

Experimental aspects of jet physics at a future EIC

B. S. Page¹, Xiaoxuan Chu¹, and E. C. Aschenauer¹

Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA



(Received 6 February 2020; accepted 2 April 2020; published 16 April 2020)

In this work, we present an overview of experimental considerations relevant to the utilization of jets at a future electron-ion collider, a subject which has been largely overlooked up to this point. A comparison of jet-finding algorithms and resolution parameters is presented along with a detailed analysis of basic jet quantities, such as multiplicities and kinematic distributions. A characterization of the energy in the event not associated with a jet is also made. In addition, detector requirements and the effects of realistic detector resolutions are discussed. Finally, an example analysis is presented in which dijets are used to access the gluon helicity contribution to the spin of the proton.

DOI: [10.1103/PhysRevD.101.072003](https://doi.org/10.1103/PhysRevD.101.072003)

I. INTRODUCTION

Jet observables have proven to be powerful tools for the exploration of the subatomic world in high energy collider environments (see for example [1]). Early jet measurements at e^+e^- colliders established the spin-1/2 nature of quarks as well as the existence and spin properties of gluons and helped solidify quantum chromodynamics (QCD) as the correct theory of the strong interaction. At hadron-hadron colliders (as well as the lepton-proton collider HERA), jets have become indispensable tools for studies ranging from the determination of both polarized and unpolarized parton distribution functions (PDFs) to the exploration of the electroweak sector to beyond the standard model searches; for examples please see [2–11] and references therein. Advances in background subtraction and substructure techniques have also made jets attractive probes of the hot dense medium created in heavy ion collisions [12,13]. Jet data coming from modern colliders such as the LHC are being matched by ever more sophisticated theoretical understanding, with calculations for many channels reaching next-to-next-to-leading order (NNLO) in α_s , and including the resummation of relevant large logarithms. The combination of advanced experimental techniques and theoretical power make jet observables true precision probes.

Given the utility of jets in other collider settings, it is logical to explore their potential applications at the proposed electron-ion collider (EIC). Several topics important to the physics goals of an EIC have been identified which

may benefit from jet analyses, including accessing the gluon Wigner distribution [14,15], probing the linearly polarized Weizsäcker-Williams gluon transverse momentum dependent distributions [16–19] and the gluon Sivers function [20], exploring the (un)polarized hadronic structure of the photon [21], constraining (un)polarized quark and gluon PDFs at moderate to high momentum fraction (x) values [22], and studies of hadronization and cold nuclear matter properties [23,24]. There are two features inherent to jets which make them attractive probes for the above topics. The first being that jets are good surrogates for the scattered partons due to the fact that they contain many of the final state particles that arise as the parton hadronizes and thus more accurately represent the parton kinematics than single particle observables. The second property is that jets have substructure, which characterizes the distribution of energy within the jet in a rigorous way [25–27]. Studying how substructure is modified between $e + p$ and $e + A$ collisions could provide information about how partons lose energy in the cold nuclear medium [28].

Although several studies of the impact jets may have at an EIC for specific topics have been performed, no dedicated exploration of the experimental issues surrounding jet finding at an EIC has been done. And, while jets were studied extensively at HERA, the lower center-of-mass energy envisioned for an EIC means that many of the jet properties observed at HERA will be different at an EIC. Therefore, this paper systematically details several experimental aspects of jet physics as they are expected to manifest at an EIC, as well as outlining an example analysis. We hope this paper will serve as a resource for those interested in exploring the use of jets at a future EIC for a wide range of physics topics.

The remainder of the paper is organized as follows: Sec. II describes the Monte Carlo setup used to generate the $e + p$ events we analyze. In Sec. III, we discuss several

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. 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.

aspects of jet finding, such as the choice of algorithm and recombination parameter as well as some basic kinematic properties of jets at an EIC. Section IV characterizes energy and particle distributions from the underlying event, as provided by our Monte Carlo, which will add additional energy to the reconstructed jets. In addition the simulated underlying event activity is compared to similar measures from $p + p$ events at $\sqrt{s} = 200$ GeV performed by the STAR experiment. The effect that a realistic detector will have on the energy resolution of reconstructed jets is explored in Sec. V by smearing the momentum and energy of incoming particles according to a specific detector model. Special attention is given to the role of hadronic calorimetry at midrapidity. Section VI presents a method for accessing gluon polarization by measuring the longitudinal double-spin asymmetry for dijet final states. Strategies for isolating the appropriate partonic subprocesses utilizing cuts on the dijet kinematics are discussed, and the value and statistical precision of the expected asymmetries are shown and compared with our current knowledge of the polarized parton distribution functions. Finally, we summarize our findings in Sec. VII.

II. MONTE CARLO SAMPLE

To facilitate the studies presented in this paper, a large sample of $e + p$ pseudodata was generated using PYTHIA-6.4 [29]. Two proton PDFs were used as input, depending on the Q^2 range being simulated. For Q^2 greater than 1.0 GeV^2 , the CTEQ6.1 PDF [30] set was used, while the CTEQ5m set [31] was used for $10^{-5} < Q^2 < 1.0 \text{ GeV}^2$. The SAS PDF set [32,33] was used for cases when the partonic structure of the photon was relevant. The choice of the lower Q^2 limit was driven by the acceptance of a proposed low Q^2 -tagger, which will reside near the beampipe, outside of the central detector. The CTEQ5m PDF is used for the proton, because contrary to modern PDFs (i.e., CT, NNPDF, HERAPDF, MSTW) its value is not frozen at its input scale Q_0^2 , but allows description of the partonic structure of the proton at $Q^2 \leq Q_0^2$. The scale for the evaluation of PDFs was chosen to be the sum of the partonic transverse momentum squared and one-half times Q^2 . The requirement that the event contain a jet with $p_T > 5 \text{ GeV}/c$ ensures that the scale is perturbative even in the photoproduction region. The PYTHIA version, tune parameters, and PDF sets used here have been shown to reproduce HERA jet and hadron cross sections [20,21] as well as jet shapes [27] and are therefore appropriate for these studies. Future work would benefit from comparisons to other generators to explore sensitivities to different shower, hadronization, and underlying event models.

Events were generated at center-of-mass energies (\sqrt{s}) of 45 GeV and 141 GeV, corresponding to (electron x proton) beam energies of $10 \times 50 \text{ GeV}$ and $20 \times 250 \text{ GeV}$, respectively. These energies represent lower and higher

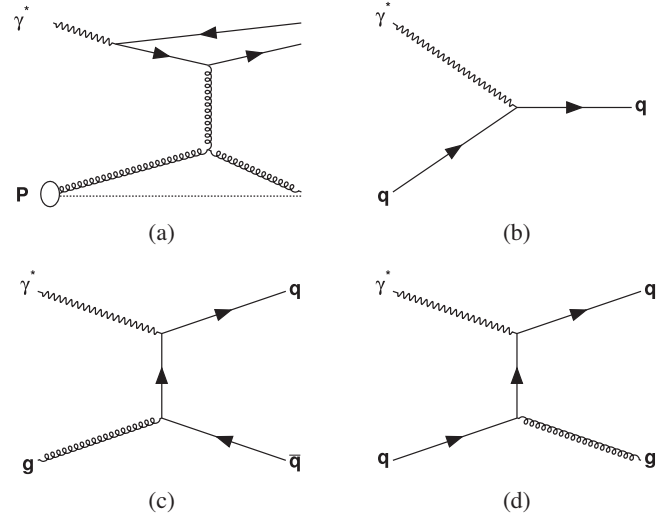


FIG. 1. Feynman diagrams for the subprocesses considered in this analysis: (a) Resolved (b) $\mathcal{O}(\alpha_s^0)$ LO DIS, (c) photon-gluon fusion (PGF) and (d) QCD Compton scattering (QCDC). Figure is from [37].

ranges generally considered for an EIC [34]. As of this writing, EIC machine designs [35] are still being finalized, so the ultimate maximum and minimum beam energies may deviate somewhat from those assumed in this paper but should not greatly affect the conclusions drawn here. It is also envisioned that the EIC will be able to operate at a number of \sqrt{s} values between these bounds.

In this paper, we only consider neutral current (NC) events, a first study on the capabilities of an EIC for charged current (CC) events can be found in [36]. The NC events fall into two categories: resolved and direct. Resolved processes [see Fig. 1(a)] are those in which the virtual photon interacts via the hadronic component of its wave function, contributing a quark or gluon to a hard-scattering with a parton from the nucleon. The resolved category includes the $qq \rightarrow qq$, $q\bar{q} \rightarrow q\bar{q}$, $q\bar{q} \rightarrow gg$, $gg \rightarrow qq$, $gg \rightarrow q\bar{q}$, $gg \rightarrow gg$ subprocesses and plays a significant role in the production of high- p_T particles at low Q^2 . In direct processes, the photon interacts as a pointlike particle with the partons of the nucleon. The subprocesses which comprise the direct category include leading order DIS 1(b) (L.O. subprocess), photon-gluon fusion (PGF) 1(c), and QCD Compton (QCDC) 1(d). Direct processes contribute more at high Q^2 values while resolved processes dominate at $Q^2 < 1 \text{ GeV}^2$. The higher order resolved, QCDC, and PGF subprocesses (hereafter referred to as the H.O. subprocesses) involve two high- p_T partons separated in azimuth and thus often give rise to dijet events. Nonperturbative minimum-bias subprocesses such as elastic scattering and soft diffraction were also generated but did not contribute any jets which passed selection criteria.

III. JET FINDING AND JET PROPERTIES

The jets used in this study were formed using stable final state particles generated by the PYTHIA Monte Carlo described in Sec. II. Here, stable refers to particles which would not normally decay in the volume of a detector, such as charged pions, kaons, protons, and neutrons. Neutral pions were allowed to decay, and the resulting (predominately) photons were passed to the jet finder. To match expected detector acceptances, only particles having transverse momenta with respect to the beam greater than 250 MeV/ c and pseudorapidity between ± 4.0 were considered candidates for inclusion in jets; the scattered lepton was not allowed to be part of the jet. The effects of using a higher minimum particle p_T cut were also explored (see Sec. III E).

A. Reference frames

At hadron colliders, analyses are carried out almost exclusively in the reference frame of the detector, the laboratory frame, as the kinematics of the interacting partons are not generally known. For DIS, however, the scattering kinematics are known event-by-event which makes it possible to boost to other frames. A particularly useful frame for jet analyses is the Breit or “brick wall” frame [38]. The Breit frame is oriented such that for the lowest order DIS process $\gamma^* q \rightarrow q'$, the virtual photon and interacting quark collide head-on along the z -axis and is boosted such that the only nonzero component of the virtual photon four-momentum is $p_z = -Q$. A consequence of this boost is that the z -momentum of the incoming quark is $Q/2$ while the scattered quark has z -momentum $-Q/2$ (hence the name “brick wall” frame) and the proton remnant has a z -momentum of $(1-x)Q/(2x)$. This leads to a natural separation between jets associated with the struck quark and those associated with the proton remnant. Working in the Breit frame has the effect of suppressing contributions from the L.O. subprocess, as the scattered quark has zero transverse momentum by construction (although, as will be seen in Sec. III C, jets can arise from the L.O. process at large Q^2 due to final state radiation). A dedicated study of L.O. jets in the laboratory frame can be found in [24]. The Breit frame is also advantageous for the study of jets arising from the H.O. subprocesses as their transverse momenta will be taken with respect to the photon-quark axis, which is the relevant quantity in these interactions. This is especially important at large values of Q^2 where the angle between the photon-quark and beam axes is significant. At low Q^2 , jet analyses are often carried out in the laboratory frame (see [20,21,27]) as the angle between the photon-quark and beam axes is negligible, meaning the Breit and laboratory frames are simply related by a longitudinal boost. In order to make the jet-finding procedure consistent, this analysis uses the Breit frame for all Q^2 . This choice does not affect our results at low Q^2 . Boosts to other frames

such as the photon-quark center-of-mass or proton rest frames can also be performed but are not covered here.

For this analysis, all particle four-momenta were boosted into the Breit frame and then passed to the jet finder for clustering. This avoids any changes in the particle content which may arise due to variations in clustering between the two frames. When it is necessary to present a jet quantity with respect to the detector, the Breit frame thrust axis of the jet was simply boosted back to the laboratory frame.

B. Jet definitions

While the idea of a jet as a collimated spray of particles is conceptually easy to grasp and jets are often easy to identify “by eye” in event displays, a well-defined method of mapping a set of particles into a set of jets is required for jets to be useful in experimental and theoretical analyses. The collection of rules that determine how particles are grouped into jets is known as a jet algorithm, while the prescription for merging the momenta of individual particles to form the overall jet momentum is known as a recombination scheme. The combination of jet algorithm, recombination scheme, and any additional parameters controlling the behavior of the jet algorithm is known as a jet definition and, as the name implies, fully defines a jet for purposes of an analysis [39].

There are a number of jet algorithms on the market [39,40] and the determination of which algorithm to utilize will depend on the requirements of the specific analysis being performed. As this manuscript is meant to give a general overview of jets at an EIC, the choice of a specific algorithm is not critical as long as there are not significant differences between algorithms for basic jet properties. To confirm this is the case, a number of jet quantities including yields, particle content, energy profile, transverse momenta, and rapidities were compared using the anti- k_T [41], k_T [42], and SIScone [43] algorithms. These were chosen as they represent the two broad categories of algorithms, sequential recombination and cone, and have seen use at both HERA and hadron colliders such as the LHC. Differences in the studied quantities ranged from negligible to a couple of percent, indicating that the choice of algorithm will have little impact on the conclusions drawn here. Because it produces jets with regular boundaries that are slightly more collimated, the anti- k_T algorithm will be used for all subsequent studies in this paper. It should be noted that several recombination schemes also exist, but only the E_Scheme [44], in which particles are combined simply by adding their four-momenta, will be considered here.

