Measurement of beauty quark production in deep inelastic scattering at HERA

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Vincenzo Chiochia
aus Turin

Hamburg
2003

Gutachter der Dissertation:	Prof. Dr. R. Klanner Prof. Dr. P. Schleper
Gutachter der Disputation:	Prof. Dr. E.Lohrmann Prof. Dr. B.Naroska
Datum der Disputation:	15.04.2003
Vorsitzender des Prüfungsausschusses:	Dr. E. Fretwurst
Vorsitzender des Promotionsausschusses:	Prof. Dr. G.Huber
Dekan des Fachbereichs Physik:	Prof. Dr. Fr.-W. Büßer

To my parents

Ästhetik ist eine spezielle Art der Datenverarbeitung
und deswegen ein Teilgebiet der Informatik.
(Frieder Nake)

Beauty is ten, nine of which is dressing.
(Azerbaijani Proverb)

Ugliness is in a way superior to beauty
because it lasts.
(Serge Gainsbourg)

Even beauties can be unattractive.
If you catch a beauty in the wrong light
at the right time, forget it.
I believe in low lights and trick mirrors.
I believe in plastic surgery.
(Andy Warhol)

Abstract

In this dissertation the measurement of *beauty* quark (or b-quark) production in deep inelastic *electron-proton* collisions at HERA is presented. The measurement is based on the data collected by the ZEUS experiment during the years 1999-2000 and corresponds to an integrated luminosity of 60 pb^{-1}. The reaction is

$$e^+ \; p \to e^+ \; b \; \bar{b} \; X \to e^+ \; \mu \; \text{Jet} \; X,$$

where at least one of the b-quarks decays semi-leptonically into a muon and a jet. Events are selected containing the scattered positron, a hard jet reconstructed in the Breit frame and a muon traversing the barrel and rear muon chambers of the ZEUS experiment. The fraction of events from b-decays is extracted on a statistical basis using the distribution of the muon transverse momentum relative to the closest jet.

QCD predictions to the next-to-leading order (NLO) accuracy are obtained for this process and a detailed analysis of the NLO QCD uncertainties is carried out. Both the total and single differential cross sections are almost a factor 1.4 above the NLO QCD predictions but compatible within the experimental and theoretical uncertainties. The result is also compatible with the prediction of the `CASCADE` Monte Carlo program, based on k_T-factorization and CCFM QCD evolution equation. Other Monte Carlo predictions lie below the data.

A more precise measurement of beauty quark production will be possible after the luminosity upgrade of the HERA collider. To fully benefit from the new experimental conditions, the ZEUS detector was upgraded by an improved tracking system consisting of a Straw Tube Tracker and a silicon Micro Vertex Detector (MVD). The latter device will allow a precise reconstruction of the heavy meson decays.

In this dissertation the results of test beam measurements performed on the MVD half modules are presented. The charge collection, the signal/noise ratio and the hit resolution are measured for different incidence angles using existing and newly developed hit reconstruction algorithms. Detailed studies of the charge sharing mechanism are performed on the regions of the MVD sensors corresponding to the bonding pads and on the overlap between single sensors.

Zusammenfassung

Diese Dissertation befasst sich mit der Messung der Beauty (b)-Produktion in Tief-Inelastischen Streuprozessen bei HERA. Analysiert wurden die Messergebnisse der ZEUS-Kollaboration der Jahre 1999-2000; die analysierte Datenmenge entspricht einer integrierten Luminosität von 60 pb^{-1}. Untersucht wurde folgender Prozess:

$$e^+ \; p \to e^+ \; b \; \bar{b} \; X \to e^+ \; \mu \; \text{Jet} \; X.$$

Hierbei muss mindestens eines der b-Quarks semileptonisch in ein Muon und einen Jet zerfallen.

Ausgewählte Ereignisse enthalten das gestreute Positron, einen im Breit-Bezugssytem rekonstruierten Jet und ein Muon, das den "barrel" und die hinteren Muon-Kammern durchläuft. Der Anteil der b-Zerfälle wird unter Zuhilfenahme der Verteilung des Transversalimpulses der Muonen relativ zum nahegelegensten Jet auf statistischer Basis bestimmt.

Für diesen Prozess konnten QCD-Vorhersagen der Ordnung α_s^2 (NLO) gemacht werden und eine detaillierte Analyse der NLO QCD-Unsicherheiten durchgeführt werden. Desweiteren wurden die Messungen mit Monte-Carlo-Programmen verglichen, die auf Matrixelementen in niedrigster Ordnung Stëurungstheorie basieren und verschiedene QCD-Entwicklungsgleichungen benützen.

Der Umbau des HERA-Beschleunigers zu Erhöhung der Luminosität wird präzisere Messungen ermöglichen. Um die neuen experimentellen Bedingungen nach diesem Umbau voll ausschöpfen zu können, wurde im ZEUS-Detekor ein verbessertes Spur-Rekonstruktions-System eingebaut. Dieses System besteht aus einem *Straw Tube Tracker* und einem Silizium-*Micro Vertex Detector* (MVD). Der MVD wird eine genaue Rekonstruktion der Zerfälle schwerer Mesonen ermöglichen.

Neben der Analyse des obig dargestellten Zerfallsprozesses schwerer b-Quarks werden in dieser Dissertation die Ergebnnisse von Messungen an den MVD-Halbmodulen vorgestellt. Gemessen wurden Ladung, das Signal zu Rauschen-Verhältniss sowie die Auflösung der "Hits" unter verschiedenen Einfallswinkeln. Hierzu wurden sowohl bestehende als auch neu entwickelte "Hit"-Rekonstruktions-Algorithmen verwendet.

In den Bereichen der MVD-Sensoren, die "bonding pads" und überlappenden einzelnen Sensoren entsprechen, wurden detaillierte Untersuchungen des Ladungsteilungs-Mechanismus durchgeführt.

Contents

Introduction

The actual knowledge of the structure of matter is contained in the so called *Standard Model* (SM) of particle physics. According to the SM the universe we live in is composed of elementary particles: the *quarks*, the *leptons* (as the electron and the muon) and the *bosons* (as the photon). The quarks and leptons interact exchanging bosons. The latter mediate the four fundamental interactions of nature; from the weakest to the strongest they are: *gravity*, the *weak* interaction, *electromagnetism* and the *strong* interaction.

All particles predicted by the SM have been detected by the experiments except the so called *Higgs* boson. The search for this particle as well as other particles appearing in extensions of the SM is the main motivation for the ongoing construction of the collider with the highest center-of-mass energy ever realized, the *Large Hadron Collider* at CERN in Geneva.

According to the SM there are six quarks (and the respective anti-quarks). From the lightest to the heaviest they are called: *up, down, strange, charm, beauty* (or *bottom*) and *top*. The quarks are not detected as free particles but they are always bound with an anti-quark to form *mesons* (as the *pion*) or with two other quarks to form *baryons* (as the *proton* and the *neutron*). The quark masses range from few MeV (i.e. a million electron-Volts) for the *up* quark to hundreds of GeV (or billion electron-Volts) for the *top* quarks. The values of the quark masses are not predicted by the SM.

Nuclei, atoms, and molecules are composed by protons, neutrons and electrons. Protons and neutrons are composed by *up* and *down* quarks. The *charm*, *beauty* and *top* quarks are generally called heavy quarks, since their mass is larger than the energy scale at which the quarks are confined into hadrons, $\Lambda \sim 250$ MeV.

In nature, heavy quarks are produced only in high energy collisions, e.g. interaction of cosmic rays with ordinary matter, and their existence is ephemeral. Particles containing heavy quarks decay after a fraction of a second. Therefore, to study their properties they are produced in a controlled environment in which their decay products can be detected, e.g. particle accelerators and detectors.

The first observation of the *beauty* quark was done during the summer of 1977. A team of physicists, led by Leon M. Lederman and working on experiment 288 in the proton center beam line of the Fermilab fixed target areas, discovered the Υ (Upsilon) meson. This discovery was eventually understood as being the bound state of the *beauty* quark and its anti-quark. Their result was confirmed in experiments conducted in 1978 at Fermilab, the CERN ISR in Geneva, Switzerland and DORIS at DESY in Hamburg, Germany. The clarity of that discovery combined with the 1974 discovery of the J/ψ particle, a bound state of a *charm* with an anti-*charm* quark, made even the most skeptical physicists believe in quarks.

Nowadays, *beauty* quarks (or b-quarks) are produced in several experiments of high energy physics. Both the production and decay properties of b-quarks are extremely interesting. Decays of b-quarks (as well as the decay of c-quarks) are used to measure the violation of CP

symmetry, a combination of two Lorentz transformations: the *charge conjugation (C)* and the *parity transformation (P)*. Indeed two new 'factories' of B-mesons, PEP-II at SLAC (USA) and KEK-B at KeK (Japan), with the goal of finding CP violation in the B system have begun to operate and indeed confirmed CP violation.

On the other hand, the measurement of b-quark production is an important test of the theory of strong interactions or *Quantum Chromodynamics* (QCD). Generally speaking, heavy quark production in deep inelastic (i.e. at high virtualities of the exchanged photon) electron-proton scattering is calculable in QCD and provides information on the gluonic content of the proton which is complementary to that obtained from other measurements (as from the measurements of the scaling violation of the proton structure function). In addition, the scale of the hard scattering may be large relative to the mass of the heavy quark, thus allowing to study whether and when to treat the heavy quark as a massless parton.

In the last years the study of b-quark production has been an intriguing topic since several experiments at *proton-antiproton* colliders (like UA1, CDF and D0) have measured production rates higher than was predicted by the most precise QCD calculations available at that time. Similar measurements are also performed in $\gamma\gamma$ processes at *electron-positron* colliders (as LEP) and at the HERA *electron-proton* collider. As we discuss in this dissertation, the correct description in terms of QCD of the b-quark production rate is still an open problem.

The purpose of this dissertation is to present the measurement of b-quark production in deep inelastic *electron-proton* collisions at HERA. The corresponding reaction is

$$e\ p \to e\ b\ \bar{b}\ X$$

where at least one of the b-quarks decays semi-leptonically into a muon via a charged current interaction. The measurement is performed in a phase space region where the virtuality Q^2 of the photon exchanged between the electron and the proton is greater than 2 GeV^2. The measurement is based on the data collected by the ZEUS experiment during the years 1999-2000, corresponding to an integrated luminosity of 60 pb^{-1}. This is the first measurement of this kind performed at the ZEUS experiment. Furthermore, the single differential cross sections are measured at HERA for the first time.

To improve the existing measurements of heavy quark production, and more generally the tracking capabilities of the experiment, the ZEUS detector was recently equipped with a silicon Micro Vertex Detector (MVD). This device will allow a precise reconstruction of the heavy meson decays. At the same time the HERA collider was upgraded to reach higher luminosities. In this dissertation we discuss the future prospects of heavy flavour physics at HERA and we present the beam test measurements performed on the MVD half modules. These measurements indicate that the MVD can achieve a superb precision in the hit reconstruction. To this purpose, several reconstruction algorithms are discussed and their performance are compared.

***

This dissertation is organized as follows. In **Chapter 1** we give an introduction to the theory of heavy quark production in deep inelastic *electron-proton* (*ep*) collisions. The emphasis of this chapter is pedagogical. The main steps of the calculation are shown at the lowest order of perturbative QCD and the features of higher order corrections are discussed. In the discussion we will emphasize the implementation of the QCD calculation into software tool such as Monte Carlo programs and NLO QCD cross section calculation programs. The former are generally used in high energy physics as a common tool to correct the measurements for detector effects. The latter are a versatile way to produce QCD predictions for the measured observables with a NLO accuracy. In this Chapter we will also discuss the different QCD evolution equations, proposed to describe the dependency of the parton distribution functions on the QCD scales, and the main concepts of jet physics in the context of *ep* collisions. The fragmentation of heavy quarks into hadrons will conclude the theoretical overview. At the end of the chapter we present the recent measurements of *beauty* quarks performed at HERA.

In **Chapter 2** the HERA collider and the ZEUS detector are presented. The detector components used in our analysis and the trigger system are described in details.

Chapter 3 covers the event reconstruction. The cross section measurement is based on the selection of events containing three ingredients: a scattered positron detected in the calorimeter, a hard jet of particles and a muon detected in the muon chambers close to the jet. In addition we describe the different methods used to reconstruct the event kinematics.

Chapter 4 is dedicated to the description of the event selection. Cuts are applied to select a sample of events from inclusive deep inelastic scattering. Further selection cuts are applied on the reconstructed muon and jet. The motivation of the selection cuts are discussed in detail.

Once the final sample of events is selected we calculate the fraction of events corresponding to decays of *beauty* quarks and distinguish them from other particle decays appearing in the sample. The procedure used to extract this fraction and to calculate the corresponding cross section is described in **Chapter 5**. In addition, the kinematic distributions of the final event sample are compared to the prediction of Monte Carlo programs, using the extracted fractions. We conclude the chapter presenting the results of the cross section measurements and describing the related uncertainties.

In **Chapter 6** the QCD predictions for our measurement are discussed. The NLO QCD predictions are presented and their uncertainties are described in details. The experimental results are compared to the NLO QCD predictions and to Monte Carlo programs implementing different QCD evolution equations.

The luminosity upgrade of the HERA collider and of the ZEUS detector is presented in **Chapter 7**. We discuss some selected topics of heavy flavour production in the light of a large integrated luminosity and we describe the modifications in the HERA interaction regions. In this chapter we describe the ZEUS Micro Vertex Detector (MVD) and we present the results of the beam test measurements performed on the MVD half modules. The performance as a function of the bias voltage are shown and the hit resolution is measured as a function of the incidence angle. The performance of the MVD sensors are investigated in different regions, i.e. where probe or bond pads are present and in the overlap region between single sensors. We conclude this chapter presenting the test of a MVD full module with prototype units of the final read-out chain.

1 Heavy quark production in deep inelastic e-p collisions

In this chapter we discuss the theoretical aspects of heavy quark production in electron-proton (ep) collisions. After a brief introduction to the kinematics of the ep scattering we will show the main steps in the calculation of the heavy quark contribution to the proton structure function up to the order $\mathcal{O}(\alpha_s^2)$. Later, the role of QCD evolution equations will be clarified and the main aspects of their implementation will be discussed. The theoretical principles of jet physics in deep inelastic scattering will be presented. We will conclude our theoretical overview describing the fragmentation of heavy quarks into hadrons and their role in the evaluation of QCD predictions. The chapter ends with a review of the recent measurements of *beauty quarks* in ep collisions.

1.1 Kinematics of ep scattering

In the Standard Model the interaction between an electron (or positron) and a proton is realized through the exchange of a vector boson. If the boson is a photon (γ) or a Z^0 the process is called *neutral current* scattering (NC, Figure 1.1-a). In addition, a charged $W^\pm$ boson can be exchanged. In this case we call the reaction *charged current* scattering (CC, Figure 1.1-b).

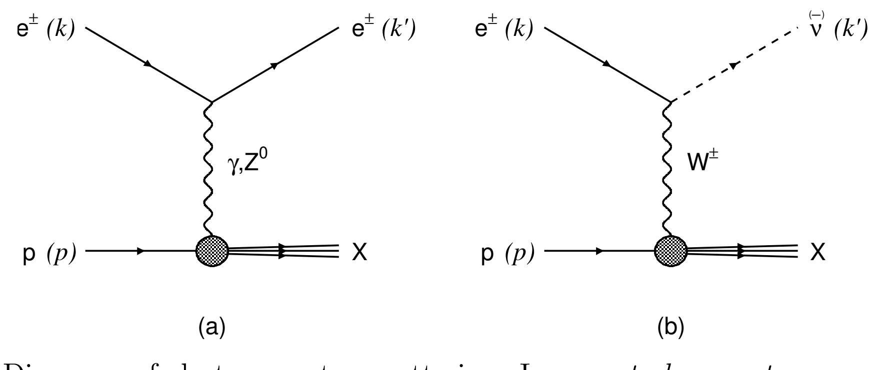

Figure 1.1 Diagrams of electron-proton scattering. In a *neutral current* process a photon or a Z^0 is exchanged (a). In a *charged current* process the exchanged boson is a charged $W^\pm$ (b),

Let k, k' and P be the four-momenta of the incoming electron, the scattered electron and proton, respectively. The four-momentum q of the boson exchanged between the electron and the proton is given by

$$q = k - k'. \tag{1.1}$$

The scattering process is characterized by the following Lorentz scalars[1]:

$$\begin{aligned} Q^2 &= -q^2 = (k-k')^2 && (1.2)\\ s &= (k+P)^2 \simeq 2k\cdot P && (1.3)\\ y &= \frac{P\cdot q}{P\cdot k} \simeq \frac{2P\cdot q}{s} && (1.4)\\ x &= \frac{Q^2}{2P\cdot q} && (1.5) \end{aligned}$$

where s is the square of the total center-of-mass energy, y and x are the *Bjorken scaling variables*. The variable y measures the energy fraction from the electron transferred to the interaction in the proton rest frame. Q^2 is the virtuality of the exchanged boson and corresponds to the resolution of the probe. A structure in the proton is resolved if its dimension is of the order of the probe wavelength $\lambda \simeq \hbar/\sqrt{Q^2} \simeq 0.197/\sqrt{Q^2}$ fm. The variables x, y and Q^2 are related by the square of the center of mass energy

$$Q^2 = sxy. \tag{1.6}$$

In addition we introduce the variable

$$W^2 = (q+P)^2 = q^2 + P^2 + 2q\cdot P \simeq -Q^2 + ys, \tag{1.7}$$

which is the invariant mass of the system recoiling against the scattered lepton and can be interpreted as the center of mass energy of the virtual boson-proton system.

1.2 The neutral current Deep Inelastic Scattering cross section

We now concentrate on the case in which the electron-proton scattering is realized through the exchange of a photon. In addition, the words *deep* and *inelastic* refer to the conditions $Q^2 \gg M^2$ and $W^2 \gg M^2$, respectively, where M is the mass of the proton. These conditions ensure that perturbative QCD is applicable for the cross section calculation and that we are above the region of nucleon resonance production. In fact, for low values of Q^2 one expects just to excite the proton into a Δ-state, e.g. as in the reaction $ep \to e\Delta^+ \to ep\pi^0$. In these processes, the invariant mass of the photon-proton system is $W^2 \simeq M_\Delta^2$.

In a deep inelastic scattering (DIS) process the double differential cross section for the reaction $ep \to eX$ is given by

$$\frac{d\sigma}{dy dQ^2} = \frac{1}{k\cdot P}\frac{\pi\alpha^2}{Q^4} L_{\mu\nu}W^{\mu\nu} \tag{1.8}$$

where α is the electromagnetic coupling constant and $L_{\mu\nu}$ and $W^{\mu\nu}$ are the leptonic and hadronic tensors respectively, given by [1]

$$\begin{aligned} L_{\mu\nu} &= 2k_\mu k'_\nu + 2k_\nu k'_\mu + q^2 g_{\mu\nu} && (1.9)\\ MW_{\mu\nu} &= \left(\frac{q_\mu q_\nu}{q^2} - g_{\mu\nu}\right) F_1(x,Q^2) + \left(P_\mu - \frac{P\cdot q}{q^2}q_\mu\right)\left(P_\nu - \frac{P\cdot q}{q^2}q_\nu\right) F_2(x,Q^2). && (1.10) \end{aligned}$$

[1] The relations 1.3 and 1.4 are obtained neglecting the proton and electron mass.

In the latter, $F_{1,2}$ are arbitrary functions of the kinematic variables x and Q^2 describing the $\gamma p \to X$ vertex. These functions, called *structure functions*, have to be obtained from the experiment and parameterize the structure of the proton. Substituting the expressions of the leptonic and hadronic tensors in Eq. 1.8 we obtain the relation

$$\frac{d\sigma}{dydQ^2} = \frac{4\pi\alpha^2}{yQ^4}[(1-y)F_2(x,Q^2) + y^2 x F_1(x,Q^2)]. \tag{1.11}$$

In order to better clarify the meaning of the structure functions, it is convenient to rewrite Eq. 1.11 as a product of the flux of virtual photons, $\Gamma(y,Q^2)$, with the cross section of the photon-proton interaction, $\gamma^* p \to X$:

$$\frac{d\sigma}{dydQ^2} = \Gamma(y,Q^2)[\sigma_T(\gamma^* p \to X) + \epsilon\sigma_L(\gamma^* p \to X)] \tag{1.12}$$

where ϵ is the ratio of the flux of longitudinal and transverse polarized photons

$$\epsilon = \frac{1-y}{1-y+y^2/2} \tag{1.13}$$

and the flux of virtual photons is given by

$$\Gamma(y,Q^2) = \frac{\alpha}{yQ^2\pi}\left(1-y+y^2/2\right). \tag{1.14}$$

The cross sections $\sigma_{L,T}$ contain the physics of the hard scattering process and can be calculated using a perturbative approach. They are related to the structure function F_2 through the relation

$$F_2(y,Q^2) = \frac{Q^2}{4\pi^2\alpha}[\sigma_T(\gamma^* p \to X) + \sigma_L(\gamma^* p \to X)]. \tag{1.15}$$

The computation of the γ^*-proton cross section is performed by assuming the proton to be composed of *point-like partons*, where point-like means a structure-less Dirac particle. In order to obtain the exact kinematics, we also have to assume the proton to move with infinite momentum, $|p| \gg M$, so that all masses can be neglected. In this frame, the relativistic time dilation slows down the rate at which the partons interact with each others. Therefore, the parton is essentially a free particle when interacting with the virtual photon.

An important consequence is that we can imagine the γ^*-proton cross section a superimposition of *probabilities* of scattering from single free particles. These probabilities are given by γ^*-parton cross sections, convoluted with the probability to pick up that parton from the proton in a certain range of momentum

$$\sigma_{\gamma^* P} = \sum_{i=partons} \int_x^1 \sigma_{\gamma^* i} \cdot f_i(\xi)\ d\xi \tag{1.16}$$

where ξ is the proton momentum fraction carried by the incoming parton. Figure 1.2 shows the first terms appearing in the sum of amplitudes representing the proton structure function F_2.

The diagrams in the first row represent the scattering process $\gamma q \to q$ and the related QCD virtual corrections. This process is usually called *quark parton model* (QPM), since it describes the photon-proton interaction as the sum of incoherent photon-quark elastic scatterings. The

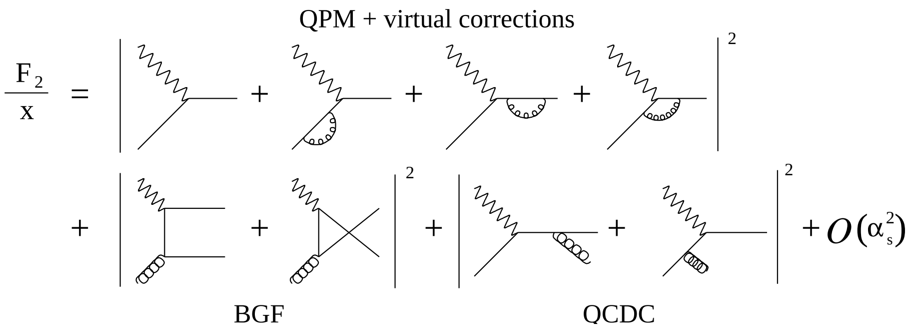

Figure 1.2 Expansion of the proton structure function F_2 up to the order $\mathcal{O}(\alpha\alpha_s)$. The Feynman diagrams show the contributions of the quark-parton-model and its virtual corrections, the boson-gluon-fusion and QCD-Compton.

quarks are assumed to be the pointlike constituents of the proton. The model predicts the scaling of structure functions $F_{1,2}$ or, in other words, the fact that $F_{1,2}$ are functions of a single variable, the Bjorken scaling variable x. Since $F_{1,2}$ do not depend on Q^2 in the QPM model the proton looks the same for all values of the probe resolution.

With the advent of QCD it became clear that the quarks cannot be considered completely freely moving in the proton when struck by the photon, but they can exchange gluons, the carriers of the strong force. The quark parton model is, therefore, only the zero-th order approximation in a perturbative expansion in the strong coulping constant α_s. The next terms appearing in the perturbative series of F_2 correspond to the process $\gamma^* g \to q\bar{q}$, in which the photon interacts with a gluon from the proton (Figure 1.2-b). The process is called *photon-gluon-fusion* or, more generically, *boson-gluon-fusion* (BGF). An additional $\mathcal{O}(\alpha_s)$ correction is given by the process in which a gluon is emitted from the quark line before or after the scattering with the photon ($\gamma^* q \to qg$). This process is called *QCD Compton-scattering* (QCDC).

The inclusion of the QCD corrections to the QPM implies that the scaling of the structure functions $F_{1,2}$ is violated. Figure 1.3 shows the proton struction function F_2 as a function of Q^2, for different values of the Bjorken variable x. For $x \geq 0.1$, the structure function is independent of Q^2 as predicted by the QPM. For $x \ll 0.1$ the structure function increases for increasing values of Q^2, that is, the number of partons resolved by the photon which share the total proton's momentum increases with Q^2. There is an increased probability of finding a quark at small x and a decreased probability of finding one at high x, because high-momentum quarks lose momentum by radiating gluons [1]. The dependence of the structure functions on Q^2 is described by the QCD evolution equations as the Altarelli-Parisi equations. In Section 1.5 the QCD evolution equations will be explained in more details.

1.3 The heavy quark contribution to F_2

Two different production mechanism have been proposed to describe the production of heavy quarks in neutral current DIS reactions[2]. According to the first mechanism, usually called *intrinsic heavy quark production* one assumes that, besides the light quarks u, d and s and the gluon g the wave function of the proton also consists of the heavy quarks c, b, t [2, 3]. In the

[2]We call *heavy* the quarks with mass bigger than the QCD scale, $m_Q > \Lambda_{QCD} \sim 250$ MeV.

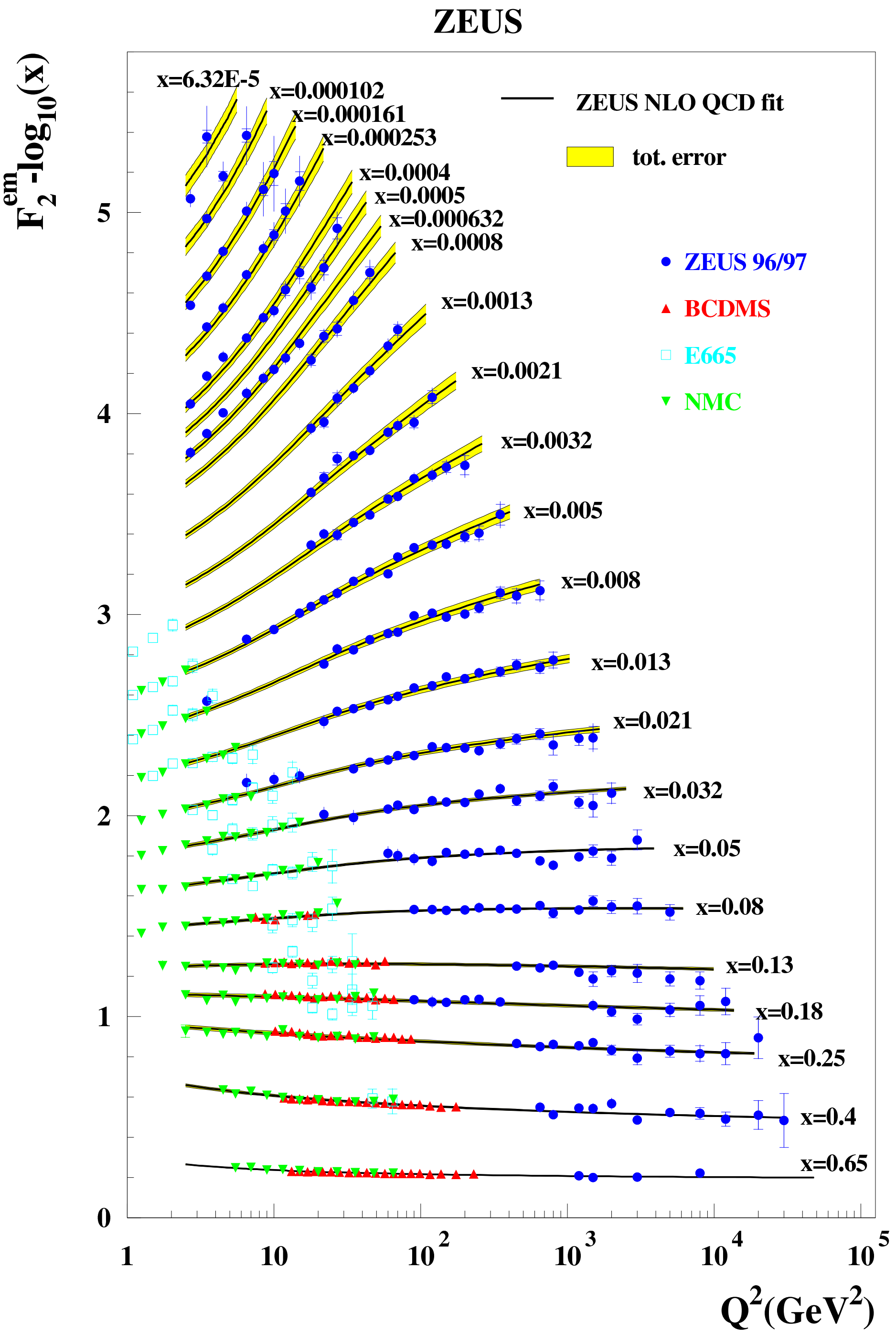

Figure 1.3 The proton structure function F_2 as function of the virtuality of the exchanged photon, for different values of the Bjorken scaling variable x. The experimental points are measured by the ZEUS collaboration and several fixed target experiments. They are compared with a fit performed using next-to-leading order QCD predictions. The error bands on the fit represent the total experimental uncertainty.

context of the QCD improved quark parton model this means that the heavy quark emerges directly from the proton and interacts with the virtual photon. The consequence is that the cross section depends on a heavy quark distribution $f_Q(z)$ in the proton. Some diagrams appearing in the calculation are shown in Figure 1.4. The lowest order process is $\gamma^* + Q \to Q$ whereas including the QCD radiation we also have $\gamma^* + Q \to Q + g$, where Q is a heavy quark.

