Heavy Flavour Production at HERA

Kenan Mujkic

September 24, 2010

Abstract

The HERA accelerator (Hadron Elektron Ring Anlage) was the world's first and to date only lepton-proton collider. Located at the Deutsches Elektronen Synchrotron (DESY) in Hamburg, Germany, it operated between May 1992 and July 2007 and provided a unique opportunity for resolving the structure of the proton. During the 2006/07 data taking HERA provided collisions between a positron beam of $E_e = 27.5\,\text{GeV}$ and a proton beam of $E_p = 920\,\text{GeV}$, corresponding to a centre-of-mass energy of $\sqrt{s} \approx 318\,\text{GeV}$. In the present analysis charm and beauty production in deep inelastic scattering in the kinematic region $5 < Q^2 < 400\,\text{GeV}^2$, $0.02 < y < 0.7$ have been measured with the ZEUS detector. The beauty and charm content in events with a jet have been extracted using the decay length significance and invariant mass of secondary vertices. Finally, a single-differential charm jet production cross section as a function of Q^2 has been measured.

1 Introduction

The production of heavy quarks in deep inelastic scattering is dominated by the boson gluon fusion (BGF) process, where the virtual photon interacts with a gluon from the proton. Thus, measuring the heavy quark contribution to the proton structure functions constrains the density of gluons and heavy quarks themselves in the proton, complementing the results from inclusive deep inelastic scattering. Furthermore, mass effects are relevant in a large part of the kinematic region for heavy quark production at HERA. The precise measurement of charm and beauty production could therefore help to distinguish between different theoretical approaches to include mass effects in perturbative QCD calculations.

Various experimental techniques are applicable for tagging heavy flavours, and most of them are tailored to a specific hadronic final state. However, if no decay channel is specified it is possible to exploit the long life time of hadrons containing heavy quarks and determine their decay length through the reconstruction of secondary vertices. In the present analysis the heavy quark content of the data sample was extracted using the decay length defined as the vector between the primary vertex and the secondary vertex, projected onto the momentum vector of the decayed hadron. Finally, the single-differential charm jet

production cross section was calculated as a function of Q^2, and compared to the theoretical prediction.
As a preface chapter two gives an overview of the accelerator and the relevant parts of the ZEUS detector, whilst chapters three and four review the principles of deep inelastic scattering and heavy flavour production at HERA. Chapter five is reserved to the event selection, including the discussion of relevant cuts for deep inelastic scattering, the track selection and reconstruction of jets as well as a discussion of the secondary vertex selection. Chapter six explains the signal extraction including the method of mirroring and subtracting the decay length significance distribution and fitting the different MC distributions to the data. In chapter seven the measurement of the single-differential charm jet production cross section is presented, and chapter eight discusses the results and possible extensions of the analysis.

2 HERA and the ZEUS Detector

2.1 HERA

The HERA accelerator (Hadron Electron Ring Anlage) was located at the Deutsches Elekronen Synchrotron (DESY) in Hamburg, and operated between May 1992 and July 2007. High energetic collisions of leptons and protons were studied at various experiments located in experimental halls around the HERA ring.
The HERA ring itself was composed of two independent storage rings, one for protons and the other for electrons (or positrons), which were located 10 - 25 m below ground and had a circumference of 6 km. Electrons and protons were accelerated in several pre-accelerators and finally injected into HERA in bunches, each bunch containing about 10^{10} particles. Inside HERA electrons and protons were accelerated from 12 GeV and 40 GeV to their final energies of 27.5 GeV and 920 GeV, respectively. At the centre-of-mass energy of $\sqrt{s} \approx 318$ GeV the electron and proton beam were brought to collision each 96 ns at the two interaction points, where the particle detectors H1 (north hall) and ZEUS (south hall) were located (see figure 1).
The data-taking at HERA is divided into two periods, HERA I from 1996 - 2000 and HERA II from 2003 - 2007. During the shutdown from 2000 to 2002 several machine and detector upgrades were undertaken. Spin rotator pairs were installed in front of the interaction regions of H1 and ZEUS, and the integrated luminosity could be increased in the course of the HERA upgrade. The luminosity was measured indirectly by two lead-scintillator calorimeters, which detect photons from the Bether-Heitler process $ep \rightarrow e^{'}p\gamma$. The luminosity is then calculated from the number of detected photons and the well-known cross section for the process. In total $\mathcal{L} \approx 0.5\ fb^{-1}$ of data was acquired.

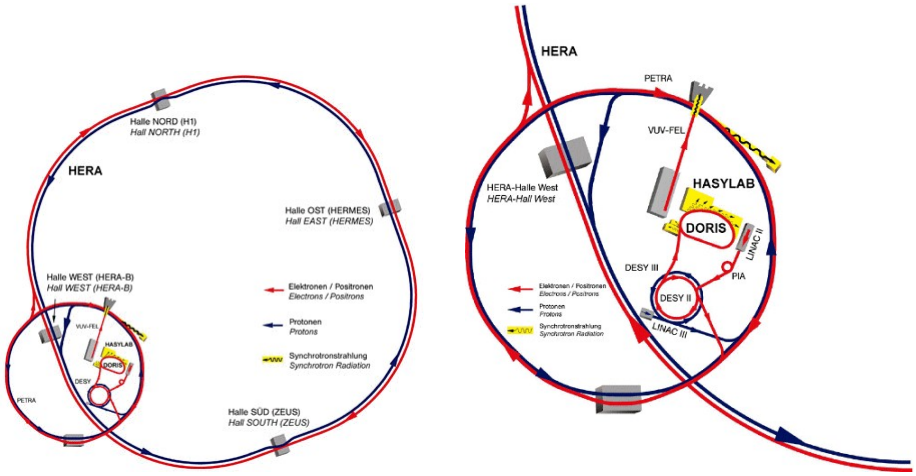

Figure 1: Schematic of HERA and pre-accelerators

2.2 The ZEUS detector

The ZEUS detector [1] was one of two multipurpose particle detectors designed to measure final state particles from lepton - proton collisions, the other one being H1. Figure 2 shows a cutaway of the detector with the most important parts of the detector being labeled.

