Hadron-level NLO predictions for QCD observables in photo-production at the Electron-Ion Collider

Peter Meinzinger[*] and Frank Krauss[†]

Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, United Kingdom

We present the first next-to-leading order accurate hadron-level predictions for inclusive QCD and jet observables in photo-production for electron-proton beams at the upcoming electron-ion collider. Our predictions show large K factors of about 60% and a dominating real correction.

I. INTRODUCTION

The upcoming electron-ion collider (EIC) aims at precision measurements of a variety of phenomena related to strong interactions and the impact of a nuclear environment on them. The scientific program [1] includes determinations of the strong coupling constant and its scaling behavior, observations of saturation effects, and measurements of the three-dimensional structure of nucleons as encoded in parton distribution functions (PDFs), generalized parton distributions (GPDs), or transverse-momentum dependent parton distributions functions (TMDs). This ambitious program rests, among others, on high-quality theory predictions, both in the form of analytic calculations and of Monte Carlo simulations, also known as event generators.

Some of the first analyses at the EIC will investigate the production of inclusive hadronic final states; their properties are usually described through a combination of event shape and jet observables, particle multiplicities and distributions. From previous experiments, for example at the HERA collider at DESY, it has become customary to classify scattering events by the virtual mass squared, or virtuality, Q^2, of the photon exchanged between the incident electron and nucleon (or, in the case of the EIC, nucleus). Broadly speaking, essential parts of the HERA scientific programme, namely the determination of proton structure functions and PDFs, were nearly exclusively driven by the regime of nonzero Q^2 associated with deeply-inelastic scattering or electroproduction. The same will certainly be also true for EIC-based efforts of improved PDF determination.

However, as the spectrum of the exchanged photons scales like $1/Q^2$ it is clear that the electron-proton cross section is dominated by the photo-production regime of relatively low $Q^2 \approx 0$. This necessitates the careful treatment of such photo-production events, which so far have not attracted the same attention as, for example, event generation for the LHC, which has reached a very satisfying level of theoretical accuracy and maturity [2]. This is due to the fact that in this regime the photon has a non-negligible hadronic component, essentially fluctuating into states with quantum numbers of neutral vector mesons such as the ρ^0 or similar. At higher resolution scales, this non-perturbative component of the photon structure is further augmented by perturbative splittings of the photon into quark–anti-quark pairs—a feature that is not present in the more familiar proton PDFs. As a consequence, the corresponding γp collisions will not be characterized by a pointlike photon interacting with partons, but rather look like hadron-hadron collisions involving the PDFs of photons, with parametrizations usually dating back two decades. This in itself suggests an interesting physics programme related to photo-production, evidenced by the breadth and number of published analysis by the HERA collaborations, which cover a wide range of exclusive and inclusive final states.

In this paper we build on previous work [3] and apply the SHERPA framework for the simulation of hadron-level photo-production events at NLO accuracy to the EIC. The quality of SHERPA's agreement with HERA data in this publication gives us confidence that our results presented here will provide a first essential, high-quality baseline for future studies of inclusive QCD final states at this experiment, based on the well-established framework of collinear factorization. The code underpinning our work will be made publicly available in a forthcoming release of SHERPA.

II. EVENT GENERATION

In contrast to high-Q^2 scattering photo-production is characterized by vanishing photon virtualities of $Q^2 \approx 0$.

---
[*]peter.meinzinger@durham.ac.uk
[†]frank.krauss@durham.ac.uk

For the simulation of such events, two different processes have to be considered: the photon can either enter the hard-processes directly, $\gamma j \rightarrow jj$, the "direct" component, or it can first undergo a transition to its hadronic component or split into a quark-antiquark pair, followed by the usual QCD scale evolution, leading to the process $jj \rightarrow jj$, the so-called "resolved" component. The total cross section therefore decomposes as

$$\sigma_{\text{tot}} = \sigma_{\gamma j \rightarrow X} + \sigma_{jj \rightarrow X}, \qquad (1)$$