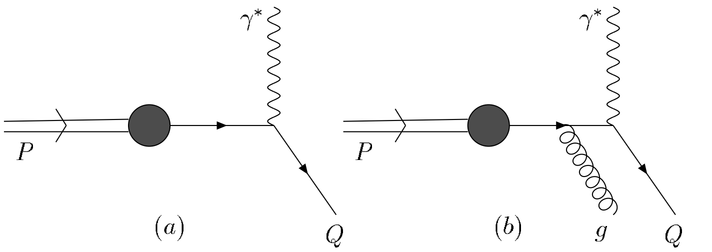

Figure 1.4 Intrinsic heavy quark production. The diagrams of the $\gamma^* + Q \to Q$ process (a) and of the $\gamma^* + Q \to Q + g$ process (b) are shown.

The second mechanism is called *extrinsic heavy quark production*. In this case the proton wave function does not contain the heavy quark components. In the lowest order perturbation theory the heavy quark and heavy anti-quark appear in pairs and are produced via BGF as shown in Figure 1.5. Here the gluon emerges from the proton in a similar way as the heavy quark in Figure 1.4. In a BGF process the quarks can be *heavy* if the squared center-of-mass

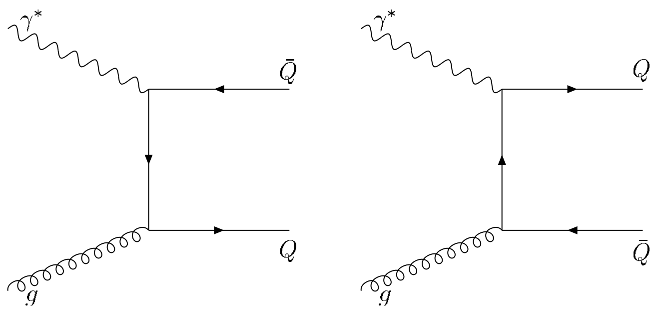

Figure 1.5 Extrisic heavy quark production via Boson-Gluon-Fusion

energy of the $\gamma^* g$ system, $\hat{s}$, is

$$\hat{s} = (\gamma^* + g)^2 = (q + \xi P)^2 > (2m_Q)^2, \tag{1.17}$$

where m_Q^2 is the mass of the heavy quark and the photon and gluon four-momenta are respectively γ and $g = \xi P$. In this context the process

$$\gamma^* + g \to Q + \bar{Q} \tag{1.18}$$

represents the *lowest order diagram in which heavy quarks can be produced in deep inelastic ep scattering* or, in other words, the BGF process is the lowest order heavy quark contribution, $F_{1,2}^{heavy}$ to the proton structure function $F_{1,2}$.

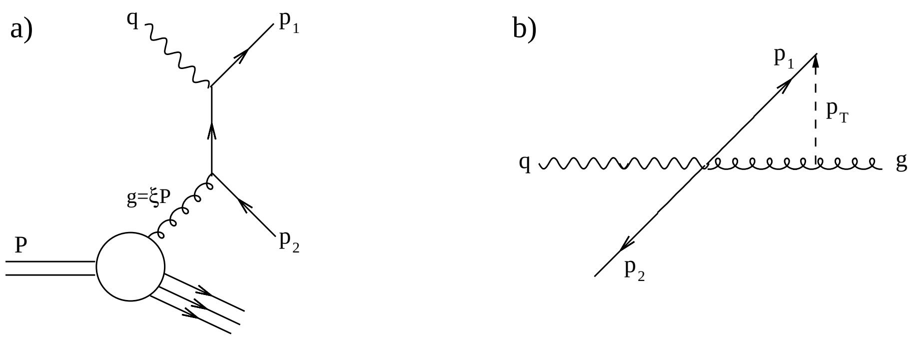

Figure 1.6 a) Feynman diagram of the photon-gluon-fusion process. b) Sketch of photon-gluon-fusion kinematics in the $\gamma^* g$ center-of-mass frame.

The main difference between the two production mechanisms is that for extrinsic heavy quark production two heavy particles are produced in the final state instead of one as in the case of the intrinsic heavy quark approach. This implies a different transverse momentum distribution of the two cases. So far, the experiments performed at HERA confirm the spectrum of transverse momentum predicted by the extrinsic mechanism, therefore, in our discussion we will only cover this approach.

In what follows, we will summarize the main steps of the BGF calculation for massless quarks and we will introduce the quark masses at the end of our discussion. In addition, we will discuss how the calculation is implemented in the programs used in this analysis.

1.3.1 Boson-Gluon-Fusion for light quarks

The BGF diagram with the relative particle four-momenta are shown in Figure 1.6-a. We introduce the Mandelstam variables

$$\begin{aligned} \hat{u} &= (g-p_1)^2 = g^2 + p_1^2 - 2gp_1 \simeq -2\xi P p_1 & (1.19)\\ \hat{t} &= (g-p_2)^2 = g^2 + p_2^2 - 2gp_2 \simeq -2\xi P p_2 & (1.20)\\ \hat{s} &= (q+g)^2 = q^2 + g^2 + 2qg \simeq -Q^2 + 2q\xi P. & (1.21) \end{aligned}$$

The three variables are related by the following expression

$$\hat{s} + \hat{t} + \hat{u} = -q^2 = Q^2. \quad (1.22)$$

The scattering amplitudes of the subprocess $\gamma^* g \to q\bar{q}$ are calculated by applying the Feynman rules to the diagram in Figure 1.6-a, to its crossed diagram and to the interference term.

Following the prescription discussed in the previous section, we distinguish the cross sections for transverse and longitudinally polarized photons, respectively named $\hat{\sigma}_T$ and $\hat{\sigma}_L$. The differential cross sections are obtained from the scattering amplitudes, $\mathcal{M}_{L,T}$, and the following relation

$$\frac{d\hat{\sigma}_{L,T}}{d\hat{t}} = \frac{1}{16\pi} \frac{1}{(\hat{s}+Q^2)^2} |\mathcal{M}_{L,T}|^2. \quad (1.23)$$

thus [4]

$$\begin{aligned}
\frac{d\hat{\sigma}_T}{d\hat{t}}(\gamma^* g \to q\bar{q}) &= \frac{\pi}{(\hat{s}+Q^2)^2} e_q^2 \alpha\alpha_s \left[\frac{\hat{u}}{\hat{t}} + \frac{\hat{t}}{\hat{u}} - 2\frac{\hat{s}Q^2}{\hat{u}\hat{t}} + \frac{4\hat{s}Q^2}{(\hat{s}+Q^2)^2}\right] &(1.24)\\
\frac{d\hat{\sigma}_L}{d\hat{t}}(\gamma^* g \to q\bar{q}) &= \frac{8\pi}{(\hat{s}+Q^2)^2} e_q^2 \alpha\alpha_s \left[\frac{Q^2\hat{s}}{(\hat{s}+Q^2)^2}\right]. &(1.25)
\end{aligned}$$

where e_q is the charge of the outgoing quark. The differential cross sections Eqs. 1.24 and 1.25 must be integrated over the variable $\hat{t}$. The longitudinal cross section σ_L gives a finite result for all $\hat{t}$ values but the transverse cross section diverges for $\hat{t} \to 0$ or $\hat{u} \to 0$. The origin of these divergences can be clarified by expressing the transverse momentum of any of the outgoing quarks, p_T, in the $\gamma^* g$ center of mass frame and using the Mandelstam variables (see Figure 1.6-b)

$$p_T^2 = \frac{\hat{s}\hat{t}\hat{u}}{(\hat{s}+Q^2)^2} \tag{1.26}$$

If $\hat{t} \to 0$ also $p_T \to 0$, therefore, the differential cross section Eq. 1.24 diverges when the outgoing quark is collinear to the $\gamma^* g$ axis. The singularity, usually called *infrared* is regularized using the *massive gluon* scheme or the *dimensional regularization* scheme[3]. In most of the Monte Carlo implementations the integration is performed by introducing a cut-off value for the transverse momentum, $p_{T,min}$. In this case, from Eqs. 1.24 and 1.25 we obtain

$$\begin{aligned}
\hat{\sigma}_T(\gamma^* g \to q\bar{q}) &= \frac{\pi e_q^2 \alpha\alpha_s}{(\hat{s}+Q^2)} \left[(z^2 + (1-z)^2)\log\frac{Q^2(1-z)}{p_{T,min}^2 z} - 2(1-2z(1-z))\beta\right] &(1.27)\\
\hat{\sigma}_L(\gamma^* g \to q\bar{q}) &= \frac{4\pi e_q \alpha\alpha_s}{(\hat{s}+Q^2)}[z(1-z)\beta] &(1.28)
\end{aligned}$$

where we have introduced the variables

$$\begin{aligned}
z &= \frac{x}{\xi} = \frac{Q^2}{\hat{s}+Q^2}, &(1.29)\\
\beta &= \sqrt{1 - \frac{4p_{T,min}^2}{\hat{s}}}. &(1.30)
\end{aligned}$$

It must be noticed that the infrared singularity does not appear when the calculation is performed for heavy quarks since the non zero quark mass already implies a minimum p_T for the outgoing parton.

We are now ready to calculate the contribution to the structure function F_2 derived from the BGF process with massless quarks. The last step consists in transforming the partonic cross sections Eqs. 1.27 and 1.28, calculated in the $\gamma^* g$ frame, to $\gamma^* P$ cross sections, $\sigma_{L,T}$. The latter are obtained by convoluting the partonic cross sections with the *parton distribution functions*

$$\sigma_{L,T} = \int_{ax}^{1} d\xi\; g(\xi)\hat{\sigma}_{L,T}, \qquad a = 1 + \frac{4p_{T,min}^2}{Q^2} \tag{1.31}$$

where the function $g(\xi)$ represents the probability of finding a gluon in the proton with momentum fraction within ξ and $\xi + d\xi$. The contribution to the structure function F_2 can be

[3] A detailed discussion of the regularization schemes for this process can be found in [5].

obtained applying Eq. 1.15

$$F_2^{BGF,light} = \frac{Q^2}{4\pi^2\alpha}\int_{ax}^{1} d\xi\ g(\xi)[\hat{\sigma}_T(\gamma^* g \to q\bar{q}) + \hat{\sigma}_L(\gamma^* g \to q\bar{q})] \tag{1.32}$$

$$= \frac{\alpha_s}{4\pi}\sum_{i=u,d,s} e_i^2 \int_{ax}^{1} d\xi\ g(\xi)\ z \cdot C_g^{light}\left(\frac{x}{\xi}, Q^2, \beta\right) \tag{1.33}$$

where

$$C_g^{light}(z) = \left[(z^2 + (1-z^2))\log\frac{Q^2(1-z)}{zp_{T,min}^2} - 2\beta + 8z(1-z)\beta\right] \tag{1.34}$$

The function C_g^{light} is called *coefficient function* and represents the hard scattering process.

1.3.2 From light to heavy quarks

The calculation of the BGF contribution to the structure function F_2 with a quark mass $m_Q \neq 0$ leads to a coefficient function containing additional terms proportional to m_Q [6]

$$C_g^{heavy}(z) = \left[(2P_{qg}(z) + 4\epsilon z(1-3z) - 8\epsilon^2 z^2)\log\frac{1+\beta}{1-\beta} + (8z(1-z) - 1 - 4\epsilon z(1-z))\beta\right] \tag{1.35}$$

where $\epsilon = m_Q/Q^2$, $\beta = \sqrt{1 - 4m_Q^2/\hat{s}}$ and we have introduced the *Altarelli-Parisi splitting function* for the process $g \to q\bar{q}$

$$P_{qg} = \frac{1}{2}\left[z^2 + (1-z)^2\right]. \tag{1.36}$$

The function represents the probability that a gluon annihilates into a $q\bar{q}$ pair such that the quark has a fraction z of the gluon momentum. As discussed in the last Section, the dependence on the minimum requirement on the quark transverse momentum does not appear any more in the definition of β since the quark mass provides a natural cut-off.

In the limit $Q^2 \gg m_Q^2$ we have

$$\epsilon = m_Q^2/Q^2 \to 0 \tag{1.37}$$

$$\beta^2 = 1 - \frac{4m_Q^2}{\hat{s}} = 1 - \frac{4m_Q^2}{Q^2}\frac{z}{1-z} = 1 - 4\epsilon\frac{z}{1-z} \to 1. \tag{1.38}$$

where we have used Eq. 1.29. Thus, in the high Q^2 limit the heavy quark coefficient function Eq. 1.35 reduces to

$$C_{g,limit}^{heavy}(z) = 2P_{qg}(z)\log\frac{1}{\epsilon} + 2P_{qg}(z)\log\frac{1-z}{z} + 8z(1-z) - 1. \tag{1.39}$$

From Eq. 1.39 we conclude that as Q^2 becomes large compared to m_Q^2 the perturbative series contains terms of the form $\log(Q^2/m_Q^2)$ which could spoil the convergence of the series. This problem has lead to the introduction of several calculation schemes which resum the logarithmic terms to all orders. In Section 7.1.1 we will describe these calculation schemes and their comparison to the HERA measurement of the charm contribution to the proton structure function, F_2^c.

1.3.3 Heavy quark structure function up to $\mathcal{O}(\alpha_s^2)$

The generalized form of the heavy quark contribution to the structure function, $F_{k,Q}$ $(k = 2, L)$, can be written as

$$\begin{aligned} F_{k,Q} &= e_Q^2 \int_x^1 \frac{d\xi}{\xi} f_Q(\frac{x}{\xi}, \mu^2) C_{k,Q}(\xi, Q^2, m_Q^2, \mu^2) && (1.40) \\ &= e_Q^2 f_Q(\mu^2) \otimes C_{k,Q}(Q^2, m_Q^2, \mu^2) && (1.41) \end{aligned}$$

where f_Q is the parton distribution, $C_{k,Q}$ is the coefficient function representing the partonic scattering processes and the symbols $\otimes$ stands for the convolution

$$f \otimes g(x) = \int_x^1 \frac{d\xi}{\xi} f(x/\xi) g(\xi). \tag{1.42}$$

In the Eqs. 1.40-1.41 we have introduced an explicit dependence of the coefficient functions and of the parton density functions on the scale μ [4]. The origin of this dependence will be clarified later in this section. The calculation of the structure function is performed through the expansion of the heavy quark coefficient functions in the strong coupling constant α_s

$$C_{k,Q} = \sum_{n=0}^{\infty} a_s^n C_{k,Q}^{(n)} = C_{k,Q}^{(0)} + a_s C_{k,Q}^{(1)} + a_s^2 C_{k,Q}^{(2)} + ... \tag{1.43}$$

with $a_s \equiv \alpha_s/4\pi$. A calculation performed up to a certain term of the perturbative series Eq. 1.43 is called a *fixed order calculation.* In the previous sections we have shown the calculation of the terms appearing in the expansion of the coefficient functions up to the order $\mathcal{O}(\alpha_s)$, showing that a heavy quark contribution appears for the first time at $n = 1$. This corresponds to the lowest level of accuracy in a fixed order calculation and it is usually called the *leading order* (LO). We will now discuss a *next-to-leading order* (NLO) calculation, i.e. the one obtained by truncating the perturbative series Eq. 1.43 after the term $n = 2$.

The NLO corrections can be classified into two categories of processes [7]:

1. *The virtual photon interacts with a gluon from the proton ($\gamma^* g$).* In this case the corrections consist in the calculation of all processes containing an extra particle in the final state, called *real corrections*, and the one-loop *virtual corrections* to the BGF process. The real corrections are represented by the reaction

$$\gamma^* + g \to Q + \bar{Q} + g. \tag{1.44}$$

   where the outgoing gluon can be radiated from one of the quarks lines (see Figure 1.7-a,b,c) or from the incoming gluon (Figure 1.7-d). The virtual corrections consist in a BGF diagram in which an additional gluon is emitted by a quark or gluon line and then reabsorbed. In addition, one has to calculate the gluon self-energy graphs containing a gluon, a ghost or a quark loop[5]. Examples of Feynman diagrams for the virtual corrections are shown in Figure 1.8.

[4] We assume $\mu = \mu_R = \mu_F$ where μ_F and μ_R are the factorization and renormalization scales respectively.
[5] A detailed discussion of the one-loop corrections of non-Abelian gauge theories can be found in Ref. [8].

2. *The virtual photon interacts with a light quark (or anti-quark) in the proton ($\gamma^* q$).* At NLO appears for the first time the reaction

$$\gamma^* + q(\bar{q}) \to q(\bar{q}) + Q + \bar{Q}. \tag{1.45}$$

in which the parton from the proton is a light quark. The reaction is represented by Feynman diagrams in which the virtual photon can interact either with the heavy quark or with the light quark. The former case resembles the diagram of Figure 1.7-d, with the incoming gluon substituted by a light quark. The latter are Compton-like processes in which a gluon is radiated before or after the $\gamma^* q$ interaction. The radiated gluon, subsequently splits into a pair of heavy quarks.

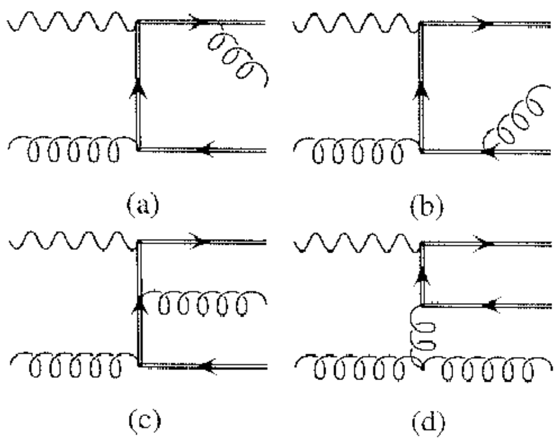

Figure 1.7 Real corrections to the process $\gamma^* g \to Q\bar{Q}$.

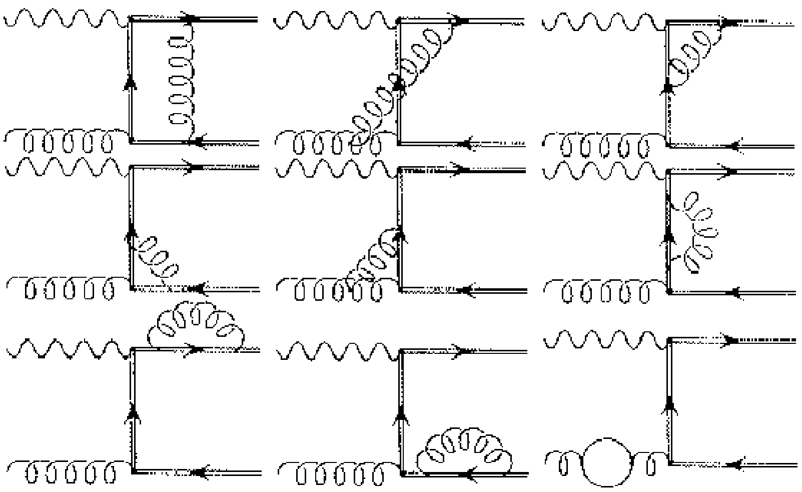

Figure 1.8 Virtual corrections to the process $\gamma^* g \to Q\bar{Q}$.

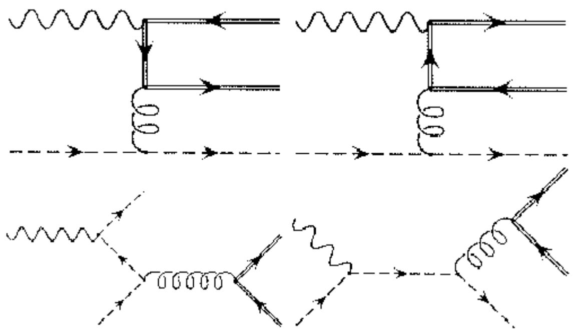

Figure 1.9 Diagrams of the process $\gamma^* g \to qQ\bar{Q}$.

The calculation of the Feynman graphs in Figure 1.7-1.8 encounter several kinds of singularities which have to be regularized and subtracted in order to obtain a finite result. The singularities can be *infrared* (IR), *ultraviolet* (UV) or *collinear* (C). The infrared divergences arise from the radiation of soft gluons. The gluon bremsstrahlung diagrams are divergent in the limit where the energy of the radiated gluon tends to zero. These singularities are canceled when we include the virtual corrections to the BGF process. The latter are also singular when the gluon energy goes to zero, due to the $k \to 0$ end of the loop-momentum integrals.

The diagrams in Figure 1.8 involve an integration over the undetermined loop momentum, and in each case the integral is divergent in the $k \to \infty$ or *ultraviolet region*. The UV divergences, regularized by N-dimensional regularization[6], are removed by mass and coupling constant renormalization. The mass renormalization removes the UV divergence by replacing the bare mass m_Q by the renormalized mass $\bar{m}_Q$ via

$$\bar{m}_Q = m_Q \left[1 + \alpha_s \delta_0 \frac{2}{\epsilon} + ... \right] \tag{1.46}$$

where the UV pole is indicated by $1/\epsilon$ and $\epsilon = N - 4$.

The origin of the collinear divergences is explained by the diagram in Figure 1.7-d. The diagram contains the propagator

$$\frac{1}{(k_1 - k_2)^2} = \frac{1}{2\omega_1\omega_2(1 - \cos\theta)} \tag{1.47}$$

where $k_1 = (\omega_1, 0, 0, \omega_1)$ and $k_2 = (\omega_2, \omega_2 \sin\theta, 0, \omega_2\cos\theta)$ are the four-momenta of the incoming and radiated gluon respectively and θ is the scattering angle between them. The propagator Eq. 1.47 diverges for $\theta \to 0$. This kind of singularity only shows up if a diagram contains a vertex with three massless particles. It is the case of a vertex with three gluons or a gluon attached to a quark line, provided that the quark is approximately massless. If the gluon is radiated by a heavy quark, the mass prevents the denominator in Eq. 1.47 to vanish.

We now consider the unrenormalized coefficient function representing the process Eq. 1.44

[6] In the dimensional regularization scheme the divergences are regulated by considering the scattering process to occur in $N = 4 + \epsilon$ rather than 4 space-time dimensions.

and the virtual corrections to the BGF process

$$\bar{C}^{(2)}_{k,g} = \bar{a}_s^2 \left[\left(\frac{1}{\epsilon_C} + \frac{1}{2} \ln \frac{m_Q^2}{\mu^2} \right) P^{(0)}_{gg} \otimes C^{(1)}_{k,g} - \beta_0 \left(\frac{2}{\epsilon_{UV}} + \ln \frac{m_Q^2}{\mu^2} \right) C^{(1)}_{k,g} \right] + \text{finite terms.} \quad (1.48)$$

The latter expression is obtained after regularizing the collinear divergences by N-dimensional regularization. The ultraviolet and collinear poles are the terms $1/\epsilon_{UV}$ and $1/\epsilon_C$ respectively. The UV pole in Eq. 1.48 is removed by coupling constant renormalization by adding the term $\bar{a}_s C^{(1)}_{k,g}$ to the r.h.s. of Eq. 1.48 and replacing the bare coupling constant $\bar{a}_s$ by the renormalized one, represented by $a_s(\mu^2)$

$$\bar{a}_s = a_s(\mu^2) \left[1 + a_s(\mu^2) \beta_0 \frac{2}{\epsilon} + ... \right]. \quad (1.49)$$

Finally, the collinear divergences are removed by *mass factorization* [9]. The bare coefficient function as

$$\bar{C}_{k,i} \left(\frac{1}{\epsilon_C}, Q^2, m^2 \right) = \Gamma_{ji} \left(\frac{1}{\epsilon_C}, \mu^2 \right) \otimes C_{k,j} \left(Q^2, m_Q^2, \mu^2 \right) \quad (1.50)$$

where Γ_{ji} represents the function which removes all collinear divergences from the bare coefficient functions. Furthermore, we need to introduce the notion of *bare* (or unrenormalized) *parton distribution* $\bar{f}_i$. The structure function Eq. 1.41 can be rewritten as:

$$\begin{aligned} F_{k,Q} &= e_Q^2 \bar{f}_i \otimes \bar{C}_{k,i} = e_Q^2 \bar{f}_i \otimes \Gamma_{ji} \otimes C_{k,j} && (1.51) \\ &= e_Q^2 f_j \otimes C_{k,j} && (1.52) \end{aligned}$$

where $f_j(\mu^2)$ is the renormalized parton density, defined as

$$f_j(\mu^2) = \Gamma_{ji} \left(\frac{1}{\epsilon_C}, \mu^2 \right) \otimes \bar{f}_i. \quad (1.53)$$

Therefore, we have absorbed the collinear pole by redefining the parton distribution function, which now depend on the factorization scale μ [7]

In the case of the collinear divergency presented in Eq. 1.47, the correspondent pole is removed by adding $C^{(1)}_{k,g}$ to $\bar{C}^{(2)}_{k,g}$ and choosing the Γ function as

$$\Gamma_{gg} \left(\frac{1}{\epsilon_C}, \mu^2 \right) = \delta(1-z) + a_s n_c \left[\frac{1}{\epsilon_C} P^{(0)}_{gg} \right] \quad (1.54)$$

where n_c is the number of colors.

1.4 The NLO program HVQDIS

To allow the evaluation of the NLO QCD prediction for heavy flavour electroproduction the virtual-photon-proton cross sections of [10, 11] have been recalculated by Harris and Smith in an exclusive fashion [12]. In other words, in the calculation one does not integrate over

[7] In our discussion we have chosen the mass factorization scale to be equal to the renormalization scale μ for simplicity.

the kinematic phase space of the outgoing particles. This enables us to study the single and double differential distributions and correlations among all outgoing particles in the reaction

$$\gamma^* + P \to Q + \bar{Q} + X \tag{1.55}$$

where $X = 0$ or 1 particle and to easily incorporate experimental cuts. The computation of the matrix elements is carried out using the *subtraction method* which is based on the replacement of divergent terms (as the soft and collinear singularities discussed is Section 1.3.3) in the squared matrix elements by *generalized plus distributions*. The essence of the method is nicely described and compared to the *phase-space slicing method* in the introduction of the paper by Kunszt and Soper [13]. The calculation was implemented in the Fortran program `HVQDIS`.

In order to compare the NLO QCD predictions with the measurement, the partonic distribution produced by the `HVQDIS` program (e.g. relative to the produced heavy quark anti-quark pair and eventually to the radiated gluon) must be transformed into the distribution of the heavy meson and of the semi-leptonic decay muon. This is achieved through a three step calculation which is described in details in Section 6.1. In what follows, we give an example of how the plus distributions are used in the calculation, following the explanation given in [12].

Let us consider the collinear singularity contained in the first diagram in Figure 1.9. This gives a contribution to the cross section

$$d\sigma(\gamma^* q \to qQ\bar{Q}) \propto \frac{d\Gamma_3}{(k_1 - k_2)^2} \propto \frac{(1-y^2)^{\epsilon/2} dy}{(1+y)} = (1-y)^{\epsilon/2}(1+y)^{-1+\epsilon/2} \tag{1.56}$$

where $d\Gamma_3$ is the three body phase space in $N = 4 + \epsilon$ space-time dimensions, k_1 and k_2 are respectively the momenta of the incoming and scattered light quark and y is the cosine of the angle between $\vec{q}$ and $\vec{k}_2$ in the $\gamma^* q$ center of mass frame. The calculation of $d\sigma$ contains the integration of the term $(1+y)^{-1+\epsilon/2}$ over the range $-1 < y < 1$, which diverges for $y = -1$. The factor $(1+y)^{-1+\epsilon/2}$ is replaced using the relation

$$(1+y)^{-1+\epsilon} \sim \left(\frac{1}{1+y}\right)_\omega + \delta(1+y)\left(\frac{1}{\epsilon} + \ln\omega\right) + \mathcal{O}(\epsilon) \tag{1.57}$$

where we have introduced the plus distribution

$$\int_{-1}^{-1+\omega} dy\ f(y)\left(\frac{1}{1+y}\right)_\omega = \int_{-1}^{-1+\omega} dy\ \frac{f(y) - f(-1)}{1+y}. \tag{1.58}$$

and $0 < \omega \leq 2$. We can thus write the cross section 1.56 as

$$\begin{aligned} d\sigma(\gamma^* q \to qQ\bar{Q}) \ &\propto \ \int_{-1}^{1} dy\ f(y)\left(\frac{1}{1+y}\right)_\omega & (1.59)\\ &= \int_{-1}^{-1+\omega} dy\ f(y)\left(\frac{1}{1+y}\right)_\omega + \int_{-1+\omega}^{1} dy\ f(y)\left(\frac{1}{1+y}\right)_\omega & (1.60)\\ &= \int_{-1}^{-1+\omega} dy\ \frac{f(y)-f(-1)}{1+y} + \int_{-1+\omega}^{1} dy\ \frac{f(y)}{1+y} & (1.61)\\ &= \int_{-1}^{1} dy\ \frac{f(y)}{1+y} - \int_{-1}^{-1+\omega} dy\ \frac{f(-1)}{1+y}. & (1.62)\end{aligned}$$

The divergency encountered at the integration limit $y = -1$ is canceled in the sum of the two integrals. In practice, the implementation of the integrals in Eq. 1.62 in the FORTRAN program is performed using standard Monte Carlo integration techniques and using $-1+\delta$ as a lower integration limit, where $\delta \sim 10^{-7}$. The upper limit of the second integral gives a contribution $f(y)\ln\omega$ which cancels against the second term appearing in Eq. 1.57. The first integral of Eq. 1.62 is called the *event* and has a positive definite weight. The second integral plus its cancellation term is called *counter-event* and it may have either positive or negative weight.

1.5 QCD evolution equations

The calculation of parton distributions cannot be achieved from first principles in perturbation theory. Attempts to use non-perturbative methods, for example numerical calculations based on lattice gauge theory, have been so far unable to achieve a level of precision necessary to make a meaningful comparison with the experimental measurements. However what can be calculated perturbatively is how they depend on the factorization scale μ^2. In the following section we will discuss this dependence in different phase-space regions of the *ep* scattering process.

