# Measurement of charm production in deep inelastic scattering using lifetime tagging for $D^{\pm}$ meson decays with the ZEUS detector at HERA

Dissertation
zur Erlangung des Doktorgrades
des Departments Physik
der Universität Hamburg

vorgelegt von
Mykhailo Lisovyi
aus Berditschev / Kiew (Ukraine)

Hamburg
2011

| | |
|---|---|
| Gutachter/in der Dissertation: | Prof. Dr. R. Klanner<br>Dr. A. Geiser |
| Gutachter/in der Disputation: | Prof. Dr. J. Haller<br>Dr. T. Schörner-Sadenius |
| Datum der Disputation: | 17.10.2011 |
| Vorsitzender des Prüfungsausschusses: | Dr. G. Steinbrück |
| Vorsitzender des Promotionsausschusses: | Prof. Dr. P. Hausschildt |
| Leiter des Departments Physik: | Prof. Dr. D. Pfannkuche |
| MIN-Dekan des Fachbereichs Physik: | Prof. Dr. H. Graener |

**Abstract**

A measurement of charm production in deep inelastic scattering at $\sqrt{s} = 318\,\mathrm{GeV}$ at HERA is presented in this thesis. The analysed data were collected with the ZEUS detector during 2005 – 2007, corresponding to the integrated luminosity of $323\,\mathrm{pb}^{-1}$. Charm production events were identified by the reconstruction of $D^{\pm}$ mesons in the $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ decay channel. The phase space of the analysis is defined by $5 < Q^2 < 1000\,\mathrm{GeV}^2$, $0.02 < y < 0.7$, $1.5 < p_{\mathrm{T}}(D^{\pm}) < 15\,\mathrm{GeV}$ and $|\eta(D^{\pm}| < 1.6$, where $Q^2$ is the photon virtuality, $y$ is the inelasticity and $p_{\mathrm{T}}(D^{\pm})$ and $\eta(D^{\pm})$ are the $D^{\pm}$ transverse momentum and pseudorapidity. Lifetime information was exploited to substantially reduce combinatorial background originating from light flavour production. Single– and double–differential cross sections were measured and compared to next–to–leading order QCD predictions as well as to previously published ZEUS measurements. The charm contribution to the proton structure function $F_2$, $F_2^c$, was extracted.

**Kurzfassung**

Diese Dissertation beschreibt die Charm–Produktion in der tiefunelastischen Streuung bei einer Schwerpunktsenergie von $\sqrt{s} = 318\,\mathrm{GeV}$ bei HERA. Die analysierten Daten wurden in den Jahren 2005 – 2007 mit dem ZEUS–Detektor aufgezeichnet, was einer integrierten Luminosität von $323\,\mathrm{pb}^{-1}$ entspricht. Die Charm–Ereignisse wurden durch die Rekonstruktion von $D^{\pm}$ Mesonen im Zerfallskanal $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ identifiziert. Der kinematische Bereich der Auswertung wird durch $5 < Q^2 < 1000\,\mathrm{GeV}^2$, $0.02 < y < 0.7$, $1.5 < p_{\mathrm{T}}(D^{\pm}) < 15\,\mathrm{GeV}$ und $|\eta(D^{\pm}| < 1.6$ definiert, wobei $Q^2$ die Virtualität des Photons, $y$ die Inelastizität und $p_{\mathrm{T}}(D^{\pm})$ und $\eta(D^{\pm})$ der Transversalimpuls und die Pseudorapidität des $D^{\pm}$ Mesons sind. Information über die Lebensdauer wird angewendet, um den aus der Produktion leichter Quarks stammenden kombinatorischen Untergrund erheblich zu reduzieren. Einfach– und doppeltdifferentielle Wirkungsquerschnitte wurden gemessen und mit QCD–Vorhersagen der nächstführenden Ordnung sowie mit früheren ZEUS–Messungen verglichen. Der Charm–Beitrag zur Protonstrukturfunktion $F_2$, $F_2^c$, wurde extrahiert.

# Contents

# Introduction

At the beginning of the XX century E. Rutherford established an intrinsic structure of the atom by scattering $\alpha$ particles on a thin gold foil. Rutherford concluded that an atom is built of a small massive nucleus and electrons orbiting around it. Since then the idea of using scattering of a probing particle to reveal the structure of a target has been exploited in many experiments in high energy physics, such as at SLAC, DESY, CERN or FNAL.

The world-only electron[1]–proton collider HERA was built to provide a deep insight into the structure of the proton. With the available photon virtuality, $Q^2$, going up to $\sim 30000\,\mathrm{GeV}^2$, the best spatial resolution is of the order of $10^{-18}\,\mathrm{m}$, which corresponds to 1/1000 of the proton size. In the modern theory of the strong interactions, Quantum Chromodynamics (QCD), it is very hard to predict soft dynamics inside of the proton, since at low energy scales the corresponding strong coupling constant, $\alpha_S$, becomes large and usual perturbative techniques are inapplicable. Therefore the probability distributions to find a quark or gluon in the proton, parton density functions (PDFs), have to be extracted from data. In this respect measurements at HERA are particularly valuable and nowadays every PDF fit includes HERA data.

So far gluon PDFs at low parton momentum fraction $x$ are predominantly constrained by measurements of the Bjorken scaling violations of the inclusive structure function $F_2$. On the other hand, heavy quark ($c$ or $b$) production provides a direct probe of the gluon distribution, since it appears in the initial state for this process. In addition, measurements of charm quark production (beauty production is significantly suppressed due to larger mass and smaller electric charge) allow to test and tune various models that are used for the heavy flavour treatment in the PDF fits. Besides that, charm production is not only important to constrain the soft QCD dynamics in the proton, but also allows to test the hard QCD regime. At high energy scales the strong coupling constant is small, but not small enough that terms above

[1]Hereafter both electrons and positrons are referred to as electrons.

the leading order can be neglected. Therefore, the contribution of higher orders is important and the heavy quark production cross sections allow to test predictions of the perturbation series. Moreover, the test is even more important due to multiple hard scales in the process (photon virtuality, quark transverse momentum and mass), resulting in possibility of spoiling the series convergence due to large logarithms.

To study open-charm final states a full reconstruction of the $D^+ \to K^-\pi^+\pi^+$ decay[2] in the ZEUS detector was used to tag charm production events. This channel was chosen due to a significant lifetime of $D^+$ mesons, which allowed to reconstruct the secondary decay vertices to suppress the light flavour (*u, d, s*) background in the data drastically, keeping selection efficiency for the signal events high. A similar analysis was already performed on $134\,\mathrm{pb}^{-1}$ of the data collected with the ZEUS detector in 2005 [1]. The current analysis benefits from a $\sim 2.5$ times larger data sample as well as better control of systematic uncertainties, profiting from an improved tracker alignment and calibration.

This thesis is devoted to the analysis of charm production. For the technical task, the author also contributed to the Monte Carlo production system in ZEUS in general and to the usage of the GRID resources in particular. This involved maintenance of the existing system as well as software development and planning of future strategies. These results will be covered in Chapter 4.

The thesis is organised in the following way: in Chapter 1 an overview of the theoretical framework for lepton-proton scattering is presented together with an overview of the recent results on charm production from ZEUS and H1. In the following Chapter 2 the next-to-leading-order (NLO) QCD calculations for charm production are described. Chapter 3 is dedicated to the HERA accelerator and the ZEUS detector. Chapter 4 concentrates on the Monte Carlo simulations in general and a contribution of the author to the ZEUS Monte Carlo production system in particular. Event reconstruction is described in Chapter 5. Chapter 6 is devoted to the event and $D^+$ selection. The measured cross sections together with the systematics studies are presented in Chapter 7. In Chapter 8 a detailed review of the extraction of the charm component of the structure function, $F_2^c$, is given. Finally, conclusions are drawn in Chapter 9.

[2]Charge conjugation is always implied.

# Chapter 1

# Heavy quark production in Deep Inelastic Scattering at HERA

In this Chapter the theoretical framework as well as the previous measurements of charm production are presented. After a short introduction to the Standard Model (SM) and Quantum Chromodynamics, the kinematics of *ep* interactions is described. Then the development of the proton structure description is reviewed. After an overview of the approaches to the charm production treatment, focus is moved towards experimental results of charm production at HERA. At the end the impact of the charm data on the proton structure is discussed.

## 1.1 The Standard Model

A widely accepted and well proved theory of elementary particle physics is the Standard Model.

According to the Standard Model all matter consists of 6 quarks (up ($u$), down ($d$), strange ($s$), charm ($c$), beauty or bottom ($b$) and ($t$)) and 6 leptons (electron ($e$) and electron neutrino ($\nu_e$), muon and muon neutrino ($\nu_\mu$), tau lepton ($\tau$) and tau neutrino ($\nu_\tau$)) and the interactions between them are mediated by 4 gauge bosons ($Z^0$, $W^\pm$, $\gamma$, $g$). Quarks and leptons are fermions with spin 1/2, which are grouped into 3 families. The Standard Model comprises 3 types of fundamental forces: electromagnetic, weak and strong. The carrier of gravitation, which is not described in the SM, is supposed to be

the so-called graviton, a boson with spin 2, but it wasn't found in an experiment so far. The electromagnetic interactions act on electrically charged particles and are mediated by electrically-neutral massless photons. These processes are described within the framework of Quantum Electrodynamics (QED). The gluon is the mediator of the strong force and it couples to particles with a colour charge. In contrast to photons, gluons themselves carry a colour charge. In analogy to QED the theory of strong interactions is called Quantum Chromodynamics (QCD). The two remaining gauge bosons, $Z^0$ and $W^{\pm}$, represent the weak interactions. The Higgs mechanism of the local symmetry breaking had to be introduced in the Standard Model to produce mass for the $Z^0$ and $W^{\pm}$ bosons. This results in one additional boson with spin 0 – the Higgs boson, H; interactions with the Higgs field are supposed to give particles their masses. So far it was not observed in experiments at the LEP[1] [2], TEVATRON [3] and the recently started LHC[2] [4, 5] colliders.

The Standard Model provides us with precise and robust predictions for various processes that were verified in many high energy physics (HEP) experiments. However the SM seems not to be "the final theory of everything", because there are observations that go beyond the scope of the SM, such as: neutrino oscillations have shown that neutrinos are not massless [6], the forth interaction, gravitation, can not be included in the theory, composition of the universe (dark matter [7] and dark energy [8]), etc. This led to the development of many Beyond the Standard Model (BSM) theories (e.g. see [9] for a review of the BSM search strategies with the early LHC data), but these are out of the scope of this thesis.

Quantum Chromodynamics is the quantum field theory of the strong interactions, which is based on the $SU(3)$ group symmetry. In contrast to QED its mediator, the gluon, carries colour charge, resulting in unique properties of the theory. Two important consequences of the gluon self-interactions are:

- Confinement: Isolated quarks and gluons can not be observed. With increasing distance between two colour charges, the energy stored in the gluon field increases and at some point it reaches the threshold of quark-antiquark pair production. This process goes on until there is no colour connections between objects that fly apart. Thus only colourless hadrons – combinations of 2 or 3 quarks – are found as free particles.
- Asymptotic freedom: On the other hand, the strength of the strong interaction decreases with decreasing length scales. This leads to the fact, that at small distances quarks and gluons become quasi free particles.

[1] Large Electron Positron collider at CERN
[2] Large Hadron Collider at CERN

These features are also reflected in the running of the strong coupling constant, which, in the leading order approximation, is given by:

$$\alpha_s(\mu_R^2) = \frac{12\pi}{(33 - 2n_f) \cdot ln(\mu_R^2/\Lambda_{QCD}^2)}, \tag{1.1}$$

where $\mu_R$ is called the renormalisation scale, $n_f$ is the number of active flavours at the given energy scale $\mu_R$, and $\Lambda_{QCD}$ is the so-called *QCD energy scale*. Thus with $\mu_R$ decreasing and approaching $\Lambda_{QCD}$, the strong coupling constant becomes large and calculations in the perturbative QCD (pQCD) framework do not converge. The parameter $\Lambda_{QCD}$ is a fundamental constant that should be extracted from experimental data, but, since the calculations for extraction are done up to a fixed order in pQCD within some scheme, the value might depend on the calculation. Typical values of $\Lambda_{QCD}$ are of the order of $\approx 200\,\text{MeV}$.

## 1.2 Deep Inelastic *ep* Scattering

In the SM there are two classes of *ep* interactions depending on the exchanged boson:

- if an electron interacts with a proton by exchanging an electrically-neutral $Z^0$ or $\gamma$ boson, such processes are called Neutral Current (NC). In this case the scattered electron can be detected in the experiment, if the scattering angle is sufficiently large.
- if the interaction occurs via the exchange of a charged $W^{\pm}$ boson, such processes are called Charged Current (CC). In this situation the electron transforms into a neutrino, which escapes the detector invisibly.

Due to the high masses of $W^{\pm}$ and $Z^0$ bosons [10], in most of the HERA kinematic phase space the scattering occurs by the exchange of $\gamma$ bosons.

Figure 1.1 explains the kinematics of NC *ep* interactions. On this diagram $k$ and $k'$ are the four-momenta of the incoming and the scattered electrons, respectively, whereas $p$ is the four-vector of the incoming proton. The variable $q$ denotes the four-momentum of the intermediate boson and can be expressed as $q = k - k'$. To describe the scattering 4 Lorentz-invariant scalars are used:

$$Q^2 = -q^2 = -(k - k')^2, \tag{1.2}$$

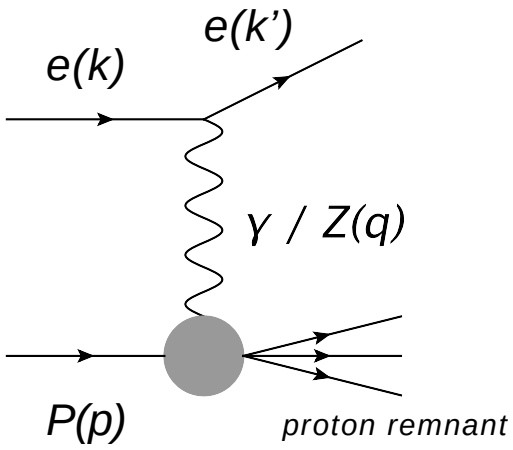


**Figure 1.1:** *A typical diagram of the NC* ep *scattering.*

$$x = \frac{Q^2}{2\,p \cdot q}, \tag{1.3}$$

$$y = \frac{q \cdot p}{k \cdot p}, \tag{1.4}$$

$$s = (k+p)^2, \tag{1.5}$$

where $Q^2$ is called *boson virtuality* and determines the distance scale on which the proton is probed, $s$ is the total *ep* center-of-mass energy squared, $y$ defines the fraction of the lepton's energy transfered to the proton in the proton's rest frame and is called *inelasticity* and $x$ is the *Bjorken scaling variable*. In the quark-parton model it defines the fraction of the proton's momentum carried by a parton. These four variables are correlated and only 3 of them are independent. For the kinematics of the HERA accelerator one can neglect the proton and electron masses ($\approx 1\,\mathrm{GeV}$ and $\approx 0.0005\,\mathrm{GeV}$ respectively) in comparison with their momenta($\approx 920\,\mathrm{GeV}$ and $\approx 27.5\,\mathrm{GeV}$ respectively). Thus from (1.5) using (1.3) and (1.4) one can derive a simple relation:

$$s = (k+p)^2 \approx 2\,k \cdot p = 2\,\frac{q \cdot p}{y} = \frac{Q^2}{x\,y} \tag{1.6}$$

Since the center-of-mass energy is fixed for the fixed beam energies, the event kinematics is completely described by two variables only. Typically the pairs $(Q^2, y)$ or $(Q^2, x)$ are chosen.

The *ep* scattering phase space is subdivided into two regions: for $Q^2 > 1\,\mathrm{GeV}^2$ processes are referred to as *Deep Inelastic Scattering (DIS)*. Events with $Q^2 \ll 1\,\mathrm{GeV}^2$ are called *Photoproduction (PHP)*.

## 1.3 Quark-Parton Model

In the simple *quark-parton model (QPM)* [11], proposed by R. Feyman, the proton is built of a set of point-like objects, *partons*, which do not interact with each other. In this model *ep* scattering can be presented as the elastic scattering of the electron on one of the partons. One can derive a simple relation for the fraction of the proton momentum carried by a parton, $\xi$, using momentum conservation and neglecting the proton and parton masses, $M$ and $m$ respectively:

$$0 \approx m^2 = (\xi\, p + q)^2 = \xi^2\, M^2 + 2\,\xi\, p \cdot q + q^2 = \xi \frac{Q^2}{x} - Q^2; \Rightarrow \xi = x$$

Hence within the QPM approximation the Bjorken $x$ corresponds to the fraction of the proton momentum carried by the parton in the hard scattering. In the QPM Bjorken has predicted that the proton structure functions $F_1$ and $F_2$ (see Section 1.4) depend only on $x$ and are independent of $Q^2$, which is known as *Bjorken scaling* [12]:

$$F_1(x) = \frac{1}{2} \sum_i e_i^2 f_i(x), \tag{1.7}$$

$$F_2(x) = \sum_i e_i^2 x f_i(x), \tag{1.8}$$

where the sum runs over all partons in the proton, $e_i$ is the charge of the i-th parton in units of the elementary charge and $f_i$ are the *parton distribution functions (PDFs)*, which define the probability to find in a proton a parton of the i-th type with the momentum fraction $x$. For the first time this was proved in experiments on DIS at SLAC [13]. Partons in this model correspond to the *quarks* suggested by Gell-Mann [14].

However, this approximation was too simplified to fully describe nature. If there would be only three so–called *valence* quarks in the proton ($uud$), then, according to the momentum conservation law, the sum of their fractional momenta would be equal to unity:

$$\sum_i \int_0^1 f_i(x) x dx = 1. \tag{1.9}$$

In contrast, it was found experimentally that this quantity is close to 0.5, hence approximately 50% of the proton momentum is carried by neutral particles [15]. These turned out to be the gluons of QCD. The first direct

experimental evidence of gluons was found in events with three hadronic jets in the $e^+e^-$ collisions at PETRA [16]. The discovery of gluons established QCD as the theory of strong interations. In the QCD framework protons are filled with low-$x$ gluons, which can split into quark-antiquark pairs producing a *quark sea*. Consequently, the Bjorken scaling is only approximate and it is broken in particular in the low-$x$ region (see Section 1.4).

## 1.4 Proton structure

The NC DIS cross section can be written in terms of the structure functions, $F_2$, $F_L$ and $F_3$:

$$\frac{d\sigma(e^{\pm}p)}{dx\,dQ^2} = \frac{2\pi\alpha^2}{x\,Q^4}\Big((1+(1-y)^2)\,F_2 - y^2\,F_L \mp x\,(1-(1-y)^2)\,F_3\Big), \quad (1.10)$$

where $\alpha$ is the electromagnetic coupling constant. The structure function $F_2$ is the dominant contribution to the neutral current scattering in the kinematic regime of the measurement presented here. $F_3$ describes the parity violation arising from $\gamma Z^0$ interference, therefore for $Q^2 \ll M_Z^2$ this contribution is suppressed and can be neglected. The *longitudinal structure function* $F_L$ parametrises the contribution from coupling to the longitudinally polarised photons. The contribution of $F_L$ to the NC *ep* cross section is suppressed for $y^2 \ll 1$, but can be up to a few percent in the kinematic region of the current measurement ($0.02 < y < 0.7$) and thus can not be neglected.

As it was mentioned in Section 1.3, in contrast to the QPM the structure functions in QCD are no longer functions of a single variable $x$. Due to continuous emissions, absorptions and splitting of gluons the proton is a dynamic system. At large energies of the intermediate boson the probed length-scale is decreasing and thus more partons at small $x$ become visible. This feature is know as *scaling violation* [17]. As a consequence, the structure function $F_2$ has a logarithmic dependence on $Q^2$ at small $x$ [18]. This dependence on $Q^2$ was confirmend by experiments at HERA [19]. Figure 1.2 shows the recent measurements of the reduced cross section for the positron-proton NC scattering, $\sigma^+_{r,\mathrm{NC}}$. For both electrons and positrons the reduced cross section is defined as

$$\sigma^{\pm}_{r,\mathrm{NC}}(x,Q^2) = \frac{d\sigma(e^{\pm}p)}{dx\,dQ^2}\cdot\frac{x\,Q^4}{2\pi\alpha^2 Y_+} = F_2 - \frac{y^2}{Y_+}\,F_L \mp \frac{Y_-}{Y_+}\,xF_3, \quad (1.11)$$

where $Y_\pm = (1\pm(1-y)^2)$, as a function of $Q^2$ and $x$. The data at the medium-$x$ region ($x \sim 0.1$) agree with the Bjorken scaling hypothesis, whereas data at

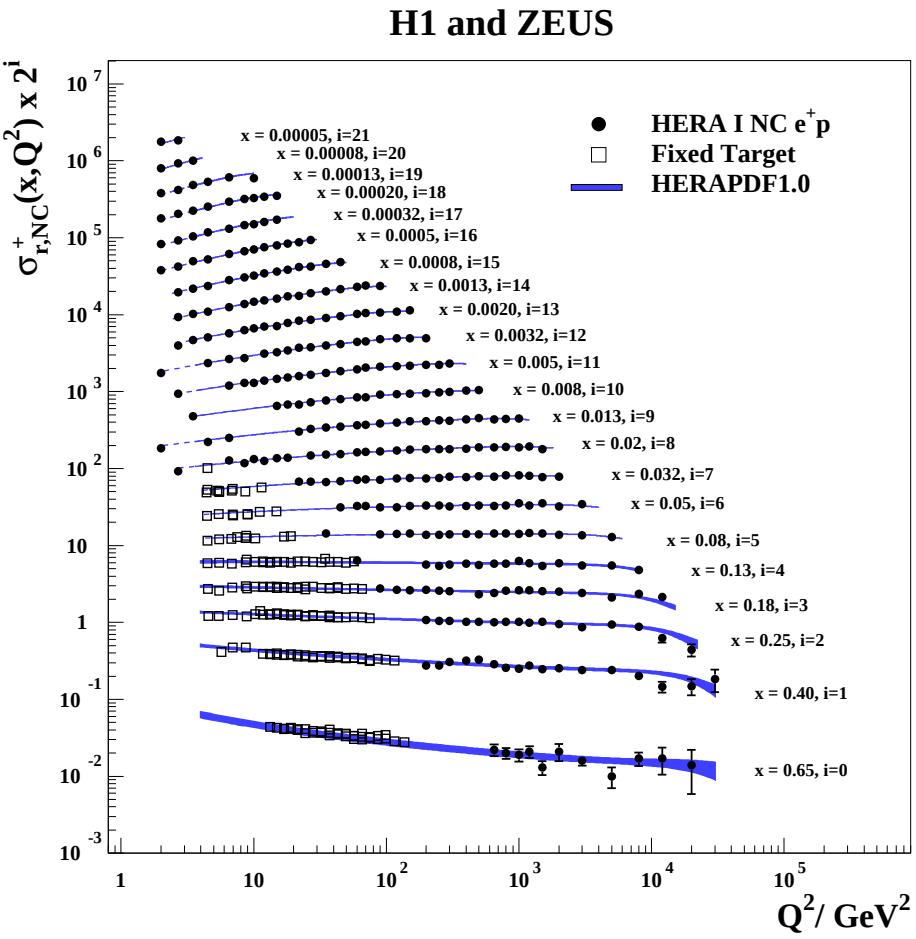


**Figure 1.2:** *The neutral current ep scattering cross section as a function of $Q^2$ at the fixed values of Bjorken variable $x$ [20]. The combination of the H1 and ZEUS measurements (black dots) is compared to the measurements at the various fixed-target experiments (open squares) and overlaid with the NLO QCD predictions based on HERAPDF1.0 [20]. Note that for the i-th value of $x$ the measured and predicted values of $\sigma_r$ are scaled by $2^i$.*

low-$x$ and high-$x$ values show clear signatures of scaling violation. Compared to experiments colliding electron on a fixed target, the HERA data cover a much broader range in $x$ and reach higher values of $Q^2$.

According to the *factorisation theorem of QCD* [21] the cross section of the neutral current DIS process $ep \to e' + X$ can be subdivided on the level

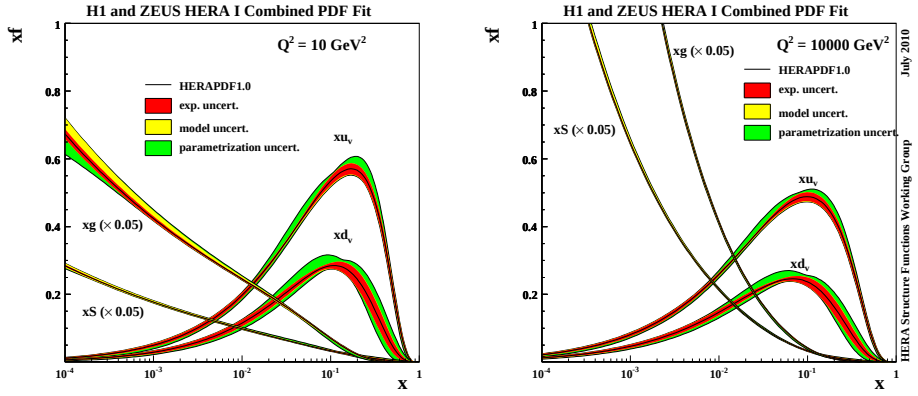


**Figure 1.3:** *The parton density functions extracted in the HERA next-to-leading-order QCD fit [20] at $Q^2 = 10\,GeV^2$ (left) and $Q^2 = 10000\,GeV^2$ (right). The distributions for the valence quarks ($u_v$ and $d_v$) are shown together with the quark sea and gluon densities, the latter two were scaled down by a factor of 20 to make the plot more informative. Different sources of uncertainties are represented as bands of different colours.*

of partons into a hard partonic cross section, $\tilde{\sigma}_i$, that can be calculated up to a fixed order in pQCD, and a soft-QCD component parametrised by the parton density functions (PDFs), $f_i(x, Q^2)$, which define the probability to find a parton of type $i$ with the momentum fraction $x$ in a proton at the given energy scale $Q^2$. Thus the cross section can be written in the form:

$$\sigma(ep \to e' + X) = \sum_{i=q,\bar{q},g} \int_0^1 dx\, \tilde{\sigma}_i(x, \mu_F) \otimes f_i(x, \mu_F^2), \qquad (1.12)$$

where the sum runs over all quarks and gluons in the proton. Both $\tilde{\sigma}_i$ and $f_i$ depend on the *factorisation scale*, $\mu_F$, that sets the energy cutoff between the soft and the hard regimes. The PDFs are universal and thus can be extracted from one process and applied to another. To describe the energy dependence of the partonic densities, e.g. to transfer the results from HERA to the TEVATRON or LHC kinematic regime, the *QCD evolution equations* can be used. The NLO QCD calculations and the Monte Carlo simulations used in this thesis are based on the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [22], which predict the evolution in $Q^2$. Figure 1.3 shows the PDFs as extracted from the combined ZEUS and H1 deep inelastic scattering data at the different energy scales [20]. As expected, the valence quarks dominate at high $x$, however the gluon and the quark sea

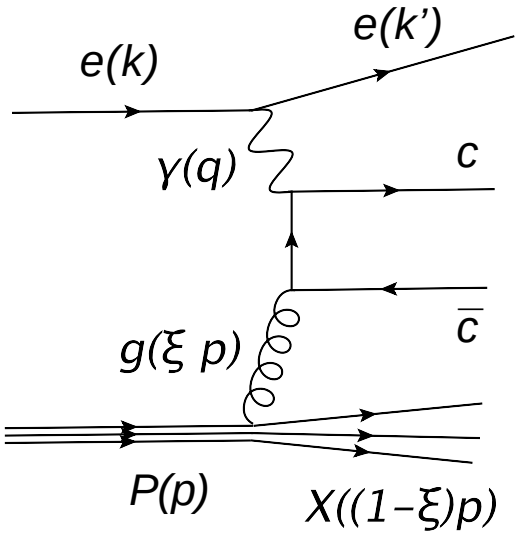


**Figure 1.4:** *The leading-order charm production diagram in NC DIS.*

contributions start to dominate at $x < 10^{-1}$. At higher $Q^2$ the valence quark distributions barely change, whereas the number of the low-$x$ gluons and sea quarks has significantly increased – the origin of the scaling violation.

## 1.5 Heavy flavour treatment

The dominant $c\bar{c}$ production mechanism in neutral current DIS is the so-called *boson-gluon fusion (BGF)*. Figure 1.4 shows the Feynman diagram in leading-order pQCD for this process. One of the gluons from the proton participates in the hard interaction with the momentum fraction of $\xi$. Thus charm production is directly sensitive to the gluon density inside the proton. And consequently, charm measurements have constraining power on the gluon PDFs.

There are various approaches to the heavy quark, e.g. charm quark, production treatment in the QCD calculations:

- *Fixed Flavour Number Scheme (FFNS)*: The charm quarks are treated as massive and are produced only perturbatively in the hard scattering. Consequently, there are only light quarks ($u$,$d$,$s$, in the case of $b\bar{b}$ production the $c$ quark is also treated as "light") and gluons inside the proton. Due to collinear gluon radiation, the convergence of the perturbation series could be spoiled at high $Q^2$ values by the presence

of terms proportional to $log(\frac{Q^2}{m_c^2})$. Thus this approximation is expected to be the most reliable in the region $Q^2 \sim m_c^2$.

- *Zero-Mass Variable Flavour Number Scheme (ZM-VFNS)*: The $c$ quark mass is neglected, but treated by a threshold function – charm production is suppressed for $Q^2 < \mu_c^2$ (the usual choice is in the range $m_c^2 < \mu_c^2 < 4m_c^2$), from $Q^2 = \mu_c^2$ onward charm is treated as an active massless flavour. The large logarithms are resumed in this approach. Thus this approximation is correct for $Q^2 \gg m_c^2$ and has a rather incorrect behaviour in the threshold region.

- *General-Mass Variable Flavour Number Scheme (GM-VFNS)*: This scheme has the behaviour of the FFNS at low $Q^2$ and follows the ZM-VFNS at high values of $Q^2$. At intermediate energy scales it interpolates between the two. At NLO various GM-VFNS have significant ambiguities in modelling the interpolation region leading to difference in the corresponding cross section predictions. Therefore, heavy quark cross section measurements are very important to constrain the charm production treatment in the GM-VFNS.

Experimentally it was proven that for the kinematic regime of HERA the predictions based on the ZM-VFNS treatment do not describe the data [23, 24]. In contrast, the FFNS calculations have shown nice agreement with measurements in the full accessible phase space [1, 23–30]. The charm production NLO QCD predictions in the GM-VFNS exist only for the inclusive cross sections and are in nice agreement with the data [23, 24, 29, 31].

## 1.6 Heavy flavour tagging

Various experimental techniques have been exploited by the HERA experiments to tag heavy quark production. Predominantly, large invariant mass and relatively long lifetime are used as a signature of the heavy quark events. All techniques that have been used can be subdivided into three classes: full reconstruction of the open-charm mesons, identification of semileptonic decays and inclusive tracking techniques.

- When open-charm meson, e.g. $D^+, D^0, D^{*+}, D_s$, decays are explicitly reconstructed, a clear peak is observed in the spectrum of invariant mass, see Fig. 1.5 for examples. This method has the smallest statistics compared to others, as a consequence of the selection of a specific

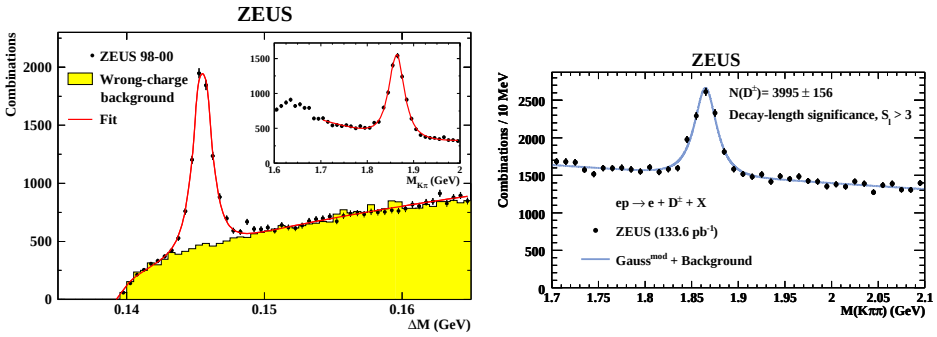


**Figure 1.5:** *Reconstructed invariant mass difference $M(D^*)-M(D^0)$ in $D^* \to D^0\pi^+$ decays [26] (left) and invariant mass spectra for $D^+ \to K^-\pi^+\pi^+$ decays [1] (right). A fit of the signal and the background functions to the data is shown as a smooth line.*

fragmentation and decay chain. The typical probabilities for a charm quark to fragment into a specific D meson, $f(c \to D)$, are $\approx 0.15-0.25$ and the branching fractions, $BR$, for the commonly used decay channels are $\approx 0.05-0.10$, thus only $1-2.5\%$ of all $c$ quarks can be tagged. The advantages of this method are its simplicity and independence of the light flavour background modelling.

- Reconstruction of the semileptonic decays comprises the identification of electrons or muons together with a signature of neutrino (imbalanced or "missing" transverse momentum $p_T^{miss}$) and a jet in an event. A fit of contributions of various flavours is performed on a set of discriminative variables. In particular charm events have distinguishable distributions of the missing transverse momentum parallel to the lepton flight direction, $p_T^{miss\|l}$, and the signed impact parameter for the lepton track, $\delta$. The impact parameter is defined as the distance of closest approach of the track to the primary point of interaction and its sign is determined by the location of the lepton track with respect to the jet it is assigned to. An example of such a measurement is presented in Fig. 1.6. Identification of the semileptonic decays is also sensitive to beauty production, hence both charm and beauty cross sections are extracted simultaneously. In addition, the distribution of the lepton transverse momentum relative to the jet, $p_T^{rel}$, was used to differentiate between charm and beauty. This approach is sensitive to the light flavour background modelling and has rather low separation power between charm and light flavours.

- Inclusive methods identify heavy flavour production by measuring dis-

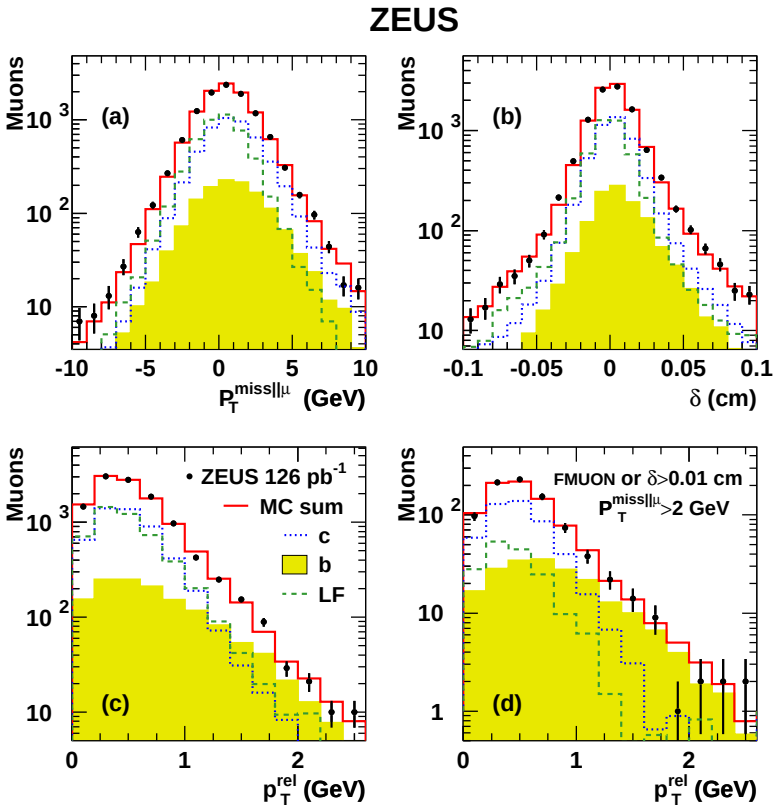


**Figure 1.6:** *Distributions of the discriminating variables from [28]: (a) missing transverse momentum parallel to the muon direction, $p_T^{miss||\mu}$, (b) signed impact parameter, $\delta$, (c) transverse momentum relative to a jet, $p_T^{rel}$, (d) $p_T^{rel}$ distribution for a heavy-flavour-enriched sample. The data (points) are overlaid with the Monte Carlo simulations (solid line). The charm, beauty and light flavour contributions are shown separately.*

placed secondary vertices with large reconstructed mass and long decay length or a certain number of tracks that have large impact parameters. These two approaches use the distributions of signed significances of either 2-dimensional decay length, $S_l$, which will be discussed in detail in Section 6.3.2, or the impact parameter, $S$. The significance is defined as the ratio of a value by its uncertainty. The light flavour distribution is (almost) symmetric and originates from the detector resolution, whereas the charm and beauty distributions have large asymmetric tail for positive values of $S_l$ and $S$. Thus the subtraction of the mirrored

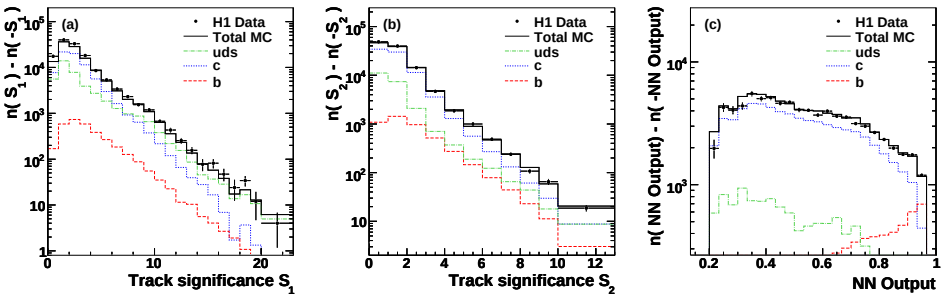


**Figure 1.7:** *The mirrored distributions of the impact parameter significances for two tracks with the highest impact parameter significances, $S_1$ and $S_2$, together with the neural network output distribution, which takes as an input various tracking information (for details see [32]). Separate flavour components were normalised according to the template fit output and are shown as histograms of different styles and colours.*

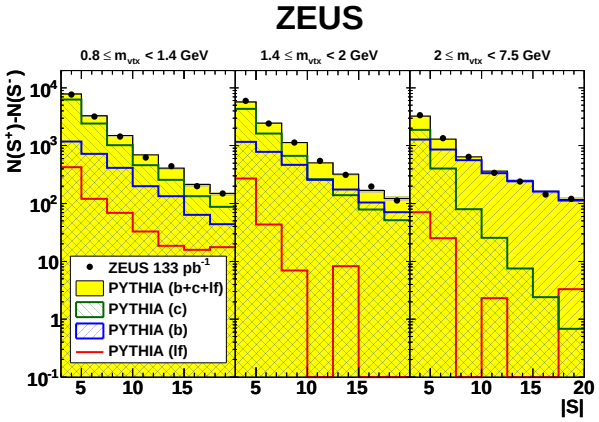


**Figure 1.8:** *Distribution of the subtracted decay length significance in three mass bins [33]. The data (black points) are compared to the fitted sum of beauty, charm and light flavour contributions (filled yellow histogram) as well as the distributions from the separate subsamples.*

negative part from the positive tail significantly reduces the dependence on the light flavour background. Examples of measurements with this technique are presented in Fig. 1.7 and Fig. 1.8. These techniques are similar to those used in the LHC experiments for $b$-tagging (e.g. [34]). The the inclusive techniques provide access to the largest statistics.

These methods are complementary as far as systematic uncertainties are concerned. Therefore all measurements performed with different techniques are important and they can be combined to further reduce statistical and systematic uncertainties.

## 1.7 Selected charm measurements in DIS

There have been many measurements of heavy quark production at HERA. Here the emphasis will be put on the charm measurements in deep inelastic scattering.

The most widely used charm tag is based on $D^{*\pm}$ mesons in the so-called *golden decay channel*. This is a strong decay mode $D^* \to D^0\pi^+$ with the subsequent $D^0$ decay $D^0 \to K^-\pi^+$. It features a narrow peak in the $M(D^*) - M(D^0)$ distribution (see Fig. 1.5). This decay channel was used by both the ZEUS and H1 collaborations in the analyses of HERA I[3] data [25–27] and HERA II[4] data [23, 24]. For these measurements it is important to have a reliable reconstruction of the low-momentum pion from the $D^*$ decay. On the other hand, one can use other D mesons exploiting the significant lifetime due to weak decay. This strategy was used in the analysis of $D^+$ and $D^0$ production performed on $134\,\mathrm{pb}^{-1}$ of data in ZEUS [1] (see Fig. 1.5). The selection of candidates with large decay length significance allowed to enrich the sample with charm. The kinematic phase space of charm meson measurements is defined in terms of $p_T(D)$ and $\eta(D)$. A compilation of the measured $D^{*+}$ and $D^+$ meson production cross sections is presented in Fig. 1.9. The data are well described by the NLO QCD predictions in the FFNS. Figure 1.9 *a)* also shows that the data are not consistent with the NLO QCD calculations in the ZM-VFNS.

Muons have also been used to measure charm at HERA. Figure 1.6 shows a selection of the kinematic variables used in [28] to disentangle charm and beauty events from the light-flavour background. The charm component predicted by the Monte Carlo simulations has a distinct shape in Fig. 1.6 *a, b*, which enables the extraction of the heavy flavour signal. Figure 1.10 shows the measurement of charm and beauty cross sections using muons. The kinematic region for the measurement was $Q^2 > 20\,\mathrm{GeV}^2$, $0.01 < y < 0.7$, $p_T^\mu > 1.5\,\mathrm{GeV}$ and $-1.6 < \eta^\mu < 2.3$. The charm cross sections are in good agreement with the NLO QCD calculations in the FFNS.

[3]The data taking period from 1996 – 2000; more details in Section 3.1
[4]The data taking period from 2003 – 2007; more details in Section 3.1

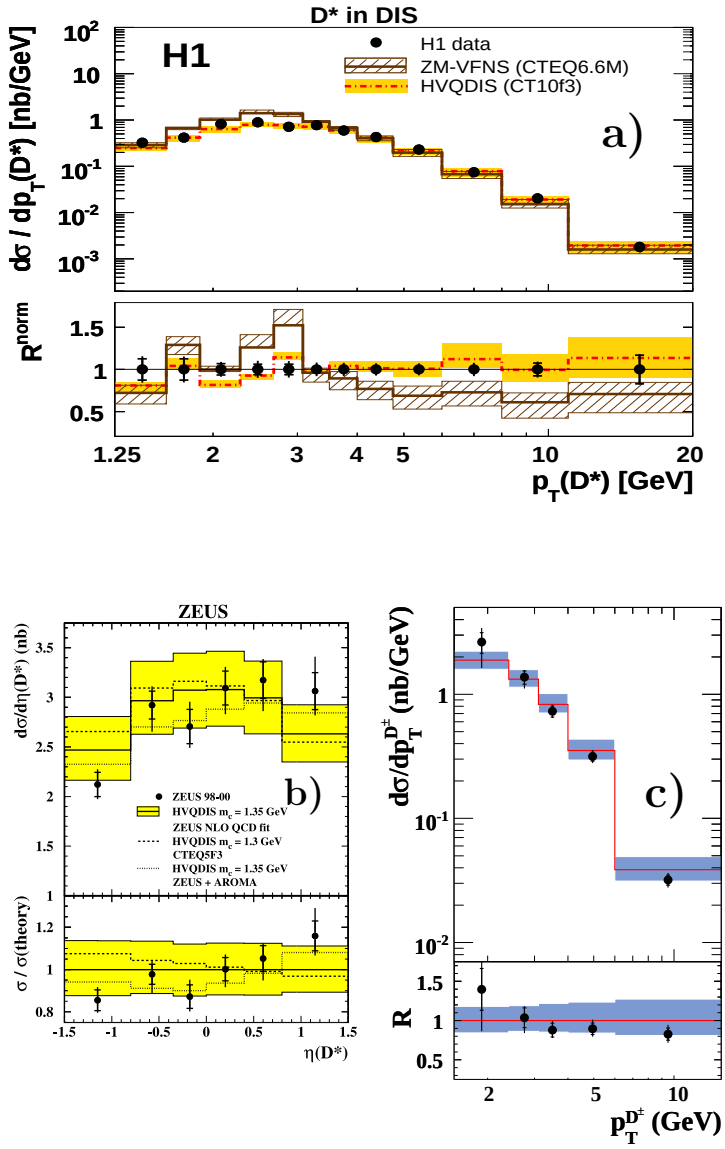


**Figure 1.9:** *Measured differential charm production cross section in (a) $p_T(D^{*\pm})$ [24], (b) $\eta(D^{*\pm})$ [26] and (c) $p_T(D^+)$ [1]. The data (black points) are compared to the NLO QCD predictions (filled bands). In addition, for each measurement the ratio of the measured to the predicted cross sections is shown.*

The inclusive techniques have been used by the H1 collaboration in publications based on the HERA I [29] and HERA II data samples [30, 32]. The method used in these analyses is based on a selection of jets with a certain

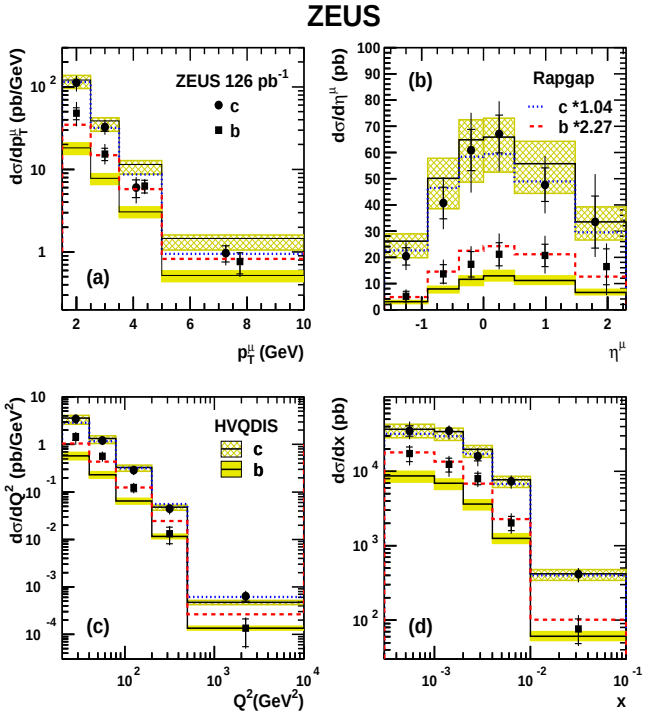


**Figure 1.10:** *Differential muon cross sections for c and b production as a function of (a) $p_T^\mu$, (b) $\eta^\mu$, (c) $Q^2$ and (d) x [28]. The bands show the NLO QCD predictions and the uncertainties. The differential cross sections from* RAPGAP*, scaled by the fit result (1.04 for c and 2.27 for b), are also shown.*

number of tracks with large impact parameter significance ($S = \delta/\sigma_\delta$). Figure 1.7 shows the distributions of the important discriminating variables. A simultaneous template fit of the predicted Monte Carlo components to the data was made to extract the charm and beauty fractions. Differential cross sections of the charm and beauty jet production were measured in the laboratory and Breit[5] frames. This measurement was done in a restricted phase space in the transverse energy of a jet in the laboratory and Breit frames, $E_T^{jet}$ and $E_T^{*jet}$ and in the pseudorapidity of the charm jet, $\eta^{jet}$. Figure 1.11 shows the comparison of the charm jet measurement with the NLO QCD predictions. The sensitivity to the scale choice was tested against the data. In ZEUS an inclusive analysis has been done in the photoproduction regime

[5]The Breit frame is defined by $2x\vec{p}+\vec{q}=0$, where $x$ is the Bjorken scaling variable, $\vec{p}$ and $\vec{q}$ are the three-momenta of the incoming proton and the exchanged boson, respectively.

