PHYSICAL REVIEW D **96,** 114005 (2017)

Nuclear structure functions at a future electron-ion collider

E. C. Aschenauer,[1,*] S. Fazio,[1,†] M. A. C. Lamont,[1,‡] H. Paukkunen,[2,3,§] and P. Zurita[1,∥]

[1]*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2]*University of Jyvaskyla, Department of Physics, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland*
[3]*Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland*

(Received 21 August 2017; published 7 December 2017)

The quantitative knowledge of heavy nuclei's partonic structure is currently limited to rather large values of momentum fraction x—robust experimental constraints below $x \sim 10^{-2}$ at low resolution scale Q^2 are particularly scarce. This is in sharp contrast to the free proton's structure which has been probed in Deep Inelastic Scattering (DIS) measurements down to $x \sim 10^{-5}$ at perturbative resolution scales. The construction of an electron-ion collider (EIC) with a possibility to operate with a wide variety of nuclei, will allow one to explore the low-x region in much greater detail. In the present paper we simulate the extraction of the nuclear structure functions from measurements of inclusive and charm reduced cross sections at an EIC. The potential constraints are studied by analyzing simulated data directly in a next-to-leading order global fit of nuclear Parton Distribution Functions based on the recent EPPS16 analysis. A special emphasis is placed on studying the impact an EIC would have on extracting the nuclear gluon parton distribution function, the partonic component most prone to nonlinear effects at low Q^2. In comparison to the current knowledge, we find that the gluon parton distribution function can be measured at an EIC with significantly reduced uncertainties.

DOI: 10.1103/PhysRevD.96.114005

I. INTRODUCTION

The deep-inelastic-scattering (DIS) experiments at the HERA collider have yielded versatile, very accurate information on the partonic structure of the free proton in a wide kinematic range [1] and contributed significantly to the theoretical advances in the sector of Quantum Chromo-Dynamics (QCD). The reach in Bjorken's x—the fraction of longitudinal momentum of the nucleon carried by the parton—goes almost down to 10^{-5} in the region of high four-momentum transfer $Q^2 \gtrsim 1\ \mathrm{GeV}^2$, where the perturbative QCD (pQCD) is applicable. The HERA experiments performed several measurements of neutral and charged-current reactions [1], as well as jet [2] and heavy-flavor cross sections [3]. These data, in varying combinations, form the backbone of all the modern global fits of free-proton Parton Distribution Functions (PDFs) [4,5]. In turn, reliable PDFs are a crucial ingredient in interpreting the measurements in hadron colliders like the Relativistic Heavy-Ion Collider (RHIC) at the Brookhaven National Laboratory (BNL) and the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN). High-precision PDFs are also essential to distinguish signals from processes beyond the Standard Model.

Notwithstanding the remarkable phenomenological success of QCD, a detailed understanding of the partonic structure of bound nuclei is still lacking. In the collinear factorized approach to pQCD, these particles are described by nuclear Parton Distribution Functions (nPDFs) [6]. Describing the fundamental constituents of the elements that make up the world we know, nPDFs are interesting in their own right. Furthermore, they are a key input for the theoretical interpretations of a large variety of ongoing and future experiments on high-energy nuclear physics, such as heavy-ion ($A + A$) and proton-nucleus (p + A) collisions at RHIC [7] and the LHC [8–10], deep-inelastic neutrino-nucleus interactions [11] and high-energy cosmic-ray interactions in the atmosphere [12]. In these cases the nPDFs characterize the initial state before the collisions and, if known accurately, can lead to the discovery of new phenomena. Moreover, a precise knowledge of nPDFs will be crucial when searching for the transition between linear and non-linear scale evolution of the parton densities [13,14]. The latter regime, known as "saturation" [15,16], occurs at low x and low interaction scale Q^2 where the recombination of low-x gluons becomes increasingly important. In lepton-nucleus ($\ell + A$) scattering such non-linearities are predicted to be more pronounced than in lepton-proton ($\ell + p$) interactions [17]. Establishing non-linear effects, is one of the key physics goals of an EIC. Altogether, the nPDFs are, and will continue to be a crucial issue in many areas of high-energy nuclear physics.

Analogous to the free proton case, $\ell + A$ scattering has a huge potential to offer information on the nPDFs [18].

---

[*] elke@bnl.gov
[†] sfazio@bnl.gov
[‡] maclamont@gmail.com
[§] hannu.paukkunen@jyu.fi
[∥] mzurita@bnl.gov

Despite some considerable effort [19,20], the HERA collider was never operated with nuclear beams and thus the kinematic reach of currently available cross-section measurements in $\ell + A$ DIS is much more restricted than in the case of protons—the existing fixed-target measurements do not reach x much below 10^{-2} in the perturbative region. As a consequence, the nPDFs are significantly less constrained than the proton PDFs.

Recently, the first global analysis of nPDFs to include LHC p + Pb Run-I data, EPPS16 [21], appeared. From the LHC data available at the time of the EPPS16 fit, the CMS dijet measurements [22] had clearly the largest impact providing additional constraints on the large-x gluons. Also data from electroweak boson production in p + Pb, collisions were used but their inclusion did not lead to significant improvements due to their limited statistical precision. The Run-II data with significantly higher luminosities are expected to provide much better constraints in the near future. However, theoretically robust LHC observables are limited to rather high Q^2 (e.g. in the case of W and Z bosons production the typical interaction scale is $Q^2 \sim 10^4$ GeV^2) and it is particularly challenging to obtain reliable constraints at the low-x, low-Q^2 domain. As already mentioned, this is the important region when it comes to differentiating linear vs nonlinear scale evolution and, in general, particularly significant for bulk observables in heavy-ion collisions, as around 90% of the particles produced at midrapidity at both RHIC ($0.002 \lesssim x \lesssim 0.4$) and the LHC ($x \lesssim 10^{-3}$) come from low-$Q^2$ processes.

To obtain gluon constraints at small x and low Q^2 from p + A collisions at the LHC or RHIC, one has to, in general, rely on observables at low transverse momentum (e.g. open charm) for which theoretical uncertainties are significant. In order to have a cleaner probe of the partonic structure of nuclei and to extend the current measurements down to smaller x, a next-generation DIS experiment is called for. To this end, two possibilities have been entertained: the LHeC collider at CERN [23] and an EIC in the United States [24]. In this paper, we will focus on the EIC project and its potential to improve the precision of nuclear PDFs. This work is organized as follows: in Sec. II we present some technical details of an EIC, relevant for the present analysis. Sections III and IV are dedicated to discuss the quantities that can be used to further the knowledge on nPDFs and showing simulation results for these. In Sec. V the impact of these measurements on the nPDFs and establishing nonlinear OCD phenomena is discussed, finally in Sec. VI our findings are summarized.

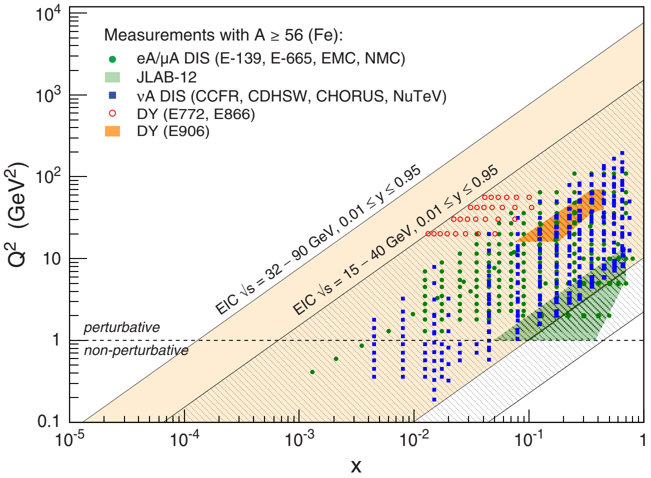

FIG. 1. The kinematic acceptance in x and Q^2 of an EIC compared to completed fixed target $\ell + A$ DIS and Drell-Yan (DY) experiments.

II. THE ELECTRON-ION COLLIDER PROJECT

Currently, there are two proposals to construct an EIC in the United States. One option would involve the addition of a hadron-accelerator complex to the existing CEBAF electron facility at the Thomas Jefferson National Laboratory (JLAB), the so-called JLEIC project [25]. The other option would be to add an electron accelerator to the existing RHIC facility at BNL, a project know as eRHIC [26]. Despite the two proposals and strategies for an EIC, the overriding goal is the same: to build a high-luminosity collider, which is flexible in terms of ion species (proton to uranium) and center-of-mass (c.o.m.) energies. Both proposals plan for a final per-nucleon c.o.m. energies ranging from 20 GeV to 90 GeV for large nuclei with an even larger range (up to 145 GeV) for polarized electron + proton ($e^- +$ p) collisions. The wide kinematic coverage of an EIC, shown in Fig. 1 in the (x, Q^2)-plane, is very important to effectively constrain nuclear PDFs. Only the eRHIC proposal for an EIC could eventually be capable of reaching top c.o.m. energy on “day 1”, whereas the JLEIC version would require a significant upgrade to reach the full c.o.m. energy. Therefore, JLEIC would stage its measurements in c.o.m. energies, starting with scanning the high and mid x region up to high Q^2 values. Both of the proposed accelerators would also be capable of reaching peak luminosities larger than 10^{34} cm^{-2} s^{-1}, three orders of magnitude higher than what was achieved at HERA. Only the JLEIC version of an EIC would be capable of reaching the peak luminosity on “day 1”, whereas eRHIC would build up its luminosity over time after upgrading the facility with hadron beam cooling. While a very large instantaneous luminosity may be required for other EIC key physics programs, this is not equally crucial for measuring structure functions. As will be described later, our study proves that, assuming collected integrated luminosity of 10 fb^{-1}, these measurements are—for the most part—not statistically limited, but rather by the associated systematic uncertainties. Therefore, a crucial aspect of this new accelerator complex is to match the high performance of a collider with a specially designed and built comprehensive DIS-specific detector in order to control systematic effects. The detector requirements come directly from the broad EIC science case. Some of the key capabilities such a detector must have are

(i) Hermetic coverage in a wide pseudorapidity range: $\sim|\eta| \leq 4$

(ii) Good scattered lepton identification and momentum resolution: in almost all cases, the DIS kinematics (x and Q^2) of the collision are most accurately calculated from the scattered electron [27]. Therefore, in order to measure these quantities as precisely as possible, an excellent particle identification as well as momentum, angular resolution and good energy resolution at very backward rapidities are required for the scattered lepton.

(iii) Good hadronic particle identification: for semi-inclusive measurements, one is also interested in identifying the hadrons produced coincidentally with the scattered lepton in the collisions. There are various techniques, which can be utilized to identify protons, pions and kaons at different momentum intervals. At low momenta, these can be identified through their specific ionization (or dE/dx) in a time projection chamber (TPC). At higher momenta, Cherenkov detectors are most widely used.

(iv) Good secondary vertex resolution: for measurements which involve heavy quarks (charm, bottom) a high resolution μ-vertex detector is essential in order to reconstruct the displaced vertices of the heavy-quark hadrons produced.

(v) High resolution and wide acceptance forward instrumentation: a Roman-pot spectrometer with almost 100% acceptance and a wide coverage in scattered proton four-momentum is crucial for studies of diffractive physics in $e+\mathrm{p}$ and $e+\mathrm{A}$ collisions. Furthermore, for $e+\mathrm{A}$ collisions, a zero-degree calorimeter (ZDC) with sufficient acceptance is a key feature vetoing on the nucleus break-up and determining the impact parameter of the collision [28].

III. REDUCED CROSS SECTION AND LONGITUDINAL STRUCTURE FUNCTION

The inclusive DIS process is a hard interaction between a lepton and a nucleon, in which the latter breaks up, the invariant mass of the hadronic final state being much larger than the nucleon mass. This is depicted in the left diagram of Fig. 2. All the relevant kinematic variables that describe the interaction are defined in Table I.

The direct observable used for constraining the nPDF is the cross section (σ), which is customarily expressed as a dimensionless quantity known as "reduced" cross section σ_{r}, defined as

$$\sigma_r \equiv \left(\frac{d^2\sigma}{dxdQ^2}\right)\frac{xQ^4}{2\pi\alpha_{\mathrm{em}}^2[1+(1-y)^2]}, \qquad (1)$$

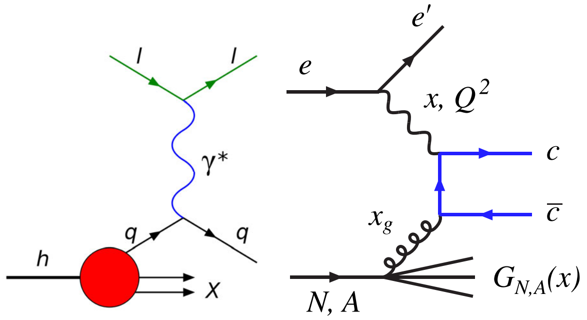

FIG. 2. Left: A depiction of inclusive DIS. Right: $c\bar{c}$ production through photon-gluon fusion.

where α_{em} is the QED fine-structure constant. At small x, the reduced cross section can be approximately expressed in terms of the structure function F_2 and the longitudinal structure function F_L as

$$\sigma_r = F_2(x,Q^2) - \frac{y^2}{1+(1-y)^2}F_L(x,Q^2). \qquad (2)$$

While F_2 is sensitive to the momentum distributions of (anti)quarks, and to gluons mainly through scaling violations, F_L has a larger direct contribution from gluons [29]. In most of the kinematical space covered by the old fixed-target DIS experiments, σ_r is dominated by F_2, to the extent that the older data were presented solely in terms of F_2, largely disregarding F_L. Therefore the information on F_L and, consequently, the direct access to the nuclear gluon are not currently available. At an EIC, the high luminosity and wide kinematic reach will enable the direct extraction of F_L and thereby more information on the behavior of the nuclear gluons can be obtained. In addition, an EIC will offer possibilities to constrain the gluon density in nuclei via measurements of the charm (bottom) structure function on which only one prior measurement exists [30]. Heavy quarks, due to their large mass, are mainly produced through photon-gluon fusion (as illustrated in the right diagram of Figure 2), the measurement of the corresponding reduced cross section $\sigma_r^{c\bar{c}}$ provides complementary information on the gluon distribution in nuclei. Also, the

TABLE I. Relevant kinematical variables in a DIS process.