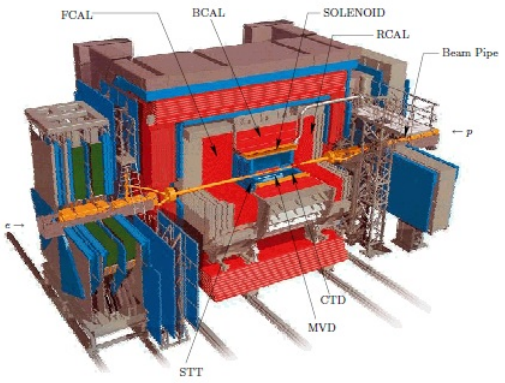

Figure 2: Cutaway of the ZEUS detector

Like most detectors in particle physics ZEUS was constructed onion like, and provided almost 4 π solid angle coverage about the interaction region. The microvertex detector (MVD) was installed in order to reconstruct secondary vertices using precisely measured tracks close to the primary vertex. A superconducting solenoid around the dift chamber provided a to a large extent

homogeneous magnetic field of 1.43 T in the proton direction, and was a crucial component for the track measurement performed by the central tracking detector (CTD). In forward direction followed the forward detector (FDET), consisting of three planar drift chambers (FTD) as well as of the two modules of the straw-tube tracker (STT), and in rear direction the rear tracking detector (RTD). The uranium scintillator calorimeter was constructed in order to measure the energy deposit of the penetrating particles, and was divided into three geometrical parts: the forward calorimeter (FCAL), the central part (BCAL) and the rear calorimeter (RCAL). Each of the three parts was subdivided into an inner electromagnetic section (EMC) and in the case of the RCAL one, otherwise two outer hadronic sections (HAC). A particle that deposited its whole energy in the electromagnetic section could be identified as an electron or photon, particles which reached or passed the hadronic section had to be identified with additional criteria. The asymmetric design was another important characteristic of the ZEUS detector, which accounted for the difference in beam energies and the resulting boost of the centre-of-mass system. Consequently particle physicists introduce the Lorentz-invariant pseudorapidity:

$$\eta = -\ln\,(\tan\frac{\theta}{2}), \tag{1}$$

and substitute the polar angle θ with the physical observable η.

2.3 Microvertex detector (MVD)

The microvertex detector (MVD) [2, 3] was one of the components that were installed during the shutdown, and consisted of a barrel (BMVD) and a forward (FMVD) section. Its main purpose was the improvement of the vertex finding and track reconstruction in the immediate vicinity of the interaction point. The BMVD consisted of 30 ladders, which were placed cylindrically around the beampipe, and were made up of 5 layers consisting of two half modules each. Each half module was made up of two silicon sensors, whose strips were perpendicular to each other, one sensor measured the position in z-direction and the other provided the r - ϕ information. There were 512 strips per sensor that were read out with a distance of 120 μm between them. The FMVD consisted

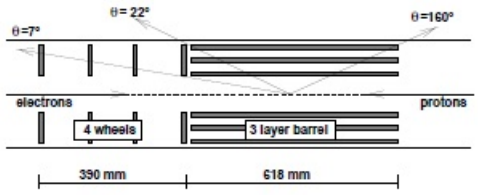

Figure 3: Side-view of the microvertex detector

of four wheels, which were mounted perpendicular to the beampipe. They were

composed of 14 sectors, each containing an inner and an outer sensor, which were mounted back-to-back. The FMVD sensors were wedge-shaped and the angle between an inner and an outer sensor was $180^0/14$. The MVD was able to detect tracks with polar angles of $7 - 160^0$, and had a design resolution of $\approx$ 10 μm for track and vertex finding.

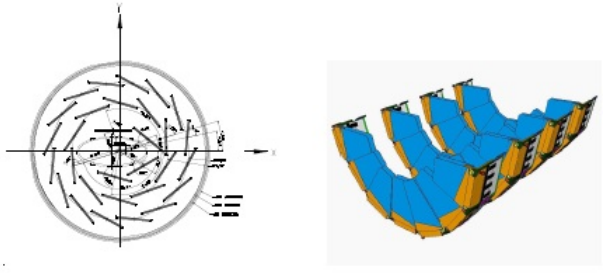

Figure 4: The barrel MVD and lower half of the forward MVD

2.4 Central tracking detector (CTD)

The central tracking detector (CTD) [4, 5, 6] consisted of 72 cylindrical layers grouped in 9 superlayers. Each superlayer was divided into 32 - 96 cells with the first superlayer having 32 cells and each further superlayer having 8 cells more than the previous one. In the odd numbered superlayers the signal wires ran parallel to the z-axis, in the even numbered with an alternating angle of 5^0 with respect to the z-axis. The orientation of the wires was chosen in such a way to allow for the measurement of the z-position of a signal. The main purpose of the CTD was the reconstruction of tracks and the measurement of the charges, momenta and creation points of particles produced in lepton - proton collisions. The track reconstruction was considered to be reliable if the particle had crossed three superlayers, such that the CTD covered an angular range of $20^0 < \theta < 160^0$. The spacial resolution achieved was $\approx 300\mu$m in r - ϕ and 1 - 5 mm in z.

2.5 Calorimeter (CAL)

The calorimeter (CAL) [7, 8] covered 99.7 % of the 4π solid angle, and consisted of the forward (FCAL), barrel (BCAL) and rear (RCAL) section. Each section was further divided into towers consisting of electromagnetic (EMAC) and hadronic (HAC) cells. Apart from their differing dimensions the electromagnetic and hadronic cells were identical in design, both had alternating layers of depleted uranium and plastic scintillator, and the response of the CAL to electromagnetic and hadronic showering was the same. The energy resolution of the calorimeter was measured to be $\frac{\sigma(E)}{E} = \frac{35\%}{\sqrt{E}}$ for hadrons and $\frac{\sigma(E)}{E} = \frac{18\%}{\sqrt{E}}$ for electrons.

