Universität Bonn

Physikalisches Institut

Measurement of beauty production in deep inelastic scattering at HERA using decays into electrons

Ramoona Shehzadi

The production of beauty quarks in ep interactions has been studied with the ZEUS detector at HERA for exchanged four-momentum squared $Q^2 > 10\,\mathrm{GeV}^2$, using an integrated luminosity of $363\mathrm{pb}^{-1}$. The beauty events were identified using electrons from semileptonic b decays with a transverse momentum $0.9 < p_T^e < 8\mathrm{GeV}$ and pseudorapidity $|\eta^e| < 1.5$. The fraction of events containing b quarks was extracted from a likelihood fit using variables sensitive to electron identification as well as to semileptonic decays. Cross sections for beauty production were measured and compared with leading-order plus parton-shower Monte Carlo predictions and with next-to-leading-order QCD calculations. The beauty contribution to the proton structure function F_2 was extracted from the double-differential cross section as a function of Bjorken-x and Q^2.

Physikalisches Institut der
Universität Bonn
Nußallee 12
D-53115 Bonn

BONN-IR-2011-01
January 2011
ISSN-0172-8741

Universität Bonn
Physikalisches Institut

Measurement of beauty production in deep inelastic scattering at HERA using decays into electrons

Ramoona Shehzadi

Dieser Forschungsbericht wurde als Dissertation von der
Mathematisch-Naturwissenschaftlichen-Fakultät der Universität Bonn
angenommen und ist auf der ULB Bonn `http://hss.ulb.uni-bonn.de/diss_online`
elektronisch publiziert.

Referent:	Prof. Dr. Ian C. Brock
Koreferent:	Prof. Dr. Jochen Dingfelder
Angenommen am:	08.11.2010
Tag der Promotion:	14.01.2011

To the Memory of My Father

Contents

Chapter 1

Introduction

Particle physics is concerned with the study of the basic elements of matter and the forces acting among them. It allows us to address questions such as "what is the world made of?" and "what holds the world together?". It aims to determine the fundamental laws that control the make-up of matter and the physical universe.

The idea of fundamental building blocks started from the time of the Greeks, with their designation of the four elements: earth, air, fire and water. Over centuries of thoughts and experiments, there has gradually arisen a powerful belief in a simple frame-work underlying the seeming complexity of the universe. For example, until the late 19^{th} century, atoms were thought to be the fundamental indivisible building blocks of all forms of matter. Simplifications came from the identification of the electron, proton and neutron, which were considered as the fundamental particles of nature, when we learned through the experiment of Rutherford in 1909 that atoms consist of mostly empty space with electrons surrounding a dense central nucleus made up of protons and neutrons. The electrons are bound into the atoms by electromagnetism, since they are attracted by the opposite charge of the nuclear protons. But in order to ensure that the nuclei does not disintegrate because of the mutual repulsion of their constituent protons, a new short-range force, the strong nuclear interaction, is required.

This simple picture did not last long. Antiparticles, which have properties such as equal mass and other properties such as opposite charge to particles, were predicted and later discovered. Experiments on radioactive beta decay, where one element spontaneously converts into another with the emission of an electron, appeared to violate the laws of conservation of energy and momentum. To conserve these quantities it was postulated that an additional particle, named the neutrino, was emitted which carried away the missing energy and momentum.

The new particle, the neutrino, neither feels the electromagnetic nor the strong force, and hence escaped undetected in these beta-decay experiments. It is created by another new short-range force, the weak interaction, which is so feeble that neutrinos on average can penetrate light-years of material such as iron before having a significant chance of

interaction; they were therefore thought to be undetectable. However, in the 1950's the huge flux of neutrinos coming from interactions in nuclear reactors led to their detection.

The science of particle physics surged forward with the invention of particle accelerators that could accelerate electrons or protons to high energies and smash them into nuclei. To the surprise of the physicists, accelerator experiments revealed that the world of particles was very rich; many more particle types similar to protons and neutrons, called baryons, and a whole new family of particles, called mesons, were discovered. By the early 1960's, as accelerators reached higher energies, a hundred or more types of particles had been identified. The discovery of these particles complicated the simple structure of the atom, consisting of three particles. The question arose, could all of these then be the new fundamental particles? Confusion reigned until the discovery of quarks.

In 1964, two physicists, Murray Gell-Mann [1] and George Zweig [2], independently hit upon the idea that neutrons and protons and all those new particles could be built up by a few types of yet smaller objects; Gell-Mann called them quarks. They could explain all the observed baryons and mesons with just three types of quarks, now called up, down, and strange, and their antiquarks.

In many experiments since then, the quark idea has been confirmed. It is now part of the Standard Model of fundamental particles and interactions [3, 4]. The discovery of further particles has shown that there are six types of quarks, namely up, down, strange, charm, bottom, and top, in order of increasing mass. Also, there are six types of leptons, the electron, muon and tau, and their neutrinos. The Standard Model incorporates the quarks and leptons as well as their interactions through the strong, weak and electromagnetic forces.

Quarks are not observable on their own due to the strong force acting between them. The strong force increases linearly with the distance between two quarks and binds the quarks into observable bound states called hadrons[1]. The charm, bottom and top quarks are called the heavy quarks. They have masses heavier than the energy scale at which quarks are confined into hadrons ($\Lambda \sim 250$ MeV). The remaining three quarks are called light flavour quarks. There are two classes of hadrons: baryons, which contain three quarks, and mesons, which contain one quark and one antiquark. The proton is the lightest baryon consisting of two up quarks and one down quark.

The theory describing the strong interactions of the Standard Model is Quantum Chromodynamics (QCD). In QCD, gluon exchange is responsible for the interactions between quarks. QCD can be tested by scattering electrons off protons. If the energy transfer is large enough, the electron scatters off a quark inside the proton. When the quark is kicked out of the proton, the process is called deep inelastic scattering. If the energy involved in the collision is sufficiently high, a heavy quark-antiquark pair can be produced.

The first deep inelastic scattering experiments were done in 1969 at the Stanford Linear

[1]The top quark decays as $t \to bW^+$; because the top quark decay time is much shorter than the hadron formation time, no top hadrons are expected to be observed.

Accelerator (SLAC). Electrons of 7 GeV were collided with a proton in a hydrogen target [5, 6]. At these experiments the structure of the proton was measured for the first time.

To probe the proton even deeper a new particle accelerator, HERA, was built in Hamburg. At HERA electrons or positrons and protons were accelerated up to 27.5 GeV and 920 GeV, respectively, and brought into collision at two points, such that the electrons (positrons) were scattered off the constituents of the protons. The detectors H1 in the North Hall and ZEUS in the South Hall were set up around the collision points. At HERA, data were taken between 1992 and 2007. The scattered lepton as well as the other particles produced in a hard scattering process could provide information about the internal structure of the proton. This information can be extracted by means of the parton density functions (PDFs), which represent the probabilities for finding a certain parton in the proton. Parton is the generic name for quarks and gluons. The determination of PDFs was one of the major aspects of the HERA physics programme.

The lepton-proton interactions can be split into three main categories; neutral current (NC) deep inelastic scattering (DIS), charged current (CC) DIS and photoproduction[2]. Within these interactions, many areas of physics were studied at HERA: structure functions which are related to parton density functions, tests of QCD, heavy flavour production, electroweak processes, search for physics beyond the Standard Model, exotic states, rare phenomena and diffractive processes.

The production of heavy quarks, beauty (b)[3] and charm (c), is a rich testing ground for the QCD sector of the Standard Model. The dominant production mechanism for heavy quarks is boson-gluon fusion, which gives a direct handle on the gluon density in the proton. At the ZEUS experiment, heavy quark cross sections have been measured using different approaches with complementary systematic uncertainties.

Several analysis of charm used the reconstruction of $D^{*\pm}$ mesons, which are produced in large numbers at HERA, and give a clean signal. Beauty production is harder to tag owing to the larger mass and lower charge of the beauty quark which suppresses its production. Its dominant decay is to charm hadrons, which then decay further, making it hard to explicitly reconstruct the decay of the beauty hadrons.

Recently, an inclusive method, which was not restricted to a specific decay channel, has been developed to look for heavy quark production [7, 8]. This method exploits the lifetime information of the heavy quarks. This method has been used for the data-taking period 2003–2007 (HERA II data), where the presence of the microvertex detector in the ZEUS experiment allowed the identification of events containing secondary vertices originating from the decay of particles with long lifetime ($c\tau \geq 100\,\mu$m).

An easier way to tag beauty production is to look for the semileptonic decays to muons [9, 10]. This method has also been used for the measurement of charm production [10]. An alternative method to tag charm and beauty is to look for the semileptonic decays to

[2]The properties of these interactions are discussed in Section 2.2.

[3]The b quark is usually referred to as the bottom quark, while the term *beauty* is the nomenclature typically used at HERA experiments.

electrons. This can provide a complementary measurement of heavy quark production to the methods already outlined, and is a very similar decay mechanism to the semileptonic decay to muons. In most of the analyses at ZEUS, the semileptonic muon channel was used as muons are relatively easy to identify, whereas the electron candidates from the semileptonic decay of heavy hadrons suffer from substantial background from electrons that are produced copiously from other processes such as from photon conversions and Dalitz decays. Also electrons are not particularly easy to detect as they can be faked by other particles, mostly pions, kaons and protons, which are produced in much larger numbers than electrons.

An advantage of using electrons rather than muons is that electrons are detected using the inner components of the detector (tracking chamber and calorimeter) whereas the outer muon chambers are used to detect muons. Muons have to traverse the rest of the detector material to reach the muon chambers, therefore restricting the measurement only to muons with a high transverse momentum of at least $\sim 2\,\mathrm{GeV}$. The momentum needed for electrons to be measured in the inner detectors is much less and hence allows the decay lepton to be measured at lower transverse momentum. The use of the electron decay channel is interesting to complement the muon analyses in terms of the kinematic range and also with respect to the systematic uncertainties.

The aim of the analysis described in this thesis is the measurement of beauty production in DIS using decays into electrons and also exploiting the lifetime information. For this, the $e^{\pm}p$ collision data at HERA recorded with the ZEUS detector in the years 2004–2007 are used. The extraction of the beauty signal is done using a likelihood ratio hypothesis test which allows the electron identification information to be combined with the semileptonic decay kinematics. The production of beauty is measured, and the beauty contribution to the proton structure function is calculated.

The structure of this thesis is described in the following:
Chapter 2 gives an overview of the theoretical concepts relevant for lepton-proton interactions, with an emphasis on those needed for the analysis of beauty production in deep inelastic scattering. The leading order and next-to-leading order theoretical predictions which are compared to the measurements at the end, are also introduced in this chapter.
Chapter 3 presents the published results for beauty production in deep inelastic scattering from the H1 and ZEUS collaborations at HERA.
Chapter 4 introduces briefly the HERA collider and the experimental setup for the ZEUS detector. The detector components that are particularly important for this analysis are described in more detail.
Chapter 5 describes the reconstruction of the events, with special emphasis on the reconstruction of particle tracks, the scattered electron and the hadronic system. In addition different methods to reconstruct the event kinematics are discussed.
Chapter 6 summarises the data and Monte Carlo samples used in this analysis followed by the details of different cuts for event and candidate selection. Some control distributions necessary to investigate the validity of the Monte Carlo simulation are also shown.
Chapter 7 introduces several discriminating variables which were used to distinguish be-

tween signal and background events.
Chapter 8 defines the likelihood method used for signal extraction. Some studies showing the likelihood performance are also presented.
Chapter 9 gives a description of the ingredients needed for the cross section determination. The corrections applied to the Monte Carlo are listed before the fitting method for the signal extraction is described. Finally some control distributions for different selection stages are shown.
Chapter 10 presents a description of the determination of the systematic uncertainties. A list of systematic and consistency checks is summarised. A comparison between the statistical errors and the sum of all systematic uncertainties is shown.
Chapter 11 shows the results of the analysis: the total and differential cross sections are compared with the leading order and next-to-leading order QCD predictions. The extraction of the beauty contribution to the proton structure function and its comparison with published results as well as different theoretical predictions is also presented.
Chapter 12 summarises the essential aspects and results of this thesis.

Chapter 2

Theoretical Overview

In this chapter the theoretical basis relevant for this analysis is presented. After a short introduction to the Standard Model of particle physics, the kinematics and properties of lepton-proton scattering, in particular the *deep inelastic scattering* (DIS) and calculation of its cross section are described. A short review of QCD, the theory of strong interactions follows. The chapter then focuses on heavy quark production in DIS, describing the decay channel used in this analysis. Afterwards, the characteristics of the event simulation approach used to calculate the acceptance corrections are presented. The chapter ends with an introduction to the HVQDIS program, which is used to extract the *next-to-leading order* (NLO) QCD predictions.

2.1 The Standard Model

The *Standard Model* (SM) [3, 4] of particle physics is a successful theory of the elementary particles and the interactions governing them, which provides our current understanding of the universe. In the SM, all matter is made of two kinds of elementary particles: quarks and leptons. These are spin-$\frac{1}{2}$ particles that satisfy Fermi-Dirac statistics and are known as fermions. Quarks and leptons are further grouped into three generations which are listed in Table 2.1. Each generation of these particles comes in a doublet. The first particle of each lepton doublet (electron neutrino, muon neutrino and tau neutrino) is not electrically charged while the second particle (electron, muon and tau) has charge $-1e$. For each quark doublet, the upper element (up, charm and top) has fractional charge $+\frac{2}{3}e$, while the lower element (down, strange and bottom) has fractional charge $-\frac{1}{3}e$. Masses of quarks and leptons increase with the generation. The charm (c), bottom (b) and top (t) quarks, having masses of $\mathcal{O}$(GeV) are categorised as *heavy* quarks, whereas the up (u), down (d) and strange (s) quarks with masses of $\mathcal{O}$(MeV) belong to the category of *light* quarks. The analysis presented in this thesis deals with the second heaviest quark called bottom (or beauty). All particles listed in Table 2.1 have their corresponding antiparticles with the same mass and energy but oppositely signed charge-like quantum numbers.

	Leptons			Quarks		
Generations	Flavour	Q_{EM}	Mass (MeV)	Flavour	Q_{EM}	Mass (MeV)
1st	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}$	0 -1	< 0.003 0.511	$\begin{pmatrix}u\\ d\end{pmatrix}$	$+2/3$ $-1/3$	1.5 to 3.3 3.5 to 6
2nd	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}$	0 -1	< 0.19 105.66	$\begin{pmatrix}c\\ s\end{pmatrix}$	$+2/3$ $-1/3$	1160 to 1340 70 to 130
3rd	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}$	0 -1	< 18.2 1776.8	$\begin{pmatrix}t\\ b\end{pmatrix}$	$+2/3$ $-1/3$	171300 ± 1200 4130 to 4370

Table 2.1: The fundamental particles in the Standard Model ordered in three generations. The electromagnetic charge Q_{EM} is given in multiples of the elementary charge unit.

The elementary particles undergo fundamental interactions in nature by exchanging gauge particles. These interactions are gravitational, electromagnetic, weak and strong. The gauge particles are the graviton for the gravitational force, the photon (γ) for the electromagnetic force, $W^{\pm}$ and Z bosons for the weak force and the gluon for the strong force. They carry spin-1 quantum numbers (except the graviton which carries spin-2) and satisfy Bose-Einstein statistics. Since the spin-1 field transforms as a vector under the Lorentz transformation, they are known as vector bosons. The SM has the ability to describe the properties of three of the four known forces of nature: electromagnetic, weak and strong (see Table 2.2). For elementary particles, the gravitational force is negligible and need not to be considered. Within the SM, the electromagnetic and weak interactions are uni-

Interaction	Theory	Mediator	Symmetry	Charge	Coupling (α_i)
Strong	QCD	8 gluons	$SU(3)_C$	colour	1
Electromagnetic	QED	γ	$U(1)_Y$	electric charge	10^{-2}
Weak	GSW	$W^{\pm}, Z$	$SU(2)_L$	weak isospin	10^{-6}

Table 2.2: The fundamental interactions in the Standard Model. The strength of the interactions is given by their couplings, α_i, at very low energies, $E \ll m_p$.

fied and explained through the electroweak theory [3], while *Quantum Chromodynamics* (QCD) [11] is used to describe the strong interactions. The SM is based on a relativistic gauge field theory and the three interactions can be described as theories of local gauge symmetry i.e. in terms of unitary groups of different dimensions.

The electroweak force is described by a gauge theory based on the $SU(2)_L \otimes U(1)_Y$ symmetry group. The gauge bosons of the weak interaction ($W^{\pm}, Z$) get their masses ($M_W = 80.43$ GeV and $M_Z = 91.19$ GeV) from spontaneous symmetry breaking which introduces a new scalar particle, the Higgs Boson, via the so-called *Higgs Mechanism*. The gauge boson of the electromagnetic interaction (γ) remains massless. Quarks and leptons

also acquire mass via interaction with the Higgs field, which has not been observed so far in the experiments.

The strong interaction is described by the $SU(3)_c$ gauge symmetry. The six quarks and the corresponding antiquarks are the interacting spinors and the resulting gauge bosons (gluons) are massless and electrically neutral. Both quarks and gluons carry colour charge, such that each quark comes in three colours (red, blue or green), whereas there are eight colour combinations of the gluons. Because they carry colour charge, gluons may also interact among themselves. Quarks are not individually observable due to the strong force linearly increasing with the distance between two quarks. Due to this confinement, all particles observed are colourless doublets (quark and antiquark) or triplets (three quarks). The doublets are called *mesons* while the triplets are denoted as *baryons*. Leptons do not carry any colour and are not directly affected by the strong interaction. The weak interaction affects all quarks and leptons.

The strength of the strong interactions between particles is proportional to the strong coupling, α_s. QCD is an asymptotically free theory i.e. the value of the coupling, α_s, decreases as the energy of the interaction increases or equivalently at shorter distances. As a consequence of this, at high energies quarks and gluons behave as almost free particles. QCD will be described in more detail in Section 2.3.3.

2.2 Kinematics of Lepton-Proton Scattering

In the SM, the fundamental lowest order process in lepton-proton scattering is mediated by the electroweak force through the exchange of a vector boson. If the exchanged boson is a photon, γ, or a Z, the process is called *neutral current* scattering (NC); in case the boson is a $W^{\pm}$, the interaction is called *charged current* scattering (CC). In NC scattering a scattered lepton appears in the final state. At HERA such a process is described by:

$$e^{\pm} + p \to e^{\pm} + X, \tag{2.1}$$

where X denotes the hadronic final state. In the other type of process i.e. CC scattering, the final state lepton is a neutrino:

$$e^{+}(e^{-}) + p \to \bar{\nu}(\nu) + X. \tag{2.2}$$

The generic Feynman diagrams of both processes including the particle's momenta are shown in Figure 2.1. The kinematics of ep collisions can be explained with the help of these diagrams as follows.
The incoming lepton, $e^{\pm}$, having a four-momentum $k(E_e,\vec{k})$ interacts with an incoming proton of four-momentum $P(E_p,\vec{P})$. The four-momentum of the resulting scattered lepton is $k'(E'_e,\vec{k'})$ and that of the exchanged boson is $q(E_\gamma,\vec{q})$, where

$$q = k - k'. \tag{2.3}$$

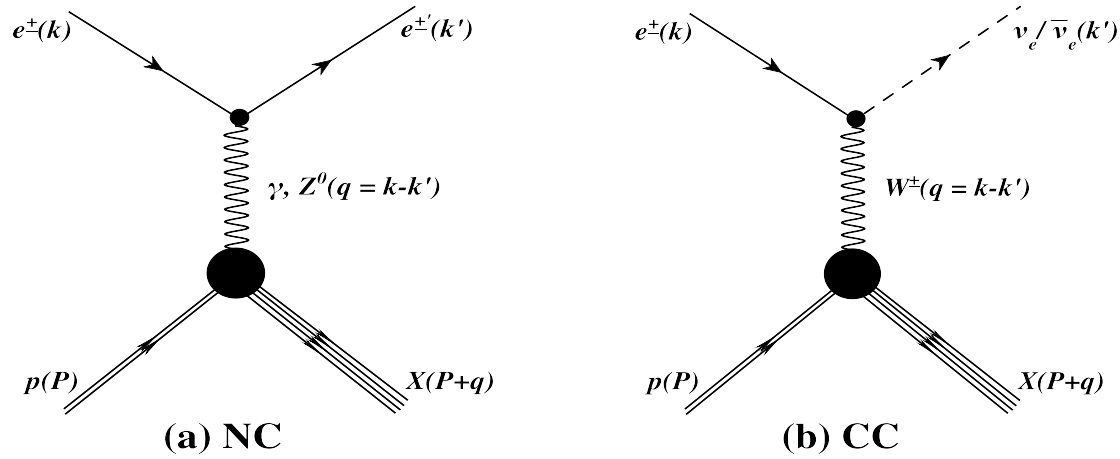


Figure 2.1: Electron-proton scattering: (a) in neutral current (NC), and (b) charged current (CC). The four vectors of the particles are given in parentheses.

The scattering is described by the following Lorentz-invariant variables:

$$s = (k+P)^2, \tag{2.4}$$
$$Q^2 = -q^2 = -(k-k')^2,\ 0 \leq Q^2 \leq s, \tag{2.5}$$
$$y = \frac{P\cdot q}{P\cdot k},\ 0 \leq y \leq 1, \tag{2.6}$$
$$x = \frac{-q^2}{2P\cdot q},\ 0 \leq x \leq 1, \tag{2.7}$$

where s is the total centre-of-mass energy squared and is given by the incoming lepton and proton beam energies. At HERA, with beam energies of $E_p = 920\,\text{GeV}$ and $E_e = 27.5\,\text{GeV}$ (cf. Chapter 4), the value of centre-of-mass energy, $\sqrt{s}$, is 318 GeV. Q^2 is the negative square of the four-momentum transfer of the exchanged boson which defines its virtuality. In the proton rest frame the inelasticity, y, describes the relative energy transfer from the lepton to the hadronic system and x is called Bjorken scaling variable. In the quark parton model (cf. Section 2.3.2) x can be interpreted as the fraction of the proton momentum carried by the struck quark. If all masses are neglected, Equations 2.4 and 2.6 can be simplified to:

$$s = (k+P)^2 \simeq 2k\cdot P, \tag{2.8}$$
$$y = \frac{P\cdot q}{P\cdot k} \simeq \frac{2P\cdot q}{s}. \tag{2.9}$$
$$\tag{2.10}$$

and all the four quantities can be related by:

$$Q^2 = sxy. \tag{2.11}$$

Since the centre-of-mass energy is fixed at HERA, the event kinematics is completely described by the knowledge of two additional variables. Typically the (x,Q^2) or the (x,y)

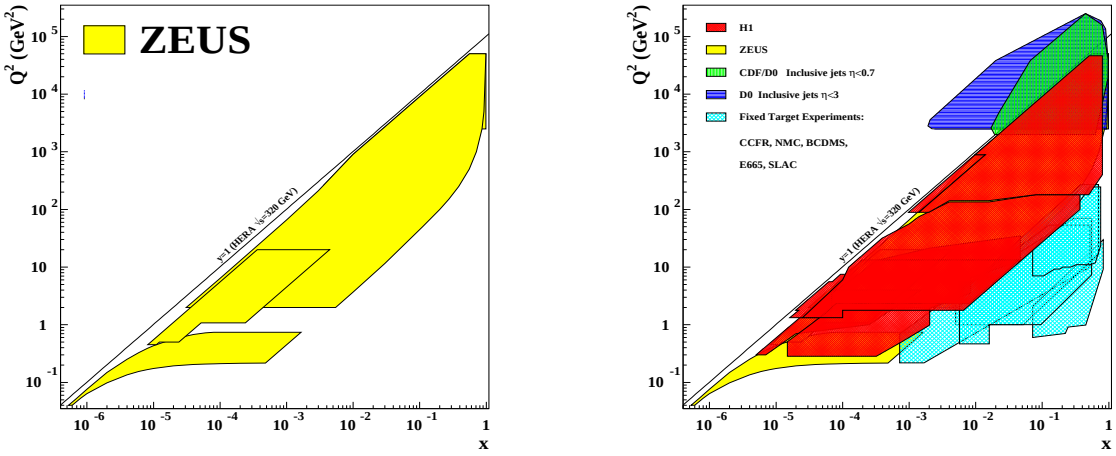


Figure 2.2: x–Q^2 plane covered by ZEUS (left) in comparison with H1, CDF, D0 and various fixed target experiments (right).

couple is chosen. The (x,Q^2) coverage of the HERA experiments is depicted in Figure 2.2. The variables can also be combined into the effective centre-of-mass energy, W, of the $\gamma(Z,W^{\pm})p$ system as:

$$W^2 = (P + q^2) = (P^2 + q^2 + 2P \cdot q) \sim sy - Q^2. \tag{2.12}$$

The virtuality, Q^2, gives the scale of the interaction. It can be thought as the "spatial" resolution of HERA to resolve the inner structure of the proton. Quantitatively it is related to the quantum wavelength, λ, of the boson which is looking into the proton, via $\lambda = h/\sqrt{Q^2}$. At HERA, the Q^2 range of up to $\sim 40000\,\mathrm{GeV}^2$ is equivalent to a resolution of $\lambda \sim 10^{-18}\,\mathrm{m}$ ($\sim \frac{1}{1000}$ of the proton radius). Since the photon exchange cross section falls rapidly as a function of Q^2, at low Q^2 the photon exchange dominates the cross section over the weak bosons. Only when Q^2 is sufficiently large, are the contributions from the Z and $W^{\pm}$ bosons significant. Indeed, when $Q^2 \geq M^2_{Z,W}$, the neutral and charged current cross sections are found to be of comparable size. The convergence of the NC and CC cross sections as seen in Figure 2.3 is an illustration of unification in the electroweak theory.

The variable Q^2 is used to split the kinematic plane into two separate regimes, the deep inelastic scattering (DIS) and the photoproduction (PhP) regime. Events with high photon virtuality, $Q^2 \geq 1\,\mathrm{GeV}^2$, and large hadronic centre-of-mass energy, W, are referred to as DIS events. For these events the incoming lepton is deflected by some measurable angle and can be identified in the detector. The events with very low photon virtuality $Q^2 \approx 0$ are characterised by the exchange of a quasi-real photon and are known as PhP events. In these events the incoming lepton is deflected at a very small angle and can not be observed in the detector. In DIS events the hard scale is provided by Q^2 and thus allows the perturbative calculations to be performed. For PhP events Q^2 is not a hard scale;

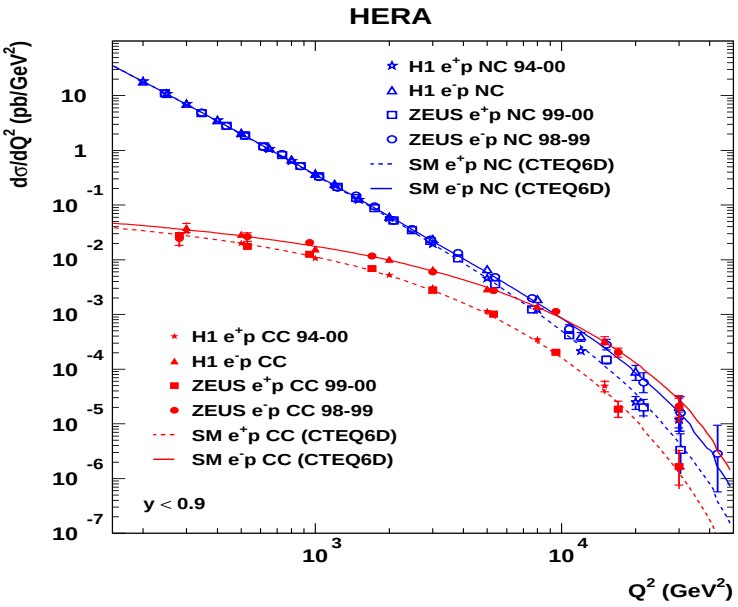


Figure 2.3: Inclusive differential NC and CC cross sections as a function of Q^2 [12]. Data points are measurements from the ZEUS and H1 experiments, the lines show the SM predictions for the specific process.

however the transverse momentum of jets or the mass of the heavy quarks produced may define a hard scale in the event. The measurements presented in this thesis are performed within the DIS regime which are therefore explained in more detail in the following section.

2.3 Deep Inelastic Scattering

As described in the last section, lepton-proton scattering with a large momentum transfer, Q^2, is called deep inelastic scattering. The term *deep* refers to $Q^2 \gg M_p^2$, while *inelastic* means $W^2 > M_p^2$, M_p being the mass of proton. This thesis focuses on heavy flavour production in *neutral current deep inelastic scattering* (NC DIS) and therefore the properties of this particular case will be discussed further.

2.3.1 Neutral Current DIS Cross Section

The general form of the lepton-proton scattering cross section can be written as a convolution of a leptonic part and a hadronic part:

$$d\sigma \sim L_{\mu\nu} W^{\mu\nu}, \tag{2.13}$$

where $L_{\mu\nu}$ and $W^{\mu\nu}$ are the leptonic and hadronic tensors, respectively [4]. $L_{\mu\nu}$ is calculable in *Quantum Electrodynamics* (QED) and for a purely electromagnetic exchange is given by:

$$L_{\mu\nu} = 2(k'_\mu k_\nu + k'_\nu k_\mu - (k' \cdot k) g_{\mu\nu}). \tag{2.14}$$

$W^{\mu\nu}$ parametrises the hadronic current and may be written as:

$$W^{\mu\nu} = W_1 \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2} \right) + \frac{W_2}{M_p^2} \left(P^\mu - \frac{P \cdot q}{q^2} q^\mu \right) \left(P^\nu - \frac{P \cdot q}{q^2} q^\nu \right), \tag{2.15}$$

where M_p is the proton mass and $W_{1,2}$ are scalar functions of x and Q^2 reflecting the proton structure. This form of the hadronic current with only two independent scalar functions is only valid for a parity-conserving interaction.
The functions W_1 and W_2 can be redefined as the structure functions F_1 and F_2 which depend on two Lorentz-invariant kinematic variables of x and Q^2:

$$F_1(x,Q^2) = M_p W_1(x,Q^2), \tag{2.16}$$

$$F_2(x,Q^2) = \frac{P \cdot q}{M_p} W_2(x,Q^2). \tag{2.17}$$

Therefore the NC DIS cross section for pure photon exchange can be written as:

$$\frac{d^2\sigma^{ep}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4} \left[\frac{y^2}{2} 2xF_1(x,Q^2) + (1-y) F_2(x,Q^2) \right], \tag{2.18}$$

ignoring mass terms. The structure function F_1 is proportional to the transverse component of the cross section, whereas the relation $F_L = F_2 - 2xF_1$, gives the longitudinal part of the cross section. Rewriting Equation 2.18 yields:

$$\frac{d^2\sigma^{ep}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4} \left[Y_+ F_2(x,Q^2) - y^2 F_L(x,Q^2) \right], \tag{2.19}$$

where

$$Y_+ = 1 + (1-y)^2. \tag{2.20}$$

2.3.2 Quark Parton Model

The *quark parton model* (QPM) [13, 14] is a simple model which is useful to understand many aspects of deep inelastic scattering. In this model, the proton is assumed to be composed of free (non-interacting) point-like spin$-\frac{1}{2}$ objects called partons. The inelastic scattering of an electron off a proton is then seen as the elastic scattering of an electron off a parton within the proton, as shown in Figure 2.4. The total ep cross section is given by the incoherent sum of the electron-parton scattering processes. In the infinite momentum frame of the proton, $|P| \gg M_p$, the partons can be assumed to have zero

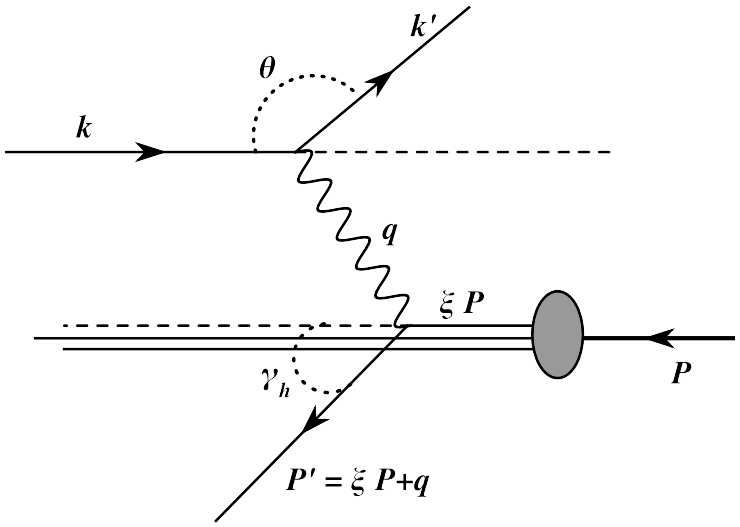


Figure 2.4: Schematic representation of the deep inelastic scattering. The photon with momentum q interacts with the parton which carries a fraction, ξ, of the proton momentum, P.

transverse momentum. Therefore neglecting the proton and parton masses and requiring four-momentum conservation the following relation can be obtained:

$$0 \approx m^2 = (\xi P + q)^2 = \xi^2 P^2 - Q^2 + 2\xi P \cdot q, \tag{2.21}$$

$$\Rightarrow \xi = \frac{Q^2}{2P \cdot q} = x, \tag{2.22}$$

where ξ is the fraction of the proton momentum carried by the struck parton and P, q and x are previously defined variables (cf. Section 2.2). Thus in the infinite momentum frame x can be interpreted as the longitudinal momentum fraction of the proton, ξ, carried by the interacting parton.
To embed partons into the DIS cross section, the probability $f_i(x)$ of finding a parton with a certain momentum inside the proton must be considered along with the cross section for the elastic scattering of the electron off a proton with that momentum. The differential cross section of such process $eq \to eq$ (see Figure 2.4) exchanging a virtual photon, γ^*, for a parton with momentum, $P = \xi P$, is given by:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2s} e_i^2 \frac{4 + (1 + \cos\theta)^2}{(1 - \cos\theta)^2}, \tag{2.23}$$

where e_i is the charge of parton, i, θ is the scattering angle of the electron in the eq centre-of-mass frame and $d\Omega = d\cos\theta d\phi$. If Q^2 and y are expressed in terms of the scattering angle, θ, and scattering energy, E_e, one gets:

$$Q^2 = 2E_e^2(1 + \cos\theta), \tag{2.24}$$

$$y = \sin^2\frac{\theta}{2}. \tag{2.25}$$

Using these two relations in Equation 2.23 gives:

$$\frac{d^2\sigma^{eq\to eq}}{dxdQ^2} = \frac{2\pi\alpha^2}{Q^4} e_i^2 \left[1 + (1-y)^2\right], \tag{2.26}$$

and taking into account such contributions for each parton:

$$\frac{d^2\sigma^{ep\to eq}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4} \sum_i e_i^2 \left[1 + (1-y)^2\right] (xf_i(x)), \tag{2.27}$$

where $f_i(x)$ is the probability of the parton, i, to have momentum fraction, x, of the proton and is also known as parton distribution function (PDF). The sum runs over all partons inside the proton. Comparison of Equation 2.19 and 2.27 yields:

$$F_L = 0 \qquad \text{and} \qquad F_2(x,Q^2) = \sum_i e_i^2 x f_i(x). \tag{2.28}$$

This means that F_2 depends only on x and has no Q^2 dependence. It is called Bjorken scaling [15] and its observation at SLAC [16] was a convincing piece of evidence that the proton contained partons and that those partons were in fact quarks. Considering Equation 2.28 and $F_L = F_2 - 2xF_1$ one also gets:

$$F_2(x) = 2xF_1(x). \tag{2.29}$$

which is known as Callan Gross relation [17], valid for spin$-\frac{1}{2}$ partons.
The structure functions are not yet calculable from first principles but have to be measured in experiments. As mentioned above, the QPM predicts the independence of the structure function from the virtuality of the photon, Q^2, at high energies and that $F_{1,2}$ depends only on the scaling variable, x. Later violation of this prediction was observed (see Figure 2.10). Also if the QPM is correct and the only partons in the proton are indeed quarks, then the sum of their fractional momenta should be equal to unity:

$$\sum_i \int_0^1 dx f_i(x) x = 1. \tag{2.30}$$

However, experimentally this value was measured to be approximately 0.5 [18], i.e. only half of the momentum of the proton is carried by the valence quarks. In the QPM, the proton consists of three (valence) quarks (uud). The answers to these questions were later explained by Quantum Chromodynamics (see Section 2.3.3). Taking into account the QCD description of the strong force between quarks, the simple proton picture of the QPM has to be modified and will be explained in Section 2.3.6.

2.3.3 Perturbative Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory to describe strong interactions through the exchange of gluons. The gluons and quarks couple via colour charges. Since the quarks

exchange gluons, they cannot be treated as free particles. Perturbative QCD (pQCD) offers a method to calculate cross sections as power series in the coupling strength of the strong interaction, α_s. The 0th order are QPM-like processes. Higher orders include gluon and quark loops called virtual corrections (see Figures 2.5 (a) and (b)). To calculate cross sections, integration over the full phase space of virtual and real quarks and gluons is needed. But the presence of the virtual particle loops which have infinite momenta cause divergences in the integration. These divergences are known as *virtual* divergences. Another type of divergences are the so-called *infrared* (IR) divergences which arise from the soft or collinear emission of massless gluons (see Figure 2.5 (c)).

In order to get rid of the divergences a procedure called renormalisation is applied.

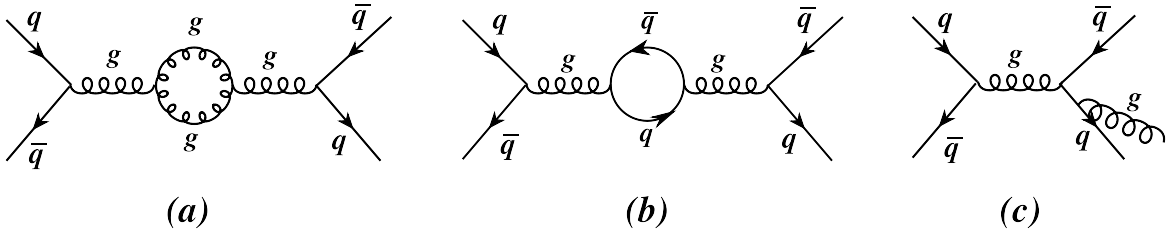


Figure 2.5: Examples of divergences present at NLO: (a) gluon loop, (b) quark loop, (c) emission of a collinear gluon.

This procedure replaces the divergent integrals by finite expressions. This requires the introduction of a new dimensional parameter called the *renormalisation* scale, μ_R. All renormalised quantities in the theory, such as the strong coupling, α_s, will depend explicitly on μ_R. This dependence of the strength of the interaction on the chosen scale is translated into a dependence of α_s on the energy ("running" coupling, $\alpha_s(Q)$), which is given by the so-called *renormalisation group* equation:

$$\mu^2 \frac{d\alpha_s}{d\mu^2} = \beta(\alpha_s), \tag{2.31}$$

where the $\beta-$function is a perturbative expansion in α_s, covering the dependency of α_s on the scale, μ^2:

$$\beta(\alpha_s) = \frac{-\beta_0}{4\pi}\alpha_s^2 - \frac{\beta_1}{8\pi^2}\alpha_s^3 - \ldots, \tag{2.32}$$

with

$$\beta_0 = 11 - \frac{2}{3}n_f \tag{2.33}$$

$$\beta_1 = 51 - \frac{19}{3}n_f \tag{2.34}$$

$$\vdots \quad \vdots \quad \vdots \tag{2.35}$$

At *leading order* (LO) the solution to Equation 2.31 is given by:

$$\alpha_s(\alpha_s^0,\mu) = \frac{4\pi}{\beta_0 \ln(\mu^2/\Lambda^2_{QCD})}, \tag{2.36}$$

Λ_{QCD} is the scale at which α_s becomes strong as μ^2 is decreased. It means that the Λ_{QCD} is the scale that sets the limit on the applicability of pQCD. The value of Λ_{QCD} at leading order was found to be ≈ 200 MeV. Figure 2.6 (right) shows the dependence of α_s on the

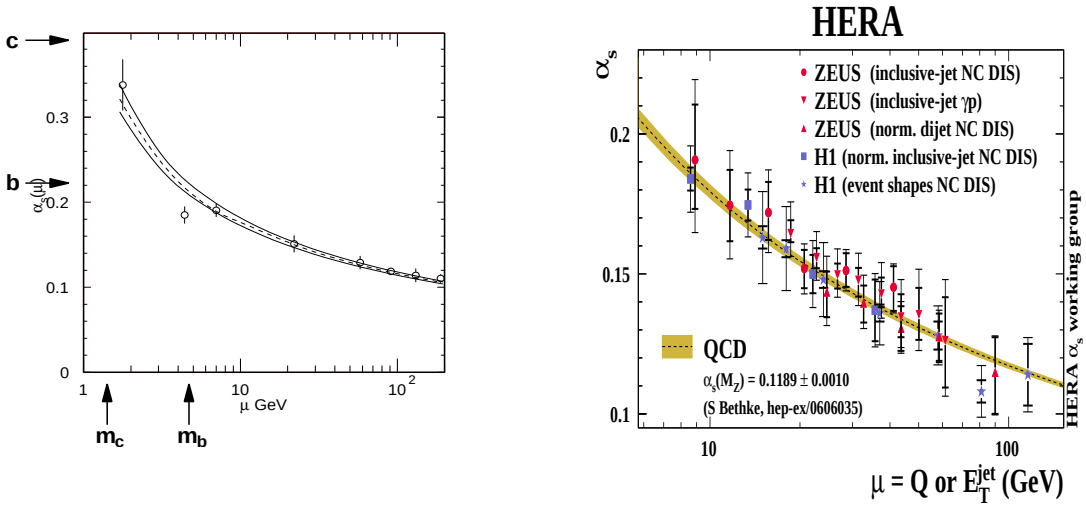


Figure 2.6: Running of the strong coupling constant, α_s, with the renormalisation scale, μ_R, including the $\pm 1\sigma$ limits denoted by the solid lines (left [19]). Running of α_s with the photon virtuality or the jet energy as the scaling parameter, μ. The data measured at HERA are compared with QCD prediction (shaded band) [20].

scale, μ. At large scale, μ, the value of α_s is small and the quarks are quasi-free. In this region of *asymptotic freedom*, pQCD is applicable. The masses of the heavy quarks m_b and m_c, are sufficiently large to provide a hard scale (see Figure 2.6 (left)). The rise of α_s for small values of μ corresponds to soft interactions and large distances which results in confinement. In this region the perturbative calculations become invalid.

2.3.4 Factorisation

Factorisation consists of the splitting of the $ep-$scattering process into two parts. One part is the interaction of high energy particles (hard sub-process) which can be described within pQCD and the second part is the long range part of the low energy processes (soft sub-process) which is not covered by pQCD. According to the factorisation theorem [21], the structure function, F_2, can be expressed as the convolution of a perturbative part, the coefficient functions (Wilson coefficients), C_2^i, and the non-perturbative parton density function (PDFs), $f_i(\xi,\mu_f,\mu)$:

$$F_2 = \sum_{i=parton} \int_x^1 C_2^i \left(\frac{x}{\xi}, \frac{Q^2}{\mu^2}, \frac{\mu_f^2}{\mu^2}, \alpha_s(\mu) \right) f_i(\xi,\mu_f,\mu) d\xi, \tag{2.37}$$

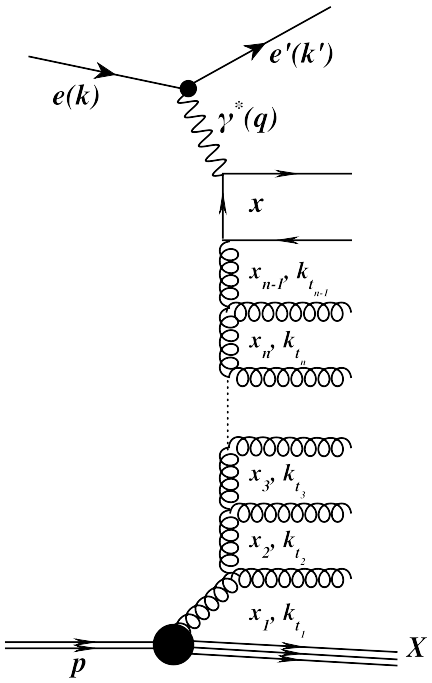


Figure 2.7: Diagram of the k_t ladder. A quark from the proton interacts with a virtual photon from the electron after radiating n gluons. Each gluon is characterised by a longitudinal momentum fraction x_i and transverse momentum k_{t_i}.

where μ is the QCD renormalisation scale and μ_f is known as factorisation scale and is an additional scale introduced by the factorisation theorem. μ_f defines the boundary between the perturbative and non-perturbative regime. The coefficient functions, C_2^i, are calculable while the PDFs can not be calculated from first principles and have to be extracted from experimental structure function data using the parton evolution models which describe the evolution of the quark and gluon momentum distribution in Q^2.

The leading order predictions used for the analysis presented in this thesis are based on the Dokshitzer Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equation [22, 23, 24, 25].

2.3.5 QCD Evolution Equations

The DGLAP evolution equations describe the way in which the quark and gluon momentum distributions inside a hadron evolve in Q^2. They have the form:

$$\frac{df_{qi}(x,Q^2)}{d\ln(Q^2)} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left[P_{qq}(x/\xi)f_{qi}(\xi,Q^2) + P_{qg}(x/\xi)f_g(\xi,Q^2)\right], \quad (2.38)$$

$$\frac{df_g(x,Q^2)}{d\ln(Q^2)} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left[P_{gq}(x/\xi)f_{qi}(\xi,Q^2) + P_{gg}(x/\xi)f_g(\xi,Q^2)\right], \quad (2.39)$$

where $f_{qi}(x,Q^2)$ is the quark density function for each quark flavour, i, and $f_g(x,Q^2)$ is the gluon density function. The functions $P_{ba}(x/\xi)$ are the DGLAP splitting functions. They describe the probability of a parton, a, with momentum fraction, ξ, to emit a parton, b, of momentum fraction, x, in the interval $Q^2 \to Q^2 + d\ln Q^2$.

The DGLAP equations are formally derived in the Leading Logarithm Approximation (LLA), where the terms of $(\alpha_s(\ln Q^2))^n$ are summed up to all orders. These $(\alpha_s(\ln Q^2))^n$ terms correspond to the ladder diagram with n gluon emission as shown in Figure 2.7. The LLA approximation is that the emissions are strongly ordered by transverse momentum of the gluon, k_T, as:

$$Q^2 \gg k^2_{T_n} \gg \cdots \gg k^2_{T_2} \gg k^2_{T_1}. \tag{2.40}$$

The approximation is valid for the large Q^2 region and not too small x region, where $\alpha_s(Q^2) \ln \frac{1}{x} \ll \alpha_s(Q^2) \ln Q^2$. There are other approaches to describe the low x physics.

At moderately small x, the Double Leading Logarithm Approximation (DLLA) can be used. In the DLLA, the terms of $\left(\alpha_s \ln \frac{1}{x} \ln Q^2\right)^n$ are summed up. To describe even smaller x, the Balitskii-Fadin-Kuraev-Lipatov (BFKL) [28, 29, 30] equation was developed. In the BFKL equation, the ordering by k_T is no longer needed but a strong ordering by x, the momentum fraction of emitted gluon is required:

$$x_1 \gg x_2 \gg \cdots \gg x. \tag{2.41}$$

to sum up the terms of $\left(\alpha_s \ln(\frac{1}{x})\right)^n$. The BFKL equation predicts a steeper increase of F_2, hence gluons, at low x than the DGLAP equation. The different approaches are summarised in Figure 2.8.

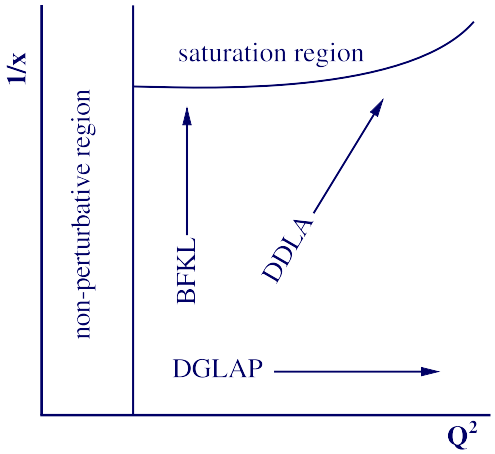


Figure 2.8: Different approaches of QCD evolution equations displayed on Q^2–$\frac{1}{x}$ plane [33].

At very low Q^2, perturbation can not work due to the large value of $\alpha_s(Q^2)$. The high density of the gluon distribution at very low x may saturate due to recombination or annihilation of gluons. The attempt of a unification of the DGLAP evolution, based on Q^2 ordering, and the BFKL evolution, based on ordering in x, led to the development of the Ciafalani-Catani-Fiorani-Marchesin (CCFM) evolution equations [31, 32].

2.3.6 QCD Improved Quark Parton Model

Taking into account the QCD description of the strong forces between quarks, the simple proton picture of the QPM has to be modified. In detail, the representation of the proton by freely moving quarks quantified by their PDFs is too simple and the internal dynamic structure of the proton has to be taken into account: the proton description has to include the gluons in the proton. In addition to the QPM contribution (Figure 2.9 (a)), diagrams

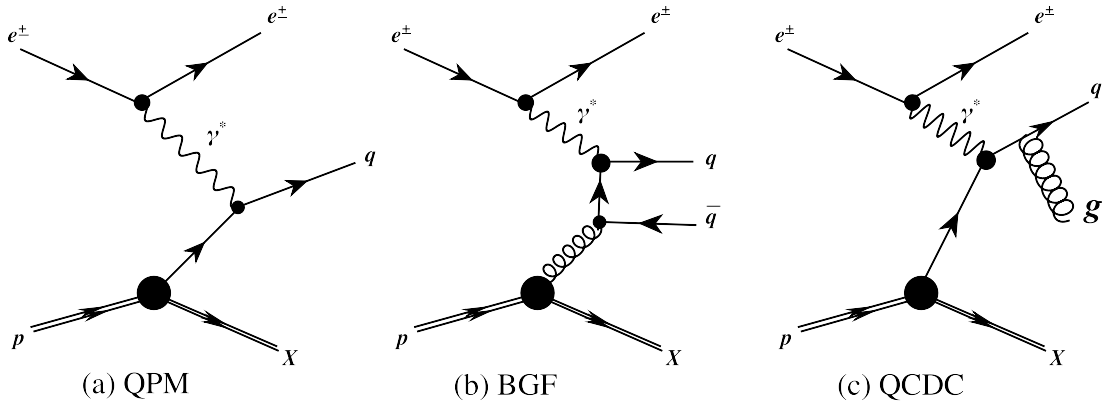


Figure 2.9: Feynman diagram of an electron-proton interaction in the quark parton model (QPM) (a). Leading order diagrams for boson-gluon fusion (BGF) (b) and QCD Compton scattering (c) where the quark radiates a gluon after the interaction.

with a gluon from the proton interacting with the electron have to be included. One of the leading-order contributions in α_s is given by the process $\gamma^* g \to q\bar{q}$. This process is called *boson-gluon fusion* (BGF) (see Figure 2.9 (b)). This graph can be described as the fluctuation of a gluon from the partonic substructure of the proton into a quark/antiquark pair. One of the quarks interacts electromagnetically with the electron by the exchange of a photon. An additional $\mathcal{O}(\alpha_s)$ correction is given by the process in which a gluon is emitted from the quark line before or after the scattering with the photon ($\gamma^* \to qg$). This process is called QCD Compton scattering (QCDC) (see Figure 2.9 (c)). The inclusion of QCD corrections to the QPM implies that scaling of structure function $F_{1,2}$ is violated.

The proton structure function F_2 as a function of Q^2 for different values of Bjorken scaling variable, x is depicted in Figure 2.10. For $x \sim 0.1$ the structure function is independent on Q^2 as predicted by QPM. For $x > 0.1$ it decreases with Q^2, i.e., there is less probability to find a quark due to gluon radiation. For $x \ll 0.1$ the structure function increases with Q^2. This indicates that at higher Q^2 the proton momentum is shared by more partons than at lower Q^2 mainly due to the contribution by the gluons in the proton splitting into $q\bar{q}$ pairs.

According to the factorisation theorem of QCD the PDFs are process independent, thus predictions can be made for processes other than the ones used for their determination. Figure 2.11 shows the dependence of PDFs on x, determined by the ZEUS and H1 collaborations from DIS data for $Q^2 = 10\,\mathrm{GeV}^2$.

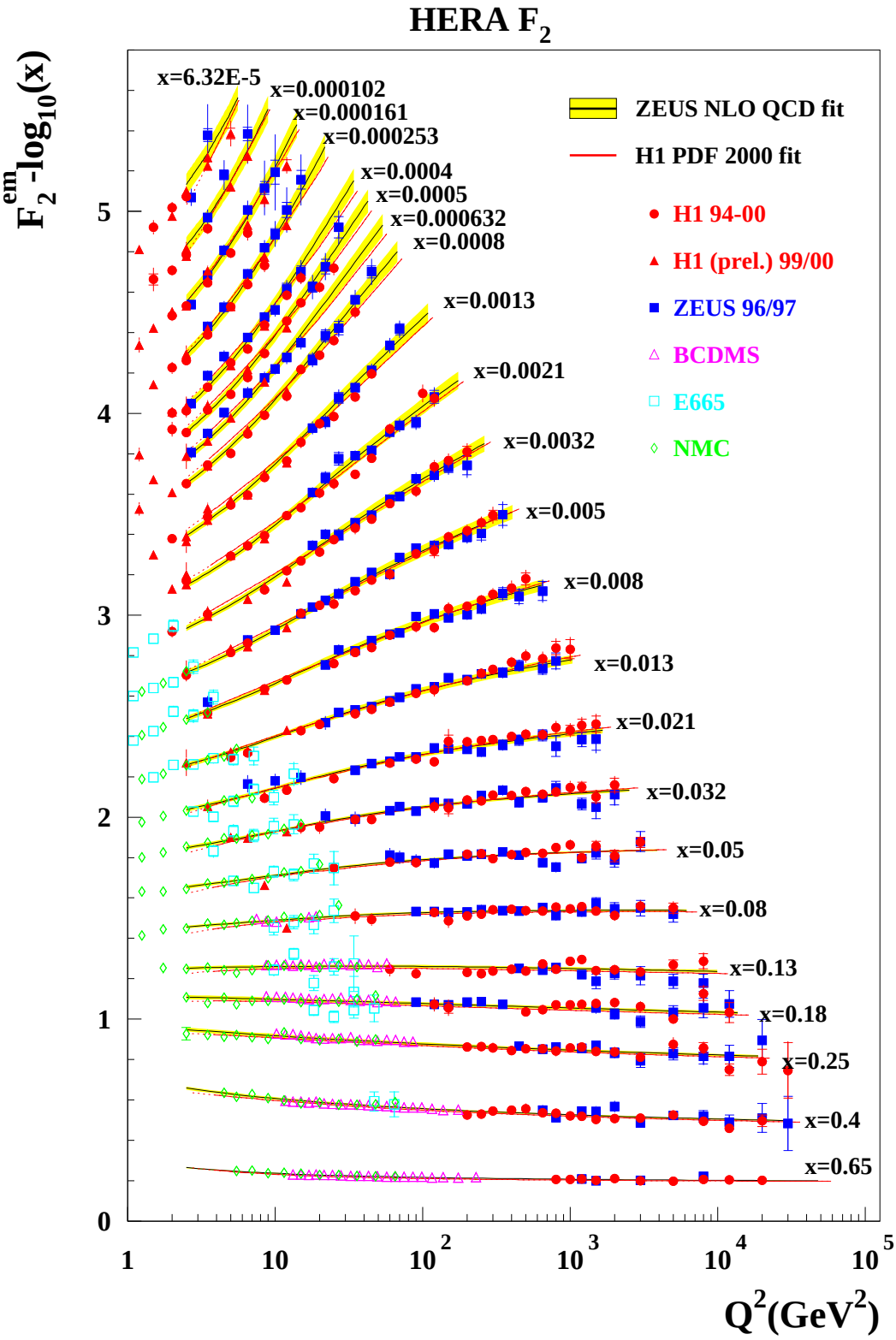


Figure 2.10: The proton structure function, F_2, as a function of the virtuality of the exchanged photon at different values of x measured at HERA and fixed target experiments.

It can be observed that the valence quarks u and d are dominant at high x. The gluon and sea quark PDFs are scaled by 0.05 and dominate at low x. The PDFs on the left were determined from ZEUS data and a corresponding fit alone [26], while the ones on right show the ZEUS-H1 combination [27]. Combining the ZEUS and H1 data considerably reduced the uncertainties.

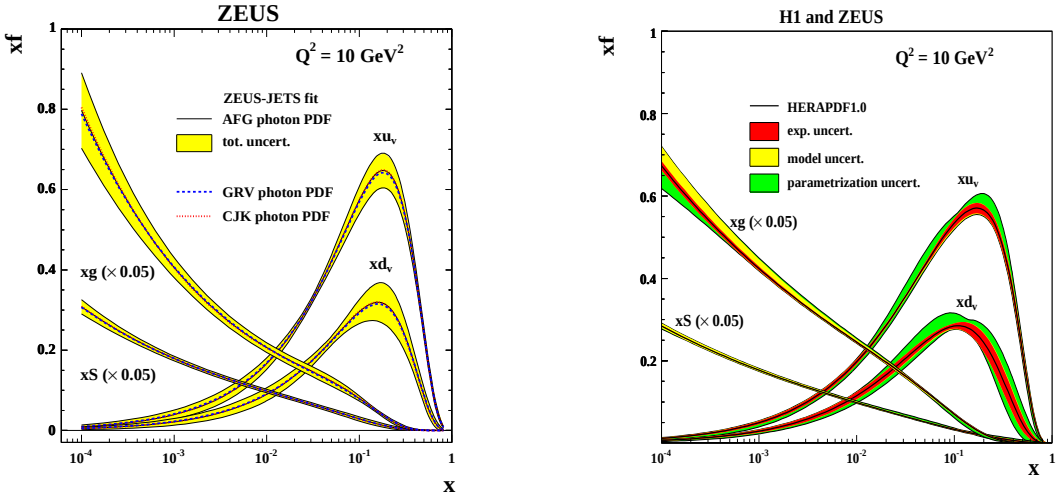


Figure 2.11: Parton density functions of the valance u and d quark, the gluons g and the sea S quark content of the proton, determined from a ZEUS next-to-leading order QCD fit (left) and from a combined H1 and ZEUS fit (right).

2.4 Heavy Quark Production in DIS

At HERA, the lowest order process of heavy quark production is the boson-gluon fusion process (BGF) (Figure 2.9 (b)). In such a process, a heavy $q\bar{q}$ pair can be produced if the squared centre-of-mass energy of the $\gamma^* g$ system, $\hat{s}$, is:

$$\hat{s} = (\gamma^* + g)^2 = (q + \xi P)^2 > (2m_q)^2, \tag{2.42}$$

where m_q is the mass of the heavy quark, and q and $g = \xi P$ are the photon and gluon four-momenta, respectively. The high quark mass sets a hard scale for the process and a reliable description by pQCD calculations should be possible, due to the low value of the running coupling constant, α_s, at a scale corresponding to heavy flavour mass (see Figure 2.6). Thus the heavy flavour production is an excellent test of pQCD. Furthermore, due to the dominant photon-gluon production process, heavy flavour measurements provide insights into the gluon contribution in the proton. Two kinematic regions can be distinguished for heavy quark production. In the regime $Q^2 \leq (2m_q)^2$, BGF is the lowest order production

process of a quark-antiquark pair of mass $2m_q$. For high $Q^2 \gg (2m_q)^2$ the splitting of a gluon into a virtual $q\bar{q}$ can be reinterpreted to occur inside the proton and the QCD improved QPM picture is applicable for the production mechanism (cf. Section2.3.6).

2.4.1 BGF Cross Section

The leading order cross section for the production of a heavy $q\bar{q}$ pair in BGF can be calculated as [34]:

$$\hat{\sigma}_{BGF} = \frac{\pi e_q^2 \alpha \alpha_s}{\hat{s}} \left[(2 + 2\omega - \omega^2) \ln \frac{1+\chi}{1-\chi} - 2\chi(1+\chi) \right], \tag{2.43}$$

where e_q denotes the electromagnetic charge of the heavy quark and the variables ω and χ are defined as:

$$\omega = 4m_q^2/\hat{s}, \qquad \chi = \sqrt{1-\omega}. \tag{2.44}$$

and $\hat{s}$ is the squared centre-of-mass energy of the $q\bar{q}$ pair. Due to the lower charge of the beauty quark and its higher mass, the factor $\frac{e_q^2}{\hat{s}}$ is smaller for beauty than for charm leading to a strong suppression of the beauty quark cross section with respect to the charm quark cross section. At the energy of the HERA collider, the beauty quark is mainly produced near the mass threshold. In this kinematic region the cross section of charm quark production is about two orders of magnitude larger than the beauty cross section.

2.5 Parton Hadronisation

As a consequence of the colour confinement in QCD, the heavy quarks (coloured) produced in the hard interaction can not be observed directly in the experiment. They form colourless bound states called hadrons which are formed from the original quark pair. This process can be described as a series of different steps. In a first perturbative step additional partons can be emitted from an initial-state parton or from the heavy quark. This is known as parton showering, which is followed by a second non-perturbative step using phenomenological models to form hadrons from these partons and is referred to as hadronisation. In the following sections both of these stages are explained.

2.5.1 Parton Showers

In the parton showering process, the partons from the partonic structure of the incoming particles and the partons originating from the hard sub-process can radiate gluons ($q \to qg$, etc.) or split into $q\bar{q}$ pairs ($g \to q\bar{q}$), also called initial- and final-state radiation. Each daughter produced in the above process may emit additional partons. The model of parton showers thus approximates multiple parton emission by a series of successive parton

splittings and contributes to higher order corrections not taken into account by the LO hard sub-processes.

Initial- and final-state parton showers are treated in different ways. Both are started at a scale, $Q^2_{\max}$, which defines the transition between the initial- and final-state radiation and the hard sub-process. The initial-state parton shower algorithm starts from a parton in the proton and models the radiation of space-like parton showers until the parton reaches the hard sub-process having a scale $Q^2_{\max}$. The simulation of this process is carried out in the backward evolution scheme starting from the scale $Q^2_{\max}$ and then tracing the showers backward in time in a sequence of decreasing Q^2 down to the point where a cut off $Q^2_o \approx 1\ \mathrm{GeV}^2$ is reached. This process is based on the usage of DGLAP evolution equations. The final-state radiation splits the final-state partons into time-like showers starting at $Q^2_{\max}$. The evolution continues until it reaches the cut-off value of Q^2_o. Branching on both sides are interleaved in a common sequence of decreasing Q^2 values. At each Q^2 value the splitting function, $P_{ba}(\xi/x)$, of the DGLAP evolution equations describe the probability of splitting of a parent parton, a, with momentum fraction, ξ, to a parton, b, having momentum fraction, x.

2.5.2 Fragmentation/Hadronisation

The fragmentation or hadronisation of the heavy quark into a heavy-flavoured hadron is described by the factorisation theorem through the following equation [35]:

$$\frac{d^3\sigma_h(k)}{d^3k} = \int D(z)\frac{d^3\sigma_q(\hat{k})}{d^3\hat{k}}\delta^3(\vec{k} - z\hat{\vec{k}})\ d^3\hat{k}\ dz, \tag{2.45}$$

where h is the heavy-flavoured hadron with momentum, k, $\hat{k}$ is the momentum of the quark and σ_q and σ_h are the parton-level and hadron-level cross sections, respectively. $D(z)$ is the non-perturbative fragmentation function which gives the probability that a parton fragments into a hadron with momentum fraction, z, of the original parton.
There are different hadronisation models which describe the formation of the hadrons from partons. Among them, the most common models used in the event generators are cluster fragmentation and string fragmentation.

- **Cluster Fragmentation:**
  The cluster model groups the partons into colour neutral objects (see Figure 2.12 (right)). The gluons are split into light (u or d) quark-antiquark or diquark-antidiquark pairs. The splitting into diquarks is suppressed with respect to that into quarks. These are subsequently clustered into colourless objects via colour connections generated in the parton shower. Then each of these clusters is fragmented into two hadrons or the lightest hadron of its flavour if the cluster is not massive enough

- **String Fragmentation:**
  In this model the $q\bar{q}$ pairs are connected by colour flux tubes, called strings. As the

q and $\bar{q}$ move apart from their common production vertex, the string is stretched between the colour charge and anticolour charge with transverse dimension of typical hadronic size ($\sim 1\,\text{fm}$). The string has a uniform energy per unit length, corresponding to a linear quark confining potential. Therefore the potential rises with increasing distance. The more q and $\bar{q}$ separate, the more potential energy is stored and the string can break up producing a new $q\bar{q}$ pair. This process (see Figure 2.12 (left)) continues as long as the invariant mass of the string pieces exceeds the on-shell mass of a hadron.

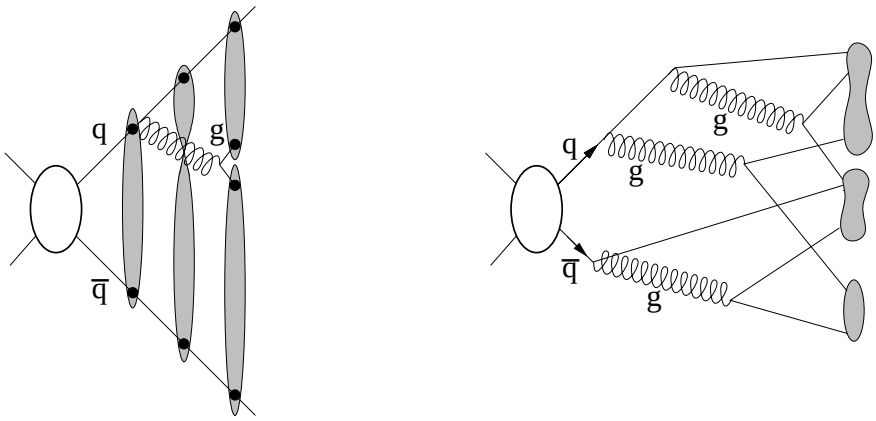


Figure 2.12: Parton fragmentation models: string fragmentation (left), cluster fragmentation (right).

