DESY Report 19-107

Probing the strange content of the proton with charm production in charged current at LHeC

xFitter Developers' team: Hamed Abdolmaleki[1], Valerio Bertone[2], Daniel Britzger[3], Stefano Camarda[4], Amanda Cooper-Sarkar[5], Achim Geiser[6], Francesco Giuli[7], Alexander Glazov[6], Agnieszka Luszczak[8], Ivan Novikov[9], Fred Olness[10], Andrey Sapronov[9], Oleksandr Zenaiev[11]

[1]Faculty of Physics, Semnan University, 35131-19111 Semnan, Iran
[2]Dipartimento di Fisica, Università di Pavia and INFN, Sezione di Pavia Via Bassi 6, I-27100 Pavia, Italy
[3]Max-Planck-Institut für Physik, Föhringer Ring 6, D-80805 München, Germany
[4]CERN, CH-1211 Geneva 23, Switzerland
[5]Particle Physics, Denys Wilkinson Bdg, Keble Road, University of Oxford, OX1 3RH Oxford, UK
[6]Deutsches Elektronen-Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany
[7]University of Rome Tor Vergata and INFN, Sezione di Roma 2, Via della Ricerca Scientifica 1, 00133 Rome, Italy
[8]T. Kosciuszko Cracow University of Technology, PL-30-084, Cracow, Poland
[9]Joint Institute for Nuclear Research, Joliot-Curie 6, Dubna, Moscow region, Russia, 141980
[10]SMU Physics, Box 0175 Dallas, TX 75275-0175, United States of America
[11]Hamburg University, II. Institute for Theoretical Physics, Luruper Chaussee 149, D-22761 Hamburg, Germany
Received: date / Accepted: date

arXiv:1907.01014v1 [hep-ph] 1 Jul 2019

Abstract We study charm production in charged-current deep-inelastic scattering (DIS) using the xFitter framework. Recent results from the LHC have focused renewed attention on the determination of the strange-quark parton distribution function (PDF), and the DIS charm process provides important complementary constraints on this quantity. We examine the current PDF uncertainty and use LHeC pseudodata to estimate the potential improvement from this proposed facility. As xFitter implements both fixed-flavor- and variable-flavor-number schemes, we can compare the impact of these different theoretical choices; this highlights some interesting aspects of multi-scale calculations. We find that the high-statistics LHeC data covering a wide kinematic range could substantially reduce the strange PDF uncertainty.

Contents

1 Introduction **2**

2 Theoretical predictions for CC charm production at the LHeC **2**
2.1 The heavy-flavor schemes 3
2.2 The reduced cross section 4
2.3 xFitter implementation 4

3 Comparison of theoretical predictions **4**
3.1 Comparison of theoretical predictions in the FFNS A and FONLL-B schemes 4
3.2 Additional comparisons 5
3.3 Contributions from different partonic subprocesses 8

4 PDF constraints from charm CC pseudodata **8**
4.1 The CC charm pseudodata 9
4.2 The profiled PDFs 10

5 Discussion and summary **12**

Appendix A: F_2^c Beyond leading-order **14**

[0]Correspondence: olness@smu.edu

1 Introduction

The deep-inelastic-scattering (DIS) experiments traditionally have provided important tests of perturbative QCD (pQCD) and are essential to precisely determine the parton distribution functions (PDFs) of the nucleon. In addition to the numerous dedicated fixed-target DIS experiments that have been performed so far, the HERA accelerator used colliding beams of leptons (electrons and positrons) and protons to investigate the nucleon structure. The broad kinematic coverage of the HERA charge-current (CC) and neutral-current (NC) DIS data in terms of the negative virtuality Q^2 of the exchanged vector boson and the Bjorken variable x_{Bj} is such that these data have significant impact on the determinations of the PDFs [1–5].

In the Standard Model (SM), the charm quark plays an important role in the investigation of the nucleon structure [6–9]. In the NC case, the photon-gluon fusion process for charm production was calculated at $\mathcal{O}(\alpha_s^2)$ with the full heavy-quark mass dependence included in the DIS hard cross sections [10, 11]. The heavy-quark mass effects in the CC process have been calculated to $\mathcal{O}(\alpha_s)$ in Refs. [12–16], and the recent work of Ref. [17] provides results up to $\mathcal{O}(\alpha_s^2)$. The large-Q^2 contributions of heavy flavors to the xF_3 structure function had already been computed in Ref. [18]. In many of the posited models which extend the SM, the coupling to "new physics" is proportional to the particle mass; hence, the heavy quarks will have an enhanced coupling and provide an optimal testing ground for these searches.

Heavy quarks also play a critical role in helping us fully characterize the SM, and the charm quark is especially useful in this respect as it can provide us direct access to the strange-sea quark distribution. The strange sea has been extensively investigated in a number of experiments including the associated production of a W boson with a charm-jet final state, which (at LO) arises from strange–gluon initial states [19–26]. Additionally, charm production in neutrino/antineutrino-nucleon DIS has been studied by a number of experiments including: CCFR [27], NuTeV [28], CHORUS [29], CDHSW [30] and NOMAD [31]. With a sign-selected beam ($\nu/\bar{\nu}$), these experiments can separately probe the strange $s(x)$ and anti-strange $\bar{s}(x)$ distributions. While the neutrino DIS experiments provide detailed information on the shape of the strange distribution, the normalization is a challenge, as that is tied to the beam flux. Separately, the HERMES collaboration used charged-lepton DIS production of charged kaons to provide a complementary extraction of $s(x)+\bar{s}(x)$ at LO [32]. Recently, charm production in CC DIS was measured for the first time in $e^{\pm}p$ collisions by ZEUS [33].

Additionally, charm production mediated by electroweak gauge boson at hadron colliders provides important information on the strange- and charm-quark distributions, and is complementary to the DIS final-state charm-quark experiments [34]. The Tevatron measured the charm-quark cross section in association with a W boson at CDF [19, 35, 36] and D0 [20], but these results were limited by low statistics.

In lieu of significant experimental constraints, many global QCD analyses tie the strange distribution to the light-sea quarks via the relation $s=\bar{s}=r_s\bar{d}$. While in principle r_s depends on both x_{Bj} and Q^2, it is often set to a fixed value [37, 38].

Using inclusive leptonic decays of W and Z bosons, the ATLAS experiment has obtained a value of $r_s=1.19\pm0.16$ at $x=0.023$ and $Q_0^2=1.9$ GeV^2 [39]. Additionally, using the cross section ratio for $W^{\pm}+c$ final states they also find a comparably large value for r_s [23]. In contrast, CMS results generally prefer lower r_s values [22, 24]. However, a recent analysis using both ATLAS and CMS data suggests that the LHC data support unsuppressed strangeness in the proton. While the result is dominated by ATLAS, this is not in contradiction with the CMS data [22, 23, 39, 40].

Looking to the future, it is clearly important to reduce the uncertainty of the strange-quark PDF as we strive to make increasingly precise tests of the SM and search for what might lie beyond. The proposed Large Hadron Electron Collider (LHeC) program has the ability to provide high statistics measurements of electrons on both protons and nuclei across a broad kinematic range to address many of these outstanding questions.

In this investigation, we make use of the xFitter tools [41] (version 2.0.0) to study the present constraints on the strange-quark PDFs, and then use LHeC pseudodata [42] to infer how these might improve. Furthermore, as xFitter implements both fixed-flavor- and variable-flavor-number schemes, we can examine the impact of these different theoretical choices.

This paper is organized as follows. In Sect. 2 we outline the theoretical details of the different heavy-flavor schemes. In Sect. 3 we compare the theoretical predictions of the different schemes across the kinematic range, and examine the individual partonic contributions. In Sect. 4 we study the impact of the LHeC pseudodata on the PDFs using a profiling technique. In Sect. 5 we provide some discussion and summarize the results. Finally, in Appendix A we discuss some of the more subtle theoretical issues that we encounter at higher orders.

2 Theoretical predictions for CC charm production at the LHeC

The proposed Large Hadron Electron Collider (LHeC) [42] would collide a newly built electron beam with the LHC

hadron beam at a center of mass energy of $\sqrt{s} = \sqrt{4E_eE_p}$; thus the 7 TeV proton beam on a 60 GeV electron beam provide $\sqrt{s} \sim 1.3$ TeV. Compared to HERA, the LHeC extends the covered kinematic range by an order of magnitude in both x_{Bj} and Q^2 with a nominal design luminosity of $10^{33} cm^{-2}s^{-1}$.

Theoretical predictions are calculated for electroweak charged-current (CC) charm production in ep collisions at the LHeC at centre-of-mass energy $\sqrt{s} = 1.3$ TeV, using a variety of heavy-flavor schemes. The predictions are provided for unpolarized beams in the kinematic range $100 < Q^2 < 100000$ GeV^2, $0.0001 < x_{\mathrm{Bj}} < 0.25$. They are calculated as reduced cross sections at different Q^2, x_{Bj} and inelasticity (y) points. The covered y range is $0.0024 < y < 0.76$.

Experimentally, however, not charm quarks but charmed hadrons (or rather their decay products) are registered in the detectors. Therefore, extrapolation to the inclusive charm-production cross section has to be carried out in a model-dependent way. Furthermore, CC production of charm quarks in the final state can happen via both electroweak and QCD processes. The former leads to an odd number of charm quarks in the final state with the W boson having the same electric charge as the sum of the electric charges of final-state charm quarks, while the latter creates an even number of charm quarks with total electric charge equal to zero. If the electric charge of the tagged charm quark can be accessed experimentally (*e.g.* when reconstructing D mesons), the QCD contribution can be subtracted by taking the difference of the yields in the events with odd and even numbers of charm quarks, otherwise the QCD contribution can be estimated only in a model-dependent way.

The CC charm process directly depends on the CKM matrix [50]. Here, the CKM matrix elements V_{cd} and V_{cs} are particularly relevant and we use the values $V_{cd} = 0.2252$ and $V_{cs} = 0.9734$. Three different heavy-flavor schemes are employed, all including a full treatment of charm-mass effects up to NLO, *i.e.* $\mathcal{O}(\alpha_s)$; in the following we describe them in detail for the particular application to CC electron-proton reactions.

2.1 The heavy-flavor schemes

The standard "A" variant of the fixed-flavor number scheme (FFNS), which we identify as **FFNS A**, uses three light flavors in both PDFs and α_s evolution for all scales, while heavy flavors (here, charm) are produced exclusively in the matrix-element part of the calculation. This scheme has been used for the PDF determinations and cross section predictions of the ABM(P) group [4, 43–45], as well as in the FF3A variant of the HERAPDF analysis [2], and implemented in xFitter through the OPENQCDRAD package [46].

Next, the "B" variant of the FFNS (**FFNS B**), known as the "mixed" or "hybrid" scheme [6] is also used. In this scheme, the number of active flavors is still fixed to three in the PDFs, relying exclusively on $\mathcal{O}(\alpha_s)$ fully massive matrix elements for charm production, while the number of flavors is allowed to vary in the virtual corrections of the α_s evolution. Corrections to the α_s evolution involving heavy-flavor loops are thus included and resummed to all orders, while no resummation is applied to other higher order corrections. This procedure will catch a fraction of the "large logs" which might spoil the fixed-flavor scheme convergence at very high scales, and is possible since the masses of the charm and beauty quarks provide natural cutoffs for infrared and collinear divergences. This scheme was used in the HERAPDF FF3B variant [2] and in applications of the HVQDIS program [6]. In general, the transition from the FFNS A to the FFNS B requires a readjustment of the treatment of matrix elements involving heavy-flavor loops. In the specific case of CC production, no such loops occur up to NLO (at NNLO they do), so that the same matrix elements can be used for both schemes; thus the only difference is in the α_s evolution.

Finally, for the variable-flavor-number scheme (VFNS) we use the "B" variant of the fixed-order-next-to-leading-log scheme (**FONLL-B**) [47] which combines the NLO $\mathcal{O}(\alpha_s)$ massive matrix elements of the FFNS with the $\mathcal{O}(\alpha_s)$ massless results of the zero-mass variable-flavor-number scheme (ZM-VFNS), allowing the number of active flavors to vary with scale, and all-order next-to-leading log resummation of (massless) terms beyond NLO. It thus explicitly includes charm and beauty both in the PDFs and in the evolution of the strong coupling constant. Whenever terms would be double-counted in the merging of the two schemes, the massless terms are eliminated in favour of the massive ones. The FONLL scheme is commonly used by the NNPDF group [5] and implemented in xFitter through the APFEL package [48].

In summary, the schemes used are:

- **FFNS A:** a NLO FFNS with $n_f = 3$ at all scales, used with the ABMP16 [45] or HERAPDF2.0 FF3A [2] NLO PDF sets.
- **FFNS B:** a NLO FFNS with $n_f = 3$ for the PDFs and variable n_f for α_s, used with the HERAPDF2.0 FF3B [2] NLO PDF set.
- **FONLL-B :** a VFNS used with the NNPDF3.1 NLO PDF set [5].

The PDF sets are available via the LHAPDF interface (version 6.1.5) [49].