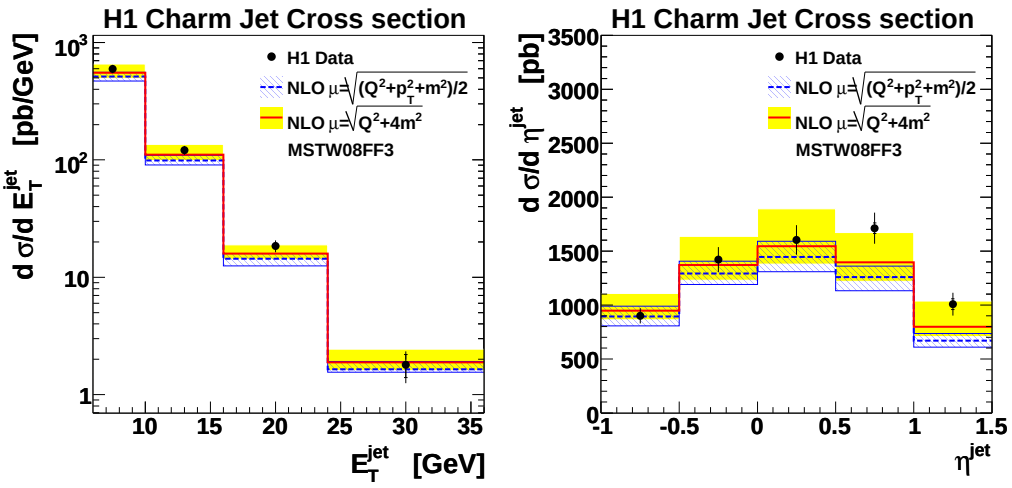


**Figure 1.11:** *The differential cross sections for the highest transverse energy jet in the laboratory frame as a function of $E_T^{jet}$ and $\eta^{jet}$ [30]. The measurements were made for the kinematic range $E_T^{jet} > 6\ GeV$, $-1 < \eta < 1.5$, $Q^2 > 6\ GeV^2$ and $0.07 < y < 0.0625$. The data are compared with the NLO QCD predictions where the bands indicate the theoretical uncertainties.*

[33], exploiting reconstruction of the displaced secondary vertices with high values of the mass and the decay length significance. A similar analysis is ongoing in DIS [35].

Overall the charm measurements in DIS at HERA are in good agreement with the NLO QCD predictions. The natural step is to combine the measurements obtained with the different analysis techniques to reach an ultimate precision and to gain constraining power on non-perturbative QCD effects, e.g. PDFs, and to cross-check consistency of the results. For this purpose an extrapolation to a common phase space is required. It is convenient to choose the full phase space to make the result independent of a particular hadronic final state. Usually the combined result is expressed in terms of either charm reduced cross sections or the charm contribution to the structure function, $F_2^c$, which is defined in analogy to the inclusive structure functions $F_2$ and $F_L$ (see (1.10)). In analogy to the inclusive structure functions, for $c$ quark production the charm contributions to the structure functions, $F_2^c$ and $F_L^c$, can be introduced, neglecting a small contribution from $F_3^c$:

$$\frac{d\sigma^{c\bar{c}}(e^{\pm}p)}{dx\,dQ^2} = \frac{2\pi\alpha^2}{x\,Q^4}\Big((1+(1-y)^2)\,F_2^c - y^2\,F_L^c\Big), \qquad (1.13)$$

Similar to the inclusive case, the term arising from $F_L^c$ is small in the common phase space of measurements at HERA, i.e. $y < 0.7$.

The values of $F_2^c$ extracted in each analysis can be combined. Figure 1.12 shows the individual measurements, which were used in the combination, together with the averaged $F_2^c$ values [31]. As a result of the combination, not only statistical, but also systematic uncertainties were reduced. These combined charm data provide a direct probe of the gluon density in the proton (see [36]). On the other hand, since so far only inclusive quantities have been calculated in the GM-VFNS, the $F_2^c$ measurements can be used to constrain the heavy flavour treatment in the GM-VFNS in PDF fits (see [37]).

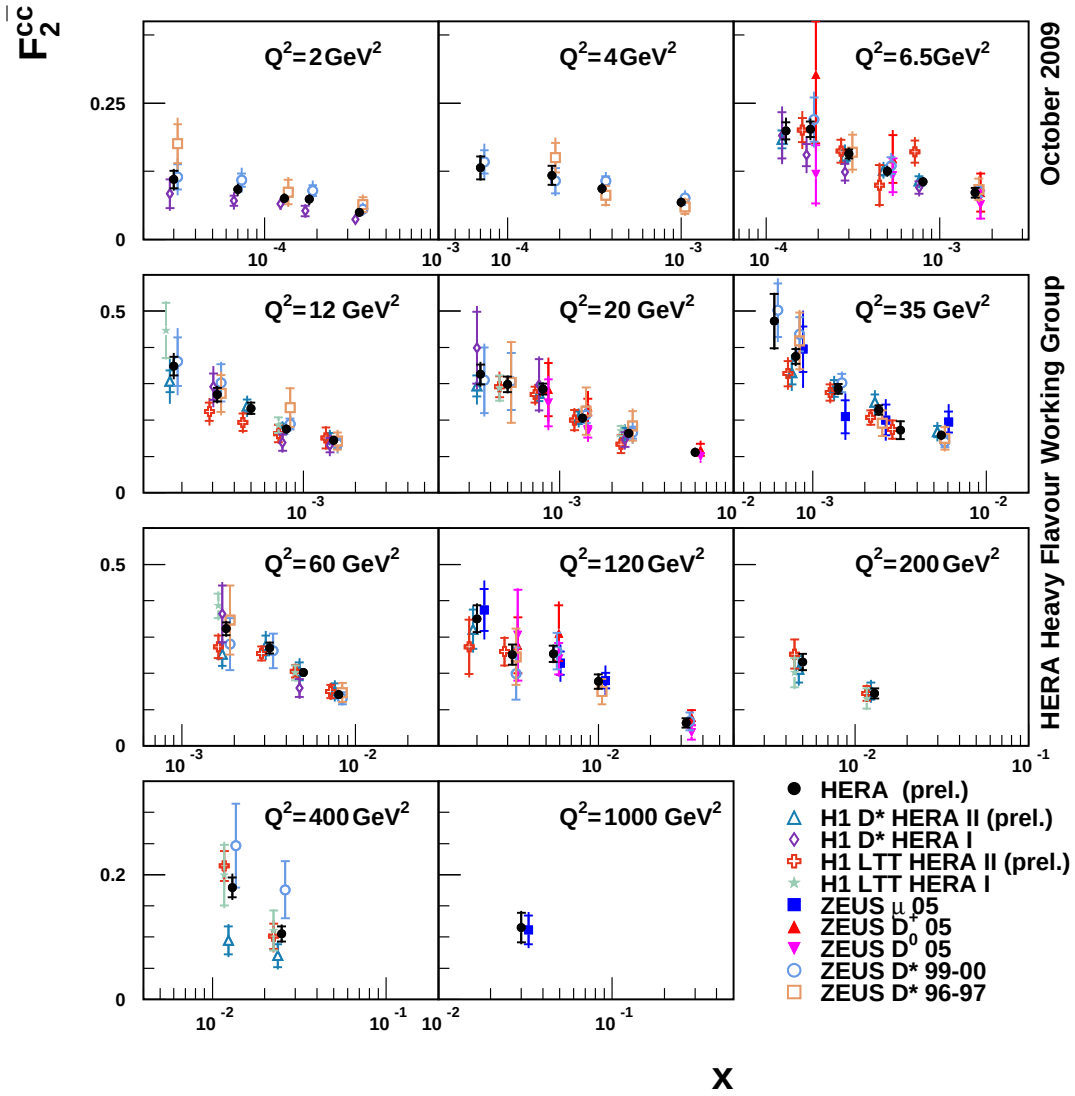


**Figure 1.12:** *Charm contribution to the proton structure function, $F_2^c$, as a function of $x$ at fixed values of $Q^2$ [31]. The averaged HERA result (black filled circles) is compared to the separate data sets from H1 and ZEUS (various coloured symbols, see the legend of the plot). The inner error bars of the averaged values represent the uncorrelated error, while the outer error bars show the total error.*

# Chapter 2

# Next-to-leading-order QCD predictions

The only available fully-differential NLO QCD predictions for heavy quark production in DIS are based on the FFNS approach (see Section 1.5) and are computed by the HVQDIS program [38]. The modified minimal subtraction, $\overline{\text{MS}}$, renormalisation scheme, which absorbs the divergent part of the NLO corrections plus a universal constant into counterterms, is used in HVQDIS to cancel out collinear and soft divergences. The program provides weighted events with either two or three outgoing partons, i.e. heavy quark pair plus possibly an additional light parton. More details on the calculation, for instance the included Feynman graphs, can be found in [39]. A detailed study of the charm fragmentation modelling was performed and results are presented in this Chapter.

## 2.1 HVQDIS setup

For the purpose of comparison to the measured cross sections the HVQDIS code was used with the following settings:

- The pole charm quark mass, $m_c$, defined as the real part of the pole in the charm quark propagator, was set to 1.5 GeV. This value was chosen to give a reasonable description of the charm data. It is also supported by the estimations made in Section 3.3 of [40] with the next-to-next-to-leading-order (NNLO) QCD relation between the $c$ quark pole mass and the $m_c(\overline{\text{MS}})$ value quoted by the Particle Data Group [10]. For the estimation of the theory uncertainty the value of $m_c$

was varied by 0.15 GeV up and down. In this way also another value (1.35 GeV) used in the past in ZEUS, e.g. [26], and the value of the pole mass (1.65 GeV), which could be extracted at NLO from the $m_c(\overline{\mathrm{MS}})$ value, are covered. The variation is similar to the one suggested in [40]. Simultaneously with the charm quark mass in the matrix element, the value used in the PDF fit was varied.

- Renormalisation and factorisation scales were set to $\mu_R = \mu_F = \sqrt{Q^2 + 4m_c^2}$. To estimate the uncertainty arising from this choice, the scales in the hard interaction were varied independently by a factor 2 up and down.

- The PDFs were extracted from the ZEUS-S NLO PDF fit [41] in the 3-flavour FFNS with the same charm quark mass as those mentioned above. In general one should use the same scales in the PDFs and in the matrix element calculations, but only fits with $\mu_R = \mu_F = Q$ are available for the central $m_c$ value. Therefore a check was performed to ensure that the effect caused by the difference in scales is small. Figure 2.1 shows a comparison of predictions using PDFs with the default $\mu_R = \mu_F = Q$ choice and with consistent $\mu_R = \mu_F = \sqrt{Q^2 + 4m_c^2}$ settings using available PDFs with $m_c = 1.35$ GeV. The experimental uncertainties from the PDF fit were propagated to the theory uncertainties.

- The $\alpha_s$ running at NLO with 3 active flavours was used setting $\Lambda^{(3)}_{QCD} = 363$ GeV.

Individual sources of the theory uncertainties, including the fragmentation uncertainties discussed in the following, were summed in quadrature to get the total theory uncertainty.

HVQDIS produces events on parton level. In order to select a specific phase space for the final-state particles ($D^+$ mesons for this study), the fragmentation process has to be modelled. A detailed discussion of the fragmentation model that was developed for this analysis is presented in the following Section 2.2.

## 2.2 Fragmentation process

The QCD factorisation theorem implies that a hadron production cross section can be factorised into a partonic cross section for the production of

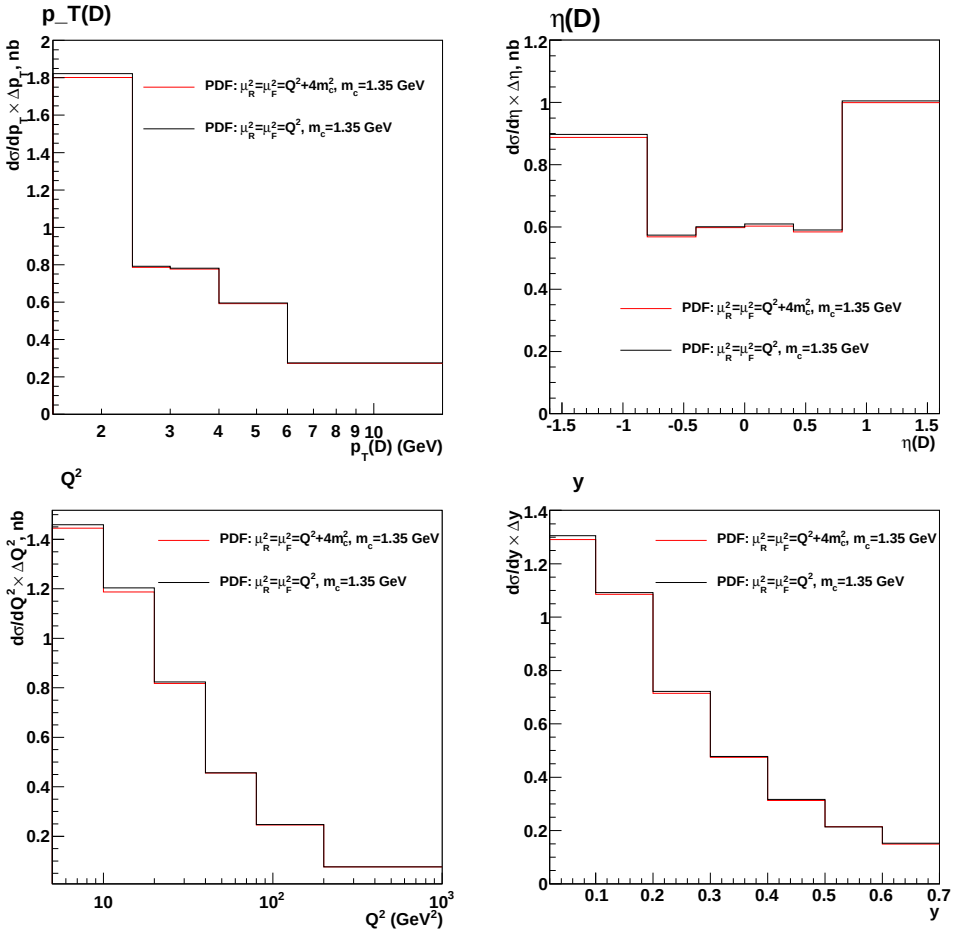


**Figure 2.1:** *A comparison of the NLO QCD predictions with the default settings and an independent Peterson fragmentation function with $\varepsilon = 0.079$ using two different scale-choices in the PDFs. The difference in the kinematic region of the current measurement is small.*

parton type $i$, $C_i$, and a *fragmentation function* from the parton $i$ into the desired hadron, $D_i$:

$$\frac{d\sigma_c}{dz'}(z', q^2, m^2) = \sum_i \int_{z'}^{1} \frac{dx}{x} C_i(x, q^2, \mu^2) D_i(\frac{z'}{x}, \mu^2, m^2), \qquad (2.1)$$

where $z'$ is the quark's momentum fraction carried by the hadron and $\mu$ is the fragmentation factorisation scale, not to be confused with $\mu_F$ – the factorisation scale for PDFs. Similar to the PDFs, the perturbative component that describes the evolution by the DGLAP equations from a starting scale,

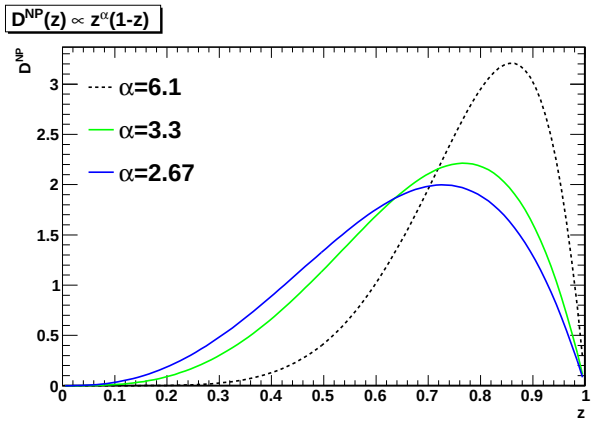


**Figure 2.2:** *The Kartvelishvili fragmentation functions measured by ZEUS [46] (dark solid line) and H1 [47] (light solid and dotted lines) with the NLO partonic cross sections. Only central values from the quoted papers are shown.*

$\mu_0$, to any given energy scale $\mu$ and a non-perturbative component given as a parametrisation at the starting scale $\mu_0$ could be isolated [42]. The perturbative part can depend on the production mechanism or on the production environment, whereas the non-perturbative component is supposed to be universal and should be measured experimentally.

The most precise measurements of the charm fragmentation function were done by the experiments BELLE [43], CLEO [44] and ALEPH [45] at $e^+e^-$ colliders. A phenomenological NLO QCD analysis of these data [42] has shown, that the ALEPH data measured at $\sqrt{s} = 91.2\,\text{GeV}$ could be reasonably described by the evolution of the non-perturbative fragmentation function extracted from the CLEO and BELLE data at $\sqrt{s} = 10.6\,\text{GeV}$.

Unfortunately HVQDIS does not calculate the perturbative part of the fragmentation function. Thus the charm fragmentation function measured by ZEUS [46] and H1 [47] in $ep$ collisions at HERA may differ from the results from $e^+e^-$ colliders. This leads to the conclusion that the results of [42] couldn't be applied in the NLO calculations by HVQDIS. Hence an empirical approach was developed[1].

The simplest approach is to neglect the perturbative part of the fragmentation function and use a fragmentation independent from the hard scattering. The common choices of the independent fragmentation functions for the non-perturbative component are:

[1]Together with Massimo Corradi.

- Peterson function [48] : $D^{\mathrm{NP}}(z) \propto \dfrac{1}{z\big(1-1/z-\varepsilon/(1-z)\big)^2}$,
- Kartvelishvili function [49] : $D^{\mathrm{NP}}(z) \propto z^{\alpha}(1-z)$,

where $\varepsilon$ and $\alpha$ are two parameters of hardness, that should be measured experimentally, and both functions should be normalised to unity. Hereafter for the D meson and c quark momenta and energies, $P_D$, $P_{c\,quark}$, $E_D$ and $E_{c\,quark}$, $z$ is defined as $\dfrac{(E+P)_D}{(E+P)_{c\,quark}}$, which is a common choice and the fragmentation function is applied in the $\gamma^* p$ (photon–proton) frame. Figure 2.2 shows the Kartvelishvili fragmentation function with the central values for the $\alpha$ parameters measured in [46, 47] in different kinematic regions. One can see, that the spread of the functions is large, and consequently, if all HERA measurements would be considered with an independent fragmentation model, the theory uncertainty would be large.

The approach used in this analysis was to fit the $\alpha$ values, extracted from the HERA data, as a function of an energy scale, e.g. the $c\bar{c}$ center-of-mass energy,$M(c\bar{c})$, or energy in the hard subprocess, $\hat{s}$. In this way the fragmentation depends on the kinematics of the hard scattering. Figure 2.3 shows such a fit by a simple $\alpha(W) = a + \dfrac{b}{W - 4m_c^2}$ function. The functional form was motivated by the phase space restriction in the limit $W \to 4m_c^2$. The fit outcome is indicated in the legend of the figure. Also shown are functions used for the estimation of the theory uncertainty.

As the energy scale the $c\bar{c}$ center-of-mass energy, $M(c\bar{c})$, was used. It is identical to $\hat{s} \equiv M(\gamma - parton)$ in leading-order (LO) QCD, but differs at the next-to-leading-order due to a possible real final-state gluon emission. It was found, that cross section in $M(c\bar{c})$ is less sensitive to the variations of the factorisation and renormalisation scales and thus the theory uncertainties arising from the fragmentation and from the scale choice become (almost) independent. This is presented in Fig. 2.4.

All the fragmentation measurements, which were discussed so far, were done with $D^{*+}$, whereas this thesis concentrates on $D^+$ production. To account for the fact, that $D^+$ mesons are produced directly from $c$ quarks, but also from decays of $D^{*+}$ and other higher states, the $D^{*\pm}$ fragmentation spectrum was corrected. The correction was applied to $\alpha$ so that the mean value of the $D^{*+}$ fragmentation spectrum was scaled down by 0.95. This coefficient was taken from the $e^+e^-$ measurements [43, 44].

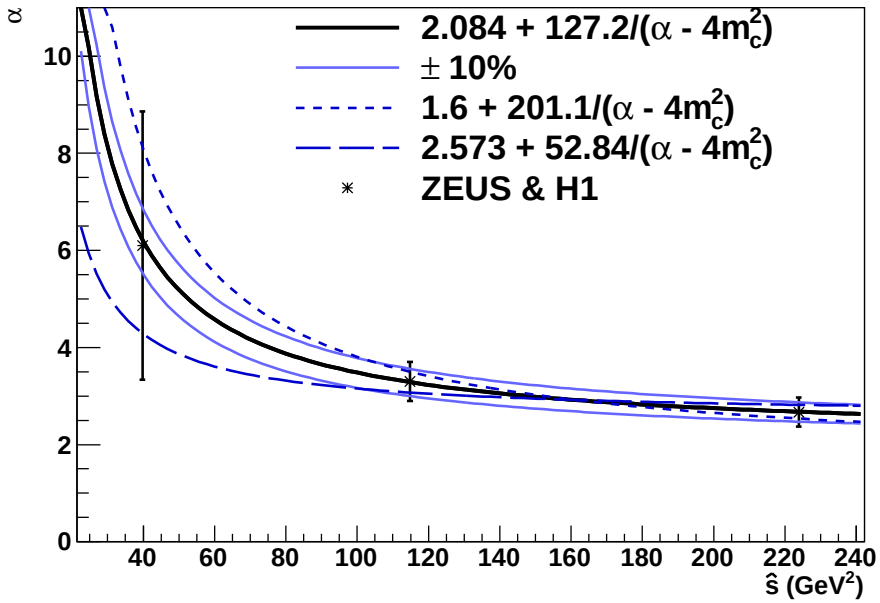


**Figure 2.3:** *A fit of the measured values of Kartvelishvili parameter $\alpha$ as a function of mean $\hat{s}$ for the ZEUS and H1 measurements [46, 47] (black stars). The values of $\hat{s}$, which correspond to each measurement, were extracted using a leading-order + parton shower Monte Carlo (see Section 4.1). The uncertainty of the measurement by H1 at the lowest $\hat{s}$ ("No $D^{*\pm}$jet sample") was scaled by $\sqrt{\chi^2/ndof} = 3.1$ to account for the badness of the fit in the paper. The fit result (solid black line) is shown together with proposed variations by 10% up and down (solid light violet lines) as well as more and less steep functions(dashed and dotted dark blue lines).*

While the fragmentation function is normalised to unity, the total probability for a certain meson to be produced in a given fragmentation process (including decays of higher states) is given by the *fragmentation fractions*. A combination of results from $e^+e^-$ colliders and from deep inelastic production at HERA was used [50], thus the value used for predictions was $f(c \to D^+) = 0.2192 \pm 0.0085$.

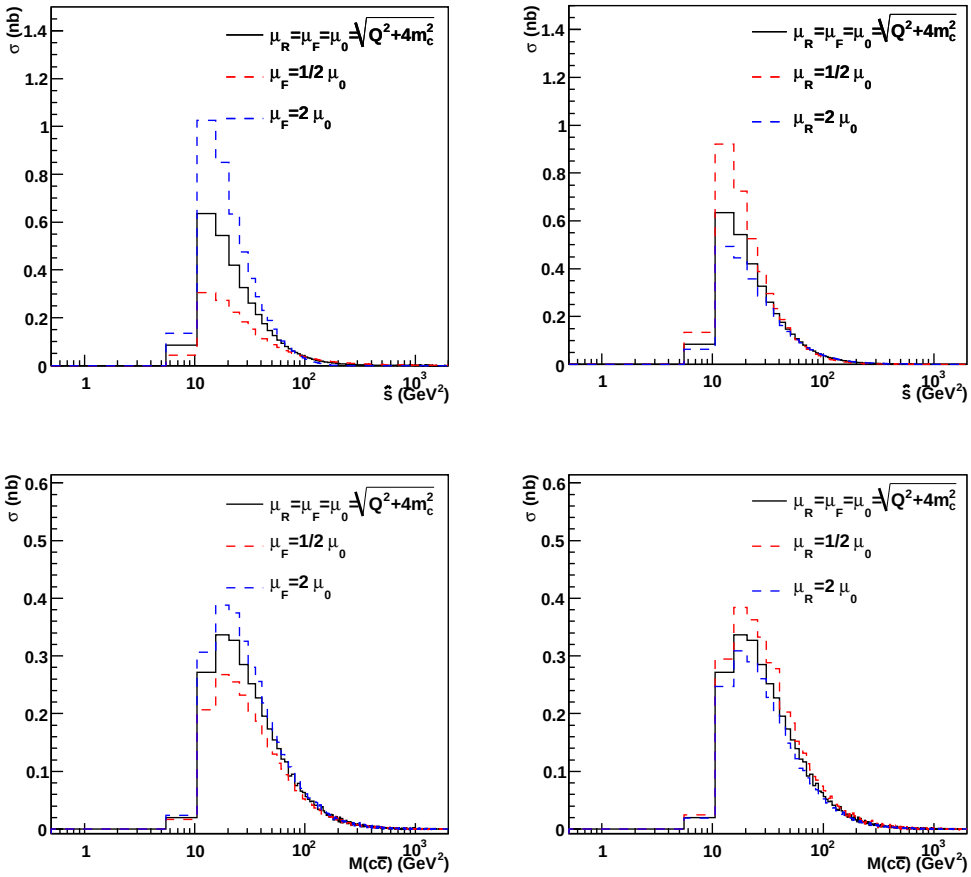


**Figure 2.4:** *Sensitivity of the charm production cross sections as a functions of $\hat{s}$ (top row) and $M(c\bar{c})$ (bottom row) to the choice of the factorisation (left column) and the renormalisation (right column) scales. Shown are NLO QCD predictions by* HVQDIS *for the kinematic region of this analysis.*

# Chapter 3

# The HERA collider and the ZEUS experiment

This Chapter is devoted to a description of HERA and the components of ZEUS, which are relevant for this analysis.

## 3.1 The HERA collider

HERA (German Hadron-Elektron Ring Anlage) [51] was the first and so far the only electron–proton collider. It was located at DESY in Hamburg, Germany. The circumference of the HERA ring was 6.3 km. The first beams were injected in the storage rings back in 1992 and the last collisions were recordered in summer 2007. The electron beam was accelerated to 27.5 GeV, whereas the proton beam energy was increased in 1998 from 820 GeV to 920 GeV. Hence, the center-of-mass energy, $\sqrt{s}$, was fixed to 300 GeV (318 GeV after 1998), which is one order-of-magnitude higher than in the previous fixed-target DIS experiments. Particles in the beams were grouped into *bunches*. After synchronisation and tuning the beams were brought into collision with an interception angle close to 0° and a bunch-crossing interval of 96 ns. For the electrons the control and focusing used normal-conducting magnets, whereas for the proton beam superconducting dipole and quadrupole magnets at 4.4 K had been used. The magnetic field to bend the electrons was up to 0.274 T, for protons up to 4.65 T.

There were 4 experimental halls at the interaction points (see Fig. 3.1). In two of them the collider experiments ZEUS [52] and H1 [53] were installed,

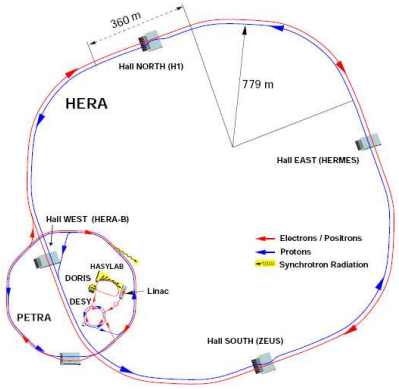


**Figure 3.1:** *Schematic view of the HERA storage rings and experiments. The figure is taken from [56].*

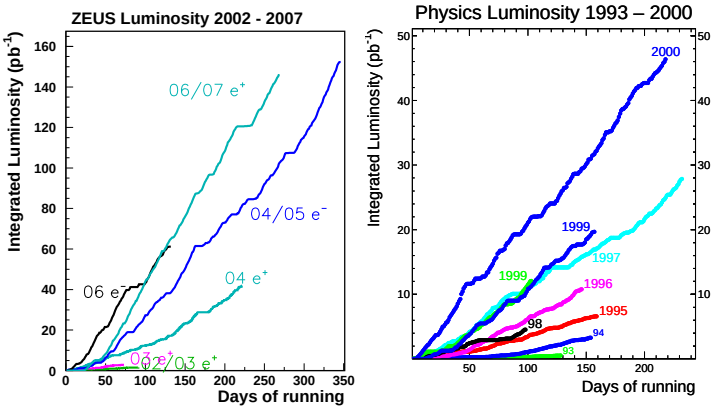


**Figure 3.2:** *Integrated luminosity collected by ZEUS for the 1993–2000 and 2002–2007 periods, shown separately for each year. The plots are taken from [57].*

in the other two the fixed target experiments HERMES [54] and HERA-B [55] were constructed.

The data harvest was subdivided into two periods: HERAI (1993–2000) and HERAII (2002-2007), with a major upgrade of the detectors and the accelerator in between. The instantaneous luminosity achieved for the HERAII period was $\approx 7.5 \times 10^{31}\,\mathrm{cm}^{-2}s^{-1}$. The integrated luminosity recordered by the ZEUS detector for two periods is shown on Fig. 3.2.

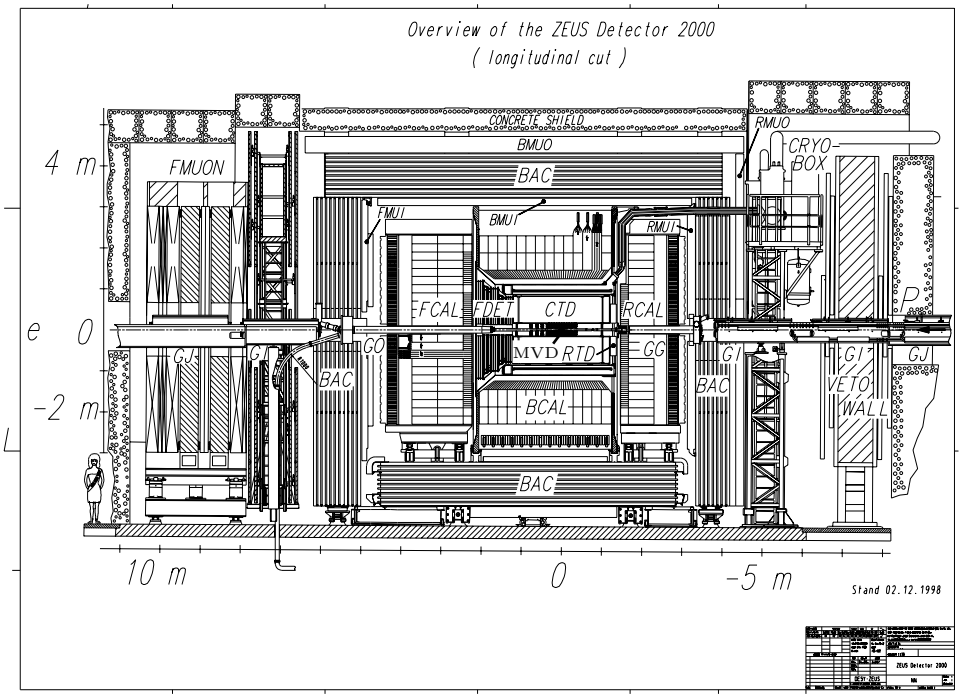


**Figure 3.3:** *A schematic view of the ZEUS detector in a vertical plane parallel to the beampipe. The figure was taken from [58].*

## 3.2 The ZEUS experiment

ZEUS was a multipurpose detector with $4\pi$–geometry designed to reconstruct the final state particles in *ep* collisions at HERA. The experimental setup of its components is shown on Fig. 3.3. The main components used in this analysis are:

- *The Micro Vertex Detector (MVD)* [59]: A silicon-strip vertexing device installed in the break between 2000 and 2002.
- *The Central Tracking Detector (CTD)* [60]: A gas drift chamber used for the particles' momentum measurement.
- *The superconducting solenoid*: Magnetic field 1.43 T was used to bend charged particles.
- *The uranium Calorimeter (CAL)* [61]: Depleted-uranium and scintillators of high resolution were used to measure energies of final-state particles.

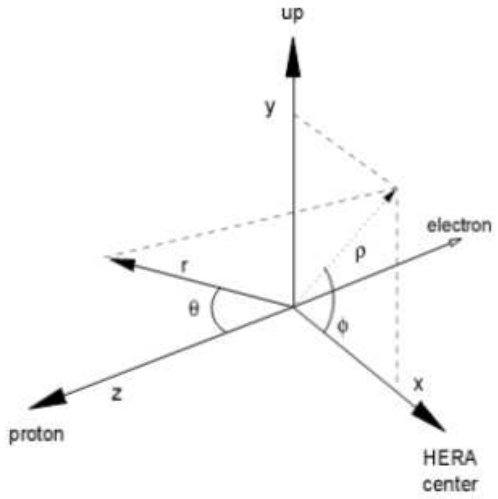


**Figure 3.4:** *The coordinate system adopted in ZEUS.*

- *The Scintillator Rear Tracking Detector (SRTD)* [62]: A scintillator-strip detector located on the inner surface of the rear endcap of the CAL around the beampipe for the measurements of the position of the scattered electron.

- *The Luminosity Monitor* [63]: Detector system to measure photons from the bremsstrahlung process, $ep \rightarrow e'\gamma p$.

These components will be described in more detail in Sections 3.4 to 3.7. A detailed description of the remaining ZEUS components can be found elsewhere [52].

## 3.3 The ZEUS coordinate system

The coordinate system used in ZEUS is an orthogonal right-handed system centred at the interaction point (see Fig. 3.4). The $Z$ axis is parallel to the beampipe and is pointing in the proton beam direction, the $X$ axis is pointing perpendicular to the beams towards the center of collider and the $Y$ axis is perpendicular to the collider plane. The polar angle $\theta$ is measured relative to the $Z$ direction, whereas the azimuthal angle $\phi$ is measured relative to $X$ in the $XY$-plane. Instead of $\theta$ frequently the *pseudorapidity*, $\eta$: $\eta = -ln(tan\frac{\theta}{2})$ is used.

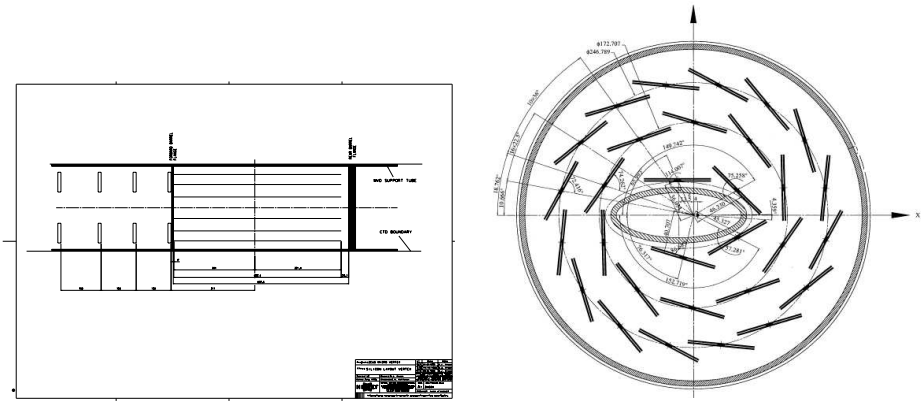

**Figure 3.5:** *Cross section of the Micro Vertex Detector in $YZ$ (left) and $XY$ (right) planes. The figure was taken from [58].*

## 3.4 The Micro Vertex detector (MVD)

The MVD [59] was an important subcomponent of the ZEUS tracking system. The Micro Vertex detector had been installed in the shutdown between 2000 and 2002 to improve the space resolution of tracks in the vicinity of the interaction point. It was subdivided into two independent components: barrel (BMVD) and forward (FMVD).

The barrel part was built of 600 silicon-strip modules, that were grouped in three cylindric layers around the beampipe (see Fig. 3.5). The modules were double-sided and strips on the opposite-side sensors were perpendicular, such that both the $r\phi$ and the $rz$ position of a hit can be measured. Thus a track could produce up to 2 hits per layer. The BMVD covered the polar angle range $30° < \theta < 150°$. The measured space resolution for a single hit was 24 µm and the two-track separation resolution was measured to be 120 µm.

The forward part consisted of 112 silicon-strip wedge-shaped sensors. The sensors were organised into 4 forward wheels, which extended the coverage in polar angle down to $\theta > 7°$.

## 3.5 The Central Tracking detector (CTD)

The Central Tracking detector [60] was used to measure the 3-momenta of charged particles with high precision as well as to estimate the mean energy losses for a particle in the active volume of the detector, $dE/dx$, which is

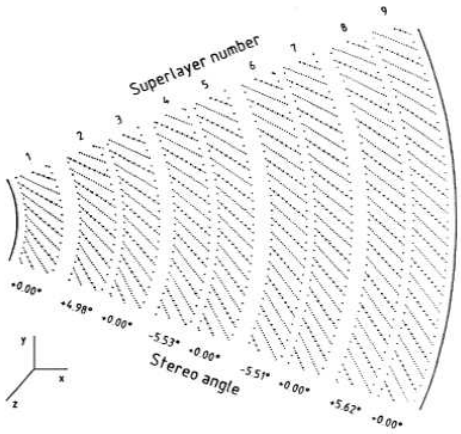


**Figure 3.6:** *Cross section of a CTD segment perpendicular to the beampipe. The larger dots indicate the sense wires. The figure was taken from [56].*

important for particle identification. The CTD was a cylindric drift chamber with an inner radius of 18.2 cm, an outer radius of 79.4 cm and a total length of 205 cm. It was filled with a mixture of 83 % $Ar$, 12 % $CO_2$ and 5 % ethane at atmospheric pressure. The CTD covered the region of $15° < \theta < 164°$ and was made of 72 layers of wires, that were grouped in 9 *superlayers (SL)* (see Fig. 3.6). Wires in the odd-numbered superlayers were parallel to the beam, whereas wires in the even-numbered superlayers had a small angle ($\pm 5°$) with respect to the $Z$ axis. This allowed to reconstruct both the $r\phi$ and the $z$ coordinates accurately. In addition, the $z$ position could be reconstructed from the time difference of the signals arriving to the opposite ends of the wires, *z–by–timing*. Due to its poor resolution ($\sim 4$ cm), this technique was used only for a fast reconstruction in the ZEUS First Level Trigger. All wires were separated into two classes– 4608 *sense wires* with the positive potential to detect the signal from the drifting electrons and 19584 *field wires* with the negative potential.

The space resolution of the CTD was 200 µm in $r\phi$ and approximately 2 mm in $z$. The combined transverse momentum resolution of MVD+CTD for the tracks fitted to the primary vertex and that have reached the 9-th superlayer was [64]:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = 0.0029\, p_{\mathrm{T}} \oplus 0.0081 \oplus \frac{0.0012}{p_{\mathrm{T}}}. \tag{3.1}$$

Here $\oplus$ indicates that the terms are added in quadrature and $p_{\mathrm{T}}$ is mea-

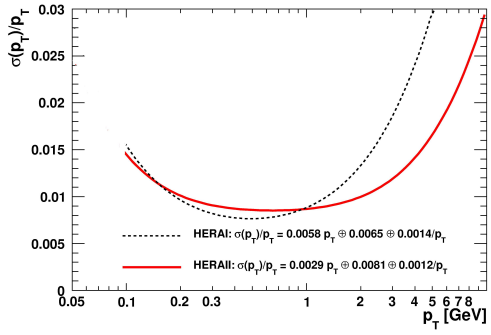


**Figure 3.7:** *Comparison of the $p_{\mathrm{T}}$ measurement resolution in the ZEUS tracking system for HERAI and HERAII periods.*

sured in *GeV*. The first term in the sum corresponds to the resolution of the hit position measurement, whereas the second and the third terms arise from the multiple scattering before and inside of the CTD, respectively. Figure 3.7 shows a comparison of the resolution obtained with the CTD+MVD in HERAII with the momentum resolution in HERAI [65]:

$$\frac{\sigma_{\mathrm{T}}}{p_{\mathrm{T}}} = 0.0058\, p_{\mathrm{T}} \oplus 0.0065 \oplus \frac{0.0014}{p_{\mathrm{T}}}. \tag{3.2}$$

The comparison shows, that the precision of the hit position measurement has been greatly improved, however the amount of multiple scattering has increased owing to the installation of the MVD in front of the CTD.

## 3.6 The uranium calorimeter (CAL)

The ZEUS calorimeter [61] was a high-resolution compensating calorimeter, which has used the depleted-uranium plates as absorber interleaved with a plastic scintillator (SCSN–38) as active material. The ratio of the scintillator and the absorber thickness was tuned to achieve the best compensation, i.e. the ratio of responses to the electromagnetic and hadronic components of a shower close to 1.

Geometrically the CAL was subdivided into tree parts (see Fig. 3.8): the forward calorimeter (FCAL) covered the range $2^\circ < \theta < 40^\circ$, the barrel calorimeter (BCAL) $37^\circ < \theta < 129^\circ$ and the rear calorimeter (RCAL) $128^\circ < \theta < 177^\circ$. Each of the three parts was subdivided into the inner electromagnetic calorimeter (EMC) and 1, for the RCAL, or 2, for the BCAL and the FCAL, hadronic calorimeter (HCAL) behind the EMC. To enable

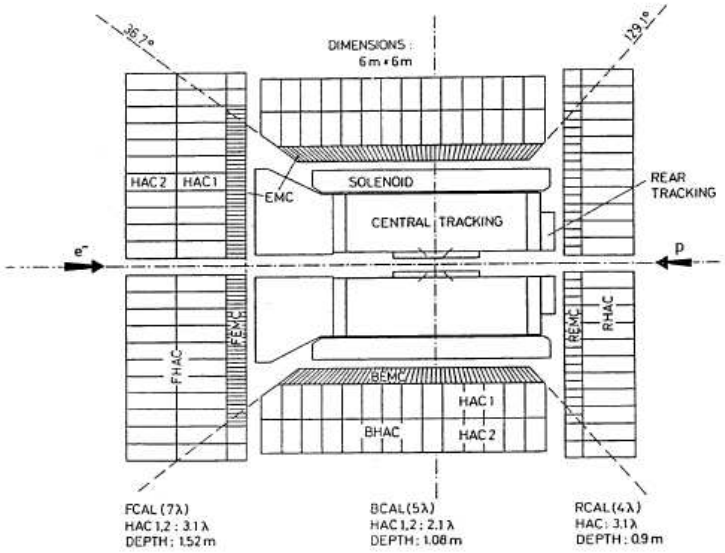


**Figure 3.8:** *Structure of the ZEUS calorimeter. The figure was taken from [66].*

position measurements both sections were subdivided into cells: the HAC was subdivided into cells of $20 \times 20\,\mathrm{cm}^2$, whereas the EMC cell size was $5 \times 20\,\mathrm{cm}^2$. The readout of each calorimeter cell was done on two sides by wavelength shifters, coupled to photomultiplier tubes. The comparison of two signals from the same cell enabled a measurement of the signal position within a cell.

The depth of the EMC in units of the radiation length was 21–25 $X_0$. The depth of the HAC in units of the absorption length varies for the RCAL, BCAL and FCAL from 3 up to 6 $\lambda$.

The noise originating from the natural radioactivity of uranium was at the level of 15 MeV for an EMC cell and 30 MeV for a HAC cell. This activity was used for an intrinsic callibration of the calorimeter with high precision. On the other hand, for measurements it produced a pile-up effect, thus energy thresholds of 60 MeV for EMC and 110 MeV for HAC had to be introduced. On average these thresholds were exceeded only in 5 EMC and 2 HAC cells out of 5918 in total.

A test beam was used to determine the energy resolution for electrons and hadrons to be $\frac{18\%}{\sqrt{E/GeV}}$ and $\frac{35\%}{\sqrt{E/GeV}}$, respectively. In addition excellent time resolution (of the order of 1 ns for energy deposits above 1 GeV)

was achieved and was actively used in all levels of the trigger system (see Section 3.8).

## 3.7 Luminosity monitor

A luminosity measurement with high precision is essential for any experiment in high energy physics. At HERA the luminosity was determined from the bremsstrahlung process $ep \to e'\gamma p$. On the theory side the cross section for this process is described by the Bethe–Heitler formula with a precision of 0.5%.

Photons emitted with an energy above a fixed energy threshold were registered in the Photon Calorimeter (PCAL) [67] installed 107 m away from the nominal point of interaction in the electron-beam direction. The luminosity could be extracted as $\mathcal{L} = N_\gamma/\sigma_{pr}$, where the predicted cross section, $\sigma_{pr}$, [68] had to be corrected for the PCAL acceptance. For the luminosity measurements a background arising from the bremsstrahlung process on nuclei of the remaining gas in the beampipe, *beam-gas interactions*, was subtracted. This admixture was measured using lepton bunches, that did not have a matching proton bunch.

During the HERA II running an additional spectrometer (SPEC) [63] was installed complementary to the PCAL system. Instead of a direct measurement of the bremsstrahlung photons the SPEC registered $e^+e^-$ pairs from the bremsstrahlung photon conversions in a copper-beryllium window. Approximately 10% of the photons converted into electron pairs, giving an order–of–magnitude reduction in the observed rate compared to the PCAL. The electron pairs were separated by the magnetic field of a dipole magnet and measured by two segmented tungsten–scintillator sampling calorimeters.

The final luminosity measurement precision achieved in ZEUS was 1.8% [69].

## 3.8 Trigger system

The bunches at HERA were crossing at the interaction points with a period of 96 ns, hence the potential collision rate was $\approx 10$ MHz. The rate of the beam-gas interactions was up to 100 kHz. These rates significantly exceed available maximum readout and storage rates of $\sim 10$ Hz. Even the rate of cosmic rays ($\sim 1$ Hz) was comparable to the rate of the useful $ep$ events

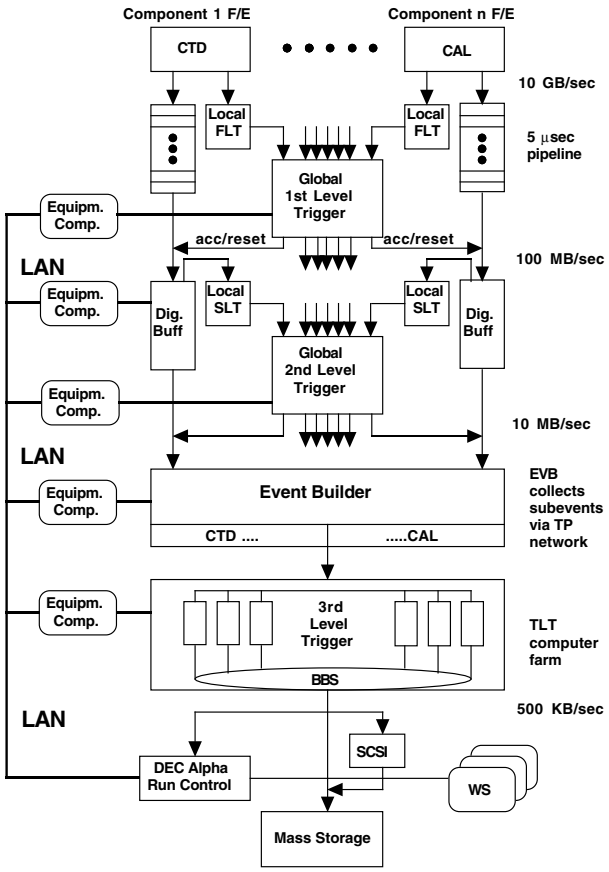


**Figure 3.9:** *A schematic view of the ZEUS trigger chain with the rates at each trigger level. The figure was taken from [56].*

($< 10\,\mathrm{Hz}$). Therefore, a three-level trigger system was used in ZEUS. The trigger on the first level was operated in a pipeline mode to avoid deadtime. Figure 3.9 shows a schematic view of a dataflow through the ZEUS trigger system.

