
$\ln(\lambda_{1,5}^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_{1,5}^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-5.00, -3.99]	0.007	2.52	17.46	6.58	7.02	8.97	6.08	4.75	4.17	6.96
[-3.99, -3.28]	0.034	1.67	8.62	3.28	2.74	4.86	3.82	2.47	1.66	2.48
[-3.28, -2.78]	0.124	1.16	3.41	0.64	0.78	1.69	1.79	1.14	0.71	1.17
[-2.78, -2.32]	0.304	0.85	3.11	0.42	0.86	0.84	0.81	0.23	2.56	0.18
[-2.32, -1.92]	0.528	0.72	1.92	0.98	0.98	0.49	0.46	0.49	0.19	0.72
[-1.92, -1.57]	0.618	0.57	1.58	0.65	0.41	0.36	1.01	0.19	0.36	0.51
[-1.57, -1.21]	0.510	0.69	1.76	0.67	0.45	0.33	0.65	0.46	0.89	0.65
[-1.21, -0.91]	0.328	1.51	3.72	0.99	0.99	1.50	0.84	1.51	1.47	1.49
[-0.91, 0.00]	0.064	2.63	5.76	1.01	2.44	2.62	2.57	1.30	1.14	1.67
$\ln(\lambda_2^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_2^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-6.00, -4.61]	0.008	2.08	14.61	5.60	7.21	7.61	5.49	4.29	0.64	4.35
[-4.61, -3.76]	0.046	1.38	6.10	0.14	0.97	2.28	1.90	1.29	1.16	4.75
[-3.76, -3.09]	0.151	0.99	3.46	0.99	0.99	0.43	0.84	0.94	2.53	0.94
[-3.09, -2.55]	0.326	0.78	3.05	0.83	1.86	0.71	0.68	0.44	1.73	0.62
[-2.55, -2.06]	0.476	0.60	1.56	0.62	0.54	0.42	0.58	0.54	0.60	0.47
[-2.06, -1.58]	0.473	0.53	1.51	0.42	0.45	0.46	0.57	0.68	0.60	0.51
[-1.58, -1.15]	0.324	1.20	2.51	0.45	0.41	1.20	0.95	1.15	0.77	0.51
[-1.15, 0.00]	0.063	2.42	5.37	0.85	2.18	2.41	2.34	0.79	1.54	1.78
$\sqrt{\lambda_0^2}$	$1/\sigma \, d\sigma/d\sqrt{\lambda_0^2}$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[0.00, 0.32]	0.276	1.43	6.84	0.27	2.61	2.66	1.26	4.93	1.39	1.72
[0.32, 0.39]	3.632	0.60	1.56	0.21	0.60	0.75	0.22	0.53	0.42	0.78
[0.39, 0.46]	3.949	0.32	2.04	0.26	0.42	0.32	0.28	0.29	1.39	1.28
[0.46, 0.55]	2.403	0.53	1.42	0.49	0.08	0.53	0.52	0.44	0.53	0.68
[0.55, 0.64]	1.071	0.97	3.02	0.97	0.82	0.90	0.94	0.83	2.02	0.27
[0.64, 0.74]	0.439	1.44	7.29	2.52	2.47	3.91	0.53	3.45	3.04	1.38
[0.74, 0.85]	0.160	2.34	15.51	6.65	4.22	8.77	5.10	7.89	2.02	1.99
[0.85, 0.94]	0.059	3.79	34.15	15.03	11.75	17.26	15.01	15.61	2.65	3.71
[0.94, 1.00]	0.025	5.14	48.64	20.24	15.04	22.85	23.17	19.35	12.84	10.28
$\tilde{\lambda}_0^1$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^1$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-1.00, -0.88]	0.013	4.29	27.48	12.72	8.62	12.57	9.19	15.17	4.01	3.45
[-0.88, -0.76]	0.037	2.93	13.46	6.40	0.70	6.31	2.72	8.88	0.30	2.25
[-0.76, -0.62]	0.088	1.97	5.36	1.03	1.60	0.58	1.94	3.70	1.30	1.31
[-0.62, -0.47]	0.204	1.34	3.60	1.29	1.76	0.82	1.60	1.07	1.34	0.45
[-0.47, -0.33]	0.432	0.84	3.90	1.22	1.65	1.91	2.36	0.72	0.76	0.07
[-0.33, -0.21]	0.771	0.59	1.56	0.33	0.58	0.47	0.76	0.43	0.56	0.58
[-0.21, -0.11]	1.147	0.52	1.34	0.48	0.43	0.50	0.52	0.52	0.20	0.52
[-0.11, -0.03]	1.473	0.46	1.62	0.34	0.15	0.26	0.40	0.46	0.21	1.34
[-0.03, 0.03]	1.669	0.34	1.87	0.65	1.38	0.80	0.43	0.19	0.10	0.43
[0.03, 0.11]	1.501	0.39	1.06	0.22	0.35	0.32	0.28	0.33	0.62	0.36
[0.11, 0.19]	1.251	0.41	1.15	0.37	0.38	0.38	0.25	0.36	0.15	0.71
[0.19, 0.29]	0.934	0.53	1.13	0.31	0.22	0.14	0.51	0.20	0.53	0.50
[0.29, 0.41]	0.600	0.78	2.03	0.55	0.82	0.67	0.71	0.76	0.65	0.76
[0.41, 0.56]	0.319	1.16	3.03	0.42	1.72	0.78	1.14	0.85	1.08	0.96
[0.56, 0.70]	0.144	1.59	5.22	1.40	1.49	1.51	0.87	1.45	1.52	3.61
[0.70, 0.84]	0.062	2.27	8.02	1.85	1.81	3.78	4.96	2.06	2.27	2.04
[0.84, 1.00]	0.019	3.55	23.22	6.19	7.92	12.31	13.36	8.59	3.27	3.33
$\tilde{\lambda}_0^0$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^0$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[1.00, 2.00]	0.032	1.71	8.77	2.22	1.71	4.57	1.52	2.45	3.63	4.87
[2.00, 3.00]	0.071	1.06	5.83	1.06	0.99	1.52	1.02	0.72	2.65	4.45
[3.00, 4.00]	0.128	0.69	5.11	0.41	0.59	0.39	0.69	0.64	2.35	4.31
[4.00, 5.00]	0.153	0.51	2.51	0.49	0.40	0.49	0.32	0.40	0.38	2.23
[5.00, 6.00]	0.162	0.39	1.23	0.28	0.23	0.42	0.25	0.65	0.65	0.36
[6.00, 7.00]	0.140	0.39	2.51	0.28	1.06	0.35	0.38	0.59	1.42	1.53
[7.00, 8.00]	0.111	0.50	3.45	0.47	0.88	0.23	0.43	0.45	1.68	2.72
[8.00, 9.00]	0.078	0.70	5.65	0.68	0.69	1.02	0.58	0.18	2.64	4.69
[9.00, 10.00]	0.052	0.98	7.63	0.74	0.21	1.78	0.73	1.40	3.15	6.40
[10.00, 11.00]	0.032	1.32	8.31	0.35	3.22	2.78	0.68	3.01	0.67	6.26
[11.00, 12.00]	0.020	1.71	8.87	0.74	4.21	3.53	0.93	4.63	1.14	4.62
[12.00, 13.00]	0.012	2.16	12.07	2.60	4.36	5.34	1.20	6.93	6.05	0.85
[13.00, 14.00]	0.007	2.65	14.30	2.87	6.60	5.94	0.77	7.49	6.03	4.18
[14.00, 15.00]	0.004	3.17	19.73	4.34	6.80	7.61	2.39	9.98	11.90	3.10

**Table C.6**Measured values of the normalized unfolded differential cross sections and uncertainties for each systematic source considered in the range  $360 < Q^2 < 866 \text{ GeV}^2$ .