The other major component of the jet definition which needs to be determined is the resolution parameter, R , which sets the effective size of the jet. As with jet algorithm, the optimal choice for R will depend on requirements driven by a particular analysis. Often, the chosen R is

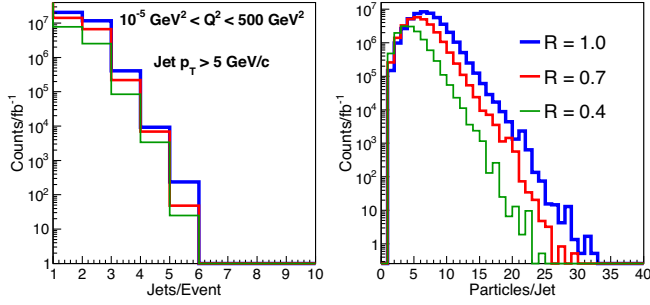


FIG. 2. Comparison of jet multiplicity (left panel) and particle multiplicity within the jet (right panel) for the anti- k_T algorithms and three resolution parameters $R = 1.0, 0.7,$ and 0.4 . The Q^2 range is between 10^{-5} GeV^2 and 500 GeV^2 and the resolved, QCDC, PGF, and leading order DIS subprocesses have been combined.

a compromise between large values that will capture more energy from the hadronizing partons and small values, which limit the contamination from underlying event (see Sec. IV for a discussion of expected underlying event activity at an EIC). The resolution parameter can be expected to affect the jet yield and particle content of jets, and because it influences how much of the energy from the hadronizing parton ends up in the jet, R should also affect how well the jet represents the underlying partonic behavior.

The jet multiplicity and number of particles within a jet are presented in the left and right panels, respectively, of Fig. 2 for R values of $1.0, 0.7,$ and 0.4 . The jets were required to have transverse momenta greater than 5 GeV/c in the Breit frame and were taken from events with Q^2 between 10^{-5} and 500 GeV^2 . The L.O. and H.O. subprocesses have been combined. It is seen that the total number of found jets increases with increasing R , which is due to the fact that larger R values admit more particles (as seen in the right panel), meaning more jets will pass the 5 GeV/c transverse momentum threshold.

One benefit of jet observables is that they can serve as proxies for hard-scattered partons because the jet will capture many of the particles emitted as the parton hadronizes. Since the size of R will affect the amount of partonic radiation included in the jet, it is necessary to assess how changes in R impact how well jets reproduce the underlying partonic kinematics. This was done by comparing the reconstructed dijet invariant mass to the parton level invariant mass. Dijets were reconstructed from H.O. events by selecting the two jets with the highest transverse momentum in the Breit frame, requiring that they be greater than 120° degrees apart in azimuth and necessitating that one jet have p_T greater than 5 GeV/c while the other has p_T greater than 4 GeV/c . The comparison between reconstructed dijet and partonic invariant mass can be seen in Fig. 3 for $R = 1.0, 0.7,$ and 0.4 . Only Q^2 values between 10 GeV^2 and 100 GeV^2 are shown as the conclusions are the same for all Q^2 values. The best agreement is seen for the largest R value and degrades as R decreases. The events at low partonic and large reconstructed dijet invariant mass arise when the one or both of the jets comprising the dijet do not match the true outgoing partons. With the exception of these “false” dijets, the fact that the reconstructed mass is consistently smaller than the partonic mass for $R < 1.0$ indicates that the smaller cones are not capturing the full energy associated with the hard-scattered partons.

Another way to assess how well jets represent the underlying partons is to measure the degree to which the jet thrust axes correspond to the directions of the partons which give rise to them. This is quantified using the distance measure ΔR , which is the quadrature sum of the difference between jet and parton rapidity and azimuthal angle ($\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$). For each jet comprising the dijet, ΔR between that jet and the two hard scattered partons is found and the minimum is taken. Figure 4 presents this minimum ΔR for both jets from the H.O. subprocesses combined for two Q^2 bins. The ΔR

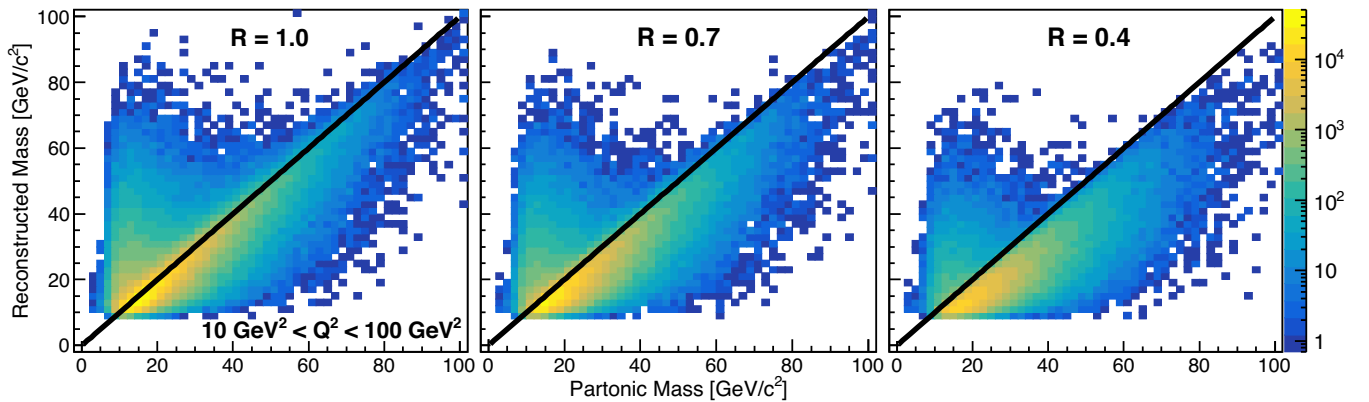


FIG. 3. Dijet invariant mass compared to the partonic invariant mass $\sqrt{\hat{s}}$ for $R = 1.0$ (left), 0.7 (middle), and 0.4 (right). Event Q^2 was required to be between 10 GeV^2 and 100 GeV^2 and the resolved, QCDC, PGF, and leading order DIS subprocesses have been combined.

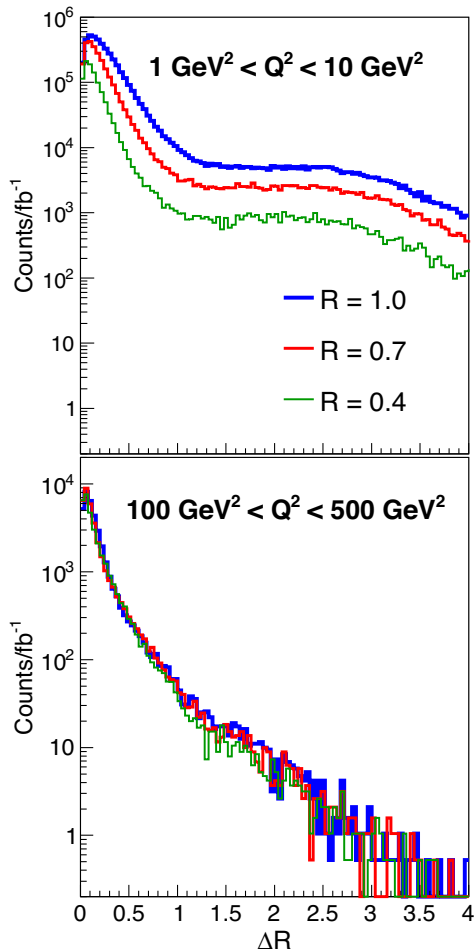


FIG. 4. Distance in rapidity-azimuth space between the jets comprising a dijet and the corresponding hard-scattered partons with Q^2 between 1 GeV^2 and 10 GeV^2 (upper panel) and 100 GeV^2 and 500 GeV^2 (lower panel) for jet radii of 1.0, 0.7, and 0.4.

distributions are larger at low Q^2 and peak closer to zero as Q^2 increases. For Q^2 between 1 and 10 GeV^2 , over 85% of the jets are within a ΔR of 0.5 to their matching parton and this jumps to 95% for Q^2 between 100 and 500 GeV^2 . Again, the tails at large ΔR arise when one or both jets in the dijet do not correspond to one of the true hard scattered partons. Surprisingly, at low Q^2 , the jet-to-parton matching is seen to improve as R decreases. This may be due to the effect of underlying event activity which will be more prevalent in larger cones and could pull the jet thrust axis slightly away from the parton direction, or it may be a selection bias effect wherein jets found with a smaller R tend to fragment in a more collimated way. Regardless, Fig. 3 shows that any benefit from the slightly better matching between the jet and parton directions at small R is overwhelmed by the parton energy missed by the smaller jet cone.

As it results in the most jets found with the largest particle content and best reproduces the partonic

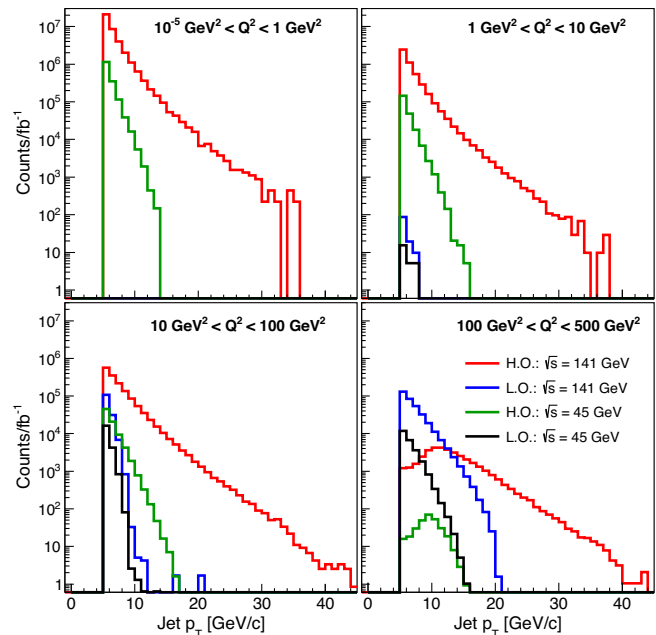


FIG. 5. Breit frame inclusive jet p_T spectra for $\sqrt{s} = 141$ and 45 GeV in Q^2 bins of 10^{-5} –1.0 GeV^2 (upper left), 1–10 GeV^2 (upper right), 10–100 GeV^2 (bottom left), and 100–500 GeV^2 (bottom right). The resolved, QCDC, and PGF subprocesses have been combined and are compared to the leading order DIS spectra and the histograms have been scaled to the counts expected for an integrated luminosity of 1 fb^{-1} .

kinematics, R will be set to 1.0 for all subsequent studies presented in this paper. Together with the choice of anti- k_T for the jet algorithm and E_{Scheme} recombination, the jet definition is fully quantified.

C. Inclusive jet kinematics

While jets arising from $e + p$ collisions were studied extensively at HERA, the lower center-of-mass energies (maximum \sqrt{s} of 141 GeV for EIC compared to 320 GeV for HERA) envisioned for an EIC make a detailed investigation of jet properties and kinematics warranted. Jets were found using the jet definition outlined in Sec. III B as implemented in FastJet-3.3.1 [45] and were required to have transverse momenta greater than 5 GeV/c in the Breit frame.

Breit frame inclusive jet p_T spectra are shown in Fig. 5 for four Q^2 ranges between 10^{-5} and 500 GeV^2 and for \sqrt{s} values of 45 and 141 GeV. The spectra have been scaled to the number of counts expected for 1 fb^{-1} of integrated luminosity. As their behaviors are similar, the H.O. subprocesses have been combined and are compared with the L.O. spectra. It is seen that high- p_T jet production is dominated by the H.O. subprocesses while the L.O. spectra are much softer. This is expected because at leading order in the Breit frame, the scattered parton moves along the z -axis and transverse momentum is generated primarily via

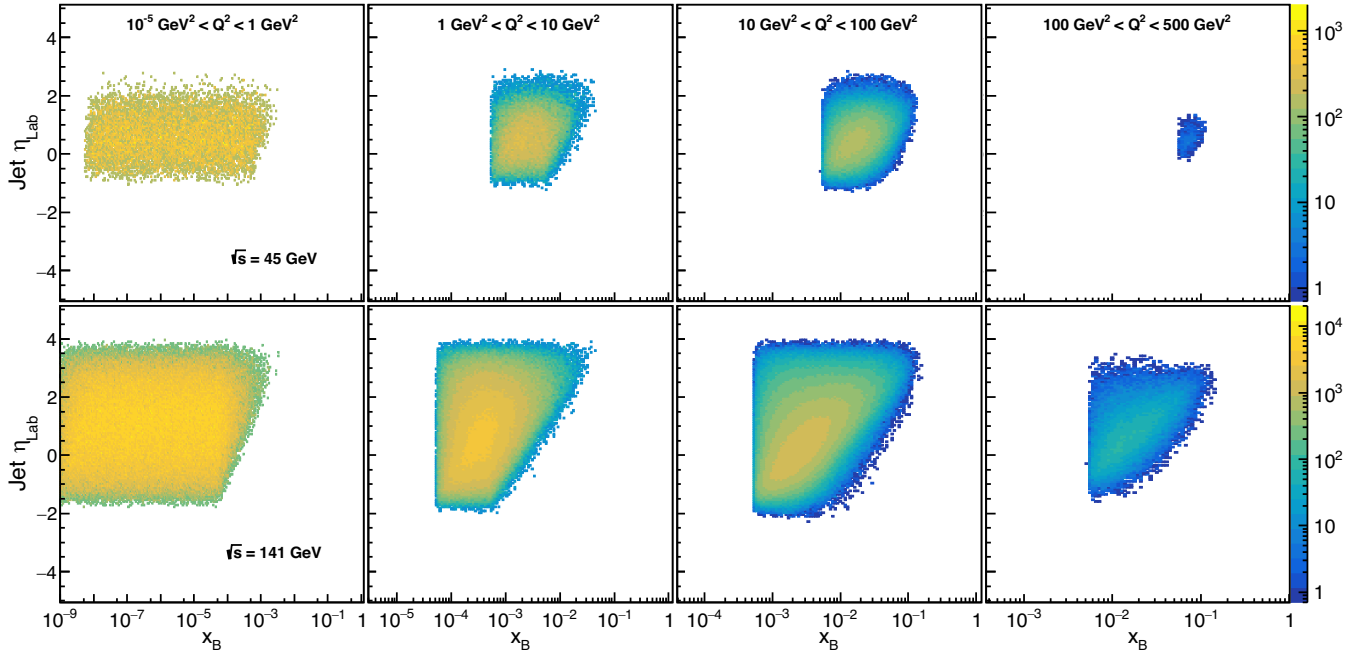


FIG. 6. Inclusive jet laboratory pseudorapidity vs x_B for Q^2 bins of 10^{-5} – 1.0 GeV^2 (left column), 1 – 10 GeV^2 (middle-left column), 10 – 100 GeV^2 (middle-right column), and 100 – 500 GeV^2 (right column) for center-of-mass energies of 45 GeV (upper row) and 141 GeV (bottom row). The resolved, QCDC, and PGF subprocesses are shown. Note that the top and bottom rows are separately scaled to the counts expected for 1 fb^{-1} of integrated luminosity.

final state radiation. The prevalence of this radiation decreases with decreasing Q^2 , so leading order DIS jets are basically absent for Q^2 below 10 GeV^2 . Note that the L.O. jet yield can increase up to 2 orders of magnitude for $Q^2 > 1 \text{ GeV}^2$ when jet-finding is performed in the lab frame, meaning this frame will be preferred for measurements focusing on L.O. jet production [24]. Figure 5 also demonstrates the critical importance of higher center-of-mass energies for jet studies as the yields can be orders of magnitude larger for $\sqrt{s} = 141 \text{ GeV}$ compared to $\sqrt{s} = 45 \text{ GeV}$. This is most pronounced at large p_T where yield differences are so great, no practical increase in luminosity could compensate.