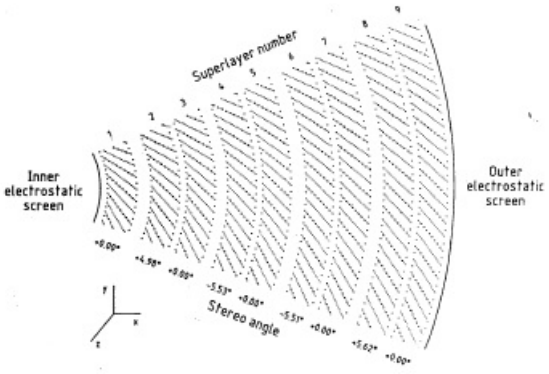

Figure 5: An octant of the central tracking detector in x - y view

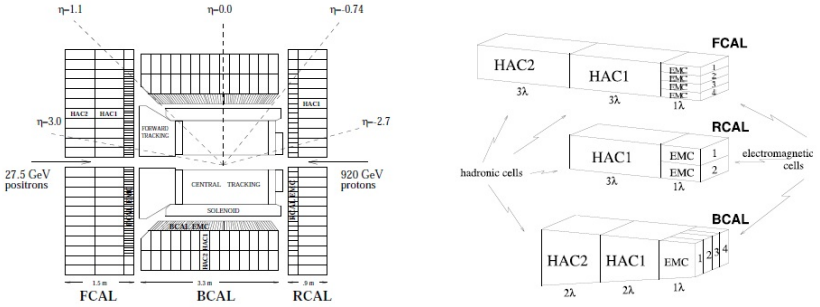

Figure 6: Schematic of calorimeter and layout of CAL tower

3 Deep Inelastic Scattering (DIS)

Deep inelastic scattering (DIS) [9] in ep interactions occurs when the electron emits an off-shell boson, that interacts with the constituents of the proton causing the proton to break up. One distinguishes two different cases: in case the exchanged boson is a photon or a Z^0 one speaks of neutral current, otherwise in case the exchanged particle is a $W^{+/-}$ of charged current DIS. As the virtual boson interacts with the internal constituents of the proton, deep inelastic scattering provides us with information about the proton structure.
Figure 7 shows the Feynman diagrams for neutral current and charged current DIS, the quantities $k, k^{'}$ and p represent the four-momenta of the incoming and outgoing electron and the incoming proton, respectively. The four-momentum of the exchanged boson, q, is given by the difference in the four-momentum of the electron before and after the boson's emission. The scattering process is completely described by the following Lorentz - scalars:

$$Q^2 = -q^2 = -(k - k^{'})^2, \tag{2}$$

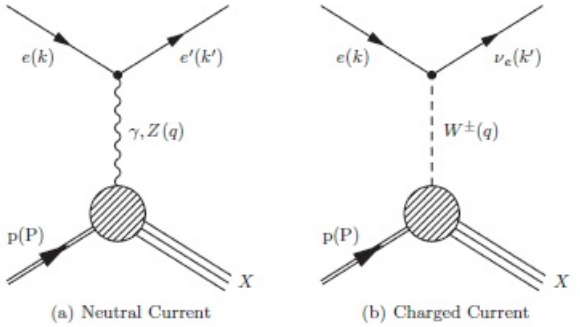

Figure 7: Feynman diagrams of neutral and charged current DIS

$$\sqrt{s} = \sqrt{(k+P)^2}, \tag{3}$$

$$x = \frac{Q^2}{2P \cdot q}, \tag{4}$$

$$y = \frac{P \cdot q}{P \cdot k}, \tag{5}$$

where Q^2 denotes the *photon virtuality*, $\sqrt{s}$ is the *centre-of-mass energy* of the *ep* system, and x and y represent the so called *Björken scaling variables*. The Björken variable x represents the fraction of the momentum of the incoming proton carried by the struck parton, and the variable y can be seen as the inelasticity of the event; it respresents the fraction of the electron energy involved in the reaction measured in the rest frame of the proton. The scalar Q^2 can be seen as a measure of the resolving power of the event with higher values of Q^2 corresponding to higher momentum transfer and therefore greater resolving power. The four variables are related by the equation:

$$Q^2 = s \cdot x \cdot y. \tag{6}$$

For $Q^2 \gg 1GeV^2 ep$ events are refered to as deep inelastic scattering, while events with low Q^2 are characterised by the exchange of quasi-real photons and refered to as photoproduction.

4 Heavy Flavours

At HERA the dominant production mechanism for heavy quarks was boson-gluon fusion (BGF) [9]. In direct BGF a photon or Z^0 boson emitted by the incoming lepton interacts with a gluon from the proton producing a heavy quark pair ($c\bar{c}$ or $b\bar{b}$). The photon may also fluctuate into $q\bar{q}$ pairs creating hadronic structures. This process, in which the heavy quark pair is produced via gluon-gluon fusion, is called resolved boson-gluon fusion, and leads to the concept of photon structure functions. A single quark can also originate directly from the

proton or from a hadron produced by the photon via excitations, which are usually classified as resolved boson-gluon fusion.

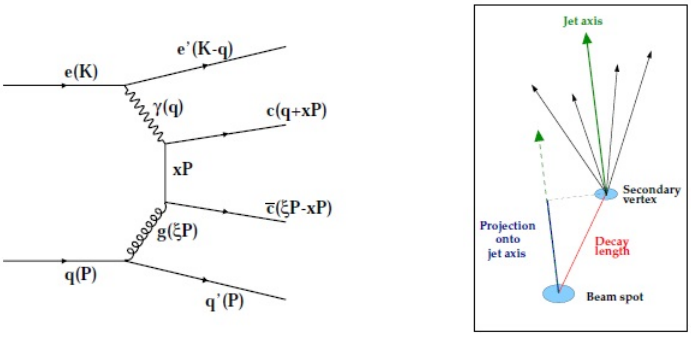

Figure 8: Direct boson gluon fusion and decay length

Various experimental techniques are applicable for tagging heavy flavours, and most of them are tailored to a specific hadronic final state. However, if no decay channel is specified it is possible to exploit the long life time of hadrons containing heavy quarks and determine their decay length through the reconstruction of secondary vertices. In the present analysis the charm and beauty content was extracted using the decay length defined as the vector between the primary vertex and the secondary vertex, projected onto the momentum vector of the decayed hadron [10]. Studies showed that there was no noticeable difference between the usage of the beam-spot and the event-primary vertex as the reference point for the decay-length calculation. In order to avoid systematic effects originating from the unprecise measurement of the z-position of the beam-spot the decay length $\vec{d}$ was projected onto the x - y plane:

