JLAB-XXXXX, SMU-HEP-20-05

Charm jets as a probe for strangeness at the future Electron-Ion Collider

Miguel Arratia,[1,2] Yulia Furletova,[2] T. J. Hobbs,[3,4] Fredrick Olness,[3] and Stephen J. Sekula[3,*]

[1] *Department of Physics and Astronomy, University of California, Riverside, CA 92521, USA*
[2] *Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA*
[3] *Department of Physics, Southern Methodist University, Dallas, TX 75275, USA*
[4] *Jefferson Lab, EIC Center, Newport News, VA 23606, USA*

(Dated: June 21, 2020)

We explore the feasibility of the measurement of charm-jet cross sections in charged-current deep-inelastic scattering at the future Electron-Ion Collider. This channel provides clean sensitivity to the strangeness content of the nucleon in the high-x region. We estimate charm-jet tagging performance with parametrized detector simulations. We show the expected sensitivity to various scenarios for strange parton distribution functions. We argue that this measurement will be key to future QCD global analyses, so it should inform EIC detector designs and luminosity requirements.

I. INTRODUCTION

The future Electron-Ion Collider (EIC) [1] will herald a new era for the study of nucleon structure by producing a unique data sample of deep-inelastic scattering (DIS) measurements off protons, deuterium, and helium with high polarization of both beams. The EIC will undertake a comprehensive mapping of the nucleon's multidimensional *tomography*, with the broad goal of unlocking the proton's partonic substructure. For collinear quantities such as the unpolarized and nucleon-helicity dependent parton distribution functions (PDFs), $f(x,Q)$ and $\Delta f(x,Q)$, this entails heightened precision for unraveling flavor and kinematical (*i.e.*, x,Q) dependence.

The EIC will possess unprecedented capabilities to address these issues, owing mainly to high luminosities (100–1000 times the HERA instantaneous luminosity). Moreover, its coverage will extends to very high x — relative to HERA [2], reaching a factor of 10 higher x for a specified Q^2 — given EIC's lower center-of-mass energy. As such, the EIC will be well-disposed to exploring not only the gluon-dominated region at lower x, but also the high-x frontier.

This access to high x will allow the EIC to resolve long-standing questions regarding the precise balance of quark flavors contributing to the proton's valence-region structure. Disentangling high-x flavor dependence in PDFs poses a challenge due to the rapid decline of even the valence quark distributions beyond $x \gtrsim 0.1$ and the comparatively small normalization of the d-quark and sea PDFs relative to the u-quark density. Extractions of the d-type PDFs are further complicated by nuclear corrections needed for DIS off deuterium. From the perspective of nonperturbative dynamics, the flavor decomposition of the proton (and of other light hadrons) carries signatures of QCD's patterns of symmetry breaking, including dynamical chiral symmetry breaking and the low-energy violation of flavor-SU(3) symmetry. This latter issue has stimulated considerable effort in separating nucleon strangeness, $s,\bar{s}(x,Q)$, from the rest of the nucleon's light-quark sea, $\bar{u},\bar{d}(x,Q)$. The strange PDF is often examined in terms of its fractional size relative to the total SU(2) quark sea,[1] R_s,

$$R_s(x,Q) = \frac{s(x,Q)+\bar{s}(x,Q)}{\bar{u}(x,Q)+\bar{d}(x,Q)} \,. \tag{1}$$

A primary source of information in contemporary determinations of the strange PDF is supplied by fixed-target neutrino DIS experiments involving heavy nuclear targets. The interpretation of data from these experiments is complicated by a subtle interplay of effects arising from nuclear, target-mass, and other power-suppressed corrections, as well as potential contamination from target fragmentation [3, 4]. These effects present a serious challenge to rigorously quantifying the uncertainty of the subsequent PDF extraction.

Sensitivity to the strange PDF can also be gained with measurements of identified hadrons in the SIDIS approach [5]. This method has multiple associated challenges as well, including a strong dependence of the extracted PDF upon the associated fragmentation function or hadronization model. This issue has prompted

* corresponding author: ssekula@smu.edu

[1] For reference, other quantities used may be the fractional size relative to the d-quark sea $r_s(x,Q) = (s+\bar{s})/(2\bar{d})$, or the integrated momentum ratio $\kappa(Q) = [\int x(s+\bar{s})dx]/[\int x(\bar{u}+\bar{d})dx]$.

efforts to perform simultaneous determinations of PDFs and fragmentation functions [6]. A feature of these studies is the observation that the entanglement of the nonperturbative PDF and fragmentation in a single measurement leads to strong correlations between them.

We note that dimuon measurements in neutrino DIS experiments [7–11] typically prefer a low value of R_s in Eq. (1), whereas kaon semi-inclusive DIS (SIDIS) measurements prefer an even lower value [12, 13]. In contrast, recent electroweak boson measurements at the LHC prefer a value of R_s consistent with unity [14–16]. New data are required to understand the apparent tension between these measurements. We discuss the preferences of different contemporary data sets on the unpolarized strangeness in further detail in Sec. IV A.

In the spin-polarized sector, knowledge of the strange helicity distribution — a quantity even less constrained than the unpolarized strange PDF — is crucial to elucidating the origin of the nucleon spin, which remains an unresolved problem [17]. In addition to fundamental knowledge of nucleon structure, the strangeness content could also illuminate the dynamics of core-collapse supernova explosions by constraining neutrino-nucleon elastic cross sections [18]. Studies of strange helicity have relied on kaon measurements in SIDIS measurements from HERMES [12] and COMPASS [13]. Like the corresponding analyses in the unpolarized sector, these measurements were prone to biases and ambiguities from the needed input from fragmentation functions.

As we shall demonstrate in this analysis, the availability of channels in which the internal structure of the free proton is directly probed by electroweak currents at DIS collider kinematics, the EIC has the potential to avoid many of the complications described above. EIC measurements thus represent a unique opportunity to open a new era of sensitivity to intrinsic strangeness in the proton, including the strange helicity. In particular, an alternative way to achieve flavor sensitivity *without* fragmentation functions and nuclear corrections is charged-current (CC) DIS (Fig. 1).

Inclusive charged-current and neutral-current (NC) DIS at HERA have been instrumental in proton structure studies [19]. Single jet production in inclusive CC DIS was measured by the ZEUS collaboration [20, 21]. Gehrman *et al.* described the ZEUS data with N^{3}LO calculations [22]; this work showed that the inclusion of higher-order pQCD corrections stabilized scale variations to the (sub)percent-level. More recently, the ZEUS collaboration published the first measurement of charm-tagged events in CC DIS [23]. While limited in precision, the ZEUS work demonstrated that this channel is a clean way to access the strange PDF. In addition, the charm

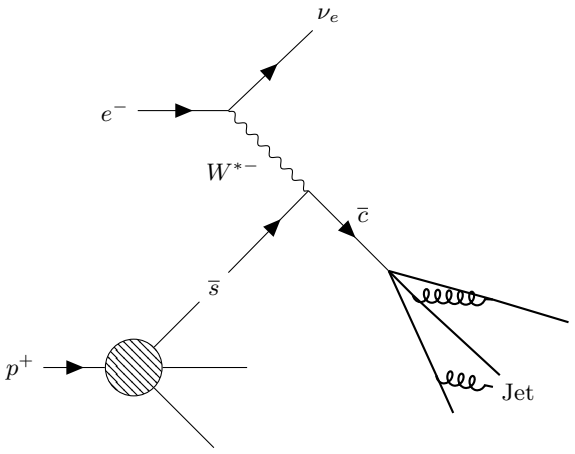

Figure 1. Leading-order diagram for the production of final-state charm in charged-current electron-proton DIS.

cross section in unpolarized and longitudinally polarized DIS [24] is known at NLO accuracy and is a promising way to access the strange helicity while being free unencumbered by many of the issues mentioned above for the neutrino DIS and SIDIS approaches.

Previous feasibility studies of jet measurements at the EIC focused on NC DIS [25–27] and photo-production [28, 29]. The feasibility of inclusive CC DIS measurements at the EIC has been studied by Aschenauer et al. [30]. Here, we specifically focus on charm-jet production in CC DIS and its potential sensitivity to the strange quark sea. This work also differs from recent work by Abdolmaleki et al [31], which emphasized the low-x region that would be provided by the Large Hadron-Electron Collider (LHeC) [32]. In contrast, we focus on the valence region, which remains poorly constrained.

The remainder of this article is as follows: in Sec. II, we describe the details of the Pythia8 and Delphes simulations upon which this analysis is based. Sec. III discusses the specifics of the charm-jet tagging essential to identifying final-state charm in our CC DIS simulations, while Sec. IV describes the strange-sea PDF inputs we employ to test the event-level discriminating power of charm-jet measurements. In Sec. V, we outline specific detector recommendations and requirements to optimize the sensitivity of CC DIS charm-jet production before concluding in Sec. VI.

II. SIMULATION

We use PYTHIA8 [33] to generate 20 million CC DIS events in unpolarized electron-proton collisions with beam energies of 10 GeV and 275 GeV respectively, which is the configuration that maximizes the luminosity in the nominal eRHIC design [34]. We do not include QED radiative corrections, which are relevant in some kinematic regions [30], as these do not drastically affect the focus of our work: the projected precision for strangeness measurements. Moreover, proper treatment of QED radiative corrections requires detailed simulations of detector response that are outside the scope of this work.