Variable	Description
η	Pseudorapidity of particle
x	Fraction of the nucleon momentum carried by the struck parton
y	Inelasticity, fraction of the lepton's energy lost in the nucleon rest frame.
$\sqrt{s}$	Center-of-mass energy
Q^2	Squared momentum transferred to the lepton, equal to the virtuality of the exchanged photon Note the relation $Q^2 \approx xys$.

TABLE II. Properties of the observables.

	σ_r	F_L	$\sigma_r^{c\bar{c}}$	$F_L^{c\bar{c}}$
Kinematic coverage	Wide	Limited	Wide	Limited
Access to gluons	Mainly via scale evolution	Direct	Direct	Direct

so far unmeasured charm contribution to F_L will be measurable at an EIC.

In the production of heavy quarks, the effects of quark mass m_q require a careful treatment to preserve the genuine, dynamical effects of m_q in the partonic processes at low-Q^2 region ($Q^2 \lesssim m_q^2$), but also to have a well defined asymptotic limit ($Q^2 \gg m_q^2$). This has lead to the development of the so-called general-mass variable flavor number scheme (GM-VFNS) which is nowadays routinely implemented in proton- and nuclear-PDF extractions. The implementation of GM-VFNS is not unambiguous, but inherently contains certain scheme dependence and several versions of the GM-VFNS can be found in the literature, see. e.g. Ref. [31]. Furthermore the theory can be formulated in terms of the running of the pole mass [32,33]. The possibility of a precise measurement of heavy-flavor observables at an EIC, in particular the so far unmeasured charm contribution to F_L, will offer an opportunity to benchmark different schemes with an unprecedented precision. In addition, an EIC will take the possibilities to constrain the intrinsic heavy-flavor components in PDFs onto a completely new level.

In Table II, we summarize some properties of the observables we have discussed. For the reduced cross sections the kinematic reach is always wide and they can be measured practically everywhere within the regions indicated in Figure 1. In the case of longitudinal structure functions the kinematic range is more restricted as their extraction requires measurements at fixed x and Q^2 with several c.o.m. energies. However, the smallest values of x can only be reached at the top c.o.m. energy and thus no F_L measurement can be performed there. This will be further discussed in the next section. Also, the sensitivity to the gluon PDFs is indicated.

IV. MONTE CARLO SIMULATIONS

A. Inclusive reduced cross sections

To estimate the statistical uncertainties in measuring σ_r in $e + A$ collisions, we simulated events using the PYTHIA 6.4 [34] Monte Carlo (MC) generator with EPS09 [35] nuclear PDFs, for different beam-energy configurations corresponding to a range in c.o.m. energy from 30 to 90 GeV. We assumed the following c.o.m. energies: $\sqrt{s} = 31.6$, 44.7, and 89.4 GeV. In doing so we simulated a data collection of 2 fb^{-1} integrated luminosity at $\sqrt{s} = 31.6$ GeV and 4 fb^{-1} at $\sqrt{s} = 44.7$, 89.4 GeV respectively, corresponding to a combined 10 fb^{-1}. We divided our phase space in 5×4 bins per decade in x and Q^2. For the purpose of this study we conservatively assumed a bin-by-bin systematic uncertainty of 1.6% based on what has been achieved at HERA. We also consider an additional overall 1.4% systematic uncertainty originating from the luminosity measurement. Figure 3 shows the fraction of the statistical uncertainty over the total one (with systematics added in quadrature), per each bin in x and Q^2 for σ_r. One can see that the σ_r determination is generally dominated by the systematic uncertainties. Nevertheless, reducing the statistical uncertainty may become relevant when extending the investigation to high values of x at very high Q^2, where collecting a data sample of a significantly higher integrated luminosity may be required for precision measurements.

Figure 4 (left) shows σ_r for $e + \text{Au}$ collisions plotted versus Q^2 at different x values, for the three c.o.m. energies. A comparable precision can be achieved using any other nucleus in a similar kinematical range. The current experimental DIS-data coverage for large nuclei (A ≥ Fe) is also shown and, for clarity, σ_r is offset by subtracting $\log_{10}(x)$ and points corresponding to different energies are horizontally offset in Q^2. The bin-by-bin statistical and systematical uncertainties are added in quadrature, whereas the overall systematic uncertainty of 1.4% on the luminosity determination is not shown. The central values for the data

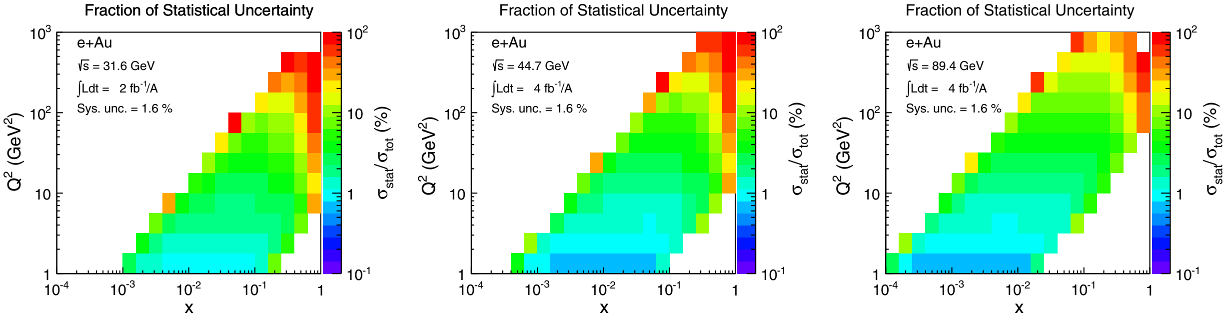

FIG. 3. Fraction of statistical uncertainty over the total uncertainty of the simulated reduced cross section measurement at an EIC, for each x and Q^2 bin, at different c.o.m. energies, assuming a combined collected luminosity of 10 fb^{-1}.

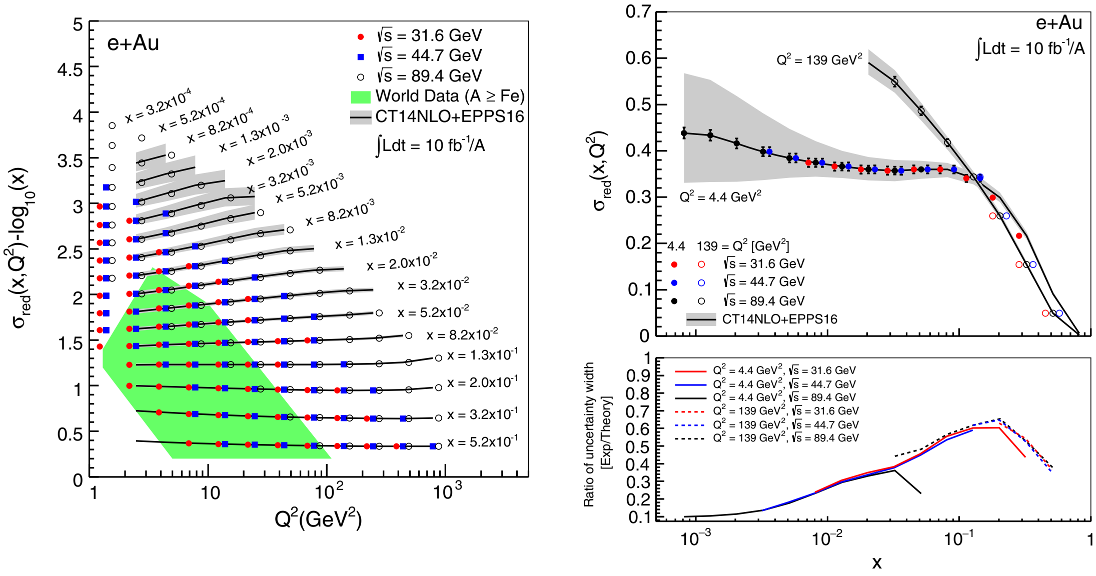

FIG. 4. The reduced cross section (left) in $e + \mathrm{Au}$ collisions at an EIC is plotted as a function of Q^2 and x, the kinematic space covered by currently available experimental data is marked on the plot by the green area. The measured reduced cross section points are shifted by $-\log_{10}(x)$ for visibility. Two examples of the σ_r (right) at Q^2 values of 4.4 GeV^2 and 139 GeV^2 are plotted versus x, with the ratio between the widths of the experimental and theoretical uncertainties shown in the bottom panel. In both plots the statistical and systematic uncertainties are added in quadrature and compared to the theoretical uncertainty (*gray bands*) from CT14NLO+EPPS16. The overall 1.4% systematic uncertainty on the luminosity determination in not shown on the plots. Points that correspond to different energy configurations are horizontally offset in Q^2 for visibility.

points have been adjusted to a next-to-leading order (NLO) calculations with CT14NLO [36] free proton PDFs supplemented with the latest nuclear modifications from EPPS16 [21].

Two examples of the σ_r as a function of x at $Q^2 = 4.4$ and 139 GeV^2 are shown in Fig. 4 (right) and compared with the theoretical uncertainties from EPPS16 and CT14NLO. The bottom panel shows the ratios between the full widths of the experimental and theoretical uncertainties versus x for the different c.o.m energies. At small x, and small Q^2 in particular, the expected uncertainties on inclusive cross-section measurements at an EIC are much smaller than those from the prediction based on EPPS16 and CT14NLO (grey band). Towards larger values of x, the existing constraints from old fixed-target experiments (SLAC and NMC in particular) do already provide stringent constraints for nPDFs and thus the advantage of EIC measurements on σ_r lies predominantly at small x. The estimated impact that these inclusive EIC data will have on the current knowledge of nuclear PDFs will be discussed in Sec. V.

B. Reduced cross section in charm production

Within the simulated data sample we have also selected $c\bar{c}$ production events by tagging K mesons which are decay products of the D mesons produced in the charm fragmentation. Figure 5 shows the momentum distribution of the decay kaons in a charm production events as a function of pseudorapidity (top plot) and the distribution of the vertex position of kaons in inclusive DIS compared with charm production events (bottom plot). One can see that kaons with a displaced vertex are coming predominantly from $c\bar{c}$ decay and are mainly produced at $\eta \leq |3|$ with momenta below 10 GeV.

Based on the vertex distribution in Fig. 5 to suppress the background from noncharm events we have requested the K to come from a vertex displaced between 0.01 and 3 cm with respect to the interaction point. Additional selection requirements on the K momentum (p_K), have been imposed to account for the η-acceptance of the particle identification (PID) detectors integrated in the EIC detector as shortly described in Sec. II. We have assumed the following K PID technologies to be in place: At mid-rapidity ($-1 < \eta < 1$), energy loss (dE/dx) in the central tracker (i.e. a time-projection chamber), and a proximity focusing Aerogel Ring-Imaging Cherenkov (RICH) detector covering the K momentum ranges $0.2\ \mathrm{GeV} < p_K < 0.8\ \mathrm{GeV}$ and $2\ \mathrm{GeV} < p_K < 5\ \mathrm{GeV}$, respectively. We considered at forward rapidities ($1 < \eta < 3.5$) a dual radiator RICH covering the kaon momentum range $2\,\mathrm{GeV} < p_K < 40\,\mathrm{GeV}$, and at backward rapidities ($-3.5 < \eta < -1$) an Aerogel RICH covering $2\ \mathrm{GeV} < p_K < 15\ \mathrm{GeV}$.

