Measurement of Charm and Beauty Production in Deep Inelastic Scattering at HERA and Test Beam Studies of ATLAS Pixel Sensors

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Vladyslav Libov
aus Uzhhorod (Ukraine)

Hamburg
2013

Gutachter der Dissertation:	PD Dr. Olaf Behnke Prof. Dr. Robert Klanner
Gutachter der Disputation:	Prof. Dr. Johannes Haller PD Dr. Achim Geiser
Datum der Disputation:	27. November 2012
Vorsitzender des Prüfungsausschusses:	Dr. Georg Steinbrück
Vorsitzender des Promotionsausschusses:	Prof. Dr. Peter Hauschildt
Dekan der MIN Fakultät:	Prof. Dr. Heinrich Graener

Abstract

A measurement of charm and beauty production in Deep Inelastic Scattering at HERA is presented. The analysis is based on the data sample collected by the ZEUS detector in the period from 2003 to 2007 corresponding to an integrated luminosity of $354\,\mathrm{pb}^{-1}$. The kinematic region of the measurement is given by $5 < Q^2 < 1000\,\mathrm{GeV}^2$ and $0.02 < y < 0.7$, where Q^2 is the photon virtuality and y is the inelasticity. A lifetime technique is used to tag the production of charm and beauty quarks. Secondary vertices due to decays of charm and beauty hadrons are reconstructed, in association with jets. The jet kinematics is defined by $E_T^{\mathrm{jet}} > 4.2(5)\,\mathrm{GeV}$ for charm (beauty) and $-1.6 < \eta^{\mathrm{jet}} < 2.2$ for both charm and beauty, where E_T^{jet} and η^{jet} are the transverse energy and pseudorapidity of the jet, respectively. The significance of the decay length and the invariant mass of charged tracks associated with the secondary vertex are used as discriminating variables to distinguish between signal and background. Differential cross sections of jet production in charm and beauty events as a function of Q^2, y, E_T^{jet} and η^{jet} are measured. Results are compared to Next-to-Leading Order (NLO) predictions from Quantum Chromodynamics (QCD) in the fixed flavour number scheme. Good agreement between data and theory is observed. Contributions of the charm and beauty production to the inclusive proton structure function, $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$, are determined by extrapolating the double differential cross sections using NLO QCD predictions.

Contributions to the test beam program for the Insertable B-Layer upgrade project of the ATLAS pixel detector are discussed. The test beam data analysis software package EUTelescope was extended, which allowed an efficient analysis of ATLAS pixel sensors. The USBPix DAQ system was integrated into the EUDET telescope allowing test beam measurements with the front end chip FE-I4. Planar and 3D ATLAS pixel sensors were studied at the first IBL test beam at the CERN SPS.

Kurzfassung

In dieser Arbeit wird eine Messung von Charm und Beauty Produktion in tief-unelastischer ep-Streuung bei HERA vorgestellt. Die Analyse basiert auf Daten, die im Zeitraum von 2003 bis 2007 mit dem ZEUS-Detektor aufgezeichnet wurden, und einer integrierten Luminosität von $354\,\mathrm{pb}^{-1}$ entsprechen. Der kinematische Bereich der Messung wird durch $5 < Q^2 < 1000\,\mathrm{GeV}^2$ und $0.02 < y < 0.7$ definiert, wobei Q^2 die Virtualität und y die Inelastizität ist. Um die Ereignisse mit Charm- und Beauty-Quarks zu identifizieren wurde die endliche Lebensdauer der durch die schwache Kraft zerfallenden Grundzustände der Charm- und Beauty-Hadronen ausgenutzt. Sekundäre Zerfallspunkte von Charm- und Beauty-Hadronen, die mit Jets assoziiert sind, werden rekonstruiert. Die Kinematik der Jets wird auf den folgenden Bereich eingeschränkt: $E_T^{\mathrm{jet}} > 4.2(5)\,\mathrm{GeV}$ für Charm (Beauty) und $-1.6 < \eta^{\mathrm{jet}} < 2.2$ für Charm und Beauty, wobei E_T^{jet} und η^{jet} die transversale Energie bzw. Pseudorapidität ist. Die Signifikanz der Zerfallslänge sowie die Masse der geladenen Spuren vom Zerfallsvertex werden als diskriminierende Variablen benutzt, um zwischen dem Charm-Signal, dem Beauty-Signal und dem Untergrund zu unterscheiden. Differenzielle Wirkungquerschnitte von Jet-Produktion in Charm- und Beauty-Ereignissen werden als Funktion von Q^2, y, E_T^{jet} and η^{jet} gemessen. Die Ergebnisse werden mit Next-to-Leading Order (NLO) Vorhersagen der Quantenchromodynamik (QCD) im "Fixed Flavour Number Scheme" verglichen. Die Theorie stimmt gut mit den Daten überein. Die Beiträge von Charm- und Beauty-Produktion zur inklusiven Protonstrukturfunktion, $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$, werden durch Extrapolation von doppelt differenziellen Wirkungsquerschnitten mit Hilfe der NLO QCD Vorhersagen bestimmt.
Beiträge zur Teststrahlmessung-Program von Insertable B-Layer Upgrade-Projekt für ATLAS Pixeldetektor werden vorgestellt. Die Teststrahlmessung-Analysesoftware EUTelescope wurde erweitert, was eine effiziente Analyse von ATLAS-Pixelsensoren erlaubt hat. Das USBPix DAQ-System wurde in das EUDET-Teleskop integriert, was Teststrahlmessungen mit dem Front-End-Chip FE-I4 ermöglicht. Planar- und 3D- ATLAS Pixelsensoren wurden während der ersten IBL-Teststrahlmessung am CERN SPS studiert.

To my Family

Contents

1 Heavy Flavour Production at HERA **5**
- 1.1 Deep Inelastic Scattering at HERA . . . 5
- 1.2 Quantum Chromodynamics . . . 7
- 1.3 Heavy Flavour Production at HERA . . . 9
- 1.4 Heavy Flavour Measurements in DIS at HERA . . . 10
  - 1.4.1 Charm Production . . . 11
  - 1.4.2 Beauty Production . . . 16
- 1.5 Monte Carlo Generators . . . 19

2 Experimental Setup **21**
- 2.1 HERA collider . . . 21
- 2.2 Overview of the ZEUS Detector . . . 23
- 2.3 Central Tracking Detector . . . 25
- 2.4 Microvertex Detector . . . 26
- 2.5 Uranium Calorimeter . . . 29
- 2.6 Luminosity Monitor . . . 31
- 2.7 Trigger and Data Acquisition System . . . 31

3 Event Reconstruction **33**
- 3.1 Scattered Electron Identification . . . 33
- 3.2 Track reconstruction . . . 35
- 3.3 Vertex reconstruction . . . 36
- 3.4 Energy Flow Objects . . . 37
- 3.5 Jet Reconstruction . . . 39
- 3.6 Kinematic variables . . . 40
  - 3.6.1 Electron Method . . . 40
  - 3.6.2 Jacquet-Blondel Method . . . 40
  - 3.6.3 Double-Angle Method . . . 41

4 Tracking Inefficiency due to Hadronic Interactions 43
4.1 Method . . . 43
4.2 Data and Monte Carlo Samples . . . 45
4.3 Selection . . . 46
4.4 Monte Carlo Reweighting . . . 50
4.5 Control Distributions . . . 52
4.6 Results . . . 58
4.7 Summary . . . 62

5 Event Selection and Signal Extraction 63
5.1 Charm and Beauty Tagging with Secondary Vertices . . . 63
5.2 Data and Monte Carlo Samples . . . 64
5.3 Event Selection . . . 66
5.4 Secondary vertex reconstruction . . . 69
5.5 Signal Extraction . . . 73
5.6 Corrections . . . 76
5.7 Control Distributions . . . 82

6 Systematic Uncertainties 91
6.1 The Scan Technique . . . 91
6.2 Sources of Uncertainty . . . 93
6.3 Summary . . . 105

7 Results 107
7.1 NLO QCD Predictions . . . 107
7.2 Differential Jet Cross Sections in Charm Events . . . 109
7.3 Differential Jet Cross Sections in Beauty Events . . . 115
7.4 $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ Determination . . . 119
7.5 Summary . . . 120

8 Test Beam Studies of the ATLAS Pixel Sensors 123
8.1 Physics Goals of the Large Hadron Collider . . . 123
8.2 The LHC Machine . . . 124
8.3 The ATLAS Experiment . . . 126
8.4 ATLAS Pixel Detector . . . 128
8.4.1 Limitations of the Pixel Detector . . . 130
8.5 Insertable B-Layer . . . 132
8.5.1 Sensor Technology Choice . . . 132
8.6 EUDET Telescope . . . 133
8.6.1 Overview of the System . . . 133
8.6.2 The Sensor . . . 135

8.6.3 Triggering . . . 136
8.6.4 EUDAQ - Data Acquisition Software . . . 137
8.6.5 Performance of the System . . . 138
8.7 EUTelescope Data Analysis Framework . . . 139
8.7.1 Raw Data Conversion and Clustering . . . 140
8.7.2 Alignment with MillepedeII . . . 141
8.7.3 Tracking . . . 142
8.7.4 DUT Characterisation . . . 142
8.8 Modification of EUTelescope . . . 142
8.8.1 Integration of ATLAS Pixel Data Type . . . 143
8.8.2 Improvement of Alignment with the Correlation Information . . . 144
8.8.3 Iterative DUT Alignment . . . 146
8.8.4 ATLAS Pixel Analysis Processor . . . 148
8.8.5 Summary . . . 148
8.9 First IBL Test Beam Campaign . . . 149
8.9.1 Experimental Setup and Samples . . . 149
8.9.2 Data Analysis . . . 150
8.9.3 Results . . . 151
8.9.4 Conclusions . . . 153
8.10 Integration of USBPix into EUDET Telescope . . . 157
8.10.1 Overview of the USBPix System . . . 157
8.10.2 Integration to the EUDET Telescope . . . 158
8.10.3 System Test at the October 2010 IBL Test Beam . . . 160
8.10.4 Summary . . . 162
8.11 Summary and Outlook . . . 162

A Mirrored Significance **165**

B Cross Section Tables **169**

Introduction

Colliding high energy elementary particles has been a most fruitful way to study the structure of matter. In the beginning of the XX century, experiments on scattering of alpha particles from a radioactive source on gold atoms by Rutherford, Geiger and Marsden [1, 2] showed that the atoms must have a massive, positively charged core, much smaller than the atom itself. The Thomson model which stated that the positive charge is distributed evenly throughout the whole atom, was in sharp contradiction with the observed angular distribution of scattered alpha particles. Thus, the existence of the *atomic nucleus* was established. With the advent of particle accelerators, it became possible to perfom scattering at higher energies. In 1954, Hofstadter and McAllister performed elastic electron-proton scattering experiments [3, 4], which revealed the electric charge distribution in the proton and showed that they are extended objects with a size of around 1 fm. *Inelastic* electron-proton scattering experiments showed that the proton can be excited to a higher mass state [5]. Hence, it was established that the proton is not an elementary and pointlike particle, as it was believed before, but a composite object. Starting from 1967, similar experiments were carried out by Friedman, Kendall and Taylor, however the incident electron energy was substantially increased thanks to the new two mile linear accelerator at SLAC [6]. This allowed to study the *deep inelastic scattering* for the first time, i.e. scattering at high values of the momentum transfer from electrons to protons. The observations were suggesting that scattering occurs on point-like charged objects inside the proton; the assumption of the smooth charge distribution in the proton was contradicting the data, very similarly as in the Rutherford experiment the model of smooth charge distribution in the atom was ruled out. Thus, *partons* were discovered, which were then identified as *quarks* and *gluons*, the smallest building blocks of matter known today (together with the *leptons*, *gauge bosons* and, most likely, the *Higgs Boson*) [7].

In 1992 the HERA accelerator [8] located at the research centre DESY[1] (Hamburg, Germany) started its operation. Electrons (positrons) were collided with protons at an unprecedented centre-of-mass energy of more than 300 GeV. HERA was and still remains the world only lepton-hadron collider. For 15 years, experiments here were collecting unique data, which allowed in-depth studies of the partonic structure of the proton. The

[1] **D**eutsches **E**lektronen-**Sy**nchrotron.

parton density functions (PDF) of the proton, characterising the fractions of the proton longitudinal momentum carried by the quarks and gluons in the proton were determined with remarkable precision. In particular, one of the interesting findings was a dominant contribution from the sea (virtual) quarks and from gluons in the domain of low parton momenta [9].

Besides being a tool to reveal the proton structure, HERA also serves as a unique laboratory to study the *strong interaction*, one of the three forces known in Nature (along with the electroweak and the gravitational force). The theory that describes it, Quantum Chromodynamics (QCD) [10], has been extensively tested at HERA. One of the interesting domains in this field is the production of *heavy quarks*, i.e. of *charm* and *beauty* quarks (production of the *top* quark is not allowed kinematically). The theoretical description of this process is challenging due to the large mass of the quarks. On one hand, the mass provides a hard scale allowing perturbative calculations to be made, on the other hand two more hard scales are available at HERA: the quark transverse momentum p_T and the photon virtuality Q^2. This complicates the calculations and several approaches (*schemes*) exist to treat this problem. Therefore, the HERA data can be useful to test and improve these schemes. At leading order of the perturbative QCD, the production of heavy quarks proceeds via the interaction of the virtual photon emitted by the incoming electron with a gluon from the proton, the so called Boson-Gluon Fusion (BGF). Thus, this process is directly sensitive to the distribution of the gluons in the proton and along with QCD tests, heavy flavour measurements may help to further improve the knowledge of the proton structure. Finally, heavy flavour measurements allow the determination of the heavy quark masses.

In this work, a measurement of charm and beauty quarks production in Deep Inelastic Scattering (DIS) with the ZEUS detector at HERA was performed. A secondary vertex technique to tag heavy flavour production was applied to a large data sample corresponding to an integrated luminosity of $354\,\mathrm{pb}^{-1}$, collected from 2003 to 2007 (denoted as the HERA II running period). The method was developed in earlier works [11–13] and was applied for the first time to DIS in Ref. [14]. Among the goals set specifically for this thesis were: a further development of the method, detailed studies of systematic uncertainties, evaluation of the tracking efficiency of the detector and extraction of the final results for the publication. The tagging technique relies heavily on the precise silicon tracking close to the interaction point. In order to contribute to future heavy flavour measurements, some work was devoted to the silicon pixel detector in ATLAS which is one of the next generation experiments at the *Large Hadron Collider*, the highest energy particle collider. The activities were related to beam tests of pixel sensors.

The report is organised in the following way. Chapter 1 presents a short introduction to the physics of heavy quarks in Deep Inelastic Scattering at HERA. In Chapter 2, details are given on the HERA collider and on the ZEUS detector. In the following chapter, algorithms employed in this thesis for reconstruction of high level objects (such as particle tracks or hadronic jets) are discussed. Afterwards, contributions from this thesis

are presented. In Chapter 4, a study of the tracking efficiency of the ZEUS detector is described. In the following, details are given on the charm and beauty measurement. The event selection and the signal extraction technique are presented in Chapter 5. Next, systematic uncertainties studies are discussed in Chapter 6. The results of the measurement are presented and confronted to theoretical predictions in Chapter 7. The findings are compared to previous measurements. In Chapter 8 studies of the ATLAS pixel sensors are described. Finally, the conclusions are given.

Chapter 1

Heavy Flavour Production at HERA

This chapter introduces the physics of heavy quarks in deep inelastic scattering (DIS) at HERA. First, a general overview of DIS is given and the proton structure functions are introduced. Afterwards, the main features of the theory of strong interactions, Quantum Chromodynamics (QCD) are described. The scope then moves to the charm and beauty quark production at HERA. The theoretical description of this process by the QCD is discussed and recent measurements are briefly overviewed. The chapter is closed with the discussion on the Monte Carlo generators which are used in this thesis to simulate signal and background processes.

1.1 Deep Inelastic Scattering at HERA

The term *Deep Inelastic Scattering* (DIS) is used to denote the scattering of a lepton on a nucleon (or nuclei) at a high centre-of-mass energy, when a large squared momentum (above 1 GeV^2) is transferred from the lepton to the nucleon [15]. The reaction can be written as:

$$l(k) + N(p) \rightarrow l'(k') + X(p_X),$$

where l (l') denote the initial (final) state lepton, N is the initial state nucleon and X is an arbitrary hadronic final state. The notation for the four momentum vector of the corresponding object is given in parentheses. In the leading order, this reaction proceeds via exchange of an electroweak gauge boson. In case of a neutral boson exchange (γ, Z^0) one speaks about Neutral Current (NC) while for a charged boson exchange ($W^{\pm}$) the process is called Charged Current (CC).

At HERA, DIS occurs in collisions of electrons (or positrons) with an energy of $E_e = 27.5\,\mathrm{GeV}$ with protons of energy $E_p = 920\,\mathrm{GeV}$. The nucleon N stands for the proton, the initial lepton l' is the electron (positron); the final state lepton is the electron (positron) for NC while it becomes the neutrino (antineutrino) for the CC process. In

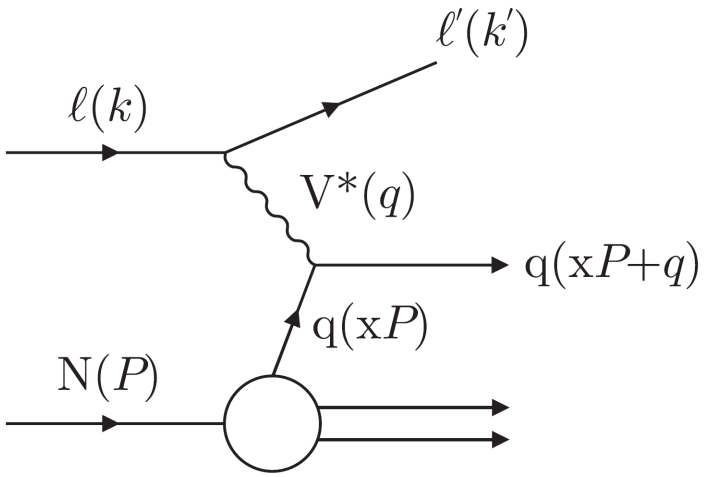


Figure 1.1: The leading order of lepton-nucleon scattering in the naive Quark Parton Model [16]. The lepton scatters off a quark in the proton via exchange of a vector boson V^*.

this thesis, the Neutral Current reaction is studied, hence the following discussion omits the CC.

The process is described by the following kinematic variables which are Lorentz invariants:

- $s = (k + p)^2$ – centre-of-mass energy ($\sqrt{s} = 318\,\text{GeV}$ for HERA);
- $Q^2 = -q^2 = -(k - k')^2$ – the negative four momentum squared of the exchanged boson, or its *virtuality*;
- $y = p \cdot q / p \cdot k$ – the energy transferred from the lepton to the proton in the proton rest frame or the *inelasticity*;
- $x = Q^2/2p \cdot q$ – Bjorken x variable (see below for interpretation).

The following relation holds (neglecting the masses of the electron and the proton):

$$Q^2 = s\,x\,y. \tag{1.1}$$

Thus, at fixed s there are only two independent variables (e.g. Q^2 and x).

In the leading order (one boson exchange), the double differential cross section of NC DIS as a function of Q^2 and x is given by:

$$\frac{d^2\sigma^{e^\pm p \to e^{\pm\prime} X}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[Y_+ F_2(x,Q^2) - y^2 F_L(x,Q^2) \mp Y_- x F_3(x,Q^2)\right], \tag{1.2}$$

where $Y_\pm = 1 \pm (1-y)^2$ and $F_2(x,Q^2)$, $F_L(x,Q^2)$ and $F_3(x,Q^2)$ are the proton *structure functions*. The F_L function is called the longitudinal structure function; F_3 is negligible for $Q^2 \ll M_Z^2$ where M_Z is the mass of the Z^0 boson.

In the *naive quark parton model* (QPM) [17], the proton consists of three *valence* quarks (two up quarks and one down quark) and of *sea* quarks. The DIS in this model is viewed as the interaction of the lepton with one of the quarks in the proton (*struck quark*), see Fig. 1.1. The Bjorken x is then interpreted as the fraction of the proton momentum carried by the struck quark[1]. In this case the cross section reduces to (neglecting Z^0 exchange):

$$\frac{d^2\sigma^{e^\pm p \to e^{\pm\prime} X}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[(1+(1-y)^2)F_2(x,Q^2)\right].$$

Here $F_2(x,Q^2)$ depends only on x and reflects the distributions of quarks in the proton:

$$F_2(x) = \sum_i e_i^2 x f_i(x),$$

where the sum runs over all quark flavours in the proton (valence and sea) and $f_i(x)$ is the probability to find a quark i with a proton momentum fraction of x, or the quark *parton distribution function* (PDF). Hence, in this simple picture, the inclusive NC DIS cross-section is directly sensitive to the quark PDFs in the proton.

However, the general case of DIS involves also the strong interaction This complicates the interpretation of the structure functions in terms of PDFs, since the inclusive scattering contains various subprocesses, but it also turns HERA into a unique laboratory for studies of the strong force. The Quantum Chromodynamics (QCD), the theory of the strong interaction, is the topic of the next section.

1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory of the strong interaction, the force which is responsible for holding the nucleons together inside nuclei and for binding the quarks inside hadrons. Together with the electroweak theory, it is part of the Standard Model, the most successful particle physics theory. QCD describes interactions of objects possessing the *colour* degree of freedom, quarks and gluons. Quarks can be thought of as matter constituents (similarly to the electron), while gluons are the *field quanta* mediating the interaction (as the photons mediate the electromagnetic force in Quantum Electrodynamics). QCD is a *gauge* quantum field theory: its Lagrangian obeys certain symmetries, i.e. is invariant under certain *gauge transformations*, which defines the properties of the theory. For QCD this is the symmetry under the $SU(3)$-group transformation in the colour space. The Non-Abelian nature of this group leads to the fact that field quanta

[1] It is assumed that quarks have no transverse momentum with respect to the proton momentum.

(gluons) can interact between themselves, unlike photons in Quantum Electrodynamics (QED). This leads to a number of properties that distinguish QCD from QED.

Among the most striking features of QCD are *asymptotic freedom* and *confinement*. Both are related to the *running* of the coupling constant of QCD, α_s, i.e. its dependence on the scale of a process (e.g. the momentum transfer). In terms of Feynman diagrams, the running of the coupling appears due to higher order virtual corrections (*loops*) to the leading order diagrams (e.g. to a qqg vertex). Including diagrams with loops in a calculation leads to infinite amplitudes. In order to overcome this problem, the *renormalisation* procedure is invoked which is the redefinition of the original theory parameters (the coupling α_s in this case) in such a way that they are finite at some arbitrary scale μ_R called *renormalisation scale*, while the infinities are absorbed into bare (non-renormalised) parameters. This introduces a dependence of observed quantities on an arbitrary parameter μ_R, which is unphysical; requiring that observables are independent of μ_R leads to the *renormalisation group equation*. Its solution yields a relation between the coupling at an arbitrary scale Q to that at the scale μ_R. For the one-loop calculation one gets [18]:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda^2)},$$

where Q^2 is the scale of the process, n_f is the number of active flavours (appearing in the loops) at the scale Q^2, Λ is a parameter of QCD depending on particular renormalisation procedure (scheme) chosen and on μ_R as well as on the number of active flavours n_f. This expression shows that α_s decreases with increasing scale. At large scales (or at small distances), the interaction is weak, the property called asymptotic freedom, whereas at small scales (large distances) the interaction becomes strong. The latter leads to confinement – the fact that coloured objects (e.g. a quark and an antiquark) cannot be taken apart by distances larger than around 1 fm (quarks are confined inside hadrons). In contrast, the coupling constant of QED increases with increasing the scale (decreasing the distance).

The smallness of α_s at high scales (small distances) allows to apply perturbative approach to the calculations, i.e. to expand the cross section in series of α_s. At small scales, when the coupling is large, this is not anymore possible, since the series does not converge. The transition between two regimes happens at scales of the order of the Λ parameter which has a value around the typical hadronic mass, $\Lambda \sim 0.1 \div 0.5\,\mathrm{GeV}$.

Quarks and gluons are not directly observable due to confinement. Only colourless hadrons, inherently non-perturbative objects, can be present in the initial or final state of any reaction. Hence, generally speaking, perturbative techniques are not applicable to any real process involving the strong force. The *Factorisation Theorem* [19] rescues the situation. It states that under certain conditions one can separate the hard process that is characterised by a large scale and soft processes which are associated to initial and final state hadrons. The hard process can be calculated with perturbative techniques, while

the soft part is usually taken from experiment. For the cross section σ of a certain QCD process in ep-collisions this can be illustrated by the following simplified formula:

$$\sigma = \int dx\, f(x, \mu_F)\, \hat{\sigma}(x, \mu_F),$$

where f is the parton density function (PDF) of a parton that is involved in the process, taken from experiment, $\hat{\sigma}$ is the calculable in the perturbative approach parton level cross section of the process. The cross section has to be integrated over the fractional momentum of the parton, x. A dependence of the PDF and of the hard cross section on a parameter μ_F, called *factorisation scale*, was indicated. The factorisation scale represents a threshold at which splitting into hard and soft processes is done: at scales above μ_F, the process is treated as perturbative; processes at lower scales are absorbed in the PDF. The dependence of the PDFs on the scale is given by the DGLAP [20–23] evolution equations.

1.3 Heavy Flavour Production at HERA

One of the domains of QCD studies at HERA is the production of charm and beauty quarks. The leading order Feynman diagram of this process (Boson-Gluon Fusion) is shown in Fig. 1.2. There is a gluon in the initial state, hence the process is directly sensitive to the gluon density in the proton.

The mass of the heavy quarks, m, represents a hard scale, hence enabling the usage of perturbative techniques. However, the situation is complicated due to the fact that in

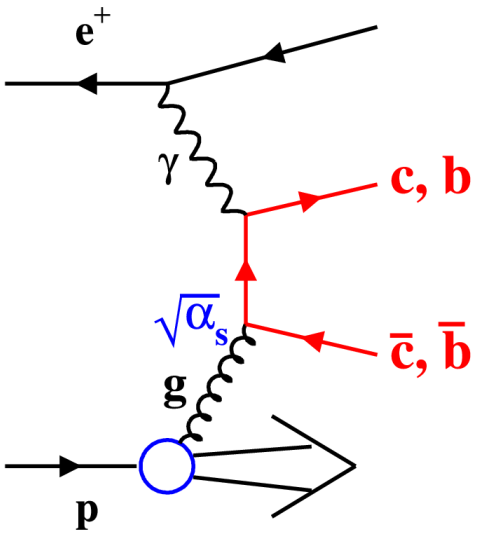


Figure 1.2: Leading order Feynman diagram for heavy quarks production at HERA.

some regions of phase space, other hard scales can be present in addition to the mass, such as the photon virtuality Q^2 and the quark transverse[2] momentum p_T. This is the so-called *multiple scale problem*. It might spoil the convergence of the perturbative series, due to the presence of the terms protortional $[\alpha_s \log(p_T^2/m^2)]^n$ or $[\alpha_s \log(Q^2/m^2)]^n$, where n is the order of a calculation. These terms represent collinear gluon radiation from the heavy quarks.

Several schemes for heavy flavour calculations are available. The *fixed flavour number scheme* (FFNS) or *massive* scheme [24–26] is a rigorous QCD calculation which fully takes heavy quark masses into account. It is expected to be appropriate at small scales, $p_T \sim m$ and $Q^2 \sim m^2$, while at very high scales it may break down due to the multiple scale problem. A calculation program called HVQDIS [27] is available; it provides NLO ($\mathcal{O}(\alpha_s^2)$) QCD predictions in this scheme for any differential cross sections of heavy flavour production at HERA.

An alternative approach is the Zero-Mass Variable Number Scheme (ZM-VFNS) [28, 29] or *massless* scheme. Here, the heavy quarks are treated as massless and they appear as active sea quarks in the proton. In this scheme it is possible to resum the large logarithms mentioned above at all orders of perturbation theory, hence the multiple scale problem does not affect predictions. However, since this approach neglects the mass, it is expected to work only far above the threshold of heavy quark production, i.e. for $Q^2 \gg m^2$ or $p_T \gg m$.

The Generalised Variable Flavour Number Schemes (GM-VFNS) [30–38] combines the massive and massless schemes. It converges to FFNS and to ZM-VFNS at small and large scales, respectively. In the intermediate region it performs a certain interpolation between the two schemes. There is some arbitrariness in the treatment of this interpolation and various approaches are used.

Measurements of heavy flavour production at HERA can help to distinguish between different theoretical schemes. They can also provide valuable information on charm and beauty quark masses, since predictions are very sensitive to these parameters. In the next section, several measurements of charm and beauty production in DIS at HERA are discussed which address these questions.

1.4 Heavy Flavour Measurements in DIS at HERA

The production of charm and beauty quarks has been studied extensively at HERA [39–69], and is still an active field of research. Various techniques were used to identify events with heavy flavours (*tagging methods*). In this section, selected results on charm and beauty production in DIS by the H1 and ZEUS experiments are discussed.

[2] A right-handed cartesian coordinate system is used, with the Z axis pointing along the incident proton direction. The transverse momentum p_T of a particle is given by $p_T = p \sin\theta$, where θ is the particle polar angle in this coordinate system and p is its momentum. The pseudorapidity η is defined as $\eta = -\ln\tan(\theta/2)$.

1.4.1 Charm Production

In a measurement by the ZEUS Collaboration [44], charm quarks were identified by reconstruction of $D^*(2010)$ mesons decaying via the *golden channel*: $D^* \to D^0\pi \to K\pi\pi$. The restrictions on the transverse momentum of D^* mesons, $p_T(D^*)$, and on their pseudorapidity, $\eta(D^*)$, were $1.5 < p_T(D^*) < 15\,\text{GeV}$ and $|\eta(D^*)| < 1.5$, respectively. The kinematic range was $1.5 < Q^2 < 1000\,\text{GeV}^2$ and $0.02 < y < 0.7$. Figure 1.3(a) shows the differential cross section as a function of the photon virtuality Q^2. The data are compared to NLO QCD predictions in the FFNS scheme. A good agreement is found in the full Q^2 range.

The same technique of charm tagging was applied in a measurement by the H1 Collaboration [56]. The phase space for the D^* mesons was $p_T(D^*) > 1.25\,\text{GeV}$ and $|\eta(D^*)| < 1.8$. The kinematic range was $5 < Q^2 < 100\,\text{GeV}^2$ and $0.02 < y < 0.7$. For the cross sections

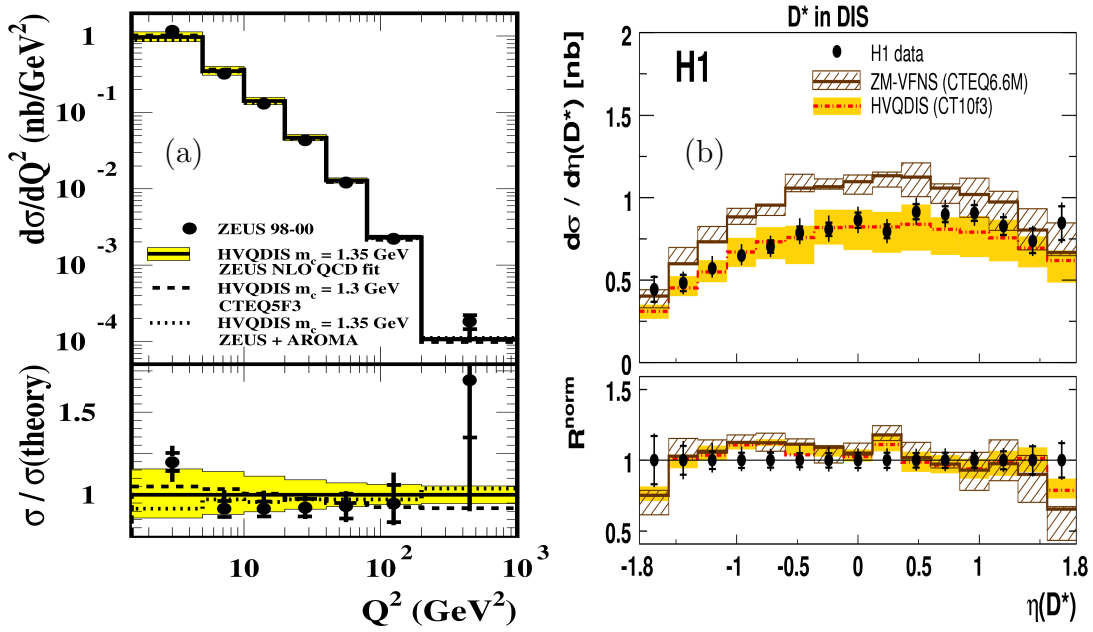


Figure 1.3: (a) Differential cross section of $D^*(2010)$ meson production as a function of Q^2 from a measurement by the ZEUS Collaboration [44]. The points represent the data, the lines represent NLO QCD predictions in the FFNS scheme, obtained with different PDF sets and hadronisation models. (b) Differential cross section of $D^*(2010)$ meson production as a function of $\eta(D^*)$ from the H1 experiment [56]. The points show the data, the dot dashed and the solid lines represent NLO QCD predictions in the FFNS and ZM-VFNS schemes, respectively. For both plots, the inner (outer) error bars show the statistical (total) uncertainties of the measurements while the bands indicate uncertainties of the predictions. The kinematic regions of the measurements are given in the main text.

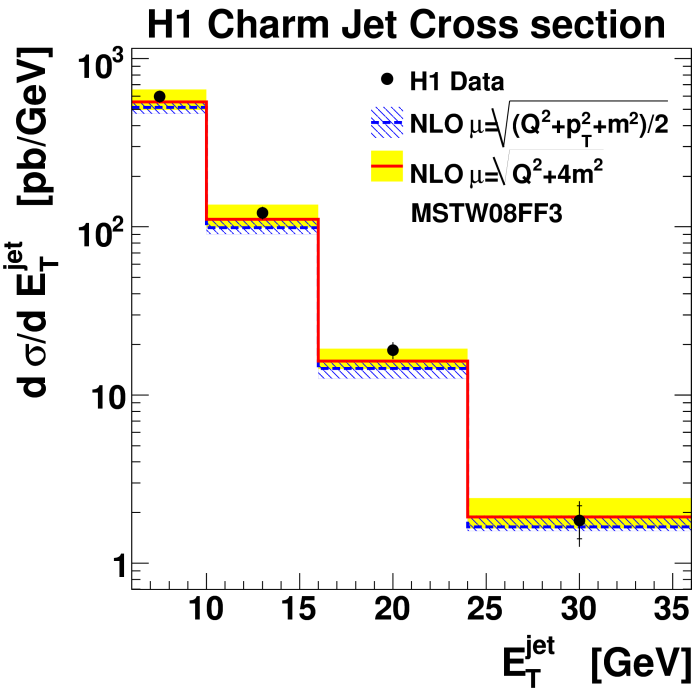


Figure 1.4: Differential cross section of charm jet production as a function of the jet transverse energy E_T^{jet} in the inclusive lifetime analysis by the H1 Collaboration [55]. The points represent the measurement, the lines show the NLO QCD predictions in the FFNS scheme, the bands indicate uncertainties of the predictions. The kinematic region of the measurement is given in the main text.

presented here, an additional cut on the D^* transverse momentum in the virtual photon-proton ($\gamma^* p$) centre-of-mass frame, $p_T^*(D^*) > 2.0\,\text{GeV}$, was applied in order to allow comparisons to predictions in the ZM-VFNS scheme. Figure 1.3(b) shows the differential cross section as a function of the D^* meson pseudorapidity $\eta(D^*)$. Data are compared to NLO QCD predictions in the FFNS and ZM-VFNS schemes. It can be concluded, that the FFNS predictions describe the data well, while the calculations in the ZM-VFNS approach overshoot the data.

Although a full reconstruction of charmed mesons provides a good signal to background ratio, this method suffers from low branching ratios of decays into a particular final state. The H1 collaboration applied an *inclusive lifetime technique* to tag heavy quarks [55]. It makes use of the fact that the ground states of the hadrons containing heavy quarks have typically long lifetimes. No specific decay modes are required in the method. Charm quarks were measured in association to jets. The phase space region in the jet transverse energy, E_T^{jet}, and jet pseudorapidity, η^{jet}, was $E_T^{\text{jet}} > 6\,\text{GeV}$ and $-1.0 < \eta^{\text{jet}} < 1.5$, respectively. The kinematic range was $Q^2 > 6\,\text{GeV}^2$ and $0.07 < y < 0.625$. Figure 1.4

shows the differential cross section for production of jets containing a charm quark, as a function of the jet transverse energy E_T^{jet}. Data are compared to NLO QCD predictions in the FFNS scheme. A good agreement is found within uncertainties.

In conclusion, measurements show that NLO QCD predictions in the fixed flavour number scheme provide a good description of charm production at HERA in the probed kinematic region. NLO QCD predictions based on the ZM-VFNS scheme which neglects the charm quark mass, fail to describe the data.

Recently, a combination of various charm measurements by H1 and ZEUS was performed [70]. For every measurement, the *reduced* charm cross section $\sigma_{\text{red}}^{c\bar{c}}$ was determined before the combination. Up to a kinematic factor, it represents the inclusive double differential cross section of charm quark-antiquark pair production, as a function of x and Q^2:

$$\sigma_{\text{red}}^{c\bar{c}} = \frac{d^2\sigma^{ep\to c\bar{c}X}}{dxdQ^2}\frac{xQ^4}{2\pi\alpha^2(1+(1-y)^2)}.$$

In order to determine $\sigma_{\text{red}}^{c\bar{c}}$, an extrapolation procedure needs to be performed, since all measurements are done in a restricted region of charm quark kinematics (e.g. due to transverse momentum and pseudorapidity cuts on D^* mesons; see also a discussion in Section 7.4). The extrapolation is done with the NLO QCD in the FFNS scheme. The same procedure is used for all measurements. After extrapolation, the combination is performed. A procedure which takes into account correlated systematic uncertainties between different measurements is used. It allows a reduction of systematic uncertainties, in addition to obvious reduction of statistical uncertainties. Figure 1.5 shows the result of the combination, together with various measurements that enter into the combination. It is readily seen that uncertainties reduce drastically after the combination, which increases the potential of the data to distinguish between various theory schemes. As an example, Figure 1.6 shows a comparison of the combined data to NLO and NNLO QCD predictions from the NNPDF group (see Ref. [70] for details). One is lead to the conclusion that the NNLO calculation agrees to data better than the NLO predictions.

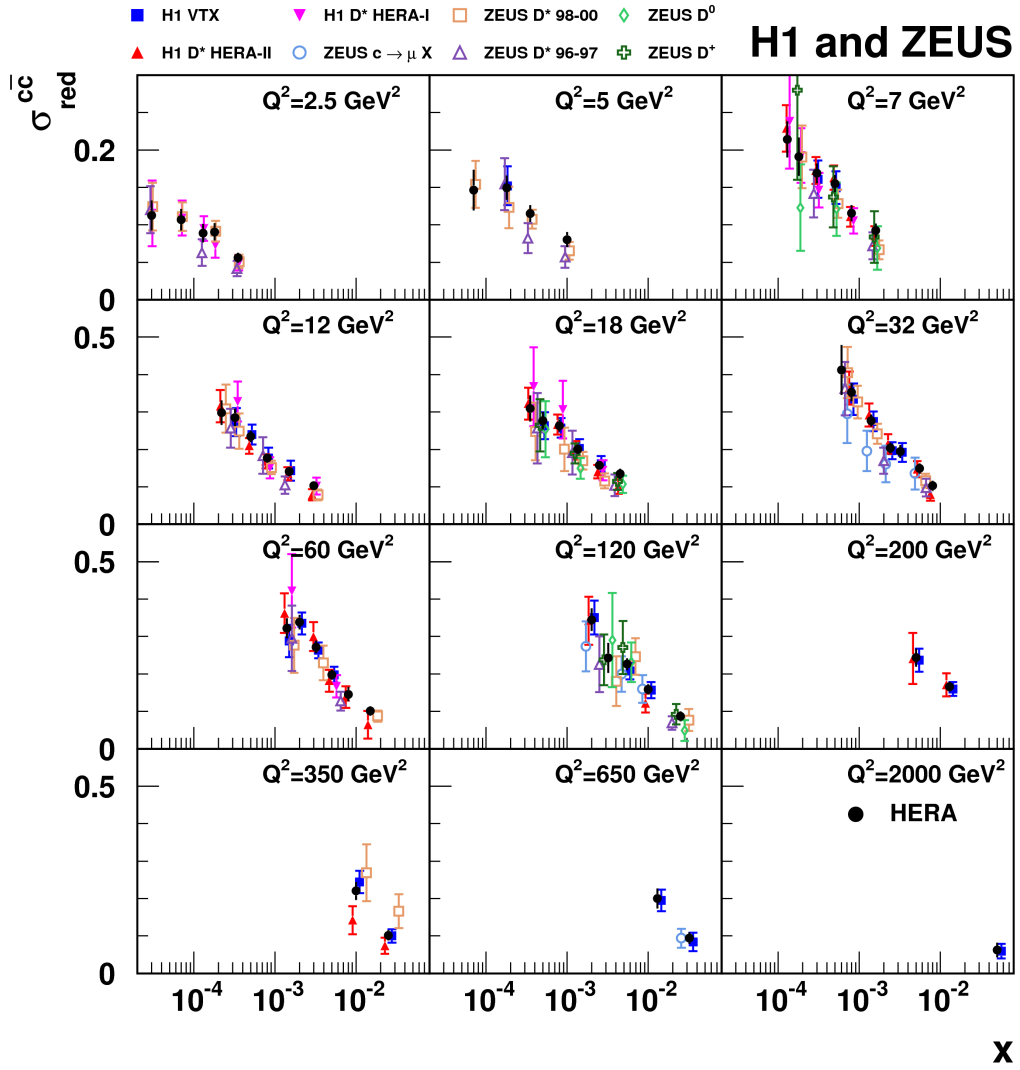


Figure 1.5: The combined reduced charm cross sections (full circles) and various measurements entering the combination (different style points). Taken from Ref. [70].

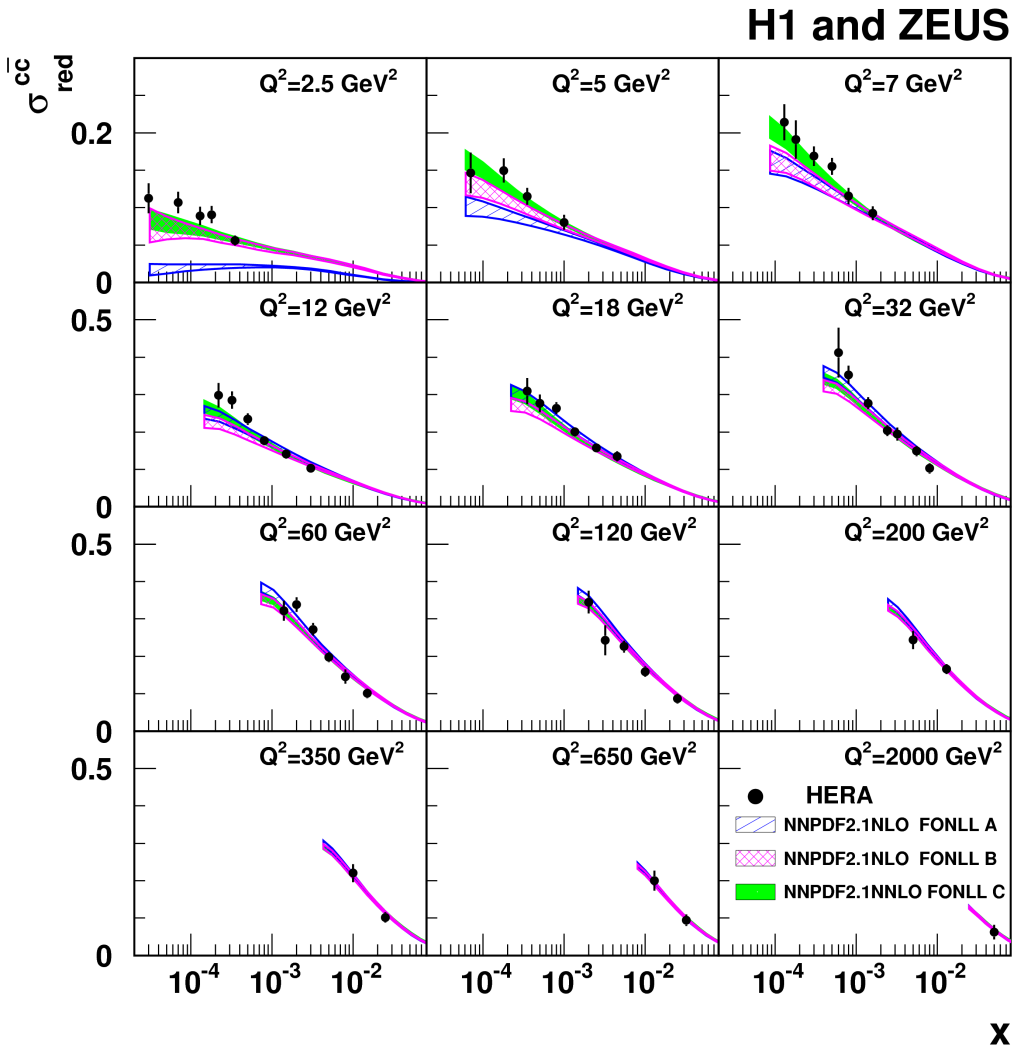


Figure 1.6: A comparison of the combined reduced charm cross sections (points) to NLO (hatched areas) and NNLO (filled area) QCD predictions from the NNPDF group. Taken from Ref. [70].

1.4.2 Beauty Production

In the inclusive lifetime analysis by the H1 Collaboration [55] mentioned in the previous subsection, the beauty quark production was measured. Figure 1.7 shows the differential cross section for production of jets containing a beauty quark, as a function of the jet transverse energy E_T^{jet}. Data are compared to NLO QCD predictions in the FFNS scheme. A good agreement is found within uncertainties.

The ZEUS Collaboration performed a measurement of the beauty quark production, where it was identified via semi-leptonic decays with a muon in the final state [57]. Muons were associated to jets which were selected in the following kinematic region: $E_T^{\text{jet}} > 5\,\text{GeV}$ and $-2.0 < \eta^{\text{jet}} < 2.5$. The phase space was defined by $Q^2 > 2\,\text{GeV}^2$ and $0.05 < y < 0.7$. Figure 1.8 shows the differential cross section of beauty jet production as a function of the photon virtuality Q^2. It is compared to the NLO QCD predictions in the FFNS scheme. As can be seen, the theory underestimates the data. The significance of the excess is around two standard deviations.

All measurements are done in a restricted region of the beauty quark kinematics (defined by the jet transverse energy and pseudorapidity cuts). Similarly to the charm case (see previous subsection), it is useful to perform an extrapolation to the full phase space, i.e. to determine cross section of beauty quark production with no cuts on its kinematics (see also a discussion in Section 7.4). The convenient observable is the *beauty contribution to the proton structure function*, $F_2^{b\bar{b}}$. It is defined in analogy to the inclusive structure functions Eq. (1.2) but a beauty quark-antiquark pair is required in the final state. Up to small contributions from the structure functions F_L and F_3, it coincides with the reduced beauty cross section $\sigma_{\text{red}}^{b\bar{b}}$ which was defined for the charm case in the previous subsection. Figure 1.9 shows a compilation of all available $F_2^{b\bar{b}}$ measurements at HERA with various QCD predictions. A good agreement between data and theory is observed; with given precision of the measurements, data are not yet able to discriminate between various predictions, contrary to the case of charm production.

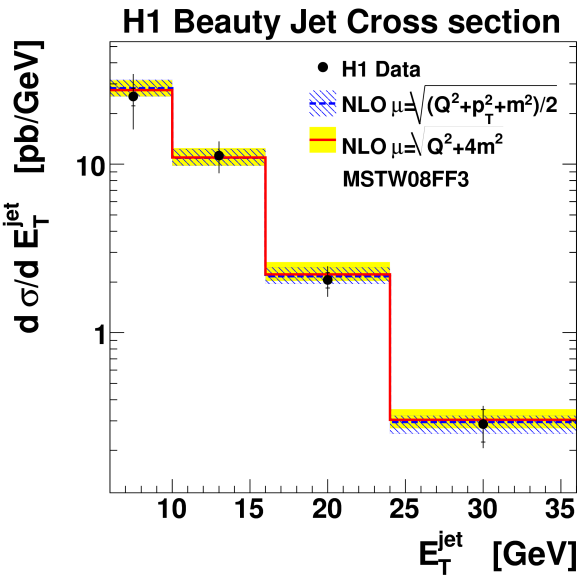


Figure 1.7: Differential cross section of beauty jet production as a function of the jet transverse energy E_T^{jet} in the inclusive lifetime analysis by the H1 Collaboration [55]. The points represent the measurement, the lines show the NLO QCD predictions in the FFNS scheme, the bands indicate uncertainties of the predictions.

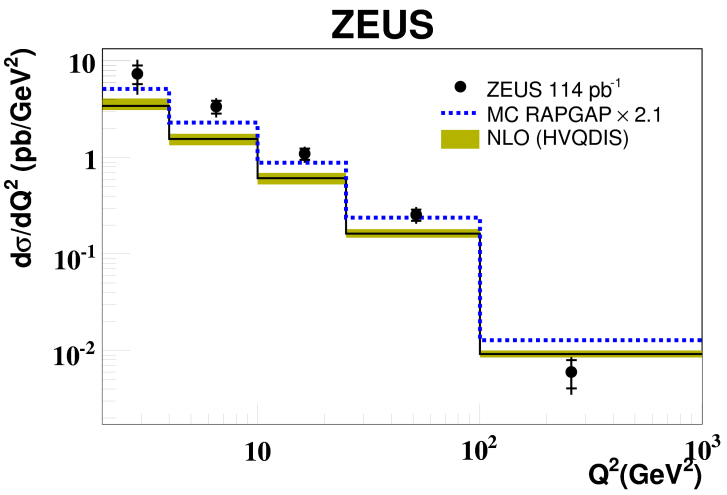


Figure 1.8: Differential cross section of beauty quark production as a function of the photon virtuality Q^2, from a measurement by the ZEUS Collaboration [57] where beauty quarks were identified via semi-leptonic decays with a muon in the final state. The points represent the data, the solid line is the NLO QCD prediction in the FFNS scheme, the band indicates its uncertainty. The dotted line represents LO QCD prediction from the RAPGAP MC, scaled by a constant factor.

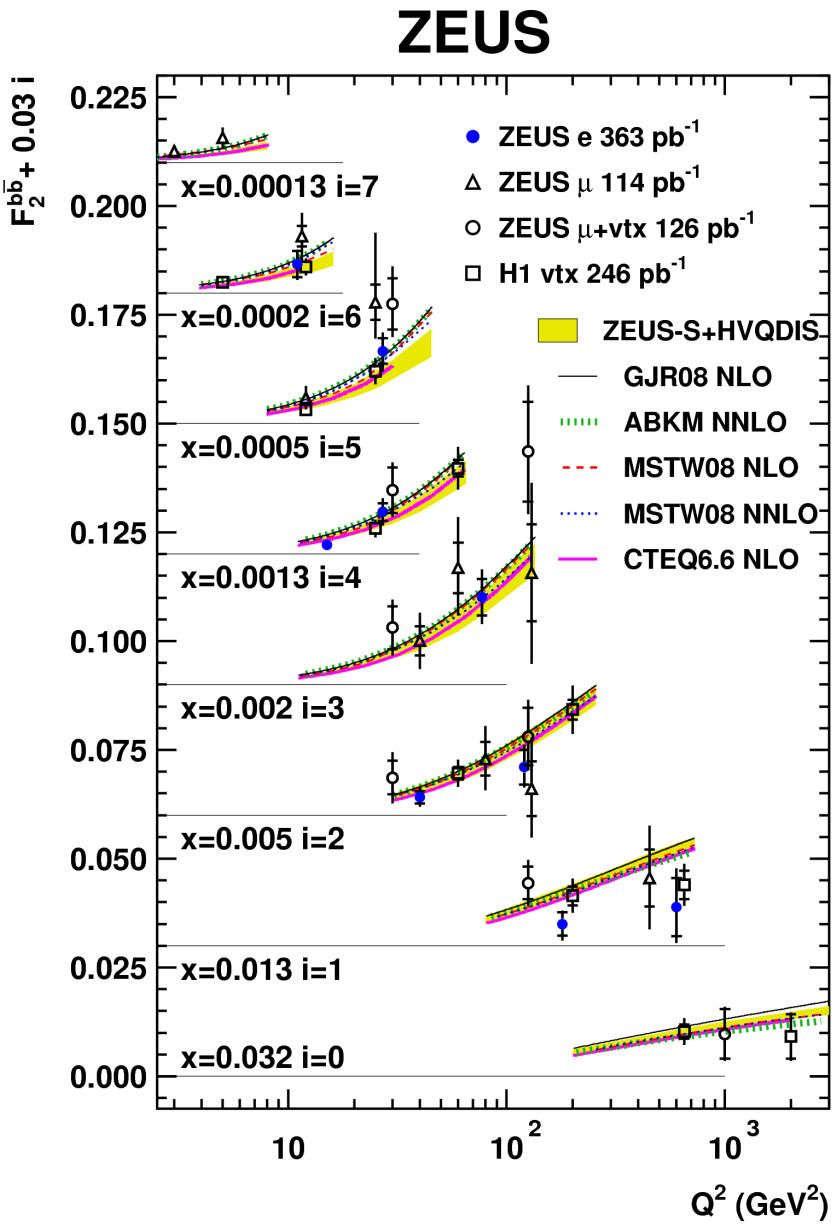


Figure 1.9: A summary of $F_2^{b\bar{b}}$ measurements [69] at HERA. The points represent various measurements from H1 and ZEUS. The lines show different theoretical predictions.

1.5 Monte Carlo Generators

At high energy colliders complex particle detectors, trigger chains and reconstruction algorithms are involved which have to deal with complicated final states emerging in collisions. A detailed understanding of the performance of these components is essential for any measurement. For example, in the analysis presented in this thesis one needs to determine the *detector acceptance*; it defines the probability to detect events of various types (e.g. charm quark production, beauty quark production) and hence allows extraction the corresponding cross sections. Also, the *events properties* need to be known, e.g. the transverse momentum, p_T, and pseudorapidity, η, spectra of produced charm or beauty quarks. This is required both for proper determination of the acceptance (since it usually depends strongly on these variables), as well as for prediction of shapes of certain distributions in order to test the consistency of the data with various models or to extract their relative contribution to the given data sample. These problems are solved with help of the *Monte Carlo* (MC) simulation. It consists of two major steps: the *event simulation* and the *detector simulation*. In the first step, virtual particle collision events are produced; every event contains full information about all particles that emerged in the collision and about their decay products. In the second step, the response of the detector, readout electronics and trigger to these particles is simulated. Information is then stored in the same format as real data, which allows to apply the same reconstruction algorithms.

Events are simulated with Monte Carlo *generators*. The basic steps of *ep*-event generation are schematically shown in Fig. 1.10. The implementation of the *hard process* depends on particular generator. In this thesis, the RAPGAP [71, 72] generator is used to simulate the signal. In the hard interaction it produces a pair of charm (or beauty)

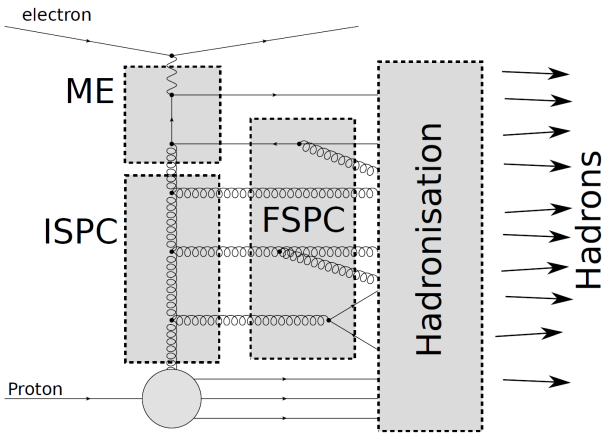


Figure 1.10: Simulation of *ep*-collison events. Different steps are shown by boxes. Indicated are: hard process (ME), initial and final parton showers (ISPC, FSPC) and hadronisation. The result of the generator is a set of outgoing hadrons.

quark-antiquark. The kinematics of the process (e.g. p_T and η distributions of the quarks) is calculated in the Leading Order perturbative QCD (the Feynman diagram is shown in Fig. 1.2). Since the photon directly enters the hard interaction, this process is also called *direct*. In RAPGAP, also *resolved* processes are available, where the photon fluctuates into a hadronic system before the hard interaction. In order to simulate higher order QCD processes, additional radiation in the initial or final state needs to be included. This is done with *parton showers*: quarks and gluons can radiate additional gluons; gluons can split into a quark-antiquark pair.

Due to confinement, the resulting partons are subject to *hadronisation* or *fragmentation*, the transition from coloured partons to colourless hadrons. This cannot be calculated in perturbative QCD and models have to be invoked. The most common is the *Lund* or *string* fragmentation model. In this model, a so-called string is attached to a pair of coloured objects (e.g. between the struck quark and the proton remnant). If these objects move apart, the energy of the string increases (due to confinement) and this ultimately results in creation of a new quark-antiquark pair, which might combine with original quarks (antiquarks) to form colourless hadrons. The process is repeated iteratively until only colourless hadrons are present. The longitudinal fragmentation is characterised by *fragmentation functions*. It gives the probability to transfer a certain fraction z of momentum of the original parton to the hadron which contains it. The Bowler [73] fragmentation function is given by:

$$f(z) \propto \frac{1}{z^{1+r_Q b m_Q^2}} z^{a_\alpha} \left(\frac{1-z}{z}\right)^{a_\beta} \exp\left(-\frac{b m_\perp^2}{z}\right), \tag{1.3}$$

where m_Q is the mass of the quark, $m_\perp$ is the transverse mass of the hadron, $m_\perp^2 = m^2 + p_T^2$ (m is the hadron mass, p_T is the mean transverse momentum of the quark); r_Q, b, a_α, a_β are model parameters taken from experiment. The Peterson [74] fragmentation function is defined in the following way:

$$f(z) \propto \frac{1}{z\left(1 - \frac{1}{z} - \frac{\epsilon}{1-z}\right)^2}, \tag{1.4}$$

where ϵ is the only function parameter; it is taken from experiment. In the last step, decays of particles are simulated. The result of this is passed to the detector simulation which is typically done with the GEANT package [75, 76].

Chapter 2

Experimental Setup

The measurement presented in this thesis is based on the data collected with the ZEUS detector which was located at the HERA collider. This chapter introduces these instruments. After a description of most important features of HERA, the relevant components of ZEUS for this analysis are discussed.

2.1 HERA collider

HERA[1] is the world only electron[2]-proton collider [8, 77]. Its task was to accelerate, store and collide beams of electrons and protons. It was located in Hamburg, Germany, at the research centre DESY. HERA was built from 1984 to 1991 and operated from 1992 to 2007. There was a shutdown in 2000-2002 when an upgrade of HERA was accomplished [78]. Electrons and protons circulated in separate rings. Each ring consisted of four 90° arcs with a radius of 797 m and of four straight sections of 360 m length, as shown in Fig. 2.1. The tunnel where the rings were placed has a total length of 6.3 km and resides 15–30 m below the ground.

The energy of the protons was 920 GeV (820 GeV before the upgrade mentioned above) while the electron energy was 27.5 GeV. This resulted in a centre-of-mass energy of $\sqrt{s} = 318\,\text{GeV}$. The maximum instantaneous luminosity that was reached is $\mathcal{L} = 5 \times 10^{31}\,\text{cm}^{-2}\,\text{s}^{-1}$ ($\mathcal{L} = 1.5 \times 10^{31}\,\text{cm}^{-2}\,\text{s}^{-1}$ before the upgrade).

A strong magnetic field is needed to provide the bending power for keeping the protons on the circular orbit in the arcs. At HERA this was supplied by 422 dipole superconducting magnets operated at a temperature of around 4 K yielding a magnetic field of more than 5 T. Due to the lower electron energy, a much weaker bending magnetic field is required for the electron ring. Hence, conventional non-superconducting magnets were used;

[1] **H**adron-**E**lektron **R**ing**a**nlage.

[2] By electron we mean the lepton which constitutes the beam. It can be either electron or positron depending on the running period, see Table 5.1.

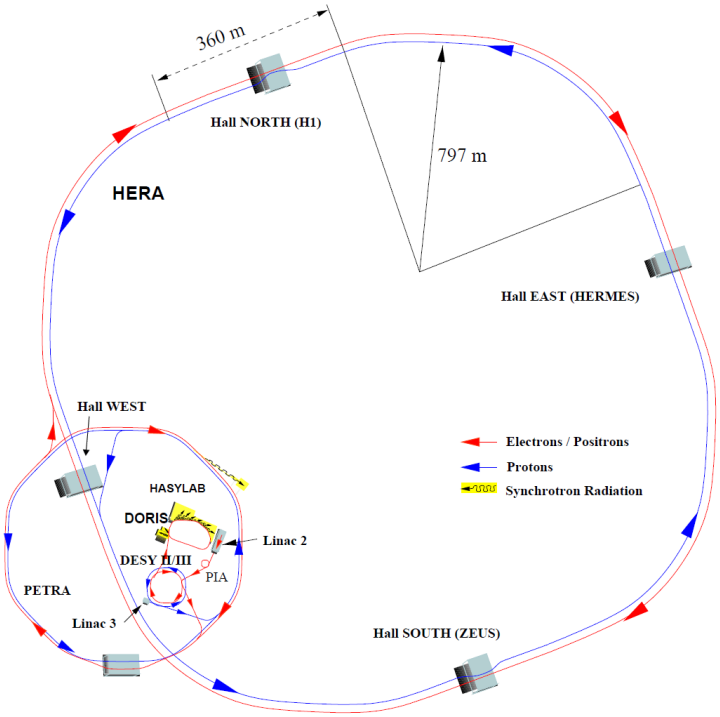


Figure 2.1: The HERA collider and its injection chain. The red and blue lines show the trajectories of electrons and protons, respectively. Locations of four collision points, where the experiments were installed, are indicated.

456 dipoles providing a magnetic field of 0.17 T were installed. Due to the low electron mass, bremsstrahlung energy losses are large for the electrons circulating in HERA and constitute around 100 MeV per turn. RF accelerating cavities, installed in the straight tunnel segments, compensated this effect.

A chain of pre-accelerators injected electrons and protons into HERA as indicated in Fig. 2.1. Protons started their path as negative hydrogen ions (H^-). They were accelerated in LINAC III to the energy of 50 MeV and were directed onto a foil in order to strip off the electrons. Before injection to HERA, they were sequentially injected to the DESY III storage ring and to PETRA II, where they were accelerated to 7.5 GeV and 40 GeV, respectively. Electrons were preaccelerated by LINAC II to the energy of 450 MeV and were afterwards injected to DESY II which boosted them to 7 GeV. Then they were directed to PETRA II (12 GeV), and finally to HERA.

Protons (electrons) moved in HERA in groups of around 10^{11} (3×10^{10}) particles (so-called *bunches*) with the distance of around 28.8 m between them, which corresponds to 96 ns time spacing. A maximum of 210 bunches was filled. Some bunches were left empty, which allowed studies of non-*ep* backgrounds.

Four experiments were performed at HERA: H1, ZEUS, HERMES and HERA-B. The

location of the experiments is shown in Fig. 2.1. H1 and ZEUS were collider experiments, i.e. they studied collisions of electrons and protons. The main goals were to investigate the interior of the proton to an unprecedented precision and to search for new phenomena. HERMES and HERA-B were fixed target experiments. At HERMES, interactions of polarised electrons with hydrogen, deuterium or helium targets were studied, the main goal being to understand the spin structure of the nucleon. At HERA-B, the protons provided by HERA were scattered off various targets, such as aluminum or carbon. The original goals were searches for CP violation in the B-meson systems, however, the experiment shifted its focus to the study of B-production cross sections and to the production of charmonium states.

2.2 Overview of the ZEUS Detector

The ZEUS experiment was one of the two collider experiments performed at the HERA accelerator. Among its main goals were studies of the structure of the proton, searches for a substructure of quarks or leptons, searches for leptoquarks, etc. The purpose of ZEUS was to detect and to measure the properties of particles that were emerging in the electron-proton collisions delivered by HERA. It was a hermetic 4π detector, meaning that particles moving in almost any direction (except those flying very close to the initial electron or proton direction) could be potentially detected. It had overall dimensions of $12 \times 11 \times 20\,\mathrm{m}^3$ and weighted around 3600 t. ZEUS is described in detail in Ref. [79]. In this section, a brief overview of the detector is given, while the following sections discuss various subsystems in more detail.

A right-handed cartesian coordinate system is used at ZEUS. The origin is in the nominal ep-interaction point. The X-axis points to the centre of the HERA machine, the Y-axis upwards while the Z-coordinate is along the incident proton direction. An overall view of the detector in the YZ-plane, i.e. along the beam is given in Fig. 2.2. The XY cross-section (perpendicular to the beam) is shown in Fig. 2.3. The ZEUS detector consisted of the tracking system, located closest to the beam, the calorimeter which surrounded it, and the muon system, placed outermost. The purpose of the tracking system was to detect charged particles and to measure their position and momentum. Its main components were the Central Tracking Detector (CTD) and the Microvertex Detector (MVD)[3]. The tracking system was surrounded by a thin solenoid. It provided a homogeneous magnetic field along the Z-direction which bends charged particles, thus allowing a measurement of their momenum by the tracking system. Outside the solenoid, the Calorimeter (CAL) was located. Its task was to detect particle showers and to measure their energies. The calorimeter consisted of an electromagnetic (EMC) and a hadronic (HAC) part which measured electromagnetic and hadronic showers, respectively. It was subdivided to the forward, barrel and rear sections (FCAL, BCAL, RCAL). Outside CAL,

[3]The MVD was installed for the HERA II period (2003-2007) whose data area analysed in this thesis

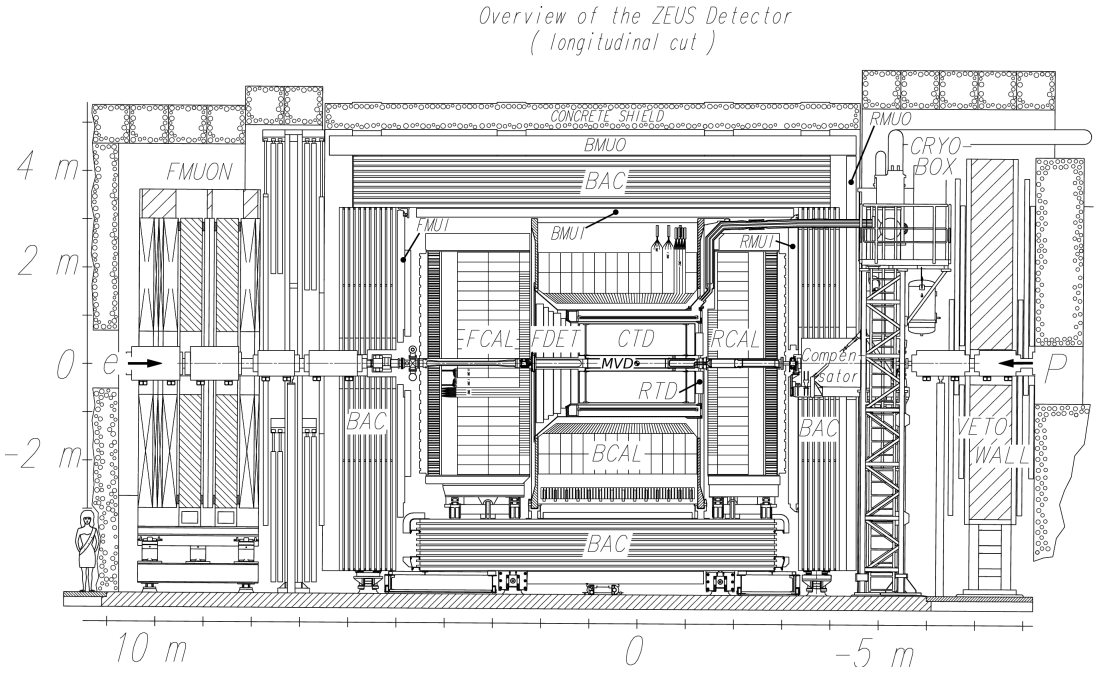


Figure 2.2: The YZ-view of the ZEUS detector [79].

