Measurement of

Charm Production

in

Deep Inelastic Scattering

at HERA II

ISBN/EAN: 978-90-6488-030-8

Coverpage drawings: courtesy of Alex Fleisig, http://www.alexfleisig.com/.
The work presented in this thesis is part of the research program of 'het Nationaal Instituut voor Kernfysica en Hoge-Energie Fysica' (NIKHEF) in Amsterdam, the Netherlands. The author was financially supported by the 'Stichting voor Fundamenteel Onderzoek der Materie' (FOM).

UNIVERSITEIT VAN AMSTERDAM

Measurement of Charm Production in Deep Inelastic Scattering at HERA II

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college
voor promoties ingestelde commissie,
in het openbaar te verdedigen
in de Aula der Universiteit

op vrijdag 28 maart 2008, te 10:00 uur

door

Tiberiu Gabriel Grigorescu

geboren te Pitesti, Roemenië

promotores: prof. dr. P. Kooijman
prof. dr. E. Koffeman

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

"Physics is like sex. Sure, it may give some practical results,
but that's not why we do it." - R. Feynman

Contents

Chapter 1

Introduction

Matter and forces in nature

Through the eye of the particle physicist, nature is regarded as matter continuously interacting with itself within a flat space-time. The matter interactions are mediated by three distinct forces: the electromagnetic, the weak nuclear and the strong nuclear forces. The fourth force known in nature, gravity, is significantly weaker at sub-atomic scales. Atoms of different chemical elements consist of the same elementary building blocks: a nucleus composed of neutrons and protons and a cloud of electrons orbiting the nucleus. The neutrons and protons are themselves complex compounds, made of quarks.

Quarks are elementary particles which as yet have no further internal structure. They can never be observed directly but only confined in hadrons. In total, there are six different quarks. Their anti-particles also exist and are called anti-quarks. Baryons, like the neutron or the proton, are bound states of three quarks (qqq). Bound states of a quark and an anti-quark are called mesons ($q\bar{q}$).

The three forces mediating interactions can be described by the exchange of particles: the photon, the three gauge bosons and the eight gluons are the forces' *messengers* through which matter bits interact with each other. A classification of elementary particles is given in Table 1.1. Both the lepton and the quark sectors are organized in three families which feature similar quantum numbers but very different masses. The lightest of the three families is formed by the electron, the electron neutrino and the up and down quarks. These are the particles which constitute the matter around us. Particles belonging to the other two families are formed in high-energy interactions and quickly decay to particles of the first family.

FERMIONS					
Leptons			Quarks		
$\nu_{\mathbf{e}}$ (e-neutrino)	ν_μ (μ-neutrino)	ν_τ (τ-neutrino)	**u** (up)	**c** (charm)	**t** (top)
e (electron)	μ (muon)	τ (tau)	**d** (down)	**s** (strange)	**b** (beauty)
BOSONS					
Forces			**Carrier particle**		
Electromagnetic			γ (photon)		
Weak nuclear			$\mathbf{W^+}$,$\mathbf{W^-}$,$\mathbf{Z^0}$ (gauge bosons)		
Strong nuclear			**g** (gluon)		

Table 1.1. Classification of elementary particles. The fermions constitute matter. The bosons are the force carriers.

The Standard Model is the most advanced quantum theory known yet which describes how elementary particles interact. It incorporates all elementary particles as well as the electromagnetic, weak and strong nuclear forces in one mathematical framework. *The Standard Model* has been confirmed experimentally to great accuracies. Nevertheless, there are many questions for which this theory cannot provide an analytic answer and, for few other questions, no answer at all. For instance, it is not known why elementary particles have precisely the masses that we observe they have.

The proton

The most abundant element in the Universe is hydrogen, which constitutes about 3/4 of the luminous matter[1]. The proton is the nucleus of the hydrogen atom. Protons were created in large numbers at about 10^{-6} seconds after the Big Bang. They are believed to be stable[2]. The proton is built from three valence quarks (uud) which carry the proton quantum numbers. These quarks are kept together by the strong force. The quarks constantly radiate and absorb gluons which, in turn, can split into pairs of quarks and anti-quarks or other gluons. This virtual sea

[1]Only 4% of the total energy density in the Universe consists of luminous matter. The rest is thought to be dark matter (22%) and dark energy (74%).

[2]The experimental lower limit on the proton's lifetime is 10^{35} years.

of quarks and gluons determines many of the proton's properties. Quantum chromo dynamics (QCD) is an advanced quantum theory, itself part of *the Standard Model*, which describes interactions of quarks and gluons.

The inner structure of the proton can be studied in great detail with electron-proton scattering. HERA, an $e^{-}p$ collider, accelerates both protons and electrons to extremely high energies and then collides them against each other. From the resulting collision fragments, information can be inferred about the structure of the proton, while at the same testing predictions of QCD.

Chapter 2

Structure of the proton

2.1 Neutral current deep inelastic scattering

Deep inelastic scattering, in short DIS, is the process in which an incoming lepton collides with a constituent parton from inside the proton. Neutral current, in short NC, refers to those interactions mediated by a neutral boson, either a photon or a Z gauge boson. Interactions mediated by charged gauge bosons $W^{\pm}$, which have a neutrino in the final state, are known as charged current interactions. The higher the virtuality of the gauge boson, the smaller the distances resolved inside the proton.

Event kinematics

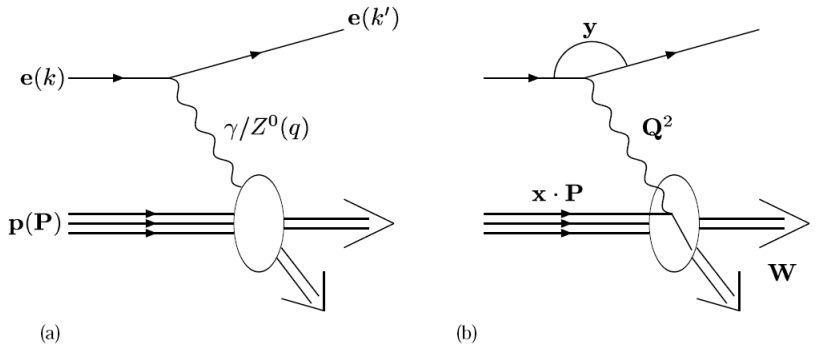


Figure 2.1. Feynman diagrams for a NC DIS collision. 4-vectors of incoming and outgoing particles (a) and Lorentz invariant scalars defining the event (b)

Figure 2.1 shows an electron proton scattering interaction schematically. Taking **k** and **P** to be the 4-momenta of the colliding lepton and proton respectively, the

NC DIS interaction is written as:

$$e(\mathbf{k}) + p(\mathbf{P}) \rightarrow e'(\mathbf{k}') + X(\mathbf{P}') \tag{2.1}$$

where $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ defines the 4-momentum transfer from the lepton to the proton from which follows that $\mathbf{P}' = \mathbf{P} + \mathbf{q}$. In Eq. 2.1, X denotes any final state obeying energy momentum and quantum number conservation. In this interaction, the exchanged boson only interacts with charged constituents of the proton. These charged constituents can be identified with quarks (and anti-quarks). In this way, one can rewrite Eq. 2.1 as:

$$e(\mathbf{k}) + q(x\mathbf{P}) \rightarrow e'(\mathbf{k}') + q(x\mathbf{P} + \mathbf{q}) \tag{2.2}$$

where x is the fraction of the proton four-momentum carried by the struck quark.

The kinematics of an event can be described using the two variables x and Q^2:

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

$$x = \frac{-\mathbf{q}^2}{2\mathbf{P} \cdot \mathbf{q}}$$

which are Lorentz scalars. The virtuality of the photon, Q^2, defines the scale of the interaction. Two other variables, which are not independent of x and Q^2, are frequently used:

- The inelasticity $y = \frac{\mathbf{P} \cdot \mathbf{q}}{\mathbf{P} \cdot \mathbf{k}}$
- The hadronic final state center of mass energy W, defined through $W^2 = (\mathbf{P} + \mathbf{q})^2$.

The following simple formulas allow transformation between variables:

$$W^2 = Q^2 \frac{1 - x}{x} + m_p^2 \tag{2.3}$$

$$Q^2 = sxy \tag{2.4}$$

where s is the center-of-mass energy squared $s = (\mathbf{k} + \mathbf{P})^2$ of the event. A combination of any these two variables define the event kinematics.

2.2 Structure functions

The most general form for the unpolarized cross section for positron proton scattering can be written as $\sigma \sim L_{\mu\nu}W^{\mu\nu}$ where $W^{\mu\nu}$ represents the hadronic tensor and $L_{\mu\nu}$ the leptonic tensor. The hadronic part can be reduced to defining three independent functions, called structure functions, which parametrize the structure of the proton as seen by the virtual boson:

$$\frac{d^2_{NC}\sigma^{e^+p}}{dQ^2dx} = \frac{2\pi\alpha^2}{xQ^4}\left(2xyF_1(x,Q^2) + 2\left(1-y\right)F_2(x,Q^2) - \left(2y+y^2\right)xF_3(x,Q^2)\right) \tag{2.5}$$

with α the fine structure constant $\alpha = e^2/4\pi\epsilon_0 \simeq 1/137$. All the detailed physics of the proton is contained within the structure functions. Often, in the literature the following regrouping $F_L(x) = F_2(x) - 2xF_1(x)$ is performed on Eq. 2.5 to obtain:

$$\frac{d^2_{NC}\sigma^{e^+p}}{dQ^2dx} = \frac{2\pi\alpha^2}{xQ^4}\left(Y_+F_2(x,Q^2) - y^2F_L(x,Q^2) - Y_-xF_3(x,Q^2)\right) \tag{2.6}$$

with $Y_\pm = 1 \pm (1-y)^2$. In this way, F_L describes the coupling of the proton only to longitudinally polarized photons. The structure function F_1 is proportional to the transverse component of the cross-section and F_2 is the sum of both transverse and longitudinal parts of the cross-section. F_3 is the parity violating term due to the Z_0 exchange. F_2 is the dominant structure function. The contribution of F_L to the total cross-section can only become appreciable at y near 1. F_3 only becomes relevant when Q^2 is comparable to the mass squared of the heavy gauge bosons. As this thesis will deal only with Q^2 values that are significantly lower, the F_3 term will be neglected from now on.

When comparing Eq. 2.5 with the cross section for the interaction between a positron and a free spin 1/2 quark the structure function can be identified with:

$$F_2(x) = \sum_q e_q^2 x f_i(x)$$

where e_q is the charge of the quark and $f_i(x)$ are quark densities inside the proton, i.e. the probability of finding a parton i with the momentum fraction x in the proton. The summation is carried out over the different quark flavors. In the case of massless spin 1/2 quarks, the conservation of helicity in the interaction precludes an interaction with a longitudinally polarized photon. In this static model of the

proton, $F_L = 0$. This is commonly referred to as the "Callan Gross" relationship. This static model of the proton predicts that the structure functions are independent of Q^2. This behavior was confirmed at SLAC at $Q^2 \simeq 4\mathbf{GeV}$ [1] and lead to the general acceptance of the quark model of hadrons [2, 3].

Subsequent measurements revealed that the total four-momentum carried by the quarks was far below the value of 1, predicted by the static model described above. The solution to this problem arrived with the introduction of Quantum Chromodynamics as the theory of strong interaction. In this theory, the strong interaction, which holds the quarks together inside the proton, is mediated by the exchange of gluons. These gluons are neutral particles and so do not directly participate in the DIS interaction. They do however carry a substantial fraction of the proton momentum. Experimental evidence for the existence of the gluons was given by the TASSO, JADE and MarkJ experiments at the PETRA collider at DESY in 1979 [4].

2.3 QCD and *ep* interactions

Quantum Chromdynamics (QCD) is a non-abelian gauge theory, based on the SU(3) color symmetry group. It describes the interactions of quarks and gluons. Quarks manifest one of the six possible color charges (three fundamental colors and three anti-colors) and interact by exchanging gluons. Gluons carry one color and one anti-color and can therefore interact with each other as well. This is a direct consequence of the non-abelian nature of the underlying symmetry group. The strong force decreases at small distances, a phenomenon called asymptotic freedom. *The Standard Model* describes the strength of the quark-gluon interaction in terms of the strong coupling constant α_s. The value of α_s varies with Q^2. As Q^2 rises, smaller distances can be resolved inside the proton and smaller values of α_s are measured. A world summary of measurements of α_s[5] is shown in Fig. 2.2. Theoretically, the running of α_s is expressed as:

$$\alpha_s(\mu_r^2) = \ln\left(\frac{12\pi}{(33-2N_f)\ln\left(\frac{\mu_r^2}{\Lambda_{QCD}^2}\right)}\right) \tag{2.7}$$

where μ_r is the renormalization scale, N_f is the number of active quark flavors in the interaction and Λ_{QCD} is a parameter which is determined experimentally. As μ_r^2 increases, for $\mu_r^2 >> \Lambda_{QCD}^2$ it holds that $\alpha_s << 1$. In this regime, QCD can be

described completely perturbatively such that each higher order Feynman diagram involving a higher power of α_s will contribute less to the total cross-section than the previous orders in the calculation. In the case when $\mu_r^2 \sim \Lambda_{QCD}^2$, $\alpha_s \sim \mathcal{O}(1)$, higher order terms in the perturbative expansion may not converge and therefore non-perturbative approaches are taken, such as phenomenological models based on measurement or numerical methods. In DIS, μ_r^2 can be equated to Q^2.

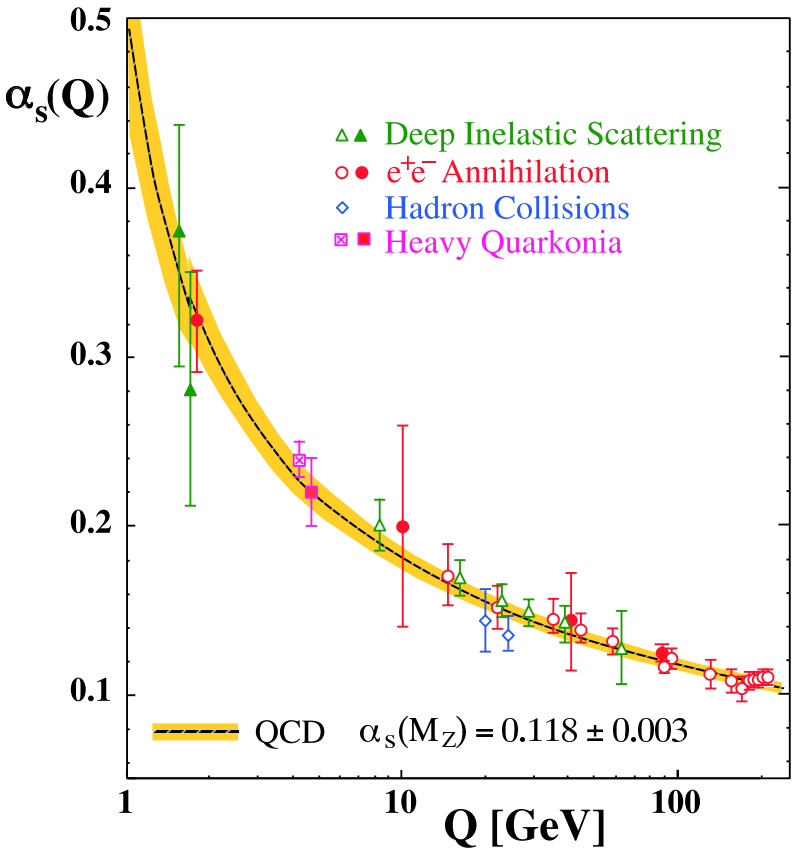


Figure 2.2. World summary of measurements of the running coupling constant α_S, as measured in DIS, e^+e^-annihilation, hadronic collisions and heavy quarkonia, as a function of Q^2. The curves are the QCD predictions for the combined world average value of $\alpha_s(M_{Z^0})$ (2006).

2.4 QCD dynamics and evolution

Inside the proton, quarks continuously exchange gluons and gluons fluctuate in $q_i\bar{q}_i$ pairs or even interact among themselves (a direct consequence of the fact that gluons also carry color charge). The number of quarks and gluons changes depending on the scale of interaction. Also, the strong force decreases rapidly at small distances:

this is called asymptotic freedom. This allows DIS to be described as a photon interaction with a free quark. Processes like a gluon radiating a quark or a quark splitting into a quark and a gluon can be computed in QCD. The calculation of the cross-section $\gamma^* q \to qg$ yields:

$$\frac{d\sigma^{\gamma^* q \to qg}}{dp_T^2} = \frac{4\pi\alpha^2}{s} e_q^2 \frac{1}{p_T^2} \frac{\alpha_s}{2\pi} P_{qq}(z) \tag{2.8}$$

where z is the momentum fraction of the outgoing quark (w.r.t. its incoming momentum) and p_T is its transverse momentum. The function $P(z)$ is called the splitting function and is proportional to the probability for the quark to split into a quark with a momentum fraction z by radiating a gluon [6]:

$$P_{q\to q}(z) = P_{qq}(z) = \frac{4}{3}\left(\frac{1+z^2}{1-z}\right) \tag{2.9}$$

The divergence at $z = 1$ is due to radiation of very soft gluons (very low energy) and is canceled by virtual loop contributions. Integrating Eq. 2.8 over p_T, one obtains:

$$\sigma^{\gamma * q \to qq} = \frac{4\pi\alpha^2}{s} e_q^2 \frac{\alpha_s}{2\pi} P_{qq}(z) \log\frac{Q^2}{\mu^2} \tag{2.10}$$

with μ some minimum momentum of the outgoing quark. The contribution of this process to the structure functions is Q^2 dependent. $F_2(x, Q^2)$ becomes:

$$\frac{F_2\,(x, Q^2)}{x} = \sum_{q,\bar{q}} e_q^2 \int_x^1 \frac{dy}{y} q(y) \left(\delta\left(1 - \frac{x}{y}\right) + \frac{\alpha_s}{2\pi} P_{qq}(\frac{x}{y}) \log\frac{Q^2}{\mu^2}\right) \tag{2.11}$$

where y is the fraction of the incoming quark w.r.t. the proton momentum and x is the momentum fraction of the secondary quark w.r.t. the proton $z = x/y$. The integral runs over all possible momentum fractions y larger than x.

One can also include quark contributions coming from gluon splitting into quarks or other gluon contributions from gluon splitting and so on. The parton (quark and gluon) densities will evolve as a function of the probe scale due to the increase in detail with increasing Q^2. This evolution of the PDF's is given by the following equation:

$$\frac{\partial}{\partial \ln Q^2}\begin{pmatrix} q_i(x, Q^2) \\ g_i(x, Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dy}{y} \mathbf{P_1}\left(\frac{x}{y}\right) \begin{pmatrix} q_i(y, Q^2) \\ g_i(y, Q^2) \end{pmatrix} \tag{2.12}$$

where q_i denotes all the active (light) quark flavors. This equation is known as the DGLAP equation, after its authors Dokshitzer, Gribov and Lipanov and Altarelli and Parisi[6]. It describes the coupled evolution of quark and gluon densities in the proton. The kernel $\mathbf{P_1}$ is given by:

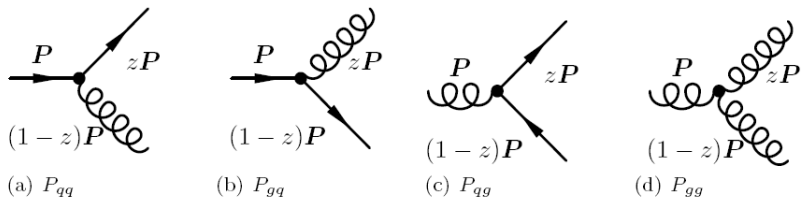
$$\mathbf{P_1}(z) = \begin{pmatrix} P_{qq}(z) & P_{qg}(z) \\ P_{gq}(z) & P_{gg}(z) \end{pmatrix} \tag{2.13}$$

where one of the entries was made explicit in Eq. 2.9. Each of the splitting functions that enter the kernel are represented schematically in Fig. 2.3.

Through the QCD induced evolution, the parton distributions depend on Q^2. By inspecting the proton with finer and finer probes (higher Q^2) more and more quarks and gluons can be resolved. We have taken into account so far only leading order (LO) splitting functions. These can be extended to next to leading order (NLO):

$$\frac{\partial}{\partial \ln Q^2}\begin{pmatrix} q_i(x,Q^2) \\ g_i(x,Q^2) \end{pmatrix} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y}\mathbf{P_1}\left(\frac{x}{y}\right)\begin{pmatrix} q_i \\ g_i \end{pmatrix} + \frac{\alpha_s^2}{4\pi^2}\int_x^1 \frac{dy}{y}\mathbf{P_2}\left(\frac{x}{y}\right)\begin{pmatrix} q_i \\ g_i \end{pmatrix} + ...$$

and so on. The splitting functions for higher orders gain in complexity as they involve more diagrams. The NLO splitting functions are known for some time [6] and recently NNLO splitting functions have been calculated numerically [7]. Possible contributions from NLO diagrams are shown in Fig. 2.7.



(a) P_{qq} (b) P_{gq} (c) P_{qg} (d) P_{gg}

Figure 2.3. Leading order contribution of different splitting functions

Processes involving higher order QCD interactions give rise to a non-zero coupling between the longitudinally polarized incoming photon and quarks off the mass-shell and consequently $F_L \neq 0$.

2.5 The improved quark parton model

It was shown that QCD can describe the evolution of the parton distributions inside the proton with Q^2. It is however not possible to predict the values of the parton distributions and these must therefore be determined experimentally.

In practice, the DIS structure function data is utilized for this. The method used is to parametrize all parton densities (quark and gluon) at a single value of Q^2, Q_0^2, as a function of x. Given the parton densities, QCD is used to predict the structure functions over a large kinematic range. By comparing the predictions with the measured structure functions, the parameters describing the parton density distributions at Q_0^2 can be fitted.

This method has been used by the ZEUS collaboration to determine the parton distributions. The inclusive structure functions measured by ZEUS and several fixed target experiments are given in Fig. 2.5, together with the QCD fit. Excellent agreement is obtained. The parton distributions extracted from these fits are shown in Fig. 2.4. The results are compared to PDFs from several PDF fitting groups, MRST [8] and CTEQ [9].

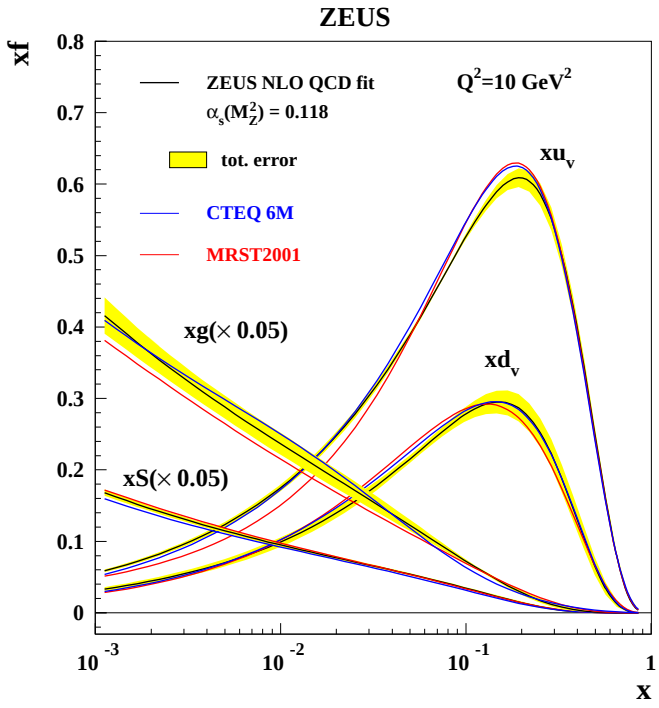


Figure 2.4. Standard ZEUS NLO QCD fit for $Q^2 = 10\,\mathbf{GeV^2}$. The gluon sea and the up and down valence quark distributions are shown. The shaded band represents the uncertainty. For comparison, MRST and CTEQ fits are also shown.

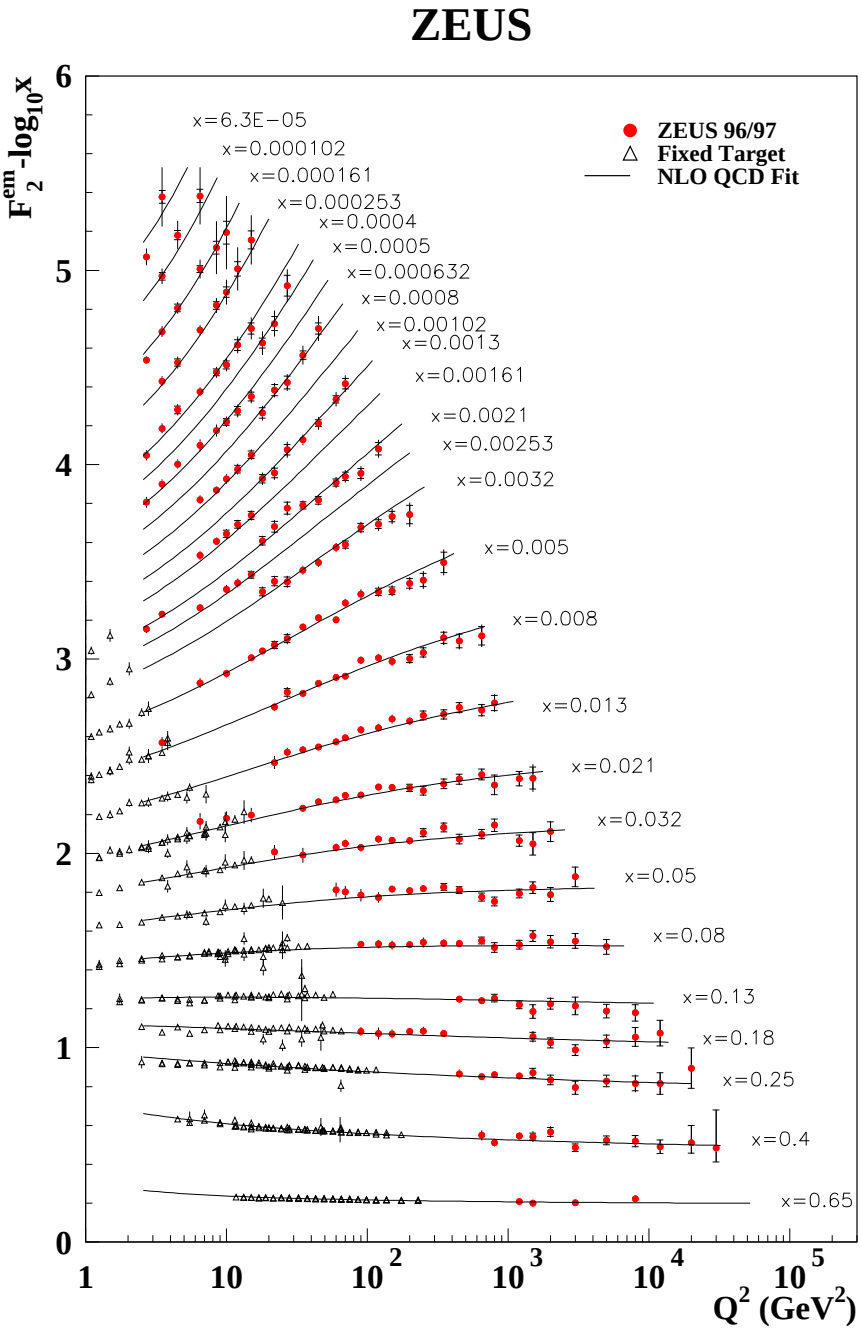


Figure 2.5. F_2 as a function of Q^2 for several values of x. ZEUS data are compared to fixed-target experiments NMC, BCDMS and E665 as well as to the ZEUS NLO QCD fits.

2.6 Heavy quark production

Production of heavy quarks can be determined at least at moderate Q^2 perturbatively, as their masses are larger than Λ_{QCD}. The focus will be on c and b quark production as t production is beyond the reach of the HERA accelerator.
Above the charm threshold, effectively $Q^2 \sim (2m_c)^2$, $c\bar{c}$ pair production increases steeply, contributing more and more to the F_2 structure function with increasing Q^2 and decreasing x. For instance, at $x \sim 0.01$ and $Q^2 \sim 100\ \mathbf{GeV^2}$, charm contributes approximately 25% to the total NC DIS cross-section.

The dominating process for creating such $c\bar{c}$ pair at HERA is known as boson gluon fusion and schematically represented in Fig. 2.6. In the pQCD inspired picture of the proton, the gluon splits (at leading order) in a off mass-shell $c\bar{c}$ pair which subsequently interacts with the photon. The interaction transfers enough energy to the quark system such that the charm quarks can become on mass-shell. Leading order and next to leading order diagrams are shown in Fig. 2.7.

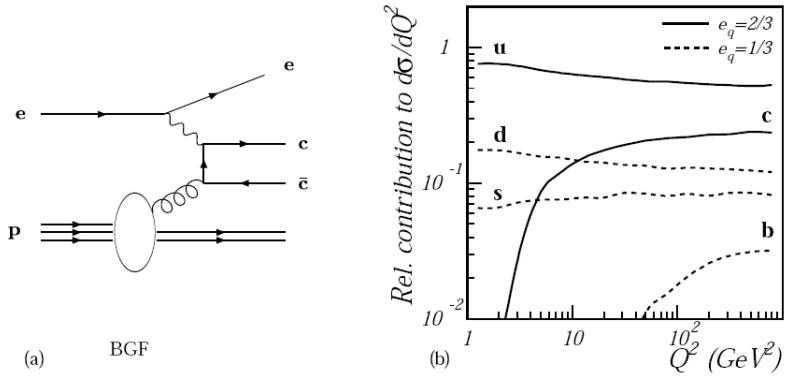
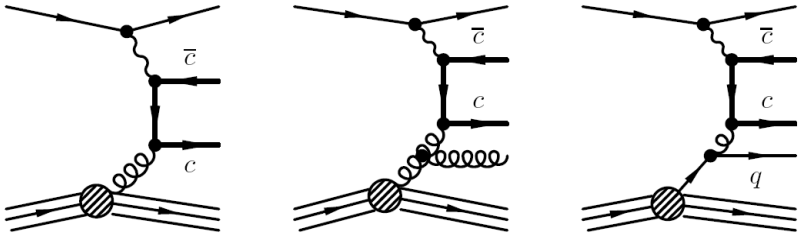


Figure 2.6. (a) The boson gluon fusion process. (b) Different flavor contributions to the NC ep cross section for $0.005 < x < 0.02$. Charm and beauty production decreases steeply at low Q^2

The charm leading-order contribution to the structure function is directly proportional to the gluon density in the proton. The next-to-leading-order contribution also contains a term that follows the gluon density. Therefore, charm production is directly sensitive to gluon density inside the proton. The cross-section for charm in DIS follows from eq. (2.5) (restricted to charm only):

$$\frac{d^2\sigma^{c\bar{c}}}{dQ^2dx} = \frac{2\pi\alpha^2}{xQ^4}\left((1+(1-y)^2)F_2^{c\bar{c}}(x,Q^2) - y^2F_L^{c\bar{c}}(x,Q^2)\right) \tag{2.14}$$



(a) Leading order (b) NLO gluon (c) NLO quark

Figure 2.7. (a) Leading order contribution to charm production in NC DIS. (b), (c) contribution of the NLO gluon ($\sim 20\%$), quark ($\sim 5\%$)

The charm structure functions dependence on the parton densities are known to the next-to-leading order [10]. The b quark is heavier: its production threshold is effectively $Q^2 = (2M_b)^2 \simeq 100\,\mathbf{GeV^2}$. As F_2 is proportional to the the square of the quark charge, the beauty production is a factor of $(e_c/e_b)^2 = 4$ suppressed w.r.t. charm production, at infinite Q^2. For a $\sqrt{(s)} = 318\mathbf{GeV}$, the cross-sections for $ep \to ec\bar{c} + X$ and $ep \to eb\bar{b} + X$ are about $0.5\,\mu b$ and $1\,nb$ respectively.

2.7 Charm hadrons

The calculations of cross-sections for b and c quark production are performed in perturbative QCD. Nevertheless, due to color confinement in QCD[1], the experimentalist measures colorless hadrons (mesons and baryons) which are produced in the fragmentation process. Hadronization cannot be described by pQCD. Therefore, theoretical predictions at hadron level depend on data already collected at other experiments and on empirical models.

Charm hadrons are hadronization products of the charm quark. Some of their properties are listed in the Table 2.1. The lighter mesons, with masses below $2\,\mathbf{GeV}$, are pseudo-scalars. The heavier mesons, containing a "*" in their names, represent excited states of their pseudo-scalar counterparts and are vector mesons with the same quark content but different quark spin alignment. Pseudo-scalars decay weakly: for instance, the two main contributing diagrams in D^0 decay are given in Fig. 2.9. There, a charm meson decays into non-charm mesons due to the flavor changing property of the weak decay. The weak decay also dictates the timescale for this decay to happen. In Fig. 2.8, the two main contributions to D^{*+} vector me-

[1]Free quarks or gluons cannot exist individually. The process of formation of hadrons out of free quarks or gluons is called hadronization.

meson	D^0	D^+	D_s^+	D^{*+}	D^{*0}	D_s^{*+}
quarks	$c\bar{u}$	$c\bar{d}$	$c\bar{s}$	$c\bar{d}$	$c\bar{u}$	$c\bar{s}$
$m\,(\mathbf{GeV})$	1.864	1.869	1.968	2.010	2.007	2.112
$c\tau\,(\mathbf{s})$	$4.1 \cdot 10^{-13}$	$10.4 \cdot 10^{-13}$	$5.0 \cdot 10^{-13}$	$\mathcal{O}(10^{-18})$	$\mathcal{O}(10^{-18})$	$\mathcal{O}(10^{-18})$

Table 2.1. Charm mesons and their properties as listed by Particle Data Book 2006: the valence quarks content, the mass (in GeV) and the mean lifetime (in seconds) are given.

son are depicted. Only electromagnetic and strong interaction forces are involved. The associated time scale is much shorter. Nowadays, particle detectors tend to have spatial resolutions for vertex reconstruction of the order of tens-hundreds of microns. For high momentum particles, above $1\,\mathbf{GeV}$, one can reconstruct charm meson decay vertices: the secondary vertices can be separated from the primary interaction point and therefore charm decay can be tagged and properties such as lifetime can be measured. Vector mesons decay much too fast and from the experimentalist's point of view their decay happens "at the primary vertex". We will come to this issue in later chapters as it is a critical tool for improving filtering out combinatorial background w.r.t. the genuine signal.

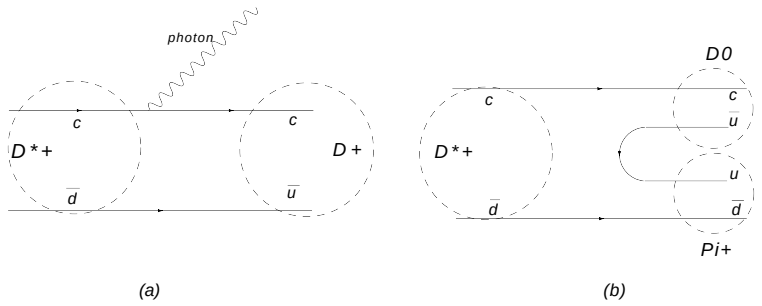


Figure 2.8. Two main contributions to the D^* branching ratios: D^{*+} decays electromagnetically (a) $D^{*+} \to D^+ + \gamma$ or strongly (b) $D^{*+} \to D^0 + \pi^+$. The time scale of these decays is much shorter than the weak interaction timescale. (roughly 10^{-18} sec. for e.m. and 10^{-23} sec. for strong interaction). The out-coming particles are still charm mesons.

The probability of charm to hadronize to a particular charm hadron is described by the charm fragmentation fractions $f(c \to D\,,\,\Lambda)$. The fragmentation fractions are assumed to be universal[11]. The fragmentation of the c quark is shown schematically in Fig. 2.10. The branching ratio's represent world averages, as reported by

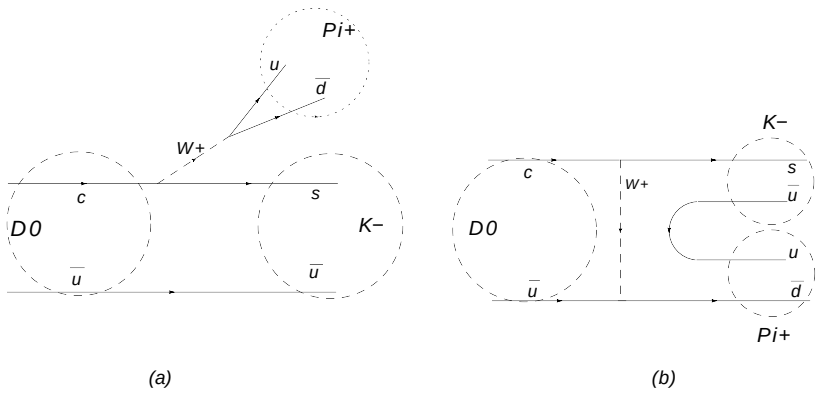
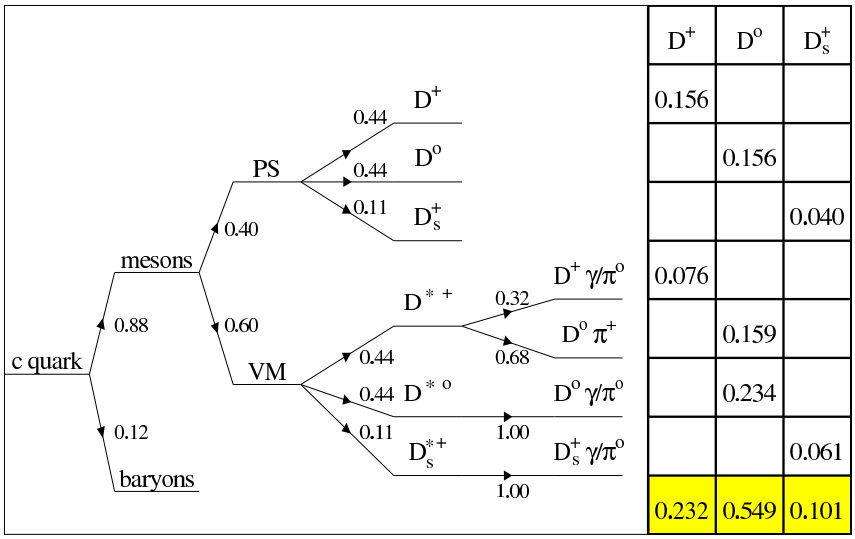


Figure 2.9. Main contributions to the $D^0 \to K^- + \pi^+$ weak decay. (a) The c quark emits a W^+which solely decays into a π^+ (b) an internal W^+ exchange and a creation of $u\bar{u}$ pair out of the vacuum. As these two diagrams interfere constructively, the cross-section for $D^0 \to K^- + \pi^+$ increases and the D^0 lifetime decreases.

the Particle Data Group [12]. A recent ZEUS measurement [11] is shown in Fig. 2.11. After the subsequent strong/electromagnetic decays of D^{*0} and D^{*+}, the D^0 charm meson is the most abundant.



D^+	D^o	D_s^+
0.156		
	0.156	
		0.040
0.076		
	0.159	
	0.234	
		0.061
0.232	0.549	0.101

Figure 2.10. The charm fragmentation tree into vector and pseudo-scalar charm mesons. The numbers indicate the world average values for the fragmentation fractions[12]. The table on the right sums the different branching contributions to the fragmentation fractions of charm to D^+, D^0 and D_s^+.

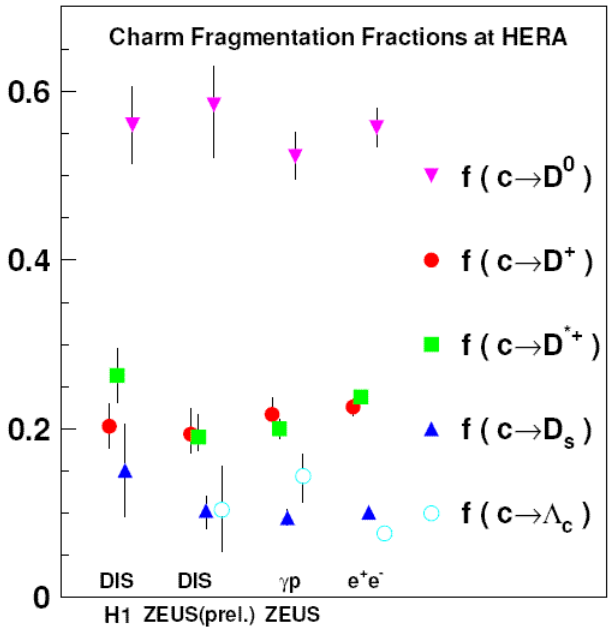


Figure 2.11. Charm fragmentation fractions as measured by ZEUS in photo-production and DIS, compared to H1 and e^+e^- results. More than half of charm quarks fragment into D^0 mesons. [11]

Chapter 3

The ZEUS detector

3.1 The HERA accelerator

The HERA (Hadron-Elektron Ring Anlage) collider, located at the Deutches Elektronen Synchrotron (DESY) research center in Hamburg, Germany, is the world's first accelerator to collide beams of electrons and protons. The HERA ring measures a total of 6.3 **km** in circumference and is situated $10 - 25$ **m** under ground. The layout of the collider is shown in Fig. 3.1.

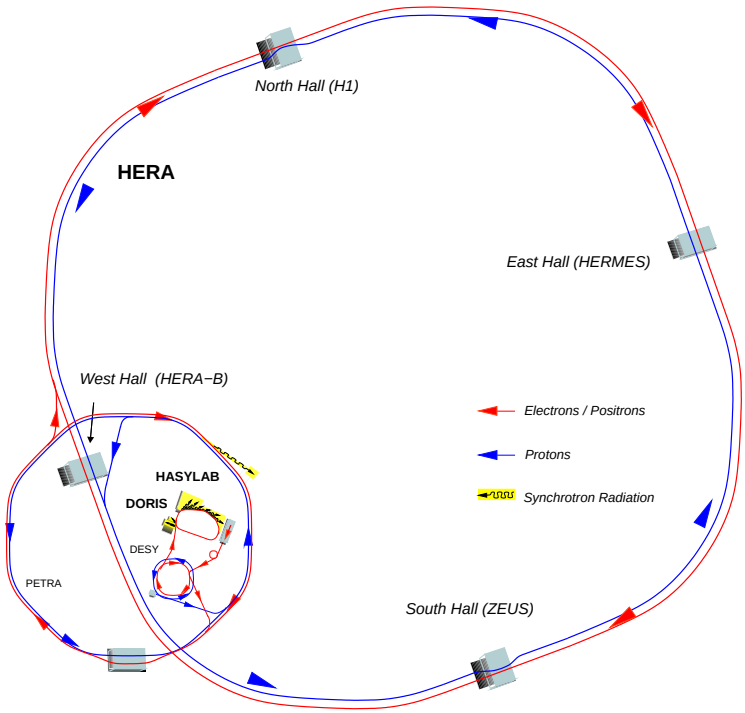


Figure 3.1. Layout of the HERA collider and pre-accelerators

Data taking at HERA began in spring 1992 and ended in summer 2007. Protons and electrons (or positrons) are stored in two independent storage rings. The collider features four straight sections, each 360 **m** in length; in three of them interactions occur (at ZEUS, H1 and HERMES), while the fourth is used for injection. Before being injected into the main ring, protons and electrons are pre-accelerated. Electrons get an initial momentum of 200 **MeV** in the linear accelerator after which they are transferred to the DESY II synchrotron and accelerated to 7.5 **GeV**. After transfer to the PETRA pre-accelerator, they are brought up to 14 **GeV** after which they are injected into HERA. Here, after the final acceleration, the electrons reach 27.5 **GeV** and are ready for collision. Protons also undergo several stages of pre-acceleration: starting off as H^- ions at 50 **MeV** in the proton LINAC, they are accelerated to 7.5 **GeV** in DESY III and further to 40 **GeV** in PETRA II before being injected in HERA. Once in HERA, protons are accelerated to a momentum of 920 **GeV**. Protons are kept on their orbit by superconducting dipole and quadrupole magnets. The electron beam is guided by warm (non-superconducting) magnets. Both beams are accelerated to their nominal momenta in HERA by radio-frequency cavities.

Along the interaction regions, guiding magnets deflect the particle beams until they run inside the same vacuum pipe. Collisions happen at zero crossing angle. After passing the interaction regions, the protons are deflected back to the proton ring.

Both the electron and the proton beams are divided in particle bunches. There is a separation of 28.8 **m** between each pair of bunches with a total capacity of 220 bunches orbiting around the HERA ring. The time between bunch crossings is 96 ns. Table 3.1 summarizes the main HERA parameters.

The main parameter of particle colliders is the *luminosity* $\mathcal{L}$. The luminosity is defined as the number of collisions per unit of area and per unit of time ($cm^{-2}s^{-1}$). The *specific luminosity* is defined as the luminosity divided by the beam currents ($cm^{-2}s^{-1}A^{-2}$). It quantifies the intrinsic quality of the colliding beams. The *integrated luminosity* $\mathcal{L}_{int}$ is the integral of the luminosity over a period of time (cm^{-2}). In particle physics, cross-sections are expressed in *pico-barns*, with $1\,pb = 10^{-36}cm^{-2}$. The number of occurrences of a certain process with a corresponding cross-section σ found in a data sample of an integrated luminosity $\mathcal{L}_{int}$ is given by:

Tunnel circumference	$6336.83\,m$
Bunch spacing (in time)	$96\,ns$
Bunch spacing (in distance)	$28.8\,m$
Number of buckets	220
Beam crossing angle	$0\,mrad$
Center of mass energy	$318\,GeV$
Number of colliding bunches	174
Peak luminosity	$7.5 \times 10^{31}\,cm^{-2}s^{-1}$
Specific luminosity	$1.8 \times 10^{30}\,cm^{-2}s^{-1}A^{-2}$
max. I_p	$140\,mA$
max. I_e	$58\,mA$
Beam size at I.P. in x	$0.110\,mm$
Beam size at I.P. in y	$0.030\,mm$

Table 3.1. Hera collider main parameters.

$$N = \sigma \times \mathcal{L}_{int} \tag{3.1}$$

3.2 The ZEUS detector

The ZEUS detector [13] has been designed as a multi-purpose particle detector that measures final state particles in ep collisions. It has an almost 4π solid angle coverage around the interaction point, leaving out only two small holes for the beam pipe to enter and exit. The detector is shown from different angles in Fig. 3.2, 3.3 and 3.4.

Fig. 3.2 shows a cross-section along the beampipe of the complete detector. The major detector components are labeled explicitly: forward, rear and barrel sections of the uranium calorimeter (the FCAL, RCAL and BCAL), the central tracking detector (CTD), the microvertex detector (MVD), the forward detector (STT), the rear tracking detector (RTD) and the small angle rear tracking detector SRTD as well as the muon detectors BMUON , FMUON and RMUON. The backing calorimeter (BAC) forms the outer shell of the ZEUS detector. Between the CTD and the uranium calorimeter, there is a thin superconducting solenoid which produces a $1.43\,\mathbf{T}$ magnetic field. The iron yoke outside the calorimeter functions as a return path for the magnetic field lines. Several detectors were built to veto events arising

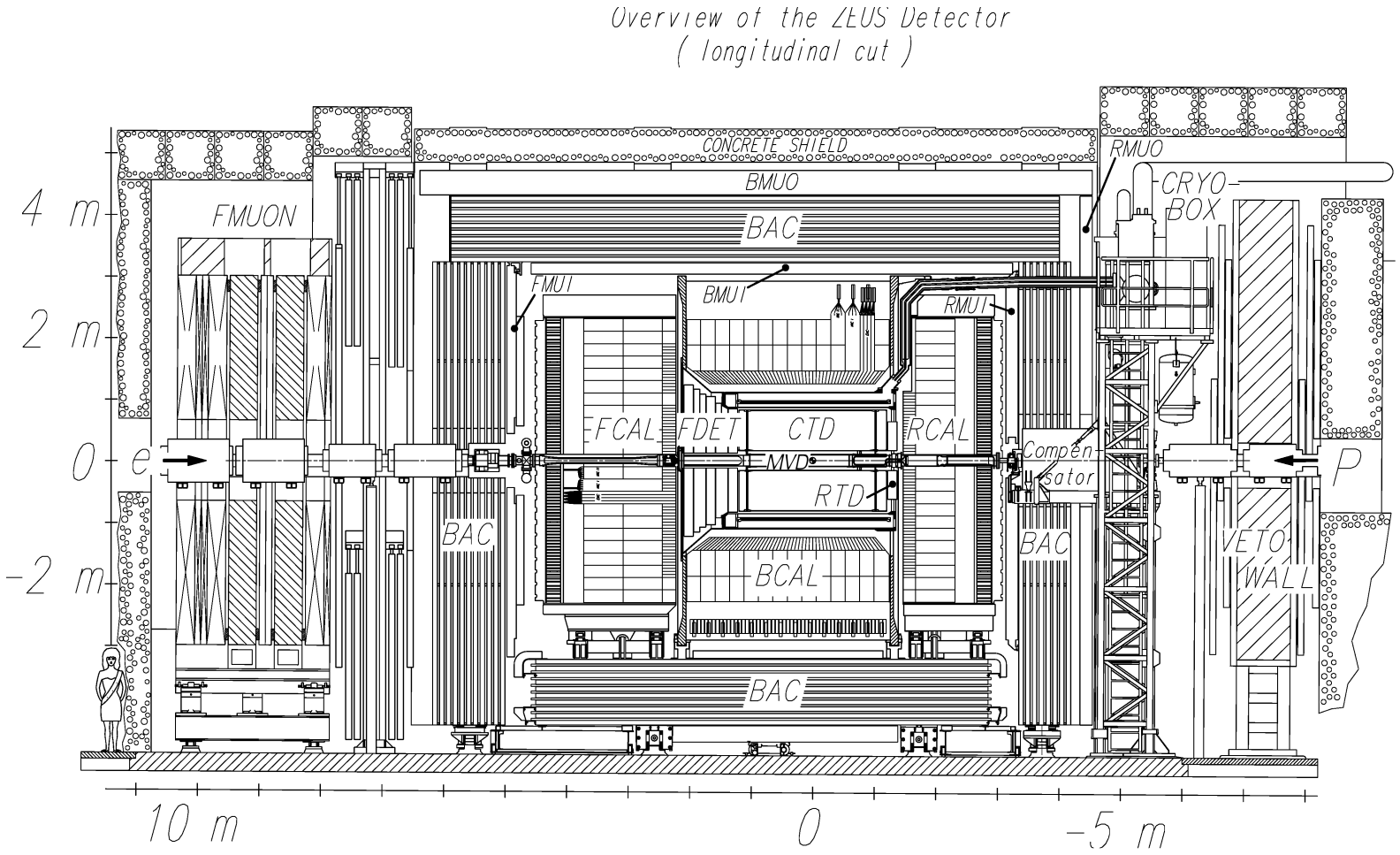


Figure 3.2. The ZEUS detector ($z-y$ view). The electrons enter the detector from the left. Major components are labeled explicitly.

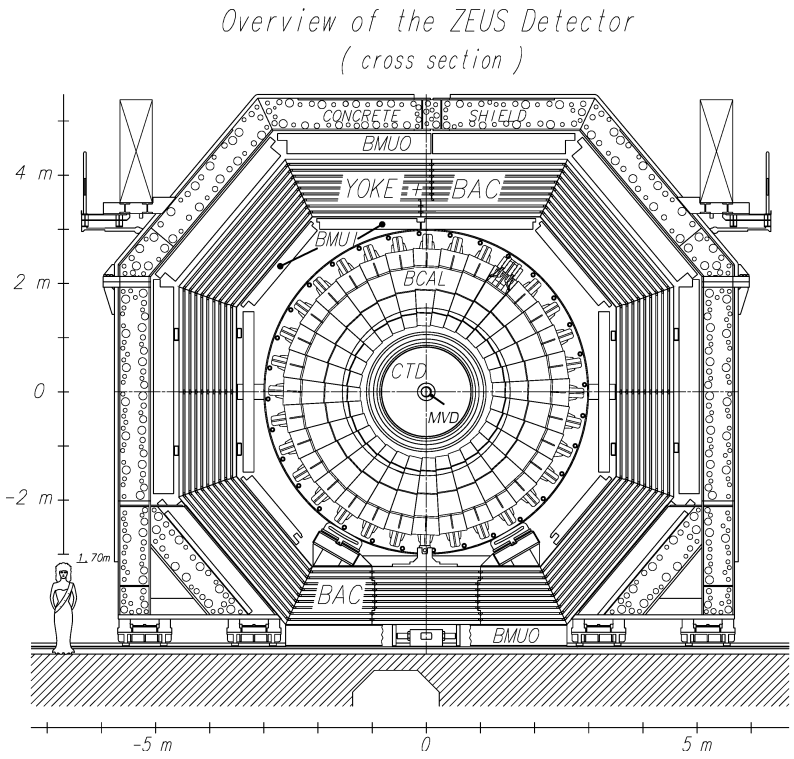


Figure 3.3. The ZEUS detector in XY cross-section view. The beampipe runs perpendicular through the center of the figure.

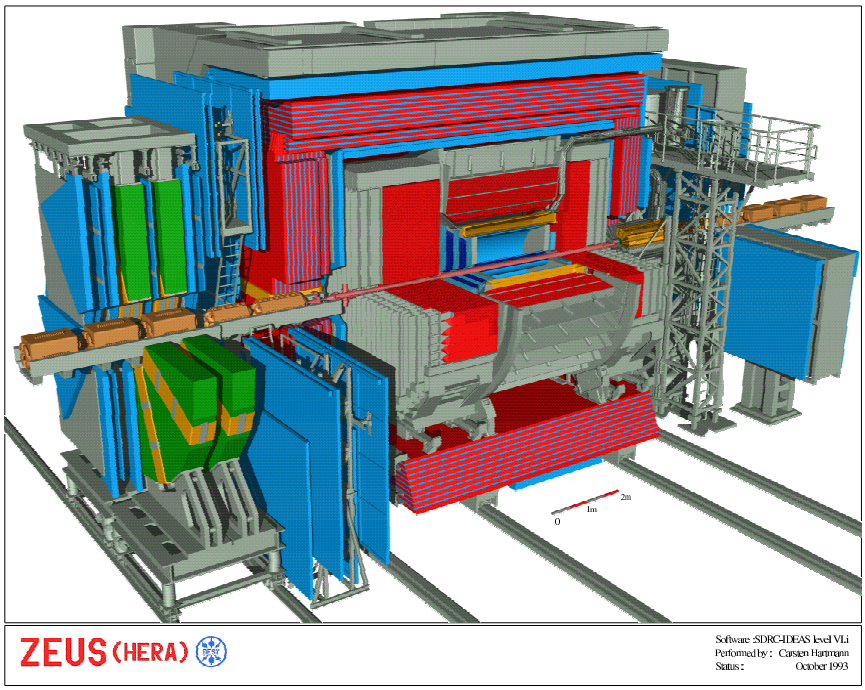


Figure 3.4. The ZEUS detector.

from beam-gas collisions and secondary collisions: the veto wall detector VETO and the C5 counter are two of these detectors. The luminosity monitor LUMI is located 107 **m** away from the interaction point.