The *First Level Trigger (FLT)* reduced the rate from $\sim 10\,\mathrm{MHz}$ to $\sim 1\,\mathrm{kHz}$. Signals from the detector sub-components (CAL, CTD, muon chambers, etc.) underwent a fast hardware reconstruction, which was used to prepare information for a trigger decision from each component. All these signals were then sent to the *Global First Level Trigger (GFLT)*, which made

the trigger decision by combining the information from the individual FLT slots. The other two important tasks of the GFLT were to synchronise subdetectors with the HERA bunch-crossing clock and to produce a fast decision to reject beam gas events ("Fast Clear").

The purpose of the *Second Level Trigger (SLT)* was to further reduce the event rate down to 100 Hz. This was achieved by a more detailed analysis of events, accepted by the GFLT, using information about tracks in the CTD, primary vertex location and more detailed analysis of the CAL (e.g. reconstructed electron candidates, hadronic clusters, etc.). Similar to the first level structure, the *Global Second Level Trigger (GSLT)* was used to combine the SLT trigger information and to pass an event to the next trigger level. The total time window available to make a decision was about 8 ms.

On the last trigger layer a full event reconstruction (close to an *offline*-reconstruction), selection and classification was done. The software-based *Third Level Trigger (TLT)* was running on a dedicated computer cluster. The low output rate from the GSLT enabled a selection using kinematic and topological criteria. For example, in this analysis inclusive DIS triggers were used, which applied cuts on the position and the parameters of the reconstructed scattered electron and common kinematic variables (see Section 6.2.1). Triggered events had a size of 150 kB/event and were transfered for the final storage on tapes in the DESY computing center at the rate of a few Hz.

The final data were stored in the format of ADAMO ("Aleph Data Model") tables [70], which could be processed by the *offline*-reconstruction program ZEPHYR.

# Chapter 4

# Monte Carlo simulations

All measured quantities in a high energy physics event are on the so-called *detector level*, which means that underlying physics is folded with the detector response. Usually the latter can not be calculated analytically. To model it simulations by the *Monte Carlo (MC) method* [71] are used.

This Chapter starts with an overview of possible usage scenarios as well as a brief description of the MC simulation steps. Then particular generators used in this thesis are reviewed. As the next logical step, a simulation of the detector response to generated events is described in detail since the author was involved deeply in this procedure. An automatic MC production system and its development are reviewed.

## 4.1 Monte Carlo events

In the simulations the QCD factorisation theorem is exploited to separate a hard interaction calculated up to a fixed order in pQCD from soft processes, which are modelled phenomenologically. Most of the simulations are based on LO predictions and higher orders are partially modelled by showers of soft partons.

Figure 4.1 illustrates the simulation steps for a BGF event. For the hard scattering the matrix element for the $\gamma g \to c\bar{c}$ process is calculated up to the first order in pQCD[1]. Soft gluon emissions, *parton showers (PS)*, are modelled to better describe the kinematics. The differences between the NLO calculations and the LO MC with one gluon emission in the shower are

[1] An attempt to introduce NLO BGF matrix elements in the photoproduction regime into MC generators is described in [72].

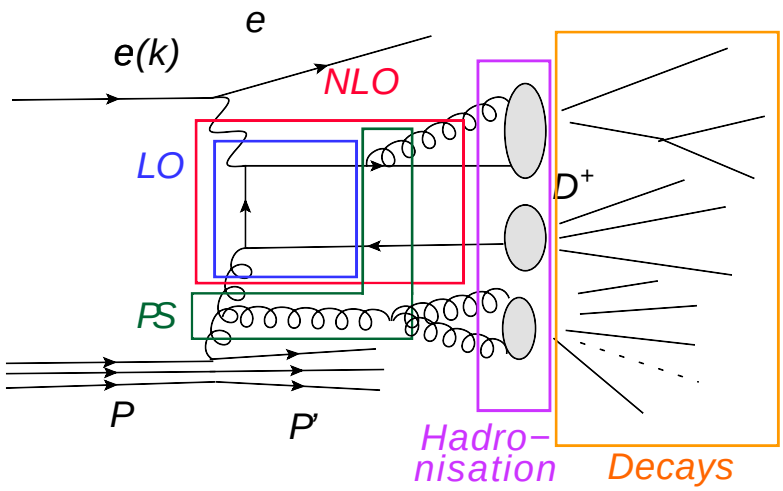


**Figure 4.1:** *An example of simulation stages followed in a standard MC generator. Shown are leading-order Feyman graph for the hard scattering (LO), soft parton showers (PS), parton hadronisation and decays into final-state particles. Also shown is how the LO+PS components simulate higher orders (NLO). The figure was inspired by Fig. 4.1 in [73].*

that the gluon can be harder and that some of the virtual contributions are accounted for in the former and absent in the latter. Thus the kinematics of the charm quarks is slightly different. After a parton shower an event is on the so-called *parton level*. The next step is to model the process of formation of colourless hadrons from coloured partons – *hadronisation*. Hadronisation is a soft-QCD process, hence only models can be applied. Finally, the decays of unstable particles are modelled to the level when decay products have a sufficiently long lifetime to interact with the detector, e.g. charged pions and kaons. Different experiments can be compared on the level of hadrons, *hadron level*, since it is affected neither by the detector setup nor by the data taking conditions. This level will be often referred to as the *true level* in this thesis.

Monte Carlo simulations are used for different purposes:

- optimisation of an event selection aiming for the largest ratio of a signal to the square root of the signal plus background, i.e the smallest statistical uncertainty for the signal (see Section 6.3.2);
- simulations of possible signatures of predicted, but not yet discovered, physics (e.g. searches of Higgs at LHC [4, 5]);

- transition from the detector level to the hadron level by detector acceptance corrections (see Section 7.1.1);
- transition from the parton level to the hadron level via hadronisation corrections.

## 4.2 Monte Carlo samples

Different generators were used to simulate different physical processes, for which they are most suited.

### 4.2.1 ARIADNE

The ARIADNE Monte Carlo generator [74] was used to simulate production of inclusive light flavours in DIS. A hard electron-parton scattering was simulated in DJANGON 1.1 [75, 76], which provides an interface between LEPTO 6.5.1 [77] and HERACLES 4.6.1 [75, 78]. A LO matrix element was generated with LEPTO, while electromagnetic corrections were calculated in HERACLES. A QCD cascade was simulated in the ARIADNE model of a colour dipole [74]. In this scheme the interaction between colour charges is replaced by the colour-dipole and parton emissions appear from the dipole as a whole, rather than from the separate charges. For this sample PYTHIA 6.146 [79, 80] has been used to simulate hadronisation and decays.

LEPTO does not include massive matrix elements for the heavy quark production, therefore ARIADNE was used only in studies that involved the light flavour background, e.g. detector resolutions and efficiencies.

### 4.2.2 RAPGAP

The RAPGAP generator [81] was used to simulate heavy flavour production in DIS. It also uses HERACLES 4.6.1 to calculate the electromagnetic corrections and PYTHIA 6.146 for the hadronisation of partons and following decays. A QCD cascade is modelled inside of RAPGAP itself.

For the hard scattering the massive matrix element of the BGF process was used. The default RAPGAP 3.0 settings were used (see [81] for more information).

## 4.3 Monte Carlo production system

Nowadays computations in high energy physics become so CPU- and disk-space-consuming, that a single PC and even one computer cluster can not deal with them on a reasonable time-scale for an output. This leads to the combination of computer resources into networks like *GRID*. At the beginning of the ZEUS data harvest general solutions were not developed yet, thus a special network of PCs from the participating institutes was established – *FUNNEL*. Currently FUNNEL encompasses PC farms in Germany, Japan and Italy. In 2004 the EGEE GRID [82] resources were also available to ZEUS [83] (currently its successor, the EGI network, is in use). Important criteria for such a system are high reliability, autonomy and flexibility: efficiency of job completion is above 95%, manual interuptions to FUNNEL are very rare and jobs can be run on various operating systems on the UNIX platform. The production rate of FUNNEL is $\approx$ 12 million events per day, which can be compared to the averaged production rate in 2006 of 2 – 3 millions/day.

### 4.3.1 Detector simulation

Event generation is much less time consuming than the detector modelling, e.g. the average time to generate 1 BGF event in RAPGAP is $\sim 0.005$ s whereas the average time for the simulation of the ZEUS detector responce is $\sim 1$ s. Thus events on the hadron level can be generated on a local PC, whereas the simulation of the detector as well as the offline-like reconstruction are performed on FUNNEL.

The output of the event generator is stored in the format of ADAMO tables. A conversion from the most common LUJETS, PYJETS and MCHEP formats is done within the AMADEUS package. The simulation of the ZEUS detector response is performed in three steps. First, the program package MOZART [84], which is based on GEANT 3.21 [85], is used to simulate the penetration of the generated hadrons through a detailed model of the ZEUS detector. The ZEUS data taking period is subdivided into smaller subperiods, referred to as the *data sub-periods*, with specific MOZART settings to provide a realistic evolution of the detector status. At the next level the three-level trigger system of ZEUS is simulated in the programme called CZAR. This component is less flexible compared to MOZART, because many parameters, e.g. trigger thresholds, are hardcoded in the program. Thus for each subperiod an individual CZAR executable exists. After this stage events have exactly the same format as the raw data events apart from the available

generator information. Therefore ZEPHYR is run afterwards with the same set-up as for the data.

### 4.3.2 Quality of the detector simulation

Clearly, an important issue is to get the detector model as close as possible to the state of the detector during data taking. Even 3 years after the shutdown of HERA simulations are under development. Modifications are introduced in a format of a new *FUNNEL version*, which comprises a frozen set-up of MOZART, CZAR and ZEPHYR. In particular, the author was involved in the introduction of:

- time-dependent CTD pulse height;
- new distributions of the primary vertex Z position;
- a simplified scheme of the division into sub-periods with different trigger setup.

From the beginning, only a couple of the data taking runs from each sub-period were used to tune a CTD response. Recently it was discovered that the effect of the CTD pulse height time–dependence, owing to a variation of gas pressure in the drift chamber with variation of atmospheric pressure, is much larger than it was expected [86]. In general this feature should have a small effect on the detector efficiency. However, in HERA II due to closeness of the CTD pulse height to the hit threshold level, even small variations of the CTD pulse height significantly affected the hit efficiency. This effect is visible in the efficiency of the CTD FLT requirements and hence in the efficiency of all FLT slots, that use CTD information. Figure 4.2 illustrates the dependence of the FLT CTD tracking efficiency[2] on the run number. Efficiency in the old MC versions was constant in time, whereas the data showed a clear time-dependence, which coincides in shape with the $dE/dx$ measurements in the CTD DQM [87]. The chosen solution was to store the values of $dE/dx$ in a calibration table and to choose the CTD gas gain for each event based on a simulated run number assigned to the event. The plot also shows, that a new MC version is able to simulate this feature well.

Another task was to introduce new vertex distributions for special runs with reduced proton-beam energy, which were taken at the end of the HERA

[2]Efficiency is measured versus a reference trigger that doesn't use any tracking information on the FLT level. FLT30 was chosen, because it will be also used in the following analysis. Efficiency is defined as $N^{\text{events}}(\text{FLT30}\,\&\,\text{FLTtracking})/N^{\text{events}}(\text{FLT30})$.

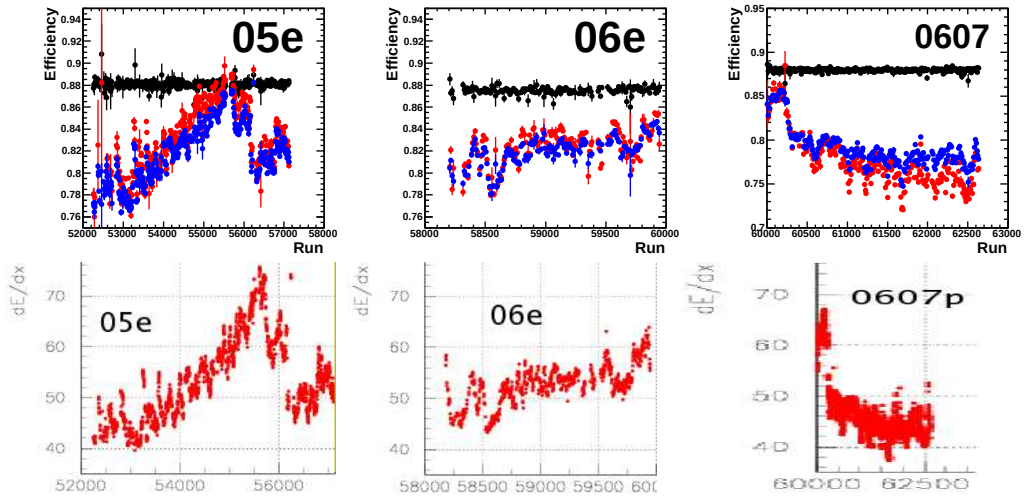


**Figure 4.2:** *Top: the dependence of the FLT tracking efficiency vs data taking run number. Light red points are data, black points – an old MC version with a fixed CTD gas gain, dark blue points – a new MC version with the gas gain tuned to the dE/dx in the CTD. Binomial errors are shown. Bottom: a measurement of mean dE/dx in the online CTD data acquisition.*

data taking. Up to now vertex distributions from the preceding period of 2006-2007 have been used. Hence complicated procedures of the vertex distribution "reweighting" on the level of a particular analysis were needed.

As it was already mentioned in Section 4.3.1, the whole HERA data taking period is divided into sub-periods with an (almost) uniform detector and trigger set-up. In particular, those sub-periods that had the same detector setup and differed only by the trigger settings, i.e. in which CZAR binaries were also identical, were merged. Therefore a specific trigger setup was chosen based on an event-dependent flag in the new calibration tables. This reduces the number of FUNNEL versions (roughly by a factor 2) and simplifies the distribution of events between the versions for users.

### 4.3.3 MC production on FUNNEL

A detailed description of the FUNNEL structure can be found in [88], here only a brief overview will be given.

FUNNEL administration scripts are run on two main servers at DESY – *leila* and *alien*. Figure 4.3 illustrates how FUNNEL works. The *mega.send*

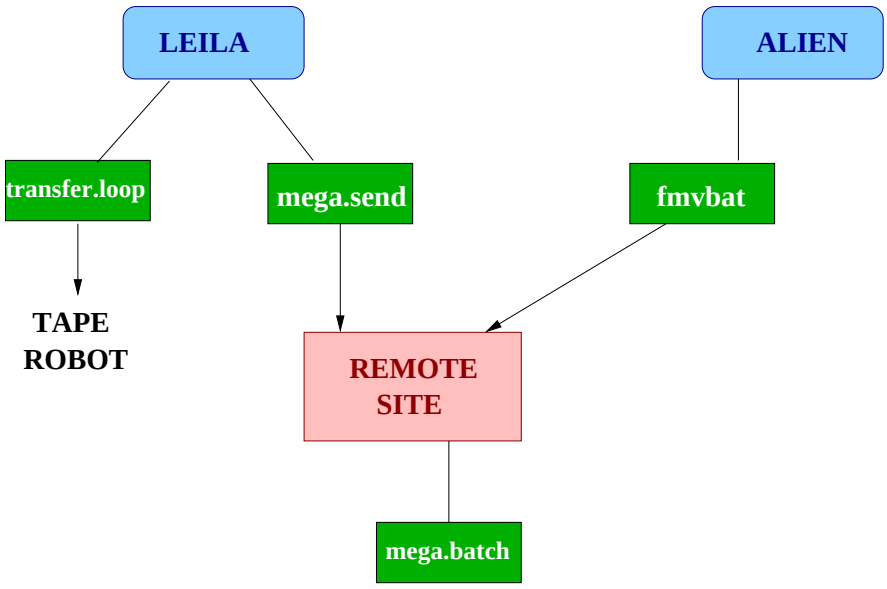


**Figure 4.3:** *A schematic view of the FUNNEL system. The dark (green) boxes correspond to the executed processes, the light (pink and blue) boxes stand for computers. Lines indicate where processes are executed, arrows- connection and direction of connection. The figure was taken from [88].*

script transfers input files for each MC job, which include an output and a log-file from an event generator, from the main servers to remote computer clusters, hereafter referred to as *sites*. Each site regularly runs the *mega.batch* script that starts jobs locally and also plays the role of administrator of available local resources. After a successful completion of a job, another script *fmvbat* collects output files together with the log-files of the MOZART, CZAR and ZEPHYR execution at DESY. Finally, the *transfer.loop* script transfers all output files to tapes in the DESY dCache system [89]. The status of the system is monitored by the *mega.stat* script. It is responsible for a variety of tasks: starting with an update of the status of jobs on the remote sites and ending with an update of the database of the finished jobs. Since the system was developed a long time ago, information about the status is mostly stored in formated text-files. Connection is always made from DESY to a remote site for all operations.

An overview of a job execution procedure on a site is shown on Fig. 4.4. As it was said, the main administration script on a site is *mega.batch*. If there are free computers it starts a process called *chain_run*, which starts and controls the job execution. First it starts the MOZART, CZAR and ZEPHYR

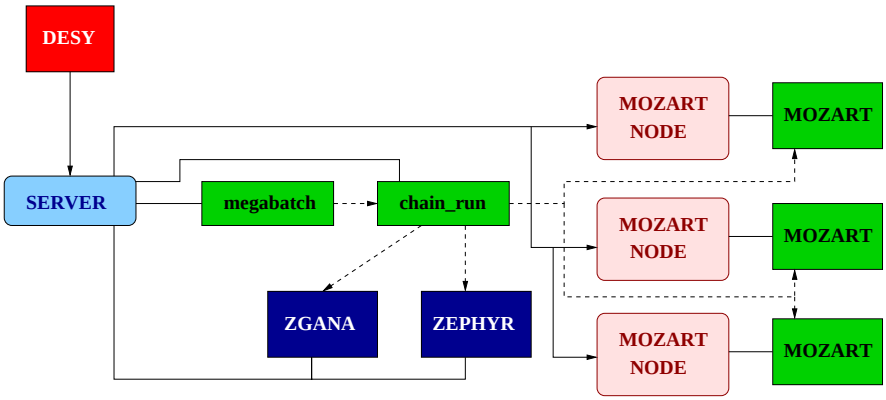


**Figure 4.4:** *A schematic view of a Monte Carlo job execution procedure. Solid lines show which process is run where, whereas solid arrows show connections between PCs. Dashed arrows indicate which processes are started by others. The figure was taken from [88].*

instances and then mounts a pipeline between them. Since MOZART is much more time-consuming than CZAR and ZEPHYR ($O(1\,\mathrm{s/event})$ for the former versus $O(0.1\,\mathrm{s/event})$ for the latter two), its execution is run in parallel streams. For this task both separate computers and individual processor cores could be used. For the execution of CZAR and ZEPHYR one standard modern PC is sufficient.

## 4.4 Monte Carlo production on GRID

In 2002 a dedicated project for the LHC computing was launched – the *Worldwide LHC Computing Grid (WLCG)*. Since 2004 ZEUS has used the European infrastructure of the WLCG, *Enabling Grids for e-science (EGEE)*, to run MC jobs [83]. This increased the production power of the FUNNEL system dramatically and by now most of the events – 80% on average – are produced on GRID.

### 4.4.1 GRID infrastructure

An overview of the EGEE GRID structure is depicted in Fig. 4.5. A user identifies himself/herself to the GRID system with an X.509 certificate. It consists of a private and a public key, which are used to decrypt and encrypt messages. One of them is stored at the *User Interface machine (UI)* and

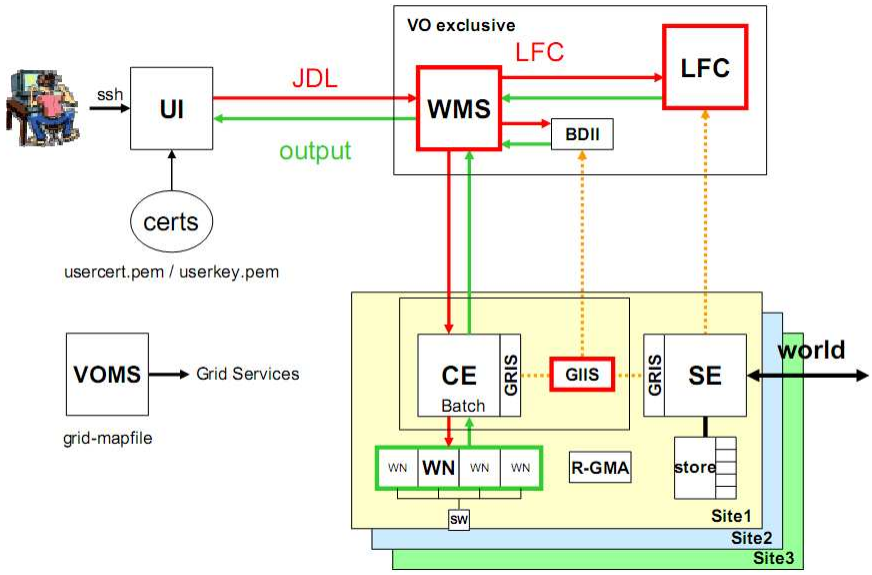


**Figure 4.5:** *A schematic view of the EGEE GRID network. The arrows show connections between the elements of the system.*

another is sent to a host to establish a connection. A job description, which can include small-sized files if needed, is sent to a *Workload Managment System, WMS,* in the *Job Description Language (JDL)* format [90]. The JDL file includes information about collections of requested input and output files, conditions to a cluster to be used for the job execution, etc. Depending on this information as well as on the data about the resources available to the user, which is stored in the *Berkley Database Information Index (BDII)* [91], the WMS makes a decision on which site, *Computing Element (CE)*, the job will be sent for execution. In the following a local batch system controls the job execution on a single element of the CE, a *Working Node (WN)*. Upon the end of the job the CE notifies the WMS and sends back the requested output files. For storage of large size files each CE has a dedicated *Storage Element (SE)*. A set of *lcg-utils* [92] was used to provide a uniform access to different underlying SE technologies. Moreover, a *Logical File Catalog (LFC)* [93] was exploited to map physical file names on SE's to a logical file name in the dedicated data base installed at DESY. The user can retrieve the job history on each step of execution.

### 4.4.2 ZEUS MC production on GRID

The MC job execution scheme on a WN is similar to the standard FUNNEL scheme described in Section 4.3.3. The major differences are that the MOZART, CZAR and ZEPHYR processes are run in series rather than in parallel and each input MC job is split into smaller sub-jobs to speed-up production. Consequently, the local job monitoring system is much simpler.

For each MC job the main Perl-script is sent for execution on a CE together with compatibility libraries for 32-bit binaries (all ZEUS software is compiled on the 32-bit architecture, whereas most of the WN's in the EGEE GRID are nowadays based on 64-bit Scientific Linux 5). To prepare for execution on a WN the script downloads from the nearest SE the input file ($\sim 100$ MB in size), an archive of calibration constants ($\sim 270$ MB) and the requested versions of the MOZART, CZAR and ZEPHYR binaries ($\sim 20$ MB each). After successfull completion of the job an output file ($\sim 600$ MB for each of $\approx 10$ sub-jobs) is copied to the DESY SE, whereas log files are retrieved back through a WMS. The log files are automatically checked on the ZEUS GRID UI to monitor potential errors during production. If any problems would be spotted the job would be resubmitted to GRID by the monitoring system, otherwise the output is transfered to the final storage destination on tapes and an entry in the dedicated MySQL database is created.

In ZEUS the production rate on GRID is limited by the submission time. The processing time for a job submission request at the DESY WMS's is $\approx 15$ s. Therefore the total number of jobs that can be submitted to the system is $\approx 5750$ jobs/day. There is an intrinsic restriction on the size of the MOZART output file, hence restricting the number of events processed in a sub-job. Up to now ZEUS has used 2500 Monte Carlo events per job, which is very close to the truncation boundary and for special classes of events with a large number of particles the boundary is crossed even with the current set-up. Thus a natural maximum for the ZEUS production rate on GRID is $\approx 14$ million events per day. To speed-up production the whole system was migrated to a compressed event output between the simulation steps. The tests have shown that this will increase the number of events in a sub-job at least by 1000. Thus in the nearest future this update will increase the ZEUS Monte Carlo production on GRID up to $\sim 20$ million events/day.

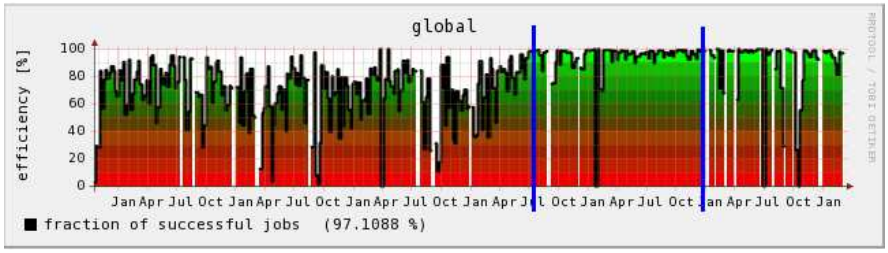


**Figure 4.6:** *The job efficiency for the Monte Carlo on GRID production in ZEUS for the period of October 2004 – March 2011. The dips are caused by both system problems and empty job queues.*

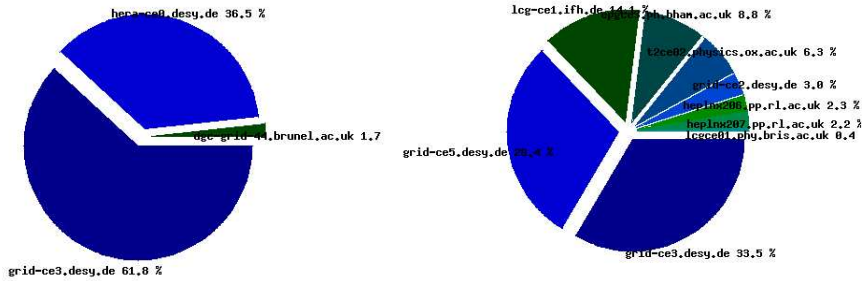


**Figure 4.7:** *A pie chart showing the relative contributions from the different CE's to the ZEUS MC production on GRID. The left pie chart shows a summary for the period 01/08/2008–01/10/2008, the right chart for 28/08/2009–28/09/2009.*

### 4.4.3 Statistics of the GRID usage

The contribution of GRID resources to the total FUNNEL production is vital ($\approx$ 82%). Hence the stability of GRID resources availability is of particular importance.

Figure 4.6 illustrates the efficiency of the job production on GRID, which is defined as the fraction of events, that were successfully finished in a unit of time. Lines indicate the time period, when the author was in charge of the ZEUS MC production on GRID. Over the last 2 years the system has been very stable and the averaged job efficiency has been at the level of $90-95\%$.

An another important characteristics is the number of CE's included in the production. A large number of sites makes the production independent of CE shutdowns, site hardware or software problems and high load of an individual site. The author has revived usage of the Service Availability

Monitoring (SAM) tests [94] to check access to sites looking for new resources and to identify problems with the used computer clusters. Figure 4.7 shows the result of author's contribution in the diversity of CE's in use. The left plot shows the dominant contribution from the two DESY CE's with a small admixture from one site from the UK, whereas the right plot indicates multiple sites from Germany and the UK.

# Chapter 5

# Event reconstruction

In the analysis presented in this thesis the $D^+$ signal is extracted from a peak in the mass spectrum of track combinations. Hence precise tracking is essential. Moreover, since both the track momentum measurement and the reconstruction of a displaced secondary vertex have been exploited, the ZEUS tracking algorithms as well as the ZEUS vertexing procedure will be reviewed in this Chapter. In the following a procedure to match tracks with the calorimeter information is described. This will be used to increase the precision of the reconstructed kinematic parameters of the hadronic system in an event. Next the procedures of the scattered electron identification and event kinematics reconstruction are discussed. The performances of different methods are compared.

## 5.1 Track reconstruction

Information about parameters of charged particles produced in the interaction is extracted from their *tracks*. Tracks are the flight trajectories in the detector that are reconstructed from the individually detected interactions of the particle with the active volume of the detector, *hits*. Moreover, the ZEUS track reconstruction procedure takes into account the uncertainties of the hit position measurements, the dead material distribution and effects of multiple scattering. The track type depends on the tracking detectors, that were used for track reconstruction. In this thesis the so-called *ZTT tracking* was used as the most global and precise approach. It combines information from the CTD and the MVD and refines the trajectory by application of the Kalman filter technique [95]. Tracks are reconstructed in two stages:

- **Pattern recognition.** The first stage is performed in multiple steps by the VCTRACK package [96]. It starts from the outermost tracking detector layer, which is the CTD SL9 for the central $\eta$ region, where the track density is lower than close to the interaction point. Combinations of three CTD hits from axial, odd-numbered, CTD superlayers form the tracking *seeds*. A track seed is extrapolated inward, gathering additional hits with increasing precision as the trajectory parameters are updated. A very broad "virtual" hit is added at the beam line to guide the trajectory. After a *"road"* of hits from the CTD through the MVD to the interaction point has been created, a least-squares fit of the track is done using the selected hits on the road in order to determine the helix[1] parameters at the beginning of the helix. In general the tracking reconstruction is not restricted to tracks with hits in all tracking devices; it also accounts for the so-called CTD-only and MVD-only tracks with hits in only one sub-detector.

- **Trajectory refinement.** A track fit is performed with the Kalman filter to improve precision of the helix parameters in the vicinity of the interaction point [97]. As input it takes the fit output from the pattern recognition stage. The track fit is applied recursively in three steps: *prediction, filtering* and *smoothing*. At the prediction step the present state $i$ (i.e. hits that have already been used for the trajectory estimation) is used to predict the position of the next $(i+1)$-th hit on the next detector sensor (could be a CTD wire or an MVD sensor). At the following filtering step the predicted and the measured values for the $(i+1)$-th hit positions are combined. At the last step a smoothing of the whole trajectory is performed and the covariance matrix is updated.

The Kalman filter examines the hits individually, whereas the least-squares fit evaluates all hits simultaneously. Removing wrong hits from the least-squares fit requires all fits to be redone from the start, which results in a large calculation time as the number of hits increases. Therefore the effect of multiple scattering on a trajectory is easier to incorporate in the Kalman filter fit.

Detailed information on the mathematical framework of the track fitting with the ZEUS detector is provided in [98].

[1] A track of a charged particle in a homogeneous magnetic field has a shape of helix due to the Lorentz force in the plane perpendicular to the magnetic field.

## 5.2 Vertex reconstruction

Proper identification of both the primary point of interaction and the $D^+$ decay vertex in an event is of particular importance for this analysis. Their position is reconstructed as the *primary* and *secondary vertices* fitted to sets of tracks. In analogy to the track fitting procedure, the vertex pattern recognition is performed first with the VCTRACK package and further refinement is applied later.

The vertex pattern recognition starts with a loose constraint that the primary vertex should be found along the proton beam line. Track pairs that are compatible with this soft constraint as well as with a common vertex are combined with other track pairs. The final choice of the primary vertex position after the pattern recognition stage is the vertex with the best overall $\chi^2$. To improve the precision of the vertex position measurement the Deterministic Annealing Filter (DAF) [99] has been introduced in ZEUS. In the chosen approach the vertex position is measured iteratively by calculating a weighted sum of the $\chi^2$ contributions from individual tracks to the vertex. The weights, $w$, are parametrised as a function of $\chi^2$ and they differ between iterations by a value of parameter $T$, which is known as *"temperature"*:

$$w(\chi^2, T) = \frac{1}{1 + \exp(\frac{\chi^2 - \chi^2_{\text{cut}}}{2T})} \tag{5.1}$$

In this way a hard $\chi^2_{\text{cut}}$ cut on outliers is replaced by a smooth temperature-dependent weight function. The DAF vertexing starts with a high value of $T$ and then the temperature is gradually reduced in the subsequent iterations until either convergence of the $\chi^2$ value for the vertex had been found or a certain minimal $T$ value had been reached. Between iterations the helix parameters and covariance matrices for the fitted tracks are recalculated at the reconstructed vertex position.

For the primary vertex fit a further improvement in precision is possible by introduction of a contraint on the vertex position to be close to the averaged interaction point, the so-called *beamspot*. The beamspot is defined as the overlap region of the colliding beams. It has a width of roughly $80 \times 20$ μm in the $XY$ plane [100] and it is too broad in the $Z$ direction to use this information as a constraint. The beamspot position varies in time and depends on $Z$. The former occurs due to the magnet movement, temperature effects as well as luminosity, background and polarisation tuning, whereas the latter arises from the fact, that the detector and the beam axes are not parallel. Thus the beamspot position is measured every 2000 events to account for

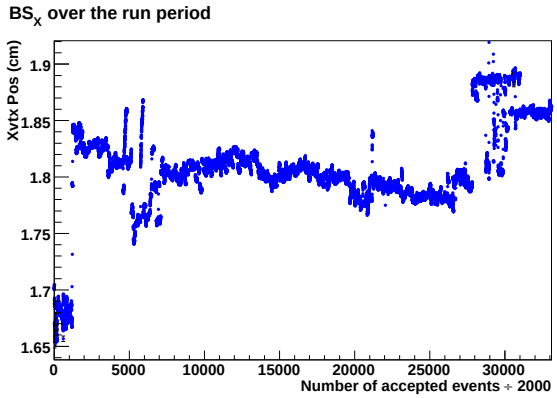


**Figure 5.1:** *The measured $x$ position of the beamspot during the period* 2006 – 2007 *[101]. The $X$ axis is shown in units of 2000 events.*

the time dependence. As an example, Fig. 5.1 shows the measured time dependence of the beamspot $x$ position in the period 2006 – 2007. The beam slope with respect to the $Z$ axis, *beam tilt*, was measured for a whole fill of the accelerator with the beams, as it was not seen to vary as much as the beamspot position. The beam tilt was evaluated as a slope of the mean values of the $x$ and $y$ vertex profiles fitted by the Gaussian function in several $z$ intervals. More information on the determination of the beamspot can be found in [101]. Figure 5.2 shows the effect on the $x$ coordinate reconstruction of the beamspot constraint in the DAF vertex fit.

In the case of the secondary vertices, e.g. the $D^+$ decay vertex, the fit is made with the same algorithm skipping the step of the pattern recognition, since the combination of tracks is chosen based on its compatibility with the $D^+$ mass. For each secondary vertex the corresponding *reduced primary vertex* was recalculated removing the secondary-vertex tracks and repeating the standard primary vertex fit.

## 5.3 Hadronic system reconstruction

To get the most precise hadronic energy measurement, information from the calorimeter and the tracking detectors can be combined into the so-called ZEUS Unidentified Flying Objects (ZUFOs)[2]. The energy resolution of the CAL develops for higher particle energies as $\sigma(E)/E \sim 1/\sqrt{E}$, while the tracking momentum resolution, parametrised by $\sigma(p_T)/p_T = a \cdot p_T \oplus b \oplus c/p_T$,

[2]ZUFOs are also referred to as Energy Flow Objects (EFOs) in ZEUS publications.

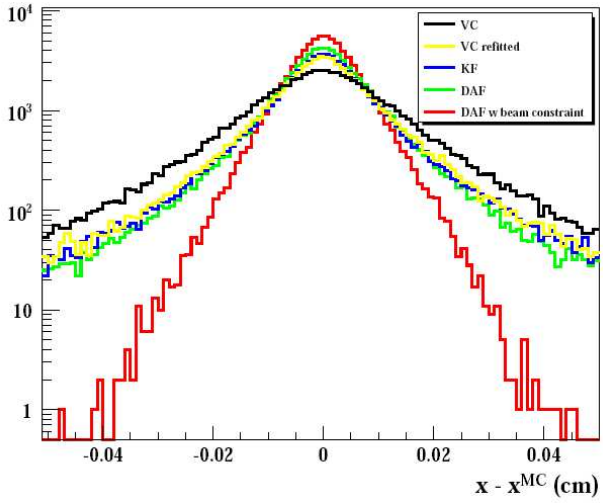


**Figure 5.2:** *A comparison of the DAF vertex reconstruction with various vertexing algorithms [102]. The DAF vertexing by itself (medium-dark blue line) brings only slight improvement in precision but increases efficiency. DAF vertexing with the beamspot constraint (light red line) significantly improves the position measurement. The Y axis is shown in arbitrary units.*

gives a better energy estimate for lower particle momenta (see Fig. 5.3). For neutral particles only CAL information could be used, whereas for charged particles the tracking information is mainly used below 10 GeV while calorimeter energy is used above.

ZUFOs are constructed in the following steps:

1. Adjacent cells[3] in the EMC, HAC1 and HAC2 sections are separately clustered into *cell islands* (see Fig. 5.4). The combination algorithm provides a unique association of a cell to its highest energy neighbour.

2. The cell islands from the previous stage are used as an input to clustering in the $(\theta, \phi)$ space. The procedure starts from the outermost layer of the CAL and goes inward by calculating the angular separation between neighbouring cell islands. The separation is used to associate the input islands together to form 3D energy clusters called *cone islands*. The position of a cone island is determined by the logarithmic

[3]The algorithm that connects the "nearest neighbours" was used, which means that cells at the corners are not combined.

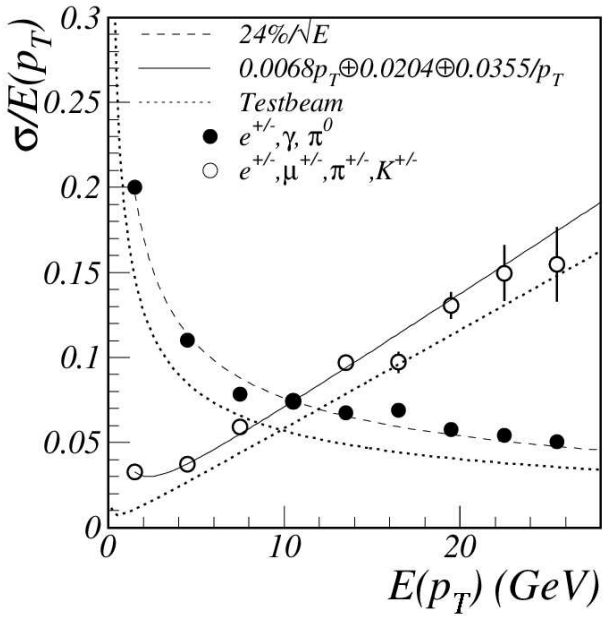


**Figure 5.3:** *The resolution of the electromagnetic energy in the EMC (filled circles) and of the $p_{\mathrm{T}}$ measurement in the CTD (open circles), obtained from a single-particle Monte Carlo sample[103].*

center-of-gravity of the energy deposit[4].

3. Charged tracks, which were fitted to a vertex and which pass certain requirements, are extrapolated to the surface of the CAL taking into account the magnetic field. The selected tracks should fullfill the following requirements:

   - $0.1 < p_{\mathrm{T}} < 20\,\mathrm{GeV}$ for tracks with hits in at least 4 CTD superlayers;
   - $0.1 < p_{\mathrm{T}} < 25\,\mathrm{GeV}$ for tracks with hits in at least 7 CTD superlayers;

   A track and a cluster are matched if either the *distance of closest approach (DCA)* between the track and the position of the cone island is less than 20 cm or the track lies within the area of the island. As a result of this procedure, groups of cone islands and tracks, which will be referred to as ZUFOs, are formed.

[4]Using logarithmic weights instead of the linear sum takes into account the exponential falloff of the transverse shower energy distribution from the shower maximum.

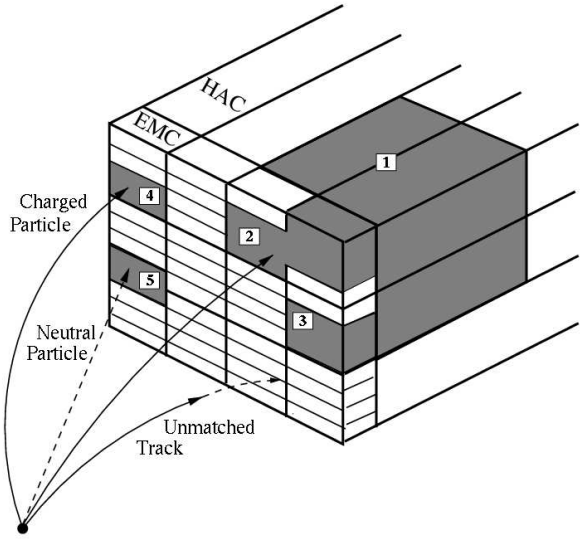


**Figure 5.4:** *This schematic picture shows how calorimeter cells are clustered into four EMC cell islands and one HAC cell island [103]. The EMC cell islands labelled 2 and 3 are joined with the HAC cell island 1 to form a cone island. The obtained cone island as well as the island 4 are matched to tracks, which means that these were charged particles. No track is found for island 5, which is treated as a neutral particle. The unmatched track corresponds to a low-momentum particle.*

4. The combination of the information from the CAL and the tracking system is carried out in the following way:

   - if one track is matched to one cone island the ZUFOs energy is taken either from the CAL cluster or from the matched track momentum depending on which measurement has better resolution;
   - for good tracks that were not associated to islands the energy is derived from the momentum measurement with the assumption that the particle is a charged pion;
   - cone islands that were not matched to any track are treated as neutral particles and the CAL energy is used;
   - cone islands with more than three associated tracks are treated as jets and energy is taken from CAL;
   - if a track is matched to multiple islands or two tracks are matched to one or two island the algorithm is similar to the one-to-one matching, but using the sum of energies or momenta instead.

Further corrections had to be applied to account for: the material budget of the detector, the inefficiency in the regions of cracks between the CAL sections, the presence of muons (muons do not release all their energy in the CAL, thus if the CAL info is used the energy would be underestimated) and the imbalance in the compensation effect for low momentum ($\sim 1\,\text{GeV}$) hadrons. Detailed information on the ZUFO reconstruction is available in [103].

In this analysis the reconstructed ZUFOs have been used to determine the kinematics of the hadronic system as well as in the DIS kinematics calculation.

## 5.4 Electron identification

The identification of the scattered electron is essential for the NC DIS event selection. The scattered electron leaves a clear signature which differentiates the neutral current DIS events from the charged current DIS, where the neutrino escapes undetected, and PHP, where the scattered electron escapes through the beam hole. There have been two main electron finders developed in ZEUS – the neural-network-based SINISTRA95[5] [104] and the probabilistic EM [105]. The former is tuned for the kinematic region of the measurement presented in this thesis, whereas the latter is better for the high-$Q^2$ region, where the electron is reconstructed in the BCAL.

SINISTRA candidate clusters are formed using the *next-to-nearest neighbour* algorithm[6] on calorimeter towers to produce islands and then merging the islands from different calorimeter sections. The position of the energy deposit within a cell is reconstructed from the imbalance of two readout photomultipliers in the cell. These energy deposits are used to calculate the longitudinal and transverse energy distributions in the original shower and this information is passed to the neural network. The neural network has been trained using Monte Carlo simulated hadronic and electromagnetic clusters in the RCAL. As an output SINISTRA returns a number between 0 and 1, which represents the probability of the cluster to be the scattered electron. In the following only the candidate with the highest probability will be considered. The identified electron is assigned the energy of the reconstructed CAL cluster.

[5] Often also referred to as SINISTRA.

[6] This algorithm allows diagonal associations in contrast to the nearest neighbour algorithm described in Section 5.3

## 5.5 Kinematic variables

After reconstructing the scattered electron and the hadronic system in an event, it is possible to calculate the kinematic variables $Q^2$, $x$ and $y$. To do this there are four quantities available: energy and polar angle of the scattered electron, $E'_e$ and $\theta_e$, as well as the total hadronic transverse momentum, $P_{\mathrm{T,had}}$, and the hadronic angle, $\gamma_{\mathrm{had}}$. The latter two are defined as:

$$P_{\mathrm{T,had}} = \sqrt{\sum_i (P^i_{x,\mathrm{had}})^2 + \sum_i (P^i_{y,\mathrm{had}})^2}, \tag{5.2}$$

$$\gamma_{\mathrm{had}} = \frac{\sum_i E^i_{\mathrm{had}} cos\theta^i}{\sum_i E^i_{\mathrm{had}}} = \frac{P^2_{\mathrm{T,had}} - \delta^2_{\mathrm{had}}}{P^2_{\mathrm{T,had}} + \delta^2_{\mathrm{had}}}, \tag{5.3}$$

where $\delta_{\mathrm{had}} = E - P_z = \sum_i (E^i_{\mathrm{had}} - P^i_{z,\mathrm{had}})$ and the sum runs over whole hadronic state except the scattered electron. The hadronic angle $\gamma_{\mathrm{had}}$ defines the struck quark flight direction in the naive QPM model.

In ZEUS there are three main methods used to reconstruct the kinematic variables depending on which combination of the electron and hadronic parameters is used.

### 5.5.1 Electron method.

The *electron method* (el) uses only electron energy and scattering angle. The kinematic variables can be calculated as following:

$$Q^2_{el} = 2E_e E'_e (1 + cos\theta_e), \tag{5.4}$$

$$y_{el} = 1 - \frac{E'_e}{2E_e}(1 - cos\theta_e), \tag{5.5}$$

$$x_{el} = \frac{Q^2_{el}}{s\, y_{el}}, \tag{5.6}$$

where $s$ is the $ep$ center-of-mass energy.

### 5.5.2 Jacquet-Blondel method.

The so-called *Jacquet-Blondel method* (JB) [106] relies exclusively on the parameters of the hadronic final state. The advantage of this method is that

it can be used even if the scattered electron can not be detected (e.g. charged current DIS events). The kinematic parameters are expressed as:

$$Q^2_{JB} = \frac{P_{\mathrm{T,had}}}{1 - y_{JB}}, \tag{5.7}$$

$$y_{JB} = \frac{\delta_{\mathrm{had}}}{2E_e}, \tag{5.8}$$

$$x_{JB} = \frac{Q^2_{JB}}{s\, y_{JB}}, \tag{5.9}$$

### 5.5.3 Double-angle method.

The *double-angle method* (da) [107] uses information from both the scattered electron and the hadronic system. As inputs it takes the hadronic angle, $\gamma_{\mathrm{had}}$, and the electron scattering angle, $\theta_e$. The kinematic variables can be calculated using the relations:

$$Q^2_{da} = 4E_e^2 \frac{\cot\frac{\theta_e}{2}}{\tan\frac{\theta_e}{2} + \tan\frac{\gamma_{\mathrm{had}}}{2}}, \tag{5.10}$$

$$y_{da} = \frac{\tan\frac{\gamma_{\mathrm{had}}}{2}}{\tan\frac{\theta_e}{2} + \tan\frac{\gamma_{\mathrm{had}}}{2}}, \tag{5.11}$$

$$x_{da} = \frac{Q^2_{da}}{s\, y_{da}}, \tag{5.12}$$

The advantages of this method is that, typically, the angular resolution is better than energy resolution and that it is only weakly dependent on the CAL calibration. Thus the double-angle method can lead to a more precise measurement of the kinematics in a large part of the phase space investigated in this analysis.