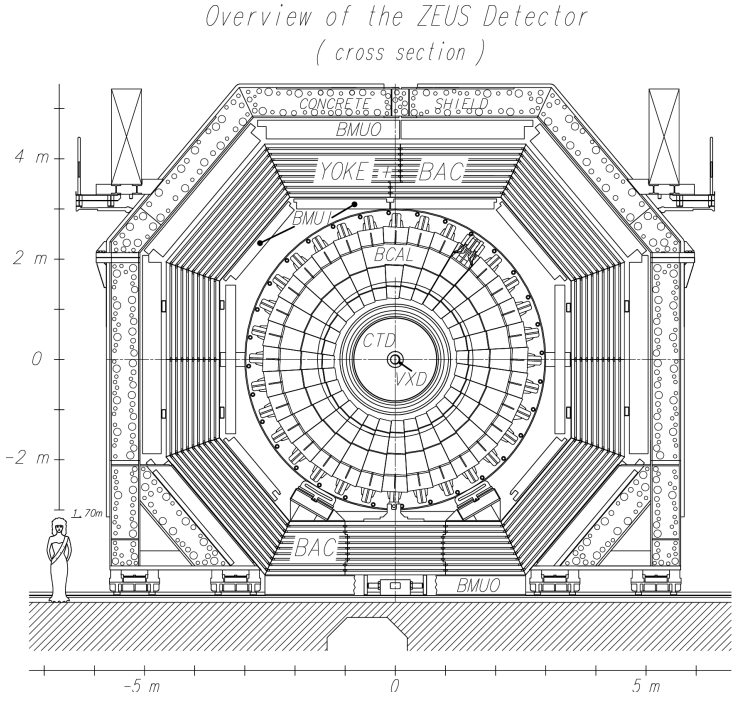


Figure 2.3: The XY-view of the ZEUS detector [79].

the muon detection system was installed; other particles than muons are in most cases stopped before it. There were additional tracking detectors in the forward regions, the Forward Tracking Detector (FDET) and the Straw Tube Tracker (STT)[4]. The luminosity was monitored by a dedicated system positioned outside the detector at around $Z \sim -100\,\mathrm{m}$.

As discussed in Section 2.1, the distance between two bunches of electrons or protons corresponds to 96 ns. This leads to a crossing rate of 10 MHz, at the ZEUS nominal interaction point. However it is not possible to store the information from all detector components at such a rate. Furthermore, only a very small fraction of bunch crossings leads to a genuine *ep*-scattering (few Hz for DIS events), while the background process of interactions of the beam protons with the residual gas in the beam pipe caused the dominating rate (10–100 kHz). In order to select only good *ep*-events, a three-level trigger system was implemented. For every bunch crossing, it made a decision whether an event should be stored for offline processing; events were accepted for storage at a rate of a few Hz.

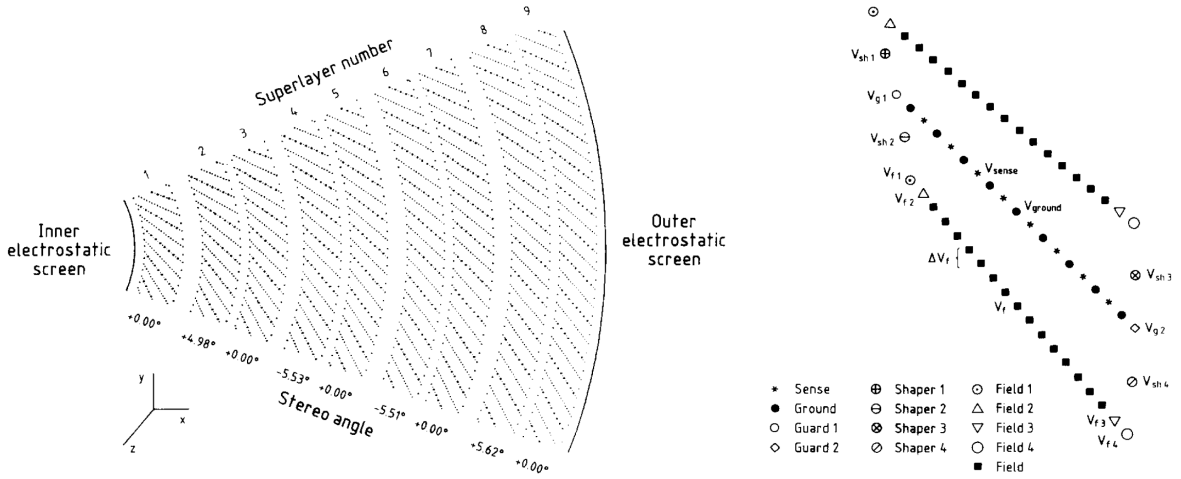
2.3 Central Tracking Detector

The Central Tracking Detector (CTD) [80–83] is the main tracking system in ZEUS. It detects charged particles and measures their trajectories (*tracks*). In conjunction with the magnetic field of 1.43 T, provided by the thin superconducting solenoid which is outside the CTD, this allows the measurement of the charged particle transverse momentum. Also the particle energy losses dE/dx are measured, which makes possible their identification. At the trigger level, the CTD reconstructs the position of the *ep*-interaction point (*primary vertex*) in order to discriminate against beam-gas interactions.

The CTD is a cylindrical drift chamber with the axis of symmetry coinciding with the beam line. Inside the chamber, wires are placed, parallel to the beam direction or at a small angle to it. The interior of the CTD is filled with a gas mixture of argon, ethane and carbon dioxide (the relatvie proportions are $Ar/CO_2/C_2H_6$: 90/8/2). A charged particle moving through it ionises the gas; the released charges drift to the *sense wires* in a constant electric field, hence with a constant velocity. Close to the wire, electrons are accelerated due to a rapid increase of the field. This leads to an avalanche and therefore to a measurable electric pulse. The time of the drift is used to determine the distance from the ionisation point to the wire.

The inner radius of the active volume is 18.2 cm, while the outer radius is 79.4 cm. The endplates are placed at $z \approx \pm 100\,\mathrm{cm}$, leading to a length of the active volume of around 2 m. The layout of wires is shown in Fig. 2.4(a) for one octant. They are organised in nine *superlayers*. A basic building block of the superlayer is a *cell*. The layout of a cell

[4]The FDET was installed only for the HERA I period (1992-2000) while the STT was available for HERA II.



(a) (b)

Figure 2.4: (a) The layout of wires in the CTD for one of the octants [81]. Larger dots represent the sense wires. Positions of the wires are shown at the endplates. (b) The layout of a cell [81]. Filled circles denote sense wires. Other wire types are also shown.

is shown in Fig. 2.4(b). It consists of eight sense wires which collect drifting electrons, as well as of the field, ground and shaping wires, which create the necessary drift field and the surface field (leading to an avalanche). There are 4608 sense wires in the CTD. Five of the superlayers have wires parallel to the Z-axis, while four of them are placed with a small angle ($\pm 5^\circ$) to it. This allows a measurement of the z-coordinate.

The CTD is capable of measuring tracks with polar angles of $15^\circ < \theta < 164^\circ$ while covering the full azimuthal angle. The single hit resolution is around 280 µm in the XY-plane, while the precision of the Z-coordinate measurement is around 1.4 mm.

2.4 Microvertex Detector

The Microvertex Detector (MVD) [84] is the innermost component of the ZEUS detector. It was installed during the HERA shutdown in 2001. By providing precise measurements of charged particles trajectories close to the primary vertex, it dramatically improved the track impact parameter resolution and as a consequence the capabilities to reconstruct the primary ep-interaction vertex as well as displaced secondary vertices due to decays of long-lived particles.

In a nutshell, the sensitive elements of the MVD are reversely biased pn-junctions. A charged particle penetrating the depleted region of the junction creates free electrons and holes; they move in the electric field and lead to an electric pulse which is detected by the readout electronics.

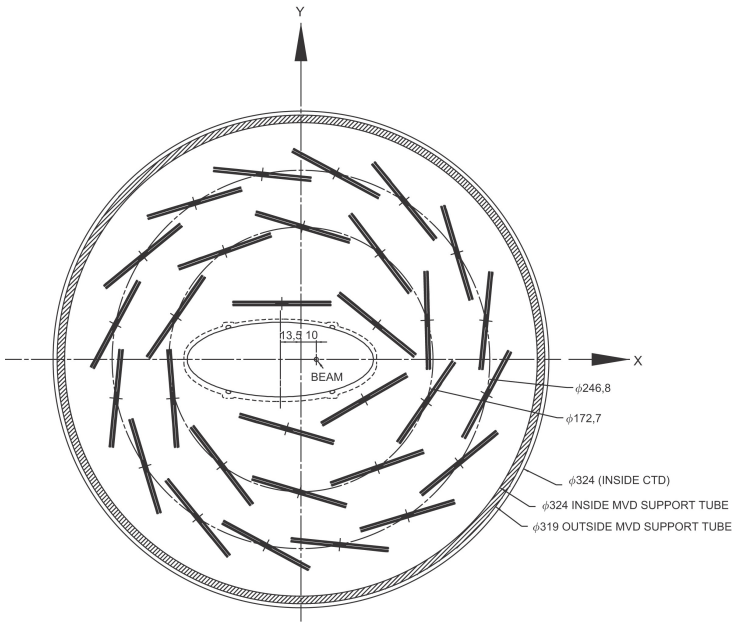


Figure 2.5: The layout of the ZEUS Microvertex Detector in the XY-plane [84]. Three concentric layers around the beam pipe (BMVD) are visible which reduce to two for azimuthal angles around $\phi = \pi$.

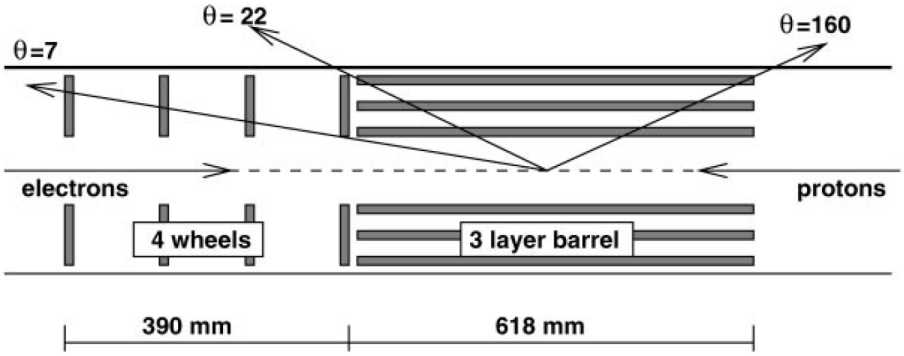


Figure 2.6: The layout of the ZEUS Microvertex Detector in the YZ-plane [85]. Three layers in the barrel part (BMVD) and four wheels in the forward direction (FMVD) are visible.

The MVD consists of *modules* of silicon microstrip *sensors* organised in three concentric cylindric layers in the barrel part (BMVD) and in four circular disks (*wheels*) perpendicular to the beam in the forward direction (FMVD), as it is shown in Figures 2.5 and 2.6. The sensor is a 300 µm thick n-type silicon substrate with p^+-implants in form of parallel strips on one side. The transverse cross-section of the sensor is given in Fig. 2.7. The strips are 14 µm wide and are placed at a pitch of 20 µm. One out of six strips is

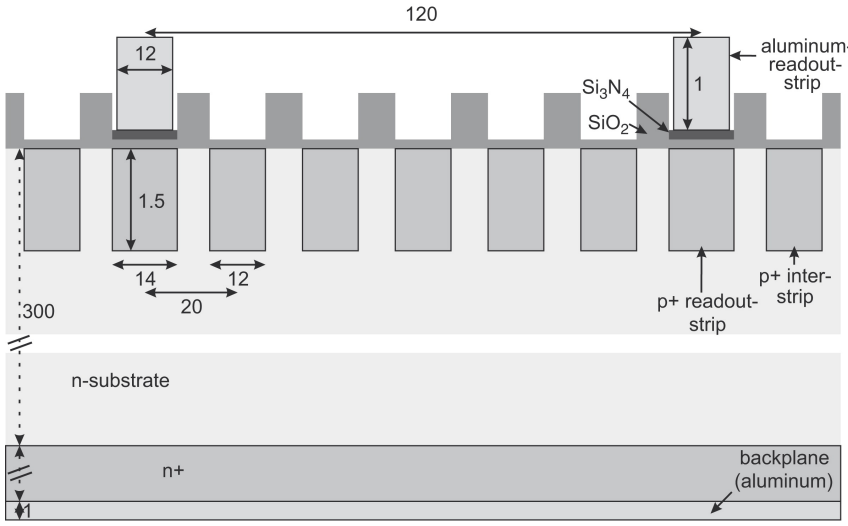


Figure 2.7: Cross section of an MVD sensor [84]. The long side of strips points perpendicularly to the plane of the drawing.

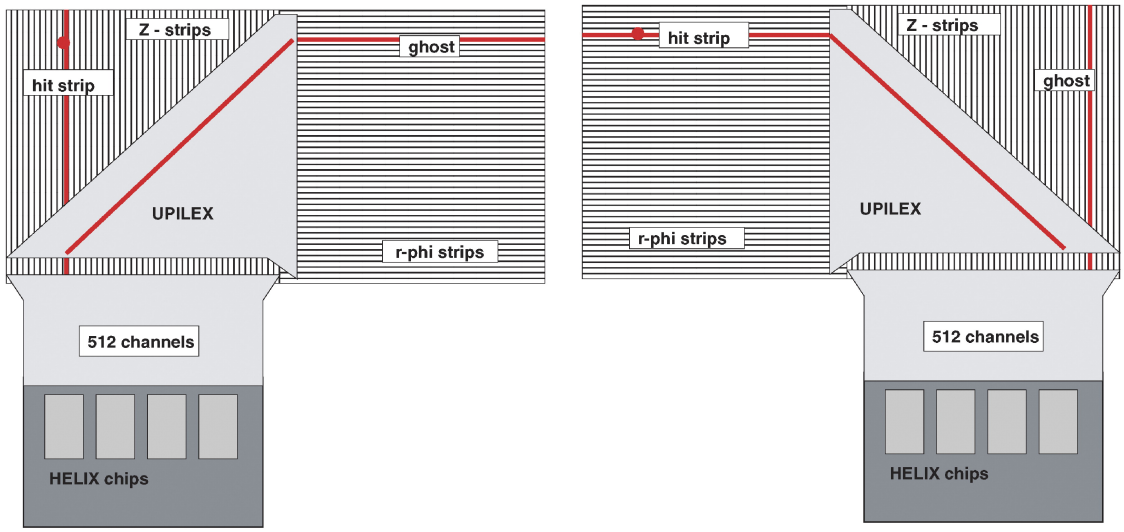
connected to the readout electronics (*readout strips*). On the backside, there is an n^+-implant and an aluminum metallisation which provide an ohmic contact for biasing the *pn*-junctions and depleting the sensor bulk.

Slightly different sensor shapes and number of strips are used in the barrel and in the forward parts. The BMVD sensors have a squared shape with a breadth of 64.2 mm. The length of the p^+-strips is 62.2 mm; there are 512 readout strips on the sensor. Two sensors are glued together side-by-side to form a *half-module*; both are connected to the same readout electronics chip, in such a way that one readout channel is connected to two strips, one on each sensor, leading to 512 readout channels per half-module. Strips are oriented perpendicularly in two sensors, as it is shown in Fig. 2.8. Two half modules are glued on top of each other to form a *full module* with 1024 readout channels. Modules are then mounted onto support structures in the barrel region. Hence, each layer provides a measurement of both $r\phi$- and z-coordinates. The FMVD sensors are wedge shaped and have 480 readout strips. Apart from different shapes and number of readout channels, the FMVD sensors are identical to BMVD sensors. Two FMVD sensors are mounted back-to-back on the circular support structures, hence a measurement of $r\phi$ is provided.

The polar angle coverage of the barrel part is $20° < \theta < 160°$ (for tracks passing all three barrel layers) while the forward part extends it down to 7°. After alignment, the single hit resolution is 23.6 µm. The combined transeverse momentum resolution of the CTD and MVD, for tracks that pass all nine CTD superlayers is:

$$\sigma(p_T)/p_T = 0.0029 \cdot p_T \oplus 0.0081 \oplus 0.0012/p_T,$$

where $\oplus$ denotes the quadratic sum and transverse momentum p_T is measured in GeV.



(a) (b)

Figure 2.8: Sensor arrangement in half modules. An example of a half module (a) and of its mirror-image (b) are shown. Two such half-modules are glued on top of each other to form modules.

The first term represents the intrinsic resolution, the second and third terms are due to multiple scattering. The transverse distance of closest approach (DCA) of tracks to the nominal vertex in the XY-plane was measured to have on average a resolution of $46 \oplus 122/p_T$.

2.5 Uranium Calorimeter

The purpose of the ZEUS Calorimeter (CAL) [86–89] was to measure energies of charged and neutral particles emerging in the *ep*-interaction. Most importantly, it was used for the reconstruction of hadronic jets and for the identification of the scattered electron.

The calorimeter consisted of alternating layers of depleted uranium (absorber) and plastic scintillator (active material). A high energy particle entering the absorber creates secondary particles due to nuclear (electromagnetic) interactions and thus starts a hadronic (electromagnetic) shower which is developed further in the following absorber layers. Charged particles from the shower excite molecules of the scintillator which then emit light upon de-excitation. The amount of light is proportional to the energy of particles in the shower, hence enabling a measurement of the shower energy. A crucial feature of the CAL is *compensation*: the response to electromagnetic and hadronic showers is equal. This was achieved by choosing appropriate absorber material (uranium) and scintillator material and tuning their thickness. The compensation leads to better hadronic

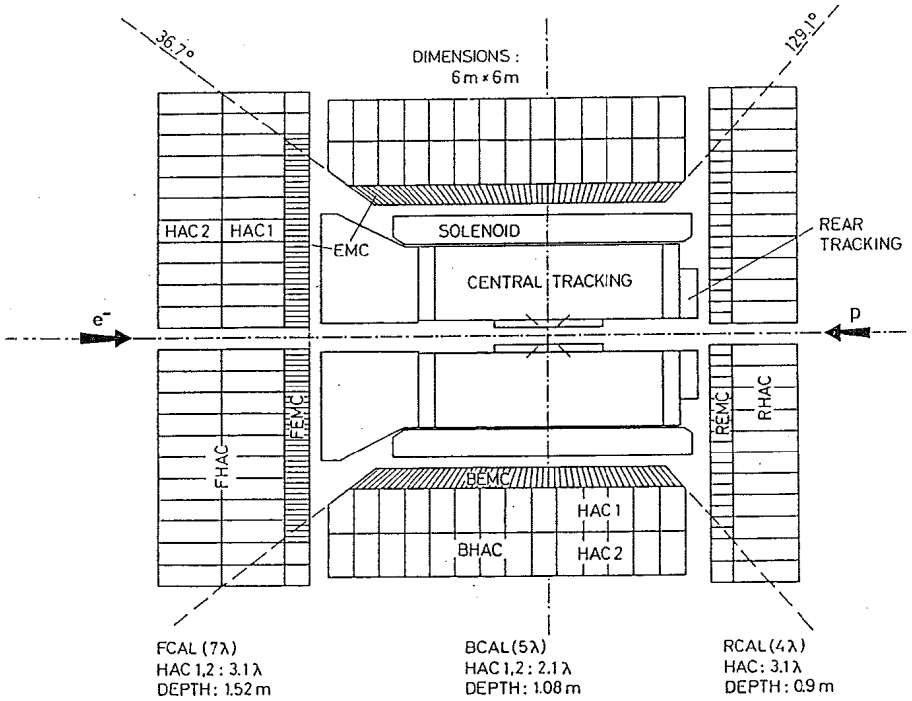


Figure 2.9: The layout of the ZEUS Calorimeter [86]. The radiation length of each part is indicated in the bottom.

energy resolution and linearity of the response to the hadronic component.

The calorimeter covered almost the full solid angle. Its overall layout is shown in Fig. 2.9. It was divided into forward (FCAL), barrel (BCAL) and rear (RCAL) sections, with the polar angle coverage of $2.2° < \theta < 39.9°$, $36.7° < \theta < 129.2°$ and $128.1° < \theta < 176.5°$, respectively, and the full azimuthal angle coverage. Each of these parts consisted of the electromagnetic section (EMC) and of one or two hadronic sections (HAC), as it is indicated in Fig. 2.9. The depth (thickness) of EMC corresponds to one interaction length (1λ) while that of HAC varies from $\sim 6\lambda$ in the forward direction to $\sim 3\lambda$ in the rear part. The thicker HAC section in the forward direction was chosen due to higher hadronic activity and energies compared to the rear hemisphere.

The calorimeter consisted of so-called *modules*. There were 24 rectangular modules in FCAL and RCAL and 32 wedge-shaped modules in BCAL. Modules were further segmented into EMC and HAC *towers*. A tower was a stack of alternating uranium and scintillator layers. The thickness of a single uranium layer was 3.3 mm, while the thickness of the scintillator layer was 2.6 mm. The transverse size of a layer in HAC towers for RCAL and FCAL was $20 \times 20\,\mathrm{cm}^2$. The EMC tower size was $5 \times 20\,\mathrm{cm}^2$ ($10 \times 20\,\mathrm{cm}^2$) in FCAL (RCAL) in the regions where it is not covered by BCAL; otherwise, the size was equal to that of the HAC tower. All scintillators in a tower were readout as a single, independent

unit (*cell*). This was done by connecting wavelength shifters to all scintillators in a cell. The wavelength shifters were then directed to photomultipliers.

The resolution of CAL was $35\,\%/\sqrt{E}$ and $18\,\%/\sqrt{E}$ for hadronic and electromagnetic showers, respectively (E is measured in GeV) as measured under test beam conditions. The timing of the pulses could be determined to a nanosecond precision which allowed the usage of CAL for triggering purposes.

2.6 Luminosity Monitor

The luminosity delivered by the HERA collider was determined by measurement of the rate of the Bethe-Heitler process, $ep \to e\gamma p$. The cross section of this process is large and well-known theoretically [90] (to an accuracy of 0.5 %). Two detectors were installed outside ZEUS, the Photon Calorimeter (PCAL) [91] and the Spectrometer (SPEC) [92]. Both PCAL and SPEC detected photons which were produced in the Bethe-Heitler reaction. The luminosity is given by:

$$L = \frac{N}{\sigma A},$$

where N is the number of observed events, σ is the cross-section of the process and A is the acceptance of the corresponding detector (PCAL or SPEC).

The PCAL was a lead-scintillator sampling calorimeter, placed at $z = -107\,\mathrm{m}$. It was shielded against the synchrotron radiation with carbon filters. Since these filters worsened the energy resolution, two Cherenkov detectors [93] were installed to measure a signal related to energy deposits in the filters. The photon energy resolution was $25\,\%/\sqrt{E}$ in HERA II data taking period, relevant for this thesis.

The SPEC detected the photons from the Bethe-Heitler via the pair conversion process, $\gamma \to e^+e^-$. The conversion took place in the exit window from the beam pipe. The SPEC used a dipole magnet with a field of 0.5 T to separate the electron and the positron. After the separation, their energy was measured by two tungsen/scintillator sandwich calorimeters.

2.7 Trigger and Data Acquisition System

At ZEUS, a three level trigger system [94–96] was employed in order to select events at a rate of a few Hz which could be written to tape. The architecture of the system is shown in Fig. 2.10. Every detector component had its own first level trigger (FLT). Digital or analogue data, read out for each bunch crossing (every 96 ns) were stored into a corresponding pipeline for around 5 µs. During this time, the FLT took a decision whether this event is accepted or not. Obviously this should be done in parallel for each event in a pipeline. Each FLT component performed calculations for a fixed amount of time

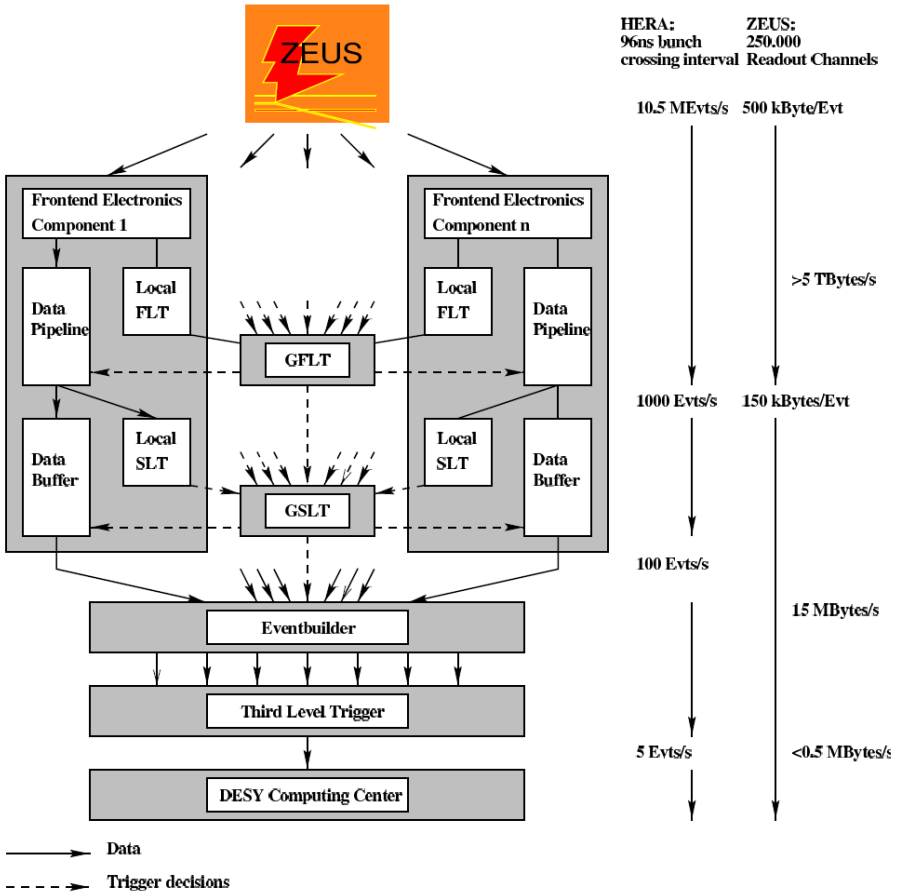


Figure 2.10: The ZEUS trigger and the data acquisition system.

(maximum 2.5 µs which corresponds to 26 bunch crossings) and afterwards sent the results to the Global First Level Trigger (GFLT) which took a decision within 1.9 µs (20 bunch crossings). If the event was accepted, it was passed to the Second Level Trigger (SLT). The FLT reduced the rate to below than 1 kHz. Similarly to the FLT, every component had its own SLT. They performed a first rough event reconstruction and sent information to the Global Second Level Trigger (GSLT) which took the decision in few ms. The resulting rate was below 100 Hz. In case of a positive decision, the event was built and sent to the third level trigger (TLT) which was a CPU farm. It performed more precise event reconstruction and made the final decision. If accepted, the event was stored to tape, at a rate of around 5 Hz.

Chapter 3

Event Reconstruction

After an event was accepted by the ZEUS trigger system, it was stored to tape for offline analysis. This chapter describes algorithms which were employed to reconstruct high level energy objects such as tracks, vertices or jets, based on the information provided by individual components of the ZEUS detector. First, the SINISTRA algorithm is discussed, which was used in the measurement presented in this thesis to identify the scattered electron. Track and vertex reconstruction algorithms are discussed afterwards. The concept of energy flow objects, which combine tracking and calorimeter information in order to obtain an optimal measurement of the hadronic final state, is then introduced. Finally, the jet reconstruction algorithm and three methods for the determination of the main event kinematic variables are presented.

3.1 Scattered Electron Identification

A main signature of the Neutral Current Deep Inelastic Scattering (NC DIS) events, which are studied in this thesis, is the presence of the scattered electron in the detector active volume. In contrast, photoproduction leads to a small electron scattering angle, and the electron escapes through the beam pipe not leaving any signals in the detector. In the Charged Current Deep Inelastic Scattering, the electron transforms to a neutrino which leaves the detector undetected. Thus, in order to select NC DIS events it is necessary to identify the scattered electron. Furthermore, the electron information (scattering angle and energy) may be used for determination of the event kinematic variables (see Section 3.6). In this thesis, the SINISTRA [97] algorithm was employed to identify the scattered electron.

Electrons are detected primarily by the high resolution uranium calorimeter (see Section 2.5) where they leave energy deposits. The main task is to separate deposits of the electron from those of single hadrons or jets. This is done based on showering properties of the electromagnetic particles and hadrons. Electromagnetic showers in RCAL result in

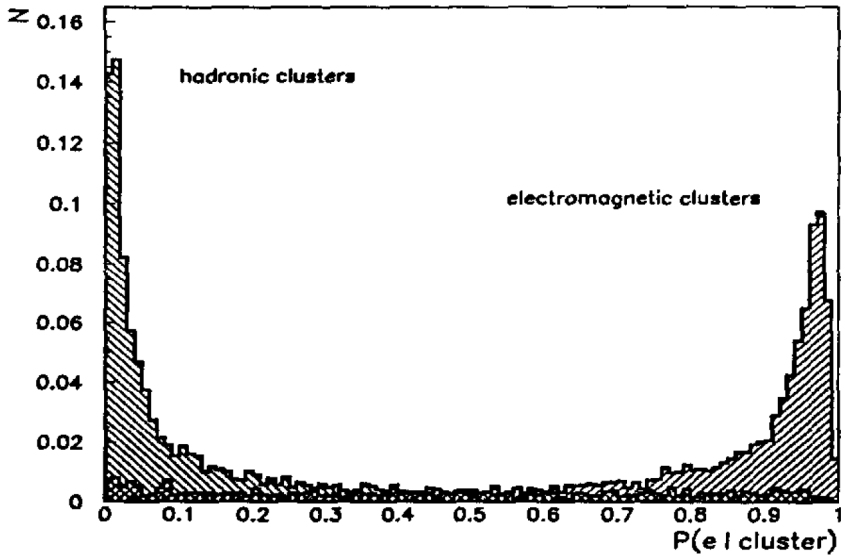


Figure 3.1: Example of the SINISTRA probability distribution for electromagnetic and hadronic clusters [97].

deposits in about three cells, with a tail to more cells due to interactions in the inactive material before the CAL. Hadronic showers are typically transversely much broader and longitudinally deeper. For example, a 10 GeV pion typically leads to deposits in about seven EMC cells and six HAC cells.

In clusters due to a single electromagnetic shower most of the energy is deposited in the EMC part; transversely the shower is typically contained within a window of 3×3 towers (i.e. $60 \times 60\,\mathrm{cm}^2$) centred on the tower with highest energy deposition. This leads to a maximum of 54 CAL variables which can describe the shower (9 HAC cells and 18 EMC cells, every cell is read out by two photomultipliers). In order to improve the separation power, a neural network (NN) approach is used in SINISTRA. All 54 variables are used by the NN, and additionaly the incidence angle of the particle causing the deposit, calculated from the coordinates of the cluster and from the primary vertex position. For every cluster the NN returns the probability p for it to be an electromagnetic cluster ($p \sim 0$ for hadronic clusters, $p \sim 1$ for electromagnetic ones). The NN was trained on a Monte Carlo sample of low-Q^2 NC DIS events. The distribution of the probability for hadronic and electromagnetic clusters in this sample is shown in Fig. 3.1. For electromagnetic clusters, the distribution is indeed peaking towards $p = 1$. The electron energy is obtained by summing up the energy deposits in the cells associated with the cluster. It is corrected for the energy losses in the inactive material between the interaction point and the calorimeter.

3.2 Track reconstruction

A charged particle traversing the ZEUS tracking detector produces signals in the Microvertex Detector (MVD) and in the Central Tracking Detector (CTD). In the forward region, it may also produce signals in the Straw Tube Tracker (STT). Track reconstruction is the process of determination of the particle trajectory based on the information from these subdetectors. Since CTD and MVD are placed in the magnetic field provided by the solenoid, the curvature of the trajectory may be employed for the particle transverse momentum measurement.

The first step is the *pattern recognition*. This is the process of assigning particular hits from the subdetectors to a track. It is performed by the VCTRACK package [98]. First, track *seeds* are identified. They are groups of hits in the outermost layers of the CTD or in the STT for the forward region. An STT seed is required to have at least eight hits, while for the CTD at least three hits (in the axial layers) are required. Afterwards, hits from the inner layers are added to the seeds to form a track candidate. The nominal primary vertex position is used to guide the hit search.

In the second step, the *track fit* is performed, based on the hit assignment from the previous step. This is done by the RTFIT program [99]. It uses a Kalman Filter technique [100] that takes into account multiple scattering effects and the residual inhomogenities of the magnetic field.

The track fit employs a *helix* parametrisation specified by five parameters (see Fig. 3.2 for an illustration):

- ϕ_H – the angle tangent to the helix in the XY-plane;
- Q/R – the ratio of the charge Q to the radius in the XY-plane, R;
- QD_H – the charge times the distance of closest approach in the XY-plane with respect to the reference point (0,0);
- Z_H – the Z-coordinate of the track at $(X, Y) = (0, 0)$;
- $\cot\theta$ – where θ is the polar angle of the track.

The first three parameters specify a circle in the XY-plane (Fig. 3.2, left) while the other two parameters define the location and the pitch of the track trajectory in Z (Fig. 3.2, right). The transverse momentum p_T of the track is determined from its curvature R according to $p_T = 0.3\,Q\,B\,R$ [101], where B is the magnetic field in Tesla ($B = 1.43\,\mathrm{T}$ in the ZEUS tracking system), R is measured meters, Q is given in units of the positron charge, and p_T in GeV/c. The momentum is calculated as $p = p_T/\sin\theta$.

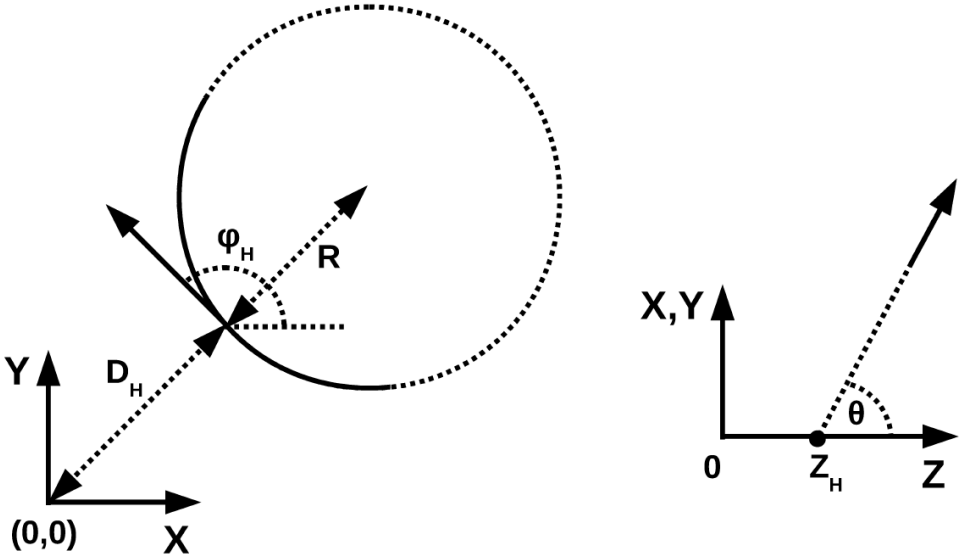


Figure 3.2: Illustration of the helix parametrisation in the XY-plane (left) and in the YZ-plane (right).

3.3 Vertex reconstruction

The position of the *ep*-interaction point (*primary vertex*) is determined in two steps. In the first step, the *vertex finding* is performed, i.e. tracks are associated to the vertex. This is done by the VCTRACK package [98]. Pairs of tracks that are loosely consistent with a common vertex close the Z-axis are searched for. Such pairs are combined and a χ^2-fit is performed. A combination of tracks that gives the best overall χ^2 in the vertex fit is considered as a set of primary tracks. In the second step, the vertex coordinates and track parameters at the vertex are refined by usage of the *Deterministic Annealing Filter* (DAF) [102–105]. The main idea behind is that instead of a fixed χ^2 cut, a weight is applied to each track according to the Fermi function:

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

where T is called a temperature and χ^2_{cut} is a parameter. At zero temperature, the weight function $w(\chi^2, T = 0)$ is a step function with $w = 1$ for $\chi^2 < \chi^2_{\text{cut}}$ and $w = 0$ for $\chi^2 > \chi^2_{\text{cut}}$, which corresponds to a fixed χ^2 cut. An iterative fitting procedure is performed. In the first iteration, the values of χ^2 for each track are taken from the VCTRACK vertex fit. A large value for the temperature T is assigned and the weight for each track is calculated. The fit is then performed taking into account the weights, and the values of track χ^2 are updated. Afterwards, the temperature T is reduced and the procedure is repeated, with new values of χ^2 as obtained in the previous iteration. This is done until T reaches a

small value. The described procedure leads to a better vertex determination, compared to the fixed χ^2-cut approach, in particular it is well suited for fits with outliers. The knowledge of the primary vertex position in the XY-plane can be further improved by averaging it over a set of events. The result is referred to as the *beam spot*. At ZEUS it was calculated for sets of ~ 2000 good events [106]. The beam spot size for the HERA II period is around $90\,\mu m$ ($25\,\mu m$) in the X- (Y-) direction. Additionally, the *beam tilt* – the dependence of the beam spot position on the Z-coordinate of the primary vertex, was determined.

The DAF algorithm is also used to reconstruct the secondary vertices due to decays of heavy flavour hadrons, in order to tag charm and beauty events. The track selection for these vertices is described in detail in Section 5.4.

3.4 Energy Flow Objects

The ZEUS calorimeter provides energy measurements of charged and neutral hadrons as well as of electrons and photons. The tracking system supplies the momentum measurement of only charged particles. The energy resolution of the CAL improves with energy as $\propto 1/\sqrt{E}$, while the transverse momentum resolution of the tracking system degrades as $\propto p_T$ (see Sections 2.4 and 2.5). Hence, at high energies the calorimeter information is more precise than that from tracking, and vice versa at low energies. Often, a particle is measured by both systems. Therefore, by combining the tracking and calorimeter information it is possible to improve the measurement of the hadronic system. This concept is known as *Energy Flow Objects* (EFO) [107, 108]. The output of the algorithm are EFOs with the properties determined either from CAL or from the tracking system, or combining them. In this section, the procedure of EFO reconstruction is briefly described.

As a first step, the clustering of the energies in the calorimeter cells is performed. This procedure is required since a particle leads typically to energy deposits in several neighbouring cells (see also discussion in Section 3.1). First, an *island clustering* is performed, separately for EMC, HAC1 and HAC2 sections (only for one HAC section in RCAL) in each part of the calorimeter (FCAL, BCAL, RCAL). If a cell has a larger energy deposit than all of its four direct neighbours (corner cells are not considered), it is called the *local maxima*. Otherwise, the cell is linked to the highest energy neighbouring cell. The procedure results in two dimensional clusters of cells, the *cell islands*. Figure 3.3(a) illustrates this algorithm. Then, a *cone clustering* is performed which is a linking of cell islands from EMC, HAC1 and HAC2 (only HAC1 for RCAL) into *cone islands*. The association is based on the distance between cell islands in the $\theta - \phi$ space. It is converted to a probability according to a distribution obtained from a single pion Monte Carlo. The positions of the resulting cone islands are determined using the logarithmically weighted centre-of-gravity of the shower. Figure 3.3 (b) illustrates the cone clustering algorithm.

After the clustering, tracks are matched to the cone islands. For this purpose tracks

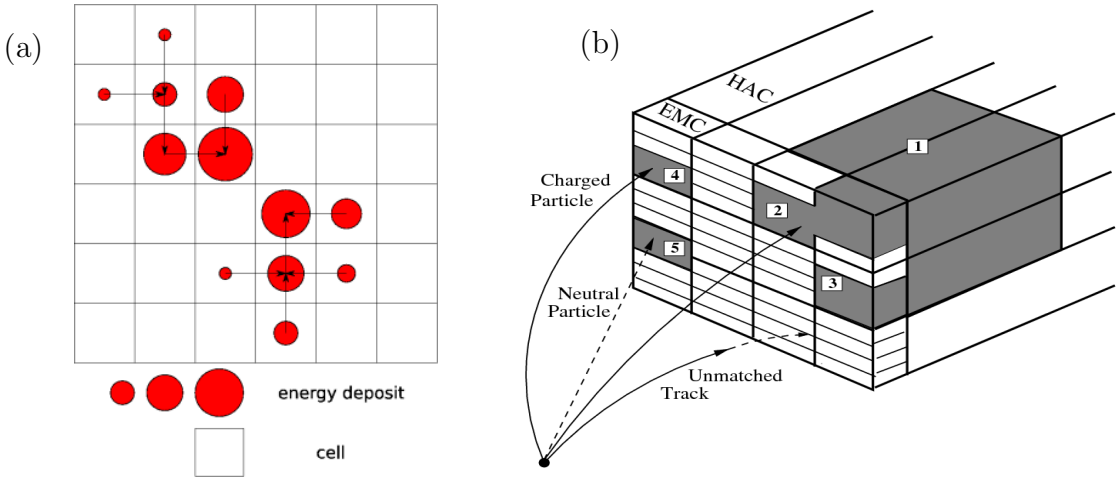


Figure 3.3: (a) An illustration of the cell clustering algorithm [107]. (b) The cone clustering algorithm [108]. Initially there are four EMC cell islands and one HAC cell island. EMC islands 2 and 3 are merged with HAC island 1 to form a single cone island. Matching of tracks to cone islands is also illustrated.

are extrapolated to the inner surface of CAL. Only tracks with at least 4 superlayers in the CTD and with transverse momentum of $0.1 < p_T < 20\,\text{GeV}$ are considered. If a track passed more than 7 superlayers, the maximum p_T is increased to 25 GeV. A track must be associated to the primary vertex. The track is considered as matched to an island, if the distance of closest approach between the track and the island position is less than 20 cm or it is less than the maximum island radius in the plane perpendicular to the ray joining the primary vertex and the island.

The energy of EFO is determined in the following way:

- if a track that passes the above criteria is not associated to any calorimeter island, the tracking information is used; the pion mass is assumed for the energy calculation;
- if a calorimeter object has no track counterpart, the calorimeter information (energy) is used; the object is assumed to be massless;
- if a calorimeter object is associated with more than 3 tracks, the calorimeter information is used; the object is assumed to be massless;
- if there was a match of a single track to a single cluster (1-to-1 match), the track information will be used if two conditions are fulfilled: $E_{\text{cal}}/p < 1.0 + 1.2\,\sigma(E_{\text{cal}}/p)$ and $\sigma(p)/p < \sigma(E_{\text{cal}})/E_{\text{cal}}$, where E_{cal} is the energy of the calorimeter cluster, p is the track momentum (both are measured in GeV), and σ is the uncertainty on the quantity given in brackets. These requirements ensure that the calorimeter cluster is

due to the track alone, and that the tracking information is more precise. Otherwise, the CAL information is used.

More complicated cases as well as various corrections to the EFO energy are explained in detail in Ref. [108]. In this thesis, the EFOs were used to reconstruct jets (see the next section) as well as to calculate the event kinematic quantities (see Section 3.6) and the $E - p_Z$ variable.

3.5 Jet Reconstruction

A *jet* is a collimated flow of particles which arises due to the hadronisation of a quark or gluon emerging in a hard QCD process. From early days of collider physics, jets served as a tool to study properties of the underlying hard process. A *jet algorithm* is a procedure to reconstruct jets from individual particles observed in a detector (or those on the Monte Carlo true level). It defines the rules of assignment of particles to jets and the way the four-momentum of the jet is constructed from the four-momenta of particles constituting it (*recombination*).

In this analysis, the longitudinally invariant inclusive k_T-clustering algorithm [109–111] with the E recombination scheme was used. The procedure consists of the following steps:

1. A list of all original particles, and a list of jets (empty at the beginning) are created;
2. For each entry i from the particle list, the parameter d_i is defined as:
$$d_i = E_{T,i}^2,$$
where $E_{T,i}$ is the transverse energy of the particle;
3. For each pair of particles i and j, the distance $d_{i,j}$ is defined as:
$$d_{i,j} = \min(E_{T,i}^2, E_{T,j}^2)[(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2]/R^2,$$
where ϕ and θ are azimuthal and polar angles of the particles, respectively, and R is a parameter (which is set to $R = 1$ here);
4. The smallest number among all d_i and $d_{i,j}$ is determined and labelled as $d_{\min}$;
5. If $d_{\min}$ is a $d_{i,j}$, the particles i and j are merged into a new (pseudo-) particle k and are removed from the list. The four-momentum of the new particle is given by the sum of the original particles four-momenta (E recombination scheme):
$$p_k = p_i + p_j. \tag{3.1}$$
This procedure of combining two particles leads to *massive* jets.

6. If $d_{\min}$ is a d_i, it is removed from the list of particles and is added to the list of jets;
7. The procedure is repeated, starting from the step 2, until the list of particles is empty.

The algorithm described is *infrared safe*, i.e. it is insensitive to long distance effects [110]. In this thesis, the Energy Flow Objects (see Section 3.4) served as the input to the jet algorithm; the jet selection criteria are described in Section 5.3. The same algorithm was also employed for theoretical calculations which were used for comparisons with the results of this measurement (see Section 7.1 for details).

3.6 Kinematic variables

This section describes three methods for the reconstruction of main kinematic variables characterising the deep inelastic scattering process, namely the photon virtuality Q^2, the inelasticity y and Bjorken x.

3.6.1 Electron Method

This method is based purely on the scattered electron information. The kinematic variables are given by:

$$Q^2_{\text{el}} = 2E_e E'_e(1+\cos\theta),$$
$$y_{\text{el}} = 1 - \frac{E'_e}{2E_e}(1-\cos\theta),$$
$$x_{\text{el}} = \frac{Q^2_{\text{el}}}{s\, y_{\text{el}}},$$

where θ is the polar angle of the scattered electron (in the ZEUS coordinate system), E_e (E'_e) is the energy of the initial (final) state electron. The method has a poor resolution for x and y at low values of the inelasticity y.

3.6.2 Jacquet-Blondel Method

This method relies on the measurement of the hadronic system only [112]. It is assumed that hadrons that are not detected escape through the beam pipe in the proton direction and their transverse momentum can be neglected. The formulae for the kinematic variables are as follows:

$$y_{\text{JB}} = \frac{1}{2\,E_e}\sum_i \left(E_i - p_{Z,i}\right),$$
$$Q^2_{\text{JB}} = \frac{(\sum_i p_{X,i})^2 + (\sum_i p_{Y,i})^2}{1-y_{JB}},$$

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

where the sums run over all reconstructed energy flow objects (EFO); $p_{X,i}$, $p_{Y,i}$, $p_{Z,i}$ and E_i are the components of the momentum and the energy of the i^{th} EFO. The Q^2 resolution is somewhat poorer than for the electron method, while the x and y measurements are better, especially at low values of y.

3.6.3 Double-Angle Method

The Double-Angle Method [113] for reconstruction of the kinematic variables is intermediate between the electron method and the Jacquet-Blondel method described above, in the sense that it employs both electron and hadronic information. It uses the electron scattering angle and an angle γ_{h}, characterising the longitudinal and transverse momentum flow of the hadronic system [114] which is defined in the following way:

$$\cos\gamma_{\mathrm{h}} = \frac{\left(\sum_i p_{X,i}\right)^2 + \left(\sum_i p_{Y,i}\right)^2 - \left(\sum_i (E_i - p_{Z,i})\right)^2}{\left(\sum_i p_{X,i}\right)^2 + \left(\sum_i p_{Y,i}\right)^2 + \left(\sum_i (E_i - p_{Z,i})\right)^2}.$$

The kinematic variables are then given by:

$$Q^2_{\mathrm{da}} = \frac{4E_e^2 \sin\gamma_{\mathrm{h}}(1+\cos\theta)}{\sin\gamma_{\mathrm{h}} + \sin\theta - \sin(\gamma_{\mathrm{h}}+\theta)},$$

$$x_{\mathrm{da}} = \frac{E_e}{E_p}\,\frac{\sin\gamma_{\mathrm{h}} + \sin\theta + \sin(\gamma_{\mathrm{h}}+\theta)}{\sin\gamma_{\mathrm{h}} + \sin\theta - \sin(\gamma_{\mathrm{h}}+\theta)},$$

$$y_{\mathrm{da}} = \frac{\sin\theta(1-\cos\gamma_{\mathrm{h}})}{\sin\gamma_{\mathrm{h}} + \sin\theta - \sin(\gamma_{\mathrm{h}}+\theta)},$$

where E_p is the energy of the incoming proton and other variables were already defined. The advantage of this method is that it does not depend on the electromagnetic scale (as the electron method) and on the hadronic scale (as the Jacquet-Blondel method). Overall it gives the best performance compared to the Jacquet-Blondel and the electron methods.

In this thesis, all three approaches are employed. The double-angle method is used for Q^2 and x determination in the whole phase space of the measurement. For the calculation of the inelasticity y, the electron and the Jacquet-Blondel methods are used at high and at low y, respectively.

Chapter 4

Tracking Inefficiency due to Hadronic Interactions

The charm and beauty measurement performed in this thesis relies on a proper detector simulation, since it is used to determine the acceptance as well as shapes of signal and background distributions. In particular, one of the most important requirements is a precise modelling of the tracking efficiency. There are several effects that lead to a tracking inefficiency, such as the detector efficiency, the pattern recognition efficiency or the disappearance of tracks due to hadronic interactions with detector matter nuclei. The latter effect is a dominant contribution to the tracking inefficiency at ZEUS. A large uncertainty was associated to it. In order to reduce the resulting cross section systematic uncertainty, a dedicated study was performed within this thesis. It is described in detail in the following.

4.1 Method

The goal of the study is to check how well the probability of hadronic interactions[1] is reproduced by the Monte Carlo (MC) simulation. This can be done by measuring a quantity which is sensitive to this effect and comparing it between data and MC. Any observed differences would point to an imperfect simulation of interactions. A correction can then be derived and applied to the MC.

Interactions in the transition region between the MVD and the CTD, i.e. in the outer wall of the MVD or the inner wall of the CTD, as well in the outer layers of the MVD, were considered. If no interaction occurs, a particle will leave hits in both MVD and CTD, hence a full-length track can be reconstructed. However, if a particle interacts hadronically with a nucleus in this detector region, it will disappear as a result of the corresponding reaction or will be significantly deflected from its original flight

[1] Shortly *interactions* in what follows.

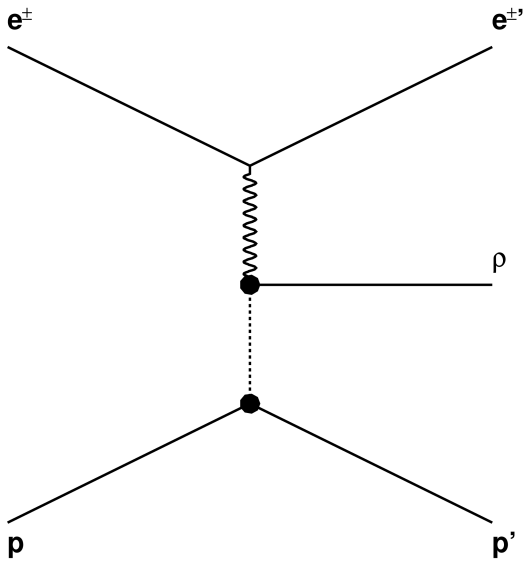
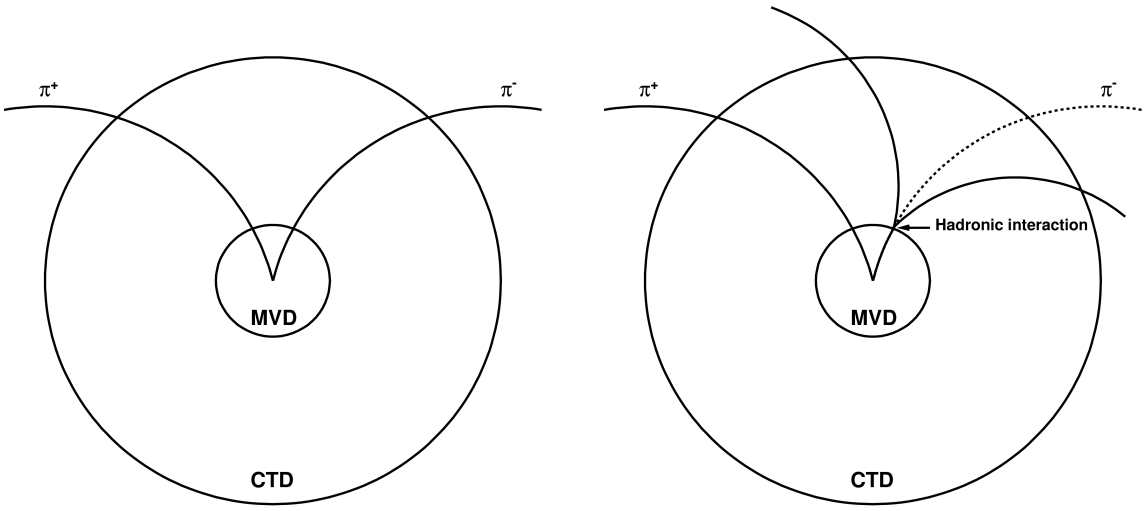


Figure 4.1: Feynman diagram of exclusive ρ production at HERA

direction, and hence generally will not be detected by the CTD. However it still can be reconstructed in the MVD. We will denote such tracks, based only on MVD hits without a CTD counterpart, as MVDSA tracks (MVD Stand-Alone tracks). The relative fraction R of such tracks with respect to all reconstructed tracks is directly sensitive to the probability of hadronic interaction in the MVD-CTD transition region. It is also related to the MVD and CTD acceptance which is assumed to be well simulated by the MC, thus it cancels out in the data-to-MC ratio if the assumption is correct. The other effect that can cause disappearance of tracks are in-flight particle decays. However, it is expected to be small compared to hadronic interactions.

Exclusive (diffractive) production of the $\rho(770)$-meson with a subsequent decay to two charged pions ($ep \to e'\rho\, p'$, $\rho \to \pi^+\pi^-$), was chosen as the type of events to be studied. In this process (see Fig. 4.1 for a Feynman diagram) the proton remains intact or dissociates into a low mass excited state and escapes the detector through the beam-pipe, hence only two charged pions, and possibly a scattered electron are detected. This is a very clean topology with practically no background (see Section 4.3), which minimises possible systematic effects.

If no interactions occur, there will be exactly two tracks (apart from the electron) coming from the primary vertex (*primary* tracks) which are well measured in the CTD (*long* tracks), as illustrated in the Figure 4.2(a). However, if one of the pions interacts in the transition region, it can be reconstructed as an MVDSA track but not as a full-length track. Hence one expects one long primary track and one MVDSA track. Additionally, there might be other tracks due to particles that arise in the interaction, see Fig. 4.2(b).



(a) (b)

Figure 4.2: An illustration of the method. (a) Class I event topology. Both pions did not interact and hence are well measured both in the MVD and the CTD (b) Class II event topology. One of the pions did not interact and is detected both in the MVD and the CTD. The other pion interacted in the MVD-CTD transition region and is reconstructed as an MVDSA track with no CTD counterpart (the original trajectory of the pion is shown as a dashed line). Additional non-primary tracks emerging from the interaction point are also shown.

However they cannot be attributed to the primary vertex, thus are easily distinguishable. For simplicity, we will call the first event type (with no interaction) as *class I*, while the second type (one of the pions interacts) will be denoted as *class II*. Cases when both pions interact are rare and were not considered.

4.2 Data and Monte Carlo Samples

The full HERA II data sample, corresponding to an integrated luminosity of around $354\,\mathrm{pb}^{-1}$ was used for these studies. The DIS regime was chosen, since it provides a clean trigger signature thanks to the presence of the scattered electron. This is in contrast to the photoproduction mode where the electron is not detected, hence triggering of exclusive ρ production events relies on the tracking information of the pions from the ρ decay.

The ZEUSVM [115] Monte Carlo generator was used to generate $ep \to e\rho^0(\to \pi^+\pi^-)p$ events in DIS. Exclusive production and decay of the ρ-meson are characterised by event kinematic variables, the photon virtuality Q^2, the hadronic system mass W and the four-

Period	Lepton Type	Integrated Luminosity, pb^{-1}	Number of MC events
2003-2004	e^+	30.61	740 289
2005	e^-	133.14	1 702 402
2006	e^-	52.69	567 459
2006-2007	e^+	137.30	2 224 481

Table 4.1: Details on data and ZEUSVM MC samples used in this chapter.

momentum exchange at the proton vertex t and by three helicity angles which are defined in the $\gamma^* p$-frame and in the s-channel helicity frame [116,117]. The Q^2 and W distributions in ZEUSVM are generated according to the Born cross-section for the process $\gamma^* p \to \rho^0 p$, where γ^* denotes a virtual photon [118]. The minimum cut on Q^2 was $1.5\,\mathrm{GeV}^2$. The t distribution is an exponential function with a slope set to $b = 5\,\mathrm{GeV}^{-2}$. Flat distributions of helicity angles are generated. ZEUSVM was interfaced to HERACLES [119] in order to take initial and final state QED radiation into account.

Table 4.1 shows the number of events that were used for each data taking period. The MC events are processed in the same way as the data, i.e. the same reconstruction and selection procedures are applied.

4.3 Selection

DIS events are characterised by a large scattering angle of the electron, which allows its detection in the main calorimeters. Hence, in order to select a DIS sample, a reconstructed electron candidate was required for each event. The SINISTRA algorithm (see Section 3.1) was used to search and identify electrons. The following criteria, which are standard for ZEUS, were applied:

- The presence of at least one SINISTRA electron candidate is required;
- Electron probability $> 90\,\%$ – for each candidate, the algorithm assigns a certain probability for it to be caused by an electromagnetic cluster. Only the most probable candidate is considered. The cut on its probability ensures a good purity of electromagnetic showers identification;
- Energy of the scattered electron $E'_e > 10\,\mathrm{GeV}$ – suppresses fake electron candidates coming from e.g. pions decaying into photons and ensures good efficiency of the finder.

Further cleaning cuts were imposed:

- A reconstructed primary vertex with $|Z_{\mathrm{VTX}}| < 30\,\mathrm{cm}$ – ensures that tracks emerging in collisions are in the CTD and MVD acceptance regions and reduces non-ep background;

- $35 < E - p_Z < 70\,\text{GeV}$, where $E - p_Z = \sum_i (E - P_Z)_i$, the sum runs over all ZUFOs. For a fully reconstructed DIS event one expects $E - p_Z = 2\,E_e = 55\,\text{GeV}$ [2]. The given cut reduces photoproduction and non-ep background.

No other explicit cuts were applied to select DIS events. In particular, no specific trigger chains were required and no kinematic (Q^2, W, t) cuts were imposed, in order to maximise the statistics. Within the selected DIS sample, events with class I and class II topologies were searched for.

For class I events, two *long* tracks of opposite charge flagged as primary were required; definition of the long track is as follows:

- A track starts in the first CTD superlayer or in the MVD;
- A track reaches the third CTD superlayer;
- $\delta < 0.2\,\text{cm}$, where δ is the transverse impact parameter, that is the distance of closest approach between the track trajectory and the beam spot (see Section 3.3), in the plane transverse to the beam direction (XY-plane). This cut ensures that a track is consistent with coming from the primary vertex.

No other tracks apart from the electron were allowed in the event.

For the class II, only one primary long track is allowed, with the same criteria as for class I. Exactly one MVDSA track is required with the following properties:

- Number of hits (sum of $r\phi$ and rz hits) in the Barrel MVD (BMVD) ≥ 5 – ensures that the track is measured in at least three MVD layers[3];
- Transverse impact parameter $\delta < 0.2\,\text{cm}$.

This ensures a good measurement by the MVD and that the track is not a secondary one. Additionally, any number of tracks not flagged as primary is allowed for the class II, since they might come from the hadronic interaction. For both classes, the kinematic range for selected primary tracks (i.e. pion candidates) was:

- $p_T > 0.2\,\text{GeV}$ – low momenta tracks might not reach the CTD due to bending in the magnetic field;
- $0.44 < \theta < 2.7\,\text{rad}$ – selects tracks in the acceptance region of the CTD. At higher or lower values of the polar angle, a track cannot pass three CTD layers.

[2] Due to the energy-momentum conservation, the quantity $E - p_Z$ is equal in the initial and final states. Its calculation is straightforward in the former case: the proton moves along the Z axis, therefore it does not contribute to $E - p_Z$ (since $E_p \approx (p_Z)_p$, neglecting the mass), while the electron moves in the opposite direction and $E_e \approx -(p_Z)_e$. Hence, in the initial state $(E - p_Z)_{\text{in}} \approx 2E_e$.

[3] As discussed in Section 2.4, each MVD layer provides one $r\phi$- and one rz-measurement (hit).

In order to suppress the contamination from exclusive production of $\phi(1020) \to K^+K^-$ where kaons are misidentified as pions, events fulfilling $1.012 < M(\phi) < 1.028\,\text{GeV}$, where $M(\phi)$ is the mass of the two primary tracks (pion candidates) assuming kaon mass, are rejected.

Although the given criteria allow selection of a reasonably clean sample of exclusive ρ events, there is still background, e.g. events when other particles in addition to the two pion candidates are produced in the *ep*-collision. For example, the exclusive production of the $\omega(782)$-meson with a subsequent decay $\omega \to \pi^+\pi^-\pi^0$ will also lead to only two reconstructed tracks in the event, since the neutral pion escapes detection by the tracking system. Hence these events can be accepted with the selection described above. Another example is non-exclusive production of the ρ-meson, when other particles in addition to the ρ are produced.

In order to supress events with additional neutral particles, the so called *elasticity cut* is applied. It makes use of calorimeter information. In particular, if there are clusters of energy in the calorimeter which cannot be attributed to one of the two pions, such events are rejected. Matching of clusters to tracks is done by extrapolating a track to the calorimeter in order to obtain the impact position and the flight direction at the inner calorimeter surface. A straight track trajectory is assumed in CAL, since it is placed outside the solenoid and the magnetic field is much lower than in the tracking detectors (ranging typically from 0 to 0.3 T, but with local maxima of 0.8 T) [120]. Clusters closer to the track trajectory than 30 cm are considered as belonging to the track, i.e. *matched*. Clusters which are more distant are considered as not coming from the track or *unmatched*. Only clusters with energy deposit exceeding the calorimeter threshold of 500 MeV are considered.

For the class I, calorimeter clusters are allowed around both long tracks. If a cluster is found which is not matched to one of the two tracks, an event is rejected. For the class II, the situation is more complicated since hadronic interactions may lead to production of other tracks which would leave energy deposit in the calorimeter. In order to take this into account, calorimeter clusters which satisfy $\Delta R = \sqrt{(\eta^{\text{trk}} - \eta^{\text{clus}})^2 + (\phi^{\text{trk}} - \phi^{\text{clus}})^2} < 0.5$ are allowed, where η^{trk} (η^{cluster}) is the pseudorapidity of the MVDSA track (cluster) and ϕ^{trk} (ϕ^{clus}) is the azimuthal angle of the MVDSA track (cluster). Angles of the clusters are determined using the reconstructed primary vertex. For the long (non-interacted) track, clusters are matched in the same way as for class I.

After this selection, ρ-meson candidates are formed. For the class I, the ρ-candidate is reconstructed from two long primary tracks, while for the class II it is based on the long primary track and the MVDSA track.

Figure 4.3(a) shows the mass distribution for data for class I events after the event selection described above. A clear peak which can be attributed to exclusive production of the $\rho(770)$-meson is observed. The asymmetric form is due to the Breit-Wigner shape of the ρ-meson. Events in the range of $M(\pi^+\pi^-) \in [0.6,\ 1.1]\,\text{GeV}$ were selected for further analysis. In this region, background from exclusive production of $\phi(1020) \to K^+K^-$ where

kaons are misidentified as pions or $\omega(782) \to \pi^+\pi^-\pi^0$ is expected [117] to be negligible. However, it contributes to the region of smaller $M(\pi^+\pi^-)$ which leads to an excess on the left of the peak. The non-resonant background is expected to be at a level of 1 % [117].

Figure 4.3(b) shows the mass distribution for data for class II events. A clear peak can be observed with similar features as for class I, despite the fact that one of the pions is measured by the MVD only, hence the ρ mass resolution is inferior compared to class I. The statistics is worse than for class I and this is a limiting factor of the method.