The ZEUS coordinate system is a right-handed Cartesian system with the Z axis pointing in the flight direction of the proton beam (the forward direction). The X axis points towards the center of the HERA ring and the origin of the coordinate system is at the nominal interaction point.

The analysis described in this thesis relies on the calorimeter and the tracking detectors CTD and MVD [1]. They will be described in the following sections.

3.2.1 The uranium calorimeter

The uranium scintillator calorimeter (CAL) [14, 15, 16] covers 99.8 % of the full solid angle and has been designed to stop all particles except muons and neutrino's. The calorimeter weighs 700 tons (see Fig 3.6). It is subdivided into 80 modules which form the forward (FCAL), central (BCAL) and rear (RCAL) calorimeters. The FCAL, RCAL and BCAL are further transversely divided into towers and longitudinally into one electromagnetic section (EMC) and one (in RCAL) or two (in BCAL, FCAL) hadronic sections. Each module has a layered structure consisting of 3.3 mm thick depleted uranium (DU) plates wrapped in stainless steel foils (200 or 400 μm thick) as absorber material interleaved with 2.6 mm thick sheets of plastic scintillator as detector. The thicknesses of the absorber and detector plates were selected to achieve equal electron and hadron response. Up to 185 layers of absorber and scintillator form the calorimeter modules. A schematic cross-section of a BCAL tower is shown in Fig. 3.5.

Light generated by charged particles crossing the scintillator planes is collected by plastic wavelength shifters and guided to photomultiplier tubes. The pulse height in each cell and the arrival time are read out, with a precision better than one nanosecond (whereas the bunch crossing time is 96 ns).

The calorimeter provides a linear response for electrons and hadrons (such as protons or pions) up to the highest energies. Test beam measurements revealed a relative energy resolution of $\sigma(E)/E = 0.35/\sqrt{E}$ for hadrons and jets and $\sigma(E)/E = 0.17/\sqrt{E}$ for electrons (with E in **GeV**).

Different particles leave different shower signatures in the calorimeter. This is illustrated in Fig. 3.7. Information from both longitudinal and transverse shapes of

[1] The forward detector STT did not function during 2005.

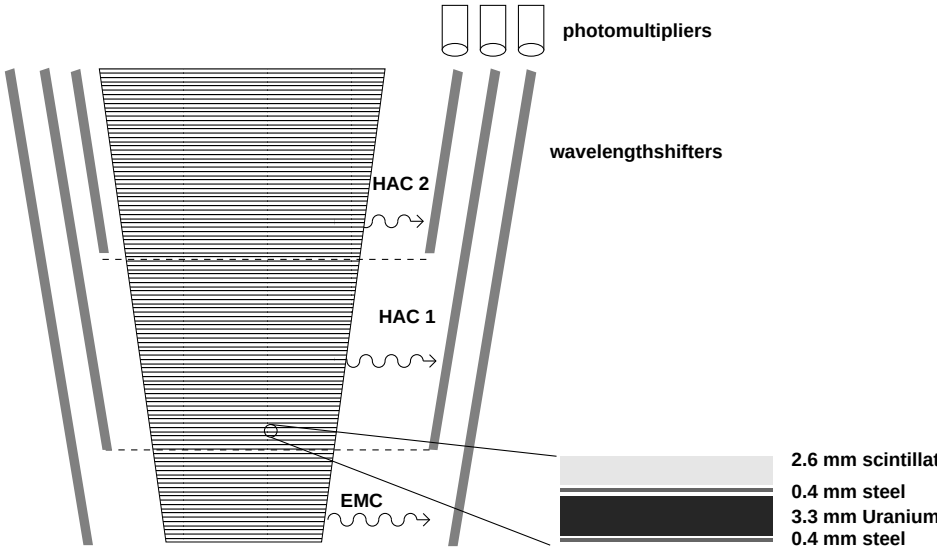


Figure 3.5. Cross-section view through BCAL. The electromagnetic section and the hadronic sections are shown. The electromagnetic section is closest to the interaction point. On the right, the layered structure of the calorimeter module is enlarged.

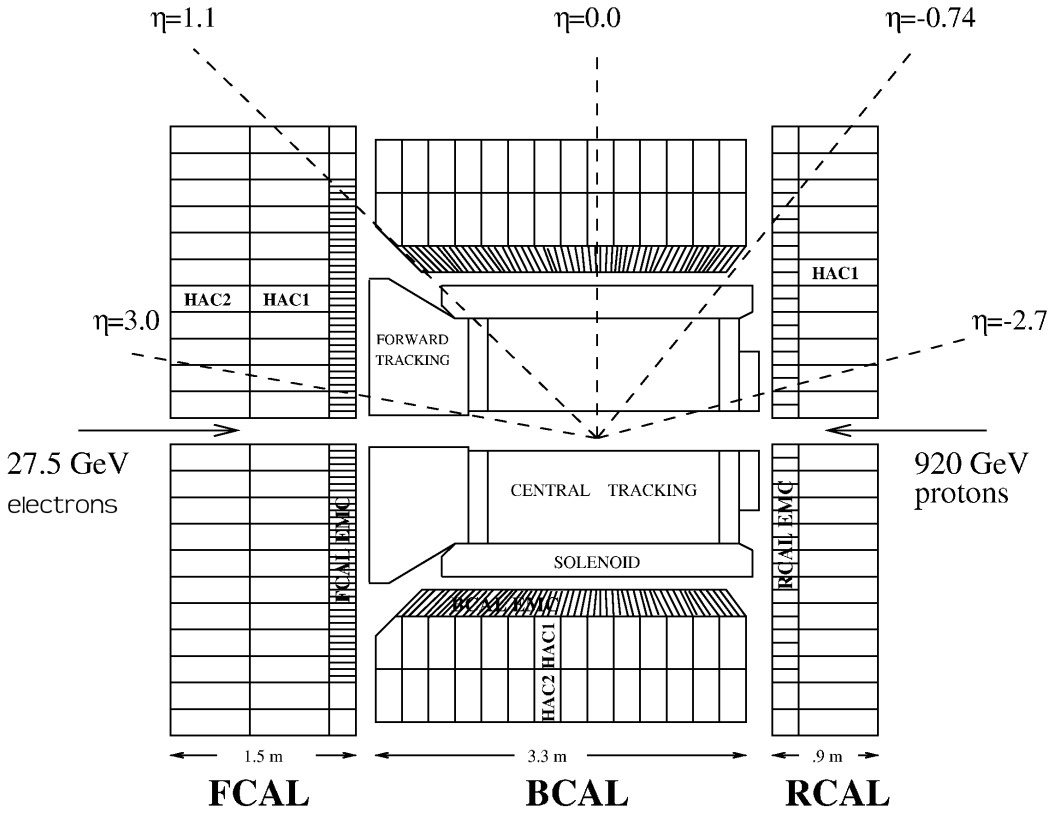


Figure 3.6. The uranium calorimeter within the ZEUS detector ($x - y$ view).

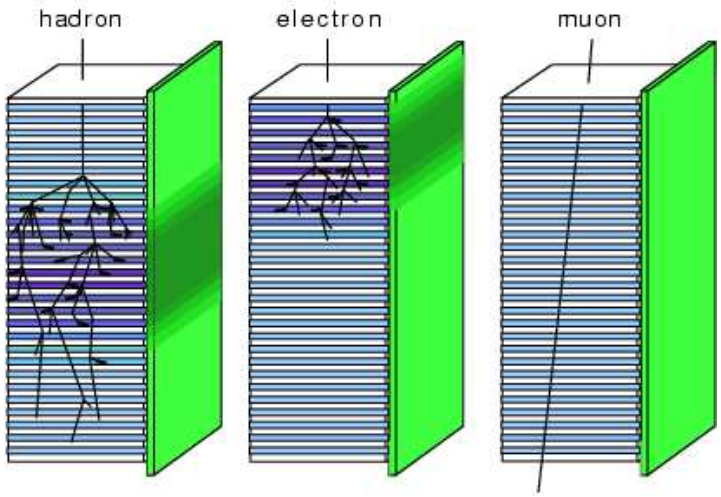


Figure 3.7. Different calorimeter shower signatures produced by a hadron, an electron and a muon are shown. In this figure, particles originating at the interaction point enter the detector from above.

the showering in the calorimeter allows for identification of three different particle types with the CAL:

- electrons, positrons and photons start to shower immediately after entering the CAL. The shower is limited in size, longitudinally and transversely.

- hadrons start showering only deeper into the calorimeter. The showers extend more in depth and width.

- muons completely traverse the calorimeter, depositing an energy loss (through ionization) that is mostly independent of the muon energy.

The calorimeter is calibrated using the natural radioactivity of the depleted uranium. The gain of the photomultiplier tubes is calibrated few times a month resulting in a uniformity of about 1%. At the end of each HERA fill, bad channel lists are archived for corrections in the subsequent offline reconstruction. During data taking, the shift crew carefully monitors calorimeter parameters in combination with trigger rates. Especially noisy cells in the calorimeter can spoil an efficient data taking and therefore should be *masked* by the calorimeter first level trigger expert as soon as they are discovered. Bad channels in the calorimeter are repaired regularly, during short periods of detector access.

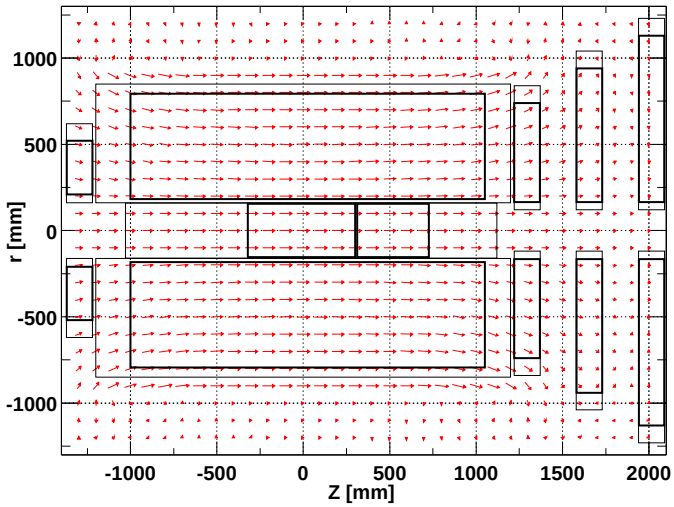


Figure 3.8. The flow of the magnetic field lines inside the detector space. The arrows follow the direction of the field lines; their length is proportional to the strength of the magnetic field. (Protons enter the detector from the left.)

3.2.2 The superconducting solenoid

The tracking detectors work in a 1.43 **T** magnetic field provided by a superconducting solenoid. Inside the magnetic field, charged particles follow a helix like trajectory in response to the Lorentz force:

$$\vec{F}_L = q\vec{v} \times \vec{B} \tag{3.2}$$

with q the electric charge, $\vec{v}$ the velocity and $\vec{B}$ the magnetic field vector. The transverse momentum of the charged particle is obtained from the track curvature:

$$p_T = |kqBr| \tag{3.3}$$

where r is the radius in *meters,* $\vec{B}$ is the field strength measured in *Tesla,* p_T the transverse momentum in **GeV** and k is the curvature constant, $k = 0.2998\,\mathbf{GeV}\,T^{-1}\,m^{-1}$.

In ZEUS, the magnetic field is oriented along the Z axis with a strength of 1.43 **T** at the interaction point and remains almost constant inside the tracking volume. A map of the field lines is shown in Fig. 3.8.

3.2.3 The central tracking detector

The central tracking detector (CTD) [17, 18] is a wire chamber that covers 205 cm in Z and extends in the radial direction from 16.2 cm to 79.4 cm. It consists of 72 cylindrical drift chamber layers, organized in 9 superlayers. The CTD covers a polar angle region of $15° < \theta < 164°$ and is filled with a mixture of argon, carbon dioxide and ethane. When particles pass through the detector, they ionize gas atoms; the freed electrons drift towards the sense wires. Near a wire, ionization avalanches are generated due to large electric fields. The pulse measured with the sense wires is proportional to the energy loss of particles passing through the CTD. Particle identification is possible using measurements of the mean energy loss, dE/dx, of charged particles in the gas of the active volume. The odd numbered (axial) superlayers contain drift wires that run parallel to the z-axis. The even numbered (stereo) superlayers are oriented at a small angle with respect to the z-axis. This allows both $r\phi$ and z coordinates to be measured.

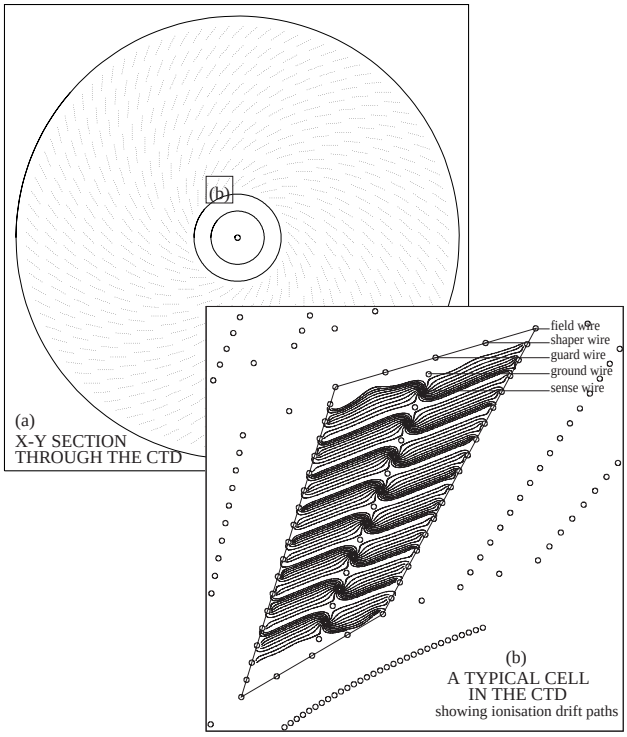


Figure 3.9. $x - y$ view through the CTD. Typical cell lay-out.

The inner three axial superlayers of the CTD are provided with a z-by-timing system. This estimates the z position of a hit by measuring the difference in pulse arrival times at the two ends of the detector, for each wire. This measurement

has a resolution of $\sim 4\,cm$, much cruder than the resolution obtained when using full stereo and axial super-layer information. The z-by-timing system is used for triggering and pattern recognition purposes.

The hits in the CTD are combined to reconstruct particle trajectories. The transverse momentum resolution for full-length CTD tracks is $\sigma(P_T)/P_T = 0.0058\,P_T \oplus 0.0065 \oplus 0.0014/P_T$ with P_T in **GeV**. Three different effects add to the total resolution: the error on the CTD hits, multiple scattering within the CTD and multiple scattering in dead material on the particle flight path before the CTD. The average resolution for a track that has passed through all superlayers is $\sim 180\,\mu m$ in $r\phi$ and $\sim 2\,mm$ in z. The event vertex measured by the CTD is accurate to about $1\,mm$ in xy and $4\,mm$ in z.

3.2.4 The luminosity monitor

At ZEUS, the delivered luminosity is measured using the Bremsstrahlung process $e\,p \rightarrow e\,p\,\gamma$. The cross-section σ_B for this process at fixed photon scattering angle θ_γ and energy E_γ is well known. The luminosity $\mathcal{L}$ is determined by measuring the photon rate N_γ:

$$\mathcal{L} = N_\gamma/\sigma_B \tag{3.4}$$

The measurement of N_γ is performed by a lead-scintillator calorimeter placed $107\,m$ away from the beamspot. It accepts photons arriving from the interaction point in an almost straight line ($\theta_\gamma < 0.5\,mrad$) with the electron beam. The error on this measurement is $\approx 3\%$. A large rate of synchrotron radiation and a high probability for multiple Bremsstrahlung photons in a single bunch crossing are two main sources that can spoil the accuracy of the luminosity measurement. In order to cope with these effects, the lead scintillator detector is shielded by three radiation lengths of uninstrumented lead, reducing the probability of detection of Bremsstrahlung photons and thus reducing the probability of photon overlap. This also shields the calorimeter from synchrotron radiation. A second independent luminosity monitor uses two separate calorimeters to measure coincidences of e^+e^- pairs coming from photon conversions $\gamma \rightarrow e^+ + e^-$ within the photon pipe exit window. Detection is low enough to avoid double counting and the calorimeters are placed outside the main synchrotron envelope. A schematic representation of the luminosity measuring system is shown in Fig. 3.10.

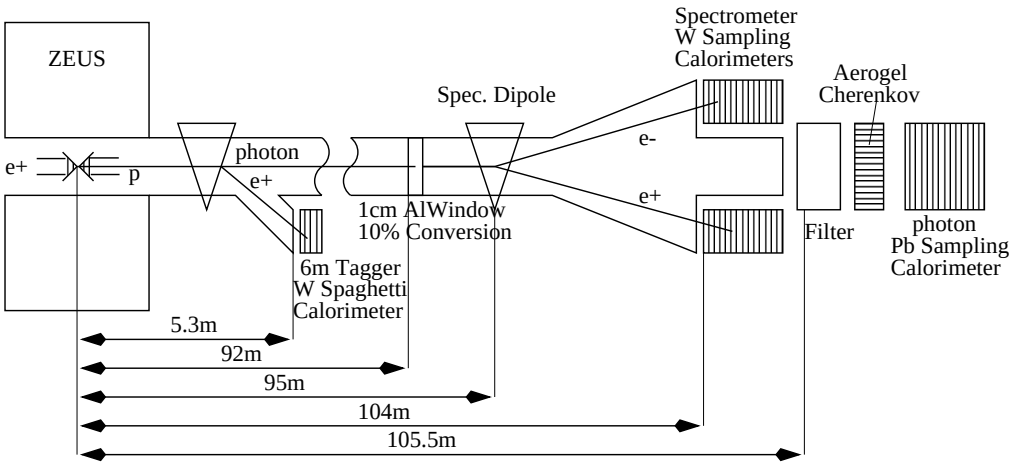


Figure 3.10. Schematic diagram of the luminosity measuring system.

3.3 The microvertex detector

3.3.1 Design

The detector closest to the interaction point is the microvertex detector (MVD). The MVD was installed during the 2000-2001 shutdown. Its main design purpose was to improve the tracking and vertexing provided by the CTD[2] such that cleaner heavy quark samples can be selected by tagging secondary decay vertices.

The MVD is a silicon strip detector. It consists of a 60 cm long barrel and 4 disks in the forward region. The MVD runs in Z along the beampipe, spanning in $r\phi$ the space between the beampipe up to the CTD inner layer. It is divided into two distinct regions: barrel (BMVD) and forward (FMVD), as seen in Fig. 3.11. The microvertex detector has the following specifications:

- a polar angle coverage of $7.6^\circ < \theta < 160^\circ$
- covering from 38 to 247 mm radially
- covering in z direction from -300 to 750 mm as measured from the ZEUS coordinate system origin
- 3(2) barrel layers and 4 forward planes
- intrinsic hit resolution of better than 20 μm

[2]The CTD in itself can recognize displaced vertexes only if they are more than $1.5-2\,\mathbf{cm}$ away from the primary vertex.

- 100 μm impact parameter resolution for completely transverse tracks ($\theta = 90°$) of at least 2 **GeV**
- noise occupancy[3] $< 10^{-3}$
- single hit efficiency $> 99\%$

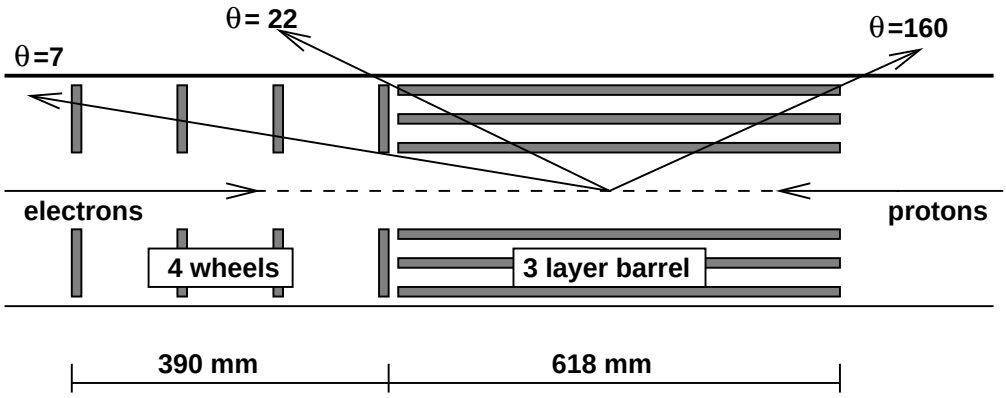


Figure 3.11. The MVD layout: a barrel region and forward wheels. Protons enter the beampipe from the right. The polar angle coverage is also shown.

The MVD is equipped with a total of 712 silicon strip sensors. The barrel region is organized into three cylinders and contains 600 sensors, each with an area of 41 cm^2. Another 112 sensors compose the four wheels. The layout of the MVD is given in Fig. 3.11. A picture of the lower half of the detector is shown in Fig. 3.12.

The semiconductor sensors are 300 μm thick and the bulk of the sensor is n-doped silicon. On one side, parallel p-doped strips are implanted at 20 μm distance from each other. The readout strips are 120 μm apart with 5 intermediate p-doped strips in between each pair. On the opposite side of the bulk, an aluminum back-plane provides a homogeneous electric field. Pairs of sensors are glued together to form a *half-module* as depicted in Fig. 3.13. The strips of the two sensors run perpendicular to one another. This is called $r\phi$ versus z ganging. When placed in the barrel MVD, one sensor provides z information and the other $r\phi$ information. Since the strips are connected with a bond wire, it cannot be determined a priori through which of the two sensors a particle has passed. This issue is resolved at the track reconstruction level. By combining information from other hits along the track as well as CTD measurements, the "ghost" hits can be distinguished from the real hits.

[3]This value changed considerably due to radiation damage.

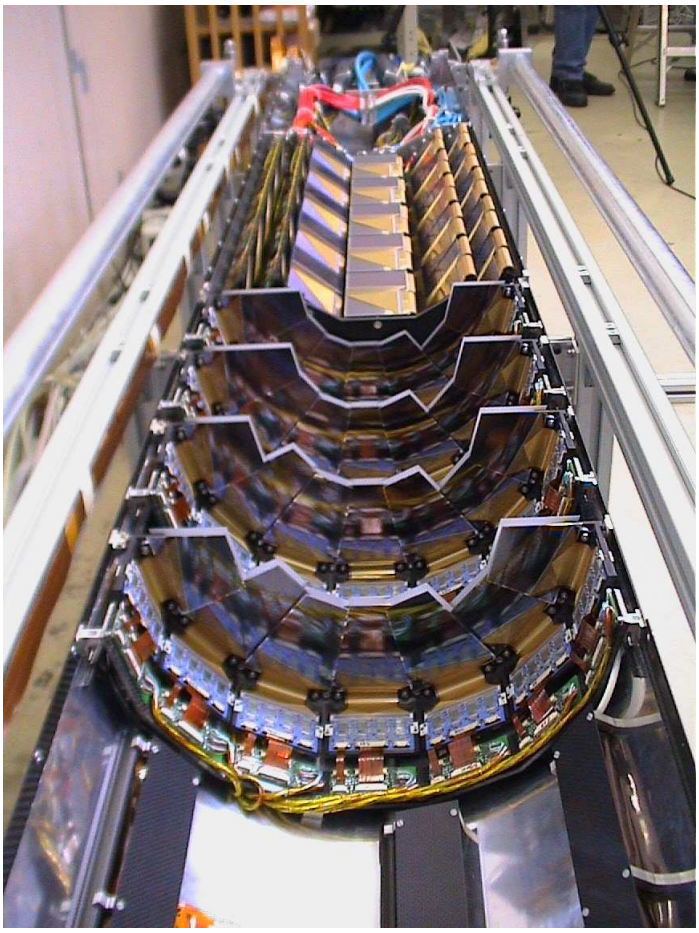


Figure 3.12. Half of the MVD detector. The different sensor layers are visible in the wheels (foreground) and in the barrel region (background). At the far end, the electric connections are seen.

By gluing half-modules together with mirror imaged half-modules, full modules are formed. With a full module it is possible to reconstruct a full three dimensional position on a particle trajectory. In the barrel MVD, 30 ladders are organized into 3 cylinders, as visible in Fig. 3.12. Five full modules are fixed to one ladder, as shown in Fig. 3.13. The ladder is made of carbon fiber composite and provides support not only for the modules but also for the electronics, the readout cables and cooling water pipes. A schematic view of the MVD barrel cylinders is presented in Fig. 3.14. The four forward wheels contain similar sensors, glued in pairs, back to back: an inner and an outer sensor form a wheel sector. The crossing angle between the strips in the inner and outer sensor is $180°/14$. The detailed design and tests of the silicon sensors are described elsewhere [19].

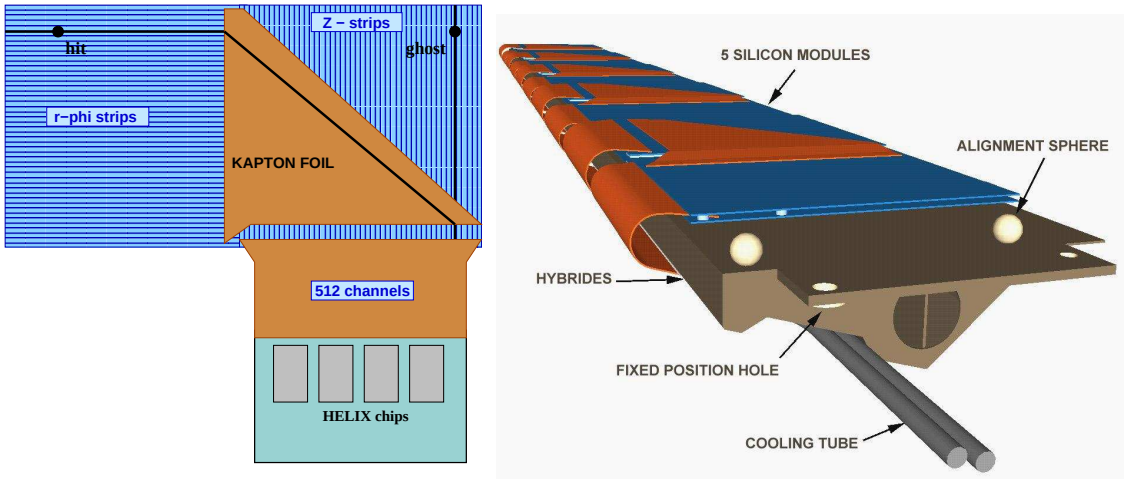


Figure 3.13. Two sensors are glued into a *half-module* (left figure). A full-module is obtained by combining one *half-module* with its mirror imaged *half-module.* In the right figure, five modules are fixed together to one ladder as part of the 30 ladders that compose the barrel MVD.

3.3.2 Detection principle and readout

The detection of charged particles crossing silicon is based on electron-hole pair generation. When a charged relativistic particle traverses matter, it looses energy. One of the main energy loss mechanisms is ionization. Inside a silicon crystal, the energy transferred to the semi-conductor is converted to electron-hole pairs[4]. The average energy necessary to produce such pair is approximately $3.6\,\mathbf{eV}$ for silicon. Typically, when crossing $300\,\mu m$ of silicon, a minimum ionizing particle (MIP) will generate about 27000 electron-hole pairs inside a $10\,\mu m$ thick tube centered around the track. The electrons and the holes will drift to opposite sides producing a readout signal. The electrons drift to the back-plane, as illustrated in Fig. 3.15.

The MVD employs Helix 3.0 readout chips. Each chip handles 128 input channels. All together, the MVD has 207360 readout channels[5]. The Helix chip amplifies the collected charge of each channel. The signal is shaped into a pulse with a height proportional to the collected charge. The width of the pulse is about $50\,ns$ and thus smaller than one bunch crossing time interval. The signal is transferred further by an analogue link to an ADC unit. The ADC unit can be programmed such that

[4]Inside semi-conductors, many more charge carriers are generated per unit of transversed distance than in a gaseous detectors. Silicon layers of hundreds of microns thickness are sufficient to produce reliable signals. Another advantage for using silicon detection is that the charge carriers move very quickly inside the silicon lattice producing a drift time of about $10\,ns$.

[5]About 5.7% of the total number of MVD channels were dead (in 2005) due to radiation damage.

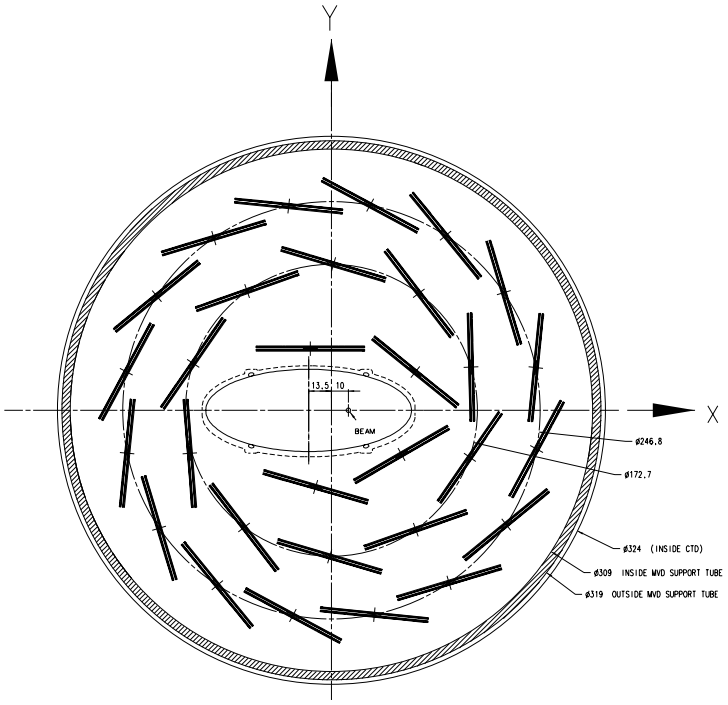


Figure 3.14. Schematic view of the MVD barrel ladders ($x - y$ cross-section).

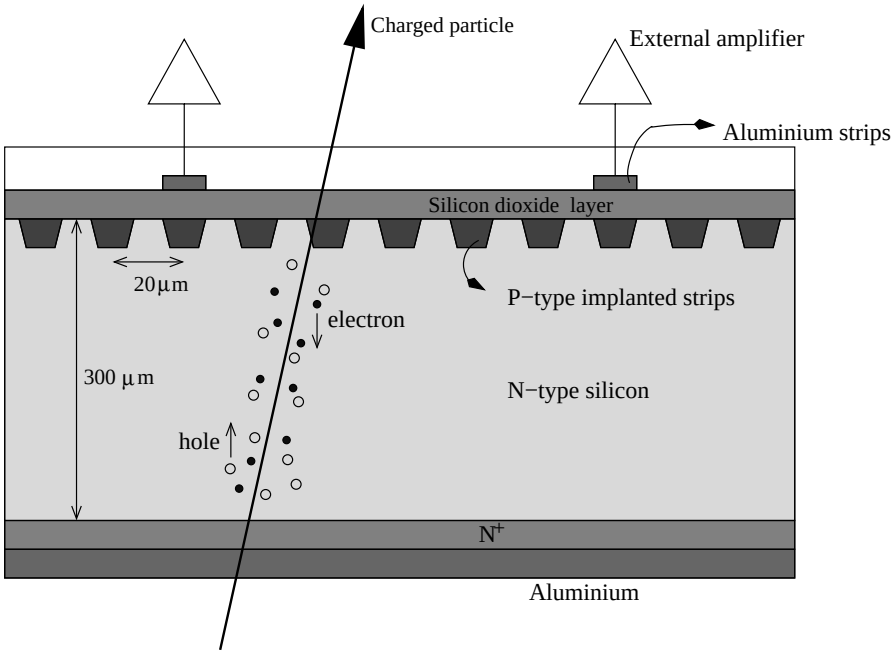


Figure 3.15. Schematic cross-section of one MVD strip sensor.

signals from individual strips or clusters are accepted only if they surpass a predefined threshold. The final data consists of strip addresses and the corresponding digitized signals.

3.3.3 Track reconstruction

ZEUS uses a combined track pattern recognition of the MVD and CTD. The pattern recognition makes full use of the track finding powers of the individual detectors and employs a multi-pass algorithm which combines the CTD, the forward and barrel MVD. This has lead to an increase in track finding efficiency for the CTD with information of the MVD. Seeds are constructed from super-layer information in both tracking detectors and the resulting trajectories are then used to find additional hits. In the MVD, separate seeds consisting of partly barrel and partly forward MVD hits are constructed to cover the transition area. Track candidates with ample hit information from both CTD and MVD are created; later in the procedure, tracks composed entirely from MVD or CTD hits are found, extending the range of track reconstruction. The information assigned to a track is then used as input to the Kalman filter based track fit[6]. The track fit computes more accurate track parameters and the corresponding covariances taking the dead material distribution of the MVD and of the beampipe into account. Outlier hits are rejected based on their χ^2 contribution to the fit (purification of the track). MVD information also improves considerably the momentum resolution at large momenta. This has a direct impact on particle mass resolution. A detailed description of the track fit procedure can be found in [22].

As particles pass through the beampipe and the MVD, they scatter in the presence of material and therefore lose energy and change direction. The average material that a particle goes through while transversing one ladder is about 3% of a radiation length. Fig. 3.16 is a scatter plot of the energy loss per unit of length dE/dx versus the track momentum, as reconstructed by the MVD. The kaon, pion and proton bands are visible. The effect of multiple scattering is mainly a change in direction, modeled by a gaussian distribution around zero angle change. The width of the distribution depends on the number of radiation lengths of the scattering medium as well as the momentum, velocity and charge of the scattered particle

[6]The Kalman filter[20] is a recursive algorithm that handles discrete data linear problems in a dynamic fashion. It is applied in many fields of science and engineering. The technique was introduced to high energy physics by R. Fruhwirth[21].

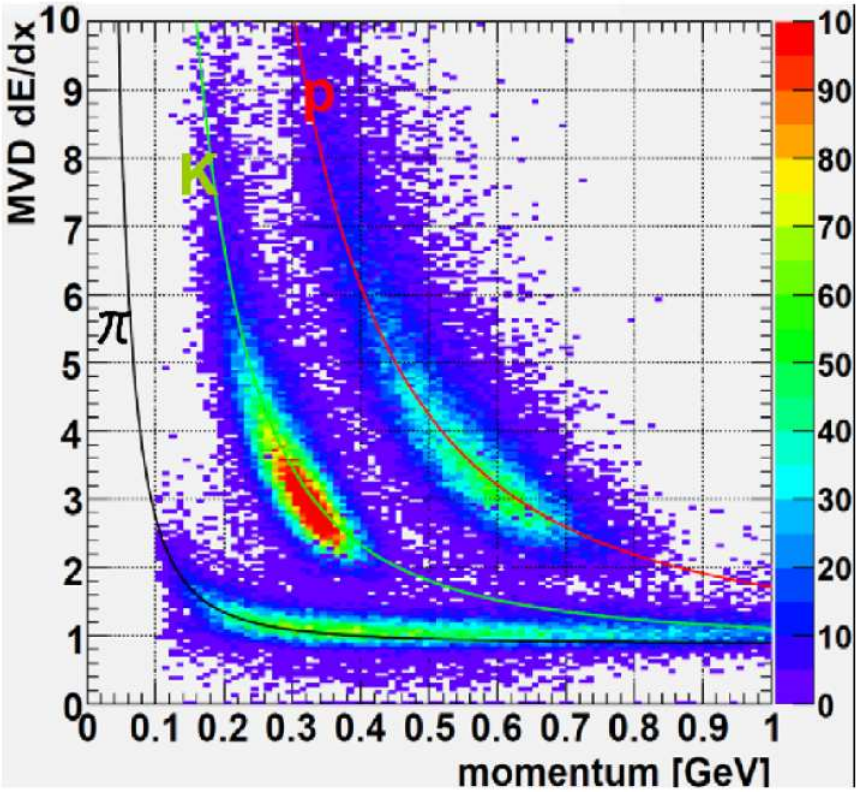


Figure 3.16. dE/dx distribution as reconstructed by the MVD. The kaon, pion and proton bands are visible.

as described in [22]. For high momentum tracks, the effects of multiple scattering become smaller. The track fit takes into account effects of both multiple scattering and ionization energy loss at each step in the fitting procedure by increasing the error matrix and modifying the track parameters. Especially the material of the beampipe leads to an unconstrained and therefore only estimated increase of these errors.

3.3.4 Alignment and resolution

Detector alignment plays a crucial role for the performance of highly resolution dependent analyses. For the MVD, a complex alignment strategy was adopted [23]. First, both the position of the half modules on the ladders and the position of the ladders inside the completed detector were surveyed with a precision of 10 μm during construction [24]. Second, a laser alignment system has been incorporated in the overall support structure to monitor the stability of the MVD during operation [25, 26]. Different stability periods are defined for more specific track based alignment

procedures, using both cosmic and *ep*-collision tracks. The data from the laser alignment system indicate that the position of the MVD as a whole is stable relative to the global ZEUS coordinate system with a precision of approximately 10 μm. It has been shown [23] that when the HERA machine magnets nearest to the ZEUS detector are operational, the associated minor movements of these magnets are also transmitted to the MVD. This effect, taken into account in the alignment, is due to the fact that the MVD service and readout cables are wrapped tightly around the final HERA magnet.

In 2003, the barrel MVD has already been aligned using cosmic muons. Cosmic muons cross both halves of the MVD therefore providing on the average 12 measurements in a very clean environment. An iterative two step method was used both to determine the global displacement of the entire MVD relative to the ZEUS coordinate system as well as to align ladders relative to each other. Statistics were not sufficient to align individual half-modules with the cosmic data. The residual of a hit is defined as the difference between the hit position as predicted by the track fit and the measured hit position. The width of the residual distributions quantifies the alignment precision of each ladder within the MVD frame. The residual distribution widths are shown for all ladders in Fig. 3.17, after cosmic alignment. The widths are about 60 μm for the $r\phi$ and and 40 μm for the rz sensors respectively. Horizontal ladders benefit the most from cosmic alignment. The forward MVD cannot be aligned with cosmics due to lack of coverage.

Recent progress was made by using tracks from the *ep* collisions recorded during 2004 and 2005. All the ladders were aligned with the same accuracy: residuals are about 25 μm. The improvement in residual distribution, both in width and shape, for ladders 8, 9, 21 and 22 is shown in Fig. 3.18. The width has been reduced from about 40 μm before the *ep*-alignment to 22 μm after *ep*-alignment. There were no significant differences found between sensors within the same ladder.

A direct consequence of the $e^{-}p$ alignment is an improvement in the impact parameter resolution of tracks with respect to the interaction point. This resolution becomes, on the average, about 100 μm. Given the long track extrapolation length, as well as multiple scattering, this value is the best possible. A radar plot in $r\phi$ of the impact parameter resolution is given Fig. 3.19. The impact parameter is plotted after cosmic alignment and after the subsequent *ep* track alignment. This is compared to the simulated impact parameter, with an ideally aligned detector, as reconstructed by the track fit routine. Vertical ladders benefit most from *ep* track

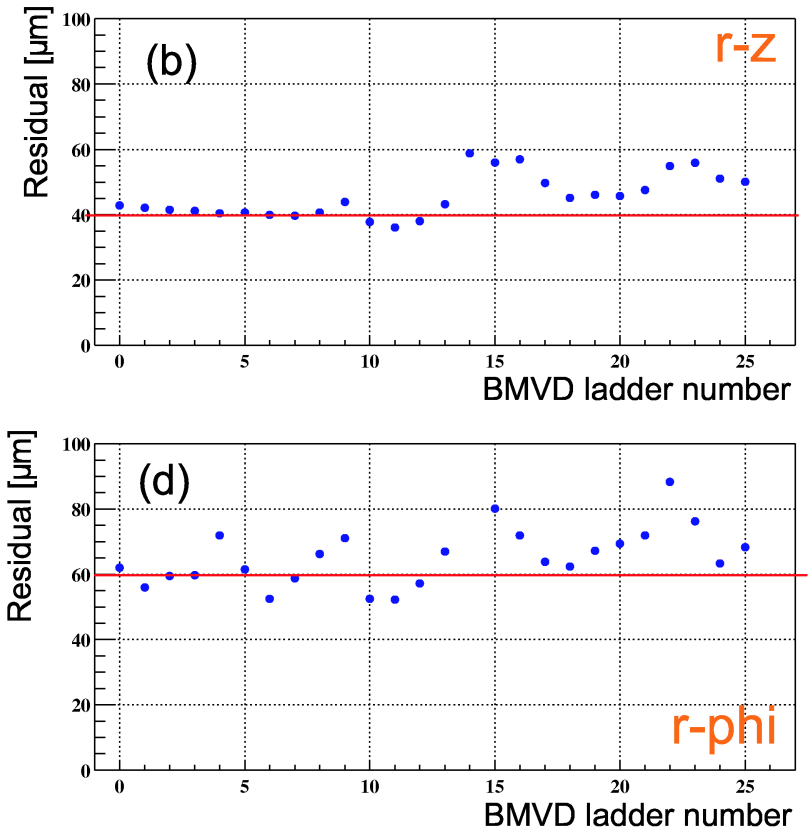


Figure 3.17. Hit residuals per ladder for all barrel MVD ladders for rz sensors (top) and $r\phi$ sensors (bottom), after applying cosmic alignment.

alignment. The regions $\phi \in [340°, 0°]$, $\phi \in [70°, 90°]$ and $\phi \in [150°, 220°]$ display a worse impact parameter resolution due to the geometry of the detector: there, only two cylinders of sensors are available, as shown in the left plot. A significant improvement in impact parameter resolution can be seen when using alignment from ep-collisions as opposed to the cosmic alignment.

The entire data set used in this thesis has been reprocessed using the final ep-alignment.

3.3.5 Vertex reconstruction and resolution

Initially, tracks are reconstructed independently. Afterwards, a search for common vertices is performed. The mathematical formalism employed for vertex reconstruction is described in detail in Chap. 4. After vertexing, tracks belong either to the primary/secondary vertices or are labeled as "non-vertex" tracks if they do

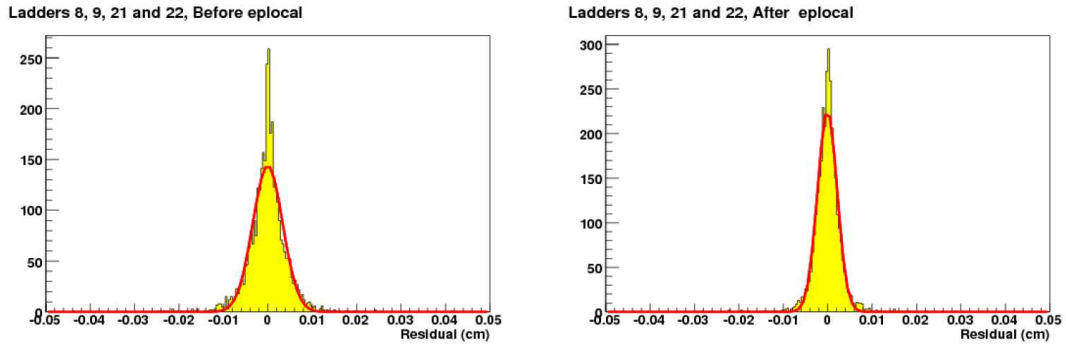


Figure 3.18. Residual distribution of four ladders near $\phi \simeq 180°$ before (left) and after alignment (right) with tracks from ep collisions. The fitted gaussian curves has a width of 40 μm (left) and 22 μm (right).

not fit to any of the reconstructed vertices. Because of the ZEUS magnetic field, the track momentum vector is meaningful only for tracks whose origins are known. This statement has great implications for charm and beauty meson tagging. The invariant mass of two tracks is calculated as:

$$m^2 = (E_1 + E_2)^2 - (\vec{P_1} + \vec{P_2})^2 \tag{3.5}$$

where $\vec{P_1}$ and $\vec{P_2}$ are the momentum vectors of the two tracks. The invariant mass of the mother particle depends strongly on the two momentum values as well as on their orientation. For an accurate invariant mass reconstruction, the momentum vectors need to be given at the decay vertex. All tracks used in invariant mass reconstruction ought to be vertexed tracks.

In ZEUS, several software packages are used for independent vertex finding and fitting. In particular, the vertex reconstruction package *kfvertex* [22] finds the best primary vertex candidate as well as secondary vertices. A standard *kfvertex* reconstruction is provided: for each event, complex information about the primary vertex, beam spot, secondary vertices, tracks associated to these, jets, calorimeter energy clusters, etc is available [27]. Seen with MVD precision, the standard vertex reconstruction is not sufficient to detect all possible event topologies automatically. For this reason, many precision vertexing analyses employ context-dependent revertexing techniques at the analysis level. The drawback is that these techniques demand high processing power and CPU time.

For this analysis, specific revertexing is employed for each event, as detailed in

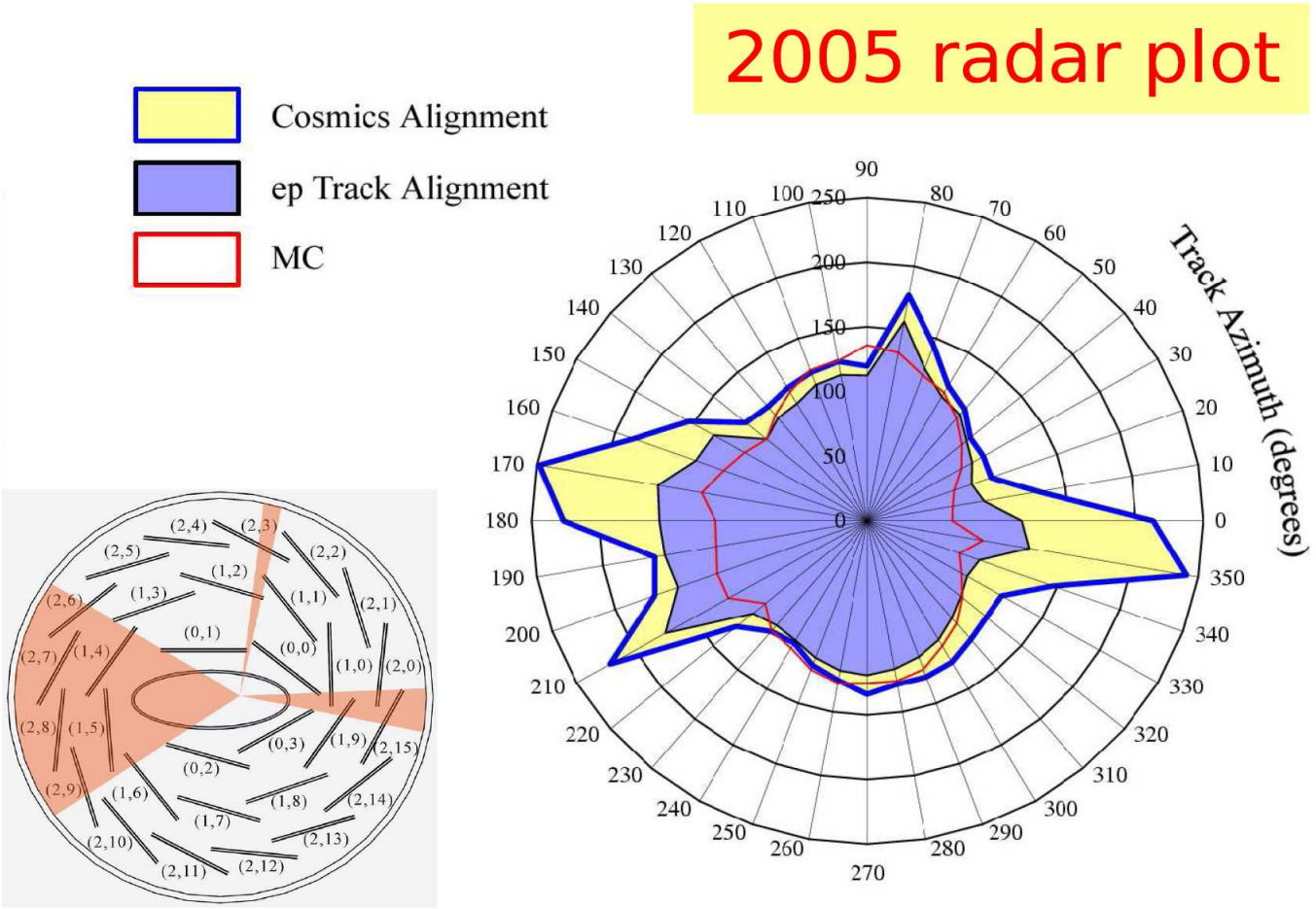


Figure 3.19. The impact parameter resolution in radar view ($r\phi$), for data recorded in 2005. The radial scale displays μm. The alignment with *ep* collision tracks improves the impact parameter resolution w.r.t. the cosmic alignment, especially for vertical ladders. The impact parameter reconstructed by the track fit routine in Monte Carlo simulations, with an ideal detector, is also shown. The asymmetry in $r\phi$ is due to the geometry of the MVD: three angular regions corresponding to two cylinders are shown in the left plot.

Chap. 6. Tracks with $p_T > 0.125\,\mathbf{GeV}$, with at least 2 MVD hits and crossing at least 2 CTD superlayers, are vertexed as follows. A search for the best two track secondary vertex candidate with a $\chi^2 < 16$ as calculated in eq. 4.6 is performed, for all track pairs. The remaining tracks in the event are vertexed separately. First, a seed for the primary vertex is chosen, by searching for track intersections in the XY plane. An attempt is made to add tracks to this vertex, one by one. This process is repeated with subsequent seeds until all tracks have been matched to the vertex or cannot be matched at all. In this way, a list of primary vertex candidates is made. The best candidate is chosen as the event vertex by considering the number of tracks that fit to it and the total χ^2 per number of degrees of freedom. The position of the primary vertex in the transverse XY plane is cross-checked with the position of the beamspot in a loose manner. The transverse positions of all reconstructed vertices for a small sample of 100k events recorded in 2005 are shown in Fig. 3.20. The mean values, also defining the beamspot position, are $13.35\,mm$ in X and $2.19\,mm$ in Y. The beamspot position can move as much as $100\,\mu m$ during the same fill. The primary vertex width is about $220\,\mu m$ in X and $180\,\mu m$ in Y. This width is dominated by the slope at which the beams traverse the detector. The secondary vertex distributions have slightly larger widths due to the fact that they were constructed from 2 tracks only whereas the primary vertex consists of 5-10 tracks. The secondary vertex distributions show also non-gaussian tails associated with genuine decays away from the primary vertex. All vertices are reconstructed with no mass constraint.

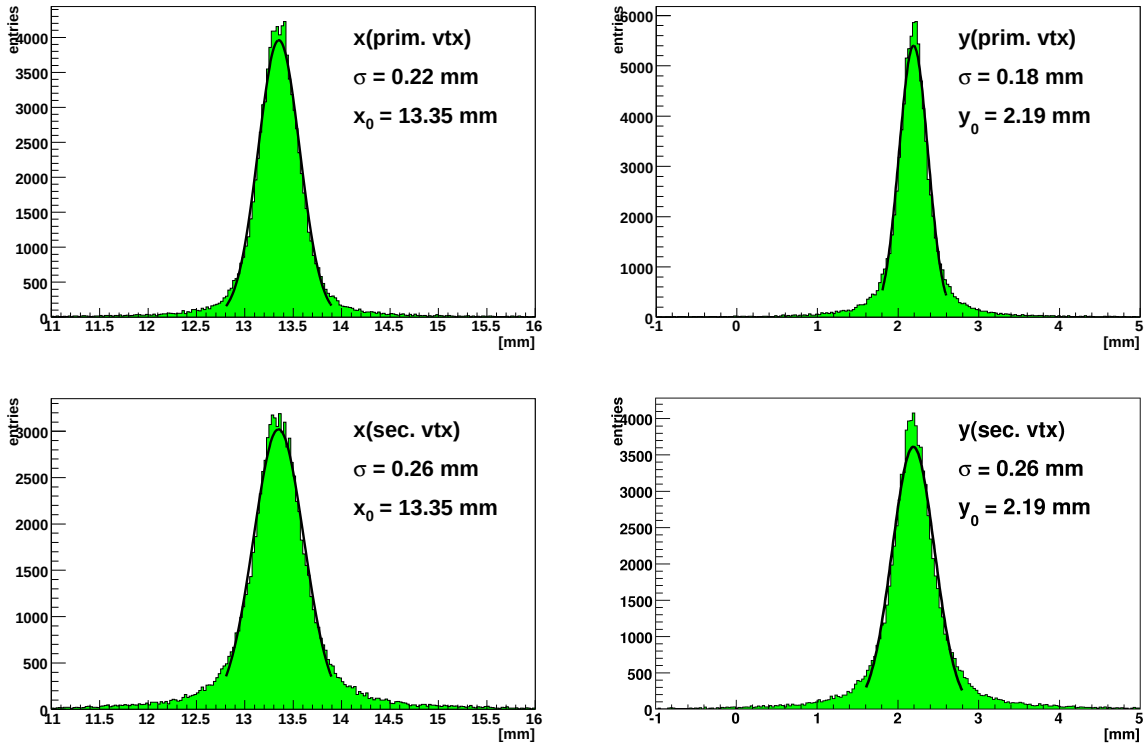


Figure 3.20. The X (left plots) and Y (right plots) positions of reconstructed primary vertices (upper plots) and of secondary vertices (lower plots) for a sample of 100k events, 2005 $e^{-}p$ data. The mean values are $13.35\,mm$ in X and $2.19\,mm$ in Y. The primary vertex width is about $220\,\mu m$ in X and $180\,\mu m$ Y. The secondary vertex distributions have larger widths and exhibit wider tails associated with genuine decays.