1.5.1 DGLAP equations

At high energies ($\hat{s} \gg -\hat{t}$) we can rewrite the BGF cross section as [1]

$$\hat{\sigma}(\gamma^* g \to q\bar{q}) \propto \frac{4\pi^2 \alpha e_f^2}{\hat{s}+Q^2}\left(\frac{\alpha_s}{2\pi} P_{qg}(z)\log\frac{Q^2}{\mu^2}\right) \tag{1.63}$$

where $\mu^2 = p^2_{T,min}$ is the cut-off value introduced to regularize the collinear singularity. Including the BGF contribution to the structure function F_2 as shown in Fig. 1.2 we obtain the expression

$$\begin{aligned}\frac{F_2(x,Q^2)}{x} = \sum_{i=u,d,s} e_i^2 \int_x^1 \frac{d\xi}{\xi} q_i(\xi)\left[\delta\left(1-\frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} P_{qq}\left(\frac{x}{\xi}\right)\log\frac{Q^2}{p^2_{T,min}}\right] + \\ + g(\xi)\frac{\alpha^2}{2\pi} P_{qg}\left(\frac{x}{\xi}\right)\log\frac{Q^2}{p^2_{T,min}}.\end{aligned} \tag{1.64}$$

The collinear singularities are absorbed into the *bare* parton distribution at a *factorization scale* μ which plays a similar role as the renormalization scale. The *renormalized* distribution $q(x,\mu^2)$ is given by

$$\begin{aligned}q(x,\mu^2) =& q(x) + \frac{\alpha_2}{2\pi}\int_x^1 \frac{d\xi}{\xi} q(x)\left[P_{qq}\left(\frac{x}{\xi}\right)\log\frac{\mu^2}{p^2_{T,min}} + C_q(\frac{x}{\xi})\right] + \\ &+ \frac{\alpha_2}{2\pi}\int_x^1 \frac{d\xi}{\xi} g(x)\left[P_{qg}\left(\frac{x}{\xi}\right)\log\frac{\mu^2}{p^2_{T,min}} + C_g(\frac{x}{\xi})\right]\end{aligned} \tag{1.65}$$

The renormalized distribution, unlike the bare distribution, are finite. It is clear from Eq. 1.65 that there is no absolute prediction for $q(x,\mu^2)$. However QCD does give the μ^2 dependence. Differentiating Eq. 1.65 we have

$$\frac{\partial}{\partial\log\mu^2} q(x,\mu^2) = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left[P_{qq}\left(\frac{x}{\xi}\right) q(\xi,Q^2) + P_{qg}\left(\frac{x}{\xi}\right) g(\xi,Q^2)\right] \tag{1.66}$$

which gives the evolution of the quark distribution $q(x,\mu^2)$. In addition, there is an equation equivalent to Eq. 1.66, giving the evolution of the gluon distribution

$$\frac{\partial}{\partial \log \mu^2} g(x,\mu^2) = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[P_{gg}\left(\frac{x}{\xi}\right) g(\xi,Q^2) + P_{gq}\left(\frac{x}{\xi}\right) q(\xi,Q^2) \right]. \tag{1.67}$$

Eqs.1.66 and 1.67 are known as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [14, 15, 16]. The solutions of the DGLAP equations can be calculated either by numerical integration in steps of $\log Q^2$ or by taking the moments of the x distribution. For simplicity we show how the latter calculation is performed, using a non-singlet quark distribution (e.g. u-$\bar{u}$) [17]. In this case Eq. 1.66 is simplifies to

$$\begin{aligned} \frac{\partial}{\partial \log \mu^2} q(x,\mu^2) &= \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} P_{qq}\left(\frac{x}{\xi}\right) q(\xi,Q^2) && (1.68) \\ &= \frac{\alpha_s(Q^2)}{2\pi} \int_0^1 d\xi \int_0^1 dz\ \delta(x-\xi z) P_{qq}\ q(\xi,Q^2). && (1.69) \end{aligned}$$

In the space of moments of the x variable the Altarelli-Parisi equation factorize and the resulting differential equation can be solved. From Eq. 1.69 the n-th moment is given by

$$\begin{aligned} \int_0^1 \frac{dx}{x} x^n \frac{\partial q(x,\mu^2)}{\partial \log \mu^2} &= \frac{\alpha_s(Q^2)}{2\pi} \int_0^1 d\xi \int_0^1 dz \int_0^1 \frac{dx}{x} x^n \delta(x-\xi z) P_{qq}(z) q(\xi,Q^2) && (1.70) \\ &= \frac{\alpha_s(Q^2)}{2\pi} \int_0^1 \frac{dz}{z} z^n P_{qq}(z) \int_0^1 \frac{d\xi}{\xi} \xi^n q(\xi,Q^2). && (1.71) \end{aligned}$$

In more compact form

$$\frac{\partial M_n}{\partial \log \mu^2} = \frac{\alpha_s(Q^2)}{2\pi} A_n M_n(Q^2) = \frac{1}{b \log(Q^2/\Lambda)} A_n M_n, \tag{1.72}$$

where

$$A_n \equiv \int_0^1 \frac{dz}{z} z^n P_{qq}(z), \qquad M_n \equiv \int_0^1 \frac{d\xi}{\xi} \xi^n q(\xi,Q^2). \tag{1.73}$$

The solution of Eq. 1.72 is of the form

$$\begin{aligned} M_n(Q^2) &= c_n \exp\left[\frac{A_n}{b} \log\left(\log \frac{Q^2}{\Lambda^2}\right)\right] && (1.74) \\ &= c_n \sum_r \frac{1}{r!} \left(\frac{A_n}{b} \log\left(\log \frac{Q^2}{\Lambda^2}\right)\right)^r, && (1.75) \end{aligned}$$

where the c_n is determined by experiment. It can be shown that, in the proper gauge[8], the r-th term of Eq. 1.75 is equivalent to a *ladder diagram* with *r-rungs*. The result was shown by Dokshitzer in Ref. [16]. We call a ladder diagram the Feynman graph in which a quark (or a gluon) from the proton interacts with a virtual photon after the emission of one or more gluons (see Figure 1.10). In this graph, each emitted gluon represents a *rung* in the ladder. The

[8]The description is valid in an axial gauge, in which the gluon has only two physical polarization states. In such a gauge we do not need unphysical *ghost* contributions which in general would be required to cancel the scalar polarization component of the gluon.

leading $\log(Q^2)$ terms (LL(Q^2)) come from the region of phase space that is *strongly ordered* in the transverse momentum of the emitted gluon, k_T^2

$$Q^2 \gg k_{T,n}^2 \gg ... \gg k_{T,1}^2. \tag{1.76}$$

In the integration over the transverse momenta each emitted gluon corresponds to a nested integral

$$\int^{Q^2} \frac{dk_{T,n}^2}{k_{T,n}^2} \frac{\alpha(k_{T,n}^2)}{2\pi} \int^{k_{T,n}^2} \frac{dk_{T,n-1}^2}{k_{T,n-1}^2} \frac{\alpha(k_{T,n-1}^2)}{2\pi} ... \int^{k_{T,2}^2} \frac{dk_{T,1}^2}{k_{T,1}^2} \frac{\alpha(k_{T,1}^2)}{2\pi}. \tag{1.77}$$

Each integral corresponds to a $\log(\log(Q^2/\Lambda^2))$ term, which gives the power behaviour of Eq. 1.75.

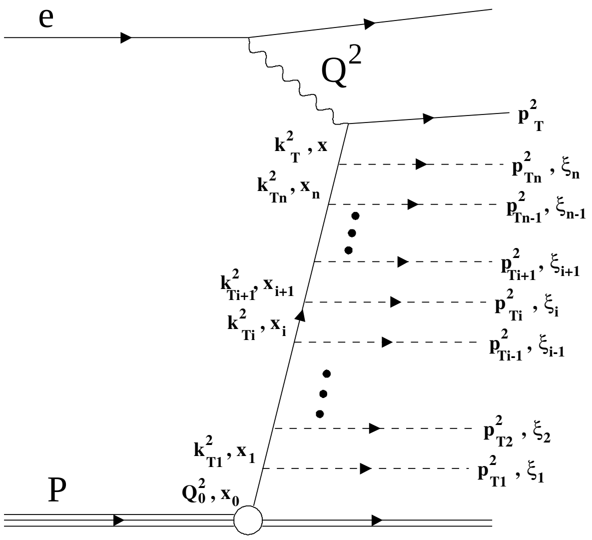

Figure 1.10 Ladder diagram with n-rungs. A quark from the proton interacts with a virtual photon after the emission of n gluons (*dashed lines*). Each quark propagator in the ladder is characterized by a longitudinal momentum fraction x_i and a transverse momentum fraction $k_{T,i}$. The strong ordered described by the DGLAP evolution equation corresponds to the relations $Q^2 \gg k_{T,n}^2 \gg ... \gg k_{T,1}^2$.

1.5.2 The small-x limit: BFKL equations and k_T factorization

We now turn our attention to the deep inelastic scattering at high center-of-mass energies of the virtual photon-proton system, W. In this context, high energy scattering means $W^2 \gg Q^2$ and it is generally referred to as small-x scattering because Bjorken-x is $Q^2/W^2 \ll 1$. To understand the type of ladder that dominates at small x we need to look at the splitting functions. As discussed above, the quark ladder with gluon rungs involves iteration of the P_{qq} splitting function[9]

$$P_{qq} = \frac{C_F}{2\pi}\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right] \tag{1.80}$$

[9]The splitting functions contain the distribution $1/(1-x)_+$ which agrees with the function $1/(1-x)$ for all values of x less than 1 and has a singularity at $x = 1$ such that the integral of this distribution with any

while a gluon ladder with gluon rungs involves the P_{gg} splitting function

$$P_{gg} = \frac{C_A}{\pi}\left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right] + \delta(1-z)\beta_0. \tag{1.81}$$

where $C_F = 4/3$, $C_A = 3$ and $\beta_0 = (11/6)C_A - (1/3)n_f$, with n_f the number of quark flavours.

At small z, P_{qq} is constant while P_{gg} grows as $1/z$. So at small x, gluon ladders with repeated iterations of $P_{gg}(z \ll 1)$ dominate, *i.e.* we have strong ordering in z [18].

At HERA Q^2 is not so large at very small x and it convenient to consider also the case in when the ends of the chain have similar transverse momenta, $k_{T,1}^2 \simeq Q^2$. In this regime we must resum the leading logarithm of x

$$\left(\alpha_s \log\frac{1}{x}\right)^n, \tag{1.82}$$

keeping the full Q^2 dependence. The DGLAP evolution does not include all the leading terms in this limit: it neglects those terms which contain the leading power of $\log(1/x)$ which are not accompanied by the leading power of $\log(Q^2)$ [19]. Therefore at moderate Q^2 and small x the DGLAP evolution should break down. To some surprise, the DGLAP evolution has been quite successful in describing the strong rise of the cross section with decreasing x. For some non-inclusive observables there are, however, clear discrepancies as recently summarized in Refs. [20, 21].

The resummation of terms proportional to $\alpha_s \log(1/x)$ to all orders, retaining the full Q^2 dependence is accomplished by the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation [22, 23, 24]. Since we must relax the strong ordering of the k_T's which gave the LL(Q^2) behaviour, we introduce the *unintegrated* gluon distribution, $f(x, k_T^2)$, in which we unfold the k_T^2 integration of the gluon distribution $g(x, Q^2)$ that we have been using so far

$$xg(x, Q^2) = \int^{Q^2} \frac{dk_T^2}{k_T^2} f(x, k_T^2). \tag{1.83}$$

In analogy to the DGLAP evolution (Eq. 1.67), we write the differential form of the BFKL equation[10]

$$\frac{\partial f(x, k_T^2)}{\partial \log(1/x)} = \int dk_T'^2\ K(k, k_T') f(x', k_T^2) \tag{1.84}$$

where $K(k, k_T')$ is the BFKL (or Lipatov) kernel

$$K(k_T, k_T') = \frac{3\alpha_s}{\pi} k_T^2 \left[\frac{1}{k_T'^2 |k_T^2 - k_T'^2|} - \beta(k_T^2)\delta(k_T^2 - k_T'^2)\right] \tag{1.85}$$

smooth function $f(x)$ gives

$$\int_0^1 dx\ \frac{f(x)}{(1-x)_+} = \int_0^1 dx\ \frac{f(x) - f(1)}{(1-x)}. \tag{1.78}$$

The factor 3/2 multiplied by the δ-function is obtained by normalizing the parton distribution to 1 and calculating the integral

$$\int_0^1 dx\ \frac{1+x^2}{(1-x)_+} = \int_0^1 dx \frac{x^2-1}{1-x} = -\frac{3}{2}. \tag{1.79}$$

A derivation of the Altarelli-Parisi splitting functions at the lowest order can be found in Ref. [8].

[10] See, for example, the textbook introduction to the dipole formulation of the BFKL equation given in [19].

with

$$\beta(k_T^2) = \int \frac{dk_T'^2}{k_T'^2} \left[\frac{1}{|k_T^2 - k_T'^2|} - \frac{1}{(4k_T'^4 - k_T^4)^{1/2}} \right]. \tag{1.86}$$

Eq. 1.84 relates the n-rung contribution, f_n, to the $(n-1)$-rung contribution, f_{n-1}, of a gluon ladder. For fixed α_s the BFKL equation can be solved analytically. The solutions show that, since there is no ordering in k_T, there will be a 'random walk' in k_T as we proceed along the chain. In addition, the evolution to smaller x is accompanied by a diffusion in k_T. The diffusion pattern is a Gaussian distribution in $\log(k_T^2)$ with a width which grows as $(\log(1/x))^{1/2}$ as x decreases.

In the framework of the unintegrated gluon distribution, predictions for the measured cross sections are calculated using the k_T-factorization theorem [25]. Cross sections are factorized into an off-shell (k_T dependent) partonic cross section $\hat{\sigma}(x/\xi, k_T^2)$ and a k_T-unintegrated parton distribution

$$\sigma = \int \frac{d\xi}{\xi} d^2k_T \; \hat{\sigma}\left(\frac{x}{\xi}, k_T^2\right) f(\xi, k_T^2) \tag{1.87}$$

In the k_T-factorization the partons entering the hard scattering matrix element are free to be off-shell, in contrast to the collinear approach which treat all incoming partons as massless (see Figure 1.11). This has some additional advantages. In the conventional approach a $2 \to 3$ process with an additional gluon in the final state (like $\gamma^* g \to gQ\bar{Q}$), in which the transverse momentum of the gluon is of the order of that of the quarks, requires the calculation of the full $\mathcal{O}(\alpha_s^2)$ matrix element. Using the k_T-factorization such processes are naturally included to the leading logarithmic accuracy, since the k_t of the incoming gluon is only restricted by kinematics and therefore can acquire a virtuality similar to the ones in a complete fixed order calculation [20].

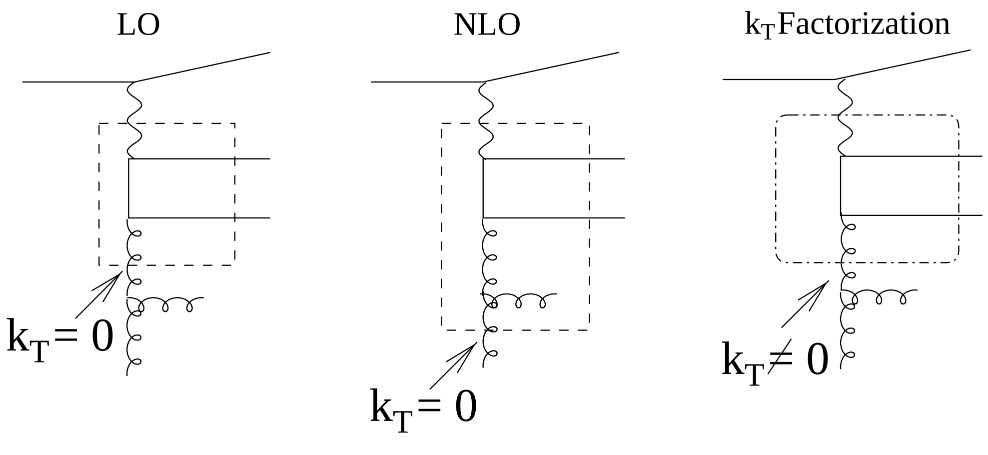

Figure 1.11 BGF diagrams at the leading (*left*), next-to-leading order (*center*) and in the k_T-factorization (*right*). In the latter case, the gluon entering the hard process (*dot-dashed box*) is free to be off-shell.

1.5.3 CCFM equation

Starting from the late eighties a group of Italian theorists developed a new QCD evolution equation, which is now known as the Catani-Ciafaloni-Fiorani-Marchesini (CCFM) equation [26, 27, 28, 29]. As a consequence of their work we now see that both DGLAP and BFKL evolution are essentially equivalent to the ordered evolution in the angles of the emitted partons. In the DGLAP collinear approximation the angle increases due to the growth of k_T, while in BFKL the angle ($\theta \simeq k_T/k_L$) grows due to the decreasing longitudinal momentum fraction as we proceed along the emission chain from the proton. The angular ordering is a consequence of the interference of soft gluon radiation (or *coherence*), which strongly effects the hard parton scattering close to the kinematical boundaries $x \to 0$ and 1. Here we will limit our discussion to the main features of the evolution equation. As for the other cases, we write the differential form of the CCFM equation [29]

$$\bar{q}^2 \frac{\partial}{\partial \bar{q}^2} \frac{x\mathcal{A}(x, k_T, \bar{q})}{\Delta_s(q, \bar{Q}_0)} = \int dz \frac{d\phi}{2\pi} \frac{\tilde{P}(z, (\bar{q}/z)^2, k_T)}{\Delta_s(\bar{q}, Q_0)} x' \mathcal{A}(x', k'_T, \bar{q/}z), \tag{1.88}$$

where $\mathcal{A}(x, k_T^2, \bar{q}^2)$ is the gluon density unintegrated in k_T. It depends also on x and on the evolution variable $\bar{q}^2$. The latter is given by

$$\bar{q} = \frac{p_{T,i}}{1 - z_i} \tag{1.89}$$

where $p_{T,i}$ is the transverse momentum of the emitted gluon i while $z_i = x/x'$ is the ratio of the energy fractions in the branching $(i-1) \to i$. The angular ordering is equivalent to the relation

$$z_{i-1}\bar{q}_{i-1} < \bar{q}_i. \tag{1.90}$$

The function $\tilde{P}$ is the gluon splitting function and for the branching i it is given by

$$\tilde{P}_g(z_i, q_i^2, k_{T,i}^2) = \frac{\bar{\alpha}_s(q_i^2(1-z_i)^2)}{1 - z_i} + \frac{\bar{\alpha}_s(k_{T,i}^2)}{z_i} \Delta_{ns}(z_i, q_i^2, k_{T,i}^2) \tag{1.91}$$

where $\bar{\alpha}_s = 3\alpha_s/\pi$ and Δ_{ns} is the *non-Sudakov* form factor. The latter appears since the splitting functions (as well as the coefficient function) are no longer evaluated in fixed order perturbation theory but supplemented with all-loop resummation of the $\alpha_s \log(1/x)$ contribution at small x. This all-loop resummation shows up in the Regge form factor for BFKL or in the non-Sudakov form factor Δ_{ns} for CCFM [20]. Both sides of Eq. 1.88 are divided by the Sudakov form factor, Δ_s, given by

$$\Delta_s(\bar{q}^2, Q^2) = \exp\left(- \int_{Q_0^2}^{\bar{q}^2} \frac{dq^2}{q^2} \int_0^{1-Q_0/q} dz \ \frac{\bar{\alpha}_s(q^2(1-z)^2)}{1-z} \right). \tag{1.92}$$

The meaning of the Sudakov form factor is clarified in Section 1.6.1 From the CCFM equation both the DGLAP and BFKL limits can be obtained when x (and therefore z) is not singular. For instance, the DGLAP equation is obtained by taking the ordering variable $\bar{q} = Q$, neglecting the non-Sudakov form factor which is not singular for finite z and replacing the integrand $\bar{q}/z$ with $\bar{q}$ [11] [29].

[11] Clearly the resulting equation must be also integrated over the transverse momenta k_T to obtain the usual DGLAP equations. In this procedure we have neglected the finite terms in z in the splitting function Eq. 1.67.

1.6 QCD evolution and event generators

The description of hadronic final states in high energy collisions requires the calculation of multi-parton emission in QCD. Since the actual implementations of fixed order pQCD calculations beyond the LO are not suited for fast event generation, higher order corrections are partially simulated using phenomenological models. The multi-partonic final states is generally pictured as a combination of particles with high transverse momenta, generated in a hard subprocess, and softer partons radiated before or after the interaction. In Monte Carlo event generators the hard scattering is represented by the $\mathcal{O}(\alpha_s)$ matrix elements discussed in Section 1.3. On the other hand, QCD evolution equations provide a theoretical framework for the description of soft parton radiation, and their implementation leads to the *Parton Shower* (PS) approach. After an introduction to the PS we will discuss an alternative phenomenological approach to multi-parton production, the *Colour Dipole Model* (CDM).

1.6.1 The parton shower

The *parton shower* (PS) approach realizes the evolution equation as a sequence of particle branchings $a \to bc$. The PS algorithms are divided into initial (or *space-like*) showers, where the branchings occur before the photon vertex, and final state (or *time-like*) showers, where they occur after the photon vertex. This separation neglects the interference terms and, therefore, is not gauge invariant. In the DGLAP evolution a and b are on the main chain of increasing space-like, while $m_c^2 \geq 0$. The sequence of branching is typically realized in a 'backward' evolution scheme where the space-like parton shower is reconstructed from the hard interaction backwards, i.e. in falling Q^2 sequence. The DGLAP equation express the branching probability in a concise form. For instance, Eq. 1.67 and 1.66 express that during a small increase $d\log Q^2$ there is a probability for a parton a with momentum fraction ξ to be resolved into a parton b at $x = z\xi$ and a parton c at $\xi - x = (1-z)\xi$. Summing up the cumulative effect of many small changes $d\log Q^2$, the probability for no radiation exponentiates. The Sudakov form factor is nothing else but this sum, giving the probability that a parton b remains at x from $(logQ^2)_{max}$ to $logQ^2 < (logQ^2)_{max}$. A knowledge of the Sudakov factor is enough to trace the evolution backwards. The sequence of branchings is repeated until a virtuality is chosen below some small cutoff scale Q_0^2, which is typically taken to be 1 GeV^2 [30].

Recently a significant progress has been made to implement the CCFM evolution equations and efficiently generate unweighted events using the backward evolution. This has been successfully achieved in the Monte Carlo program `CASCADE` [31]. As discussed, the CCFM evolution is based on the k_T-factorization which requires matrix elements with off-shell incoming particles. Off-shell matrix elements of heavy flavour lepto and hadroproduction process have been calculated in [25]. As for the DGLAP case, Eq. 1.88 gives the non-normalized probability that during a small decrease of $\bar{q}$, a t-channel gluon k' with momentum fraction x' becomes resolved into a t-channel gluon k with momentum fraction $x = zx'$ and an emitted gluon q_i. From this expression we obtain the probability for no radiation in the angular ordered region between $\bar{q}$ and q

$$\mathcal{P}_{no\ rad}(\bar{q}, q) = \exp\left(-\int_q^{\bar{q}} \frac{dq'^2}{q'^2} \int dz\, \frac{d\phi}{2\pi} \tilde{P}(z, q'/z, k_T) \frac{x'\mathcal{A}(x', k'_T, q'/z)}{\mathcal{A}(x, k_T, q')}\right). \qquad (1.93)$$

In practice the numerical calculation of the integrals in the effective form factor Eq. 1.93 would be too time consuming, since it involves the non-Sudakov form factor in the splitting function and the ratio of the gluon distribution. The solution comes with the application of the *veto algorithm*, as explained in Ref. [31].

The unintegrated gluon density $x\mathcal{A}(x, k_T^2, \bar{q})$ used in `CASCADE` was obtained from a forward evolution procedure as implemented in the program `SMALLX` [32, 33]. The gluon density was extracted through a fit to the proton structure function F_2 measured by the H1 Collaboration, as described in [31]. The fit was performed in the range $x < 10^{-2}$ and $Q^2 > 5$ GeV^2. Figure 1.12 shows a comparison of the structure function F_2 as obtained from `CASCADE` with the H1 data from [34]. Reasonable agreement with the data is observed at small x and Q^2 values. Deviations from the data are seen at larger x and Q^2 were contributions from quark initiated processes can be expected. These processes are not yet implemented in `CASCADE`.

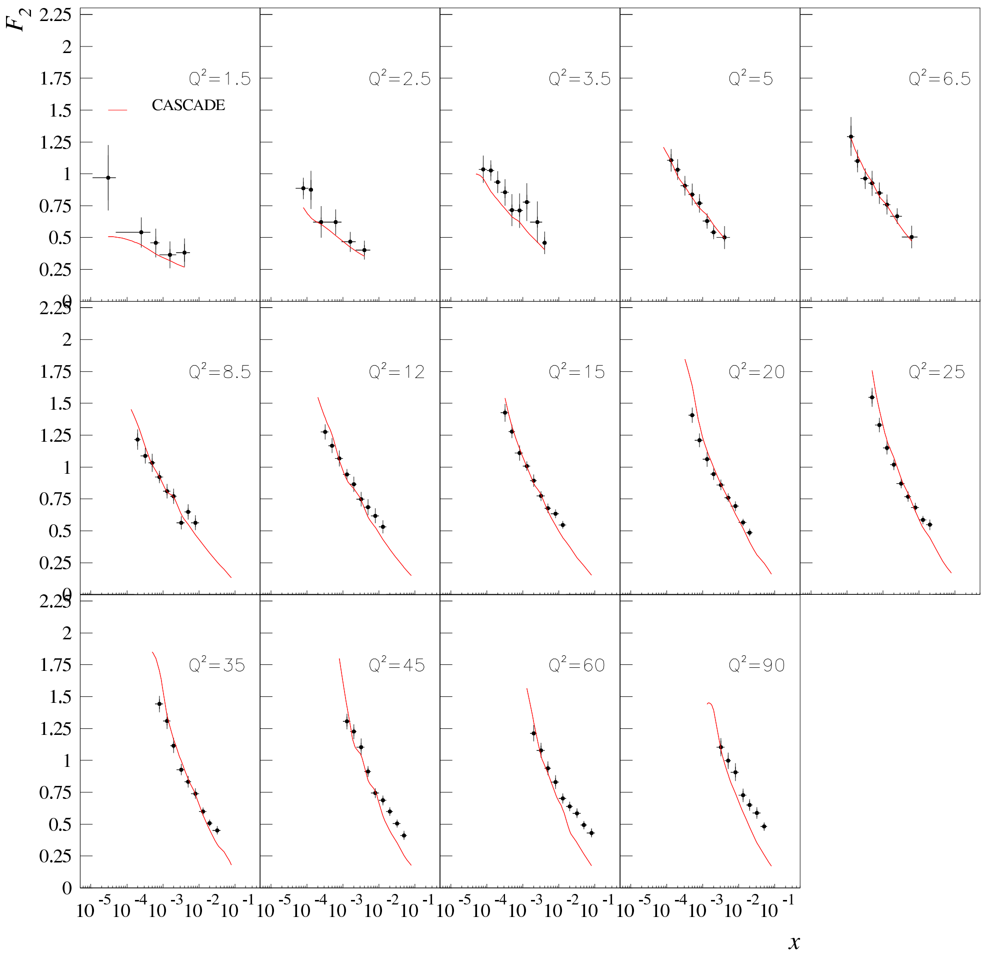

Figure 1.12 Comparison of the proton structure function F_2 measured by the H1 Collaboration (*full points*) with the `CASCADE` prediction (*line*) for different values of Q^2.

1.6.2 The colour dipole model

The CDM does not divide the parton radiation process into an initial and final state as for the PS. It is assumed that the QCD radiation can be described by the colour dipole formed between the struck quark and the proton remnant. In the process $e^+e^- \to q\bar{q}$ both the q and the $\bar{q}$ are treated as point-like particles, but in DIS only the final state quarks can be considered point-like while the proton remnant is an extended object. The CDM model, therefore, only a fraction $(\mu/p_T)^\alpha$ of the proton remnant takes part in the the emission of a gluon with transverse momentum p_T, where μ represents the size and α the dimension of the proton remnant.

1.7 Jet clustering in the Breit Frame

Quarks and gluons are not directly observable through an experiment since the confinement of colour charge recombines them into colourless hadrons. Experiments at high energy colliders have shown that quark and gluons appear as collimated flows of hadrons, generally called *jets*, which can be measured by particle detectors and that are correlated to the original partons emerging from the interaction. In order to allow a quantitative comparison between theoretical predictions and experimental results the particle jet must be a well defined observable. This is achieved through the so called *jet algorithms* which unambiguously define a jet both in the theory and in the experiment.

In deep inelastic scattering a jet algorithm should distinguish as clearly as possible between the proton remnant and the *hard jets* produced in a short-distance DIS process. This leads one to consider the properties of jet clustering algorithms with respect to *factorization* between the beam fragmentation and the hard process [35]. As discussed in Section 1.3.3, factorization means that the proton structure function is a convolution of perturbatively calculable coefficient functions and parton distributions. Also in case of cross sections of a one or multi-jet final state the factorized form Eq. 1.40 must be considered.

The coefficient function in Eq. 1.40 depend on the Bjorken variable x and the momentum fraction ξ of the parton involved in the hard subprocess only via the ratio $z = x/\xi$. As a consequence, they can be calculated at the parton level, without any explicit reference to the incoming hadron momentum. To preserve this feature, the jet algorithm should not introduce any other form of dependence on x or ξ. Therefore a jet algorithm that is well defined from the theoretical viewpoint should be based on the coefficient functions

$$C_{P,i}\left(\frac{x}{\xi}, \alpha_S(\mu^2), Q^2, \mu^2, d_{cut}\right) \tag{1.94}$$

where $P = (T, L)$ is the polarization of the virtual photon, i is the incoming parton from the proton and d_{cut} is a jet resolution scale that will be discussed later. The form 1.94 is due to the fact that the coefficient function must be Lorentz-invariant and can only depend on i, P and the incoming momenta ξp and q:

$$C_{P,i} = C_{P,i}(\xi p \cdot q^2, q^2) \tag{1.95}$$

In case of jet coefficient functions an auxiliary vector v must be introduced in order to fully define the jet clustering algorithm. Therefore, the jet coefficient function can be written as

$$C^{(n)}_{P,i} = C^{(n)}_{P,i}(\xi p \cdot q, q^2, \xi p \cdot v, q \cdot v, v^2, d_{cut}). \tag{1.96}$$

In order to preserve the y dependence in the total cross section, the auxiliary vector should be expressible as a linear combination of p and q, i.e. $v = \alpha p + \beta q$, where the coefficients α and β can be any function of x and Q^2. Substituting into Eq. 1.96 we find

$$C_{P,i}^{(n)} = C_{P,i}^{(n)} \left(\frac{Q^2}{2x/\xi}, -Q^2, \frac{\beta Q^2}{2x/\xi}, \frac{\alpha Q^2}{2x}, \beta Q^2 (\frac{\alpha}{x} - \beta), d_{cut} \right). \tag{1.97}$$

To obtain a function of the variables x/ξ, Q^2 and d_{cut} we must require $\alpha = x \cdot f(Q^2)$ and $\beta = g(Q^2)$. In other words, the jet algorithm will be factorizable if the auxiliary vector has the form

$$v = x\ f(Q^2)p + g(Q^2)q. \tag{1.98}$$

The simplest example of a suitable auxiliary vector is

$$v = 2xp + q. \tag{1.99}$$

From Eq. 1.99 it follows that $v^2 = Q^2$ and $v \cdot p = 0$, i.e. the virtual photon momentum is purely space-like in the rest frame of v, which is called the *Breit frame*.

