Charm Jets at EIC with Examples from ATHENA

Presented at the Second Workshop on Jets for 3D Imaging at the EIC

(With thanks and gratitude to the EIC User Community, especially the ATHENA proto-collaboration and its

Jets, Heavy Flavor, Electroweak, and Beyond-the-Standard Model Physics Working Group)

Stephen Sekula[1]

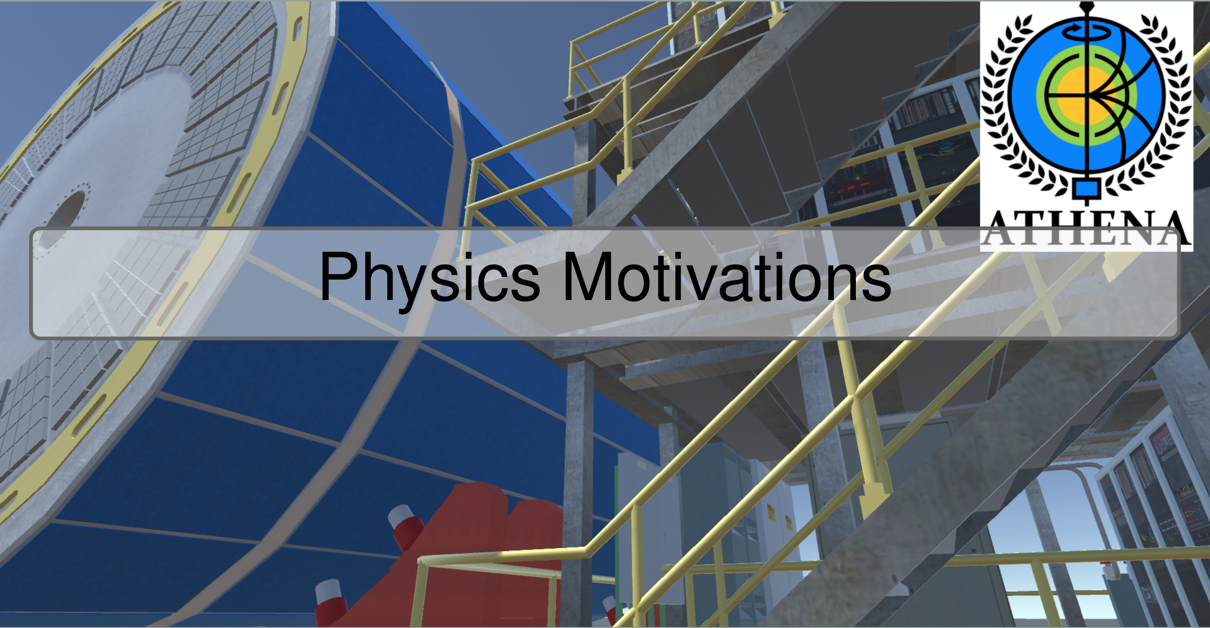
[1] Southern Methodist University, Dallas, TX;

September 28, 2021

SMU | DEDMAN COLLEGE OF HUMANITIES & SCIENCES

ATHENA

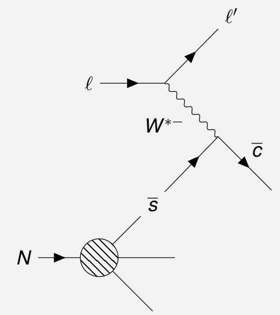
Outline



Physics Motivations

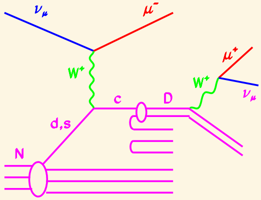
Physics Process of Interest - Charged-Current Scattering

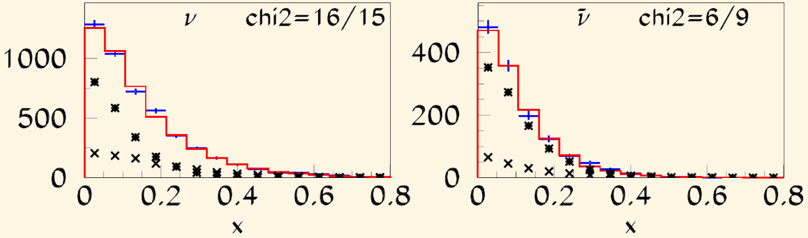
Feynman Diagram for CC DIS

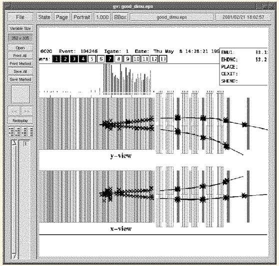


A greatly improved understanding of the quark sea in nucleons, especially at high x, is a key science driver for the EIC program. Charged-Current (CC) Deep Inelastic Scattering (DIS) offers one tool for probing the sea. An obvious target here is the strange quark, whose contributions (e.g. $s(x, Q^2)$ and helicity) are still poorly constrained (especially at high x).

EXAMPLE: Neutrino scattering on a fixed target (NuTeV) [1]







The dimuon signal is very experimentally clean, and thanks to Cabibbo suppression the inference that the preceding charm production is dominated by $s \to c$ transitions is safe; the interpretation of the underlying strange sea is challenging due to nuclear corrections and fragmentation/hadronization functions that are convoluted with the underlying PDFs.

Experimental data from this and many other similar experiments have been a key part of PDF fits in recent decades.