Chapter 4

Vertexing

Vertexing is a central issue of this analysis. As explained in Sec. 3.3, the momentum of a helix track is meaningful only at a given point along the helix. Because the invariant mass of two tracks depends directly on the value and orientation of each of the two track momenta, a rigorous approach to vertex reconstruction is taken. In this chapter, the mathematical formalism of vertexing will be described. The material presented here is described in greater detail in [28]. This formalism has been "coded" in the ZEUS vertexing package *kfvertex*. Also, tagging variables such as decay length, decay length significance, impact parameter and lifetime are defined.

4.1 Track parametrization

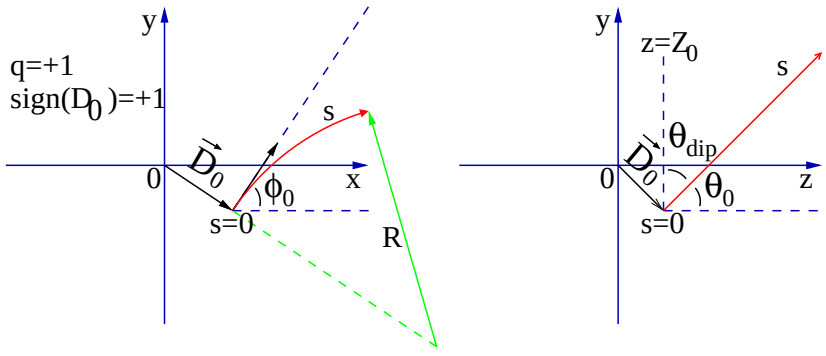


Figure 4.1. Projection of a helix track onto the XY plane (left) and YZ plane(right). The helix parameters are explicitly shown.

The trajectory of a charged particle inside an uniform magnetic field is a helix. The velocity $v_{\parallel}$ of the particle along the field lines determines the *step* of the helix;

the radius of the helix circle is proportional to $v_\perp$, the component of the particle velocity perpendicular to magnetic field lines. ZEUS employs a right handed coordinate system with colliding particles moving along the OZ axis, with the protons advancing in the positive direction. A uniform magnetic field of strength $B_z \simeq 1.43$ Tesla is aligned along the OZ axis: $\vec{B} = (0, 0, B_z)$.
The trajectory of each track in the global Cartesian (x, y, z) frame is parametrized by a 5-vector helix:

$$\vec{p} = (W, T, \phi_0, D_0, Z_0) \tag{4.1}$$

W is the signed curvature of the track: $W = q/R$, with q the charge and R the helix radius. The parameter T is defined as $tan(\theta_{dip})$ where θ_{dip} is the complement of the polar angle θ_0. The Cartesian vector $\vec{D_0}$ points to the point on the helix closest to the z-axis. The scalar D_0 is then the distance of closest approach to the z-axis and Z_0 the z coordinate of the helix at this point. Projecting the helix onto a circle in the xy plane, ϕ_0 is the angle that $\vec{t}$, the tangent of the (projected) track at the point of closest approach, makes with the x axis. D_0 is signed. The sign is defined by the sign of the cross-product $\vec{D_0} \times \vec{t}$. Fig. 4.1 shows the projection of a helix track onto the xy plane (left plot) and yz plane (right plot) for a positively charged particle.

One can now recover the position (x, y, z) of any point along the helix using the above parametrization:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \vec{D_0} + \begin{pmatrix} \frac{1}{W}\cos(\phi_0)\sin(Ws_\perp) + \frac{1}{W}\sin(\phi_0)[1-\cos(Ws_\perp)] \\ \frac{1}{W}\sin(\phi_0)\sin(Ws_\perp) - \frac{1}{W}\cos(\phi_0)[1-\cos(Ws_\perp)] \\ Ts_\perp \end{pmatrix} \tag{4.2}$$

Here, $s_\perp = s\sin(\theta)$ is the distance along the helix, measured from the point of closest approach.

The *trackfit* routine gives the best possible estimate of five track parameters and their uncertainties, based on a collection of hits inside the tracking detectors. A detailed description of how *trackfit* is implemented can be found in [22]. The estimated uncertainties of each of the five helix parameters are organized in a 5×5 covariance matrix V. Once reconstructed, a track is completely described by a 5-vector $\vec{p}$ in the helix parametrization and its covariance matrix V and can be further manipulated as such.

4.2 Kalman filter and vertex reconstruction

4.2.1 The mathematical formalism

One starts always with an estimate of the vertex[1]. Then, this estimate is compared to the information from a random[2] track in the event. If the track is consistent with coming from the vertex, the track information is added, yielding a new estimate. This procedure is then iterated for each track in the event and that leads to a final vertex estimate. This technique is called *filtering*. By *smoothing*, the momentum of each particle is recomputed at the final vertex position. A notation is defined:

x_k = estimate of the vertex position after *adding* k tracks to vertex

x^t = true vertex position

C_k = cov(x_k)

q_k = estimate of the momentum of particle k at x_k

D_k = cov(q_k)

E_k = cov(x_k,q_k)

m_k = measurement vector (five measured helix parameters)

v_k = measurement noise (disturbance in measuring m_k)

V_k = cov(v_k)

$G_k = V_k^{-1}$ = weightmatrix of particle k

The measurement equation is a map h of the true vertex position x^t and true momentum q_k^t of track k to the measured parameters m_k distorted by the noise v_k:

$$m_k = h_k(x^t, q_k^t) + v_k \tag{4.3}$$

[1] A vertex estimate can come from various sources. One good start could be the beam spot. In ZEUS, the *kfvertex* routine uses overlap points of track pairs, in the xy plane, as seeds for primary vertex finding.

[2] By random track it is understood here any track of a given event that passed certain quality cuts depending on the transverse momentum, number of MVD hits or CTD superlayers, η distribution, etc.

Kalman filtering can be applied only to linear measurement equations. Therefore, $h_k(x^t, q^t)$ is approximated by a first order Taylor expansion:

$$h_k(x^t, q_k^t) \approx h_k(x^{(0)}, q_k^{(0)}) + A_k(x^t - x_k^{(0)}) + B_k(q_k^t - q_k^{(0)}) = c_k^{(0)} + A_k x^t + B_k q_k^t \quad (4.4)$$

with

$$\begin{aligned} A_k &= \tfrac{\partial h_k}{\partial x^t}\big|_{(x_k^{(0)}, q_k^{(0)})}, B_k = \tfrac{\partial h_k}{\partial q_k^t}\big|_{(x_k^{(0)}, q_k^{(0)})} \\ c_k^{(0)} &= h_k(x^{(0)}, q_k^{(0)}) - A_k x_k^{(0)} - B_k q_k^{(0)} \end{aligned} \quad (4.5)$$

The vertex reconstruction proceeds as follows: choose a starting value x_0 for the vertex position and a covariance matrix. This can depend on information that one has before hand. A blind choice would be the origin of the coordinate system and a 3×3 covariance matrix with infinite quantities on the diagonal. A more fortunate choice is the beam spot and its standard deviations. Next, recompute an estimate x by adding the weighted information of track 1 [3]. This is done by the so called χ^2 *minimizing* method.

$$\begin{aligned} \chi^2_{KF}(x, q) &= (x - x_0)^T (C_0)^{-1} (x - x_0) + \\ &+ (m_1 - c_1^{(0)} - A_1 x - B_1 q)^T G_1 (m_1 - c_1^{(0)} - A_1 x - B_1 q) \end{aligned} \quad (4.6)$$

The position x and the momentum q that minimize this χ^2_{KF} are the first guesses for x_1 and q_1 at the new vertex. One solves $\partial\chi^2_{KF}/\partial x = 0$ and $\partial\chi^2_{KF}/\partial q = 0$ and finds:

$$\begin{aligned} x_1 &= C_1 \left[(C_0)^{-1} x_0 + A_1^T G_k^B (m_1 - c_1^{(0)}) \right] \\ q_1 &= W_1 B_1^T G_1 (m_1 - c_1^{(0)} - A_1 x_1) \\ C_1 &= \left((C_0)^{-1} + A_1^T G_1^B A_1 \right)^{-1} \\ D_1 &= W_1 + Q_1 B_1^T G_1 A_1 C_1 A_1^T G_1 B_1 W_1 \\ E_1 &= -C_1 A_1^T G_1 B_1 W_1 \end{aligned} \quad (4.7)$$

[3] Note that the point $x^{(0)}$ around which we Taylor expand need not be the same as the vertex start value x_0. One choice is the point on the track closest to x_0.

with

$$W_1 = (B_1^T G_1 B_1)^{-1} \quad , \quad G_1^B = G_1 - G_1 B_1 W_1 B_1^T G_1$$
$$cov(x_1) = C_1 \ , \ cov(q_1) = D_1 \ , \ cov(x_1, q_1) = E_1$$

D_1 is called momentum covariance and E_1 cross-covariance.
The new (updated) vertex is located at x_1 with the covariance C_1. We can now add track 2 and so on. In this way, an iterative procedure adding track k will lead to the same answers as in Eq. (4.7) but with the lower index 1 replaced by k (and 0 by $k-1$). *Smoothing* track k to the new vertex changes its momentum:

$$q_k^s = W_k B_k^T G_k (m_k - c_k^{(0)} - A_k x_k) \tag{4.8}$$

where s upper index stands for *smoothed*. This step is essential for accurately determining the invariant mass at the vertex. After smoothing all tracks to the final vertex position, the covariance matrix Q for the correlation between the smoothed momenta is recovered. For instance, the cross-momentum correlation covariance matrix between tracks i and j after smoothing has the form:

$$Q_{i,j} = cov(q_i^s, q_k^s) = W_i B_i^T G_i A_i C_f A_j^T G_j B_j W_j \tag{4.9}$$

with C_f being the final vertex covariance matrix. Note that $Q_{i,j} = Q_{i,j}^T$.

So far, the formalism is valid for any track parametrization. In ZEUS, charged tracks parametrized as in Eq. 4.1 are vertexed in this way.

4.3 Expanding the tracking package

The *trackfit* routine can only reconstruct tracks that have crossed the detector. The *kfvertex* routine has been modified for this analysis such that new tracks can be constructed, associated to particles which decayed before reaching the MVD. A pseudotrack is defined as a composite object made from a vertex and several charged daughter tracks emerging from that vertex. The pseudotrack represents the mother particle which decayed at that vertex. A careful treatment of the errors involved leads to a completely determined object parametrized by a 5-vector in the helix parametrization and its covariance matrix. Furthermore, one could employ

the pseudotrack (a real particle) to search for new decay vertices at which the reconstructed mother particle had participated as a daughter/decay product.

The method will become more transparent if a specific parametrization is chosen. The main focus of this thesis is the neutral pseudotrack (neutral mother particle decaying in two charged tracks). The treatment of charged pseudotracks is similar [29](to be published).

4.3.1 Neutral Pseudotracks

As the curvature is not defined for neutral tracks moving in straight lines in a magnetic field, the parametrization from Eq. 4.1 presents a problem. Therefore, the following parameters are chosen for the measurement vector m :

$m(1) = p$ $(0, +\infty)$ absolute value of the total momentum
$m(2) = \phi$ $(-\pi, \pi)$ azimuth angle at DCA
$m(3) = \theta$ $(0, \pi)$ polar angle
$m(4) = dca$ $(-\infty, +\infty)$ distance of closest approach
$m(5) = z$ $(-\infty, +\infty)$ z at DCA

The geometrical momentum vector $\vec{q}$ is defined as follows:

$q(1) = p$ $(0, +\infty)$ absolute value of the total momentum
$q(2) = \phi$ $(-\pi, \pi)$ azimuth angle at DCA
$q(3) = \theta$ $(0, \pi)$ polar angle

The vector $\vec{p} = (p_x, p_y, p_z)$ is called the kinematical momentum of a particle.

Although the geometry of the neutral track does not depend on its momentum, its errors do. This happens essentially because the errors of the daughter track momenta translate into an error on the vertex position which will feed into the error of the pseudotrack. The calculation of the pseudotrack parameters is done in three steps:

(a) Determine the kinematical momentum of the mother from the fitted geometrical momenta of the daughters at the vertex

(b) Calculate the geometrical momentum of the mother particle from its kinematical momentum. This is necessary as it is part of the neutral track parametrization.

(c) Calculate the measurement parameters of the mother particle at the DCA as well as its covariance matrix.

In step (a), for each of the two charged daughters, one has:

$$p_x = p_t \cos(\phi) \ , \ p_y = p_t \sin(\phi) \ , \ p_z = p_t \cot(\theta) \tag{4.10}$$

So a mapping $\vec{p}_{a,b} = f(\vec{q}_{a,b})$ exists. The Jacobian propagation matrix is defined as $W_{a,b} = (\partial f / \partial q_{a,b})|_{q^{(0)}_{a,b}}$. With $\vec{p}_a$ and $\vec{p}_b$ calculated, the pseudotrack momentum is simply:

$$\vec{p}_{tot} = \vec{p}_a + \vec{p}_b \tag{4.11}$$

In step (b) the geometrical momentum of the mother particle is computed, starting from $\vec{p}_{tot}$. Using $\vec{p}_{tot} = (p_x, p_y, p_z)$, the mapping $\vec{q}_{tot} = g_n(p_{tot})$ takes the form:

$$\begin{aligned} p = \sqrt{p_x^2 + p_y^2 + p_z^2} \ , \quad \phi = \left(\frac{p_x}{\sqrt{p_x^2 + p_y^2}} \right) \ , \\ \theta = \arccos \left(\frac{p_x}{\sqrt{p_x^2 + p_y^2 + p_z^2}} \right) \end{aligned} \tag{4.12}$$

The Jacobian matrix W_n^- for the error propagation and its inverse are defined as:

$$W_n^- = \left. \frac{\partial \vec{g_n}}{\partial \vec{p}} \right|_{p^{(0)}_{tot}} = (W_n)^{-1} \tag{4.13}$$

Finally, all entries of the neutral track "measurement" vector m_n are known. Fig. 4.2 illustrates how parameters measured at the decay vertex P_0 are translated into a measurement vector given w.r.t. the point of closest approach. Using the notation of Fig. 4.2, it holds that:

$$\begin{aligned} m_n = h_n(x_0, y_0, z_0; p_0, \phi_0, \theta_0) \\ p = p_0 \ , \quad \phi = \phi_0 \ , \quad \theta = \theta_0 \ , \quad dca = r_\perp \sin(\xi) \\ z = z_0 - r_\perp \cos(\xi) \frac{1}{\tan(\theta)} \end{aligned} \tag{4.14}$$

where:

$$r_\perp = \sqrt{x_0^2 + y_0^2} \ , \ \phi_0' = \arccos \left(\frac{x_0}{r_\perp} \right) \ , \ \xi = \phi_0 - \phi_0' \tag{4.15}$$

The covariance matrix of the pseudotrack has to be calculated by a proper error propagation. Let us write down again the names and the physical meaning of each of the matrices involved:

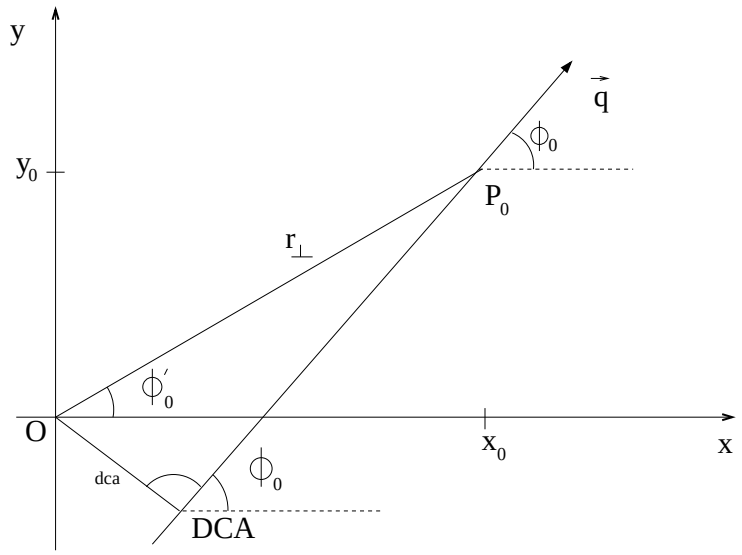


Figure 4.2. Pseudotrack projection on the xy plane. Parameters measured at decay vertex P_0 are expressed in Eq. 4.14, 4.15 in terms of parameters expressed at DCA.

$C_f = cov(x_f)$**,** covariance matrix of the smoothed vertex position

$E_i = cov(x_f, q_i^s)$**,** for daughter tracks

$Q_{i,j} = cov(q_i^s, q_j^s)$**,** correlation between tracks , $i \neq j$

$Q_{i,i} = cov(q_i^s, q_i^s) = D_i$**,** momentum covariance for each daughter track

A**,** B first order Taylor expansion matrices for the pseudotrack w.r.t. the point of closest approach, as given in Eq. 4.5.

We define:

$E = \sum_i E_i W_i^T$

$Q = \sum_{i,j} W_i Q_{i,j} W_j^T$

Then, the covariance matrix of the mother particle is:

$$C_{pseudo} = AC_f A^T + AE(W^{-1})^T B^T + BW^{-1}E^T A^T + BW^{-1}Q(W^{-1})^T B^T \quad (4.16)$$

The 5×5 covariance matrix C_{pseudo} is symmetric and positive definite.

4.3.2 The decay length significance

A neutral particle created at the (primary) vertex $\vec{v_1}$ decays later at the secondary vertex $\vec{v_2}$ into two oppositely charged daughter particles, as pictured in Fig. 4.3. The decay length vector is defined as:

$$\vec{L} = \vec{v_2} - \vec{v_1} \tag{4.17}$$

The decay length is then $L = |\vec{L}|$. Let $\vec{1_L}$ be a unit vector in the direction of $\vec{L}$: $|\vec{1_L}| = 1$. The error of the decay length is then computed by projecting the covariance matrices C_1 and C_2 of $\vec{v_1}$ and $\vec{v_2}$ respectively along the unit vector $\vec{1_L}$:

$$\Delta L = \sqrt{(\vec{1_L})^T \cdot C_1 \cdot \vec{1_L} + (\vec{1_L})^T \cdot C_2 \cdot \vec{1_L}} \tag{4.18}$$

Due to measurement errors, the momentum of the mother particle at the decay vertex, $\vec{p_M} = \vec{p_1} + \vec{p_2}$ does not always point in the same direction as the decay length vector $\vec{L}$. In order to separate poorly reconstructed decays in which the mother particle seems to be coming back towards the origin vertex $\vec{v_1}$, equivalent to $\vec{L} \cdot \vec{p_M} < 0$, from the more "physical" decays in which the mother particle moves away from $\vec{v_1}$, with $\vec{L} \cdot \vec{p_M} > 0$, the decay length is signed:

$$L \rightarrow Sign(\vec{L} \cdot \vec{p_M}) \cdot L \tag{4.19}$$

The decay length significance is the ratio of the signed decay length and its error:

$$\sigma_{D.L.} = Sign(\vec{L} \cdot \vec{p_M}) \frac{L}{\Delta L} \tag{4.20}$$

4.3.3 Impact parameter significance

The impact parameter d of the pseudotrack w.r.t. the primary vertex is a useful tool which discriminates between particles created at the interaction vertex or elsewhere. It is defined as the distance of closest approach of the pseudotrack momentum line to the primary vertex v_1, as shown in Fig. 4.3. The following holds:

$$\cos\gamma = \frac{\vec{L} \cdot \vec{p_M}}{|L|\,|p_M|} \tag{4.21}$$

with γ the angle opposing the impact parameter. Then, d is computed as:

$$d^2 = L^2 - \frac{(\vec{L} \cdot \vec{p_M})^2}{|p_M|^2} \tag{4.22}$$

where the relations $\sin\gamma = d/L$ and $\sin^2\gamma + \cos^2\gamma = 1$ were used.

If $\vec{1_d}$ is a unit vector pointing in the same direction as the impact parameter, then the impact parameter error is : $\Delta d = \sqrt{(\vec{1_d})^T \cdot C_1 \cdot \vec{1_d} + (\vec{1_d})^T \cdot C_2 \cdot \vec{1_d}}$ with C_1, C_2 covariance matrices of the primary and secondary vertices. The significance of the impact parameter is simply: $\sigma_d = d/\Delta d$ and is a positive quantity.

4.3.4 Lifetime

The lifetime of a particle that has the momentum $\vec{p}$, velocity $\beta = v/c$ and decay length $\vec{L}$ in the laboratory frame is:

$$c\tau = \frac{L}{\beta\gamma} = m\frac{L}{|p|} = m\frac{L_{xy}}{P_T} \tag{4.23}$$

where $\gamma = 1/\sqrt{1+\beta^2}$, P_T is the transverse momentum and L_{xy} the projection of the decay length vector $\vec{L}$ on the xy plane. The last equality exploits the fact that the ZEUS detector has a better resolution in the transverse plane than along the beam axis.

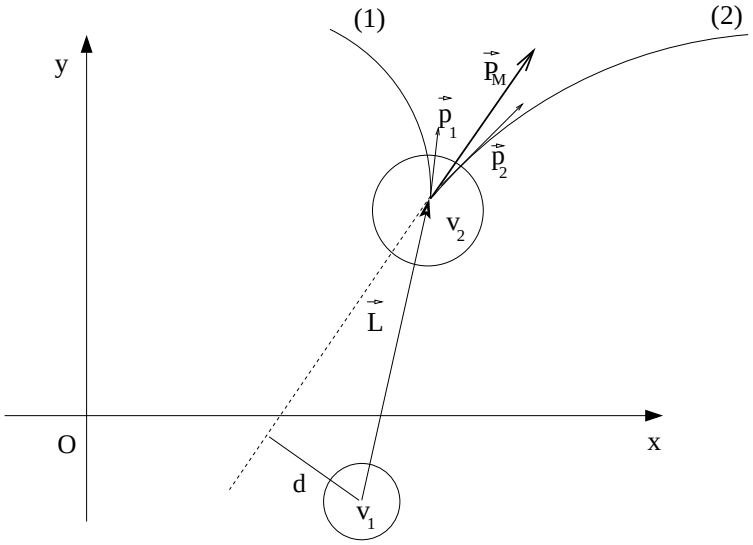


Figure 4.3. A neutral particle moves away from vertex v_1 and decays later on at vertex v_2. The particle momentum $\vec{p_M}$ at v_2, the decay length vector $\vec{L}$ and the impact parameter d of the particle w.r.t. to (creation) vertex v_1 are shown.

Chapter 5

Event selection in DIS

In this chapter, the ZEUS data acquisition flow is explained. The signature of a DIS event is described. Various methods employed to reconstruct the event kinematics are compared. The DIS event selection cuts and data selection utilizing the three level ZEUS trigger chain are presented.

5.1 Data acquisition flow

The ZEUS data acquisition uses a three level trigger system in order to reduce the huge volumes of data streaming from the detector. The HERA bunch crossing rate is high: every $96\,ns$, a proton bunch collides with an electron bunch possibly giving an $e^{-}p$ interaction. On average, one in one hundred bunch crossings gives a detectable event. Therefore a fast readout is essential. The raw data coming from $\sim$ 250.000 channels of the detector is approx. 0.5 Mb in size and, given that final event data is written out at speeds close to 1 Mb/sec, the total amount of data needs to be drastically reduced in early stages of data acquisition. Treated as a black box, the trigger chain effectively reduces the event rate from a 100 kHz input down to 5 Hz such that end storage facility can handle the data flow. The challenge is to identify the 5 most interesting events from hundreds of thousands of detectable collisions every second. At each level in the trigger chain, data volume is reduced, giving the system time to compute more characteristics of each accepted event at the next level of the trigger. Each subsequent level can therefore perform more sophisticated filtering of the data. The schematic description of the data flow and event rate reduction is given in Fig. 5.1.

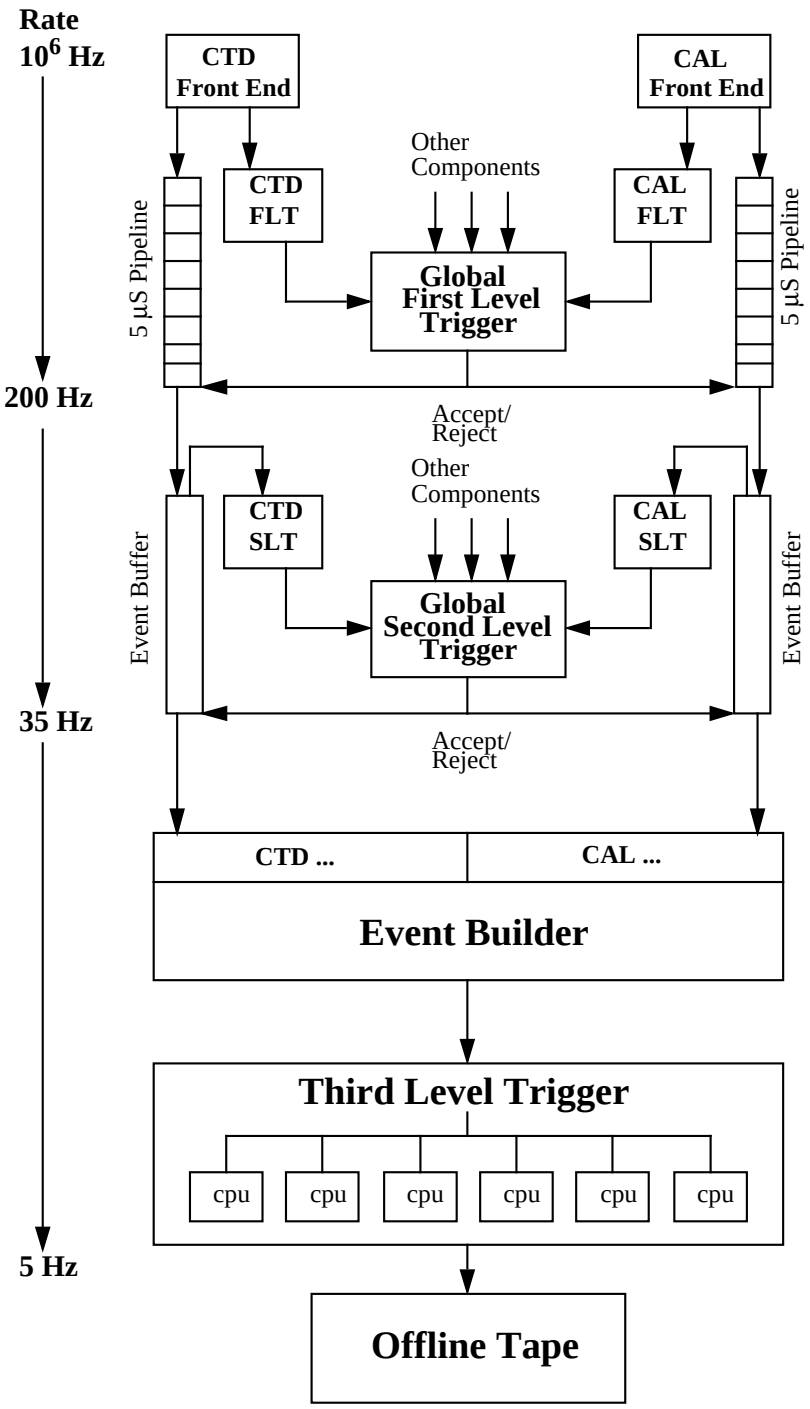


Figure 5.1. Data acquisition flow at ZEUS. The left vertical scale depicts the event rate reduction at each trigger level.

5.1.1 First level trigger

The first level trigger (FLT) is a hardware based trigger employing programmable logic to accept or reject events. It consists of component FLT's, mostly local to each component, and the global first level trigger (GFLT). At first level trigger, all data are stored by the detector components locally, in hardware buffers. Because the trigger cannot take decisions within the time between two bunch crossings, the data are moved down a pipeline for a maximum delay of about $5\,\mu s$ to allow a trigger decision to be taken. The individual component decisions use only a subset of the component's entire data and are taken within $1 - 2.5\,\mu s$. The component trigger decisions are then fed to the GFLT algorithm which combines the local decisions to make the global first level trigger decision. The output of the this global trigger arrives with a latency of $4.4\,\mu s$.

The most important components are the CAL and the CTD. Typical information used by GFLT is the vertex position of the event, total transverse energy and energy sums in sections of the calorimeter. Events that have a signal in SRTD for which the timing is inconsistent with particles arriving from the event vertex are rejected as beam gas events. A NC DIS event is identified at the FLT level by tagging the scattered electron. The FLT reduces the event rate from 100 kHz to below 1 kHz by selecting events with large energy deposits in the calorimeter and good tracks from the CTD FLT.

For the current analysis, only those events are kept which, at the FLT level, passed *at least one* of the following criteria:

- an isolated electron identification (isolated EMC clusters in the calorimeter) with a minimum energy of $3.9\,\mathbf{GeV}$ in at least one RCAL EMC cluster. (Near the beam pipe, this threshold is raised to $5\,\mathbf{GeV}$.)

- an isolated electron identification (isolated EMC clusters in the calorimeter), with a minimum energy of $2\,\mathbf{GeV}$ in at least one RCAL EMC cluster AND a track pointing to it.

- the total energy in the CAL EMC cells is larger than $20\,\mathbf{GeV}$.

- for higher Q^2, the electron scatters at larger angles, crossing the BCAL. To keep those events, a cluster with minimum BCAL energy of $4.7\,\mathbf{GeV}$ AND a reconstructed track pointing to this cluster are required.

For all the events, additional cuts on the timing information from C5, veto wall and SRTD were required such that background events like beam gas interaction and cosmics are rejected.

5.1.2 Second level trigger

All events accepted by GFLT are fully digitized and then copied to local component second level trigger (SLT) processors. The component SLT's use the entire data from each component in order to send processed information to the global SLT (GSLT). The algorithms are more complex than the GFLT ones: objects like track momenta, event vertex and calorimeter clusters are reconstructed.

The ZEUS *Global Tracking Trigger (GTT) Barrel Algorithm* integrates track information from the CTD and MVD to obtain a global picture of the track topology in the barrel region (-1.5 $< \eta <$ 1.5) of the ZEUS detector at the SLT stage. Algorithm processing is performed on a farm of Linux PCs and, to avoid unacceptable dead-time in the ZEUS readout system, must be completed within the strict requirements of the ZEUS trigger system. The GTT Barrel Algorithm greatly improves the vertex resolution and the track finding efficiency of the ZEUS SLT.

At the SLT level, events are vetoed using timing obtained from the FCAL, RCAL or the upper and lower halves of the BCAL. Timings are calibrated such that a $t = 0$ corresponds to collisions at the nominal interaction point. If the arrival time reported by RCAL, in absolute value, is larger than 8 ns or the time difference between FCAL and RCAL arrival times is larger than 8 ns, the event is vetoed as proton beam gas event. If the lower BCAL half reports a time delay longer than 10 ns w.r.t. the upper BCAL half, the event is vetoed as a cosmic event. A cut on $\sum_{cells,i}^{n} E_i^{cells} - P_{z,i}^{cells} < 75\,\mathbf{GeV}$ is also required in order to further remove background and reduce rate.

The GSLT takes its decision after ≈6 ms. The SLT reduces the event rate from 1 kHz to a typical 30-100 Hz.

For each event that passed the GSLT, the data of the event is sent to the event builder (EVB) where it is combined into a single record of ADAMO database tables[30]. This is also the data structure used in subsequent offline analysis . It is further distributed to the TLT processor nodes.

5.1.3 Third level trigger

The third level trigger (TLT) algorithms run on a farm of processors. Each workstation individually analyzes each single event with a custom simplified version of the offline reconstruction software and uses the full event information to calculate the event kinematics. Detailed tracking and jet finding algorithms are performed. Events that pass specific trigger criteria get a tag associated with these criteria added. These tags are known as the TLT flags and the pattern of these flags is used to select specific types of events. Typically, about 5 events pass the TLT criteria each second. The TLT accepted events are sent further to the event repository through a dedicated connection (FLINK) and to the online cluster for monitoring. During data taking, the ZEUS crew on shift monitors the trigger rates carefully for an optimal quality of data being recorded.

All data used in this analysis passed the TLT criteria, described by the following bit names:

- TLT_HFL02: a filter for charm mesons in DIS, it requires a DIS scattered electron OR a HFM trigger bit to be set. Essentially, the HFM trigger bits filter events with at least one loose charm meson candidate:
  - HFM01 tags events in which the decay $D^* \rightarrow K^- + \pi^+ + \pi_s^+$ has been loosely reconstructed ($1.65 < m(K,\pi) < 2.1\,\mathbf{GeV}$; $P_T(D^*) > 1.35\,\mathbf{GeV}$; $P_T(K,\pi) > .035\,\mathbf{GeV}$; $P_T(\pi_s) > 0.1\,\mathbf{GeV}$)
  - HMF02 and HFM03 are similar to HFM01 but tag the decays $D^* \rightarrow K^- + \pi^+ + \pi^- + \pi^+ + \pi_s^+$ and $D^* \rightarrow K^0 + \pi^+ + \pi^- + \pi_s^+$ respectively
  - HFM04 tags the decay $D^0 \rightarrow K^- + \pi^+$. The D^0 candidate is reconstructed in the mass window $[1.6\,,\,2.2]\,\mathbf{GeV}$ with the $P_T(D^0) > 2.8\,\mathbf{GeV}$and $P_T(K,\pi) > 0.7\,\mathbf{GeV}$
  - HFM05 through HFM17 reconstruct other charm meson decay channels such as $D^0 \rightarrow K + 3\pi$, $D^0 \rightarrow K^0 + 2\pi$, $D^+$and D_s decays.
- TLT_DIS03: a medium Q^2 trigger, it requires an electron of at least $4\,\mathbf{GeV}$, found by either *SINISTRA* or *EM* packages [31, 32]. For electrons scattered into the RCAL, a minimal radius around the beampipe of at least 35 cm is also required. Furthermore, $30\,\mathbf{GeV} < E_{total} - p_Z$ is necessary as well as a combination of FLT and SLT bits (FLT 28,30,40,41,43,44,46,47 and SLT 1,2,3,4,5,6,7,8) to be on.

- TLT_SPP02: an inclusive DIS trigger. It requires $30 < E_{total} - p_Z < 100\,\mathbf{GeV}$, an electron of at least $4\,\mathbf{GeV}$ and excludes a region of $12 \times 12\,\mathbf{cm}$ around the rear beampipe for the impact position of the electron (box cut).

5.1.4 Offline reconstruction

Typically, within a couple of days after acquisition, the data are reconstructed offline using not only the full event information but also calibration information regarding the status of the detector, such as calorimeter noise, MVD alignment, etc. so that corrections to event variables are performed for a more accurate measurement. The more CPU intensive parts of the reconstruction are implemented here rather than at TLT level. Samples of similar events are selected by associating each event with a specific code called a DST bit.

DST bits 9,10 and 11 (tagging good NC events or events with a well reconstructed electron) were also required for all selected events. Bit 9 requires a scattered electron with an energy of at least $4\,\mathbf{GeV}$ found by at least 4 of the electron finder packages, bit 10 filters events with a well reconstructed primary vertex and bit 11 requires $E - P_Z > 30\,\mathbf{GeV}$.

5.2 Data from Hera II

After a long shutdown, HERA resumed operation in 2001. A summary of the luminosity recorded by ZEUS since 2002 until present day is given in Fig. 5.2. The luminosity values corresponding to different data taking periods are also given in Table 5.1. In the years 2002 and 2003 positrons were used. During the second half of 2004 and the entire 2005, electrons were used. This analysis uses a data sample of 127.35 pb^{-1}, corresponding to 129714544 $e^- p$ collisions recorded by ZEUS in 2005. In this thesis, the incoming/scattered lepton will be referred to as *the electron*. The following sections look at the selection criteria for obtaining a clean NC DIS sample. The quality of the selected events is scrutinized. All the runs have microvertex detector information available.

5.3 Event reconstruction

The final state of a DIS event consists of the scattered electron and the hadronic system. The hadronic system refers to all event attributes not related to the scat-

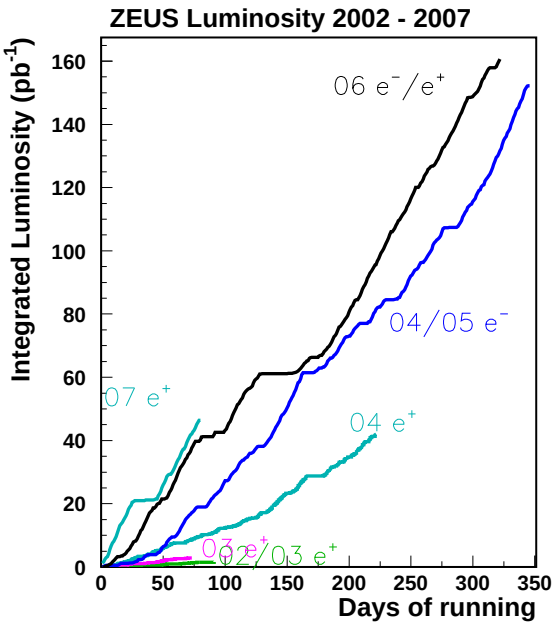


Figure 5.2. Integrated luminosity gated by ZEUS since 2002 until presently. Different positron/electron periods are plotted separately. The data set used in this analysis is shown as "04/05 e^- ".

tered electron: the jet, which is the result of the struck quark hadronizing, and the proton remnant. An event is determined in terms of kinematics by any two of the variables Q^2, x and y, as explained in Chap. 2. Different methods are employed in order to compute the event kinematics: one can use only information from the hadronic system, scattered electron or a combination of the two.

Period	Beams	Recorded luminosity
2007 (MER)	e^+p	7.77 pb^{-1}
2007 (LER)	e^+p	13.18 pb^{-1}
2006/2007	e^+p	145.90 pb^{-1}
2006	e^-p	61.23 pb^{-1}
2004/2005	e^-p	**152.26** pb^{-1}
2004	e^+p	43.74 pb^{-1}
2003	e^+p	2.87 pb^{-1}
2002/2003	e^+p	1.78 pb^{-1}

Table 5.1. A summary of different data taking periods and corresponding luminosities, as recorded by ZEUS. The top two entries (LER, MER = low and respectively medium energy runs) refer to periods for which the center of mass energy was altered. This analysis uses a sub-sample of 127 pb^{-1} of the 2004/2005 data taking period.

5.3.1 Jacquet-Blondel method

The Jacquet-Blondel method uses only hadronic information to reconstruct the event kinematics. First, the following are defined:

$$\delta_h = \sum_i^N (E_i - p_{z,i}) \tag{5.1}$$

$$p_{T,h}^2 = \left(\sum_i^N p_{x,i}\right)^2 + \left(\sum_i^N p_{y,i}\right)^2 \tag{5.2}$$

where i runs over all calorimeter clusters not associated with the scattered electron. The polar angle γ_h of the struck quark, at leading order, is given by:

$$\cos\gamma_h = \frac{p_{T,h}^2 - \delta_h^2}{p_{T,h}^2 + \delta_h^2} \tag{5.3}$$

The kinematic variables of the event are reconstructed as:

$$y_{JB} = \frac{\delta_h}{2E_{e\,beam}} \tag{5.4}$$

$$Q_{JB}^2 = \frac{p_{T,h}^2}{1 - y_{JB}} \tag{5.5}$$

where $E_{e\,beam}$ is the nominal beam energy of the incoming electron. The hadronic x is obtained from $Q^2 = sxy$. Because a large fraction of the hadronic system escapes undetected as it goes down the beampipe, the measured $p_{T,h}^2$ can be distorted. For this reason, the resolution of this method is somewhat poor, especially for low Q^2 and moderate x. The Jacquet-Blondel method is mainly used in charge current interactions where there is no information available from the final state lepton: neutrino's cross the ZEUS detector completely undetected.

5.3.2 Electron method

This method uses solely the measurement of the scattered electron. The kinematic variables are given by the following equations:

$$y_e = 1 - \frac{E_e}{2E_{e\,beam}}(1 - \cos\theta_e) \tag{5.6}$$

$$Q_e^2 = 2E_e E_{e\,beam}\,(1+\cos\theta_e) \tag{5.7}$$

where θ_e and E_e are the angle and the energy of the scattered electron respectively. Both reconstructed variables depend on the accurate reconstruction of E_e and θ_e. Although the electron angle can be measured well with the help of the SRTD at low angles and by tracking at high angles, accurately determining the electron energy is problematic due to some amount of dead material in front of the calorimeter at low angles. At moderate x, $x \in [10^{-4}, 10^{-2}]$, this method has a poor resolution. A more detailed description of different reconstruction methods and their discriminative power for different regions of the phase-space can be found in previous ZEUS theses [33, 34].

5.3.3 Double angle method

This method relies on the angles of both the hadronic system and the scattered electron [35].

$$y_{DA} = \frac{\sin\theta_e(1-\cos\gamma_h)}{\sin\gamma_h + \sin\theta_e - \sin(\gamma_h - \theta_e)} \tag{5.8}$$

$$Q^2_{DA} = 4E^2_{e^-beam}\frac{\sin\gamma_h(1+\cos\theta_e)}{\sin\gamma_h + \sin\theta_e - \sin(\gamma_h - \theta_e)} \tag{5.9}$$

Neither y_{DA} nor Q^2_{DA} depend on any measured energies. Since angles are in general more precisely measured than energies with the ZEUS detector, this method leads to accurate results in larger regions of the phase-space. These regions are characterized by substantial hadronic activity at large angles and a well measured electron track. These coincide with the region where charm can be measured well. The double angle method will therefore allow for an optimum reconstruction of the kinematic variables and so will be used in this thesis for the determination of the kinematics.

5.4 DIS selection

Present ZEUS analyses use a complex library of reconstruction software which translates the raw timed hits in the detector into physical objects, as explained in detail in previous chapters. Each event that has passed the trigger is reconstructed using programs from this library. The reconstructed events containing tracks, covariance

matrices and calorimeter energy deposits are written to the output. ROOT[36] is an object oriented data analysis framework which serves as an interface between the data and the particle reconstruction algorithms. All data used in this thesis have been processed within the Heavy Flavor analysis group at ZEUS in the form of ROOT files. This large data sample represents the raw data. A preselection is made from these events. First, events passing the criteria for deep inelastic scattering are selected. This is done using the following criteria:

DIS selection:

- z position of the primary vertex: $-50\,\mathbf{cm} < z < 50\,\mathbf{cm}$. The z position of the vertex is restricted to this range to ensure a high (and better understood) acceptance for the calorimeter as well as for the central trackers. Also, events for which the interaction point could not be reconstructed properly are rejected.

- Energy distribution in the calorimeter: $38 < \delta < 65\,\mathbf{GeV}$, with $\delta = E - p_z$ and E being the event total energy. For a perfectly contained and measured DIS event, it follows from energy and momentum conservation that $\delta = 2E_{e\,beam} = 55\,\mathbf{GeV}$. Particles escaping down the beam-pipe in the forward direction have energies almost equal to their momenta in the z direction and therefore their contribution to δ is negligible. The same is not true in photo-production: the electron escapes the detector through the rear beam-pipe and therefore E^e and p_z^e do not cancel, effectively lowering the measured δ. This cut is therefore effective at removing photo-production background events while keeping DIS events.

- Scattered electron : at least one candidate with probability larger than 0.9 according to the SINISTRA electron finder [31] and energy in the calorimeter of at least $10\,\mathbf{GeV}$ is required. This electron is most likely to be the scattered electron.

- $y_e < 0.95$. This cut will remove fake candidates (such as neutral pions) mistaken for scattered electrons by the ZEUS reconstruction software.

- $y_{JB} > 0.02$. This cut is a convenience cut as in charm analysis well reconstructed tracks are required and thus reconstructable charm events will always pass this criterion. This cut is stronger than the cut used in inclusive DIS

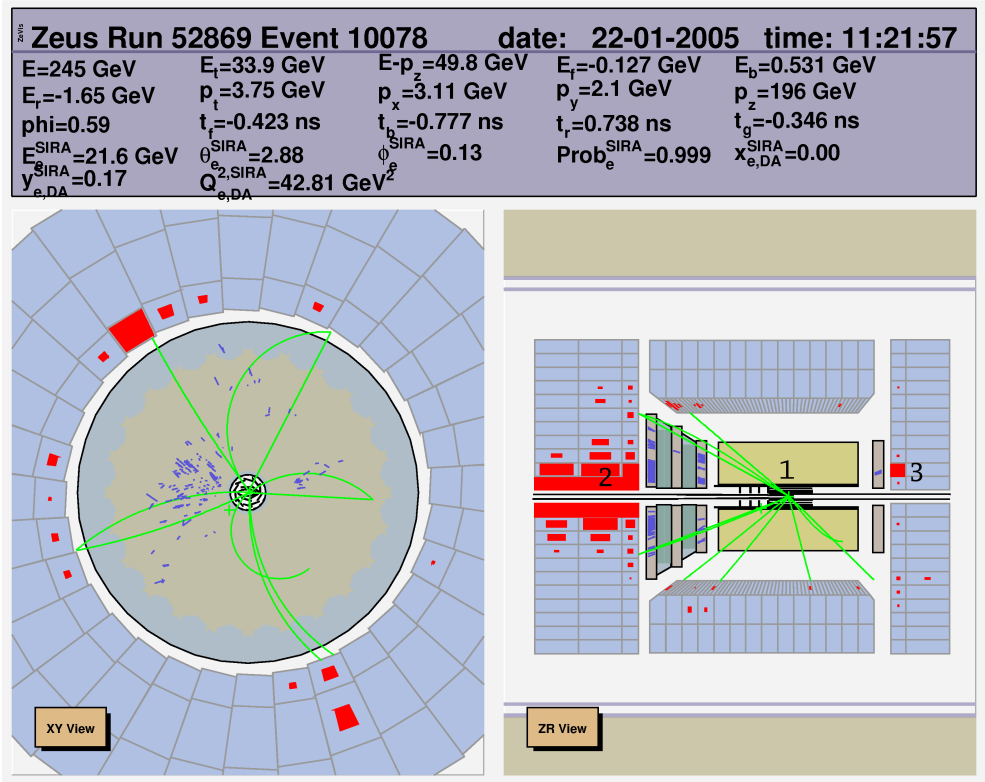


Figure 5.3. This DIS event with $Q^2 \simeq 43\,\mathbf{GeV^2}$ was recorded in January 2005. The detector is seen in $r\,\phi$ view on the left and $z\theta$ view on the right. The interaction point (1) and the proton remnant (2) are marked on the right figure. Also an electron of $21\,\mathbf{GeV}$ is clearly seen in the calorimeter (3)

analyses. It removes events where the full hadronic system is contained near the beampipe in the proton direction.

- A restriction is made on the kinematic domain: $1 < Q^2_{DA} < 1000\,\mathbf{GeV^2}$ and $0.03 < y_{DA} < 0.7$. For this, double angle variables are used.

This DIS selection is a standard selection for any ZEUS NC DIS analysis and therefore very well understood. In Fig. 5.3 and 5.4 two typical DIS events of $Q^2 \simeq 43\,\mathbf{GeV}$ and $Q^2 \sim 336\,\mathbf{GeV}$ are presented: the electron is found in the RCAL and the proton remnant flies in the forward direction, leaving energy deposits in the FCAL. In both figures, the calorimeter energy deposits associated with the scattered electron (3) and the proton remnant (2) are clearly visible. The interaction point is also explicitly labeled (1). In Fig. 5.4, the region around the interaction point has been enlarged (left figure). The beampipe and the first cylinder of barrel MVD ladders are visible. The primary interaction point (1) is pointed out, as well as a

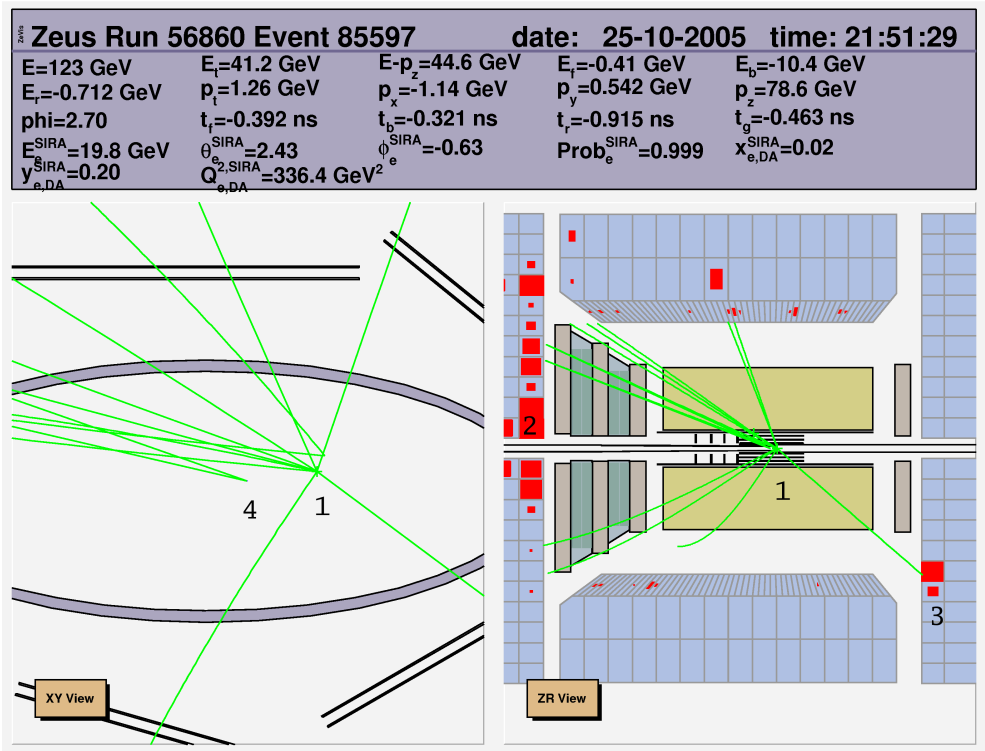


Figure 5.4. A DIS event with $Q^2 \simeq 336\,\mathbf{GeV^2}$ recorded in October 2005. On the right plot: the scattered electron (3) was found in RCAL with an energy of $19.8\,\mathbf{GeV}$. The interaction point (1) and proton remnant (2) are also shown. On the left plot: the region around the interaction point has been enlarged. The beampipe and the first four ladders of the barrel MVD are visible. The primary interaction point can be seen (1) as well as a secondary vertex consisting of two tracks (4).

two track secondary vertex (4).

For a subset of 240k events of raw data, the distributions in $E - p_z$, Z vertex position, energy of the scattered electron energy, the angle θ of the scattered electron, $log_{10}(Q^2)$ and $log_{10}(x)$ of the event are shown in Fig. 5.5. This is shown both excluding and including the DIS selection.

When searching for charm, it is necessary to identify the secondary vertices associated with the decay of the charm particle. For these vertices, high resolution is achieved if the tracks to be vertexed are well defined. This implies that a certain selection of "good" tracks should be performed for each event. Distributions of track characteristics such as momentum, number of MVD hits, η and number of crossed superlayers in the CTD are presented in Fig. 5.6, after applying the DIS selection. The η distribution is asymmetric, with more reconstructed tracks in the forward direction, due to the large proton boost along the positive z axis. Well re-

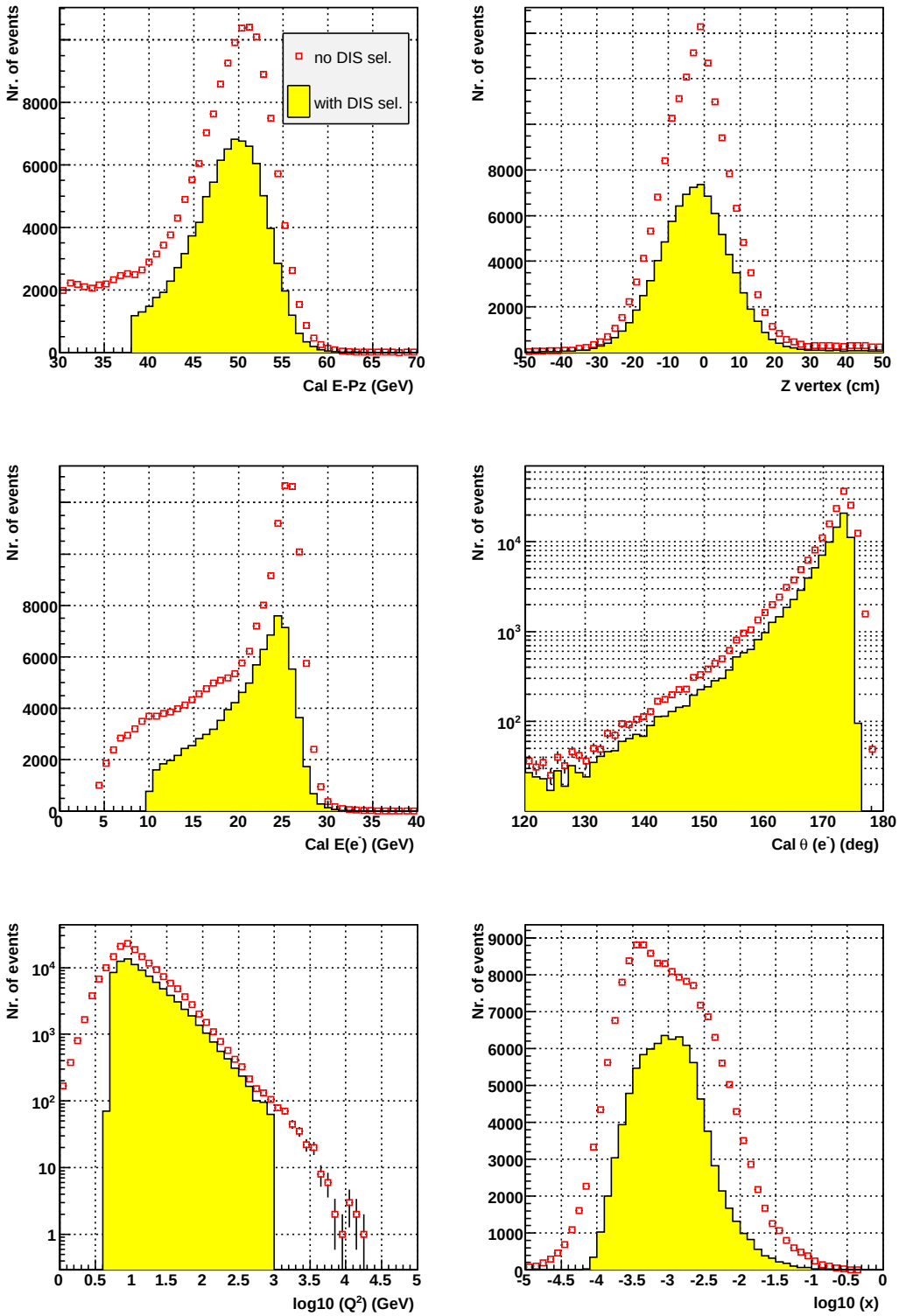


Figure 5.5. Data plots: distributions of $E - P_Z$, Z vertex position, the energy of the scattered electron as reported by the CAL, the angle θ of the scattered electron, $log_{10}(Q^2)$ and $log_{10}(x)$. The open squares represent distributions before the DIS selection is applied, the filled histograms are the distributions after DIS selection. The figures correspond to 240k events recorded during 4 runs in 2005.