The $q\bar{q}$ pairs are created according to the probability of a quantum mechanical tunnelling process $\left(e^{(-\frac{\pi m_{q,\perp}^2}{k})}\right)$ which depends on the transverse mass squared $m_{q,\perp}^2 = m_q^2 + p_{q,\perp}^2$ and the string tension, $k \approx 1\,\text{GeV/fm}$. Due to the dependence on the parton mass, the production of the heavy quark hadrons is suppressed in the hadronisation. The string fragmentation function for the light quarks has the form:

$$D(z) \sim \frac{1}{z}(1-z)^a \exp\left(-\frac{bm_h^2}{z}\right). \tag{2.46}$$

which describes the probability for producing a hadron, h, with a mass, m_h, and a momentum fraction, $z = (E + P_\parallel)_h/(E + P)_q$ where $P_\parallel$ is the momentum of the formed hadron, h, along the direction of the quark. a and b in Equation 2.46 are free parameters which have to be determined from measurements.
The heavy quark fragmentation functions are expected to peak near 1 because the produced hadron should retain a large fraction of the initial quark momentum. A commonly used fragmentation function for the heavy flavour production is *Peterson* fragmentation function [36] which has the form:

$$D(z) \propto \frac{1}{z}\left(1 - \frac{1}{z} - \frac{\epsilon}{1-z}\right)^{-2}, \tag{2.47}$$

where the parameter ϵ depends on the heavy flavour considered. For beauty quark production it has been determined to $\epsilon_b \approx 0.0035$ [37] but the exact value depends on the treatment of parton showering. The Lund string model used by RAPGAP MC combines the string fragmentation approach (cf. Section 2.5.2) for light flavour quarks with the Peterson fragmentation parametrisation for heavy flavours.

2.6 Beauty Hadrons and Decays

The hadronisation process of b quarks produced in the hard interaction results in the formation of beauty hadrons. The weakly decaying beauty hadrons that are relevant at HERA are $B^{\pm}$ and $B^0/\bar{B}^0$, followed by B_s^0 and the Λ_b^0 baryons. Some properties of these

Hadron	Quark content	Mass (MeV)	Lifetime (ps)
$B^{\pm}$	$u\bar{b}/\bar{u}b$	5279.15 ± 0.31	1.643 ± 0.010
$B^0/\bar{B}^0$	$d\bar{b}/\bar{d}b$	5279.53 ± 0.33	1.527 ± 0.008
$B_s^0/\bar{B}_s^0$	$s\bar{b}/\bar{s}b$	5366.30 ± 0.60	1.454 ± 0.040
$\Lambda_b^0/\bar{\Lambda}_b^0$	$udb/\bar{u}\bar{d}\bar{b}$	5620.20 ± 1.60	1.288 ± 0.065

Table 2.3: The quark content, mass and lifetime of the most frequently produced B-hadrons [19].

hadrons are listed in Table 2.3. Due to their short lifetime (~ 1 ps) beauty hadrons can not be observed directly. They can only be detected through their decay particles which reach the detector.

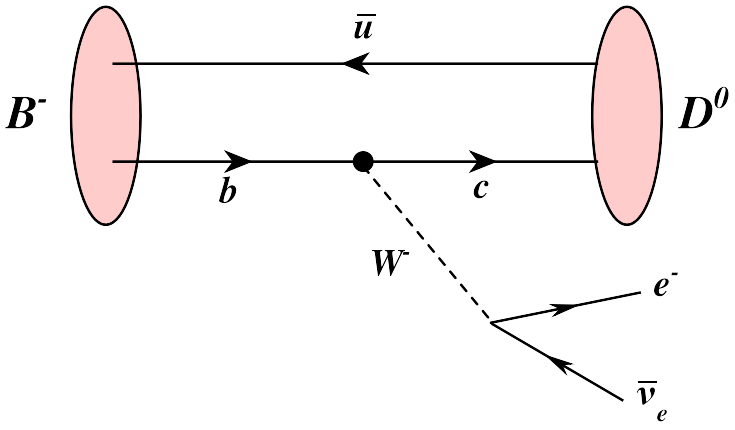


Figure 2.13: Semileptonic decay of B hadron; $B^- \to e^- \nu_e D^0$, in the spectator model.

The simplest model to describe the decay of beauty hadrons is given by the *spectator* model which is shown in Figure 2.13, where the b quark in the hadron decays weakly and the light

quark continues in the final state. This model is based on the assumption that the light quark in the heavy hadron has a negligible influence on the other quark and thus the QCD effects between the quarks such as binding effects and gluon radiation can be neglected. This assumption is justified by the large mass of b quark compared to the other light quark in the hadron which is mainly a u or d quark.

The b quark most likely decays into a virtual W boson and c quark, where the W boson decays leptonically or hadronically producing a pair of leptons, $l\nu$, or a pair of quarks, $q\bar{q}$, respectively. The decay channel considered in this analysis is the semileptonic decay into an electron, e and an electron antineutrino, $\bar{\nu_e}$ and its charge conjugate process ($\bar{b} \to e^+\nu_e$). The branching ratio of this process is about 10.9%. In addition to the direct electron

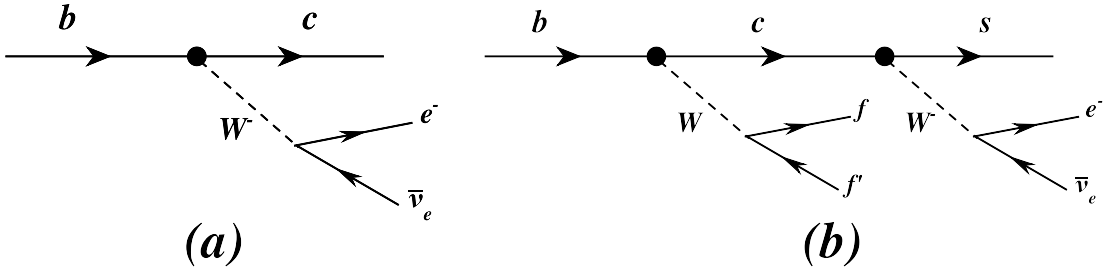


Figure 2.14: Beauty quark decay: (a) prompt decay, and (b) cascade decay via a charm quark.

production in beauty quark decay ($B \to eX$), electrons can be produced from beauty quark decays through cascade processes where the charm quark resulting from the beauty quarks further decays semileptonically into an electron (see Figure 2.14). About 9% of all b quarks produce an electron from a charm quark in the cascade decay ($B \to cX \to e\nu sX$). The contribution from other decay modes of beauty quarks, e.g., ($B \to J/\psi X \to eeX$) or ($B \to \tau X \to eX'$) are much smaller but will also be considered in this analysis.

2.7 Monte Carlo Simulation

Simulation programs based on Monte Carlo (MC) techniques have become an essential tool to understand the complexity of high energy physics processes and of particle detectors. MC programs are used to simulate the physics events, which through a comparison with real data provide a better understanding of the detector performance and the response of different detector components. They are also used to extract efficiency and acceptance corrections necessary for unfolding cross sections.
At HERA, the event simulation is done in two main steps. First the *ep*-scattering process is simulated using an "event generator" which provides the complete list of the four-momenta of the final-state particles. In the second step these events are passed through a simulation of the detector. In the following sections the general structure of an event generator,

the specific MC programs used in this analysis and the detector simulation program are described.

2.7.1 Event Generators

QCD MC event generators make use of the factorisation theorem [21]. Following the assumptions of the theorem, an ep scattering process can be divided into several different stages as shown in Figure 2.15. Although some of the processes of these stages have already been described in detail, here just for completeness, a brief outline of those is given too.

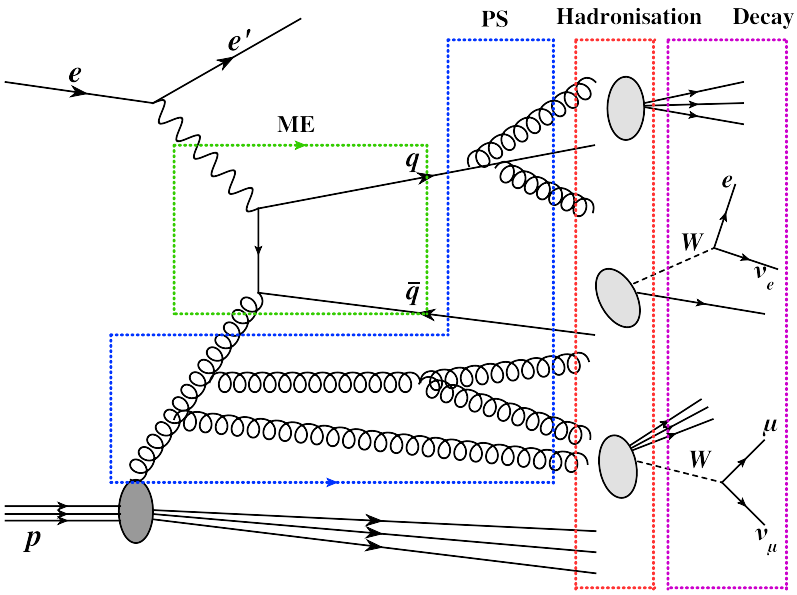


Figure 2.15: Heavy flavour production and evolution as modelled in ep event generators in a boson-gluon fusion process.

- **Hard sub-process:**
  This is the main part of the event simulation program and describes the interaction between incoming beam particles. In our case it is the interaction between a parton extracted from the proton and the exchanged photon. The flavour and momentum of the incoming parton are chosen according to the selected parton distributions (PDFs) and are used as an input to the calculation. This part of the process can be calculated in fixed order (FO) perturbative expansion because it involves a hard scale. The hard momentum transfer scale, μ, sets the boundary condition for the initial-state and final-state parton showers.

- **Initial-state and final-state radiation:**
  Partons (also leptons and photons) can emit other partons in $a \rightarrow bc$ processes, during the initial and final state (before and after the hard scatter) which can have

a strong influence on the topology of an event. These perturbative corrections which are not taken into account by the LO hard sub-process are modelled by the so-called *parton shower* method (cf. Section 2.5.1).

- **Hadronisation:**
  This is the process in which colourless hadrons are formed out of the coloured partons. It is a non-perturbative phenomenon which is not well understood and is described by phenomenological models. Some of the most commonly used hadronisation models were described in Section 2.5.2.

- **Particle decay:**
  In this process the unstable hadrons formed during the hadronisation process decay according to their branching fractions.

2.7.2 DIS MC

In this analysis $ep-$scattering events in the DIS regime were generated with the RAPGAP [38, 39] and DJANGOH [40] event generators. The RAPGAP 3.00 Monte Carlo program in the massive mode was used to generate beauty and charm events, while DJANGOH was used to simulate the light flavour events.

DJANGOH is an event generator which includes both QED and QCD radiative effects. It is an interface between the HERACLES [41, 42] and ARIADNE [43] MC programs. The HERACLES program takes into account a complete set of one-loop EW radiative corrections and radiative scattering. The hard scattering between the parton and photon is simulated according to the Standard Model cross sections and the proton PDFs. The parametrisation of the PDFs is chosen according to the CTEQ5D [44] set of proton PDFs. ARIADNE uses the colour dipole model (CDM) [45, 46, 47] and simulates QCD cascades. In the CDM, the struck quark in DIS is connected to the proton remnant via a colour dipole which can radiate a gluon, setting up two independent dipoles between the partons. Gluon radiation from the dipoles continues, creating more dipoles until the hadronisation scale is reached. The effects of initial-state radiation are described in terms of final-state gluon radiation from colour dipoles produced in the hard interaction. The output from the ARIADNE program serves as an input to the hadronisation process. The hadronisation is performed using the JETSET program [48] which applies the Lund string fragmentation model [49] (cf. Section 2.5.2).
RAPGAP also uses the HERACLES program to simulate the QED radiative effects from the incoming and scattered lepton. The first order QCD processes (BGF, QCDC) are simulated using the exact matrix elements. For higher order corrections QCD parton showers, based on the *leading log* DGLAP [22, 23, 24, 25] splitting functions are used. They can occur before and after the hard sub-process. The fragmentation and decay channels are done with PYTHIA [48] which uses the Lund string model.

2.7.3 Detector Simulation

All the event generators supported at ZEUS (like RAPGAP, DJANGOH, PYTHIA) are gathered in a software package called AMADEUS (ZEUS interface to MC Generator). AMADEUS converts the output of the event generator such as the four-momenta of the particles produced in the *ep* interaction and all the relevant kinematic variables to a suitable format (ADAMO format) for the following simulation of the ZEUS detector response and the trigger system.

For simulating the detector response the events are processed through MOZART (Monte Carlo for ZEUS Analysis Reconstruction and Triggering). MOZART is based on the GEANT [50] package which contains a description of all the detector components, including the material they are made of, their geometrical shapes and positions. The program traces the particles through the detector simulating its response and taking into account physics processes such as energy loss, multiple scattering and particle decays. The output of MOZART is then fed into the CZAR (Complete Zgana Analysis Routine) [51] package which simulates the trigger logic as implemented in data taking. In the next step, which is common for both MC simulated events and the ZEUS raw data, the offline reconstruction is performed using the reconstruction package ZEPHYR (ZEUS Physics Reconstruction). All the ZEUS data are organised using the ADAMO [52] (Aleph DAta MOdel) data management system used for the data storage in memory or on an external media (tape or disk) and for their documentation. Finally the user can access the data using EAZE or ORANGE [53] programs, where both MC simulated events and data events are analysed in the same way. A diagram for the ZEUS reconstruction chain of data and MC is shown in Figure 2.16.

2.8 Next-to-Leading Order Predictions

As stated above, in the heavy quark production process (mainly BGF) the large mass of the heavy quarks provides a hard scale so that calculations in pQCD are expected to be reliable. Perturbative QCD calculations of heavy quark production are available also in next-to-leading order (NLO). However in addition to the hard scale set up by the heavy quark mass the simultaneous presence of a competing hard scale, such as the transverse momentum (p_T) of the heavy quark or the virtuality of the exchanged photon (Q^2) causes some complications in theoretical calculations, due to terms in perturbative expansion which depend logarithmically on the ratio of these scales. Since the perturbative expansions can not be optimised for all scales at once, different approaches to the calculations have been developed assuming a single hard scale in each. In the following the different approaches restricted to DIS region are described.

- **The next-to-leading order (NLO) massive approach:**
  This approach assumes that there is no intrinsic heavy quark (beauty or charm) in

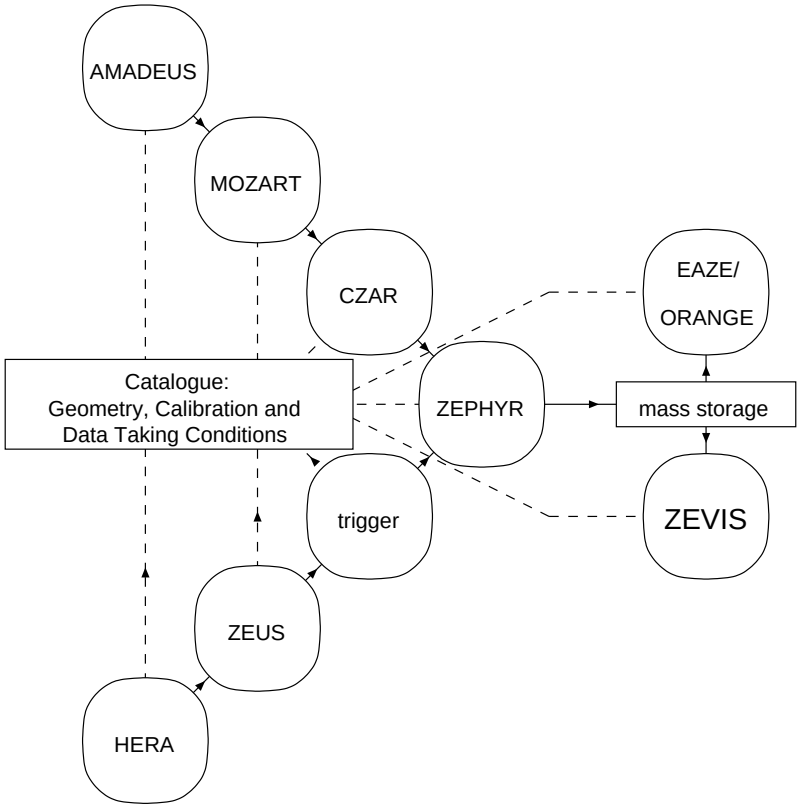


Figure 2.16: A schematic diagram of the ZEUS Data and Monte Carlo reconstruction chain; full lines with arrows indicate the data flow; dashed line indicate the communication with the catalogue.

the proton (or photon). The heavy quarks are only produced dynamically in the hard scattering. This approach is expected to work best when all relevant hard scales, e.g., the photon momentum transfer, Q^2, are of the order of heavy quark mass, (m_q). This scheme is also known as Fixed Flavour Number Scheme (FFNS).

- **The NLO massless approach:**
  For $Q^2 \gg m_q^2$, large $\log(Q^2/m_q)$ terms could in principle spoil the reliability of the predictions. In this case, it might be preferable to switch to the massless scheme, in which m_q is neglected kinematically. The potentially large logarithms can then be re-summed to all orders (next-to-leading log (NLL) summation). Since such an approach is obviously not applicable when $Q^2 \sim m_q^2$, schemes have been designed which make a continuous transition between the FO massive and NLL massless scheme. This is often referred to as the GM-VFNS (General Mass Variable Flavour Number Scheme) [54].

In this thesis, the measurements will be compared with NLO calculations in the massive scheme using the HVQDIS program which is described in the following section. The

NLO contributions to the cross sections are found to be significant. Feynman graphs of real corrections (emission of gluons) and virtual corrections (interference term with LO) contributing to the NLO cross sections are shown in Figure 2.17 and 2.18, respectively. Comparing to LO diagrams (BGF process 2.9 (b)), the main difference relies on additional radiation of hard gluons and the interference with virtual corrections.

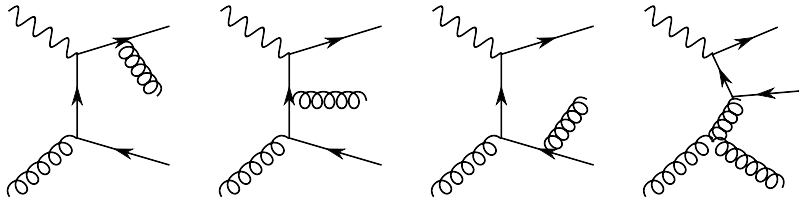


Figure 2.17: Real NLO QCD contribution to heavy quark production.

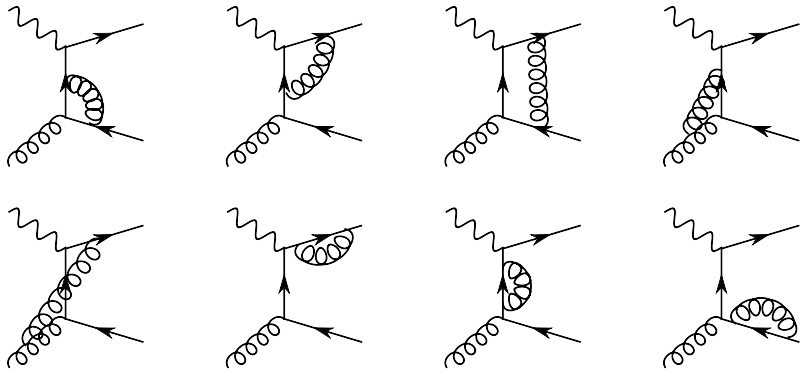


Figure 2.18: Virtual NLO QCD contributions to heavy quark production.

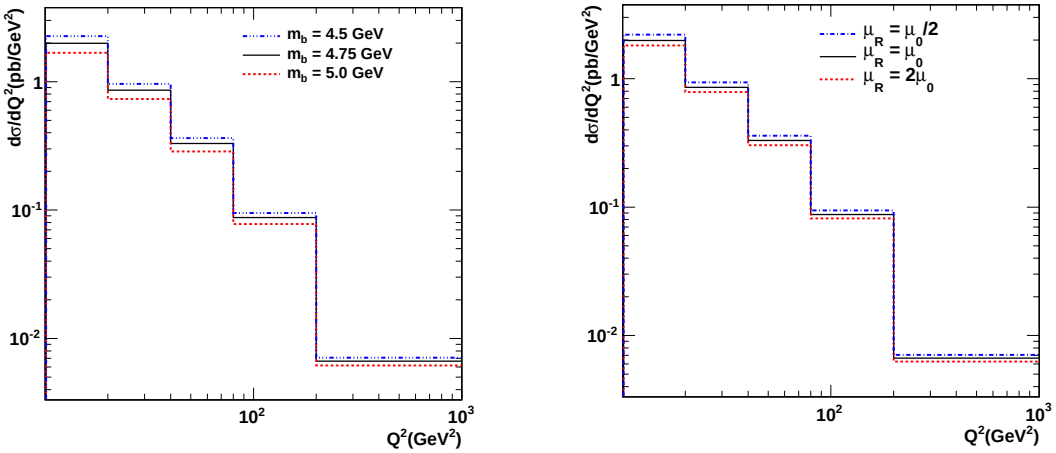
2.8.1 NLO Programme HVQDIS

The HVQDIS programme [55, 56] was used to extract the NLO QCD predictions. It is based on the FFNS (cf. Section2.8) in which only light flavours are present in the proton and heavy quarks are produced in the interaction. A number of input variables to the program must be specified to obtain predictions for the production of electrons from semileptonic decays at NLO using HVQDIS. The following set of parameters was used for the central values of the cross section prediction:

- the parton density functions were obtained by repeating the ZEUS-S [57] PDF fit in the FFNS with quark masses set to the same values as in the HVQDIS calculation;
- the beauty quark mass was set to $m_b = 4.75\,\text{GeV}$;
- the renormalisation, μ_R, and factorisation, μ_F, scales were chosen to be equal and set to $\mu_R = \mu_F = \sqrt{Q^2 + 4m_b^2}$;

- the Peterson fragmentation function [36], with $\epsilon_b = 0.0035$ was used to produce beauty hadrons from the heavy quarks;
- for beauty, both the contributions from prompt and from cascade decays, including $b \to \tau \to e$ and $b \to J/\psi \to e^+e^-$, were taken into account in the effective branching fraction, which was set to 0.217 [19].

The uncertainty on the theoretical predictions was evaluated by independently varying μ_R and μ_F by a factor two; by varying the beauty quark mass simultaneously within $m_b \in [4.5, 5.0]$ GeV in the calculation and in the PDF fit; by varying the fragmentation parameter, ϵ_b, within an uncertainty of 0.002 and by varying the proton PDFs within their experimental uncertainty. The total theoretical uncertainty was obtained by adding in quadrature the effects of each variation, where the dominant contribution to the uncertainty originates from the mass and μ_R scale variation. As an example differential cross sections as a function of Q^2 for different variations of mass and μ_R are displayed in Figure 2.19(a) and 2.19(b), respectively.



(a) Mass variation **(b)** μ_R variation

Figure 2.19: The variation of (a) mass, and (b) renormalisation scale in the HVQDIS calculation, which are the dominant contributions to the overall NLO uncertainty.

Chapter 3

Beauty Production in DIS at HERA

The study of beauty quark production is one of the main research topics at HERA II. Beauty quarks are predominantly produced via the boson-gluon fusion process (cf. Section 2.4) and provide an important tool to investigate our present understanding of the theory of Quantum Chromodynamics. The large mass of the b quark provides a hard scale which ensures that the cross sections are perturbatively calculable. The measured beauty cross sections can be used to extract the beauty contribution, $F_2^{b\bar{b}}$, to the inclusive proton structure function, F_2. In the following a brief review of the previous measurements of beauty production in the DIS regime that have been made at HERA is presented.

A short introduction to the beauty quark identification techniques is given in Section 3.1. The measurements of the beauty quark production cross section from the ZEUS and H1 collaborations are presented in Sections 3.2 and 3.3, respectively. A comparison of the measurements of the beauty quark contribution to the proton structure function from the two experiments is shown in Section 3.4.

3.1 Beauty Quark Identification Techniques

Various methods (variables) for beauty quark identification are used for the measurement of the b-quark production cross section. A detailed explanation of these techniques used in this thesis is given in Chapters 7 and 8. In the following a brief introduction to the identification methods used in the previous measurements presented in Sections 3.2 and 3.3 is given. The different algorithms are often combined to enhance the separation of the beauty quark signal from charm and light flavour background.

- Beauty quarks can be identified by their semileptonic decay into a muon or an electron using the transverse momentum of the lepton relative to the axis of the jet to which they are associated ($p_{\mathrm{T}}^{\mathrm{rel}}$ method). Because of the large mass of b quark, the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution of beauty quark is harder than for charm and light flavour quarks

and hence can be used to separate the beauty signal from charm and light flavour backgrounds (cf. Section 7.2.1).

- The signed impact parameter, δ, which is defined as the distance of closest approach of a track to the primary vertex, reflects the lifetime of the weakly decaying heavy hadrons and hence can be used to discriminate between beauty and charm decays and the decays of light quarks. The sign of the impact parameter allows a statistical disentanglement of detector resolution effects from the effects of the decay lifetime of the heavy hadrons (cf. Section 7.3.1).

- The distribution of the component of the missing transverse momentum parallel to the lepton direction, $p_{\mathrm{T}}^{\mathrm{miss},||l}$, has a positive tail for events containing semileptonic heavy quark decays. This variable can be combined with other variables to enhance the separation of semileptonic beauty or charm decays due to the presence of the neutrino (cf. Section 7.3.3).

3.2 ZEUS Collaboration Measurements of Beauty

The ZEUS collaboration has published two results on beauty production in DIS using semileptonic decays into muons [9, 10].

The first measurement [9] used 114 pb^{-1} of data collected by the ZEUS detector in the years 1996-2000. The kinematic region of this analysis was defined by: $Q^2 > 2\,\mathrm{GeV}^2, 0.05 < y < 0.7$ and at least one jet with $E_{\mathrm{T}}^{\mathrm{jet}} > 5$ GeV and $-2.0 < \eta^{\mathrm{jet}} < 2.5$ including a muon of $p_{\mathrm{T}}^{\mu} > 1.5$ GeV and $\eta^{\mu} > -1.6$ inside a cone of $\Delta R < 0.7$ to the jet axis.

The fraction of events from b-hadron decays in the data sample was extracted by fitting the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution of the data using the Monte Carlo predictions for the processes producing beauty, charm and light quarks. In this analysis, the visible cross section, $\sigma_{b\bar{b}}$, and differential cross sections as a function of photon virtuality, Q^2, the transverse momentum of the muon, p_{T}^{μ}, and its pseudorapidity, η^{μ}, as well as the transverse momentum of the jet, $p_{\mathrm{T}}^{\mathrm{jet}}$ and its pseudorapidity, η^{jet}, were measured. They were compared to leading order (LO) plus parton shower (PS) Monte Carlo predictions and NLO QCD calculations.

Figure 3.1 shows the differential cross section as a function of p_{T}^{μ} and η^{μ} compared to the HVQDIS NLO calculation and the Rapgap Monte Carlo prediction scaled to the data. In shape, both the MC and the NLO QCD calculations reasonably describe the data. The largest fraction of the observed difference in normalisation to the NLO prediction is about 2 standard deviation in the low p_{T}^{μ} region.

The second measurement [10] used the dataset from the HERA II running period collected in 2005 corresponding to an integrated luminosity of 126 pb^{-1}. Charm and beauty quarks were identified through their decays into muons in a kinematic region of $Q^2 > 20\ \mathrm{GeV}^2, 0.01 < y < 0.7$, with at least one muon in the event with $p_{\mathrm{T}}^{\mu} > 1.5$ GeV and $-1.6 < \eta^{\mu} < 2.3$. A loose jet selection of $p_{\mathrm{T}}^{\mathrm{jet}} > 2.5$ GeV and $-3 < \eta^{\mathrm{jet}} < 3$ was applied.

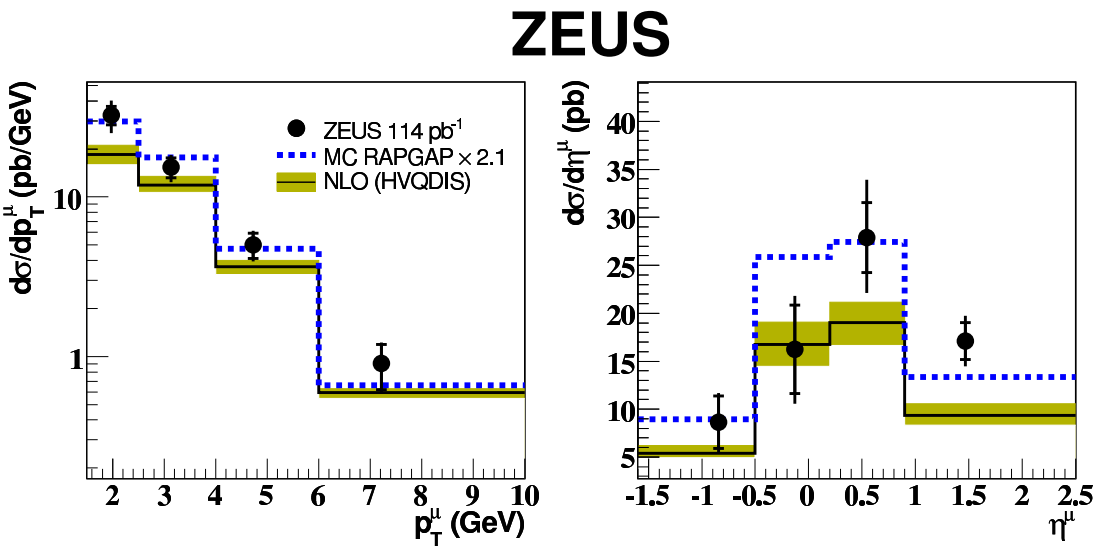


Figure 3.1: Differential beauty cross sections as a function of p_{T}^{μ} and η^{μ}, for events with at least one jet and one muon, compared to the HVQDIS NLO QCD calculations and to the scaled RAPGAP MC. The errors on the data points correspond to the statistical uncertainty (inner error bars) and to the statistical and systematical uncertainty added in quadrature (outer error bars). The shaded bands show the uncertainty of the theoretical prediction. Plots taken from [9].

The fractions of muons originating from charm, beauty or light flavour events were determined from a simultaneous fit of three discriminating variables sensitive to different aspects of heavy quark decays. These variables were $p_{\mathrm{T}}^{\mathrm{rel}}$, impact parameter and $p_{\mathrm{T}}^{\mathrm{miss},||\mu}$ (cf. Section 3.1). The visible cross sections, $\sigma_{b\bar{b}}$, and differential cross sections as a function of the kinematic variables, Q^2, x and muon variables, p_{T}^{μ} and η^{μ} were calculated.

The measured differential beauty (and charm) cross sections compared with the NLO QCD calculation based on HVQDIS and RAPGAP scaled Monte Carlo predictions are shown in Figure 3.2. The charm cross sections are in good agreement with the HVQDIS calculations. The tendency of the beauty cross sections to lie above the central NLO prediction is concentrated at low p_{T}^{μ} and Q^2. The RAPGAP Monte Carlo gives a good description of the shape of all differential cross sections.

In both analyses presented above, double differential cross sections were also measured in bins of x and Q^2, $d^2\sigma/dxdQ^2$, and were used to extract the heavy quark contribution to the proton structure function (see Section 3.4).

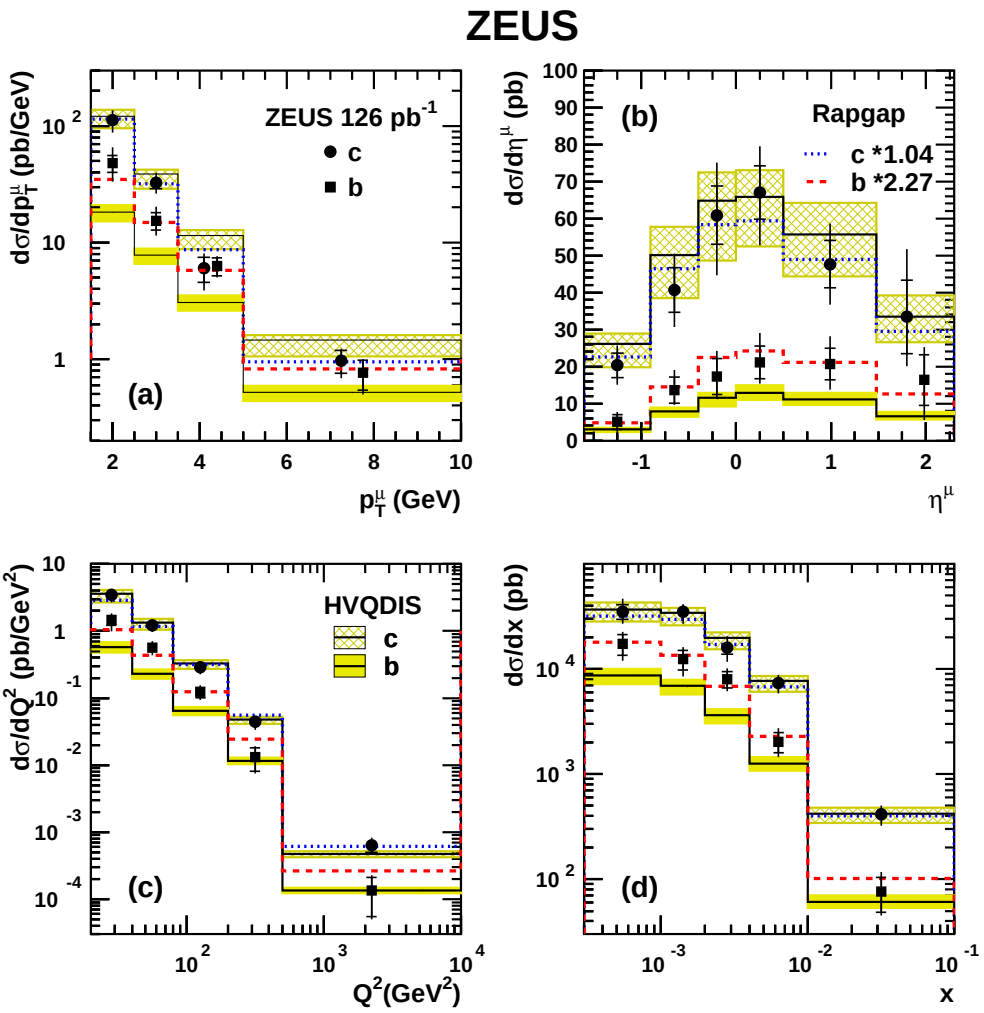


Figure 3.2: Differential muon cross sections for c and b quarks as a function of (a) p_{T}^{μ}, (b) η^{μ}, (c) Q^2 and (d) x. The inner error bars show the statistical uncertainty while the outer error bars show the systematic and statistical uncertainties added in quadrature. The bands show the NLO QCD predictions obtained with the HVQDIS program and the corresponding uncertainties. The differential cross sections from RAPGAP, scaled by the factors corresponding to the result of the fit, are also shown. Plots taken from [10].

3.3 H1 Collaboration Measurements of Beauty

A brief review of three measurements from the H1 collaboration for inclusive beauty cross sections is presented[1]. In these analyses the number of beauty events were determined

[1]Inclusive charm cross sections were also measured in these analyses, but are not discussed here.

using variables reconstructed by the H1 vertex detector including the impact parameter of tracks to the primary vertex and the reconstructed position of the secondary vertex.

Two measurements [58, 59] were based on the HERA I dataset, corresponding to an integrated luminosity of 57.4 pb^{-1} taken in the years 1999-2000. These measurements were performed in the low and high Q^2 regimes, restricted to the kinematic regions of $12 \leq Q^2 \leq 60\,\mathrm{GeV}^2, 0.0002 \leq x \leq 0.005$ and $Q^2 > 150\,\mathrm{GeV}^2, 0.1 < y < 0.7$, respectively. The third measurement [60] used a dataset with an integrated luminosity of $189\,\mathrm{pb}^{-1}$ which was about three times greater than in the HERA I measurements. The data were recorded in the HERA II running running period during 2006 and 2007. The analysis was performed in the kinematic range of $5 \leq Q^2 \leq 2000\,\mathrm{GeV}^2$ and $0.0002 \leq x \leq 0.05$ and hence covers both lower and higher Q^2 than the previous measurements.

The HERA I measurements made use of the impact-parameter significance distributions (ratio of impact parameter to its error). The quantities S_1, S_2 and S_3 defined as the significance of the track with the highest[2], second highest and third highest absolute significance were calculated[3]. In order to substantially reduce the uncertainty due to the resolution of impact parameter and the light quark normalisation, the contents of the negative significance bins were subtracted from the corresponding positive significance bins. The resulting distributions are dominated by c quark events, with a b-quark fraction increasing with

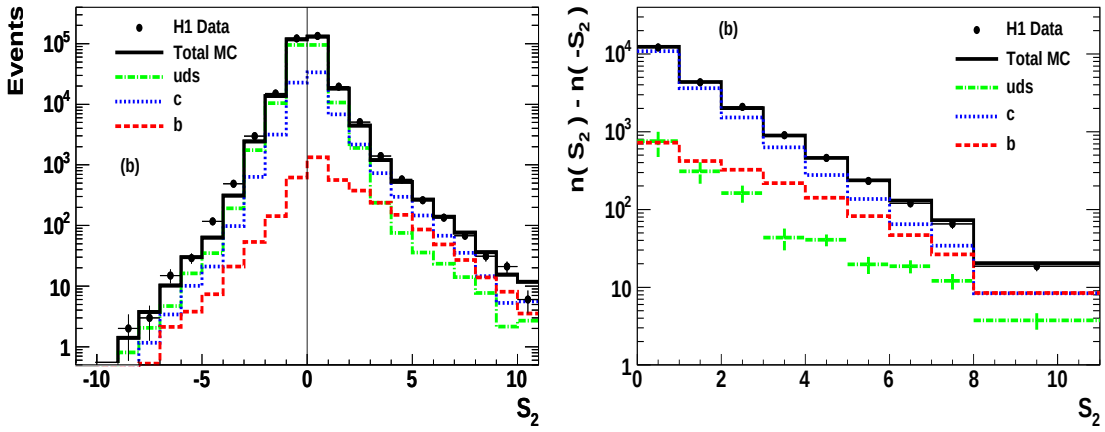


Figure 3.3: The significance, $\delta/\sigma(\delta)$, distribution of the track with the second highest absolute signicance, S_2 (left), and the subtracted S_2 (right). Included in the figure is the expectation from the DJANGOH Monte Carlo simulation for light quarks and that from the RAPGAP Monte Carlo simulation for c and b quarks. The contributions from the various quark flavours are shown after applying the scale factors obtained from the fit to the subtracted significance distributions of the data. Plots taken from [58].

[2]The definition of S_1 is different in [59], where the S_1 distribution only contained events with one reconstructed CST track and is simply the significance of the track.

[3]For the high Q^2 measurement, only S_1 and S_2 were used.

significance (see for example S_2 distribution in Figure 3.3 before and after subtraction). The data were split into bins in Q^2 and x and the contributions of beauty and charm were determined separately in each bin using a least squares simultaneous fit to the subtracted S_1, S_2 and S_3 distributions (only S_1 and S_2 for the high Q^2 analysis).

In the HERA II measurement, the reconstructed position of a secondary vertex in the transverse plane was also used in addition to the transverse displacement of tracks from the primary vertex. For events with three or more tracks in the vertex detector the reconstructed variables were used as an input to an artificial neural network. This method has better discrimination between c and b quark compared to HERA I measurements which used only the transverse displacement of tracks from the primary vertex. The measurement is made differentially by dividing the data into discrete $y - Q^2$ intervals. The fractions of c, b and light quarks in the data were extracted in each $y - Q^2$ interval using a least squares simultaneous fit to the subtracted S_1, S_2 and NN distributions and the total number of inclusive events before any track selection.

The results of fit in each $y - Q^2$ interval are converted to a measurement of the 'reduced b cross section' defined from the differential cross section as:

$$\tilde{\sigma}^{b\bar{b}}(x, Q^2) = \frac{d^2\sigma^{b\bar{b}}}{dxdQ^2} \frac{xQ^2}{2\pi\alpha^2(1 + (1 - y)^2)}. \tag{3.1}$$

The measurement of $\tilde{\sigma}^{b\bar{b}}$ as a function of x for fixed Q^2 values is depicted in Figure 3.4. Also shown are the HERA I data extracted using measurements in low and high Q^2 regions and the combined H1 data. The HERA I and HERA II datasets were combined for each $x - Q^2$ point where there were two measurements. The $\tilde{\sigma}^{b\bar{b}}$ data from HERA I and HERA II and the combined H1 data show a good agreement for all measured x and Q^2 values. The HERA II measurement is better and covers a wider kinematic range than the HERA I measurement.

3.4 Beauty Contribution to F_2

The measurements based on different tagging techniques presented in Sections 3.2 and 3.3 were used to extract the beauty contribution, $F_2^{b\bar{b}}$, to the proton structure function, F_2.

$F_2^{b\bar{b}}$ can be extracted from the reduced cross section or inclusive double differential cross sections as a function of x and Q^2 (for details see Section 11.3) by extrapolating the cross sections from the measured range to the full kinematic phase space. The extrapolation factors are calculated from the NLO QCD predictions.

A summary of the $F_2^{b\bar{b}}$ measurements as a function of Q^2 for different x values is shown in Figure 3.5. Both ZEUS measurements and the combined H1 measurement are compared with NLO and NNLO QCD predictions, based on fixed-flavour and variable-flavour number schemes. The HVQDIS calculation was done in FFNS using the ZEUS-S proton PDF. Among other theory predictions based on different parameters are, predictions by the

CTEQ [61] and MSTW [62] groups which use NLO calculations based on the general-mass variable-flavour-number scheme with different treatments of the flavour-threshold region [63]. The MSTW prediction is also available in a variant partially including NNLO terms [62]. The NLO prediction of GJR [64] is based on the FFNS. The prediction of ABKM [65] is based on a partial NNLO FFNS calculation which is almost complete in the threshold region $Q^2 \sim m_b^2$. The measurements from the two experiments are compatible within the uncertainties. The predictions from different theoretical approaches agree fairly well with each other as well as with the measurements.

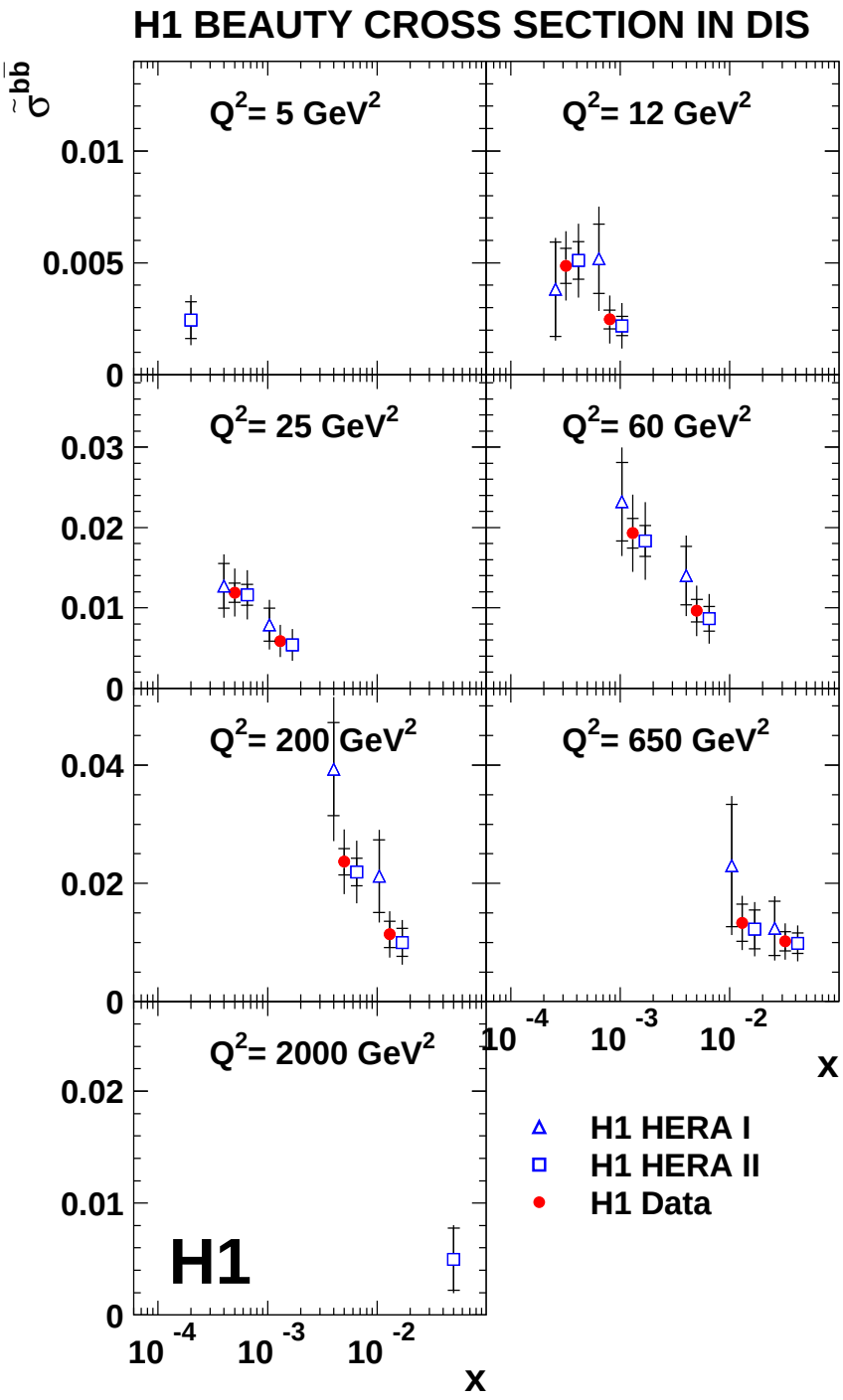


Figure 3.4: The measured reduced cross section $\tilde{\sigma}^{b\bar{b}}$ as a function of x for different Q^2 values. The inner error bars show the statistical error, the outer error bars represent the statistical and systematic errors added in quadrature. The HERA II measurements are compared with those from HERA I. The combined H1 data are also shown. The x values of the HERA I and HERA II data are shifted for visual clarity. Plot taken from [60].

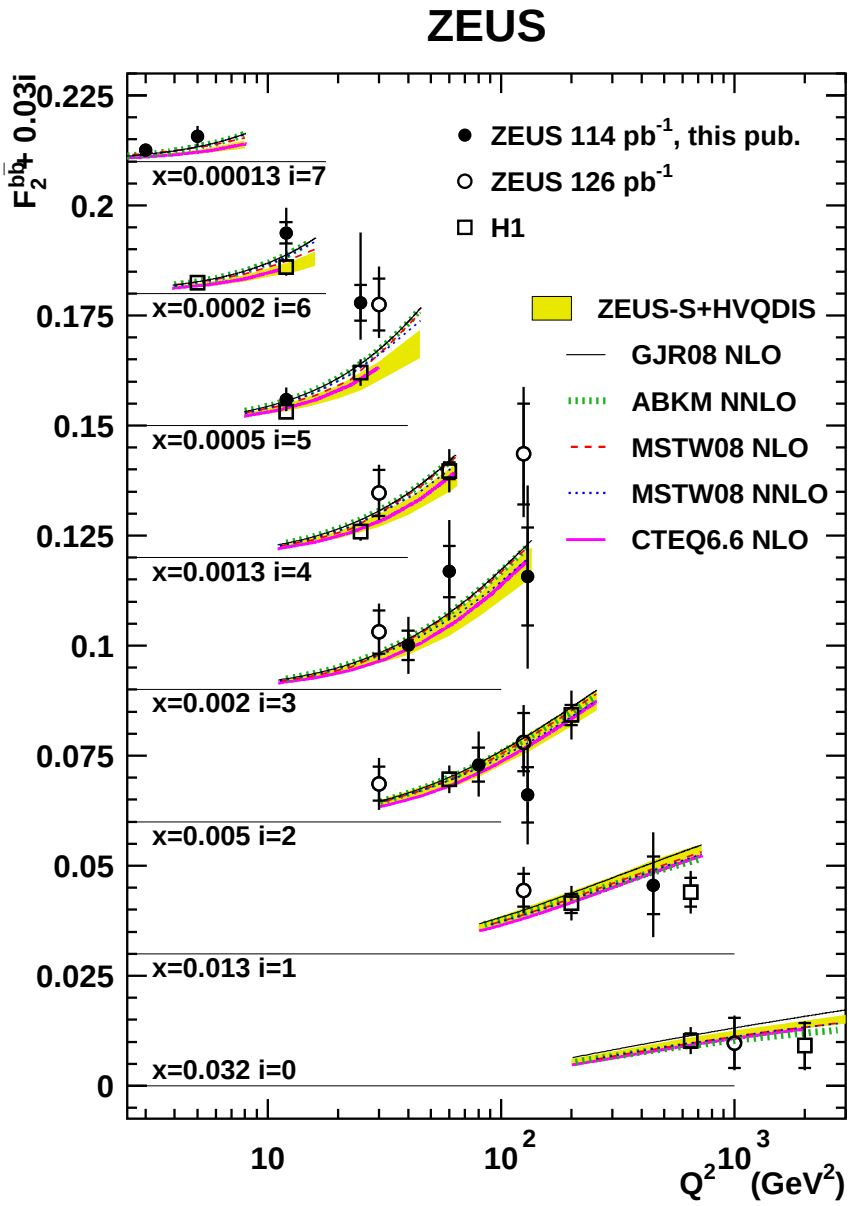


Figure 3.5: The structure function $F_2^{b\bar{b}}$ as a function of Q^2 for fixed values of x. The errors on the data points correspond to the statistical uncertainty (inner error bars) and to the statistical and systematical uncertainty added in quadrature (outer error bars). The horizontal lines indicate the zero-line for each series of measurements. Results from different measurements are compared with different QCD predictions (lines and band). The measurements from the two experiments are compatible within the errors and in agreement with the theory. Plot taken from [9]

Chapter 4

HERA and the ZEUS Detector

In this chapter the HERA collider and the ZEUS detector are introduced briefly, giving particular emphasis on the parts of the detector relevant for the analysis described in this thesis. In addition an overview of the ZEUS trigger and data acquisition system is given at the end of the chapter.

4.1 The HERA Collider

HERA (Hadron-Elektron Ring-Anlage) was the first particle accelerator for the study of high energy lepton-proton collisions [66]. It was built at the DESY (Deutsches Elektron Synchrotron) laboratory in Hamburg, Germany between 1984 and 1991, and was in operation from 1992 to 2007. It had a circumference of 6.3 km and was located in a tunnel 15-30 m underground. The HERA accelerator used to collide electrons/positrons, accelerated to an energy of 27.5 GeV, with 920 (820) GeV protons (the energy of the proton beam was 820 GeV at the beginning and was increased to 920 GeV in 1998). The resulting centre-of-mass energy, $\sqrt{s_{ep}}$, was 318 (300) GeV, more than an order of magnitude higher than the previous fixed-target lepton-nucleon experiments, which allowed a wider kinematic range in x and Q^2 to be explored.

Four experiments were in operation at HERA. The positions of the four different halls along the HERA ring where four experiments were located can be seen in Figure 4.1. The two particle beams, accelerated and stored in two separate rings, were brought into collision at zero crossing angle at two interaction points, one in the South Hall where the ZEUS experiment was located and the other in the North Hall where there was the H1 experiment. The other two experiments were the fixed target experiments and made use of only one beam. In the East Hall the HERMES experiment studied the spin structure of the nucleon using the collisions of longitudinally polarised leptons on an internally polarised gas target (H, D or He). The HERA-B experiment located in the West Hall was used until 2003 to collide the proton beam halo with wire targets and was designed to study CP-violation in the B sector.

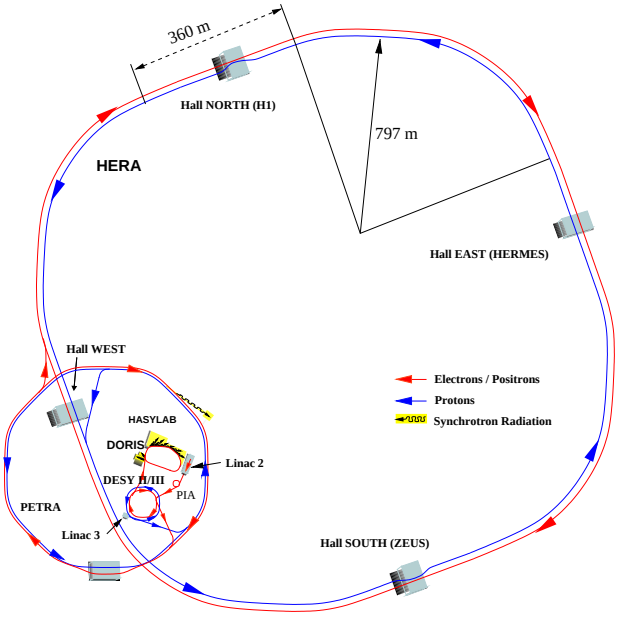


Figure 4.1: Schematic view of the HERA collider and its pre-accelerator system.

Figure 4.1 shows a schematic drawing of the HERA collider and its pre-accelerator systems. It had two different injection systems for the beams. The proton acceleration chain started with negative hydrogen ions (H^-) accelerated in a LINAC to an energy of 50 MeV. The electrons were then stripped off the H^- ions to obtain protons which were injected into the proton synchrotron DESY III and accelerated up to 7.5 GeV in 11 bunches with a temporal gap of 96 ns, the same as in the main HERA ring; these bunches were then transferred to PETRA, where they were accelerated to 40 GeV. Finally they were injected into the HERA proton storage ring and the injection stopped when the ring was filled with 210 bunches. In the HERA ring the proton beam was then accelerated up to 920 GeV.

The leptons were first accelerated to an energy of 250 MeV (electrons) and 450 MeV (positrons) in the linear accelerators LINAC I and LINAC II, respectively. They were then injected to the DESY II synchrotron, accelerated to 7.5 GeV and then transferred to PETRA II, where they reached an energy of 14 GeV in bunches separated by 96 ns gaps. Finally they were injected into the HERA ring, until it contained 210 bunches and accelerated to the nominal lepton beam energy of 27.5 GeV. In order to study the background conditions, some of the bunches were kept empty, they are known as pilot bunches. When both bunches were empty the non-beam related background, such as cosmic rays, could be studied. When either the lepton or the proton bunch was empty, it was used to study the beam-related background originating from the interaction of the lepton or proton beam with the residual gas in the beam pipe.

The HERA collider started operation in 1992 in its initial configuration with 820 GeV protons and 26.7 GeV electrons. The energy of the electron beam, E_e, was slightly raised

to 27.5 GeV in 1994 and that of the proton beam, E_p, was increased to 920 GeV in 1998. Most of the time HERA used the positron beam to get more stable running conditions. The electron beam was only used during the years 1998, part of 1999, 2005 and part of 2006. In the end there was similar luminosity for electron and positron running periods.

HERA had two phases of operation, known as HERA I and HERA II. The first phase of operation, HERA I, lasted from 1992 to 2000 and the total delivered luminosity during that period was $193\,\mathrm{pb}^{-1}$. During the shutdown in 2000/2001, the HERA collider was upgraded to deliver about a factor five time higher specific luminosity at the interaction point [67]. In addition spin rotators were included to rotate the spin of the leptons such that the lepton beam was longitudinally polarised at all interaction regions. The experiments also took the opportunity of the shutdown to upgrade their detectors to achieve higher precision measurements, especially in the context of heavy flavour and high Q^2 physics. After the detector and accelerator upgrades, the second phase of data taking, HERA II, began in 2003. From October 2003, until the end of 2006, HERA provided stable beam operations and delivered a total luminosity of $556\,\mathrm{pb}^{-1}$. Before the end of data taking (June 2007), special data, using two different lower proton beam energies 460 GeV (Low Energy Run) and 575 GeV (Medium Energy Run) were taken. The corresponding luminosities for the LER and MER were $15.8\,\mathrm{pb}^{-1}$ and $8.1\,\mathrm{pb}^{-1}$, respectively. This data was taken in order to measure the longitudinal structure function, F_L.

A summary of some of the HERA parameters for different running periods is given in Table 4.1 and a comparison of the HERA integrated luminosity during the two running periods (HERA I and HERA II) is shown in Figure 4.2.

Running period	Lumi. ($\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	Energy (GeV)		$\sqrt{s_{ep}}$ (GeV)	Beam current (mA)	
		Lepton	Proton		Lepton	Proton
1993 - 1997	$1.6 \cdot 10^{31}$	27.5	820	300	43	163
1998 - 2000	$1.6 \cdot 10^{31}$	27.5	920	318	43	163
2003 - 2006	$7.0 \cdot 10^{31}$	27.5	920	318	58	140
2007 LER	$1.47 \cdot 10^{31}$	27.5	460	225	58	140
2007 MER	$1.47 \cdot 10^{31}$	27.5	575	251	58	140

Table 4.1: Typical parameters of the HERA collider.

4.2 ZEUS Detector

The ZEUS detector [68] was a multi-purpose detector designed to study a wide range of physics involving high energy lepton-proton collisions at HERA. It had a size of $12\times11\times20$ m^3 and a weight of 3600 tons. It was a quasi-hermetic detector as it covered most of the 4π

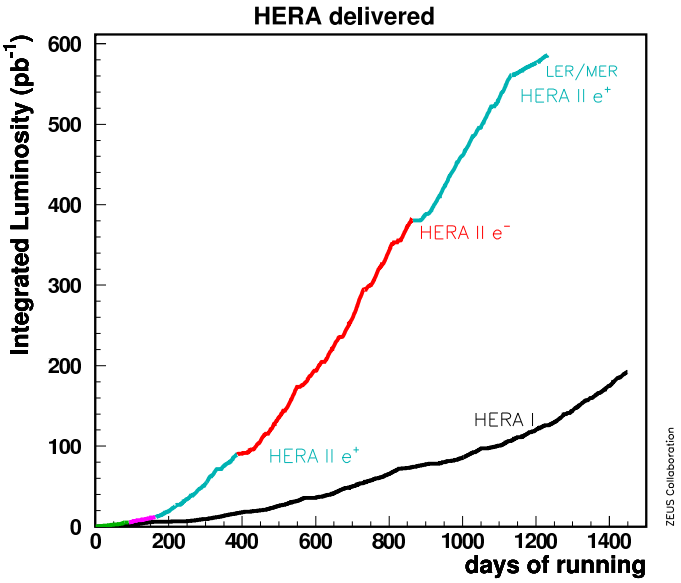


Figure 4.2: HERA delivered integrated luminosity for each running period.

solid angle with the exception of very small regions around the beam pipe. As a result of the energy difference between the electron and proton beams, most of the final state particles were boosted in the proton beam direction. Thus the ZEUS detector was specially instrumented with additional sub-detectors in the proton beam direction.

The ZEUS coordinate system [69], depicted in Figure 4.3 is a right-handed Cartesian system with the origin at the nominal interaction point. The Z–axis points in the direction of the proton beam and is often referred to as the "forward direction". The X–axis is perpendicular to the beam direction pointing towards the centre of the HERA ring and the Y–axis points upwards. The polar angle, θ, and the azimuthal angle, ϕ, are measured relative to the Z–axis and X–axis, respectively. In general, the pseudorapidity, $\eta = -\ln\tan(\theta/2)$, is used instead of the angle θ, as a difference in η is longitudinally Lorentz-invariant.

The ZEUS detector had the typical design of high energy physics experiments. A schematic overview of the detector in the Z–Y and the X–Y planes can be seen in Figure 4.4 and 4.5, respectively. A short description of the main detector components is given below followed by the sections where a more detailed description of the components most relevant for the analysis is given.

Starting from the interaction point and moving radially outwards, the ZEUS detector consisted of charged particle tracking detectors surrounding the beam pipe. The innermost detector was the vertex detector (VXD) which was removed during the 1995-1996 shutdown and, in 2001, a silicon microvertex detector (MVD) was installed in its place to improve the tracking and especially vertexing. The MVD was surrounded by the central tracking detector (CTD) which was a cylindrical drift chamber. The CTD was supplemented in the forward region by three sets of planar drift chambers (FTD) with interleaved transition

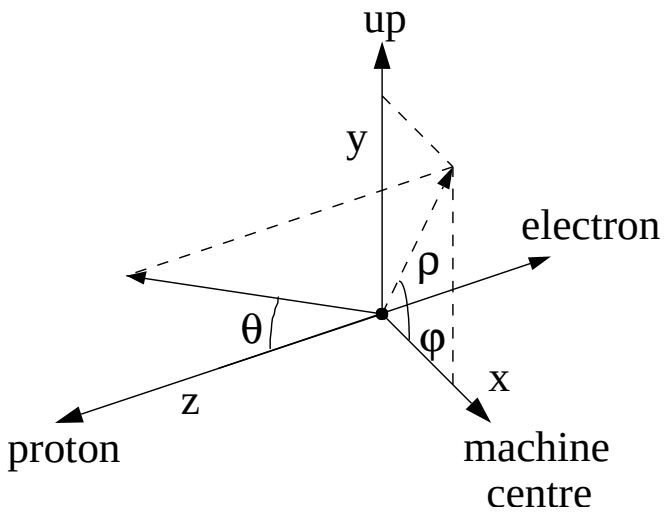


Figure 4.3: The ZEUS coordinate system.

radiation detectors (TRD) and in the rear direction with one planar drift chamber consisting of three layers (RTD). In order to improve the forward tracking, in 2001, the TRD system was replaced by a straw-tube tracker (STT) [70] consisting of two modules. The last tracking detector was the small-angle rear tracking detector (SRTD) which helped to improve the position resolution for particles. All the tracking components combined, offered an angular acceptance of $7.5° < \theta < 170°$ for charged particles. The tracking system was enclosed by a superconducting solenoid providing a magnetic field of 1.43 T for the determination of charge and momentum of the charged particles.

A high resolution compensating uranium-scintillator calorimeter (CAL) surrounded the tracking detectors and the magnet. It was the main instrument to measure the energy of electrons and hadrons and was subdivided into three sections, the forward (FCAL), the barrel (BCAL) and the rear (RCAL) calorimeters. Inside of the FCAL and RCAL the hadron-electron separator (HES) [71] was placed, which helped to distinguish electrons from hadrons. In front of the CAL, scintillating tiles, called the presampler were used for the detection of pre-showering particles improving the energy measurement. Further, the CAL was surrounded by the iron yoke which provided the return path for the magnetic field flux and was instrumented with proportional tubes serving as a calorimeter (backing calorimeter, (BAC)) for the particles which had not been observed by the CAL. The BAC was also used as a muon detection system. Dedicated muon identification chambers were located inside (FMUI, BMUI, RMUI) and outside (FMUO, BMUO, RMUO) the iron yoke.

Other detectors were located several metres away from the main detector along the beam axis. The veto wall was located in the rear direction at about $Z = -7.5$ m from the interaction point.

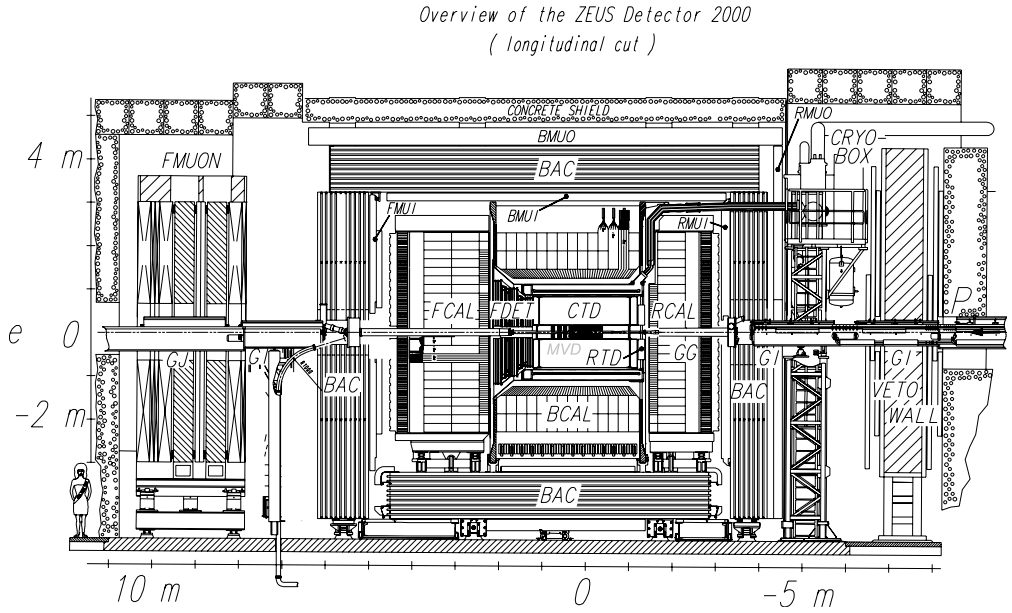


Figure 4.4: The ZEUS detector in a Z–Y cross section.

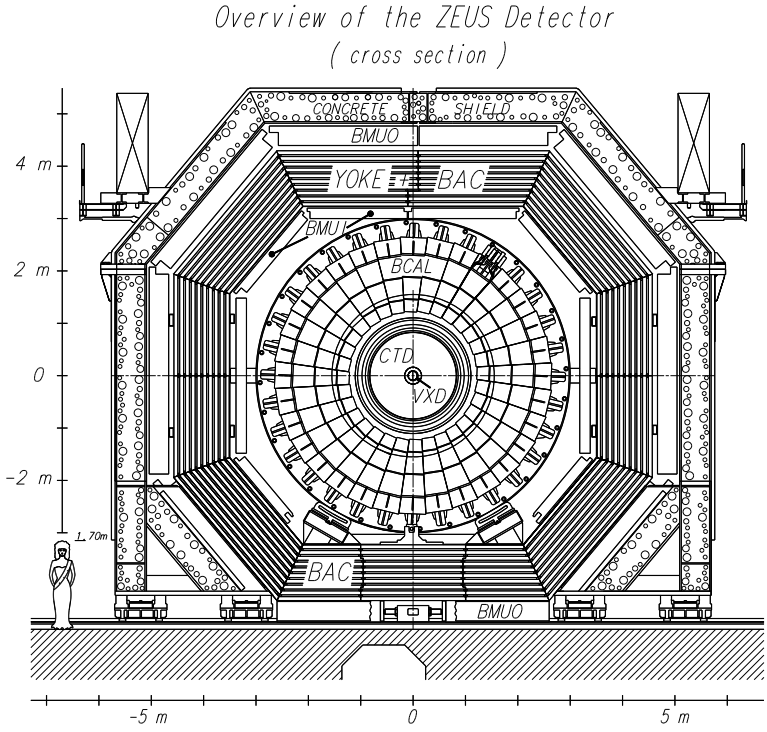


Figure 4.5: The ZEUS detector in an X–Y cross section.

It consisted of an iron wall (7.6×0.9 m) covered on both sides with scintillation counters and was used to reject background from beam gas interactions. In addition, two kinds of luminosity monitor were placed at $Z \simeq 100$ m, namely the photon calorimeter and the spectrometer for the luminosity measurement.

4.2.1 Microvertex Detector

As described in Section 4.2, during the HERA upgrade shutdown, in 2001, a silicon microvertex detector (MVD) was installed in the ZEUS experiment in the gap between the beam pipe and the inner volume of the CTD. The MVD was designed to improve the overall precision of the tracking system and to allow the identification of events containing secondary vertices originating from the decay of particles with long lifetime ($c\tau \geq 100\mu$m) [72]. Some of the design specifications of the MVD were:

- polar angular coverage of $10° - 160°$;
- measurement of three points for each track in two independent projections;
- 20 μm intrinsic hit resolution;
- two-track separation of 200 μm.

