Charm and Beauty Production in $D^*\mu$ Events at H1/HERA

Dissertation

zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von

ANGELA ISABELA LUCACI-TIMOCE

aus Năsăud

Hamburg
2007

Gutachter/Gutachterin der Dissertation:	Prof. Dr. E. Elsen Dr. A. Geiser
Gutachter/Gutachterin der Disputation:	Prof. Dr. E. Elsen Prof. Dr. R.-D. Heuer
Datum der Disputation:	18.07.2007
Vorsitzende/Vorsitzender des Prüfungsausschusses:	Dr. H. D. Rüter
Dekan der Fakultät Mathematik, Informatik und Naturwissenschaften	Prof. Dr. A. Frühwald

Abstract

Charm and beauty photoproduction is investigated with the H1 detector at HERA using events with a reconstructed D^* meson and a muon. Data taken during the years 1999-2000 and 2004-2006, corresponding to a total integrated luminosity of 320 pb^{-1}, are analysed. The D^* mesons with transverse momentum $p_t(D^*) > 1.5$ GeV and pseudorapidity $|\eta(D^*)| < 1.5$ are reconstructed via the decay channel $D^{*\pm} \to \overset{(-)}{D}{}^0\pi_s^\pm \to (K^\mp\pi^\pm)\pi_s^\pm$. In addition, muons with momentum $p(\mu) > 2$ GeV and $|\eta(\mu)| < 1.735$ are selected. With this selection, the contribution of light quark initiated events is negligible. The fractions of charm and beauty events in data are extracted exploiting the charge and azimuthal angle correlation between the D^* meson and the muon. The charm and beauty cross sections were measured in photoproduction, i.e. for photon virtualities $Q^2 < 1$ GeV^2, and for inelasticities $0.05 < y < 0.75$:

$$\begin{aligned}\sigma_{vis}(ep \to ec\bar{c}X \to eD^{*\pm}\mu X) &= 314 \pm 33\ (\text{stat}) \pm 101\ (\text{syst})\ \text{pb} \\ \sigma_{vis}(ep \to eb\bar{b}X \to eD^{*\pm}\mu X) &= 90 \pm 20\ (\text{stat}) \pm 26\ (\text{syst})\ \text{pb}.\end{aligned}$$

The measurements are compared to leading order QCD calculations supplemented with parton showers as predicted by the PYTHIA Monte Carlo event generator. A ratio data/theory of 1.2 ± 0.1 (stat) ± 0.4 (syst) and of 1.3 ± 0.3 (stat) ± 0.4 (syst) is obtained for charm and beauty, respectively. Differential cross sections for the variables describing the $D^*\mu$ system like transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, invariant mass $M(D^*\mu)$ and inelasticity $y(D^*\mu)$ are measured and compared to the PYTHIA and the CASCADE Monte Carlo event generators. The theoretical models describe the shape of the differential distributions. Transverse momentum studies of the $D^*\mu$ pair indicate a sensitivity to differing implementations of QCD dynamics in the simulation programs.

Kurzfassung

In dieser Analyse wird die Produktion von Charm- und Beauty-Quarks am Elektron-Proton-Speicherring HERA untersucht. Hierzu werden Ereignisse mit einem D^*-Meson und einem Myon selektiert. Die Analyse beruht auf Daten die in den Jahren 1999-2000 und 2004-2006 vom H1-Detektor aufgenommen wurden. Die analysierte Daten entsprechen einer integrierten Luminosität von 320 pb^{-1}.

D^*-Mesonen mit einem Transversalimpuls $p_t(D^*) > 1.5$ GeV und einer Pseudorapidität $|\eta(D^*)| < 1.5$ werden unter Verwendung des Zerfallskanals $D^{*\pm} \to \overset{(-)}{D}{}^0\pi_s^\pm \to (K^\mp\pi^\pm)\pi_s^\pm$ rekonstruiert. Zusätzlich wird ein Myon mit einem Impuls $p(\mu) > 2$ GeV und einer Pseudorapidität $|\eta(\mu)| < 1.735$ gefordert. Mit dieser Selektion, ist der Anteil an leichten Quarks vernachlässigbar. Der Anteil an Charm- und Beauty-Ereignissen wird anhand der Ladungs- und Winkelkorrelationen zwischen dem D^*-Meson und dem Myon bestimmt. Die Charm- und Beauty-Wirkungsqueschnitte werden im Photoproduktionsbereich, d.h. für $Q^2 < 1$ GeV^2, und für eine Inelastizität $0.05 < y < 0.75$ gemessen:

$$\begin{aligned}\sigma_{vis}(ep \to ec\bar{c}X \to eD^{*\pm}\mu X) &= 314 \pm 33\ (\text{stat}) \pm 101\ (\text{syst})\ \text{pb} \\ \sigma_{vis}(ep \to eb\bar{b}X \to eD^{*\pm}\mu X) &= 90 \pm 20\ (\text{stat}) \pm 26\ (\text{syst})\ \text{pb}.\end{aligned}$$

Die Messungen werden mit QCD-Rechnungen in führender Ordnung, wie sie vom PYTHIA-Generator vorhergesagt werden, verglichen.

Das Verhältnis zwischen gemessenen und vorhergesagten Wirkungsquerschnitten beträgt 1.2 ± 0.1 (stat) ± 0.4 (syst) für Charm- und 1.3 ± 0.3 (stat) ± 0.4 (syst) für Beauty-Produktion. Differentielle Wirkungsquerschnitte werden als Funktion des Transversalimpulses $p_t(D^*\mu)$, der Pseudorapidität $\eta(D^*\mu)$, der invarianten Masse $M(D^*\mu)$ und der Inelastizität $y(D^*\mu)$ gemessen, und mit den Vorhersagen von den Ereignissgeneratoren PYTHIA und CASCADE verglichen. Die theoretischen Modelle beschreiben den Verlauf der differentiellen Verteilungen. Studien zum Transversalimpuls des $D^*\mu$-System zeigen, dass diese Grösse von den unterschiedlichen QCD-Rechnungen zur Produktion schwerer Quarks beinflusst ist.

Abstract

În această lucrare se studiază producerea de cuarci 'charm' şi 'beauty' cu detectorul H1 la acceleratorul HERA, folosind evenimente cu un mezon D^* şi un miuon. Au fost analizate datele din anii 1999-2000 şi 2004-2006, corespunzând unei luminozităţi de 320 pb^{-1}. Mezonii D^* cu impuls transversal $p_t(D^*) > 1.5$ GeV şi cu pseudorapiditate $|\eta(D^*\mu)| < 1.5$ au fost reconstruiţi folosind canalul de dezintegrare $D^{*\pm} \to \overset{(-)}{D}{}^0\pi_s^\pm \to (K^\mp\pi^\pm)\pi_s^\pm$. În plus, sunt selectaţi miuoni cu impuls $p(\mu) > 2$ GeV şi cu pseudorapiditate $|\eta(\mu)| < 1.735$. Cu această selecţie, contribuţia datorită cuarcilor uşori poate fi neglijată. Fracţiunile de evenimente conţinând cuarci 'charm' si 'beauty' au fost extrase pe baza corelaţiilor sarcinii electrice şi a diferenţei unghiulare azimutale ale mezonului D^* si ale miuonului. Au fost măsurate următoarele secţiuni eficace pentru 'charm' şi 'beauty':

$$\begin{aligned} \sigma_{viz}(ep \to ec\bar{c}X \to eD^{*\pm}\mu X) &= 314 \pm 33 \text{ (stat)} \pm 101 \text{ (sist) pb} \\ \sigma_{viz}(ep \to eb\bar{b}X \to eD^{*\pm}\mu X) &= 90 \pm 20 \text{ (stat)} \pm 26 \text{ (sist) pb.} \end{aligned}$$

Măsurătorile au fost comparate cu predicţii în ordinul întâi ale cromodinamicii cuantice obţinute cu programul Monte Carlo PYTHIA.

A fost obţinut un raport date/teorie de 1.2 ± 0.1 (stat) ± 0.4 (sist) pentru 'charm' şi de 1.3 ± 0.3 (stat) ± 0.4 (sist) pentru 'beauty'. Secţiunile eficace diferenţiale au fost măsurate ca funcţie de variabilele care descriu sistemul $D^*\mu$, şi anume impuls transversal $p_t(D^*\mu)$, pseudorapiditate $\eta(D^*\mu)$, masă invariantă $M(D^*\mu)$ şi inelasticitate $y(D^*\mu)$, şi comparate cu programele Monte Carlo PYTHIA şi CASCADE. Modelele teoretice descriu forma distribuţiilor diferenţiale. Studii ale impulsului transversal $p_t(D^*\mu)$ au indicat că această variabilă este dependentă de diferitele predicţii ale cromodinamicii cuantice pentru producerea de cuarci grei implementate în programele Monte Carlo.

Contents

Introduction

The constituents of matter and their interactions can be studied in the energetic collisions of particles which take place in particle accelerators. These studies are the object of *high energy physics*. The theory describing the elementary particles that make up matter and forces between them (strong, weak and electromagnetic[1]) is called the *Standard Model*. An important component of the Standard Model is *Quantum Chromodynamics* (QCD), which encloses the theory of the strong interaction between quarks and gluons, both carrying colours. In this theory, the cross section of a process may be calculated based on perturbative expansion in the coupling constant. Perturbative calculations can be applied if there is a large scale compared to the QCD parameter $\Lambda_{QCD} = 0.26$ GeV [1]. The heavy quarks, i.e. c, b and t, which have masses $m \gg \Lambda_{QCD}$, provide such a scale.

The subject of this thesis is the study of heavy quark photoproduction with the H1 detector at the HERA accelerator, at DESY, Hamburg (Germany). At HERA energies, the heavy quarks are mostly charm and beauty. In the leading order picture, they are produced via boson-gluon fusion.

This thesis is structured as follows. In the first chapter, theoretical considerations of heavy quark production in ep collisions, the general ideas of QCD and the evolution equations which are implemented in Monte Carlo models, are introduced. Next, the experimental methods of reconstructing heavy quarks are discussed, with the emphasis on the double tagging method using a D^* meson and a muon. The H1 detector components relevant for this analysis are described in chapter 2. The event and trigger selection, the D^* meson and the muon reconstruction are treated extensively in chapter 3. A comparison of the kinematic variables of interest in data to PYTHIA Monte Carlo is done in chapter 4. Based on the charge and azimuthal angular correlation between the D^* meson and the muon, charm and beauty fractions in data are determined. The method used is presented in chapter 5. Next, the total charm and beauty cross sections are measured in the visible range of the analysis. Chapter 7 presents the measurement of the differential cross sections in bins of combined $D^*\mu$ variables, and their comparison to the leading order QCD PYTHIA and CASCADE generators. Studies of the mean $D^*\mu$ transverse momentum, which could be sensitive to the different evolution schemes implemented in Monte Carlo models, are discussed in chapter 8.

The results are summarised in chapter 9.

In this thesis, the natural system of units, in which $\hbar = c = 1$, is used.

[1] The gravitational force is not included in the Standard Model.

Chapter 1

Heavy Quark Production at HERA

In this chapter a theoretical and experimental survey of heavy quark production at the HERA collider is given. First, the kinematic variables used to describe the *ep* interaction are introduced. Then the basic ideas of QCD, the theory of strong interaction, are discussed. Finally, the experimental methods to reconstruct the heavy quark pair and existing measurements are reviewed.

1.1 Electron-Proton Scattering

At HERA accelerator, beams of electrons[1] collide with beams of protons. In this section, the kinematic variables necessary to understand the theoretical description of heavy quark production are presented.

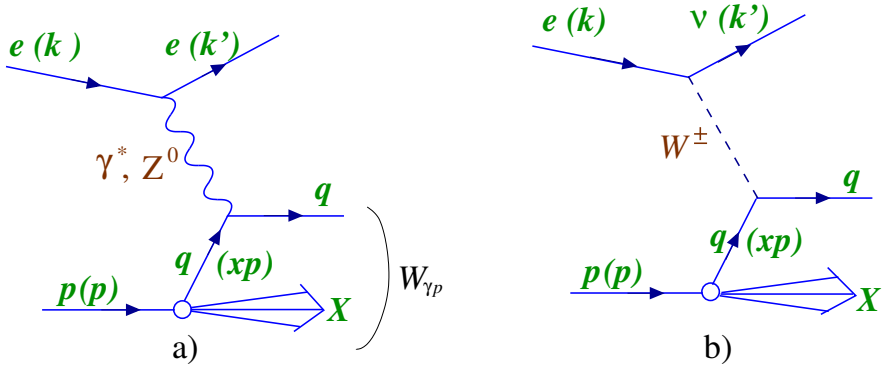

Figure 1.1: *Feynman diagram of the ep scattering: a) neutral current, b) charged current events.*

In leading order (LO), the electron-proton scattering $ep \to \ell X$ ($\ell = e$ or ν) takes place via the exchange of a gauge boson between the electron and a quark in the proton. Depending on the type of the boson, which can be *neutral* (a virtual photon γ^* or a Z^0) or

[1] HERA can work both with electrons and positrons, so 'electron' generically refers to both species in this thesis.

charged ($W^{\pm}$), one speaks about *neutral current (NC)* or *charged current (CC)* events (see figure 1.1).

The initial and final state particles are characterised by the four-momenta:

- incoming electron: $k = (E_e, \vec{k})$;
- proton: $p = (E_p, \vec{p})$;
- outgoing electron: $k' = (E_{e'}, \vec{k'})$.

Neglecting the electron and proton masses, which are small compared to the involved energies, the Lorentz invariant variables that describe the event kinematics are:

$$s = (k+p)^2 \simeq 4E_eE_p; \tag{1.1}$$
$$Q^2 = -q^2 = -(k-k')^2,\ 0 \leq Q^2 \leq s; \tag{1.2}$$
$$W^2_{\gamma p} = (q+p)^2 \simeq -Q^2 + y \cdot s; \tag{1.3}$$
$$x = \frac{Q^2}{2 \cdot p \cdot q},\ 0 \leq x \leq 1; \tag{1.4}$$
$$y = \frac{p \cdot q}{k \cdot p},\ 0 \leq y \leq 1; \tag{1.5}$$

where:

- $\sqrt{s}$ is the centre-of-mass energy[2];
- q is the four-momentum of the exchange boson;
- $W_{\gamma p}$ is the boson-proton centre-of-mass energy;
- $x,\ y$ are the Bjorken variables. These are constrained to take values between 0 and 1 as a result of four-momentum conservation. In *quark parton model* (QPM) x is the fraction of the proton's momentum carried by the struck quark, and y is the fraction of the incident electron's momentum carried by the exchange boson in the rest frame of the proton.

The above quantities are related by:

$$Q^2 = x \cdot y \cdot s. \tag{1.6}$$

For fixed beam energies, only two of these kinematic variables are independent.

The value of Q^2 is a measure of the *virtuality* of the exchange boson and provides an interaction resolution scale. The size of the proton is of the order $\mathcal{O}(10^{-15}\text{m})$, equivalent to $Q^2 \approx 1\ \text{GeV}^2$. At $Q^2 \gtrsim 1\ \text{GeV}^2$, the virtual photon probes the structure of the proton, enabling the internal constituents to be investigated. This type of interaction is called

[2] At HERA, the electron beam-energy is $E_e = 27.5$ GeV and the proton-beam energy: $E_p = 920$ GeV, such that $\sqrt{s} = 318$ GeV.

deep inelastic scattering[3] (DIS), and it is dominated by NC photon exchange events. For $Q^2 < 1\,\mathrm{GeV}^2$, the exchange photon is considered to be quasi-real, or *on mass shell*. This interaction is termed **photoproduction**.

In QPM, the proton is seen as a set of smaller point-like particles, called 'partons' (identified as the valence quarks which are required by the SU(3) symmetry), each carrying definite fractions of the proton momentum. In this model, one assumes that the nucleon has a very high momentum, such that it is Lorentz contracted in the longitudinal direction. Due to time dilation, the electron, in the short time it needs to cross the nucleon, sees the partons 'frozen', i.e. they do not interact with each other.

The general form of the *ep* cross section is:

$$d\sigma_{ep} \propto L_{\mu\nu} W^{\mu\nu}, \tag{1.7}$$

where $L_{\mu\nu}$ and $W^{\mu\nu}$ are leptonic and hadronic tensors. The latter can be expressed [2] using the *structure functions* $F_i(x,\ Q^2),\ i = 1,\ 2$. In QPM, the structure functions are independent of Q^2 (*Bjorken scaling*), and are directly related to the *parton distribution functions*[4], $f_i(x)$ by:

$$F_2(x) = 2xF_1(x) = \sum_i e_i^2 \cdot x \cdot f_i(x). \tag{1.8}$$

The sum runs over all partons with charge e_i.

The QCD description modifies the simple QPM picture: in addition to valence quarks, the proton contains a sea of quarks and gluons. The Bjorken scaling is violated, i.e. F_2 depends on Q^2: $F_2 = F_2(x,\ Q^2)$.

1.2 Equivalent Photon Approximation

To understand the emission of quasi-real photons by electrons, one considers the field of a fast charged particle to be similar with the electromagnetic radiation. The latter is interpreted as a flux of photons with energy distribution $n(y)$, y being the energy fraction of the photons relative to the initial electron energy.

In the *equivalent photon approximation*, the *ep* cross section is composed of the γp cross section times the flux factor [3]:

$$d\sigma_{ep}(y,\ Q^2) = \sigma_{\gamma p}(y) \cdot dn(y,\ Q^2), \tag{1.9}$$

with Q^2 the photon virtuality.

Neglecting the details of Q^2 dependence and terms involving the longitudinal photon polarisation, Weizsäcker and Williams calculated the equivalent number of photons as:

$$dn(y,\ Q^2_{max}) = f_{\gamma/e}(y,\ Q^2_{max})dy, \tag{1.10}$$

[3] 'Deep' refers to the high value of the momentum transferred from the electron to the constituents of the nucleon, and 'inelastic' because the final state contains in addition other particles, apart from the scattered electron and the proton.

[4] The parton distribution function $f_i(x)$ gives the probability to find a parton i, with a momentum fraction x, at scale Q^2, in the proton.

with the photon flux given by:

$$f_{\gamma/e} = \frac{\alpha_{em}}{2\pi} \cdot \left[\frac{1+(1-y^2)}{y} \cdot \ln \frac{Q^2_{max}}{Q^2_{min}} - 2m_e^2 y \cdot \left(\frac{1}{Q^2_{min}} - \frac{1}{Q^2_{max}} \right) \right]. \tag{1.11}$$

The kinematical limits [4] on Q^2 are: $Q^2_{min} \approx \frac{y^2}{1-y} \cdot m_e^2$ and $Q^2_{max} \approx (1-y)s$.

1.3 Quantum Chromodynamics

The theory which describes the strong interaction is called *Quantum Chromodynamics*, QCD[5] [5]. The quarks have a quantum number called *colour*. They can come in three colours: red, green and blue, denoted symbolically by R, G and B, respectively. The observed experimental bound states $q\bar{q}$ (mesons) and qqq (baryons) are considered to be colourless or white.

The basic features of QCD are *asymptotic freedom* and *quark confinement*: for very small separations, the quarks behave as free, non-interacting particles, whereas at large distances they are confined and cannot be separated. Confinement has not been proven. Experimentally, free quarks have never been observed.

1.3.1 Asymptotic Freedom

David J. Gross, H. David Politzer and Frank Wilczek received the Nobel Prize in 2004 for their discovery of asymptotic freedom [6]. This feature can be explained based on the dielectric properties of the vacuum and on the self-interaction of the gluon field[6].

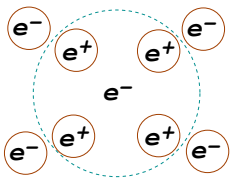

Figure 1.2: *In QED, the electron charge is screened by positron charges. If one places a test charge closer to the electron, the measured charge is larger.*

In quantum electrodynamics (QED), an electron emits photons, which may subsequently form electron-positron pairs before they are reabsorbed. These pairs constitute electric dipoles which align with the electron in an energetically favourable configuration, as illustrated in figure 1.2. The vacuum around the electron has become a polarised medium.

The Coulomb force experienced by a test charge depends on the distance to the electron: with decreasing distance, there are less positrons which shield the electron, and the measured charge is larger. This effect is known as *charge screening* [5].

In QCD, virtual quark-antiquark pairs also tend to screen the colour charge. However, there is a major difference compared to QED: in contrast to photons, which do not carry electric charge, the gluons carry colour charge and may interact with themselves (see figure 1.3).

[5] The Greek word 'chroma' means colour.

[6] The asymptotic freedom can be also described as a paramagnetic effect due to the spin of the gluons (see [7]).

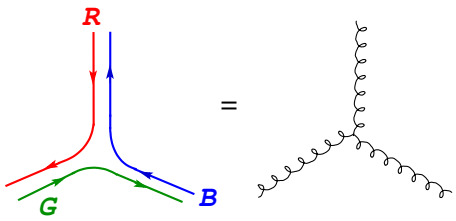

Figure 1.3: *Gluon self-coupling [5].*

A gluon carries colour away from the test quark, which consequently alters its colour, weakening the original colour content. One says that the gluon *anti-screens* the quark.

The virtual quarks and gluons have different effects on the quark: screening and anti-screening, respectively. As long as there are no more than 16 flavours of quarks, without counting the anti-quarks separately, the anti-screening prevails [8]. This leads to a reduction of the colour force at small distances.

1.3.2 Ultraviolet and Infrared Divergences

The interaction processes between particles may be represented graphically by Feynman diagrams. In perturbative QCD, cross sections are calculated, based on Feynman rules, as power series in the strong coupling constant α_s. Because the momentum of particles in loops contained in the diagrams is not constrained by the requirement of energy-momentum conservation, divergences occur when integrating over the momentum of internal particles. These divergences are of two types: **ultraviolet divergencies**[7] which are associated with infinite momenta, i.e. with very high energy approaching infinity, or equivalently, very short distances, and **infrared divergencies** that correspond to very small energies (approaching zero), or equivalently very long distances. These divergences are not specific to QCD, but general in field theory, and several techniques were developed to 'regularise' them.

To handle the ultraviolet divergencies, one renormalises the divergent integrals by introducing cut-offs, e.g. an ultraviolet momentum cut-off. For the divergent diagram, an arbitrary scale μ_R is introduced, which is not present in the original Lagrangian of the theory. The new parameter is called the *renormalisation scale*. A set of rules, called *renormalisation scheme*, is needed to determine this scale. The procedure is not unique. The physical result does not depend on the choice, but a theoretical dependence is obtained in QCD, since the perturbative expansion is done only up to a given order. One can use for example the *minimal subtraction scheme*, in which μ is chosen to be the same for every divergent diagram [10].

The infrared divergencies can be removed by applying the factorisation theorem (see below).

1.3.3 Factorisation Theorem

Consider an interaction in which the building blocks of the hadronic matter, the quarks and the gluon, are involved. In general, the cross section of such a process is a combination

[7] The term 'ultraviolet' was borrowed from classical mechanics, where it was used in context of the *ultraviolet catastrophe*. This refers to a prediction from the beginning of the twentieth century that an ideal black body at thermal equilibrium will radiate infinite amounts of energy at high frequencies (on the ultraviolet side of the visible spectrum). The prediction was unrealistic, and the problem was solved later by Plank, who postulated that electromagnetic energy could only be emitted in discrete packets of energy proportional to the frequency [9].

of short and long distance contributions, and thus is not computable directly in QCD.

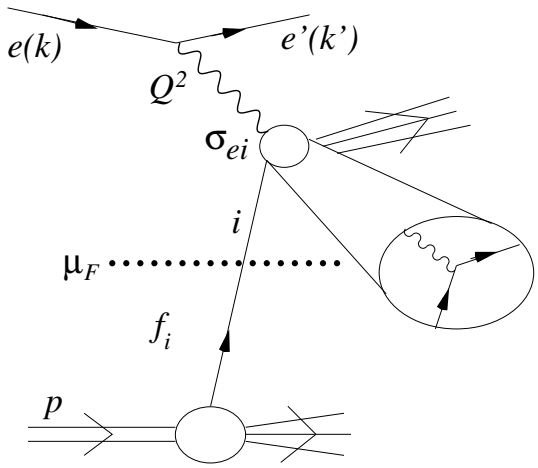

Figure 1.4: *Illustration of factorisation theorem for deep inelastic ep scattering [11].*

The *factorisation theorem* of QCD allows to derive predictions for these cross sections by separating (factorising) the two components in a systematic fashion. In deep inelastic ep scattering for example (see figure 1.4), the long distance effects (which are not perturbatively calculable) are factorised into *parton distribution functions* f_i, giving the density of partons i in the proton p. These functions can be measured experimentally and they are supposed to be universal, i.e. if they are measured for one process, they can be used for any other process [12].

The short distance effects are considered in the perturbatively calculable lepton-parton scattering cross section σ_{ei}. As in the renormalisation case, a new parameter, called *factorisation scale* (μ_F), is introduced. This parameter defines the separation between short- and long-distance effects.

The cross section is then given by [11]:

$$\sigma_{ep} = \sum_i (f_i(\mu_F) \otimes \sigma_{ei}(\mu_F)) \tag{1.12}$$

The choice of factorisation scale is arbitrary. For $\mu_F^2 = Q^2$, the function f_i is interpreted in the parton picture as the parton density in the proton seen by a photon with virtuality (resolving power) Q^2.

Note that the factorisation theorem has been proven to date to apply only for a limited number of processes, like DIS [13, 14] and Drell-Yan. For other processes, for example diffraction in hadron-hadron scattering, this theorem is not applicable. Experimentally, such an ansatz is seen to describe many other processes.

1.3.4 The Strong Coupling α_s

Due to asymptotic freedom, the QCD coupling α_s is small only at high energies, such that only in this domain precision tests can be performed using perturbation theory. The coupling α_s depends on the renormalisation scale. This dependence is controlled by the β functions, which can be calculated in QCD. If only the term containing β_0, the first β function, is kept, α_s is given by [15]:

$$\alpha_s(\mu_R) = \frac{4\pi}{\beta_0 \cdot \ln(\mu_r/\Lambda_{QCD})}, \tag{1.13}$$

where $\beta_0 = 11 - 2/3 \cdot n_f$, n_f being the number of active flavours, i.e. quarks with mass less than the energy scale μ_R. The QCD scale Λ_{QCD} is defined as the energy scale where α_s diverges to infinity, such that perturbative QCD is no more applicable. At the scale

$\mu_R = M_Z$, where M_Z is the mass of Z^0 boson, the strong coupling is $\alpha_s(M_Z) = 0.1205 \pm 0.004$ from NLO calculations [15].

The scale Λ_{QCD} is calculated in lattice QCD to be around 260 MeV [1]. Λ_{QCD} plays an important role, as it sets the size of hadrons and thus gives also the scale of the kinetic energy of quarks confined within these hadrons. Compared to Λ_{QCD}, *u, d* are *light quarks*, *s* is a borderline case, while *c, b* and *t* are *heavy quarks* [16].

1.3.5 Evolution Equations

The theory does not predict the perturbative parton distributions, they have to be measured. Remarkably, as a consequence of the factorisation theorem, once they are determined at a scale μ, one can predict them at any scale μ', for a given order, using QCD *evolution equations*.

1.3.5.1 DGLAP Equation

Considering gluon radiation and gluon splitting, as in figure 1.5, the DGLAP[8] equations [17] are obtained. The splitting function P_{gq}, for example, gives the probability to find a gluon g with momentum fraction z in a quark q.

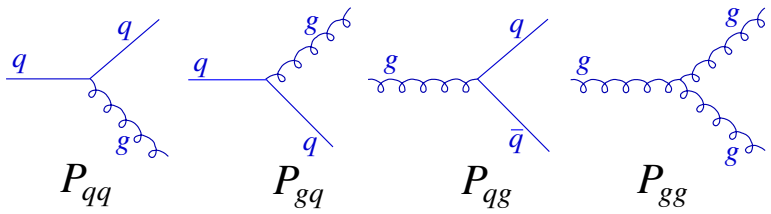

Figure 1.5: *Feynman diagrams corresponding to DGLAP splitting functions in LO.*

Because gluons may emit quark pairs, the DGLAP evolution equations for the quark density q_i for flavour i and the gluon density g_i are coupled:

$$\frac{dq_i(x,\ Q^2)}{d\ln Q^2} = \frac{\alpha_s}{2\pi} \cdot \int_x^1 \frac{dz}{z} \cdot \left[q_i(z,Q^2) \cdot P_{qq}\left(\frac{x}{z}\right) + g(z,Q^2) \cdot P_{qg}\left(\frac{x}{z}\right) \right], \tag{1.14}$$

$$\frac{dg_i(x,\ Q^2)}{d\ln Q^2} = \frac{\alpha_s}{2\pi} \cdot \int_x^1 \frac{dz}{z} \cdot \left[\sum_i q_i(z,Q^2) \cdot P_{gq}\left(\frac{x}{z}\right) + g(z,Q^2) \cdot P_{gg}\left(\frac{x}{z}\right) \right]. \tag{1.15}$$

In the DGLAP approach, the partons evolve according to the following scenario: the incoming parton from the proton (see figure 1.6), initially with low transverse momentum k_{t0} and carrying a fraction x_0 of proton's momentum, moves to higher k_t's and lower momentum fractions by successive small angle parton emissions, ending with a maximum transverse momentum given by the $q\bar{q}$ pair, or with the virtuality of the photon, Q^2.

[8] Dokshitzer, Gribov, Lipatov, Altarelli, Parisi.

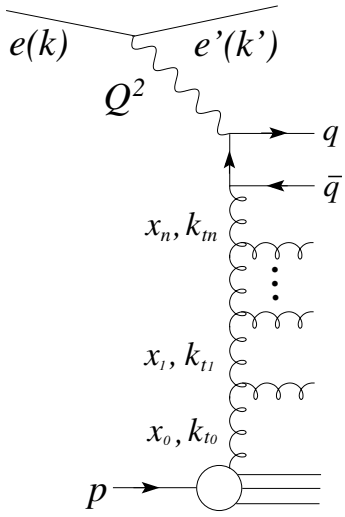

Figure 1.6: *Ladder diagram showing successive emissions of partons in DGLAP approach.*

The gluons emissions are therefore strongly ordered in k_t:

$$Q_0^2 \ll ... \ll k_{t,\,i}^2 \ll k_{t,\,i+1}^2 \ll ... \ll Q^2$$

which implies a strong ordering also for the emitted quanta.

When evaluating the ladder with n rungs, expressions proportional to $\alpha_s^n(Q^2) \cdot [\ln(Q^2/Q_0^2))]^n$ are obtained. Since each power n in α_s is accompanied by the same (maximal) power of $\ln(Q^2/Q_0^2)$, this is called *leading log approximation* (here in $\ln(Q^2/Q_0^2)$) [11].

The DGLAP approximation is expected to be valid for sufficiently large Q^2 and not too small x, in order not to produce too large logarithms:

$$\alpha_s(Q^2) \cdot \ln\frac{1}{x} \ll \alpha_s(Q^2) \cdot \ln\frac{Q^2}{Q_0^2} < 1. \qquad (1.16)$$

The DGLAP approach is referred to as the **collinear approach**, since the strong ordering implies that the virtuality of the partons entering the hard scattering matrix elements can be neglected compared to the large hard scale [18].

1.3.5.2 BFKL Equation

For small x, where the DGLAP approximation is not expected to be valid anymore, the BFKL[9] equation has been developed. In this scheme, the partons emissions follow strong ordering in fractional momenta: $x_0 \ll ... \ll x_i \ll x_{i+1} \ll ... \ll x$, but with no ordering in transverse momenta. The region of validity is:

$$\alpha_s(Q^2) \cdot \ln\left(\frac{Q^2}{Q_0^2}\right) \ll \alpha_s(Q^2) \cdot \ln\left(\frac{1}{x}\right) < 1. \qquad (1.17)$$

The BFKL [19] equation contains an explicit k_t dependence and is expressed in terms of the function $f(x, k_T^2)$, which is called **'unintegrated' gluon density**, and which is related to the DGLAP gluon density $g(x, Q^2)$ by:

$$xg(x, Q^2) \backsimeq \int_0^{Q^2} \frac{dk_T^2}{k_T^2} \cdot f(x, k_T^2). \qquad (1.18)$$

In the BFKL regime, a high-energy cross section may be factorised into an off mass-shell partonic cross section and an unintegrated parton density, both k_t dependent. This is referred to as **$\mathbf{k_t}$ factorisation** or **semi-hard approach**[10].

[9] Balitsky, Fadin, Kuraev, Lipatov.

[10] The word 'semi-hard' refers to the situation in which the involved hard scale is large compared to the Λ_{QCD} parameter, but smaller than the total energy, $\sqrt{s}$.

1.3.5.3 CCFM Equation

The CCFM[11] equation [20] unifies the BFKL and DGLAP approaches and takes into account *coherence effects* by angular ordering: during the initial cascade, the parton emission takes place in a region with increasing opening angle towards the hard scale.

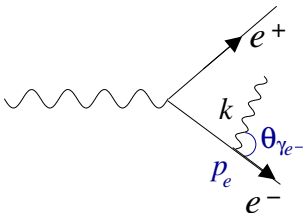

Figure 1.7: *Production of an e^+e^- pair and the emission of a soft photon, k (adapted from [21]).*

To understand the physical origin of angular ordering [21], one can consider the QED analogy of a relativistic e^+e^- pair, with e^- having the virtual mass $M = \sqrt{(p+k)^2}$ and radiating a soft photon, as in figure 1.7. To find out to what extent e^+ and e^- may emit independently photons, one estimates the formation time, i.e. the time needed for the photon to be radiated:

$$t_{form} \propto \frac{1}{M} \cdot \frac{|p_e|}{M} = \frac{p_e}{(p_e+k)^2} \backsimeq \frac{1}{k \cdot \theta^2_{\gamma e^-}}, \quad (1.19)$$

with $\theta_{\gamma e^-}$ the angle between the emitted photon and the electron. Knowing the transverse wavelength of the radiated photon:

$$\lambda_t = \frac{1}{k_t} = \frac{1}{k \cdot \theta_{\gamma e^-}}, \quad (1.20)$$

one obtains the formation time:

$$t_{form} \approx \frac{\lambda_t}{\theta_{\gamma e^-}}. \quad (1.21)$$

During this time, the e^+e^- pair separate a transverse distance:

$$d_t^{e^+e^-} \approx \theta_{e^+e^-} \cdot t_{form} \backsimeq \lambda_t \cdot \frac{\theta_{e^+e^-}}{\theta_{\gamma e^-}}. \quad (1.22)$$

For large angle photon emissions: $\theta_{\gamma e^-} \gg \theta_{e^+e^-}$ (or $\theta_{\gamma e^+} \gg \theta_{e^+e^-}$), the transverse distance is $d_t^{e^+e^-} < \lambda_t$, such that the emitted photon cannot resolve the internal structure of the e^+e^- pair, probing only its total electric charge, which is zero. In conclusion, for $\theta_{\gamma e^-} \gg \theta_{e^+e^-}$ one expects the photon emission to be strongly suppressed, whereas for $\theta_{\gamma e^-}$ $(\theta_{\gamma e^+}) < \theta_{e^+e^-}$, i.e. $d_t^{e^+e^-} \gg \lambda_t$, the e^- and e^+ may emit photons independently. This is referred to as the **Chudakov effect** [21].

A similar situation occurs in QCD, where the coherence effects refer to soft-gluon emission. Consider a process of the type $q_0 \to qg$. The quark in the final state shares its colour with the gluon, such that q and g cannot emit subsequent gluons incoherently. Let g' be a soft gluon emitted by the $q-g$ system at large angles ($\theta_{q_0 g'} > \theta_{qg}$), as in figure 1.8.

Since this soft gluon corresponds to a large (transverse) wavelength, it cannot resolve the separate charge of q or the g, but only the net charge carried by q_0 . Therefore this soft gluon can be thought of as being radiated by q_0 rather than the $q-g$ system (coherence between the emitted soft gluon and the initial state). Then, in the soft region, the partons emitted by q and g radiate, to a good approximation, independently of each other, inside regions of decreasing angle ($\theta_{qg'} < \theta_{qg}$ and $\theta_{gg'} < \theta_{qg}$).

[11] Ciafaloni, Catani, Fiorani, Marchesini.

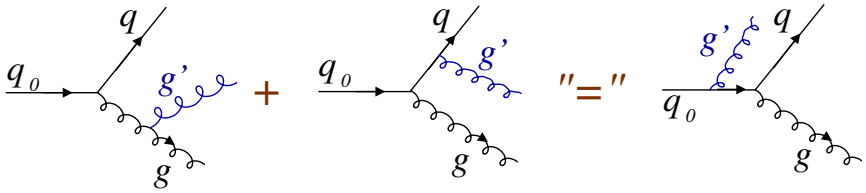

Figure 1.8: *Illustration of coherence principle in QCD: the soft gluon, g', which is emitted at wide angles, cannot resolve the $q-g$ system, and it acts as if would come from the initial quark (adapted from [21]).*

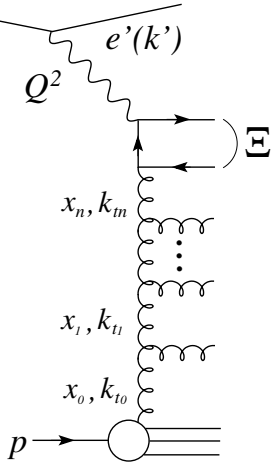

Figure 1.9: *Partons emissions in the CCFM approach: the maximum evolution angle Ξ is given by the $q\bar{q}$ pair.*

The CCFM equation reduces to the DGLAP equation at moderate x, where the angular ordering becomes an ordering in the gluon transverse momenta. At very small x, there is no constraint on the transverse momenta, and the BFKL equation is satisfied.

The differential form of the CCFM equation is the following [22]:

$$\bar{q}_t^{\,2} \cdot \frac{d}{d\bar{q}_t^2} \cdot \frac{x\mathcal{A}(x,\ k_t,\ \bar{q}_t)}{\Delta_s(\bar{q}_t,\ Q_0)} = \int dz \cdot \frac{d\Phi}{2\pi} \cdot \frac{\tilde{P}(z,\ \bar{q}_t/z,\ k_t)}{\Delta_s(\bar{q}_t,\ Q_0)} \cdot x' \cdot \mathcal{A}(x',\ k_t',\ \bar{q}_t/z), \tag{1.23}$$

where:

- $\bar{q}_t$ is the evolution variable, related to the maximum angle Ξ via

$$\bar{q}_t = x_n \cdot \sqrt{\Xi} \cdot s, \tag{1.24}$$

  with $s = (p_e + p_p)^2$ the squared centre-of-mass energy;
- $\mathcal{A}(x,\ k_t,\ \bar{q}_t)$ is the unintegrated gluon density;
- $\tilde{P}(z,\ \bar{q}_t/z,\ k_t)$ is the splitting function[12];

[12]This splitting function is different from the one in the DGLAP case, which depends only on z.

- Δ_s is the Sudakov form factor:

$$\Delta_s = \exp\left(-\int_{Q_0^2}^{\bar{q}_t} \frac{dq^2}{q^2} \cdot \int_0^{1-Q_0/q} dz \cdot \frac{\frac{3\alpha_s \cdot (q^2(1-z)^2)}{\pi}}{1-z}\right); \tag{1.25}$$

- Φ is the azimuthal angle.

1.4 Photoproduction at HERA

The gauge boson of *Quantum Electrodynamics* (QED), the photon, is regarded as being without structure. However, when probing the quantum fluctuations of the field, one may speak about photon 'structure'. For example, by interacting with a Coulomb field, the photon could materialise in an e^+e^- pair. With increasing energies, the possibility that the photon may transform to a vector meson like ρ^0, ω or Φ has to be considered. This is referred to as the *vector meson dominance model* (VDM) [23]. Later, e^+e^- experiments showed that at sufficiently high energies the photon may fluctuate into a $q\bar{q}$ pair, but without the formation of a hadronic bound state (see figure 1.10). In this analysis, the VDM contribution is irrelevant, and only the direct and the anomalous resolved contributions are considered.

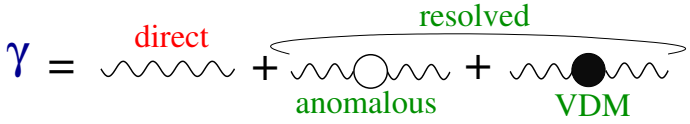

Figure 1.10: *Possible photon states: the photon may interact directly, or through its resolved states, when it fluctuates into a $q\bar{q}$ pair without forming a hadronic bound state (anomalous), or forming a vector meson (VDM)* [3].

In case of photoproduction events at HERA, one distinguishes between several contributions. If the photon interacts directly with the constituents of the proton, one speaks about **direct** production or about a 'point-like' photon (figure 1.11, a). If the photon fluctuates into a hadronic state before interacting, the photon is referred to as being **resolved** (figure 1.11, b-d). Processes in which a gluon or light quark from the photon takes part in the interaction are labelled as **normal resolved** (figure 1.11, b). In case the photon fluctuates into a heavy $q\bar{q}$ pair, and one of the quarks interacts with a parton from the proton, one speaks about **heavy quark excitation** processes (figure 1.11, c-d).

In the present analysis, the contribution of the normal resolved process is of the order of a few percent, and therefore neglected (see section 6.9), such that only the direct and excitation processes will be considered.

Note that at next to leading order (NLO), the distinction between direct and resolved processes becomes ambiguous.

The resolved processes are expected to contribute significantly in photoproduction regime [24]. As a consequence, the charm and beauty photoproduction cross sections are sensitive to the parton content of the proton and of the photon.

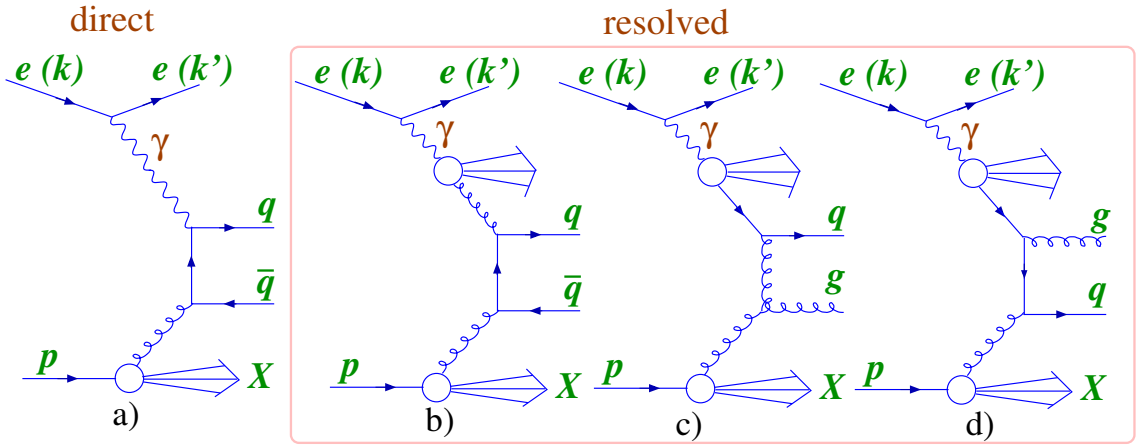

Figure 1.11: *Leading order Feynman diagrams for heavy quark production: a) direct process (γg fusion), b) normal resolved, and heavy quark excitation processes: c) with a gluon propagator, d) with a quark propagator. Because the gluon-gluon coupling is stronger than the quark-gluon coupling, the excitation process with a gluon propagator dominates over the excitation with a quark propagator.*

1.5 QCD Calculation Schemes for Heavy Quark Production

At HERA energies, heavy quarks are charm and beauty. NLO calculations for heavy quark production are available in several schemes. All approaches assume that Q^2 and the heavy quark mass m_q provide a hard enough scale, such that perturbative QCD can be applied and the factorisation theorem is valid.

Massive scheme: In this scheme[13], the charm and beauty quarks are treated fully massive. Some of the LO and NLO diagrams in this approach are shown in figure 1.12. This scheme is expected to work for $Q^2,\ p_t^2 \backsim m_q^2$, where Q^2 is the photon virtuality and p_t^2 is the transverse momentum of the heavy quark. The generator programs in this scheme, which include NLO QCD calculations, are FMNR [25] for photoproduction and HVQDIS [26] for DIS.

Massless scheme: In this scheme[14], the heavy quarks are treated as being infinitely massive below some scale $\mu < m_q$ and completely massless above scale μ. The LO and NLO processes included in this scheme are shown in figure 1.13. Note that the NLO processes in this approach are of the order α_s^1, compared to the massive scheme, in which the NLO processes are of the order α_s^2. This scheme works well for $Q^2,\ p_t^2 \gg m_q^2$.

Mixed schemes: To obtain an uniform description over the whole range in μ, composite schemes have been developed. They provide the transition from the massive ap-

[13]The massive scheme is also referred to as the 'Fixed Flavour Number Scheme' (FFNS), since the number of heavy flavours is fixed for all scales.

[14]The massless scheme is sometimes labelled as the 'Zero Mass Variable Flavour Number Scheme' (ZMVFNS).

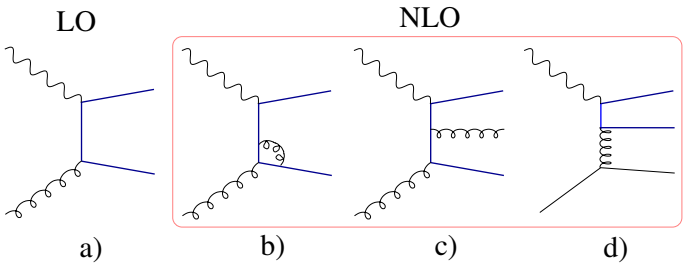

Figure 1.12: *Feynman diagrams for the processes contributing to heavy quark production in* ***massive scheme****: a) order α_s^1 boson-gluon fusion, $\gamma^* g \rightarrow q\bar{q}$; b) order α_s^2 virtual corrections to a); c) order α_s^2 real corrections involving gluon and d) light quark initial states* [27].

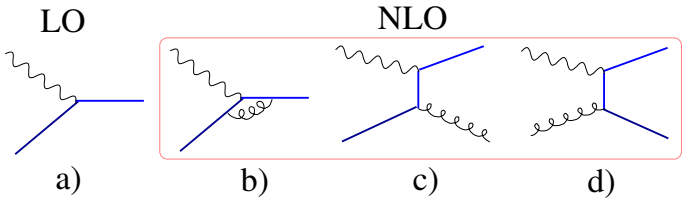

Figure 1.13: *Feynman diagrams for processes contributing to heavy quark (q) production in* ***massless scheme****: a) order α_s^0 heavy flavour excitation, $\gamma^* q \rightarrow q$, b) order α_s^1 virtual corrections to a); c) order α_s^1 real corrections to the heavy flavour excitation mechanism; d) order α_s^1 heavy flavour creation mechanism* [27].

proach (at $\mu \backsimeq m_q$) to the massless approach (at $\mu \gg m_q$). Such a mixed scheme, FFNLO ('Fixed-order Plus Next to Leading Logarithms') [28], is available for charm photoproduction at HERA, with D^* mesons in the final state.

1.6 Physics Simulation

The Monte Carlo programs generate hard scattering processes according to theoretical models and simulate the detector response to physics events . This way, the theory can be tested within the finite resolution and acceptance of the detector.

1.6.1 Event Generation

The steps of a leading order QCD generator for an *ep* scattering are depicted in figure 1.6.1. The emission of the virtual photon by the electron is described by a matrix element which can be calculated in QED. The matrix element of the hard interaction between the photon and the parton from the proton is calculated in leading order QCD. The partons are evolved from the proton side down to a cut-off scale according to QCD evolution equations (see section 1.3.5). The sequences of branchings lead to parton showers.

The result of events generation is a list of particles identified by a particle-type and their four-momentum vector. This stage is referred to as **generator level**.

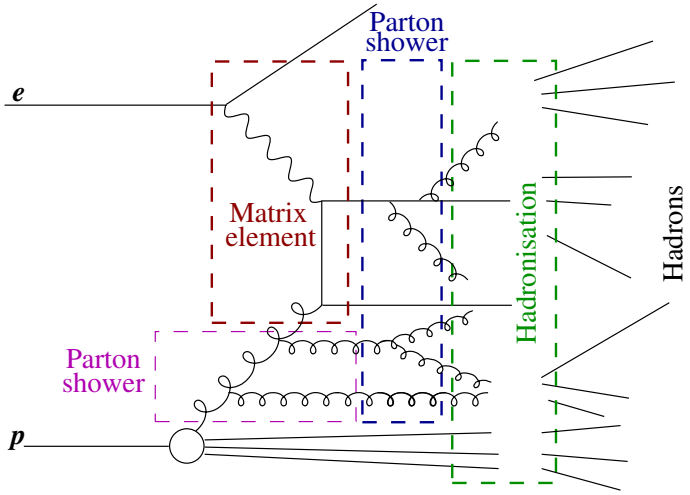

Figure 1.14: *Elements of an event generator for an ep scattering.*

1.6.2 Fragmentation and Hadronisation

At large distances, where perturbative QCD is no more applicable, the coloured partons transform into colourless hadrons, a process which is called **fragmentation**. The combination of fragmentation and the subsequent decay of unstable particles is termed **hadronisation**.

One phenomenological fragmentation model is the *Lund string model* [29, 30]. There the colour field between a massless $q\bar{q}$ pair is modelled by a massless[15] and relativistic string, which has a string tension κ. This tension gives a linear potential, similar to the QCD potential:

$$V(r) = -\frac{4\alpha}{3r} + \kappa \cdot r, \tag{1.26}$$

where r is the distance between the two colour charges. When the quarks move apart, the string stretches, and the potential energy stored in it increases. As a result, the string may split up into two colour-singlet systems: $q\bar{q}'$ and $q'\bar{q}$. If the invariant mass of the string pieces is high enough, further breaks can occur. This proceeds until there are only on-mass-shell hadrons.

Heavy quarks production is strongly suppressed ($u : d : s : c \approx 1 : 1 : 0.3 : 10^{-11}$ [31]) in this model. They are not produced in the soft fragmentation, but in perturbative parton-shower branchings $g \to q\bar{q}$.

The quark pair formation is described by fragmentation functions $f(z)$, where z is the fraction of $(E \pm p_z)$ taken by each new particle. Several parametrisations for these functions are available. If it is requested that the fragmentation process should be the same, irrespective

[15] The string is not really massless, since κ corresponds to a 'mass density' along the string.

if the flavour iteration started from the q end, or from the $\bar{q}$ end, the choice is unique, and consists in the **Lund symmetric fragmentation function**:

$$f(z) \propto z^{-1} \cdot (1-z)^a \cdot \exp(-bm_t^2/z), \tag{1.27}$$

where a and b are tunable parameters, and m_t is the hadron transverse mass[16].

For heavy quark fragmentation, harder functions are needed. This is the case for the Peterson function [32], which describes the formation of a hadron H from a heavy quark Q and a light quark q:

$$D_q^H(z) = \frac{N}{z \cdot [1 - (1/z) - \varepsilon_q/(1-z)]^2}, \tag{1.28}$$

where z is the energy fraction $z = E_H/E_q$, N a factor which normalises the total probability for hadron formation to one, and ϵ_q is a parameter that describes the hardness of the fragmentation process for a quark of type q.

1.6.3 PYTHIA

PYTHIA is a general purpose generator for ep, $\bar{p}p$ and e^+e^- events. Specific for PYTHIA is the *'backwards evolution'* [31], i.e. one starts with the hard interaction and successively reconstructs the preceding branchings. This way, the Monte Carlo evolution is done in terms of decreasing virtuality or transverse momentum, and increasing momentum fraction x. The DGLAP equation (section 1.3.5.1) describes the initial parton showers evolution and the matrix elements are calculated on-shell. The Lund symmetric fragmentation model is used for the fragmentation of light quarks, whereas the Peterson fragmentation function describes the heavy quark fragmentation. Apart from direct processes, PYTHIA generates also resolved and excitation ones (see section 1.4).