$$d_{xy} = \frac{(\vec{S}_{xy} - \vec{B}_{xy}) \cdot \vec{p}}{p_T}, \tag{7}$$

where $\vec{S}_{xy}$ and $\vec{B}_{xy}$ denote the position of the secondary vertex and the beam-spot in the transverse plane, $\vec{p}$ represents the momentum vector of the decaying hadron, and p_T denotes its transverse momentum.
The sign of the decay length was assigned using the axis of the jet associated to the secondary vertex. In case the decay length $\vec{d}$ was in the same hemisphere as the jet axis $\vec{j}$ a positive sign was assigned; otherwise the sign of the decay length was negative. In a perfect detector this sign convention would result in the decay length of heavy flavour decays being strictly positive and those of half of the combinatorial background being negative. The finite detector resolution translates into a distribution with some heavy flavour decays smeared out into the negative region. Further the unprojected decay length had a minimum at zero due to the contribution of vertices for which the decay length is approximately perpendicular to the jet direction and the sign of the decay length flips.

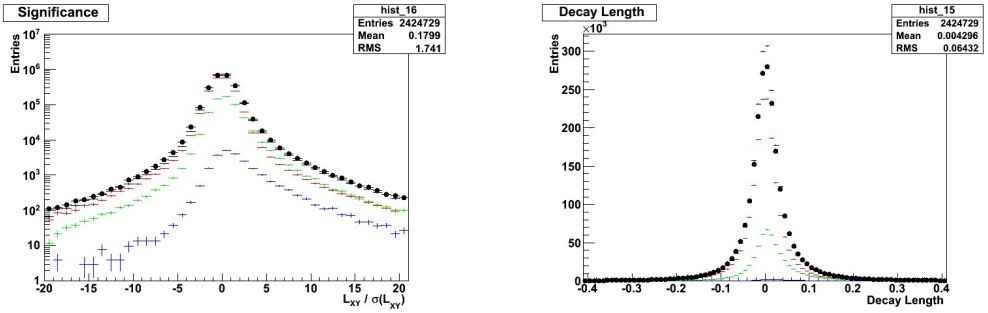

Figure 9: Decay length significance and projected 2 D decay length d_{xy} [cm]

The projection of the two-dimensional decay length onto the axis of the associated jet assured a Gaussian like distribution.
The decay length significance S is defined as the decay length d_{xy} divided by the error δd_{xy} on the decay length and provides a powerful distinguishing variable for separating combinatorial background from heavy flavour decays [10]:

$$S = \frac{d_{xy}}{\delta d_{xy}} \tag{8}$$

S is a measure for the probability of the vertex being correctly reconstructed at a considerable distance from the interaction point. The significance distributions for beauty, charm and light flavours are shown in figure 9. Here and in all following figures blue, green and red histograms denote beauty, charm and light flavours, respectively, while the data are denoted by black points. Compared to the decay length distributions, the significances reveal even larger asymmetries for beauty and charm, which implies that the secondary vertices associated to heavy flavour jets could be determined with a higher precision than those associated to light flavour jets. The distribution of the two-dimensional decay length projected onto the axis of the associated jet is shown separately for beauty, charm and light flavours in figure 9.

5 Event and candidate selection

For the present analysis data taken in the 2006/07 e^+p running period with an integrated luminosity of $\mathcal{L} = 145.9\ pb^{-1}$ was available. A corresponding set of Monte Carlo (MC) samples of charm and beauty generated with the Rapgap [11] and of light flavours generated with the Ariadne [12] program was accessible. Since the beauty and charm samples contained signal events, which were used to measure production cross sections, they were required to be available in a sufficient quantity in order to assure statistical precision. Therefore roughly 18 times the data statistics was produced for the beauty and 3.5 times the data statistics for the charm samples. For the light flavours an inclusive sample was

used, which contained the background events and corresponded approximately to the luminosity of the data.
In order to select events from Deep Inelastic Scattering (DIS) and discart events in the photoproduction regime in the present analysis it was required that a candidate for the scattered positron was reconstructed with an impact point on the surface of the RCAL outside a region of ± 13 cm in x and y around the beampipe. Candidates for the scattered positron were identified with the Sinistra programme and assigned a probability $P_{e'}$. For the final selection events with the most probable candidate having a sinistra probability $P_{e'} \leq 0.9$ and an energy $E_{e'} \leq 10$ GeV were rejected. In order to reject further events from photoproduction the following cuts were applied:

$$y_e < 0.7, \tag{9}$$

$$y_{JB} > 0.02, \tag{10}$$

$$Q^2_{DA} < 5\,\mathrm{GeV}^2, \tag{11}$$

$$44.0\,\mathrm{GeV} < E - p_z < 65.0\,\mathrm{GeV}, \tag{12}$$

where y_{JB} and y_e are estimators for the inelasticity, and Q^2_{DA} for the virtuality of the event.
For small values of y, the Jacquet-Blondel estimator [13] $y_{JB} = (E - p_y)/(2E_e)$ was used, where $E - p_z = \sum_i E^i - p^i_z$ and the sum runs over all energy flow objects (EFOs). EFOs [14] combine information from the calorimetry and tracking, corrected for energy loss in dead material. On the other hand, for large values of the inelasticity, the estimator y_e was calculated with the electron method, which uses the scattering angle θ_e, and the energies before and after the collision, $E_{e/e'}$, for the determination of the kinematic variables:

$$y_e = 1 - \frac{E_{e'}}{2E_e}(1 - \cos\theta_e). \tag{13}$$

The estimator Q^2_{DA} for the virtualiy was determined with the double-angle method [15], which exploits the scattering angle θ_e of the outgoing lepton and the effective angle φ_h of the hadronic system to reconstruct the kinematics of the event. In terms of these variables the estimator for the virtuality assumes the following form:

$$Q^2_{DA} = 4E^2_e \frac{\sin\varphi_h(1 + \cos\theta_e)}{\sin\varphi_h + \sin\theta_e - \sin(\varphi_h + \theta_e)}. \tag{14}$$