The assumed bin-by-bin systematic uncertainty for the measurement of $\sigma_r^{c\bar{c}}$ is 3.5% and it is added in quadrature to the statistical one. The point to point systematics is higher to account for the additional challenge to positively identify the kaon in the particle ID detectors. An overall 1.4%

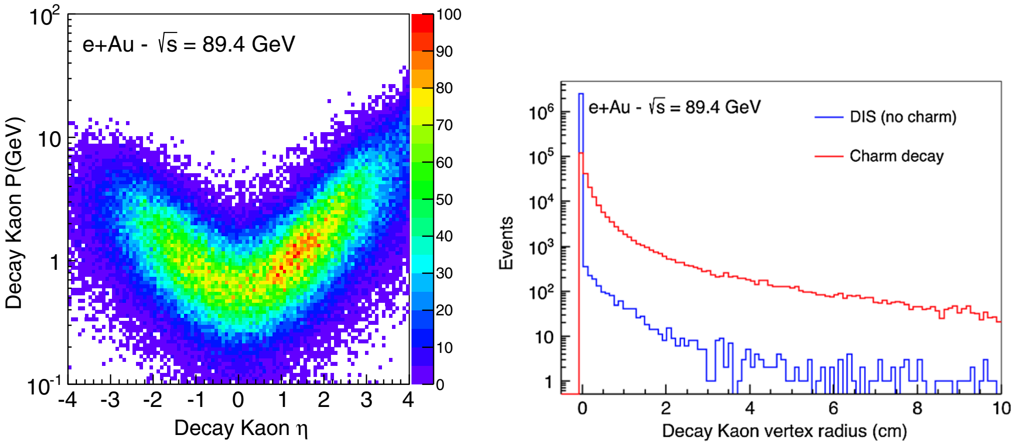

FIG. 5. Left: The distribution of the momentum of a decay K from $c\bar{c}$ production events versus pseudorapidity. Right: The vertex position of K in inclusive DIS (blue line) compared to $c\bar{c}$ production events (red line).

systematic uncertainty originating from the measurement of luminosity is also assumed.

Figure 6 (left) shows $\sigma_r^{c\bar{c}}$ plotted versus Q^2 at different x values, for the selected c.o.m. energies used earlier. For clarity, $\sigma_r^{c\bar{c}}$ is offset by subtracting $\log_{10}(x)/10$ and points that correspond to different energy configurations are horizontally offset in Q^2. Figure 6 (right) also shows two examples of the $\sigma_r^{c\bar{c}}$ as a function of x at $Q^2 = 4.4$ and 139 GeV^2. The bottom panel shows the ratios between the full widths of the experimental and theoretical uncertainties versus x for the different c.o.m energies. The assumed overall uncertainty on the luminosity is not shown on the plots. As done for the inclusive study, also for the charm production the data points have been rescaled onto the predictions from CT14NLO+EPPS16. The expected uncertainties on $\sigma_r^{c\bar{c}}$ at an EIC are much smaller than the prediction based on EPPS16 (grey band). Unlike in the inclusive case, for charm production the theoretical uncertainties clearly exceed the projected experimental ones also at large x.

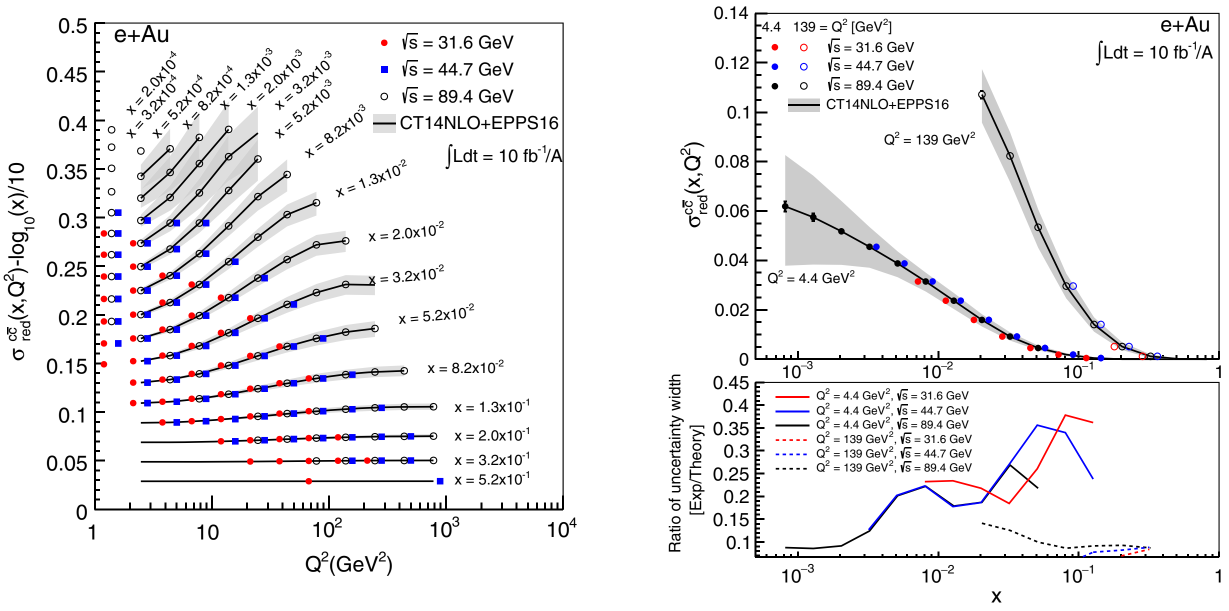

FIG. 6. The reduced cross section (*left*) of $c\bar{c}$ production in e+Au collisions at an EIC is plotted as a function of Q^2 and x. The points are shifted by $-\log_{10}(x)/10$ for visibility. Two examples of the $\sigma_r^{c\bar{c}}$ (*right*) at Q^2 values of 4.4 GeV^2 and 139 GeV^2 are plotted versus x, with the ratio between the widths of the experimental and theoretical uncertainties shown in the bottom panel. In both plots the statistical and systematic uncertainties are added in quadrature and compared to the theory uncertainty (gray bands) from CT14NLO+EPPS16. The overall 1.4% systematic uncertainty on the luminosity measurement in not shown on the plots. Points that correspond to different energy configurations are horizontally offset in Q^2 for visibility.

The efficiency of selecting $c\bar{c}$ production events has been evaluated as the ratio between the number of selected charm events and the number of all charm events simulated within the kinematical acceptance of an EIC. The overall charm selection efficiency has been estimated to be ~30% with no significant c.o.m. energy dependence. A slight rise with x was also found, but it is not significant at very small Q^2 values and becomes a little more pronounced at higher Q^2.

In order to be confident that the selection criteria used in the present study yield a sufficiently clean sample of charm production events, we studied possible background contaminations. The ratio between the number of background events with kaons in the final state passing the whole selection but not coming from a charm decay, and the signal containing only charm events has been studied. The overall background over signal ratio (B/S) has been estimated to be 0.95% ($\sqrt{s} = 31.6$ GeV), 0.98% ($\sqrt{s} = 44.7$ GeV), and 1.16% ($\sqrt{s} = 89.4$ GeV), thus showing a slight c.o.m. energy dependence, respectively. B/S has been also studied as a function of x at different Q^2 values for the selected energies and it was found to never significantly exceed 2%.

C. QED corrections

Cross section measurements with a precision as anticipated from an EIC need to account for all processes, which could alter the relation of measured to true event kinematics. The radiation of photons and the corresponding virtual corrections (QED corrections) from the incoming and outgoing lepton can cause significant effects on the reconstruction of the reduced cross section. The correction of these radiative effects can be either done through Monte Carlo techniques or including the QED effects directly in the PDF analysis.

For neutral-current $l + A$ scattering, there exists a gauge-invariant classification into leptonic, hadronic and interference contributions. The dominant correction comes from the leptonic contribution, where the photons are emitted collinear with the leptons and give rise to large logarithmic terms $\propto \log(Q^2/m_\ell^2)$, where m_ℓ is the lepton mass. In comparison to the case with no radiation, the momentum carried by the radiated photons will alter the values of x and Q^2 measured from the scattered lepton. Since the PDFs are typically very steep functions of x, even small changes can lead to large variation in the cross sections. Also the initial- and final-state quarks may radiate photons giving rise to large logarithmic terms, which are nowadays often resummed to photonic component in the PDFs. However, these corrections do not alter the event kinematics and are therefore much smaller than the contributions coming from the radiation off the leptons.

The effect of the QED radiation off the incoming and outgoing lepton can be quantified by a correction factor

$$R_C = \frac{\sigma_r(\mathcal{O}(\alpha_{\mathrm{em}}))}{\sigma_r(\mathrm{born})} - 1, \qquad (3)$$

where $\sigma_r(\mathrm{born})$ and $\sigma_r(\mathcal{O}(\alpha_{\mathrm{em}}))$ are the reduced cross section at born-level and including the first-order radiative corrections, respectively. To compute the above correction factors for σ_r and $\sigma_r^{c\bar{c}}$ for the EIC kinematics, a sample of events were generated using the DJANGO simulator [37]. The DJANGO Monte-Carlo generator was recently expanded to simulate $\ell + A$ collisions including $\mathcal{O}(\alpha_{\mathrm{em}})$ radiative effects. The simulations show that most of the radiative real photons have an energy much below 1 GeV, as shown in Fig. 7 (left). These radiative photons are typically emitted at very rear angles (in the electron going direction), see Fig. 7 (right), and are uniformly distributed in azimuthal angle.

Figure 8 shows the radiative correction factor versus the inelasticity, y, due to QED radiation in e + Au collisions at $\sqrt{s} = 89.4$ GeV for different Q^2 values, in the case of the inclusive (left plot) and charm (right plot) reduced cross sections. These values are compatible with earlier predictions [38]. In the photon-nucleon center-of-mass frame, the maximum energy of the radiated photon, $E_\gamma^{\max}$, is given by

$$E_\gamma^{\max}/\sqrt{s} \approx \sqrt{y\left(1 - \frac{Q^2}{sy}\right)} \qquad (4)$$

One can see that as the inelasticity y grows, larger QED corrections are expected, on the other hand if Q^2 grows, $E_\gamma^{\max}$ decreases and we can anticipate that the corrections get smaller. This behavior can be verified from Fig. 8 (left) in the case of the inclusive cross section. Towards small y and large Q^2 the phase space available for the photon emission becomes more and more restricted, and the correction factor falls strongly and becomes finally negative. This is a typical behavior if the phase space for photon emission becomes restricted and negative virtual corrections dominate (incomplete cancellation of infrared divergences).

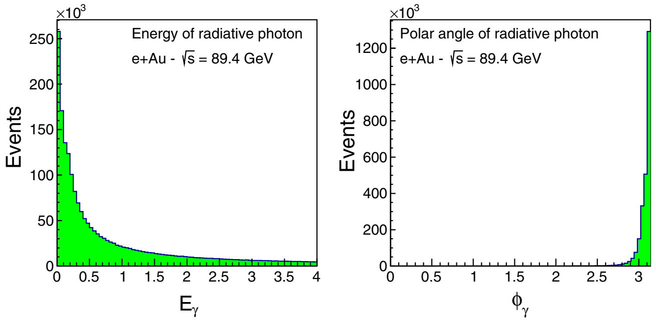

FIG. 7. The energy (left) and polar angle (right) distribution of radiative photons emitted in e + Au collision events.

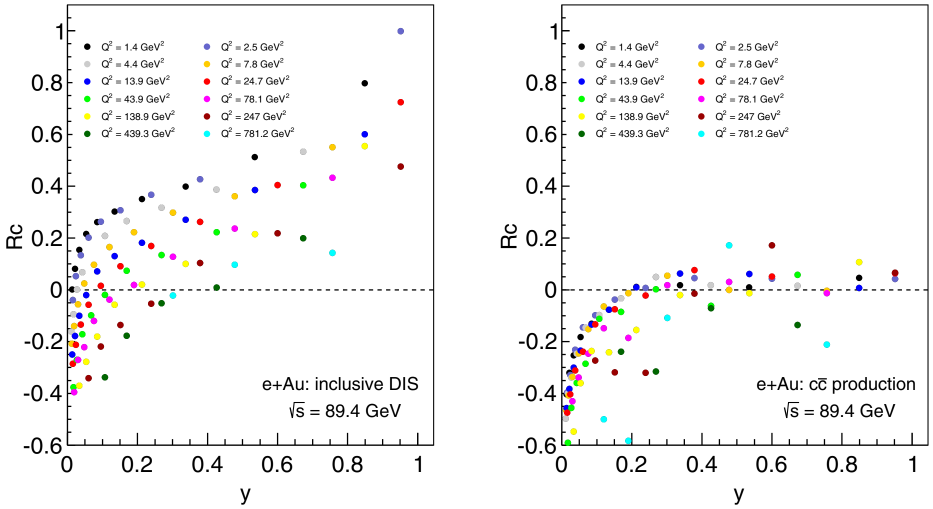

FIG. 8. The radiative correction factor to the inclusive reduced cross section (left) and the reduced cross section for charm production (right) as a function of the inelasticity, estimated for 20 GeV electrons off 100 GeV Au-ions collision events, at different Q^2 values.

The size of radiative corrections can be reduced utilizing the information about the hadronic final state. Increasing the invariant mass of the hadronic final state leads to narrower phase space available for photon emission. This is the reason why the radiative correction factor for $\sigma_r^{c\bar{c}}$ shown in Fig. 8 is significantly reduced at high y and high Q^2. For σ_r a simple cut on the invariant mass of the hadronic final state W_{had} will reduce R_C. A similar effect can be achieved cutting on $E - p_z$ (p_z: longitudinal momentum of hadronic final state particles) from the Jacquet-Blondel method [39,40]. The reduction of the radiative corrections will be considerable at largest y and at small x, but probably not yet sufficient at larger values of x.