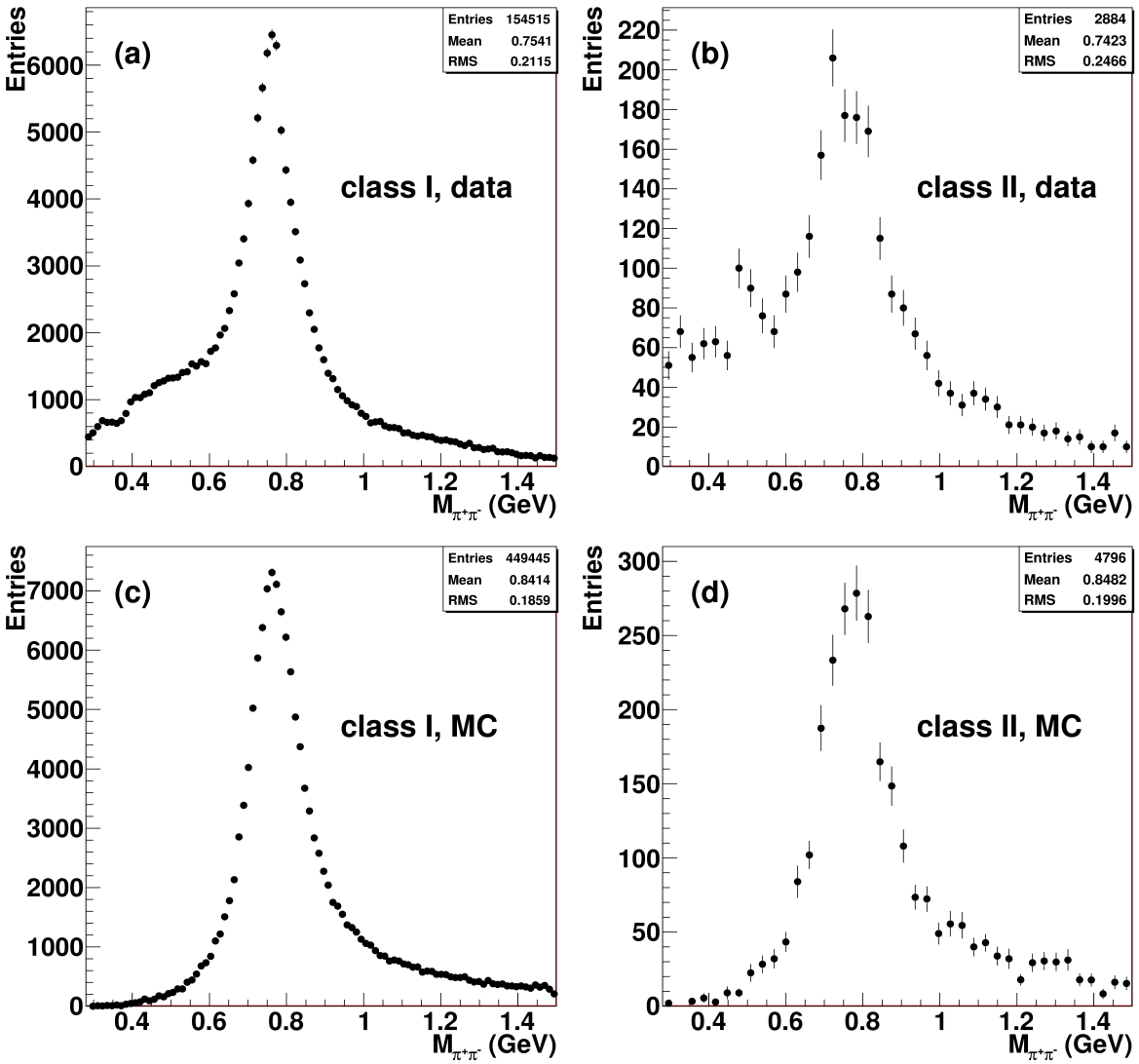


Figure 4.3: Mass distributions of $\rho(770)$ candidates for class I events (a), (c) and class II events (b), (d). Top row distributions are for data, bottom row plots are obtained from the MC.

Figures 4.3(c)-(d) show the ρ mass distributions for MC for class I and class II samples. Since the ω and ϕ background was not simulated, no excess is present at $M(\pi^+\pi^-) \lesssim 0.6\,\text{GeV}$, as for the data.

4.4 Monte Carlo Reweighting

The quality of the modelling of the ρ-meson decay products kinematics by the MC can influence results obtained with the method described in this chapter. For example, it was mentioned above that one of the assumptions is a correct description of the MVD acceptance for MVDSA tracks. Since it depends crucially on the track kinematics, the latter has to be well described by the MC.

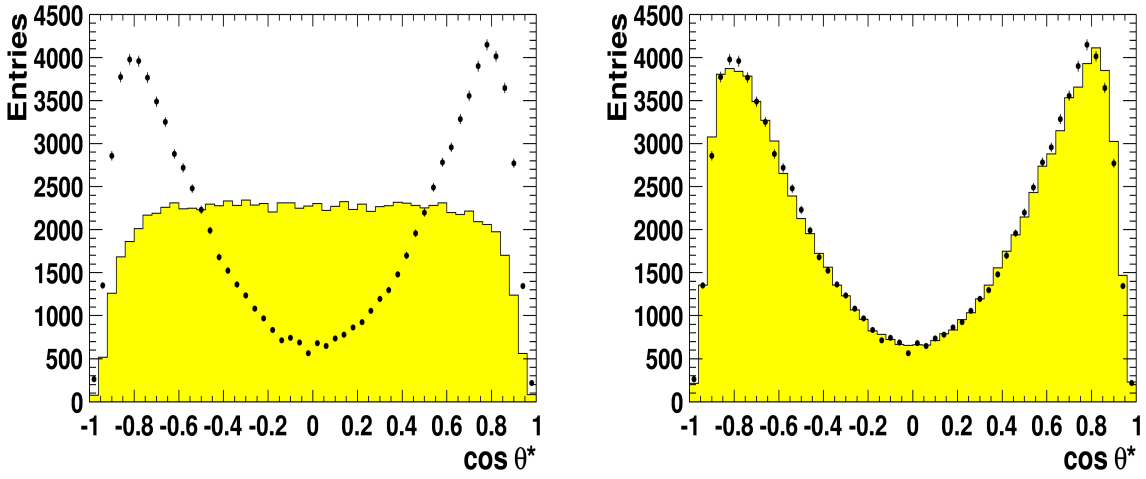
However it is known that the ZEUSVM MC does not provide a good description of the process kinematics. As discussed in Section 4.2, flat distributions for helicity angles are generated, which do not correspond to the measured ones [117]. Event kinematic variables (Q^2, W, t) are not well described as well [117]. Both effects spoil the description of the daughter products kinematics.

In order to reach a better agreement between data and MC, it is necessary to reweight the MC. The reweighting parametrisations are determined from the high statistics class I sample and are then applied both for class I and class II MC samples. One possibility is to reweight the helicity angles using the spin-density matrix elements as a function of Q^2 [117]. However for our purpose it is sufficient to describe only the daughter products kinematics of the ρ-meson, since the method described in this chapter is based on the detection of these particles. Hence, no attempt is made to perfectly describe the event kinematics and all of the helicity angles but rather to describe the ρ-meson kinematics and decay angles in its rest frame. This is sufficient to provide a good description of the daughter pion kinematics. In contrast, the helicity angles approach requires in general also a good understanding of the Q^2, W and t variables since the angles are defined in the $\gamma^* p$-frame.

We define θ^* as the angle between the positive track and the ρ direction, in the ρ rest frame. It is obtained as follows: the laboratory frame is rotated such that its z axis coincides with the ρ direction, then a boost along the z direction to the ρ rest frame is performed. Figure 4.4 shows the $\cos\theta^*$ distribution for data and MC, for class I events. The MC was normalised to data area. One is forced to conclude that the data and the MC have very different shapes, hence the MC is reweighted. The following reweighting function was obtained from this distribution:

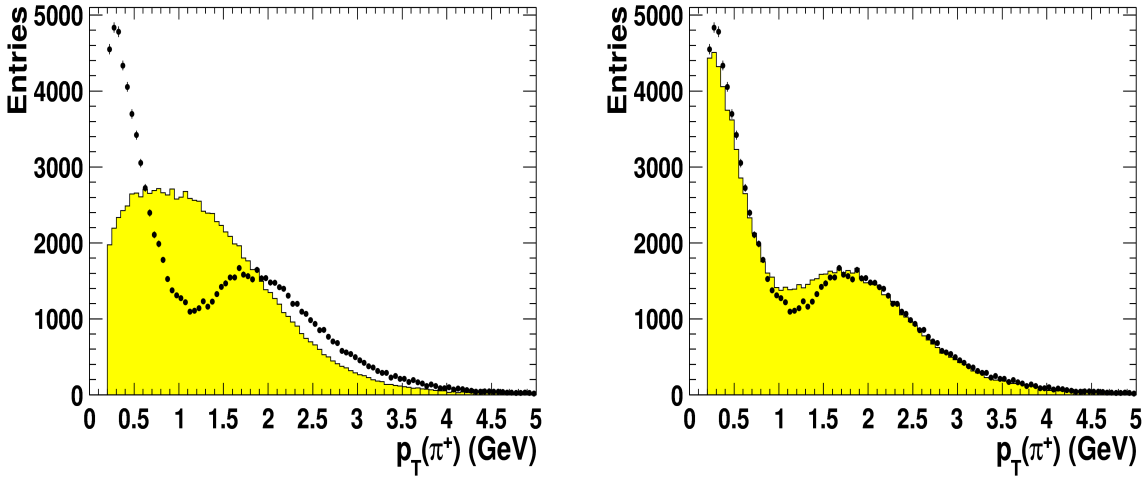
$$\zeta(\cos\theta^*) = 0.284 + 0.00138\cos\theta^* + 2.67(\cos\theta^*)^2$$

The $\cos\theta^*$ distribution after application of this reweighting is shown in Fig. 4.4(b). A good agreement is observed. The observable θ^* has direct influence on the kinematics of



(a) (b)

Figure 4.4: The $\cos\theta^*$ distribution for class I before (a) and after (b) reweighting. The points show the data, the histogram represents the MC. MC was normalised to data area.



(a) (b)

Figure 4.5: The distributions of the transverse momentum of the positive pion, $p_T(\pi^+)$, for class I events before the $\cos\theta^*$ reweighting (a) and after it (b). The points show the data, the histogram represents the MC. MC was normalised to data area.

the daughter products in the laboratory frame. Fig. 4.5 shows the transverse momentum distribution of the positive pion for class I before and after $\cos\theta^*$-reweighting. The agreement becomes much better after reweighting.

The azimuthal angle of the scattered electron as well as that of the ρ meson are not well described by the MC. One of the possible reasons is that in the MC, no trigger requirements were imposed, while data must be always accepted by certain trigger slots, mostly by the inclusive DIS triggers. The fact that the efficiency of these triggers might depend on the azimuthal angle, especially at low electron scattering angles (due to the electron box cut), may lead to such a discrepancy. In order to correct for this, a reweighting in the azimuthal angle ϕ of the ρ-meson is performed with the following parametrisation:

$$\xi(\phi) = 1.28 - 0.33\,\phi + 0.080\,\phi^2 + 0.0123\,\phi^3 - 0.050\,\phi^4 - 0.0011\,\phi^5 + 0.0039\,\phi^6$$

The high order of the polynomial is necessary to account for the shape of the data-to-MC ratio. After this reweighting, the azimuthal distribution of the scattered electron, the ρ meson and its decay products is significantly improved (see Section 4.5).

In summary, each MC event gets a weight $w = \zeta(\cos\theta^*)\,\xi(\phi)$ with θ^* and ϕ defined on the true MC level.

4.5 Control Distributions

As discussed in the previous section, it is necessary to check whether the MC describes the data reasonably, and to understand, whether remaining differences (if any) influence the results. In this section, various distributions are compared between data and MC. For all plots, MC was normalised to the data.

Figure 4.6 shows the distributions of the transverse momentum p_T, the polar angle θ and the azimuthal angle ϕ of the ρ meson and of its positive decay product (π^+) for class I events. All quantities are calculated in the laboratory frame. The distributions for π^- are very similar to those of π^+. A reasonable agreement between the data and the MC simulation is observed. However some discrepancies are present for $p_T(\rho)$, Fig. 4.6(a), for $\theta(\rho)$, Fig. 4.6(b) and for $p_T(\pi^+)$, Fig. 4.6(d). They can arise due to the imperfect simulation of the ρ meson kinematics and as well as of the helicity angles (see Section 4.4). The deficit of the MC at low $p_T(\rho)$, Fig. 4.6(a), can be also associated to additional background in data. It was checked that results obtained in this chapter do not change significantly if these differences are removed by reweighting of a corresponding variable or by removing the region below $p_T(\rho) = 1.3\,\text{GeV}$, where the excess of the data is present (see Section 4.6).

Figure 4.7 shows the kinematic variables (p_T, θ and ϕ) of the ρ meson and of the long primary track for the class II sample. Similarly to class I, the MC undershoots the data at low values of the ρ transverse momentum. No other significant discrepancies are observed within statistical uncertainties. Figures 4.8(a)-(c) show the kinematic variables

of the MVDSA track in class II events. Somewhat different shapes are observed compared to the long tracks distributions, which can be attributed to the MVD acceptance. The MC describes the data well. Figures 4.8(d)-(e) show the number of hits in the Barrel MVD (BMVD) for MVDSA tracks in class II events. Excellent agreement of the data and the MC suggests good understanding of the MVD hit efficiency which was assumed so far. Figure 4.9 shows the distribution of the transverse and z-impact parameters for the MVDSA track for class II events. Similar shapes are observed for data and MC, hence the selected MVDSA tracks are consistent with real primary tracks and are not due to background. There is also some discrepancy which could arise due to the fact that not the latest alignment corrections were applied for the MVDSA track reconstruction algorithm [121].

Figure 4.10(a) shows the multiplicity distribution for the long tracks (those starting at or before the first CTD superlayer and passing at least three superlayers) for class II, including the non-interacted pion. The MC reasonably reproduces the distribution, in particular the fact that there are often several non-primary tracks. Some discrepancies are visible for cases with one and two long tracks. This can happen if the probability of certain hadronic reactions is not well simulated. This is exactly an effect searched for and does not pose a problem. The transverse impact parameter distribution for the long tracks (excluding the non-interacted pion) is shown in Fig. 4.10(b). It is very broad as expected for secondary tracks. The description by the MC is reasonable, indicating that there is no significant pollution from e.g. primary tracks to this sample.

In summary, the MC describes reasonably the properties of the ρ meson and of its decay products for both class I and class II samples. The remaining differences in track kinematics do not affect the results of this chapter significantly, which was verified by several checks (see Section 4.6).

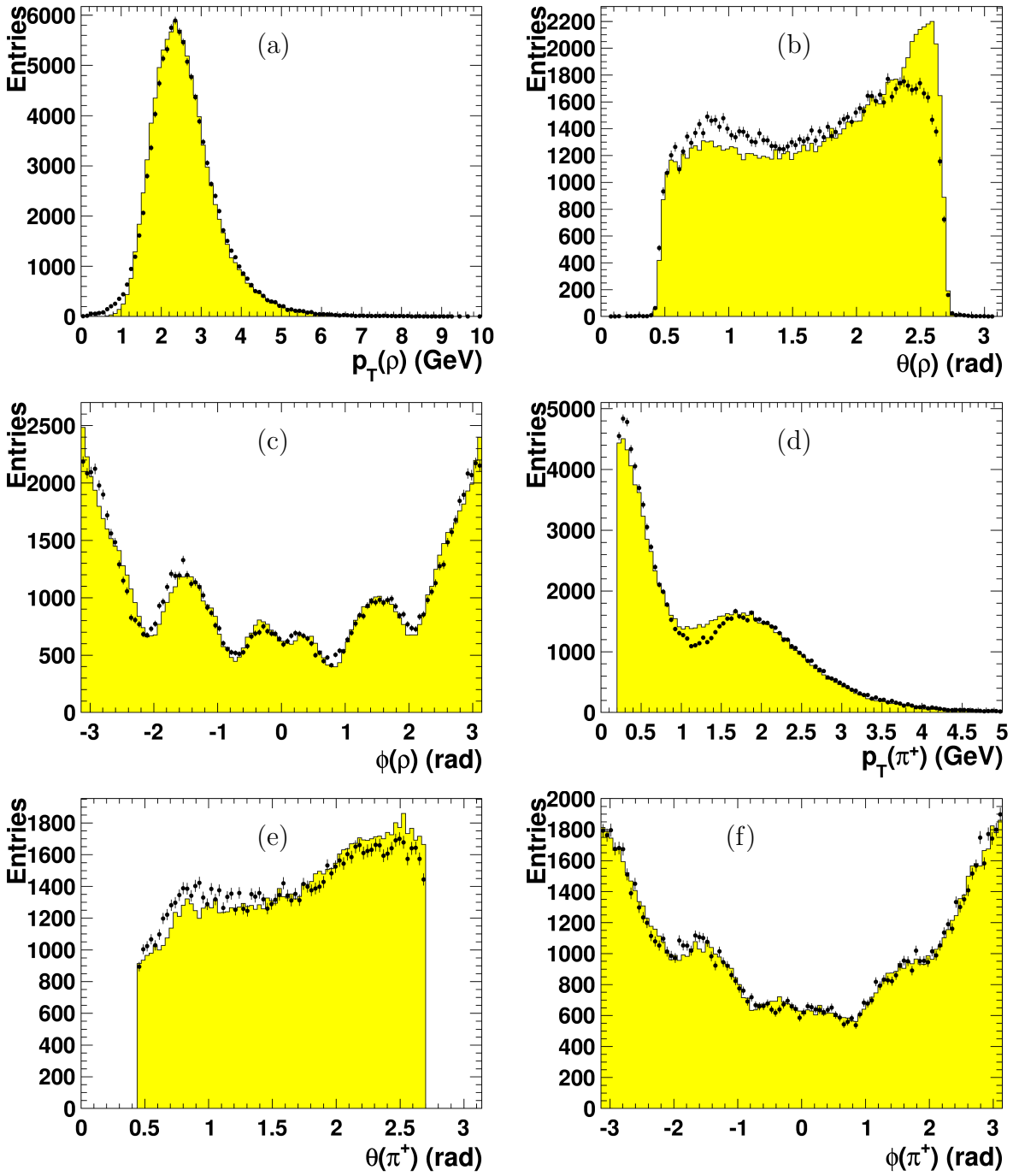


Figure 4.6: Control plots for the class I sample: the transverse momentum (a), the polar angle (b) and the azimuthal angle (c) of the ρ candidate; the transverse momentum (d), the polar angle (e) and the azimuthal angle (f) of the positive pion. The points show the data, the histogram represents the MC.

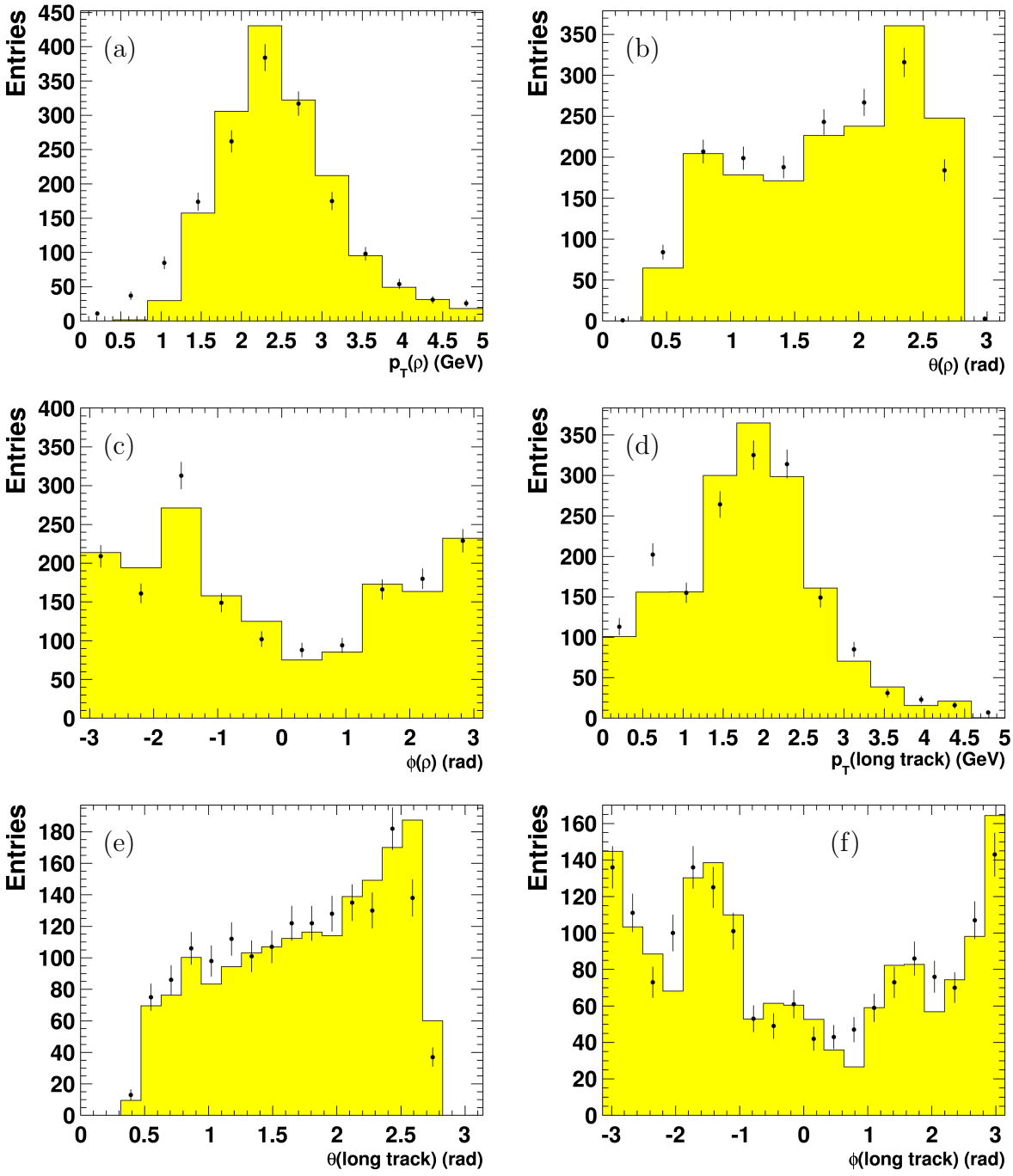


Figure 4.7: Control plots for the class II sample: the transverse momentum (a), the polar angle (b) and the azimuthal angle (c) of the ρ candidate; the transverse momentum (d), the polar angle (e) and the azimuthal angle (f) of the pion which did not interact (long track). The points show the data, the histogram represents the MC.

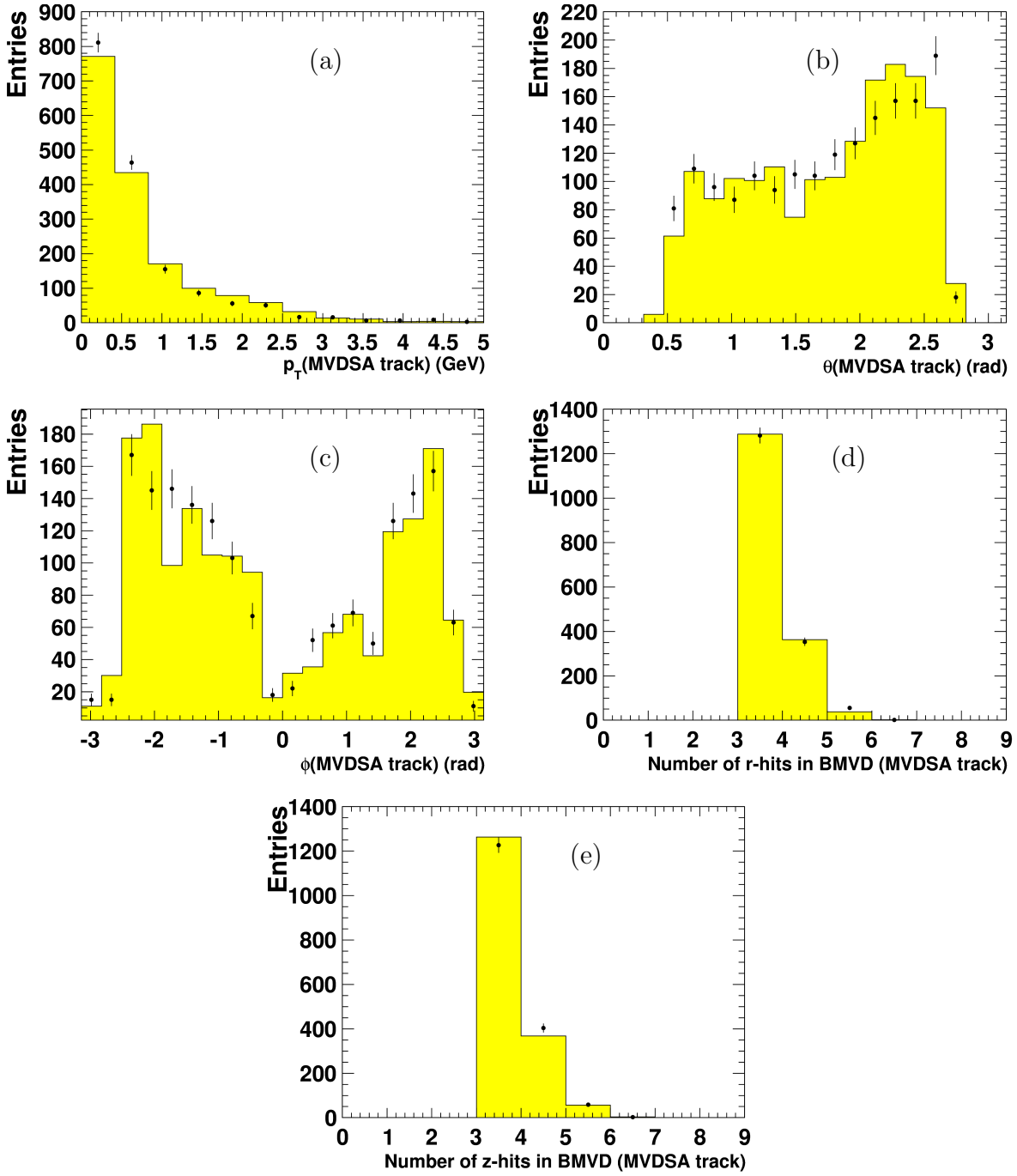


Figure 4.8: Control plots for the MVDSA track, class II. Shown are the transverse momentum (a), the polar angle (b), the azimuthal angle (c), the number of r-hits in the Barrel MVD (BMVD) (d) and the number of z-hits in BMVD (e). The points show the data, the histogram represents the MC.

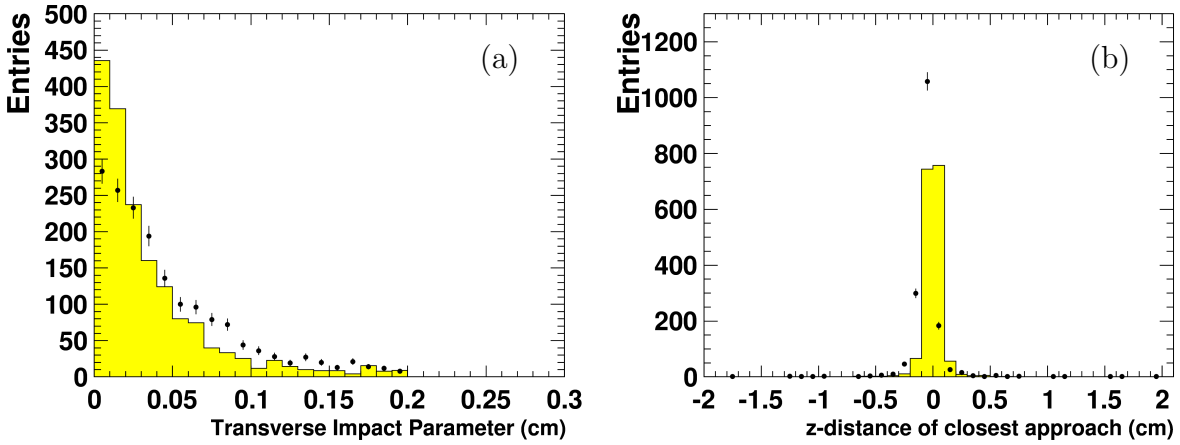


Figure 4.9: Control plots for the MVDSA track, class II. Shown are the transverse impact parameter with respect to the beam spot (a) and the z-impact parameter (b). The points show the data, the histogram represents the MC.

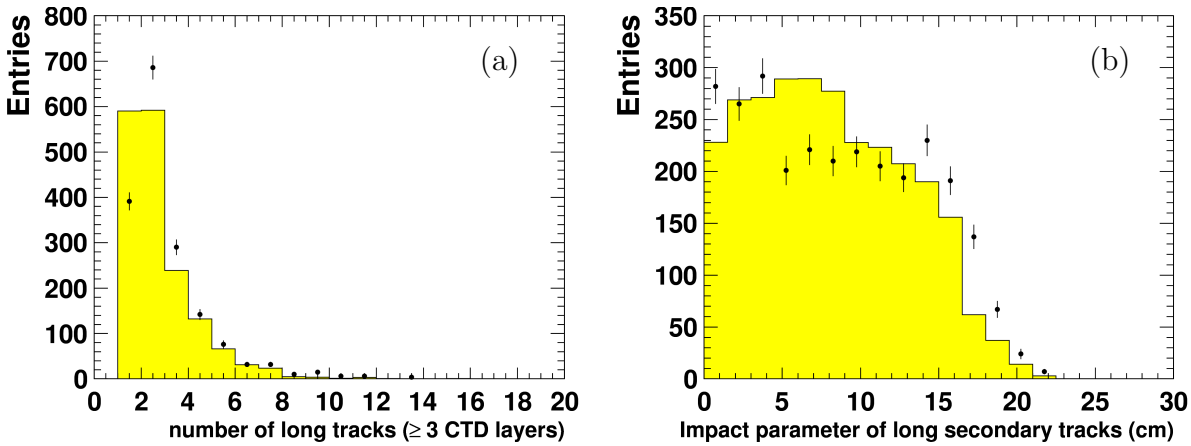


Figure 4.10: The number of long tracks including the non-interacted pion (a) and the transverse impact parameter of all secondary long tracks, excluding the non-interacted pion (b). Both distributions are for the class II sample. The points show the data, the histogram represents the MC.

4.6 Results

This section presents results of the hadronic interactions investigation. Two types of observables are constructed and studied: the *single ratio* and the *double ratio*.

The *single ratio*, R, is defined in the following way:

$$R(x_i) = \frac{N_{\mathrm{MVDSA}}(x_i)}{N_{\mathrm{long,\ primary}}(x_i) + N_{\mathrm{MVDSA}}(x_i)}.$$

Here $N_{\mathrm{MVDSA}}(x_i)$ is the number of MVDSA tracks within the class II which are in the bin i of some track or event variable x, e.g. track p_T; $N_{\mathrm{long,\ primary}}(x_i)$ is the total number of long primary tracks in the bin i both from class I and class II. If the variable x is not required to be in the bin i (or it is an event variable), there will be two long primary tracks and no MVDSA tracks in class I events ($N_{\mathrm{long,\ primary}} = 2$, $N_{\mathrm{MVDSA}} = 0$) while for class II events there will be one long primary and one MVDSA track ($N_{\mathrm{long,\ primary}} = 1$, $N_{\mathrm{MVDSA}} = 1$).

The single ratio R determined in this way is proportional to the probability for a track to interact hadronically in the MVD-CTD transition region and to the MVD acceptance for MVDSA tracks. Assuming that the latter is well modelled by the Monte Carlo, it is useful to construct the *double ratio*:

$$\epsilon(x) = \frac{R_{\mathrm{DATA}}(x)}{R_{\mathrm{MC}}(x)},$$

where $R_{\mathrm{DATA}}(x)$ and $R_{\mathrm{MC}}(x)$ are the single ratio obtained for data and Monte Carlo respectively. This quantity is a direct measure of how well the hadronic interaction probability is simulated by the MC and is the main observable in this method.

Figures 4.11(a)-(f) show results for the single ratio R for data and MC as a function of the track azimuthal angle ϕ, polar angle θ and transverse momentum p_T. Generally, a similar behaviour is observed in data and MC. In the azimuthal angle distribution, there are dips around $\phi = 0\,\mathrm{rad}$ and $\phi = \pm\pi\,\mathrm{rad}$ which are regions of reduced MVD acceptance. The θ distributions are smooth and show a slow rise towards the backward direction. Finally, the p_T distribution shows a steep rise towards low momenta.

Figure 4.12 shows the double ratio ϵ for the same kinematic variables. It follows that on average the Monte Carlo underestimates the interaction probability. The effect is consistent with a constant as a function of track angles ϕ and θ. It has a more pronounced structure as a function of p_T. A fit with a constant is performed for the whole range in ϕ and θ. The fit yields $\epsilon_\phi = 1.29 \pm 0.04$ and $\epsilon_\theta = 1.33 \pm 0.04$ for ϕ and θ respectively. The values are consistent with each other. Based on the observations from the p_T distribution, the following model is proposed: the variable ϵ is consistent with constants for $p_T < 1.5\,\mathrm{GeV}$ and for $p_T > 1.5\,\mathrm{GeV}$, however the constants are different in these momenta regions. Hence, the p_T distribution is fitted with a constant separately for $p_T < 1.5\,\mathrm{GeV}$

and for $p_T > 1.5\,\mathrm{GeV}$; a good quality of the fit suggests that this model is consistent with data. The fit yields $\epsilon_{p_T<1.5\,\mathrm{GeV}} = 1.36 \pm 0.05$ and $\epsilon_{p_T>1.5\,\mathrm{GeV}} = 0.96 \pm 0.09$ respectively. The latter number is consistent with unity with a 10 % uncertainty, hence at this level of precision no deviation of MC from data is observed. However, at transverse momenta below 1.5 GeV, the MC underestimates the probability of hadronic interactions.

Several checks were performed in order to assess the stability of the result. Distributions of $p_T(\rho)$ and of $\theta(\rho)$ were reweighted in the MC, such that they reproduced the data, in order to take into account improper description of the ρ kinematics; a cut of $p_T(\rho) > 1.3\,\mathrm{GeV}$ was applied in order to remove the region where there is a discrepancy between data and MC, Fig. 4.6(a); the calorimeter noise threshold for the elasticity cut was lowered to 300 MeV; the maximum allowed distance between the track and the island was lowered to 20 cm; the cone radius around the MVDSA track for the elasticity cut, ΔR, was varied by 0.3 (i.e. lowered to 0.2 and increased to 0.8); a cut of $Q^2_{\mathrm{da}} > 5\,\mathrm{GeV}^2$ was applied; the lower cut on the ρ mass was increased to 0.65 GeV. The biggest change of the result comes from the variation of the calorimeter noise threshold ($\epsilon_{p_T<1.5\,\mathrm{GeV}} = 1.25$ and $\epsilon_{p_T>1.5\,\mathrm{GeV}} = 0.8$) and of the cone radius for the elasticity cut ($\epsilon_{p_T<1.5\,\mathrm{GeV}} = 1.24$ and $\epsilon_{p_T>1.5\,\mathrm{GeV}} = 0.86$ for $\Delta R = 0.2$ and $\epsilon_{p_T<1.5\,\mathrm{GeV}} = 1.51$ and $\epsilon_{p_T>1.5\,\mathrm{GeV}} = 1.06$ for $\Delta R = 0.8$). Other variations do not affect the result significantly. Based on this, the uncertainty on both $\epsilon_{p_T<1.5\,\mathrm{GeV}}$ and $\epsilon_{p_T>1.5\,\mathrm{GeV}}$ is estimated to be around ± 0.2.

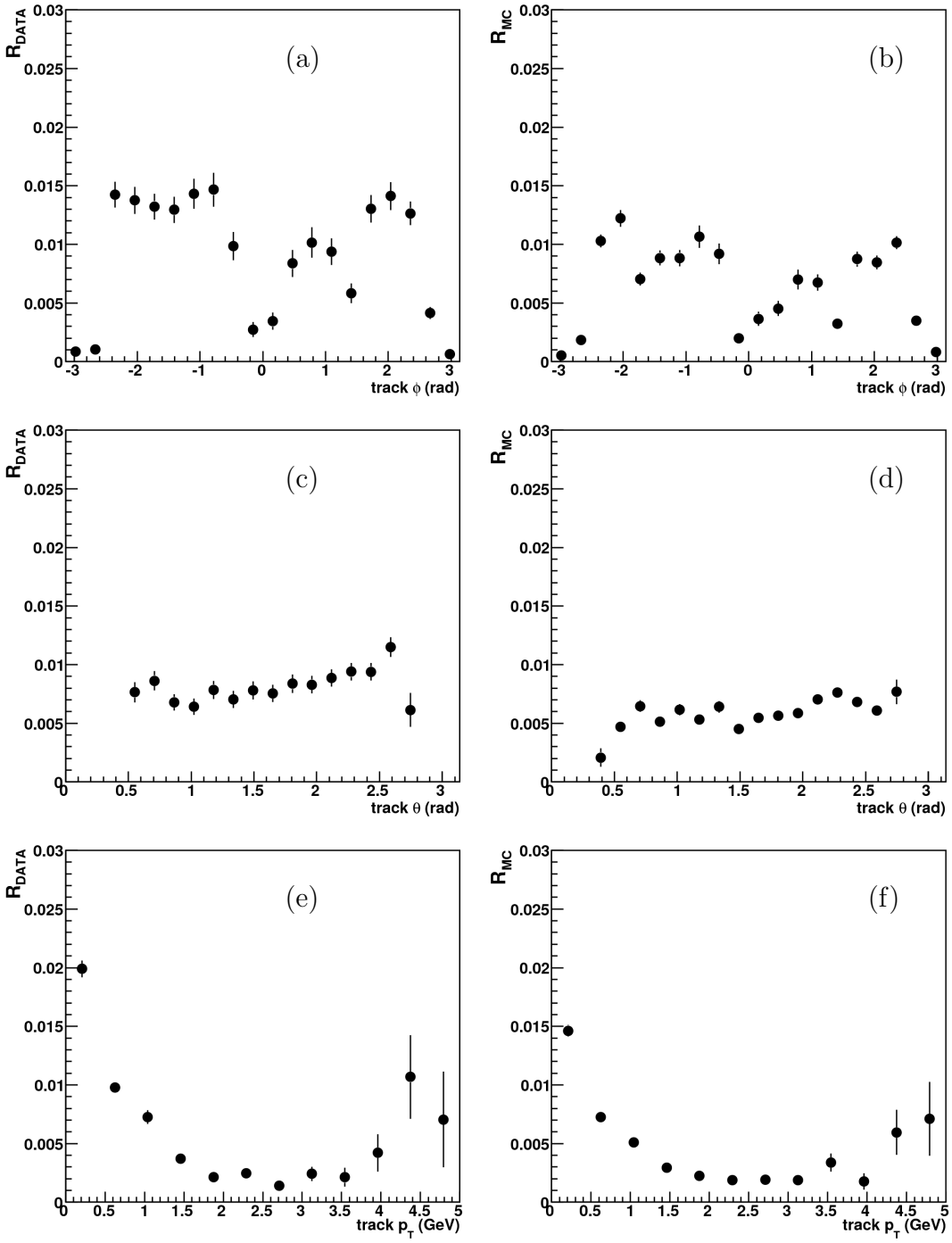


Figure 4.11: Single ratio R (see text for the definition) for data (left column) and MC (right column) as a function of the azimuthal angle ϕ (a)-(b), the polar angle θ (c)-(d) and the transverse momentum p_T (e)-(f) of the track.

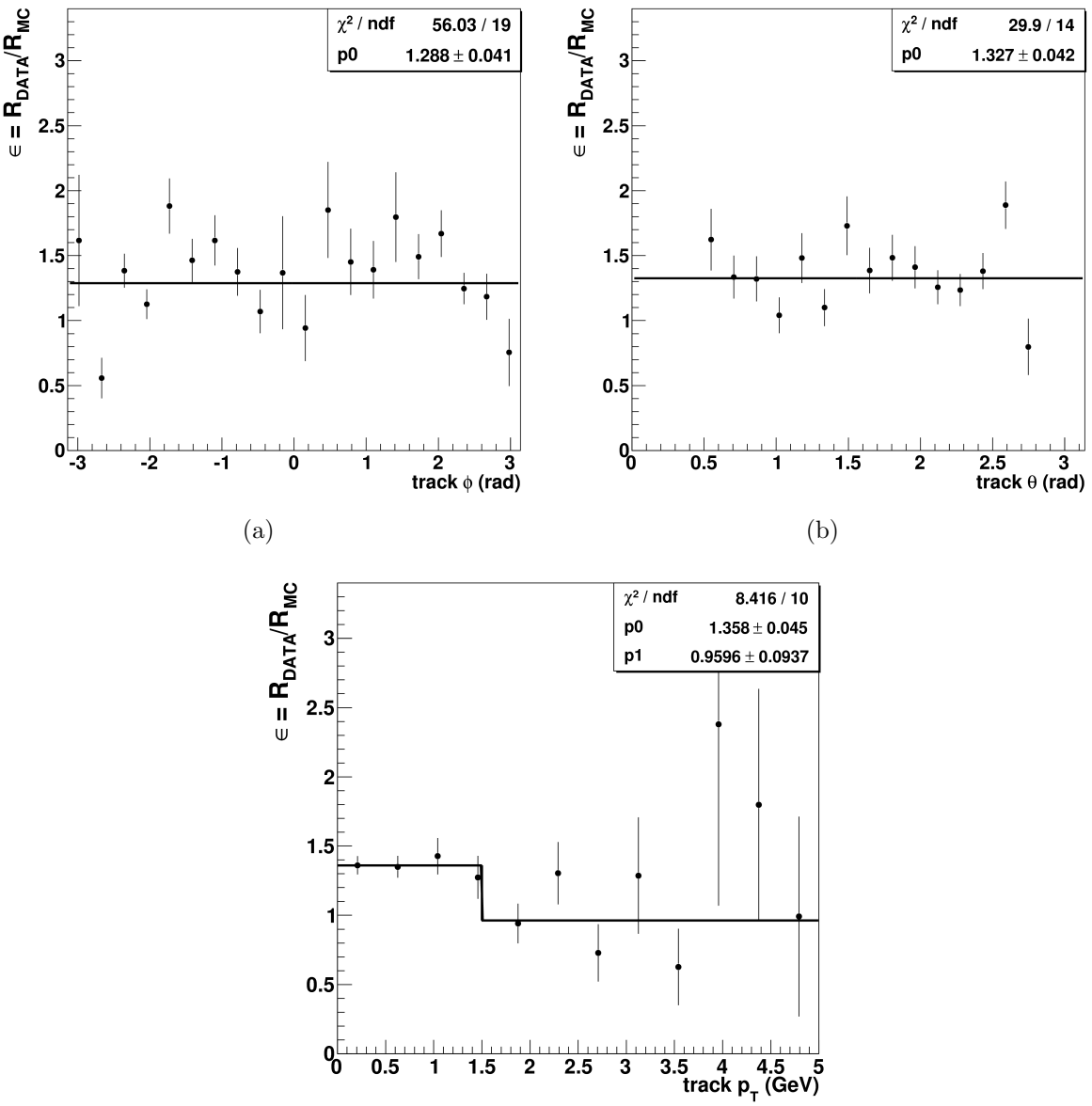


Figure 4.12: Double ratio ϵ (see text for the definition) as a function of the polar angle ϕ (a), the azimuthal angle θ (b) and the transverse momentum p_T (c) of the track.

4.7 Summary

In this chapter, the quality of the hadronic interactions modelling by the MC was studied. It was found that on average the MC underestimates the probability of interactions. The effect depends on transverse momentum p_T of the track. At low momenta of $p_T < 1.5\,\text{GeV}$, it is underestimated by a factor of ~ 1.4, while for $p_T > 1.5\,\text{GeV}$ a good description is observed.

One of the reasons for underestimation of hadronic interaction probability might be an imperfect simulation of cross-sections of certain reactions. An example could be the production of the $\Delta(1232)$ resonance:

$$\pi + p \rightarrow \Delta(1232)$$

The highest cross-section for this reaction is at the centre-of-mass energy of the pion-proton system close to the mass of the Δ-baryon, i.e. for pion momenta below 1 GeV if the proton is at rest. If MC underestimates the cross-section, this will lead to an underestimation of the interaction probability for pion momenta below 1 GeV. The other reason could be that the cross-section parametrisation used in GEANT [122, 123] is valid only at incident particle momentum of more than 2 GeV.

In order to take the observed difference into account, a correction has to be applied to the hadronic interaction probability in the MC. According to the results, a p_T-dependent correction must be used. No correction is applied for momenta above 1.5 GeV. A correction of 1.36 is applied for $p_T < 1.5\,\text{GeV}$. An uncertainty of ± 0.2 is assigned to the correction in both momentum regions, based on the effects of the procedure variations.

Chapter 5

Event Selection and Signal Extraction

This chapter describes the analysis procedure for the charm and beauty measurement performed in this thesis. It is organised as follows: the general idea of the secondary vertex method that was employed for heavy flavour tagging is discussed first. Data and Monte Carlo samples that were used are described next. A detailed list of event selection criteria is then given. The procedures of signal extraction and of charm and beauty cross sections determination are presented afterwards. Various corrections that were applied to the MC, which ensure a good description of the data, are then discussed. Comparisons of the data to the MC, showing the quality of the simulations, are presented in the following section.

5.1 Charm and Beauty Tagging with Secondary Vertices

In order to distinguish charm and beauty quark production from background processes, which we collectively call *light flavour* background, it is necessary to perform a certain tagging procedure. In this thesis, a lifetime tagging method is used. It is based on the fact that ground-state hadrons containing heavy quarks[1] have typically long lifetimes, due to their weak decays. Hence, heavy hadrons travel a certain distance before they decay[2], which leads to a *secondary* vertex. In contrast, in light flavour events, particles originate mostly from the primary vertex. If there are at least two charged particles among decay

[1] Shortly *heavy hadrons* in what follows

[2] On average, this distance (decay length) is given by $c\tau p/m$, where c is the speed of light in vacuum, τ is the average lifetime of the particle in its rest frame, p is the particle momentum and m is the particle mass. For ground states of charm hadrons, typical values $c\tau$ are around 100 µm, while for beauty hadrons $c\tau$ is in the order of 400 µm [101].

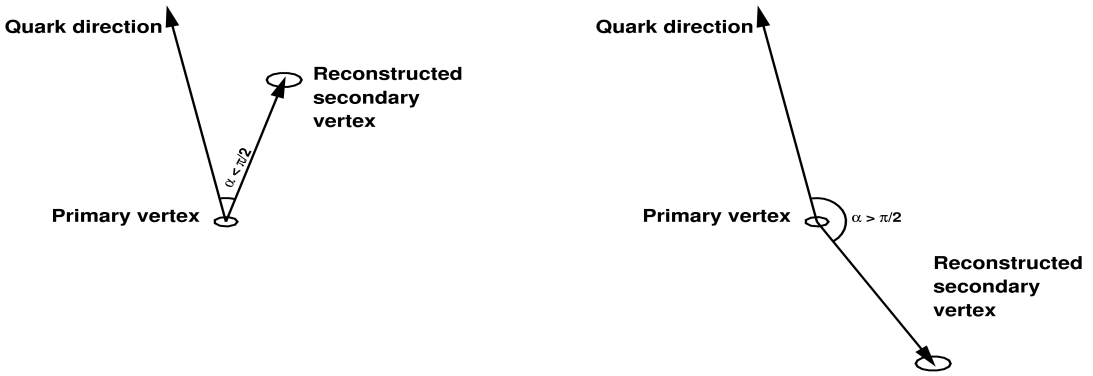


Figure 5.1: Positive (left) and negative (right) projected decay length.

products of a heavy hadron, a corresponding secondary vertex can be reconstructed. The *decay length*, which is the distance between the primary and secondary vertex, can serve as a discriminant variable between charm, beauty and light flavour production processes. The fact that the shapes of the discriminant variable distributions are different for each of these processes is exploited to determine their relative contributions.

In order to suppress background and to enhance the sensitivity to heavy flavours, it is useful to have an estimate of the initial charm or beauty quark flight direction. This gives a possibility to construct *signed* variables, e.g. a *projection* of the decay length on the quark direction. Large values of the angle between the decay length vector and the quark direction, in particular negative values of a projection ($\alpha > \pi/2$) are unphysical since the hadron flight direction should be close to the initial quark direction, see Figure 5.1. This can happen either due to combinatorial background (fake secondary vertex) or due to resolution effects.

High energy hadron jets provide an excellent tool for estimating initial quark properties, since they originate from hard partons and preserve information about them. In this thesis secondary vertices are associated to jets. By construction, at most one secondary vertex can be reconstructed for each jet.

5.2 Data and Monte Carlo Samples

The dataset analysed in this thesis was collected in the period from 2004 to 2007 and corresponds to an integrated luminosity of about $354\,\mathrm{pb}^{-1}$. The proton beam energy

during this period was 920 GeV, while the electron[3] energy was 27.5 GeV. Table 5.1 shows more details about various subperiods.

Period	Lepton Type	Integrated Luminosity, pb^{-1}
2003-2004	e^+	30.60
2005	e^-	133.14
2006	e^-	52.69
2006-2007	e^+	137.30

Table 5.1: Details on data samples analysed in this thesis.

Monte Carlo simulations were used to obtain shapes of signal and background distributions (templates) and to determine acceptance corrections. Inclusive production of charm and beauty quarks in DIS was used for the signal MC. Similarly, an inclusive light flavour DIS sample was generated for the background MC.

The RAPGAP 3.0 generator [71, 72] was used to simulate the heavy quark production. Two main production mechanisms are available in the generator: the Boson-Gluon Fusion (BGF) and resolved photon processes (see Section 1.5). The BGF mode was selected for the main samples. Additionally, resolved photon samples were generated in order to study the sensitivity of results to the underlying Monte Carlo model. The charm and beauty masses were set to 1.5 GeV and 4.75 GeV respectively. The CTEQ5L [124] set was used to parametrise the proton parton densities. For the resolved photon sample, the GRV-G LO [125, 126] PDF set was chosen to parametrise the photon parton densities. The number of flavours was set to 5 and Λ_{QCD} was set to 0.146 GeV consistent with the PDF. No explicit cuts on the quark momentum in the photon-gluon rest frame were applied. Parton fragmenation in RAPGAP is done with the Lund string model as implemented in the PYTHIA MC generator [127]. The longitudinal fragmentation function was set to the Bowler [73] parametrisation, see equation (1.3).

RAPGAP generates events according to the leading order (LO) QCD calculation ($\mathcal{O}(\alpha_s)$), i.e. it is a LO matrix element generator. In order to simulate higher order QCD effects, the initial and final state parton showers were switched on. They are based on the leading logarithm DGLAP [20–23] splitting functions in leading order of α_s [72]. Initial and final state real photon emission, as well as virtual QED corrections were simulated by the HERACLES generator [119] interfaced to RAPGAP.

Table 5.2 shows the number of events generated for the charm and beauty BGF mode samples and equivalent integrated luminosity normalised to the data integrated luminosity. The corresponding luminosities of the resolved photon samples are very similar.

The light flavour background was simulated with the DJANGOH 1.6 generator [128] which was interfaced to ARIADNE 4.12 [129]. The CTEQ5D [124] PDFs were used and

[3]By electron we mean the lepton which constitutes the beam. It can be either electron or positron depending on the running period, see Table 5.1.

Period	Charm, $Q^2 > 1.5\,\mathrm{GeV}^2$		Charm, $Q^2 > 4.0\,\mathrm{GeV}^2$		Beauty, $Q^2 > 1\,\mathrm{GeV}^2$	
	Events	$L_{\mathrm{MC}}/L_{\mathrm{data}}$	Events	$L_{\mathrm{MC}}/L_{\mathrm{data}}$	Events	$L_{\mathrm{MC}}/L_{\mathrm{data}}$
2003-2004	2 478 580	1	4 400 000	3	1 060 000	27
2005	9 000 000	1	10 000 000	2	1 920 000	18
2006	3 325 000	1	5 900 000	3	840 000	16
2006-2007	8 657 678	1	17 600 000	3.5	4 680 000	18

Table 5.2: Number of generated events and corresponding integrated luminosity relative to data for the BGF mode charm and beauty Monte Carlo samples used in this thesis.

quarks were taken to be massless. The minimum Q^2 cut was set to $4\,\mathrm{GeV}^2$. The number of generated events corresponds to the data luminosity. DJANGOH simulates processes of both light and heavy quark production. Hence, events with heavy flavours were removed in order to extract a light flavour sample.

The generated MC samples were passed through a simulation of the ZEUS detector based on GEANT3 [122]. The resulting events were processed in the same way as data, i.e. the same trigger, reconstruction and selection algorithms were applied.

5.3 Event Selection

This section describes main steps for selection of a Neutral Current (NC) DIS events sample for the measurement. First, details on the online selection, i.e. requirements of relevant triggers are given. The offline selection is explained afterwards. Finally, the jet reconstruction and selection criteria are described.

Online Selection

As discussed in Section 2.7, events were selected online by means of a three-level trigger system. At the analysis level, only events accepted by certain *trigger slots* were considered.

One of the following *first level* trigger slots was required (the logical *or*): *FLT28*, *FLT30*, *FLT40*, *FLT41*, *FLT43*, *FLT44*, *FTL46* or *FLT62*. They demand typically a good track or an energy deposit in the calorimeter [14].

At the *second level*, events accepted by the *SPP1* trigger slot were selected. *SPP1* has the following criteria:

- at least one of the eight first level triggers slots listed above;
- $E-p_Z > 30\,\mathrm{GeV}$, where $E-p_Z = \sum_i (E - P_Z)_i$, the sum runs over all ZUFOs. For a fully reconstructed DIS event one expects $E-p_Z = 2E_e = 55\,\mathrm{GeV}$ (see Section 4.3). The given cut reduces photoproduction and non-ep background;
- One of the following conditions (logical *or*):

  - Electromagnetic energy (EMC) in RCAL greater than 2.5 GeV;
  - EMC in BCAL greater than 2.5 GeV;
  - EMC in FCAL greater than 10 GeV;

- Energy of the scattered electron, E'_e, greater than 5 GeV.

The following third level trigger slots were required:

- *SPP02* for 2003-2004 (e^+) and 2005 (e^-) periods;
- *SPP09* or *HFL17* or *HPP31* for 2006 (e^-) and 2006-2007 (e^+) periods.

The criteria of these trigger slots include:

- *SPP01* second-level trigger slot;
- a cut on the minimum energy E'_e of the scattered electron;
- *the electron box cut* - the electron position is required to be outside a rectangular region of certain size ($a \times b\,\mathrm{cm}^2$) centred at ($X$, Y)=(0, 0) on the surface of the RCAL;
- restrictions on the maximum and minimum values of $E - p_Z$;
- The *HPP31* slot additionally includes cuts on tracks, on the primary vertex and on the Q^2 quantity.

Particular values of the cuts are vary for different slots and are summarised in the Table 5.3.

Requirement	**SPP02**	**SPP09**	**HFL17**	**HPP31**
SLT slot	SPP1	SPP1	SPP1	SPP1
$E'_{e,MIN}$ (GeV)	4	4	-	7
Electron box cut (cm^2)	12 x 12	15 x 15	-	12 x 12
$(E - p_Z)_{MIN}$ (GeV)	30	30	-	34
$(E - p_Z)_{MAX}$ (GeV)	100	100	-	75
$Q^2_{TLT,MIN}$ (GeV^2)	-	-	-	6
Tracking	-	-	-	1 CTD track with $p_T > 0.2$ GeV number of bad tracks <6
Primary vertex	-	-	-	$-60 < Z_{\mathrm{VTX}} < 60$ cm

Table 5.3: Criteria of the third-level triggers that were used for the analysis.

During data taking, the status of various detector subsystems was continuously monitored and stored to special flags attached to each event. Most important are the EVTAKE and MVDTAKE flags. EVTAKE shows whether the CTD, CAL and luminosity monitor were functioning properly. Similarly, MVDTAKE flag indicates whether the MVD was in a good state. Only events flagged as good by EVTAKE and MVDTAKE are selected for further analysis.

Offline NC DIS Selection

A more stringent NC DIS event selection was applied offline. Stronger cuts on the scattered electron, on $E - p_Z$ and on the primary vertex position were applied and the phase space of the measurement was defined. These cuts are listed below:

- The presence of at least one SINISTRA (see Section 3.1) electron candidate is required;
- Electron probability $> 90\,\%$ – for each candidate, the algorithm assigns a certain probability for it to be caused by an electromagnetic cluster. Only the most probable candidate in each event is considered. The cut on its probability ensures a good purity of electromagnetic showers identification;
- Energy of the scattered electron $E'_e > 10\,\mathrm{GeV}$ – suppresses fake electron candidates coming from e.g. pions decaying into photons and ensures a good efficiency of the finder;
- Electron box cut: $26 \times 26\,\mathrm{cm}^2$ – this is the region around the beam-pipe where the detection of the scattered electron is not possible (*beam-pipe hole*) or has reduced acceptance;
- bad cell rejection: events are rejected if $7.515 \leq X_e \leq 31.845\,\mathrm{cm}$ and $7.90 \leq Y_e \leq 31.90\,\mathrm{cm}$, where X_e and Y_e are the coordinates of the scattered electron – these are regions where the calorimeter did not perform well;
- $44 < E - p_Z < 65\,\mathrm{GeV}$;
- $-30 < Z_{\mathrm{VTX}} < 30\,\mathrm{cm}$ – further reduces non-ep background and ensures high CTD and MVD acceptance for tracks coming from the primary vertex and from the region close to it.

The kinematics of ep-scattering, for the fixed $\sqrt{s}$, is defined by two independent kinematic variables. In this thesis, the following phase space region in terms of Q^2 and y was considered:

- $5 < Q^2_{\mathrm{da}} < 1000\,\mathrm{GeV}^2$, where Q^2_{da} is the photon virtuality reconstructed with the double-angle method [113];
- $y_{\mathrm{JB}} > 0.02$, where y_{JB} is the inelasticity reconstructed using the Jacquet-Blondel method [112] which has the best resolution at low-y;
- $y_{\mathrm{el}} < 0.7$, where y_{el} is the inelasticity reconstructed with the electron method [113] which has the best resolution at high-y.

The lower Q^2 cut is related to the inability to reconstruct the electron close to the beam-pipe (similarly to the electron box cut). The SINISTRA algorithm is expected to work reliably up to certain values of Q^2, which determines the upper Q^2 cut. The lower y cut ensures the presence of hadronic activity in the detector; the upper y cut reduces the photoproduction background.

Jet Selection

In the selected DIS events sample, jets are reconstructed and selected. The k_T-clustering algorithm [109] in the longitudinally invariant mode with the E_T-recombination scheme [110] was used for the jet reconstruction (see Section 3.5). All reconstructed energy flow-objects (ZUFOs) were used as input to the algorithm. Jets containing the scattered electron are rejected. Jets with the following kinematic cuts are selected for the further analysis:

- $E_T^{\mathrm{jet}} > 4.2(5)\,\mathrm{GeV}$ for charm (beauty);
- $-1.6 < \eta^{\mathrm{jet}} < 2.2$.

Different E_T^{jet} cuts were used for extraction of charm and beauty cross sections. A cut of $E_T^{\mathrm{jet}} = 5\,\mathrm{GeV}$ is sufficiently low for the beauty quark to be measured down to very low transverse momenta, thanks to its large mass ($m_b \sim 4.75\,\mathrm{GeV}$). In contrast, the charm quark is much lighter ($m_c \sim 1.5\,\mathrm{GeV}$), hence this cut does not allow a measurement of the charm quark below a certain momentum. Therefore, it was lowered in order to increase the phase space for the charm measurement[4]. A cut of $E_T^{\mathrm{jet}} > 4.2\,\mathrm{GeV}$ is used to obtain all plots shown in this chapter, except for the beauty enriched sample (Section 5.7) where $E_T^{\mathrm{jet}} > 5.0\,\mathrm{GeV}$ was applied.

5.4 Secondary vertex reconstruction

After the selection described above, a search for secondary vertices is performed. Vertices are reconstructed from charged tracks that come from jets. Each jet can have one

[4]By lowering the cut further, one can perform a measurement in a larger phase space region; however, this was not considered in this thesis.

associated secondary vertex. Charged tracks that belong to a jet are selected with a *cone* cut: a track is considered as coming from a jet if $R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 1$, where $\Delta\eta = \eta^{\text{jet}} - \eta^{\text{trk}}$ and $\Delta\phi = \phi^{\text{jet}} - \phi^{\text{trk}}$ are distances between the track and the jet axis in η and ϕ respectively. The following cuts are imposed on tracks:

- $p_T > 500\,\text{MeV}$;
- at least 4 MVD hits[5].

These cuts ensure an excellent impact parameter resolution for tracks, which is crucial for secondary vertex reconstruction: the transverse momentum cut reduces amount of multiple scattering, while the cut on number of hits assures that MVD provides enough spatial information close to the secondary vertex.

If at least two tracks with the above criteria were associated to a jet, a secondary vertex is fitted from the selected tracks, otherwise the jet is discarded. The Deterministic Annealing Filter (DAF) which was described in Section 3.3 was used for the fit. The fit yields the coordinates of the vertex with full covariance matrix, the χ^2-value of the fit, as well as the mass of the vertex assuming the pion mass for all tracks. Figure 5.2 shows the distribution of the χ^2-value of the fit per degree of freedom, n_{dof}, for data and light flavour, charm and beauty MC samples; the MC contributions were normalised by scaling factors obtained from the fit which is discussed in Section 5.5. Data tend to have a slightly worse χ^2/n_{dof}. Several effects might lead to this, such as a not perfect simulation of the number of primary tracks by the MC which are picked up in the secondary vertex fit. It was checked that in the regions relevant for signal extraction, there are no indications for this discrepancy (see Section 5.7). The following quality cuts are applied on vertices:

- $\chi^2/n_{\text{dof}} < 6$ - ensures a good fit quality;
- $|Z_{\text{SV}}| < 30\,\text{cm}$, where Z_{SV} is the Z-coordinate of the secondary vertex - ensures a good CTD and MVD acceptance for tracks that belong to the vertex.

The two-dimensional decay length is defined as the distance between the secondary vertex and the primary vertex in the XY-plane (i.e. in the plane perpendicular to the beam). The beam spot position (see definition in Section 3.3) is taken for the primary vertex coordinates. This is more robust than using the reconstructed primary vertex on the event basis, since it can be biased by tracks coming from the secondary vertex due to the other heavy quark in the event (charm or beauty quarks are created in pairs). It may also introduce additional systematic uncertainties, e.g. due to the modelling of the track multiplicity by the MC. The *projected decay length* L_{XY} is defined as the projection of the two-dimensional decay length on the jet axis in the XY-plane, see Fig. 5.3(a) for

[5]As discussed in Section 2.4, each MVD layer provides one $r\phi$- and one rz-measurement (hit). Thus, requiring at least four MVD hits usually selects tracks which are measured in two MVD layers.

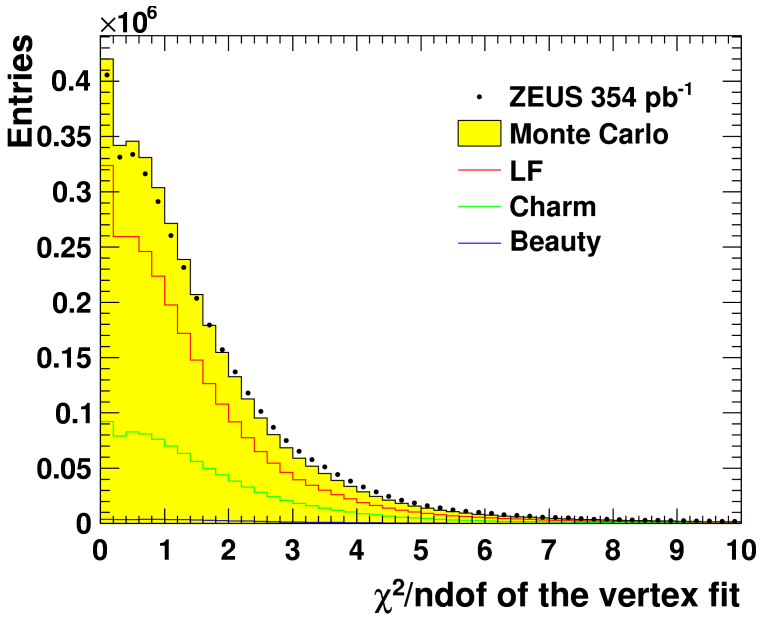
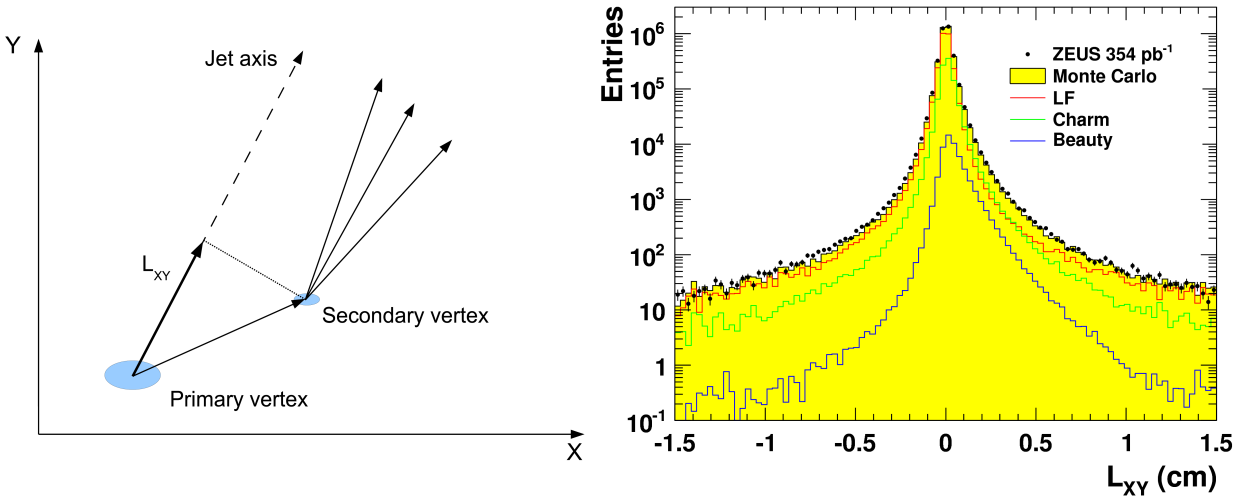


Figure 5.2: Distribution of $\chi^2/n_{\rm dof}$ of the secondary vertex fit. Points show the data; solid lines represent different MC contributions while the shaded histogram shows their sum. Different MC are scaled according to the fit results (see Section 5.5).



(a) (b)

Figure 5.3: Definition of the projected decay length L_{XY} (a) and its distribution (b). Other details are as in the Fig. 5.2.

an illustration. As it was discussed in Section 5.1, the association of jets with secondary vertices allows the construction of signed variables. The projected decay length is defined to have a positive sign, if the angle between the decay length and the jet axis in the XY-plane is less than 90° and it is negative otherwise.

Figure 5.3(b) shows the projected decay length distribution. The contribution from light flavours is mostly symmetric and reflects the detector resolution effects. However, charm and beauty distributions show an asymmetry towards positive decay lengths, which is related to the long lifetime of heavy flavour hadrons. A cut $L_{XY} < 1\,\text{cm}$ was imposed in order to reject long-lived light flavour hadrons, such as K^0_S-mesons or Λ-baryons. Additionally, a cut on the mass of the vertex of $1 < m_{\text{vtx}} < 6\,\text{GeV}$ was applied which suppresses K^0_S and Λ contributions. If a good vertex was found, this vertex is accepted for further analysis.