Finally, standard inclusive DIS triggers were applied on analysis level, and only events with a well-reconstructed primary vertex within $|Z| < 30$ cm around the interaction point were selected in order to reject further non-ep interactions. Control distributions for Z_{prm} and Q^2_{DA} are shown in figure 10, and reasonable agreement between data and MC was found.
As described in the previous section the charm and beauty quarks that are produced in boson-gluon fusion hadronise and subsequently decay with the decay

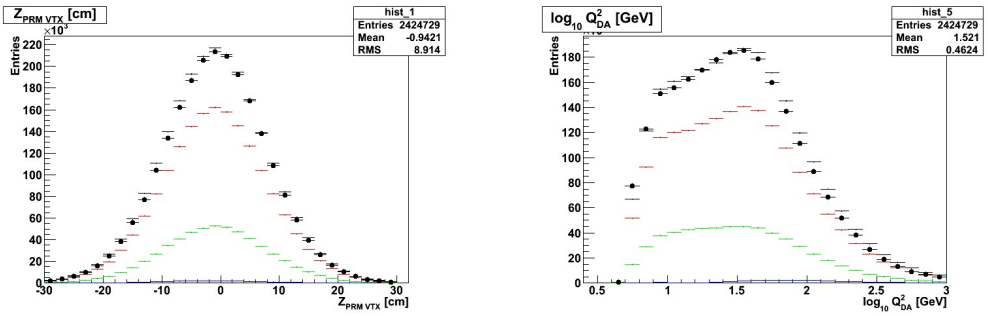

Figure 10: Primary Vertex Z_{prm} [cm] and Virtuality Q^2_{DA} [GeV^2]

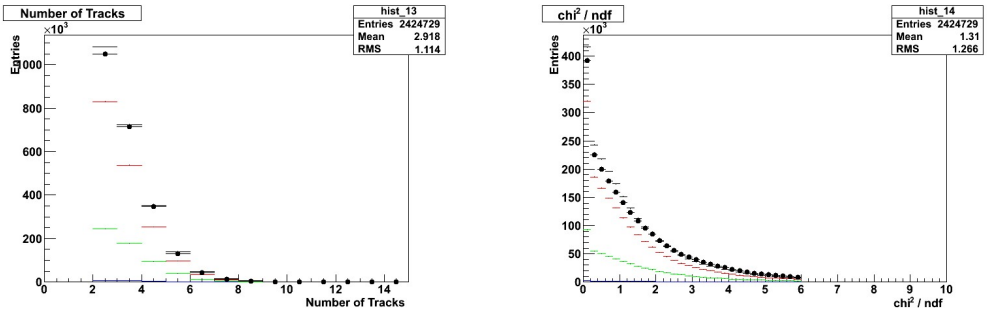

Figure 11: Track multiplicity and fit significance χ^2/ndf

products appearing in the detector as jets of final state particles. Due to the long lifetime of charm and beauty hadrons it is of importance to reconstruct secondary vertices precisely in order to separate heavy flavour from light flavour signals and further background. The reconstruction is based on the principle of fitting a vertex for each jet using well-reconstructed tracks that could be associated to it geometrically [16, 17]. For the track selection the following cuts were applied:

- $p_t \geq 0.5$ GeV,
- # MVD hits ≥ 4,
- # CTD superlayers ≥ 3 of # STT hist ≥ 1 .

The tracks had to be associated to a jet with the maximal distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 1$ to the closest jet. If two or more such tracks were associated to a jet, a candidate vertex was fitted from the coresponding tracks using a weighted χ^2 fit. The weights of tracks not well fitting to the secondary vertex were reduced by a deterministic annealing filter [18]. Figure 11 shows the track multiplicity and fit significance for selected secondary vertices; in both distributions reasonable agreement between data and MC was found. The strong peak at very small values of χ^2/ndf is caused by vertices with only two fitted tracks.

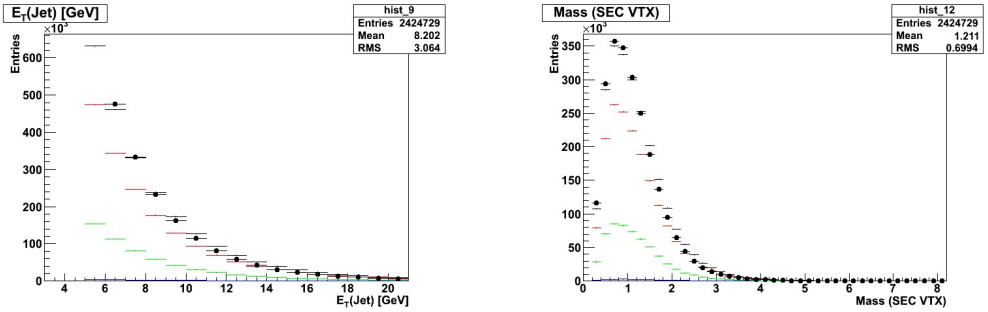

Figure 12: Transverse momentum E_T [GeV] of jet and M_{Vtx} [GeV] of secondary vertex

This is expected, since it is much more likely to obtain a well-defined vertex from fitting two tracks than from a fit of three or more tracks.
At analysis level jets reconstructed with the k_t cluster algorithm [19] were used to apply the final selection. Jets were selected if they fulfilled the requirements $E_T^{jet} > 5$ GeV and $-1.6 < \eta^{jet} < 2.2$, and jets containing the DIS electron rejected. For the reconstruction of jets with high p_T several algorithms are available, whose input elements can be calorimeter cells, objects combining calorimeter and tracking information (EFOS) or even hadrons or partons for the reconstruction of jets on truth level. The k_t algorithm is based on the principle of successive combination of close object pairs into clusters. For each pair (i, j) a distance d_{ij} is calculated:

$$d_{ij} = min(E_T^i, E_T^j)^2 \Delta R^2, \tag{15}$$

where E_T^k denotes the transverse energies of the two objects. Further for each input object i a distance d_{ib} to the beam particles is defined:

$$d_{ib} = (E_T^i)^2 R^2. \tag{16}$$

In case the minimum of all distances between paired objects is smaller that the minimum of all distances between single objects and the beam particles, i.e.