The Breit frame is usually defined such that the auxiliary vector is zero in all components

$$0 = 2x\vec{p} + \vec{q}. \tag{1.100}$$

The z axis of the Breit frame is taken as the direction of the incoming proton. In this convention the four-momentum of the exchanged photon is

$$q = (0, 0, 0, -Q) \tag{1.101}$$

Figure 1.13 shows the QPM, QCDC and BGF processes in the Breit frame. In a QPM process the incoming quark collides head-on with the virtual boson. The quark is back-scattered along the z axis with no transverse momentum. If p and p' are the four-momenta of the incoming and scattered quark respectively we have

$$p = (Q/2, 0, 0, Q/2), \qquad p' = (Q/2, 0, 0, -Q/2). \tag{1.102}$$

On the other hand, particles with non-zero transverse momentum are produced in the Breit frame in case of QCDC and BGF processes. This property is extremely important when measuring QCD processes in deep inelastic scattering and, in our case, heavy quark production. In the Breit frame jets with non-zero transverse momentum can only be produced by strong interactions. Furthermore, according to the extrinsic production model (see Section 1.3), heavy quarks can be produced in DIS only through BGF or higher order QCD processes. It is, therefore, advisable to measure the production rate of heavy quarks in DIS counting events with hard jets reconstructed in the Breit frame rather than in the laboratory frame[12]. In the latter case the transverse momentum of the reconstructed jet is not necessarily a signature of a hard interaction. For instance, in the laboratory frame the outgoing quark jet of a QPM event can have very high transverse momenta since it balances the transverse momentum of the scattered electron.

In Section 3.4 we will discuss the algorithm used in this analysis for jet reconstruction in the Breit frame whereas in Section 4.2.4 we will describe the selection cuts applied to select events with hard jets.

[12]The measurement is clearly combined with an additional tag as the semi-leptonic decay of heavy mesons or the reconstruction of secondary vertices.

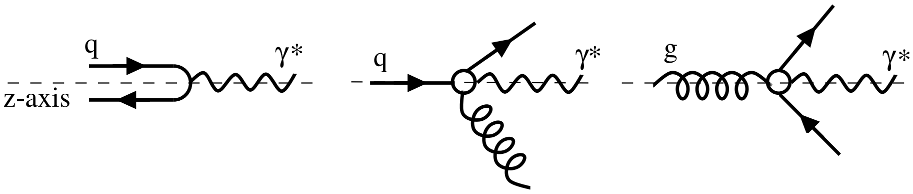

Figure 1.13 *From left to the right:* QPM, QCDC and BGF processes in the Breit frame.

1.8 Heavy quark fragmentation functions

Let us imagine an experiment observing a B meson in ep collisions. The production mechanism which leads to the observed final state is sketched in Figure 1.14-a. A bottom anti-bottom quark pair is produced in the hard process, describable by perturbative QCD. For simplicity we discuss the case in which no gluon is produced in the hard scattering. In this case, in the $\gamma^* g$ center-of-mass frame, the produced quark (b) and anti-quark ($\bar{b}$) separate with equal and opposite momentum. Once the quark pair is separated by a distance of about 1 fm, the strong coupling constant α_s becomes very large and strong color forces pull on the separating b and $\bar{b}$. The pair stretches the color lines of force until the increasing potential energy is sufficient to create one or more pairs $q\bar{q}$ of light quarks ($q\bar{q}$). The process continues until clusters of quarks and gluons are formed, each of one with zero net color and low internal momentum and, therefore, bound by very strong color forces. The strong color forces turn the bound quark pairs into hadrons. The formed hadron containing the beauty quark carries only a fraction z of the original b quark momentum. In particular, the attachment of a light anti-quark $\bar{q}$ to the heavy b quark produces a B meson.

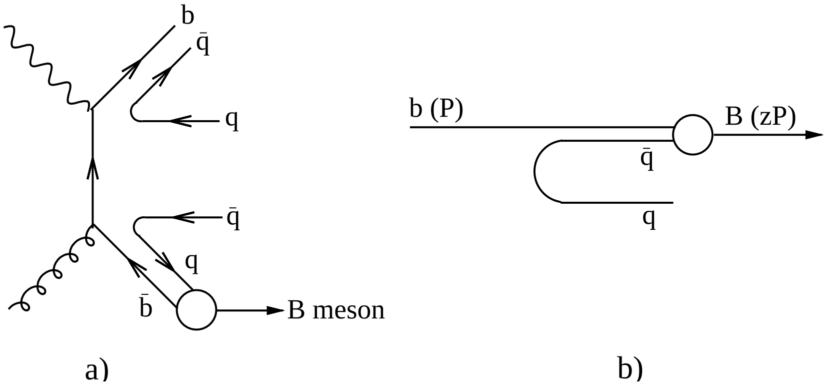

Figure 1.14 Production of b-quarks in γg collisions and their fragmentation into B-mesons (*left*). Kinematic variables of the fragmentation process (*right*): the incoming b-quark has a four-momentum P and the outgoing B-meson carries a fraction z of the original four-momentum.

The model described here can be used to quantify the amplitude for a rapidly moving b quark of momentum P to fragment into a hadron $B = b\bar{q}$ of momentum zP and a light quark

Fragmentation function	Functional form	Parameters
Kartvelishvili et al. [38]	$Nz^{\alpha_b}(1-z)$	α_b
Bowler [39]	$N\frac{1}{z^{1+bm_\perp^2}}(1-z)^a \exp(-\frac{bm_\perp^2}{z})$	a, $bm_\perp^2$
Lund symmetric [40]	$N\frac{1}{z}(1-z)^a \exp(-\frac{bm_\perp^2}{z})$	a, $bm_\perp^2$
Peterson et al. [36]	$N\frac{1}{z}(1-\frac{1}{z}-\frac{\varepsilon_b}{1-z})^{-2}$	ε_b
Collins-Spiller [41]	$N(\frac{1-z}{z}+\frac{(2-z)\varepsilon_b}{1-z})(1+z^2)(1-\frac{1}{z}-\frac{\varepsilon_b}{1-z})^{-2}$	ε_b

Table 1.1 Heavy quark fragmentation functions. In the functional forms N is a normalization constant, different for each fragmentation function.

q by calculating the energy difference before and after fragmentation (see Fig. 1.14-b) [36]:

$$\begin{aligned}\Delta E &= E_b - (E_B + E_q) &(1.103)\\ &= \sqrt{m_b^2+P^2} - \sqrt{m_B^2+z^2P^2} - \sqrt{m_q^2+(1-z)^2P^2} &(1.104)\\ &\simeq \frac{m_b^2}{2P}\left(1-\frac{1}{z}-\frac{\epsilon_b}{1-z}\right) &(1.105)\end{aligned}$$

where $\epsilon_b = m_{\bar{q}}^2/m_Q^2$ and it is assumed $m_B = m_b$ for simplicity. The matrix element can be estimated using the quantum mechanical recipe

$$|\mathcal{M}(b \to B\ q)|^2 \propto \frac{1}{\Delta E^2}. \tag{1.106}$$

The probability $D_b^B(z)$ to find a B meson in the debris of a b quark carrying a fraction z of the b quark energy follows multiplying the amplitude Eq. 1.106 by the factor $1/z$ which takes in account the longitudinal phase space:

$$D^B(z) = \frac{N}{z\left(1-\frac{1}{z}-\frac{\epsilon_b}{1-z}\right)^2} \tag{1.107}$$

where the normalization N is fixed by the sum over all hadrons containing the heavy b quark:

$$\sum_{hadrons} \int dz\ D^B(z) = 1. \tag{1.108}$$

The function Eq. 1.107 is generally known as the *Peterson fragmentation function*. It is a phenomenological way to describe a non-perturbative process, but has no fundamental meaning. Although the Peterson function is commonly used to parameterize the fragmentation process it was shown that has a wrong dependence on the QCD scale and on the heavy quark mass [37]. Other functional forms have been proposed to model the fragmentation process. In Table 1.8 we show a summary of the functional forms and the relative parameters. In the next section we will introduce a description of the fragmentation function based on the heavy quark mass expansion and discuss the application of fragmentation function to the QCD predictions of the process under study.

1.8.1 Fragmentation functions and QCD predictions

Fragmentation functions of b-quarks have been measured by several experiments in e^+e^- collisions using weakly decaying B mesons. These functions are generally parameterized in terms of the scaled energy x_b, defined as the ratio of the heavy hadron energy to the beam energy:

$$x_b \equiv \frac{E_{had}}{E_{beam}} = \frac{2E_{had}}{\sqrt{s}}, \tag{1.109}$$

where $\sqrt{s}$ is the available center-of-mass energy. The z variable discussed in the previous section is not accessible experimentally, therefore a direct reconstruction of the $D_b^B(z)$ is not possible. A collection of recent measurements of the scaled momentum fraction x_b is shown in Fig. 1.15.

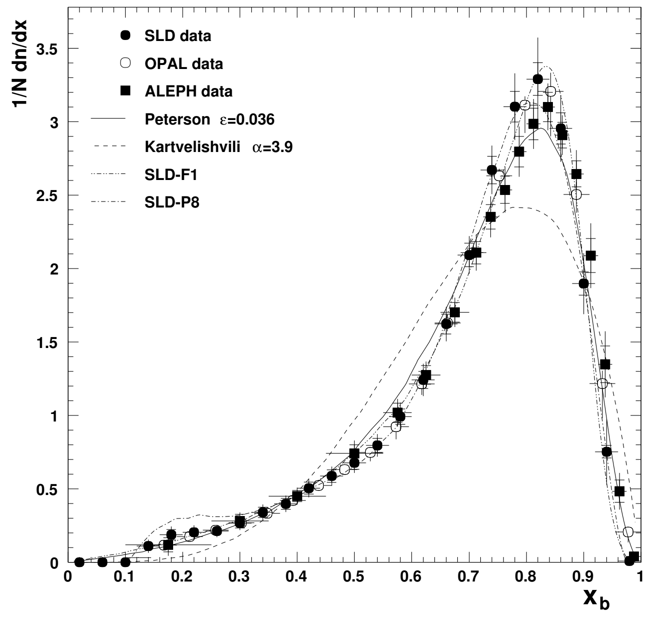

Figure 1.15 Scaled momentum distribution of the energy of a hadron containing the b quark with respect to the b quark energy. The measurements are obtained by several experiments in e^+e^- collisions.

The fragmentation functions measured in e^+e^- collisions can in principle be used to make prediction for other processes as e.g. the production of B mesons at $p\bar{p}$ and ep colliders. Once the cross section σ_b for the production of the b-quark has been calculated using pQCD, the respective B meson cross section is obtained by folding it with a non-perturbative component:

$$\sigma_B = \sigma_b \otimes D^{np}. \tag{1.110}$$

Here we should stress that the component D^{np} is not an observable quantity and it cannot be determined in absolute terms. The definition of the non-perturbative part depends on the calculation scheme used to estimate the partonic process and its parameters (Λ_{QCD}, m_Q, etc.). Consequently, the non-perturbative part should be used only together with perturbative calculations of the same kind (LO, NLO, resummed, etc.) as the one used in the fit to the e^+e^- data.

Although analytical predictions cannot be achieved for the non-perturbative part, its general features have been established using the heavy quark effective theory (HQET) [37, 42]. The main assumption underlying the HQET is that matrix elements can be expanded in powers of λ/m_Q, where λ is a typical hadronic scale ($\lambda \sim 0.4$ GeV) and m_Q is the mass of the heavy quark, and where the coefficient in the expansion are independent of m_Q. The calculation is performed in the space of Mellin moments, which turn convolutions into products. The Nth-moment of the non-perturbative part is defined as:

$$D_N^{np} \equiv \int_0^1 dx\, x^{N-1} D^{np}(x). \tag{1.111}$$

In the space of moments, the leading terms of the power series of D_N^{np} as predicted by HQET are:

$$D_N^{np} = 1 - (N-1)\frac{\lambda}{m_Q} + \mathcal{O}\left(\frac{\lambda^2}{m_Q^2}\right). \tag{1.112}$$

This result can be compared with the measured fragmentation functions once the moments are calculated. The moments of the b-quark fragmentation function measured by the ALEPH Collaboration have been calculated and fitted with a one parameter form [43]:

$$D(z,\alpha) = (\alpha+1)(\alpha+2)z^{\alpha}(1-z) \tag{1.113}$$

The form was originally proposed by Kartvelishvili et al. in Ref. [38] and it is consistent with the leading power correction if one replaces α with $2m_Q/\lambda$. To get the best suited parameter for α a fit to the moments of the fragmentation function obtained in e^+e^- by ALEPH data [44], shown in Fig. 1.16, was optimized to describe the moments around $N \sim 4$. The motivation is given by the dependence of the cross section for heavy quark production in $p\bar{p}$ and ep collisions on the transverse momentum of the b-quark. The latter can be expressed as $d\sigma_b/dp_t \propto 1/p_T^4$, therefore using the Eq. 1.110 the cross section for the B meson is:

$$\frac{d\sigma_B}{dp_T} = \frac{d\sigma_b}{dp_T} \otimes D^{np} \tag{1.114}$$

$$\propto \int dz\, d\bar{p}_T\, \frac{1}{\bar{p}_T^4} D^{np}(z)\delta(z\bar{p}_T - p_T) = \frac{D_4^{np}}{p_T^4}. \tag{1.115}$$

In addition, the fit has been performed using a calculation for the perturbative part resummed with NLL accuracy, i.e. a fixed order calculation up to the order α_s^2 in which remaining large logarithmic terms are resummed to all orders [45, 46].

Following the prescriptions discussed above, the extracted D^{np} term was combined with a NLL calculation in order to describe the production of B^+ meson in $p\bar{p}$ collisions at Tevatron. Figure 1.17 shows the differential cross section of B^+ mesons as a function of the transverse momentum measured by the CDF Collaboration. The dashed line and the band represent the prediction of a fixed order NLL calculation including the fragmentation function optimized with the fit to the ALEPH data. The measured points are generally above the QCD prediction but in agreement within the experimental and theoretical uncertainties. The QCD prediction obtained using a Peterson function to parameterize the fragmentation process with $\epsilon = 0.006$ is below the data.

In case of deep inelastic ep collisions, the cross sections for the hard process σ_b has been calculated up to the order α_s^2, as discussed in Section 1.3.3, but no additional resummation is included. It is not clear whether NLL resummation is required in the range of transverse momenta of the b-quarks produced at HERA.

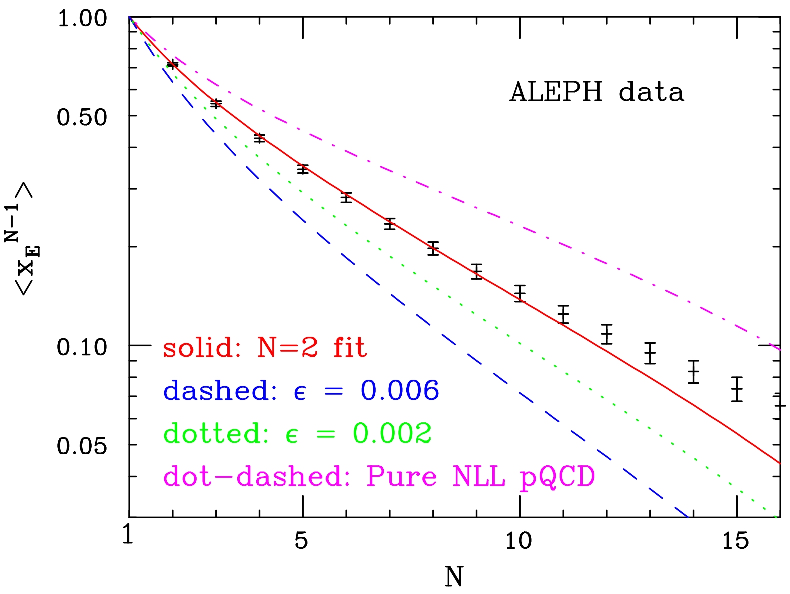

Figure 1.16 Moments of the b-quark fragmentation function measured by the ALEPH Collaboration in e^+e^- collisions. The moments are compared with a pure perturbative theoretical prediction and with pQCD folded with various models for the non-perturbative part. The solid line represents a fit to the data points using the functional form Eq. 1.113 [43].

1.9 Measurements of beauty production in ep collisions

In this section we review the experimental results obtained previously obtained at the HERA collider in both the photoproduction ($Q^2 \sim 0$ GeV^2) and the Deep Inelastic Scattering ($Q^2 > 1$ GeV^2) kinematic regimes.

The experimental procedure often relies on the measurement of the transverse momentum p_T^{rel} of the muon produced in the semi-leptonic decay with respect to the axis of the closest jet. This spectrum is harder for b quarks than for c quarks and, therefore, allows a statistical separation of the signal from the background. Additional information such as lifetime measurements provided by silicon detectors can also be used. An alternative tag is given by the coincidence of D^* mesons and muons which provide sensitivity to the region of low transverse b quark momenta and is less affected by background.

1.9.1 Open beauty in photoproduction

When the exchanged photon has small virtuality (Q^2) the time-scale of the interaction is such that its hadronic structure can be revealed. Photoproduction at HERA can therefore be similar to the processes at hadron colliders and supplies complementary information. Both ZEUS and H1 already published results on b-quark photoproduction [47, 48, 49].

The H1 Collaboration has measured the production of b-quarks in the photoproduction regime using a sample of 6.6 pb^{-1}. Events were selected by requiring at least two jets with transverse energy $E_T > 6$ GeV and a muon in the region of polar angle $35° < \theta_\mu < 130°$ and

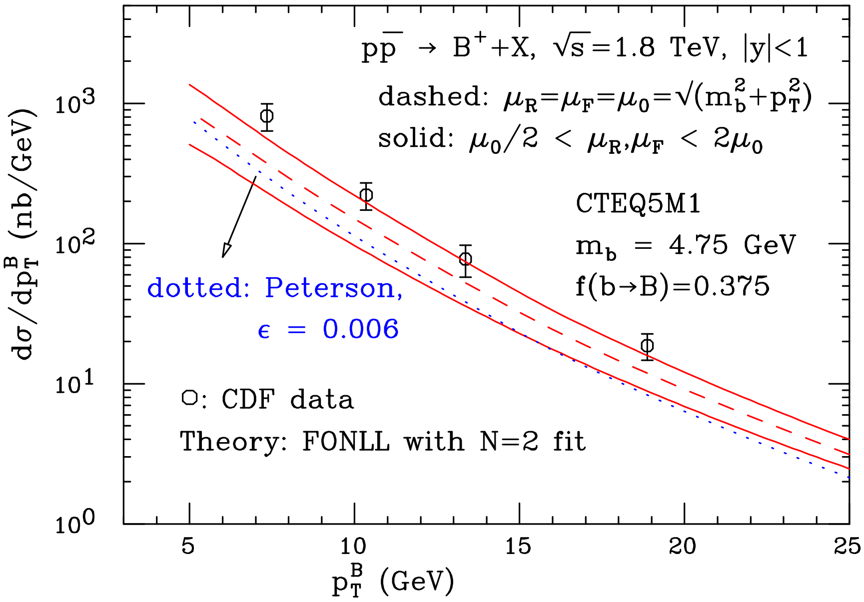

Figure 1.17 Differential cross section of B^+ meson production as a function of the transverse momentum of the B^+ meson measured by the CDF Collaboration (*open circles*). The measured points are compared with a fixed order NLL calculation including the fragmentation function extracted from a fit to the ALEPH data (*dashed line*). The band corresponds to the variation of the renormalization and factorization scale. The theoretical expectation obtained with a Peterson function and $\epsilon = 0.006$ is also shown for comparison [43].

transverse momentum $p_T^\mu > 2$ GeV. The fraction of events from b-decays was determined on a statistical basis, using the transverse momentum of the muon relative to the thrust axis of the jet which contains it. Using this fraction the cross section was measured in the kinematic range defined by $Q^2 < 1$ GeV^2, $0.1 < y < 0.8$, $p_T^\mu > 2$ GeV and $35° < \theta_\mu < 130°$

$$\sigma(ep \to b\bar{b}X \to \mu X') = 0.176 \pm 0.016\ (stat.)\ {}^{+0.026}_{0.017}\ (syst.)\ \text{nb}. \tag{1.116}$$

The QCD prediction was calculated using the calculation of Frixione et al, implemented in the program `FMNR` [50]. The b-quark mass was set to 4.75 GeV, the renormalization and factorization scales were set to the transverse mass $m_T = \sqrt{m_b^2 + p_T^2}$. The parton densities for the proton and the photon were respectively set to MRSG and GRV-HO. The Peterson function was used to describe the fragmentation of the b-quark to hadrons. The value predicted by the `FMNR` program is 0.054 ± 0.9 nb.

The ZEUS experiment has recently measured the differential cross sections of beauty photoproduction using events with at least two jets and a muon in the final state [51]. The luminosity used is almost three times larger than in the previous measurements. The fraction of events from b decays has been extracted using the p_T^{rel} method. The kinematic region is defined by $Q^2 < 1$ GeV^2, $0.2 < y < 0.8$, $p_T^{Jet1(2)} > 7(6)$ GeV, $|\eta^{Jet1(2)}| < 2.5$, $p_T^\mu > 2.5$ GeV and $-1.6 < \eta^\mu < 2.3$ [13]. Figure 1.18 shows a comparison between the measured differential

[13] $\eta = -\ln(\tan\theta/2)$ is the pseudorapidity, where θ is the polar angle measured with respect to the proton beam

cross section and a NLO QCD calculation for different regions of the muon pseudorapidity and transverse momentum. The QCD prediction was calculated using the program `FMNR`. The hadronization is modeled by a Peterson function and the spectrum of the semi-leptonic muon momentum was extracted from `PYTHIA` [52, 53]. The bands around the NLO prediction show the results obtained by varying the b-quark mass as well as the renormalization and factorization scales. The measured cross sections are a factor 1.4 larger than the central NLO prediction but compatible with it within the experimental and theoretical uncertainties.

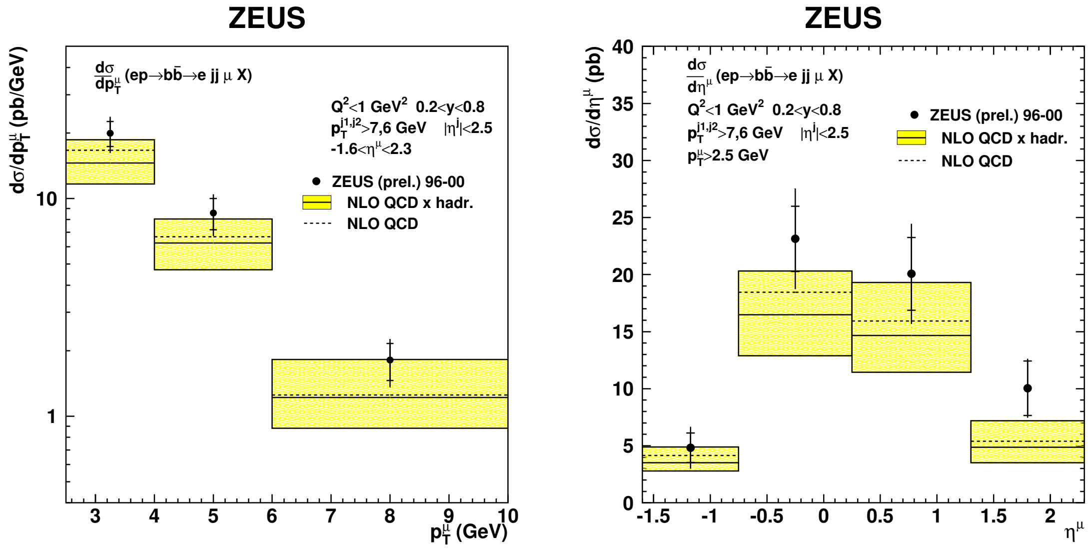

Figure 1.18 Differential cross section of beauty photoproduction as a function of the muon transverse momeuntum (*left*) and of the muon pseudorapidity (*right*) compared with NLO QCD calculations.

In addition, a dijet cross section ($ep \to b\bar{b}X \to \ jet\ jet\ X$) has been determined using `PYTHIA` to extrapolate to the unmeasured part of the muon kinematics and to correct for the branching ratio. For this measurement a different data sample has been used with looser cuts on the transverse momentum of the muon at large pseudorapidities. The result is $\sigma^{dijet} = 733 \pm 61 \pm 104$ pb while the NLO QCD prediction is 381^{+117}_{-78} pb, corresponding to a data excess of a factor of two.

1.9.2 Open beauty in DIS

Both the H1 and ZEUS Collaborations have released preliminary results on beauty production in the DIS region. In the H1 result the analysed sample corresponds to 10.5 pb^{-1} of data collected during the year 1997 [54]. Events were selected with the scattered electron detected at a polar angle $\theta_e < 177°$ relative to the direction of the incoming proton, and with energy $E'_e > 9$ GeV. In addition, two jets were required with transverse energy $E_T > 5$ GeV. The cross section was determined in the kinematic range $2 < Q^2 < 100$ GeV^2, $0.05 < y < 0.7$, $p_T^\mu > 2$ GeV^2 and $35° < \theta^\mu < 130°$. The jet kinematic range was extrapolated to the full

direction.

phase space using the AROMA Monte Carlo program. The fraction of events from b-decays was determined by a fit to the two dimensional (δ, p_T^{rel}), where δ is the impact parameter of the muon track (see Figure 1.19). The resulting fraction in the data was $f_b = 43 \pm 8\%$. From this value the cross section of beauty production in DIS was extracted

$$\sigma(ep \to bX \to \mu X) = 39 \pm 8(stat.) \pm 10\ (syst.)\ \text{pb}. \tag{1.117}$$

The NLO QCD prediction was obtained using the HVQDIS program. The Peterson function was used to describe the b-quark fragmentation. The GRV98 parton distributions were used, the quark mass was set to 4.75 GeV and the renormalization and factorization scales to $\mu = \sqrt{Q^2 + m_b^2}$. The QCD prediction is (11 ± 2) pb, where the error was estimated by varying the scale μ by factor of two, the Peterson parameter ϵ between 0.0016 and 0.0069 and m_b between 4.5 and 5 GeV.

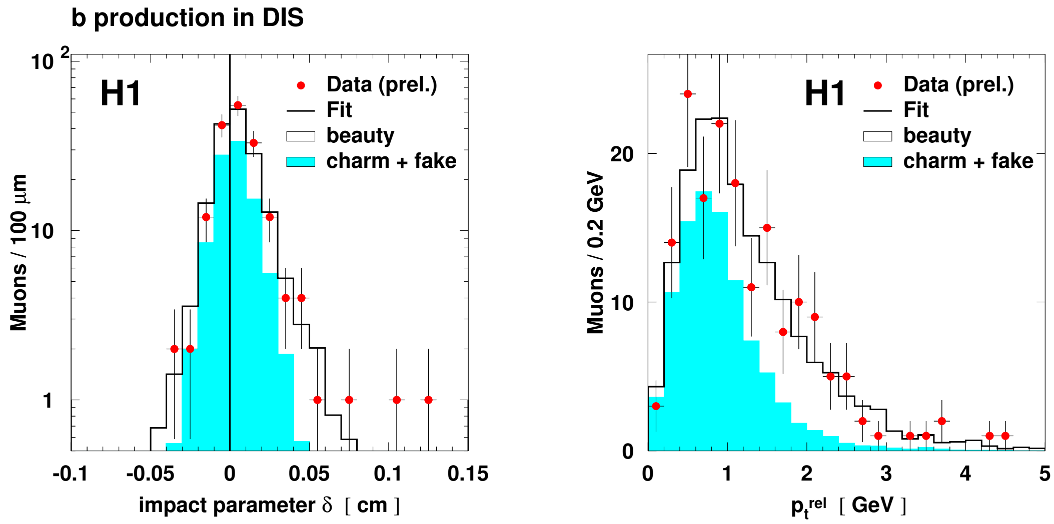

Figure 1.19 Distributions of the muon impact parameter (*left*) and transverse momentum (*right*) measured by the H1 Collaboration in the DIS regime. The fraction of events from b-decays is extracted with a likelihood fit to the data. The distributions relative to the b-decays and to the combination of charm and light quark decays are shown separately.

Thanks to the high luminosity in the new ZEUS measurement differential distributions have been measured for the first time in DIS [55]. The experimental results as well as the QCD prediction are the subject of this thesis and are discussed in details in the next chapters of this thesis.

1.9.3 $D^* - \mu$ correlations

The separation of charm and beauty contributions to the signal can also be performed by exploiting the charge and angular correlations of the D^* meson and of the muon in the reaction $ep \to eb\bar{b}X \to eD^*\mu X$. Of particular interest is the configuration in which the muon and the

D^* originate from the same parent B-meson yielding unlike charge sign $D^* - \mu$ pairs produced in the same hemisphere and constrained not to exceed the B-meson mass (Figure 1.20-a). Another important contribution arises from charm-pair production, where one charm quark fragments into a D^* and the other decays into a muon. This again yields unlike-sign D^*-muon pairs. Furthermore, production of b-quark pairs in which the D^* and muon originate from different beauty quarks can yield both like and unlike-sign D^*-muon combinations, depending on whether the muon is from the decay of the primary beauty quark (Figure 1.20-b), or from a secondary charm quark (Figure 1.20-c), and whether B^0-$\bar{B}^0$ mixing has occurred. Beauty production is the only source of genuine like-sign D^*-muon pairs.