	Parameter
Proton PDF	CTEQL [33]
Photon PDF	GRV-LO [34]
$B^0 - \overline{B}^0$ mixing	$x_d = \Delta m_{B^0}/\Gamma_{B^0} = 0.73$, $x_s = \Delta m_{B_s^0}/\Gamma_{B_s^0} = 18$
Charm mass	$m_c = 1.5$ GeV
Beauty mass	$m_b = 4.8$ GeV
Heavy quark fragmentation function	Peterson $\epsilon_c = 0.078$ (0.04) $\epsilon_b = 0.008$ (0.002)

Table 1.1: *Parameters used in PYTHIA Monte Carlo program. The first values are valid for PYTHIA version 6.1, and the values in the parentheses for version 6.2. If only one value is quoted, it is valid for both versions. 'PDF' stands for parton density function.*

The parameters used in the PYTHIA Monte Carlo program are given in table 1.1 (for more details, see appendix G). Due to technical reasons, PYTHIA version 6.1 (6.2) was used for the HERA I (II)[17] run period.

[16] The transverse mass m_t of a particle is given by $m_t^2 = m^2 + p_t^2$, where p_t is the transverse momentum. If the z-axis is chosen as the beam direction, $p_t^2 = p_x^2 + p_y^2$.

[17] In this thesis, 'HERA I' refers to the data of the years 1999-2000, and 'HERA II' to 2004-2006.

1.6.4 CASCADE

CASCADE is a full Monte Carlo event generator for ep (DIS and photoproduction) and $\bar{p}p$ processes. It generates first the hard scattering process and calculates the hard cross section using off-shell matrix elements [22]. Then the initial state cascade is generated according to the CCFM equation (see section 1.3.5.3), in a backward evolution approach. Finally, the Lund string fragmentation model is used for the hadronisation processes, as in PYTHIA. The unintegrated gluon densities are obtained beforehand from a CCFM fit to the measured structure function $F_2(x,\ Q^2)$.

The specified parameters of the PYTHIA Monte Carlo were also implemented for CASCADE (see table 1.1). The CASCADE version 1.0 (1.2) was used for HERA I (II) run period.

1.6.5 Events Simulation and Reconstruction

After the generation of events, the interactions of the generated particles with the detector are modelled using a simulation package (H1SIM [35]) which is based on GEANT. The list of generated particles is fed into the GEANT program, that gives as output the x, y, z coordinates of particle trajectories through the tracking detectors and energy depositions in the calorimeters. Further, the detector response is digitised, and the trigger response is simulated.

Finally, the events are fully reconstructed (i.e. with calibrated detectors, linked tracks, event kinematics, etc.) with the H1 reconstruction package, H1REC [36]. This final step is termed **reconstruction level**. To enable comparisons with real data, the same reconstruction package is applied to data and Monte Carlo.

To reconstruct the hadronic final state particles, the *Hadroo2* algorithm is used. This algorithm creates a particle candidate from either track or calorimeter information. In case both types of information are available for a particle, the one with the best resolution is used. Special care is taken to avoid the energy double counting, i.e. if the track measurement is taken to make a particle candidate, the calorimetric energy has to be suppressed. More details about the algorithm can be found in [37].

The difference between the generated and the reconstructed level provides the corrections for detector acceptances and for resolution effects in data.

1.7 Experimental Methods and Results

At HERA, heavy quarks are produced mainly in the boson-gluon fusion process. However, due to confinement (see section 1.3), one never sees 'free' quarks, they are bound with other quarks to produce hadrons. Therefore one either observes the decay particles of these hadrons, or the 'jets'[18] which contain the hadrons coming from the heavy quarks.

[18] A 'jet' in particle physics is a group of particles having approximately the same direction, and consisting of the decay products of a member of a $q\bar{q}$ pair produced in a collision.

The most important single[19] and double tagging[20] methods used at HERA are mentioned in table 1.2. In the following, these methods, together with the experimental results, are presented in more detail.

Tagged flavour	Method
Single tagging	
c	Reconstruction of $D^{*\pm}$ in 'golden' channel $D^{*\pm} \to \overset{(-)}{D}{}^0\pi_s^\pm \to (K^\mp\pi^\pm)\pi_s^\pm$
c	Reconstruction of D mesons + lifetime tag
b	Lepton + p_t^{rel}+lifetime tag+jets
c, b	Inclusive lifetime tagging
Double tagging	
c, b	$D^*\mu$ reconstruction
b	$\mu\mu$ events

Table 1.2: *Heavy flavour tagging methods used at HERA.*

1.7.1 Single Tagging

1.7.1.1 Charm Tagging with a D^* Meson and a Jet

Information about c quarks is most often obtained at HERA by using the $D^{*\pm}$ decay channel ('golden' channel):

$$D^{*\pm} \to \overset{(-)}{D}{}^0\pi_s^\mp \to K^\mp\pi^\pm\pi_s^\pm, \tag{1.29}$$

which is depicted at the quark level in figure 1.15. The global branching ratio of this decay channel is [15]:

$$\mathcal{BR}(D^* \to K\pi\pi_s) = \mathcal{BR}(D^* \to D^0\pi_s) \cdot \mathcal{BR}(D^0 \to K\pi) = (2.57 \pm 0.05)\%. \tag{1.30}$$

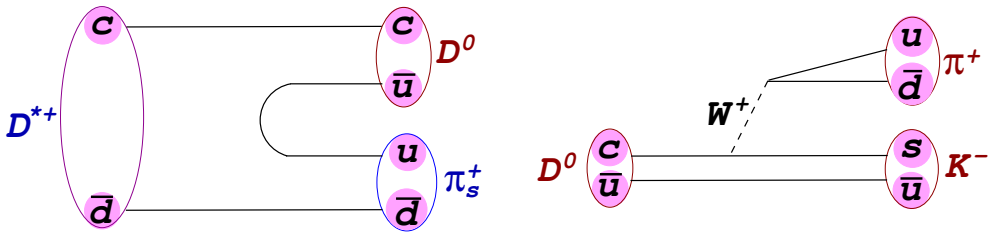

Figure 1.15: *Quark level diagram for $D^{*+} \to D^0\pi^+$ and $D^0 \to K^-\pi^+$ decays.*

In addition to the D^* meson, a jet which does not contain the D^* may be selected. This jet may come from the other heavy quark, or from a light parton of a higher order processes, such that the production dynamics in heavy flavour events can be studied in more detail.

[19] *Single tagging* refers to the situation in which only the decay particles coming from one quark are reconstructed.

[20] In case of *double tagging*, information about both heavy quarks is potentially obtained.

Jets are usually reconstructed using the inclusive k_t algorithm [38]. Two input objects are merged into a resulting jet based on distance parameters, i.e. momentum weighted distances in the $\eta - \phi$ plane. The 4-vector of the resulting jet can be obtained either by adding the 4-vectors of the input objects (E recombination scheme), or the jet is made massless, by setting its energy to be equal to the magnitude of its 3-momentum, and only the 3-vector components of the input objects are added (p_t recombination scheme).

Charm production, using events with a D^* meson and a jet defined by the k_t algorithm and in the p_t recombination scheme, has been studied for example in [39]. Noticeable differences between the cross sections as a function of pseudorapidity[21] η between the D^* and the jet were observed: the D^* distribution falls with increasing values of η, while the jet distribution is almost flat (see figure 1.16). These differences are a sign of hard non-charm partons in the forward regions. i.e. $\eta > 0$, for which the dominant mechanism is hard gluon radiation off the gluon from the proton.

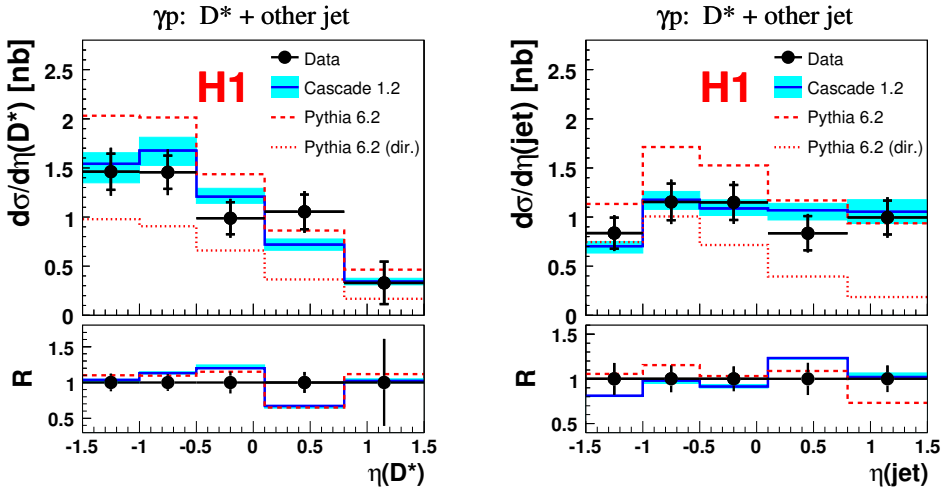

Figure 1.16: *The cross section distributions, in photoproduction (γp), as a function of pseudorapidity η for D^* meson (left) and for the jet that does not contain the D^*, i.e. other jet (from [39]). R is the data/theory ratio.*

Since in the centre-of-mass system of the interacting partons the two outgoing partons are produced back-to-back, the difference in ϕ angle between the D^* and the jet:

$$\Delta\phi(D^*,\ jet) = \begin{cases} |\phi_{D^*} - \phi_{jet}| & \text{if } |\phi_{D^*} - \phi_{jet}| \leq 180^\circ \\ 360^\circ - |\phi_{D^*} - \phi_{jet}| & \text{if } |\phi_{D^*} - \phi_{jet}| > 180^\circ \end{cases}$$

can be used to test the correct description of higher order effects in the calculations. From the cross section as a function $\Delta\Phi(D^*,\ jet)$ (see figure 1.17), it has been concluded that only approximatively 25% events originate from a back-to-back configuration (configuration predicted by the collinear approximation, in leading order). The majority of events can be described only if theoretical models with higher order QCD effects are considered.

[21] The pseudorapidity is defined as $\eta = -\ln(\tan(\theta/2))$, where θ is the polar angle.

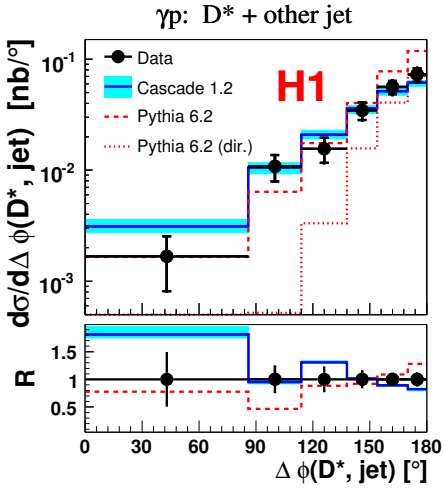

Figure 1.17: *The cross section distribution as a function of the azimuthal angular difference, $\Delta\Phi(D^*,\ jet)$ (from [39]). R is the data/theory ratio.*

1.7.1.2 Charm Tagging Using Lifetime

The charm quarks can be tagged by reconstructing also other charm hadrons than the $D^{*\pm}$ mesons, for example D^+, D^0 and D_s^+. However, the background is much higher in this case. To improve the situation, the sizable lifetime of the D mesons can be exploited[22]. The combinatorial background is dominated by light quarks which have most of the tracks originating from the primary vertex. Therefore the combinatorial background can be reduced by applying a cut on the vertex separation significance $S_l = l/\sigma_l$, where l is the radial decay length, i.e. the separation between the D meson production vertex and the decay vertex (measured by the H1 central silicon tracker), and σ_l is its error. This has been done in [41], where several decay channels of charmed mesons were used. The corresponding invariant masses are shown in figure 1.18.

In perturbative QCD, the D meson cross sections are computed as a convolution of a parton level hard scattering cross section with the *fragmentation function* $D_D^{(c)}(z)$, which describes the transition of an on-shell c quark of momentum p into a hadron D carrying a fraction z of the c quark momentum.

The differential cross sections for the production of different charm mesons can be seen in figure 1.19. The shapes of the distributions are consistent within the errors, suggesting that the fragmentation functions are similar for each type of meson (apart from a normalisation factor).

1.7.1.3 Beauty Tagging with Leptons and the Impact Parameter Method

Beauty quarks can be tagged by muons from the semileptonic decays of B hadrons inside jets from the hadronisation of b quarks. By requesting that each event has at least two jets,

[22] Charm mesons with lifetime between 0.4 and 1 psec have a typical separation between their production vertex and the decay vertex of 120 to 315 μm [40].

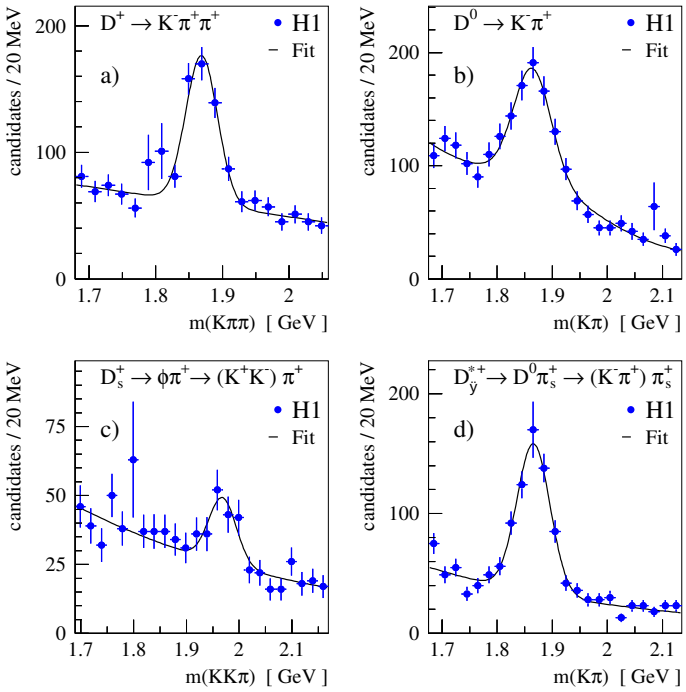

Figure 1.18: *Invariant mass distributions for the D meson candidate decays. The signals are fitted with a Gaussian function (from [41]).*

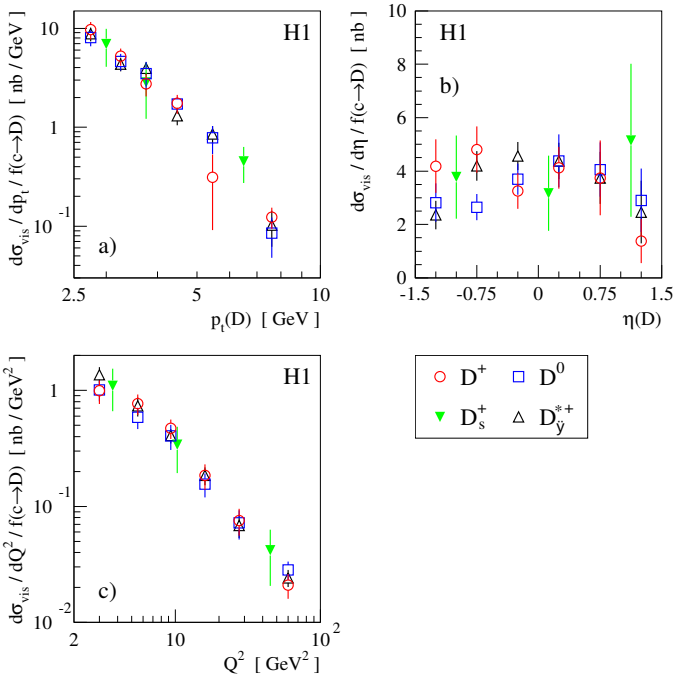

Figure 1.19: *D mesons production cross sections, divided by their respective measured fragmentation factors, as a function of a) D meson transverse momentum $p_t(D)$, b) pseudorapidity $\eta(D)$ and c) photon virtuality Q^2 (from [41]).*

from which at least one should contain a muon, a sample of events enriched in b quarks is obtained.

Beauty production measurements in events with a muon and one jet in DIS, and two jets in photoproduction, have been done in [42]. To discriminate beauty events from the charm or light quark background, one exploited two distinct properties of the B hadrons: the large mass and the long lifetime. The B hadrons large mass is reflected in a broad distribution of the transverse momentum p_t^{rel} of muons relative to the beauty quark jet (see figure 1.20). The long lifetime of B hadrons induces a large impact parameter $\delta \sim 200$ μm of the decay muon tracks relative to the primary vertex.

The transverse momentum of the muon track p_t^{rel}, measured with respect to the direction of the associated jet, is given by:

$$p_t^{rel} = \frac{|\vec{p}_\mu \times (\vec{p}_{jet} - \vec{p}_\mu)|}{|\vec{p}_{jet} - \vec{p}_\mu|}, \tag{1.31}$$

with $\vec{p}_\mu$ ($\vec{p}_{jet}$) the momentum vectors of the muon (jet) in the laboratory frame.

The *impact parameter* δ is given by the transverse distance of closest approach of the track to the primary vertex point. If the angle between the axis of the associated jet and the line joining the primary vertex to the point of closest approach of the track is less than 90°, the signed impact parameter is taken to be positive.

Depending on the lifetime of the produced hadrons, one can distinguish between different quark flavours: tracks produced at the primary vertex result in a symmetric distribution around $\delta = 0$ (see figure 1.20), while tracks from decays of long lived particles have mainly a positive δ. Negative impact parameters are mainly due to detector resolution.

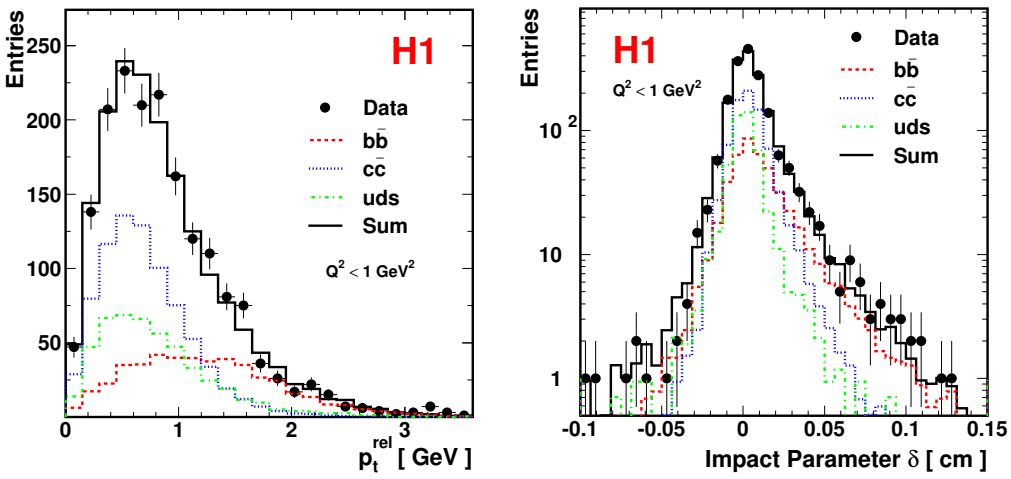

Figure 1.20: *The relative transverse momentum p_t^{rel} and the signed impact parameter δ distributions for photoproduction events (from [42]).*

Note that for the p_t^{rel} and δ variables, the direction of the jet, with which the muon is associated, needs to be reconstructed with sufficient precision. This can be done for jets

with p_t larger than 5 GeV [43], which sets a lower limit on the transverse momentum of the probed b quark.

To estimate the fraction of the photon energy entering the hard interaction, in LO QCD, the following variable is defined:

$$x_\gamma^{obs} = \frac{\sum_{Jet1}(E - p_z) + \sum_{Jet2}(E - p_z)}{\sum_h (E - p_z)}, \tag{1.32}$$

where the sum in the numerator runs over all particles associated with the two jets, and that in the denominator over all hadronic final state particles. E denotes the particle energy, and p_z the z-component of the momentum. In direct photon events (at parton level), the final state consists only of the two jets plus the proton remnant in forward direction, but since the latter does not contribute much to $\sum_h (E - p_z)$, x_γ^{obs} approaches unity. Resolved events correspond to low x_γ^{obs}.

The differential cross section as a function of x_γ^{obs} is presented in figure 1.21. Direct events contribute at x_γ^{obs} close to one, while resolved processes are present at smaller values. The best description is given here by PYTHIA Monte Carlo, which includes resolved processes. Although the NLO calculation, performed with the FMNR program, has large uncertainties due to scale variations, it describes the data reasonably well.

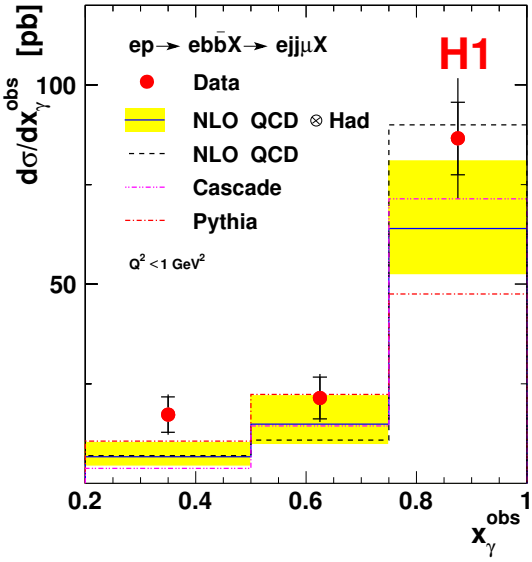

Figure 1.21: *Differential cross section as a function of x_γ^{obs} for data and Monte Carlo (from [42]).*

1.7.1.4 Inclusive Lifetime Tagging

A simultaneous charm and beauty production measurement in photoproduction has been performed in the inclusive dijet analysis [44], based on the fact that the long lifetime of c

and b flavoured hadrons leads to displacement of tracks from the primary vertex, such that a separation between heavy and light quarks is possible.

For statistical precision, the impact parameter significances, $S = \delta/\sigma(\delta)$ were defined: S_1 for events with only one selected track associated to the jet, and S_2, for jets with at least two associated tracks (see figure 1.22). To reduce the uncertainty due to δ resolution and the light quark normalisation, in the S_1 and S_2 distributions the negative bins were subtracted from the positive ones (see figure 1.22). The reference distribution for c, b and light quarks

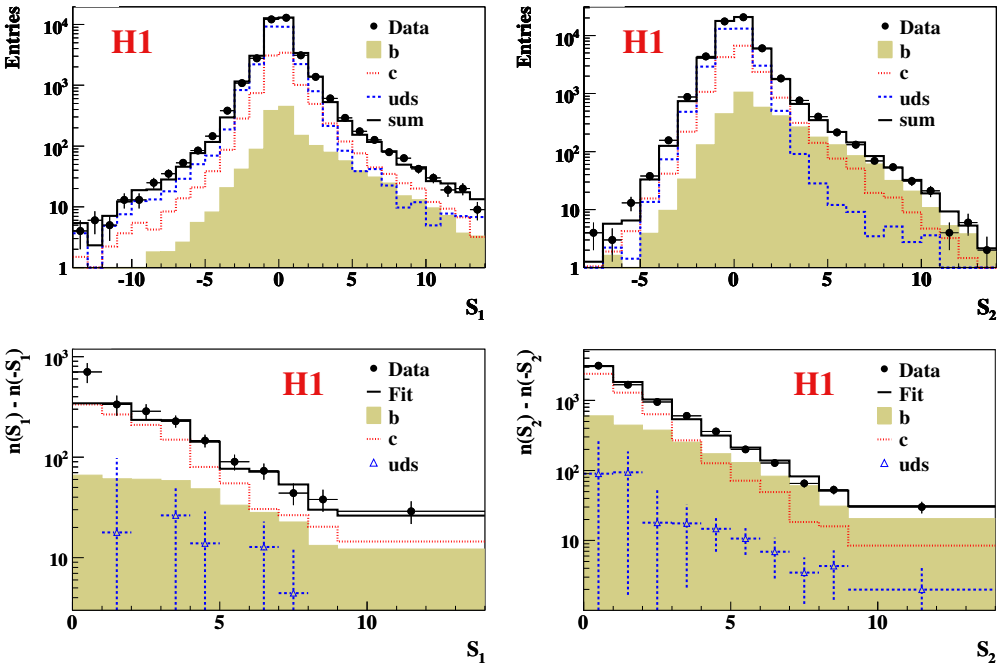

Figure 1.22: *The signed impact parameter significances S_1 and S_2 (top) and the corresponding subtracted distributions (bottom) [44]. These distributions are dominated by charm quarks at small values of significance, and by beauty quarks at larger values. The light quark contributions are small.*

were taken from PYTHIA. By fitting them to the measured S_1 and S_2 distributions, the c, b and light quark fractions in the data were obtained.

In figure 1.23, the measured beauty production cross section is compared to NLO QCD predictions in the *massive scheme*, in which only the u, d and s are the active flavours in the proton and the photon, while charm and beauty are produced dynamically in the hard scattering. The H1 data agree with the ones from ZEUS, while the NLO predictions tend to underestimate the data.

1.7.2 Double Tagging

Events in which both heavy quarks are tagged can be used to test higher order QCD effects. Double tagging analyses at HERA are done by selecting events with a D^* meson and a muon in the final state, or with two muons.

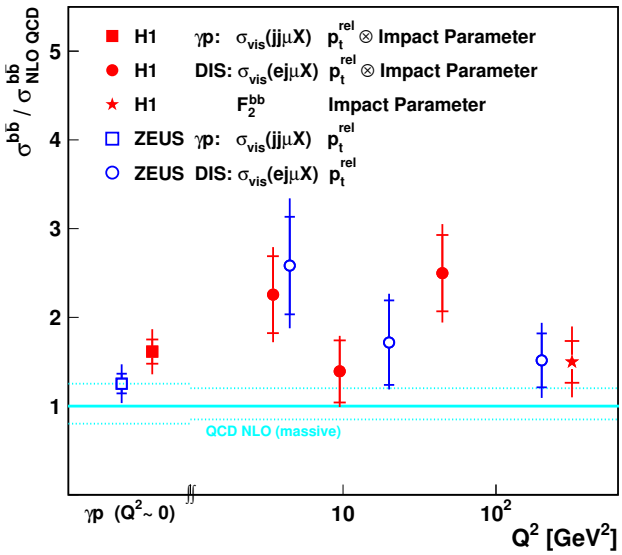

Figure 1.23: *Ratio of beauty production cross section measurements at HERA to NLO QCD predictions (massive scheme) [42], performed in photoproduction (γp) events with the FMNR program, and in DIS with HVQDIS.*

1.7.2.1 Double Tagging with a D^* Meson and a Muon

The subject of this thesis is the study of charm and beauty production in $D^*\mu$ events, where the D^* meson is reconstructed in the 'golden' decay channel $D^* \to K\pi\pi_s$. The various production configurations are shown in figure 1.24.

The branching ratios related to the D^* meson which are used in this analysis are presented in table 1.3 (from [15]).

Branching ratio [%]	
Meson	**Quarks**
$\mathcal{BR}(D^{*+} \to D^0\pi_s^+) = 67.7 \pm 0.5$	$f(c \to D^*X) = 25.5 \pm 1.5 \pm 0.8$
$\mathcal{BR}(D^0 \to K^-\pi^+) = 3.80 \pm 0.07$	$f(b \to D^*X) = 17.3 \pm 2.0$
$\mathcal{BR}(D^{*+} \to K^-\pi^+\pi_s^+) = 2.57 \pm 0.05$	$f(b \to D^*\mu X) = 2.75 \pm 0.19$

Table 1.3: *Branching ratios of the chosen D^* decay channel and of the single quarks to D^* [15]. If two errors are quoted, the first is statistical and the second is systematic.*

In case of $c\bar{c}$ production, the D^* and the muon have opposite electric charges and come from different quarks, as depicted in figure 1.24, d).

For beauty, there are several possibilities (see table 1.4), since the muon may come from a B or D meson decay. If the D^* and the muon come from different b quarks, they are in opposite hemispheres, and may have either opposite charge (figure 1.24, a), or the same charge (figure 1.24, b). If the D^* and the muon come from the same b quark, they have

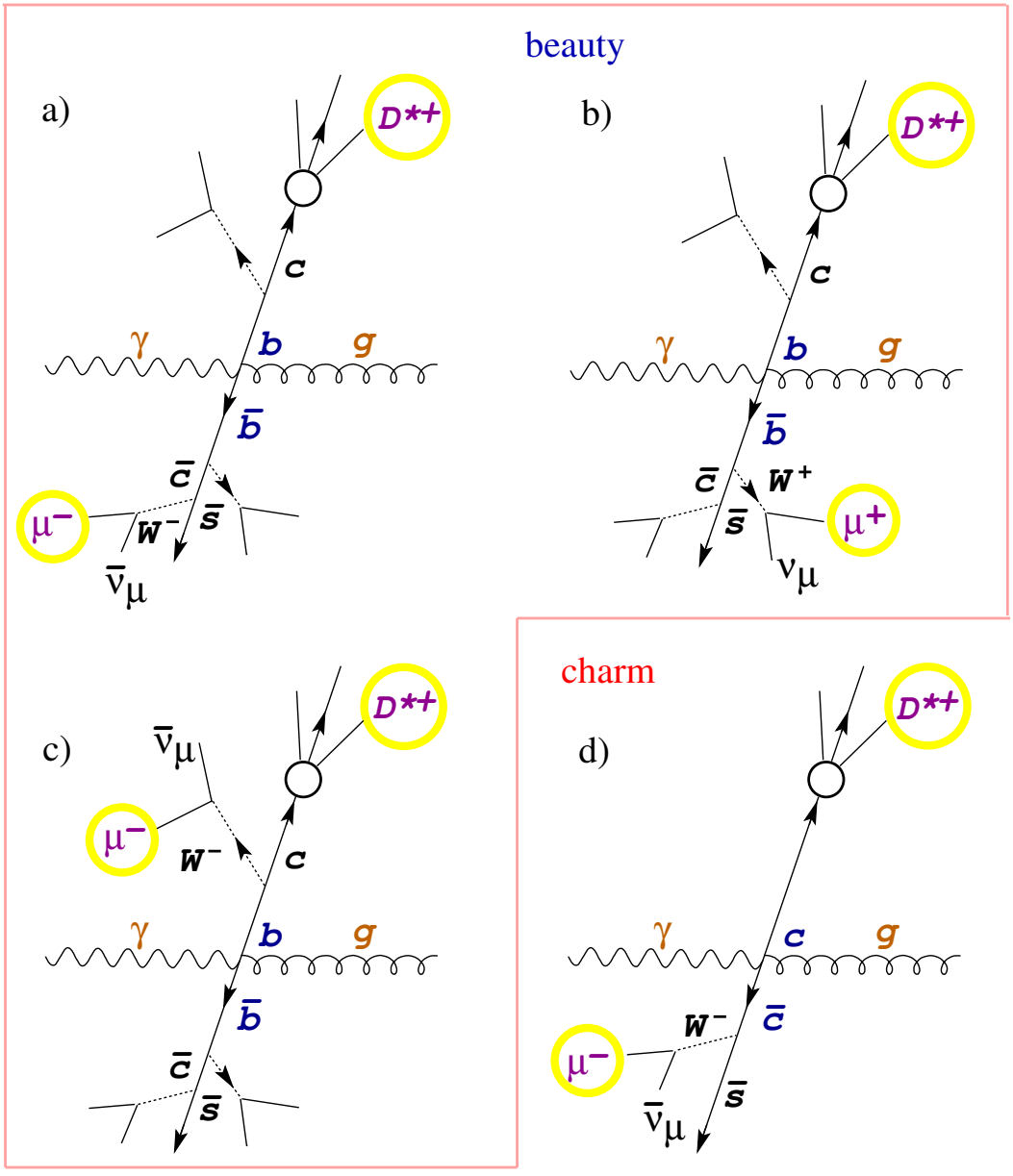

Figure 1.24: *Possible configurations with a D^* and a muon in case of beauty (a, b, c), and charm production (d). Similar configurations are present for the charge conjugate states.*

opposite charges and lie in the same hemisphere (figure 1.24, c). Note that figure 1.24 contains no $B^0 - \overline{B}^0$ mixing contributions (for details, see appendix B).

Decay	**Global branching ratios [%]**	
$\mu^- \leftarrow \bar{c}c \rightarrow D^{*+}$	$2 \cdot f(c \rightarrow D^*) \cdot \mathcal{BR}(D^* \rightarrow K\pi\pi_s) \cdot$ $\cdot \mathcal{BR}(c \rightarrow \mu)$	0.1075 ± 0.0093
$\mu \leftarrow \bar{b}b \rightarrow c(\sim D^*)$	$2 \cdot f(b \rightarrow D^*) \cdot \mathcal{BR}(D^* \rightarrow K\pi\pi_s) \cdot$ $\cdot \mathcal{BR}(b \rightarrow \mu)$	0.3262 ± 0.0228
$\mu \leftarrow X \leftarrow \bar{b}b \rightarrow c(\sim D^*)$	$2 \cdot f(b \rightarrow D^*) \cdot \mathcal{BR}(D^* \rightarrow K\pi\pi_s) \cdot$ $\cdot \mathcal{BR}(b \rightarrow X \rightarrow \mu)$	
$X \leftarrow \bar{b}b \rightarrow \mu^- c(\sim D^{*+})$	$2 \cdot \mathcal{BR}(b \rightarrow D^* \mu X) \cdot \mathcal{BR}(D^* \rightarrow K\pi\pi_s)$	

Table 1.4: *Global branching ratios for charm and beauty production [15]. The factor 2 accounts for charge conjugate states. In case of beauty, the given branching ratio is the sum of the branching ratios of the different beauty scenarios.*

Branching ratio [%]	
Charm	**Beauty**
$\mathcal{BR}(c \rightarrow \mu) = 8.2 \pm 0.5$	$\mathcal{BR}(b \rightarrow \mu) = 10.95^{+0.29}_{-0.25}$
	$\mathcal{BR}(b \rightarrow c \rightarrow \mu) = 8.02 \pm 0.19$
	$\mathcal{BR}(b \rightarrow \bar{c} \rightarrow \mu) = 1.6^{+0.4}_{-0.5}$
	$\mathcal{BR}(b \rightarrow \tau) = (2.48 \pm 0.26)$
	$\mathcal{BR}(\tau \rightarrow \mu) = (17.36 \pm 0.05)$
	$\mathcal{BR}(b \rightarrow \tau \rightarrow \mu) = 0.43 \pm 0.05$

Table 1.5: *Muon branching ratios [15].*

If the muon is obtained directly from the heavy hadrons decay, one speaks about *direct decay* (i.e. *prompt muons*). If intermediary states are present, one refers to *cascade decays*. With the values given in table 1.5, the total branching ratio to produce a muon in cascade decays in this analysis is:

$$\begin{aligned} \mathcal{BR}(b \rightarrow X \rightarrow \mu) &= \mathcal{BR}(b \rightarrow c \rightarrow \mu) + \mathcal{BR}(b \rightarrow \bar{c} \rightarrow \mu) + \mathcal{BR}(b \rightarrow \tau \rightarrow \mu) \\ &= (10.05 \pm 0.24)\%, \end{aligned} \tag{1.33}$$

where $\mathcal{BR}(b \rightarrow \tau \rightarrow \mu) = \mathcal{BR}(b \rightarrow \tau) \cdot \mathcal{BR}(\tau \rightarrow \mu) = (0.43 \pm 0.05)\%$.

In figure 1.25, the correlations of the D^* mesons and of the muon with the initial charm quark, as predicted by PYTHIA Monte Carlo, are shown. Due to its larger mass compared with the muon, the D^* meson is better correlated in η and ϕ with the initial heavy quark.

The D^* meson provides a clean heavy flavour tag, but the small branching ratio for the chosen D^* decay channel results in a low statistics sample. Therefore, by selecting a D^* meson and a muon a compromise is made between a large sample, but with low purity, on the muon side, and a small sample, but with high purity, on the D^* side.

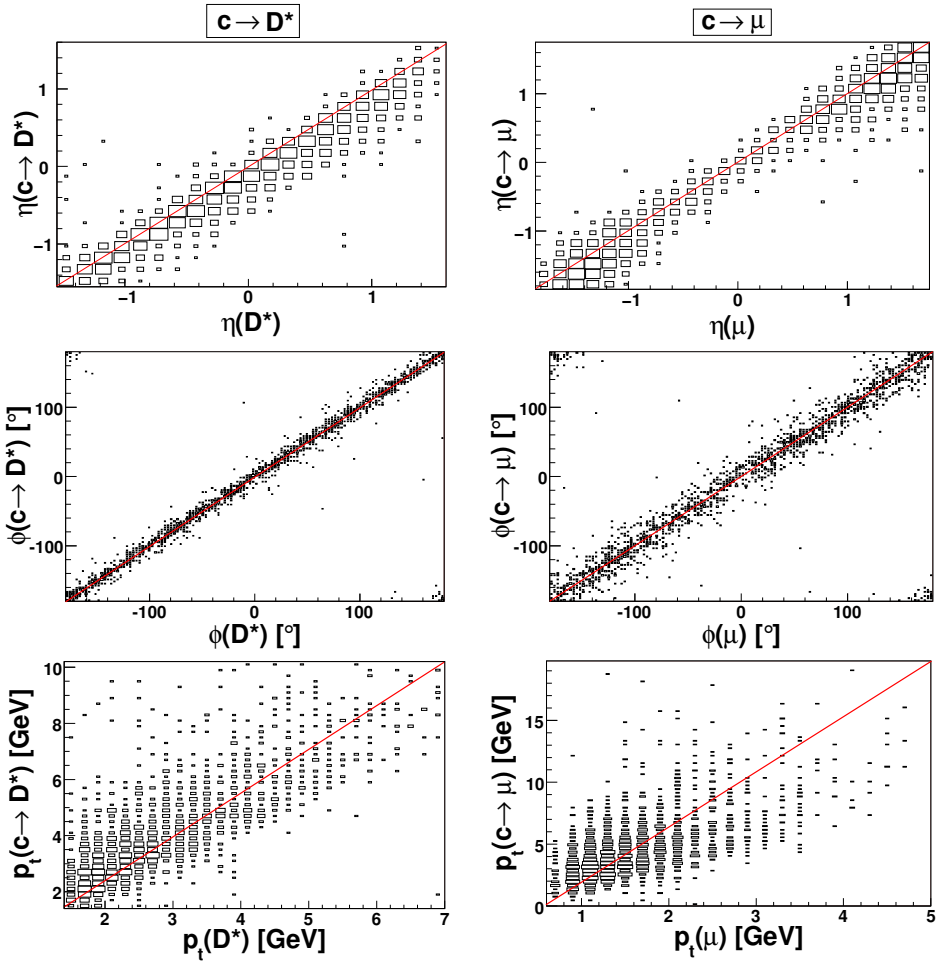

Figure 1.25: *Correlations of the D^* meson (left) and of the muon (right) with the corresponding charm quark, as predicted by the PYTHIA Monte Carlo, direct component only. The distributions of pseudorapidity η, azimuthal angle ϕ and of transverse momentum p_t are shown.*

1.7.2.2 $D^*\mu$ Variables

To approximate the heavy quark pair kinematics, the following $D^*\mu$ quantities are defined:

- transverse momentum:

$$p_t(D^*\mu) = \sqrt{\left(p_x^{D^*} + p_x^{\mu}\right)^2 + \left(p_y^{D^*} + p_y^{\mu}\right)^2}; \tag{1.34}$$

- azimuthal angular difference:

$$\Delta\phi(D^*\mu) = \begin{cases} |\phi_{D^*} - \phi_\mu| & \text{if } |\phi_{D^*} - \phi_\mu| \leq 180°, \\ 360° - |\phi_{D^*} - \phi_\mu| & \text{if } |\phi_{D^*} - \phi_\mu| > 180°; \end{cases}$$

- invariant mass:

$$M(D^*\mu) = \sqrt{m_{D^*}^2 + m_\mu^2 + 2 \cdot (E_{D^*} \cdot E_\mu - \vec{p}_{D^*} \cdot \vec{p}_\mu)}; \tag{1.35}$$

- rapidity:

$$y = \frac{1}{2} \cdot \ln \frac{(E_{D^*} + E_\mu) + \left(p_z^{D^*} + p_z^\mu\right)}{(E_{D^*} + E_\mu) - (p_z^{D^*} + p_z^\mu)}; \tag{1.36}$$

- pseudorapidity:

$$\eta(D^*\mu) = -\frac{1}{2} \cdot \ln\left(\frac{1-\cos\theta}{1+\cos\theta}\right) = -\ln(\tan(\theta/2)), \tag{1.37}$$

with $\cos\theta = (p_z^{D^*} + p_z^\mu)/p^{D^*\mu}$, $p^{D^*\mu} = \sqrt{(p_x^{D^*} + p_x^\mu)^2 + \left(p_y^{D^*} + p_y^\mu\right)^2 + (p_z^{D^*} + p_z^\mu)^2}$.

Although these quantities are not independent, they help in understanding different properties of the $D^*\mu$ pair.

In the LO picture, the heavy quarks are produced via the **boson-gluon fusion** (figure 1.11, a). In the photon-gluon rest frame, they are produced in a back-to-back configuration (see figure 1.26, a). If the interacting partons do not acquire a transverse momentum k_t with respect to the photon and the proton, in photoproduction the laboratory frame and the photon-gluon rest frame differ only by a longitudinal Lorentz boost along the z-axis. Since the boost does not change the topology in the plane transverse to the boost axis, the heavy quarks are still back-to-back in the laboratory frame. Their azimuthal angular difference $\Delta\phi(q\bar{q})$ is expected to peak at 180^0 and their combined transverse momentum $p_t(q\bar{q})$ to vanish.

However, gluon radiation (figure 1.11, b) and a possible initial transverse momentum of the incoming partons lead to $\Delta\phi(q\bar{q}) < 180^0$ and $p_t(q\bar{q}) > 0$ GeV. These $q\bar{q}$ variables which are sensitive to NLO effects are approximated by combined $D^*\mu$ quantities: $p_t(D^*\mu)$ and $\Delta\phi(D^*\mu)$

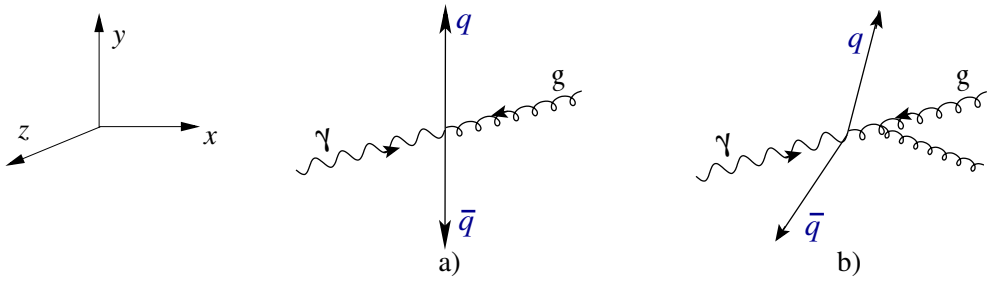

Figure 1.26: *Representation of heavy quarks in photon-gluon rest frame: a) LO; b) gluon radiation in NLO picture. The axes of the coordinate system are also shown.*

The mass $M(D^*\mu)$ is used as an approximation of the heavy quark pair invariant mass $M(q\bar{q})$ which, in LO, corresponds to the centre-of-mass energy of the initial partons. Similar, the rapidity $y(D^*\mu)$ approximates the variable $y(q\bar{q})$, which gives the longitudinal direction of the heavy quark pair. The invariant mass $M(q\bar{q})$ and the inelasticity $y(q\bar{q})$ are related to the fraction of proton momentum carried by the gluon x_g [45]:

$$x_g = \frac{M(q\bar{q})}{\sqrt{s}} \cdot \exp\left(-y(q\bar{q})\right). \tag{1.38}$$

Analogous to the pseudorapidity $\eta(q\bar{q})$, the $D^*\mu$ quantity $\eta(D^*\mu)$ is defined. This variable, together with $p_t(D^*\mu)$, will be used to study the QCD dynamics implemented in PYTHIA and CASCADE Monte Carlo (see chapter 8).

The $D^*\mu$ variables distributions will be presented and discussed in chapter 7.

1.7.2.3 Definition of the $D^*\mu$ Correlation Regions

Based on the $D^*\mu$ correlations between their electric charges, $Q(D^*)$ and $Q(\mu)$, and their azimuthal angular difference $\Delta\phi(D^*\mu)$, four correlation regions are defined in table 1.6 (see also figure 1.24). The distributions of events in these regions will be later used to extract the fractions of charm and beauty events in the data (see chapter 5).

Region	$\Delta\phi$	Charge relation
1	$< 90°$	$Q(D^*) = Q(\mu)$
2	$\geq 90°$	$Q(D^*) = Q(\mu)$
3	$< 90°$	$Q(D^*) \neq Q(\mu)$
4	$\geq 90°$	$Q(D^*) \neq Q(\mu)$

Table 1.6: *Definitions of the $D^*\mu$ charge and angle correlations regions.*

1.7.2.4 Tagging with Two Leptons

Beauty production in events with two muons in the final state has been measured by ZEUS [46]. Similar to the $D^*\mu$ double tagging, the muons may come from the same parent b hadron or from different beauty quarks of a $b\bar{b}$ pair.

The advantages compared to the $D^*\mu$ analysis are: bigger statistics due to larger branching ratios, wider rapidity coverage and low charm background, which allows significant measurement of $b\bar{b}$ correlations. In addition, since the background is suppressed, lower transverse momentum cuts can be applied, compared to the single tagging of beauty with leptons, so that this method is sensitive to b quark production towards $p_t = 0$.

Because the difference between like- and unlike-sign dimuon contributions (figure 1.27) is essentially free from fake muon background, it can be used to measure the beauty contribution, if the other sources of background are known. The background from $c\bar{c}$ production, where both charm quarks decay into a muon, has not been measured directly in the dimuon data, but normalised to the charm contribution to the similar $D^*\mu$ sample. Background muons from quarkonium decays and Bethe-Heitler processes[23] are not directly accompanied by hadronic activity, so they have an isolated muon signature. To reduce this type of background, a cut on the quadratic sum of the total energy deposited in a cone around each muon direction was applied.

[23] In the Bethe-Heitler process, also known as 'photon-photon' process, the proton radiates a photon which interacts with the photon from the incoming electron or positron, generating an $l\bar{l}$ pair, where l stands for lepton (e or μ).

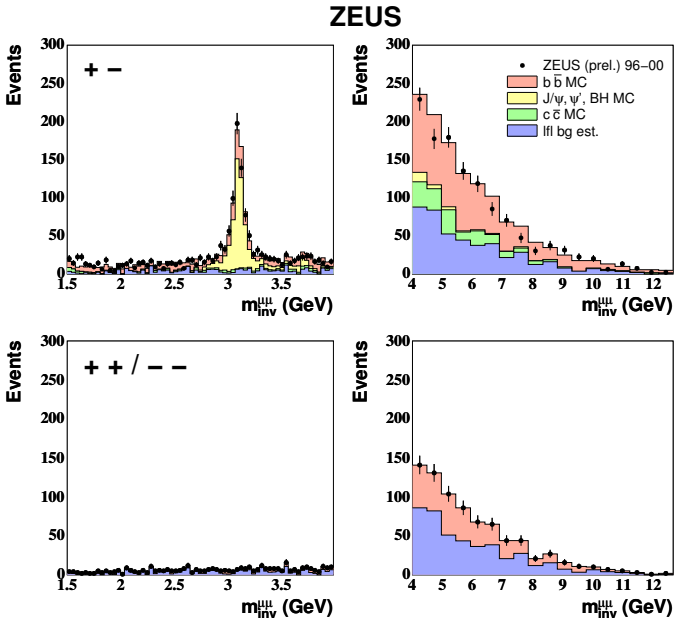

Figure 1.27: *Dimuon invariant mass $m^{\mu\mu}_{inv}$ distributions of unlike- (like-) sign dimuon pair in the low mass (top) and high mass sample (bottom) (from [46]). In the low mass region, the J/ψ production, not originating from b decays, contributes significantly, while the high mass region is strongly beauty enriched.*

A summary of the beauty production measurements at the HERA collider, from double tag analyses, is presented in figure 1.28. It can be seen that all H1 and ZEUS data exceed the NLO predictions.

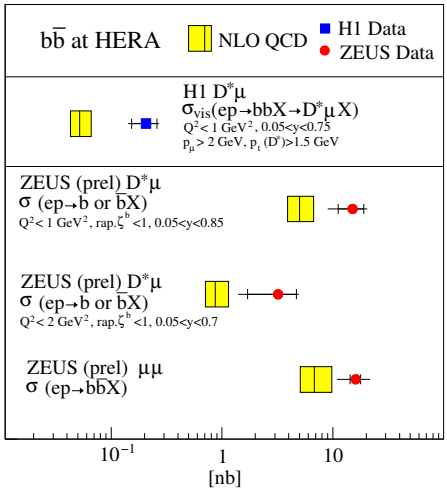

Figure 1.28: *Comparison of beauty production cross section from **double tag** analyses to NLO prediction (from [43]). From top to bottom, the results are shown from: H1 $D^*\mu$ [47], ZEUS $D^*\mu$ in photoproduction and DIS [48], and ZEUS $\mu\mu$ [46].*

More details about the heavy quark tagging methods applied at HERA can be found in the reviews [43] and [40].

Chapter 2

The H1 Detector at HERA

2.1 The H1 Detector

The data used in this analysis were taken with the H1 experiment at the Hadron Electron Ring Accelerator (HERA) at the Deutsches Elektronen-Synchrotron (DESY) in Hamburg, Germany.

HERA is the first lepton-proton collider in the world. It is a double ring accelerator complex. Its most important operating parameters for the year 2006 are shown in table 2.1.

Parameter	**Unit**	**Electrons**	**Protons**
Beam energy	GeV	27.5	920
Average beam currents	mA	22	79
Particles per bunch	10^{10}	≤ 3.68	≤ 8.75
Number of filled bunches		156	150
Bunch length	cm	0.9	12
Beam lifetime in collision	h	10-15	>200
Longitudinal polarisation	%	30-45	-
Circumference	km	6.3	
Time between bunch crossings	ns	96	
Peak luminosity	10^{31} cm^{-2} s^{-1}	3-5	
Average luminosity	pb^{-1} d^{-1}	1-2.5	

Table 2.1: *The HERA parameters for the year 2006 [49].*

The beams are brought into collision at two interaction points: in the hall North, where the H1 experiment is located, and in the hall South, at the ZEUS experiment. In the East, there is a fixed target experiment, HERMES, which uses the longitudinally polarised electron beams to study the spin structure of the nucleons. Until 2003, the West hall was occupied by the HERA-B experiment, which was designed to measure CP violation in

decays of B mesons.

The H1 detector is a general purpose detector with nearly full solid angle coverage. Because of the large momentum imbalance between the lepton and the proton beam, the centre-of-mass for the ep collisions is boosted along the proton direction. As a consequence, the detector is asymmetric and highly segmented in this direction. A longitudinal view of the H1 detector can be seen in figure 2.1.

