

Constraints on gluon distribution functions in the nucleon and nucleus from open charm hadron production at the Electron-Ion Collider

Matthew Kelsey,^{1,2,*} Reynier Cruz-Torres,² Xin Dong,² Yuanjing Ji,²
Sooraj Radhakrishnan,^{3,2} and Ernst Sichtermann²

¹Wayne State University, Detroit, Michigan 48202, USA

²Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

³Kent State University, Kent, Ohio 44242, USA

 (Received 11 July 2021; accepted 4 August 2021; published 2 September 2021)

The Electron-Ion Collider (EIC) at Brookhaven National Laboratory will be a precision quantum chromodynamics machine that will enable a vast physics program with electron+proton/ion collisions across a broad center-of-mass range. Measurements of hard probes such as heavy flavor in deep inelastic scatterings will be an essential component to the EIC physics program and are one of the detector R&D driving aspects. In this paper we study the projected statistical precision of open charm hadron production through exclusive hadronic channel reconstruction with a silicon detector concept currently being developed using a PYTHIA-based simulation. We further study the impact of possible intrinsic charm in the proton on projected data, and estimate the constraint on the nuclear gluon parton distribution function (PDF) from the charm structure functions $F_2^{c\bar{c}}$ in $e + \text{Au}$ collisions using a Bayesian PDF reweighting technique. Our studies show the EIC will be capable of delivering an unprecedented measurement of charm hadron production across a broad kinematic region and will provide strong constraints to both intrinsic charm and nuclear gluon PDFs.

DOI: [10.1103/PhysRevD.104.054002](https://doi.org/10.1103/PhysRevD.104.054002)

I. INTRODUCTION

The Electron-Ion Collider (EIC), a US-based facility planned to be constructed at the current Relativistic Heavy-Ion Collider (RHIC) facility at Brookhaven National Laboratory, will be a state-of-the-art quantum chromodynamics (QCD) laboratory [1–3]. The EIC will be capable of delivering high-luminosity ($10^{34} \text{ cm}^{-2} \text{ s}^{-1}$) collisions of electrons on (polarized) protons, light and heavy ions over a large span of center-of-mass energies ranging from $\sqrt{s} = 20$ to 141 GeV [4]. Some of the flagship EIC measurements include studies of the spin structure of protons and light ions, the partonic structure of light and heavy ions, partonic transport in nuclear matter, and hadronization.

Measurements of heavy flavor hadrons (hadrons containing a charm or bottom quark) in deep inelastic scatterings (DIS) are valuable since at leading order (LO) heavy quark pairs are produced via photon-gluon fusion as shown in Fig. 1, and have direct and clean access to the gluonic structure of nucleons/nuclei. (Nuclear) parton

distribution functions (PDF) are an essential ingredient in understanding measurements of nuclear collisions and are of broad interest in the particle and nuclear physics communities.

Previous generations of DIS experiments have measured the reduced charm cross sections in lepton-proton collisions over a limited kinematic range with mostly $x_B < 0.1$, with x_B defined as the Bjorken scaling variable, and have been well described by perturbative QCD (pQCD) [5]. At the EIC, kinematic coverage of measurements of charm hadrons will extend to $x_B > 0.1$ as shown in Fig. 2 for $e + p$ collisions (and also to very low $x_B < 10^{-4}$ for higher beam energies). Due to this coverage, measurements of charm hadron production will be able to constrain PDFs at large parton longitudinal momentum fraction, x_p . This will also provide information to elucidate the EMC effect and short range correlations from $e + A$ collisions at large x_p [6]. Compared to inclusive and semi-inclusive light hadron data, heavy flavor production will shed light on the understanding of the gluonic role to the short range correlations in this region.

The broad kinematic coverage will also be interesting as it probes the potential intrinsic charm (IC) contribution to the proton wave function, which was first proposed in the paper by Brodsky *et al.* [7]. Intrinsic heavy quarks, referring to a heavy $Q\bar{Q}$ pair which couples to the valance quarks in the proton forming a five quark Fock state

*mkelsey@wayne.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

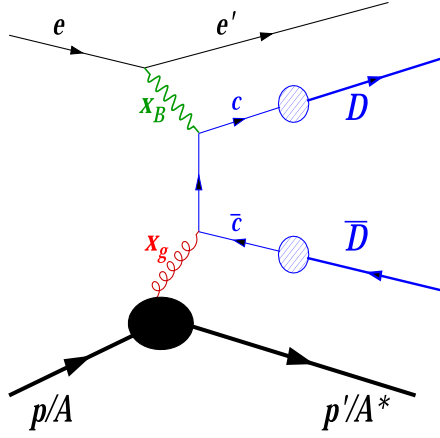


FIG. 1. Leading-order diagram for charm and anticharm pair production in $e + p$ /ion deep inelastic scatterings (DIS).

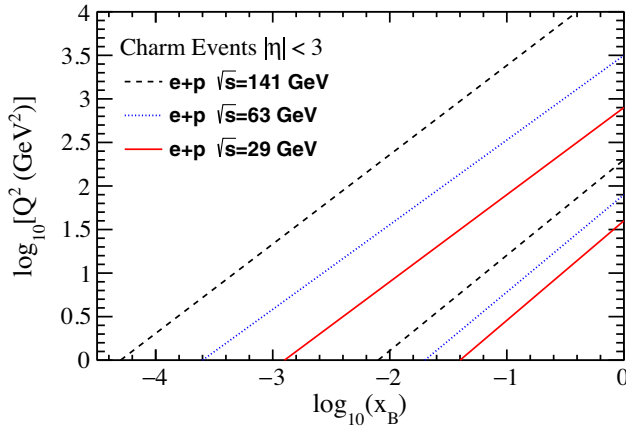


FIG. 2. Q^2 vs x_B coverage of charm events with at least one charm hadron with $|\eta| < 3$ for three beam-energy configurations.

$|uudQ\bar{Q}\rangle$, have been of considerable interest for several decades and are not experimentally confirmed (see e.g., the review [8] and references therein). Separate from the “extrinsic” charm component, arising from perturbative gluon splittings, the IC component is nonperturbative in nature and cannot be calculated with pQCD. Therefore, an understanding of the IC PDF which will be important for precise and accurate (n)PDF determinations from global fits, must come from experimental data. Most analyses find an IC component at large values of x_p stemming from the large charm quark mass [7–12] (also for intrinsic bottom [13]). This limits the constraints from existing DIS data. More recent studies of data from the Large Hadron Collider (LHC) have been able to place limits on IC [14]. There recently have been additional proposed measurements to probe IC in hadron collisions at RHIC and LHC [15], and SeaQuest [16], which will provide valuable insight into IC before the EIC era. However, as we show in this paper, charm hadron production at the EIC will be well positioned

to study IC with high precision over a broader kinematic range with respect to previous DIS experiments.

Charm hadron production measurements at an EIC have been explored in previous studies [17–21]. In Ref. [17] charm events were tagged by secondary charged kaon tracks while Ref. [18] reported charm mesons decays with secondary vertex displacement. In Ref. [21] charm jets in charge current DIS are studied by utilizing a displaced track counting method. Recently there has been a rapid development of detector R&D for the EIC including precision tracking detectors [20,22]. In this paper we report the statistical projections of charm production in $e + p/\text{Au}$ collisions with exclusive hadronic decays utilizing the secondary vertexing capabilities provided from a silicon vertex detector aimed for future EIC experiments. We additionally report the impact on these measurements from intrinsic charm and the constraints these measurements will have on nuclear gluon PDFs.

In this analysis, we have not considered the final state effects on the charm hadron production, e.g., the possible cold nuclear medium impact on the charm quark energy loss and hadronization, which are largely uncertain at the present stage. We envision that future EIC experiments with significant statistics over different collision energies and nuclei species will allow us to conduct extensive measurements on various differential charm hadron production, e.g., anticharm-to-charm ratio, charm hadrochemistry, and charm fragmentation in jets, to offer insights towards the understanding of these final state effects.

This paper is organized as follows: Sec. II is dedicated to the simulation setup used to study charm reconstruction with an EIC detector. In Sec. III we will provide the projections for the charm hadron reduced cross sections and structure functions $F_2^{c\bar{c}}$ with the expected integrated luminosities. In Sec. IV we study the impact of an IC component in the proton on $F_2^{c\bar{c}}$ from more recent global PDF fits. In Sec. V we will show the constraint on the nuclear gluon PDF from the projected $F_2^{c\bar{c}}$. In Sec. VI we will summarize our conclusions. We note here that in Secs. IV and V the PDF sets used were generated from the LHAPDF 6 library [23].