### 5.5.4 Comparison of reconstruction methods.

A comparison of resolutions of the different kinematic reconstruction methods was done on an inclusive neutral current DIS sample generated with ARIADNE. The bias and resolutions of the reconstruction methods for some variable $A$ were determined from the mean and root-mean-square

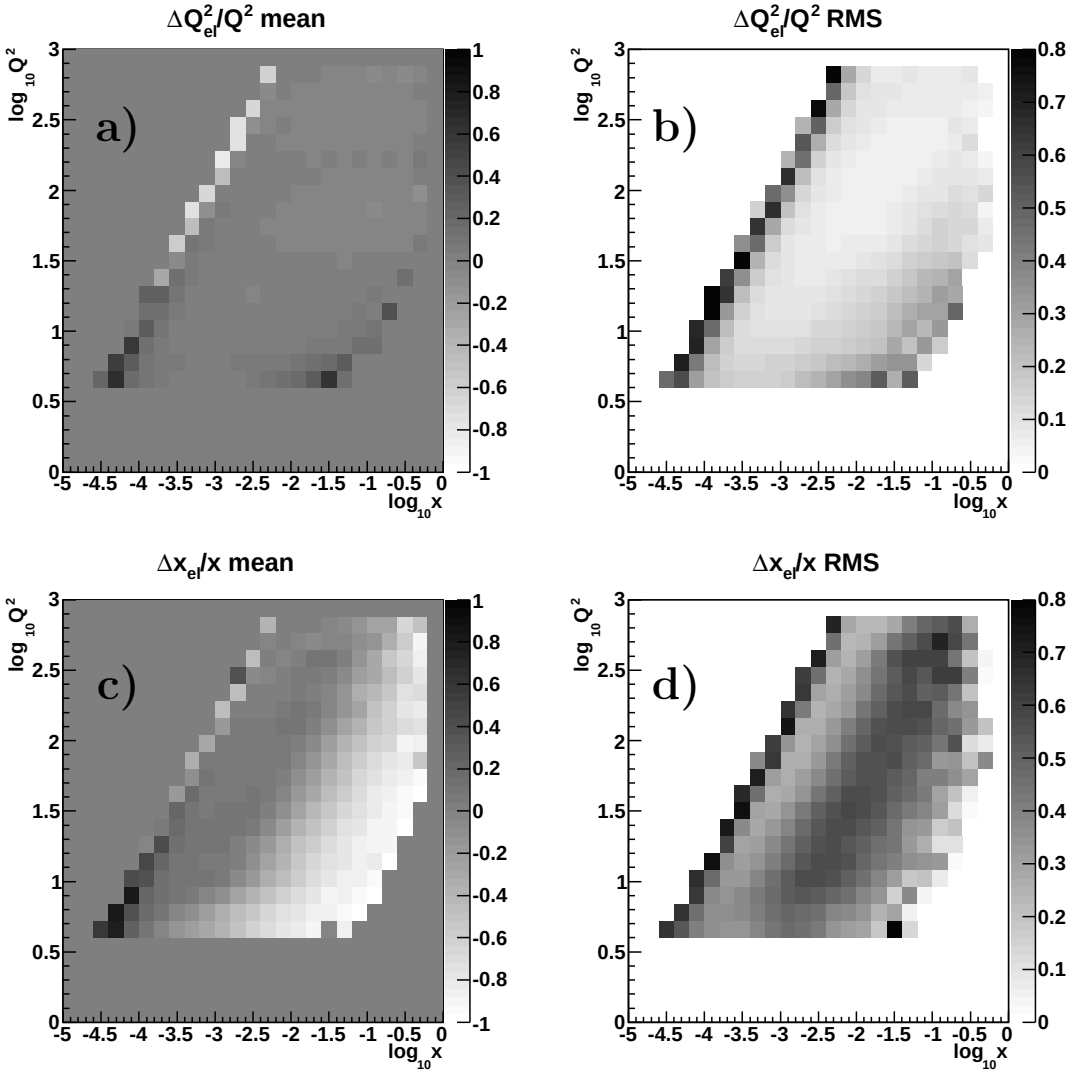


**Figure 5.5:** *The mean values (a and c) and the RMS (b and d) for $Q^2$ (a and b) and $x$ (c and d) using the electron method. The colour scheme is shown individually for each plot.*

(RMS) of the relative difference between the reconstructed and true values for $A$:

$$\frac{A_{rec} - A_{true}}{A_{true}} \tag{5.13}$$

Figures 5.5, 5.6 and 5.7 show the mean values and RMS for $Q^2$ and $x$ with the electron, double-angle and Jacquet-Blondel methods, respectively. The

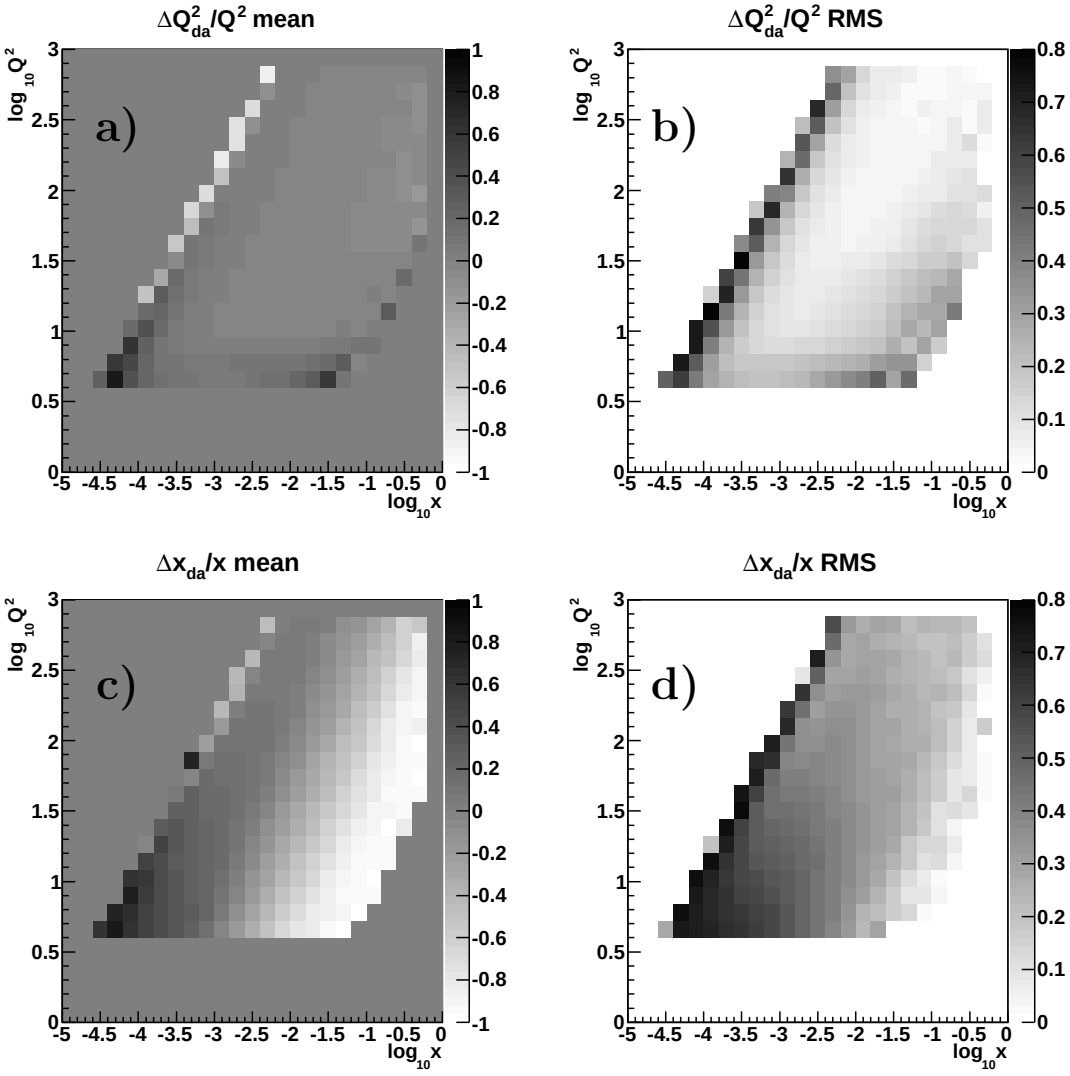


**Figure 5.6:** *The mean values (a and c) and the RMS (b and d) for $Q^2$ (a and b) and $x$ (c and d) using the double-angle method. The colour scheme is shown individually for each plot.*

study was done double-differentially in $Q^2$ and $x$ to examine the dependence of the performance on phase space. The $Q^2$ and $x$ variables were chosen to be directly comparable to [107]. One can immediately see that the Jacquet-Blondel method has the poorest resolution in most of the kinematic plane. In general, the electron and double-angle methods show similar performances. The former is better at low $Q^2$ and low $x$ (and hence high $y$), whereas the

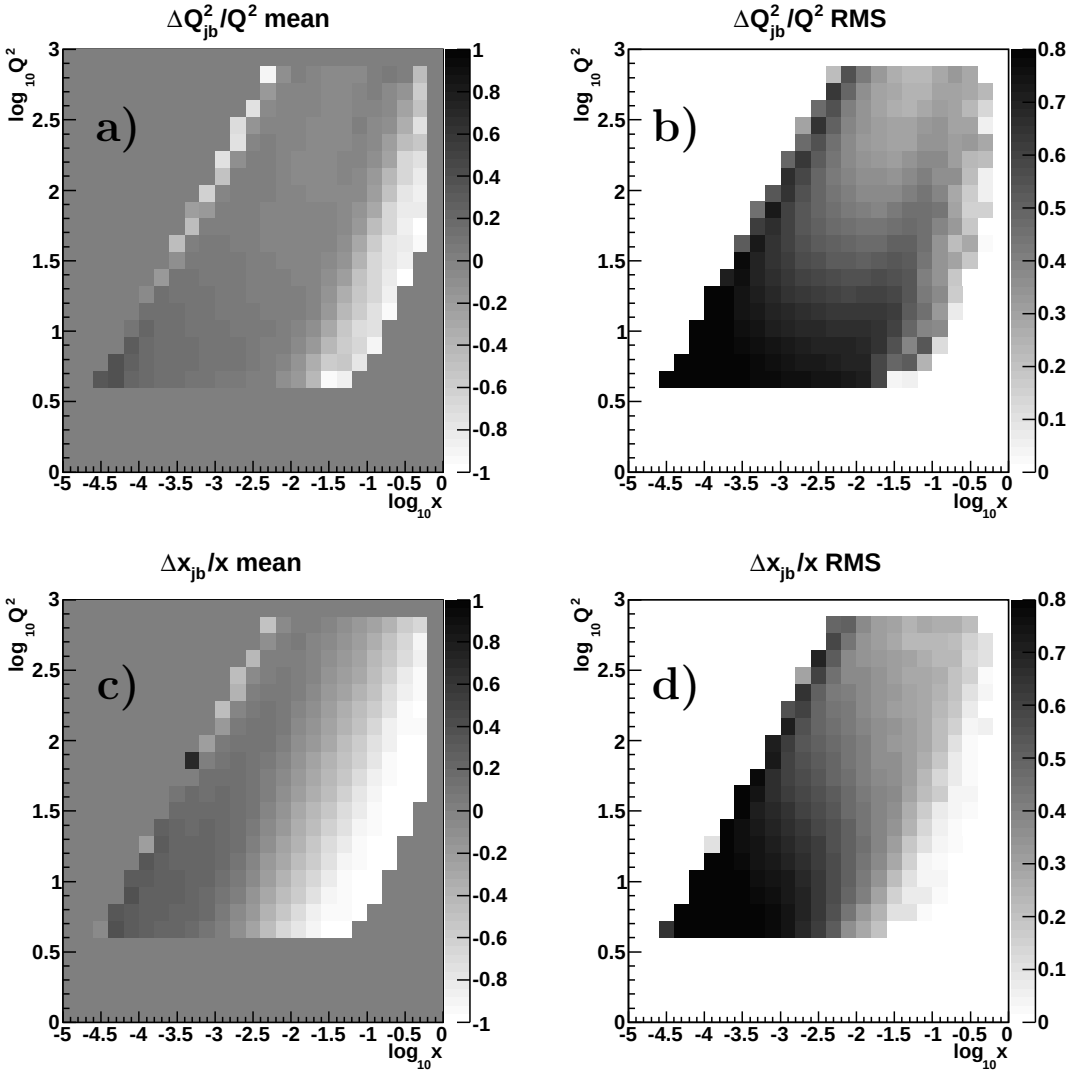


**Figure 5.7:** *The mean values (a and c) and the RMS (b and d) for $Q^2$ (a and b) and $x$ (c and d) using the Jacquet-Blondel method. The colour scheme is shown individually for each plot.*

latter is slightly better at medium and high $Q^2$ and high $x$ (and thus low $y$), especially for the $x$ reconstruction. The double-angle method performs better than the other two methods in larger part of the kinematic plane of this analysis. Therefore it was used to reconstruct the kinematic variables.

# Chapter 6

# Event and $D^+$ candidate selection

This Chapter describes the selection of deep inelastic scattering events as well as $D^+$ candidates. The usage of secondary vertices for the $D^+$ signal extraction is stressed, since it considerably improves the separation between the heavy and light flavours. The Monte Carlo distributions are validated versus data and corrections for the Monte Carlo simulations are extracted.

## 6.1 Data and Monte Carlo samples

For this measurement data collected by ZEUS in 2005–2007 with the center-of-mass energy 318 GeV have been used. Both electron-proton and positron-proton collisions were used, because the charm neutral current DIS cross sections at low $Q^2$ (1.13) are invariant with respect to the lepton charge. The positron–proton collision data collected in the 2003 – 2004 period were not included in the analysis due to technical reasons. Table 6.1 shows beam parameters and the integrated luminosity for the used sub-periods, $\mathcal{L}$. The total luminosity sums up to $323\,\mathrm{pb}^{-1}$.

The ARIADNE Monte Carlo generator was used to simulate inclusive NC DIS events (see Section 4.2.1). This sample was used to test the general detector performance and to study trigger efficiencies (e.g. see Sections 5.5 and 6.2.6). A summary of the generated samples for each period is given in Table 6.2. For the comparison with the data, the luminosity of the sub-samples for each year were normalised to the corresponding data luminosity.

| **Year** | **Beams** | $\sqrt{s}$, GeV | $\mathcal{L}$, pb$^{-1}$ | $\sigma_{\mathcal{L}}$, % |
|---|---|---|---|---|
| 2005 | $e^- p$ | 318 | 134 | 1.8 |
| 2006 | $e^- p$ | 318 | 52 | 1.8 |
| 2006–2007 | $e^+ p$ | 318 | 137 | 1.8 |
| total | | | 323 | 1.8 |

**Table 6.1:** *The data samples collected by ZEUS in 2005–2007.*

| **MC type** | $Q^2_{\rm MC}$ **region** | Integrated luminosity, pb$^{-1}$ | | |
|---|---|---|---|---|
| | | 2005 $e^-$ | 2006 $e^-$ | 2006–2007 $e^+$ |
| ARIADNE | $Q^2_{\rm MC} > 4\,{\rm GeV}^2$ | 133 | 55 | 142 |
| RAPGAP $c$ inclusive | $Q^2_{\rm MC} > 4\,{\rm GeV}^2$ | 283 | 165 | 498 |
| | $1.5 < Q^2_{\rm MC} < 4\,{\rm GeV}^2$ | 147 | 55 | 142 |
| RAPGAP $b$ inclusive | $Q^2_{\rm MC} > 1\,{\rm GeV}^2$ | 2115 | 925 | 2578 |

**Table 6.2:** *The Monte Carlo samples used in the analysis. All samples were generated with* $\sqrt{s} = 318\,GeV$.

The RAPGAP generator was used to simulate production of the charm and beauty quarks (see Section 4.2.2). To achieve a satisfactory statistical precision, a significant amount of these events is required. Due to the steep $Q^2$ dependence of the charm cross section ($d\sigma/dQ^2 \sim 1/Q^4$ according to (1.13)), the charm sample was split into two parts: $Q^2_{\rm MC} > 4\,{\rm GeV}^2$ [1] and $1.5 < Q^2_{\rm MC} < 4\,{\rm GeV}^2$. The former was generated with $\approx 2-3$ times the integrated luminosity of the data, whereas the latter corresponds to 1 times the integrated luminosity of the data. Since the beauty cross section is much smaller it was possible to generate a large number of beauty events with $Q^2_{\rm MC} > 1\,{\rm GeV}^2$ and an integrated luminosity of 18 times the data. The luminosity numbers for the individual periods are tabulated in Table 6.2.

[1] Hereafter $Q^2_{\rm MC}$ corresponds to the $Q^2$ value at the generator level calculated from the incoming and outgoing lepton.

## 6.2 DIS selection

### 6.2.1 Online selection

The events studied in this thesis had to be triggered online by one of the inclusive DIS TLT slots:

**SPP02** Inclusive low-$Q^2$ trigger with small box cut (only in 2005):

- $30 < E - P_z < 100\,\text{GeV}$ (measured in the CAL);
- $E'_e > 4\,\text{GeV}$ (for the best candidate);
- Small box cut[2]: $|x| > 12\,\text{cm}$, $|y| > 12\,\text{cm}$.

**SPP09** Inclusive low-$Q^2$ trigger with medium box cut (since 2006):

- $30 < E - P_z < 100\,\text{GeV}$ (measured in the CAL);
- $E'_e > 4\,\text{GeV}$ (for the best candidate);
- Box cut: $|x| > 15\,\text{cm}$, $|y| > 15\,\text{cm}$.

**HFL17** Inclusive NC DIS with two tracks (since 2006):

- the same selection as in SPP02;
- 2 tracks in CTD;

**HPP31** Inclusive low-$Q^2$ trigger with small box cut and electron (since 2006):

- $34 < E - P_z < 75\,\text{GeV}$ (measured in the CAL);
- $E'_e > 7\,\text{GeV}$ (for the best candidate);
- $Q^2_{\text{TLT}} > 6\,\text{GeV}^2$;
- 1 track in CTD with $p_{\text{T}} > 0.2\,\text{GeV}$;
- a vertex in an event with $-60 < Z_{\text{vtx}} < 60\,\text{cm}$;
- Box cut: $|x| > 12\,\text{cm}$, $|y| > 12\,\text{cm}$.

[2]The "box cut" refers to a condition on the scattered electron position on the RCAL surface to be outside of a rectangle around the beampipe.

The SPP02 slot is the most inclusive trigger, but it was heavily prescaled[3] in 2006 – 2007. Thus events in 2005 were selected with SPP02 only, whereas events from the 2006 – 2007 period were selected using the logical OR of the remaining three triggers.

All TLT slots require exclusively SPP1 on the SLT level. The SLT SPP1 slot required any of FLT28, FLT30, FLT36, FLT40, FLT43, FLT44, FLT46, FLT47 to be triggered. These FLT bits have a large overlap region with low exclusive rates for each trigger; the highest unique rate was found for FLT30 (3 – 5% depending on the $Q^2$ region). All first level slots but FLT30 used some FLT tracking requirements in their logic. The Monte Carlo simulations were not able to reproduce the FLT tracking efficiency in the data and thus this had to be corrected for at the analysis level (see Section 6.2.6).

### 6.2.2 Scattered electron selection

The key aspect of the deep inelastic scattering selection is the identification of the scattered electron in the detector. The applied selection leads to good quality candidates and to the reduction of various backgrounds:

- **SINISTRA probability** $> 0.9$ **:** Ensures high purity for the electron candidate sample.
- **Electron energy** $E'_e > 10\,\mathbf{GeV}$ **:** This requirement grants high efficiency for the electron candidate selection while reducing the background from photoproduction events, where electrons can be "faked" by photons from the neutral pion decays.
- **Electron isolation criteria** $E^{\mathrm{cone}}_{\mathrm{non\,e}} < 5\,\mathbf{GeV}$ **:** A cut on the energy deposit around and not originating from the electron candidate in the CAL provided by the electron finder. The energy activity is calculated in a cone centered around the electron with a radius of 0.8 in the $(\eta, \phi)$ plane: $\sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.8$.
- **Electron position :** A cut on the electron position on the CAL surface, called *box cut*, was applied to remove events in which the electron passed through poorly described edges of the CAL around the beampipe: $|x| > 13\,\mathrm{cm}$, $|y| > 13\,\mathrm{cm}$. This cut is slightly larger than the one in the trigger logic (see Section 6.2.1) due to a difference between the online and offline versions of SINISTRA. In addition a box

[3] A prescaling factor of n means that only one out of n events, in which the slot was fired, is triggered by the slot.

on the RCAL surface that turned out not to be described by the MC simulations [108] was removed: $11 < x < 27\,\text{cm}$, $10.5 < y < 27\,\text{cm}$. The last cut was applied only on specific sub–set of data, which is set by run ranges 59600 – 60780, 61350 – 61580 and 61800 – 63000.

- **Geometry cuts :** Remove regions of the calorimeter, where reconstruction of the scattered electron is poor and/or hard to simulate due to their geometrical complexity:

  - electrons in the overlap region between the RCAL and BCAL ($\sqrt{x^2+y^2} > 175\,\text{cm}$ on the RCAL surface) were rejected;
  - electrons in the regions of cracks between the RCAL, BCAL and FCAL, *super-cracks*, were rejected: $164 < z < 174\,\text{cm}$, $-104 < z < -98.5\,\text{cm}$;
  - the region of gaps between halves of the RCAL were removed from the analysis: $6.5 < x < 12\,\text{cm}$ if $y > 0$ and $-14 < x < -8.5\,\text{cm}$ if $y < 0$;
  - "chimney" cut to remove a region in the RCAL, where cooling tubes and supply cables for the solenoid were mounted: $|x| < 12\,\text{cm}$ if $y > 80\,\text{cm}$

### 6.2.3 Background suppression

An additional selection was applied to suppress dominant backgrounds:

- $\boldsymbol{E - P_z}$: Energy and momentum conservation imply that the $E - P_z$ quantity is conserved. The total energy of an event before collision is equal to $E = E_p + E_e$, whereas the total longitudinal momentum component $P_z \approx E_p - E_e$. Thus their difference for an ideal NC DIS event should be equal $2E_e = 55\,\text{GeV}$. Particles that leave the detector unregistered through the forward beam hole ($E \approx p_z$) contribute only a little to the total measured $E - P_z$ value, whereas particles that escape through the rear beam hole ($E \approx -p_z$) reduce this quantity by twice their energy. Thus for PHP events, when the scattering angle of the electron is close to 0, the value of $E - P_z$ is lower– typically below $30\,\text{GeV}$. Therefore a cut $E - P_z > 40\,\text{GeV}$ had to be imposed to reduce the photoproduction contamination. A cut of $E - P_z < 65\,\text{GeV}$ was introduced to reject events of cosmic ray background and overlapping interactions. The value of $E - P_z$ is obtained summing $\sum_i E^i - P_z^i$ over all ZUFOs in an event.

- $\boldsymbol{-30 < Z_{\mathrm{vtx}} < 30}$ cm: Events of beam-gas interactions as well as cosmic ray events are suppressed by the requirement of the primary vertex within this narrow range along the $Z$ axis. In addition this condition cuts away "satellite bunch interactions"– interactions with the following or the previous bunch in the beam.

- $\boldsymbol{y_{JB} > 0.02}$: To reject events, where the hadronic system was not measured precisely, a condition on the $y$ variable reconstructed from the hadronic final state (Jaquet-Blondel method) was imposed.

### 6.2.4 Kinematic region

The double-angle method was used to reconstruct the kinematic parameters $Q^2$ and $y$, since it has shown the best performance in the kinematic region of the measurement (see Section 5.5):

- $5 < Q^2_{da} < 1000\,\mathrm{GeV}^2$;

- $0.02 < y_{da} < 0.7$.

### 6.2.5 Inclusive comparison with Monte Carlo.

In order to ensure good simulation of inclusive DIS events a comparison of kinematic distributions between data and Monte Carlo was done. Such kind of plots are often called "control plots" because they allow to control the level of agreement between the data and simulations. Figure 6.1 shows the control plots for important event variables. The data are reasonably well described by the MC apart from the scattered electron energy, $E - P_z$ and the fraction of tracks that are fitted to the primary vertex. The same observations were made in other analyses. The first two features originate from the necessity of corrections to the EMC and HAC calibrations at analysis level. The effect of the missing corrections will be accounted for in the systematic uncertainty evaluation. The origin of the discrepancy in the distribution of the fraction of the primary tracks is unknown. To check if this has an effect on on the results presented in this thesis the distribution in the MC simulations was weighted to describe the data. The effect on the $D^+$ cross sections was found to be negligible.

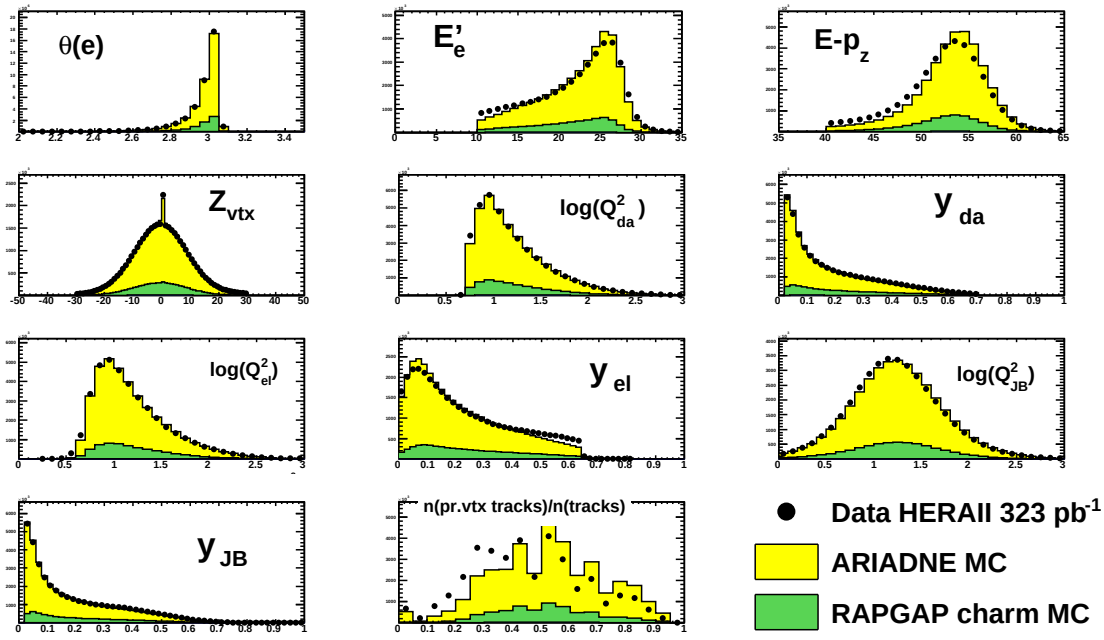


**Figure 6.1:** *Distributions for important parameters of DIS kinematics. The data (black points) are compared to the sum of light flavour production generated by* ARIADNE *(light yellow histogram) and charm production by* RAPGAP *(dark green histogram). The Monte Carlo simulations were normalised to the number of events in the data.*

### 6.2.6 FLT efficiency corrections

In previous analyses it has been discovered that the FLT tracking efficiency is not well described in Monte Carlo simulations [86]. For example, in Section 4.3.2 a recent update of the simulations has been presented, but this appeared only recently and it fixes only the overall normalisation. That is why corrections on analysis level had to be applied here.

All events at the FLT level are categorised with respect to the tracking information depending on the total number of tracks, $N(\mathrm{trk})$, and the number of tracks that were fitted to the primary vertex, $N(\mathrm{vertex{-}fitted\ trk})$. Figure 6.2 shows the so-called *"track class"* assigned to an event depending on these two quantities. The track classes that are relevant for the selection in the FLT triggers, which were used in this analysis, are:

**0 :** no tracks;

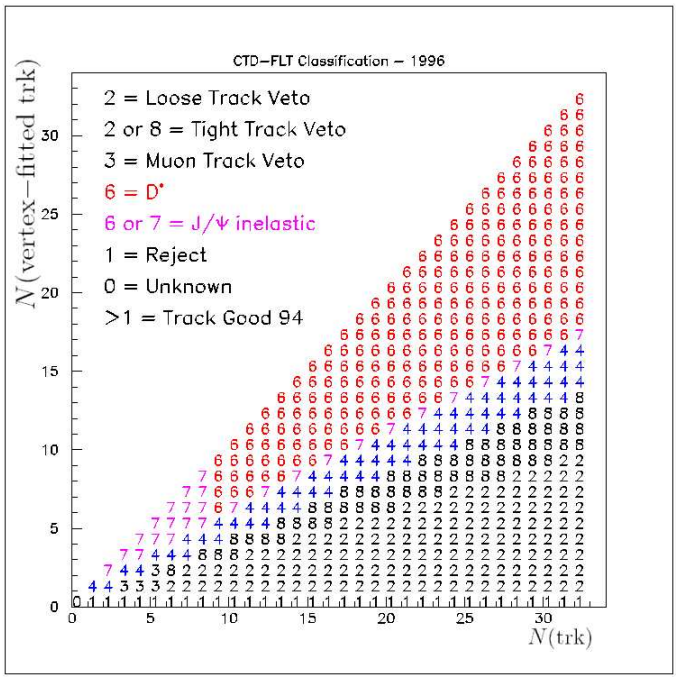


**Figure 6.2:** *Definition of the track classes that are assigned to events on the FLT level.*

**1 :** no vertex–fitted tracks;

**2 :** $N(\text{vertex}-\text{fitted trk})/N(\text{trk}) < 0.25$;

**8 :** $N(\text{vertex}-\text{fitted trk})/N(\text{trk}) < 0.38$;

All tracking requirements at the FLT level can be subdivided into five categories:

- **Loose tracking veto (LOOSE).** Reject events with tracking class 2: trk_class $\neq 2$;
- **Loose tracking veto + at least one vertex track (LOOSE + q95b).** Reject events with either tracking class 2 or tracking class 1: trk_class $\neq 2$ & trk_class $\neq 1$;
- **Semi-loose tracking veto (SEMILOOSE).** Reject events with either tracking class 2 or tracking class 8 and at least 26 tracks: trk_class $\neq 2$ & (trk_class $\neq 8$ & $N(\text{trk}) \geq 26$) ;

- **Semi-loose tracking veto + at least one vertex track (SEMILOOSE + q95b).** Reject events with either tracking class 2, or tracking class 1 or class 8 with at least 26 tracks: trk_class $\neq 2$ & trk_class $\neq 1$ & (trk_class $\neq 8$ & $N$(trk) $\geq 26$);

- **Tight tracking veto + at least one vertex track (TIGHT + q95b).** Reject events with either tracking class 2, or tracking class 1 or class 8: trk_class $\neq 2$ & trk_class $\neq 1$ & trk_class $\neq 8$;

The efficiency of the FLT tracking requirements has to be tested versus a reference FLT bit, which does not apply any tracking cuts and uses exclusively CAL information. FLT30[4] was chosen since it is the only such FLT bit that is included in the analysis trigger chain. The efficiency was defined as the ratio of the number of events that pass FLT30 and a specific FLT tracking requirement over the total number of events that pass FLT30:

$$eff = \frac{N(\text{FLT30 \& tracking})}{N(\text{FLT30})}. \tag{6.1}$$

The efficiency was determined for events that pass the standard DIS selection described in Sections 6.2.1 to 6.2.4. The measured efficiencies in the data and MC could be directly compared and the ratio of the two defines a correction, which one should apply on the MC to reproduce the tracking efficiency in the data. The correction was derived for each data taking period separately to account for the dependence of the CTD pulse height (e.g. see Fig. 4.2). The correction was parametrised as a function of the hadronic angle, $\gamma_{\text{had}}$. Figure 6.3 shows the dependence of the FLT tracking efficiency as measured in the data and in the MC simulations at low $Q^2$ in one of the data taking periods. The MC does not reproduce the shape in the data. An example of the correction obtained at low $Q^2$ in the corresponding data taking periods is presented in Fig. 6.4. It was observed that the correction for the LOOSE and SEMILOOSE tracking requirements as well as for the LOOSE+q95b and SEMILOOSE+q95b tracking requirements are the same and those were treated together. The largest correction is needed for the TIGHT+q95b requirements. The correction was fitted by polynomial functions of the first order for the LOOSE and SEMILOOSE tracking requirements, the fifth order for the LOOSE+q95b and SEMILOOSE+q95b tracking requirements and the forth order for the TIGHT tracking requirements. The fit output in

[4]The FLT30 slot requires: an isolated electron to be found in the RCAL, the total energy in the first ring of cells around the beampipe if the EMC section of the RCAL above 15 GeV, the total energy in the EMC section of the RCAL excluding the first ring above 2 – 4 GeV depending on the run number.

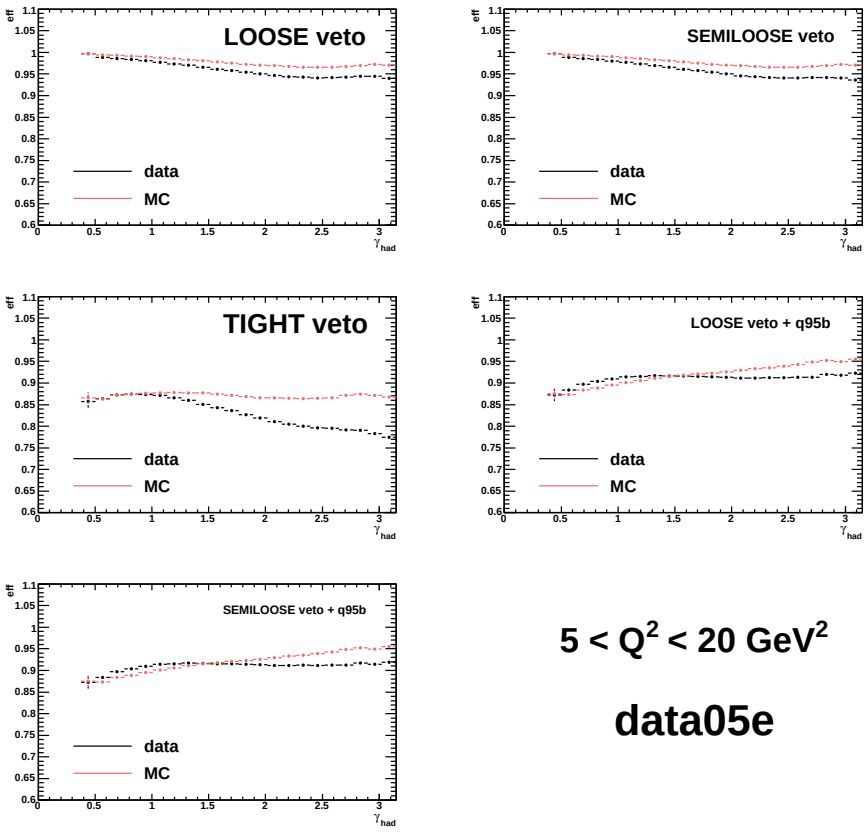


**Figure 6.3:** *The FLT tracking efficiency measured in events triggered by FLT30 in the data (black points) and in the Monte Carlo simulations (light red points). The binomial uncertainties are shown.*

various $Q^2$ bins is presented in Appendix A separately for different data taking periods.

The corrections had to be applied as a function of $\gamma_{\rm had}$ the following logic:

- if an event was triggered by the FLT30 slot $\Rightarrow$ no correction was needed;
- if FLT30 was not triggered, then if at least one of the slots with LOOSE or SEMILOOSE requirements (i.e. FLT36, FLT40, FLT47 for this analysis) was triggered $\Rightarrow$ LOOSE correction had to be applied;
- if neither FLT30 nor slots with LOOSE or SEMILOOSE tracking requirements were triggered, then if at least one of the slots with LOOSE+ q95b or SEMILOOSE+q95b requirements (i.e. FLT28, FLT43, FLT46

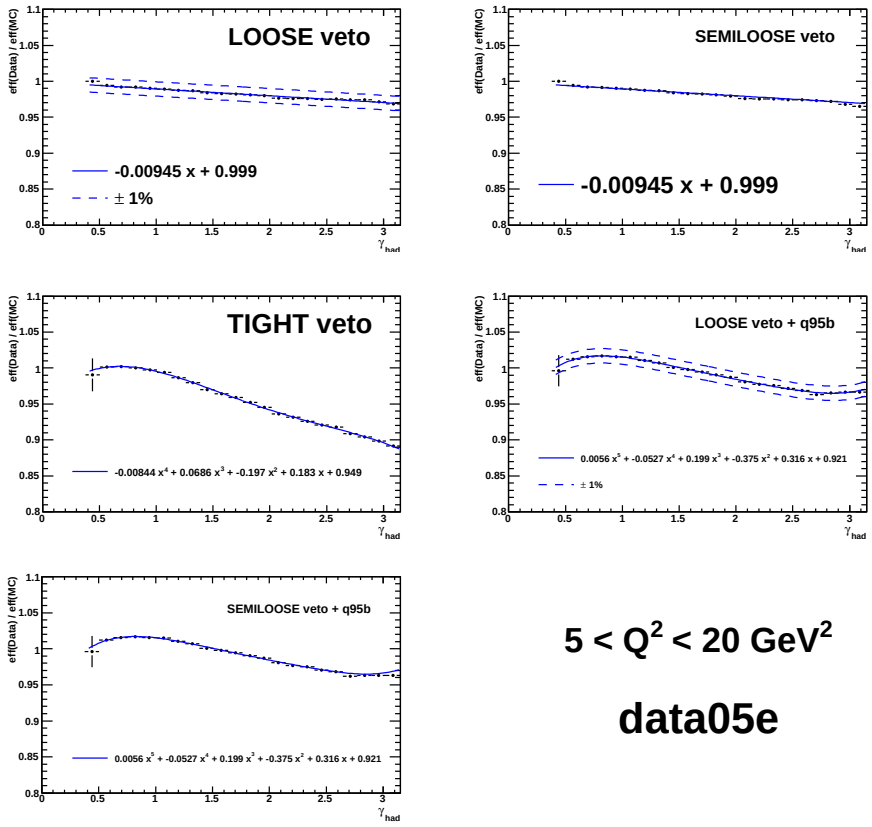


**Figure 6.4:** *The FLT tracking efficiency correction at low $Q^2$ for the 2005e data taking period. The corrections for the LOOSE, LOOSE+q95b and TIGHT+q95b requirements are fitted by polynomial functions (solid blue lines) see Appendix A for details. The functions for SEMILOOSE and SEMILOOSE+q95b requirements were taken from the LOOSE and LOOSE+q95b fits, respectively. For LOOSE and LOOSE+q95b, ±1% variations of the function normalisation are also shown (dashed blue lines).*

for the current analysis) was triggered ⇒ LOOSE+q95 correction had to be applied;

- if an event was not triggered by any of the mentioned FLT bits (i.e. was triggered only by FLT44 with TIGHT+q95b requirements) ⇒ TIGHT+q95 correction had to be applied.

Each step of this procedure has tighter requirements as far as FLT tracking is concerned.

The suggested procedure is general enough that the obtained corrections can be used in any DIS analysis[5]. An additional cross-check of the sensitivity to the event topology might be needed, e.g. a similar study using a LOOSE tracking slot as a reference and comparison of the output with a ratio of $eff(LOOSE + q95)/eff(LOOSE)$ or $eff(TIGHT + q95)/eff(LOOSE)$.

## 6.3 $D^+$ candidate selection

The underlying principle of full particle reconstruction is the same for all analyses. Combinations of tracks with proper charges are made and tracks are assigned the expected masses of the decay products. The invariant mass of the combination is calculated from the energy and momentum conservation rules and if it is found to be close to the expectation, then the combination is accepted as a *candidate*.

The decay channel used in this thesis is: $D^+ \to K^-\pi^+\pi^+$. It has a large branching fraction: $\mathcal{B}(D^+ \to K^-\pi^+\pi^+) = 9.4 \pm 0.4\%$ [10]. Thus combinations are made of three tracks assuming two charged pions of the same charge and one charged kaon of the opposite charge.

### 6.3.1 $D^+$ selection cuts

The following selection has been applied:

- $\boldsymbol{1.5 < p_{\mathrm{T}}(D^+) < 15}$ GeV : the kinematic region in the $D^+$ transverse momentum. At lower values of $p_{\mathrm{T}}(D^+)$ the acceptance drops quickly to zero, whereas background contamination increases. For higher transverse momenta the production cross section is too low to be measured with the available luminosity. Once the charm component of the structure function, $F_2^c$, will be extracted it will be crucial to have as inclusive kinematic region as possible (see Section 8.1);
- $\boldsymbol{-1.6 < \eta(D^+) < 1.6}$ : the kinematic region in the pseudorapidity of $D^+$;
- $\boldsymbol{p_{\mathrm{T}}}(\mathrm{K}) > \boldsymbol{0.5}$ GeV, $\boldsymbol{p_{\mathrm{T}}}(\pi) > \boldsymbol{0.35}$ GeV : the cuts on the kaon and pion track transverse momenta were chosen to keep a reasonably small number of background candidates, while still having sizeable acceptance at

[5]One should mention, that these corrections differ between MC versions, e.g. see Fig. 4.2. These particular corrections have been extracted for the FUNNEL versions num05t2.0, num06t3.0 and num07t3.1

low $p_{\mathrm{T}}(D^+)$. Usage of the mean energy losses in the CTD will not give further improvement in the particle identification in this region;

- $\boldsymbol{-1.75 < \eta(K) < 1.75}$, $\boldsymbol{-1.75 < \eta(\pi) < 1.75}$ **:** the restriction comes from the coverage of the tracking system;

- **"good" tracks :** each track has to start in the MVD and to pass at least 3 CTD superlayers to ensure good momentum resolution. In addition, each track was required to have at least 2 $r\phi$ and 2 $rz$ hits in the MVD to guarantee good spatial resolution close to the primary vertex.

- $D^{*+}$ **"reflection" suppression :** the $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$ decay has an identical final state to the decay under study and could produce an additional peak at 2010 MeV in the invariant mass spectrum. It was suppressed by dropping one of the pions and requiring that the mass difference is not in the range $143 < M(K\pi\pi) - M(K\pi) < 148$ MeV.

- $D_s^+$ **"reflection" suppression :** the $D_s^+ \to \phi\pi^+ \to K^-K^+\pi^+$ decay has a final state similar to the used one and could produce an asymmetric peak in the invariant mass spectrum. It was suppressed by assuming that one of the pions is a kaon and requiring that the invariant mass of the kaon pair lies outside of the range $1.0115 < M(KK) < 1.0275$ GeV.

- $1.7 < M(D^+) < 2.1$ **GeV.**

This selection is similar to the one used in [1].

To further improve the signal-to-background ratio the fact that the $D^+$ meson has a significant lifetime ($c\tau = 311.8 \pm 2.1$ μm [10]) has been exploited. This results in the possibility to reconstruct a decay vertex that is spatially separated from the primary vertex.

### 6.3.2 Selection of secondary vertices

For every $D^+$ candidate the secondary vertex as well as the reduced primary vertex were fitted in 3D. Since the primary vertex position measurement along the $Z$ axis is less precise than in the $XY$ plane, the projection of the decay length on the $XY$ plane was used. Figure 6.5 shows how the decay length information was exploited. The decay length of a secondary vertex is defined as the length of the vector, $\vec{r}_{XY}$, from the reference point of interaction to the secondary vertex. As the reference point a reduced DAF

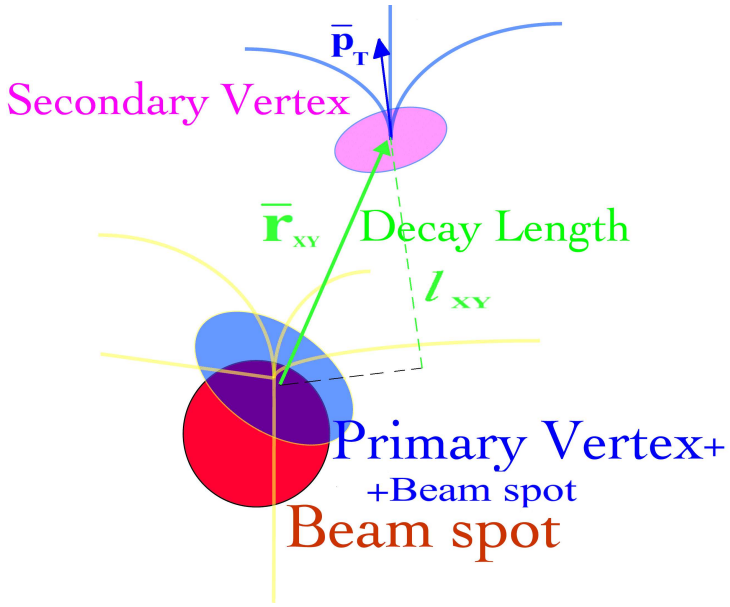


**Figure 6.5:** *Definition of the projected decay length.*

vertex with the beamspot constraint (see Section 5.2) was used in this thesis as it has better spatial resolution than the beamspot itself. To improve the sensitivity to particles with long lifetimes it is important to introduce the projected decay length, which is defined as the projection of the decay length on the particle's flight direction:

$$l_{XY} = \frac{\vec{r}_{XY} \cdot \vec{p}_{\mathrm{T}}}{|\vec{p}_{\mathrm{T}}|} \tag{6.2}$$

In an ideal detector for a real $D^+$ decay vertex this quantity would be positive and equal $|\vec{r}_{XY}| \neq 0$. However, due to the finite detector resolution the projected decay length gets smeared, but the corresponding distribution still has an asymmetric positive tail. On the other hand, candidates constructed from the light quarks either do not have a significant decay length or their lifetime is much larger than those considered here (e.g. $K^0_S$, $\Lambda^0$). Thus after detector effects they produce a distribution with a symmetric core centered at 0 and asymmetric tail at very high values of the decay length.

To achieve a better sensitivity to long-living particles it is possible to combine the information about the projected decay length and its error, $\sigma_l$, to reject fakes with large reconstructed decay length but with large uncertainty. The *signed 2D decay length significance*, $S_l$, is defined as a ratio of

the projected decay length and its uncertainty:

$$S_l = \frac{l_{XY}}{\sigma_l} \tag{6.3}$$

The error on the projected decay length is determined by the projections of the errors of the primary and secondary vertices:

$$\sigma_l = \sqrt{(\sigma_{sec.vtx.}^{projected})^2 + (\sigma_{prim.vtx.}^{projected})^2} \tag{6.4}$$

Each of these errors can be expressed in terms of the covariance matrix elements, $C_{ij}^{vtx}$, for the corresponding vertex:

$$\begin{aligned} \sigma_{vtx}^{projected} &= (\vec{\nabla} l_{XY})\, C\, (\vec{\nabla} l_{XY})^{\mathrm{T}} \\ &= (\frac{\partial l_{XY}}{\partial x})^2 C_{xx}^{vtx} + 2\,\frac{\partial l_{XY}}{\partial x}\,\frac{\partial l_{XY}}{\partial y}\,C_{xy}^{vtx} + (\frac{\partial l_{XY}}{\partial y})^2 C_{yy}^{vtx} \end{aligned} \tag{6.5}$$

Using (6.2) one can get that $\frac{\partial l_{XY}}{\partial x} = \frac{\partial((xp_x + yp_y)/|p_{\mathrm{T}}|)}{\partial x} = cos\phi$ and in analogy $\frac{\partial l_{XY}}{\partial y} = sin\phi$, where $\phi$ is the azimuthal angle of the D meson momentum direction. Thus one can write a simple expression for the error on the projected decay length:

$$\sigma_l = \sqrt{C_{xx}^{prim.+sec.}\, cos^2\phi + 2\, C_{xy}^{prim.+sec.}\, cos\phi\, sin\phi + C_{yy}^{prim.+sec.}\, sin^2\phi} \tag{6.6}$$

where $C_{ij}^{prim.+sec.} = C_{ij}^{prim.vtx.} + C_{ij}^{sec.vtx.}$.

Figure 6.6 shows the comparison of the significance distributions for $D^+$ candidates in the light flavour, charm and beauty MC samples. Both the charm and beauty distributions feature clear asymmetric tails to positive $S_l$ values, whereas the distribution for the light flavour background is roughly symmetric.