2.2 The reduced cross section

The reduced CC charm-production cross sections can be expressed as a linear combination of structure functions:

$$\sigma^{\pm}_{\text{charm,CC}} = \frac{1}{2}\left(Y_+ F_2^{\pm} \mp Y_- x F_3^{\pm} - y^2 F_L^{\pm}\right), \tag{1}$$

with

$$Y_{\pm} = 1 \pm (1-y)^2 . \tag{2}$$

In the quark-parton model, when we neglect the gluons, the structure functions become:

$$\begin{aligned} F_2^+ &= xD + x\overline{U}, \\ F_2^- &= xU + x\overline{D}, \\ F_L &= 0, \\ xF_3^+ &= xD - x\overline{U}, \\ xF_3^- &= xU - x\overline{D}. \end{aligned} \tag{3}$$

The terms xU, xD, $x\overline{U}$ and $x\overline{D}$ denote the sum of parton distributions for up-type and down-type quarks and anti-quarks, respectively.[1] The $\pm$ superscript on σ and F corresponds to the sign of $W^{\pm}$. Below the b-quark mass threshold, these sums are related to the quark distributions as follows:

$$\begin{aligned} xU &= xu + xc, \\ x\overline{U} &= x\overline{u} + x\overline{c}, \\ xD &= xd + xs, \\ x\overline{D} &= x\overline{d} + x\overline{s}. \end{aligned} \tag{4}$$

In the FFNS the charm-quark densities are zero. In the phase-space corners $y \to 0$ and $y \to 1$ and using the same quark-parton model approximation, we have the following asymptotic relations:

$$\begin{aligned} y \to 0: &\quad \sigma^{\pm}_{\text{charm,CC}} = F_2^{\pm} = xD(x\overline{D}) + x\overline{U}(xU), \\ y \to 1: &\quad \sigma^{\pm}_{\text{charm,CC}} = \frac{1}{2}(F_2^{\pm} \mp xF_3^{\pm}) = x\overline{U}(xU). \end{aligned} \tag{5}$$

Thus the contribution from the strange-quark PDF is suppressed at high y.

2.3 xFitter implementation

All calculations are interfaced in xFitter and available with $\overline{\text{MS}}$ heavy-quark masses. The reference value of the $\overline{\text{MS}}$ charm mass is set to $m_c(m_c) = 1.27$ GeV [50], and α_s is set to the value used for the corresponding PDF extraction: $\alpha_s(M_Z) = 0.1191$ for ABMP16 and $\alpha_s(M_Z) = 0.118$ for NNPDF3.1. The renormalization and factorization scales are chosen to be $\mu_r^2 = \mu_f^2 = Q^2$.

To estimate theoretical scale uncertainties, μ_r and μ_f are simultaneously varied up and down by a factor of two. In the case of the FONLL-B calculations, also the independent μ_r and μ_f variations are checked. Furthermore, the PDF uncertainties are propagated to the calculated theoretical predictions, while the uncertainties arising from varying the charm mass $m_c(m_c) = 1.27 \pm 0.03$ GeV by one standard deviation are smaller than 1% and therefore neglected. In the FONLL-B scheme, as a cross check, the calculation was performed with the pole charm mass $m_c^{\text{pole}} = 1.51$ GeV which is consistent with the conditions of the NNPDF3.1 extraction [5]. The obtained theoretical predictions differ from the ones calculated with $m_c(m_c) = 1.27$ GeV by less than 1%. The total theoretical uncertainties are obtained by adding in quadrature scale and PDF uncertainties.

3 Comparison of theoretical predictions

We now provide some numerical comparisons of the heavy-flavor schemes using their separate input conditions and associated PDF sets. Caution is necessary in these comparisons as the PDF sets are extracted with different input assumptions, data sets, and tolerance criteria; this is, in part, why we shall separately display the μ_r, μ_f and PDF uncertainties in the following.

3.1 Comparison of theoretical predictions in the FFNS A and FONLL-B schemes

Figs. 1, 2 and 3 show theoretical predictions for the FFNS A and FONLL-B schemes calculated as described in the previous sections with their total uncertainties. The FFNS A and FONLL-B results agree reasonably well within uncertainties in the bulk of the phase space. However, in phase-space corners such as $Q^2 \gtrsim 10000$ GeV^2 or small y the predictions in the two schemes differ by more than 50%, exceeding the theoretical uncertainties.

To examine these differences further, in Fig. 4 we separately compute PDF and scale uncertainties (setting $\mu_r = \mu_f = \mu$) of the charm CC cross section as a function of Q^2 for different values of x_{Bj} calculated in the FFNS A and FONLL-B scheme.

Comparing the two schemes, the larger variation of the FONLL-B scheme reflects the larger PDF uncertainty of the underlying PDF sets used: ABMP16 for FFNS A and NNPDF3.1 for FONLL-B. This difference is most evident in Fig. 4 which specifically separates out the PDF uncertainty, and reflects the independent inputs and assumptions used in the different PDF extractions.

[1] In these expressions, we neglect the CKM mixing for brevity, but it is fully contained in the calculations.

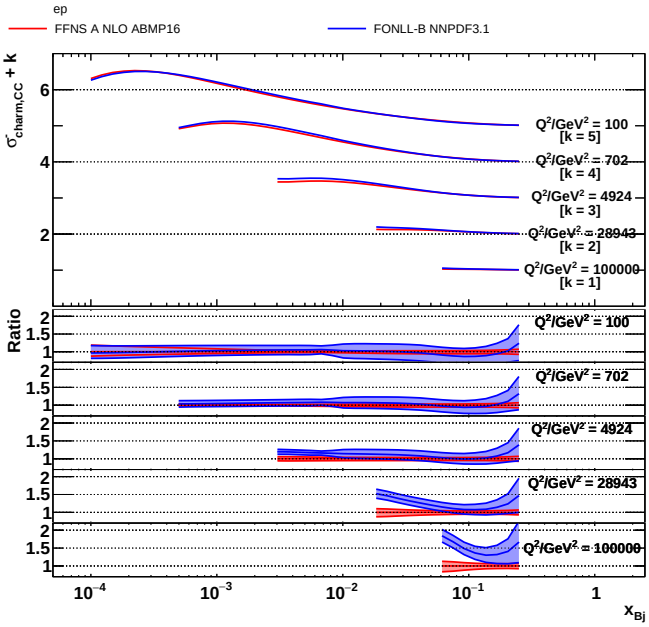

Figure 1 The theoretical predictions with their total uncertainties for charm CC production at the LHeC as a function of x_{Bj} for different values of Q^2 calculated in the FFNS A and FONLL-B schemes. The bottom panels display the theoretical predictions normalized to the nominal values of the FFNS A predictions.

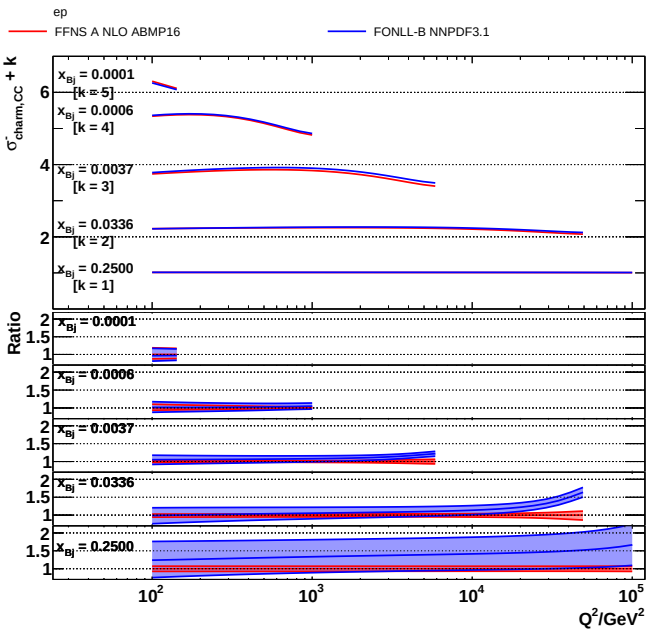

Figure 2 The theoretical predictions with their total uncertainties for charm CC production at the LHeC as a function of Q^2 for different values of x_{Bj} calculated in the FFNS A and FONLL-B schemes. The bottom panels display the theoretical predictions normalized to the nominal values of the FFNS A predictions.

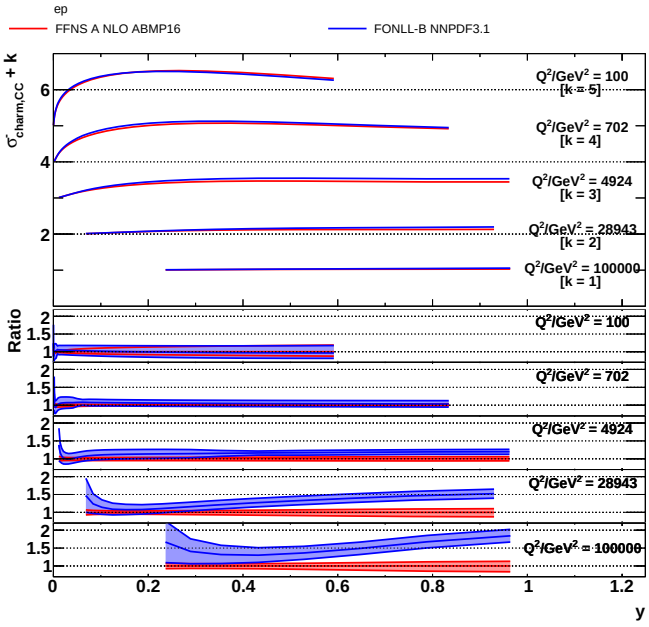

Figure 3 The theoretical predictions with their total uncertainties for charm CC production at the LHeC as a function of y for different values of Q^2 calculated in the FFNS A and FONLL-B schemes. The bottom panels display the theoretical predictions normalized to the nominal values of the FFNS A predictions.

Examining the results of Fig. 4, we also observe some other interesting features. For both of the calculations, the PDF uncertainties are relatively stable across the Q^2 range for fixed x_{Bj}, but tend to increase at larger x_{Bj} values. As is well known, in pQCD calculations the effect of scale variations is indicative of the convergence of the series. We observe that the scale uncertainties for the FONLL-B scheme uniformly decrease with increasing Q^2. For the FFNS A scheme, the scale uncertainties decrease for small x_{Bj} values but increase with Q^2 at intermediate values of x_{Bj}. Additional details are shown in Fig. 5 where we separately vary μ_{r} and μ_{f} for the FONLL-B scheme. Here we note that the uncertainty associated to μ_{r} is very small and the total scale uncertainty is dominated by the variations of μ_{f} which is tied to the PDFs, $f_i(x, \mu_{\mathrm{f}})$. For the FFNS A in XFITTER, it is not possible to separately vary μ_{r} and μ_{f} in the current implementation, so the separate uncertainties can only be inferred by comparison to the FONLL-B case.

3.2 Additional comparisons

To further explore whether the differences between the two sets of theoretical predictions are due to the different treatment of heavy quarks or to the different PDF sets, theoretical calculations in FFNS A and FONLL-B are repeated with the HERAPDF2.0 PDF sets extracted from the HERA DIS data [2]. Predictions in the FFNS B scheme are

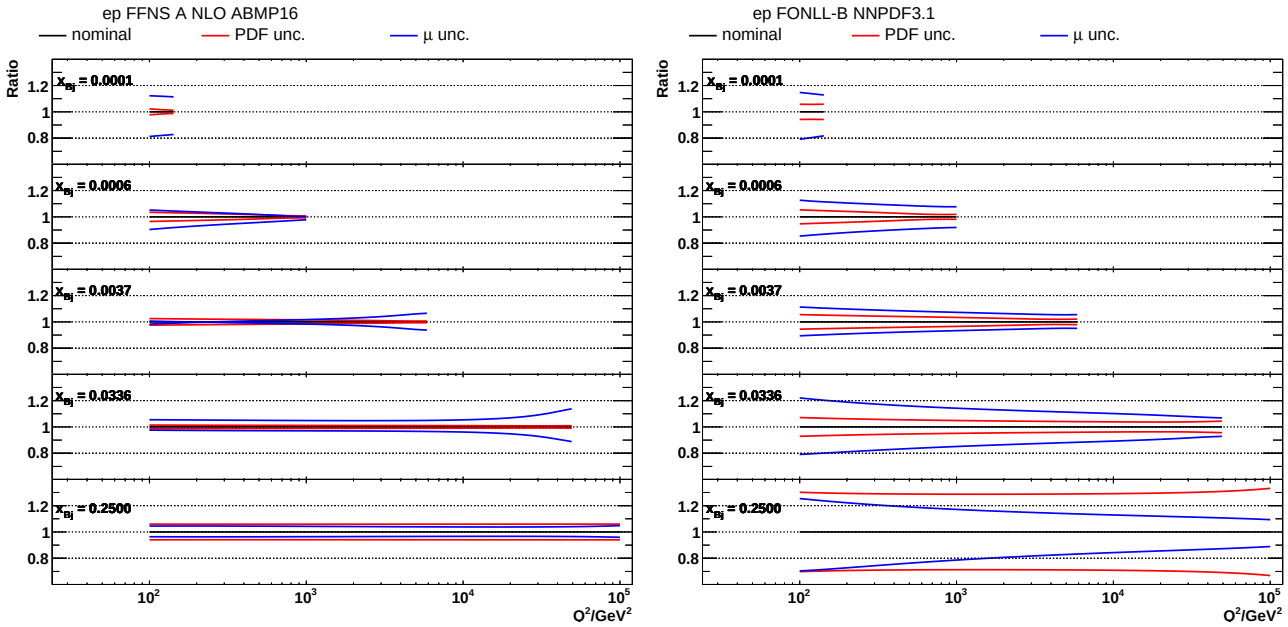

Figure 4 Relative theoretical uncertainties of charm CC predictions for the LHeC as a function of Q^2 for different values of x_{Bj} calculated in the FFNS A and FONLL-B schemes. The PDF and scale uncertainties are shown separately.