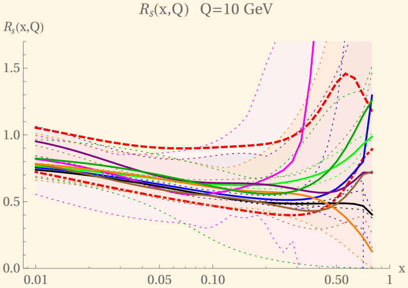
Constraints on our understanding of $s(x, Q^2)$

Some Definitions

$$r_s = \frac{s + \overline{s}}{2\overline{d}}$$

$$R_s = \frac{s + \overline{s}}{\overline{u} + \overline{d}}$$

The above are functions of x and Q^2

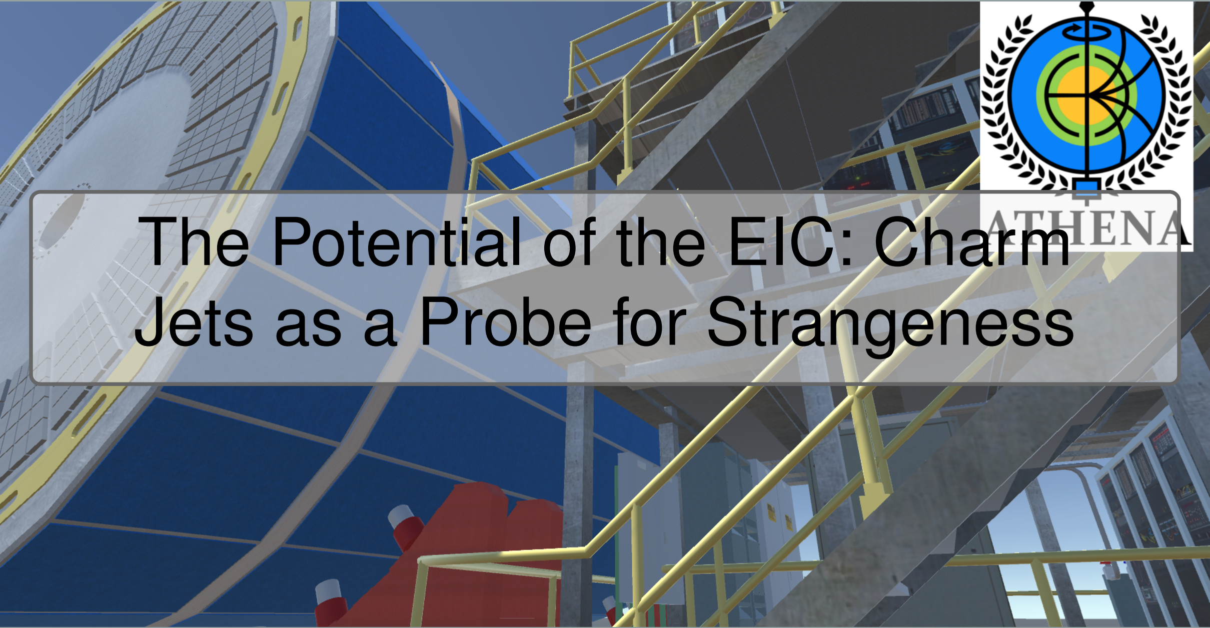


Example: PDF sets and uncertainty on $R_s(x, Q)$ vs. x [2]

PDF Set	$\kappa(Q) = \int R_s(x, Q)$	color
Rs-Low	0.37	red
CJ15nlo	0.43 ±0.01	black
CT18NNLO	0.44 $^{+0.15}_{-0.11}$	brown
MSTW2008nnlo68cl	0.48 $^{+0.02}_{-0.03}$	blue
EPPS16_CT14nlo_Pb208	0.49 $^{+0.21}_{-0.19}$	orange
nCTEQ15FullNuc_208_82	0.50 ±0.01	light green
HERAPDF20_NLO_VAR	0.57 $^{+0.19}_{-0.42}$	green
NNPDF31_nnlo_as_0118	0.59 $^{+0.30}_{-0.30}$	magenta
CT18A NNLO	0.63 $^{+0.23}_{-0.16}$	purple
Rs-High	0.96	red

Legend for R_s plot.

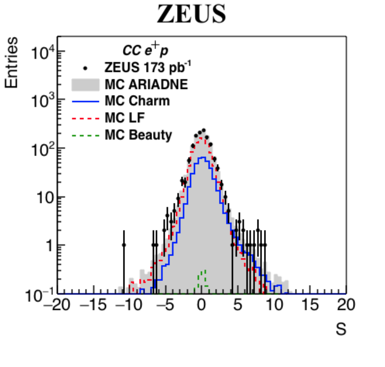
Dimuon measurements in ν DIS typically prefer a low value of R_s; kaon semi-inclusive DIS typically prefers an even lower value; LHC electroweak boson measurements prefer a value consistent with unity $\rightarrow$ **while theoretical improvements are definitely being made, we as a community benefit immensely from new data to fuel additional progress.**



The Potential of the EIC: Charm Jets as a Probe for Strangeness

Charm Jets: The ZEUS Experimental Approach [3]

- HERA-II $e^{\pm}p$ collision data, $\sqrt{s} = 318$ GeV, $\mathcal{L}_{e^+} = 173\text{pb}^{-1}$, and $\mathcal{L}_{e^-} = 185\text{pb}^{-1}$
- Tracking capabilities included a microvertex detector in addition to the central tracking detector
- Calorimetry based on a uranium-scintillator design with EM and HAD capabilities



Reconstruct jets (k_T, R=1) from energy flow objects and identify charm jets using displaced secondary vertex (SVX) counting. Fit to flight information of SVXs.

$$\sigma^+_{cEW} = 8.5 \pm 5.5(\text{stat.}) \pm ^{+0.2}_{-1.3}(\text{syst.})\text{ pb}$$
$$\sigma^-_{cEW} = -5.7 \pm 7.2(\text{stat.}) \pm ^{+1.0}_{-1.2}(\text{syst.})\text{ pb}$$

Statistical uncertainties dominate → expect that to be vastly reduced by the high-luminosity program at the EIC.

Dominant systematic uncertainties were from secondary vertex corrections and predictions of the QCD charm contribution; sub-dominant systematics arose from light-flavor background estimation, selection efficiency, and jet energy scale.