$\ln(\lambda_1^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_1^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-4.00, -3.15]	0.006	3.66	26.20	12.04	5.86	14.89	9.92	12.29	3.90	2.67
[-3.15, -2.59]	0.036	2.31	8.89	3.86	0.81	5.54	2.33	4.48	1.41	0.12
[-2.59, -2.18]	0.160	1.48	3.34	1.41	1.38	0.60	1.40	1.13	1.05	0.62
[-2.18, -1.86]	0.423	0.92	1.90	0.62	0.78	0.87	0.52	0.07	0.41	0.77
[-1.86, -1.58]	0.704	0.80	4.36	2.04	1.51	1.08	2.04	1.38	1.61	1.44
[-1.58, -1.29]	0.792	0.82	2.80	1.27	1.21	0.19	1.72	0.76	0.69	0.16
[-1.29, -1.05]	0.648	0.72	2.23	0.44	0.56	0.70	0.61	1.31	1.05	0.53
[-1.05, -0.81]	0.449	1.37	5.70	2.60	2.74	1.21	2.76	0.90	2.01	1.52
[-0.81, -0.61]	0.276	2.39	9.72	3.68	3.15	2.02	4.59	2.39	3.55	4.67
[-0.61, 0.00]	0.047	3.34	11.48	1.35	1.48	2.99	3.84	2.42	3.95	8.45
$\ln(\lambda_{1,5}^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_{1,5}^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-5.00, -3.99]	0.009	2.80	14.70	6.54	3.29	8.85	4.64	7.21	0.66	1.53
[-3.99, -3.28]	0.052	1.80	5.10	0.72	1.67	2.24	1.21	0.64	1.98	2.95
[-3.28, -2.78]	0.184	1.17	2.82	0.79	0.58	1.17	1.04	1.03	0.94	1.10
[-2.78, -2.32]	0.400	0.82	2.94	1.33	1.53	0.36	1.02	0.85	0.80	1.14
[-2.32, -1.92]	0.570	0.78	3.12	1.65	1.18	0.69	1.89	0.62	0.76	0.16
[-1.92, -1.57]	0.550	0.70	1.75	0.50	0.62	0.65	0.59	0.46	0.70	0.69
[-1.57, -1.21]	0.403	1.01	3.89	1.88	1.92	0.82	1.78	1.33	0.99	0.56
[-1.21, -0.91]	0.247	2.07	8.27	3.52	3.39	1.72	4.13	2.02	2.49	3.15
[-0.91, 0.00]	0.042	3.15	11.23	1.71	0.63	2.80	3.91	2.37	4.28	8.11
$\ln(\lambda_2^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_2^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-6.00, -4.61]	0.012	2.24	8.58	2.64	1.56	4.72	0.94	3.37	2.71	4.18
[-4.61, -3.76]	0.067	1.45	5.33	1.45	1.42	0.98	1.44	1.43	2.22	3.48
[-3.76, -3.09]	0.207	0.96	2.85	1.09	1.66	0.77	0.43	0.80	0.69	1.17
[-3.09, -2.55]	0.388	0.75	2.56	1.31	1.41	0.26	1.27	0.71	0.25	0.12
[-2.55, -2.06]	0.477	0.66	2.14	1.00	0.47	0.04	1.43	0.66	0.44	0.49
[-2.06, -1.58]	0.399	0.73	2.16	0.81	0.74	0.63	0.32	1.50	0.23	0.36
[-1.58, -1.15]	0.245	1.71	6.83	3.11	2.73	1.38	3.40	1.41	2.15	2.56
[-1.15, 0.00]	0.042	2.96	10.01	1.77	1.64	2.64	3.76	2.32	3.35	6.92
$\sqrt{\lambda_0^2}$	$1/\sigma \, d\sigma/d\sqrt{\lambda_0^2}$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[0.00, 0.32]	0.334	1.49	7.95	1.46	1.01	1.96	1.49	4.32	5.18	2.51
[0.32, 0.39]	3.847	0.77	2.79	0.30	1.27	0.60	0.57	1.02	1.59	1.11
[0.39, 0.46]	3.870	0.47	3.31	0.59	0.46	0.22	0.41	0.35	2.76	1.49
[0.46, 0.55]	2.242	0.72	2.73	0.72	0.37	0.67	0.62	0.42	1.86	1.35
[0.55, 0.64]	0.976	1.43	3.35	1.40	0.61	1.24	0.93	1.44	1.04	1.13
[0.64, 0.74]	0.402	2.10	8.88	2.68	2.59	4.39	2.09	5.65	1.62	1.50
[0.74, 0.85]	0.149	3.30	18.46	7.80	2.46	10.30	3.85	11.20	3.19	2.55
[0.85, 0.94]	0.058	5.18	35.41	16.05	6.62	19.16	14.35	17.76	3.30	5.12
[0.94, 1.00]	0.025	6.93	56.26	24.31	11.43	28.18	26.18	25.75	13.08	9.04
$\tilde{\lambda}_0^1$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^1$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-1.00, -0.88]	0.012	6.40	36.35	19.74	5.44	15.96	14.10	18.59	5.55	5.60
[-0.88, -0.76]	0.034	4.39	14.90	7.00	2.66	3.76	3.77	8.75	3.55	5.38
[-0.76, -0.62]	0.083	2.79	6.82	2.52	2.73	2.57	2.44	1.83	1.46	2.61
[-0.62, -0.47]	0.189	1.62	6.08	0.92	0.60	3.06	4.31	1.62	0.80	1.39
[-0.47, -0.33]	0.412	0.92	5.68	1.47	0.81	2.89	3.21	0.98	2.96	0.52
[-0.33, -0.21]	0.742	0.65	2.90	0.99	0.49	1.45	1.76	0.92	0.54	0.65
[-0.21, -0.11]	1.139	0.56	1.74	0.50	0.14	0.52	0.34	0.61	0.39	1.24
[-0.11, -0.03]	1.476	0.50	1.77	0.28	1.11	0.37	0.39	0.49	0.16	1.02
[-0.03, 0.03]	1.688	0.39	2.37	0.73	0.26	1.12	0.62	0.62	1.64	0.39
[0.03, 0.11]	1.547	0.41	1.56	0.51	0.27	0.74	0.74	0.15	0.82	0.37
[0.11, 0.19]	1.280	0.46	1.44	0.32	0.42	0.33	0.54	0.36	0.94	0.41
[0.19, 0.29]	0.954	0.60	1.43	0.60	0.50	0.60	0.17	0.50	0.30	0.60
[0.29, 0.41]	0.608	0.87	2.45	0.25	1.32	0.17	0.84	0.51	0.83	1.32
[0.41, 0.56]	0.321	1.27	3.55	1.14	1.77	1.27	0.17	1.25	0.74	1.68
[0.56, 0.70]	0.144	1.92	4.78	1.67	1.70	1.34	0.93	1.28	1.83	2.42
[0.70, 0.84]	0.061	2.94	11.96	1.81	3.82	6.46	5.16	6.03	2.86	1.90
[0.84, 1.00]	0.020	4.60	27.36	7.15	6.11	16.75	13.17	12.42	3.33	4.39
$\tilde{\lambda}_0^0$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^0$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[1.00, 2.00]	0.025	2.38	8.55	1.27	1.55	4.67	2.32	3.44	4.13	2.70
[2.00, 3.00]	0.057	1.54	5.82	1.52	1.23	1.59	1.18	1.95	3.18	3.13
[3.00, 4.00]	0.108	1.03	5.05	1.04	1.02	1.02	0.73	0.83	3.00	3.32
[4.00, 5.00]	0.137	0.69	3.81	1.10	0.57	0.49	0.83	0.67	2.41	2.31
[5.00, 6.00]	0.155	0.45	1.50	0.25	0.21	0.43	0.37	0.45	0.98	0.67
[6.00, 7.00]	0.142	0.42	1.11	0.42	0.42	0.41	0.41	0.42	0.13	0.41
[7.00, 8.00]	0.121	0.58	2.29	0.63	0.58	0.52	0.74	0.40	0.99	1.50
[8.00, 9.00]	0.090	0.78	3.51	0.29	1.19	0.76	0.55	0.77	1.98	2.20
[9.00, 10.00]	0.063	1.01	7.86	0.91	0.99	1.00	1.00	0.97	5.39	5.19
[10.00, 11.00]	0.042	1.27	8.34	0.51	1.25	0.88	0.75	0.77	5.74	5.59
[11.00, 12.00]	0.027	1.56	10.70	1.08	1.51	0.77	1.28	2.36	7.04	7.16
[12.00, 13.00]	0.017	1.88	10.20	1.77	1.59	1.44	1.85	3.81	5.20	6.91
[13.00, 14.00]	0.010	2.20	10.86	2.20	1.16	2.36	2.19	5.04	4.28	7.27
[14.00, 15.00]	0.006	2.53	9.72	2.26	3.92	4.04	2.08	6.59	1.31	1.29

**Table C.7**Measured values of the normalized unfolded differential cross sections and uncertainties for each systematic source considered in the range  $866 < Q^2 < 2080 \text{ GeV}^2$ .