where j denotes partons coming from either the proton or the photon. Our results, presented below, are obtained by running SHERPA [4,5], where in each run we generated 10 million events each at MC@NLO accuracy and at leading order for comparison. We used the implementation of the photon flux and the photon PDFs as described in [3]. Matrix elements were calculated with AMEGIC [6] and COMIX [7,8] for tree-level and OpenLoops [9] for one-loop order. QCD Bremsstrahlung was modeled by the CSShower parton shower [10] and matched to the NLO matrix elements with the SHERPA realization [11] of the MC@NLO method [12,13]. Multiple-parton interactions have been added through an implementation of the Sjostrand-van Zijl model [14,15] and the partons were hadronized with the cluster fragmentation of [16]. Analogously to [3], we used the SAS1M and SAS2M sets as photon PDFs and averaged over these two sets as suggested by their authors [17,18]. For the proton we used the PDF4LHC21_40_pdfas set [19] interfaced through LHAPDF [20] and set the value and running of α_S accordingly. For the scales we chose $\mu_F = \mu_R = \max\{H_T/2, 4 \text{ GeV}\}$, and estimated the associated uncertainties with a 7-point scale variation. We used RIVET [21] for the event analysis.

The beam setup was chosen to be an electron and a proton beam with energies of 18 and 275 GeV, respectively, in accordance with the highest energy scenario at the planned EIC facilities [1]. Similarly to studies in [22], particles were analyzed in the laboratory frame using the anti-k_T jet algorithm with $R = 1.0$ and a minimum transverse energy of $E_T > 6$ GeV and a pseudorapidity $|\eta| < 4$, demanding one jet in the event.

III. PREDICTIONS FOR JET PRODUCTION

In Fig. 1, we present the inclusive jet transverse momentum distribution and the observable x_γ, which is defined through

$$x_\gamma = \frac{E_T^{(1)} e^{-\eta^{(1)}} + E_T^{(2)} e^{-\eta^{(2)}}}{2yE_e}. \qquad (2)$$

Here, $E_T^{(1,2)}$ and $\eta^{(1,2)}$ denote the jet transverse energy and pseudorapidity of the leading and sub-leading jet, respectively, and y denotes the fraction of the electron energy E_e carried by the photon. It had previously been observed that the x_γ observable is especially suitable to discern different photon PDF parametrizations [23].

The spectrum of the jet transverse energies is driven by the resolved component at $E_T < 10$ GeV, with the direct component taking over with increasing jet E_T. In the x_γ-distribution the two components are more clearly separated: for low values of $x_\gamma \lesssim 0.6$ the resolved component dominates, however at $x_\gamma > 0.6$ the cross section of the direct component exceeds the resolved one. This is significantly lower than the value of $x_\gamma > 0.75$ that had been used as a cut at HERA [23,24] and is a consequence of the lower collision energy, decreasing the contribution of the resolved component relative to the direct cross section.

Overall we see a large K factor in these two observables of about 60%. This can be attributed to the asymmetric jet cuts which leads to a large part of the phase space being left unfilled in the leading order calculation. By means of the real correction this part of the phase space is filled up,
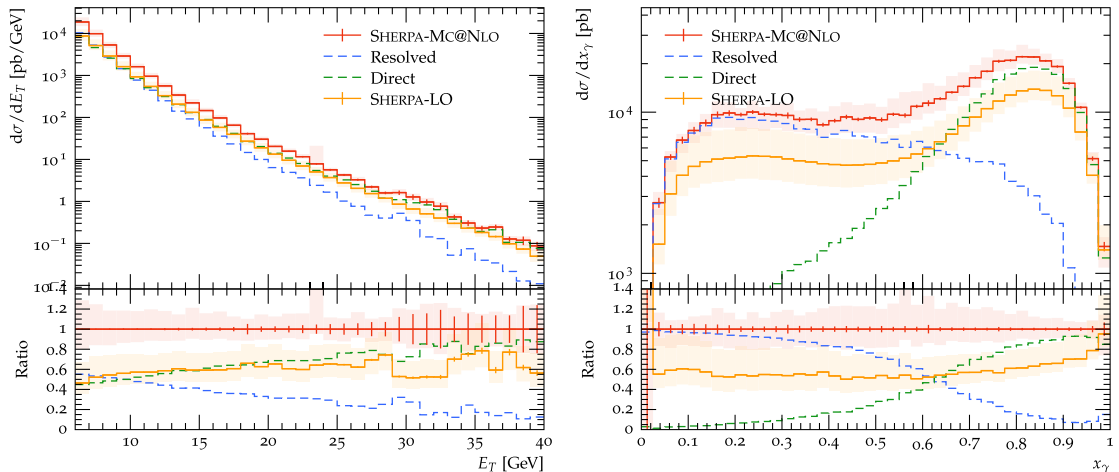