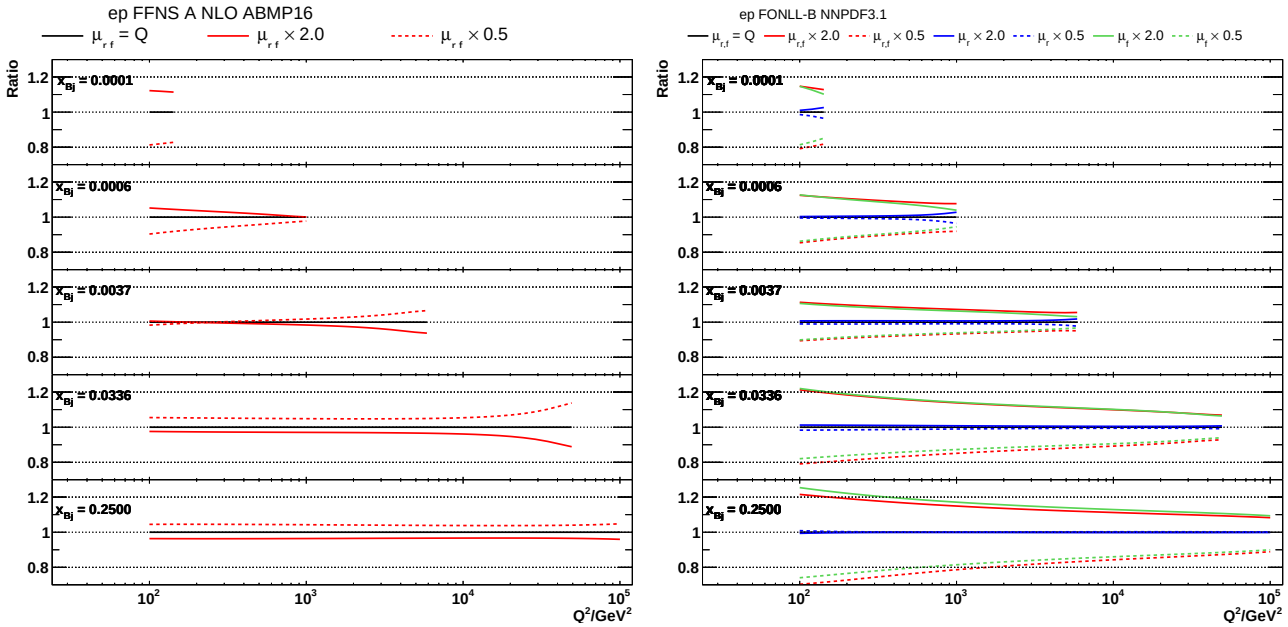

Figure 5 The impact of separate scale variations on charm CC predictions for the LHeC as a function of Q^2 for different values of x_{Bj} calculated in the FFNS A and FONLL-B schemes.

also produced using the HERAPDF2.0 FF3B PDF set and the FFNS B matrix elements, which are equivalent to the FFNS A matrix elements at NLO for CC charm production. The results are displayed in Fig. 6. The differences between FFNS A and FONLL-B are similar to those displayed in Figs. 1-3 and demonstrate that these differences arise from the different treatment of the heavy quarks in the two schemes. The FFNS B predictions lie between the FFNS A and FONLL-B predictions, indicating that a large part of the difference is due to the different treatment of heavy quarks in the running of α_s at high x_{Bj} or low y.

Furthermore, to investigate the impact of the NNLO corrections available at $Q \gg m_c$ for the FFNS calculation, approximate NNLO predictions are obtained using the ABMP16 NNLO PDF set [4]. The results for the cross section as a function of Q^2 for different values of x_{Bj} are shown in Fig. 7, where they are compared to the NLO FFNS A predictions from Fig. 2. The NNLO corrections do

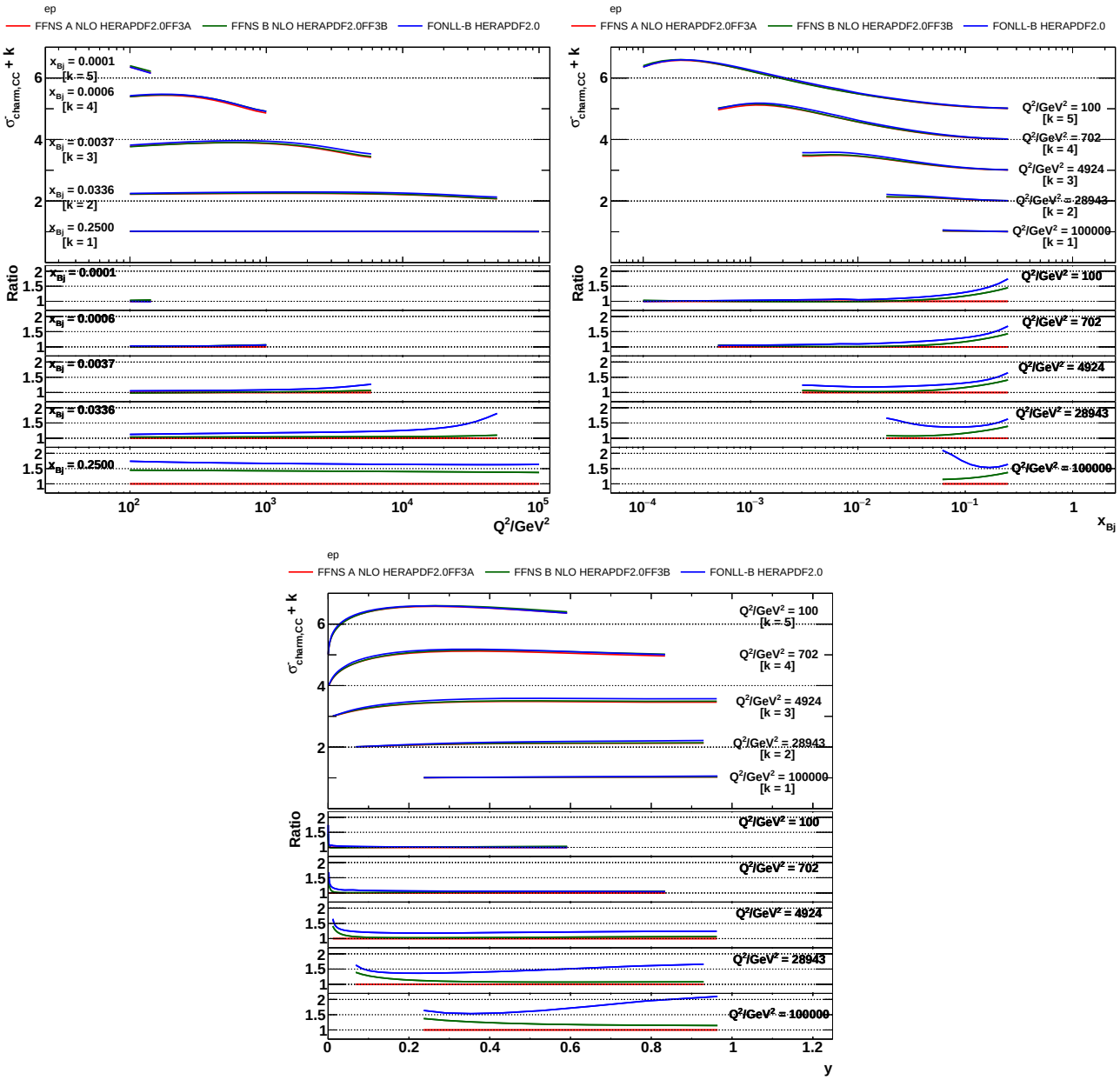

Figure 6 The theoretical predictions for CC charm production at the LHeC as a function of Q^2 (x_{Bj}, y) for different values of x_{Bj} (x_{Bj}, Q^2) obtained using the HERAPDF2.0 PDF sets in the FFNS A, FFNS B and FONLL-B schemes. The bottom panels display the theoretical predictions normalized to the nominal values of the FFNS A predictions.

not exceed 10% and thus cannot account for the differences between the FFNS A and FONLL-B theoretical predictions. Similar results are observed for the cross sections as functions of other kinematic variables.

To better understand the differences between the FFNS and VFNS calculations, Fig. 6 is particularly instructive. We see that at low Q^2 the FFNS (FFNS A and FFNS B) and VFNS (FONLL-B) results agree within uncertainties (as demonstrated in Fig. 2). When the scale μ is below the charm-threshold scale μ_c (typically taken to be equal to $m_c(m_c)$) the charm PDFs vanish and the FFNS and VFNS reduce to the same result.[2] For increasing scales, the VFNS resums the $\alpha_s \ln(\mu^2/\mu_c^2)$ contributions via the DGLAP evolution equations and the FFNS and VFNS will slowly diverge logarithmically. This behavior is observed in Fig. 6 and is consistent with the characteristics demonstrated in Ref. [52].

More precisely, Ref. [52] used a matched set of $n_f = 3$ and $n_f = 5$ PDFs to study the impact of the scheme choice at

[2]Note that while the charm-threshold scale μ_c is commonly set to the charm quark mass $m_c(m_c)$, the choice of μ_c is arbitrary and amounts to a renormalization scheme choice [51].

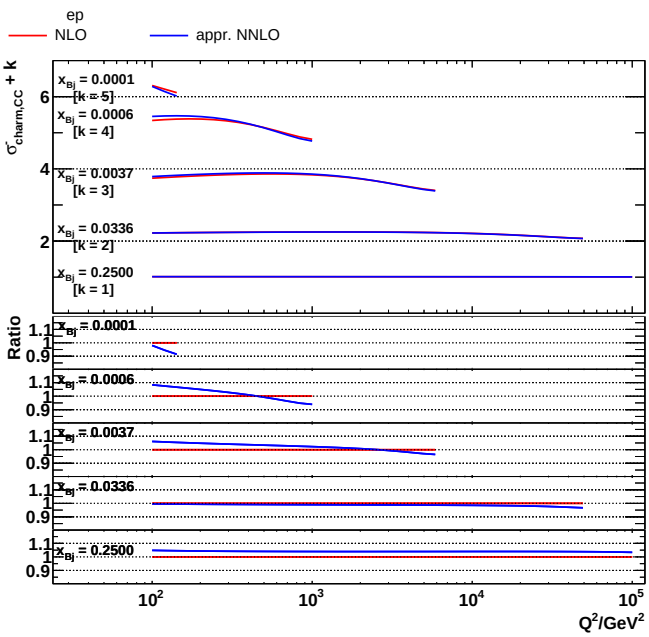

Figure 7 The theoretical predictions for CC charm production at the LHeC as a function of Q^2 for different values of x_{Bj} calculated in the FFNS A scheme at NLO and approximate NNLO. The bottom panels display the theoretical predictions normalized to the nominal values of the FFNS A NLO predictions.

large scales. They found that the resummed contributions in the VFNS yielded a larger cross section than the FFNS (the specific magnitude was x-dependent), and that for Q^2 scales more than a few times the quark mass, the differences due to scheme choice exceeded the differences due to (estimated) higher-order contributions. Thus, we have identified the source of the scheme differences at large Q^2.

The source of the scheme differences at large x_{Bj} is a bit more subtle. The VFNS includes a resummation of higher-order logarithms of the form $\alpha_s \ln(\mu^2/\mu_c^2)$. In Fig. 18 of the Appendix we display the separate contributions of the VFNS for a choice of $\{x_{\mathrm{Bj}}, Q^2\}$; the difference between the LO and SUB curves is indicative of the additional contribution of the resummed logarithms. This contribution depends on the particular x_{Bj} value and we find (*c.f.,* Fig. 11 of Ref. [52]) that these terms increase for larger x_{Bj} values. Figure 6 indicates that a large fraction of these seems to be caught by the FFNS B scheme. Thus, we have identified the source of the scheme differences at large x_{Bj}.

3.3 Contributions from different partonic subprocesses

The fundamental difference between the FFNS and the VFNS is the treatment of the heavy partons, the charm in particular. In the FFNS the charm is not included in the PDFs as an active parton, so charm quarks only arise from gluon splitting, $g \to c\bar{c}$. In contrast, the VFNS does include the charm as an active partonic flavor, and thus allows for charm-initiated subprocesses. To better appreciate these differences, we will study the individual partonic contributions to the cross section as functions of the kinematic variables x_{Bj}, Q^2, and y.

Figs. 8, 9 and 10 show the contributions from separate partonic subprocesses to the CC charm production cross section in the FFNS A and FONLL-B schemes as a function of: x_{Bj} for different values of Q^2, Q^2 for different values of x_{Bj}, and y for different values of Q^2, respectively.