II. SIMULATION SETUP

For the studies reported in this paper we utilize the PYTHIA 6.4 event generator [24] with the settings outlined in [25] to generate $e + p$ collisions. Events are generated with vector-meson diffractive and resolved processes, semihard QCD $2 \rightarrow 2$ scattering, neutral boson scatterings off heavy quarks within the proton, and photon-gluon fusion. The latter two processes are predominantly responsible for heavy flavor quark production.

The kinematic variables used through this paper are defined as follows: In the one-photon-exchange approximation, an incoming electron of four momentum e scatters into a final state e' via the emission of a virtual photon of four momentum $q = e - e'$, which subsequently interacts

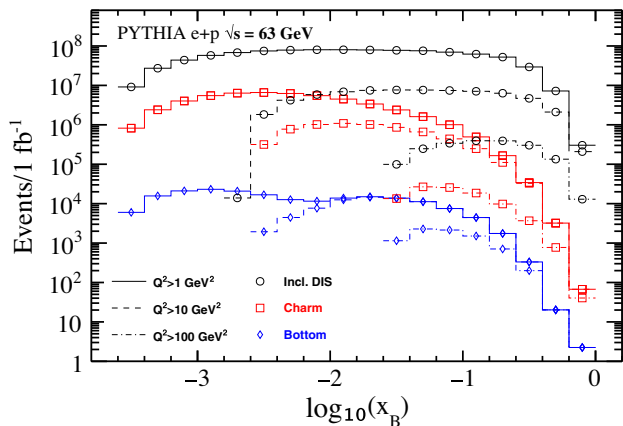


FIG. 3. Event yields per 1 fb^{-1} integrated luminosity in $e + p$ $\sqrt{s} = 63 \text{ GeV}$ PYTHIA collisions for inclusive DIS (open black circles), charm (open red squares), and bottom (open blue diamonds), for three lower bounds of Q^2 as depicted by the various line styles.

with the hadron beam with four momentum p . We follow the convention that the hadron beam momentum is along the positive z direction. The Bjorken scaling variable is defined as $x_B \equiv Q^2/(2p \cdot q)$ and $Q^2 \equiv -q^2$ is minus the square of the four momentum transfer. The inelasticity is defined as $y \equiv p \cdot q/(p \cdot e)$. For the purposes of these studies we do not include any radiative corrections to the incoming/scattered lepton.

The inclusive yields from PYTHIA with a 10 GeV electron beam colliding with a 100 GeV proton beam ($\sqrt{s} = 63$) per 1 fb^{-1} integrated luminosity are shown in Fig. 3 for DIS, charm and bottom production. Integrated above $Q^2 = 1 \text{ GeV}^2$ charm events are about 5.2% of the total DIS cross section, and bottom events are about 0.02%. As previously mentioned, for the expected $e + p$ beam energy configurations charm hadron production will probe a broad kinematic range in x_B and Q^2 as shown in Fig. 2 for 18×275 ($\sqrt{s} = 141 \text{ GeV}$), 10×100 ($\sqrt{s} = 63 \text{ GeV}$), and 5×41 ($\sqrt{s} = 29 \text{ GeV}$) GeV $e + p$ collisions. Here we place a kinematic cut on the pseudorapidity η of the charm hadron to be within ± 3 , which represents the approximate acceptance of an EIC central tracker detector. In Fig. 4, we show the momentum and polar angle distributions for D^0 (top panel) and $D^0 \rightarrow \pi$ (middle panel), where the polar angle is defined with respect to the beam axis, for $\sqrt{s} = 63 \text{ GeV}$ $e + p$ collisions. In the bottom panel we show the decay pion distributions with an event level cut of $x_B > 0.1$. Other charm hadrons have qualitatively similar distributions. With the current binning the bin edges between the first and second polar sectors at 5° and 175° show approximately $\eta = \pm 3$, with positive η defined as the hadron beam direction. One can observe in DIS collisions charm hadrons are produced predominantly within $-3 < \eta < 3$. For events with larger x_B , the decay hadrons are more populated at large η . Therefore, the quality of charm hadron measurements

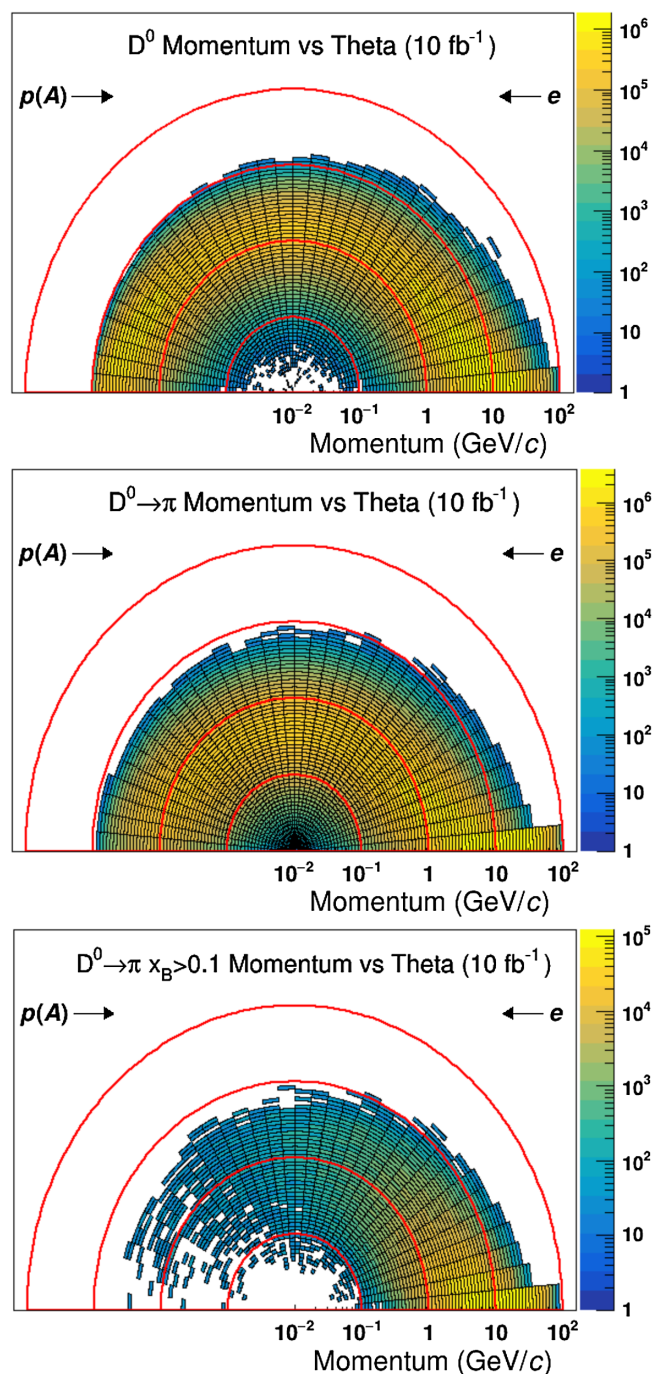


FIG. 4. Kinematic distributions, in polar coordinates, of D^0 mesons (top) and decay pions (middle and bottom) in $\sqrt{s} = 63 \text{ GeV}$ electron-proton collisions generated with PYTHIA 6. Each red semicircle shows the absolute momentum scale at each order of magnitude as indicated by the x -axis intercept. The z axis denotes the yield scaled to 10 fb^{-1} . The bottom panel shows the decay-pion distributions after applying an event-level selection of $x_B > 0.1$.

at large x_B will be highly impacted by the forward tracking system.

Charm hadron reconstruction in the EIC environment is studied using a fast simulation procedure where we smear

TABLE I. Smearing parameters used in fast simulation in different η bins: momentum resolution with two sets of magnetic-field configurations, $\text{DCA}_{r\phi}$ pointing resolution and particle identification momentum upper limits. All p and p_T values are in the unit of GeV/c .