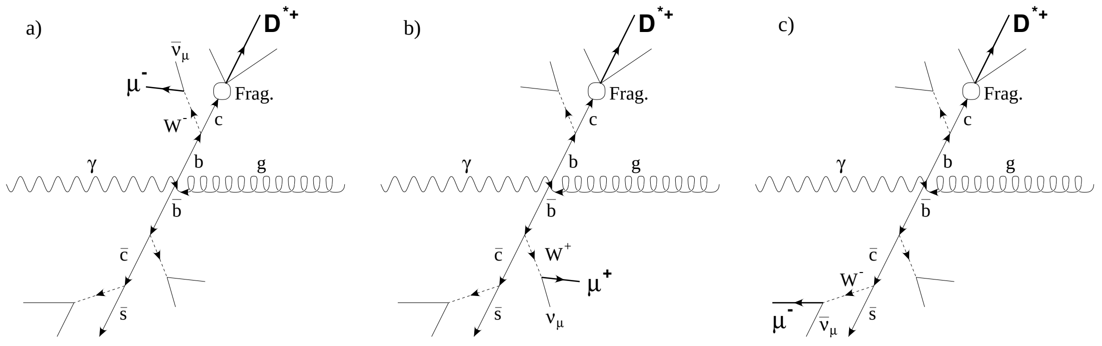

Figure 1.20 Production and decay of b-quark pairs in γg collisions. The diagrams show some of the decays which generate a D^* and a muon in the final state (*see text*).

Using this strategy and performing a likelihood fit on the kinematic distributions H1 has extracted the cross sections of beauty and charm production in the kinematic region defined by $p_T^{D^*} > 1.5$ GeV, $|\eta^{D^*}| < 1.5$, $p_T^{\mu} > 1$ GeV, $|\eta^{\mu}| < 1.74$ and $0.05 < y < 0.75$ [56]. The measured values, which confirm previous results, are respectively

$$\sigma^b_{vis} = 380 \pm 120\ (stat.)\ \pm 130\ (syst.)\ \text{pb}, \tag{1.118}$$
$$\sigma^c_{vis} = 720 \pm 115\ (stat.)\ \pm 245\ (syst.)\ \text{pb} \tag{1.119}$$

which are well above LO+parton shower Monte Carlo expectations.

A similar analysis has been conducted by ZEUS [57] whose selection has been optimized for decays of b-quarks. Figure 1.21 shows the distributions of the mass difference (D^*-D^0) for D^*-muon pairs reconstructed in the same hemisphere ($\Delta R < 2$) and in the two opposite hemispheres of the event ($\Delta R > 2$) [14]. The corresponding number of beauty, charm, and fake-muon candidates is indicated in each case. In the unlike-sign part, the beauty and charm contributions are clearly separated.

The beauty cross section, measured in the phase space defined by $p_T^{D^*} > 1.9$ GeV, $|\eta^{D^*}| < 1.5$, $p_T^{\mu} > 1.4$ GeV, $-1.75 < \eta^{\mu} < 1.3$ is

$$\sigma^b_{vis} = 214 \pm 52\ (stat.)\ ^{+96}_{-84}\ (syst.)\ \text{pb}. \tag{1.120}$$

The result is in good agreement with the H1 measurement once the same cuts are applied. In order to compare the measured cross section with NLO QCD predictions a photoproduction-enriched sample has been selected by applying the cuts $Q^2 < 1$ GeV^2 and $0.05 < y < 0.85$.

[14] $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is the distance between the D^* and the muon in the $\eta - \phi$ plane.

Furthermore, the measurement was restricted to a b-quark rapidity range $\zeta^b < 1$ where the distributions of the Monte Carlo program used to extrapolate agree with the respective FMNR spectra within ±15%. The result for the extrapolated cross section is $\sigma\gamma p \to b(\bar{b})X = 15.1 \pm 3.9^{+3.8}_{-4.7}$ nb while the NLO prediction of FMNR for this reaction is only $5.0^{+1.7}_{-1.1}$ nb.

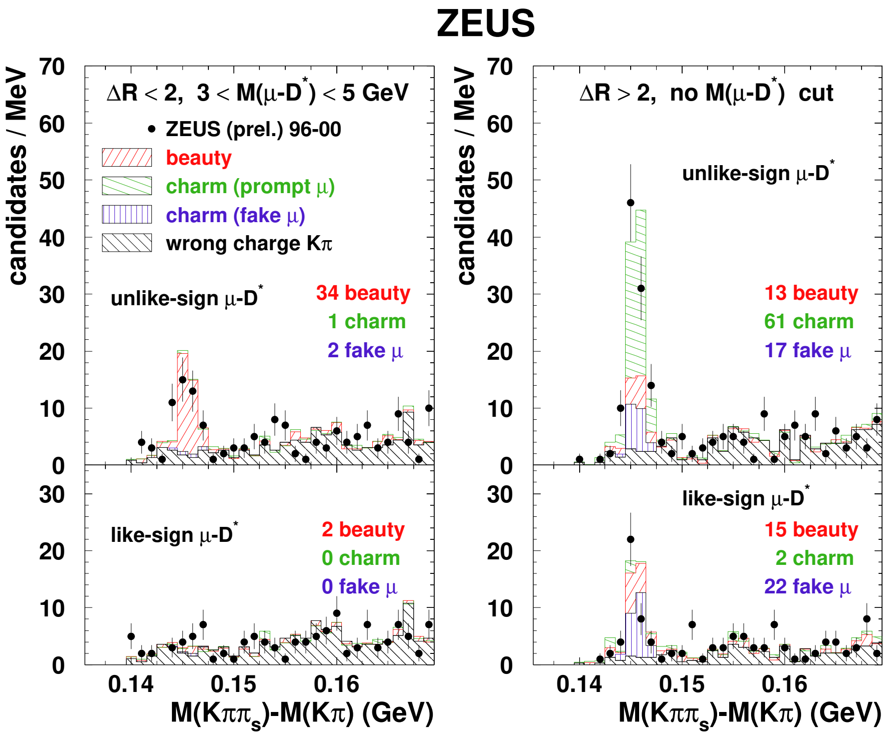

Figure 1.21 $\Delta M(D^* - D^0)$ distributions for four different regions of phase space. The data are shown as full circles. The beauty and charm contributions, and the fake muon combinatorial D^* background are indicated as shown in the legend. The number of beauty, charm and fake-muon candidates is indicated in each case.

1.9.4 Summary and perspectives

The understanding of the b-quark production mechanism is an outstanding puzzle in QCD. Recently, a set of visible cross sections of beauty production, defined close to the detector acceptance, have been measured at HERA. The results are about a factor 1.4 higher than the NLO QCD predictions but consistent within the experimental and theoretical uncertainties both in the photoproduction and the DIS regions. An excess of measured cross sections over the NLO QCD prediction is observed when attempts are made to extrapolate cross sections to regions which are not directly measured by the detectors.

Figure 1.22 shows the ratio of the b-quark production cross sections measured at HERA to the QCD NLO predictions. The plot collects the results in both the photoproduction and DIS kinematic regions. In the photoproduction region, the ZEUS preliminary result (*full square*) refers to the result obtained after the extrapolation of the muon kinematics to the full phase space (see Section 1.9.1). This is performed folding the measured visible cross section with a correction factor extracted from the PYTHIA Monte Carlo program. It must be noticed that the differential cross sections measured in the region where there is a detector are in agreement with the QCD NLO predictions within the experimental and theoretical uncertainties. The H1 results in the photoproduction regime (*full star and circle*), as well as the result in DIS (*full circle*) also rely on extrapolation factors extracted from Monte Carlo programs. In this case the jet kinematics is extrapolated to the full range of pseudorapidity and transverse energy. The ZEUS photoproduction measurement obtained studying the D^*-muon correlations (*full triangle*) has been obtained extrapolating the measured cross section to the b-quark level. This measurement is more sensitive to the region of low transverse momentum of the b-quark but it is limited by statistics.

Considering the different experimental approach and theoretical assumption included in each measured cross section it is difficult to draw a unique conclusion. It is desirable, in the future to provide measurement of visible cross sections, which do not rely on theoretical assumption, together with the more inclusive extrapolated results. This approach would facilitate the interpretation of the experimental results, i.e. disentangling whether the existing discrepancies with QCD prediction are produced by missing higher order corrections in the fixed-order calculation, by inappropriate description of the fragmentation process and so on.

The HERA collider is now starting a new phase of operation at higher luminosities. Together with the enhanced b-tagging capabilities of the new H1 and ZEUS vertex detector and tracking triggers more precise and differential measurements can be expected within the next five years. In Section 7.1.1 we will discuss some selected topics of heavy flavour production in the light of a large integrated luminosity foreseen at HERA II.

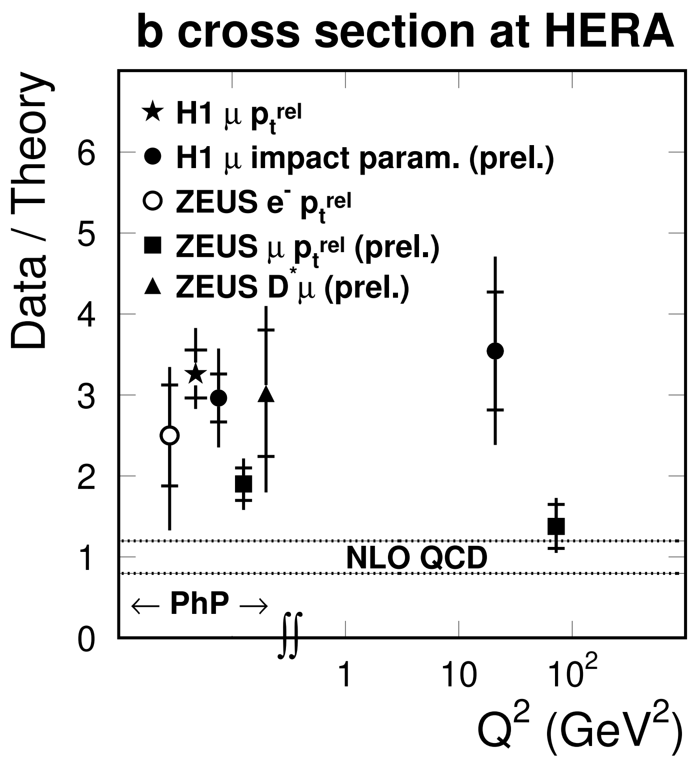

Figure 1.22 Ratio of the measured b-production cross section at HERA to NLO QCD expectations as a function of Q^2. The full squares correspond to the new ZEUS preliminary results for DIS and photoproduction with the p_T^{rel} method. The triangles represents the photoproduction measurement using the D^*-muon tag. The error bars on the data point indicate the statistical and systematic uncertainties of the measurement added in quadrature (*outer error bars*). The dotted band corresponds to the NLO prediction with a generic uncertainty of 20%.

2 The HERA collider and the ZEUS detector

In this chapter we will briefly describe the HERA collider and the ZEUS detector. The detector components used in the analysis described in this thesis are discussed in more details. A complete description of the ZEUS detector can be found in [58, 59].

2.1 The HERA collider

HERA (Hadron Electron Ring Anlage) [60] is the only lepton-proton collider in the world. It is located at the DESY (Deutsches Electronen SYnchrotron) laboratory, in Hamburg, and operates since 1992. The HERA machine collides electrons or positrons, accelerated to an energy of 27.5 GeV, with 820(920) GeV protons (the energy of the proton beam was changed at the beginning of 1998 from 820 to 920 GeV). The resulting center-of-mass energy is 300(318) GeV, more than an order of magnitude higher than at previous fixed–target experiments. As a consequence a new and wider kinematic region is accessible at HERA (see Figure 2.1).

The HERA tunnel is 6.3 km long and it is located between 15 and 30 m below the ground level. Electrons (or positrons) and protons are accelerated in two different rings. The beam line of the lepton ring consists of conventional magnets with maximum field of 0.165 T, while the proton beam is made of superconducting magnets with a maximum field of 4.65 T.

The experiments are located in four experimental halls along the HERA ring. Lepton-proton collisions occur in two interactions points, one in the North Hall where the H1 experiment is located, the other in the South Hall where the ZEUS experiment is placed. In the East Hall the HERMES experiment is located, which studies the spin structure of the nucleon using the collisions of polarized leptons on an internal polarized gas target. The HERA–B experiment, located in the West Hall, was built to use collisions of the proton beam halo with a wire target to produce B–mesons to study the CP-violation in the $B - \bar{B}$ system.

Figure 2.2 shows a layout of the HERA facility and of its pre–accelerators system. The proton acceleration chain starts with negative hydrogen ions (H^-) accelerated in a LINAC to 50 MeV. The electrons are extracted from the H^- ions to produce protons, which are injected into the proton synchrotron DESY III, accelerated up to 7.5 GeV, and then transferred to PETRA, where they are accelerated to 40 GeV. Finally they are injected into the HERA proton storage ring, where they reach the nominal beam energy of 920 GeV.

The electron (or positron) pre-acceleration chain starts in a linear accelerator, LINAC I (LINAC II), where the leptons are accelerated up to 450 MeV. The leptons are then injected into DESY-II, accelerated to 7 GeV and then transferred to PETRA II, where they reach an energy of 14 GeV. They are then injected into HERA where they reach the nominal lepton beam energy of 27.5 GeV. HERA can be filled with a maximum of 210 bunches of each leptons and protons spaced by 96 ns. Some of these bunches are kept empty (*pilot bunches*) in order to study the background conditions. When either the lepton or the proton bunch is empty, the

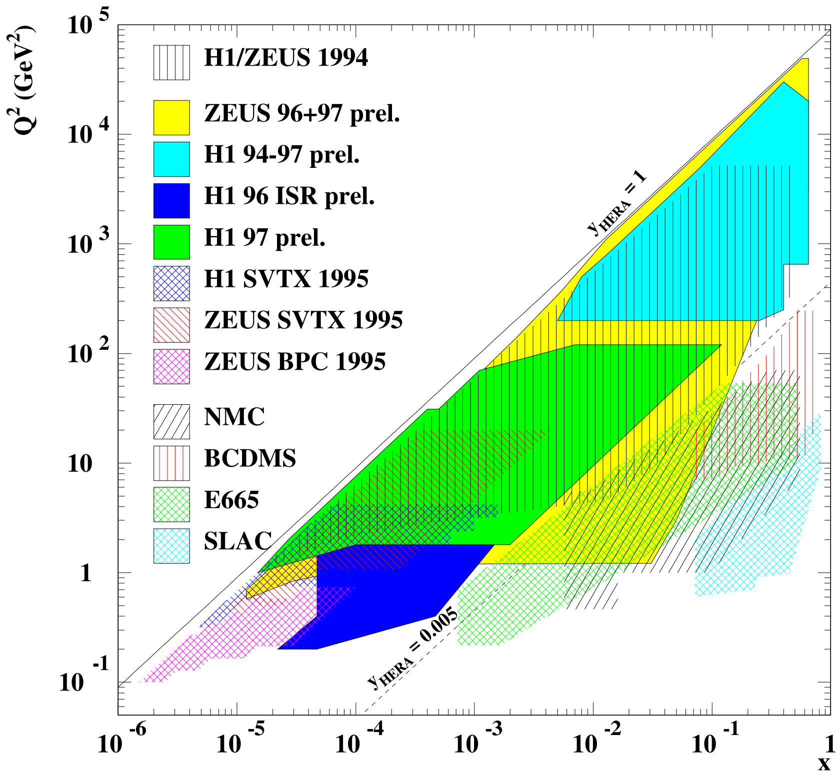

Figure 2.1 The kinematic range in the (x, Q^2) plane accessible by fixed target and HERA experiments.

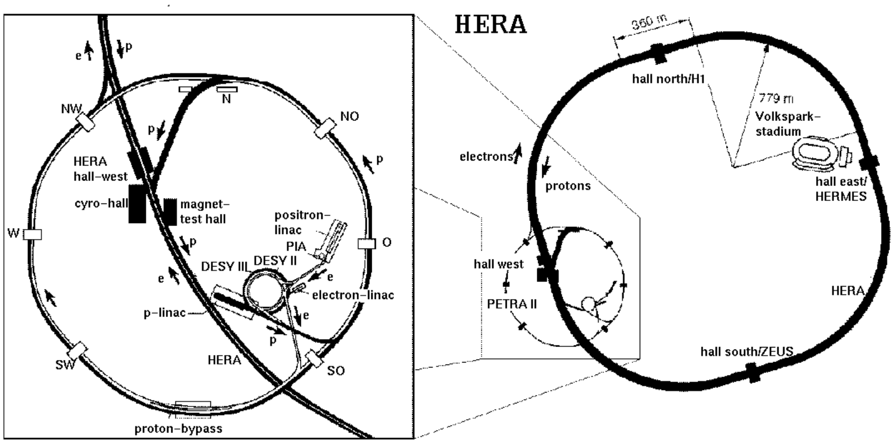

Figure 2.2 The HERA storage ring (*right*) with its pre-accelerators system (*left*).

Main General Design Parameters of HERA		
Construction time	May 1984–November 1990	
Circumference of the HERA tunnel	6335.83 m	
Depth underground	15–30 m	
Number of pre–accelerators for HERA	6 (LINAC I, LINAC II, DESY II, PETRA II, H–LINAC, DESY III)	
Number of interaction points	4	
The HERA beams	Lepton	Proton
Center–of–mass energy (actual)	318 GeV	
Nominal energy (actual)	27.52 GeV	920 GeV
Magnetic field	0.165	4.65
Relative energy spread $\Delta E/E$	10^{-3}	10^{-4}
Injection energy	14 GeV	40 GeV
Luminosity per interaction point	$1.6 \times 10^{31} \mathrm{cm}^{-2}\mathrm{s}^{-1}$	
Design Average Current	58 mA	163 mA
Particles per bunch	3.65×10^{10}	10^{11}
Maximum number of bunches	210	210
Beam crossing angle	head–on collisions, 0 mrad	
Bunch distance	96 ns (28.8 m)	
Beam length at maximum energy (σ_z)	0.85 cm	19 cm
Beam width at maximum energy (σ_x)	0.286 mm	0.28 mm
Beam height at maximum energy (σ_y)	0.06 mm	0.058 mm
Synchrotron radiation loss per turn	125 MeV	6×10^{-6} MeV
Polarization time at 30 GeV	35 min	–
Filling time	15 min	20 min

Table 2.1 Main design parameters of the HERA collider [60].

beam–related background, originating from the interaction of the lepton or the proton beam with the residual gas in the beam pipe, can be studied, whereas when both bunches are empty the non–beam–related background can be estimated, as the rates of cosmic rays. The main design parameters of HERA are summarized in Table 2.1.

The luminosity delivered by HERA in the last years is shown in Figure 2.3 and summarized in Table 2.2. The running operations began in 1992 with an electron beam, but in 1994 it was realized that the electron beam current was limited by positively ionized dust particles getting in the beam pipe through the pumps, reducing the lifetime of the beam. For this reason HERA switched to positrons in July 1994, achieving a more stable lepton beam and a significant increase in the integrated luminosity of the collected data. During the 1997-98 shutdown period, new pumps were installed in the lepton beam to improve the electron beam lifetime, and therefore during 1998 and part of 1999 HERA was running again with electrons. Furthermore, during the 1998 the energy of the proton beam was raised from 820 to 920 GeV.

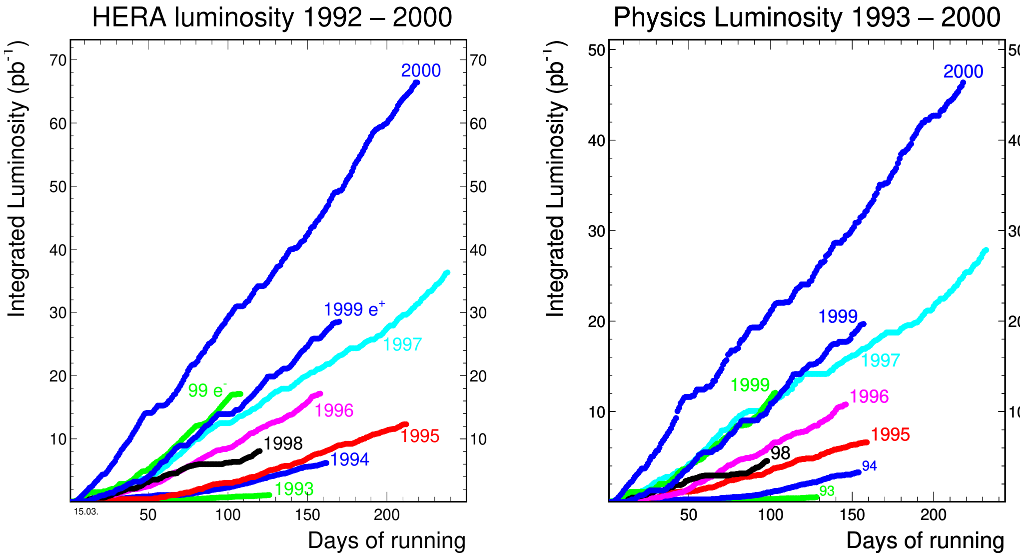

Figure 2.3 Integrated luminosity delivered by HERA *(left)* and recorded by the ZEUS experiment *(right)* in the 1993-2000 running period.

Year	HERA Luminosity (pb^{-1})
e^- 92 − 94	2.19
e^- 94 − 97	70.92
e^- 98 − 99	25.20
e^- 99 − 00	94.95

Table 2.2 Summary of the luminosity delivered by HERA from 1992 to 2000.

2.2 The ZEUS detector

The ZEUS detector is a multipurpose detector designed to study lepton–proton scattering at HERA. It is a quasi–hermetic detector since it covers most of the 4π solid angle, with the exception of small regions around the beam pipe. ZEUS was commissioned and upgraded during the years of data-taking taking into account the needs of the physics and the technical understanding gained during the running period.

The layout of the detector was dictated by the topology of the processes produced at HERA. The detector can measure energies from few MeV to hundreds of GeV in the forward region. For low–momentum particles the tracking in the magnetic field is very precise, while high energy particles are well measured by the calorimetric system[1].

Particle identification is needed in a wide momentum range: in NC DIS events the scattered lepton has to be identified and measured with high precision; the identification of electrons,

[1] A precise measurement of the hadronic final state in an event is achieved combining the information of both detectors. In Section 3.3.1 we describe how this is done.

positrons and muons is needed in order to study the semi-leptonic decay of heavy quarks and exotic processes involving leptons. In CC DIS processes a hermetic detector is needed in order to reconstruct the missing transverse momentum carried by the outgoing neutrino.

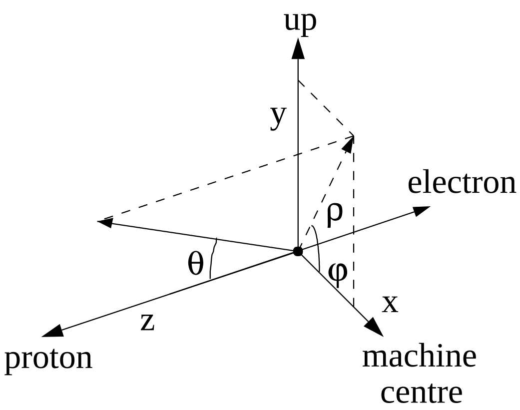

Figure 2.4 The ZEUS coordinate system.

The ZEUS coordinate system is shown in Figure 2.4. It is a right-handed, orthogonal system, with the origin at the nominal interaction point (IP), the z axis pointing in the proton direction (also referred to as the *forward* direction), the x axis pointing toward the center of HERA and the y axis pointing upward. Because of the large momentum imbalance between the lepton and the proton beam, most of the final state particles are boosted in the forward direction, and therefore the ZEUS sub-detectors are coaxial but asymmetric with respect to the IP. The detector layout is shown in Figure 2.5 and 2.6. The main detector is approximately 20 m long, 12 m large and 11 m high, and it weights around 3600 T.

The innermost detector that can be seen in Figs. 2.5 and 2.6 is the Vertex Detector (VXD), that was removed during the 1995-96 shutdown. Therefore during 1996-2000 data-taking the detector closest to the interaction point was the Central Tracking Detector (CTD), which is surrounded by a super-conducting solenoid, giving a magnetic field of 1.43 T, for the determination of the charge and momentum of the particles. The forward (FTD) and rear (RTD) tracking detectors complete the tracking system.

Outside the super-conducting solenoid the ZEUS calorimeter is located, a compensating high–resolution uranium–scintillator calorimeter, divided into forward, barrel and rear sections (respectively FCAL, BCAL and RCAL), with different thickness. The calorimeter is enclosed by an iron yoke that provides the return path for the solenoidal magnetic field flux, and serves as absorber for the Backing Calorimeter (BAC), which measures energy that escapes detection from the main calorimeter. The muon detectors are located inside (FMUI, BMUI and RMUI) and outside (FMUO, BMUO and RMUO) the iron yoke.

The Small angle Rear Tracking Detector (SRTD) is a scintillator strip detector located on the front face of the rear calorimeter, around the beam pipe, used to measure the impact point of the scattered electrons with high accuracy.

Other detectors are located several meters away from the main detector along the beam pipe. The Leading Proton Spectrometer (LPS), consisting of six stations of silicon strip detector located at distances of $24 - 90$ m from the interaction point, measures protons scattered at very small angles in elastic reactions. The VETO, located about $z = -7.5$ m before the

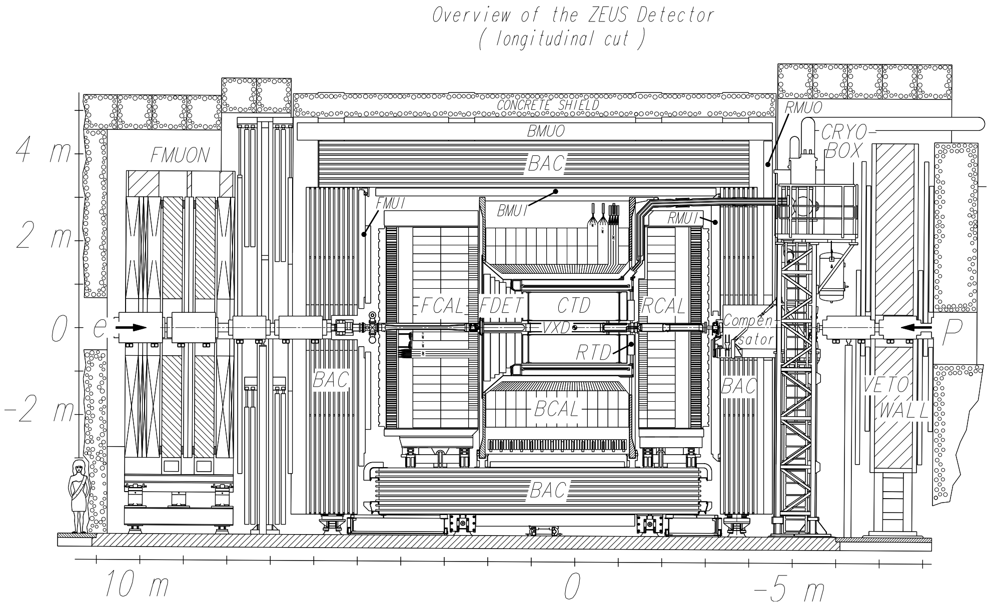

Figure 2.5 Cross section of the ZEUS detector along the beam axis.

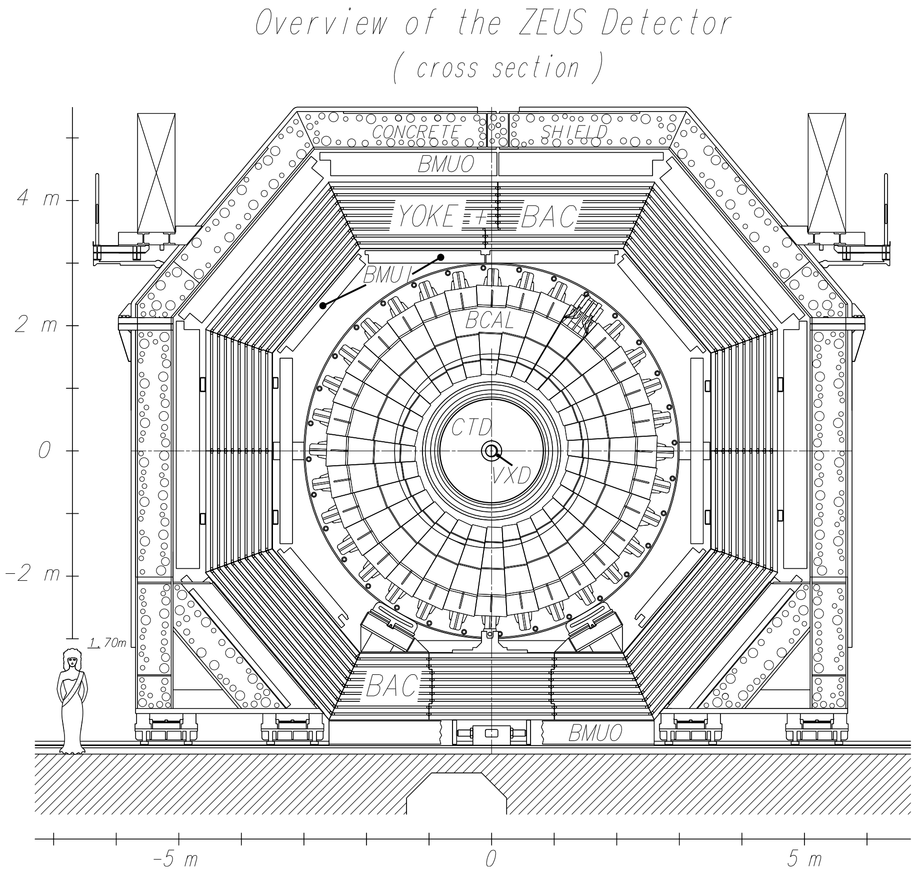

Figure 2.6 Cross section of the ZEUS detector in the xy plane.

interaction point and consisting of an iron wall supporting scintillator hodoscopes, is used to reject background from beam–gas interactions. The LUMI detector, made of two small lead–scintillators calorimeters, at $z = -35$ m and $z = -(104 \div 107)$ m, detects electrons and photons from bremsstrahlung events for the luminosity measurement.