$\ln(\lambda_1^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_1^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-4.00, -3.15]	0.008	3.17	12.73	6.62	3.15	3.86	2.56	8.49	1.32	1.75
[-3.15, -2.59]	0.069	2.29	5.70	1.77	1.00	2.12	2.19	2.34	2.17	1.91
[-2.59, -2.18]	0.282	1.54	3.75	1.53	0.23	1.04	1.12	1.54	1.53	1.49
[-2.18, -1.86]	0.564	0.97	3.40	1.43	0.82	1.00	2.19	0.78	0.74	0.97
[-1.86, -1.58]	0.694	0.93	6.42	3.12	2.45	1.45	3.80	2.24	1.76	0.21
[-1.58, -1.29]	0.647	1.01	4.69	2.28	2.44	0.87	2.38	0.45	1.62	0.76
[-1.29, -1.05]	0.518	1.14	3.58	0.24	0.42	1.01	1.82	1.08	0.37	2.37
[-1.05, -0.81]	0.366	1.89	12.67	6.03	4.70	3.28	7.96	4.13	1.95	1.89
[-0.81, -0.61]	0.223	2.71	18.11	8.65	4.46	2.77	11.39	1.34	8.08	4.63
[-0.61, 0.00]	0.033	3.37	25.76	7.48	3.33	2.06	12.11	3.16	15.70	13.34
$\ln(\lambda_{1,5}^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_{1,5}^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-5.00, -3.99]	0.015	2.57	7.51	3.85	1.06	1.11	2.39	4.88	1.16	1.35
[-3.99, -3.28]	0.093	1.91	4.98	1.84	1.49	1.58	1.65	1.73	1.91	1.91
[-3.28, -2.78]	0.280	1.27	3.09	0.84	1.12	0.61	1.16	1.03	1.27	1.27
[-2.78, -2.32]	0.466	0.86	4.28	2.05	1.25	1.13	2.83	1.32	0.24	0.86
[-2.32, -1.92]	0.518	0.93	6.26	2.99	3.22	0.87	3.45	1.69	1.87	0.08
[-1.92, -1.57]	0.445	1.00	2.87	0.86	0.96	0.56	1.00	0.90	1.19	1.43
[-1.57, -1.21]	0.323	1.57	9.10	4.05	3.27	2.55	5.71	3.42	1.28	0.82
[-1.21, -0.91]	0.196	2.50	17.04	8.48	4.90	3.61	10.98	3.00	6.18	2.71
[-0.91, 0.00]	0.029	3.24	24.81	7.17	2.94	2.01	11.59	3.08	15.29	12.79
$\ln(\lambda_2^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_2^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-6.00, -4.61]	0.019	2.19	5.17	1.87	1.60	2.01	2.18	2.38	0.92	0.74
[-4.61, -3.76]	0.105	1.63	3.69	1.52	1.38	0.39	0.57	1.61	1.50	1.20
[-3.76, -3.09]	0.275	1.05	3.56	1.28	1.04	1.28	2.26	0.88	0.55	1.00
[-3.09, -2.55]	0.414	0.78	4.96	2.35	2.14	1.00	3.04	1.30	1.16	0.75
[-2.55, -2.06]	0.423	0.84	3.24	1.68	1.52	0.81	1.79	0.16	0.21	0.84
[-2.06, -1.58]	0.323	1.23	5.28	1.83	1.70	1.71	3.13	2.17	1.17	1.19
[-1.58, -1.15]	0.192	2.25	14.74	7.50	4.34	3.66	9.80	3.60	3.49	1.57
[-1.15, 0.00]	0.028	3.11	23.68	6.71	2.45	1.90	11.05	2.99	14.73	12.19
$\sqrt{\lambda_0^2}$	$1/\sigma \, d\sigma/d\sqrt{\lambda_0^2}$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[0.00, 0.32]	0.442	1.70	7.05	0.40	1.70	1.46	1.08	2.08	5.66	2.05
[0.32, 0.39]	4.084	0.81	2.13	0.91	0.27	0.55	0.36	1.24	0.80	0.60
[0.39, 0.46]	3.667	0.53	3.84	1.40	0.79	0.30	0.27	0.30	3.17	1.27
[0.46, 0.55]	2.034	0.92	3.28	0.90	0.66	0.92	0.82	0.50	2.05	1.64
[0.55, 0.64]	0.865	1.63	4.74	1.27	1.75	0.58	1.40	2.99	1.29	1.50
[0.64, 0.74]	0.355	2.33	9.31	3.79	2.28	3.26	2.33	6.30	1.44	1.93
[0.74, 0.85]	0.130	3.58	17.10	7.45	3.57	6.31	3.54	11.13	5.11	3.02
[0.85, 0.94]	0.047	5.40	25.01	13.09	4.73	11.20	2.97	15.86	2.13	3.50
[0.94, 1.00]	0.020	7.46	50.24	24.52	18.27	23.18	14.86	26.30	5.82	7.01
$\tilde{\lambda}_0^1$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^1$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-1.00, -0.88]	0.009	7.62	30.62	20.47	6.72	11.66	4.92	13.67	5.56	6.11
[-0.88, -0.76]	0.026	5.02	12.57	6.63	5.01	4.38	2.95	5.38	2.45	0.89
[-0.76, -0.62]	0.065	3.22	8.18	2.64	3.22	2.11	4.84	1.24	0.86	3.00
[-0.62, -0.47]	0.156	1.93	6.91	1.80	1.15	3.14	3.88	1.92	1.19	3.08
[-0.47, -0.33]	0.353	1.28	5.49	1.40	1.34	3.30	3.11	1.53	0.70	1.16
[-0.33, -0.21]	0.683	0.92	4.00	1.50	0.43	2.05	2.11	1.70	0.91	0.63
[-0.21, -0.11]	1.098	0.70	1.64	0.51	0.66	0.56	0.38	0.62	0.43	0.68
[-0.11, -0.03]	1.483	0.52	1.60	0.47	0.81	0.37	0.14	0.62	0.78	0.49
[-0.03, 0.03]	1.749	0.41	1.36	0.40	0.44	0.35	0.41	0.32	0.88	0.41
[0.03, 0.11]	1.624	0.40	2.27	0.58	0.81	0.96	0.99	0.19	1.37	0.40
[0.11, 0.19]	1.355	0.50	1.61	0.02	0.48	0.65	0.92	0.42	0.66	0.48
[0.19, 0.29]	1.006	0.71	2.58	0.34	0.53	0.36	1.17	0.55	1.51	1.29
[0.29, 0.41]	0.637	1.06	3.04	1.06	0.90	1.03	0.45	1.02	1.78	0.84
[0.41, 0.56]	0.336	1.60	5.09	1.56	0.75	1.30	0.36	1.13	4.11	0.55
[0.56, 0.70]	0.151	2.41	5.74	1.68	2.09	1.38	0.97	2.96	1.76	2.28
[0.70, 0.84]	0.061	3.53	10.31	2.42	2.63	3.96	2.06	6.62	2.34	3.44
[0.84, 1.00]	0.019	5.29	23.38	8.50	6.77	10.03	2.90	15.40	5.16	5.29
$\tilde{\lambda}_0^0$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^0$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[1.00, 2.00]	0.017	2.60	7.82	1.97	2.55	1.73	1.66	5.41	2.44	1.74
[2.00, 3.00]	0.040	1.93	5.65	1.25	1.47	1.55	0.68	3.82	1.87	1.89
[3.00, 4.00]	0.080	1.43	3.82	1.38	1.27	1.23	1.20	0.53	1.37	1.97
[4.00, 5.00]	0.109	1.05	2.89	0.79	0.86	0.84	1.11	1.05	0.82	1.47
[5.00, 6.00]	0.138	0.73	2.57	1.31	1.41	1.00	0.63	0.59	0.61	0.50
[6.00, 7.00]	0.138	0.57	3.96	0.88	0.66	0.43	0.43	0.54	3.00	2.11
[7.00, 8.00]	0.131	0.61	2.06	0.93	0.56	0.52	0.35	0.60	1.28	0.54
[8.00, 9.00]	0.106	0.75	2.48	0.67	0.70	0.66	0.73	0.02	0.74	1.76
[9.00, 10.00]	0.083	1.02	2.72	0.46	0.71	0.55	0.63	1.00	1.01	1.71
[10.00, 11.00]	0.059	1.40	4.37	0.81	1.35	0.90	1.23	1.21	1.14	3.10
[11.00, 12.00]	0.042	1.84	6.56	3.36	2.10	1.45	3.87	1.71	0.80	1.82
[12.00, 13.00]	0.027	2.29	10.73	3.15	1.93	1.91	3.34	2.04	6.06	6.38
[13.00, 14.00]	0.018	2.74	9.16	4.24	2.76	2.18	4.62	2.56	3.14	2.89
[14.00, 15.00]	0.012	3.13	17.35	4.30	2.84	2.94	5.12	2.04	10.96	10.28



**Table C.8**

Measured values of the normalized unfolded differential cross sections and uncertainties for each systematic source considered in the range  $2080 < Q^2 < 5000 \text{ GeV}^2$ .