In addition, higher-order QCD effects are an important consideration for CC DIS charm production, which receives corrections at NLO from boson-fusion channels unavailable at LO. Still, we expect these higher-order corrections to not dramatically alter the kinematic properties used here for the reconstruction and sensitivity evaluations. The total t-channel CC DIS cross section for $Q^2 > 100\,\mathrm{GeV}^2$ is reported by PYTHIA8 to be 14.8 pb, which is similar to the NLO calculation in Ref. [30]. Comparing the LO vs. NLO calculations for the charged-current charm-production structure functions (*e.g.*, $F_{2c}^{W^-}$) indicates that the NLO corrections are generally relatively mild, especially in the large-x region of relevance to the measurements discussed in this study [35, 36]. This is similar to the situation for Monte Carlo-generated x, Q, η, and p_T distributions for NC DIS charm production, which suggest no substantive differences in the shapes of these distributions between LO and NLO accuracy [37]. Thus, NLO corrections should not significantly impact the present analysis, and we reserve a more detailed examination of the higher-order correction effects to future work.

A. Detector response parametrization

The basic requirements for a collider-detector experiment at the EIC have been established in order to explore the impact of possible detector choices on the realization of physics goals [38]. The baseline EIC detector consists of an inner charged particle tracking system, an electromagnetic calorimeter, a particle-identification (PID) system, and a hadronic calorimeter. The PID system is envisioned to yield at least 3σ separation of $\pi^\pm$, $K^\pm$, and $p^\pm$ for momenta between $1-50$ GeV, depending on η. Electron identification will primarily be achieved using the electromagnetic calorimeter. A dedicated muon system has not been excluded, but is not specified in the baseline.

We use the DELPHES package [39] to obtain a parameterized simulation of detector response. We show in Table I the parameterization of momentum, energy, and impact parameter resolution used as input for DELPHES. All simulated systems provide full azimuthal coverage. The inner tracker is immersed in a 1.5 T solenoidal magnetic field.

The tracking efficiency at the EIC is expected to be close to unity with negligible fake rate, given the low event multiplicity and the proposed use of redundant low-mass silicon pixel detectors [38]. We incorporate a conservative estimate of tracking inefficiency of 1–5% depending on the η region, which is also shown in Table I.

B. Jet kinematics

Jets are reconstructed with the anti-k_T algorithm [40] and $R=1.0$ as implemented in FASTJET [41]. The choice of $R=1.0$ follows the HERA experiments, which showed this minimized hadronization corrections [42]. Jets are defined both "at the generator level" and at the "reconstructed level"; the input for the generator level are final-state particles in PYTHIA8 (excluding neutrinos), whereas the input for the reconstructed level are particle-flow objects from DELPHES. Reconstructed jets are matched to generated jets with an angular distance selection of $\Delta R = \sqrt{(\phi_{jet}^{gen} - \phi_{jet}^{reco})^2 + (\eta_{jet}^{gen} - \eta_{jet}^{reco})^2} < 0.5$ (half the radius parameter). The requirement that an electron-proton collision produce a reconstructed jet within the tracking fiducial region, $|\eta| < 3.0$, is 95% efficient on CC DIS events.

Figure 2 shows the kinematics[2] of charm jets, which lie prominently at low angles to the positive z-axis ($\eta \approx 1.3$) with momenta of $p \approx 15$ GeV. However, a significant fraction of jets are produced at even shallower angles up to $\eta \approx 3$; accounting for the large radius parameter of these jets, this implies that efficient reconstruction and tagging of charm jets will require tracking and calorimeter coverage out to $\eta = 3.5-4.0$, consistent with the baseline EIC detector described above.

The inclusive and charm-jet p_T cross sections are shown in Figure 3. The ratio of charm-to-inclusive cross section is about 3.5% at $p_\mathrm{T}=10$ GeV and it decreases to less than 0.5% at 40 GeV. The jet p_T is correlated with x, so this

[2] We follow the HERA convention to define the coordinate system: the z-direction is defined along the beam axis and the electron beam goes towards negative z. The polar angle θ is defined with respect to the proton direction.

Tracking resolution		
$[-1.0, 1.0]$	$0.5\% \oplus 0.05\% \times p$	
$1.0 < \lvert\eta\rvert < 2.5$	$1.0\% \oplus 0.05\% \times p$	
$2.5 < \lvert\eta\rvert < 3.5$	$2.0\% \oplus 0.01\% \times p$	
Track Impact Parameter Resolution		
Parameter	Resolution [μm]	
d_0	20	
z_0	20	
Charged Particle Tracking Efficiency [%]		
η	$p_{\mathrm{T}} = [0.1, 1.0]$ GeV	$p_{\mathrm{T}} > 1.0$ GeV
$[-3.5, -2.5]$	95	97
$[-2.5, -1.5]$	96	98
$[-1.5, 1.5]$	97	99
$[1.5, 2.5]$	96	98
$[2.5, 3.5]$	95	97
Electromagnetic Calorimeter ($E > 0.2$ GeV)		
η	Resolution [%]	
$[-4.0, -2.0]$	$\sqrt{E} \times (2.0) \oplus E \times (1.0)$	
$[-2.0, -1.0]$	$\sqrt{E} \times (7.0) \oplus E \times (1.0)$	
$[-1.0, 1.0]$	$\sqrt{E} \times (10.0) \oplus E \times (1.0)$	
$[1.0, 4.0]$	$\sqrt{E} \times (12.0) \oplus E \times (2.0)$	
Hadronic Calorimeter ($E > 0.4$ GeV)		
η	Resolution [%]	
$[-4.0, -1.0]$	$\sqrt{E} \times (50.0) \oplus E \times (10.0)$	
$[-1.0, 1.0]$	$\sqrt{E} \times (100.0) \oplus E \times (10.0)$	
$[1.0, 4.0]$	$\sqrt{E} \times (50.0) \oplus E \times (10.0)$	
PID performance		
$K^{\pm}$, $\pi^{\pm}$	$\geq 3\sigma$ separation in the range	
$[-4.0, -1.0]$	up to 10 GeV	
$[-1.0, 1.0]$	up to 6 GeV	
$[1.0, 4.0]$	up to 50 GeV	
$e^{\pm}$, $\pi^{\pm}$	$\geq 2.4\sigma$ separation (rejection factor 100)	
$\mu^{\pm}$, $\pi^{\pm}$	$\geq 2\sigma$ separation	

Table I. Tracking momentum and impact parameter resolution, tracking efficiency, calorimetry resolution, and PID performance that are used as input for DELPHES fast simulations. These parameters are partially based on Ref. [38].

decrease reflects the faster drop of the strange PDF with respect to the valence quark PDFs.

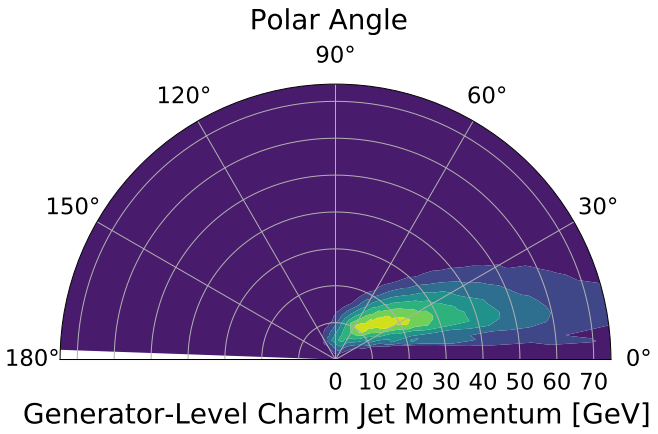

Figure 2. The kinematics (momentum, p, and polar angle, θ, with respect to the direction of the hadron beam) of generated charm jets in CC DIS with $Q^2 > 100$ GeV^2. The jets are clustered with the anti-k_{T} algorithm with $R = 1.0$.

C. Event selection, jet and missing-energy performance

Following HERA measurements [19], the tagging of charged-current DIS events is obtained requiring large missing-transverse energy (E_T^{miss}), which is defined as the magnitude of the vector sum of the transverse momenta of all DELPHES particle-flow objects. It is defined at the generator level in a similar way but using all stable generated particles, or equivalently, by neutrinos.

Figure 4 shows the E_T^{miss} performance obtained with the baseline parameters. The relative E_T^{miss} resolution ranges from 20% (23%) at 10 GeV to 6% (11%) at 40 GeV, defined with a Gaussian fit (standard deviation). The difference between the relative resolutions obtained with a Gaussian fit and the standard deviation reflect the tails of the response, which primarily come from losses due to thresholds in tracking and calorimetry.

The relative jet p_{T} resolution ranges from 18% (21%) at 10 GeV to 7% (12%) at 40 GeV, defined with a Gaussian fit (standard deviation). We have studied what happens to jets and E_T^{miss} in the case that the hadronic calorimeter provides less or no coverage in the barrel region ($|\eta| < 1.0$), such as in the BEAST [43] or JLEIC [44] detector designs. This leads to a significant and asymmetric tail in the jet and E_T^{miss} response. That feature would complicate future unfolding procedures as well as background rejection for photoproduction and NC DIS. Our results agree with Page et al. [26], who reported that

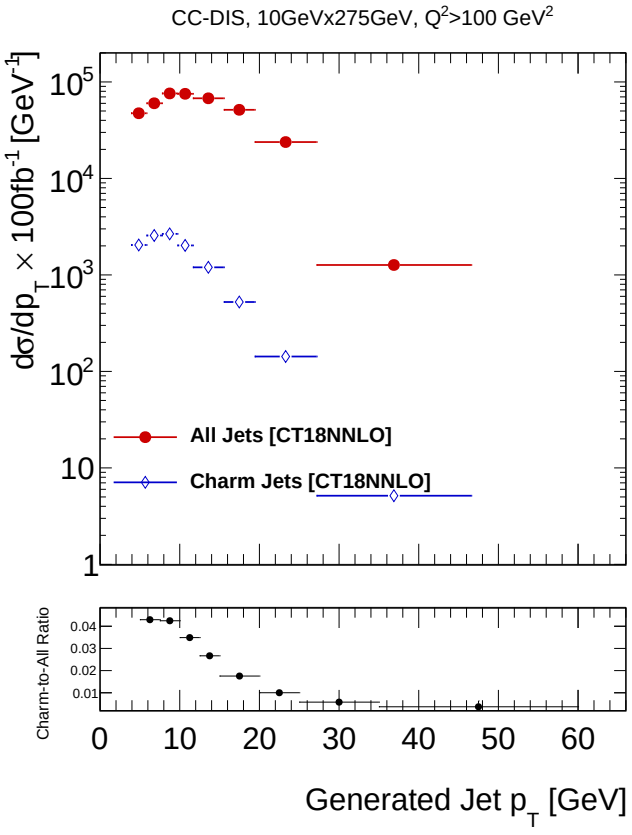

Figure 3. Inclusive and charm-jet production in charged-current DIS at generator level. The jets are reconstructed with the anti-k_T algorithm with $R = 1.0$.

the lack of barrel hadronic calorimeter leads to a severe degradation of jet performance.