In order to have a good matching with the existing detectors (CTD, FTD and RTD), the MVD was split into two sections: the barrel (BMVD) and the forward (FMVD).
The barrel section of the MVD was 64cm long and had a polar angular coverage of $30-150°$. It was sub-structured in three layers to allow high efficiency in the pattern recognition. The inner layer followed the elliptical path around the beam pipe and was placed at a variable radius between 3 and 5 cm from the CTD axis. This layer was not hermetic in ϕ due to the extra space needed for the elliptical beam pipe. The second and third layers were placed along a circular path at $r \sim 8.6$ cm and $r \sim 12.3$ cm, respectively (see Figure 4.6).

The BMVD was equipped with 600 single-sided silicon strip sensors of approximately 64×64 mm^2 size and 320 μm thickness with p^+ strips implanted into n-type bulk. Each sensor had 3082 strips with 20 μm pitch. Every sixth strip was read out resulting in 512 readout strips at a readout pitch of 120 μm. Two sensors were glued together to form a half-module as depicted in Figure 4.7(left). The strips of two sensors when glued together ran perpendicular to each other. A full module was formed by gluing a mirror imaged half module on the top of another module. Five modules were then glued on a carbon fibre structure called a ladder (see Figure 4.7(right)). When placed in the BMVD, the sensors parallel to the beam pipe gave the $r-\phi$ information while the perpendicular ones gave the Z-information. Figure 4.6 shows how the 30 ladders were located around the beam pipe.

The FMVD consisted of 4 planes (wheels) perpendicular to the beam axis. Each wheel was made of two back to back layers of 14 silicon sensors with same technical characteristics as

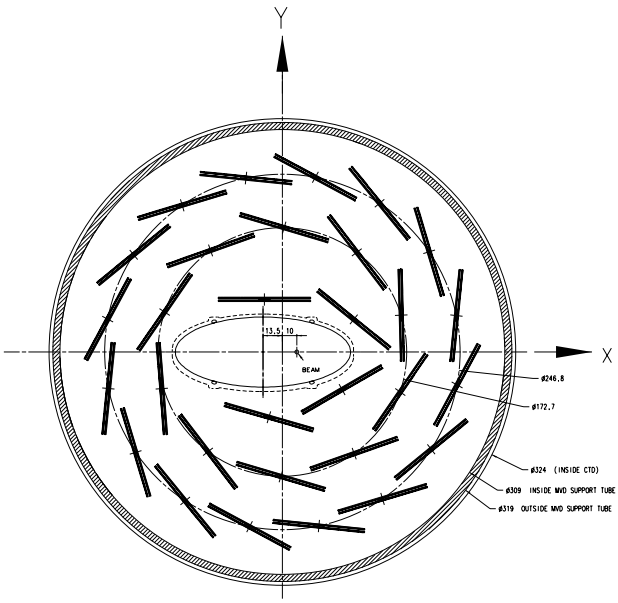


Figure 4.6: Cross section of the BMVD: three silicon layers positioned around the beam pipe are shown.

of the barrel sensors but with a trapezoidal shape (see Figure 4.8). Another difference was that the number of readout strips per sensor was 4800. The four wheels were positioned at $Z = 32, 45, 47$ and 75 cm. The FMVD allowed the polar angular coverage to be extended to 7° providing additional forward tracking capabilities. In each wheel the two layers were parallel but the strips were tilted by $\sim 13°$ in opposite directions to provide two coordinates for a particle traversing the wheel.

The MVD silicon sensors were read out using the analog chips Helix 3.0 [73] which were mounted on ceramic structures called hybrids (see Figure 4.7). A more detailed description of detection principle, layout and readout system of the MVD can be found in [74].

4.2.2 Central Tracking Detector

The central tracking detector [75, 76] was a cylindrical wire drift chamber designed to measure the direction and momentum of the charged particles as well as their energy loss, $\mathrm{d}E/\mathrm{d}x$, which provides information for particle identification. It had an overall length of 240 cm and an outer radius of 85 cm, while the active volume had a length of 203 cm with inner and outer radii of 18.2 cm and 79.4 cm, respectively. Longitudinally it covered the region from $Z = -100$ cm to $Z = 103$ cm resulting in a polar angular coverage of $15° < \theta < 164°$.

The CTD consisted of 72 concentric sense-wire layers arranged in nine super-layers as shown in Figure 4.9(a). Each super-layer consisted of cells made up of 8 sense wires with

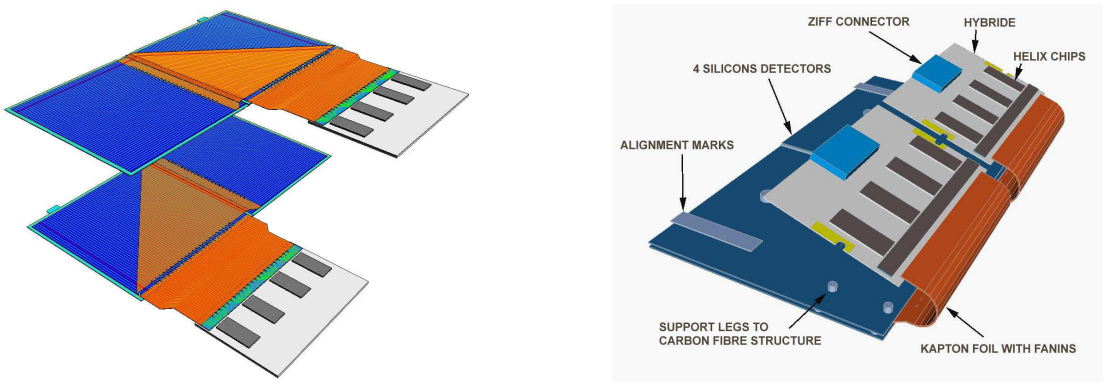


Figure 4.7: Assembly of two half modules into a module (left) and a ladder mounted on the support (right).

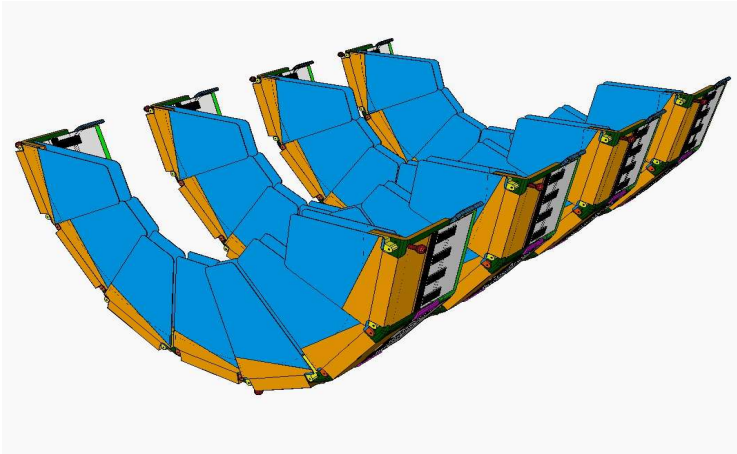


Figure 4.8: FMVD 4 wheels.

their associated field wires (see Figure 4.9(b)). The number of cells increased from 32 in the innermost super-layer to 96 in the outermost super-layer. In total there were 4608 sense wires with positive voltage and 19584 field wires with negative voltage.

In the odd numbered (axial) super-layers the wires were installed parallel to the Z–axis while the wires in the even numbered (stereo) super-layers had a small stereo angle of $\pm 5°$. With this configuration the Z-position of a track could be reconstructed with an accuracy of about 2 mm. All the wires of super-layer 1 and half the wires of super-layers 3 and 5 were additionally instrumented with a Z-by timing system. This system was used to estimate the Z-position of a hit by measuring the difference in arrival time of the pulse on the sense wires at each end of the detector. The achieved resolution of this system, 4.4 cm, was worse than the one obtained using full axial and stereo wire information, but it was faster and therefore was used predominately for trigger purposes.

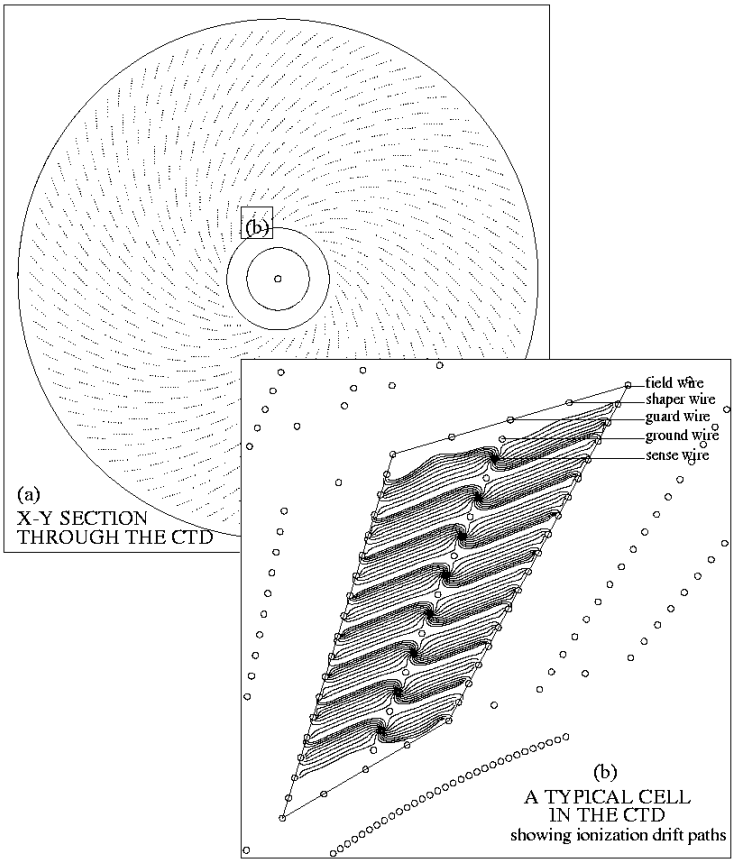


Figure 4.9: X–Y section through the CTD. (a) shows nine concentric super-layers surrounding the beam pipe. (b) shows a single CTD cell consisting of sense and field forming wires.

The CTD was filled with a gas mixture of argon (82%), ethane (13%) and carbon dioxide (5%) up to the year 2000. In that year, a trace of water ($\sim 0.15\%$) was added to the gas mixture to reduce deposits on the wires. A charged particle traversing the CTD produced ionisation of the gas in the chamber. The resulting electrons drifted towards the positive sense wires, whereas the positively charged ions were repelled and drifted towards the negatively charged field wires. The drift velocity of the electrons was approximately constant and equal to $50\,\mu$m/ns. Close to the sense wires, where the field was very strong, an avalanche-like multiplication of the electrons occured, which resulted in a signal pulse which was measured via electronic readout with a resolution of 8 bits. The measured pulse height was proportional to the energy loss of the initial particle, and was used to measure the specific ionisation. This pulse height was corrected for a number of effects [77].

To estimate the ionisation energy loss per unit length, $\mathrm{d}E/\mathrm{d}x$, of particles in the CTD, the truncated mean of the anode-wire pluse heights was calculated, which removes the lowest 10% and at least the highest 30% depending on the number of saturated hits.

The measured $\mathrm{d}E/\mathrm{d}x$ values were corrected by normalising to the average $\mathrm{d}E/\mathrm{d}x$ for tracks around the region of minimum ionisation for pions with momentum, p, satisfying $0.3 < p < 0.4\,\mathrm{GeV}$ [77].

As mentioned in Section 4.2, the CTD was enclosed in a superconducting solenoid which provided a magnetic filed of 1.43 T. This field caused charged particles to travel on a circular path of radius, R, given by

$$R = \frac{p_{\mathrm{T}}}{QB} \tag{4.1}$$

where Q is the charge of the particles (Coulomb), B is the strength of the magnetic field (Tesla) and p_{T} is the transverse momentum ($\mathrm{kg\,m\,s^{-1}}$). This equation allows p_{T} of the charged particle to be determined. The resolution on the transverse momentum, p_{T}, of tracks fitted to the interaction vertex and passing at least three CTD super-layers (with $p_{\mathrm{T}} > 1.5\,\mathrm{GeV}$) is [78]:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = 0.0058 p_{\mathrm{T}} \oplus 0.0065 \oplus \frac{0.0014}{p_{\mathrm{T}}}, \tag{4.2}$$

where p_{T} is expressed in GeV and the symbol $\oplus$ indicates the quadratic sum. The first term corresponds to the resolution of the hit positions while the second and third terms are due to multiple scattering before and inside the CTD, respectively.

After the installation of the MVD in 2001, the resolution changed. The influence of the multiple scattering was larger due to the extra material of the MVD whereas the hit resolution of combined tracking was better compared to the resolution quoted above due to the inclusion of the MVD hits. The new result for the transverse momentum resolution, after including the MVD in the global tracking reconstruction is [79]:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = 0.0026 p_{\mathrm{T}} \oplus 0.0104 \oplus \frac{0.0019}{p_{\mathrm{T}}}. \tag{4.3}$$

4.2.3 Uranium-Scintillator Calorimeter

The ZEUS calorimeter [80] was a high resolution compensating calorimeter. It was designed to provide a measurement of the energy of the particles by their total absorption and positions of the energy deposits. It completely surrounded the tracking devices and solenoid and covered 99.7% of the solid angle. It consisted of alternating layers of absorbing and active material. The absorber layers were made of 3.3 mm thick depleted uranium plates (98.1% ^{238}U, 1.7% N_b, 0.2% ^{235}U) wrapped in stainless steel foils and the active layers were made of 2.6 mm thick plastic scintillator (SCSN 38) plates. The thickness of the absorber and of the active layers was chosen in order to have linear and equal response of an electron and a hadron of the same energy passing through the detector, $e/h = 1.00 \pm 0.02$. This mechanism known as compensation allowed good resolution to be achieved in the determination of both the electromagnetic and hadronic energy. The electromagnetic and hadronic energy resolutions of the ZEUS calorimeter measured using test-beam conditions

were:

$$\frac{\sigma(E_e)}{E} = \frac{18\%}{\sqrt{E}} \oplus 2\% \tag{4.4}$$

$$\frac{\sigma(E_{\text{had}})}{E} = \frac{35\%}{\sqrt{E}} \oplus 1\% \tag{4.5}$$

where E is the particle energy measured in GeV.

The principle of operation of the CAL was that particles entering the absorber interacted with the uranium layer providing secondary particles and hence generated showers. The energy of the particles was measured via scintillating material which converted the deposited energy into light and then was measured via photomultipliers. This method ensured that both, neutral and charged particles would be detected. Due to Bremstrahlung, the typical showers of electrons are shallow while those of hadrons are deeper and broader. Muons act as minimum ionising particles and hence do not shower in the CAL. The difference of the various types of showers is depicted schematically in Figure 4.10.

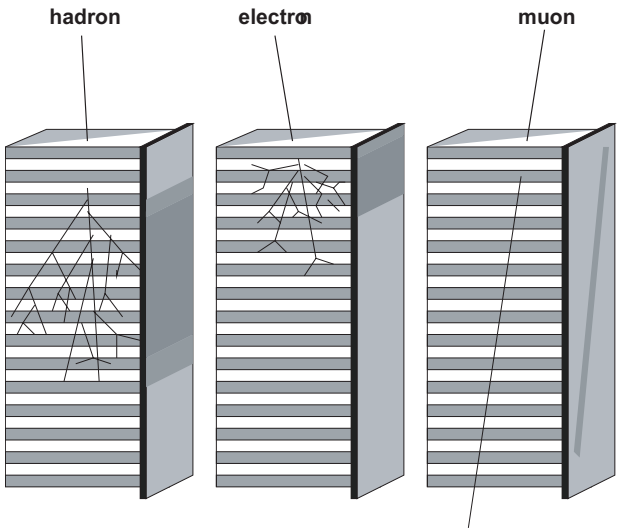


Figure 4.10: Different types of shower shapes in the CAL.

The CAL was divided into three parts: the forward (FCAL), the barrel (BCAL) and the rear calorimeter (see Figure 4.11 (left)). Since most of the final state particles in a lepton-proton interaction at HERA were boosted to the forward direction, the three parts had different thickness. The FCAL being thickest part was $\sim 7\lambda$, the BCAL $\sim 5\lambda$ and the RCAL $\sim 4\lambda$, where λ is the interaction length. The polar angular coverage of the three parts of the calorimeter is given in Table 4.2. The three parts of the calorimeter were divided into modules. The 23 FCAL modules and the 23 BCAL modules were rectangular, whereas the 32 BCAL modules which surrounded the cylindrical CTD were wedge-shaped covering 11.25° in azimuth. An FCAL module is shown in Figure 4.11 (right). Each module was segmented into $20 \times 20\,\text{mm}^2$ towers, which were further segmented longitudinally

Section	Polar angle	pseudorapidity
FCAL	$2.2° < \theta < 39.9°$	$1.0 < \eta < 4.0$
BCAL	$36.7° < \theta < 129.1°$	$-0.7 < \eta < 1.1$
RCAL	$128.1° < \theta < 176.5°$	$-3.5 < \eta < -0.7$

Table 4.2: Angular acceptance of the different parts of the CAL.

into one electromagnetic (EMC) and two hadronic (HAC) sections (RCAL had only one hadronic section). EMC sections were transversely divided into four cells (only two in the

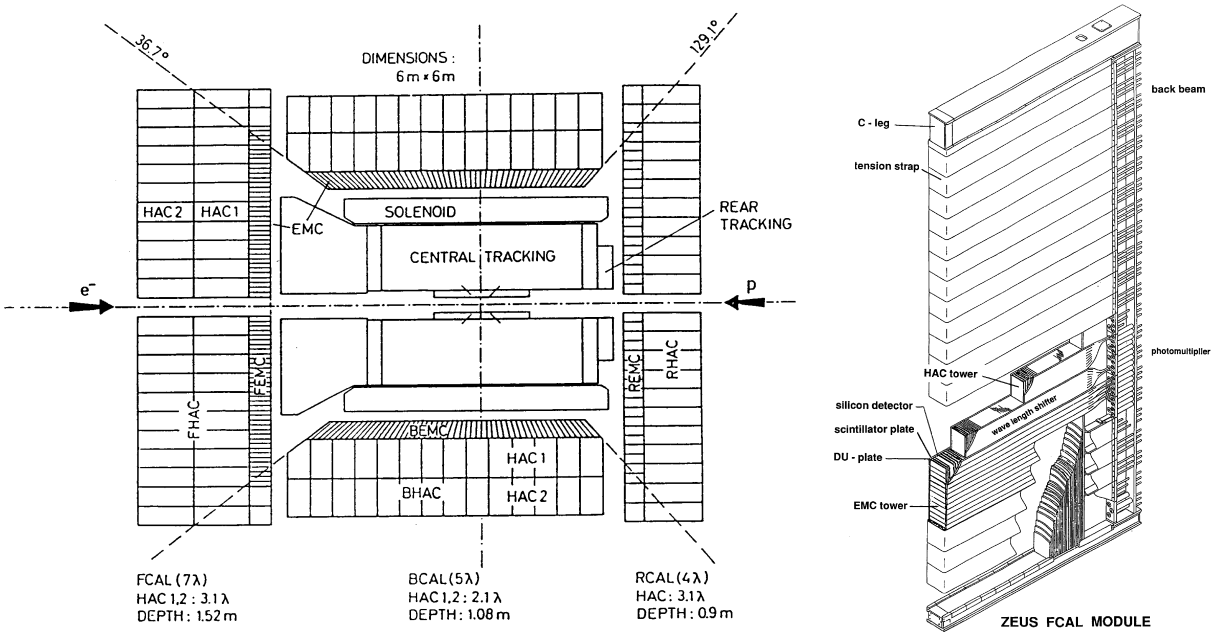


Figure 4.11: Schematic view of the CAL along the beam axis; showing the FCAL, BCAL and the RCAL with their electromagnetic (EMC) and hadronic (HAC) parts (left). Layout of an FCAL module (right).

RCAL). The size of the cell was $20 \times 5\,\mathrm{mm}^2$ for the FCAL and BCAL and $20 \times 10\,\mathrm{mm}^2$ for the RCAL. In total CAL had 5918 cells. Each cell was read out on two sides by 2 mm thick wave-length shifters, which were coupled to photomultiplier tubes. The summed information per cell was used for energy and time measurements. The CAL helped to provide accurate timing information, with a resolution of 1 ns for particles with an energy deposit greater than 1 GeV. This information was used to determine the timing of the particles with respect to the bunch-crossing time and was very useful for trigger purposes in order to reject background events.

4.2.4 Rear Tracking Detector (SRTD)

The small-angle rear tracking detector (SRTD) [82] was able to improve the angular measurement compared to the RCAL. It helped to make a rather precise measurement of the scattered electron which was needed to determine the kinematic variables of the event. SRTD had two scintillator strip layers. One layer was mounted in the vertical direction and the other one in the horizontal direction. The strips were either 24 cm or 44 cm long, 1 cm wide and 0.5 cm thick. The total covered surface was about 4200 mm^2 around the beam pipe at $Z = -148$ cm and provided an angular coverage between 162° and 176°.

4.3 Luminosity Measurement

The high precision measurement of the integrated luminosity is essential for the determination of cross sections. At ZEUS the luminosity was determined by measuring the Bethe-Heitler QED bremsstrahlung process $ep \to ep\gamma$, where the leptons and photons are scattered at very small angles. The cross section of this process is large ($\sigma_{BH} \geq 20$ mb) and can be calculated as a function of the photon energy to an accuracy of 0.5% [84]. The precise measurement of bremsstrahlung process and its high rate allowed an accurate determination of the luminosity.

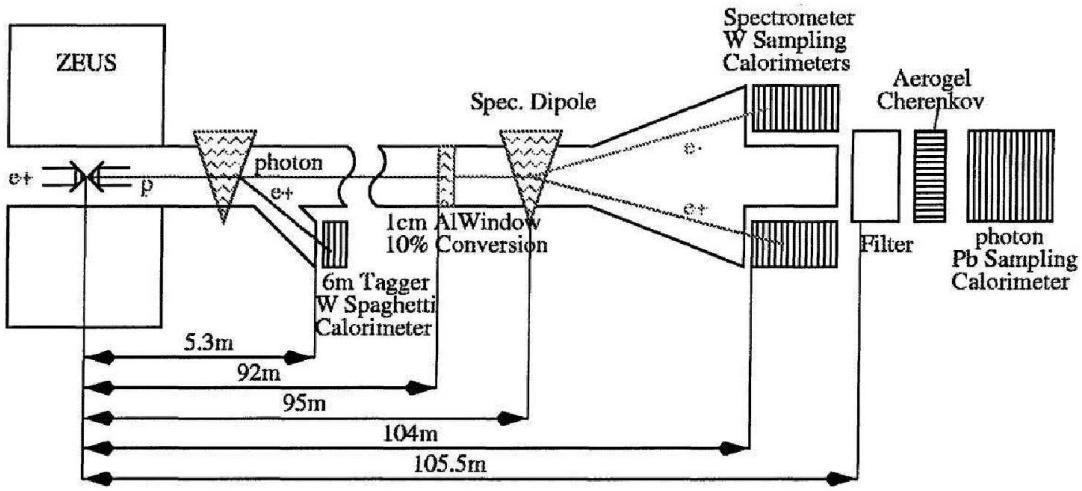


Figure 4.12: The ZEUS luminosity monitor system.

Two independent monitors were used to measure the luminosity at the ZEUS detector during the HERA II running, the *photon calorimeter* (PCAL) [83] and the *spectrometer* (SPEC) [85]. The schematic layout of both PCAL and SPEC systems is shown in Figure 4.12. The PCAL system was a sampling lead-scintillator calorimeter at $Z = -107$ m

which detected the photons exiting the beam pipe through a copper-beryllium window at $Z = -92.5$ m. To protect against synchrotron radiation, the photon calorimeter was shielded by a carbon-lead filter. The achieved calorimeter resolution was $\sigma(E)/E = 23\%/\sqrt{E}$, with E measured in GeV.

The bremsstrahlung event rate was determined by counting the number of photons above a fixed energy threshold. The luminosity was then calculated by dividing the measured rate by the bremsstrahlung cross section using the relation:

$$\mathcal{L} = \frac{N_\gamma}{\sigma_{\text{proc}}} \tag{4.6}$$

where σ_{proc} has to be corrected for the detector acceptance. The main contribution to the background was the bremsstrahlung of the leptons interacting with residual beam gas. This was estimated by using pilot bunches (lepton bunch without a matching proton bunch) evaluating for these the rate of bremsstrahlung events.

The SPEC system was installed for HERA II running as a complementary system to the PCAL with particular emphasis on combating the possible problem of pile-up arising from the increased luminosity. Pile-up occurs when one electron-proton bunch crossing produces multiple bremsstrahlung photons which cannot be distinguished from each other. The SPEC design avoided this by not measuring the bremsstrahlung photons directly but instead measuring the e^+e^- pairs arising form the pair production of bremsstrahlung photons in the exit window. The same exit window at $Z = 92.5$ m used for the PCAL was used for the SPEC and approximately 10% of the bremsstrahlung photons converted into electron pairs, giving an order of magnitude reduction in the rate of observed photons. The electron pairs were separated from the circulating beams, and then spatially split by the magnetic field of a dipole magnet. They were then measured by two segmented tungsten-scintillator sampling calorimeters which were 84 mm apart. The initial achieved luminosity precision for HERA II data was 3.5% but subsequent studies and calibrations reduced this to 2.6%.

4.4 Trigger System and Data Acquisition

The bunch structure of the beams at HERA, produced a beam crossing every 96 ns corresponding to an event rate of approximately 10 MHz. The rate of interesting physics events resulting from the $ep-$collisions varied between 0.05 Hz for NC DIS events with $Q^2 > 100\,\text{GeV}^2$ and 250 Hz for soft PhP events for an instantaneous luminosity of $2 \cdot 10^{31}\,\text{cm}^{-2}\text{s}^{-1}$. On the other hand, the rate of background events exceeded the physics events by several orders of magnitude. Apart from noise in the sensors and electronics, the main background source was beam-gas events. These are interactions of leptons or protons with the residual gas nuclei or with the beam pipe and had a typical rate of 10 kHz. An additional background was due to cosmic ray muons passing through the detector. The rate that the ZEUS data acquisition (DAQ) was able to write to tape was about 5 Hz. Thus a significant

reduction of the data rate and size was needed. For this ZEUS adopted a three-level trigger system [68].

The three levels of the ZEUS trigger system were organised such that at each level the rate was reduced giving increased time to perform more detailed calculations. A schematic diagram of the trigger and data acquisition system is shown in Figure 4.13 and a more detailed description of the different levels is given below.

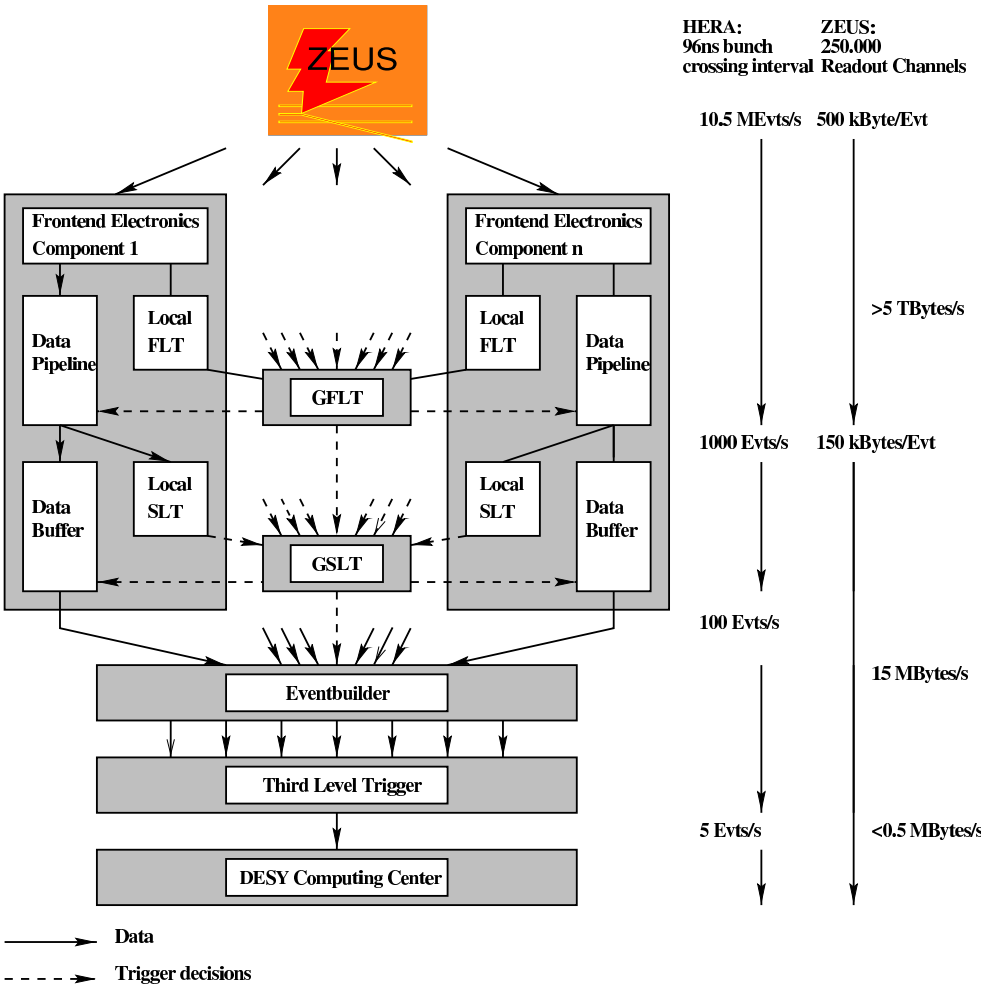


Figure 4.13: The ZEUS trigger and the data acquisition system.

4.4.1 First-Level Trigger

The first-level trigger (FLT) was a hardware-based trigger. Several ZEUS detector components (CTD, CAL, etc.) were equipped with their own FLT, implemented in hard-wired logic circuits. At this level a decision was taken on the basis of some event properties like

the transverse energy of the event and energy sums in different sections of the CAL. As the bunch crossing time (96 ns) was not enough to take a decision, the data of every bunch crossing was stored in pipelines which were able to store up to 46 bunch crossings. The FLT for each component made use of 26 bunch-crossing times (2.5 μs) to take a decision. The information from all the components was then sent to the global first-level trigger (GFLT) which made a final decision within 1.9 μs (corresponding to 20 bunch-crossing times). The overall time taken by FLT was $\sim$4.4 μs (corresponding to the 46 HERA bunch crossings stored in the pipelines). The resulting output rate of GFLT was below 1 kHz.

4.4.2 Second-Level Trigger

The second-level trigger was a software-based trigger which used a network of transputers [86], optimised for parallel processing. If the GFLT accepted the event, the data of all the components was analysed further and the event quantities were re-calculated to a higher degree of precision. Similarly to the FLT, the response from SLT processor of each component was sent to the global second level trigger (GSLT) where the event decision was made. The GSLT provided a decision after 7 ms, reducing the event rate to 50-100 Hz. The decision was based upon limited charged particle tracking, timing information from CAL, vertex determination, $E - P_z$, jet or scattered electron tagging.

In HERA II, the GSLT was supplemented by an additional tracking trigger, called the global tracking trigger (GTT). It was developed to take advantage of the extra tracking information provided by MVD. This algorithm ran on a PC farm, where the MVD hits were combined with the track segments from CTD and were then used to determine the position of the primary vertex more precisely.

For events accepted by the second-level trigger, the data from all the components was passed to the Event Builder (EB), which combined all the data of an event into a single record of ADAMO database tables [87] and made them accessible to the third-level trigger.

4.4.3 Third-Level Trigger

The third-level trigger (TLT) used a farm of processors where the events were processed in parallel. It consisted of a series of algorithms (a simplified version of the offline reconstruction software) which fully reconstructed the events using the information from several components. After this reconstruction, the trigger decision was made, based on physical quantities of the event, such as the kinematic variables, outputs of the electron finding algorithms, topologies of hadronic final state etc.

Events that passed the trigger criteria were then classified into different physics categories according to the various filters set by the physics working groups. The events accepted by one or more filters (with a size of $\sim$150 kb) were transferred to the DESY computer centre and recorded on tape for a re-processing with complete calibrations and full reconstruction

software. Trigger thresholds were adjusted such that the output rate of TLT was about 5 Hz.

Chapter 5

Event Reconstruction

The aim of this analysis is the measurement of beauty production in deep inelastic scattering (DIS), where the b quark decays semileptonically into an electron and a jet. This process is described by the relation:

$$ep \to e'bX \to e'e_{sl}\nu_e X'. \tag{5.1}$$

One example of such a beauty candidate event in the ZEUS event display (ZeVis) [88] is shown in Figure 5.1. The reconstruction of the event properties needed to identify such events consists of the reconstruction of final-state quantities including the scattered

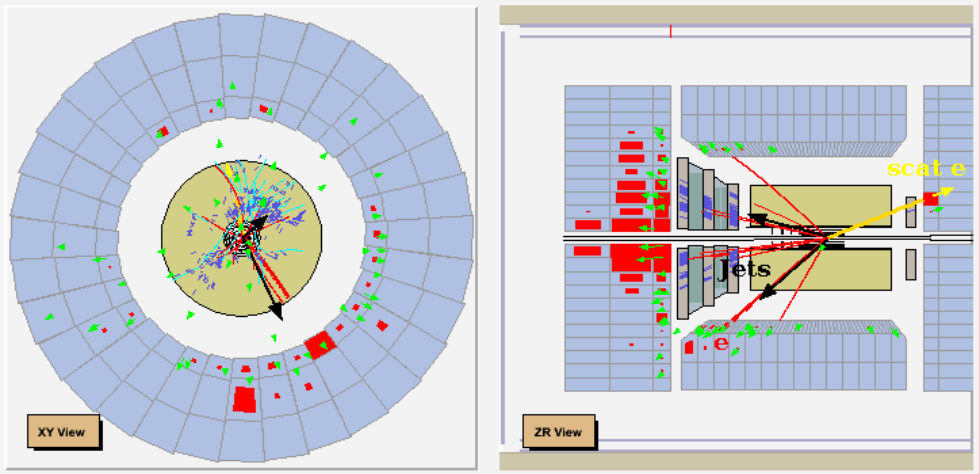


Figure 5.1: Event display of a simulated beauty candidate event in the X–Y plane (left) and in the Z–R plane (right). Shown are the reconstructed tracks (red lines), the CAL energy (red cells in the CAL) and the EFOs (green arrows). Two jets are reconstructed (black arrows) and yellow line in the RCAL is the scattered electron. An electron candidate (from semileptonic decay) associated to a jet in the BCAL can also be seen.

electron, the hadronic system of the b-quark decay and its associated electron. In this chapter the different steps to reconstruct these quantities are described.

As a first step the reconstruction of particle tracks using different tracking detectors (e.g. CTD, MVD, SRTD) is described in Section 5.1. After the track reconstruction, in Section 5.2, the vertex reconstruction is presented. Section 5.3 describes the algorithms used to reconstruct the scattered electron. Then in Section 5.4 the reconstruction of the hadronic system is presented. At the end of the chapter in Section 5.6 different methods used to reconstruct the kinematic variables are discussed.

5.1 Track Reconstruction

Tracks of charged particles are reconstructed by combining their hit information from the inner tracking detectors, taking into account the errors on the hit measurements, the detector material distribution and the effects of multiple scattering. The track reconstruction allows the momentum and charge of a particle to be measured. Furthermore the energy loss, $\mathrm{d}E/\mathrm{d}x$, along the track provides a tool for particle identification.

Depending on the information used for track reconstruction, three different track types can be found at ZEUS:

- CTD only tracks: this track type uses only CTD information and has been used mainly to analyse the HERA I dataset;
- Regular tracks: this track type is available for the HERA II running period. It combines both CTD and MVD information;
- ZTT tracks: these are the "regular tracks" in which the accuracy of the tracking (close to the vertex) has been improved using a Kalman filter technique [91].

All track types are described using a five parameter helix parametrisation [89, 90] (see Figure 5.2). This parametrisation was designed to describe particle tracks in the solenoidal magnetic field parallel to the beam axis. The first three parameters represent a circle in the X–Y plane and the latter two the location and pitch in Z. The parameters are:

1. ϕ_H: the azimuthal angle of the direction vector of the helix at the point-of-closest approach;
2. Q/R: ratio of charge Q to local radius R;
3. $Q \cdot D_H$: D_H connects the helix to the reference point $(X_{ref}, Y_{ref}) = (0,0)$;
4. Z_H: the Z coordinate of the reference point;
5. $\cot\theta = \tan(\pi/2 - \theta)$: the angle of dip in X–Y plane.

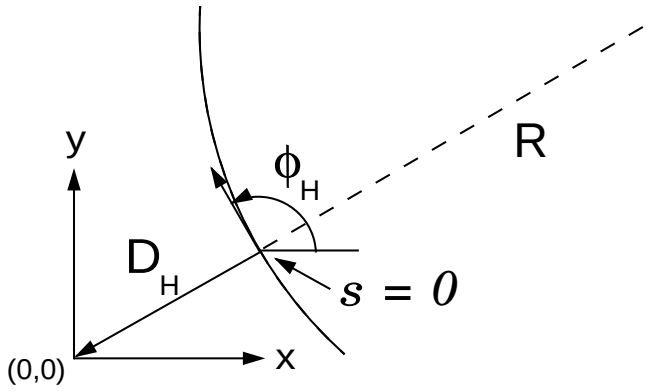


Figure 5.2: Helix in the X–Y plane for a positively charged track.

Since for this analysis the information coming from the MVD is essential to study the decay of particles containing heavy quarks, the type of tracking used is ZTT tracking. The offline reconstruction of these tracks is performed in two distinct stages. The first stage, pattern recognition, is carried out by the VCTRACK package [90]. This consists of a multi-step algorithm which combines information from different tracking detectors. In the first step, starting from the outermost super-layer of the CTD, each track candidate begins as a track seed consisting of three CTD hits from an axial super-layer. In the second step, these seed hits together with all other hits picked up in the same super-layer, called CTD segments, are assigned to initial trajectories. The next step of the algorithm combines matching track segments in the CTD with clusters of hits in the MVD. This procedure stores not only those tracks which have combined hits from the CTD and MVD, but also those which have hits in only one of the detectors. Depending on whether these tracks have hits in the CTD or MVD, they are called CTD-only and MVD-only tracks, respectively.

After the pattern recognition, the information assigned to a track by VCTRACK is passed to a track-fitting package known as the rigorous track fit (RTFIT). This track fitting program is based on Kalman filter techniques. It improves the accuracy of the tracking by taking into account effects like inhomogeneous magnetic field, multiple scattering and ionisation energy loss and then determines the track helix parameters [89, 90] and their covariances.

5.2 Vertex Reconstruction

Once the tracks have been found, the primary vertex, as well as the secondary vertices originating from the decay of long-lived particles or the interaction of particles with the material of the detector, can be reconstructed. Analogous to the track reconstruction, the vertex reconstruction consists of two main stages: vertex finding and fitting. Vertex finding involves the identification of the tracks belonging to the same decay vertex while the vertex

fitting implies the estimation of the vertex position as well as the track parameters at the vertex.

The primary vertex finding is done using the VCTRACK package. In this algorithm, it is initially assumed that a primary vertex should be found along the proton beamline; this serves as a soft constraint on the vertex position. The track pairs compatible with both a common vertex and the soft constraint are combined with other track pairs and a vertex is chosen based on the overall χ^2 of the best combination. Once the primary vertex and associated tracks have been identified, a deterministic annealing filter (DAF) [92] is applied to remove outliers before the final determination of the vertex [93].

Secondary vertices are reconstructed using the user vertex utility (VXU) algorithm. In principle the vertex fit performed by this package is same as that of the VCTRACK package for the primary vertex, except that in this case the beamline constraint is excluded. Another difference is that, in contrast to the VCTRACK package, VXU allows the user to select the tracks used as input into the algorithm. Only selected tracks satisfying the criteria: $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$, at least four hits in the MVD and passing at least three CTD super-layers are fitted to a vertex. Moreover tracks which lead to a bad χ^2 of the vertex, are discarded. If the algorithm finds a vertex, the fit procedure outputs three vertex coordinates, the covariance matrix and the χ^2.

5.2.1 Beam Spot

One of the discriminating variables used in this analysis is decay length of particles (see Section 7.2.3). Its measurement depends directly on the position of the primary and secondary vertices, hence, an accurate determination of the primary vertex is essential. The spread of the primary vertex in the transverse plane X–Y is smaller than its reconstruction resolution, so a high precision on its position can be achieved by replacing its value in each event with its average. Technically this average is known as the beam spot. It was calculated averaging the primary vertex position every $\sim$2000 good events, after applying some quality cuts on the tracks participating in the fit. A detailed description of the cuts and the procedure can be found in [94]. Typical values for the achieved resolution are $\sim 180\,\mu$m in X and $\sim 170\,\mu$m in Y.

5.3 Reconstruction of the Scattered Electron

The identification of the scattered lepton[1] is crucial to select neutral current (NC) DIS events. The presence of the scattered electron distinguishes NC events from other types of event, like photoproduction (where the e escapes through the beam hole) or charged

[1]The scattered lepton can either be an electron or a positron depending on whether electron or positron beam was used. Hereafter unless explicitly stated the scattered lepton (electron or positron) is referred to as scattered electron.

current (where instead of an e, an invisible neutrino appears in the final state). The accurate determination of the kinematics depends on the precise reconstruction of the scattered electron. In this section, the electron finding algorithm and measurement of the angle and energy of the scattered electron is presented.

5.3.1 Scattered Electron Identification

There are a number of so-called *electron finder* algorithms that are used within the ZEUS analysis framework. Each of these algorithms was developed for a special purpose, e.g. finding electrons in the RCAL or finding electrons in high Q^2 events. The algorithm used in this analysis is SINISTRA [95, 96] which is based on neural network techniques. This algorithm was trained on low Q^2 NC data and MC simulations to optimise its efficiency for differentiating between electromagnetic and hadronic showers.

A scattered electron creates an electromagnetic shower in the CAL, where most of its energy would be measured in the EMC cells with a small energy leakage to the HAC cells. Starting from a cell with a maximum energy deposit as a seed point, neighbouring cells above a certain threshold are combined to make so-called *cell islands*. A schematic view of the island clustering algorithm is depicted in Figure 5.3.

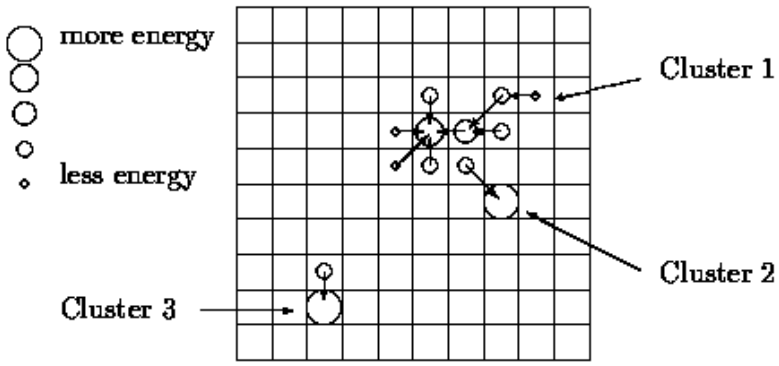


Figure 5.3: Schematic view of the island clustering algorithm in the ZEUS calorimeter.

Once the CAL energy deposits are clustered into islands, the finding of candidates for the scattered electron is done by SINISTRA. It evaluates the shower profile using a neural network based on calorimeter information. The output variable given by the algorithm is known as probability, P_{sira}, which can have a value from 0 (hadron-like) to 1 (electron-like). The probability for the low energy scattered electrons ($E < 10$ GeV) tends to be significantly smaller than 1, since for lower energy showers it gets harder to distinguish whether it comes from an electron or a hadron. $\pi^{\pm}$ contaminates the electron energy distribution at low values. In addition photons may fake an electromagnetic cluster, which is then identified as an electron.

5.3.2 Angle Reconstruction

After the scattered electron identification, the scattering angle θ_e is measured using the electron position in the calorimeter combined with the vertex position. The electron position is measured from the calorimeter cells associated to the electron candidate. The details of the position measurement can be found in [97]. The achieved resolution is $\sim 1\,\text{cm}$ in X and Y. Whenever the electron trajectory lies within the acceptance region of the CTD and SRTD, they are used to improve the resolution of its position measurement.

The low Q^2 region, corresponding to $\theta_e > 162°$, is covered by the SRTD detector whose fine segmentation (cf. Section 4.2.4) allows a resolution almost 3 times more precise than the one of the calorimeter. In the region of high Q^2 values ($\theta_e < 157°$), the position reconstruction is improved by requiring the association of the CAL cluster with a good CTD track. This track must fulfil some conditions: it must pass at least three CTD super-layers, must have a distance of closest approach (DCA) to the CAL cluster less than 10 cm and its momentum P, must be greater than 5 GeV. In the other angular range $157° < \theta_e < 162°$ only calorimeter position information is used.

5.3.3 Energy Reconstruction

In this analysis, several cuts based on electron energy, like cuts on $E'_e, E - P_z, E_t, P_t$ (see Chapter 6) are applied to select DIS events, so the level of agreement of the simulation of the energy with the real measurement may have a significant impact on the final sample. Moreover, the scattered electron energy can be used to reconstruct the kinematic variables, therefore its accurate measurement is essential.

The measured electron energy can deviate from its original value because of energy losses in the inactive material, non-uniformity of the detector and inadequate understanding of the individual response of each calorimeter cell (see Figure 5.4). Detailed studies [98, 99, 100] have been done in order to extract correction factors to overcome these effects. Here only a very short description of these corrections is given.

- **RCAL cell-by-cell energy scale correction:**
  The energy scale in data is corrected on a cell-by-cell basis. Electrons detected well within each cell were used to extract its scale factors.

- **Dead material correction:**
  A scattered electron produced at the interaction point passes through material: for example beam-pipe, tracking detectors, magnets, cables, etc. before reaching the CAL, and thus a fraction of its energy gets dissipated. To compensate for this loss of energy, the dead material correction is applied. This correction depends on the calorimeter energy, the amount of inactive material, the position of the event vertex and the scattered angle.

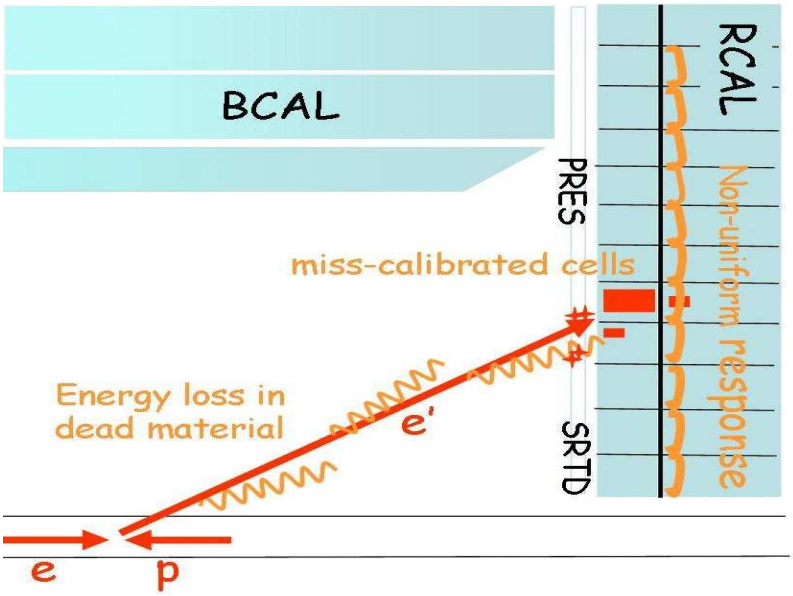


Figure 5.4: Shown are different effects which cause the deviation of energy from its original value (Courtesy of [98]).

- **Non-uniformity correction:**
  The CAL has gaps between the cells, towers and modules. These gaps cause a non-uniform response of the calorimeter and lead to a deviation of the measured energy at each edge. Such an effect in the RCAL is corrected by the product of an X-dependent correction factor and a Y-dependent factor which are derived separately for the data and MC. The X-dependent (Y-dependent) parameters are different for the upper and lower halves (left and right halves).

The correction parameters were mainly determined using two different methods, one of which used so-called *kinematic peak* events, which are characterised by the energy of the scattered electron being close to the beam energy, and the other method used the energy derived from the double angle method (cf. Section 5.6.3).

5.4 Reconstruction of the Hadronic System

Reconstruction of the hadronic system is done by combining the information of both neutral and charged particles. The energy of both types of particles is measured by the CAL whereas the momentum of charged particles is measured using the CTD. The energy resolution of the CAL is $\sigma(E)/E \propto 1/\sqrt{E}$, which improves for higher particle energies, while the resolution of the tracking system, given by $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}} \sim a.p_{\mathrm{T}} \oplus b \oplus c/p_{\mathrm{T}}$, gives a better energy estimate for lower particle energies (see Figure 5.5). For neutral particles only CAL

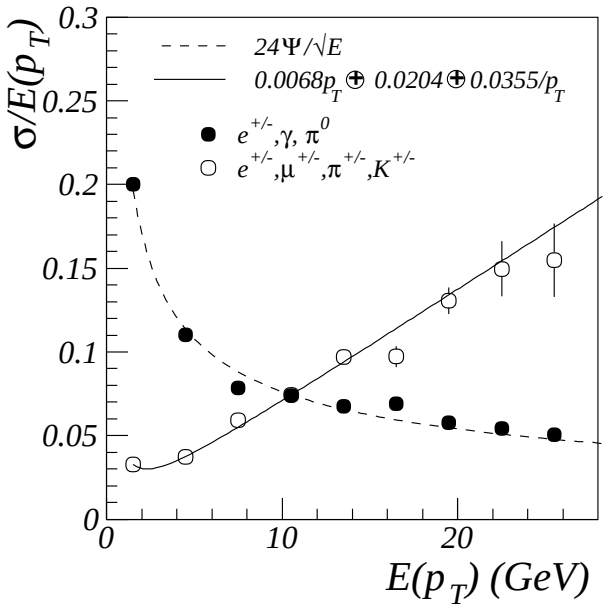


Figure 5.5: Resolution from single particle MC simulation [101]. Shown here is the track transverse momentum resolution in the CTD (open dots) and the CAL energy resolution (filled dots).

information is used, while for charged particles, to achieve the best resolution tracking information for particles having energies below 10GeV to 15GeV is combined with calorimeter information for particles with higher energies. The combination of these two pieces of information gives the energy-momentum four-vectors known as energy flow objects (EFOs), representing the oriented energy deposition of particles travelling in the detector. In the following, the methods to combine both measurements and to apply energy corrections due to inactive material are presented.

5.4.1 Reconstruction of EFOs

The reconstruction procedure of EFOs consists of several steps, including removal of the scattered electron, clustering of the calorimeter energy deposits and then combining the CAL and CTD information. In a first step, the calorimeter cells assigned to the scattered electron are removed and then the calorimeter energy deposits are clustered using a two stage algorithm [102].

In the first stage, the cell energy measurement is clustered separately in each part of the calorimeter (FCAL, BCAL, RCAL) and for each calorimeter layer (EMC, HAC1, HAC2) into cell islands (see Figure 5.3). In the second stage of the clustering procedure, the cell islands belonging to the shower of a single particle or a jet of particles are grouped together to form a so-called *cone island* (3-dimensional objects). The matching starts from

the outermost hadronic layer of the CAL and proceeds inward by calculating the angular separation to neighbouring islands in polar and azimuthal angles θ, ϕ. The separation is transformed to a matching probability according to a single pion MC simulation [102]. Afterwards all cell islands with a probability above a threshold are grouped into a cone island. The cone island centre is calculated using the logarithmic centre-of-gravity of the shower[2].

In the next step the CAL and CTD information is combined. First the tracks are extrapolated to the CAL inner surface taking into account the magnetic field geometry and then further into the CAL by a linear approximation using the track momentum vector at the CAL surface. These tracks must fulfil certain requirements: they must be vertex-fitted tracks with hits in at least four super-layers and they must have a transverse momentum p_{T} between 0.1 and 20 GeV. If the tracks have hits in five super-layers, the upper limit on p_{T} is increased to 25 GeV. The matching of a track to a cone island is then done using the distance-of-closest (DCA) approach method. A match is assigned between the extrapolated track and the cone island centre (see Figure 5.6) if the DCA is smaller than 20 cm or if the track lies inside the area of the island. In the case where one track is matched to one cone island, the EFO momentum is taken from either the calorimeter cluster or the matched

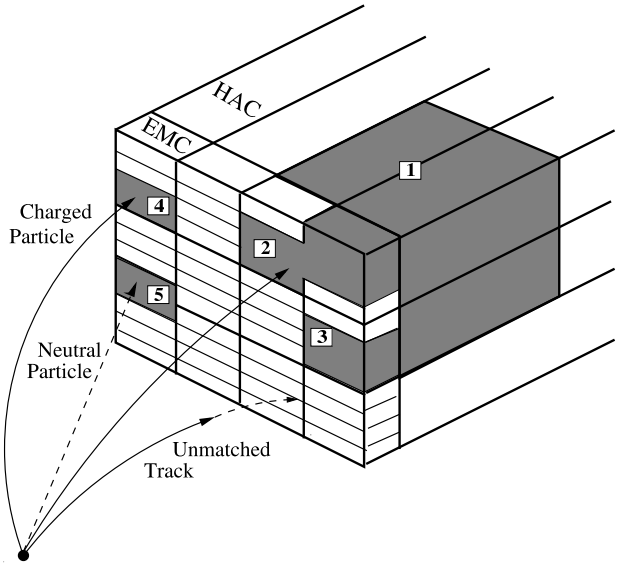


Figure 5.6: Reconstruction of the energy flow objects. EMC cell islands 2 and 3 are joined with HAC cell island 1 to from a cone island. In the next step island 4 and the cone islands (1,2,3) are matched with tracks (charged particles). Island 5 has no matched track (neutral particle). An unmatched track corresponding to a low momentum particle is also shown. Taken from [102]

[2]Using logarithmic weights instead of linear weights takes into account the exponential fall-off of the shower energy distribution from the shower maximum and avoids systematic biases due to varying cells solid-angle as seen by the vertex.

track, depending on the resolution of the track momentum and the CAL cluster energy resolution. Tracks without an associated island are treated as charged particles and the CTD information is used under the assumption that the particle is a pion. Cone islands not matched to a track are treated as neutral particles and the CAL energy measurement is used. Cone islands with more than three associated tracks are treated as jets and the CAL information is used. In more complicated types of EFOs, where one track is associated to two or three islands or two tracks to one or two islands, the sum of the island energies or track momenta is used.

The reconstructed EFOs after applying different corrections (cf. Section 5.4.2) are used for jet finding. In this analysis they are also used as an input to select the semileptonic electron candidates (see Chapter 6).

5.4.2 EFOs and Cone Island Energy Corrections

The reconstruction of the EFOs using CAL islands has shown some discrepancies between data and MC simulations [103, 104]. The following corrections have therefore been applied to EFOs:

- **Inactive material correction:**
  The presence of inactive or dead material in the detector causes energy loss and can underestimate the energy of the particles passing through it. This kind of energy loss is not sufficiently simulated and therefore an offline correction is needed. A detailed

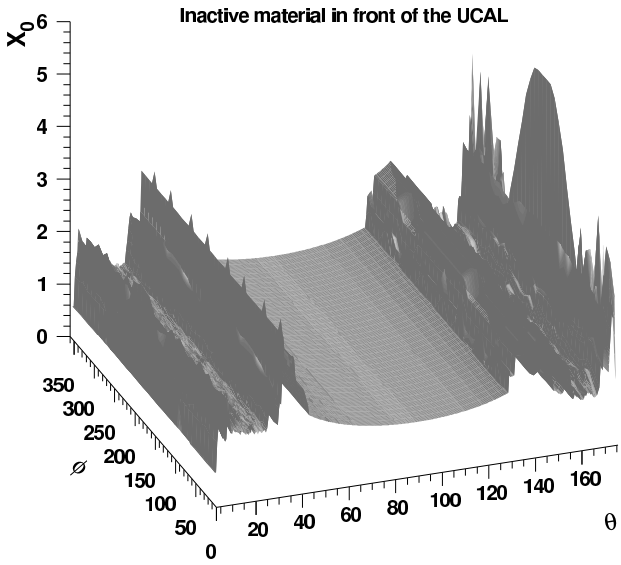


Figure 5.7: Map of the dead material distribution between the interaction region and the CAL in units of the radiation length, X_0, in the θ–ϕ plane.

dead material map (see Figure 5.7) has been used to correct the contribution to the energy loss mainly from the beam-pipe, the tracking detectors and the solenoid. The amount of the dead material between the interaction point and the calorimeter varies between 1 and 3 radiation length, X_0, depending on ϕ and θ. The energy loss is more significant for low energy particles and corrections are applied depending on the polar and azimuthal angles and the cone island energy.

- **Calorimeter geometry:**
  The regions of the cracks between the calorimeter sections are not simulated perfectly and are corrected offline.

- **Corrections for low momentum hadrons:**
  Hadrons having momenta below 1 GeV lose their energy mainly through ionisation without hadronic interactions. In this case the calorimeter compensation effect is reduced ($e/h \sim 0.6$) and causes an overestimation of the energy which has to be corrected.

5.5 Jet Reconstruction

One of the main steps in the reconstruction of the events used in this analysis is the identification of the partons from the hard interaction. The partons can not be measured directly in the detector, since colour confinement prevents them from appearing as free particles. Instead, colourless hadrons are produced through the fragmentation and hadronisation process. These hadrons appear as a collimated flow of objects called jets. The measurement of the energy and position of the hadrons in the detector makes it possible to measure quark and gluons indirectly using the strong correlation between the partons of the hard interaction and the reconstructed jets in the detector. In order to draw conclusions between theoretical predictions and experimental results, well-defined reconstruction of jets is essential. Different types of algorithm have been developed in order to reconstruct jets. To be suitable for *ep*-collisions, the jet algorithm has to fulfil following requirements:

- **Collinear safety:** the results should be independent of one parton splitting into two partons moving collinearly;

- **Infrared safety:** the results should be independent of the emission of very low energy particles;

- **Beam remnant:** correct treatment of beam remnants especially of the proton and if present of the photon;

- **Lorentz invariance:** independence from longitudinal Lorentz boosts.

Additionally, in heavy quark production, the jet algorithm should distinguish between the remnant of the proton and the jets from the hard process.

In order to meet all these requirements, in this analysis jets were reconstructed using a k_{T}-type clustering algorithm [105, 106] based on EFOs (cf. Section 5.4.1). This algorithm interactively clusters the EFOs into jets using the following steps:

1. resolution variables d_k and d_l are calculated for all pairs of EFOs, e_k and e_l:

$$d_k = (p_{\mathrm{T}}^k)^2, \tag{5.2}$$
$$d_{kl} = \min((p_{\mathrm{T}}^k)^2,(p_{\mathrm{T}}^l)^2)\Delta R_{kl}^2, \tag{5.3}$$

   where $\Delta R_{kl}^2 = \Delta\eta_{kl}^2 + \Delta\phi_{kl}^2$.

2. the smallest resolution variable from the list of d_k, d_{kl} is determined: $(d_{\min} = \min(d_k,d_{kl}))$

   - if $d_{\min} = d_{kl}$, the two EFOs, e_k and e_l are merged into a single new pseudo-particle using the recombination scheme where hadrons are treated as massive particles and their four-momenta are added as: $p_{kl} = p_k + p_l$. The resolution variable is given by:

$$d_{kl} = (p_{T,k} + p_{T,l})^2, \tag{5.4}$$
$$d_{(kl)i} = \min((p_{T,k} + p_{T,l})^2,(p_{T,i})^2)\Delta R_{(kl)i}^2 \tag{5.5}$$

   - if $d_{\min} = d_k$, the EFO e_k is removed from the EFO list and is included in the remnant jet.

3. this procedure continues until all particles and pseudo-particles not included in the remnant jet have resolution variables larger than the hard scale, d_{cut}, where $\Lambda^2 \ll d_{\mathrm{cut}} \ll s$, where Λ is the QCD scale and $\sqrt{s}$ is the centre-of-mass energy.

This procedure results in hard final-state jets and remnant jets. The use of the E recombination scheme in the heavy flavour studies has proven to be more useful compared with other recombination schemes [105, 106], as the massive approach yields a good agreement of the measurements and theoretical predictions of jets from beauty quarks [107].

In this analysis jets are preselected with a loose E_{T} cut of 2.5 GeV and $|\eta_{\mathrm{jet}}| < 2.5$. In the final selection a slightly harder cut ($p_{\mathrm{T}} > 2.5$ GeV) is applied. The lower p_{T} cut is introduced to ensure the correspondence of a jet to an initial quark, while the η cut helps to remove remnant jets and restricts the measurement to a region of the detector which is well understood.

5.6 Reconstruction of the Kinematic Variables

After reconstruction of the scattered electron and the hadronic final state, four measured quantities are available. These are the energy and the polar angle of the scattered electron, (E'_e,θ_e), and the total transverse momentum, $P_{t,\mathrm{had}}$, and δ_{had} of the hadronic system, where

$$P_{t,\mathrm{had}} = \sqrt{(\sum_{\mathrm{had}}(P_{x,\mathrm{had}})^2) + (\sum_{\mathrm{had}}(P_{x,\mathrm{had}})^2)}, \tag{5.6}$$

$$\delta_{\mathrm{had}} = \sum_{\mathrm{had}}(E_{\mathrm{had}} - P_{z,\mathrm{had}}). \tag{5.7}$$

The sum runs over all particles in the final state except the scattered electron. The hadronic angle γ_{had} can be reconstructed from δ_{had} and $P_{t,\mathrm{had}}$ as:

$$\cos\gamma_{\mathrm{had}} = \frac{P^2_{t,\mathrm{had}} - \delta^2_{\mathrm{had}}}{P^2_{t,\mathrm{had}} + \delta^2_{\mathrm{had}}}, \tag{5.8}$$

which corresponds to the angle of the scattered quark in the naive quark parton model.

The variables $E'_e, \theta_e, \gamma_{\mathrm{had}}, P_{t,\mathrm{had}}$ can now be used to reconstruct the kinematic variables x, y and Q^2. Only two of these kinematic variables are independent due to the relation $Q^2 = sxy$, so only two of the four measured variables are needed. There are several different methods to reconstruct kinematic variables depending on which of the two variables are used. Three of the methods are described here.

5.6.1 Electron Method

The so-called *electron method* (El) uses only the input from the scattered electron i.e. its energy and polar angle. The kinematic variables Q^2, y and x are expressed as:

$$Q^2_{\mathrm{el}} = 2E_eE'_e(1+\cos\theta_e), \tag{5.9}$$

$$y_{\mathrm{el}} = 1 - \frac{E'_e}{2E_e}(1-\cos\theta_e), \tag{5.10}$$

$$x_{\mathrm{el}} = \frac{Q^2_{\mathrm{el}}}{s \cdot y_{\mathrm{el}}}. \tag{5.11}$$

where E_e is the energy of the incoming electron, E'_e, θ_e are energy and angle of the scattered electron and $\sqrt{s}$ is the centre-of-mass energy .

The reconstructed values are sensitive to the electron energy scale. Equations 5.9 and 5.10 show that low Q^2 corresponds to large θ_e in the ZEUS coordinate system i.e. low scattering angle and high y corresponds to low E'_e. This method offers good resolution in the high y region, while at low y region its resolution diverges due to the $1/y$ dependence in the error propagation formula of the relative error [108]:

$$\frac{\delta y_{\mathrm{el}}}{y_{\mathrm{el}}} = \left(1 - \frac{1}{y_{\mathrm{el}}}\right)\frac{\delta E'_e}{E'_e} \oplus \left(\frac{1}{y_{\mathrm{el}}} - 1\right)\cot\left(\frac{\theta_e}{2}\right)\delta\theta_e. \tag{5.12}$$

5.6.2 Jacquet-Blondel Method

The *Jacquet-Blondel* (JB) method [109] relies exclusively on the variables of the hadronic final state, namely the transverse momentum, $P_{t,\mathrm{had}}$, and δ_{had}. The kinematic variables are expressed as:

$$Q^2_{\mathrm{jb}} = \frac{P^2_{t,\mathrm{had}}}{1-y_{\mathrm{jb}}}, \tag{5.13}$$

$$y_{\mathrm{jb}} = \frac{\delta_{\mathrm{had}}}{2E_e}, \tag{5.14}$$

$$x_{\mathrm{jb}} = \frac{Q^2_{\mathrm{jb}}}{s \cdot y_{\mathrm{jb}}}. \tag{5.15}$$

This method gives a better resolution for low y than the electron method, because the resolution does not depend on terms proportional to $1/y$:

$$\frac{\delta y_{\mathrm{jb}}}{y_{\mathrm{jb}}} = \frac{\delta E_e}{E_e} \oplus \cot\left(\frac{\theta_{\mathrm{jb}}}{2}\right) \delta\theta_{\mathrm{jb}} \tag{5.16}$$

This reconstruction method is mainly used in CC analyses and in photoproduction, where the scattered electron is absent or undetected, being scattered at low angle into the beam-pipe.

5.6.3 Double Angle Method

The *double angle* (DA) method uses the angles of the scattered electron, θ_e, and the hadronic system, γ_{had} [110]. The kinematic quantities are:

$$Q^2_{\mathrm{da}} = 4E_e^2 \frac{\sin\gamma_{\mathrm{had}}(1+\cos\theta_e)}{\sin\gamma_{\mathrm{had}} + \sin\theta_e - \sin(\gamma_{\mathrm{had}}+\theta_e)}, \tag{5.17}$$

$$y_{\mathrm{da}} = \frac{\sin\theta_e(1-\cos\gamma_{\mathrm{had}})}{\sin\gamma_{\mathrm{had}} + \sin\theta_e - \sin(\gamma_{\mathrm{had}}+\theta_e)}, \tag{5.18}$$

$$x_{\mathrm{da}} = \frac{E_e}{E_p} \frac{\sin\gamma_{\mathrm{had}} + \sin\theta_e + \sin(\gamma_{\mathrm{had}}+\theta_e)}{\sin\gamma_{\mathrm{had}} + \sin\theta_e - \sin(\gamma_{\mathrm{had}}+\theta_e)}. \tag{5.19}$$

This reconstruction method minimises the sensitivity to the absolute energy scale but the resolution of all the reconstructed variables depends strongly on the determination of the angle, γ_{had}, of the hadronic system.