$\ln(\lambda_1^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_1^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-4.00, -3.15]	0.017	3.45	13.80	2.86	5.17	2.29	2.84	3.00	6.13	9.14
[-3.15, -2.59]	0.155	2.41	8.28	2.25	0.96	1.16	2.20	2.01	3.17	6.03
[-2.59, -2.18]	0.435	1.54	4.67	1.19	1.53	1.11	1.42	1.53	0.75	3.09
[-2.18, -1.86]	0.572	1.15	5.07	2.13	2.55	0.96	3.19	0.71	0.65	1.15
[-1.86, -1.58]	0.577	1.10	11.81	3.96	2.66	1.44	4.26	2.98	5.89	7.18
[-1.58, -1.29]	0.521	1.24	11.20	1.10	0.67	0.90	0.61	0.42	5.74	9.37
[-1.29, -1.05]	0.421	1.84	9.32	1.92	2.43	2.90	4.27	1.63	0.87	6.62
[-1.05, -0.81]	0.302	2.80	15.61	6.72	3.85	4.05	9.79	3.92	6.90	0.93
[-0.81, -0.61]	0.180	3.78	27.77	11.51	2.51	4.50	14.21	3.83	16.64	10.21
[-0.61, 0.00]	0.024	4.40	35.10	12.68	4.40	3.44	16.23	3.47	21.64	16.65
$\ln(\lambda_{1,5}^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_{1,5}^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-5.00, -3.99]	0.031	2.86	7.77	2.78	1.62	1.08	2.11	2.47	1.67	5.24
[-3.99, -3.28]	0.173	2.01	5.50	1.98	1.59	0.62	1.29	1.91	0.29	3.73
[-3.28, -2.78]	0.371	1.35	3.94	0.66	1.14	1.31	2.13	1.19	1.18	1.70
[-2.78, -2.32]	0.444	1.06	6.40	3.01	2.85	1.42	3.76	1.74	1.81	0.38
[-2.32, -1.92]	0.429	1.08	11.82	3.01	1.72	0.86	2.88	2.17	6.25	8.59
[-1.92, -1.57]	0.359	1.50	8.89	1.01	1.03	2.40	1.86	1.07	2.91	7.48
[-1.57, -1.21]	0.261	2.48	13.07	5.20	5.10	3.98	8.31	3.77	2.70	2.30
[-1.21, -0.91]	0.155	3.59	25.45	10.12	0.45	3.92	13.03	3.39	15.70	9.43
[-0.91, 0.00]	0.020	4.30	33.92	12.06	4.25	3.97	15.71	3.69	20.90	15.98
$\ln(\lambda_2^1)$	$1/\sigma \, d\sigma/d\ln(\lambda_2^1)$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-6.00, -4.61]	0.034	2.49	5.25	2.35	1.09	0.75	1.72	1.99	1.83	1.98
[-4.61, -3.76]	0.164	1.76	4.80	1.72	1.54	1.10	1.19	1.76	1.72	2.43
[-3.76, -3.09]	0.331	1.19	3.85	1.58	0.30	1.42	2.68	0.46	1.10	0.43
[-3.09, -2.55]	0.387	0.94	6.57	2.92	2.95	0.93	3.41	1.87	2.38	1.83
[-2.55, -2.06]	0.354	1.11	10.15	1.16	1.10	0.93	0.71	0.39	5.36	8.31
[-2.06, -1.58]	0.256	2.05	9.49	3.57	2.46	3.78	6.25	3.04	1.70	1.23
[-1.58, -1.15]	0.148	3.34	21.48	8.40	3.67	3.45	11.36	2.84	12.83	7.25
[-1.15, 0.00]	0.019	4.21	33.32	12.50	4.21	5.42	16.01	2.96	19.92	15.07
$\sqrt{\lambda_0^2}$	$1/\sigma \, d\sigma/d\sqrt{\lambda_0^2}$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[0.00, 0.32]	0.569	2.28	5.37	0.14	2.24	2.15	1.64	1.99	1.51	2.23
[0.32, 0.39]	4.111	0.93	3.18	0.31	0.65	0.92	0.70	0.59	1.40	2.26
[0.39, 0.46]	3.413	0.64	1.46	0.40	0.59	0.64	0.43	0.54	0.49	0.30
[0.46, 0.55]	1.888	1.33	3.73	1.28	1.82	1.29	1.29	0.97	1.11	1.32
[0.55, 0.64]	0.801	2.36	5.81	2.33	1.92	2.36	2.33	1.27	2.29	1.10
[0.64, 0.74]	0.320	3.39	10.89	3.27	3.32	3.37	2.63	2.82	4.26	6.41
[0.74, 0.85]	0.117	4.68	10.32	4.26	4.54	4.62	2.86	2.41	2.49	2.08
[0.85, 0.94]	0.041	6.21	18.54	4.95	4.33	6.06	3.78	6.05	6.91	11.24
[0.94, 1.00]	0.014	7.57	38.72	16.40	25.48	4.59	7.14	18.93	6.09	7.50
$\tilde{\lambda}_0^1$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^1$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[-1.00, -0.88]	0.004	8.03	70.50	21.97	12.77	9.95	6.15	21.97	53.11	28.63
[-0.88, -0.76]	0.018	5.78	20.95	1.78	9.85	7.82	10.15	5.11	10.20	3.36
[-0.76, -0.62]	0.046	4.17	17.57	3.51	14.35	1.98	4.43	4.16	4.72	3.13
[-0.62, -0.47]	0.117	2.94	6.72	1.81	2.04	1.85	2.31	2.85	2.49	2.44
[-0.47, -0.33]	0.283	1.96	4.74	1.19	2.31	1.88	1.00	1.14	1.26	2.11
[-0.33, -0.21]	0.585	1.32	5.04	2.09	0.42	1.37	2.33	1.74	2.21	1.98
[-0.21, -0.11]	1.012	0.96	2.29	0.78	0.78	0.76	0.36	0.98	0.76	0.92
[-0.11, -0.03]	1.466	0.71	2.80	0.44	2.36	0.66	0.23	0.21	0.70	0.73
[-0.03, 0.03]	1.792	0.51	1.21	0.46	0.49	0.33	0.40	0.45	0.20	0.49
[0.03, 0.11]	1.746	0.48	3.14	0.31	0.70	0.44	0.45	0.39	1.44	2.53
[0.11, 0.19]	1.476	0.59	2.32	0.82	1.44	0.59	1.25	0.36	0.50	0.04
[0.19, 0.29]	1.098	0.80	3.47	1.34	2.05	1.05	1.63	0.83	0.61	0.77
[0.29, 0.41]	0.687	1.25	4.21	1.09	3.03	0.66	1.22	0.29	1.86	0.57
[0.41, 0.56]	0.352	2.13	5.36	1.95	2.07	2.11	2.08	1.76	2.01	0.44
[0.56, 0.70]	0.162	3.39	8.46	2.71	2.41	3.39	3.33	2.79	3.21	2.51
[0.70, 0.84]	0.064	4.82	13.93	4.59	4.61	4.70	4.13	3.50	3.81	7.90
[0.84, 1.00]	0.018	6.50	15.66	5.17	6.27	6.47	4.42	4.19	4.59	6.08
$\tilde{\lambda}_0^0$	$1/\sigma \, d\sigma/d\tilde{\lambda}_0^0$	Stat. (%)	Tot. (%)	HFS (jet) (%)	HFS (other) (%)	HFS ( $\phi$ ) (%)	Lepton (E) (%)	Lepton ( $\phi$ ) (%)	Model (%)	Closure (%)
[1.00, 2.00]	0.011	3.16	7.28	1.62	3.15	3.08	3.42	2.62	1.50	0.37
[2.00, 3.00]	0.027	2.59	6.54	1.21	2.29	2.56	2.10	3.65	1.24	1.86
[3.00, 4.00]	0.059	2.08	5.60	1.76	2.83	1.79	0.82	1.97	1.99	2.05
[4.00, 5.00]	0.085	1.58	5.15	1.57	1.18	1.45	1.12	1.74	2.68	2.57
[5.00, 6.00]	0.114	1.15	2.62	1.15	0.56	0.99	1.10	0.93	0.56	0.72
[6.00, 7.00]	0.123	0.79	2.06	0.70	0.76	0.72	0.79	0.42	0.79	0.77
[7.00, 8.00]	0.128	0.63	1.52	0.63	0.61	0.32	0.42	0.42	0.63	0.55
[8.00, 9.00]	0.113	0.73	1.76	0.40	0.86	0.72	0.73	0.20	0.50	0.59
[9.00, 10.00]	0.099	0.98	2.81	0.32	0.98	0.97	0.98	1.44	0.98	0.98
[10.00, 11.00]	0.078	1.34	4.17	1.34	3.16	0.08	0.98	0.22	1.18	1.19
[11.00, 12.00]	0.061	1.82	3.70	1.78	1.06	1.44	0.72	0.22	1.43	1.18
[12.00, 13.00]	0.044	2.31	5.12	2.27	2.28	2.05	1.32	0.04	1.10	1.81
[13.00, 14.00]	0.034	2.93	6.91	2.46	2.90	2.52	1.90	1.00	2.28	2.93
[14.00, 15.00]	0.023	3.51	7.55	2.86	0.40	3.40	2.40	2.41	2.66	2.47

- [5] S. Chekanov, et al., ZEUS, Substructure dependence of jet cross sections at HERA and determination of alpha(s), Nucl. Phys. B 700 (2004) 3–50, <https://doi.org/10.1016/j.nuclphysb.2004.08.049>, arXiv:hep-ex/0405065.
- [6] A. Altheimer, et al., Jet substructure at the tevatron and LHC: new results, new tools, new benchmarks, J. Phys. G 39 (2012) 063001, <https://doi.org/10.1088/0954-3899/39/6/063001>, arXiv:1201.0008.
- [7] A. Altheimer, et al., Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd–27th of July 2012, Eur. Phys. J. C 74 (2014) 2792, <https://doi.org/10.1140/epjc/s10052-014-2792-8>, arXiv:1311.2708.
