Open Charm and Beauty Production at HERA

Olaf Behnke for the ZEUS and H1 Collaborations

DESY, Notkestrasse 85, 22607 Hamburg, Germany

Abstract

A review is provided of open charm and beauty production at HERA and its description by perturbative QCD (pQCD). Four years after the end of the data taking there is still a steady flow of new charm and beauty results from HERA. Among the results reported here are the first combined H1 and ZEUS measurements on the contribution from charm production to deep inelastic scattering (DIS), represented by the structure function $F_{cc}$, as well as new precise results on the corresponding structure function for beauty production, $F_{bb}$. Furthermore the situation of charm and beauty production in the photoproduction kinematic regime is reviewed. Since it is a related field also the first hadroproduction results from LHC are presented. A brief outlook is given on open heavy flavour prospects at possible future $ep$ colliders, with a focus on the LHeC.

Keywords: charm, beauty, perturbative QCD, photoproduction, deep inelastic scattering

1. Introduction

Heavy quark production at HERA provides an exciting testing ground for perturbative QCD (pQCD). In leading order, heavy quarks are produced in $ep$ collisions via the Boson Gluon Fusion (BGF) process shown in Fig. 1 on the left. This process provides direct access to the gluon density in the proton. BGF type processes dominate DIS scattering towards lower values of the Bjorken scaling variable $x$, due to the large gluon density. In the limit of large photon virtualities $Q^2$, the events with charm and beauty quarks are expected to account for $\sim 36\%$ and $\sim 9\%$ of the BGF processes and hence contribute significantly to inclusive DIS. On the theoretical side, the description of heavy quark production in the framework of perturbative QCD is complicated due to the presence of several large scales like the heavy quark masses, the transverse momentum $p_T$ of the produced quarks and $Q^2$. Different calculation schemes have been developed to obtain predictions from pQCD. At low scales $p_T$ (or $Q^2$) the fixed-flavour number scheme (FFNS) [1] is expected to be most appropriate where the quark masses are fully accounted for. Calculation programs [2, 3] are available to Next-to-Leading Order (NLO), which is order $\mathcal{O}(\alpha_s^3)$ for the cross sections. An exemplary NLO diagram is shown in Fig. 2 on the left. Complete Next-to-Next-to-Leading Order (NNLO) predictions are not yet available due to difficulties to determine two loop diagrams with heavy quark lines (see Fig. 2 right). However, some important steps have been already undertaken towards an NNLO calculation, including thresh-
old resummation of soft gluon radiation, as discussed in
the talk by S. Alekhin [4]. At very high scales the NLO

\[ \frac{e^+}{p} \gamma \sqrt{\alpha_s} g e + \frac{e^+}{p} \gamma \sqrt{\alpha_s} g \]

FFNS scheme predictions are expected to break down
since large logarithms \( \ln(p_T^2/m^2) \) are neglected that re-
present collinear gluon radiations from the heavy quark
lines. These logarithms can be resummed to all orders
in the alternative Zero-Mass Variable Flavour Number
schemes (ZM-VFNS) [5]. Here the charm and beauty
quarks are treated above kinematic threshold as mass-
less and appear also as active sea quarks in the pro-
ton and in the resolved photon, as depicted in figure 1
middle and right. Most widely used are nowadays the so-called Generalised Variable Flavour Number
Schemes (GM-VFNS) [6]. These mixed schemes converge to the massive and massless schemes at low
and high kinematical scales, respectively, and apply a
suitable interpolation in the intermediate region. How-
ever, the exact modelling of the interpolation and in
general the treatment of mass dependent terms in the
perturbation series are still a highly controversial issue
among the various theory groups. The different treat-
ments have profound implications for global Parton
Distribution Function (PDF) fits and influence the fitted
densities of gluons and other quark flavours in the pro-
ton. This topic is discussed extensively in the talk by K.
Lipka [7] as well as the implications for predictions of
many important processes at the LHC, for instance for
Z and W production. As it turns out the HERA inclu-
sive charm production data in DIS allow to determine
the charm quark mass parameter in the different scheme
calculations and this helps to stabilise the PDF fits and predictions for the LHC.