The optimal value for a cut on the $S_l$ value for a candidate was found by maximising the statistical significance of the mass peak in MC ($N^{signal}/\sigma(N^{signal})$). The signal was extracted by a fit to the $D^+$ mass spectra (for details see the following Section 6.4) gradually increasing the the decay length significance cut. The uncertainty, $\sigma(N^{signal})$, was estimated by $\sqrt{N^{signal} + N^{background}}$, where the number of background candidates under the peak, $N^{background}$, was estimated also from the fit. The optimal selection of the secondary vertices was found to be almost independent of the kinematic phase space:

- $\boldsymbol{S_l > 4}$;

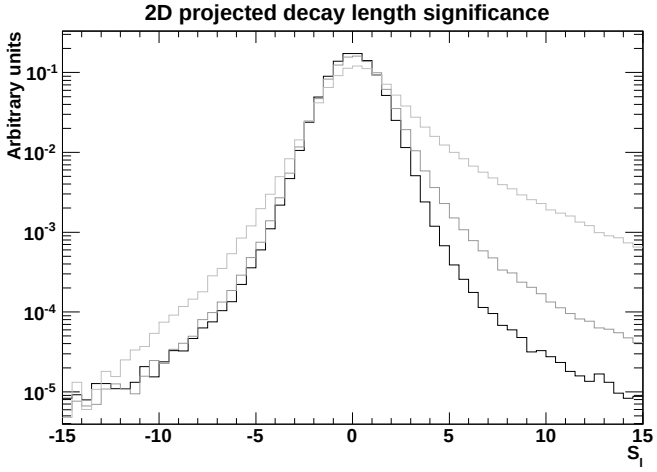


**Figure 6.6:** *Comparison of shapes of the signed decay length distributions for $D^+$ candidates for different flavour components. The distributions from events containing light flavour (black line), charm (grey line) and beauty (light grey line) quark production are shown. The distributions were normalised to unity, hence the Y axis is shown in arbitrary units.*

- $\boldsymbol{\chi^2_{sec.vtx.} < 10}$ : a cut on the $\chi^2$ of the secondary vertex fit[6] was imposed to ensure a good secondary vertex.

The reconstruction of secondary vertices from $D^+ \rightarrow K^-\pi^+\pi^+$ decays provides a significant improvement in the statistical precision of the signal by reducing the combinatorial background. The effectiveness of the significance cut is illustrated in Fig. 6.7, which shows the invariant mass spectra before and after the optimal cut on $S_l$. After the cut the combinatorial background is considerably reduced. In this way the statistical uncertainty is reduced from 6.3% to 2.3%.

## 6.4 $D^+$ signal extraction

A fitting procedure was used to extract the $D^+$ signal from the mass distribution of the candidates after selection. The function used for the extraction was a Gaussian function for the signal:

$$g(x; N, M, \sigma) = \frac{N}{\sqrt{2\pi}\sigma} exp(-\frac{(x-M)^2}{2\sigma^2}) \tag{6.7}$$

[6]The number of degrees of freedom for a vertex with 3 tracks is $2 \times 3 - 3 = 3$

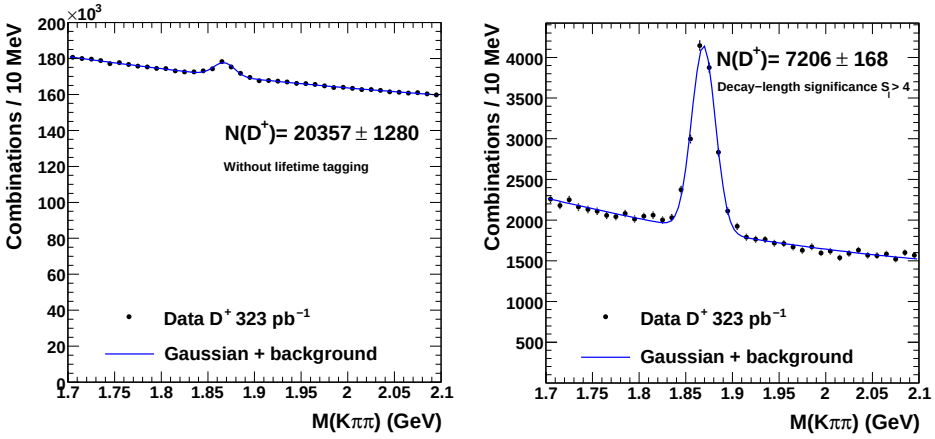


**Figure 6.7:** *Mass spectra for the reconstructed $D^+$ candidates in the data before (left) and after (right) application of the $S_l > 4$ cut.*

with a 2$^{\text{nd}}$ order polynomial to parametrise the combinatorial background:

$$b(x; p_0, p_1, p_2) = p_2 x^2 + p_1 x + p_0. \tag{6.8}$$

An example of the obtained fit function is shown in Figure 6.7. The fit was done using the least-$\chi^2$ method as implemented in the MINUIT package [109]. The integrals of the fitting function within each mass bin were used in the $\chi^2$ function. To account for possible non-linearities, the statistical uncertainty was estimated as the average of the positive and negative statistical uncertainties obtained with the MINOS function[109].

## 6.5 Signal comparison with Monte Carlo

Figure 6.8 shows the comparison of the signal distributions for the event and $D^+$ kinematics. The number of reconstructed $D^+$ mesons was extracted from the fit, as explained in Section 6.4, in each bin of the variable on the $X$ axis. The beauty normalisation was scaled by 1.5, as obtained in the previous beauty measurement in a similar kinematic region [35]. The data are reasonably well described by the Monte Carlo simulations in $y_{da}$ and $E'_e$. Nevertheless the data are poorly described in other important parameters: transverse momentum and pseudorapidity of $D^+$ as well as $Q^2_{da}$. Both momentum and pseudorapidity measurements in ZEUS are very precise, thus discrepancies in $p_{\text{T}}(D^+)$ and $\eta(D^+)$ do not originate from the detector simulation. The inclusive control distribution for $Q^2_{da}$ (Fig. 6.1) have shown reasonable agreement.

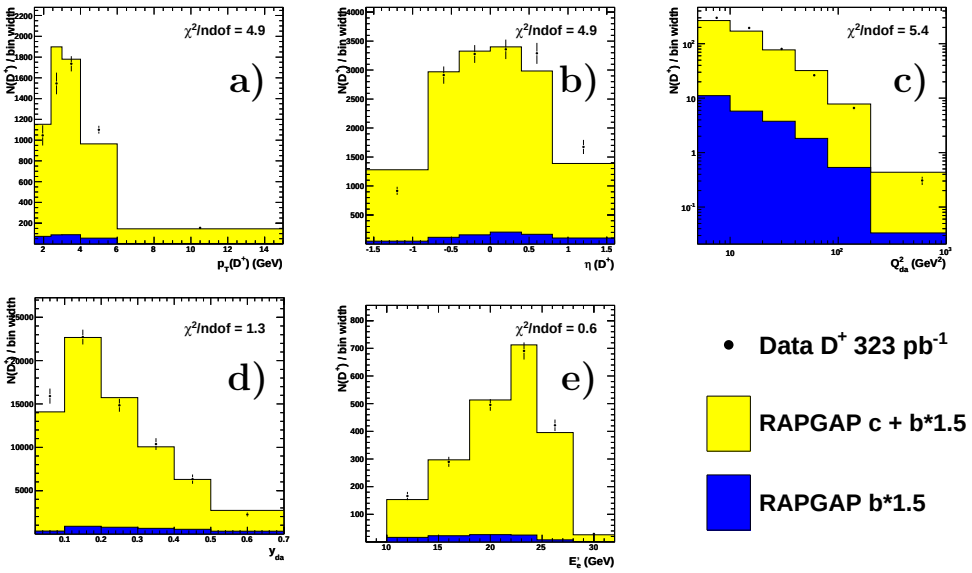


**Figure 6.8:** *Signal distributions for event and $D^+$ kinematics. Shown is the number of reconstructed $D^+$ mesons as a function of: a) $p_{\mathrm{T}}(D^+)$, b) $\eta(D^+)$, c) $Q^2_{da}$, d) $y_{da}$, e) energy of the scattered electron. The data (black points) are compared to the simulated charm+beauty MC distributions (light yellow filled area). The beauty contribution (dark blue filled area) is shown separately. The sum of charm+beauty was normalised to the data area.*

Therefore these discrepancies originate from imperfect charm cross section shapes in the Monte Carlo generator. Thus the shape of these charm production cross sections in the Monte Carlo simulations were tuned to describe the control distributions in the data (see following Section 6.6).

Figure 6.9 shows the control plots for parameters of $D^+$ vertices for the signal. These plots were done without selection of the secondary vertices and thus the minimum $p_{\mathrm{T}}(D^+)$ cut had to be raised to get a reasonable signal to background ratio. The data are well described by the Monte Carlo simulations. This will result in much smaller systematics from this source compared to the previous measurements [1].

## 6.6 Monte Carlo weighting

Figure 6.8 shows that the nominal Monte Carlo simulations do not reproduce the signal distributions in the data for some of the variables. If not

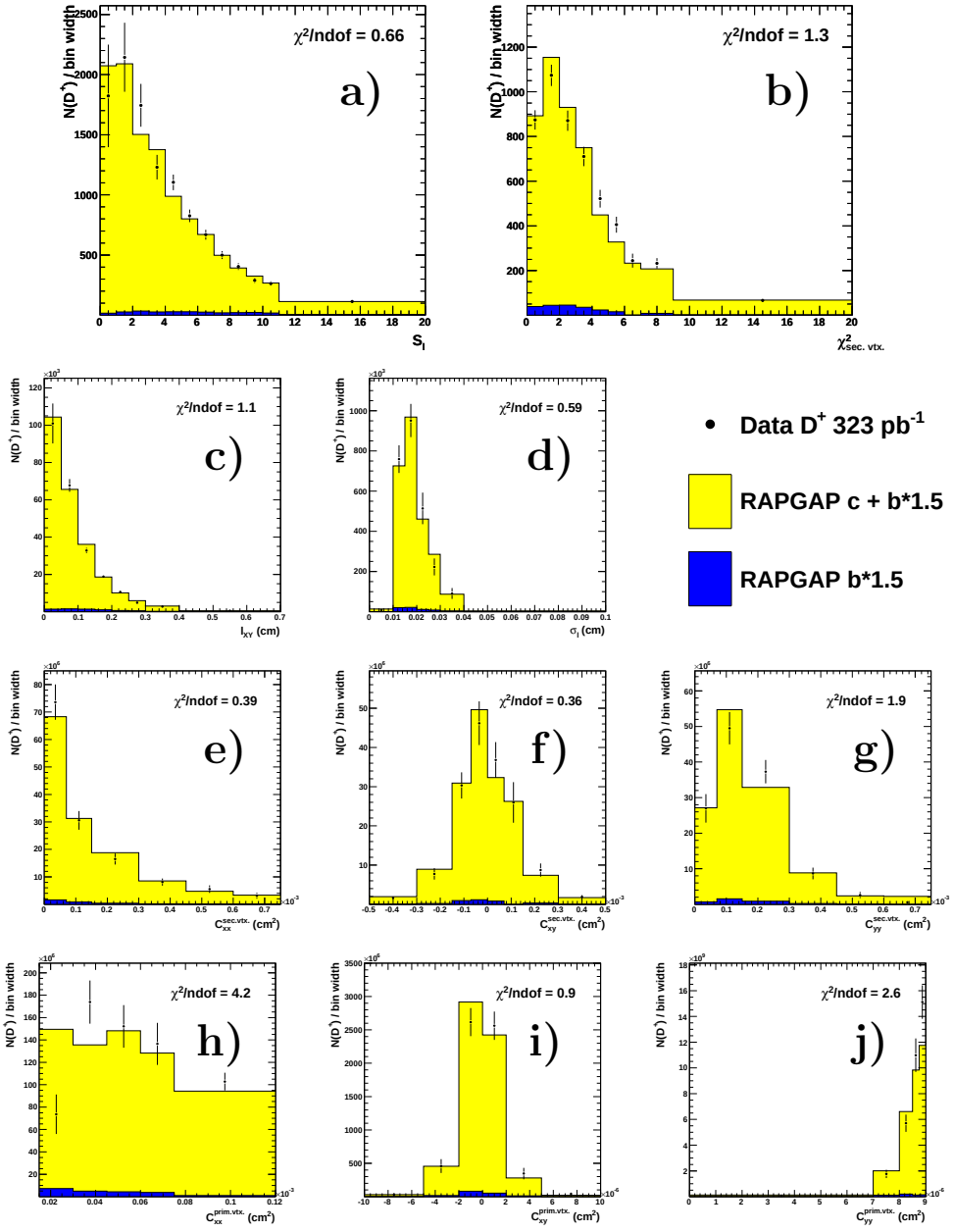


**Figure 6.9:** *Signal distributions for vertex-related parameters for $p_{\mathrm{T}}(D^+) > 3\,GeV$. Shown is the number of reconstructed $D^+$ mesons as a function of: a) $S_l$ , b) $\chi^2_{sec.vtx.}$, c) $l_{XY}$, d) $\sigma_l$, e–g) covariance matrix elements for the secondary vertex fit, $C_{ij}^{sec.vtx.}$, h–j) covariance matrix elements for the reduced primary vertex fit with a beamspot constraint, $C_{ij}^{prim.vtx.}$. In the region $S_l < 0$ no stable fit for the signal was found. Other details are the same as in Fig. 6.8.*

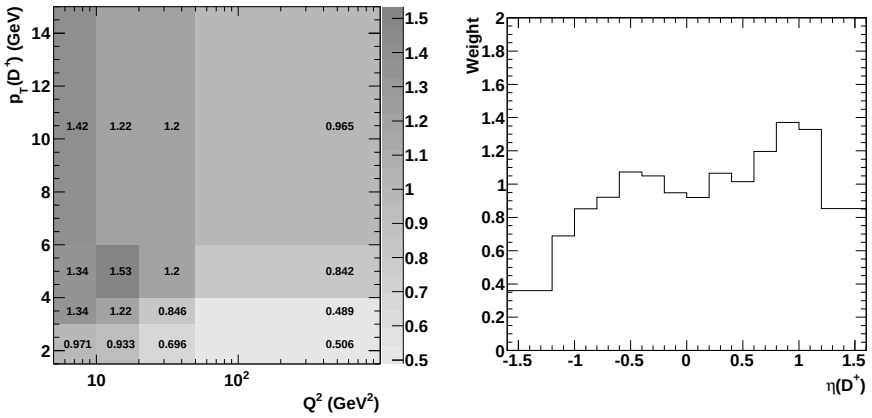


**Figure 6.10:** *Weights to be applied to the Monte Carlo simulations to describe the corresponding control plots in the data. Left: weights extracted double differentially in $p_{\mathrm{T}}(D^+)$ and $Q^2_{da}$, right: weights in $\eta(D^+)$.*

corrected, this would result in a wrong acceptance determination for the cross section calculation (see Section 7.1.1). Therefore the shape of the differential charm cross sections in the Monte Carlo simulations had to be tuned to describe the control distributions in the data.

The transverse momentum of a $D^+$ meson that was produced in the BGF process is correlated with the virtuality of the exchanged boson, whereas both are almost independent of the meson's pseudorapidity. Hence the tuning can be applied independently in $\eta(D^+)$ and $p_{\mathrm{T}}(D^+)$ or $Q^2$. Tuning of the shape in $p_{\mathrm{T}}(D^+)$ (a harder distribution is observed in the data) was found to lead to even worse description in $Q^2_{da}$. Therefore a double-differential correction in $p_{\mathrm{T}}(D^+)$ and $Q^2$ was used. The tuning was performed by applying weights to the $D^+$ particles on the hadron level of the charm Monte Carlo simulations, such that the control plots in the data were reproduced. This procedure is also often referred to as *weighting* or *reweighting*. The weights are extracted from the ratio of the data to MC in the corresponding control distribution. In the analysis presented in this thesis a binned weighting was used, hence the weighting function was not smooth, but rather histogram-like. The binning should be chosen to keep the purity, i.e. the fraction of events reconstructed in a bin which were generated in that bin, high (e.g. above 0.75 for most of the bins in $p_{\mathrm{T}}(D^+)$ and $Q^2$). For bins where migrations outside of a bin are still significant, the weighting would not give a perfect agreement between data and MC in one step. Therefore an iterative procedure should be applied in the case that agreement is not satisfactory.

Figure 6.10 shows the weighting functions that were extracted and applied hereafter in this analysis[7]. The weighting function in $\eta(D^+)$ changes fast at the borders of the kinematic region of the measurement, hence the smallest possible binning is important. For the weighting in $p_{\mathrm{T}}(D^+)$ and $Q^2$ two iterations were needed, since after the first iteration $\chi^2/ndof$[8] was slightly below 1. After this weighting perfect agreement in the corresponding control distributions was observed (for both $\chi^2/ndof < 0.1$).

Control distributions for the signal kinematics after Monte Carlo weighting are shown in Fig. 6.11. After weighting the data are very well described in the tuned parameters ($p_{\mathrm{T}}(D^+)$, $\eta(D^+)$, $Q^2_{da}$). The data are also well described in other kinematic variables.

This tuning will be applied in the following by default and a variation of the shape of the weighting function will be considered in the systematic studies.

[7] Work done together with Oleksandr Zenaev.

[8] $ndof$ = number of degrees of freedom.

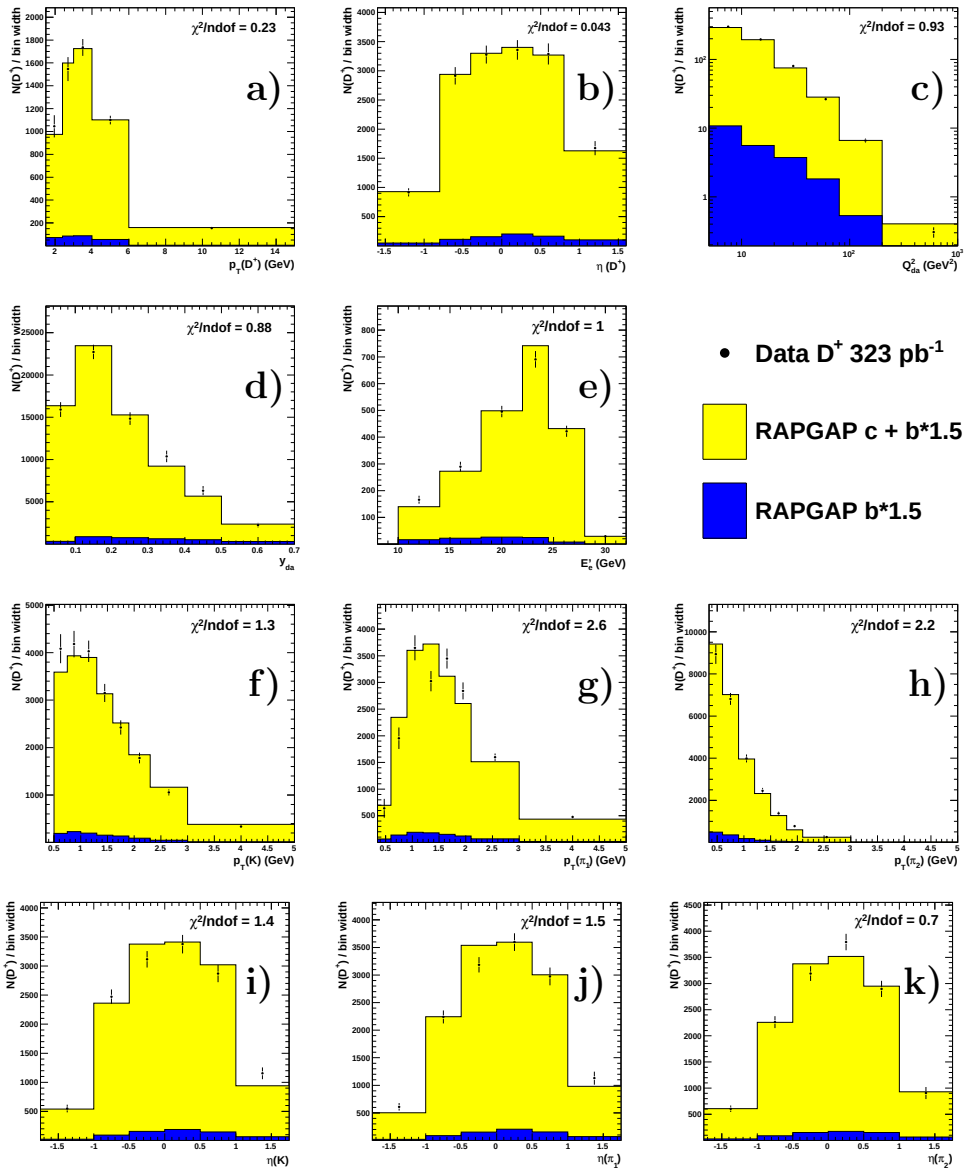


**Figure 6.11:** *Signal distributions for event and $D^+$ kinematics after the weighting procedure. Shown is the number of reconstructed $D^+$ mesons as a function of: a) $p_{\mathrm{T}}(D^+)$, b) $\eta(D^+)$, c) $Q^2_{da}$, d) $y_{da}$, e) energy of the scattered electron, f) $p_{\mathrm{T}}$ of the kaon, g–h) $p_{\mathrm{T}}$ of the pions ordered in transverse momentum, i) $\eta$ of the kaon, j–k) $\eta$ of the pions ordered in transverse momentum. Details are the same as in Fig. 6.8.*

# Chapter 7

# $D^+$ production cross section

The measurement of $D^+$ production cross sections in the kinematic region $5 < Q^2 < 1000\,\mathrm{GeV}^2$, $0.02 < y < 0.7$, $1.5 < p_{\mathrm{T}}(D^+) < 15\,\mathrm{GeV}$ and $|\eta(D^+| < 1.6$ is presented in this Chapter. First, the cross section extraction procedure is reviewed and the definition of the measured cross sections is given. This is followed by the systematic studies that were performed. Finally, the measured cross sections are presented and compared with the previous measurements in the same kinematic region as well as with NLO QCD predictions.

## 7.1 Cross section determination

Cross sections are used to express the probabilities of certain type of interactions to occur. The integrated cross section for the process that is studied can be extracted from the total number of *signal* events, $N$, and the integrated luminosity, $\mathcal{L}$, that is measured with a reference process (e.g. see Section 3.7): $\sigma = N/\mathcal{L}$. In order to obtain the total number of signal charm events the number of reconstructed events in the data, $N^{reco}_{data}$, has to be corrected to account for the detector and selection effects as well as for an admixture of other processes with the same signature:

$$\sigma = \frac{N^{reco}_{data} - N^{reco}_{MC,b}}{\mathcal{A}_c\,\mathcal{L}\,\mathcal{B}} \cdot \mathcal{C}^{\mathrm{rad}}, \tag{7.1}$$

where $\mathcal{B}$ is the branching ratio for the decay (i.e. according to [10] $\mathcal{B}(D^+ \to K^-\pi^+\pi^+) = 9.4 \pm 0.4\%$) whereas $\mathcal{A}_c$ defines the acceptance for $D^+ \to K^-\pi^+\pi^+$ in charm events in the kinematic region of the measurement, which will be discussed in Section 7.1.1.

The data contain a small fraction of events with $D^+$ produced from $B$ and $\Lambda_b$ hadron decays, $N^{reco}_{MC,b}$. These have to be subtracted from the data to obtain the $D^+$ production cross section that is directly comparable to the NLO QCD predictions for $D^+$ from charm production by the HVQDIS program. The subtraction was done using beauty production events simulated by RAPGAP. The simulated yields from the beauty MC sample was scaled by a factor 1.5 as measured in [35] and a variation of this scaling factor will be included in the systematic uncertainty.

In order to obtain cross sections that can be compared with the calculations by HVQDIS the data also had to be corrected to the Born level in QED, i.e. without real QED emissions, but with the running electromagnetic coupling constant, $\alpha$. The radiative correction factor, $\mathcal{C}^{\rm rad}$, will be discussed in detail in Section 7.1.2.

In analogy to the integrated cross section, the single differential cross section as a function of a given variable $Y$ is defined as:

$$\frac{d\sigma}{dY} = \frac{N^{reco}_{data} - N^{reco}_{MC,b}}{\mathcal{A}_c \, \mathcal{L} \, \mathcal{B} \Delta Y} \cdot \mathcal{C}^{\rm rad}, \tag{7.2}$$

where $\Delta Y$ corresponds to the width of the given bin in $Y$.

### 7.1.1 Acceptance correction

The acceptance correction for detector and reconstruction effects was extracted from the RAPGAP Monte Carlo simulations as:

$$\mathcal{A}_c = \frac{N^{reco}_{MC,c}}{N^{true}_{MC,c}}. \tag{7.3}$$

Here $N^{reco}_{MC,c}$ defines the number of reconstructed $D^+$ mesons in the Monte Carlo sample with the charm pair production in the BGF process. It has to be obtained using a selection identical to the one in the data that was detailed in Chapter 6. $N^{true}_{MC,c}$ defines the number of generated $D^+$ mesons in the analysis phase space. The good description of the data by the MC simulations after weighting (see Figs. 6.9 and 6.11) gives confidence in the use of the RAPGAP MC to correct the data.

The averaged acceptance for this measurement was found to be $\approx 5\%$. Figures 7.1 and 7.2 show the extracted acceptances in bins of $p_T(D^+)$, $\eta(D^+)$, $Q^2$ and $y$ and double differentially in $Q^2$ and $y$ [1]. Typical values are below

[1] On the true level the DIS kinematic variables, $Q^2$ and $y$, are defined at the photon vertex.

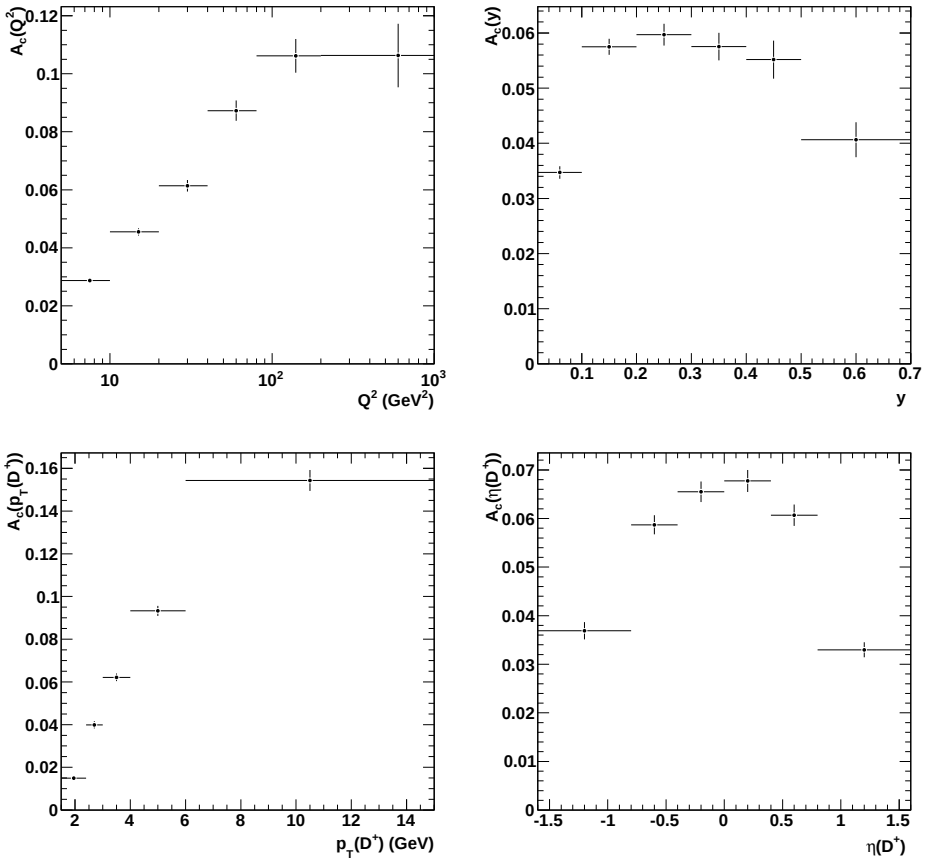


**Figure 7.1:** *Extracted acceptances for the single differential cross sections.*

those in analyses of $D^*$ production (e.g. see Sect. 8.5.1 in [110]) and the production of other D mesons without lifetime tagging (e.g. see Sect. 5.1.5 in [111]). This is due to the strong decay length significance cut imposed in the analysis presented in this thesis, which reduces the signal by a factor $\approx 3$.

### 7.1.2 Radiative corrections

The Monte Carlo sample used for the acceptance correction was generated including first-order electromagnetic effects, i.e. real and virtual emissions from both the initial and final state lepton. These sittings allow to describe the QED effect present in data. However, HVQDIS does not incorporate QED emissions and includes only the running of $\alpha$. The correction for the

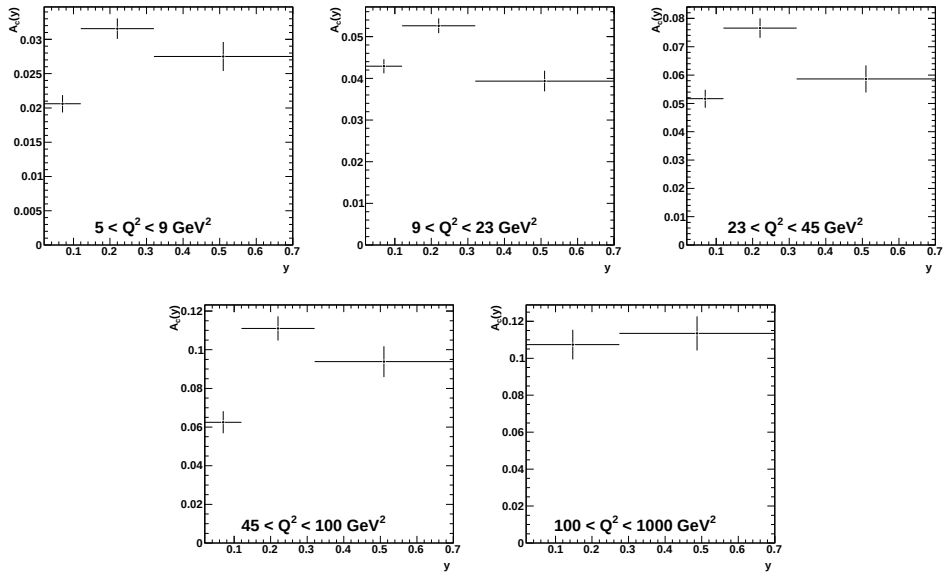


**Figure 7.2:** *Extracted acceptances for the double differential cross sections in $Q^2$ and $y$.*

data to the QED Born level was carried out using the RAPGAP MC:

$$\mathcal{C}^{\mathrm{rad}} = \sigma^{\mathrm{Born}}/\sigma, \tag{7.4}$$

where $\sigma$ is the cross section with full QED radiation settings and $\sigma^{\mathrm{Born}}$ corresponds to a MC sample with the QED corrections switched off. The corrections for the differential cross sections are presented in Figs. 7.3 and 7.4. The corrections are typically $\sim 2\%$ and go up to 10% at high $Q^2$ and hence at high $p_{\mathrm{T}}(D^+)$.

## 7.2 Systematic uncertainties

Various sources of systematic uncertainties have been accounted for by repeating the analysis with modified selection criteria or a varied simulation model. The resulting systematic uncertainties were evaluated as the difference between the output after the variation and the nominal cross section, $\Delta\sigma^{syst} = \sigma^{syst} - \sigma$. Systematic uncertainties from different sources were added in quadrature separately for positive and negative deviations from the nominal cross section[2]. The list of the systematic checks and their impact

[2]In the case of both deviations going in the same direction only the larger one was taken into account.

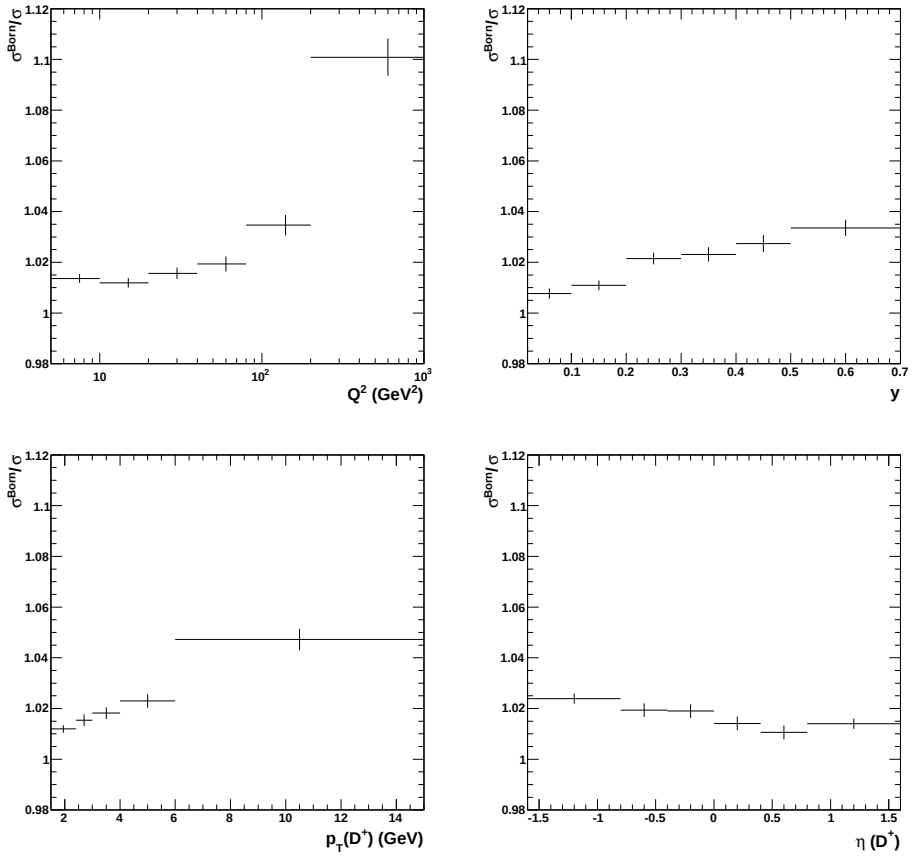


**Figure 7.3:** *The radiative corrections to the QED Born level for the single differential cross sections.*

on the measurement is given below. A breakdown of the systematic uncertainties for the differential cross sections is presented in Appendix B.

## Position of the scattered electron in the RCAL

The uncertainty from the Monte Carlo simulation of the scattered lepton position on the RCAL surface was evaluated by varying the size of the box cut around the beampipe by $\pm 1$ cm simultaneously in the $x$ and $y$ directions in both the data and the MC simulations. The averaged effect is $\pm 1\%$ and goes up to $+5\%$ at low $Q^2$. Previous studies [112] have shown, that after a dedicated alignment it is possible to reduce the variation range to 0.4 cm. However, since the effect on this measurement was found to be small, the procedure was not adopted here.

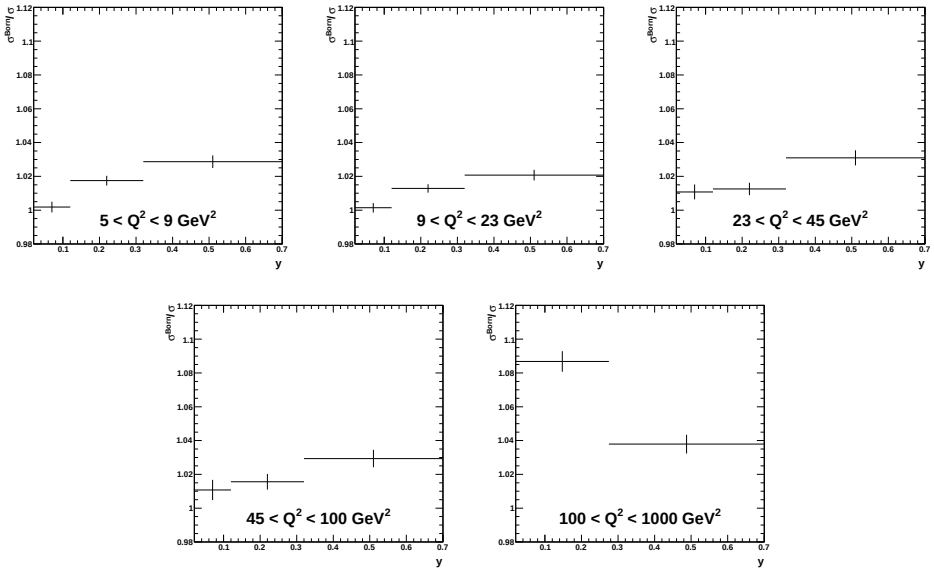


**Figure 7.4:** *The radiative corrections to the QED Born level for the double differential cross sections in $Q^2$ and $y$.*

## Energy scale in the EMC

To account for the EMC response simulation the electron energy scale was varied in the MC simulations within the known scale uncertainty of the EMC, which is $\pm 2\%$ [113]. The averaged effect is $< 1\%$ and goes up to $-4\%$ at high $y$.

## Energy scale in the HAC

To account for the HAC response simulation the ZUFO energy scale was varied within $\pm 3\%$, which is the scale uncertainty of the HAC [113]. The variation was applied to the HAC energy scale in the Monte Carlo only. This leads to a conservative estimate of the hadron energy scale uncertainty, since energy resolution for the tracking–based ZUFOs is much better at low and medium momenta (see transverse momentum resolution on Fig. 3.7) and the track energy scale is very precisely known. However, the resulting systematic uncertainty is small: the averaged effect was $< 1\%$.

## FLT tracking efficiency

The uncertainty from the simulation of the FLT tracking efficiency was evaluated by varying the correction functions that were extracted in Section 6.2.6 by ±50%. Figure 6.4 has shown that the extracted corrections lie within a variation of the fitted function by 1%, but, as it was said, the dependence of the correction on the event topology has not yet been studied. That is why a conservative variation by 50% was used to account for possible systematics. The averaged effect is below 1% and it is dependent on $Q^2$ due to the kinematic coverage of the FLT30 slot (see Appendix B).

## Signal extraction procedure

The number of the reconstructed $D^+$ mesons has been extracted from a fit. To account for possible systematics the fit procedure was changed. Two modifications were studied. To account for the background shape description the background parametrisation was changed from a second order polynomial to an exponential function: $exp(p_2x^2 + p_1x + p_0)$. The averaged effect is below 1%. To account for the signal shape description the signal parametrisation was switched from a Gaussian function to a so-called modified Gaussian function: $\frac{N}{\sqrt{2\pi}\sigma} exp(-\frac{1}{2}X^{1+\frac{1}{1+0.5X}})$, where $X = \frac{x-M}{\sigma}$ and the meaning of the parameters $N$, $M$, $\sigma$ is similar to a standard Gaussian function. This function has longer tails, compared to the nominal Gaussian function. The effect is +1% on average. All fit functions provide reasonable description of the mass spectra ($\chi^2/ndof \approx 1$).

## Reconstruction of the decay length

The inclusive secondary vertex analysis [73] has shown an indication that the tails of the inclusive decay length distribution are not simulated well. To correct for this feature a procedure of the decay length "smearing" in the Monte Carlo was suggested. Parameters of the smearing had to be tuned to describe the data. A similar procedure was exploited here to estimate the systematic effect. The parameters had to be re-tuned on the negative half of the signed 2D decay length distribution, which originates from the decay length resolution, for the $D^+$ candidates that pass the default selection without requirements on the secondary vertex, which were described in Section 6.3.2. Figure 7.5 shows the effect of the smearing procedure on the decay length distribution. After application of the re-tuned smearing function the

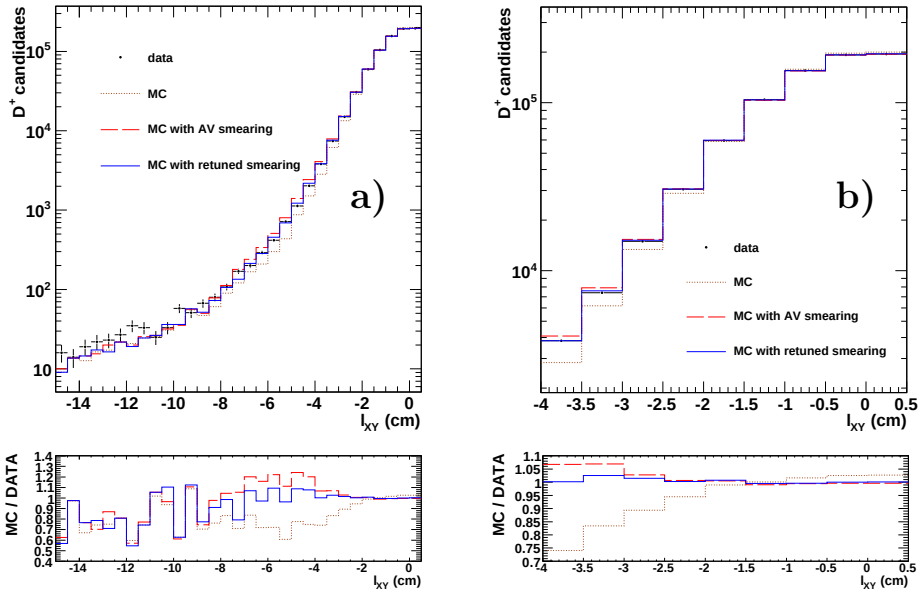


**Figure 7.5:** *Distributions for the negative side of the signed 2D decay length for the $D^+$ candidates a) in the full range, b) in the core region. The data (black points) are compared to the nominal Monte Carlo simulations (dotted orange line), the Monte Carlo simulations with the smeared resolution as in [73] (dashed red line) and Monte Carlo with the re-tuned smearing function for this analysis (solid blue line). In general, both smearing functions perform similar, although the re-tuned smearing is a bit better in the medium-$l_{XY}$ region. The nominal Monte Carlo does not provide a good description.*

negative half of the signed 2D decay length distribution in the Monte Carlo simulations reproduces the corresponding distribution in the data well. The effect of the smearing application is moderate – typically below 3%, that is why the smearing was not yet applied for the central values of cross sections. The full effect of smearing was included as a systematic uncertainty.

The size of this systematic uncertainty is smaller than that in the previous ZEUS analysis [1]. This is due to better MVD alignment and calibrations. The reasonably good description of the decay length resolution before smearing Fig. 7.5 as well as the good description of the signal decay length and significance distributions (see Fig. 6.9) confirm that this source of systematics is under control. An additional cross check was done varying the decay length significance cut in a wide range. The behaviour of the data and MC agrees within the systematic uncertainty quoted above.

### $b$ quark contribution

The normalisation factor for the Monte Carlo predictions of beauty production was taken from [35]. However, in that analysis a dependence of the normalisation factor (referred to as $k_b$ in the reference) on the region of the phase space was observed. Therefore, the beauty contribution normalisation factor was varied from 1.5 by $\pm 0.5$. The effect on the cross sections is $\pm 3\%$ on average and increases at high $Q^2$ and high $y$.

### Monte Carlo model

The uncertainties from the Monte Carlo description of the kinematic variables were evaluated by changing the shapes of the weighting functions. This is an important source, since it changes the acceptance.

For $\eta(D^+)$ the reweighting function obtained in Section 6.6 was varied by $\pm 50\%$. The averaged effect is below 3%. The size of the systematic uncertainty strongly depends on $\eta(D^+)$ and $y$, whereas it is rather constant is $Q^2$ and $p_{\mathrm{T}}(D^+)$.

The variation of the $Q^2$ and $p_{\mathrm{T}}(D^+)$ shapes was done simultaneously by altering the double-differential weighting function obtained in Section 6.6. The variation range was chosen to keep a reasonably good description of the double-differential control distribution in $Q^2$ and $p_{\mathrm{T}}(D^+)$, i.e. while $\chi^2/ndof \leq 1$. It turned out to correspond roughly to a 50% variation. Figure 7.6 shows the control distributions for the main kinematic variables after application of the two weighting functions that were used to estimate the systematic uncertainty. The estimated systematic effect on the cross sections is below 5% and strongly depends on $Q^2$ and $p_{\mathrm{T}}(D^+)$.

### Tracking efficiency

To estimate the effect of imperfections in the Monte Carlo simulation of the tracking efficiency in the data, e.g. due to imperfect modelling of the hadronic secondary interaction probability in the material of the tracking system or due to pattern recognition problems, the single-track efficiency in the Monte Carlo was degraded by 2%. The value of 2% inefficiency was estimated for the possible difference. The overall systematic uncertainty from this source was calculated as a linear sum of the single–track inefficiencies for three tracks, resulting in +6% uncertainty that is fully correlated between all bins of all cross sections. When dedicated studies will come to a conclusion,

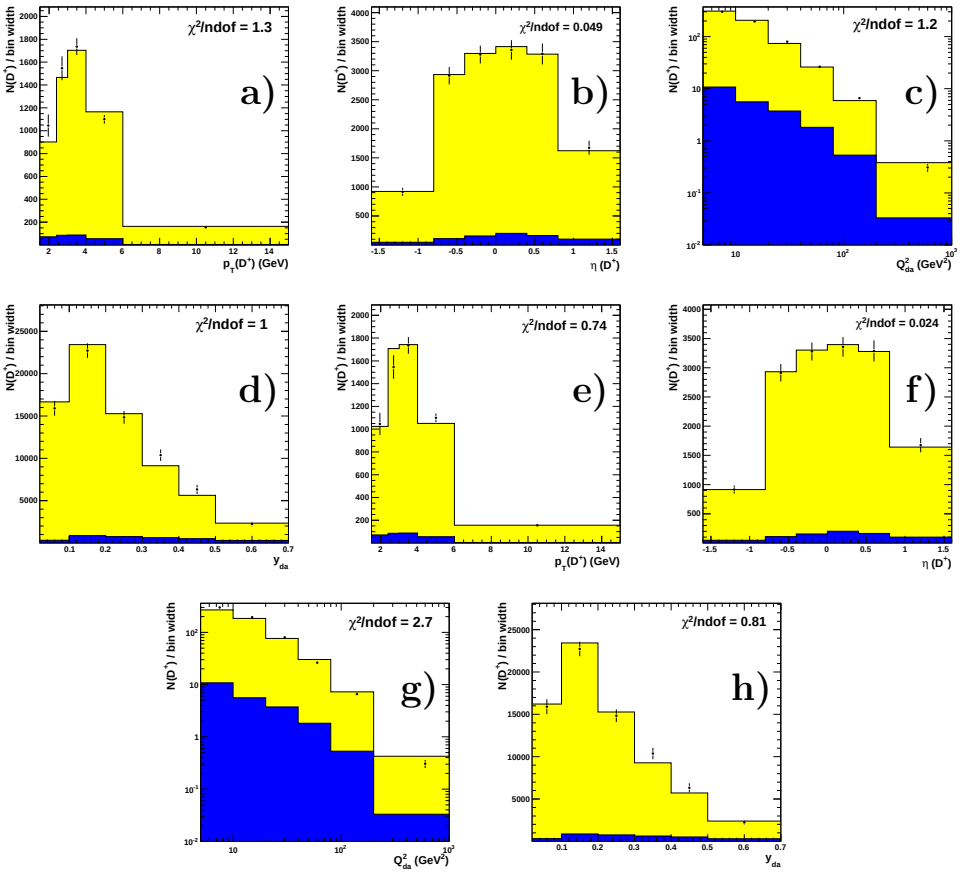


**Figure 7.6:** *Control distributions for two $Q^2$ and $p_{\mathrm{T}}(D^+)$ weighting variants that correspond to $\chi^2/ndof \approx 1$ ( a–d) and e–h) ). Shown are numbers of reconstructed $D^+$ mesons as a function of: a,e) $p_{\mathrm{T}}(D^+)$, b,f) $\eta(D^+)$, c,g) $Q^2_{da}$, d,h) $y_{da}$. Other details are the same as in Fig. 6.8.*

it will be possible to replace this uncertainty by a correction, which will be constant with respect to the kinematic variables in the phase space of this measurement (e.g. see studies on track $p_{\mathrm{T}}$ dependence in [114]), with smaller systematic uncertainty.

## MVD hit efficiency

The uncertainty from the simulation of the MVD hit efficiency was estimated as the difference between data and MC efficiencies of reconstructing tracks with at least 2 $r\phi$ and 2 $rz$ hits. The uncertainty for a single track was

found to be ±0.3%. This number has to be added linearly for the number of tracks used in the analysis. Thus the resulting systematic uncertainty is ±0.9%.