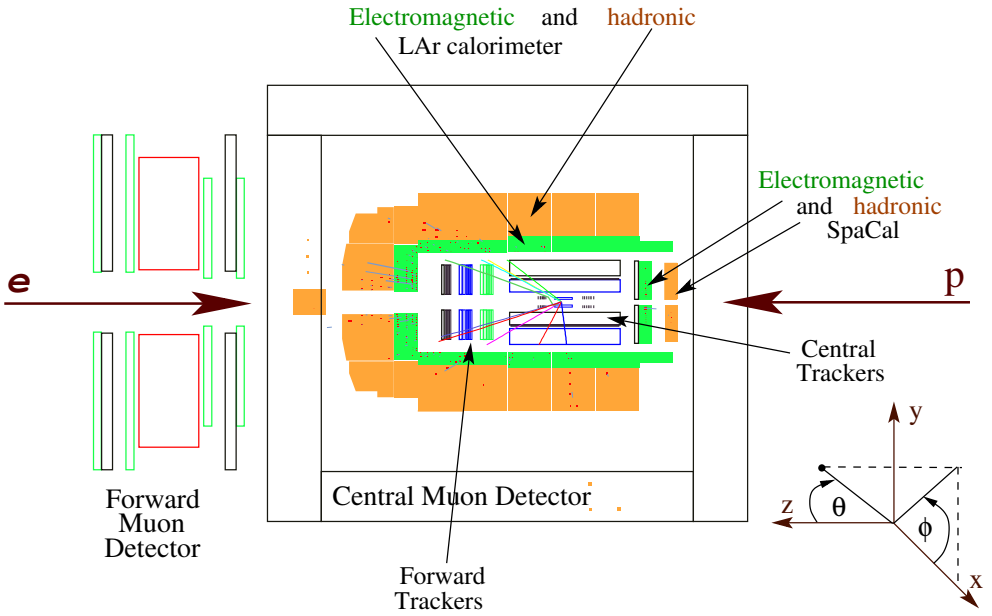

Figure 2.1: *The H1 detector (longitudinal view) and its coordinate system.*

The H1 coordinate system is right-handed, with the centre at the nominal interaction point. The z-axis points to direction of the incoming proton (i.e. *forward* direction, whereas the electron direction is referred to as the *backward* direction). The polar angle θ is defined with respect to the z-axis. The azimuthal angle ϕ is the angle with respect to the x-axis in the xy-plane.

A detailed description of the detector can be found elsewhere [50]. Here only the components relevant for this analysis are presented.

2.2 Central Tracking System

A radial view of the H1 central tracking detectors is shown in figure 2.2. The innermost detector is the *Central Silicon Tracker* (CST) which consists of two barrel layers of double sided silicon detectors. Resolutions of 12 μm in the $r-\phi$ plane and of 22 μm in z [51] allow the reconstruction of secondary vertices from decay of particles with a few hundreds μm decay lengths.

The CST is surrounded by two jet chambers, *Central Jet Chamber* 1 and 2, CJC1 and CJC2.

Jet chambers owe their name to the fact that they achieve optimal two-track resolution[1] as needed for jet reconstruction [52]. The H1 jet chambers have the sense wires parallel to the beam axis. The actual positions of the sense wires are offset from the nominal wire plane by $\pm$ 150 μm due to electrical forces, such that position is well defined and track ambiguities, due for example to wrong track segments, can be resolved.

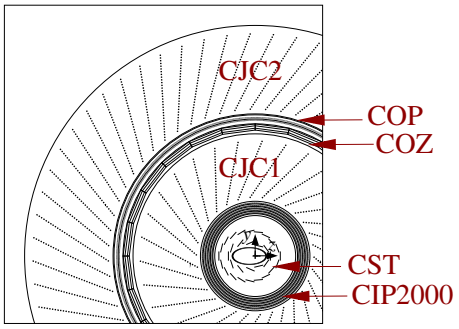

Figure 2.2: *Radial view of the H1 central trackers.*

The z-coordinate is determined with the *charge division method* with a precision of 2.2 cm (see table 2.2). The resolution in the z-coordinate measurement is improved by two orders of magnitude (see table 2.2) with the help of two z-chambers: CIZ (*Central Inner Z-chambers*), situated inside the CJC1, and COZ (*Central Outer Z-chambers*), between CJC1 and CJC2.

In order to deliver a fast timing and to get information about the position of the event vertex on the z-axis, two chambers were built: CIP (*Central Inner Proportional* chamber) and COP (*Central Outer Proportional* chamber). These are multi-wire chambers which consist of two layers of wires and pad cathodes segmented in z and azimuthal angle ϕ.

To improve the z-vertex triggering and the rejection of background events, in 2001, during the luminosity upgrade of the HERA detector, the CIP2000 detector [53] was introduced. It is a five layer cylindrical multi-wire proportional chamber with cathode readout, which replaced the CIP (two planes) and the CIZ.

The H1 drift chambers signals are digitised and processed online in order to determine the relevant parameters of the signal pulse (i.e. the drift time and the pulse integrals).

Parameter	**Unit**	CJC1	CJC2	CIZ	COZ
Inner radius	mm	203.0	530.0	173.5	460.0
Outer radius	mm	451.0	844.0	200.0	485.0
$\sigma_{r\phi}$	mm	0.170		28	58
σ_z	mm	22		0.26	0.20
Polar angular range		$11^\circ \leq \theta \leq 169^\circ$			

Table 2.2: *Parameters of the H1 central tracking chambers [50].*

[1] If two particles traverse the detector close to each other, one needs to distinguish between the hits forming the tracks of these particles.

2.3 Calorimeters

Calorimeters are devices to measure the total energy deposited by particles. For high energy electrons, the Bremsstrahlung[2] and pair production processes lead to the production of *electromagnetic showers*. The H1 detector contains a liquid argon (LAr) and a lead scintillating (SpaCal) calorimeter. Their energy resolutions and polar angle coverage are presented in table 2.3.

Calorimeter	**Energy resolution**		**Polar angle coverage**
	Electromagnetic particles	Hadrons	
LAr	$\frac{\sigma(E)}{E} \propto \frac{12\,\%}{\sqrt{E/\mathrm{GeV}}} \oplus 1\%$	$\frac{\sigma(E)}{E} \propto \frac{50\,\%}{\sqrt{E/\mathrm{GeV}}} \oplus 2\,\%$	$4^\circ < \theta < 154^\circ$
SpaCal	$\frac{\sigma(E)}{E} \propto \frac{7.5\,\%}{\sqrt{E/\mathrm{GeV}}} \oplus 1\,\%$	$\frac{\sigma(E)}{E} \propto \frac{30\,\%}{\sqrt{E/\mathrm{GeV}}} \oplus 7\,\%$	$153^\circ < \theta < 178^\circ$

Table 2.3: *Energy resolutions and polar angle coverage of the H1 calorimeters [50]. The* $\oplus$ *sign represents addition in quadrature with subsequent square root.*

The liquid argon (LAr) calorimeter is a **sampling** calorimeter. The liquid argon is the active medium (which generates the signal) while the passive medium (steel and lead) functions as an absorber. The LAr calorimeter response for electrons is about 30% higher than for hadrons (**non-compensating** calorimeter). Therefore weighting techniques are applied to equalise the response to the electromagnetic and hadronic components of a shower. The LAr calorimeter covers the regime $Q^2 \gtrsim 100\ \mathrm{GeV}^2$ for the detection of the scattered electron in deep-inelastic scattering processes.

For a precise measurement of the electron quantities in backward direction, a lead-scintillating calorimeter (SpaCal) is used. The forward region is covered by a plug calorimeter, consisting of copper plates interleaved with layers of large area silicon detectors.

To measure the hadronic energy leaking out of LAr and SpaCal calorimeters, tubular wire chambers called *limited streamer tubes* are integrated in the Central Muon Detector, forming the Tail Catcher.

2.4 Central Muon Detector

Muons are *minimum ionising particles*, i.e. they do not interact strongly, and are able to penetrate large amounts of absorbers with minor energy losses and small deviations of the direction. The muon chambers are filled with gas mixture of CO_2, argon and iso-butane, such that muon momentum and direction are measured via ionisation in gas.

The **Central Muon Detector** (CMD) is divided into four sub-detectors (see table 2.4). Every sub-detector consists of 16 modules. A module is formed out of ten iron layers, each 7.5 cm thick, interleaved with 10 layers of limited streamer tubes, as shown in figure 2.3.

[2]The term *Bremsstrahlung* comes from German and refers to radiation emission by a charged particle decelerating in the field of atomic nuclei [52].

Sub-detector	Polar angular coverage
Forward end-cap	$5° \leq \theta \leq 35°$
Backward end-cap	$130° \leq \theta \leq 175°$
Forward and backward barrel	$35° \leq \theta \leq 135°$

Table 2.4: *The H1 central muon sub-detectors and their polar angular coverage.*

In front and behind of the instrumented iron there are **muon boxes** which contain three layers: two layers with strips perpendicular to the wire direction and a third layer with pads. The wire- and strip- signals are read out digitally, so that three dimensional points can be determined. Table 2.5 contains the achieved resolutions.

Together with the trackers, the muon system is used to trigger on muons from heavy quarks decay. The muon triggers will be needed in this analysis, thus the muon trigger sectors are shown in figure 2.4.

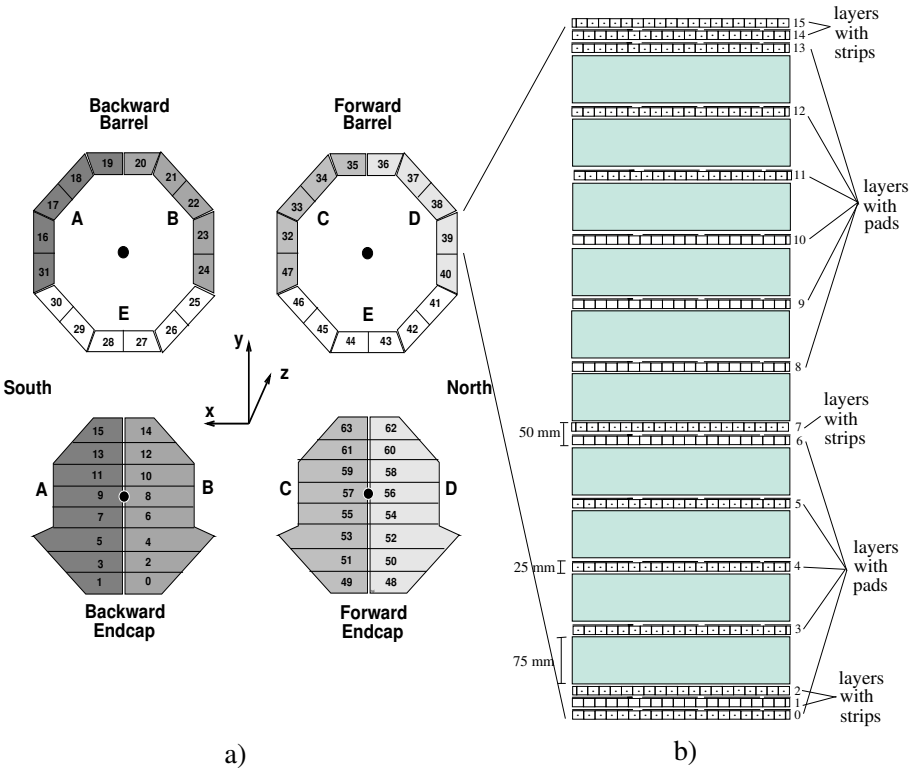

Figure 2.3: *Central Muon Detector: a) muon modules, b) cross section.*

2.4.1 Muon Reconstruction in the LAr Calorimeter

When a muon reaches the calorimeter layers, it is identified by the small energy deposition (with a minimum of the order of 10 MeV per traversed centimetres) close to the muon path of flight. The muon track is extrapolated in the LAr calorimeter and two cylinders of

Resolution	Wire	Strip	Pads
σ_x [mm]	3-4	10-15	100

Table 2.5: *Spatial resolution of the muon wires, strips and pads.*

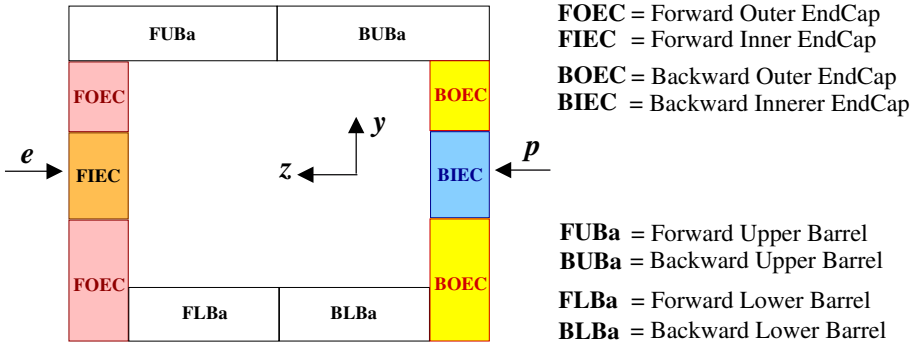

Figure 2.4: *The trigger sectors of the Central Muon Detector [54].*

radius 15 cm, which contains all signals of a muon, and 30 cm (contains on average 90% of a hadronic shower) are defined.

Q_μ^{Cal}	Meaning
0	no calorimeter muons
1	weak quality
2	good quality
3	very good quality

Table 2.6: *Calorimeter muon qualities Q_μ^{Cal} and their meaning.*

To describe the agreement of the observed energy depositions in the two cylinders with the pattern expected for a minimum ionising particle, a set of discriminating variables, like electromagnetic calorimeter energy and track length, are used. Cuts[3] implemented in fuzzy logic are applied to the above mentioned variables to discriminate muons against light hadrons, like kaons and pions, which produce hadronic showers and deposit most of their energy in the calorimeter. Several calorimeter muons qualities are defined, as in table 2.6 (for details, see [55]).

2.4.2 Muon Reconstruction in Instrumented Iron System

If it has enough energy, the muon might travel further and enter the instrumented iron system. The average energy loss in any of the ten iron plates (see section 2.4) is approximately 90 MeV. Since the muons observed with the H1 detector have energies of a few GeV, there is a big probability for the muon to leave the detector. The strips and wires of the muon detector determine the spatial position of the track. A pattern recognition program [56] is run separately for the barrel and the end-caps, such that, for a muon passing both of them, two track segments can be found. In a first step, neighbouring hits are combined into

[3]Since the muon energy deposition depends on the muon momenta and polar angle, the cuts depend on these variables.

super-hits. This way the hit resolution is reduced to 2 cm for wires and 4 cm for strips, and the combinatorial background is significantly reduced. Further, the muon modules which contain a hit are grouped into *sectors*. For each sector, a local two-dimensional coordinate system is defined. A two-dimensional histogramming method (conformal mapping) is used to parametrise each hit into a straight line or a sine curve. All hits belonging to the same track correspond to curves which intersect in one point in the parameter space. These crossings are looked for by the program. Once one is found, it is assigned to an *association* and removed from the sample. The search continues until no other crossing is found.

The resulting associations are subject to quality checks (for example, a minimal number of layers with hits is requested.) In the next step, muon tracks with at least 3 wire hits and one strip or pad signal, and with the distance between the first and the last wire hit greater than 15 cm are fitted. Due to variation of the magnetic field in the return yoke, the curvature of a track is not constant. For fitting purposes, the instrumented iron return yoke is divided into 21 regions (10 iron plates, 9 gaps between them, and the 2 muon boxes). For every region, a mean magnetic field is chosen and parabolas are fitted to wire and strips:

$$y_i(x_i) = a_i + b_i \cdot x_i + c_i \cdot x_i^2,\ i = 1 \ldots 21; \tag{2.1}$$

with $y_i,\ x_i$ the coordinates of a hit in the plane, and $a_i,\ b_i,\ c_i$ the fit parameters. The parameter c_i is given by:

$$c_i = const \cdot \frac{Q}{P_0} \cdot B_i \cdot \frac{P_0}{P_0 - dE_i}, \tag{2.2}$$

where Q is electric charge, P_0 is momentum of the particle at the beginning of the track, B_i is the magnetic field in region i and dE_i is the energy loss in plate i. The accuracy for Q/P_0 is of the order of 30% in the barrel and worse in the end-caps.

The fit parameters are obtained by minimising the χ^2 function:

$$\chi^2 = \sum_{wires} \frac{1}{\sigma^2_{wires}} \cdot [y_{wires}(x_{fit} - y_{fit})]^2 + \sum_{strips} \frac{1}{\sigma^2_{strips}} \cdot [y_{strips}(x_{fit} - y_{fit})]^2\,, \tag{2.3}$$

where $\sigma_{wires} = 0.4$ cm and $\sigma_{strips} = 1.0$ cm are the resolutions.

The track parameters are: the coordinate of the first measured point, track direction reconstructed in this point and the ratio of charge/momentum, the latter being reconstructed in the barrel.

2.4.3 Muon Track Linking

To identify an iron muon through its penetrating capability, its track is linked to a track measured in the central trackers (*inner track*) if:

$$|\theta_{inner\ track} - \theta_{iron\ track}| < 0.2 \text{ rad } (12^\circ), \tag{2.4}$$

$$-\frac{\pi}{2} < (\phi_{inner\ track} - \phi_{iron\ track}) < 0.2 \text{ rad if } \kappa > 0 \tag{2.5}$$

or

$$-\frac{\pi}{2} < (\phi_{iron\ track} - \phi_{inner\ track}) < 0.2 \text{ rad if } \kappa > 0, \tag{2.6}$$

$$p_{inner\ track} > 1\ \text{GeV}, \tag{2.7}$$

where κ is the curvature of the track. The last cut ensures that the inner track is able to reach the iron system.

Because a muon track may be linked to more than one inner track, a *link probability* is defined according to a χ^2 distribution:

$$\chi^2 = \left(\vec{X}_E - \vec{X}_I\right)^T \cdot V^{-1} \cdot \left(\vec{X}_E - \vec{X}_I\right), \tag{2.8}$$

where $\vec{X}_E$ ($\vec{X}_I$) stands for the extrapolated (inner) track parameters, and V is the covariance matrix. By selecting only the track with the highest link probability, double counting of muons is avoided.

The distribution of the muon link probability predicted by PYTHIA Monte Carlo, direct component[4], is shown in figure 2.5.

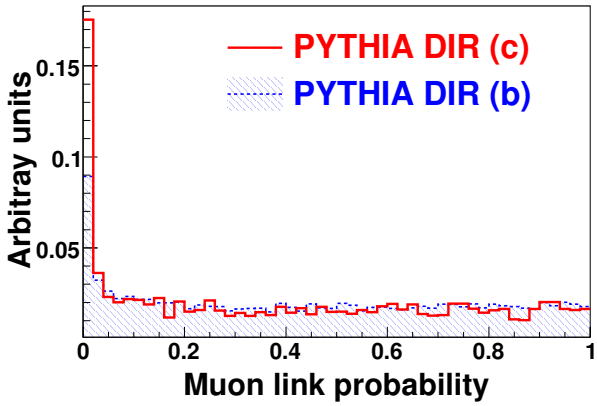

Figure 2.5: *Distribution of the muon link probability as predicted by PYTHIA charm and beauty Monte Carlo, direct component only.*

2.5 Luminosity System

The luminosity delivered by the HERA machine is determined from the rate of the Bethe-Heitler process, $ep \to ep\gamma$ [57]. This process is independent of the inner structure of the proton and its rate can be scalculated in quantum electrodynamics (QED).

The luminosity system consists of two crystal scintillating calorimeters, one at $z = -33$ m, which detects the scattered electron, and one at $z = -103$ m, for detecting the outgoing photon.

The instantaneous luminosity [50] is obtained as:

$$\mathcal{L} = \frac{R_{tot} - (I_{tot}/I_0) \cdot R_0}{\sigma_{vis}}, \tag{2.9}$$

[4] A similar behaviour is observed for the excitation component.

where R_{tot} is the total rate of Bremsstrahlung, R_0 is the rate of Bremsstrahlung events coming from non-colliding bunches, referred to as *pilot bunches*[5], I_{tot} and I_0 are the corresponding electron currents, and σ_{vis} is the visible part of the Bethe-Heitler cross section where the rate is measured.

The sources of systematic errors in the luminosity determination are well known for the different years of data taking. Here only the errors for the years used in the present analysis are presented (see table 2.7). The accuracy of the luminosity is mainly driven by the precision of the photon energy measurement and of σ_{BH}. Another contribution to the systematic error in the luminosity measurement comes from the dependence of the system acceptance on possible variations of the electron beam angle at the interaction point [50]. Overlapping $ep \to e\gamma p$ events in the same bunch crossing (pileup events) contribute with up to 0.3% in the systematics. Around 0.1% error is due to triggering. A small fraction of protons may escape from the main bunches and form parasitic satellites (p-satellites), separated by a few nsec. Because the contribution of the proton satellites to the total ep luminosity is only a few per cent [58], the systematic error due to this effect is less than 1%.

Source of systematic uncertainties	**Systematic uncertainty $\delta\mathcal{L}$ [%]**				
	HERA I	**2003-2004**	**2005**	**2006 e^-**	**2006 e^+**
Theoretical σ_{BH}	0.5				
Energy scale	0.7-1.1	1.5	0.4	0.1	0.1
PD acceptance	0.4-1.2	0.5	1.2	1.0	0.9
Pileup	0.3	0.2	0.2	0.2	0.1
Trigger	0.1	0.1	0.1	0.1	0.1
p-satellites	0.4-1.0	0.25-0.75	0.40	0.55	0.45
Total $\delta\mathcal{L}$ [%]	**1.2-1.8**	**1.8**	**1.5**	**1.3**	**1.2**

Table 2.7: *Sources of systematic errors for the H1 luminosity measurement (from [59]). σ_{BH} is the cross section of the Bethe-Heitler process. 'PD' stands for photon detector.*

2.6 Trigger System

The tasks of a trigger system are to initiate the readout of the detector, to distinguish between interesting physics events and background[6], and to reduce the rate at which the events are stored. Every 96 nsec electron and proton bunches cross at H1. This corresponds to a rate of 10.4 MHz. The resulting particles may interact with the detector material, generating signals which are afterwards stored in pipelines.

[5]The pilot bunches are used for studying the background processes, from which the electron interaction with residual gas in the beam pipe is the main source that effects the Bethe-Heitler rate determination.

[6]Interactions of the beam protons or electrons with the residual gas of the beam pipe, synchrotron radiation and cosmic rays are the main sources of background at HERA.

Trigger Level	Level 1	Level 2	Level 3	Level4
Decision time	2.3 μsec	20 μsec	$<$ 800 μsec	$\backsim$ 100 ms
Signals	Trackers, calori-, meters, muon system, time of flight	Level 1 with finer granularity	Fast Track Trigger	Power PC farm
Method	Programmable look-up tables and Boolean logic	Neuronal networks and topological processors	Software algorithms	Software algorithms
Rate	$\backsim$ 1 kHz	$\backsim$ 100-200 Hz	50 Hz	5 Hz
Actions	Stops the pipeline	Starts the readout	Starts event building	Data logging

Table 2.8: *The H1 trigger system.*

H1 uses a four-level trigger system (see table 2.8). The first and the second trigger level (L1 and L2) are synchronous to the HERA accelerator clock.

- **Level 1**: For each bunch crossing, L1 provides a trigger decision after 2.3 μsec, without causing dead time. General information about an event (i.e. z-position of the vertex, arrival time in the time of flight system, track properties) is encoded in threshold topological coincidences, forming *trigger elements*. These elements are further combined into 128 *raw subtriggers*, consisting of trigger and veto conditions.

  To account for beam conditions and to have a reasonable output rate, the L1 subtriggers may be prescaled[7]. After prescaling, the raw subtriggers decisions are called *actual subtriggers*.

  If L1 decides to keep the event, an L1Keep signal is generated. The pipelines are stopped and the dead time (in which subsequent events are lost) of the detector starts accumulating.

- **Level 2**: Topological and neuronal network triggers offer information at this level. The *L2 Neuronal Network* (L2NN) is composed of twelve networks which are trained to identify a specific *ep* event signature. The *L2 Topological Triggers* [60] use calorimeter and tracker information of level 1, but in finer granularity. An event is projected on a 16×16 $\theta - \phi$ plane. Such a projection is called a *topology*. The output of the neuronal networks of L2NN and the projections of the L2TT provide 26 trigger elements which are used to validate the L1 actual subtriggers.

  If L2 decides to keep an event, the detector readout starts. Otherwise, the trigger signals from the next bunch crossings are read in after a latency of several μsec.

- **Level 3**: To reduce the dead time, a third trigger level was implemented, but it was only used for physics analysis starting with the HERA II run period, as part of the *Fast Track Trigger* (FTT).

[7] A prescale of n means that only the n^{th} event which fulfils the subtrigger conditions will be accepted.

- **Level 4**: Is an asynchronous trigger level which performs a complete reconstruction and classification of the event. It is based on software algorithms running on a computer farm. The raw event data provided by sub-detectors is processed by the H1 reconstruction software, H1REC, which reconstructs the tracks in the drift chambers and in the muon detector, as well as the clusters in calorimeters. If an event was accepted by L4, an *L4 verified subtrigger* bit is set when the corresponding L1 raw bit is set. At this trigger level, the events are classified into **L4 classes**, such as heavy flavour and jet classes. If the events full-fill the 'hard scale' requirements (for example, high Q^2, high p_t), or if they pass one of the final state finders, they are assigned to physics classes. Otherwise, they are down-scaled by a weighting factor, the **L4 weight**.

In the following, the level 1 $DCRPh$ and the z-vertex triggers, which are going to be used in the analysis, are briefly described.

The $DCRPh$ trigger was designed to reject background events based on the distance of closest approach, DCA[8]. Tracks from background events have big DCA. More information about this trigger can be found in [61]. In 2005, the $DCRPh$ trigger was replaced by FTT (see section 2.6.1).

To reconstruct online the z-vertex position of the primary interaction, the z-vertex trigger was built. Four coincident pad signals (from CIP and COP, or CIP and first forward proportional chambers) which can be connected by a straight line pointing to the z-axis form a *ray*. A z-vertex histogram is defined around the nominal interaction region. The histogram has 16 equal bins and covers a region of about 88 cm. Each ray is assigned unambiguously to a bin of the histogram. The bin with the most entries is expected to contain the interaction vertex of the ep collision. For more details, see [62].

2.6.1 Fast Track Trigger

For a better selectivity of the exclusive final states containing heavy quarks, after the HERA upgrade in 2001, the H1 collaboration installed a three level based track trigger called FTT (*Fast Track Trigger*) [63]. FTT uses information from four groups of three layers of wires each (three inside CJC1 and one inside CJC2, as in figure 2.6) and it is able to reconstruct up to 48 tracks, which is the case for more than 98% of the events of interest.

FTT consists of two hardware levels (L1 and L2) and a third software level. At the first level, L1, the jet chamber signals are digitised and predefined track segment hit patterns are searched for. At L2, all the track segments are collected and linked to tracks, based on a fast parallel algorithm. A three dimensional helix fit is applied to tracks. At this level, event quantities, like track multiplicity, momentum sum and invariant masses for low multiplicity events are also determined. This is done in 10 μs, and the result is sent to L3, where software programs implemented on a computer farm reconstruct the event and look for particle resonances in high multiplicity events, within 100 μs.

[8] In the $r-\phi$ plane, which is perpendicular to the beam axis, tracks appear as circles. Tracks can be parametrised with a curvature κ, azimuthal angle ϕ, and the distance of closest approach DCA (the closest point of the circle with respect to the centre of coordinate system).

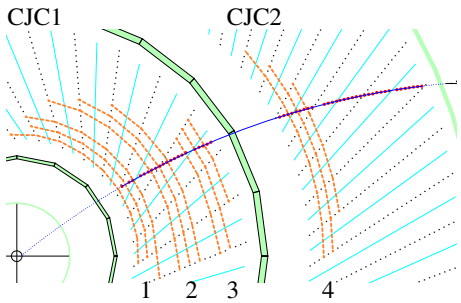

Figure 2.6: *The $r-\phi$ view of the H1 central jet chamber: a charged particle from the interaction point traverses the four trigger layers (the thick dashed lines) used by the FTT (from [63]).*

Chapter 3

Event Selection and Reconstruction

In this chapter, the selection criteria of the $D^*\mu$ sample which is used to measure the charm and beauty cross section are presented[1]. First, the general event preselection is discussed. Then the reconstruction of the D^* meson from its decay particles is explained, with emphasis on the Δm method. Next, the muon selection is described. The chapter continues with the level 4 classification of events. Finally, the triggers used for the different years of data taking are presented.

3.1 Overview of Selection Cuts

Table 3.1 contains a summary of the event selection cuts. In the following, the applied cuts are discussed in detail.

3.2 Preselection

In this analysis, data acquired before the HERA luminosity upgrade [66], i.e. HERA I run period (1999-2000), and after upgrade, i.e. HERA II (2004-2006), are used. Only *good* and *medium* quality runs[2], with acceptable beam conditions and subtrigger prescale factors, are selected. In addition, the sub-detectors necessary for this analysis, namely CJC, CIP/COP, CMD, the time of flight and the luminosity system, are required to be operational. The resulting integrated luminosities, after the detector status check and satellite bunch corrections [64], are presented in table 3.2.

Due to technical difficulties in accessing the level 4 classification, the 1997 data, corresponding to an integrated luminosity of around 19 pb^{-1}, are not considered in this analysis. This is not a considerable loss, since they amount to only 6% of the total HERA I and II luminosity. In the year 1998, the proton energy was increased from 820 GeV to 920 GeV, and HERA switched from e^+p to e^-p collisions. Because beam induced background caused

[1] This analysis was performed in the H1 Object Oriented, H1OO [64], environment. H1OO is based on the ROOT framework [65] which uses C++ as programming language.

[2] The runs are periods of data taking characterised by the same beam and trigger conditions.

Selection cuts	
Data taking period	1999, 2000, 2004, 2005, 2006
Run quality	good and medium
Detector status	CJC, CIP/COP, CMD, SpaCal, ToF, Lumi
z-vertex	$\lvert z_{vtx}\rvert < 35$ cm
Trigger selection	S19 $\parallel$ S22 $\parallel$ S56 (1999-2000) S19 $\parallel$ S23 (2004-2006)
L4 classes	class 15 (D^*) or class 16 (muon)
Kinematic selection	$E_{e'} < 8$ GeV ($Q^2 < 1$ GeV^2) $0.05 < y_{JB} < 0.75$
D^* selection	Decay channel: $D^* \to D^0\pi_s \to K\pi\pi_s$ $p_t(D^*) > 1.5$ GeV $\lvert\eta(D^*)\rvert < 1.5$ $\Delta m = m(K\pi\pi_s) - m(K\pi) \leq 0.1685$ GeV
D^* decay particles	$R(K,\ \pi,\ \pi_s) > 10$ cm $d'_{ca}(K\pi) < 1$ cm $p_t(K,\ \pi) > 0.3$ GeV $p_t(\pi_s) > 0.12$ GeV $\lvert m(K\pi) - m(D^0)\rvert \leq 80$ MeV
Muon selection	iron muon with $Q_\mu \geq 10_{new}$ $p(\mu) > 2$ GeV $\lvert\eta(\mu)\rvert < 1.735$
$D^*\mu$ selection	$f_{p_t}(D^*\mu) \geq 0.20$

Table 3.1: *Overview of the cuts applied in this analysis.*

Run period	**Year**	**Lepton**	**Luminosity [pb^{-1}]**
HERA I	1999	e^-	10.39
	1999	e^+	13.67
	2000	e^+	46.27
	$\sum$		70.33
HERA II	2004	e^+	38.83
	2005	e^-	102.04
	2006	e^-	53.75
	2006	e^+	74.54
	$\sum$		269.16

Table 3.2: *Integrated luminosities for HERA I and II run periods.*

problems in the tracking detectors, this data period is neglected as well. Runs with special trigger settings ('minimum bias runs') and runs with unstable experimental conditions which result in luminosities $\mathcal{L}_{run} < 0.1\,\mathrm{nb}^{-1}$ are omitted.

By requiring the primary vertex of the event to be reconstructed close to the nominal vertex: $|z_{vtx}| < 35$ cm, the non ep-background due to collisions of the satellite bunches with the nominal bunches is reduced.

3.3 Selection of Photoproduction Events

This analysis considers **untagged photoproduction** events, in which the electron escapes detection. These types of events are selected by requiring that there is no isolated high energy electromagnetic cluster, either in the SpaCal or the LAr calorimeter, consistent with a signal from a scattered electron:

$$E_{e'} < 8\,\mathrm{GeV}. \tag{3.1}$$

This restricts the photon virtuality to $Q^2 < 1\ \mathrm{GeV}^2$.

The inelasticity y_{JB} is calculated with the Jacquet-Blondel method [67], which is based on the reconstruction of the energy E_i and the momentum $\vec{p_i}$ of all particles i of the hadronic final state (HFS), i.e. all final state particles, except the scattered electron:

$$y_{JB} = \frac{1}{2E_e} \cdot \sum_{i \in HFS} (E_i - p_{z,\,i}). \tag{3.2}$$

The inelasticity is limited to:

$$0.05 < y_{JB} < 0.75. \tag{3.3}$$

The lower cut reduces the background events from proton-gas interactions [68]. Due to possible inefficiencies of the electron finder algorithm, the electron may be inadequately contained in the sum over the HFS particles, such that the inelasticity is overestimated with a quantity $(E'_e - p'_{e,\,z})/(2E_e)$. The upper cut on y_{JB} rejects therefore the background from DIS events.

3.4 The D^* Meson Reconstruction and Selection

The D^* mesons are reconstructed via the 'golden' decay channel:

$$D^{*\pm} \to \overset{(-)}{D}{}^0 \pi_s^{\mp} \to K^{\mp}\pi^{\pm}\pi_s^{\pm}. \tag{3.4}$$

Tracks measured in the central region of the detector, i.e. with pseudorapidity

$$-1.5 < \eta < 1.5 \tag{3.5}$$

are selected. These central tracks must fulfil the standard requirements presented in table 3.3.

Central tracks
$p_t > 0.10$ GeV
$20° < \theta < 160°$
$\lvert d'_{ca} \rvert < 2.0$ cm
$R_{start} < 50$ cm
$R > 10$ cm for $\theta \leq 150°$ $R > 5$ cm for $\theta > 150°$

Table 3.3: *Cuts used for the central tracks selection ('Lee West selection') [69]. d'_{ca} is the minimal distance of the tracks with respect to the primary vertex, R_{start} is the radial distance of the first hit with respect to the origin, and R is the radial track length.*

To reconstruct the D^0 meson, all possible combinations of two oppositely charged tracks, with $p_t > 300$ MeV, are considered. The pion and kaon mass, respectively, are assigned to the two tracks. The invariant mass of the D^0 candidate is given by:

$$m(K\pi) = \sqrt{(E_K + E_\pi)^2 - (p_K + p_\pi)^2}. \tag{3.6}$$

If the invariant mass is close to the nominal D^0 mass ($m_{D^0} = 1.865$ GeV [15]):

$$|m(K\pi) - m(D^0)| \leq 0.08 \text{ GeV}, \tag{3.7}$$

the combination of the two tracks is kept as a D^0 candidate. This suppresses the combinatorial background and contributions from D^0 decay modes other than $D^0 \rightarrow K\pi$ (*reflections*), which result in a D^0 reconstructed mass different from the nominal value. A third track, with $p_t > 120$ MeV and charge different from the charge of the kaon, is considered to be a slow pion (π_s) candidate. Then the D^* four-vector is given by summing the four-vectors of all decay particles.

The low p_t combinatorial background is suppressed by requesting:

$$p_t(D^*) > 1.5 \text{ GeV}. \tag{3.8}$$

The mass difference between the D^* and D^0 candidates, Δm, has to be:

$$\Delta m = m(K\pi\pi_s) - m(K\pi) \leq 0.1685 \text{ GeV}. \tag{3.9}$$

Figure 3.1 presents the correlations between the D^* meson and the daughter particles as predicted by PYTHIA Monte Carlo. Since the slow pion is produced almost at rest in the D^* meson rest frame, it is highly correlated with the D^* polar and azimuthal angle. The correlations with the D^0 daughters are not so pronounced.

3.4.1 The Δm Method

The advantage of the chosen D^* decay channel is that the nominal mass difference:

$$\Delta m = m_{D^*} - m_{D^0} = (145.42 \pm 0.01) \text{ MeV} \tag{3.10}$$

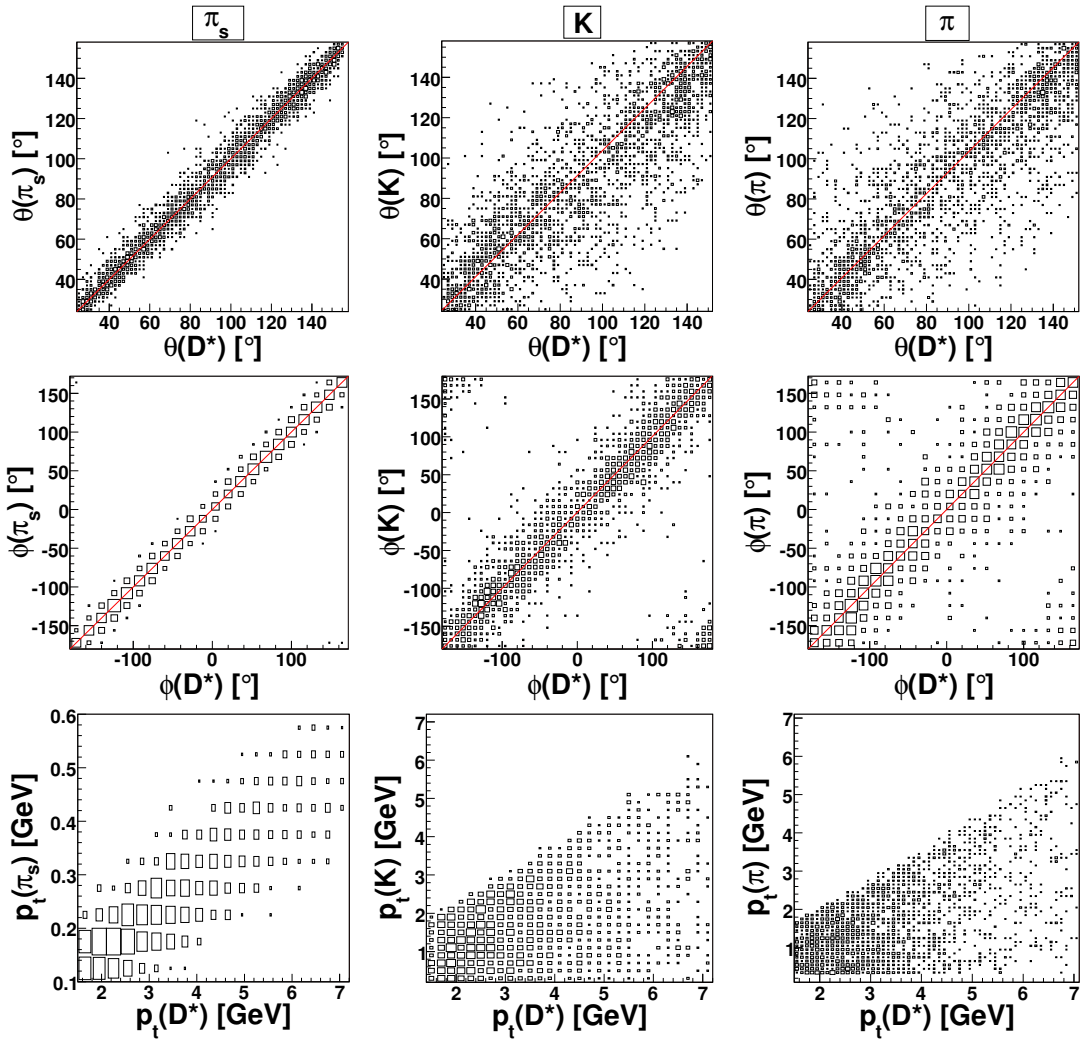

Figure 3.1: *Angular and transverse momentum correlations between the D^* meson and its daughters (π_s-left, K-middle, π-right), as predicted by charm PYTHIA Monte Carlo, direct component only.*

is close to the pion mass ($m_{\pi_s} = 139.57$ MeV), such that the phase space for combinatorial background is very limited. In addition, the Δm peak resolution is dominated by the experimental momentum resolution of the π_s track, since the K and π errors partially cancel [43].

In the D^* meson rest frame, the D^0 and π_s are produced with a low momentum:

$$p^* = \sqrt{\left(\frac{m_{D^*}^2 + m_{\pi_s}^2 - m_{D^0}^2}{2 \cdot m_{D^*}}\right)^2 - m_{\pi_s}^2} = 37.5 \text{ MeV}. \tag{3.11}$$

Since the pion from the D^* meson receives a small boost in the laboratory frame [68]:

$$\vec{p_{\pi_s}} \approx \sqrt{\frac{m_{\pi_s}^2 + p^{*2}}{m_{D^*}}} \cdot \vec{p_{D^*}} \approx \frac{1}{14}\vec{p_{D^*}} \tag{3.12}$$

it is referred to as the 'slow' pion.

The number of reconstructed D^* mesons is obtained by fitting the following function to the Δm distribution:

$$f_{total}^{RiCh}(\Delta m) = s_{bin} \cdot [f_{signal}(\Delta m) + f_{bgr}(\Delta m)]\,, \tag{3.13}$$

where s_{bin} is the bin width. The D^* signal fit function consists of a Gaussian:

$$f_{signal}(\Delta m) = \frac{N(D^*)}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left(-\frac{1}{2}\cdot\frac{(\Delta m-\mu)^2}{\sigma^2}\right). \tag{3.14}$$

$N(D^*)$ is the number of D^* mesons, σ and μ are the width and the mean value of the Gaussian distribution. The background is described by an exponential function starting at the phase space border, the pion mass. The fit results are improved if a quadratic term in Δm is considered as well:

$$f_{bgr}(\Delta m) = U_N \cdot (\Delta m - m_\pi)^{U_{exp}} \cdot (1 - U_{cor}\cdot\Delta m^2). \tag{3.15}$$

The fit parameters are: the normalisation factor U_N, the exponential U_{exp} and the correction factor U_{cor}. The maximum-likelihood method is used in the fit. The results are shown in figure 3.2, left.

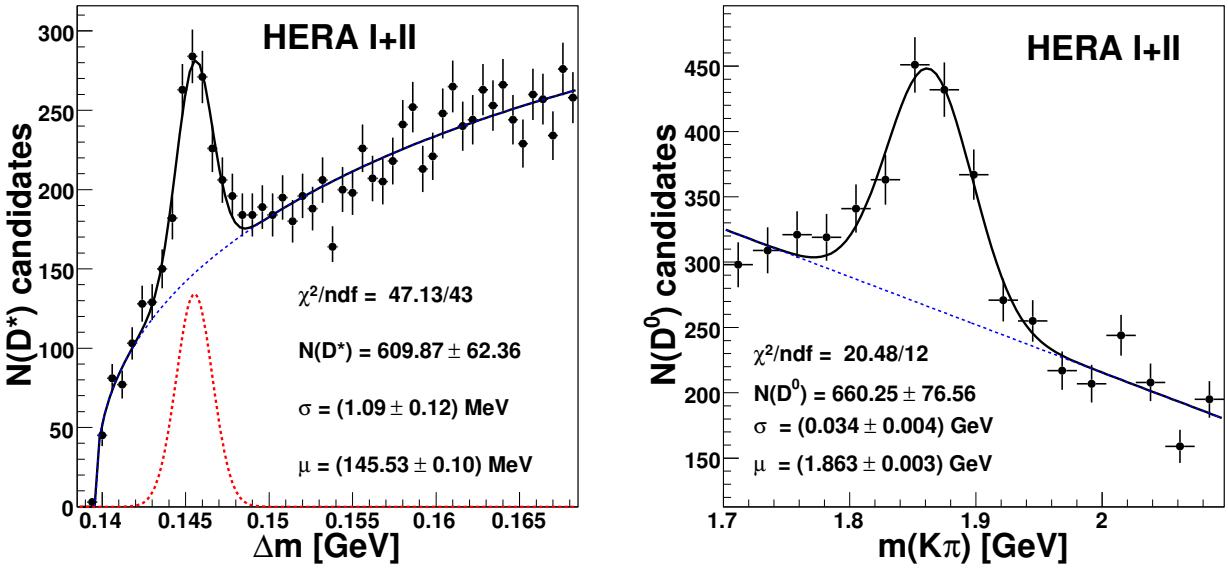

Figure 3.2: *Left: the fitted $\Delta m = m(K\pi\pi_s) - m(K\pi)$ distribution for HERA I and II run periods. Right: the fitted D^0 invariant mass $m(K\pi)$. The fit functions are described in text.*

To verify conversely the number of D^* mesons obtained from the Δm fit, the D^0 invariant mass $m(K\pi)$ distribution, for the events in the signal region:

$$0.143\ \text{GeV} < \Delta m < 0.148\ \text{GeV} \tag{3.16}$$

is fitted as well (see figure 3.2, right). The D^0 fit function consists of a Gaussian for the signal and a polynomial for the background:

$$f(m(D^0)) = s_{bin}\cdot\frac{N(D^0)}{\sqrt{2\pi}}\cdot\exp\left[-\frac{1}{2}\cdot\left(\frac{m(D^0)-\mu}{\sigma}\right)^2\right] + U_{slope}\cdot m(D^0) + U_N. \tag{3.17}$$

The obtained number of D^0, $N(D^0) = 660 \pm 77$, and D^* mesons, $N(D^*) = 610 \pm 62$, are compatible, within the errors.

3.4.2 Background Subtraction

To describe the combinatorial background, *wrong charge (WrCh)*, combinations (in contrast to the *right charge (RiCh)*) are selected. Again three tracks are considered, but this time the K and the π candidates are taken to be of the same charge, as originating from a doubly charged particle, such that in the end three tracks combinations of the type $K^+\pi^+\pi_s^-$ and $K^-\pi^-\pi_s^+$ are kept. Because in both cases, $RiCh$ and $WrCh$, two tracks having the same charge and one track of opposite charge are selected, one expects the combinatorial background in the Δm distributions to have similar shapes. However, the $WrCh$ background proves to be slightly lower. This is sometimes explained based on multiplicities (see for example [70]). Let $N_{+(-)}$ be the number of tracks for positively (negatively) charged particles, with transverse momentum $p_t > 0.3$ GeV. Similarly, $n_{+(-)}$ is defined for lower transverse momentum particles. Then the possible combinations are:

$$RiCh: \quad (K^{\pm}\pi^{\mp})\pi_s^{\mp} \rightarrow N_{RiCh} \quad \equiv \quad N_+N_-n_- + N_-N_+n_+ = N_+N_- \cdot (n_+ + n_-) \tag{3.18}$$

$$WrCh: \quad (K^{\pm}\pi^{\pm})\pi_s^{\mp} \rightarrow N_{WrCh} \quad \equiv \quad N_+ \cdot (N_+ - 1) \cdot n_- + N_- \cdot (N_- - 1) \cdot n_+ \tag{3.19}$$

On average, $N_{WrCh} < N_{RiCh}$, such that $WrCh$ combinatorial background underestimates the $RiCh$ background. However, scaling factors of about 0.7-0.8, as indicated in figure 3.3, cannot be explained alone by the multiplicity argument. Another reason may be the level 4 D^* finder (see section 3.6), which looks only for $RiCh$ combinations, such that $WrCh$ combinations are lost at this level.

To account for this, a scaling factor is introduced:

$$U_{scal} = \frac{\sum_{bin} N_{bin}^{WrCh}}{\sum_{bin} N_{bin}^{RiCh}} \quad \text{for } 0.155 < \Delta m < 0.1685, \tag{3.20}$$

which scales the $WrCh$ background such that it describes the $RiCh$ background far from the signal region, in order to be independent of the signal fit function. The scaling factor is obtained by fitting simultaneously the $RiCh$ and $WrCh$ distributions with the function:

$$f_{bgr}^{Ri+Wr} = s_{bin} \cdot \left[f'_{bgr}(\Delta m_{RiCh} + U_{scal} \cdot f'_{bgr}(\Delta m_{WrCh}) \right], \tag{3.21}$$

where:

$$f'_{bgr}(\Delta m) = U_N \cdot (\Delta m - m_\pi)^{U_{exp}}. \tag{3.22}$$

Such a fit is shown in figure 3.3. Once the background is known, it can be subtracted to obtain the signal distributions:

$$signal = RiCh - \frac{WrCh}{U_{scal}}. \tag{3.23}$$

This method of background subtraction will be later used in this analysis (see chapter 8).

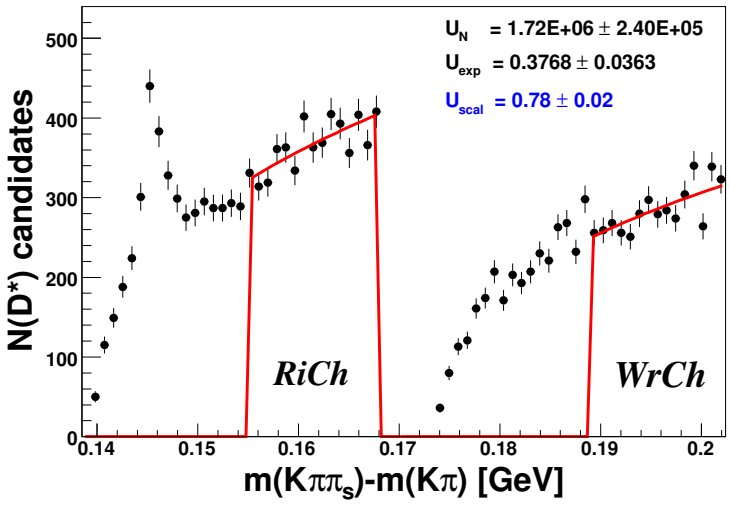

Figure 3.3: *Illustration of the simultaneous fit to the right (RiCh) and wrong charge (WrCh) distributions. For clarity, the latter is shifted with 171.5 MeV to the right. The obtained scaling factor, U_{scal}, is also indicated.*

3.5 Muon Reconstruction and Selection

Since the purity of the calorimeter muons is low [71], only muons identified in the instrumented iron system are considered in this analysis . The muon selection is performed using good tracks as defined in [69] (see table 3.3).

	Barrel	**Forward end-cap**	**Backward end-cap**
ρ_0	< 100 cm	< 100 cm	< 100 cm
z_0	< 100 cm	< 100 cm	< 100 cm
n_{layers}	≥ 2	≥ 6	≥ 3
first layer	≤ 5	≤ 5	≤ 8
last layer	≥ 2	≥ 6	≥ 3

Table 3.4: *Cuts applied to the muon iron tracks reconstructed in the barrel, forward and backward end-cap. The quantity n_{layers} is the number of layers with hits in the corresponding part of the detector.*

In this analysis, the standard cuts listed in table 3.4 are used for the iron tracks. The quantity ρ_0 is the distance of the (back-extrapolated) iron track to the beam line at $z = 0$, while z_0 is the distance of (back-extrapolated) iron track to the nominal interaction point at $z = 0$. Both variables are related to the probability that the iron track is due to a particle from an ep event: the smaller they are, the bigger is the probability. The cuts on the number of layers and the last layer are useful because the hadrons which form the non-muon background are likely to get stuck in between the iron plates.

To be sure that the muon comes from the decay of a charmed or beauty hadron, only *primary vertex fitted tracks* are selected. This reduces the muon background from misidentified particles, like kaons, which have a large enough lifetime, such that they can be

resolved in the CJC.