5.6.4 Comparison of Reconstruction Methods

In order to judge which reconstruction method should be adopted, bias and resolution of the kinematic variables for each method were compared using the DIS MC in the x and

Q^2 plane. The bias and resolutions were determined using the mean and width of the relative difference between the reconstructed and true value, which was calculated, e.g. of Q^2, using:

$$\frac{Q^2_{\text{rec}} - Q^2_{\text{gen}}}{Q^2_{\text{gen}}}, \tag{5.20}$$

where the subscript 'rec' means reconstructed value, and 'gen' means true value. Figures 5.8, 5.9 and 5.10 show the comparison of the resolutions of the three reconstruction methods for Q^2, x and y, respectively. The electron method, as already mentioned, works well in low x - low Q^2 region ($\Leftrightarrow$ high y), but introduces large bias and resolution in the high x - low Q^2 region ($\Leftrightarrow$ low y), especially for x reconstruction. The JB method does not have large bias and works well for x and y reconstruction in the low y region, however the reconstruction of Q^2 is very poor. The DA method generally performs better than the other two methods over most of the kinematic plane and hence was adopted to reconstruct the kinematic variables.

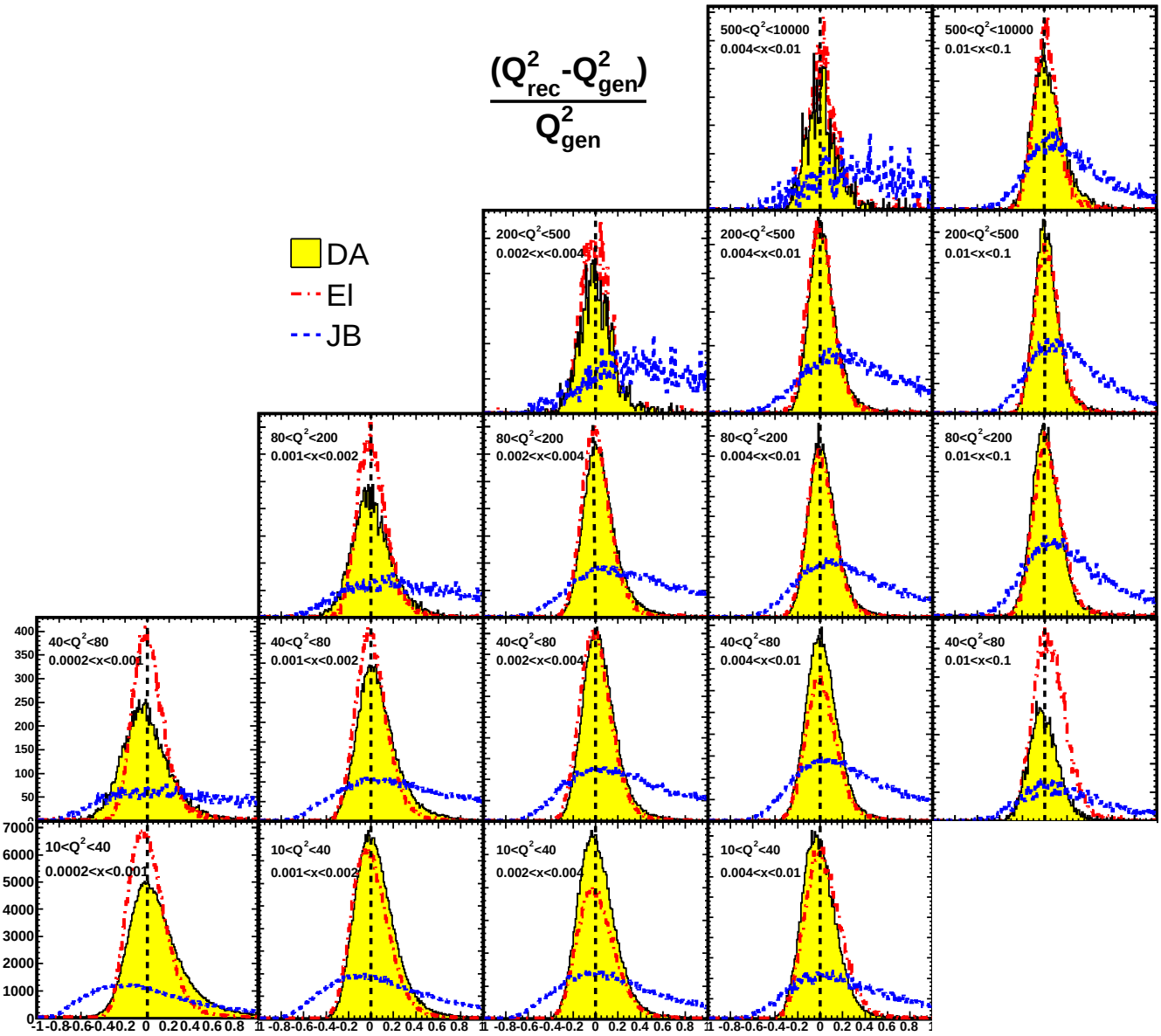


Figure 5.8: Resolution of Q^2, $(Q^2_{\text{rec}} - Q^2_{\text{gen}})/Q^2_{\text{gen}}$, for each reconstruction method in the x–Q^2 plane. The filled yellow histogram corresponds to the DA, red dashed line is the El and blue dashed line is the JB method.

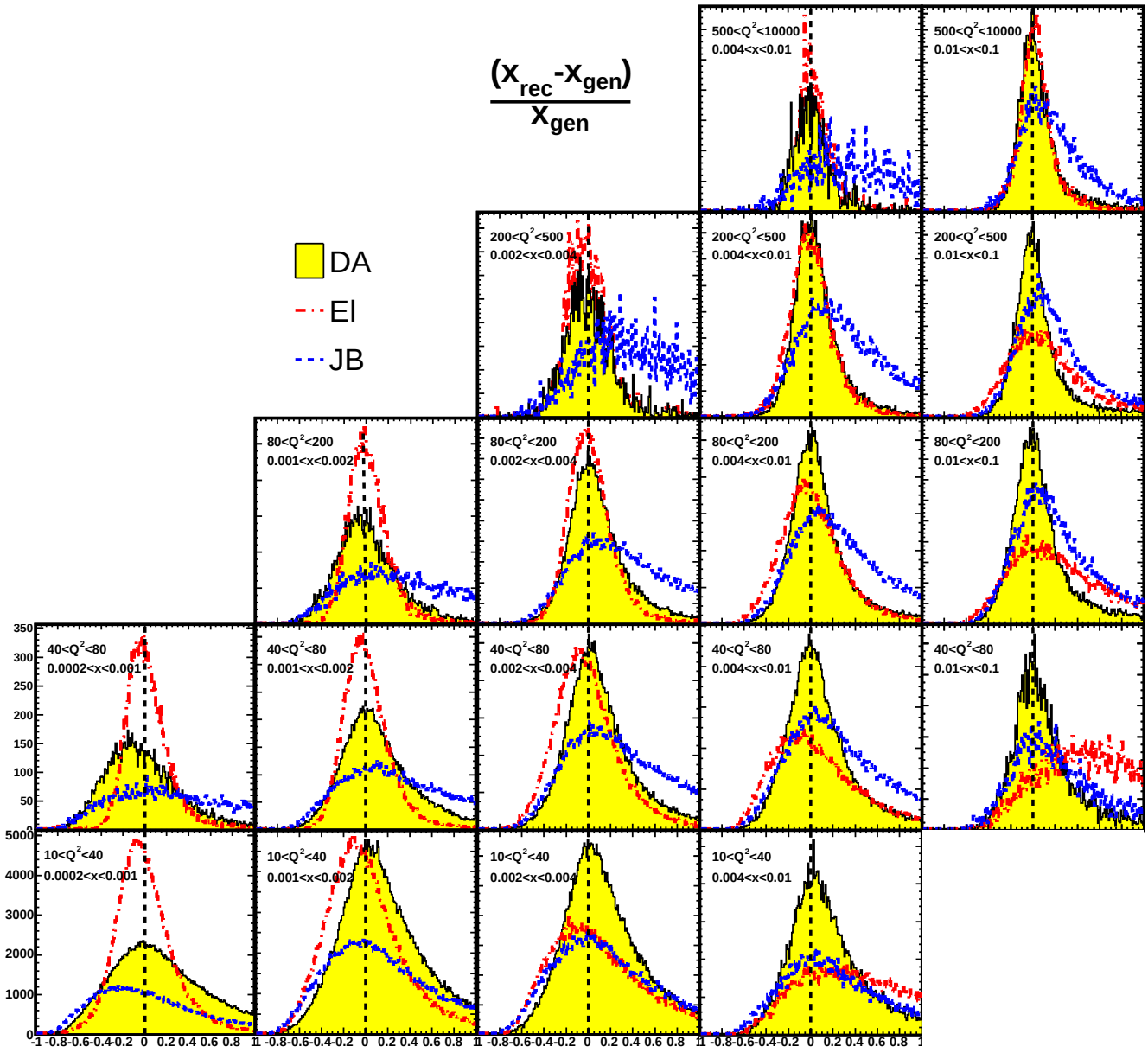


Figure 5.9: Resolution of x, $(x_{\text{rec}} - x_{\text{gen}})/x_{\text{gen}}$, for each reconstruction method in the x–Q^2 plane. The filled yellow histogram corresponds to the DA, red dashed line is the El and blue dashed line is the JB method.

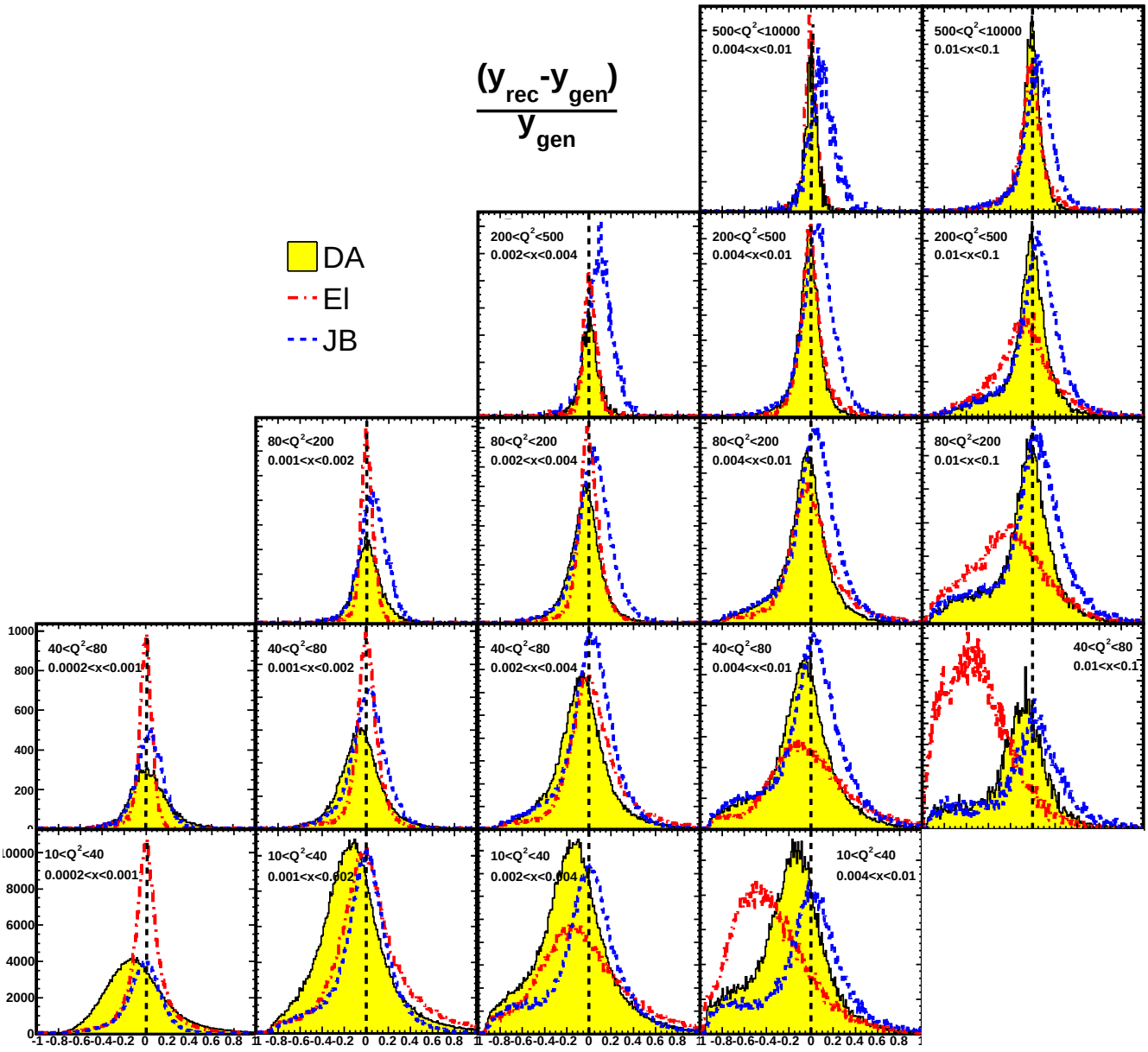


Figure 5.10: Resolution of y, $(y_{\text{rec}} - y_{\text{gen}})/y_{\text{gen}}$, for each reconstruction method in the x–Q^2 plane. The filled yellow histogram corresponds to the DA, red dashed line is the El and blue dashed line is the JB method.

Chapter 6

Event and Candidate Selection

To analyse beauty production in DIS using semileptonic electron decays, a DIS plus electron (from b–quark decay) event selection is applied to data and MC events. As a first step DIS events are identified by the detection of a scattered electron in the calorimeter. Since several background sources can also make an electromagnetic energy deposition in the CAL which may be mis-identified as a scattered electron, several cleaning cuts are applied to remove such events and to ensure a good reconstruction of the scattered electron. After obtaining a clean DIS sample, further selection cuts are applied to enrich the b-quark component by finding candidates for semileptonic decays to electrons.

In this chapter, the used data and MC samples are listed in Section 6.1 followed by the details of different cuts for event and candidate selection in the subsequent sections. A summary of the complete set of applied cuts and a set of control plots showing the comparison between the data and the MC events after the final selection is given in Section 6.5.

6.1 Data and MC Samples

For this analysis data collected by the ZEUS detector during the HERA II running period (2004-2007) were used. During this running period, protons were brought into collision with electrons or positrons at a centre-of-mass energy of 318 GeV. The runs for which the detector was not in an appropriate state, because some essential parts of it (CAL, CTD or MVD) were not working properly, were excluded[1]. A part of the dataset from the beginning of 2004 was also discarded, since the triggers were not well understood. The integrated luminosities from the ZEUS luminosity measurement for each running period are summarised in Table 6.1. In total an integrated luminosity of 363 pb^{-1} was used. This dataset was the first version of the Grand Reprocessed (GR) data processed with the

[1]This was done with the help of the EVTake [111], MVDTake databases.

Running period	Hadron	Lepton	$\sqrt{s_{ep}}$ (GeV)	$\mathcal{L}$ (pb^{-1})
2004	proton	positron	318	32
2005	proton	electron	318	132
2006	proton	electron	318	55
2006-2007	proton	positron	318	142

Table 6.1: Used datasets and their corresponding luminosities for different running periods.

software version which included latest corrections and detector alignment[2].

Monte Carlo samples mainly needed for un-folding the cross sections in the end (cf. Chapter 11) were simulated with the RAPGAP and DJANGOH MC generators, whose basic features were already described in Section 2.7.2. The MC simulation was done separately for the beauty and charm events. An inclusive DIS sample was used for light flavour background events after rejecting beauty and charm production processes. A summary of the generated processes, their cross sections and corresponding luminosities for the different running periods is given in Table 6.2.

		$\mathcal{L}$ (pb^{-1})			
Event type	σ(pb)	2004e^+	2005e^-	2006e^-	2006-2007e^+
b inclusive	907.7	925	2115	925	2555
c inclusive	35346	98	280	166	495
lf inclusive	319500	32	134	54	142

Table 6.2: Used MC sets and their corresponding luminosities for different running periods. The generated cross section predictions are also given.

To achieve a satisfactory statistical precision, a significant amount of MC events is needed. Since the cross section for beauty production is moderately low, it was possible to generate a large number of beauty events with a corresponding luminosity of 18 times the data luminosity. On the other hand, due to substantially higher production cross sections for charm and light flavour events, a similar amount of statistics was hard to generate mainly because of the non-negligible amount of time consumed for the simulation and the limitations of storage space. Therefore a compromise had to be made between a reasonable statistical precision and the available disk space. A charm sample of about 3 times data luminosity and a light flavour sample having the same luminosity as the data was produced.

[2]After the end of the data-taking, Grand Reprocessing was done where all datasets as well as the corresponding MC simulations were processed with exactly the same software making it possible to combine different datasets.

6.2 Online Event Selection

As described in Section 4.4, at ZEUS a three level online trigger system was operated to record the interesting physics events with high efficiency and to suppress the event rate to an acceptable level by rejecting background events. A suitable trigger pre-selection was applied to the data events to select DIS events which had a possible semileptonic decay of a heavy particle to an electron. Since the detector simulation takes into account the simulation of the trigger system, the same trigger pre-selection was also applied to the MC events. The trigger selection consists of at least one of the following third-level trigger slots (inclusive "or" of the six slots):

- SPP01: Inclusive medium Q^2 DIS
- SPP02: Inclusive low Q^2 DIS
- DIS03: Medium Q^2
- DIS04: High Q^2
- HFL17: Neutral current DIS
- HFL10: DIS semileptonic e

At the third-level trigger, slots consist of cuts on event quantities calculated on the third level and require combinations of first and second level trigger slots, more details are given in Appendix A.

6.3 Offline Event Selection

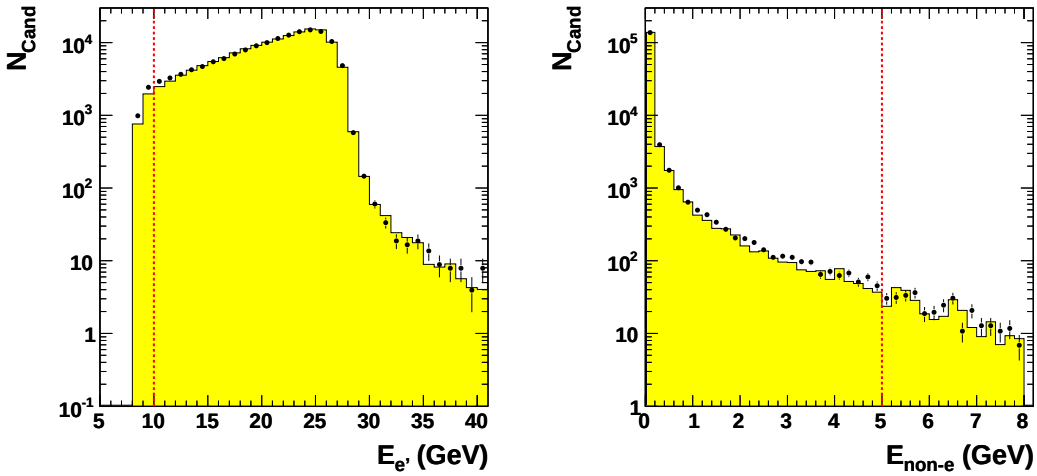
On the event sample that passed the online selection, further cuts were applied offline to select a clean DIS sample. The offline selection criteria were optimised such that they ensured the good reconstruction of the events and increased the purity of the sample. In the following the details of the selection requirement are presented. Some of the cuts which were applied on different variables are shown as superimposed lines on the distributions where the data is compared to the MC simulation. These distributions were made using the full selection criteria except for the cut under study. In all these plots the black points represent the data while the yellow shaded areas correspond to simulated events. The latter are the luminosity-weighted sum of all Monte Carlo samples and are normalised to the area of the data distribution.

6.3.1 Scattered Electron Requirement

The key component for the selection of DIS events is the presence of the scattered electron. The scattered electron candidates were required to have a good quality which also helped to reduce various background events.

- **Sinistra probability:** For the precise identification of the scattered electron, an electron candidate with the probability from Sinistra, $P_{\text{sira}} > 0.90$, was required.
- **Electron energy:** The corrected energy of the scattered electron, $E_{e'}$, (cf. Section 5.3.3) was required to be greater than 10 GeV. This requirement guarantees high electron efficiency and helps to remove photoproduction background coming from decays of low energy neutral pions into photons, 'faking' an electron signature in the CAL. The reconstructed scattered electron energy after applying all cuts except for $E_{e'} > 10$ GeV is shown in Figure 6.1(a). The superimposed line at 10 GeV represents the cut value.
- **Isolation energy:** Typically in a clean DIS event, the scattered electron is back-to-back to the hadronic system in the transverse plane and the energy deposited in the electromagnetic calorimeter is well isolated. On the other hand, in case of background events, the scattered electrons are found close to a hadronic shower in the CAL. In such events it becomes hard to determine which energy deposit belongs to the electron. To evaluate this, the electron isolation energy, $E_{\text{non-e}}$, is defined as an energy deposited in the CAL not associated with the scattered electron inside an $\eta - \phi$ cone of radius $R_{\text{cone}} = 0.8$ centred on the electron. The cone radius is defined by $R_{\text{cone}} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, where $\Delta\phi(\Delta\eta)$ is the difference of the azimuthal angle (pseudorapidity) between the scattered electron and CAL energy deposits. The isolation energy was required to be $E_{\text{non-e}} < 5$ GeV. The distribution of the isolation energy together with the threshold line at 5 GeV is shown in Figure 6.1(b).
- **Track matching:** If the candidate for the scattered electron was found to be within the acceptance region of the CTD[3], a track matched to the cluster was required. To match a track to the cluster, the distance of closest approach (DCA) between the calorimeter cluster and the end point of the extrapolated CTD track was required to be less than 10 cm. This requirement helps to reject events where a photon is wrongly associated with a track in the CTD. The electron track momentum, p_e^{trk}, was required to be greater than 3 GeV which increases the probability that the track belongs to the scattered electron and not to a low energy charged particle. Furthermore the associated track was required to pass at least 4 CTD superlayers. Also, as the energy measurements close to the calorimeter module edges can be unreliable, the events were rejected if in the BCAL, the electron track was extrapolated to be within a distance of 1.5 cm to a module edge (DME).

[3]If the polar angle of the candidate satisfies $0.6 < \theta_e < 2.5$, the electron is considered to be within the CTD acceptance [112].



(a) (b)

Figure 6.1: (a) Reconstructed energy, $E_{e'}$, and (b) isolation energy, $E_{\text{non-e}}$, of the scattered electron for both data (filled circles) and simulated events (shaded histograms) after all cuts except for the one depicted. The latter are the luminosity-weighted sum of all Monte Carlo samples and normalised to the data distributions. Vertical dashed red lines indicate cut thresholds.

6.3.2 Geometry cuts

In some regions of the calorimeter, the reconstruction of the scattered electron is poor and difficult to simulate well due to their geometrical complexity. The events in which a candidate was found in such regions were rejected to ensure good reconstruction of the scattered electrons. In the following such geometrical cuts based on the electron position are described.

- **Super-crack cuts:** The term 'super-cracks' refers to the edges of the BCAL next to the RCAL and FCAL. If the scattered electron comes to one of these regions, the energy deposit may be shared in both the BCAL and RCAL (or FCAL). The simulation of this is not precise and easily causes mis-reconstruction of events. Therefore events were rejected where the electron was found in the following regions: $-104 < Z_e < -98.5$ cm and $164 < Z_e < 174$ cm.

- **Chimney cut:** In the upper half of the RCAL central module of the ZEUS detector there was a region where EMC cells could not be mounted. Instead this region was occupied by pipes that carry helium to and from the superconducting solenoid. This area known as chimney was excluded by rejecting the events in the RCAL where the position of the candidate was found within: $|X_e| < 12$ cm and $Y_e > 80$ cm.

- **RCAL radius cut:** As the outer region of the RCAL was partially overlapped by the BCAL, the measurements from this region were found to be not well simulated.

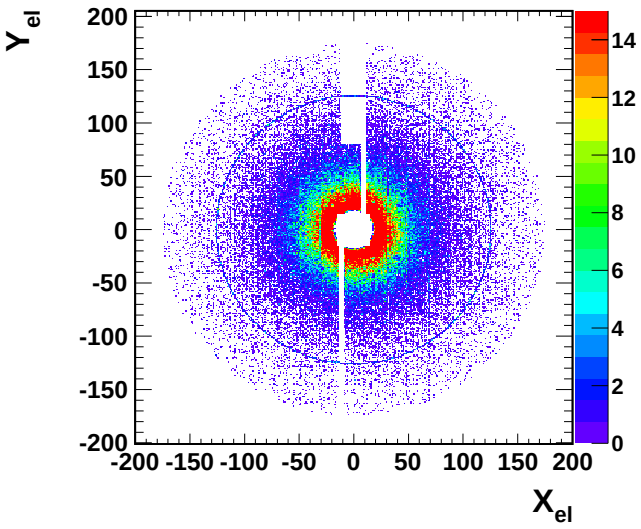


Figure 6.2: Scattered electron position distribution on the RCAL surface after the geometry cuts. The effect of RCAL radius cut, chimney cut and calorimeter gap cut is clearly visible.

Therefore events were rejected when an electron candidate was found in the RCAL with a position radius of, $R_e^{\mathrm{RCAL}} > 175$ cm.

- **Calorimeter gap cut:** To make it possible to open the ZEUS detector, the right and left halves of the RCAL could be separated during injection. The upper part of the middle module was attached to the negative side of the RCAL and the lower part was attached to the positive side. Since the simulation of this gap was not satisfactory, the events were rejected where the position of the electron candidate was found within this gap, defined by:
  $y_e > 0$ cm, $\quad 6.5 < x_e < 12$ cm
  $y_e < 0$ cm, $\quad -14 < x_e < -8.5$ cm.

- **Box cut:** There was a hole in the calorimeter for the beam pipe where the electron could not be reconstructed. Such events were rejected by requiring that the electron position in the RCAL must be outside the box defined by $|x_e| < 15$ cm and $|y_e| < 15$ cm.

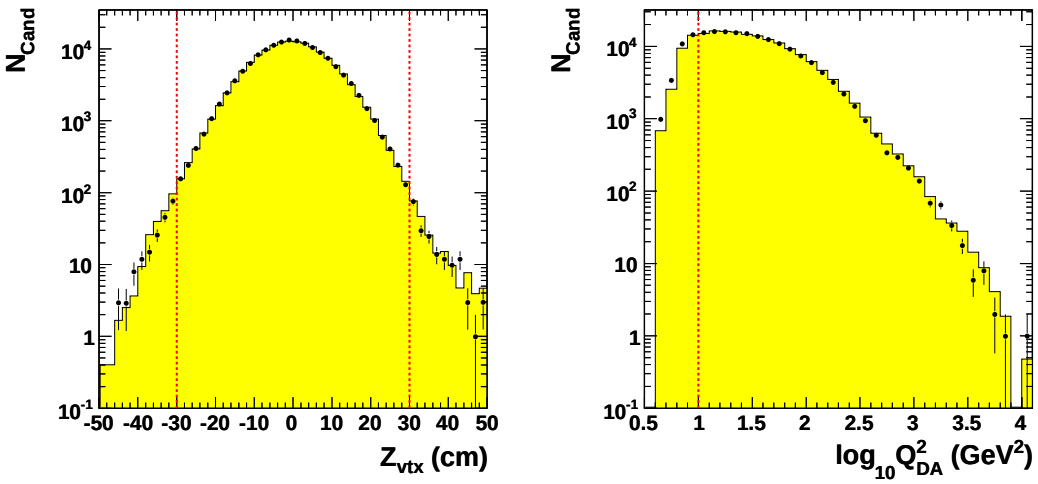
Figure 6.2 shows the electron position distribution on the RCAL surface after the geometry cuts. The effects of RCAL radius cut, chimney cut and calorimeter gap cut are clearly visible. The hole of radius ≈ 15 cm about the coordinate centre is due to the Q^2 cut described below, $Q^2 > 10\,\mathrm{GeV}^2$.

6.3.3 General cuts

- **Z_{vtx} cut:** The proton and electron beams collide at a nominal position of $Z = 0$ cm with a Gaussian spread, $\sigma \approx 11$ cm, resulting in ep events distributed around the nominal position. In order to remove the non-ep events which can occur uniformly

in Z, a cut on Z_{vtx} was required of $|Z_{\text{vtx}}| < 30\,\text{cm}$. The control distribution for Z_{vtx} after applying all other selection cuts except on Z_{vtx} is shown in Figure 6.3(a).

- **Q^2 cut:** The events were restricted to a Q^2 region greater than $10\,\text{GeV}^2$, where the kinematic variable Q^2 was reconstructed using the double angle method (cf. Section 5.6.3). This cut was applied since below this cut the acceptance dropped drastically. Also, for this Q^2, the scattered electrons are well measured in the central tracking detector and background from photoproduction can be neglected. The Q^2 distribution together with the threshold line at $10\,\text{GeV}^2$ is depicted in Figure 6.3(b).



(a) (b)

Figure 6.3: (a) Z component of the primary vertex, Z_{vtx}, and (b) photon virtuality reconstructed using the double angle method, Q^2, for both data (filled circles) and simulated events (shaded histograms) after all cuts except for the one depicted. Vertical dashed red lines indicate the cuts.

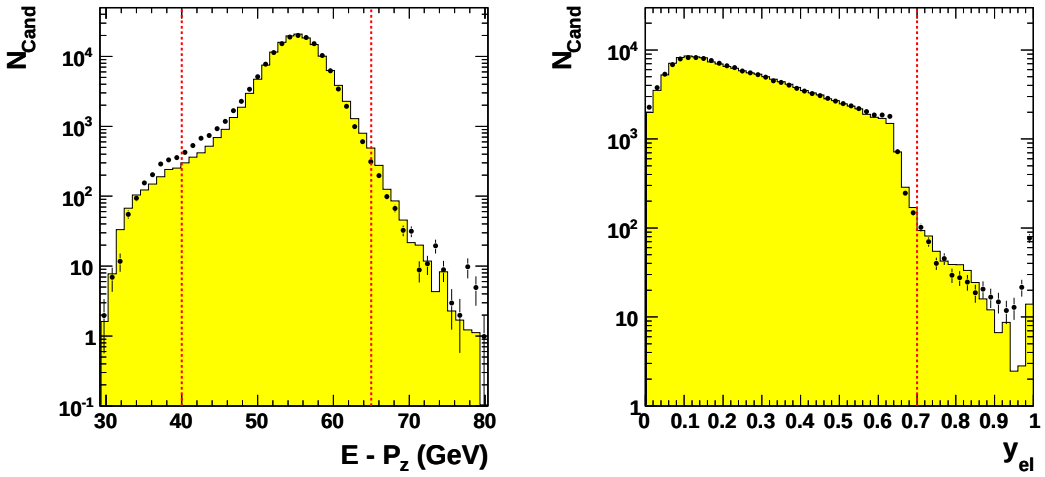
6.3.4 Background Suppression

Some cuts on different variables were applied which helped to reduce background events and increased the purity of the sample.

- **$\delta_{\text{tot}}(\mathbf{E} - \mathbf{p_Z})$ cut:** Energy and momentum conservation requires ep collisions to conserve the quantity $E - p_Z$, where E is the total energy of the event given by $E = E_p + E_e$ and p_Z is the sum of longitudinal momentum $p_Z = E_p - E_e$. Their difference, i.e. $\delta = E - p_Z = 2E_e$ yields twice the energy of the incoming lepton beam. Thus, for perfectly measured NC DIS events, $E - p_Z$ should peak at 55 GeV. Particles that escape through the forward beam hole do not contribute to the overall

δ. However, the particles escaping through the rear beam-hole reduce this quantity by two times their energy. So in photoproduction events where the scattered electron is not detected, the δ distribution peaks at lower values, typically below 30 GeV and hence a cut of $E - p_Z > 40$ GeV was used to reject such events. An upper cut at $E - p_Z < 65$ GeV was introduced to remove residual cosmic ray background and also to suppress events with an overlap between beam-gas events and physics ep scatterings. Figure 6.4(a) shows the distribution of the δ variable with the vertical dashed red lines indicating the cut values.

- **y_{el} cut:** In some PhP events, di-leptons coming from the decay of neutral pions could produce an electromagnetic shower in the CAL and be mis-identified as the scattered electron. These fake electrons are usually detected in the FCAL region, where the tracking information can not be used to compensate the electron finder. In this case the measured energy and polar angle would be low and from the relation (5.10), the inelasticity, y, reconstructed using the electron method will have a large value. Such background events were suppressed by requiring an upper cut on y_{el}, $y_{\mathrm{el}} < 0.7$ (see Figure 6.4(b)).



(a) (b)

Figure 6.4: (a) Difference of total energy and longitudinal momentum, $E - p_Z$, and (b) y reconstructed using only the electron information, y_{el}, for both data (filled circles) and simulated events (shaded histograms) after all cuts except for the one depicted. Vertical dashed red lines indicate cut thresholds.

- **y_{jb} cut:** The events where the hadronic system was not well reconstructed in the forward region were rejected by requiring a cut on the kinematic variable y reconstructed using only the information from the hadronic system. The cut was required to be $y_{\mathrm{jb}} > 0.05$ and can be seen in Figure 6.5.

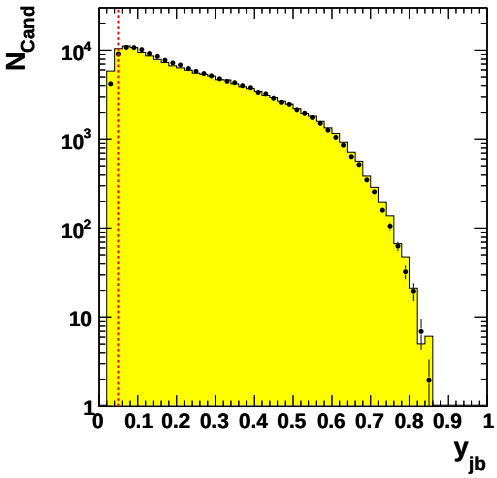


Figure 6.5: y reconstructed using only the information from the hadronic system, y_{jb}, for both data (filled circles) and simulated events (shaded histograms) after all cuts except for the one depicted. The vertical dashed red line indicates the cut threshold.

- **Transverse momentum balance:** The total transverse momentum, P_t, of a perfectly reconstructed NC event should be equal to 0 GeV. However in the ZEUS calorimeter, the resolution on the P_t measurement is proportional to $\sqrt{E_t}$. Cuts of $P_t\sqrt{E_t} < 4\sqrt{\mathrm{GeV}}$ and $P_t/E_t < 0.7$ were applied, which rejected events with a significant imbalance in P_t, usually arising from cosmic-ray events and beam-related background events.

6.4 Pre-Selection of Electron Candidates

To enrich the beauty signal from semileptonic electron decays, one has to develop ways to reduce possible background sources and to discriminate the signal from background. The beauty signal also contains the cascade decays $b \to c \to e, b \to \tau \to e$ and $b \to J/\psi \to e^+e^-$. In this analysis, in addition to semileptonic charm decays[4], there are two other types of backgrounds that need to be considered.

- **Electron background:** The main source of this type of background is electrons from photon conversions, $\gamma \to e^+e^-$, (see Appendix B). Among other non-negligible contributions are electrons from Dalitz decays of neutral pions, $\pi^0 \to \gamma e^+e^-$, and mis-reconstructed DIS electrons. The electrons from semileptonic decays have to be separated from this large amount of background.

- **Non-electron background:** Electrons can be faked by other particles, mostly pions, kaons and protons, which are produced in far greater numbers than electrons. Hence, separating electrons from non-electrons is one of the main parts of the work presented in this analysis.

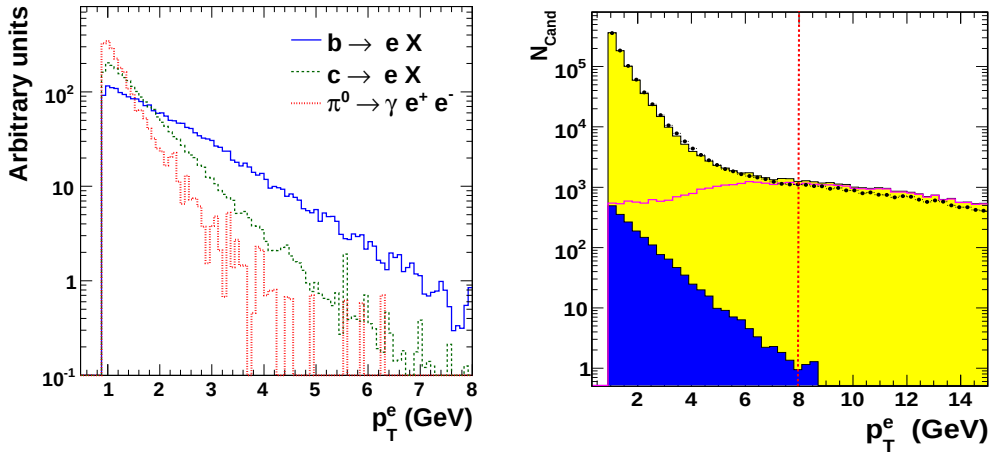
[4]Because of not enough separation power for electrons from semileptonic charm decays, they were treated as background (see Section 8.3).

Since at ZEUS most of the particle identification techniques are based on ionisation measurement and on the different showering properties of the particles in the absorber material of the detector, the tracking and calorimeter information is combined in this analysis for electron identification. The detailed strategy for identification of electrons from semileptonic decays by combining information from different variables is described in the next two chapters. In this section, a pre-selection of possible electron candidates is summarised.

6.4.1 Selection Cuts

- **Track p_T:** Tracks were required to have a minimum transverse momentum of $p_T^e > 0.9$ GeV. Below this cut, hadronic and electronic showers become nearly indistinguishable. Even dE/dx in the CTD, the only non-shower based technique, has also fundamental difficulties to separate hadrons and electrons below this momentum (see Section 7.1.1). This cut also helps to reduce the electron background from photon conversions and Dalitz decays. As the goal of this analysis is to look for electrons from semileptonic b-hadron decays, a higher transverse momentum of the track enhances the contribution from decays of beauty hadrons whilst reducing the contributions from decays of charm and light flavour hadrons.

  Figure 6.6(a) shows the distribution of p_T^e for electrons from b-hadron decays (solid



(a) (b)

Figure 6.6: (a) Normalised distributions of transverse momentum, p_T^e, for electrons from semileptonic beauty decays (blue), charm decays (green), and Dalitz decays (red). (b) p_T^e distribution for both data (filled circles) and simulated events (shaded histograms) after all cuts except for an upper cut on p_T^e. The vertical dashed red line indicates the cut threshold. The contributions from remaining DIS electron background (magenta) and the beauty signal (blue shaded area) are also shown.

blue line), c-hadron decays (green dashed line) and from Dalitz decays (red dotted line). To enable a shape comparison, the distributions for Dalitz decay and decays from charm are area-normalised to the distribution of b decay[5]. The semileptonic charm decay distribution is harder than for Dalitz decays with the distribution for semileptonic b decay being significantly harder than these two. This clearly shows that the higher in p_{T}^{e}, the larger the enhancement of decays from b hadrons. But, since a non-negligible amount of DIS electrons remains[6], which usually have a larger momentum, an upper cut at $p_{\mathrm{T}}^{e} > 8$ GeV was also applied. This cut was chosen such that it helped to suppress the DIS electron background without losing too much beauty signal. In Figure 6.6(b) the data and MC comparison is shown for p_{T}^{e} without applying an upper cut. The yellow shaded area corresponds to the sum of MC normalised to the data, which is shown in black points. A magenta line is drawn for the true remaining DIS background and the blue shaded area is the beauty signal. It is clear from this distribution that an upper cut on p_{T}^{e} at 8 GeV helps to reduce a significant amount of DIS background without losing any beauty signal.

- **Track η:** The electron candidate selection was restricted to the central region by applying a cut on the pseudo-rapidity of $|\eta_e| < 1.5$. This ensures the best possible measurement of the track kinematics by the CTD and allows $\mathrm{d}E/\mathrm{d}x$ to be used for particle identification.

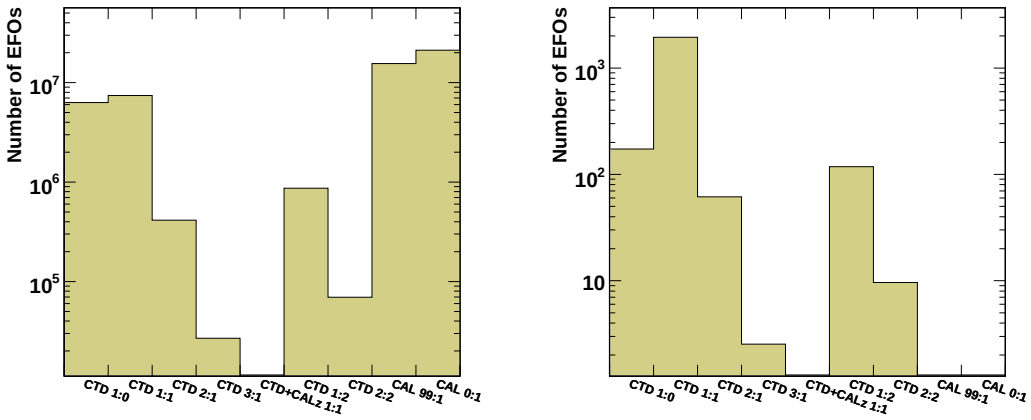
- **EFO object:** Several EFO types exist depending on the track-island relationship and on the method of energy determination. Distributions of different EFO classes for all particle types and for a sub-sample of only electrons from the simulated beauty sample are shown in Figure 6.7(a) and 6.7(b), respectively[7]. As can be seen in Figure 6.7(b) the dominant part of the true electrons are reconstructed as type 1:1, so only this type of EFOs is selected. In this type of EFO object, only one reconstructed track is matched to one island. This cut helps significantly to improve the signal-to-background ratio.

- **Energy deposit:** Electrons lose almost all their energy in the electromagnetic part of the calorimeter, while hadrons deposit a considerable amount in the hadronic part of the calorimeter. Therefore the fraction of electromagnetic energy in the EMC relative to the total energy ($E^{\mathrm{EMC}}/E^{\mathrm{CAL}}$), measured for an energy flow object in the CAL, is useful for separating electrons from hadrons. Figure 6.8 shows the shape of this variable for electrons, muons and hadrons. Most of the electrons have an EMC-fraction close to one, whereas the hadrons deposit a significant fraction of their energy in the hadronic calorimeter, which is reflected in lower values of $E^{\mathrm{EMC}}/E^{\mathrm{CAL}}$.

[5]The term "b (c) decay" refers to "b-hadron (c-hadron) decay".

[6]This implies that these scattered electrons which were identified by Sinistra were not correct.

[7]Note that for (CAL 99:1) and (CAL 0:1) it is not possible to make a matching to the Monte Carlo truth information.



(a) **(b)**

Figure 6.7: Distributions of different EFO classes: (a) for all particle types and (b) for a sub-sample of only electrons from the simulated beauty sample. The classes describe the n:m relationship of tracks and islands. Energies are either determined by the CTD or the CAL (cf. Section 5.4.1). The dominant part of the true electrons are reconstructed as type "CTD 1:1". The notations "CAL 99:1" or "CTD+CALz 1:1" are technical description for the EFOs where the association to the tracks, or CAL information, respectively, did not work properly.

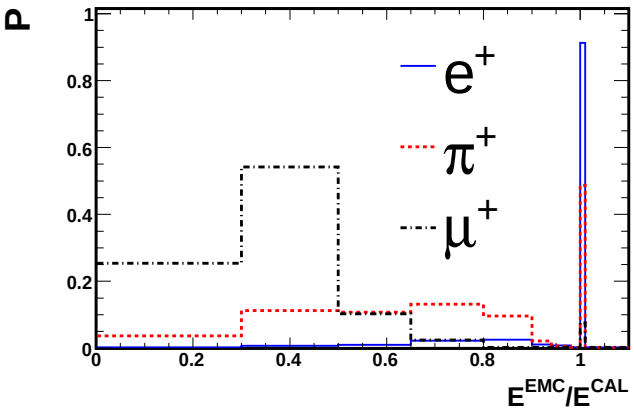


Figure 6.8: Normalised distribution of $E^{\mathrm{EMC}}/E^{\mathrm{CAL}}$ for electrons (blue), pions (red) and muons (black). Most of the electrons have an EMC-fraction close to one, whereas the hadrons deposit a significant fraction of their energy in the hadronic calorimeter. Energies of muons are split between the calorimeter and the muon system giving a mean value of $E^{\mathrm{EMC}}/E^{\mathrm{CAL}}$ around 0.4 (Taken from [120]).

The muons, which are minimal ionising particles, can pass the calorimeter, and deposit part of their energy in the muon chambers, giving a mean value of $E^{\mathrm{EMC}}/E^{\mathrm{CAL}}$ around 0.4. A hard cut of $E^{\mathrm{EMC}}/E^{\mathrm{CAL}} > 0.95$ was required to enrich the electron sample.

- **EFO energy:** A minimum energy in an EFO object, E^{CAL}, was required to be greater than 0.5 GeV. The origin of this cut is discussed in Section 7.1.2.

- **Geometry cuts:** As discussed during the event selection, some geometry cuts are needed to remove those parts of the calorimeter where the reconstruction of the scattered electron is not reliable. Similarly EFOs (and hence the electron candidates) can not be reconstructed properly in such regions and geometry cuts based on the position of EFOs were introduced. The candidates were rejected if they were found in the following regions:

$-160 < z_{\mathrm{isl}} < -110$ cm	(super crack I)
$-178 < z_{\mathrm{isl}} < -232$ cm	(super crack II)
$y_{\mathrm{isl}} > 80$ cm and $\lvert x_{\mathrm{isl}}\rvert < 12$ cm	(chimney cut)
$R_{\mathrm{isl}} > 175$ cm	(radius cut).

- **Conversion background:** Electrons from photon conversions were tagged and rejected based on the distance of closest approach of a pair of charged tracks opposite to each other in the plane perpendicular to the beam axis and on their invariant mass. Some details on the way the conversion finder works at ZEUS are given in Appendix B.

- **dE/dx measurement:** The $\mathrm{d}E/\mathrm{d}x$ measurement provides a tool for distinguishing different types of particles and was used as a main variable for particle identification. To reduce the major background of non-electrons in the candidate selection a preselection cut was applied on the electron hypothesis test $T^{e}_{\mathrm{d}E/\mathrm{d}x}$ (see Section 8.2.1). The applied cut was $-2\ln T^{e}_{\mathrm{d}E/\mathrm{d}x} < 3$.

- **Track quality:** In order to measure the mean decay length of the tracks within the jet a secondary vertex was fitted using all good tracks assigned to the jet. The good tracks were defined by a minimum transverse momentum of at least $p_{\mathrm{T}} > 0.5$ GeV, four hits in the MVD and at least three superlayers passed in the CTD.

- **Jet association:** A jet with the following characteristics was associated to the electron: $p_{\mathrm{T}}^{\mathrm{jet}} > 2.5$ GeV, $|\eta^{\mathrm{jet}}| < 2.0$.
  The electron-jet association was done by cutting on the distance between the electron candidate and the jet in the η–ϕ plane defined by $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 1.0$. A lower cut on ΔR at 0.05 was also applied to reduce DIS electron background as can be seen in Figure 6.9, where the ΔR distribution after applying an upper cut at 1 and a line representing the lower cut at 0.05 is shown. About 80% of the electron candidates were associated to a jet and kept for further analysis.

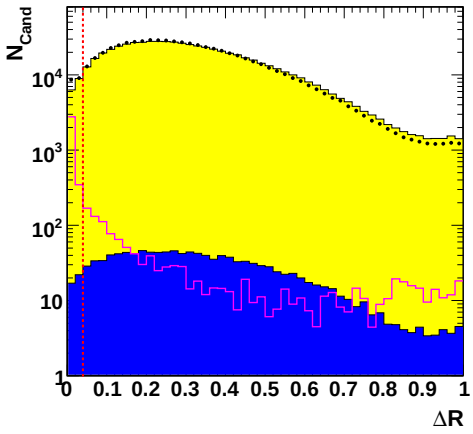


Figure 6.9: The distance between the electron candidate and the jet in the η–ϕ plane, ΔR, for both data (filled circles) and simulated events (shaded histograms) after all cuts except for a lower cut on ΔR. The vertical dashed red line indicates the cut. The contributions from remaining DIS background (magenta) and beauty signal (blue shaded area) are also shown.

- **Secondary vertex quality:** To use decay length significance as a tool for separating the long-lived beauty hadrons from charm and light flavour background, a good quality of reconstructed secondary vertices was required. Good quality was ensured by applying the following cuts: $\chi^2/\text{dof} \leq 6, |d_{XY}| < 1$ cm, $|d_Z| < 30$ cm, where d_{XY} and d_Z denote the distances of the vertices to the beam spot in X–Y and Z, respectively. These cuts were adopted from the inclusive secondary vertexing-analysis [125] where several studies were made to choose this appropriate set of cuts.

6.5 Summary of Selection

Figure 6.10 shows a scatter plot of the reconstructed data events that passed all the event and candidate selection requirements in the x–Q^2 plane. Some example lines of constant y, θ and Q^2 are also shown.

A summary of all offline cuts for event and candidate selection is given in Table 6.3

After the DIS event and electron candidate selection, the quality of the data was checked with several control plots. These plots include all selection cuts listed in the Table 6.3. Figure 6.11 shows some example distributions for the scattered electron and kinematic variables and in Figure 6.12 a selection of further control plots including some track plots is shown. All distributions, except the pseudorapidity of electron candidates, η^e, show a good agreement between the data and the Monte Carlo. Several studies were done to understand the origin of the disagreement in the η^e distribution and are presented in Chapter 9.

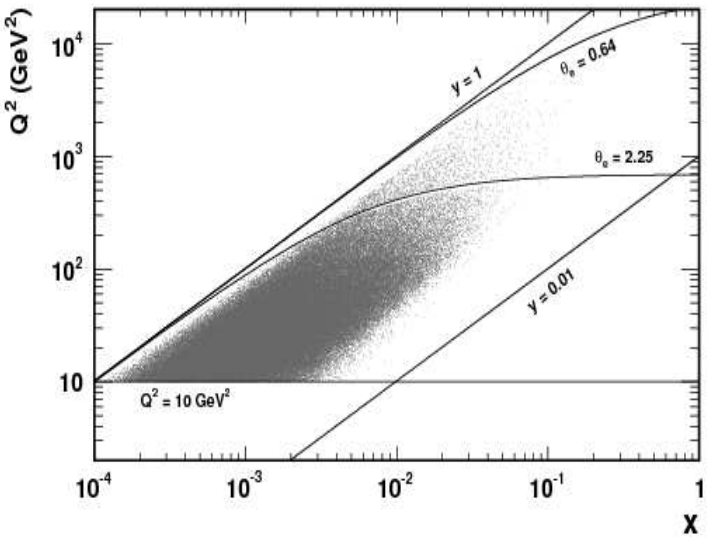


Figure 6.10: Scatter plot of the reconstructed data events that passed all the event and particle selection requirements in the x–Q^2 plane.

Event selection	**Candidate selection**
$P_{\text{sira}} > 0.90$	$0.9 < p_{\text{T}}^{e} < 8\,\text{GeV}$
$E_{e'} > 10\,\text{GeV}$	$-1.5 < \eta^{e} < 1.5$
$E_{\text{non-e}} < 5\,\text{GeV}$	1:1 track-island
Track matching to DIS e	$E^{\text{EMC}}/E^{\text{CAL}} > 0.95$
Geometry cuts	$E^{\text{CAL}} > 0.5\,\text{GeV}$
$-30 < Z_{\text{vtx}} < 30\,\text{cm}$	Geometry cuts
$Q^2 > 10\,\text{GeV}^2$	Veto on conversions
$40 < E - p_Z < 65\,\text{GeV}$	$-2\ln T^{e}_{\text{d}E/\text{d}x} < 3$
$y_{\text{el}} < 0.7$	Track quality cuts
$y_{\text{jb}} > 0.05$	$p_{\text{T}}^{\text{jet}} > 2.5\,\text{GeV}$
$P_t/E_t < 0.7$	$-2.0 < \eta^{\text{jet}} < 2.0$
	$0.05 < \Delta R < 1$
	Secondary vertex quality cuts

Table 6.3: Summary of cuts applied for offline selection of DIS events and electron candidates

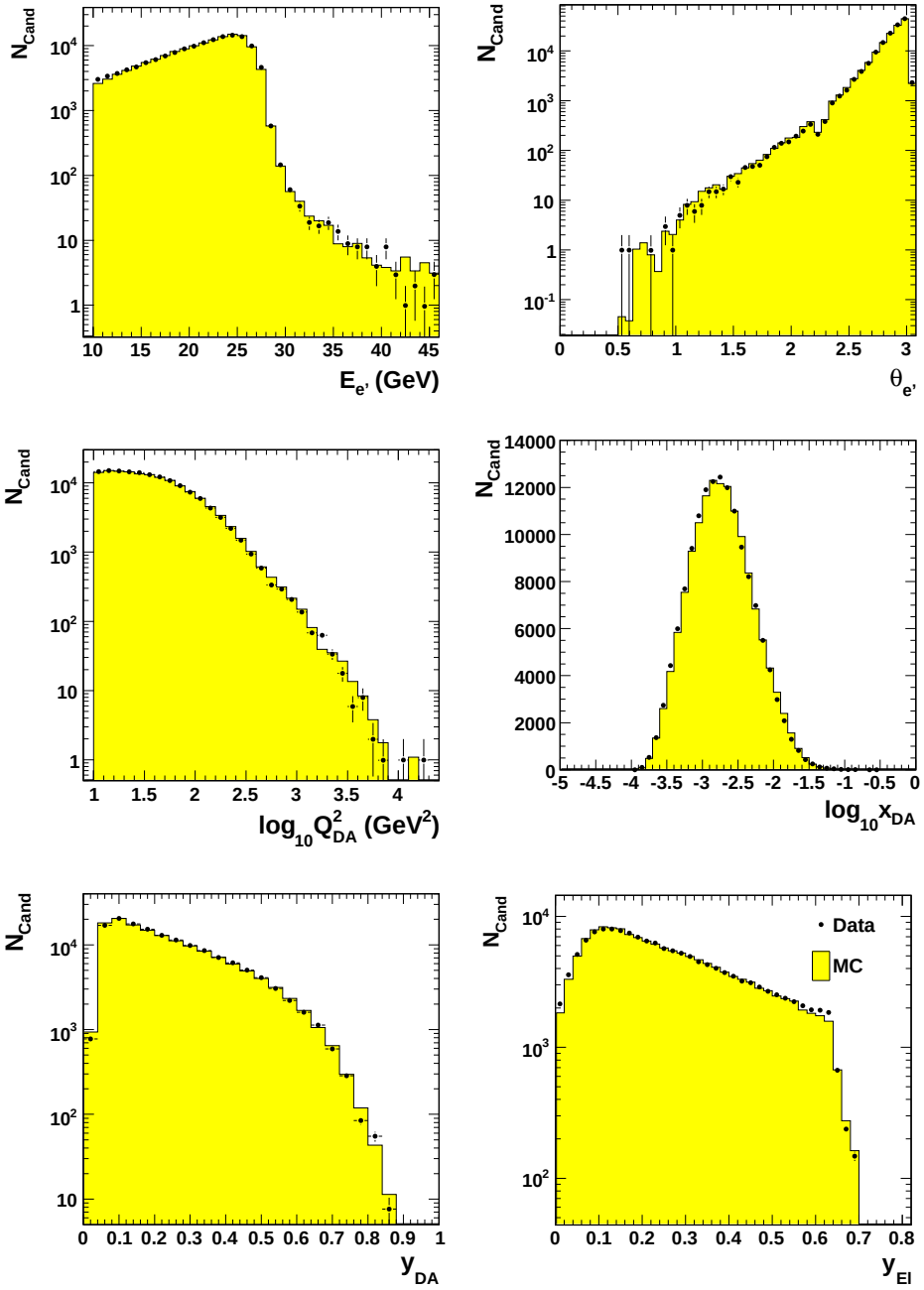


Figure 6.11: Control plots for scattered electron and kinematic variables for both data (filled circles) and simulated events (shaded histograms) after all selection cuts. (From top left to bottom right: $E_{e'}$, $\theta_{e'}$, Q^2, x, y_{DA} and y_{El}).

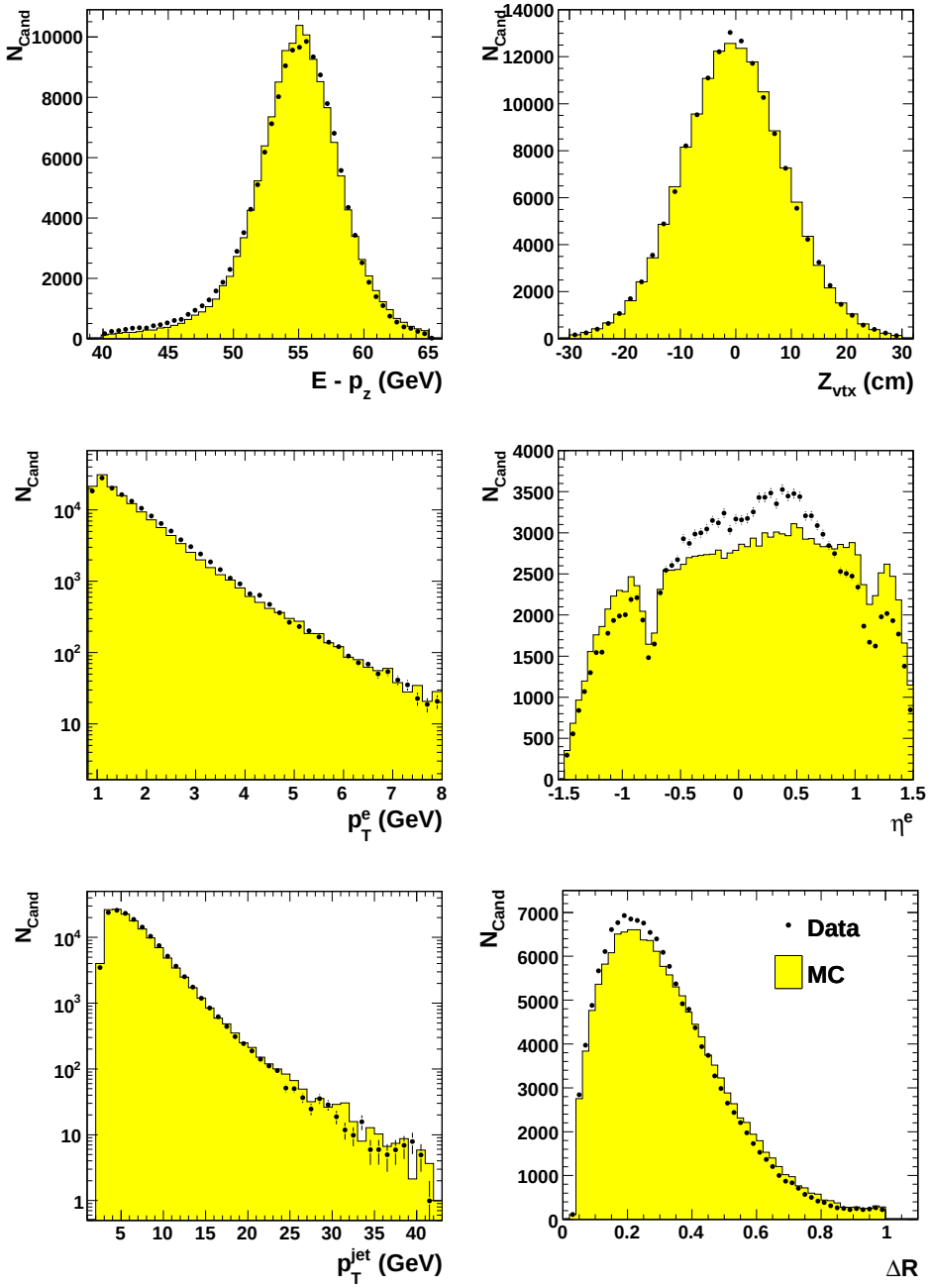


Figure 6.12: Control plots for track variables for both data (filled circles) and simulated events (shaded histograms) after all selection cuts. (From top left to bottom right: $E-p_Z$, $Z_{\rm vtx}$, $p_{\rm T}^e$, η^e, $p_{\rm T}^{\rm jet}$ and ΔR).

Chapter 7

Discriminating Observables

As described in Section 6.4, for the identification of electrons from semileptonic b decays, two types of background sources need to be considered. To extract the relatively small beauty signal from these large backgrounds, a likelihood method developed during previous analyses [113, 114] was adopted. This method involves particle and decay identification using variables sensitive to electron identification as well as to semileptonic decay kinematics and then combines the information of these variables into one discriminating variable using a likelihood hypothesis.

In the context of this analysis, several discriminating variables were studied to perform particle and decay identification. A selection of variables was done depending on their separation power and on their description of the data and the Monte Carlo. In this chapter, the variables used at the end in the likelihood hypothesis are described in Section 7.1 and 7.2. Some alternative variables are presented in Section 7.3. The way how to combine the information of different variables using the likelihood method is the subject of the next chapter.

7.1 Particle Identification

To identify electrons from other particle types ($\pi^{\pm}, K^{\pm}, p/\bar{p}$), the following variables were found to have good separation power and were used as input variables for the likelihood:

- $\mathrm{d}E/\mathrm{d}x$, the average energy loss per unit length of the track in the CTD;
- $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$, the ratio of energy deposited in the calorimeter to track momentum measured in the CTD;
- d_{cell}, the depth of the energy deposit in the calorimeter.

7.1.1 Ionisation Loss

The ionisation energy loss per distance, $\mathrm{d}E/\mathrm{d}x$, was used as a main variable for electron identification [115]. The value of $\mathrm{d}E/\mathrm{d}x$ depends on a particle's velocity and therefore particles of different masses have different $\mathrm{d}E/\mathrm{d}x$ for the same momentum. This enables different types of particles to be distinguished. The $\mathrm{d}E/\mathrm{d}x$ of a particle of unit charge is given by the Bethe-Bloch formula [116, 117].

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{4\pi N e^4}{m_e c^2 \beta^2} z^2 \left(\ln \frac{2 m_e c^2 \beta^2 \gamma^2}{I} - \beta^2 \right), \tag{7.1}$$

where x is the distance travelled by the particle through the medium, z is the charge of the travelling particle, N is the number of density of electrons in the matter traversed and I is the mean excitation energy of the atom. Equation 7.1 is only valid for particles much heavier than the scattered particles and does not take into account statistical fluctuations due to secondary tracks at high energies. Thus, in this application, the simulation of the $\beta\gamma$ dependence of the $\mathrm{d}E/\mathrm{d}x$ was done by implementing a statistical formulation using the *photo-absorption-ionisation model* by Allison and Cobb [118].

The predictions of the Bethe-Bloch formula for different particle types, in the momentum range of $0.1 - 30$ GeV and the actual measured values of CTD $\mathrm{d}E/\mathrm{d}x$ in the momentum range $0.1 - 100$ GeV are displayed in Figure 7.1. At low momentum hadron bands are

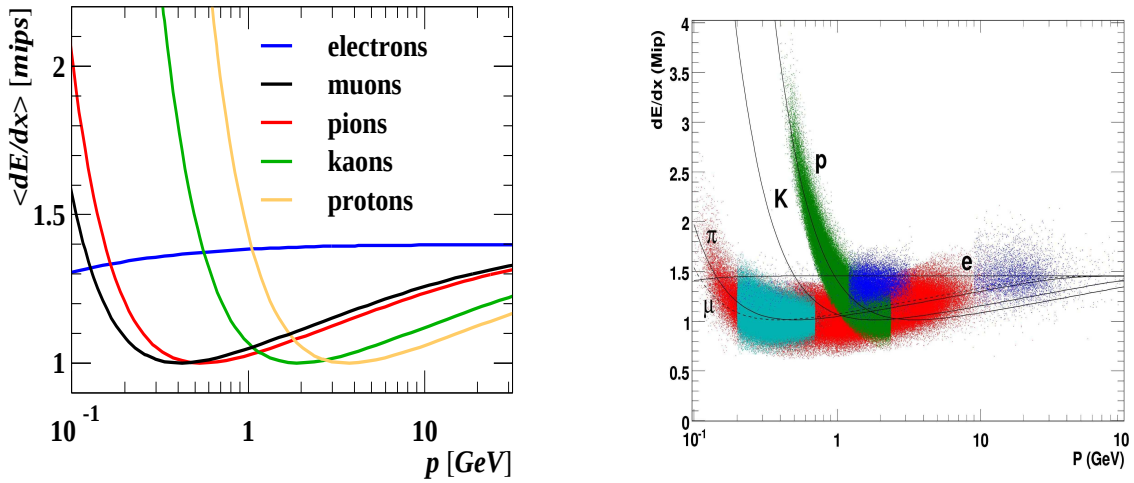


Figure 7.1: Bethe-Bloch prediction for different particle types (left), taken from [121]. $\mathrm{d}E/\mathrm{d}x$ measured in the CTD for different particle samples as a function of the momentum, p, (right), taken from [122].

clearly separated from each other but the $\mathrm{d}E/\mathrm{d}x$ for electrons crosses the hadron bands, which makes it difficult to separate the electrons from hadrons below a momentum of 1 GeV. Above this momentum, the $\mathrm{d}E/\mathrm{d}x$ of electrons is sufficiently separated from that of the hadrons rendering it an important discriminating variable for particle identification.

In the ZEUS software, variables are available where $\mathrm{d}E/\mathrm{d}x$ information was used to calculate probabilities to find a certain particle type. These probability distributions for electrons, protons, pions and kaons are shown in Figure 7.2. The distributions are well described and were used as an input to the likelihood. They were also used to calculate electron hypothesis test, $T^e_{\mathrm{d}E/\mathrm{d}x}$, (cf. Section 8.2.1) where a preselection cut at $-2\ln T^e_{\mathrm{d}E/\mathrm{d}x} < 3$ was applied to reduce the major background of non-electrons.

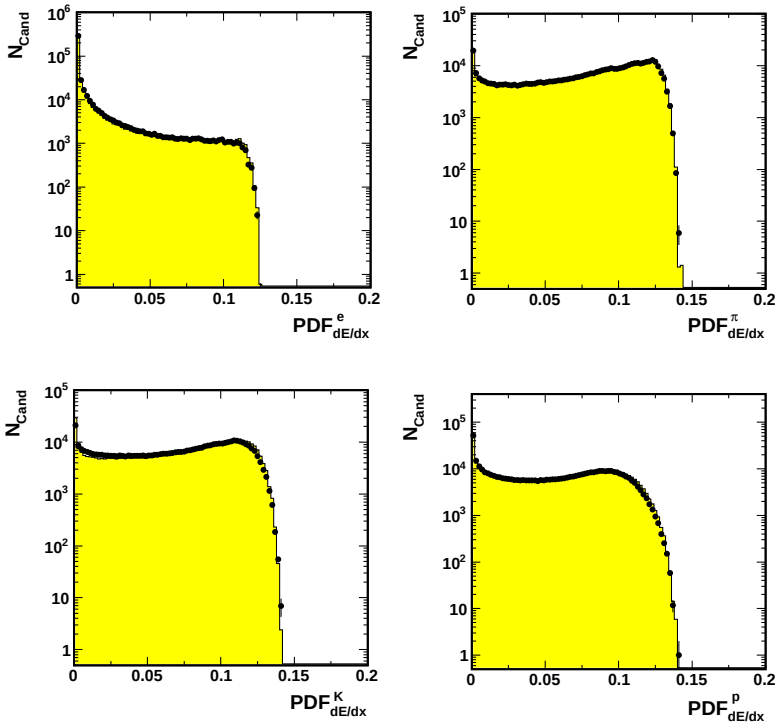


Figure 7.2: Probability distributions of a given particle and its kinematic quantities to be an electron, pion, kaon or proton. Shown are distributions for both data (filled circles) and simulated events (shaded histograms). The latter are the luminosity-weighted sum of all Monte Carlo samples and normalised to the data distributions.