In the review presented here the focus is on compar-
ing the HERA open charm and beauty production data
with pQCD calculations in the various schemes, to see
how well they are doing. A systematic “tour through
the hard scales” is taken, as indicated in Fig. 3, start-
ing with the smallest quark mass and photon virtuality
scales, charm in photoproduction and ending with the
largest scales, beauty production in DIS. In the outlook

Also some prospects for heavy quark production at pos-
sible future \( ep \) colliders are briefly given.

2. Results

2.1. Charm photoproduction

One of the most precise measurements of charm
photoproduction to date are provided by the ZEUS
\( D^* \) meson analysis [8] based on the HERA I data
set. The \( D^* \) mesons are identified via full reconstruc-
tion of the decay \( D^{*+} \rightarrow D^0 \pi^+ \) with subsequent decay
\( D^0 \rightarrow K^-\pi^+ \) or \( D^0 \rightarrow K^-\pi^+\pi^+\pi^- \). Throughout this
review the charge conjugated states and decay chains are
implicitly included. Figure 4 shows the obtained dif-
ferential production cross sections as a function of the
transverse momentum of the \( D^* \) meson. A vast region
from 1.9 GeV to 20 GeV is covered, where the data are
falling over four orders of magnitude. The ZEUS data
are compared to the FMNR [2] massive scheme NLO
calculations and to the FONLL [9] predictions which
are based on the massive scheme at NLO plus a next-
to-leading log resummation of terms of collinear ori-
gin. The FMNR prediction describes the data better than
FONLL towards the high transverse momenta. This is a
surprise since this is the region where one might expect
that the massive scheme starts to fail. At the starting
point of the spectrum, where the transverse momenta
are not much above the charm quark mass, both calcu-
lations are slightly below the data, however still within

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{mQ} & \textbf{Photoproduction} \( Q^2 \approx 0 \) & \textbf{DIS} \( Q^2 \geq 1 \text{ GeV}^2 \) \\
\hline
\textbf{c} & 1.5 GeV & 1. 3. \\
\textbf{b} & 4.75 GeV & 2. 4. \\
\hline
\end{tabular}
\caption{Ordering scheme of topics in this review, according to the
hard scales of quark mass and photon virtuality \( Q^2 \).}
\end{table}
reasonable agreement considering the estimated prediction uncertainties. These uncertainties have been obtained by varying the renormalisation and factorisation scales simultaneously by the factors two up and down from the nominal scale $\mu_{r,f} = \sqrt{m^2 + p^2_t}$. They reach a level of about 50% at small transverse momenta. This is no surprise, since $\alpha_s$, which enters the cross section prediction already at leading order (see Fig. 1 left), is large for the relatively small hard scales available in this kinematic domain. Thus the predictive power of the NLO calculations remains limited in the kinematic threshold region. The situation will only change when NNLO predictions will become available, which unfortunately cannot be expected for the near future.

Figure 5 shows the results of a similar but more recent $D^*$ meson measurement [10] from H1, in this case as a function of the pseudorapidity of the meson. The data are compared to different massive NLO plus matched parton shower calculations, MC@NLO [12] and POWHEG [13]. The spectrum itself and the quality of the description by the calculations bear a striking resemblance to the ZEUS photoproduction results shown in Fig. 4. Also for pp collisions the predictions tend to undershoot the data, but are still in reasonable agreement considering the theory uncertainties. These uncertainties are considerably larger compared to those at HERA, which demonstrates the virtue of having at HERA an electromagnetic photon probe at hand. In the future one can expect the LHC experiments to extend the phase space of charm production to much higher

![Figure 5: Differential $D^*$ photoproduction cross sections as a function of the pseudorapidity of the meson. The H1 measurement [10] is compared to the generalised variable flavour number scheme calculation (labelled GMVFNS), for more details of the calculation see [10].](image)

on $D^*$ meson differential production cross sections as a function of the meson transverse momentum. The data are compared to different massive NLO plus
transverse momenta than those covered in Fig. 6. It will be interesting to see how well the various theory models can describe this large momentum domain.