In these figures we observe that the gluon contribution to the FFNS is strikingly similar to the charm contribution to the VFNS. This is explained by the fact that in the FFNS the charm is present only in the final state and produced predominantly in the hard process $\gamma g \to c\bar{c}$. In contrast, in the VFNS the charm is present also in the initial state and mainly produced by $g \to c\bar{c}$ collinear splitting through DGLAP evolution. The fundamental underlying process is (and has to be) the same in both the FFNS and VFNS, but the factorization boundary between PDFs and hard scattering cross section, $\hat{\sigma} \otimes f$, (determined by the scale μ and the scheme choice) is different.[3]

These figures highlight another interesting feature of the QCD theory; we observe that for the VFNS the gluon contribution (green curves) can become negative in particular kinematic regions. This is because in the VFNS we combine the gluon-boson fusion process (the NLO terms of Figs. 16 and 17) with the counter-term (the SUB terms), and this combination can be negative. This behavior underscores the fact that the renormalization scale μ is simply "shuffling" contributions among the separate sub-pieces, but the total physical cross section remains positive and stable, *cf.*, Fig. 18 and Ref. [53]. This is a triumph of the QCD theory.

Next, turning our attention to the strange PDF contribution, it is notable that the FFNS and VFNS behave qualitatively very similar as functions of Q^2, x_{Bj}, and y. In particular, we observe that the strange fraction increases for x_{Bj} and decreases for Q^2 and y. In particular, at high y the strange PDF contribution drops to zero in favor of the gluon or charm quark PDFs (see Fig. 10 and Eq. (5)). Similar phenomena (although less pronounced) are observed at low x_{Bj} and/or high Q^2. In these phase-space regions, the dominant contributions to the cross section are proportional to the gluon PDF in the FFNS or to the charm-quark PDFs in the VFNS.

4 PDF constraints from charm CC pseudodata

Now we turn to examine how the LHeC can reduce the PDF uncertainties and thus improve our predictive power.

[3]Note there is a "subtraction" term which closely matches the LO process, but this $\mathcal{O}(\alpha_s)$ process is contained in the NLO gluon-initiated contribution. For details, see Appendix A

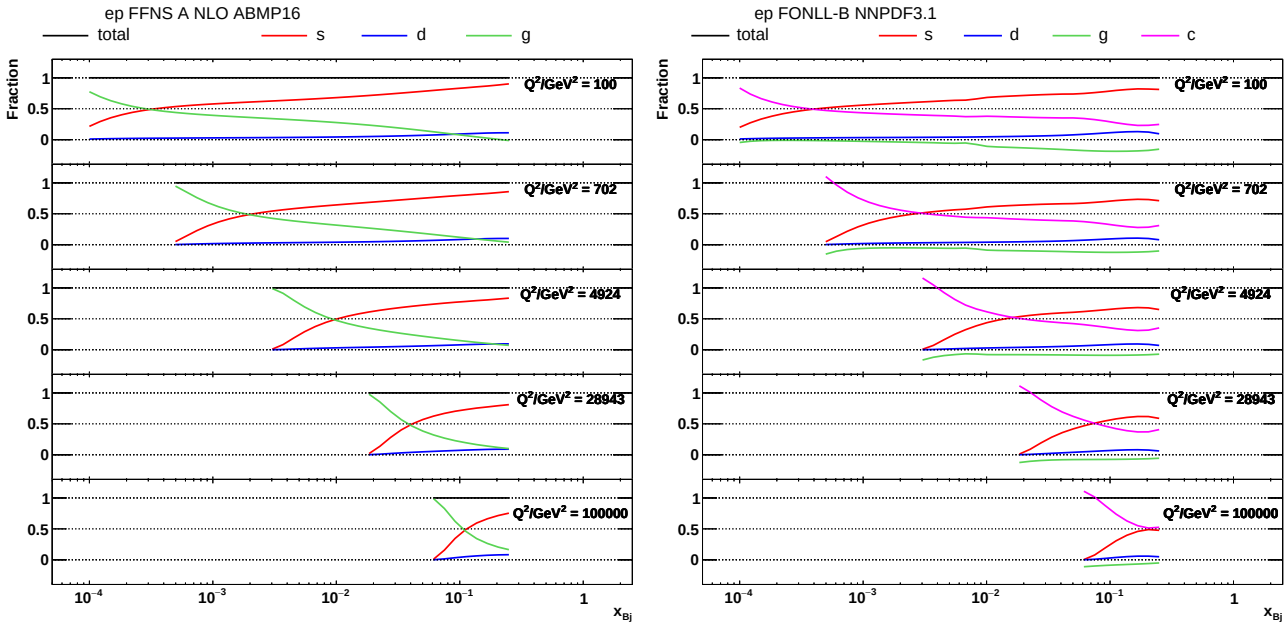

Figure 8 The partonic subprocesses for charm CC production cross sections in the FFNS A (left) and FONLL-B (right) schemes as a function of x_{Bj} for different values of Q^2.

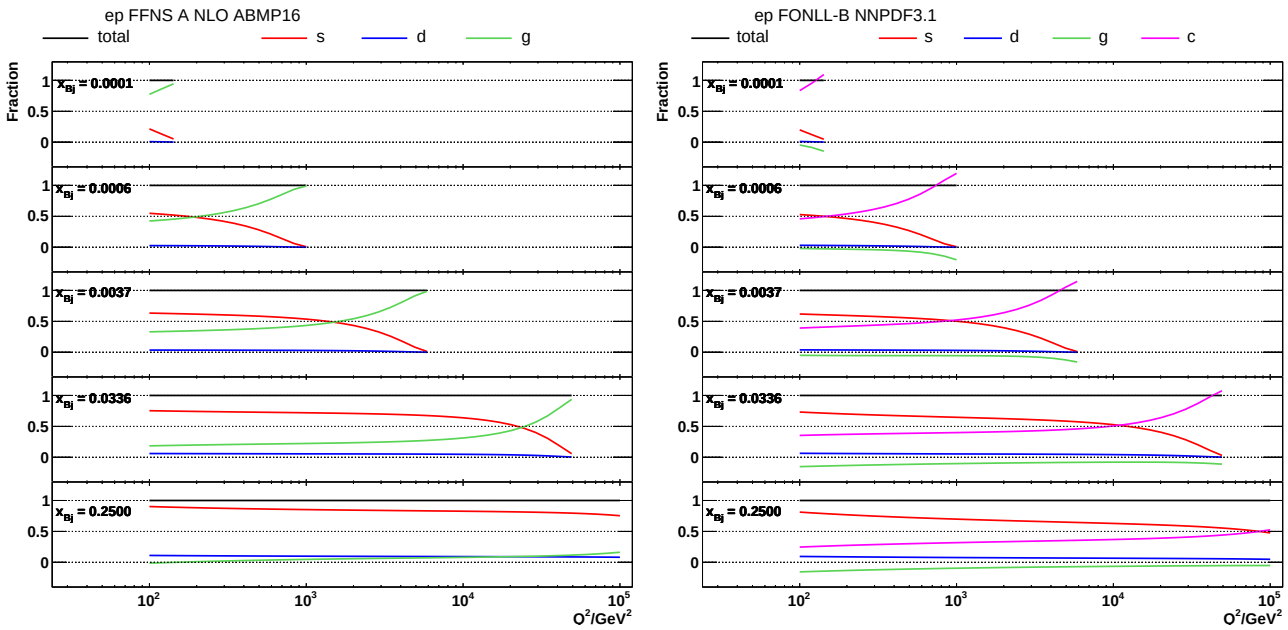

Figure 9 The partonic subprocesses for charm CC production cross sections in the FFNS A (left) and FONLL-B (right) schemes as a function of Q^2 for different values of x_{Bj}.

The impact of charm CC cross section measurements at the LHeC on the PDFs is quantitatively estimated using the profiling technique [54]. This technique is based on minimizing the χ^2 between data and theoretical predictions taking into account both experimental and theoretical uncertainties arising from PDF variations. Two NLO PDF sets were chosen for this study: ABMP16 [45] and NNPDF3.1 [5]. All PDF sets are provided with uncertainties in the format of eigenvectors. In the presence of strong constraints (the LHeC data is very precise), it is preferable to use the eigenvector representation as only a few MC replicas would survive the Bayesian reweighting.

4.1 The CC charm pseudodata

For this study, pseudodata for charm CC production cross section differential in Q^2 and x_{Bj} and corresponding to an

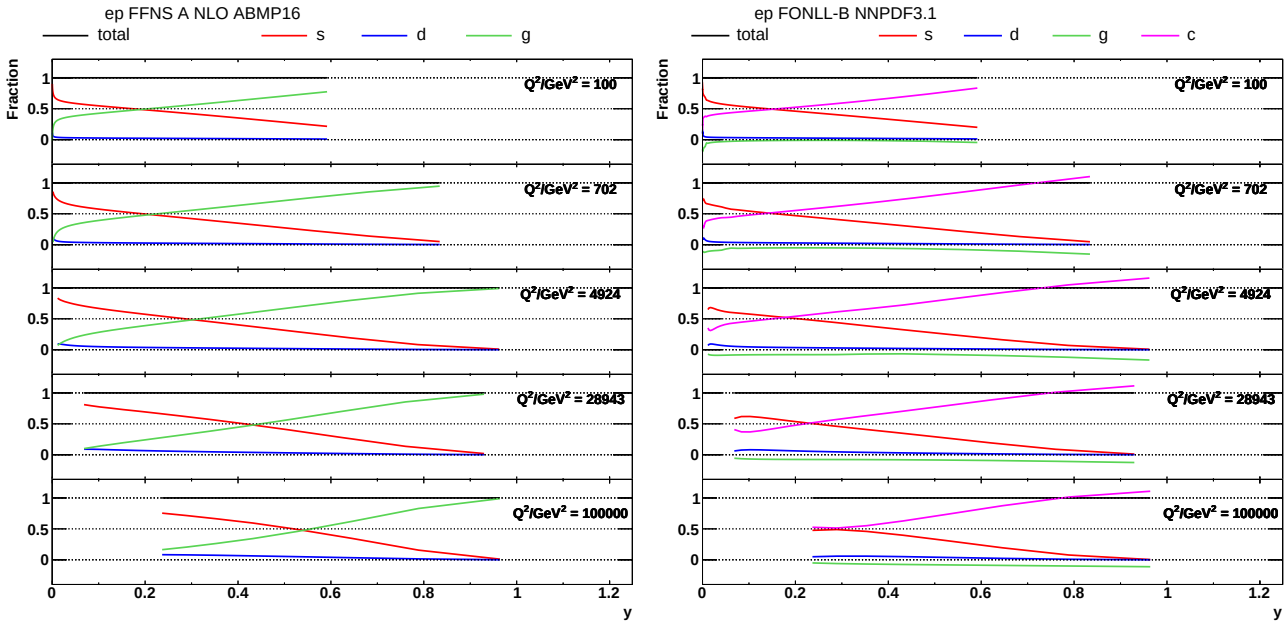

Figure 10 The partonic subprocesses for charm CC production cross sections in the FFNS A (left) and FONLL-B (right) schemes as a function of y for different values of Q^2.

integrated luminosity of 100 fb^{-1} [42, 55] and polarization $P = -0.8$ are used. Theoretical predictions are calculated at NLO in pQCD both in the FFNS A with number of active flavors $n_f = 3$ and in the FONLL-B scheme. The charm-mass reference value in the $\overline{\text{MS}}$ scheme is set to $m_c(m_c) = 1.27$ GeV and α_s is set to the value used for the corresponding PDF extraction. The renormalization and factorization scales are chosen to be $\mu_\text{r}^2 = \mu_\text{f}^2 = Q^2$.

The χ^2 value is calculated as follows:

$$\chi^2 = \mathbf{R}^T \mathbf{Cov}^{-1} \mathbf{R} + \sum_\beta b_{\beta,\text{th}}^2 \,, \qquad \mathbf{R} = \mathbf{D} - \mathbf{T} - \sum_\beta \Gamma_{\beta,\text{th}} b_{\beta,\text{th}} \,, \tag{6}$$

where $\mathbf{D}$ and $\mathbf{T}$ are the column vectors of the measured (data) and predicted (theory) values, respectively. The correlated theoretical PDF uncertainties are included using the nuisance parameters $b_{\beta,\text{th}}$ with their influence on the theory predictions described by $\Gamma_{\beta,\text{th}}$, where the index β runs over all PDF eigenvectors. For each nuisance parameter a penalty term is added to the χ^2, representing the prior knowledge of the parameter. No theoretical uncertainties except the PDF uncertainties are considered. The full covariance matrix $\mathbf{Cov}$ representing the statistical and systematic uncertainties of the data is used in the fit. The statistical and systematic uncertainties are treated as additive, *i.e.* they do not change in the fit. The systematic uncertainties are assumed uncorrelated between bins.

To treat the asymmetric PDF uncertainties of the NNPDF3.1 set, the χ^2 function in Eq. (6) is generalized assuming a parabolic dependence of the prediction on the nuisance parameters [41]:

$$\Gamma_{\beta,\text{th}} \to \Gamma_{\beta,\text{th}} + \Omega_{\beta,\text{th}} b_{\beta,\text{th}} \,, \tag{7}$$

where $\Gamma_{\beta,\text{th}} = \frac{1}{2}(\Gamma^+_{\beta,\text{th}} - \Gamma^-_{\beta,\text{th}})$ and $\Omega_\beta = \frac{1}{2}(\Gamma^+_{\beta,\text{th}} + \Gamma^-_{\beta,\text{th}})$ are determined from the shifts of the predictions corresponding to up ($\Gamma^+_{\beta,\text{th}}$) and down ($\Gamma^-_{\beta,\text{th}}$) variations of each PDF eigenvector.