FIG. 1.   Distributions of inclusive jet transverse energy E_T (left) and the x_γ (right) of the SHERPA simulation with MC@NLO accuracy, compared with results at LO.

however, owing to the low E_T cut on the jets and the Born-level order of these corrections, the cross section does not receive a significant reduction of its scale uncertainty bands.

Moving on to event shapes, we looked at similar variables as in [25], studying distributions of transverse thrust and transverse thrust minor in the upper panel of Fig. 2. In both observables, the direct and the resolved component contribute to approximately equal amounts throughout the whole parameter space. This is in contrast to the C-parameter distribution, depicted in Fig. 2 (lower left panel), where we can identify distinct regions where one of the two components dominates. While the direct component contributes mostly at the lowest value and in the tail of the distribution, the resolved component contribution reaches up to about 60% near the global maximum of $C \approx 0.1$. Looking at the hadron multiplicity distribution in Fig. 2 (lower right), we see—in agreement with expectations—that the low-multiplicity states are dominantly

produced by direct photo-production, while the high-multiplicity tail is determined by the resolved component. It is worth noting here that the effect of multiple-parton interactions is generally mostly negligible at the EIC collision energies; it does, however play a significant role in high-multiplicity states, therefore still necessitating a tuning of the modeling.

As a last pair of observables, we look at the ratio of c- and b-quark jets to light-quark jets in transverse momentum and pseudorapidity distributions in Fig. 3. Heavy-quark jets are almost exclusively produced through the direct process and in the region $\eta < 0$, where they make up almost a third of the overall activity. As expected, the heavy-quark jets are predominantly c-quark jets and their tagging will benefit from a cut on the x_γ value.