D. The longitudinal structure function

As F_L is typically very small it is a demanding quantity to study experimentally [41–43]. It is usually extracted through a Rosenbluth separation analysis. This requires measuring σ_r for at least three different c.o.m. energies and extracting F_L from a fit of σ_r as a function of $Y^+ \equiv y^2/(1+(1-y)^2)$ for each bin in x. It is clear from Eq. (2), that the slope of this distribution represents F_L. Therefore, having at hand data with enough range in c.o.m. energy to provide a good lever arm in Y^+ will be crucial for obtaining good-quality fits and extracting precise values of F_L.

To illustrate the extraction of F_L from σ_r, Figure 9 shows the simulated σ_r in $e+\mathrm{Au}$ collisions for $\sqrt{s} = 63.2$, 77.5 and 89.4 GeV at $Q^2 = 7.81$ GeV^2 as a function of Y^+ for three different x-values. In order to ensure that the fit gives reasonable results, at least 3 points within a lever-arm in Y^+ larger than 0.1 are required. To compensate the collapsing lever-arm in Y^+ with increasing x, lower electron c.o.m. energies of $\sqrt{s} = 31.6$, 38.7 and 44.7 GeV are critical in order to reach higher-x.

As in the case of reduced cross sections, also for F_L we studied the potential of an EIC to measure both the inclusive and the charm structure functions in $e+\mathrm{Au}$ collisions. The collection of our results versus x for a number of Q^2 values is shown in Fig. 10 for F_L (left) and $F_L^{c\bar{c}}$ (right). The three different c.o.m. energies used in each extraction are also indicated on the plots. Measurements performed using a 5 GeV and a 20 GeV electron beam are indicated on the plots by open and solid circles respectively. For clarity, the values are offset by adding a constant factor C. The NLO predictions using the CT14NLO free proton PDFs with the EPPS16 nuclear modifications are shown by the gray bands. One can see that, in comparison to the PDF error bands, with a combined collected luminosity of 10 fb^{-1} at each electron beam-energy configuration, an EIC can perform a very precise measurement of the inclusive F_L and $F_L^{c\bar{c}}$

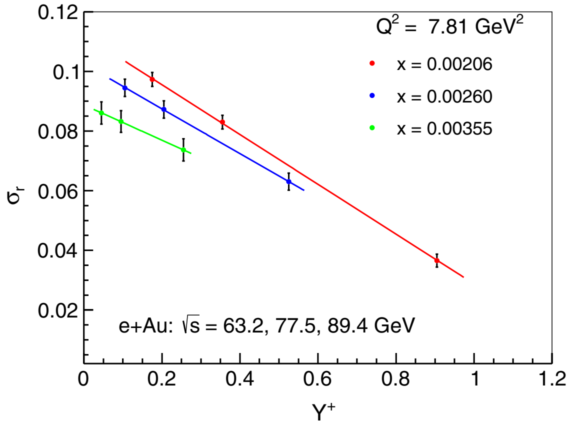

FIG. 9. The reduced cross section for $e+\mathrm{Au}$ collisions at $\sqrt{s} = 63.2$, 77.5 and 89.4 GeV versus Y^+ for $Q^2 = 7.81$ GeV^2 for three values of x. Fitting the slope of each data set with fixed x gives the negative of F_L.

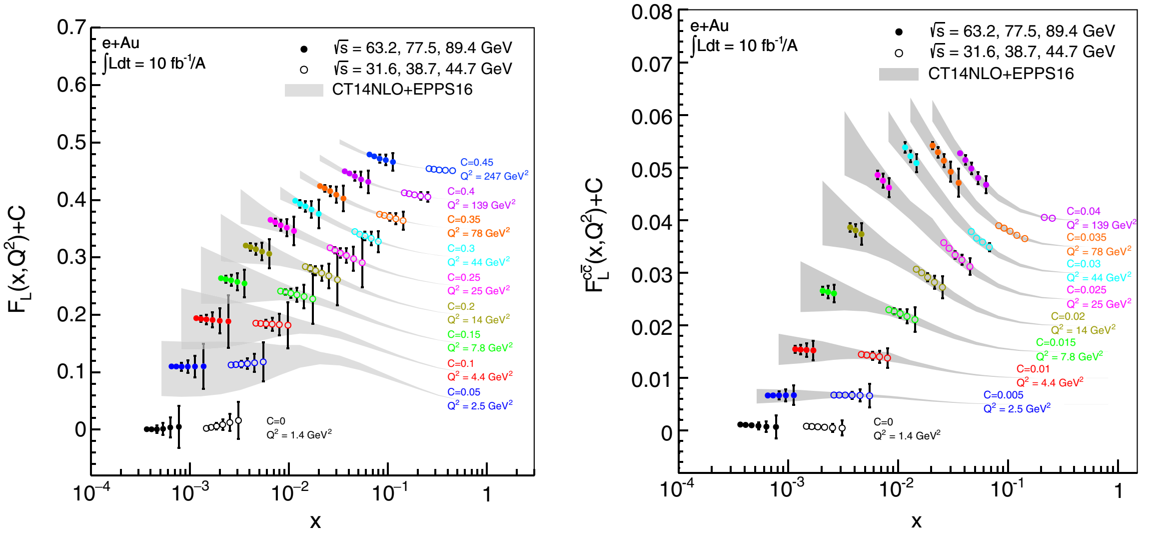

FIG. 10. Inclusive F_L (left) and $F_L^{c\bar{c}}$ (right) as a function of x for several values of Q^2. The vertical bars represent statistical and systematic uncertainties added in quadrature. The grey bands represent the theoretical predictions based on EPPS16.

in several x, Q^2 bins. With the highest c.o.m. energies, the longitudinal structure functions can be measured with a high precision down to $x \sim 7 \times 10^{-4}$ at low Q^2.

It is also important to note that an EIC can achieve a comparable precision in measuring F_L and $F_L^{c\bar{c}}$ for the proton, significantly improving the existing measurements from HERA [41,43].

V. IMPACT OF AN EIC ON NUCLEAR PDFs

A. Pseudodata for cross-section ratios

To estimate the impact an EIC would have on nuclear PDFs, we have generated a sample of pseudodata D_i for the ratios $\sigma_r^{e+\text{nucleus}}/\sigma_r^{e+\text{proton}}$. The pseudodata are based on a NLO calculation using the CT10NLO free proton PDFs and EPS09 for the nuclear modifications, denoted here by T_i^{EPS09}. The values T_i^{EPS09} were distorted in the same way as in Ref. [44] by adding Gaussian noise according to the estimated percentual point-by-point uncorrelated (δ_i^{uncorr}) and normalization uncertainties (δ_i^{norm}) as

$$D_i = T_i^{\text{EPS09}} \times [1 + \delta_i^{\text{uncorr}} r_i + \delta_i^{\text{norm}} r^{\text{norm}}], \qquad (5)$$

where r_i and r^{norm} are Gaussian random numbers with unit variance. The uncertainties from $e + A$ and $e + \text{p}$, added in quadrature, are included in δ_i^{uncorr} and δ_i^{norm}. The construction of pseudodata has been done independently for each $\sqrt{s}$ and for two nuclei, Carbon (^{12}C) and Gold (^{197}Au). We have not accounted for any experimental correlations between the pseudodata for different $\sqrt{s}$ or different nuclei, as these are difficult to estimate at this stage. The luminosity uncertainty of 1.4% is assumed to be uncorrelated for each data set with different $\sqrt{s}$ and nucleus, and it is treated in the global χ^2 minimization as in the EPPS16 analysis [21]. For clarity, in Table III we record the assumed values for systematic uncertainties discussed already earlier in Sec. IV.

B. nPDF analysis

As an EIC would extend the current kinematic reach of $e + A$ measurements to smaller values of x, it is clear that such new information would have an impact on the global extractions of nPDFs. One way to quantitatively address the improvement that an EIC would entail is to take advantage of PDF reweighting techniques [45]. However, once the new measurements probe the PDFs in a previously unconstrained kinematic range, care has to be taken that the results are not overly affected by parametrization bias. Here, our starting point is the recent global analysis of nPDFs, EPPS16 [21]. There, the nuclear modification of the proton PDF is defined as

$$R_i(x, Q^2) \equiv \frac{f_i^{\text{proton}/A}(x, Q^2)}{f_i^{\text{proton}}(x, Q^2)}, \qquad (6)$$

where $f_i^{\text{proton}/A}(x, Q^2)$ denotes the bound-proton PDF for flavor i and $f_i^{\text{proton}}(x, Q^2)$ is the corresponding free-proton PDF. The adopted x dependence was

TABLE III. The systematic uncertainties of inclusive and charm-tagged cross-section measurements. The values are in percents.

Sources of Uncertainty	Value in σ_r (%)	Value in $\sigma_r^{c\bar{c}}$ (%)
Luminosity	1.4	1.4
Electron identification and efficiency	1.6	1.6
RICH and dE/dx PID	0	3
Vertex finding	0	1

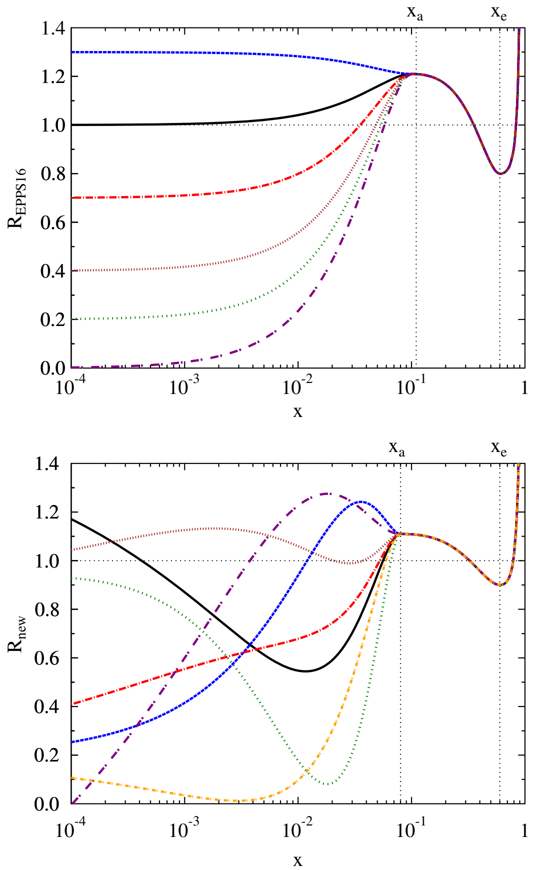

FIG. 11. Illustration of the rigidity/flexibility of the small-x fit functions used in EPPS16 analysis (upper) and in the present work (lower).

$$R_{\mathrm{EPPS16}}(x) = \begin{cases} a_0 + a_1(x - x_a)^2 & x \leq x_a \\ b_0 + b_1 x^{\alpha} + b_2 x^{2\alpha} + b_3 x^{3\alpha} & x_a \leq x \leq x_e \\ c_0 + (c_1 - c_2 x)(1 - x)^{-\beta} & x_e \leq x \leq 1. \end{cases} \tag{7}$$

In the equations above, x_a and x_e are the values of x corresponding to the assumed antishadowing maximum and EMC minimum, respectively (see Fig. 11). The rest of the parameters were adjustable but constrained such that the piecewisely defined parametrization is smooth over all x. The A dependence of the fit functions was encoded with a power-law-like parametrization at $x = x_a$, $x = x_e$, and in the case of sea quarks also in the limit $x \to 0$, see Ref. [21] for further details. Figure 11 (upper) shows some examples of how the function in Eq. (7) behaves at small x when freezing the parameters that control the region $x > x_a$. The stiffness of $R_{\mathrm{EPPS16}}(x)$ is obvious: only a monotonic decrease or increase towards $x \to 0$ is possible. Exactly the same limitation would apply also if we were to perform a PDF-reweighting study. Here, our goal is to partly release this assumption to obtain a less-biased estimate of the projected data constraints. In practice, we have replaced the EPPS16 small-x fit function in Eq. (7) by a more flexible form.

$$R_{\mathrm{new}}(x \leq x_a) = a_0 + (x - x_a)^2 \left[a_1 + \sum_{k=1}^{2} a_{k+2} x^{k/4} \right]. \tag{8}$$

Some examples of how this function can behave are shown in Fig. 11 (lower). Ideally, the same functional form should be applied to all partonic species, but in the present work we only use it for the gluons. They arguably play a special role being particularly prone to nonlinear effects at low Q^2 and also in controlling the small-x behavior of sea quarks at higher Q^2 through $g \to q\bar{q}$ splitting. In fact, an extension to all parton flavors would require a complete change in the analysis methodology which is beyond the scope of this work. This is because the Hessian method [46] that was used in the EPPS16 analysis (and is used here, too) to quantify the PDF uncertainties becomes unstable in the presence of large uncertainties and complex correlations among the fit parameters (within a single flavor and across various flavors). To overcome this limitation, Monte-Carlo techniques [47–49] should be used instead. This is left as a future work.