We use the Jacquet-Blondel method [45] to reconstruct the event kinematics: the event inelasticity is reconstructed as: $y_{JB} = \sum_i (E_i - p_{z,i})/E_e$, where the sum runs over all particles in the event (particle-flow objects) and E_e is the electron beam energy; the transfer-momentum squared is $Q^2_{JB} = (\mathrm{E_T^{miss}})^2/(1 - y_{JB})$ and Bjorken x is $x_{JB} = Q^2_{JB}/s y_{JB}$, where $s = 4E_eE_p$ and E_e (E_p) is the energy of the electron (proton) beam. We compute the "bin-survival probability" defined as $p^i = \left(N^i_{gen} - N^i_{out}\right)/\left(N^i_{gen} - N^i_{out} + N^i_{in}\right)$, where N_{gen} is the number of events generated in a bin i; N_{out} is the number of events generated in bin i but reconstructed in bin $j \neq i$; and N_{in} is the number of events generated in a bin $j \neq i$ but reconstructed in bin i.

Figure 5 shows the two-dimensional bin-survival probability, which are about 70% or better for a large region

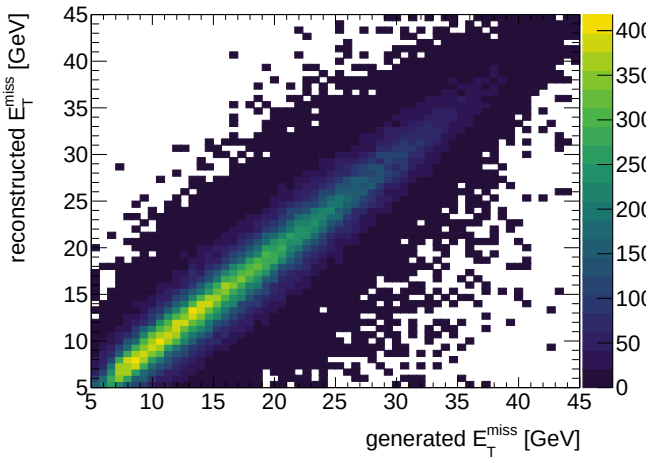

Figure 4. Missing-transverse energy (E_T^{miss}) response matrix for charged-current DIS events. The missing energy is reconstructed using DELPHES particle-flow objects.

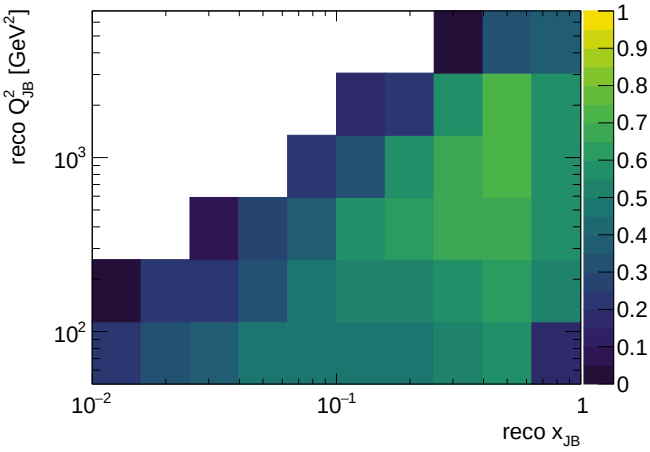

Figure 5. Bin survival probability obtained using the Jacquet-Blondel method in charged-current DIS events.

at high Q^2 and x. Similar results were presented by Aschenauer et al. [30] using the BEAST detector design parameters. This level of bin-survival probability would enable a controlled unfolding procedure in two dimensions (x and Q^2 or x and p_{T}^{jet}). In this work, we focus on one-dimensional distributions (either x, or the p_{T}^{jet} spectrum) and leave detailed unfolding studies for future work.

We select events with $E_T^{\mathrm{miss}} > 10$ GeV. From our simulation of CC DIS events with a reconstructed fiducial jet, this requirement is 87% efficient; we find it to be 75% efficient on events that contain a reconstructed, truth-

matched charm jet. Background from photo-production and NC DIS is suppressed by the $E_T^{\rm miss}$ selection and can be further suppressed by topological cuts, far-forward tagging of electrons, and kinematic constrains, as was done by the HERA experiments [19]. In the following, we assume that these backgrounds are negligible.

III. CHARM JET TAGGING

A. Displaced Track Counting

After jet reconstruction and $E_T^{\rm miss}$ selection, we use a charm-jet tagging algorithm that employs the counting of high-impact-parameter tracks. The $c\tau$ of charm hadrons varies between about 0.2-0.5mm [46], and, for typical charm jets produced at EIC energies, this results in flight lengths of up to a few millimeters from the interaction point. The decay of the charm hadron can result in one or more tracks whose impact parameter is significantly displaced from the interaction point.

We match tracks to a jet and compute the distance of closest approach to the interaction point in the $x-y$ plane (d_0) and along the z-axis (z_0). The 3-D impact parameter significance is then defined as $\mathrm{IP_{3D}} = \sqrt{(d_0/\sigma_{d_0})^2 + (z_0/\sigma_{z_0})^2}$. We assume a resolution of $\sigma_{z_0} = \sigma_{d_0} = 20\mu$m. The signed impact parameter, $\mathrm{sIP_{3D}}$ (Fig. 6), is obtained by multiplying $\mathrm{IP_{3D}}$ with the sign of the product $\vec{p}_j \cdot \vec{r}_{\rm track}$, where $\vec{p}_j$ is the parent jet momentum and $\vec{r}_{\rm track}$ is a vector that points from the interaction point to the point of closest approach on the track.

A basic optimization of the tagger parameters was performed by maximizing the Punzi figure of merit [47] for a target integrated luminosity of 100 fb^{-1} and a goal of separating light from charm jets at 5σ significance. This leads to the requirements of ≥ 2 tracks, each of which satisfies $p_{\rm T}^{track} > 1.0$ GeV; $\mathrm{sIP_{3D}} > 3.75$; and $|d_0| < 3$mm. A jet meeting these criteria is referred to as "tagged." This approach selects both long-lived charm and bottom jets. An example of such a charm-tagged jet is shown in Fig. 7.

The tagging efficiency is defined by identifying all jets matched in the simulation to either a bottom hadron, charm hadron, or light hadron (in that hierarchy), and then determining the number relative to each population that additionally pass the tagging requirement. The performance is summarized in Fig. 8. For charm jets with $p_{\rm T}^{jet} > 10$ GeV, this basic approach leads to charm-jet efficiencies ranging between 10-25% and light-jet efficiencies between $(1-3)\times 10^{-5}$.

As an example of the effect of detector performance on charm-jet tagging efficiency, we degrade the impact parameter resolution along the z axis from 20 μm to 100 μm.

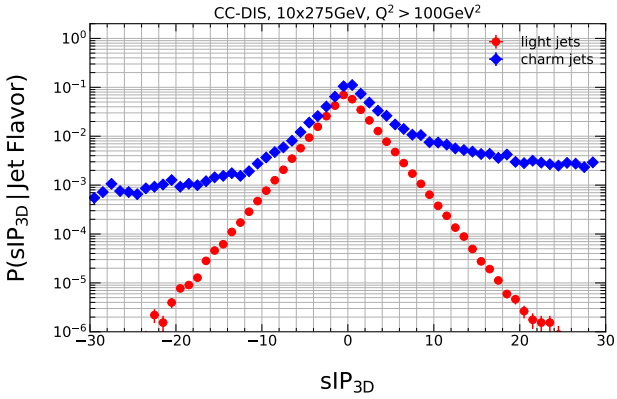

Figure 6. Signed impact parameter significance, $\mathrm{sIP_{3D}}$, probability distribution for light and charm jets.