- [8] D. Adams, et al., Towards an understanding of the correlations in jet substructure, Eur. Phys. J. C 75 (2015) 409, <https://doi.org/10.1140/epjc/s10052-015-3587-2>, arXiv:1504.00679.
- [9] R. Kogler, et al., Jet substructure at the Large Hadron Collider: experimental review, Rev. Mod. Phys. 91 (2019) 045003, <https://doi.org/10.1103/RevModPhys.91.045003>, arXiv:1803.06991.
- [10] A.J. Larkoski, I. Moult, B. Nachman, Jet substructure at the Large Hadron Collider: a review of recent advances in theory and machine learning, Phys. Rep. 841 (2020) 1–63, <https://doi.org/10.1016/j.physrep.2019.11.001>.
- [11] D.M. South, M. Steder, H1, The H1 data preservation project, J. Phys. Conf. Ser. 396 (2012) 062019, <https://doi.org/10.1088/1742-6596/396/6/062019>, arXiv:1206.5200.
- [12] D. Britzger, S. Levonian, S. Schmitt, D. South, H1, Preservation through modernisation: the software of the H1 experiment at HERA, EPJ Web Conf. 251 (2021) 03004, <https://doi.org/10.1051/epjconf/202125103004>, arXiv:2106.11058.
- [13] A.J. Larkoski, J. Thaler, W.J. Waalewijn, Gaining (mutual) information about quark/gluon discrimination, J. High Energy Phys. 11 (2014) 129, [https://doi.org/10.1007/JHEP11\(2014\)129](https://doi.org/10.1007/JHEP11(2014)129), arXiv:1408.3122.
- [14] S. Catani, G. Turnock, B.R. Webber, Jet broadening measures in  $e^+e^-$  annihilation, Phys. Lett. B 295 (1992) 269–276, [https://doi.org/10.1016/0370-2693\(92\)91565-Q](https://doi.org/10.1016/0370-2693(92)91565-Q).
- [15] P.E.L. Rakow, B.R. Webber, Transverse momentum moments of hadron distributions in QCD jets, Nucl. Phys. B 191 (1981) 63–74, [https://doi.org/10.1016/0550-3213\(81\)90286-8](https://doi.org/10.1016/0550-3213(81)90286-8).
- [16] R.K. Ellis, B.R. Webber, QCD jet broadening in hadron hadron collisions, Conf. Proc. C 860623 (1986) 74.
- [17] E. Farhi, A QCD test for jets, Phys. Rev. Lett. 39 (1977) 1587–1588, <https://doi.org/10.1103/PhysRevLett.39.1587>.
- [18] S. Chatrchyan, et al., CMS, Search for a Higgs boson in the decay channel  $H \rightarrow ZZ^{(*)} \rightarrow q\bar{q}\ell^-\ell^+$  in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, J. High Energy Phys. 04 (2012) 036, [https://doi.org/10.1007/JHEP04\(2012\)036](https://doi.org/10.1007/JHEP04(2012)036), arXiv:1202.1416.
- [19] F. Pandolfi, Search for the Standard Model Higgs boson in the  $H \rightarrow ZZ \rightarrow l^+l^-q\bar{q}$  decay channel at CMS, Ph.D. thesis, Zurich, ETH, New York, 2012, <https://doi.org/10.1007/978-3-319-00903-2>.
- [20] Pileup Jet Identification, Technical Report, CERN, Geneva, 2013, <https://cds.cern.ch/record/1581583>.
- [21] R.D. Field, R.P. Feynman, A parametrization of the properties of quark jets, Nucl. Phys. B 136 (1978) 1, [https://doi.org/10.1016/0550-3213\(78\)90015-9](https://doi.org/10.1016/0550-3213(78)90015-9).
- [22] D. Krohn, M.D. Schwartz, T. Lin, W.J. Waalewijn, Jet charge at the LHC, Phys. Rev. Lett. 110 (2013) 212001, <https://doi.org/10.1103/PhysRevLett.110.212001>, arXiv:1209.2421.
- [23] Performance of quark/gluon discrimination in 8 TeV  $pp$  data, CERN, Geneva, 2013, <https://cds.cern.ch/record/1599732>.
- [24] Discrimination of light quark and gluon jets in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector, CERN, Geneva, 2016, <https://cds.cern.ch/record/2200202>.
- [25] Quark versus gluon jet tagging using charged particle multiplicity with the ATLAS detector, CERN, Geneva, 2017, <https://cds.cern.ch/record/2263679>.
- [26] G. Arnison, et al., UA1, Analysis of the fragmentation properties of quark and gluon jets at the CERN SPS  $p\bar{p}$  collider, Nucl. Phys. B 276 (1986) 253–271, [https://doi.org/10.1016/0550-3213\(86\)90296-8](https://doi.org/10.1016/0550-3213(86)90296-8).
- [27] P. Bagnaia, et al., UA2, Bern-CERN-Copenhagen-Osney-Pavia-Saclay, Measurement of jet fragmentation properties at the CERN  $p\bar{p}$  collider, Phys. Lett. B 144 (1984) 291–296, [https://doi.org/10.1016/0370-2693\(84\)91823-9](https://doi.org/10.1016/0370-2693(84)91823-9).
- [28] P. Bagnaia, et al., UA2, Measurement of production and properties of jets at the CERN  $p\bar{p}$  collider, Z. Phys. C 20 (1983) 117–134, <https://doi.org/10.1007/BF01573214>.
- [29] W. Bartel, et al., JADE, Experimental evidence for differences in  $p_T$  between quark jets and gluon jets, Phys. Lett. B 123 (1983) 460–466, [https://doi.org/10.1016/0370-2693\(83\)90994-2](https://doi.org/10.1016/0370-2693(83)90994-2).
- [30] W. Braunschweig, et al., TASSO, Charged multiplicity distributions and correlations in  $e^+e^-$  annihilation at PETRA energies, Z. Phys. C 45 (1989) 193, <https://doi.org/10.1007/BF01674450>.
- [31] M. Derrick, et al., Comparison of charged particle multiplicities in quark and gluon jets produced in  $e^+e^-$  annihilation at 29-GeV, Phys. Lett. B 165 (1985) 449–453, [https://doi.org/10.1016/0370-2693\(85\)91263-8](https://doi.org/10.1016/0370-2693(85)91263-8).
- [32] A. Petersen, et al., Inclusive charged particle distribution in nearly threefold symmetric three jet events at  $E_{c.m.} = 29$ -GeV, Phys. Rev. Lett. 55 (1985) 1954, <https://doi.org/10.1103/PhysRevLett.55.1954>.
- [33] K. Abe, et al., SLD, Measurement of the charged multiplicity of  $Z^0 \rightarrow b\bar{b}$  events, Phys. Rev. Lett. 72 (1994) 3145–3149, <https://doi.org/10.1103/PhysRevLett.72.3145>, arXiv:hep-ex/9405004.
- [34] K. Abe, et al., SLD, Measurement of the charged multiplicities in  $b$ ,  $c$  and light quark events from  $Z^0$  decays, Phys. Lett. B 386 (1996) 475–485, [https://doi.org/10.1016/0370-2693\(96\)01025-8](https://doi.org/10.1016/0370-2693(96)01025-8), arXiv:hep-ex/9608008.
- [35] Y.-K. Kim, et al., AMY, A comparison of quark and gluon jets produced in high-energy  $e^+e^-$  annihilations, Phys. Rev. Lett. 63 (1989) 1772, <https://doi.org/10.1103/PhysRevLett.63.1772>.
- [36] M.S. Alam, et al., CLEO, Study of gluon versus quark fragmentation in  $\Upsilon \rightarrow g\bar{g}\gamma$  and  $e^+e^- \rightarrow q\bar{q}\gamma$  events at  $\sqrt{s} = 10$  GeV, Phys. Rev. D 56 (1997) 17–22, <https://doi.org/10.1103/PhysRevD.56.17>, arXiv:hep-ex/9701006.
- [37] G. Alexander, et al., OPAL, A Direct observation of quark-gluon jet differences at LEP, Phys. Lett. B 265 (1991) 462–474, [https://doi.org/10.1016/0370-2693\(91\)90082-2](https://doi.org/10.1016/0370-2693(91)90082-2).
- [38] P.D. Acton, et al., OPAL, A study of differences between quark and gluon jets using vertex tagging of quark jets, Z. Phys. C 58 (1993) 387–404, <https://doi.org/10.1007/BF01557696>.
- [39] R. Akers, et al., OPAL, A model independent measurement of quark and gluon jet properties and differences, Z. Phys. C 68 (1995) 179–202, <https://doi.org/10.1007/BF01566667>.
- [40] D. Buskulic, et al., ALEPH, Study of the subjet structure of quark and gluon jets, Phys. Lett. B 346 (1995) 389–398, [https://doi.org/10.1016/0370-2693\(95\)00037-L](https://doi.org/10.1016/0370-2693(95)00037-L).