After adopting the more flexible functional form for the gluons also the baseline, against which the effect of an EIC should be contrasted, will be different from EPPS16. Thus, we have also performed a global nPDF fit, which is otherwise equal to the EPPS16 analysis, with the exception of the more flexible functional form of Eq. (8) to parametrize the small-x gluon nuclear modifications. As in EPPS16, the nPDF uncertainties are determined via the Hessian method [46], but in the present work the Hessian matrix is computed using the linearized prescription [50]. A fixed tolerance $\Delta\chi^2 = 50$, which corresponds approximately to the 90% confidence-level of EPPS16, was employed. We have not repeated the determination of the 90% confidence-level $\Delta\chi^2$ for all fits separately, though adding new data sets, especially with a large number of data points, has been observed to influence $\Delta\chi^2$ when it is computed on the basis of dynamical tolerance criterion. For example, in the EPPS16 analysis adding ~900 data points led to a 15-unit increase. However the uncertainty bands scale as $\sqrt{\Delta\chi^2}$, and no dramatic differences are expected from corrections on $\Delta\chi^2$.

We have studied the impact of various combinations of the EIC pseudodata: grouping different $\sqrt{s}$, using only the inclusive pseudodata, and incorporating the charm-tagged observables in addition. For the inclusive case the following energy configurations were used:

$$(E_e/\mathrm{GeV}, E_{\mathrm{p,C,Au}}/\mathrm{GeV}) = (5,50), (5,75), (5,100), (20,50), (20,75), (20,100),$$

and the charm pseudodata correspond to the setups

$$(E_e/\mathrm{GeV}, E_{\mathrm{p,C,Au}}/\mathrm{GeV}) = (5, 100),\ (20, 100).$$

While also other combinations are possible, this collection already gives a good idea of the impact. In the following the data with 5 GeV electron beam are referred to as "low-energy scenario" and the 20 GeV electron beam data as "high-energy scenario."

Figure 12 shows the resulting nuclear modifications for all partonic flavors at $Q^2 = 1.69\ \mathrm{GeV}^2$ (the parametrization scale), and Fig. 13 at $Q^2 = 10\ \mathrm{GeV}^2$ (relevant e.g. for J/Ψ production). The results are given for ^{208}Pb nucleus which is the most relevant one for the current LHC heavy-ion program. The hatched bands represent the uncertainties from the baseline fit, the blue bands correspond to fits with inclusive data only, and the black bars to the analyses including the charm cross sections. The results are shown for both the low- and high-energy versions of an EIC. In the case of up and down quarks, the trends are quite clear—the more data are used the narrower the uncertainty bands get, up to a factor of two reduction at small x. In the case of valence quarks, this is a reflection of the fact that they get better constrained at $x \sim 0.1$, which also leads to smaller uncertainties at small x due to the valence-quark sum rules and form of the fit function. Neither the inclusive nor the charm cross sections are sensitive to the (anti-)strange quarks. As a result, there are no significant differences in the obtained nuclear modifications. To constrain (anti-)strange quarks at an EIC, measurements of charm production in charged-current reactions (mediated by W^-) or, perhaps, semi-inclusive kaon production should be considered. For the gluons, the widths of the uncertainty bands evolve as expected: in the baseline fit the uncertainties are rather significant at all values of x and adding the inclusive EIC pseudodata brings the uncertainties down especially at small and mid x. Finally, when the charm-tagged pseudodata are incorporated into the analysis, the mid- and large-x gluons become very well determined. It is stressed that the nucleons have been assumed to carry zero intrinsic charm at the charm mass threshold $Q^2 = 1.69\ \mathrm{GeV}^2$—all the charm quarks are generated perturbatively. Allowing a nonzero charm-quark content at the starting scale would presumably reduce the impact on the gluon uncertainty. However, the experimental evidence suggests that, in practice, the charm content of the nucleons at its mass threshold should be very small [51]. It is noted no additional theoretical uncertainties have been considered choosing a specific (SACOT implementation of the GM-VFNS) charm treatment in EPPS-16. In all cases, the nPDF extraction including the high-energy data has further reduced the uncertainties compared to the fit with low-energy data only. On one hand, the kinematic reach in (x, Q^2) is better with high energies, see Fig. 1. On the other hand, the high-energy data has approximately twice as much data points, which also explains part of the improvements.

At first sight, it may appear puzzling that the uncertainties of up- and down-sea quark distributions at small-x remain rather sizable in comparison to the typical $\sim 2\%$ uncertainty of the small-x inclusive cross section data. The reason is the significant anticorrelation between the two quark flavors, which leads to cancellations in the cross sections: a larger $\bar{u}$-density is compensated by smaller $\bar{d}$-density and vice versa. The situation can be illustrated by examining the total nuclear $\bar{u}$-quark distribution $\bar{u}^A$ (for clarity the x and Q^2 arguments are suppressed),

$$\bar{u}^A = \frac{Z}{A} R_{\bar{u}} \bar{u}^{\mathrm{proton}} + \frac{A-Z}{A} R_{\bar{d}} \bar{d}^{\mathrm{proton}}, \tag{9}$$

where A is the nuclear mass number, and Z the number of protons. We can decompose $\bar{u}^A$ in terms of the average modification $R_{\bar{u}+\bar{d}}$,

$$R_{\bar{u}+\bar{d}} \equiv (R_{\bar{u}} \bar{u}^{\mathrm{proton}} + R_{\bar{d}} \bar{d}^{\mathrm{proton}})/(\bar{u}^{\mathrm{proton}} + \bar{d}^{\mathrm{proton}}) \tag{10}$$

and the difference $\delta R_{\bar{u}-\bar{d}}$,

$$\delta R_{\bar{u}-\bar{d}} \equiv R_{\bar{u}} - R_{\bar{d}}, \tag{11}$$

as

$$\begin{aligned} \bar{u}^A = {} & R_{\bar{u}+\bar{d}} \left(\frac{Z}{A} \bar{u}^{\mathrm{proton}} + \frac{A-Z}{A} \bar{d}^{\mathrm{proton}} \right) \\ & + \delta R_{\bar{u}-\bar{d}} \left(\frac{2Z}{A} - 1 \right) \frac{\bar{u}^{\mathrm{proton}}}{1 + \bar{u}^{\mathrm{proton}}/\bar{d}^{\mathrm{proton}}}. \end{aligned} \tag{12}$$

For an isoscalar nucleus (like ^{12}C), the last term in Eq. (12) is zero and thus the cross sections are not sensitive to the flavor separation. For nonisoscalar nuclei (like ^{197}Au) the last term in Eq. (12) is nonzero, but merely a correction to the leading term proportional to $R_{\bar{u}+\bar{d}}$. Indeed, at small x, $\bar{u}^{\mathrm{proton}} \approx \bar{d}^{\mathrm{proton}}$, and the term proportional to $\delta R_{\bar{u}-\bar{d}}$ is suppressed by a factor of $(Z/A - 1/2) \approx -0.1$ for ^{197}Au in comparison to the $R_{\bar{u}+\bar{d}}$ term. As a consequence, the sensitivity to flavor decomposition is always reduced in inclusive cross sections. A similar reasoning applies also in the case of valence quarks and explains the poor flavor decomposition at large x despite the addition of high-precision EIC pseudodata. To gain a better sensitivity to the flavor decomposition, wisely chosen differences of cross sections or structure functions in neutral- and charged-current reactions could lead to a partial cancellation of the $R_{\bar{u}+\bar{d}}$ terms, thereby increasing the importance of $\delta R_{\bar{u}-\bar{d}}$ terms (or equivalent for valence quarks).

For the presence of significant anticorrelation, it can be expected that the flavor-averaged nuclear modifications for valence quarks

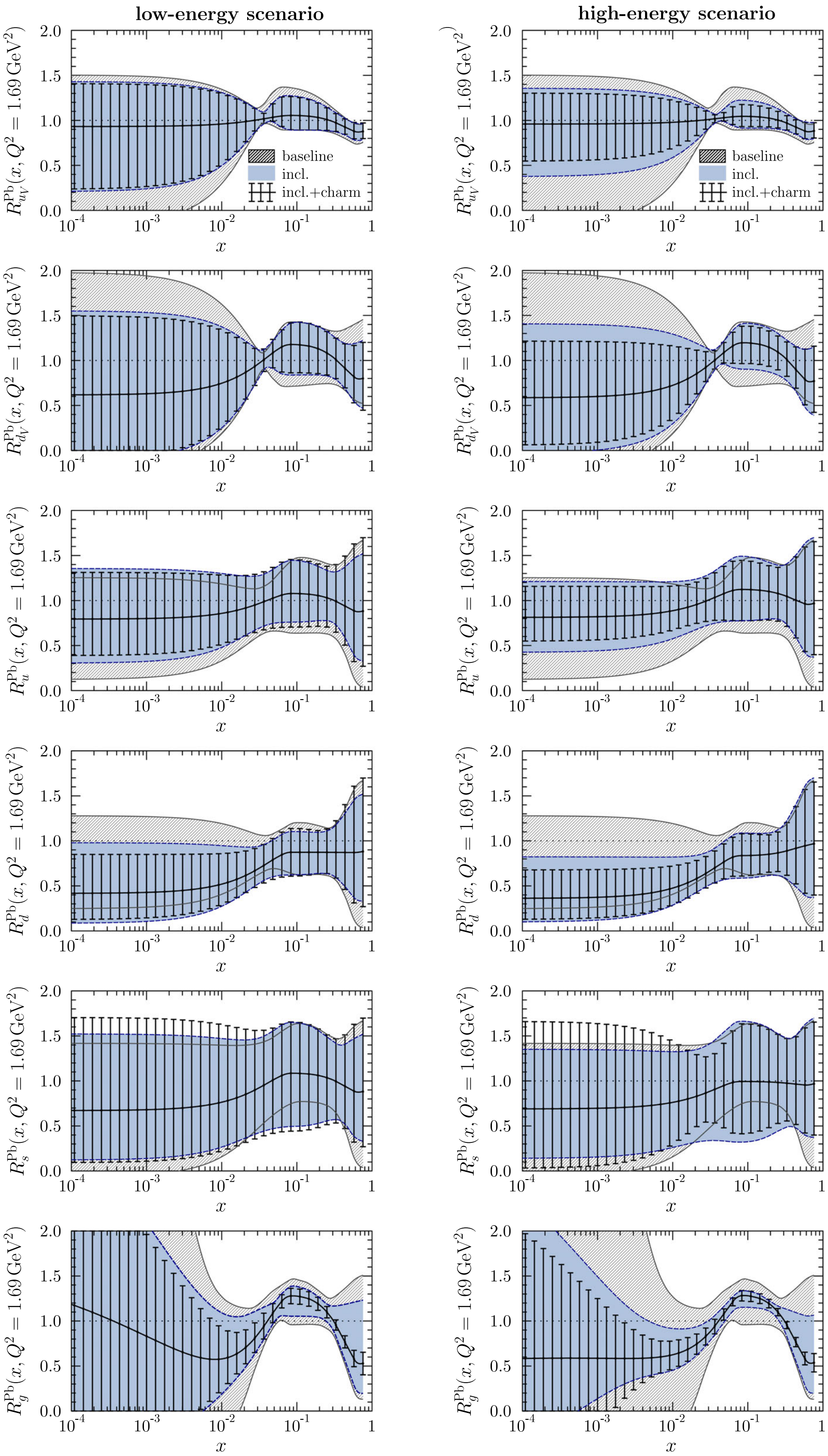

FIG. 12. Results for the nuclear modifications of Pb at $Q^2 = 1.69\ \mathrm{GeV}^2$. The hatched bands correspond to the baseline fit, the blue bands are the results from fits with no charm data included, and the black error bands denote the full analysis with inclusive and charm data.