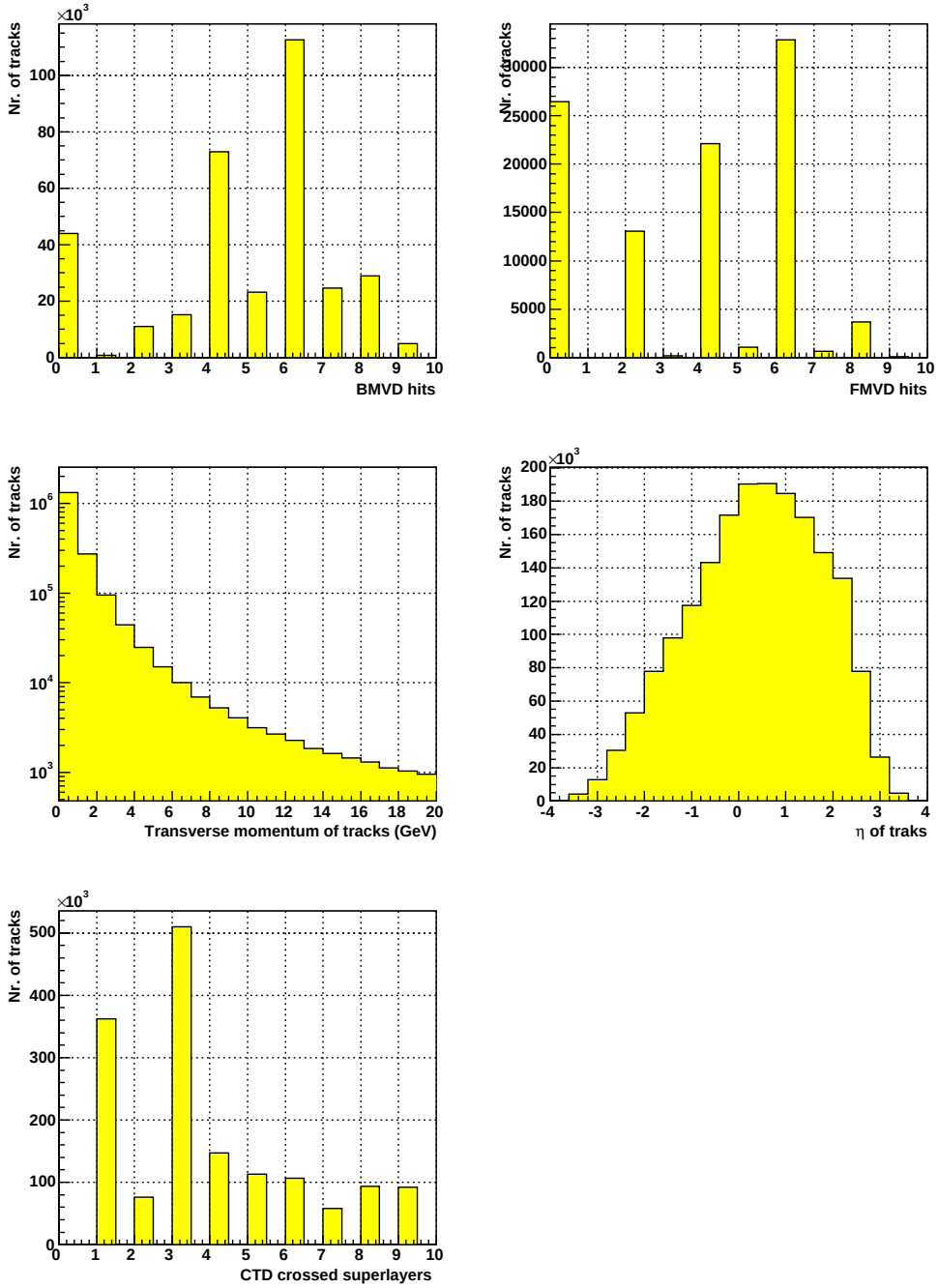


Figure 5.6. Data plots. For all tracks, the following distributions are shown (from top left to bottom right): the number of hits in barrel MVD, MVD hits in the wheels, transverse momentum, pseudorapidity η and the number of crossed superlayers in the CTD. The distributions represent the DIS selected data.

constructed CTD tracks, which crossed 5 superlayers and whose DCA z coordinates were reconstructed within 10 cm around the origin of the coordinate system, are used to extract the barrel MVD efficiency shown in first plot of Fig. 5.6. About $\sim 90\%$ of all good tracks crossing the BMVD are associated hits in the MVD, as it can be inferred from Fig. 5.6. Both the pattern recognition software efficiency and the dead MVD channels influence this value. Also, there are tracks with more than 6 barrel hits (3 $r\phi$ + 3 rz hits, in 3 different cylinders), due to small overlaps of the edges of some neighboring ladders (see Fig. 3.14). In very rare cases, the maximum of 24 hits has been observed. For the forward MVD, the selected tracks do not always pass through the detector leading to an apparent loss in efficiency in Fig. 5.6. In Fig. 5.8 it is shown that tracks crossing at least 4 superlayers in the CTD have a corresponding polar angle larger than $\theta = 22°$. For θ values around 19°, tracks will cross the first wheel of the FMVD and 2-3 CTD superlayers. This is valid for straight lines; low momentum helices can collect more hits in the FMVD and still cross few CTD superlayers. This analysis uses tracks within the pseudo-rapidity window $-1.6 < \eta(track) < 1.6$, corresponding to a polar angle range of $25.4° < \theta < 154°$. The large majority of the selected tracks will therefore leave at least 2 hits in the BMVD, ensuring a precise reconstruction of their position and momentum.

5.5 The Monte Carlo simulation

Charm quarks are generated and fragmented into a D meson. Charm decay was modelled through eight different decay modes. Transverse momentum cuts for the charm mesons were introduced in order to generate Monte Carlo events efficiently in the kinematic region required by the measurement, as summarized in Table 5.2.

For this analysis, the Monte Carlo programs HERWIG[37, 38] and PYTHIA[39], which implement leading order matrix elements followed by parton showers and hadronization, were used to model the final state. Direct and resolved events were generated separately and in proportion to the cross-sections. As input for the proton and photon parton distribution functions[1], the CTEQ5L[40] and GRV-G LO[41] were used.

After generating the final state partons using the Monte Carlo generators, events are input into the detector simulation. The detector simulation is performed by the

[1]The photon PDF's are used for the generation of resolved events.

Decay modes	P_T **cuts**
$D^{*+} \to D^0\,\pi_s^+ \to K^-\,\pi^+\,\pi_s^+$	$1.25\,\mathbf{GeV}$
$D^{*+} \to D^0\,\pi_s^+ \to K_s^0\,\pi^+\,\pi^-\,\pi_s^+$	$1.35\,\mathbf{GeV}$
$D^{*+} \to D^0\,\pi_s^+ \to K^-\,\pi^+\,\pi^+\,\pi^-\,\pi_s^+$	$2.3\,\mathbf{GeV}$
$D^0 \to K^-\,\pi^+$	$2.6\,\mathbf{GeV}$
$D_S^+ \to K^-\,K^+\,\pi^+$	$1.7\,\mathbf{GeV}$
$D^+ \to K^-\,\pi^+\,\pi^+$	$2.8\,\mathbf{GeV}$
$D^+ \to K^-\,K^+\,\pi^+$	$1.7\,\mathbf{GeV}$
$\Lambda_c \to p\,K^-\,\pi^+$	$2.8\,\mathbf{GeV}$

Table 5.2. Modelled decay channels in charm Monte Carlo and their P_T cuts at the generating level.

MOZART program which is based on GEANT 3.13 package[42]. The MOZART program simulates the interaction between particles passing through different components and materials of the detector and outputs the simulated detector response. Trigger simulation is also included. The MC events are reconstructed in precisely the same way as the data and stored into similar ROOT ntuples. These ROOT files also contain information on the generated physics event: the truth variables. The simulated sample used in this analysis has an equivalent total luminosity of $295\,\mathbf{pb^{-1}}$.

In Fig. 5.7, comparisons are made between data and simulation for the distributions of DIS variables $E - P_Z$, the primary vertex Z position, the scattered electron energy and angle, Q^2 of the event and Bjorken x. The same reconstruction software was used both in data and in Monte Carlo. The histograms were area normalized[2]. The shape of all distributions is well described by the simulation. A shift of few **GeV** to higher values is seen in the energy of the scattered electron due to imperfections in the dead material description. This also reflects in the $E - P_Z$ distribution.

[2]Luminosity normalization cannot be applied when comparing distributions of events because light flavor events were not included in the simulation.

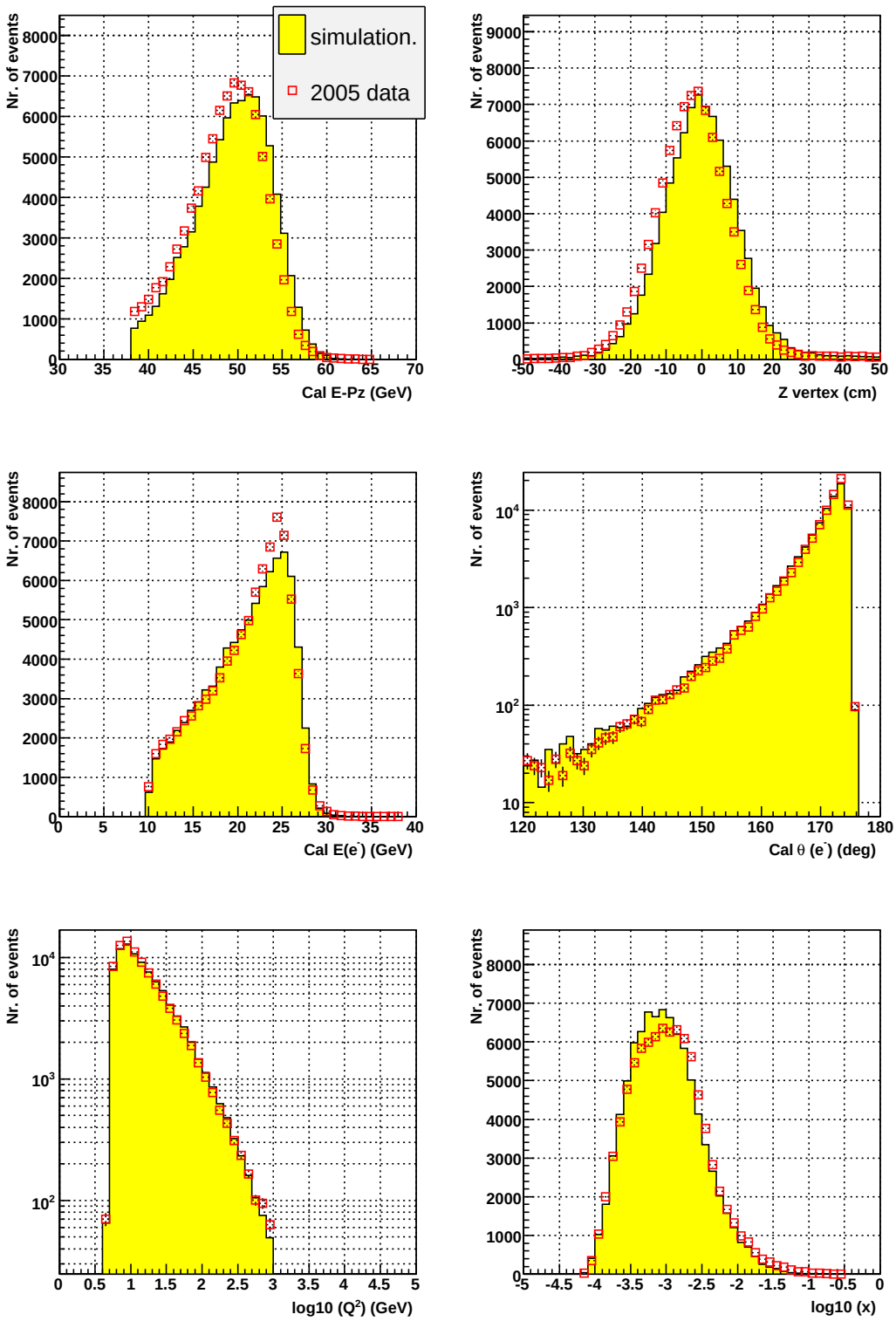


Figure 5.7. Distributions of $E - P_Z$, the primary vertex Z position, the energy and angle of scattered electron, Q^2 of the event and Bjorken x are compared in data and simulated files. The histograms are area normalized. The shape is well described overall. A shift of $\sim 2\,\mathbf{GeV}$ in the scattered electron energy is seen, producing a similar effect on $E - P_Z$.

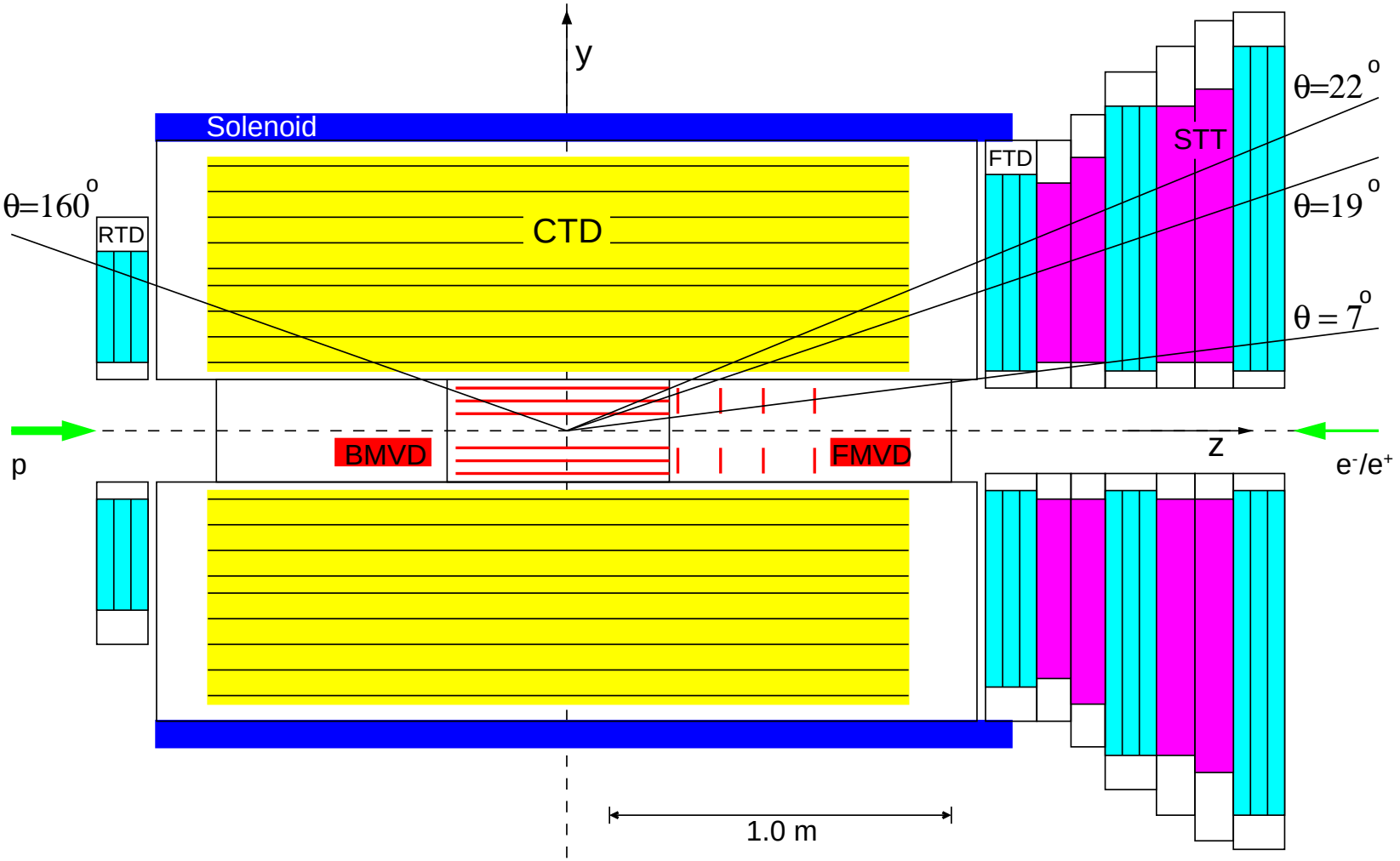


Figure 5.8. The ZEUS detector and its tracking detectors. A polar angle of $\sim 19^\circ$ or smaller ensures one or more hits in the FMVD. A polar angle bigger than 22° corresponds to 4 or more superlayers in the CTD.

Chapter 6

Charm finding

This chapter is organized as follows. First, the reconstruction of the D^0 charm meson is described. A total sample of candidates is found. Second, the total sample is further divided into two sub-samples: the tagged sample, consisting of D^0 mesons originating from a D^{*+} decay, and the untagged sample, of D^0's originating elsewhere. Based on the tagged sample, a method is developed for removing a fake reflected component of the signal in the untagged sample. Third, the cleaned signal is optimized, based on parameters computed at the D^0 decay vertex. The signal is studied as a function of the virtuality Q^2 of the event, the transverse momentum and pseudo-rapidity of the charm meson as well as Bjorken x. Last, a comparison between data and simulation is made.

6.1 Charm at Zeus

In Zeus, the charm content of the proton can be probed in various ways. Several techniques suited to investigating charm are explained in this chapter.

Free quarks or gluons cannot exist individually, due to color confinement. The process of formation of hadrons out of free quarks or gluons is called hadronization. The narrow cone of particles created by the hadronization of a single parton is called jet. Most charm quarks hadronize into charm mesons. The charmed vector mesons (D^{*+}, D^{*0} and D_s^{*+} and their antiparticles) always decay to charmed pseudo-scalars (D^+, D^0 and D_s^+ and their antiparticles). Those decays are either strong or electromagnetic and from the experimentalist's point of view the vector mesons have zero life-time: a charmed vector meson created at the interaction point will decay to a pseudo-scalar at the interaction point. The pseudo-scalar will continue its flight

for a relatively short time before decaying. D^0 is the shortest lived pseudo-scalar charmed meson: it has a lifetime of $(410 \pm 1.5) \times 10^{-15}$ seconds, as listed by Particle Data Group in 2006 [43], or , in experimental units, a "$c\tau$" of $123\,\mu m$. The slightly more massive D^+ lives almost 3 times longer: $c\tau = 311\,\mu m$. The ZEUS detector is powerful enough to resolve secondary decay vertices associated with charmed meson decays.

6.2 D^0 reconstruction

6.2.1 The algorithm

D^0 mesons will be reconstructed via the decay channel $D^0 \rightarrow K^- + \pi^+$ which accounts for 3.8 % of all D^0 decays. The daughter particles, kaons and pions, are both electrically charged and long lived. These will generate hits in the tracking detectors and their path can therefore be reconstructed.

D^0 reconstruction is performed as follows: for each event a list of high quality tracks is made. Criteria for selecting good tracks are:

- the number of hits in the MVD (at least 2 hits required)
- the number of crossed superlayers in the CTD (minimum of 3 superlayers required)
- a minimum transverse momentum of $0.2\,\mathbf{GeV}$.

These requirements ensure that the selected tracks are well defined and are constrained to the central region of the ZEUS detector where the reconstruction efficiency is higher and acceptance well understood. The criteria also remove *ghost* tracks found in the pattern recognition. Ghost solutions can be created by combining good segments from different tracks. They almost never link well with the MVD information.

For each event, all two-track combinations are made with tracks of opposite charge. No absolute value for the charge of the kaon and pion candidates is required therefore allowing the reconstruction of $\bar{D}^0$ as well. The best secondary vertex formed by the track pair is calculated. The track momenta are recomputed at the secondary vertex using the smoothing procedure described in Chap. 4. The invariant D^0 mass is calculated at the vertex. Candidates with a mass $m(D^0) \in$

$[1, 2.5]$ **GeV** are kept. In order to reduce the background, further cleaning cuts are introduced such as minimum transverse momenta for both the K and π as well as for the reconstructed D^0. Candidates are reconstructed only in the pseudo-rapidity region of $\eta \in [-1.6, 1.6]$. The reconstructed D^0's could be either directly produced in the hadronization process of the charm quark or could come from a charmed vector meson decay, such as $D^{*0} \to D^0 + \gamma$ and $D^{*+} \to D^0 + \pi^+$.

6.2.2 Signal and background.

The entire 2005 $e^- p$ dataset was used to produce the D^0 invariant mass plot shown in Fig. 6.1. The corresponding kinematic range is $5 < Q^2 < 1000\,\mathbf{GeV^2}$, $3 < P_T(D^0) < 20\,\mathbf{GeV}$, $|\eta(D^0)| < 1.6$. The signal has been enhanced by demanding that $|\eta(K,\pi)| < 1.6$ and $P_T(K,\pi) > 0.8\,\mathbf{GeV}$. The histogram is fitted with a modified Gaussian signal superimposed on a modified exponential background:

$$\mathbf{F}(m) = (am + b)\,\exp\left[c(m - 1.73)\right] + A\,\exp\left[-\left(\frac{|m - m_0|}{2\sigma}\right)^{\left(1+\frac{1}{1+|m-m_0|/(2\sigma)}\right)}\right] \tag{6.1}$$

The modified Gaussian was chosen based on the fact that the signal does not have a fixed width: its resolution depends slightly on the K and π daughter track momenta. The background function, an exponential with varying amplitude, fitted the easiest to the the regions outside the signal peak.[1]

A clear signal can be seen at the fitted value of $m_0 = 1.864\,\mathbf{GeV}$, on top of a large but smooth background[2]. There are 7440 ± 233 signal events extracted by the fit. The statistical error has been computed as $\Delta N(D^0) = \sqrt{N + B}$ where N is the number of candidates and B is the number of background combinations under the peak. The sudden rise in background for mass values below $1.73\,\mathbf{GeV}$ is due to another D^0 decay, $D^0 \to K^- + \pi^+ + \pi^0$. By leaving out the π^0 daughter track, a smeared out peak appears in the background below values of $m = m(D^0) - m(\pi^0) \sim 1.73\,\mathbf{GeV}$. This feature has been taken into account by allowing the fit function to change slope around the mentioned value (two sets of a, b and c parameters).

The candidates were divided into two sub-samples: if the D^0 candidate was found to have originated from a $D^{*+} \to D^0 + \pi^+$ decay, then the candidate was

[1] The MINUIT package [44] found the lowest χ^2 when fitting this particular background function to the data points, w.r.t. to several polynomial functions.

[2] The PDG[43] value for D^0 invariant mass is $1864.5 \pm 0.4\,\mathbf{MeV}$.

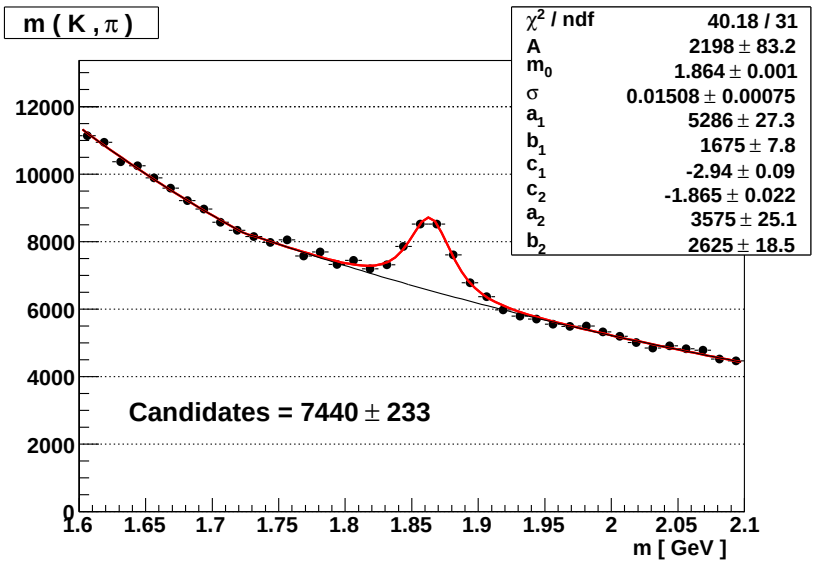


Figure 6.1. D^0 invariant mass for the kinematic range: $Q^2 > 5\,\mathbf{GeV^2}$, $3 < P_T(D^0) < 20\,\mathbf{GeV}$ and $|\eta(D^0)| < 1.6$. The plot corresponds to $127\,pb^{-1}$ of e^-p collisions recorded in 2005. The fit parameters of the function from eq. 6.1 are listed on the top right: the meson invariant mass $m_0 = 1.864\,\mathbf{GeV}$, the width of the modified gaussian, $\sigma = 15\,\mathbf{MeV}$ and so on. A total of 7440 ± 233 candidates was found.

labeled as "tagged". The rest of the candidates which were not associated with D^{*+} decays formed the "untagged " sample. Below, a detailed description is given of how the tagging procedure was carried out and what the advantages are of dividing the sample into tagged and untagged sub-samples.

6.2.3 D^0 originating from D^{*+} decays

D^{*+} can be best reconstructed using the golden decay mode $D^{*+} \rightarrow D^0 + \pi_s^+ \rightarrow K^- + \pi^+ + \pi_s^+$. Here, π_s means that the pion is "slow" or almost at rest in the D^{*+} center of mass due to phase-space considerations: the D^{*+} and D^0 masses are only $145.15\,\mathbf{MeV}$ apart and the π^+ has a mass of $139.57\,\mathbf{MeV}$.

For each event, a search for D^0 candidates is performed as described previously. A third track is then added as the slow pion, with the constraint that its electric charge be opposite to the charge of the assumed kaon. A minimum transverse momentum for the slow pion was required: $p_T(\pi_s) > 0.125\,\mathbf{GeV}$. Let $\delta m = m(D^*) - m(D^0)$. Then, by keeping only those D^0 candidates for which a slow pion was found such that $0.143 < \delta m < 0.148\,\mathbf{GeV}$, a clean sample of D^0's coming from D^* decays is selected. This sample will be referred to as the "tagged" sample. All other D^0 candidates which fail the δm cut form the "untagged" sample. The invariant mass

difference δm is plotted in Fig. 6.2. The signal has been greatly enhanced by additionally requiring a "genuine" D^0: $1.8 < m(D^0) < 1.92\,\mathbf{GeV}$ (right plot).

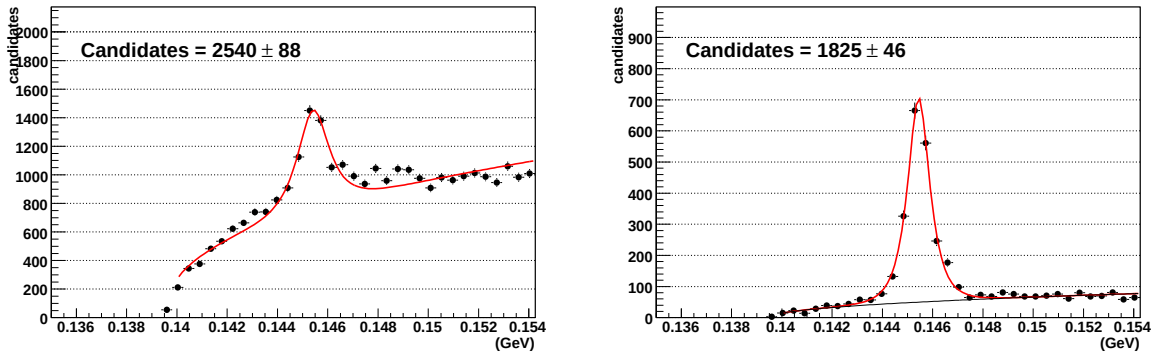


Figure 6.2. The invariant mass difference $\delta m = m(D^*) - m(D^0)$ is shown (left). The signal has been enhanced greatly by requiring $1.8 < m(D^0) < 1.92\,\mathbf{GeV}$(right). The fitted peak value is $0.1454\,\mathbf{GeV}$, identical to the PDG value [43].

6.2.4 The reflected signal

As the slow pion's charge is selected opposite to the kaon charge, the kaon and the pion are identified without ambiguity. For the candidates belonging to the untagged sample, there is no way to uniquely identify the kaon and the pion; therefore, the invariant mass distribution in Fig. 6.1 contains not only the signal and combinatorial background but also a component from the signal with incorrectly assigned particle identifications. Wrongly assigned combinations are unavoidable within the untagged sample. The wrongly assigned candidates add both to the background and to the signal of the invariant mass distribution but with different weights. For very high values of the D^0 momentum, the kaon and pion momenta contribute significantly more than their masses when reconstructing the mass of the D^0 mother and therefore wrongly assigned candidates will also exhibit a peak at the same mass value as the D^0 meson. At very low momenta of the D^0 meson, wrongly assigning masses leads to a well spread invariant mass distribution which does not contribute to the signal more than to the background. In general, the wrongly assigned candidates produce a so-called "reflected signal". The tagged sample is in this respect reflection-free. Fig. 6.3 shows the invariant mass distribution of both the tagged and the untagged samples. There were 1668 ± 46 tagged candidates found and 5747 ± 228 untagged candidates found.

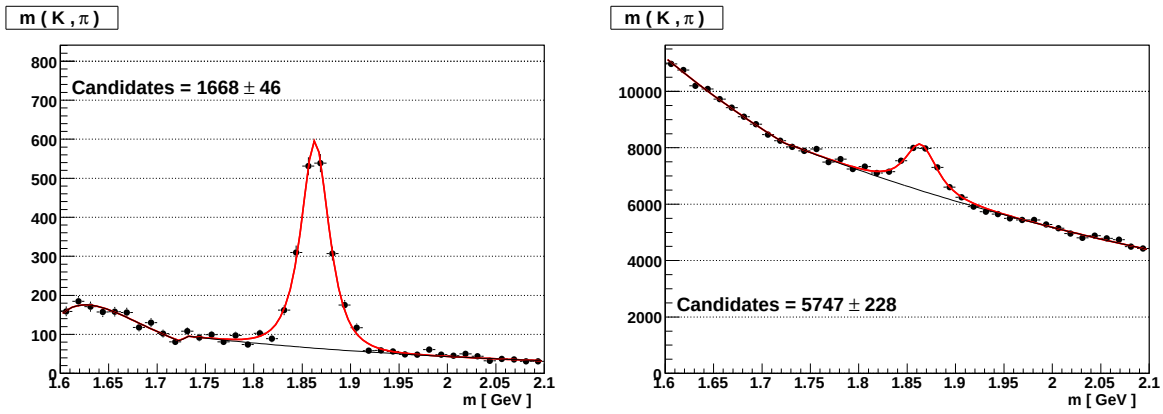


Figure 6.3. Left plot: invariant mass distribution for tagged D^0's (coming from D^* decays). Due to phase-space and right charge selection, the background is much reduced w.r.t. Fig. 6.1. Right plot: invariant mass distribution for untagged D^0's.

One can now produce a purely reflected invariant mass distribution by wrongly assigning masses to the kaon and the pion tracks for the tagged sample. In this way, the artificially induced reflection can be studied and used further to remove the reflection within the untagged sample. The reflection produced by the tagged sample is shown in Fig. 6.4. In order to remove reflected signal from the untagged sample,

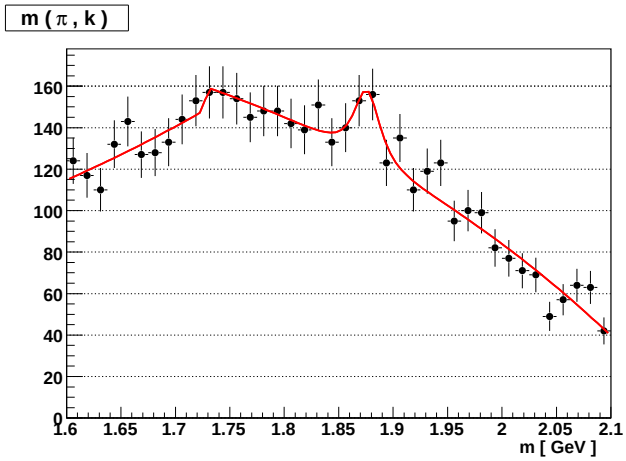


Figure 6.4. Invariant mass for the reflected signal produced by tagged D^0 sample. The number of entries in the above histogram equals the number of entries in Fig. 6.3 but the peak structure is lost.

the reflection displayed in Fig. 6.4 is scaled up and subtracted from the untagged sample invariant mass distribution. The scale factor is not known a priori. This is due to the fact that the efficiencies of the tagged and untagged sample differ as the efficiency of selecting the slow pion, which acted as a discriminant between the two

samples, is less than 100%. Therefore, the subtraction of the reflected signal is performed in an iterative fashion, also described in [45]. A good starting value for the scaling factor at the first iteration is the ratio of tagged/untagged candidates. The iteration continues until the scale factor settles to a value[3]. The mass distribution corresponding to the untagged sample, after the iterative subtraction, is presented in Fig. 6.5. The number of untagged candidates was reduced from 5747 to 5215.

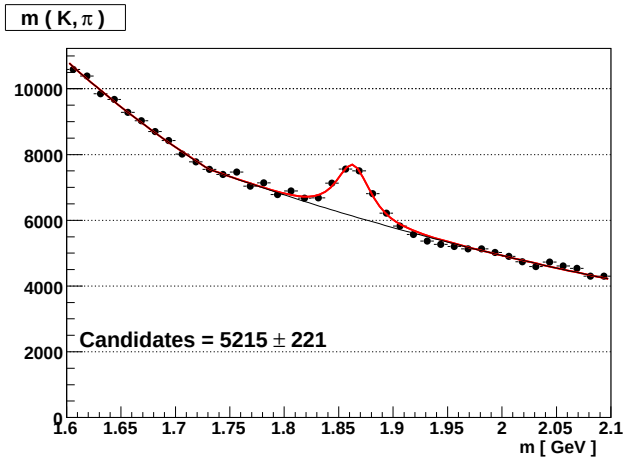


Figure 6.5. Invariant mass distribution for the untagged sample, after reflection subtraction. A total of 5215 ± 221 candidates was found. By comparing to Fig. 6.4, the reflection contributed to as much as 9 %.

The reflection contributed to as much as 9 % of the total signal.

6.2.5 Improving the signal versus background ratio

As the charmed meson decay vertex is explicitly reconstructed, several tagging variables can serve as good tuning cuts when scrutinizing the ratio $\sqrt{N+B}/N$, where N is the number of candidates and B the number of background combinations under the mass peak. This ratio represents the relative error on the number of candidates and it is this error that propagates directly into the error on the (differential) cross-sections. Cuts on the decay length, the decay length error, the "$c\tau$" of the meson candidate as well as the χ^2 of the secondary vertex need to be optimized such that the relative error on the signal is minimum.

[3]The iteration converges quickly. The first two iterative steps return: 5747(start value)→5124(1^{st} step)→5238(2^{nd} step)→etc

χ^2 of the secondary vertex

χ^2 is a measure of how well the vertexing procedure has performed. The decay vertex quality had been optimized at the vertexing level, by keeping only vertices with $\chi^2 < 16$. Two helix tracks of 5 parameters each have 10 degrees of freedom in total. After vertexing the two tracks to one vertex, there are only 9 degrees of freedom left: the 3-dimensional vertex position and two 3-dimensional directional momentum vectors of the two tracks at the vertex. Therefore, the $\chi^2/ndof = 16$ is equivalent to a 4σ confidence level.

The relative error on the signal is studied while varying the χ^2 cut: $\chi^2 < \chi^2_{cut}$, with $\chi^2_{cut} \in \{1, 2, 4, 6, 8, 16\}$. The results are presented in Table 6.1. The smallest relative error was obtained for $\chi^2_{cut} = 8$ beyond which the relative error remains constant. Setting $\chi^2_{cut} = 8$ is equivalent with accepting almost all secondary vertices[4]. Accepting candidates with vertices whose χ^2 is larger than this value would not change the relative error on the signal but would further lower the purity of the sample.

χ^2cut	**Nr. candidates**	**Relative Error**
1	4773	3.81 %
2	5568	3.42 %
4	6318	3.21 %
6	6603	3.17 %
8	6695	3.13 %
16	6951	3.13 %

Table 6.1. The variation on the number of candidates when varying χ^2_{cut}. Each row corresponds to a fitted signal value. The smallest relative error is obtained when $\chi^2_{cut} \leq 8$.

Primary vertex χ^2 increase due to pseudo-track fitting

A neutral pseudotrack has been reconstructed for each D^0 candidate following the method described in Sec. 4.3.1. Furthermore, the pseudotrack was refitted to the primary vertex. A new primary vertex position and its corresponding χ^2 increase were computed. The signal and its relative error is studied for various cuts on the $\chi^2_{increase}$ induced by the new fit to the primary vertex and the results are given in Table 6.2. The relative error on the signal increases rapidly when the sample

[4]A cut of $\chi^2 = 9$ is equivalent to a 3 σ confidence level: only 1% of the genuine signal is lost.

$\chi^2_{increase}$	Nr. candidates	Error	Relative error
1	1542	113	7.4 %
4	4015	179	4.4 %
6	4810	195	4.0 %
10	5815	209	3.5 %
20	6729	221	3.2 %

Table 6.2. The variation of the number of candidates with varying the $\chi^2_{increase}$ cut in $\chi^2 < \chi^2_{increase}$, after fitting the D^0 neutral pseudotrack to the primary vertex as described in section 4.3.1. Each row corresponds to a fitted signal. The smallest relative error is achieved when using the entire data sample (no $\chi^2_{increase}$ cut).

size is reduced. The $\chi^2_{increase}$ proves to be a poor discriminant between signal and background. Nevertheless, the $\chi^2_{increase}$ cut can be a useful tool when searching for D^0 candidates not coming from the interaction point, such as in B meson decays.

Decay length. Significance of the decay length

The decay length is the distance between the secondary vertex and primary interaction point: $l = |\vec{l}| = |\vec{x_2} - \vec{x_1}|$. The decay length is signed, as explained in Sec. 4.3. The significance of the decay length is defined as the signed decay length divided by its error:

$$\sigma_l = \frac{l}{\Delta l} \cdot sign(\vec{l} \cdot \vec{P}) \tag{6.2}$$

Pure combinatorial background as well as light flavor decays produce symmetric distributions of the signed decay length and its significance. In Fig. 6.6 the significance distribution of all combinations is plotted. The plot exhibits an excess for positive σ_l with respect to its mirrored negative region. The excess is due to heavy flavor decays, among which charm decays contribute most. The central saddle point is a phase-space effect: as $l \to 0$ so does the $4\pi l^2 dl$ volume element and therefore the number of combinations per volume element in the "decay length" space becomes smaller.

The $m(K, \pi)$ invariant mass distribution has been fitted for various significance cuts $\sigma \geq \sigma_{cut}$ with $\sigma_{cut} \in \{-2, -1, 0, 1, 2, 3, 4, 5, 6, 7, 8\}$. The results are shown in Table 6.3. A relative error of 3.2 % is achieved at $\sigma_{cut} = 1$. This cut also enhances the charm purity within the selected sample by increasing the signal over

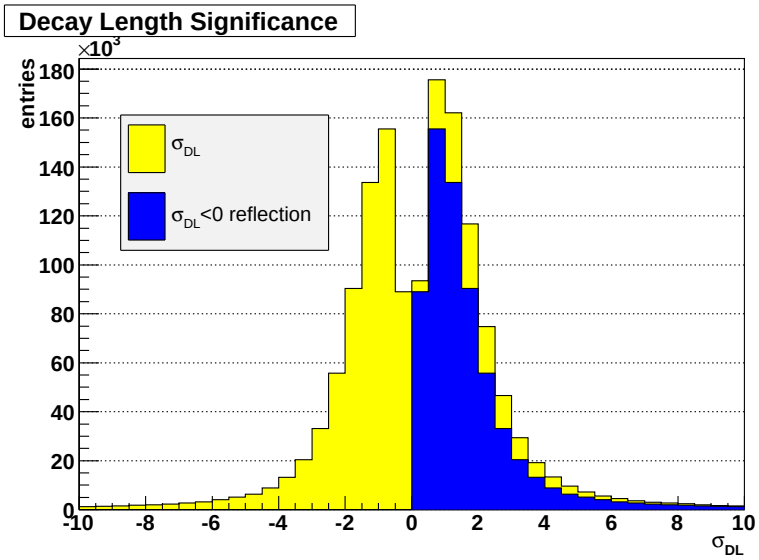


Figure 6.6. The signed decay length significance for all combinations exhibits excess on the positive side due to genuine heavy flavor decays. The negative left side of the plot is also shown as mirrored on the right side.

background ratio.

σ_{cut}	Nr. of candidates	Error	Relative error
-2	6768	267	3.9 %
-1	6250	201	3.2 %
0	5454	177	3.2 %
1	4527	148	3.2 %
2	2707	106	3.9 %
3	1350	76	5.6 %
4	659	59	8.9 %
5	348	48	13.7 %
6	192	42	21.8 %
7	158	38	24.0 %

Table 6.3. Variation of the number of candidates with varying the cut $\sigma_l \geq \sigma_{cut}$. Each row corresponds to a fitted signal. The relative error is smallest when $\sigma_l = 1$.

$c\tau$ of the meson candidate

The proper time is reconstructed as:

$$c\tau = \beta\gamma l = \frac{m}{|\vec{P}|} l = m\frac{l_{xy}}{P_T} \tag{6.3}$$

where $\beta = v/c$ and $\gamma = 1/\sqrt{1-\beta^2}$, l is the signed decay length and $\vec{P}$ the charm meson momentum. The signal versus background ratio has been studied for various cuts of $c\tau > c\tau_{cut}$ with $c\tau_{cut} \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ **mm**. The results are tabulated in Table 6.4. The smallest relative error is when $c\tau_{cut} = 0$ which is implicit when cutting on the D.L. significance: $\sigma_{cut} = 1$.

$c\tau_{cut}$	**Nr. of candidates**	**Error**	**Relative Error**
0.0	5454	177	3.2 %
0.1	4419	151	3.4 %
0.2	2894	117	4.0 %
0.3	1770	92	5.1 %
0.4	969	75	7.7 %
0.5	588	64	10.8 %
0.6	391	56	14.3 %
0.7	250	50	20.0 %
0.8	185	46	24.8 %
0.9	127	42	33.0 %

Table 6.4. The variation on the number of candidates when varying $c\tau_{cut}$. Each row corresponds to a fitted signal value. The smallest relative error is obtained when $c\tau_{cut} = 0$.

So far, the cuts we explored produced larger relative errors when applied, except the decay length significance cut. Unfortunately, the D^0 meson proves to be a difficult particle to track due to its short lifetime. As the error on the secondary vertex reconstruction is of the order of hundred microns, the same as the charm meson $c\tau$, tools that discriminate between charm and light flavor decay and which worked extremely well when reconstructing other (longer lived) charm mesons [46] fail when applied to D^0 decays because at large enough distances/proper times, where the background combinations are very few, the charm meson signal is also depleted by the rapid decay.

6.3 Unfolding the signal

The invariant mass distributions for $m(K,\pi)$ will be shown in bins of transverse momentum P_T and pseudo-rapidity η of the D^0 meson as well as x and Q^2 of the interaction, for the untagged sample. The kinematic range and all the selection cuts are given below:

- *Kinematic range*
  - $Q^2_{DA} > 5\,\mathbf{GeV^2}$
  - $3\,\mathbf{GeV} < P_T(D^0) < 20\,\mathbf{GeV}$
  - $-1.6 < \eta(D^0) < 1.6$
  - $0.02 < y_{DA} < 0.7$
- *DIS selection*
  - primary vertex position: $-30\,\mathbf{cm} < Z(vtx) < 30\,\mathbf{cm}$
  - electron energy: $E(e^-) > 10\,\mathbf{GeV}$
  - $E - P_Z = \delta$ of the event: $40\,\mathbf{GeV} < \delta < 65\,\mathbf{GeV}$
  - $y_{JB} > 0.02$
  - $y_{el} < 0.95$
- *D^0 selection cuts*
  - $P_T(K) > 0.8\,\mathbf{GeV}$, $P_T(\pi) > 0.8\,\mathbf{GeV}$
  - $|\eta(K)| < 1.6$
  - $|\eta(\pi) < 1.6$
  - at least 3 superlayers in the CTD for both K and π
  - at least 2 hits in the MVD for both K and π
  - applying reflection correction for wrongly assigned tracks (reflection subtraction).

A total number of 3156 ± 136 candidates was found.

6.3.1 P_T spectrum

The signal has been reconstructed for the following bins in transverse momentum of the charmed meson: $P_T \in \{3, 3.5, 4.5, 6, 20\}$. The different signals are shown in Fig. 6.7. The change in slope at around 1.73 **GeV** becomes more evident at higher P_T. The width of the signal goes through a minimum value after which increases again, with increasing P_T. The fit values and their errors, for each P_T bin, are given in Table 6.5.

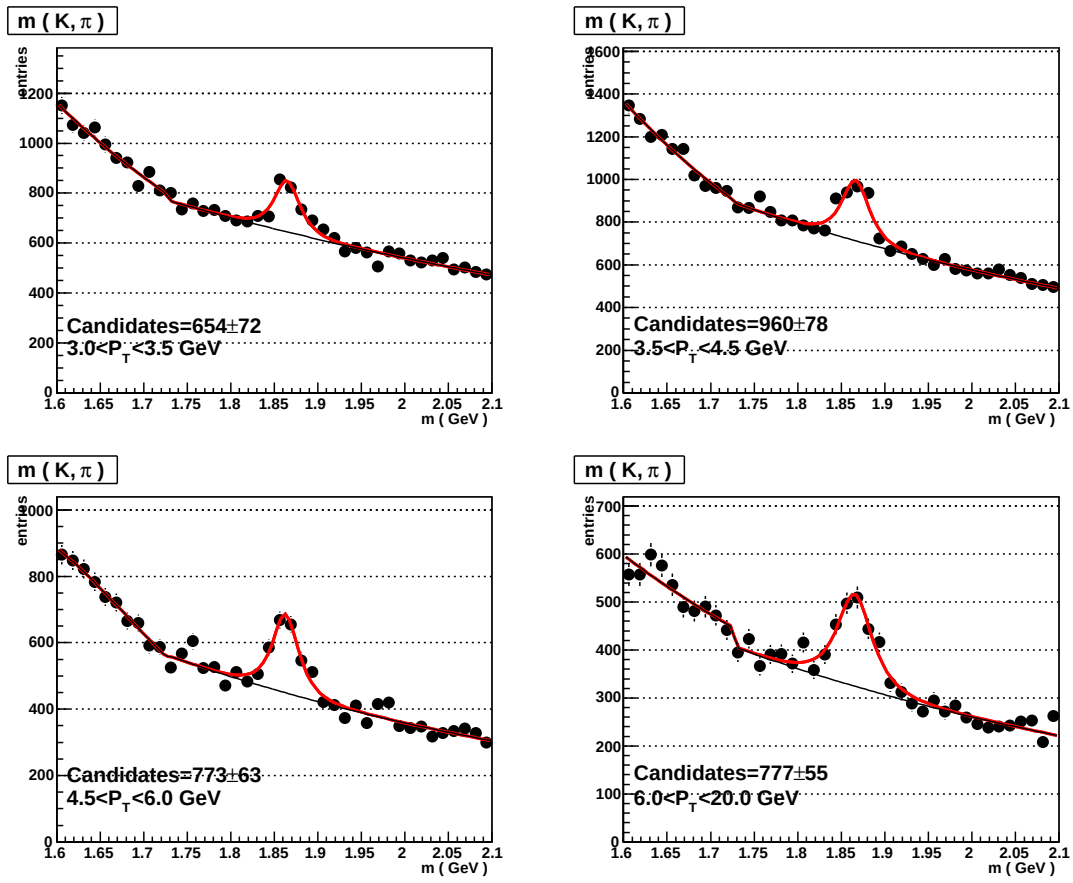


Figure 6.7. $m(K, \pi)$ distribution for 4 transverse momentum bins. The signal corresponds to the untagged sample, reflection subtracted.

6.3.2 η spectrum

The signal has been reconstructed for the following bins in pseudo-rapidity of the charmed meson: $\eta \in \{-1.6, -0.8, -0.4, 0, 0.4, 0.8, 1.6\}$. The different fitted distributions are given in Fig. 6.8. The background is higher in the forward region as

Bin edges [GeV]	Signal	Statistical Error
$3 - 3.5$	654	± 72
$3.5 - 4.5$	960	± 78
$4.5 - 6$	773	± 63
$6 - 20$	777	± 55

Table 6.5. The signal is unfolded in 4 bins of $P_T(D^0)$. The bin edges and the absolute statistical error are given.

there are more forward tracks due to the high momentum of the proton. As a consequence, the (relative) error on the signal is higher in the forward region. The fit values and their errors, for each η bin, are given in Table 6.6.

Bin edges	Signal	Statistical Error
$[-1.6] - [-0.8]$	348	± 43
$[-0.8] - [-0.4]$	590	± 48
$[0.4] - 0$	595	± 56
$0 - 0.4$	654	± 59
$0.4 - 0.8$	590	± 59
$0.8 - 1.6$	397	± 63

Table 6.6. The signal is unfolded in 6 bins of $\eta(D^0)$. The bin edges and the absolute statistical error are given.

6.3.3 Q^2 spectrum

The total Q^2 range has been divided as: $Q^2 \in \{5, 15, 40, 100, 1000\}$. The fitted signals are presented in Fig. 6.9. The fit values and their errors are tabulated in Table 6.7. The highest Q^2 bin exhibits a significantly wider signal width.

6.3.4 x spectrum

The x range of the event has been partitioned as: $x \in \{10^{-4}, 5 \cdot 10^{-4}, 10^{-3}, 3.5 \cdot 10^{-3}, 10^{-1}\}$.The fitted histograms are shown in Fig. 6.10. The fitted signal values and their errors are tabulated in Table 6.8.

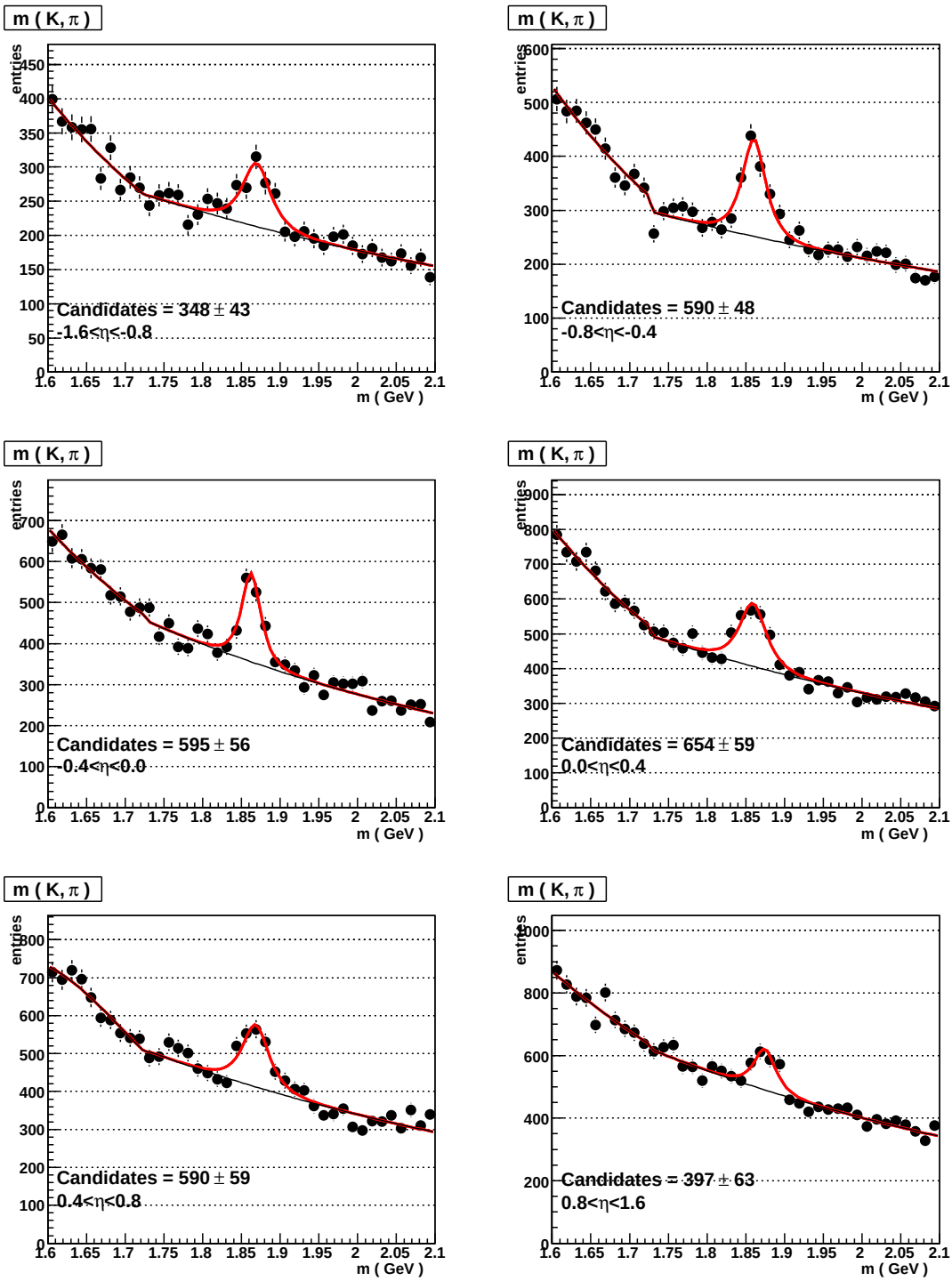


Figure 6.8. $m(K,\pi)$ distribution for 6 pseudo-rapidity bins. The signal corresponds to the untagged sample, reflection subtracted.