The following muon quality (see section 2.4.1) is assigned to iron muons:

$$Q_{\mu}^{iron} = Q_{\mu}^{Cal} + 10. \tag{3.24}$$

As in [71], a new iron muon quality, 10_{new}, is defined. This new quality corresponds to quality 10, with additional requirements: for polar angle[3] $\theta_{\mu} > 135°$, the number of layers with hits[4] should be $n_{layers} > 5$ and the first layer with hits has to be less than or equal to 5. In this analysis, only iron muons with the following qualities are selected:

$$Q_{\mu}^{iron} = 10_{new},\ 11,\ 12 \text{ and } 13. \tag{3.25}$$

Only muons with a momentum of at least 1.5 GeV are able to traverse the LAr calorimeter and reach the central muon detector [55]. To be sure that they traverse also the iron layers, muons with:

$$p(\mu) > 2 \text{ GeV} \tag{3.26}$$

are selected. In addition, their pseudorapidity is required to be:

$$|\eta(\mu)| < 1.735. \tag{3.27}$$

The last cut corresponds to the one applied to the central tracks: $20° < \theta < 160°$. Similar to the D^* meson daughters, a cut on the radial track length is applied:

$$R(\mu) > 10 \text{ cm}. \tag{3.28}$$

3.6 Level 4 Classification

On trigger level 4 (L4), the events are subject to an event classification. If they are not assigned to a class, they are down-scaled by the L4 weight. The distribution of these weights for the run periods used in this analysis is shown in figure 3.4, left. In total, 12% of the events have L4 weights. These events need to be included in the distributions with the corresponding weight. However, there are problems when fitting the Δm distributions with reduced statistics and large weights. Another possibility, which is chosen here, is to select only events which were assigned to an L4 class, since in this case the weights are 1.

The L4 classes distribution for the HERA I and II run periods can be seen in figure 3.4, right. Most of the events belong to classes 24 ('muons'), 15 ('open heavy flavour') and 16 ('hidden heavy flavour')[5]. The latter two classes, 15 and 16, are used in this analysis.

[3] At this angle, $\theta \backsimeq 135°$, the LAr calorimeter ends and more hadrons could reach the iron system and fake muons.

[4] The inner and outer muon boxes layers are not considered, since there no material is traversed, so they do not improve the muon signature.

[5] In charm production, the term 'open' refers to the events in which the c and the $\bar{c}$ quark fragment into different hadrons, such that the final particles get a c-quantum number different from zero (e.g. $D^{*+} = c\bar{d}$.). In the opposite case, 'hidden' charm production, both c quarks fragment into the same hadron, which has therefore a net c-quantum number equal to zero (e.g. $J/\psi = c\bar{c}$).

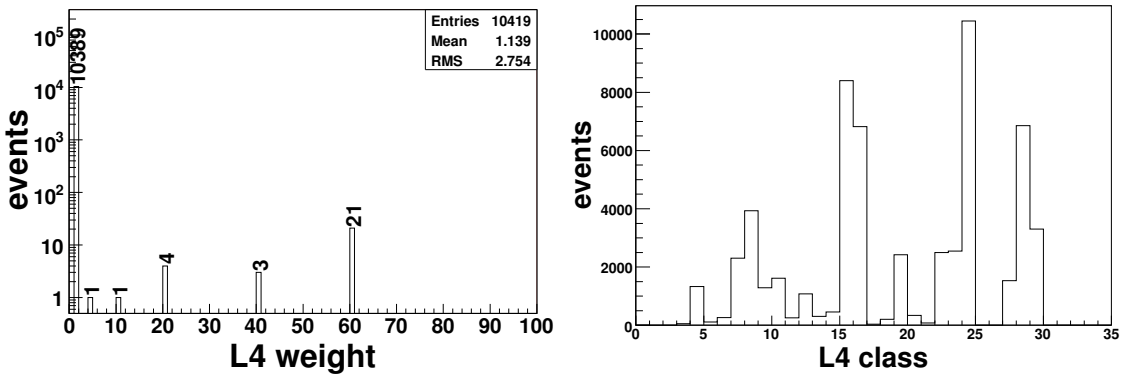

Figure 3.4: *Left: distribution of L4 weights for HERA I and II run periods. Right: distribution of L4 classes.*

The events which belong to **class 15** are found by a Heavy Quark Selection routine HQSEL [72] which searches also in other D^* decay channels than the one used in this analysis. To reduce the L4 trigger rates, several cuts (see table 3.5) are already applied in this routine.

Subtriggers	1, 2, 4, 43, 44, 61, 82, 83, 100 or 108
$K,\ \pi$	$\lvert d'_{ca}\rvert < 4$ cm track length > 10 cm $\chi^2_{xy} < 50$ $p_t > 150$ MeV
π_s	track length > 10 (15) cm $p_t > 70$ (100) MeV
D^0 D^*	$1.364\ (1.764) < m(K\pi) < 2.364\ (1.864)$ $p_t > 1.4$ GeV $\Delta m = m(K\pi\pi_s) - m(K\pi) < 170$ MeV

Table 3.5: *Cuts of the L4 open charm finder HQSEL, for untagged photoproduction events with D^* mesons from the decay channel $D^* \to D^0\pi_s \to K\pi\pi_s$. χ^2_{xy} is the χ^2 of the track fit in the $x-y$ plane. The harder cuts shown in brackets have been applied until mid of 2005. Apart from the cut on χ^2_{xy}, all the other cuts are lower than the ones applied in this analysis (see table 3.1).*

Class 16 contains events accepted by the 'high mass finder' [73]. This finder runs on all events, independent of subtriggers, and searches for muons and electrons from heavy flavours. The selected events are assigned to 6 subclasses, as indicated in table 3.6. From figure 3.5 it can be seen that most of the events belong to subclass AOPEN.

	Subclass	Content
1	AOPEN	One iron or forward muon from b or c decays
2	ATWOMU	One high quality muon (from elastic $J/\psi \to \mu\mu$) and one track with invariant mass $M > 2$ GeV
3	ATWOEL	One high quality electron (from elastic $J/\psi \to ee$) and one track with invariant mass $M > 2$ GeV
4	AJPSI	Two high quality muons (from inelastic $J/\psi \to \mu\mu$) with invariant mass $M > 2$ GeV
5	2ELEC	At least two identified electrons (from inelastic $J/\psi \to ee$) with good associated tracks and invariant mass $M > 1.5$ GeV
6	2PRONG	Two central tracks with invariant mass $M > 2$ GeV; no identified leptons required

Table 3.6: *Classification of the events assigned to class 16.*

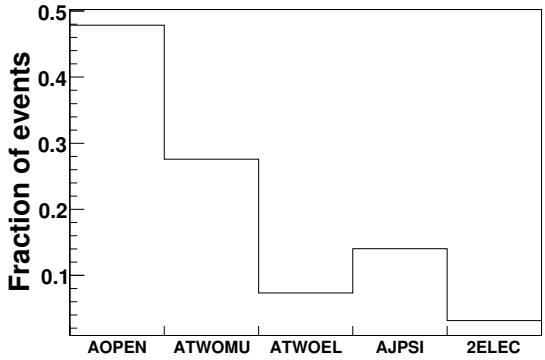

Figure 3.5: *Distribution of events in class 16. Only the subclasses in which at least one lepton is identified are shown.*

3.7 $D^*\mu$ Cut

To reduce the combinatorial background, a cut on the ratio of the $D^*\mu$ transverse momentum to the transverse momentum of all particles in the hadronic final state $f_{p_t}(D^*\mu)$ is applied. Since the D^* mesons resulting from heavy quark production carry most of the transverse momentum, the $D^*\mu$ system is required to carry a fraction of 20% of the total transverse momentum of the hadronic final state particles:

$$f_{p_t}(D^*\mu) = \frac{p_t(D^*) + p_t(\mu)}{\sum_{\theta>10^\circ} p_t} > 0.20. \tag{3.29}$$

By selecting only particles with polar angle $\theta > 10^\circ$, contributions from the proton remnant[6] are reduced. The $f_{p_t}(D^*\mu)$ distribution, together with the corresponding efficiency (see

[6]The term 'proton remnant' refers to the particles scattered at small angles which are lost down the beam pipe.

section 6.6) and signal to background ratio, are shown in figure 3.6. The *signal to background ratio* is calculated as the ratio of the integral of the Gaussian and the background function within 3σ around the mean value of the Gaussian. This ratio increases rapidly, as the cut is tightened. However, with the applied cut, a compromise between reducing the background, and keeping as many events as possible is made (see also appendix C). At the transverse momentum fraction of 20%, an efficiency of 90% and a signal to background ratio of 0.64 are obtained.

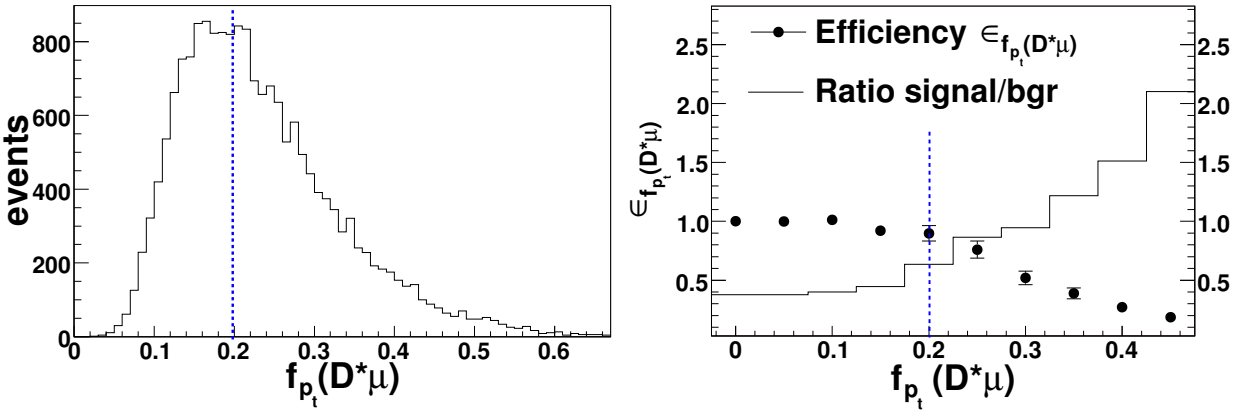

Figure 3.6: *Left: distribution of $f_{p_t}(D^*\mu)$ variable for HERA I and II data (not background subtracted). Right: the markers show the efficiency of the $f_{p_t}(D^*\mu)$ cut, i.e. number of D^* mesons for every bin of the variable, divided by the number of the D^* mesons obtained without any cut on the $D^*\mu$ transverse momentum fraction. The distribution presents the signal to background ratio as a function of the same variable. The dotted lines indicate the cut applied in the analysis.*

3.8 Subtrigger Selection

The H1 trigger system was introduced in section 2.6. In this section, the selection of the analysis subtriggers for HERA I and II run periods is presented. This is done separately for the two periods due to varying trigger definitions.

3.8.1 HERA I

In figure 3.7, the *actual subtriggers*[7] distribution for the $D^*\mu$ events in the signal region[8] (from the HERA I data sample) can be seen. The subtriggers with the highest response are: 15, 19, 22, 37, 56, 60, 71, 83 and 87. Out of them, the subtriggers 19 and 22, which trigger events with at least one high transverse momentum muon, and 56, which selects events with an untagged D^* meson, were chosen. These subtriggers are logical combinations of level 1 (L1) trigger elements (see table 3.7).

[7] The actual subtriggers are a logical 'AND' combination between actual L1 subtriggers and verified L4 subtriggers (see section 2.6).

[8] The *signal region* is defined by $0.143\ \text{GeV} < \Delta m < 0.148\ \text{GeV}$, i.e. $\pm 3\,\sigma$ around the nominal Δm value.

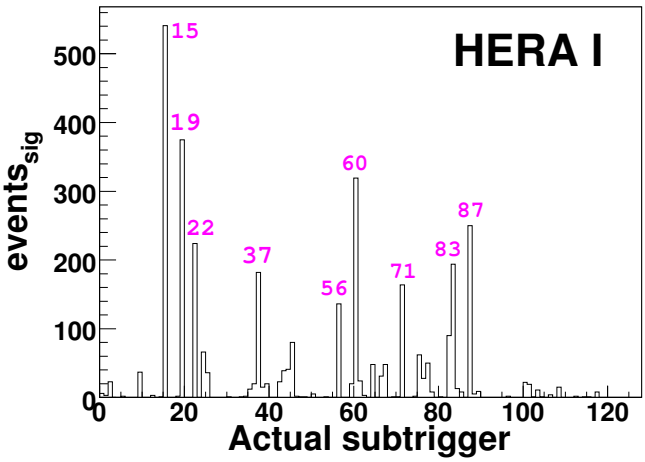

Figure 3.7: *Distribution of actual subtriggers for events with a D^* meson in signal region and a muon (HERA I data sample).*

Subtrigger	Definition	Average prescale
19	Mu_Bar && DCRPh_CNH && (zVtx_sig > 1)	1.03
22	Mu_BIO && DCRPh_CNH && (zVtx_sig > 1)	1.04
56	Mu_Any && DCRPh_Ta && (SPCLe_IET > 1 || SPCLe_IET_Cen_2)	1.00

Table 3.7: *Definition of subtriggers used for HERA I run period (the veto trigger elements are not shown). The following notations are used: '&&' for logical AND, '||' for logical OR. The meaning of the trigger elements is given in appendix D. The luminosity averaged prescale factors for the selected subtriggers are also shown.*

Subtrigger S19 is a logical combination of muon, track and z-vertex trigger elements. Mu_Bar fires if a muon signature in the barrel part of the muon system is found. The notation DCRPh_CNH stands for a combination of trigger elements which fire if there are at least three tracks with transverse momentum $p_t \geq 450$ MeV, one negative track with $p_t \geq 450$ MeV and one track with $p_t \geq 800$ MeV. In addition, zVtx_sig > 1 is required in the definition of subtrigger 19. This is the equivalent of a significant maximum in the z-vertex histogram (see section 2.6).

Subtrigger S22 is similar to S19. The difference is in the muon trigger element, which fires if a muon signature is found either in the backward or forward outer cap, or if there are at least two muon signatures in the backward or outer cap.

Subtrigger S56 is a combination of muon, track and SpaCal trigger elements. Mu_Any fires if there is a muon signature in any part of the central muon detector (end-caps or barrel). The trigger element DCRPh_Ta is on if a track with $p_t \geq 450$ MeV is found. The SpaCal trigger elements fire in case an electron with energy $E \geq 2$ GeV or $E \geq 6$ GeV is measured in the central region of the SpaCal.

The remaining subtriggers are designed to trigger: events with inelastic J/ψ (S15, S37 and

S60), isolated leptons (S71) and D^* jets in tagged photoproduction (S83 and S87). These subtriggers contribute only up to 5% to the $D^*\mu$ sample, so they are not considered in this analysis.

3.8.2 HERA II

The HERA II actual subtrigger distribution is shown in figure 3.8. Most of the events are triggered by S15, S19 and S23. Subtrigger 15, which is designed for inelastic J/ψ events, contributes with less than 7% to the total $D^*\mu$ events, and it has a high average prescale factor (1.75 for 2004 and 2005). Therefore, only subtriggers S19 and S23 are used for the HERA II run period. Their definition at the level 1 (L1) is given in table 3.8. In 2006, level 2 and 3 trigger conditions were added. The L2 condition requires at least one reconstructed track at L2 with $p_t \geq 400$ MeV. The L3 condition checks if at least one high p_t track matches the triggered muon module.

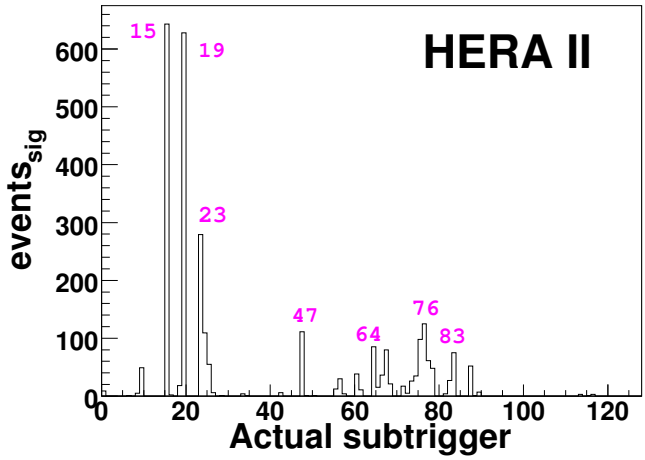

Figure 3.8: *Distribution of actual subtriggers for events with a D^* meson in signal region and a muon (HERA II data sample).*

3.8.3 Event Yield

To check if the event selection is stable over the whole data taking period, the *event yield*, i.e. the number of photoproduction events which contain a D^* meson and a muon per unit of integrated luminosity is plotted as a function of run number for each year (see figures 3.9 and 3.10). Note that the low value in some of the last bins is just due to binning effects. If the detector readout functioned properly and the luminosity was accurately measured, the event yield should be constant, apart from changes in the trigger definitions which are documented.

In table 3.9 the mean event yield is shown. In the beginning of 2005, the event yield for S19 was smaller compared to the other years. This was due to hardware problems in the muon trigger electronics. In general, the subtriggers have comparable event yields for the

Sub-trigger	Year	Definition	Average prescale
19	2004, 2005 (run number $<$ 401627)	Mu_Bar && DCRPh_CNH	1.03
	2005 (run number $\geq$ 401627, run number $<$ 427872)	Mu_Bar && FTT_mul_Tc $>$ 2 && FTT_mul_Td $>$ 0	
	2005, 2006 (run number $\geq$ 427872)	CIP_sig $>$ 1 && Mu_Bar && FTT_mul_Tc $>$ 2 && FTT_mul_Td $>$ 0	
23	2004,2005 (run number $<$ 401617)	Mu_BIO && DCRPh_CNH	1.13
	2005 (run number $\geq$ 401617 run number $<$ 427872)	Mu_BIO && FTT_mul_Tc $>$ 2	
	2005, 2006 run number $\geq$ 427872	CIP_sig $>$ 1 && Mu_BIO && FTT_mul_Tc $>$ 2	

Table 3.8: *Definition of subtriggers used for HERA II run period on trigger level 1 (the veto trigger elements are not shown). The definitions changed in the year 2005, with the introduction of the FTT trigger (see section 2.6.1).*

different years, with the exception of S19 and S23, which improved significantly in 2006. The variations in the event yield due to hardware problems should be described by the Monte Carlo. This will be checked in section 6.3.

Mean event yield				
Run period		**S19**	**S22**	**S56**
HERA I	1999	73 ± 4	40 ± 2	31 ± 2
	2000	97 ± 5	55 ± 3	33 ± 4
		S19	**S23**	
HERA II	2004	82 ± 6	-	
	2005	74 ± 4	60 ± 3	
	2006	115 ± 4	103 ± 4	

Table 3.9: *Mean event yield, i.e. number of selected events divided by luminosity, for HERA I and II subtriggers.*

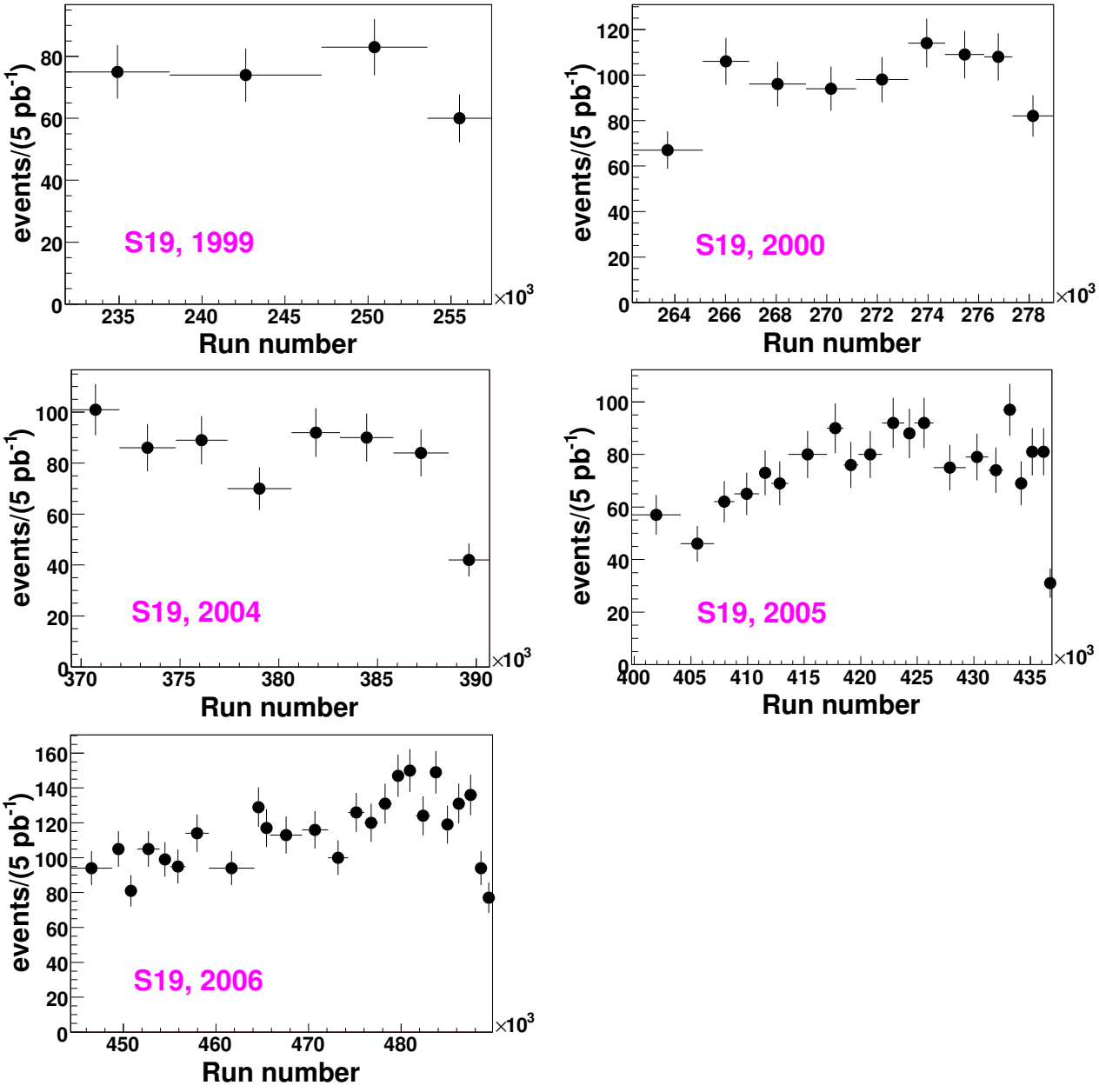

Figure 3.9: *The number of $D^*\mu$ events triggered by **S19** as a function of run number. The event yields for the different years are shown separately. Each bin corresponds to an integrated luminosity of 5 pb^{-1}.*

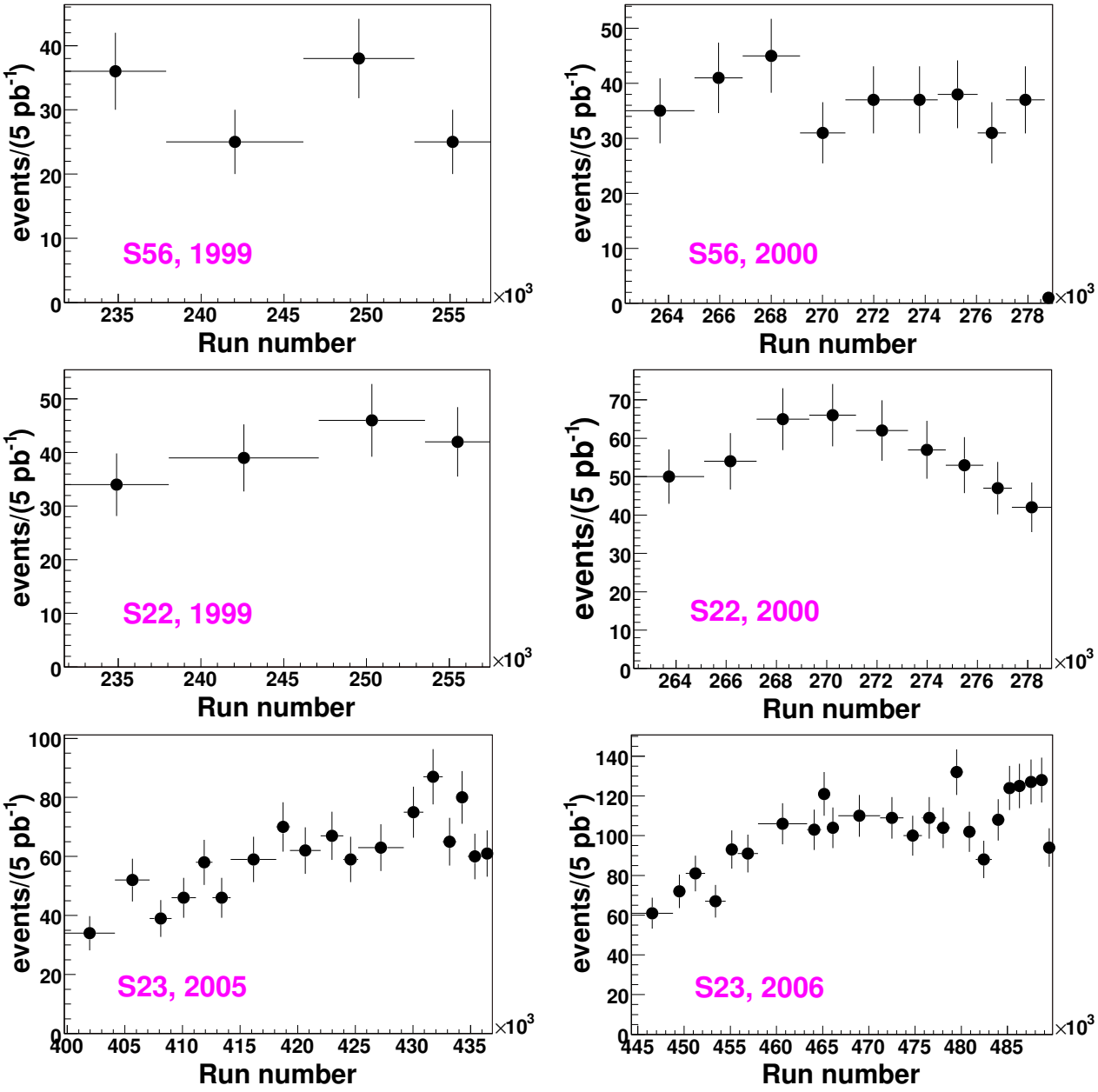

Figure 3.10: *The number of $D^*\mu$ events triggered by **S56** (top), **S22** (middle) and **S23** (bottom) as a function of run number. The event yields for the different years are shown separately. Each bin corresponds to an integrated luminosity of 5 pb^{-1}.*

Chapter 4

Monte Carlo Description of Data

The goal of the analysis is to study charm and beauty production at HERA. To obtain the fractions of c and b events in data, distributions as predicted by PYTHIA Monte Carlo will be used (see chapter 5). Further, the efficiencies and the muon background corrections needed for the cross section measurement (see chapter 6) will also be determined from Monte Carlo. For reliable results, the Monte Carlo has to describe the data in all aspects. To check if this is the case, a comparison of data to PYTHIA Monte Carlo[1] will be presented in this chapter. First, the corrections of the Monte Carlo due to used out-dated branching ratios and to missing prescale factors will be explained. Then, signal distributions of the kinematic variables interesting for the analysis, in data and in Monte Carlo (after corrections), will be shown. Finally, a study of the muon background using events from the $D^0 \to K\pi$ decay will be presented.

4.1 Monte Carlo Corrections

4.1.1 Branching Ratios

As mentioned in section 1.6.3, due to technical reasons, two different versions of PYTHIA generator were used. In the two cases, the events were generated according to version dependent branching ratios which differ from the most recent measurements [15]. Therefore, correction factors are determined for the branching ratios important for this analysis. They are given in table 4.1 (4.2) for PYTHIA version 6.1 (6.2).

For charm production, where only one configuration is possible (see section 1.7.2.1), the correct branching ratios play a role only in the Monte Carlo cross section measurement.

For beauty production, on the other hand, the branching ratios influence in addition the contributions of $D^*\mu$ events coming from one of the three possible beauty scenarios. To ensure the correct contributions, the fraction of beauty events from the different scenarios, according to the measured branching ratios (f^{meas}) are calculated (see tables 4.3 and 4.4).

[1] CASCADE Monte Carlo will be later used to determine the systematic uncertainty due to the model dependence (see section 6.8.1).

Channel	$\mathcal{BR}^{MC}$ [%]	$\mathcal{BR}^{meas}$ [%]	$f_{cor} = \mathcal{BR}^{meas}/\mathcal{BR}^{MC}$
$D^* \to K\pi\pi_s$	2.5	2.57 ± 0.05	1.03 ± 0.02
$c \to D^*$	30.0	$25.5 \pm 1.5 \pm 0.8$	0.85 ± 0.05
$c \to \mu$	9.4	8.2 ± 0.5	0.87 ± 0.05
$1/2 \cdot c\bar{c} \to D^*\mu$	2.82	2.09 ± 0.18	$\mathbf{0.74 \pm 0.06}$
$b \to D^*$	30.0	17.3 ± 2.0	0.58 ± 0.07
$b \to \mu$	10.5	10.95 ± 0.27	1.04 ± 0.03
$b \to X \to \mu$	10.7	10.03 ± 0.64	0.94 ± 0.06
$b \to D^*\mu$	2.19	2.75 ± 0.19	1.26 ± 0.09
$1/2 \cdot \mu \leftarrow X \leftarrow \bar{b}b \to D^*$	3.15	1.74 ± 0.35	0.55 ± 0.11
$1/2 \cdot \mu \leftarrow \bar{b}b \to D^*$	3.21	1.89 ± 0.22	0.59 ± 0.07
$b \to D^*\mu$	2.19	2.75 ± 0.19	1.26 ± 0.09
$1/2 \cdot b\bar{b} \to D^*\mu$	8.55	6.38 ± 0.46	$\mathbf{0.75 \pm 0.05}$

Table 4.1: *Comparison between the branching ratios used in **PYTHIA version 6.1**, $\mathcal{BR}^{MC}$, and the measured ones, $\mathcal{BR}^{meas}$ [15]. The correction factors f_{cor} marked with bold font, will be applied when determining the Monte Carlo cross sections.*

Channel	$\mathcal{BR}^{MC}$ [%]	$\mathcal{BR}^{meas}$ [%]	$f_{cor} = \mathcal{BR}^{meas}/\mathcal{BR}^{MC}$
$D^* \to K\pi\pi_s$	2.49	2.57 ± 0.05	1.03 ± 0.02
$c \to D^*$	25.4	$25.5 \pm 1.5 \pm 0.8$	1.0 ± 0.06
$c \to \mu$	10.1	8.2 ± 0.5	0.81 ± 0.05
$1/2 \cdot c\bar{c} \to D^*\mu$	2.57	2.09 ± 0.18	$\mathbf{0.81 \pm 0.05}$
$b \to D^*$	27.2	17.3 ± 2.0	0.64 ± 0.07
$b \to \mu$	10.5	10.95 ± 0.27	1.04 ± 0.03
$b \to X \to \mu$	9.8	10.03 ± 0.64	1.02 ± 0.07
$b \to D^*\mu$	2.6	2.75 ± 0.19	1.06 ± 0.09
$1/2 \cdot \mu \leftarrow X \leftarrow \bar{b}b \to D^*$	2.67	1.74 ± 0.35	0.65 ± 0.13
$1/2 \cdot \mu \leftarrow \bar{b}b \to D^*$	2.42	1.89 ± 0.22	0.78 ± 0.09
$b \to D^*\mu$	2.6	2.75 ± 0.19	1.06 ± 0.07
$1/2 \cdot b\bar{b} \to D^*\mu$	7.69	6.38 ± 0.46	$\mathbf{0.83 \pm 0.06}$

Table 4.2: *Comparison between the branching ratios used in **PYTHIA version 6.2**, $\mathcal{BR}^{MC}$, and the measured ones, $\mathcal{BR}^{meas}$ [15]. The correction factors f_{cor} marked with bold font, will be applied when determining the Monte Carlo cross sections.*

A similar fraction is defined based on the Monte Carlo branching ratios (f^{MC}). By dividing the two fractions, correction factors for beauty scenarios were obtained. At **generator** level, where the decay chain of the D^* meson and of the muon can be traced back exactly, all the events were weighted with the corresponding correction factor. At **reconstructed** level, the particles needed to be matched with the generated ones.

Channel	f^{MC} [%]	f^{meas} [%]	$f_{cor} = f^{meas}/f^{MC}$
$\mu \leftarrow X \leftarrow \bar{b}b \rightarrow D^*$	37	27 ± 8	**0.73±0.21**
$\mu \leftarrow \bar{b}b \rightarrow D^*$	38	30 ± 11	**0.79±0.29**
$b \rightarrow D^*\mu$	25	43 ± 16	**1.72±0.64**

Table 4.3: *Contribution of the different beauty scenarios. The indicated correction factors f_{cor} are applied in* ***PYTHIA version 6.1****.*

Channel	f^{MC} [%]	f^{meas} [%]	$f_{cor} = f^{meas}/f^{MC}$
$\mu \leftarrow X \leftarrow \bar{b}b \rightarrow D^*$	35	27 ± 8	$\mathbf{0.77 \pm 0.23}$
$\mu \leftarrow \bar{b}b \rightarrow D^*$	31	30 ± 11	$\mathbf{0.97 \pm 0.35}$
$b \rightarrow D^*\mu$	34	43 ± 16	$\mathbf{1.27 \pm 0.47}$

Table 4.4: *Contribution of the different beauty scenarios. The indicated correction factors f_{cor} are applied in* ***PYTHIA version 6.2****.*

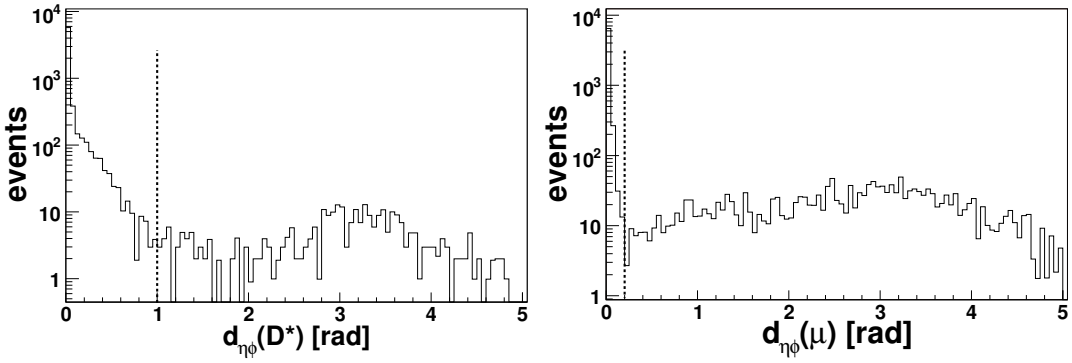

Figure 4.1: *The distance $d_{\eta\phi}$ in $(\eta - \phi)$ plane for the D^* meson (left) and for the muon (right), obtained with PYTHIA Monte Carlo. The dashed lines indicate the cuts applied to ensure a good correlation between the reconstructed and the generated particles.*

For this, the distance in the $\eta - \phi$ plane[2] was defined:

$$d_{\eta\phi} = \sqrt{(\eta_{rec,\,X} - \eta_{gen,\,X})^2 + (\phi_{rec,\,X} - \phi_{gen,\,X})^2}, \text{ where } X = D^* \text{ or } \mu. \tag{4.1}$$

From figure 4.1, it can be seen that for the D^* meson, the particle at the reconstructed level is correlated with the one at the generated level for $d_{\eta\phi}(D^*) < 1$. For muon, the correlation is present for $d_{\eta\phi}(\mu) < 0.2$. Therefore, the branching ratio corrections at **reconstructed level** are applied for these distances $d_{\eta\phi}$.

After these corrections, it was studied how are the beauty events distributed in the four correlation regions (see table 4.5). The correlation regions 1 and 2 are dominated by beauty events from the decay $\mu \leftarrow \bar{b}b \rightarrow D^*$. However, apart from the B^0 oscillations, processes of the type $b \rightarrow c\bar{c}s$ result in the D^* meson and the muon having the same charge, although

[2] If the reconstructed particle corresponds to the generated one, the distance $d_{\eta\phi}$ between the two should be small.

Correlation regions	Beauty scenarios [%]		
	$\mu \leftarrow X \leftarrow \bar{b}b \rightarrow D^*$	$\mu \leftarrow \bar{b}b \rightarrow D^*$	$b \rightarrow D^*\mu$
1	5	95	-
2	9	91	-
3	1	3	96
4	21	38	41
all	9	44	47

Table 4.5: *Corrected contributions of the different beauty scenarios to the correlations regions, as predicted by PYTHIA Monte Carlo. The values are obtained with D^* meson and muon at generator level. The cuts which define the visible range of the analysis are applied.*

they come from different quarks. In region 3, the decay $b \rightarrow D^*\mu$ contributes mostly. Surprisingly, the largest contribution to region 4 does not come from $\mu \leftarrow \bar{b}b \rightarrow D^*$, but from $b \rightarrow D^*\mu$. This may be explained by the fact that the muons from cascade decay have softer momenta, and are therefore suppressed by the cut on the muon momenta.

4.1.2 Level 1 Prescales

The used Monte Carlo programs do not include the L1 prescale factors. To make a comparison of data to the theoretical models implemented in Monte Carlo, the prescale factors are introduced by weighting the events as described bellow.

If subtrigger i has a prescaling factor d_{ik} in run k, the probability that this subtrigger actually fires in the event number j, in run k, is [74]:

$$P_{jk}^{L1} = 1 - \prod_{i=1}^{N_{subtr}} \left(1 - \frac{1}{d_{ik}},\right) \tag{4.2}$$

where N_{subtr} is the number of subtriggers, from a subtrigger combination (see tables 4.6 and 4.7), which fired.

One can use the inverse of this probability as a weight, i.e. $w_{jk} = 1/P_{jk}^{L1}$, for each event to obtain the correct subtrigger rates. However, since the prescale factors may vary from run to run, events triggered by the same subtriggers combination are assigned different weights, which results in large statistical errors.

To reduce them, the data were divided into run ranges with approximately unchanged trigger definitions, i.e. for HERA I the periods 1999e^- and 1999e^+/2000 were considered, and for HERA II the years 2004, 2005 and 2006. The *luminosity normalised prescale*

weights, $\overline{w}$, were calculated as [74]:

$$\overline{w} = \frac{\sum_{k=1}^{N_{runs}} L_k}{\sum_{k=1}^{N_{runs}} L_k \cdot P_{jk}}, \tag{4.3}$$

with L_k the integrated luminosity of run k.

The weight $\overline{w}$ is applied to all $D^*\mu$ events in Monte Carlo to account for the L1 prescaling. The obtained weights for HERA I and II run periods are shown in tables 4.6 and 4.7.

Subtriggers combination	S19	S22	S56	$\overline{w}$	
				1999e^-	**1999e^+/2000**
1	on	off	off	0.96	0.97
2	off	on	off	0.96	0.97
3	off	off	on	1.00	1.00
4	on	on	off	0.96	0.98
5	on	off	on	1.00	1.00
6	off	on	on	1.00	1.00
7	on	on	on	1.00	1.00

Table 4.6: *Possible combinations of the used subtriggers and the luminosity normalised prescale weights $\overline{w}$ for 1999/2000 run period.*

Subtriggers combination	S19	S23	$\overline{w}$		
			2004	**2005**	**2006**
1	off	on	0.99	0.85	0.91
2	on	off	0.99	0.96	0.95
3	on	on	1.00	0.97	0.97

Table 4.7: *Possible combinations of the used subtriggers and the luminosity normalised prescale weights $\overline{w}$ for 2004-2006 run period.*

4.2 Description of the Kinematic Variables

In this section, the distributions of the number of D^* mesons $N(D^*)$, as a function of the kinematic variables of interest, are presented. The number of D^* mesons $N(D^*)$ is obtained by fitting the Δm distribution (see section 3.4.1) in bins of the considered variable. Due to reduced statistics in data, the mean and sigma of the Gaussian function in the fit are fixed to the values obtained by fitting the total data sample. For Monte Carlo, where statistics is not an issue, all the fit parameters are free.

Apart from direct processes, PYTHIA contains also resolved ones (see section 1.4). These two types of processes are generated separately. For the comparison with data, the number of D^* mesons obtained for the direct and the resolved case are normalised to the individual cross sections of the processes, and then added.

The data are compared to the sum of charm and beauty events, added according to the fractions:

$$f_c = (71 \pm 7)\% \quad (4.4)$$
$$f_b = (26 \pm 6)\%. \quad (4.5)$$

The method used to extract these fractions will be described in chapter 5. The statistical error band is obtained with the error propagation formula and considering the correlation factor between the two fractions[3].

The important variable in untagged photoproduction is the inelasticity y_{JB}, which in this analysis is calculated with the Jacquet-Blondel method (see section 3.3). The distribution of this variable is shown in figure 4.2. Monte Carlo describes reasonably the data.

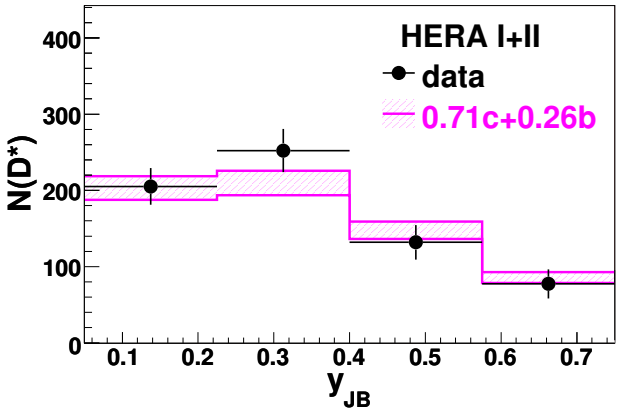

Figure 4.2: *Distribution of inelasticity y_{JB}: comparison of data (black points) to the sum of charm and beauty PYTHIA Monte Carlo (hashed regions), for HERA I and II run periods. Monte Carlo is normalised to data. The band indicates the statistical errors.*

When selecting D^* meson candidates, cuts were applied on their pseudorapidity and transverse momentum. Their distributions are presented in figure 4.3. The D^* meson is reconstructed via its decay products. Since the D^* is highly correlated (see figure 3.1) with the slow pion π_s, its distributions are shown in figure 4.4.

The distributions of the muon kinematic variables on which cuts are applied during selection, the momentum and the pseudorapidity, can be seen in figure 4.5. Similar to π_s, the muon track length, the distance of closest approach $d'_{ca}(\mu)$ of the corresponding non-vertex fitted track with respect to the primary vertex, and the azimuthal angle $\phi(\mu)$ are presented.

The combined $D^*\mu$ quantities: transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, azimuthal angular difference $\Delta\Phi(D^*\mu)$, transverse momentum fraction of the $D^*\mu$ pair,

[3] If N_{data} is given by $N_{data} = f_b \cdot N_b + f_c \cdot N_c$, where f_b is the beauty fraction in data, the corresponding error is: $\sigma^2_{N_{data}} = (N_b \cdot \sigma_{f_b})^2 + (f_b \cdot \sigma_{N_b})^2 + (N_c \cdot \sigma_{f_c})^2 + (f_c \cdot \sigma_{N_c})^2 + 2 \cdot cov(f_b,\ f_c) \cdot N_b \cdot \sigma_{f_b} \cdot N_c \cdot \sigma_{f_c}$.

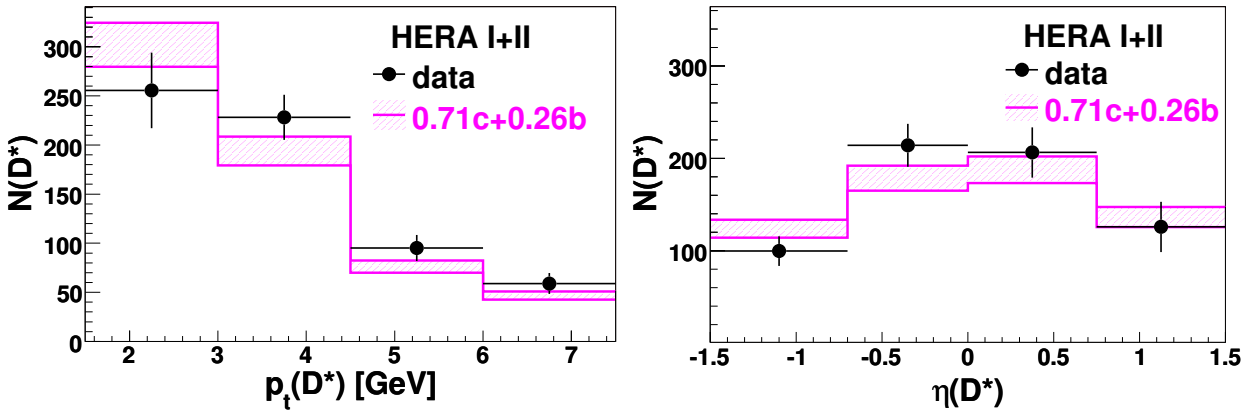

Figure 4.3: *Comparison of* D^* ***meson*** *transverse momentum* $p_t(D^*)$ *and pseudorapidity* $\eta(D^*)$ *in data (black points) to the sum of charm and beauty PYTHIA Monte Carlo (hashed regions), for HERA I and II run periods. Monte Carlo is normalised to data. The bands indicate the statistical errors.*

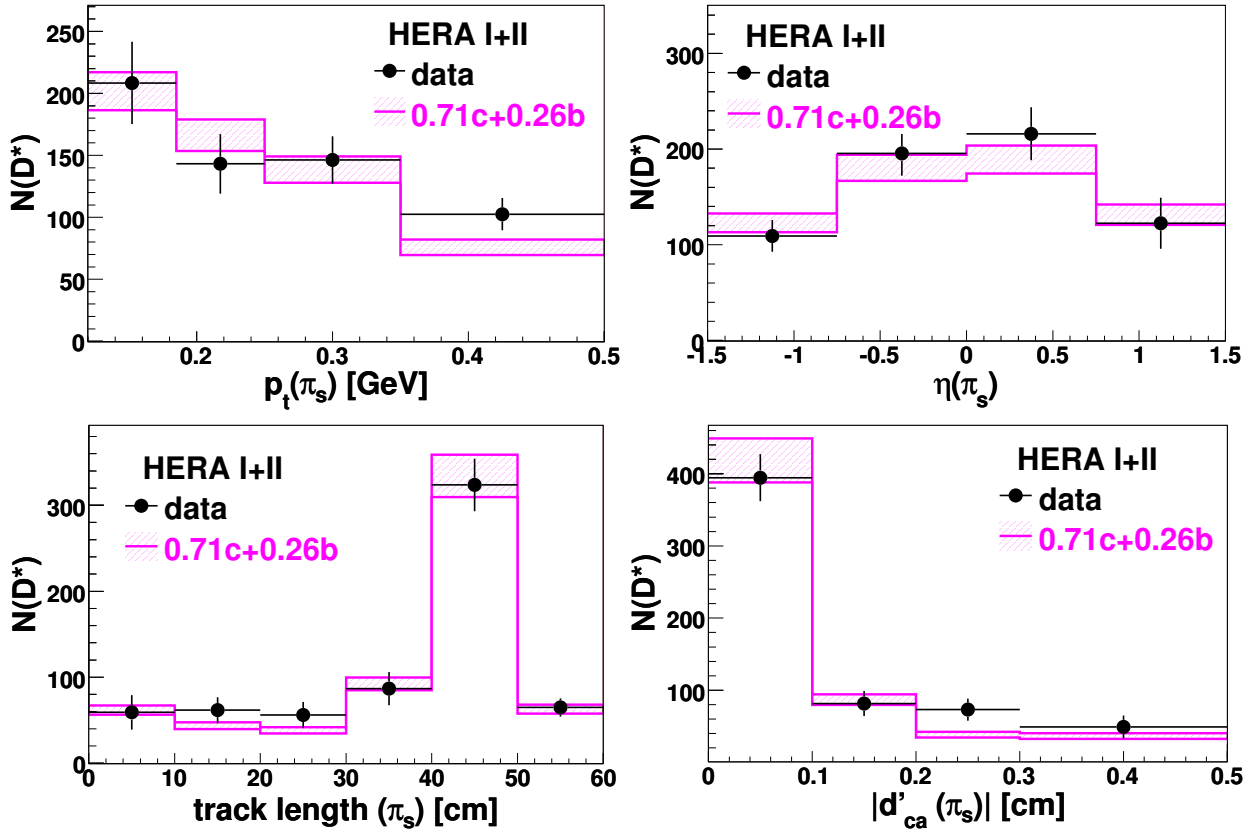

Figure 4.4: *Comparison of* ***slow pion*** (π_s) ***variables*** *in data (black points) to the sum of charm and beauty PYTHIA Monte Carlo (hashed regions), for HERA I and II run periods. The slow pion transverse momentum* $p_t(\pi_s)$*, pseudorapidity* $\eta(\pi_s)$*, radial track length and the distance of closest approach* $d'_{ca}(\pi_s)$ *are shown. The Monte Carlo distributions are normalised to data. The bands indicate the statistical errors.*

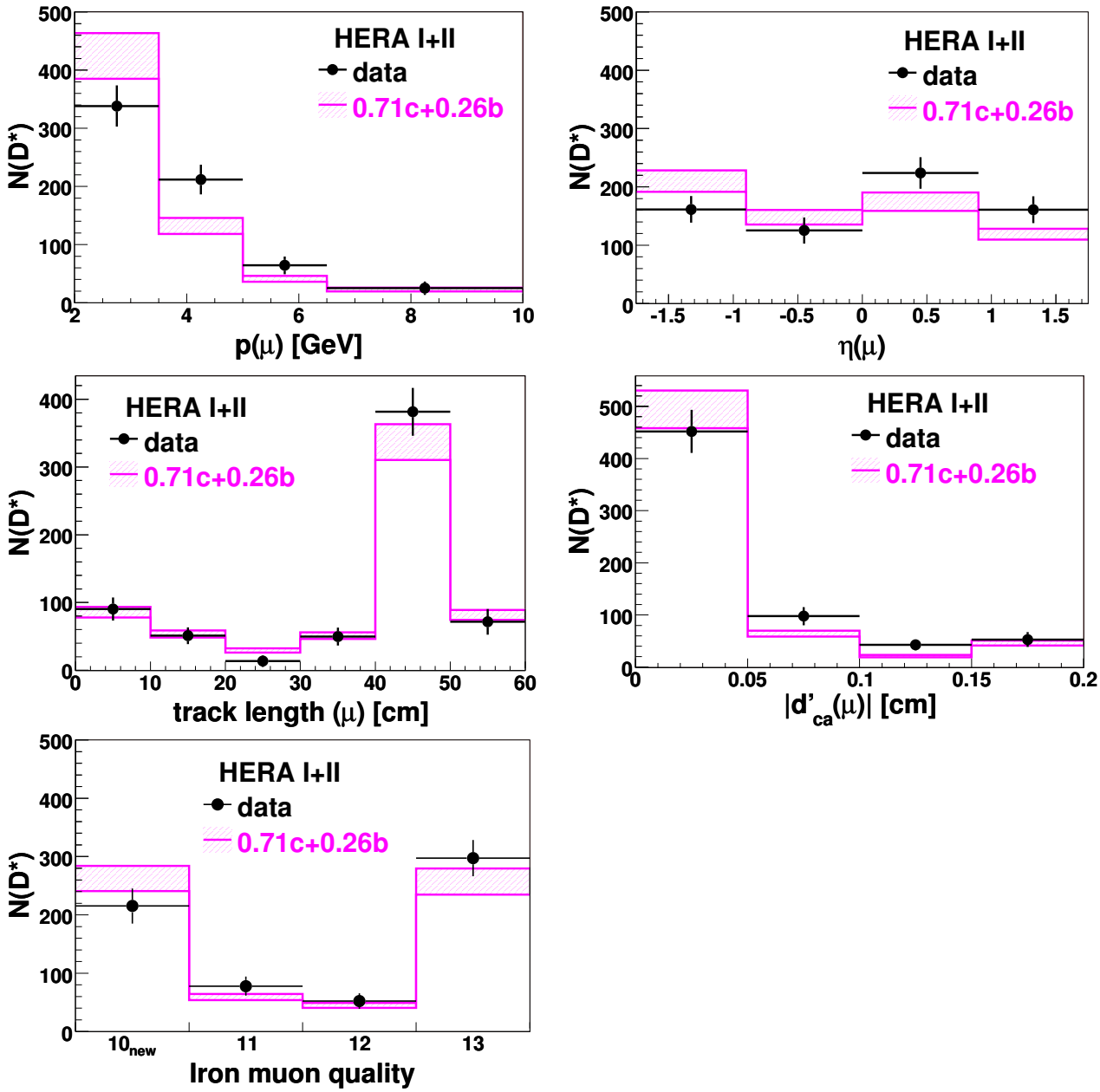

Figure 4.5: *Comparison of the* ***muon variables*** *in data (black points) to the sum of charm and beauty PYTHIA Monte Carlo (hashed regions), for HERA I and II run periods. The muon momentum $p(\mu)$, pseudorapidity $\eta(\mu)$, radial track length, the distance of closest approach $d'_{ca}(\mu)$, and the iron muons qualities are shown. The Monte Carlo distributions are normalised to data. The bands indicate the statistical errors.*

$f_{p_t}(D^*\mu)$, with respect to the transverse momentum of all hadronic final state particles with a polar angle $\theta > 10°$, rapidity $y(D^*\mu)$ and invariant mass $M(D^*\mu)$ are shown in figure 4.6.