### Luminosity measurement

The uncertainty on the luminosity measurement was constant in time and was equal ±1.8%. This uncertainty is fully correlated between cross section bins and between various measurements in ZEUS. Following the general convention in ZEUS, it was not added to the differential cross sections.

### Measurement of $\mathcal{B}(D^+ \to K^-\pi^+\pi^+)$

According to [10], the uncertainty due to the branching ratio for the chosen decay is ±4.3%. This normalisation uncertainty was also not added to the differential cross sections, following the ZEUS conventions.

## 7.3 Charm production cross sections

The charm production cross sections for the process $ep \to e'c\bar{c}X \to e'D^+X$ were measured in the kinematic region:

$$5 < Q^2 < 1000\,\mathrm{GeV}^2,$$

$$0.02 < y < 0.7,$$

$$1.5 < p_{\mathrm{T}}(D^+) < 15\,\mathrm{GeV},$$

$$-1.6 < \eta(D^+) < 1.6$$

using $323\,\mathrm{pb}^{-1}$ of the HERA data collected at $\sqrt{s} = 318\,\mathrm{GeV}$.

The single–differential cross sections were measured as a function of the $D^+$ meson transverse momentum, $p_{\mathrm{T}}(D^+)$, and pseudorapidity, $\eta(D^+)$, as well as of the exchange boson virtuality, $Q^2$, and inelasticity, $y$. The double–differential cross sections were measured as a function of $Q^2$ and $y$. The numbers of the signal $D^+$ candidates in data and Monte Carlo that are needed to calculate the cross section using (7.2) were extracted applying the fit procedure explained in Section 6.4 in each bin. The resulting fit distributions are presented in Appendix C.

The measured single–differential cross sections are shown in Figs. 7.7 and 7.8 and the values are tabulated in Table 7.1; the inner error bars represent the statistical uncertainty whereas the outer error bars correspond to the statistical and systematic uncertainties added in quadrature. Statistical uncertainties from both the data and the Monte Carlo samples were propagated to the statistical uncertainty of the measurement. The systematic sources were discussed in detail in Section 7.2. The systematic uncertainties due to the luminosity and branching ratio measurements were not included. The dominant systematic uncertainties are due to the Monte Carlo model, i.e. variations of $p_{\mathrm{T}}(D^+)$, $\eta(D^+)$ and $Q^2$ shapes, as well as the tracking efficiency simulation. The final systematic uncertainties are asymmetric, since the uncertainty due to the tracking efficiency is one–sided.

The cross sections are compared to the previously published measurement based on the 2005 data sample [1] in the same phase space. The previous analysis includes a small fraction of $D^+$ mesons originating from b quark production and the previous measurement had to be slightly corrected to the binning of the new measurement using Monte Carlo simulations. The results obtained in this thesis agree with the previous measurement and have significantly reduced statistical and systematic uncertainties. The improvement in the statistical uncertainty is due to the larger sample that was used in this analysis as well as due to better background suppression by the decay length significance cut owing to better MVD alignment and calibration. Improvement in the MVD alignment and calibrations also lead to a reduction of the systematic uncertainty. Thus the measurement presented in this thesis is the most precise measurement of charm production with $D^+$ mesons. The measurements were also compared to the NLO QCD predictions that were detailed in Chapter 2. The measured cross sections are well described by the NLO QCD calculations.

In addition to the cross sections discussed above, Fig. 7.9 shows the measured double–differential cross sections as a function of $y$ in five bins of $Q^2$. The values are summarised in Table 7.2. The data are well described by the NLO QCD predictions. Due to a larger sample it is possible to measure cross sections more differentially and thus a comparison to the previously published ZEUS result is possible only for $5 < Q^2 < 9\,\mathrm{GeV}^2$. The same conclusion as for the single-differential cross sections can be drawn – the measurements are in agreement and the new analysis has improved precision. These cross sections were used to extract the charm contribution to the structure function $F_2$, $F_2^c$, as described in the following Chapter 8.

Preliminary results of this analysis have been released by the ZEUS collaboration for conferences [115]. Since then a better procedure of the MC

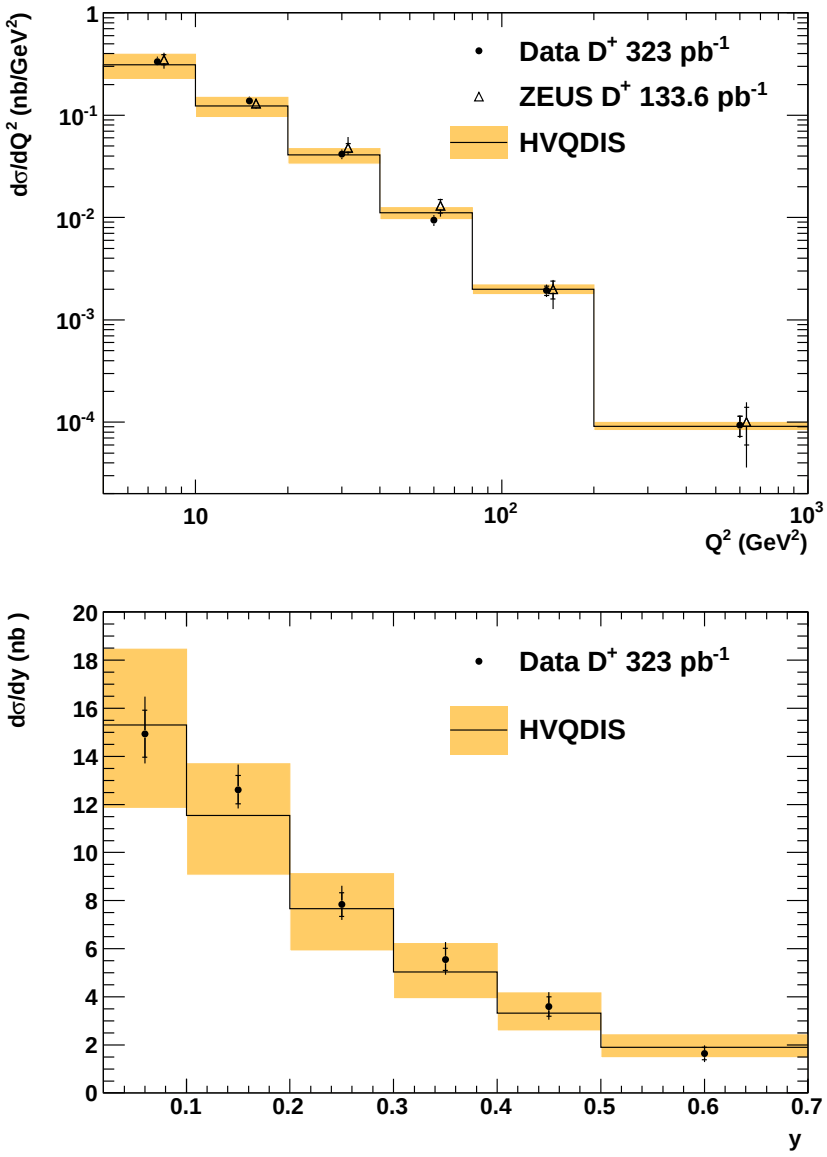


**Figure 7.7:** *Differential cross sections for $D^+$ meson production as a function of $Q^2$ (top) and $y$ (bottom). The cross sections are given in the kinematic region $5 < Q^2 < 1000\ GeV^2$, $0.02 < y < 0.7$, $1.5 < p_T(D^+) < 15\ GeV$, $|\eta(D^+| < 1.6$. The measurements obtained in this analysis (filled points) are compared to the previous ZEUS measurement (open triangles), where possible. The earlier measurements are slightly shifted along the $X$ axis to improve the visibility. The inner error bars correspond to the statistical uncertainty, while the outer error bars represent the statistical and systematic uncertainties added in quadrature. Additional uncertainties of 1.8% due to the luminosity measurement and 4.3% due to the branching ratio measurement are not included. The solid line and the shaded band represent the NLO QCD predictions with their uncertainties.*

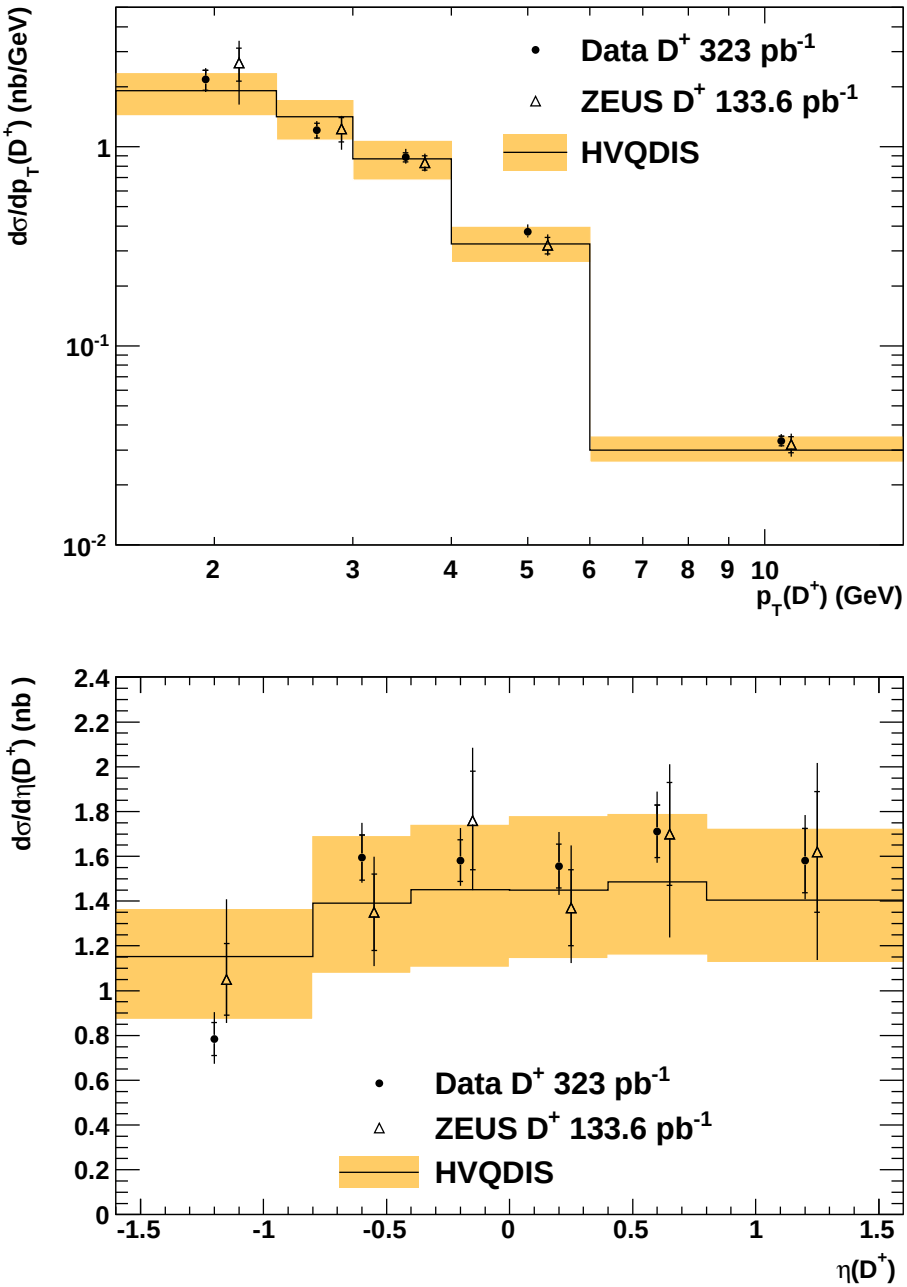


**Figure 7.8:** *Differential cross sections for $D^+$ meson production as a function of $p_{\mathrm{T}}(D^+)$ (top) and $\eta(D^+)$ (bottom). The cross sections are given in the kinematic region $5 < Q^2 < 1000\ GeV^2$, $0.02 < y < 0.7$, $1.5 < p_{\mathrm{T}}(D^+) < 15\ GeV$, $|\eta(D^+| < 1.6$. Details are the same as in Fig. 7.7.*

| $Q^2$ (GeV$^2$) | $d\sigma/dQ^2$ | $\Delta_{stat}$ ($nb$/GeV$^2$) | $\Delta_{syst}$ | $\mathcal{C}^{\rm rad}$ |
|---|---|---|---|---|
| 5 , 10 | 0.326 | $\pm$0.019 | $^{+0.032}_{-0.015}$ | 1.014 |
| 10 , 20 | 0.139 | $\pm$0.0072 | $^{+0.0120}_{-0.0085}$ | 1.012 |
| 20 , 40 | 0.0416 | $\pm$0.0025 | $^{+0.0041}_{-0.0037}$ | 1.016 |
| 40 , 80 | 0.00941 | $\pm$0.00072 | $^{+0.00093}_{-0.00088}$ | 1.019 |
| 80 , 200 | 0.00194 | $\pm$0.00020 | $^{+0.00018}_{-0.00017}$ | 1.035 |
| 200 ,1000 | 0.000093 | $\pm$0.000021 | $^{+0.000009}_{-0.000009}$ | 1.101 |
| $y$ | $d\sigma/dy$ | $\Delta_{stat}$ ($nb$) | $\Delta_{syst}$ | $\mathcal{C}^{\rm rad}$ |
| 0.02, 0.1 | 14.90 | $\pm$0.97 | $^{+1.20}_{-0.75}$ | 1.008 |
| 0.1 , 0.2 | 12.60 | $\pm$0.59 | $^{+0.85}_{-0.53}$ | 1.011 |
| 0.2 , 0.3 | 7.84 | $\pm$0.49 | $^{+0.62}_{-0.43}$ | 1.021 |
| 0.3 , 0.4 | 5.55 | $\pm$0.46 | $^{+0.57}_{-0.45}$ | 1.023 |
| 0.4 , 0.5 | 3.59 | $\pm$0.41 | $^{+0.43}_{-0.36}$ | 1.027 |
| 0.5 , 0.7 | 1.63 | $\pm$0.26 | $^{+0.22}_{-0.23}$ | 1.034 |
| $p_{\rm T}(D^+)$ (GeV) | $d\sigma/dp_{\rm T}(D^+)$ | $\Delta_{stat}$ ($nb$) | $\Delta_{syst}$ | $\mathcal{C}^{\rm rad}$ |
| 1.5 , 2.4 | 2.17 | $\pm$0.25 | $^{+0.19}_{-0.15}$ | 1.012 |
| 2.4 , 3.0 | 1.21 | $\pm$0.10 | $^{+0.10}_{-0.06}$ | 1.015 |
| 3.0 , 4.0 | 0.887 | $\pm$0.047 | $^{+0.072}_{-0.045}$ | 1.018 |
| 4.0 , 6.0 | 0.373 | $\pm$0.017 | $^{+0.029}_{-0.015}$ | 1.023 |
| 6.0 ,15.0 | 0.0333 | $\pm$0.0018 | $^{+0.0022}_{-0.0010}$ | 1.047 |
| $\eta(D^+)$ | $d\sigma/d\eta(D^+)$ | $\Delta_{stat}$ ($nb$) | $\Delta_{syst}$ | $\mathcal{C}^{\rm rad}$ |
| -1.6, -0.8 | 0.784 | $\pm$0.074 | $^{+0.096}_{-0.084}$ | 1.024 |
| -0.8, -0.4 | 1.59 | $\pm$0.10 | $^{+0.12}_{-0.05}$ | 1.019 |
| -0.4, 0.0 | 1.58 | $\pm$0.09 | $^{+0.11}_{-0.06}$ | 1.019 |
| 0.0 , 0.4 | 1.56 | $\pm$0.10 | $^{+0.12}_{-0.08}$ | 1.014 |
| 0.4 , 0.8 | 1.71 | $\pm$0.12 | $^{+0.13}_{-0.08}$ | 1.011 |
| 0.8 , 1.6 | 1.58 | $\pm$0.14 | $^{+0.15}_{-0.10}$ | 1.014 |

**Table 7.1:** *Differential cross sections for $D^+$ meson production in bins of $Q^2$, $y$, $p_{\rm T}(D^+)$ and $\eta(D^+)$. Statistical and systematic uncertainties are presented separately. Further uncertainties of* 1.8% *and* 4.3% *due to the luminosity and the branching ratio measurements were not included in $\Delta_{syst}$. Also shown are the correction factors to the QED Born level.*

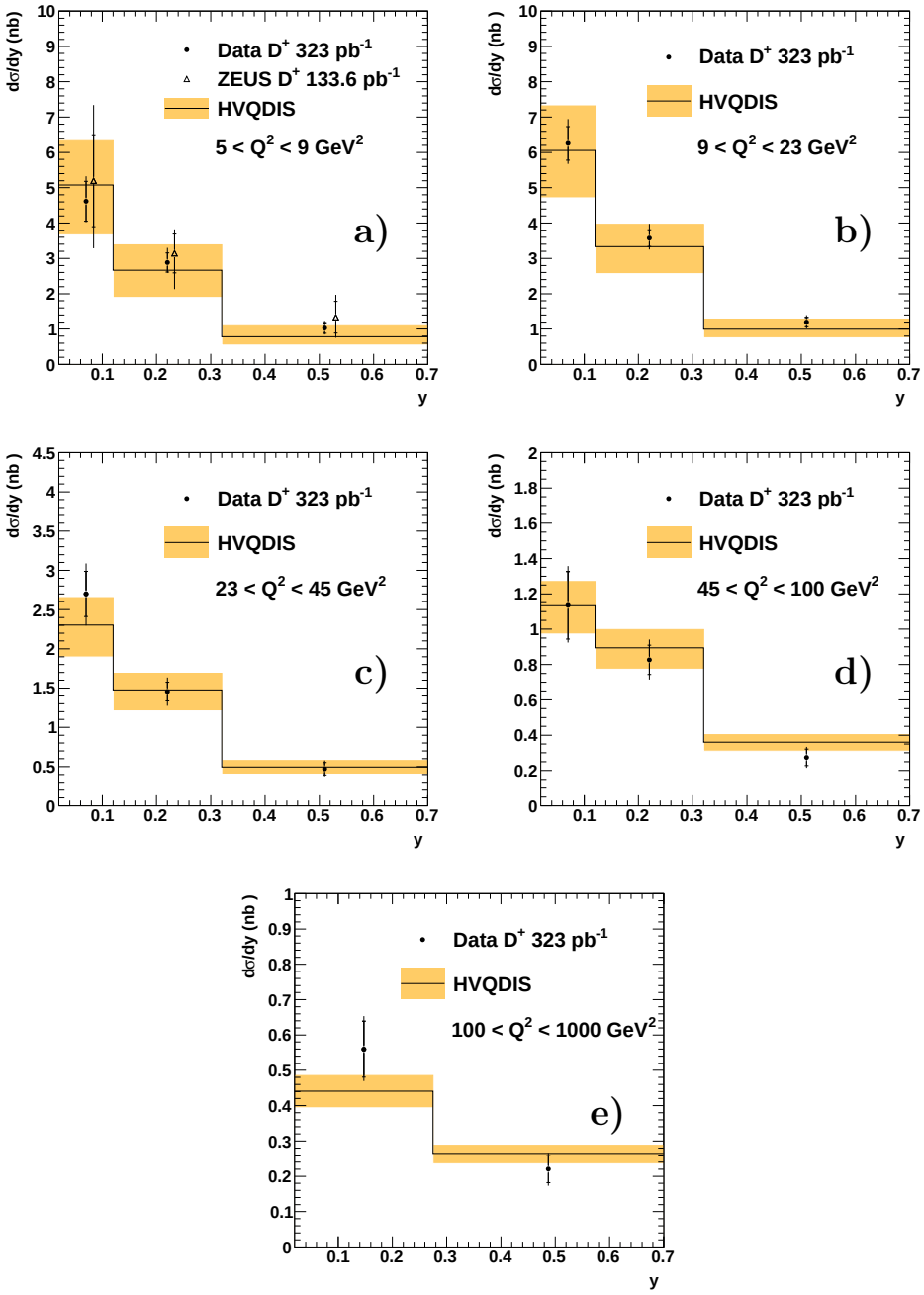


**Figure 7.9:** *Differential cross sections for $D^+$ meson production as a function of y for different regions of $Q^2$: a) $5 < Q^2 < 9\ GeV^2$, b) $9 < Q^2 < 23\ GeV^2$, c) $23 < Q^2 < 40\ GeV^2$, d) $40 < Q^2 < 100\ GeV^2$ and e) $100 < Q^2 < 1000\ GeV^2$. Details are the same as in Fig. 7.7.*

| Bin | $Q^2$ (GeV$^2$) | $y$ | $d\sigma/dy$ | $\Delta_{stat}$ | $\Delta_{syst}$ | $d\sigma^{NLO}/dy$ | $\Delta_{theo}$ | $\mathcal{C}^{\mathrm{rad}}$ |
|---|---|---|---|---|---|---|---|---|
| | | | | ($nb$) | | ($nb$) | | |
| 1 | 5 , 9 | 0.02 , 0.12 | 4.62 | $\pm 0.56$ | $^{+0.42}_{-0.18}$ | 5.1 | $^{+1.3}_{-1.4}$ | 1.002 |
| 2 | | 0.12 , 0.32 | 2.88 | $\pm 0.27$ | $^{+0.32}_{-0.13}$ | 2.66 | $^{+0.72}_{-0.77}$ | 1.017 |
| 3 | | 0.32 , 0.70 | 1.03 | $\pm 0.15$ | $^{+0.14}_{-0.11}$ | 0.78 | $^{+0.31}_{-0.24}$ | 1.029 |
| 4 | 9 , 23 | 0.02 , 0.12 | 6.26 | $\pm 0.47$ | $^{+0.49}_{-0.33}$ | 6.1 | $^{+1.3}_{-1.3}$ | 1.001 |
| 5 | | 0.12 , 0.32 | 3.58 | $\pm 0.24$ | $^{+0.32}_{-0.22}$ | 3.33 | $^{+0.63}_{-0.77}$ | 1.013 |
| 6 | | 0.32 , 0.70 | 1.20 | $\pm 0.13$ | $^{+0.14}_{-0.13}$ | 1.00 | $^{+0.29}_{-0.24}$ | 1.021 |
| 7 | 23 , 45 | 0.02 , 0.12 | 2.72 | $\pm 0.29$ | $^{+0.26}_{-0.28}$ | 2.30 | $^{+0.35}_{-0.41}$ | 1.011 |
| 8 | | 0.12 , 0.32 | 1.46 | $\pm 0.12$ | $^{+0.13}_{-0.14}$ | 1.48 | $^{+0.21}_{-0.27}$ | 1.013 |
| 9 | | 0.32 , 0.70 | 0.476 | $\pm 0.079$ | $^{+0.068}_{-0.057}$ | 0.495 | $^{+0.084}_{-0.09}$ | 1.031 |
| 10 | 45 , 100 | 0.02 , 0.12 | 1.13 | $\pm 0.19$ | $^{+0.11}_{-0.09}$ | 1.13 | $^{+0.14}_{-0.16}$ | 1.011 |
| 11 | | 0.12 , 0.32 | 0.828 | $\pm 0.083$ | $^{+0.079}_{-0.076}$ | 0.90 | $^{+0.10}_{-0.12}$ | 1.016 |
| 12 | | 0.32 , 0.70 | 0.276 | $\pm 0.047$ | $^{+0.037}_{-0.037}$ | 0.36 | $^{+0.04}_{-0.05}$ | 1.029 |
| 13 | 100 , 1000 | 0.020,0.275 | 0.56 | $\pm 0.08$ | $^{+0.05}_{-0.04}$ | 0.441 | $^{+0.044}_{-0.047}$ | 1.087 |
| 14 | | 0.275,0.700 | 0.220 | $\pm 0.038$ | $^{+0.025}_{-0.027}$ | 0.265 | $^{+0.023}_{-0.029}$ | 1.038 |

**Table 7.2:** *Differential cross sections for $D^+$ meson production as a function of $y$ in five regions of $Q^2$. Statistical and systematic uncertainties are presented separately. Further uncertainties of* 1.8% *and* 4.3% *due to the luminosity and the branching ratio measurements were not included in $\Delta_{syst}$. Also shown are the NLO QCD predictions with associated uncertainties, which will be used in the following Chapter 8. The correction factors to the QED Born level are tabulated as well.*

weighting has been introduced resulting in small shifts of the central values as well as in a better estimate of the systematic uncertainty due to the Monte Carlo model.

# Chapter 8

# Extraction of $F_2^c$

In order to compare the charm measurements obtained with different analysis techniques as well as to combine them one has to extrapolate the measurements to a common phase space. It is convenient to choose the full phase space, since many theoretical calculations, for example all GM-VFNS predictions, have been done for inclusive quantities only (e.g. [116–118]). Therefore the measured double–differential cross sections in $Q^2$ and $y$ were used to extract the charm contribution to the structure function $F_2$, denoted as $F_2^c$ or $F_2^{c\bar{c}}$. This Chapter presents the extraction procedure together with the results obtained in this analysis.

## 8.1 $F_2^c$ extraction procedure

The measurements of the charm cross sections are always done with a specific final state and only a restricted phase space in $p_{\mathrm{T}}$ and $\eta$ accessible due to detector limitations. Therefore an extrapolation from the measured visible cross sections to the inclusive charm cross sections is required:

$$F^c_{2,\mathrm{meas}}(x_i, Q_i^2) = \frac{\sigma_{i,\mathrm{meas}}(ep \to e' c\bar{c}X \to e' D^+ X)}{\sigma_{i,\mathrm{theo}}(ep \to e' c\bar{c}X \to e' D^+ X)} F^c_{2,\mathrm{theo}}(x_i, Q_i^2) \qquad (8.1)$$

Here $\sigma_{i,\mathrm{meas}}$ and $\sigma_{i,\mathrm{theo}}$ are the measured and predicted, respectively, cross sections in the $i$-th bin of $Q^2$ and $y$, whereas $Q_i^2$ and $x_i$ lie within this bin. For the calculation of $\sigma_{i,\mathrm{theo}}$ and $F^c_{2,\mathrm{theo}}$ the same setup in HVQDIS was used as for the comparison with the differential cross sections, see Chapter 2 for details. The uncertainty on the extrapolation procedure was estimated

by simultaneously varying the settings of the $\sigma_{i,\text{theo}}$ and $F^c_{2,\text{theo}}$ calculations and adding the resulting uncertainties in quadrature. Values of $F^c_{2,\text{theo}}$ were extracted from the inclusive charm quark production cross sections in the vicinity of each point of $Q_i^2$ and $x_i$ using (1.13). The contribution from $F_L^c$ had to be subtracted in HVQDIS for the $F^c_{2,\text{theo}}$ calculation. The size of the $F_L^c$ contribution is negligible on average and rises up to 3% at the high values of $y$. Running of $\alpha$ with $Q^2$ was taken into account in the $F^c_{2,\text{theo}}$ extraction.

The extrapolation is based on theory, therefore a good agreement between the data and NLO QCD predictions in the measured phase space is essential. Figures 7.7 and 7.8 have shown that the NLO QCD predictions by HVQDIS provide good description of the measured cross section.

Nevertheless, to rely less on the theoretical calculations and thus to reduce the theoretical uncertainties it is desirable to maximise the kinematic coverage in $p_\text{T}$ and $\eta$. A measure of inclusiveness is the extrapolation factor, $\mathcal{F}$, which is defined as the ratio of the inclusive meson production cross section to the cross section in a certain kinematic region. For example, for $p_\text{T}(D^+) > 3\,\text{GeV}$, which can be accessed without lifetime information, e.g. see [119], the extrapolation factors are $\approx 9$ at low $Q^2$ and $\approx 2.5$ at high $Q^2$, whereas for the kinematic region of this analysis, i.e. $p_\text{T}(D^+) > 1.5\,\text{GeV}$, the extrapolation factors range between $\approx 3$ and $\approx 1.5$ at low and high $Q^2$, respectively. The restriction on $\eta$ manifests itself in the dependence of the extrapolation factor on $x$, since going to very low or very high values of $x$ the $c\bar{c}$ system gets boosted either to the rear or to the forward direction and thus can escape detection.

## 8.2 $F_2^c$ results

The values of the measured and predicted cross sections in bins of $Q^2$ and $y$ were given in Table 7.2. Values of $Q_i^2$ and $x_i$, at which $F_2^c$ was extracted, were chosen to coincide with the grid adopted by the ZEUS and H1 $F_2^c$ combination group [31] to simplify inclusion of this data in the combination in the future.

The $F_2^c$ measurement is presented in Fig. 8.1 as a function of $x$ for fixed values of $Q^2$. The values and the corresponding uncertainties are given in Table 8.1. The theoretical uncertainty, $\Delta_{theo}$, is significant compared to the experimental uncertainties at low $Q^2$ and its fractional contribution decreases as $Q^2$ increases. A comparison of these results to the inclusive $F_2$ measurements, e.g. [20], reveals that the contribution of charm production to the

| $Q^2$ (GeV$^2$) | $x$ | $F_2^c$ | $\Delta_{stat}$ | $\Delta_{syst}$ | $\Delta_{theo}$ | $\mathcal{F}$ | $\mathcal{C}_L$ |
|---|---|---|---|---|---|---|---|
| 6.5 | 0.00016 | 0.218 | $\pm 0.031$ | $^{+0.030}_{-0.023}$ | $^{+0.030}_{-0.038}$ | 3.04 | 0.986 |
| | 0.00046 | 0.131 | $\pm 0.012$ | $^{+0.014}_{-0.006}$ | $^{+0.022}_{-0.012}$ | 2.15 | 0.999 |
| | 0.00202 | 0.0644 | $\pm 0.0078$ | $^{+0.0059}_{-0.0025}$ | $^{+0.0091}_{-0.0078}$ | 3.05 | 1.000 |
| 20.4 | 0.0005 | 0.336 | $\pm 0.037$ | $^{+0.040}_{-0.037}$ | $^{+0.029}_{-0.046}$ | 2.59 | 0.978 |
| | 0.00135 | 0.211 | $\pm 0.014$ | $^{+0.019}_{-0.013}$ | $^{+0.015}_{-0.013}$ | 1.85 | 0.998 |
| | 0.0025 | 0.159 | $\pm 0.012$ | $^{+0.013}_{-0.009}$ | $^{+0.016}_{-0.011}$ | 2.61 | 1.000 |
| 35 | 0.0008 | 0.307 | $\pm 0.051$ | $^{+0.044}_{-0.037}$ | $^{+0.019}_{-0.022}$ | 2.4 | 0.971 |
| | 0.0014 | 0.257 | $\pm 0.021$ | $^{+0.023}_{-0.024}$ | $^{+0.015}_{-0.013}$ | 1.62 | 0.992 |
| | 0.0034 | 0.210 | $\pm 0.022$ | $^{+0.020}_{-0.022}$ | $^{+0.013}_{-0.020}$ | 2.12 | 1.000 |
| 60 | 0.0015 | 0.244 | $\pm 0.041$ | $^{+0.033}_{-0.032}$ | $^{+0.014}_{-0.012}$ | 2.48 | 0.975 |
| | 0.0032 | 0.212 | $\pm 0.021$ | $^{+0.020}_{-0.019}$ | $^{+0.010}_{-0.010}$ | 1.55 | 0.995 |
| | 0.008 | 0.144 | $\pm 0.024$ | $^{+0.014}_{-0.011}$ | $^{+0.014}_{-0.010}$ | 1.78 | 1.000 |
| 200 | 0.005 | 0.215 | $\pm 0.037$ | $^{+0.024}_{-0.026}$ | $^{+0.013}_{-0.008}$ | 2.04 | 0.976 |
| | 0.013 | 0.184 | $\pm 0.026$ | $^{+0.016}_{-0.014}$ | $^{+0.011}_{-0.0087}$ | 1.53 | 0.997 |

**Table 8.1:** *The values of the extracted $F_2^c$ at each $Q^2$ and $x$ point. The statistical ($\Delta_{stat}$), systematic ($\Delta_{syst}$) and theoretical ($\Delta_{theo}$) uncertainties are given separately. Further uncertainties of* 1.8% *and* 4.3% *due to the luminosity and the branching ratio measurements were not included in $\Delta_{syst}$. The theoretical uncertainty, $\Delta_{theo}$, represents the uncertainty due to the extrapolation. Also quoted are the kinematic extrapolation factors, $\mathcal{F}$, and the correction for $F_L^c$, $\mathcal{C}_L = \sigma^{\pm}_{r,c\bar{c}}/F_2^c$.*

inclusive NC DIS scattering rises from $\sim 10\%$ at low $Q^2$ and high $x$ up to $\sim 20\%$ at the high values of $Q^2$.

The measurements are compared with the combined ZEUS and H1 $F_2^c$ measurements [31]. The data are in good agreement. The measurements are also well described by the NLO QCD predictions in the GM-VFNS based on HERAPDF1.0 [20] with the associated uncertainty due to the charm quark mass parameter in the PDF fit. The agreement is even more remarkable, if one considers that neither $F_2^c$ nor $F_L$ data were included in the HERAPDF1.0 fit, hence the gluon PDF is purely constrained by the scaling violations of the inclusive $F_2$ measurements.

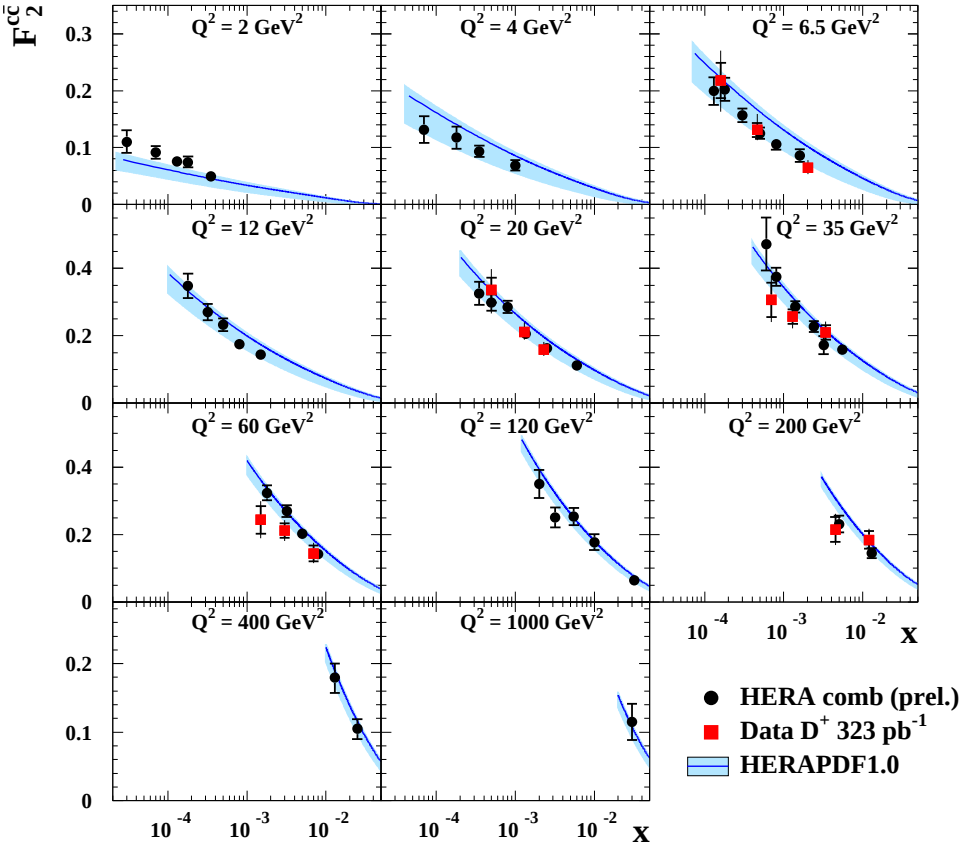


**Figure 8.1:** *The structure function $F_2^c$ as a function of $x$ for various values of $Q^2$. The results obtained in this analysis (filled red squares) are compared to the combination of $F_2^c$ measurements with various techniques from both ZEUS and H1 [31] (filled black circles). For results presented in this thesis the inner error bars correspond to the statistical uncertainty, while the outer error bars represent the statistical, systematic and theoretical uncertainties added in quadrature. Also shown are predictions in the GM-VFNS based on HERAPDF1.0 with a variation of the charm quark mass parameter in the range of 1.3 GeV to 1.7 GeV (filled band).*

# Chapter 9

# Conclusions

Charm production in deep inelastic scattering has been measured using the reconstruction of $D^+$ mesons in the kinematic region $5 < Q^2 < 1000\,\mathrm{GeV}^2$, $0.02 < y < 0.7$, $1.5 < p_\mathrm{T}(D^+) < 15\,\mathrm{GeV}$ and $|\eta(D^+| < 1.6$. The measurement was performed on $323\,\mathrm{pb}^{-1}$ of data collected with the ZEUS detector during the 2005 – 2007 $e^\pm p$ running at the center-of-mass energy $\sqrt{s} = 318\,\mathrm{GeV}$.

A key feature of this analysis is the accurate reconstruction of the displaced secondary vertices from the $D^+$ decays. This leads to a significant suppression of the combinatorial background from light flavour production and to an extension of the kinematic region to lower transverse momentum of the $D^+$ meson compared to the previously published $D^+$ results from HERA I. Lifetime tagging of heavy quark production became possible with the installation of the Micro Vertex Detector.

The results obtained in this thesis are in agreement with the previously published ZEUS $D^+$ measurement in the same phase space [1]. The comparison has shown that in addition to the reduced statistical errors owing to a roughly 2.5 times larger data set and better background suppresion, a significant improvement in the systematic uncertainties was achieved, in particular due to better reconstruction of the secondary vertices.

The measured single– and double–differential cross sections have also been used to test the next-to-leading order QCD predictions. Good agreement between the cross sections in the data and NLO QCD calculations was found. This allowed to use the predictions to extrapolate outside of the kinematic region in $p_\mathrm{T}$ and $\eta$ and to extract the charm contribution to the structure function $F_2$.

$F_2^c$ is an important input to the parton density fits – not only it is directly sensitive to the gluon PDF in the proton due to the BGF production

mechanism but also it allows to constrain the heavy flavour treatment in the GM-VFNS PDF fits [37]. The $F_2^c$ measurements presented in this thesis are in agreement with the combination of the ZEUS and H1 measurements using various experimental techniques [31]. The new measurement is based on a complementary technique and has different systematic uncertainties and is thus a valuable input to the combination in the future.

There are still some possible improvements to further exploit the potential of this analysis. The most straightforward is to include the remaining 10% of the HERA II statistics collected in 2004, but this would require dedicated trigger studies for the period to be added. One of the dominant systematics in this analysis originates from the track efficiency simulation. This is expected to be improved as soon as corresponding studies will come to a conclusion. Further extending the kinematic coverage would allow to rely less on the theory in the $F_2^c$ extraction. Therefore, one more interesting study would be to investigate the feasibility of going further down in $p_{\mathrm{T}}(D^+)$. However, this is difficult as efficient supression of the combinatorial background requires a strong cut on the $D^+$ lifetime, which is impossible without a significantly boosted $D^+$ rest frame.

# Appendix A

# FLT tracking efficiency corrections

This appendix is devoted to the corrections for FLT tracking requirements. Section 6.2.6 describes the procedure to derive the corrections. Obtained distributions were fitted with monotonic functions of the hadronic angle, $\gamma_{\mathrm{had}}$. The dependences for the LOOSE and SEMILOOSE tracking requirements as well as LOOSE+q95b and SEMILOOSE+q95b were found to be the same. The distributions were fitted by:

- LOOSE: $p_1 x + p_0$;
- LOOSE+q95b: $p_5 x^5 + p_4 x^4 + p_3 x^3 + p_2 x^2 + p_1 x + p_0$;
- TIGHT+q95b: $p_4 x^4 + p_3 x^3 + p_2 x^2 + p_1 x + p_0$.

The fit output in each $Q^2$ range for the separate data taking periods is presented in Table A.1.

| Period | $Q^2$(GeV$^2$) | LOOSE | | LOOSE+q95 | | | | | | TIGHT+q95 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $p_0$ | $p_1$ | $p_0$ | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ | $p_0$ | $p_1$ | $p_2$ | $p_3$ | $p_4$ |
| 2005 $e^-$ | (5;20) | 0.999 | -0.00945 | 0.921 | 0.316 | -0.375 | 0.199 | -0.0527 | 0.0056 | 0.949 | 0.183 | -0.197 | 0.0686 | -0.00844 |
| | (20;80) | 0.999 | -0.0112 | 0.953 | 0.215 | -0.276 | 0.146 | -0.0364 | 0.00354 | 0.97 | 0.117 | -0.162 | 0.0622 | -0.00814 |
| | (80;200) | 0.998 | -0.0103 | 0.959 | 0.135 | -0.184 | 0.106 | -0.0291 | 0.00305 | 0.959 | 0.0874 | -0.134 | 0.0569 | -0.00838 |
| | (200;1000) | 0.994 | -0.0091 | 0.933 | 0.212 | -0.275 | 0.149 | -0.0352 | 0.00274 | 0.928 | 0.163 | -0.226 | 0.104 | -0.0168 |
| 2006 $e^-$ | (5;20) | 0.997 | -0.0095 | 1.03 | 0.0322 | -0.0636 | 0.0265 | -0.0051 | 0.000471 | 0.992 | 0.116 | -0.17 | 0.0673 | -0.00903 |
| | (20;80) | 0.998 | -0.0106 | 0.939 | 0.287 | -0.38 | 0.211 | -0.0551 | 0.00558 | 0.966 | 0.138 | -0.187 | 0.0726 | -0.00959 |
| | (80;200) | 0.998 | -0.00993 | 0.941 | 0.221 | -0.342 | 0.232 | -0.0742 | 0.00898 | 0.956 | 0.0803 | -0.118 | 0.0487 | -0.0073 |
| | (200;1000) | 0.993 | -0.00873 | 0.887 | 0.47 | -0.762 | 0.548 | -0.182 | 0.0225 | 0.908 | 0.245 | -0.34 | 0.161 | -0.0259 |
| 2006 – 2007 $e^+$ | (5;20) | 1 | -0.0247 | 1.08 | -0.262 | 0.337 | -0.242 | 0.0785 | -0.00914 | 0.947 | 0.18 | -0.259 | 0.0976 | -0.0122 |
| | (20;80) | 1 | -0.0264 | 0.949 | 0.23 | -0.34 | 0.19 | -0.0498 | 0.0052 | 0.981 | 0.0647 | -0.157 | 0.0575 | -0.00632 |
| | (80;200) | 1 | -0.0278 | 0.925 | 0.232 | -0.335 | 0.203 | -0.06 | 0.00711 | 0.934 | 0.125 | -0.209 | 0.0786 | -0.00949 |
| | (200;1000) | 0.994 | -0.028 | 0.815 | 0.71 | -1.1 | 0.742 | -0.233 | 0.0277 | 0.828 | 0.479 | -0.666 | 0.309 | -0.0491 |

**Table A.1:** *Parameters of the fitting functions that parametrise the FLT tracking efficiency corrections.*

# Appendix B

# Individual systematic uncertainties

Contributions from individual systematic sources are presented in this appendix. Tables B.1 and B.2 give a breakdown of the systematic uncertainties in bins of the double– and single–differential cross sections. Uncertainties arising from the tracking and MVD hit efficiency simulation as well as from the lumi and branching ratio measurements are not tabulated, since those are constant over the kinematic phase space. Comparison of the individual systematic sources with the statistical uncertainties is shown in Figs. B.1–B.9.