The dominant production mechanism for charm in photoproduction at HERA is the BGF process (Fig. 1 left), where the photon enters directly the hard interaction. However, it is known since long that there is a sizeable production component which can be explained by resolved photon processes. A leading order diagram for such a process is shown in Fig. 1 right. Here the photon emitted by the electron fluctuates hadronically before the hard interaction, and a parton from this fluctuation enters the hard process. In the depicted reaction, which is called ‘charm excitation’, this parton is a charm quark. The charm excitation process exists only in the massless picture since in the massive scheme only light partons can be constituents of the resolved photon structure, and the charm quarks can only be produced in the hard interaction. However, in the massive calculation, excitation like processes (but with massive charm quarks) can appear at NLO. The main experimental hints on contributions from excitation like processes were obtained by analyses selecting events containing a reconstructed $D^*$ meson and at least two jets. The kinematic information from the two leading jets allows to reconstruct (in the leading order picture) the fraction $x_T$ of the photon energy which is carried into the hard interaction. Recent H1 results [14] on this observable are shown in Fig. 7 and are compared with predictions by the MC@NLO massive scheme programme. The calculation provides a good description of the data for values of $x_T$ close to unity, where direct processes are dominating, but undershoots the data by far in the small $x_T$ region. This demonstrates that the aforementioned massive NLO contribution to excitation like processes cannot account for the charm event yields in this region. A better description of the data in this region (not shown) is obtained by the PYTHIA [15] Monte Carlo model containing a massless charm excitation component. Also, as has been reported in previous Ringberg workshops, numerous other HERA measurements involving $D^*$ mesons and jets have added evidence for the presence of a sizeable excitation like component at small $x_T$, for instance in the analysis [16], where the dijet angular distributions were investigated.

2.2. Beauty photoproduction

The relatively large beauty quark mass makes one believe that the pQCD predictions for beauty quark production should be very reliable. Thus it was a surprise when the very first results on beauty production in photoproduction at HERA and also in $p\bar{p}$ collisions at TEVATRON, about fifteen years ago, indicated some excesses of the data over the NLO calculations. However, since then the situation has improved. Nowadays one can state that there is in general a reasonable agreement of data and NLO predictions, both for HERA and TEVATRON. The main reasons for the situation change is that much more precise data have become available and simultaneously the theory models have advanced.

Figure 8 shows a compilation of all the available beauty photoproduction results at HERA as a function of the transverse momentum of the beauty quark.

![Figure 7: Differential cross sections for $D^*$ meson plus dijet production in photoproduction as a function of the observable $x_T$. The H1 data [14] are compared with the MC@NLO [12] massive scheme predictions.](image)

![Figure 8: Summary plot of beauty photoproduction results at HERA as a function of the transverse momentum of the beauty quark. The H1 and ZEUS data points, based on various beauty hadron tagging techniques, are compared to the FMNR [2] massive NLO predictions, for more details of the calculation see [17].](image)
The plot contains measurements based on various experimental techniques. The latest two additions are the ZEUS measurement labelled “ZEUS 133 pb$^{-1}$ b jet” [17] and the H1 analysis labelled “H1 (prel) $b \rightarrow ee$” [18]. In the new ZEUS measurement the events containing beauty quarks are tagged by exploiting the signatures of long lifetime and large mass of the beauty hadrons. Tracks associated to the beauty jet candidate are fitted to a secondary vertex. The significance of the decay length from primary to secondary vertex and the secondary vertex mass are used as observables to separate the beauty, charm and light flavour contributions to the event yields. With this technique an almost background free subsample of more than thousand beauty quark events is obtained. The measurement delivers the highest $p_T$ data point in Fig. 8 plus four other points. The new H1 measurement [18] is complementary, it is based on tagging events where both beauty quarks decay semileptonically with electrons in the final state. This double tagging technique allows to select events with very small beauty quark transverse momenta down to threshold. All the measurements shown in Fig. 8 are compared to the FMNR massive scheme NLO calculation. Both the new ZEUS and H1 measurements are reasonably well described by the calculation as is the large majority of the other measurements. The calculation is also shown (dashed line) with a factor two smaller renormalisation and factorisation scale $\mu$. It should be noted that this smaller scale was for many years the standard choice and using it gave the impression that there might be some trend of theory undershooting data. However, since the scale choice is arbitrary such a conclusion is not warranted.