2.3 The Central Tracking Detector

The Central Tracking Detector (CTD) [61, 62] is a cylindrical wire drift chamber used to measure the direction and momentum of the charged particles and to estimate the energy loss dE/dx to provide information on particle identification. The inner radius of the chamber is 18.2 cm, the outer is 79.4 cm, and its active region covers the longitudinal interval from $z = -100$ cm to $z = 104$ cm, resulting in a polar angle coverage of $15° < \theta < 164°$. The CTD is filled with a mixture of argon (Ar), carbon dioxide (CO_2) and ethane (C_2H_6) in the proportion 85:5:1. It consists of 72 radial layers of sense wires, divided in groups of eight into nine *superlayers* (SL). A group of eight radial sense wires in a superlayer, with the associated field wires, makes up a *cell*. The sense wires are 30 μm thick, while the field wires have different sizes. A total of 4608 sense wires and 19584 field wires is contained in the CTD. The field wires are tilted of 45° with respect to the radial direction, in order to compensate the Lorentz angle of 45° due to the electric and magnetic field.

A charged particle crossing the CTD produces ionization of the gas in the chamber. The electrons from the ionization drift toward the sense wires (positive), whereas the positively charged ions drift toward the negative field wires. The drift velocity of the electron is approximately constant and equal to 50μm/ns; close to the wire an avalanche effect occurs, giving an amplification factor on the electrons of $\sim 10^4$, so that a measurable pulse is induced on the sense wires.

The superlayers are numbered so that the number 1 is the innermost SL, whereas the outermost is number 9. Odd numbered SLs have wires parallel to the z direction (*axial superlayers*), while wires in even numbered SLs are at a small stereo angle of $\pm 5°$ (*stereo superlayers*) to achieve a better resolution in z. The achieved resolution is ~ 200 μm in the $r - \phi$ plane and ~ 2 mm in the z coordinate.

For trigger purposes, the three inner axial superlayers (SL1, SL3, SL5) are equipped with a system that determines the z positions using information on the arrival time of the particle (z–by–timing system). The resolution achieved on the z coordinate with this system is ~ 4 cm.

The resolution on the transverse momentum, p_T, for tracks fitted to the interaction vertex and passing at least three CTD superlayers, and with $p_T > 150$ MeV, is given by:

$$\frac{\sigma(p_T)}{p_T} = 0.0058 \cdot p_T \oplus 0.0065 \oplus \frac{0.0014}{p_T} \tag{2.1}$$

where the symbol $\oplus$ indicates the quadratic sum. The first term is the hit position resolution, while the second and the third depend on the multiple scattering inside and before the volume of the chamber, respectively.

2.4 The Uranium-scintillator Calorimeter

The ZEUS calorimeter (UCAL) [63, 64, 65, 66] is a high-resolution compensating calorimeter. It surrounds completely the tracking devices and the solenoid, and covers 99.7% of the 4π solid angle. It consists of 3.3 mm thick depleted uranium plates (98.1% U^{238}, 1.7% Nb, 0.2% U^{235}) as absorber alternated with 2.6 mm thick organic scintillators (SCSN–38 polystyrene) as active material. The thickness of the absorber and of the active material have been chosen in order to have the same response for an electron or a hadron of the same energy ($e/h = 1.00 \pm 0.02$) passing through the detector. This mechanism is called *compensation*, and allows to achieve good resolution in the determination of both the electromagnetic and the hadronic energy. The achieved electromagnetic resolution is

$$\frac{\sigma(E)}{E} = \frac{18\%}{\sqrt{E}} \oplus 2\%, \tag{2.2}$$

while the hadronic resolution is

$$\frac{\sigma(E)}{E} = \frac{35\%}{\sqrt{E}} \oplus 1\% \tag{2.3}$$

where E is the particle energy, measured in GeV.

The UCAL is divided into three regions: the forward (FCAL), barrel (BCAL) and rear (RCAL) calorimeter (see Figure 2.7). Since most of the final state particles in a lepton-proton interaction at HERA are boosted in the forward (proton) direction, the three parts are of different thickness, the thickest one being the FCAL ($\sim 7\ \lambda$), then the BCAL ($\sim 5\ \lambda$) and finally the RCAL ($\sim 4\ \lambda$), where λ is the interaction length. Each region of the calorimeter

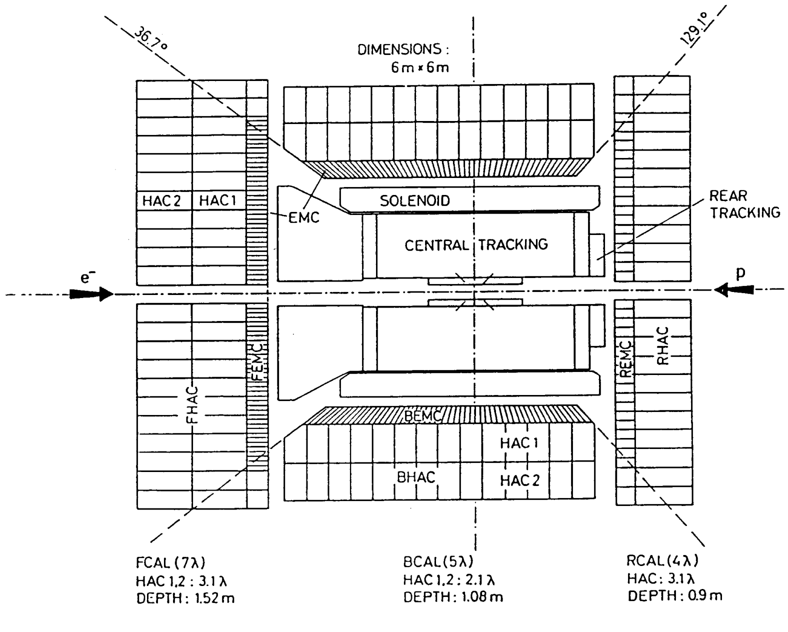

Figure 2.7 Schematic view of the UCAL along the beam axis.

is divided into *modules*, and each module is divided into one electromagnetic (EMC) and two (one in RCAL) hadronic (HAC) sections. These sections are made up of *cells*, whose sizes depend on the type (EMC or HAC) and position (in FCAL, BCAL or RCAL) of the cell, as reported in Table 2.3.

The FCAL consists of one EMC (first 25 uranium-scintillator layers) and two HAC (remaining 160 uranium-scintillator layers) sections. The electromagnetic section has a depth of

UCAL Part	Angular coverage	EMC	HAC
FCAL	$2.5^o - 39.9^o$	20×5 cm^2	20×20 cm^2
BCAL	$36.7^o - 129.2^o$	20×5 cm^2	20×20 cm^2
RCAL	$128.1^o - 178.4^o$	20×10 cm^2	20×20 cm^2

Table 2.3 Angular coverage of the UCAL parts and dimensions of the cells.

26 X_0, while each hadronic section is 3.1 λ deep. The EMC and HAC cells are superimposed to form a rectangular module (see Figure 2.8). 23 of these modules make up the FCAL.

The BCAL consists of one EMC and two HAC sections, the EMC being made of the first 21 uranium–scintillator layers, the two HACs of the remaining 98 layers. The resulting depth is 21 X_0 for the electromagnetic section, and 2.0 λ for each hadronic section. The cells are organized in 32 wedge-shaped modules, each covering 11.25^o in azimuth.

The RCAL is made up of 23 modules similar to those in the FCAL, but it consists of one EMC and only one HAC section. Therefore its depth is 26 X_0 for the EMC part and 3.1 λ for the HAC part.

The light produced in the scintillators is read by 2 mm thick wavelength shifter bars at both sides of the module, and brought to one of the 11386 photomultiplier tubes (PMT) where it is converted into an electrical signal. These information are used for energy and time measurement. The UCAL provides accurate timing information, with a resolution of the order of 1 ns for particles releasing an energy deposit greater than 1 GeV. These information can be used to determine the timing of the particle with respect to the bunch-crossing time, and it is very useful for trigger purposes in order to reject background events, as we will discuss in Section 2.8.

The stability of the PMTs and of the electronics is monitored with lasers and charge pulses. In addition, the small signal coming from the natural radioactivity of the depleted uranium (wrapped in stainless steel foils not to interfere with the physical particles measurements) gives a very stable signal, also used for the calibration. The achieved accuracy is better than 1%.

2.5 The muon detectors

The main aim of the muon detectors is the measurement of the tracks produced in the interaction region (*pointing tracks*) and which cross the whole calorimeter thickness and the iron yoke (*penetrating tracks*). Muons can traverse large amount of material without being absorbed since they release energy only by ionization.

The muon detection system, as the other detectors in ZEUS, has to cope with the boost of the particles in the forward direction. The momenta of the muons can be very different depending on their polar angle: in the forward region muons with more than 10 GeV momentum are easily produced, while in the barrel and rear regions the average momentum of the muons is expected to be much smaller. Therefore the muon detection system is split into two subdetectors, designed and realized in different ways: the Forward Muon Detector (FMUON) and the Barrel and Rear Muon Detector (BMUON and RMUON respectively).

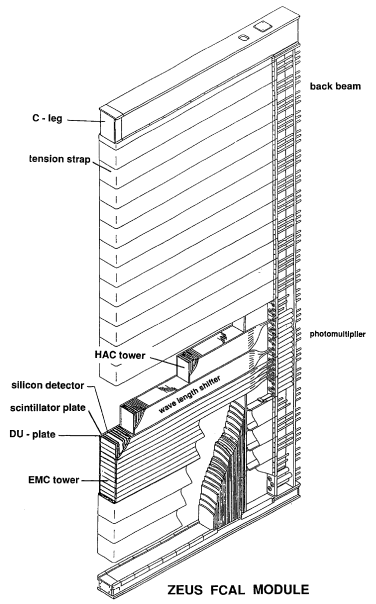

Figure 2.8 A module of the forward calorimeter.

2.5.1 The Forward Muon Detector

The Forward Muon Detector is divided in two regions (see Figure 2.9): one is located between the FCAL and the BAC (FMUI), the other is positioned outside the BAC (FMUO). The FMUON detector consists of:

- a system of four planes of limited streamer tubes [67] trigger planes (LST1÷LST4), with digital ρ and ϕ readout;
- two coverage planes of limited streamer tubes with digital (ρ, ϕ) and analog ρ readout, in the large polar angle region (LW1, LW2);
- four planes of drift chambers (DC1÷DC4);
- two large toroidal iron magnets providing a magnetic field of 1.7 T for the momentum separation and measurement in the angular region $5^\circ < \theta < 16^\circ$.

The first plane of LST and drift chambers make up the FMUI detector, while the rest of the system is placed outside the iron yoke.

The Limited Streamer Tubes planes

The aim of the limited streamer tubes (LST) planes is to trigger on muon candidates and to reconstruct their position in terms of the azimuthal and radial coordinates of the track.

A trigger plane is made of four LST chambers, grouped by two in two half-planes. A quadrant consists of two layers of LST, positioned horizontally inside a plastic sheet. The tubes of the two planes are slightly displaced (0.5 cm) in order to achieve a complete geometrical acceptance. Each quadrant is contained in an aluminum air tight box. The signals generated by the LST

Figure 2.9 Schematic view of the forward muon detector along the beam axis.

are induced on copper strips with polar geometry, glued on the outer side of the plastic sheet. The number of radial ρ strips is 132, each strip is 1.9 cm wide. They are divided along the bisector of the quadrant so that the simplest unity of the trigger plane to be read is the octant. The number of ϕ strips is 32 per octant: each strip covers an interval of 1.4° in the azimuthal angle.

The Drift Chambers

The drift chambers are needed in order to obtain a good momentum resolution. Each plane consists of four chambers, grouped two by two in two half planes, fixed on a sustain panel. The basic constituent of the chamber is the *cell*, made of four sense wires and of the layers needed to generate the appropriate electric field. The four sense wires measure the radial coordinate. The information gathered by the wires are sent to a TDC, which converts them into a time distance, connected to the space distance by a known relation.

The Large Angle Coverage planes

The two large angle coverage planes (LW) are needed in order to achieve the desired geometrical acceptance also in the region left uncovered by the toroids ($16° < \theta < 32°$). Each plane consists of eight aluminum tight wrappings that contain a LST layer. The LST signal is induced on copper strips with radial geometry, spaced of $0.7°$ in the ϕ coordinate and of 1.8 cm in the ρ coordinate. The number of ϕ strips is 64 per octant, while the ρ strips are 192 per octant. The achieved resolution in the ρ coordinate, using a center of gravity algorithm, is ~ 1 mm.

2.5.2 The Barrel and Rear Muon Detector

The Barrel and Rear Muon Detector [68] covers a very large area, of the order of 2000 m^2, and it has a modular structure. The basic element is the *chamber*. The chambers covering the inner barrel part, between the CAL and the iron yoke, are called BMUI, while the chambers situated outside the yoke, are denoted as BMUO. In a similar way, in the rear region the detector is divided into RMUI and RMUO chambers. In Figure 2.10 an layout of the B/RMUON detector is shown.

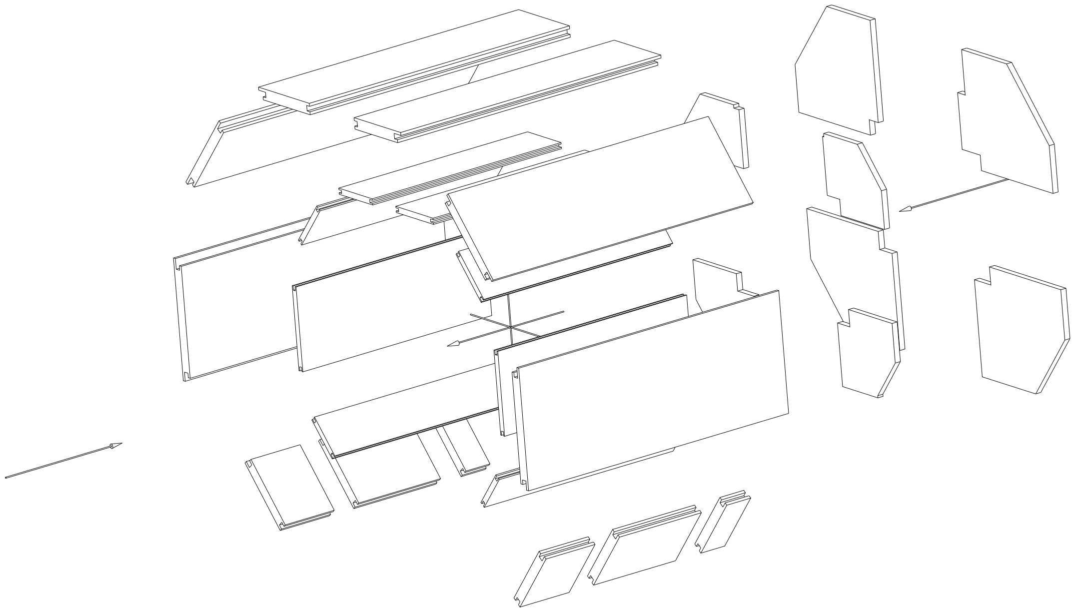

Figure 2.10 Layout of the Barrel and Rear Muon Detector.

The chambers have different shapes and dimensions depending on their location, but their internal structure is always the same. The supporting structure of each chamber is an aluminum honeycomb frame, 20 cm thick in the rear chambers, 40 cm in the barrel ones. Two planes of LST are placed on both sides of the honeycomb, as shown in Figure 2.11. The two layers on the same side of the chamber are staggered by 8.3 mm in order to minimize dead areas for particles traversing at 90° with respect to the wire plane.

Each LST is made of a plastic profile with eight cells. In each cell a copper-berillium wire of 100 μm diameter is disposed. The distance between two sense wires is 1 cm. The LST's are made out of Noryl ENV509 [2], a plastic material that does not contain chlorine and hence presents much less of a safety hazard in case of a fire than previously used PVC types. The varnish chosen for the Noryl substrate is a carbon based one whose resistivity is in the range $5 \div 15$ kΩ/□.

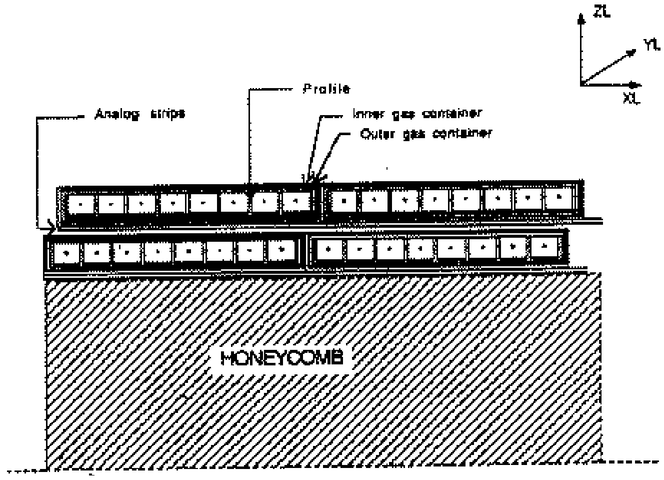

Figure 2.11 Transverse section of a BMUON chamber, showing the honeycomb structure and two layers of LST glued on one side of the frame.

During the data–taking the sense wires are brought to $\sim$ 4500 V and they constitute the anodes, while the inner cell surfaces, act as cathodes. The cells are fluxed with a mixture of carbon dioxide (CO_2), argon and isobuthane (C_4H_{10}). Each LST plane is equipped on one side by read-out strips that run orthogonal to the wires. These strips are obtained by 1.6 thick G10 boards, machined on one side to form 13 mm wide strips with 15 mm pitch. The boards are glued to the LST gas container and their relative position is then fixed in the mounting phase. In the BMUI and BMUO chambers the LSTs are parallel to the beam direction, while in RMUI and RMUO they are horizontal (parallel to the ZEUS x direction).

The wire read-out is based on a module which collects signals from 32 adjacent wires capable of recording the pattern of hit wires in relation to a generic trigger. When a muon candidate is detected by the trigger the information recorded is more complete, consisting in addition of an on-board multi-hit TDC output.

The strip read-out is of analog type [69]. It is based on a module that processes signals from 32 strips. In the event of a generic trigger not involving muons, data produced by this module consist of digital information: the strip signals are OR-ed in group of 4 and the resulting pattern of the eight OR's correlated to the trigger is stored. In this case the resolution on the single hit is $1/\sqrt{12}$ cm in the direction orthogonal to the wires, $6/\sqrt{12}$ cm along the wires. In the event a muon global trigger the information is integrated by the A/D conversion of all the actual amplitudes of the 32 strip signals correlated to the correct trigger. This information allows to reconstruct the position of the track in this coordinate with a resolution better than 1 mm.

There are two types of trigger signals in an event: when the muon chambers identify a

[2]Noryl is a product of General Electric Plastics.

penetrating track, giving activity in the inner and outer muon chambers, and *pointing* to the central tracking devices, the trigger signal is called *strong*; on the other hand, if the muon chambers find activity in at least one of the inner chambers, the trigger signal is called *weak*.

The trigger signal is determined by the analysis of the wires and strips hit in the event. First of all, the adjacent wires in the same LST plane are OR'ed in groups of 16, and the same is done with the strips. The signal obtained from these groups is further analysed in order to obtain information on the polar and azimuthal position of the wires and strips that have been hit. Coincidence matrices designed for this purposes analyse the information coming from the inner and the outer muon chambers, and select configurations generated by tracks coming from the volume corresponding to the inner tracking detectors (*pointing tracks*). In case of positive response by the matrices, a strong trigger signal is present.

In case of a strong or weak trigger signal in the event the following information are recorded: the addresses of the hit wires, the drift distance wire by wire, the addresses of the hit strips and the charge collected by each strip.

2.6 The Small Angle Rear Tracking Detector

The precise determination of the electron impact position is required for the reconstruction of the kinematic variables. The angular resolution of the rear calorimeter is unfavourable at low deflection angles of the scattered electron. The Small Angle Rear Tracking Detector (SRTD) was designed to improve the position measurement and identification of charged particles within its acceptance, identify pre-showering electrons and upstream proton-beam-gas interactions [70].

The SRTD consists of two layers of scintillators strips, 24 or 44 cm long. It covers a surface of $\sim 4200\ \mathrm{cm}^2$ and it surrounds the beam pipe hole of the RCAL (see Figure 2.12). The strip cross section is 5.0 ± 0.2 mm by 9.80 ± 0.05. The scintillation light is detected by photomultipliers (PMTs) via light guides. Space for housing the PMTs and front-end electronics was available on the sides of the RCAL at approximately 2.5 m distance from the active detector. The SRTD read-out electronics is similar to that used for the ZEUS calorimeter. A detailed description can be found in Ref. [65].

The position resolution was determined from the data using DIS electrons with impact position near the boundary between two calorimeter cells [70]. In this restricted region, selected using the energy imbalance between two neighboring cells, the calorimeter position resolution is better than 0.1 cm although the average resolution is only 1 cm. The measured SRTD position resolution is approximately 0.3 cm in both coordinates (see Figure 2.13).

2.7 The luminosity measurement

The luminosity measurement at ZEUS is done by studying the bremsstrahlung events $ep \to ep\gamma$, where the lepton and the photon are scattered at very small angle. The cross section of this process is described by the Bethe–Heitler formula

$$\frac{d\sigma}{dk} = 4\alpha r_e^2 \frac{E'}{kE}\left(\frac{E}{E'} + \frac{E'}{E} - \frac{2}{3}\right)\left(\ln\frac{4E_pEE'}{m_e m_p k} - \frac{1}{2}\right) \tag{2.4}$$

where k, E and E' are respectively the energies of the photon, the incoming and outgoing electron. The theoretical prediction for the process is known with an accuracy of $\sim 0.5\%$,

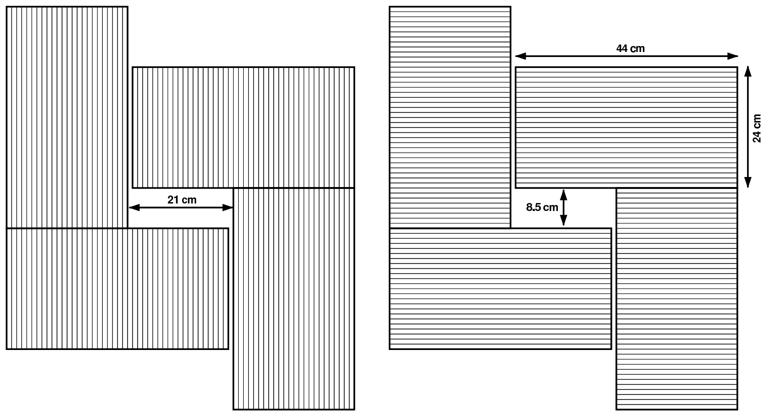

Figure 2.12 Schematic view of the SRTD layers. Each layer is divided in four quadrants of scintillator strips.

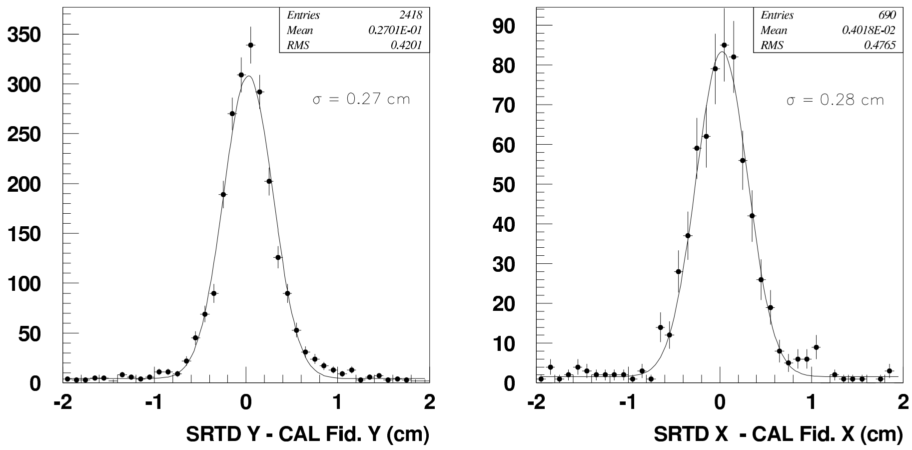

Figure 2.13 Distribution of the difference in the y-coordinate reconstructed by the SRTD and RCAL in the fiducial boundary region between two RCAL cells (*left*). The same distribution is shown for the x-coordinate (*right*).

allowing a precise determination of the luminosity.

The luminosity monitor consists of a photon and a lepton calorimeter, located along the beam line at $z = -(104 \div 107)$ m and $z = -35$ m, respectively (Fig. 2.14). To protect the photon calorimeter against synchrotron radiation, it was shielded by a carbon-lead filter. The resulting calorimeter resolution is $\sigma(E)/E = 23\%/\sqrt{E}$. The bremsstrahlung event rate is determined by counting the number of the photon above a fixed energy threshold, and not by the simultaneous identification of the lepton and the photon, because of the dependence of the lepton calorimeter acceptance on the beam position and angle. The luminosity is then calculated by dividing the evaluated rate by the bremsstrahlung cross section, corrected for the detector acceptance.

The main contribution to the background is given by the bremsstrahlung of leptons on the residual gas in the beam pipe. This can be estimated using *pilot* bunches, i.e. lepton bunches with no associated proton bunches, evaluating for these the rate of bremsstrahlung events. The achieved precision on the luminosity measurement is of the order of 1.5–2%.

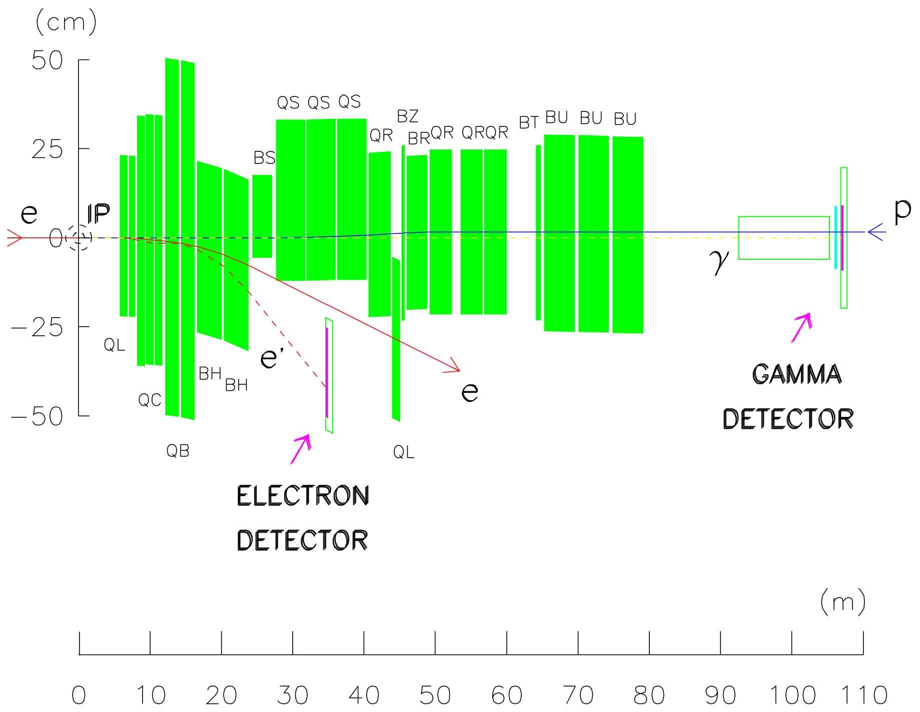

Figure 2.14 Schematic view of the luminosity monitor.

2.8 The ZEUS trigger system

The ZEUS trigger is structured in three levels [71]. The First Level Trigger has the task to set a trigger rate of 1 kHz by eliminating most of the beam-gas and beam-halo background. Programmable logic is applied at this level. Results of the local processors are combined in the global first level trigger (GFLT). At the Second Level commercially available microprocessors (mainly INMOS Transputers) analyze the digitized data of the components. Combining the results of the local processors in the global second level trigger (GSLT) results in a further reduction of the trigger rate to 100 Hz. After a positive decision the data of all sub-detectors is gathered and sent to a Third Level processor farm of Silicon Graphics stations.

2.8.1 The First Level Trigger

The First Level Trigger strongly suppress the following background events:

- events coming from interactions of leptons or protons with the residual gas in the beam pipe, near the interaction point. The estimated rate of this kind of events, assuming a sensitive region of 100 m before and after the interaction point, with the nominal beam currents and a vacuum of $\sim 10^{-9}$ Torr, is 50 kHz;

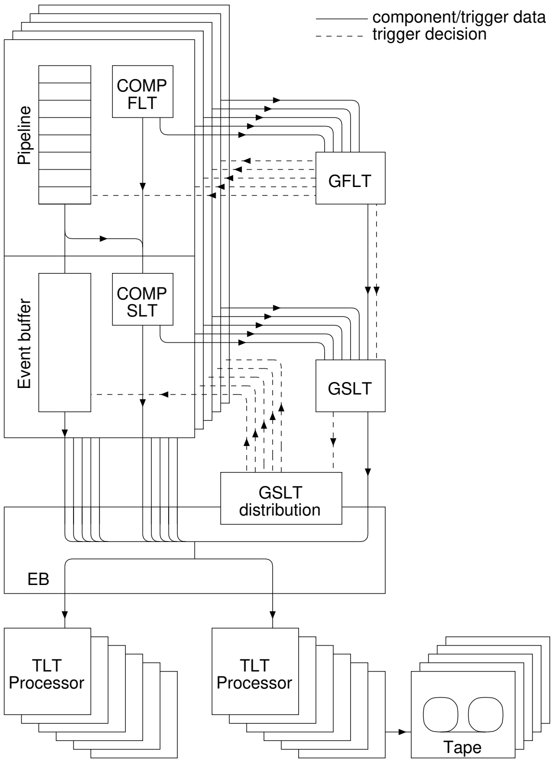

Figure 2.15 Diagram of the ZEUS trigger and data acquisition chain.

- events coming from interactions of the protons in the beam halo with the collimators, that can produce secondary hadrons decaying into high energy muons, crossing all the detector. In these events, however, the muons are typically parallel to the proton beam direction and can easily be distinguished from *ep* events;

- cosmic ray muons, with an incoming rate of the order of 1 kHz.

For every bunch crossing all data are stored in a pipeline, clocked at 96 nsec, for 100 ms while the first level trigger calculations are being performed and the first level trigger signal is propagating back to the component. For some components, such as the calorimeter, this pipeline is a switched capacitor array, which is an analog device that stores charge. For other components, such as the tracking detectors, the information is already digital and stored in a digital pipeline. The trigger processing for all components is pipelined and dead-timeless, accepting data from a new crossing every 96 nsec.