As indicated above, the x_γ distribution is particularly important for tagging and discerning different event types, and it also acts as an excellent discriminator of resolved and unresolved photon interactions. Given the relative cross
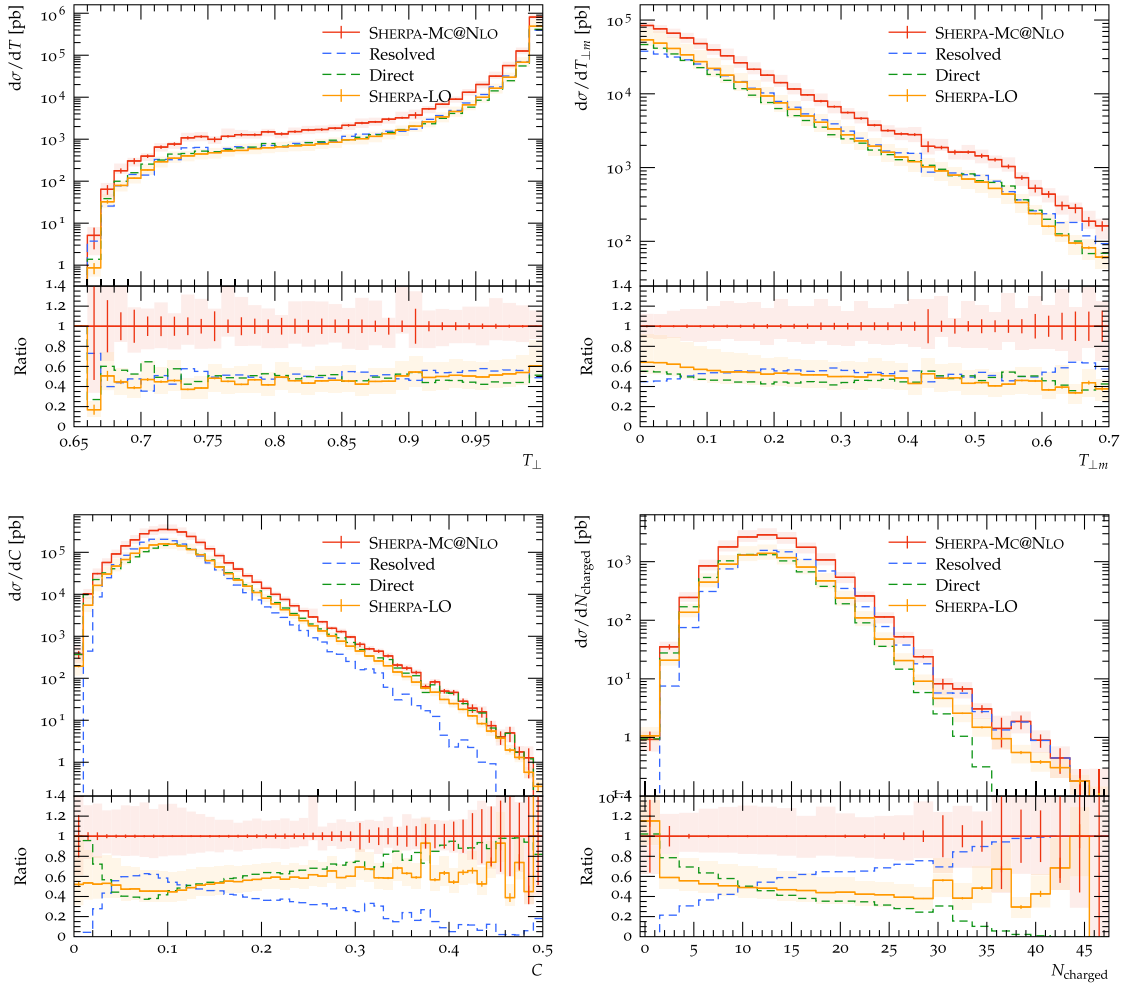

FIG. 2. Event shapes observables: Transverse thrust major (upper left) and minor (upper right) distributions, C parameter (lower left) and charged particle multiplicity distribution N_{charged} inside the detector acceptance of $|\eta| < 4$ (lower right), all obtained with the SHERPA simulation at MC@NLO accuracy, compared with results at LO.
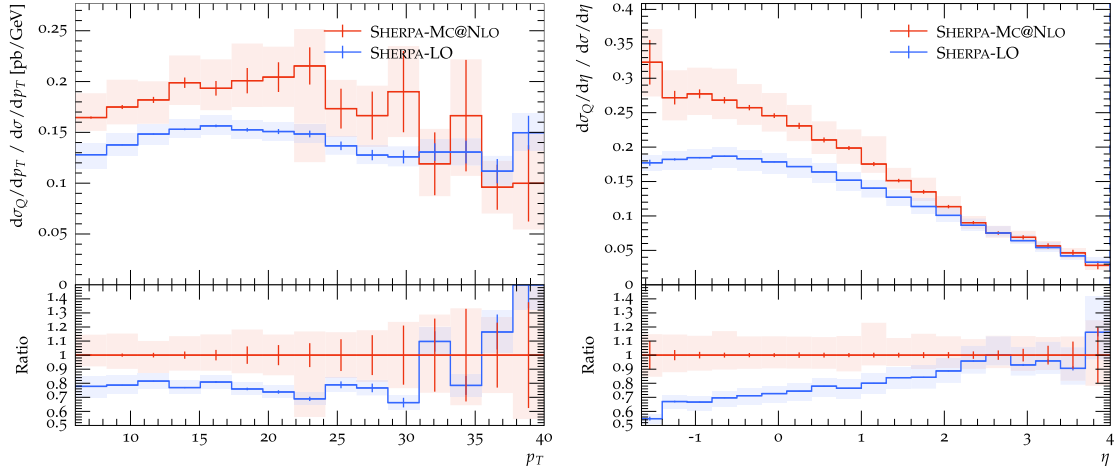
FIG. 3.   Distributions of heavy quark jet transverse momentum p_T (left) and pseudorapidity η (right) of the SHERPA simulation with MC@NLO accuracy, compared with results at LO.