Figure 9 shows from the same ZEUS analysis [17] as discussed above the measured cross sections as a function of the pseudorapidity of the beauty tagged jet. Again a good description is observed by the FMNR massive NLO calculation, for both PDF sets tested.

As for charm also for beauty production the very first results from LHC have become recently available. Figure 10 shows the cross sections for b-jet production as measured [19] by CMS, using an inclusive secondary vertex tag method. The results are presented double differentially as function of the transverse momentum and the rapidity of the tagged jet. In the covered region of transverse momenta from 20 GeV to about 350 GeV the cross sections fall over seven orders of magnitude. The MC@NLO calculation, which is also shown, provides a good description over most of the phase space, with the exception at highest $p_T$ in the not so central rapidity bins where it exceeds the data.
2.3. Charm production in DIS

For charm production in deep inelastic scattering at HERA, in general a good description of the data by the massive NLO scheme predictions has been observed. Figure 11 shows the recent H1 measurements [20] of $D^*$ meson production as a function of $Q^2$. Over the whole kinematic range, $5 \text{ GeV}^2 < Q^2 < 1000 \text{ GeV}^2$, the data are well described by the HVQDIS [3] massive NLO calculations performed with two different PDF sets. Figure 12 presents for the same H1 analysis the obtained cross sections as a function of $z(D^*)$, which denotes the fraction of the photon energy transferred in the proton rest frame to the $D^*$ meson. At low $z$ the HVQDIS prediction clearly undershoots the data. This deficiency is the most significant one observed for charm production in DIS and has been observed in many analyses. The interpretation of this effect is involved since the $z$ observable is sensitive to both higher order perturbative corrections and to the hardness of the fragmentation process. Figure 13 shows results from the same H1 analysis as a function of the event inelasticity $y$. Additional cuts have been applied here on the $D^*$ transverse momentum in the $\gamma p$ centre-of-mass frame $p_{T,\gamma}(D^*) > 2.0 \text{ GeV}$. This facilitates a comparison of the data also to a prediction [21] based on the zero mass variable flavour number scheme. As one can see this calculation predicts significantly too high cross sections at low values of $y$, corresponding to the threshold region. The massive NLO scheme calculation provides a much better description of the data. Further interesting new measurements [22] of charm and beauty jets in deep inelastic scattering using a secondary vertex tag are available from H1. They show that the massive scheme NLO QCD predictions (and testing several PDF sets used for the calculation) describe the data also in the presence of an additional hard scale provided by the jet.

From all charm analyses at HERA, the determination of the contribution of charm production to the total DIS rates is the one which is most in the focus of interest. This contribution is usually represented by the structure function $F_2^c$, which is defined as the part of $F_2$ due to events with charm quarks in the final state. Experimentally one can measure charm production only in the acceptance range of the detectors, which is limited to rather central rapidities and usually to some minimum
transverse momentum of the tagged charmed hadron. Thus, in order to determine $F_2^{c\bar{c}}$, the measurements are extrapolated from the visible to the total phase space. The extrapolation causes inevitably a further systematic uncertainty. The extrapolation factors are determined using the NLO theory calculations and are typically of order 1.5 or larger. Recently, the H1 and ZEUS collaborations combined [23] their available $F_2^{c\bar{c}}$ measurements based on various tagging techniques (using fully reconstructed charm mesons, muons from semileptonic decays or inclusive secondary vertex tags). For the combination a method of weighted averaging was applied, where parameters representing experimental systematic uncertainties (e.g. calorimeter energy scales) are also fitted. This leads to an effective cross calibration of the two experiments and thus to greatly reduced uncertainties. Figure 14 shows the $F_2^{c\bar{c}}$ input data and also the combined results, as a function of the Bjorken scaling variable $x$ for various $Q^2$ values. The improvement of the various predictions. A dominant source of uncertainty for the calculations is the value of the charm quark mass parameter, which affects the $F_2^{c\bar{c}}$ predictions especially at low $Q^2$. As discussed in the talk [7] and documented in [24], the combined $F_2^{c\bar{c}}$ data can be used to determine for the various schemes their optimal charm quark mass parameter values. This turns out to be helpful to stabilise the PDF fits to the HERA inclusive neutral and charged current data such that the different scheme predictions for important LHC processes, e.g. W production, become much more consistent with each other.