The FLT operates on a subset of the full data. Each component completes its internal trigger calculations and passes information for a particular crossing to the GFLT between 1.0 and 2.5 μs after the crossing occurred. The components data are processed and combined in parallel in eight Trigger Logic Modules, and 64 individual sub-triggers (*slots*) are generated. The GFLT decision is a logical OR of these 64 sub-triggers. The GFLT calculations take a minimum of 20 crossings (1.9 μs) additional time after receiving information from the individual components. The GFLT is issued exactly 46 crossings, or 4.4 μs, after the crossing that produced it. If a global first level trigger is not issued for a crossing, the component data are discarded.

The disposition of every event and some basic information about it is retained for analysis of trigger performance. There is additional processing of calorimeter trigger data by the Fast Clear between arrivals of global first level triggers. The Fast Clear aborts events before processing by the second level trigger. The design goal of the first level trigger, including the Fast Clear, is to reduce the input interaction crossing rate to below 1 KHz.

2.8.2 The Second Level Trigger

The issuing of a first level trigger causes component data to be transferred to buffers for processing by the second level trigger. The second level trigger processor functions as an asynchronous pipeline, i.e. a series of parallel processors. It has access to a large fraction of the full data for the event. For example, in the case of the calorimeter, while the first level trigger would examine groups of cells with a reduced digitization accuracy, the second level trigger performs calculations on the individual cells with the full dynamic range. The second level trigger is able to perform iterative calculations that are not possible in the pipelined structure of the first level trigger. These features enable the second level trigger to achieve a reduction of the first level trigger rate from 1 kHz to 100 Hz.

The SLT software is executed on a network of transputers. The analysis of the events is done in parallel so that the available processing time is much larger than at the FLT, of the order of some milliseconds. Each detector component has its own SLT processor, and all the information from the single components are sent to the GSLT after the processing.

The information the GSLT uses to distinguish between *ep* and background events are based mainly on the time of arrival of the particles to the calorimeter. As discussed in Section 2.4, the UCAL can give timing information with a resolution of the order of 1 ns. The time is calibrated so that a physics event originating from the interaction point has zero time in the whole calorimeter, when the produced particles arrive at it. A proton-gas event, originating upstream of the detector, produces particles reaching the RCAL before the FCAL, with a time difference of ~ 10 ns. In the same way, events produced in lepton-gas interactions downstream the detector will have particles reaching the FCAL before the RCAL, with a time difference larger than the UCAL timing resolution. Also cosmic events and electronic noise will appear as asynchronous to the HERA time and would therefore be suppressed.

The timing information from the UCAL is available when at least one cell above threshold (200 MeV) has been read by PMTs on both sides. The times are calculated by a weighted average on all the cells above threshold, for the different regions, with a bigger weight for the more energetic cells.

2.8.3 The Third Level Trigger

If the event is accepted by the GSLT, all the components send their information to the event builder (EVB), which combines their data, writes them in a standard format, and makes them accessible to the Third Level Trigger (TLT). The input rate to the TLT is 100 Hz, while the output rate is $3 \div 5$ Hz. The aim of the TLT is not only to reject background events but also to select the particular classes of physics processes produced in the *ep* interactions.

The TLT consists of a set of algorithms executed on a farm of SGI computers, The main information that the reconstruction program sends to the selection algorithms are the location and the energy of the calorimeter cells above threshold, together with the variables (momentum and position) of the tracks reconstructed by the CTD. The resolution on the energy and on the track variables is lower than what is achieved in the final reconstruction program. The data analysis detailed in this thesis uses a TLT algorithm that selects events with an electron and a muon in the final state. The algorithm is described in more details in Section 4.1.2.

After having accepted an event, the TLT sends the data via an optical link to the DESY computing center, where the events are written to disk to be available for further offline reconstruction and data analysis.

3 Event Reconstruction

The aim of this analysis is the measurement of b-quark production in deep inelastic scattering using the semi-leptonic decay into muons. The reaction is

$$e\ p \rightarrow e\ b\ \bar{b}\ X \tag{3.1}$$

where at least one of the b-quarks decays semi-leptonically in the reaction

$$b(\bar{b}) \rightarrow \mu^{\pm}\nu c \rightarrow \mu^{\pm}\ Jet. \tag{3.2}$$

Each event is therefore characterized by the presence of a positron scattered in the detector and of the decay products of the b-quark (or its anti-quark $\bar{b}$). Semi-leptonic b-decay produce muons which are very close to the fragmented hadron. The muon identification requires dedicated algorithms which can cope with high track multiplicities around the muon track and can exploit the tracking capabilities of the muon detectors. The remaining products of the b-quark fragmentation are reconstructed inclusively through the application of *jet algorithms*.

An event selected by the procedure described in Chapter 4 is shown in Figure 3.1. The left side of the figure represents a cross section of the ZEUS detector across the zy plane. The positron and proton bunches are coming from the left and right side of the figure respectively. After the collision, the scattered positron traverses the central tracker and releases its energy in the rear calorimeter. The two b-quarks fragment into jets of particles with high transverse momentum and back-to-back with respect to the beam line. A muon, escaping the calorimeter, traverses the inner and outer muon chambers and its trajectory is slightly deflected by the presence of a magnetic field in the iron joke. In the upper right part of the figure a cross section across the xy plane is shown enhanced around the muon segments. In the lower right section a schematic representation of the central tracker is drawn.

In this chapter we discuss the reconstruction of all the physical quantities needed to characterize a DIS event with heavy quarks in the final state.

3.1 Reconstruction of the kinematic variables

The kinematic of a DIS event is defined by two independent variables once the energy of the incoming lepton E_e and proton E_p are known. The physical quantities measured in the detector are the energy and the polar angle of the scattered lepton (E'_e, θ_e) and the four-momentum of the particles belonging to the hadronic final state ($E_h, p_{h,x}, p_{h,y}, p_{h,z}$). From the latter the hadronic system can be characterized by the scalar quantity δ_{had}, the total transverse

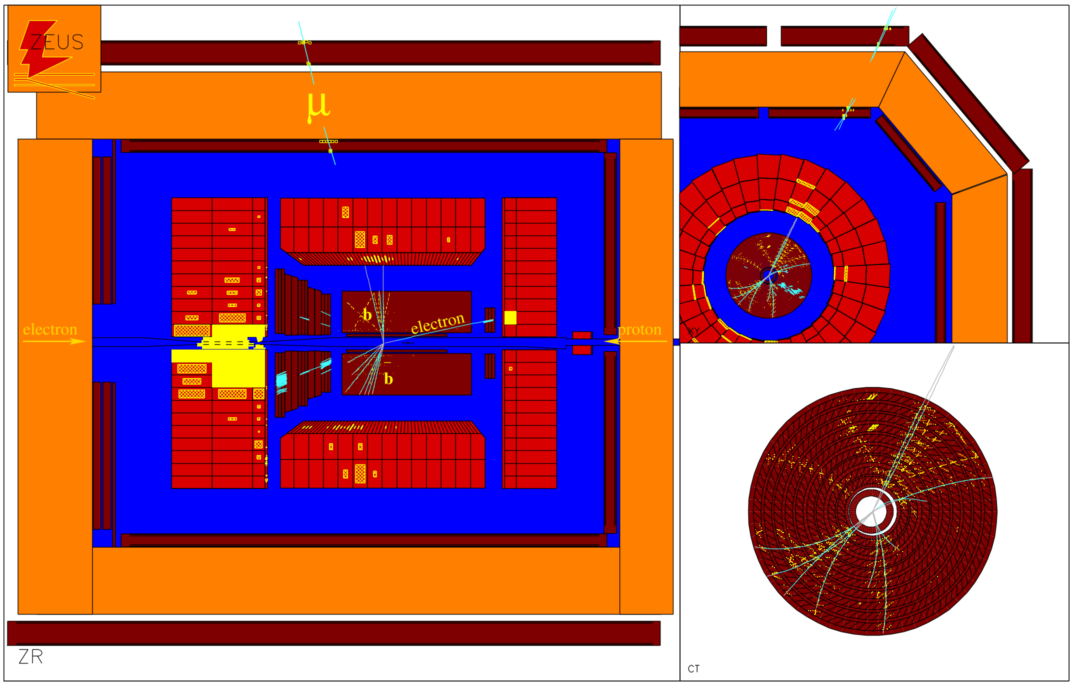

Figure 3.1 An event selected by this analysis. *On the left:* cross section of the ZEUS detector across the zy plane showing the two particle jets, the scattered electron and the muon traversing the inner and outer muon chambers. *On the right, top:* Particular of the detector cross section across the xy plane. *On the right, bottom:* Tracks reconstructed in the CTD and projected on the xy plane; two particle jets are reconstructed back-to-back with respect to the beam line.

momentum $p_{T,had}$ or the inclusive angle γ_{had}

$$\delta_{had} = \sum_{hadrons} (E_h - p_{h,z}) \tag{3.3}$$

$$p_{T,had} = \sqrt{\left(\sum_{hadrons} p_{h,x}\right)^2 + \left(\sum_{hadrons} p_{h,y}\right)^2} \tag{3.4}$$

$$\tan\frac{\gamma_{had}}{2} = \frac{\delta_{had}}{p_{T,had}} \tag{3.5}$$

The quantity δ_{had} is by construction minimally affected by the losses of particles scattered outside the detector acceptance in the forward direction, which are produced by the fragmentation of the proton remnant. Those particles have generally high energies but small transverse momenta. On the contrary, $p_{T,had}$ is more sensitive to the losses in the forward direction. The inclusive angle γ_{had} is a function of $p_{T,had}$ in which these effects are reduced and the uncertainties on the energy scale are canceled by the ratio. In conclusion, a sensible kinematic reconstruction can be obtained combining any pair of variables of the set $E_e, \theta_e, \delta_{had}, \gamma_{had}$.

3.1.1 Electron method

This method was used in fixed target experiments and it is certainly the easiest one since it only requires the measurement of one particle. The kinematic variables are expressed in terms of the energy and polar angle of the scattered lepton

$$Q_e^2 = 2E_e E'_e(1+\cos\theta_e) \tag{3.6}$$

$$y_e = 1 - \frac{E'_e}{2E_e}(1-\cos\theta_e) \tag{3.7}$$

$$x_e = \frac{Q^2}{s y_e} \tag{3.8}$$

where E_e is the energy of the lepton beam and s is the center of mass energy. The method is precise at high y, but its resolution diverges towards $y_e \to 0$ due to the $1/y$ dependence in the error propagation formula of the relative error [72]

$$\frac{\delta y_e}{y_e} = (1 - \frac{1}{y_e})\frac{\delta E_e}{E_e} \oplus (\frac{1}{y_e} - 1)\cot\left(\frac{\theta_e}{2}\right)\delta\theta_e. \tag{3.9}$$

As will discuss in Section 3.2.2, the electron angle can be measured with good precision taking advantage of the SRTD detector at low angles and of the tracking system at high angles, but the electron energy poses some problems due to the considerable amount of inactive material (1-5 X_0) between the interaction region and parts of the calorimeter surface.

3.1.2 Jacquet-Blondel method

The event kinematics can also be reconstructed using only the hadronic final state variables δ_{had} and $p_{T,had}$ [73]:

$$y_{JB} = \frac{\delta_{had}}{2E_e} \tag{3.10}$$

$$Q_{JB}^2 = \frac{p_{T,had}^2}{1-y_{JB}} \tag{3.11}$$

$$x_{JB} = \frac{Q^2}{s \cdot y_{JB}}. \tag{3.12}$$

This method can improve the y measurement based on the scattered lepton in the low y region since the relative error does not contain any term $\sim 1/y$:

$$\frac{\delta y_{JB}}{y_{JB}} = \frac{\delta E_{JB}}{E_{JB}} \oplus \cot\left(\frac{\theta_{JB}}{2}\right)\delta\theta_{JB} \tag{3.13}$$

3.1.3 Double angle method

This method relies on the electron polar angle and the angle γ_{had} of the hadronic system [74]

$$y_{\theta\gamma} = \frac{\sin\theta_e(1-\cos\gamma_{had})}{\sin\gamma_{had} + \sin\theta_e - \sin(\theta_e+\gamma_{had})} \tag{3.14}$$

$$Q_{\theta\gamma}^2 = 4E_e^2\frac{\sin\gamma_{had}(1+\cos\theta_e)}{\sin\gamma_{had} + \sin\theta_e - \sin(\theta_e+\gamma_{had})} \tag{3.15}$$

$$x_{\theta\gamma} = \frac{E_e}{E_p}\frac{\sin\gamma_{had} + \sin\theta_e + \sin(\theta_e+\gamma_{had})}{\sin\gamma_{had} + \sin\theta_e - \sin(\theta_e+\gamma_{had})} \tag{3.16}$$

This reconstruction method does not require a precise knowledge of the energy scales and the corrections due to the photon radiation are small. Nevertheless the resolution of all the variables depends strongly on the polar angle θ_e. Considering only the uncertainty due to the angular measurement we have:

$$\frac{\delta y_{\theta\gamma}}{y_{\theta\gamma}} \simeq -\frac{1-y_{\theta\gamma}}{\sin\theta_e}\delta\theta_e, \qquad \frac{\delta Q^2_{\theta\gamma}}{Q^2_{\theta\gamma}} \simeq -\frac{2}{\sin\theta_e}\delta\theta_e, \qquad \frac{\delta x_{\theta\gamma}}{x_{\theta\gamma}} \simeq -\frac{1+y_{\theta\gamma}}{\sin\theta_e}\delta\theta_e. \tag{3.17}$$

Therefore, at low Q^2, i.e. if θ_e is greater than about 175° the terms in Eqs. 3.17 dominate the total error and the electron or the Jacquet-Blondel methods give a more precise reconstruction [75].

3.1.4 The Σ and $e\Sigma$ methods

The Σ methods allows the determination of y and Q^2 independently of initial state photon radiation by reconstructing the incident electron energy [76]. The kinematic variables are given by the expressions

$$y_\Sigma = \frac{\delta_{had}}{\delta_{had} + E'_e(1-\cos\theta_e)} \tag{3.18}$$

$$Q^2_\Sigma = \frac{E_e^2 \sin^2\theta_e}{1-y_\Sigma} \tag{3.19}$$

$$x_\Sigma = \frac{E_e^2 \sin^2\theta_e}{s\cdot y_\Sigma(1-y_\Sigma)}. \tag{3.20}$$

The error propagation on y_Σ gives

$$\frac{\delta y_\Sigma}{y_\Sigma} = (1-y)\left(\frac{\delta(\delta_{had})}{\delta_{had}} \oplus \frac{\delta E_e}{E_e} \oplus \frac{\delta\theta_e}{\tan\theta_e/2}\right). \tag{3.21}$$

At low y there is no divergence whereas at high y the error increases as $(1-y)$ but is still dominated by the error on δ_{had}.

The $e\Sigma$ method combines the electron method and the Σ method. The Q^2 variable is calculated using the electron method (Eq. 3.6) whereas the y variable is calculated according to the Σ prescription, therefore

$$Q^2_{e\Sigma} = Q^2_e \tag{3.22}$$

$$y_{e\Sigma} = y_\Sigma \tag{3.23}$$

$$x_{e\Sigma} = \frac{Q^2_e}{s y_\Sigma}. \tag{3.24}$$

This combination optimizes the Q^2 resolution, while being quite insensitive to the initial photon radiation.

3.2 Identification and reconstruction of the scattered lepton

In the measurement of DIS processes the identification of the scattered lepton is of vital importance. The kinematics of the scattered electron in DIS events are summarized in Figure

3.2 where lines of constant electron energies and angles in the (x, Q^2) phase space are drawn. Due to the $1/Q^4$ dependence of the lepton–nucleon cross section most interactions involve small momentum transfer, i.e. the lepton is scattered at small angles and can eventually escape through the beam pipe, represented by the grey region in Figure 3.2 (on the right side).

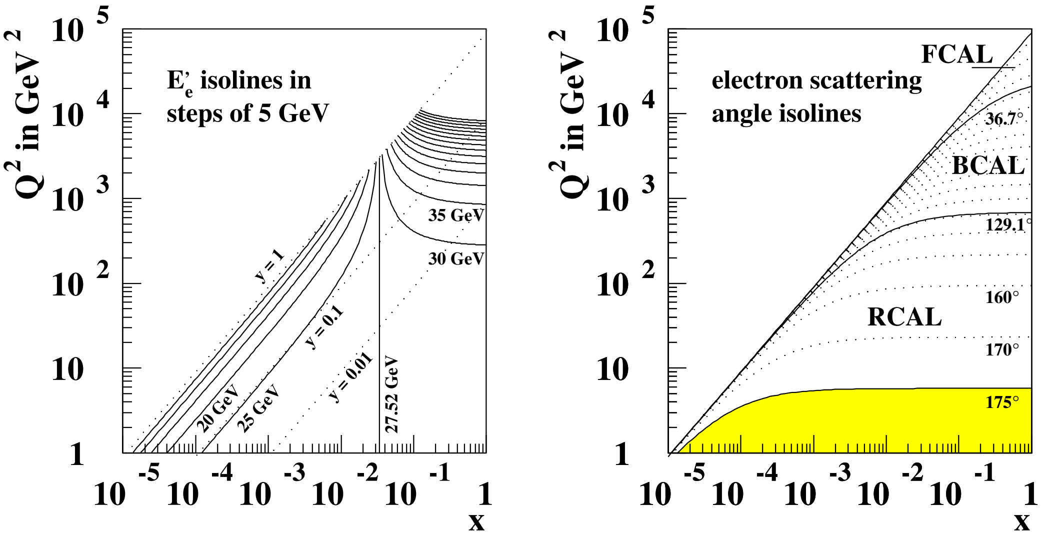

Figure 3.2 Contour lines of fixed electron energy (*left*) and angle (*right*) in the HERA kinematic plane. On the left plot the isolines corresponding to $y = 0.01$, 0.1 and 1 are shown. On the right plot the regions covered by the different sections of the calorimeter are separated by solid lines. The grey area corresponding to values of the polar angles $\theta_e > 175°$ which are outside the acceptance of the rear calorimeter.

3.2.1 Electron identification

The strategy followed by the ZEUS collaboration for the lepton identification is based on a neural network trained on patterns given by the showering properties of electrons and hadrons in the segmented calorimeter [77, 78]. The energy distribution of a shower is described in terms of a discrete set of points in the phase space and it is reconstructed in several steps.

Initially the various calorimeter cells are grouped into so-called *islands* by a clustering algorithm. Since the uranium CAL exhibits a varying granularity and projectivity the shower is then transformed into a coordinate space corresponding to a projective geometry. The number of points vary from shower to shower and cannot be used as an input of a *feed–forward* neural network. It is thus expanded in terms of moments of an infinite series. The moments characterize the shape of the shower developed in the calorimeter. Typically only terms up to a certain order are important: the higher order terms are sensitive to the tails of the energy distribution which may not be properly simulated in the Monte Carlo. The series is then truncated and only moments up to the second order are used for the transverse coordinates. The longitudinal shape of the shower is probed up to the third moment as it represents a strong discriminator between electromagnetic and hadronic showers. The one-dimensional output of this network can be interpreted as the probability $\mathcal{P}$ that the shower was produced by an electron, therefore the island is called an electron candidate if $\mathcal{P} > 90\%$.

Almost 83% of the final event sample selected by this analysis is characterized by Q^2 values below 100 GeV^2. Dedicated studies based on Monte Carlo simulations have shown that in the region defined by $\theta_e \geq 2.2$ rad, $E'_e \leq 30$ GeV and $Q^2 \leq 100$ GeV^2 the efficiency to identify the correct DIS electron is $\sim 85\%$ at 10 GeV reaching the full efficiency at ~ 17 GeV [79].

3.2.2 Reconstruction of the electron scattering angle

Once the electron has been identified in the calorimeter, the scattering angle θ_e can be measured combining the impact position at the calorimeter with the event vertex. The impact position is measured from the calorimeter cells associated to the electron candidate, but the CTD and SRTD detectors are used to improve the measurement whenever the electron trajectory lies within the respective acceptance region.

For the reconstruction of the electromagnetic shower in the RCAL only those cells are used, which are connected by at least one corner to the cell of maximum energy [80]. The y–coordinate is calculated through the logarithmic average of three cells to ensure a uniform resolution over the entire y-range of a cell. The x–coordinate is instead calculated starting from the identification of the shower center in each cell, given by the *imbalance* of the left and right energy measurements in one cell. From these three x–positions the final position is calculated using weights proportional to the logarithm of each energy. The position resolution achieved by this algorithm is ~ 1 cm in both coordinates.

The low Q^2 region corresponding to $\theta_e > 162°$ is covered by the SRTD detector whose fine segmentation allows a resolution almost three times more precise than the one of the calorimeter. The reconstruction of the impact position proceeds in three steps [81]. Signals above threshold of consecutive scintillator strips are grouped into clusters in each plane. The algorithm then pairs clusters of horizontal and vertical strips into hits and calculates the hit position with a center of gravity of the three central strips in the shower maximum. About 70% of the events selected in this analysis are reconstructed using this procedure.

In the region at higher Q^2 values ($\theta_e < 157°$) the reconstruction is improved by requiring an association of the calorimeter cluster with a good CTD track. The track must traverse at least three superlayers of the CTD, have a distance of closest approach to the calorimeter cluster of less than 10 cm and momentum greater than 5 GeV. In the angular region $157° < \theta_e < 162°$ the calorimeter position is used.

3.3 Reconstruction of the hadronic system

The energy resolution of a sampling calorimeter is mainly determined by sampling fluctuation in the shower development. Since shower phenomena are driven by statistical processes the fractional resolution improves with increasing energies scaling inversely as the square root of the incident energy E:

$$(\sigma/E) \propto 1/\sqrt{E} \tag{3.25}$$

Additional limitation at low energy may be introduced by instrumental effects whose relative contribution also decreases with E.

In case of multi–wire tracking chambers the precision of the momentum measurement depends on the curvature radius R of the particle trajectory in the magnetic field. It is straightforward to show that the error on the radius is of the form $\sigma(R) \sim \sigma_{hit}R^2$ and, since the transverse momentum p_T of the track is proportional to R, the form $\sigma(p_T) \sim p_T^2$ follows immediately. The fractional resolution is thus:

$$(\sigma/p_T) \propto p_T \tag{3.26}$$

The fractional resolutions measured in the ZEUS EMC calorimeter and the central tracker are drawn in Figure 3.3. From the picture it is easy to conclude that a sensible reconstruction of all particles belonging to the hadronic final state can be achieved combining the CTD measurement in the low energy region and the CAL measurement in the region starting from $10-15$ GeV.

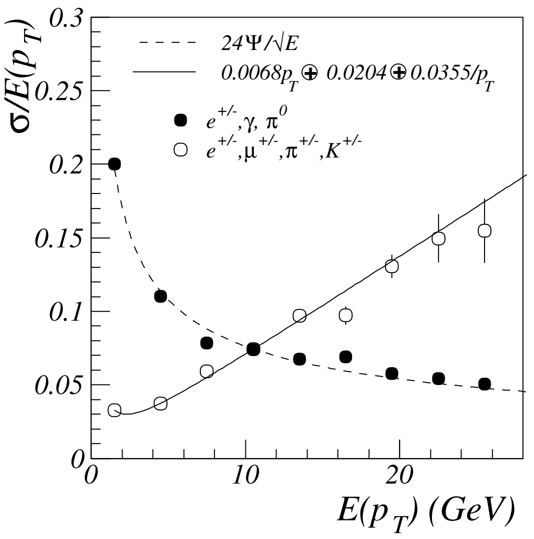

Figure 3.3 CTD and CAL electro-magnetic resolution [82]

In the next paragraphs we will discuss in detail how the particle reconstruction is achieved in the calorimeter and in the tracking system. In addition, we will explain why additional corrections are needed.

3.3.1 Reconstruction of the hadronic system using Energy Flow Objects

In Figure 3.4 a schematic representation of calorimeter and tracking information is shown. As a first step, after removing the calorimeter cells assigned to the scattered electron, the calorimeter energy deposits are clustered using a two stage algorithm. The reason for introducing a two stage algorithm is dictated by the structure of the ZEUS calorimeter. As mentioned in Section 2.4 it is divided into three, spatially separated main sections. This spatial separation implies several complications for a local clustering algorithm in handling the energy deposits of a single particle which is not confined to one calorimeter section, since the energy will be split into two or more clusters. Therefore, the first stage of the clustering procedure is performed in each calorimeter layer separately. Each cell with sufficient energy is considered a candidate to be connected with one of its neighbours. The connection is made with the nearest neighbour with the highest energy or with the highest energy cell next to the neighbour. This procedure is repeated for each cell and produces a unique assignment of a cell to a so called *cell island*.

The second stage of the clustering procedure collects the cell islands belonging to a shower of a single particle or a jet of particles into the so called *cone islands*. The matching of cell islands starts from the outermost hadronic layer of the calorimeter and proceeds inwards. The angular separation between cell islands of different layers is calculated in the θ-ϕ space and translated to a probability according to a distribution determined from a single pion MC [83]. Links with highest probability are accepted provided that the probability is larger than a threshold. Once the linking procedure has been completed, the cone islands are generated by combining all calorimeter cells which point to the same cone island in the electromagnetic layer.

Tracks are reconstructed independently using the ZEUS reconstruction software described in Ref. [84]. Each track is first extrapolated to the calorimeter surface and then further into the calorimeter, assuming a linear trajectory given by the track momentum unit vector at the end of the helix swim. The distance of closest approach is then calculated, defined as the shortest distance between the track trajectory and the calorimeter island.

The track-island association is dictated by the properties of the hadronic final state being analyzed. In an inclusive sample of DIS events the average multiplicity of particles is approximately proportional to $\ln(W^2)$. In these events some calorimeter islands are due to energy deposits of single particles whereas others are due to energy deposits of several particles. A sensible quantity to perform the track-island association is, therefore, the lateral extension of the calorimeter island. Thus, a track and a cone island with a distance of closest approach d are associated if the following condition is satisfied:

$$d \leq \max(r_{min}, r_{island}) \tag{3.27}$$

where r_{island} is the radius of the cone island and $r_{min} = 20$ cm is a parameter extracted from a MC simulation by maximizing the track-island matching efficiency for single particle calorimeter clusters [83].

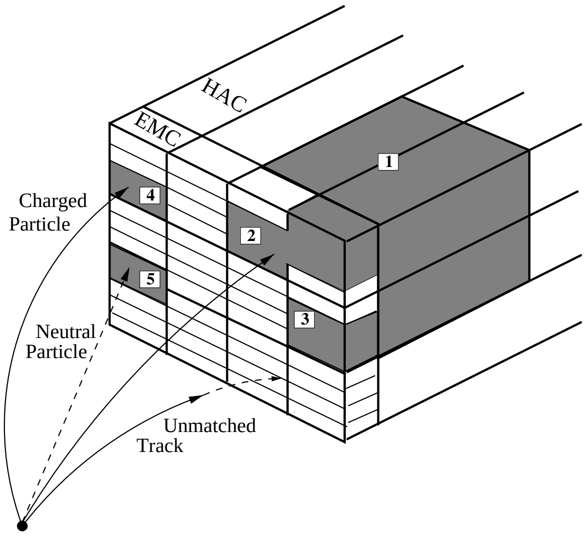

Figure 3.4 Schematic representation of an EFO object.

The set of associated track–islands, known as Energy Flow Objects (EFOs), are processed according to the following criteria:

- Charged tracks without any associated calorimeter object are considered as pions. The energy is calculated from the CTD measurement and the pion mass.
- Calorimeter islands without any associated track are considered as neutral particles.
- Calorimeter islands associated to more than three tracks are considered as one jet of particles.

In case of a one-to-one correspondence the substitution of the track momentum with the calorimeter island is based on two conditions. As a first check, the energy E_{cal} of the calorimeter island and the momentum p of the associated track must satisfy the condition:

$$E_{cal}/p < 1 + 1.2 \cdot \delta(E_{cal}/p) \tag{3.28}$$

where $\delta(E_{cal}/p)$ is the error on the energy/momentum ratio. In addition, the resolution of the momentum measurement must be smaller than the resolution on the respective energy measurement, $\delta p/p < \delta E_{cal}/E_{cal}$.

The track momentum is also preferred if the object satisfies the following requirements: $E_{cal} < 5$ GeV, $E_{cal}/p < 0.25$, $p < 30$ GeV. These conditions are typically fulfilled by muons, which normally traverse the calorimeter thickness generating an energy deposit of few GeV. As we will discuss in Section 3.3.3, further corrections are applied whenever an EFO is associated to a muon candidate.

3.3.2 Corrections on the calorimeter islands

Detailed studies on data and Monte Carlo simulation have highlighted the need to apply the following corrections on the islands reconstructed in the calorimeter [85]:

- *Energy loss in the inactive material between the interaction point and the calorimeter.* The contribution of the beam pipe, tracking detector and the solenoid to the energy loss varies with the polar angle between 1 and 3 radiation lengths X_0 (see Figure 3.5). Energy losses are more significant for low energy particles and are generally difficult to include in the detector simulation. Additional correction factor are therefore parameterized as function of the polar angle and of the island energy and applied to the calorimetric EFOs.

- *Energy overestimation of low momentum hadrons.* Protons and pions with momentum under 1 GeV loose their energy mainly through ionization without hadronic interactions. For those particles the calorimeter is not compensating, i.e. $e/h \sim 0.6$. This effect implies an overestimation of low energy islands.

3.3.3 EFOs associated to muons

As discussed in the introduction to this Chapter the final state of the events selected by this analysis contains most of the times muons close to decay products of a meson. Since the energy release of a muon traversing the calorimeter is dominated by ionization, the measured energy is not proportional to the momentum of the incoming particle. The inclusion of the calorimetric measurement of the muon four-momentum in the particle jet would generally lead to an underestimation of the jet energy. In this context the use of EFOs instead of the bare

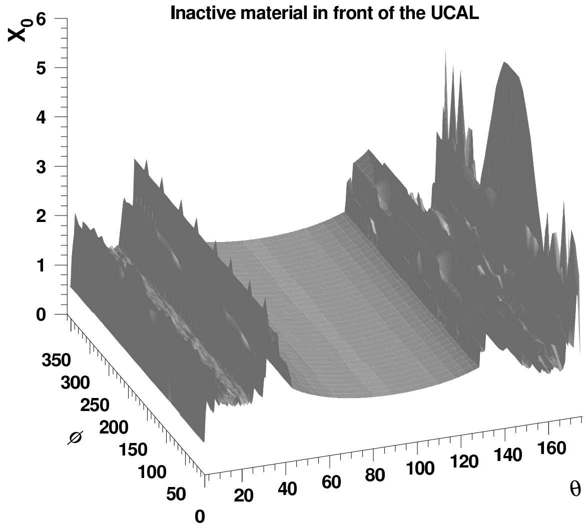

Figure 3.5 Distribution of inactive material between the interaction point and the ZEUS calorimeter as function of the polar and azimuthal angles of a particle produced at the nominal interaction point.

calorimeter cells helps to correctly evaluate the particle momentum using the tracking informations. Nevertheless when the muon EFO is a calorimetric island some additional corrections are required.