sections, it is also clearly important to improve the theoretical predictions obtained in resolved photo-production. However, at the moment, this domain suffers from large uncertainties due to the somewhat outdated photon PDFs, obtained about 2 decades earlier, and with proton PDFs from the same period for the description of photo-production at HERA. This is further highlighted in Fig. 4, where we generated one million parton-level events at LO for 11 photon PDFs sets each, including a 7-point scale variation, using the same settings as before. Creating an envelope from the minimal/maximal values per bin among the different PDF sets, we find that the variations are

of similar size as the scale uncertainties and reach up to 50% of the binned cross section. This finding also suggests that a renewed fitting exercise of photon PDFs is of great importance for the upcoming EIC and the full exploitation and understanding of its data.

IV. SUMMARY

In this paper we presented the first fully differential hadron-level study at next-to-leading order accuracy for the photo-production of inclusive QCD final states at the EIC. For our predictions, obtained with SHERPA, we assumed beam energies of 18 (275) GeV for the electron (proton) beam and analyzed the events with inclusive-jet as well as event-shape and heavy-jet observables. Photo-production represents an interesting laboratory to study a large variety of sometimes subtle QCD effects, by providing access to a wide range of exclusive and inclusive production channels. Our results will also be instrumental for renewed efforts to fit PDFs for the photon in collinear factorization.

We identified the photon PDFs as probably the most obvious bottleneck in precise photo-production phenomenology for the EIC, as they are based on fits from 2 decades ago, using contemporary proton PDFs and simulation tools of that period. An updated fit based on modern methodology including error estimates is urgently needed: a coherent and qualitatively satisfying theoretical description of resolved photo-production necessitates a PDF fit based on modern proton PDFs, with consistent settings, for example for α_S.

This is an urgently needed input to arrive at state-of-the-art baseline predictions for inclusive QCD final-state production in its dominant channel and it will ultimately provide us with a theoretical sound baseline for inclusive QCD studies at the EIC.
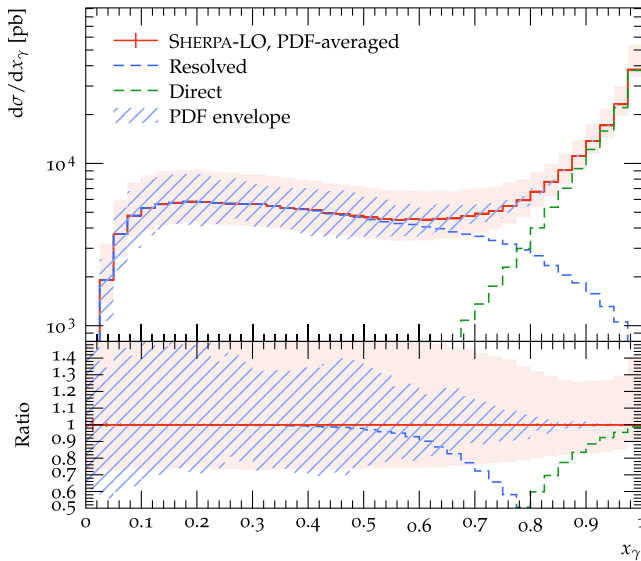

FIG. 4.   Distributions of x_γ of the SHERPA simulation at parton-level with LO accuracy and without MPI effects, averaged over 11 photon PDF sets. The PDF envelope shows the per-bin variation among the PDF sets.