Collider Configuration and Physics Goals (c.f. [7])

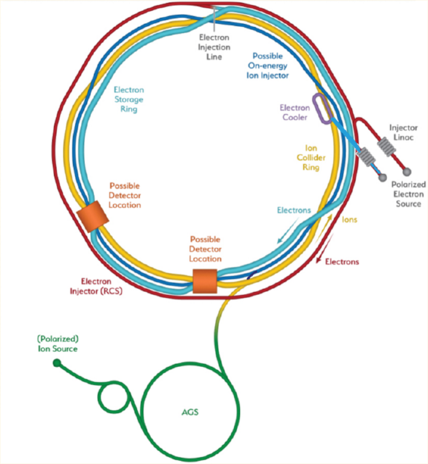
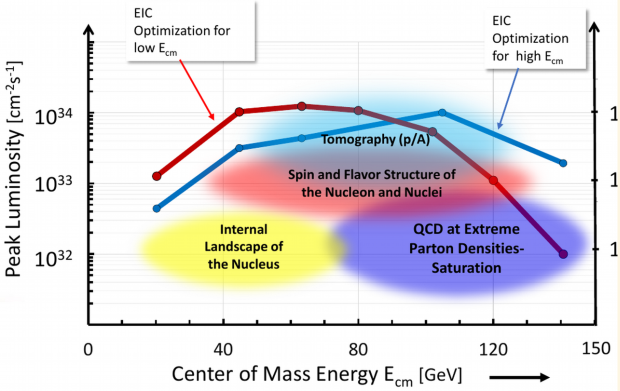
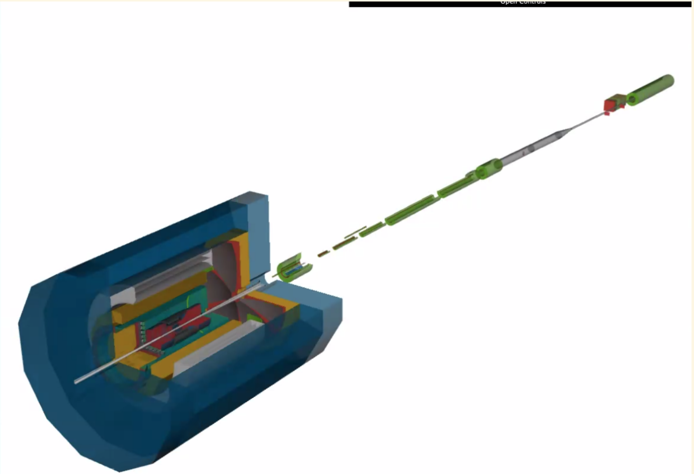


Image from Ref. [4]

The electron-hadron beam model employed for the studies shown here is an e^- and p^+ collision with $E_e = 10\,\mathrm{GeV}$ and $E_p = 275\,\mathrm{GeV}$. This achieves $\sqrt{s} = \sqrt{4E_eE_p} \approx 105\,\mathrm{GeV}$, high instantaneous luminosity, and can yield $\approx 100\,\mathrm{fb}^{-1}$/year [5, 6].



ATHENA-Like Implementation in Fast Simulation



Utilizing Delphes detector implementation [8] based on ATHENA concept, here shown using GEANT4.

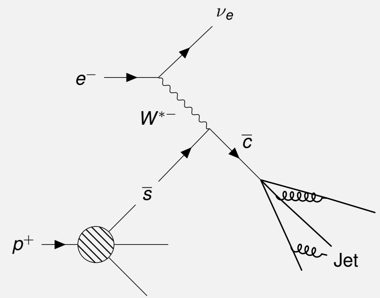
- Barrel: 3T magnet, All-Silicon Tracker (no MPGD) + HP-DIRC, EMCAL + HCAL (Fe/Sc)
- Hadron-going direction: Si-Disks + GEM Layer + dRICH + EMCAL(Tungsten Powder/ScFi) + HCAL (Fe/Sc) + B0(MAPS) + Off-Momentum + Roman Pots(AC-LGAD) + ZDC
- Electron-going direction: Si-Disks + GEM Layer + mRICH + iEMCAL($PbWO_4$) + oEMCAL + HCAL (Fe/Sc) + Low-Q^2 Tagger

Delphes implements 3T field, track resolution/smearing/coverage, as well as an EMCAL and HCAL, based on the above concepts. The mRICH, barrel DIRC, and mRICH performance is taken from EIC User Group PID performance estimates for the Yellow Report [7] to implement PID maps (e.g. $\pi \to K$). Where possible, full G4-based simulation has been used to inform performance maps/smearing/efficiency in Delphes (tracking, calorimetry)

Modeling CC DIS

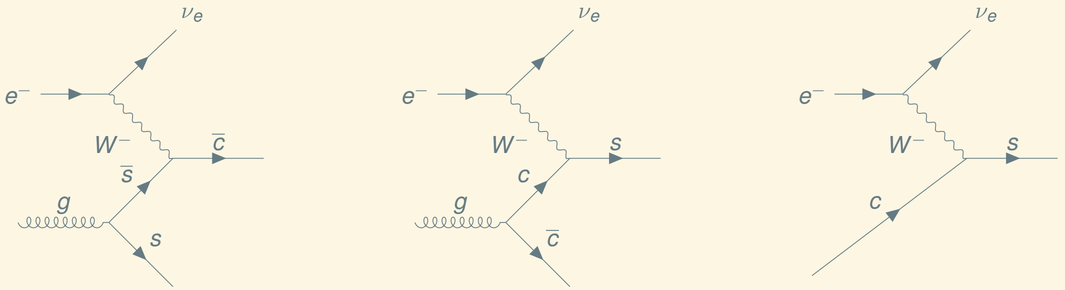
- Charged-Current (CC) DIS is simulated using PYTHIA8 [9, 10]
  - Process: `WeakBosonExchange:ff2ff(t:W)`
  - $Q^2 > 100\ \mathrm{GeV}^2$
  - $\sigma^{PYTHIA8}_{Q^2>100\ \mathrm{GeV}^2} = 14.76\,\mathrm{pb}$
- PDF set: CT18 from LHAPDF; baseline is CT18NNLO
- Generated 20M events for study (using varying amounts of this for the parts of this talk)
- EIC beamspot model employed in PYTHIA8, including crossing angle of 25 mrad.[11]

Feynman Diagram for CC DIS



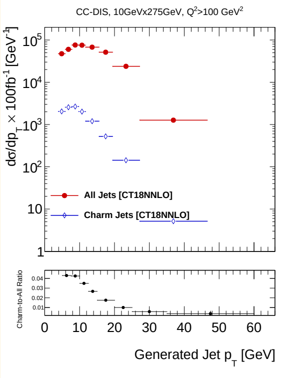
Broader View of Charm Jet Production at the EIC

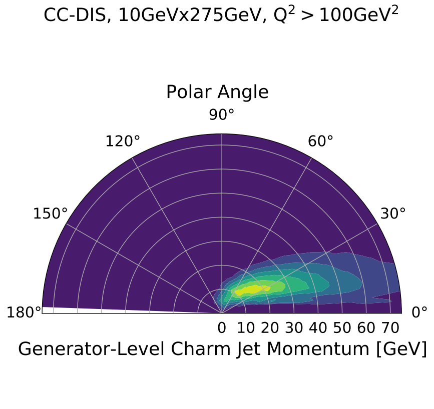
The CC DIS process $W^- \overline{s} \to \overline{c}$ is, of course, not the only contributor to charm jet production at an electron-hadron machine [3].