7.1.2 Calorimeter Energy over Track Momentum

As described in Section 4.2.3, the ZEUS calorimeter was a high resolution compensating calorimeter which provided equal signals for hadrons and electromagnetic particles of the same initial energy, i.e. $e/h = 1$, where e and h are responses to electrons and hadrons, respectively. This, however, is best achieved only for energies above 10 GeV. Below 10 GeV, there is some deviation from $e/h = 1$, especially below 3 GeV and there are differences in the response to different particle types at the same momentum. As also the resolution in

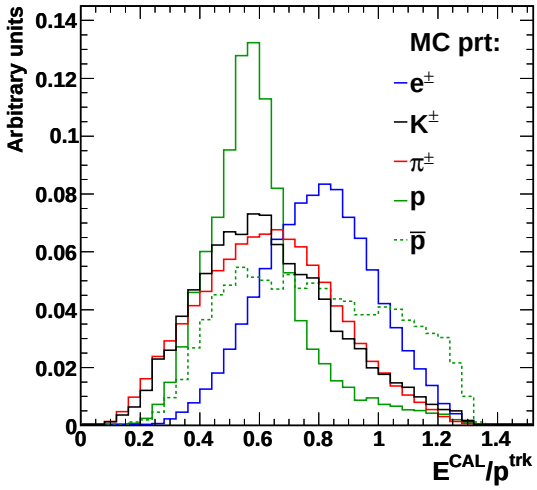


Figure 7.3: Normalised distributions of $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ for electrons, kaons, pions, protons and anti-protons extracted from the Monte Carlo samples used in this analysis. For the first three types of particles, the samples of particles and anti-particles are combined, as no significant difference was observed.

the CAL differs for hadrons and electrons of the same energy, it results in different energy distributions for the different particle types.

In the energy range $1 < E < 10$ GeV, the momentum resolution of the CTD is better than the CAL energy resolution (cf. Section 5.4) and gives a good reference for the tracks matched to calorimeter energy deposits. By taking the ratio of energy and momentum measurements, $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$, the energy dependence is reduced and hence this variable can be used for particle identification.

This variable was mainly introduced to separate anti-protons [119], since due to the annihilation processes within the detector material, the distribution of $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ is shifted to larger values than that for other particles, where $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ is usually expected to be one. However, since the calorimeter compensation is a function of particle's energy, the mean values of the $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ distributions are shifted depending on the particle's mass. This can be seen in Figure 7.3, where the distributions of heavy particles are moved to lower $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ values, while the lighter particles are shifted to high values. In this figure the normalised $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ distributions for all particles in the Monte Carlo samples used for this analysis are presented. The distributions were also checked separately for positive and negative charged particles and were found to have very similar shapes except for protons and anti-protons.

To check the quality of the variable description, the data was compared with the Monte Carlo as shown in Figure 7.4(a). The Monte Carlo does not reproduce the data with large discrepancies at lower $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ values. As such behaviour was not observed in the HERA I analysis [115], some studies were done to find the origin of the disagreement. It was found that with the higher tracking precision for the HERA II data, the tracks which were matched to EFO objects with not much energy associated, caused imperfections at lower values of $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$. Such candidates were removed by applying a cut on minimum energy in an EFO object, $E^{\mathrm{CAL}} > 0.5$ GeV. The $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ distribution after this cut is shown in

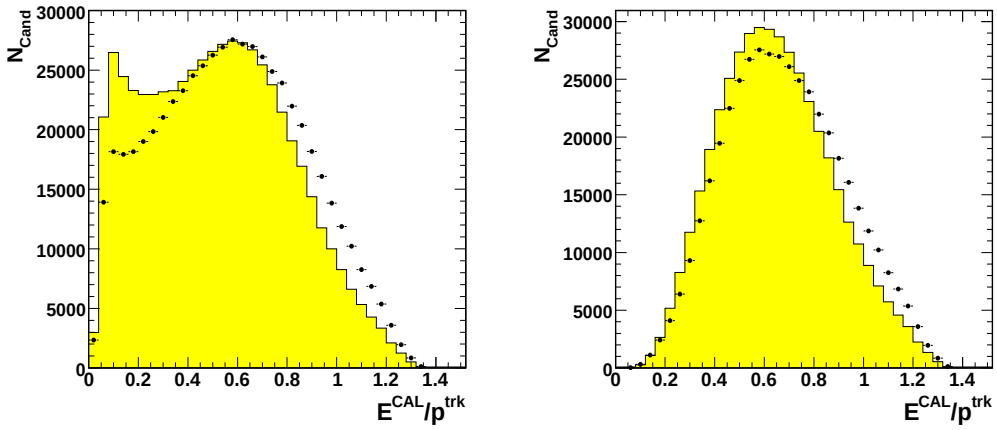


Figure 7.4: $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ distribution of electron candidates before and after applying a cut on minimum value of E^{CAL}. Shown are distributions for both data (filled circles) and simulated events (shaded histograms).

Figure 7.4(b). Still the description is not good, showing a shift in the data with respect to the Monte Carlo. Later studies showed that the remaining discrepancies are due to not well simulated background and hence including this variable in the likelihood does not cause large systematic effects.

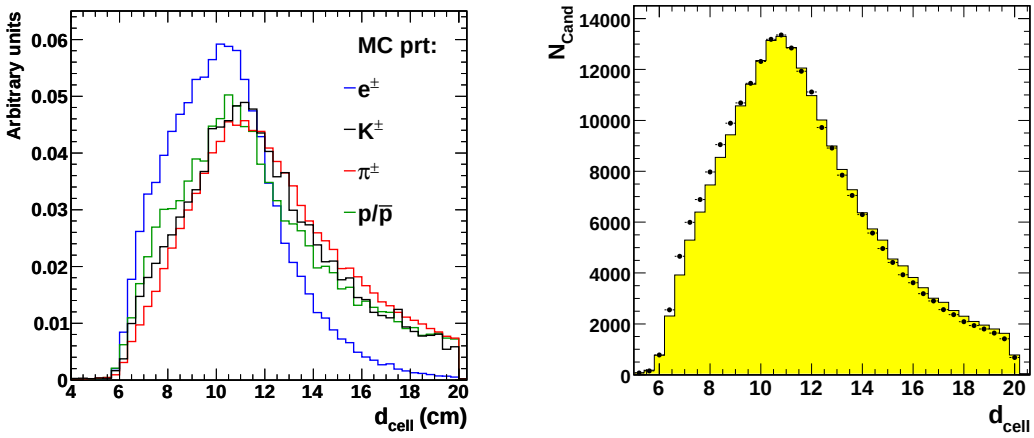
7.1.3 Depth of Energy Deposit

Previous studies performed in PhP analyses [113, 120] showed that the depth of the energy deposit in the calorimeter, d_{cell}, can be used to separate electrons and hadrons. The normalised d_{cell} distributions of all particles in the Monte Carlo samples used in this analysis are presented in Figure 7.5(a). Although the separation power of this variable is small compared to $\mathrm{d}E/\mathrm{d}x$[1], it was included in the likelihood as it is expected to have a small systematic uncertainty. Also this variable shows quite a good description between the data and the Monte Carlo (see Figure 7.5(b)).

7.2 Decay Identification

For the identification of electron candidates originating from different decay sources, the following variables were found to have good separation power and were combined with the particle identification variables using a likelihood method:

[1]The separation is weak as already applied a cut on $E^{\mathrm{EMC}}/E^{\mathrm{CAL}}$.



(a) (b)

Figure 7.5: (a) Normalised distributions of d_{cell} for electrons, pions, kaons and protons extracted from the Monte Carlo samples used in this analysis. (b) d_{cell} distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms).

- $p_{\text{T}}^{\text{rel}}$, the transverse momentum component of the electron candidate relative to the direction of the jet axis;
- $\Delta\phi$, the azimuthal angle difference between the electron and the neutrino, both originating from the semileptonic decay;
- $d/\delta d$, the signed decay-length significance of the electron candidate.

7.2.1 Using Mass Information

The large mass of the beauty quarks can be exploited to differentiate them from light flavour quarks. Since beauty quarks are significantly more massive than charm and light flavour quarks, they show different decay kinematics. A suitable observable which reflects the characteristic of the quark masses in semileptonic decays into electrons is the transverse momentum of the electron relative to its parent quark, $p_{\text{T}}^{\text{rel}}$. Due to the large beauty quark mass, the $p_{\text{T}}^{\text{rel}}$ spectra of semileptonic b decays is harder than the $p_{\text{T}}^{\text{rel}}$ spectrum for charm and light quarks.

As quarks are not experimentally observable, jets are reconstructed to represent the momentum of the quark and to define the reference for $p_{\text{T}}^{\text{rel}}$ calculation which is a reasonable estimate of the direction of the parent quark. The transverse momentum is then calculated using the following relation:

$$p_{\text{T}}^{\text{rel}} = \frac{|\vec{p}^{\,\text{jet}} \times \vec{p}_e|}{|\vec{p}^{\,\text{jet}}|}, \tag{7.2}$$

where $\vec{p}^{\,\text{jet}}$ and $\vec{p}_e$ are the momentum vectors of the jet and the electron candidate. A sketch of the reconstruction of the variable is depicted in Figure 7.6. This variable which

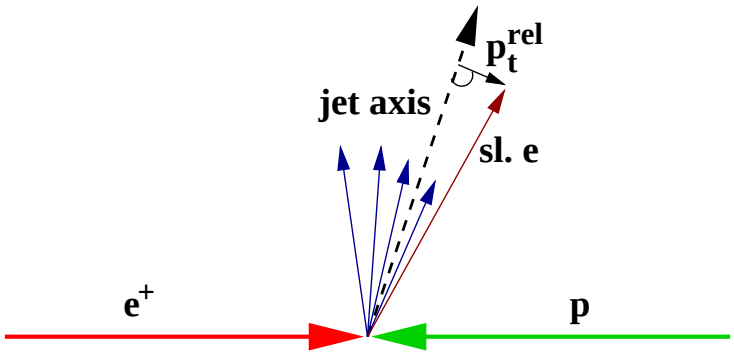


Figure 7.6: Definition of transverse momentum of electron with respect to jet axis, $p_{\text{T}}^{\text{rel}}$.

serves as an important variable for beauty separation has already been used in previous heavy flavour analyses at ZEUS using semileptonic decays to both muons and electrons (e.g. [94, 123, 115]).

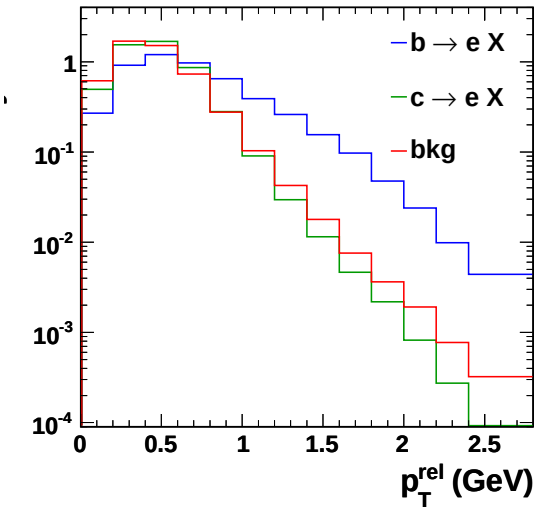
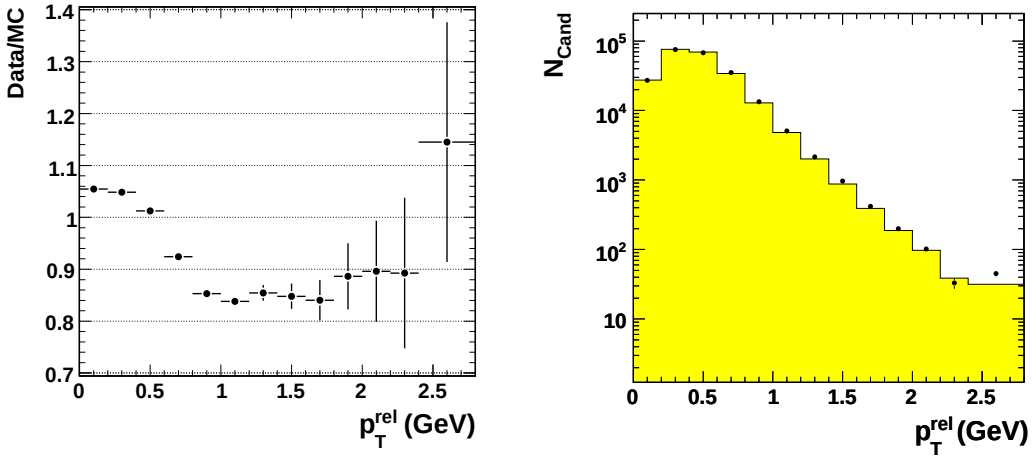


Figure 7.7: Normalised distributions of $p_{\text{T}}^{\text{rel}}$ for semileptonic electrons from beauty decays (blue), charm decays (green), and other decays (red). The distributions were extracted from the Monte Carlo samples used in this analysis.

Figure 7.7 shows $p_{\text{T}}^{\text{rel}}$ spectrum for true semileptonic beauty and charm decays as well as for other decays labelled as background. The distributions were made using the Monte Carlo samples used in this analysis after applying all the DIS selection cuts and preselection of electron candidates (cf. Chapter 6). The distributions were normalised to one to compare the shape for different samples. As can be seen the $p_{\text{T}}^{\text{rel}}$ spectrum from beauty decays has a harder spectrum than for charm and other decays and hence a good separation of the beauty signal is possible.

In order to use the shape of the $p_{\text{T}}^{\text{rel}}$ variable for beauty identification in the likelihood, a good description of the data by the simulated events is necessary. In previous analyses [123,

94, 113] deviations in the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution have been observed. Studies showed that the discrepancy is visible in the background but not in the signal-enriched region. In order to avoid any systematic bias in the signal region, the correction factors were determined from a background dominated sample. The pure background sample was extracted by cutting on the test function $T^{e}_{\mathrm{d}E/\mathrm{d}x}$. The correction factor shown in Figure 7.8(a) was determined by computing the ratio in every bin of $p_{\mathrm{T}}^{\mathrm{rel}}$ between the data and the Monte Carlo. The correction is of the order of $\sim 5-15\%$. The $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution after applying the corrections is depicted in Figure 7.8(b) and shows a very good agreement between the data and the Monte Carlo.



(a) (b)

Figure 7.8: (a) Correction factor for the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution determined by taking the ratio of the data to the Monte Carlo from a background enriched sample. (b) $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms) after applying the corrections.

7.2.2 Using Neutrino Information

Neutrinos originating from semileptonic b decays can provide a source of information for beauty identification. Since neutrinos are not directly observable in the detector, the missing transverse momentum vector, $\vec{\not{p}}_{\mathrm{T}}$, can be utilised for neutrino reconstruction. $\vec{\not{p}}_{\mathrm{T}}$ is defined as the negative sum of the EFO momenta transverse to the beam axis,

$$\vec{\not{p}}_{\mathrm{T}} = -\left(\textstyle\sum_i p_x^i, \sum_i p_y^i\right), \tag{7.3}$$

where the sum runs over all EFOs. The vector $\not{p}_{\mathrm{T}}$ is used as an estimator to represent the direction of the neutrino escaped from the semileptonic decay. Since the transverse momenta of the neutrinos from semileptonic decays are of the order of few GeV only, their

precise reconstruction is not possible. Previous studies [113, 115] however have shown that their azimuth can be determined reliably. Therefore the azimuthal angle difference, $\Delta\phi$, between the outgoing neutrino and the electron defined by:

$$\Delta\phi = |\phi(\vec{p}_e) - \phi(\vec{\not{p}}_{\mathrm{T}})|, \tag{7.4}$$

was calculated as a measure of the decay signature of the neutrino. Figure 7.9 shows a sketch of $\Delta\phi$ reconstruction. Contrary to the $p_{\mathrm{T}}^{\mathrm{rel}}$ calculation, where the jet axis was chosen

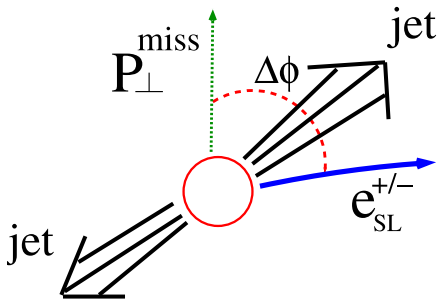
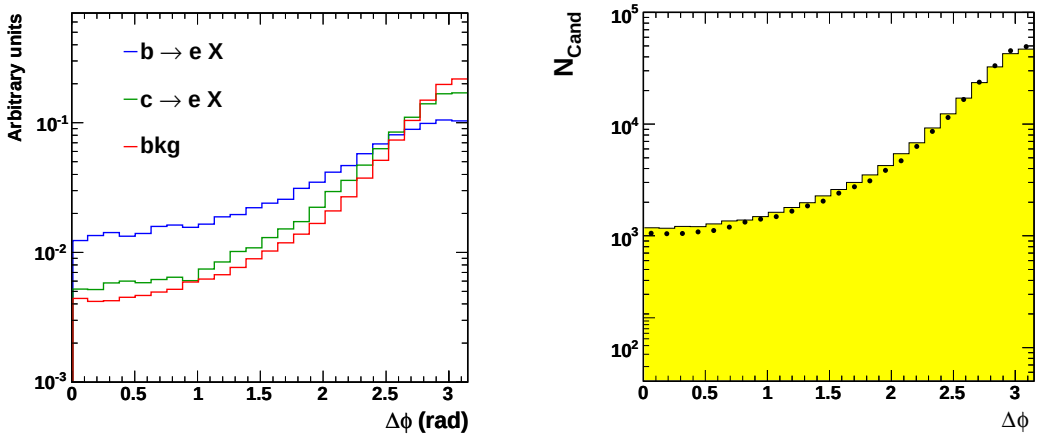


Figure 7.9: Schematic sketch of $\Delta\phi$, the azimuthal angular difference between the electron and the neutrino, both originating from the semileptonic decay.

to roughly describe the direction of the hadronic decay products, the electron direction used in $\Delta\phi$ is much more precise. Thus the $\Delta\phi$ variable is in some way a complementary variable to $p_{\mathrm{T}}^{\mathrm{rel}}$.



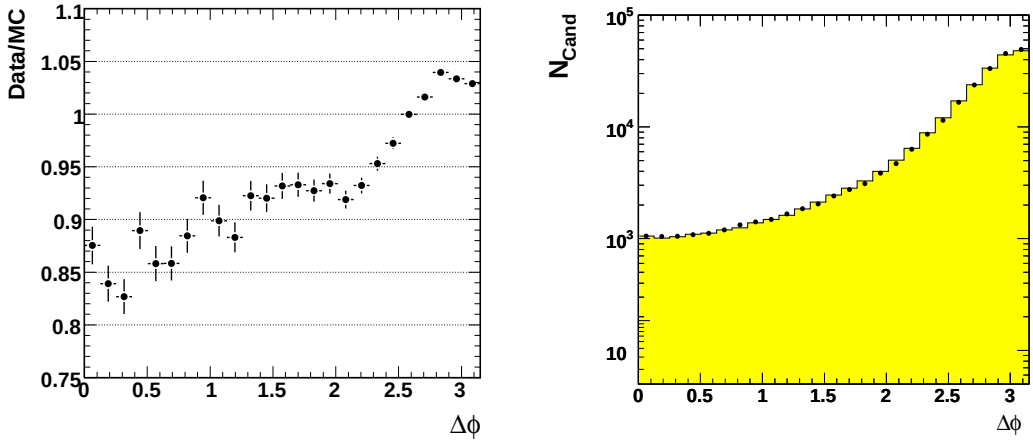
(a) (b)

Figure 7.10: (a) Normalised distributions of $\Delta\phi$ for semileptonic electrons from beauty decays (blue), charm decays (green), and other decays (red). (b) $\Delta\phi$ distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms).

The normalised $\Delta\phi$ distributions for electrons originating from semileptonic beauty and charm decays and all other electrons are shown in Figure 7.10(a). The distributions show

a clear separation of beauty from charm and light flavour events. In contrast to PhP analyses [115, 120], the separation of charm from light flavour events is worse because of the low jet momenta.

A comparison of the $\Delta\phi$ distribution for the data and the Monte Carlo, shown in Figure 7.10(b), shows some discrepancies at low $\Delta\phi$ values. A similar method as used for the $p_{\mathrm{T}}^{\mathrm{rel}}$ correction was applied to extract the correction factor for light flavour events. The correction factor is shown in Figure 7.11(a) and is of the order of $5-15\%$. The effect of this correction was found to be much smaller than for $p_{\mathrm{T}}^{\mathrm{rel}}$ and the systematic uncertainties obtained from the correction variation were treated as a consistency check (cf. Chapter 10). The $\Delta\phi$ distribution, after applying the correction to the light flavour events, can be seen in Figure 7.11(b) and shows a very good agreement between the data and the Monte Carlo.



(a) **(b)**

Figure 7.11: (a) Correction factor for the $\Delta\phi$ distribution determined by taking the ratio of the data to the Monte Carlo from a background enriched sample. (b) $\Delta\phi$ distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms) after applying the corrections.

7.2.3 Using Lifetime Information

The long lifetime of b hadrons can be used to separate the beauty signal from charm and light flavour background. A measurable quantity closely related to the lifetime information is the decay length, d, which is calculated as the distance between the primary vertex (or beamspot) and the secondary vertex. Due to the higher precision on its position measurement, it was decided to use the beamspot $(X_{\mathrm{bsp}}, Y_{\mathrm{bsp}})$ as a reference point for the decay-length calculation. The decay length was restricted to two dimensions by calculating

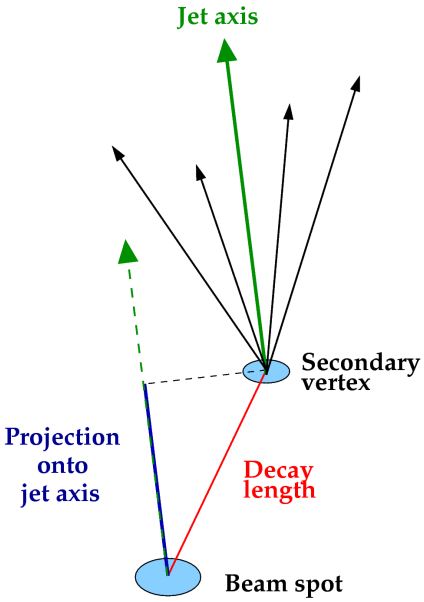


Figure 7.12: Sketch of the decay-length calculation. The distance between the beamspot and the refitted secondary vertex is calculated in X–Y plane and then is projected onto the axis of the associated jet. (Taken from [125])

its projection onto the X–Y plane and is given by:

$$d = \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} = \begin{pmatrix} X_{\mathrm{vtx}} - X_{\mathrm{bsp}} \\ Y_{\mathrm{vtx}} - Y_{\mathrm{bsp}} \end{pmatrix}. \tag{7.5}$$

The sign of the decay length was assigned by determining the relative position of the secondary vertex with respect to jet axis. If the secondary vertex was in the same hemisphere as the jet axis, a positive sign was assigned to the decay length and a negative sign otherwise. Negative decay lengths which originate from the secondary vertices reconstructed opposite to the direction of associated jet are unphysical and are caused purely by detector resolution effects. To avoid discontinuities for the vertices where the decay vector was approximately perpendicular to the jet direction, the two dimensional decay length was further projected onto the jet axis resulting in a modification of Equation 7.5 to:

$$d = \begin{pmatrix} \Delta X \\ \Delta Y \end{pmatrix} \cdot \frac{\vec{j}_{2D}}{|\vec{j}_{2D}|}. \tag{7.6}$$

The determination of the decay length and its projection onto the jet axis is depicted in Figure 7.12

As expected from the lifetime difference between charm hadrons and beauty hadrons, the shape of the decay-length distribution is different for semileptonic beauty decays from those for the charm decays and light flavour background. This can be seen in Figure 7.13(a), where normalised decay-length distributions are calculated for different Monte Carlo samples after applying all selection cuts. For light flavours there is a symmetric distribution

around zero, while an asymmetry with an enhancement towards large decay lengths for beauty and charm, due to longer lifetimes is visible. It should be noted that the decay-length significance, $d/\delta d$, has larger separation power (see Figure 7.13(b)) and hence was used in this analysis. δd is the error on decay length and was calculated from the covariance matrices of the beamspot and the secondary vertex [125].

The reliability of the use of decay-length significance depends on the Monte Carlo descrip-

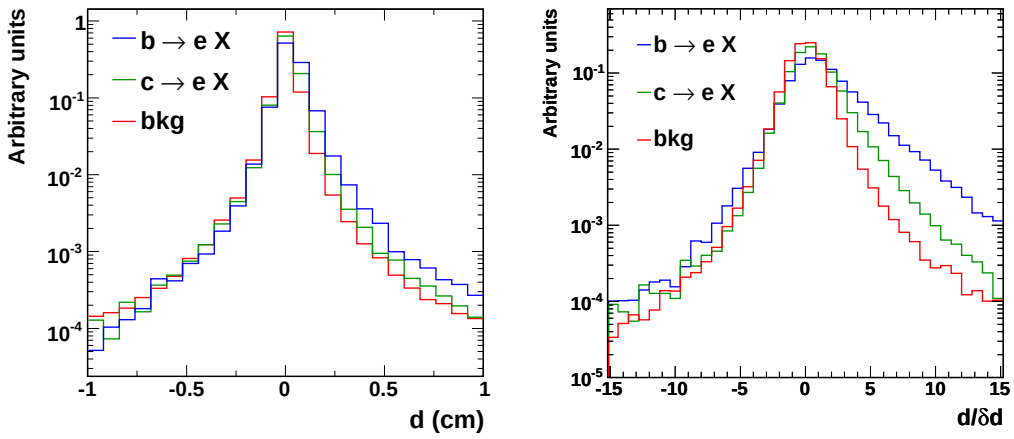


Figure 7.13: Normalised distributions of decay length, d, and decay-length significance, $d/\delta d$, for semileptonic electrons from beauty decays (blue), charm decays (green), and other decays (red).

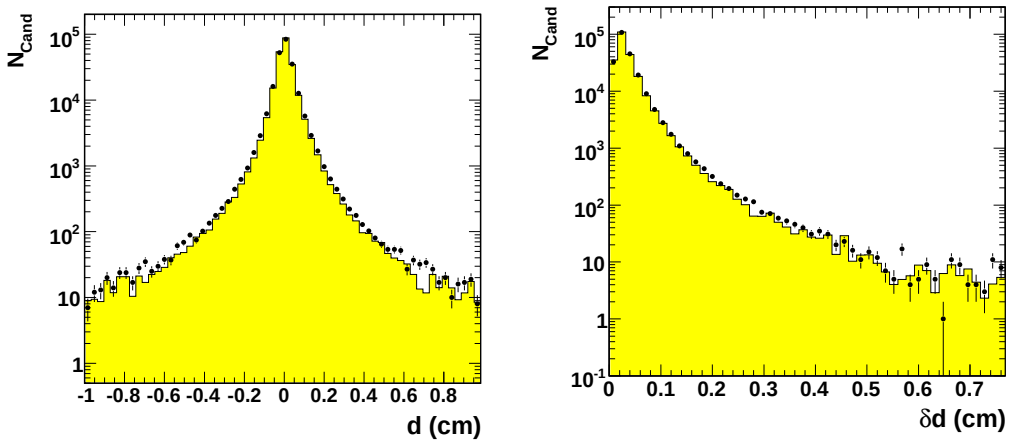


Figure 7.14: Control plots for decay length, d, and its error, δd, for the data (filled circles) and the simulated events (shaded histograms).

tion of decay length and its error. A comparison of the decay length and its error distributions between the data and the MC is shown in Figure 7.14. While the decay-length error is reasonably well described, the decay length in itself shows some discrepancies; the data distribution is a little bit wider than the MC distribution. To correct for small deficits in the description of the decay-length distribution, a procedure developed in [125] was adopted to smear the values of the decay length and improve the overall description. The decay-length distribution in Monte Carlo was corrected using an empirical function which smeared a fraction of all candidates by a Gaussian function which had the main effect for significances below five (central region of $d/\delta d$). The fraction of candidates for which the decay length was smeared was set to 5% and the width of the Gaussian was chosen to be $1.8 \cdot \delta d$. In addition, a smaller fraction of 1% of the candidates in the intermediate region (significances below 10) was smeared by another Gaussian of width $2.3 \cdot \delta d$ and in the tails a fraction of $\delta d \cdot 10\%$ of decay lengths were smeared with an exponential function.

Figure 7.15 shows a comparison of the decay-length significance distributions before and after applying the above mentioned smearing to the decay lengths in the Monte Carlo. A clear improvement in the agreement between the data and the Monte Carlo is visible. The smearing procedure was varied to estimate the systematic uncertainties as discussed in Chapter 10.

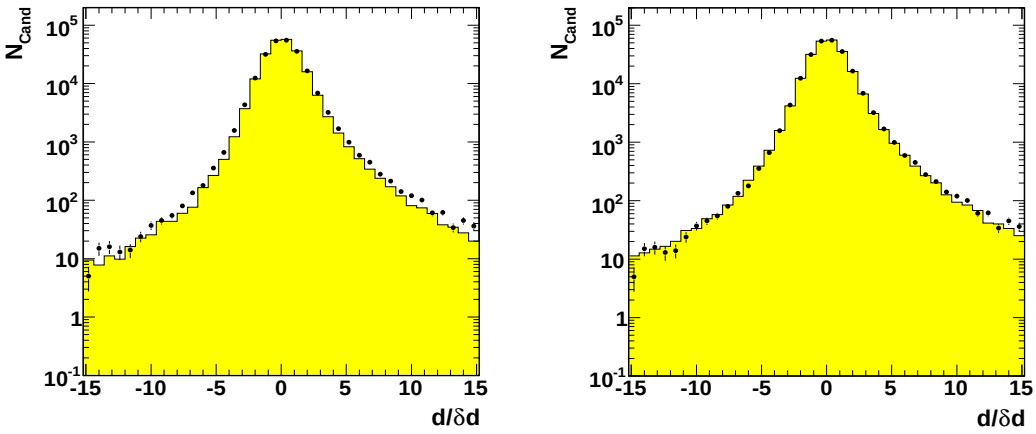


Figure 7.15: Control plots for decay-length significance, $d/\delta d$, for the data (filled circles) and the simulated events (shaded histograms), (a) before and (b) after applying smearing.

7.3 Alternative Observables

In addition to the variables already discussed, several other variables were studied which could be useful for separating semileptonic b decays from c decays and light flavour back-

ground. Some of these variables are described briefly below, but they were not used in the likelihood at the end, e.g. because of introducing large systematics due to not well described distributions.

7.3.1 Impact Parameter Significance

In addition to the decay length, another approach to use lifetime information uses the measured impact parameter. It is a track-based variable and is defined as the distance of closest approach of a track to the primary vertex (or beamspot). The impact parameter of a track is assigned a positive sign if the angle between the jet direction and the line connecting the vertex and the point of closest approach of the track is less than 90° and a negative sign otherwise (see Figure 7.16). The positive signed impact parameters reflect the lifetime of

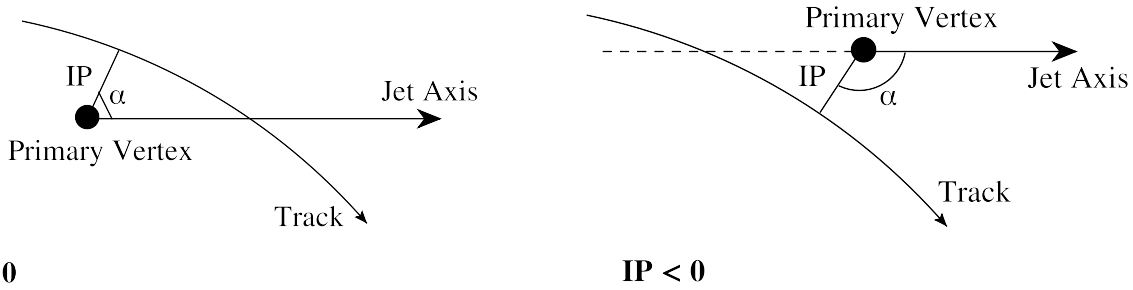


Figure 7.16: Schematic sketch of the impact-parameter variable and the definition of its sign.

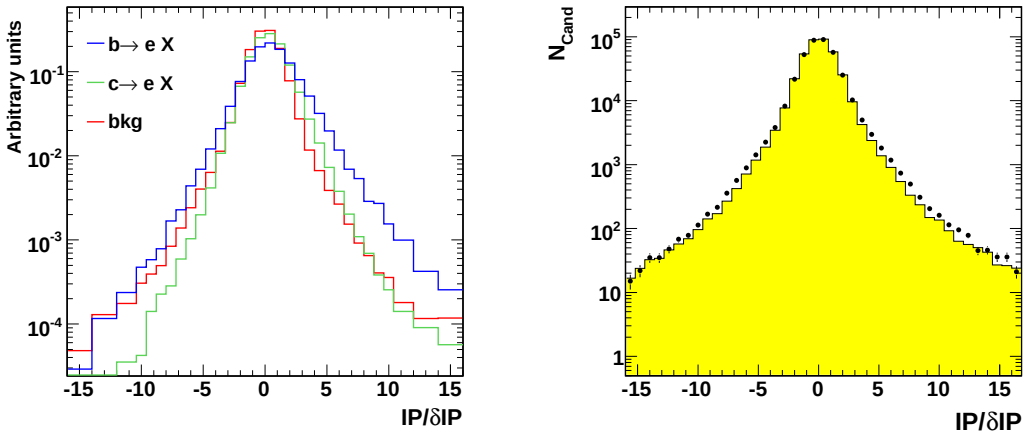
heavy hadrons and the negative signed impact parameters are caused by resolution effects of tracks of particles coming from the beamspot. To have a better control of resolution effects, the impact-parameter significance, IP/δIP, is calculated. δIP is the uncertainty on the impact parameter calculated from the beamspot and the helix uncertainties.

The use of this method has an advantage over the decay length as it increases the statistics[2]. However, the decay-length significance was used at the end as it has a larger separation power than the impact-parameter significance (see Figures 7.13(b) and 7.17(a)). Also the comparison of the data to the Monte Carlo revealed larger discrepancies in the impact-parameter significance distribution (see Figure 7.17(b)), which caused a larger systematic bias.

7.3.2 Invariant Mass

The reconstruction of the secondary vertices provide another variable that could be used to separate light flavour and heavy flavour contributions. It is known as the invariant

[2]The efficiency of reconstructing a secondary vertex is much lower than the impact parameter and is about 60%.



(a) (b)

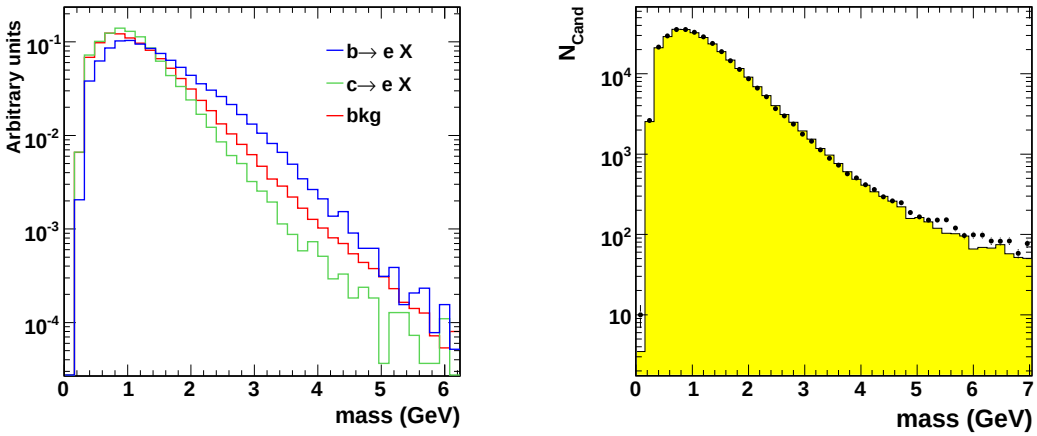
Figure 7.17: (a) Normalised distributions of IP/δIP for semileptonic electrons from beauty decays (blue), charm decays (green), and other decays (red). (b) IP/δIP distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms).

mass of the secondary vertex and is reconstructed using the four momenta of the tracks that enter the secondary vertex fit. In inclusive secondary vertex analyses [124, 125], this variable was combined with the decay-length significance to enhance the discriminating power. The decay-length significance distribution was divided into different mass bins to obtain beauty, charm and light flavour enriched regions. Such an approach requires large samples and it was not possible to use it in this analysis due to limited statistics. The other possibility of including this variable in the likelihood also did not help much, as the separation power between beauty, charm and light flavour contributions was not good due to low jet momenta (see Figure 7.18(a)). Although the mass distribution (see Figure 7.18(b)) shows a good agreement between the data and the Monte Carlo, it was not used in the likelihood at the end to avoid an additional source of systematic uncertainty.

7.3.3 $p_{\mathrm{T}}^{\mathrm{miss},||e}$

As described in Section 7.2.2, to benefit from the knowledge that the events of interest contain an undetected neutrino, a variable $\Delta\phi$ was used in this analysis. In another analysis [10], this variable was replaced by, $p_{\mathrm{T}}^{\mathrm{miss},||e}$, the component of the missing transverse momentum parallel to the lepton direction. This new variable was also studied in the context of this analysis.

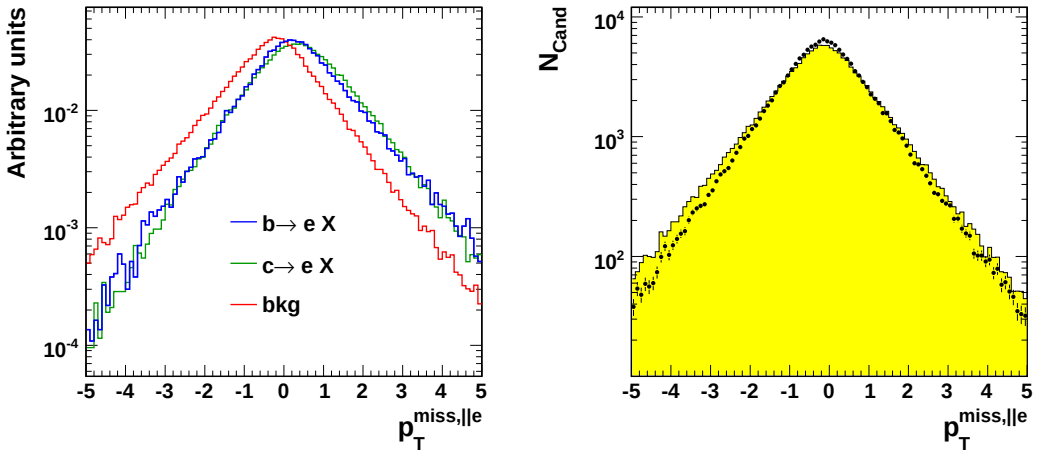
Figure 7.19(a) shows normalised distributions for electrons from b decay, from c decay and other decays and a comparison of the $p_{\mathrm{T}}^{\mathrm{miss},||e}$ distribution between the data and the



(a) (b)

Figure 7.18: (a) Normalised distributions of invariant mass of the selected secondary vertices for electrons from semileptonic beauty decays (blue), charm decays (green), and other decays (red). (b) Invariant mass distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms).

Monte Carlo is depicted in Figure 7.19(b). The $p_{\mathrm{T}}^{\mathrm{miss},||e}$ variable has a positive tail of events containing semileptonic heavy quark decays. Although this variable yields a little bit better



(a) (b)

Figure 7.19: (a) Normalised distributions of $p_{\mathrm{T}}^{\mathrm{miss},||e}$ for semileptonic electrons from beauty decays (blue), charm decays (green), and other decays (red). (b) $p_{\mathrm{T}}^{\mathrm{miss},||e}$ distribution of electron candidates for both data (filled circles) and simulated events (shaded histograms).

separation power than $\Delta\phi$, the quality of the variable description is worse. A very complicated set of corrections would be needed [10] to overcome the discrepancies which resulted in much larger systematic effects than for $\Delta\phi$ and hence the variable was used only as systematic cross check.

Chapter 8

Signal Extraction

The likelihood method used for the signal extraction is presented in this chapter. This method provides a simple way to combine the information from various sub-detectors. In our case the discriminating observables (cf. Chapter 7) based on the information of tracking detectors and the calorimeter are taken as an input to the likelihood. Since the tracking detectors and the calorimeter operate best in different momentum ranges, the measurements from these detectors are complementary. Combining these different measurements using the likelihood method helps to maximise the separation power between signal and background.

The likelihood method is described in Section 8.1, with an introduction to the needed ingredients in the subsequent Sections 8.1.1 and 8.1.2. Some studies to understand the performance of the likelihood method using the single input variables with a particular emphasis on the separation power of the $\mathrm{d}E/\mathrm{d}x$ measurement are discussed in Section 8.2. At the end of the chapter in Section 8.3, a combined likelihood showing the influence of combining the different variables into one global discriminating variable is presented.

8.1 Likelihood Ratio Test

The *likelihood ratio test* is a powerful statistical method which is used to identify particles combining all available information. It is defined as the ratio of the likelihood hypothesis to test (the particle being of the type in question) to the sum of the likelihood hypothesis and anti-hypothesis (the particle being of any other type).

The likelihood hypothesis, $\mathcal{L}_i$, for a particle type i is given by:

$$\mathcal{L}_i = \alpha_i \cdot \prod_j \mathcal{P}_i(d_j), \tag{8.1}$$

with $i \in \{e, \pi, K, p\}$[1] and α_i denoting the abundances (see Section 8.1.1) for particle type i. $\mathcal{P}_i(d_j)$ is the probability density for a particle type i being observed with some discriminating observable, d_j (see Section 8.1.2). The discriminating variables used for particle identification in this analysis are $\mathrm{d}E/\mathrm{d}x$, $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ and d_{cell} (cf. Section 7.1). Using these variables in Equation 8.1, the likelihood hypothesis for electrons, $\mathcal{L}_e$, which is the particle of interest in this analysis, can be written as:

$$\mathcal{L}_e = \alpha_e \cdot \mathcal{P}_e(\mathrm{d}E/\mathrm{d}x) \cdot \mathcal{P}_e(E^{\mathrm{CAL}}/p^{\mathrm{trk}}) \cdot \mathcal{P}_e(d_{\mathrm{cell}}). \tag{8.2}$$

To achieve the best separation power, the test function, T_e, for a particle being an electron is defined by taking the ratio of two likelihoods:

$$T_e = \frac{\mathcal{L}_e}{\sum_{i\in\{e,\pi,K,p\}} \mathcal{L}_i}, \tag{8.3}$$

where the numerator corresponds to the likelihood hypothesis for an electron and the denominator is the sum of all hypotheses for different particle types.

To further differentiate electrons originating from different decay sources, from semileptonic decay of b or c hadrons or from other decays, the variables $p_{\mathrm{T}}^{\mathrm{rel}}$, $\Delta\phi$ and $d/\delta d$ (cf. Section 7.2), sensitive to different aspects of heavy quark decays are combined in the likelihood. The likelihood hypothesis for electrons, $\mathcal{L}_e$, given in Equation 8.2, extended using the additional information of its decay source, which in our case is semileptonic b decay, is given by:

$$\begin{aligned} \mathcal{L}_{e,b} \;=\; & \alpha_e \cdot \mathcal{P}_e(\mathrm{d}E/\mathrm{d}x) \cdot \mathcal{P}_e(E^{\mathrm{CAL}}/p^{\mathrm{trk}}) \cdot \mathcal{P}_e(d_{\mathrm{cell}}) \cdot \\ & \beta_b \cdot \mathcal{P}_b(p_{\mathrm{T}}^{\mathrm{rel}}) \cdot \mathcal{P}_b(\Delta\phi) \cdot \mathcal{P}_b(d/\delta d), \end{aligned} \tag{8.4}$$

where β_b denotes the abundances for electrons from semileptonic b decays. In an analogous way the extension of the test function, T_e, given in Equation 8.3 to a combination of a test for a particle being an electron and a second test for an electron candidate originating from a semileptonic b decay $T_{e,b}$ is given as:

$$T_{e,b} = \frac{\mathcal{L}_{e,b}}{\sum_{i,j} \mathcal{L}_{i,j}}, \tag{8.5}$$

where i belongs to different particle types and j stands for different decay sorts ($e^{\pm}$ from semileptonic beauty decays, $e^{\pm}$ from semileptonic charm decays and background). In the following the combined beauty test function for the electron hypothesis, $T_{e,b}$, is denoted by T and for the test function using a single variable, the notation T_{d_j} is used.

[1] In this analysis, the contribution of muons was found to be negligible after the selection cuts and is therefore not listed.

8.1.1 Particle Abundances

As defined in Equation 8.4, the abundances for different particle and decay types is one of the ingredients needed for the calculation of the likelihood test function. The relative abundances for different particle types are extracted from the Monte Carlo samples (cf. Section 6.1) and vary over the kinematic region of interest. Hence, the particle abundances for $e^{\pm}, \pi^{\pm}, p/\bar{p}$ and $K^{\pm}$ are binned in the pseudorapidity, η, and the transverse momentum, p_{T}, of the tracks. To avoid binning effects in the likelihood distribution at the end, a fine binning for the abundances is chosen. Bins with low statistics are combined afterwards and the entries are scaled with the bin width. Further, the distributions are normalised such that in each bin the sum of the different particle types is one (see Figure 8.1). The complicated dependence of the relative abundances as a function of p_{T} and η can be seen.

The relative abundances of electrons originating from different sources; from the semilep-

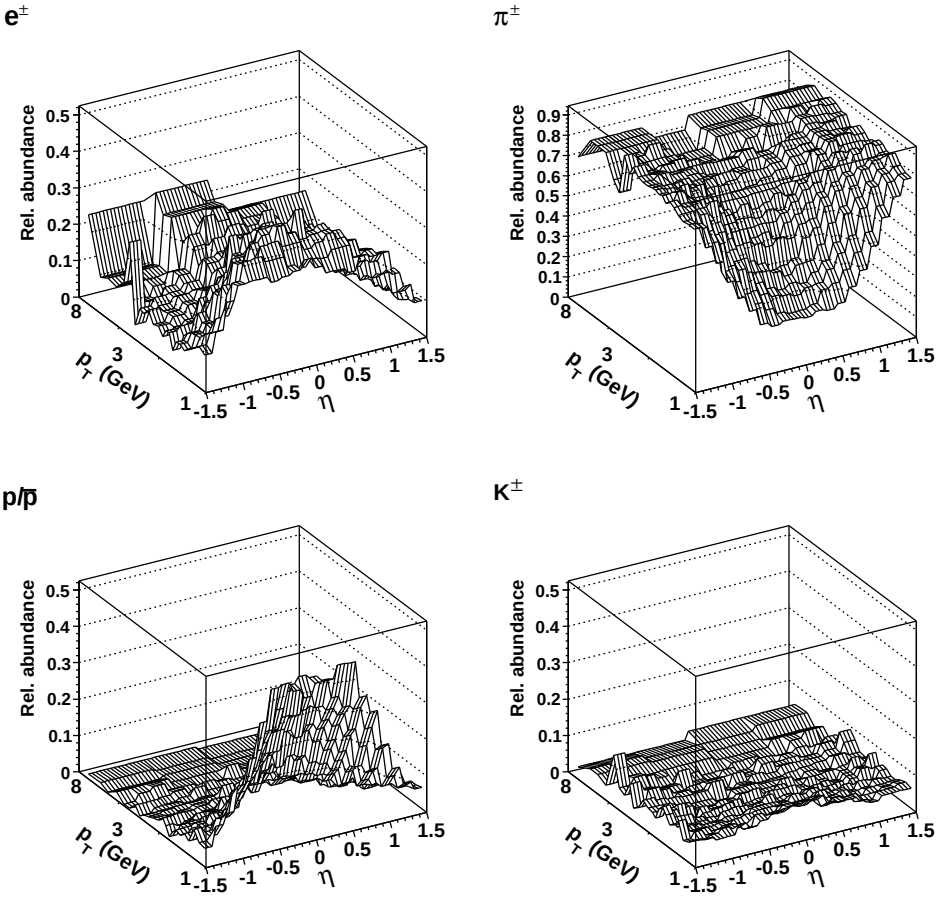


Figure 8.1: Relative abundances for different particle types $e^{\pm}, \pi^{\pm}, p/\bar{p}$ and $K^{\pm}$ as a function of η and p_{T}.

tonic b decays, c decays and all other sources are calculated only in bins of p_{T} of the tracks because no η dependence was observed. Like particle abundances, the numbers are taken again from the Monte Carlo samples. The three distributions are shown in Figure 8.2. Due to the large mass of b quark, the momentum spectrum of electrons from semileptonic b decays is harder than for semileptonic c decays which helps to differentiate between the two decay types. For electrons from other sources, at low momentum the contributions are mainly from electrons from photon conversions and Dalitz decays while the distribution at high momentum is dominated by mis-identified DIS electrons.

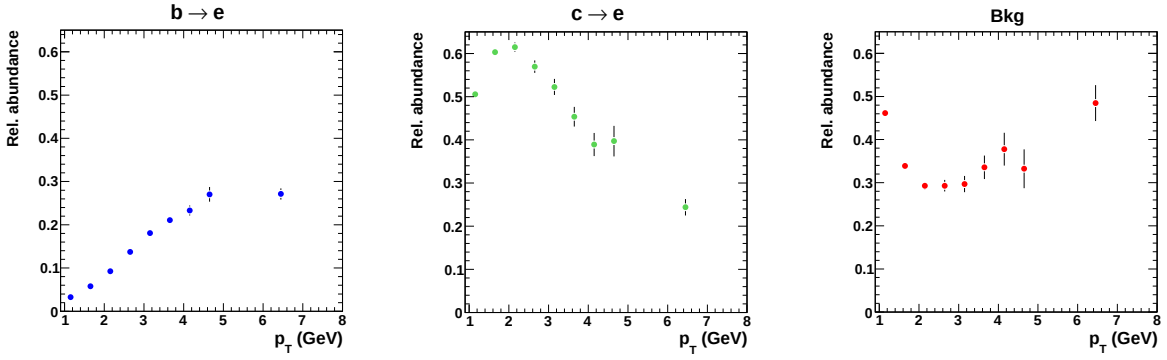


Figure 8.2: Relative abundances for different decay types as a function of p_{T}: for semileptonic beauty decays, from semileptonic charm decays and all other sources of electrons (from left to right). The numbers are obtained with the help of true information in all Monte Carlo events.

8.1.2 Probability Density Functions

The second contribution to the likelihood test function is the probability density functions, PDFs, of the variables used for the particle and decay identification. The PDFs reflect the relative probability of a particle track being observed at a given value of a selected variable. PDFs can be created either from binned histograms or analytical functions [114]. In this analysis binned PDFs obtained from the Monte Carlo simulations are used. Care has been taken that the observables used for the PDFs in the likelihood are well simulated in the Monte Carlo so that they accurately describe the data.

Binned PDFs (normalised distributions) of the variables to be used in the likelihood function were already shown in Chapter 7. The binning of the PDFs was chosen such that bins are not too wide and that there is an adequate number of entries in each bin. Wide bins were avoided because they would dilute the shape of the PDFs and lose important discriminating information for use in the likelihood.

8.2 Likelihood for Single Variables

To check the separation power of the input variables, the likelihood test function was computed separately for each variable by combining the probability density function for that variable with the relative abundances. As an example the likelihood test function for the electron hypothesis for a variable for particle identification (e.g. for $\mathrm{d}E/\mathrm{d}x$ measurement) is calculated as:

$$T_{\mathrm{d}E/\mathrm{d}x} = \frac{\alpha_e \cdot \mathcal{P}_e(\mathrm{d}E/\mathrm{d}x)}{\alpha_e \cdot \mathcal{P}_e(\mathrm{d}E/\mathrm{d}x) + \alpha_\pi \cdot \mathcal{P}_\pi(\mathrm{d}E/\mathrm{d}x) + \alpha_p \cdot \mathcal{P}_p(\mathrm{d}E/\mathrm{d}x) + \alpha_K \cdot \mathcal{P}_K(\mathrm{d}E/\mathrm{d}x)}. \quad (8.6)$$

The test functions for the other variables for particle identification are computed in an analogous way by replacing the $\mathrm{d}E/\mathrm{d}x$ PDFs by the PDFs of the other variables. Figure 8.3 shows the normalised distribution of the hypothesis test function for the three variables $\mathrm{d}E/\mathrm{d}x$, $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ and d_{cell}. Note that $-2\ln T$ has been drawn rather than T. It is common to draw $-2\ln T$, since it converges towards the χ^2-distribution for Gaussian shaped probability densities or a very large number of uncorrelated measurements. The likelihood distributions are computed from the Monte Carlo samples and are shown for electrons (blue line) and non-electrons (red dotted line).

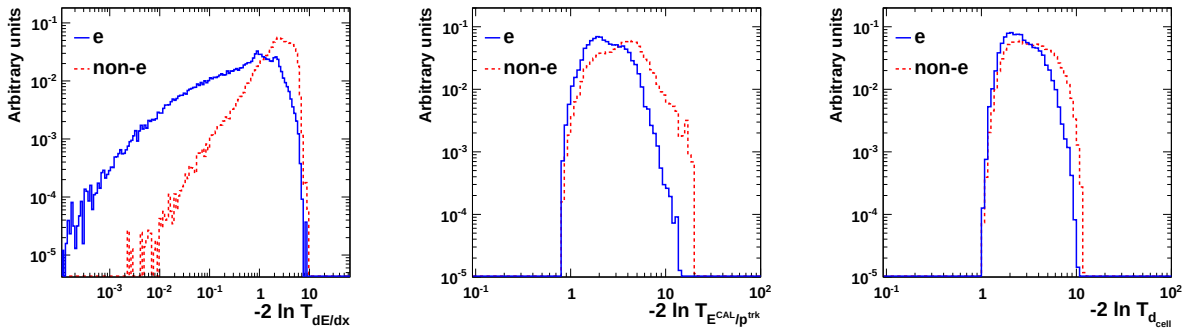


Figure 8.3: Normalised distributions of test function for a single input variable for electron identification: (a) $\mathrm{d}E/\mathrm{d}x$, (b) $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$, and (c) d_{cell}. The distributions were extracted from the Monte Carlo samples used in this analysis. Shown are the contributions for electrons (blue line), and non-electrons (red dotted line).

Figure 8.3(a) illustrates the importance of the $\mathrm{d}E/\mathrm{d}x$ measurement for electron identification. Electrons are shifted towards lower values of the test function, while non-electrons are at higher values, making it possible to have a good separation between electrons and non-electrons. With decreasing separation power, the other two variables, $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ and d_{cell} also help in distinguishing electrons from non-electrons.

Instead of differentiating between different particle types, the test function for a variable for decay identification was computed using three different decay sources. For example, the test function for beauty hypothesis using the $p_{\mathrm{T}}^{\mathrm{rel}}$ information is computed as:

$$T_{p_{\mathrm{T}}^{\mathrm{rel}}} = \frac{\beta_b \cdot \mathcal{P}_b(p_{\mathrm{T}}^{\mathrm{rel}})}{\beta_b \cdot \mathcal{P}_b(p_{\mathrm{T}}^{\mathrm{rel}}) + \beta_c \cdot \mathcal{P}_c(p_{\mathrm{T}}^{\mathrm{rel}}) + \beta_o \cdot \mathcal{P}_o(p_{\mathrm{T}}^{\mathrm{rel}})}, \tag{8.7}$$

where β_b, β_c and β_o are the abundances for the three decay sources.

Normalised distributions of the likelihood test function for the three decay variables $p_{\mathrm{T}}^{\mathrm{rel}}$, $\Delta\phi$ and $d/\delta d$ are shown in Figure 8.4. These distributions show a similar trend as was observed in the normalised distributions of these variables (cf. Chapter 7). The likelihood distribution of $p_{\mathrm{T}}^{\mathrm{rel}}$ has a different shape for electrons from semileptonic b decays, while the distributions for electrons from charm decays and other decays look similar. The variable $\Delta\phi$, which was mainly used to differentiate between heavy quark decays and other decays in the dijet analysis [113, 120], has very similar shape for the different decay types. To understand the lower separation power of $p_{\mathrm{T}}^{\mathrm{rel}}$ and $\Delta\phi$ in this analysis compared to [113, 120], some studies were performed. The distributions were made after varying the jet momentum cut in the preselection. It was observed that the poor separation power is due to the lower jet momenta (2.5 GeV compared to 6(7) GeV in the dijet selection). The shape of the likelihood distribution for $d/\delta d$ differs for the three decay types, with more separation power for beauty than for charm.

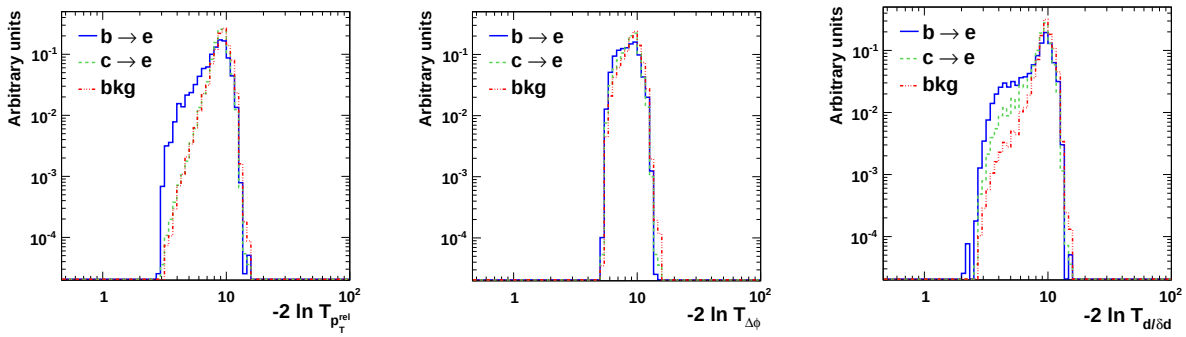


Figure 8.4: Normalised distributions of the test function for a single input variable for decay identification: (a) $p_{\mathrm{T}}^{\mathrm{rel}}$, (b) $\Delta\phi$, and (c) $d/\delta d$. The distributions were extracted from the Monte Carlo samples used in this analysis. Shown are the contributions for electrons from semileptonic b decays (blue line), semileptonic c decays (green dotted line) and all other sources (red dotted-dashed line).

8.2.1 dE/dx Likelihood

The variable dE/dx is not only able to distinguish between electrons and non-electrons, but can also be used to separate different particle types among non-electrons. This can be achieved by using dE/dx likelihood for different particle hypotheses. The dE/dx likelihood test function for different particle hypotheses was calculated using:

$$T^{i}_{\mathrm{d}E/\mathrm{d}x} = \frac{\mathcal{P}_i}{\sum_{i\in\{e,\pi,K,p\}} \mathcal{P}_i} \tag{8.8}$$

where in contrast to Equation 8.6, abundances are not used in the calculation. Figure shows the distributions of the dE/dx likelihood test function for the electron hypothesis as well as for the pion, kaon and proton hypotheses for different particle types. For these distributions the candidates were selected with all the cuts listed in Table 6.3 except the cut, $-2\ln T^{e}_{\mathrm{d}E/\mathrm{d}x} < 3$.

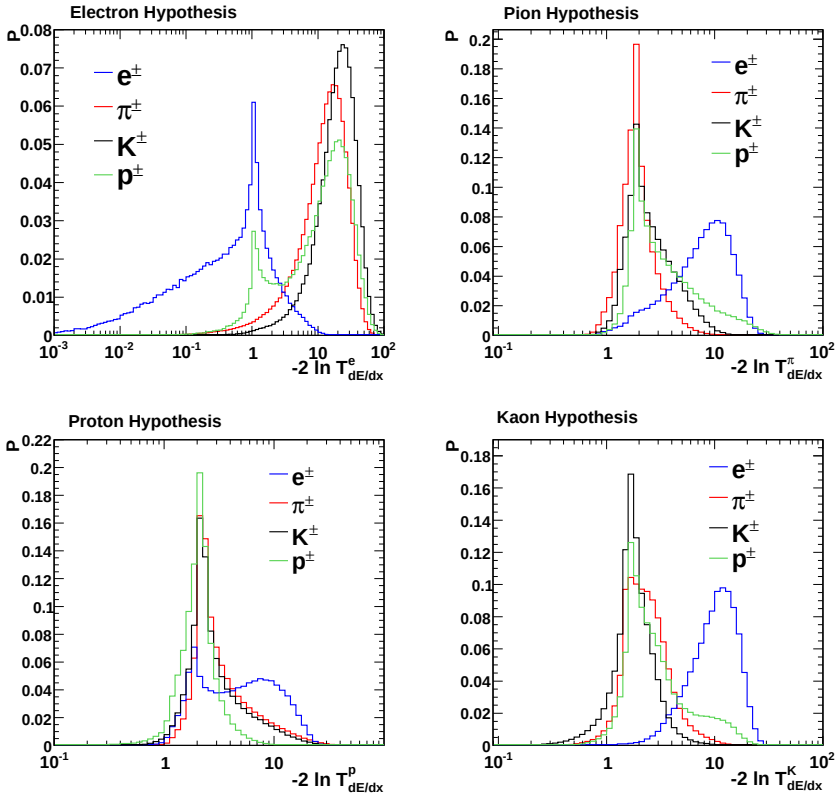


Figure 8.5: Normalised distribution of dE/dx likelihood test function for $e^{\pm}, \pi^{\pm}, p/\bar{p}$ and $K^{\pm}$. Four likelihood distribution from top left to bottom right are using the electron hypothesis, pion hypothesis, proton hypothesis and kaon hypothesis.

The upper left plot shows the separation for the electron hypothesis, which is able to distinguish between electrons and pions or kaons. The proton distribution has two peaks,

one in the pure background region and one close to the electron peak, resulting in a separation limited to the sub-sample of the right peak. The origin of the double-peak structure are the crossing points of the bands in the $\mathrm{d}E/\mathrm{d}x$–p plane (see Figure 7.1). The upper right and lower left plots show that it is also possible to make pion and proton enriched samples. Due to the overlaying structure of the $\mathrm{d}E/\mathrm{d}x$ bands the possibility to obtain a sample with a reasonable kaon purity is limited.

8.3 Combined Likelihood

To maximise the separation power between the signal and background, the information of different variables was combined into one discriminating variable using Equation 8.5. The normalised distribution of the combined likelihood test function, T, is shown in Figure 8.6. The distribution on left is shown for electrons from semileptonic b decays, c decays and other decays. This distribution shows a very good separation for electrons from b decays. The limited separation power of single variables as depicted in Figure 8.4 is much more enhanced for beauty. However, combining different variables did not help in having enough separation power for charm, since $d/\delta d$ was the the only variable to discriminate charm.

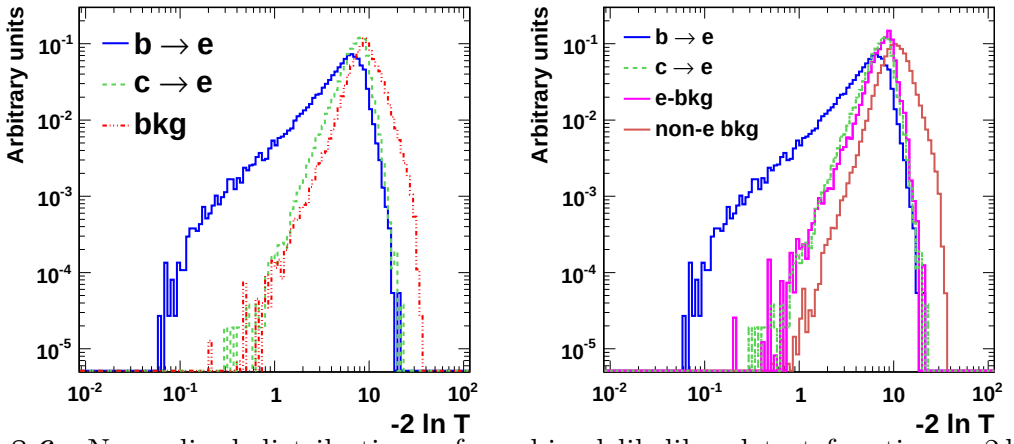


Figure 8.6: Normalised distributions of combined likelihood test function, $-2\ln T$, for electrons from semileptonic b decays (blue), semileptonic c decays (green dotted line), and other decays (red dotted-dashed line) in left plot. The background from other decays has been split into electron background (magenta) and non-electrons (brown) in right plot. The distributions were extracted from the Monte Carlo samples used in this analysis.

As a further check the background from other decays was split into electron and non-electron background (see right plot in Figure 8.6). The shapes of the charm signal and electron background are very similar. Hence if one would try to fit charm, the fit would be sensitive to electron background fraction. To avoid large systematic uncertainties for

charm results, the charm and electron background contributions were combined and are defined as other electrons in following. The distribution of the combined likelihood test function was then used only to extract beauty cross sections by fitting the contributions for electrons from b decays, other electrons and non-electrons (see Section 9.4).

Chapter 9

Cross Section Determination

In this chapter the ingredients needed for the determination of the cross section are described. The cross section in itself is defined in Section 9.1. A list of corrections that were applied to the Monte Carlo before extracting the signal are summarised in Section 9.2 including the details of their origin. In Section 9.3, the acceptance and QED corrections needed for unfolding cross sections are described. Afterwards, in Section 9.4 the method used in this analysis for the determination of the beauty signal from the selected data sample is presented, where the likelihood test function defined in the last chapter is used to extract the beauty fraction by fitting the Monte Carlo predictions to the data. At the end of the chapter in Section 9.5 some control distributions after using the fit results are shown for different selection stages.