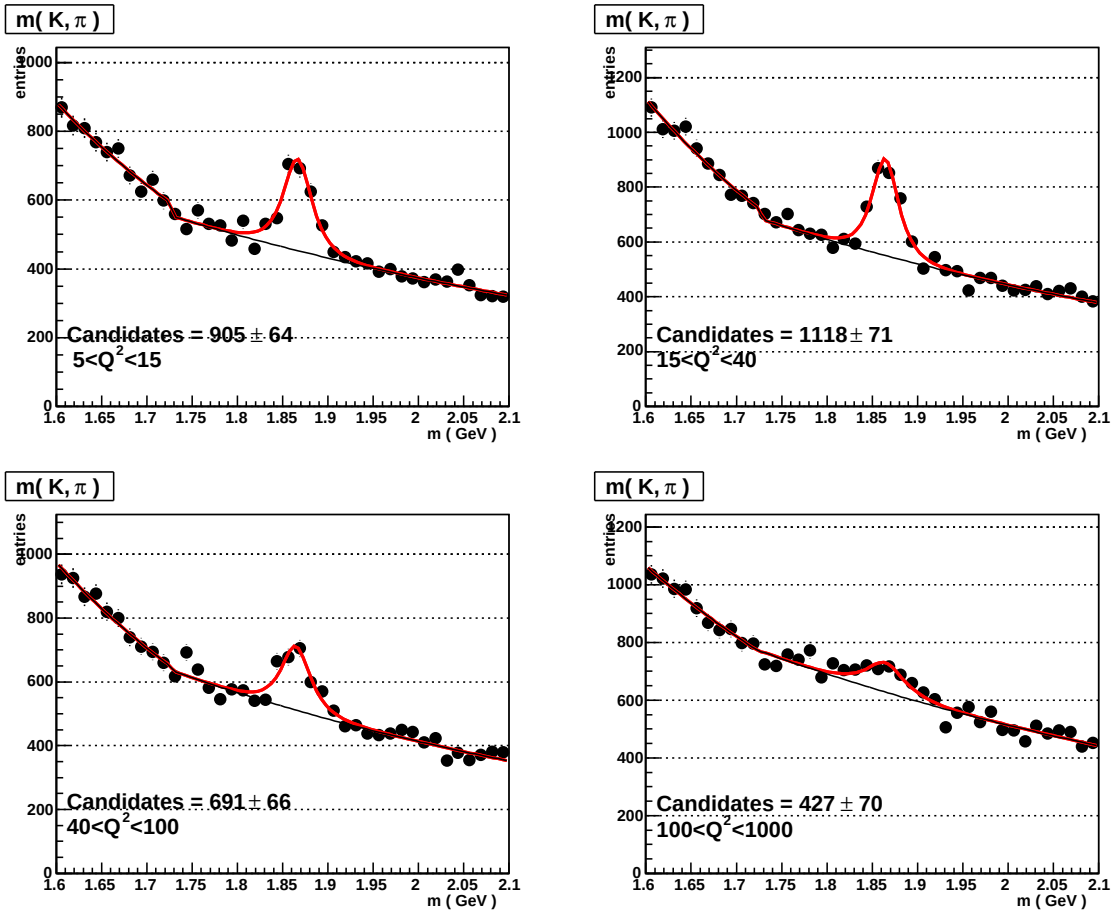


Figure 6.9. $m(K,\pi)$ distribution for 4 Q^2 bins. The signal corresponds to the untagged sample, reflection subtracted.

Bin edges [GeV^2]	Signal	Statistical Error
$5-15$	905	± 64
$15-40$	1118	± 71
$40-100$	691	± 66
$100-1000$	427	± 70

Table 6.7. The signal is unfolded in 4 bins of Q^2 of the event.

Bin edges	Signal	Statistical Error
$10^{-4} - 5\cdot 10^{-4}$	625	± 54
$5\cdot 10^{-4} - 10^{-3}$	647	± 53
$10^{-3} - 3.5\cdot 10^{-3}$	1153	± 78
$3.5\cdot 10^{-3} - 10^{-1}$	708	± 79

Table 6.8. The signal is unfolded in 4 bins of Bjorken x.

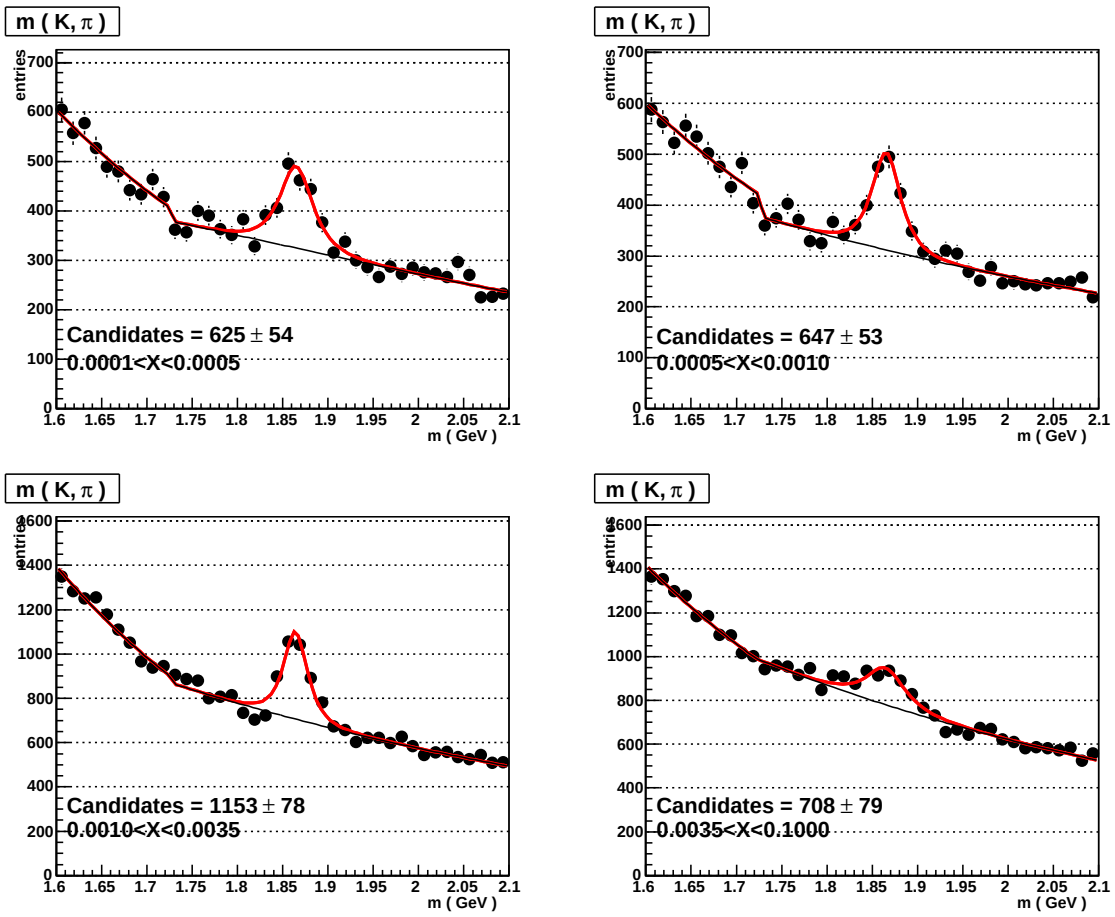


Figure 6.10. $m(K,\pi)$ distribution for 4 x bins. The signal corresponds to the untagged sample, reflection subtracted.

6.4 Comparison between simulation and data

A description of the detector simulation used for the present measurement is given in Sec. 5.5. The same reconstruction code has been run both on data and simulation. As an example, the simulated signal is shown in four bins of Q^2 in Fig. 6.11. Because the simulation does not include light flavors, the simulated signal is reconstructed more easily than in data.

Distributions of signal events as functions of different variables are compared in the data and simulated files. The distributions are area normalized. Each bin corresponds to a fitted signal value. Luminosity normalization was not used as the luminosity normalization factor $\mathcal{L}_{data}/\mathcal{L}_{MC}$ is $\sim 10-15\%$ lower than the area normalization factors for each of the unfolding variables. This effect is caused by the fact that the Monte Carlo simulation employs only leading order calculations.

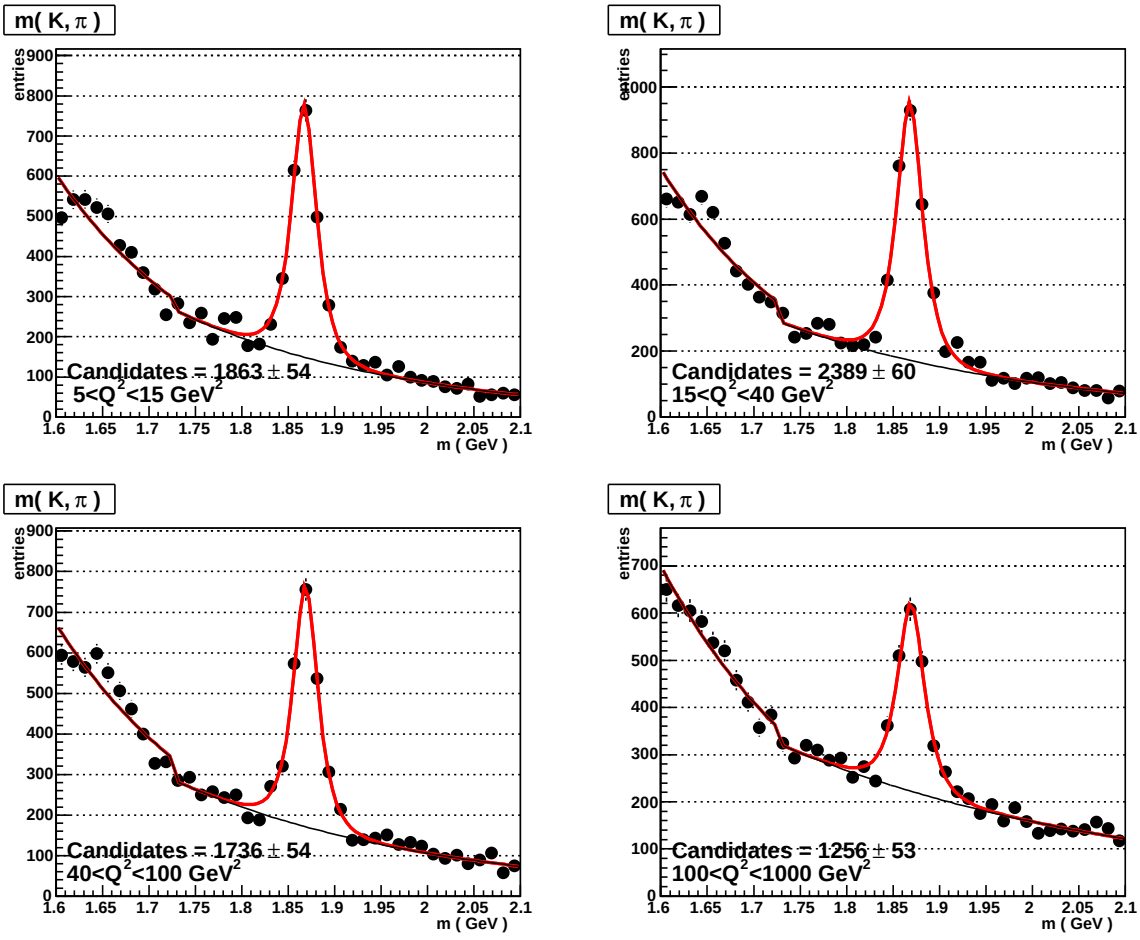


Figure 6.11. The reconstructed signal in the simulated files, in bins of Q^2. This corresponds to a simulated luminosity of $295\,pb^{-1}$, as described in Sec. 5.5.

The signal has been unfolded in:

- bins of the event kinematic variables x, y, Q^2, Z position of the primary vertex, electron energy and $E - P_Z$. Results are presented in Fig. 6.12. There is good agreement in shape. The electron energy distribution exhibits a slight shift of few GeV to higher values in the simulation w.r.t. the data. This also causes the shift seen in the $E - P_Z$ distribution. The DIS cuts are placed safely outside of the affected regions. Also, the simulation presents a slight shift to higher values in the y distribution, but within the statistical errors.

- bins of $P_T(K)$, $P_T(\pi)$, $\eta(K)$ and $\eta(\pi)$ as well as a function of the number of MVD hits of the kaon and the pion tracks. The fit results are shown in Fig. 6.13. The data and the MC samples agree within the errors. The

momentum distributions are well described. The pion pseudo-rapidity is less accurately described in the simulated files, especially at the low and high end of the pseudo-rapidity spectrum. Regarding the number of MVD hits the kaon and the pion tracks collect, a clear shift is seen in the simulation: there are relatively more simulated kaons and pions with more MVD tracks than in the data. This is due to the fact that the MVD suffered from radiation damage (up to 5.7% of the total number of channels were dead in 2005) and this aspect is not simulated at all.

- bins of the decay tagging variables: χ^2 of the secondary vertex, the decay length, the decay length error and the decay length significance. The fit values are shown in Fig. 6.14. The decay length error is the least well described variable. The resolution of the detector is over-estimated in Monte Carlo. This will introduce a significant systematic error in the cross-section measurement. The χ^2 and decay length distributions are well described.

- bins of $P_T(D^0)$ and $\eta(D^0)$. The corresponding distributions of the number of meson candidates are given in Fig. 6.15. The data and the simulation show good agreement in shape. In the central region, the pseudo-rapidity of the charm meson is less accurately described but well within the statistical errors.

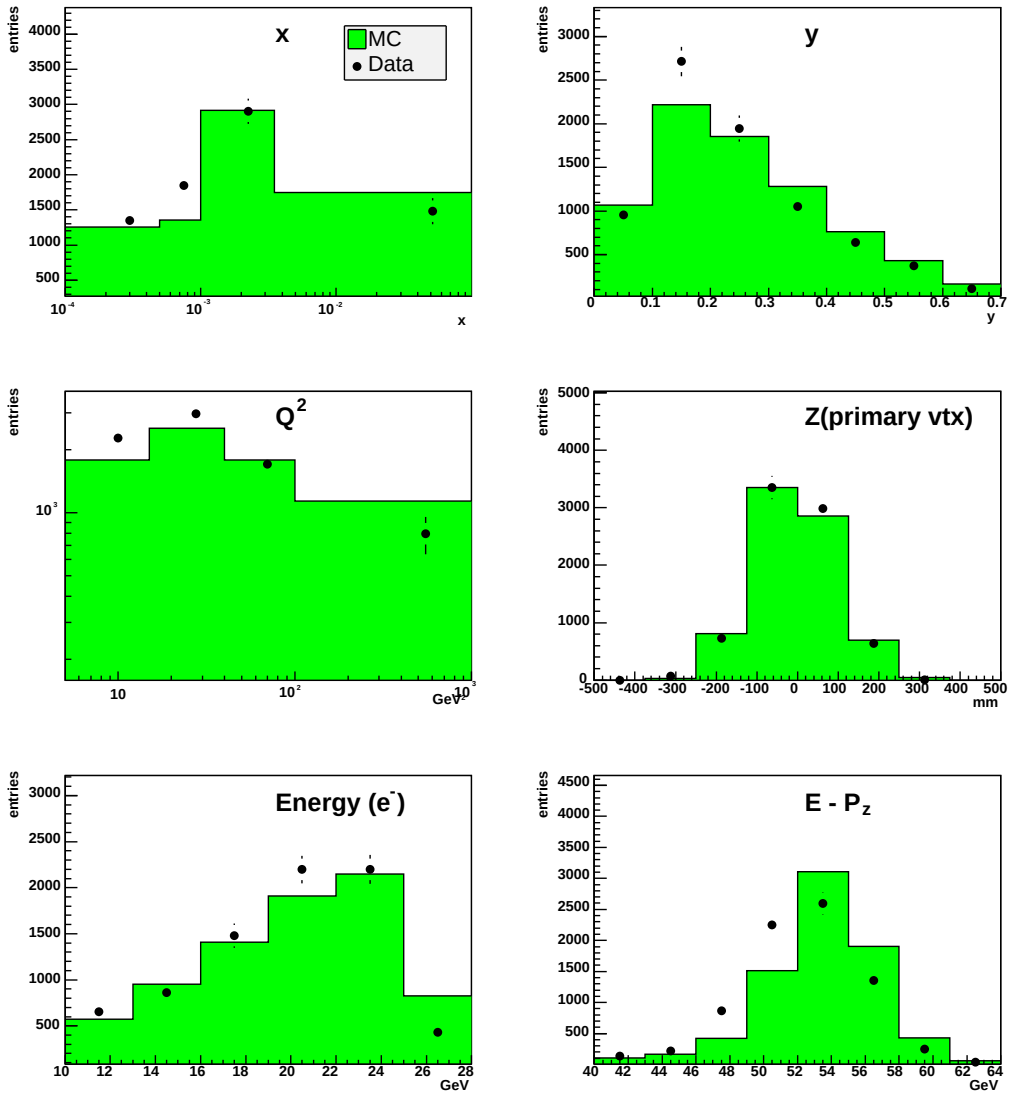


Figure 6.12. Comparison between MC (fill) and data (points): the variation of the number of candidates as a function of x (top left), y (top right), Q^2 (center left), Z_{VTX} (center right), electron energy (bottom left) and $E - P_Z$ (bottom right). Each bin corresponds to a fitted signal value. In each plot, the data distribution has been area-normalized with the Monte Carlo.

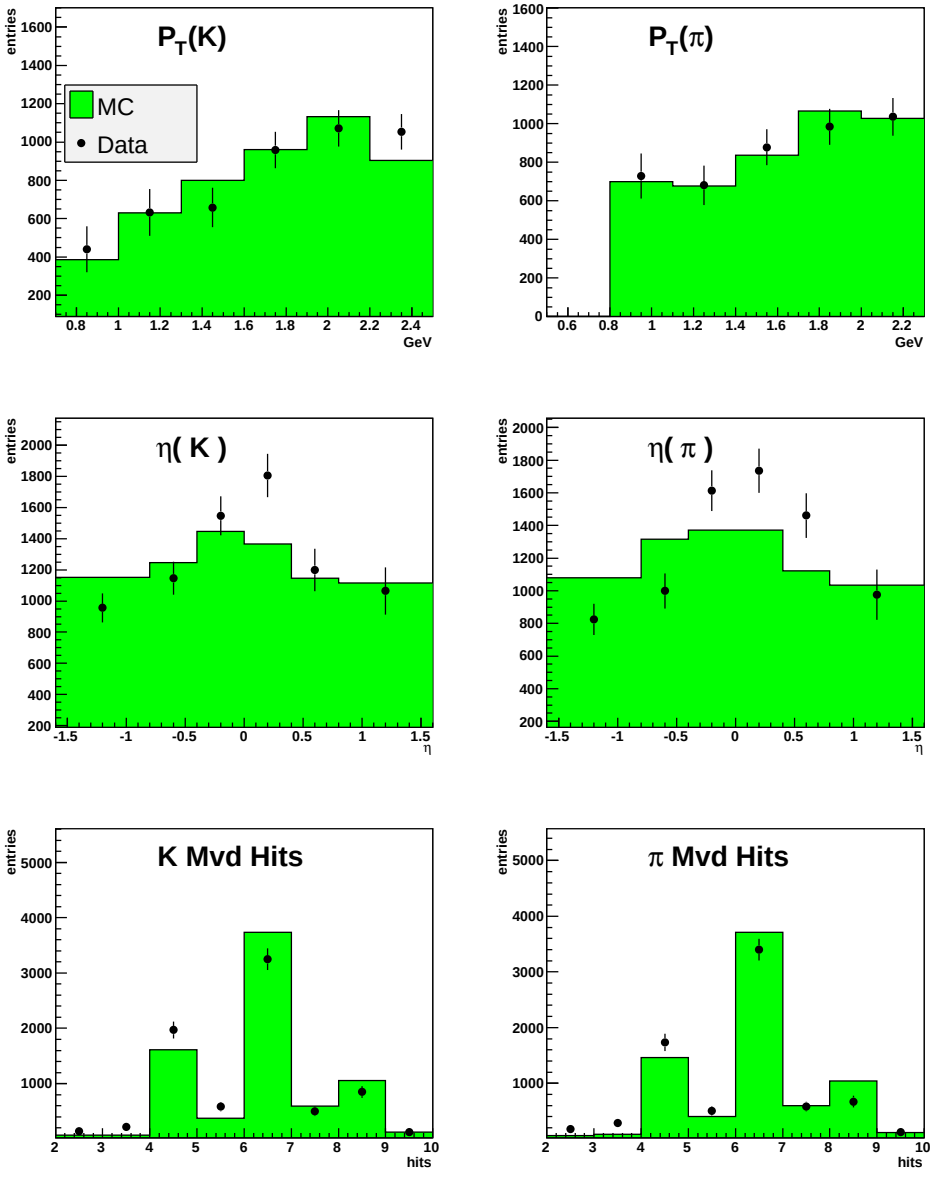


Figure 6.13. Comparison between MC (fill) and data (points): the variation of the number of candidates as a function of $P_T(K)$(top left), $P_T(\pi)$(top right), $\eta(K)$ (center left), $\eta(K)$ (center right), nr. of MVD hits of the kaon (bottom left) and nr. of MVD hits of the pion (bottom right). Each bin corresponds to a fitted signal value. In each plot, the data distribution has been area-normalized with the Monte Carlo.

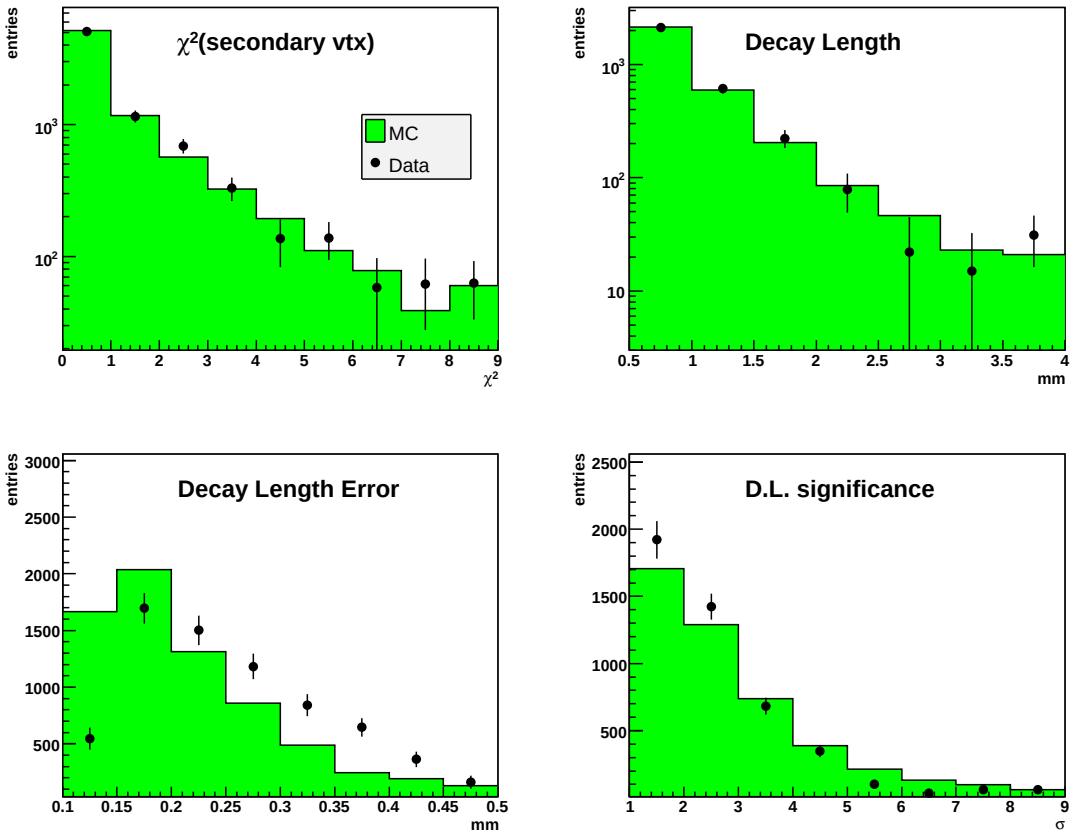


Figure 6.14. Comparison between MC (fill) and data (points): the variation of the number of candidates as a function of χ^2 of the secondary vertex (top left), decay length (top right), decay length error (bottom left) and decay length significance (bottom right). Each bin corresponds to a fitted signal value. In each plot, the data distribution has been area-normalized with the Monte Carlo.

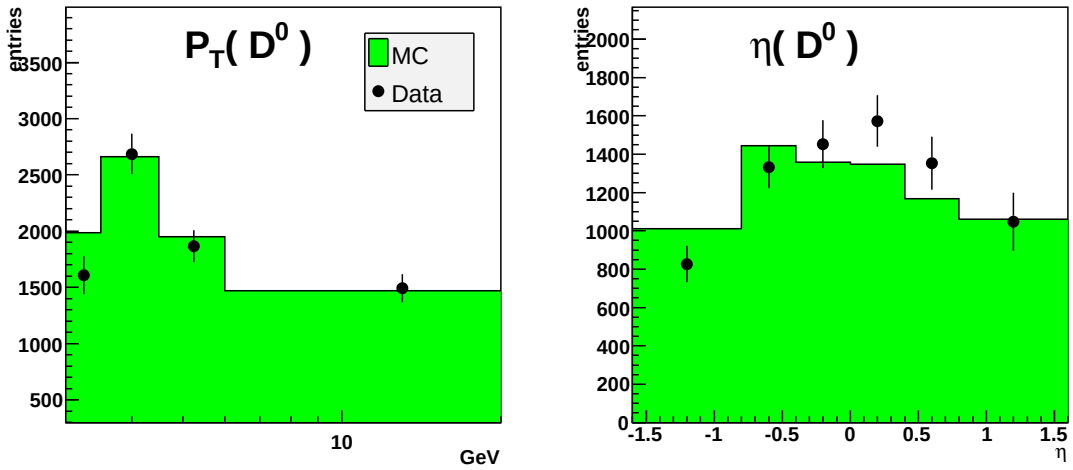


Figure 6.15. Comparison between MC (fill) and data (points): the variation of the number of candidates as a function of $P_T(D^0)$ (left) and $\eta(D^0)$ (right). Each bin corresponds to a fitted signal value. In each plot, the data distribution has been area-normalized with the Monte Carlo.

Chapter 7

Charm cross-sections

In this chapter, the measurements of events containing an untagged D^0 meson (not coming from a D^* decay) in the final state are presented. The charm cross-sections are presented as functions of photon virtuality Q^2 of the event, transverse momentum P_T of the charm meson, pseudo-rapidity η of the charm meson and Bjorken x. The kinematic region has been chosen such that the results can be compared to previous ZEUS measurements performed on HERA I data.

7.1 Definition of the cross-sections

The measured cross-sections are $d\sigma/dQ^2$, $d\sigma/dP_T(D^0)$, $d\sigma/d\eta(D^0)$ and $d\sigma/dx$. The measurement was performed in the following kinematic region:

- $5\,\mathbf{GeV^2} < Q^2 < 1000\,\mathbf{GeV^2}$
- $0.02 < y < 0.7$
- $3 < P_T(D^0) < 20\,\mathbf{GeV}$

As the measurement required a D^0 meson in the final state which did not come from a D^* decay as the only criterion, the cross-sections include events both involving direct charm decay or charm decay that followed from a beauty meson decay.

The differential cross-section $d\sigma/d\alpha$ for a given observable α was computed as:

$$\frac{d\sigma}{d\alpha} = \frac{N^{data}}{\mathcal{A}\,\mathcal{L}\,\mathcal{B}\,\Delta\alpha} \tag{7.1}$$

where $\mathcal{A}$ is the acceptance in the corresponding α bin, $\mathcal{L}$ is the total luminosity of the data set, $\mathcal{B}$ is the branching ratio for the particular decay channel and $\Delta\alpha$ is

the bin width. The 2005 e^-p data set has a luminosity of $\mathcal{L} = 127.35\,\mathbf{pb^{-1}}$. The branching ratio quoted by the Particle Data Group in 2006 for the $D^0 \to K^- + \pi^+$ decay is $\mathcal{B} = 3.82 \pm 0.07\%$. The detector acceptance $\mathcal{A}$ within a given bin $\Delta\alpha$ was determined from Monte Carlo as the ratio between the number of reconstructed candidates and the number of generated mesons within that bin: $\mathcal{A} = N^{recon}/N^{gen}$ and also includes effects of bin to bin migration.

7.2 Systematic uncertainties

To accurately account for the errors induced by the uncertainties on variables used to select the final event sample, the thresholds of various cuts were varied both to larger and smaller values within the resolution of these variables. This was done both in data and Monte Carlo. The variations on individual cuts are listed in Table 7.1. For some variables, the threshold can be varied only in one direction: varying

Variable	Nominal Thres.	Variation
$z_{vtx}(\mathbf{cm})$	30	± 5
$E - P_Z(\mathbf{GeV})$	$40 - 65$	$\to 45 - 60$
$E(e^-)(\mathbf{GeV})$	10	$\to 11$
$Y_{el.}$	0.95	$\to 0.9$
Y_{JB}	0.02	$\to 0.03$
Box Cut ($\mathbf{cm^2}$)	15×15	$\to 16 \times 16$
MVD Hits (K)	2	$\to 3$
MVD Hits (π)	2	$\to 3$
$P_T(K)(\mathbf{GeV})$	0.8	± 0.1
$P_T(\pi)(\mathbf{GeV})$	0.8	± 0.1
$\lvert\eta(K)\rvert$	1.6	± 0.05
$\lvert\eta(\pi)\rvert$	1.6	± 0.05
$\chi^2_{sec}/ndof$	8	± 1
$\sigma(D.L.)$	1	$\uparrow 2$ and $\downarrow -1$

Table 7.1. List of cuts which generate systematic uncertainties. The nominal thresholds and their variations are shown. For certain variables, variations in only one direction were allowed; the resulting systematic error in these cases were symmetrized.

in the opposite direction would spoil the data set by allowing background events into the sample. To fully take into account the effects of the cut variations, a

detailed description of the background would be necessary. When this is the case, the obtained error is symmetrized.

A nominal cross-section σ_0 is calculated, corresponding to the nominal cuts. Each variation of a given variable i leads to a new cross-section value σ_i . Defining $\delta_i = \sigma_i - \sigma_0$, the upper and lower systematic errors Δ_i^+ and Δ_i^- on the cross-section due to source i are computed as:

$$\Delta_i^+ = \begin{cases} \delta_i \ \ if \ \ \delta_i > 0 \\ 0 \ \ if \ \ \delta_i < 0 \end{cases}$$

$$\Delta_i^- = \begin{cases} \delta_i \ \ if \ \ \delta_i < 0 \\ 0 \ \ if \ \ \delta_i > 0 \end{cases}$$

The total error is then computed by summing all the individual errors in quadrature:

$$\Delta^+ = \sqrt{\sum_i (\Delta_i^+)^2} \tag{7.2}$$

and

$$\Delta^- = \sqrt{\sum_i (\Delta_i^-)^2} \tag{7.3}$$

The systematic error induced by each of the variations from Table 7.1 are shown as a function of Q^2, $P_T(D^0)$, $\eta(D^0)$ and x in Fig. 7.9, 7.10, 7.11 and 7.12. These errors will be discussed in detail below.

7.2.1 DIS selection

The errors due to the DIS selection are small with respect to other sources. To account for these errors, the z_{vtx} of the primary vertex, $E - P_z$, the electron energy, Y_{JB} and Y_{el}, as well as the box cut around the beam pipe were varied. A significant contribution came from the box cut variation at low Q^2 (the electron hits the RCAL closer to the beampipe at low Q^2). The shift between the data and the Monte Carlo electron energy displayed by control plots in Fig. 6.12 also introduces a noticeable systematic error. The size of these errors is estimated to be $\sim 5\%$.

7.2.2 Track momentum and pseudo-rapidity

The uncertainty in track momentum was accounted for by varying the momenta of both the kaon and the pion tracks by $\pm 0.1\,\mathbf{GeV}$. The effect on the final cross-section was negligible. The pseudo-rapidity cut was also varied by ± 0.05 units for both tracks. This effect was less than 8%, despite the slight disagreement in pseudo-rapidity seen in control plots in Fig. 6.13.

7.2.3 Uncertainty in the MC beauty fraction

In addition to charm mesons directly produced at the interaction point, the simulated sample contains a small number of charm candidates coming from beauty meson decays. In total, 1147 *true* D^0 generated mesons originated from beauty meson decays, compared to 33679 D^0 mesons generated at the interaction point. Both numbers correspond to an equivalent luminosity of $295\,pb^{-1}$. The charm mesons originating from beauty decay form *the beauty fraction*. In order to determine how sensitive the measurement is to the beauty fraction, the Monte Carlo beauty fraction was doubled and its induced error on the cross-section symmetrized. The effect was not significant with respect to other error sources, indicating that the measurement is not sensitive to beauty production.

7.2.4 Luminosity

The luminosity at ZEUS, for the 2005 data set, is known within 2 % accuracy. This error is not added to the total systematic error.

7.2.5 Number of MVD hits

The control plots in Fig. 6.13 revealed that in the Monte Carlo more signal events are reconstructed where both the K and the π daughter tracks have more hits in the MVD when compared to tracks in the data. This is a known flaw of the Monte Carlo and it is caused by some dead strips in the MVD not being simulated in Monte Carlo. Therefore, the simulation contains relatively more tracks, on the average, with a higher number of MVD hits. In this analysis, a minimum of 2 hits in the MVD was required for both the K and the π. The systematic error due to this cut was computed by changing the threshold to 3 hits for each of the two tracks. This error is significant and is at the level of 5%.

7.2.6 Signal extraction

A significant statistical error appears in the final cross-section measurement due to the signal extraction. The systematic error associated with signal extraction has also been investigated. In particular, due to the fact that the Monte Carlo does not contain light flavor decays, the signal over background ratio is much enhanced in Monte Carlo. This leads to an inaccurate modeling of the background w.r.t. the data sample. The fit for extracting the number of candidates is performed by the MINUIT package [44] within ROOT. The fit function is given in Eq. 6.1. As a systematic variation, the background component of the function was changed to a linear function. The error introduced by changing the background function is not significant, being at the level of 2-3% for most of the bins.

7.2.7 Reflected signal subtraction

Wrong mass assignment to the daughter track pair pollutes the signal. An iterative procedure of subtracting the reflected signal was explained in detail in Chap. 6. Despite the fact that the procedure is applied independently in data and in the Monte Carlo, significant errors are introduced because statistical fluctuations are scaled up during the iterative process. To account for this effect, all cross-sections were recomputed without subtracting the reflected signals for any of the fits. The errors were symmetrized. Although in some bins the errors reach the 10-15% level, especially in the η bins, they are within the statistical errors.

7.2.8 Uncertainty due to the D.L. significance

Tagging the charmed meson decay and using decay related variables to enhance the charm purity of the sample is another source of systematic errors. When comparing data with simulation, as in Fig. 6.14, an overall good agreement is seen for the entire spectrum of variables such as χ^2 of the secondary vertex and the signed decay length itself. A disagreement is seen when comparing the data and simulated signals in bins of the decay length error, namely there are relatively more candidates in the Monte Carlo with small decay length errors. Errors are underestimated in the Monte Carlo for various reasons: inappropriate treatment of dead material, alignment issues, dead cells in the MVD or noisy detector cells not correctly accounted for and so on. Therefore, event selection by decay significance has a strong dependence on the decay length error. This is investigated further.

One defines the signal efficiency associated with the decay length significance cut as:

$$\epsilon = \frac{N_\sigma}{N_R} \tag{7.4}$$

where N_σ is the number of signal candidates at a minimal particular decay significance cut σ_{DL} and N_R is a reference signal, extracted without a significance cut. In Fig. 7.1, the signal efficiency in data and Monte Carlo are compared. Differences are visible. The relative data efficiency ϵ_{rel}, defined below, quantifies these differences.

$$\epsilon_{rel} = (\epsilon_{data} - \epsilon_{MC})/\epsilon_{MC} \tag{7.5}$$

ϵ_{rel} is also shown in Fig. 7.1. The simulation describes well the data up to a decay

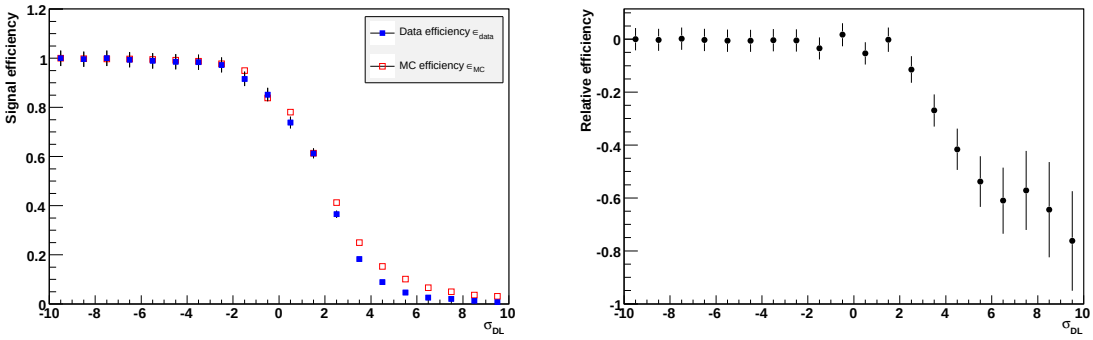


Figure 7.1. Left plot: Signal efficiency as a function of the D.L. significance cut, comparison between data, in colored squares, and Monte Carlo, in open squares. Right plot: The relative difference of the two efficiencies as defined in Eq. 7.5.

significance of $\sigma_{DL} = 2$ which is larger than the nominal cut used in this analysis. As the significance has per definition a pull of 1 and because the nominal significance cut is within the range for which the Monte Carlo describes the data, thresholds of one unit away from the nominal value, $\sigma_{DL}^{low} = 0$ and $\sigma_{DL}^{high} = 2$, can be used for calculating the systematic error.

A more conservative approach can be taken. The efficiency ϵ was plotted in Fig. 7.2 as a function of decay length error Δ_{DL}, both for data and for Monte Carlo. The right plot denotes the relative difference between the data and the simulated sample efficiencies. An overall variation of 75% in the relative efficiency difference can be seen (in one direction only) when the entire decay length error spectrum is spanned.

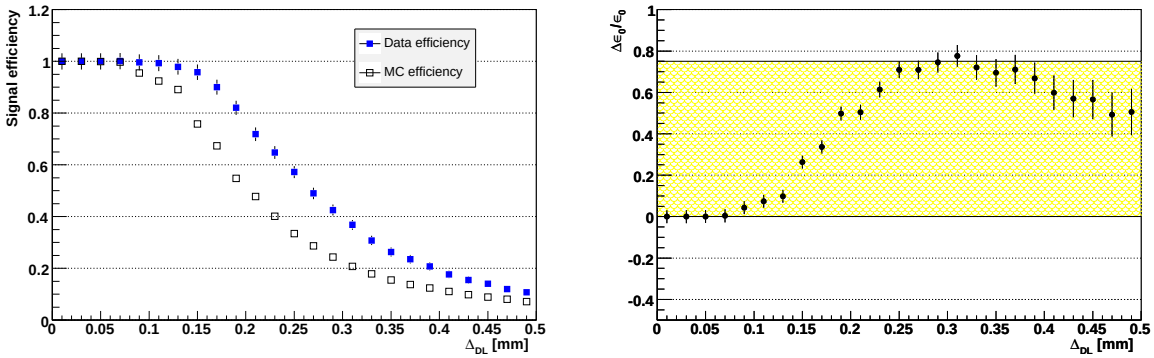


Figure 7.2. Left plot: signal efficiency in data (solid squares) and MC (open squares) as a function of decay length error Δ_{DL}. Right plot: the relative difference of the two signal efficiencies. An overall variation of 75 % in relative efficiency difference is sufficient within the errors to contain the entire efficiency spectrum.

The efficiency variation extracted from the decay length error spectrum in Fig. 7.2 is propagated to the efficiency plot in Fig. 7.3. The horizontal shaded area denotes the 75% efficiency error which has been centered on the nominal cut of $\sigma_{DL} = 1$. A modified arc-tangent function, fitted along the efficiency slope, reaches the outer edges of the error band at of $\sigma_{DL}^{low} = -1$ and $\sigma_{DL}^{high} = 2$. These are the variational thresholds employed in this analysis for calculating the induced systematic error. In conclusion, both the "one unit away from the nominal value" rule and the conservative approach taken above lead to very similar estimates for the variation of the D.L. significance employed in calculating the systematic error. This is confirmed by the fact that the simulation and the data agree well in efficiency for all values of D.L. significance below $\sigma_{DL} = 1$, as shown in Fig. 7.1 (right plot) and that the number of candidates in the region of disagreement, in relation to the total number of candidates, is minute.

It is worth mentioning that a solution for reducing significantly the systematic errors due to σ_{DL} can be found by rescaling the decay length error in the Monte Carlo such that it describes the data as accurately as possible. As an exercise, a simple rescaling of the form $\Delta_{DL}^{new} = 120\%\Delta_{DL}$ was performed only on the simulated samples. The new relative efficiency difference is plotted once more in Fig. 7.4. Only a 20% variation in relative efficiency difference is seen after scaling up the Monte Carlo decay length error. This could be improved further.

The decay significance systematic error contributes the most to the final error with respect to any other sources. Nevertheless, this error is comparable to the

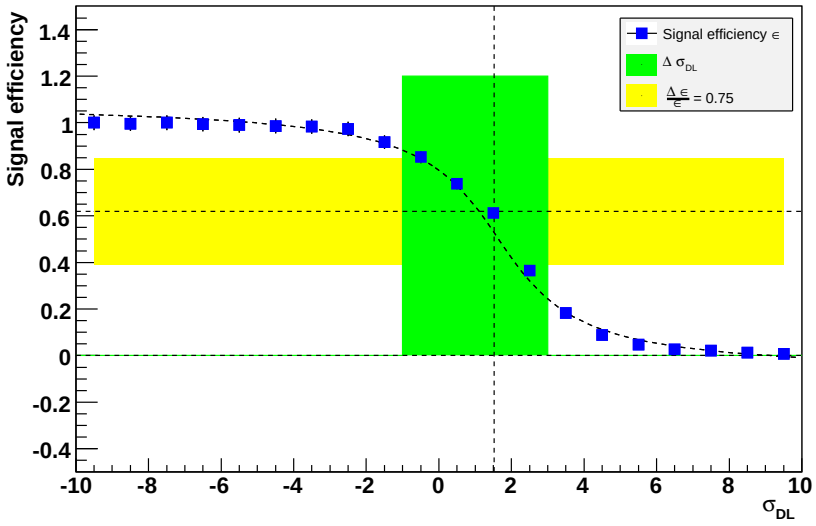


Figure 7.3. The signal efficiency in data as a function of the decay length significance is shown in solid squares. The nominal cut used in this analysis is $\sigma_{DL} = 1$. An arc-tangent function is fitted to the efficiency curve. The horizontal shaded band centered on the nominal cut represents the 75 % variation in relative efficiency extracted from the decay length error spectrum (see Fig. 7.2). A systematic shift in reconstructing σ_{DL} by $\pm\Delta\sigma_{DL}$ is equivalent to a systematic shift in the signal efficiency. Therefore, a variation in significance is computed. The arc-tangent fit intersects the outer edges of the horizontal shaded error band at $\sigma_{DL}^{low} = -1$ and at $\sigma_{DL}^{high} = 2$. These are the two σ_{DL} up and down threshold values used to compute the systematic errors induced by the decay length tag.

statistical error of the measurement.

7.3 D^0 cross-sections

7.3.1 Differential cross-section w.r.t. Q^2: $d\sigma/dQ^2$

The differential cross-section w.r.t. the photon virtuality Q^2 of the event for the process $e^-p \to e^- + D^0 + X$ has been measured for the bins: $Q^2 \in \{5, 15, 40, 100, 1000\}$ **GeV²**. The results are presented in Fig. 7.5 and tabulated in Appendix B, Table B.1. The inner error bars (where visible) denote the statistical errors, the outer error bars represent the systematic errors. The measurement is compared to a second ZEUS analysis performed recently by John Loizides and Dan Nicholass from University College London. The second analysis measured within the same kinematic range as the one defined in Sec. 6.3 but was performed independently, using different tracking and vertexing packages, a slightly larger data set and some different cleaning

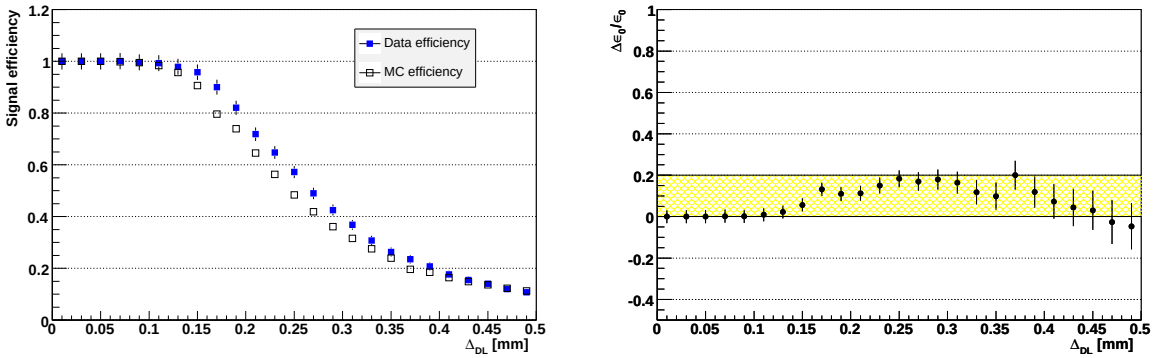


Figure 7.4. Left plot: signal efficiency in data (solid squares) and simulation (open squares) as a function of decay length error Δ_{DL}, after rescaling the Monte Carlo decay length error by a 20% increase. Right plot: the new relative difference of the two signal efficiencies. An overall variation of only 20 % in relative efficiency difference is seen after rescaling.

cuts. The results of a previous ZEUS measurement, using HERA I data [47], are also shown, for the first two Q^2 bins only. The next to leading order theoretical prediction is also displayed, together with its error. Details on how the NLO prediction was calculated are given in Sec. 7.4. The relative difference between the measurement and the theoretical prediction, as defined in Eq. 7.6, is shown separately (bottom plot). The error bars correspond to the relative total error (statistic and systematic); the shaded band is the relative error of the theoretical prediction. There is good agreement in shape between the data and the theoretical prediction but the overall normalization is off by $\sim 15-25\%$. The systematic errors on $d\sigma/dQ^2$ due to the sources listed in Table 7.1 are displayed in Fig. 7.9. The total systematic error (lower plot in Fig. 7.9) is calculated by adding all systematic errors in quadrature. It is worth mentioning that for each systematic variation, the signal extraction is subject to a statistical fit error which cannot be decoupled from the intrinsic systematic error. Therefore, low statistics bins, such as the high Q^2 bin, display larger systematic errors due to the fact that the background and signal fits suffer from statistical fluctuations.

7.3.2 Differential cross-section w.r.t. $P_T(D^0)$: $d\sigma/dP_T$

The differential cross-section as a function of the charm meson transverse momentum $P_T(D^0)$ for the process $e^-p \to e^- + D^0 + X$ has been measured for the following bins: $P_T \in \{3, 3.5, 4.5, 6, 20\}$ **GeV**. The results are presented in Fig. 7.6 and tab-

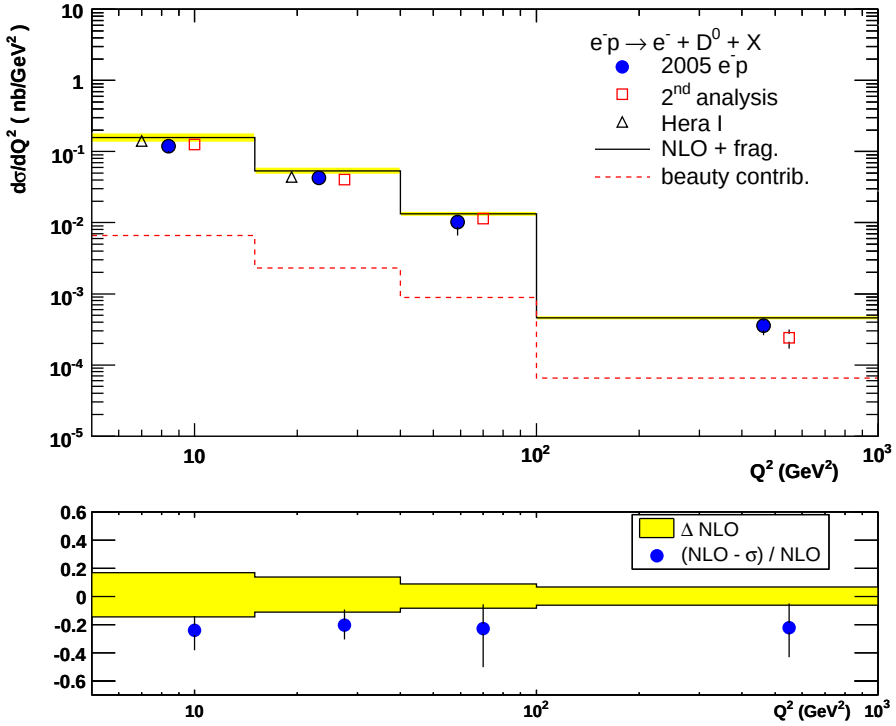


Figure 7.5. The differential cross-section $d\sigma/dQ^2$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to a second independent ZEUS measurement (only statistical error bars shown), to a measurement performed on Hera I data (the first 2 bins only) and to the NLO theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. Bottom plot: the relative difference between the measurement and the NLO prediction is shown, together with its total error. The shaded band represents the error on the prediction.

ulated in Appendix B, Table B.2. The measurement is compared to the second independent analysis as well as to the NLO theoretical prediction. The beauty contribution to the cross-section, computed purely from the simulated sample, is also shown. The relative difference between the measurement and the NLO prediction is shown separately (bottom plot). The error bars correspond to the relative total error (statistical and systematic); the shaded band gives the relative error on the theoretical prediction. There is good agreement in shape but the measured cross-section values are $20 - 30\%$ lower. In the third P_T bin, the measurement agrees the least with the second analysis. The systematic errors on $d\sigma/dP_T(D^0)$ due to uncertainties in measuring the variables listed in Table 7.1 are displayed in Fig. 7.10, for each source independently. The total systematic error was obtained by

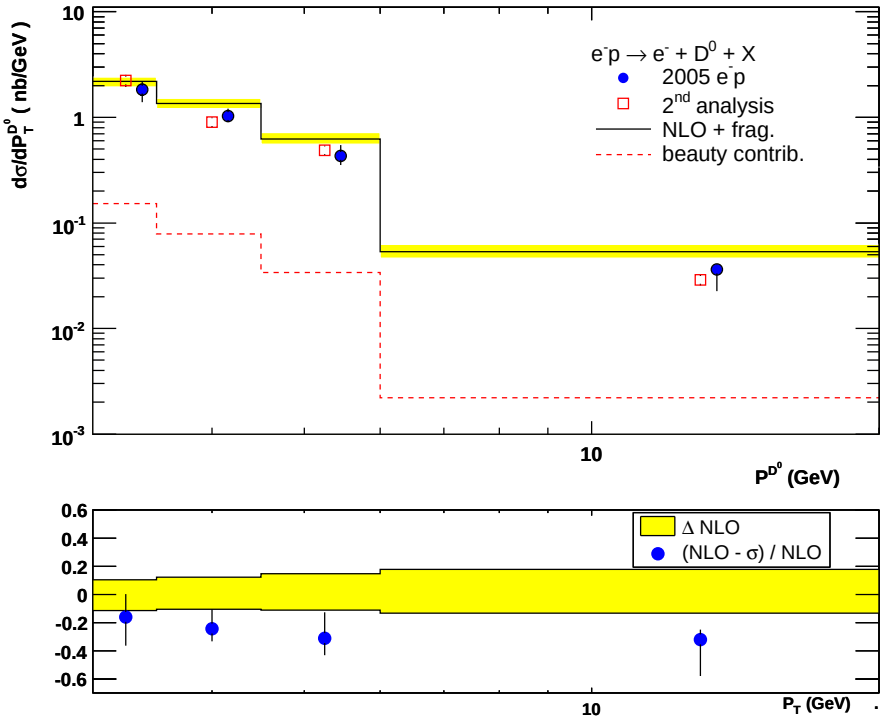


Figure 7.6. The differential cross-section $d\sigma/dP_T(D^0)$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical errors shown) and to the next to leading order theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. The relative difference between the measurement and the NLO prediction is shown separately (bottom plot). The error bars denote the relative total errors (statistical and systematic) whereas the shaded band is the relative error on the theoretical prediction.

adding the individual errors in quadrature. The measured cross-section has a large systematic error in the highest P_T bin due to the systematic error induced by the variation of D.L. significance.

7.3.3 Differential cross-section w.r.t. $\eta(D^0)$: $d\sigma/d\eta$

The differential cross-section as a function of the charmed meson pseudo-rapidity $\eta(D^0)$ for the process $e^-p \to e^- + D^0 + X$ has been measured for the following bins: $\eta \in [-1.6, -0.8, -0.4, 0, 0.4, 0.8, 1.6]$. The results are presented in Fig. 7.7 and tabulated in Appendix B, Table B.3. The measurement is compared to the second independent analysis as well as to the NLO theoretical prediction. The two

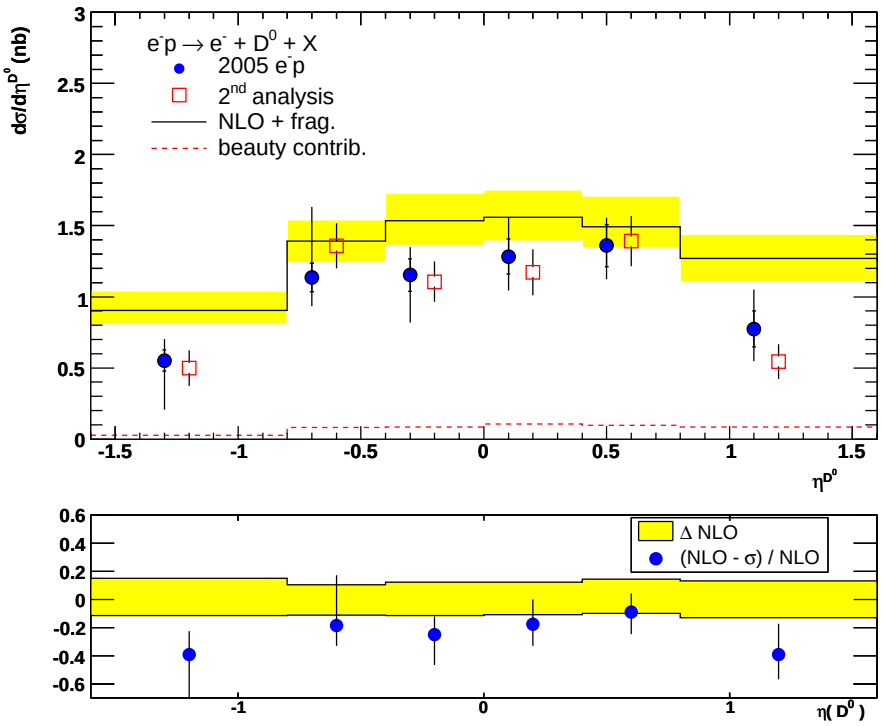


Figure 7.7. The differential cross-section $d\sigma/d\eta(D^0)$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical errors shown) and to the next to leading order theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. The relative difference between the measurement and the prediction is shown separately (bottom plot). The shaded band represents the relative error on the NLO prediction whereas the error bars correspond to the relative total error.

analyses agree within the statistical errors. The beauty contribution to the cross-section, computed purely from the simulated sample, is also shown. The relative difference between the measurement and the NLO prediction is shown separately (bottom plot). Both the relative total errors on the measurement (error bars) and the error on the prediction (shaded band) are also displayed. The NLO calculation overestimates the cross-section by some $10-25\%$ in the central region and almost 40% at high and low η. The overall shape is reasonably reproduced. The systematic errors on $d\sigma/d\eta(D^0)$ due to uncertainties in measuring the variables listed in 7.1 are displayed in Fig. 7.11, for each source independently. The total systematic error was obtained by adding the individual errors in quadrature.