In general, a good agreement between data and Monte Carlo is observed. Only for the muon momentum (figure 4.5) deviations are observed, since PYTHIA predicts a softer spectrum than in data.

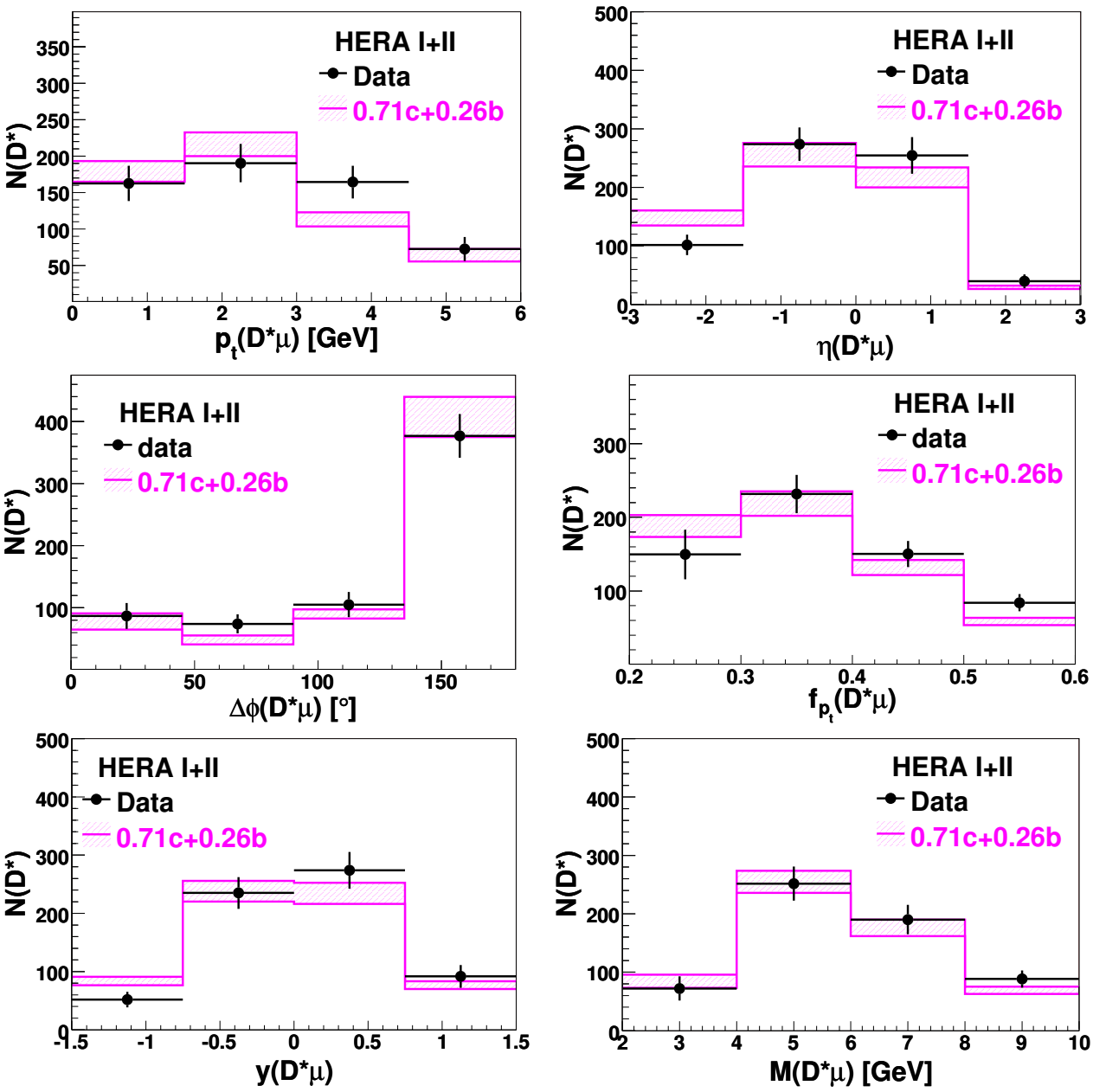

Figure 4.6: *Comparison of the* ***combined $D^*\mu$ variables*** *in data (black points) with the sum of charm and beauty PYTHIA Monte Carlo (hashed regions), for HERA I and II run periods. The transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, azimuthal angular difference $\Delta\phi(D^*\mu)$, transverse momentum fraction $f_{p_t}(D^*\mu)$, rapidity $y(D^*\mu)$ and invariant mass $M(D^*\mu)$ are shown. The Monte Carlo distributions are normalised to data. The bands indicate the statistical errors.*

4.3 Muon Background

The number of signal D^* events are obtained by fitting the Δm distribution, as described in section 3.4.1. This way, almost only charm events are accepted, and the background induced by light quarks (u, d, s) can be neglected. The remaining background in $D^*\mu$ events consists in **fake muons**, i.e. hadrons mis-identified as muons and muons from light quarks decay. The fake muons may result from different processes:

- The hadrons resulting from an ep interaction are most often pions and kaons. They are unstable particles, decaying predominantly into a muon and a neutrino, which escapes undetected [15]:

$$\begin{aligned} \pi^{\pm} &\rightarrow \mu^{\pm}\nu, \ (99.98770 \pm 0.00004)\%, \\ K^{\pm} &\rightarrow \mu^{\pm}\nu, \ (63.43 \pm 0.17)\%. \end{aligned}$$

  Because the decay length of these mesons is large ($c\tau = 7.8$ m for π, and $c\tau = 3.7$ m for K), they are usually stopped in the LAr calorimeter before decaying. Nevertheless, it may happen that they decay into muons in the inner detector region, before reaching the calorimeter. This probability is up to 3% for pions, depending on momentum[4] and polar angle, and up to 7% for kaons [55].

  The resulting muon track from the **in-flight decays** can be interpreted as a muon from the ep interaction point.

- Usually, the final state hadrons are stopped in the LAr, where they deposit their entire energy. However, there are cases when the hadronic energy leaks out of the calorimeter, reaches the instrumented iron and passes a few layers. The muon reconstruction program tries to identify these showers which have typically a shallow distribution [75]. This type of contribution is denoted as **punch-through**.

- The probability for a particle to interact with the detector is $P(d) = e^{-d/\lambda}$, with d the traversed distance in the material, and λ the nuclear interaction length of the medium. Therefore there is a non-zero probability for a hadron to traverse the central trackers and the calorimeters without strong interaction. At H1, the material in front of the barrel muon system varies between 4.5 λ (at $\theta = 90°$) and 9 λ (at $\theta = 35°$) [75]. As a result, the maximum probability for a hadron to reach the iron without strong interaction is approximately 0.6%. The muon reconstruction algorithms do not always distinguish between these hadronic showers and single penetrating muons. This is referred to as **sail-through** background.

For a particle with mass m, lifetime τ and momentum p, the probability to decay at distance d is:

$$P(d) = 1 - e^{-\frac{d}{\beta\gamma\tau}} = 1 - e^{-\frac{d \cdot m}{\tau \cdot |\vec{p}|}}. \tag{4.6}$$

If d and p are kept fixed, the relevant parameter is m/τ.

[4] Contributions from in-flight decays are expected to be important only at low momenta, whereas at high momenta they die out exponentially [75].

	Pion	Kaon
Mass	$m_\pi = 0.1396$ GeV	$m_K = 0.494$ GeV
Lifetime	$\tau_\pi = 2.6 \cdot 10^{-8}$ s	$\tau_K = 1.24 \cdot 10^{-8}$ s

Table 4.8: *The pion and kaon mass and lifetime [15].*

Considering the values given in table 4.8, one obtains:

$$\frac{m_K}{\tau_K} \approx 8 \cdot \frac{m_\pi}{\tau_\pi}, \tag{4.7}$$

so that the kaon has a larger probability to decay into a muon, compared with the pion, therefore also the muon mis-identification probability is larger for kaons than for pions.

However, the muons coming from these decays have mostly low momenta, and therefore they are rejected by the cut $p(\mu) > 2$ GeV.

4.3.1 Fake Muons from Pions and Kaons via $D^0 \to K\pi$ Decay

In practice it is difficult to disentangle the different contributions to the fake muons (this has been done for example in [76]). In the present thesis, the fake muons contributions are studied using the decay particles of the D^0 meson ($D^0 \to K^-\pi^+$) from the inclusive D^* untagged photoproduction sample, since this sample has similar event topology as the final data set.

Basically, the same D^* and D^0 mesons cuts as in table 3.1 were used. In addition, one of the D^0 daughters was requested to full-fill the same transverse momentum requirement as the muon: $p(K) > 2$ GeV or $p(\pi) > 2$ GeV. Kaon or pion candidates are selected and Δm distributions in bins of p, θ and ϕ of this initial sample are plotted. By fitting as described in section 3.4.1, the numbers of hadrons in the Δm peak, $N(hadrons)$, is obtained for every bin of the chosen variables distributions. Further, it is checked if the kaon or the pion are falsely reconstructed as a muon. If so, the same distributions as for the initial hadron sample are plotted, and the number of mis-identified hadrons, $N(hadrons \to \mu)$, is obtained.

The **hadron mis-identification probability** is defined as the ratio:

$$\mathcal{P}(hadrons \to \mu) = \frac{N(hadrons \to \mu)}{N(hadrons)}, \text{ where } hadrons = K,\ \pi. \tag{4.8}$$

The resulting mis-identification probabilities distributions for data and PYTHIA Monte Carlo are shown in figures 4.7 and 4.8. The probabilities depend only weakly on the type of heavy quark, the small differences being explained by the differences in the initial distributions of momentum p and polar angle θ. For pions, PYTHIA predicts a mis-identification probability approximatively two times lower than in data. Studies were performed to check if this behaviour is dependent on the 'fake' muon quality or its electric charge, but no explanation was found. The kaon probabilities are reasonably described by PYTHIA Monte Carlo.

In the following, the PYTHIA Monte Carlo will be used to estimate the muon background. To account for the discrepancies between data and Monte Carlo in the pion case, a relative systematic uncertainty of 20% will be attached to the muon background estimation (see section 6.8.1).

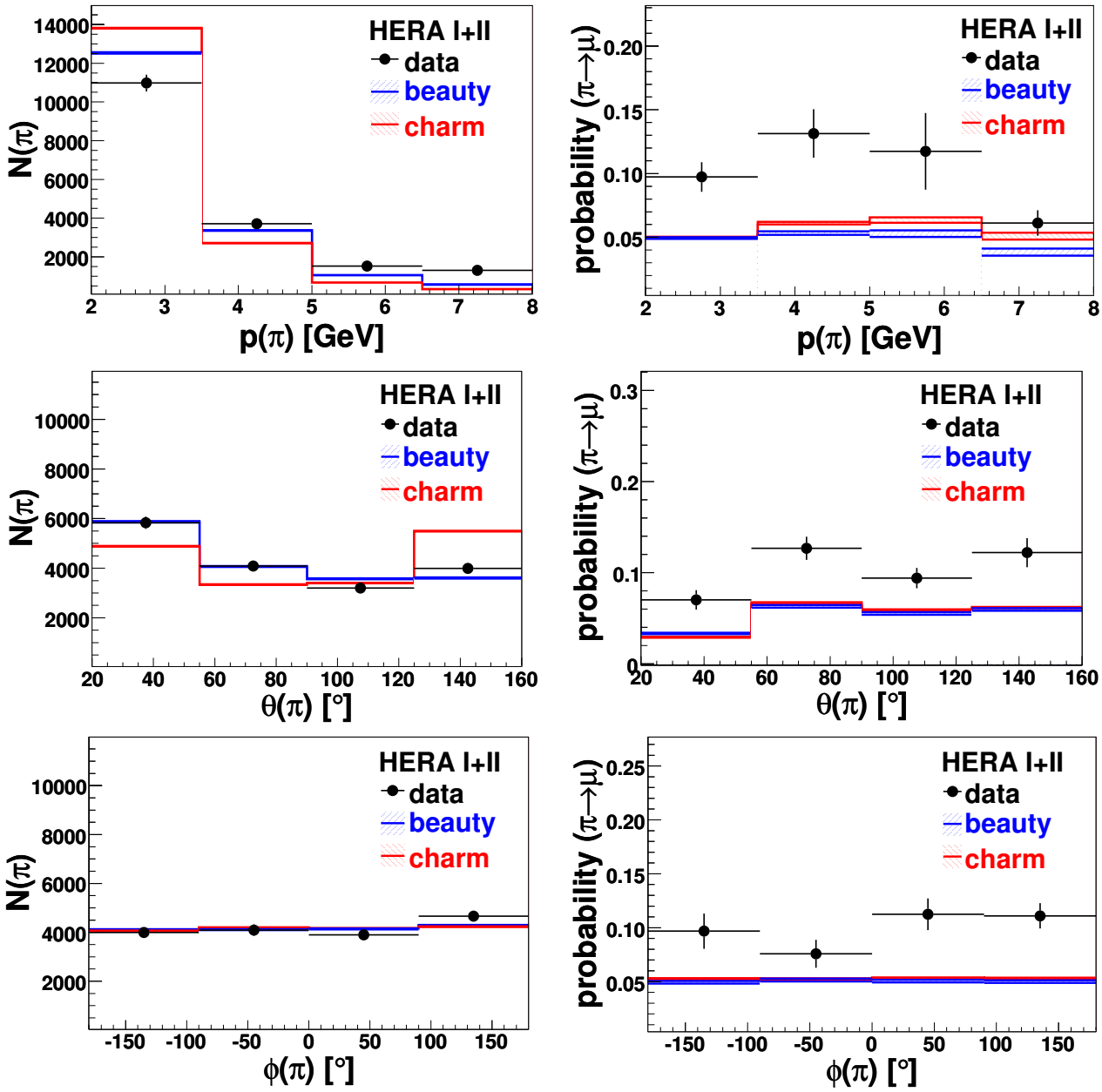

Figure 4.7: *Left: distributions of the number of* ***pions*** *(π) in the initial sample as a function of momentum p, polar angle θ and azimuthal angle ϕ. The PYTHIA Monte Carlo distributions are normalised to data. Right: probabilities for a pion to be mis-identified as a muon as a function of the same variables.*

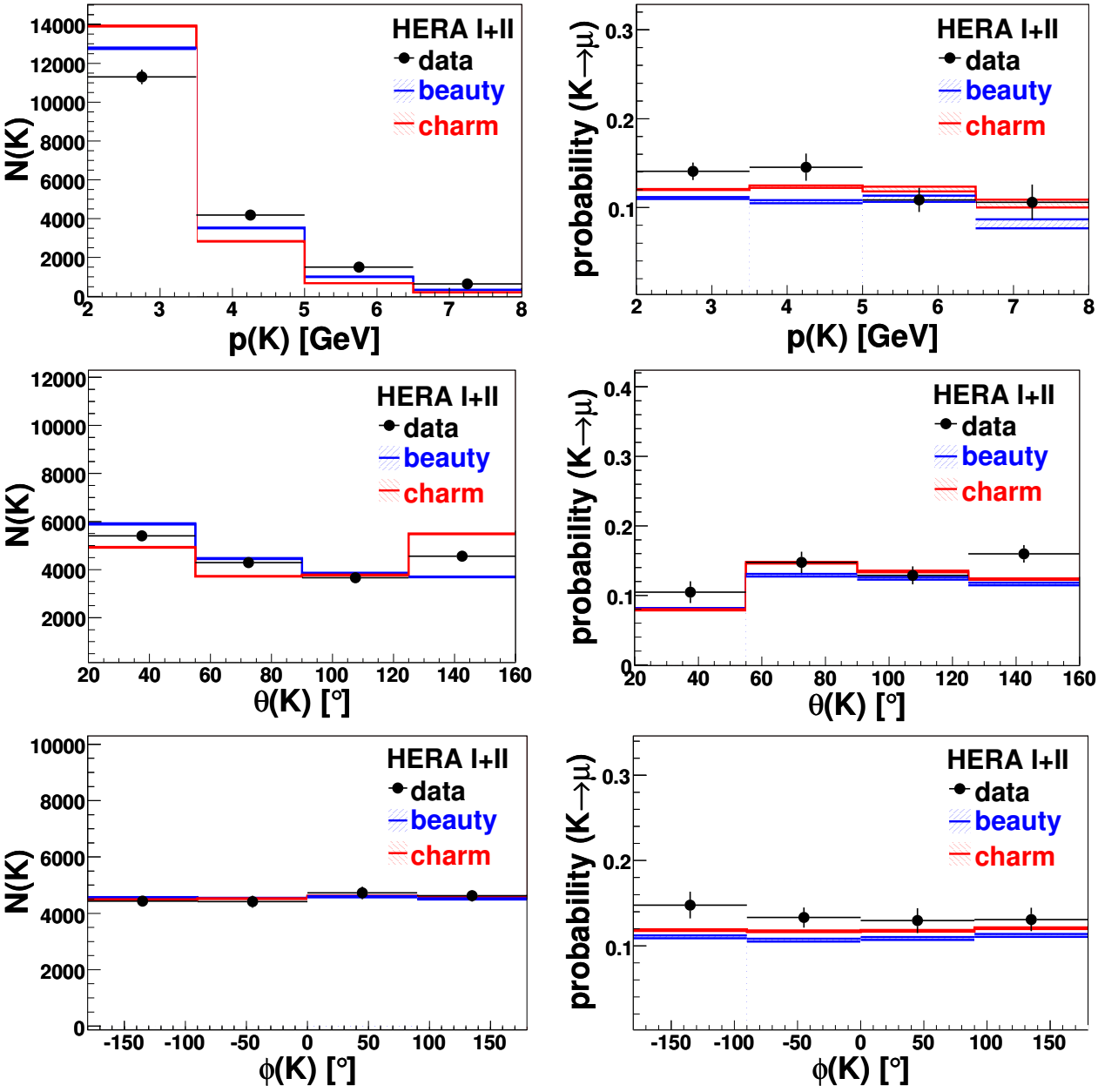

Figure 4.8: *Left: distributions of the number of* ***kaons*** *(K) in the initial sample as a function of momentum p, polar angle θ and azimuthal angle φ. The PYTHIA Monte Carlo distributions are normalised to data. Right: probabilities for a kaon to be mis-identified as a muon as a function of the same variables.*

Chapter 5

Quark Flavour Separation

In this chapter, the method used to determine the charm and beauty fractions is discussed. The method is based on the $D^*\mu$ charge and azimuthal angular correlations. An alternative procedure is presented as a cross-check.

5.1 Fit Procedure

A χ^2 fit is used to obtain the charm and beauty fractions. Due to the $D^*\mu$ selection, the background from light quark events (u, d, s) is expected to be small, and hence neglected. Therefore it is assumed that the data contain only events from charm and beauty quarks. Input to the fit are the data, charm and beauty histograms for the four correlation regions, obtained by fitting the corresponding Δm distributions as described in section 3.4.1 (see figure 5.1, left). The statistics in the Monte Carlo distributions are set to the one in data:

$$\sum_{reg\ i=1}^{4} N_{c,\,MC}^{reg\ i} = \sum_{reg\ i=1}^{4} N_{b,\,MC}^{reg\ i} = \sum_{reg\ i=1}^{4} N_{data}^{reg\ i}. \tag{5.1}$$

The χ^2 is defined as:

$$\chi^2 = \sum_{reg\ i=1}^{4} \left(\frac{N_{data}^{reg\ i} - F(f_c,\ N_{MC}^{reg\ i})}{\sigma_{N_{data}^{reg\ i}}} \right)^2, \tag{5.2}$$

The sum runs over the four correlation regions. The quantity N_{data} is the number of events in the data distribution and $\sigma_{N_{data}^{reg\ i}}$ is the corresponding error[1]. The function F depends on the charm and beauty Monte Carlo distributions obtained with the Δm fit:

$$F(f_c,\ N_{MC}^{reg\ i}) = f_c \cdot N_{c,\,MC}^{reg\ i} + f_b \cdot N_{b,\,MC}^{reg\ i}. \tag{5.3}$$

The charm and beauty fractions are obtained by minimisation: $\partial\chi^2/\partial f_q = 0$, where q stands for the quark type: charm or beauty.

[1] The uncertainty on the number of events in Monte Carlo is neglected, since it can be reduced significantly by using a large enough sample of events.

Although it was assumed that the data contain only charm and beauty events, the sum $f_c + f_b$ is not constrained to be equal to one. This is done in order to allow for the independent fluctuations of the numbers of charm and beauty events, N_c and N_b. If the purpose would be to check the partitioning of data in heavy quark events, then the fitting function given in equation 5.3 would reduce to:

$$F'(f_c,\ N_{MC}^{reg\,i}) = f_c \cdot N_{c,\,MC}^{reg\,i} + (1 - f_c) \cdot N_{b,\,MC}^{reg\,i}, \tag{5.4}$$

that is $f_b = 1 - f_c$, which introduces full correlation between the independent components. However, the intention is to determine the heavy quark fractions based on the selected $D^*\mu$ sample. If the experiment would be repeated, and new data would be fitted to the same model as the old data, slightly different results would be obtained. The fitting function given in equation 5.3 allows for independent fluctuations, such that it may happen that the sum $f_c + f_b$ is not necessarily one.

The fit results are shown in figure 5.1, right. The obtained heavy quark fractions and the corresponding numbers of events are[2]:

$$f_c = (71 \pm 7)\%, \quad N_c = 477 \pm 47 \tag{5.5}$$

$$f_b = (26 \pm 6)\%, \quad N_b = 175 \pm 39 \tag{5.6}$$

The number of heavy quark events is given by:

$$N_q = N_{data} \cdot f_q, \tag{5.7}$$

where q stands for the quark type: charm or beauty. The associated errors are obtained taking into account the errors on f_q: $\delta_{N_q} = N_{data} \cdot \delta_{f_q}$, where δ is the statistical error.

The χ^2 probability is calculated based on the χ^2 value and on the number of degrees of freedom[3]. It is the probability that a repeated experiment will return a chi-squared statistic greater than χ^2. The fit is good if the probability is of the order of 10% [77]. If the fit probability is larger than 0.001, the fit may be acceptable if the errors are non-normal, or have been moderately underestimated. Otherwise, if the fit probability is less than 0.001, the model and/or the estimation procedure are not reliable.

To determine the number of heavy quark events in the four correlation regions, the relative distribution of the initial Monte Carlo events (figure 5.1, left) is used:

$$N_q^{reg\,i} = N_q \cdot \frac{N_{q,\,MC}^{reg\,i}}{\sum_i N_{q,\,MC}^{reg\,i}} \equiv N_q \cdot x_{q,\,MC}^{reg\,i} \tag{5.8}$$

where $x_{q,\,MC}^{reg\,i}$ is the proportion of events of flavour q in region i as predicted by the Monte Carlo (see table 5.1). The fraction of charm events in region i is (see table 5.2):

$$f_{c,\,data}^{reg\,i} = \frac{N_c^{reg\,i}}{N_c^{reg\,i} + N_b^{reg\,i}}. \tag{5.9}$$

[2]Here the fit results for the HERA I+II data sample are presented. For a separate treatment of HERA I and II data, see appendix H.

[3]The number of degrees of freedom of a fit is defined as the number of fitted data points minus the number of fit parameters. In this case, the fit parameters are two: the charm and beauty fractions, f_c and f_b.

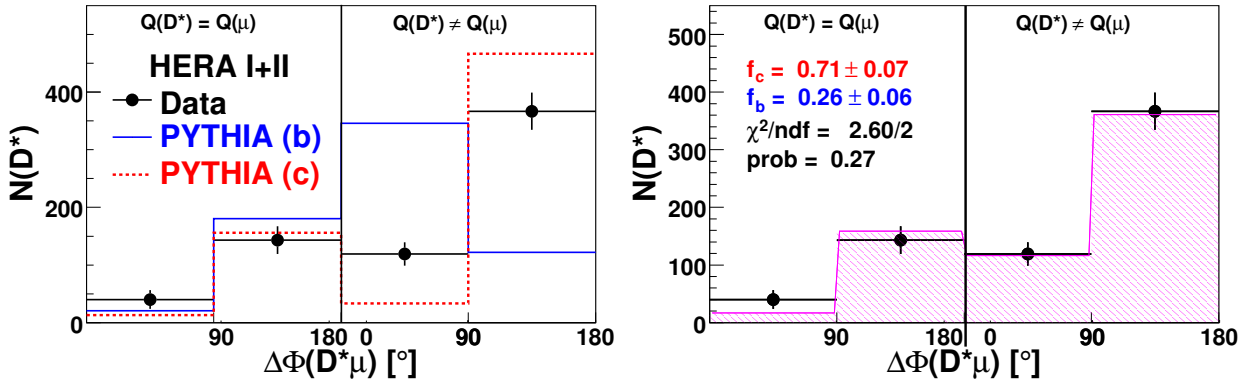

Figure 5.1: *Left: distributions of events in the four correlations regions for data (black points) and initial beauty and charm PYTHIA Monte Carlo which are input to the fit procedure. The Monte Carlo distributions are individually normalised to data. Right: distribution of events in the four correlation regions determined with the χ^2 fit (hashed region) compared with data (black points). The obtained χ^2/ndf, where 'ndf' is the number of degrees of freedom, and the fit probability are also shown.*

Region	1	2	3	4	$\sum$
$x^{reg\,i}_{c,\,MC}$ [%]	2.3	23.3	5.4	69.0	100
$x^{reg\,i}_{b,\,MC}$ [%]	3.1	27.0	51.7	18.2	100

Table 5.1: *Relative distribution of all muons events in region i, $x^{reg\,i}$, as predicted by charm and beauty PYTHIA Monte Carlo.*

Region	1	2	3	4	$\sum$
$N_c^{reg\,i}$	11 ± 1	111 ± 11	26 ± 3	329 ± 32	477 ± 47
$N_b^{reg\,i}$	5 ± 1	47 ± 11	91 ± 20	32 ± 7	175 ± 39
$f^{reg\,i}_{c,\,data}$ [%]	67 ± 5	70 ± 5	22 ± 4	91 ± 2	73 ± 5

Table 5.2: *Distribution of charm and beauty events in the four correlation regions. The fraction of charm events is also shown.*

5.2 Signal Muons

The distributions of $N(D^*\mu)$ in the different correlation regions contain also background events, coming from different sources. By fitting the Δm distribution, only signal D^* events were considered, but there are also contributions due to mis-identified muons (see section 4.3.1).

To approximate the *signal muons* contributions, the Monte Carlo samples are used. This is a good assumption as long as the Monte Carlo describes the muon background in data. This has been checked in section 4.3. The observed discrepancies in the muon background description will be accounted for when estimating the systematic uncertainties associated

with the measurement (section 6.8.1). In Monte Carlo, one can make at generator level a distinction between muons of interest[4], i.e. muons coming from heavy hadrons, and other muons. If at reconstructed level the event contains signal muons, the event is labelled 'signal'. If only signal $D^*\mu$ events are considered:

$$N_q^{sig,\,reg\,i} = N_q \cdot x_{q,\,MC}^{reg\,i} \cdot \frac{N_{q,\,MC}^{sig,\,reg\,i}}{N_{q,}^{reg\,i}} \equiv N_q \cdot x_{q,\,MC}^{reg\,i} \cdot x_{q,\,MC}^{sig,\,reg\,i} \tag{5.10}$$

with $x_{q,\,MC}^{sig,\,reg\,i}$ the proportion of events of flavour q from region i which contain signal muons.

Figure 5.2 shows the predicted signal muon contributions to the four correlation regions for HERA I and II run periods. For charm production it is expected to have signal $D^*\mu$ events only in region 4. In this region, the particle having a charge different from the one of the D^* meson is in most cases a pion, which still can be mis-identified as a muon. In region 2, kaons, which have the same charge as the D^* meson, are mostly mis-identified. Since the muon mis-identification probability is larger for kaons than for pions (see section 4.3), the muon background is larger in region 2 than in region 4. For beauty production, signal muon events are expected in regions 2, 3 and 4. PYTHIA Monte Carlo predicts a much smaller contribution of the muon background to beauty than to charm events.

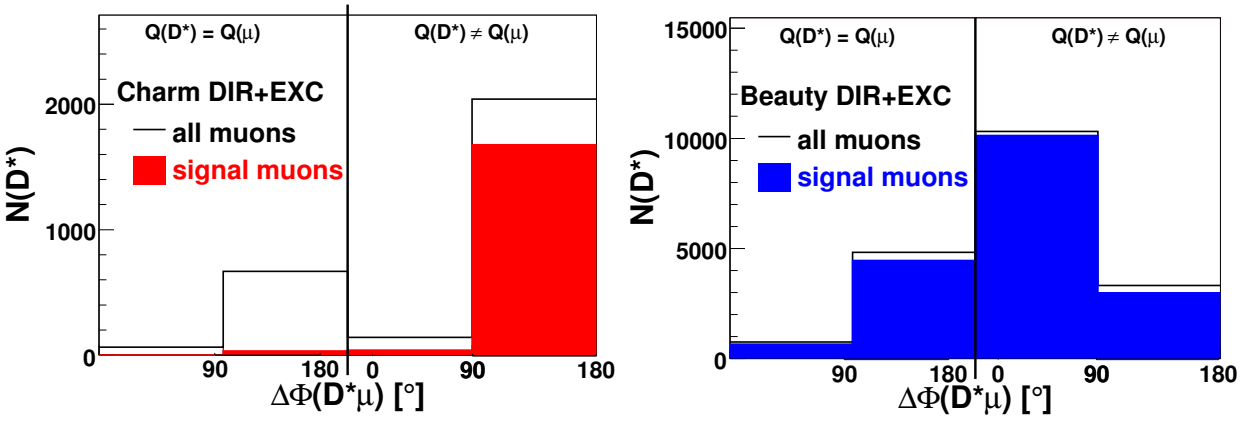

Figure 5.2: *Signal muon contributions to correlation regions in case of charm (left) and beauty (right) production as predicted by PYTHIA Monte Carlo. The label 'DIR' ('EXC') stands for direct (excitation) component in PYTHIA.*

Region	1	2	3	4	$\sum$
$x_{c,\,MC}^{sig,\,reg\,i}$ [%]	0	5.1	31.4	81.9	59.3
$x_{b,\,MC}^{sig,\,reg\,i}$ [%]	86.1	92.7	97.2	92.5	94.8

Table 5.3: *Relative distribution of signal muons events in region i, $x^{sig,\,reg\,i}$, i.e. how many of all muons are signal muons, for charm and beauty PYTHIA Monte Carlo.*

[4] According to the Monte Carlo particle numbering scheme [78], all charm and beauty mesons have cipher 3 and 4, respectively, on position three, counting from left to right, in their PDG index (for example, D^* (2010) has PDG index 413, B^0 meson 511 etc). Therefore muons which result from the decay of this kind of particles are considered to be 'signal muons'.

Region	1	2	3	4	$\sum$
$N_c^{sig,\,reg\,i}$	0	6 ± 1	8 ± 1	270 ± 27	283 ± 28
$N_b^{sig,\,reg\,i}$	5 ± 1	44 ± 10	88 ± 20	30 ± 7	166 ± 37
$f_{c,\,data}^{sig,\,reg\,i}$ [%]	0	11 ± 2	8 ± 2	90 ± 2	63 ± 3

Table 5.4: *Distribution of charm and beauty events in data, in the four correlation regions, after the muon background correction. The corresponding charm fractions are also shown.*

The obtained relative distributions are given in table 5.3. Around 96% of the signal $D^*\mu$ events from charm production are in region 4, and 53% of the total beauty events in region 3. After muon background correction, the total charm fraction decreased from 73% to 63% (see table 5.4). The number of charm and beauty events in the four correlation regions, after the muon background correction, will be later used to measure the corresponding cross sections (see chapter 6).

5.3 Comparison with ΔR Method

To check the obtained charm and beauty fractions, instead of the correlation regions, the following $D^*\mu$ variable (see ZEUS analysis [76]) was used:

$$\Delta R(D^*\mu) = \sqrt{\Delta\Phi^2_{D^*\mu} + (\eta_\mu - \eta_{D^*})^2}. \tag{5.11}$$

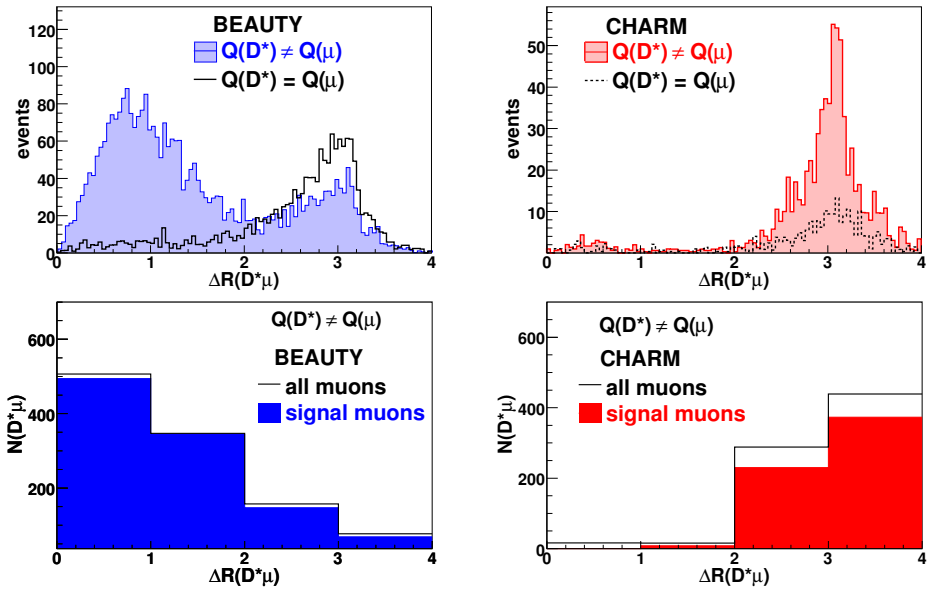

Figure 5.3: *Top: ΔR distributions for beauty (left) and charm (right) production. Note that the distributions are not background subtracted. Bottom: ΔR distributions for beauty (left) and charm (right) production, with the D^* meson and the muon having different charges.*

The ΔR distributions are shown in figure 5.3, top. For beauty production, a peak is observed at low ΔR values for oppositely charged D^* mesons and muons. This is due to the case in which both particles come from the same b quark. Higher ΔR values are mainly produced when the two particles come from different b quarks, and the muon is the result of a D meson decay [76]. For charm, high ΔR values correspond to D^* mesons and muons coming from different c quarks. It is obvious that the best separation power is obtained if the opposite charge sample is considered. The initial beauty and charm distributions, with the corresponding signal muon content, are presented in figure 5.3, bottom.

The same χ^2 fit is performed this time on the ΔR distribution for the opposite charge sample. The result of the fit can be seen in figure 5.4.

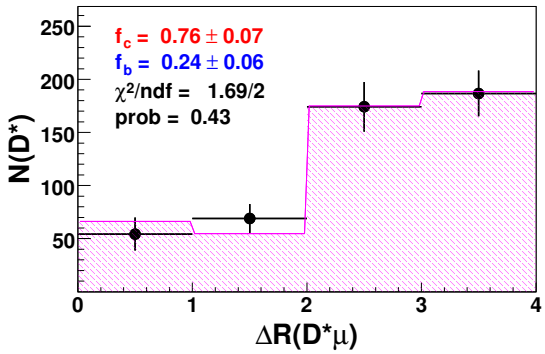

Figure 5.4: *Result of the χ^2 fit in case of ΔR distributions with the D^* meson and the muon carrying opposite charges.*

The charm fractions obtained by fitting the ΔR distribution for opposite charge $D^*\mu$ pairs, corresponding to correlations regions 3 and 4, are compared with the default fit in table 5.5. An agreement, within the errors, is observed between the two results. However, since by using the correlation regions, charm and beauty fractions for all data sample, and not only for the opposite charge case, are obtained, the former method is preferred. It is used in the next chapter to measure total visible charm and beauty cross sections.

Fit	**Correlation region**	N_c	N_b	$f_{c,\,data}$ [%]
Before muon background correction				
Correlation	all	477 ± 47	175 ± 39	73 ± 5
region	3+4	353 ± 32	123 ± 21	$\mathbf{74 \pm 4}$
ΔR	3+4	366 ± 35	118 ± 27	$\mathbf{76 \pm 4}$
After muon background correction				
Correlation	all	283 ± 28	166 ± 37	63 ± 3
region	3+4	278 ± 27	118 ± 21	$\mathbf{70 \pm 4}$
ΔR	3+4	300 ± 29	115 ± 26	$\mathbf{72 \pm 5}$

Table 5.5: *Comparison between different fit methods, before and after the muon background correction.*

Chapter 6

Total Charm and Beauty Cross Sections

The goal is to measure the cross sections of the processes $ep \to eq\bar{q}X \to eD^*\mu X$, where $q = c$ or b. The analysis is restricted to the kinematic range selected for the reconstruction of particles (*visible range*, see table 6.1) and no extrapolation to the full phase space is attempted.

First, the quantities needed to determine the cross sections in data are discussed. The measurements are presented in section 6.8. The sources of systematic uncertainties are discussed in section 6.8.1. The determination of the visible cross section in PYTHIA Monte Carlo is presented in section 6.9. Finally, a comparison of the measurements with other H1 and ZEUS analyses is done.

Photoproduction	$Q^2 < 1$ GeV^2 $0.05 < y < 0.75$
D^*	$p_t(D^*) > 1.5$ GeV $\lvert\eta(D^*)\rvert < 1.5$
μ	$p(\mu) > 2.0$ GeV $\lvert\eta(\mu)\rvert < 1.735$

Table 6.1: *Definition of the visible range in this analysis.*

6.1 Cross Section Definition

The total charm and beauty cross sections are given by the formula:

$$\sigma^q_{vis}(ep \to eq\bar{q}X \to eD^*\mu X) = \frac{N^q(D^*\mu)}{\epsilon_{L4} \cdot \epsilon^q \cdot \mathcal{BR}(D^* \to K\pi\pi_s) \cdot \mathcal{L}_{presc}} \quad (q = c \text{ or } b), \qquad (6.1)$$

where:

- $N^q(D^*\mu)$ is the number of $D^*\mu$ events from charm or beauty production, obtained after muon background subtraction;
- ϵ_{L4} is the L4 efficiency;
- the total flavour dependent efficiency ϵ^q:

$$\epsilon^q = \epsilon^q_{trig} \cdot \epsilon^q_{rec} \cdot \epsilon^q_{f_{p_t}(D^*\mu)}, \tag{6.2}$$

  is the product of the trigger, reconstruction and the $f_{p_t}(D^*\mu)$ cut efficiencies;
- $\mathcal{BR}(D^* \to K\pi\pi_s)$ is the branching ratio of the decay channel $D^* \to D^0\pi_s \to K\pi\pi_s$;
- $\mathcal{L}_{presc}$ is the prescale-corrected luminosity (see section 6.2).

Dividing by the branching ratio, the cross section is extended to inclusive D^* meson production, thus accounting for the experimentally observed decay channel. To account for the finite performance of the triggers and of the L4 classification, the trigger and the L4 efficiencies are introduced in the definition of the cross section. Due to acceptance and limited resolution of the detector, not all the generated events are also reconstructed in the visible range. This is accounted for by the reconstruction efficiency. By considering the efficiency of the $f_{p_t}(D^*\mu)$ cut, the cross section is extrapolated to the visible range.

6.2 Prescale-Corrected Luminosity

As explained in section 2.6, the level 1 subtriggers may be prescaled. This needs to be considered when determining the luminosity for the chosen subtriggers.

In Monte Carlo, the level 1 triggers prescales are not simulated. This is corrected for by applying average prescale weights, depending on the subtrigger combinations (see section 4.1.2). Suppose n is the number of possible subtriggers combinations (see tables 4.6 and 4.7). Then, the number of $D^*\mu$ events before L1 weighting is:

$$N_{before\ L1\ weight} = \sum_n N_n, \tag{6.3}$$

with N_n the number of events[1] for a certain trigger combination n. After L1 weighting:

$$N_{after\ L1\ weight} = \sum_n \frac{N_n}{\overline{w}_n}, \tag{6.4}$$

where $\overline{w}_n$ are the average prescale weights.

The effect of the L1 prescale weighting can be quantified by introducing an average prescale factor $\bar{d}$:

$$\bar{d} = \frac{N_{before\ L1\ weight}}{N_{after\ L1\ weight}}. \tag{6.5}$$

[1]Note that, by construction, the trigger combinations do not overlap.

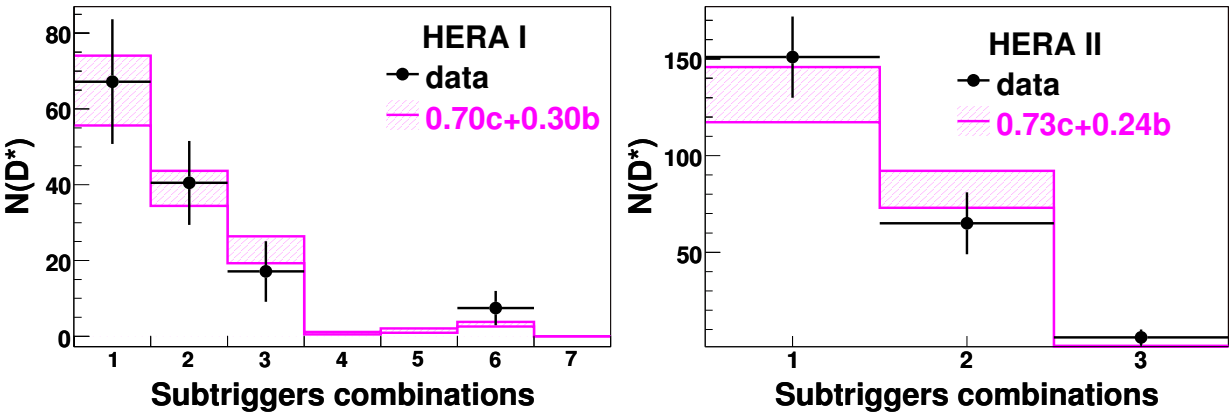

Figure 6.1: *Distribution of events with the different subtriggers combinations (see tables 4.6 and 4.7) for HERA I and II data and PYTHIA generator. The error band indicates the statistical errors. The Monte Carlo is normalised to data.*

	Average prescale $\bar{d}$	Corrected luminosity $\mathcal{L}_{corr} = \mathcal{L}/\bar{d}$ [pb^{-1}]
HERA I	1.01	69.6
HERA II	1.07	251.6
$\sum$		321.2

Table 6.2: *Average prescale factors obtained with PYTHIA charm Monte Carlo, for HERA I and II run periods, and the corresponding prescale-corrected luminosities.*

This approach is valid only if the used Monte Carlo describes the different subtrigger combinations. This was verified separately for HERA I and II run periods, since the triggers selected in the two periods are either different, or they have changed definition. The distribution of the number of D^* mesons obtained with the different subtriggers combinations is shown in figure 6.1. The data is compared to the sum of charm and beauty Monte Carlo, according to fractions[2] obtained with the method explained in chapter 5.

Using PYTHIA Monte Carlo, the average prescale factors shown in table 6.2 were obtained. Since the data sample is dominated by charm, the factors from charm Monte Carlo were used, but similar values were obtained with beauty Monte Carlo (1.03 and 1.02 for HERA I and II, respectively).

[2]For HERA I the following charm and beauty fractions were obtained: $f_c = 0.70 \pm 0.15$ and $f_b = 0.30 \pm 0.12$. Similar, for HERA II: $f_c = 0.73 \pm 0.08$, $f_b = 0.24 \pm 0.07$ (see appendix H). These values are compatible, within the errors.

6.3 Comparison of Trigger Efficiencies in Data and Monte Carlo

Before determining the trigger efficiency for the cross section measurement, one has to verify that the Monte Carlo describes the trigger efficiencies obtained in data. For this purpose, the subtrigger efficiency is calculated using **monitor subtriggers**, which are designed to be independent of the selected triggers. Since there is no such subtrigger, the efficiency is calculated with respect to the trigger elements (TE) forming the considered subtrigger. It is formed as the ratio of events triggered by both, the monitor[3] and the studied trigger element, to the events triggered by the monitor trigger alone:

$$\epsilon_{trig} = \frac{N(D^*)^{signal}_{monitor \;\&\&\; TE}}{N(D^*)^{signal}_{monitor}}. \tag{6.6}$$

The error of the trigger efficiency is the binomial statistical one:

$$\sigma_{\epsilon_{trig}} = \sqrt{\frac{\epsilon_{trig} \cdot (1 - \epsilon_{trig})}{N(D^*)^{signal}_{monitor}}}. \tag{6.7}$$

Due to reduced statistics in data, instead of taking the number of D^* mesons obtained from fitting the Δm distribution, the number of events with a D^* meson candidate in the *signal* region, i.e. $0.143\ \text{GeV} < \Delta m < 0.148\ \text{GeV}$, and a muon were considered in determining the trigger elements efficiencies[4]. The same method is applied in Monte Carlo, where the combinatorial background is small.

6.3.1 HERA I

The monitor triggers should be independent of the signal triggers. Since S0 requires only energy depositions in the SpaCal, it suits as a monitor for all considered trigger elements, except for the SpaCal trigger elements. S61 is based on the SpaCal and CJC's trigger elements, and on the z-vertex significance, so it can be used as monitor for the central muon detector trigger elements. S83 and S87 trigger photoproduction events where the electron is scattered at very small angles and detected in the electron tagger detector. They require energy deposition in the electron tagger and use signals of the CJC's triggers. In addition, S83 has a z-vertex significance condition, and S87 requires a muon in the central muon detector. Therefore S83 suits as a monitor for the muon trigger elements, and S87 for zVtx_sig.

The efficiencies are listed in table 6.3. Note that, when more than one monitor subtrigger is defined, they are used in a logical OR combination.

[3] In data, both the actual L1 and the verified L4 subtriggers are required for the monitor, whereas in Monte Carlo only the actual subtriggers are used, because the L4 trigger bits are not simulated. The L4 efficiency is separately corrected for.

[4] With this method, the background events in the given Δm region are also considered. It is assumed that the signal to background ratio is the same for the two samples needed to determine the trigger efficiency.

Only events with a muon in the barrel of the muon system were considered in determining the efficiency of the Mu_Bar trigger element. Similar, for Mu_BIO events with muons in the end-caps were taken. Because not all photoproduction events have an energy deposition in the SpaCal, an electromagnetic energy deposition of at least[5] 2.5 GeV was required for calculating the efficiency of the trigger element (SPCLe_IET > 1 || SPCLe_IET_Cen_2).

Trigger element	Monitor subtrigger	ϵ_{data} [%]
zVtx_sig > 1	35, 39, 56, 82, 85, 87, 101	89 ± 2
DCRPh_Ta	24, 25, 35, 39, 64,	99 ± 1
DCRPh_CNH	67, 76, 77, 101	94 ± 2
Mu_Bar	2, 9, 35, 37, 39, 61, 64, 67,	83 ± 3
Mu_BIO	71, 75, 76, 77, 82, 83, 84,	28 ± 3
Mu_Any	85, 100, 101, 108	76 ± 2
SPCLe_IET > 1 || SPCLe_IET_Cen_2	19, 22, 24, 25, 37, 39, 67, 75, 77, 83, 84	91 ± 4

Table 6.3: *HERA I trigger elements efficiencies.*

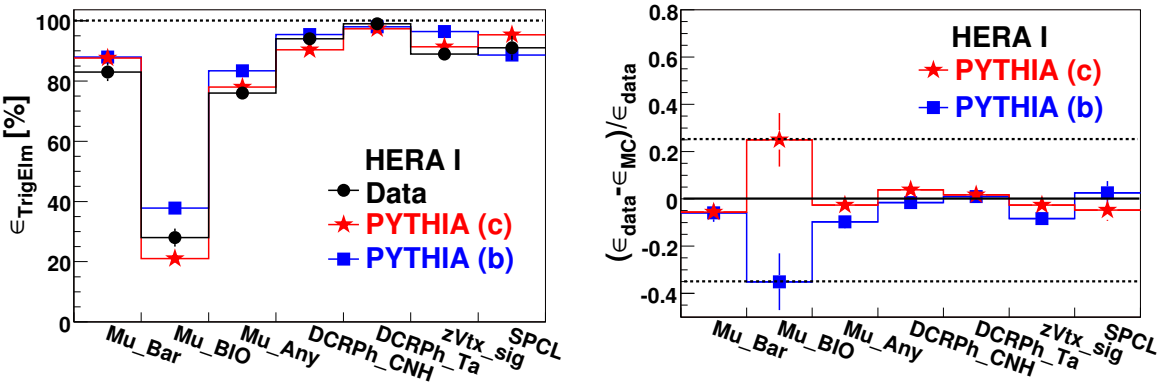

Figure 6.2: *Left: trigger elements efficiencies for HERA I data and PYTHIA Monte Carlo. The definition of trigger elements is given in appendix D. 'SPCL' is a short form of SPCLe_IET > 1 || SPCLe_IET_Cen_2. Right: relative deviations of the Monte Carlo efficiencies from the trigger elements efficiencies in data.*

Assuming the veto trigger elements are 100 % efficient [80], the subtriggers efficiencies in data are calculated as follows:

$$\epsilon_{S19} = \epsilon_{\text{Mu\_Bar}} \cdot \epsilon_{\text{DCRPh\_CNH}} \cdot \epsilon_{(\text{zVtx\_sig}>1)} = (69 \pm 3)\ \%, \quad (6.8)$$

$$\epsilon_{S22} = \epsilon_{\text{Mu\_BIO}} \cdot \epsilon_{\text{DCRPh\_CNH}} \cdot \epsilon_{(\text{zVtx\_sig}>1)} = (23 \pm 3)\ \%, \quad (6.9)$$

$$\epsilon_{S56} = \epsilon_{\text{Mu\_Any}} \cdot \epsilon_{\text{DCRPh\_Ta}} \cdot \epsilon_{(\text{SPCLe\_IET}>1\ \|\ \text{SPCLe\_IET\_Cen\_2})} = (68 \pm 4)\ \%. \quad (6.10)$$

[5] In [79] it was shown that the SpaCal trigger elements start to be efficient only from electromagnetic cluster energies above 2 GeV.