The values of the nuisance parameters at the minimum, $b^{\text{min}}_{\beta,\text{th}}$, are interpreted as optimized, or profiled, PDFs, while uncertainties of $b^{\text{min}}_{\beta,\text{th}}$ determined using the tolerance criterion of $\Delta\chi^2 = 1$ correspond to the new PDF uncertainties. The profiling approach assumes that the new data are compatible with the theoretical predictions using the existing PDFs, such that no modification of the PDF fitting procedure is needed. Under this assumption, the central values of the measured cross sections are set to the central values of the theoretical predictions.

4.2 The profiled PDFs

The profiling study is performed using two sets of LHeC charm CC pseudodata:

- the full set,
- a restricted set with data points for which the difference between the FFNS A and FONLL-B are smaller than the present PDF uncertainties. The latter is taken for simplicity as the sum of the ABMP16 and NNPDF3.1 uncertainties, but for the most data points it is dominated by the NNPDF3.1 uncertainties (see Fig. 4).

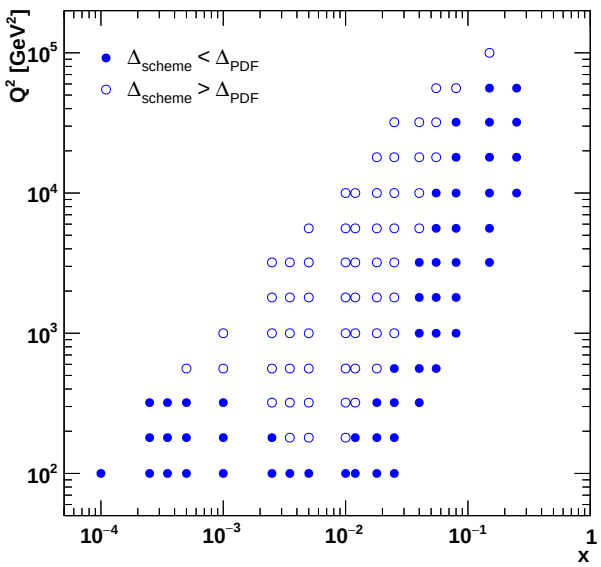

Figure 11 The full ($\Delta_{\rm scheme} < \Delta_{\rm PDF}$, $\Delta_{\rm scheme} > \Delta_{\rm PDF}$) and restricted ($\Delta_{\rm scheme} < \Delta_{\rm PDF}$) sets of data points which are used for PDF profiling.

Given the sizable differences observed between the FFNS A and FONLL-B predictions, the study with the restricted data set (also referred to as 'with cuts') aims to check whether or not model independent constraints on the strange PDF can be extracted using the charm CC reaction at LHeC. The two sets of data points are shown in Fig. 11 as functions of Q^2 and $x_{\rm Bj}$.

The original and profiled ABMP16 and NNPDF3.1 PDF uncertainties are shown in Figs. 12–15. The uncertainties of the PDFs are presented at the scales $\mu_{\rm f}^2 = 100$ GeV^2 and $\mu_{\rm f}^2 = 100000$ GeV^2. A strong impact of the charm CC pseudodata on the PDFs is observed for both PDF sets. In particular, the uncertainties of the strange PDF are strongly reduced once the pseudodata are included in the fit. Also the gluon PDF uncertainties are decreased. Furthermore, in the case of the NNPDF3.1 set, the charm PDF uncertainties are reduced significantly. For all PDF sets, only small differences can be noticed between the PDF constraints obtained using the full or restricted set because the whole $x_{\rm Bj}$ range is covered in both cases (see Fig. 11) despite the fact that the number of data points in the restricted set is roughly half of the total number of data points.

Additionally, in the case of the NNPDF3.1 set, it is possible to check the constraints on the strange quark and anti-quark distributions separately, because no assumption $s = \bar{s}$ is used in NNPDF3.1. The LHeC e^-p pseudodata provide direct constraints only on $\bar{s}$. Nevertheless due to the apparently strong correlation between s and $\bar{s}$ in the NNPDF3.1 fit, quite strong constraints are present on both the s and $\bar{s}$ distributions once the direct constraints on $\bar{s}$

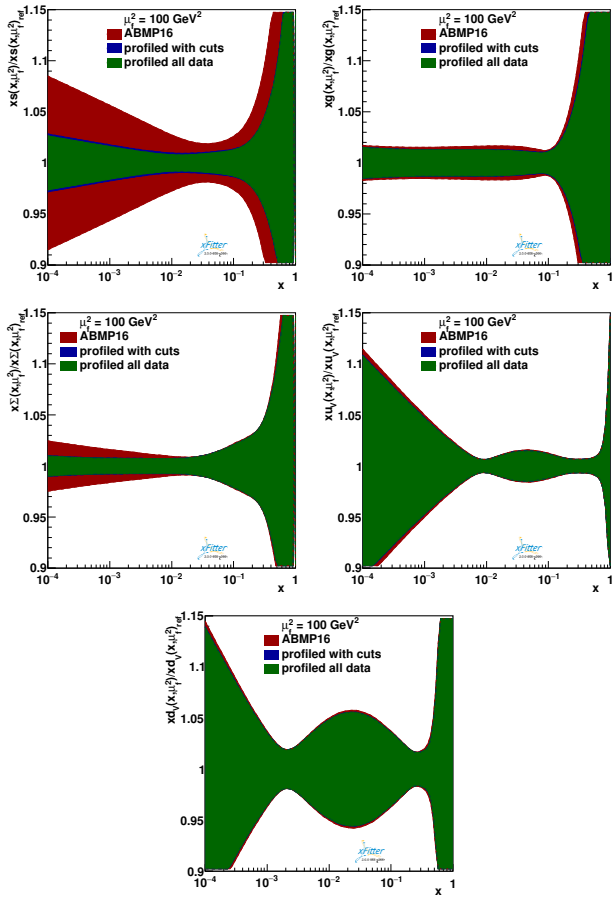

Figure 12 The relative strange (top left), gluon (top right), sea quark (middle left), u valence quark (middle right) and d valence quark (bottom) PDF uncertainties at $\mu_{\rm f}^2 = 100$ GeV^2 of the original and profiled ABMP16 PDF set.

are provided by the LHeC pseudodata. However, only mild constraints are put on the ratio $s/\bar{s}$. This indicates that for precise determination of $s/\bar{s}$ both e^-p and e^+p data will be needed.

Comparing the results of profiled PDFs in the FFNS and the VFNS, we find both analyses are able to significantly improve the constraints on the strange quark PDF. This result gives us confidence that the general features we observe here are independent of the details of the heavy flavor scheme.

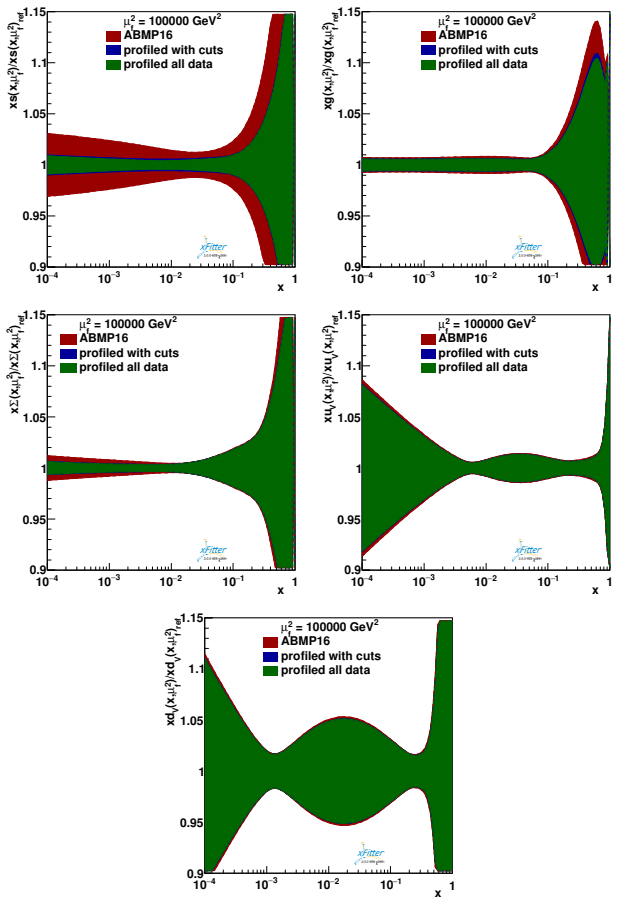

Figure 13 The relative strange (top left), gluon (top top right), sea quark (middle left), u valence quark (middle right) and d valence quark (bottom) PDF uncertainties at $\mu_f^2 = 100000$ GeV^2 of the original and profiled ABMP16 PDF set.

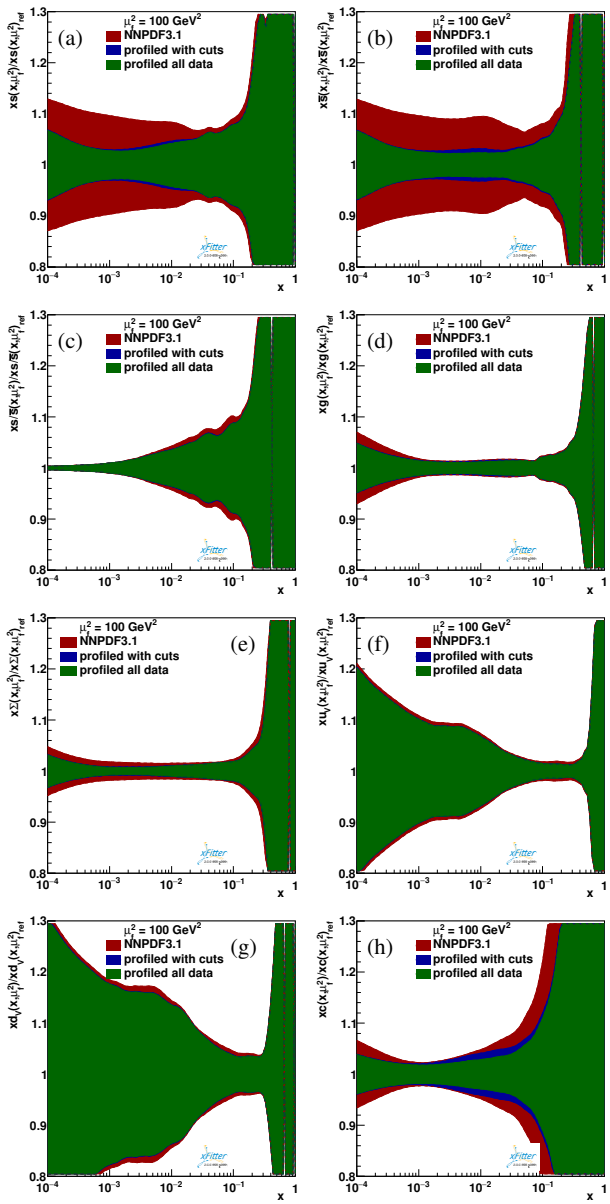

Figure 14 The relative strange quark (a), strange anti-quark (b), and ratio $s/\bar{s}$ (c), gluon (d), sea quark (e), u valence quark (f), d valence quark (g) and charm quark (h) PDF uncertainties at $\mu_f^2 = 100$ GeV^2 of the original and profiled NNPDF3.1 PDF set.

5 Discussion and summary

The recent performance of the LHC has exceeded expectations and produced an unprecedented number of precision measurements to be analyzed; thus, it is essential to improve the theoretical calculations to match. The uncertainty for many of these precision measurements stems primarily from the PDFs. Hence, our ability to measure fundamental parameters of the Standard Model (SM), such as the W boson mass and $\sin^2\theta_W$, ultimately comes down to how accurately we determine the underlying PDFs [56]. Additionally, our ability to characterize and constrain SM processes can indirectly impact beyond-standard-model (BSM) signatures.

We have focused on the strange-quark distribution which, at the LHC, can have a significant impact on the W/Z cross section: one of the "standard candle" measurements. If we can reduce the uncertainty for these predictions, we can set stringent limits on any admixture of physics at higher scales. Unfortunately, at present the strange PDF has a comparably large uncertainty because measurements from the LHC and HERA, as well as older fixed-target experiments, do not seem to provide a definitive result for this flavor component.

This situation has prompted us to examine the CC DIS charm production at the LHeC to determine the impact of this data set on the PDF uncertainty. We considered the LHeC as this high-energy ep/A facility could potentially run in parallel with the LHC and provide insights into these issues at low x and high Q^2 in advance of a FCC program.