η	σ_p/p —3.0 T (%)	$\sigma(\text{DCA}_{r\phi})$ (μm)	$p_{\text{max}}^{\text{PID}}$ (GeV/c)
(−3.0, −2.5)	$0.1 \cdot p \oplus 2.0$	$60/p_T \oplus 15$	10
(−2.5, −2.0)	$0.02 \cdot p \oplus 1.0$	$60/p_T \oplus 15$	10
(−2.0, −1.0)	$0.02 \cdot p \oplus 1.0$	$40/p_T \oplus 10$	10
(−1.0, 1.0)	$0.02 \cdot p \oplus 0.5$	$30/p_T \oplus 5$	6
(1.0, 2.0)	$0.02 \cdot p \oplus 1.0$	$40/p_T \oplus 10$	50
(2.0, 2.5)	$0.02 \cdot p \oplus 1.0$	$60/p_T \oplus 15$	50
(2.5, 3.0)	$0.1 \cdot p \oplus 2.0$	$60/p_T \oplus 15$	50

the particle level momentum and vertex position using parametrized momentum and pointing resolutions. Additionally, a particle identification (PID) scenario, tracking efficiencies, and primary vertex (PV) resolution are included in our fast simulation. The detector concept that guides the tracking and PID performance used here is explained in more detail in Ref. [22] and references therein. The general design consists of a barrel detector covering $|\eta| < 1$ with six silicon pixel layers, and five tapered silicon pixel planes in each the forward and backward regions covering approximately $1 < |\eta| < 3$. The radial extent of the outer barrel layer and largest forward/backward planes is 43.2 cm. The inner radii of the barrel and forward/backward planes are limited by the beam pipe radius and thickness, which in the nominal interaction region is 3.17 cm and 760 μm , respectively. At large $|z|$ the beam pipe radius fans out to take into account the finite beam crossing angle, and therefore the inner-most barrel layer and disks have radii of 3.30 and 3.20 cm, respectively. We note that we do not explicitly study the effects of a finite beam crossing angle. The inner radii of the detector planes gradually increase at larger $|z|$ to reach a maximum of 5.91 and 4.41 cm for the forward and backward regions, respectively. The total z extent of the tracking system is 242 cm.

The relevant momentum and pointing resolution parametrizations, and PID performance used in the fast simulation are tabulated in Table I for a uniform longitudinal magnetic field strength of 3 T. We do not consider particles that are produced with $|\eta| > 3$. The PV resolution was studied using a GEANT4-based full simulation in the Fun4All framework [26–28]. In this study we embed PYTHIA events with an event level selection of $Q^2 > 1 \text{ GeV}^2$ and with at least one produced D^0 in an all-silicon tracker and fit for the PV position using all reconstructed tracks within the detector acceptance ($|\eta| < 3$). Figure 5 shows the PV resolution in all three dimensions as a function of track multiplicity and has a resolution around 30 μm per dimension at very low multiplicity and gradually reduces to about 10 μm per dimension at a multiplicity greater than 15 tracks. For charm events the average track multiplicity is nine corresponding to an average PV

resolution of around 18 μm in each dimension. A complete study of the tracking efficiency is not yet available, but as a proxy we use the pion pseudotracking efficiency in the same Fun4All simulation using truth level track seeding as shown in Fig. 6 as a function of track transverse momentum (p_T) for three different η regions. At $|\eta| < 1$ we note the low p_T efficiency is worse compared to more forward/backward tracks, and is directly related to the minimum p_T threshold needed for a track to reach the outer-most barrel layer. For $|\eta| > 1$, the efficiency loss at higher p_T is due to edges in acceptance at the barrel-to-plane transition region and also the beam pipe openings in the silicon planes at small radii. Particles with $|\eta| < 3$ in the fast simulation are randomly removed from an event according to these efficiencies. Only tracks which pass are used to determine a total event multiplicity for the event-by-event multiplicity-dependent PV position smearing.

Acquiring a high-purity charm hadron dataset requires the identification of the secondary decay vertex to reduce combinatorial backgrounds. For these studies we consider only D^0 mesons, which have a $c\tau$ of approximately 120 μm [29], with the exclusive hadronic $D^0 \rightarrow K^- \pi^+$ and charge conjugate decays. The pertinent topological variables are the distance between the $K\pi$ and primary vertices

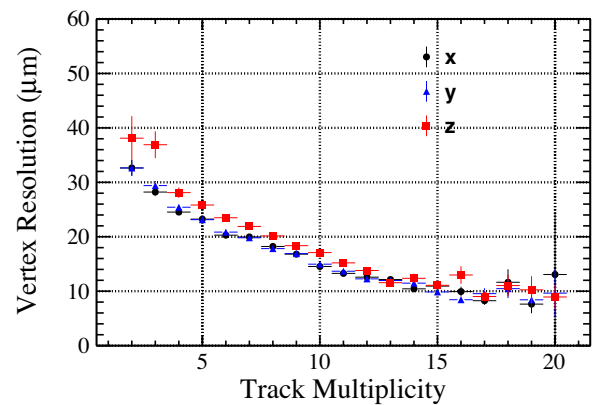


FIG. 5. Primary vertex resolution determined by fitting reconstructed tracks with $|\eta| < 3$ in the full simulation setup with PYTHIA $e + p$ events at $\sqrt{s} = 141 \text{ GeV}$ collisions with a uniform 3.0 T magnetic field and event level selection of $Q^2 > 1 \text{ GeV}^2$.

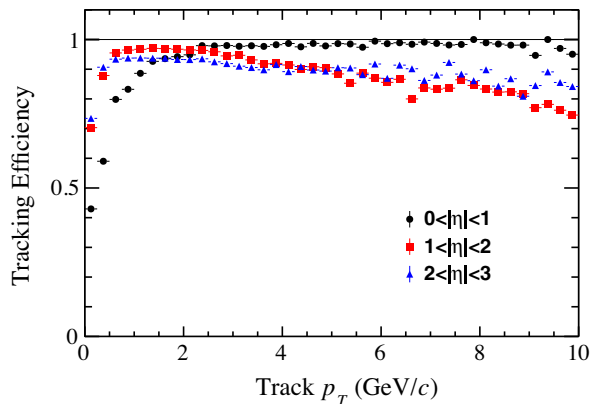


FIG. 6. Tracking efficiency determined in the full simulation in a uniform 3.0 T magnetic field for three different η regions.

(decay length), the distance-of-closest-approach (DCA) between the $K\pi$ pair, and the cosine of the angle, θ , between the D^0 momentum and primary-to-secondary vertex vectors. We note here that the pointing resolution is expected to be significantly better in the $r - \phi$ dimensions than the z dimension in the forward and backward rapidity regions, and therefore we only consider topological variables in the transverse plane. In Fig. 7 we show all three topological distributions for D^0 signal (black points) and background (red histogram) with $p_T < 2$ GeV/c and $|\eta| < 3$, and an invariant $K\pi$ mass within 3σ of the mass

peak. Here both pointing and PV resolutions are folded into the simulation.

As a baseline selection we utilize the following requirements to select D^0 candidates: (i) Radial decay length $> 40 \mu\text{m}$; (ii) pair DCA $< 150 \mu\text{m}$; (iii) $\cos(\theta) > 0.98$. We note that a greater signal significance can be achieved with an optimized or more sophisticated selection (e.g., boosted decision tree) and utilizing the longitudinal dimension; this is left to future studies. The impact of these selections on the signal mass peaks in $\sqrt{s} = 63$ GeV $e + p$ collisions is shown in Fig. 9 for various kinematic regions. In general the signal-to-background (S/B) ratio is improved with respect to no secondary vertex selections. Particularly, this improvement is seen at higher values of p_T . This in effect will provide a data sample with higher signal significance and will reduce systematic uncertainties associated with the signal extraction.

As a systematic check on the fast simulation procedure, we perform the same smearing routine using an alternative detector performance that is directly extracted from the particular GEANT-based simulation described above, and compare the topological variables between the fast and full simulations. These results for an example region of kinematic space are shown in Fig. 8 with the fast simulation shown as the blue histograms and full simulation as the points (similar quantitative comparisons are seen for all kinematic regions). Within the available statistics of the computationally expensive full simulation there is quite

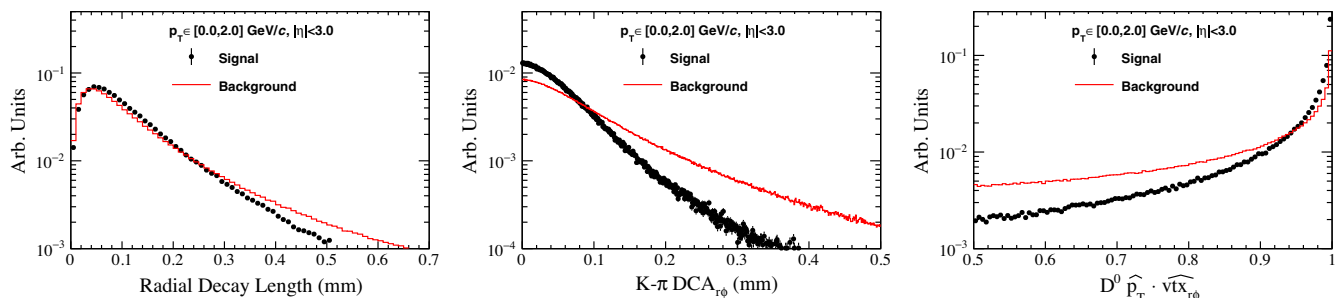


FIG. 7. Example fast simulation distributions for the D^0 radial decay length (left), DCA between kaon and pion tracks (middle), and cosine of the D^0 radial decay vertex pointing angle (right) for signal (black points) and background (red histograms).