The systematic sources are indexed in the order they were introduced in Section 7.2:

$\delta_1$ **:** Position of the scattered electron in the RCAL;

$\delta_2$ **:** Energy scale in the EMC;

$\delta_3$ **:** Energy scale in the HAC;

$\delta_4$ **:** FLT tracking efficiency;

$\delta_5$ **:** Signal extraction procedure(background parametrisation);

$\delta_6$ **:** Signal extraction procedure(signal parametrisation);

$\delta_7$ **:** Reconstruction of the decay length;

$\delta_8$ **:** $b$ quark contribution;

| Bin | $\delta_1$ | $\delta_2$ | $\delta_3$ | $\delta_4$ | $\delta_5$ | $\delta_6$ | $\delta_7$ | $\delta_8$ | $\delta_9$ | $\delta_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $^{+6.4\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+0.0\%}_{-0.4\%}$ | $^{+1.2\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.8\%}$ | $^{+1.1\%}_{-1.1\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+1.6\%}_{-2.0\%}$ |
| 2 | $^{+7.6\%}_{-1.8\%}$ | $^{+0.0\%}_{-0.1\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+1.3\%}_{-1.3\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+2.7\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.1\%}$ | $^{+1.8\%}_{-1.8\%}$ | $^{+1.8\%}_{-1.8\%}$ | $^{+3.4\%}_{-2.8\%}$ |
| 3 | $^{+4.9\%}_{-0.0\%}$ | $^{+0.2\%}_{-0.0\%}$ | $^{+0.6\%}_{-0.5\%}$ | $^{+2.7\%}_{-2.7\%}$ | $^{+0.6\%}_{-0.0\%}$ | $^{+3.9\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.0\%}$ | $^{+2.8\%}_{-2.8\%}$ | $^{+9.6\%}_{-9.6\%}$ | $^{+1.8\%}_{-1.5\%}$ |
| 4 | $^{+1.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.1\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+0.0\%}_{-0.7\%}$ | $^{+2.1\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.4\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+4.1\%}_{-5.0\%}$ |
| 5 | $^{+0.0\%}_{-0.6\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.1\%}$ | $^{+1.2\%}_{-1.2\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+4.1\%}_{-0.0\%}$ | $^{+0.6\%}_{-0.0\%}$ | $^{+1.4\%}_{-1.4\%}$ | $^{+2.1\%}_{-2.1\%}$ | $^{+4.4\%}_{-5.4\%}$ |
| 6 | $^{+0.3\%}_{-0.0\%}$ | $^{+2.2\%}_{-1.6\%}$ | $^{+0.1\%}_{-0.2\%}$ | $^{+1.3\%}_{-1.3\%}$ | $^{+0.0\%}_{-0.5\%}$ | $^{+1.2\%}_{-0.0\%}$ | $^{+0.0\%}_{-3.5\%}$ | $^{+2.9\%}_{-2.9\%}$ | $^{+8.8\%}_{-8.8\%}$ | $^{+3.3\%}_{-4.0\%}$ |
| 7 | $^{+0.0\%}_{-0.2\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.4\%}$ | $^{+0.1\%}_{-0.1\%}$ | $^{+0.0\%}_{-1.3\%}$ | $^{+4.8\%}_{-0.0\%}$ | $^{+0.0\%}_{-7.5\%}$ | $^{+1.0\%}_{-1.0\%}$ | $^{+1.3\%}_{-1.3\%}$ | $^{+5.5\%}_{-6.7\%}$ |
| 8 | $^{+0.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.4\%}_{-0.2\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+0.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-3.0\%}$ | $^{+0.0\%}_{-3.8\%}$ | $^{+3.3\%}_{-3.3\%}$ | $^{+1.2\%}_{-1.2\%}$ | $^{+5.8\%}_{-7.1\%}$ |
| 9 | $^{+0.0\%}_{-0.6\%}$ | $^{+1.3\%}_{-1.5\%}$ | $^{+1.4\%}_{-0.0\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+0.0\%}_{-1.7\%}$ | $^{+6.3\%}_{-0.0\%}$ | $^{+2.8\%}_{-0.0\%}$ | $^{+4.7\%}_{-4.7\%}$ | $^{+7.9\%}_{-7.9\%}$ | $^{+6.1\%}_{-7.4\%}$ |
| 10 | $^{+0.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.4\%}_{-0.9\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+5.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.8\%}$ | $^{+1.2\%}_{-1.2\%}$ | $^{+2.3\%}_{-2.3\%}$ | $^{+5.3\%}_{-6.4\%}$ |
| 11 | $^{+0.1\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.4\%}_{-0.6\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.9\%}$ | $^{+0.0\%}_{-0.3\%}$ | $^{+3.3\%}_{-3.3\%}$ | $^{+0.7\%}_{-0.7\%}$ | $^{+6.5\%}_{-7.9\%}$ |
| 12 | $^{+0.0\%}_{-0.3\%}$ | $^{+0.2\%}_{-1.4\%}$ | $^{+1.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+1.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+6.5\%}_{-6.5\%}$ | $^{+6.6\%}_{-6.6\%}$ | $^{+7.7\%}_{-9.4\%}$ |
| 13 | $^{+0.0\%}_{-0.5\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.7\%}_{-0.3\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+1.8\%}_{-0.0\%}$ | $^{+3.6\%}_{-0.0\%}$ | $^{+0.0\%}_{-5.0\%}$ | $^{+3.2\%}_{-3.2\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+3.7\%}_{-4.6\%}$ |
| 14 | $^{+0.1\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.2\%}$ | $^{+0.9\%}_{-1.3\%}$ | $^{+0.5\%}_{-0.5\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+0.0\%}_{-3.3\%}$ | $^{+0.0\%}_{-5.5\%}$ | $^{+6.9\%}_{-6.9\%}$ | $^{+3.2\%}_{-3.2\%}$ | $^{+5.5\%}_{-6.7\%}$ |

**Table B.1:** *Breakdown of the systematic uncertainties for the double–differential cross sections in bins of y for various ranges of $Q^2$. The bin order goes along with Table 7.2.*

$\delta_9$ **:** Monte Carlo model (variation of the $\eta(D^+)$ shape);

$\delta_{10}$ **:** Monte Carlo model (variation of the $p_{\mathrm{T}}(D^+)$ and $Q^2$ shapes);

The bins are numbered following the order in Tables 7.1 and 7.2.

| Bin | $\delta_1$ | $\delta_2$ | $\delta_3$ | $\delta_4$ | $\delta_5$ | $\delta_6$ | $\delta_7$ | $\delta_8$ | $\delta_9$ | $\delta_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $Q^2(1)$ | $^{+5.6\%}_{-0.8\%}$ | $^{+0.2\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+1.0\%}_{-1.0\%}$ | $^{+0.0\%}_{-0.3\%}$ | $^{+3.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.5\%}$ | $^{+1.9\%}_{-1.9\%}$ | $^{+3.0\%}_{-3.0\%}$ | $^{+0.0\%}_{-0.0\%}$ |
| $Q^2(2)$ | $^{+0.3\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.2\%}$ | $^{+0.1\%}_{-0.2\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+2.2\%}_{-0.0\%}$ | $^{+0.2\%}_{-0.0\%}$ | $^{+1.5\%}_{-1.5\%}$ | $^{+3.5\%}_{-3.5\%}$ | $^{+3.8\%}_{-4.6\%}$ |
| $Q^2(3)$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+0.5\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+1.9\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.9\%}$ | $^{+2.4\%}_{-2.4\%}$ | $^{+3.1\%}_{-3.1\%}$ | $^{+6.3\%}_{-7.7\%}$ |
| $Q^2(4)$ | $^{+0.0\%}_{-0.2\%}$ | $^{+0.2\%}_{-0.6\%}$ | $^{+0.6\%}_{-0.1\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.1\%}$ | $^{+0.0\%}_{-0.8\%}$ | $^{+0.0\%}_{-2.6\%}$ | $^{+3.7\%}_{-3.7\%}$ | $^{+2.8\%}_{-2.8\%}$ | $^{+6.2\%}_{-7.6\%}$ |
| $Q^2(5)$ | $^{+0.0\%}_{-0.1\%}$ | $^{+0.0\%}_{-0.4\%}$ | $^{+0.6\%}_{-0.5\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.3\%}$ | $^{+0.0\%}_{-3.2\%}$ | $^{+4.4\%}_{-4.4\%}$ | $^{+2.4\%}_{-2.4\%}$ | $^{+5.3\%}_{-6.5\%}$ |
| $Q^2(6)$ | $^{+0.2\%}_{-0.0\%}$ | $^{+0.7\%}_{-1.3\%}$ | $^{+0.1\%}_{-1.4\%}$ | $^{+0.7\%}_{-0.7\%}$ | $^{+1.9\%}_{-0.0\%}$ | $^{+0.0\%}_{-4.5\%}$ | $^{+0.0\%}_{-3.6\%}$ | $^{+6.0\%}_{-6.0\%}$ | $^{+1.8\%}_{-1.8\%}$ | $^{+4.1\%}_{-5.0\%}$ |
| $y(1)$ | $^{+2.8\%}_{-0.1\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.7\%}_{-0.0\%}$ | $^{+0.1\%}_{-0.1\%}$ | $^{+0.0\%}_{-0.4\%}$ | $^{+2.4\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.4\%}$ | $^{+1.1\%}_{-1.1\%}$ | $^{+1.7\%}_{-1.7\%}$ | $^{+3.1\%}_{-3.8\%}$ |
| $y(2)$ | $^{+0.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+0.4\%}_{-0.4\%}$ | $^{+0.0\%}_{-0.3\%}$ | $^{+1.2\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.8\%}$ | $^{+2.0\%}_{-2.0\%}$ | $^{+0.5\%}_{-0.5\%}$ | $^{+1.6\%}_{-2.0\%}$ |
| $y(3)$ | $^{+0.6\%}_{-1.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.5\%}_{-0.1\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+0.1\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.1\%}$ | $^{+2.7\%}_{-2.7\%}$ | $^{+3.4\%}_{-3.4\%}$ | $^{+2.1\%}_{-2.6\%}$ |
| $y(4)$ | $^{+0.4\%}_{-0.1\%}$ | $^{+0.0\%}_{-0.1\%}$ | $^{+0.1\%}_{-0.1\%}$ | $^{+0.8\%}_{-0.8\%}$ | $^{+0.0\%}_{-0.1\%}$ | $^{+2.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+3.3\%}_{-3.3\%}$ | $^{+6.9\%}_{-6.9\%}$ | $^{+1.2\%}_{-1.5\%}$ |
| $y(5)$ | $^{+0.6\%}_{-0.0\%}$ | $^{+1.1\%}_{-0.2\%}$ | $^{+1.1\%}_{-0.0\%}$ | $^{+0.8\%}_{-0.8\%}$ | $^{+0.0\%}_{-0.9\%}$ | $^{+3.4\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.8\%}$ | $^{+4.5\%}_{-4.5\%}$ | $^{+8.6\%}_{-8.6\%}$ | $^{+1.9\%}_{-2.3\%}$ |
| $y(6)$ | $^{+0.0\%}_{-0.4\%}$ | $^{+1.2\%}_{-4.4\%}$ | $^{+0.5\%}_{-0.2\%}$ | $^{+0.7\%}_{-0.7\%}$ | $^{+0.5\%}_{-0.0\%}$ | $^{+0.6\%}_{-0.0\%}$ | $^{+0.0\%}_{-6.1\%}$ | $^{+7.8\%}_{-7.8\%}$ | $^{+8.9\%}_{-8.9\%}$ | $^{+1.8\%}_{-2.2\%}$ |
| $p_T(D^+)(1)$ | $^{+3.4\%}_{-2.3\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+0.0\%}_{-0.7\%}$ | $^{+0.6\%}_{-0.6\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+0.0\%}_{-4.0\%}$ | $^{+3.7\%}_{-3.7\%}$ | $^{+3.0\%}_{-3.0\%}$ | $^{+2.2\%}_{-1.8\%}$ |
| $p_T(D^+)(2)$ | $^{+2.2\%}_{-0.0\%}$ | $^{+0.1\%}_{-0.1\%}$ | $^{+1.0\%}_{-0.0\%}$ | $^{+0.6\%}_{-0.6\%}$ | $^{+0.3\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.5\%}$ | $^{+0.0\%}_{-1.2\%}$ | $^{+2.9\%}_{-2.9\%}$ | $^{+2.9\%}_{-2.9\%}$ | $^{+0.1\%}_{-0.1\%}$ |
| $p_T(D^+)(3)$ | $^{+0.4\%}_{-0.8\%}$ | $^{+0.1\%}_{-0.3\%}$ | $^{+0.6\%}_{-0.0\%}$ | $^{+0.6\%}_{-0.6\%}$ | $^{+0.0\%}_{-0.7\%}$ | $^{+2.9\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.1\%}$ | $^{+2.7\%}_{-2.7\%}$ | $^{+2.7\%}_{-2.7\%}$ | $^{+2.4\%}_{-2.0\%}$ |
| $p_T(D^+)(4)$ | $^{+1.0\%}_{-0.0\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+0.3\%}_{-0.1\%}$ | $^{+0.5\%}_{-0.5\%}$ | $^{+0.1\%}_{-0.0\%}$ | $^{+2.6\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.4\%}$ | $^{+2.6\%}_{-2.6\%}$ | $^{+2.2\%}_{-2.2\%}$ | $^{+1.7\%}_{-1.4\%}$ |
| $p_T(D^+)(5)$ | $^{+0.7\%}_{-0.0\%}$ | $^{+0.3\%}_{-0.7\%}$ | $^{+0.4\%}_{-0.1\%}$ | $^{+0.4\%}_{-0.4\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+1.1\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.8\%}$ | $^{+1.5\%}_{-1.5\%}$ | $^{+0.6\%}_{-0.6\%}$ | $^{+0.9\%}_{-0.8\%}$ |
| $\eta(D^+)(1)$ | $^{+0.8\%}_{-0.0\%}$ | $^{+1.7\%}_{-1.2\%}$ | $^{+0.7\%}_{-0.0\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+0.0\%}_{-1.5\%}$ | $^{+2.2\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.9\%}$ | $^{+2.7\%}_{-2.7\%}$ | $^{+9.6\%}_{-9.6\%}$ | $^{+1.3\%}_{-1.6\%}$ |
| $\eta(D^+)(2)$ | $^{+0.5\%}_{-0.2\%}$ | $^{+0.4\%}_{-0.8\%}$ | $^{+0.5\%}_{-0.0\%}$ | $^{+0.9\%}_{-0.9\%}$ | $^{+0.4\%}_{-0.0\%}$ | $^{+3.1\%}_{-0.0\%}$ | $^{+0.1\%}_{-0.0\%}$ | $^{+2.0\%}_{-2.0\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+1.3\%}_{-1.6\%}$ |
| $\eta(D^+)(3)$ | $^{+1.9\%}_{-0.6\%}$ | $^{+0.2\%}_{-0.0\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+0.6\%}_{-0.6\%}$ | $^{+0.2\%}_{-0.0\%}$ | $^{+0.0\%}_{-0.7\%}$ | $^{+0.0\%}_{-2.1\%}$ | $^{+2.5\%}_{-2.5\%}$ | $^{+0.0\%}_{-0.0\%}$ | $^{+1.5\%}_{-1.8\%}$ |
| $\eta(D^+)(4)$ | $^{+1.3\%}_{-0.2\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+0.4\%}_{-0.1\%}$ | $^{+0.5\%}_{-0.5\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+1.6\%}_{-0.0\%}$ | $^{+0.0\%}_{-3.2\%}$ | $^{+3.2\%}_{-3.2\%}$ | $^{+0.1\%}_{-0.1\%}$ | $^{+2.0\%}_{-2.4\%}$ |
| $\eta(D^+)(5)$ | $^{+0.4\%}_{-0.1\%}$ | $^{+0.2\%}_{-0.1\%}$ | $^{+0.2\%}_{-0.1\%}$ | $^{+0.3\%}_{-0.3\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+3.2\%}_{-0.0\%}$ | $^{+0.0\%}_{-1.7\%}$ | $^{+2.6\%}_{-2.6\%}$ | $^{+0.6\%}_{-0.6\%}$ | $^{+2.4\%}_{-3.0\%}$ |
| $\eta(D^+)(6)$ | $^{+2.6\%}_{-0.6\%}$ | $^{+0.0\%}_{-0.2\%}$ | $^{+0.8\%}_{-1.1\%}$ | $^{+0.2\%}_{-0.2\%}$ | $^{+0.0\%}_{-0.5\%}$ | $^{+3.5\%}_{-0.0\%}$ | $^{+0.0\%}_{-2.2\%}$ | $^{+3.2\%}_{-3.2\%}$ | $^{+3.5\%}_{-3.5\%}$ | $^{+2.1\%}_{-2.6\%}$ |

**Table B.2:** *Breakdown of the systematic uncertainties for the single–differential cross sections in bins of $Q^2$, $y$, $p_{\mathrm{T}}(D^+)$ and $\eta(D^+)$. Bins are the same as in Table 7.1.*

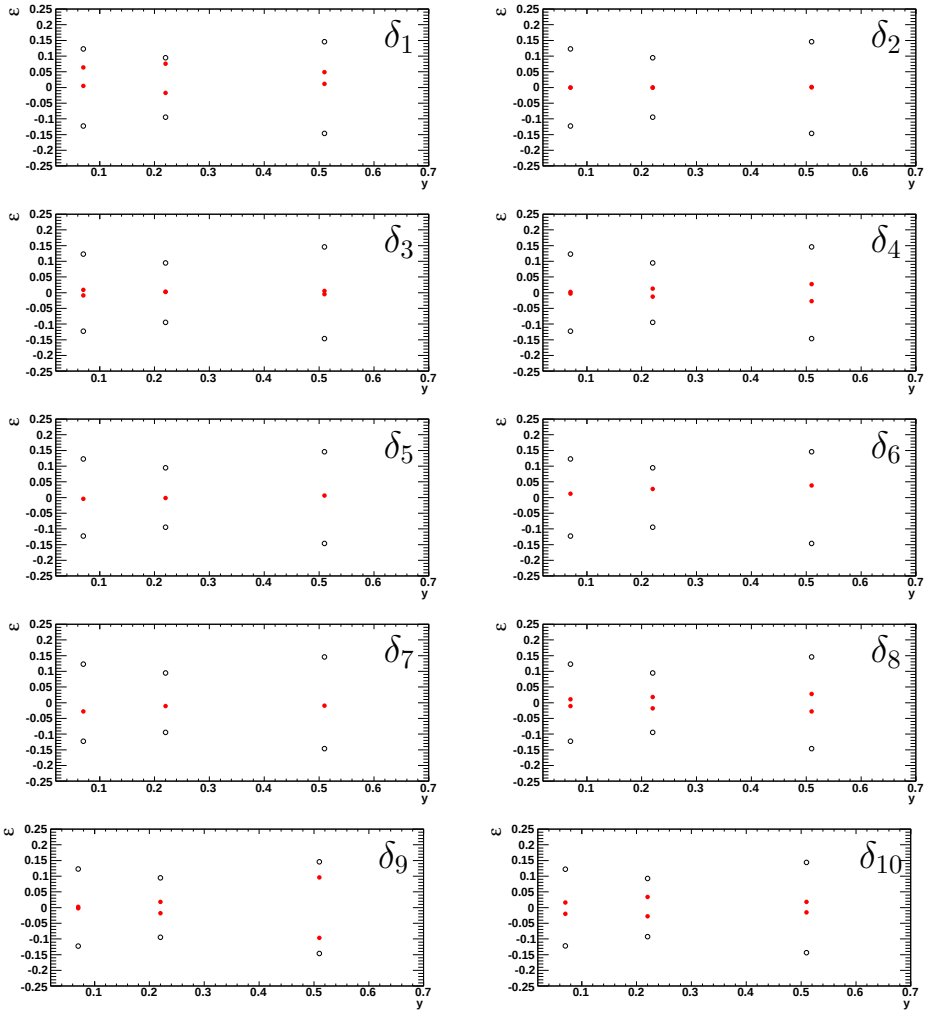


**Figure B.1:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of y for* $5 < Q^2 < 9\ GeV^2$.

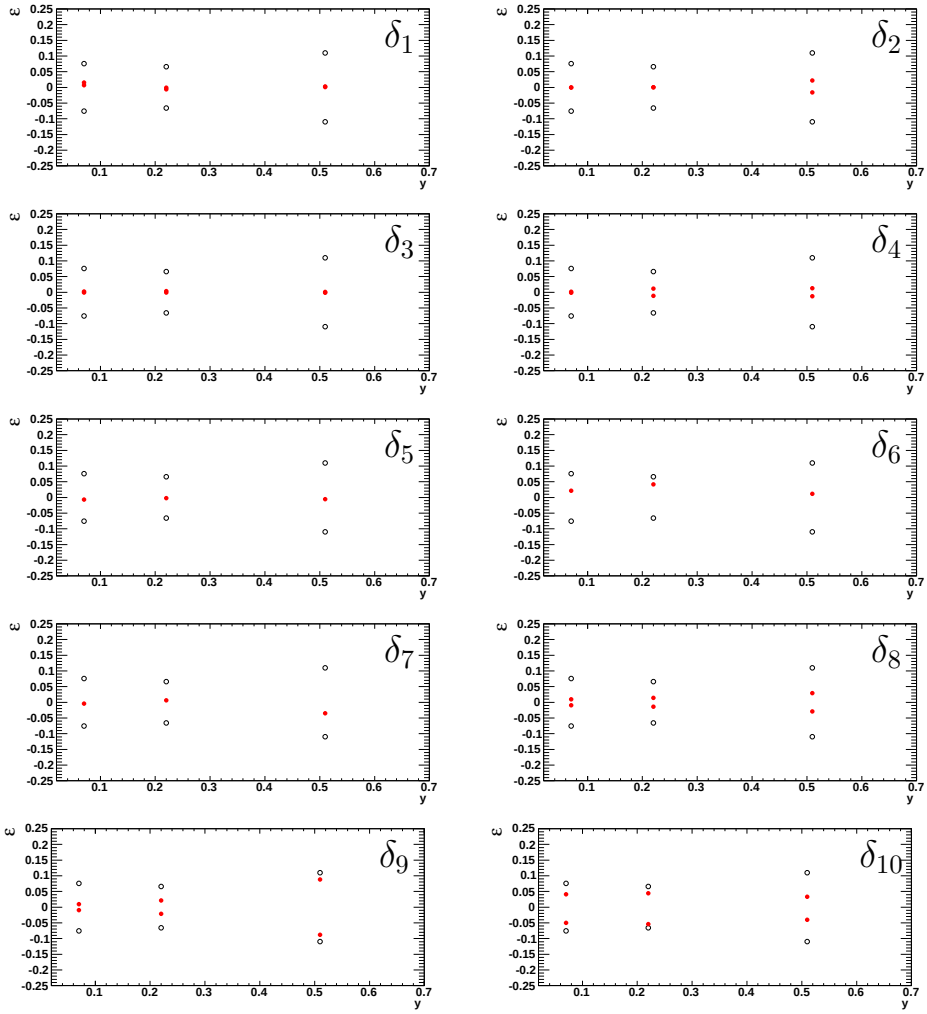


**Figure B.2:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of $y$ for $9 < Q^2 < 23\ GeV^2$.*

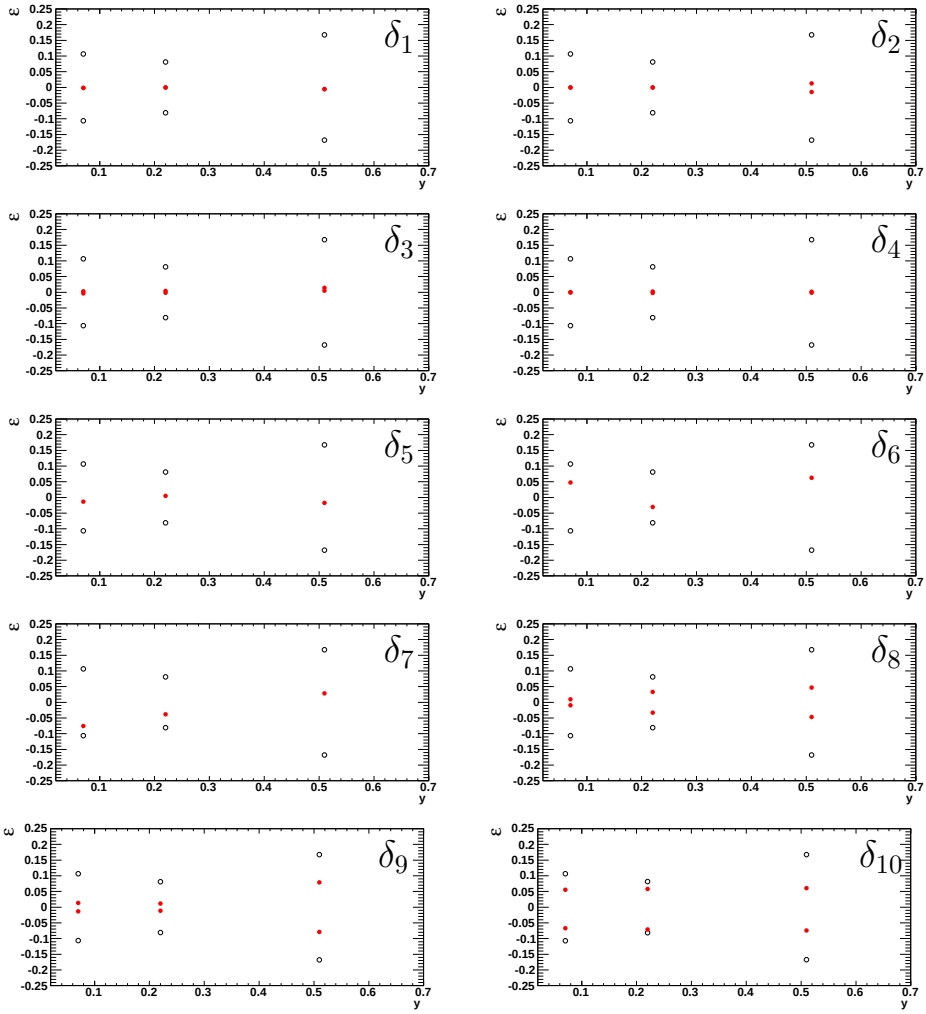


**Figure B.3:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of y for* $23 < Q^2 < 45\ GeV^2$.

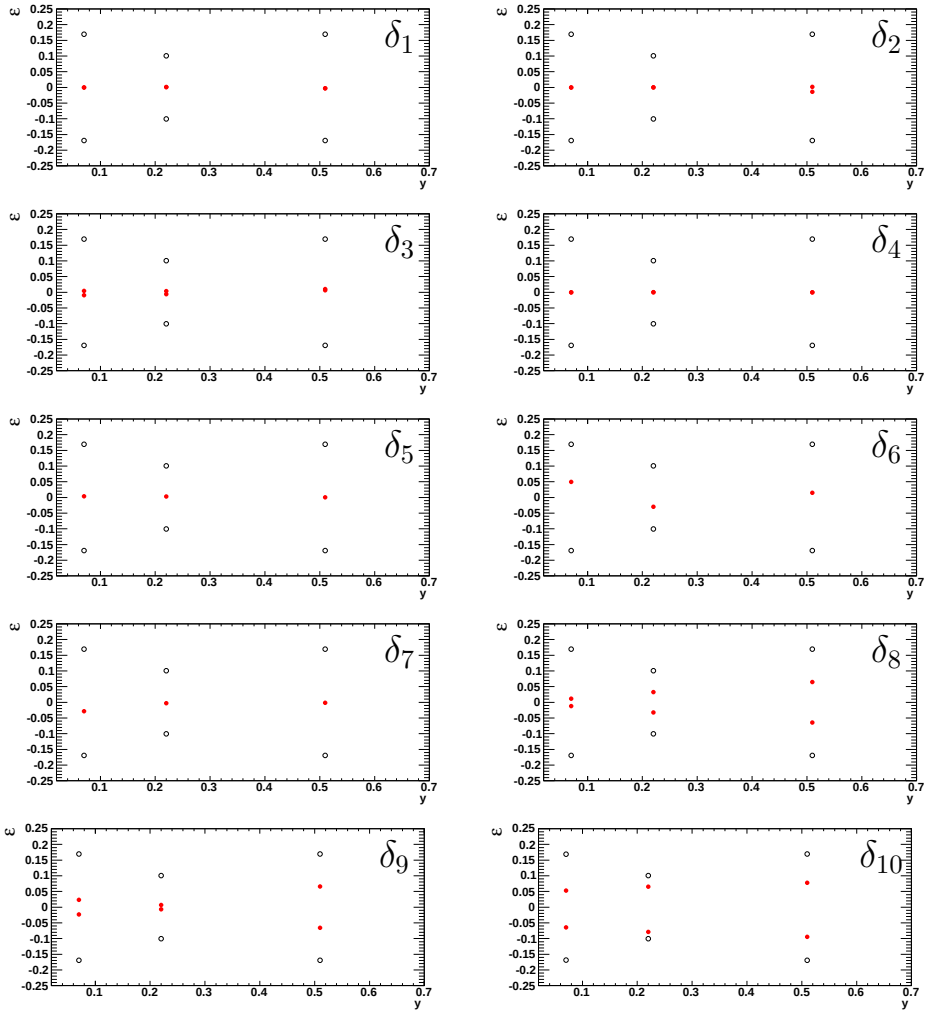


**Figure B.4:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of y for* $45 < Q^2 < 100\ GeV^2$.

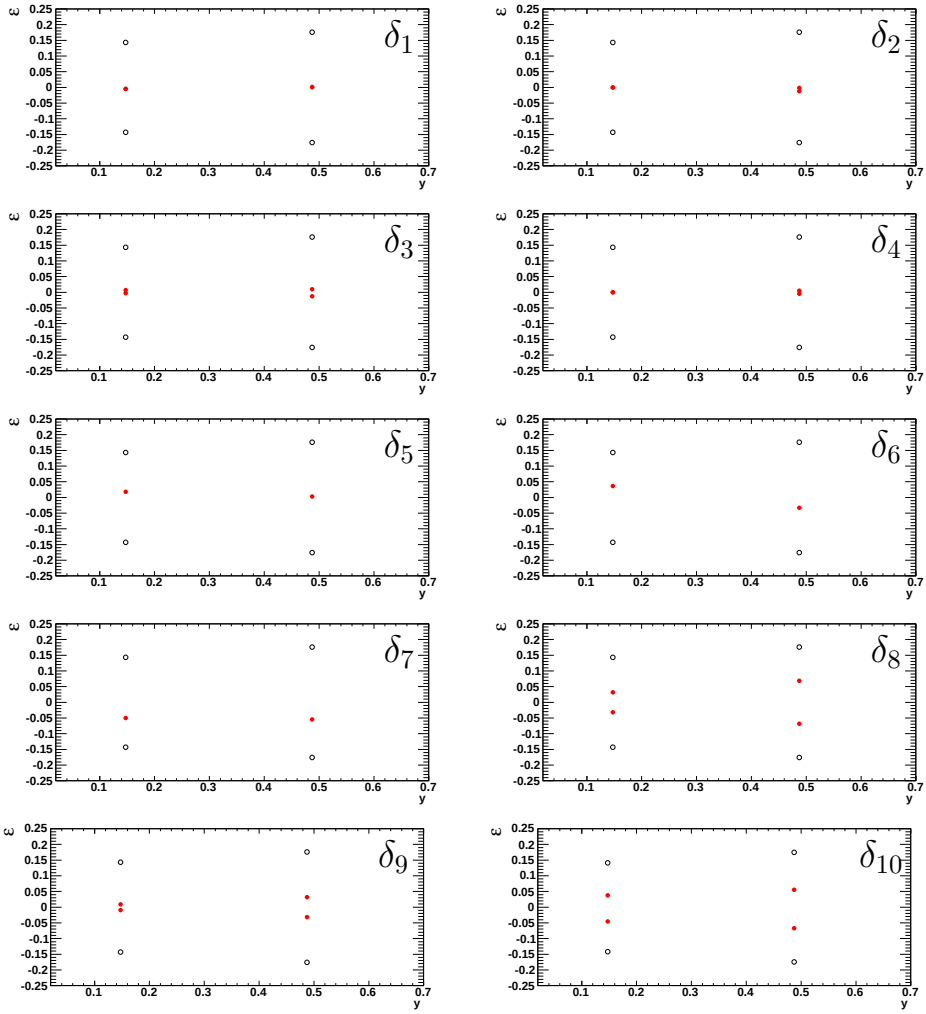


**Figure B.5:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of y for* $100 < Q^2 < 1000\ GeV^2$.

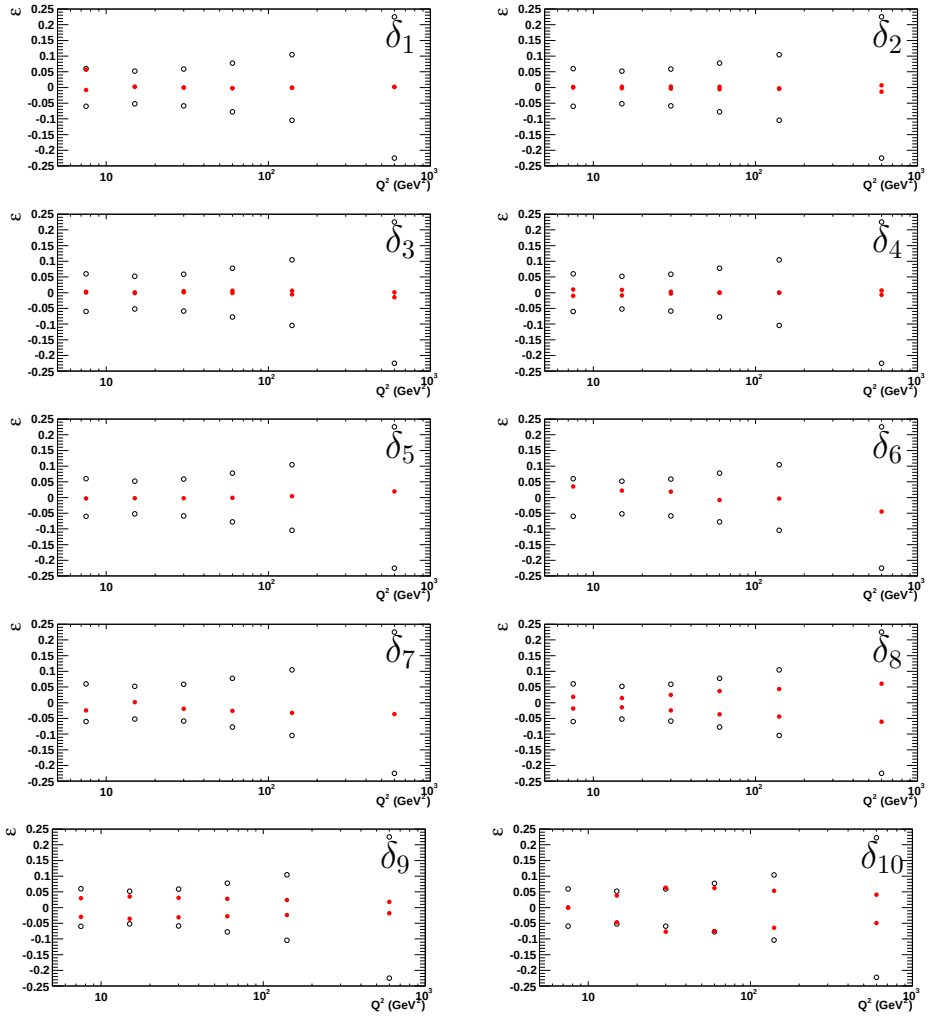


**Figure B.6:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of $Q^2$.*

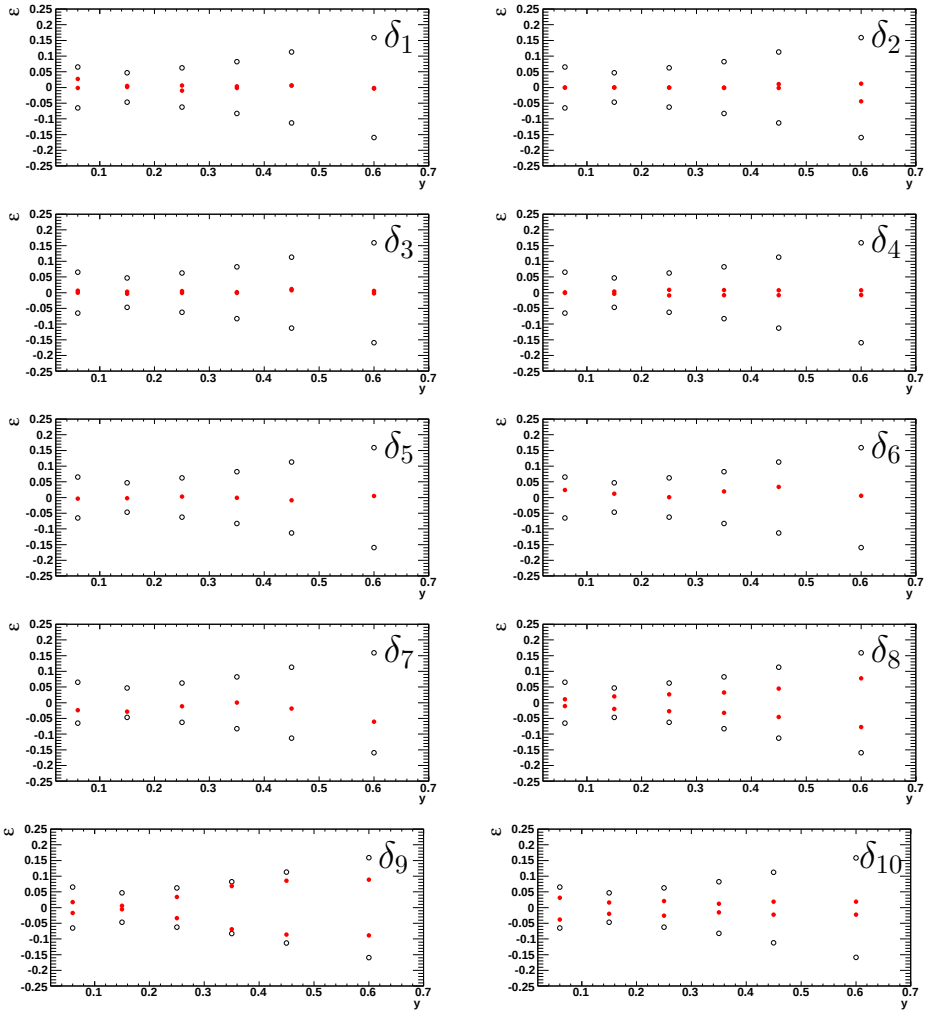


**Figure B.7:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of $y$.*

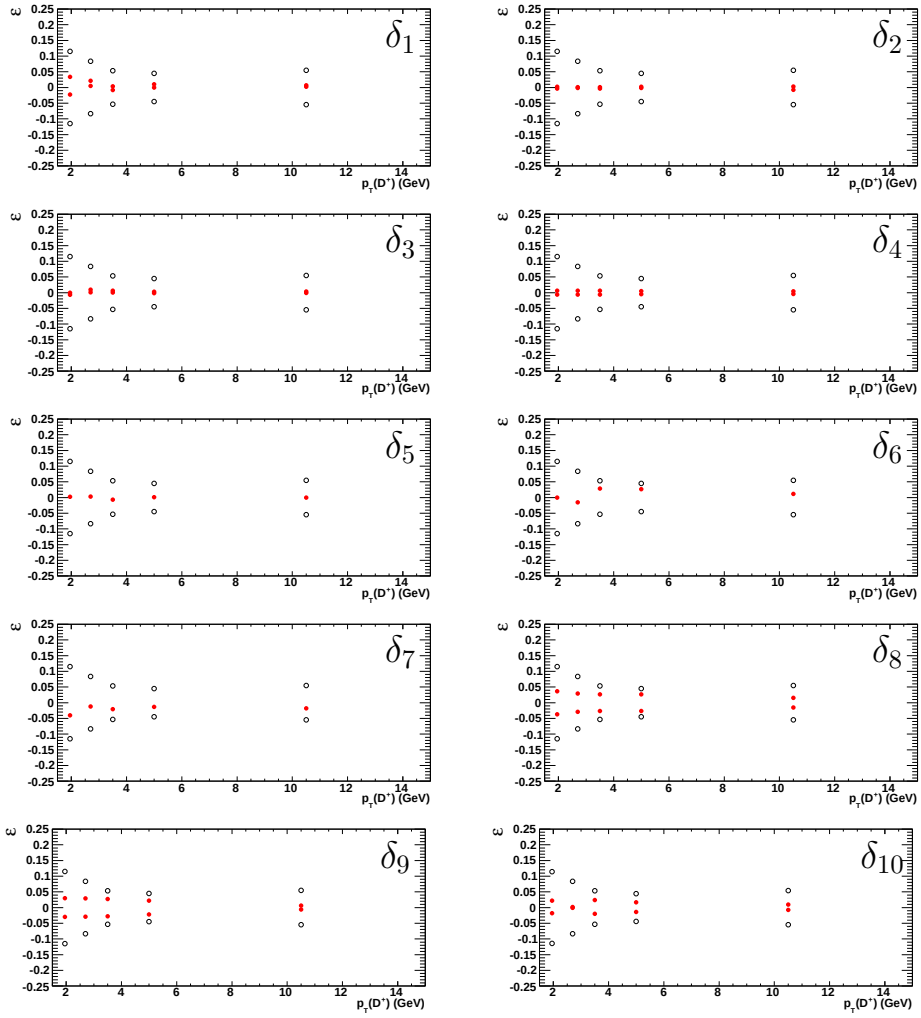


**Figure B.8:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of* $p_{\mathrm{T}}(D^+)$.

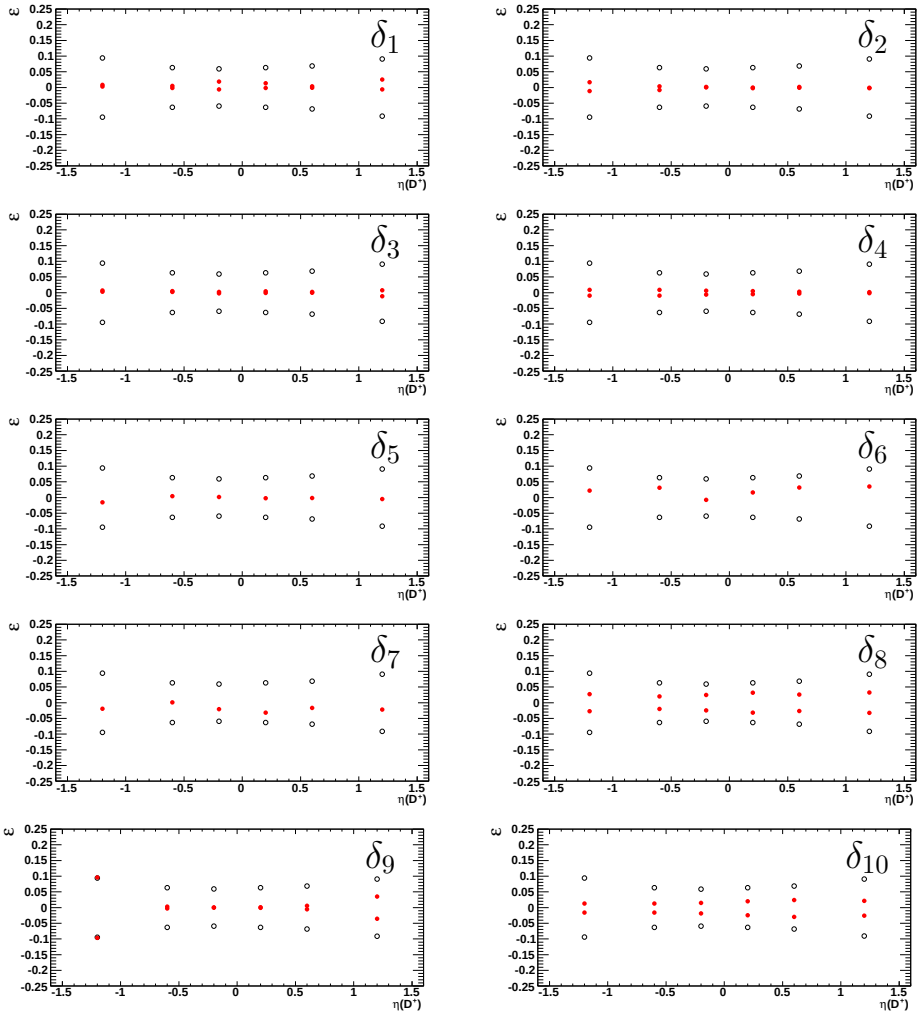


**Figure B.9:** *Relative size of the individual systematic uncertainties (filled red circles) are compared with the relative statistical uncertainty in each bin (open black circles) as a function of* $\eta(D^+)$.

# Appendix C

# Mass spectra fits in data and Monte Carlo

In this appendix the invariant mass spectra for data and Monte Carlo are presented, which are used in the single and double differential cross section calculation. Both charm and beauty Monte Carlo fits are presented separately.

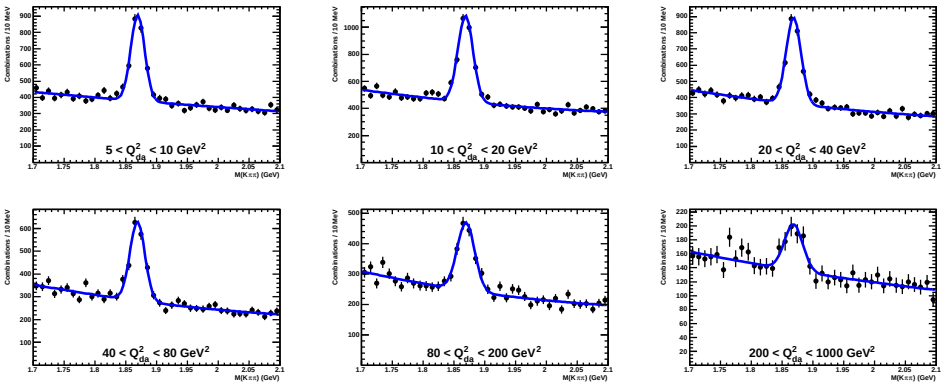


**Figure C.1:** *The reconstructed invariant mass of $D^+$ candidates in the data (black dots) in bins of $Q^2_{da}$. The fit output is also shown (solid blue line).*

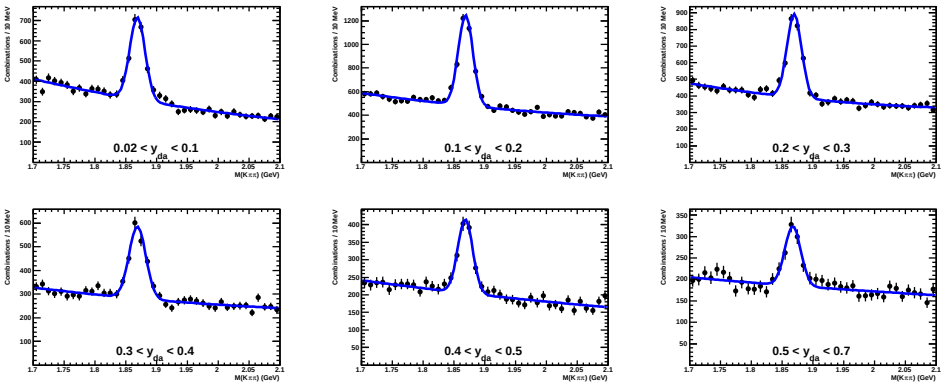


**Figure C.2:** *The reconstructed invariant mass of $D^+$ candidates in the data (black dots) in bins of $y_{da}$. The fit output is also shown (solid blue line).*

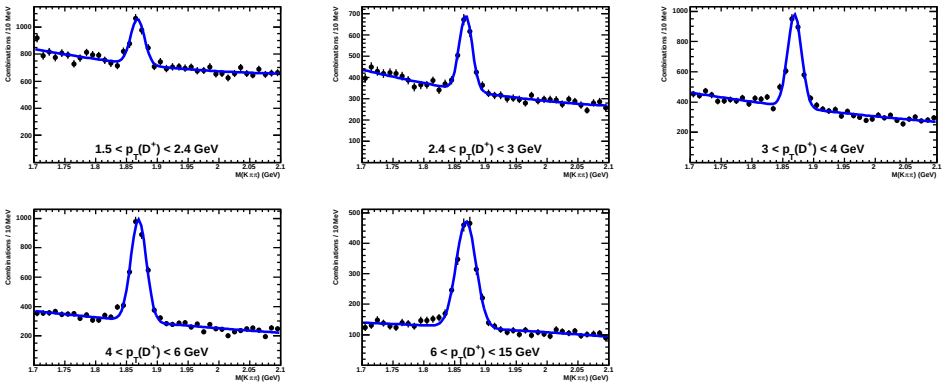


**Figure C.3:** *The reconstructed invariant mass of $D^+$ candidates in the data (black dots) in bins of $p_{\mathrm{T}}(D^+)$. The fit output is also shown (solid blue line).*

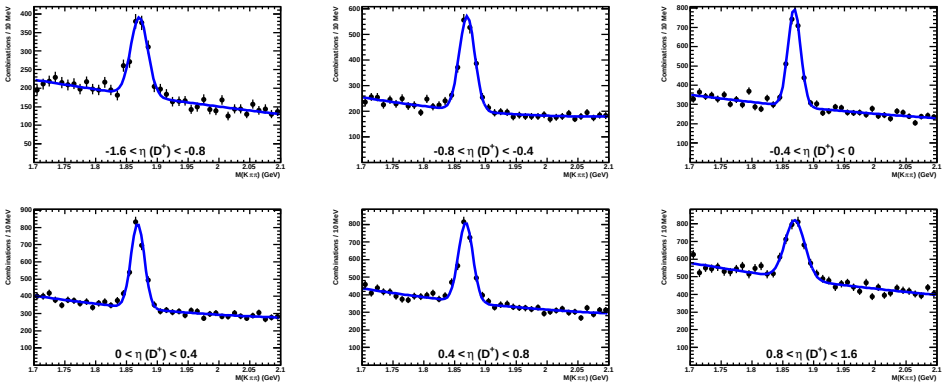


**Figure C.4:** *The reconstructed invariant mass of $D^+$ candidates in the data (black dots) in bins of $\eta(D^+)$. The fit output is also shown (solid blue line).*

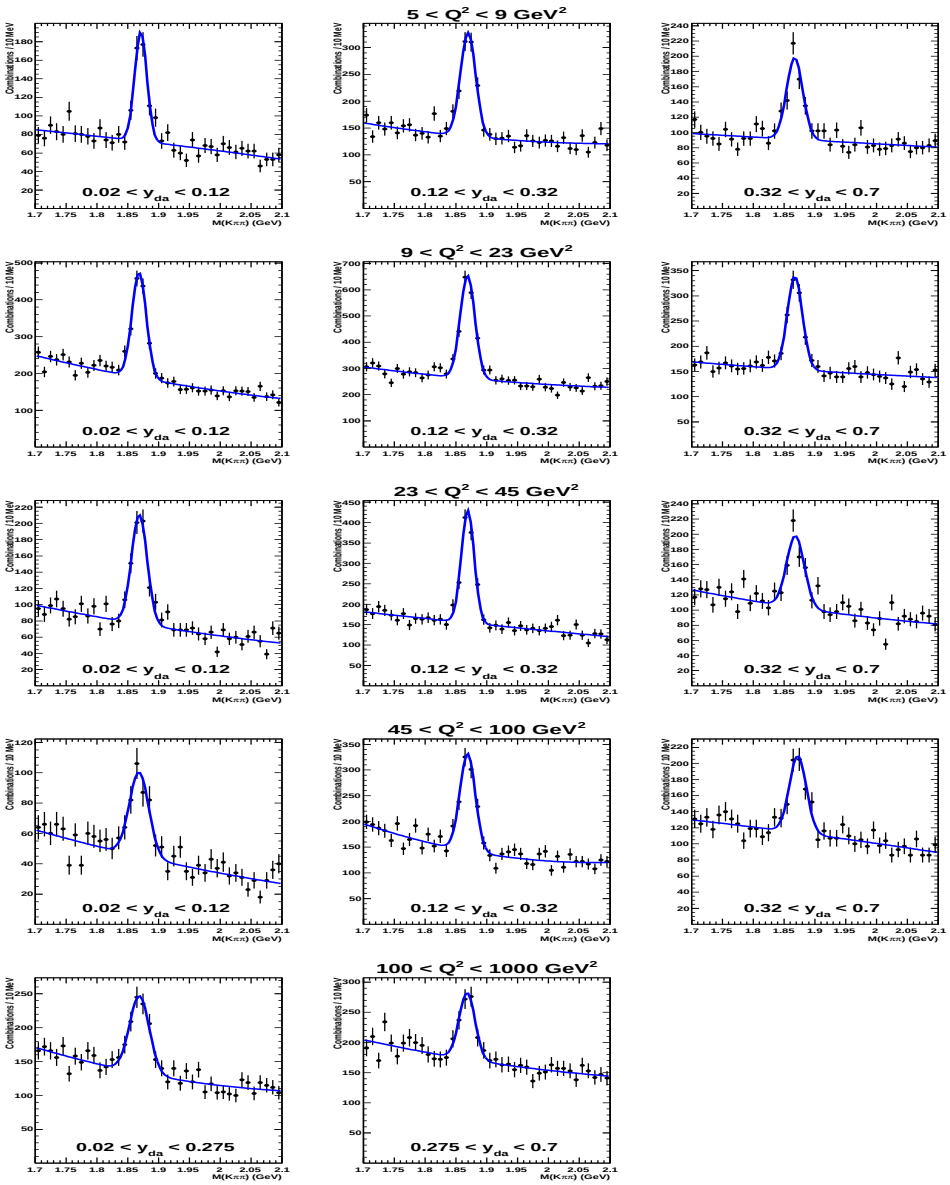


**Figure C.5:** *The reconstructed invariant mass of $D^+$ candidates in the data (black dots) in bins of $y_{da}$ for various ranges of $Q^2_{da}$. The fit output is also shown (solid blue line).*