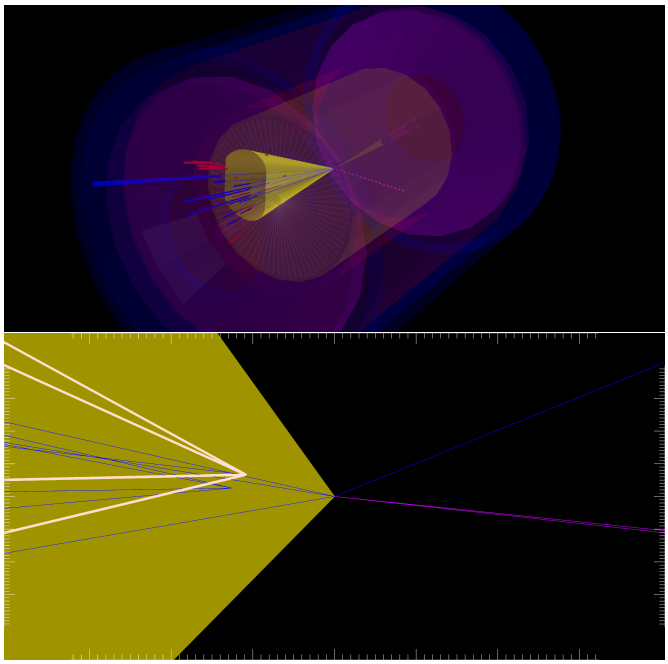

Figure 7. A pair of event displays of a single CC DIS event simulated with PYTHIA8 and reconstructed with DELPHES. A reconstructed jet is represented as a yellow cone; blue bars are hadronic calorimeter energy deposits, and red bars are electromagnetic calorimeter energy deposits. Tracks are indicated by blue lines; the yellow-highlighted tracks originate from a displaced decay vertex. The zoomed-in view shows high-impact-parameter tracks in the jet originating from displaced vertices.

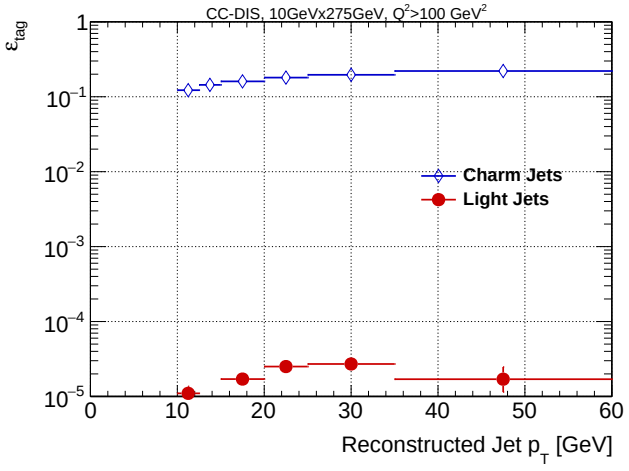

Figure 8. Jet tagging efficiency for light- and charm jets as a function of the jet p_{T}. The high-impact parameter track-counting approach is used to obtain these results.

We then re-optimize the tagging approach, holding the light-jet efficiency constant at $\approx 1\times10^{-5}$ (the nominal optimized value), to see if compensation for the degradation is possible by re-tuning the requirements on the hyperparameters. The average charm tagging efficiency, with re-optimization, degrades from 12% to 9.4% while maintaining the same light-jet efficiency. The re-optimization generally maintains all hyper-parameter requirements except that on the minimum flight significance, which is loosened to maintain background rejection and signal efficiency as the flight significance distribution is diluted by the increased resolution. This change in performance represents a 22% loss in charm-jet tagged yield as a result of this degradation of z_0 resolution. If we additionally degrade the d_0 resolution to 100μm, we observe a further dilution-induced loosening of the flight significance requirement and a corresponding decline in re-optimized tagging efficiency to 8.0% (for the same light-jet efficiency). This would represent an overall loss of 34% of tagged charm jets from the baseline scenario.

We also assessed a more optimistic scenario in which the tracking system permits an improved impact parameter resolution, $\sigma_{d_0} = \sigma_{z_0} = 10\mu$m, over the EIC baseline. A re-optimization under this case results in maintaining the threshold for the minimum number of high-impact tracks as well as the track p_{T} threshold, while tightening the threshold on $\mathrm{sIP_{3D}}$ to 7σ. This yields about the same light-jet efficiency while increasing charm-jet efficiency to 12.6%, a 5% gain.

B. Single-Track PID Jet-Tagging Approaches

Dedicated PID approaches are anticipated as part of the baseline EIC detector. For example, calorimeter-only methodologies (ECAL/HCAL) can be used to separate electrons from pions or other hadrons, as well as using responses from other sub-components like Cherenkov, preshower, or transition-radiation detectors. An additional dedicated muon system could be employed to separate muons especially from pions; a Cherenkov radiation detector could be used to separate kaons from pions. We considered the potential of such systems for charm-jet tagging.

We studied the impact on charm jet tagging if we employed searches in jets for single, high-impact parameter, well-identified kaons, muons, or electrons. Since about 88% of actual charm jets reconstructed in the detector simulation are untagged by the approach in Section III A, we explored the additional tagging efficiency that might be recovered.

We consider only tracks with $p_{\mathrm{T}} > 1$ GeV and with $\mathrm{sIP_{3D}} \geq 3$. We emulate a future PID approach or system by selecting true charged kaons, electrons, pions, or muons contained in a reconstructed jet and applying the following conservative efficiencies/mis-identification rates:

- 90% kaon identification efficiency and a 0.44% pion mis-identification rate (3σ kaon-pion separation). This is consistent with the EIC detector baseline.
- 90% electron identification efficiency and a 2% pion mis-identification rate; this corresponds to a rejection factor of 50 for pions, or a 2.4σ electron-pion separation.
- 95% muon identification efficiency and a 5.4% pion mis-identification rate (2σ muon-pion separation).

Using these approaches, and using the charm jets untagged by the $\mathrm{sIP_{3D}}$ approach, we found an additional 2% charm-jet tagging efficiency gain using just electrons; an additional 3% efficiency gain using just muons; and an additional 9% efficiency gain using just kaons. Combining all three in a logical "OR" resulted in tagging an additional 15% of charm jets previously un-tagged by the track-counting approach. Using all methods together brought the total charm-jet tagging efficiency to 26%. We note that we did not optimize the hyperparameters (track momentum, $\mathrm{sIP_{3D}}$) for this PID-based study, but picked reasonable values given the kinematics involved in these events.

The caveat is that these single-track PID approaches brought a significant increase in background from light-jet contamination. The mis-tagging rate of light jets using the described methods to select single electrons, muons, or kaons was 4×10^{-4}, 5×10^{-4}, and 2×10^{-3}, respectively. These are 1-2 orders of magnitude worse than the light-jet contamination levels from the track-counting approach. This indicates the need to improve and optimize approaches relying on single-track PID (*e.g.*, combining single-track information together into a multivariate discriminant). It also suggests that significantly lower PID performance than employed here would lead to even more challenging background levels from such approaches.

IV. SENSITIVITY TO STRANGENESS

At leading order in α_s, final-state charm production in CC DIS is driven by the flavor excitation process shown in Fig. 1, in which an initial-state (anti)strange quark absorbs a $W^{\mp}$ boson to excite (anti)charm [35]. For this reason, leading-order charm jet production has direct sensitivity to the proton's strange-quark content. In the present analysis, we examine the event-level impact of varying the input strangeness within a set of extremal bounds determined within the CT18 global fit [48].

A. Theory inputs: extreme R_s scenarios

As a proxy for different behaviors of the light-quark sea, we extremize inputs for the high-x behavior of the strange suppression factor, defined in Eq. (1). This quantity has in general received significant attention, as it quantifies the extent to which hadronic-scale QCD interactions lead to the violation of the flavor symmetry commonly assumed in the earliest PDF analyses, *i.e.*, $s = \bar{s} = \bar{u} = \bar{d}$. Historically, PDF fits assumed a suppressed intrinsic strangeness by fixing $R_s = 0.5$, such that the x dependence of the proton's s-PDF was entirely determined by that of the $\bar{u}$, $\bar{d}$ anti-quark PDFs. As noted in Sec. I, such choices have primarily been made given the sparsity of data with direct sensitivity to nucleon strangeness, including semi-inclusive kaon production and dimuon production in neutrino-nucleus DIS. In fact, even modern PDF fits that exclude this latter data, including CJ15, do not actively fit nucleon strangeness, instead taking $R_s = \kappa = 0.4$ [50].

At the same time, an independent strange component of the nucleon wave function has long been the subject of modeling efforts in nonperturbative QCD [18, 51–53],

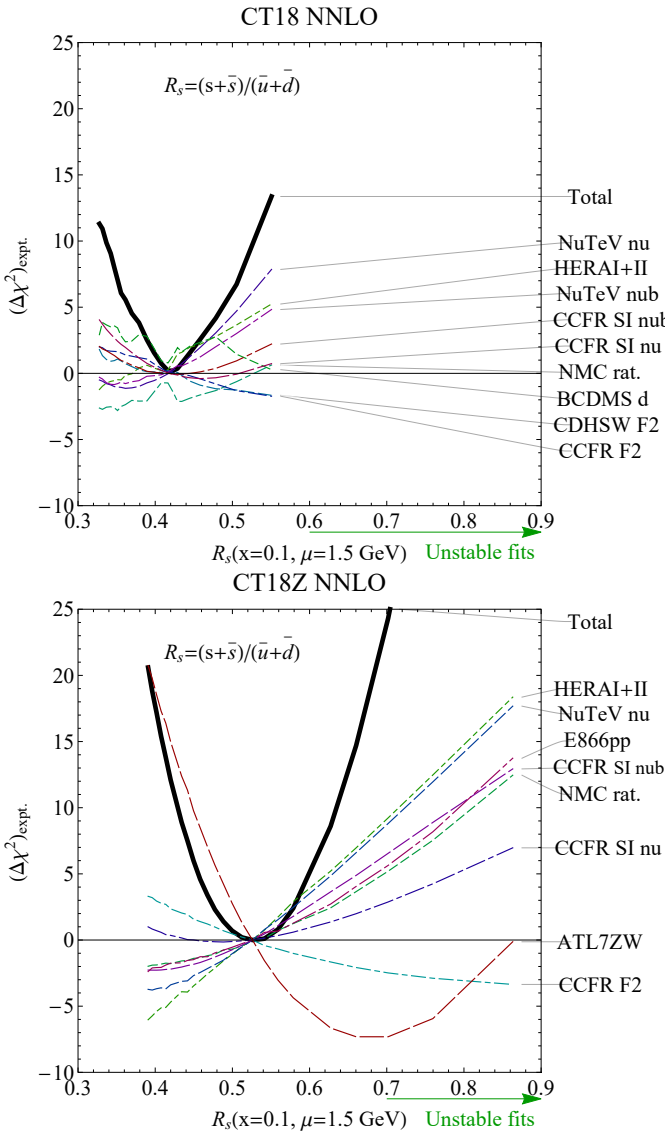

Figure 9. To gauge the event-level sensitivity of charm-jet production, we perform simulations with two extreme inputs for the behavior of the strangeness suppression ratio, R_s, as defined in Eq. (1), taken from the recent CT18 NNLO global PDF analysis [48]. The upper panel corresponds to a LM scan over values of R_s at high $x = 0.1$ in the primary CT18 baseline fit, while the lower panel was obtained for the alternative CT18Z fit, which included the ATLAS 7 TeV W/Z data [49], in addition to a number of other modifications. In both panels, the PDF scale is the factorization scale, $Q = \mu = 1.5$ GeV.