- [41] G. Alexander, et al., OPAL, Test of QCD analytic predictions for the multiplicity ratio between gluon and quark jets, Phys. Lett. B 388 (1996) 659–672, [https://doi.org/10.1016/S0370-2693\(96\)01319-6](https://doi.org/10.1016/S0370-2693(96)01319-6).
- [42] D. Buskulic, et al., ALEPH, Quark and gluon jet properties in symmetric three jet events, Phys. Lett. B 384 (1996) 353–364, [https://doi.org/10.1016/0370-2693\(96\)00849-0](https://doi.org/10.1016/0370-2693(96)00849-0).
- [43] P. Abreu, et al., DELPHI, Energy dependence of the differences between the quark and gluon jet fragmentation, Z. Phys. C 70 (1996) 179–196, <https://doi.org/10.1007/s002880050095>.
- [44] K. Ackerstaff, et al., OPAL, Multiplicity distributions of gluon and quark jets and tests of QCD analytic predictions, Eur. Phys. J. C 1 (1998) 479–494, <https://doi.org/10.1007/s100520050097>, arXiv:hep-ex/9708029.
- [45] P. Abreu, et al., DELPHI, The scale dependence of the hadron multiplicity in quark and gluon jets and a precise determination of  $C_A/C_F$ , Phys. Lett. B 449 (1999) 383–400, [https://doi.org/10.1016/S0370-2693\(99\)00112-4](https://doi.org/10.1016/S0370-2693(99)00112-4), arXiv:hep-ex/9903073.
- [46] G. Abbiendi, et al., OPAL, Experimental properties of gluon and quark jets from a point source, Eur. Phys. J. C 11 (1999) 217–238, <https://doi.org/10.1007/s100520050628>, arXiv:hep-ex/9903027.
- [47] M.Z. Akrawy, et al., OPAL, A study of coherence of soft gluons in hadron jets, Phys. Lett. B 247 (1990) 617–628, [https://doi.org/10.1016/0370-2693\(90\)91911-T](https://doi.org/10.1016/0370-2693(90)91911-T).
- [48] G. Abbiendi, et al., OPAL, Scaling violations of quark and gluon jet fragmentation functions in  $e^+e^-$  annihilations at  $\sqrt{s} = 91.2$  GeV and 183–209 GeV, Eur. Phys. J. C 37 (2004) 25–47, <https://doi.org/10.1140/epjc/s2004-01964-4>, arXiv:hep-ex/0404026.
- [49] T. Affolder, et al., CDF, Charged particle multiplicity in jets in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  TeV, Phys. Rev. Lett. 87 (2001) 211804, <https://doi.org/10.1103/PhysRevLett.87.211804>.
- [50] G. Aad, et al., ATLAS, Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy  $\sqrt{s} = 7$  TeV with the ATLAS detector, Phys. Rev. D 84 (2011) 054001, <https://doi.org/10.1103/PhysRevD.84.054001>, arXiv:1107.3311.
- [51] G. Aad, et al., ATLAS, Measurement of the jet fragmentation function and transverse profile in proton-proton collisions at a center-of-mass energy of 7 TeV with the ATLAS detector, Eur. Phys. J. C 71 (2011) 1795, <https://doi.org/10.1140/epjc/s10052-011-1795-y>, arXiv:1109.5816.
- [52] S. Chatrchyan, et al., CMS, Shape, transverse size, and charged-hadron multiplicity of jets in  $pp$  collisions at  $\sqrt{s} = 7$  TeV, J. High Energy Phys. 06 (2012) 160, [https://doi.org/10.1007/JHEP06\(2012\)160](https://doi.org/10.1007/JHEP06(2012)160), arXiv:1204.3170.
- [53] ATLAS Collaboration, Measurement of the charged particle multiplicity inside jets from  $\sqrt{s} = 8$  TeV  $pp$  collisions with the ATLAS detector, Eur. Phys. J. C 76 (2016) 1, <https://doi.org/10.1140/epjc/s10052-016-4126-5>, arXiv:1602.00988.
- [54] ATLAS Collaboration, Properties of jet fragmentation using charged particles measured with the ATLAS detector in  $pp$  collisions at  $\sqrt{s} = 13$  TeV, Phys. Rev. D 100 (2019) 052011, <https://doi.org/10.1103/PhysRevD.100.052011>, arXiv:1906.09254.
- [55] A. Capella, I.M. Dremin, J.W. Gary, V.A. Nepochailo, J. Van Tran Thanh, Evolution of average multiplicities of quark and gluon jets, Phys. Rev. D 61 (2000) 074009, <https://doi.org/10.1103/PhysRevD.61.074009>, arXiv:hep-ph/9910226.
- [56] I.M. Dremin, J.W. Gary, Energy dependence of mean multiplicities in gluon and quark jets at the next-to-next-to-leading order, Phys. Lett. B 459 (1999) 341–346, [https://doi.org/10.1016/S0370-2693\(99\)00713-3](https://doi.org/10.1016/S0370-2693(99)00713-3), arXiv:hep-ph/9905477; Erratum: Phys. Lett. B 463 (1999) 346.
- [57] ATLAS Pythia 8 tunes to 7 TeV data, Technical Report, CERN, Geneva, 2014, <https://cds.cern.ch/record/1966419>.

- [58] ATLAS Collaboration, Measurement of jet charge in dijet events from  $\sqrt{s} = 8$  TeV  $pp$  collisions with the ATLAS detector, Phys. Rev. D 93 (2016) 052003, <https://doi.org/10.1103/PhysRevD.93.052003>, arXiv:1509.05190.
- [59] A.M. Sirunyan, et al., CMS, Measurements of jet charge with dijet events in  $pp$  collisions at  $\sqrt{s} = 8$  TeV, J. High Energy Phys. 10 (2017) 131, [https://doi.org/10.1007/JHEP10\(2017\)131](https://doi.org/10.1007/JHEP10(2017)131), arXiv:1706.05868.
- [60] W.J. Waalewijn, Calculating the charge of a jet, Phys. Rev. D 86 (2012) 094030, <https://doi.org/10.1103/PhysRevD.86.094030>, arXiv:1209.3019.
- [61] C. Frye, A.J. Larkoski, M.D. Schwartz, K. Yan, Factorization for groomed jet substructure beyond the next-to-leading logarithm, J. High Energy Phys. 07 (2016) 064, [https://doi.org/10.1007/JHEP07\(2016\)064](https://doi.org/10.1007/JHEP07(2016)064), arXiv:1603.09338.
- [62] C. Frye, A.J. Larkoski, M.D. Schwartz, K. Yan, Precision physics with pile-up insensitive observables, arXiv:1603.06375, 2016.
- [63] S. Marzani, L. Schunk, G. Soyez, The jet mass distribution after Soft Drop, Eur. Phys. J. C 78 (2018) 96, <https://doi.org/10.1140/epjc/s10052-018-5579-5>, arXiv:1712.05105.
- [64] S. Marzani, L. Schunk, G. Soyez, A study of jet mass distributions with grooming, J. High Energy Phys. 07 (2017) 132, [https://doi.org/10.1007/JHEP07\(2017\)132](https://doi.org/10.1007/JHEP07(2017)132), arXiv:1704.02210.
- [65] Z.-B. Kang, K. Lee, X. Liu, F. Ringer, The groomed and ungroomed jet mass distribution for inclusive jet production at the LHC, J. High Energy Phys. 10 (2018) 137, [https://doi.org/10.1007/JHEP10\(2018\)137](https://doi.org/10.1007/JHEP10(2018)137), arXiv:1803.03645.
- [66] Z.-B. Kang, K. Lee, X. Liu, F. Ringer, Soft drop groomed jet angularities at the LHC, Phys. Lett. B 793 (2019) 41–47, <https://doi.org/10.1016/j.physletb.2019.04.018>, arXiv:1811.06983.
- [67] J.R. Andersen, et al., Les Houches 2017: physics at TeV colliders Standard Model working group report, arXiv:1803.07977, 2018.
- [68] E.-C. Aschenauer, K. Lee, B.S. Page, F. Ringer, Jet angularities in photoproduction at the Electron-Ion Collider, Phys. Rev. D 101 (2020) 054028, <https://doi.org/10.1103/PhysRevD.101.054028>, arXiv:1910.11460.
- [69] Z.-B. Kang, X. Liu, S. Mantry, D.Y. Shao, Jet charge: a flavor prism for spin asymmetries at the Electron-Ion Collider, Phys. Rev. Lett. 125 (2020) 242003, <https://doi.org/10.1103/PhysRevLett.125.242003>.
- [70] P. Agostini, et al., LHeC, FCC-he Study Group, The Large Hadron–Electron Collider at the HL-LHC, J. Phys. G 48 (2021) 110501, <https://doi.org/10.1088/1361-6471/abf3ba>, arXiv:2007.14491.
- [71] A. Abada, et al., FCC, FCC physics opportunities: future circular collider conceptual design report volume 1, Eur. Phys. J. C 79 (2019) 474, <https://doi.org/10.1140/epjc/s10052-019-6904-3>.