In addition to their typical transverse momenta, it is important to understand where jets are located in the detector and how that correlates to the x_B and Q^2 of the event. Figure 6 shows the inclusive jet pseudorapidity distributions, in the laboratory frame, as a function of x_B for four Q^2 bins ranging from 10^{-5} GeV^2 to 500 GeV^2 and \sqrt{s} values of 45 GeV and 141 GeV , for the H.O. subprocesses. Figure 7 shows the same for the L.O. subprocess in the Q^2 ranges 10 – 100 GeV^2 and 100 – 500 GeV^2 . As expected from basic DIS kinematics, smaller x_B values are probed by larger center-of-mass energies and visa versa. Comparing Figs. 6 and 7, it is seen that the leading order DIS jet pseudorapidity is more strongly correlated with x_B than that of the resolved, QCDC, or PGF jets. Because there is only one outgoing parton in DIS, jet pseudorapidity should be

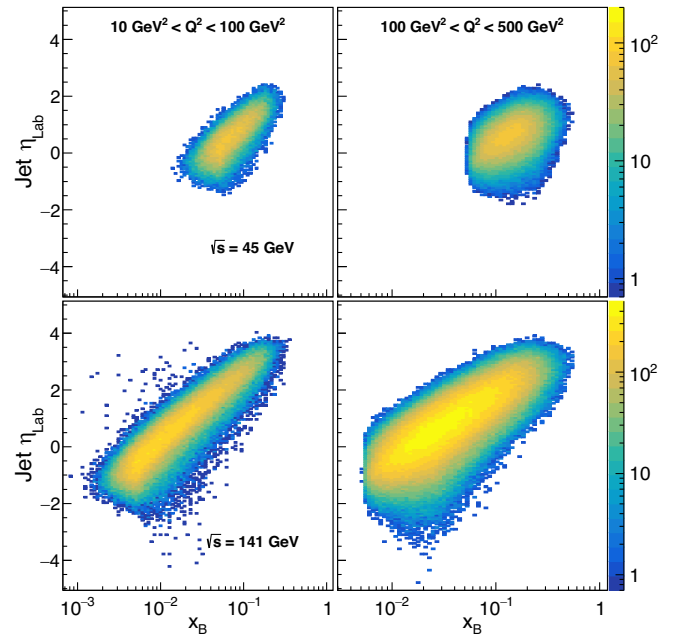


FIG. 7. Inclusive jet laboratory pseudorapidity vs x_B for Q^2 bins of 10 – 100 GeV^2 (left column) and 100 – 500 GeV^2 (right column) and center-of-mass energies of 45 GeV (upper row) and 141 GeV (bottom row). Only the leading order DIS subprocess is shown. Note that the top and bottom rows are separately scaled to the counts expected for 1 fb^{-1} of integrated luminosity.

strongly determined by the event x_B and Q^2 with the observed width of the distributions in Fig. 7 due to the finite Q^2 ranges and, more importantly, final state radiation altering the trajectory of the outgoing quark. The presence of a second hard parton in H.O. events breaks the strong relationship between x_B , Q^2 , and η and allows the resultant jets to fill the kinematically allowed phase space. The importance of hadron beam energy to jet position can also be seen in Figs. 6 and 7 by contrasting the distributions at the two \sqrt{s} values for given x_B and Q^2 . Larger hadron beam momenta impart more of a boost to final state particles, so jets at $\sqrt{s} = 141$ GeV will be pushed to higher pseudorapidities compared to jets from collisions at lower \sqrt{s} . Thus, good forward tracking and calorimetry capabilities will be needed to utilize jets at large \sqrt{s} .

D. Dijet kinematics

So far, only inclusive jet quantities have been considered, yet as stated above, the H.O. subprocesses naturally give rise to correlated two jet final states (dijets). By measuring the properties of both jets in coincidence, dijets can provide information on the leading order kinematics of the hard scattering event, such as the momentum fraction contributed by the virtual photon. Several studies have already explored the utility of dijet measurements at the EIC [16,20,21], and a further study will be presented in Sec. VI. As before, dijets were selected by identifying the two jets with the largest transverse momenta in the Breit frame and requiring them to be greater than 120° apart in azimuth. It was further required that one jet have p_T greater than 5 GeV/ c while the other have p_T greater than 4 GeV/ c .

The scale relevant for a dijet is its invariant mass, which is simply $\sqrt{(\mathcal{P}_1 + \mathcal{P}_2)^2}$ where \mathcal{P}_1 and \mathcal{P}_2 are the four-momenta of the two jets. The dijet invariant mass spectra are shown in Fig. 8 for four Q^2 ranges and center-of-mass energies of 45 and 141 GeV. The H.O. subprocesses are combined and compared to the L.O. spectra. While the leading order DIS subprocess results in only one outgoing quark, at high Q^2 , the proton remnant can receive enough transverse momentum to produce a second jet which will satisfy the dijet conditions. These L.O. dijets can be effectively separated from the H.O. dijets via a cut on the ratio of dijet mass over Q (see Fig. 22). As was the case with inclusive jet p_T , the dijet cross section is significantly larger for \sqrt{s} of 141 GeV than for 45 GeV, and the spectra extend to much higher mass values.

To characterize the location of a dijet in the detector, the pseudorapidities of both jets need to be recorded simultaneously as in Fig. 9. As before, the H.O. subprocesses have been combined, and now dijets arising from the L.O. subprocess are not shown. Only $\sqrt{s} = 141$ GeV events are shown as the $\sqrt{s} = 45$ GeV distributions are just shifted to lower pseudorapidity for a given x_B - Q^2 bin due to the smaller boost from the less energetic hadron

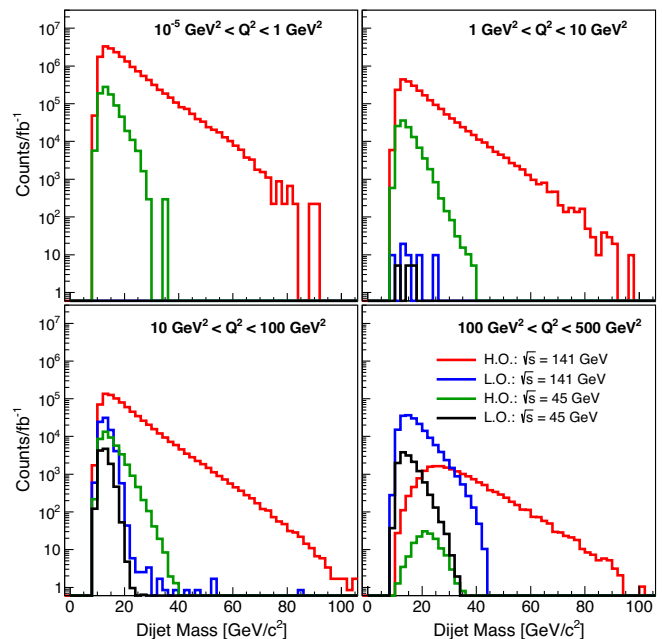


FIG. 8. Breit frame dijet invariant mass spectra for $\sqrt{s} = 141$ and 45 GeV in Q^2 bins of 10^{-5} – 1.0 GeV 2 (upper left), 1 – 10 GeV 2 (upper right), 10 – 100 GeV 2 (bottom left), and 100 – 500 GeV 2 (bottom right). The resolved, QCDC, and PGF subprocesses have been combined and are compared to the leading order DIS spectra, and the histograms have been scaled to the counts expected for an integrated luminosity of 1 fb $^{-1}$.

beam. As was the case for inclusive jets, jet pseudorapidities increase as x_B is increased at a fixed Q^2 , and for a fixed x_B bin, jet pseudorapidities decrease as Q^2 is increased.

While the absolute pseudorapidities of the two jets comprising a dijet depend on the event x_B and Q^2 as shown in Fig. 9, the relative pseudorapidity is connected to the dijet invariant mass. Expanding the four-vector expression given above, the dijet invariant mass can be approximated (ignoring the individual jet masses) as

$$M \approx \sqrt{2p_{T1}p_{T2}(\cosh(\Delta\eta) - \cos(\Delta\phi))}, \quad (1)$$

where p_{T1} and p_{T2} are the transverse momenta of the two jets and $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and azimuthal angle differences, respectively, between the two jets. In this form, it is apparent that dijets can acquire a large mass if their constituent jets have large transverse momenta, and/or if the pseudorapidity difference between the two jets is large. The interplay between jet p_T and $\Delta\eta$ is made explicit in Fig. 10 which shows the difference in pseudorapidity and average jet transverse momenta of the constituent jets for five invariant mass bins and Q^2 ranges of 10^{-5} – 1.0 GeV 2 and 100.0 – 500.0 GeV 2 . At low Q^2 , the average jet p_T remains small even for the largest mass bins, meaning that larger invariant masses are driven by greater pseudorapidity separations. As Q^2 increases, the average jet p_T increases,

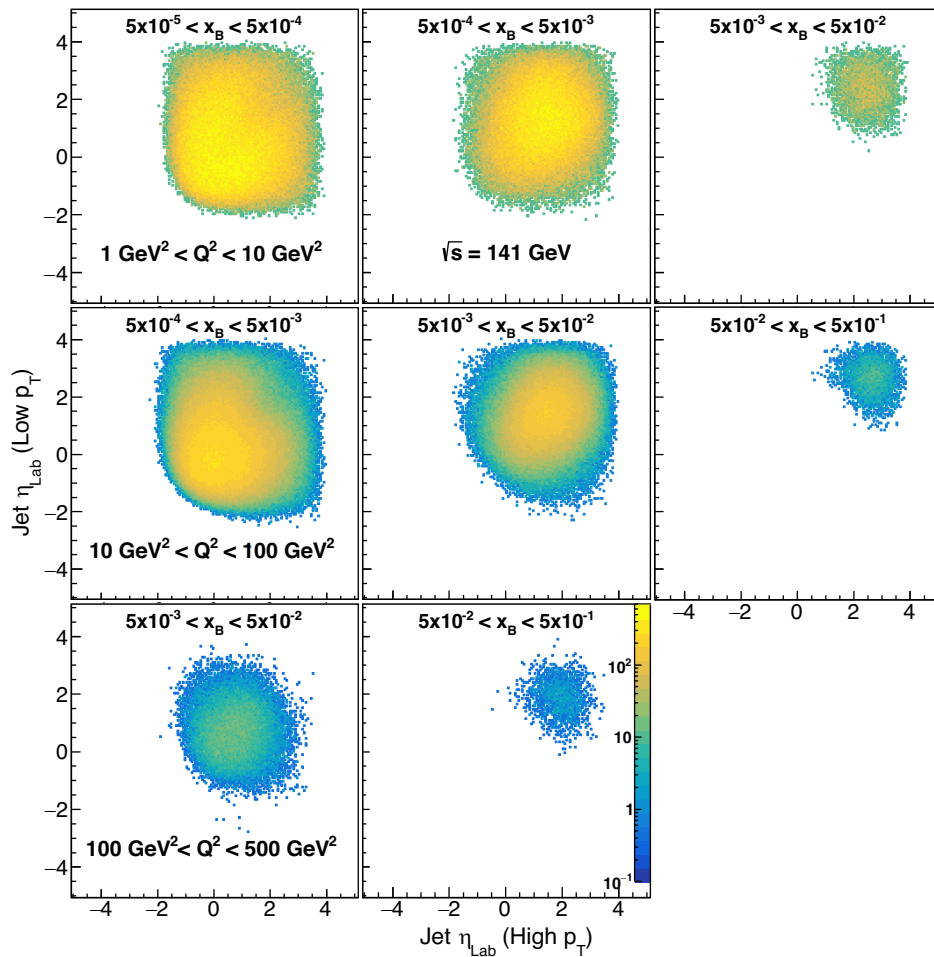


FIG. 9. Laboratory frame jet $\eta - \eta$ correlations for dijets in select x_B bins for Q^2 ranges of 1 – 10 GeV^2 (top row), 10 – 100 GeV^2 (middle row), and 100 – 500 GeV^2 (bottom row). Only the 141 GeV center-of-mass energy is shown. The resolved, QCDC, and PGF subprocesses have been combined, and leading order DIS events are not included. Note that all panels share the same scale, which has been set to represent the number of counts expected for an integrated luminosity of 1 fb^{-1} .

and $\Delta\eta$ does not need to be as large to produce a high invariant mass dijet. Thus, even for high mass dijets, it will be important to have good jet energy resolution to low p_T and large detector acceptance.

E. Minimum particle transverse momentum

The jets used in the above discussion were created from particles which had a transverse momentum of at least 250 MeV/c with respect to the beam. This value was chosen as it is slightly higher than the typical cutoff used in $p + p$ jet finding at STAR (see for example [46]), which when running at $\sqrt{s} = 200$ GeV should have relatively similar particle p_T spectra as can be expected at an EIC. The particle p_T cutoff is largely driven by detector considerations, with the magnetic field strength often the dominant factor. While detector designs for the EIC are still in active development, many include a relatively large solenoidal magnetic field of 2 to 3 T in order to provide good p_T resolution over the full pseudorapidity range. This

will limit the acceptance for low p_T charged particles as they will bend so severely in the magnetic field that they will not reach the calorimeters and will be displaced significantly from any neutral particles which arise from the hadronizing parton.