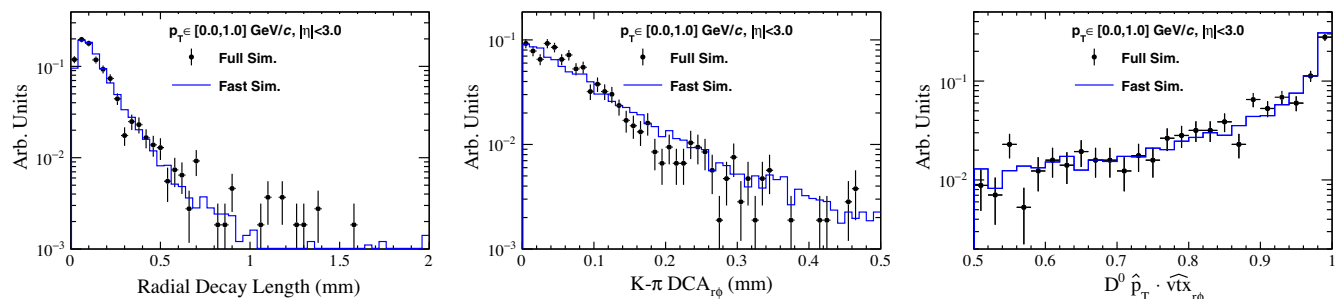


FIG. 8. Distributions for D^0 radial decay length (left), DCA between kaon and pion tracks (middle), and cosine of the D^0 radial decay vertex pointing angle (right) for the full GEANT4-based simulation (black points) and fast simulation (blue histograms).

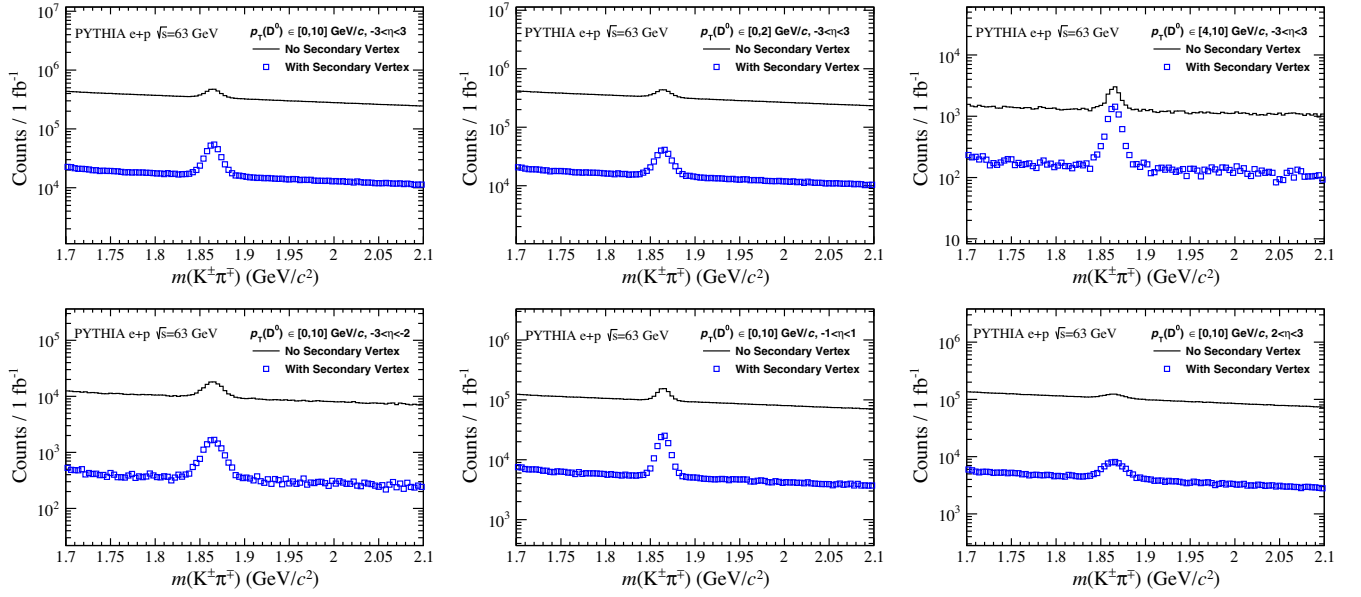


FIG. 9. D^0 invariant mass distributions from PYTHIA events with (blue squares) and without (black histogram) secondary vertex selection. Each panel shows different D^0 kinematic regions as specified in the legends.

good agreement with the fast simulation and we conclude the fast simulation smearing routine is adequate for our projections within the kinematic regions we study.

III. PROJECTION FOR REDUCED CHARM CROSS SECTIONS AND STRUCTURE FUNCTIONS

We first provide the projections for $e + p$ collisions as a baseline, and here we focus our studies on $\sqrt{s} = 63$ and $\sqrt{s} = 29$ GeV energies. The reduced charm cross sections are explicitly calculated as

$$\sigma_r^{c\bar{c}}(x_B, Q^2) = \frac{dN(D^0 + \bar{D}^0)/2}{\mathcal{L} \cdot \epsilon \cdot \mathcal{B}(D^0 \rightarrow K\pi) \cdot f(c \rightarrow D^0) \cdot dx_B dQ^2} \times \frac{x_B Q^4}{2\pi\alpha^2 [1 + (1-y)^2]}, \quad (1)$$

where y is the inelasticity, \mathcal{L} is the integrated luminosity, ϵ is the total efficiency (tracking, PID, reconstruction, and acceptance), $\mathcal{B}(D^0 \rightarrow K\pi)$ is the D^0 branching ratio to $K\pi$, and $f(c \rightarrow D^0)$ is the D^0 fragmentation fraction in PYTHIA (56.6%). As can be observed from the latter quantity, for the purposes of these calculations we scale the measured D^0 yield to get the total charm cross section. We choose a binning in $\log_{10}(Q^2)$ and $\log_{10}(x_B)$ that is five equal bins per decade along each dimension.

The number of $D^0 + \bar{D}^0$ candidates is determined by counting the number of true $D^0 \rightarrow K\pi$ decays (postselection) with invariant mass within $\pm 3\sigma$ of the peak; The background is calculated in the same fashion and is used to define the signal significance in each bin. We scale all statistical uncertainties to a nominal integrated luminosity

of 10 fb^{-1} per center-of-mass energy. The reduced charm cross sections for $\sqrt{s} = 63$ and $\sqrt{s} = 29$ GeV $e + p$ collisions in bins of Q^2 and x_B are shown in Fig. 10 and the uncertainties shown are purely statistical. As can be seen there will be particularly good Q^2 and x_B overlap between the two center-of-mass energies.

To extract the charm structure function $F_2^{c\bar{c}}$ we use the Rosenbluth technique [30] and take the cross sections at $\sqrt{s} = 63$ and $\sqrt{s} = 29$ GeV at fixed x_B and Q^2 and fit the linear form

$$\sigma_r^{c\bar{c}}(x_B, Q^2) = F_2^{c\bar{c}}(x_B, Q^2) - \frac{y^2}{Y^+} F_L^{c\bar{c}}(x_B, Q^2), \quad (2)$$

where $Y^+ \equiv 1 + (1-y)^2$. An example fit for four slices of Q^2 and x_B is shown in Fig. 11. The extracted $F_2^{c\bar{c}}$ central values and uncertainties from the fits are shown in Fig. 12.

Compared to the work in Refs. [17,18], our simulation studies represent a more realistic description of charm-reconstruction capabilities with the EIC detector as we have included PID, momentum and single track pointing resolutions, and primary vertex resolution guided by ongoing detector development/requirements and a full GEANT-based simulation. Furthermore, we have included for the first time the primary vertex resolution in the topological reconstruction of $D^0 \rightarrow K\pi$ decays in an EIC simulation.