In Fig. 16 further recent $F_2^{c\bar{c}}$ measurements [25] by ZEUS are shown, which have not yet been used in the combination with H1. The data are compared to the predictions based on the HERAPDF1.0 [26] PDF set. This set was obtained from fitting the flavour inclusive neutral and charged current data from the HERA I period, i.e. without using any $F_2^{c\bar{c}}$ measurements. Within the theory uncertainties, dominated by the charm quark mass value, the prediction describes the $F_2^{c\bar{c}}$ data very well. In the first place this demonstrates the universality of the gluon density obtained from the scaling violations of $F_2$ with the one that drives charm production. This universality is expected from the QCD factorisation theorem. From the new $F_2^{c\bar{c}}$ results one can ex-

Figure 14: The structure function $F_2^{c\bar{c}}$ as a function of $x$ for various $Q^2$ values. The HERA combined results are shown as well as the various separate input data from H1 and ZEUS used for the combination. For better visual clarity the different input data sets are offset horizontally from each other by small amounts.

Figure 15: The structure function $F_2^{c\bar{c}}$ as a function of $x$ for various $Q^2$ values. The HERA combined results [23] are compared to various pQCD based predictions (for further details and references for the predictions see also [23]).
Figure 16: The structure function $F_{cc\bar{c}}^2$ as a function of $x$ for various $Q^2$ values. The new preliminary ZEUS data [25], based on $D^*$ and $D^+$ meson tags are shown together with the HERA combined results presented already in Fig. 15. Also shown are the pQCD predictions based on the HERAPDF1.0 [26] set.

expect another significant improvement towards the final HERA combined $F_{cc\bar{c}}^2$ data which will provide one of the most important legacies from HERA in the domain of heavy flavour physics.

2.4. Beauty production in DIS

The domain of beauty production in DIS concludes our survey of open heavy flavour production at HERA. Figure 17 shows new ZEUS results [27] on beauty jet production, as a function of the jet transverse momentum. The analysis is based on the same inclusive secondary vertex tagging method as used in the photoproduction measurement (see Fig. 9). The resulting measurements are the most precise ones so far. As shown in Fig. 17 the massive scheme HVQDIS calculation describes the data adequately over the large momentum range covered.

Also for beauty production the contribution to the total DIS cross section, represented by the structure function $F_{bb\bar{b}}^2$, is at the centre of attention. Of particular interest are high values of $Q^2$ where one can measure with $F_{bb\bar{b}}^2$ an effective beauty sea quark density in the proton. This can be used for predictions of many interesting processes at LHC with beauty quarks in the initial state. For instance, as discussed in [29], in the minimal supersymmetric extension of the standard model the production of the neutral Higgs boson $A$ is driven by $bb\to A$ and for the calculation of this process the PDF uncertainties dominate over the theoretical uncertainties of the perturbative calculation. Figure 18 shows a compilation [30] of the HERA $F_{bb\bar{b}}^2$ results, based on semileptonic and/or inclusive secondary vertex tags of beauty hadrons in the final state. The data are described by the various model predictions. It is a most important task, to be yet performed, to combine the various $F_{bb\bar{b}}^2$ data from H1 and ZEUS. This will allow to achieve the best precision and thus to test the theory predictions at a new level of quality.