$$min(d_{ij}) < min(d_{ib}), \tag{17}$$

the objects i and j are merged into a cluster. Otherwise object i is complete and marked as a jet. This procedure is iterated until no more pairs with an energy smaller than a certain fraction of the total energy in the event can be found.
In order to select secondary vertices of good quality and reject questionable candidates further cuts were applied:

- $\chi^2/ndf \leq 6$,
- $d_{xy} < 1$ cm,

- $d_z < 30$ cm,

where d_{xy} and d_z denote the distances to the beam-spot in the x-y and the z direction, and χ^2/ndf represents the fit significance.
The cut values were chosen according to previous studies, which confirmed a considerable suppression of background vertices after imposing the quality cuts [20]. In addition, the selected vertices were required to have an invariant mass of the fitted tracks $M_{Vtx} < 7.5$ GeV, since heavy quark signals at significantly higher masses than the beauty mass are not meaningful. The same event and candidate selection was applied to Monte Carlo samples of beauty and charm events generated with the RAPGAP and for light flavour events generated with the ARIADNE program.

6 Signal Extraction

As described in the previous section events containing secondary vertex candidates that fulfilled the criteria were selected for the signal extraction. The vertex candidates were defined as beauty or charm signal according to whether the event contained at least one beauty or charm quark. This included heavy quarks from the hard subprocess as well as those originating from gluon splitting in light flavour events.
For each candidate vertex the two-dimensional decay length, d_{xy}, projected onto the axis of the associated jet was calculated. The heavy flavour content of the sample was then determined using the decay-length significance, $S = d_{xy}/\delta d_{xy}$, and the invariant mass, M_{Vtx}, of the tracks contributing to the secondary vertex fit [21]. The negative side of the significance distribution ($S^-, S < 0$), which is unphysical and caused by detector-related effects, was mirrored onto and subtracted from the positive side ($S^+, S > 0$). This procedure is supposed to minimise the light flavour contribution, since the unmirrored decay-length significance is to a great extent symmetric around zero for light flavour, in contrast to the same distributions for heavy flavours. The decay-length significance distribution was then divided into three different mass bins, such that the highest massbin was dominated by beauty, while the lower two massbins bins were dominated by charm [21]. Figures 13 - 15 show the mirrored and subtracted significance distributions for the three different mass bins; one can observe that the light flavour fraction has decreased significantly.
In order to extract the contributions from beauty, charm and light flavours a binned χ^2 fit of the three MC distributions to the data was performed simultaneously for all three massbins [21]. The overall normalisation of the Monte Carlo was constrained to the normalisation of the data in the unmirrored significance distributions. The fit function to be minimised was defined in the following way:

$$\chi^2 = \sum_{i=1}^{N_S} \sum_{j=1}^{N_M} \frac{(N_{ij}^{data} - k_b \cdot N_{ij}^{b} - k_c \cdot N_{ij}^{c} - k_{lf} \cdot N_{ij}^{lf})^2}{(\delta_{ij}^{data} - k_b \cdot \delta_{ij}^{b} - k_c \cdot \delta_{ij}^{c} - k_{lf} \cdot \delta_{ij}^{lf})^2} \qquad (18)$$

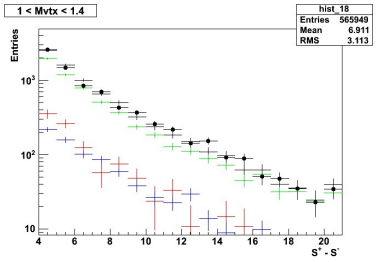

Figure 13: Mirrored Significance, 1.0 GeV $< M_{vtx} <$ 1.4 GeV

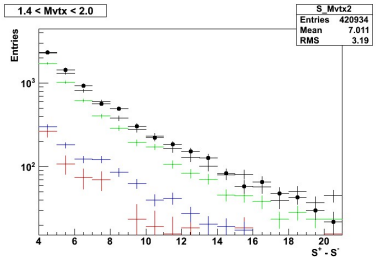

Figure 14: Mirrored Significance, 1.4 GeV $< M_{vtx} <$ 2.0 GeV

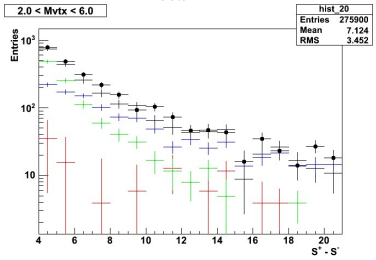

Figure 15: Mirrored Significance, 2.0 GeV $< M_{vtx} <$ 6.0 GeV

$$+\frac{(N_{tot}^{data} - k_b \cdot N_{tot}^{b} - k_c \cdot N_{tot}^{c} - k_{lf} \cdot N_{tot}^{lf})^2}{(\delta_{tot}^{data} - k_b \cdot \delta_{tot}^{b} - k_c \cdot \delta_{tot}^{c} - k_{lf} \cdot \delta_{tot}^{lf})^2} \tag{19}$$

The sums run over all mirrored and subtracted decay-length significance bins i and the three M_{Vtx} bins j. N_{ij}^{l} denotes the number of entries for the sample $l \in \{data, b, c, lf\}$ in bin (i, j) and δ_{ij} represents the corresponding errors. For the MC samples the number of entries was obtained by adding up the MC subsamples weighted according to their integrated luminosities and normalising the sum to the data luminosity. The free parameters of the fit consist in the three scaling factors k_n for the beauty, charm and light flavour MC samples. Due to the normalisation to the data luminosity they directly correspond to the scaling factors that will appear in the cross sections. For the constraint of the overall nomalisation the number of entries and the error in a given bin have been replaced by the total number of entries in the unmirrored distribution and its error.