In Ref. [17] the longitudinal charm structure functions $F_L^{c\bar{c}}$ are derived from simulation using charm events tagged by the identification of a displaced kaon vertex, and contain background levels that are less than 2%. Comparing the kinematic coverage in Q^2 and x_B of $F_L^{c\bar{c}}$, our derived $F_2^{c\bar{c}}$ has slightly better coverage, particularly in the high- x_B

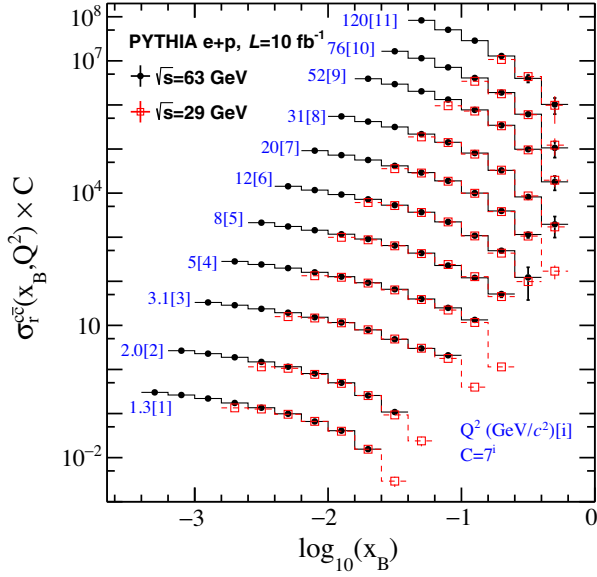


FIG. 10. The reduced charm cross section in bins of $\log_{10}(x_B)$ and $\log_{10}(Q^2)$ for $\sqrt{s} = 63$ GeV (closed black circles) and $\sqrt{s} = 29$ GeV (open red squares) electron + proton collisions in PYTHIA 6 with the charm hadron reconstructed using the all-silicon detector. The vertical values in each Q^2 bin are scaled by the constant terms C defined in the plot for clarity. The $\sqrt{s} = 63$ GeV data is placed at the x axis bin centers while the $\sqrt{s} = 29$ GeV is displaced along the x axis for clarity. The statistical uncertainties are scaled to 10 fb^{-1} .

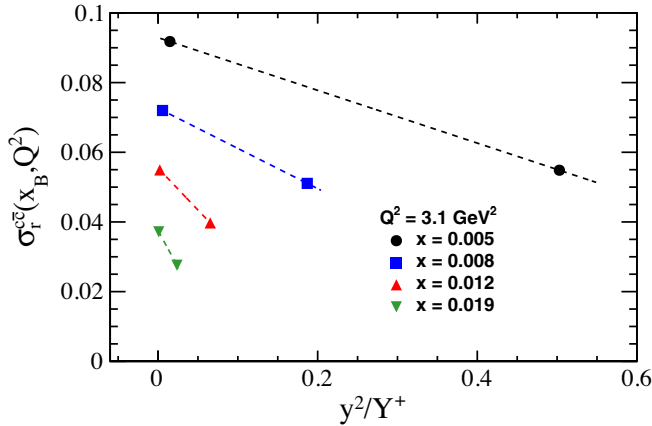


FIG. 11. Example linear fits to the reduced charm cross sections versus y^2/Y^+ in four different slices of Q^2 and x_B .

region (>0.1). This difference is likely driven by the choice of beam energies used in the simulations and also from the need for three energies to reasonably constrain $F_L^{c\bar{c}}$, as opposed to two energies used for $F_2^{c\bar{c}}$. However, we note that in Ref. [17] single track and primary vertex resolutions are not folded into the kaon distributions. Incorporating these resolutions could significantly smear the charm and background kaon vertex distributions, and consequentially reduce the charm event purity and limit the kinematic

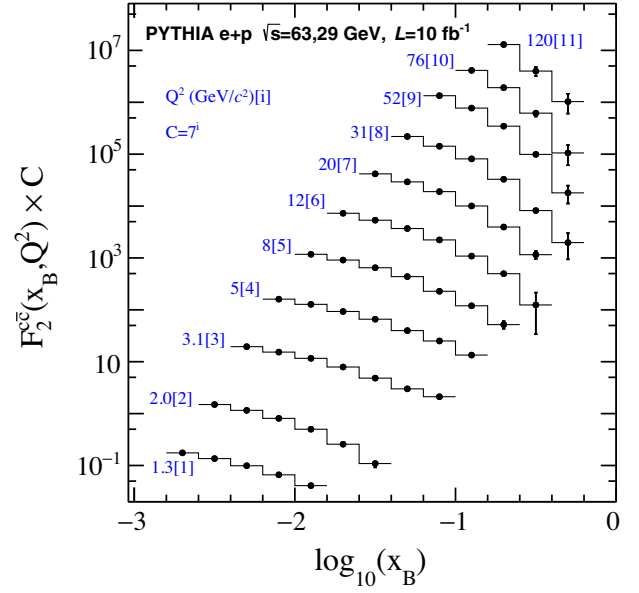


FIG. 12. The projected $F_2^{c\bar{c}}$ in bins of $\log_{10}(x_B)$ and Q^2 . The data points in each Q^2 bin are scaled by a factor of C for clarity. The statistical uncertainties are scaled to 10 fb^{-1} .

coverage. Therefore, in corroboration with Ref. [17], our studies show that with a detector response measurements of charm structure functions will be possible across a broad kinematic range at the EIC. We also note that a precise measurement of $F_L^{c\bar{c}}$ will still be feasible with a full detector response, but as shown in Fig. 2 with the needed three center-of-mass energies the overlap in Q^2 and x_B will be limited.

IV. IMPACT OF INTRINSIC CHARM IN THE PROTON

To estimate the impact on the charm structure functions $F_2^{c\bar{c}}$ with the existence of IC, we replace the default PDFs in PYTHIA 6.4 with those from the recent CT14 analysis [10] with and without IC components. There are several IC models used in the CT14 analysis (and hence different PDF sets), and for our studies we pick two which encapsulate two valiant features of the different models: A scenario in which the IC is “sealike”(Sea1) and another as “valencelike”(BHPS1). Both Sea1 and BHPS1 IC models are described in more detail and compared to data in [10]. The former sealike model has an IC PDF that is spread across all values of x_p , while the latter valencelike produces an IC PDF that is concentrated at values of $x_p > 0.1$. In both models, the average total momentum fraction of the IC is on the order of 1%.

The effect of including IC in the CT14 PDFs on our $F_2^{c\bar{c}}$ projections can be seen in Fig. 13 for two representative bins of Q^2 , and in Fig. 14 across all bins of Q^2 and x_B . Here the black data points correspond to our projections without any IC, and the black dashed line and red solid line show

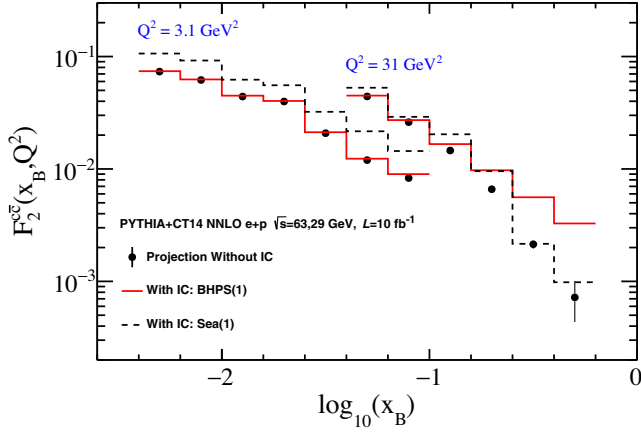


FIG. 13. The projected $F_2^{c\bar{c}}$ with the baseline CT14 PDF (black points), CT14 with sealike intrinsic charm (dashed black lines), and CT14 with BHPS-like intrinsic charm (red lines), for two representative values of Q^2 . The statistical uncertainties are scaled to 10 fb^{-1} .

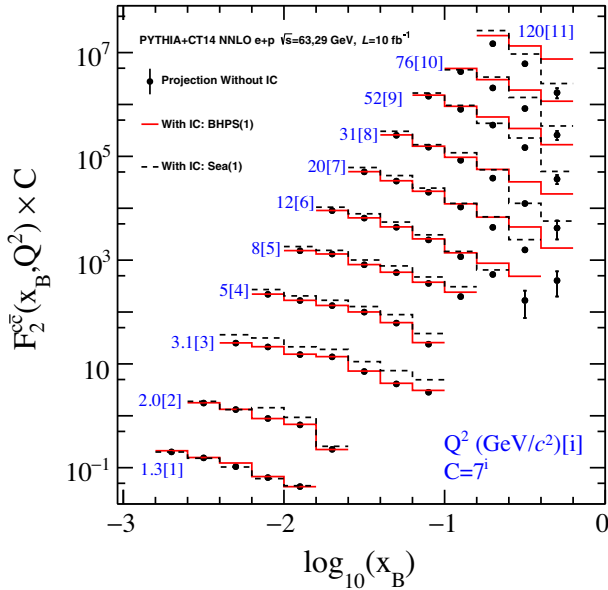


FIG. 14. The projected $F_2^{c\bar{c}}$ with the baseline CT14 PDF (black points), CT14 with sealike intrinsic charm (dashed black lines), and CT14 with BHPS-like intrinsic charm (red lines). The data points in each Q^2 bin are scaled by a factor of C for clarity. The statistical uncertainties are scaled to 10 fb^{-1} .

the enhancement from including a sealike and valancelike IC PDF, respectively. As expected the sealike model produces an enhancement across all values of x_B as the IC PDF is finite across a broad x_p range, and is about a factor of 2 larger with respect to the projection without IC. The valancelike model produces a significant enhancement at $x_B > 0.1$, which is about a factor of 7 at very large x_B . No enhancement is observed at lower values of Q^2 for the valancelike model because in this region the data do not probe values of x_p where the valancelike IC PDF is significant.