- [72] I. Abt, et al., H1, The H1 detector at HERA, DESY-93-103, 1993.
- [73] B. Andrieu, et al., H1 Calorimeter Group, The H1 liquid argon calorimeter system, Nucl. Instrum. Methods A 336 (1993) 460–498, [https://doi.org/10.1016/0168-9002\(93\)91257-N](https://doi.org/10.1016/0168-9002(93)91257-N).
- [74] I. Abt, et al., H1, The H1 detector at HERA, Nucl. Instrum. Methods A 386 (1997) 310–347, [https://doi.org/10.1016/S0168-9002\(96\)00893-5](https://doi.org/10.1016/S0168-9002(96)00893-5).
- [75] I. Abt, et al., H1, The tracking, calorimeter and muon detectors of the H1 experiment at HERA, Nucl. Instrum. Methods A 386 (1997) 348–396, [https://doi.org/10.1016/S0168-9002\(96\)00894-7](https://doi.org/10.1016/S0168-9002(96)00894-7).
- [76] R.D. Appuhn, et al., H1 SPACAL Group, The H1 lead/scintillating fiber calorimeter, Nucl. Instrum. Methods A 386 (1997) 397–408, [https://doi.org/10.1016/S0168-9002\(96\)01171-0](https://doi.org/10.1016/S0168-9002(96)01171-0).
- [77] D. Pitzl, et al., The H1 silicon vertex detector, Nucl. Instrum. Methods A 454 (2000) 334–349, [https://doi.org/10.1016/S0168-9002\(00\)00488-5](https://doi.org/10.1016/S0168-9002(00)00488-5), arXiv:hep-ex/0002044.
- [78] B. Andrieu, et al., H1 Calorimeter Group, Beam tests and calibration of the H1 liquid argon calorimeter with electrons, Nucl. Instrum. Methods A 350 (1994) 57–72, [https://doi.org/10.1016/0168-9002\(94\)91155-X](https://doi.org/10.1016/0168-9002(94)91155-X).
- [79] B. Andrieu, et al., H1 Calorimeter Group, Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations, Nucl. Instrum. Methods A 336 (1993) 499–509, [https://doi.org/10.1016/0168-9002\(93\)91258-0](https://doi.org/10.1016/0168-9002(93)91258-0).
- [80] F.D. Aaron, et al., H1, Determination of the integrated luminosity at HERA using elastic QED Compton events, Eur. Phys. J. C 72 (2012) 2163, <https://doi.org/10.1140/epjc/s10052-012-2163-2>, arXiv:1205.2448; Erratum: Eur. Phys. J. C 74 (2014) 2733.
- [81] C. Adloff, et al., H1, Measurement and QCD analysis of neutral and charged current cross-sections at HERA, Eur. Phys. J. C 30 (2003) 1–32, <https://doi.org/10.1140/epjc/s2003-01257-6>, arXiv:hep-ex/0304003.
- [82] F.D. Aaron, et al., H1, Inclusive deep inelastic scattering at high  $Q^2$  with longitudinally polarised lepton beams at HERA, J. High Energy Phys. 09 (2012) 061, [https://doi.org/10.1007/JHEP09\(2012\)061](https://doi.org/10.1007/JHEP09(2012)061), arXiv:1206.7007.
- [83] V. Andreev, et al., H1, Measurement of multijet production in  $ep$  collisions at high  $Q^2$  and determination of the strong coupling  $\alpha_s$ , Eur. Phys. J. C 75 (2015) 65, <https://doi.org/10.1140/epjc/s10052-014-3223-6>, arXiv:1406.4709.
- [84] V. Andreev, et al., H1, Measurement of jet production cross sections in deep-inelastic  $ep$  scattering at HERA, Eur. Phys. J. C 77 (2017) 215, <https://doi.org/10.1140/epjc/s10052-017-4717-9>, arXiv:1611.03421; Erratum: Eur. Phys. J. C 81 (2021) 739.
- [85] U. Bassler, G. Bernardi, On the kinematic reconstruction of deep inelastic scattering at HERA: the Sigma method, Nucl. Instrum. Methods A 361 (1995) 197–208, [https://doi.org/10.1016/0168-9002\(95\)00173-5](https://doi.org/10.1016/0168-9002(95)00173-5), arXiv:hep-ex/9412004.
- [86] M. Peetz, Search for deviations from the standard model in high transverse energy processes at the electron proton collider HERA, PhD thesis, Lyon 1 U., 2003, <https://doi.org/10.3204/DESY-THESIS-2003-023>.
- [87] S. Hellwig, Untersuchung der  $D^* - \pi_{low}$  Double Tagging Methode in Charm-analysen, Master's thesis, Hamburg U., 2004.
- [88] B. Porthault, First measurement of charged and neutral current cross sections with the polarized positron beam at HERA II and QCD-electroweak analyses, PhD thesis, Paris XI Orsay U., 2005.
- [89] M. Cacciari, G.P. Salam, G. Soyez, FastJet user manual, Eur. Phys. J. C 72 (2012) 1896, <https://doi.org/10.1140/epjc/s10052-012-1896-2>, arXiv:1111.6097.
- [90] M. Cacciari, G.P. Salam, Dispelling the  $N^3$  myth for the  $k_t$  jet-finder, Phys. Lett. B 641 (2006) 57–61, <https://doi.org/10.1016/j.physletb.2006.08.037>, arXiv:hep-ph/0512210.
- [91] S. Catani, Y.L. Dokshitzer, M.H. Seymour, B.R. Webber, Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions, Nucl. Phys. B 406 (1993) 187–224, [https://doi.org/10.1016/0550-3213\(93\)90166-M](https://doi.org/10.1016/0550-3213(93)90166-M).
- [92] S.D. Ellis, D.E. Soper, Successive combination jet algorithm for hadron collisions, Phys. Rev. D 48 (1993) 3160–3166, <https://doi.org/10.1103/PhysRevD.48.3160>, arXiv:hep-ph/9305266.
- [93] 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, [https://doi.org/10.1016/0010-4655\(94\)90086-8](https://doi.org/10.1016/0010-4655(94)90086-8).
- [94] H. Jung, Hard diffractive scattering in high-energy  $ep$  collisions and the Monte Carlo generator RAPGAP, Comput. Phys. Commun. 86 (1995) 147–161, [https://doi.org/10.1016/0010-4655\(94\)00150-Z](https://doi.org/10.1016/0010-4655(94)00150-Z).
- [95] H. Spiesberger, et al., Radiative corrections at HERA, in: Workshop on Physics at HERA, 1992.
- [96] 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, <https://doi.org/10.1007/BF01558572>.
- [97] 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, [https://doi.org/10.1016/0010-4655\(92\)90136-M](https://doi.org/10.1016/0010-4655(92)90136-M).
- [98] 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, J. High Energy Phys. 07 (2002) 012, <https://doi.org/10.1088/1126-6708/2002/07/012>, arXiv:hep-ph/0201195.
- [99] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Parton fragmentation and string dynamics, Phys. Rep. 97 (1983) 31–145, [https://doi.org/10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7).
- [100] S. Schael, et al., ALEPH, Bose-Einstein correlations in  $W$ -pair decays with an event-mixing technique, Phys. Lett. B 606 (2005) 265–275, <https://doi.org/10.1016/j.physletb.2004.12.018>.
- [101] L. Lönnblad, ARIADNE version 4: a program for simulation of QCD cascades implementing the color dipole model, Comput. Phys. Commun. 71 (1992) 15–31, [https://doi.org/10.1016/0010-4655\(92\)90068-A](https://doi.org/10.1016/0010-4655(92)90068-A).
- [102] R. Brun, F. Bruyant, M. Maire, A.C. McPherson, P. Zanzarini, GEANT3, Manual, CERN, Geneva, 1987, <https://cds.cern.ch/record/1119728>.
- [103] T. Sjöstrand, S. Mrenna, P.Z. Skands, PYTHIA 6.4 physics and manual, J. High Energy Phys. 05 (2006) 026, <https://doi.org/10.1088/1126-6708/2006/05/026>, arXiv:hep-ph/0603175.
- [104] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C.O. Rasmussen, P.Z. Skands, An introduction to PYTHIA 8.2, Comput. Phys. Commun. 191 (2015) 159–177, <https://doi.org/10.1016/j.cpc.2015.01.024>, arXiv:1410.3012.
- [105] W.T. Giele, D.A. Kosower, P.Z. Skands, A simple shower and matching algorithm, Phys. Rev. D 78 (2008) 014026, <https://doi.org/10.1103/PhysRevD.78.014026>, arXiv:0707.3652.
- [106] W.T. Giele, L. Hartgring, D.A. Kosower, E. Laenen, A.J. Larkoski, J.J. Lopez-Villarejo, M. Ritzmann, P. Skands, The VINCIA parton shower, PoS DIS2013 (2013) 165, <https://doi.org/10.22323/1.191.0165>, arXiv:1307.1060.