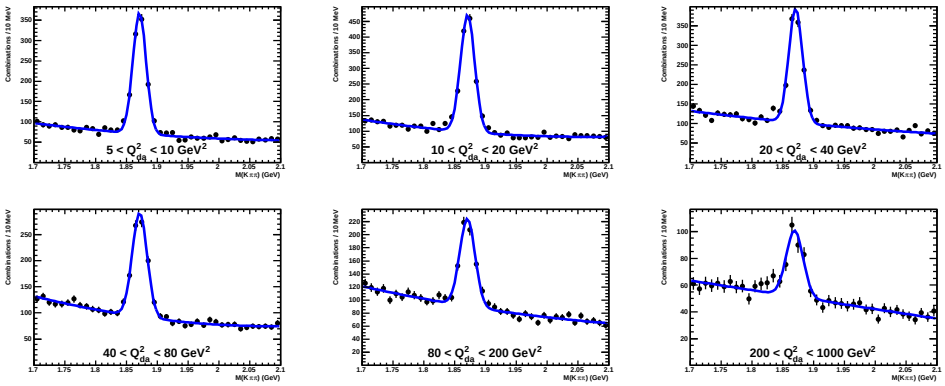


**Figure C.6:** *The reconstructed invariant mass of $D^+$ candidates in the charm Monte Carlo sample (black dots) in bins of $Q^2_{da}$. The fit output is also shown (solid blue line).*

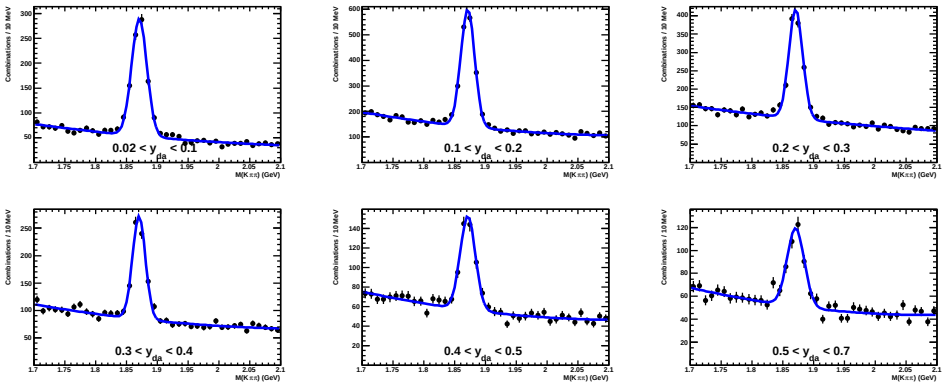


**Figure C.7:** *The reconstructed invariant mass of $D^+$ candidates in the charm Monte Carlo sample (black dots) in bins of $y_{da}$. The fit output is also shown (solid blue line).*

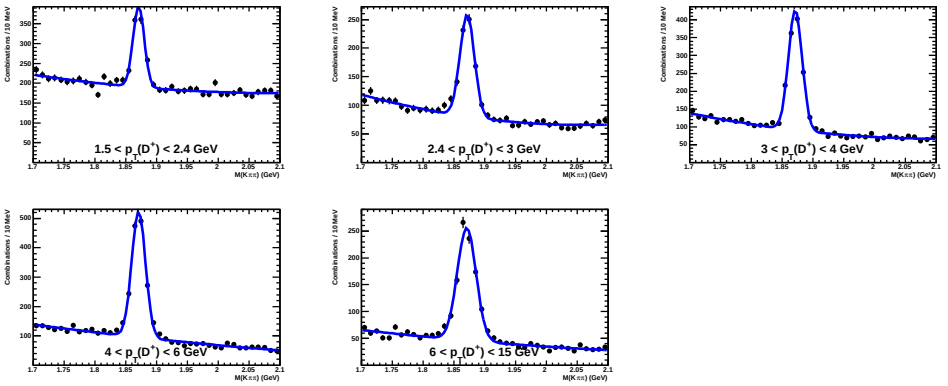


**Figure C.8:** *The reconstructed invariant mass of $D^+$ candidates in the charm Monte Carlo sample (black dots) in bins of $p_{\mathrm{T}}(D^+)$. The fit output is also shown (solid blue line).*

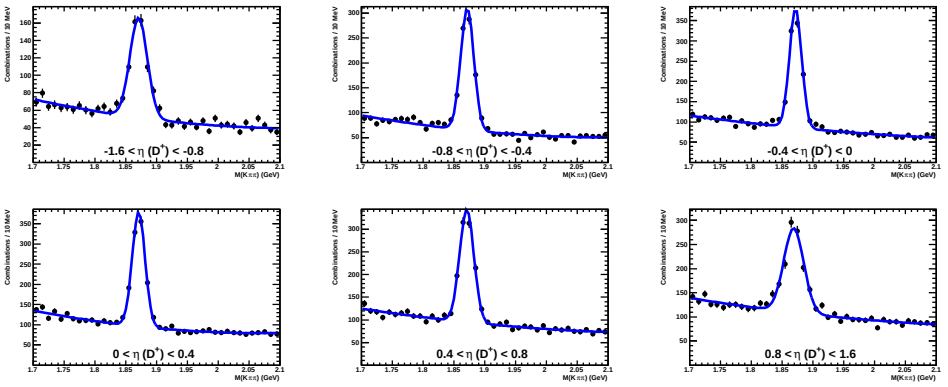


**Figure C.9:** *The reconstructed invariant mass of $D^+$ candidates in the charm Monte Carlo sample (black dots) in bins of $\eta(D^+)$. The fit output is also shown (solid blue line).*

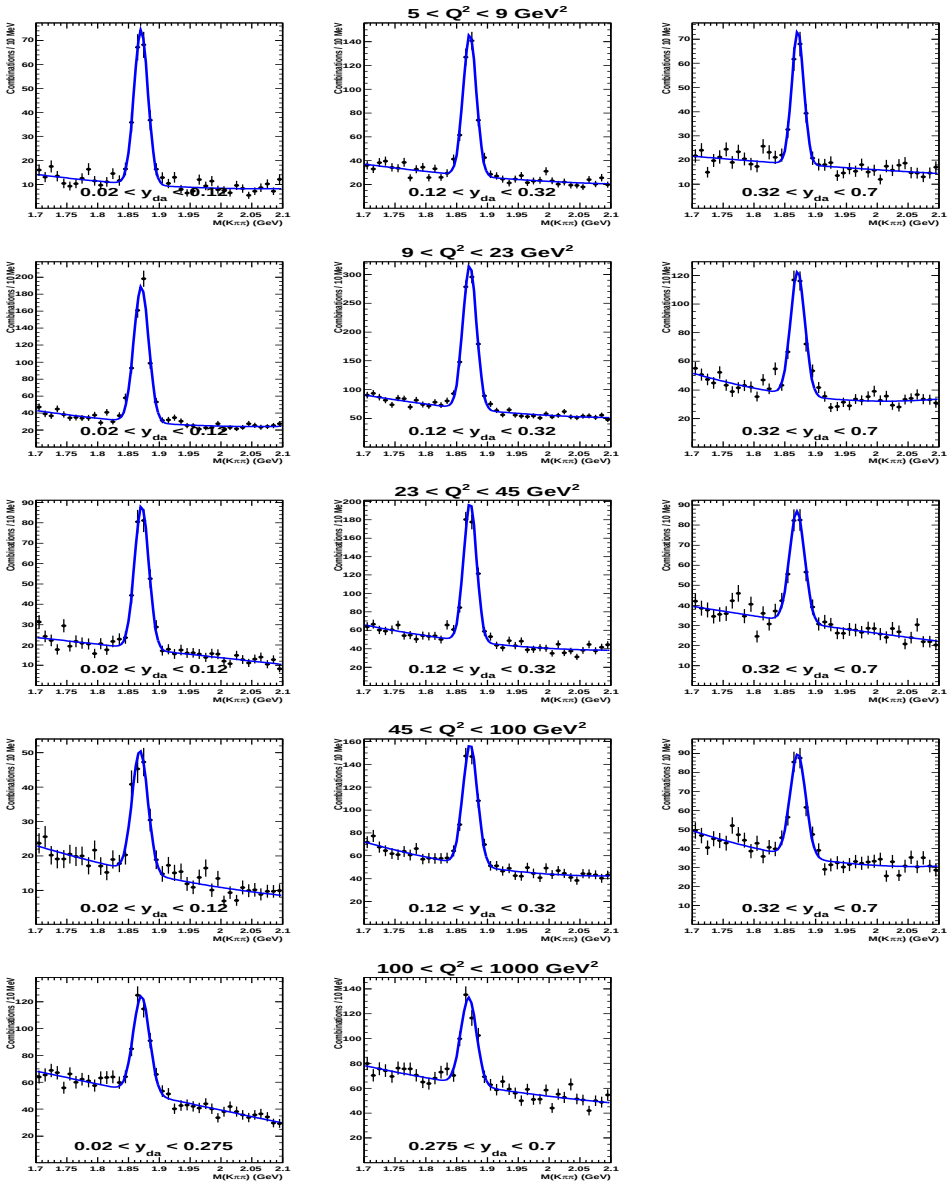


**Figure C.10:** *The reconstructed invariant mass of $D^+$ candidates in the charm Monte Carlo sample (black dots) in bins of $y_{da}$ for various ranges of $Q^2_{da}$. The fit output is also shown (solid blue line).*

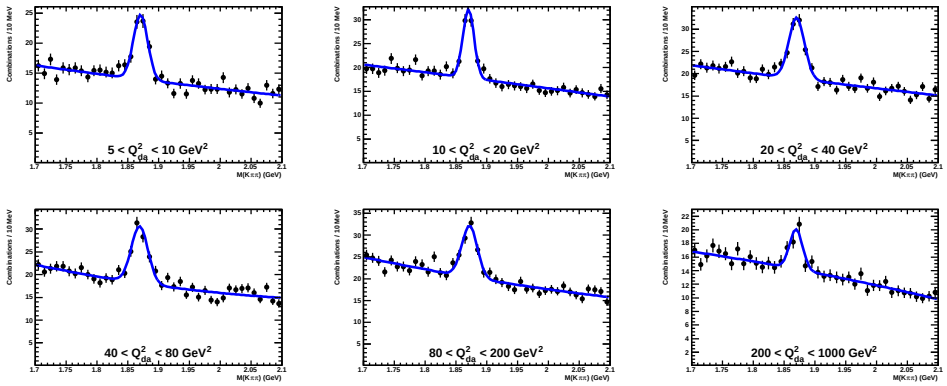


**Figure C.11:** *The reconstructed invariant mass of $D^+$ candidates in the beauty Monte Carlo sample (black dots) in bins of $Q^2_{da}$. The fit output is also shown (solid blue line).*

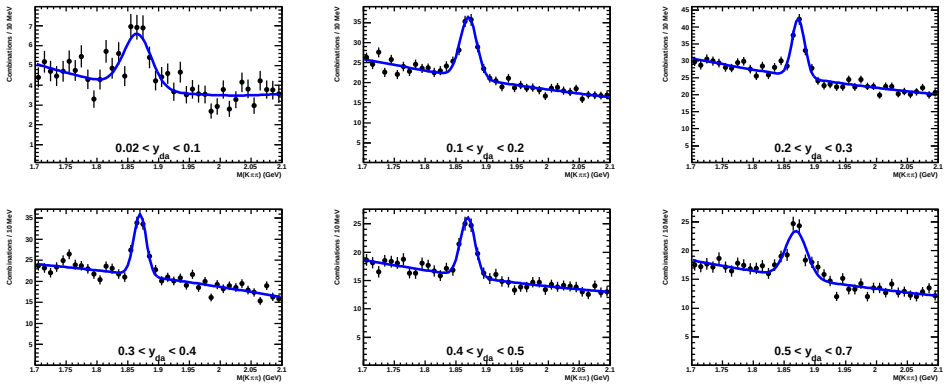


**Figure C.12:** *The reconstructed invariant mass of $D^+$ candidates in the beauty Monte Carlo sample (black dots) in bins of $y_{da}$. The fit output is also shown (solid blue line).*

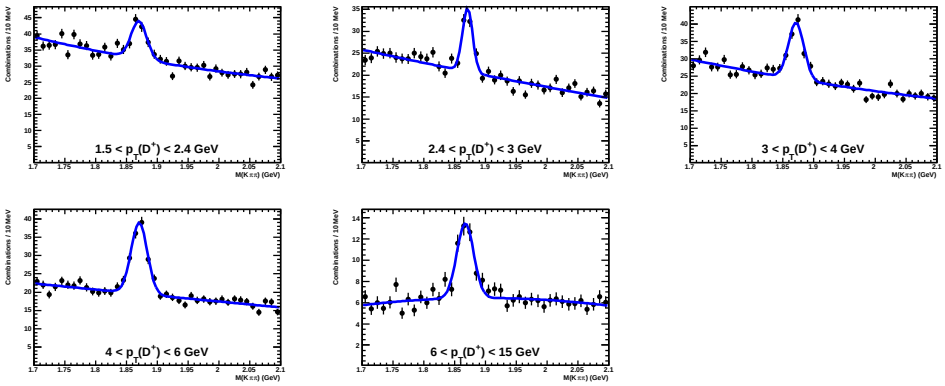


**Figure C.13:** *The reconstructed invariant mass of $D^+$ candidates in the beauty Monte Carlo sample (black dots) in bins of $p_{\mathrm{T}}(D^+)$. The fit output is also shown (solid blue line).*

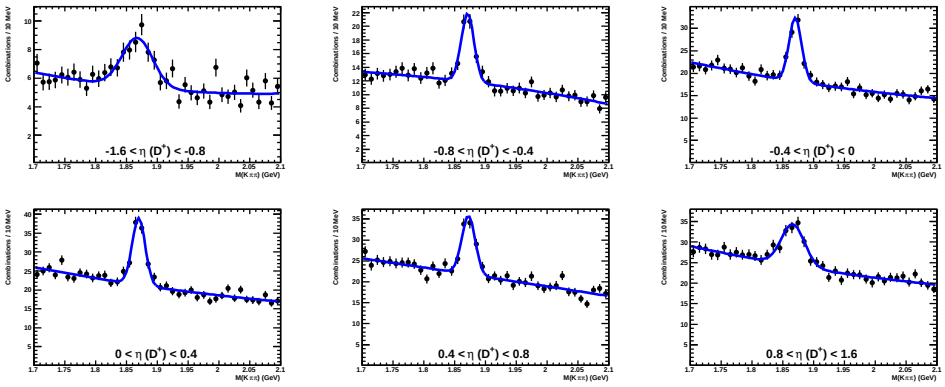


**Figure C.14:** *The reconstructed invariant mass of $D^+$ candidates in the beauty Monte Carlo sample (black dots) in bins of $\eta(D^+)$. The fit output is also shown (solid blue line).*

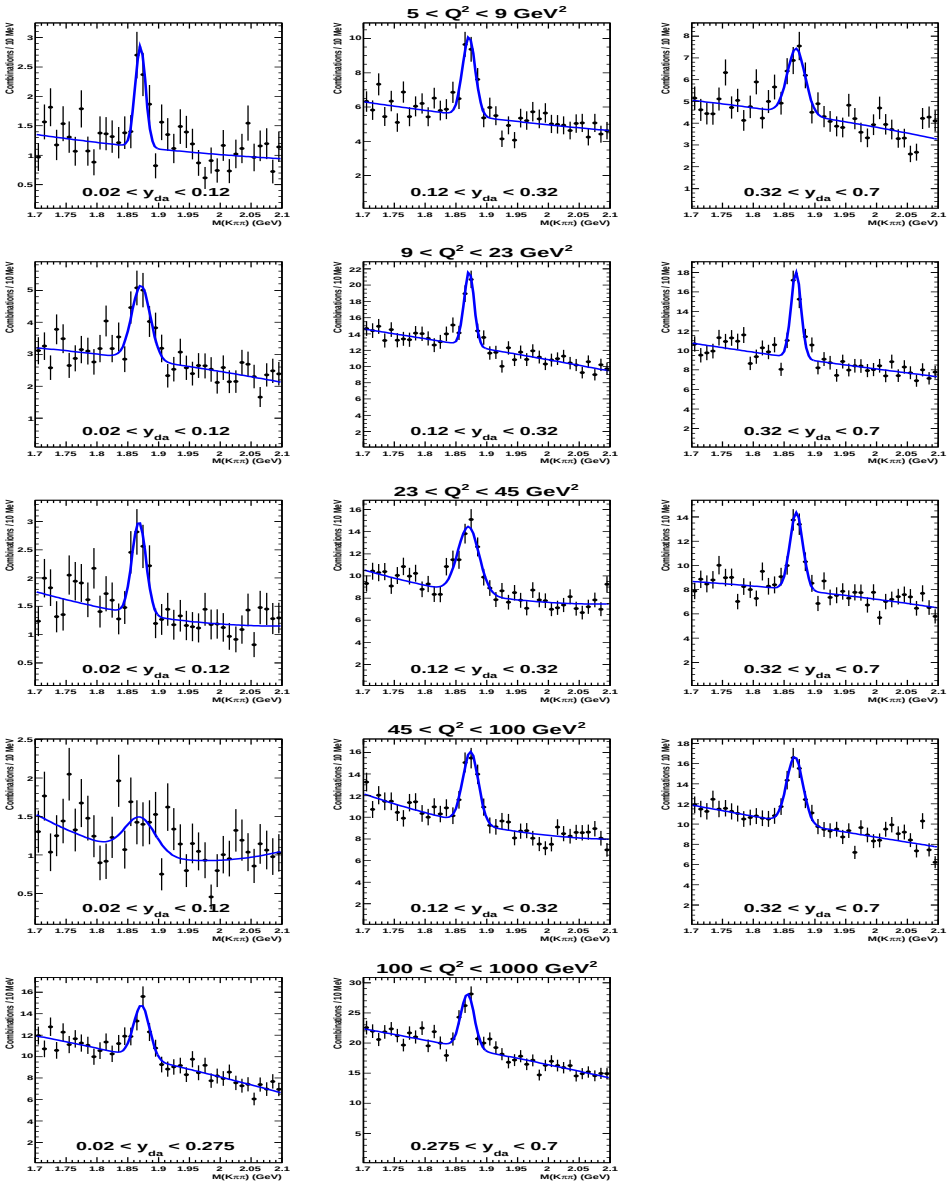


**Figure C.15:** *The reconstructed invariant mass of $D^+$ candidates in the beauty Monte Carlo sample (black dots) in bins of $y_{da}$ for various ranges of $Q^2_{da}$. The fit output is also shown (solid blue line).*

# References


[1] ZEUS Collaboration; S. Chekanov et al., *Measurement of $D^{\pm}$ and $D^0$ Production in Deep Inelastic Scattering Using a Lifetime Tag at HERA*. Eur. Phys. J. **C 63**, 171 (2009).

[2] ALEPH, DELPHI, L3 and OPAL Coll., *Search for the Standard Model Higgs boson at LEP*. Phys. Lett. **B 565**, 61 (2003).

[3] CDF and D0 Collaborations, *Combined CDF and D0 upper limits on Standard Model Higgs boson production with up to* $8.2\,fb^{-1}$ *of data*. Preprint hep-ex/1103.3233, 2011.

[4] ATLAS Collaboration, *Combined Standard Model Higgs boson searches in pp collisions at* $\sqrt{s} = 7\ TeV$ *with the ATLAS Experiment at the LHC* (unpublished). ATLAS-CONF-2011-112, 2011.

[5] CMS Collaboration, *Search for Standard Model Higgs boson in pp collisions at* $\sqrt{s} = 7\ TeV$ (unpublished). CMS-PAS-HIG-11-011, 2011.

[6] KamLAND Coll., A. Abe et al., *Precision measurement of neutrino oscillation parameters with KamLAND*. Phys. Rev. Lett. **100**, 221803 (2008). ArXiv:hep-ex/0801.4589;
SNO Coll., B. Aharmim et al., *Low-energy-threshold analysis of the Phase I and Phase II data sets of the Sudbury Neutrino Observatory*. Phys. Rev. **C81**, 055504 (2010). ArXiv:nucl-ex/0910.2984;
MINOS Coll., P. Adamson et al., *Measurement of the neutrino mass splitting and flavor mixing by MINOS*. Phys. Rev. Lett. **106**, 181801 (2011). ArXiv:hep-ex/1103.0340.

[7] G. Bertone, D. Hooper, J. Silk, *Particle dark matter: evidence, candidates and constrains*. Phys. Rep. **405**, 279 (2005).

[8] D. N. Spergel et al., *Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology*. Astrophys. J. Suppl. **170**, 377 (2007).

[9] D. Alves et al., *Simplified models for LHC new physics searches* (unpublished). An official summary of results from "Topologies for early LHC searches" workshop (SLAC, September 2010). ArXiv:hep-ph/1105.2838, 2011.

[10] K. Nakamura et al.; Particle Data Group, *The review of particle physics.* J. Phys. **G 37**, 075021 (2010).

[11] R.P. Feynman, *Very High-Energy Collisions of Hadrons.* Phys. Rev. Lett. **23**, 1415 (1969).

[12] J.D. Bjorken, *Asymptotic Sum Rules at Infinite Momentum.* Phys. Rev. **179**, 1547 (1969).

[13] G. Miller et al., *Inelastic Electron–Proton Scattering at Large Momentum Transfers.* Phys. Rev. **D 5**, 528 (1972).

[14] M. Gell-Mann, *A Schematic Model of Baryons and Mesons.* Phys. Lett. **8**, 214 (1964).

[15] T. Eichten et al., *Measurement Of The Neutrino - Nucleon Anti-Neutrino - Nucleon Total Cross-Sections.* Phys. Lett. **B 46**, 274 (1973).

[16] R. Brandelik et al., *Evidence for Planar Events in $e^+e^-$ Annihilation at High–Energies.* Phys. Lett. **B 86**, 243 (1979).

[17] D.J. Fox et al., *Test of scale invariance in high-energy muon scattering.* Phys. Rev. Lett. **33**, 1504 (1974).

[18] R. K. Ellis, F. Hautmann, B. R. Webb, *QCD scaling violation at small $x$.* Phys. Lett. **348**, 582 (1995).

[19] ZEUS Coll., M. Derrick et al., *Measurement of the Proton Structure Function $F_2$ in ep Scattering at HERA.* Phys. Lett. **B 316**, 412 (1993);
H1 Coll., I. Abt et al., *Measurement of the Proton Structure Function $F_2(x, Q^2)$ in the Low $x$ Region at HERA.* Nucl. Phys. **B 407**, 515 (1993).

[20] ZEUS, H1 Collaborations, F.D.Aaron et al., *Combined measurement and QCD analysis of the inclusive $e^{\pm}p$ cross sections at HERA.* JHEP **2010**, 1 (2010). ArXiv:0911.0884; doi: 10.1007/JHEP01(2010)109.

[21] R.K. Ellis et al., *Perturbation Theory and the Parton Model in QCD.* Nucl. Phys. **B 152**, 285 (1979).

[22] V.N. Gribov and L.N. Lipatov, *Deep Inelastic ep Scattering in Perturbation Theory.* Sov. J. Nucl. Phys. **15**, 438 (1972);
G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language.* Nucl. Phys. **B 126**, 298 (1977);
Yu.L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics [in Russian].* Sov. Phys. JETP **46**, 641 (1977).

[23] H1 Collaboration; F. D. Aaron et al., *Measurement of the D* Meson Production Cross Section and $F_2^{c\bar{c}}$, at High $Q^2$, in ep Scattering at HERA.* Phys. Lett. **B686**, 91 (2010), hep-ex/0911.3989.

[24] H1 Coll.; F. D. Aaron et al., *Measurement of $D^{*\pm}$ meson production and determination of $F_2^{c\bar{c}}$ at low $Q^2$ in Deep-Inelastic Scattering at HERA*, redreport DESY-11-066, DESY, 2011. Submitted to EPJ C.

[25] ZEUS Coll., J. Breitweg et al., *Measurement of $D^{*\pm}$ Production and the Charm Contribution to $F_2$ in Deep Inelastic Scattering at HERA.* Eur. Phys. J. **C 12**, 35 (2000), hep-ex/9908012.

[26] ZEUS Coll., S. Chekanov et al., *Measurement of $D^{*\pm}$ Production in Deep Inelastic $e^{\pm}p$ Scattering at DESY HERA.* Phys. Rev. **D 69**, 012004 (2004), hep-ex/0308068.

[27] H1 Collaboration; Aktas, A. et al., *Production of $D^{*\pm}$ mesons with dijets in deep-inelastic scattering at HERA.* Eur. Phys. J. **C51**, 271 (2007), hep-ex/0701023.

[28] ZEUS Coll., S.Chekanov et al., *Measurement of Charm and Beauty Production in Deep Inelastic ep Scattering from Decays into Muons at HERA.* Eur. Phys. J. **65**, 65 (2010). ArXiv: 0904.3487.

[29] H1 Collaboration; A. Aktas et al., *Measurement of $F_2^c$ and $F_2^b$ at low $Q^2$ and x using the H1 vertex detector at HERA.* Eur. Phys. J. **C45**, 23 (2006), hep-ex/0507081.

[30] H1 Collaboration; F. D. Aaron et al., *Measurement of charm and beauty jets in deep inelastic scattering at HERA.* Eur. Phys. J. **C71**, 1509 (2011), hep-ex/1008.1731.

[31] K.Daum on behalf of the H1 and ZEUS Coll., *Combined $F_2^c$ measurement at HERA*. Proceedings of the XVIII International Workshop on Deep-Inelastic Scattering and Related Subjects, Florence, 19–23 April 2010. PoS(DIS 2010)153, 2010, available on `http://pos.sissa.it/archive/conferences/106/153/DIS%202010_153.pdf`. ZEUS-prel-09-015, H1prelim-09-171.

[32] H1 Collaboration; F. D. Aaron et al., *Measurement of the charm and beauty structure functions using the H1 vertex detector at HERA*. Eur. Phys. J. **C65**, 89 (2010), hep-ex/0907.2643.

[33] ZEUS Collaboration; S. Chekanov et al., *Measurement of heavy-quark jet photoproduction at HERA*, redreport DESY-11-067, DESY, 2011. To be submitted to EPJ C.

[34] CMS Collaboration, *Algorithms for b jet identification in CMS*, Technical Report CMS PAS BTV-09-001, CERN, 2009.

[35] P. Roloff, *Measurement of charm and beauty production in deep inelastic scattering at HERA*. Dissertation, Universitaet Hamburg, Report Thesis in preparation, 2011.

[36] ZEUS and H1 Collaborations, *QCD analysis using charm data* (unpublished). ZEUS-prel-10-009, H1prelim-10-045, 2010.

[37] ZEUS and H1 Collaborations, *Charm mass parameter in the QCD analysis and implications for the LHC* (unpublished). ZEUS-prel-10-019, H1prelim-10-143, 2010.

[38] B.W. Harris and J. Smith, *Charm Quark and $D^{*\pm}$ Cross Sections in Deeply Inelastic Scattering at DESY HERA*. Phys. Rev. **D 57**, 2806 (1998).

[39] B.W. Harris and J. Smith, *Heavy-Quark Correlations in Deep-Inelastic Electroproduction*. Nucl. Phys. **B 452**, 109 (1995).

[40] A. Martin et al., *Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions*. Eur. Phys. J. **C 70**, 51 (2010). ArXiv:hep-ph/1007.2624.

[41] ZEUS Coll., S. Chekanov et al., *ZEUS next-to-leading-order QCD analysis of data on deep inelastic scattering*. Phys. Rev. **D 67**, 012007 (2003), hep-ex/0208023.

[42] M. Cacciari, P. Nason and C. Oleari, *A Study of heavy flavored meson fragmentation functions in $e^+e^-$ annihilation*. JHEP **0604**, 006 (2006), hep-ph/0510032.

[43] Belle Coll., R. Seuster et al., *Charm hadrons from fragmentation and B decays in $e^+e^-$ annnihilation at $\sqrt{s} = 10.6\,GeV$*. Phys. Rev. **D 73**, 032002 (2006).

[44] CLEO Coll., M. Artuso et al., *Charm meson spectra in $e^+e^-$ annnihilation at 10.5 GeV center of mass energy*. Phys. Rev. **D 70**, 112001 (2004).

[45] ALEPH Coll., A. Heister et al., *Study of the fragmentation of b quarks into B mesons at the Z peak*. Phys. Lett. **B 512**, 30 (2001).

[46] ZEUS Collaboration; S. Chekanov et al., *Measurement of the charm fragmentation function in D photoproduction at HERA*. JHEP **04**, 082 (2009). Redreport DESY-08-209.

[47] H1 Coll., F. D. Aaron et al., *Study of charm fragmentation into $D^{*\pm}$ mesons in deep inelastic scattering at HERA*. Eur. Phys. J. **C59**, 589 (2009), hep-ex/0808.1003.

[48] C. Peterson et al., *Scaling Violations in Inclusive $e^+e^-$ Annihilation Spectra*. Phys. Rev. **D 27**, 105 (1983).

[49] V. G. Kartvelishvili, A. K. Likhoded, V. A. Petrov, *On the fragmentation functions of heavy quarks into hadrons*. Phys. Lett. **B78**, 615 (1983).

[50] Erich Lohrmann, *Summary of charm fractions* (unpublished), 2011, available on `https://www.desy.de/h1zeus/heavy-flavour/internal/documents/CharmFragmentationFraction.pdf`. ZEUS and H1 only.

[51] *A proposal for a large electron-proton colliding beam facility at DESY*. DESY–HERA 81–10, 1981.

[52] ZEUS Coll., U. Holm (ed.), *The ZEUS Detector*. Status Report (unpublished), DESY (1993), available on `http://www-zeus.desy.de/bluebook/bluebook.html`.

[53] H1 Coll., I. Abt et al., *The H1 Detector at HERA*. Nucl. Inst. Meth. **A 386**, 310 (1997);
H1 Coll., I. Abt et al., *The tracking, calorimeter and muon detectors of the H1 experiment*. Nucl. Inst. Meth. **A 386**, 348 (1997).

[54] K. Ackerstaff et al., *The HERMES Spectrometer*. Nucl. Inst. Meth. **A 417**, 230 (1998), hep-ex/9806008.

[55] HERA-B Coll., *HERA–B: An Experiment to Study CP Violation in the B System Using an Internal Target at the HERA Proton Ring*, Design Report DESY-PRC 95/01, DESY, 1995.

[56] Available on `http://www-zeus.desy.de/~vschoenb/`.

[57] Available on `http://www-zeus.desy.de/physics/lumi/`.

[58] Available on `http://www.nikhef.nl/pub/departments/mt/projects/zeus/vertex/website.html`.

[59] A. Polini et al., *The design and performance of the ZEUS microvertex detector*. Nucl. Inst. Meth. **A 581**, 656 (2007), arXiv:0708.3011.

[60] N. Harnew et al., *Vertex Triggering Using Time Difference Measurements in the ZEUS Central Tracking Detector*. Nucl. Inst. Meth. **A 279**, 290 (1989);
B. Foster et al., *The Performance of the ZEUS Central Tracking Detector z-by-Timing Electronics in a Transputer Based Data Acquisition System*. Nucl. Phys. Proc. Suppl. **B 32**, 181 (1993);
B. Foster et al., *The Design and Construction of the ZEUS Central Tracking Detector*. Nucl. Inst. Meth. **A 338**, 254 (1994).

[61] M. Derrick et al., *Design and Construction of the ZEUS Barrel Calorimeter*. Nucl. Inst. Meth. **A 309**, 77 (1991);
A. Andresen et al., *Construction and Beam Test of the ZEUS Forward and Rear Calorimeter*. Nucl. Inst. Meth. **A 309**, 101 (1991);
A. Caldwell et al., *Design and Implementation of a High-Precision Readout System for the ZEUS Calorimeter*. Nucl. Inst. Meth. **A 321**, 356 (1992);
A. Bernstein et al., *Beam Tests of the ZEUS Barrel Calorimeter*. Nucl. Inst. Meth. **A 336**, 23 (1993).

[62] A. Bamberger et al., *The Small Angle Rear Tracking Detector of ZEUS*. Nucl. Inst. Meth. **A 401**, 63 (1997).

[63] M. Helbich et al., *The Spectrometer System for Measuring ZEUS Luminosity at HERA*. Nucl. Inst. Meth. **A 565**, 572 (2006), physics/0512153.

[64] Available on `https://www-zeus-data.desy.de/tracking/ZEUS_ONLY/info/resolution.html`. ZEUS only.

[65] R. Hall-Wilton et al., *The CTD Tracking Resolution* (unpublished). ZEUS-99-024, internal ZEUS-note, 1999.

[66] Available on `http://www.desy.de/~cflt/info/NIM/figures.html`.

[67] J. Andruszków et al., *Luminosity Measurement in the ZEUS Experiment*. Acta Phys. Pol. **B 32**, 2025 (2001).

[68] T. Haas, V. Makarenko, *Precision calculation of processes used for luminosity measurement at the ZEUS experiment*. Eur. Phys. J. **C71**, 1574 (2011), hep-ph/1009.2451.

[69] V. Drugakov, *Luminosity measurement* (unpublished). Talk at ZEUS Analysis Forum, 20 May 2010. ZEUS only.

[70] S. M. Fisher, P. Palazzi, *Using a data model from software design to data analysis: What have we learned?* Comp. Phys. Comm. **57**, 169 (1989).

[71] N. Metropolis, S. Ulam, *The Monte Carlo method*. "Jour. of the Amer. Stat. Assoc." **44**, 335 (1949).

[72] T. Toll, *MC@NLO for heavy flavour photoproduction at HERA*. Dissertation, Universität Hamburg, Report DESY-THESIS-2010-004, 2010, available on `http://www-h1.desy.de/psfiles/theses/h1th-519.pdf`.

[73] V. Schönberg, *Measurement of beauty and charm photoproduction using inclusive secondary vertexing with the ZEUS detector at HERA*. Dissertation, Universität Bonn, Report Bonn-IR-2010-05, 2010, available on `http://hss.ulb.uni-bonn.de/2010/2127/2127.pdf`.

[74] L. Lönnblad, ARIADNE *Version 4 – A Program for Simulation of QCD Cascades Implementing the Colour Dipole Model*. Comp. Phys. Comm. **71**, 15 (1992).

[75] H. Spiesberger, HERACLES *and* DJANGOH*: Event Generation for ep Interactions at HERA Including Radiative Processes*, 2005, available on `http://wwwthep.physik.uni-mainz.de/~hspiesb/djangoh/djangoh.html`. MZ-TH/05-15.

[76] G.A. Schuler and H. Spiesberger, DJANGO*: The Interface for the Event Generators* HERACLES *and* LEPTO, *Proc. Workshop on Physics at HERA*, W. Buchmüller and G. Ingelman (eds.), Vol. 3, p. 1419. Hamburg, Germany, DESY (1991).

[77] G. Ingelman, A. Edin and J. Rathsman, LEPTO 6.5*: A Monte Carlo Generator for Deep Inelastic Lepton–Nucleon Scattering*. Comp. Phys. Comm. **101**, 108 (1997).

[78] A. Kwiatkowski, H. Spiesberger and H.-J. Möhring, HERACLES*: An Event Generator for ep Interactions at HERA Energies Including Radiative Processes*. Comp. Phys. Comm. **69**, 155 (1992). Also in *Proc. Workshop Physics at HERA*, eds. W. Buchmüller and G.Ingelman, (DESY, Hamburg, 1991).

[79] T. Sjöstrand et al., *High-Energy-Physics Event Generation with* PYTHIA *6.1*. Comp. Phys. Comm. **135**, 238 (2001), hep-ph/0010017.

[80] T. Sjöstrand, S. Mrenna, P. Skands, PYTHIA 6.4 *physics and manual*. JHEP **2006**, 581 (2006). ArXiv:hep-ph/0603175.

[81] H. Jung, *Hard Diffractive Scattering in High Energy ep Collisions and the Monte Carlo Generator* RAPGAP. Comp. Phys. Comm. **86**, 147 (1995).

[82] *EGEE portal: Enabling Grids for E–sciencE*, available on `http://www.eu-egee.org/`.

[83] A. Gellrich et al., *Installing and operating a GRID infrastructure at DESY*. Proceedings of Computing in High Energy Physics and Nuclear Physics, Interlaken (Switzerland), 1 October 2004, 2004. CERN-2005-002, Vol. 2.

[84] T. Haas, *Generating Monte Carlo events with MOZART* (unpublished). ZEUS-92-021, internal ZEUS note, 1992.

[85] R. Brun et al., GEANT3, Technical Report CERN-DD/EE/84-1, CERN, 1987.

[86] A. Geiser, *Trigger and rate studies with inclusive low $Q^2$ DIS* (unpublished). Talk at ZEUS NC $e^+p$ task force meeting, 08 September 2009. ZEUS only.

[87] *CTD DQM Archive* (unpublished), available on `http://www-zeus.desy.de/~tcbol/archive.php3`. ZEUS only.

[88] A. Parenti, *2009 Guide to Funnel: The ZEUS Monte Carlo Production Facility* (unpublished). ZEUS-09-002, internal ZEUS note, 2009.

[89] dCache Team, P. Fuhrmann et al., *dCache, the Overview*, 2005, available on `http://www.dcache.org/manuals/dcache-whitepaper-light.pdf`; dCache Team, P. Fuhrmann et al., *The dCache Book*, available on `http://www.dcache.org/manuals/Book-1.9.11/`.

[90] Fabrizio Pacini et al., *Job Description Language attributes specification*, available on `https://edms.cern.ch/file/590869/1/WMS-JDL.pdf`.

[91] *Berkeley Database Information Index V5*, available on `https://twiki.cern.ch/twiki/bin/view/EGEE/BDII`.

[92] Stephen Burke et al., *File and replica management client tools*, in *gLite 3.2 User Guide*, pp. 92–99. WLCG, 2011. CERN–LCG–GDEIS–722398.

[93] Stephen Burke et al., *File Catalogue in gLite 3.2*, in *gLite 3.2 User Guide*, pp. 87–92. WLCG, 2011. CERN–LCG–GDEIS–722398.

[94] Domenico Vicinanza, *SAM users guide*, available on `http://sam-docs.web.cern.ch/sam-docs/docs/MANUserManual.pdf`.

[95] R. E. Kalman, *A new approach to linear filtering and prediction problems.* "Jour. of Basic Engineering" **D 82**, 35 (1960); R. Frühwirth, *Application of Kalman Filtering to Track and Vertex Fitting.* Nucl. Inst. Meth. **A 262**, 444 (1987).

[96] G.F. Hartner et al., *VCTRAK(3.07/04): Offline Output Information* (unpublished). ZEUS-97-064, internal ZEUS Note, 1997; G.F. Hartner, *VCTRAK Briefing: Program and Math* (unpublished). ZEUS-98-058, internal ZEUS Note, 1998.

[97] E. Maddox, *A Kalman Filter Track Fit for the ZEUS Microvertex Detector* (unpublished). ZEUS-03-008, internal ZEUS note, 2003.

[98] A. Spiridonov, *Mathematical framework for fast and rigorous track fit for the ZEUS detector* (unpublished). DESY-08-182. Available at arXiv:physics.data-an/0812.1245, 2008.

[99] R. Frühwirth and A. Strandlie, *Track fitting with ambiguities and noise: A study of elastic tracking and nonlinear filters*. Comp. Phys. Comm. **120**, 197 (1999).

[100] R. Mankel, *ZEUS tracking tutorial* (unpublished), available on `http://www-zeus.desy.de/lectures/index.php3`.

[101] Available on `http://www-zeus.desy.de/~dnicho/ZEUS_ONLY/index.php3`. ZEUS only.

[102] H. Stadie, *Global fitting of vertex final states* (unpublished). Talk at the ZEUS Collaboration meeting, October 2006. ZEUS only.

[103] N. Tuning, *ZUFOs: Hadronic Final State Reconstruction with Calorimeter, Tracking and Backsplash Correction* (unpublished). ZEUS-01-021, internal ZEUS Note, 2001.

[104] H. Abramowicz, A. Caldwell and R. Sinkus, *Neural Network Based Electron Identification in the ZEUS Calorimeter*. Nucl. Inst. Meth. **A 365**, 508 (1995).

[105] A. Kappes, *Measurement of $e^- p \to e^- X$ differential cross sections at high $Q^2$ and of the structure function $xF_3$ with ZEUS at HERA*. Ph.D. Thesis, Universität Bonn, Bonn, Germany, Report BONN-IR-01-16, 2001, available on `http://www-zeus.physik.uni-bonn.de/german/phd.html`.

[106] F. Jacquet and A. Blondel, *Detection of the Charged Current Event – Method II*, *Proceedings of the Study for an ep Facility for Europe*, U. Amaldi (ed.), p. 391. Hamburg, Germany (1979). Also in preprint DESY 79/48.

[107] S. Bentvelsen, J. Engelen and P. Kooijman, *Reconstruction of $(x, Q^2)$ and Extraction of Structure Functions in Neutral Current Scattering at HERA*, *Proc. Workshop on Physics at HERA*, W. Buchmüller and G. Ingelman (eds.), Vol. 1, p. 23. Hamburg, Germany, DESY (1992).

[108] A. Geiser, *GR Tracking, Trigger and Physics studies with inclusive low $Q^2$ DIS* (unpublished). Talk at the ZEUS Heavy Flavour group meeting, 01 July 2009. ZEUS only.

[109] F. James, M. Roos, *Minuit: a system for function minimization and analysis of the parameter errors and correlations.* Comp. Phys. Comm. **10**, 343 (1975);
F. James, *Minuit v94.1, CERN Program Library Long Writeup D506* (unpublished), available on
`http://wwwinfo.cern.ch/asdoc/minuit/minmain.html`.

[110] W. Dunne, *Measurements of $D^*$ Production in Deep Inelastic Scattering at HERA.* PhD, University of Glasgow, 2007.

[111] M. Zambrana, *Measurement of D mesons production in Deep Inelastic Scattering at HERA.* PhD, Universidad Autonoma de Madrid, 2007.

[112] S. Shimizu, *Measurement of the longitudinal structure function $F_L$ at HERA.* PhD, University of Tokyo, 2009.

[113] Y. Ri, *Measurement of $e^-p$ neutral-current deep inelastic scattering cross sections with longitudinally polarized electron beams at $\sqrt{s} = 318\,GeV$.* PhD, Tokyo Metropolitan University, 2009.

[114] O. Behnke, *Tracking efficiency tests with $K^0_S$ decays* (unpublished). Talk at ZEUS Heavy Flavour group meeting, 25 March 2010. ZEUS only.

[115] ZEUS Collaboration, *Measurement of charm production with $D^\pm$ in DIS at HERAII* (unpublished). ZEUS-prel-10-005, 2010.

[116] M.A.G. Aivazis et al., *Leptoproduction of heavy quarks I – General formalism and kinematics of charged current and neutral current production processes.* Phys. Rev. **D 50**, 3085 (1994), hep-ph/9312318.

[117] M.A.G. Aivazis et al., *Leptoproduction of heavy quarks II – A unified QCD formulation of charged and neutral current processes from fixed target to collider energies.* Phys. Rev. **D 50**, 3102 (1994), hep-ph/9312319.

[118] R.G. Roberts and R.S. Thorne, *Ordered analysis of heavy flavor production in deep inelastic scattering.* Phys. Rev. **D 57**, 6871 (1998).

[119] ZEUS Collaboration, S. Chekanov et al., *Measurement of D Mesons Production in Deep Inelastic Scattering at HERA*. JHEP **2007**, 074 (2007).

# Acknowledgements

And now, when the thesis is completed, I would like to use this opportunity to thank all people who contributed in various ways. If something or somebody will not be mentioned they will not be forgoten and I hope they will forgive me.

At the first place, I would like to thank my supervisor Dr. Achim Geiser for giving me the opportunity to work on my Ph.D. in ZEUS. Without your guidance, support and advises this work wouldn't have a happy end. I have benefited a lot from your outstanding understanding of the fundamental physics as well as feeling of the "common sense". Moreover, your patience and constructive feedback have made a significant contribution into the structure and the style of this thesis.

I would also like to thank Prof. Robert Klanner, Prof. Johannes Haller, Dr. Tomas Schörner-Sadenius and Dr. Georg Steinbrück for agreeing to participate in the thesis disputation as well as for their interest and time. I would especially like to thank Robert Klanner for reading the manuscript and making useful comments and suggestions. I'm also grateful to Tomas for organising within the "Physics at the Terascale" alliance multiple schools on a broad variety of topics, from which I've learnt a lot.

I'm also grateful to Olaf Behnke for thoroughly reading this thesis and his suggestions. As always, your creative ideas brought my attention to the details that I haven't thought about.

There have been many people, who made their contributions in endless discussions, tones of e-mails and constructive criticism. Most of all I would like to thank Philipp Roloff with whom I started my first analysis in high energy physics. The experience that I've gained working with you is unique. The next round of acknowledgements goes to Massimo Corradi and Sasha Zenaev. It was a pleasure to cooperate with Massimo because he was always glad to help with an advise and a constructive criticism. I'm also glad that we worked together with Sasha. Although I was supposed to be play a role

of an adviser, instead I learnt a lot from our cooperation. There have been many people, like Slava, Lena, Friederike, Verena, Sebastian, Danny, Inna, Denis, Dan, Katerina, Julia and many-many more, who were always open for questions and who have helped me on various stages of this work. My sincere thanks goes to Tobias Haas for the constant support and interest in this analysis. I'm very grateful to Andrea for a structured view on the GRID and FUNNEL systems that I've gained from him. In addition, I would also like to thank Janusz, Krzysztof, Corinna, Yuriy, another Janusz and whole ZEUS offline group for a deep insight into the ZEUS software and help with various technical issues. My journey into physics started 11 years ago[1] in the Ukrainian Physics and Mathematics Lyceum of Kiev University (UPML). I would like to thank my teachers for inspiration and my classmates and friends for constant support and the unique atmosphere.

Last but not least, I would like to thank my parents for support and love over all years. My deepest gratitude also goes to my beloved girlfriend, Natasha. You have always helped me to relaxe, shared my life interests, but on the other hand understood and supported me, when I had to stay concentrated on my work.

[1] Oh my God! Time goes so fast!..