To study the effect that the loss of low p_T particles will have on jet quantities, the jet finding was rerun with the low p_T particle cutoff doubled to 500 MeV/c . The most obvious effects of raising the cutoff are a reduction in jet/dijet yields and the average number of particles in a jet. The jet and dijet yields are reduced by roughly 37% for $10^{-5} < Q^2 < 1 \text{ GeV}^2$ and 20% for $10^{-5} < Q^2 < 500 \text{ GeV}^2$. The effect of the minimum p_T cut on jet particle content can be seen in Fig. 11 for all particles as well as charged hadrons only.

Removing low p_T particles may also affect how well jets reproduce the kinematics of the underlying partons due to the loss of energy contributed by these particles. The impact of this loss was studied in the same way as the R dependence in Sec. III B, by comparing the reconstructed dijet mass to the diparton invariant mass and by measuring

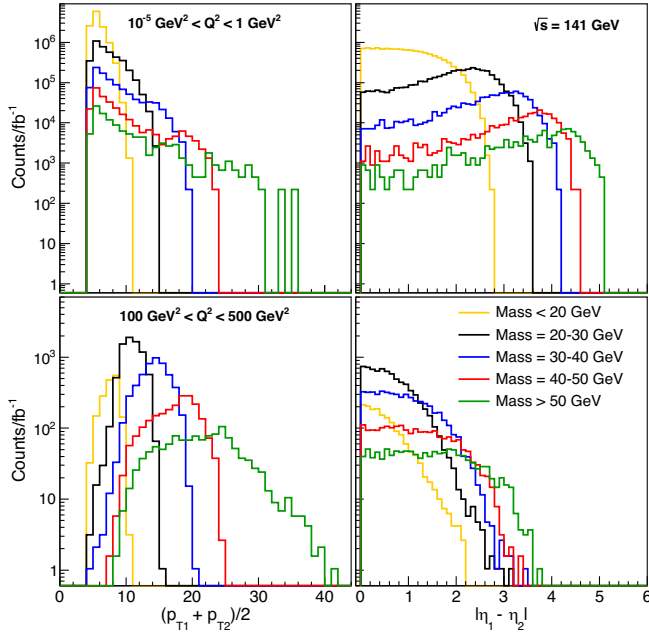


FIG. 10. Average jet p_T (left column) and rapidity difference (right column) for jets comprising a dijet separated in bins of dijet invariant mass for Q^2 ranges of 10 – 100 GeV^2 (top row) and 100 – 500 GeV^2 (bottom row). Only the 141 GeV center-of-mass energy is shown. The resolved, QCDC, and PGF subprocesses have been combined and leading order DIS events are not included. Histograms have been scaled to the counts expected for an integrated luminosity of 1 fb^{-1} .

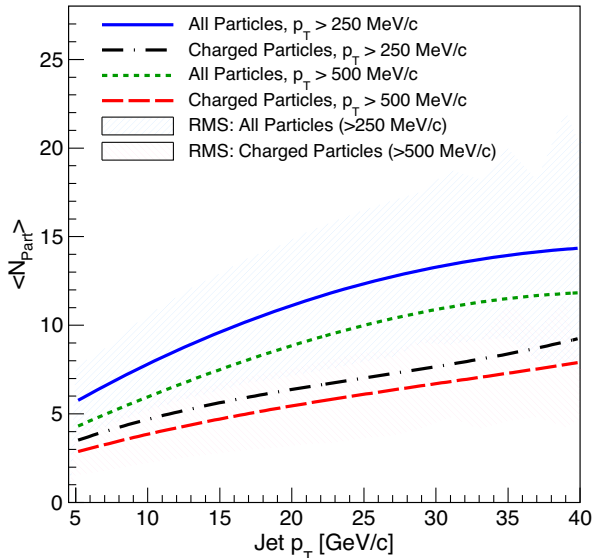


FIG. 11. The average number of particles in a jet as a function of the transverse momentum of the jet for all stable particles and only charged particles for minimum particle p_T s of 250 and 500 MeV/c . Also shown are the rms variations for all particles with $p_T > 250 \text{ MeV}/c$ and charged particles with $p_T > 500 \text{ MeV}/c$.

ΔR between the jet and parton directions. The higher minimum particle p_T cut slightly reduces the reconstructed dijet mass versus the true diparton invariant mass, much like what was seen when reducing R in Fig. 3, although the magnitude of the effect is not as great. There was no visible change in the ΔR distributions. As detector designs become more advanced, further studies will need to be made to ensure that there is sufficient acceptance for low p_T particles.

IV. UNDERLYING EVENT PROPERTIES

The underlying event activity, which contributes background energy to jet signals is quantified in this section for jets produced from the H.O. subprocesses in the Breit frame. “Underlying event” (UE) refers to those particles, which do not arise from the outgoing hard-scattered partons and can contain contributions from initial and final state radiation (ISR, FSR), beam remnants, and multiple parton interactions (MPI). As the QCDC, and PGF subprocesses proceed via a direct $\gamma +$ parton interaction, there is no contribution from MPI, and the only beam remnant arises from the hadron side. On the other hand, the resolved subprocess is defined by parton + parton scattering where one parton is supplied by the photon (see Fig 1), and because the photon behaves hadronically, it will contribute to the beam remnant and allow for MPI. The effects from MPI were modeled for photonproduction at HERA in, for example, [47], and were found to be significant in certain kinematic regions. The size of the MPI contribution was found to depend on the γ -proton center of mass energy ($\sqrt{y} \times \sqrt{s}$). The average γ -proton energy for the EIC sample is 42 GeV, which is a factor of 2.7 smaller than the minimum energy considered for the HERA results. In the model, this corresponds to a nearly factor of 4 reduction in the probability of having two scatterings in a single event. The MPI was also seen to depend on the poorly known gluon density inside the photon, which makes it difficult to disentangle MPI and PDF effects in the absence of data at EIC energies. Because of these factors, multiple parton interactions have been omitted from this study.

We utilize two methods to analyze UE effects in $e + p$ collisions, the “region method” [48] and the “off axis cone method” [49]. In the region method, the azimuthal angle of the highest p_T jet in each dijet event is selected as the reference angle and particles are grouped into one of three regions based on their azimuthal angle relative to this reference, $\Delta\phi \equiv (\phi - \phi_{\text{ref jet}})$. The particle candidate pool is identical to that used in the jet-finding. The “toward” region is defined as $|\Delta\phi| < 60^\circ$ and contains the reference jet, while the “away” region has $|\Delta\phi| > 120^\circ$ and generally contains the lower p_T , or associated, jet of the dijet. The “transverse” region is defined as $60^\circ < |\Delta\phi| < 120^\circ$ and the activity here is dominated by the UE. Figure 12 illustrates the definition of the three regions.

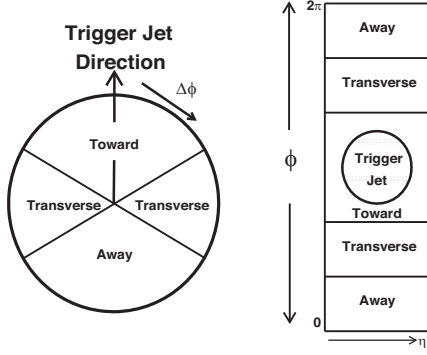


FIG. 12. Illustration of the “toward”, “away”, and “transverse” regions as defined relative to the highest p_T jet in the event. The angle $\Delta\phi \equiv \phi - \phi_{\text{ref jet}}$ is the azimuthal angle between a charged particle and the highest p_T jet, from $-\pi$ to π . The toward region is defined as $|\Delta\phi| < 60^\circ$, while the away region is $|\Delta\phi| > 120^\circ$. The transverse region is defined as $60^\circ < |\Delta\phi| < 120^\circ$. The plot is from [48].

Three observables are used to characterize UE activity: the average charged particle multiplicity ($\langle N_{\text{ch}} \rangle$), the average charged particle scalar p_T sum ($\langle \text{sum} p_T \rangle$), and average charged particle p_T ($\langle p_T \rangle$). Figure 13 presents $\langle N_{\text{ch}} \rangle$ (top) and $\langle \text{sum} p_T \rangle$ (bottom) as a function of $|\Delta\phi|$ for particles with $p_T > 250$ MeV/c and $-4 < \eta < 4$. Reference jets with $p_T > 5$ GeV/c (associated jet $p_T > 4$ GeV/c) and

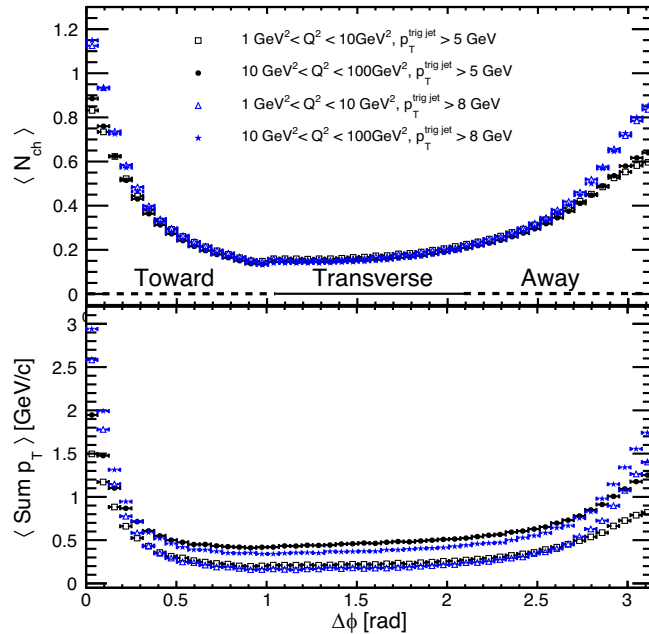


FIG. 13. Average number of charged particles (top) and average charged particle scalar p_T sum (bottom) as a function of the azimuthal angle, $\Delta\phi$, between the particle and the reference jet for $p_T^{\text{ref jet}} > 5$ GeV or 8 GeV, and $Q^2 < 10$ GeV² or $10 < Q^2 < 100$ GeV². Each point corresponds to the $\langle N_{\text{ch}} \rangle$ in a 3.6° bin.

$p_T > 8$ GeV/c (associated jet $p_T > 7$ GeV/c) were compared along with jets from two Q^2 regions: $1 \text{ GeV}^2 < Q^2 < 10 \text{ GeV}^2$ and $10 \text{ GeV}^2 < Q^2 < 100 \text{ GeV}^2$. There is a strong dependence on jet p_T in the toward and away regions for both $\langle N_{\text{ch}} \rangle$ and $\langle \text{sum} p_T \rangle$, which is expected as these regions are dominated by the jets. In the transverse region, no dependence on jet p_T is seen for the average number of charged particles while a mild difference is seen in the p_T sum for the higher Q^2 range. Interestingly, it is the lower p_T jets which show a higher p_T sum in the transverse region, seemingly indicating that jets with higher p_T leave less energy available for underlying event activity. No Q^2 dependence is seen for $\langle N_{\text{ch}} \rangle$ while larger Q^2 s lead to greater sum p_T s in all regions.

The dependence of $\langle N_{\text{ch}} \rangle$ and $\langle p_T \rangle$ on the trigger jet p_T in the toward, away, and transverse regions is made more explicit in Fig. 14 for Q^2 between 1 GeV² and 10 GeV². It is seen that both the particle density (top) and average p_T (bottom) in the toward and away region depend strongly on the trigger jet p_T as is to be expected. Conversely, both quantities show a weak anticorrelation with trigger jet p_T in the transverse region, in agreement with Fig. 13. The effect of initial and final state radiation (ISR/FSR) on the observables can be seen in the right-hand column of Fig. 14 where the radiation effects have been disabled. The presence of ISR/FSR leads to an increase in $\langle N_{\text{ch}} \rangle$ for all three regions while surprisingly, the average charged particle p_T is seen to increase somewhat without ISR/FSR effects.

Unlike the identical species configurations that are often run at colliders, the collisions at an EIC will be asymmetric in both particle type and beam energy. This will lead to an asymmetric η dependence in particle production and UE activity as seen in Fig. 15. Here, the average charged particle multiplicity densities and average charged particle p_T sum densities in the Transverse region are shown as a function of reference jet p_T for particles in backward ($-4 < \eta < -1$), mid ($-1 < \eta < 1$), and forward ($1 < \eta < 4$) pseudorapidity ranges (as defined in the laboratory frame) from the region method (filled symbols). The reference jet was required to be within the same pseudorapidity range as the particles with the added restriction that the jet η must be 0.4 units away from a range boundary in order to facilitate comparisons to the off axis cone UE characterization method. It is seen that the UE charged particle density is higher in the Forward (hadron-going) direction which is expected as any beam remnant contribution will generally follow the struck hadron. Also, particle density in the forward region will be higher due to the boost from the more energetic hadron beam.