ACKNOWLEDGMENTS

---

[1] R. Abdul Khalek *et al.*, Science requirements and detector concepts for the electron-ion collider: EIC Yellow report, Nucl. Phys. **A1026**, 122447 (2022).

[2] Andy Buckley *et al.*, General-purpose event generators for LHC physics, Phys. Rep. **504**, 145 (2011).

[3] Stefan Hoeche, Frank Krauss, and Peter Meinzinger, Resolved photons in SHERPA, Eur. Phys. J. C **84**, 178 (2024).

[4] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F Siegert, and J. Winter, Event generation with SHERPA 1.1, J. High Energy Phys. 02 (2009) 007.

[5] Enrico Bothmann *et al.*, Event generation with SHERPA 2.2, SciPost Phys. **7**, 034 (2019).

[6] Frank Krauss, Ralf Kuhn, and Gerhard Soff, AMEGIC++ 1.0: A matrix element generator in C++, J. High Energy Phys. 02 (2002) 044.

[7] Tanju Gleisberg and Stefan Höche, COMIX, a new matrix element generator, J. High Energy Phys. 12 (2008) 039.

[8] Tanju Gleisberg and Frank Krauss, Automating dipole subtraction for QCD NLO calculations, Eur. Phys. J. C **53**, 501 (2008).

[9] Federico Buccioni, Jean-Nicolas Lang, Jonas M. Lindert, Philipp Maierhöfer, Stefano Pozzorini, Hantian Zhang, and Max F. Zoller, OpenLoops 2, Eur. Phys. J. C **79**, 866 (2019).

[10] Steffen Schumann and Frank Krauss, A parton shower algorithm based on Catani-Seymour dipole factorisation, J. High Energy Phys. 03 (2008) 038.

[11] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert, A critical appraisal of NLO + PS matching methods, J. High Energy Phys. 09 (2012) 049.

[12] Stefano Frixione and Bryan R. Webber, Matching NLO QCD computations and parton shower simulations, J. High Energy Phys. 06 (2002) 029.

[13] S. Frixione, P. Nason, and B. R. Webber, Matching NLO QCD and parton showers in heavy flavour production, J. High Energy Phys. 08 (2003) 007.

[14] Torbjörn Sjöstrand and Maria van Zijl, A multiple-interaction model for the event structure in hadron collisions, Phys. Rev. D **36**, 2019 (1987).

[15] Gerhard A. Schuler and Torbjorn Sjostrand, Hadronic diffractive cross-sections and the rise of the total cross-section, Phys. Rev. D **49**, 2257 (1994).

[16] Gurpreet Singh Chahal and Frank Krauss, Cluster hadronisation in SHERPA, SciPost Phys. **13**, 019 (2022).

[17] Gerhard A. Schuler and Torbjörn Sjöstrand, Low- and high-mass components of the photon distribution functions, Z. Phys. C **68**, 607 (1995).

[18] Gerhard A. Schuler and Torbjörn Sjöstrand, Parton distributions of the virtual photon, Phys. Lett. B **376**, 193 (1996).

[19] Richard D. Ball *et al.*, The PDF4LHC21 combination of global PDF fits for the LHC Run III, J. Phys. G **49**, 080501 (2022).

[20] Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordström, Ben Page, Martin Rüfenacht, Marek Schönherr, and Graeme Watt, LHAPDF6: Parton density access in the LHC precision era, Eur. Phys. J. C **75**, 132 (2015).

[21] Christian Bierlich *et al.*, Robust independent validation of experiment and theory: RIVET version 3, SciPost Phys. **8**, 026 (2020).

[22] J. Adam *et al.*, ATHENA detector proposal—a totally hermetic electron nucleus apparatus proposed for IP6 at the electron-ion collider, J. Instrum. **17**, P10019 (2022).

[23] S. Chekanov *et al.*, Dijet photoproduction at HERA and the structure of the photon, Eur. Phys. J. C **23**, 615 (2002).

[24] H. Abramowicz *et al.*, Inclusive-jet photoproduction at HERA and determination of alphas, Nucl. Phys. **B864**, 1 (2012).

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