The trigger element efficiencies obtained in data are compared to the ones obtained in PYTHIA Monte Carlo (see figure 6.2, left). The relative deviations of the Monte Carlo trigger elements efficiencies from the efficiencies in data are shown in figure 6.2, right. The largest deviation, of +35% for beauty, and of −25% for charm, is observed for the trigger element Mu_BIO, which belongs to S22 (S23). The other trigger elements are described within ±10% for beauty and within ±5% for charm. It was verified what is the influence of these deviations on the final trigger efficiency. It was found that 21% (38%) of the beauty (charm) events are triggered by trigger combinations which contain S22. The resulting averaged deviations from the efficiencies in data are of 13% (15%) for charm (beauty). This will be later considered in the systematic errors associated with the trigger efficiency.

6.3.2 HERA II

The level 1 trigger element efficiencies in data, for HERA II run period, are given in table 6.4. A comparison of the results to the Monte Carlo efficiencies is shown in figure 6.3, left. The deviations of the Monte Carlo efficiencies from the efficiencies in data are presented in figure 6.3, right.

Trigger element	Monitor	$\epsilon^{\mathrm{L1}}_{data}[\%]$
Mu_Bar	9, 10, 57, 61, 115	87 ± 3
Mu_BIO		38 ± 3
DCRPh_CNH	10, 24, 57, 72, 115	100
CIP_sig > 1	9, 24, 25, 61, 64, 65	98 ± 1
FTT_mul_Tc > 2	9, 24, 25, 64, 65, 115	100
FTT_mul_Td > 0		100

Table 6.4: *HERA II level 1 subtrigger elements efficiencies.*

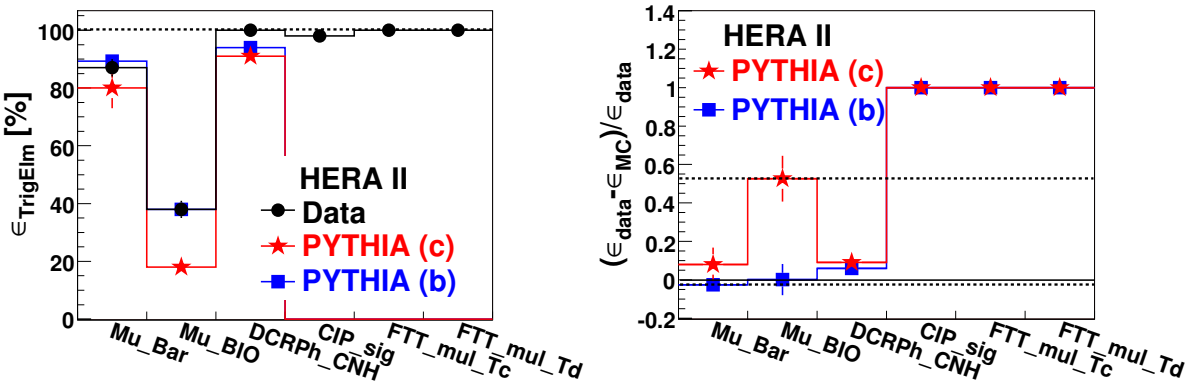

Figure 6.3: *Left: trigger elements efficiencies for HERA II data and PYTHIA Monte Carlo. Right: relative deviations of the Monte Carlo efficiencies from the trigger elements efficiencies in data.*

As in HERA I run period, the largest deviation is observed for the Mu_BIO trigger element,

Subtrigger	Monitor	$\epsilon^{\mathrm{L2}}_{trig}$ [%]	$\epsilon^{\mathrm{L3}}_{trig}$ [%]
19	9, 76, 87,	100	100
23	122 (run number $<$ 477285)	100	100

Table 6.5: *Level 2 and 3 efficiencies for the subtriggers 19 and 23 in the year 2006.*

but this time only in the charm case (+52.6%). The other trigger elements in charm Monte Carlo are described up to +9%. About 55% of the charm events are triggered by S23 alone, such that the averaged deviation of the charm trigger efficiency for the HERA II run period is 33%. For beauty, the maximal deviation of 6%, for DCRPh_CNH, is considered.

In 2005, the DCRPh-trigger was replaced by FTT (see section 2.6.1). In Monte Carlo, the subtriggers are constructed from the corresponding level 1 (L1) trigger elements. The simulation of the FTT trigger elements has not been completed by the time of performing this analysis, such that they are not included in the definition of the subtriggers. In data, they are 100% efficient[6] (see table 6.4). Thus their exclusion from the trigger definition is not expected to have any effect on the trigger efficiencies.

The CIP_sig trigger element appears in the definition of the HERA II selected subtriggers at the end of the year 2005. Since the Monte Carlo events for that period do not pass the selection cuts, no CIP_sig efficiency is shown for Monte Carlo.

In 2006, apart from the level 1 (L1) trigger conditions, level 2 and 3 conditions were implemented for subtriggers S19 and S23 (see section 3.8.2). The trigger efficiencies for the different levels are calculated as follows:

$$\epsilon^{\mathrm{L2}}_{trig} = \frac{N(D^*\mu)^{signal}|_{(monitor\ \&\&\ \mathrm{L1}trig)\ \&\&\ \mathrm{L1L2}trig}}{N(D^*\mu)^{signal}|_{monitor\ \&\&\ \mathrm{L1}trig}} \tag{6.11}$$

$$\epsilon^{\mathrm{L3}}_{trig} = \frac{N(D^*\mu)^{signal}|_{[(monitor\ \&\&\ \mathrm{L1}trig)\ \&\&\ \mathrm{L1L2}trig]\ \&\&\ \mathrm{L1L3}trig}}{N(D^*\mu)^{signal}|_{(monitor\ \&\&\ \mathrm{L1}trig)\ \&\&\ \mathrm{L1L2}trig}}. \tag{6.12}$$

The sign '&&' stands for logical 'AND'. The L1L2*trig* is on if the raw subtriggers at both levels, L1 and L2, are on. To avoid the influence of prescales, the *raw* subtriggers were considered. For the monitor triggers (*moni*), both the actual L1 subtriggers and the verified L4 subtriggers are required. Because of reduced statistics, the number of $D^*\mu$ events is determined not from the Δm fit, but by counting the events in the **signal** region. The resulting efficiencies are 100%, as it is shown in table 6.5. These high efficiency are expected for the selected data sample, since the L2 condition should be always fulfilled for muons from an *ep* interaction, and the L3 validation has an efficiency larger than 98% for isolated muons [81].

The level 1 (L1) trigger efficiencies were determined as the product of the corresponding trigger elements efficiencies. The trigger elements efficiencies for the HERA II data sample, calculated using equation 6.6, are shown in table 6.4.

[6] This is expected, since FTT_mul_Td requires at least one track with transverse momentum $p_t \geq 900$ MeV (see appendix D), which is the case for the muon track. The other FTT trigger element, FTT_mul_Tc, is on if there are at least three tracks with $p_t \geq 400$ MeV. The D^* and the muon track fulfil this p_t condition.

With the trigger definition from table 3.8 and knowing that the other trigger elements efficiencies are almost 100 %, the L1 subtrigger efficiencies in data are:

$$\epsilon_{S19} \simeq \epsilon_{\text{Mu\_Bar}} = (87 \pm 3)\%, \text{ and } \epsilon_{S23} \simeq \epsilon_{\text{Mu\_BIO}} = (38 \pm 3)\%. \tag{6.13}$$

6.4 Total Trigger Efficiencies

Having shown that the trigger element efficiencies are described, within the errors, by the Monte Carlo, the total trigger efficiencies are determined as the ratio between the number of reconstructed $D^*\mu$ pairs with a positive trigger decision, $N^{rec,\,trig}(D^*\mu)|_{vis}$, and the number of reconstructed $D^*\mu$ pairs, $N^{rec}(D^*\mu)|_{vis}$, in the visible range of the analysis, defined in table 6.1:

$$\boxed{\epsilon_{trig} = \frac{N^{rec,\,trig}(D^*\mu)|_{vis.\,range}}{N^{rec}(D^*\mu)|_{vis.\,range}}} \tag{6.14}$$

Since the purpose is to measure charm and beauty cross sections, flavour specific efficiencies are needed. In addition, the efficiencies for the c (b) dominated region 4 (3) are determined. For every flavour, direct and resolved processes are generated separately. The corresponding efficiencies are summed according to the cross section of the individual processes. The results are shown in table 6.6. The efficiencies shown in bold will be used for measuring the cross sections.

Trigger Efficiency [%]			
	Correlation Region		
	1-4	**3**	**4**
c DIR	35.8 ± 0.7	-	38.4 ± 0.9
c EXC	35.7 ± 1.0	-	37.8 ± 1.3
c DIR+EXC	$\mathbf{35.8 \pm 0.6}$	-	$\mathbf{38.1 \pm 0.8}$
b DIR	64.8 ± 0.4	69.4 ± 0.5	-
b EXC	69.7 ± 0.5	71.9 ± 0.6	-
b DIR+EXC	$\mathbf{68.2 \pm 0.4}$	$\mathbf{71.1 \pm 0.6}$	-

Table 6.6: ***Trigger efficiency*** *(in percent) for HERA I and II PYTHIA Monte Carlo. The label 'DIR' ('EXC') stands for direct (excitation) processes. The statistical errors are calculated as explained in appendix F.*

6.5 Reconstruction Efficiency

The reconstruction efficiency ϵ_{rec} is determined from PYTHIA Monte Carlo as the ratio:

$$\boxed{\epsilon_{rec} = \frac{N^{rec}(D^*\mu)|_{vis.\,range}}{N^{gen}(D^*\mu)|_{vis.\,range}}} \tag{6.15}$$

$N^{gen}(D^*\mu)$ is the number of generated $D^*\mu$ pairs, within the visible range. $N^{rec}(D^*\mu)$ is the number of reconstructed $D^*\mu$ pairs coming from heavy quarks, i.e. **without muon background**, within the visible range, but regardless if the pair is also at the generator level in the visible range or not. This way, possible migration effects, due for example to events which at generator level are not in the visible kinematic range, but on reconstructed level are, or vice-versa, are taken into account.

The obtained reconstruction efficiencies are given in table 6.7.

Reconstruction Efficiency [%]			
	Correlation Region		
	1-4	**3**	**4**
c DIR	35.5 ± 0.6	-	33.7 ± 0.6
c EXC	36.5 ± 0.9	-	33.6 ± 1.0
c DIR+EXC	$\mathbf{35.9 \pm 0.5}$	-	$\mathbf{33.7 \pm 0.5}$
b DIR	42.9 ± 0.2	44.8 ± 0.3	-
b EXC	44.5 ± 0.3	46.8 ± 0.4	-
b DIR+EXC	$\mathbf{44.0 \pm 0.2}$	$\mathbf{46.2 \pm 0.3}$	-

Table 6.7: ***Reconstruction efficiency*** *(in percent) for HERA I and II run periods. The label 'DIR' ('EXC') stands for direct (excitation) processes. The indicated statistical errors are binomial.*

6.5.1 Resolution Effects

The quality of the reconstruction of a variable x can be studied by using the **resolution**:

$$resolution(x) = \frac{x^{rec} - x^{gen}}{x^{gen}}, \tag{6.16}$$

where x^{rec} is the reconstructed variable, and x^{gen} the generated one. The resolutions of the D^* meson transverse momentum and pseudorapidity, and muon momenta and pseudorapidity can be seen in figure 6.4. The Gaussian fits with mean close to zero show that the mentioned variables are well correlated at generated and reconstruction level.

On the other hand, the y_{JB} variable is not so well reconstructed (see figure 6.5). This is due to the fact that y_{JB} is calculated from the energies of all hadronic final state particles within the detector, and it is therefore sensitive to energy calibrations [82] and particle losses.

The resolution of a certain variable x on which cuts were applied may affect the phase space of the analysis. This was studied by varying separately the cuts on the variables which define the visible range. The cuts were tightened with one σ, where σ is given by the Gaussian fit to the corresponding resolution distributions[7].

[7]Since charm is dominant in data, the resolutions predicted by charm PYTHIA Monte Carlo were considered. However, PYTHIA predicts similar resolutions for beauty production.

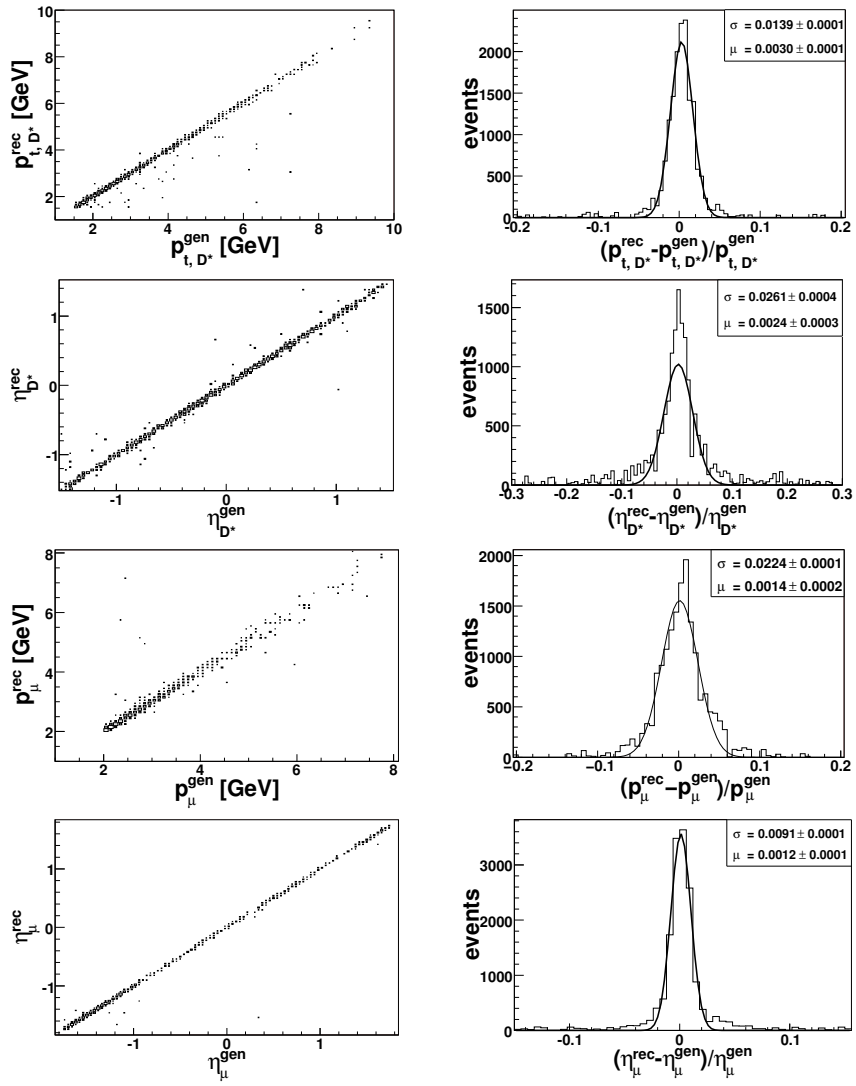

Figure 6.4: *Left: correlation between reconstructed and generated D^* meson and muon variables, as predicted by PYTHIA charm Monte Carlo. Right: D^* meson and muon resolutions distributions, fitted with a Gaussian.*

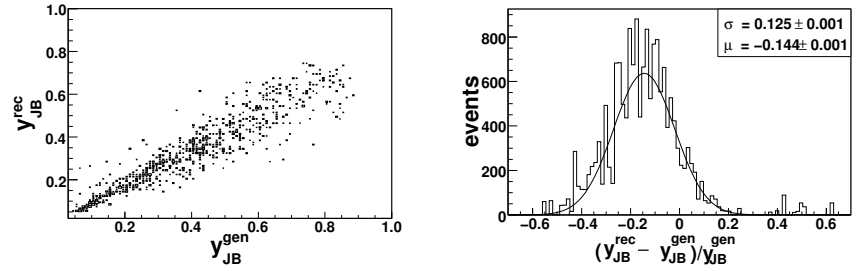

Figure 6.5: *Left: correlation between reconstructed and generated y_{JB} variable as predicted by PYTHIA charm Monte Carlo. Right: y_{JB} resolution distribution, fitted with a Gaussian.*

The relative variation of the number of $D^*\mu$ events, corrected for the muon background, was calculated as:

$$\delta_{rel}(x) = \frac{N(D^*\mu) - N(D^*\mu)|_{x,\,1\,\sigma}}{N(D^*\mu)}, \tag{6.17}$$

where $N(D^*\mu)$ is the number of $D^*\mu$ events obtained with no variations. $N(D^*\mu)|_{x,\,1\,\sigma}$ is the number resulted after tightening the cut on the reconstructed variable x with one σ.

The variations are presented in table 6.8. They are largest in case of the y_{JB} variable. This was expected, due to worse resolution, compared with the others variables.

Applied Cut	Relative Variation δ_{rel} [%]	
	Charm	Beauty
$0.056 < y_{JB} < 0.656$	8	10
$p_t(D^*) > 1.52$ GeV	4	4
$\lvert\eta(D^*)\rvert < 1.46$	3	6
$p(\mu) > 2.045$ GeV	1	3
$\lvert\eta(\mu)\rvert < 1.719$	3	2

Table 6.8: *Relative variations of the number of $D^*\mu$ events after the one σ variations of the kinematic cuts.*

6.6 Efficiency of the $f_{p_t}(D^*\mu)$ Cut

The efficiency of the $f_{p_t}(D^*\mu)$ cut is calculated as the ratio between the number of $D^*\mu$ events which survived the cut, and the ones without the cut applied:

$$\epsilon_{f_{pt}(D^*\mu)>0.2} = \frac{N(D^*\mu)|_{f_{pt}(D^*\mu)>0.2}}{N(D^*\mu)}. \tag{6.18}$$

The efficiencies are given in table 6.9.

6.7 Trigger Level 4 Efficiency

The L4 efficiency is determined from data, since the L4 classification is not simulated in Monte Carlo. It is defined as follows:

$$\epsilon_{L4} = \frac{N(D^*\mu)|_{L4\,class}}{N(D^*\mu)|_{L4\,weighted}} = (90.2 \pm 3.2)\%. \tag{6.19}$$

$N(D^*\mu)|_{L4\,class}$ is the number of $D^*\mu$ pairs for the events in which the presence of the L4 classes 15 or 16 (see section 3.6) was requested, whereas $N(D^*\mu)|_{L4\,weighted}$ is determined for the L4 weighted events (see figure 6.6). The error is calculated as explained in appendix F.

Efficiency of the $f_{p_t}(D^*\mu)$ cut [%]			
	Correlation Region		
	1-4	**3**	**4**
c DIR	99.3 ± 0.1	-	99.1 ± 0.6
c EXC	87.3 ± 1.2	-	90.1 ± 1.0
c DIR+EXC	$\mathbf{94.3 \pm 0.5}$	-	$\mathbf{95.3 \pm 0.5}$
b DIR	88.5 ± 0.3	88.0 ± 0.3	-
b EXC	80.1 ± 0.6	85.6 ± 0.6	-
b DIR+EXC	$\mathbf{82.7 \pm 0.4}$	$\mathbf{86.3 \pm 0.4}$	-

Table 6.9: ***Efficiency of the $f_{p_t}(D^*\mu)$ cut*** *(in percent) for HERA I and II PYTHIA Monte Carlo. The label 'DIR' ('EXC') stands for direct (excitation) processes. The indicated errors are calculated as explained in appendix F.*

The L4 efficiency is determined only for the total sample, and not separately for the correlation regions 3 and 4, since only in this case the Δm fit for the L4 weighted events gives reliable results[8].

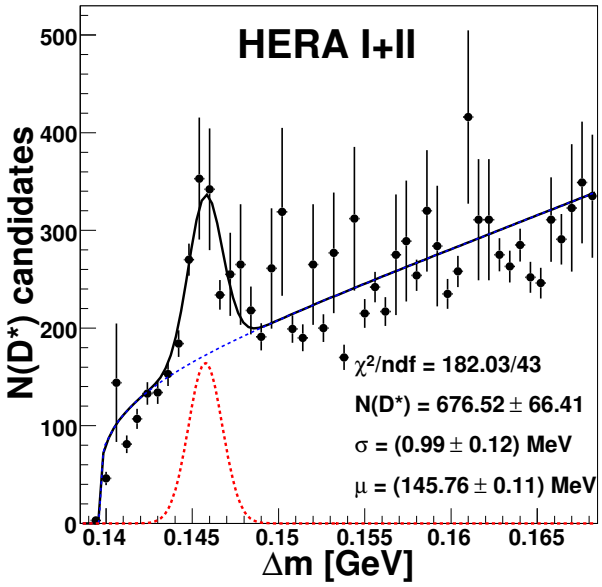

Figure 6.6: *The Δm distribution weighted with the corresponding L4 weights (no L4 class requested).*

[8]There are statistical limitations due to the weighted events. Let N be the number of events with weights w_i, $i = 1 \ldots N$. The error on the weighted events is given by $\sqrt{\sum_{i=1}^{N} w_i^2}$. For example, if there are 100 events, each with weight 1, $\sum w_i = 100$ and the associated error is 10. But if there is only one event with weight 20, $\sum w_i = 159$ and the error is 22.33, i.e. 19% relative error. In principle, the larger the weight, the bigger the error, such that a large statistics of unweighted events is needed to compensate for a single large weight.

6.8 Charm and Beauty Cross sections

Having all the necessary ingredients, the visible beauty and charm cross section are calculated (see table 6.10). The quoted errors are statistical.

	Charm		**Beauty**	
	Correlation Region		Correlation Region	
	1-4	4	1-4	3
$N(D^*\mu)$	283 ± 28	270 ± 27	166 ± 37	88 ± 20
ϵ_{rec}	0.359 ± 0.005	0.337 ± 0.005	0.440 ± 0.002	0.462 ± 0.003
ϵ_{trig}	0.358 ± 0.006	0.381 ± 0.008	0.682 ± 0.004	0.711 ± 0.004
$\epsilon_{f_{p_t}(D^*\mu)}$	0.943 ± 0.005	0.953 ± 0.005	0.827 ± 0.004	0.863 ± 0.004
ϵ_{L4}	0.902 ± 0.032			
$\mathcal{L}$ [pb^{-1}]	321.2			
$\mathcal{BR}(D^* \to K\pi\pi_s)$	0.0257 ± 0.0005			
σ [pb]	314 ± 34	296 ± 33	90 ± 20	42 ± 10

Table 6.10: *Total visible $D^*\mu$ cross sections for charm and beauty production.*

6.8.1 Systematic Uncertainties

The sources of systematic uncertainties associated with the cross section measurement are discussed below.

Definition of the Correlation Region:
Using the charge and azimuthal angular difference between the D^* and the muon, four correlation regions were defined (see table 1.6). The systematic error associated with this definition was studied by moving the cut on $\Delta\phi$ from 90° by ±10°. The analysis was performed again with the modified $\Delta\phi$ cut and the final cross sections were recalculated. This led to a relative systematic uncertainty of **3%** for charm and of **2%** for beauty cross section.

The Model Dependence:
The model dependence was studied by repeating the analysis with the CASCADE Monte Carlo model, which implements a different perturbative QCD ansatz (see section 1.6). The systematic uncertainties were calculated as $(\sigma_{PYTHIA} - \sigma_{CASCADE})/\sigma_{PYTHIA}$, where $\sigma_{PYTHIA\,(CASCADE)}$ is the cross section obtained using PYTHIA (CASCADE) Monte Carlo. The uncertainties amounted to **18%** for charm and **14%** for beauty production.

Uncertainties of the Branching Ratios Measurements
The corrections factors for the beauty scenarios (see section 4.1.1) were calculated based on the latest branching ratios [15] which are not measured exactly. The effect of their uncertainties on the cross sections measurements was studied by varying the correction factors, one at a time, with their corresponding errors given in tables 4.1 and 4.2. This resulted in a systematic error of **3%** for charm cross section, and **4%** for beauty.

Estimation of the Muon Background
The Monte Carlo simulation was used (see section 5.2) to estimate the muon background. The muon background plays a role in determining the number of signal beauty and charm events in data, and in the reconstruction efficiency. Assuming the reconstruction efficiency is the same for signal and background muons, the muon background is involved only in the determination of the number of events. Due to the imperfect description of the muon background (see section 4.3.1), a systematic uncertainty of **20%** was assumed for charm and beauty production, respectively.

Δm Fit
The number of D^* mesons in the four correlation regions was obtained by fitting the corresponding Δm distribution, but with the mean and the width of the Gaussian function fixed to the values obtained when fitting the total sample. To study the effect of fixing the parameters, the width was varied with its error of 0.1 MeV, and the cross sections were recalculated. With this, a systematic uncertainty of **4% (3%)** was obtained for charm (beauty) production.

Particle Reconstruction
The reconstruction efficiency calculated with equation 6.15 depends on the following elements.

D^* meson and μ track reconstruction efficiency: Since the track reconstruction efficiency increases with the p_t of the particles, a relative uncertainty of 4% is assumed for the slow pion, and of 3% for the kaon, the pion and the muon [70]. Conservatively, a maximal correlation was assumed between the efficiencies of the the four particles, and an uncertainty of **13%** was obtained.

Muon identification efficiency: This was studied in [80] using elastic J/ψ mesons and it was found to be **1.5%**.

Contributions From Other Decay Channels
The number of reconstructed D^* mesons decaying in the channel $D^* \to D^0\pi_s \to (K\pi)\pi_s$ is determined by fitting the $\Delta m = m(K\pi\pi_s) - m(K\pi)$ distribution.

The D^0 meson is reconstructed from its decay daughters, by selecting two oppositely charged tracks and assigning them the pion and kaon mass, respectively (see section 3.4). Since no particle identification is done, the mass hypothesis may be wrong. As an example, for $D^0 \to K^-K^+$, the pion mass may be assigned to one of the kaons, such that the resulting D^0 meson invariant mass will be shifted from the nominal value, according to the difference between the K and π masses (see figure 6.7).

Possible sources of such *reflections* are listed in table 6.11. In [83], a contribution to the Δm signal of around 3% from $D^0 \to K^-K^+$ and $D^0 \to \pi^+\pi^-$ was found. Approximately 0.5% were attributed to $D^0 \to \pi^+\pi^-\pi^0$ and $D^0 \to K^-\mu^+\nu_\mu$. A final reflection contribution of $(3.5 \pm 1.5)\%$ is quoted. The effect of these reflections is accounted for in this analysis by a relative systematic uncertainty of **5%** [71].

Trigger Efficiencies
Deviations of the Monte Carlo trigger element efficiencies from the efficiencies in data were

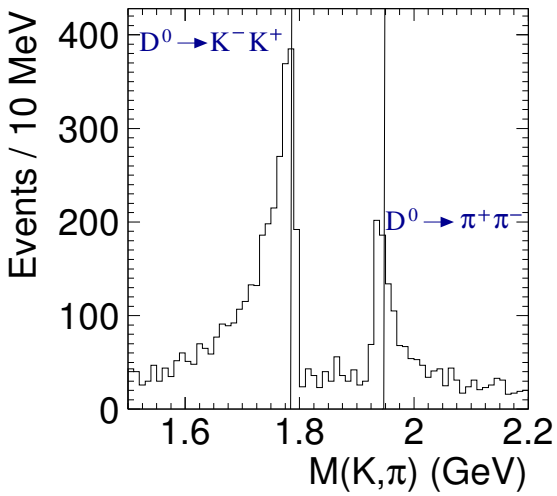

Figure 6.7: *Invariant mass $M(K,\ \pi)$ for Monte Carlo events containing $D^0 \to K^-K^+$ and $D^0 \to \pi^+\pi^-$ decays (from [83]). These events pass the cut applied on the D^0 mass window. Note that the distribution does not contain any resolution effect.*

Decay mode	**Branching ratio [%]**
$D^0 \to K^-\pi^+$	3.80 ± 0.07
$D^0 \to K^-K^+$	0.384 ± 0.010
$D^0 \to \pi^+\pi^-$	0.1364 ± 0.0032
$D^0 \to K^-\pi^+\pi^0$	14.1 ± 0.5
$D^0 \to \pi^+\pi^-\pi^0$	1.31 ± 0.06
$D^0 \to K^-e^+\nu_e$	3.51 ± 0.11
$D^0 \to K^-\mu^+\nu_\mu$	3.19 ± 0.16
$D^0 \to \bar{K}^*(892)e^+\nu_e$	2.17 ± 0.16

Table 6.11: *D^0 meson decay modes and the corresponding branching ratios [15]. The first decay mode is chosen in this analysis, but the other decays may also contribute, as explained in text.*

observed for HERA I and II run period (see table 6.12 and section 6.3), mainly due to the imperfect description of the muon triggering. The total systematic uncertainties attributed to the trigger efficiencies are shown in bold in table 6.12. They were obtained by weighting the HERA I and II efficiencies with the corresponding Monte Carlo luminosities. When

Systematic Uncertainties on ϵ_{trig} [%]		
Run period	Charm	Beauty
HERA I	13	15
HERA II	33	6
Luminosity weighted sum	**21**	**8**

Table 6.12: *Systematic uncertainties on the trigger element efficiencies.*

variating the trigger efficiencies according to their systematic uncertainties, a systematic error of **5%** (**1%**) was obtained for the charm (beauty) cross section.

Normal Resolved Contribution
The normal resolved component was neglected in PYTHIA Monte Carlo. This amounts to 3% (5%) for charm (beauty) production (see section 6.9).

Luminosity Contribution
In table 2.7, the systematic errors in the luminosity measurement for the years used in this analysis were presented. Since most of the sources are correlated for the different years, an average systematic error of **1.5%** is quoted.

Branching Ratio
The branching ratio of the D^* meson decay channel is $\mathcal{BR}(D^* \to K\pi\pi_s) = (2.57 \pm 0.05)\%$. Its relative error is therefore **2%**.

Other Possible Sources of Systematic Errors
Since the level 1 (L1) trigger prescales are not simulated in Monte Carlo, this is corrected for, as explained in section 4.1.2. The effect of this corrections is shown in figure 6.8, left, for charm events, and right, for beauty events. Because the prescale factors are small, their associated systematic uncertainty on the cross section measurement is found to be negligible and not considered here.

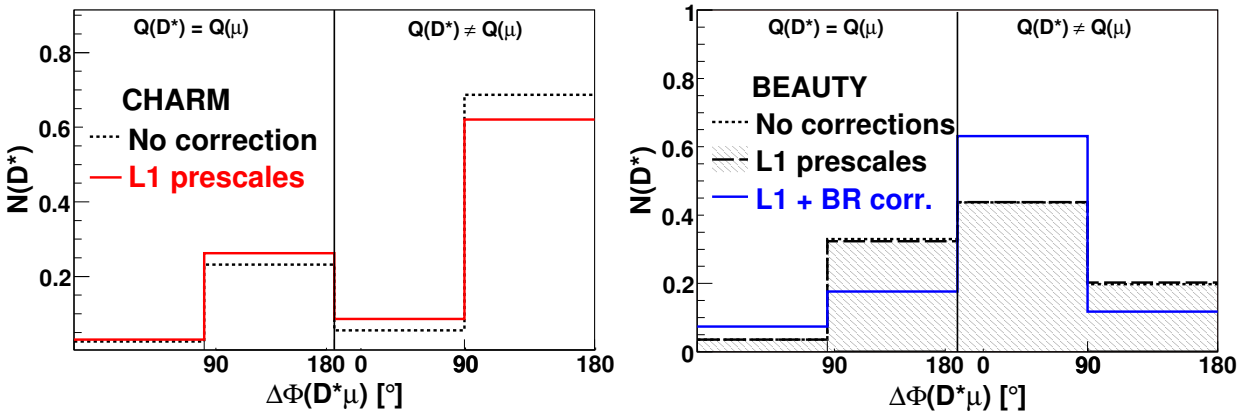

Figure 6.8: *Distributions of Monte Carlo events in the four correlations regions before and after the applied corrections. 'L1 prescales' refers to the level 1 triggers prescales, and 'BR corr.' to the beauty corrections due to the out-dated branching ratios. The distributions are normalised to one.*

Summary of the Systematic Uncertainties
A summary of the final systematic uncertainties is given in table 6.13. The largest contribution comes from the muon background estimation (20%), which is however only a rough approximation, followed by the uncertainties due to the D^* meson and muon track reconstruction. More detailed Monte Carlo studies of the kaon and pion mis-identification probability would lead to a considerable reduction of the systematic uncertainties. Significant contributions (up to 18%) are obtained also due to the model dependence.

	Source	Variation	ϵ_{syst} [%]	
			Charm	**Beauty**
1	Definition of the correlation regions	$\Delta\phi(D^*\mu) \pm 10°$	3	2
2	Model dependence	Cascade	18	14
3	Branching ratios for b scenarios	PDG	3	4
4	Δm fit	$\sigma \pm 0.1$ MeV	4	3
5	Trigger efficiency		5	1
6	Normal resolved contribution		3	5
7	D^* and μ track reconstruction		13	
8	μ identification		1.5	
9	Estimation of the μ background		20	
10	D^0 meson reflections		5	
11	Luminosity		1.5	
12	$\mathcal{BR}(D^* \to K\pi\pi_s)$		2	
Total systematic error			**32**	**29**

Table 6.13: *Sources of systematic errors for the total visible charm and beauty cross sections.*

6.9 Monte Carlo Cross Sections

The Monte Carlo cross section is determined as:

$$\sigma^q_{vis,\,MC}(ep \to eq\bar{q}X \to eD^*\mu X) = \frac{N^q_{MC}(D^*\mu)}{\mathcal{L}^q_{MC}} \cdot f^{cor}(q\bar{q} \to D^*\mu) \quad q = c \text{ or } b \qquad (6.20)$$

where $N^{MC}(D^*\mu)$ is the number of $D^*\mu$ events obtained after the kinematic cuts which define the visible range of the analysis, with the muon coming from heavy quarks, $\mathcal{L}^{MC}$ is the luminosity of the corresponding Monte Carlo, and $f^{cor}(q\bar{q} \to D^*\mu)$ is the correction factor determined in section 4.1.1.

The obtained HERA I and II PYTHIA Monte Carlo cross sections are shown in tables 6.14 and 6.15. Since the predicted cross sections for the normal resolved component are about 3% for charm production, and 5% for beauty, this component is neglected in the present analysis. This is taken into account in the systematic uncertainties indicated in table 6.13.

Selection	Cross section [pb]		
	DIRECT	**EXCITATION**	**RESOLVED**
inclusive $c\bar{c}$	601189	372958	78924
D^*	265732	136000	33103
$+\ 0.05 < y < 0.75$	129704	97201	21250
$+\ p_t(D^*) > 1.5$ GeV	42417	41130	5710
$+\ \lvert\eta(D^*)\rvert < 1.5$	26570	26592	1662
$+\ \mu$	2641	2975	200
$+\ p(\mu) > 2.0$ GeV	389	414	65
$+\ \lvert\eta(\mu)\rvert < 1.735$	**147**	**115**	**7**
$+\ \Delta\Phi(D^*\mu) > 90°$, $Q(D^*) \neq Q(\mu)$ (region 4)	**133**	**93**	**5**

Table 6.14: *Inclusive **charm** cross sections predicted by PYTHIA Monte Carlo (HERA I and II). The effect of the kinematic cuts which define the visible region is shown (the $Q^2 < 1\ GeV^2$ cut is applied already in the beginning). The cross section in the charm dominated region, i.e. region 4, is also indicated.*

Selection	Cross section [pb]		
	DIRECT	**EXCITATION**	**RESOLVED**
inclusive $b\bar{b}$	3837	1223	709
D^*	2089	528	379
$+\ 0.05 < y < 0.75$	1586	457	284
$+\ p_t(D^*) > 1.5$ GeV	842	285	161
$+\ \lvert\eta(D^*)\rvert < 1.5$	712	188	53
$+\ \mu$	308	59	17
$+\ p(\mu) > 2.0$ GeV	71	44	7
$+\ \lvert\eta(\mu)\rvert < 1.735$	**56**	**15**	**4**
$+\ \Delta\Phi(D^*\mu) < 90°$, $Q(D^*) \neq Q(\mu)$ (region 3)	**23**	**11**	**1**

Table 6.15: *Inclusive **beauty** cross sections predicted by PYTHIA Monte Carlo (HERA I and II). The effect of the kinematic cuts which define the visible region is shown (the $Q^2 < 1\ GeV^2$ cut is applied already in the beginning). The cross section in the beauty dominated region, i.e. region 3, is also indicated.*

6.10 Comparison with Other Measurements

In table 6.16, the measured charm and beauty cross sections are compared to PYTHIA Monte Carlo. The results are consistent with the leading order QCD predictions, within the errors. In figure 6.9, the obtained visible charm and beauty cross sections are compared with recent H1 and ZEUS measurements.

	Total Visible Cross Section [pb]	
	Charm	**Beauty**
Data	314 ± 33 (stat) ± 101 (syst)	90 ± 20 (stat) ± 26 (syst)
PYTHIA	262	71
Data/theory	1.2 ± 0.1 (stat) ± 0.4 (syst)	1.3 ± 0.3 (stat) ± 0.4 (syst)

Table 6.16: *The measured charm and beauty photoproduction cross sections in the kinematic range $Q^2 < 1\ GeV^2$, $0.05 < y < 0.75$, $p_t(D^*) > 1.5\ GeV$, $|\eta(D^*)| < 1.5$, $p(\mu) > 2\ GeV$ and $|\eta(\mu)| < 1.735$ in comparison to predictions from leading order QCD PYTHIA Monte Carlo. Events from the years 1999-2000 (HERA I) and 2004-2006 (HERA II) were used.*

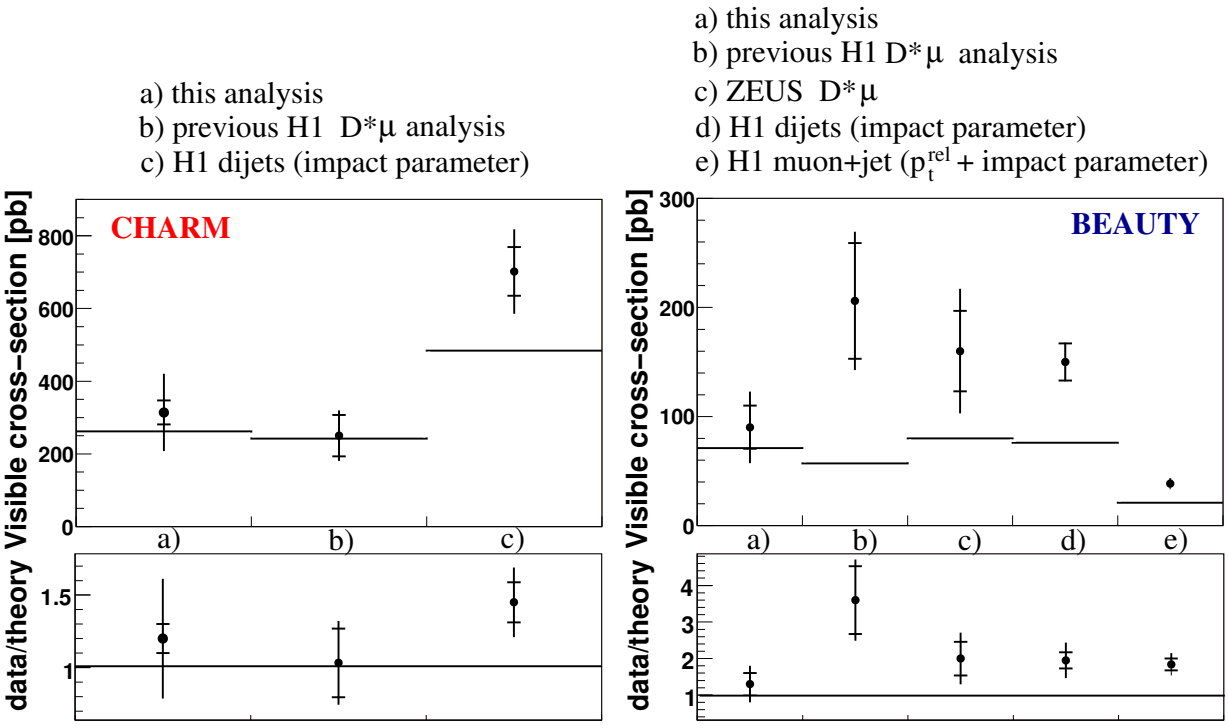

Figure 6.9: *Comparison of the total cross section of charm (left) and beauty (right) with other H1 and ZEUS analyses (see text for details). The points indicate the data, the lines the PYTHIA Monte Carlo predictions. The error bars correspond to the statistical errors. The total error is given by the statistical and systematic errors added in quadrature. The ratio data/theory is also shown.*

The H1 analyses using dijet [44] and muons plus jet [42] samples were discussed in section 1.7. Results from the previous H1 $D^*\mu$ [47] and the similar ZEUS analysis [84] are

shown as well. Note that the mentioned analyses (apart from the ZEUS result) are done in photoproduction regime, $Q^2 < 1$ GeV^2, but for different visible ranges.

In [47], the H1 collaboration measured the charm and beauty cross sections using $D^*\mu$ events from the years 1997 and 1999-2000. For a direct comparison, the present measurement was done separately for the 1999-2000 data sample[9]. The results are presented in table 6.17. The measured charm cross sections are consistent between the two analyses. A large difference is observed for the beauty cross section. The present result is closer to the predicted value, although the same tendency is observed, as for all analyses, that the Monte Carlo underestimates the data.

HERA I Cross Section [pb]		
	Charm	**Beauty**
This analysis	280 ± 64	105 ± 41
Previous H1 $D^*\mu$ analysis	250 ± 57	206 ± 53
PYTHIA	242	57
Data/theory Ratio		
This analysis	1.2 ± 0.3	1.8 ± 0.7
Previous H1 $D^*\mu$ analysis	1.0 ± 0.2	3.6 ± 0.9

Table 6.17: *Comparison between results of this analysis and the previous H1 $D^*\mu$ [47] measurement of charm and beauty cross sections for HERA I data sample. The indicated errors are statistical. The data/theory ratio is also shown.*

In this thesis, the statistical errors are reduced compared to the previous H1 $D^*\mu$ result. The PYTHIA predictions are however not identical due to the different correction factors for the relevant branching ratios in PYTHIA version 6.1, used in the previous analysis, and version 6.2.

The ZEUS beauty cross section [84] extrapolated to the same kinematic range as the one of the present analysis is:

$$\sigma_{vis,\,\gamma p}(ep \to eb\bar{b}X \to eD^{*\pm}\mu X) = 135 \pm 33\ (\text{stat})^{+24}_{-31}\ (\text{syst})\ \text{pb}, \tag{6.21}$$

which is higher than the present result, but in agreement within errors. The ZEUS measurement included the normal resolved component which was neglected here. In addition, the b quarks did not fragment directly into D^* mesons, but the kinematics of the $b \to B \to D^*$ chain was simulated in detail. Another difference consists in the inclusion of secondary muon branching fractions for D^* and μ from the same b quark (i.e. $b \to D^*D$, where $D \to \mu$).

[9] The 1997 data was not used in this analysis due to technical reasons.

Chapter 7

Differential Cross Sections

In this chapter, the normalised differential cross sections in bins of the $D^*\mu$ transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, rapidity $y(D^*\mu)$ and invariant $M(D^*\mu)$ are measured. First, the distributions of these variables for charm and beauty production are discussed. Then the normalised differential cross sections are presented. The method is identical to the one used for the total cross sections, but extending the measurement to differential variables.

7.1 Combined $D^*\mu$ Variables

The $D^*\mu$ combined variables are used as an approximation of the corresponding $q\bar{q}$ variables. The transverse momentum $p_t(q\bar{q})$, which is sensitive to NLO effects, is approximated by $p_t(D^*\mu)$. Similar to the invariant mass $M(q\bar{q})$, which in leading order corresponds to the centre-of-mass energy $\sqrt{\hat{s}}$ of the incoming photon and the gluon, the variable $M(D^*\mu)$ is defined. The quantities $y(q\bar{q})$ and $\eta(q\bar{q})$ are associated with the direction of the quark pair, and approximated by $y(D^*\mu)$ and $\eta(D^*\mu)$.

The correlations of the reconstructed $D^*\mu$ variables with the $q\bar{q}$ quantities are presented in figure 7.1. The best correlation is observed for the inelasticity y, followed by pseudorapidity η and invariant mass M. In case of transverse momentum p_t, the correlation is strongly diluted due to fragmentation effects.

The distributions of the $D^*\mu$ variables are shown in figure 7.2. The beauty Monte Carlo predicts a harder $p_t(D^*\mu)$ spectrum compared to charm Monte Carlo. The $D^*\mu$ pairs coming from c quarks are present in central and forward regions of pseudorapidity $\eta(D^*\mu)$, whereas $D^*\mu$ pairs from b quarks are concentrated only in the central region. These differences are due to the production kinematics and to the muon selection criteria. In charm production, the muon has a softer momentum spectrum compared to beauty production. As a result, the cut of 2 GeV on the muon momentum suppresses the muons in the central region of pseudorapidity [71].

No significant differences between charm and beauty production are observed for the inelasticity $y(D^*\mu)$.

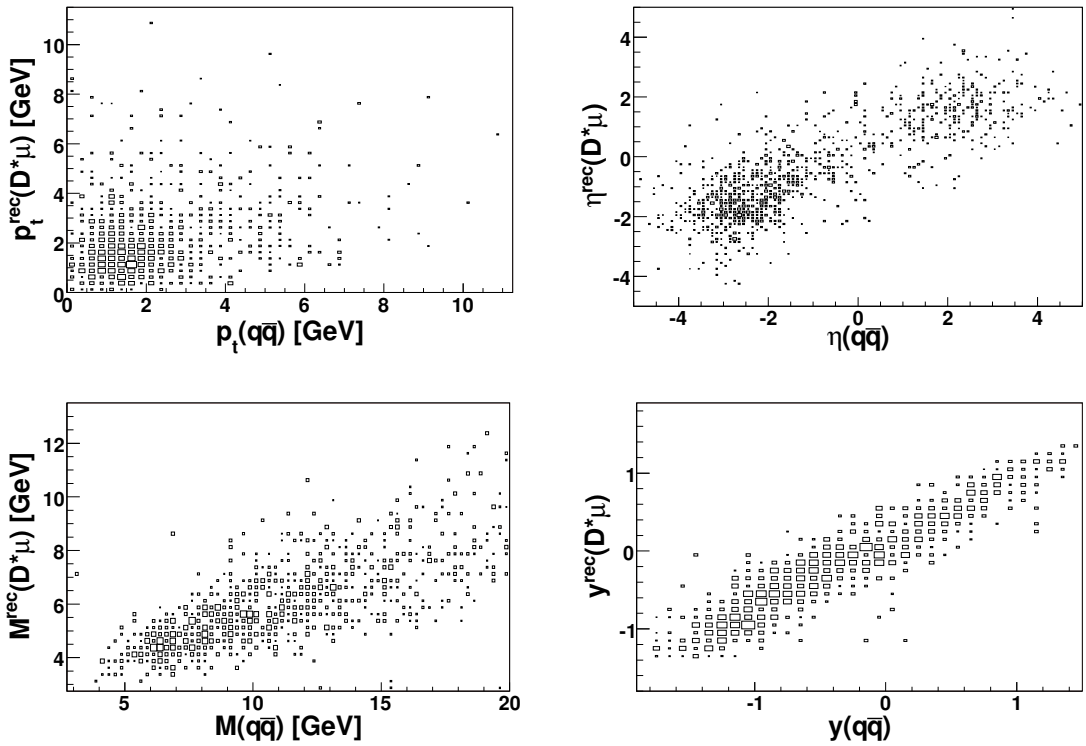

Figure 7.1: *Correlation of the reconstructed $D^*\mu$ variables with the generated $q\bar{q}$ quantities, as predicted by PYTHIA charm Monte Carlo. The distributions of transverse momentum p_t, pseudorapidity η, invariant mass M and inelasticity y are shown. Note that the distributions are not background subtracted.*

According to PYTHIA predictions (see table 5.1), around 52% of beauty events correspond to correlation region 3, i.e. D^* and μ have opposite charge and come from the same quark. This happens via decays of the type $B^0 \to D^{*-}\mu^+\nu_\mu$. Since the neutrino is not reconstructed, the invariant mass of the $D^*\mu$ pairs from beauty quarks ranges approximately from the mass of the D^* meson (2.01 GeV) to the B^0 mass (5.28 GeV). The $D^*\mu$ pairs from charm quarks have higher invariant masses.

7.2 Definition of Differential Cross Section

The differential cross section in bins of a variable x, in data, is calculated as:

$$\boxed{\frac{d\sigma^q_{vis}}{dx} = \frac{\Delta N^q(D^*\mu)}{\mathcal{L}_{presc} \cdot \mathcal{BR}(D^* \to K\pi\pi_s) \cdot \epsilon^q_{trig}(x) \cdot \epsilon^q_{rec}(x) \cdot \epsilon^q_{f_{pt}}(x) \cdot \epsilon_{L4}} \cdot \frac{1}{\Delta x}} \quad q = c \text{ or } b. \quad (7.1)$$

$\Delta N^q(D^*\mu)$ is the number of $D^*\mu$ events, **corrected for muon background**, obtained by applying the charm and beauty fractions in the considered bin. $\epsilon^q(x)$ are the flavour and bin dependent efficiencies. Δx is the width of the considered bin.