LEFT DIAGRAM: modelled in PYTHIA8 for this study. Reported cross-section is $\sigma^{PYTHIA8}_{Q^2>100\ \mathrm{GeV}^2} = 14.76\,\mathrm{pb}$.
MIDDLE AND RIGHT DIAGRAMS: Gluon splitting and charm-initiated graphs are also important contributors, being sensitive to $g(x, Q^2)$ and $F_2^{c\overline{c}}$ (including a possible nonperturbative charm component); these were not modeled in this study. Also (not shown) need to consider final-state gluon radiation, $g \to c\overline{c}$. All of these play essential roles in EIC physics (e.g. Sivers function using charm jets).

Differential Cross Sections and Yields for 100fb^{-1}

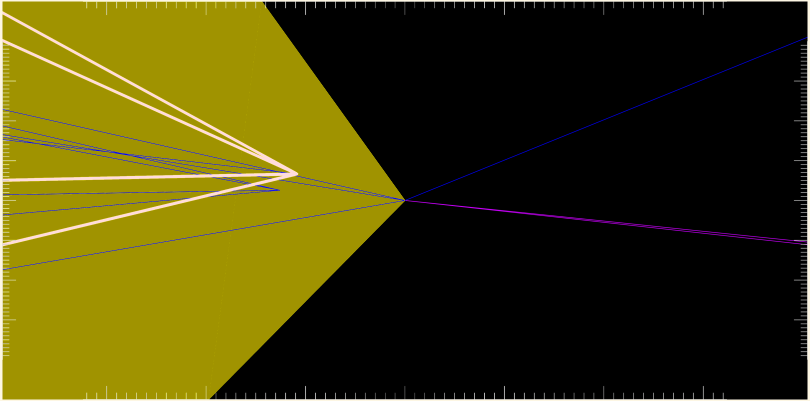


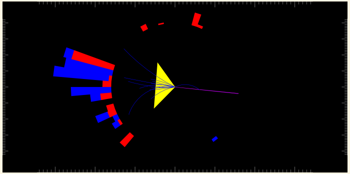


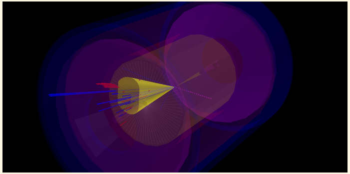
Theoretical expectation is for thousands of charm jets produced in 100fb^{-1} from CC DIS (a few percent of the total CC DIS jet yield).

Charm jets (anti-k_T, R=1) preferentially produced at low-angle ($\theta \approx 35^\circ$ or $\eta \approx 1.2$, approximately the transition region between barrel and hadronic endcap) but span out to $\eta \approx 3.5 - 4.0$. Need good tracking and calorimetry coverage (and hermeticity!) down to low forward angles to insure reliable reconstruction of these jets.

Example: Event Display of Charm Jet, 8 high-IP tracks, $E_T^{miss} = 24.7\ \mathrm{GeV}$, $p_T = 24.5\ \mathrm{GeV}$, $\eta = 1.6$

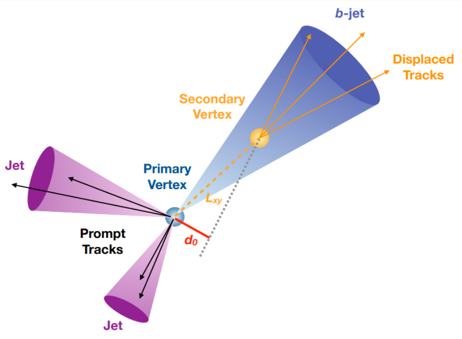






The displaced vertices in this particle flow jet (there are 2; 1 is highlighted) are about 0.5mm from the IP in $x - y$ plane. The jet was truth-matched to a charm jet and charm-tagged. Two true kaons are present, one emerging from each displaced vertex. Good case for jet-level tagging but also inclusion of track-level particle identification (PID) in an inclusive tagging algorithm.

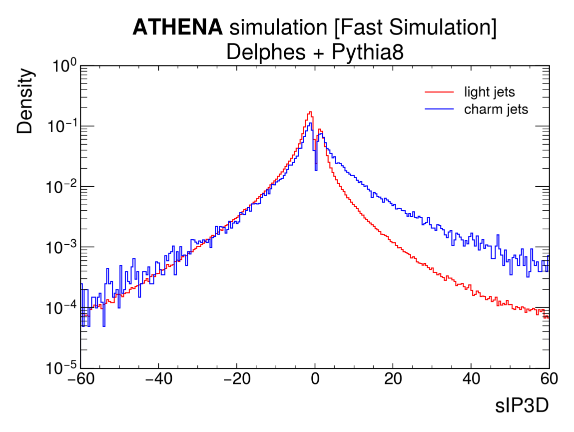
Flavor-Tagging Jets: High-Impact Parameter Track Counting [Work in Progress!]