Due to a finite detector resolution, vertices with large values of the projected decay length L_{XY} might be reconstructed in light flavour events, as Figure 5.3(b) shows. However, a large uncertainty is typically associated to L_{XY} in such cases, since the reconstructed secondary vertex is still consistent with the primary vertex. On the other hand, secondary vertices reconstructed from the decay products of a heavy hadron are expected to lead often to a smaller uncertainty compared to L_{XY}, reflecting the non-zero decay length of the hadron. Hence, in order to enhance the separation power between signal and background, the *significance* of the projected decay length is introduced. It is defined as:

$$S = \frac{L_{XY}}{\sigma(L_{XY})},$$

where $\sigma(L_{XY})$ is the uncertainty on L_{XY}. It is calculated using the secondary vertex covariance matrix, i.e. the correlations between the X and Y coordinates are fully taken into account. The beam spot size (see Section 3.3) is used for the uncertainty of the primary vertex. The beam tilt correction is small and the uncertainty which it introduces to the beam spot position is neglected.

Figure 5.4 shows the significance distribution. It has similar features as the L_{XY} distribution: charm and beauty are clearly asymmetric toward high positive significance values, while the light flavour background is mostly symmetric. However, the separation between the flavours is better than for the L_{XY} variable: at high values of the significance ($S \sim 10 \div 20$), charm and beauty contributions are similar to the light flavour, while in the Figure 5.3(b), they are much smaller at high values of L_{XY} where one expects the best separation.

The fact that the shapes of the significance distribution are different for charm, beauty and light flavour is exploited for determination of relative contribution of each component, as it is described in the following section.

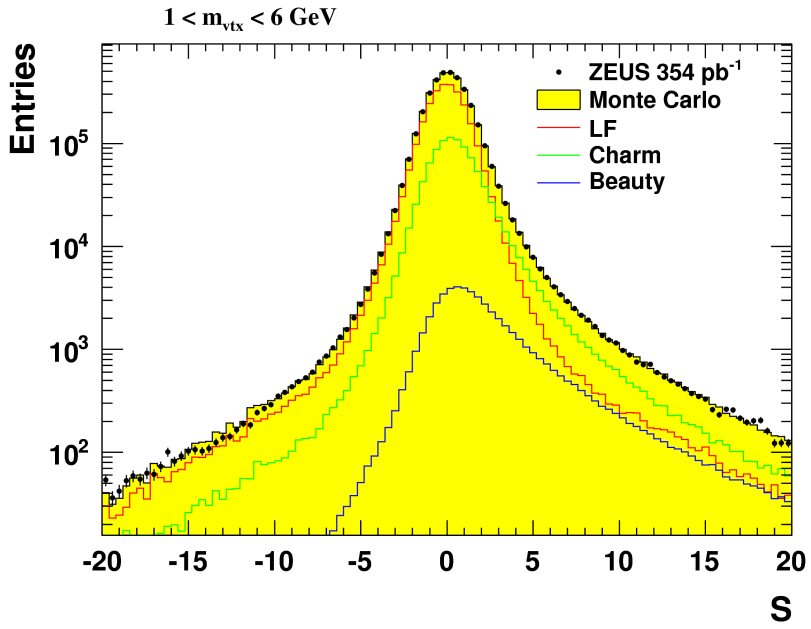


Figure 5.4: Distribution of the projected decay length significance. Other details are as in the Fig. 5.2.

5.5 Signal Extraction

The projected decay length significance S is used as the main discriminating variable between the light and heavy quarks production, thanks to its sensitivity to the lifetime of produced hadrons. In order to enhance the separation power to identify beauty events, the vertex mass m_{vtx} is used in addition: the sample is splitted into three mass bins: 1–1.4, 1.4–2 and 2–6 GeV. Figure 5.5 shows the significance distribution in the three mass bins. As can be seen, beauty is more pronounced in the third bin thanks to the fact that beauty hadrons are much heavier than charm or light flavour hadrons.

However, these distributions are still dominated by the light flavour background. In order to reduce it, the *negative subtraction*, or *mirroring* is performed: the negative part of the significance distribution ($S < 0$) is mirrored onto and subtracted from the positive one ($S > 0$). Since the light flavour significance distribution is mostly symmetric, it is expected to become very small after subtraction. On the other hand, heavy flavours have a positive asymmetry, hence one expects a large contribution from them.

Figure 5.6 shows the result of the negative subtraction procedure. The light flavour is dramatically reduced, and the distributions are dominated by charm and beauty production. In the highest mass bin the beauty contribution is particularly enhanced.

In order to extract contributions of charm and beauty quark production, a binned

χ^2-fit is performed. Before the fit, the MC contributions are normalised according to the data integrated luminosity. The mirrored significance in three mass bins and the integral of the unmirrored distribution are used in the fit simultaneously. A cut on maximum absolute value of significance, $|S| < 20$, is applied, since light flavour background due to long-lived strange hadrons increases at $S > 20$. Additionally, only the region of $|S| > 4$ is used in the mirrored significance distributions, as shown in Fig. 5.6. This suppresses the contamination of primary tracks to the secondary vertex, i.e. leads to a better separation between the secondary and the primary vertex. Besides, a dependence of the cross section on the jet azimuthal angle, ϕ^{jet}, was observed if the region of $|S| < 4$ was included in the

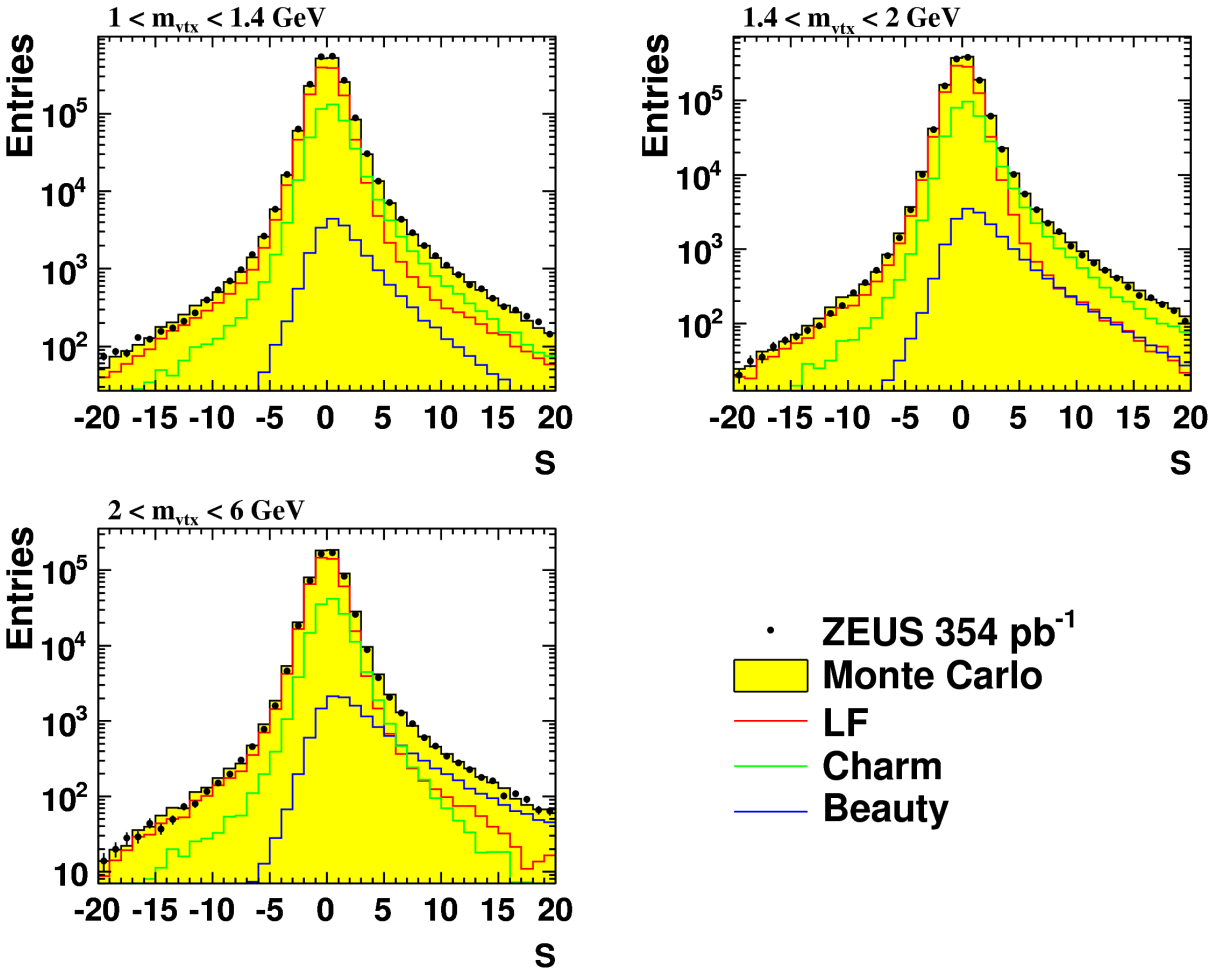


Figure 5.5: Decay length significance in three mass bins. Other details are as in the Fig. 5.2.

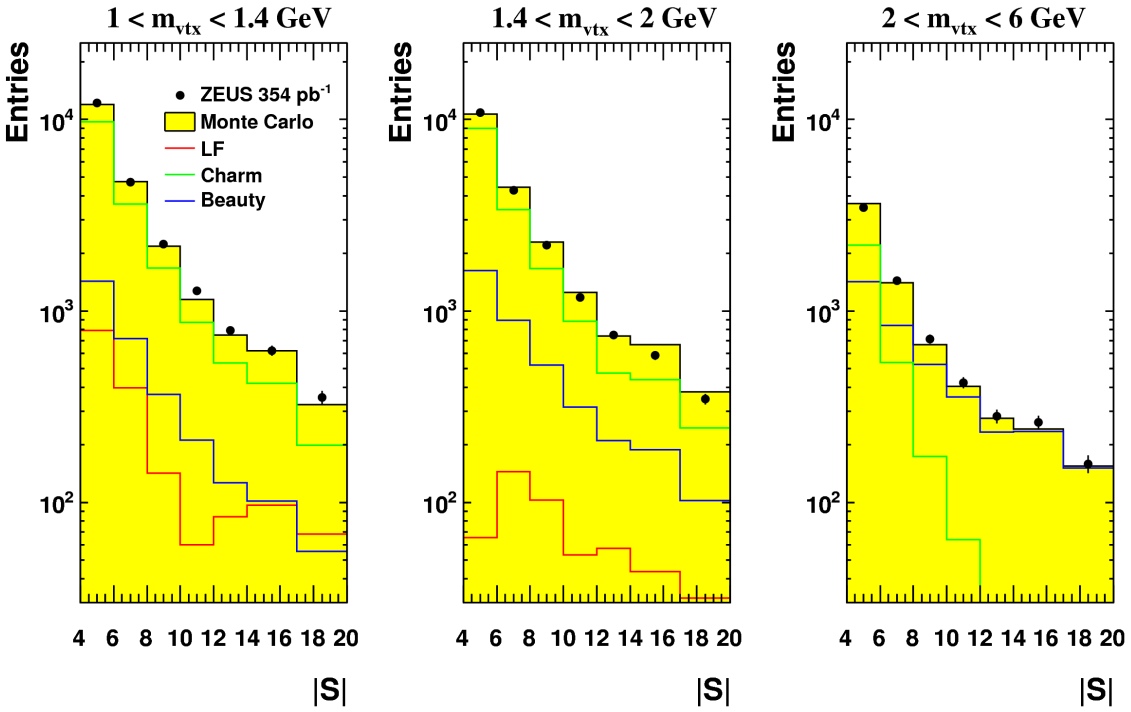


Figure 5.6: Negative subtracted significance in three mass bins. Other details are as in Fig. 5.2.

fit [14], which is unphysical.

The χ^2 of the fit is defined as follows:

$$\chi^2 = \sum_i \frac{(N_i^{\rm d} - k_b\, N_i^b - k_c\, N_i^c - k_{uds}\, N_i^{uds})^2}{(\sigma_i^{\rm d})^2 + k_b^2(\sigma_i^b)^2 + k_c^2(\sigma_i^c)^2 + k_{uds}^2(\sigma_i^{uds})^2} + \frac{(N_{\rm tot}^{\rm d} - k_b\, N_{\rm tot}^b - k_c\, N_{\rm tot}^c - k_{uds}\, N_{\rm tot}^{uds})^2}{(\sigma_{\rm tot}^{\rm d})^2 + k_b^2(\sigma_{tot}^b)^2 + k_c^2(\sigma_{\rm tot}^c)^2 + k_{uds}^2(\sigma_{\rm tot}^{uds})^2}, \tag{5.1}$$

where the sum is running over all bins in the mirrored distribution; $N_i^{\rm d}$, N_i^b, N_i^c and N_i^{uds} are the number of entries in data, beauty, charm and light flavour MC in the bin i; $\sigma_i^{\rm d}$, σ_i^b, σ_i^c, σ_i^{uds} are their uncertainties; $N_{\rm tot}^{\rm d}$, $N_{\rm tot}^b$, $N_{\rm tot}^c$ and $N_{\rm tot}^{uds}$ are the total number of entries in the unmirrored distribution (without splitting into mass bins) for data, beauty, charm and light flavour MC and $\sigma_{\rm tot}^{\rm d}$, $\sigma_{\rm tot}^b$, $\sigma_{\rm tot}^c$, $\sigma_{\rm tot}^{uds}$ are their uncertainties. The fit yields values of the scaling factors k_b, k_c and k_{uds} which have to be applied to the beauty, charm and light flavour MC respectively, in order to give the best description of data. The results from the inclusive fit are as follows:

- $k_b = 1.32 \pm 0.05$ (4.1 % relative uncertainty);
- $k_c = 0.940 \pm 0.014$ (1.5 % relative uncertainty);
- $k_{uds} = 0.915 \pm 0.005$ (0.5 % relative uncertainty);
- fit quality: $\chi^2/n_{\mathrm{dof}} = 21.9/19$=1.15;
- correlation coefficients: $\rho_{b,\,uds} = 0.580$, $\rho_{c,\,uds} = -0.983$, $\rho_{b,\,c} = -0.678$.

The MC contributions shown in Fig. 5.6 were scaled by these factors. A good description is observed, as also suggested by the χ^2/n_{dof}. It also follows from the figure that the first and the second mass bins are dominated by charm production, while the third mass bin at $|S| > 8$ is dominated by beauty. This suggests a procedure of obtaining charm and beauty enhanced regions: one has to mirror the distributions and to require additionally:

- Charm: $|S| > 4$ and $1 < m_{\mathrm{vtx}} < 2\,\mathrm{GeV}$;
- Beauty: $|S| > 8$ and $2 < m_{\mathrm{vtx}} < 6\,\mathrm{GeV}$.

Figures in Appendix A show mirrored significance distributions for different kinematic regions (low, mid and high Q^2; backward, central and forward directions). Good description is observed. The beauty fraction increases from low to high Q^2. This is expected, since at low Q^2 beauty production is kinematically suppressed compared to charm due to larger beauty mass.

Cross section definition

The differential cross section of jet production in events with charm or beauty quarks as a function of a certain variable X is given by the following formula:

$$\frac{d\sigma^{c,b}(X)}{dX} = \frac{k_i^{c,b}\, N_i^{c,b}}{\Delta X_i L},$$

where $k_i^{c,b}$ is the charm or beauty scaling factor in the bin i, $N_i^{c,b}$ is the number of jets from the MC obtained by running the jet algorithm on the true level hadrons before decays of charm and beauty hadrons, ΔX_i is the width of the bin and L is the total integrated luminosity.

5.6 Corrections

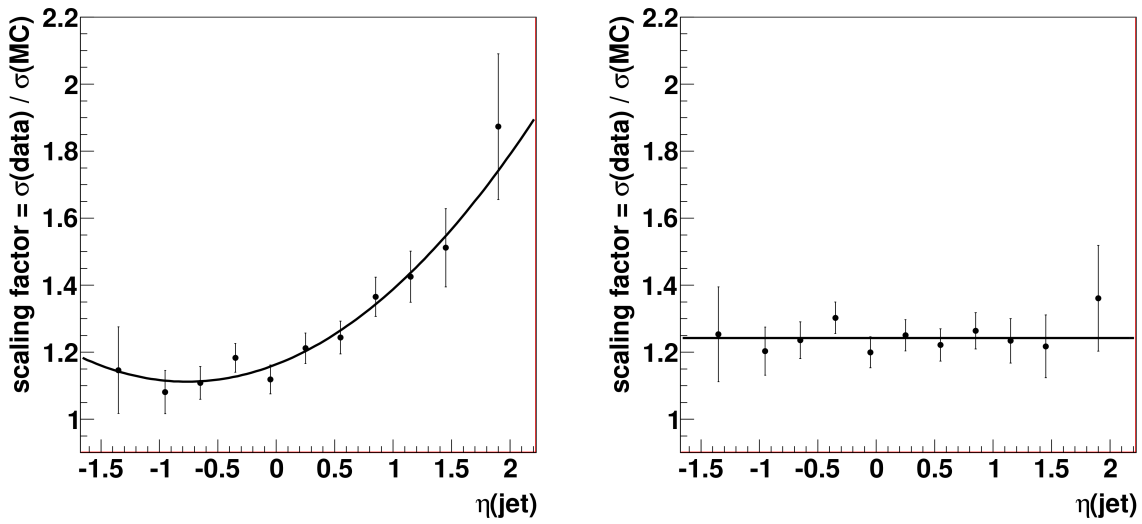
A reliable simulation of the underlying physics as well as of the detector response by the Monte Carlo is essential for a proper cross section measurement, since these effects

directly influence the shapes of signal and background templates and the acceptance determination. This section describes various corrections that were applied to the MC prior to the signal extraction procedure discussed in Section 5.5. They ensure that the MC describes the data.

Reweighting

The acceptance depends strongly on the jet pesudorapidity η^{jet} and on the jet transverse energy E_T^{jet}, hence a good modelling of these variables is necessary. However, discrepancies between data and MC were observed for these distributions for charm. Fig. 5.7(a) shows the charm scaling factors k_c obtained from the fits in different bins of the jet pseudorapidity (see Section 5.5). They represent the ratio of the measured cross sections and the MC cross sections; the fact that the distribution is not constant as a function of η^{jet} indicates that the η^{jet}-distribution is not well modelled in the Monte Carlo sumilation. In order to correct for this effect, a reweighting is applied. The corresponding reweighting function was determined by a second-order polynomial fit to this distribution. The following parametrisation was obtained[6]:

$$0.937 + 0.109\,\eta^{\text{jet}} + 0.0715\,(\eta^{\text{jet}})^2. \tag{5.2}$$



(a) (b)

Figure 5.7: Charm scaling factors as a function of jet pseudorapidity before (a) and after (b) η^{jet}-reweighting. Solid lines show a second-order polynomial (a) and a constant (b) fits.

[6]The function was normalised by a factor of 1.24. Total normalisation of reweighting function does not affect the results.

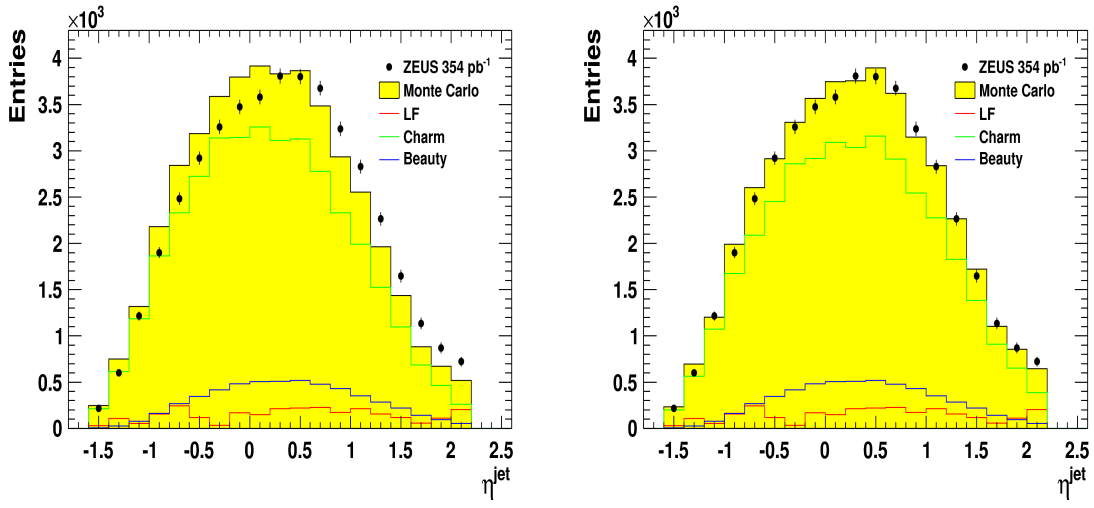


Figure 5.8: Jet pseudorapidity distribution in the charm enhanced region before (a) and after (b) η^{jet}-reweighting. Other details are as in the Fig. 5.2.

After it is applied to the MC, the scaling factors become flat as can be seen from Fig. 5.7(b). A zero-order polynomial is fitted to this distribution in order to demonstrate consistency with a constant. Fig. 5.8 shows the jet pseudorapidity distribution in the charm enhanced region (see Section 5.5 for charm enhanced region definition) before and after the reweighting. One concludes that the description is improved significantly. The application of the η^{jet}-reweighting leads to an increase of the total charm cross section by around 5 % which is significant.

A similar procedure was applied for the jet transverse energy E_T^{jet}. The resulting reweighting function parametrisation determined from the scaling factors as a function of E_T^{jet} is as follows:

$$0.753 + 0.189\sqrt{E_T^{\text{jet}}}, \tag{5.3}$$

where E_T^{jet} is measured in GeV. The application of this E_T^{jet}-reweighting leads to a decrease of the total charm cross section by around 3 %.

The Q^2 distribution for both charm and beauty is not well described by the RAPGAP simulation as well. The corresponding reweighting was determined in [14]:

- Charm: $\exp(-0.486 - 0.0158\,Q^2) + 0.781$;
- Beauty: $\exp(-0.599 - 0.00389\,Q^2) + 0.631$.

The effect of the reweighting is around +4 % and +10 % on the total charm and beauty beauty cross section.

Additionally, the light flavour MC was reweighted in E_T^{jet}, since it overshoots the data significantly at high values of E_T^{jet}. The weighting factors were determined on a bin-by-bin basis from the inclusive E_T^{jet} distribution. The effect on the measured charm and beauty cross sections turned out to be negligible.

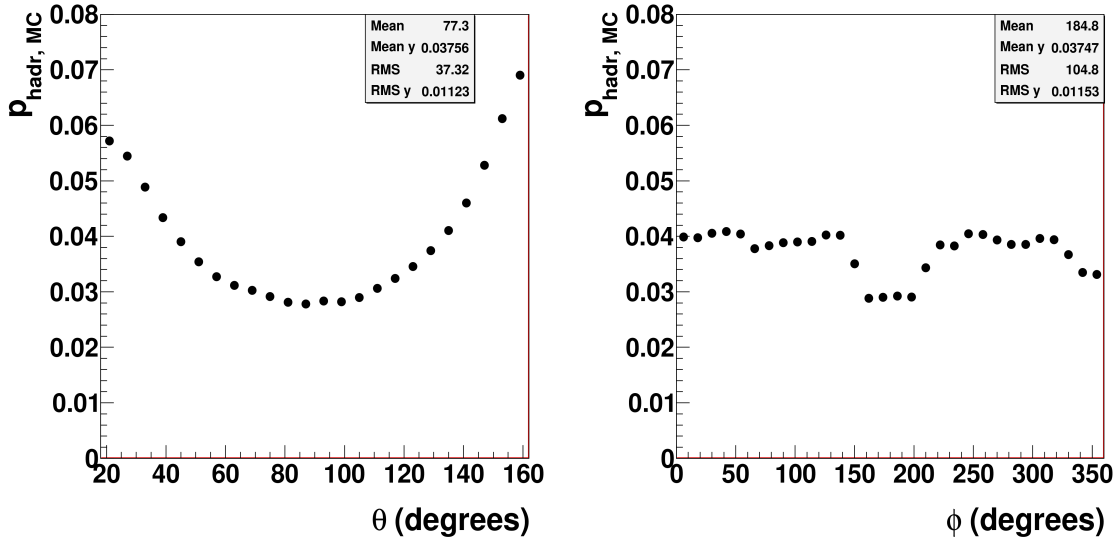
Tracking efficiency correction

The quality of the tracking efficiency modelling by the MC directly affects the cross section measurement. The loss of tracks due to hadronic interactions is one of the dominant contributions to the tracking inefficiency. A dedicated study of this effect was done in this thesis, see Chapter 4. It was found that on average the MC underestimates the probability of track loss due to hadronic interactions p_{hadr}, which leads to an overestimation of the total tracking efficiency. In turn, the measured cross section is underestimated.

In order to take this into account, a correction on the interaction probability in MC has to be applied. The correction factor is defined in the following way:

$$(\epsilon - 1)\frac{p_{\text{hadr, MC}}}{1 - p_{\text{hadr, MC}}}, \tag{5.4}$$

where $p_{\text{hadr, MC}}$ is the probability of hadronic interactions in the MC and ϵ is the ratio of the interaction probability between data and MC ($\epsilon = p_{\text{hadr, DATA}}/p_{\text{hadr, MC}}$). This gives the difference between the actual tracking efficiency and the one simulated by the MC.



(a) (b)

Figure 5.9: Average tracking inefficiency due to hadronic interactions in the MC as a function of track θ (a) and track ϕ (b), for tracks selected for forming secondary vertices.

$p_{\mathrm{hadr,\,MC}}$ was determined by evaluating in the MC how often tracks are lost due to the interaction. This quantity was studied differentially as a function of track momentum p and track angles θ and ϕ. Fig. 5.9 shows the distribution of p_{hadr} as a function of track θ and ϕ for the tracks that form vertices after the selection described above. The inefficiency increases towards the forward an backward direction, as expected from the fact that more material is passed. The azimuthal distribution shows a characteristic pattern with a dip around 180° which is due to the MVD geometry: in this region there are only two layers instead of three. For tracks that form the selected vertices the average probability is around 3.8 %. The quantity ϵ was determined in Chapter 4. It is consistent with being a constant as a function of θ and ϕ, while there is a p_T dependence: $\epsilon \approx 1.4$ for $p_T < 1.5\,\mathrm{GeV}$ and $\epsilon \approx 1$ for $p_T < 1.5\,\mathrm{GeV}$.

Based on these figures, a correction of around 1.5 % has to be applied to the tracking efficiency at transverse momenta below 1.5 GeV. It is performed by discarding tracks randomly with a probability that is calculated with equation (5.4). After track rejection, the secondary vertex fit is redone, the cross sections are reevaluated and the correction on the cross sections is determined. It amounts to +1.1 % and +3.4 % for the total charm and beauty cross section, respectively.

Significance smearing

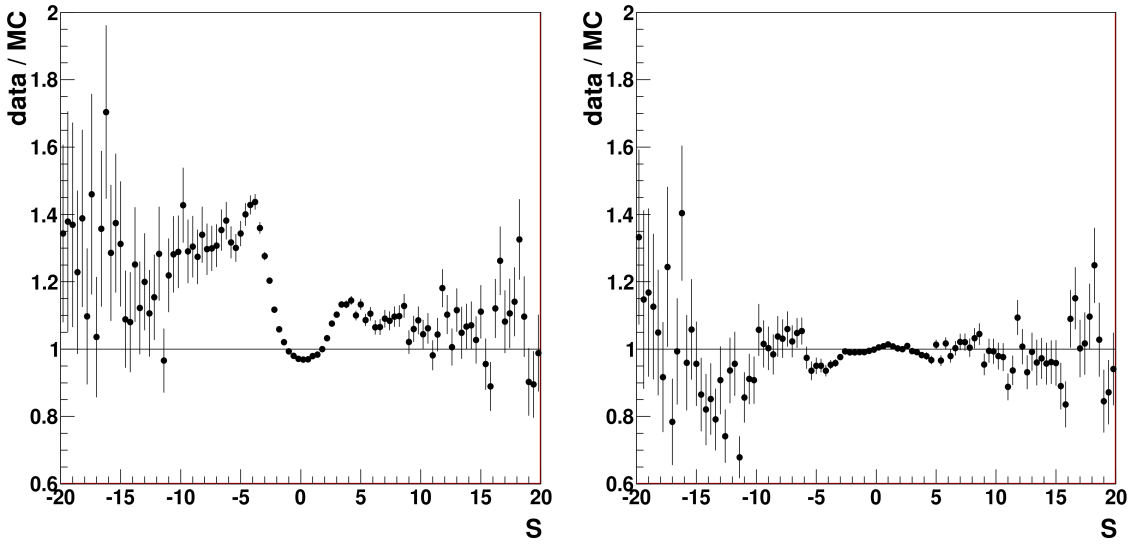
As discussed in Section 5.5, the projected decay length significance is used as a discriminating variable for the signal extraction. Hence its accurate modelling by the Monte Carlo is essential. However, it was found not to be well reproduced by the MC. This readily follows from the Fig. 5.10(a) which shows the data-to-MC ratio for the significance distribution. One of the reasons for this discrepancy might be an improper simulation of the ZEUS tracking system. In order to improve the description, the significance distribution is *smeared* in the Monte Carlo simulation: for a certain fraction of randomly chosen vertices, the significance is modified according to:

$$S' = S + \alpha,$$

where S and S' are the unsmeared and smeared significance, respectively, and α is a random number distributed according to a gaussian. In particular,

- 7 % of vertices are smeared by a gaussian with a width of $\sigma = 1.6$ and
- 0.2 % of vertices are smeared by a gaussian with a width of $\sigma = 7$.

The first gaussian corrects the core part of the distribution, while the second one modifies the tails. The parameters were tuned based on the $S < 0$ part of the significance distribution which is dominated by detector resolution. The quality of the description after application of this procedure was checked differentially as a function of various kinematic variables. While in most cases the description was reasonable, some residual discrepancies



(a) before smearing (b) after smearing

Figure 5.10: Data-to-MC ratio for the projected decay length significance distribution.

were observed in the forward region. It was found, that they are due to the fact that the projected decay length L_{XY} distribution in the MC is shifted as a whole with respect to the data. In order to correct for this, a shift of the decay length was applied in the MC in this η^{jet}-region:

$$L'_{XY} = L_{XY} + \Delta,$$

where Δ has values shown in Table 5.4. Additionally, the fraction of smeared vertices was slightly modified for the region of $1.6 < \eta^{\text{jet}} < 2.2$, which leads to a better description in that region: 9 % of vertices were smeared with a narrow gaussian and 0.6 % of vertices were smeared with a wide gaussian (instead of 7 % and 0.2 %, respectively, as in other regions).

η^{jet}-range	Δ, cm
[1.6, 2.2]	−0.00219
[1.3, 1.6]	−0.00158
[1.0, 1.3]	−0.000766

Table 5.4: Shifts of the projected decay length in the Monte Carlo.

Fig. 5.10(b) shows the data-to-MC ratio for the significance distribution after application of smearing and shifts. The description is improved dramatically after the procedure. The effect on the total charm cross section is −1.5 % while that on beauty +1 %. The residual discrepancy between data and MC is taken into account in the systematic uncertainty (see Section 6).

Branching ratios and fragmentation fractions reweighting

The values of branching ratios of charm hadron decays as well as fragmentation fractions of charm quarks were found to be outdated in the standard ZEUS simulation package. A dedicated investigation of this was performed in Ref. [130]. A C++ routine was developed, which corrects fragmentation fractions and branching ratios to the world average values [131,132] by performing a reweighting. It removes a discrepancy between data and MC in the vertex mass spectrum for the charm enhanced region [130] which is particularly sensitive to the final states composition in charm hadrons decays. The effect of the correction on the total charm cross sections is $-5\,\%$, while on the beauty it is $-9\,\%$.

5.7 Control Distributions

In this section various control plots are presented after the final selection and corrections[7]. Figures 5.11(a)-(f) show the distributions of event kinematic variables Q^2, x, y, $E - p_Z$ as well as of the Z-coordinate of the primary vertex. The distributions are dominated by the light flavour background contribution. The charm contribution is also significant, while the beauty component is very small. All distributions are described reasonably. In the $E - p_Z$ distribution, the MC is shifted with respect to the data, which is taken into account in the systematics. Figures 5.11(g)-(i) show distributions of the scattered electron quantities which are important for the reconstruction of event kinematics. A reasonable agreement is found. Figures 5.11(j)-(l) show the jet kinematic variables. Again, a reasonable description is observed. The slight discrepancy in $\eta^{\rm jet}$ comes mainly from the light flavour, since it does not appear in the $\eta^{\rm jet}$-distribution for the charm enhanced region (Fig. 5.12(h)). It is not expected to affect the results, since the light flavour contribution is significantly suppressed after the mirroring[8].

Figure 5.12 shows similar distributions in the charm enriched region which was obtained with the procedure described in the end of Section 5.5, i.e. requiring $1 < m_{\rm vtx} < 2\,{\rm GeV}$, $|S| > 4$ and mirroring of the distributions. The distributions are indeed dominated by charm production. Figures 5.12 (a)-(f) show event kinematics variables Q^2, x, y, $E - p_Z$ as well as the Z-coordinate of the primary vertex. A good agreement is observed. A shift of the data with respect to the MC can be seen in the $E - p_Z$ distribution, similarly to the inclusive case (Fig. 5.11(e)). Figures 5.12 (g)-(i) show the jet variables. Excellent agreement is observed, as expected after the reweighting procedures in $E_T^{\rm jet}$ and $\eta^{\rm jet}$. Secondary vertex variables are shown in Figures 5.12 (j)-(l). For Figures 5.12 (j) and (l), the $m_{\rm vtx}$ and $\chi^2/n_{\rm dof}$ cuts, respectively, were not applied. Again, a good description is found. The discrepancy seen in the $\chi^2/n_{\rm dof}$ distribution for the inclusive case (Fig. 5.2)

[7] The tracking efficiency correction was not applied for these plots.

[8] The effect on the cross sections due to the reweighting of light flavour in $E_T^{\rm jet}$ was negligible as discusssed in Section 5.6.

is not observed for the charm enriched region (Fig. 5.12 (l)), hence the problem is related mostly to the light flavour modelling.

Analogously, Figure 5.13 shows control plots for the beauty enhanced sample, obtained with the mirroring and requiring $|S| > 8$ as well as $2 < m_{\rm vtx} < 6\,{\rm GeV}$. The distributions demonstrate that this region is indeed dominated by the beauty production. Overall, a good agreement between data and MC is found. However, the statistics is smaller compared to charm or inclusive cases, hence it is more difficult to spot small discrepancies, if any. Figures 5.13(a)-(f) show the event kinematic variables. Shapes of the Q^2 and y is somewhat different compared to charm, due to the fact that the topology of beauty events is somewhat different. Figures 5.13 (g)-(i) show the jet variables. Again, a good agreement is observed. A small shift of the MC towards lower values of $\eta^{\rm jet}$ is seen in the Fig. 5.13(h) which is, however, not significant. It was checked in Ref. [14], that reweighting of this variable does not lead to a significant change in the cross sections. Figures 5.13 (j)-(l) show the vertex quantities. Again, a good description is provided by the MC. The track multiplicity for vertices is on average higher than for charm, as expected. Due to this fact, there are more degrees of freedom in the fit and the $\chi^2/n_{\rm dof}$ has a different shape.

Figures 5.14, 5.15 and 5.16 show the event displays for light flavour, charm and beauty event candidates, respectively. The light flavour event candidate was selected from the region of $1 < m_{\rm vtx} < 1.4\,{\rm GeV}$ and $|S| < 4$; the charm event candidate satisfies $1 < m_{\rm vtx} < 2\,{\rm GeV}$ and $S > 4$, while the beauty event candidate was required to fullfill $2 < m_{\rm vtx} < 6\,{\rm GeV}$ and $S > 8$. As expected, for the light flavour event, the reconstructed secondary vertex is consistent with the primary; for the charm event, it is separated from the primary vertex, while for beauty the separation improves further, thanks to higher track multiplicity and longer lifetime of the ground state beauty hadron.

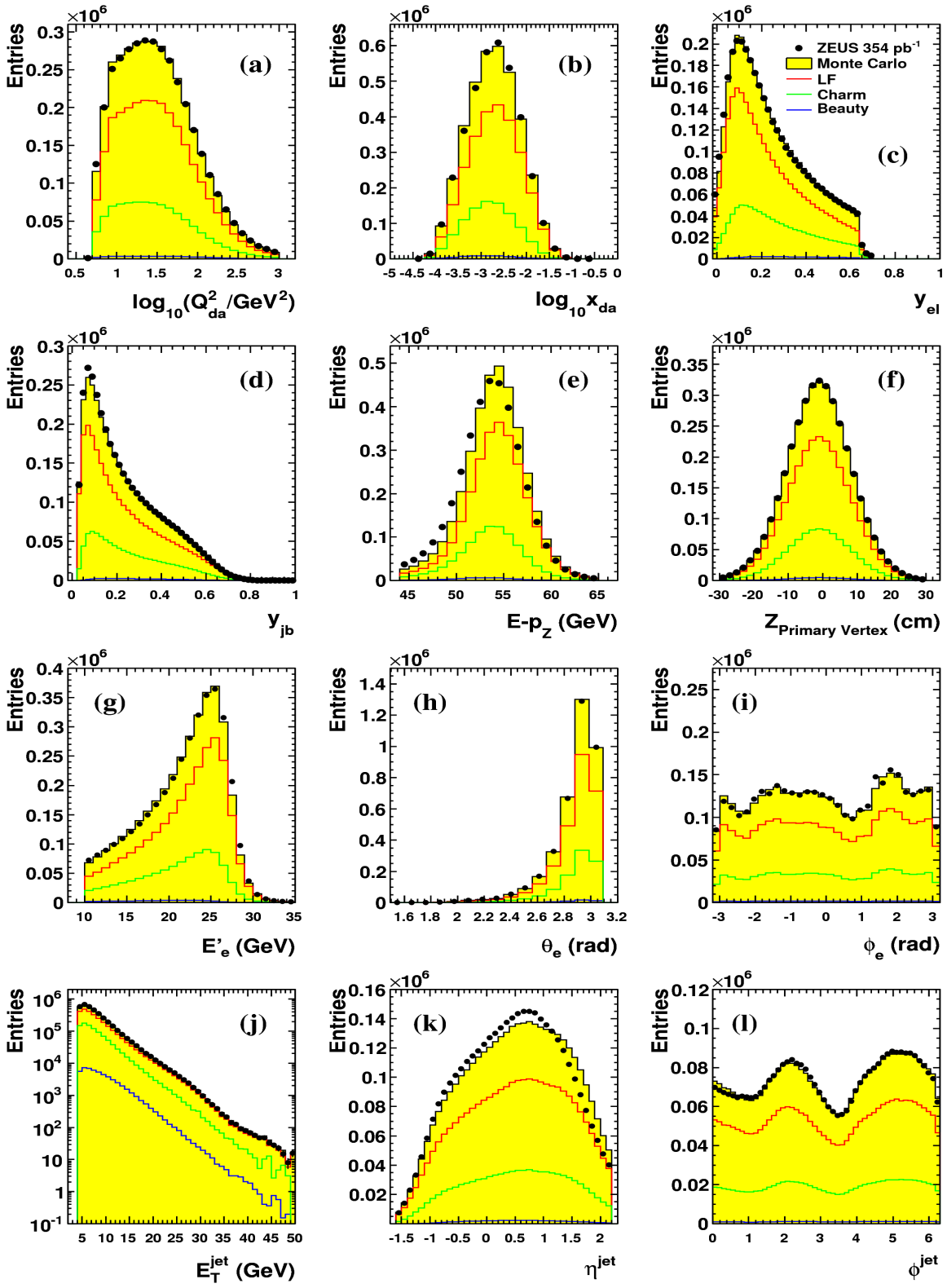


Figure 5.11: Control plots for event kinematics (a)-(f), electron kinematics (g)-(i) and jet kinematics (j)-(l). Other details are as in the Fig. 5.2.

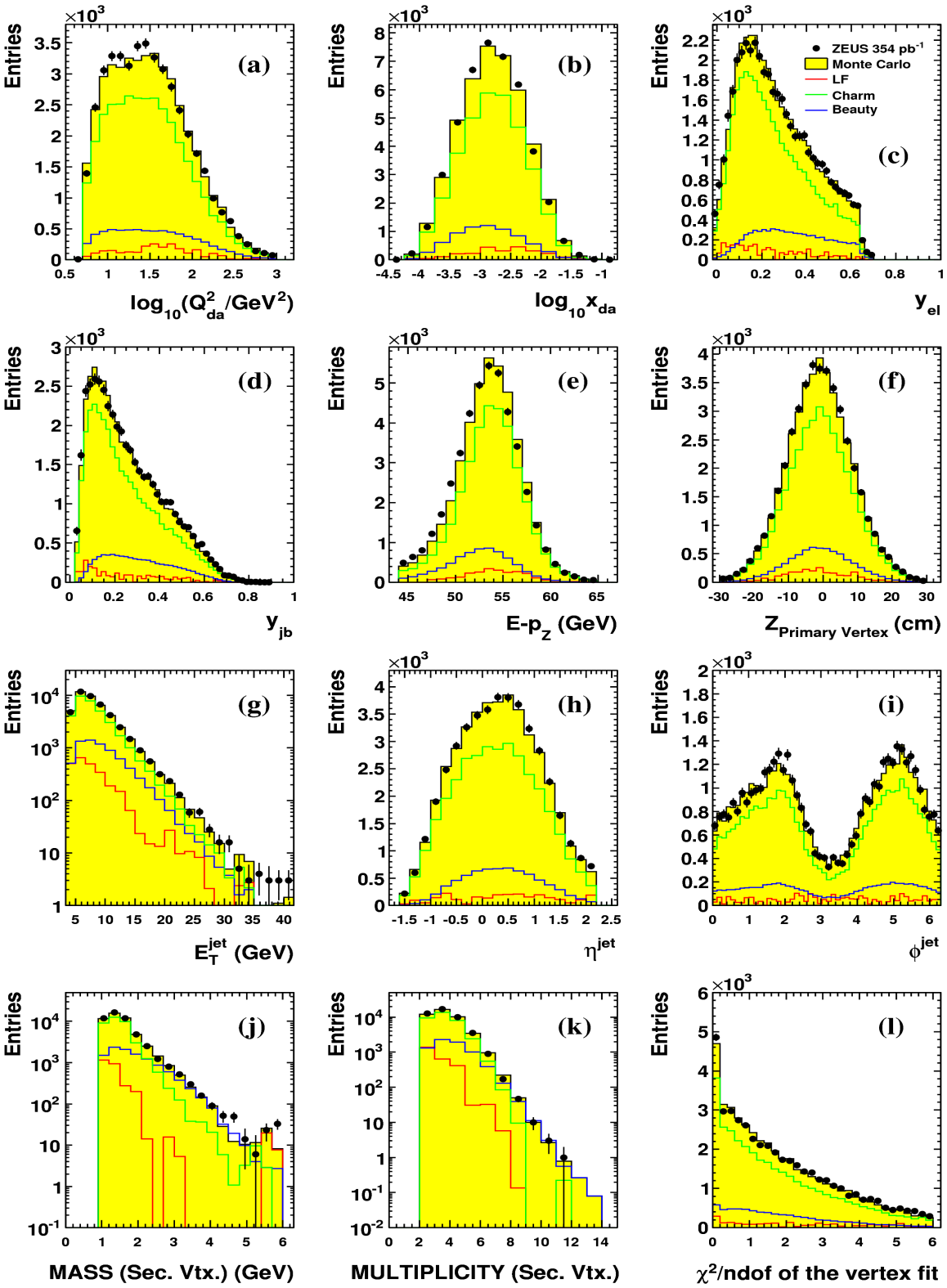


Figure 5.12: Control plots for the charm enriched region: event kinematics (a)-(f), jet kinematics (g)-(i) and vertex properties (j)-(l). Other details are as in the Fig. 5.2.

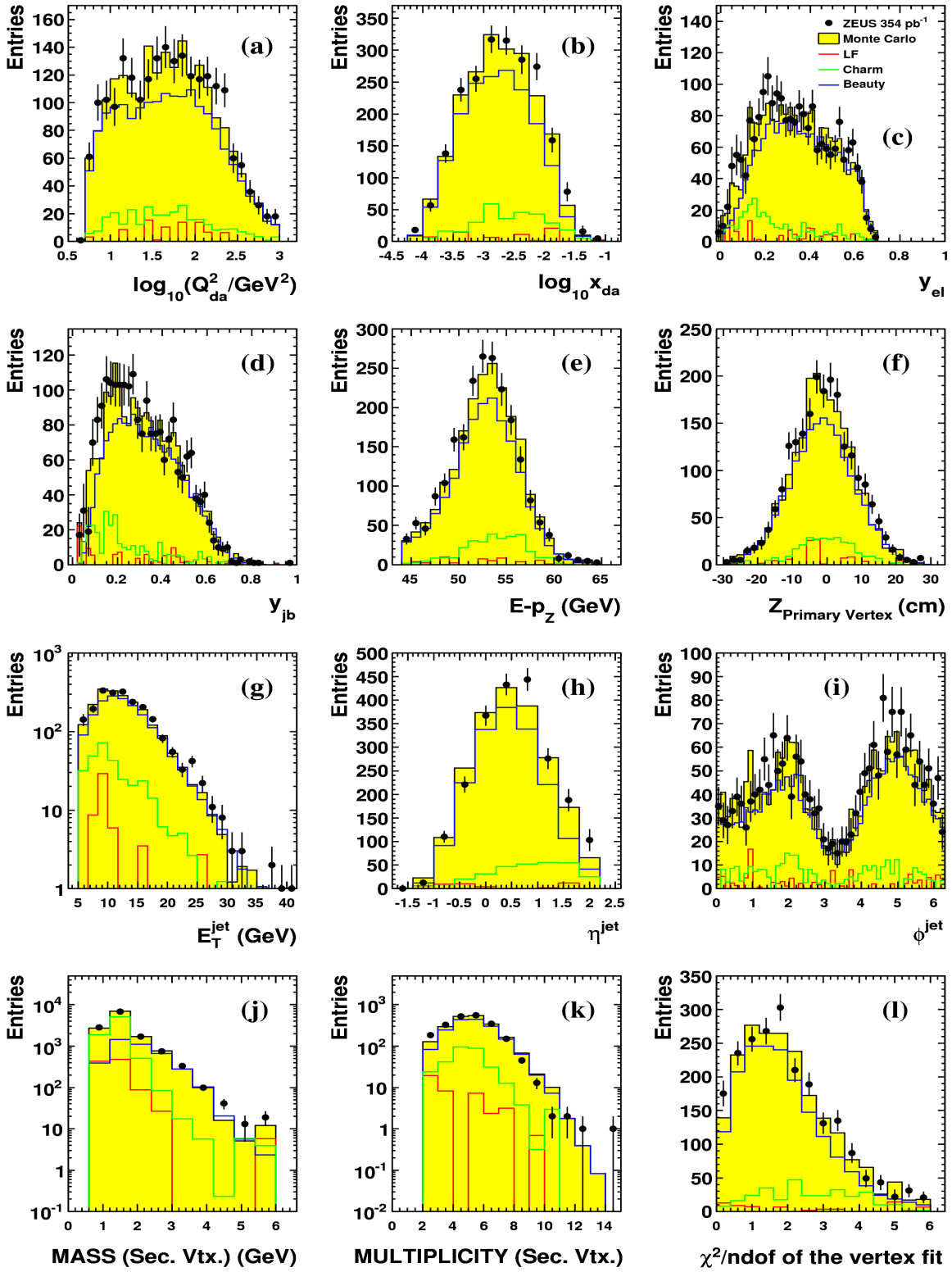


Figure 5.13: Control plots for the beauty enriched region: event kinematics (a)-(f), jet kinematics (g)-(i) and vertex properties (j)-(l). Other details are as in the Fig. 5.2.

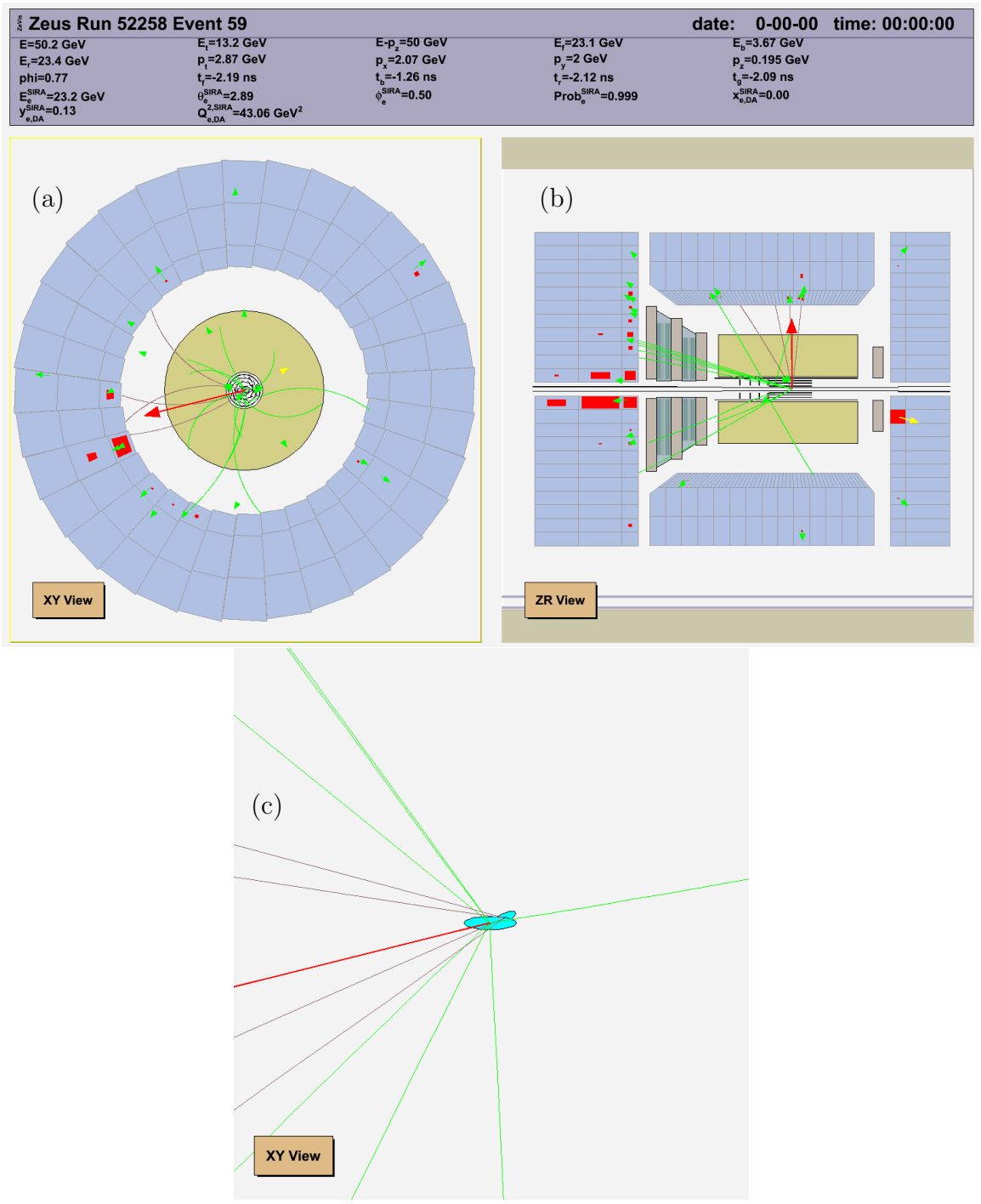


Figure 5.14: A light flavour background event candidate in the XY view (a), the ZR view (b) and a zoomed XY view (c). Energy deposits in the calorimeter are shown by red boxes with the box size representing the deposited energy. Green arrows show the Energy Flow Objects (see Section 3.4). The yellow arrow shows the scattered electron. The red arrow shows the reconstructed jet. Grey lines are tracks associated to the secondary vertex, green lines are other tracks. The bigger ellipse represents the beam spot, the smaller ellipse shows the position and the uncertainty of the secondary vertex.

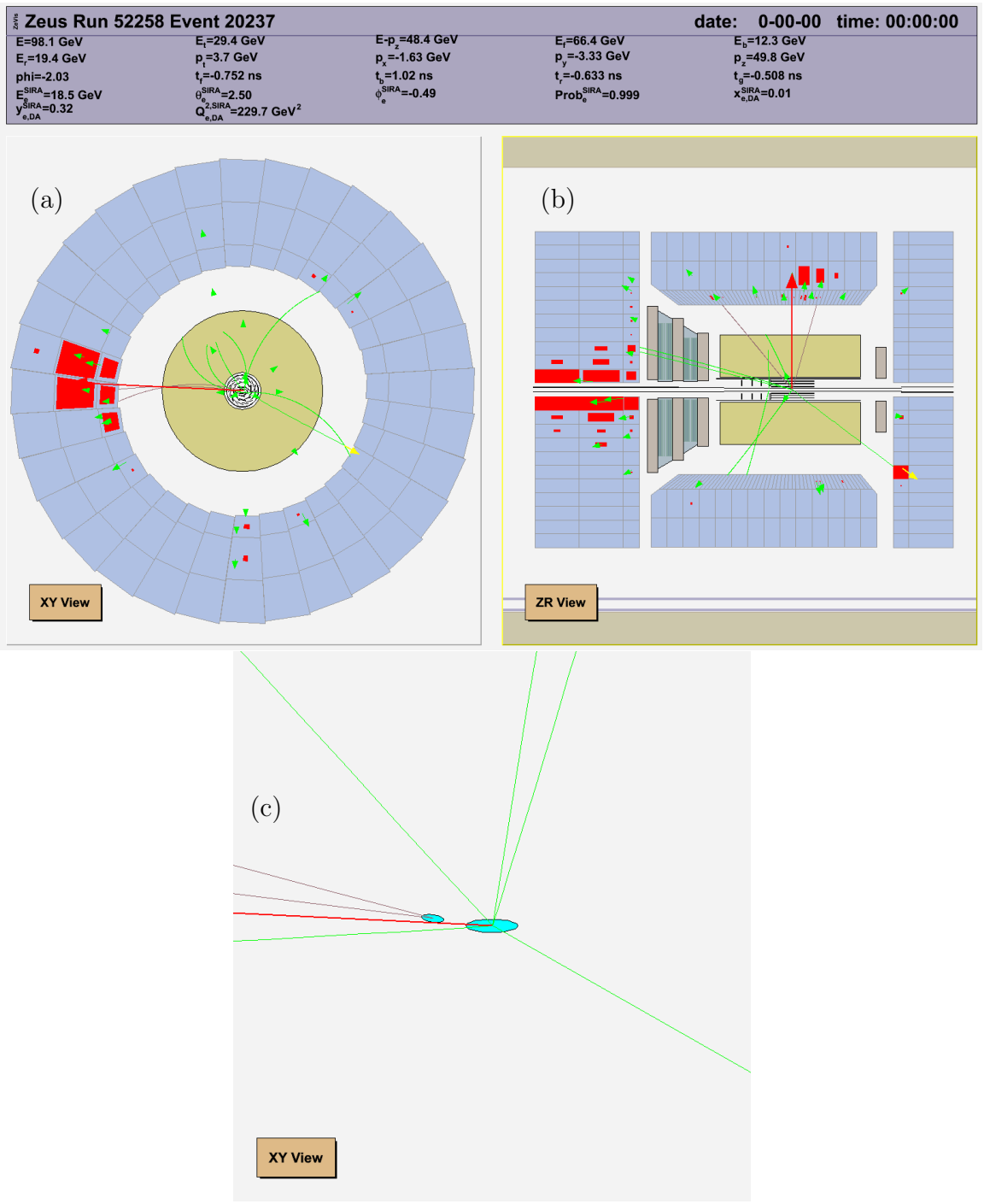


Figure 5.15: A charm event candidate in the XY view (a), the ZR view (b) and a zoomed XY view (c). Energy deposits in the calorimeter are shown by red boxes with the box size representing the deposited energy. Green arrows show the Energy Flow Objects (see Section 3.4). The yellow arrow shows the scattered electron. The red arrow shows the reconstructed jet. Grey lines are tracks associated to the secondary vertex, green lines are other tracks. The bigger ellipse represents the beam spot, the smaller ellipse shows the position and the uncertainty of the secondary vertex.

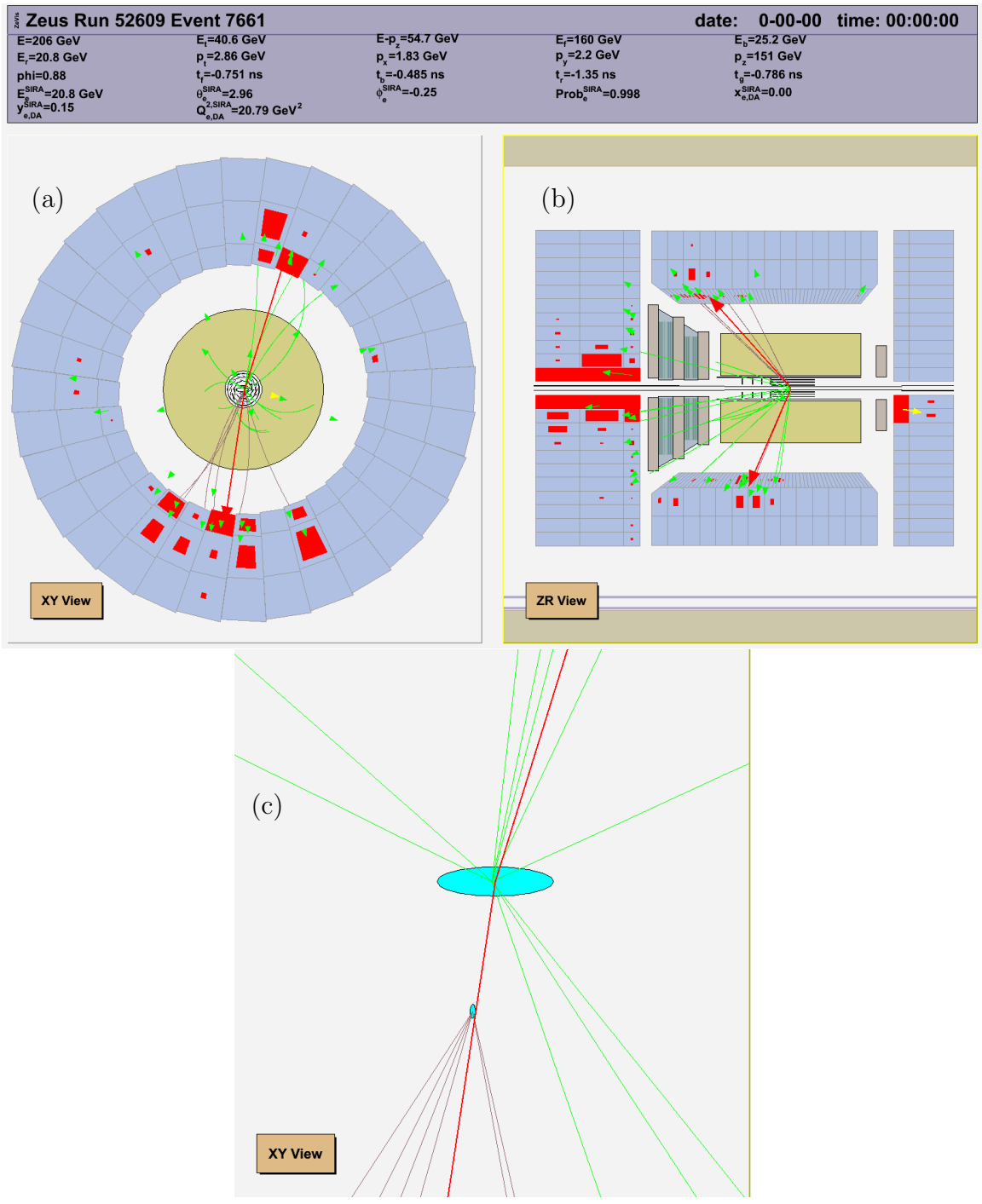


Figure 5.16: A beauty event candidate in the XY view (a), the ZR view (b) and a zoomed XY view (c). Energy deposits in the calorimeter are shown by red boxes with the box size representing the deposited energy. Green arrows show the Energy Flow Objects (see Section 3.4). The yellow arrow shows the scattered electron. The red arrow shows the reconstructed jet. Grey lines are tracks associated to the secondary vertex, green lines are other tracks. The bigger ellipse represents the beam spot, the smaller ellipse shows the position and the uncertainty of the secondary vertex.

Chapter 6

Systematic Uncertainties

Every physics measurement suffers from a number of systematic uncertainties. One of the most important tasks for an experimentalist is to identify them and to quantify the influence of these effects on the physics results. A detailed investigation of systematic uncertainties for the charm and beauty measurement presented in this thesis was performed and is discussed in this chapter. First, the general method that was used for determination of uncertainties is described. Afterwards, all identified sources of uncertainty are discussed.

6.1 The Scan Technique

Often, in order to evaluate a systematic uncertainty, a certain parameter or procedure is modified, and the result is reextracted. The difference Δ between the default result and the one obtained with an altered parameter is called a systematic uncertainty. However this approach is not protected from statistical fluctuations. For the measurement presented in this thesis, template fits have to be redone each time the analysis procedure is changed. Variations of certain parameters cause event migration from and to neighboring bins; due to the fact that MC templates are statistically limited, this introduces fluctuations to the fit results, which alters the difference Δ. Hence the obtained systematic uncertainty contains a statistical component and differs from the actual systematic effect.

In order to overcome this problem and hence to achieve a better systematic uncertainty determination, a *scan technique* was used in this thesis. The idea of the method is that the relevant parameter is not changed only once, but in steps in a certain range. For each variation, the template fit is redone and results are reextracted. The corresponding dependence of cross sections on the parameter represents the systematic effect. For small parameter variations, the resulting dependence is expected to be linear and can be approximated with a first-order polynomial. In this case, a systematic effect is related to the slope of the line. From the parameter scan one can also check if there is any

significant nonlinear dependence present in the range of the uncertainty of the parameter which would lead to asymmetric uncertainties. With the scan technique it is possible to see directly whether there is a systematic trend and to reduce the influence of statistical fluctuations on the resulting systematic uncertainty.

The method is illustrated in Fig. 6.1, for the case of the tracking efficiency systematics as an example (see Section 5.6 for description of the tracking efficiency correction and the next section for details on parameter variations for the systematic uncertainty determination). The horizontal axis denotes the value of a relevant parameter α that has to be varied (in this case, the variation of a parameter ϵ which enters the probability of track discarding in the MC). The vertical axis denotes the total charm cross section. The default value of the parameter is denoted by the vertical dashed line and corresponds to $\alpha = 0$. The dot-dashed line shows the upward variation of this parameter at which systematics has to be determined ($\alpha = 0.2$). Several variations of α are performed, and

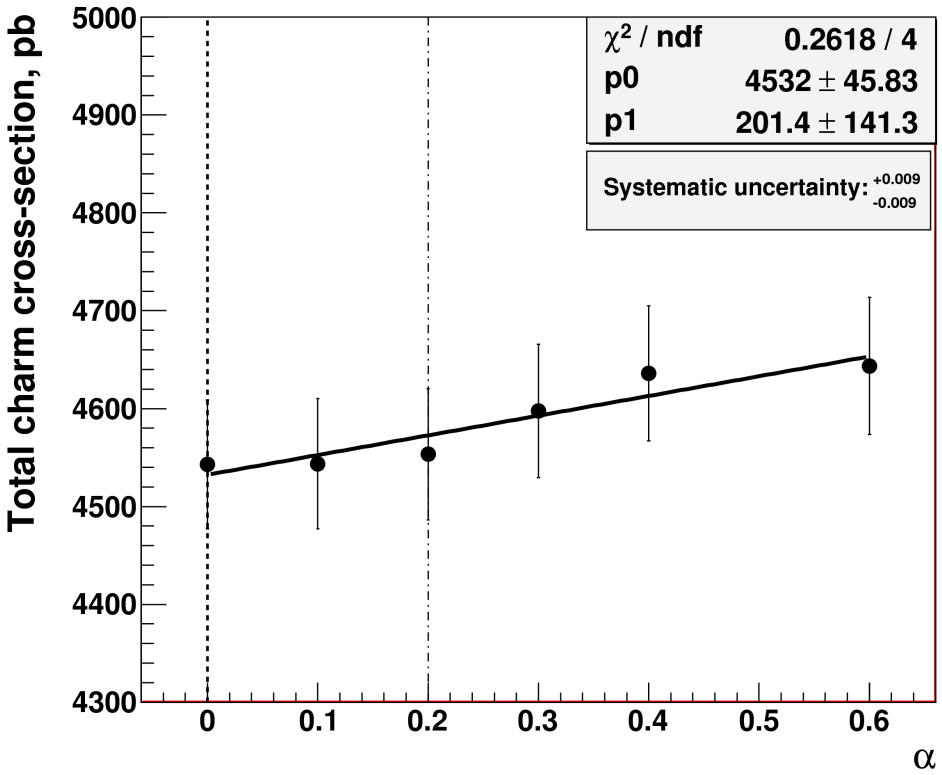


Figure 6.1: Illustration of the scanning technique for the systematic uncertainty determination. Points show the total charm cross section; only statistical uncertainties are indicated. The dashed line shows the default value of a parameter, while the dot-dashed line represents its variation at which the systematic uncertainty has to be determined. The solid line shows a straight line fit.

for every value of α, the cross section is reevaluated. One concludes that there is a systematic increase of the cross section with the increase of α. The dependence is fitted with a first-order polynomial. Due to the fact that the data points are highly correlated, the resulting $\chi^2/n_{\mathrm{dof}} = 0.2618/4$ is much smaller than unity. The uncertainties on the fitted parameters are therefore meaningless. The slope of the fitted line corresponds to the systematic uncertainty per unit of parameter variation. By multiplying the slope by the difference between the default parameter and the variation ($\Delta\alpha = 0.2$), the systematic uncertainty is determined. It follows from Fig. 6.1 that the determination of the systematic uncertainty with a single variation (i.e. the difference Δ between the cross sections at $\alpha = 0.2$ and $\alpha = 0$) would significantly underestimate it.

In this way, uncertainties were determined for each identified source. They were evaluated separately for total charm and beauty cross sections as well as for every bin in all differential cross sections and added quadratically to obtain the total systematic uncertainty, separately for positive and negative variations.

6.2 Sources of Uncertainty

Various effects that influence the charm and beauty measurement presented in the thesis have been identified and studied. This section describes these effects as well as the procedures to evaluate the corresponding systematic uncertainties.