7.3.4 Differential cross-section w.r.t. Bjorken x: $d\sigma/dx$

The differential cross-section as a function of the Bjorken x ,for the process $e^-p \to e^-+D^0+X$, has been measured for the following bins: $x \in \{10^{-4}, 5\times10^{-4}, 10^{-3}, 3.5\times 10^{-3}, 10^{-1}\}$. The results are presented in Fig. 7.8 and tabulated in Appendix B, Table B.4 . The cross-section is concentrated at low values of x. The measurement

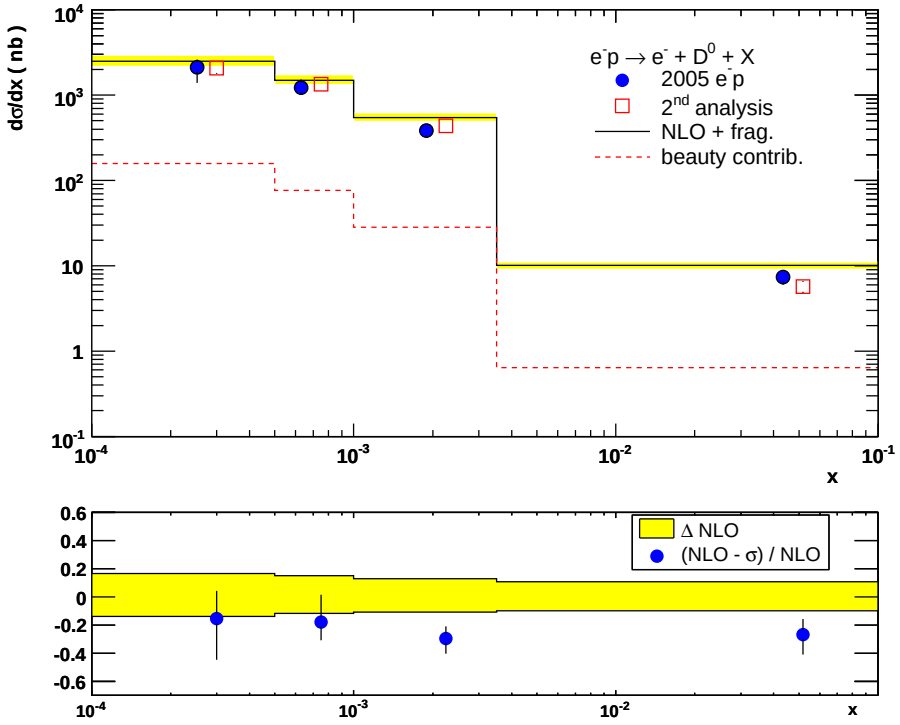


Figure 7.8. The differential cross-section $d\sigma/dx$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical errors shown) and to the NLO theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference.The relative difference between the measurement and the prediction is shown separately (bottom plot). The shaded band represents the relative theoretical error whereas the error bars correspond to the relative total error on the measured cross-section.

is compared to the second analysis and to the next to leading order theoretical prediction. The beauty contribution to the cross-section, computed purely from the simulated sample, is also shown. The second analysis and the present analysis agree well, given the different tracking, vertexing and cut optimization used. The relative difference between the measurement and the theoretical prediction is displayed separately (bottom plot). A good agreement in shape with the NLO is seen but

the overall normalization is $15-30\%$ higher for the predicted cross-section. The systematic errors on $d\sigma/dx$ due to uncertainties in measuring the variables listed in 7.1 are displayed in Fig. 7.12, for each source independently. The total systematic error was obtained by adding the individual errors in quadrature.

7.4 The theoretical prediction

The theoretical predictions for the differential cross-sections, displayed in Fig. 7.5, 7.6, 7.7 and 7.8, were computed using the HVQDIS package [48]. HVQDIS calculates differential cross-sections of charm production in DIS in both leading order and next to leading order. Peterson fragmentation [49] is applied to both the charm and the anti-charm quarks at the hadron level. The Peterson fragmentation function depends only on one parameter, ϵ_P, determined by experiment [50]. Other inputs to HVQDIS are the renormalization and factorization scales μ_R and μ_F, the mass of the charm quark m_c and a particular choice of parton distribution functions. These input values together with their variations are given in Table 7.2.

Quantity	**Value**	**Variation**
Renormalization scale μ_R	$\mu_R = \sqrt{Q^2 + 4M_C^2}$	$2\sqrt{Q^2 + 4M_C^2}$
Factorization scale μ_F	$\mu_F = \mu_R$	larger of: $\frac{1}{2}\sqrt{Q^2 + 4M_C^2}$ or $2M_C$
Peterson parameter ϵ_P	0.035	± 0.030
Charm mass m_C	$1.35\,\mathbf{GeV}$	$\pm 0.15\,\mathbf{GeV}$
Input PDF's	ZEUS NLO PDF	ZEUS PDF: Upper and lower predictions

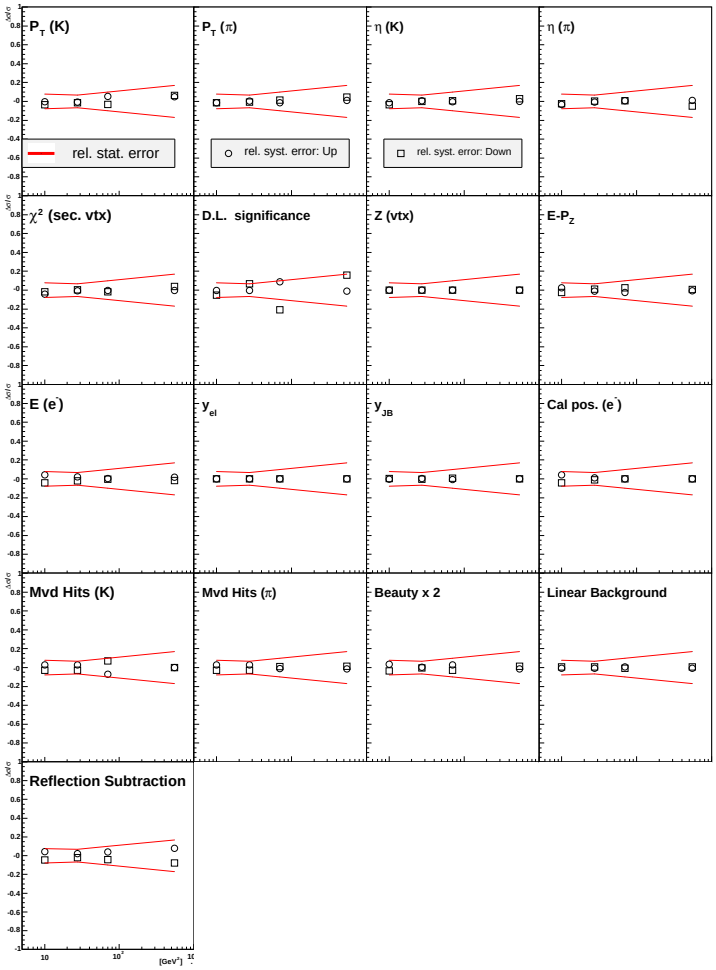
Table 7.2. The list of input parameters to the HVQDIS package used in calculating the central value. The input parameter variations are used to calculate the errors on the central value.

The differences between the measured (σ_{exp}) and theoretical (σ_{th}) differential cross-sections are quantified by $\delta\sigma_{rel}$:

$$\delta\sigma_{rel} = \frac{\sigma_{exp} - \sigma_{th.}}{\sigma_{th}} \tag{7.6}$$

$\delta\sigma_{rel}$ has been plotted as a function of Q^2, $P_T(D^0)$, $\eta(D^0)$, and Bjorken x in Fig. 7.5, 7.6, 7.7 and 7.8 respectively. The error bars on $\delta\sigma_{rel}$ denote both statistical

and systematic errors on the measurement, divided by the NLO central value. The relative error on the prediction is shown as a shaded band. The theoretical predicted cross-sections are on the average 20-30% higher than what the measurement finds. The relative difference is somewhat larger in η, where it reaches 40% at high and low η values.



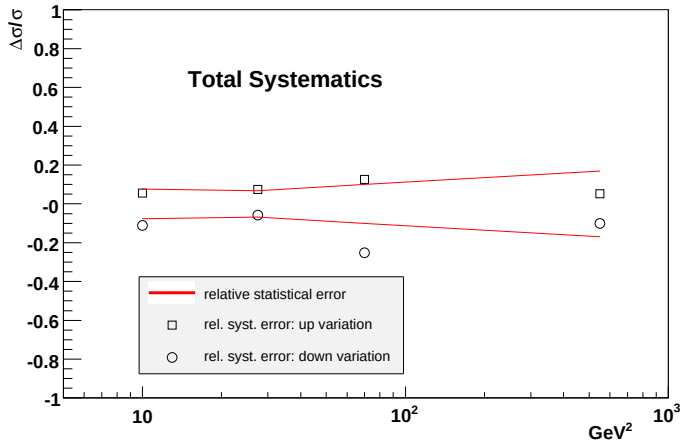
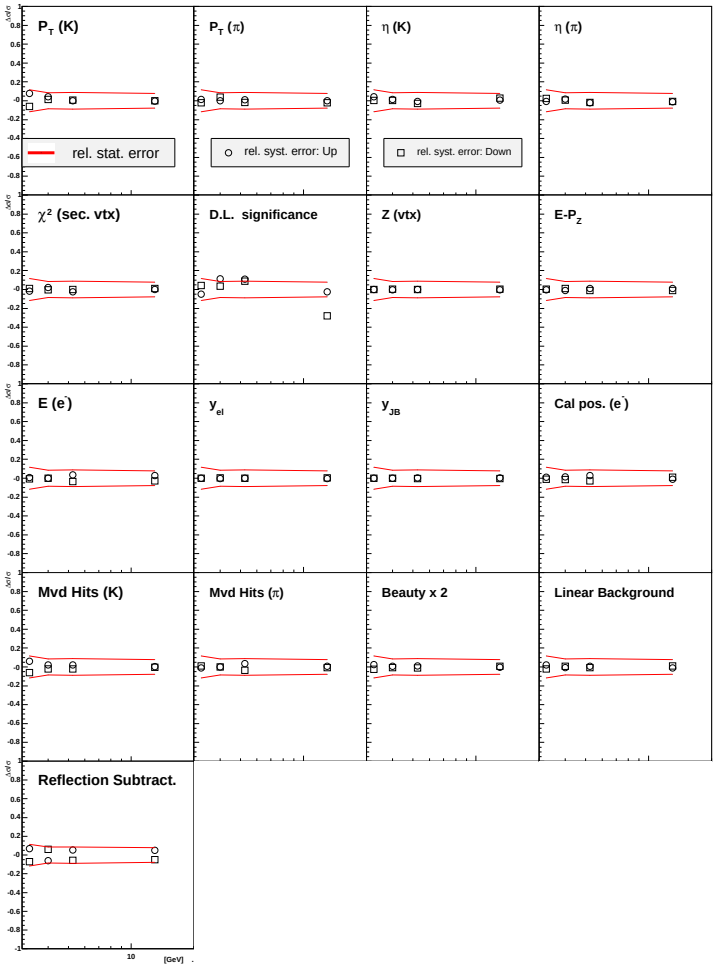


Figure 7.9. The systematic errors, calculated and labelled for each source separately, are displayed as a function of Q^2. A large contribution for all Q^2 bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction. The total systematic error as a function of Q^2 (lower plot) is computed by adding in quadrature all systematic errors. The total systematic error is consistent with the statistical error.



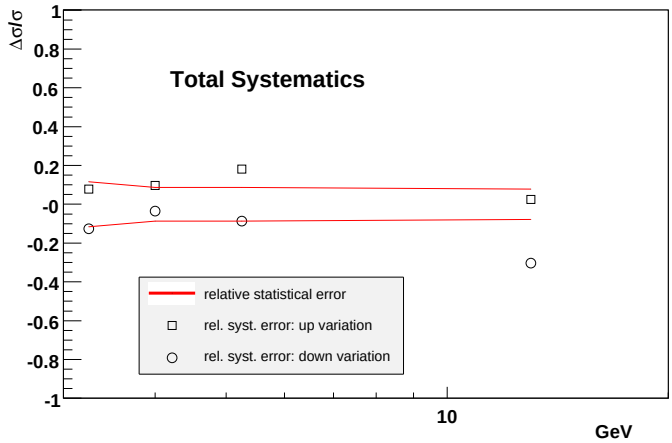
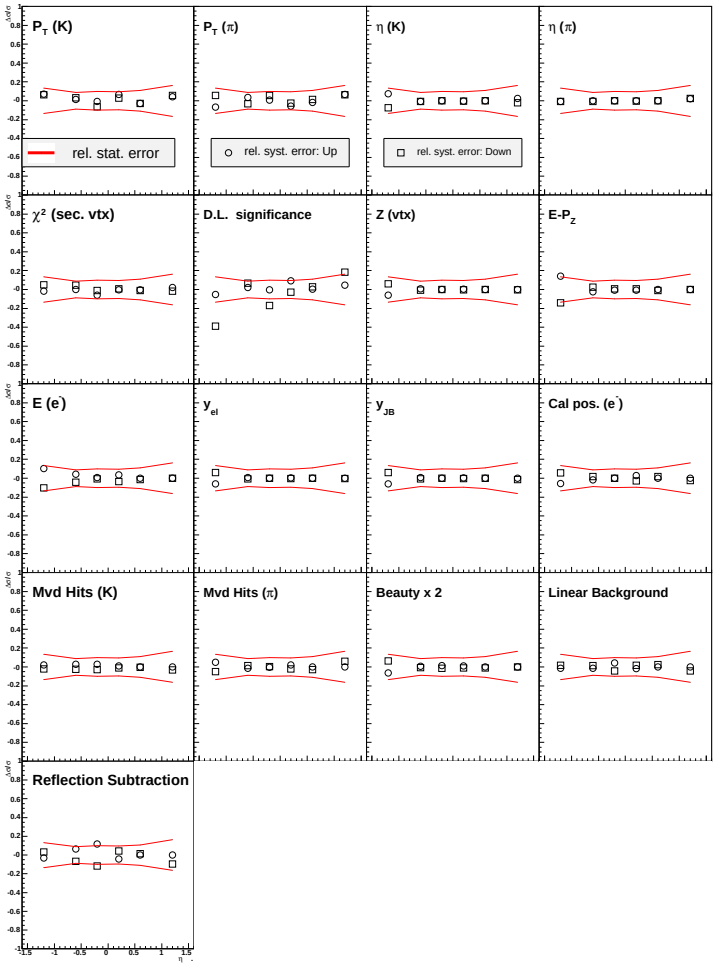


Figure 7.10. The systematic errors, calculated and labelled for each source separately, are displayed as a function of P_T of the charm meson. A large contribution for all P_T bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction. The total systematic error as a function of P_T (lower plot) is computed by adding in quadrature all systematic errors. The total systematic error is consistent with the statistical error.



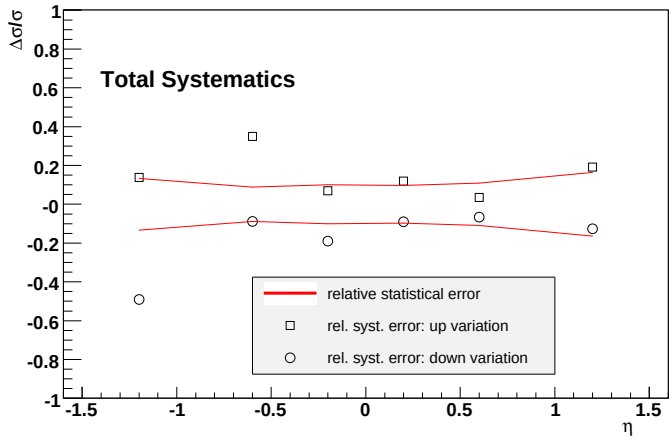
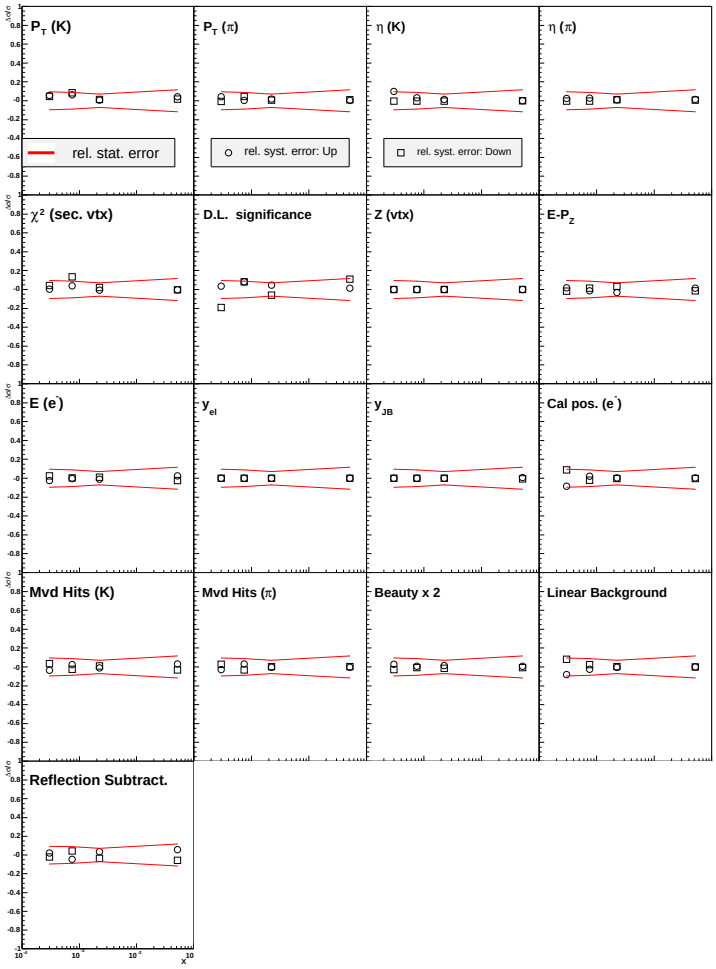


Figure 7.11. The systematic errors, calculated and labelled for each source separately, are displayed as a function of η of the charm meson. A large contribution for all η bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction. The total systematic error as a function of η (lower plot) is computed by adding in quadrature all systematic errors. The total systematic error is consistent with the statistical error.



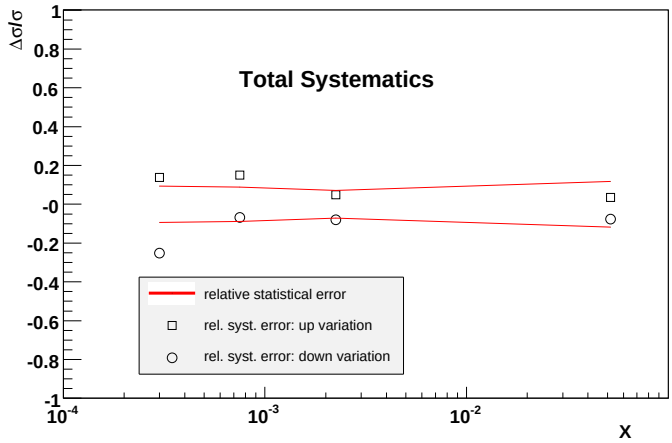


Figure 7.12. The systematic errors, calculated and labelled for each source separately, are displayed as a function of proton momentum fraction x of the struck quark. A large contribution for all x bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction. At very low x, the systematic errors due to cuts on number of MVD hits and changing the background fit function become significant as well. The total systematic error as a function of x (lower plot) is computed by adding in quadrature all systematic errors. The total systematic error is consistent with the statistical error.

Chapter 8

Expanding the kinematic range

Measurements of D^0 production using HERA I data sets have been performed only for the kinematic range $P_T(D^0) > 3\,\mathbf{GeV}$. The measurement presented in the previous chapter is performed for the same kinematic range, using HERA II data. The agreement between the present measurement and previous HERA I published results confirmed that the upgraded ZEUS detector is well understood. After the upgrade, tracking and vertexing improved significantly. Using Hera II data sets, it is now possible to reconstruct D^0 mesons with transverse momenta less than $3\,\mathbf{GeV}$. The aim of this analysis is to measure charm production in a phase-space region unexplored so far at ZEUS, while at the same time taking advantage of the increased statistics. This will allow for a finer granularity of the measurement.

8.1 Reconstructing D^0 mesons at low P_T

Fig. 8.1 shows the invariant mass distribution for candidates with $1.5 < P_T(D^0) < 3\,\mathbf{GeV}$. A clear signal is visible at a fitted mass value of $1.861\,\mathbf{GeV}$. A total of 5483 ± 473 candidates were found at these low transverse momenta. The background is large as no cuts on the decay variables were yet applied. The plot corresponds to a luminosity of $127.35\,\mathbf{pb^{-1}}$, $e^- p$ collisions, recorded by ZEUS in 2005. The cleaning cuts applied on the transverse momenta of the daughter tracks were reduced: $P_T(K) > 0.7\,\mathbf{GeV}$ and $P_T(\pi) > 0.4\,\mathbf{GeV}$. The cut asymmetry is due to the fact that the kaon is more massive.

The kinematic range specified in Sec. 6.3 can therefore be expanded to lower $P_T(D^0)$:

- $Q^2 > 5\,\mathbf{GeV^2}$

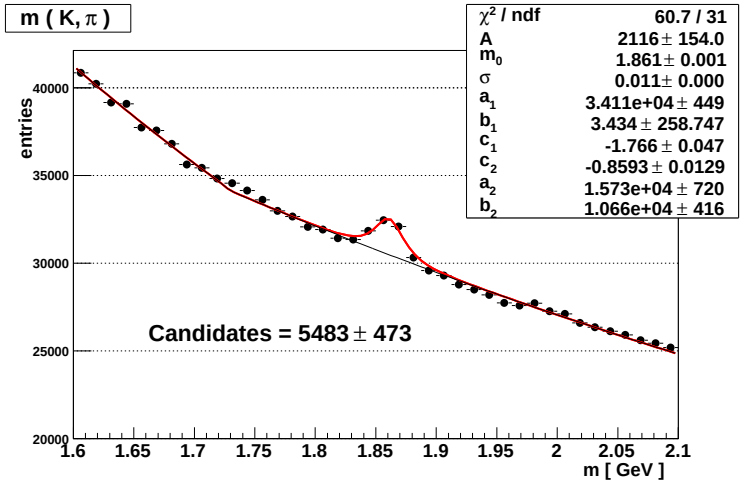


Figure 8.1. Invariant mass distribution for untagged D^0 candidates with $1.5 < P_T(D^0) < 3\,\mathbf{GeV}$.

- $1.5\,\mathbf{GeV} < P_T(D^0) < 20\,\mathbf{GeV}$
- $0.02 < y < 0.7$.

8.2 Optimizing the signal

A similar prescription to the one detailed in Sec. 6.2.5 is applied for optimizing the relative error of the signal in the kinematic range defined above. Table 8.1 samples the variation of the extracted signal and its error when varying the decay length significance cut. The smallest relative error is found for $\sigma_{DL} \in \{-1, 0, 1\}$. The nominal cut $\sigma_{DL}^{cut} = 1$ is chosen as this The secondary vertex selection has already been optimized at the vertexing level, by keeping only good two track vertices with a $\chi^2_{sec} < 16$. When minimizing the relative error, an optimum χ^2 associated with the D^0 decay vertex is found at $\chi^2_{sec} < 8$. The relative error on the signal extraction is minimum when no cut on the proper lifetime is set.

8.3 Data versus simulation

New Monte Carlo files were produced in which the D^0 charm mesons were generated with no minimal transverse momentum cut[1]. Events in the simulation were reweighted so that the decay length significance is well described. In Fig. 8.2, data

[1]Note that the D^0 production is suppressed for $P(D^0)$ values lower than the charm quark mass.

σ_{cut}	**Nr. of candidates**	**Error**	**Relative error**
-2	13694	519	3.8%
-1	12433	466	3.7%
0	11229	411	3.7%
1	9306	344	3.7%
2	5246	237	4.5%
3	2545	166	6.5%
4	1375	129	9.4%
5	856	109	13%
6	542	96	18%
7	429	87	20%

Table 8.1. Variation of the number of candidates with varying the cut $\sigma_{DL} \geq \sigma_{DL}^{cut}$. Each row corresponds to a fitted signal. The relative error is smallest when $\sigma_{DL}^{cut} = 1$.

are compared to simulation. The signal distributions as a function of x, y, Q^2 of the event, Z_{vtx} position of the primary vertex, scattered electron energy and $E - P_Z$ of the event are presented. The MC histograms are area normalized to the data points. Each bin represents a fitted signal value. A good agreement in shape is seen overall. The simulated electron energy exhibits a shift by few GeV to higher values than in the data due to imperfect dead material description in the simulation. This reflects also in the shift in $E - P_Z$ of the event. Although the nominal cuts on the electron energy and $E - P_Z$ of the event are placed safely at the edges of the two distributions, the induced systematic error cannot be neglected.

In Fig. 8.3, the signal is plotted as a function of $P_T(K)$, $P_T(\pi)$, $\eta(K)$, $\eta(\pi)$ and the number of MVD hits of the kaon and pion tracks. Each bin corresponds to a fitted signal value. In general, a good agreement in shape is seen. A systematic shift is visible in the number of MVD hits: both the kaon and the pion have more MVD hits, on the average, in the simulation than in the data. This introduces a significant systematic error.

In Fig. 8.4, the data are compared to the simulation in bins of χ^2 of the secondary vertex, the decay length, its error and its significance. The decay length error is the least well described. The simulated distributions exhibit an excess at low values of the decay length error. The significance distributions agree well due to the fact that the simulated events were generated such that the data is well described for this variable in particular. In Fig. 8.5, the number of candidates is plotted a function

of $P_T(D^0)$ and $\eta(D^0)$. The data and the simulation agree.

In order to investigate whether the disagreement between the reconstructed D.L. errors in data and simulation has any effect on the D.L. significance distributions, in different regions of the phase-space, fitted signal values have been plotted as a function of the D.L. significance in different bins of Q^2 and x of the event, as well as in bins of $P_T(D^0)$ and $\eta(D^0)$. The results are presented in Fig. 8.6, 8.7, 8.8 and 8.9. The D.L. significance is well simulated for all bins. Some bins, such as high Q^2 or high x bins, do not have enough statistics for a proper unfolding of the signal. It is safe to select events using the D.L. significance.

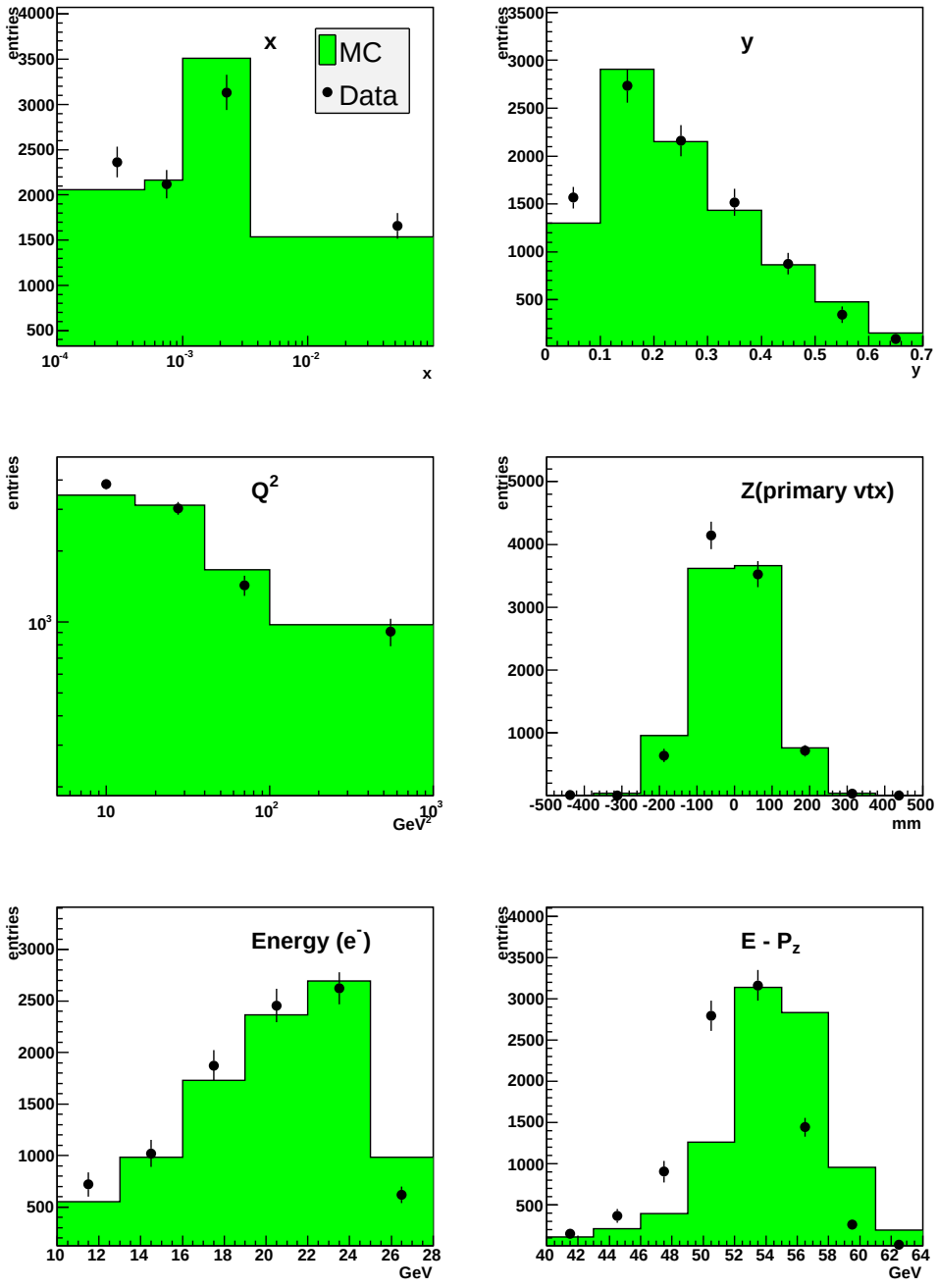


Figure 8.2. Comparison between data and simulation: the number of D^0 candidates as a function of x, y, Q^2 of the event, Z_{vtx} position of the primary vertex, the energy $E(e^-)$ of the scattered electron and $E - P_z$ of the event. An overall good agreement in shape is seen. The simulated electron energy and implicitly the simulated $E - P_z$ exhibit a shift to higher values by few GeV due to imperfections in the dead material description. The simulation histograms have been area-normalized to the data.

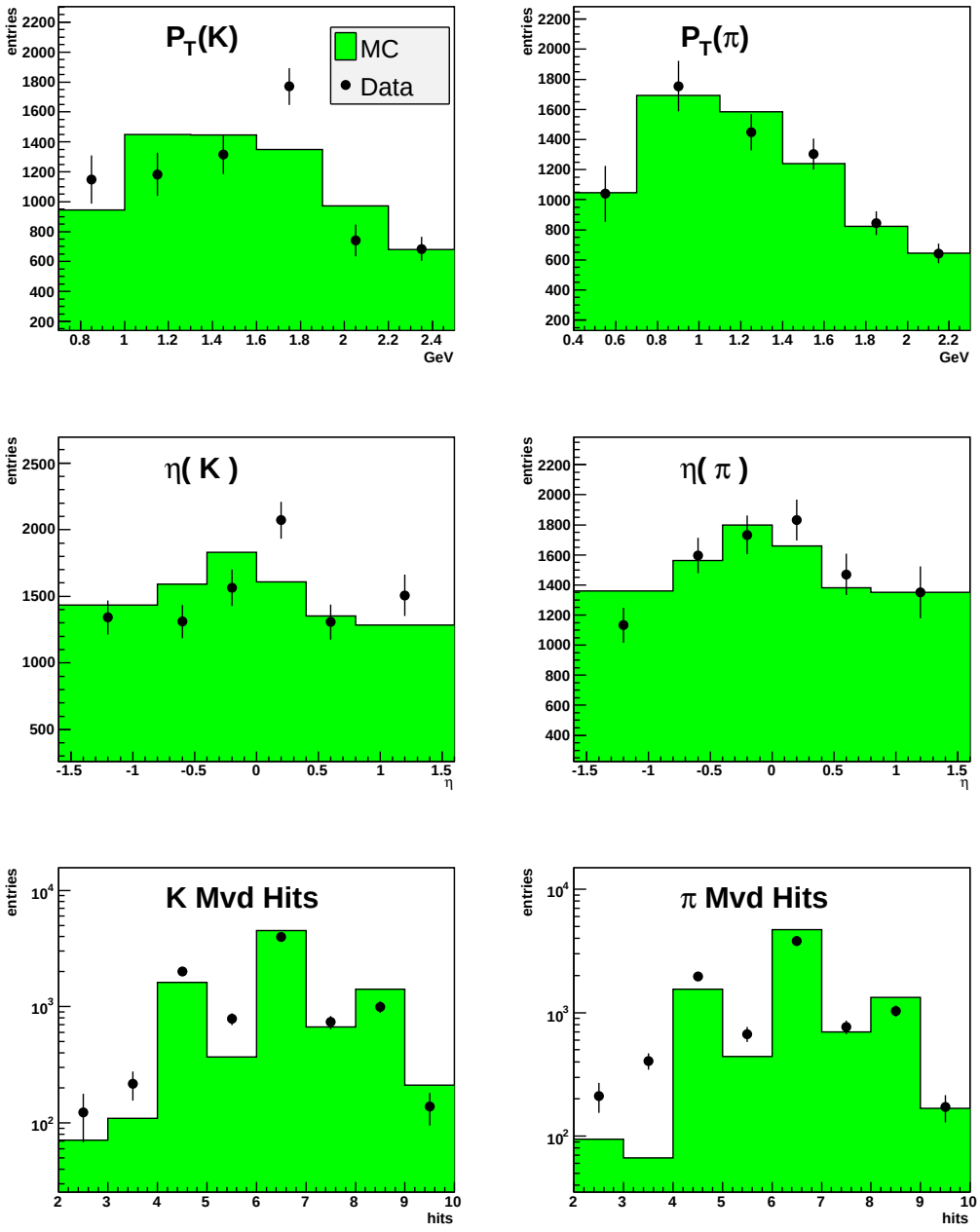


Figure 8.3. Comparison between data and simulation: the number of D^0 candidates as a function of $P_T(K)$, $P_T(\pi)$, $\eta(K)$, $\eta(\pi)$ and the number of MVD hits of the kaon and pion track. An overall good agreement in shape is seen. In the simulated files, there are relatively more (kaon and pion) tracks with higher number of MVD hits than in the data. The simulation histograms have been area-normalized to the data.

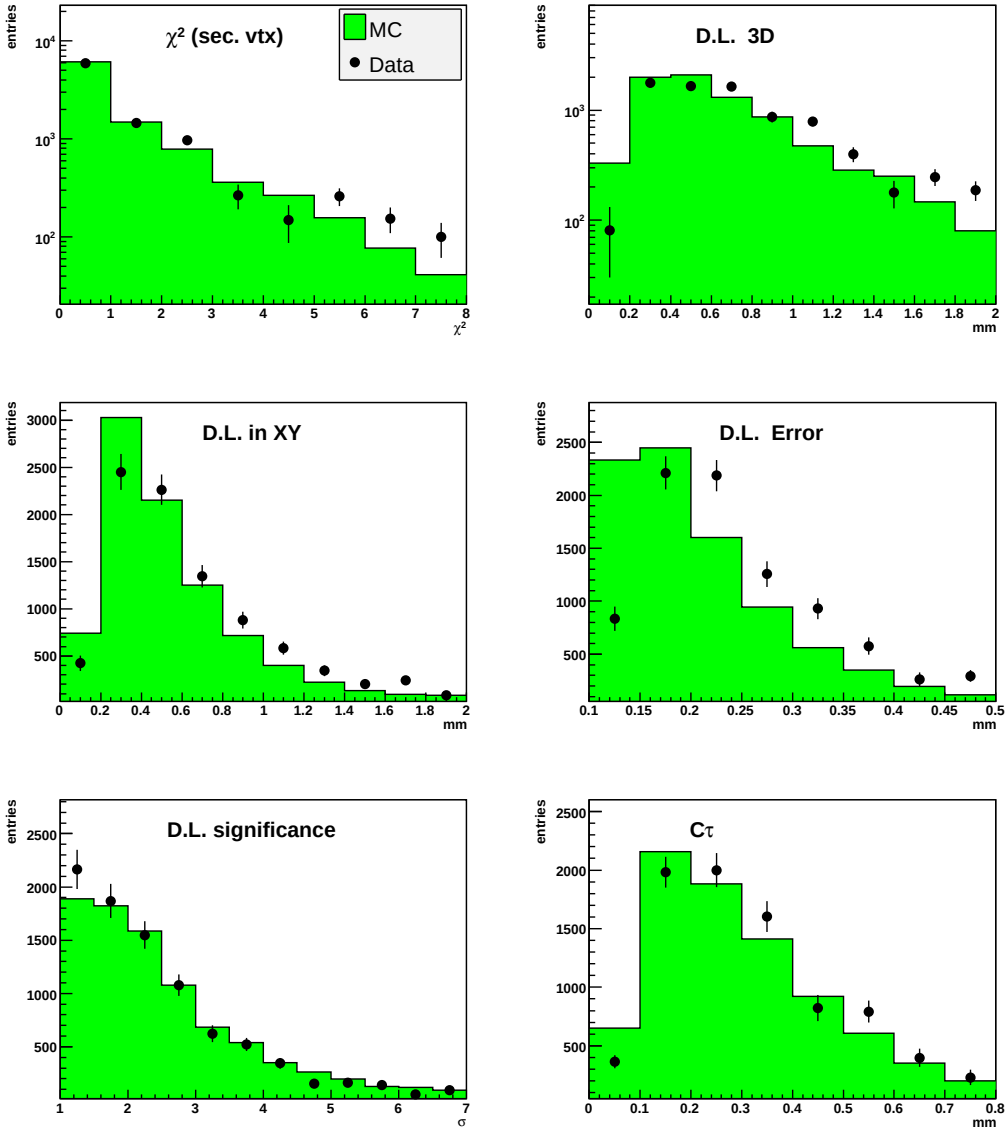


Figure 8.4. Comparison between data and simulation: the number of D^0 candidates as a function of χ^2 of the secondary vertex, the three dimensional decay length, the two dimensional decay length (projection in XY), the two dimensional decay length error, the decay length significance and the proper lifetime. The decay length error is the least well described by the simulation. The simulated events were generated such that the decay length significance describes the data well. The simulation histograms have been area-normalized to the data.

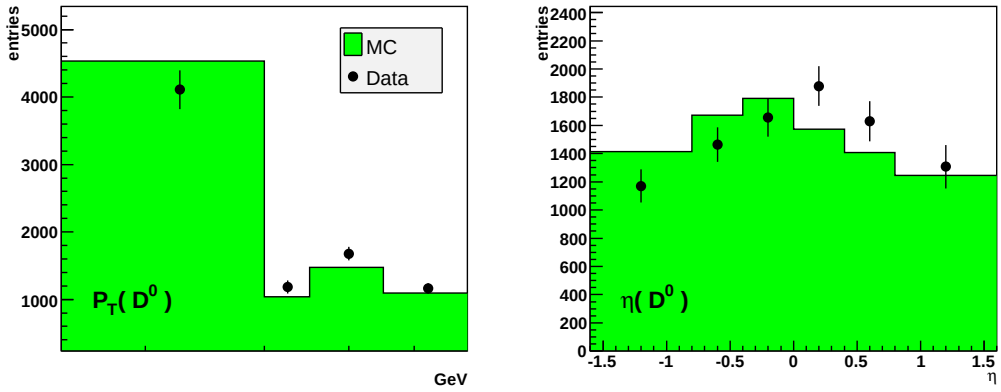


Figure 8.5. Comparison between data and simulation: the number of D^0 candidates as a function of $P_T(D^0)$ and $\eta(D^0)$. Good agreement in shape is seen. The Monte Carlo histograms were area normalized to the data points. Each bin represents a fitted signal value.

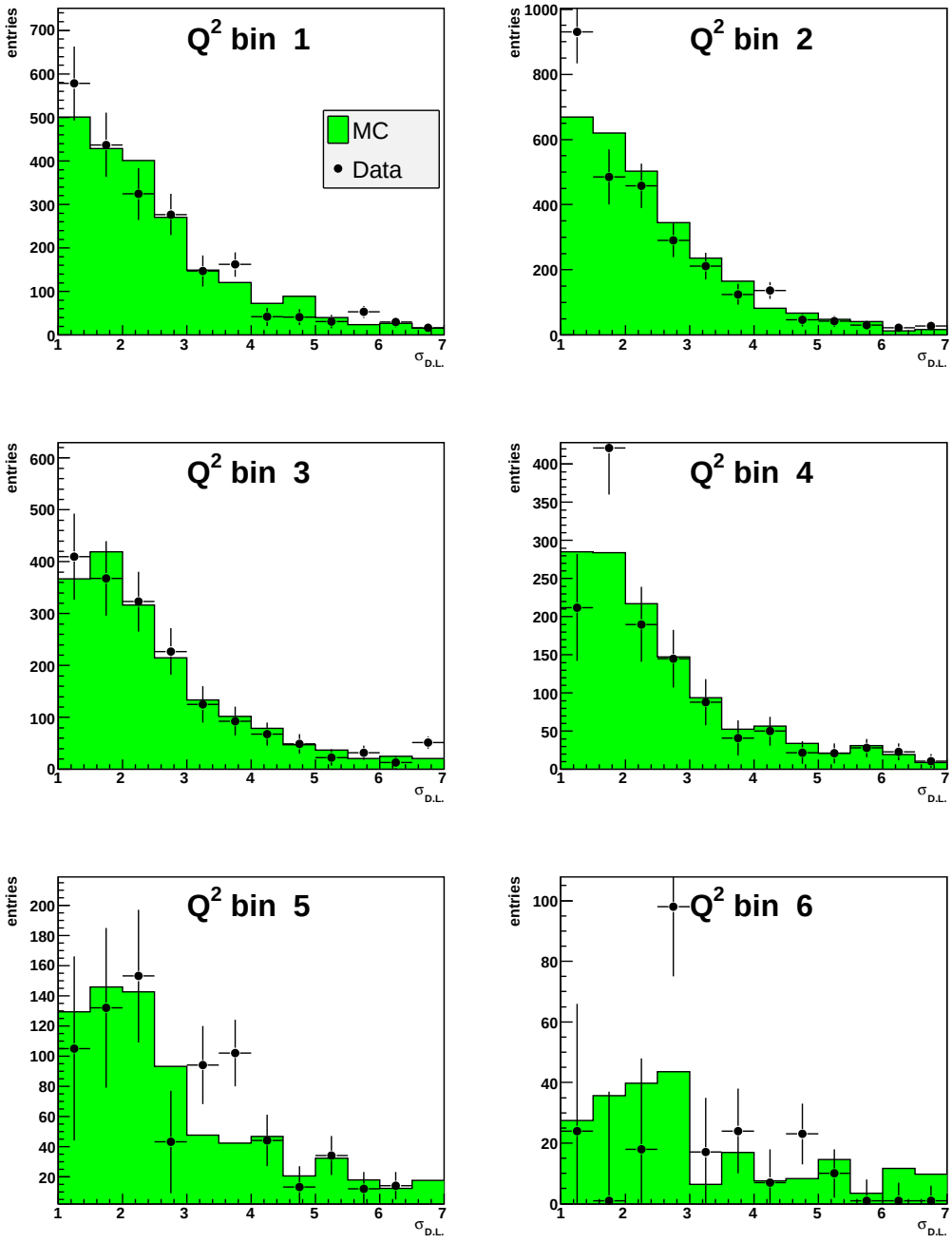


Figure 8.6. Comparison between data and simulation: the number of D^0 candidates as a function of D.L. significance in 6 bins of $Q^2 \in [5, 10, 20, 40, 80, 200, 1000]$ $\mathbf{GeV^2}$ of the event. The histograms are area-normalized. Each bin represents a fitted signal value.

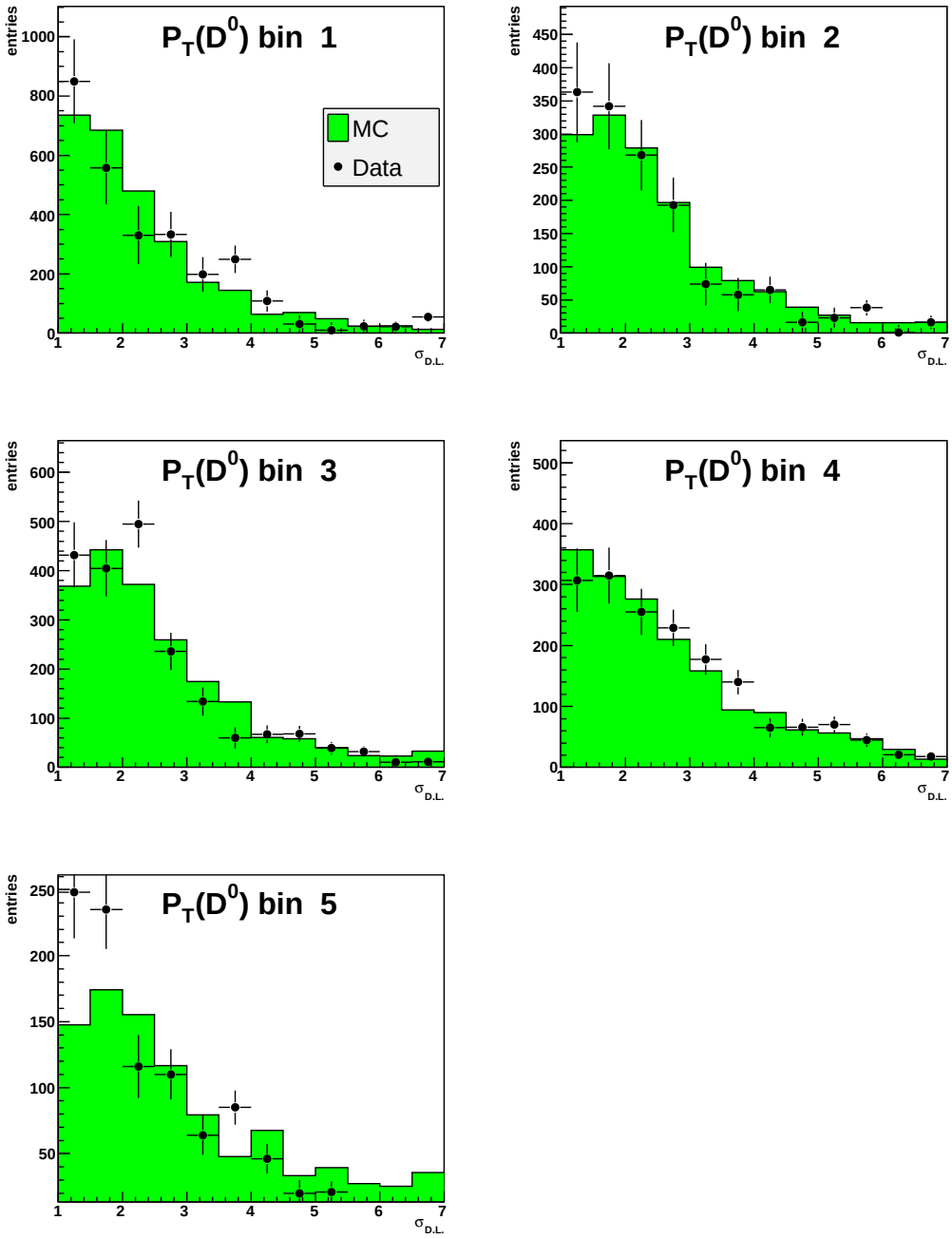


Figure 8.7. Comparison between data and simulation: the number of D^0 candidates as a function of D.L. significance in 5 bins of $P_T(D^0) \in [1.5, 2.4, 3, 4, 6, 20]\,\mathbf{GeV}$. The Monte Carlo histograms were area normalized to the data points. Each bin represents a fitted signal value.

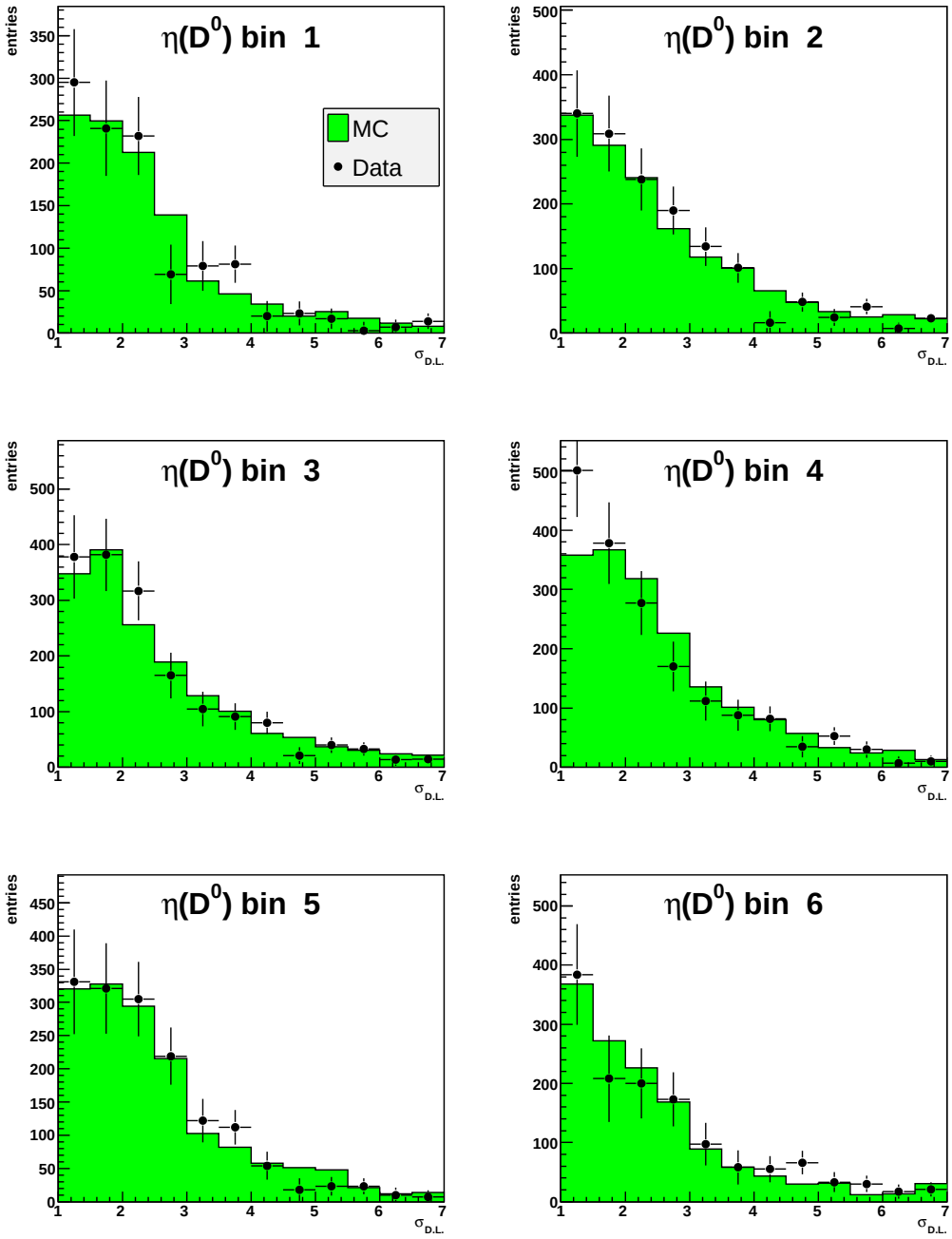


Figure 8.8. Comparison between data and simulation: the number of D^0 candidates as a function of D.L. significance in 6 bins of pseudo-rapidity $\eta(D^0) \in [-1.6, -0.8, -0.4, 0., 0.4, 0.8, 1.6]$. The Monte Carlo histograms were area normalized to the data points. Each bin represents a fitted signal value.