We have also repeated this exercise with the NNPDF3 PDF set¹ where the intrinsic charm content of the proton is determined using a “fitted” charm method [31]. The IC PDF from the NNPDF analyses is comparable to the BHPS1 model in the CT14 analysis and peaks at slightly large values of x_p , and has a total momentum fraction also on the order of 1%. We observe that the NNPDF3 IC produces an enhancement in the projected data for $x_B > 0.1$ that is qualitatively similar to the CT14 valancelike model. The absolute enhancement is a factor of 10 for $x_B > 0.1$ (compared to a factor of seven for the CT14 BHPS1 model). We find no enhancement in the projected data with NNPDF3 PDFs below a Q^2 value of 8 GeV^2 where the data probe values of $x_B < 0.1$.

With both IC models the observed enhancement with respect to the no-IC scenario is well above the projected statistical uncertainties of the baseline data. These studies show at the EIC measurements of charm structure functions will be extremely sensitive to the existence of IC and to different IC models.

V. CONSTRAINTS TO THE NUCLEAR GLUON PARTON DISTRIBUTION FUNCTIONS

To project the statistical precision of a charm structure function measurement in $e + \text{Au}$ collisions we utilize the CT14+EPPS16 Au PDFs [32,33] in our PYTHIA simulation and repeat the same procedure outlined in Sec. III. CT14+EPPS16 is used as an example here but we also repeat the exercise for nCTEQ15 [34] and nNNPDF2.0 [35] PDFs for Au nuclei. (The respective proton baseline PDFs used are CT14 [32] and NNPDF3.1 [36] of the same order in the strong coupling constant and with the same value of the strong coupling constant.) The projected ratios of $F_2^{c\bar{c}}$ in $e + \text{Au}$ collisions with respect to $e + p$ are shown in Fig. 15 as the black data points. The data here have been randomly displaced from the central value using the statistical uncertainties in both $e + \text{Au}$ and $e + p$ collisions, and we note the total integrated luminosity per energy and colliding system here is $1 \text{ fb}^{-1}/\text{nucleon}$. The gray hashed curves show the uncertainty on the $F_2^{c\bar{c}}$ ratio coming from the EPPS16 gluon² nPDFs uncertainties only, and are in general much larger than the projected data uncertainties.

To estimate the impact on the nuclear gluon PDF uncertainties we utilize the Bayesian PDF reweighting

¹We explicitly use the NNPDF3_IC_nlo_as_0118_mcpole_1470 and NNPDF3_nIC_nlo_as_0118_mcpole_1470 PDF sets provided by the NNPDF Collaboration.

²Our goal is only to study the impact on the nuclear gluon PDF and we are not performing a complete QCD fit; therefore, we only select charm hadrons in PYTHIA that originate from the photon-gluon fusion process to be used in the Bayesian PDF reweighting procedure. We then weight this subsample of data to conserve the total yield and uncertainty from all processes in each bin of Q^2 and x_B .

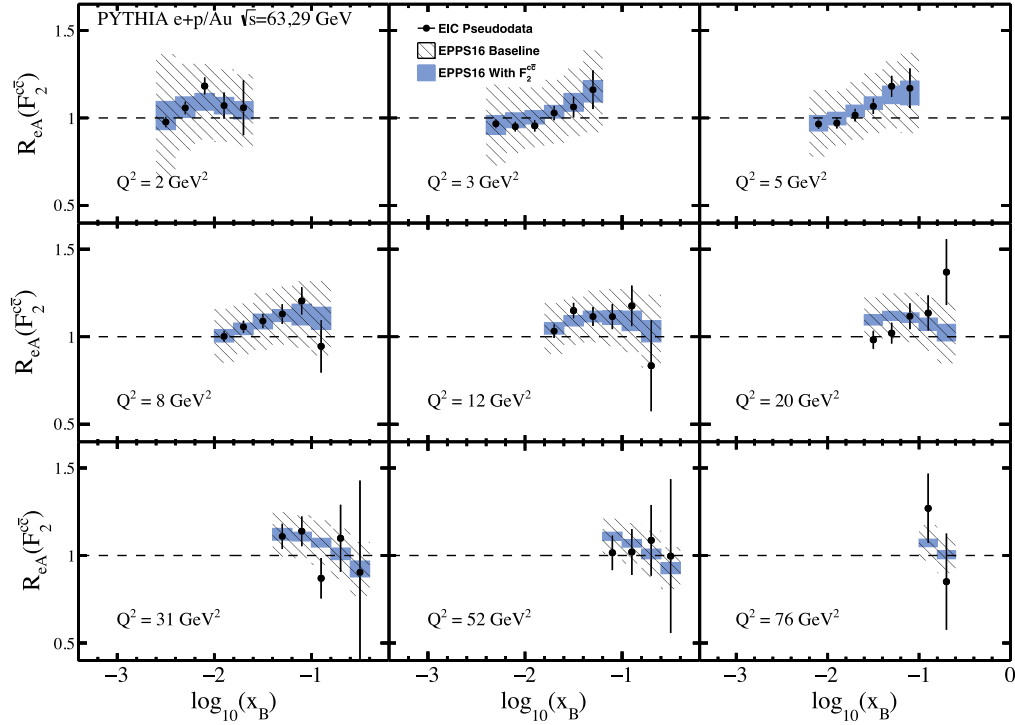


FIG. 15. The ratios of $F_2^{c\bar{c}}$ in $e + \text{Au}$ to $e + p$ collisions in bins of Q^2 and x_B with integrated luminosities of $1 \text{ fb}^{-1}/\text{nucleon}$. The hashed gray curves show the EPPS16 [33] baseline uncertainties while the solid blue curves show the reweighted uncertainties.

procedure [37]. In the case of EPPS16 and nCTEQ15 PDFs, where the PDF errors are Hessian, we generate a large sample of PDF replicas using

$$f_k = f_0 + \sum_i \left(\frac{f_{i,+} + f_{i,-}}{2} \right) r_{k,i}, \quad (3)$$

where f_0 is the PDF central value, the index i runs over all eigen PDF error sets, and $f_{i,+/-}$ are the deviations from the central value of the i th plus/minus PDF error. $r_{k,i}$ is a random Gaussian number centered at zero with unit variance. The reweighted (Hessian) PDFs are then constructed by

$$f_{\text{new}} = f_0 + \sum_i \left(\frac{f_{i,+} + f_{i,-}}{2} \right) \left[\frac{1}{N_{\text{rep}}} \sum_k w_k r_{k,i} \right], \quad (4)$$

with the index k running over all replicas with weights w_k . Here we choose the weights proposed by Giele and Kelle [38]

$$w_k = \frac{\exp[-\chi_k^2/2]}{\frac{1}{N_{\text{rep}}} \sum_k \exp[-\chi_k^2/2]}. \quad (5)$$

In the case of the nNNPDF2.0 PDFs we use the set of already produced Monte Carlo replicas and the weights advocated for by the NNPDF Collaboration [39,40]

$$w_k = \frac{(\chi_k^2)^{\frac{1}{2}(n-1)} \exp[-\chi_k^2/2]}{\frac{1}{N_{\text{rep}}} \sum_k (\chi_k^2)^{\frac{1}{2}(n-1)} \exp[-\chi_k^2/2]}, \quad (6)$$

where n is the number of data points in the fit.