3. Conclusion and outlook

In this review it was shown that open charm and beauty production provide a most intriguing testing ground for perturbative QCD. The extra hard scale pro-
Figure 18: The structure function $F_{2}^{b\bar{b}}$ as a function of $Q^2$ for various $x$ values. The measurements [30] from H1 and ZEUS, based on various beauty hadron tagging techniques, are compared to different pQCD predictions (for references see also [30]).

The measurements of the charm and beauty quark production contributions to the total DIS cross sections, represented by the structure functions $F_{2}^{c\bar{c}}$ and $F_{2}^{b\bar{b}}$, provide the most important legacy of all heavy flavour measurements at HERA. They contain most valuable information for the PDF fits at HERA, for instance $F_{2}^{c\bar{c}}$ is important for determining values of the charm quark mass parameters used in the fits.

Across the full kinematic range, the massive scheme NLO calculations provide a reasonable description of the charm and beauty production data with exceptions in some phase space corners. That the massive calculations have not yet been observed to break down in the HERA kinematic region was anticipated by many theorists, see for instance the discussion in [31]. Unfortunately, to this day, NLO programs are available (FMNR and HVQDIS) only for the massive scheme, which can calculate differential cross sections for any kinematic configuration of the outgoing hard partons (up to three partons at NLO). The predictions based on these parton level calculations have one weakness that should be mentioned here. So far the hadronisation process for these calculations has been modelled by rather simple models, such as the Peterson [32] fragmentation function with fixed parameters independent of the hard scale.

The present calculations in the massless or generalised variable flavour number schemes can either provide predictions for total DIS cross sections (see next paragraph) or for single inclusive particle spectra. Where applicable, such NLO or NNLO predictions did in most cases not lead to an improvement but provided a similar or sometimes even worse description of the HERA data.

After the very successful physics program at HERA there are ideas for future electron-proton colliders. One of them is the Electron-Ion Collider (EIC) [33] which is at present under discussion in the US. The EIC is planned to be operated with a factor 3 or lower centre-of-mass energy compared to HERA, but with 100 times larger luminosity and with polarised protons and nuclei. Heavy flavour production at the EIC will be helpful as a specific hard probe that can be used to test parton saturation effects that are expected at low $x$, in particular in heavier nuclei. Another intriguing project is the LHeC [34] collider. Here the idea is to collide the LHC protons with electron beams of energies between 50 and 150 GeV, depending on the accelerator design. The LHeC would be a worthy successor of HERA with a factor 3-5 higher centre-of-mass energy and also much larger integrated luminosity. On top of this the heavy flavour measurements will greatly benefit from the advanced detector design at LHeC with high precision (Silicon or similar) trackers all over the place. At HERA, the tagging of heavy flavours was restricted to central rapidities and e+e− efficiencies take the background pollution into account. It is defined as the efficiency of an equivalent background free sample with the same signal precision as that obtained in the data.

\footnote{The effective efficiency takes the background pollution into account. It is defined as the efficiency of an equivalent background free sample with the same signal precision as that obtained in the data.}
only 0.1% (few%) for charm (beauty) were reached. At LHeC much higher efficiencies can be expected. Figure 19 shows the LHeC expected results for the structure function $F_2^{cc}$, obtained with the RAPGAP [28] Monte Carlo program. The projected LHeC data are presented as points with error bars which (where visible) indicate the estimated statistical uncertainties. For the open points the detector acceptance is assumed to cover the whole polar angular range. For the grey shaded and black points, events are only accepted if at least one charm quark is found with polar angles $\theta_c > 2^0$ and $\theta_c > 10^0$, respectively. The combined HERA results from H1 and ZEUS are shown as triangles with error bars representing their total uncertainty. For further details see the caption of Fig. 19.

LHeC will be also the first $e\nu$-collider where the production of top quarks can be studied. For more physics study results see the web pages [34].

4. Bibliography

References


[23] H1 and ZEUS Collaborations, H1prelim-09-171, ZEUS-prel-09-015, web page: https://www.desy.de/h1zeus/combined_results/heavy_flavours/h1_zeus_f2c_prelim.pdf.