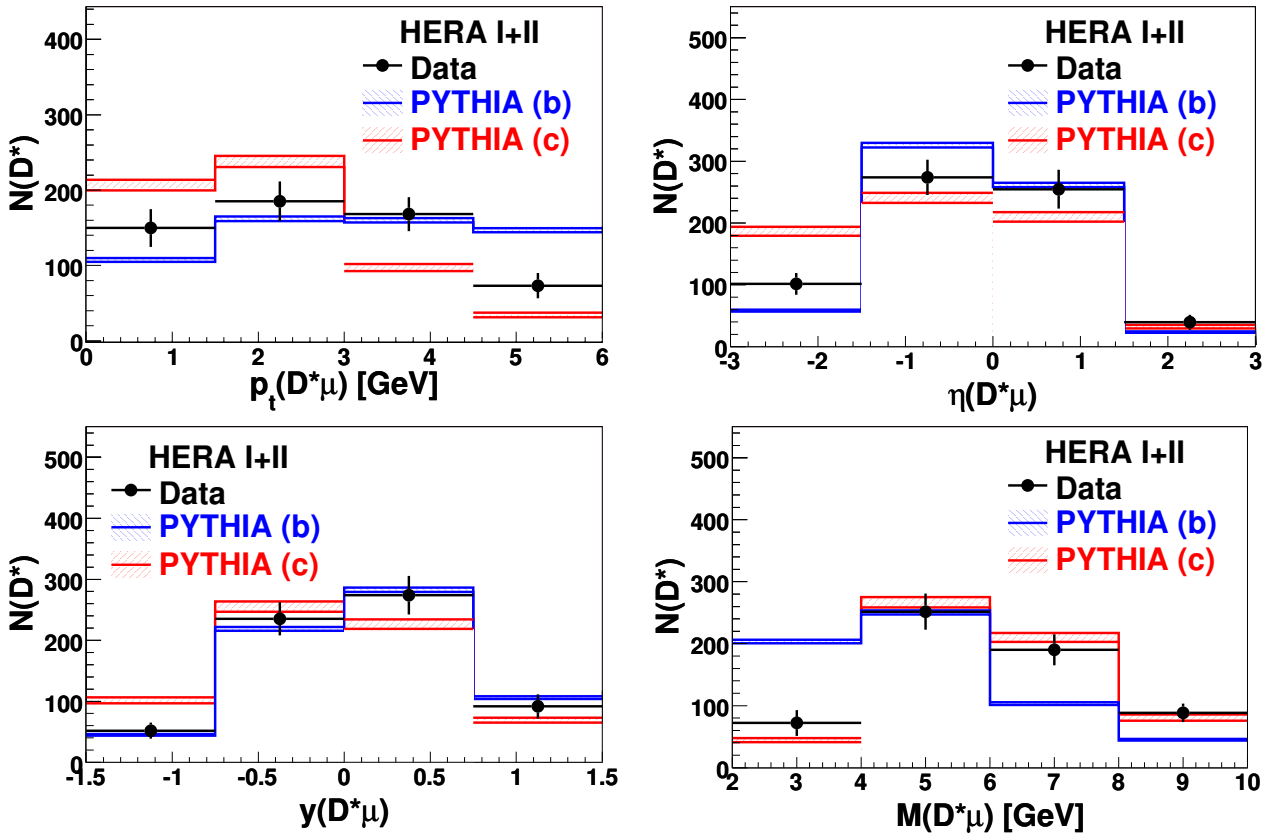

Figure 7.2: *Comparison of the* ***combined $D^*\mu$ variables*** *in data (black points) to charm and beauty PYTHIA Monte Carlo (hashed regions), for HERA I and II run periods. The transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, inelasticity $y(D^*\mu)$ and the invariant mass $M(D^*\mu)$ are shown. The Monte Carlo distributions are normalised to data. The error band indicates the statistical errors.*

Extraction of the Number of Events

For differential cross sections, the $D^*\mu$ variables are divided into two bins. This limitation is imposed by the fit procedure, which requires the presence of both charm and beauty events with sufficient statistics in the correlation regions, in order to give reliable results.

The fit procedure is identical to the procedure discussed in chapter 5, just that this time the fit is performed simultaneously in the two bins, in order to have a consistent normalisation factor. First, the Monte Carlo distributions of the four correlation regions, in the two bins, are fitted to data (see figure 7.4, left). The results of the fit are shown in figure 7.4, right. In case of empty bins, the statistics was too low to determine the number of D^* mesons with the Δm fit.

The fractions of charm (f_c) and beauty (f_b) events in data are obtained. The total number of heavy quark events in data is calculated as:

$$N_q = N_{data} \cdot f_q \quad (q = c \text{ or } b), \tag{7.2}$$

where N_{data} is the total number of events in data. The relative distributions of Monte Carlo

events are calculated:

$$x_q^{bin} = \frac{\sum_{i=1}^{4} N_{q,\,bin}^{reg\,i}}{\sum_{bin=1}^{2}\sum_{i=1}^{4} N_{q,\,bin}^{reg\,i}},\ q = c \text{ or } b. \tag{7.3}$$

$N_{q,\,bin}^{reg\,i}$ is the number of Monte Carlo events in correlation region i in one bin. The sums in the denominator run over the four correlation regions and over the two bins of the kinematic variables of interest.

The numbers of c and b events in a bin of a variable x are given by:

$$N_q^{bin} = x_q^{bin} \cdot N_q. \tag{7.4}$$

Similar, the numbers of c and b *signal* events, i.e. corrected for the muon background, is:

$$N_{q,\,sig}^{bin} = x_{q,\,sig}^{bin} \cdot N_q^{bin}, \tag{7.5}$$

where

$$x_{q,\,sig}^{bin} = \frac{\sum_{i=1}^{4} N_{q,\,bin}^{reg\,i,\,sig}}{\sum_{bin=1}^{2}\sum_{i=1}^{4} N_{q,\,bin}^{reg\,i,\,sig}},\ q = c \text{ or } b. \tag{7.6}$$

The obtained numbers of signal events, which are used to calculate the differential cross sections in data, are presented in table 7.1.

Quantity	**Range**	$\boldsymbol{N_{c,\,sig}^{bin}}$	$\boldsymbol{N_{b,\,sig}^{bin}}$
$p_t(D^*\mu)$	0 - 5 GeV	263 ± 26	96 ± 22
	5 - 10 GeV	12 ± 1	53 ± 12
$\eta(D^*\mu)$	-3 - 1	172 ± 18	131 ± 27
	1 - 4	77 ± 8	40 ± 8
$M(D^*\mu)$	2 - 6 GeV	121 ± 12	110 ± 26
	6 - 12 GeV	144 ± 14	39 ± 9
$y(D^*\mu)$	-1.5 - 0	126 ± 13	72 ± 15
	0 - 1.5	122 ± 13	107 ± 22

Table 7.1: *Number of signal events in the two bins of the $D^*\mu$ combined variables, as determined with the χ^2 fit.*

Efficiencies

The reconstruction, trigger and $f_{p_t(D^*\mu)}$ efficiencies were calculated for every bin (see figure 7.5), according to the formulas presented in chapter 6. The efficiency of the f_{p_t} cut is mostly stable in the bins of the considered variables, whereas the reconstruction and trigger efficiencies show variations.

Differential Cross Section in Monte Carlo
In Monte Carlo, the differential cross sections are calculated as:

$$\boxed{\frac{d\sigma^q_{vis,\,MC}}{dx} = \frac{\Delta N^q_{MC}(D^*\mu)}{\mathcal{L}_{MC}} \cdot \frac{f^{cor}(q\bar{q} \to D^*\mu)}{\Delta x}} \qquad q = c \text{ or } b. \tag{7.7}$$

$\Delta N^q_{MC}(D^*\mu)$ is the number of generated $D^*\mu$ pairs in the visible range of the analysis, in the considered bin. $\mathcal{L}_{MC}$ is the luminosity of the Monte Carlo sample, and $f^{cor}(q\bar{q} \to D^*\mu)$ is the correction factor determined in section 4.1.1.

7.3 Migration Effects

In case of differential cross sections, possible migrations of events from one bin to another, due to detector resolution, need to be corrected for. In this analysis, the corrections for migration effects are included in the reconstruction efficiency as defined in equation 6.14. However, it is important to ensure that the migration effects are small, in order to reduce the sensitivity to the correct modelling of events. This is done by choosing an appropriate binning.

In this context, the following two quantities are defined: the **purity** $\mathcal{P}$ and **stability** $\mathcal{S}$. They are given by:

$$\mathcal{P} = \frac{N_{rec\,\&\,gen}}{N_{rec}} \text{ and } \mathcal{S} = \frac{N_{rec\,\&\,gen}}{N_{gen}}. \tag{7.8}$$

N_{rec} (N_{gen}) is the number of events found within a given bin, on reconstructed (generated) level. $N_{rec\,\&\,gen}$ is the number of events which are generated and reconstructed within the same bin.

The purity is a measure of the fraction of reconstructed events which are also at generated level in the same bin. The stability gives the fraction of generated events which remain also at reconstructed level in the same bin. Usually, the purity and stability are required to be larger than 30% [85].

The purity and stability distributions for the $D^*\mu$ variables are shown in figure 7.3. The purities are high (above 71%). The stabilities are in general over 30%, with the exception of the last bin in $\eta(D^*\mu)$, where the stability amounts to 28%.

7.4 Systematic Uncertainties

In principle, the same systematic uncertainties which were determined for the total cross sections (see section 6.8.1) are valid also here, only that they are calculated for every bin of a variable. For normalised differential cross sections ($1/\sigma \cdot d\sigma/dx$) some of the systematic errors cancel. Examples are the errors of the D^* and μ track reconstruction, of the muon identification, D^* reflections, branching ratio and of the luminosity.

The systematics due to the Δm fit were studied by variating the sigma of the Gaussian function in the Δm fit (see section 3.4.1) with its error, i.e. ± 0.1 MeV. This resulted in

systematic uncertainties up to 5%, which are however smaller than the statistical errors.

7.5 Normalised Differential Cross Sections

The differential cross sections for the $D^*\mu$ combined variables, normalised to the total cross section, $1/\sigma \cdot d\sigma_{vis}/dx$, in the visible range, for beauty and charm production, are presented in figure 7.6. The shape of the $D^*\mu$ distributions in data are compared to the leading order QCD PYTHIA and CASCADE Monte Carlo.

The $D^*\mu$ pairs coming from b quarks are produced with a larger transverse momentum than the ones from c quarks. The CASCADE generator is based on the CCFM evolution equation, in which the gluon taking part in the hard interaction may have a non-vanishing transverse momentum k_t. As a result, the $p_t(D^*\mu)$ is slightly harder than in the PYTHIA case.

According to PYTHIA and CASCADE, in charm production the $D^*\mu$ pairs travel mostly in the backward region, i.e. negative $\eta(D^*\mu)$.

The two generators predict the same shape for the invariant mass $M(D^*\mu)$ in beauty production. In the $c\bar{c}$ case, the CASCADE spectrum is harder.

Differences between PYTHIA and CASCADE are observed also for the inelasticity $y(D^*\mu)$ in charm production. PYTHIA predicts almost a flat behaviour, whereas according to PYTHIA the $D^*\mu$ pairs tend to have large inelasticities $y(D^*\mu) > 0$.

In general, PYTHIA offers the better description. However, in view of the large errors, it is not possible to state whether the data prefer the PYTHIA or the CASCADE predictions.

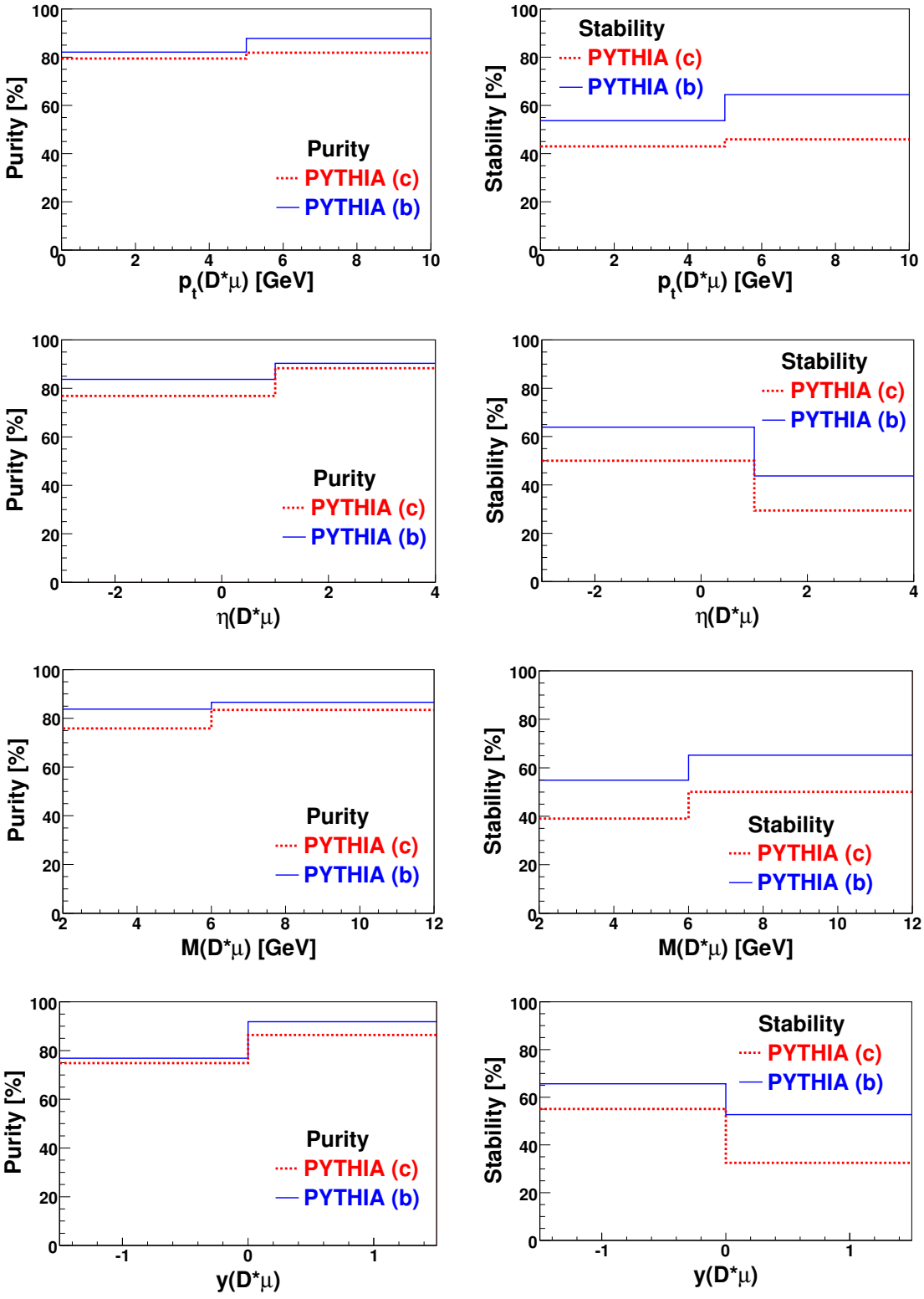

Figure 7.3: *Purity (left) and stability (right) distributions for $D^*\mu$ combined variables: transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, invariant mass $M(D^*\mu)$ and inelasticity $y(D^*\mu)$, as predicted by PYTHIA charm (c) and beauty (b) Monte Carlo.*

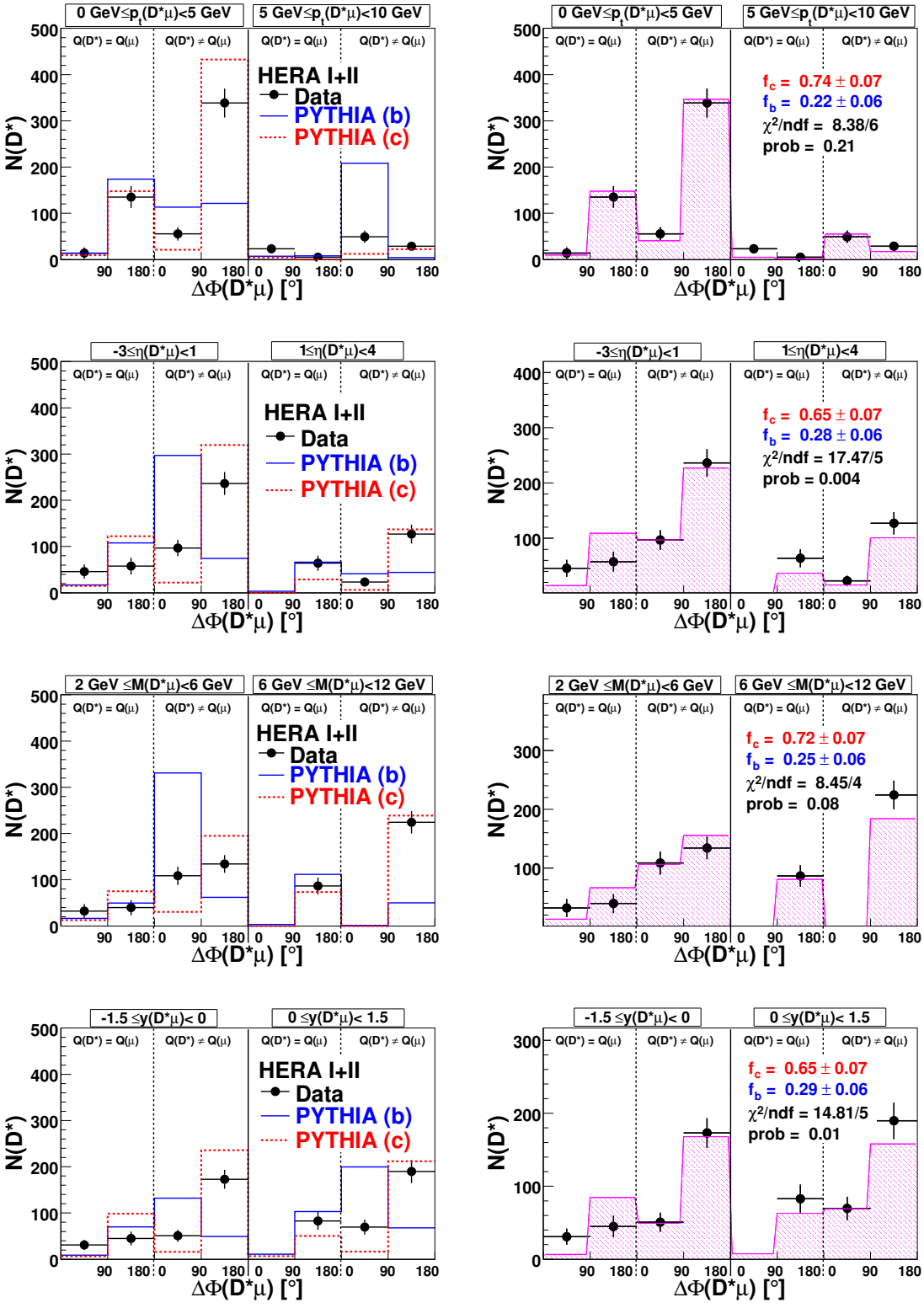

Figure 7.4: *Left: distributions of events in the four correlations regions for the two bins of combined $D^*\mu$ variables. The distributions of the transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, invariant mass $M(D^*\mu)$ and inelasticity $y(D^*\mu)$ are shown. The black points indicate the data, and the lines the PYTHIA beauty (b) and charm (c) Monte Carlo. The Monte Carlo is normalised to data. Left: results of the χ^2 fit for the same variables.*

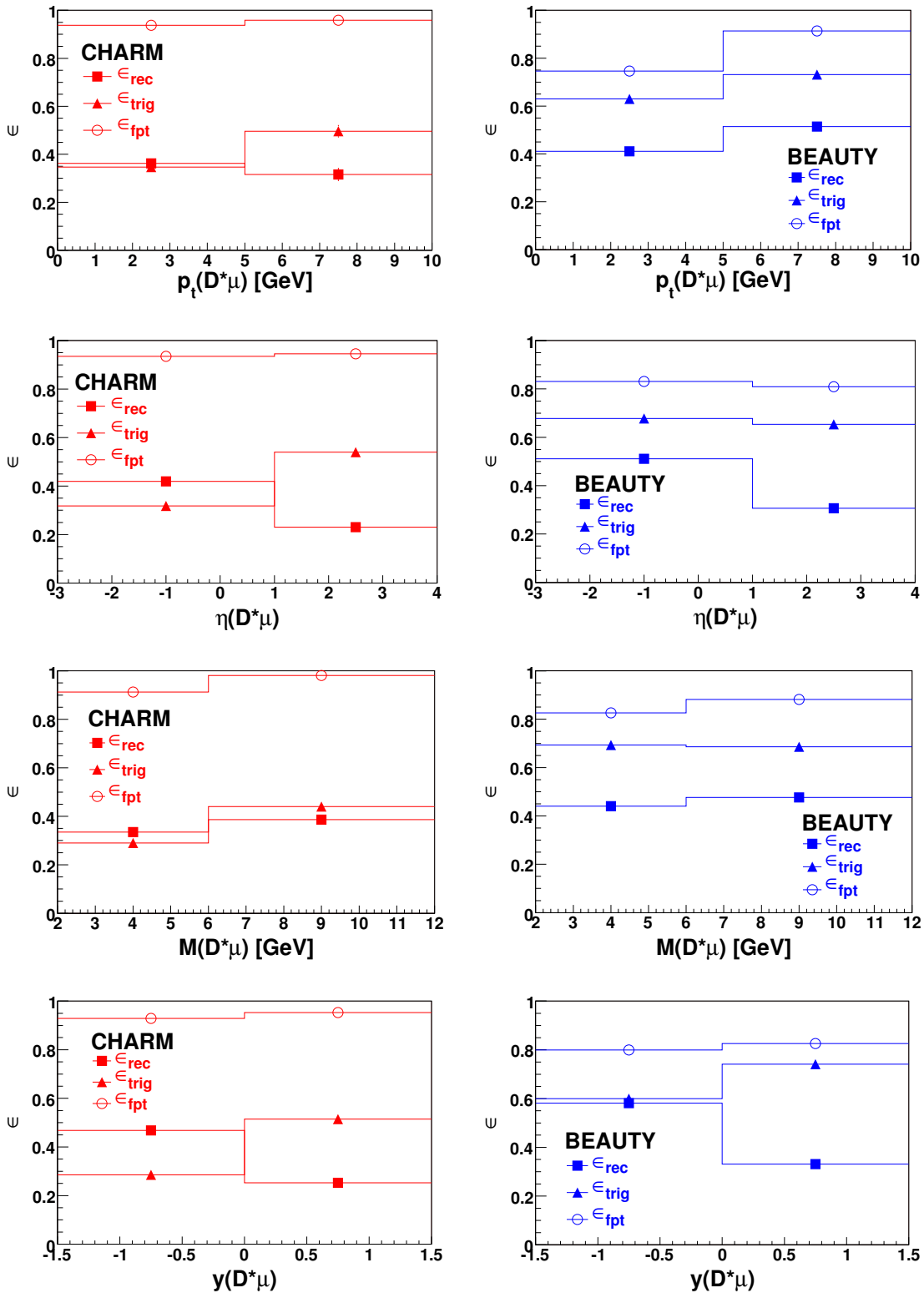

Figure 7.5: *Efficiencies as a function of transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, invariant mass $M(D^*\mu)$ and inelasticity $y(D^*\mu)$ for charm (left) and beauty (right) production. The reconstruction ϵ_{rec}, trigger ϵ_{trig} and $f_{p_t(D^*\mu)}$ cut ϵ_{fpt} efficiencies are shown.*

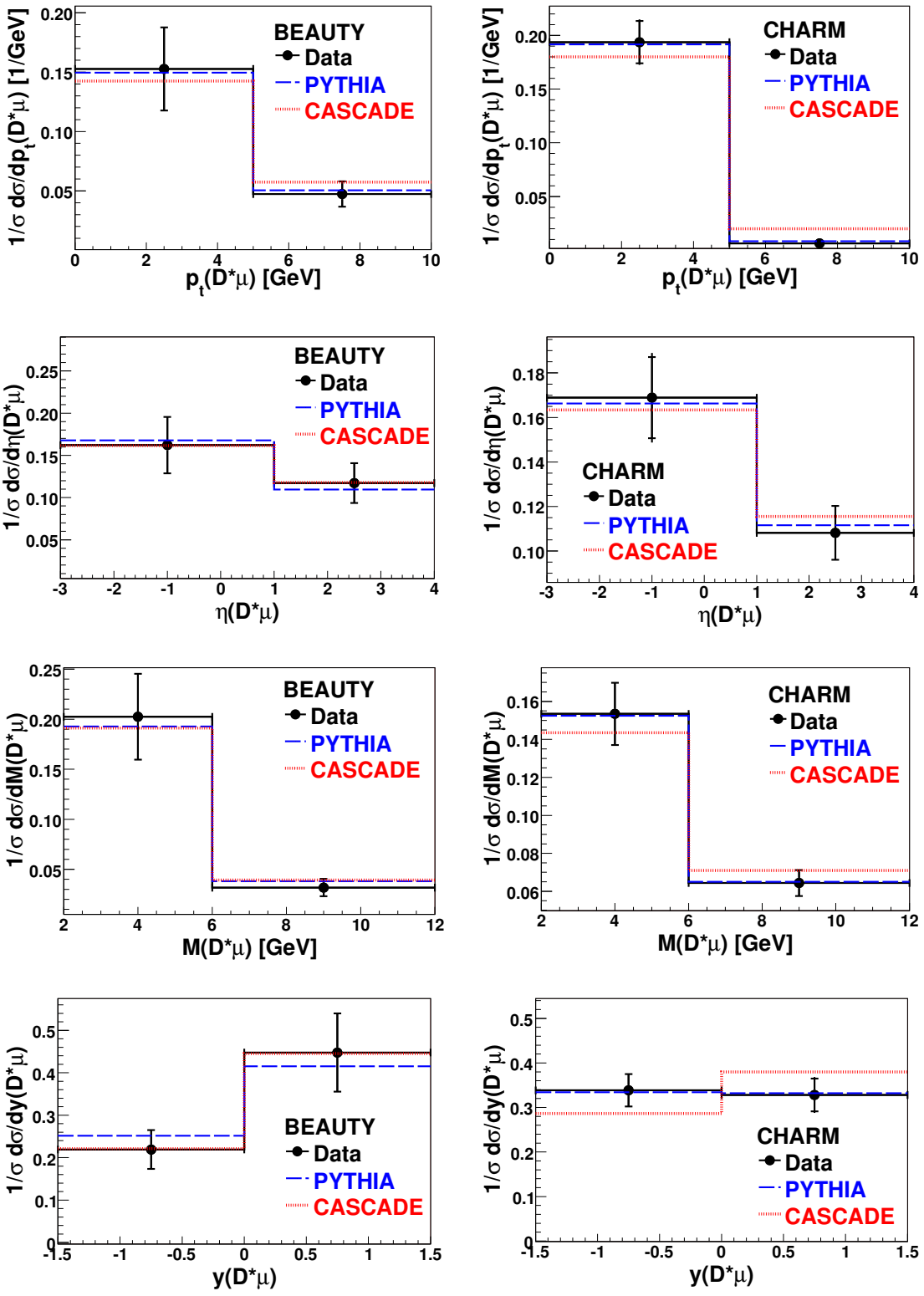

Figure 7.6: *Normalised differential cross sections for the transverse momentum $p_t(D^*\mu)$, the pseudorapidity $\eta(D^*\mu)$, the invariant mass $M(D^*\mu)$ and the inelasticity $y(D^*\mu)$: left - beauty, right - charm production. The data (black points) are compared to PYTHIA and CASCADE Monte Carlo generators (dotted lines). The error bars indicate the statistical uncertainties. The total errors are given by the statistical and systematic errors added in quadrature.*

Chapter 8

Mean $D^*\mu$ Transverse Momentum Studies

In this chapter studies of the mean $D^*\mu$ transverse momentum, $\langle p_t(D^*\mu)\rangle$, in data and Monte Carlo, will be presented. First, a motivation will be given, then the results will be introduced.

To study the QCD dynamics of heavy quark production via boson-gluon fusion, the PYTHIA and CASCADE generators were used. PYTHIA is based on the DGLAP evolution equation (see section 1.3.5.1), which assumes an ordering in the transverse momenta of the emitted gluons. CASCADE implements the CCFM equation (section 1.3.5.3), in which an angular ordering is considered. Initial transverse momentum of the incoming partons lead to a non-vanishing transverse momentum of the heavy quark pair (see section 1.7.2.2). The latter is approximated by the $D^*\mu$ variable $p_t(D^*\mu)$. Although there are smearing effects due to fragmentation of charm (or beauty) into the D^* meson, this variable is expected to be sensitive to possible intrinsic momenta of the incoming partons. This proved to be the case for the mean $D^*\mu$ transverse momenta $\langle p_t(D^*\mu)\rangle$[1] as a function of pseudorapidity $\eta(D^*\mu)$.

8.1 Measurement of $\langle p_t(D^*\mu)\rangle$

To be able to compare data to Monte Carlo, signal distributions were obtained analogous to the subtraction method presented in section 3.4.2. As an example, the initial $p_t(D^*\mu)$ distributions in data, for the $RiCh$ and $WrCh$ cases, are presented in figure 8.1 (right), for correlation region 3, and in figures 8.2 and 8.3 (right) for region 4. The value of $\langle p_t(D^*\mu)\rangle$ is taken to be the mean of the $p_t(D^*\mu)$ signal distribution. The procedure may fail due to statistical fluctuations which lead to negative entries. In these cases, the corresponding mean transverse momenta are further neglected. This happens only at extreme values of pseudorapidity $\eta(D^*\mu)$, whereas the differences between PYTHIA and CASCADE Monte Carlo are observed in central regions.

[1] By taking the mean of the momenta, normalisation effects can be neglected.

As explained in section 3.7, the $D^*\mu$ pair is required to carry more than 20% of the transverse momentum of all hadronic final state particles, $f_{p_t}(D^*\mu) > 0.20$. The influence of this variable on the mean transverse momentum of the $D^*\mu$ pair has been studied (see figure 8.4). Up to momentum fractions of 20%, almost no change in the mean p_t value is observed. A statistical deviation is seen at a fraction of 15%. At higher momentum fractions, the mean p_t value is strongly affected.

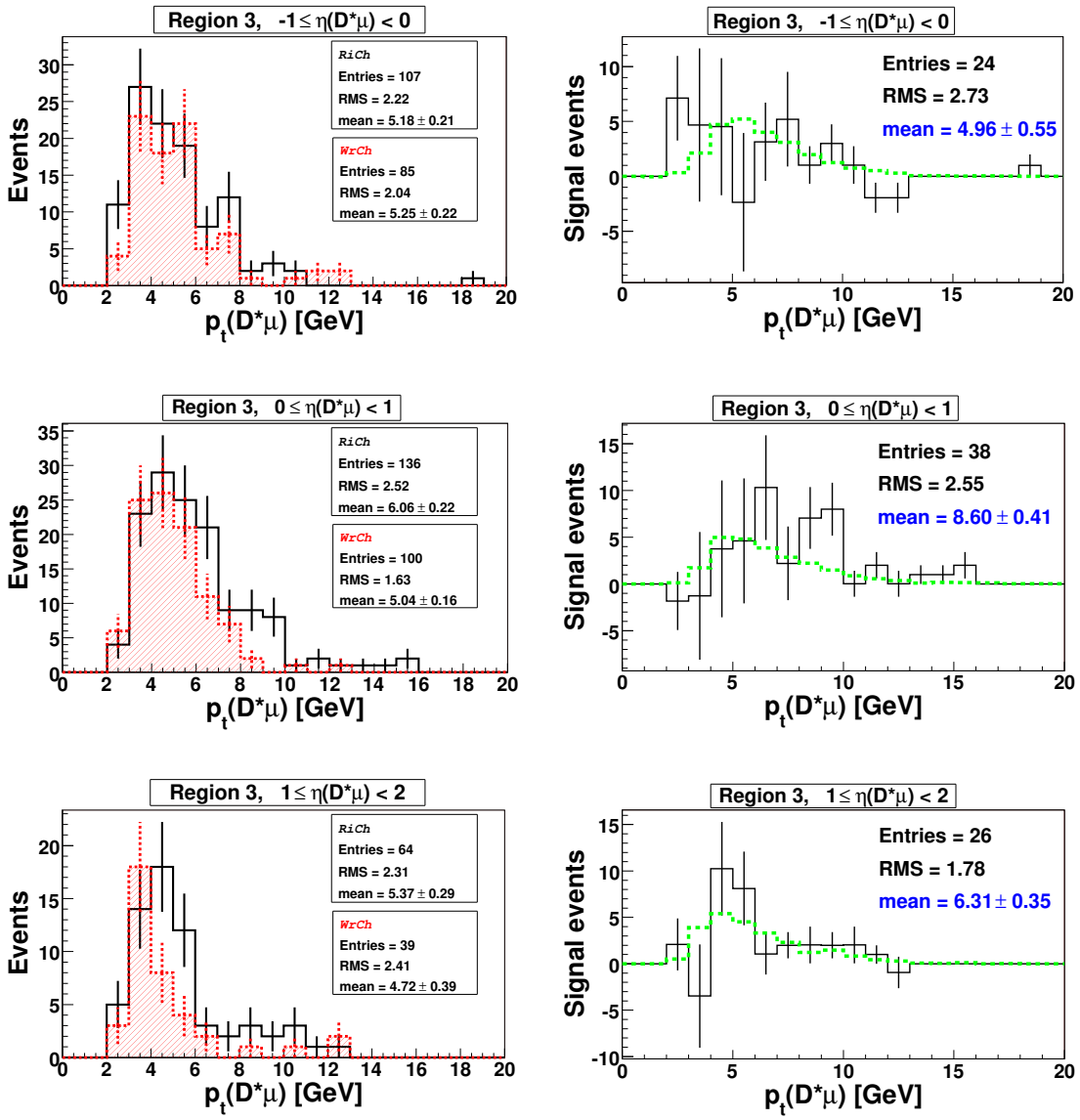

Figure 8.1: *Distributions of transverse momentum $p_t(D^*\mu)$ in* ***data****, for* ***correlation region 3****. Left: right charge (RiCh) and wrong charge ($WrCh$) $p_t(D^*\mu)$ distributions. Right: signal p_t distribution with resulting statistical errors. The $WrCh$ histograms were scaled with a factor of 1.03 in order to describe the RiCh background. The dotted line indicates the behaviour of CASCADE beauty Monte Carlo.*

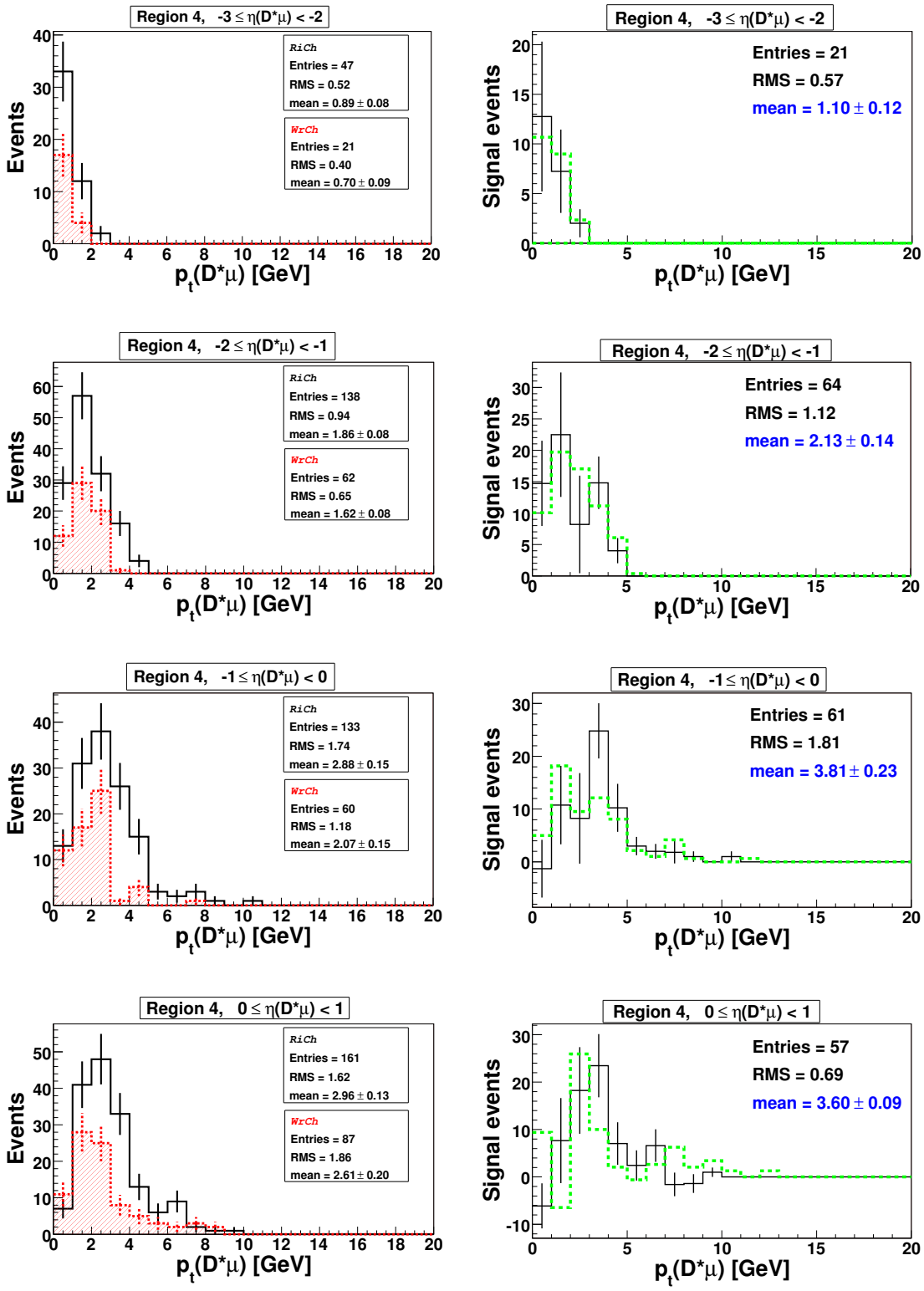

Figure 8.2: *Distributions of transverse $p_t(D^*\mu)$ in* **data**, *for* ***correlation region 4***. *Left: right charge (RiCh) and wrong charge ($WrCh$) $p_t(D^*\mu)$ distributions. Right: signal p_t distribution with resulting statistical errors. The $WrCh$ histograms were scaled with a factor of 0.84 in order to describe the RiCh background. The dotted line indicates the behaviour of PYTHIA charm Monte Carlo.*

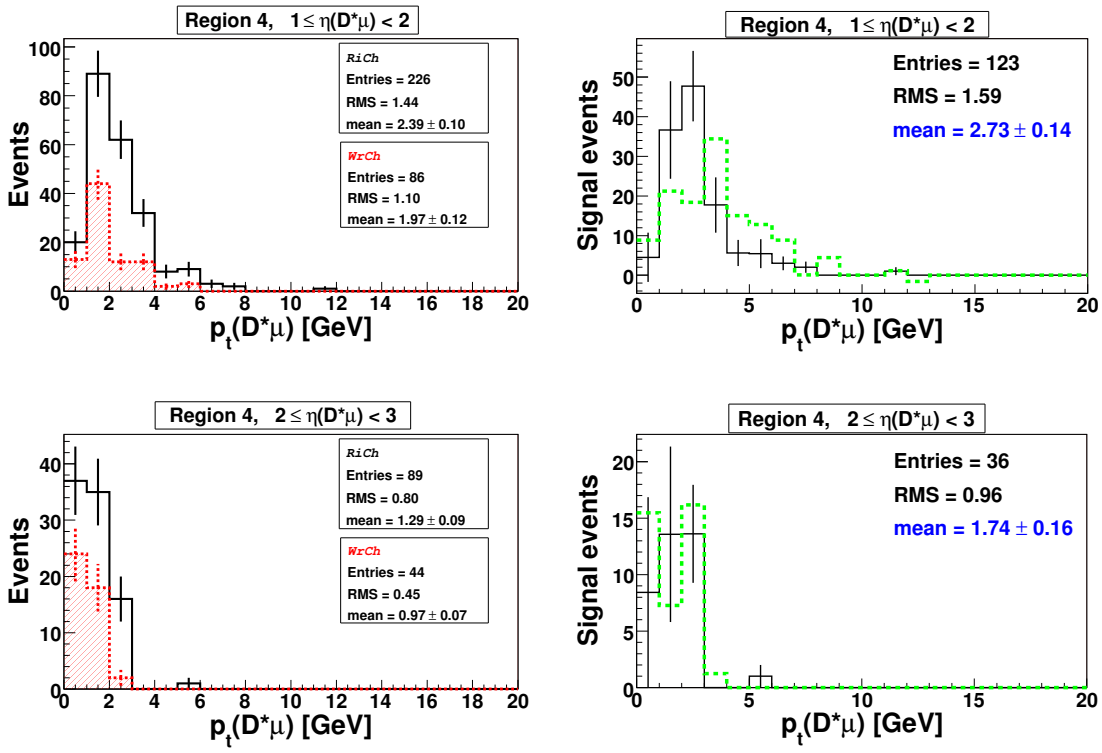

Figure 8.3: *Distributions of transverse $p_t(D^*\mu)$ in* ***data****, for* ***correlation region 4*** *(continued). Left: right charge (RiCh) and wrong charge (WrCh) $p_t(D^*\mu)$ distributions. Right: signal p_t distribution with resulting statistical errors. The WrCh histograms were scaled with a factor of 0.84 in order to describe the RiCh background. The dotted line indicates the behaviour of PYTHIA charm Monte Carlo.*

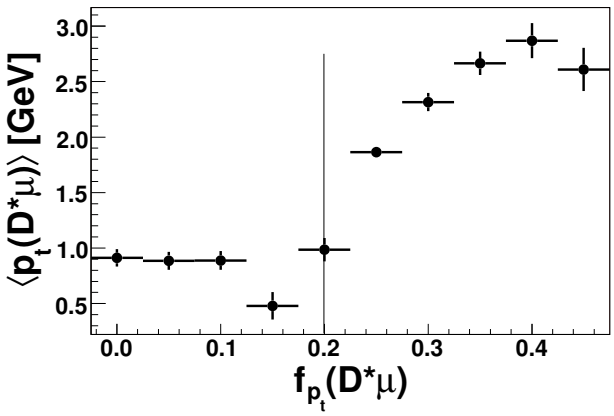

Figure 8.4: *Mean $D^*\mu$ transverse momentum $\langle p_t(D^*\mu)\rangle$ as a function of momentum fraction $f_{p_t}(D^*\mu)$. The line indicates the cuts applied in this analysis.*

8.2 Results

The obtained mean transverse momenta distributions for charm and beauty production in the four correlation regions are shown in figure 8.5. Important are the charm and beauty dominated regions, i.e. region 4 and 3. For completeness, distributions in the other regions are shown as well.

According to PYTHIA (see table 5.2), region 4 contains 91% events[2] coming from c quarks, so charm Monte Carlo is expected to describe the data. A mean $p_t(D^*\mu)$ difference of around 0.8 GeV is observed between CASCADE, which is favoured by data, and PYTHIA prediction.

Correlation region 3 contains 80% events from b quarks. CASCADE and PYTHIA predict similar $D^*\mu$ transverse momentum values. In this region, the D^* meson and the muon come from the same b quark, and the sensitivity to the possible intrinsic k_t momenta of the initial partons is lost. Since the data points fluctuate strongly, no statement, whether PYTHIA or CASCADE describes the data, is possible.

8.3 Systematic Uncertainties

The mean transverse momenta distributions, $\langle p_t(D^*\mu)\rangle$, in bins of $\eta(D^*\mu)$, are obtained by subtracting the $WrCh$ histograms scaled by a factor U_{scal}, which is given by a simultaneous fit to the $RiCh$ and $WrCh$ distributions, away from the signal region (see section 8.1). However, when subtracting, the scaling factor is treated as a constant, although it has an associated error given by the fit. By modifying that factor with its error, systematic uncertainties up to 6% were obtained. Note that they are negligible compared to the statistical errors.

[2]Similar contributions are predicted by CASCADE.

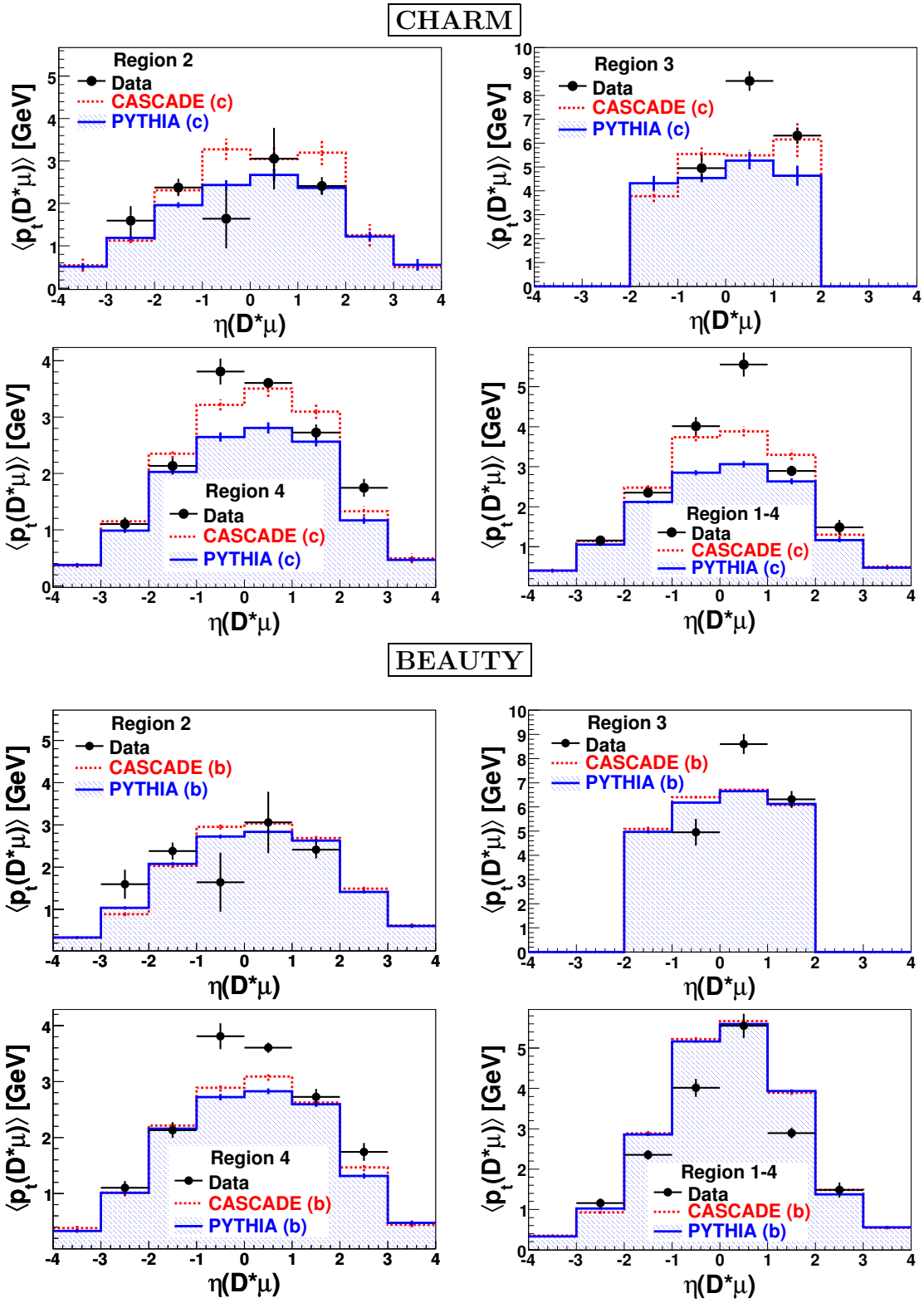

Figure 8.5: *Distributions of the mean transverse momentum $\langle p_t(D^*\mu)\rangle$ as a function of pseudorapidity $\eta(D^*\mu)$, for charm and beauty production, in the four correlation regions (defined in table 1.6). The data (points) are compared to PYTHIA and CASCADE generators (hashed regions). For the cases in which the data points are missing, the mean transverse momenta could not be measured reliably due to the background subtraction procedure.*

Chapter 9

Conclusions and Outlook

In this thesis, the production of charm and beauty events at the ep collider HERA is studied using events with a D^* meson and a muon. Data taken by the H1 experiment in the years 1999-2000 (HERA I) and 2004-2006 (HERA II) are analysed. The charm (or beauty) quark is tagged by reconstructing D^* mesons with transverse momentum $p_t(D^*) > 1.5$ GeV and pseudorapidity $|\eta(D^*)| < 1.5$ in the decay channel $D^{*\pm} \to K^{\mp}\pi^{\pm}\pi_s^{\pm}$. Muons with momentum $p(\mu) > 2.0$ GeV and pseudorapidity $|\eta(\mu)| < 1.735$ are selected as well. In case of charm production, the muon comes mostly from the second c quark, such that the heavy quark pair is almost completely reconstructed. In beauty production, the muon comes predominantly from the same b quark as the D^* meson.

The $D^*\mu$ charge and azimuthal angular correlations are exploited to extract the fractions of charm and beauty events[1] in data:

$$f_c = (71 \pm 7)\%, \tag{9.1}$$
$$f_b = (26 \pm 6)\%. \tag{9.2}$$

The obtained fractions are corrected for muon background using the leading order QCD Monte Carlo program PYTHIA. With these fractions, the charm and beauty cross sections are measured in photoproduction, i.e. at photon virtualities $Q^2 < 1$ GeV^2, and for inelasticities $0.05 < y < 0.75$, corresponding to photon-proton centre-of mass energies 71 GeV $< W_{\gamma p} <$ 275 GeV:

$$\sigma_{vis}(ep \to ec\bar{c}X \to D^*\mu X) = 314 \pm 33\ (\text{stat}) \pm 101\ (\text{syst})\ \text{pb} \tag{9.3}$$
$$\sigma_{vis}(ep \to eb\bar{b}X \to D^*\mu X) = 90 \pm 20\ (\text{stat}) \pm 26\ (\text{syst})\ \text{pb}. \tag{9.4}$$

The measurements are compared to the leading order QCD calculations as implemented in the Monte Carlo simulation program PYTHIA. This program is based on the DGLAP evolution equation, which considers k_t ordered gluon radiation in the parton evolution. Both charm and beauty cross sections are found to be in agreement, within the errors, with PYTHIA predictions: a ratio data over theory of 1.2 ± 0.1 (stat) ± 0.4 (syst) is obtained for charm, and of 1.3 ± 0.3 (stat) ± 0.4 (syst) for beauty.

[1] The sum $f_c + f_b$ is less than one, due to the independent fluctuations of the numbers of charm and beauty events in the χ^2 fit (for details, see chapter 5). The difference of 3% is accounting for normalisation effects.

The kinematic properties of the heavy quark pair are studied using combined $D^*\mu$ variables: transverse momentum $p_t(D^*\mu)$, pseudorapidity $\eta(D^*\mu)$, invariant mass $M(D^*\mu)$ and rapidity $y(D^*\mu)$. The differential cross sections are measured as a function of these variables and compared to PYTHIA and CASCADE simulation programs. CASCADE is a leading order QCD Monte Carlo based on the CCFM evolution equation, which considers gluon radiation ordered not in transverse momentum k_t, but in emission angle. Taking into account the statistical and systematic uncertainties, the differential distributions are in general consistent with the leading order QCD predictions.

The distribution of the mean $D^*\mu$ transverse momentum $\langle p_t(D^*\mu)\rangle$ as a function of pseudorapidity $\eta(D^*\mu)$ indicates a sensitivity to the two different QCD models implemented in PYTHIA and CASCADE. Differences of around 0.8 GeV in transverse momenta are observed between the two predictions, if the D^* meson and the muon come from different c quarks. In this case, data tend to favour CASCADE. The sensitivity is washed out in beauty production, since there D^* and μ come mostly from the same quark.

In summary, the advantages of the double tagging method using $D^*\mu$ events are:

- The full reconstruction of the D^* decay products allows an accurate determination of the D^* kinematics, and hence of the kinematics of the mother charm quark, since the D^* is well correlated with the initial quark;
- Via the combined $D^*\mu$ variables, one has access to the kinematics of both heavy quarks, and a more detail study of the production process, compared to single tagging, is possible;
- The number of signal $D^*\mu$ events can be easily determined by fitting the Δm distribution;
- Since the muons are selected with momentum $p(\mu) > 2$ GeV, the beauty quarks in this case can be probed at lower momenta than in events with a muon and jets (see section 1.7.1.3).

The disadvantages are:

- The total tagging efficiency is low: around 610 $D^*\mu$ events are found in a sample corresponding to 320 pb^{-1};
- The kinematic and geometric acceptance of the three tracks of the D^* decay particles restrict the study to the central region of the detector;
- Since the correction for the fake muon background is based on the Monte Carlo simulation, a good understanding of this background is necessary, which is not always easy to get;
- The muon tagging is not very suited for charm production, where the muons have in general low transverse momenta (less than 1 GeV).