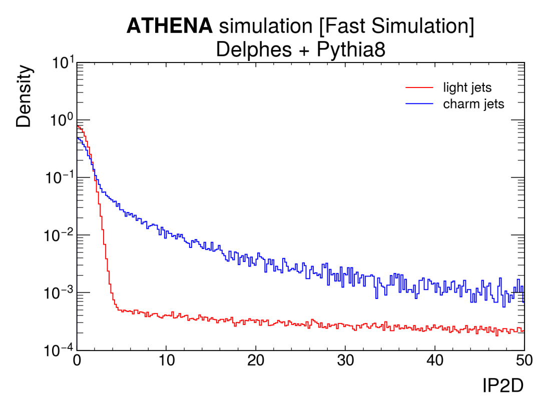
Graphic from Ref. [12]

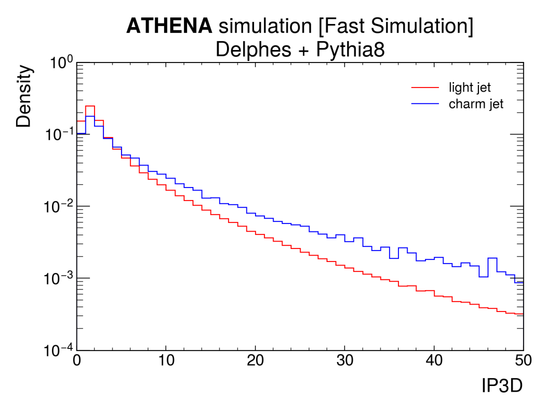
A basic flavor tagging algorithm: high-impact-parameter track counting inside a jet object, e.g. *signed 3-D impact-parameter* defined as:

$$\mathrm{sIP_{3D}} = \mathrm{sgn}(\vec{p}_j \cdot \vec{L}) \times \sqrt{(d_0/\sigma_{d_0})^2 + (z_0/\sigma_{z_0})^2} \quad (1)$$



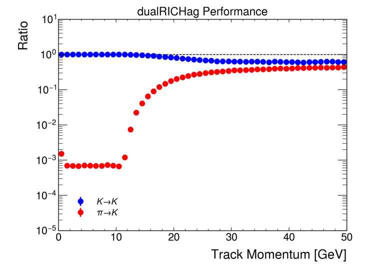
Flavor-Tagging Jets: High-Impact Parameter Track Counting [Work in Progress!]

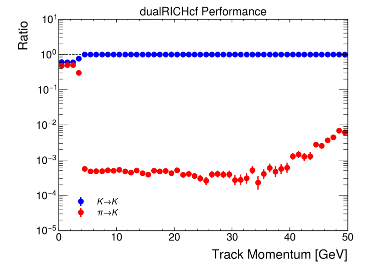


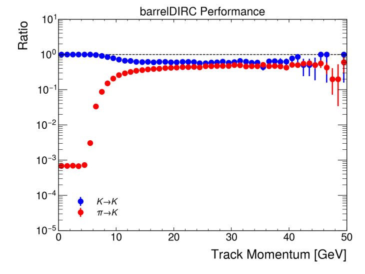


We optimize hyper-parameters by minimizing the uncertainty on the final light-jet-subtracted charm jet yield: track $p_T \geq 0.5$ GeV and $\mathrm{IP}_{2D} > 3$; ≥ 3 such tracks. This yields 17% (0.9%) efficiency on charm (light) jets.

Flavor-Tagging Jets: Displaced Kaon Counting [Work in Progress!]

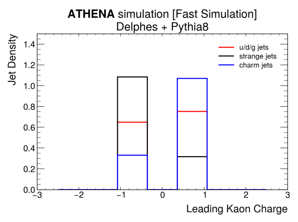


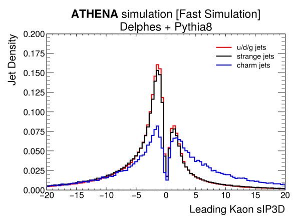


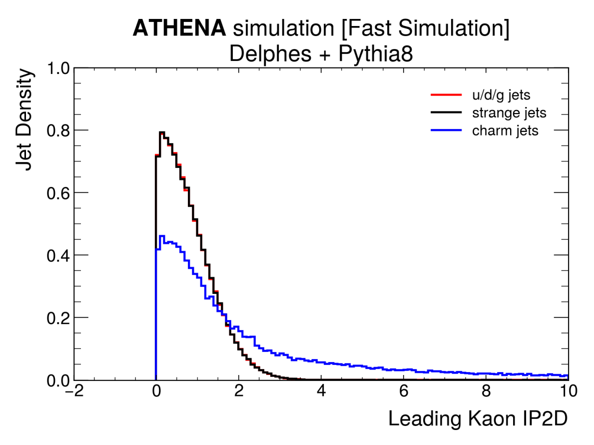


Kaon identification performances are entirely based on EIC UG PID group estimates of performance of various systems (e.g. aerogel or C_2F_6-based RICH in dRICH configuration, or barrel DIRC) implemented as identification selection efficiency maps in Delphes. No first-principles full simulation (so far) is employed for this part of the work.

Flavor-Tagging Jets: Displaced Kaon Counting [Work in Progress!]

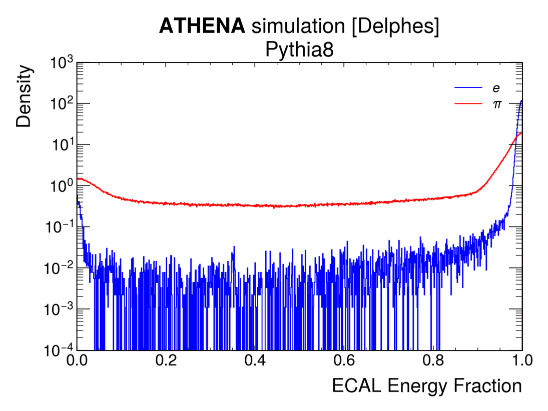






A quick optimization emphasizing a maximization of the charm-to-light jet ratio suggests reasonable cuts on the leading (by p_T) identified kaon: $q_K \times q_e < 0$, $\mathrm{sIP}_{3D} > 2.4$, and $IP_{2D} > 4.0$. This yields 7% (0.02%) efficiency on charm (light) jets.

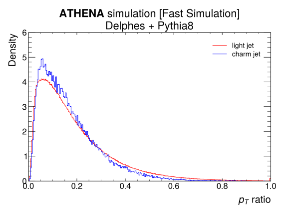
Flavor-Tagging Jets: Displaced Calo-Electron Counting [Work in Progress!]