The second technique used to investigate UE effects is the off axis cone method [49], developed by the ALICE Collaboration. The off axis cone method studies the UE on a jet-by-jet level, as opposed to the region method which is designed to study the UE on the event level. For every

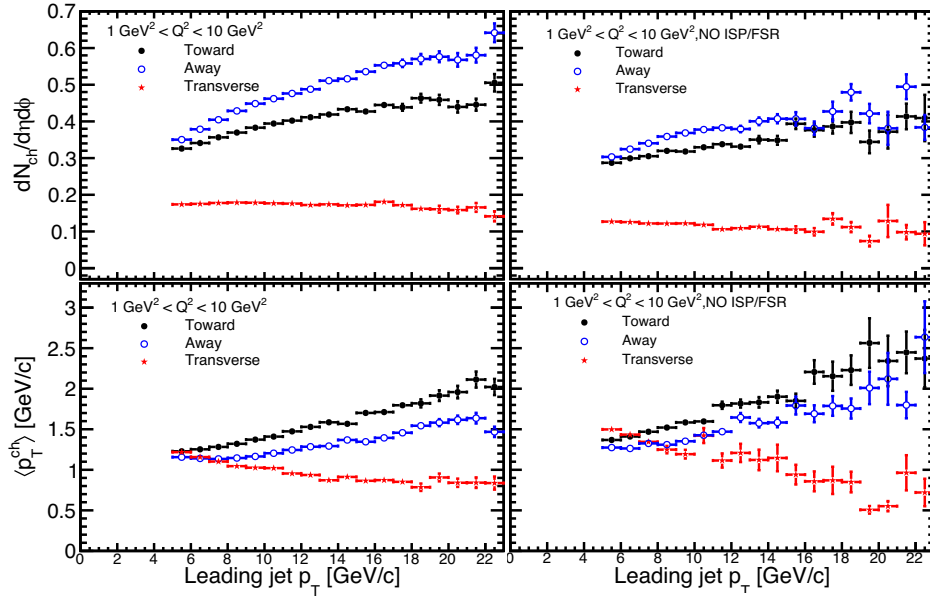


FIG. 14. Charged particle density (top row) and mean p_T (bottom row) as a function of the transverse momentum of the reference jet for the toward, away, and transverse regions. The left column is the standard simulation, while the right column has had initial and final state radiation disabled. Reference jets were required to have $p_T > 5$ GeV/c and laboratory pseudorapidity between ± 4 while Q^2 was selected to be between 1 GeV^2 and 10 GeV^2 .

reconstructed jet, two off axis cones [cone(-) and cone(+)] are defined, each of which is centered at the same η as the jet but $\pm\pi/2$ away in ϕ from the jet ϕ , as shown in Fig. 16, and the particles which fall inside these cones are used to

characterize the underlying event. The cone radius was chosen to be 0.4 so as not to overlap with the primary jet. The multiplicity density is defined as the average number of charged particles inside each cone, $\langle N_{\text{ch}} \rangle$, divided by the

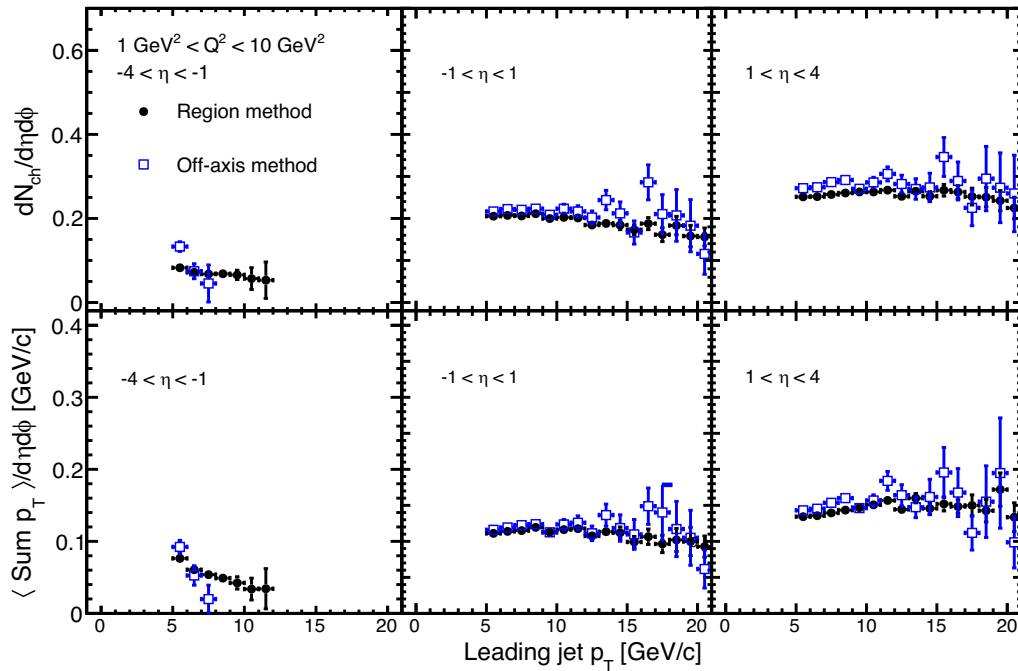


FIG. 15. Average charged particle density (top) and charged particle p_T sum densities (bottom) in the transverse region as a function of trigger jet p_T for three pseudorapidity ranges and Q^2 between 1 GeV^2 and 10 GeV^2 . Results from the regions method (closed circles) and off axis cone method (open squares) are compared. The displayed pseudorapidity ranges apply to the particles used in the analysis while the reference jets were required to be more than 0.4 units of pseudorapidity from a boundary.

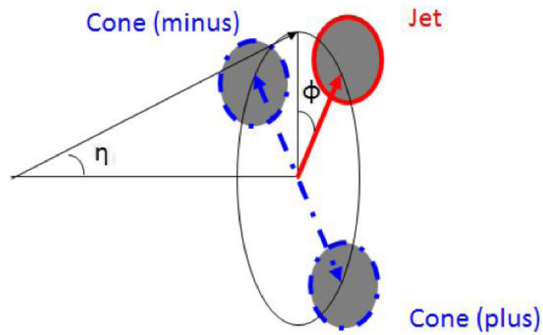


FIG. 16. Illustration of two off axis cones relative to a jet.

cone area while the $\langle p_{T\text{sum}} \rangle$ density is defined as the average off axis cone p_T divided by the cone area.

The UE results from the off axis method (open symbols) are compared to the region method (closed symbols) in Fig. 15. As with the reference jets from the region method, the jets from the off axis method were required to be more than 0.4 away from a range boundary in η so that the full off axis cone would fit in the indicated pseudorapidity bin. With the jet and particle pseudorapidities defined in this way, a faithful comparison between the region and off axis methods can be made, and the agreement is seen to be very good. The pseudorapidity dependence of the UE activity seen in Fig. 15 means the off axis cone method will be important when correcting jet quantities for underlying event contamination as it will be necessary to measure that component at the pseudorapidity of the jet.

The results described above characterize the expected underlying event at an EIC as generated in our Monte Carlo. To get a better feeling for the size of these effects and provide a sanity check on the simulation, it is instructive to compare the simulated $e + p$ results with $p + p$ data at a similar center-of-mass energy. The STAR experiment [50] at the Relativistic Heavy Ion Collider (RHIC) has performed a similar analysis of UE activity as presented here on $p + p$ data taken at $\sqrt{s} = 200$ GeV, which provides an opportunity for such a comparison. The STAR analysis [51] measured both $\langle N_{\text{ch}} \rangle$ and $\langle p_T \rangle$ for all charged particles in the midrapidity region ($-1 < \eta < 1$) with $p_T > 0.2$ GeV/ c . The STAR average charged particle density varies from 0.8 to 0.5 for reference jets with p_T of 5 GeV/ c to 40 GeV/ c . This is a factor of roughly 2 to 4 greater than what is observed at mid or forward rapidity in Fig. 15. This is not surprising as the STAR result involves the collision of two protons. However, the average charged particle p_T measured by STAR is relatively flat with a value of 0.6 GeV/ c , which is at least a factor of 2 lower than the result presented in Fig. 14. The larger $\langle p_T \rangle$ in the transverse region at the EIC is due to the boost into the Breit frame. While the particles which participate in the hard scattering processes initially move along the photon-parton axis, the underlying event particles arise largely from the proton and are thus more aligned with the beam axis. When measured with respect to

the photon-parton axis, as is done in the Breit frame, these underlying event particles acquire, on average, larger transverse momenta. When analyzed in the laboratory frame, $\langle p_T \rangle$ in the transverse region is roughly 0.6 GeV/ c , in agreement with the STAR result.

V. DETECTOR EFFECTS

The jets used for the results presented in Secs. III and IV were reconstructed at “particle level,” taking as input the exact four-momenta of all generated final state particles. These jets do not include distortions which will arise from the finite energy and momentum resolutions and inefficiencies of any real detector. Because the entirety of the EIC physics program requires high resolution calorimetry and tracking over a wide acceptance range, the induced distortions are expected to be small. Nevertheless, it is important to investigate how jets will be affected by a realistic detector environment. In order to quantify how these distortions will affect jet reconstruction, the energy and momenta of input particles were smeared based on a model EIC detector before being clustered into jets. These smeared jets were then compared to the corresponding particle level jets to study detector effects.

A. Smearing generator and detector model

Generally, detector effects are investigated by propagating simulated events through a detailed detector model which reproduces the relevant energy and momentum resolutions, efficiencies, material budgets, and readout responses of the actual device. As such detailed models for prospective EIC detectors are only starting to be developed, and key detector technology choices are still in flux, the effects of finite resolution and acceptance on jet-finding were explored using a smearing generator, which alters a particle’s energy or momentum based on a specific resolution function. While not a substitute for a full detector simulation, this smearing method has the benefit of being much faster computationally, making it easy to investigate different subdetector configurations and resolutions.

The smearing generator used allows a user to define “devices” which encode the behavior of individual or collections of detector subsystems. A single device will smear the energy, momentum, or direction of all particles which fall into its acceptance. Here, acceptance not only refers to the spatial extent of the device, but also to particle properties such as charge and how the particle interacts with the detector material (hadronically or electromagnetically). Three particle charge and interaction types are used in the smearing performed here: charged hadronic, neutral hadronic, and electromagnetic. Charged hadronic particles (assumed to be detected using a tracker) have their momenta and trajectories smeared while neutral hadronic and electromagnetic particles (assumed to be detected with calorimeters) have their energies smeared. Because a device

will only smear either the energy or momentum component of a particle, the energy-momentum-mass relationship of the smeared 4-vector will be broken. To address this, after the smearing was performed, the charged hadron energies were altered to match their momenta assuming the particles had a pion mass. Similarly, the momenta of neutral hadrons and particles interacting electromagnetically were set equal to the smeared energy, which is equivalent to assuming the particle was massless. This simplistic compensation will be inadequate for particles with significant mass, such as protons and neutrons, but is sufficient for the purpose of this study. When more complete detector simulations are developed, efforts should be made to determine the utility of the calorimeter systems as well as particle identification for more accurate particle four-momentum reconstruction.

For this study, the smearing generator devices were defined such that they would reproduce the projected behavior of BeAST, Brookhaven’s “green field” detector proposal. BeAST is built around a 3 T solenoidal magnet and will include high precision tracking detectors spanning a pseudorapidity range of $|\eta| < 3.5$, electromagnetic calorimetry covering the range $|\eta| < 4.0$, and hadron calorimetry in the forward and backward regions $1 < |\eta| < 4.0$. BeAST will also have good vertex detection and particle identification capabilities as well as instrumentation to detect particles scattered at small angles, both in the hadron and lepton beam direction, such as Roman pots and a system to tag low Q^2 electrons. However, these systems do not directly affect jet reconstruction and were therefore not included in this simulation. The calorimeter resolutions assumed for different detector regions can be found in Table I while the tracking resolution for different particle momenta as a function of pseudorapidity can be seen in Fig. 17. Several modifications to the baseline BeAST configuration were also considered, including the introduction of a track finding inefficiency factor of 5% and the addition of a midrapidity hadron calorimeter assuming a high and low energy resolution.

TABLE I. Assumed energy resolutions and pseudorapidity ranges for the electromagnetic and hadron calorimeters included in the detector smearing model.

Component	Pseudorapidity range	Resolution
Back EMCAL	$-4.0 < \eta < -2$	$\frac{1.5\%}{\sqrt{E}} \oplus 1\%$
Midback EMCAL	$-2 < \eta < -1$	$\frac{7\%}{\sqrt{E}} \oplus 1\%$
Mid EMCAL	$-1 < \eta < 1$	$\frac{10\%}{\sqrt{E}} \oplus 1\%$
Fwd EMCAL	$1 < \eta < 4.0$	$\frac{10\%}{\sqrt{E}} \oplus 1\%$
Fwd/back HCal	$1 < \eta < 4.0$	$\frac{50\%}{\sqrt{E}} \oplus 10.0\%$
Lo res mid Hcal	$-1 < \eta < 1$	$\frac{75\%}{\sqrt{E}} \oplus 15\%$
Hi res mid Hcal	$-1 < \eta < 1$	$\frac{35\%}{\sqrt{E}} \oplus 2\%$

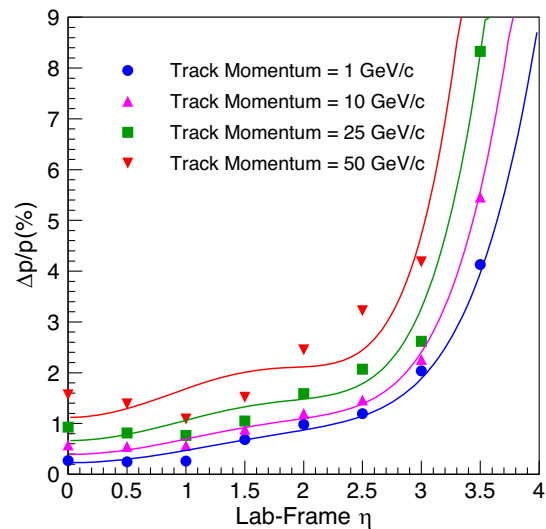


FIG. 17. Track momentum resolution assumed for the smearing generator as a function of track pseudorapidity. The points represent extractions of resolution at specific momenta and pseudorapidity from simulation of a model BeAST tracking detector and the curves are instances of the function used to fit the points that was passed to the smearing generator.

B. Smearing results

Using the smearing procedure and resolution parameters presented above, individual particle 4-momenta were altered and then passed to the jet-finder to be clustered into jets using the same procedure as for unaltered particles. Thus, for each event, there will be a set of unaltered particle level jets and a set of smeared jets. In order to evaluate how the smearing procedure has modified the properties of a given particle level jet, an association must be made between that particle level jet and a particular smeared jet. This is done by finding the smeared jet which minimizes the quantity $\Delta R = \sqrt{(y_{\text{Particle}} - y_{\text{Smeared}})^2 + (\phi_{\text{Particle}} - \phi_{\text{Smeared}})^2}$ for each particle level jet, with y being the rapidity and ϕ the azimuthal angle of the jet. Particle level and smeared jets were required to have $\Delta R < 1.0$ in order to be considered associated.