7 Cross Sections

The total cross section, σ, of a given process is defined as the number of events, N, in the sample containing the process divided by the integrated luminosity $\mathcal{L}$. The formula for the inclusive charm production jet cross section can therefore

be written as:

$$\sigma_c = \frac{N_c^{rec}}{\mathcal{A} \cdot \mathcal{L}}, \tag{20}$$

where $\mathcal{A}$ refers to the charm acceptance and N_c^{rec} to the number of reconstructed charm jets in the data, which have to be determined from the fit using:

$$N_c^{rec} = k_c \cdot N_c^{MC}, \tag{21}$$

where N_c^{MC} denotes the number of reconstructed charm events in the Monte Carlo simulation. If we define the acceptance as $\mathcal{A} = N_c^{MC}/N_c^{HL}$ equation 20 can be written as:

$$\sigma_c = \frac{k_c \cdot N_c^{HL}}{\mathcal{L}}. \tag{22}$$

Here k_c denotes the charm scaling factor obtained from the fit and N_c^{HL} the number of generated charm jets on hadron level. Hadron level jets are obtained by running the k_t cluster algorithm on all final-state MC particles before the decay of the beauty and charm hadrons. True beauty or charm jets could therefore be defined as all hadron level jets containing a b or c hadron, respectively, including all beauty and charm baryons and excitations. In case the true hadron jet contains a b hadron as well as a c hadron originating from a cascade decay, the candidate is counted as a beauty signal. Signatures with beauty and charm hadrons resulting from gluon splitting were als incorporated in the respective signal independently from the hard subprocess.
The single-differential charm jet production cross section as a function of Q^2 is therefore defined as:

$$\frac{d\sigma_c}{dQ^2} = \frac{k_c \cdot N_c^{HL}}{\mathcal{L} \cdot \Delta Q^2}, \tag{23}$$

where ΔQ^2 refers to the width of the given Q^2 bin. In order to determine the charm cross section the simultaneous χ^2 fit of the three different MC distributions to the data for all three mass bins had to be performed for every single bin of Q^2. Figure 16 shows the single differential charm jet production cross section, where the lines with error bars represent the data, while the histogram corresponds to the MC normalised to the area of the data. One can observe a slight discrepancy between the Monte Carlo simulation and the data. However, at this stage the analysis does not include systematic effects, and the measured cross section needs to be compared to NLO QCD predictions.
Furthermore, the measurement of the charm cross section does not include the kinematic region $Q^2 > 400\text{GeV}^2$. The reason is that the mirrored significance distributions provide low statistics for the high Q^2 bins, and the fit of the different Monte Carlo distributions to the data loses its meaning in this kinematic region. The extension of the analysis to the whole data sample with an integrated luminosity of 354 pb^{-1} will lead to a significant decrease in statistical uncertainty. In order to extend the analysis it will also be crucial to improve the vertex and track reconstruction at high Q^2.

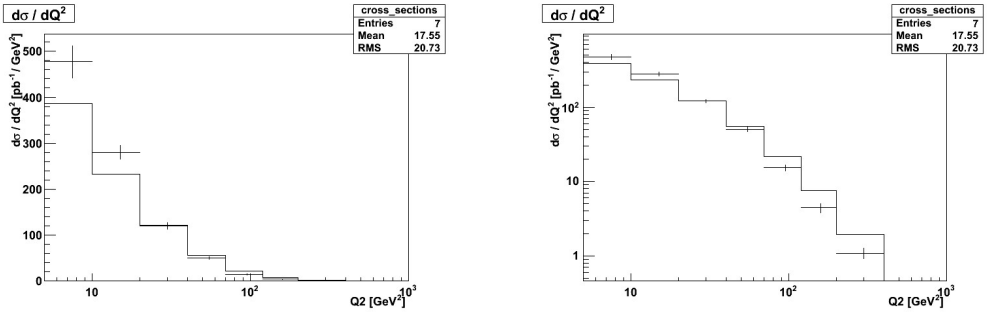

Figure 16: Single-differential charm jet cross section $\mathrm{d}\sigma/dQ^2$ $[pb^{-1}/\mathrm{GeV}^2]$

8 Summary and Outlook

The differential charm jet production cross section shown in figure 16 comprises a slight discrepancy between data and Monte Carlo. However, the comparison with NLO QCD predictions and the correction of systematic effects would improve this picture significantly. The first step in the extension of the present analysis will therefore be the study of the systematics.

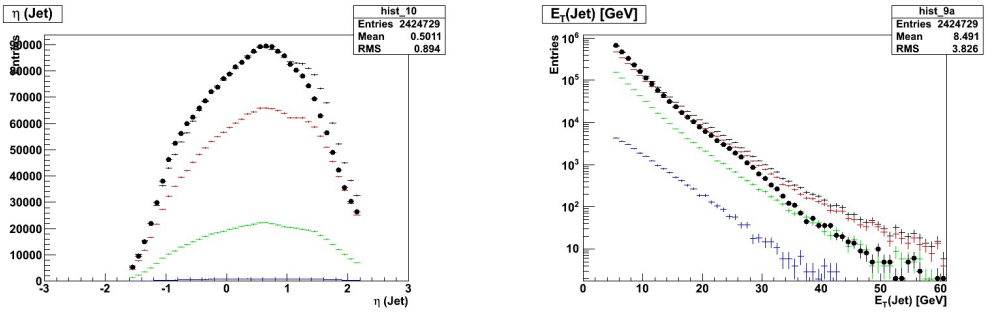

Figure 17: Transverse Energy E_T^{Jet} and pseudorapidity η^{Jet}

Figure 17 shows the pseudorapidity and transverse energy E_T of the selected jets on a linear and logarithmic scale, respectively. In both distributions discrepancies between data and Monte Carlo can be observed, most significantly for high transverse momenta $E_T > 30$GeV and pseudorapidity $\eta > 1.2$. In order to study the effects of this disagreement on the cross sections the light flavour contribution needs to be reweighted according to the differences between data and Monte Carlo assuming a correct description of the beauty and charm contribution. A similar re-weighting needs to be done for the distribution of the invariant mass M_{Vtx} of the secondary vertices and the virtualiy Q^2. For the reweighting with respect to Q^2 one can expect an effect on the total cross section of up to 9 %, it will be most significant for differential cross sections as functions of η and E_T.