9.1 Definition of the Cross section

The cross section, σ, of a given process is defined as the number of events produced in this process, N_{events}, per integrated luminosity $\mathcal{L}$:

$$\sigma = \frac{N_{\text{events}}}{\mathcal{L}}. \tag{9.1}$$

In order to measure the cross section for the process:

$$ep \to e'bX \to e'e_{\text{sl}}\nu_e X' \tag{9.2}$$

in the DIS regime, the fraction of events in the data sample coming from this process has to be determined. The cross section for beauty quark production, σ_b, is therefore determined using the following relationship:

$$\sigma_b = \frac{N^e_{\text{data}} \cdot f_b}{\mathcal{L} \cdot \mathcal{A}_b} C_r, \tag{9.3}$$

where $N^e_{\rm data}$ denotes the number of selected events in the data sample and f_b is the fraction of beauty events in the data sample determined using a fit of a likelihood test function as described in Section 9.4. $\mathcal{A}_b$ is the beauty acceptance and C_r is the QED radiative correction and are defined in Section 9.3. The single differential cross section as a function of a given variable, v, is defined accordingly:

$$\frac{d\sigma_b}{dv} = \frac{N^e_{\rm data} \cdot f_b}{\mathcal{L} \cdot \mathcal{A}_b \cdot \Delta v}\, C_r, \tag{9.4}$$

where Δv refers to the width of the given v bin.

9.2 Monte Carlo Corrections

In addition to the corrections already discussed in Chapter 7, some other corrections were applied to the Monte Carlo samples, either to only signal or background or to both, depending upon the nature and origin of the correction. In the following a list of these corrections together with their effect on the total cross section is given. Possible systematic uncertainties after applying corrections to cover the remaining discrepancies between data and Monte Carlo were also estimated and are discussed in next chapter.

9.2.1 Charm Spectrum Reweighting

The electron energy spectrum from charm decays was reweighted to agree with the measured spectrum from CLEO data [129]. The code for reweighting was adopted from the $b \to \mu$ analysis [10]. The momentum spectrum of leptons in the centre-of-mass system of charmed hadrons for PYTHIA Monte Carlo is given in Figure 9.1 and is compared with the measured spectrum of D^0 and D^+ decays. The spectrum extracted from PYTHIA shows good agreement with the measurements from CLEO data. The lepton spectrum from beauty decays was found to be in good agreement [130] with that determined from e^+e^- data and hence did not need any correction. The comparison of the total beauty cross sections with and without applying charm spectrum reweighting yielded an effect of $\sim -6\%$.

9.2.2 Lifetime Correction

It was found that the lifetimes of weakly decaying hadrons used in the Monte Carlo generators were not quite correct. The used values were taken from Jetset/PYTHIA [48] and were different from the world average values [131]. A comparison of the Jetset/PYTHIA values and PDG values for various beauty hadrons is given in Table 9.1.

In the existing Monte Carlo files a reweighting was applied for each of these hadrons using a reweighting factor $e^{-(\lambda_{\rm PDG}-\lambda_{\rm PYTHIA})\cdot ct}$, where $\lambda = 1/c\tau$ and ct is the distance between

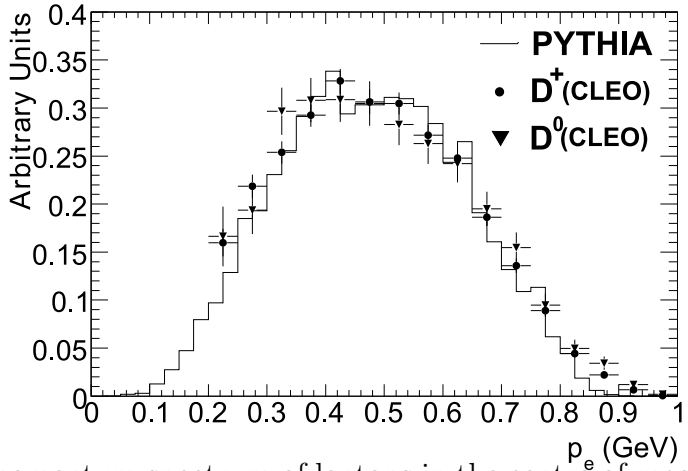


Figure 9.1: The momentum spectrum of leptons in the centre-of-mass system of charmed hadrons for PYTHIA Monte Carlo. The spectrum is compared with the spectra for D^0 and D^+ decays measured by the CLEO collaboration [129].

Hadron	$c\tau$(PYTHIA)	$c\tau$(PDG)
B^0	468 μm	457 μm
$B^{\pm}$	462 μm	491 μm
$B^{\pm}$	483 μm	441 μm

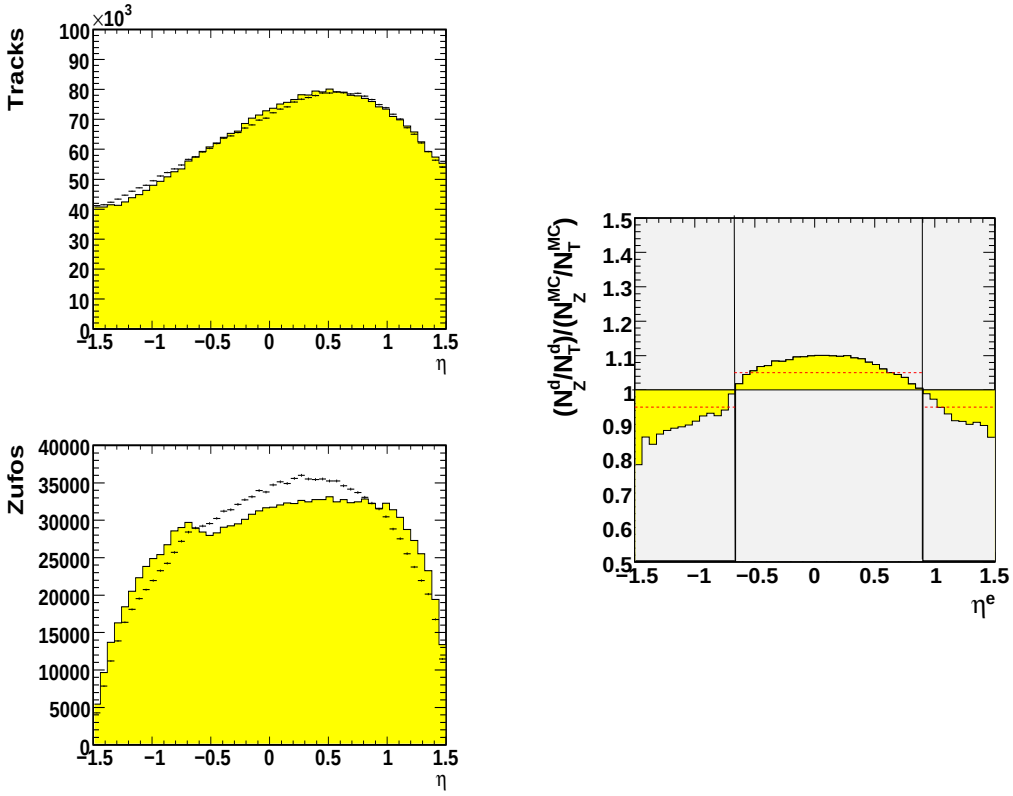
Table 9.1: Comparison of lifetime's simulated values and PDG values for different beauty hadrons.

production and decay vertex. The effect of a wrongly simulated B hadron lifetime on the total cross section was found to be $\sim -1\%$.

9.2.3 η Reweighting

As a large discrepancy between data and Monte Carlo was observed in the distribution of the electron pseudorapidity (see Figure 6.12), studies were done to check at which stage of the analysis this disagreement appears. Eta plots were made during the cut flow, where the different cuts for the pre-selection of electron candidates described in Chapter 6 were applied one by one. It was observed that the general trend of the disagreement is visible at a very early stage, where the 1:1 track-island association is required. In order to investigate whether this disagreement is caused by non-simulated background or whether it is a deficit in the description of the electron efficiency, the distribution was also checked for different stages of enrichment.

To calculate the disagreement in the island finding efficiency, the eta distribution was compared starting from track objects and then comparing this distribution with tracks



(a) **(b)**

Figure 9.2: (a) η^e distribution for all tracks (top), for tracks associated to a single island (bottom). (b) The ratio of the island finding efficiency calculated using the plots on the left. From [122].

associated to one single island (1:1 EFOs). This is shown in Figure 9.2(a). The ratio of the island finding efficiency in the data and the Monte Carlo shown in Figure 9.2(b) was used to correct the Monte Carlo. As a central correction, a factor 0.95 was taken for the regions of FCAL and RCAL and 1.05 for the BCAL region which is also shown by red dashed lines in Figure 9.2(b). This correction factor was applied to both signal and background. The effect of this correction on the total cross section was $\sim -4\%$. It was already studied in another reweighting procedure, that a reweighting of only background had a small effect on the cross section compared to the reweighting applied on both signal and background and hence for a more conservative estimate the corrections were applied to both signal and background as default.

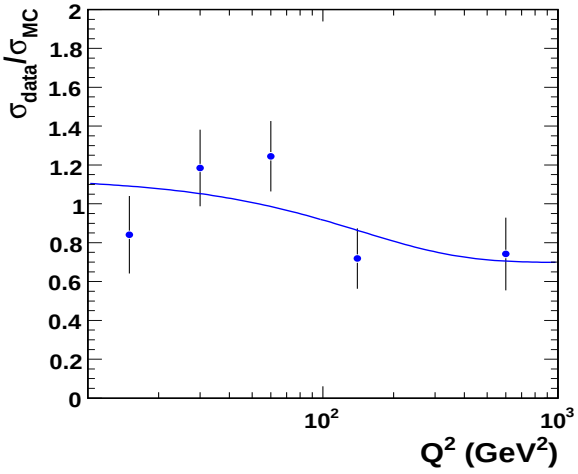


Figure 9.3: Ratio of measured cross section to LO cross section in bins of Q^2 (blue points). Fitted weighting function for beauty (blue curve).

9.2.4 Q^2 Reweighting

It was observed in this and other analyses that the measured differential cross section as a function of Q^2 in data is somewhat steeper than in the RAPGAP Monte Carlo [8]. Due to the differences between the simulation and the measured values, the Q^2 distribution was reweighted for the cross section determination. In Figure 9.3, the points show the ratio of the measured cross section to the LO cross section in bins of Q^2.

This measurement was not precise enough to decide on a proper correction since with this precision one could even fit a constant with a reasonable χ^2. Hence, the fit parametrisation was taken from the inclusive secondary vertexing analysis [8] where the precision was much better than in this case. The adopted weighting function for beauty was $\exp(-0.8298 - 0.006915Q^2) + 0.698$ and is shown by the blue curve in Figure 9.3. The effect of this correction on the total cross section was $\sim$+4%.

9.3 Acceptance and QED Corrections

After applying all the corrections mentioned above, the acceptance corrections to the measured cross sections were determined with the help of Monte Carlo simulations. The acceptance of the detector and the reconstruction software is determined using the definition:

$$\mathcal{A} = \frac{N_e^{\mathrm{rec}}}{N_e^{\mathrm{gen}}}, \tag{9.5}$$

where N_e^{rec} is the number of electrons from semileptonic decays reconstructed in the Monte Carlo sample satisfying the selection criteria detailed in Chapter 6, and N_e^{gen} is the number of electrons from semileptonic decays produced in the signal process that satisfy the kinematic requirements of the cross section definition (see Section 11.1) using the Monte Carlo information at the generator level. For the total cross section, the overall acceptance was found to be 12%.

The purities are calculated in a similar way. Here, the ratio of the number of particles which have been both reconstructed and generated in the same bin, $N_e^{\text{gen+rec}}$, to the number of reconstructed particles, N_e^{rec}, is taken as purity, $\mathcal{P}$,

$$\mathcal{P} = \frac{N_e^{\text{gen+rec}}}{N_e^{\text{rec}}}. \tag{9.6}$$

To measure differential cross sections in bins of a variable in question, the acceptance in each bin is needed for the cross section computation. In addition, the acceptances and purities give a good guidance for the choice of the binning. The binning was chosen in a way, that the typical purities are of the order of $\sim$70% and do not drop much below $\sim$50% and the acceptances are more or less flat in different bins.

The acceptances for the bins of single differential cross sections are shown in Figure 9.4. The acceptances are usually between $10-15\%$ without having too large steps. The graphs for purities for all bins of differential cross sections and acceptances for double differential cross sections can be found in Appendix C.

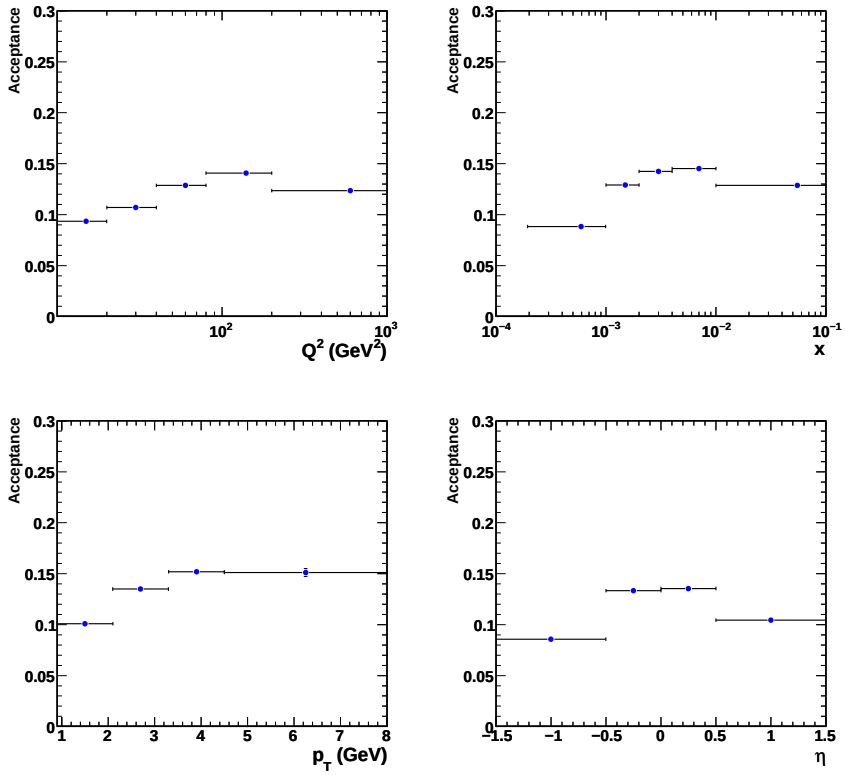


Figure 9.4: Acceptances in bins of the kinematic variables Q^2 and x (top) and in bins of the decay electron variables p_{T}^e and η^e (bottom) in beauty events.

The cross sections were also corrected to the QED Born level, calculated using a running coupling constant, α_{em}, such that they can be compared directly to the NLO QCD

predictions by HVQDIS (cf. Section 2.8.1). The radiative corrections were obtained using the RAPGAP MC as, $C_r = \sigma_{\rm born}/\sigma_{\rm rad}$, where $\sigma_{\rm rad}$ is the cross section with full QED corrections (as used in the standard MC samples) and $\sigma_{\rm born}$ was obtained with the QED corrections turned off. The corrections are typically $C_r \approx 1.05$ and go up to $C_r \approx 1.1$ for the high Q^2 region. The obtained correction factors in bins of different variables are shown in Figures C.4 and C.5 in Appendix C.

9.4 Extraction of Beauty Fraction

As described in Section 8.3, the electron candidates in the Monte Carlo samples were classified as electrons from semileptonic beauty decays, electrons from other sources[1] and non-electrons. The fraction of beauty signal was determined by performing a simultaneous fit of beauty, other electron and non-electron Monte Carlo predictions to the real data. The fit was done with the distribution of the beauty test function, T, which was computed separately for these three samples and for data. The fit range of the test function was restricted to $-2\ln T < 10$ to remove the region dominated by background and where the test function falls rapidly. The aim of the fit was to determine the relative contributions of the three subsamples f_b, f_e, f_o to the data such that:

$$N^e_{\rm data} = N^e_{\rm data} \cdot f_b + N^e_{\rm data} \cdot f_e + N^e_{\rm data} \cdot f_o, \tag{9.7}$$

$$\simeq k_b \cdot N^b_{MC} + k_e \cdot N^e_{MC} + k_o \cdot N^o_{MC}, \tag{9.8}$$

where k_b, k_e, k_o are the scaling factors with which the Monte Carlo templates have to be multiplied so that the Monte Carlo distributions sum up to the data distribution. The fractions f_b, f_e, f_o are varied within the fit so that in each bin, i, the total number of scaled MC entries is close to the number of entries for the real data. Since some bins can have very low statistics, a *binned maximum likelihood fit* was used. Additionally, the limited statistics of the Monte Carlo samples was also taken into account. For this the method developed by Barlow et. al. [126], where the shape of MC samples is not completely fixed but is allowed to vary within its uncertainty, has been used.[2]

The result of the fit is shown in Figure 9.5 and corresponds to a scaling of the cross section predicted by the beauty Monte Carlo by a factor of 1.32 ± 0.11. For the other two samples, the scaling factors were determined to be ~ 1.1 for the electron background and ~ 1.3 for the other sources. The obtained values for the fractions of all three Monte Carlo contributions and their corresponding absolute numbers: $N^{b,e}_{\rm data} = f_b \cdot N^e_{\rm data}$, $N^{e,e}_{\rm data} = f_e \cdot N^e_{\rm data}$ and $N^{o,e}_{\rm data} = f_o \cdot N^e_{\rm data}$ are given in Table 9.2. The value of χ^2 divided by the number of degrees of freedom (ndf) is 18/28 and the correlation matrix for three parameters is given in Table 9.3. It is visible that the beauty signal is well separated from the electron background and other sources. The large correlation of the electron background and other sources is due to the hard cut on $T^e_{{\rm d}E/{\rm d}x}$.

[1] Electrons from other sources also include semileptonic charm decays.

[2] This fitting method is implemented in the ROOT TFractionFitter [127, 128].

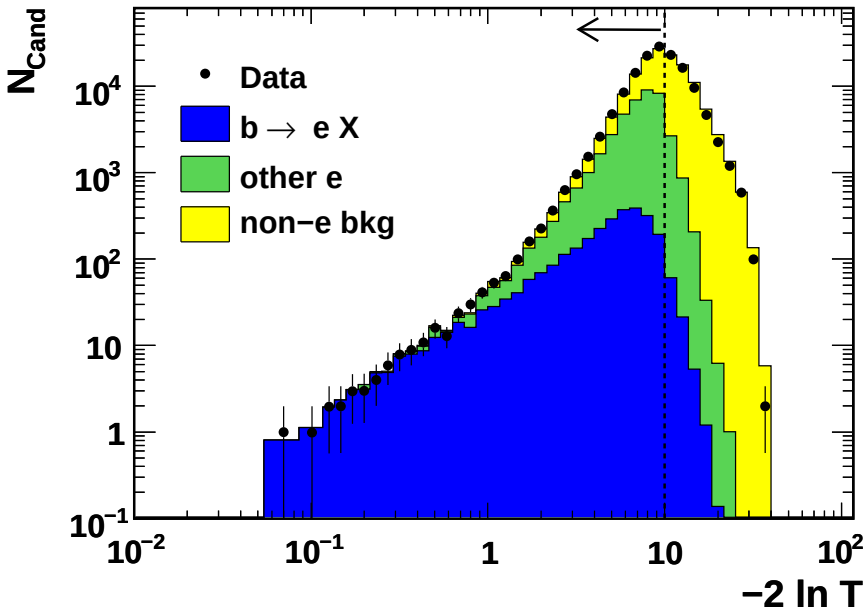


Figure 9.5: The distribution of the likelihood test function, $-2\ln T$, using the beauty hypothesis, for electron candidates, N_{Cand}, in data compared to the Monte Carlo expectation after the fit described in the text. The arrow indicates the region included in the fit ($-2\ln T < 10$). The shaded areas show the fitted contributions from b quarks (blue), the background from electrons (green) and non- electrons (yellow).

The same fit as displayed in Figure 9.5 was done in several different bins of Q^2, x, p_T^e and η^e. The corresponding fit distributions can be found in Figures E.1-E.3 in Appendix E. The obtained fit results are used for the determination of total and differential cross sections which are presented in Chapter 11

9.5 Control Plots

In the following some examples of control plots are shown to check the description for background and signal. At this stage of selection (including $-2\ln T < 10$ to the previously described selection cuts), the sample was still background dominated and hence the distribution of all electron candidates quantifies the quality of the background description. To check the quality of the signal description, the same distributions were made for a signal-enriched sample selected by applying a cut on the likelihood hypothesis at $-2\ln T < 1.5$. With this cut it was possible to obtain a beauty fraction of $\sim$70% with a remaining main contribution of electron background of $\sim$ 20%. Many distributions for the above stated two selection regions were made. Here only a selection of some plots is shown. Additional plots can be found in Appendix D.

The first set of variables (Fig 9.6) are the likelihood input variables, $p_T^{rel}, \Delta\phi$ and $d/\delta d$.

Fractions		Number of events	
f_b	0.032 ± 0.002	$N^{b,e}_{\text{data}}$	2798 ± 224
f_e	0.415 ± 0.012	$N^{e,e}_{\text{data}}$	35687 ± 1008
f_o	0.552 ± 0.011	$N^{o,e}_{\text{data}}$	47515 ± 923

Table 9.2: Fit results for f_b, f_e and f_o and the corresponding absolute numbers $N^{b,e}_{\text{data}}$, $N^{e,e}_{\text{data}}$ and $N^{o,e}_{\text{data}}$. The error of the latter is given by the statistical uncertainty of the fit.

	$b \to eX$	Other e	Bkg.
$b \to eX$	1.000	-0.570	0.389
other e	-0.570	1.000	-0.880
Bkg.	0.389	-0.880	1.000

Table 9.3: Correlation matrix for the three fit parameters.

The plots on the left are drawn for $-2\ln T < 10$ while the plots on the right include a cut on $-2\ln T < 1.5$. In general a reasonable agreement between data and Monte Carlo is observed. Differences in the shapes for different selection stages show the regions of sensitivity as well as the influence of the variable on the likelihood. The next plots shown in Figure 9.7 are the track variables, where an indirect influence of the selection is visible, e.g in p^e_{T}, where the signal electrons clearly have a harder spectrum than background particles. These variable do not enter directly into the likelihood but are of considerable interest as they define the kinematic phase space. Figure 9.7 also shows the pseudorapidity of electron candidates after applying the already discussed eta reweighting. As expected the agreement between data and Monte Carlo for the looser cuts is still not good. Since the signal enriched distribution would start to look worse for larger reweighting factors, the remaining differences are probably caused by not well described background and are taken into account as a systematic uncertainty (cf. Section 10.1.5).

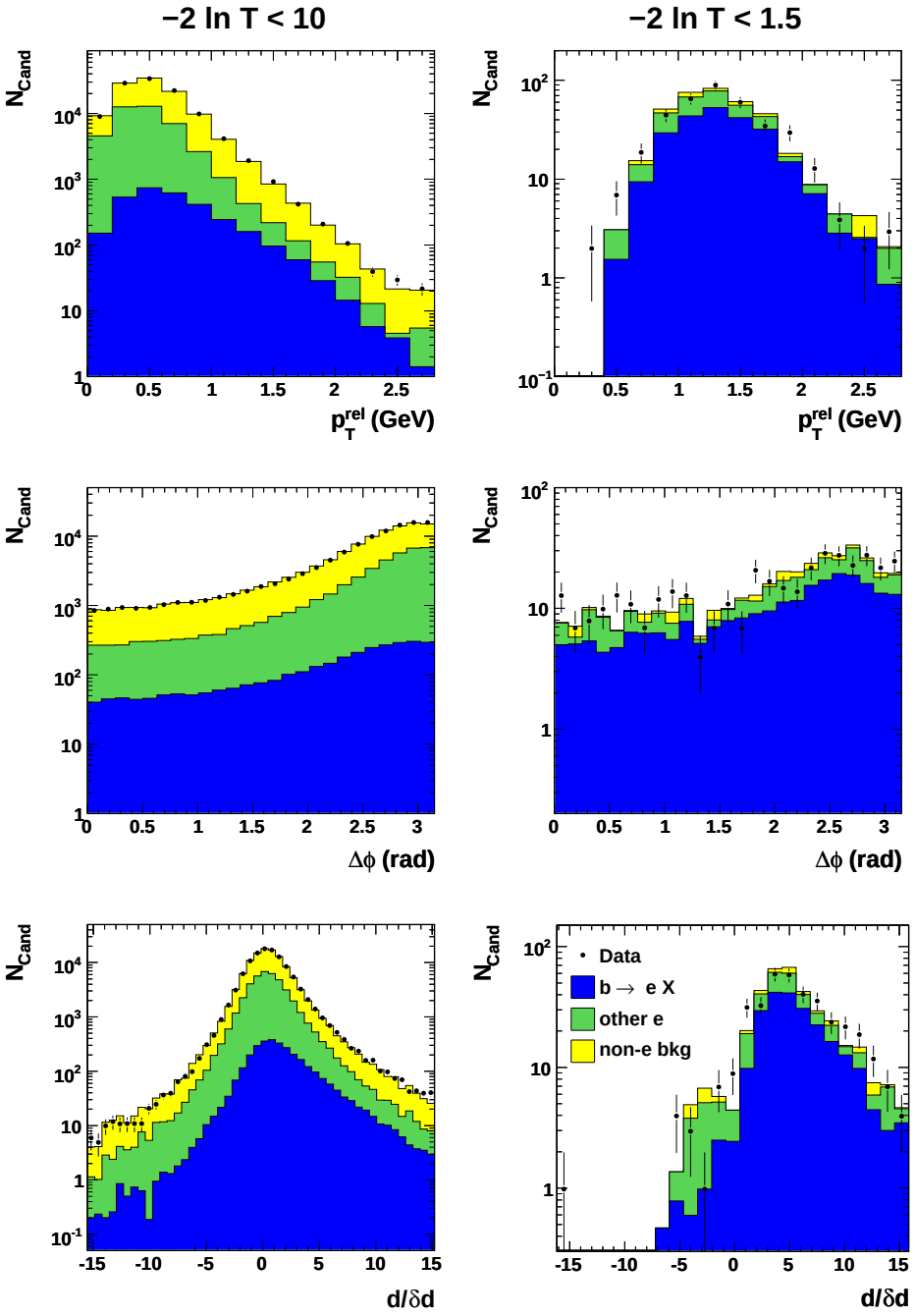


Figure 9.6: Fit variables: (a) for all candidates that enter the fit satisfying $-2\ln T < 10$, and (b) for the beauty enriched region ($-2\ln T < 1.5$). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) as denoted in the figure after applying the scale factors from the fit. The summed distribution is compared to the data distribution shown by the black points. Variables shown from top to bottom are $p_{\mathrm{T}}^{\mathrm{rel}}$, $\Delta\phi$ and $d/\delta d$.

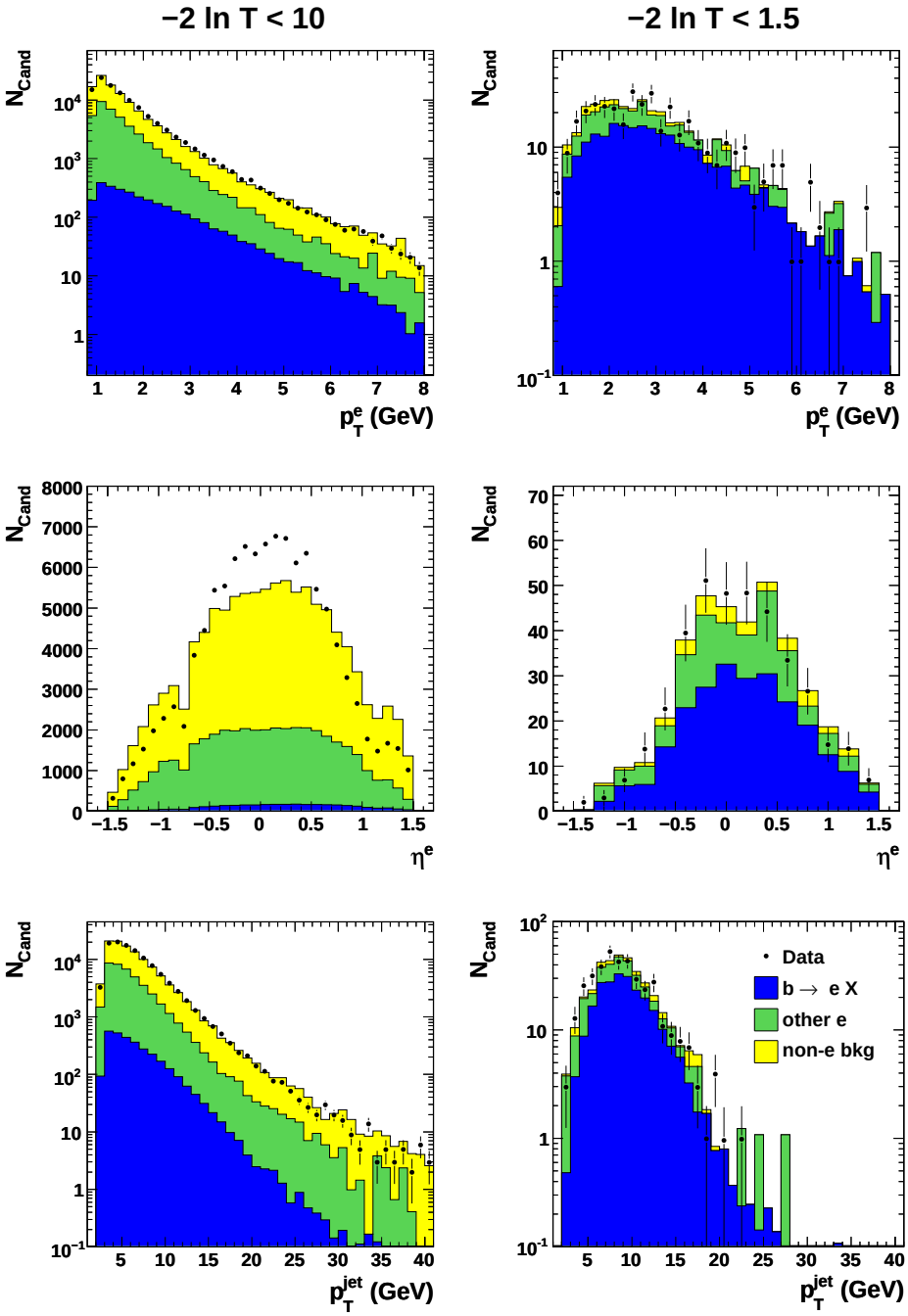


Figure 9.7: Track variables: (a) for all candidates that enter the fit satisfying $-2\ln T < 10$, and (b) for the beauty enriched region ($-2\ln T < 1.5$). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) as denoted in the figure after applying the scale factors from the fit. The summed distribution is compared to the data distribution shown by the black points. Variables shown from top to bottom are p_{T}^{e}, η^{e} and $p_{\mathrm{T}}^{\mathrm{jet}}$.

Chapter 10

Systematic Studies

To evaluate the total uncertainty on the measurements, two potential sources of error, the statistical error and systematic uncertainty have to be combined. Unlike the statistical uncertainty which depends solely on the number of selected events, N, systematic uncertainties require a thorough understanding of the detector and possible error sources. These errors could depend on the selection criteria, models used to reconstruct variables, estimation of the background events and the methods used to obtain the results. Since, the likelihood method used for the signal extraction relies on accurate modelling of the data by the Monte Carlo, differences of the input variables between the data and the Monte Carlo would also lead to systematic biases in the obtained results.

In order to estimate the systematic uncertainty for the possible sources, the analysis was repeated by varying the different parameters like selection cuts, likelihood variables and background sources within their uncertainties and the cross sections were extracted by redoing the fit to the likelihood distribution. Then the relative difference between the new cross section, σ_{sys}, and the nominal value, σ_{nom}, gives a measure of the relative systematic uncertainty, $\Delta\sigma/\sigma$, on the cross section. The changes in the cross section due to the individual contributions were added in quadrature, separately for the positive and negative deviations from the nominal cross section values, to obtain the total systematic uncertainty. The same procedure was also applied for the differential cross sections to estimate the systematic uncertainty for each analysis bin.

In the following, the details of the systematic checks performed within this analysis are presented. Evaluated relative systematic uncertainties for different sources for each cross section bin, if not shown here, can be found in Appendix F.

10.1 Systematic Uncertainties

In this section, only the systematic checks that had a significant contribution to the overall systematic uncertainty are listed. Other systematics which had a small effect on the cross

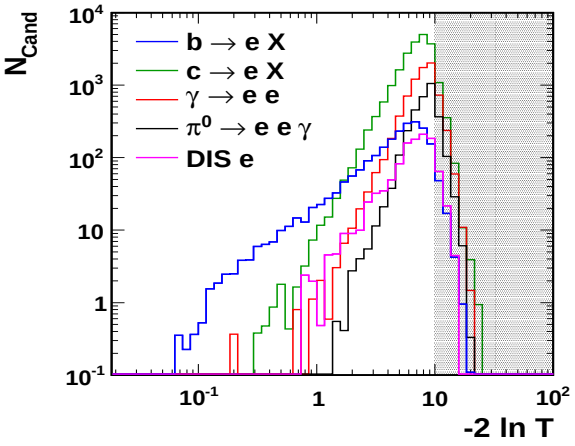


Figure 10.1: Likelihood test function, $-2\ln T$, for electrons from semileptonic b-hadron decays (blue), c-hadron decays (green), photoconversions (red), Dalitz decays (black) and mis-reconstructed DIS electrons (magenta). The distributions were extracted from the Monte Carlo samples used in this analysis.

section were treated as consistency checks and are described in Section 10.2.

10.1.1 Electron Background Variation

The likelihood distribution for electrons from semileptonic decays of b hadrons and for the four main contributions to the electron background are shown in Figure 10.1. The relative contributions of the different electron sources were varied separately to study possible systematic biases. Since the electron background has a larger probability than the non-electrons in the likelihood hypothesis, resulting in a more signal-like shape, the change in the relative contribution of the electrons in the background sample directly influences the signal fraction.

The fraction of electrons from photon conversions was varied by $\pm 25\%$. This value was adapted from photoproduction analysis [115], where this number was estimated after making studies of the control plots for the conversion candidates. The deviation in the total cross section amounted to $\pm 1.3\%$. In an analogous way the contributions from Dalitz decays and DIS electrons were varied by $\pm 25\%$ and resulted in systematic uncertainties of $\pm 0.6\%$ and $^{+1.8}_{-1.6}\%$, respectively. The quadratic sum of all these three sources of systematic uncertainties led to a $^{+2.3}_{-2.1}\%$ effect on the total cross section.

The resulting systematic uncertainty in bins of single and double differential cross sections is shown in Figure 10.2. The points show the relative deviation of the cross section from the central value for a given bin. This deviation can be compared with the red dashed lines showing the statistical uncertainty, which is typically of the order of $\sim 15-30\%$ for single differential cross sections and $\sim 30-60\%$ for the double differential cross sections. The left plot shows the deviation of the cross section as a function of bins in Q^2, x, p_{T} and η, while the right plot is for the bins of differential cross section as a function of x for four different regions of Q^2 (for the details of the binning, see Chapter 11). Overall, there is almost a constant effect in all cross section bins, except in the lowest and highest p_{T} bins where a somewhat larger deviation is visible. The low p_{T} region is sensitive to the variation of

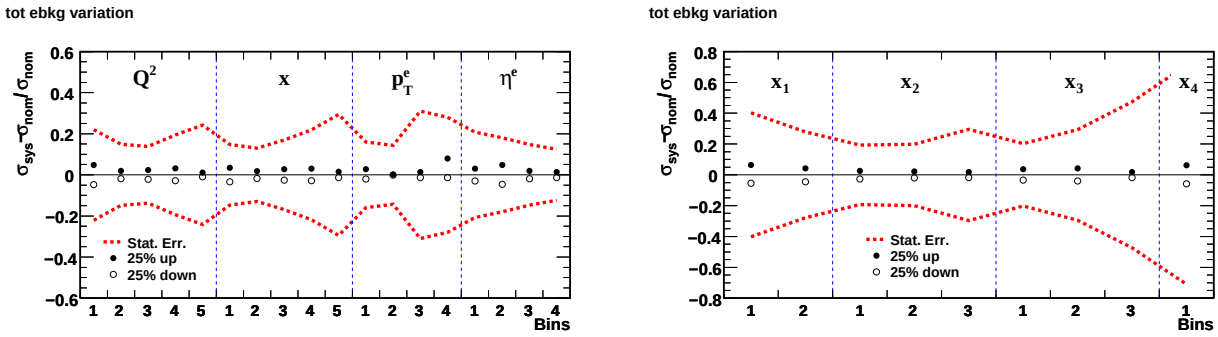


Figure 10.2: Quadratic sum of the systematic uncertainties for the variation of three electron background sources (photon conversions, Dalitz decays and DIS electron background) for single and double differential cross sections. The left plot shows the systematic uncertainty as a function of the cross-section bins for the variables, Q^2, x, p_{T}^e and η^e, and the right plot as a function of x bins for four different regions of Q^2, denoted as x_1, x_2, x_3 and x_4. The points show the relative deviation of the cross section from the central measured values for a given bin and the dashed (red) line shows the statistical error on the measurement.

conversion electrons and the highest p_{T} bin to the variation of DIS electron background.

Charm Contribution

The relative contribution of charm in the electron background sample was varied by $\pm 25\%$. The effect of this variation on the total cross section was $^{+0.9}_{-1.1}\%$ and went up to $\sim 5\%$ in a few bins of the differential cross sections (in one bin the negative deviation was of the order of 12% (see Figure 10.3), note that this bin also has a large statistical uncertainty.)

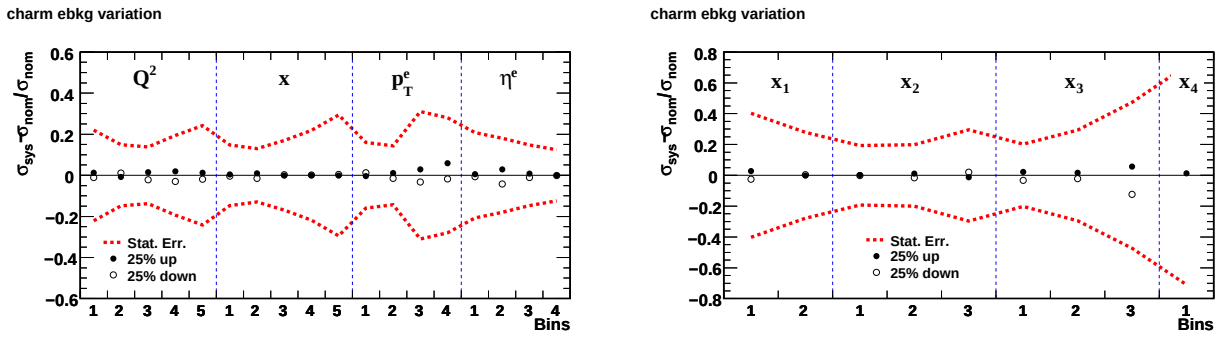


Figure 10.3: Systematic uncertainty for the variation of charm background for single and double differential cross sections. Other details as in the caption of Figure 10.2.

10.1.2 DIS Event Selection

Several selection cuts described in Section 6.3 to identify DIS events were varied. The only cuts that had a significant effect are listed here.

- **Variation of $E_{e'}$ cut:** The cut on the scattered electron energy was varied within $E_{e'} \in$ [9,11]GeV. The uncertainty on the total cross section for this variation was determined to be $^{+1.4}_{-1.0}\%$.
- **Variation of $E - p_Z$ cut:** As discussed in Section 6.3.4 a cut of $E - p_Z > 40$ GeV was used to reject PhP background and $E - p_Z < 65$GeV to reduce the overlay events. To evaluate the systematic uncertainty of these estimations, the $E - p_Z$ window was widened and tightened by ± 4 GeV. The effect on the total cross section was $^{+1.3}_{-1.3}\%$.
- **Variation of y_{jb} cut:** The cut on the inelasticity was varied within $y_{\mathrm{jb}} \in [0.04, 0.06]$ and resulted in an uncertainty of $^{+0.9}_{-0.6}\%$.

The quadratic sum of the systematic uncertainties arising from the variation of DIS selection cuts for single and double differential cross sections can be seen in Figure F.1, where some bins show quite large effects ($\sim 15 - 20\%$). Overall, the effect is substantially smaller than the statistical error.

10.1.3 Likelihood Variables

All the inputs to the likelihood were varied in a reasonable range to estimate the systematic uncertainty. The only effects having a measurable influence on the likelihood are described below. Other uncertainties and effects that were considered as consistency checks are included in Section 10.2.

Decay Length Smearing

As already shown in Chapter 7, the tails of the decay length (significance) distribution were not fully described by the Monte Carlo. A small correction to the decay lengths in the Monte Carlo was applied in order to reproduce the data, by smearing 5% of the tracks in the central region and applying an additional smearing to the tracks in the tails of the decay-length distribution. To determine a reasonable variation of the smearing, the ratio of the decay-length significance of data to Monte Carlo was checked by taking different fractions of the tracks in the central region for which the decay length was smeared. These ratio plots are shown in Figure 10.4. The largest discrepancies are visible for the first plot where no smearing is applied and the last plot where 10% of the decay lengths were smeared. Taking these variations would clearly overestimate the systematic effect. A reasonable variation range was defined to be $3 - 7\%$ smearing, where discrepancies are already visible (see 3^{rd} and 5^{th} plot). The evaluated systematic uncertainty of the smearing procedure

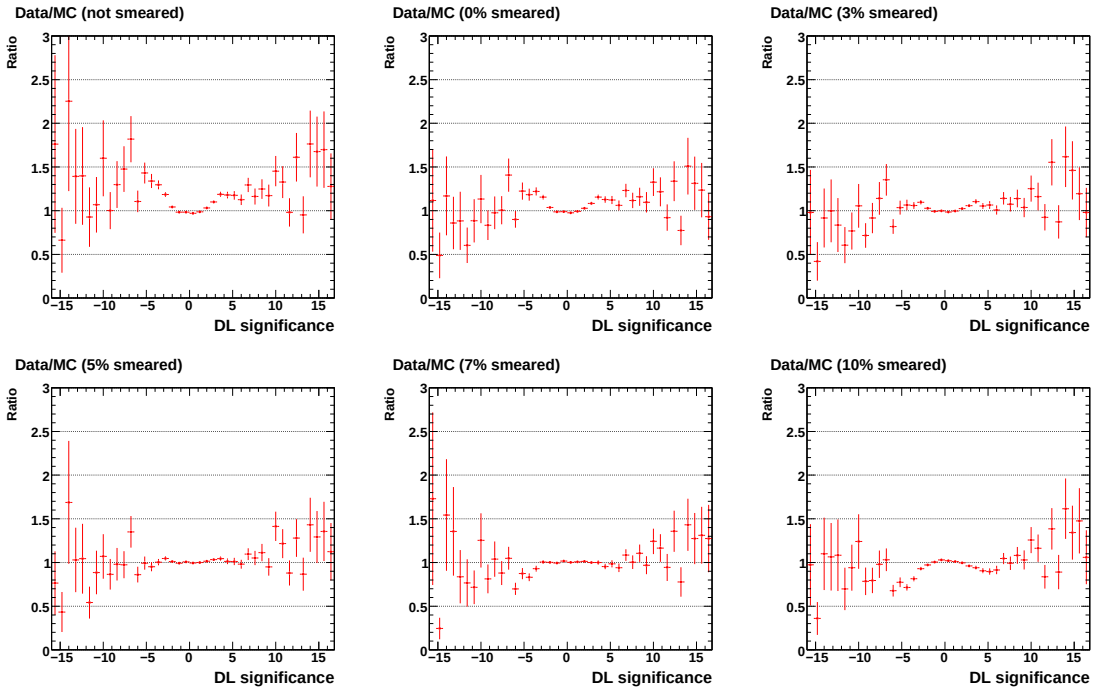


Figure 10.4: Ratios of data to Monte Carlo for the decay length significance for different smearing. For the first plot no smearing is applied, whereas in all other plots smearing is applied while taking different fractions of the decay lengths smeared by a Gaussian in the central region. From the second to the last plot fractions of $0\%, 3\%, 5\%, 7\%$ and 10% of the decay lengths were smeared. (Note that as default 5% smearing was applied (cf. Section 7.2.3)).

for the total cross section by varying the fraction of the smeared decay lengths by $\pm 2\%$ was $^{+2.6}_{-2.0}\%$. In bins of differential cross sections the uncertainties were typically $3-7\%$ as shown in Figure F.1.

$p_{\mathrm{T}}^{\mathrm{rel}}$ Shape Correction Variation

As discussed in Chapter 7, the differences between the $p_{\mathrm{T}}^{\mathrm{rel}}$ shape of the data and the Monte Carlo were corrected by applying a reweighting to the light flavour MC sample. To estimate an uncertainty due to the remaining differences between the $p_{\mathrm{T}}^{\mathrm{rel}}$ distributions, the correction factor was varied within a certain range. To define this range which covers possible discrepancies without overestimating the uncertainty, the $p_{\mathrm{T}}^{\mathrm{rel}}$ distributions and the corresponding ratio plots of data and Monte Carlo were checked after applying different amounts of the correction. As an example a comparison of the ratio between data and Monte Carlo without applying any correction (red dots) to the ones after applying four different steps between -100% and $+100\%$ to the default correction (black points) is shown

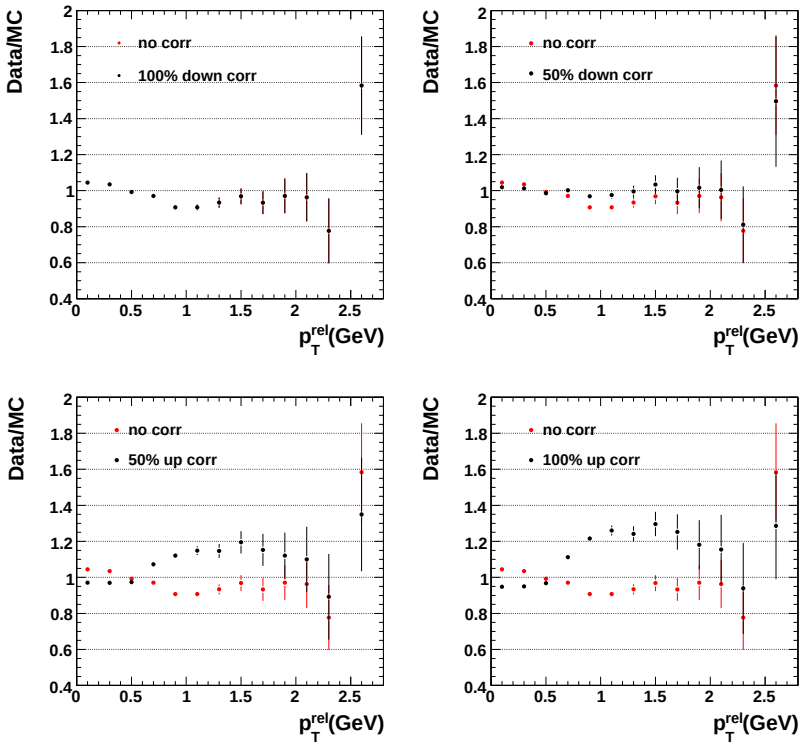


Figure 10.5: Comparison of the ratios of data to Monte Carlo for the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution without applying any correction (red dots) and after applying different amounts of correction between −100% and +100% of the default correction (black dots).

in Figure 10.5. It is clearly visible that a factor of two larger correction (4$^{\mathrm{th}}$ plot) would overestimate the uncertainty, whereas a factor of two down (2$^{\mathrm{nd}}$ plot) used in previous measurements [9, 10] might underestimate the uncertainty in some bins. It was therefore decided to take the −100% and +50% (1$^{\mathrm{st}}$ and 3$^{\mathrm{rd}}$ plots) as the variation to determine the uncertainty. The resulting uncertainty on the total cross section was $^{-1.5}_{-2.4}$% with deviations in the same direction for both variations. In such cases, to avoid a double counting and for a more conservative estimate of the uncertainty, only the larger deviation was considered. In differential cross sections no clear trend was observed (see Figure F.2), with sometimes large deviations in the same direction.

dE/dx Simulation

Both the mean and width of the Bethe Bloch distribution were varied in the Monte Carlo separately and simultaneously by one standard deviation. These two variations were then combined maximising the uncertainty on the dE/dx test function. The uncertainty corresponding to this variation was found to be ±0.4% for the total cross section. For differential

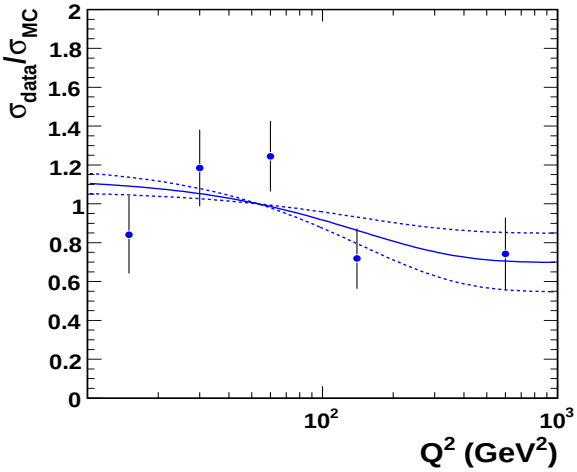


Figure 10.6: Ratio of measured cross section to LO cross section in bins of Q^2 (blue points). Fitted weighting function for beauty (solid blue curve) including the distributions for the variation (dashed blue curves).

cross sections the effect was mostly of the order of $\sim 1-3\%$ with a larger effect ($\sim 10\%$) in some bins of p_{T}^e and η^e.

10.1.4 Q^2 Reweighting

The Q^2 reweighting introduced in Section 9.2.4 was varied by a factor of two in both directions. Figure 10.6 shows the Q^2 reweighting function (solid blue curve) including the distributions for the variations by a factor of two (dashed blue curves). The systematic uncertainty on the total cross section due to this variation was $^{+2.0}_{-1.9}\%$ with an almost constant effect in differential cross sections except in high x-bins where uncertainty is of the order of 4%.

10.1.5 η Reweighting

As described in Section 9.2.3, to account for possible deficits in the simulation of the electron efficiency, the candidates were reweighted in the Monte Carlo by a factor of 0.95 in the FCAL and RCAL regions and by a factor of 1.05 in the BCAL region. These weights were varied between 0% and 10% in oder to estimate the systematic uncertainty due to the modelling of this deficit. This variation yielded an uncertainty of $^{+4.0}_{-3.7}\%$ on the total cross section. The effect in different bins of differential cross sections is quite constant but a little bit smaller in the first and last η bin. Since the first (last) bin has entries both from the RCAL (FCAL) and the BCAL and as they basically compensate the efficiency, the effect is not large in these two bins.

10.1.6 Charm Spectrum Reweighting

The reweighting of the electron spectrum to the CLEO data as described in Section 9.2.1 was varied by $\pm 50\%$, where this variation range was adopted from a previous analysis [10].

This leads to a systematic uncertainty of $^{+3.4}_{-2.9}\%$ for the total cross section. For differential cross sections the effect was of the order of $2-7\%$.

10.1.7 Tracking Efficiency

To check the effect of possible deficits in the tracking efficiency, e.g. from an imperfect simulation of the hadronic secondary interaction probability in the material of the tracking system, a reconstruction uncertainty per track of -2% in the data with respect to the Monte Carlo was assumed. The overall uncertainty due to this tracking inefficiency was then determined by randomly rejecting 2% of all tracks in the Monte Carlo and subsequently repeating the secondary vertex finding and recalculating all cross sections. The resulting uncertainty on the total cross section was -3.4%.

A similar systematic check has been made in inclusive secondary vertexing analysis [125], where some studies were done in order to justify the value of the reconstruction uncertainty of 2%. The track multiplicity in the signal enriched sample was checked for inefficiencies of $0\%, 2\%$ and 4% as shown in Figure 10.7. These plots taken from [132] show that the choice of 2% reconstruction uncertainty per track is reasonable.

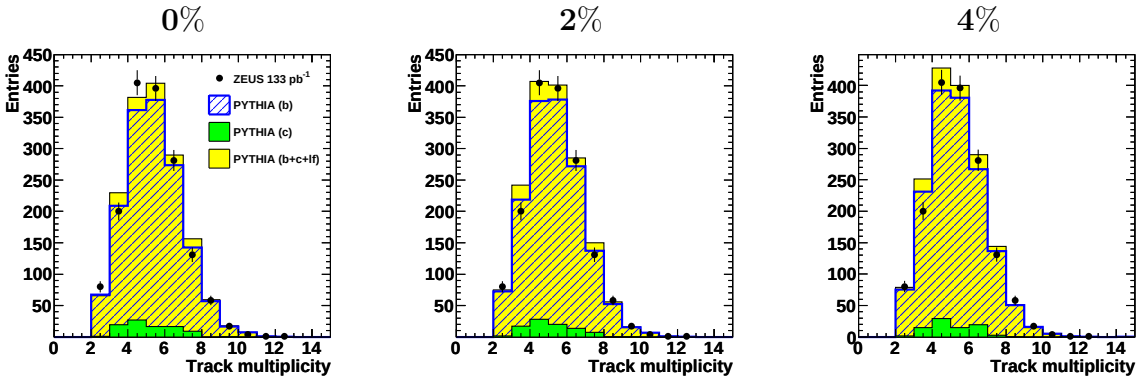


Figure 10.7: Track multiplicity for beauty enriched sample for inefficiencies of $0\%, 2\%$ and 4% (from left to right). The shaded histograms show the contributions from b quarks (blue), c quarks (green) and light flavour quarks (yellow). The summed distribution is compared to the data distribution shown by the black points. Courtesy of [132].

10.1.8 Signal Extraction

The uncertainty originating from the signal extraction procedure was evaluated by varying the range of the likelihood test function included in the fit. To find a reasonable variation of the fit range, the dependence of the cross section and the uncertainty was studied as a function of the cut on $-2\ln T$. It was possible to fit the distribution by including 3 bins

on the right and excluding 3 bins on the left from the central value. The uncertainties were found to get larger by making the hard cut on $-2\ln T$, whereas loosening the cut caused a drastic increase in χ^2/ndf. A reasonable variation of fit range was chosen where both quantities, the uncertainties and χ^2/ndf were not too large. This range was $-2\ln T \in [8.5, 11.5]$. The uncertainty on the total cross section for this variation was found to be $^{-2.4}_{-2.3}$%. The effect in bins of differential cross sections was of the order of $\sim$5% and went up to $\sim$15% in some of the bins.

10.1.9 Global Energy Scale

Previous studies [133, 134] have shown that the relative uncertainty on the measurement of the scattered electron energy and the hadronic energy scale is 2%. Therefore the final simulated electron and hadronic energy were varied simultaneously by $\pm$2% in the Monte Carlo. This systematic variation changes the shape of $E - p_Z, P_t$ and of all other variables which used the electron or hadron energy in their calculation. The effect on the total cross section was $^{-1.0}_{+1.2}$%. No clear trend was found in bins of differential cross sections, sometimes with large variations in the same direction. To estimate a reasonable uncertainty, the effect was symmeterized using the relation, $\pm|\frac{\Delta\sigma^+ - \Delta\sigma^-}{2}|$ (see Figure F.3).

10.1.10 Jet Energy Scale and Cut

The transverse jet momentum in the Monte Carlo was varied by $\pm$3% resulting in an uncertainty of $^{+1.7}_{+0.7}$% on the total cross section. This variation was different than the variation of hadronic energy scale which mainly affected $E - p_Z$ distribution. As another systematic check, the cut on $p_{\mathrm{T}}^{\mathrm{jet}}$ was varied by $\mp 0.5\,\mathrm{GeV}$ in both data and Monte Carlo. This caused an uncertainty of $^{+1.6}_{-0.04}$%. The corresponding plots to see the effect on differential cross sections can be found in Appendix F.

10.1.11 Trigger Correction

The Monte Carlo trigger efficiency was corrected for some deficits in the trigger. The Monte Carlo was reweighted assuming 5% trigger inefficiency for the events which were not fired by FLT30 and the cross sections were extracted again. The effect on the total cross section amounted to +1.2%. The differences in bins of differential cross sections as a function of Q^2 and x were a bit larger in high and low Q^2 and x bins. As expected no dependence was found in bins of p_{T}^e and η^e.

10.1.12 Luminosity Uncertainty

The uncertainty on the luminosity measurement has been determined for different data taking periods. For the data-taking in 2004 the uncertainty is 3.5% and 1.8% for 2005-

2007 data. Thus a ±2.0% overall normalisation uncertainty associated with the luminosity measurement was included. This uncertainty was applied only to the total cross section.

10.1.13 Overall Systematic Uncertainty

The total positive systematic uncertainty was calculated in a given bin by summing in quadrature all the contributions which result in an increase of the cross section[1]. Similarly the total negative systematic uncertainty was a sum in quadrature of all the contributions decreasing the cross section. An overall systematic uncertainty of $^{+7.6}_{-8.1}$% for the total cross section was found. A summary of the studied systematic effects on the total beauty cross sections is given in Table 10.1.

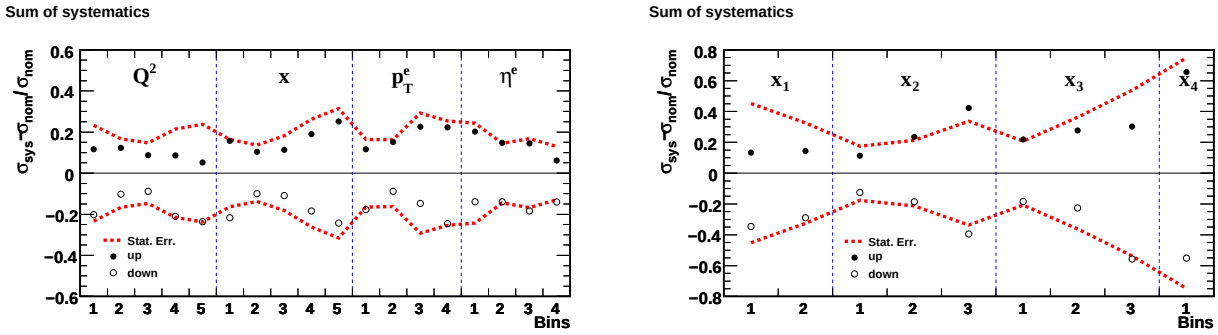


Figure 10.8: Total systematic uncertainty for single and double differential cross sections. The left plot shows the systematic uncertainty as a function of the cross-section bins for the variables, Q^2, x, p_{T}^e and η^e, and the right plot as a function of x bins for four different regions of Q^2, denoted as x_1, x_2, x_3 and x_4. The points show the relative deviation of the cross section from the central measured values for a given bin and the dashed (red) line shows the statistical error on the measurement. (For the details of the binning, see Chapter 11).

The total systematic uncertainty and its comparison to statistical error for each bin of differential cross section are depicted in Figure 10.8. In general the total systematic uncertainty is of the same order as the statistical uncertainty. Note that statistical uncertainties in Figure 10.8 are slightly different from the ones shown in Figures 10.2,10.3 and Appendix F. This was due to different reference values after applying the missing corrections (Q^2 reweighting, η reweighting and life time correction) which were discovered and corrected during the systematic studies.

[1]Note that for the cases, where both variations of a certain cut (or correction) caused a deviation in the same direction, only the larger deviation was taken into account.

Systematic sources and variations	Effect on $\sigma^{\text{vis}}_{b\to e}$
Electron background variation ($\pm 25\%$)	$^{+2.3}_{-2.1}\%$
Charm background variation ($\mp 25\%$)	$^{+0.9}_{-1.1}\%$
DIS selection ($E_{e'}, E-p_Z, y_{\text{jb}}$)	$^{+1.7}_{-1.5}\%$
Decay length smearing ($\pm 50\%$)	$^{+2.6}_{-2.0}\%$
$p_{\text{T}}^{\text{rel}}$ shape correction variation ($-100/+50\%$)	$^{-1.5}_{-2.4}\%$
$\mathrm{d}E/\mathrm{d}x$ simulation (1 σ mean and width)	$^{+0.4}_{-0.4}\%$
Q^2 reweighting ($\pm 50\%$)	$^{+2.0}_{-1.9}\%$
Eta reweighting ($\pm 100\%$)	$^{+4.0}_{-3.7}\%$
Charm spectrum reweighting ($\pm 50\%$)	$^{+3.4}_{-2.9}\%$
Tracking efficiency (2% tracks removed)	-3.4%
Signal extraction procedure (fit range ± 1.5)	$^{-2.4}_{-2.3}\%$
Global energy scale ($\mp 2\%$)	$^{+1.2}_{-1.0}\%$
Jet energy scale ($\pm 3\%$)	$^{+0.7}_{+1.7}\%$
Jet cut variation (∓ 0.5 GeV)	$^{+1.60}_{-0.04}\%$
Trigger Efficiency (-5% FLT30 in MC)	$+1.2\%$
Luminosity measurement	$^{+2.0}_{-2.0}\%$

Table 10.1: Summary of systematic uncertainties for the total beauty cross section. The first column shows different sources of systematics with the range of variation in parentheses. The second column shows the effect on the total cross section.

10.2 Consistency Checks

In addition to the systematic checks discussed above, several consistency checks were performed to test the stability of the measurement. These are described in the following.

- **Charge independence:** the independence of the signal extraction method from the particle charges was checked. For this, the cross sections were extracted separately for the positively and negatively charged particles and were found to be consistent with the default ones. As a further check, the sample was split into electron and positron running periods and then again into positively and negatively charged particles. This ensured that there was no scattered electron treated as semileptonic electron candidate in the positively charged candidates for the electron running and in the negatively charged candidates for the positron running. These four sub-samples (which were statistically independent) were fitted separately. A comparison of the beauty scaling factors for the different samples is depicted in the Table 10.2. Note that the default scaling factor listed here, differs from the final result (cf. Section 9.4),

due to the missing corrections (Q^2 reweighting, η reweighting and life time correction), which were discovered later during the systematic studies. Due to the lower statistics in the sub-samples the uncertainty was larger than for the default fit, but within the uncertainty the results were consistent.

$e^{\pm}p$ data		$e^{\pm}p$ data	
Default		pos.	neg.
1.50 ± 0.11		1.58 ± 0.16	1.48 ± 0.17
$e^{-}p$ data		$e^{+}p$ data	
pos.	neg.	pos.	neg.
1.66 ± 0.23	1.38 ± 0.25	1.36 ± 0.22	1.58 ± 0.25

Table 10.2: Comparison of the beauty scaling factors for the different samples. First value corresponds to the default value extracted from the complete sample. The second and the third values are extracted separately for the sub-samples of positively and negatively charged particles. The last four values are determined from the sub-samples of electron and positron running periods split into positively and negatively charged particles.

- **Momentum cut:** the stability of the cut on the transverse momentum of the electrons, p_{T}^{e}, was tested by varying the cut by $\pm 3\%$ in the data and Monte Carlo. The resulting effect on the total cross section was less than 1%.

- **Jet association:** as described in Section 6.4.1, the electron candidates were associated to a jet using a minimal distance $\Delta R < 1$ in the η–ϕ plane. To check the stability of this association, the cut on ΔR was varied within $[0.8, 1.5]$. For a looser cut of $\Delta R = 1.5$, the effect was not that large due to the long flat tail of the distribution (see Figure 6.9). For a cut of $\Delta R = 1.5$, it is more likely to find more than one jet which can be associated to the electron candidate. Additional studies showed that the rate of electrons associated to a wrong jet increases above $\Delta R \sim 1.3$. Therefore it was safer to set the default cut to $\Delta R = 1$. When tightening the ΔR cut, the relative effect on the cross section was a bit larger ($\sim -1\%$).

- **Selection cuts:** some selection cuts, like the Z-vertex position, box cut, cut on P_t/E_t and the pre-selection cut on the $\mathrm{d}E/\mathrm{d}x$ test function were varied within a reasonable range before repeating the analysis. The effects of all these variations on the total cross section were less than 1% and hence were not included into the systematic errors.

- **$\Delta\phi$ shape variation:** the $\Delta\phi$ correction factor applied to light flavour Monte Carlo to correct the difference in the shape between the data and the Monte Carlo was varied by a factor of two. The evaluated uncertainty was found to be $\sim \pm 0.5\%$ and was treated as a consistency check.

- **Dropping $E^{\mathrm{CAL}}/p^{\mathrm{trk}}, d_{\mathrm{cell}}$:** the particle identification method was checked by dropping discriminating variables from the likelihood hypothesis. This was done separately for $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ and d_{cell}. The resulting cross sections had somewhat larger uncertainties and were found to deviate by $\sim$2% and $\sim$1% from the default values.

Chapter 11

Cross Sections and $F_2^{b\bar{b}}$ Results

The results of the total and differential cross sections for the measurement of beauty quark production using the semileptonic electron decay channel are presented in this chapter. The measurements are compared to both RAPGAP LO+PS predictions and next-to-leading-order QCD predictions calculated with the HVQDIS programme. The extraction of the beauty contribution to the proton structure function, $F_2^{b\bar{b}}$, from the double-differential cross sections as a function of Bjorken-x and Q^2 is described in Section 11.3. At the end of the chapter a comparison of the $F_2^{b\bar{b}}$ measurement from this analysis with previous measurements and theoretical predictions is shown.

The results presented in this chapter for the beauty production measurement were approved and published as preliminary results by the ZEUS collaboration. A paper on this analysis, after small updates, has also been accepted by the collaboration and will be submitted for publication in November 2010.

11.1 Total Cross Section

As described in Chapter 9, the number of events with semileptonic decays of beauty quarks into electrons in DIS has been determined by fitting the Monte Carlo templates to the data. Using Equation 9.3, the numbers given in Table 9.2 can be directly converted into corresponding cross sections. The cross sections have been measured in the kinematic range:

$$\begin{aligned} &Q^2 > 10\,\text{GeV}^2, && 0.05 < y < 0.7 \\ &0.9 < p_T^e < 8\,\text{GeV}, && -1.5 < \eta^e < 1.5 \end{aligned}$$

using the complete HERA II data set $(2004-2007)$ corresponding to an integrated luminosity of $363\,\text{pb}^{-1}$. The total visible cross section for electrons from semileptonic decays of beauty quark was found to be:

$$\sigma_{b\to e}^{\text{vis}} = (71.8 \pm 5.5(\text{stat.})^{+5.4}_{-5.8}(\text{syst.}))\text{pb}. \tag{11.1}$$

The errors given correspond to the statistical and total systematic uncertainties (cf. Chapter 10). The measurement can be compared to the NLO QCD predictions calculated with the HVQDIS program for the same kinematic region, using the specifications introduced in Section 2.8.1:

$$\sigma_{b\to e}^{\mathrm{NLO}} = (67^{+10}_{-11})\mathrm{pb}. \tag{11.2}$$

This value agrees well with the measured cross section, which is a factor 1.3 higher than the RAPGAP leading order prediction of 54 pb.

11.2 Differential Cross Sections

Differential cross sections were determined for the kinematic variables, Q^2, x and for the decay electron variables, p_T^e and η^e using Equation 9.4. The fractions needed to calculate the differential cross sections were determined by performing the fit procedure described in Section 9.4 in bins of these variables. The resulting fit distributions are depicted in Figures E.1-E.3 in Appendix E. The results of these differential cross sections with statistical and systematic errors and their comparison with the NLO QCD predictions and RAPGAP LO predictions are shown in Figures 11.1 and 11.2.