This case study of the CC DIS charm production at the LHeC provides a practical illustration of the many features of xFitter. As the xFitter framework is designed

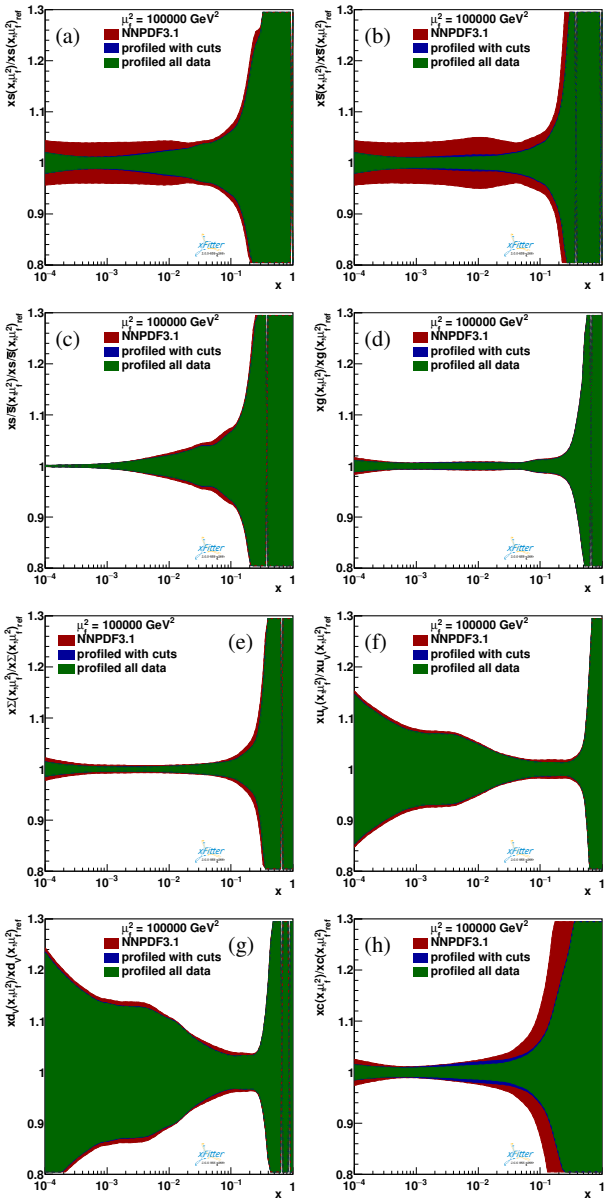

Figure 15 The relative strange quark (a), strange anti-quark (b), and ratio $s/\bar{s}$ (c), gluon (d), sea quark (e), u valence quark (f), d valence quark (g) and charm quark (h) PDF uncertainties at $\mu_f^2 = 100000\,\mathrm{GeV}^2$ of the original and profiled NNPDF3.1 PDF set.

to be a versatile open-source software framework for the determination of PDFs and the analysis of QCD physics, we can readily adapt this tool to address the impact and influence of new data sets. Furthermore, as both FFNS and VFNS calculations are implemented, we can use XFITTER as a theoretical "laboratory" to study the resummation of large logarithms and multi-scale issues. We have outlined some of these issues in the Appendix. In particular, the CC DIS charm production involves a flavor-changing $W^{\pm}$ boson, multiple quark masses enter the calculation, and this introduces some subtle theoretical issues to properly address the disparate mass and energy scales.

Using the XFITTER framework, we find that the LHeC can provide strong constraints on the strange-quark PDF, especially in the previously unexplored small-x_{Bj} region.[4] A large reduction of uncertainties is observed also when restricting the input data to the kinematic range where the differences between the FFNS A and FONLL-B schemes are not larger than the present PDF uncertainties, indicating that the obtained PDF constraints are stable and independent of the particular heavy-flavor scheme. As noted above, a reduction of the strange-PDF uncertainties influences the W/Z production, and thus the Higgs production; hence, the LHeC CC DIS charm production data represent a valuable addition for the future global PDF fits.

However, since charm CC production in e^-p collisions mostly probe $\bar{s}$, only mild constraints are put on the ratio $s/\bar{s}$ using the NNPDF3.1 PDF set as reference; therefore for a precise determination of this ratio, both e^-p and e^+p data will be needed.

In conclusion, we find that CC DIS charm production at the LHeC can provide strong constraints on the strange PDF which are complementary to the current data sets. As the PDF uncertainty is the dominant factor for many precision analyses, a reduction of these uncertainties will allow for more accurate predictions which can be used to constrain both SM and BSM physics processes.

Acknowledgements We would like to thank Max Klein for providing the pseudodata, and John C. Collins, Aleksander Kusina, Ted C. Rogers, Ingo Schienbein, George Sterman, for useful discussion. The work of O. Z. has been supported by Bundesministerium für Bildung und Forschung (contract 05H18GUCC1). This work F. O. was supported by the U.S. Department of Energy under Grant No. DE-SC0010129.

[4] In this study we have focused exclusively on the LHeC result for CC DIS charm production; however, the LHeC has a broad multi-faceted program which is described in Refs. [57, 58].

Appendix A: F_2^c Beyond leading-order

The multi-scale problem: The CC DIS charm production process involves some interesting issues that we will explore here in detail. In particular, there are multiple mass and energy scales which span a wide kinematic range, and it becomes an intricate puzzle to treat them all properly.

For this current illustration, we will focus on the contribution to the DIS F_2^c structure function from the process involving the strange and charm quark; other quark combinations can be addressed in a similar manner. The fully inclusive F_2 can be studied using the energy and angle of the outgoing lepton; in contrast, F_2^c also requires information about the final hadronic state, and this introduces some subtleties. In particular, we will show that as we go to higher orders the F_2^c structure function must be defined carefully so that: i) theoretically it is free of divergences and independent of the renormalization scales when calculated to all orders, and ii) experimentally it matches what is measured by the detector.

The mass scales: What makes this process complex is that we encounter a number of different mass scales. Furthermore, there is no fixed hierarchy for the mass scales, and we will need to compute both in the low-Q region, where $Q \lesssim m_c$, as well as in the high-Q region, $Q \gg m_c$.

The Q scale is related to the invariant mass of the virtual-boson probe (W^+ in this case), and can be expressed in terms of the energy and angle of the lepton; this is a physically measurable kinematic variable.

In contrast, the scale μ is an unphysical scale which implements the separation between the PDF and the hard-scattering cross section, and the scale at which α_s is evaluated; thus, the physics should be insensitive to a variation of μ. As our calculations typically involve the dimensionless combination $\ln(\mu/Q)$, we generally choose $\mu \sim Q$ to avoid large logarithms.

The strange quark is a "light" active parton with an associated PDF $s(x)$ and mass $m_s < \Lambda_{\rm QCD}$. The strange-quark mass is comparable to or less than other hadronic scales which are neglected; as such, it serves only as a regulator and plays no physical role. Effectively, we can take $m_s \to 0$ if we choose. We treat the up and down quarks masses $m_{u,d}$ in a similar manner.

The charm quark is a "heavy" object; its associated mass $m_c > \Lambda_{\rm QCD}$ does play a physical role and cannot generally be neglected. There may or may not be a PDF associated with the charm. In a $n_f = 3$ FFNS scheme, we will assume the charm PDF to be zero.[5] In a VFNS there is a charm PDF only when the μ scale is above the scale where the charm PDF is activated; we call this the matching scale, μ_c. It is common[6] to set $\mu_c = m_c$, but this is not required.[7] In this study, however, we will adopt this common choice.

Because there are two different quark masses involved (m_s and m_c) in the CC DIS process, we can examine the mass singularities of the t-channel and u-channel separately. This separation is particularly useful to understand how the individual mass singularities are addressed, and how the FFNS and the VFNS organize the contributions to the total structure function.

The $n_f = 3$ FFNS: To be specific, we will consider CC DIS production of a charm quark. We first compute this in the $n_f = 3$ FFNS where $\{u,d,s\}$ are light "active" partons in the proton, and the charm c is considered an external "heavy" particle. This can be implemented in the ACOT scheme [53] for example by using a CWZ renormalization [61] where the light "active" partons are renormalized with normal $\overline{\rm MS}$, and the "heavy" quarks use a zero-momentum subtraction. In this scheme, the **leading-order** (LO) process is $sW^+ \to c$ as illustrated in Fig. 16. At **next-to-leading-order** (NLO), we then include $gW^+ \to c\bar{s}$ which has both t-channel (Fig. 16) and u-channel (Fig. 17) contributions.[8]

t-channel: The t-channel process has an intermediate s-quark exchanged, and if we use the strange quark mass m_s to regulate the singularities, this will yield a contribution proportional to $\ln(Q/m_s)$. This mass singularity arises from the region of phase space where the exchanged s-quark becomes collinear and close to the mass shell; that is, when the phase space of the $gW^+ \to c\bar{s}$ process begins to overlap with that of the $sW^+ \to c$ process. This "double counting" is resolved by a **subtraction** (SUB) counter-term given by:

$$(SUB) \sim f_g \otimes \widetilde{f}_{g\to s} \otimes \sigma_{sW^+\to c}\,.$$

Here, $\widetilde{f}_{g\to s}$ is the perturbative splitting of the gluon into an $s\bar{s}$ pair; the leading term is proportional to:[9]

$$\widetilde{f}_{g\to s}(x,\mu) \sim \frac{\alpha_S(\mu)}{2\pi} P^{(1)}_{g\to s}(x)\,\ln\left(\frac{\mu^2}{m_s^2}\right) + \mathcal{O}(\alpha_s^2)$$

[5] It is possible to extend this to incorporate an intrinsic-charm PDF.

[6] The choice of matching scale $\mu_c = m_c$ is common because at NLO the $\overline{\rm MS}$ matching conditions on the PDFs are proportional to the DGLAP kernel times $\ln(\mu/m_c)$. As an explicit calculation shows, the constant term vanishes. Therefore, by choosing $\mu_c = m_c$ we have the simple boundary condition $f_c(x,\mu = m_c) = 0$. At NNLO, the constant term is non-zero and this yields $f_c(x,\mu = m_c) \neq 0$. See Ref. [59] and references therein.

[7] By displacing the matching scale to larger values $\mu_c > m_c$, one can have the advantage of avoiding delicate cancellations in the region $\mu \sim m_c$; this flexibility was explored in Refs. [51, 60].

[8] Note, there are also corresponding quark-initiated processes; we will focus on the gluon-initiated processes as this is sufficient to illustrate our points. Both the gluon- and quark-initiated contributions are included in our calculations.

[9] The scale of the SUB term is μ as the relevant scale here is the renormalization scale of the PDF: $f(x,\mu) \otimes \hat{\sigma}(x,Q,\mu)$.

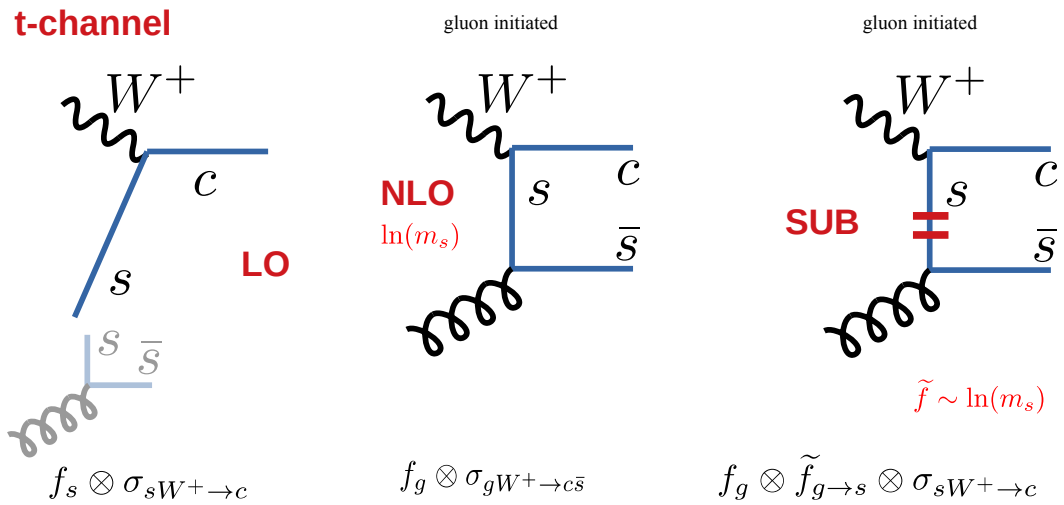

Figure 16 The t-channel processes up to $\mathcal{O}(\alpha_S^1)$. Note we sum the combination (NLO−SUB) to obtain the complete $\mathcal{O}(\alpha_S^1)$ correction; we find it useful to study these terms separately. The higher-order quark-initiated contributions are not show, but are included in the calculation.

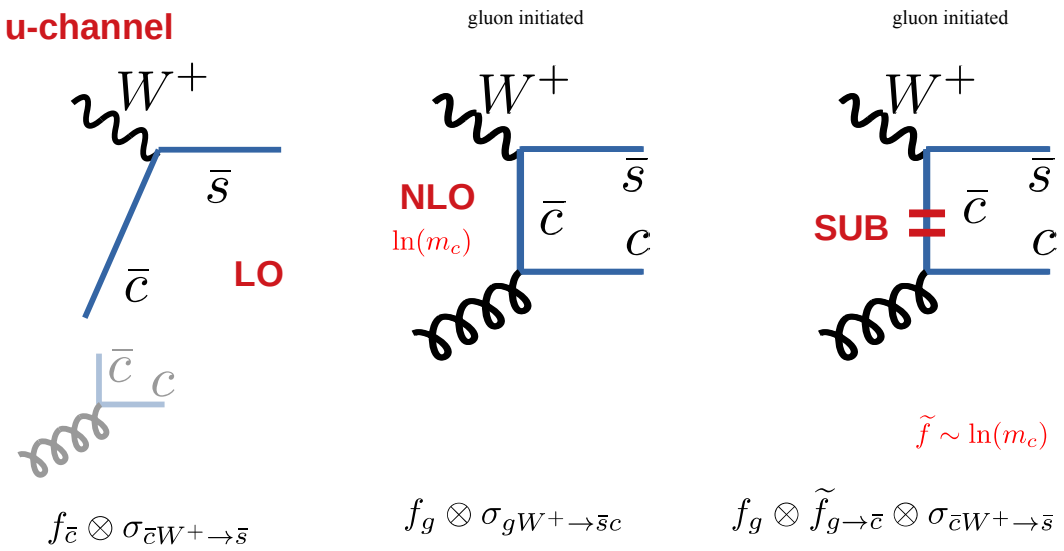

Figure 17 The u-channel processes up to $\mathcal{O}(\alpha_S^1)$. Note the NLO t-channel and u-channel terms are combined coherently at the amplitude level. The higher-order quark-initiated contributions are not show, but are included in the calculation.

where $P^{(1)}_{g\to s}(x)$ is the $\mathcal{O}(\alpha_s)$ DGLAP splitting kernel for $g \to s$.