lattice studies [54, 55], and dedicated global PDF analyses [6, 56]. For these reasons, as well as the importance of detailed knowledge of the nucleon sea's flavor structure for precision phenomenology at hadron colliders like the LHC, PDF fitting efforts like the CTEQ-TEA (CT) Collaboration have had a sustained interest in the constraints high-energy data place on $s(x,Q)$ and $R_s(x,Q)$. The most recent iteration of PDF fits developed within the CT global analysis framework — CT18 — were recently released in Ref. [48]. This latest fit examined implications of the recent LHC Run-1 data for the s-PDF, which extended the sensitivity of the global data set beyond that driven by the legacy data included in older fits; among these legacy data are the fixed-target neutrino DIS experiments such as CCFR [8] and NuTeV [57], which still provide the dominant PDF pulls in the strange sector.

A particular subtlety explored in Ref. [48] is the theoretical description of the recent ATLAS 7 TeV W/Z production data [49], which generally prefer an enlarged strange PDF and comparatively larger value of R_s. For instance, at more intermediate $x = 0.023$ and $Q^2 = 1.9\,\mathrm{GeV}^2$, the ATLAS collaboration reported a value of $R_s = 1.13 \pm 0.05(\mathrm{exp}) \pm 0.02(\mathrm{mod})^{+0.01}_{-0.06}(\mathrm{par})$ based upon an internal fit, suggesting unsuppressed strangeness along the lines of the earlier 2012 ATLAS result [14]. The preference of the ATLAS W/Z data for enlarged strangeness was confirmed by the CT18 global analysis as well as other recent studies [58–60]. A detailed discussion of these data and challenges associated with their theoretical description is presented in App. A of Ref. [48]. Ultimately, the ATLAS 7 TeV data were not treated in the CT18 main fit, but rather in an alternative fit, CT18Z, which included these data in addition to several other alternate choices for theory settings and data set selections.

This brings us to the central question of this analysis: *can high-precision charm-tagged data obtained from CC DIS jet production at the EIC help resolve the still ambiguous size and x dependence of the strange-quark sea*? If possible, the EIC would then play a pivotal role via the charm-jet CC DIS channel in navigating the apparent tensions between the pulls of the νA DIS and W/Z hadroproduction data on R_s and nucleon strangeness. For the present feasibility study, we explore this question by examining the event-level discriminating power of CC DIS jet simulations upon widely-separated theory inputs for R_s. Given the especially strong resolving power of CC DIS jet measurements at high x, we therefore examine whether such hypothetical data might be sensitive to two extreme sets for different behavior of R_s at $x = 0.1$ near the nonperturbative starting scale. We note that the current study concentrates on total strangeness; in CT18, which took $s = \bar{s}$, this is therefore $2s = s + \bar{s}$. In principle,

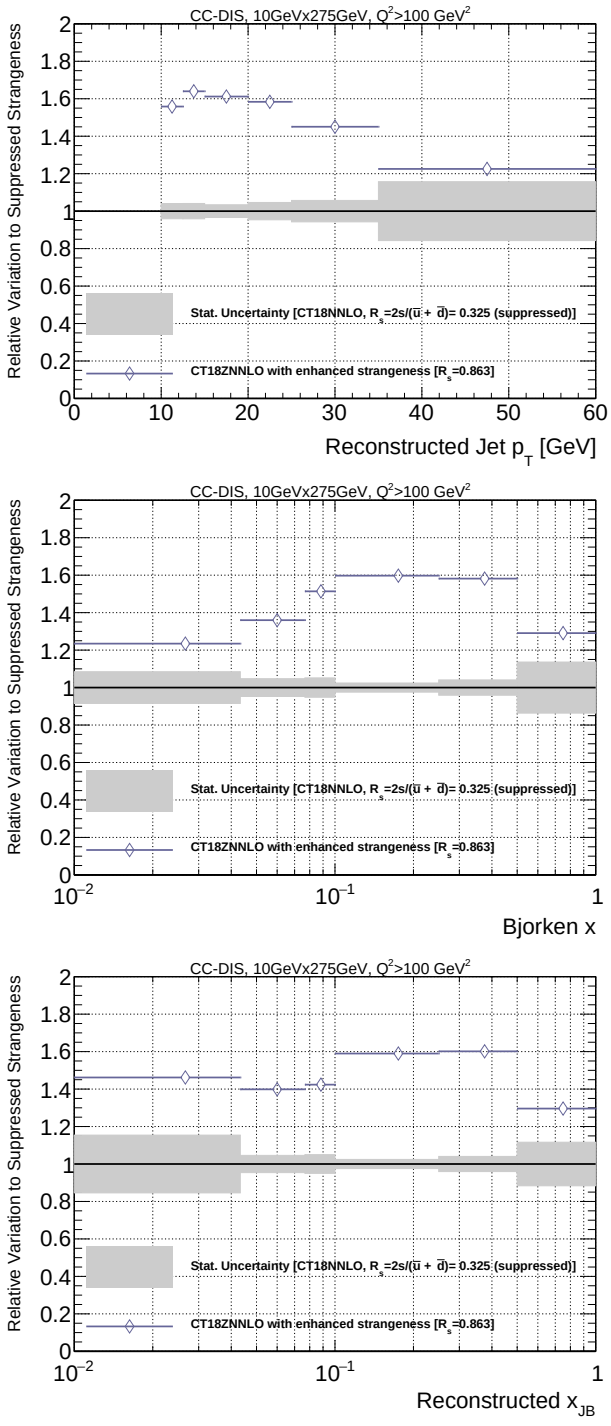

Figure 10. A comparison of two cases: $R_s = 2s/(\overline{u}+\overline{d}) = 0.325$ (CT18 NNLO with suppressed strangeness) and $R_s = 0.863$ (CT18Z NNLO with enhanced strangeness). The gray band indicates the expected statistical error on the reconstructed and tagged charm jet p_T (top), Bjorken x (middle), and x_{JB} (bottom) spectrum in $100\,\mathrm{fb}^{-1}$ of data. The blue points indicate the relative difference in expected yields between the enhanced and suppressed strangeness cases, $1 + (N_{0.863} - N_{0.325})/N_{0.325}$.

an EIC with positron beams could provide an advantageous setting to test $s \neq \bar{s}$. We leave this question to future work.

Among the most robust techniques for exploring the constraints and allowed ranged for specific PDFs in a global analysis is the Lagrange Multiplier (LM) technique [61]. This method proceeds by constraining a PDF to maintain a given numerical value while refitting the other tunable parameters within a global fit. By continuously varying the chosen PDF away from its fitted value, it is then possible to quantify the corresponding variation in χ^2 for the full fit as well as individual data sets. In Fig. 9, we plot the result of this procedure as explored in Ref. [48] for R_s at $x = 0.1$ and $Q = 1.5\,\text{GeV}$. Based upon the LM scans over R_s shown in this figure, we select two extreme PDF sets: one associated with strongly suppressed nucleon strangeness (*i.e.*, small $R_s < 0.5$), identified with the extreme leftmost boundary of the upper LM scan obtained under CT18 in Fig. 9; and one associated with relatively unsuppressed strange ($R_s \sim 1$), corresponding to the rightmost boundary of the lower LM scan based on CT18Z. We therefore examine at what level CC DIS charm jets can discriminate two theory/PDF inputs; namely, PDFs corresponding to

$$\begin{aligned} R_s &= 0.325 \quad (\text{CT18}, \quad \text{suppressed strange}) \\ R_s &= 0.863 \quad (\text{CT18Z}, \quad \text{enhanced strange}) \end{aligned} \tag{2}$$

B. Event-level sensitivity of charm-jet production

For the rest of this paper, we employ only the high-impact parameter track-counting jet-tagging approach (Sec. III A), to maintain charm-jet purity while sacrificing overall statistical precision. We believe this offers a reasonable, if still conservative, baseline for estimating sensitivity to the physics of interest: intrinsic strangeness in the proton.

The EIC beam configuration studied here is expected to result in $\mathcal{O}(1000)$ events in $100\,\text{fb}^{-1}$ of integrated luminosity after charm tagging. We show in Fig. 10 the expected precision of the tagged charm-jet spectrum. Across much of the jet p_T region, or indeed as a function of x_{JB} or true Bjorken x, such uncertainties would be at the level of 10%. This is likely conservatively overestimated, since we have demonstrated that it's possible to enhance charm tagging efficiency with modest additional effort (*e.g.*, single-track PID), even while we have neglected other experimental effects (triggers, knowledge of the jet energy scale, *etc.*) in this study. This statistical uncertainty is to be compared to the range of variation in knowledge of the strangeness PDF from the two scenarios outlined earlier: suppressed and enhanced strangeness contributions (the statistical uncertainties are derived from the suppressed scenario, making them additionally conservative). There is strong evidence that the use of charm-tagged jets at EIC will provide new constraints on the strangeness PDF and should be part of a global analysis of strangeness within the EIC program.