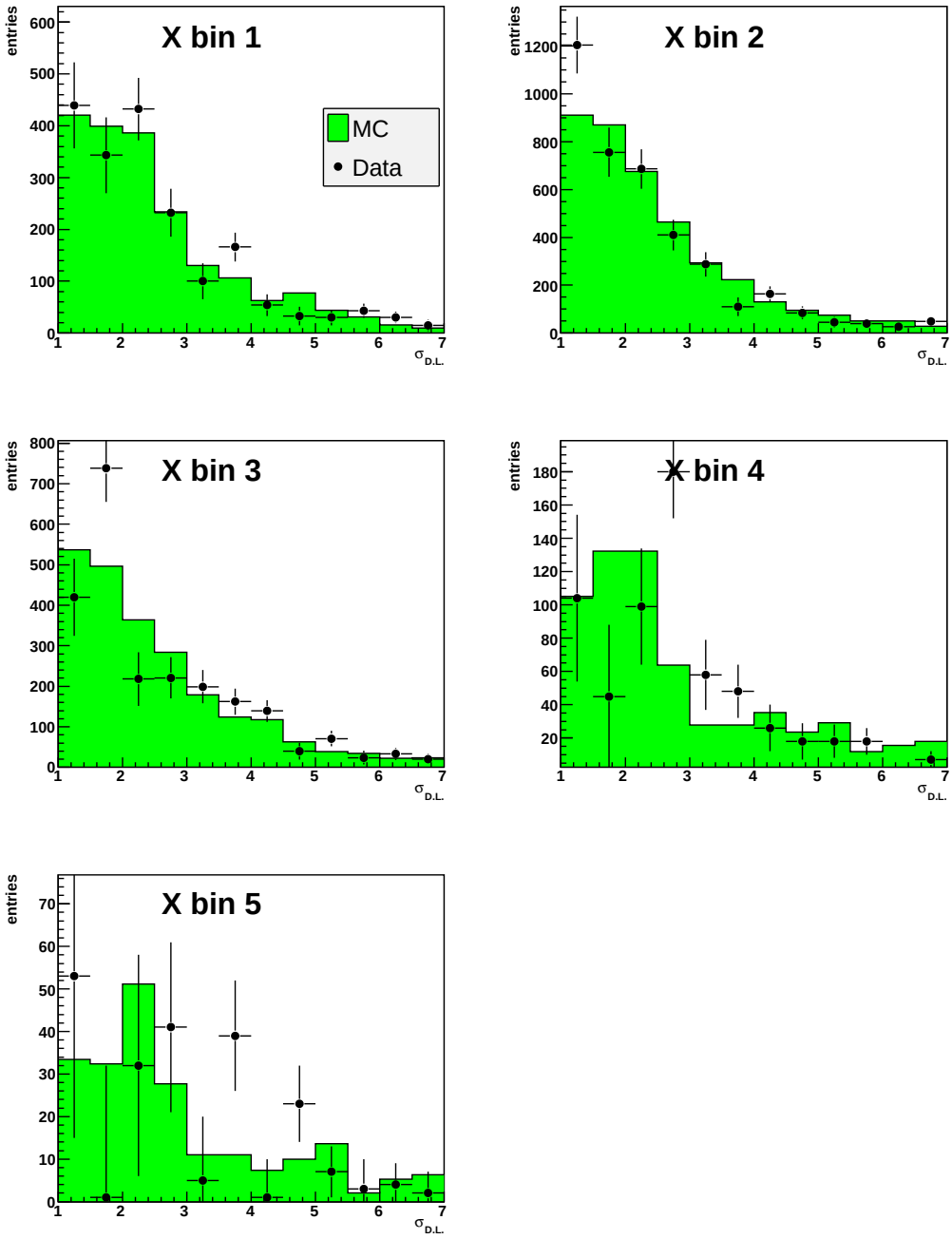


Figure 8.9. Comparison between data and simulation: the number of D^0 candidates as a function of D.L. significance in 5 bins of $x \in [8 \cdot 10^{-5}, 4 \cdot 10^{-4}, 1.5 \cdot 10^{-3}, 5 \cdot 10^{-3}, 10^{-2}, 10^{-1}]$ of the event. The Monte Carlo histograms were area normalized to the data points. Each bin represents a fitted signal value.

8.4 Systematic uncertainties

The approach taken to estimate the systematic errors induced by various selection criteria was described in detail in Sec. 7.2. The same approach was used for the measurement performed in the extended kinematic range. A summary of the nominal cuts and their systematic variations is given in Table 8.2. For convenience,

Variable	Nominal Thres.	Variation
$z_{vtx}(\mathbf{cm})$	30	± 5
$E - P_Z(\mathbf{GeV})$	$40 - 65$	$\rightarrow 45 - 60$
$E(e^-)(\mathbf{GeV})$	10	$\rightarrow 11$
$Y_{el.}$	0.95	$\rightarrow 0.9$
Y_{JB}	0.02	$\rightarrow 0.03$
Box Cut ($\mathbf{cm^2}$)	15×15	$\rightarrow 16 \times 16$
MVD Hits (K)	2	$\rightarrow 3$
MVD Hits (π)	2	$\rightarrow 3$
$P_T(K)(\mathbf{GeV})$	0.7	± 0.1
$P_T(\pi)(\mathbf{GeV})$	0.4	± 0.1
$\lvert\eta(K)\rvert$	1.6	± 0.05
$\lvert\eta(\pi)\rvert$	1.6	± 0.05
$\chi^2_{sec}/ndof$	8	± 1
$\sigma(D.L.)$	1	$\uparrow 2$ and $\downarrow 0$

Table 8.2. List of cuts which generate systematic uncertainties. The nominal thresholds and their variations are shown. For certain variables, variations in only one direction were allowed; the resulting systematic errors in these cases were symmetrized.

the errors on the differential cross-sections are shown for each source separately in Appendix A, in figures A.1, A.2, A.3 and A.4. The most significant sources are summarized below:

- The largest systematic error on the cross-section was induced by the D.L. significance cut. The relative systematic error is within $20 - 25\%$ for most of the bins. The error was computed using a variation of ± 1 unit away from the nominal cut.
- The subtraction of the reflected signal induces a systematic error of about 20%.
- The systematic error induced by changing the background function to a linear function can be as high as $5 - 15\%$ of the differential cross-section for some Q^2 bins.

- Non-negligible errors come also from varying the track momenta and the track number of MVD hits. At low momentum, the effects of multiple scattering increase. These are slightly underestimated in the Monte Carlo.

8.5 Single differential cross-sections

The single differential cross-sections as a function of Q^2, $P_T(D^0)$, $\eta(D^0)$and Bjorken x were measured in the extended kinematic range defined in Sec. 8.1. They are presented below.

8.5.1 Differential cross-section w.r.t. Q^2: $d\sigma/dQ^2$

The differential cross-section as a function of the interaction Q^2 of the event for the process $e^- p \to e^- + D^0 + X$ has been measured for the following bins: $Q^2 \in [5, 10, 20, 40, 80, 200, 1000]\,\mathbf{GeV^2}$. The results are presented in Fig. 8.10 and tabulated in Appendix B, Table B.5. The inner error bars (where visible) denote the statistical errors, the outer error bars represent the systematic errors. The measurement is compared to the second ZEUS analysis and to the NLO theoretical prediction. Details on how the NLO prediction was calculated are given in Sec. 7.4. The relative difference between the measurement and the theoretical prediction, as defined in Eq. 7.6, is shown separately (bottom plot). The error bars correspond to the relative total error (statistic and systematic); the shaded band corresponds the relative error of the theoretical prediction. There is good agreement in shape between the data and the theoretical prediction but the overall normalization is off by $\sim 25\%$. The systematic errors on $d\sigma/dQ^2$ due to the sources listed in Table 8.2 are displayed in Appendix A, Fig. A.1. The total systematic error, shown in Fig. 8.11, top left plot, is calculated by adding all systematic errors in quadrature. Due to low statistics, the highest Q^2 has large statistical errors.

8.5.2 Differential cross-section w.r.t. $P_T(D^0)$: $d\sigma/dP_T$

The differential cross-section as a function of $P_T(D^0)$ for the process $e^- p \to e^- + D^0 + X$ has been measured for the following bins: $P_T \in [1.5, 2.4, 3, 4, 6, 20]\,\mathbf{GeV}$. The results are presented in Fig. 8.12 and tabulated in Appendix B, Table B.6. The measurement is compared to the second ZEUS analysis and to the NLO theoretical prediction. The relative difference between the measurement and the theoretical

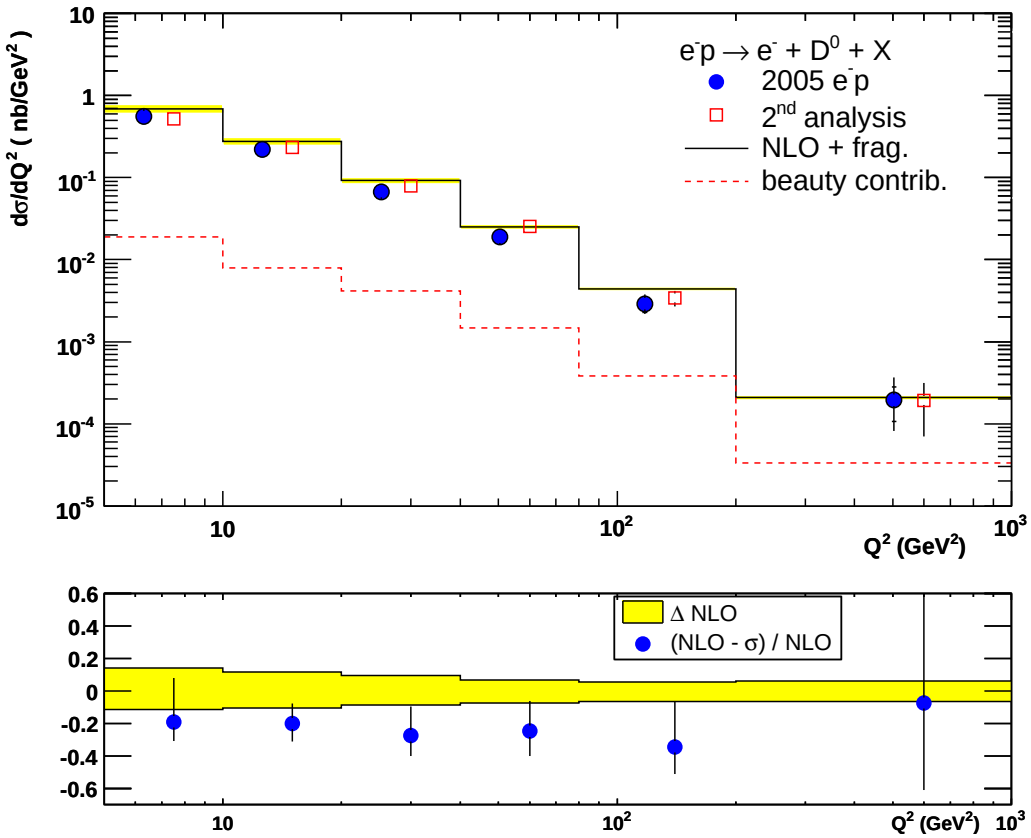


Figure 8.10. The differential cross-section $d\sigma/dQ^2$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical error bars shown) and to the NLO theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. Bottom plot: the relative difference between the measurement and the NLO prediction is shown, together with its total error. The shaded band represents the error on the prediction.

prediction is shown separately (bottom plot). The error bars correspond to the relative total error (statistic and systematic); the shaded band corresponds the relative error of the NLO prediction. There is again good agreement in shape between the data and the theoretical prediction but the overall normalization is off by $15 - 30\%$. The systematic errors on $d\sigma/dP_T(D^0)$ due to the sources listed in Table 8.2 are displayed in Appendix A, Fig. A.2. A large contribution for all P_T bins comes from varying the decay length significance cut and , at low P_T, from varying the calorimeter box cut and the transverse momenta of the daughter tracks. The total systematic errors, shown in Fig. 8.11, are within 15% of the measured

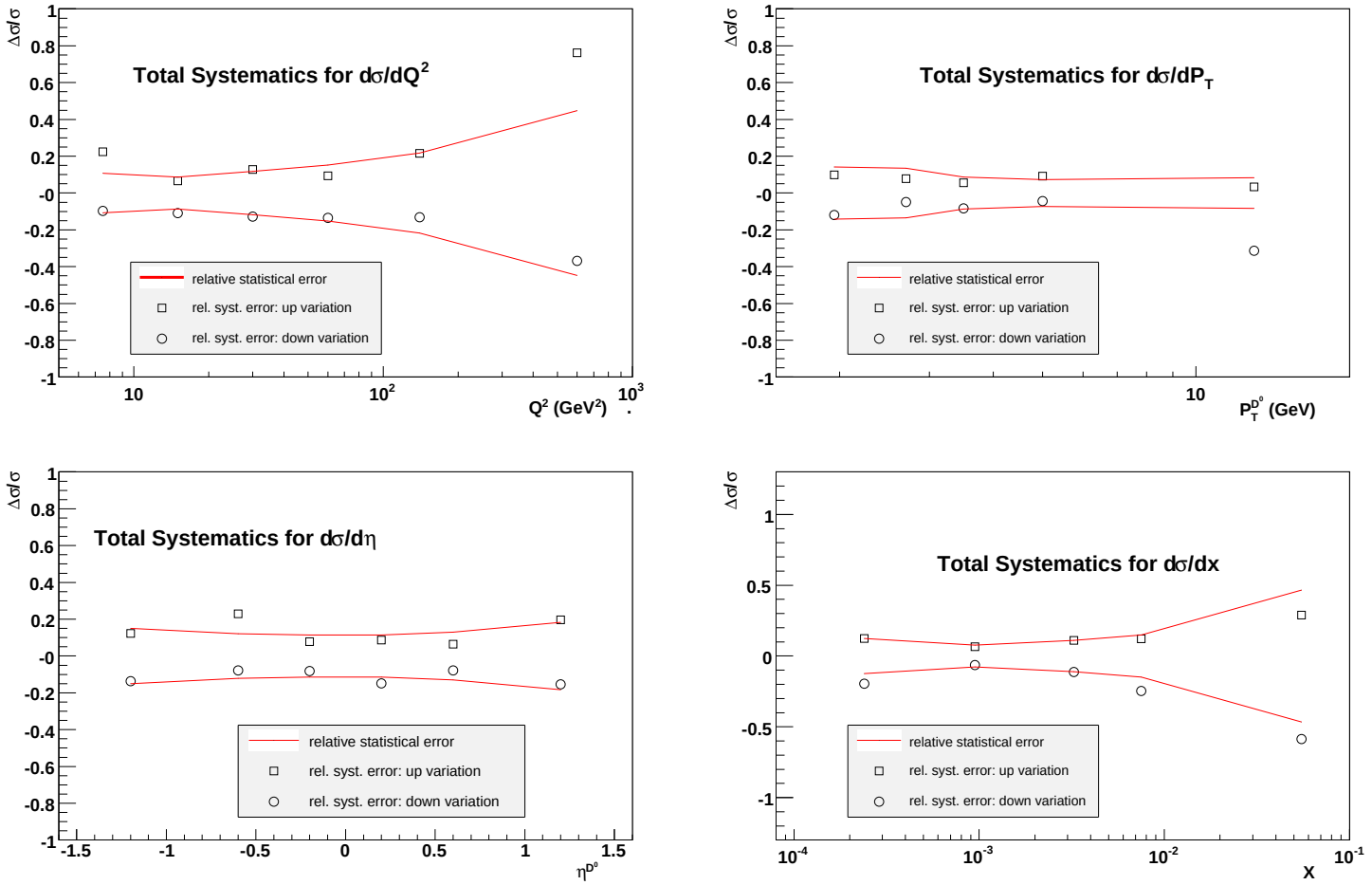


Figure 8.11. The total systematic errors as a function of Q^2 (top left), $P_T(D^0)$ (top right), $\eta(D^0)$ (bottom left) and Bjorken x (bottom right) are computed by adding in quadrature all systematic errors induced by each source. The total systematic errors are of the same size with the statistical errors. Low statistics bins, such as the high Q^2 or high x bins display larger stat. and syst. errors.

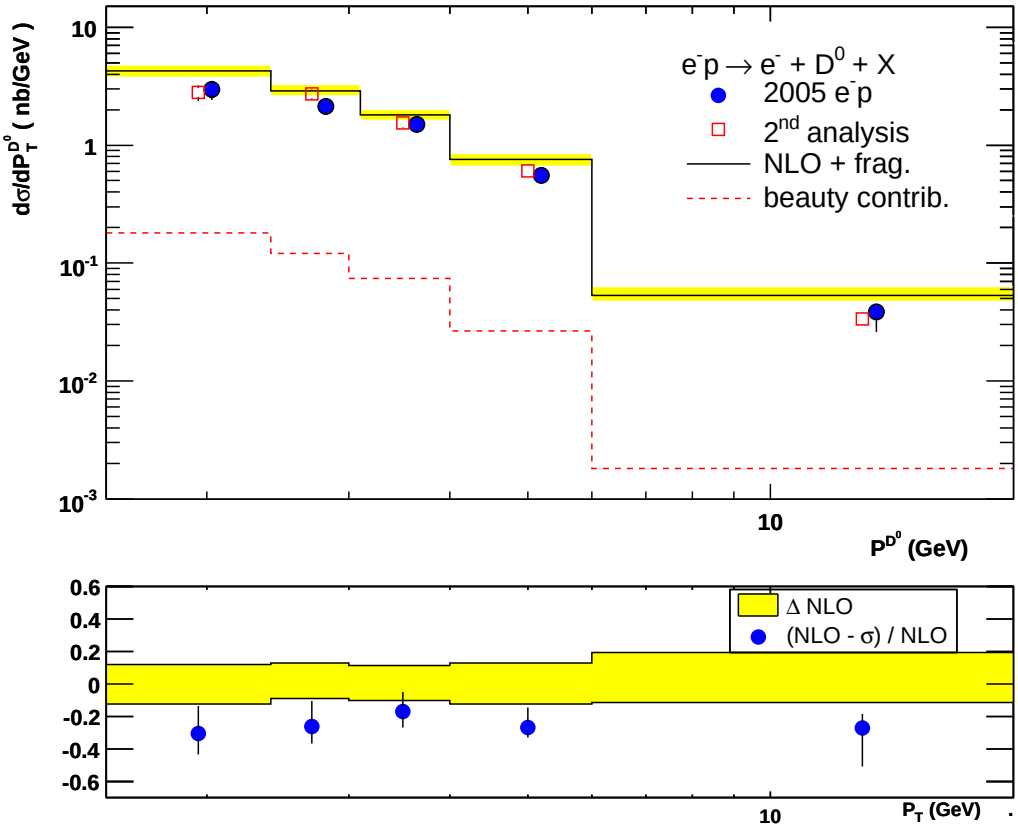


Figure 8.12. The differential cross-section $d\sigma/dP_T(D^0)$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical error bars shown) and to the NLO theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. Bottom plot: the relative difference between the measurement and the NLO prediction is shown, together with its total error. The shaded band represents the error on the prediction.

cross-section, except the highest P_T bin where they reach 30%.

8.5.3 Differential cross-section w.r.t. $\eta(D^0)$: $d\sigma/d\eta$

The differential cross-section as a function of $\eta(D^0)$ for the process $e^-p \to e^- + D^0 + X$ has been measured for the following bins: $\eta \in [-1.6, -0.8, -0.4, 0, 0.4, 0.8, 1.6]$. The results are presented in Fig. 8.13 and tabulated in Appendix B, Table B.7. The measurement is compared to the second ZEUS analysis and to the NLO theoretical prediction. The relative difference between the measurement and the theoretical prediction is shown underneath. There is reasonable agreement in shape between

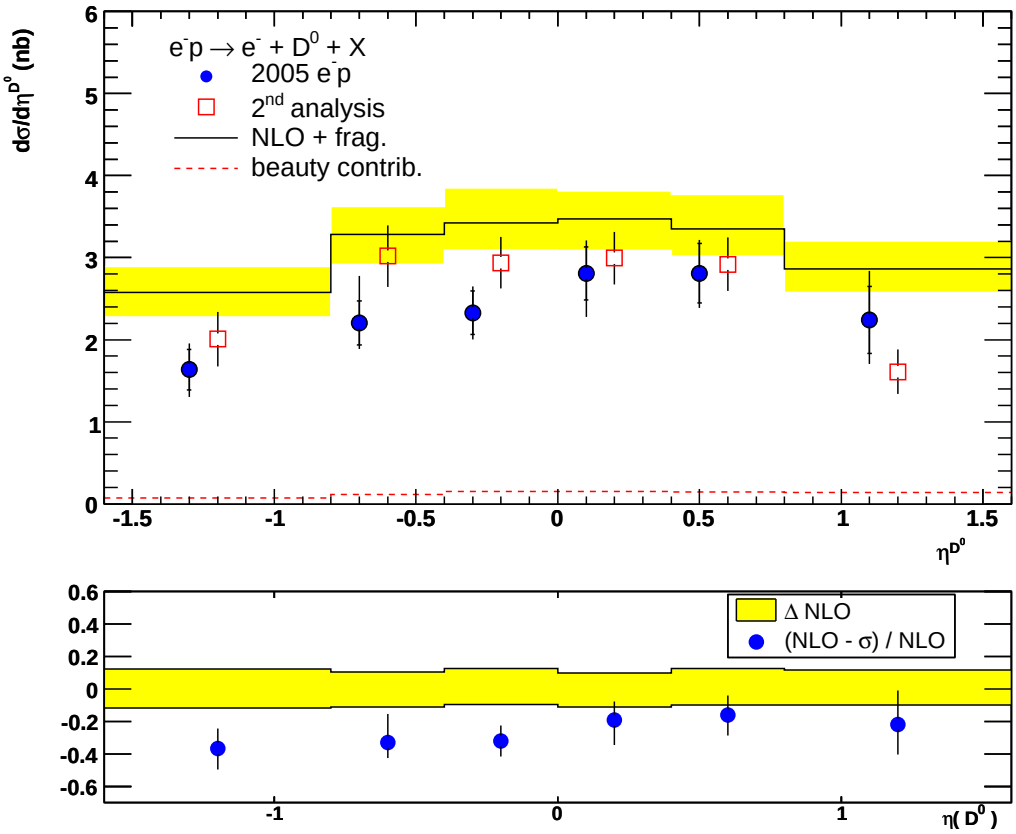


Figure 8.13. The differential cross-section $d\sigma/d\eta(D^0)$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical error bars shown) and to the NLO theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. Bottom plot: the relative difference between the measurement and the NLO prediction is shown, together with its total error. The shaded band represents the error on the prediction.

the data and the theoretical prediction. The theory overestimates the cross-section by 20-30% and there in an indication that the data are concentrated more at larger values of η: the lowest η shows the largest disagreement of almost 40%. All the systematic errors on $d\sigma/dP_T(D^0)$ are displayed in Appendix A, Fig. A.3, for each source. The total systematic errors, shown in Fig. 8.11, are within $20-25\%$ of the measured cross-section and of the same size as the statistical errors. Large systematic errors for all η bins come from varying the decay length significance cut and from turning off the reflected signal subtraction. Also, errors induced by the background fit function and the cuts on the daughter track momenta are significant.

8.5.4 Differential cross-section w.r.t. Bjorken x: $d\sigma/dx$

The differential cross-section as a function of Bjorken x for the process $e^-p \to e^- + D^0 + X$ has been measured for the following bins: $x \in [8 \cdot 10^{-5}, 4 \cdot 10^{-4}, 1.5 \cdot 10^{-3}, 5 \cdot 10^{-3}, 10^{-2}, 10^{-1}]$. The results are presented in Fig. 8.14 and tabulated in Appendix B, Table B.8. The measurement is compared to the second ZEUS analysis

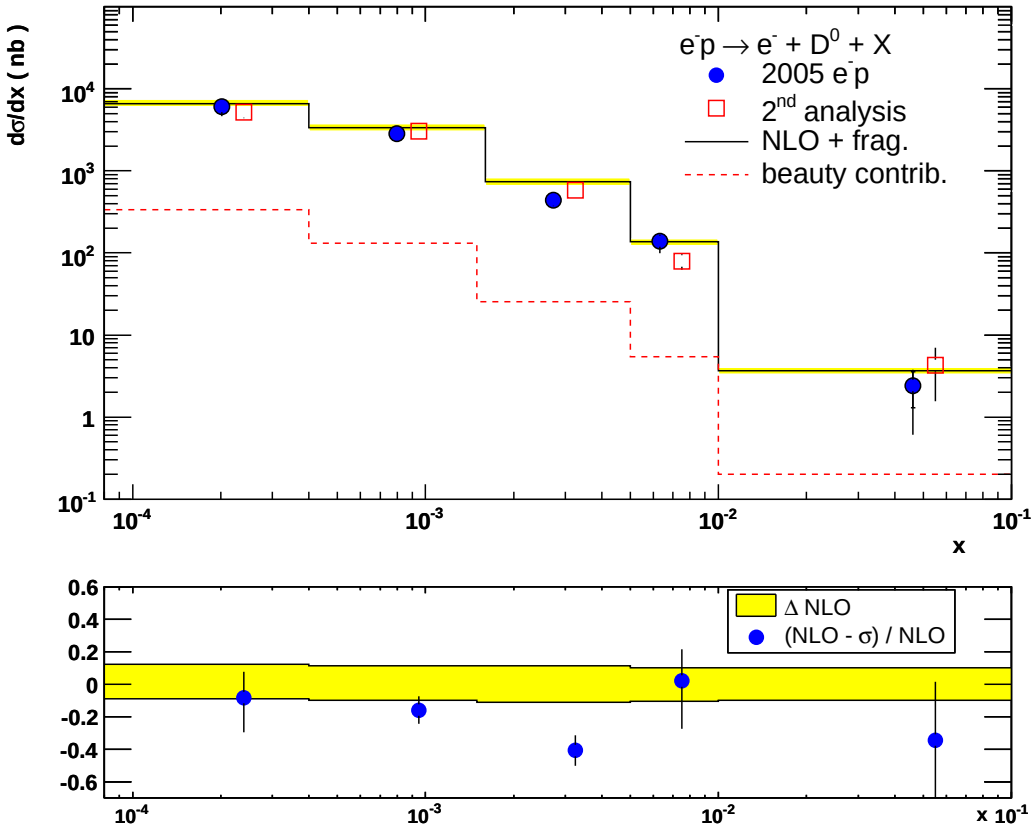


Figure 8.14. The differential cross-section $d\sigma/dx$ for the process $e^-p \to e^- + D^0 + X$ was measured on a 2005 e^-p data sample of $127\,pb^{-1}$(top plot). The inner error bars (where visible) represent statistical errors, the outer bars denote the systematic errors. The results are compared to the second independent ZEUS measurement (only statistical error bars shown) and to the NLO theoretical prediction. The beauty cross-section in the simulated files is also displayed, for reference. Bottom plot: the relative difference between the measurement and the NLO prediction is shown, together with its total error. The shaded band represents the error on the prediction.

and to the NLO theoretical prediction. There is good agreement between the two analyses. The relative difference between the measurement and the theoretical prediction is shown separately (bottom plot). The shape is reasonably described although there is a tendency for the data to drop relative to the prediction towards

higher x values. The statistics precludes any further conclusion. All the systematic errors on $d\sigma/dx$ are displayed in Appendix A, Fig. A.4, for each source. The total systematic errors, shown in Fig. 8.11, are within $20-25\%$ of the measured cross-section and of the same size as the statistical errors, except for the highest x bin where both statistical and systematic errors reach 50%. The D.L. significance, the background fit function and the reflection subtraction induce the largest systematic errors.

8.6 Double differential cross-section

The double differential cross-section as a function of Q^2 and y , $\sigma(Q^2_{low}, Q^2_{high}, y_{low}, y_{high})$, for the process $e^- p \to e^- + D^0 + X$, has been measured *in the bin*[2] for the following bins: 3 Q^2bins , $Q^2 \in [5, 9, 44, 1000]\,\mathbf{GeV}^2$; each Q^2 bin was further divided in 3 y bins: $y \in [0.02, 0.12, 0.3, 0.7]$. The kinematic range for this measurement is defined in Sec. 8.1. The results are presented in Fig. 8.15 and tabulated in Appendix B, Table B.9.

The systematic errors associated with the double differential cross-section suffered significantly due to the intrinsic statistical errors on the signal fits. Because of the lower statistics in each double differential bin, especially at low and high $Q^2$and y, the effect was much more pronounced than in the case of single differential cross-sections. This issue has been dealt with by subtracting an estimated statistical error from the systematic error, for each source separately. A value of the statistical error estimate for each particular cut variation was computed in the simulated files, by looking at the efficiency of the cut variation on the signal:

$$\epsilon_i = \frac{N_i^{MC}}{N_{nom}^{MC}} \tag{8.1}$$

where N_i^{MC} represents the number of meson candidates reconstructed in the simulation after varying cut i, and N_{nom}^{MC} is the number of meson candidates reconstructed at nominal cuts. The estimated statistical error for source i that feeds into the induced systematic error is then given by:

$$\Delta_i^{feed} = \sqrt{\epsilon_i(1-\epsilon_i)} \cdot \Delta^{stat} \tag{8.2}$$

where Δ^{stat} was the statistical error on the reconstructed signal in data, at the

[2]This means that the cross-section values were not divided by the double bin widths.

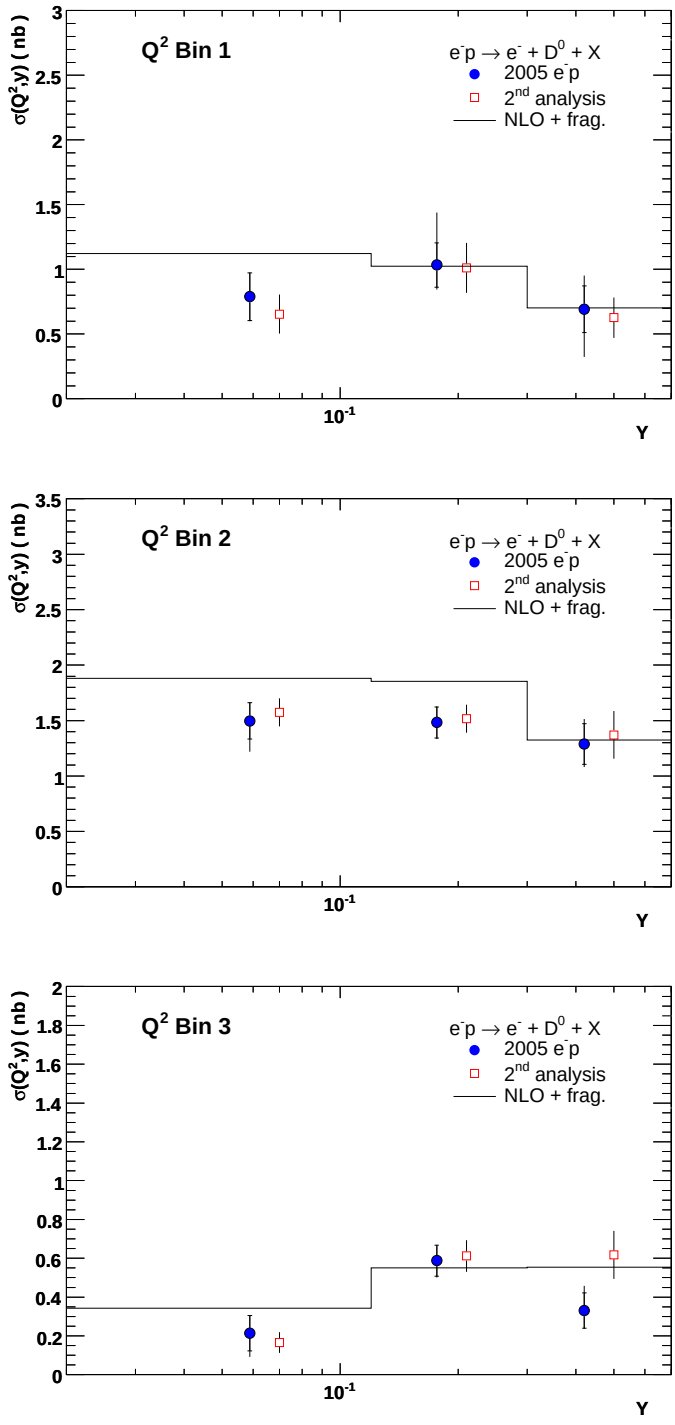


Figure 8.15. The double differential cross-section $\sigma(Q^2_{low}, Q^2_{high}, y_{low}, y_{high})$ is shown for 3 different Q^2 bins, $Q^2 \in [5, 9, 44, 1000]\,\mathbf{GeV}^2$ in the top, central and bottom figures. The cross-section is given in the bin (the values have not been divided by the bin width). The measurement is compared to the NLO theoretical prediction and the second analysis (only statistical errors shown). Larger errors are associated to bins in which the statistics was low.

nominal cuts. If the polluted (original) systematic error is Δ^{sys}, the corrected systematic error becomes:

$$\Delta_i^{sys.cor} = \begin{cases} \sqrt{(\Delta_i^{sys})^2 - (\Delta_i^{feed})^2} \ \ if \ \ \Delta_i^{sys} > \Delta_i^{feed} \\ 0 \ \ if \ \ \Delta_i^{sys} < \Delta_i^{feed} \end{cases}$$

Finally, the original sign of Δ^{sys} is passed to the $\Delta_i^{sys.cor}$. It is worth noting that this subtraction has little effect when statistics are high.

The systematic errors for each source listed in Table 8.2 are given in Appendix A, figures A.5, A.6 and A.7 corresponding to first, second and third Q^2 respectively. Once again, the largest contribution comes from varying the D.L. significance. The reflection subtraction induces also a large systematic error, especially In the highest y bins, for all 3 Q^2 bins, reaching the level of 30% in the first Q^2, first y bin. The systematic errors in most of the bins are within the statistical error. The total systematic errors corresponding to each of the three Q^2 bins are displayed in Fig. 8.16. The lowest total error of about $15-20\%$ is found for the central Q^2 bin, in all 3 y bins, as a larger fraction of the signal is concentrated here.

8.7 Measuring $F_2^{c\bar{c}}$

It was shown in Fig. 8.10, 8.12, 8.13, 8.14 and 8.15 that the NLO prediction reproduces the shape of the measured cross-section distributions well in all measured variables. Assuming this holds outside the measured kinematic range, we can use the NLO calculation to extrapolate to the unmeasured kinematic area. The NLO calculation is also used to perform the bin centering corrections, as explained below. From the measured visible cross-section in a bin in Q^2 and y, the charm structure function $F_2^{c\bar{c}}$ can be extracted at any value of x and Q^2 that falls within the kinematic limits corresponding with the y and Q^2 bins through the equation:

$$F_{2,meas}^{c\bar{c}}(x,Q^2) = \frac{\sigma_{meas}^{bin}(y_{low}, y_{high}, Q^2_{low}, Q^2_{high})}{\sigma_{th}^{bin}(y_{low}, y_{high}, Q^2_{low}, Q^2_{high})} F_{2,th}^{c\bar{c}}(x,Q^2) \tag{8.3}$$

where $F_{2,th}^{c\bar{c}}$ is the NLO prediction for the structure function, σ_{th} is the NLO theoretical prediction for the cross-section and σ_{meas} is the measured cross-section. To depend least on any mismatch of the cross-section shape within the bin, the bin center of gravity is chosen for the Q^2 value at which $F_2^{c\bar{c}}$ is quoted. Also, the y bin edges were translated to x bin edges through the relation $Q^2 = sxy$ and a central x

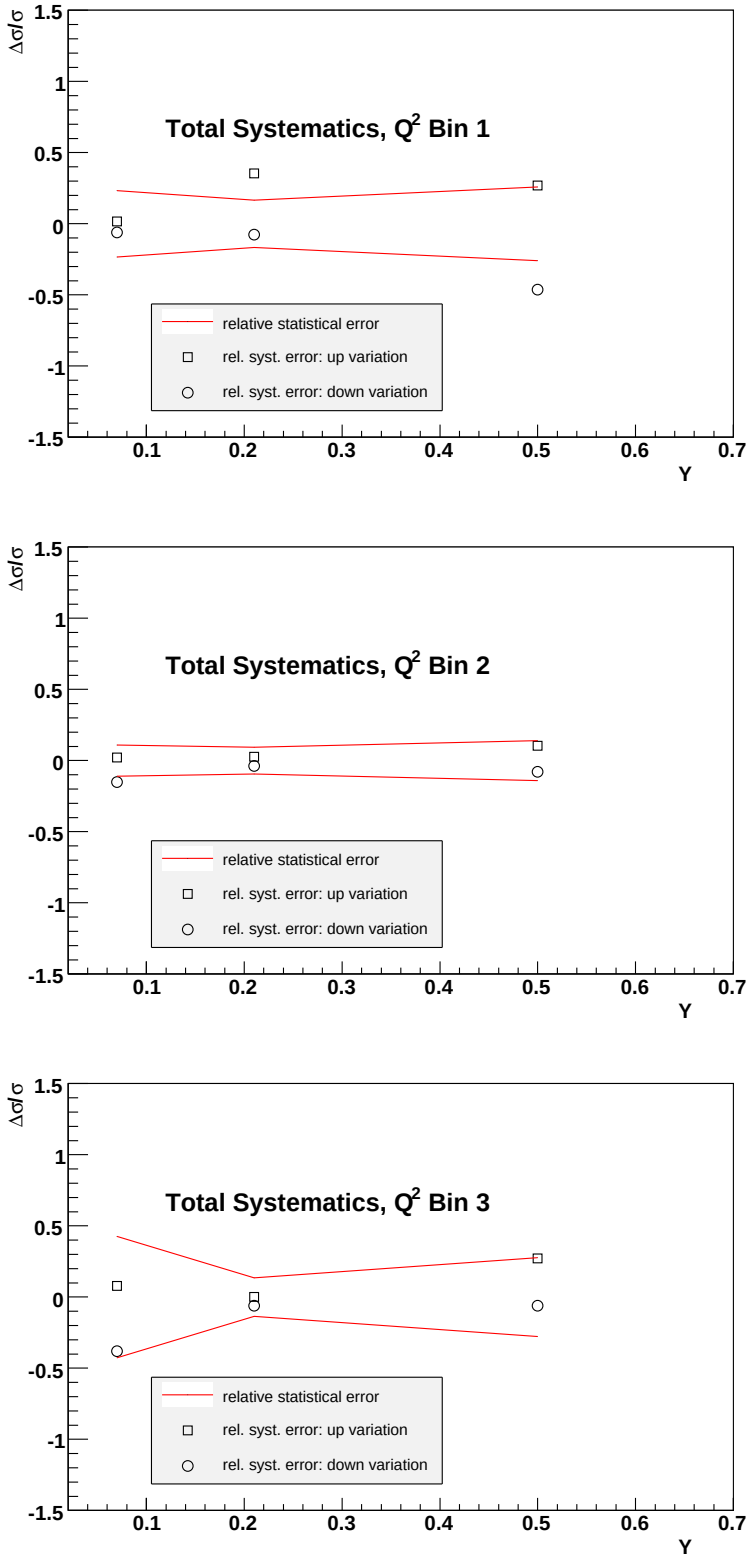


Figure 8.16. The total systematic errors on the double differential cross-section $d\sigma/dQ^2dy$ are shown for 3 Q^2 bins: $Q^2 \in [5, 9, 44, 1000]\,\mathbf{GeV}^2$. The total systematic errors were computed by adding individual errors in quadrature.

value was chosen at which to quote the structure function.

The measured $F_2^{c\bar{c}}$ is plotted in Fig. 8.17 and tabulated in Table B.10 of Appendix B. The results are compared to NLO theoretical prediction, which is independent of the kinematic range of the measurement as the extrapolation procedure was performed to the full phase-space. Also, the results of the second independent analysis are shown. The measured shape is well reproduced by the NLO prediction, except for the lowest x bin at high Q^2 where the measured value is lower but within 2σ of the NLO central value.

The ratio between the measured $F_2^{c\bar{c}}$ presented above and the ZEUS published F_2^{em}[51], at the same Q^2 and Bjorken x values, has been plotted in Fig. 8.18. Its shape is compared to the NLO/F_2^{em} and it is found to be well reproduced. There is an increase in charm contribution to the total F_2 at higher Q^2. The ratio decreases with increasing x, proving that charm production is dominated by gluon distribution at low x. Regarding the overall normalization, the NLO predicts a higher ratio of $F_2^{c\bar{c}}/F_2^{em}$ by about 20% w.r.t. the measured value. If this deficit is actually due to a gluon density mismatch, then any cross-section dependent on gluon-gluon fusion at the Large Hadron Collider will be less than the expected value by 40%.

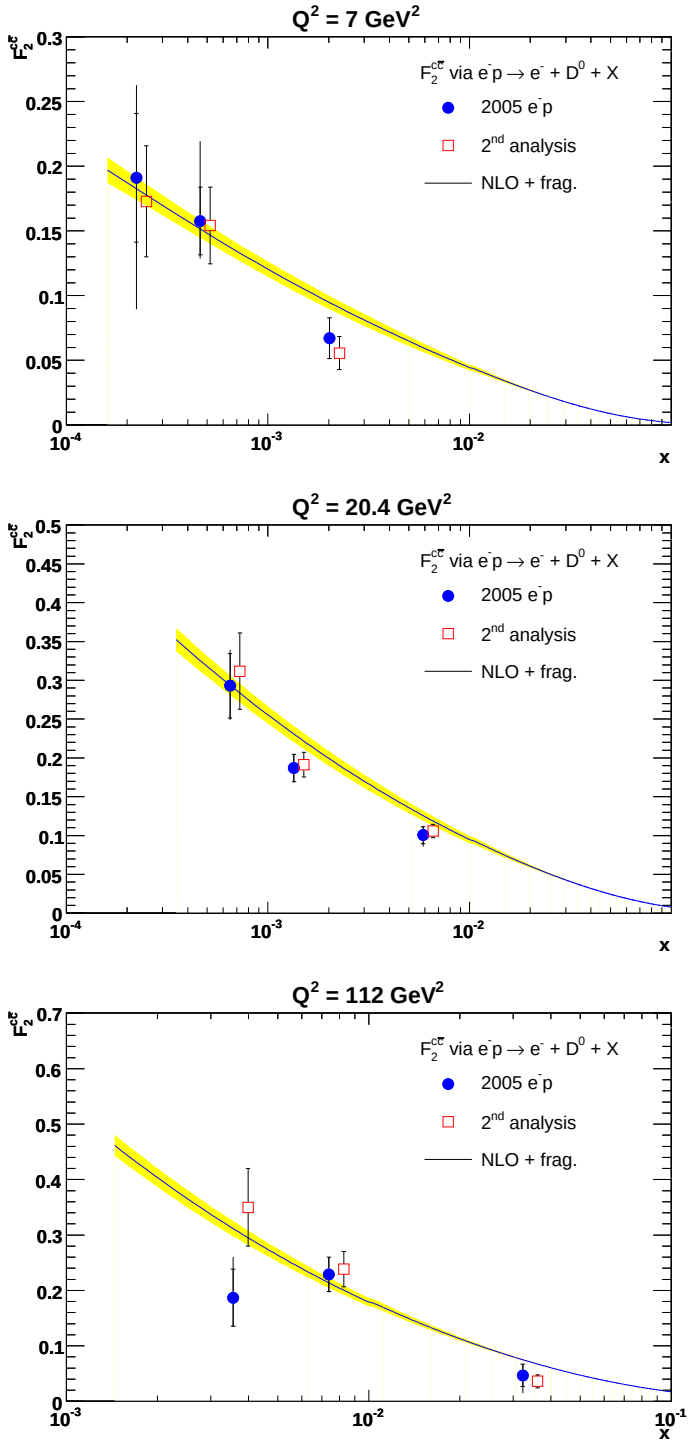


Figure 8.17. The measured $F_2^{c\bar{c}}$ as a function of Bjorken x for three fixed values of Q^2 is presented. The results are compared to the second independent analysis and to the theoretical NLO prediction.

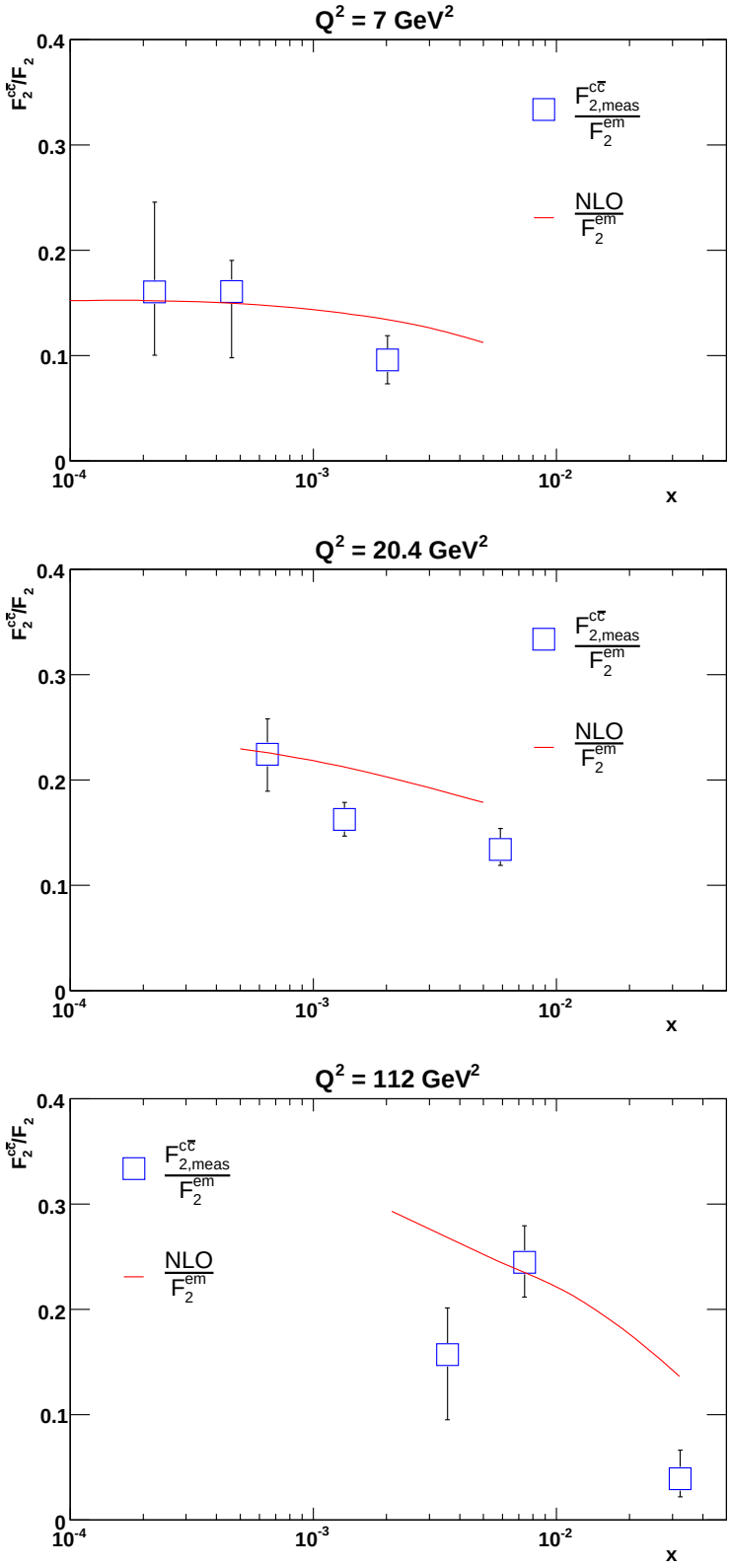


Figure 8.18. The ratio $F_2^{c\bar{c}}/F_2^{em}$ is compared to the ratio NLO/F_2^{em} at 3 different Q^2 values. Charm contributes to as much as $\sim 25\%$ to the total F_2 at high Q^2. The shape is well reproduced by the NLO.

Chapter 9

Conclusions

The aim of the analysis presented in this thesis is the measurement of the production of charm quarks in deep inelastic scattering. The measurement was performed on data recorded by the ZEUS collaboration in 2005, having a total luminosity of 127.35 pb^{-1}. The present analysis relies both on the performance of the ZEUS microvertex detector as well as on analysis-dependent vertexing techniques.

More than 95% of the HERA II data recorded in 2005 has MVD information. The vertex detector improves greatly the quality of the reconstructed tracks and vertices. A key issue for a good MVD performance is the MVD alignment. Great progress has been made by using tracks from e^-p collisions to align all ladders of the MVD with equal accuracy. Previously, only alignment based on cosmic tracks was used. The new alignment procedure based on e^-p tracks improves hit resolution ($\sim 25 - 30\,\mu m$ for all ladders now) and impact parameter w.r.t. the interaction point ($\sim 100 - 150\,\mu m$). This analysis takes advantage of these improvements in the alignment.

Charm was reconstructed using the decay channel $D^0 \to K + \pi$ which accounts for 3.8% of charm D^0 meson decays. As more than 50% of charm hadronizes to D^0 mesons either directly or via D^* decays (a large fraction of D^* decay to D^0), the decay channel $D^0 \to K + \pi$ is a prolific channel for charm signal. There are disadvantages, however, such as the short D^0 lifetime and the fact that the K and the π are not uniquely identified, leading to a reflected signal.

For this analysis, specific vertexing was performed, by vertexing separately combinations of two tracks of opposite charge. Good vertices were selected by the vertex χ^2. The D^0 invariant mass at the vertex was reconstructed. The best reduced primary vertex was created using all other tracks in the event. In this way, the charm

meson decay was tagged. Information about the decay length and decay length significance, the momentum vector at the decay vertex and the charm meson lifetime were optimized for producing the smallest relative error on the signal and the highest purity. A neutral pseudotrack was built using knowledge of the decay vertex, the K and π tracks and their errors. The pseudotrack was then refitted to the primary vertex, producing a new primary vertex position and a χ^2 increase. An attempt was made to further improve the relative error on the signal by selecting events using the primary vertex χ^2 increase cut. This provided no improvement essentially because the D^0 meson is a difficult particle to tag. It is the shortest lived particle whose decay can be tagged by the ZEUS detector. Its decay length is comparable to the flightpath resolution which makes it difficult to distinguish the genuine decays from background combinations, which also decrease in number with increasing decay lengths. In total, there were 7440 ± 233 D^0 candidates found for the kinematic range $5 < Q^2 < 1000\,\mathbf{GeV^2}$, $3 < P_T(D^0) < 20\,\mathbf{GeV}$. Among them, 1668 ± 46 candidates were additionally tagged as coming from a D^{*+} decay, $D^{*+} \rightarrow D^0 + \pi_s$. The rest were either directly produced in the hadronization process of the charm quark or came from a strong/electromagnetic D^{*0} decay. The improvement in track and vertex resolution, due to the microvertex detector, allowed signal reconstruction of D^0 mesons at transverse momenta less than $3\,\mathbf{GeV}$. An additional number of 5483 ± 473 D^0 candidates were reconstructed with a transverse momentum in the range of $1.5 < P_T(D^0) < 3\,\mathbf{GeV}$. This was the first charm meson signal reconstructed at ZEUS at $P_T(D^0) < 3\,\mathbf{GeV}$ and it increased significantly the statistics.

The extracted signal was used to perform two similar measurements, for the kinematic ranges of $1.5 < P_T(D^0) < 20\,\mathbf{GeV}$ and $3 < P_T(D^0) < 20\,\mathbf{GeV}$ respectively. The higher P_T measurement acted as a check with respect to previous HERA I measurements while the measurement in the expanded kinematic range increased statistics. The differential cross-sections as functions of Q^2 of the event, $\eta(D^0)$, $P_T(D^0)$ and Bjorken x were measured. They were compared to the next to leading order theoretical prediction. For both measurements, the differential cross-section shapes were well reproduced, for all variables, but the predicted overall normalization was higher by a factor of $20 - 30\%$. It was shown that the NLO prediction describes the data well at low transverse momenta of the charm meson. The main input parameters to the theoretical prediction software package HVQDIS are: the ZEUS parton density functions, the charm mass $M_c = 1.35\,\mathbf{GeV}$, the Peterson frag-

mentation parameter $\epsilon = 0.035$ and the renormalization and factorization scales $\mu_R = \mu_F = \sqrt{Q^2 + 4M_c^2}$. Also, the double differential cross-section $d^2\sigma/dQ^2dy$ was measured. This cross-section was used to extract the charm structure function $F_2^{c\bar{c}}$. A good agreement with the theoretical prediction of $F_2^{c\bar{c}}$ was found. When measuring $F_2^{c\bar{c}}$, the increased statistics allowed finer granularity and reduced the extrapolation by a factor of ~ 3.

The vertexing techniques used in this analysis can be applied with great success to charm or beauty meson tagging. Selecting charm candidates by their decay length significance and/or by the χ^2 increase of the charm pseudotrack on the primary vertex have already proved to be great tools for enhancing the signal purity and reducing the signal relative error in the tagging of D^+ mesons at ZEUS. Also, the accuracy of the measurement presented here will be increased by including the entire HERA II data set with a luminosity of $\approx 400\, pb^{-1}$.

This measurement confirms that theoretical predictions performed with gluon distributions extracted from inclusive DIS are certainly trustworthy within $\sim 25\%$.

Appendix A

Systematic uncertainties

In this appendix, the relative systematic errors on the differential cross-sections are shown for each source separately, for the following cross-sections:

- single differential cross-sections as a function of Q^2, $P_T(D^0)$, $\eta(D^0)$ and Bjorken x, displayed in figures A.1, A.2, A.3 and A.4 respectively.
- double differential cross-section $\sigma(Q^2_{low}, Q^2_{high}, y_{low}, y_{high})$ in figures: A.5, A.6 and A.7.