The complete contribution to the structure function is given by:

$$F_2^c \sim TOT = LO + (NLO - SUB)$$

The complete $\mathcal{O}(\alpha_s)$ contribution is the combination $(NLO - SUB)$; our separation into NLO and SUB is simply to illustrate the interplay of these components. Both the NLO and SUB terms have $\ln(m_s)$ divergences, but these precisely cancel and yield a well-defined result even if we take the $m_s \to 0$ limit.[10]

[10]In fact, we could have taken $m_s = 0$ initially and used dimensional regularization to compute the contributions.

u-channel: We next examine the u-channel NLO contribution to the $gW^+ \to c\bar{s}$ process. This has an intermediate c-quark exchanged and is proportional to $\ln(Q/m_c)$. In the FFNS where the charm is a "heavy" non-parton, there is no counter-term for this graph, and the resulting observables will retain the $\ln(Q/m_c)$ dependence. In principle, this means that when we go to large Q scales, these terms will begin to degrade the convergence of the perturbative series. In practice, while this degradation only grows logarithmically, at large scales (such as at the LHC energies) we do find it convenient to treat the charm on an equal footing as the u, d, s partons.

The VFNS: We now turn to the VFNS scheme where we include the charm quark as an "active" parton and compute its associated PDF.

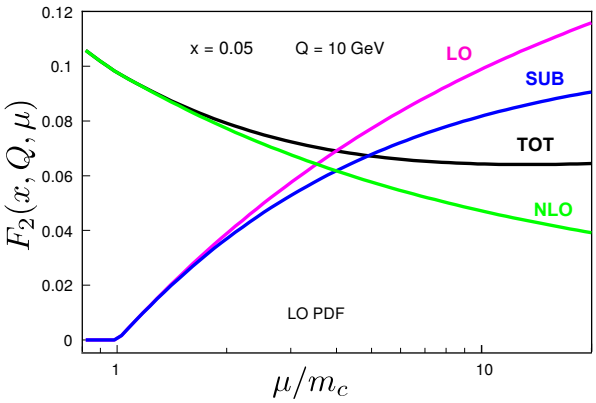

Figure 18 Calculation of F_2^c vs. μ in the VFNS illustrating the cancellation of the LO ($\bar{c}W^+ \to \bar{s}$) and the SUB ($g \to \bar{c}$)$\otimes$($\bar{c}W^+ \to \bar{s}$) contributions in the region $\mu \sim m_c$. The Q scale is fixed at 10 GeV and the charm PDF is matched at $\mu_c = m_c$ such that $f_c(x, \mu = m_c) = 0$.

In this case, there is a u-channel counter-term (SUB) given by $f_g \otimes \widetilde{f}_{g\to\bar{c}} \otimes \sigma_{\bar{c}W^+\to\bar{s}}$ which is proportional to $\ln(\mu/m_c)$. The NLO u-channel contribution will have a $\ln(Q/m_c)$ factor, so the combination $(NLO - SUB)$ is also free of mass singularities.[11]

What is less obvious is that we must also include the LO process $\bar{c}W^+ \to \bar{s}$. There are two ways we can understand why this is necessary.

Explanation #1: matching of LO and SUB: Recall that in the t-channel case, the subtraction term SUB removed the double counting between the LO $sW^+ \to c$ and NLO $gW^+ \to c\bar{s}$ subprocesses.

The u-channel case is analogous in that this subtraction term removes the double counting between the LO $\bar{c}W^+ \to \bar{s}$ and NLO $gW^+ \to c\bar{s}$ subprocesses; both contributions are required to ensure that the resulting cross section is insensitive to the scale μ.

This is apparent in Fig. 18 where we plot the individual terms versus μ for fixed values of $x_{\rm Bj}$ and Q. In the region $\mu \sim m_c$, the charm PDF $f_c(x,\mu)$ (and hence, the LO contribution) rises very quickly as the DGLAP evolution is driven by the very large gluon distribution via $g \to c\bar{c}$ splitting, and combined with a large $\alpha_s(\mu)$. The SUB subtraction also rises quickly as this is driven by the logarithmic term $\ln(\mu^2/m_c^2)$. The difference $(LO - SUB)$ is the physical contribution to the total $[TOT = LO + NLO - SUB]$, and it is this combination that is smooth across the "turn on" of the charm PDF at the matching scale $\mu_c = m_c$. We now see that if we neglect the LO ($\bar{c}W^+ \to \bar{s}$) contribution, we lose the cancellation between LO and SUB in the region $\mu \sim m_c$, and our structure function (or cross section) would have an anomalous shift at the arbitrarily location (μ_c) where we turn on the charm PDF.

As we vary the unphysical scale μ, we are simply shifting contributions between the separate $\{LO, NLO, SUB\}$ terms which individually exhibit a large μ-dependence. However, the total combination (TOT), which represents the physical observable, is relatively insensitive to μ (up to higher orders), and this property is evident in Fig. 18.

Explanation #2: removing "double counting:" A second way to understand why we require the LO process $\bar{c}W^+ \to \bar{s}$ is to consider the regions of phase space covered by each of the subprocesses. The singularity of the u-channel NLO $gW^+ \to c\bar{s}$ processes arises from the phase-space region where the intermediate $\bar{c}$-quark becomes collinear and close to the mass shell.[12] This is precisely the phase-space region of the LO process $\bar{c}W^+ \to \bar{s}$ where the partonic $\bar{c}$-quark is collinear to the hadron. The SUB term then removes the "double counting" between the LO and NLO contributions; hence, all three contributions $\{LO, NLO, SUB\}$ are necessary to cover the full phase space.

This is also apparent if we consider the transverse momentum (p_T) of the final-state charm in the Breit frame. For the LO $\bar{c}W^+ \to \bar{s}$ process in the Breit frame, the incoming W^+ and $\bar{c}$ are collinear, and the produced $\bar{s}$ must have zero p_T in this frame.

For the NLO $gW^+ \to c\bar{s}$ process, we integrate over the complete phase space for the exchanged $\bar{c}$ quark, and this will include the region where the $\bar{c}$-quark is emitted nearly collinear to the gluon and nearly on-shell; in this region the $\bar{c}$-quark will have $p_T \sim 0$ and we encounter a singularity from the internal $\bar{c}$-quark propagator. The $p_T \sim 0$ region is precisely that subtracted by the SUB counter term[13] and this ensures that the combination $(NLO - SUB)$ is free of divergences.

Recap: To recap, i) the combination of the LO and SUB terms ensure a minimal μ variation at low μ, and ii) the combination of SUB and NLO ensures that the mass singularities are cancelled at high μ.

This interplay of terms illustrates some of the intricacies of QCD, especially since this exchange is across different orders of α_s.

Furthermore, note that in the u-channel for both the LO and SUB contributions, the charm quark is collinear to the incoming hadron, and thus exits in the hadron remnants. While this may be experimentally difficult to observe, because we are asking for a *"fully inclusive"* F_2^c, these

[11] Specifically, the combination $(NLO - SUB)$ is free of mass singularities and finite in the limit $m_c \to 0$. Note that the VFNS fully retains the charm quark mass m_c and (in contrast to some claims in the literature) the factorization holds up to $\mathcal{O}(\Lambda^2/Q^2)$ corrections; all terms of order (m_c^2/Q^2) are fully included [62].

[12] For example, the c-quark is off-shell by the order of its mass m_c; this is independent of the scale Q and *does not* assume any $Q \gg m_c$ limit.

[13] Specifically, the incoming W^+ and g are collinear and the gluon then emits a collinear $c\bar{c}$ pair so the final $\bar{s}$ has zero p_T.

contributions cannot be simply ignored. We will discuss this further in the following section.

Defining F_2^c: The LO u-channel $\bar{c}W^+ \to \bar{s}$ process foreshadows difficulties that we encounter if we try and extend the concept of "fully inclusive" F_2^c to higher orders. We note that in Ref. [62] Collins extended the proof of factorization to include heavy quarks such as charm and bottom for an inclusive structure function F_2; analysis of a *"fully inclusive"* F_2^c is more complex for a number of reasons. Whereas F_2 only requires measurement of the outgoing lepton energy and angle, F_2^c also requires information on the hadronic final state. At the parton level, this introduces complications including when the charm is in the hadronic remnants and brings in both fragmentation and fracture functions.

To characterize the theoretical issues involved in constructing F_2^c, we can imagine starting from the (well-defined) inclusive F_2, and then dividing the contributions into two sets: one for F_2^c for the "heavy" charm quark, and the rest into $F_2^{u,d,s}$ for the "light" quarks. We will show that this theoretical procedure encounters ambiguities.

The LO u-channel $\bar{c}W^+ \to \bar{s}$ process does not have any "apparent" charm quark in the final state, but this contribution is essential to balance with the SUB process $f_g \otimes \widetilde{f}_{g\to\bar{c}} \otimes \sigma_{\bar{c}W^+\to\bar{s}}$. Note that for the SUB process the charm quark arises from a gluon splitting into a collinear $c\bar{c}$ pair which is then part of the hadron remnants. For the LO process, presumably our $\bar{c}$ quark also came from a gluon splitting into a collinear $c\bar{c}$ pair. Thus, our F_2^c must include those cases where the charm is contained in the hadron remnants.

This issues touches on the fact that, because the charm parton ultimately fragments into a charmed hadron (typically a D meson), we must introduce a set of fragmentation functions (FFs) which are scale-dependent and will factorize final-state singularities in a similar manner as the PDFs factor the initial-state singularities.[14] Specifically, we may also allow for the possibility that a gluon or a light quark fragments into a charmed hadron.

The bubble diagram: Some of the theoretical intricacies of defining a *"fully inclusive"* F_2^c are illustrated in Fig. 19 which shows a higher-order DIS process with a quark-antiquark loop.

Let us compute this diagram in the $n_f = 3$ FFNS where the internal loop is a massive $c\bar{c}$-pair and the external quark is a light quark $\{u,d,s\}$. If the final state is represented by Cut-A, then we have charm quarks in the final state, and this should be included in F_2^c.

However, if we instead use Cut-B as a final state, there is no charm in the final state, so this should not be included

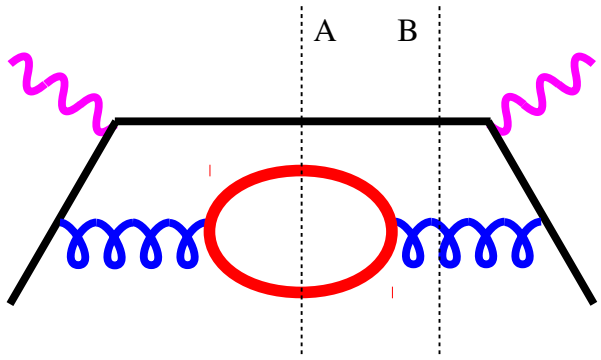

Figure 19 A higher order Feynman graph illustrating the complications in defining a "fully inclusive" F_2^{charm}. A light quark (q) scatters from a vector boson (V) with a $c\bar{c}$ in the internal loop. If we cut the amplitude at "A" we have charm in the final state and this must be included in F_2^{charm}. If we cut the amplitude with cut "B" there is no charm in the final state. Additionally, since this diagram contributes to the beta function, this highlights the complications of using an α_S and hard scattering $\hat{\sigma}$ with differing $N_{\rm eff}$.

in F_2^c. [More precisely, when we renormalize the charm loop with zero-momentum subtraction, this contribution effectively decouples.] Thus, the contribution from Cut-A will be included in F_2^c, but the contribution from Cut-B will not.

This diagram generates additional complications in that multiple quark flavors are involved. For example, the bubble diagram involves quarks of both $q = \{u,d,s\}$ and c flavors, so this contribution cannot be uniquely assigned to F_2^q or F_2^c. We can introduce theoretical definitions to make the choice, but then we have to be careful about double-counting contributions and introducing uncancelled singularities. For example, the bubble diagram of Fig. 19 is encountered in the F_2^c heavy-quark calculations of Refs. [63, 64]; here, an additional scale Δ is introduced to subdivide the contributions.

The running of α_s in the FFNS: The bubble diagram of Fig. 19 also highlights the difficulty of using a $n_f = 3$ FFNS with a VFNS running of α_s. In a $n_f = 3$ FFNS, internal $c\bar{c}$ loops decouple from the theory and are not included in the calculation;[15] however, the β-function with $n_f = 4$ requires precisely these $c\bar{c}$ loop contributions. This deficiency can be patched order by order by expanding the β-function and inserting the required terms at each order [65–67]. Once again, we cannot unambiguously divide the inclusive F_2 into separate "light" and "heavy" quantities.