The charm-jet yields described above do not include contributions from gluon-initiated processes, which is not simulated in PYTHIA8. Gluon-initiated background was also estimated to be small in the recent ZEUS measurement [23] and is expected to be even smaller at the EIC (with its larger x reach); so, we neglect these backgrounds in the present study. In the electron-proton channel, the background from sea-d quarks can be safely neglected. Although we neglect these backgrounds for this study, we recognize that at a future EIC detector experiment these will have to be quantified, characterized, and subtracted in order to interpret the data.

In addition to the beam energies discussed here, we have also explored lower-energy configurations (*e.g.*, 10×100 GeV electron-on-proton). As expected, the charm jet yields decline strongly due to decreasing production cross section combined with lower jet p_T. At such energies, the expected charm-jet yield is $\mathcal{O}(10 - 100)$ events.

We have also explored kinematics that will be available at the EIC for nuclear beams. The per-nucleon energy of the nuclear beams is reduced by a factor of Z/A, which is about 0.4-0.5 for most nuclei considered at EIC. As discussed above, lower energies leads to a rapid decline in the expected statistics, so we consider the highest center-of-mass energy can be reached for nuclear beams, which is 110 GeV per-nucleon and a 18 GeV beam electron (center-of-mass energy of about 90 GeV). While the cross sections for hard processes in electron-nucleus collisions gets enhanced by a factor of A, the expected luminosity for nuclear beams is approximately a factor A smaller, which leads to similar expected rates. The higher electron energy (18 GeV instead of 10 GeV) comes at the cost of reduced luminosity due to power limitations, which is roughly a factor of 5 [34]. Thus, the expected statistical uncertainty for nuclear beams is expected to be roughly a factor of $\sqrt{5}$ larger than our nominal studies [3], for equal running time. We thus conclude that the prospects for CC DIS charm-jet studies with nuclear beams are promising.

[3] Note that the lower hadron energy leads to a less boosted kinematics with respect to what we show in Fig. 2. Given that the hadronic final state at mid-rapidity increases [25], the role of the barrel hadronic calorimeter in the E_T^{miss} measurements is enhanced with respect to the higher hadron beam configuration.

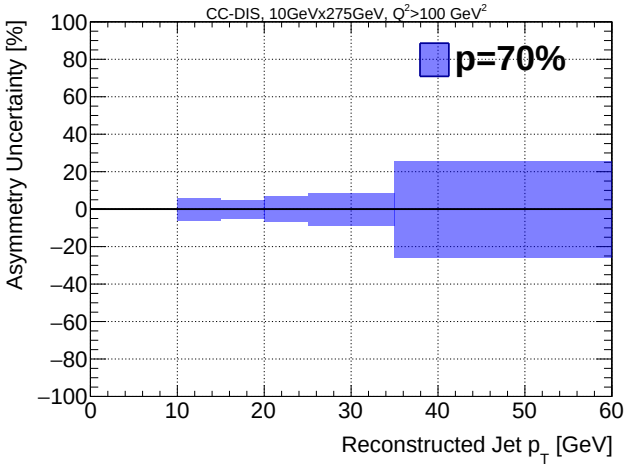

Figure 11. Projected statistical precision on the spin-asymmetry measurement, assuming the polarization of each beam is 70%.

CC DIS measurements with nuclear beams would yield additional flavour sensitivity, for example by using deuterium or helium-3 beams, or to study nuclear effects with heavier nuclei. Given that preliminary studies of heavy ion LHC W/Z production suggest an enhanced strange for the nuclear as well as the proton PDFs, the EIC's ability to explore a variety of nuclear beams could prove illuminating [60, 62]. This would allow us to test the flavour dependence of anti-shadowing and EMC effects, which remains an open question (see Refs. [3, 63–72]). We reserve these studies for future work.

The measured percent error on the spin asymmetry (A) is estimated [24], as a function of the yield of events (N) and the polarization of each beam (P_e, P_p), to go as

$$\delta A/A \approx \frac{1}{P_p}\left(\frac{1}{2}N(1+P_e)\right)^{-\frac{1}{2}} \quad (3)$$

The projected precision on charm-jet spin asymmetries is shown in Fig. 11. Recent work by Borsa *et al.* [73] showed that jet production in polarized NC DIS, which they calculated to NNLO accuracy, is sensitive to quark helicity. When extended to CC DIS, those results could be compared with the precision we estimate to gauge the sensitivity to strange helicity.

Reaching sensitivity to the nucleon's strange helicity content is likely to require a long-term campaign at the EIC, probably requiring an integrated luminosity on the order of 500 fb^{-1}, for asymmetries at the 1%-level. However, there are several ways in which one could aim at increasing the sensitivity. First, our estimate for charm-jet efficiency is rather conservative, and could be improved to at least the level of modern charm-taggers in collider experiments such as the LHC that routinely yield 20–50% [74]. We have also demonstrated the basic gains and challenges that can be expected from the use of displaced leptons and kaons from the charm-meson decays. A mature multivariate analysis would combine all the information including displaced tracks, PID, leptons, and topological (*e.g.*, secondary vertex) information.

A complementary way to increment the statistical power of this channel is to lower the Q^2 cut. Given that the cross section decreases as $1/Q^4$, a relaxing of the selection of > 100 GeV^2 to > 50 GeV^2 would increment the yield substantially and gain sensitivity at lower-x, which further reduces the light-flavor jet background from valence quarks. The challenges associated with measuring low-Q^2 charged-current DIS are manifold, including rejection to backgrounds from photo-production and misidentified neutral-current DIS, as well as increased background from gluon-initiated processes. While most HERA studies imposed a selection on $Q^2 > 200$ GeV^2 [19], studies by Aschenauer *et al.* [30] showed that a lower limit of $Q^2 > 100$ GeV^2 is feasible at the EIC. Future dedicated studies should explore the limit on low Q^2, which most likely will demand highly hermetic detector systems with low thresholds.

V. DETECTOR REQUIREMENTS

In this section, we summarize the main detector requirements to measure charm-jets in charged-current DIS:

- The reconstruction of charm jets with large radius parameter ($R = 1.0$) requires tracking and calorimeter coverage extending in the positive-z direction out to at least $\eta = 3.5 - 4.0$ (Fig. 2). A high tracking efficiency will be essential to reconstruct and tag these jets.

- Given the jet kinematics (Fig. 2) are centered around the barrel-endcap transition region of a typical collider detector, the dead-areas, material budget, and geometry have to be optimized to avoid drastically degrading the detector performance for these jets. An example of a design that achieved this is given by the ZEUS calorimeter [75].

- As jet production is typically in the forward direction and at lower angles to the hadron beam direction, vertex or impact parameter resolution in both the $x - y$ plane and along the z direction will be essential to flavor-tagging approaches such

as those described here. Significant degradation of resolution beyond the baseline of 20 μm is observed to cause significant loss of charm-tagged jet yields for a fixed light-jet efficiency. We also note that a simple optimization of tagging hyper-parameters tended to prefer track momentum thresholds down to 0.75 GeV, illustrating the need to have high efficiency for low-momentum tracks even for the purpose of selecting signal jets.

- We explored the use of single-track PID to enhance charm-jet tagging performance. While we saw significant gains in charm-jet efficiency using baseline EIC detector PID guidance, or reasonable assumptions where such guidance was not present (*e.g.*, electrons and muons), we also saw significant increases in light-jet mis-tagging rates. This suggests that optimization and multi-variable approaches are necessary to include such single tracks in a larger approach to charm-jet tagging. It also suggests, however, that assuming worse PID efficiency and misidentification scenarios than we employed here will only increase the challenge of using such information in a future EIC detector. Dedicated PID system coverage and detector granularity will need to extend well into the forward region, defined above, given the size of the jets.

- This work implies the need for an hermetic detector, with full calorimetry coverage to reach as low a Q^2 (corresponding to low $E_T^{\rm miss}$) as possible, while ensuring background suppression to photo-production and NC DIS. This also demands low thresholds for both tracking and calorimetry, as well as calorimetry resolution (the tracking resolution is subdominant). One example of the importance of hermeticity and its impact on jet resolution will be the trigger; while we did not explore trigger algorithms in this work, an $E_T^{\rm miss}$-based CC DIS trigger algorithm will benefit from strong coverage and finer granularity, as any trigger decisions will necessarily use lower resolution than is available in a fully-calibrated offline environment. The efficiency of the $E_T^{\rm miss}$ requirement (Sec. II C) is sufficiently below unity that degrading this resolution further has strong implications for potential trigger efficiency.

VI. SUMMARY AND CONCLUSION

We have explored the experimental feasibility of charm-jet cross-section measurements in charged-current DIS at the future Electron-Ion Collider. We use parametrized detector simulations with the DELPHES package with baseline parameters for the EIC detectors. We estimated the performance of an high-impact-parameter track-counting algorithm to tag charm jets. We also explored the potential of particle identification to increment tagging efficiency. Our feasibility studies suggest that the prospects for constraining unpolarized nucleon strangeness are rather promising in this channel, while the strange helicity is more demanding yet likely feasible in the long-term. These goals represent a challenge that demands high luminosity as well as a well-designed EIC detector with good capabilities for measuring displaced vertices, particle ID, jets, and missing-transverse energy. As such, it represents a robust platform on which to inform the design of the EIC detectors.