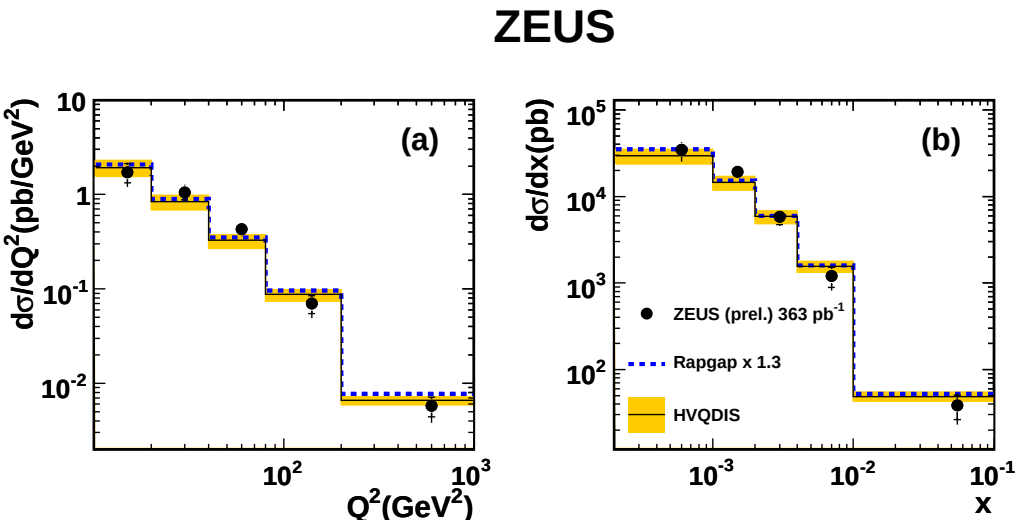


Figure 11.1: Differential cross sections for electrons from semileptonic b-quark decays as a function of the kinematic variables (a) Q^2, and (b) x. The cross sections are given for $Q^2 > 10\,\mathrm{GeV}^2$, $0.05 < y < 0.7$, $0.9 < p_{\mathrm{T}}^e < 8\,\mathrm{GeV}$ and $|\eta^e| < 1.5$. The measurements are shown as points. The inner error bar shows the statistical uncertainty and the outer error bar shows the statistical and systematic uncertainties added in quadrature. The solid line shows the NLO QCD prediction, with the theoretical uncertainties indicated by the band; the dashed line shows the scaled prediction from RAPGAP.

Q^2 (GeV^2)	$d\sigma_{b\to e}/dQ^2$ ($\mathrm{pb/GeV}^2$)	$d\sigma^{\mathrm{NLO}}_{b\to e}/dQ^2$ ($\mathrm{pb/GeV}^2$)
10 : 20	$1.73\pm0.40^{+0.20}_{-0.35}$	$1.93^{+0.37}_{-0.37}$
10 : 40	$1.05\pm0.18^{+0.13}_{-0.11}$	$0.84^{+0.13}_{-0.15}$
40 : 80	$0.428\pm0.063^{+0.037}_{-0.038}$	$0.327^{+0.050}_{-0.057}$
80 : 200	$0.070\pm0.015^{+0.006}_{-0.015}$	$0.087^{+0.011}_{-0.013}$
200 : 1000	$0.0057\pm0.0014^{+0.0003}_{-0.0014}$	$0.0066^{+0.0006}_{-0.0007}$

x	$d\sigma_{b\to e}/dx$ (pb)	$d\sigma^{\mathrm{NLO}}_{b\to e}/dx$ (pb)
0.0002 : 0.0010	$34800\pm5700^{+5500}_{-7600}$	29700^{+5400}_{-6100}
0.0010 : 0.0020	$19400\pm2700^{+2000}_{-1900}$	14700^{+2400}_{-2800}
0.0020 : 0.0040	$5800\pm1100^{+700}_{-600}$	5900^{+900}_{-1100}
0.0040 : 0.0100	$1200\pm310^{+230}_{-220}$	1560^{+220}_{-230}
0.0100 : 0.1000	$38.4\pm12.1^{+9.7}_{-9.4}$	$48.5^{+6.2}_{-5.7}$

Table 11.1: Differential cross sections for electrons from b-quark decays as a function of Q^2 and x. The cross sections are given for $Q^2 > 10\,\mathrm{GeV}^2$, $0.05 < y < 0.7$, $0.9 < p^e_{\mathrm{T}} < 8\,\mathrm{GeV}$ and $|\eta^e| < 1.5$. The first uncertainty is statistical and the second is systematic. In addition, the NLO QCD prediction and its uncertainty is given.

For a better comparison of the shape the Rapgap LO MC predictions were normalised according to the result of the global fit and are shown as the blue dashed lines. Very good agreement between the measurements and both predictions is observed. The details of the bins used in the different variables and summaries of all the cross section values and their uncertainties are listed in Tables 11.1 and 11.2.

In addition to the single-differential cross sections discussed above, double-differential cross sections as a function of x, split into four different Q^2 regions, were also extracted. The corresponding fit distributions can be found in Figures E.4-E.5 in Appendix E. The resulting cross sections with statistical and systematic errors and their comparison to the LO and NLO predictions are given in Figure 11.3. In general the data are in good agreement with both predictions, although there is a tendency for the predictions to be above the measured cross sections at high x. The cross section values and their uncertainties are summarised in Table 11.3. These cross sections were used to extract $F_2^{b\bar{b}}$ as described in the next section.

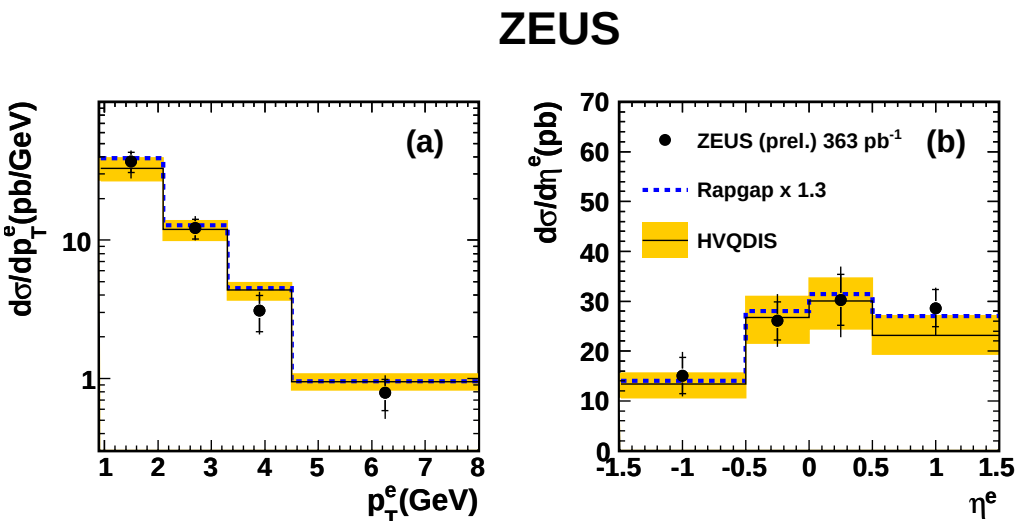


Figure 11.2: Differential cross sections for electrons from semileptonic b-quark decays as a function of the decay electron variables (a) p_{T}^e, and (b) η^e. Other details as in the caption of Figure 11.1.

p_{T}^e (GeV)	$d\sigma_{b\to e}/dp_{\mathrm{T}}^e$ (pb/GeV)	$d\sigma_{b\to e}^{\mathrm{NLO}}/dp_{\mathrm{T}}^e$ (pb/GeV)
0.9 : 2.1	$36.9\pm6.3^{+4.3}_{-6.5}$	$33.1^{+6.1}_{-6.3}$
2.1 : 3.2	$12.2\pm2.0^{+1.8}_{-1.1}$	$12.0^{+1.8}_{-2.0}$
3.2 : 4.5	$3.08\pm0.90^{+0.70}_{-0.45}$	$4.36^{+0.59}_{-0.67}$
4.5 : 8.0	$0.78\pm0.19^{+0.18}_{-0.19}$	$0.95^{+0.13}_{-0.12}$

η^e	$d\sigma_{b\to e}/d\eta^e$ (pb)	$d\sigma_{b\to e}^{\mathrm{NLO}}/d\eta^e$ (pb)
-1.5 : -0.5	$15.1\pm3.7^{+3.1}_{-2.1}$	$13.4^{+2.3}_{-2.7}$
-0.5 : 0.0	$26.0\pm3.8^{+3.9}_{-3.6}$	$26.7^{+4.3}_{-5.1}$
0.0 : 0.5	$30.3\pm5.1^{+4.4}_{-5.6}$	$30.0^{+4.7}_{-5.6}$
0.5 : 1.5	$28.6\pm3.7^{+1.8}_{-4.0}$	$23.2^{+3.9}_{-3.9}$

Table 11.2: Differential cross sections for electrons from semileptonic b-quark decays as a function of the decay electron variables (a) p_{T}^e, and (b) η^e. Other details as in the caption of Table 11.1

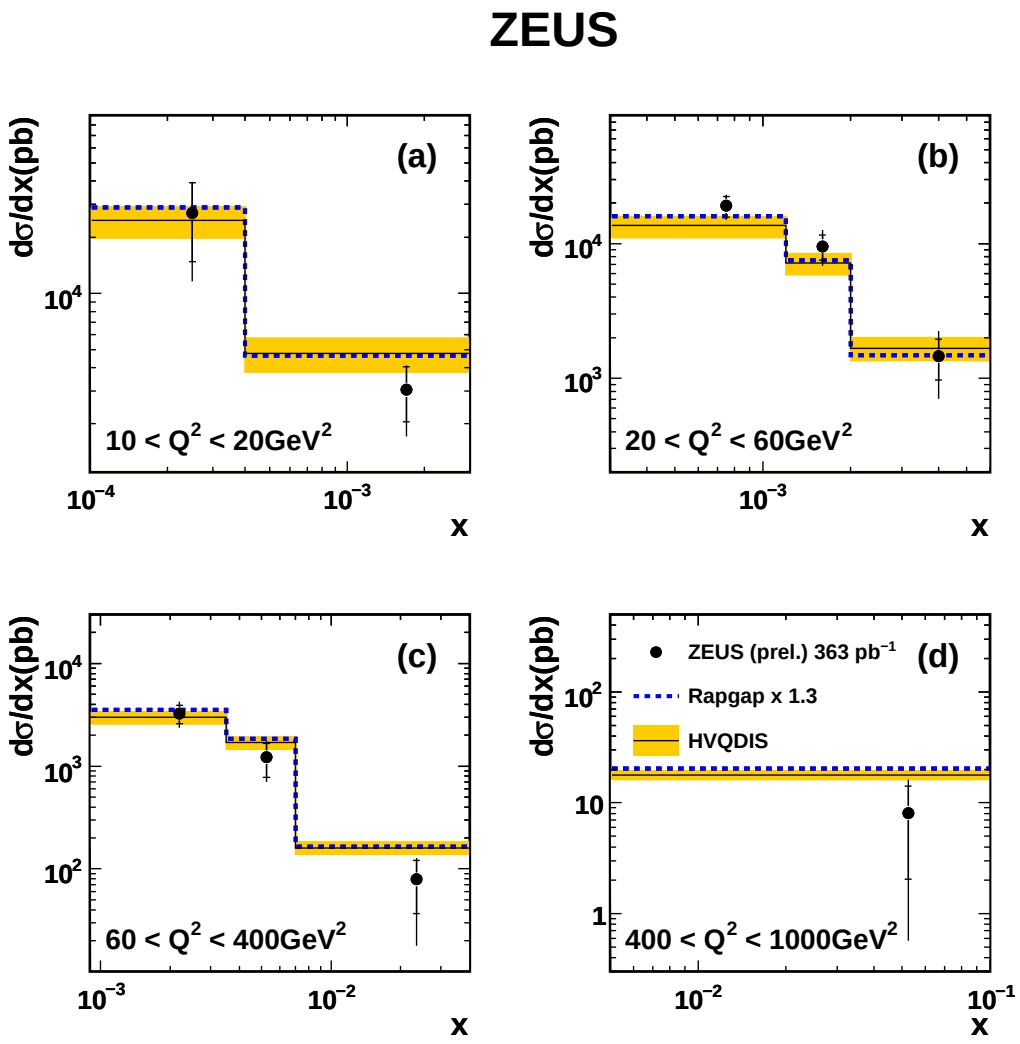


Figure 11.3: Differential cross sections for electrons from semileptonic b-quark decays as a function of x for different regions of Q^2. The Q^2 regions are $10 < Q^2 < 20\,\mathrm{GeV}^2$ for (a), $20 < Q^2 < 60\,\mathrm{GeV}^2$ for (b), $60 < Q^2 < 400\,\mathrm{GeV}^2$ for (c) and $400 < Q^2 < 1000\,\mathrm{GeV}^2$ for (d). Other details as in the caption of Figure 11.1.

Bin	Q^2 (GeV^2)	x	$d\sigma_{b\to e}/dx$ (pb)	$d\sigma^{\mathrm{NLO}}_{b\to e}/dx$ (pb)
1	10 : 20	0.0001 : 0.0004	$27000\pm12000^{+4000}_{-9000}$	25000^{+4000}_{-5000}
2	10 : 20	0.0004 : 0.0030	$3000\pm1000^{+400}_{-900}$	4800^{+1000}_{-1000}
3	20 : 60	0.0003 : 0.0012	$19100\pm3400^{+2100}_{-2400}$	13700^{+2100}_{-2600}
4	20 : 60	0.0012 : 0.0020	$9600\pm2000^{+2200}_{-1800}$	7200^{+1200}_{-1300}
5	20 : 60	0.0020 : 0.0060	$1460\pm490^{+620}_{-580}$	1660^{+320}_{-320}
6	60 : 400	0.0009 : 0.0035	$3260\pm670^{+710}_{-600}$	3010^{+340}_{-450}
7	60 : 400	0.0035 : 0.0070	$1220\pm440^{+340}_{-280}$	1700^{+200}_{-240}
8	60 : 400	0.0070 : 0.0400	$79\pm42^{+24}_{-44}$	160^{+22}_{-22}
9	400 : 1000	0.0050 : 0.1000	$8.1\pm6.1^{+5.3}_{-4.5}$	$17.7^{+1.4}_{-1.6}$

Table 11.3: Differential cross sections for electrons from semileptonic b-quark decays as a function of x for four different Q^2 ranges. The cross sections are given for $Q^2 > 10\,\mathrm{GeV}^2$, $0.05 < y < 0.7$, $0.9 < p_{\mathrm{T}}^e < 8\,\mathrm{GeV}$ and $|\eta^e| < 1.5$. The first uncertainty is statistical and the second is systematic. In addition, the NLO QCD prediction and its uncertainty is given.

11.3 Extraction of $F_2^{b\bar{b}}$

The beauty contribution to the proton structure function, F_2, denoted as, $F_2^{b\bar{b}}$, can be defined in terms of the inclusive double-differential cross section as a function of x and Q^2:

$$\frac{d^2\sigma_b}{dxdQ^2} = \mathcal{K}[F_2^{b\bar{b}}(x,Q^2) - \frac{y^2}{Y_+}F_L^{b\bar{b}}(x,Q^2)], \tag{11.3}$$

where $\mathcal{K} = Y_+\frac{2\pi\alpha_{em}^2}{xQ^4}$ and $Y_+ = 1+(1-y)^2$. The measured cross sections were corrected for the longitudinal structure function, $F_L^{b\bar{b}}$, to extract the structure function $F_2^{b\bar{b}}$. The small correction for $F_L^{b\bar{b}}$ was taken into account in the HVQDIS prediction.

The inclusive beauty cross section is obtained by measuring the visible beauty cross section using semileptonic electron decay channel in the kinematic range ($0.9 < p_T^e < 8\,\mathrm{GeV}, |\eta^e| < 1.5$), extrapolating to the full kinematic phase space and then employing the semileptonic beauty fraction $f(b\to e)$ to derive the total cross section.

The electron cross sections, $\sigma_{b\to e}$, measured in bins of x and Q^2 (cf. Figure 11.3) were used to extract $F_2^{b\bar{b}}$ at a reference point in the x–Q^2 plane using:

$$F_2^{b\bar{b}}(x_i,Q_i^2) = \sigma_{b\to e}\frac{F_2^{b\bar{b},\mathrm{NLO}}(x,Q^2)}{\sigma^{\mathrm{NLO}}_{b\to e}}, \tag{11.4}$$

where the theoretical cross section, $\sigma_{b\to e}^{\mathrm{NLO}}$ and the values of $F_2^{b\bar{b}}$ were calculated in the FFNS using the HVQDIS program. In this calculation, the same parton densities, beauty mass ($m_b = 4.75\,\mathrm{GeV}$), and factorisation and renormalisation scales ($\sqrt{Q^2+4m_b^2}$) were used as for the NLO predictions for the single- and double-differential cross sections. The uncertainty on the extrapolation from the measured range to the full kinematic phase space was estimated by varying the settings of the calculation (see Section 2.8.1) for $F_2^{b\bar{b},\mathrm{NLO}}/\sigma_{b\to e}^{\mathrm{NLO}}$ and adding the resulting uncertainties in quadrature.

Figure 11.4 shows the kinematic plane in x–Q^2 for this analysis. The boxes correspond to the x–Q^2 binning used to extract the cross sections. For each bin a reference point in x and Q^2 was defined (see red dots in Figure 11.4) to calculate the structure function. The reference points used in previous ZEUS publications [10, 9] are also shown.

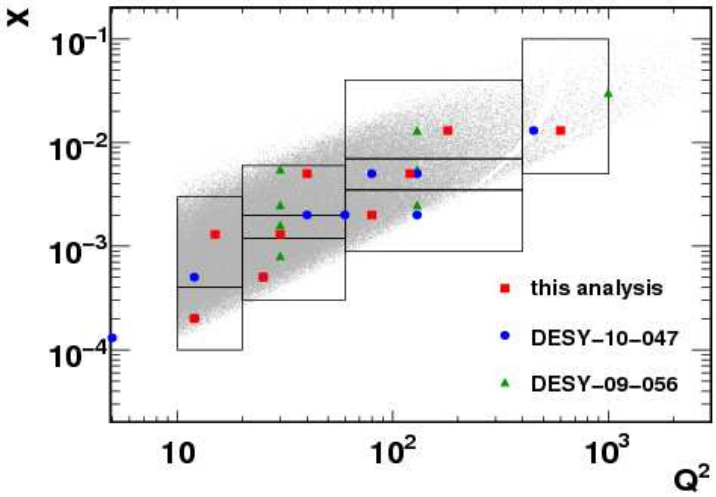


Figure 11.4: Kinematic plane in x–Q^2 for this analysis. The boxes correspond to the x–Q^2 binning used to extract the cross sections and the red dots refer to the selected reference points for the definition of $F_2^{b\bar{b}}$. The blue and green points correspond to the reference points of the previous ZEUS publications [9] and [10], respectively.

The measurements of $F_2^{b\bar{b}}$ as a function of x for nine different values of Q^2 are shown in Figure 11.5. The values and the corresponding uncertainties are given in Table 11.4. The measurements are compared with the NLO QCD prediction and the previous measurements discussed in Chapter 3 [9, 10, 60]. To compare the results presented here, with the published measurements, the earlier results were extrapolated to the Q^2 values chosen for this analysis. For $Q^2 > 10\,\mathrm{GeV}^2$, this measurement represents the most precise determination of $F_2^{b\bar{b}}$ by the ZEUS collaboration. It is in good agreement with previous ZEUS analyses and the H1 measurement. The NLO QCD prediction describes the data well.

The structure function is also presented in Figure 11.6 as a function of Q^2 for fixed values

of x. This figure was taken from a previous ZEUS measurement [9] in order to compare the new results with the published ZEUS and H1 points and different theory predictions (see Figure 11.6). For the HVQDIS prediction, a different scale parametrisation, $\mu = \frac{1}{2}\sqrt{Q^2 + p_{\mathrm{T}}^2 + m_b^2}$, was used as in [9].

The presented analysis is not sensitive to the low-Q^2 region covered by the HERA I muon analysis but in the region of $10 \lesssim Q^2 \lesssim 100\,\mathrm{GeV}^2$ the precision of the measurement is better than the previous ZEUS analyses, and in very good agreement with the H1 results. For higher values of Q^2 the measured points are on the lower side of the theories but are still consistent with the other measurements. All the theoretical predictions shown provide reasonable description of the data.

Bin	Q^2 (GeV)	x	$F_2^{b\bar{b}}$
1	12	0.0002	$0.0072 \pm 0.0033^{+0.0010+0.0012}_{-0.0026-0.0015}$
2	15	0.0013	$0.0021 \pm 0.0007^{+0.0003+0.0004}_{-0.0006-0.0004}$
3	25	0.0005	$0.0152 \pm 0.0027^{+0.0017+0.0025}_{-0.0019-0.0029}$
4	30	0.0013	$0.0110 \pm 0.0023^{+0.0026+0.0019}_{-0.0020-0.0021}$
5	40	0.005	$0.0041 \pm 0.0014^{+0.0018+0.0009}_{-0.0016-0.0007}$
6	80	0.002	$0.0208 \pm 0.0043^{+0.0045+0.0029}_{-0.0038-0.0032}$
7	120	0.005	$0.0110 \pm 0.0040^{+0.0030+0.0015}_{-0.0025-0.0015}$
8	180	0.013	$0.0050 \pm 0.0027^{+0.0015+0.0006}_{-0.0028-0.0006}$
9	600	0.013	$0.0089 \pm 0.0067^{+0.0058+0.0008}_{-0.0049-0.0008}$

Table 11.4: The structure function $F_2^{b\bar{b}}$ given for nine different values of Q^2 and x. The first error is statistical, the second systematic and the last is the extrapolation uncertainty.

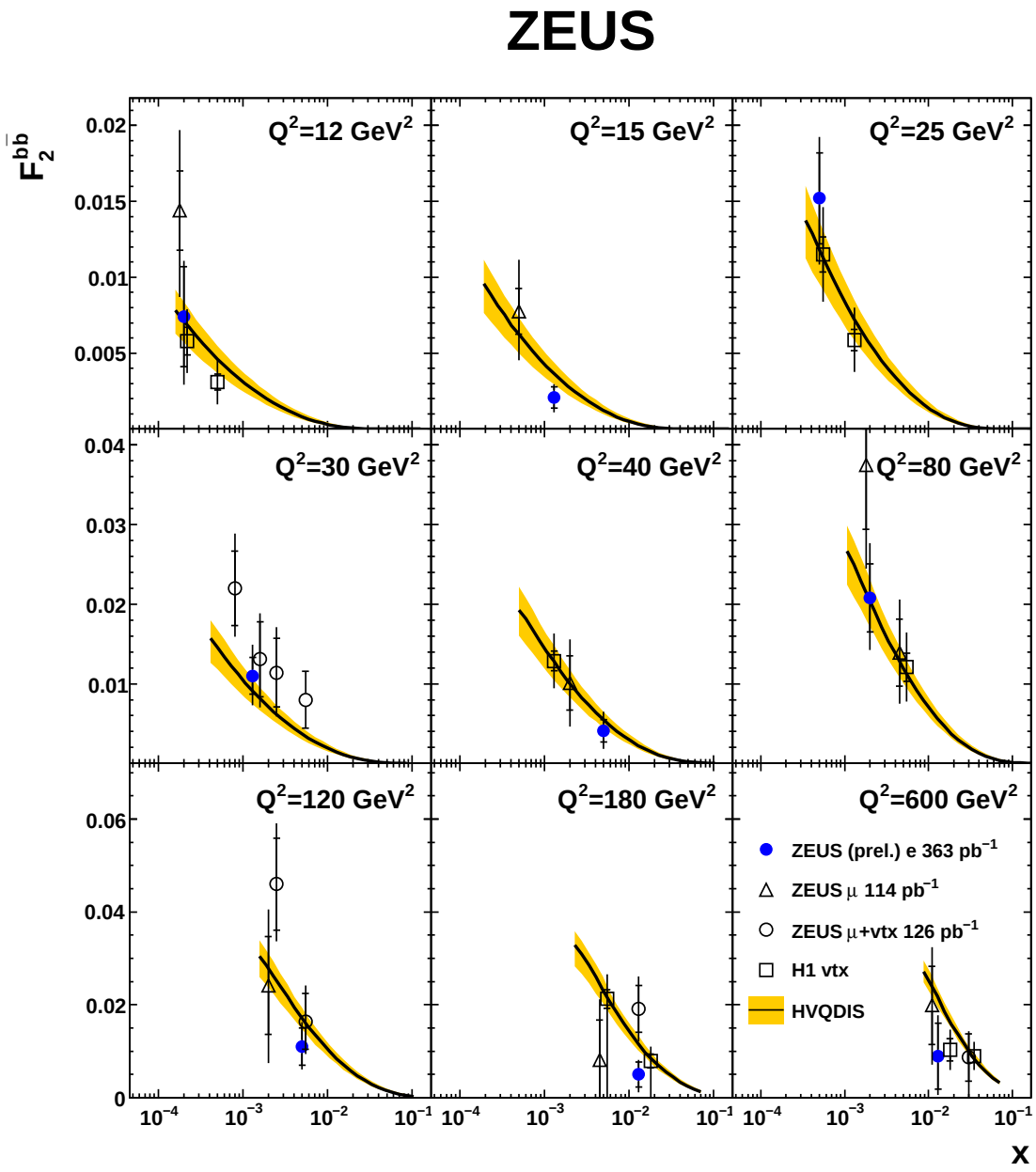


Figure 11.5: The structure function $F_2^{b\bar{b}}$ (filled symbols) as a function of x for nine different values of Q^2 compared to previous results (open symbols). The inner error bars are the statistical uncertainty while the outer error bars represent the statistical, systematic and extrapolation uncertainties added in quadrature. The band represents the uncertainty on the NLO QCD prediction. Previous data have been corrected to the reference Q^2 range of this analysis, given in Figure 11.3.

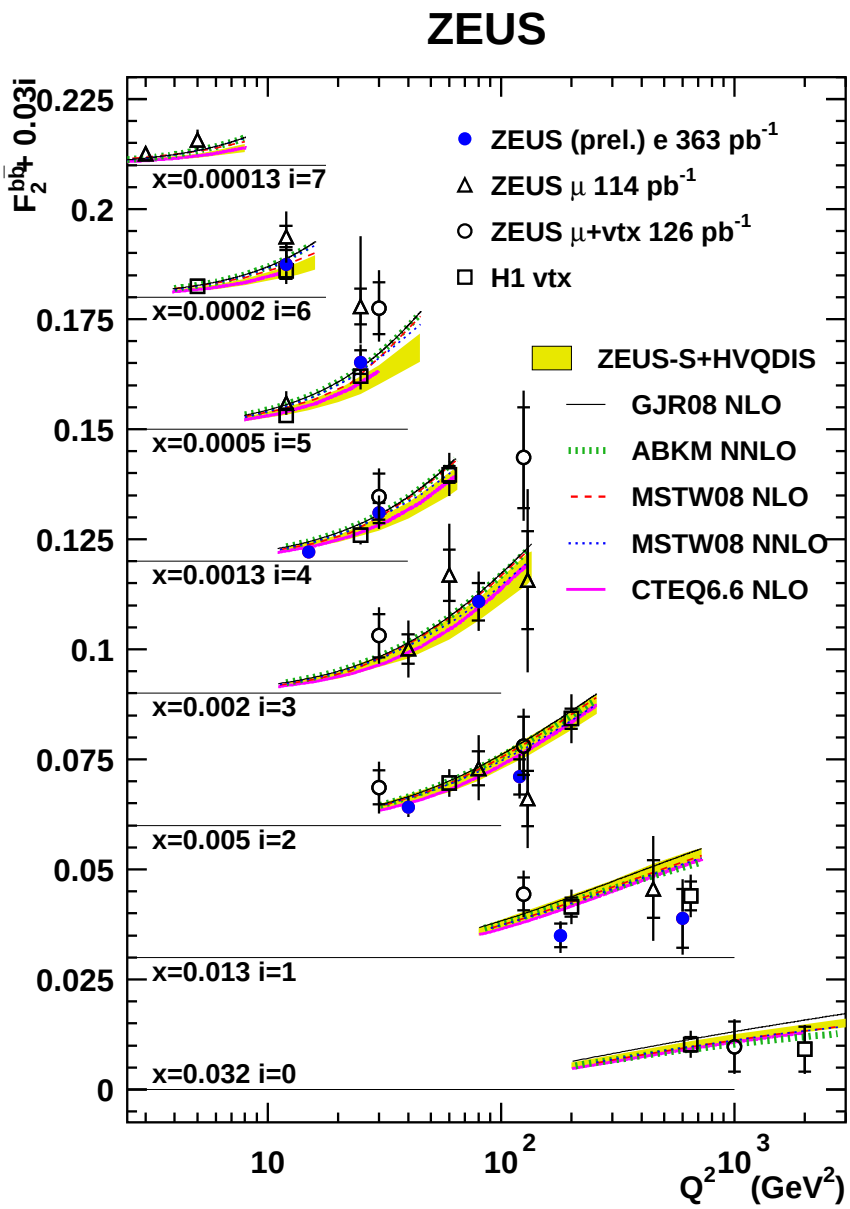


Figure 11.6: The structure function $F_2^{b\bar{b}}$ (filled symbols) as a function of Q^2 for fixed values of x compared to previous results (open symbols). The inner error bars are the statistical uncertainty while the outer error bars represent the statistical, systematic and extrapolation uncertainties added in quadrature. The data have been corrected to the same reference x of the previous analysis [9]. The measurements are compared to several NLO and NNLO QCD predictions (see text for details).

Chapter 12

Summary and Conclusions

The main objective of this thesis was the measurement of beauty production in ep collisions at HERA at a centre-of-mass energy $\sqrt{s} = 318\,\text{GeV}$. The analysis was performed using $e^{\pm}p$ collision data recorded with the ZEUS detector in the period 2004–2007, corresponding to an integrated luminosity of $\mathcal{L} = 363\,\text{pb}^{-1}$. For the identification of beauty events, the semileptonic decays of the b hadrons into electrons or positrons were used. Kinematically, the analysis was restricted to deep inelastic scattering with four momentum exchange $Q^2 > 10\,\text{GeV}^2$ which, in addition to the large mass of the b quark, provides a hard scale for the comparison to NLO QCD calculations.

In a first step, DIS events were selected which required a scattered electron to be detected in the calorimeter. Several cleaning cuts were applied to increase the purity of the sample. After getting a clean DIS sample, further selection cuts were applied to enrich the b-quark component by finding candidates for semileptonic decays to electrons. Using the electron channel allows the decay lepton to be measured at lower transverse momentum and provides a complementary method to previous analyses using decays into muons. The muon decay channel was widely used in previous measurements at ZEUS, because of the simpler muon identification. However, the muons have to be energetic enough to penetrate the calorimeter and reach the muon chambers and are therefore required to have transverse momentum larger than $\sim 2\,\text{GeV}$. The accurate determination of muon efficiency as a function of transverse momentum is difficult. The electrons are detected using the inner components of the detector (tracking chamber and calorimeter) and the momentum needed to measure electrons is much less. In the analysis presented here, electrons with transverse momentum down to 0.9 GeV were accepted.

To extract the beauty content of the sample, a likelihood method developed during a previous photoproduction analysis [115] was used and adapted to the different kinematics of the DIS regime. This method involves particle and decay identification using variables sensitive to electron identification as well as to semileptonic decay kinematics and then combines the information of these variables into one discriminating variable using a likelihood hypothesis. In the context of this analysis, several discriminating observables were

studied to perform the particle and decay identification. A selection of variables was done depending on their separation power and on the quality of description of the data by the Monte Carlo simulation.

The variable, $\mathrm{d}E/\mathrm{d}x$ was found to be a powerful tool for the particle identification and was combined with $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ and d_{cell} as an input to the likelihood test function to separate electrons from non-electrons. To distinguish the different origins of electron candidates, the likelihood test function was extended by using variables sensitive to different aspects of heavy quark decays. Two of these variables were $p_{\mathrm{T}}^{\mathrm{rel}}$ and $\Delta\phi$, which were already used in the previous photoproduction analysis [115] based on the HERA I dataset. The presence of the MVD in the HERA II running period improved the overall precision of the tracking system and allowed the lifetime information to be used. Several variables related to the lifetime information were tested and the significance of the measured decay length of weakly decaying b hadrons, $d/\delta d$, was used as an additional variable in the likelihood test function. This variable was the most important variable for the beauty identification. The fraction of beauty signal was determined by fitting the Monte Carlo templates to the real data distribution of the likelihood test function.

The total visible cross section for semileptonic decays to electrons from the decay of b hadrons was measured in the kinematic region $Q^2 > 10\,\mathrm{GeV}^2$, $0.05 < y < 0.7$, and $0.9 < p_{\mathrm{T}}^e < 8\mathrm{GeV}$, $|\eta^e| < 1.5$. This yielded $\sigma_{b\to e} = 71.8 \pm 5.5(\mathrm{stat.})^{+5.4}_{-5.8}(\mathrm{syst.})\mathrm{pb}$. This cross section agrees well with the theoretical HVQDIS NLO prediction for the same kinematic region of $67^{+10}_{-11}\mathrm{pb}$, but is a factor 1.3 higher than the one predicted by the Rapgap MC using the LO matrix elements with DGLAP parton showers. This underestimation of beauty production is typical for this kind of MC generator and was also seen in other analyses [10, 9].

A comparison of the error on the total visible cross section from this measurement to the previous measurements using the muon decay channel [10, 9] reveals that, while the three measurements have similar statistical precision, the systematic uncertainty for the measurement using the electron decay channel, is about half of that of the muon measurements, resulting in an improvement in overall precision for the beauty production cross section. The dominant contributions to the larger systematic uncertainties for the muon measurements were from the variation of $p_{\mathrm{T}}^{\mathrm{rel}}$ shape correction and muon efficiency. The use of the lifetime information and the combination of several variables in the likelihood hypothesis, in the electron analysis, reduced the uncertainty due to the variation of $p_{\mathrm{T}}^{\mathrm{rel}}$ shape correction, considerably.

Differential cross sections as a function of the kinematic variables Q^2, x and the decay electron variables p_T^e and η^e were measured. The scaled Rapgap Monte Carlo provides a good description of the shape of the differential distributions. The results are well described by the NLO QCD predictions calculated using the HVQDIS program. This is different from one of the previous measurements [10] of beauty production using semileptonic decays into muons, where the beauty cross sections were about a factor two above the central NLO QCD predictions, albeit with large uncertainties.

The differential cross sections as a function of x split into different Q^2 regions were also

measured. In all distributions the data are well described by both the predictions from the LO+PS Monte Carlo simulations as well as the NLO QCD calculations.

These cross sections were used to extract the structure function, $F_2^{b\bar{b}}$. The measurement is in agreement with the results obtained from previous analyses using different techniques. For $Q^2 > 10\,\mathrm{GeV}^2$, this measurement represents the most precise determination of $F_2^{b\bar{b}}$ by the ZEUS collaboration. The results were also compared to several NLO and NNLO QCD calculations, which provide a reasonable description of the data.

The results presented in this thesis for the beauty production measurement have been made public by the ZEUS collaboration and been shown at international conferences. A paper on this analysis, after small updates, has also been approved by the collaboration and is currently almost through the publication procedure.

This measurement is the first measurement for beauty production in DIS using semileptonic decays into electrons using the full HERA II dataset. This is a complementary analysis technique and provides a very valuable measurement of $F_2^{b\bar{b}}$ with independent systematics and a cross-check of other analyses. This measurement will serve as an important input for the long-term aim to produce a combined result for the $F_2^{b\bar{b}}$ measurement including the results obtained with different methods from both HERA experiments, ZEUS and H1.

Appendix A

Trigger Definitions

In this appendix, trigger logic at the first-level and second-level triggers are briefly described. The definitions of the third-level trigger slots used to select DIS events in this analysis are given in Section A.3.

A.1 First-Level Trigger

At the first-level trigger a selection of different slots, mainly based on calorimeter information, were used. Over 10 FLT slots were used by each SLT slot.

Generally an FLT slot (e.g. FLTDIS) used to select DIS events comprised a logical OR of a number of different quantities, such as the EMC energy in the calorimeter, an isolated electron candidate (selected using an isolated energy deposit with certain amounts of energy in the EMC and HAC), the total transverse energy in the CAL, and a track in the CTD. The time measurement of the energy deposit in the C5 counters, the veto wall and the SRTD were also used to veto events.

Different slots need different combinations of these quantities with different cuts, optimised to ensure an efficient selection of DIS events while rejecting as much background as possible.

A.2 Second Level Trigger

At the second-level trigger, events had to pass the inclusive DIS slot SPP1 or one of the eight SLT slots (SLT 01, 02, 03, 04, 05, 06, 07 or 08) for further consideration by the TLT slots. The requirements for an SLT slot e.g. SPP1 are:

SPP1: Inclusive DIS

- FLTDIS or FLT36

- $E - p_Z > 30\,\text{GeV}$
- $E_{\text{REMC}} > 2.5\,\text{GeV}$ or $E_{\text{BEMC}} > 2.5\,\text{GeV}$ or $E_{\text{FEMC}} > 10\,\text{GeV}$ or $E_{\text{FHAC}} > 10\,\text{GeV}$
- $E_e > 5\,\text{GeV}$

The requirements for the other SLT slots were similar to SPP1. They all required a minimum value of $E - p_Z$, required some energy in the RCAL or BCAL and needed to pass a selection of FLT slots. Further requirements for different slots included cuts on the total transverse energy of the event, E_T, the event vertex, the scattered electron energy and some basic track requirements.

A.3 Third-Level Trigger

The third-level trigger slots consist of cuts on event quantities calculated at the third level and require combinations of different second-level trigger slots. The trigger slots used in the analysis presented in this thesis are defined in the following.

SPP01: Inclusive medium Q^2 DIS

- SLT SPP1
- $30\,\text{GeV} < E - p_Z < 100\,\text{GeV}$
- $E_{e'} > 4\,\text{GeV}$
- $R > 30\,\text{cm}$ for RCAL electrons

SPP02/SPP09: Inclusive low Q^2 DIS

- SLT SPP1
- $30\,\text{GeV} < E - p_Z < 100\,\text{GeV}$
- $E_{e'} > 4\,\text{GeV}$
- Box cut $12 \times 12\,\text{mm}^2/15 \times 15\,\text{mm}^2$

DIS03: Medium Q^2

- Pass one of the SLT DIS slots
- $E_{e'} > 4\,\text{GeV}$
- $R > 35\,\text{cm}$ for RCAL electrons

- $30\,\text{GeV} < E - p_Z < 100\,\text{GeV}$

DIS04: High Q^2

- Pass one of the SLT DIS slots
- $E_{e'} > 7\,\text{GeV}$
- $40 \times 40\,\text{mm}^2$ box cut for RCAL electrons
- $30\,\text{GeV} < E - p_Z < 100\,\text{GeV}$

HFL17: NC DIS

- SLT SPP1
- NC DIS with 2 tracks

HFL10: DIS semileptonic e

- Pass one of SLT DIS slots
- $|Z_{\text{vtx}}| < 80\,\text{cm}$
- At least 2 tracks
- $30\,\text{GeV} < E - p_Z < 100\,\text{GeV}$
- DIS electron with $E_{e'} > 7\,\text{GeV}$
- Find a track with $p_{\text{T}} > 1.0\,\text{GeV}$, $0.6 < \theta < 2.55$
- Track within 30 cm of an island ($E > 0.5\,\text{GeV}, 90\%$ of energy in EMC)
- 2nd electron candidate (from $\text{d}E/\text{d}x$)

Appendix B

Conversion Finder

As described in Section 6.4, in addition to non-electrons like pions, protons and kaons, electrons from sources other than semileptonic decays also contribute to the background. One of the main source of this type of background is electrons from photon conversions. Photons are produced copiously over a wide energy range, mostly from the dominant decay of π^0 mesons into two photons. These photons traverse a significant amount of material in the detector before entering the active volume of the CTD; the beam pipe, the MVD and the inner wall of the CTD. Therefore there is a reasonable probability that these photons will convert into an electron-positron pair. At low momenta, these electrons constitute a major background in the search for electrons from semileptonic decays. The reliable detection and suppression of these electrons is therefore important.

An electron-positron pair system arising from a converted photon has a distinctive signature; the electron and positron have a zero distance of closest approach, opening angle and invariant mass (neglecting the mass of the electron itself). A topological conversion finder, CONVERT2 [135], which exploits this distinctive signature, examining combinations of oppositely charged track pairs was used to identify those electrons in our sample which were from photon conversions. The topology of the system is illustrated in Figure B.1 showing a photon converting into an e^+e^- pair in the X–Y and R–Z views.

The distance between the tracks, when the tangents of the two curves are parallel, ΔXY, and the difference in polar angle, $\Delta\theta$, each provide quality tests of the conversion candidate. As a preselection $\Delta XY < 0.4$ cm and $\Delta\theta < 0.1$ rad were applied. Combining these two variables with their resolutions a dimensionless quality factor, D, is defined:

$$D = \sqrt{\left(\frac{\Delta XY}{\sigma_{XY}}\right)^2 + \left(\frac{\Delta\theta}{\sigma_\theta}\right)^2}, \tag{B.1}$$

where σ_{XY} and σ_θ are the resolutions of ΔXY and $\Delta\theta$ estimated separately for the data and the Monte Carlo. In the previous HERA I analyses, e.g. in [115], the following values

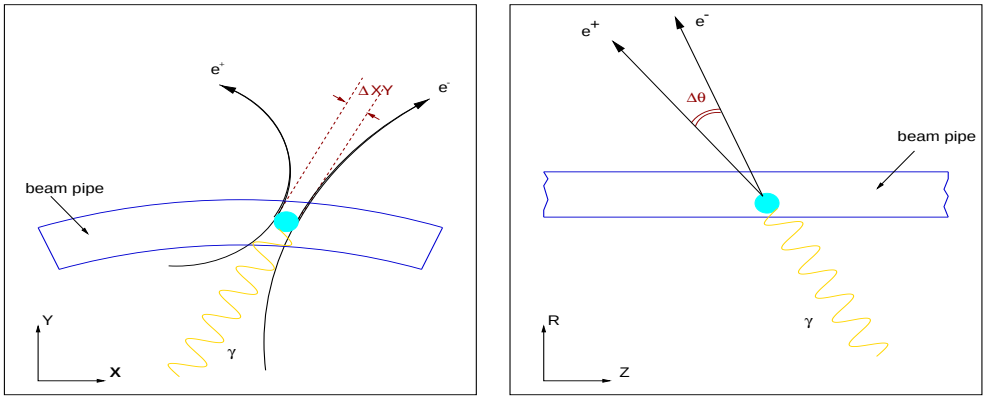


Figure B.1: A schematic view of a photon converting into an electron positron pair in detector material. The left figure shows the X–Y view and the right the R–Z view. The angle $\Delta\theta$ denotes the opening angle between the two tracks and ΔXY is the minimal distance between the two tracks in X–Y plane.

based on VCTRHL tracking had been used.

$$\sigma_{XY}^{\text{Data}} = (0.08 \pm 0.01)\,\text{cm}, \tag{B.2}$$
$$\sigma_{XY}^{\text{MC}} = (0.11 \pm 0.02)\,\text{cm}, \tag{B.3}$$
$$\sigma_{\theta}^{\text{Data}} = (0.017 \pm 0.0001)\,\text{rad}, \tag{B.4}$$
$$\sigma_{\theta}^{\text{MC}} = (0.017 \pm 0.0002)\,\text{rad}. \tag{B.5}$$

By cutting on D, one can either select a pure sample of photon conversions with a low efficiency, or apply a loose cut ($D < 15$) in order to identify as many conversions as possible, albeit with a higher number of misidentified conversions.

In the context of this analysis the CONVERT2 routine was extended to take the improved tracking of the HERA II data into account. A small sample from 2005 data corresponding to 14 pb^{-1} luminosity and an inclusive light flavour MC sample ($\mathcal{L} \simeq 25\,\text{pb}^{-1}$) were used to tune the conversion finder. The events were selected by searching for $\gamma \to e^+e^-$ using cuts: $D < 5$, $M_\gamma < 0.025\,\text{GeV}$ and $\Delta XY < 0.15\,\text{cm}$. These events included real conversions (opposite charge) and equally charged ones (for background estimation). The quality factor D was calculated using the old resolution parameters (cf. Equations B.2-B.5).

The additional MVD information in HERA II data, yielded into more track classes than in the HERA I data. New resolution numbers for ΔXY and $\Delta\theta$, based on ZTTRHL tracking, were extracted in an iterative procedure for five different track classes. The five different classes for tracks were selected, depending on the number of hit superlayers in the CTD (N_{CTD}), the number of reconstructed MVD hits (N_{MVD}), and the radial position of the conversion vertex inside the first MVD layer (in) or outside the first MVD layer (out). The five selected classes were:

1. CTD ($N_{\text{CTD}} \geq 3$ & $N_{\text{MVD}} < 4$);

2. MVD (in) ($N_{\mathrm{CTD}} < 4$ & $N_{\mathrm{MVD}} \geq 4$);
3. MVD (out) ($N_{\mathrm{CTD}} < 4$ & $N_{\mathrm{MVD}} \geq 4$);
4. CTD+MVD (in) ($N_{\mathrm{CTD}} \geq 4$ & $N_{\mathrm{MVD}} \geq 4$);
5. CTD+MVD (out) ($N_{\mathrm{CTD}} \geq 4$ & $N_{\mathrm{MVD}} \geq 4$).

For each of the samples, the width of the ΔXY and $\Delta\theta$ distributions were determined. The procedure was iterated until the resolution numbers got stable. Table B.1 shows the new resolution numbers for the five subsamples, calculated separately for the data and the Monte Carlo. Most of the classes show similar numbers for the data and the MC (the significant differences in the $\Delta\theta$ resolution in the inner MVD region are still not understood). The best resolution in $\Delta\theta$ could be achieved for the MVD tracks where the precision is significantly better than for the CTD tracks and also better than the values determined for HERA I, based on VCTRHL tracking. For the resolution in ΔXY, the differences were less pronounced but have similar trends.

	ΔXY (cm)		$\Delta\theta$ (rad)	
	Data	MC	Data	MC
CTD	0.070	0.066	0.021	0.019
MVD (in)	0.066	0.061	0.013	0.019
MVD (out)	0.060	0.060	0.012	0.014
CTD+MVD (in)	0.051	0.045	0.018	0.017
CTD+MVD (out)	0.060	0.050	0.019	0.017

Table B.1: Resolutions of ΔXY and $\Delta\theta$ for the five different track classes defined in the text. The resolutions have been determined separately for data and MC.

Figure B.2 shows the ΔXY and $\Delta\theta$ distributions for $\gamma \to e^+e^-$, using old and new resolution parameters. An improvement of the conversion finder using the new tracking and the re-determined resolution values can be clearly seen.

To estimate the impact of the improvements in the conversion finder on the analysis, it was observed how many true electrons from photon conversions are left in the sample after cutting on the candidates found by the conversion finder. Figure B.3 shows the distribution of the electron likelihood hypothesis (see Chapter 8) for all true electrons (yellow), for a subsample of electrons originating from photon conversions (blue) and the remaining conversion electrons, after cutting on the conversion finder, using the old resolution parameters based on VCTRHL tracking (green) and using the new resolution parameters based on ZTTRHL tracking (red). The improved conversion finder helped to suppress the conversion background by more than a factor of two. The updated conversion finder can detect $\sim$70% of the conversions in the data sample.

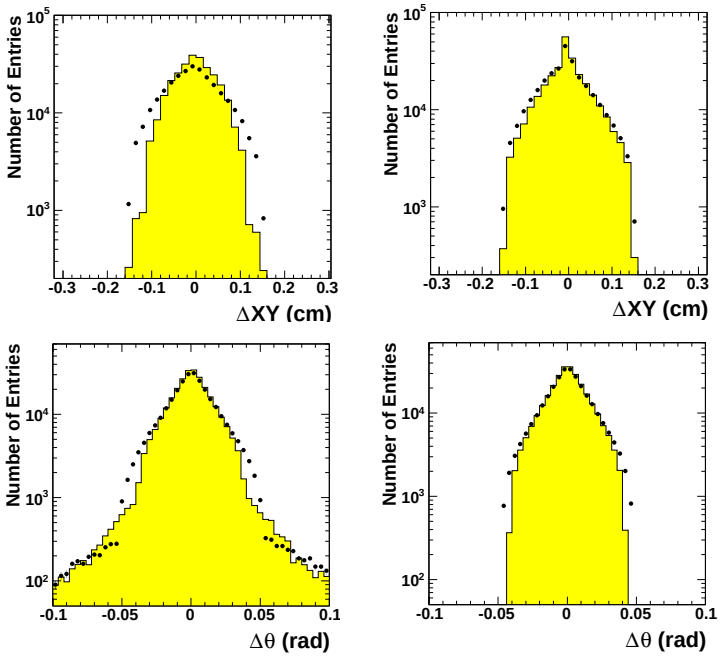


Figure B.2: The minimal distance, ΔXY, and the opening angle, $\Delta\theta$, for data (black points) and Monte Carlo (yellow histogram): using the old resolution parameters (left column), and using the new resolution parameters (right column).

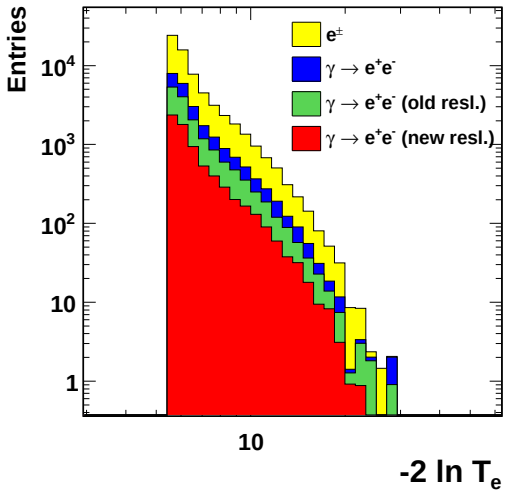


Figure B.3: The electron likelihood hypothesis, $-2\ln T_e$, (see Chapter 8) for all true electrons (yellow). The blue distribution shows the subsample of electrons originating from photon conversions. After the cut on the conversion finder, using the old resolution parameters based on VCTRHL tracking, the green contribution remains, which can be further suppressed by using the new resolution parameters based on ZTTRHL tracking. The remaining electrons from photon conversions are shown in red.

Appendix C

Acceptances and QED Corrections

In this appendix, the acceptances, purities and QED radiative corrections in the bins of the variables for which differential cross sections are measured are presented.

The acceptances for the bins of single-differential cross sections were already shown in Chapter 9. Figure C.1 shows the acceptances for double-differential cross sections as a function of x and Q^2. The acceptances are usually in the range 10–15%. The purities shown in Figures C.2 and C.3 are of the order of 70–90% for Q^2, p_{T}^e and η^e bins, while in some x bins the purity drops to 50% because of migrations between these bins.

The QED radiative corrections needed to correct the measured cross sections to the Born level are shown in Figures C.4 and C.5. The corrections are typically $C_r \approx 1.05$ and go up to $C_r \approx 1.1$ for the high Q^2 region.

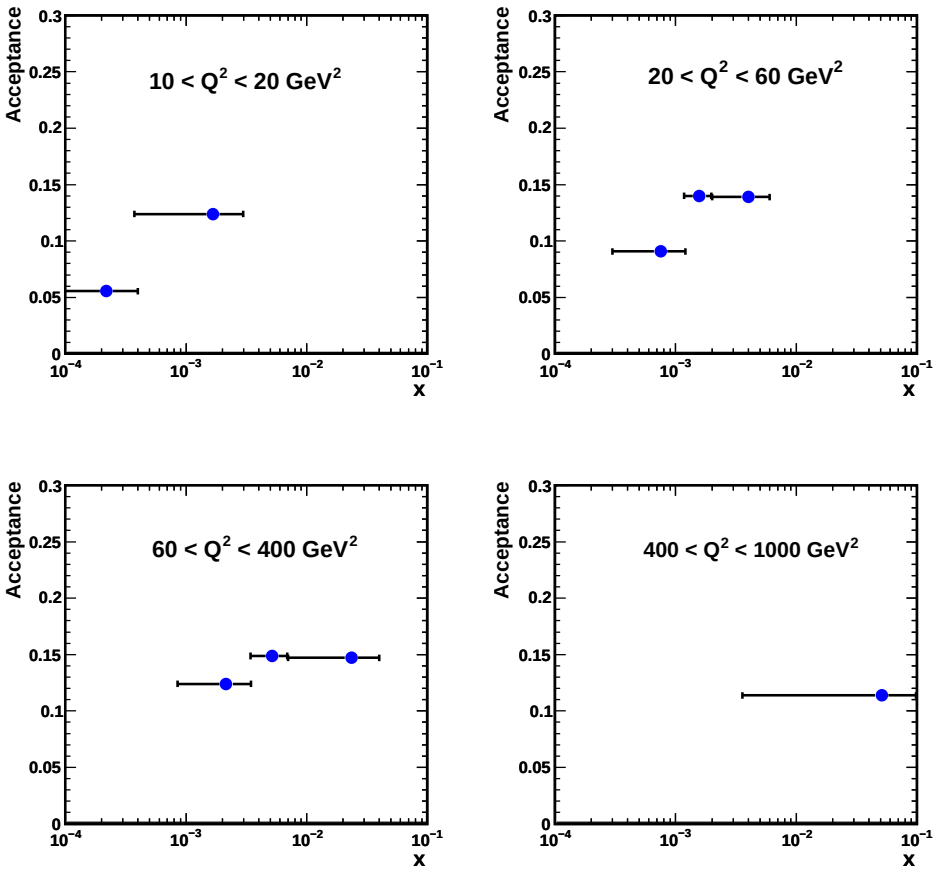


Figure C.1: Acceptances in bins of x for four Q^2 regions in beauty events.

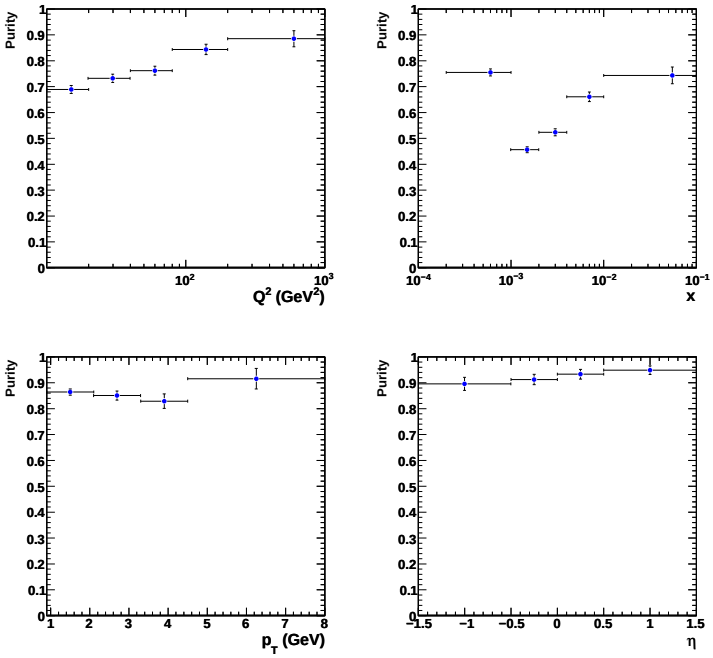


Figure C.2: Purities in bins of the kinematic variables Q^2 and x (top) and in bins of the decay electron variables p_{T}^e and η^e (bottom) in beauty events.

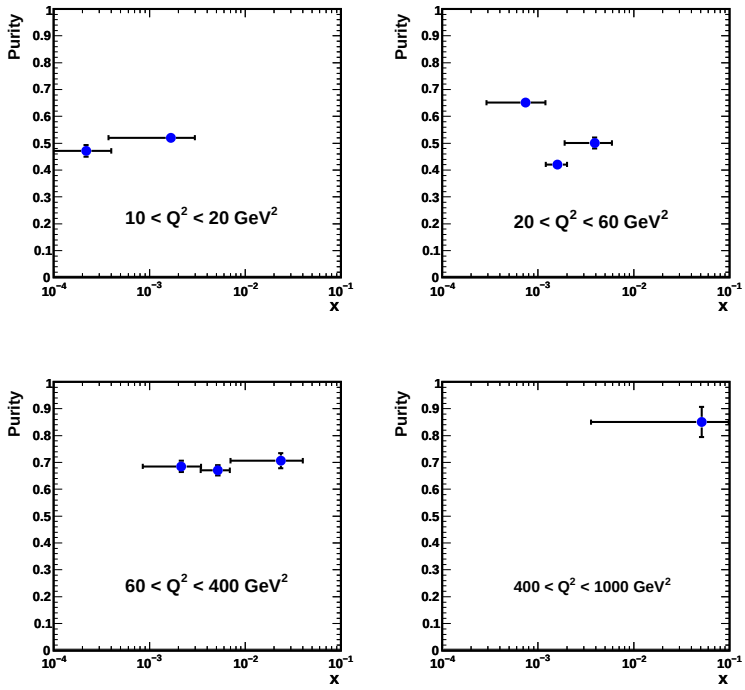


Figure C.3: Purities in bins of x for four Q^2 regions in beauty events.

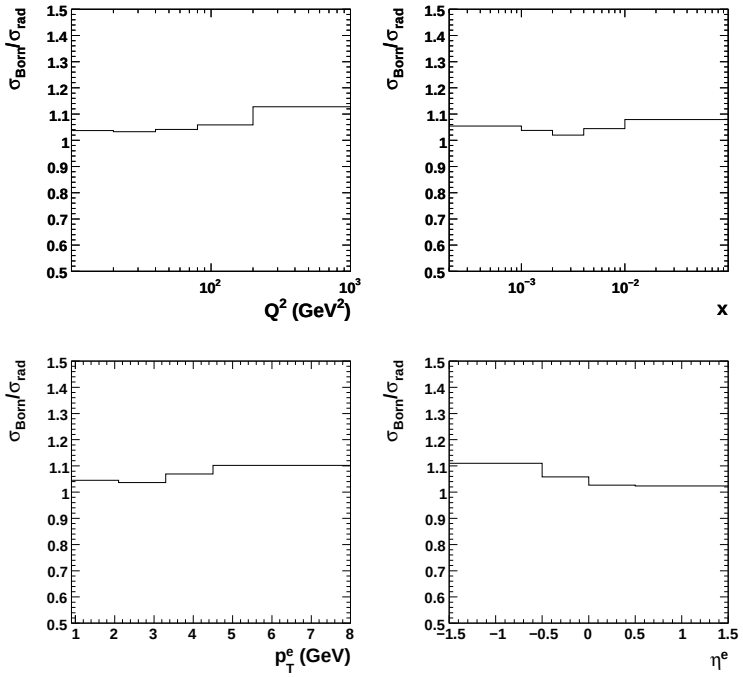


Figure C.4: QED corrections in bins of the kinematic variables Q^2 and x (top) and in bins of the decay electron variables p_{T}^e and η^e (bottom) in beauty events.

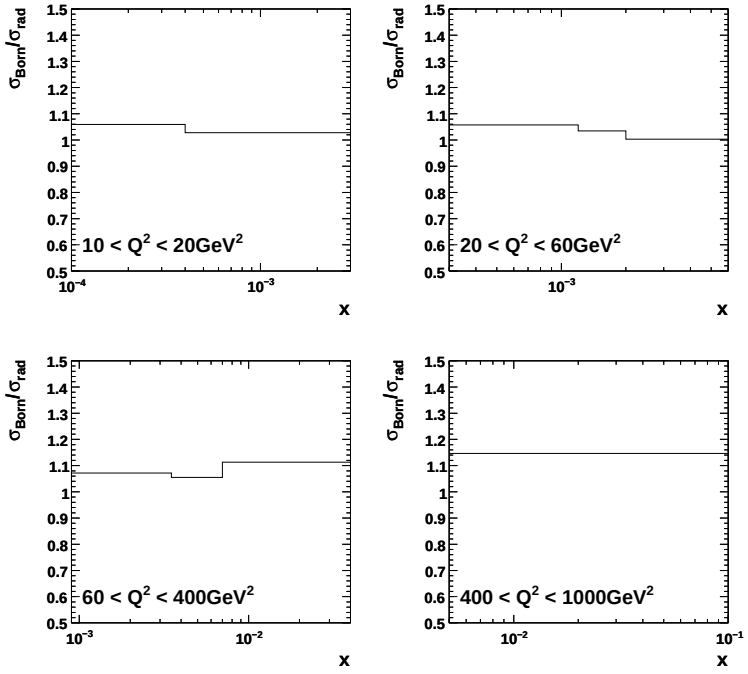


Figure C.5: QED corrections in bins of x for four Q^2 regions in beauty events.

Appendix D

Control Plots

In order to check the quality of the signal and background descriptions, the distributions of the variables relevant for this analysis were made for the signal-enriched and background-dominated samples, as described in Chapter 9. In addition to all the selection cuts listed in Table 6.3, the background-dominated sample includes a cut on the likelihood hypothesis at $-2\ln T < 10$, while the signal-enriched sample was selected by applying a hard cut at $-2\ln T < 1.5$. Some of the distributions for these two selection regions were already shown in Section 9.5. In the following several additional distributions are shown. In all Figures D.1-D.5, the plots on the left are drawn for $-2\ln T < 10$ while the plots on the right include a cut on $-2\ln T < 1.5$. Note that, with the exception of one plot (ΔR distribution), left plots are shown on logarithmic scale, while the plots on the right side are shown on linear scale, as statistics is limited.

The first set of variables (Figure D.1) are the scattered electron variables $E_{e'}, \theta_{e'}$ and $\eta_{e'}$. The variables show a good description for both selection regions. Figure D.2 shows the kinematic variables, photon virtuality, Q^2, Bjorken-x and inelasticity, y, calculated using the double-angle method. In general a reasonable agreement between data and Monte Carlo is observed. $E - p_Z$, $Z_{\text{prm-vtx}}$ and ΔR distributions are shown in Figure D.3. Some discrepancies are visible for the background-dominated sample, especially for the $E - p_Z$ and ΔR distributions. The description is significantly improved for the hard selection (signal-enriched). The energy deposited in the calorimeter, E^{CAL}, and two of the variables used for electron identification in the likelihood hypothesis, $E^{\text{CAL}}/p^{\text{trk}}$ and d_{cell}, are shown in Figure D.4. The discrepancy in $E^{\text{CAL}}/p^{\text{trk}}$ is mostly visible in the background enriched region while the other two variables are described quite well for both selections. In the last Figure D.5, the distributions of number of MVD hits, $n_{\text{mvd hits}}$, impact parameter significance, $IP/\delta IP$ and invariant mass of the secondary vertex, $M_{\text{sec-vtx}}$ are shown. $n_{\text{mvd hits}}$ show a fair agreement between data and Monte Carlo. The last two variables also show a reasonable description, but were not used at the end in this analysis (cf. Section 7.3).

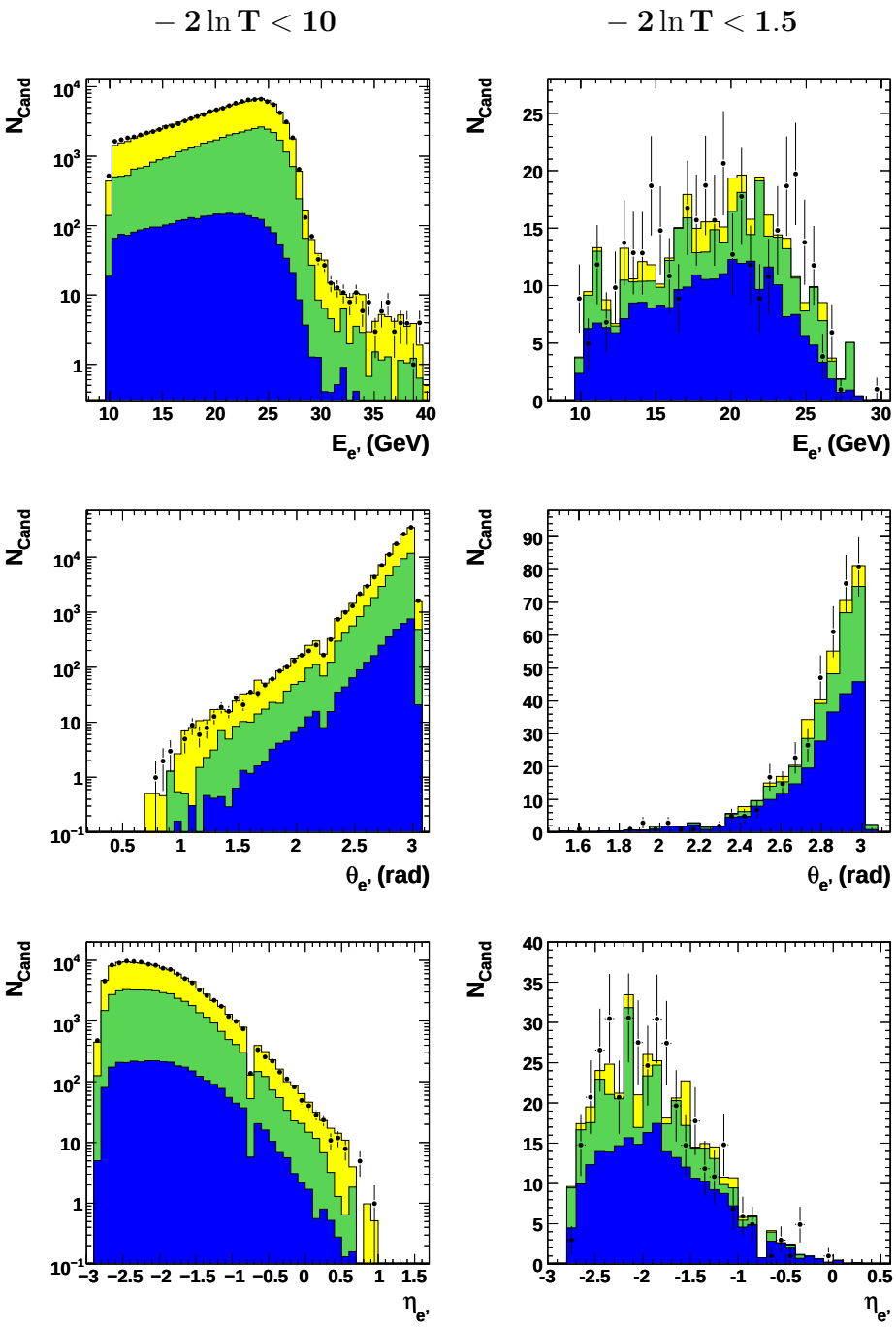


Figure D.1: Scattered electron variables: for all candidates that enter the fit satisfying $-2\ln T < 10$ (left), and for the beauty-enriched region, $-2\ln T < 1.5$ (right). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) after applying the scale factors from the fit. The summed distributions are compared to the data distributions shown by the black points. Variables shown from top to bottom are $E_{e'}$, $\theta_{e'}$ and $\eta_{e'}$.