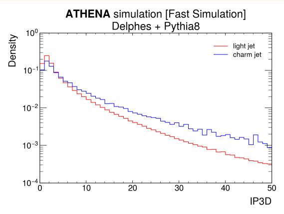


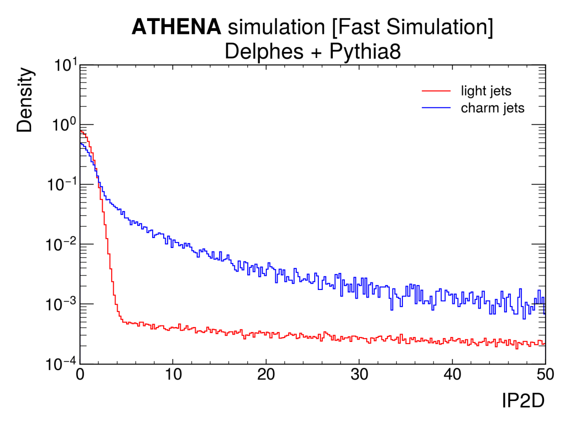
Rather than employ an electron ID parameterization, use full simulation of EMCAL and HCAL and single-particle events to model a very basic calorimeter-only electron identification approach. This approach uses a crude method that can be easily enhanced, selecting a track as an electron if $f_{EM} \equiv E_{EM}/(E_{EM} + E_{HAD}) > 0.991$. This achieves about 90% (20%) electron (pion) efficiency.

Even a single-variable ID approach like this, using full simulation, allows us to investigate the use of identified electrons in a tagging approach. We know multi-purpose detectors like ATHENA can do better than electron ID using a single observable, obviously! $\rightarrow$ a conservative look at this approach.

Flavor-Tagging Jets: Displaced Calo-Electron Counting [Work in Progress!]







A quick optimization emphasizing a maximization of the charm-to-light jet ratio suggests reasonable cuts on the leading (by p_T) identified electron: $p_T^e/p_T^{jet} < 0.54$, $\mathrm{sIP_{3D}} > 3.4$, and $IP_{2D} > 3.2$. This yields 6% (0.4%) efficiency on charm (light) jets.

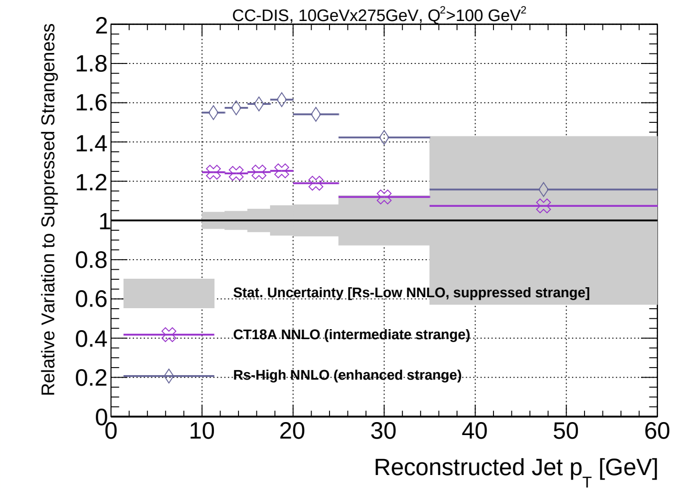
Flavor Tagging Efficiency - a preliminary synthesis

	Inclusive Charm Efficiency	Exclusive Charm Efficiency	Inclusive Light Efficiency	Exclusive Light Efficiency
CharmIPXDTagger	17.3	14.6	0.861	0.814
KTagger	6.86	4.69	0.0187	0.0141
ETagger	6.17	3.98	0.426	0.377

The combined charm (light) jet tagging efficiency from these early investigations is expected to be at the level of 23% (1.4%). These very preliminary investigations don't yet include key approaches, like secondary vertex finding, that are obviously commonplace for detectors with precision trackers.

We know from experience at running experiments that charm tagging using (displaced) tracks and vertices, especially combined using deep neural networks (c.f. Ref. [13, 14, 15]), can achieve 40% charm jet efficiency for 1-5% light jet efficiencies. Detectors like ATHENA will bring single-particle PID to the table, enhancing the potential for the EIC.

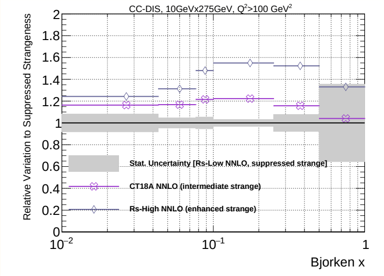
Statistical Uncertainty ($100\,\mathrm{fb}^{-1}$) vs. current theory uncertainty on Strangeness in the Proton

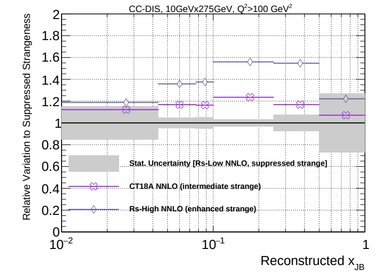


This is from a Yellow Report-era projection [2] using only displaced track-counting approaches (no PID or vertexing) with 20% (0.4%) charm (light) jet efficiency. Compare two variations of the CT18 PDFs: CT18NNLO with suppressed strangeness ($R_s = \frac{2s}{\bar{u}+\bar{d}} = 0.325$) and CT18ZNNLO with enhanced strangeness ($R_s = 0.863$). The variation in charm jet yields promises to vastly exceed statistical uncertainty from a "worst-case" (suppressed) scenario.

Even considering obvious missing uncertainties like jet energy scale ($\sim 10\%$), selection and other detector systematics, and the fact that this charm tagging efficiency is likely conservatively underestimated, this has the promise of greatly improving our knowledge of the strange PDF.