Further the tails of the decay length significance distribution shown in figure 9 are not fully described by the Monte Carlo simulation. In order to investigate the impact of this discrepancy on the cross section measurement the MC needs to be smeared out to reproduce the data. An empirical smearing function needs to be developed, which would consist of a Gaussian for the peak region and a logarithmic function for smearing the tail. To evaluate the systematic uncertainties the smearing procedure needs to be applied and subsequently the fits need to be re-done.
Previous studies have shown that the first level trigger (FLT) was not equally efficient in data and MC, and therefore a trigger efficiency correction needs to be applied. For the latest version of reconstruction the luminosity measurement is another source of uncertainty, since certain triggers have not been fired for certain runs, and correspondingly the luminosity measurement needs to be corrected.
Further, the uncertainty originating from the fit procedure can be estimated by varying the fit range in the mirrored and subtracted significance distribution. However, it is known that the cross section is not flat around the beam axis without the application of the cut $S^+ - S^- > 4$, and therefore one should not expect significant improvements due to a variation of the fit range.
The single-differential charm jet production cross section shown in figure 16 does not comprise the kinematic region of $Q^2 > 400\,\mathrm{GeV}$, since the low statistics does not allow for a meaningful fit of the MC distributions to the data. The extension of the analysis to the whole data sample will lead to a significant improvement of the statistical uncertainty in the high Q^2 bins. In order to extend the analysis it will further be crucial to improve the vertex and track reconstruction at high values of Q^2.
Apart from the studies of systematic effects the analysis will be extended to single-differential charm production jet cross sections as functions of E_T, η and x. In order to obtain these cross sections the mirrored and subtracted significance distributions need to be determined for every single E_T, η and x bin. The scaling factors for charm, beauty and light flavours are then extracted by a simultaneous fit of the MC distributions to the mirrored significance distributions of the data for all three mass bin in each bin of E_T, η and x. The single-differential charm production jet cross section as a function of a variable v can then be obtained in analogy to equation 23:

$$\frac{d\sigma_c}{dv} = \frac{k_c \cdot N_c^{HL}}{\mathcal{L} \cdot \Delta v}, \tag{24}$$

where v represents the kinematic variables E_T, η and x, and Δv refers to the width of the bin.
In order to extract the structure functions F_2^{cc} the double-differential charm production cross section $d\sigma/dxdQ^2$ needs to be measured, which can be interpreted as a single-differential cross section $d\sigma/dx$ as a function of x for different intervals of Q^2. To extract F_2^{cc} an extrapolation to the full kinematic phase space needs to be performed using the HVQDIS package, which calculates cross

sections for charm production in DIS in leading and next-to-leading order:

$$F^{cc}_{2,meas}(x_i, Q^2_i) = \frac{\sigma_{meas,i}}{\sigma_{theo,i}} \times F^{cc}_{2,theo}(x_i, Q^2_i). \tag{25}$$

The prediction $F^{cc}_{2,theo}$ for the structure function from HVQDIS is multiplied by the ratio of the measured, $\sigma_{meas,i}$, to the predicted, $\sigma_{theo,i}$, visible cross section in a given bin i. Measuring the heavy quark contributions to the structure functions constrains the gluons and heavy quarks in the proton, complementing the results from deep inelastic scattering. In this project the most precise and definite measurements will be made at high momentum transfer, Q^2, the region most interesting for the LHC.

References

[1] ZEUS Collaboration, *The ZEUS Detector*, Status Report 1993

[2] A. Polini et al., Nucl. Inst. Meth. **A581**, 656 (2007)

[3] E. Koffeman, ZEUS-Note 00-028 (2000)

[4] N. Harnew et al., Nucl. Inst. Meth. **A 279**, 290 (1989)

[5] B. Foster et al., Nucl. Phys. Proc. Suppl. **B 32**, 181 (1993)

[6] B. Foster et al., Nucl. Inst. Meth. **A 338**, 254 (1994)

[7] M. Derrick et al., Nucl. Inst. Meth. **A 309**, 77 (1991)

[8] A. Andresen et al., Nucl. Inst. Meth. **A 309** 101 (1991)

[9] R. Devenish and A. Cooper-Sarkar, *Deep Inelastic Scattering*, Oxford University Press (2003)

[10] D. Nicholass, *The Study of $D^{\pm}$ and D^0 Meson Production in Deep Inelastic Scattering at HERA II with the ZEUS Detector*, PhD Thesis, University College London (2008)

[11] H. Jung, Comp. Phys. Comm. **86**, 147 (1995)

[12] L. Lönnblad, Comp. Phys. Comm. **71**, 15 (1992)

[13] F. Jacquet and A. Blondel, *Proc. the Study of an eP Facility for Europe*, U Amaldi (ed.), p. 391 (1979)

[14] T. Tuning, *ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction*, ZEUS-Note 01-021 (2001)

[15] S. Bentvelsen, J. Engelen and P. Kooijman, *Proc. Workshop on Physics at HERA*, W. Buchmüller and G. Ingelmann (eds.), Vol. 1, p. 23. Hamburg, Germany, DESY (1992)

[16] P. Billoir, R. Fruehwirth, M. Regler, Nucl. Inst. Meth. **A 241**, 115 (1985)

[17] E. Maddox, *Study of heavy quark production at HERA using the ZEUS microvertex detector*, PhD Thesis, NIKHEF, Amsterdam, Netherlands (2004)

[18] H. Stadie, *New Vertex Tools and How to Use Them* ZEUS presentation (2006) (unpublished)

[19] S. Catani, Y.L. Dokshitzer, B.R. Webber, *The K-perpendicular clustering algorithm for jets in deep inelastic scattering and hadron collisions.* Phys. Lett. **B 285** 291 (1992)

[20] A.G. Yagües Molina, *Study of Beauty Photoproduction with the ZEUS Experiment at the Electron-Proton Collider HERA*, PhD Thesis, Humboldt-University Berlin (2008)

[21] ZEUS Collaboration, *Measurement of beauty production from inclusive secondary vertices in DIS and $F_2^{b\bar{b}}$ extraction at ZEUS*, Conference Proceeding, XXXVth International Conference on High Energy Physics, Paris, France (2010)