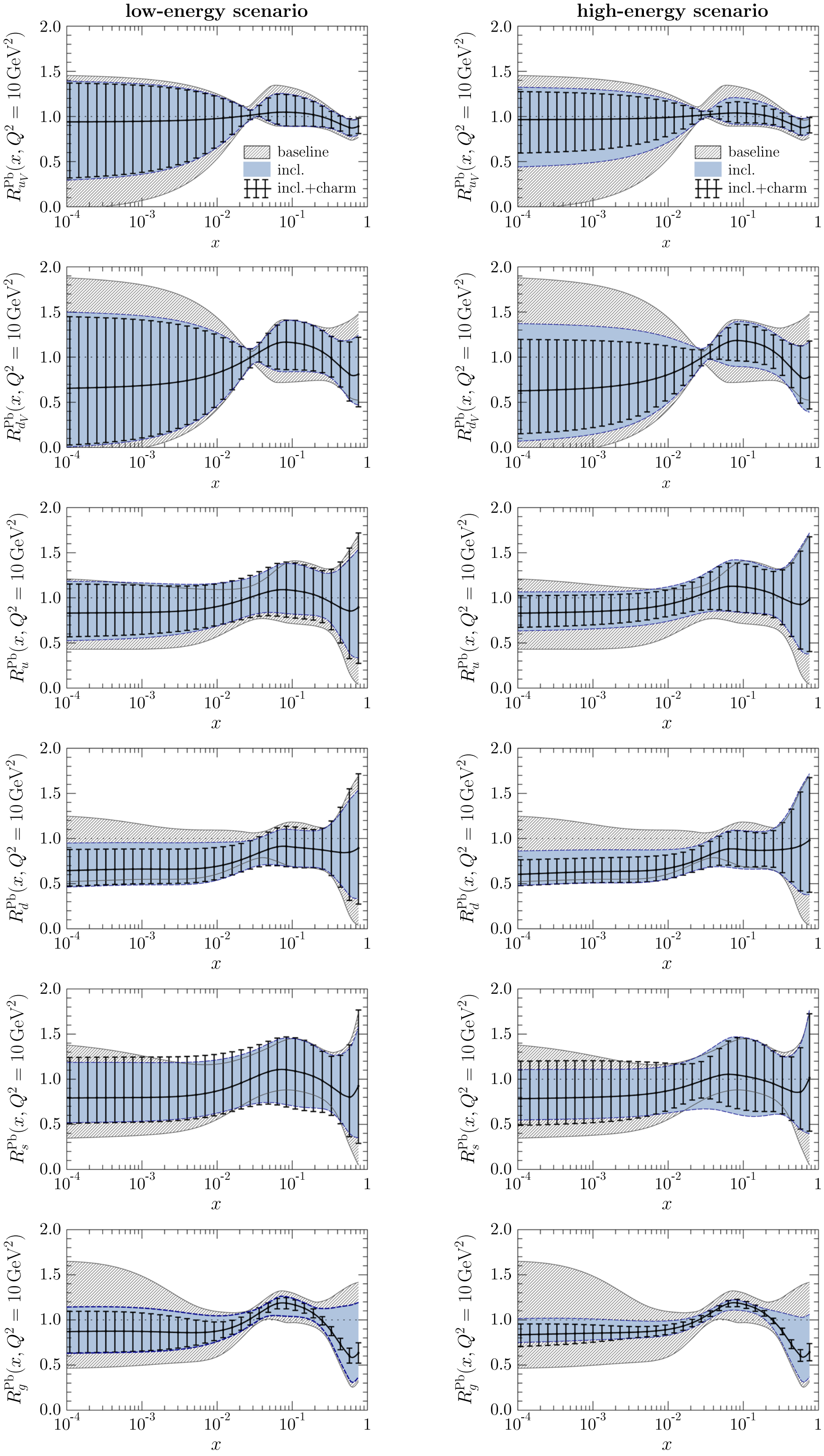

FIG. 13. As Fig. 12 but at $Q^2 = 10\ \mathrm{GeV}^2$.

$$R_V^{\mathrm{Pb}}(x,Q^2) \equiv \frac{f_{u_{\mathrm{valence}}}^{\mathrm{proton}/A} + f_{d_{\mathrm{valence}}}^{\mathrm{proton}/A}}{f_{u_{\mathrm{valence}}}^{\mathrm{proton}} + f_{d_{\mathrm{valence}}}^{\mathrm{proton}}}, \qquad (13)$$

and for light sea quarks,

$$R_S^{\mathrm{Pb}}(x,Q^2) \equiv \frac{f_{\bar{u}}^{\mathrm{proton}/A} + f_{\bar{d}}^{\mathrm{proton}/A} + f_{\bar{s}}^{\mathrm{proton}/A}}{f_{\bar{u}}^{\mathrm{proton}} + f_{\bar{d}}^{\mathrm{proton}} + f_{\bar{s}}^{\mathrm{proton}}}, \qquad (14)$$

will be much better constrained. We note that these two flavor-independent functions (plus the gluon modification) are what the nPDF fits (e.g. EPS09 [35], DSSZ [52]) have traditionally parametrized. In the latest global analyses, nCTEQ15 and EPPS16, this practice has been abandoned as being too restrictive. Presumably this will be the case for all the future global fits of nPDFs. The results are shown in Fig. 14, which presents the flavor-averaged quark nuclear modifications. For R_V, the differences between the baseline and EIC fits remain always quite modest. We recall that the rather small uncertainty at small x is a pure parametrization bias as the functional form was made more flexible only for gluons. For R_S the impact of EIC data is larger, especially at small x, Q^2 above the parametrization scale. In fact, for the baseline fit, the uncertainty is clearly larger at $Q^2 = 10\ \mathrm{GeV}^2$ than at the parametrization scale. This results from the very large gluon uncertainty in the baseline fit at $Q^2 = 1.69\ \mathrm{GeV}^2$, which partially transmits to sea quarks through the partonic scale evolution. Therefore, the effect of EIC pseudodata is to suppress the small-x uncertainty of R_S up to a factor of four at $Q^2 = 10\ \mathrm{GeV}^2$, even though the improvement is less sizable at the parametrization scale $Q^2 = 1.69\ \mathrm{GeV}^2$. However, it should be kept in mind that,

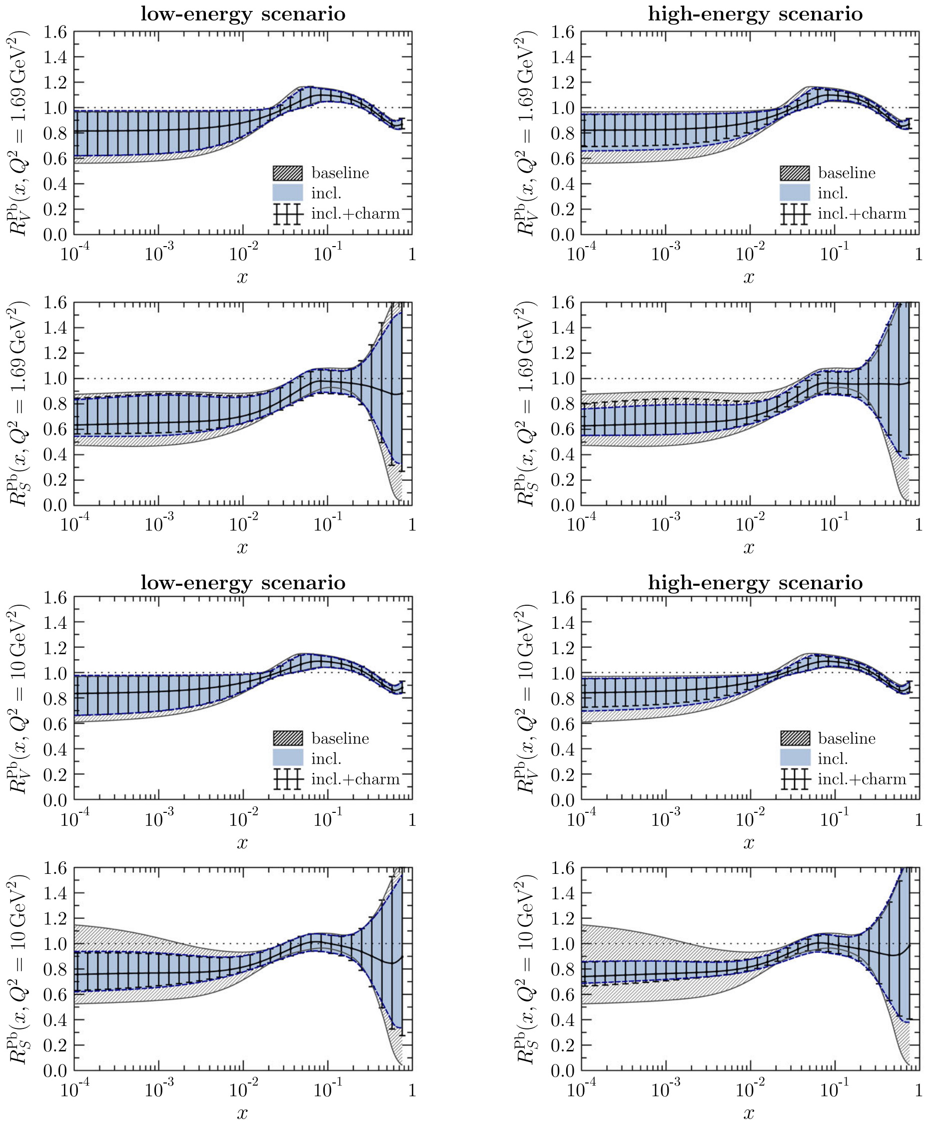

FIG. 14. As Fig. 12, but for average valence (upper panels) and average light sea quarks (lower panels). The upper set of four panels corresponds to $Q^2 = 1.69\ \mathrm{GeV}^2$ and the lower set of panels to $Q^2 = 10\ \mathrm{GeV}^2$.

similarly to the case of R_V, the small-x uncertainties of R_S, particularly at the parametrization scale $Q^2 = 1.69\ \mathrm{GeV}^2$, are artificially small due to the stiff original functional forms.

C. Impact on theoretical predictions

Henceforth the impact of the significantly improved nPDF uncertanties on theoretical predictions of experimental observables is discussed. Figure 15 shows examples of the ratios of inclusive reduced cross sections $\sigma_r(e+\mathrm{C})/\sigma_r(e+\mathrm{p})$ and $\sigma_r(e+\mathrm{Au})/\sigma_r(e+\mathrm{p})$ for $\sqrt{s} \approx 89.4$ GeV. They are compared with the predictions from the baseline fit, and with both the fits using only the low-energy pseudodata, and including also the high-energy EIC pseudodata. In comparison to the baseline fit, the low-energy EIC fit leads to clearly reduced uncertainties at small x. The inclusion of high-energy data reduces the uncertainties further by another factor of two at smallest values of x. Towards large values of x, the impact of EIC pseudodata gradually decreases as the constraints from the old fixed-target data start to dominate.

As already discussed, the inclusion of charm-tagged cross sections clearly improves the gluon constraints at large x. While the charm data nominally reaches equally small values of x as the inclusive data, the produced charm quarks always originate from "parent" gluons (via $g \to c\bar{c}$ splittings) with clearly higher x. Furthermore, the charm measurements range up to $x \sim 0.3$. Thus, it is not surprising that it is predominantly the large-x region for gluons that gets better determined by the charm data. Examples of the ratios of charm reduced cross sections corresponding to $\sqrt{s} \approx 89.4$ GeV are shown in Fig. 16. The data are

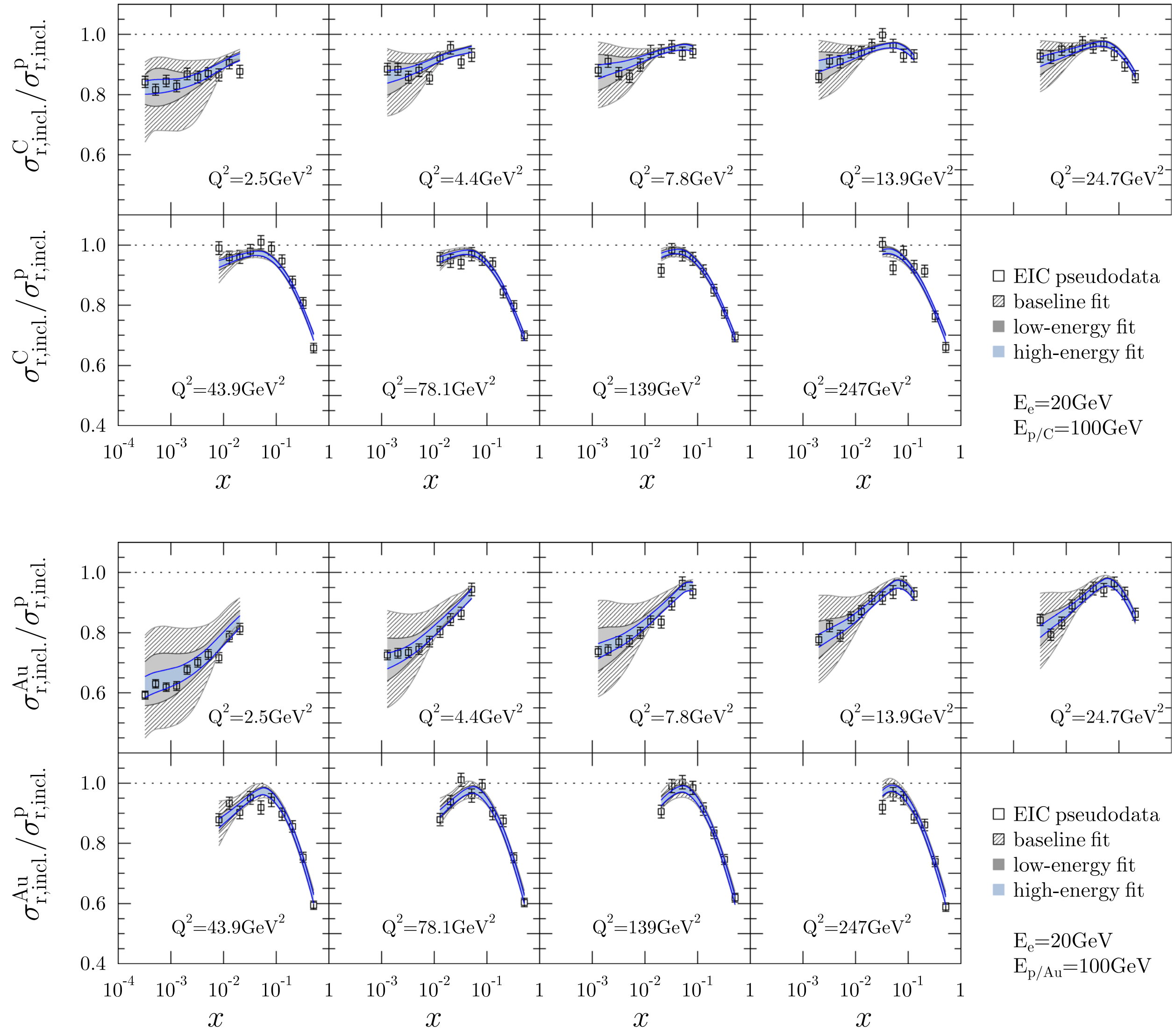

FIG. 15. The inclusive EIC pseudodata (in $E_e = 20$ GeV, $E_{\mathrm{p,C,Au}} = 20$ GeV setup) for carbon (upper panels) and gold (lower panels) compared with the baseline fit (hatched bands), the fit with inclusive low-energy data only (gray bands), and the fit with the inclusive low- and high-energy data (blue bands). The assumed overall 1.4% data normalization uncertainty is not shown.