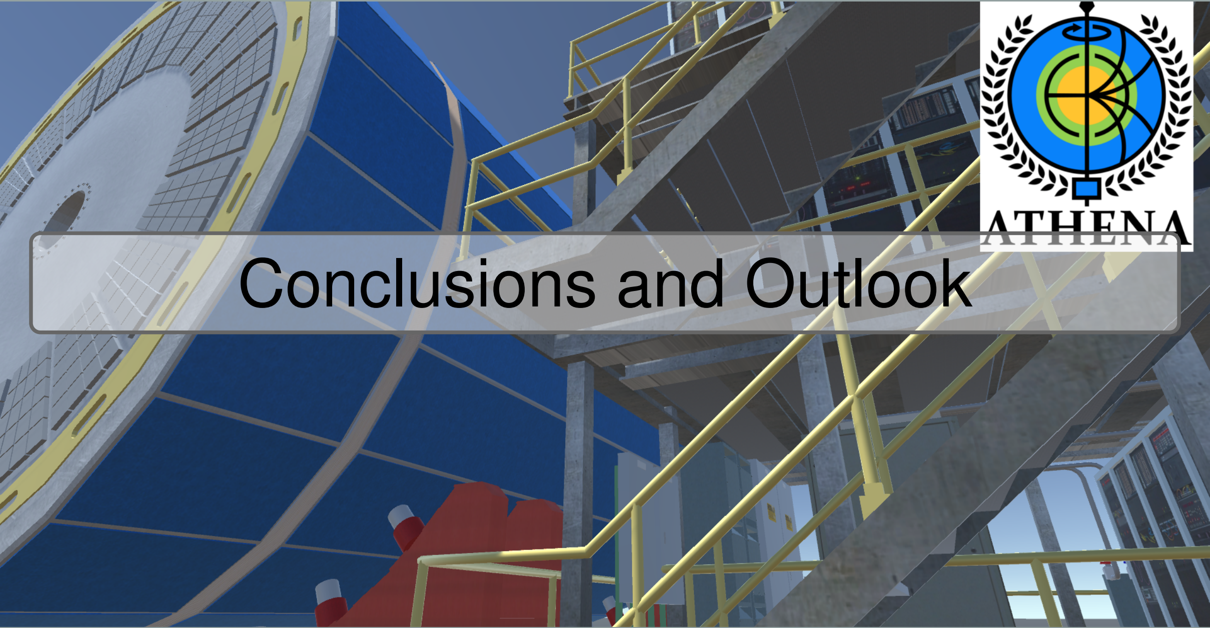
Statistical Uncertainty ($100\,\mathrm{fb}^{-1}$) vs. current theory uncertainty on Strangeness in the Proton





A comparison (again from Ref. [2]) of yield uncertainty (grey, suppressed strangeness) to the variation between enhanced and suppressed strangeness in the proton (blue markers), but for true Bjorken *x* and for the experimentally inferred "Jacquet-Blondel" *x* [16].

The best experimental sensitivity to the differences between these extrema is anticipated in the region of $x_{JB} \approx x = [0.05, 0.5]$.



Conclusions and Outlook

Conclusions and Outlook

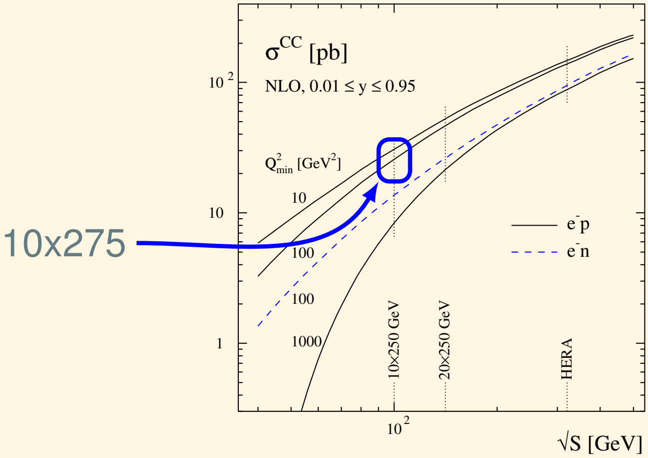
- CC DIS has been for a long time (and still is) a vital tool in probing instrinsic strangeness in the proton.
- The neutrino beam experiments, kaon semi-inclusive DIS, and LHC data (for example) have improved our knowledge of $s(x, Q^2)$ but there is so much more to be learned, especially at high-x.
- Ultimately, a **global analysis that combines approaches** will be the best way forward for $s(x, Q^2)$ and our community as a whole $\rightarrow$ take advantage of the strengths and weaknesses of all available approaches while gaining from the statistically independent methodologies!
- There are many tools in the arsenal of modern flavor taggers: displaced tracks, single-particle PID (combining calo, tracking, and dedicated PID information), vertexing, and multivariate classification $\rightarrow$ here, in the context of the ATHENA detector, I've teased some of these approaches and their net effect $\rightarrow$ likely still conservative.

A Future Tagger

- Displaced and other tracks
- Full secondary vertex reconstruction
- Sequential decay finding
- Jet substructure
- Single-particle PID information

Appendix

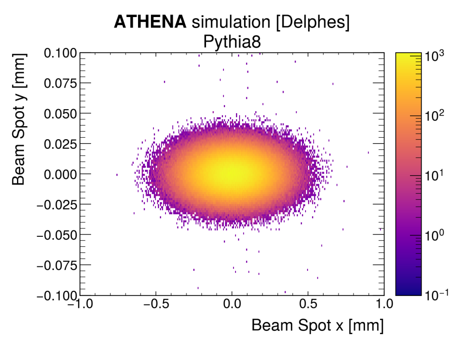
CC DIS NLO Cross-Section Prediction

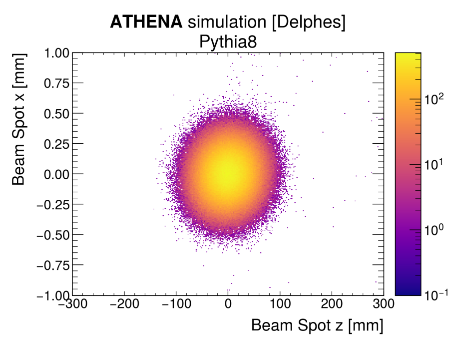


Total CC DIS rate at EIC with 10x275 configuration ($\sqrt{s} = 105\ \mathrm{GeV}$) is about 20pb.

Figure is from Ref. [17].