The measurement was performed in the kinematic range:

- $5\,\mathbf{GeV^2} < Q^2 < 1000\,\mathbf{GeV^2}$
- $0.02 < y < 0.7$
- $1.5 < P_T(D^0) < 20\,\mathbf{GeV}$

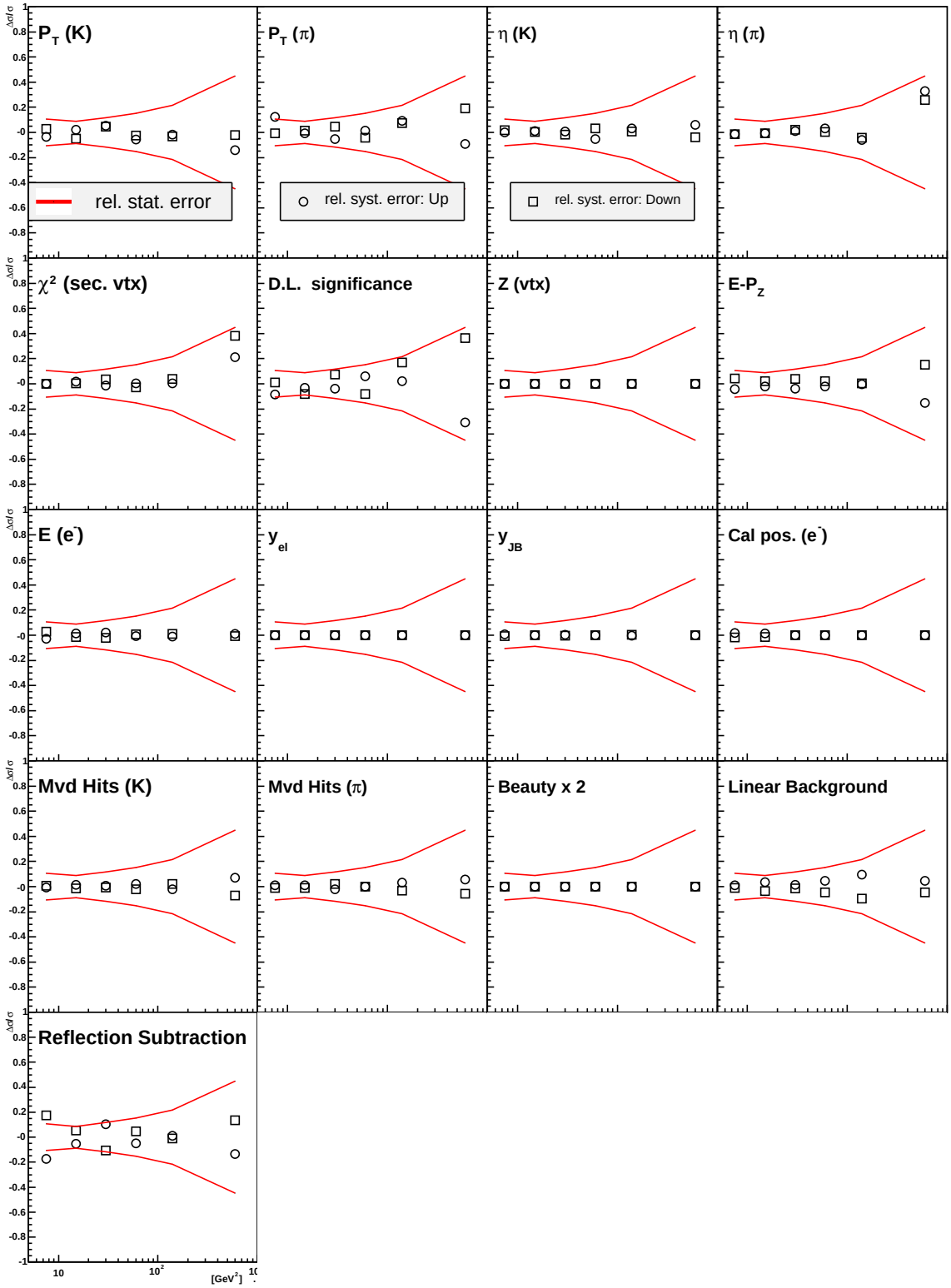


Figure A.1. The systematic errors, calculated and labelled for each source separately, are displayed as a function of Q^2. A large contribution for all Q^2 bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction. The highest Q^2 bin suffers from low statistics and therefore large errors.

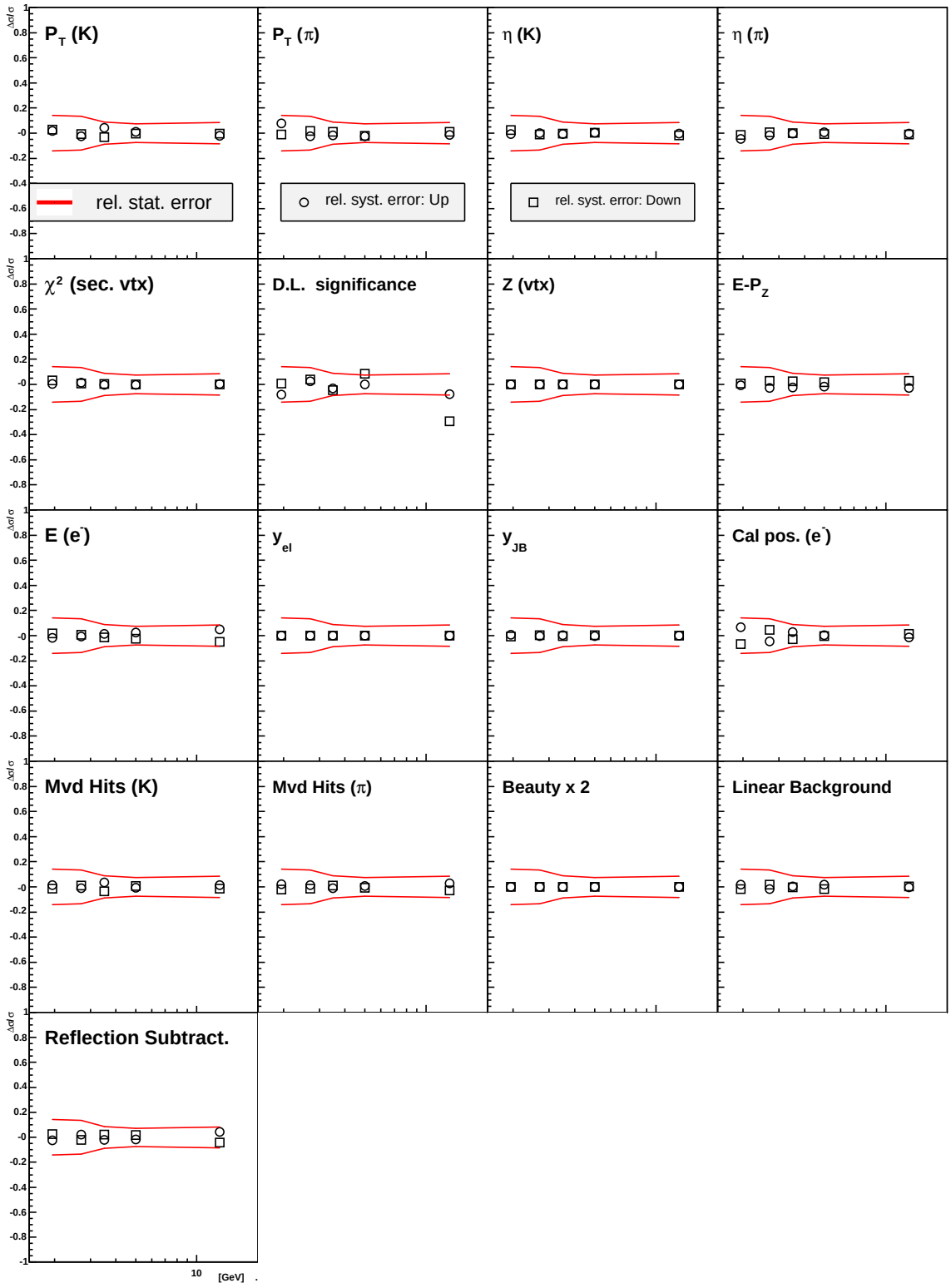


Figure A.2. The systematic errors, calculated and labelled for each source separately, are displayed as a function of $P_T(D^0)$. A large contribution for all P_T bins comes from varying the decay length significance cut and , at low P_T, from varying the calorimeter box cut and the transverse momenta of the daughter tracks.

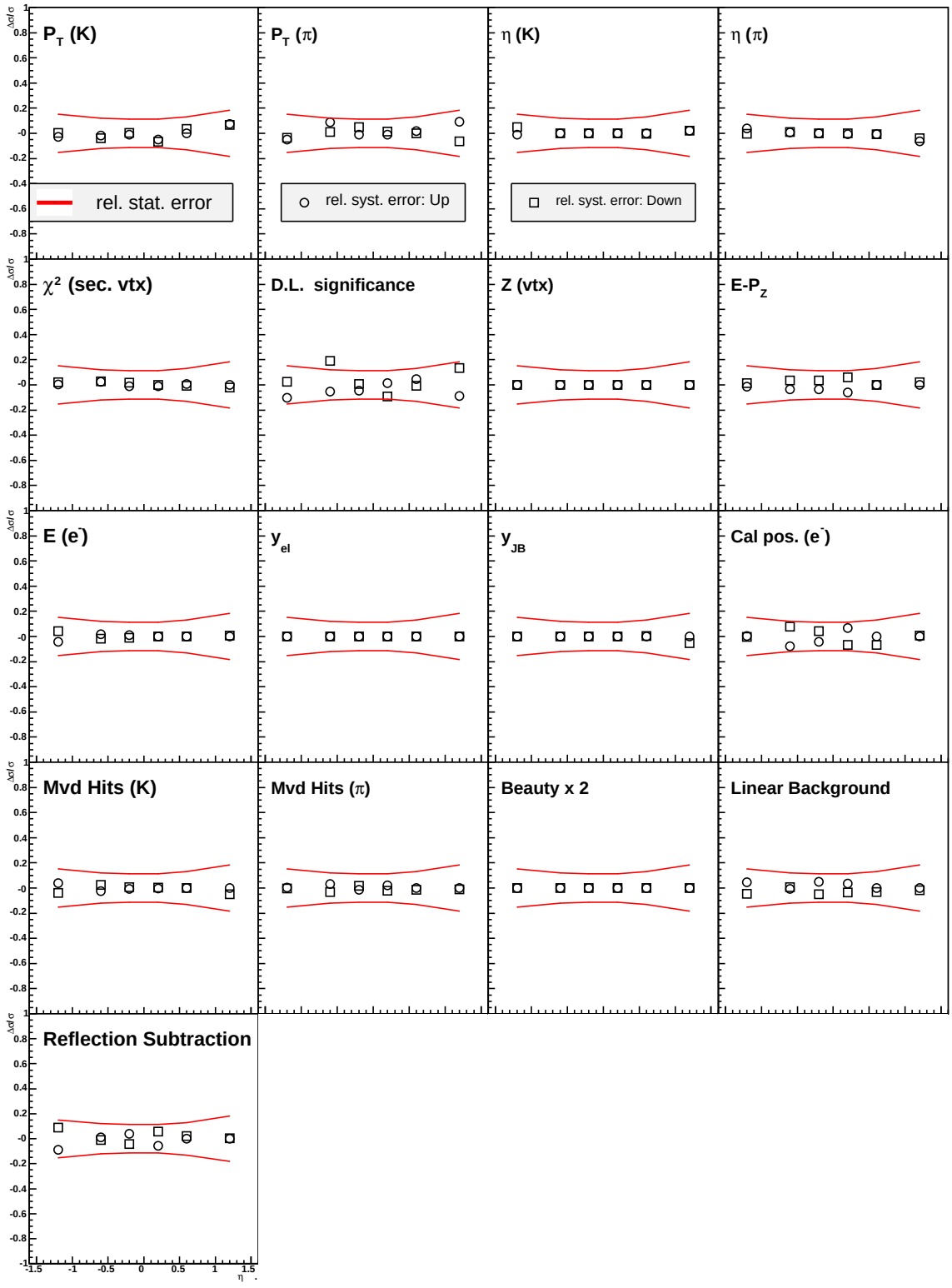


Figure A.3. The systematic errors, calculated and labelled for each source separately, are displayed as a function of $\eta(D^0)$. A large contribution for all η bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction. Also, errors induced by the background fit function and the cuts on the daughter track momenta are significant.

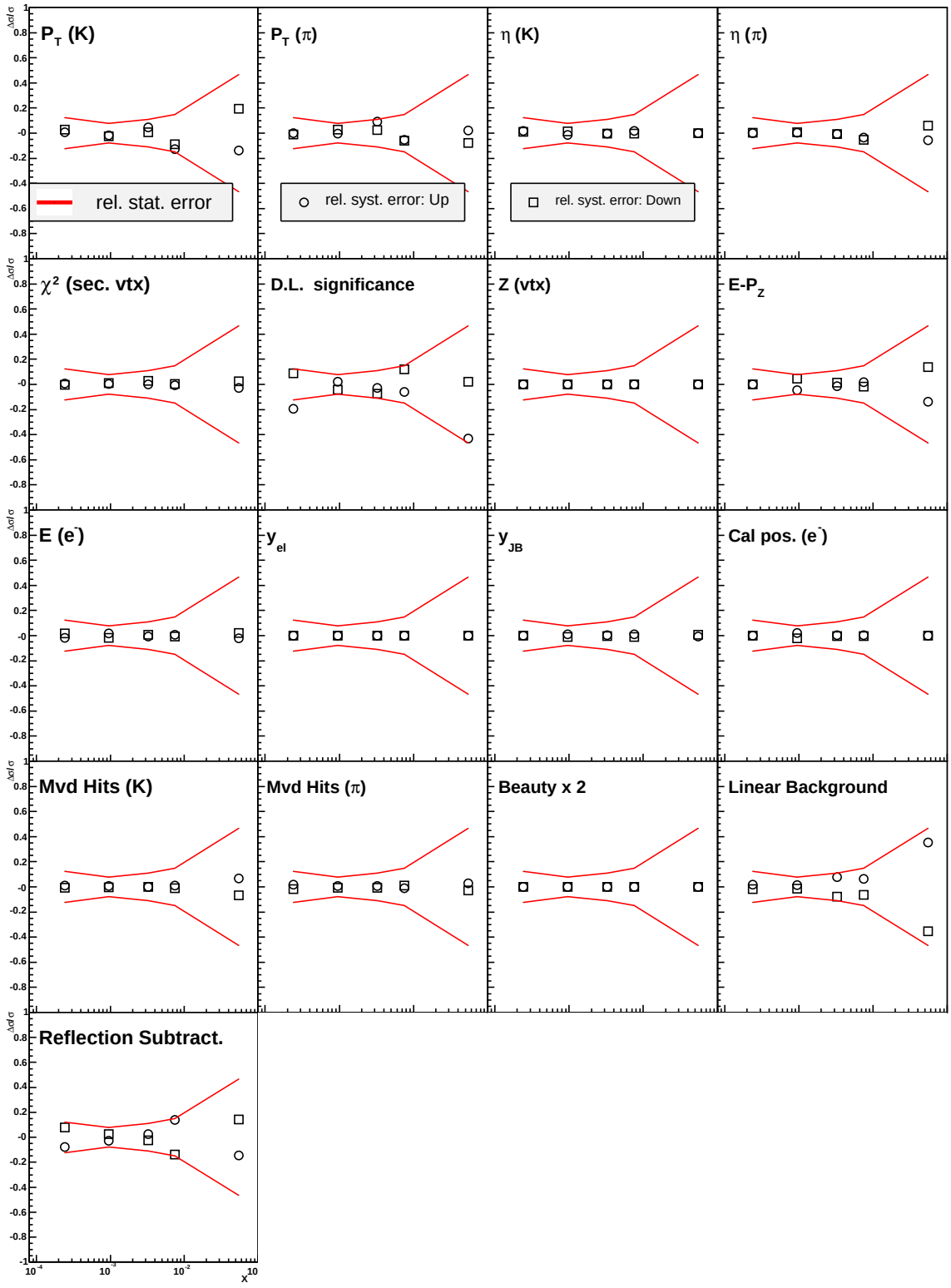


Figure A.4. The systematic errors, calculated and labelled for each source separately, are displayed as a function of Bjorken x. A large contribution for all x bins comes from varying the decay length significance cut and from turning off the reflected signal subtraction.

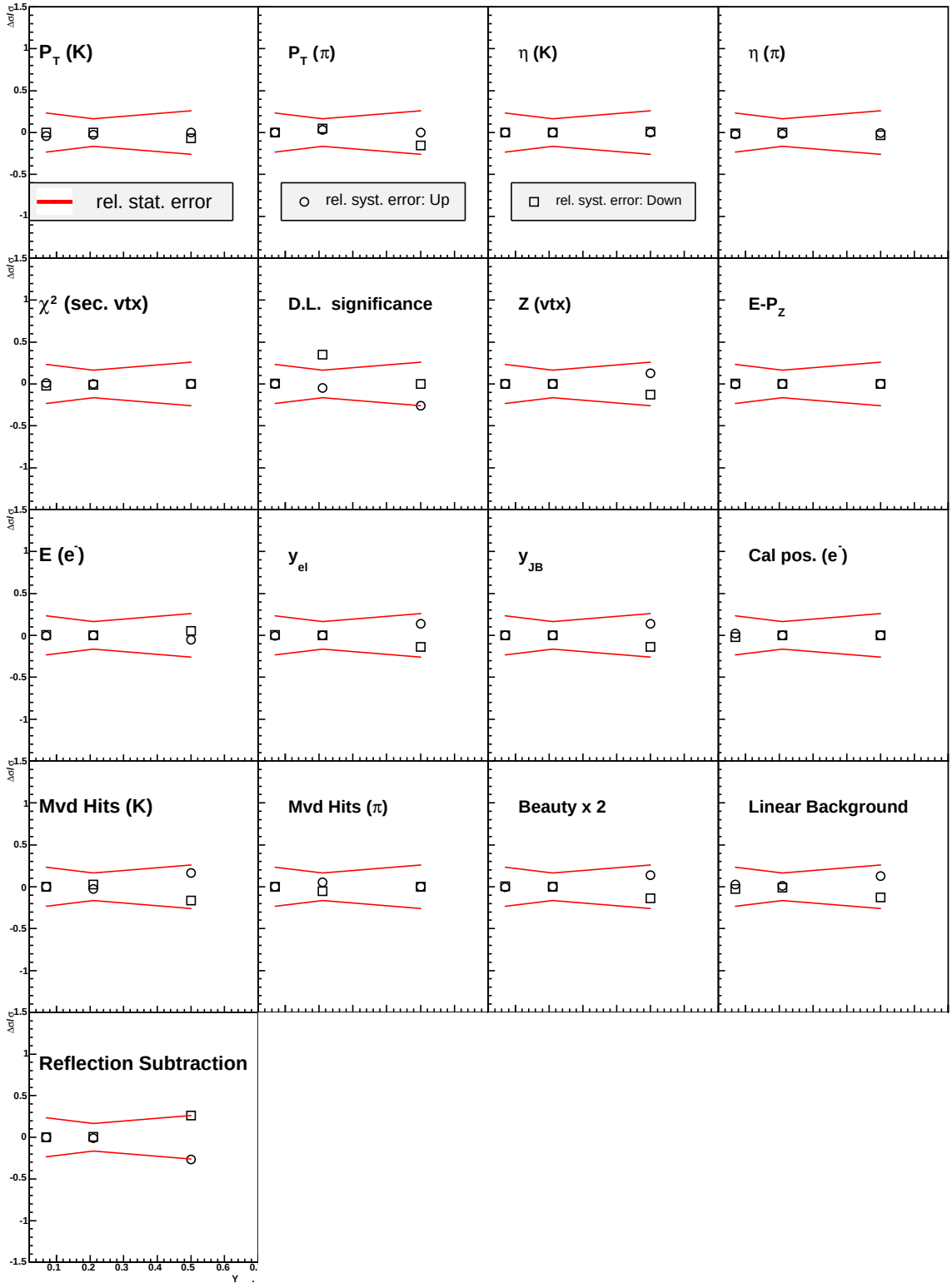


Figure A.5. The systematic errors on the double differential cross-section $d\sigma/dQ^2dy$, for $5 < Q^2 < 9\,\mathbf{GeV^2}$, are shown for each source.

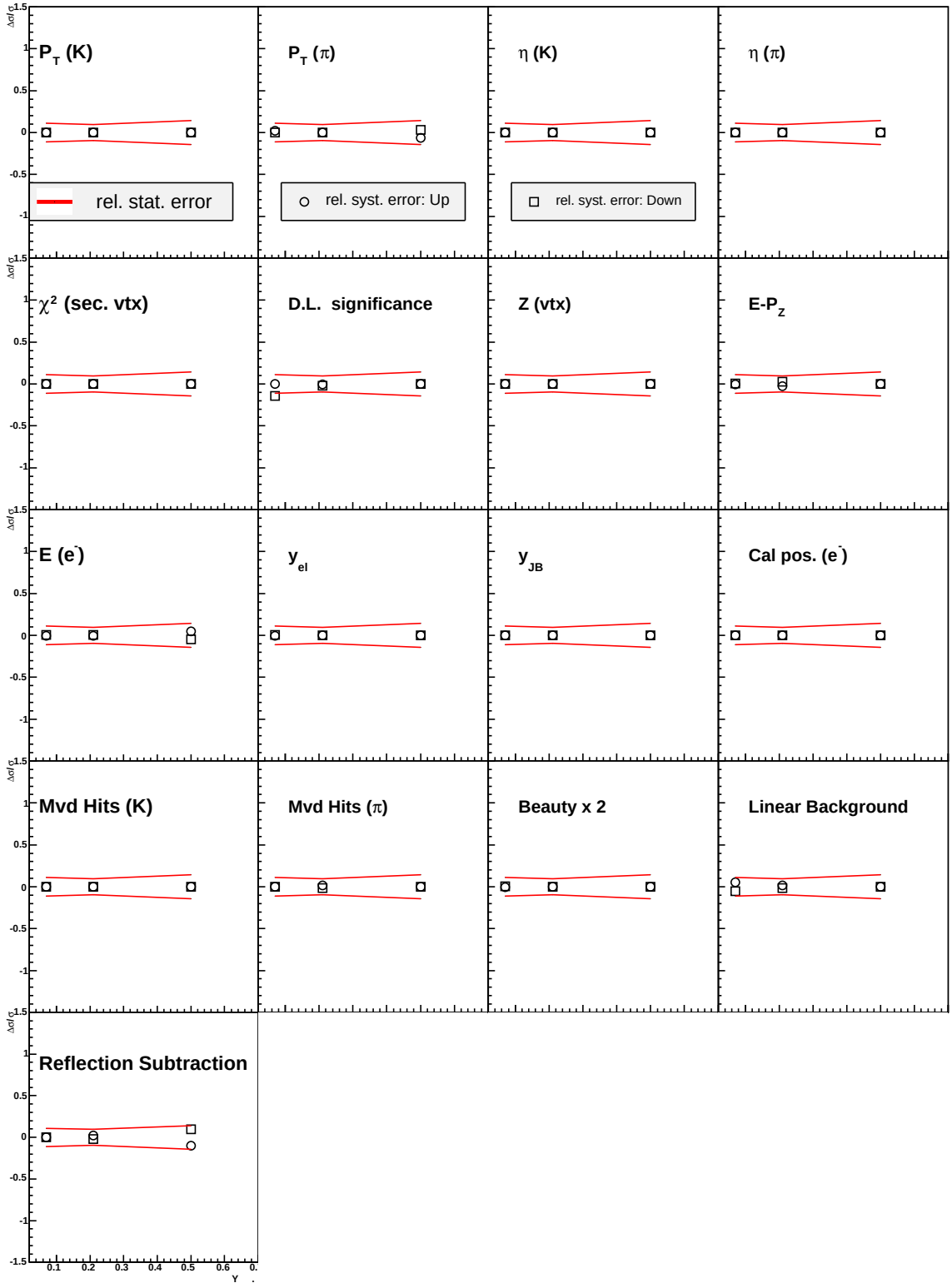


Figure A.6. The systematic errors on the double differential cross-section $d\sigma/dQ^2dy$, for $9 < Q^2 < 44\,\mathbf{GeV^2}$, are shown for each source.

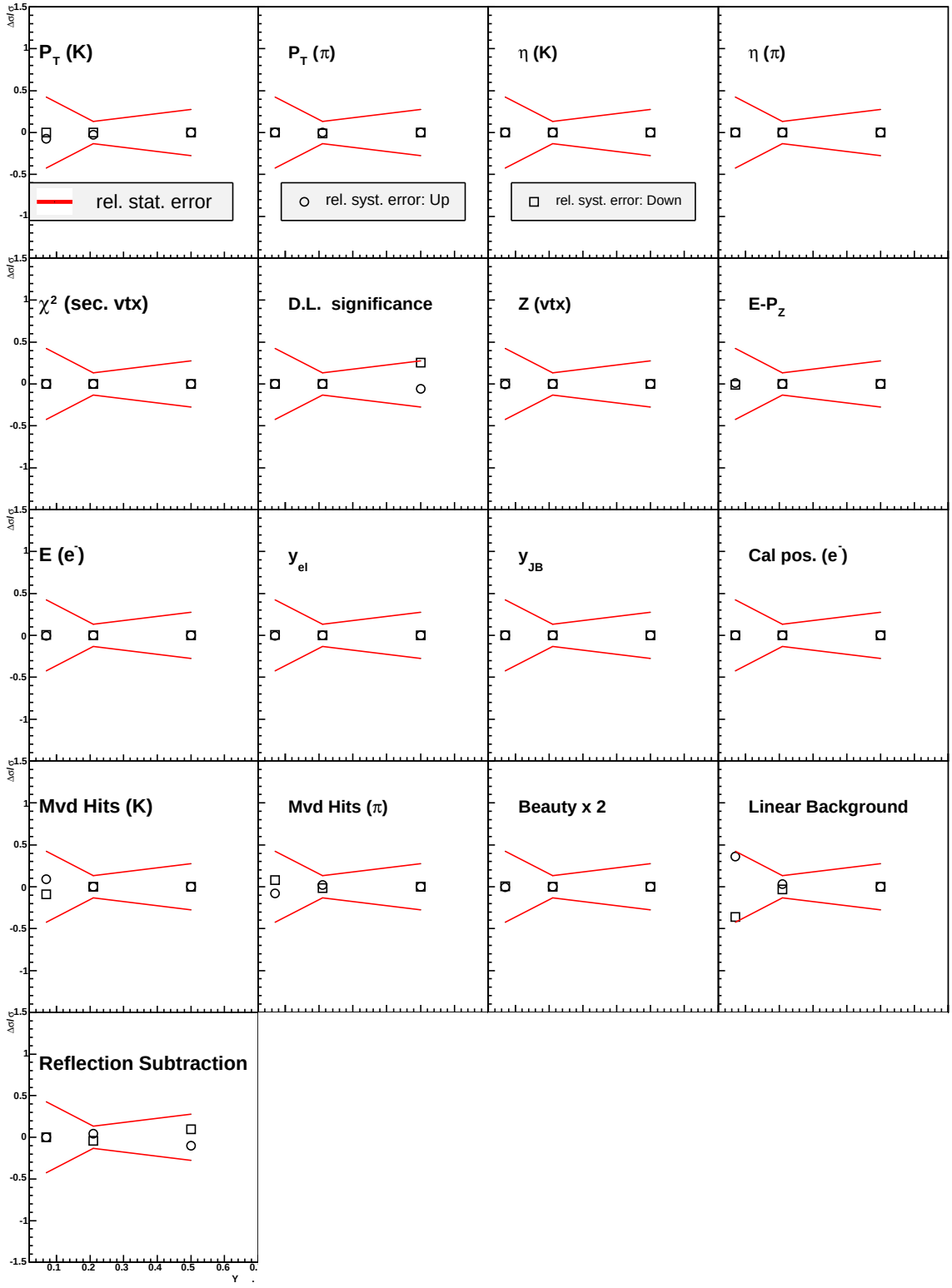


Figure A.7. The systematic errors on the double differential cross-section $d\sigma/dQ^2dy$, for $44 < Q^2 < 1000\,\mathbf{GeV^2}$, are shown for each source.

Appendix B

Cross-section tables

In this appendix, the measured cross-section values are listed in tables. The two different measurements were presented in detail in Chap. 7 and 8.

$P_T(D^0) > 3\,\mathbf{GeV}$

For the process $e^- p \to e^- + D^0 + X$, differential cross-sections as a function of Q^2, $P_T(D^0)$, $\eta(D^0)$ and Bjorken x have been measured in the kinematic range:

- $5\,\mathbf{GeV^2} < Q^2 < 1000\,\mathbf{GeV^2}$
- $0.02 < y < 0.7$
- $3 < P_T(D^0) < 20\,\mathbf{GeV}$

The cross-section values are given in Tables B.1 , B.2, B.3 and B.4.

$P_T(D^0) > 1.5\,\mathbf{GeV}$

For the process $e^- p \to e^- + D^0 + X$, differential cross-sections as a function of Q^2, $P_T(D^0)$, $\eta(D^0)$ and Bjorken x have been measured in the kinematic range:

- $5\,\mathbf{GeV^2} < Q^2 < 1000\,\mathbf{GeV^2}$
- $0.02 < y < 0.7$
- $1.5 < P_T(D^0) < 20\,\mathbf{GeV}$

The cross-section values are given in Tables B.5, B.6, B.7 and B.8.

The double differential cross-section in the bin $\sigma(Q^2_{low}, Q^2_{high}, y_{low}, y_{high})$ has been measured for the same kinematic region, of $1.5 < P_T(D^0) < 20\,\mathbf{GeV}$, and is given in Table B.9. The measured structure function $F_2^{c\bar{c}}$ for the same kinematic range is presented in Table B.10.

Bin edges Q^2(GeV^2)	Cross-section (nb/GeV^2)	Stat. error	Syst. (up)	Syst. (down)
5 - 15	0.119	0.009	0.006	0.013
15 - 40	0.0426	0.0029	0.0031	0.0024
40 - 100	0.0105	0.0010	0.0012	0.0025
100 - 1000	0.000357	0.000060	0.000018	0.000035

Table B.1. The measured cross-section $d\sigma/dQ^2$ in bins of Q^2. The values correspond to the kinematic range defined in Sec. 6.3, for the process $e^-p \to e^- + D^0 + X$.The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges P_T(GeV)	Cross-section (nb/GeV)	Stat. error	Syst. (up)	Syst. (down)
3 - 3.5	1.82	0.21	0.14	0.23
3.5 - 4.5	1.011	0.086	0.1	0.034
4.5 - 6	0.435	0.037	0.078	0.037
6 - 20	0.0361	0.0028	0.0008	0.01

Table B.2. The measured cross-section $d\sigma/dP_T(D^0)$ in bins of the transverse momentum $P_T(D^0)$ of the charmed meson. The values correspond to the kinematic range defined in Sec. 6.3, for the process $e^-p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges η	Cross-section (nb)	Stat. error	Syst. (up)	Syst. (down)
(-1.6) - (-0.8)	0.55	0.07	0.07	0.27
(-0.8) - (-0.4)	1.14	0.1	0.39	0.1
(-0.4) - (0)	1.15	0.11	0.08	0.21
(0) - (0.4)	1.29	0.12	0.15	0.11
(0.4) - (0.8)	1.36	0.14	0.04	0.09
(0.8) - (1.6)	0.77	0.12	0.14	0.1

Table B.3. The measured cross-section $d\sigma/d\eta(D^0)$ values in bins of the pseudo-rapidity η of the charmed meson. The values correspond to the kinematic range defined in Sec. 6.3, for the process $e^-p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges x	Cross-section (nb)	Stat. error	Syst. (up)	Syst. (down)
10^{-4} - $5 \cdot 10^{-4}$	2124	200	290	530
$5 \cdot 10^{-4}$ - 10^{-3}	1230	109	184	83
10^{-3} - $3.5 \cdot 10^{-3}$	386	27	19	31
$3.5 \cdot 10^{-3}$ - 10^{-1}	7.38	0.86	0.26	0.55

Table B.4. The measured cross-section $d\sigma/dx$ values are given in bins of the Bjorken x. The values correspond to the kinematic range defined in Sec. 6.3, for the process $e^-p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges Q^2(GeV^2)	Cross-section (nb/GeV^2)	Stat. error	Syst. (up)	Syst. (down)
5 - 10	0.557	0.060	0.124	0.054
10 - 20	0.221	0.018	0.014	0.024
20 - 40	0.0669	0.0078	0.0085	0.0085
40 - 80	0.0188	0.0028	0.0017	0.0025
80 - 200	0.00290	0.00062	0.00062	0.00040
200 - 1000	0.000193	0.000090	0.000147	0.000071

Table B.5. The measured cross-section $d\sigma/dQ^2$ values in bins of Q^2. The values correspond to the kinematic range defined in Sec. 8.1, for the process $e^-p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are given for each bin.

Bin edges P_T(GeV)	Cross-section (nb/GeV)	Stat. error	Syst. (up)	Syst. (down)
1.5 - 2.4	2.97	0.42	0.30	0.35
2.4 - 3	2.13	0.28	0.16	0.10
3 - 4	1.50	0.13	0.08	0.12
4 - 6	0.554	0.040	0.051	0.025
6 - 20	0.0386	0.0032	0.0012	0.0121

Table B.6. The measured cross-section $d\sigma/dP_T(D^0)$ values in bins of the transverse momentum $P_T(D^0)$ of the charmed meson are given. The values correspond to the kinematic range defined in Sec. 8.1, for the process $e^-p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges η	Cross-section (nb)	Stat. error	Syst. (up)	Syst. (down)
(-1.6) - (-0.8)	1.63	0.24	0.20	0.22
(-0.8) - (-0.4)	2.21	0.26	0.50	0.17
(-0.4) - (0)	2.33	0.26	0.18	0.18
(0) - (0.4)	2.81	0.32	0.24	0.40
(0.4) - (0.8)	2.81	0.36	0.18	0.21
(0.8) - (1.6)	2.24	0.40	0.43	0.34

Table B.7. The measured cross-section $d\sigma/d\eta(D^0)$ in bins of the pseudo-rapidity η of the charmed meson. The values correspond to the kinematic range defined in Sec. 8.1, for the process $e^-p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges x	Cross-section (nb)	Stat. error	Syst. (up)	Syst. (down)
$8 \cdot 10^{-5} - 4 \cdot 10^{-4}$	6081	748	758	1192
$4 \cdot 10^{-4} - 1.5 \cdot 10^{-3}$	2848	222	187	179
$1.5 \cdot 10^{-3} - 5 \cdot 10^{-3}$	440	48	49	50
$5 \cdot 10^{-3} - 10^{-2}$	139	20	16	34
$10^{-2} - 10^{-1}$	2.42	1.12	0.70	1.42

Table B.8. The measured cross-section $d\sigma/dx$ in bins of Bjorken x. The values correspond to the kinematic range defined in Sec. 8.1, for the process $e^- p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

Bin edges $Q^2(\mathrm{GeV}^2)$	Bin edges y	Cross-section (nb)	Stat. error	Syst. (up)	Syst. (down)
$5-9$	$0.02-0.12$	0.79	0.18	0.01	0.04
	$0.12-0.3$	1.03	0.17	0.36	0.07
	$0.3-0.7$	0.69	0.18	0.18	0.32
$9-44$	$0.02-0.12$	1.49	0.16	0.03	0.22
	$0.12-0.3$	1.48	0.14	0.04	0.05
	$0.3-0.7$	1.28	0.18	0.13	0.10
$44-1000$	$0.02-0.12$	0.214	0.091	0.016	0.081
	$0.12-0.3$	0.587	0.079	0.	0.036
	$0.3-0.7$	0.330	0.091	0.090	0.019

Table B.9. The measured cross-section *in the bin* $d^2\sigma/dQ^2dy$. The values correspond to the kinematic range defined in Sec. 8.1, for the process $e^- p \to e^- + D^0 + X$. The statistical and the upper and lower systematic errors are shown for each bin.

$Q^2(\mathrm{GeV}^2)$	x	$F_2^{c\bar{c}}$	ΔF_{stat}	ΔF_{sys}^{up}	ΔF_{sys}^{down}
7	$2.22 \cdot 10^{-4}$	0.191	0.049	0.051	0.088
	$4.61 \cdot 10^{-4}$	0.157	0.026	0.055	0.012
	$2.01 \cdot 10^{-3}$	0.067	0.015	0.001	0.004
20.39	$6.48 \cdot 10^{-4}$	0.293	0.041	0.019	0.015
	$1.34 \cdot 10^{-3}$	0.186	0.017	0.004	0.005
	$5.88 \cdot 10^{-3}$	0.100	0.011	0.001	0.010
112	$3.55 \cdot 10^{-3}$	0.186	0.051	0.052	0.011
	$7.38 \cdot 10^{-3}$	0.229	0.030	0	0.009
	$3.23 \cdot 10^{-2}$	0.046	0.019	0.005	0.025

Table B.10. Measured $F_2^{c\bar{c}}$ values at fixed Q^2 and x. The statistical and systematic errors are also shown.

Bibliography

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

[2] M. Gell-Man, *A schematic model of baryons and mesons*, Phys. Lett. **8** (1964) 214.

[3] G. Zweig, *An SU3 model for strong interaction symmetry and its breaking*, 1964, CERN TH-412 (unpublished).

[4] M. Althoff et al.[TASSO collaboration], *Measurement of the Structure Function F2(γ) At Q^2 from $7\,GeV^2$ to $70\,GeV^2$*, Z. Phys. **C 31** (1986) 527.

[5] S. Bethke, *Experimental Tests of Asymptotic Freedom*, [hep-ex/0606035 v2].

[6] G. Altarelli and G. Parisi, *Asymptotic Freedom in parton language*, Nucl. Phys. **B 126** (1977) 298.

[7] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt, *Update of parton distributions at NNLO*, Phys. Lett. **B 652** (2007) 292, [arXiv:0706.0459].

[8] A.D. Martin et al., *MRST2001: Partons and α_s from precise deep inelastic scattering and Tevatron jet data*, Eur. Phys. J. **C 73** (2002) 73, [hep-ph/0110215].

[9] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **207** (2002) 12, [hep-ph/0201195].

[10] E. Laenen, S. Riemersma, J. Smith, and W.L. van Neerven, *Complete $\mathcal{O}(\alpha_s)$ corrections to heavy flavor structure functions in electroproduction*, Nucl. Phys. **B 392** (1993) 162.

[11] S. Chekanov et al. [ZEUS Collab.], *Measurement of Charm Fragmentation Ratios and Fractions in Photoproduction at HERA*, Eur. Phys. J. **C 44** (2005) 351, [hep-ex/0508019].

[12] D. E. Groom et al.[Particle Data Group Collab.], *Review of Particle Physics*, Eur. Phys. J. **C 15** (2000) 1, [http://pdg.lbl.gov].

[13] U. Holm (ed.) ZEUS collab., *The ZEUS detector*, Status report (unpublished), DESY 1993 [http://www-zeus.desy.de/bluebook/bluebook.html].

[14] M. Derrick et al., *Design and construction of the ZEUS barrel calorimeter*, Nucl. Instr. Meth. **A 309** (1991) 77.

[15] A. Andresen et al.[ZEUS Calorimeter Group Collab.], *Construction and beam test of the ZEUS forward and rear*

calorimeter, Nucl. Inst. Meth. **A 309** (1991) 101.

[16] A. Bernstein et al.[ZEUS Barrel Calorimeter Group Collab.], *Beam tests of the ZEUS barrel calorimeter*, Nucl. Inst. Meth. **A 336** (1993) 23.

[17] N. Harnew et al., *Vertex triggering using time difference measurements in the ZEUS Central Tracking Detector*, Nucl. Inst. Meth. **A 279** (1989) 290.

[18] B. Foster et al. [ZEUS collab.], *The design and construction of the ZEUS Central Tracking Detector*, Nucl. Inst. Meth. **A 338** (1994) 254.

[19] D. Dannheim et al., *Design and tests of the silicon sensors for the ZEUS microvertex detector*, 2003, [hep-ex/0211026].

[20] R. E. Kalman, *A new approach to linear fitting and prediction problems*, Transactions of the ASME-Journal of Basic Engineering **82** (1960) 35.

[21] R. Fruchwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research **A 262** (1987) 444.

[22] E. Maddox, *Study of Heavy Quark Production at HERA using the ZEUS microvertex detector*, Ph.D. thesis, NIKHEF, 2004.

[23] ZEUS collaboration, *The design and performance of the ZEUS Micro Vertex detector*, Nucl. Inst. Meth. **A 581** (2007) 656, [arXiv:/abs/0708.3011].

[24] Nikhef Institute, *Mechanical Design of the Micro Vertex Detector*, [http://www.nikhef.nl/pub/departments/mt/projects/zeus/vertex/website.html].

[25] T. Matsushita et al., *Optical Alignment System for the ZEUS Micro Vertex Detector*, Nucl. Inst. Meth. **A 466** (2001) 383.

[26] K. Korsak-Gorzo et al., *The optical alignment system of the ZEUS microvertex detector*, Nucl. Inst. Meth. **A 580** (2007) 1227, [doi:10.1016/j.nima.2007.06.046].

[27] Zeus collab., *ZEUS Lectures*, [http://www-zeus.desy.de/lectures/].

[28] R. Luchsinger and C. Grab, *Vertex reconstruction by means of the method of Kalman filtering*, Computer Physics Communications **76** (1993) 263.

[29] A. Keramydas, Ph.D. thesis, NIKHEF, to be published.

[30] S.M. Fisher and M. Palazzi, *The ADAMO data system*, 1993.

[31] H. Abramowicz, A. Caldwell, and R. Sinkus, *Neural network based electron identification in the Zeus calorimeter*, Nucl. Inst. Meth. **A 365** (1995) 508.

[32] A. Lopez-Duran Viani and S. Schlenstedt, *Electron finder efficiencies and impurities. A comparison between SINISTRA95, EM and EMNET*, [ZEUS note 99-077, 2000].

[33] S. Schagen, *Charm in the proton*, Ph.D. thesis, NIKHEF, 2004.

[34] N. Tuning, *Proton Structure Functions at HERA*, Ph.D. thesis, NIKHEF, 2001.

[35] 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 **Vol. 1** (1992) , (eds.) W. Buchmuller and G. Ingelman, Hamburg, Germany, DESY.

[36] http://root.cern.ch.

[37] G. Corcella et al., *HERWIG 6.3 release note*, [hep-ph/0107071] .

[38] G. Corcella et al., *HERWIG 6.1 release note*, [hep-ph/9912396] .

[39] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, and E. Norrbin, *High-energy-physics event generation with PYTHIA 6.1*, 2001, [hep-ph/9601282].

[40] H. L. Lai et al.[CTEQ Collaboration], *Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions*, 2000, [hep-ph/9903282].

[41] M. Gluck, E. Reya, and A. Vogt, *Parton structure of the photon beyond the leading order*, Phys. Rev. **D 45** (1992) 3986.

[42] R. Brun, F. Bruyant, M. Maire, A.C. McPherson, and P. Zanarini, *GEANT3*, Internal report CERN DD/EE/84-1.

[43] *The Particle Data Group*, [http://pdg.lbl.gov/].

[44] The ROOT Math library team, *The Minuit Package*, [http://project-mathlibs.web.cern.ch/project-mathlibs/sw/Minuit2/html/index.html].

[45] S. Chekanov et al. [ZEUS collab.], *Measurement of charm fragmentation ratio's and fractions in photoproduction at HERA*, Eur. Phys. J. **C 44** (2005) 351.

[46] ZEUS collab., *Measurement of the D+ lifetime with the ZEUS Micro-Vertex-Detector at HERA-II*, submitted to ICHEP 2006[http://www-zeus.desy.de/physics/hfla /public/PublicResults/].

[47] ZEUS collab., *Measurements of charm fragmentation fractions and ratios in deep inelastic scattering at HERA*, JHEP **7** (2007) 74, [doi:10.1088/1126-6708/2007/07/074].

[48] B. W. Harris and J. Smith, *Heavy-quark correlations in deep-inelastic electroproduction*, Nucl. Phys. **B 452** (1995) 109, [hep-ph/9503484v1].

[49] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, *Scaling violations in inclusive e^+e^- annihilation spectra* , Phys. Rev. **D 27** (1983) 105.

[50] R. Akers et al.[OPAL Collab.], *A measurement of the Production of D^* Mesons on the Z^0 resonance*, Z. Phys. **C 67** (1995) 27.

[51] S. Chekanov et al. [ZEUS Collab.], *Measurement of the neutral current cross section and F2 structure function for deep inelastic e^+p scattering at HERA*, Eur. Phys. J. **C 21** (2001) 443.

Samenvatting

Het meestvoorkomende element in het heelal is waterstof. Het vormt meer dan 75% van de zichtbare materie[1]. Het proton is de kern van het waterstof atoom. Protonen werden in groten getale gecreëerd ongeveer 10^{-6} seconde na de Big Bang. Protonen worden stabiel verondersteld[2].

In het midden van onze zon fuseren protonen met elkaar tot helium kernen. Dit vindt plaats bij een druk die honderd miljard keer hoger is dan de atmosferische druk van de aarde en een temperatuur van rond vijftien millioen graden. Deze kern reactie voorziet de zon van energie die constant vrijkomt aan het oppervlak in de vorm van, onder andere, fotonen. Dit is de energie die het leven op aarde mogelijk maakt en onderhoudt.

Na de ontdekking van het elektron door J.J.Thomson in 1897, werd het duidelijk dat atoom kernen positief geladen zijn, zodat de atomen zelf neutraal zijn. In 1918 voerde Ernest Rutherford verstrooiïngs experimenten uit met helium kernen, ook wel alpha deeltjes genoemd, op stikstof gas. Hij bemerkte dat de scintillatie detector signalen van waterstof kernen detecteerde. Rutherford bepaalde dat deze kernen alleen van het stikstofgas afkomstig konden zijn en zo concludeerde hij dat stikstof kernen uit waterstof kernen bestaan. Dit impliceert dat de waterstof kern een elementair deeltje is, genaamd het proton.

Rutherford postuleerde het bestaan van nog een ander deeltje, het neutron genaamd, dat gevormd kan worden als een proton een electron vangt. Het neutron is in massa vrijwel gelijk aan het proton, maar het is elektrisch neutraal. Het neutron werd ontdekt in 1932 door James Chadwick. Het is een instabiel deeltje met een levensduur van ongeveer 15 minuten voordat het vervalt tot een proton, een elektron en een neutrino.

Als neutronen zich samen met protonen in een atoomkern bevinden, zijn zij

[1] Slechts 4% van de totale energie dichtheid bestaat uit zichbare materie. De rest wordt gevormd door donkere materie (22%) en donkere energie (74%).

[2] De experimentele ondergrens van de levensduur van een proton is 10^{35} jaar.

echter stabiel. Twee protonen en twee neutronen vormen de heliumkern; 92 protonen en 146 neutronen vormen de uraniumkern, het zwaarste element op aarde.

De ontwikkeling van deeltjes versnellers in de jaren 50 leidde tot de ontdekking van verschillende deeltjes, hadronen genaamd. Het leek dat zulke grote aantallen deeltjes niet elementair konden zijn. Het idee van quarks kwam voort uit een klassificatie van hadronen: het quark model was onafhankelijk ontwikkeld door Murray Gell-Mann and Kazuhiko Nishijima in 1961. Er bestaan 6 verschillende soorten quarks: up, down, charm, strange, top and bottom. Ze hebben elk hun antideeltje, dat een identieke massa maar tegenovergestelde quantumgetallen heeft. Het wordt verondersteld dat het proton, het neutron en alle anderen hadronen samengesteld zijn uit quarks. De ontdekking van Δ^{++}, een hadron dat bestaat uit 3 identieke *up* quarks, was in tegenspraak met de theorie: Pauli's *uitsluitings principe* ging hier niet op.

Een oplossing voor dit problem werd gevonden in 1965 door Young Han, Yoichiro Nambu and Oscar W. Greenberg die onafhankelijk van elkaar voorstelden dat quarks een aanvullende *SU(3) ijk vrijheidsgraad* hadden, ook wel kleur lading genoemd. In de jaren 70 werd een geunificieerde theorie ontwikkeld, het Standaard Model genaamd. Deze theorie bracht drie fundamentele krachten onder één noemer: de electromagnetische, de sterke en zwakke kern krachten.

Volgens het Standaard Model werken nieuwe deeltjes, ijkbosonen genaamd, als kracht mediatoren en ontstaan in de ijk symmetrieën van de elementaire bouwstenen van de theorie, de quarks en leptonen.

Quantum chromo-dynamica is een quantum velden theorie, onderdeel van het Standaard Model, dat de sterke interactie via gluonen, *de vector bosonen van het kleuren veld*, beschrijft. Gluonen hebben zelf kleur lading en kunnen zodoende met elkaar een wisselwerking aangaan. Experimenteel bewijs van gluonen werd in 1979 gegeven bij de Petra collider in Duitsland. Vanuit het perspectief van de chromo-dynamica, werd het proton een erg dynamisch en complex object, bestaande uit 3 quarks (*uud*) en een eeuwig veranderende zee van virtuele quarks en gluonen.

De structuur van het proton is uitvoerig bestudeerd door de ZEUS collaboratie met behulp van de HERA deeltjes versneller in Hamburg, Duitsland. Daar werden protonen en elektronen tot zeer hoge energieën gebracht en vervolgens frontaal tegen elkaar gebotst. Door deze botsing valt het proton uiteen en komen er quarks en/of gluonen vrij. De quarks en gluonen kunnen zich niet volledig vrijmaken maar combineren tot deeltjes zodat er een cascade van deeltjes de ZEUS detector doorkruist.

Uit deze deeltjes kan informatie afgeleid worden over de oorspronkelijke structuur van het proton. De HERA deeltjes versneller kan gezien worden als een gigantische elektronen microscoop die de onderdelen van het proton tot in detail onderzoekt.

De huidige analyse meet productie van charm quarks in *diepe inelastische verstrooiïng* door reconstructie van D^0 charm mesonen, die vervallen via $D^0 \rightarrow K + \pi$. De meting is sterk afhankelijk van informatie van de microvertex detector (MVD) en van nauwkeurige spoor- en vertexreconstructie. Op het niveau van de detector is het recent mogelijk geworden de microvertex detector nauwkeuriger te calibreren. Dit is mogelijk door het gebruik van sporen afkomstig van de $e^- p$ botsing, wat leidde tot een verbeterd oplossend vermogen in de MVD van ongeveer 20 micron.

Dit vertaalt zich tot een oplossend vermogen voor de impact parameter voor sporen ten opzichte van het $e^- p$ interactie punt van ongeveer 100 micron. Dit is de theoretisch best mogelijke waarde, gegeven de lange extrapolatie lengte en het dode material in de MVD. Om optimaal gebruik te kunnen maken van het oplossend vermogen van de MVD, werd een analyse-afhankelijke benadering van vertexreconstructie gebruikt: voor elke gebeurtenis, werden secundaire vertices van paren sporen met tegengestelde lading geconstrueerd, terwijl de rest van de sporen in de gebeurtenis de positie van de primaire interactie bepaalden. Karakteristieke vervals variabelen werden gereconstrueerd in de buurt van de vervalsvertex. Deze variabelen zijn bijvoorbeeld vervalslengte, significantie van de vervalslengte, de eigen levensduur en een pseudospoor dat wordt geassocieerd met het oorspronkelijke vervallende deeltje. Bovendien werden deze variabelen gebruikt als *filter* parameters om het D^0 meson signaal te versterken. Dit leidde tot een nauwkeurigere meting. Ook stond deze reconstructie van het verval toe een signaal bij lage waarden van de transversale impuls van het meson te detecteren. Dit resulteerde in een significante toename van de statistiek, wat een betere granulariteit van de meting mogelijk maakte.

Charm quark productie werd gemeten als een functie van foton virtualiteit Q^2, transversale impuls van het D^0 meson, pseudo-rapidity van het D^0 meson alsmede de Bjorken x waarde van de gebeurtenis. Deze differentiële werkzame doorsneden werden vervolgens vergeleken met de *next-to-leading-order* theoretische voorspelling van QCD. Overeenstemming werd gevonden in de vorm van alle verdelingen van de werkzame doorsneden. Gemeten werd dat de algemene normalisatie 20-25% lager was dan de voorspelde theoretische waarde. Ook werd de bijdrage van de charm quark aan de totale F_2 structuur functie bepaald. Er werd aangetoond dat

de charm productie gedomineerd wordt door gluons bij lage Bjorken x. Als de discrepantie tussen de theoretische en gemeten normalisaties geïnterpreteerd wordt als een discrepantie in de gluon dichtheid, dan zal elke doorsnede afhankelijk van gluon-gluon fusie bij de *Large Hadron Collider* met ongeveer 40% lager uitvallen.

Acknowledgements

The work presented in this book would have never been completed without the help and support of many people, inside and outside the physics world. The word "thank" will be abused in the following paragraphs but genuinely meant every time.

First, I would like to thank the ZEUS group at NIKHEF for pointing me in the right direction from the very beginning. Sven Schagen had the initiative to help our analysis framework get started. Erik Maddox gave often precious advice and help while he was at NIKHEF as well as afterwards. Henk Tiecke has always been interested in my research, looking for (and often finding) little problems with my analysis that only an experienced eye can spot at once. He is also the person who gave me and Avraam the first tours of DESY, ZEUS control room, the first lecture on the calorimeter, etc. I would like to thank Nicola Coppola, the NIKHEF post-doc at DESY, for going with me through all the horrible (or so it seemed then) scripts and codes, setting up the computing frame that I used for my analysis, as well as for keeping me entertained while at DESY. This includes the detailed discussions about the subtlety of the German and Italian languages.

My thanks also go to the Heavy Flavor Group members at DESY. John Loizides is both a working mate and a good friend who could often make me laugh with his positive approach to everything, from tedious physics to love adventures. John has been a constant help and more than once have I called him at late hours because I was stuck with my reasearch. Recently, under John's guidance, Dan Nicholass and I, both performing similar analyses, exchanged loads of material and information regarding our results, which generated interesting debates. John and Dan, I couldnt have made it without your help, guys!

I would like to thank my promotor Els Koffeman for her constant interest in the status of my work, for the help she offered and for the sense of responsibility that sometimes a young researcher needs to be reminded of. Els made me give my first talk on K^0's, two weeks after I started my PhD. She organized video conferences

between NIKHEF and DESY while I was in Hamburg. In this way, Erik and Henk could follow our work (mine and Avraam's). Although not a ZEUS member for more than a year now, Els continued to "check on how and what I was doing", offering good advice and solutions to my physics troubles. She has helped me greatly in writing this manuscript.

A special thanks goes to my promotor Paul Kooijman, the coolest supervisor I had and will probably ever have. Paul taught me to think and work independently. He got me involved into expanding the vertexing package, a central issue to my analysis. He provided detailed help during my work on the neutral pseudo-track. His suggestions regarding the direction of my research often payed off. Also, helping Paul teach second year students at Utrecht University has been a fun thing to do. Thanks for making my PhD a very pleasant experience, Paul.

I want to thank also my Greek colleague and friend, Avraam Keramydas, for all those wonderful times in Hamburg and Utrecht, for the good laughs and the talks, driving around or waiting for the Domino pizza deliveries, making electronic music or doing physics. I wish you good luck with finishing your PhD, malaka!

I am also grateful to all the smokers at NIKHEF for keeping me company through all those coffee breaks. So, Gordon, Ertan, Peter, Mikolaj, Fabian, Manouk, cheers guys and I hope we'll all quit one day.

I would like to thank the entire Utrecht gang, my dear bunch of international students, for being my social support, my critics and my crying shoulder. Many thanks to good friends and fantastic musicians Arnaud Koetsier and Catalin Rusu, for playing together with me. I hope our collaboration will not stop here. Damien, cheers for taking care that I did "socialize" sometimes in Amsterdam, for the poker nights or the live music event hunts.

Last but not least, I would like to say few words to the ones dearest to me, my girlfriend and my family.

Thank you, Fietje, for being there for me so many times. For defining my moral scale, for the great support throughout my stressy moments. For all the trips and adventures we went through during these last four years. Good luck with your PhD and with moving in together ;)

As vrea sa multumesc familiei mele, parintilor mei si fratelui meu Nutu, pentru tot ceea ce fac si tot ceea ce inseamna pentru mine. Pentru ca m-au incurajat in interesul meu pentru stiintele exacte, pasiunea pentru muzica, pentru faptul ca au fost alaturi de mine, fizic sau mental, la toate intamplarile din viata mea.

ISBN/EAN: 978-90-6488-030-8