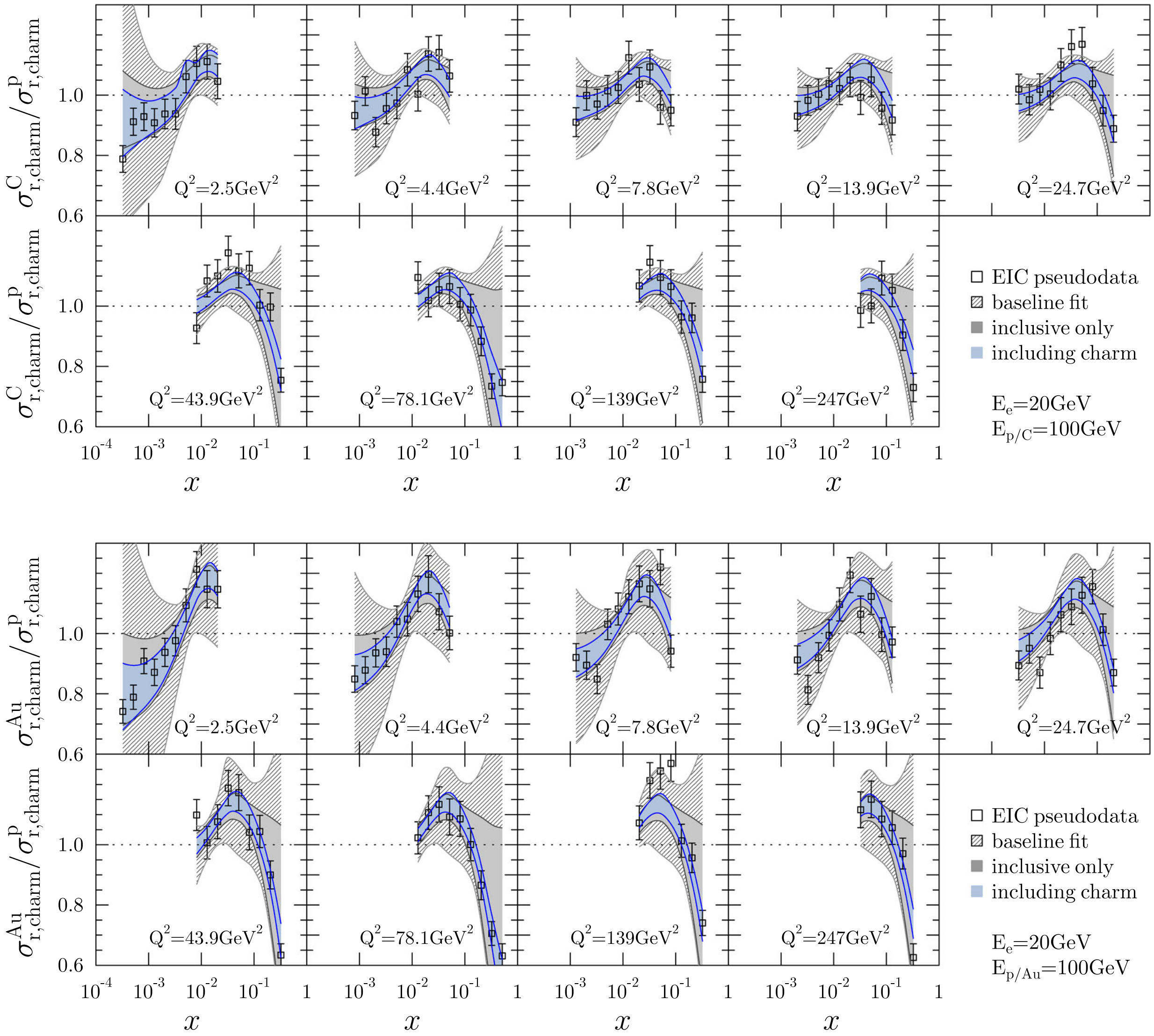

FIG. 16. The charm-tagged EIC pseudodata (in $E_e = 20$ GeV, $E_{\mathrm{p,C,Au}} = 20$ GeV setup) for carbon (upper panels) and gold (lower panels) compared with the baseline fit (hatched bands), the fit with inclusive EIC data included (gray bands), and charm data included (blue bands). The overall 1.4% normalization uncertainty is not shown.

compared with the baseline fit including only inclusive data, and the full analysis with the charm data. The baseline-fit errors (hatched bands) clearly exceed the estimated data uncertainties and already the addition of inclusive EIC data reduces the uncertainties quite a bit (gray bands). The inclusion of charm data shrinks the uncertainties further especially at large x (blue bands).

While the jet production at the LHC is known to constrain high-x gluons [53], it is unlikely that a precision like the one obtained here could be reached. Potential constraints on nPDFs through jet production at an EIC have been recently investigated [54]. However, at large-x, the jets in DIS originate predominantly from valence quarks. This is in contrast to the charm cross sections in which the contributions of valence quarks start to appear only at next-to-next-to leading order in pQCD. Thus, the charm production will be one of the key measurement for the large-x gluons and will shed light on the size of intrinsic charm component in heavy nuclei.

The EPPS16 analysis is currently the only available parametrization to include constraints from the LHC Run-1 p + Pb data. We would like to point out that there is a significant complementarity between these LHC measurements and measurements at an EIC. To illustrate this point we present in Fig. 17 the LHC p + Pb data on W [55], Z [56,57], and dijet [22] production included in the EPPS16 fit (thereby also in the fits presented here). The inclusive pion production data measured by PHENIX [58] at RHIC are shown as well. The data are compared with the baseline fit and the full high-energy EIC analysis. The reduction of the uncertainties upon including the EIC pseudodata is quite dramatic and concretely demonstrates how an EIC

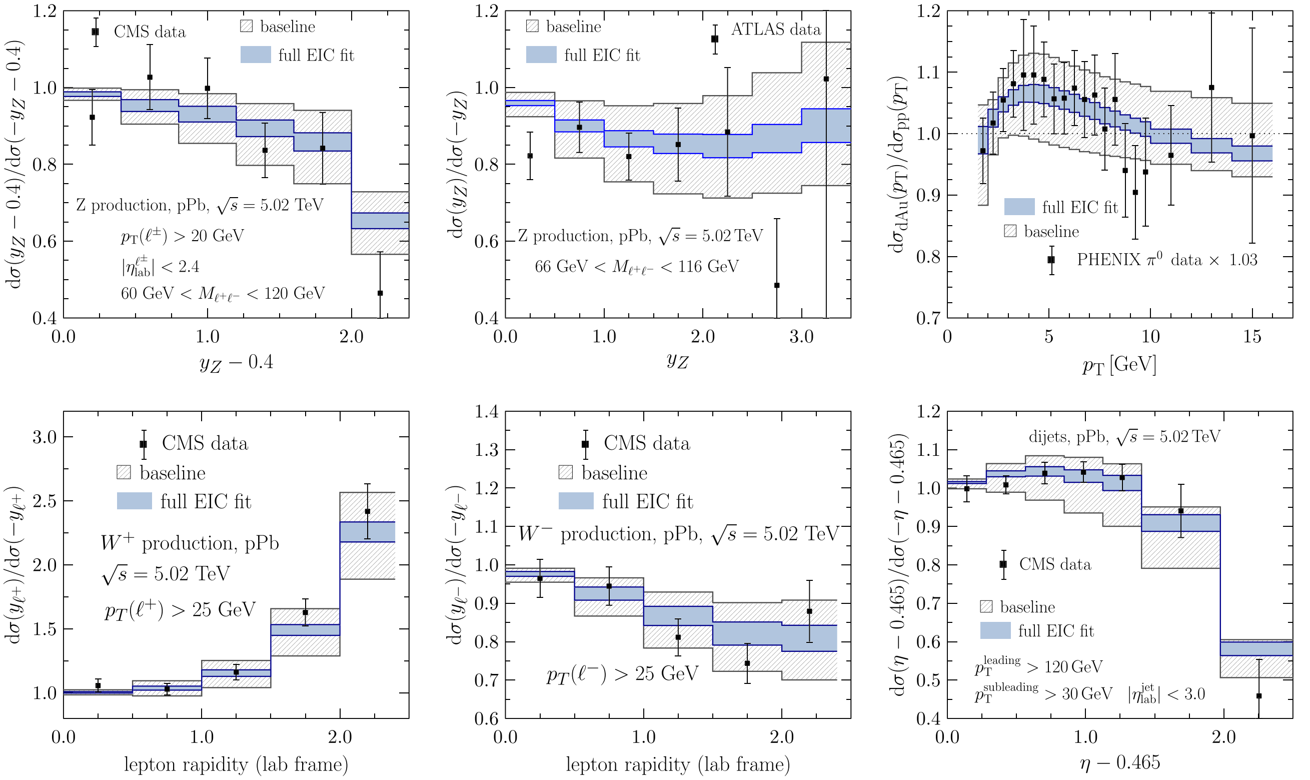

FIG. 17. The LHC p–Pb and RHIC D–Au data for Z (two upper left panels), W$^\pm$ (two lower left panels), dijet production (lower right panel), and inclusive pion production [58] (upper right panel) compared with the baseline fit (hatched bands) and the full high-energy EIC analysis (blue bands).

and the LHC can complement each other. This is important in order to truly and precisely address the universality of nPDFs. It should be stressed that the Q^2 in typical LHC p + Pb observables is much higher than the ones probed at an EIC. Thus, the constraints on observables at low Q^2 from these LHC measurements are only scarce. The theoretical uncertainties on LHC observables that probe the low-Q^2 and low-x domain (e.g. open charm, exclusive J/Ψ) are always bound to be large and in order to obtain reliable constraints a DIS experiment like an EIC is crucial.

D. Relation to nonlinear QCD

As discussed in Sec. I, establishing nonlinear QCD phenomena, such as gluon saturation, is one of the key physics goals of an EIC. It has been shown [24,59] that to unambiguously identify saturation at an EIC requires several complementary measurements. These include the measurement of dihadron correlations [60], exclusive diffractive vector meson production [61] as well as the difference in production cross section for inclusive diffractive cross sections [24,59]. Predictions based from current saturation-models show nonlinear effects should set in for a gold nucleus at $x < 10^{-3}$ for $Q^2 = 1$ GeV^2 (see Fig. 3.9 in Ref. [24]). Recent calculations [62,63] demonstrate that the inclusion of small-x resummation to NLO and NNLO fixed-order calculations leads to a quantitative improvement in the perturbative description of the HERA inclusive and charm-production reduced cross sections. The underlying physics mechanism behind the resummation is the BFKL dynamics [64] which, in comparison to fixed-order QCD, may also have similar effects as nonlinear QCD phenomena. Reference [14] describes a study accessing nonlinear effects at an EIC by comparing the behaviour of nPDFs extracted in different kinematic regions with and without sizable nonlinear effects. The authors find that F_2 at low x and Q^2 would be sensitive to nonlinear effects. At the values of x at an EIC, the predictions of saturation models for F_L are already distinctively different from those of collinear factorization. Unfortunately, the authors used F_L values, which did not cover the full $x - Q^2$ range accessible at an EIC as well as the total experimental uncertainties of F_L are significant larger in certain $x - Q^2$ bins than the ones presented in this paper. The study concluded that F_L, despite its direct sensitivity to gluons, would not be sensitive to saturation effects. It would be interesting to repeat this analysis based on the F_L values presented here. An alternative way to study nonlinear effects is a phenomenological study of the QCD scale evolution of DIS cross sections within the framework of physical anomalous dimensions. There one would observe deviations from the scale evolution governed by the physical anomalous dimensions, which will unambiguously quantify the size and relevance of non- linear effects caused by an abundance of gluons with small momentum fractions [65]. A experimentally driven way to study nonlinear effects is studying the ratio of F_2^A/F_2^p and other structure functions as function of A and tagging the centrality of the collision [28]. A particularly strong impact-parameter dependence of the ratio $F_2^A/F_2^{\rm p}$ in central collisions towards small x and low Q^2 could be a sign of nonlinear effects becoming dominant, as the estimated impact-parameter dependence of nPDFs is there only rather weak [66].