Tracking efficiency

The tracking efficiency directly affects the acceptance, hence its modelling by the Monte Carlo is essential for the cross section determination. It was studied in detail in this thesis, in particular it was found that the tracking efficiency is on average overestimated in the MC, which leads to an underestimation of the measured cross sections, see Chapter 4. The corresponding correction is described in Section 5.6. It is performed by randomly rejecting tracks in the MC according to a probability, which depends on the data-to-MC ratio of the hadronic interaction rate, ϵ, and the tracking inefficiency due to hadronic interactions in the MC, $p_{\mathrm{hadr,\,MC}}$, see equation (5.4).

The ϵ parameter is known with moderate precision; in order to evaluate the resulting systematics, it is varied within its uncertainties and the procedure of rejecting tracks is repeated. According to the results of Chapter 4, there is a dependence of ϵ on the transverse track momentum p_T, while there is no significant dependence on other track parameters: $\epsilon \approx 1.4$ for $p_T < 1.5\,\mathrm{GeV}$ and $\epsilon \approx 1$ for $p_T > 1.5\,\mathrm{GeV}$. An uncertainty of 0.2 was assigned to the ϵ parameter in both momentum regions. Hence, in order to evaluate the systematic uncertainty, the ϵ parameter is varied simultaneously to 1.2 and 1.6 at low momenta, and to 0.8 and 1.2 at high p_T.

With the method of track rejection, it is not possible to perform the correction if $\epsilon < 1$,

since this means that the tracking efficiency is underestimated in the MC, and one would need to add tracks to the MC rather than reject, which cannot be easily done. Hence, one-sided (positive) variations are performed and the resulting systematic uncertainty is symmetrised. Simultaneous variations in both momenta regions are performed which leads to a more conservative uncertainty (since the systematic effect is maximised):

$$\epsilon = \begin{cases} 1.4 + \alpha, & p_T < 1.5\,\text{GeV}; \\ 1 + \alpha, & p_T > 1.5\,\text{GeV}, \end{cases}$$

where α is the variation of the ϵ parameter.

Figures 6.1 and 6.2 show the total charm and beauty cross sections, respectively, for different values of α. As discussed in the previous section, the systematic uncertainty is determined from a straight line fit to these distributions. The resulting uncertainty on the total charm cross section is $\pm 0.9\,\%$. The uncertainty on the beauty cross section is $\pm 2.8\,\%$ which is one of its dominant uncertainties.

Prior to this study, the tracking efficiency was not well-known. A systematic uncertainty of $-2\,\%$ was assigned to it, irrespectively of track parameters [11]. The resulting systematic uncertainty was $+5\,\%$ on the total charm cross section and $+8\,\%$ on the

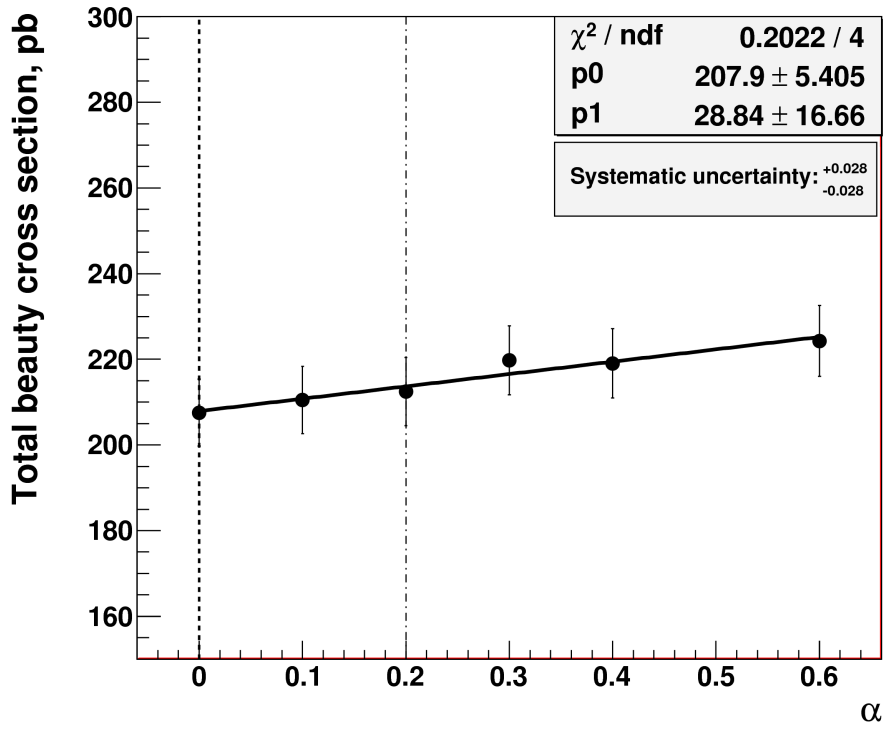


Figure 6.2: Total beauty cross section as a function of the ϵ parameter variation. Other details are as in Fig. 6.1.

beauty [14]. Hence, a substantial reduction of the uncertainty due to tracking efficiency was achieved in this thesis.

Charm Fragmentation Function

Due to the fact that the acceptance for the vertex reconstruction may depend on the momentum of the corresponding charm or beauty hadron, the measured cross section is sensitive to modelling of the fragmentation hardness in the Monte Carlo. This process is parametrised by the charm (beauty) quark longitudinal fragmentation function, which defines the probability to transfer a certain longitudinal momentum fraction from a charm (beauty) quark to a charm (beauty) hadron (see a more detailed discussion in Section 1.5). For MC samples used in this analysis, the Bowler parametrisation [73] was chosen, as described in Section 5.2. In order to evaluate the sensitivity of the results to the fragmentation function, a different parametrisation – the Peterson fragmentation function [74] – was alternatively tried.

Figure 6.3 shows the Bowler and Peterson functions, normalised to unit area. Definitions of the functions are given in Section 1.5. The same settings as for the default MC samples were used to plot the Bowler function, namely: $a_\alpha = a_\beta = 0.3, b = 0.58, r_Q = 1, m = 2.01\,\text{GeV}$, $m_Q = 1.5\,\text{GeV}$, $p_T = 0\,\text{GeV}$. The Peterson parametrisation is controlled by a single parameter, ϵ. It was shown in an analysis of D^*-meson photoproduction by

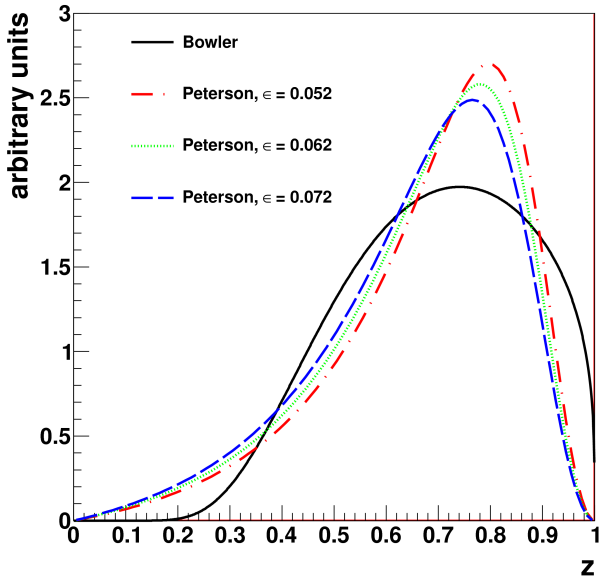
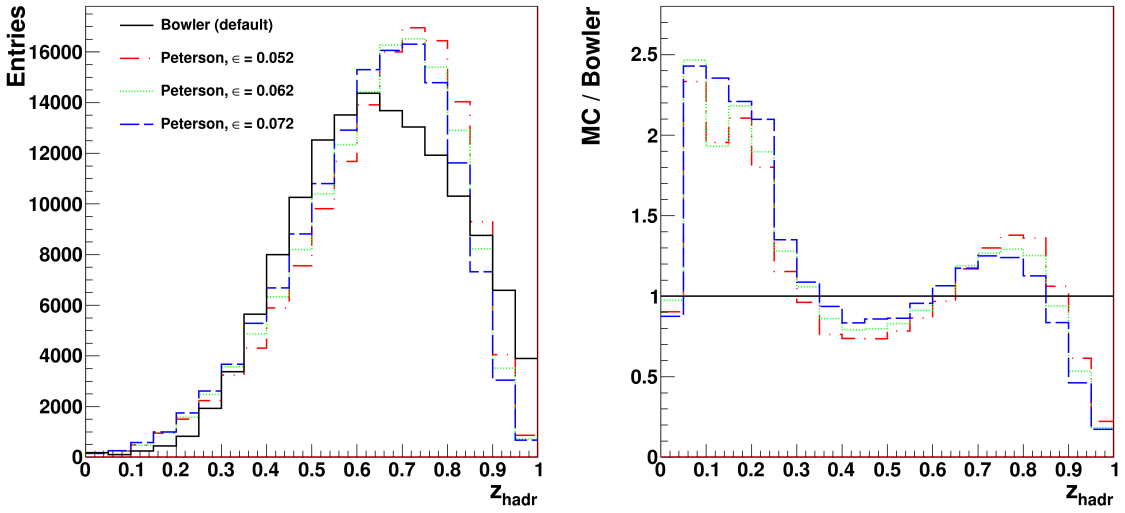


Figure 6.3: Bowler and Peterson fragmentation functions. The solid line shows the Bowler parametrisation (see text for details on the values of parameters), while the dashed, dotted and dot-dashed lines represent Peterson functions with different values of the ϵ parameter.



(a) (b)

Figure 6.4: Distributions of z_{hadr} for MC samples with different fragmentation functions (a) and corresponding reweighting functions (b).

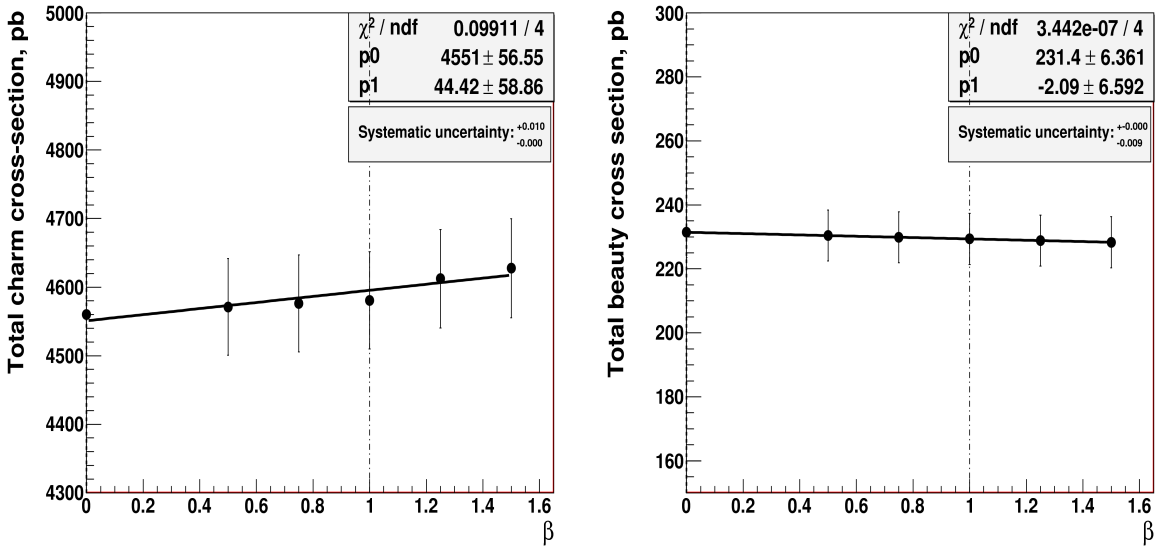
the ZEUS Collaboration [133], that the PYTHIA fragmentation model gives the best description of data for $\epsilon = 0.062 \pm 0.007^{+0.008}_{-0.004}$. In Fig. 6.3, the Peterson function was plotted with three values of ϵ: 0.052, 0.062 and 0.072. Generally Bowler and Peterson functions behave similarly, however Bowler is somewhat wider.

In order to modify the fragmentation function in the MC simulation, a reweighting of a variable, representing the fraction of the longitudinal momentum transferred from the charm (beauty) quark to the charm (beauty) hadrons, z_{hadr}, was performed. This variable is sensitive to the longitudinal fragmentation function and has different shapes for different parametrisations. It was defined in the following way:

$$z_{\mathrm{hadr}} = \frac{(E + p_{||})_{\mathrm{hadr}}}{E_{\mathrm{string}}},$$

where E_{hadr} is the energy of the charm (beauty) hadron, $p_{||,\,\mathrm{hadr}}$ is the projection of its momentum onto the parent charm (beauty) quark momentum, E_{string} is the energy of the string which fragments to a given hadron (see Section 1.5 for more details on the string fragmentation model). All quantities are calculated in the rest frame of the string. Only charm or beauty hadrons that are produced in the fragmentation (not in decays) are considered.

In order to obtain the shape of z_{hadr} for the Peterson function, dedicated RAPGAP Monte Carlo samples were generated. All settings were as for the default samples, except for the fragmentation function which was chosen to be the Peterson function with ϵ-parameter values of 0.052, 0.062 and 0.072. Fig. 6.4(a) shows the distribution of z_{hadr} for



(a) (b)

Figure 6.5: Effect of the charm fragmentation function reweighting on the total charm (a) and beauty (b) cross sections. Other details are as in Fig. 6.1.

these samples as well as for the default one, with the Bowler function. The distribution illustrates that z_{hadr} is indeed sensitive to the fragmentation function and behaves similarly to the parametrisations shown in Fig. 6.3. By reweighting of z_{hadr} in the main samples (i.e. with the Bowler function), it is possible to mimic a change of the fragmentation function. Fig. 6.4(b) shows the resulting weighting functions from the Bowler to the Peterson parametrisation. For the evaluation of the fragmentation function systematics only the case of $\epsilon = 0.062$ was used for the reweighting.

In order to allow systematic uncertainty determination with the scan technique, the weight is defined in the following way:

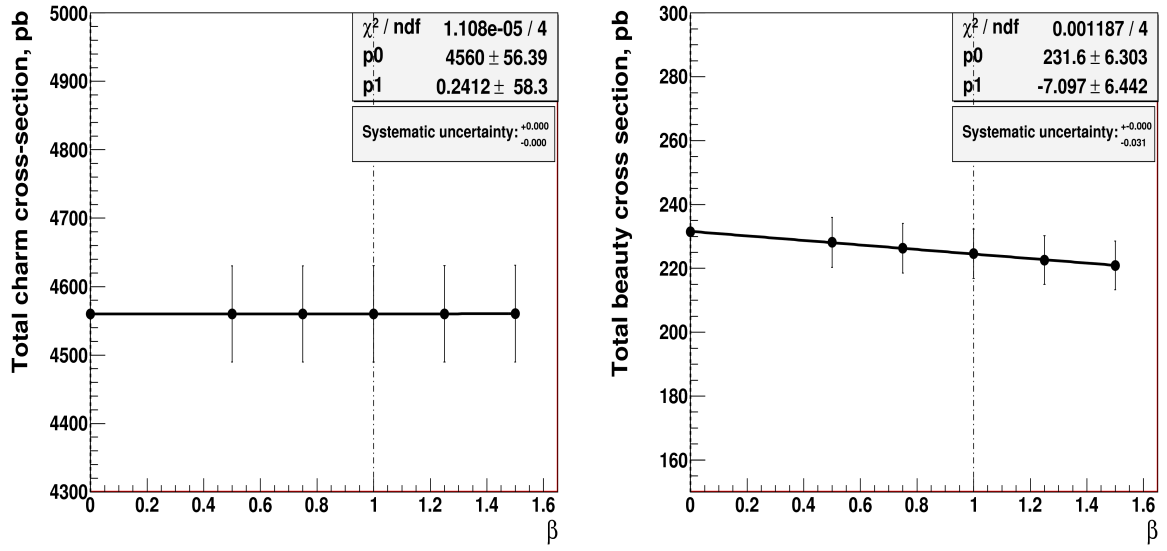
$$\text{weight} = 1 - \beta[1 - \omega(z)],$$

where $\omega(z)$ is the weight as a function of z shown in Fig. 6.4(b)[1], $\beta \in [0, 1]$ is the amount of the reweighting: for $\beta = 1$, reweighting is fully performed, while $\beta = 0$ means no reweighting. Figure 6.5 shows the dependences of the total charm and beauty cross sections on the β parameter. The resulting systematic effect on the total charm cross section was found to be $+1\,\%$, while for the beauty it is $-0.9\,\%$. Figure 6.5(a) indicates that the usage of the scan technique for charm improves the determination of the systematic uncertainty which would be underestimated in case of single variation only.

[1]The total event weight is defined as a product of weights of all charm (beauty) hadrons produced in the event.

Beauty Fragmentation Function

A similar procedure was also performed for the beauty fragmentation function systematics. The Peterson ϵ-parameter is different for beauty: the value of $\epsilon = 0.0041$ measured by the OPAL experiment at LEP [134] was used. Figure 6.6 shows the effect of the reweighting of the beauty fragmentation function from Bowler to Peterson on the total charm and beauty cross section, respectively. The effect on the charm cross section is negligible. The systematic uncertainty on the beauty is $-3\,\%$ and is one of its dominant uncertainties.



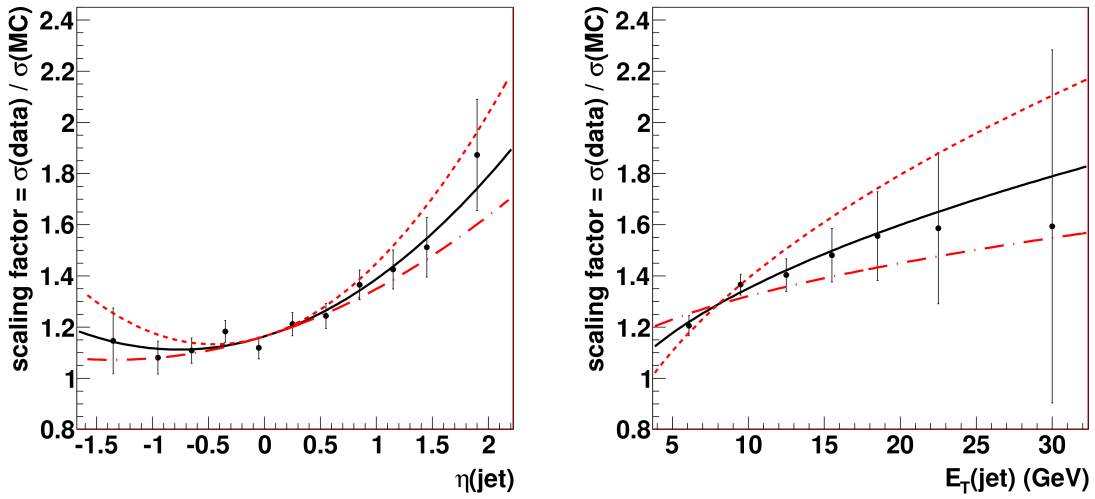
(a) (b)

Figure 6.6: Effect of the beauty fragmentation function reweighting on the total charm (a) and beauty (b) cross sections. Other details are as in Fig. 6.1.

Model dependence: charm η^{jet}- and E_T^{jet}-spectrum

As discussed in Section 5.6, the acceptance is sensitive to the modelling of the jet pseudorapidity, η^{jet}, and of the jet transverse energy, E_T^{jet}. Spectra of these distributions were reweighted in the MC so that it gives a good description of the data. In order to evaluate the corresponding systematic uncertainties, the η^{jet}- and E_T^{jet}-reweighting procedures were modified and the results were reextracted.

Figure 6.7(a) shows the scaling factors for charm, k_c, as a function of η^{jet} before the η^{jet}-reweighting as well as the default reweighting function, obtained from this distribution (see Section 5.6). The variations of the function for the systematics were determined by modifying its parameters until it is still consistent with the data (i.e. that each data point



(a) (b)

Figure 6.7: Variations of reweighting functions for determination of the systematic uncertainty on η^{jet}-reweighting (a) and E_T^{jet}-reweighting (b) for charm. The solid lines show the default reweighting functions; the dashed and dot-dashed lines are their extreme upward and downward variations respectively.

is not further from the curve than one standard deviation)[2]. The resulting functions are shown in Fig. 6.7(a). They are obtained by the variation of the coefficient p_2 of the quadratic term of the reweighting function, equation (5.2), and correspond to p_2=0.121 (upward variation) and to p_2=0.0403 (downward variation), while the default value is p_2=0.0715. The resulting systematic uncertainty on the total charm cross section is $^{+1.5\,\%}_{-1\,\%}$ and is negligible for the beauty. The fact that the variation of p_2 is asymmetric with respect to its default value leads to an asymmetric uncertainty. It is interesting to note that the usage of the scan technique does not improve the results in this case. This can be explained by the fact that only a reweighting is performed and there are no event migrations.

The systematics due to modelling of the E_T^{jet}-distribution was determined in a similar way. Figure 6.7(b) shows the default reweighting function as well as its extreme variations. The systematic uncertainty on the total charm cross section is $^{+2.2\,\%}_{-1.7\,\%}$, while on beauty it is $^{+1.7\,\%}_{-1.3\,\%}$.

[2]This approach leads to a more conservative uncertainty estimate, than requiring a change of χ^2 by 1.

Light Flavour Background

The mirroring of the projected decay length significance (Section 5.5) suppresses the light flavour background contribution almost completely, as can be seen by comparing Fig. 5.5 and Fig. 5.6. However it is still non-negligible, i.e. there is a certain asymmetry of the light flavour which remains in the mirrored distributions. One of the sources are decays of long-lived hadrons such as K^0_S-mesons or Λ-baryons [135].

In order to evaluate the sensitivity to this asymmetry, the light flavour contribution in the mirrored distributions was varied before the fit. The amount of the variation was $\pm 30\,\%$, according to studies by the H1 Collaboration [52, 55]; this number represents the uncertainty on the asymmetry in the MC, deduced from the comparison of K^0_S events in data and MC.

After the asymmetry was scaled by 30 % up (down), the χ^2/n_{dof} of the fit changed from the default value of 1.152 to 1.089 (1.397), while the light flavour scaling factor changed by $+0.6\,\%$ ($-0.6\,\%$). The resulting systematic uncertainty on the charm cross section is $\pm 2\,\%$ and is among the dominant effects. The effect on beauty is small ($< 0.5\,\%$). The difference between the effects on the charm and on the beauty is due to the fact that at high vertex masses where the beauty is enhanced, the light flavour contribution is very small and does not affect the results. However, for smaller m_{vtx}, where the charm dominates, the contribution of the light flavour is not negligible. These findings are consistent with the previous study [14].

Significance Smearing

The influence of the significance distribution modelling on the results was evaluated by the variation of the smearing procedure. In particular, the amount of smearing, i.e. the probability for a vertex to be smeared, was varied. This was done independently for the core and the tails, keeping the default parameters for other part. The ranges of the variation were determined conservatively as extremes at which the MC starts to deviate from the data dramatically. The probability for the core gaussian was varied to 5 % and 9 % (default: 7 %), while for the tail gaussian to 0 % (i.e. the smearing of the tail part was switched off) and 0.4 % (default: 0.2 %)[3].

Figure 6.8 shows the data-to-MC ratio of the significance distribution for the default smearing parameters as well as for these variations. It makes evident that the discrepancy at the variations is large, hence an uncertainty determined in this way is conservative

[3]For this study, additional corrections to the significance distribution (see Sect. 5.6) were undone, i.e. shifts of the decay length in the three most forward η^{jet}-bins were switched off, and smearing parameters in the most forward η^{jet}-bin were always set to same values as in the other bins instead of having their own values. The relative uncertainty was then applied to the cross sections with all corrections switched on. Thanks to linearity of the cross section as a function of smearing parameters (see Fig. 6.9), the uncertainty determined for the highest η^{jet}-bin corresponds to variation of $\pm 2\,\%$ (core) and $\pm 0.2\,\%$ (tail), as in the other regions.

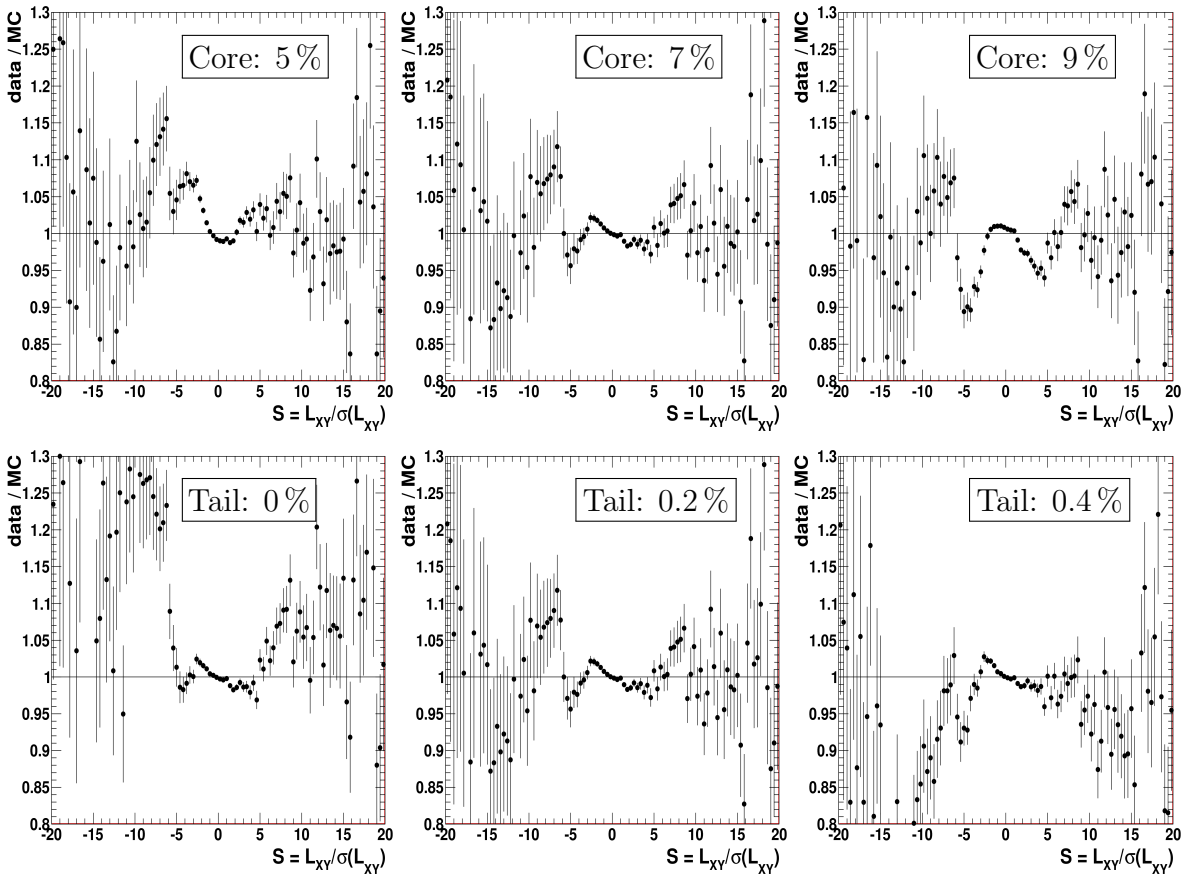
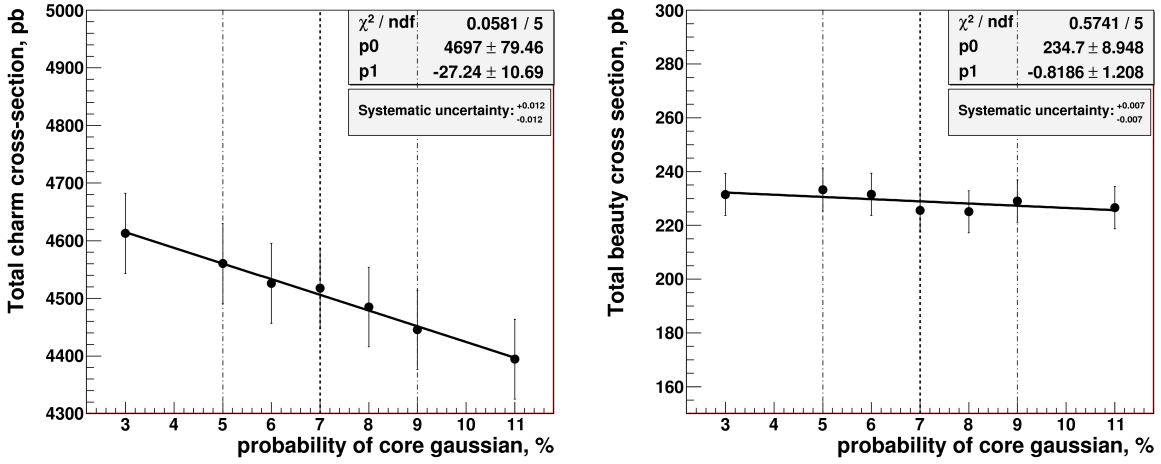


Figure 6.8: Data-to-MC ratio of the significance distribution, for different smearing parameters (shown in the boxes). Top row: variation of the core smearing parameters; bottom row: variation of the tail smearing.



(a) (b)

Figure 6.9: Effect of the variation of the core smearing probability on the total charm (a) and beauty (b) cross sections. Other details are as in Fig. 6.1.

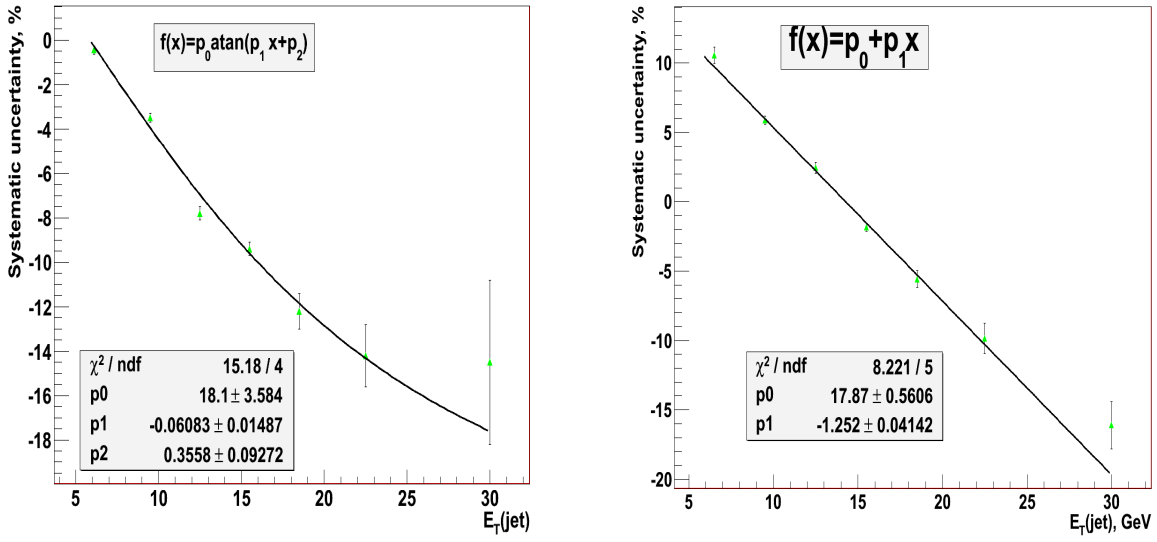
and covers the small disagreement between data and MC in the default spectrum (see Section 5.6).

Figure 6.9 shows the total charm and beauty cross sections as a function of the smearing probability for the core part. The resulting systematic uncertainty on the total charm cross section is $\pm 1.2\,\%$, while it is $\pm 0.7\,\%$ on the beauty. The uncertainty on the tail variation was deduced in a similar way. The resulting systematics on the total charm and beauty cross sections is $< 0.5\,\%$ and $\pm 1.1\,\%$, respectively.

For the core variation, the effect is largest for charm, while the tail variation affects mostly beauty. This can be explained by the fact that beauty is more pronounced in the region of high absolute significance (tail), while the charm populates mostly the region of lower $|S|$ (tail) as Fig. 5.6 shows. Hence, the variation of the core (tail) affects mostly the charm (beauty) cross sections.

Hadronic Energy Scale

The hadronic energy scale uncertainty arises due to the non-perfect calibration of the ZEUS hadronic calorimeter (HAC). The calibration precision is $3\,\%$ for jets with low transverse energy, $E_T^{\text{jet}} < 10\,\text{GeV}$ [136]. Hence, in order to evaluate the resulting sys-



(a) (b)

Figure 6.10: Systematic effect of the variation of the hadronic energy scale by $+3\,\%$ on the charm (a) and beauty (b) E_T^{jet}-differential cross section. The solid lines show the fits by an inverse tangent (a) and a first-order polynomial (b) functions, for illustration purposes.

tematic uncertainty, the energy measured in HAC is varied by $\pm 3\,\%$ in the MC only. As discussed in Section 5.3, jets were reconstructed from ZUFOs which include both the calorimeter and the tracking information. Therefore, for the hadronic energy scale systematics, only those ZUFOs were considered which are attributed to calorimeter energy deposits and contain no tracking information; their transverse energies were summed up to obtain the calorimeter contribution $E_T^{\text{jet, CAL}}$ to the jet transverse energy and $E_T^{\text{jet, CAL}}$ was varied.

The effect on the total charm cross section is $\pm 0.9\,\%$ while on the beauty total cross section it is $\pm 0.7\,\%$. A strong effect was observed for the differential cross section as a function of E_T^{jet}. Figure 6.10 shows the effect of the modification of the hadronic energy scale by $+3\,\%$ as a function of E_T^{jet}. For both charm and beauty, it increases up to $\sim 15\,\%$ at highest E_T^{jet} and is the dominant systematic uncertainty in that region. It is interesting to note that for charm, the effect is negative (i.e. the cross section decreases) for the whole region of E_T^{jet} (at $+3\,\%$ variation of the scale), while for beauty the effect is positive at low E_T^{jet} and it is negative at high E_T^{jet}. This difference between charm and beauty can be explained by different shapes of the E_T^{jet} spectrum: for charm production, a region above the threshold is probed, where the energy spectrum is monotonically falling, while for beauty production the threshold region is also probed, where the energy spectrum is rising at low values of E_T^{jet} and, after reaching the maximum, falls monotonically.

No strong dependences on the differential cross sections as a function of other variables were observed.

Electromagnetic Energy Scale

A similar uncertainty as the hadronic energy scale uncertainty arises due to the non-perfect calibration of the electromagnetic calorimeter. The calibration of the electromagnetic calorimeter is more precise than of the hadronic part and is at $1\,\%$ level at ZEUS [136]. The most affected observable is the scattered electron energy. Hence, in order to evaluate the effect of the electromagnetic energy scale, the reconstructed energy of the scattered electron was raised and lowered by $1\,\%$ in the MC only. The systematic uncertainty on both charm and on beauty total cross section is small and corresponds to $< 1\,\%$.

Signal Extraction

As discussed in Section 5.5, the fit is performed for the region of $|S| \in [4, 20]$. In order to cross-check the influence of the fit range on the results, the lower cut on $|S|$ was modified by ± 1 and the fit was repeated. The χ^2/n_{dof} of the fit changes from 1.152 to 1.167 (0.803) after a decrease (increase) of the lower $|S|$ cut by 1.

A modest effect of $\pm 0.8\,\%$ was observed both for charm and beauty total cross section.

DIS Selection

The influence of the event selection on the results was checked by modifying the selection criteria both in data and MC. The low-y cut was changed to 0.01 and 0.03. The cut on E'_e was changed to 9 GeV and 11 GeV. The $E - p_Z$ cut was modified to 42 GeV and 46 GeV. The latter also assesses a possible bias due to the non-perfect description of the $E - p_Z$ distribution by the Monte Carlo. Small effects of $\lesssim 1\,\%$ were found for both charm and beauty total cross sections.

Some systematic uncertainty sources were studied in the previous studies [14, 130] and there was no attempt made to improve them in this thesis. They are briefly described below.

Charm fragmentation fractions and branching ratios

As discussed in Section 5.6, the values of the charm fragmentation fractions and of the branching ratios were found to be outdated in the MC simulations. In order to correct them to the world average values [131, 132], a dedicated study was performed [130]; a reweighting procedure was developed and applied to the analysis. The corresponding systematic uncertainty was deduced by variations of the fragmentation fractions and the branching ratios within their uncertainties, repeating the reweighting and the fits.

Variations of the branching ratios yielded a $^{+3.3\,\%}_{-2.4\,\%}$ uncertainty on charm and $^{+1.8\,\%}_{-2.1\,\%}$ on beauty, while the effect from fragmentation fractions variation was $^{+1.0\,\%}_{-1.1\,\%}$ on charm and $\pm 0.3\,\%$ on beauty.

FLT Efficiency

At the first level trigger, except for the *FLT30* slot (see Section 5.3 for description of the first level trigger slots used in this analysis), events were vetoed using the tracking information (*track veto*). It is known that this effect is not well reproduced by the MC simulation; the difference between data and MC was estimated to be not more than 5 % [14]. Hence, in order to estimate the systematic uncertainty due to this effect, events not triggered by *FLT30* were weighted down by 5 % in the MC, and the analysis procedure was repeated [14].

The resulting systematic uncertainty on the total charm cross-section is $+1\,\%$, while it is $+2\,\%$ on beauty.

Model dependence: charm and beauty Q^2 spectrum

As discussed in Section 5.6, the Q^2 spectra were corrected for the charm and beauty MC by performing a reweighting. The resulting systematic uncertainty was evaluated by

variation of the reweighting by ±50 % [14]. The uncertainty on the charm total cross-section is ±2 % and it is ±3 % on the beauty total cross section.

Luminosity

The uncertainty on the luminosity determination at ZEUS is 2 %. This uncertainty was not included in the differential cross sections.

6.3 Summary

Detailed studies of systematic uncertainties were performed in this thesis. Table 6.1 summarises the uncertainties on the total charm and beauty cross sections from the sources that have been studied. The procedure has been repeated for every bin in the differential distributions. The uncertainties from all the sources have been added in quadrature in order to obtain the total systematic uncertainty.

Source	Charm systematics	Beauty systematics
Tracking efficiency	±0.9 %	±2.8 %
Charm fragmentation function	+1 %	−0.9 %
Beauty fragmentation function	negligible	−3 %
Hadronic energy scale	±0.9 %	±0.7 %
Significance smearing (core)	±1.2 %	±0.7 %
Significance smearing (tails)	<1 %	±1 %
Light flavour background	±2 %	<1 %
Model dependence: charm η^{jet}-reweighting	$^{+1.5\,\%}_{-1\,\%}$	<0.5 %
Model dependence: charm E_T^{jet}-reweighting	$^{+2.2\,\%}_{-1.7\,\%}$	$^{+1.7\,\%}_{-1.3\,\%}$
Model dependence: Q^2-reweighting	±2 %	±3 %
Signal extraction	±0.8 %	±0.8 %
DIS Selection (y_{JB} cut)	<1 %	±0.7 %
DIS Selection (E'_e cut)	<1 %	±1.0 %
DIS Selection ($E - p_Z$ cut)	±0.7 %	±0.7 %
EM scale	<1 %	<1 %
FLT efficiency	+1 %	+2 %
Luminosity	±2 %	±2 %
TOTAL (no Luminosity)	$^{+6.1\,\%}_{-5.2\,\%}$	$^{+6.0\,\%}_{-6.5\,\%}$

Table 6.1: Summary of the systematic uncertainties on the total charm and beauty cross sections.

Chapter 7

Results

This chapter presents the results of the measurement. First, details are given on theoretical predictions which are compared to the data. Measured differential cross sections of jet production in events containting charm and beauty quarks, together with theory calculations are then presented. Finally, the procedure to determine the charm and beauty contributions $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ to the proton structure function F_2 is discussed and corresponding results are shown.

7.1 NLO QCD Predictions

Theoretical predictions for comparisons with the cross sections measured in this thesis were obtained with the HVQDIS program. It provides Next-to-Leading Order (NLO) QCD calculations for charm and beauty quark production in deep inelastic electron-proton scattering in the Fixed Flavour Number Scheme (FFNS). HVQDIS works in a similar fashion as Monte Carlo generators, i.e. it produces events and provides full information about them. However, in contrast to MC programs, no hadronisation is performed: only partons created in the hard interaction (charm or beauty quarks and possibly an additional gluon) are present. This section describes the program settings which were used for the calculations and for the estimation of their uncertainties as well as the corrections that have to be applied to the predictions in order to allow direct comparisons with data.

The HERAPDF 1.0 FFNS PDF set [70] was used to parametrise the parton densities in the proton. The three-flavour variant of the PDF set was used for both charm and beauty cross section predictions. The strong coupling constant was set to the same value as in the PDF fit, namely to $\alpha_s^{n_f=3}(M_Z) = 0.105$ (this corresponds to $\alpha_s^{n_f=5}(M_Z) = 0.116$). The masses of the charm and beauty quarks were set to $m_c = 1.5\,\text{GeV}$ and $m_b = 4.75\,\text{GeV}$ respectively. Both renormalisation and factorisation scales were chosen to be $\mu_R = \mu_F = \sqrt{Q^2 + 4m_{b,c}^2}$.

In order to allow comparisons with data, measured and calculated cross sections have

to be defined in a consistent way. For this, hadronisation and real photon emission effects need to be taken into account. Hadronisation is accounted for in two steps. First, a jet algorithm[1] is applied to the outgoing partons in HVQDIS and the *parton level jet cross section* $\sigma^{\text{jet}}_{\text{part}}$ is calculated. In the second step, the parton level jet cross section is multiplied by a correction factor C_{hadr} which is obtained from the RAPGAP MC simulation in the following way:

$$C_{\text{hadr}} = \frac{\sigma_{\text{hadr, MC}}}{\sigma_{\text{part, MC}}},$$

where $\sigma_{\text{hadr, MC}}$ is the cross section of jet production in MC with the jet algorithm being applied to true level hadrons and $\sigma_{\text{part, MC}}$ is the same quantity obtained by running the jet algorithm on MC partons before the hadronisation. For the definition of $\sigma_{\text{hadr, MC}}$, the charm and beauty ground state hadrons are input to the jet algorithm and not their daughter particles.

HVQDIS simulates no real photon emissions from the incoming or outgoing lepton, which affects resulting cross sections (however, virtual corrections at the photon vertex are included). Hence, in order to allow comparisons to measured cross sections, *radiative corrections* are applied. Correction factors are obtained with the RAPGAP MC simulation:

$$C_{\text{rad}} = \frac{\sigma_{\text{rad, MC}}}{\sigma_{\text{no rad, MC}}},$$

where $\sigma_{\text{rad, MC}}$ is the jet cross section with simulation of the initial and final state real photon emission, while $\sigma_{\text{no rad, MC}}$ is the same quantity without the real photon emission. The cross section $\sigma_{\text{rad, MC}}$ was obtained from default samples, while for the $\sigma_{\text{no rad, MC}}$ quantity dedicated samples were generated with all setings identical to the default samples except for the photon radiation which was switched off.

Charm and beauty differential cross sections were calculated in the same phase space and with the same binning as the measured ones. Corrections were determined in Ref. [14], for each bin in the differential cross sections, separately for charm and beauty. In summary, cross sections theory predictions, directly comparable to data are given by:

$$\sigma = C_{\text{hadr}} \times C_{\text{rad}} \times \sigma^{\text{jet}}_{\text{part}}.$$

The uncertainties of the theory predictions were estimated by varying the HVQDIS settings. The charm and beauty quark masses and α_s were varied within their uncertainties:

- m_c was changed to 1.35 GeV and 1.65 GeV (charm cross sections);
- m_b was changed to 4.5 GeV and 5.0 GeV (beauty cross sections);
- α_s was changed to 0.103 and 0.107.

[1] Throughout this section, the same jet algorithm is used as for the reconstruction of detector level jets, namely the inclusive k_T-algorithm.

In order to estimate effects from missing higher orders, the factorisation and renormalisation scales were modified:

- μ_F was changed to $\frac{1}{2}\sqrt{Q^2+4m_{b,c}^2}$ and $2\sqrt{Q^2+4m_{b,c}^2}$;
- μ_R was changed to $\frac{1}{2}\sqrt{Q^2+4m_{b,c}^2}$ and $2\sqrt{Q^2+4m_{b,c}^2}$.

For variations of the charm quark mass, of α_s and of μ_F, a consistent PDF, i.e. obtained with the same value of the corresponding parameter in the QCD fit, was used. For each variation, a deviation from the default prediction was calculated. Effects from all sources were added in quadrature separately for positive and negative variations to obtain the total uncertainty on the prediction.

In order to assess the sensitivity to PDFs, the ABKM NLO PDF set [137] was used alternatively. In this case m_b was set to 4.5 GeV, consistent with the PDF and a modified version of HVQDIS was used, in which a missing term of the perturbative series was added [14]. Only the central value of predictions were produced, since uncertainties are expected to be very similar. As for the HERAPDF set, the three-flavour variant was used both for charm and beauty cross sections.

7.2 Differential Jet Cross Sections in Charm Events

This section presents measured differential cross sections of jet production in charm events. The cross sections are obtained by repeating the fit procedure described in the Section 5.5 for every bin in the corresponding variable. The measurement is performed in the following phase space region: $5 < Q^2 < 1000\,\mathrm{GeV}^2$, $0.02 < y < 0.7$, $E_T^{\mathrm{jet}} > 4.2\,\mathrm{GeV}$ and $-1.6 < \eta^{\mathrm{jet}} < 2.2$.

Figures 7.1(a)-(d) show the measured cross sections as a function of the jet pseudo-rapidity η^{jet}, jet transverse energy E_T^{jet}, photon virtuality Q^2 and Bjorken-x as well as corresponding theory predictions. The overall precision of the data is around 7 % on average and up to 20 % in some bins. Uncertainties of the theory predictions depend on the kinematic region and range from 10 % to 40 %. Generally, theory is consistent with data within uncertainties. The central values of the predictions are however typically 20 % to 30 % lower than the measured values. In most bins, the data are more precise than the theoretical predictions, hence they might have constraining power on these predictions, e.g. on their input parameters. The predictions obtained with the ABKM PDF set are very similar to those with the HERAPDF 1.0 set, hence the uncertainties due to PDF are not very large.

The cross section has largest values in the cental region, $\eta^{\mathrm{jet}} \approx 0$, and reduces towards positive and negative η^{jet}. The agreement between data and theory is good in the region $-0.2 < \eta^{\mathrm{jet}} < 1.6$ within uncertainties. In the backward region $-1.6 < \eta^{\mathrm{jet}} < -0.2$ the

theory slightly undershoots the data: the measured values are one standard deviation above the upper edge of the predictions. In the most forward bin ($1.6 < \eta^{\text{jet}} < 2.2$) the prediction is below the data as well, however both experimental and theoretical uncertainties are large. The E_T^{jet}-cross section falls off by three orders of magnitude monotonically from lowest values of $E_T^{\text{jet}} \approx 4.2\,\text{GeV}$ to the highest accessible $E_T^{\text{jet}} \approx 35\,\text{GeV}$. The predictions describe this dramatic drop-off well. In the region of $4.2 < E_T^{\text{jet}} < 8\,\text{GeV}$ the theory slightly underestimates the data. Since jets are a direct probe of the underlying charm quarks, the good agreement of data and theory suggests a good understanding of the dynamics of charm quark production at HERA by the NLO QCD. A large kinematic range is probed in Q^2 and x, where cross sections fall by three orders of magnitude. A reasonable description of this behaviour by the NLO QCD is observed for both Q^2 and x. However, by a closer inspection one notices that in most Q^2 bins, the theory undershoots the data, especially in the region of $10 < Q^2 < 70\,\text{GeV}^2$. Similarly, for the region of $0.0006 < x < 0.05$, predictions are below the data.

It is interesting to compare these observations to previous measurements of charm production at HERA. In a measurement of charm-jet cross sections by the H1 Collaboration [55], a similar kinematic region was probed, with somewhat more restricted η^{jet}-range. The same settings on HVQDIS were used for the predictions. Generally, a good agreement of data with theory was found within uncertainties. However, in regions of low-E_T^{jet} and of low-Q^2, central values of the predictions are below the data, consistently to observations of this analysis. In an investigation of $D^*(2010)$-production by the ZEUS Collaboration [44], a somewhat wider region in Q^2 was probed than in this measurement. The agreement between data and central values of the theoretical predictions is good. However, a value of the charm mass of $m_c = 1.35\,\text{GeV}$ was used for theory calculations which is different from $m_c = 1.5\,\text{GeV}$ used in this analysis. A decrease of m_c leads to an increase of predicted cross section; hence, the theory would slightly undershoot the data if a value of $m_c = 1.5\,\text{GeV}$ was used. In the $D^*(2010)$-measurement by the H1 Collaboration [56], a good agreement between data and theory was found; however, in certain kinematic regions, such as low-Q^2, low-x or low momentum of the D^*-meson, the central values of predictions are mostly below the data. In summary, similar observations were made in earlier publications: albeit the general agreement of data and theory is reasonable within the theory uncertainties, the central values of the predictions tend to underestimate the data in some regions of phase space such as at low Q^2 or at low charm quark momentum.

As it was discussed in Section 1.2, fixed-order QCD calculations in the fixed flavour number scheme (FFNS) can potentially suffer from the multiple scale problem in regions of high photon virtuality Q^2 or high quark transverse momentum squared p_T^2, compared to the charm quark mass squared m_c^2, due to presence of terms proportional to $[\alpha_s \log(p_T^2/m_c^2)]^n$ or $[\alpha_s \log(Q^2/m_c^2)]^n$ in the perturbative series, where n is the order of the calculation. On the other hand, they should be reliable for $(E_T^{\text{jet}})^2$ and Q^2 comparable to m_c^2. Up to highest probed Q^2, no deviations from FFNS predictions were observed so far in previous H1 and ZEUS measurements, see Section 1.4 and Ref. [138]. Nevertheless it

is instructive to check whether the same conclusion may be reached in this measurement, since it covers both regions and hence is sensitive to the multiple scale problem. Indeed, the E_T^{jet}-cross section spans from $(E_T^{\text{jet}})^2 \approx 18\,\text{GeV}^2$ which is above but comparable to the squared charm quark mass, $m_c^2 \approx 2.25\,\text{GeV}^2$, to $(E_T^{\text{jet}})^2 \approx 1200\,\text{GeV}^2$ which is very much above m_c^2. The NLO QCD prediction provides a reasonable description up to the highest E_T^{jet}, similarly as at low-E_T^{jet}. Hence, the multiple scale problem does not manifest itself for charm production in the probed kinematic region confirming the reliability of FFNS predictions. Similarly, the photon virtuality is probed from $Q^2 = 5\,\text{GeV}^2 \sim m_c^2$, up to $Q^2 = 1000\,\text{GeV}^2 \gg m_c^2$. A good description is observed up to highest values of Q^2.

Figs. 7.2 and 7.3 show the differential cross sections as a function of x in bins of Q^2. Again, an agreement between data and theory is observed within uncertainties. However, the central values of the predictions are mostly below the measured cross sections, especially in the region of $20 < Q^2 < 60\,\text{GeV}^2$, consistently to observations in the differential cross section in Q^2, Fig. 7.1 (c). These double-differential cross sections are the basis for the determination of the charm contribution to the proton structure function, $F_2^{c\bar{c}}$ (Sect. 7.4).

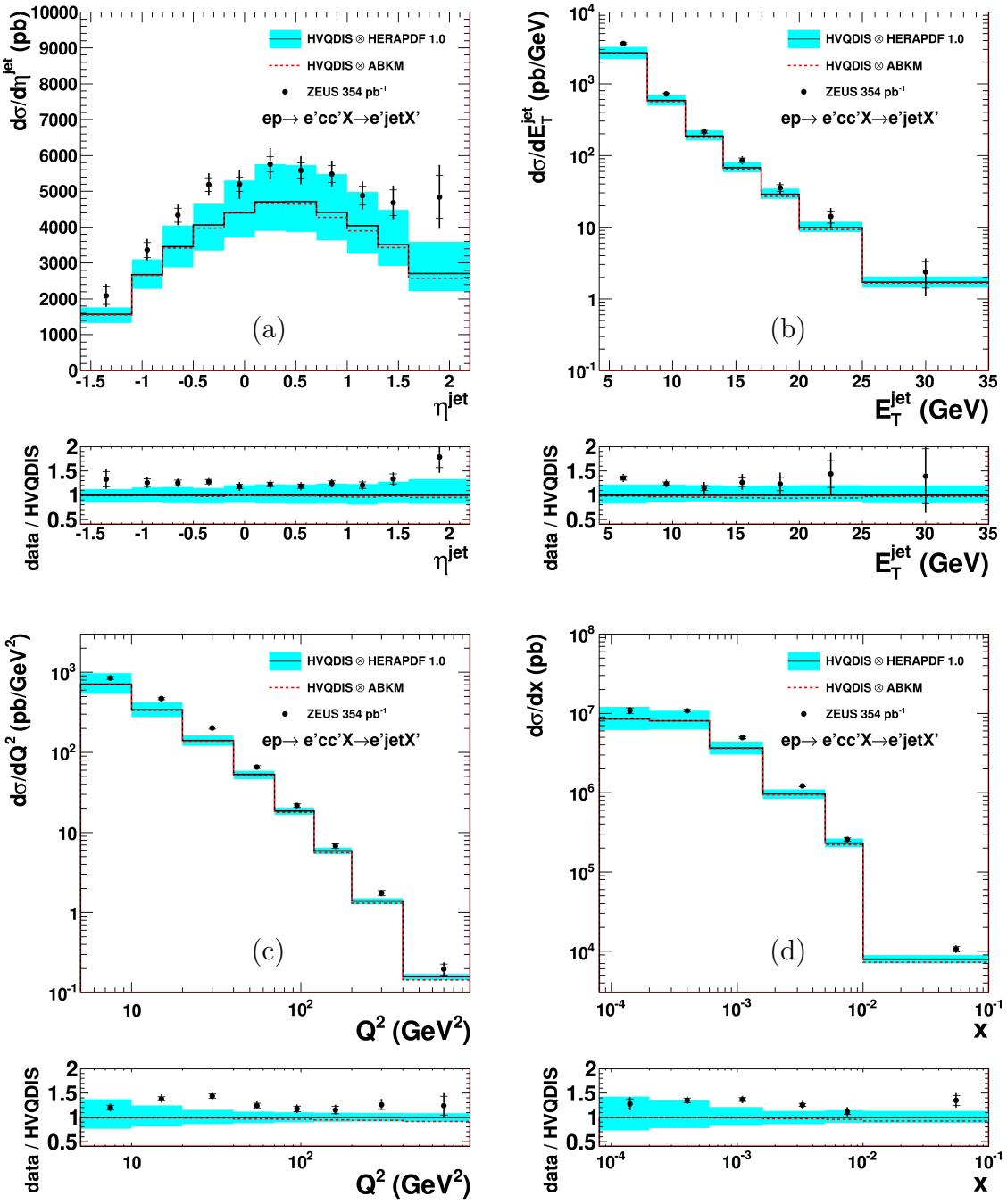


Figure 7.1: Differential cross sections of jet production in charm events as a function of the jet pseudorapidity η^{jet} (a), transverse jet energy E_T^{jet} (b), photon virtuality Q^2 (c) and Bjorken-x (d) for the following phase space region: $5 < Q^2 < 1000\,\text{GeV}^2$, $0.02 < y < 0.7$, $E_T^{\text{jet}} > 4.2\,\text{GeV}$ and $-1.6 < \eta^{\text{jet}} < 2.2$. The inner error bars represent statistical uncertainties, while the outer error bars show statistical and systematic uncertainties added in quadrature. The solid lines represent NLO QCD predictions obtained with the HERAPDF 1.0 PDF set, while the blue bands show their uncertainties. The red dashed lines denote NLO QCD predictions obtained with the ABKM NLO PDF set. The small inserts below each cross section plot show the corresponding data-to-theory ratio.

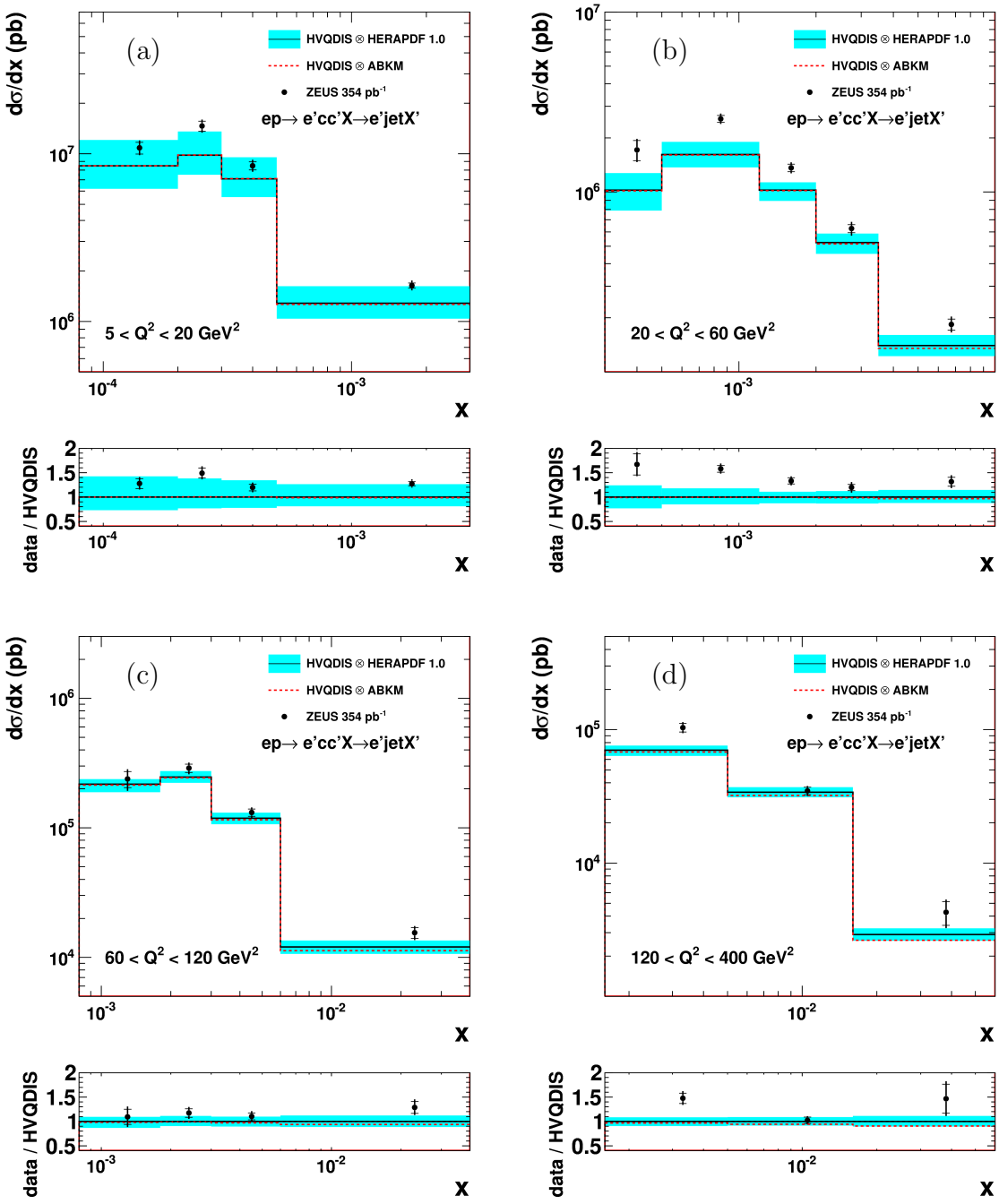


Figure 7.2: Differential cross sections of jet production in charm events for $0.02 < y < 0.7$, $E_T^{\rm jet} > 4.2\,{\rm GeV}$ and $-1.6 < \eta^{\rm jet} < 2.2$ as a function of x in the following regions of Q^2: $5 < Q^2 < 20\,{\rm GeV}^2$ (a), $20 < Q^2 < 60\,{\rm GeV}^2$ (b), $60 < Q^2 < 120\,{\rm GeV}^2$ (c), $120 < Q^2 < 400\,{\rm GeV}^2$ (d). Other details are as in the Fig. 7.1.

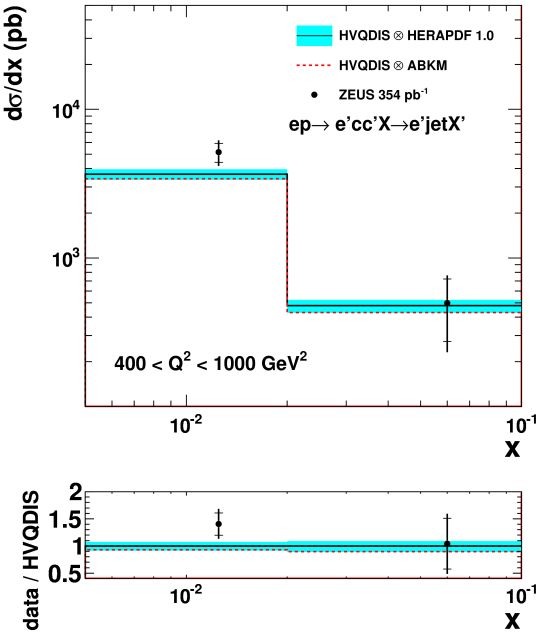


Figure 7.3: Differential cross sections of jet production in charm events for $0.02 < y < 0.7$, $E_T^{\text{jet}} > 4.2\,\text{GeV}$ and $-1.6 < \eta^{\text{jet}} < 2.2$ as a function of x for the Q^2 region of $400 < Q^2 < 1000\,\text{GeV}^2$. Other details are as in the Fig. 7.1.

7.3 Differential Jet Cross Sections in Beauty Events

In this section, jet cross sections in beauty events are presented. The same set of cross sections was obtained as for charm (see the previous section). The phase space region is defined by $5 < Q^2 < 1000\,\mathrm{GeV}^2$, $0.02 < y < 0.7$, $E_T^{\mathrm{jet}} > 5.0\,\mathrm{GeV}$ and $-1.6 < \eta^{\mathrm{jet}} < 2.2$ which is the same as for the charm measurement except for the jet transverse momentum region.

Figs. 7.4(a)-(d) show differential cross sections as a function of η^{jet}, E_T^{jet}, Q^2 and x. All bins are the same as for charm except for the η^{jet}-differential cross section, where the two lowest-η^{jet} bins were merged in order to reach a satisfactory statistical precision. The relative uncertainties of the measured beauty cross sections are larger than for charm and are around 10–15 % on average. In contrast, the relative uncertainties on the theoretical predictions are smaller than for the charm case, thanks to the large mass of the beauty quark; they range from 10 % to 20 %. The agreement between data and theory is good which points to a good reliability of NLO QCD predictions for beauty production at HERA. As for the charm case, the description remains good up to highest values of E_T^{jet} and Q^2. A similar consistency of data and theory for beauty cross sections was observed in recent measurements with a lifetime tag by the H1 Collaboration [55] and with an electron tag by the ZEUS Collaboration [69]. The measurement presented here covers a larger range in the pseudorapidity than the previous studies [55,69]: region up to $\eta^{\mathrm{jet}} = 2.2$ is probed, whereas previous analyses were limited to around $\eta = 1.5$.

Figures 7.5 and 7.6 show differential cross sections as a function of x in bins of Q^2. Compared to charm, the highest-x bin for $120 < Q^2 < 400\,\mathrm{GeV}^2$ was removed due to low statistical precision. Again, a good agreement between data and theory is observed. These cross sections are used for the $F_2^{b\bar{b}}$ determination (Sect. 7.4).

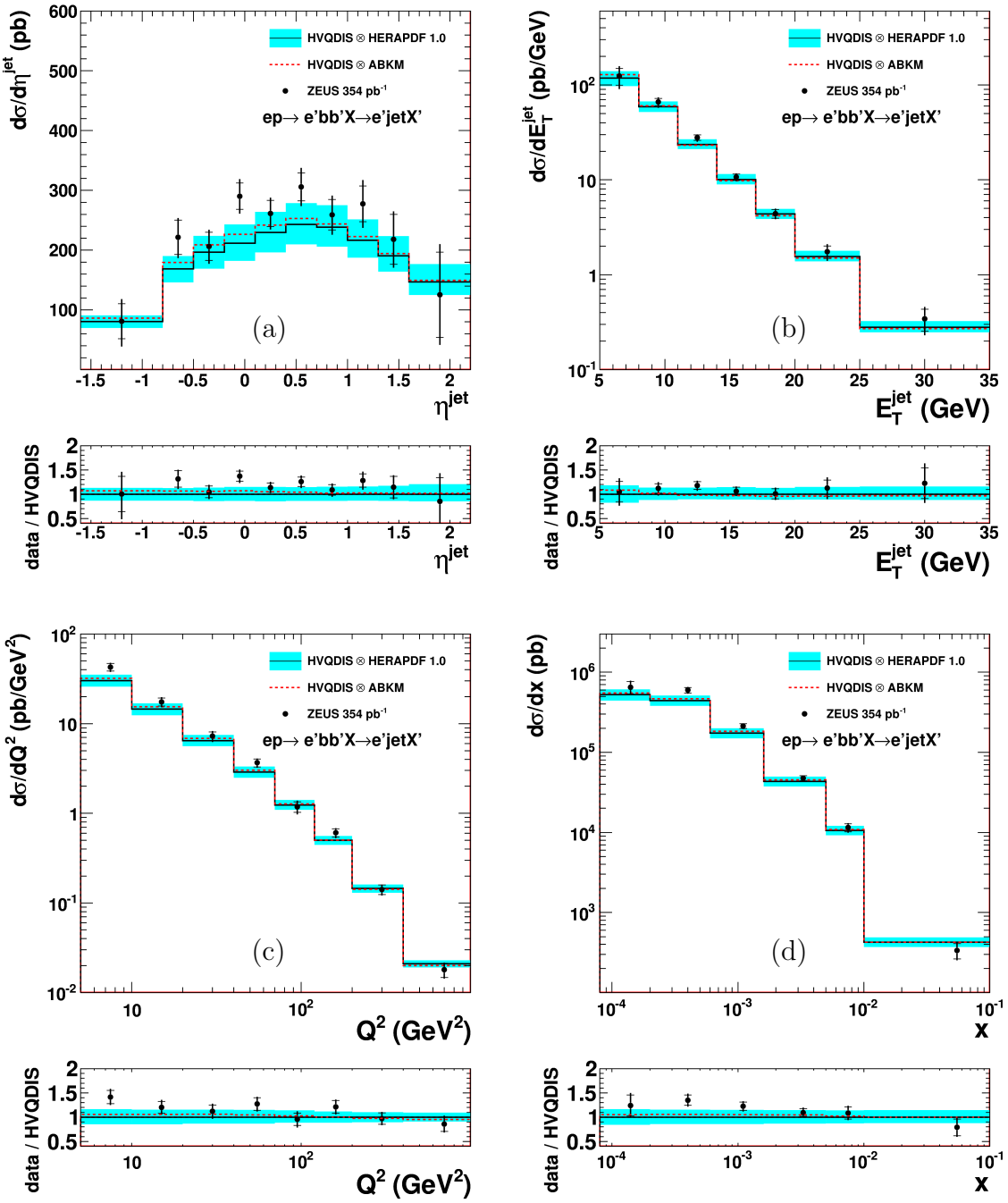


Figure 7.4: Differential cross sections of jet production in beauty events as a function of the jet pseudorapidity $\eta^{\rm jet}$ (a), transverse jet energy $E_T^{\rm jet}$ (b), photon virtuality Q^2 (c) and Bjorken-x (d) for the following phase space region: $5 < Q^2 < 1000\,{\rm GeV}^2$, $0.02 < y < 0.7$, $E_T^{\rm jet} > 5.0\,{\rm GeV}$ and $-1.6 < \eta^{\rm jet} < 2.2$. Other details are as in the Fig. 7.1.