The relationship between the transverse momenta of associated particle level and smeared jets can be seen in Fig. 18 for the baseline BeAST design, as well as the 5% track finding inefficiency and midrapidity hadron calorimeter scenarios. As the baseline design does not include a hadron calorimeter at midrapidity, neutrons and K_L^0 's in this region are not detected, meaning smeared jets will tend to have lower transverse momentum than their corresponding particle level jets. The population and extent of this tail depends on the average number of neutral hadrons in the event sample and the amount of transverse momentum they carry. Removing 5% of charged particles increases somewhat the number of events which populate the off diagonal tail. The inclusion of a hadron calorimeter at midrapidity captures the remaining neutral energy making

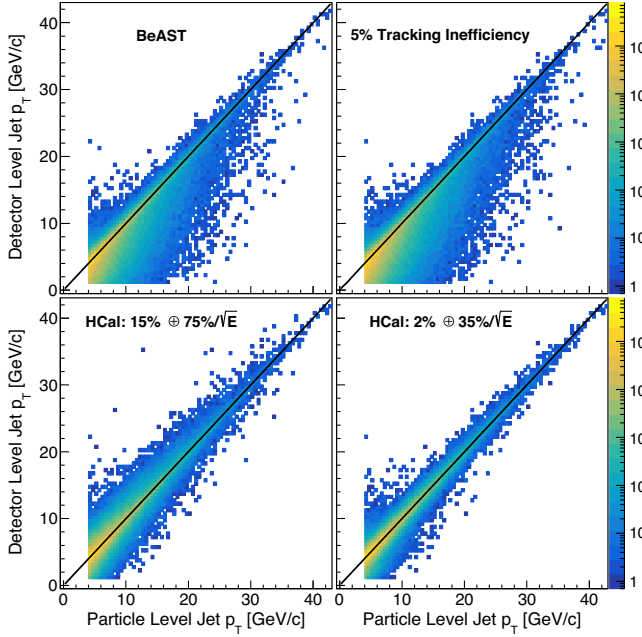


FIG. 18. Correlation between particle level and smeared jet p_T for the BeAST detector setup (upper left), BeAST assuming a 5% track finding inefficiency (upper right), and BeAST assuming a midrapidity hadron calorimeter with resolution $15\% \oplus \frac{75\%}{\sqrt{E}}$ (lower left) and resolution $2\% \oplus \frac{35\%}{\sqrt{E}}$ (lower right).

the correlation between particle level and smeared jet p_T more symmetric around the diagonal. The width of the distribution is then determined by the resolution of the hadron calorimeter.

A more detailed comparison of the relationships shown in Fig. 18 can be obtained by taking projections onto the particle level axis for narrow slices of smeared jet p_T (or vice versa). Figure 19 presents three such projections with smeared jet transverse momenta of 7, 10, and 13 GeV/c for the baseline BeAST and BeAST with two midrapidity

hadron calorimeter configurations and essentially shows how different particle level p_T values contribute to a given smeared jet p_T . It is evident that the high resolution hadron calorimeter (green) substantially improves the jet resolution, however, it is less clear that the low resolution calorimeter (red) provides much advantage over the baseline design (blue). Smeared jets found with the low resolution calorimeter option have less contribution from particle level jets with larger p_T than in the baseline design; however, these smeared jets obtain a large contribution from lower p_T particle level jets due to the large energy distortion introduced by the calorimeter. Implications of this observation will be discussed in the next section.

In addition to transverse momentum, smearing of the rapidity and azimuthal angle of the jet thrust axes were also investigated, and very good agreement between particle level and smeared jets was seen. It should be noted, however, that the position resolutions inherent to the calorimeters were not considered in this exercise as they depend on details such as material, tower size, and readout which have not been finalized. This should be revisited when more complete detector simulations are available and will be critically important for future work investigating the utility of jet shape observables.

C. Hadron calorimetry

The decision not to include a midrapidity hadron calorimeter in the BeAST design was based on several considerations including the low energies of produced particles, the modest fraction of total energy carried by neutral hadrons, the use of streaming readouts which do not require a trigger, and finally, the significant cost of such a detector. Figure 19 makes it clear that a hadron calorimeter with sufficiently high resolution can markedly improve jet energy measurements. Unfortunately, calorimeter cost increases with resolution, meaning the inclusion of such a high resolution midrapidity hadron calorimeter (which

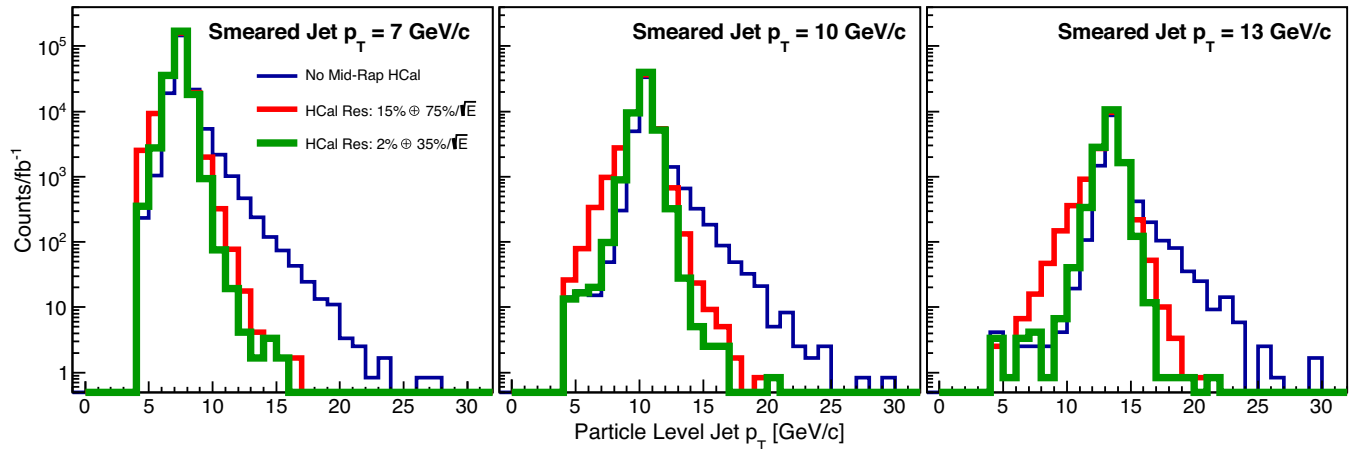


FIG. 19. Particle jet p_T spectra for smeared jet p_T values of 7 (left), 10 (middle), and 13 GeV/c (right). The baseline BeAST design (blue) and BeAST plus two midrapidity hadron calorimeter configurations (red and green) are compared.

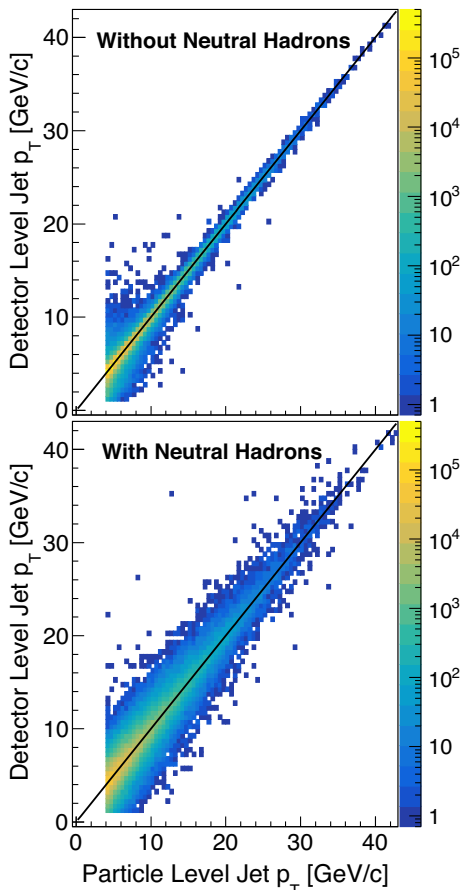


FIG. 20. Relationship between particle level and smeared jet transverse momenta for jets which do not (upper panel) and do (lower panel) contain neutral hadrons.

must cover a large volume) may be infeasible. While the corrections to jet energy needed in the absence of a hadron calorimeter will be modest, it is worth considering the benefits that could be provided by a more economical lower resolution calorimeter.

One such benefit would be the ability to implement an unbiased jet trigger. While the current plan calls for a data acquisition system capable of recording all interactions, this ability has not been demonstrated, which means that the capability of triggering on events with jets in an unbiased way could be necessary. Even if such a streaming readout is possible, a traditional trigger system including a hadron calorimeter may be more economically feasible. A hadron calorimeter would also provide *in situ* measurements of neutral hadron abundances and energies, which would reduce the uncertainty in any Monte Carlo based corrections to the jet energy. Figure 18 shows that even a low resolution hadron calorimeter would reduce the number of jets reconstructed at significantly lower transverse momenta, which would mitigate the loss of jets which would otherwise fail a minimum p_T cut. The largest benefit, though, would likely come from the ability to

differentiate between jets which do and do not contain neutral hadrons.

A hadron calorimeter should make it possible to separate jets containing neutral hadrons from those which do not by identifying energy deposits which do not have a corresponding charged particle track. The energy resolutions of the roughly 65% of jets which do not contain a neutral hadron will be dominated by the high precision tracker and electromagnetic calorimeters. The superior resolution for jets which do not contain neutral hadrons versus those which do, can be seen in Fig. 20. Separating jets in this way would allow a much smaller correction to be applied to the majority of jets while reserving the larger corrections for the 35% of jets which contain energy from neutral hadrons. This scheme should improve overall jet energy resolution much more than what would be possible considering only the energy recorded by the calorimeter.

VI. JET APPLICATION: TAGGING PHOTON-GLUON FUSION

Previous sections have focused on technical aspects of jet finding at an EIC without discussion of potential applications. As stated above, the utility of dijets at an EIC has been explored recently in the context of accessing the gluon Sivers function [20] and Weizsäcker-Williams gluon distributions [16] as well as determining polarized and unpolarized photon structure functions [21]. This section will present a related measurement in which dijets are used to tag photon-gluon fusion events for the purpose of exploring the gluon contribution to the spin of the proton, ΔG , via the longitudinal double spin asymmetry A_{LL} at leading order.

A. Kinematics and tagging

One of the signatures of the PGF process is the production of particles with large momenta transverse to the photon-proton interaction axis which are back-to-back in azimuth, meaning the observation of a dijet in the Breit frame can be used to tag possible PGF events. Unfortunately, both the resolved and QCD compton processes, which are background to a ΔG measurement, also produce such dijets. While a global analysis could likely handle these background contributions in a consistent way, it is worth exploring what can be done experimentally to isolate the PGF process.

Because the dijet kinematics approximate those of the outgoing partons, they can be used to reconstruct properties of the event which will help separate the PGF process from resolved and QCDC events. In this analysis, the two variables used for this purpose are x_γ and x_p , which are the momentum fractions carried by the parton originating from the photon and the parton coming from the proton, respectively. These quantities are reconstructed from the dijet kinematics as follows:

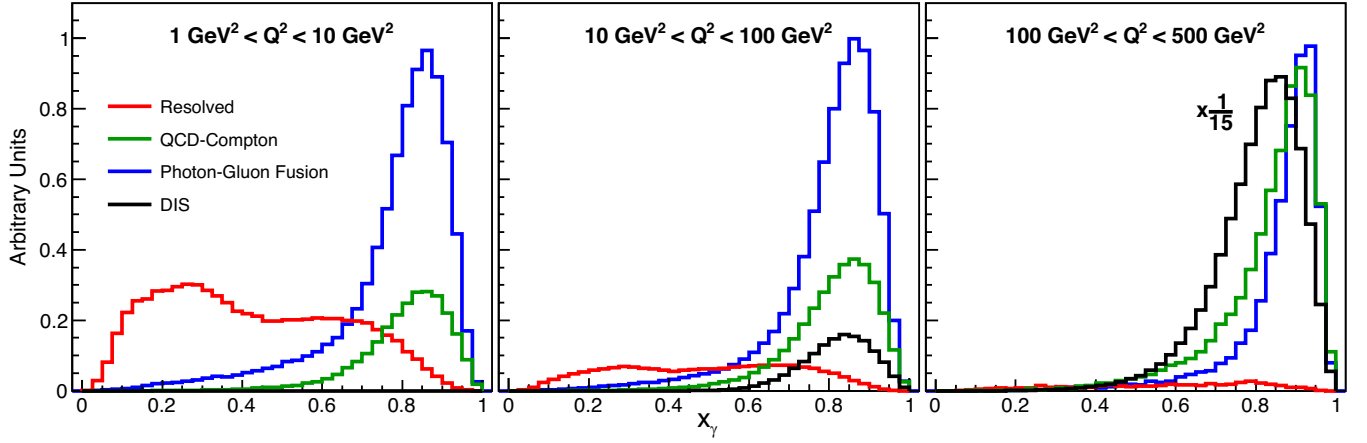


FIG. 21. Reconstructed x_γ for the resolved, QCDC, PGF, and DIS subprocesses in Q^2 bins of 1 – 10 GeV^2 (left), 10 – 100 GeV^2 (middle), and 100 – 500 GeV^2 (right). Note that each panel has been scaled separately an arbitrary amount and in the right panel, the DIS curve has been scaled down by an additional factor of 15.

$$x_\gamma = \frac{1}{2E_e y} (m_{T1} e^{-Y_1} + m_{T2} e^{-Y_2}) \quad (2)$$

$$x_p = \frac{1}{2E_p} (m_{T1} e^{Y_1} + m_{T2} e^{Y_2}), \quad (3)$$

where E_e and E_p are the energies of the incoming electron and proton beams, respectively, y is the inelasticity, m_T is the jet transverse mass defined as the quadrature sum of the jet mass and p_T , and Y is the jet rapidity, in the laboratory frame. The correlation between generated and reconstructed x_γ and x_p is quite good (see [21]). Dijets were reconstructed using the same method as described in Sec. III D.