Outlook

The present measurements can be improved by reducing the statistical and especially the systematic uncertainties. For this purpose, detailed studies of the hadrons mis-identification probabilities, in particular for pions, are necessary. Further, the reconstruction efficiency of the D^* daughters and of the muon track need to be investigated extensively in order to reduce the corresponding systematic uncertainty. Another important point is the Monte Carlo description of the muon kinematic variables, specifically of the muon momentum. Towards a perfect description, correction factors need to be determined and applied.

To get more insight into the heavy quark production mechanisms, a comparison with NLO QCD predictions can be done.

For double tagging the heavy quarks, other particles, for example D^*e, may be considered.

Appendix A

Example Events in Correlation Regions

Event displays with $D^*\mu$ pairs in correlation[1] regions 2 and 3 are shown in figures A.1 and A.2. The kinematic quantities of the D^* decay particles and of the muon, for the two cases, are given in table A.1.

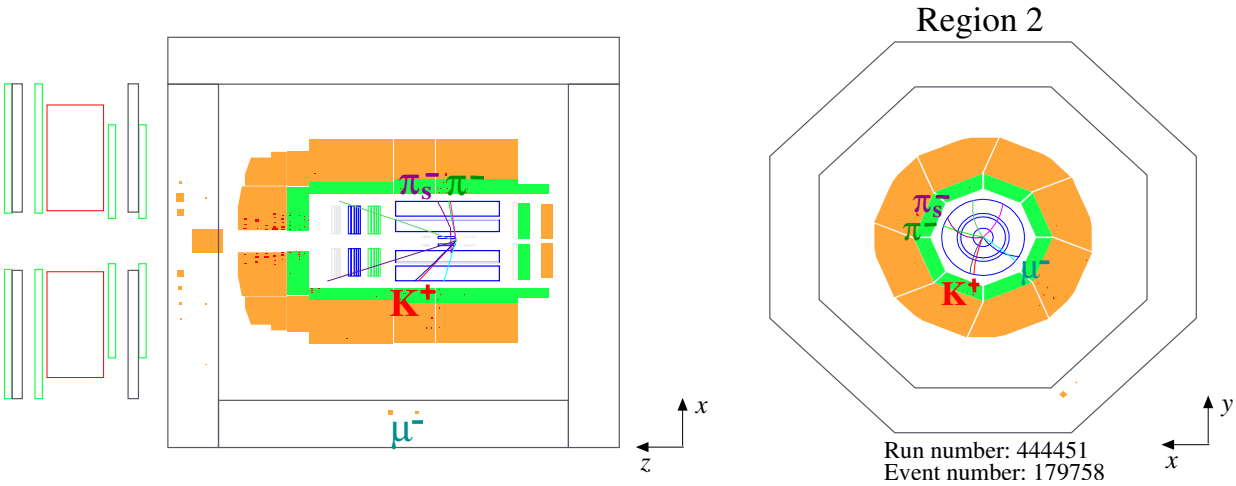

Figure A.1: *Side- and radial view of the H1 detector in case of an event with a D^* meson and a muon in kinematic* ***region 2****, i.e the azimuthal angular difference between the two particles is larger than 90° and they have the same charge.*

[1]For a definition of the correlation regions see table 1.6.

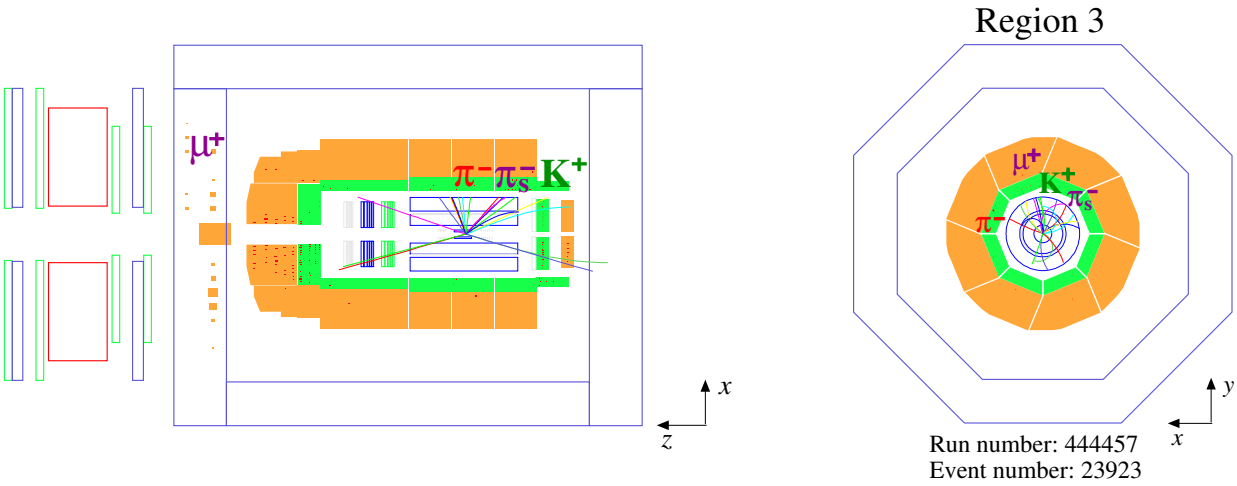

Figure A.2: *Side- and radial view of the H1 detector in case of an event with a D^* meson and a muon in kinematic* ***region 3****, i.e the azimuthal angular difference between the two particles is less than 90° and they have opposite charge.*

Figure	Particle	p_t [GeV]	θ [°]	ϕ [°]	$\Delta\phi(D^*\mu)$ [°]
A.1	K^+	0.91	42.12	-68.18	130.51
	π^-	1.61	79.27	11.81	
	π_s^-	0.21	66.97	-13.94	
	μ^-	2.04	73.19	-145.41	
A.2	K^+	1.0	140.29	94.44	30.35
	π^-	1.68	128.49	18.98	
	π_s^-	0.17	128.68	67.29	
	μ^+	3.47	20.74	77.52	

Table A.1: *Kinematic variables of the particles in the two events, in the correlation regions 2 and 3, presented in figure A.1 and A.2. Note that the charge of the D^* meson is given by π_s.*

Appendix B

$B^0 - \overline{B}^0$ Mixing

When presenting the possible configurations with a D^* meson and a muon for beauty and charm production (see figure 1.24), contributions from the $B_d^0 - \overline{B}_d^0$ or $B_s^0 - \overline{B}_s^0$ mixing were neglected.

The time probabilities that these B mesons, in each system, remain unchanged (+) or oscillate into each other (−) are proportional to [15]:

$$|g_{\pm}(t)|^2 = \frac{\exp\left(-\Gamma_q \cdot t\right)}{2} c \cdot \left[\cosh\left(\frac{\Delta\Gamma_q}{2} \cdot t\right) \pm \cos\left(\Delta m_q \cdot t\right)\right], \tag{B.1}$$

where $\Gamma_q = (\Gamma_H + \Gamma_L)/2$ is the total decay width, and $\Delta m_q = m_H - m_L > 0$ is the mass difference. 'H' and 'L' indicate the heavy and light mass eigenstates, q stands for d or s quark. Without CP violation, the time-integrated mixing probability is given by:

$$\chi_q = \frac{\int |g_-(t)|^2 \cdot dt}{\int |g_-(t)|^2 \cdot dt + \int |g_+(t)|^2 \cdot dt} = \frac{x_q^2 + y_q^2}{2 \cdot (x_q^2 + 1)}, \tag{B.2}$$

with $x_q = \Delta m_q / \Gamma_q$ and $y_q = \Delta\Gamma_q / (2\Gamma_q)$.

In case of $D^*\mu$ pairs which come from different quarks (see figure 1.24, a) and b), the $B^0 - \overline{B}^0$ mixing may induce a migration of the two beauty cases, by flipping the charge of the particles from one of the two quark branches (see figure B.1). The beauty quarks have the following branching ratios [15]:

$$\mathcal{BR}(b \to B_d^0) = (39.8 \pm 1.2)\%, \tag{B.3}$$

$$\mathcal{BR}(b \to B_s^0) = (10.3 \pm 1.4)\%. \tag{B.4}$$

The time integrated mixing probabilities are [15]:

$$\chi_d = 0.188 \pm 0.003 \ \text{ and } \ \chi_s > 0.49878 \ \ (\text{CL} = 95\%). \tag{B.5}$$

Therefore, although the $B_s^0 - \overline{B}_s^0$ mixing is maximal, the probability to obtain a B_s^0 meson from a b quark is low, whereas in the $B_d^0 - \overline{B}_d^0$ system is the opposite situation: the branching ratio is larger, but the mixing probability is smaller, such that the resulting oscillation fractions are comparable ($\sim 5\%$ to $\sim 8\%$). Therefore, migrations between like

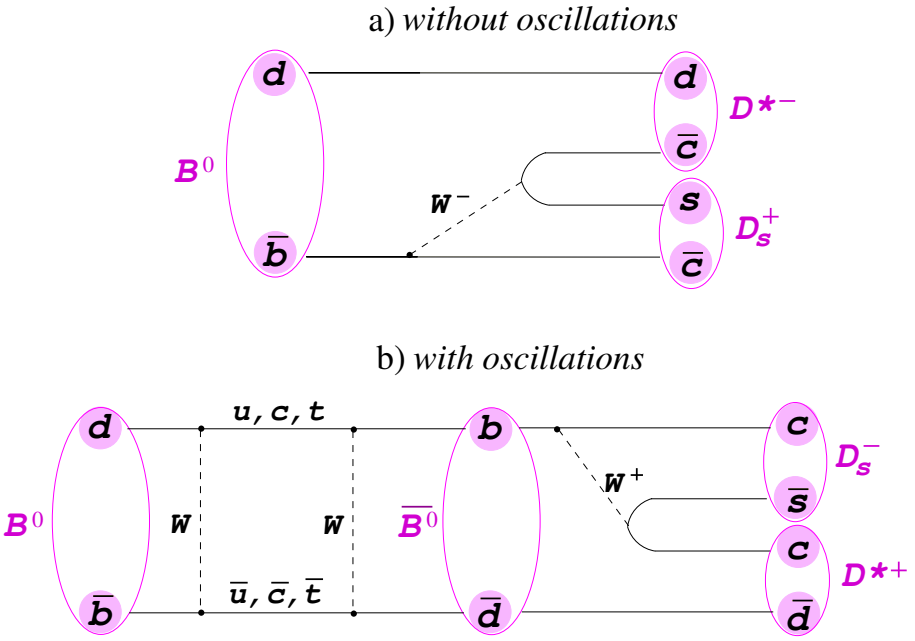

Figure B.1: *Illustration of $B^0 - \overline{B}^0$ mixing: a) the B^0 meson decays directly into a D^{*-} and a D_s^+ meson, without oscillating into a $\overline{B}^0$; b) B^0 first 'oscillates' into its antiparticle via quark transitions before decaying.*

and unlike sign $D^*\mu$ events coming from different quarks are expected in approximatively 13% of the cases.

According to PYTHIA predictions (see table 5.4), 53% of beauty events are in correlation region 3, and 18% in region 4. Consider region 3. With an oscillation fraction of 5%, 95% of the b events from region 3 stay in this region, whereas 5% of the b events in region 4 migrate in region 3. This results into an overall fraction of 51% of b events in region 3, which is close to the initial fraction of 53%.

The $B^0 - \overline{B}^0$ is included in the used Monte Carlo programs.

Appendix C

$f_{p_t}(D^*\mu)$ Cuts

To reduce the combinatorial background, the transverse momentum fraction $f_{p_t}(D^*\mu)$ was used (see section 3.7). The effect of the $f_{p_t}(D^*\mu)$ cuts on the Δm distribution is shown in figures C.1 and C.2.

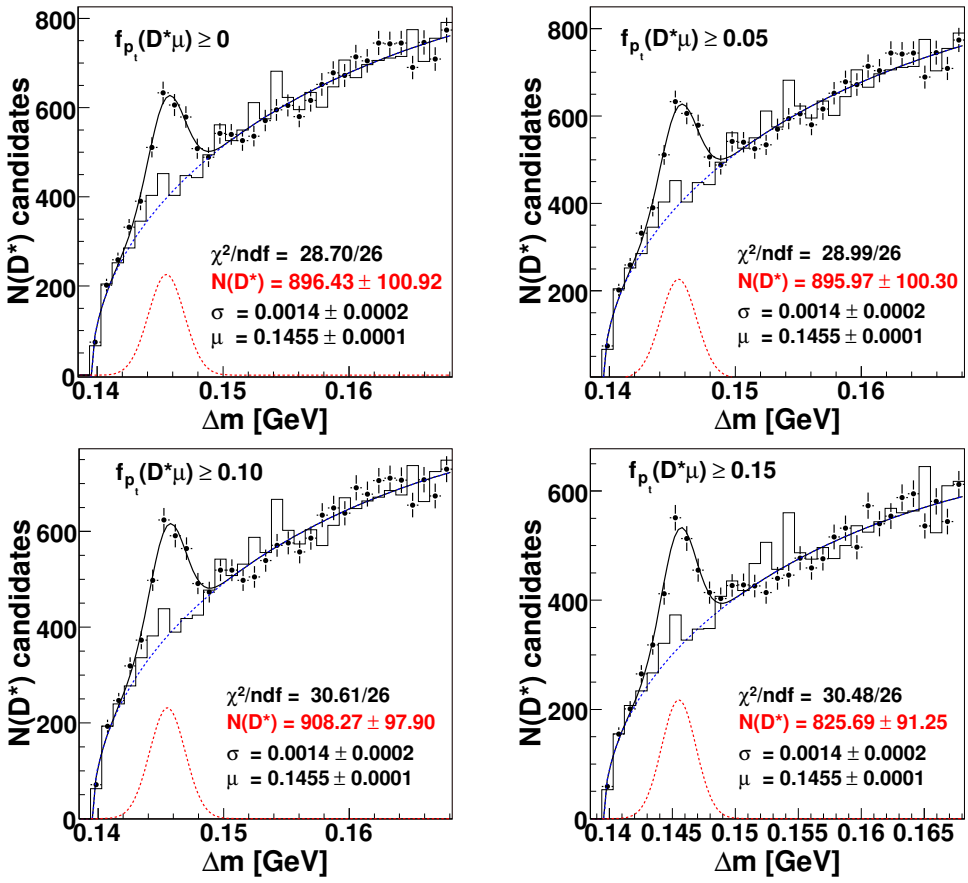

Figure C.1: *Distributions of the invariant mass difference* $\Delta m = m(K\pi\pi_s) - m(K\pi)$ *with different* $f_{p_t}(D^*\mu)$ *cuts for HERA I and II run periods.*

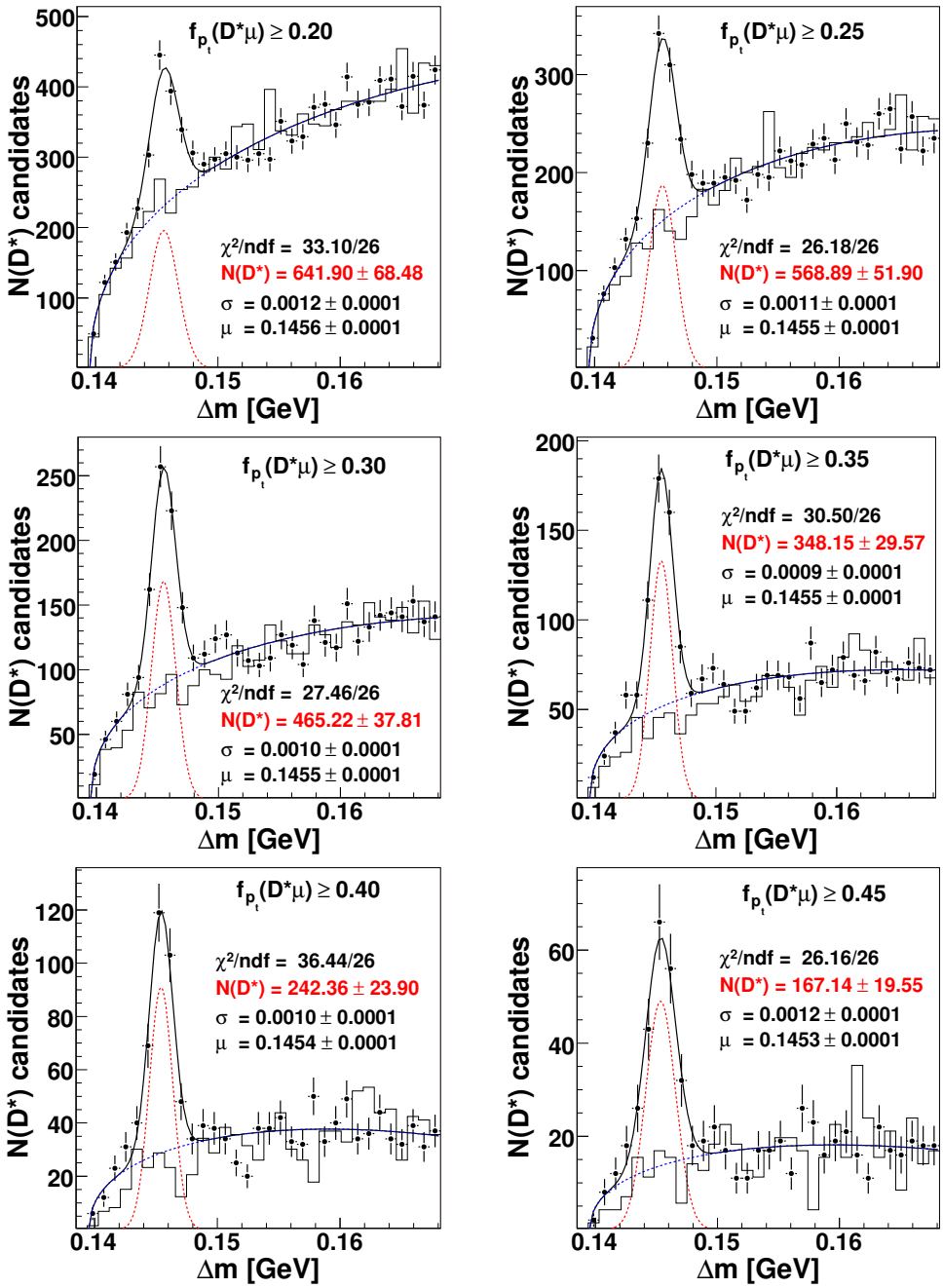

Figure C.2: *Distributions of the invariant mass difference $\Delta m = m(K\pi\pi_s) - m(K\pi)$ with different $f_{p_t}(D^*\mu)$ cuts for HERA I and II run periods (continued).*

Appendix D

Trigger Elements Definition

Trigger element (TE)	Definition
TE of the Central Muon Detector	
Mu_Bar	muon candidate in the barrel
Mu_ECQ	muon candidate in the endcaps, but not in the forward inner: Mu_FOEC ‖ Mu_BOEC ‖ Mu_BIEC
Mu_FOEC	muon candidate in the Forward Outer Cap
Mu_FIEC	muon candidate in the Forward Inner Cap
Mu_BOEC	muon candidate in the Backward Outer Cap
Mu_BIEC	muon candidate in the Backward Inner Cap
Mu_Any	muon candidate in the endcaps or the barrel: Mu_FIEC ‖ Mu_FOEC ‖ Mu_BOEC ‖ Mu_BIEC ‖ Mu_Bar
Mu_2_BIoOEC	at least 2 muon candidates in BIEC or in BOEC
Mu_BIO	shortcut for Mu_BOEC ‖ Mu_2_BIoOEC ‖ Mu_FOEC

Table D.1: *Definition of the central muon detector trigger elements used for building the subtriggers selected in this analysis (see section 3.8).*

Trigger element (TE)	Definition
TE of CJC1 and CJC2	
DCRPh_CNH	shortcut for DCRPh_Tc && DCRPh_TNeg && DCRPh_THig
DCRPh_Tc	at least three fired track masks with $p_t \geq 450$ MeV
DCRPh_TNeg	at least one fired negative track mask with $p_t \geq 450$ MeV
DCRPh_THig	at least one fired track mask with $p_t \geq 800$ MeV
DCRPh_Ta	at least one fired track mask with $p_t \geq 450$ MeV
TE for vertex significance	
zVtx_sig	significant maximum in z-vertex histogram
TE of SpaCal	
SPCLe_IET > 1	measured electron in the SpaCal, $E \geq 2$ GeV
SPCLe_Cen_2	similar to SPCLe_IET > 1, but for the central region in SpaCal
SPCLe_IET > 2	measured electron in the SpaCal, $E \geq 6$ GeV
SPCLe_Cen_3	similar to SPCLe_IET > 2, but for the central region in SpaCal
TE of CIP2000 system	
CIP_mul < 11	CIP multiplicity
TE of FTT	
FTT_mul_Td > 0	at least one track with $p_t \geq 900$ MeV
FTT_mul_Tc > 2	at least three tracks with $p_t \geq 400$ MeV

Table D.2: *Definition of the CJC1, CJC2, vertex significance, SpaCal, CIP2000 system and FTT trigger elements used for building the subtriggers selected in this analysis (see section 3.8).*

Appendix E

Correction for Level 1 Prescaling of Subtriggers

In general case, the cross section of a process is defined as:

$$\sigma = \frac{N}{\mathcal{L} \cdot \epsilon}, \tag{E.1}$$

with N the number of selected events, $\mathcal{L}$ the integrated luminosity of the sample and ϵ the efficiency. However, due to level 1 prescaling of the subtriggers, the obtained number of events may differ from the initial number of events obtained from the considered physical process: for example, in case of a prescaling of 2, out of 10 events only 5 would be accepted. Therefore, to obtain a right cross section, corrections need to be applied.

Depending on the subtriggers, there are two possibilities[1]:

- **One subtrigger**: If only one subtrigger ST is used, the integrated luminosity is corrected by the corresponding lowest prescale factor d_i^{ST} in run i:

$$\mathcal{L}_{corr} = \sum_{i}^{runs} \frac{\mathcal{L}_i}{d_i^{ST}}. \tag{E.2}$$

- **Several subtriggers**: If several subtriggers are used, from which one or more may be prescaled, for every trigger combination prescale weights are calculated, as described in section 4.1.2, and applied to the Monte Carlo events. By comparing the events before and after prescaling, an average prescale factor is determined (see section 6.2). Finally, the luminosity is corrected by this factor, such that the correct cross section is obtained.

In this analysis, three subtriggers are used for the HERA I run period, and two subtriggers for the HERA II. The complexity added by considering several subtriggers is justified by the significant increase in statistics. For example, in figure E.1 the Δm fit for the events triggered only by S19 (HERA I) can be seen. The obtained number of D^* mesons is to

[1] A third possibility would be to accept only events with no prescale.

be compared with the case in which all three HERA I subtriggers are used, for which $N(D^*) = 124 \pm 28$. Therefore, approximatively 44% of events are gained by using the other subtriggers.

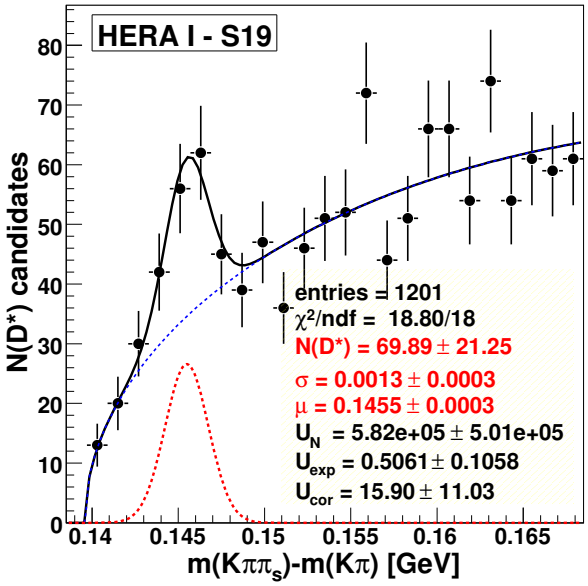

Figure E.1: *Distribution of the invariant mass difference $\Delta m = m(K\pi\pi_s) - m(K\pi)$ for the events triggered only with subtrigger 19 in the 1999-2000 (HERA I) run period.*

Appendix F

Efficiency of a Cut: Definition and Error Calculation

In the course of an analysis, it is often necessary to apply cuts on a variable, typically to reduce the background and to improve the signal to background ratio. Events are lost, and the uncertainty associated with these events has to be quantified. Let N_1 be the number of events which pass the cut, N_2 the number of events which fail the cut, and σ_{N_1}, σ_{N_2} their corresponding errors. The **efficiency of the cut** is defined as:

$$\epsilon_{cut} = \frac{N_1}{N_1 + N_2} \equiv \frac{N_1}{N}, \tag{F.1}$$

with N the total number of events. For calculating the error on this efficiency, the approach presented in [86] is followed.

Usually, the error of an efficiency is taken to be the binomial one:

$$\sigma_{\epsilon_{cut}} = \sqrt{\frac{\epsilon_{cut} \cdot (1 - \epsilon_{cut})}{N}}. \tag{F.2}$$

However, if N_1 and N are the result of two separate fits consisting in separate signal plus background functions, a difficulty appears, since the background fluctuations introduce correlations between the signal samples before and after the cut.

One possible solution is to produce a third histogram of events which fail the cut, and fit to obtain N_2. Then, one can use the fact that N_1 and N_2 are uncorrelated, since no event is present both in the histogram for events which pass the cut and for the events which fail the cut. The **statistical error on the efficiency** is therefore obtained by error propagation:

$$\sigma^2_{\epsilon_{cut}} = \left(\frac{\partial \epsilon_{cut}}{\partial N_1} \cdot \sigma_{N_1}\right)^2 + \left(\frac{\partial \epsilon_{cut}}{\partial N_2} \cdot \sigma_{N_2}\right)^2 \tag{F.3}$$

$$= \left[\frac{N_2}{(N_1 + N_2)^2} \cdot \sigma_{N_1}\right]^2 + \left[\frac{N_1}{(N_1 + N_2)^2} \cdot \sigma_{N_2}\right]^2. \tag{F.4}$$

Assuming that all involved fits behave properly, in addition to $N = N_1 + N_2$ one can also assume that $\sigma^2_N = \sigma^2_{N_1} + \sigma^2_{N_2}$, such that a third histogram with the events failing the cut is

not even needed, and the above equations can be reduced to:

$$\sigma^2_{\epsilon_{cut}} = \left[\frac{N - N_1}{N^2} \cdot \sigma_{N_1}\right]^2 + \left(\frac{N_1}{N^2}\right)^2 \cdot (\sigma^2_N - \sigma^2_{N_1}) \quad \text{(F.5)}$$

$$= (1 - 2\epsilon_{cut}) \cdot \frac{\sigma^2_{N_1}}{N^2} + \epsilon^2_{cut} \cdot \frac{\sigma^2_N}{N^2}. \quad \text{(F.6)}$$

Note that for ideal fits, which perfectly separate signal from background, one expects $\sigma_{N_1} = \sqrt{N_1}$ and $\sigma_N = \sqrt{N}$, and the equation F.6 simplifies to the formulae expected from the binomial statistics (see equation F.2).

This method of determining the error of an efficiency was used in this analysis for the trigger (section 6.4), $f_{p_t}(D^*\mu)$ cut (section 6.6) and the L4 efficiencies (section 6.7).

An extensive discussion about the treatment of errors when calculating efficiencies can be found for example in [87].

Appendix G

Monte Carlo Data Sets

CHARM Monte Carlo				
Run period	Generator	$\mathcal{L}$ [pb^{-1}]	PDF p/γ	Main cuts
HERA I	PYTHIA 6.1 DIR	1574.68	CTEQ5L	$0\ \text{GeV}^2 < Q^2 < 1\ \text{GeV}^2$
	PYTHIA 6.1 EXC	832.18	GRV-G LO	$0 < y < 1$
	CASCADE 1.0	702.36		$p_t(D^*) > 1$ GeV
				$15^\circ < \theta(D^*) < 165^\circ$
HERA II	PYTHIA 6.2 DIR	895.016	GRV-LO	$0\ \text{GeV}^2 < Q^2 < 5\ \text{GeV}^2$
	PYTHIA 6.2 EXC	954.562	GRV-G LO	$0.01 < y < 0.9$
	CASCADE 1.2	432.127		$p_t(D^*) > 1.3$
				$\lvert\eta(D^*)\rvert < 2.5$
BEAUTY Monte Carlo				
HERA I	PYTHIA 6.1 DIR	15634.15	CTEQ5L	$0\ \text{GeV}^2 < Q^2 < 1\ \text{GeV}^2$
	PYTHIA 6.1 EXC	13871.89	GRV-G LO	$0 < y < 1$
	CASCADE 1.0	7796.71		$p_t(D^*) > 1$ GeV
				$15^\circ < \theta(D^*) < 165^\circ$
HERA II	PYTHIA 6.2 DIR	32774	GRV-LO	$0\ \text{GeV}^2 < Q^2 < 5\ \text{GeV}^2$
	PYTHIA 6.2 EXC	92863.2	GRV-G LO	$0.01 < y < 0.9$
	CASCADE 1.2	26550.8		$p_t(D^*) > 1.3$
				$\lvert\eta(D^*)\rvert < 2.5$

Table G.1: *Monte Carlo data sets used in this analysis. The label 'DIR' ('EXC') stands for the direct (excitation) component in PYTHIA. For every Monte Carlo, events with* $c\bar{c}$ $(b\bar{b}) \to D^* \to K\pi\pi_s$ *were generated. The heavy quark masses were set to* $m_c = 1.5$ *GeV and* $m_b = 4.8$ *GeV.*

Appendix H

Quark Flavour Separation

In chapter 5, the method used to extract the charm and beauty fractions for HERA I+II run periods data sample was presented. For completeness, the fit results are presented here separately for HERA I (see figure H.1 and table H.1) and HERA II run periods (figure H.2 and table H.2). The fractions obtained for the two run periods are compatible, within the errors.

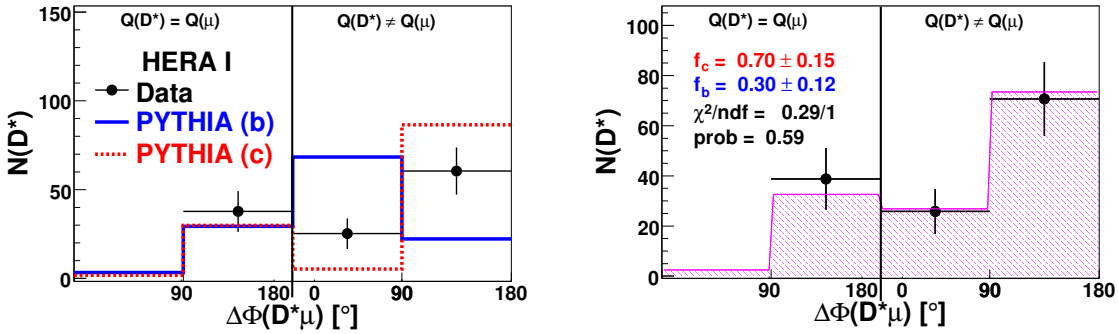

Figure H.1: *Left: distributions of events in the four correlations regions for* ***HERA I*** *data (black points) and initial beauty and charm PYTHIA Monte Carlo which are input to the fit procedure. The Monte Carlo distributions are individually normalised to data. Right: distribution of events in the four correlation regions determined with the* χ^2 *fit (hashed region) compared with data (black points). The obtained* χ^2/ndf*, where 'ndf' is the number of degrees of freedom, and the fit probability are also shown.*

Region	1	2	3	4	$\sum$
Before muon background correction					
$N_c^{reg\,i}$	1.3 ± 0.3	23 ± 5	4 ± 1	66 ± 5	94 ± 21
$N_b^{reg\,i}$	1.1 ± 0.4	10 ± 4	23 ± 9	7 ± 3	41 ± 16
$f_c^{reg\,i}$ [%]	54 ± 11	70 ± 10	15 ± 6	90 ± 4	$\mathbf{70 \pm 10}$
After muon background correction					
$N_c^{sig,\,reg\,i}$	0	0.7 ± 0.2	1.1 ± 0.2	53 ± 12	54 ± 12
$N_b^{sig,\,reg\,i}$	1.0 ± 0.4	9 ± 4	22 ± 9	7 ± 3	39 ± 15
$f_c^{sig,\,reg\,i}$ [%]	0	7 ± 3	5 ± 2	88 ± 5	$\mathbf{58 \pm 11}$

Table H.1: *Distribution of charm and beauty events in the four correlation regions for* ***HERA I*** *run period, before and after muon background correction. The corresponding fractions of charm events are also shown.*

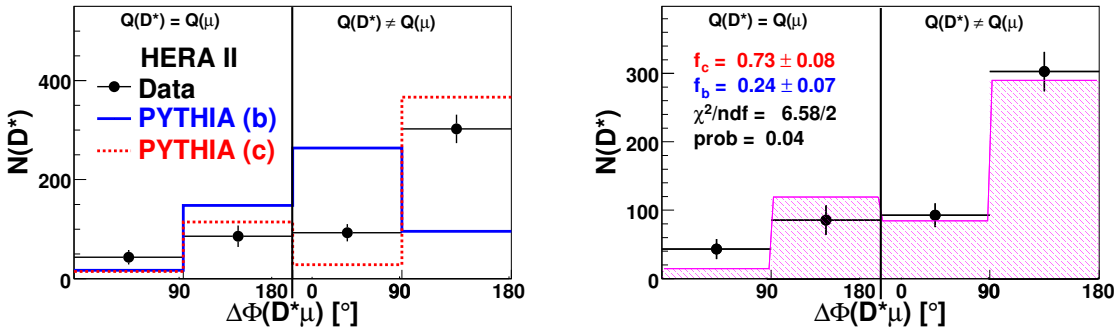

Figure H.2: *Left: distributions of events in the four correlations regions for* ***HERA II*** *data (black points) and initial beauty and charm PYTHIA Monte Carlo which are input to the fit procedure. The Monte Carlo distributions are individually normalised to data. Right: distribution of events in the four correlation regions determined with the χ^2 fit (hashed region) compared with data (black points). The obtained χ^2/ndf, where 'ndf' is the number of degrees of freedom, and the fit probability are also shown.*

Region	1	2	3	4	$\sum$
Before muon background correction					
$N_c^{reg\,i}$	11 ± 1	84 ± 10	20 ± 2	267 ± 31	381 ± 42
$N_b^{reg\,i}$	4 ± 1	36 ± 10	64 ± 17	23 ± 6	127 ± 34
$f_c^{reg\,i}$ [%]	72 ± 6	70 ± 6	24 ± 5	92 ± 2	$\mathbf{75 \pm 5}$
After muon background correction					
$N_c^{sig,\,reg\,i}$	0	5 ± 1	7 ± 1	228 ± 27	239 ± 27
$N_b^{sig,\,reg\,i}$	4 ± 1	34 ± 9	62 ± 17	21 ± 6	120 ± 32
$f_c^{sig,\,reg\,i}$ [%]	0	12 ± 3	10 ± 3	92 ± 2	$\mathbf{67 \pm 6}$

Table H.2: *Distribution of charm and beauty events in the four correlation regions for* ***HERA II*** *run period, before and after muon background correction. The corresponding fractions of charm events are also shown.*

List of Figures

List of Tables

References

[1] M. Gockeler *et al.*, "A Determination of the Lambda Parameter from Full Lattice QCD", *Phys. Rev.* **D73** (2006) 014513, hep-ph/0502212. 10, 17

[2] B. Foster, "Lectures on HERA physics", hep-ex/0206011. 13

[3] M. Erdmann, *The Partonic Structure of the Photon: Photoproduction at the Lepton-Proton Collider HERA*. Springer Verlag, 1997. 13, 21

[4] C. Friberg and T. Sjöstrand, "Jet Production by Real and Virtual Photons", (1999) hep-ph/9906407. 14

[5] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and Sons, 1984. 14, 15

[6] http://nobelprize.org/nobel_prizes/physics/laureates/2004/. 14

[7] R. K. Ellis, W. J. Stirling and B. R. Webber, *"QCD and Collider Physics"*. Cambridge University Press, 1996. 14

[8] B. P. Roe, *Particle Physics at the New Millennium*. Springer Verlag, 1996. 15

[9] M. Mansfield and C. O. Sullivan, *Understanding Physics*. John Wiley & Sons, 1998. 15

[10] G. Sterman *et al.*, "Handbook of Perturbative QCD", *Rev. Mod. Phys.* **67** (January, 1995) 157–248. 15

[11] M. Kuhlen, *QCD at HERA, The Hadronic Final State in Deep Inelastic Scattering*, vol. 150. Springer Verlag, 1999. 16, 18

[12] J. C. Collins, D. E. Soper and G. Sterman, "Factorization of Hard Processes in QCD", *Adv. Ser. Direct. High Energy Phys.* **5** (1988) 1–91, hep-ph/0409313. 16

[13] J. C. Collins, L. Frankfurt and M. Strikman, "Proof of Factorisation for Exclusive Deep-Inelastic Processes", hep-ph/9709336. 16

[14] J. C. Collins, "Proof of Factorisation for Diffractive Hard Scattering", *Phys. Rev.* **D57** (1998) 3051–3056, hep-ph/9709499. 16

[15] **Particle Data Group** Collaboration, "Review of Particle Physics", *J. Phys. G: Nucl. Part. Phys* **33** (2006). http://pdg.lbl.gov. 16, 17, 27, 34, 36, 56, 70, 71, 79, 80, 101, 103, 130

[16] H. Quinn, "Heavy Flavour Physics", *European School of High Energy Physics* (2002). 17

[17] Yu. L. Dokshitzer, Sov. Phys. JETP 46 (1977) 641; V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 438, 675; G. Altarelli and G. Parisi, Nucl. Phys. 126 (1977) 297. 17

[18] A. V. Lipatov and N. P. Zotov, "Unintegrated Gluon Gistributions in $D^{*\pm}$ and Dijet Associated Photoproduction at HERA", hep-ph/0512013. 18

[19] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Sov. Phys. JETP 45 (1977) 199; Y.Y. Balitsky and L.N. Lipatov, Sov. J. Nucl. Phys. 28 (1978) 282. 18

[20] M. Ciafaloni, Nucl. Phys. B296 (1988) 49; S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. B234 (1990) 339; Nucl. Phys. B336 (1990) 18. 19

[21] Y. L. Dokshitzer, V. A. Khoze and S. I. Troyan, "QCD Coherence in High-energy Reactions", *Rev. Mod. Phys.* **60** (1988) 373–388. 19, 20

[22] H. Jung, "The CCFM Monte Carlo generator CASCADE", *Comput. Phys. Commun.* **143** (2002) 100–111. 20, 26

[23] T. H. Bauer, R. D. Spital and D. R. Yennie, "The Hadronic Properties of the Photon in High Energy Interactions", *Rev. Mod. Phys.* **50** (1978) 261–436. 21

[24] A. V. Lipatov and N. P. Zotov, "Unintegrated Gluon Distributions in $D^{*\pm}$ and Dijet Associated Photoproduction at HERA", hep-ph/0512013. 21

[25] Frixione, Stefano and Nason, Paolo and Ridolfi, Giovanni, "Differential Distributions for Heavy Flavor Production at HERA", *Nucl. Phys.* **B454** (1995) 3–24, hep-ph/9506226. 22

[26] Harris, B. W. and Smith, J., "Heavy Quark Correlations in Deep Inelastic Electroproduction", *Nucl. Phys.* **B452** (1995) 109–160, hep-ph/9503484. 22

[27] Amundson, James and Schmidt, Carl and Tung, Wu-Ki and Wang, Xiaoning, "Charm Production in Deep Inelastic Scattering from Threshold to High Q^2", *JHEP* **10** (2000) 031, hep-ph/0005221. 23

[28] Cacciari, Matteo and Frixione, Stefano and Nason, Paolo, "The p_T Spectrum in Heavy-Flavor Photoproduction", *JHEP* **03** (2001) 006, hep-ph/0102134. 23

[29] B. Andersson *et al.*, "Parton Fragmentation and String Dynamics", *Phys. Rep.* **97** (1983) 31–145. 24

[30] B. Andersson, S. Mohanty and F. Söderberg, "Recent Developments in the Lund Model", hep-ph/0212122. 24

[31] T. Sjostrand, S. Mrenna and P. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) hep-ph/0603175. 24, 25

[32] C. Peterson, D. Schlatter, I. Schmitt and P. M. Zerwas, "Scaling Violations in Inclusive e^+e^- Annihilation Spectra", *Phys. Rev.* **D** (1983) 105–111. 25

[33] H. L. Lai *et al.*, "Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions", *Eur. Phys. J.* **C12** (2000) 375–392, hep-ph/9903282. 25

[34] M. Glück, E. Reya and A. Vogt, "Dynamical Parton Distributions Revisited", *Eur. Phys. J.* **C5** (1998) 461–470, hep-ph/9806404. 25

[35] **H1** Collaboration, "Guide for the H1 Simulation Program, H1SIM", *H1 Internal Note, 03-11/89* (1989). 26

[36] **H1** Collaboration, "H1 Reconstruction Program, H1REC",. https://www-h1.desy.de/icas/imanuals/h1rec/h1rec9/doc.html. 26

[37] M. Peez, B. Portheault and E. Sauvan, "An Energy Flow Algorithm for Hadronic Reconstruction in OO: Hadroo2", *H1 Internal Note, H1-01/05-616* (2005). 26

[38] S. D. Ellis and D. E. Soper, "Successive Combination Jet Algorithm for Hadron Collisions", *Phys. Rev.* **D48** (1993) 3160–3166, hep-ph/9305266. 28

[39] **H1** Collaboration, "Inclusive $D^{*\pm}$ Meson Cross Sections and $D^{*\pm}$-Jet Correlations in Photoproduction at HERA", hep-ex/0608042. 28, 29

[40] A. B. Meyer, "Heavy Quark Production at HERA." Univ. Hamburg, Habilitation, 2005. 29, 40

[41] **H1** Collaboration, "Inclusive Production of D^+, D^0, D_s^+ and D^{*+} Mesons in Deep Inelastic Scattering at HERA", *Eur. Phys. J.* **C38** (2005) 447–459, hep-ex/0408149. 29, 30

[42] **H1** Collaboration, A. Aktas *et al.*, "Measurement of Beauty Production at HERA using Events with Muons and Jets", *Eur. Phys. J.* **C41** (2005) 453–467, hep-ex/0502010. 31, 32, 34, 107

[43] O. Behnke, "Production of Charm and Beauty Quarks at HERA." Ruprecht-Karls-Univ, Habilitation, 2005. 32, 40, 57

[44] **H1** Collaboration, A. Aktas *et al.*, "Measurement of Charm and Beauty Dijet Cross Sections in Photoproduction at HERA using the H1 Vertex Detector", *Eur. Phys. J.* **C47** (2006) 597–610, hep-ex/0605016. 32, 33, 107

[45] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, "On the Determination of the Gluon Density of the Proton from Heavy Flavor Production at HERA", *CERN-TH.6864/93* (1993) hep-ph/9304289. 38

[46] **ZEUS** Collaboration, "Measurement of Beauty Production from Dimuon Events at HERA." Contributed paper no. 269 at the XXII International Symposium on Lepton-Photon Interactions at High Energy, Uppsala, Sweden, 2005. 39, 40

[47] **H1** Collaboration, A. Aktas *et al.*, "Measurement of Charm and Beauty Photoproduction at HERA using $D^*\mu$ Correlations", *Phys. Lett.* **B621** (2005) 56–71, hep-ex/0503038. 40, 107, 108

[48] **ZEUS** Collaboration, "Measurement of Open Beauty Production in Deep Inelastic Scattering at HERA using a $D^* + \mu$ Tag." Contributed paper no. 575 at International Europhysics Conference On High Energy Physics (EPS 2003), Aachen, Germany, 2003. 40

[49] F. Willeke, "HERA and the Next Generation of Lepton-Ion Colliders." Proceedings of EPAC 2006, Edinburgh, Scotland. 41

[50] **H1** Collaboration, "The H1 Detector at HERA", *Nucl. Instr. Meth.* **A386** (1997) 310–347. 42, 43, 44, 48, 49

[51] D. Pitzl *et al.*, "The H1 Silicon Vertex Detector", *Nucl. Instrum. Meth.* **A454** (2000) 334–349, hep-ex/0002044. 42

[52] R. K. Bock and A. Vasilescu, *The Particle Detector Briefbook*. Springer, 1998. 43, 44

[53] S. Lüders, "CST/CIP 2000 Stand-alone Tracking and Vertexing", *H1 Internal Note, H1-04/99-570* (1999). 43

[54] H. Itterbeck, *Techniques and Physics of the Central-Muon-Trigger System of the H1-Detector at HERA*. Ph.D. Dissertation, RWTH Aachen, September, 1997. 46

[55] B. Naroska, S. Schiek and G. Schmidt, "Lepton Identification in the H1 Detector at Low Momenta", *H1 Internal Note, H1-05/97-518* (1997). 46, 61, 79

[56] C. Kleinwort and U.-P. Krueger, "Track Reconstruction in the Iron", *H1 Software Note 92-035* (1992). 46

[57] N. Gogitidze and S. Levonian, "An Offline Luminosity Determination for the 1995 e^+p Data", *H1 Internal Note, H1-02/96-471* (1996). 48

[58] S. Levonian and A. Panitch, "Treatment of the Proton Satellite Bunches in 1994 Data", *H1 Internal Note, H1-09/95-454* (1995). 49

[59] S. Levonian, "The Status of H1 Luminosity Precision." Talk given in the H1 Collaboration Meeting, DESY, Hamburg, Feb., 2007. 49

[60] M. Jacquet, "The Steering of the H1 Level 2 Topological Trigger", *H1 Internal Note, H1-06/97-519* (1997). 50

[61] J. Riedlberger, "The H1 Trigger with Emphasis on Tracking Triggers", *H1 Internal Note, H1-01/95-419* (1995). 51

[62] B. Hanspeter, "Principles and Operation of the z-Vertex Trigger", *H1 Internal Note, H1-05/96-479* (1996). 51

[63] A. Baird *et al.*, "A Fast High Resolution Track Trigger for the H1 Experiment", *IEEE Trans. Nucl. Sci.* **48** (2001) 1276–1285, hep-ex/0104010. 51, 52

[64] The H1OO Group, "The H1 OO Physics Analysis Project",. available from the H1 Collaboration. 53

[65] http://root.cern.ch/. 53

[66] D. Pitzl, "The H1 Detector for HERA II." ICHEP, Amsterdam, 2002. 53

[67] F. Jacquet and A. Blondel, "Proceedings of the Study for an *ep* Facility for Europe." ed. U.Amaldi. DESY 79/48 (1979) 391. 55

[68] U. Kathage, *Photoproduktion von $D^{*\pm}$ Mesonen bei HERA*. Ph.D. Dissertation, Christian-Albrechts-Universität zu Kiel, April, 1997. 55, 57

[69] L. West, "How to Use the Heavy Flavour Working Group Track, Muon and Electron Selection Code", *H1 Internal Note* (1997). 56, 60

[70] W. Erdmann, *Untersuchung der Photoproduktion von D^*-Mesonen am ep-Speichering HERA*. Ph.D. Dissertation, Univ. Hamburg, 1996. DESY FH1K-96-02. 59, 102

[71] J. Wagner, *Charm and Beauty Production at HERA with $D^*\mu$ Events*. Ph.D. Dissertation, Hamburg Univ., 2004. DESY-THESIS-2004-022. 60, 61, 102, 109

[72] Heavy Flavour Working Group, H1 Collaboration, "Open Charm Finder for Level 4: HQSEL", 1999. https://www-h1.desy.de/h1/iww/iwork/ihq/sw-doc/hqsel99.html. 62

[73] P. Merkel (H1 Collaboration), "HMASS Finder on L4 for 1999 e^+ Running", 1999. https://www-h1.desy.de/h1/iww/iwork/ihq/sw-doc/hqsel99.html. 62

[74] S. Egli, E. Elsen, V. Lemaitre, K. Mueller, H. Rick and H.-C. Schultz-Coulon, "Calculating Event Weights in Case of Downscaling on Trigger Levels 1-4", *H1 Internal Note, H1-04/97-517* (1997). 73, 74

[75] U. Langenegger and G. Tsipolitis, "On Muon Identification in the Barrel of the H1 Detector", *H1 Internal Note, H1-04/98-542* (1998). 79

[76] A. Longhin, *Measurement of Beauty Production at HERA with a $D^* + \mu$ Tag*. Ph.D. Dissertation, Padova University, Science Faculty, October, 2003. DESY-THESIS-2004-050. 80, 87, 88

[77] H. P. William *et al.*, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988-1992. 84

[78] K. Hagiwara *et al.*, "Review of Particle Physics", *Phys. Rev. D.* **66** (2002) 010001+. http://pdg.lbl.gov. 86

[79] K. Krüger, *Photoproduction of J/ψ Mesons at Medium and Low Elasticities at HERA*. Ph.D. Dissertation, Univ. Hamburg. 93

[80] P. Fleischmann, *Elastic J/ψ Production at HERA*. Ph.D. Dissertation, Univ. Hamburg. DESY-THESIS-2004-013. 93, 102

[81] H1 Trigger Hypernews Forum, message 478, http://www-h1.desy.de/icgi-hypernews/get/Detector/Trigger.html. 95

[82] D. Meer, *Heavy Quarks in Photoproduction at HERA and a Linking Algorithm for the Fast Track Trigger*. Ph.D. Dissertation, Swiss Federal Institute of Technology Zürich, 2005. Diss. ETH No. 15875, ETHZ-IPP Intern. Report 2005-01. 97

[83] S. Hengstman, *A Measurement of Diffraction Charm Production at HERA*. Ph.D. Dissertation, Univ. Zürich, 2000. 102, 103

[84] **ZEUS** Collaboration, S. Chekanov *et al.*, "Measurement of Open Beauty Production at HERA in the $D^*\mu$ Final State", *Eur. Phys. J. C* **50** (2007) 299–314. 107, 108

[85] M. Martisikova, *Jet Shapes in Charm Photoproduction at HERA*. Ph.D. Dissertation, 2005. DESY-THESIS-2005-047. 113

[86] **CDF** Collaboration, "Questions Submitted to the CDF Statistics Committee." http://www-cdf.fnal.gov/physics/statistics/statistics_faq.html#ssl4. 138

[87] T. Ullrich and Z. Xu, "Treatment of Errors in Efficiency Calculations", physics/0701199. Jan, 2007. 139

Acknowledgements

I thank my supervisor Dr. Eckhard Elsen, for the three years in which he found time to answer my questions about Physics (and not only about that), for all the evening explanations at the blackboard, and especially for taking away my anxiety over deadlines.

I thank the referees, Dr. Achim Geiser and Prof. Dr. R.-D. Heuer, for agreeing to evaluate my work.

I thank the persons who helped in proofreading of the thesis: Dr. Hannes Jung, Zlatka Staykova and Lluis Marti. Special thanks to my colleague, Dr. Martin Göttlich, for all the useful comments, and for the encouragements when it was difficult for me to see the light at the end of the tunnel.

I thank Prof. Dr. Beate Naroska and Dr. Carsten Niebuhr for offering me the possibility to work in the H1 collaboration.

My thanks go to the office colleagues: Svetlana Vinokurova - for being a good friend, Michal Deak, Dr. Krzysztof Kutak, Julia Ruiz Tabasco - for supporting me.

I also thank my Professor at the University of Bucharest, Ionel Lazanu, for making me like particle Physics and for supporting me to go this way.

I thank Renate Wilke, for welcoming me in her home, for her singing and her smile, and above all for being the perfect fairy-tale grandmother.

I thank Adrian, for showing me the Universe through his eyes, and for all the other miracles which my words cannot describe.