The charm-tagging performance studies advanced in this work have the potential to extend the rapidly emerging field of jet studies for the future EIC [25–29, 73, 76–110]. In particular, charm-jet tagging approaches could be applied to neutral current boson-gluon fusion (*e.g.*, see pp. 289 of Ref. [111]) or photo-production processes. Exploiting CC DIS charm-jet measurements to constrain the nucleon's quark-gluon structure will also require continued advances in precision QCD and global analyses in order to ensure the stability of the eventual PDF extractions we envision.

CODE AVAILABILITY

The Delphes configuration file for the EIC general-purpose detector used in this work can be found in: `https://github.com/miguelignacio/delphes_EIC`

ACKNOWLEDGEMENTS

We are grateful to Pavel Nadolsky for helpful inputs, especially related to PDF choices and theory implications. We thank Oleg Tsai for insightful discussions on calorimetry technology for EIC detectors. We thank the members of the EIC User Group for many insightful discussions during the Yellow Report activities. M.A and Y. F acknowledge support through DOE Contract No. DE-AC05-06OR23177 under which JSA operates the Thomas Jefferson National Accelerator Facility. T.J.H., S.J.S., and F.O. acknowledge support through US DOE grant DE-SC0010129. T.J.H. also acknowledges support from a JLab EIC Center Fellowship. We gratefully acknowledge SMU's Center for Research Computation for their support and for the use of the SMU ManeFrame II high-performance computing cluster, which enabled a portion

of the simulation and analysis work in this paper.

[1] A. Accardi *et al.*, Eur. Phys. J. **A52**, 268 (2016), arXiv:1212.1701 [nucl-ex].

[2] H. Abramowicz *et al.* (ZEUS), Phys. Rev. D 89, 072007 (2014), 10.1103/PhysRevD.89.072007.

[3] I. Schienbein, J. Yu, K. Kovarik, C. Keppel, J. Morfin, F. Olness, and J. Owens, Phys. Rev. D **80**, 094004 (2009), arXiv:0907.2357 [hep-ph].

[4] K. Kovarik, I. Schienbein, F. Olness, J. Yu, C. Keppel, J. Morfin, J. Owens, and T. Stavreva, Phys. Rev. Lett. **106**, 122301 (2011), arXiv:1012.0286 [hep-ph].

[5] E. C. Aschenauer, I. Borsa, R. Sassot, and C. Van Hulse, Phys. Rev. **D99**, 094004 (2019), arXiv:1902.10663 [hep-ph].

[6] N. Sato, C. Andres, J. Ethier, and W. Melnitchouk (JAM), Phys. Rev. D **101**, 074020 (2020), arXiv:1905.03788 [hep-ph].

[7] A. Bazarko *et al.* (CCFR), Z. Phys. C **65**, 189 (1995), arXiv:hep-ex/9406007.

[8] M. Goncharov *et al.* (NuTeV), Phys. Rev. D **64**, 112006 (2001), arXiv:hep-ex/0102049.

[9] F. Olness, J. Pumplin, D. Stump, J. Huston, P. M. Nadolsky, H. Lai, S. Kretzer, J. Owens, and W. Tung, Eur. Phys. J. C **40**, 145 (2005), arXiv:hep-ph/0312323.

[10] O. Samoylov *et al.* (NOMAD), Nucl. Phys. B **876**, 339 (2013), arXiv:1308.4750 [hep-ex].

[11] A. Kayis-Topaksu *et al.*, New J. Phys. **13**, 093002 (2011), arXiv:1107.0613 [hep-ex].

[12] A. Airapetian *et al.* (HERMES), Phys. Lett. B **666**, 446 (2008), arXiv:0803.2993 [hep-ex].

[13] M. Alekseev *et al.* (COMPASS), Phys. Lett. B **680**, 217 (2009), arXiv:0905.2828 [hep-ex].

[14] G. Aad *et al.* (ATLAS), Phys. Rev. Lett. **109**, 012001 (2012), arXiv:1203.4051 [hep-ex].

[15] S. Chatrchyan *et al.* (CMS), JHEP **02**, 013 (2014), arXiv:1310.1138 [hep-ex].

[16] G. Aad *et al.* (ATLAS), JHEP **05**, 068 (2014), arXiv:1402.6263 [hep-ex].

[17] A. Deur, S. J. Brodsky, and G. F. De Téramond, Rept. Prog. Phys. **82** (2019), 10.1088/1361-6633/ab0b8f, arXiv:1807.05250 [hep-ph].

[18] T. Hobbs, M. Alberg, and G. A. Miller, Phys. Rev. C **93**, 052801 (2016), arXiv:1601.01729 [astro-ph.HE].

[19] H. Abramowicz *et al.* (H1, ZEUS), Eur. Phys. J. C **75**, 580 (2015), arXiv:1506.06042 [hep-ex].

[20] S. Chekanov *et al.* (ZEUS), Eur. Phys. J. **C31**, 149 (2003), arXiv:hep-ex/0306018 [hep-ex].

[21] S. Chekanov *et al.* (ZEUS), Phys. Rev. **D78**, 032004 (2008), arXiv:0802.3955 [hep-ex].

[22] T. Gehrmann, A. Huss, J. Niehues, A. Vogt, and D. M. Walker, Phys. Lett. **B792**, 182 (2019), arXiv:1812.06104 [hep-ph].

[23] I. Abt *et al.* (ZEUS), JHEP **05**, 201 (2019), arXiv:1904.03261 [hep-ex].

[24] S. Kretzer and M. Stratmann, Eur. Phys. J. C **10**, 107 (1999), arXiv:hep-ph/9902426.

[25] M. Arratia, Y. Song, F. Ringer, and B. Jacak, (2019), arXiv:1912.05931 [nucl-ex].

[26] B. Page, X. Chu, and E. Aschenauer, Phys. Rev. D **101**, 072003 (2020), arXiv:1911.00657 [hep-ph].

[27] L. Zheng, E. Aschenauer, J. Lee, B.-W. Xiao, and Z.-B. Yin, Phys. Rev. D **98**, 034011 (2018), arXiv:1805.05290 [hep-ph].

[28] E.-C. Aschenauer, K. Lee, B. Page, and F. Ringer, Phys. Rev. D **101**, 054028 (2020), arXiv:1910.11460 [hep-ph].

[29] X. Chu, E.-C. Aschenauer, J.-H. Lee, and L. Zheng, Phys. Rev. **D96**, 074035 (2017), arXiv:1705.08831 [nucl-ex].

[30] E. C. Aschenauer, T. Burton, T. Martini, H. Spiesberger, and M. Stratmann, Phys. Rev. **D88**, 114025 (2013), arXiv:1309.5327 [hep-ph].

[31] H. Abdolmaleki *et al.* (xFitter Developers' Team), Eur. Phys. J. C **79**, 864 (2019), arXiv:1907.01014 [hep-ph].

[32] J. Abelleira Fernandez *et al.* (LHeC Study Group), J. Phys. G **39**, 075001 (2012), arXiv:1206.2913 [physics.acc-ph].

[33] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008), arXiv:0710.3820 [hep-ph].

[34] BNL, "An electron-ion collider study," `https://wiki.bnl.gov/eic/upload/EIC.Design.Study.pdf` (2020).

[35] M. Aivazis, F. I. Olness, and W.-K. Tung, Phys. Rev. Lett. **65**, 2339 (1990).

[36] R. Thorne and R. Roberts, Eur. Phys. J. C **19**, 339 (2001), arXiv:hep-ph/0010344.

[37] B. Harris and J. Smith, Phys. Rev. D **57**, 2806 (1998), arXiv:hep-ph/9706334.

[38] E. A. et al., "Electron-ion collider detector requirements and r&d handbook," `http://www.eicug.org/web/sites/default/files/EIC_HANDBOOK_v1.2.pdf` (2020).

[39] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3), JHEP **02**, 057 (2014), arXiv:1307.6346 [hep-ex].

[40] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), arXiv:0802.1189 [hep-ph].

[41] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. **C72**, 1896 (2012), arXiv:1111.6097 [hep-ph].

[42] P. Newman and M. Wing, Rev. Mod. Phys. **86**, 1037 (2014), arXiv:1308.3368 [hep-ex].

[43] E. Aschenauer *et al.*, (2014), arXiv:1409.1633 [physics.acc-ph].

[44] V. Morozov *et al.*, in *10th International Particle Accelerator Conference* (2019) p. WEPGW123.

[45] U. Amaldi *et al.*, in *ECFA Study of an ep Facility for Europe* (1979) pp. 377–414.

[46] M. Tanabashi *et al.* (Particle Data Group), Phys. Rev. D **98**, 030001 (2018).

[47] G. Punzi, eConf **C030908**, MODT002 (2003), arXiv:physics/0308063.

[48] T.-J. Hou *et al.*, (2019), arXiv:1912.10053 [hep-ph].

[49] M. Aaboud *et al.* (ATLAS), Eur. Phys. J. C **77**, 367 (2017), arXiv:1612.03016 [hep-ex].

[50] A. Accardi, L. Brady, W. Melnitchouk, J. Owens, and N. Sato, Phys. Rev. D **93**, 114017 (2016), arXiv:1602.03154 [hep-ph].

[51] A. Signal and A. W. Thomas, Phys. Lett. B **191**, 205 (1987).

[52] W.-C. Chang and J.-C. Peng, Phys. Rev. Lett. **106**, 252002 (2011), arXiv:1102.5631 [hep-ph].

[53] T. Hobbs, M. Alberg, and G. A. Miller, Phys. Rev. C **91**, 035205 (2015), arXiv:1412.4871 [nucl-th].

[54] P. Junnarkar and A. Walker-Loud, Phys. Rev. D **87**, 114510 (2013), arXiv:1301.1114 [hep-lat].