As they are largely a low Q^2 phenomenon, a significant fraction of resolved events can be eliminated simply by requiring that $Q^2 > 1 \text{ GeV}^2$. However, because requiring two high- p_T jets significantly biases the event sample against leading order DIS and toward higher-order processes, a non-negligible resolved contribution remains for $Q^2 > 1 \text{ GeV}^2$ (see Fig. 21). This remaining resolved component can be greatly reduced by requiring that the reconstructed x_γ be close to unity. As explained in [21], the virtual photon behaves as a point particle for the PGF and QCDC processes, meaning it contributes 100% of its momentum to the interaction and thus should have $x_\gamma = 1$. Conversely, for a resolved event, the photon behaves as a composite particle and only a fraction of its momentum contributes to the interaction, meaning x_γ will have a broad distribution of values less than unity. Figure 21 presents the reconstructed x_γ distributions for dijets from the resolved, QCDC, PGF, and leading order DIS processes for three Q^2 ranges. It is clear that a cut on x_γ can effectively remove a large fraction of the resolved contribution while preserving most of the direct events. Cuts on x_γ of 0.75 and 0.60 were placed for the Q^2 ranges

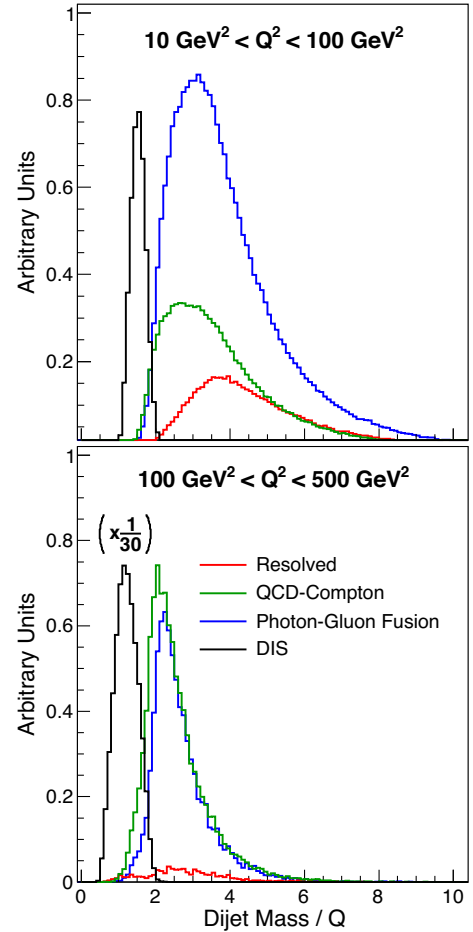


FIG. 22. Ratio of dijet invariant mass over Q for the resolved, QCDC, PGF, and DIS subprocesses in Q^2 bins of 10 – 100 GeV^2 (upper panel), and 100 – 500 GeV^2 (lower panel). This ratio allows the separation of dijet events arising from the leading order DIS subprocess from all others. Note that each panel has been scaled an arbitrary amount and in the lower panel, the DIS curve has been scaled down by an additional factor of 30.

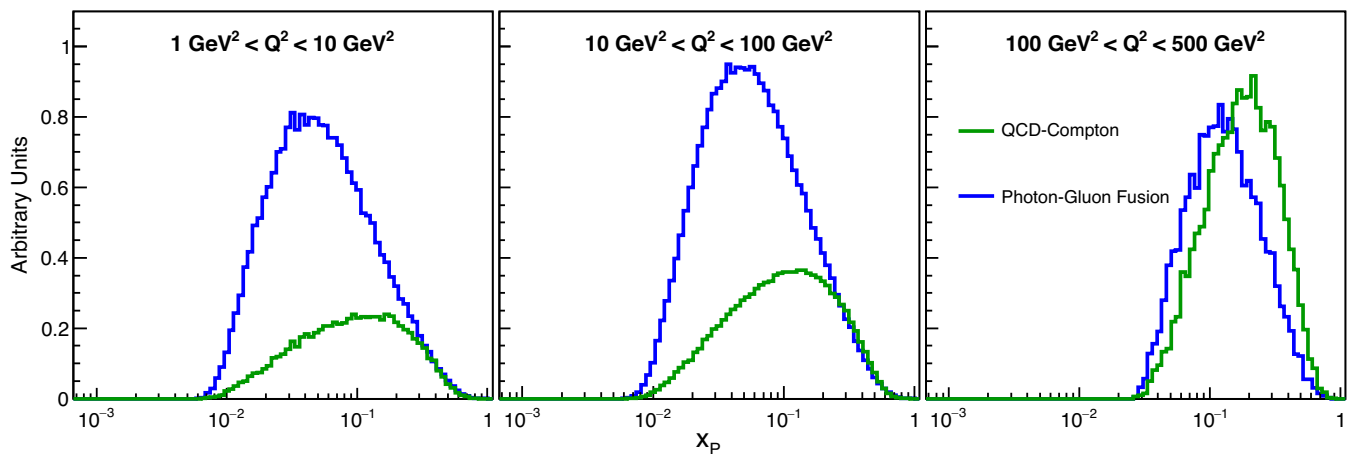


FIG. 23. [color online] Reconstructed momentum fraction of the parton arising from the proton for the QCDC and PGF subprocesses in Q^2 bins of $1 - 10 \text{ GeV}^2$ (left), $10 - 100 \text{ GeV}^2$ (middle), and $100 - 500 \text{ GeV}^2$ (right).

$1 - 10 \text{ GeV}^2$ and $10 - 100 \text{ GeV}^2$, respectively, which reduce the resolved component to less than 10% of the remaining PGF contribution. No cut is placed on the $100 - 500 \text{ GeV}^2$ bin. It should be noted that the dijets from leading order DIS which appear at larger Q^2 values arise when the target remnant receives a large enough transverse kick to form a jet which passes the selection criteria. Such events can be eliminated, with minimal loss to PGF and QCDC yields, by cutting on the ratio of dijet mass to Q (the ratio was required to be greater than 2.0) as shown in Fig. 22. For the following, residual contributions from resolved and leading order DIS events were omitted for simplicity as neither subprocess was found to contribute significantly to the expected asymmetry.

Removing the QCDC contribution is not as straightforward as eliminating resolved events because the PGF and QCDC processes have very similar event topologies. However, as can be seen in Fig. 23, the PGF cross section peaks at lower values of x_p relative to QCDC events. Thus, at least at Q^2 below 100 GeV^2 where the QCDC cross section is small compared to PGF, x_p can be used to select regions of high or low signal-to-background. It should be noted that placing an upper x_p cut will restrict the maximum accessible dijet mass as seen in Fig. 24. The relationship between x_p and dijet mass (at leading order) is given by the Eq. (4),

$$x_p = x_B + \frac{M_{jj}^2}{s y}, \quad (4)$$

where x_B is Bjorken- x , M_{jj} is the invariant mass of the dijet system, s is the center-of-mass energy, and y is the inelasticity. Equation (4) shows that the x_p values accessible to this measurement are driven by the center-of-mass energy and that for a minimum dijet mass of 10 GeV^2 , s of 20000 GeV^2 , and a maximum inelasticity of 0.95, the

lowest x_p available is roughly 5×10^{-3} . Figure 24 presents accessible x_p values as a function of dijet invariant mass as well as curves delineating the available phase-space for center-of-mass energies of 141 and 45 GeV.

B. Expected asymmetry

The method used here to determine the behavior of A_{LL} is the same as in [21], which was adapted from [52]. For each simulated event, a weight was calculated using the subprocess and kinematic information from PYTHIA as well as external (un)polarized PDFs. The asymmetry is then

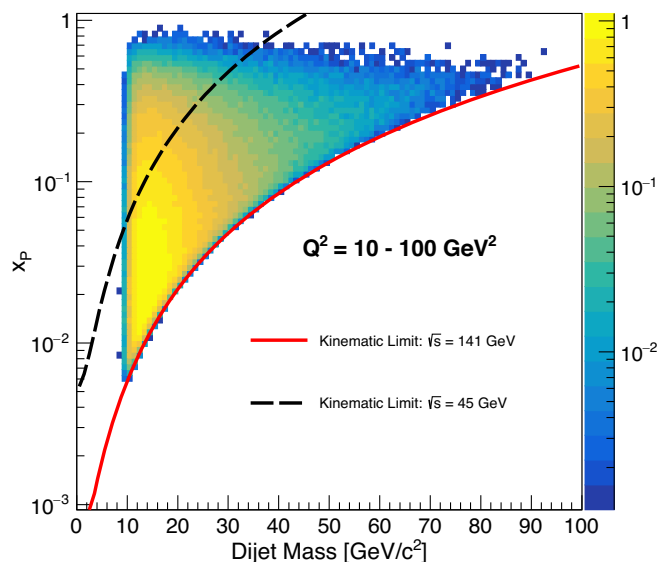


FIG. 24. Reconstructed momentum fraction of the parton arising from the proton vs the invariant mass of the resulting dijet for the QCDC and PGF processes combined and Q^2 between 10 and 100 GeV^2 . The solid red and dashed black lines denote the allowed phasespace for $\sqrt{s} = 141$ and 45 GeV , respectively.

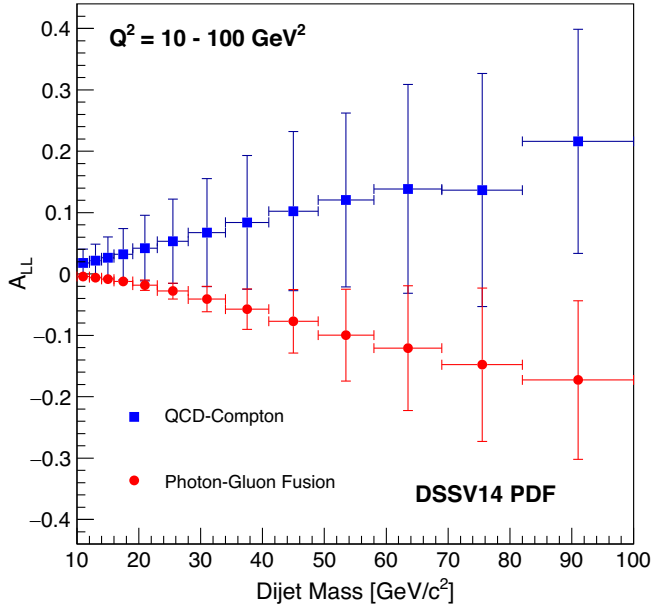


FIG. 25. DiJet A_{LL} as a function of dijet invariant mass for the QCDC and PGF subprocesses in the 10 – 100 GeV^2 Q^2 bin.

found as the average over these weights. The weights are calculated according to

$$w = \hat{a}(\hat{s}, \hat{t}, \mu^2, Q^2) \frac{\Delta f_a^{\gamma^*}(x_a, \mu^2)}{f_a^{\gamma^*}(x_a, \mu^2)} \frac{\Delta f_b^N(x_b, \mu^2)}{f_b^N(x_b, \mu^2)}, \quad (5)$$

where $\hat{a}(\hat{s}, \hat{t}, \mu^2, Q^2)$ is the subprocess dependent parton level asymmetry, $(\Delta)f_a^{\gamma^*}(x_a, \mu^2)$ is the (polarized) PDF for the virtual photon, and $(\Delta)f_b^N(x_b, \mu^2)$ is the (polarized) PDF for the proton. The leading order expressions for \hat{a} were taken from [52] and include the appropriate depolarization factors. The DSSV14 [53] and NNPDFpol1.1 [3] sets were used to describe the polarized proton and were normalized by the MSTW2008 [54] and NNPDF2.3 [55] unpolarized PDFs, respectively. Because only direct events were considered, the $\Delta f_b^N(x_b, \mu^2)/f_b^N(x_b, \mu^2)$ term is identically unity.

Figure 25 shows A_{LL} as a function of dijet invariant mass for the QCDC and PGF subprocesses obtained using the DSSV set (NNPDF is similar) for Q^2 between 10 and 100 GeV^2 . The error bars represent the rms of the distribution of weights in each mass bin. The width of the weight distribution is larger for the QCDC process because the sign of the weight can change due to different asymmetry signs for up and down quarks. For the PGF process on the other hand, the gluon asymmetry is positive everywhere in the relevant kinematics and the \hat{a} term is always negative, meaning the weight is always negative and therefore the spread in weights is smaller. It is seen that the asymmetries for each subprocess grow with dijet mass and become sizable, reaching values of 20% for the highest

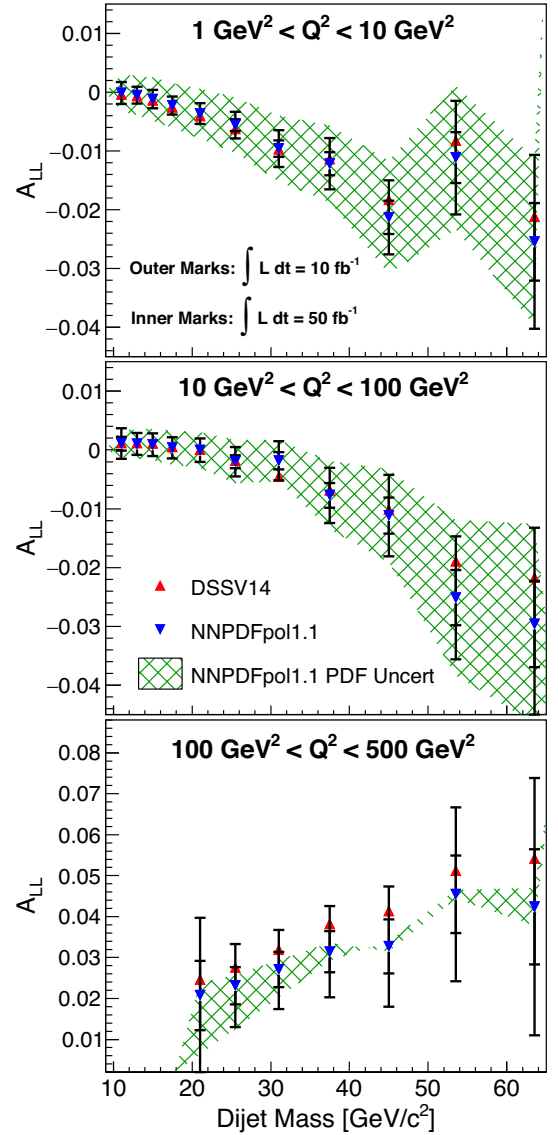


FIG. 26. DiJet A_{LL} as a function of dijet invariant mass for the combined QCDC and PGF subprocesses using the DSSV14 and NNPDF1.1 polarized PDFs in the 1 – 10 GeV^2 (top), 10 – 100 GeV^2 (middle), and 100 – 500 GeV^2 (bottom) Q^2 bins. Note that projected statistical uncertainties for the DSSV14 points are not shown for clarity, but are nearly identical to those from NNPDF1.1.

masses. However, because the QCDC and PGF asymmetries are roughly equal in magnitude but opposite in sign, one can expect that the total asymmetry will be significantly smaller than the asymmetry for either individual subprocess.