The EPPS16, nCTEQ15, and nNNPDF2.0 nuclear gluon ratios before and after reweighting are shown in Fig. 16 for $Q^2 = 2 \text{ GeV}^2$. Figures 17 and 18 show the same for $Q^2 = 20 \text{ GeV}^2$ and $Q^2 = 120 \text{ GeV}^2$, respectively. The improvement of the gluon nPDF uncertainties, shown in the bottom panels, is estimated to be about a factor of 5 for the EPPS16 and nCTEQ15 nPDFs across all values of gluon momentum fraction. For the nNNPDF2.0 gluon nPDF, the improvement is on average about a factor of 2, 4, and 3 at values of $x_g < 0.01$, $x_g \sim 0.01 - 0.1$, and $x_g > 0.1$, respectively. The reweighted EPPS16 uncertainties for the predicted $F_2^{c\bar{c}}$ ratios are shown in Fig. 15 as the shaded blue bands.

We have also studied the effect of using different weights for the EPPS16 and nCTEQ15 reweighting procedure. We have first performed the analyses with the weights advocated for by the NNPDF Collaboration shown in Eq. (6). In general, we find these weights produce a comparable reduction in the gluon nPDF uncertainties compared to the nominal weights. It has been argued that the Giele-Keller weights should be scaled by taking $\chi_k^2 \rightarrow \chi_k^2 / \Delta\chi^2$ in Eq. (5), which takes into account the tolerance $\Delta\chi^2$ criteria used in the PDF error analyses [37]. Scaling the Giele-Keller

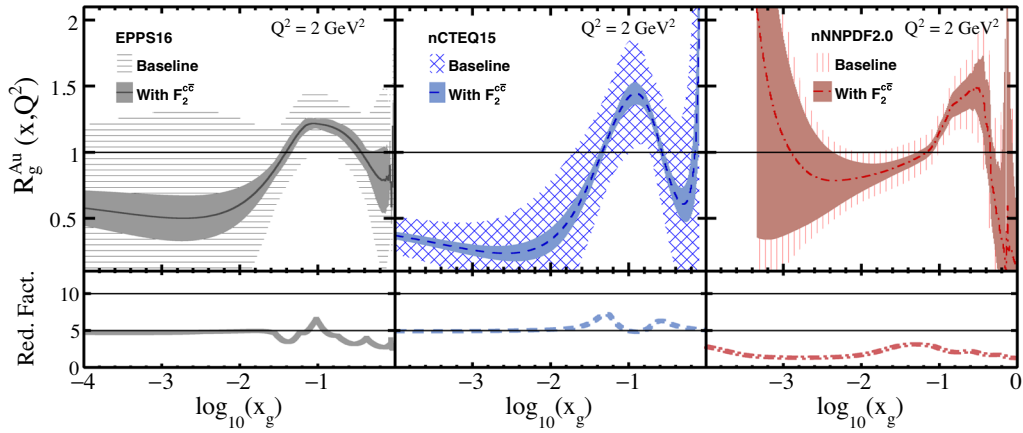


FIG. 16. Top: nuclear gluon ratios as a function of x_g for Au nuclei for the EPPS16 [33] (left), nCTEQ15 [34] (middle), and nNNPDF2.0 [35] (right) PDF sets. The hashed curves show the baseline uncertainties and solid curves show the reweighted uncertainties using the projected charm data with integrated luminosities of 1 fb^{-1} /nucleon. Bottom: reduction factor of the nPDF uncertainties with the inclusion of the EIC charm data.

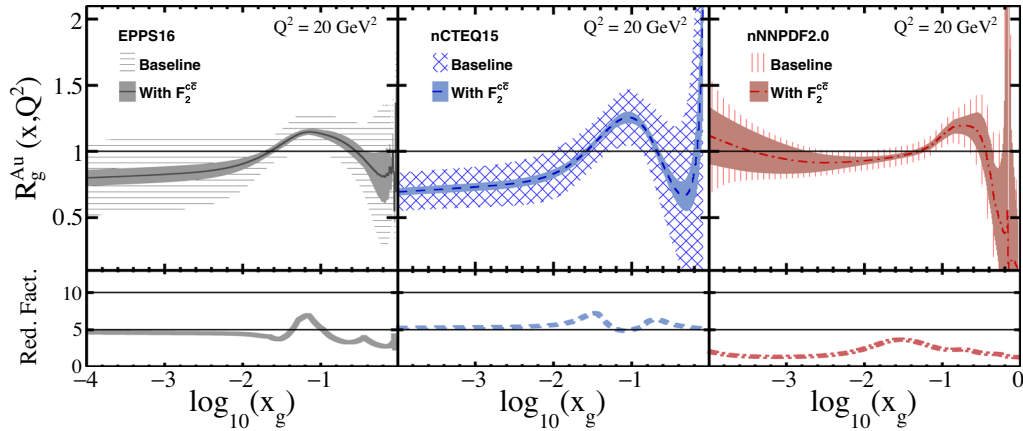


FIG. 17. Same as Fig. 16 but for $Q^2 = 20 \text{ GeV}^2$.

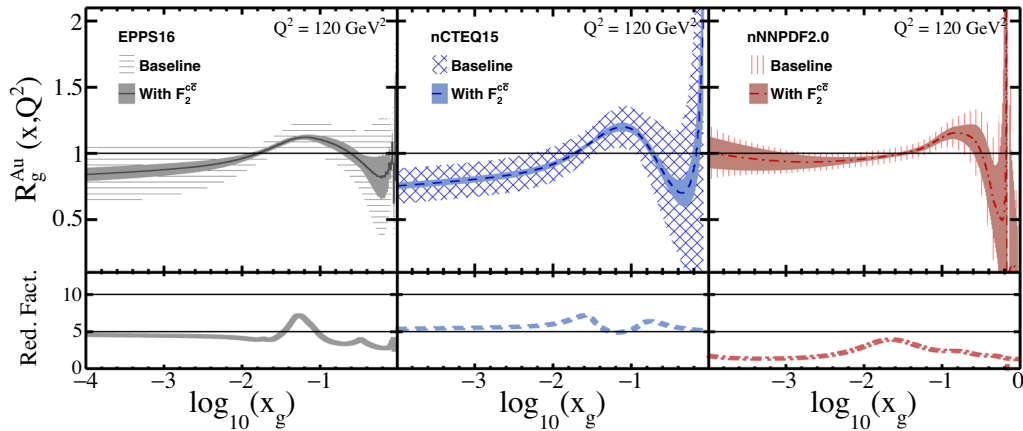


FIG. 18. Same as Fig. 16 but for $Q^2 = 120 \text{ GeV}^2$.

weights in this way broadens the weight distribution versus χ_k^2 . We have also repeated the analysis with the scaled weights and find the reduction in the nPDF uncertainties is

improved rather than deteriorated. We attribute this to the fact that the χ_k^2 distribution normalized by the number of degrees of freedom (*n.d.f.*) has long tails at large values,

and for example EPPS16 has $\langle\chi_k^2\rangle/n.d.f. = 19$. In effect, this scaled weighting procedure includes more information utilizing a significant fraction of replicas that have large deviation with respect to the data, hence producing an improvement in the reweighted nPDF uncertainties. However, this highlights that the new pseudodata have more constraint than existing data used in the global fits potentially indicating a full refit of the data is needed. Therefore, we consider our improved constraints shown in Figs. 16, 17, and 18 as a lower limit.

It has been shown in Ref. [17] that the low- x_g region will already be well constrained from inclusive measurements, with little improvement from charm measurements. Our studies are in corroboration with those in Ref. [17], which use a more flexible EPPS16* nPDF set (particularly at regions of partonic x_p less than the antishadowing peak), and show the high- x_g region will be significantly constrained from open charm measurements. Therefore, confirming that the inclusive and charm measurements will be complementary to constraining the nuclear gluon PDF across all x_g .

VI. CONCLUSIONS

To summarize, we have studied the estimated precision of charm hadron production measurements using a fast simulation that includes detector performance from ongoing detector development. Using the exclusive $D^0 \rightarrow K^- \pi^+$ decay channel we have estimated the expected statistical precision of the reduced charm cross sections in $\sqrt{s} = 63$ and $\sqrt{s} = 29$ GeV $e + p$ collisions, and $F_2^{c\bar{c}}$ in a dataset with an integrated luminosity of 10 fb^{-1} per energy. With the detector resolutions used we show the identification of the charm hadron decay vertex will significantly reduce the backgrounds in a EIC environment. This will improve

the signal significance and reduce systematic uncertainties associated with the signal extraction.

We studied the effect of using proton PDFs that include several IC models on the observed $F_2^{c\bar{c}}$ in PYTHIA and find $F_2^{c\bar{c}}$ is enhanced well above the projected statistical uncertainties and those coming from the baseline PDFs. For the IC PDFs determined in the CT14 BHPS1 model and NNPDF3 fitted charm, this enhancement is found to be nearly an order of magnitude at $x_p > 0.1$. These data will be important for understanding the proton charm PDFs and for an accurate determination of all the proton PDFs.