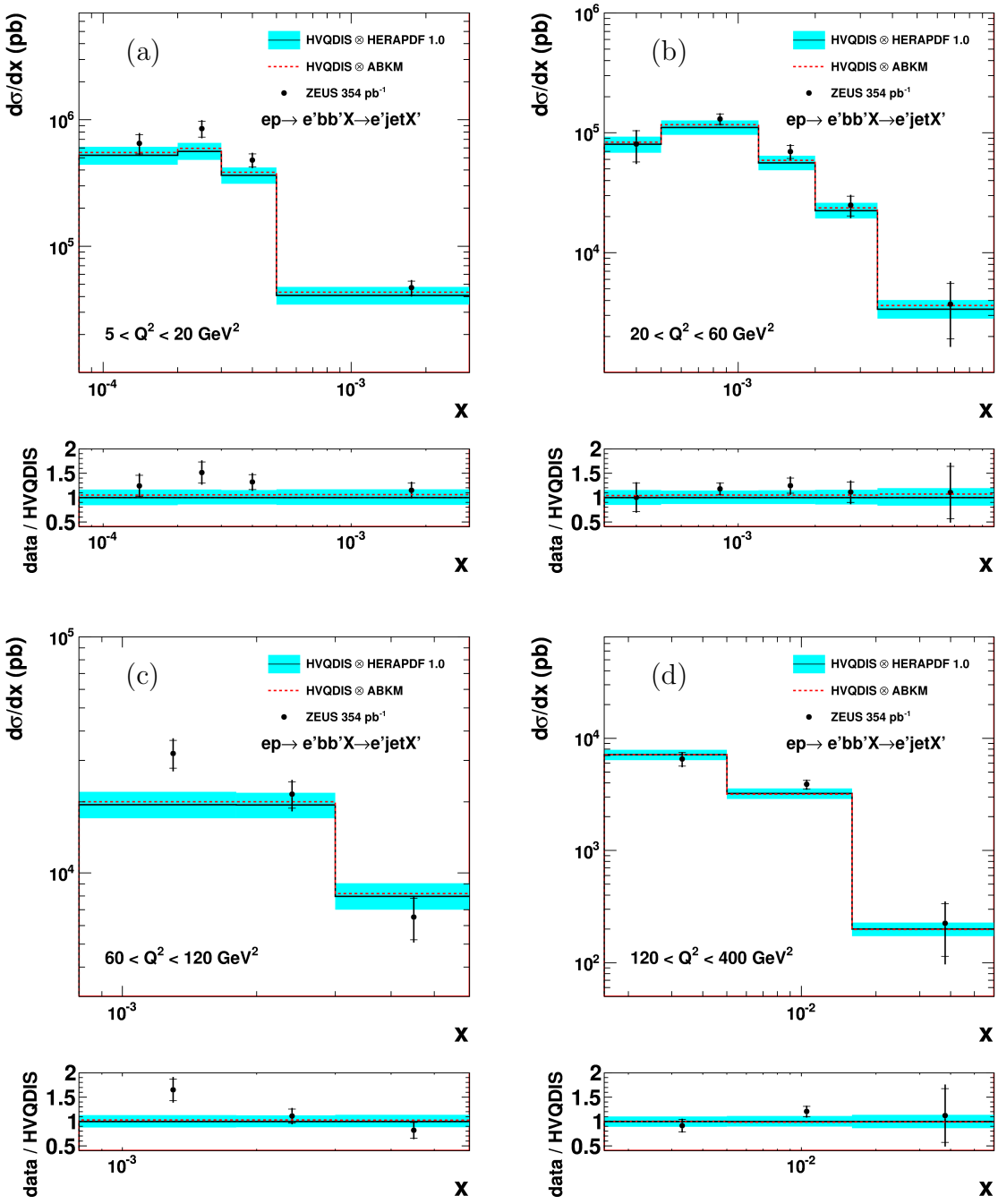


Figure 7.5: Differential cross sections of jet production in beauty events for the phase space region of $0.02 < y < 0.7$, $E_T^{\text{jet}} > 5.0\,\text{GeV}$ and $-1.6 < \eta^{\text{jet}} < 2.2$ as a function of x in the following regions of Q^2: $5 < Q^2 < 20\,\text{GeV}^2$ (a), $20 < Q^2 < 60\,\text{GeV}^2$ (b), $60 < Q^2 < 120\,\text{GeV}^2$ (c), $120 < Q^2 < 400\,\text{GeV}^2$ (d). Other details are as in the Fig. 7.1.

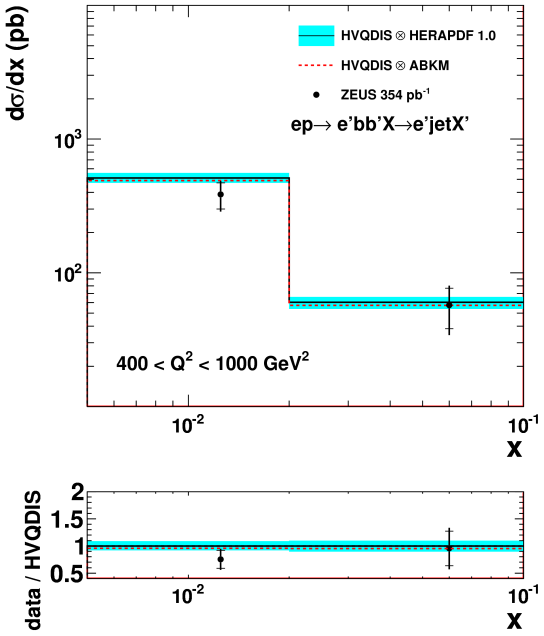


Figure 7.6: Differential cross sections of jet production in beauty events for $0.02 < y < 0.7$, $E_T^{\text{jet}} > 5.0\,\text{GeV}^2$ and $-1.6 < \eta^{\text{jet}} < 2.2$ as a function of x for the Q^2 region of $400 < Q^2 < 1000\,\text{GeV}^2$. Other details are as in the Fig. 7.1.

7.4 $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ Determination

In the previous sections, the differential jet cross sections in charm (beauty) events were presented. Results were compared to theory calculations based on the HVQDIS program, which provides NLO QCD predictions in the fixed flavour number scheme. Predictions in other schemes than FFNS and higher order calculations are not available for such differential cross sections.

A more convenient observable is the charm (beauty) contribution to the proton structure function, $F_2^{c\bar{c}}$ ($F_2^{b\bar{b}}$). It is closely related to the double-differential cross section of charm (beauty) quark-antiquark pair production as a function of Q^2 and x, $d^2\sigma^{q\bar{q}}/dQ^2dx$, where $q = c, b$. The $F_2^{q\bar{q}}$ is a function of x and Q^2 and is defined in analogy to the inclusive DIS case (neglecting Z^0 exchange):

$$\frac{d^2\sigma^{q\bar{q}}}{dQ^2dx} = \frac{2\pi\alpha^2}{Q^4x}[(1+(1-y)^2)F_2^{q\bar{q}}(x,Q^2) - y^2F_L^{q\bar{q}}(x,Q^2)],$$

where $F_L^{q\bar{q}}(x,Q^2)$ is the contribution of charm (beauty) to the longitudinal proton structure function F_L. QCD calculations of this quantity are available in different schemes both at NLO and (partially) at NNLO. In addition, $F_2^{q\bar{q}}(Q^2,x)$ is independent of particular final states and allows direct comparisons between measurements with different charm (beauty) tagging techniques as well as their combination.

In order to extract the $F_2^{q\bar{q}}$, it is necessary to determine $d^2\sigma^{q\bar{q}}/dQ^2dx$. However, due to the fact that detectors always have only a limited geometrical acceptance, it is not possible to directly measure the latter, and only double-differential cross sections $d^2\sigma/dQ^2dx$ with certain cuts on the quark kinematics are accessible. In this measurement these are the restrictions on the jet pseudorapidity η^{jet} and on the jet transverse energy E_T^{jet}. Hence, an extrapolation must be performed from the visible phase space (with cuts) to the full phase space (without any cuts).

In this thesis, the extrapolation is done with the HVQDIS NLO QCD calculations. The value of $F_2^{q\bar{q}}(x_i,Q_j^2)$ at a certain point in the kinematic plane (x_i, Q_j^2) is obtained in the following way:

$$F_2^{q\bar{q}}(x_i,Q_j^2) = \frac{\left(\frac{d^2\sigma}{dQ^2dx}\right)^{\text{meas}}_{i,j}}{\left(\frac{d^2\sigma}{dQ^2dx}\right)^{\text{NLO}}_{i,j}} \; F_2^{q\bar{q}}(x_i,Q_j^2)^{\text{NLO}},$$

where i and j denote a certain bin in x-Q^2 plane, the point (x_i, Q_j^2) is chosen inside the bin; $\left(\frac{d^2\sigma}{dQ^2dx}\right)^{\text{meas}}_{i,j}$ is the measured charm (beauty) double-differential cross section in this bin (e.g. one of those shown in Figs. 7.2 or 7.5 divided by the bin width in Q^2); $\left(\frac{d^2\sigma}{dQ^2dx}\right)^{\text{NLO}}_{i,j}$ is the NLO prediction of this quantity from HVQDIS; $F_2^{q\bar{q}}(x_i,Q_j^2)^{\text{NLO}}$ is the

HVQDIS prediction for the $F_2^{q\bar{q}}$ at (x_i, Q_j^2). The same set of $(x_i,\, Q_j^2)$ points for extraction of $F_2^{q\bar{q}}$ was chosen as in Ref. [14].

The HVQDIS settings used for extraction of $F_2^{q\bar{q}}$ were as described in Section 7.1. The uncertainty of the extrapolation was estimated by varying the settings and redoing the procedure. The same variations were used as described in Section 7.1 except for the scale variations: for the extraction, a *simultaneous* variation of μ_R and μ_F was performed, following Ref. [70].

A reliable shape description of cross sections as a function of η^{jet} and E_T^{jet} is required to ensure a proper extrapolation from the visible to the full phase space in these variables. As it was shown in Figs. 7.1 (a), (b) for charm and in Figs. 7.4 (a), (b) for beauty, the shapes are indeed reasonably well described by the theory in the visible region.

Figures 7.7 and 7.8 show results for the extracted $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ respectively, as a function of x for different values of Q^2. An agreement between data and NLO QCD predictions is observed both for charm and beauty within uncertainties; for $Q^2 = 25\,\text{GeV}^2$ and $Q^2 = 30\,\text{GeV}^2$ theory are below the data for charm. However one has to be cautious since for the prediction uncertainties (which are given by the blue bands in Figures 7.7 and 7.8) a *simultaneous* variation of factorisation and renormalisation scales was performed; normally this is done separately and results in larger uncertainties. It is evident that for charm at low values of Q^2, the total uncertainty is significantly larger than the statistical one. This comes from large extrapolation uncertainties due to a relatively high cut on the jet transverse energy, compared to the charm quark mass m_c. For beauty the extrapolation uncertainties are small thanks to the large beauty quark mass: the E_T^{jet} cut is comparable to m_b and allows a measurement down to low values of its transverse momentum p_T, hence less extrapolation is needed.

7.5 Summary

In this chapter, the results of the charm and beauty quark production measurement were presented. Differential cross sections as functions of the jet pseudorapidity η^{jet}, jet transverse energy E_T^{jet}, photon virtuality Q^2 and Bjorken-x were measured and compared to NLO QCD predictions in the fixed flavour number scheme. Both charm and beauty results agree reasonably to theory in the whole kinematic region. However, the central values of charm predictions tend to underestimate the measured cross sections; the upper margins of the predictions are mostly consistent with the data though. These observations are generally consistent to previous publications.

The double-differential cross sections as a function of x and Q^2 were used to extract charm and beauty contributions to the F_2 proton structure function, $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$. Again, a reasonable agreement between data and theory is observed. The $F_2^{c\bar{c}}$ results are competitive at mid- and high-Q^2, while the $F_2^{b\bar{b}}$ measurement is the most precise among published measurements. These data will serve as input to future combinations of various measurements of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ by H1 and ZEUS.

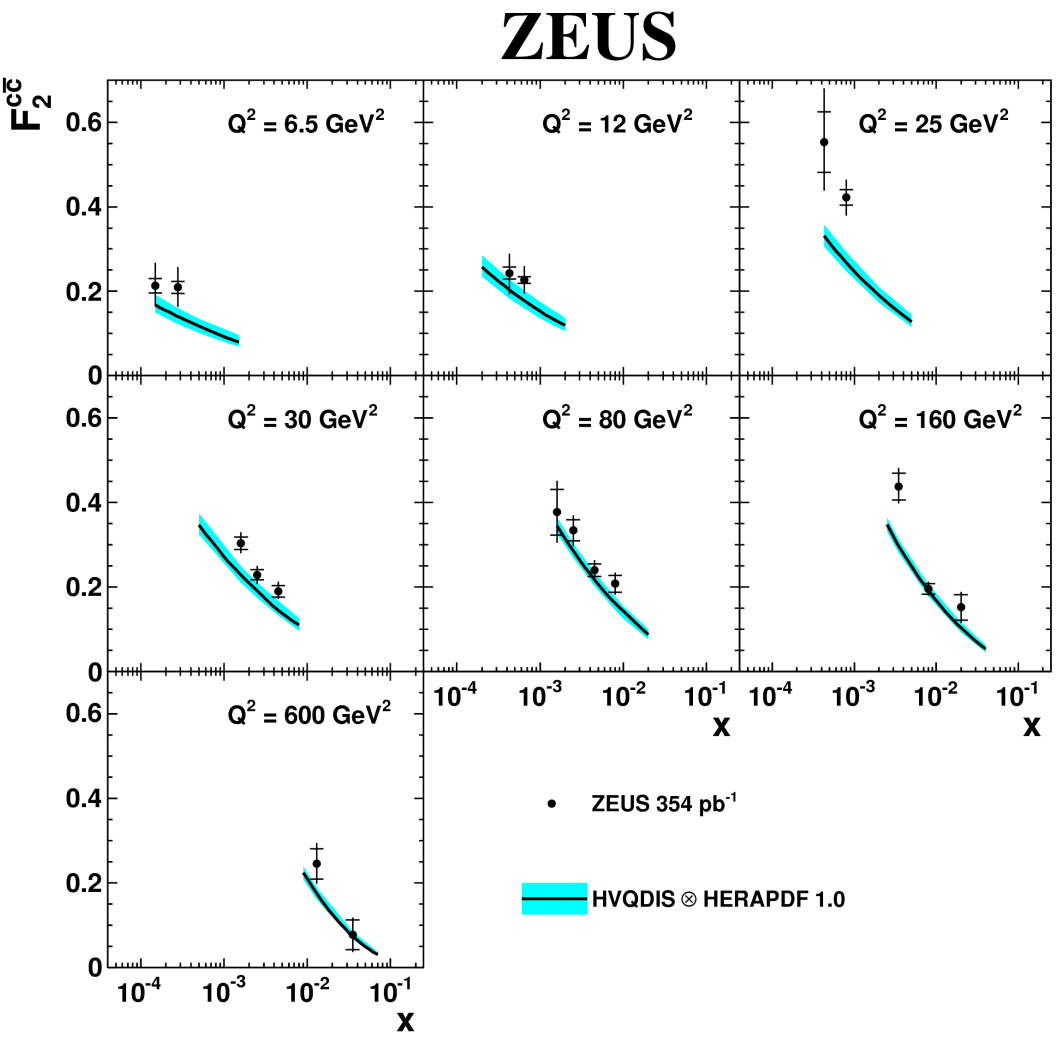


Figure 7.7: The charm contribution to the proton structure function, $F_2^{c\bar{c}}$, as a function of x for different values of Q^2. The inner error bars represent statistical uncertainties, while the outer error bars show statistical, systematic and extrapolation uncertainties added in quadrature. The solid line represents NLO QCD prediction obtained with the HERAPDF 1.0 PDF set, while the blue band show its uncertainty. A simultaneous variation of factorisation and renormalisation scales was performed for evaluating the uncertainty (see text).

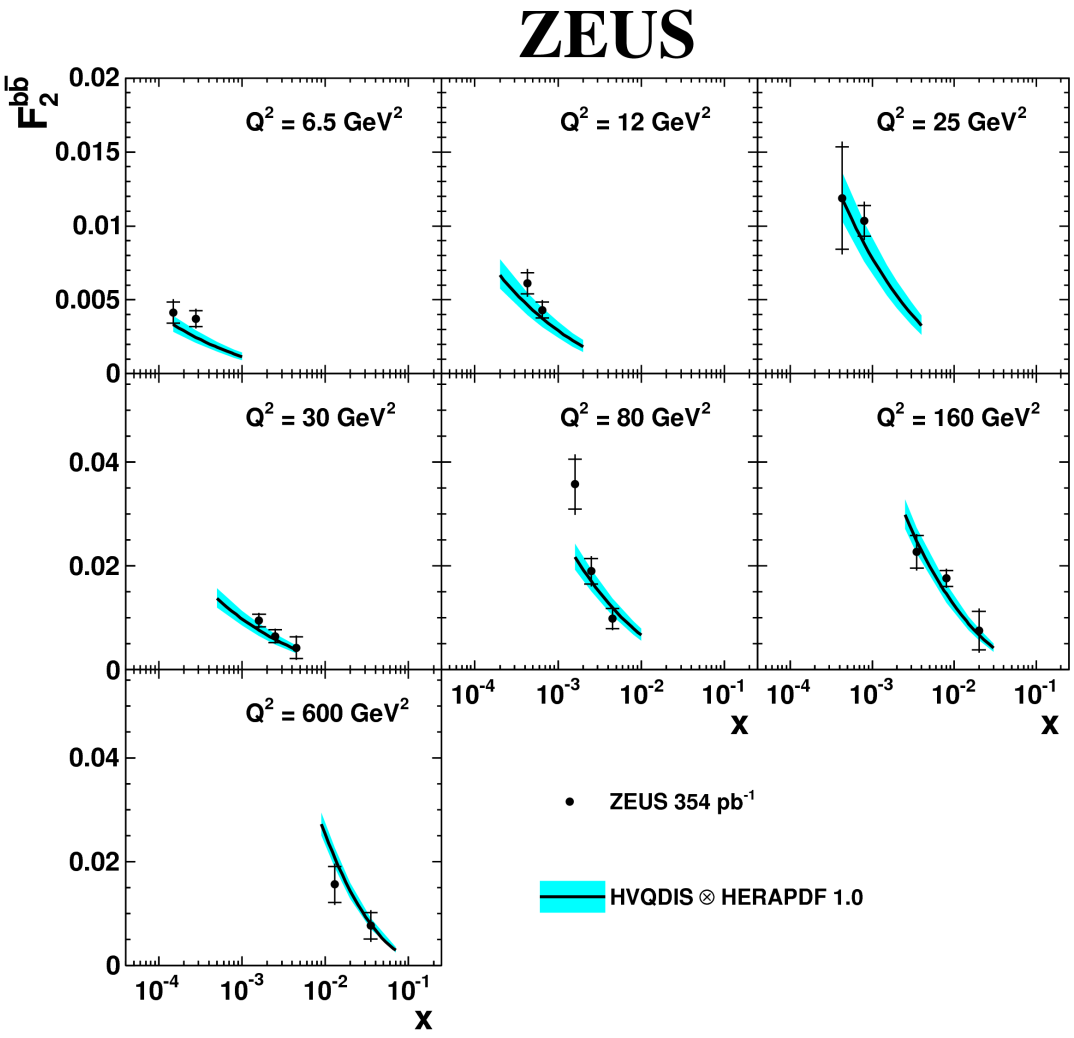


Figure 7.8: The beauty contribution to the proton structure function, $F_2^{b\bar{b}}$, as a function of x for different values of Q^2. Other details are as in the Fig. 7.7.

Chapter 8

Test Beam Studies of the ATLAS Pixel Sensors

This chapter describes ATLAS Pixel detector upgrade activities, which were performed within this thesis. It is organised as follows: a general introduction to the physics goals of the Large Hadron Collider (LHC) is given first, the LHC machine and the ATLAS experiment are briefly described next. The focus then moves to the present ATLAS Pixel detector, the need of its upgrade and the Insertable B-layer project. Afterwards the test beam program for the IBL pixel sensors characterisation and the relevant tools are introduced. In what follows, the contributions of the thesis to this program are presented. They include extension of the EUTelescope test beam data analysis software, the first IBL test beam campaign and data analysis as well as integration of the USBPix DAQ system into the EUDET Telescope.

8.1 Physics Goals of the Large Hadron Collider

The Standard Model (SM) is the very successfull theory of particle interactions. It has survived all experimental tests so far, primarily at LEP, SLC, Tevatron and HERA colliders. It had predicted existence of several particles, such as heavy gauge bosons, $W^{\pm}$ and Z^0, which have been discovered experimentally. By now a new particle, consistent with the Higgs Boson which was also predicted theoretically, was discovered [139, 140]. This all definitely gives confidence in SM. However, there are also several phenomena that it is not able to describe [7]:

- **Gravity.** So far it was not possible to construct a renormalizable quantum theory of gravitational interaction, hence the Standard Model is incomplete. The common opinion is that our understanding of space and time has to be completely reviewed before a relativistic quantum theory of gravity can be developed.

- **Dark Matter.** Astrophysical observations suggest that conventional baryonic matter constitutes only a small fraction of matter in the Universe. There must be a different component, called *Dark Matter*, which cannot be composed of the particles that are included in SM.

- **Dark Energy.** It is known that the expansion velocity of the Universe is increasing with time. However if only usual gravity would be at work, the expansion should decelerate. Hence there is an unknown mechanism, conventionally called *Dark Energy*, which is responsible for this acceleration.

- **Hierarchy Problem.** One expects, that quantum effects for gravity become important at the Planck scale of $M \sim 10^{19}\,\mathrm{GeV}$. If particles with such masses exist, they would inevitably occur in the self-energy loops and would lead to quadratic corrections to it. The empirical success of the Standard Model relies on a very precise cancellation of this large correction which is unnatural. This is termed as the *Hierarchy Problem*.

These are the main reasons why it is widely believed that the SM is not the final theory of Nature, despite its tremendous success. There are also purely aesthetic reasons, which may lead to this conclusion. One example is the fact that SM contains many free parameters which cannot be predicted and can only be taken from the experiment. It would be nice to have fewer constants in the final theory and to be able to predict all the other parameters.

New theories and models have been proposed in the past decades, such as String theory, Supersymmetry (SUSY) or Technicolor, which address some of the problems listed above [7]. For example, SUSY is designed to cancel quadratic divergences in scalar particle loop corrections, hence it provides a solution to the Hierarchy Problem. SUSY also predicts existence of new particles, which could constitute Dark Matter.

A search for new phenomena, which would unambiguosly show that the SM breaks down, as well as search for the missing piece of the SM, the Higgs Boson, and study of its properties are the main objectives of the *Large Hadron Collider* (LHC). It is briefly described in the following section.

8.2 The LHC Machine

The purpose of the Large Hadron Collider [141] is to store, accelerate and bring into collisions beams of protons or ions. The design centre-of-mass energy in the proton mode is 14 TeV while the design luminosity constitutes $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which makes LHC the most powerful particle accelerator in the world. The corresponding parameters for the ion operation are the energy of 5.6 TeV per nucleon pair and the luminosity of $\mathcal{L} = 10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

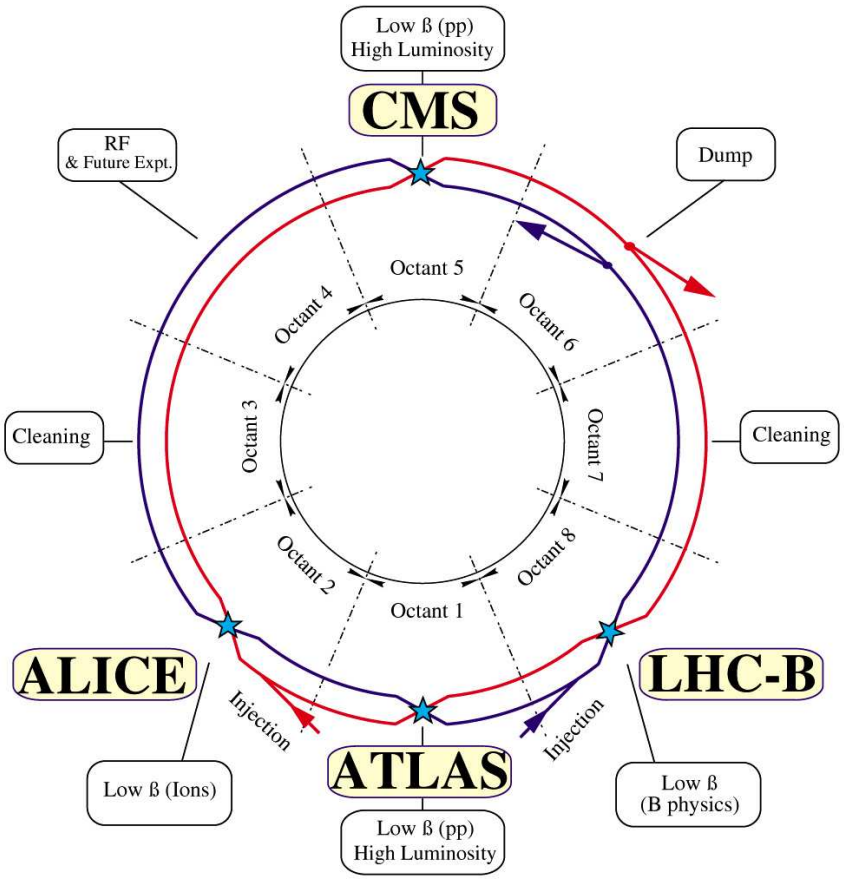


Figure 8.1: The LHC layout [141].

LHC is located at CERN (Geneva, Switzerland) in the 27 km circumference LEP tunnel, 45 m to 170 m below the ground. The layout of the machine is given in Fig. 8.1. It has two rings, for clockwise and anticlockwise rotating beams, shown in red and blue respectively. The tunnel consists of eight arcs and eight straight sections. In the arcs, superconducting dipole magnets which provide the bending field are placed. Beams are brought into collision in four straight sections. The remaining four straight sections contain important machine systems, such as accelerator cavities, cleaning or dump systems.

Particles constituting the counter-rotating beams are of the same charge, hence the direction of the bending field must be opposite for both rings. Consequently, there is an independent magnet system for each ring in contrast to the Tevatron, where the same magnets could be used both for proton and antiproton beams. Because of the space limitation in the tunnel, the twin-bore magnet design was chosen: both magnet systems and rings share the same mechanical structure and cryostat [141].

In order to reach the design energy with the existing tunnel, a magnetic field of 8.36 T was required. In order to achieve it, the nominal operation temperature of the super-

conducting magnets was set to 1.9 K, while for the previous large accelerator complexes (Tevatron, HERA, RHIC) it was around 4.5 K. This imposes more demands on the magnet cable and coil assembly.

For the nominal operation, 2808 bunches of 1.15×10^{11} protons are stored in the LHC. The bunch spacing is 25 ns, which results in the collision rate of 40 MHz. The injection chain of the LHC consists of the Linac 2, the Proton Synchrotron Booster, the Proton Synchrotron and the Super Proton Synchrotron accelerators. They provide protons of the energy of 50 MeV, 1.4 GeV, 25 GeV and 450 GeV respectively [141].

Experiments are located at four beam intersection points (Fig. 8.1). ATLAS and CMS are general-purpose experiments with 4π detector geometry. LHCb is a forward detector designed to study B-physics while the focus of ALICE is on ion-ion collisions.

8.3 The ATLAS Experiment

ATLAS [142] is a general purpose particle physics experiment at the Large Hadron Collider. Its task is to detect and measure particles that emerge in collisions delivered by the LHC. The detector design was driven by benchmark physics processes such as searches for the Higgs boson in various decay channels, SUSY and new heavy gauge bosons searches, tests of extra dimension models as well as of quark and lepton compositeness models, and precision measurements of Standard Model processes.

The general layout of ATLAS is similar to other collider detectors and is shown in Fig. 8.2. Closest to the interaction point a tracking system is located. It detects charged particles and measures their charge and momentum. It is surrounded by the solenoid magnet which provides a 2 T magnetic field and allows momentum measurement by the tracking system. Outside the solenoid, the calorimeter system is placed. It measures the energy of particles and jets and consists of electromagnetic and hadronic sections. The outermost component is the muon system which enables detection of muons. In the following, more details on these subsystems are given.

- **Tracking system** (Inner Detector, ID) includes the Pixel detector, the Semiconductor Tracker and the Transition Radiation Tracker. The Pixel detector is the innermost component of the ID. It consists of three cylindric pixel layers (barrel region) and three wheels at each end of the barrel part (endcap region) thus providing three measurements of $R - \phi$ and z coordinates[1] per track for $|\eta| < 2.5$. A more detailed description of the Pixel detector is given in chapter 8.4.

[1]The ATLAS coordinate system is a right-handed cartesian system with its origin in the nominal interaction point. The X-axis is pointing to the LHC ring centre while the Y-axis is pointing upwards. The Z-axis is along the beam direction. The ϕ and θ angles are defined as usual in spherical coordinates. The radial distance from the beam line is denoted as $R = \sqrt{X^2 + Y^2}$. The pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$.

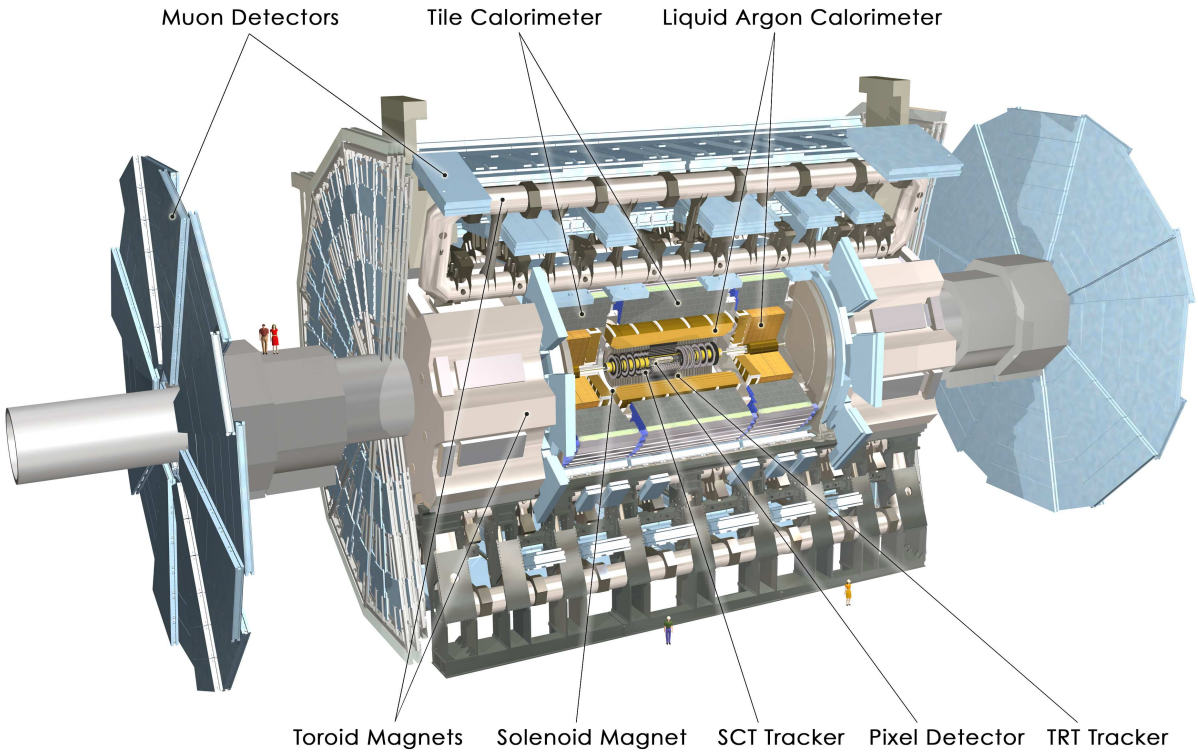


Figure 8.2: The ATLAS detector [142].

The Semiconductor Tracker (SCT) surrounds the Pixel detector. It consists of eight cylindric strip layers (barrel region) and nine end-cap wheels on either side of the barrel region providing at least four space-point measurements per track. In the barrel part, there are 40 mrad stereo strips allowing measurement of both $R-\phi$ and z coordinates and also a set of strips running parallel to the beam direction in each layer which provides only $R-\phi$ measurement. In the end-cap wheels there is a set of strips running radially and a set of 40 mrad stereo strips. The pitch in both barrel and end-cap regions is approximately 80 µm. This gives the resolution per module of 17 µm ($R-\phi$) and 580 µm (z) in the barrel region, and 17 µm ($R-\phi$) and 580 µm (R) in the end-caps. The total number of channels is around 6.3 million. The SCT covers the region of $|\eta|$ <2.5.

The outermost part of the tracking system, the Transition Radiation Tracker (TRT), is composed of 4 mm straw tubes. It provides only $R-\phi$ information with an accuracy of 130 µm per straw. In the barrel region, straws are parallel to the beam, while in the endcap part they are placed radially. The angular coverage is $|\eta| < 2$.

- **LAr Electromagnetic Calorimeter** provides energy measurements of electrons

and photons. It is divided into a barrel part ($|\eta|$ <1.475) and two end-cap components (1.375< $|\eta|$ <3.2). The EM calorimeter is a lead-LAr detector with accordion-shaped kapton electrodes and lead absorber plates over its full coverage.

- **Hadronic Calorimeter** is located directly outside the EM calorimeter. Its barrel covers the region $|\eta|$ <1.0 and its two extended barrels the range 0.8< $|\eta|$ <1.7. It is a sampling calorimeter using steel as the absorber and scintillating tiles as active material.

- **Muon System** consists of an air-core toroidal magnet and muon detectors. The magnet provides strong bending field without the degradation of resolution due to multiple scattering effects. Muon chambers are organised in three layers in the barrel region and three wheel layers in the endcap region. They cover the region of $|\eta| < 2.7$.

8.4 ATLAS Pixel Detector

The Pixel detector [142–145] is the innermost component of ATLAS. It consists of three layers in the barrel part and of six endcap wheels, three on either side of the interaction point (IP). Thus, it supplies three high-precision space-point measurements per track. The Pixel detector is crucial for the tracking performance of ATLAS. Thanks to the high granularity and closeness to the IP it provides the necessary pattern recognition capabilities in the high particle multiplicity environment of LHC.

One of the crucial ingredients to reach the goals outlined in Section 8.1 are the b-tagging capabilities. For example, production of SUSY particles might result in several high-p_T b-jets [144], hence efficient b-tagging is essential. Another requirement is good secondary vertex reconstruction. It is especially important for the B-physics program: search for CP-violation in B-decays, B_S^0-mesons mixing, rare B-decays and B-hadron spectroscopy [144]. The b-tagging and secondary vertex capabilities are largely determined by the Pixel detector. Additionally it provides excellent reconstruction of the primary vertices. At nominal luminosity, there are on average 25 pile-up interactions for every bunch crossing, each potentially resulting in a separate primary vertex. The Pixel detector allows resolving these vertices.

The principle of a semiconductor pixel detector is shown in Fig. 8.3. A reversely-biased pn-junction acts as a sensitive element (sensor). A charged particle traversing the sensor releases free charges due to ionisation. They drift in electric field and lead to an electric pulse which is detected by the readout circuitry. The readout electronics chip is connected to the sensor with a metal bond (bump-bond). This type of pixel is called *hybrid* due to the fact that readout chip and sensor wafers are built separately and later connected with a bump-bonding process.

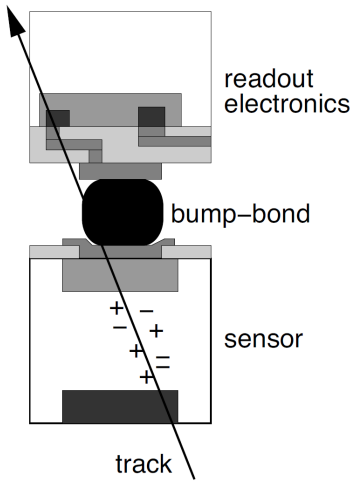


Figure 8.3: The detection principle of the hybrid pixel detectors [143]. The electronic readout chip is connected to the sensitive element via bump-bonding.

Fig. 8.4 shows the overall layout of the detector. The three cylindric layers in the barrel region have average radii of 50.5 mm, 88.5 mm and 122.5 mm. The inner layer is also called the B-layer. The six endcap disks, three on each side of the IP, are placed perpendicularly to the Z-axis at $Z = \pm 495\,\text{mm}$, $Z = \pm 580\,\text{mm}$ and $Z = \pm 650\,\text{mm}$. The basic building block of the detector is the so-called module. Identical modules are mounted on staves in the barrel region and on sectors in the endcap. A module consists of three major elements [143]:

- **The sensor** is a high resistivity n-type silicon wafer with implants on both sides. The thickness of the wafer is 256 µm. On the front side it contains a grid of 47232 n^+ implants organised in 144 columns and 328 rows. The implant size in 128 columns corresponds to a pixel pitch of $400 \times 50\,\mu\text{m}^2$, while in other 16 columns the pitch is $600 \times 50\,\mu\text{m}^2$ (*long* pixels). The back side of the wafer contains a p^+ implant which provides a pn-junction.

- **The readout electronics** consists of 16 identical FE-I3 [146] front-end chips. Each chip has identical 2880 readout channels organised as a matrix of 18×160 pixels of $400 \times 50\,\mu\text{m}^2$.

- **The flex-hybrid** is a flexible printed circuit board with a thickness of around 100 µm. It contains elements that control operation of the front-end chips, such as the Module Control Chip (MCC) [147] which reads out data from the chips, builds events and sends them outside the detector via the opto-links.

The three elements are are connected to each other as shown in Fig. 8.5. Front-end chips are bump-bonded to the sensor tile on the negative implant side, the flex-hybrid is

glued to it on the other. There are 1456 identical modules in the barrel region and 288 in the endcap resulting in more than 80 million readout channels.

8.4.1 Limitations of the Pixel Detector

Despite its superior features, the Pixel Detector has also certain limitations. They include the occupancy problem, radiation damage and hardware lifetime. These issues are briefly explained in this section.

The readout system of the Pixel detector was designed for the nominal luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. At higher luminosities, the number of pile-up events per bunch crossing and hence the particle multiplicity increase, which leads to an occupancy problem: the front-end chip FE-I3 is not able to process as many hits anymore, and data loss occurs. Simulations show that the B-layer inefficiency due to this effect is around 9 % for luminosities of three times the nominal one, and more than 95 % at ten times the nominal one [148,149]. Additionally, the link between the Module Controller Chips (MCC) and the off-detector electronics will cause data losses at two or more times the nominal luminosity due to the limited bandwidth [147].

Another effect deteriorating the Pixel detector performance is radiation damage. The components that are most affected by this include the sensors, the front-end chips and the MCC [149]. The sensor is sensitive to the displacement damage which is conventionally expressed as equivalent damage of a fluence of 1 MeV neutrons ($\mathrm{n_{eq}/cm^2}$). It leads to increase of the noise due to increased leakage current, decrease of the charge collection efficiency due to traps and increase of the depletion voltage. The readout electronics is affected due to appearance of positive charges at the gate oxide (surface damage). It is sensitive to the total ionising dose deposited in the relevant component (expressed in Gy). The estimated lifetime of the B-layer in terms of the integrated luminosity is around $300\,\mathrm{fb}^{-1}$ at 14 TeV which corresponds to 0.66 MGy and $1.6 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ accumulated at this layer [149]. Effects of radiation damage were already observed in 2011 [150, 151].

Both occupancy and radiation damage problems affect mostly the B-layer due to its proximity to the interaction region. This has dramatic consequenses on the tracking performance, hence on the physics potential of ATLAS. Additional problem is the hardware lifetime: occasional failures of single pixels, front-end chips or modules will happen with time.

Clearly, an upgrade of the Pixel detector is needed in order to fully profit from high instantaneous and integrated luminosities that LHC will provide.

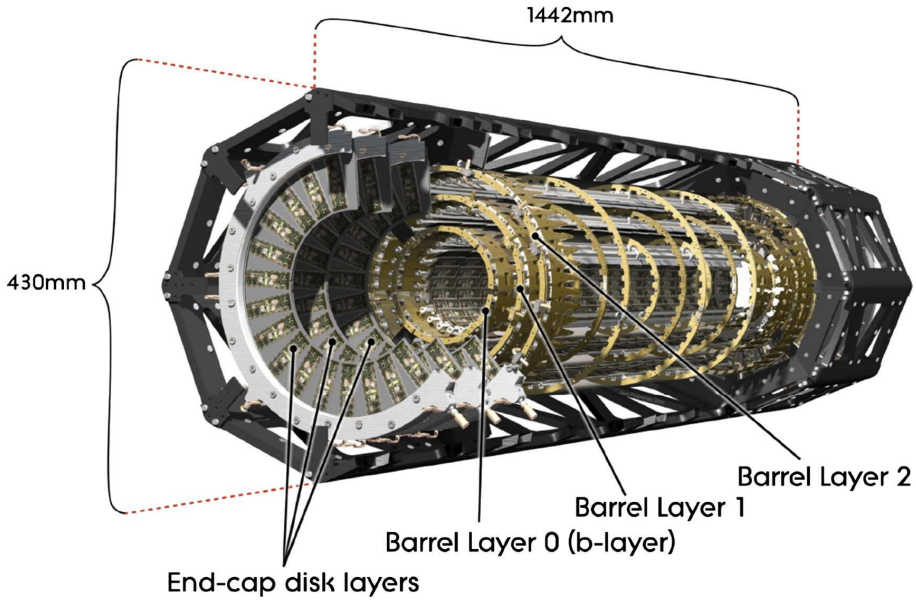


Figure 8.4: The ATLAS Pixel detector [143].

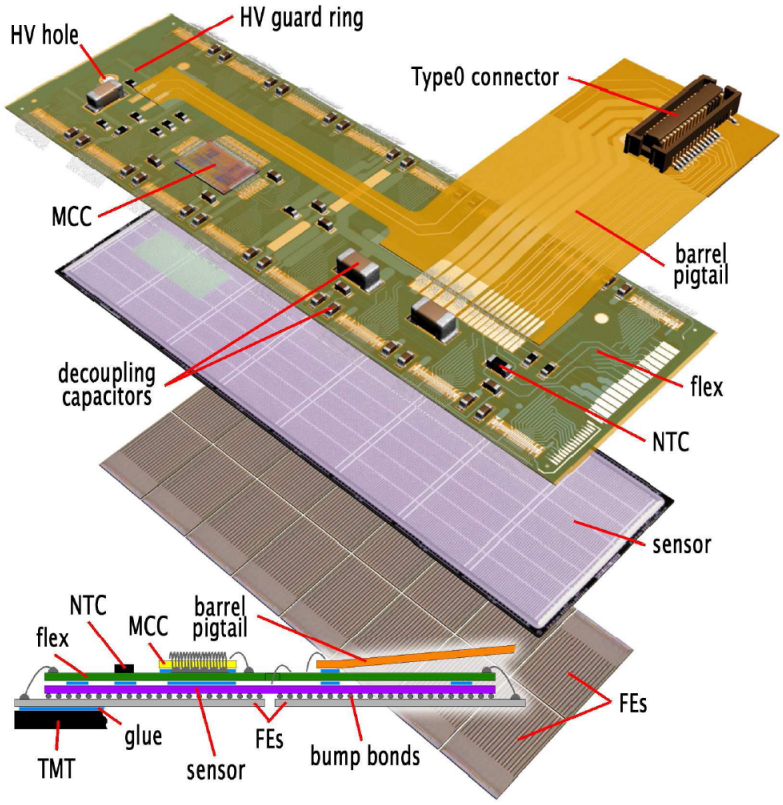


Figure 8.5: The building blocks of a module [143].

8.5 Insertable B-Layer

The Insertable B-Layer (IBL) [149] is an upgrade project of the present ATLAS Pixel detector which is designed to overcome the limitations of the present Pixel detector listed in the previous section and to further improve tracking capabilities. IBL implies the insertion of one additional hybrid pixel layer between the existing Pixel detector and the new, smaller beam pipe. It is expected to be accomplished during the first LHC Long Shutdown (LS1) in 2013.

By providing another measurement point close to the IP it restores the tracking performance in case of complete or partial failure of the current B-layer. It will have a finer granularity than the present pixel detector which leads to better resolution and lower occupancy. The sensors will be read out by the new front-end chip, the FE-I4 [152], which has a different architecture compared to FE-I3 and remains fully efficient even at high occupancies.

At the time of this study, the sensor technology was not yet chosen. Three candidates were considered [149]:

- **Planar Sensor.** The principle is the same as for the present Pixel detector. It is a well proven technology with high production yield. The disadvantage of this sensor type is that it requires high bias voltage (up to 1 kV) after irradiation to ensure full depletion. The reference design is n-on-n as for the present detector, but other options were also proposed [149].

- **3D Sensor** is a novel technology. It has a more complicated electrode structure than a planar sensor: they are vertically implanted deeply in the sensor bulk. As a consequence, lower bias voltage is needed to deplete the sensitive volume (below 100 V before the irradiation and around 200 V after a dose of $10^{16}\,\mathrm{n_{eq}/cm^2}$ [149]). Moreover, the dead area at the edge of the sensors is reduced (active edges). The disadvantage of the technology is that it is not yet mature, and high production yield has not been demonstrated yet.

- **Diamond Sensor** is very different from the planar or 3D: instead of a semiconductor, it uses diamond crystal as a sensitive element. Diamond has larger gap which leads to lower signals, but the noise is also reduced. This technology was the least mature of the three in terms of production yield at the time of this study.

8.5.1 Sensor Technology Choice

A wide program of sensor characterisation was defined in order to provide input to the decision on the sensor type to be used for IBL [149]. An important ingredient of this program were the test beam measurements. At the test beam one can study the

response of detector prototype to the beam of high energy particles. Test beams allow measurements of pixel detector properties such as single hit efficiency, single hit resolution, charge collection efficiency, charge sharing, cluster size, Lorentz angle etc. These quantities have direct influence on the tracking performance of the final detector and hence are very important. The simultaneous characterisation of different technologies in the same experimental setup and conditions assures comparability of results.

The test beam program that was defined for the IBL project consisted of three main campaigns: the first combined test beam with FE-I3 chip at CERN SPS, the first beam test of FE-I4 chip at DESY, and the final test beam campaign with magnetic field at CERN SPS. The EUDET Telescope (see Section 8.6 for detailed description) was chosen as external tracking device, which is an excellent tool for such applications.

This thesis contributed to preparations for the test beam program, namely to the extension of the data analysis software package EUTelescope (which is introduced in Section 8.7), to the integration of the USBPix Data Acquisition System (see Section 8.10) into the EUDET telescope as well as to the first IBL test beam data taking and data analysis.

8.6 EUDET Telescope

The EUDET Telescope is a tracking device designed for detector prototypes characterisation at the test beam. This section describes the telescope in its final configuration, as it was used in the measurements of this thesis. For descriptions of the system at earlier stages of the development see references [153–155].

8.6.1 Overview of the System

The active area of the telescope consists of six Mimosa26 [156, 157] pixel sensors (see section 8.6.2) placed perpendicularly to the beam, hence up to six measurement points per track are provided. A sensor resides in a mechanical structure called plane. The planes are grouped into two arms, with three planes in each arm, see Fig. 8.6. The distance between the arms and between planes within an arm can be adjusted independently. A detector prototype to be characterised - Device Under Test (DUT) - can be placed between the arms and also behind the telescope.

The principal structure of the system is shown in Fig. 8.7. Data from the sensors are read out by custom VME cards called EUDRB [158, 159]. There is one card for each sensor. The EUDRBs send data to a single board computer via the VME backplane, where data from all sensors are merged and sent to the Data Acquisition Computer (EUDAQ PC) for storage. Data readout and storage is triggered by a dedicated Trigger Logic Unit (TLU) [160]. Four fast scintillators, two upstream and two downstream of the telescope, are used as input. The system is controlled by the DAQ software called EUDAQ [161]

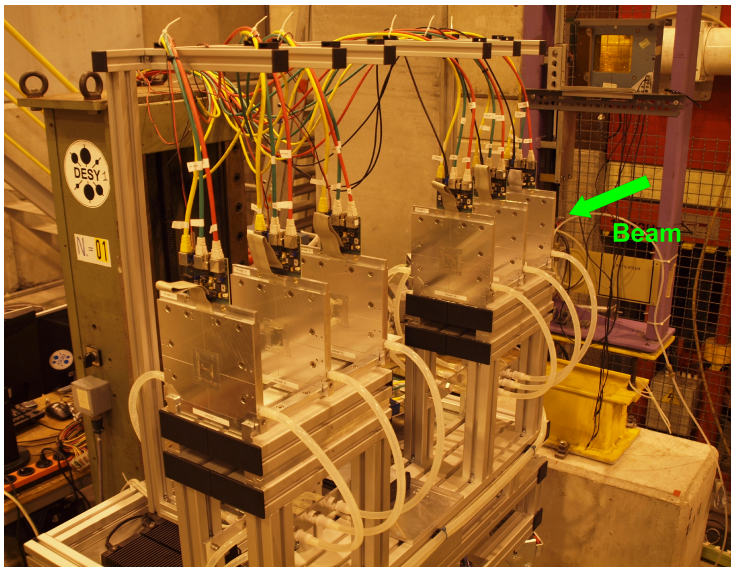


Figure 8.6: The EUDET Telescope at CERN SPS. Three planes in each of the two arms are visible. A sensor is in the middle of a plane. The beam comes from the right perpendicularly to the planes and penetrates all the sensors.

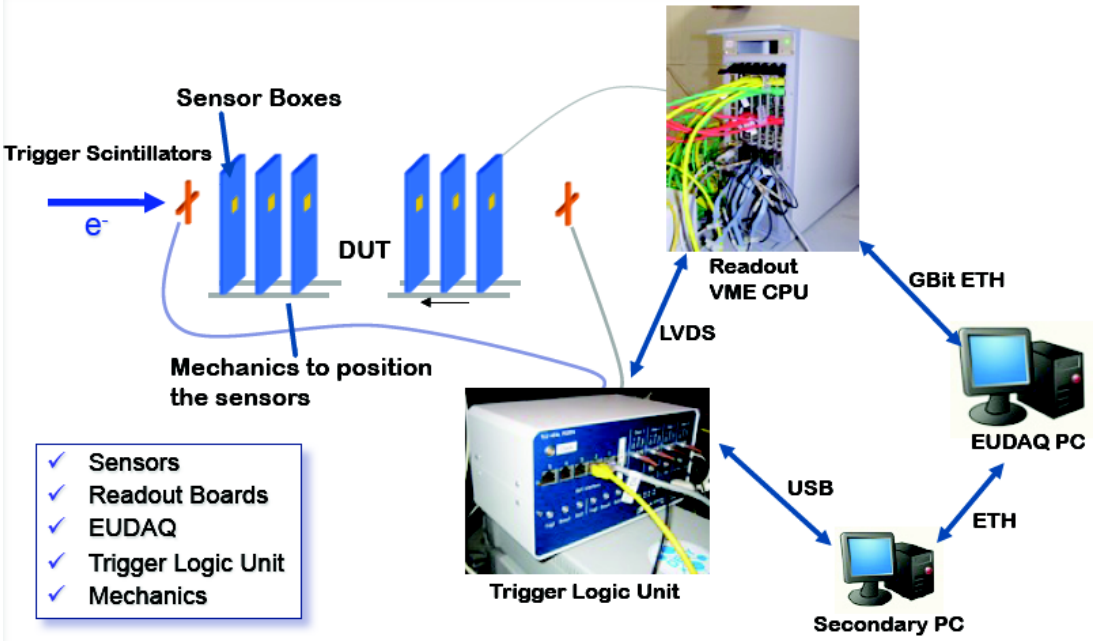


Figure 8.7: Main components of the EUDET Telescope.

which runs on the EUDAQ PC. It configures the hardware, starts and stops the DAQ, receives data from the telescope sensors and DUTs, builds events[2] and stores them to disk. The telescope components most relevant for this thesis are described in more detail below.

8.6.2 The Sensor

The Mimosa26 [156, 157] sensor acts as a sensitive element of the telescope. It is a Monolitic Active Pixel Sensor (MAPS) [162]. In MAPS, the front-end electronics and sensitive volume are built on the same silicon substrate, in contrast to hybrid pixel detectors where front-end is built separately and then connected to the sensitive part with bump-bonding technique. Charges released by ionising radiation diffuse thermally to the collecting electrodes.

The Mimosa26 chip consists of the pixel array itself and periphery logic as shown in Fig. 8.8. The pixel array contains 1152 columns of 576 square pixels with a pitch of 18.4 µm, resulting in an active area of $21 \times 10.6\,\mathrm{mm}^2$. Each pixel contains three transistors: for diode reset, pixel switch and a source follower to buffer the collected charge. Each column is terminated by the end-of-column discriminator.

The sensor is read out in the so-called rolling-shutter mode. Row by row is constantly selected, and all columns within are read out simultaneously. The preamplified signal from each column is sent to the end-of-column discriminator, where it is compared to a threshold. After the whole pixel array is read out, the zero-suppression logic on the

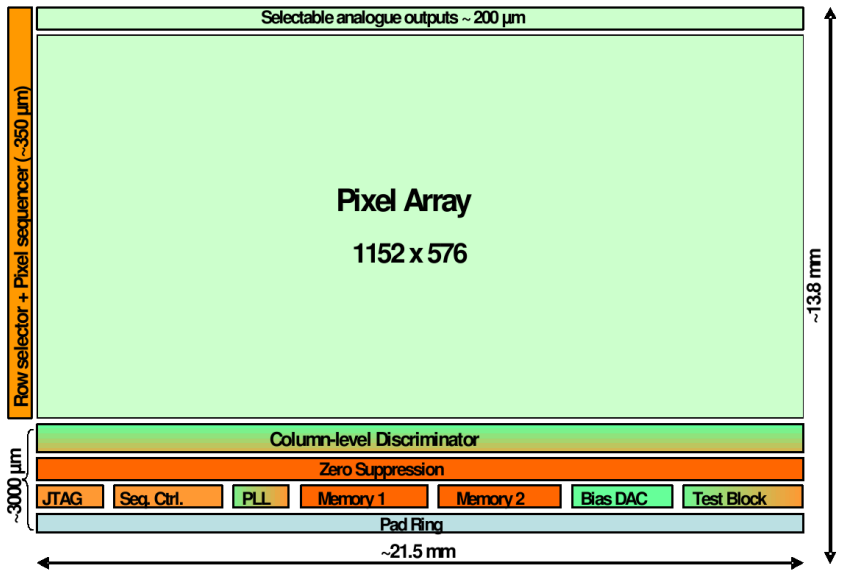


Figure 8.8: The Mimosa26 chip architecture [156].

[2]See Section 8.6.3 for the definition of event.

periphery searches for the hit patterns. This allows a data reduction factor in the order of 10 to 1000. The sparsified data are stored in two memory blocks before being sent out of the chip for storage. The readout time of the whole pixel array is 112.5 µs.

8.6.3 Triggering

The data readout and storage is initiated (*triggered*) by the Trigger Logic Unit (TLU) [160], a custom-built device designed specifically for the EUDET test beam telescope. A charged particle traversing the active area of the telescope leads to signals in the upstream and downstream scintillators which are then sent to the TLU. An arbitrary logical function of these signals can be chosen as a trigger condition, e.g. the coincidence of all four scintillators or the coincidence of at least three of them. If such a condition is fulfilled, the TLU triggers the read out and storage to disk; data that were read out between two successive triggers form an *event*.

Upon a trigger receipt, the TLU starts communication with detectors that take data (the telescope sensors and the DUTs), the so-called *handshake*. Each detector must be connected to one of the six 8-pin RJ45 LVDS interfaces on the TLU. An interface provides TRIGGER, TRIGGER CLOCK, BUSY and RESET lines. TRIGGER and RESET are output lines to a detector, while TRIGGER CLOCK and BUSY are inputs from it. The exact meaning of the signals on these lines is defined by the handshake mode. Three modes are possible:

- **No handshake.** This is the most simple communication mode, in which the TLU does not expect any response from a device. Upon a trigger receipt, the TLU raises the TRIGGER line for a fixed amount of time. After it is de-asserted, the system is ready for triggers again.

- **Simple handshake.** In this case the TLU expects a response from the detector – it waits while it finishes the read out, i.e. accounts for possible variations of the dead time. Once a trigger is encountered, the TLU raises the TRIGGER line. The device receives this and in turn asserts the BUSY line. On receipt the TLU de-asserts TRIGGER. As soon as the detector finishes the read out it de-asserts BUSY and the system is ready for triggers again.

- **Trigger data handshake.** This is the most advanced communication mode. Additionally to the functionality of the simple handshake, the TLU also sends the trigger number to the detector. This ensures synchronisation between different subdetectors, e.g. between the telescope and the DUT. After a trigger is received, the TLU raises the TRIGGER line, the device asserts BUSY and TLU de-asserts TRIGGER, just as in the simple handshake. Then the 15-bit trigger number is transferred. In order to obtain each bit, the detector must toggle the TRIGGER CLOCK line. On receipt of a pulse on this line, the TLU sends one bit via the TRIGGER line. After

all bits are received, the detector de-asserts BUSY, the TLU sets the TRIGGER to logical low and the system is ready for triggers again.

Figure 8.9 illustrates the simple and the trigger data handshake modes. If each subdetector works in one of these two modes, the telescope will be ready for a new trigger only when all devices finish their readout, hence the slowest system determines the maximum overall event rate. Communication between the TLU and the telescope sensors DAQ is implemented in the trigger data handshake. This mode is also recommended for the integration of new systems such as the FE-I4 readout, since it is the most robust of the three.

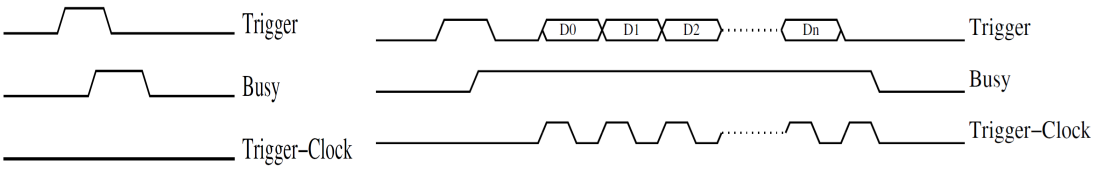


Figure 8.9: Illustration of the simple handshake (left) and the trigger data handshake (right) [160].

8.6.4 EUDAQ - Data Acquisition Software

EUDAQ [161] is a platform independent data acquisition software package developed for the EUDET Telescope. It is written in an object-oriented way in C++, has a modular structure and allows an easy integration of the DUT. The EUDAQ consists of several components as shown in Fig. 8.10. Each component is an independent application which can run on the same or different computers. Applications communicate with each other with the TCP/IP protocol.

- **Run Control** is the main component of the package. It steers all other modules by sending *commands* to them. Typical examples are *configure*, *start run* and *stop run* commands. A user can issue them via Graphical User Interface (GUI).

- **Producer** is an application that is associated to a piece of hardware that produces data, e.g. telescope or the DUT. Its task is to read data from a detector, convert it into EUDAQ internal data type and send it to the Data Collector (see below). Each detector has its own Producer. There is also a TLU Producer which for every event sends only a trigger number.

- **Data Collector** is responsible for the data storage. It receives data from all Producers, merges them to form events and stores to disk in binary format.

- **Monitor** serves for online data quality monitoring. It reads the collected data from disks, performs on-line reconstruction and displays it in a GUI.
- **Log Collector** receives the log messages from all EUDAQ modules and writes them to disk. Thus, all the logging information is stored centrally and can be easily retreived offline.

Typically Run Control, Data Collector, Log Collector and Monitor are launched on the same computer (EUDAQ PC) which is in the control room. The TLU Producer runs on a PC which TLU is connected to (denoted as Secondary PC in Fig. 8.7). The producer that reads data from the telescope sensors runs on the VME single board computer which collects data from EUDRBs. The DUT producers typically run on dedicated computers where their DAQ software operates.

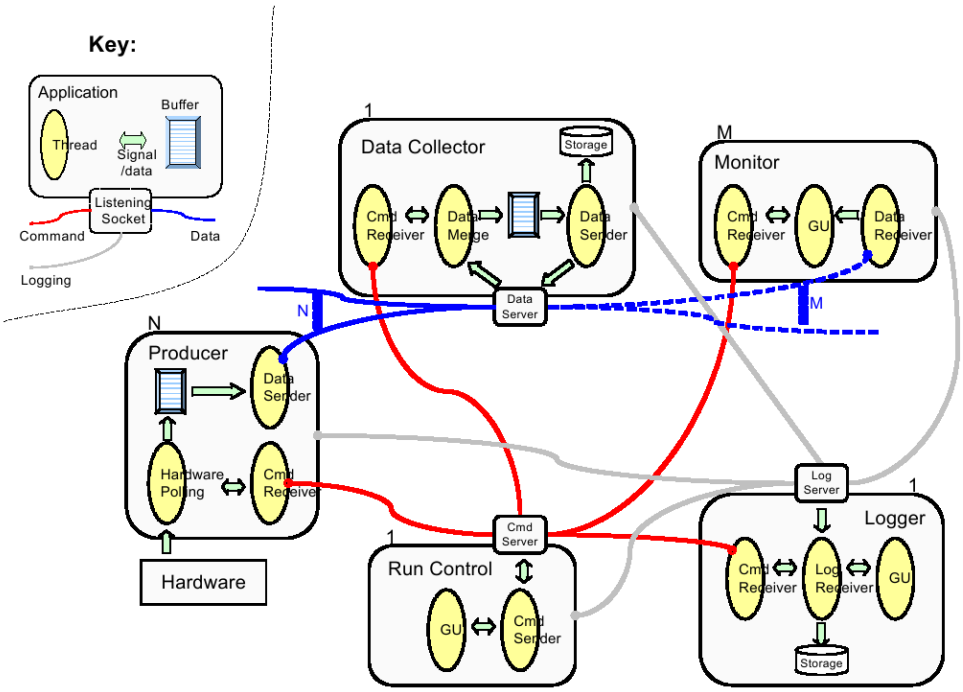


Figure 8.10: Schematics of EUDAQ [161].

8.6.5 Performance of the System

The parameters of the Mimosa26 sensor depend on discriminator threshold value, which is usually given in terms of the noise RMS, i.e. a pixel is accepted for processing if

$$S > kN \text{ or } S/N > k,$$

where S is signal amplitude, N is noise RMS and k is an integer number. Its typical values are $k \in [5, 10]$.

For the threshold of $S/N = 8$ the hit resolution of the Mimosa26 sensor is around 4.5 µm [156]. The track resolution of the telescope which combines measurements from up to six sensors is around 2 µm between the telescope arms. The efficiency of the sensor is about 98 % for this threshold value [156]. The gives a total tracking efficiency of $0.98^6 = 0.89$ if hits in all six planes are required. It can be increased by allowing tracks with missing hits in one or more sensors.

The event rate is limited by the VME readout of the telescope. If all EURDB boards are in a single crate, the rate is around 400 Hz. It increases by a factor of two if the boards are distributed among two VME crates.

8.7 EUTelescope Data Analysis Framework

The offline analysis of data collected with the EUDET Telescope is performed with a dedicated software framework called EUTelescope [163]. It performs all the analysis steps which are necessary to convert raw data stored to disk by EUDAQ to high level objects such as tracks measured by the telescope. It also allows the DUT characterisation. Since EUDET Telescope was initially designed as infrastructure for the ILC detector R&D, the EUTelescope was implemented as part of the ILC software, called ILCSOFT [164]. This dictates the data model that is used within the framework and its modular structure.

The data model is based on the LCIO persistency framework [165, 166]. It defines the basic objects such as detector components, hits, clusters, tracks, calorimeter islands and provides the file format to store these objects.

The modular structure means that every independent task, such as a track fit, is implemented as a separate module and is not embedded in a single program. Every module is a so called *Marlin Processor*, which is a C++ class with a well defined structure. Its methods are called internally by the Marlin [167] application. For every event, a processor accepts input data in LCIO format, such as a set of hits. The algorithm being executed by the processor, e.g. a track fit, is implemented in the *ProcessEvent* method of this class. The resulting objects - tracks in this example - are stored to the output file. Execution is steered by a set of configuration options stored in form of an XML file, called a steering file. It contains all important settings, such as number of events to be processed or the maximum allowed χ^2 of a track.

This modular approach allows better flexibility, e.g. one can easily try various options of a reconstruction algorithm without the need to repeat the previous steps. It also makes debugging easier, since one can directly identify the problematic module. Finally, the structure of all modules is the same which simplifies code development and maintanence.

The analysis steps that the EUTelescope performs, each being a separate Marlin Processor, are described in what follows.

8.7.1 Raw Data Conversion and Clustering

The data collected by the EUDET Telescope are stored in the EUDAQ native binary format. However, the EUTelescope employs the LCIO format, as discussed above. Hence, the first step of the data analysis is the format conversion.

The output of this processor is a set of objects of *TrackerData* type, which is one of the basic data types in the LCIO model [166]. A single object contains information that was read out from a single Mimosa26 sensor in a given event. Hence, six TrackerData objects will be stored to the output file for each event. The unique identification number of the sensor (sensor ID) is attached to this object which allows to distinguish data from different sensors during further analysis steps.

The pixel array data are stored to the *ChargeADC* member of TrackerData which is an array of floats. Three numbers are stored for each fired hit: its row number, column number and collected charge. Since Mimosa26 is a digital sensor with no charge information, the third number is irrelevant and is always set to unity. The numbers are stored sequentially to the array.

The next analysis step is the so called clustering. When a high energetic charged particle traverses the sensor, charges are released in the pixel that is traversed by the particle. However due to charge sharing, signals may occur in the neighboring pixels as well, hence several adjacent pixels may be fired due to a single ionising particle. Moreover, if a particle impinges not perpendicularly to the sensor it can penetrate and release charges in several pixels. Therefore a clustering algorithm has to be applied to a raw pixel map which merges neighboring pixels to a cluster. The latter is considered as a hit in further steps, and not every single pixel. Two clustering algorithms are available within the EUTelescope:

- **Fixed Frame Clustering** searches for clusters of hits of a certain size $m \times n$ pixels. A cluster cannot exceed that size.
- **Sparse Clustering** assigns hits to clusters based on the vicinity criterion: pixels are merged to a cluster if distance between them does not exceed certain value. A cluster can have arbitrary shape and size.