The combined QCDC and PGF A_{LL} obtained using both the DSSV and NNPDF PDFs can be seen in Fig 26 as a function of dijet invariant mass in three Q^2 ranges. The error bars represent expected statistical uncertainties calculated according to

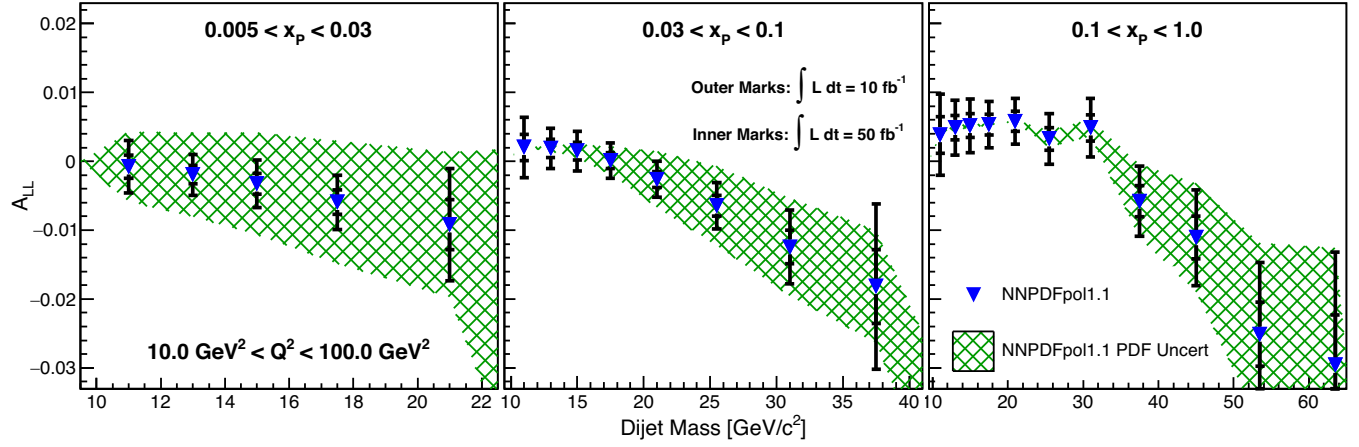


FIG. 27. Dijet A_{LL} as a function of dijet invariant mass for the combined QCD and PGF subprocesses using the NNPDF1.1 polarized PDFs in the 10 – 100 GeV² Q^2 bin for partonic momentum fractions of $0.005 < x_p < 0.03$ (left), $0.03 < x_p < 0.1$ (middle), and $0.1 < x_p < 1.0$ (right).

$$\sigma = \frac{1}{P_e P_p} \sqrt{\frac{1}{N} - \frac{A_{LL}^2}{N}}, \quad (6)$$

where P_e and P_p are the electron and proton beam polarizations (taken as 80% and 70%, respectively) and N is the number of expected events assuming an integrated luminosity of 10 fb⁻¹ or 50 fb⁻¹. The green bands represent the uncertainty on the NNPDFpol1.1 polarized PDF. As expected, the QCD and PGF asymmetries cancel to a large degree, resulting in maximum asymmetries of a few percent.

In order to better isolate the gluon contribution, it would be helpful to reduce the fraction of QCD events, which carry information on the quarks. As mentioned above, the reconstructed momentum fraction carried by the parton from the proton can be used to select kinematic regions where the PGF subprocess is dominant. Figure 27 presents dijet A_{LL} as a function of invariant mass for the bin $Q^2 = 10\text{--}100$ GeV² for three x_p slices: $0.005 < x_p < 0.03$, $0.03 < x_p < 0.1$, and $0.1 < x_p < 1.0$, with the ratio of PGF to QCD events decreasing with increasing x_p . Note that only NNPDF1.1 results are shown for clarity. The bars again show expected statistical uncertainties assuming 10 and 50 fb⁻¹ and the green bands are the uncertainty on the NNPDFpol1.1 polarized PDFs. The effects of slicing in x_p are modest, but do shift the asymmetries to more negative values and increase the ratio of PDF to statistical uncertainties.

The PDF uncertainties presented in Figs. 26 and 27 represent the current state of knowledge on the helicity structure of the proton. These uncertainties will shrink substantially with the addition of inclusive g_1 measurements, which will be the golden channel for the constraint of $\Delta g(x, Q^2)$. Figures 26 and 27 show that substantial integrated luminosities will be needed in order for the dijet

measurements to improve on our current knowledge of $\Delta g(x, Q^2)$ meaning it will be unlikely the dijet measurement can compete directly with g_1 in constraining the gluon contribution to the proton spin. The benefit of the dijet measurement will likely be in its complementarity to g_1 as the dijets arise from different subprocesses and will have different associated systematics than inclusive observables.

VII. SUMMARY AND OUTLOOK

Jet observables have proven their utility as probes of the subatomic realm at virtually all high energy colliders operated to date, while recent experimental and theoretical advances, spurred by the success of modern colliders such as the LHC, have seen jets become true precision probes. This success behooves those interested in the science an EIC will address to explore the potential benefits that jets could provide. To that end, this paper has systematically explored a number of topics relevant to the experimental analysis of jets at an EIC.

The first issues addressed were particulars of the actual jet finding. There was no significant dependence seen on the choice of jet algorithm, but jets with larger radii were found to better reproduce the underlying partonic kinematics, and the anti- k_T algorithm with $R = 1.0$ was chosen for all subsequent studies. Next, jet kinematic distributions were quantified, comparing inclusive jet p_T , dijet mass, and pseudorapidity spectra for both inclusive jets and dijets for a range of Q^2 values, subprocesses, and center-of-mass energies. It was seen that higher center-of-mass energies produce greater yields of jets/dijets, especially at larger p_T /mass. The pseudorapidity of jets was also seen to increase with \sqrt{s} , driven by the larger boost imparted by higher hadron beam energy. The energy contribution from underlying event activity was also studied using two different methods and was found to be small, although it will need to be considered when dealing with low p_T jets,

where even small underlying event contributions can have a fractionally larger effect. Distortions of the jet p_T due to realistic detector resolutions were investigated using a smearing program tuned to replicate the BeAST detector design and were found to be minor. Special attention was given to the role of hadron calorimetry at midrapidity with high resolution, low resolution and no calorimeter options explored. A scheme to use a low resolution calorimeter as a neutral hadron veto system with the goal of improving the overall jet energy resolution was also discussed. Finally, an example analysis was presented in which dijets were used to tag photon-gluon fusion events for the purpose of constraining the gluon helicity contribution to the proton spin. Methods for reducing background and isolating PGF events were demonstrated and the expected asymmetries and their uncertainties were found and compared to current knowledge of gluon polarizations.

While this paper provides a solid introduction to the experimental reality of jet physics at an EIC, the topic is still relatively new and more detailed follow-up studies will be needed to build a robust EIC jet program. Areas of future

study include potentially fruitful topics such as jet substructure and the use of jets in $e + A$ collisions, which were not addressed here at all. In addition, more realistic detector simulation and modeling will be needed in order to inform detector performance requirements. We hope this paper will serve as a valuable resource and jumping off point for both theorists and experimentalists who wish to further pursue jet topics at the EIC.

ACKNOWLEDGMENTS

We would like to thank Felix Ringer, Kyle Lee, Kolja Kauder, Miguel Arratia, Barbara Jacak, and Frank Petriello for helpful discussions. We are grateful to Alexander Kiselev for providing details on proposed tracking resolutions for the BeAST detector design. B. P. is supported by the Program Development program at Brookhaven National Laboratory, while E. C. A and X. C. acknowledge support from the U.S. Department of Energy under Contract No. de-sc0012704.

-
- [1] A. Ali and G. Kramer, *Eur. Phys. J. H* **36**, 245 (2011).
 - [2] D. De Florian, G. A. Lucero, R. Sassot, M. Stratmann, and W. Vogelsang, *Phys. Rev. D* **100**, 114027 (2019).
 - [3] E. R. Nocera, R. D. Ball, S. Forte, G. Ridolfi, and J. Rojo (NNPDF Collaboration), *Nucl. Phys.* **B887**, 276 (2014).
 - [4] T.-J. Hou *et al.*, arXiv:1912.10053.
 - [5] A. Bhatti and D. Lincoln, *Annu. Rev. Nucl. Part. Sci.* **60**, 267 (2010).
 - [6] S. Marzani, G. Soyez, and M. Spannowsky, *Lect. Notes Phys.* **958** (2019).
 - [7] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), *Eur. Phys. J. C* **75**, 580 (2015).
 - [8] C. Adloff *et al.* (H1 Collaboration), *Eur. Phys. J. C* **1**, 97 (1998).
 - [9] S. Aid *et al.* (H1 Collaboration), *Phys. Lett. B* **358**, 412 (1995).
 - [10] M. Derrick *et al.* (ZEUS Collaboration), *Phys. Lett. B* **348**, 665 (1995).
 - [11] M. Derrick *et al.* (ZEUS Collaboration), *Phys. Lett. B* **384**, 401 (1996).
 - [12] S. Sadhuon (ALICE Collaboration), *MDPI Proc.* **10**, 43 (2019).
 - [13] M. Connors, C. Nattrass, R. Reed, and S. Salur, *Rev. Mod. Phys.* **90**, 025005 (2018).
 - [14] Y. Hatta, B.-W. Xiao, and F. Yuan, *Phys. Rev. Lett.* **116**, 202301 (2016).
 - [15] Y. Hagiwara, Y. Hatta, and T. Ueda, *Phys. Rev. D* **94**, 094036 (2016).
 - [16] A. Dumitru, V. Skokov, and T. Ullrich, *Phys. Rev. C* **99**, 015204 (2019).
 - [17] J. Zhou, *Phys. Rev. D* **94**, 114017 (2016).
 - [18] D. Boer, S. J. Brodsky, P. J. Mulders, and C. Pisano, *Phys. Rev. Lett.* **106**, 132001 (2011).
 - [19] C. Pisano, D. Boer, S. J. Brodsky, M. G. A. Buffing, and P. J. Mulders, *J. High Energy Phys.* **10** (2013) 024.
 - [20] L. Zheng, E. C. Aschenauer, J. H. Lee, B.-W. Xiao, and Z.-B. Yin, *Phys. Rev. D* **98**, 034011 (2018).
 - [21] X. Chu, E.-C. Aschenauer, J.-H. Lee, and L. Zheng, *Phys. Rev. D* **96**, 074035 (2017).
 - [22] R. Boughezal, F. Petriello, and H. Xing, *Phys. Rev. D* **98**, 054031 (2018).
 - [23] X. Liu, F. Ringer, W. Vogelsang, and F. Yuan, *Phys. Rev. Lett.* **122**, 192003 (2019).
 - [24] M. Arratia, Y. Song, F. Ringer, and B. Jacak, arXiv:1912.05931.
 - [25] R. Kogler *et al.*, *Rev. Mod. Phys.* **91**, 045003 (2019).
 - [26] A. J. Larkoski, I. Moulton, and B. Nachman, arXiv:1709.04464.
 - [27] E.-C. Aschenauer, K. Lee, B. S. Page, and F. Ringer, *Phys. Rev. D* **101**, 054028 (2020).
 - [28] S. Cao, T. Luo, G.-Y. Qin, and X.-N. Wang, *Phys. Lett. B* **777**, 255 (2018).
 - [29] T. Sjostrand, S. Mrenna, and P. Z. Skands, *J. High Energy Phys.* **05** (2006) 026.
 - [30] D. Stump, J. Huston, J. Pumplin, W.-K. Tung, H. L. Lai, S. Kuhlmann, and J. F. Owens, *J. High Energy Phys.* **10** (2003) 046.
 - [31] H. L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. I. Olness, J. F. Owens, J. Pumplin, and W. K. Tung (CTEQ Collaboration), *Eur. Phys. J. C* **12**, 375 (2000).
 - [32] G. A. Schuler and T. Sjostrand, *Z. Phys. C* **68**, 607 (1995).

- [33] G. A. Schuler and T. Sjostrand, *Phys. Lett. B* **376**, 193 (1996).
- [34] A. Accardi *et al.*, *Eur. Phys. J. A* **52**, 268 (2016).
- [35] E. C. Aschenauer *et al.*, [arXiv:1409.1633](https://arxiv.org/abs/1409.1633).
- [36] E. C. Aschenauer, T. Burton, T. Martini, H. Spiesberger, and M. Stratmann, *Phys. Rev. D* **88**, 114025 (2013).
- [37] L. Zheng, E. C. Aschenauer, J. H. Lee, and B.-W. Xiao, *Phys. Rev. D* **89**, 074037 (2014).
- [38] J. A. Rinehimer and G. A. Miller, *Phys. Rev. C* **80**, 015201 (2009).
- [39] G. P. Salam, *Eur. Phys. J. C* **67**, 637 (2010).
- [40] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Tonnesmann, *Prog. Part. Nucl. Phys.* **60**, 484 (2008).
- [41] M. Cacciari, G. P. Salam, and G. Soyez, *J. High Energy Phys.* **04** (2008) 063.
- [42] S. D. Ellis and D. E. Soper, *Phys. Rev. D* **48**, 3160 (1993).
- [43] G. P. Salam and G. Soyez, *J. High Energy Phys.* **05** (2007) 086.
- [44] G. C. Blazey *et al.*, QCD and weak boson physics in Run II. in *Proceedings, Batavia, USA, 1999* (2000), pp. 47–77 [[arXiv:hep-ex/0005012](https://arxiv.org/abs/hep-ex/0005012)].
- [45] M. Cacciari, G. P. Salam, and G. Soyez, *Eur. Phys. J. C* **72**, 1896 (2012).
- [46] J. Adam *et al.* (STAR Collaboration), *Phys. Rev. D* **100**, 052005 (2019).
- [47] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Z. Phys. C* **72**, 637 (1996).
- [48] T. Affolder *et al.* (CDF Collaboration), *Phys. Rev. D* **65**, 092002 (2002).
- [49] B. B. Abelev *et al.* (ALICE Collaboration), *Phys. Rev. D* **91**, 112012 (2015).
- [50] K. H. Ackermann *et al.* (STAR Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **499**, 624 (2003).
- [51] J. Adam *et al.* (STAR Collaboration), *Phys. Rev. D* **101**, 052004 (2020).
- [52] A. Airapetian *et al.* (HERMES Collaboration), *J. High Energy Phys.* **08** (2010) 130.
- [53] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, *Phys. Rev. Lett.* **113**, 012001 (2014).
- [54] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Eur. Phys. J. C* **63**, 189 (2009).
- [55] R. D. Ball *et al.*, *Nucl. Phys.* **B867**, 244 (2013).