VI. SUMMARY

We have studied the inclusive and charm cross-section measurements at an EIC and especially their impact on the global in the framework of EPPS16 using projected pseudodata for an EIC. A special attention was paid on the determination of gluon densities for which an extended small-x parametrization was used. It was shown that an EIC will have an enormous impact on the global extractions of nPDFs, particularly on the gluon which is currently only weakly constrained. At low resolution scale $Q^2 = 1.69\ \mathrm{GeV}^2$ the gluon distribution can be determined well down to $x \sim 10^{-2}$ but towards higher Q^2, the small-x uncertainties quickly shrink across all the small-x domain. The inclusion of the charm-tagged cross-section measurements decreases the gluon uncertainties substantially at large x. For the quark sector our study is somewhat more limited as our current analysis methodology does not permit the use of more flexible parametrizations for all the quark flavors simultaneously. Despite this limitation, our results indicate, especially at Q^2 above the parametrization scale, a significant reduction of the sea-quark uncertainties.

The high precision and the wide kinematic coverage in x and Q^2 achievable for different observables at an EIC will allow for stringent tests of the nPDF universality. Ultimately, such an endeavor requires a combination of complementary results from the LHC, RHIC and elsewhere.

ACKNOWLEDGMENTS

We are very grateful to the EIC group at BNL whose ongoing efforts made this analysis possible. E. C. A., S. F., and P. Z. acknowledge the support by the U.S. Department of Energy under Contract No. DE-SC0012704. The work of H. P. is currently supported by the Academy of Finland, Project No. 308301. In addition, H. P. acknowledges the funding from the Academy of Finland, Project No. 297058; the European Research Council Grant No. HotLHC ERC-2011-StG-279579; Ministerio de Ciencia e Innovación of Spain and FEDER, Project No. FPA2014-58293-C2-1-P; Xunta de Galicia (Conselleria de Educacion)—he has been part of the Strategic Unit AGRUP2015/11.

---

[1] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data, Eur. Phys. J. C **75**, 580 (2015).

[2] V. Andreev *et al.* (H1 Collaboration), Measurement of jet production cross sections in deep-inelastic ep scattering at HERA, Eur. Phys. J. C **77**, 215 (2017).

[3] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination and QCD analysis of charm production cross section measurements in deep-inelastic ep scattering at HERA, Eur. Phys. J. C **73**, 2311 (2013).

[4] S. Forte and G. Watt, Progress in the determination of the partonic structure of the proton, Annu. Rev. Nucl. Part. Sci. **63**, 291 (2013).

[5] J. Rojo *et al.*, The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II, J. Phys. G **42**, 103103 (2015).

[6] H. Paukkunen, Status of nuclear PDFs after the first LHC p-Pb run, Nucl. Phys. **A967**, 241 (2017).

[7] E. C. Aschenauer *et al.*, The RHIC cold QCD plan for 2017 to 2023: A portal to the EIC, arXiv:1602.03922.

[8] N. Armesto and E. Scomparin, Heavy-ion collisions at the Large Hadron Collider: A review of the results from Run 1, Eur. Phys. J. Plus **131**, 52 (2016).

[9] A. Dainese *et al.*, CERN Yellow Report, 2017, p. 635.

[10] C. A. Salgado and J. P. Wessels, Proton–lead collisions at the CERN LHC, Annu. Rev. Nucl. Part. Sci. **66**, 449 (2016).

[11] L. Alvarez-Ruso *et al.*, NuSTEC white paper: Status and challenges of neutrino-nucleus scattering, arXiv:1706.03621.

[12] A. Bhattacharya, R. Enberg, Y. S. Jeong, C. S. Kim, M. H. Reno, I. Sarcevic, and A. Stasto, Prompt atmospheric neutrino fluxes: perturbative QCD models and nuclear effects, J. High Energy Phys. 11 (2016) 167.

[13] J. L. Albacete, J. G. Milhano, P. Quiroga-Arias, and J. Rojo, Linear vs non-linear QCD evolution: From HERA data to LHC phenomenology, Eur. Phys. J. C **72**, 2131 (2012).

[14] C. Marquet, M. R. Moldes, and P. Zurita, Unveiling saturation effects from nuclear structure function measurements at the EIC, Phys. Lett. B **772**, 607 (2017).

[15] J. Jalilian-Marian and Y. V. Kovchegov, Saturation physics and deuteron-gold collisions at RHIC, Prog. Part. Nucl. Phys. **56**, 104 (2006).

[16] J. L. Albacete and C. Marquet, Gluon saturation and initial conditions for relativistic heavy ion collisions, Prog. Part. Nucl. Phys. **76**, 1 (2014).

[17] A. H. Mueller and J. W. Qiu, Gluon recombination and shadowing at small values of x, Nucl. Phys. **B268**, 427 (1986).

[18] M. Arneodo, Nuclear effects in structure functions, Phys. Rep. **240**, 301 (1994).

[19] M. Arneodo, A. Bialas, M. W. Krasny, T. Sloan, and M. Strikman, Nuclear beams in HERA, hep-ph/9610423.

[20] T. Alexopoulos *et al.*, Reports No. DESY-03-194, No. DESY-PRC-03-02, No. H1-04-03-609.

[21] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, EPPS16: Nuclear parton distributions with LHC data, Eur. Phys. J. C **77**, 163 (2017).

[22] S. Chatrchyan *et al.* (CMS Collaboration), Studies of dijet transverse momentum balance and pseudorapidity distributions in pPb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV, Eur. Phys. J. C **74**, 2951 (2014).

[23] J. L. Abelleira Fernandez *et al.* (LHeC Study Group Collaboration), A large hadron electron collider at CERN: Report on the physics and design concepts for machine and detector, J. Phys. G **39**, 075001 (2012).

[24] A. Accardi *et al.*, Electron ion collider: The next QCD frontier: Understanding the glue that binds us all, Eur. Phys. J. A **52**, 268 (2016).

[25] S. Abeyratne *et al.*, MEIC design summary, arXiv:1504.07961.

[26] E. C. Aschenauer *et al.*, eRHIC design study: An electron-ion collider at BNL, arXiv:1409.1633.

[27] U. Bassler and G. Bernardi, Structure function measurements and kinematic reconstruction at HERA, Nucl. Instrum. Methods Phys. Res., Sect. A **426**, 583 (1999).

[28] L. Zheng, E. C. Aschenauer, and J. H. Lee, Determination of electron-nucleus collision geometry with forward neutrons, Eur. Phys. J. A **50**, 189 (2014).

[29] N. Armesto, H. Paukkunen, C. A. Salgado, and K. Tywoniuk, Nuclear effects on the longitudinal structure function at small x, Phys. Lett. B **694**, 38 (2010).

[30] J. J. Aubert *et al.* (European Muon Collaboration), Production of charmed particles in 250-GeV μ^+—iron interactions, Nucl. Phys. **B213**, 31 (1983).

[31] R. S. Thorne and W. K. Tung, PQCD formulations with heavy quark masses and global analysis, arXiv:0809.0714.

[32] S. Alekhin and S. Moch, Heavy-quark deep-inelastic scattering with a running mass, Phys. Lett. B **699**, 345 (2011).

[33] A. Accardi *et al.*, A critical appraisal and evaluation of modern PDFs, Eur. Phys. J. C **76**, 471 (2016).

[34] T. Sjostrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 physics and manual, J. High Energy Phys. 05 (2006) 026.

[35] K. J. Eskola, H. Paukkunen, and C. A. Salgado, EPS09: A new generation of NLO and LO nuclear parton distribution functions, J. High Energy Phys. 04 (2009) 065.

[36] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C.-P. Yuan, New parton distribution functions from a global analysis of quantum chromodynamics, Phys. Rev. D **93**, 033006 (2016).

[37] G. A. Schuler and H. Spiesberger, DJANGO: The interface for the event generators HERACLES and LEPTO, *Proceedings of Workshop on Physics at HERA*, edited by W. Buchmüller and G. Ingelman (DESY, Hamburg, Germany, 1991), p. 1419.

[38] D. Boer *et al.*, Gluons and the quark sea at high energies: Distributions, polarization, tomography, arXiv:1108.1713.

[39] F. Jacquet and A. Blondel, Report No. DESY 79/48 1979; in *Proceedings of the Study of an ep Facility for Europe*, edited by U. Amaldi, *Conf. Proc. C* 790402 1, (1979).

[40] U. Bassler and G. Bernardi, On the kinematic reconstruction of deep inelastic scattering at HERA, Nucl. Instrum. Methods Phys. Res., Sect. A **361**, 197 (1995).

[41] S. Chekanov *et al.* (ZEUS Collaboration), Measurement of the longitudinal proton structure function at HERA, Phys. Lett. B **682**, 8 (2009).

[42] H. Abramowicz *et al.* (ZEUS Collaboration), Deep inelastic cross-section measurements at large y with the ZEUS detector at HERA, Phys. Rev. D **90**, 072002 (2014).

[43] V. Andreev *et al.* (H1 Collaboration), Measurement of inclusive ep cross sections at high Q^2 at $\sqrt{s} = 225$ and 252 GeV and of the longitudinal proton structure function F_L at HERA, Eur. Phys. J. C **74**, 2814 (2014).

[44] D. d'Enterria, K. Krajczr, and H. Paukkunen, Top-quark production in proton–nucleus and nucleus–nucleus collisions at LHC energies and beyond, Phys. Lett. B **746**, 64 (2015).

[45] H. Paukkunen and P. Zurita, PDF reweighting in the Hessian matrix approach, J. High Energy Phys. 12 (2014) 100.

[46] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, Uncertainties of predictions from parton distribution functions. 2. The Hessian method, Phys. Rev. D **65**, 014013 (2001).

[47] W. T. Giele and S. Keller, Implications of hadron collider observables on parton distribution function uncertainties, Phys. Rev. D **58**, 094023 (1998).

[48] W. T. Giele, S. A. Keller, and D. A. Kosower, Parton distribution function uncertainties, hep-ph/0104052.

[49] G. Watt and R. S. Thorne, Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs, J. High Energy Phys. 08 (2012) 052.

[50] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Parton distributions for the LHC, Eur. Phys. J. C **63**, 189 (2009).

[51] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017).

[52] D. de Florian, R. Sassot, P. Zurita, and M. Stratmann, Global analysis of nuclear parton distributions, Phys. Rev. D **85**, 074028 (2012).

[53] H. Paukkunen, K. J. Eskola, and C. Salgado, Dijets in $\mathrm{p}+\mathrm{Pb}$ collisions and their quantitative constraints for nuclear PDFs, Nucl. Phys. **A931**, 331 (2014).

[54] M. Klasen, K. Kovarik, and J. Potthoff, Nuclear parton density functions from jet production in DIS at the EIC, Phys. Rev. D **95**, 094013 (2017).

[55] V. Khachatryan *et al.* (CMS Collaboration), Study of W boson production in pPb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV, Phys. Lett. B **750**, 565 (2015).

[56] V. Khachatryan *et al.* (CMS Collaboration), Study of Z boson production in pPb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV, Phys. Lett. B **759**, 36 (2016).

[57] G. Aad *et al.* (ATLAS Collaboration), Z boson production in $\mathrm{p}+\mathrm{Pb}$ collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV measured with the ATLAS detector, Phys. Rev. C **92**, 044915 (2015).

[58] S. S. Adler *et al.* (PHENIX Collaboration), Centrality Dependence of π^0 and eta Production at Large Transverse Momentum in $\sqrt{s_{NN}} = 200\,\mathrm{GeV}\,d+\mathrm{Au}$ Collisions, Phys. Rev. Lett. **98**, 172302 (2007).

[59] E. C. Aschenauer *et al.*, The electron-ion collider: Assessing the energy dependence of key measurements, arXiv:1708.01527.

[60] L. Zheng, E. C. Aschenauer, J. H. Lee, and B. W. Xiao, Probing gluon saturation through dihadron correlations at an electron-ion collider, Phys. Rev. D **89**, 074037 (2014).

[61] T. Toll and T. Ullrich, Exclusive diffractive processes in electron-ion collisions, Phys. Rev. C **87**, 024913 (2013).

[62] M. Bonvini, S. Marzani, and C. Muselli, Towards parton distribution functions with small-x resummation: HELL 2.0, arXiv:1708.07510.

[63] R. D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo, and L. Rottoli, Parton distributions with small-x resummation: Evidence for BFKL dynamics in HERA data, arXiv:1710.05935.

[64] J. R. Forshaw and D. A. Ross, *Quantum Chromodynamics and the Pomeron*, Cambridge Lecture Notes in Physics, 1st ed. (Cambridge University Press, 1997), p. 268.

[65] M. Hentschinski and M. Stratmann, On the practical application of physical anomalous dimensions, arXiv:1311.2825.

[66] I. Helenius, K. J. Eskola, H. Honkanen, and C. A. Salgado, Impact-parameter dependent nuclear parton distribution functions: EPS09s and EKS98s and their applications in nuclear hard processes, J. High Energy Phys. 07 (2012) 073.