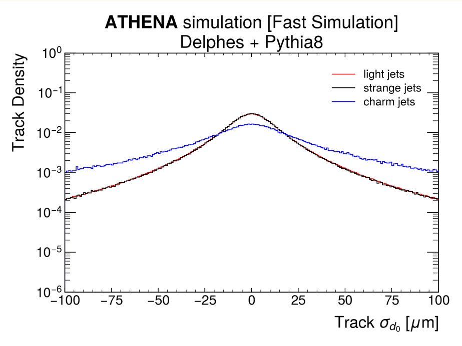
Beam Spot Modelling [11]

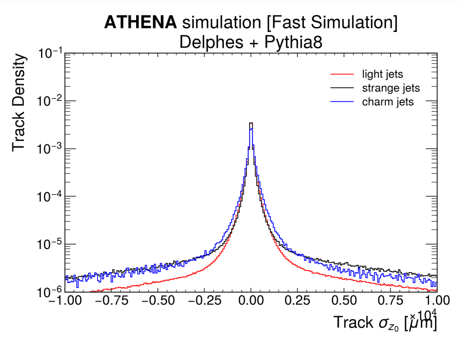




The EIC beam spot is modeled by configuring PYTHIA8 to move around the interaction point according to both a spread in three dimensions, as well as a tilt due to the beam crossing angle. Note the horizontal range changes from the left to the right plot, as the beam would be quite extended (e.g. many cm instead of many microns) in *z*.

Track DOCA (xy, z) Modeling





The modeled track d_0 and z_0 resolutions, which are derived in fast simulation from an independent simulation of the performance of an all-silicon tracking detector. These are tracks from jets in CC DIS. Pay close attention to the change in power of 10 on the x-axis, as the σ_{z_0} is smeared by two effects: the beam spot profile in z and the tracker resolution. Differences between jets are washed out in z at the track level by these effects. Vertexing should help recover this.

Thoughts on Charm Tagger Calibration for the EIC

Draw from lessons learned in LHC experiences in calibrating advanced jet taggers:

- Light Jets: cultivate a sample enriched in these jets. For example, use the negative-signed jets in sIP_{3D}, artificially flip their sign, and evaluate tag efficiency on these jets in data vs. simulation.
- Charm Jets: cultivate a sample of jets with exclusive, high-purity D meson final states fully reconstructed inside them. These yield "high confidence" that the parent jet is initiated by charm. Execute the tagger on these jets and compare performance in data vs. simulation.

Considerations: contamination of light by charm, charm by light, etc.; combination of SIDIS, exclusive, and jet-based analyses at end given potential overlap between jet approach and exclusive approach due to use of exclusive D states in calibration of tagger. We should welcome these challenges - it means we've succeeded in operating the EIC and its detector(s)!

References I

[1] **NuTeV** Collaboration, M. Goncharov *et al.*, "Precise Measurement of Dimuon Production Cross-Sections in ν_μ Fe and $\bar{\nu}_\mu$ Fe Deep Inelastic Scattering at the Tevatron.," *Phys. Rev. D* **64** (2001) 112006, `arXiv:hep-ex/0102049`.

[2] M. Arratia, Y. Furletova, T. J. Hobbs, F. Olness, and S. J. Sekula, "Charm jets as a probe for strangeness at the future Electron-Ion Collider," *Phys. Rev. D* **103** (2021) no. 7, 074023, `arXiv:2006.12520 [hep-ph]`.

[3] **ZEUS** Collaboration, I. Abt *et al.*, "Charm production in charged current deep inelastic scattering at HERA," *JHEP* **05** (2019) 201, `arXiv:1904.03261 [hep-ex]`.

[4] "The EIC Machine." `https://www.bnl.gov/eic/machine.php`.

[5] E. C. Aschenauer, M. D. Baker, A. Bazilevsky, K. Boyle, S. Belomestnykh, I. Ben-Zvi, S. Brooks, C. Brutus, T. Burton, S. Fazio, A. Fedotov, D. Gassner, Y. Hao, Y. Jing, D. Kayran, A. Kiselev, M. A. C. Lamont, J. H. Lee, V. N. Litvinenko, C. Liu, T. Ludlam, G. Mahler, G. McIntyre, W. Meng, F. Meot, T. Miller, M. Minty, B. Parker, R. Petti, I. Pinayev, V. Ptitsyn, T. Roser, M. Stratmann, E. Sichtermann, J. Skaritka, O. Tchoubar, P. Thieberger, T. Toll, D. Trbojevic, N. Tsoupas, J. Tuozzolo, T. Ullrich, E. Wang, G. Wang, Q. Wu, W. Xu, and L. Zheng, "eRHIC Design Study: An Electron-Ion Collider at BNL," 2014.

References II

[6] Ferdinand Willeke, "EIC accelerator and IR design status," in *2nd EIC Yellow Report Workshop at Pavia University*. 2020. https://indico.bnl.gov/event/8231/contributions/37060/.

[7] R. Abdul Khalek *et al.*, "Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report," arXiv:2103.05419 [physics.ins-det].

[8] M. Arratia and S. Sekula, "A Delphes card for the EIC yellow-report detector," arXiv:2103.06886 [physics.ins-det].

[9] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to pythia 8.1,"*Computer Physics Communications* **178** (Jun, 2008) 852867. http://dx.doi.org/10.1016/j.cpc.2008.01.036.

[10] T. Sjöstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual,"*Journal of High Energy Physics* **2006** (May, 2006) 026026. http://dx.doi.org/10.1088/1126-6708/2006/05/026.

[11] B. Page, "eicsimubeameffects," 2021. https://github.com/bspage912/eicSimuBeamEffects.

[12] **ATLAS** Collaboration, G. Aad *et al.*, "Configuration and performance of the ATLAS *b*-jet triggers in Run 2," arXiv:2106.03584 [hep-ex].

References III

[13] **CMS** Collaboration, A. M. Sirunyan *et al.*, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV,” *JINST* **13** (2018) no. 05, P05011, `arXiv:1712.07158 [physics.ins-det]`.

[14] **ATLAS** Collaboration, M. Aaboud *et al.*, “Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment,” *Phys. Rev. Lett.* **120** (2018) no. 21, 211802, `arXiv:1802.04329 [hep-ex]`.

[15] **ATLAS** Collaboration, “Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration,”.

[16] U. Amaldi *et al.*, “REPORT FROM THE STUDY GROUP ON DETECTORS FOR CHARGED CURRENT EVENTS,” in *ECFA Study of an ep Facility for Europe*, pp. 377–414. 1, 1979.

[17] E. C. Aschenauer, T. Burton, M. Stratmann, T. Martini, and H. Spiesberger, “Prospects for charged current deep-inelastic scattering off polarized nucleons at a future electron-ion collider,”*Physical Review D* **88** (Dec, 2013) . `http://dx.doi.org/10.1103/PhysRevD.88.114025`.