We have also studied the constraint on the gluon nPDFs coming from the EIC charm measurements using a Bayesian PDF reweighting procedure. We have incorporated several nuclear PDFs in our simulation, and have calculated the ratio of $F_2^{c\bar{c}}$ in $e + \text{Au}$ to $e + p$ with 1 fb^{-1} /nucleon worth of data per energy and colliding system. Incorporating the EIC charm data in a Bayesian PDF reweighting procedure shows that relative gluon nPDF uncertainties can be reduced by a factor of about 3 to 7 depending on the nPDF used and kinematic region. This will be particularly important in the high- x_p region where inclusive measurements will have little constraint, thus highlighting the necessity of these measurements.

ACKNOWLEDGMENTS

We thank Feng Yuan for invaluable discussions throughout the course of these studies. We also thank the RCF facility at Brookhaven National Laboratory and the NERSC center and Lawrence Berkeley National Laboratory for the computational resources used for this work. This work was supported in part by the Office of Nuclear Physics within the U.S. DOE Office of Science and the U.S. National Science Foundation.

-
- [1] 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).
 - [2] E. C. Aschenauer *et al.*, The Electron-Ion Collider: Assessing the energy dependence of key measurements, *Rept. Prog. Phys.* **82**, 024301 (2019).
 - [3] R. A. Khalek *et al.*, Science requirements and detector concepts for the Electron-Ion Collider: EIC yellow report, [arXiv:2103.05419](https://arxiv.org/abs/2103.05419).
 - [4] Electron-Ion Collider eRHIC: Pre-Conceptual Design Report, <https://wiki.bnl.gov/eic/upload/EIC.Design.Study.pdf> (2019).
 - [5] H. Abramowicz *et al.*, Combination and QCD analysis of charm and beauty production cross-section measurements in deep inelastic ep scattering at HERA, *Eur. Phys. J. C* **78**, 473 (2018).
 - [6] S. Malace, D. Gaskell, D. W. Higinbotham, and I. Cloet, The challenge of the EMC effect: Existing data and future directions, *Int. J. Mod. Phys. E* **23**, 1430013 (2014).
 - [7] S. J. Brodsky, P. Hoyer, C. Peterson, and N. Sakai, The intrinsic charm of the proton, *Phys. Lett.* **93B**, 451 (1980).
 - [8] S. Brodsky, V. A. Bednyakov, G. I. Lykasov, J. Smiesko, and S. Tokar, The physics of heavy quark distributions in hadrons: Collider tests, *Prog. Part. Nucl. Phys.* **93**, 108 (2017).
 - [9] H. Abdolmaleki and A. Khorramian, Parton distribution functions and constraints on the intrinsic charm content of the proton using the Brodsky-Hoyer-Peterson-Saka approach, *Phys. Rev. D* **99**, 116019 (2019).

- [10] T.-J. Hou, S. Dulat, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, C. Schmidt, J. Winter, K. Xie, and C.-P. Yuan, Ct14 intrinsic charm parton distribution functions from cteq-tea global analysis, *J. High Energy Phys.* **02** (2018) 59.
- [11] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Parton distributions for the LHC, *Eur. Phys. J. C* **63**, 189 (2009).
- [12] P. Jimenez-Delgado, T. J. Hobbs, J. T. Londergan, and W. Melnitchouk, New Limits on Intrinsic Charm in the Nucleon from Global Analysis of Parton Distributions, *Phys. Rev. Lett.* **114**, 082002 (2015).
- [13] F. Lyonnet, A. Kusina, T. Jezo, K. Kovarik, Fred Olness, I. Schienbein, and J.-Y. Yu, On the intrinsic bottom content of the nucleon and its impact on heavy new physics at the LHC, *J. High Energy Phys.* **07** (2015) 141.
- [14] V. A. Bednyakov, S. J. Brodsky, A. V. Lipatov, G. I. Lykasov, M. A. Malyshev, J. Smiesko, and S. Tokar, Constraints on the intrinsic charm content of the proton from recent atlas data, *Eur. Phys. J. C* **79**, 92 (2019).
- [15] V. Goncalves, F. Navarra, and T. Ullrich, Looking for intrinsic charm in the forward region at BNL RHIC and CERN LHC, *Nucl. Phys.* **A842**, 59 (2010).
- [16] R. Vogt, Limits on intrinsic charm production from the seaquest experiment, *Phys. Rev. C* **103**, 035204 (2021).
- [17] E. Aschenauer, S. Fazio, M. Lamont, H. Paukkunen, and P. Zurita, Nuclear structure functions at a future electron-ion collider, *Phys. Rev. D* **96**, 114005 (2017).
- [18] E. Chudakov, D. Higinbotham, C. Hyde, S. Furlotov, Y. Furlotova, D. Nguyen, M. Stratmann, M. Strikman, and C. Weiss, Heavy quark production at an Electron-Ion Collider, *J. Phys. Conf. Ser.* **770**, 012042 (2016).
- [19] C.-P. Wong, X. Li, M. Brooks, M. J. Durham, M. X. Liu, A. Morreale, C. da Silva, and W. E. Sondheim, A proposed forward silicon tracker for the future electron-ion collider and associated physics studies, [arXiv:2009.02888](https://arxiv.org/abs/2009.02888).
- [20] C.-P. Wong *et al.*, A proposed forward silicon tracker for the future electron-ion collider and associated physics studies, [arXiv:2009.02888](https://arxiv.org/abs/2009.02888).
- [21] M. Arratia, Y. Furlotova, T. J. Hobbs, F. Olness, and S. J. Sekula, Charm jets as a probe for strangeness at the future electron-ion collider, *Phys. Rev. D* **103**, 074023 (2021).
- [22] J. Arrington *et al.*, Eic physics from an all-silicon tracking detector, [arXiv:2102.08337](https://arxiv.org/abs/2102.08337).
- [23] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, LHAPDF6: Parton density access in the LHC precision era, *Eur. Phys. J. C* **75**, 132 (2015).
- [24] T. Sjstrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, *J. High Energy Phys.* **05** (2006) 026.
- [25] PYTHIA 6.4.28, <https://eic.github.io/software/pythia6.html>.
- [26] Fun4All, https://wiki.bnl.gov/sPHENIX/index.php/EIC_SPHENIX_FUN4ALL.
- [27] C. Pinkenburg (PHENIX Collaboration), Analyzing ever growing datasets in PHENIX, *J. Phys. Conf. Ser.* **331**, 072027 (2011).
- [28] C. Pinkenburg, Online monitoring and online calibration/reconstruction for the PHENIX experiment, in *14th International Conference on Computing in High-Energy and Nuclear Physics* (2005) pp. 127–129, <http://dx.doi.org/10.5170/CERN-2005-002.127>.
- [29] P. Zyla *et al.* (Particle Data Group), Review of particle physics *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).
- [30] M. N. Rosenbluth, High energy elastic scattering of electrons on protons, *Phys. Rev.* **79**, 615 (1950).
- [31] R. D. Ball, V. Bertone, M. Bonvini, S. Carrazza, S. Forte, A. Guffanti, N. P. Hartland, J. Rojo, and L. Rottoli, A determination of the charm content of the proton, *Eur. Phys. J. C* **76**, 647 (2016).
- [32] 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).
- [33] 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).
- [34] K. Kovarik *et al.*, nCTEQ15: Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework, *Phys. Rev. D* **93**, 085037 (2016).
- [35] R. A. Khalek, J. J. Ethier, J. Rojo, and G. van Weelden, NNNPDF2.0: Quark flavor separation in nuclei from LHC data, *J. High Energy Phys.* **09** (2020) 183.
- [36] R. D. Ball *et al.*, Parton distributions from high-precision collider data, *Eur. Phys. J. C* **77**, 663 (2017).
- [37] H. Paukkunen and P. Zurita, Pdf reweighting in the Hessian matrix approach, *J. High Energy Phys.* **12** (2014) 100.
- [38] W. T. Giele and S. Keller, Implications of hadron collider observables on parton distribution function uncertainties, *Phys. Rev. D* **58**, 094023 (1998).
- [39] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali (NNPDF Collaboration), Reweighting NNPDFs: The W lepton asymmetry, *Nucl. Phys.* **B849**, 112 (2011); Erratum, *Nucl. Phys.* **B854**, 926 (2012); Erratum, *Nucl. Phys.* **B855**, 927 (2012).
- [40] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, Reweighting and unweighting of parton distributions and the LHC W lepton asymmetry data, *Nucl. Phys.* **B855**, 608 (2012).