The output of the clustering processor is a set of cluster objects. Such entity keeps row and column numbers of pixels that form a cluster and provides the possibility to calculate its centre, i.e. the most probable incidence point of the ionising particle which caused this cluster. For the digital Mimosa26 sensor, where no charge information is available, the geometrical centre of a cluster is used. For sensors with charge measurement, a centre-of-gravity point [168] or the so called η-algorithm [169] can be used in order to improve the spatial resolution.

After cluster reconstruction the so called Hitmaker Processor is invoked. It translates cluster centre positions in terms of pixel rows and columns (possibly fractional) into space

points - hits, which are used later for track reconstruction. At this point, the geometry information is used, such as the pixel pitch or distance between the planes.

8.7.2 Alignment with MillepedeII

Even though the mechanical structures that hold the sensors (planes) are very robust, there are always small misalignments of the sensors, i.e. shifts or rotations with respect to their nominal positions. In order to fully profit from the excellent resolution of the Mimosa26 sensor, they have to be precisely aligned, i.e. alignment corrections (shifts and rotations) have to be determined and taken into account in the transformation from the row and column number to the space-point coordinate.

The so called track based alignment is performed in the EUTelescope. This method is based on a minimisation of track residuals with respect to alignment corrections. A track residual is the distance between the fitted track position and the measured hit position (see also sect. 8.8.2). If there is no misalignment, this distribution is centred at zero and its width reflects the resolution of the sensors. However, in case of misalignment, the distribution will be shifted from zero and its shape distorted. By taking into account alignment corrections the distribution can be restored to nominal one. Hence, corrections can be determined as those that give the residuals with minimal width and peak position closest to zero.

The tool that is used for this is called Millepede II [170]. It is a least square fit program that was designed specifically for alignment of high energy physics tracking detectors. It has been used extensively in large scale experiments such as CMS, H1, ZEUS or HERA-B [171–173]. The crucial feature of the algorithm is that it is able to perform a least-squares fit with very large number of parameters. In the track based alignment, there are two sets of parameters that have to be defined simultaneously in the fit: the track parameters (also called *local* parameters since they appear only in a subset of measurements – here for a single track measurement) and the alignment corrections (*global* parameters – they appear for all tracks). There is a large number of local parameters due to a large number of tracks that are needed for statistically presice results. The idea of the Millepede is to employ a special structure of the equation matrix and to redefine the problem in such a way that only global parameters are determined, since the local ones are not of interest [170].

Several alignment modes are available in the EUTelescope depending on the number of alignment corrections to be determined. For the Mimosa26 sensors, the shifts in a plane perpendicular to the beam and the rotation in this plane are determined. This approach neglects possible rotations around axes perpendicular to the beam but it is sufficient for a normal incidence angle. Additionally, rotations around other axes can be also defined. This is more important for tilted sensor studies.

After the alignment parameters are determined, the hit coordinates obtained during the Hitmaker Processor are corrected to take the misalignment into account.

8.7.3 Tracking

The *EUTelTestFitter* processor reconstructs telescope tracks. For every event it gets a set of measured hits in all sensors, corrected for misalignment, searches for patterns and determines the track parameters. Its output is a set of reconstructed tracks in form of *Track* objects of LCIO.

A dedicated track fitting method that takes into account multiple scattering effects [163, 174, 175] is used in the processor. For a given hypothesis, i.e. a set of hits that might potentially form a track, a χ^2-ansatz that contains additional terms allowing direction change in a layer due to multiple scattering is constructed. Thanks to a few simplifying assumptions, such as small beam spread and small scattering angles, the problem of χ^2 minimisation can be reduced to an analytical solving of a matrix equation.

The processor is controlled by several parameters, such as maximum χ^2 or minimum number of hits that belong to a track. In order to assure that secondary scattering is properly taken into accout, it is important to specify the geometry of the setup, in particular the thickness in radiation lengths of all layers, and the beam energy.

8.7.4 DUT Characterisation

The last analysis step is the DUT characterisation which is performed by a dedicated processor called *EUTelDUTHistograms*. As input it takes a set of reconstructed telescope tracks and a list of hits in the DUT. Its output is a set of histograms which show most important DUT properties.

For every event, the processor searches for hits in the vicinity of the track position at the DUT. If found, the distance between the track position and the measured hit (the residual) is calculated and filled into a histogram. The distribution directly shows the DUT resolution, assuming it is much worse than that of the telescope (which is usually the case). By comparing number of cases where a hit was found around a track to those where there was no hit, it is possible to measure the efficiency of the DUT.

8.8 Modification of EUTelescope

Several modifications to the EUTelescope framework were implemented during this thesis. They include:

- Integration of the ATLAS Pixel data type to the framework;
- Automation of the alignment processor;
- Development of the ATLAS Pixel specific DUT analysis processor.

These extensions allow the analysis of the ATLAS Pixel sensors within the framework and a more efficient workflow. They are described in more detail in this section.

8.8.1 Integration of ATLAS Pixel Data Type

The EUTelescope was designed as a generic tool that enables analysis of any Device Under Test (DUT). However, every DUT has its own specific features, which have to be accounted for in the framework. Even for the universal class of pixel detectors, there are distinctive properties for every particular sensor type like pixel pitch, matrix size etc. Especially the ATLAS pixels have certain specifics such as standard and long pixels (Sect. 8.4), which has to be considered at all stages of the analysis. Hence, the first step of the thesis was the insertion of the ATLAS Pixel data type into the EUTelescope framework.

In the EUTelescope, each pixel type can have a separate class that accounts for its unique features. In particular, it defines the data that are stored for every pixel. An instance of this class holds the information about a single pixel hit and is created whenever some actions on the pixel data are performed. It must inherit the *EUTelBaseSparsePixel* class. This simplifies the implementation, unifies its structure and makes possible to profit from inheritance features of C++, namely that large parts of the analysis code can be developed independently of the pixel type. For every hit, the ATLAS front-end FE-I3 provides the row number, the column number, Time-over-Threshold (ToT) which is related to the collected charge (see Section 8.9.3) and the Level-1 trigger ID (LV1). The class that handles Mimosa26 pixels (*EUTelSimpleSparsePixel*) has only three data members and hence is inadequate for the ATLAS Pixel. *EUTelAPIXSparsePixel* class was implemented for the ATLAS Pixel type which can manage all the information provided by the front-end chip.

In order to distinguish the pixel type, a unique identifier given by the *SparsePixelType* enumeration is attached to every object that handles pixel information. For example, a cluster object will contain this identifier in order to specify which pixel type it is composed of. Depending on the type, different algorithms must be invoked, e.g. reconstruction of the cluster centre position. A *kEUTelAPIXSparsePixel* variable was added to the *SparsePixelType* enumeration in order to denote the ATLAS Pixel type. It was propagated to all parts of code, which are specific to the pixel type.

As for the Mimosa26 case, the hitmap for every event for the ATLAS Pixel sensors is stored as a *TrackerData* object of LCIO (see Section 8.7.1). Due to more data members per pixel, the decoding of the underlying *ChargeADC* array, that is extraction of the pixel information from a sequence of numbers, had to be changed. This decoding is handled by the templated *EUTelSparseDataImpl* and *EUTelSparseData2Impl* classes. Depending on the pixel class that is passed to these decoders, a different decoding function is invoked which treats the array differently. Corresponding functions for the ATLAS Pixel type were implemented.

The fact that there are different pixel types (standard and long) has to be taken into account in the reconstruction of the cluster centre. A special cluster class *EUTelAPIXSparseClusterImpl* inherited from *EUTelSparseClusterImpl* was created, which properly reconstructs the centre-of-gravity of a cluster containing a long pixel.

8.8.2 Improvement of Alignment with the Correlation Information

The alignment is one of the most crucial parts of the EUDET Telescope data analysis. Due to the very high resolution of the Mimosa26 sensor, even a small deviation of few microns from the nominal position of a sensor leads to a significant reduction of the tracking precision. Hence it is necessary to determine exact layer positions, before the track fit is performed.

The most important issue for the alignment is the selection of real tracks with small combinatorial background as input for the Millepede algorithm (Section 8.7.2). In the standard EUTelescope, this is done by adjusting the *cone* and the *residual* cuts for every layer that needs to be aligned. The procedure is illustrated in Fig. 8.11. Track candidates are formed by considering all possible hit combinations with exactly one hit from every plane. For every hit in a given combination, except one in the last layer, a cone around the Z-axis[3] is defined (Fig. 8.11(a)). If the hit in the next layer is outside this cone, the combination is rejected. For each combination that remains, a straight line track is fitted. The residuals – distances between the fitted track position and the hit position (see Fig. 8.11(b)) are calculated and filled into the histograms separately for each layer and for both directions (X and Y).

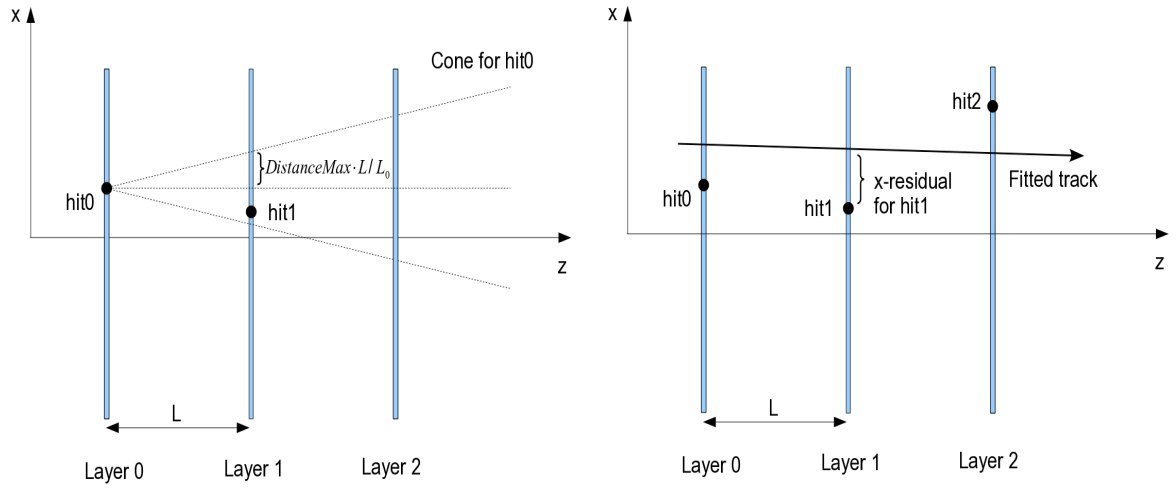
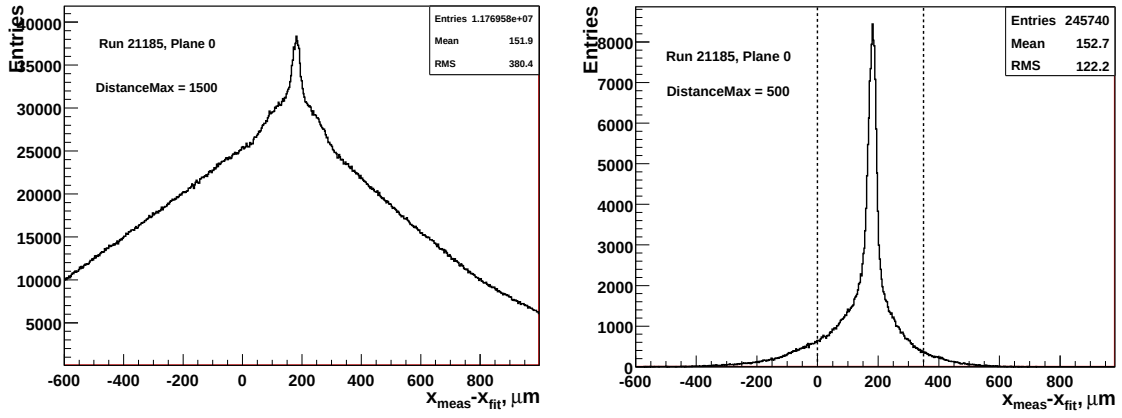


Figure 8.11: Illustration of the cone cut selection (a) and track residual definition (b).

[3]The coordinate system of the EUDET Telescope is usually defined in the following way: the surface of the first Mimosa26 sensor (Layer 0) is taken as an XY plane. X-axis points along the columns and Y-axis – along the rows of the sensor. The Z-axis is perpendicular to this plane and is close to the beam direction.

The cone cut is defined by the *DistanceMax* option which is (up to a constant, L_0) the tangent of the cone half angle (see Fig. 8.11(a)). It has a dramatic effect on the amount of background among the selected track candidates. Fig. 8.12 shows typical residual distributions in the first layer for the X-direction for cone cut values of DistanceMax=1500 and DistanceMax=500 (corresponding to the cone half angle of around 0.9° and 0.3°, respectively). With the looser cut, combinatorial background dominates the distribution, while with the tighter one, background is significantly suppressed and a peak corresponding to genuine tracks becomes pronounced. In order to further suppress the background, it is necessary to impose residual cuts, as determined from the peak position and its width (in this case approximately from 0 to 350).



(a) (b)

Figure 8.12: Example of residual distribution with loose cut (a) and tight DistanceMax cut (b). Residual cuts as determined from this distribution are shown as dashed lines.

Obviously this procedure of track selection is very inefficient: the cone cut as well as residual cuts for each layer and for both X and Y directions have to be determined in order to assure good track selection. Moreover, the positions of the sensors may change with time during a test beam due to changes in setup. Determination of the residual cuts has to be redone in this case. If one has to modify the setup frequently, the data analysis is too time consuming.

However, one can greatly improve the alignment procedure with the usage of the so-called *correlation plots*. For any pair of two sensors and a given direction (X or Y) a correlation plot is obtained by considering all pairs of hits in these two sensors and filling a two dimensional histogram with hit coordinates from the first plane on one of the axes and from the second plane on the other one. If hits are only due to real tracks, a clear correlation appears. Additionaly, there will be a smooth distribution due to combinatorial background. From such a plot, it is immediately visible whether detectors are aligned or not. In case of a perfect alignment, the correlation line points to the coordinate origin

(0, 0) and has a slope of 45°. A deviation from this line is a measure of misalignment. Fig. 8.13 shows an example of the correlation plot between two telescope planes. A clear correlation which corresponds to real tracks is observed. It can be readily deduced that there is no significant misalignment.

Based on such a correlation plot for two sensors, the allowed hit position in one of them can be predicted from the hit position in the other one. This may serve as a replacement for cone and residual criteria for the track selection. Such a procedure was implemented to the alignment processor *EUTelMille*: only those hit combinations which fall within the *correlation band*[4] in both X- and Y-directions are considered. Correlation plots are obtained for combinations of all sensors with the first one, and based on the hit position in the first plane, the allowed positions in all subsequent planes are determined.

In this way it is possible to select tracks for alignment without choosing the cone and residual cuts. Fig. 8.14 shows the residual distribution obtained with this method. It shows very clear peak and almost no background, which suggests that a sample of real tracks was obtained. No manual adjustment of residual cuts or the DistanceMax cut was done.

A similar procedure was also implemented in the fitter processor. In the standard EUTelescope, the hit selection was based on considering all possible hit combinations. With high occupation, the number of combination increases greatly and slows down the fitting procedure. However if one only considers hits that fall into the correlation band, the number of combinations reduces greatly.

8.8.3 Iterative DUT Alignment

The Millipede II is a powerful tool which allows a simultaneous alignment of the telescope and the DUT sensors. However, there are several effects that can spoil this one-step procedure. The resolution of the DUTs is typically much worse than of the telescope. Hence, the alignment precision will be reduced if the DUT sensors are included in the fit or special care has to be taken to account for the different resolution. The more important issue is that some of the DUTs might be problematic and not show any signal, or be badly aligned with the telescope sensors. In this case no or very few tracks will be selected for the alignment, which also spoils its precision or makes it impossible.

An iterative alignment procedure was developed in order to solve these problems. As a first step, telescope sensors are aligned with no DUT sensors included in the track selection, and resulting alignment corrections are applied. Afterwards, each DUT sensor is aligned separately. During this stage, the telescope alignment constants are not altered

[4]The correlation band is defined as the area between two straight lines which envelop the concentration of points in a correlation plot, such as shown in Fig. 8.13. The vertical separation between the envelopes (the *width* of the correlation band) can be specified during the analysis, the default value was set to 400 µm. A procedure which automatically determines the position of the correlation band from a correlation plot, based on maximisation of entries within the band, was developed.

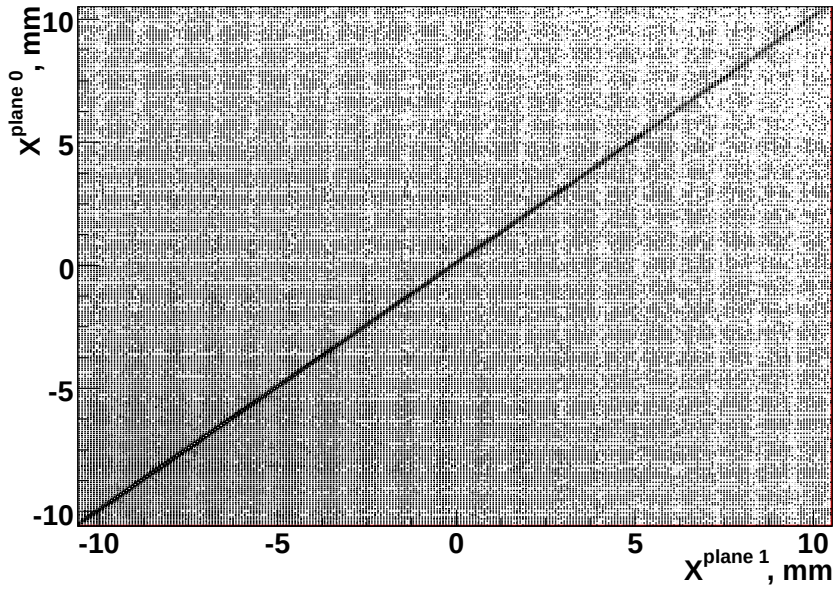


Figure 8.13: Correlation plot for two adjacent telescope sensors (plane 0 and plane 1).

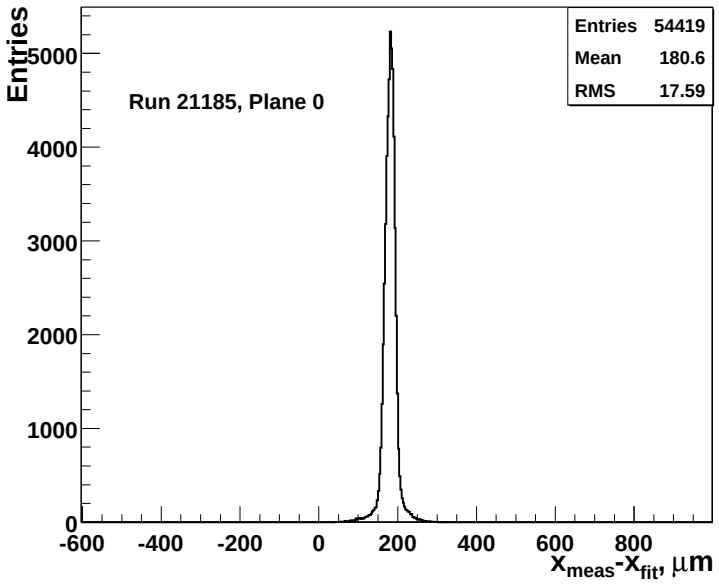


Figure 8.14: Residual distribution for track selection based on the correlation band method.

– they are kept fixed in the Millepede fit. In this way DUTs do not influence the alignment of the telescope. After each step, the alignment constants are applied to correct the hit coordinates of the DUT sensor under consideration. Ultimately, a set of aligned telescope and DUT hits is obtained which is ready for further analysis steps. A dedicated BASH[5] script which performs this task step by step for any number of the DUT sensors was created. The script also performs sequentially all necessary analysis steps described in Sect. 8.7 and hence automates the entire procedure.

8.8.4 ATLAS Pixel Analysis Processor

Finally, the DUT analysis processor *EUTelDUTHistograms* was extended in order to allow characterisation of the ATLAS Pixel sensors within the framework. Various histograms, specific to this sensor type, such as hitmaps, ToT, LV1 or cluster size distributions were added to the processor.

The extrapolated track position at the DUT is given in the telescope coordinate system (see the footnote on page 144 for the definition). Hence the transformation from this frame of reference to the local coordinate system of each sensor was implemented, which gives a possibility to find the exact track position within the sensor and also within the pixel. In order to perform this transformation, it is in particular important to keep track of all alignment corrections that were applied to every sensor, since they have to reverted. The telescope resolution is much better than the pitch of the ATLAS Pixel sensors, hence various *in-pixel* distributions, i.e. dependence of certain quantities on the track impact position within the pixel can be studied. In-pixel distributions such as the collected charge or charge sharing probability were implemented in the processor.

8.8.5 Summary

The EUTelescope data analysis software package was extended within this thesis, which allowed an efficient analysis of the ATLAS Pixel test beam data collected with the EUDET Telescope. The improvements include the integration of the ATLAS pixel type to all parts of the framework, automation of alignment and performance improvement of the fitter by usage of the correlation information as well as development of the histogramming processor, which calculates and visualises important detector quantities such as efficiency or charge sharing probability specifically for the ATLAS Pixel sensors.

This extension is the basis for data analysis presented in the following section. The code was also partially used for the official IBL test beam data analysis software release.

[5]BASH - a command processor and a scripting language for UNIX operating systems.

8.9 First IBL Test Beam Campaign

The first IBL test beam campaign took place in October/November 2010 at CERN SPS. It was a starting point for a wide IBL test beam program (see Sect. 8.5.1). The main goal of this campaign was to test whether all hardware and software tools work as expected and to collect the necessary experience from planning to results, that is to prepare for the final IBL test beam. The test beam time was divided in two phases. In the first one (Phase I), unirradiated Planar Sensor and 3D Sensor (both with FE-I3 front-end chip, since FE-I4 was not yet available) were tested simultaneously in a single experimental setup and under the same external conditions. The second phase, Phase II, was devoted to a first test beam with the USBPix DAQ system (see section 8.10) and to more detailed studies of various sensor designs within the PPS and 3D R&D groups.

This section describes the Phase I of the test beam. The experimental setup, offline data analysis and results are discussed.

8.9.1 Experimental Setup and Samples

The beam test was performed in the H6B hall at CERN SPS where the EUDET Telescope was installed. Fig. 8.15 shows the experimental setup. The ATLAS Pixel devices under test were placed inbetween two telescope arms. The pion beam of 120 GeV comes from the right.

Two sensors of different technology were tested:

- 3D sensor, full 3D, Stanford production, 200 µm thick (denoted as 3D)
- Planar sensor, n-in-p, CiS production, 300 µm thick (denoted as PPS)

Additionally, a third sensor of Planar type was inserted in the beam to serve as an extra reference plane for timing.

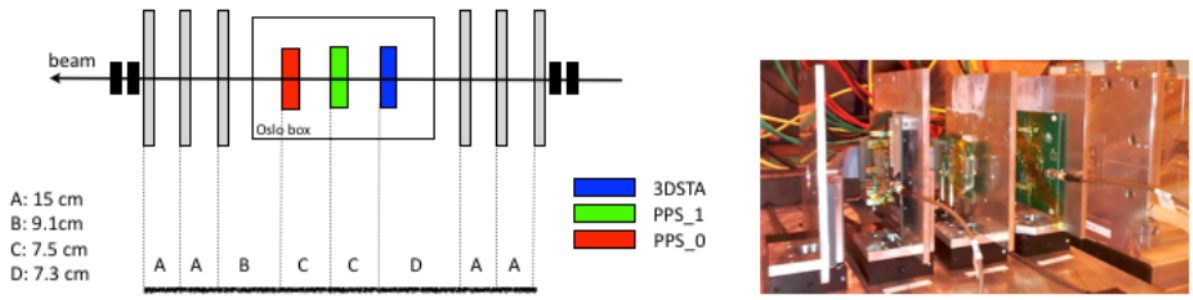


Figure 8.15: Experimental setup for the October 2010 IBL test beam. Three devices under test are placed inbetween the telescope arms: a 3D sensor (3DSTA), a Planar sensor (PPS_1) and an additional Planar device for timing purposes (PPS_0).

The EUDET Telescope configuration was as described in Sect. 8.6. The trigger condition was a coincidence of all four scintillators. The distances between the sensors are shown in Fig. 8.15.

The sensors were tested at 0°, 10°, 15°, 20° sensor tilt with respect to the normal track incidence. At least 2 million events were taken for each angle. In this chapter, results for the normal incidence angle are shown.

8.9.2 Data Analysis

The modified EUTelescope (see section 8.8) was used for the offline data analysis. The clustering algorithm was chosen to be *SparseClustering2* both for the Mimosa26 and ATLAS pixel sensors. The pixels in a cluster were required to be adjacent, i.e. the maximum distance between two pixels, controlled by *SparseMinDistance* parameter, was set to 0. The alignment was done with the iterative procedure based on the correlation information as described above. At least four hits in the telescope were required for the track fit. A $\chi^2 < 40$ cut was imposed.

Figure 8.16 shows the number of tracks distribution. Around 87 % of events contain at least one reconstructed track. Due to high intensity of the beam and relatively long integration time (around 100 µs), there are often more than 1 track reconstructed.

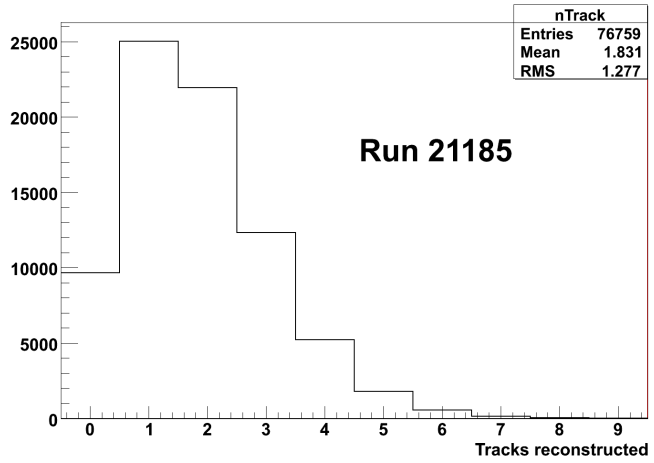


Figure 8.16: Number of reconstructed tracks per event.

Figure 8.17 shows the residuals between the track fit and measured hit position in the first telescope sensor. They are well centred at zero and are very narrow, with an RMS around 2 µm reflecting the extremely good resolution and alignment. The shape slightly deviates from a gaussian. This is due to the fact, that tracks are formed with different number of hits, from 4 to 6. It was checked that requiring strictly 6 hits per track results

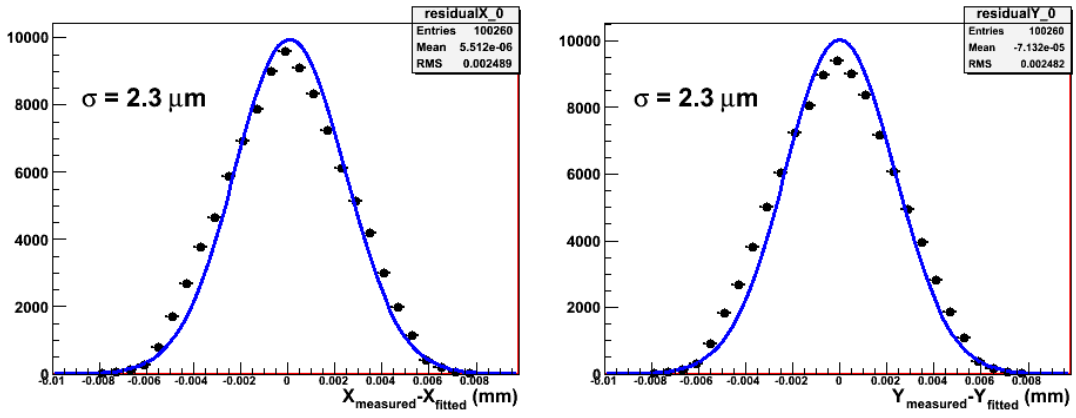


Figure 8.17: Trackfit residuals. The line is a gaussian fit.

in a gaussian distribution. The RMS is the same for X and Y direction as expected for the square pixel shape.

The expected resolution at the DUT is around 4.5 µm (see Section 8.6.5) which allows detailed in-pixel studies (the size of the ATLAS pixel is $50 \times 400\,\mu m^2$).

8.9.3 Results

Spatial resolution

Fig. 8.18 shows the residuals between the fitted telescope track and the measured hit positions in the Planar and 3D sensors. Since the DUT pixel size is much larger than the telescope tracking precision, these distributions reflect mainly the spatial resolutions of the DUTs. The residuals are centred at zero, showing that the alignment of the DUTs was successful.

For the Planar device, the RMS of the distribution in the short pixel direction is 16 µm which is close to $50\,\mu m/\sqrt{12} = 14.4\,\mu m$. In the long direction, the distribution has a box-like shape from -200 µm to $+200$ µm as expected from the long side size of 400 µm. Also small shoulders are visible from -300 µm to $+300$ µm. They correspond to a small fraction of elongated pixels of 600 µm length.

For the 3D sensor, similar resolution in the short direction is observed as for the Planar device. Similar box shape is observed for the long direction. However, in case of 3D sensor, a specific structure appears: five dips are observed in the central part of the distribution. Those correspond to the geometrical structure of the 3D sensor, in particular to the vertically implanted electrodes, see the next subsection. A similar behaviour was also observed in earlier test beam studies of 3D sensors [176].

Charge collection efficiency

For every hit, the front-end chip FE-I3 provides a measurement of the Time-over-Threshold (ToT) – the duration of time for which the pulse is above the threshold. It is measured in units of the bunch-crossing time (25 ns). ToT is related to the charge deposited by a particle in a roughly linear manner [143, 177]. Fig. 8.19 presents the ToT comparison for 3D and PPS sensors for hits (clusters) matched to tracks. The distributions have a Landau function shape, as expected for the Minimun Ionising Particles. As can be seen from the plot, the peak position is somewhat lower for the 3D sensor which can be due to its smaller thickness. The 3D distribution shows also a small bump at low values of ToT, which suggests that there are areas with lower charge collection efficiency.

In order to study the origin of the lower charge collection efficiency, an in-pixel distribution is shown in Fig. 8.20. A clear structure is observed for 3D sensor. Indeed regions of lower charge collection are present within the pixel. They correspond to the 3D implants and this is the reason for peculiar structure of the residuals in Figure 8.18. The Planar sensor shows more uniform response, however there is a small region of lower efficiency which corresponds to the bias dot.

Charge sharing probability

If charge sharing between pixels occurs, more than one pixel will have deposited charge. This effect is more pronounced at the edges of the pixel, due to smaller drifting distance to a neighboring pixel. Charge sharing is important since it may improve the spatial resolution.

The charge sharing probability is defined as the ratio of the number of clusters with more than one pixel to the total number of clusters. The in-pixel distributions of charge sharing probability are shown in Figure 8.21 for both sensors. The width of the green band indicates that charge sharing distance is around 10 µm for PPS sensor. This agrees to observations of earlier publications. Charge sharing is much less pronounced for 3D sensor, with a charge distance of $\sim$5 µm, in agreement to a previous study [176].

Hit efficiency

One of the most important quantities is the hit efficiency which defines the probability to detect a hit when a track passes the detector. Fig 8.22 shows the hit efficiency in the local frame of reference. A high efficiency of 99.5 % is observed for the PPS sensor, while the 3D sensor has a lower efficiency of 95 %. This is due to the presence of lower charge collection efficiency regions within the pixel.

8.9.4 Conclusions

The results of the first IBL test beam were presented in this section. Two different technologies were compared simultaneously: Planar and 3D. The performance of both types of sensors is generally comparable. However the efficiency of the 3D sensor is lower at normal incidence. It can be compensated by tilting the sensors [176] as will be the case for the IBL modules. The results observed agree to previous measurements.

This demonstrates understanding of the test beam tools that will be used for the final IBL test beam based on which (together with lab measurements) the technology will be chosen. The extended EUTelescope software is fully capable for ATLAS Pixel device analysis from raw data to final high level observables such as hit efficiency or charge sharing probability.

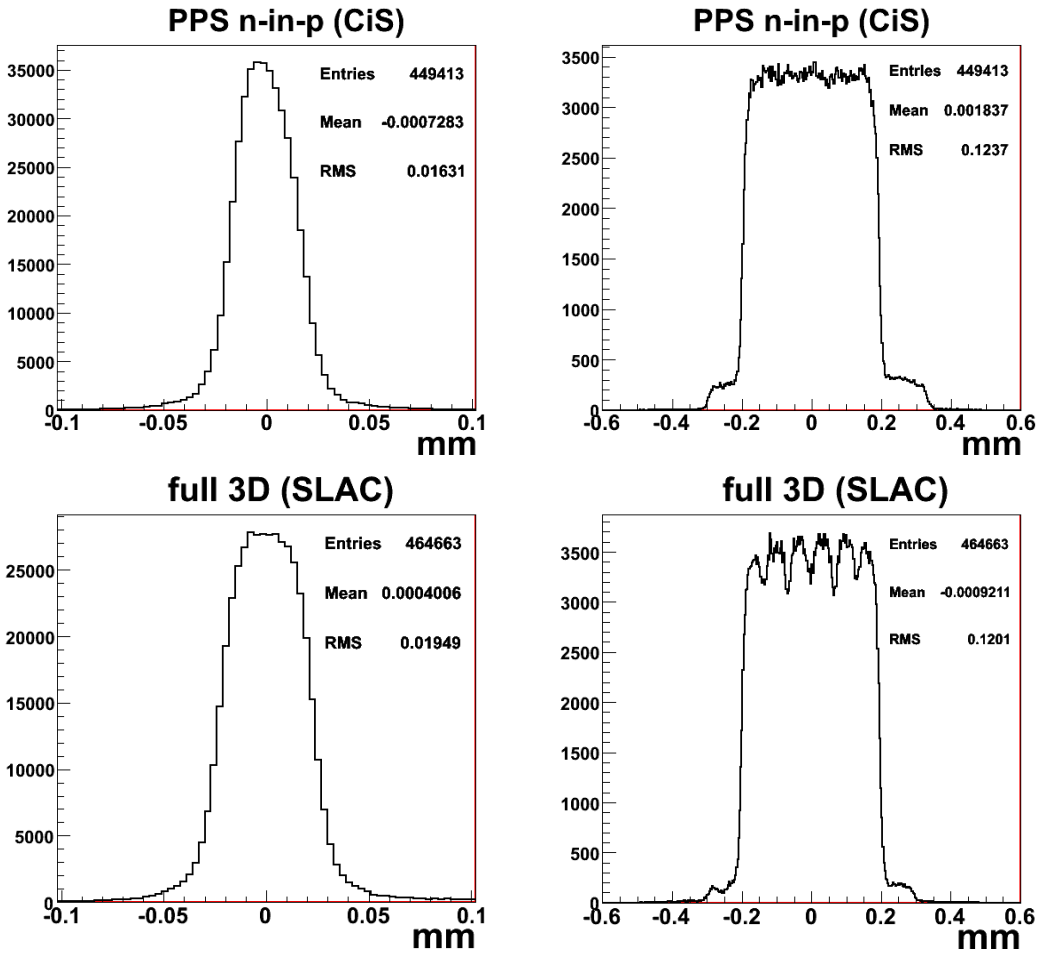


Figure 8.18: Residuals for PPS device in the short (left) and long (right) pixel direction. Residuals for 3D device in the short (left) and long (right) pixel direction

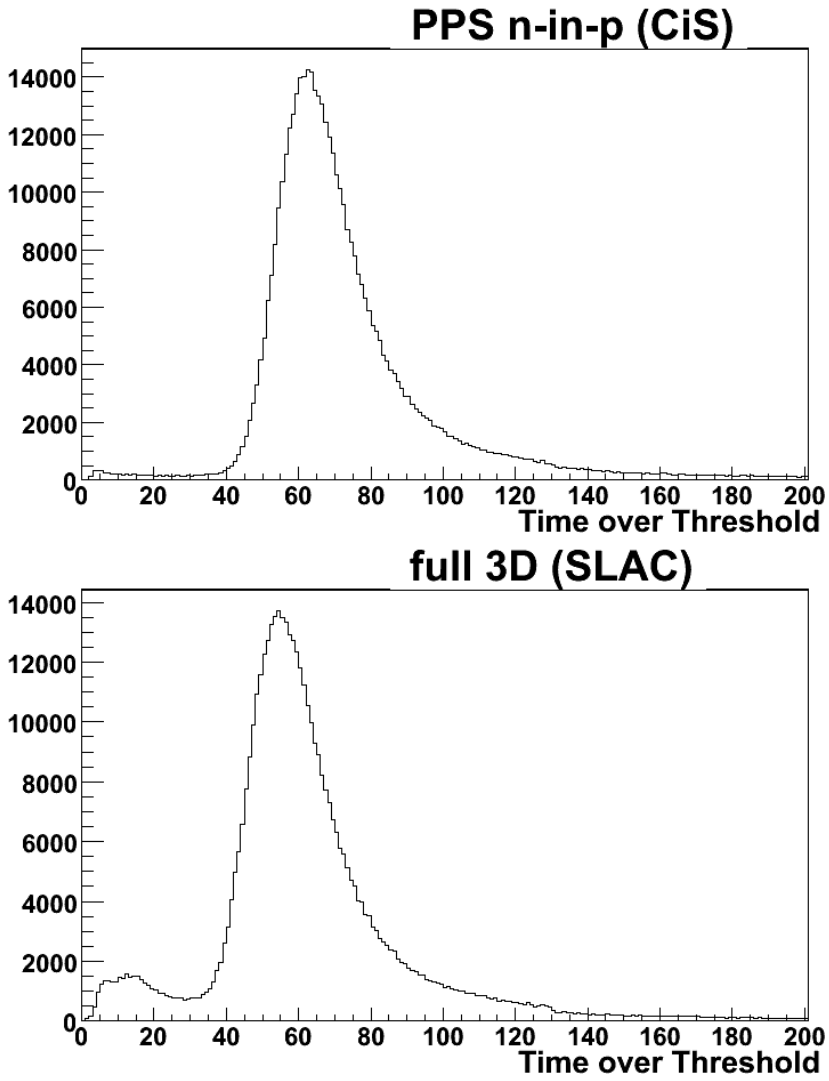


Figure 8.19: Time-over-Threshold distribution for PPS and 3D sensors.

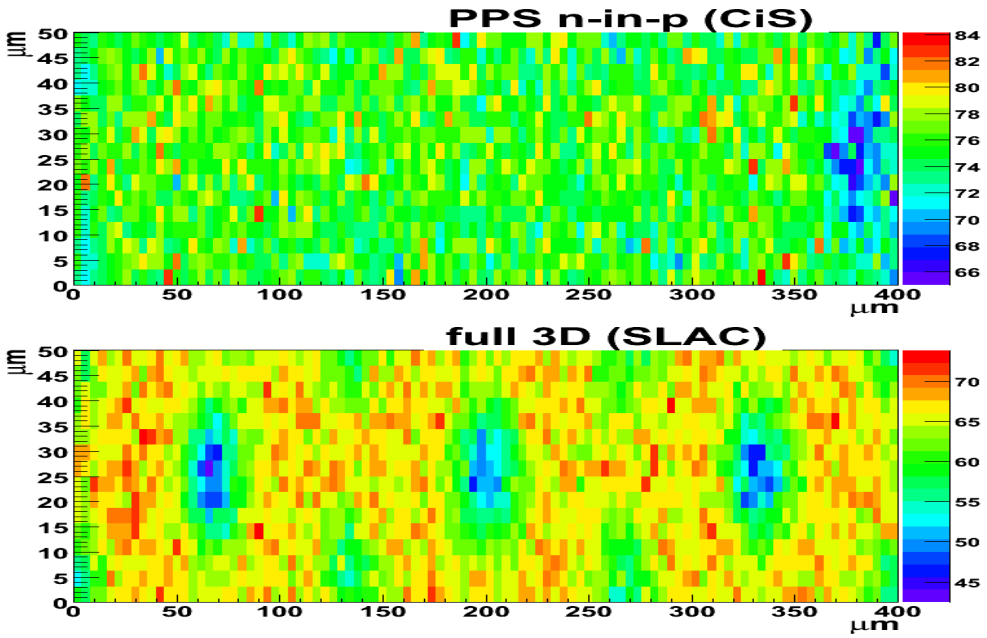


Figure 8.20: The ToT value of the pixel as a function of the incidence point inside the pixel (charge collection efficiency).

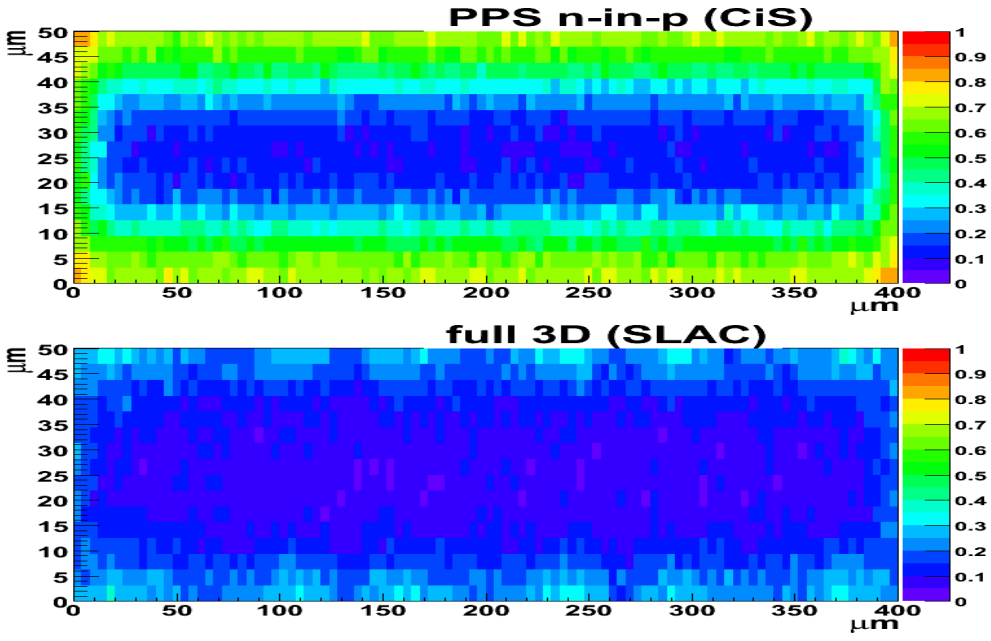


Figure 8.21: Charge sharing probability inside the pixel.

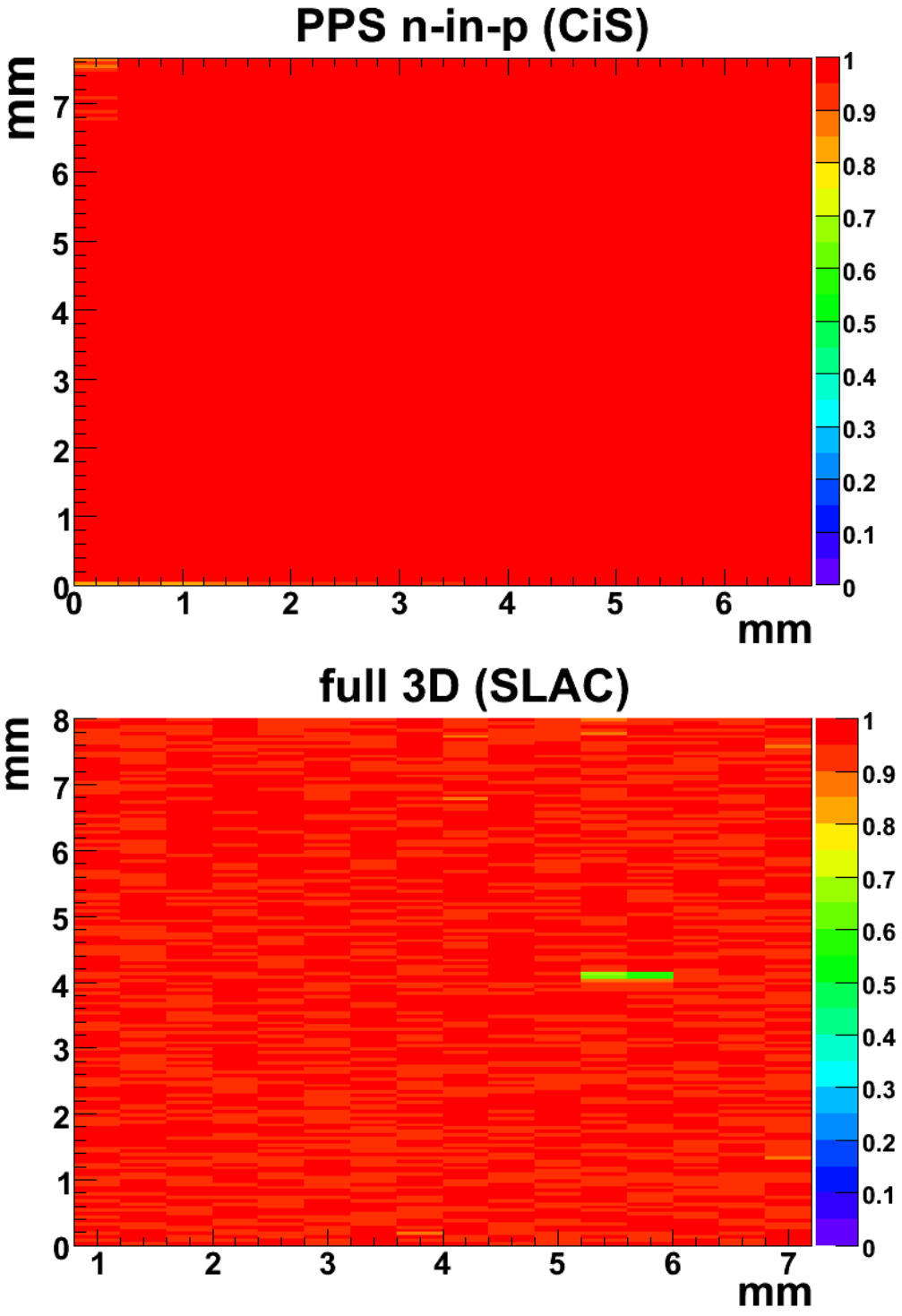


Figure 8.22: Hit efficiency for the sensor.

8.10 Integration of USBPix into EUDET Telescope

In the previous section, the first IBL test beam was described. The ATLAS Pixel sensors that were studied were operated with the FE-I3 front-end, the chip that is used in the current Pixel detector. However, in the final IBL test beam one has to characterise sensors that are read out by the next generation ATLAS Pixel front-end chip FE-I4 [152], since it will be used for IBL. The conventional ATLAS Pixel data acquisition system (DAQ), that has been used for all previous ATLAS Pixel test beams, including the one described above, is collectively called TurboDAQ. It is not suited for operation of FE-I4 and is an outdated system not maintained anymore.

In order to overcome these limitations, a new DAQ system which allows extension to the FE-I4, was developed at Bonn University – the so-called USBPix [178, 179]. It requires much less hardware than the previous system and has a better user interface. It was chosen to be the ATLAS Pixel DAQ system for the final IBL test beam with FE-I4 read-out. In order to allow data taking with the EUDET Telescope, certain steps have to be accomplished – the system needs to be *integrated* to the telescope.

This section describes the integration of the USBPix system to the EUDET Telescope on the DAQ software level. After an overview of USBPix, technical details on the integration are given. The results of the first test run of the USBPix system with the EUDET Telescope in test beam conditions at CERN SPS (IBL test beam Phase II) are presented afterwards.

8.10.1 Overview of the USBPix System

The USBPix system is shown in Fig. 8.23. It consists of two PCB boards connected together with board connectors, the Multi-IO Board and the Adapter Card. An ATLAS Pixel sensor with a front-end chip to be read out has to be placed on the so-called Single Chip Card. It is then connected via a flat ribbon cable to the USBPix board (more precisely, to the Adapter Card). The USBPix board is connected via the USB interface to a PC where the DAQ software called STcontrol is running.

Multi-IO board is the main part of the USBPix. It is based on the Xilinx FPGA[6] which controls the read out and performs the on-board analysis. There is an SRAM memory block on the board. It temporary stores the data collected from the Pixel Sensor before it is sent out. The USB controller allows communication to the PC. The Adapter Card provides the interface for the Single Chip Card.

STcontrol is a DAQ software that configures the system and controls the operation. It is based on the PixLib package which is a collection of C++ classes originally developed for ATLAS Pixel Read-out Drivers (RODs) [143]. Hardware specific parts have been adapted to access the ATLAS Pixel sensor via the USBPix board. The Graphical User

[6]Field Programmable Gate Array

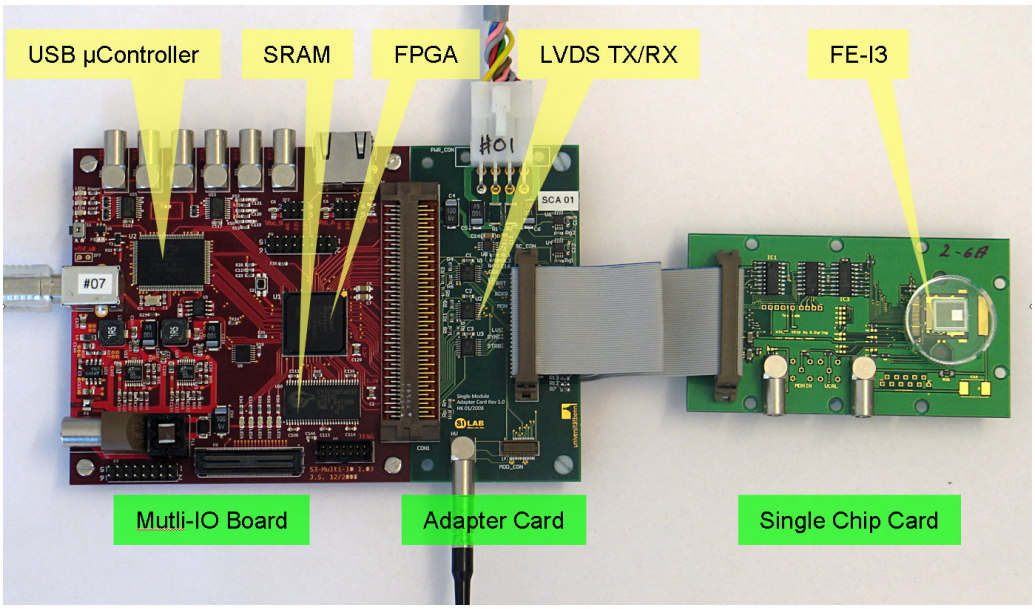


Figure 8.23: The USBPix board (left) and a Single Chip Card connected to it via a flat ribbon cable (right).

Interface is implemented in ROOT[7] and Qt[8].

8.10.2 Integration to the EUDET Telescope

The full integration of the DUT into the EUDET Telescope means implementing two communication mechanisms: DUT ↔ TLU and DUT ↔ EUDAQ PC (see Section 8.6.3 for the TLU description and Section 8.6.4 for more information on EUDAQ).

The communication with TLU is needed for the trigger handshake. Different types of handshake that are supported by the TLU were described in detail in Sect. 8.6.3. The recommended mode is the trigger data handshake: when TLU encounters a coincidence of signals from certain number of scintillators (depending on the settings), it raises the TRIGGER line to the DUT and waits while the DUT raises the BUSY signal. The TLU then de-asserts the TRIGGER line and clocks out the trigger number on it. After that DUT de-asserts the BUSY signal and the system is ready for new triggers. Such a handshake for USBPix was implemented at Bonn University [180].

The communication to the EUDAQ PC is required to get commands from the Run Control and to send data from the DUT to the central DAQ PC. The data from the telescope and from the DUTs are merged to form events and are stored to disk for offline

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

[8] Qt is a development framework which allows efficient creation of applications and user interfaces, http://qt.nokia.com

analysis. Since this type of communication was missing for the USBPix board, it was implemented in this thesis.

The EUDAQ Software foresees a unified procedure for the DUT integration. This procedure implies creating a C++ class (*Producer class*) that inherits *EUDAQ::Producer* [161]. An application that creates an instance of this class must be launched. It has to pass an IP address of the main DAQ PC where the Run Control is running. Thanks to the *EUDAQ::Producer* functionality, it then automatically establishes connection to the DAQ PC. The communication with the Run Control is done via TCP/IP protocol. This application is also called a Producer (see section 8.6.4).

On the other hand, this class has to have access to the pixel hardware, so that data can be read out. Taking this into account, the Producer class called *ProducerThread* was implemented as part of the USBPix software, STcontrol. In particular, it was realised as a friend class of the *STControlEngine* which is the highest level class of the STcontrol. The friendship is needed in order to be able to access members of *STControlEngine* that contain pixel data. An instance of *ProducerThread* is created whenever STcontrol is launched in the test beam mode. Figure 8.24 illustrates this scheme.

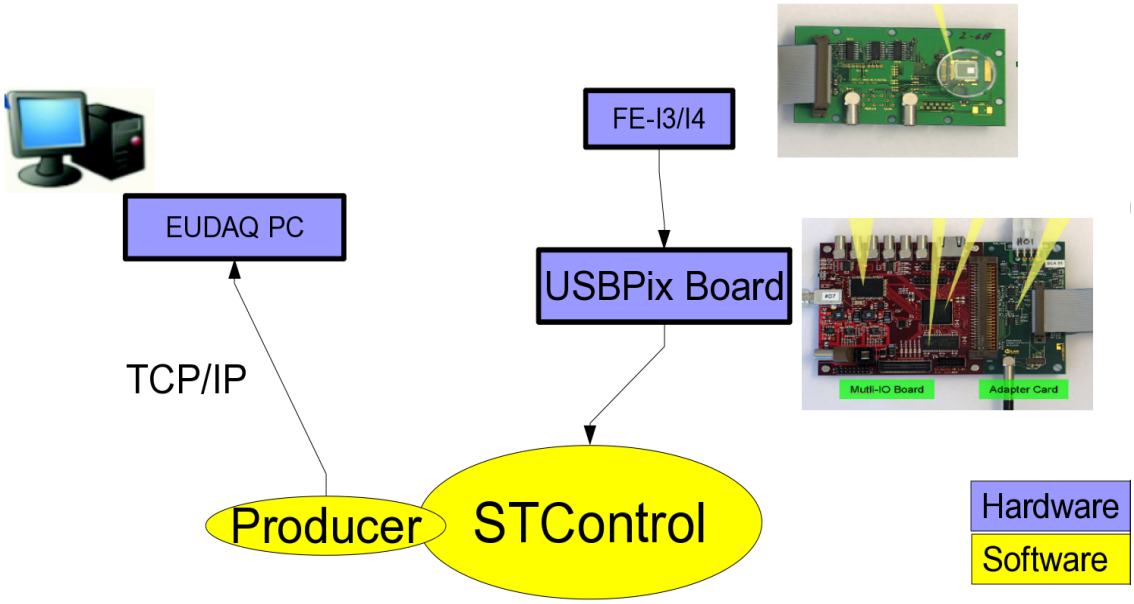


Figure 8.24: Integration of the USBPix to the EUDET Telescope on the DAQ software level. The Producer class is part of the STControl software.

Since data have to be continously read out during the run, the main routine was realised as a thread. This was achieved by inheriting *ProducerThread* from the *QThread* class of Qt, in addition to *eudaq::Producer* inheritance.

The following class methods, inherited from the mother classes, were re-implemented in order to provide the required functionality:

- *OnStartRun()* is automatically called when the *START RUN* button is pressed in the Run Control GUI. This method has to take appropriate actions to start the data taking by the USBPix. In particular it launches the *Source Scan* in the test beam mode. This is an external triggering data taking mode: on receipt of a trigger from the TLU, the USBPix reads out the hit information from the front-end chip and stores it to the SRAM.

  Also the *start()* method which is inherited from *QThread* is called within OnStartRun(). It automatically launches the main Producer thread called *run()*.

- *run()* is the main routine of the Producer. It is implemented as a thread, i.e. is executed permanently as an infinite loop. Returning from this method stops the thread. The amount of collected data in the SRAM is constantly checked in the loop. When it is almost full, the read out is initiated: events are built and sent to the EUDAQ PC for storage, afterwards the SRAM is cleared.

- *OnStopRun()* is called when a *STOP RUN* button is pressed in the Run Control GUI. It stops the source scan and sends all the data that remain in the SRAM to the DAQ PC.

- *OnConfigure()* is called when a *CONFIGURE* button is pressed in the Run Control. It has to take appropriate actions to configure the hardware. However, in the first version of the Producer no operations were performed whithin this method - the USBPix had to be configured from the STControl. In the later developments, the configuration of the USBPix in OnConfigure() was implemented [181].

In addition to the Producer, a *Data Converter Plugin* was created. This class is needed for data conversion from the EUDAQ native format to the LCIO format, i.e. for the first step of the offline analysis (Sect. 8.7.1). It also allows data quality monitoring during data taking. The class must inherit *EUDAQ::DataConverterPlugin*. The conversion is realised in the *GetLCIOSubEvent* method. The LCIO format that was adopted is identical to one used in the previous TurboDAQ system, i.e. storing the sensor hitmap for every event as a TrackerData object. This assures that the offline analysis software can be used in the same way to analyse data collected with TurboDAQ or USBPix systems.

8.10.3 System Test at the October 2010 IBL Test Beam

The USBPix system was for the first time used at the test beam in October 2010 at CERN SPS during the Phase II of the IBL test beam which was described in Sect. 8.9. A single chip card with FE-I3 front-end was connected to the USBpix and placed in the

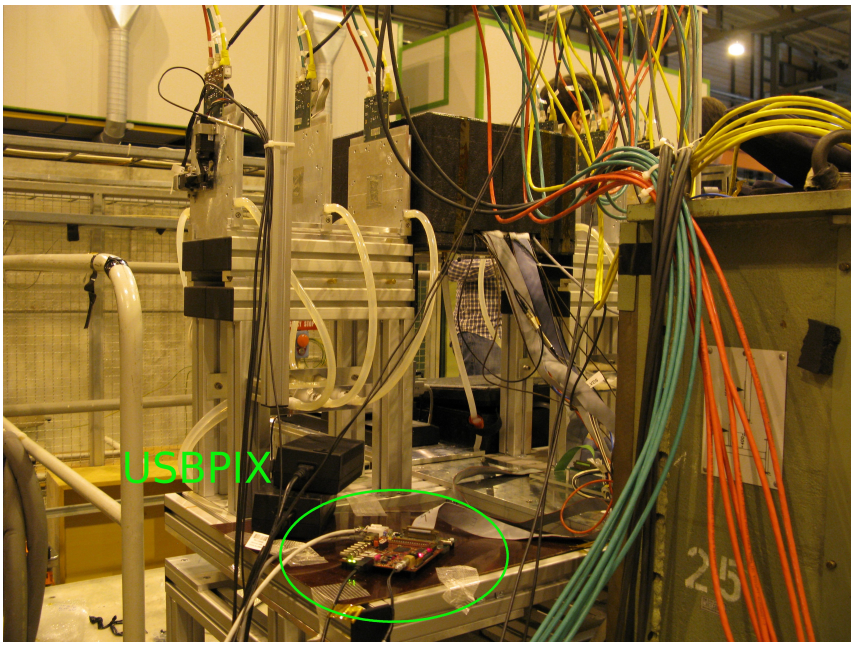


Figure 8.25: USBPix system test at CERN SPS during the IBL test beam, October 2010. The USBPix board is highlighted with the green circle. The telescope planes and the box with the ATLAS pixel sensors are also visible.

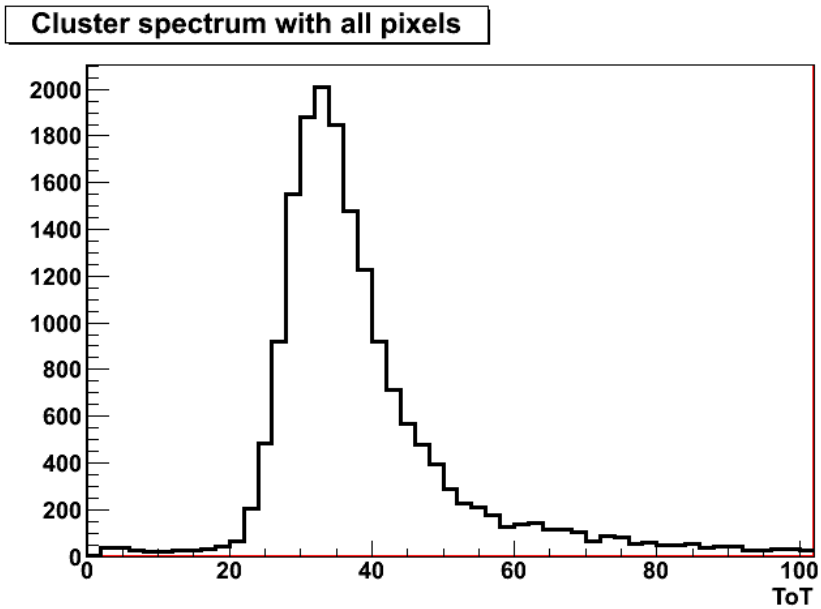


Figure 8.26: Distribution of the Time-over-Threshold value (proportional to collected charge) per cluster for events collected with the USBPix.

beam (Fig. 8.25). In parallel, the conventional TurboDAQ system was running. After a debugging session a stable operation was reached. Several million events were collected without any problems. The online monitoring, in particular the correlation plots, the ToT and LV1 spectra showed that the data collected are reasonable and correspond to the beam particles. A subset of collected data was also analysed offline with the EUTelescope. Results as anticipated were obtained. As an example, Fig. 8.26 shows the Time-over-Threshold (ToT) distribution (related to collected charge, see Sect 8.9.3). A Landau-like shape is observed as expected for minimum-ionising particles.

8.10.4 Summary

In this section the integration of the USBPix DAQ system to the EUDET Telescope was described, in particular implementation of the EUDAQ Producer. This allowed usage of the USBPix as DAQ for the ATLAS Pixel sensors at the test beam with the EUDET Telescope. The system was tested during the first IBL test beam campaign. It showed stable performance and easy operation. The online data quality monitoring as well as offline analysis show reasonable results.

8.11 Summary and Outlook

This chapter described various activities within the IBL upgrade program, related to test beam studies of the pixel sensors. The offline data anlysis framework EUTelescope was extended, which allowed an efficient analysis of the ATLAS Pixel test beam data. The data from the first IBL test beam was analysed. The results were consistent with expectations, which assured a good understanding of the tools and of the sensors and gave confidence for further test beam campaigns. The USBPix DAQ system has been integrated to the EUDET Telescope which was a missing piece for usage of this system in test beams. It was tested during the second part of the mentioned test beam and showed very satisfying performance. This development allowed final IBL test beams with the FE-I4 readout chip as well.

Based on the work presented here, many developments have taken place since it was accomplished (January 2011). Two more IBL test beams were performed, at DESY and at CERN, where irradiated sensors with the FE-I4 chip were tested. The final decision for the IBL was taken: 75 % of IBL modules will be Planar sensors and 25 % 3D sensors, in the forward region. This approach profits on one hand from high production yield of the PPS technology and on the other hand gives a possibility to the 3D community to use the sensor for a large scale project which is very important for a developing technology.

Conclusions

In this work, a measurement of charm and beauty quark production in deep inelastic scattering at HERA was presented. A data set collected by the ZEUS detector in the period from 2003 to 2007, corresponding to an integrated luminosity of $354\,\mathrm{pb}^{-1}$, was analysed. A secondary vertex technique was employed to tag heavy flavours. The method makes use of the fact that charm and beauty hadrons have long lifetimes and large decay lengths, hence they can be separated from the light flavour background. The advantage of the method is that no specific final state is required, every charm or beauty hadron decay with at least two charged tracks can be potentially reconstructed.

A detailed evaluation of the systematic uncertainties was performed in this thesis; a substantial reduction of uncertainties from certain sources was achieved, compared to the previous study [14]. As an example, the systematics due to a large uncertainty on the ZEUS tracking efficiency was reduced thanks to dedicated studies of hadronic interactions performed in this work.

Differential cross-sections of jet production in charm and beauty events were measured and compared to NLO QCD calculations in the fixed flavour number scheme. Within uncertainties, an agreement between data and theory was found. The central values of predictions for charm are however typically below the data.

The contributions $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ of charm and beauty quark production to the inclusive proton structure function F_2 were determined from the double differential cross sections. NLO QCD predictions were used to extrapolate the cross sections from the visible to the full phase space. The obtained results can be compared to theory predictions in various schemes. The extracted $F_2^{b\bar{b}}$ represent the most precise measurement among the published HERA results. The $F_2^{c\bar{c}}$ is competitive compared to the previous measurements. In future, these data will be used for the combination of H1 and ZEUS results which will be the ultimate results on $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ from HERA. A QCD analysis of them can potentially improve understanding of heavy flavour production as well as allow more precise determinations of the quark masses.

The work was also devoted to beam test studies of the ATLAS pixel sensors, within the Insertable B-layer (IBL) upgrade programme. The test beam data analysis software EUTelescope was extended, which allowed efficient analysis and characterisation of ATLAS pixel sensors. At CERN SPS, a simultaneous beam test characterisation of two sensor

types, the Planar and 3D was performed with the EUDET pixel telescope. Usage a single experimental setup ensured that all conditions are the same and hence the results are comparable. The results that were obtained such as the hit efficiency or charge collection were consistent to the previous studies, thus ensuring understanding of the new tools, which was an important step in the programme. The ATLAS pixel data acquisition system for IBL beam tests called USBPix was integrated to the EUDET Telescope, allowing its usage for final IBL testbeams.