- [107] S. Höche, S. Prestel, The midpoint between dipole and parton showers, Eur. Phys. J. C 75 (2015) 461, <https://doi.org/10.1140/epjc/s10052-015-3684-2>, arXiv:1506.05057.
- [108] R.D. Ball, et al., NNPDF, Parton distributions from high-precision collider data, Eur. Phys. J. C 77 (2017) 663, <https://doi.org/10.1140/epjc/s10052-017-5199-5>, arXiv:1706.00428.
- [109] L.A. Harland-Lang, A.D. Martin, P. Motylinski, R.S. Thorne, Parton distributions in the LHC era: MMHT 2014 PDFs, Eur. Phys. J. C 75 (2015) 204, <https://doi.org/10.1140/epjc/s10052-015-3397-6>, arXiv:1412.3989.
- [110] J. Bellm, et al., Herwig 7.0/Herwig++ 3.0 release note, Eur. Phys. J. C 76 (2016) 196, <https://doi.org/10.1140/epjc/s10052-016-4018-8>, arXiv:1512.01178.
- [111] M. Bahr, et al., Herwig++ physics and manual, Eur. Phys. J. C 58 (2008) 639–707, <https://doi.org/10.1140/epjc/s10052-008-0798-9>, arXiv:0803.0883.
- [112] J. Bellm, et al., Herwig 7.2 release note, Eur. Phys. J. C 80 (2020) 452, <https://doi.org/10.1140/epjc/s10052-020-8011-x>, arXiv:1912.06509.

- [113] L. Lönnblad, S. Prestel, Merging multi-leg NLO matrix elements with parton showers, *J. High Energy Phys.* 03 (2013) 166, [https://doi.org/10.1007/JHEP03\(2013\)166](https://doi.org/10.1007/JHEP03(2013)166), arXiv:1211.7278.
- [114] E. Bothmann, et al., Sherpa, Event generation with Sherpa 2.2, *SciPost Phys.* 7 (2019) 034, <https://doi.org/10.21468/SciPostPhys.7.3.034>, arXiv:1905.09127.
- [115] G.S. Chahal, F. Krauss, Cluster hadronisation in Sherpa, *SciPost Phys.* 13 (2022) 019, <https://doi.org/10.21468/SciPostPhys.13.2.019>, arXiv:2203.11385.
- [116] S. Schumann, F. Krauss, A parton shower algorithm based on Catani-Seymour dipole factorisation, *J. High Energy Phys.* 03 (2008) 038, <https://doi.org/10.1088/1126-6708/2008/03/038>, arXiv:0709.1027.
- [117] S. Höche, F. Krauss, S. Schumann, F. Siegert, QCD matrix elements and truncated showers, *J. High Energy Phys.* 05 (2009) 053, <https://doi.org/10.1088/1126-6708/2009/05/053>, arXiv:0903.1219.
- [118] S. Höche, F. Krauss, M. Schonherr, F. Siegert, QCD matrix elements + parton showers: the NLO case, *J. High Energy Phys.* 04 (2013) 027, [https://doi.org/10.1007/JHEP04\(2013\)027](https://doi.org/10.1007/JHEP04(2013)027), arXiv:1207.5030.
- [119] 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, <https://doi.org/10.1103/PhysRevLett.124.182001>, arXiv:1911.09107.
- [120] A. Andreassen, P.T. Komiske, E.M. Metodiev, B. Nachman, A. Suresh, J. Thaler, Scaffolding simulations with deep learning for high-dimensional deconvolution, in: 9th International Conference on Learning Representations, 2021, arXiv:2105.04448.
- [121] L.B. Lucy, An iterative technique for the rectification of observed distributions, *Astron. J.* 79 (1974) 745–754, <https://doi.org/10.1086/111605>.
- [122] W.H. Richardson, Bayesian-based iterative method of image restoration, *J. Opt. Soc. Am.* 62 (1972) 55.
- [123] G. D'Agostini, A multidimensional unfolding method based on Bayes' theorem, *Nucl. Instrum. Methods A* 362 (1995) 487–498, [https://doi.org/10.1016/0168-9002\(95\)00274-X](https://doi.org/10.1016/0168-9002(95)00274-X).
- [124] P.T. Komiske, E.M. Metodiev, J. Thaler, Energy flow networks: deep sets for particle jets, *J. High Energy Phys.* 01 (2019) 121, [https://doi.org/10.1007/JHEP01\(2019\)121](https://doi.org/10.1007/JHEP01(2019)121), arXiv:1810.05165.
- [125] H. Qu, L. Gouskos, ParticleNet: jet tagging via particle clouds, *Phys. Rev. D* 101 (2020) 056019, <https://doi.org/10.1103/PhysRevD.101.056019>, arXiv:1902.08570.
- [126] E.A. Moreno, O. Cerri, J.M. Duarte, H.B. Newman, T.Q. Nguyen, A. Periwai, M. Pierini, A. Serikova, M. Spiropulu, J.-R. Vlimant, JEDI-net: a jet identification algorithm based on interaction networks, *Eur. Phys. J. C* 80 (2020) 58, <https://doi.org/10.1140/epjc/s10052-020-7608-4>, arXiv:1908.05318.
- [127] E.A. Moreno, T.Q. Nguyen, J.-R. Vlimant, O. Cerri, H.B. Newman, A. Periwai, M. Spiropulu, J.M. Duarte, M. Pierini, Interaction networks for the identification of boosted  $H \rightarrow b\bar{b}$  decays, *Phys. Rev. D* 102 (2020) 012010, <https://doi.org/10.1103/PhysRevD.102.012010>, arXiv:1909.12285.
- [128] V. Mikuni, F. Canelli, ABCNet: an attention-based method for particle tagging, *Eur. Phys. J. Plus* 135 (2020) 463, <https://doi.org/10.1140/epjp/s13360-020-00497-3>, arXiv:2001.05311.
- [129] E. Bernreuther, T. Finke, F. Kahlhoefer, M. Krämer, A. Mück, Casting a graph net to catch dark showers, *SciPost Phys.* 10 (2021) 046, <https://doi.org/10.21468/SciPostPhys.10.2.046>, arXiv:2006.08639.
- [130] J. Guo, J. Li, T. Li, R. Zhang, Boosted Higgs boson jet reconstruction via a graph neural network, *Phys. Rev. D* 103 (2021) 116025, <https://doi.org/10.1103/PhysRevD.103.116025>, arXiv:2010.05464.
- [131] M.J. Dolan, A. Ore, Equivariant energy flow networks for jet tagging, *Phys. Rev. D* 103 (2021) 074022, <https://doi.org/10.1103/PhysRevD.103.074022>, arXiv:2012.00964.
- [132] V. Mikuni, F. Canelli, Point cloud transformers applied to collider physics, *Mach. Learn.: Sci. Technol.* 2 (2021) 035027, <https://doi.org/10.1088/2632-2153/ac07f6>, arXiv:2102.05073.
- [133] P. Konar, V.S. Ngairangbam, M. Spannowsky, Energy-weighted message passing: an infra-red and collinear safe graph neural network algorithm, arXiv:2109.14636, 2021.
- [134] C. Shimmmin, Particle convolution for high energy physics, arXiv:2107.02908, 2021.
- [135] S. Gong, Q. Meng, J. Zhang, H. Qu, C. Li, S. Qian, W. Du, Z.-M. Ma, T.-Y. Liu, An efficient Lorentz equivariant graph neural network for jet tagging, arXiv:2201.08187, 2022.
- [136] M.-H. Guo, J.-X. Cai, Z.-N. Liu, T.-J. Mu, R.R. Martin, S.-M. Hu, PCT: point cloud transformer, *Comput. Vis. Media* 7 (2021) 187–199.
- [137] F. Chollet, Keras, <https://github.com/fchollet/keras>, 2017.
- [138] 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.
- [139] D. Kingma, J. Ba Adam, A method for stochastic optimization, arXiv:1412.6980, 2014.
- [140] Perlmutter system, [https://docs.nersc.gov/systems/perlmutter/system\\_details/](https://docs.nersc.gov/systems/perlmutter/system_details/). (Accessed 4 May 2022).
- [141] A. Sergeev, M.D. Balso, Horovod: fast and easy distributed deep learning in TensorFlow, arXiv preprint, arXiv:1802.05799, 2018.
- [142] R. Kogler, Measurement of jet production in deep-inelastic  $ep$  scattering at HERA, <https://www-h1.desy.de/psfiles/theses/h1th-590.pdf>, 2011.
- [143] F.D. Aaron, et al., H1, Measurement of  $D^{*\pm}$  meson production and determination of  $F_2^{c\bar{c}}$  at low  $Q^2$  in deep-inelastic scattering at HERA, *Eur. Phys. J. C* 71 (2011) 1769, <https://doi.org/10.1140/epjc/s10052-011-1769-0>, arXiv:1106.1028; Erratum: *Eur. Phys. J. C* 72 (2012) 2252.
- [144] B. Efron, Bootstrap methods: another look at the jackknife, *Ann. Stat.* 7 (1979) 1–26, <https://doi.org/10.1214/aos/1176344552>.