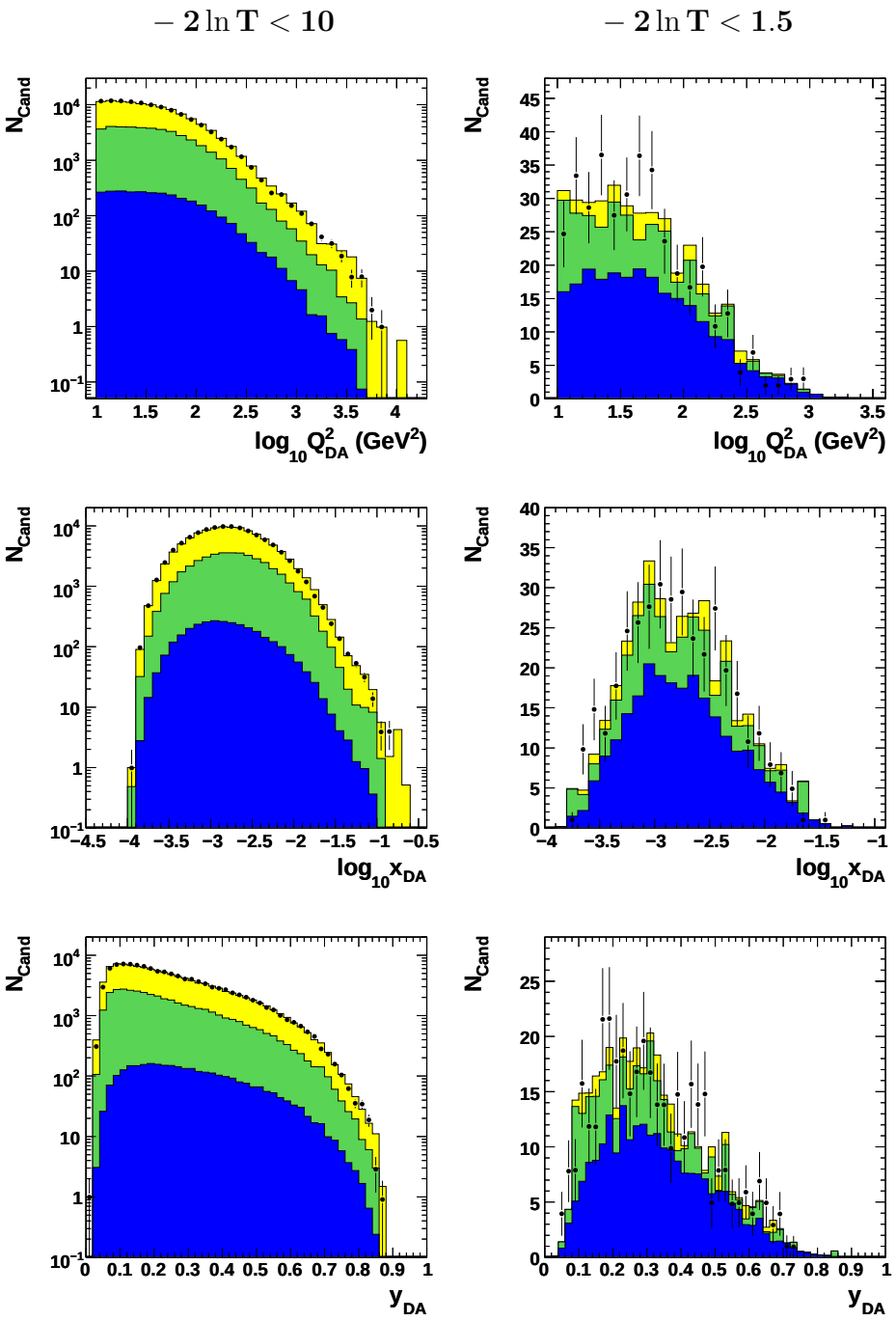


Figure D.2: Kinematic variables: for all candidates that enter the fit satisfying $-2\ln T < 10$ (left), and for the beauty-enriched region, $-2\ln T < 1.5$ (right). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) after applying the scale factors from the fit. The summed distributions are compared to the data distributions shown by the black points. Variables shown from top to bottom are Q^2, x and y.

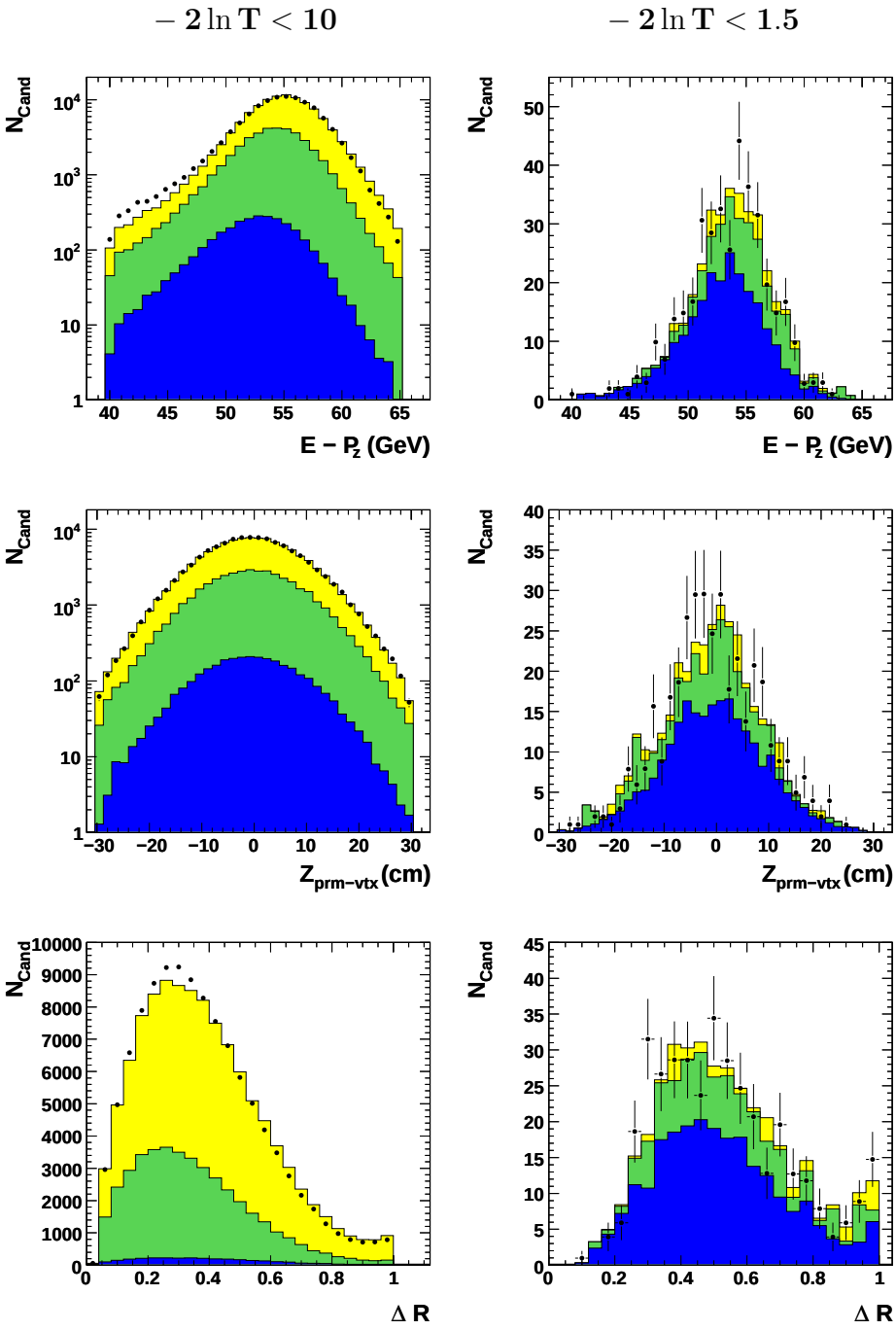


Figure D.3: $E - p_Z$, $Z_{\text{prm-vtx}}$ and ΔR: for all candidates that enter the fit satisfying $-2\ln T < 10$ (left), and for the beauty-enriched region, $-2\ln T < 1.5$ (right). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) after applying the scale factors from the fit. The summed distributions are compared to the data distributions shown by the black points.

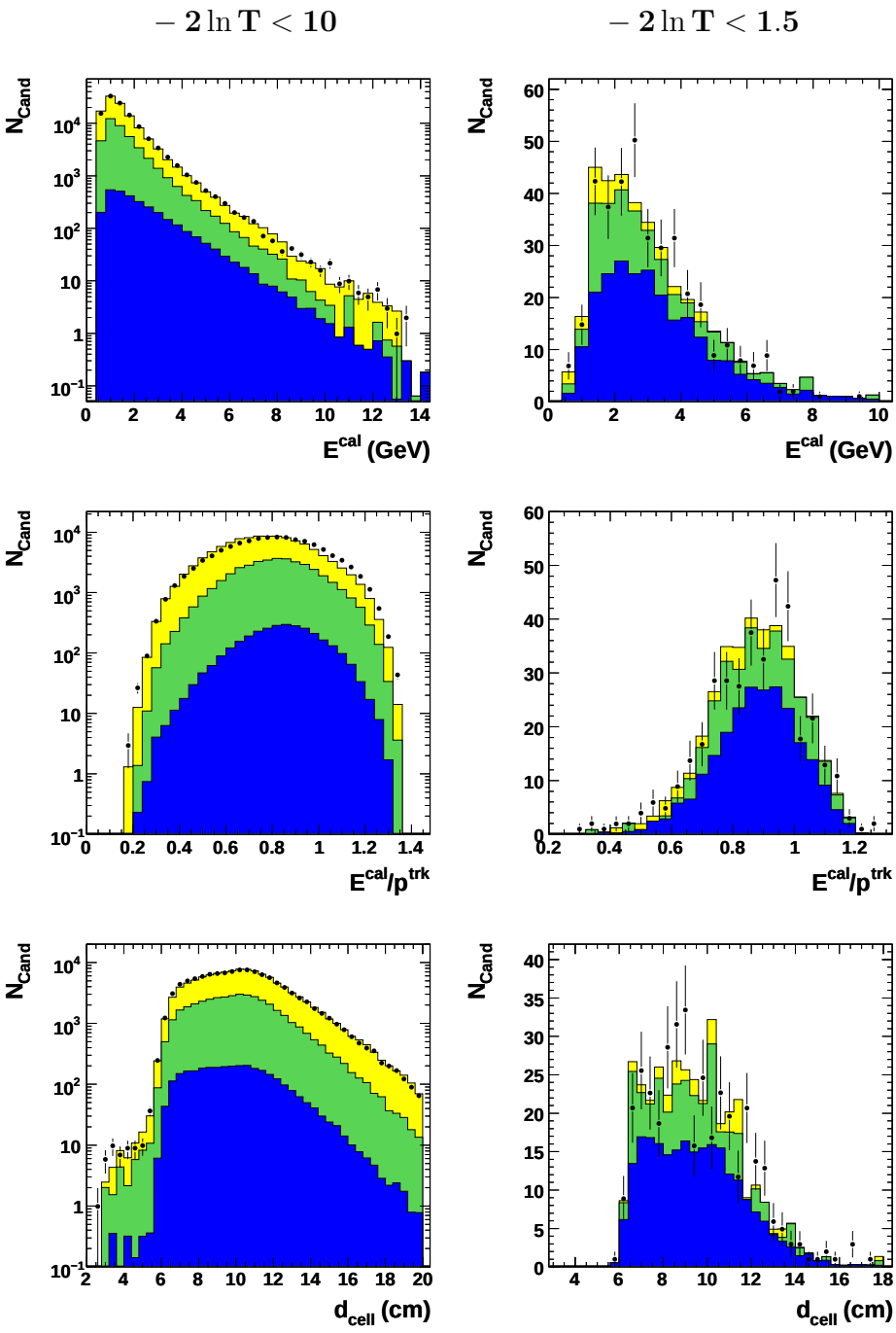


Figure D.4: E^{CAL} and fit variables for electron identification: for all candidates that enter the fit satisfying $-2\ln T < 10$ (left), and for the beauty-enriched region, $-2\ln T < 1.5$ (right). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) after applying the scale factors from the fit. The summed distributions are compared to the data distributions shown by the black points. Variables shown from top to bottom are E^{CAL}, $E^{\mathrm{CAL}}/p^{\mathrm{trk}}$ and d_{cell}.

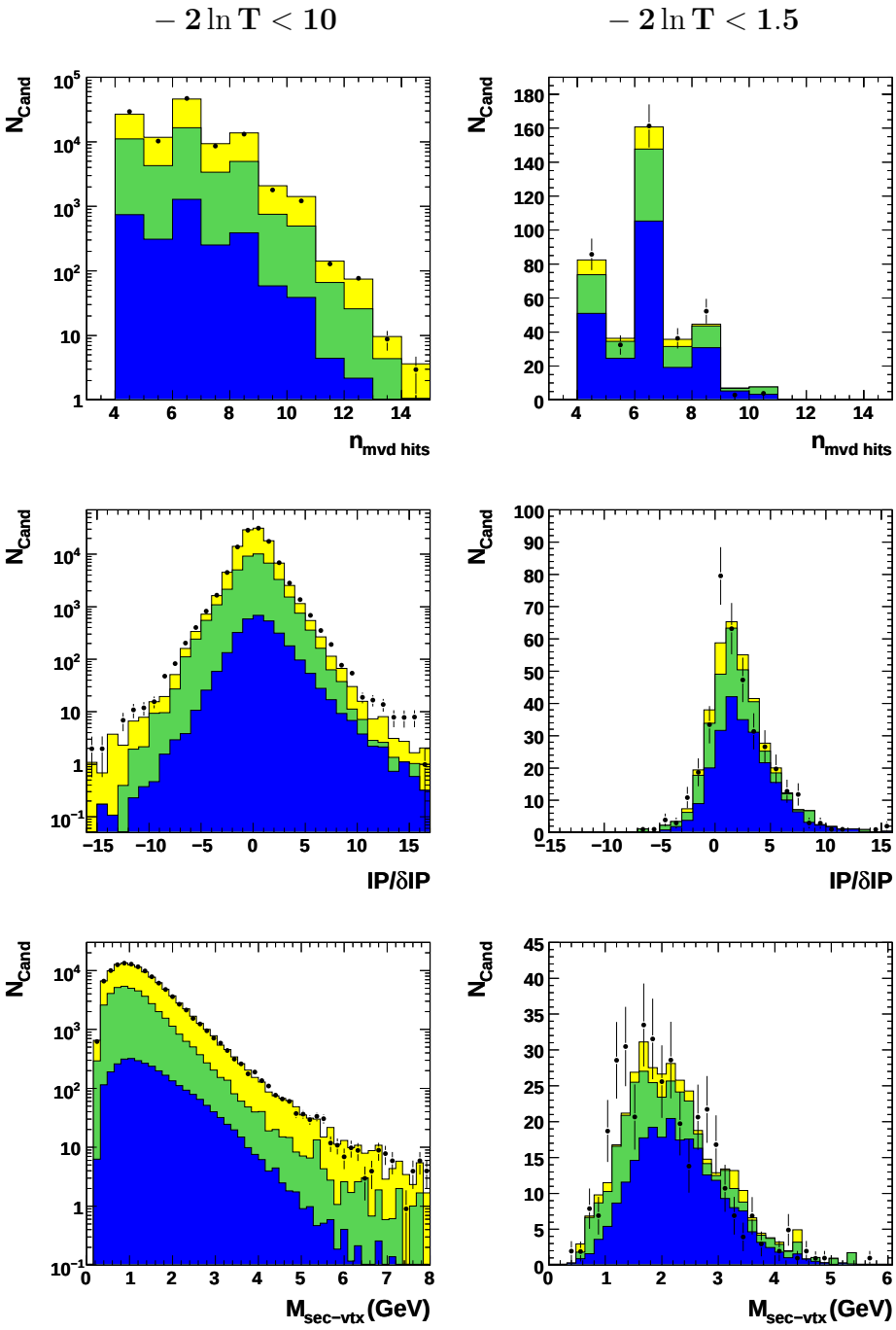


Figure D.5: $n_{\text{mvd hits}}$, $IP/\delta IP$ and $M_{\text{sec-vtx}}$: for all candidates that enter the fit satisfying $-2\ln T < 10$ (left), and for the beauty-enriched region, $-2\ln T < 1.5$ (right). The shaded areas show the contributions from b quarks (blue), electron background (green) and non-electrons (yellow) after applying the scale factors from the fit. The summed distributions are compared to the data distributions shown by the black points.

Appendix E

Fit Distributions

For the differential cross sections, the likelihood distribution was split into the bins of the variable under study. The same fit procedure as described in Chapter 9 was applied to determine the beauty fractions in each bin. In this appendix the results of the likelihood test function fits corresponding to the single-differential cross sections as a function of Q^2, x, p_{T}^e and η^e and double-differential cross sections as a function of x in four regions of Q^2 presented in Chapter 11 are shown. The data distributions are shown together with the scaled Monte Carlo contributions in Figures E.1-E.5.

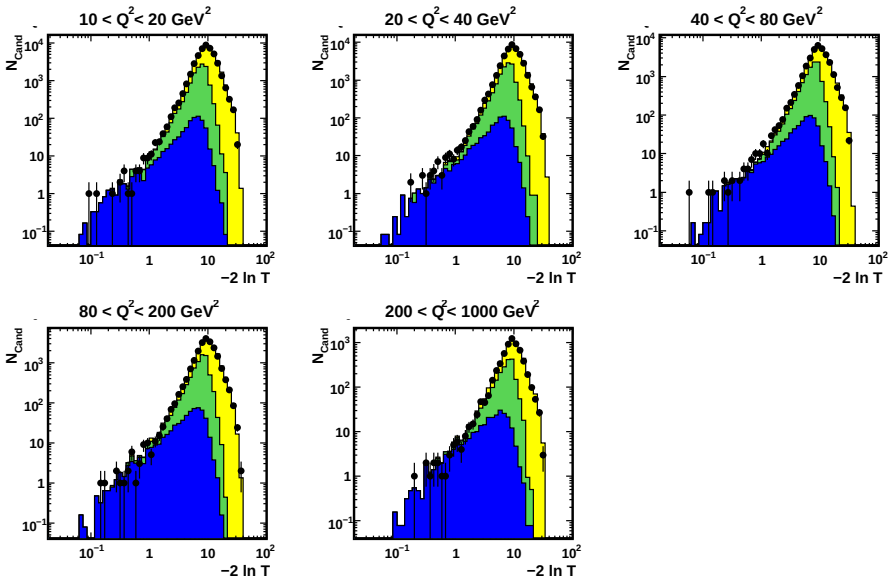


Figure E.1: Distributions of the likelihood test function, $-2\ln T$, using the beauty hypothesis, to extract the beauty fractions. The five plots show the distribution for the five bins in Q^2. The shaded areas show the fitted contributions from b quarks (blue), the background from electrons (green) and non-electrons (yellow). The sum of the Monte Carlo distributions are overlayed with the data (black points).

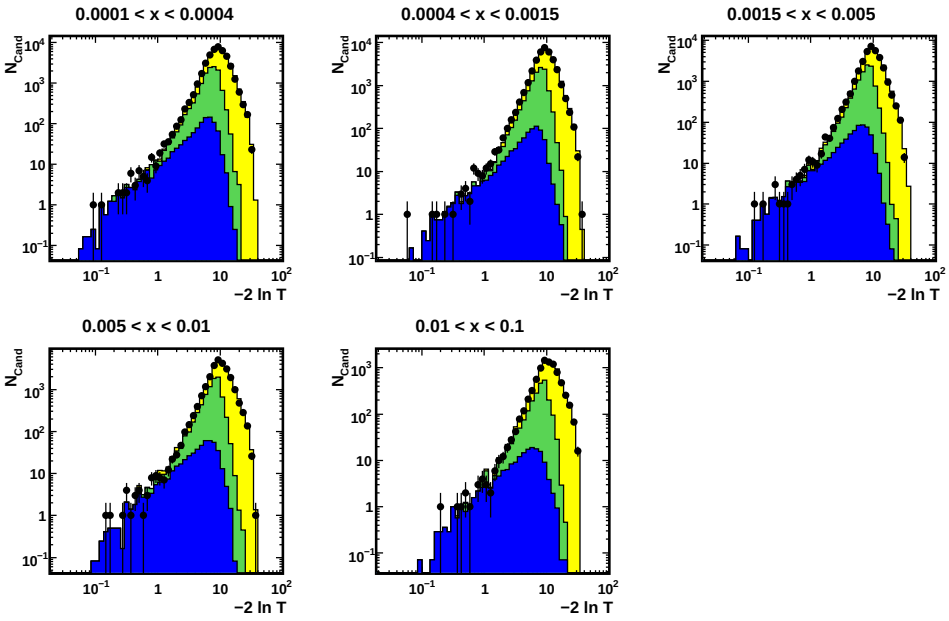
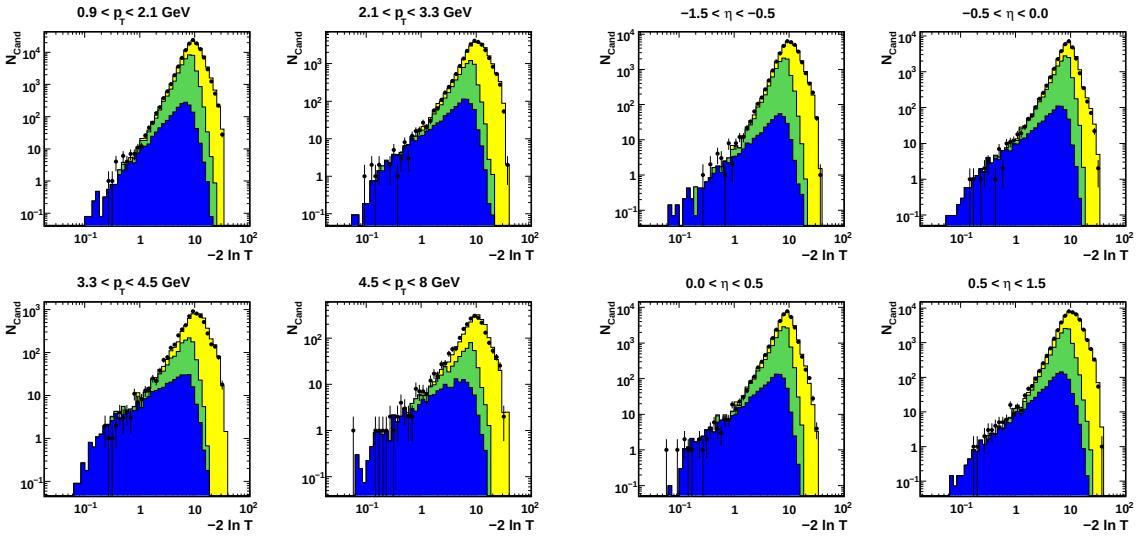


Figure E.2: Distributions of the likelihood test function, $-2\ln T$, using the beauty hypothesis, to extract the beauty fractions. The five plots show the distribution for the five bins in x. For further details see the caption of Figure E.1.



(a) Fit in bins of p_{T}^{e} **(b)** Fit in bins of η^{e}

Figure E.3: Distributions of the likelihood test function, $-2\ln T$, using the beauty hypothesis, to extract the beauty fractions. The four plots in (a) show the distribution for the four bins in p_{T}^{e} and the ones in (b) show the distribution for the four bins in η^{e}. For further details see the caption of Figure E.1.

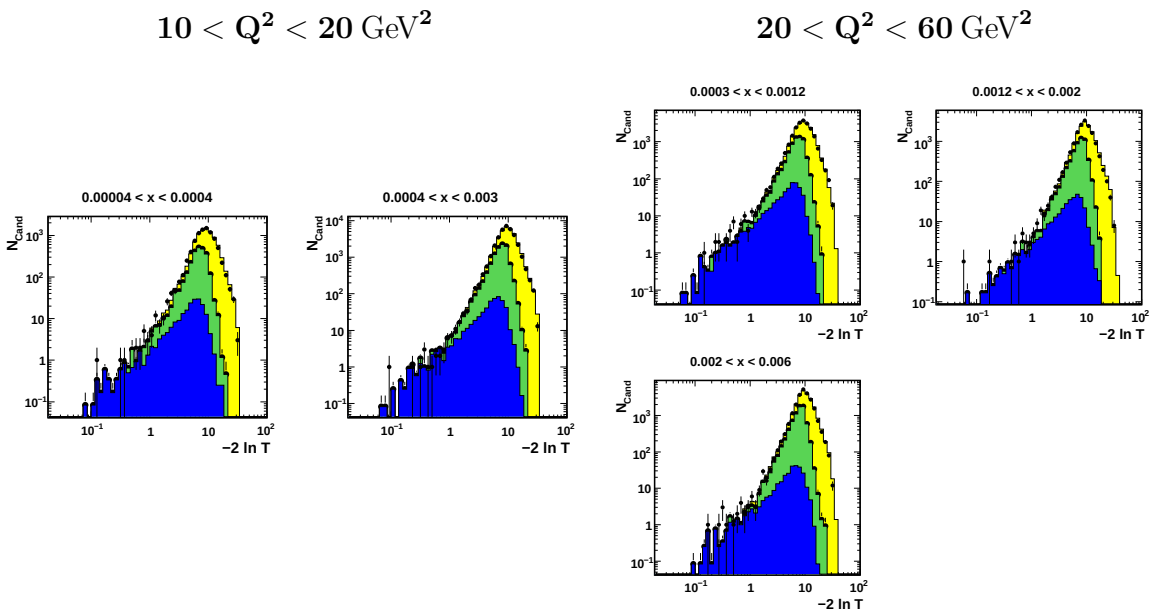


Figure E.4: Distributions of the likelihood test function to extract the beauty fractions. The two plots on left show the distribution for the two bins in x for $10 < Q^2 < 20\,\mathrm{GeV}^2$ and the ones on right show the distribution for the three bins in x for $20 < Q^2 < 60\,\mathrm{GeV}^2$. For further details see the caption of Figure E.1.

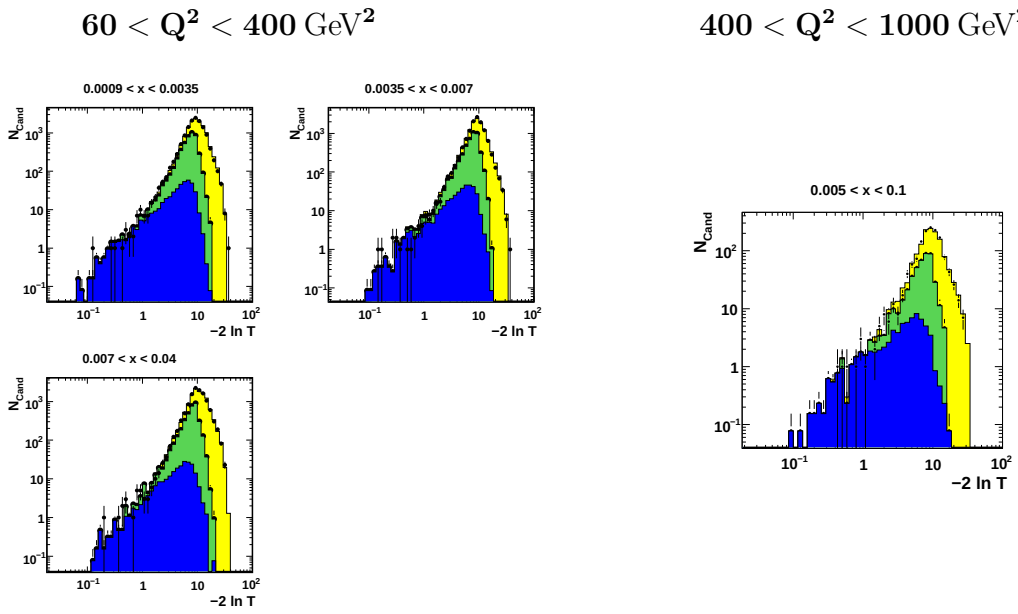


Figure E.5: Distributions of the likelihood test function to extract the beauty fractions. The two plots on left show the distribution for the three bins in x for $60 < Q^2 < 400\,\mathrm{GeV}^2$ and the ones on right show the distribution for the one bin in x for $400 < Q^2 < 1000\,\mathrm{GeV}^2$. For further details see the caption of Figure E.1.

Appendix F

Systematics

In this appendix, systematic uncertainties for each cross section bin (for the details of the binning, see Chapter 11), for all the different sources discussed in Chapter 10 are shown.

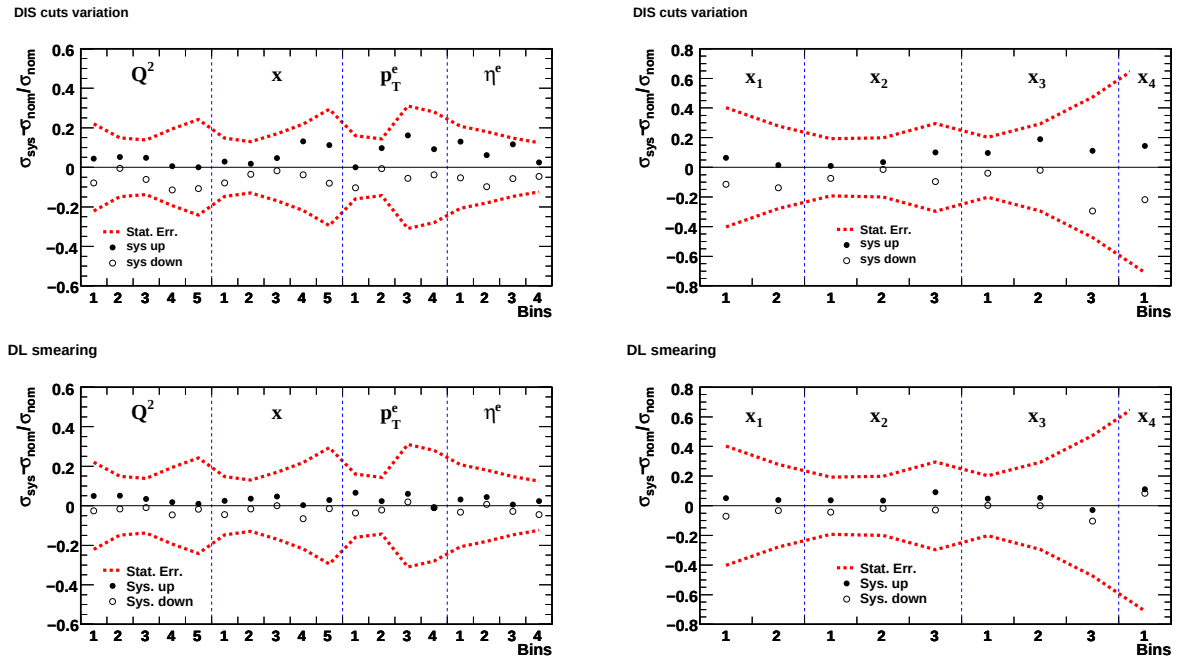


Figure F.1: Quadratic sum of the systematic uncertainties for the variation of DIS selection cuts (top) and the variation of DL smearing by ±2% (bottom) for single- and double-differential cross sections. The left plots show the systematic uncertainties for the cross-section bins of the variables, Q^2, x, p_{T}^e and η^e, and the right plots for x bins for four different regions of Q^2, denoted as x_1, x_2, x_3 and x_4. The points show the relative deviation of the cross section from the central measured value for a given bin and the dashed red line shows the statistical error on the measurement.

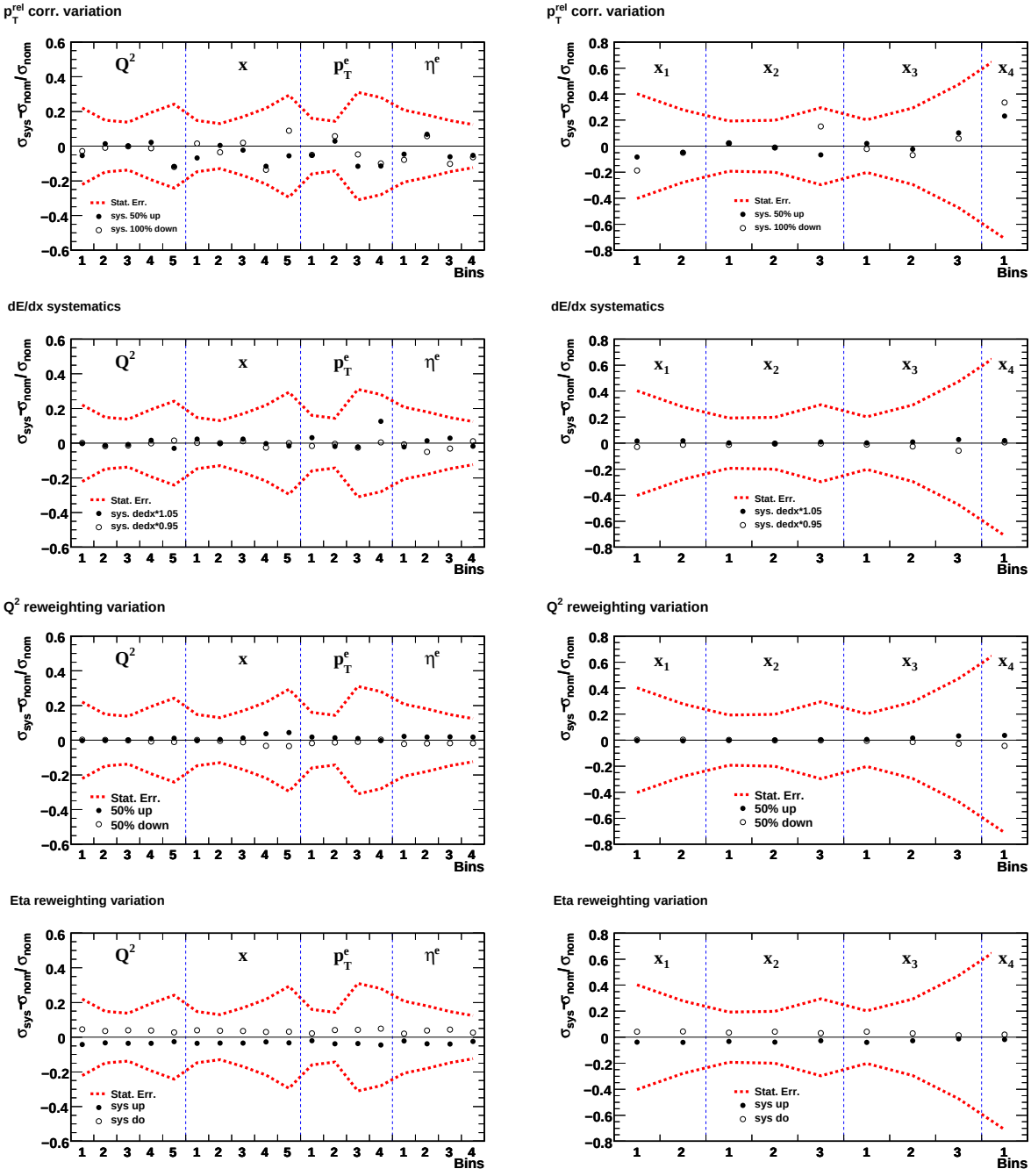


Figure F.2: Systematic uncertainties for the variation of, (from top to bottom), $p_{\mathrm{T}}^{\mathrm{rel}}$ shape correction by +50% and −100%, dE/dx simulation by varying the mean and width of Bethe Bloch distribution by one standard deviation, Q^2 reweighting by ±50% and η reweighting between 0% and 10%. For further details see the caption of Figure F.1.

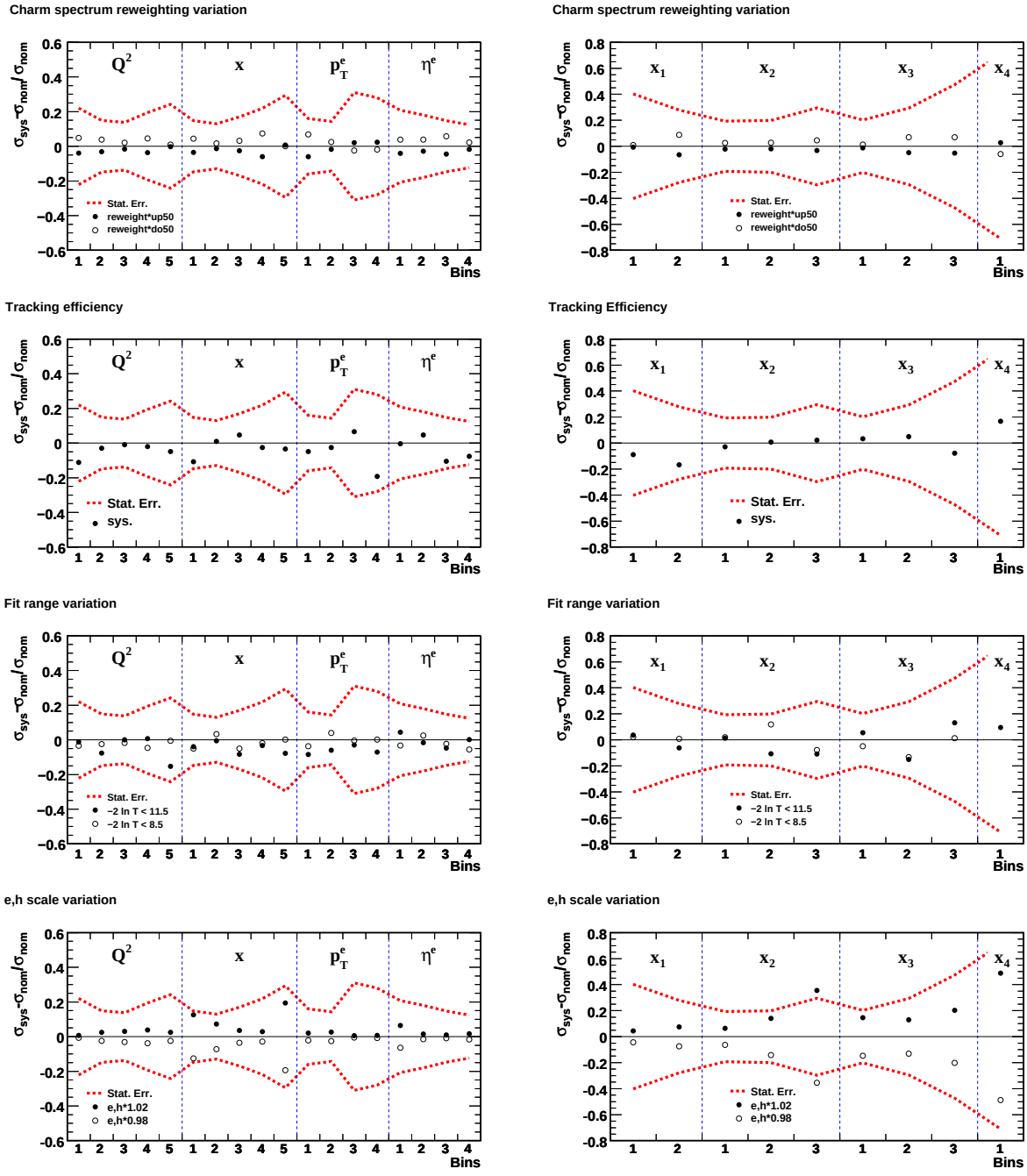


Figure F.3: Systematic uncertainties for the variation of, (from top to bottom), charm spectrum reweighting by ±50%, tracking efficiency by rejecting 2% of tracks in MC, for the signal extraction procedure by varying the fit range by ±1.5% and global energy scale by ±2%. For further details see the caption of Figure F.1.

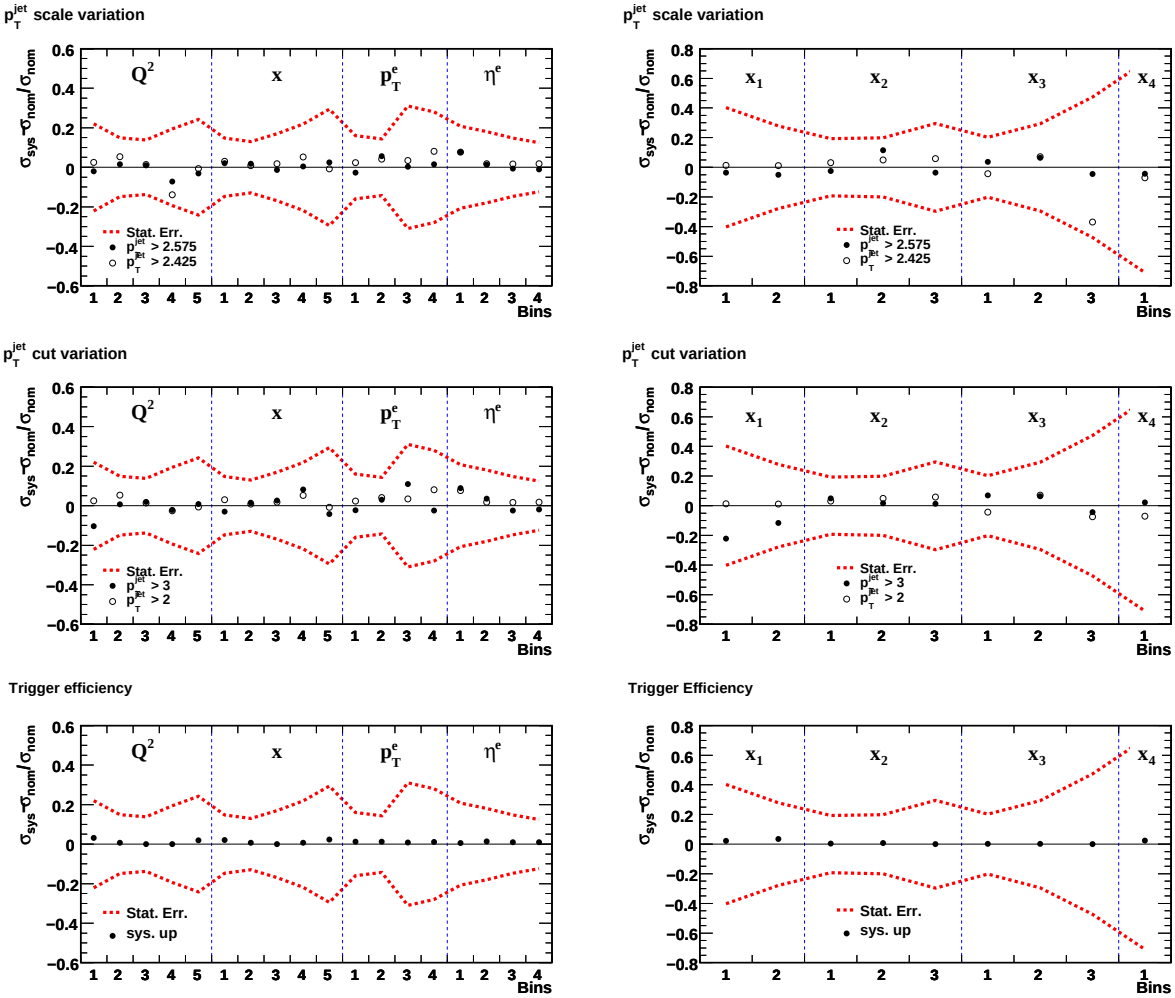


Figure F.4: Systematic uncertainties for the variation of, (from top to bottom), jet energy scale by ±3%, cut on $p_{\mathrm{T}}^{\mathrm{jet}}$ by ±0.5% and trigger efficiency by weighting down the events not fired by FLT30 by 5%. For further details see the caption of Figure F.1.

List of Figures

List of Tables

Bibliography

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

[2] G. Zweig, *An SU_3 Model for Strong Interaction Symmetry and its Breaking; Part II*, CERN Report 8419/TH.412, (1964).

[3] D.J. Griffiths, *Introduction to Elementary Particles 2nd Edition*, WILEY-VCH Verlag GmbH & Co. KGaA, (2008).

[4] F. Halzen, A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, (1984).

[5] M. Breidenbach et. al., *Observed behaviour of highly inelastic electron-proton scattering*, Phys. Rev. Lett. **23**, 935 (1969).

[6] E.D. Bloom et al., *High-Energy Inelastic ep Scattering at* 6° *and* 10°, Phys. Rev. Lett. **23**, 930 (1969).

[7] ZEUS Collab., *Measurement of beauty photoproduction from inclusive secondary vertexing at HERA II*, write-up submitted to DIS09 conference, Madrid, April 2009, ZEUS-prel-09-005.

[8] ZEUS Collab., *Measurement of beauty production from inclusive secondary vertices in DIS and $F_2^{b\bar{b}}$ extraction at ZEUS*, write-up submitted to International Conference on High Energy Physics, Paris, July 2010, ZEUS-prel-10-004.

[9] ZEUS Collab., H. Abramowicz et. al., *Measurement of beauty production in DIS and $F_2^{b\bar{b}}$ extraction at ZEUS*, 1005.3396, (2010).

[10] ZEUS Collab., S. Chekanov et. al., *Measurement of charm and beauty production in deep inelastic ep scattering from decays into muons at HERA*, arXiv:0904.3487, (2009).

[11] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge Monographs on Particle Physcis, Nuclear Physics and Cosmology, Vol. 8. Cambridge University Press, 1996.

[12] A. Gabareen Mokhtar, *Study of Neutral and Charged Current Cross Sections at High Q^2 at HERA.* Ph.D. Thesis, Tel Aviv University, Report DESY-THESIS-2006-005, 2006.

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

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

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

[16] G. Miller et al., *Inelastic Electron-Proton Scattering at Large Momentum Transfers and the Inelastic Structure Functions of the Proton*, Phys. Rev., **D 5**, 528 (1972).

[17] C.G. Callan, D.J. Gross, *High-Energy Electroproduction and the Constitution of the Electric Current*, Phys. Rev. Lett., **22**, 156 (1969).

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

[19] K. Nakamura et al., *The Review of Particle Physics.*, Journal of Physics G, **37** 075021, 2010.
`http://pdg.lbl.gov`

[20] C. Glasman, *Jet production in deep inelastic scattering at HERA.*, arXiv:hep-ex/0312011v1.

[21] G. Sterman, C. Collins, D. Soper, *Soft gluons and factorisation*, Nucl. Phys., **B 308**, 833 (1988).

[22] Y.L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbative Theory in QCD (in Russian)*, Sov. Phys. JETP **46**, 641 (1977).

[23] V.N. Gribov and L.N. Lipatov, *Deep Inelastic ep Scattering in Perturbation Theory*, Sov. J. Nucl. Phys. **15**, 438 (1972).

[24] L.N. Lipatov, *The Parton Model and Perturbation Theory*, Sov. J. Nucl. Phys. **20**, 96 (1975).

[25] G. Altarelli, G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B 126**, 298 (1977).

[26] ZEUS Collab., S. Chekanov et al., *An NLO QCD Analysis of Inclusive Cross-Section and Jet-Production Data from the ZEUS Experiment*, Eur. Phys. J. **C 42**, 1 (2005).

[27] H1 and ZEUS Collab., *Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}p$ Scattering Cross Sections at HERA*, JHEP **01** 109 (2010).

[28] E.A. Kuraev, L.N. Lipatov, V.S. Fadin, *Multi-Reggeon Processes in the Yang-Mills Theory*, Sov. Phys., JETP, **44**, 443 (1976).

[29] E.A. Kuraev, L.N. Lipatov, V.S. Fadin, *The Pomeranchuk Singularity in Nonabelian Gauge Theories*, Sov. Phys., JETP, **45**, 199 (1977).

[30] I.I. Balitzki, L.N. Lipatov, *The Pomeranchuk Singularity in Quantum Chromodynamics*, Sov. J. Nucl. Phys., **28**, 822 (1978).

[31] M. Ciafaloni, *Coherence Effects in Initial Jets at Small Q^2/s*, Nucl. Phys., **B 296**, 49 (1988).

[32] S. Catani, F. Fiorani, G.Marchesini, *QCD Coherence in Initial State Radiation*, Phys. Lett., **B 234**, 339 (1990).

[33] S. Shimizu., *Measurement of the Longitudinal Structure Function FL at HERA.*, Ph.D. Thesis., `http://www-zeus.desy.de/physics/sfe/theses/` April 2009.

[34] L.M. Jones, H.W. Wyld, *A Measurement of Beauty Production in High-Energy Positron-Proton Scattering*, Phys. Rev., **D 17**, 759 (1978).

[35] Frixione, Stefano, *Theory versus Experiments in Heavy Flavor Production*, hep-ph/0111368, 2001.

[36] C. Peterson et al., *Scaling violations in inclusive e^+e^- annihilation spectra* Phys. Rev., **D 27**, 105 (1983).

[37] P. Nason, C. Oleari, *A Phenomenological Study of Heavy Quark Fragmentation Functions in e^+e^- Annihilation*, Nucl. Phys. **B 565**, 245 (2000).

[38] H. Jung, Comp. Phys. Commun. **86** 147 (1995).

[39] H. Jung, *The RAPGAP Monte Carlo for Deep Inelastic Scattering*, version 3.10, Lund University, (2005).

[40] G.A. Schuler, H. Spiesberger, *DJANGOH: The interface for the event generators HERACLES and LEPTO*, In Hamburg 1991, Proceedings, Physics at HERA, Vol. 3, 1419-1432.

[41] A. Kwiatkowski, H. Spiesberger, H.-J. Möhring, *HERACLES 4.1 - An event generator for ep interactions at HERA including radiative processes*, Proceedings, Workshop on Physics at HERA, Vol. 3, edited by W. Buchmüller, G. Ingelman (1991), p.1924.

[42] A. Kwiatkowski, H. Spiesberger, H.-J. Möhring, *HERACLES - An event generator for ep interactions at HERA including radiative processes. version 4.6, (1996)*, `http://www.desy.de/~hspiesb/heracles.html`

[43] L. Lönnblad, *ARIADNE version 4: A Program for Simulation of QCD Cascades Implementating the Colour Dipole Model*, Comp. Phys. Commun. **71**, 15 (1992).

[44] CTEQ Collab., H.L. Lai et al., Eur. Phys. J. **C 12**, 375-392 (2000).

[45] G. Gustafson, Phys. Lett., **B 175** 453-456 (1986).

[46] G. Gustafson, U. Pettersson, Nucl. Phys., **B 306** 746-758 (1988).

[47] B. Andersson et al., Z. Phys., **C 43** 625-632 (1989).

[48] T. Sjostrand et al., *High-energy-physics event generation with PYTHIA 6.1*, Comput. Phys. Commun. **135**, 238-259 (2001).

[49] B. Andersson et al., *Parton Fragmentation and String Dynamics* Phys. Rept., **97** 31-145 (1983).

[50] R. Brun. Geant3, Technical Report CERN-DD/EE/84-1, CERN, 1987.

[51] Els de Wolf (editor) et al., *ZGANA, ZEUS Trigger Simulation Library*, available on `http://www-zeus.desy.de/components/offline/offline.html`

[52] S.M. Fisher, A. Palazzi, *The ADAMO Data System Programmers's Manual - Version 3.2.* CERN ECP and RAL, available on `http://adamo.web.cern.ch/Adamo/refmanual/Document.html`

[53] S. Dusini, A. Fox-Murphy, I. Grabowska-Bold, C. Gwenlan, J. Sztuk et al., ZEUS Collab., *ORANGE USER MANUAL*, available on `http://www-zeus.desy.de/ZEUS_ONLY/analysis/orange/index.html`

[54] M. Cacciari, S. Frixione, P. Nason, JHEP **0103** 006 (2001).

[55] B.W. Harris, J. Smith, Nucl. Phys. **B 452,** 109 (1995).

[56] B.W. Harris, J. Smith, Phys. Lett. **B 353,** 535 (1995). Erratum-ibid **B 359** (1995) 423.

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

[58] H1 Collab., A. Aktas et. al., *Measurement of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at low Q^2 and x using the H1 vertex detector at HERA.*, Eur. Phys. J. **C 45**, 23 (2006).

[59] H1 Collab., A. Aktas et. al., *Measurement of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at high Q^2 using the H1 vertex detector at HERA.*, Eur. Phys. J. **C 40**, 349 (2005).

[60] H1 Collab., Aaron, F. D. et. al., *Measurement of charm and beauty structure functions using the H1 vertex detector at HERA.*, Eur. Phys. J. **C 65**, 89 (2010).

[61] P. M. Nadolsky, H. L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, D. Stump, W. K. Tung, C. P. Yuan, *Implications of CTEQ global analysis for collider observables.*, Phys. Rev. **78**, 013004 (2008).

[62] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC.*, Eur. Phys. J. C - Particles and Fields **63**, 189 (2009).

[63] R. S. Thorne, W. K. Tung, *pQCD formulations with heavy quark masses and global analysis.*, 0809.0714 (2008).

[64] M. Glück, P. Jimenez-Delgado and E. Reya, *Dynamical parton distributions of the nucleon and very small-x physics.*, Eur. Phys. J. **C 53**, 355 (2008).

[65] S. Alekhin and S. Moch, *Higher order qcd corrections to charged-lepton deep-inelastic scattering and global fits of parton distributions*,. Physics Letters **B 672**, 166 (2009)

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

[67] U. Schneekolth, *The HERA luminosity upgrade.* DESY-HERA-98-05, 1998.

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

[69] E. Hilger, *ZEUS coordinate system* (unpublished). ZEUS-86-17, Internal ZEUS Note, 1986.

[70] ZEUS Collab., *A Straw-Tube Tracker for ZEUS* (unpublished). ZEUS-98-046, Internal ZEUS Note, 1998.

[71] A. Dwurazny et al., Nucl. Instrum. and Meth. **A 277**, 176, (1989).

[72] ZEUS Collab., *A microvertex detector for ZEUS* (unpublished). ZEUS-97-006, internal ZEUS Note, 1997.

[73] W. Fallot-Burghardt et al., *Helix user manual*, v2.1. HD-ASICS-33-0697, 1999.

[74] C. Coldewey, *The ZEUS Micro Vertex Detector*, Nucl. Instrum and Meth. **A 447**, 44, (2000).

[75] B. Foster, J. Malos and D.H. Saxon et al., *The design and construction of the ZEUS central tracking detector.* Nucl. Instrum. Meth. **A 338**, 254, (1994).

[76] C.B. Brooks et al., *Development of the ZEUS central tracking detector*. Nucl. Instrum. Meth. **A 283**, 477, (1998).

[77] D. Bartsch, *Energy-loss measurement with the ZEUS Central Tracking Detector.* Ph.D. Thesis, Physikalisches Institut der Universität Bonn, Report BONN-IR-2007-05, 2007, available on `http://www-zeus.physik.uni-bonn.de/german/phd.html`.

[78] R. Hall-Wilton, N. McCubbin, P. Nylander, M. Sutton and M. Wing, *The CTD tracking resolution*. ZEUS-99-024, Internal ZEUS Note, 1999.

[79] E. Maddox, *Study of heavy quark production at HERA using the ZEUS microvertex detector*. Ph.D. thesis, University of Amsterdam, NIKHEFF, 2004.

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

[81] A. Barbenger et al., Nucl. Instrum. Meth. **A 382**, 419, (1996).

[82] A. Bamberger et al., *The small angle rear tracking detector of ZEUS*. Nucl. Instrum. Meth. **A 401**, 63, (1997).

[83] J. Andruszkow et al., *Luminosity measurement in the ZEUS experiment*. Acta Phys. Polon., **B 32**, 2025, (2001).

[84] H. Bethe and W. Heitler, *On the stopping of fast particles and on the creation of positive electrons*. Proc. Roy. Soc. Lond. **A 146**, 83, (1934).

[85] M. Helbich et al., *The Spectrometer system for measuring ZEUS luminosity at HERA*. Nucl. Inst. Meth. A **565**, 572-588 (2006).

[86] Inmos (Bristol) Limited, *The Transputer Book*. 2nd edition, 1989.

[87] S.M. Fisher at al., *ADAMO entity-relationship programming system: Reference Manual*. CERN Programming Techniques Group, ECP Division, 1993.

[88] O. Kind et al., *A ROOT-based client-server event display for the ZEUS experiment*. In the proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California.

[89] G. Hartner et al., *VCTRAK(3.07/04): Offline output information* (unpublished). ZEUS-97-064, internal ZEUS Note, 1997.

[90] G. Hartner, *VCTRAK Briefing: Program and Math* (unpublished). ZEUS-98-058, internal ZEUS Note, 1998.

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

[92] H.Stadie, *New Vertex Tools & How to Use Them* (unpublished). 2/8/2006, internal ZEUS presentation, 2006.

[93] P. Billoir, R. Frühwirth and M. Regler, Nucl. Instrum. and Meth., **A 241**, 115 (1985).

[94] S. Miglioranzi. *Beauty photoproduction at HERA II with the ZEUS experiment.* Ph.D. thesis, University College London, 2006.

[95] H. Abramowicz, A. Caldwell and R. Sinkus, *Neural network based electron identification in the ZEUS calorimeter.* Nucl. Instrum and Meth. **A 365**, 508 (1995) [hep-ex/9505004].

[96] R. Sinkus and T. Voss, *Particle identification with neural networks using a rotational invariant moment representation.* Nucl. Instrum and Meth. **A 391**, 360 (1997).

[97] T. Doeker, A. Frey. M. Nakao, *Electron position reconstruction - Update of the ELECPO routines.* ZEUS-94-123, internal ZEUS Note, 1994.

[98] B. Reisert, *CAL energy scale calibration* (unpublished). Talk at the ZEUS F_L Monthly Meeting, DESY, January 2008.

[99] J. Ferrando, *Energy scale and EM showers in 06-07 e^+p data* (unpublished). Talk at the ZEUS Physics Coordinator Meeting, DESY, 10.07.2008.

[100] B. Reisert, *Updates on the CAL energy scale calibration* (unpublished). Talk at the ZEUS F_L Meeting, DESY, March 2008.

[101] G.M. Briskin, *Diffractive dissociation in ep deep inelastic scattering.* DESY-THESIS-98-036, 1998.

[102] N. Tuning, *ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction.* ZEUS-01-021, internal ZEUS Note, 2001.

[103] M. Turcato, *Measurement of beauty photoproduction at HERA* (unpublished). DESY-THESIS-2003-039, 2003.

[104] V. Chiochia, *Measurement of beauty quark production in deep inelastic scattering at HERA*, DESY-THESIS-2003-031, 2003.

[105] S. Catania, Yu.L. Dokshitzerb, B.R. Webber, *The $K_\perp$-clustering algorithm for jets in deep inelastic scattering and hadron collisions.* Phys.Lett., **B 285**, 291 (1992).

[106] S. Catani et al., Nucl. Phys. **B 406**, 187, (1993).

[107] M. Corradi, M. Turcato, *Bottom photoproduction measured using decays into muons in dijet events in ep collisions at $\sqrt{s}$=318 GeV.* DESY-03-212, 2004. [hep-ex/0312057]

[108] M. Klein, *Physics at HERA*, edited by W. Buchmüller and G. Ingelman, Vol. 1, p.71, Hamburg, (1991).

[109] F. Jacquet, A. Blondel, *Detection of the charged current event - method II*, Proceedings of the Study for an *ep* facility for Europe (Hamburg, Germany) (U. Amaldi, ed.), DESY 79/48, 391, (1979).

[110] S. Bentvelsen et al., *Reconstruction of* (x, Q^2) *and extraction of structure functions in neutral current scattering at HERA*. Proceedings of the Workshop on Physics at HERA, edited by W. Buchmüller and G. Ingelman, Vol. 1, p.23, Hamburg, (1991).

[111] Evtake, 2007. `http://www-zeus.desy.de/~ifhuta/ZEUS_ONLY/evtake/`.

[112] B. Straub., *The EM Electron Finder* (unpublished), 1998, available on `http://www-zeus.desy.de/~straub/ZEUS_ONLY/doc/em.ps`.

[113] O.M. Kind, *Open Beauty Production at HERA.*, Ph.D. Thesis, Physikalisches Institut der Universität Bonn, 2006, available on `http://brock.physik.uni-bonn.de/zeus_pub.php`

[114] M. Jüngst, *Measurement of beauty and charm cross sections in photoproduction using decays into electrons with ZEUS at HERA*, Ph.D. Thesis, Physikalisches Institut der Universität Bonn, 2010, available on `http://brock.physik.uni-bonn.de/zeus_pub.php`

[115] ZEUS Collab., S. Chekanov et al., *Beauty photoproduction using decays into electrons at HERA*. Phys. Rev. **D 78**, 072001, 0805.4390 (2007).

[116] H.A. Bethe, Ann. Phys., **5** 325 (1930).

[117] H.A. Bethe, Z. Phys., **76** 293 (1932).

[118] W.W.M. Allison, J.H. Cobb, *Relativistic charged particle identification by energy loss.* Annual Review in Nuclear and Particle Physics, p. 253 (1980).

[119] Graziano Bruni, *p/pion separation using the ZEUS calorimeter.* ZEUS-04-018, internal ZEUS Note, 2004.

[120] M. Jüngst, *Elektronidentifikation mit dem ZEUS-Detektor und Bestimmung des Beauty-Produktionquerschnitts.* Diploma Thesis, Physikalisches Institut der Universität Bonn, 2005, available on `http://brock.physik.uni-bonn.de/zeus_pub.php`

[121] A. Elizabeth Nunico Quiroz, private communication, 2010.

[122] Markus Jüngst, private communication, 2010.

[123] M. Turcato, *Measurement of Beauty Photoproduction at HERA*. Ph.D. Thesis, Universita degli Studi di Padova, Report DESY-THESIS-03-039, 2002, available on `http://www-libraray.desy.de/preparch/desy/thesis/desy-thesis-03-039.ps.gz`.

[124] A.G. Yagües Molina, *Study of Beauty Photoproduction with the ZEUS Experiment at the Electron-Proton Collider HERA*, Humboldt University of Berlin, DESY:6561, 2008, available on `http://www-zeus.desy.de/physics/hfla/public/Theses/theses.html`

[125] V. Schönberg, *Measurement of beauty and charm photoproduction using inclusive secondary vertexing with the ZEUS detector at HERA*, Physikalisches Institut der Universität Bonn, 2010, available on `http://brock.physik.uni-bonn.de/zeus_pub.php`

[126] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **vol 77**, 219-238 (1993).

[127] Rene Brun et al., *The ROOT System Homepage*, CERN, 1995-2006. `http://root.cern.ch`.

[128] Rene Brun et al., *The ROOT User's Guide*, CERN, 2005. `http://root.cern.ch/root/doc/RootDoc.html`.

[129] N. E. Adam, et al., *Absolute Branching Fraction Measurements for D^+ and D^0 Inclusive Semileptonic Decays*, Phys. Rev. Lett. **97**, 251801 (2006), `hep-ex/0604044v4`

[130] B. Kahle, *Measurement of beauty-production in deep inelastic scattering at HERA II*, Ph.D. Thesis, Hamburg University, Hamburg (Germany), Report DESY-THESIS-2006-011, DESY, 2006.

[131] Particle Data Group, W. M. Yao et al., *Review of particle physics*, J. Phys. G 33, 1 (2006).

[132] V. Schönberg, private communication, 2010.

[133] S. U. Noor, *Measurement of neutral current electron-proton cross sections with longitudinally polarised electrons using the ZEUS detector.*, Ph.D. Thesis, York University, 2007.

[134] Y. Ri, *Measurement of neutral current deep inelastic ep-scattering cross sections with longitudinally polarized electrons with ZEUS at HERA.*, Ph.D. Thesis, Tokyo Metropolitan University, March 2009.

[135] L. Köpke, R. van Woudenberg, *Photon Conversions in the ZEUS Detector*, ZEUS-94-016, Internal ZEUS Note, 1994.

.

Acknowledgements

First and foremost of all, I praise Allah, the almighty for providing me this opportunity and granting me the capability to proceed successfully.

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study. It is not possible to list them all here. Their support in this effort is, however, greatly appreciated.

In the first place, I would like to express my sincere respect and deepest gratitude to my thesis adviser, Prof. Ian C. Brock, for giving me the opportunity to work in the ZEUS-BONN group and accepting me as a Ph.D student. It would have been impossible for me to complete this thesis without his patience, constructive feedback and insightful advice. I have benefited a lot from his stimulating ideas and suggestions. I could not have imagined having a better adviser for my Ph.D study. And I must add, he is not only a good adviser but also a very good human being.

Many thanks to Prof. Jochen Dingfelder, Prof. H. -W. Hammer and Dr. Gregor Kirfel for having agreed to participate in the thesis committee and for their interest and time they spent on the thesis and the disputation.

Dr. Mujahid Kamran, who is currently Vice Chancellor at the University of Punjab, Pakistan, deserves a special thanks as he was the one who initially motivated me to go for the higher studies in the field of particle physics. He has been and always will be an inspiration for me in my academic career.

I am very thankful to Prof. Erwin Hilger, Prof. Ewald Paul and Markus Jüengst who read different parts of this thesis and are responsible for numerous valuable corrections.

I would like to thank the whole ZEUS collaboration and especially the current and former members of the ZEUS-BONN group, the members of the heavy flavour group, ZEUS spokesperson, Aharon Levy, the ZEUS physics coordinators, Achim Geiser and Burkard Reisert and the EB members for the paper on this analysis, Massimo Corradi, Enrico Tassi and Ana Yagues. I am cordially thankful to Markus Jüengst for supporting me in this project from the very beginning to the concluding level. He always answered patiently all my questions and really helped me a lot to develop an understanding of the subject and especially the data analysis. Without his help this work would have never seen a happy end.

Many friends have helped me stay sane through these difficult years. Their support and care helped me overcome setbacks and stay focused on my graduate study. I greatly value their friendship and I deeply appreciate their belief in me. I am also grateful to the Pakistani families that helped me adjust to a new country. Once again, I wish to pay my heartiest thanks to very honourable and affectionate Prof. Erwin Hilger, who was always eager to discuss about my culture, country and life in various circumstances. I morally obliged to mention the names of Rosi and Liz who showed their love, care and made me feeling that I am not alone, when when I was sick.

Most importantly, none of this would have been possible without the love, emotional support and patience of my family. My very special thanks to the one person to whom I owe everything I am today, my father, Muhammad Ashraf, for instilling in me the love of reading and the pursuit of knowledge from my early sapling years, and to him I dedicate this thesis. He was very enthusiastic when I told him that I had enrolled for a Ph.D and although he is no longer alive to see its completion, he is forever remembered. He always had that unswerving and almost unnerving faith in me which does not allow for anything less than success. If it were not for him, I would not be sitting here today.

My deepest gratitude goes to my beloved mother for her constant encouragement, helpful advice, prayers, and affection. I am indebted to my sweet sisters, Uzma and Nadia and to all my brothers, who always listened to all my problems and boosted me morally.

Hopefully, this Ph.D dissertation will not be the end of my journey in seeking for more knowledge to understand the meaning of life.