Extensions to bottom and top: While we have used the charm quark to illustrate these features, the same properties can, in principle, be applied to both the bottom and top

[14] For the NLO quark-initiated contributions (not shown) we will have final state singularities from processes such as $c \to cg$ which will be factorized into the FFs.

[15] More precisely, the heavy quarks are renormalized with zero-momentum subtraction and their contributions decouple; this is why we can neglect loops from the the top quark and any other heavy particle.

quark.[16] For the case of the bottom quark, the larger mass m_b yields a smaller $\alpha_s(\mu)$ for $\mu \sim m_b$ and the evolution of $f_b(x,\mu)$ is thus reduced. Nevertheless, for large-scale processes (such as at the LHC) we often find it convenient to make use of $f_b(x,\mu)$ and treat the bottom on an equal footing as the other light quarks. For the case of the top quark, the very large mass m_t yields a much smaller $\alpha_s(\mu)$ for $\mu \sim m_t$ and the evolution of $f_t(x,\mu)$ is comparatively reduced.

Summary To properly define F_2^c at higher orders, we encounter the theoretical issues discussed above: as the charm quark fragments into a charmed meson, we must be careful to ensure that the theoretical quantity matches what is actually measured experimentally. This is more complex than simply asking for the portion of F_2 has has a charm in the final state, and is an issue for both the FFNS and VFNS as we move to higher orders. We can perform the computation in the FFNS but in the large energy limit we encounter $\ln(Q^2/m_c^2)$ divergences and this, in part, contributes to the observed differences at large Q.

The VFNS includes the charm quark as an active parton for μ scales above a matching scale μ_c. For large Q scales, the mass singularities of NLO and SUB terms will cancel to yield a result free of divergences. For scales $\mu \sim m_c$, cancellation between the LO and SUB contributions ensures a minimal μ dependence; however, as this can be delicate to implement numerically, we have the option of displacing the matching scale μ_c to a larger scale where the cancellation is more stable [51, 60].

References

1. H. Abdolmaleki, et al., Eur. Phys. J. **C78**(8), 621 (2018). DOI 10.1140/epjc/s10052-018-6090-8
2. H. Abramowicz, et al., Eur. Phys. J. **C75**(12), 580 (2015). DOI 10.1140/epjc/s10052-015-3710-4
3. J. Gao, L. Harland-Lang, J. Rojo, Phys. Rept. **742**, 1 (2018). DOI 10.1016/j.physrep.2018.03.002
4. S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, Phys. Rev. **D96**(1), 014011 (2017). DOI 10.1103/PhysRevD.96.014011
5. R.D. Ball, et al., Eur. Phys. J. **C77**(10), 663 (2017). DOI 10.1140/epjc/s10052-017-5199-5
6. O. Behnke, A. Geiser, M. Lisovyi, Prog. Part. Nucl. Phys. **84**, 1 (2015). DOI 10.1016/j.ppnp.2015.06.002
7. O. Zenaiev, Eur. Phys. J. **C77**(3), 151 (2017). DOI 10.3204/PUBDB-2017-01474,10.1140/epjc/s10052-017-4620-4
8. H. Abdolmaleki, A. Khorramian, A. Aleedaneshvar, Nucl. Part. Phys. Proc. **282-284**, 27 (2017). DOI 10.1016/j.nuclphysbps.2016.12.006
9. H. Abdolmaleki, A. Khorramian, arXiv:1903.02583, accepted by Phys. Rev. **D** (2019)
10. E. Laenen, S. Riemersma, J. Smith, W.L. van Neerven, Nucl. Phys. **B392**, 162 (1993). DOI 10.1016/0550-3213(93)90201-Y
11. E. Laenen, S. Riemersma, J. Smith, W.L. van Neerven, Nucl. Phys. **B392**, 229 (1993). DOI 10.1016/0550-3213(93)90202-Z
12. T. Gottschalk, Phys. Rev. **D23**, 56 (1981). DOI 10.1103/PhysRevD.23.56
13. M. Gluck, S. Kretzer, E. Reya, Phys. Lett. **B398**, 381 (1997). DOI 10.1016/S0370-2693(97)90016-2, 10.1016/S0370-2693(97)00232-3. [Erratum: Phys. Lett.B405,392(1997)]
14. J. Blümlein, A. Hasselhuhn, P. Kovacikova, S. Moch, Phys. Lett. **B700**, 294 (2011). DOI 10.1016/j.physletb.2011.05.007
15. M. Buza, W.L. van Neerven, Nucl. Phys. **B500**, 301 (1997). DOI 10.1016/S0550-3213(97)00327-1
16. J. Blümlein, A. Hasselhuhn, T. Pfoh, Nucl. Phys. **B881**, 1 (2014). DOI 10.1016/j.nuclphysb.2014.01.023
17. E.L. Berger, J. Gao, C.S. Li, Z.L. Liu, H.X. Zhu, Phys. Rev. Lett. **116**(21), 212002 (2016). DOI 10.1103/PhysRevLett.116.212002
18. A. Behring, J. Blümlein, A. De Freitas, A. Hasselhuhn, A. von Manteuffel, C. Schneider, Phys. Rev. **D92**(11), 114005 (2015). DOI 10.1103/PhysRevD.92.114005
19. T. Aaltonen, et al., Phys. Rev. Lett. **100**, 091803 (2008). DOI 10.1103/PhysRevLett.100.091803
20. V.M. Abazov, et al., Phys. Lett. **B666**, 23 (2008). DOI 10.1016/j.physletb.2008.06.067
21. V.M. Abazov, et al., Phys. Lett. **B743**, 6 (2015). DOI 10.1016/j.physletb.2015.02.012
22. S. Chatrchyan, et al., JHEP **02**, 013 (2014). DOI 10.1007/JHEP02(2014)013
23. G. Aad, et al., JHEP **05**, 068 (2014). DOI 10.1007/JHEP05(2014)068
24. A.M. Sirunyan, et al., Eur. Phys. J. **C79**(3), 269 (2019). DOI 10.1140/epjc/s10052-019-6752-1
25. R. Abdul Khalek, S. Bailey, J. Gao, L. Harland-Lang, J. Rojo, (2019)
26. G. Kramer, B. Lampe, H. Spiesberger, Z. Phys. **C72**, 99 (1996). DOI 10.1007/s002880050228
27. W.G. Seligman, et al., Phys. Rev. Lett. **79**, 1213 (1997). DOI 10.1103/PhysRevLett.79.1213
28. M. Tzanov, et al., Phys. Rev. **D74**, 012008 (2006). DOI 10.1103/PhysRevD.74.012008
29. G. Onengut, et al., Phys. Lett. **B632**, 65 (2006). DOI 10.1016/j.physletb.2005.10.062

[16] Additionally, Collins definitively addressed the case of multiple heavy quarks which can allow for both charm and bottom in a unified framework; in contrast to some incorrect claims in the literature, there is no difficulty in including multiple heavy quarks. (*cf.* Ref. [62], Sec. IX.)

30. J.P. Berge, et al., Z. Phys. **C49**, 187 (1991). DOI 10.1007/BF01555493
31. O. Samoylov, et al., Nucl. Phys. **B876**, 339 (2013). DOI 10.1016/j.nuclphysb.2013.08.021
32. A. Airapetian, et al., Phys. Lett. **B666**, 446 (2008). DOI 10.1016/j.physletb.2008.07.090
33. I. Abt, et al., JHEP **05**, 201 (2019). DOI 10.1007/JHEP05(2019)201
34. H.L. Lai, P.M. Nadolsky, J. Pumplin, D. Stump, W.K. Tung, C.P. Yuan, JHEP **04**, 089 (2007). DOI 10.1088/1126-6708/2007/04/089
35. T. Aaltonen, et al., Phys. Rev. Lett. **110**(7), 071801 (2013). DOI 10.1103/PhysRevLett.110.071801
36. T.A. Aaltonen, et al., Phys. Rev. **D93**(5), 052012 (2016). DOI 10.1103/PhysRevD.93.052012
37. S. Kretzer, H.L. Lai, F.I. Olness, W.K. Tung, Phys. Rev. **D69**, 114005 (2004). DOI 10.1103/PhysRevD.69.114005
38. A.D. Martin, R.G. Roberts, W.J. Stirling, R.S. Thorne, Phys. Lett. **B604**, 61 (2004). DOI 10.1016/j.physletb.2004.10.040
39. M. Aaboud, et al., Eur. Phys. J. **C77**(6), 367 (2017). DOI 10.1140/epjc/s10052-017-4911-9
40. A.M. Cooper-Sarkar, K. Wichmann, Phys. Rev. **D98**(1), 014027 (2018). DOI 10.1103/PhysRevD.98.014027
41. S. Alekhin, et al., Eur. Phys. J. **C75**(7), 304 (2015). DOI 10.1140/epjc/s10052-015-3480-z
42. J.L. Abelleira Fernandez, et al., J. Phys. **G39**, 075001 (2012). DOI 10.1088/0954-3899/39/7/075001
43. S. Alekhin, J. Blümlein, S. Moch, Phys. Rev. **D86**, 054009 (2012). DOI 10.1103/PhysRevD.86.054009
44. S. Alekhin, J. Blümlein, L. Caminadac, K. Lipka, K. Lohwasser, S. Moch, R. Petti, R. Placakyte, Phys. Rev. **D91**(9), 094002 (2015). DOI 10.1103/PhysRevD.91.094002
45. S. Alekhin, J. Blümlein, S. Moch, Eur. Phys. J. **C78**(6), 477 (2018). DOI 10.1140/epjc/s10052-018-5947-1
46. S. Alekhin. "OPENQCDRAD". URL http://www-zeuthen.desy.de/~alekhin/OPENQCDRAD/
47. S. Forte, E. Laenen, P. Nason, J. Rojo, Nucl. Phys. **B834**, 116 (2010). DOI 10.1016/j.nuclphysb.2010.03.014
48. V. Bertone, S. Carrazza, J. Rojo, Comput. Phys. Commun. **185**, 1647 (2014). DOI 10.1016/j.cpc.2014.03.007
49. A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, G. Watt, Eur. Phys. J. **C75**, 132 (2015). DOI 10.1140/epjc/s10052-015-3318-8
50. M. Tanabashi, et al., Phys. Rev. **D98**(3), 030001 (2018). DOI 10.1103/PhysRevD.98.030001
51. V. Bertone, et al., Eur. Phys. J. **C77**(12), 837 (2017). DOI 10.1140/epjc/s10052-017-5407-3
52. A. Kusina, F.I. Olness, I. Schienbein, T. Jezo, K. Kovarik, T. Stavreva, J.Y. Yu, Phys. Rev. **D88**(7), 074032 (2013). DOI 10.1103/PhysRevD.88.074032
53. M.A.G. Aivazis, J.C. Collins, F.I. Olness, W.K. Tung, Phys. Rev. **D50**, 3102 (1994). DOI 10.1103/PhysRevD.50.3102
54. H. Paukkunen, P. Zurita, JHEP **12**, 100 (2014). DOI 10.1007/JHEP12(2014)100
55. J. Blümlein, M. Klein, Nucl. Instrum. Meth. **A329**, 112 (1993). DOI 10.1016/0168-9002(93)90925-8
56. ATLAS, C. Collaborations, in *HL/HE-LHC Physics Workshop: final jamboree Geneva, CERN, March 1, 2019* (2019)
57. J.L. Abelleira Fernandez, et al., arXiv:1211.5102 (2012)
58. M. Klein, in *From My Vast Repertoire ...: Guido Altarelli's Legacy*, ed. by A. Levy, S. Forte, G. Ridolfi (2019), pp. 303–347. DOI 10.1142/9789813238053_0015
59. T. Stavreva, F.I. Olness, I. Schienbein, T. Jezo, A. Kusina, K. Kovarik, J.Y. Yu, Phys. Rev. **D85**, 114014 (2012). DOI 10.1103/PhysRevD.85.114014
60. V. Bertone, et al., PoS **DIS2018**, 015 (2018). DOI 10.22323/1.316.0015
61. J.C. Collins, F. Wilczek, A. Zee, Phys. Rev. **D18**, 242 (1978). DOI 10.1103/PhysRevD.18.242
62. J.C. Collins, Phys. Rev. **D58**, 094002 (1998). DOI 10.1103/PhysRevD.58.094002
63. A. Chuvakin, J. Smith, W.L. van Neerven, Phys. Rev. **D61**, 096004 (2000). DOI 10.1103/PhysRevD.61.096004
64. A. Chuvakin, J. Smith, W.L. van Neerven, Phys. Rev. **D62**, 036004 (2000). DOI 10.1103/PhysRevD.62.036004
65. D. Napoletano, Istituto Nazionale di Fisica Nucleare | Sezione di Milano (2014)
66. I. Bierenbaum, J. Blümlein, S. Klein, Phys. Lett. **B672**, 401 (2009). DOI 10.1016/j.physletb.2009.01.057
67. F. Cascioli, P. Maierhöfer, N. Moretti, S. Pozzorini, F. Siegert, Phys. Lett. **B734**, 210 (2014). DOI 10.1016/j.physletb.2014.05.040