Appendix A

Mirrored Significance

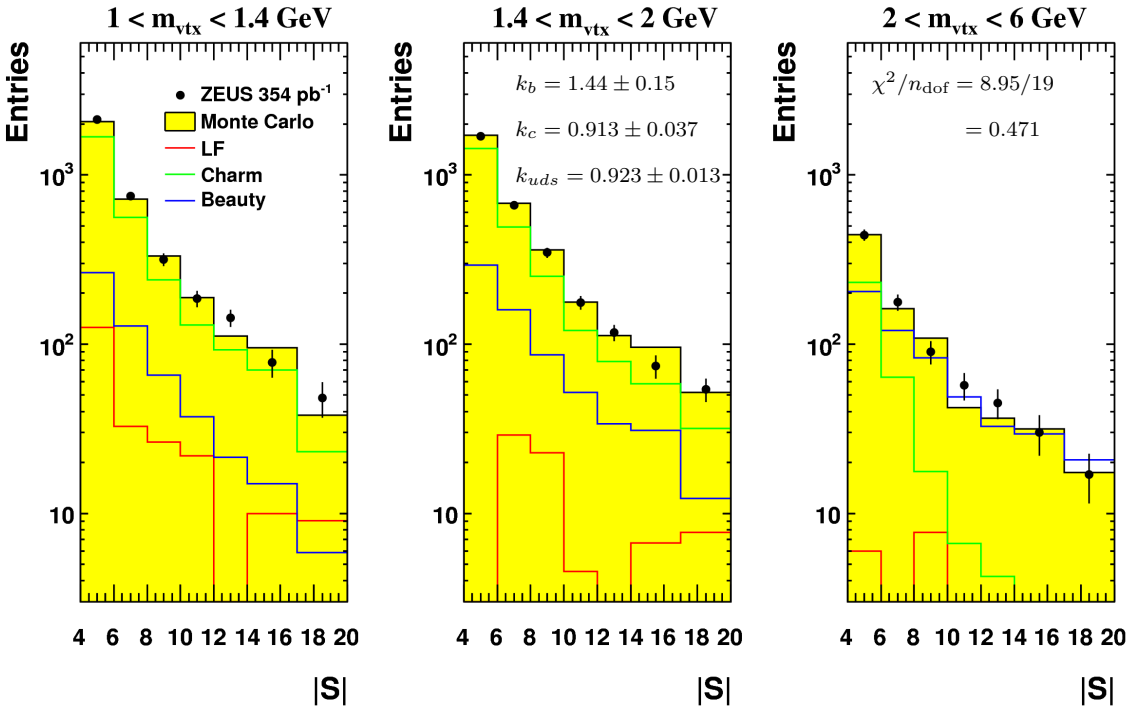


Figure A.1: Negative subtracted significance in three mass bins for the region of low Q^2 ($5 < Q^2 < 10\,\mathrm{GeV}^2$) for the charm measurement ($E_T^{\mathrm{jet}} > 4.2\,\mathrm{GeV}$). Points show the data; solid lines represent different MC contributions while the shaded histogram shows their sum. Values of the beauty, charm and light flavour scaling factors obtained from the fit of these distribution according to the procedure described in Section 5.5 are indicated. Different MC contributions are scaled by the corresponding scaling factors. The χ^2/n_{dof} of the fit is also shown.

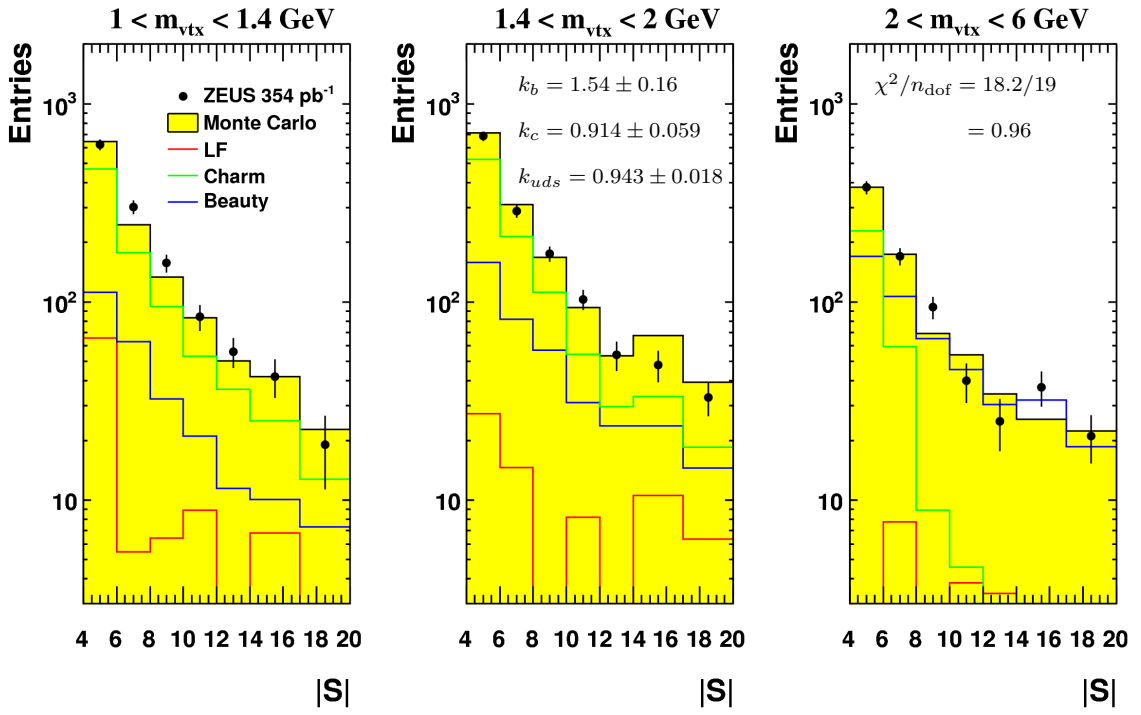


Figure A.2: Negative subtracted significance in three mass bins for the region of mid Q^2 ($120 < Q^2 < 200\,\mathrm{GeV}^2$). Other details are as in Fig. A.1.

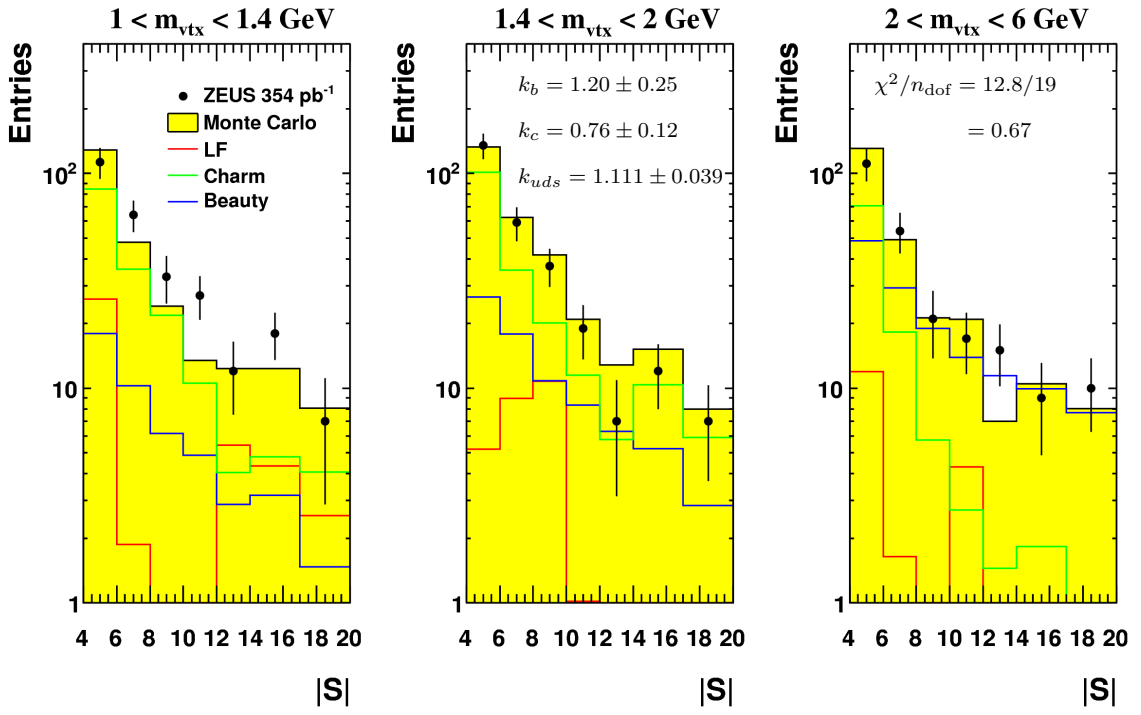


Figure A.3: Negative subtracted significance in three mass bins for the region of high Q^2 ($400 < Q^2 < 1000\,\mathrm{GeV}^2$). Other details are as in Fig. A.1.

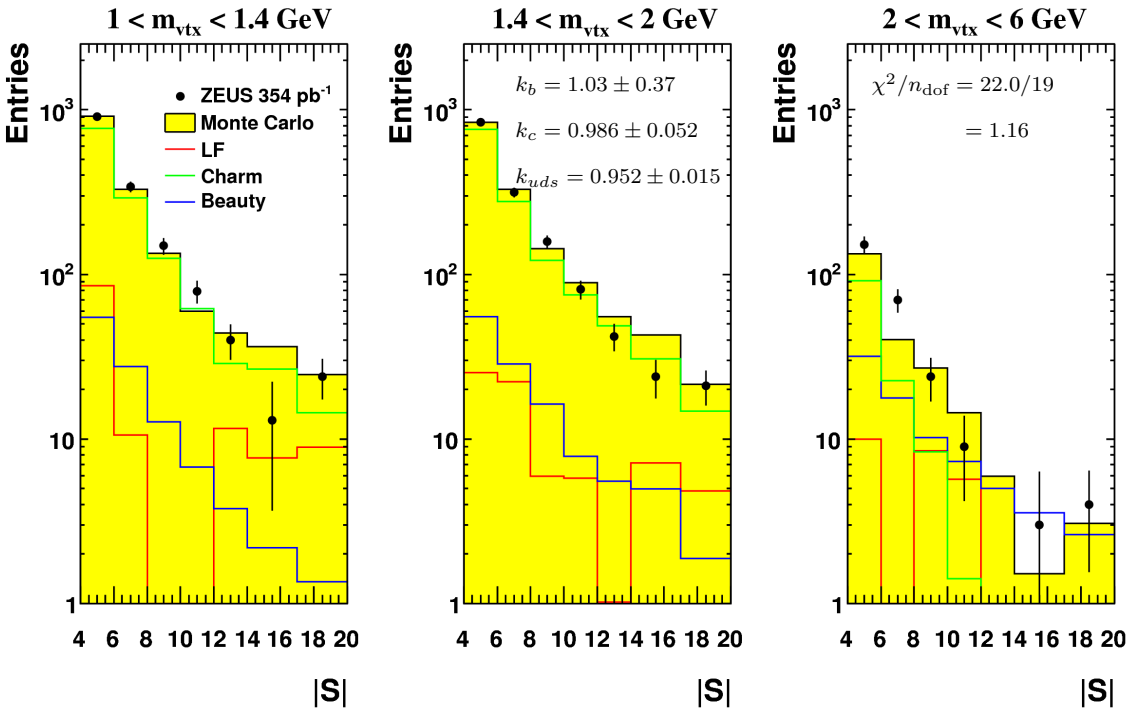


Figure A.4: Negative subtracted significance in three mass bins in the backward region ($-1.6 < \eta^{\text{jet}} < -0.8$) for the beauty measurement ($E_T^{\text{jet}} > 5.0\,\text{GeV}$). Other details are as in Fig. A.1.

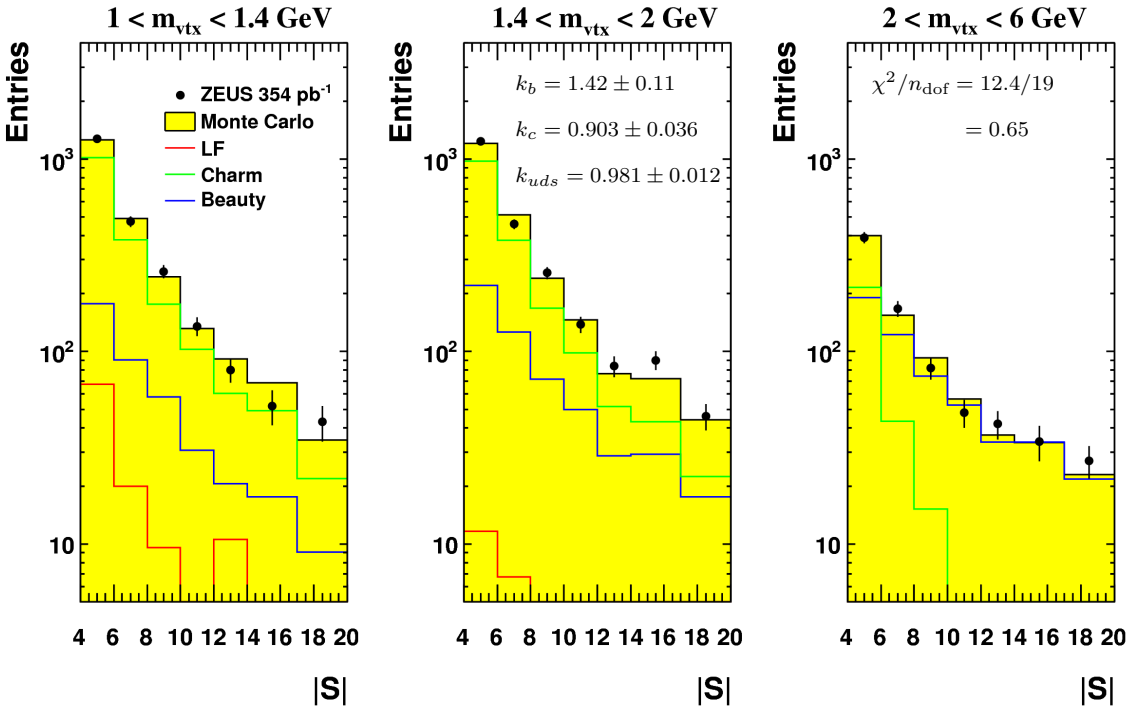


Figure A.5: Negative subtracted significance in three mass bins in the central region ($-0.2 < \eta^{\text{jet}} < 0.1$). Other details are as in Fig. A.4.

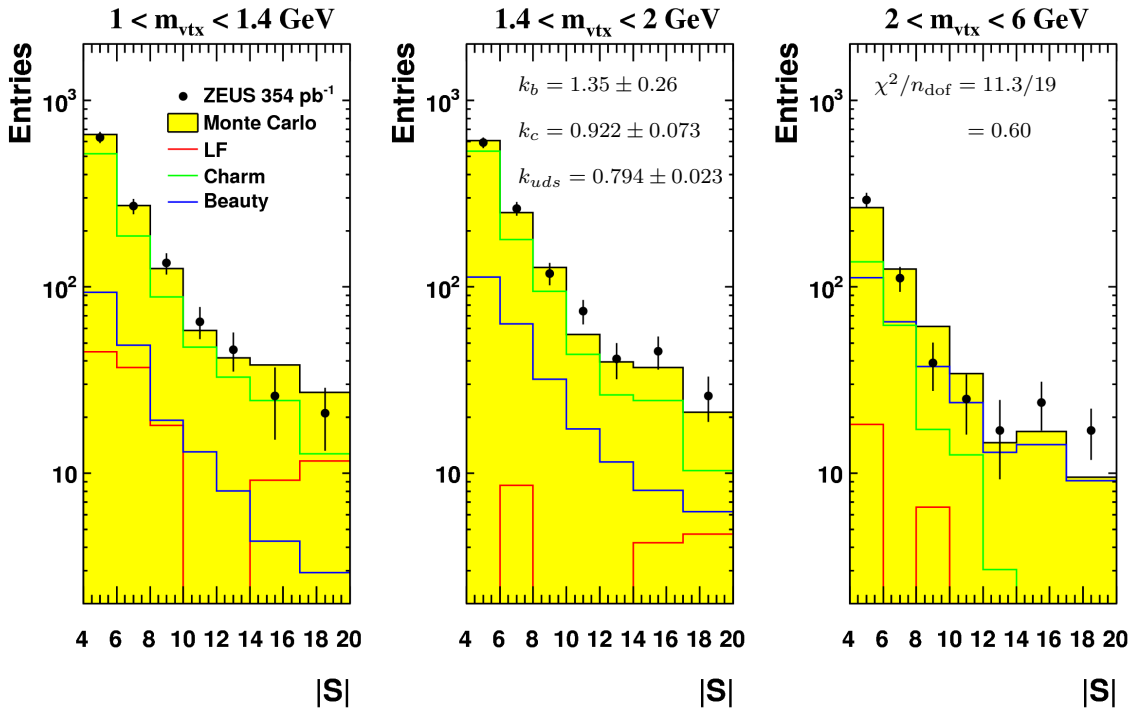


Figure A.6: Negative subtracted significance in three mass bins in the forward region ($1.3 < \eta^{\text{jet}} < 1.6$). Other details are as in Fig. A.4.

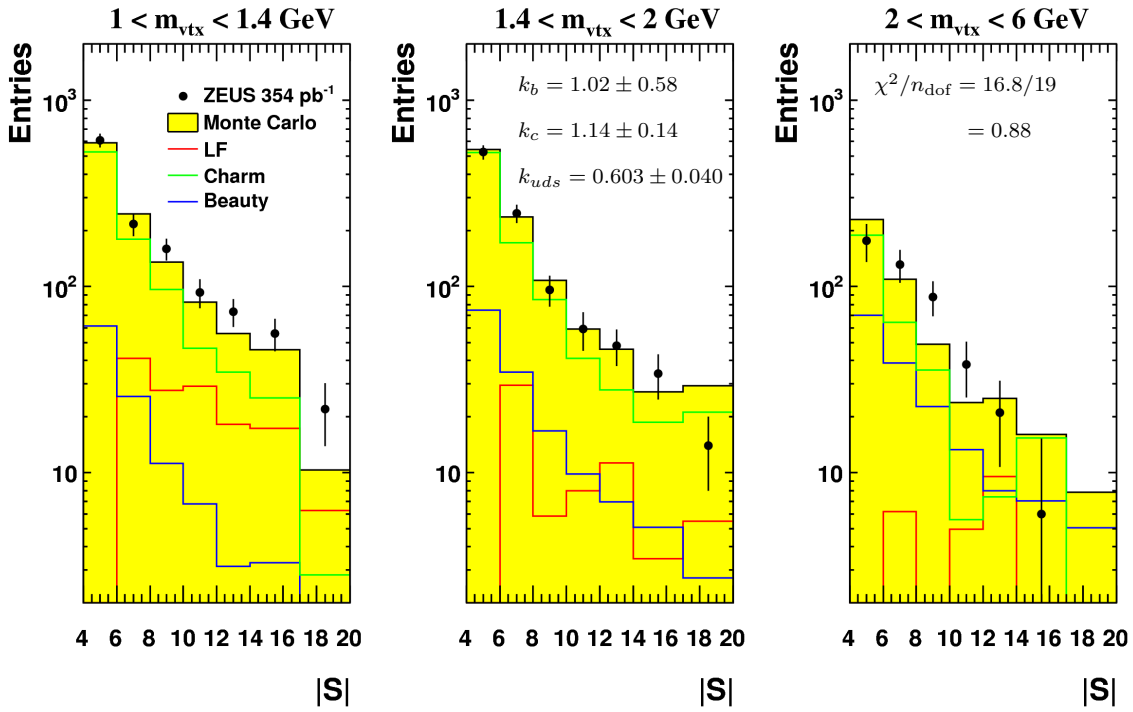


Figure A.7: Negative subtracted significance in three mass bins in the forward region ($1.6 < \eta^{\text{jet}} < 2.2$). Other details are as in Fig. A.4.

Appendix B

Cross Section Tables

E_T^{jet} (GeV)	$\mathrm{d}\sigma_b/\mathrm{d}E_T^{\text{jet}}$ (pb/GeV)
5 : 8	$124 \pm 25^{+21}_{-23}$
8 : 11	$66.2 \pm 5.7^{+5.1}_{-6.9}$
11 : 14	$27.9 \pm 1.9^{+1.3}_{-1.8}$
14 : 17	$10.70 \pm 0.84^{+0.55}_{-0.61}$
17 : 20	$4.40 \pm 0.46^{+0.33}_{-0.32}$
20 : 25	$1.75 \pm 0.26^{+0.23}_{-0.23}$
25 : 35	$0.344 \pm 0.089^{+0.073}_{-0.071}$

E_T^{jet} (GeV)	$\mathrm{d}\sigma_c/\mathrm{d}E_T^{\text{jet}}$ (pb/GeV)
4.2 : 8	$3630 \pm 120^{+200}_{-180}$
8 : 11	$720 \pm 22^{+43}_{-40}$
11 : 14	$214.6 \pm 9.7^{+21}_{-20}$
14 : 17	$85.8 \pm 6.0^{+10}_{-9.0}$
17 : 20	$35.4 \pm 3.9^{+5.5}_{-5.4}$
20 : 25	$14.1 \pm 2.7^{+3.5}_{-3.2}$
25 : 35	$2.38 \pm 0.97^{+0.85}_{-0.86}$

Table B.1: Differential cross sections for jet production in beauty events (top) and charm events (bottom) as a function of E_T^{jet}. The cross sections are given for $5 < Q^2 < 1000\,\text{GeV}^2, 0.02 < y < 0.7$, $E_T^{\text{jet}} > 5(4.2)\,\text{GeV}$ and $-1.6 < \eta^{\text{jet}} < 2.2$. The measurements are given together with their statistical and systematic uncertainties.

η^{jet}	$\mathrm{d}\sigma_b/\mathrm{d}\eta^{\text{jet}}$ (pb)
-1.6 : -0.8	$81 \pm 29^{+23}_{-30}$
-0.8 : -0.5	$222 \pm 29^{+15}_{-20}$
-0.5 : -0.2	$206 \pm 24^{+15}_{-17}$
-0.2 : 0.1	$290 \pm 22^{+18}_{-19}$
0.1 : 0.4	$262 \pm 22^{+15}_{-16}$
0.4 : 0.7	$306 \pm 23^{+22}_{-22}$
0.7 : 1.0	$259 \pm 26^{+19}_{-21}$
1.0 : 1.3	$278 \pm 30^{+26}_{-27}$
1.3 : 1.6	$218 \pm 42^{+20}_{-22}$
1.6 : 2.2	$125 \pm 71^{+46}_{-44}$

η^{jet}	$\mathrm{d}\sigma_c/\mathrm{d}\eta^{\text{jet}}$ (pb)
-1.6 : -1.1	$2090 \pm 240^{+230}_{-200}$
-1.1 : -0.8	$3360 \pm 210^{+230}_{-210}$
-0.8 : -0.5	$4340 \pm 200^{+220}_{-200}$
-0.5 : -0.2	$5190 \pm 190^{+260}_{-240}$
-0.2 : 0.1	$5200 \pm 200^{+360}_{-350}$
0.1 : 0.4	$5760 \pm 210^{+390}_{-370}$
0.4 : 0.7	$5580 \pm 210^{+330}_{-320}$
0.7 : 1.0	$5480 \pm 240^{+290}_{-280}$
1.0 : 1.3	$4880 \pm 260^{+300}_{-280}$
1.3 : 1.6	$4680 \pm 360^{+310}_{-280}$
1.6 : 2.2	$4840 \pm 600^{+660}_{-660}$

Table B.2: Differential cross sections for jet production in beauty events (top) and charm events (bottom) as a function of η^{jet}. The cross sections are given for $5 < Q^2 < 1000\,\text{GeV}^2, 0.02 < y < 0.7$, $E_T^{\text{jet}} > 5(4.2)\,\text{GeV}$ and $-1.6 < \eta^{\text{jet}} < 2.2$. The measurements are given together with their statistical and systematic uncertainties.

Q^2 (GeV^2)	$\mathrm{d}\sigma_b/\mathrm{d}Q^2$ ($\mathrm{pb}/\mathrm{GeV}^2$)
5 : 10	$42.8 \pm 4.2^{+3.0}_{-2.6}$
10 : 20	$17.5 \pm 1.8^{+1.4}_{-1.3}$
20 : 40	$7.28 \pm 0.80^{+0.62}_{-0.65}$
40 : 70	$3.67 \pm 0.38^{+0.19}_{-0.23}$
70 : 120	$1.18 \pm 0.16^{+0.13}_{-0.14}$
120 : 200	$0.607 \pm 0.065^{+0.039}_{-0.045}$
200 : 400	$0.141 \pm 0.017^{+0.010}_{-0.010}$
400 : 1000	$0.0180 \pm 0.0032^{+0.0018}_{-0.0017}$

Q^2 (GeV^2)	$\mathrm{d}\sigma_c/\mathrm{d}Q^2$ ($\mathrm{pb}/\mathrm{GeV}^2$)
5 : 10	$852 \pm 35^{+48}_{-41}$
10 : 20	$469 \pm 15^{+27}_{-22}$
20 : 40	$201.9 \pm 6.2^{+10}_{-9.2}$
40 : 70	$65.7 \pm 2.7^{+3.6}_{-3.2}$
70 : 120	$21.7 \pm 1.1^{+1.3}_{-1.1}$
120 : 200	$6.80 \pm 0.44^{+0.44}_{-0.37}$
200 : 400	$1.75 \pm 0.13^{+0.10}_{-0.08}$
400 : 1000	$0.197 \pm 0.031^{+0.028}_{-0.025}$

Table B.3: Differential cross sections for jet production in beauty events (top) and charm events (bottom) as a function of Q^2. The cross sections are given for $5 < Q^2 < 1000\,\mathrm{GeV}^2, 0.02 < y < 0.7$, $E_T^{\mathrm{jet}} > 5(4.2)\,\mathrm{GeV}$ and $-1.6 < \eta^{\mathrm{jet}} < 2.2$. The measurements are given together with their statistical and systematic uncertainties.

x	$\mathrm{d}\sigma_b/\mathrm{d}x$ (pb)
0.00008 : 0.0002	$650000 \pm 110000^{+74000}_{-71000}$
0.0002 : 0.0006	$599000 \pm 47000^{+39000}_{-37000}$
0.0006 : 0.0016	$214000 \pm 15000^{+11000}_{-12000}$
0.0016 : 0.005	$47500 \pm 3500^{+2700}_{-3000}$
0.005 : 0.01	$11500 \pm 1300^{+920}_{-930}$
0.01 : 0.1	$336 \pm 73^{+41}_{-40}$

x	$\mathrm{d}\sigma_c/\mathrm{d}x$ (pb)
0.00008 : 0.0002	$10850000 \pm 870000^{+760000}_{-650000}$
0.0002 : 0.0006	$10840000 \pm 390000^{+570000}_{-440000}$
0.0006 : 0.0016	$4970000 \pm 140000^{+260000}_{-230000}$
0.0016 : 0.005	$1217000 \pm 31000^{+69000}_{-63000}$
0.005 : 0.01	$257000 \pm 12000^{+18000}_{-17000}$
0.01 : 0.1	$10650 \pm 790^{+910}_{-820}$

Table B.4: Differential cross sections for jet production in beauty events (top) and charm events (bottom) as a function of x. The cross sections are given for $5 < Q^2 < 1000\,\mathrm{GeV}^2, 0.02 < y < 0.7$, $E_T^{\mathrm{jet}} > 5(4.2)\,\mathrm{GeV}$ and $-1.6 < \eta^{\mathrm{jet}} < 2.2$. The measurements are given together with their statistical and systematic uncertainties.

Q^2 (GeV^2)	x	$\mathrm{d}^2\sigma_b/\mathrm{d}x\mathrm{d}Q^2$ (pb/GeV^2)
5 : 20	0.00008 : 0.0002	$650000 \pm 110000^{+74000}_{-70000}$
5 : 20	0.0002 : 0.0003	$850000 \pm 120000^{+70000}_{-68000}$
5 : 20	0.0003 : 0.0005	$480000 \pm 56000^{+38000}_{-35000}$
5 : 20	0.0005 : 0.003	$47000 \pm 5900^{+4000}_{-3900}$
20 : 60	0.0003 : 0.0005	$81000 \pm 24000^{+9500}_{-10000}$
20 : 60	0.0005 : 0.0012	$131000 \pm 13000^{+7100}_{-7700}$
20 : 60	0.0012 : 0.002	$69800 \pm 8800^{+5700}_{-6100}$
20 : 60	0.002 : 0.0035	$25000 \pm 4700^{+2800}_{-2900}$
20 : 60	0.0035 : 0.01	$3700 \pm 1800^{+990}_{-1000}$
60 : 120	0.0008 : 0.0018	$32100 \pm 4300^{+2700}_{-2800}$
60 : 120	0.0018 : 0.003	$21600 \pm 2800^{+1700}_{-1800}$
60 : 120	0.003 : 0.006	$6500 \pm 1300^{+560}_{-620}$
120 : 400	0.0016 : 0.005	$6540 \pm 910^{+390}_{-460}$
120 : 400	0.005 : 0.016	$3880 \pm 340^{+210}_{-230}$
120 : 400	0.016 : 0.06	$225 \pm 110^{+62}_{-64}$
400 : 1000	0.005 : 0.02	$386 \pm 85^{+52}_{-49}$
400 : 1000	0.02 : 0.1	$58 \pm 19^{+12}_{-13}$

Table B.5: Double-differential cross sections for jet production in beauty events as a function of x for different ranges of Q^2. The cross sections are given for $5 < Q^2 < 1000\,\mathrm{GeV}^2, 0.02 < y < 0.7$, $E_T^{\mathrm{jet}} > 5\,\mathrm{GeV}$ and $-1.6 < \eta^{\mathrm{jet}} < 2.2$. The measurements are given together with their statistical and systematic uncertainties.

Q^2 (GeV^2)	x	$\mathrm{d}^2\sigma_c/\mathrm{d}x\mathrm{d}Q^2$ (pb/ GeV^2)
5 : 20	0.00008 : 0.0002	$10850000 \pm 870000^{+740000}_{-590000}$
5 : 20	0.0002 : 0.0003	$14630000 \pm 1000000^{+960000}_{-780000}$
5 : 20	0.0003 : 0.0005	$8490000 \pm 480000^{+550000}_{-480000}$
5 : 20	0.0005 : 0.003	$1639000 \pm 56000^{+100000}_{-89000}$
20 : 60	0.0003 : 0.0005	$1710000 \pm 220000^{+140000}_{-110000}$
20 : 60	0.0005 : 0.0012	$2550000 \pm 110000^{+130000}_{-120000}$
20 : 60	0.0012 : 0.002	$1364000 \pm 66000^{+71000}_{-67000}$
20 : 60	0.002 : 0.0035	$626000 \pm 33000^{+46000}_{-44000}$
20 : 60	0.0035 : 0.01	$184000 \pm 13000^{+14000}_{-13000}$
60 : 120	0.0008 : 0.0018	$238000 \pm 34000^{+30000}_{-30000}$
60 : 120	0.0018 : 0.003	$289000 \pm 22000^{+22000}_{-21000}$
60 : 120	0.003 : 0.006	$131100 \pm 8300^{+9500}_{-9100}$
60 : 120	0.006 : 0.04	$15500 \pm 1500^{+1300}_{-1100}$
120 : 400	0.0016 : 0.005	$103500 \pm 7600^{+7100}_{-5500}$
120 : 400	0.005 : 0.016	$34800 \pm 2200^{+2000}_{-1800}$
120 : 400	0.016 : 0.06	$4270 \pm 860^{+580}_{-560}$
400 : 1000	0.005 : 0.02	$5150 \pm 750^{+710}_{-650}$
400 : 1000	0.02 : 0.1	$500 \pm 220^{+140}_{-140}$

Table B.6: Double-differential cross sections for jet production in charm events as a function of x for different ranges of Q^2. The cross sections are given for $5 < Q^2 < 1000\,\mathrm{GeV}^2, 0.02 < y < 0.7$, $E_T^{\mathrm{jet}} > 4.2\,\mathrm{GeV}$ and $-1.6 < \eta^{\mathrm{jet}} < 2.2$. The measurements are given together with their statistical and systematic uncertainties.

Q^2 (GeV^2)	x	F_2^{bb}
6.5	0.00015	$0.00413 \pm 0.00072\,^{+0.00047\ +0.00033}_{-0.00044\ -0.00028}$
6.5	0.00028	$0.00373 \pm 0.00053\,^{+0.00031\ +0.00030}_{-0.00030\ -0.00026}$
12	0.00043	$0.00612 \pm 0.00072\,^{+0.00049\ +0.00043}_{-0.00044\ -0.00033}$
12	0.00065	$0.00431 \pm 0.00054\,^{+0.00037\ +0.00020}_{-0.00036\ -0.00012}$
25	0.00043	$0.0119 \pm 0.0035\,^{+0.0014\ +0.0009}_{-0.0015\ -0.0007}$
25	0.0008	$0.0103 \pm 0.0010\,^{+0.0006\ +0.0006}_{-0.0006\ -0.0005}$
30	0.0016	$0.0094 \pm 0.0012\,^{+0.0008\ +0.0004}_{-0.0008\ -0.0004}$
30	0.0025	$0.0065 \pm 0.0012\,^{+0.0007\ +0.0002}_{-0.0007\ -0.0002}$
30	0.0045	$0.0042 \pm 0.0021\,^{+0.0011\ +0.0001}_{-0.0012\ -0.0001}$
80	0.0016	$0.0357 \pm 0.0048\,^{+0.0030\ +0.0012}_{-0.0031\ -0.0012}$
80	0.0025	$0.0190 \pm 0.0024\,^{+0.0015\ +0.0005}_{-0.0016\ -0.0005}$
80	0.0045	$0.0098 \pm 0.0020\,^{+0.0008\ +0.0002}_{-0.0009\ -0.0003}$
160	0.0035	$0.0227 \pm 0.0032\,^{+0.0014\ +0.0005}_{-0.0016\ -0.0003}$
160	0.008	$0.0176 \pm 0.0015\,^{+0.0009\ +0.0004}_{-0.0011\ -0.0004}$
160	0.02	$0.0075 \pm 0.0037\,^{+0.0021\ +0.0003}_{-0.0022\ -0.0002}$
600	0.013	$0.0156 \pm 0.0035\,^{+0.0021\ +0.00014}_{-0.0020\ -0.0002}$
600	0.035	$0.0077 \pm 0.0026\,^{+0.0016\ +0.0002}_{-0.0018\ -0.0001}$

Table B.7: 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.

Q^2 (GeV^2)	x	$F_2^{c\bar{c}}$
6.5	0.00015	$0.214 \pm 0.017\,^{+0.015\ +0.049}_{-0.012\ -0.044}$
6.5	0.00028	$0.209 \pm 0.014\,^{+0.014\ +0.043}_{-0.011\ -0.043}$
12	0.00043	$0.243 \pm 0.014\,^{+0.016\ +0.041}_{-0.014\ -0.047}$
12	0.00065	$0.2268 \pm 0.0077\,^{+0.014\ +0.029}_{-0.012\ -0.029}$
25	0.00043	$0.553 \pm 0.072\,^{+0.045\ +0.096}_{-0.034\ -0.082}$
25	0.0008	$0.423 \pm 0.018\,^{+0.022\ +0.030}_{-0.020\ -0.034}$
30	0.0016	$0.304 \pm 0.015\,^{+0.016\ +0.014}_{-0.015\ -0.011}$
30	0.0025	$0.229 \pm 0.012\,^{+0.017\ +0.006}_{-0.016\ -0.008}$
30	0.0045	$0.190 \pm 0.013\,^{+0.014\ +0.010}_{-0.013\ -0.000}$
80	0.0016	$0.377 \pm 0.054\,^{+0.048\ +0.017}_{-0.048\ -0.009}$
80	0.0025	$0.334 \pm 0.025\,^{+0.025\ +0.0015}_{-0.024\ -0.015}$
80	0.0045	$0.240 \pm 0.015\,^{+0.018\ +0.004}_{-0.017\ -0.004}$
80	0.008	$0.208 \pm 0.020\,^{+0.017\ +0.000}_{-0.015\ -0.002}$
160	0.0035	$0.438 \pm 0.032\,^{+0.030\ +0.010}_{-0.023\ -0.007}$
160	0.008	$0.196 \pm 0.013\,^{+0.011\ +0.003}_{-0.010\ -0.003}$
160	0.02	$0.152 \pm 0.031\,^{+0.021\ +0.002}_{-0.020\ -0.000}$
600	0.013	$0.245 \pm 0.036\,^{+0.034\ +0.006}_{-0.031\ -0.002}$
600	0.035	$0.077 \pm 0.035\,^{+0.022\ +0.003}_{-0.022\ -0.001}$

Table B.8: The structure function $F_2^{c\bar{c}}$ 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.

Bibliography

[1] H. Geiger and E. Marsden, *On a Diffuse Reflection of the α-Particles*, *Proc. R. Soc. Lond.* **A 82** (1909) 495–500.

[2] E. Rutherford, *The Scattering of α and β Particles by Matter and the Structure of the Atom*, *Phil. Mag.* **21** (1911) 669–688.

[3] R. Hofstadter and R. W. McAllister, *Electron Scattering from the Proton*, *Phys. Rev.* **98** (1955) 217–218.

[4] R. W. McAllister and R. Hofstadter, *Elastic Scattering of 188-Mev Electrons from the Proton and the Alpha Particle*, *Phys. Rev.* **102** (1956) 851–856.

[5] W. Bartel et al., *Electroproduction of pions near the $\Delta(1236)$ isobar and the form factor $G_M^*(q^2)$ of the $(\gamma N\Delta)$-vertex*, *Phys. Lett.* **B 28** (1968) 148–151.

[6] J. I. Friedman and H. W. Kendall, *Deep inelastic electron scattering*, *Ann. Rev. Nucl. Part. Sci.* **22** (1972) 203–254.

[7] D. Perkins, *Introduction to High Energy Physics, 4th edition*. Cambridge University Press (2000).

[8] B. Wiik, *HERA, a Proposal for a Large Electron-Proton Colliding Beam Facility at DESY*, DESY-HERA-81-10 (1981).

[9] A. M. Cooper-Sarkar and R. Devenish, *The rise and fall of F_2 at low x*, *Acta Phys. Polon.* **B 34** (2003) 2911–2928.

[10] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*. Cambridge University Press (1996).

[11] **ZEUS** Collaboration, H. Abramowicz et al., *Measurement of heavy-quark jet photoproduction at HERA*, *Eur. Phys. J.* **C 71** (2011) 1659.

[12] A. G. Yagues Molina, *Study of beauty photoproduction with the ZEUS experiment at the electron-proton collider HERA*, Ph.D. thesis, Humboldt University (2007).

[13] V. E. Schoenberg, *Measurement of beauty and charm photoproduction using inclusive secondary vertexing with the ZEUS detector at HERA*, Ph.D. thesis, Bonn University, BONN-IR-2010-05 (2010).

[14] P. Roloff, *Measurement of charm and beauty production in deep inelastic scattering at HERA*, Ph.D. thesis, Hamburg University, DESY-THESIS-2011-049 (2011).

[15] R. Devenish and A. Cooper-Sarkar, *Deep Inelastic Scattering.* Oxford University Press (2004).

[16] A. M. Cooper-Sarkar, R. Devenish, and A. De Roeck, *Structure Functions of the Nucleon and their Interpretation*, *Int. J. Mod. Phys.* **A 13** (1998) 3385–3586.

[17] J. D. Bjorken and E. A. Paschos, *Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon*, *Phys. Rev.* **185** (1969) 1975–1982.

[18] F. Halzen and A. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics.* John Wiley & Sons (1984).

[19] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1988) 1–91.

[20] V. Gribov and L. Lipatov, *Deep Inelastic ep Scattering in Perturbation Theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[21] L. Lipatov, *The Parton Model and Perturbation Theory*, *Sov. J. Nucl. Phys.* **20** (1975) 94–102.

[22] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B 126** (1977) 298.

[23] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*, *Sov. Phys. JETP* **46** (1977) 641–653.

[24] E. Laenen, S. Riemersma, J. Smith, and W. van Neerven, *Complete $\mathcal{O}(\alpha_s)$ Corrections to Heavy-Flavour Structure Functions in Electroproduction*, *Nucl. Phys.* **B 392** (1993) 162–228.

[25] E. Laenen, S. Riemersma, J. Smith, and W. van Neerven, *$\mathcal{O}(\alpha_s)$ Corrections to Heavy-Flavour Inclusive Distributions in Electroproduction*, *Nucl. Phys.* **B 392** (1993) 229–250.

[26] S. Riemersma, J. Smith, and W. van Neerven, *Rates for Inclusive Deep-Inelastic Electroproduction of Charm Quarks at HERA*, *Phys. Lett.* **B 347** (1995) 143–151.

[27] B. Harris and J. Smith, *Heavy-Quark Correlations in Deep-Inelastic Electroproduction*, *Nucl. Phys.* **B 452** (1995) 109–160.

[28] B. Kniehl, M. Krämer, G. Kramer, and M. Spira, *Cross sections for charm production in ep collisions: massive versus massless scheme*, *Phys. Lett.* **B 356** (1995) 539–545.

[29] B. A. Kniehl, G. Kramer, and M. Spira, *Large-p_T photoproduction of $D^{*\pm}$ mesons in ep collisions*, *Z. Phys.* **C 76** (1997) 689–700.

[30] F. Olness and W.-K. Tung, *When is a heavy quark not a parton? Charged Higgs production and heavy quark mass effects in the QCD-based parton model*, *Nucl. Phys.* **B 308** (1988) 813–832.

[31] M. A. G. Aivazis, F. I. Olness, and W.-K. Tung, *Leptoproduction of heavy quarks. I. General formalism and kinematics of charged current and neutral current production processes*, *Phys. Rev.* **D 50** (1994) 3085–3101.

[32] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, *Leptoproduction of heavy quarks. II. A unified QCD formulation of charged and neutral current processes from fixed-target to collider energies*, *Phys. Rev.* **D 50** (1994) 3102–3118.

[33] M. Krämer, F. I. Olness, and D. E. Soper, *Treatment of heavy quarks in deeply inelastic scattering*, *Phys. Rev.* **D 62** (2000) 096007.

[34] R. S. Thorne and R. G. Roberts, *Ordered analysis of heavy flavor production in deep-inelastic scattering*, *Phys. Rev.* **D 57** (1998) 6871–6898.

[35] R. Thorne and R. Roberts, *A practical procedure for evolving heavy flavour structure functions*, *Phys. Lett.* **B 421** (1998) 303 – 311.

[36] R. Thorne and R. Roberts, *A variable flavour number scheme for charged current heavy flavour structure functions*, *Eur. Phys. J.* **C 19** (2001) 339–349.

[37] G. Kramer and H. Spiesberger, *Inclusive photoproduction of D^* mesons with massive charm quarks*, *Eur. Phys. J.* **C 38** (2004) 309–318.

[38] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, *Inclusive $D^{*\pm}$ production in $p\overline{p}$ collisions with massive charm quarks*, *Phys. Rev.* **D 71** (2005) 014018.

[39] **H1** Collaboration, C. Adloff et al., *Inclusive D^0 and $D^{*\pm}$ Production in Deep Inelastic ep Scattering at HERA*, *Z. Phys.* **C 72** (1996) 593–605.

[40] **ZEUS** Collaboration, J. Breitweg et al., *D^* production in deep inelastic scattering at HERA*, *Phys. Lett.* **B 407** (1997) 402–418.

[41] **H1** Collaboration, C. Adloff et al., *Measurement of D^* meson cross-sections at HERA and determination of the gluon density in the proton using NLO QCD*, *Nucl. Phys.* **B 545** (1999) 21–44.

[42] **ZEUS** Collaboration, J. Breitweg et al., *Measurement of $D^{*\pm}$ production and the charm contribution to F_2 in deep inelastic scattering at HERA*, *Eur. Phys. J.* **C 12** (2000) 35–52.

[43] **H1** Collaboration, C. Adloff et al., *Measurement of $D^{*\pm}$ meson production and F_2^c in deep-inelastic scattering at HERA*, *Phys. Lett.* **B 528** (2002) 199–214.

[44] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of $D^{*\pm}$ production in deep inelastic $e^{\pm}p$ scattering at DESY HERA*, *Phys. Rev.* **D 69** (2004) 012004.

[45] **H1** Collaboration, A. Aktas et al., *Inclusive production of D^+, D^0, D_s^+ and D^{*+} mesons in deep inelastic scattering at HERA*, *Eur. Phys. J.* **C 38** (2005) 447–459.

[46] **H1** Collaboration, 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** (2005) 349–359.

[47] **H1** Collaboration, 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** (2006) 23–33.

[48] **H1** Collaboration, A. Aktas et al., *Inclusive $D^{*\pm}$ meson and associated dijet production in deep-inelastic scattering at HERA*, *Eur. Phys. J.* **C 51** (2007) 271–287.

[49] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of D mesons production in deep inelastic scattering at HERA*, *JHEP* **0707** (2007) 074.

[50] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of $D^{\pm}$ and D^0 production in deep inelastic scattering using a lifetime tag at HERA*, *Eur. Phys. J.* **C 63** (2009) 171–188.

[51] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of charm and beauty production in deep inelastic ep scattering from decays into muons at HERA*, *Eur. Phys. J.* **C 65** (2010) 65–79.

[52] **H1** Collaboration, F. Aaron et al., *Measurement of the charm and beauty structure functions using the H1 vertex detector at HERA*, *Eur. Phys. J.* **C 65** (2010) 89–109.

[53] **H1** Collaboration, F. Aaron et al., *Measurement of the $D^{*\pm}$ meson production cross section and $F_2^{c\bar{c}}$, at high Q^2, in ep scattering at HERA*, *Phys. Lett.* **B 686** (2010) 91–100.

[54] **ZEUS** Collaboration, H. Abramowicz et al., *Measurement of D^+ and Λ_c^+ production in deep inelastic scattering at HERA*, *JHEP* **1011** (2010) 009.

[55] **H1** Collaboration, F. Aaron et al., *Measurement of charm and beauty jets in deep inelastic scattering at HERA*, *Eur. Phys. J.* **C 71** (2011) 1509.

[56] **H1** Collaboration, F. Aaron et al., *Measurement of $D^{*\pm}$ meson production and determination of $F_2^{c\bar{c}}$ at low Q^2 in deep-inelastic scattering at HERA*, *Eur. Phys. J.* **C 71** (2011) 1769.

[57] **ZEUS** Collaboration, H. Abramowicz et al., *Measurement of beauty production in DIS and $F_2^{b\bar{b}}$ extraction at ZEUS*, *Eur. Phys. J.* **C 69** (2010) 347–360.

[58] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of beauty photoproduction using decays into muons in dijet events at HERA*, *JHEP* **0904** (2009) 133.

[59] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of beauty production from dimuon events at HERA*, *JHEP* **0902** (2009) 032.

[60] **ZEUS** Collaboration, S. Chekanov et al., *Beauty photoproduction using decays into electrons at HERA*, *Phys. Rev.* **D 78** (2008) 072001.

[61] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of open beauty production at HERA in the $D^*\mu$ final state*, *Eur. Phys. J.* **C 50** (2007) 299–314.

[62] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of beauty production in deep inelastic scattering at HERA*, *Phys. Lett.* **B 599** (2004) 173–189.

[63] **ZEUS** Collaboration, S. Chekanov et al., *Bottom photoproduction measured using decays into muons in dijet events in ep collisions at $\sqrt{s} = 318\,\mathrm{GeV}$*, *Phys. Rev.* **D 70** (2004) 012008.

[64] **ZEUS** Collaboration, J. Breitweg et al., *Measurement of open beauty production in photoproduction at HERA*, *Eur. Phys. J.* **C 18** (2001) 625–637.

[65] **H1** Collaboration, A. Aktas et al., *Measurement of charm and beauty dijet cross sections in photoproduction at HERA using the H1 vertex detector*, *Eur. Phys. J.* **C 47** (2006) 597–610.

[66] **H1** Collaboration, A. Aktas et al., *Measurement of charm and beauty photoproduction at HERA using $D^*\mu$ correlations*, *Phys. Lett.* **B 621** (2005) 56–71.

[67] **H1** Collaboration, A. Aktas et al., *Measurement of beauty production at HERA using events with muons and jets*, *Eur. Phys. J.* **C 41** (2005) 453–467.

[68] **H1** Collaboration, C. Adloff et al., *Measurement of open beauty production at HERA*, *Phys. Lett.* **B 467** (1999) 156–164.

[69] **ZEUS** Collaboration, H. Abramowicz et al., *Measurement of beauty production in deep inelastic scattering at HERA using decays into electrons*, *Eur. Phys. J.* **C 71** (2011) 1573.

[70] **H1 and ZEUS** Collaboration, H. Abramowicz et al., *Combination and QCD analysis of charm production cross section measurements in deep-inelastic ep scattering at HERA*, *Eur. Phys. J.* **C 73** (2013) 1–26.

[71] H. Jung, *Hard diffractive scattering in high energy ep collisions and the Monte Carlo Generator* RAPGAP, *Comp. Phys. Comm.* **86** (1995) 147.

[72] H. Jung, *The* RAPGAP *Monte Carlo for Deep Inelastic Scattering. User Manual*, available at `http://rapgap.hepforge.org` (2011).

[73] M. Bowler, e^+e^- *Production of Heavy Quarks in the String Model*, *Z. Phys.* **C 11** (1981) 169.

[74] C. Peterson et al., *Scaling violations in inclusive* e^+e^- *annihilation spectra*, *Phys. Rev.* **D 27** (1983) 105.

[75] S. Agostinelli et al., GEANT*4 – a simulation toolkit*, *Nucl. Instrum. Meth.* **A 506** (2003) 250–303.

[76] J. Allison et al., GEANT*4 developments and applications*, *IEEE Trans. Nucl. Sci.* **53** (2006) 270–278.

[77] G. A. Voss and B. H. Wiik, *The Electron-Proton Collider Hera*, *Annu. Rev. Nucl. Sci.* **44** (1994) 413–452.

[78] U. Schneekloth, *The HERA luminosity upgrade*, DESY-HERA-98-05 (1998).

[79] **ZEUS** Collaboration, *The ZEUS detector: Status report 1993*, available at `www-zeus.desy.de/bluebook/bluebook.html`.

[80] C. Brooks et al., *Development of the ZEUS Central Tracking Detector*, *Nucl. Instrum. Meth.* **A 283** (1989) 477–483.

[81] B. Foster et al., *The design and construction of the ZEUS central tracking detector*, *Nucl. Instrum. Meth.* **A 338** (1994) 254.

[82] N. Harnew et al., *Vertex triggering using time difference measurements in the ZEUS central tracking detector*, *Nucl. Instrum. Meth.* **A 279** (1989) 290.

[83] B. Foster et al., *The Performance of the ZEUS Central Tracking Detector z-by-Timing Electronics in a Transputer Based Data Acquisition System*, *Nucl. Phys. Proc. Suppl.* **B 32** (1993) 181–188.

[84] A. Polini et al., *The design and performance of the ZEUS microvertex detector*, *Nucl. Instrum. Meth.* **A 581** (2007) 656.

[85] E. Koffeman, *A silicon micro vertex detector for the ZEUS experiment*, *Nucl. Instrum. Meth.* **A 453** (2000) 89–92.

[86] A. Andresen et al., *Construction and beam test of the ZEUS forward and rear calorimeter*, *Nucl. Instrum. Meth.* **A 309** (1991) 101.

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

[88] A. Caldwell et al., *Design and implementation of a high-precision readout system for the ZEUS calorimeter*, *Nucl. Instrum. Meth.* **A 321** (1992) 356.

[89] A. Bernstein et al., *Beam tests of the ZEUS barrel calorimeter*, *Nucl. Instrum. Meth.* **A 336** (1993) 23.

[90] H. Bethe and W. Heitler, *On the Stopping of Fast Particles and on the Creation of Positive Electrons*, *Proc. R. Soc. Lond.* **A 146** (1934) 83.

[91] J. Andruszkow et al., *Luminosity measurement in the ZEUS experiment*, *Acta Phys. Pol.* **B 32** (2001) 2025.

[92] M. Helbich et al., *The Spectrometer system for measuring ZEUS luminosity at HERA*, *Nucl. Instrum. Meth.* **A 565** (2006) 572–588.

[93] J. Chwastowski, J. Figiel, A. Kotarba, K. Olkiewicz, and L. Suszycki, *Aerogel Cherenkov detectors for the luminosity measurement at HERA*, *Nucl. Instrum. Meth.* **A 504** (2003) 222–227.

[94] W. Smith, K. Tokushuku, and L. Wiggers, *The ZEUS trigger system*, DESY-92-150B, C92-09-21 (1992).

[95] R. Carlin, W. Smith, K. Tokushuku, and L. Wiggers, *Experience with the ZEUS trigger system*, *Nucl. Phys. Proc. Suppl.* **44** (1995) 430–434.

[96] R. Carlin, W. Smith, K. Tokushuku, and L. Wiggers, *The trigger of ZEUS, a flexible system for a high bunch crossing rate collider*, *Nucl. Instrum. Meth.* **A 379** (1996) 542–544.

[97] H. Abramowicz, A. Caldwell, and R. Sinkus, *Neural network based electron identification in the zeus calorimeter*, *Nucl. Instrum. Meth.* **A 365** (1995) 508–517.

[98] G. Hartner, *VCTRAK Briefing: Program and Math*, ZEUS-Note 98-058 (1998).

[99] S. Spiridonov, *Mathematical Framework for Fast and Rigorous Track Fit for the ZEUS Detector*, ZEUS-Note 08-003 (2008).

[100] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, *Transactions of the ASME Journal of Basic Engineering* **82** (1960) 35–45.

[101] J. Beringer et al., *Review of Particle Physics*, *Phys. Rev.* **D 86** (2012) 010001.

[102] R. Frühwirth and A. Strandlie, *Track fitting with ambiguities and noise: A study of elastic tracking and nonlinear filters*, *Comp. Phys. Comm.* **120** (1999) 197–214.

[103] E. G. K. Rose and G. Fox, *Statistical mechanics and phase transitions in clustering*, *Phys. Rev. Lett.* **65** (1990) 945–948.

[104] K. Rose, *Deterministic annealing for clustering, compression, classification, regression, and related optimization problems*, in *Proceedings of the IEEE*, vol. 86, pp. 2210–2239, 1998.

[105] F. Didierjean, G. Duchêne, and A. Lopez-Martens, *The deterministic annealing filter: A new clustering method for γ-ray tracking algorithms*, *Nucl. Instrum. Meth.* **A 615** (2010) 188 – 200.

[106] S. Miglioranzi, *Beauty photoproduction at HERA II with the ZEUS experiment*, Ph.D. thesis, University College London (2006).

[107] G. M. Briskin, *Diffractive dissociation in ep deep inelastic scattering*, Ph.D. thesis, School of Phys. and Astronomy, Tel Aviv University, DESY-THESIS-1998-036 (1998).

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

[109] S. Catani et al., *Longitudinally-invariant $k_\perp$-clustering algorithms for hadron-hadron collisions*, *Nucl. Phys.* **B 406** (1993) 187.

[110] S. Ellis and D. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D 48** (1993) 3160.

[111] S. Catani, Y. Dokshitzer, and B. Webber, *The k_T-clustering algorithm for jets in deep inelastic scattering and hadron collisions*, *Phys. Lett.* **B 285** (1992) 291–299.

[112] F. Jacquet and A. Blondel, *Detection of the charged current event – Method II*, in *Proceedings of the Study for an ep Facility for Europe* (U. Amaldi, ed.), p. 391, 1979.

[113] J. E. S. Bentvelsen and P. Kooijman, *Reconstruction of (x, Q^2) and extraction of structure functions in neutral current scattering at HERA*, in *Proc. Workshop on Physics at HERA* (W. Buchmüller and G. Ingelman, eds.), p. 23, 1992.

[114] G. E. Wolf, *The Kinematics of deep inelastic scattering in the presence of initial state radiation*, DESY-97-047 (1997).

[115] K. Muchorowski, Ph.D. thesis, Warsaw University (1996).

[116] G. Wolf and K. Schilling, *How to analyse vector-meson production in inelastic lepton scattering*, *Nucl. Phys.* **B 61** (1973) 381.

[117] **ZEUS** Collaboration, S. Chekanov et al., *Exclusive ρ^0 production in deep inelastic scattering at HERA*, *PMC Phys.* **A 1** (2007) 6.

[118] A. Kreisel, *Exclusive ρ^0 electroproduction at HERA*, Ph.D. thesis, Tel Aviv University (2004).

[119] H. S. A. Kwiatkowski and H.-J. Möhring, HERACLES*: An event generator for ep interactions at HERA energies including radiative processes*, *Comp. Phys. Comm.* **69** (1992) 155. also in *Proc. Workshop Physics at HERA*, eds. W. Buchmüller and G.Ingelman, (DESY, Hamburg, 1991).

[120] J. Mainusch, F. Corriveau, R. Klanner, and G. Levman, *Influence of magnetic fields on the response of a uranium scintillator sampling calorimeter*, *Nucl. Instrum. Meth.* **A 312** (1992) 451–456.

[121] G. Hartner, *Private Communication.*

[122] R. Brun et al., GEANT*3*, CERN-DD-EE-84-1 (1987).

[123] R. Brun, F. Carminati, and S. Giani, GEANT – *Detector Description and Simulation Tool*, CERN-W5013, CERN-W-5013 (1994).

[124] **CTEQ** Collaboration, H. Lai et al., *Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions*, *Eur. Phys. J.* **C 12** (2000) 375–392.

[125] M. Gluck, E. Reya, and A. Vogt, *Photonic parton distributions*, *Phys. Rev.* **D 46** (1992) 1973–1979.

[126] M. Gluck, E. Reya, and A. Vogt, *Parton structure of the photon beyond the leading order*, *Phys. Rev.* **D 45** (1992) 3986–3994.

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

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

[129] L. Loennblad, *Ariadne version 4 - A program for simulation of QCD cascades implementing the colour dipole model*, *Comp. Phys. Comm.* **71** (1992) 15.

[130] O. Viazlo, *Measurement of beauty and charm cross sections at high Q^2 with the ZEUS experiment at the HERA collider*, Master's thesis, Kyiv University (2012).

[131] K. Nakamura et al., *Review of particle physics*, *J. Phys.* **G 37** (2010) 075021.

[132] E. Lohrmann, *A Summary of Charm Hadron Production Fractions*, `arXiv:1112.3757` (2011).

[133] **ZEUS** Collaboration, S. Chekanov et al., *Measurement of the charm fragmentation function in D^* photoproduction at HERA*, *JHEP* **04** (2009) 082.

[134] **OPAL** Collaboration, G. Abbiendi et al., *Inclusive analysis of the b quark fragmentation function in Z decays at LEP*, *Eur. Phys. J.* **C 29** (2003) 463–478.

[135] S. de Boer, *Backgrounds in charm and beauty secondary vertex analysis*, DESY Summer Student Report, available at `www.desy.de/f/students/2011/reports/deboer.pdf` (2011).

[136] M. Wing, *Setting the jet energy scale for the ZEUS calorimeter*, `hep-ex/0206036`.

[137] S. Alekhin, J. Blumlein, S. Klein, and S. Moch, *The 3, 4, and 5-flavor NNLO Parton from Deep-Inelastic-Scattering Data and at Hadron Colliders*, *Phys. Rev.* **D 81** (2010) 014032.

[138] A. Chuvakin, J. Smith, and B. Harris, *Variable flavor number schemes versus fixed order perturbation theory for charm quark electroproduction*, *Eur. Phys. J.* **C 18** (2001) 547–553.

[139] **ATLAS** Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B 716** (2012) 1–29.

[140] **CMS** Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B 716** (2012) 30–61.

[141] L. Evans, P. Bryant, et al., *LHC Machine*, *JINST* **3** (2008) S08001.

[142] **ATLAS** Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[143] G. Aad et al., *ATLAS pixel detector electronics and sensors*, *JINST* **3** (2008) P07007.

[144] **ATLAS** Collaboration, *ATLAS inner detector: Technical design report, Vol. 2*, CERN-LHCC-97-17.

[145] **ATLAS** Collaboration, M. Alam et al., *ATLAS pixel detector: Technical design report*, CERN-LHCC-98-13.

[146] I. Peric et al., *The FEI3 readout chip for the ATLAS pixel detector*, *Nucl. Instrum. Meth.* **A 565** (2006) 178–187.

[147] R. Beccherle et al., *MCC: The Module Controller Chip for the ATLAS pixel detector*, *Nucl. Instrum. Meth.* **A 492** (2002) 117–133.

[148] D. Arutinov et al., *Digital architecture and interface of the new ATLAS pixel front-end IC for upgraded LHC luminosity*, *IEEE Trans. Nucl. Sci.* **56** (2009) 388–393.

[149] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-019.

[150] A. L. Schorlemmer, *Monitoring radiation damage in the ATLAS Pixel Detector*, ATL-INDET-PROC-2012-014 (2012).

[151] K. Toms, *Radiation damage observations in the atlas pixel detector using the high voltage delivery system*, ATL-INDET-PROC-2011-041 (2011).

[152] M. Garcia-Sciveres et al., *The FE-I4 pixel readout integrated circuit*, *Nucl. Instrum. Meth.* **A 636** (2011) S155–S159.

[153] T. Haas, *A pixel telescope for detector R&D for an international linear collider*, *Nucl. Instrum. Meth.* **A 569** (2006) 53–56.

[154] P. Roloff, *The EUDET high resolution pixel telescope*, *Nucl. Instrum. Meth.* **A 604** (2009) 265–268.

[155] E. Corrin, *The EUDET High Resolution Beam Telescope - The Final Digital Readout*, EUDET-Report-2009-6 (2009).

[156] C. Hu-Guo et al., *First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope*, *Nucl. Instrum. Meth.* **A 623** (2010) 480–482.

[157] J. Baudot et al., *JRA-1 Milestone: Validation of the Final Sensor Equipping the EUDET Beam Telescope*, EUDET-Report-2009-25 (2009).

[158] A. C. Ramusino, *The EUDET Data Reduction Board (EUDRB)*, EUDET-Memo-2008-38 (2008).

[159] A. C. Ramusino, *The EUDET Data Reduction Board (EUDRB)*, EUDET-Memo-2007-36 (2007).

[160] D. Cussans, *Description of the JRA1 Trigger Logic Unit (TLU), v0.2c*, EUDET-Memo-2009-4 (2009).

[161] E. Corrin, *EUDAQ Software User Manual*, EUDET-Memo-2010-01 (2010).

[162] R. Turchetta et al., *A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology*, *Nucl. Instrum. Meth.* **A 458** (2001) 677–689.

[163] A. Bulgheroni et al., *EUTELESCOPE: tracking software*, EUDET-Memo-2007-20 (2007).

[164] ILCSOFT reconstruction and analysis software package for ILC, available at `http://ilcsoft.desy.de/portal/`.

[165] F. Gaede et al., *LCIO - A persistency framework for linear collider simulation studies*, `physics/0306114v1`.

[166] F. Gaede et al., "*LCIO - Users manual.*" available at `http://lcio.desy.de`.

[167] F. Gaede and J. Engels, *Marlin et al - A Software Framework for ILC detector R&D*, EUDET-Report-2007-11 (2007).

[168] L. Rossi, P. Fischer, T. Rohe, and N. Wermes, *Pixel Detectors. From Fundamentals to Applications.* Springer Verlag (2005).

[169] E. Belau, R. Klanner, G. Lutz, E. Neugebauer, H. Seebrunner, et al., *Charge collection in silicon strip detectors*, *Nucl. Instrum. Meth.* **214** (1983) 253.

[170] V. Blobel, *Millepede II - Linear Least Squares Fits with a Large Number of Parameters*, Program Manual, available at `http://www.desy.de/ blobel/Mptwo.pdf` (2007).

[171] G. Flucke, P. Schleper, G. Steinbruck, and M. Stoye, *CMS silicon tracker alignment strategy with the Millepede II algorithm*, *JINST* **3** (2008) P09002.

[172] I. Belotelov, A. Lanyov, and G. Ososkov, *A study of the millepede alignment algorithm on the Monte Carlo model of the HERA-B outer tracker*, *Phys. Part. Nucl. Lett.* **3** (2006) 261–272.

[173] V. Blobel and C. Kleinwort, *A new method for the high-precision alignment of track detectors*, *Contribution to Advanced Statistical Techniques in Particle Physics Conference* (2002).

[174] A. Żarnecki and P. Nieżurawski, *EUDET Telescope Geometry and Resolution Studies*, EUDET-Report-2007-01 (2007).

[175] A. Żarnecki, *Analytical track fittng: performance and new options*, EUDET-Memo-2010-14 (2010).

[176] M. Mathes et al., *Test Beam Characterization of 3-D Silicon Pixel Detectors*, *IEEE Trans. Nucl. Sci.* **55** (2008) 3731–3735.

[177] C. Lapoire, *Operational experience with the ATLAS Pixel Detector at the LHC*, *Nucl. Instrum. Meth.* **A 699** (2013) 149–152.

[178] M. Backhaus et al., *Development of a versatile and modular test system for ATLAS hybrid pixel detectors*, *Nucl. Instrum. Meth.* **A 650** (2011) 37–40.

[179] The USBpix test system homepage: `http://icwiki.physik.uni-bonn.de/twiki/bin/view/Systems/UsbPix`.

[180] J. Janssen, *Development of an FPGA-based FE-I3 pixel readout system and characterization of novel 3D and planar pixel detectors*, Master's thesis, Bonn University, BONN-IB-2010-08 (2010).

[181] S. Schultes, *Entwicklung eines Strahl-Teleskops mit ATLAS-Pixeldetektoren*, Master's thesis, Bonn University, BONN-IB-2012-09 (2011).

Acknowledgements

After the quest is over, I wish to sincerely thank all the people, without whom this work would never have been accomplished.

First of all, thanks goes to **Olaf Behnke**, the driving force behind the heavy flavour measurement part of this thesis. Thank you, Olaf, for all your constant support and encouragement! Your enthusiasm was always motivating, your experience was invariably guiding the way, your jokes were repeatedly helping in tough moments!

On the hardware side of the thesis, I would like to thank **Ingrid-Maria Gregor**. Thanks, Ingrid, for introducing me into the world of Particle Detectors and for giving me an opportunity to work on the ATLAS Pixel beam test program! You were always providing a lot of inspiration and stimulation to work hard, but it was always a lot of fun as well!

A number of people supported me in the course of these studies, whom I am very much pleased to thank here:

- **Prof. Robert Klanner** – for critically reviewing the work and for providing valuable comments on the manuscript;
- **Philipp Roloff** – for introducing me into the secondary vertex analysis and for tremendous help at initial stages of the work;
- **Achim Geiser** – for deep involvement in the analysis at all stages;
- **Sasha Viazlo** – for his dedicated Monte Carlo studies;
- **Ulrich Kötz** – for providing the right guidance and equipment in the laboratory;
- **Jörg Behr** – for help with the EUTelescope software;
- **Igor Rubinsky**, **Jens Weingarten**, **Matthias George** and everyone from the ATLAS Planar Pixel and IBL test beam groups – for nice team work on ATLAS pixels;
- ZEUS Heavy Flavour group and DESY-Pixel group – for many fruitful discussions.

Special acknowledgement to **Vladimir Y. Aushev** who back in 2007 proposed me to work on the ZEUS experiment for my bachelor thesis, and thus opened to me the doors of Particle Physics!

Time to recall **my friends** – my colleagues here at DESY as well as my "old" friends in the fatherland – thanks for a lot of fun during these years!

Finally, I am so happy to express heartfelt gratitude to **my lovely family**. Not to speak that your steady support and firm trust helped me all the way to finish this work – it's fair to say that I would never become a physicist without you!

THANK YOU ALL!