[55] J. Liang, M. Sun, Y.-B. Yang, T. Draper, and K.-F. Liu, (2019), arXiv:1901.07526 [hep-ph].

[56] H. Lai, P. M. Nadolsky, J. Pumplin, D. Stump, W. Tung, and C.-P. Yuan, JHEP **04**, 089 (2007), arXiv:hep-ph/0702268.

[57] D. A. Mason, *Measurement of the strange - antistrange asymmetry at NLO in QCD from NuTeV dimuon data*, Ph.D. thesis, Oregon U. (2006).

[58] R. D. Ball *et al.* (NNPDF), Eur. Phys. J. C **77**, 663 (2017), arXiv:1706.00428 [hep-ph].

[59] R. S. Thorne, S. Bailey, T. Cridge, L. A. Harland-Lang, A. Martin, and R. Nathvani, PoS **DIS2019**, 036 (2019), arXiv:1907.08147 [hep-ph].

[60] A. Kusina *et al.* (nCTEQ), (2020).

[61] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, H. Lai, and W. Tung, Phys. Rev. D **65**, 014012 (2001), arXiv:hep-ph/0101051.

[62] A. Kusina, F. Lyonnet, D. Clark, E. Godat, T. Jezo, K. Kovarik, F. Olness, I. Schienbein, and J. Yu, Eur. Phys. J. C **77**, 488 (2017), arXiv:1610.02925 [nucl-th].

[63] J. R. West, S. J. Brodsky, G. F. de Teramond, A. S. Goldhaber, and I. Schmidt, (2020), arXiv:2004.14659 [hep-ph].

[64] J. Arrington and N. Fomin, Phys. Rev. Lett. **123**, 042501 (2019), arXiv:1903.12535 [nucl-ex].

[65] S. J. Brodsky, I. Schmidt, and S. Liuti, (2019), arXiv:1908.06317 [hep-ph].

[66] J. Ryckebusch, W. Cosyn, T. Vieijra, and C. Casert, Phys. Rev. C **100**, 054620 (2019), arXiv:1907.07259 [nucl-th].

[67] E. Segarra, A. Schmidt, T. Kutz, D. Higinbotham, E. Piasetzky, M. Strikman, L. Weinstein, and O. Hen, Phys. Rev. Lett. **124**, 092002 (2020), arXiv:1908.02223 [nucl-th].

[68] R. Wang, R. Dupre, Y. Huang, B. Zhang, and S. Niccolai, Phys. Rev. C **99**, 035205 (2019), arXiv:1806.09148 [hep-ph].

[69] I. C. Cloet, W. Bentz, and A. W. Thomas, Phys. Rev. Lett. **109**, 182301 (2012), arXiv:1202.6401 [nucl-th].

[70] I. Cloet, W. Bentz, and A. Thomas, Phys. Rev. Lett. **102**, 252301 (2009), arXiv:0901.3559 [nucl-th].

[71] W.-C. Chang, I. Cloet, D. Dutta, and J.-C. Peng, Phys. Lett. B **720**, 188 (2013), arXiv:1109.3108 [hep-ph].

[72] S. J. Brodsky, I. Schmidt, and J.-J. Yang, Phys. Rev. **D70**, 116003 (2004), arXiv:hep-ph/0409279 [hep-ph].

[73] I. Borsa, D. de Florian, and I. Pedron, (2020), arXiv:2005.10705 [hep-ph].

[74] M. Aaboud *et al.* (ATLAS), Phys. Rev. Lett. **120**, 211802 (2018), arXiv:1802.04329 [hep-ex].

[75] M. Derrick, D. Gacek, N. Hill, B. Musgrave, R. Noland, E. Petereit, J. Repond, R. Stanek, and K. Sugano, Nucl. Instrum. Meth. A **309**, 77 (1991).

[76] G. Peccini, L. Moriggi, and M. Machado, (2020), arXiv:2003.13882 [hep-ph].

[77] V. Guzey and M. Klasen, JHEP **05**, 074 (2020), arXiv:2004.06972 [hep-ph].

[78] V. Guzey and M. Klasen, (2020), arXiv:2003.09129 [hep-ph].

[79] Z.-B. Kang, K. Lee, and F. Zhao, (2020), arXiv:2005.02398 [hep-ph].

[80] X. Li *et al.*, *Proceedings, 49th International Symposium on Multiparticle Dynamics (ISMD 2019): Santa Fe (NM), United States*, EPJ Web Conf. **235**, 04002 (2020), arXiv:2002.05880 [nucl-ex].

[81] D. Gutierrez-Reyes, I. Scimemi, W. J. Waalewijn, and L. Zoppi, JHEP **10**, 031 (2019), arXiv:1904.04259 [hep-ph].

[82] D. Gutierrez-Reyes, Y. Makris, V. Vaidya, I. Scimemi, and L. Zoppi, JHEP **08**, 161 (2019), arXiv:1907.05896 [hep-ph].

[83] Y.-Y. Zhang, G.-Y. Qin, and X.-N. Wang, Phys. Rev. D **100**, 074031 (2019), arXiv:1905.12699 [hep-ph].

[84] Y. Hatta, N. Mueller, T. Ueda, and F. Yuan, (2019), arXiv:1907.09491 [hep-ph].

[85] H. Mäntysaari, N. Mueller, and B. Schenke, Phys. Rev. **D99**, 074004 (2019), arXiv:1902.05087 [hep-ph].

[86] U. D'Alesio, F. Murgia, C. Pisano, and P. Taels, (2019), arXiv:1908.00446 [hep-ph].

[87] R. Kishore, A. Mukherjee, and S. Rajesh, (2019), arXiv:1908.03698 [hep-ph].

[88] D. Kang and T. Maji, PoS **LC2019**, 061 (2019), arXiv:1912.10656 [hep-ph].

[89] K. Roy and R. Venugopalan, (2019), arXiv:1911.04530 [hep-ph].

[90] F. Salazar and B. Schenke, Phys. Rev. **D100**, 034007 (2019), arXiv:1905.03763 [hep-ph].

[91] D. Gutierrez-Reyes, I. Scimemi, W. J. Waalewijn, and L. Zoppi, Phys. Rev. Lett. **121**, 162001 (2018), arXiv:1807.07573 [hep-ph].

[92] R. Boughezal, F. Petriello, and H. Xing, Phys. Rev. **D98**, 054031 (2018), arXiv:1806.07311 [hep-ph].

[93] M. Klasen and K. Kovařík, Phys. Rev. **D97**, 114013 (2018), arXiv:1803.10985 [hep-ph].

[94] A. Dumitru, V. Skokov, and T. Ullrich, Phys. Rev. **C99**, 015204 (2019), arXiv:1809.02615 [hep-ph].

[95] X. Liu, F. Ringer, W. Vogelsang, and F. Yuan, Phys. Rev. Lett. **122**, 192003 (2019), arXiv:1812.08077 [hep-ph].

[96] M. D. Sievert and I. Vitev, Phys. Rev. **D98**, 094010 (2018), arXiv:1807.03799 [hep-ph].

[97] M. Klasen, K. Kovarik, and J. Potthoff, Phys. Rev. **D95**, 094013 (2017), arXiv:1703.02864 [hep-ph].

[98] P. Hinderer, M. Schlegel, and W. Vogelsang, Phys. Rev. **D96**, 014002 (2017), arXiv:1703.10872 [hep-ph].

[99] G. Abelof, R. Boughezal, X. Liu, and F. Petriello, Phys. Lett. **B763**, 52 (2016), arXiv:1607.04921 [hep-ph].

[100] Y. Hatta, B.-W. Xiao, and F. Yuan, Phys. Rev. Lett. **116**, 202301 (2016), arXiv:1601.01585 [hep-ph].

[101] A. Dumitru and V. Skokov, Phys. Rev. **D94**, 014030 (2016), arXiv:1605.02739 [hep-ph].

[102] D. Boer, P. J. Mulders, C. Pisano, and J. Zhou, JHEP **08**, 001 (2016), arXiv:1605.07934 [hep-ph].

[103] A. Dumitru, T. Lappi, and V. Skokov, Phys. Rev. Lett. **115**, 252301 (2015), arXiv:1508.04438 [hep-ph].

[104] P. Hinderer, M. Schlegel, and W. Vogelsang, Phys. Rev. **D92**, 014001 (2015), [Erratum: Phys. Rev.D93,no.11,119903(2016)], arXiv:1505.06415 [hep-ph].

[105] T. Altinoluk, N. Armesto, G. Beuf, and A. H. Rezaeian, Phys. Lett. **B758**, 373 (2016), arXiv:1511.07452 [hep-ph].

[106] D. Kang, C. Lee, and I. W. Stewart, Phys. Rev. **D88**, 054004 (2013), arXiv:1303.6952 [hep-ph].

[107] C. Pisano, D. Boer, S. J. Brodsky, M. G. A. Buffing, and P. J. Mulders, JHEP **10**, 024 (2013), arXiv:1307.3417 [hep-ph].

[108] Z.-B. Kang, S. Mantry, and J.-W. Qiu, Phys. Rev. **D86**, 114011 (2012), arXiv:1204.5469 [hep-ph].

[109] Z.-B. Kang, A. Metz, J.-W. Qiu, and J. Zhou, Phys. Rev. **D84**, 034046 (2011), arXiv:1106.3514 [hep-ph].

[110] D. Boer, S. J. Brodsky, P. J. Mulders, and C. Pisano, Phys. Rev. Lett. **106**, 132001 (2011), arXiv:1011.4225 [hep-ph].

[111] C. A. Aidala *et al.*, (2020), arXiv:2002.12333 [hep-ph].