As a first step, the energy release of the muon in the calorimeter has been parameterized as a function of its polar angle. The dependence, obtained from a sample of isolated muons, is plotted in Figure 3.6. The average energy loss of the muon is of about 2-3 GeV, a value which is consistent with the difference between the momentum measurement of the CTD and of the muon chambers. For each muon candidate identified in the tracking and muon chambers (see Section 3.5) the corresponding EFO is searched applying the correction according to the following scheme [85]:

- *The corresponding EFO is a track.* No action is needed.
- *The corresponding EFO is a calorimeter island whose energy is compatible with a muon.* If the island energy is between 50% and 150% of the energy release expected from Figure 3.6 the EFO is substituted with the tracking measurement.
- *The corresponding EFO is a calorimeter island whose energy is bigger than the one expected from a muon.* This is case in which the muon is included in a jet. The muon EFO is then splitted in two EFOs: the first is set to the muon momentum measured with the tracking while the second is obtained by the difference of the original island and the energy loss expected by our parameterization. In this way the muon contribution to the jet is correctly reconstructed;
- *Other cases.* If the muon points to an island whose energy is much lower than what expected from the energy parameterization or no island is found the tracking measurement is substituted to the reconstructed EFO.

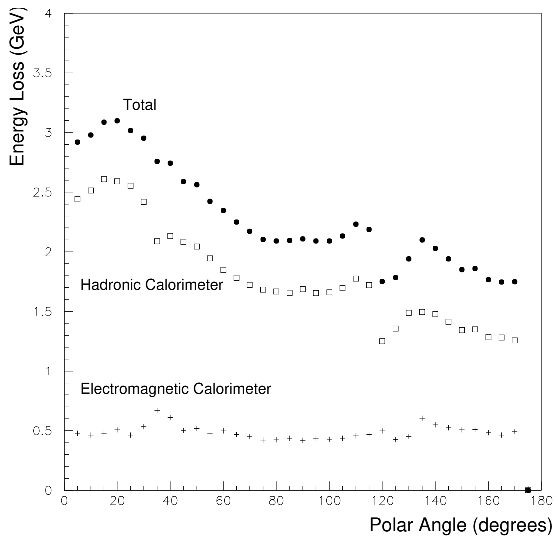

Figure 3.6 Muon energy loss in the ZEUS calorimeter as a function of the polar angle (full circles). The energy fractions released in the electromagnetic calorimeter (crosses) and in the hadronic calorimeter (open squares) is shown [85].

3.4 Jet reconstruction

The aim of the jet reconstruction is to separate soft and hard processes. This is achieved by a cut on the transverse energy in a frame where all the E_T has to be produced by the hard scattering. In deep inelastic scattering one has to deal with the proton remnant and isolate the latter from jets with high transverse momentum. In Section 1.7 we have shown that the Breit frame is the proper frame to perform the jet clustering in deep inelastic scattering.

In this section we will describe the boost and the jet clustering algorithm. In addition, the uncertainty related to the jet energy scale and the correlation of the jet energy to the energy of the incoming particles are discussed.

3.4.1 Boosting to the Breit frame

In order to perform the jet clustering the EFOs must be boosted from the laboratory frame into the Breit frame. As seen in Section 1.7, the boost to the Breit frame is defined in terms of the proton momentum p and the one of the exchanged photon, q. The latter can be calculated using the initial (k) and final (k') electron four-momenta, $q = k - k'$. The vector k' is obtained from the measurements of the energy E' and the angles (θ_e, ϕ_e) of the scattered electron using the following formulae:

$$
\begin{aligned}
k'_0 &= E'_e \\
k'_x &= E'_e \sin\theta_e \cos\phi_e \\
k'_y &= E'_e \sin\theta_e \sin\phi_e \\
k'_z &= E'_e \cos\phi_e
\end{aligned}
$$

Particular care has to be taken to avoid bias to the boost vector introduced by emission of real photons before or after the hard scattering process. A detailed discussion is given in Section 4.2.4.

3.4.2 The longitudinally invariant k_T clustering algorithm

On the basis of the discussion of Section 1.7 the so called k_T-algorithm has been originally introduced by Catani, Dokshitzer and Webber for e^+e^- collisions and later modified by Ellis and Soper to the $p\bar{p}$ case. The algorithm is defined in the Breit frame and consists of an iterative clustering procedure. Starting with the initial list of objects (the EFOs of Section 3.3.1 in our case) belonging to the hadronic final state, each one characterized by a transverse momentum $E_{T,i}$, azimuthal angle ϕ_i and pseudorapidity η_i, the algorithm proceeds as follows:

1. one defines hard scattering scales $d_i = E_{T,i}^2$. For each pair of objects one computes also
$$d_{ij} = \min(E_{T,i}^2, E_{T,j}^2)[(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2] \tag{3.29}$$
2. Find the smallest of all the d_i and d_{ij} and label it d_{min}.
3. If d_{min} is one of the d_{ij} merge the objects i and j into a new object k with:
$$E_{T,k} = E_{T,i} + E_{T,j}, \quad \eta_k = \frac{E_{T,i}\eta_i + E_{T,j}\eta_j}{E_{T,k}}, \quad \phi_k = \frac{E_{T,i}\phi_i + E_{T,j}\phi_j}{E_{T,k}}. \tag{3.30}$$
4. If d_{min} is one of the d_i, the corresponding object cannot be merged.

The procedure continues until there are no more objects to cluster. Once the clustering is completed a set of variables is calculated for each jet in the Breit frame (E_T^{Breit}, the pseudorapidity η^{Breit} and the azimuthal angle ϕ^{Breit}). As discussed, the jet transverse energy measured in the Breit frame represents the scale of the hard scattering and a cut on this observable must be applied in order to select the hard part from the soft part. In this analysis the jets are first preselected with a loose E_T^{Breit} cut of 1.5 GeV. A harder E_T cut is applied in the final event selection. In addition, the polar and azimuthal angles in the laboratory frame are required for each jet in order to restrict the measurement in a region of the detector which is well understood. The inverse of the Lorentz transformation discussed in Section 3.4.1 is therefore applied to the four-momentum of the jets in the Breit frame, which are considered to be massless. In what follows the jet kinematic variables measured in the laboratory frame are labeled as E_T^{Lab}, η^{Lab} and ϕ^{Lab}.

3.4.3 Jet energy correction

In the jet reconstruction procedure described in the previous sections we have introduced the concept of EFOs and discussed the effects of a precise description of the inactive material located between the interaction point and the calorimeter on the determination of the jet energy. It is now necessary to estimate to which degree of precision the reconstructed jet energy is correlated with the energy of the incoming particles. For this purpose, a Monte Carlo simulation of events from b-quark production is used. In what follows, we will refer to jets reconstructed from the incoming particles as *hadron level* jets and to the jet reconstructed using the EFOs as *detector level* jets. The selection of the event sample and the association between hadron and detector level jets proceeds as follows:

- Jets at hadron and detector level are selected requiring $E_T^{Breit} > 3$ GeV and $|\eta^{Lab}| < 2.5$.
- For each pair of jets, each one composed by one jet at detector level and one at hadron level, the distance ΔR in the $\eta - \phi$ plane is calculated

$$\Delta R = \sqrt{(\eta_{det} - \eta_{had})^2 + (\phi_{det} - \eta_{had})^2}. \tag{3.31}$$

  In Eq. 3.31 all variables are measured in the laboratory frame. The pair with the minimum distance is considered a good match if the distance is smaller than 0.2.

The correlation between the true and the measured jet energy is studied by measuring the ratio $E_{T,Had}^{Breit}/E_{T,Det}^{Breit}$. The mean of this distribution is shown in Figure 3.7 as a function of the jet transverse energy measured in the Breit frame and as a function of the jet pseudorapidity. The results show that the energy measured from the reconstructed jets is systematically lower than what is expected. A correction factor $\delta_{Jet} = (7.6 \pm 0.4)\%$ has been extracted performing a fit on the distribution in Figure 3.7 and applied to the reconstructed jets before applying the selection cuts.

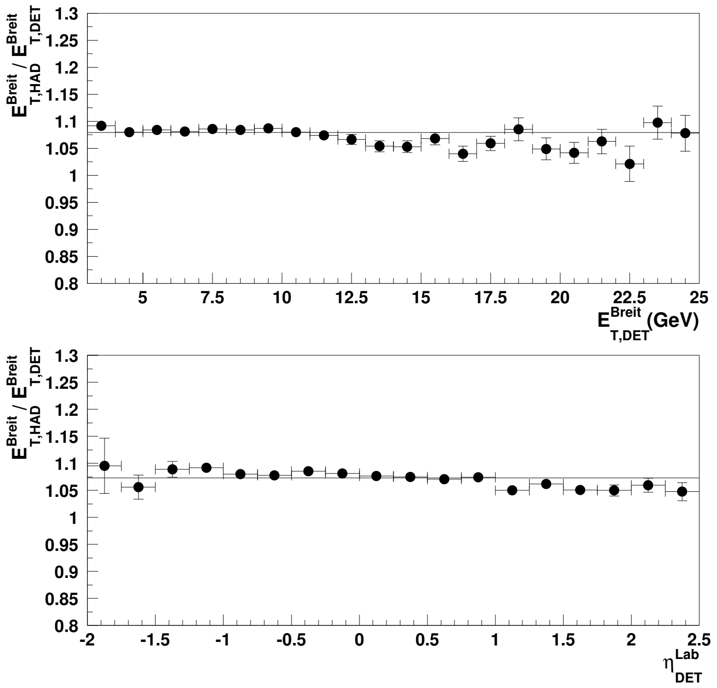

Figure 3.7 Ratio of the jet transverse energy measured at hadron and detector level as a function of the transverse energy measured in the Breit frame (*top*) and of the jet pseudorapidity in the laboratory frame (*bottom*).

3.4.4 Jet energy scale uncertainty

The uncertainty on the absolute energy scale of the calorimeter is an important contribution to the total systematic error, whenever the measurement involves hadronic final states like jets. Inclusive studies on this topic have been performed by the ZEUS Collaboration and can be found in Refs. [86, 87, 88]. As previously discussed, in our measurement the hadronic final state is reconstructed combining the calorimetric information with the tracking one. Therefore, the uncertainty on the energy scale has been evaluated with an independent study.

The following method relies on the fact that for QPM processes the transverse momenta of the leading jet and of the scattered positron are balanced in the laboratory frame. We assume that the transverse momentum carried by the proton remnant is negligible as well as the uncertainty on the measurement of the energy of the scattered positron. In our study the reconstruction methods used for the scattered positron and for the jets are the same discussed respectively in the Sections 3.2 and 3.4.2. The selection of the data and Monte Carlo simulated single-jet samples proceed according to the following steps:

- A sample of neutral current DIS events is selected in the kinematic region of this analysis following the selection cuts described in Chapter 4 both at the online and offline level.
- Successively, a sample of single-jet events is selected in the laboratory frame by applying the following cuts on the leading jet candidate: $E_T^{Lab} > 6$ GeV and $-1 < \eta^{Lab} < 2$. In addition the second leading jet is required to have $E_T^{Lab} < 5$ GeV.

Using the resulting event sample the ratio between the jet and positron transverse energy has been measured (see Figure 3.8). The distributions extracted from the data and the Monte Carlo simulation have been fitted with a Gaussian function. The fit results are also shown in Figure 3.8. The ratio between the centroids of the two Gaussian fits is

$$\frac{< E_{T,Jet}/E_{T,Electron} >_{Data}}{< E_{T,Jet}/E_{T,Electron} >_{MC}} = 1.9\% \qquad (3.32)$$

From such procedure we conclude that an uncertainty in the jet energy scale within 2% must be considered.

3.5 Reconstruction of the semi-leptonic muons

The semi-leptonic muons are reconstructed combining the informations of the tracking detector and of the muon chambers. The muon trajectory is reconstructed in the CTD and then traverses the region of the uranium calorimeter before reaching the inner muon chambers. Since the magnetic field inside the calorimeter is very low compared to the one inside the solenoid, the displacement of the trajectory from a straight line is mainly due to multiple scattering. Under these conditions, an efficient and precise reconstruction can be achieved if tracks upstream and downstream the calorimeter are joined using an accurate simulation of the particle trajectory and taking into account the full error matrices of the single reconstruction steps. For this purpose the muon chambers must be exploited not only as a muon tagger but as a real tracking device.

A specific algorithm (`BREMAT`, [89]) has been designed to associate inner detector tracks to the segments reconstructed in the barrel and rear muon chamber and provide the resulting matching χ^2. We briefly describe the structure in the following sections.

3.5.1 Track selection and extrapolation

Tracks reconstructed in the CTD without vertex refitting are selected according to the following loose requirements:

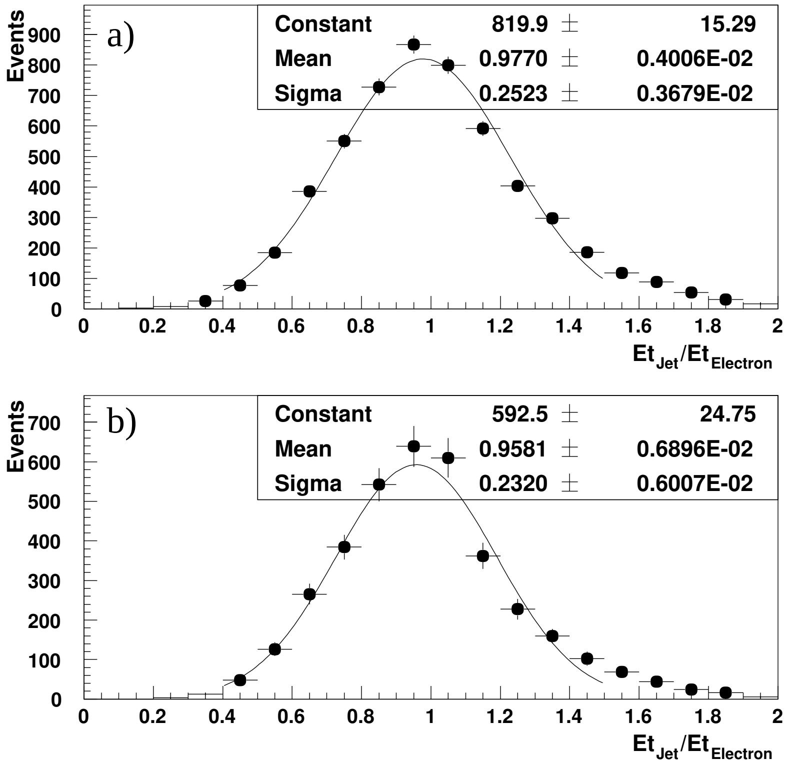

Figure 3.8 Distribution of the ratio $E_{T,Jet}/E_{T,Electron}$ (*full circles*) measured on data (a) and Monte Carlo simulation (b). The distributions are fitted with a Gaussian function (*solid line*). The results of the fit and the relative uncertainties are shown in the upper right corner of each plot.

- Track momentum $p > 1$ GeV;
- Track polar angle $\theta > 20°$;
- track starting from the CTD first superlayer and extending at least to the third superlayer;
- Impact parameter $|D_H| < 10$ cm;
- Vertex z-coordinate $|z_{vtx}| < 75$ cm;
- Good track quality $\chi^2_{track}/n.d.f. < 5$
- Distance Δ between the central point of the segment reconstructed in the muon chambers and the linear extrapolation of the CTD track at the calorimeter entry point, $\Delta < 150$ cm.

Preselected tracks are extrapolated using the `GEANE` software package [90] up to the surface of the inner muon chambers giving a predicted position and relative error matrix in a convenient Cartesian parameterization[1].

[1]The Cartesian frame is specific of each sector of the muon detector, with x and y axes being respectively the coordinates measured by the LST wires and strips and z an axis perpendicular to the chamber and pointing

3.5.2 B/RMUON segment reconstruction and matching

As seen in Section 2.5.2 each muon chamber is composed by two pairs of LST planes, for each event the wires and strip quadruplets addresses are read out. The reconstruction algorithm [91] looks for wire–strip coincidences in all planes and when a coincidence is found in at least one of the internal and external planes a muon has been identified in the chamber. This procedure reduces the amount of random coincidences and permits a first estimation of the particle direction.

Figure 3.9 shows a schematic representation of a particle traversing a muon chambers. On each LST plane of the chamber the charge is collected by the wires (parallel to the page plane) and by the strips. The signal of four consecutive strips is collected in quadruplets.

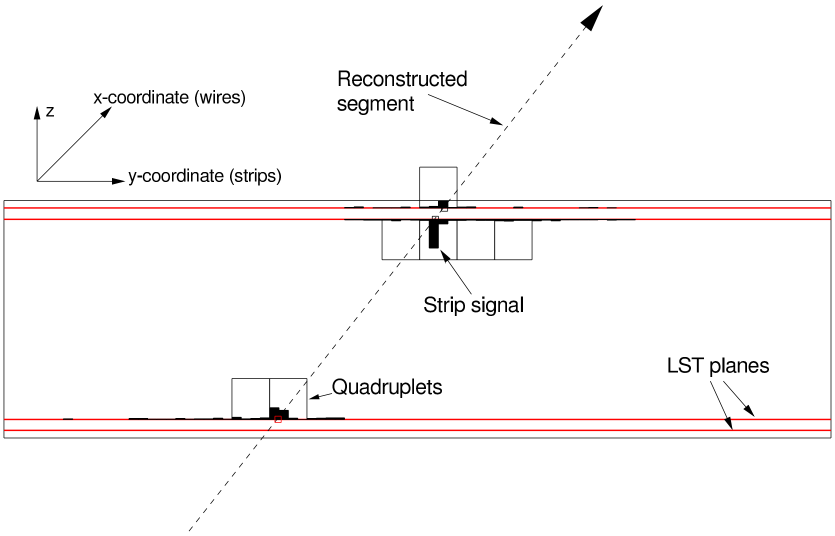

Figure 3.9 Segment reconstruction in a muon chamber. The signal collected on the strips is represented by the black histograms. The open squares represent a group of four consecutive strips, called *quadruplets*. The dashed line is the *segment* which connects the hits reconstructed on the two opposite sides of the chamber.

Naming ξ_i^M and ξ_i^P respectively the values of the measured and predicted coordinates and momenta, the residuals are simply $\delta_i = \xi_i^M - \xi_i^P$. The χ^2 of the matching procedure is therefore given by:

$$\chi^2 = \sum_{i,j}^{N} S_{ij}^{-1}\delta_i\delta_j = \delta S^{-1}\delta^T \tag{3.33}$$

where $N = 5$, the indexes i, j run over the parameter space $(x, y, x', y', Q/p)$ and S^{-1} is the inverse of the covariance matrix of the residuals

$$S_{ij} = \sum_{k.l}^{N} T_{ik}(\sigma^2_{track})_{kl}T^T_{lj} + (\sigma^2_{scattering})_{ij} + (\sigma^2_{BRMU})_{ij}. \tag{3.34}$$

Here T is the so called *transport matrix*, between the start and the stop of the swim, σ^2_{track} is the error matrix of the CTD track, $\sigma^2_{scattering}$ the error matrix of the multiple scattering

outside the detector. The direction is expressed by the slope $x' = dx/dz$, $y' = dy/dz$.

and energy loss and σ^2_{BRMU} the error matrix of the B/RMUON segment. The first two terms are calculated for each step during the GEANE extrapolation using the detailed geometry of the ZEUS detector and the magnetic field map.

As a check of the matching algorithm the *pull* distributions can be measured for each coordinate. The pull is defined as the ratio of the residual and its error, which corresponds to a diagonal element of the covariance matrix S:

$$Pull_i = \frac{\delta_i}{\sqrt{S_{ii}}}. \tag{3.35}$$

If the contributions to the error matrix written in Eq. 3.34 are correctly estimated the distribution of each variable $Pull_i$ is Gaussian with a unitary width. Pull distributions of the data sample are shown in Figure 3.10, while the distribution widths are summarized in Table 3.1 for data and different Monte Carlo simulations, detailed in Section 5.2. In these measurements only the matching candidate with the best χ^2 is considered.

Variable	Data	Light MC	Beauty MC
x	1.20±0.01	1.18±0.01	1.04±0.01
y	1.25±0.01	1.24±0.01	1.10±0.01
x'	1.15±0.01	1.20±0.01	1.03±0.01
y'	1.18±0.01	1.18±0.01	1.02±0.01
Q/P	1.49±0.01	1.56±0.02	1.29±0.02

Table 3.1 Widths of the pull distributions measured on the data and Monte Carlo samples

All distributions are well centered at zero. Moreover, since no particular selection has been applied to the data the measured pull values are nicely described by the Monte Carlo simulation of light quarks processes which are expected to dominate in the data. The values expected from the simulation of beauty decays into muons are generally closer to one since they are dominated by real muons.

In this analysis the quality of the matching is given in terms of χ^2 probability $P_{5d.o.f.}(\chi^2)$.

3.5.3 Efficiency of the muon reconstruction

The muon reconstruction efficiency has been estimated using elastic dimuon events and compared with the MC simulation. Elastic events were selected requiring two tracks and two calorimeter islands and less than 1 GeV measured in the rest of the calorimeter. The events were triggered by one of the internal muon chambers. The ratio

$$\epsilon = \frac{N_{match}}{N_{tracks}} \tag{3.36}$$

was computed, were N_{match} indicates the number of muon candidates reconstructed with the BREMAT algorithm and satisfying $P_{5d.o.f.}(\chi^2) > 5$ %, while N_{tracks} is the total number of tracks. The ratio ϵ can be considered as an estimate of the efficiency of the whole muon reconstruction chain, which includes the detection efficiency of the muon chambers, the reconstruction and matching software described in Section 3.5.2.

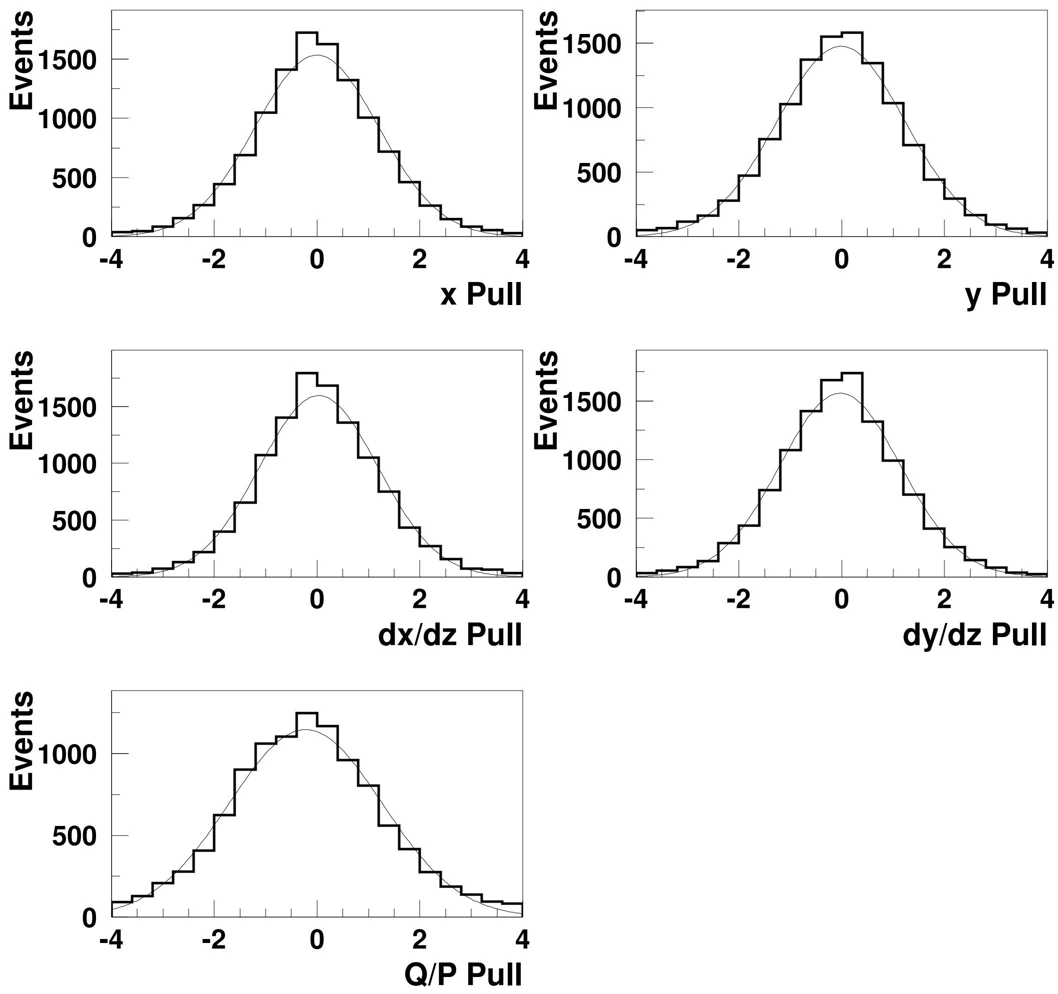

Figure 3.10 Pull distributions as a function of the BREMAT coordinates $(x, y, x', y', Q/p)$, measured on the data sample (*histogram*). A Gaussian fit (*curves*) is shown.

The measurement is compared with a simulation including elastic production of J/ψ mesons [92] and Bethe-Heitler production of muon pairs [93]. The relative normalization of the two sample is obtained by fitting the invariant mass spectrum of the muon pair to the measured one.

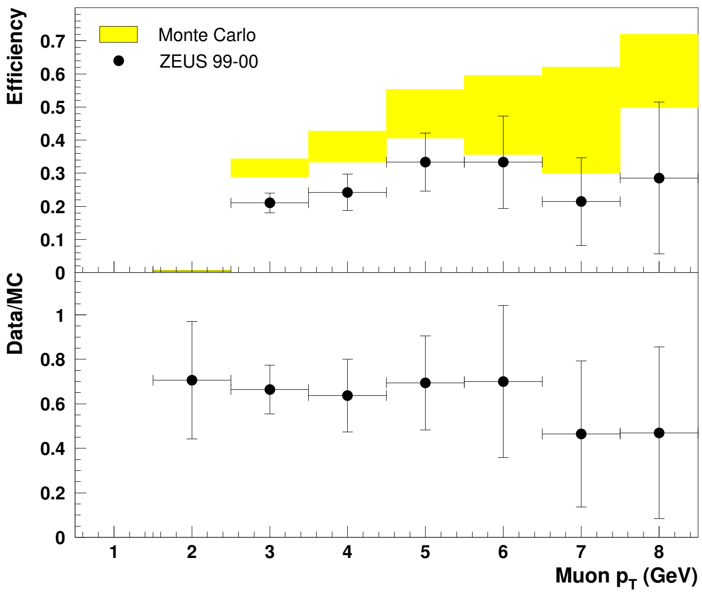

Figure 3.11 Comparison between measured and simulated muon reconstruction efficiency as a function of the muon transverse momentum.

A comparison between measurement and simulation is shown in Figure 3.11 as a function of the muon transverse momentum. The simulation overshoots the measured efficiency since it does not implement a complete list of dead LST tubes in the barrel and rear muon chambers. In addition, during the data taking the bias voltage applied to the LST tubes is lower than the one corresponding to the efficiency plateau. At lower voltages a longer detector lifetime is expected.

The ratio between data and Monte Carlo simulation, also shown in Figure 3.11, does not show any significant dependence on the muon momentum. Moreover, no significant dependence is observed as a function of the muon polar angle. Therefore, a correction factor can be extracted by averaging the ratio over the measured momentum range. The values extracted for the barrel and rear sections of the muon detector, summarized in Table 3.2 are used in the calculation of the cross section (see Section 5.4).

Detector region	Correction factor
Barrel	0.65± 0.08
Rear	0.47± 0.02

Table 3.2 Ratios between the measured and simulated muon reconstruction efficiencies.

4 Event selection

The selection of the final event sample consists of two main steps. Events are first preselected by the three levels of the ZEUS trigger. As we will see, a first rough identification of the positron and muon candidate is achieved already at this level. Successively, an offline selection is applied, which benefits from the higher precision after the final event reconstruction. In this chapter the different selection steps will be described together with the data samples used.

4.1 Trigger preselection

4.1.1 First and second level trigger

Already at the FLT it is possible to select candidates of neutral current DIS interactions by checking energy thresholds in the calorimeter. Different requirements have to be fulfilled depending on the calorimeter section:

- If the electron is scattered into the RCAL, the event is accepted if the energy in the electromagnetic section is greater than 3.75 GeV. The event is also accepted if the energy deposited outside the inner calorimeter ring (60×60 cm) is greater than 3.4 GeV.

- If the electron is scattered with $36.7° \leq \theta_e \leq 129.1°$ it is detected by the BCAL section. In this case events are accepted if the energy electromagnetic section is larger than 4.78 GeV. Since the electron always traverses more than 3 superlayers of the CTD a track is required in addition. This requirement reduces the background from beam halo muons characterized by energy deposits in the BCAL without a track in the CTD.

- The trigger strategy for electrons scattered in the FCAL is based on transverse energy requirements. This allows to achieve a high trigger efficiency in a section which has to cope with high background due to large energy deposits in the forward region. The latter are caused by the proton remnant or the color flow between the proton and particles produced in the hard scattering. This section corresponds to events at very high Q^2.

Additional background rejection is performed at the second level trigger (GSLT). At this stage the full calorimetric information is available and more sophisticated requirements can be applied to select DIS processes and reject halo or cosmic muons. The background from photo-production events is reduced exploiting energy and momentum conservation of the scattering process. The quantity $\delta = E - p_z$ is introduced, defined as the difference of the total energy and the longitudinal component of the momentum. The value of δ_i for the incoming particles is given by

$$\begin{aligned}\delta_i &= \delta_{proton} + \delta_{electron} \\ &= 920\text{ GeV} + 27.5\text{ GeV} - (920\text{ GeV} - 27.5\text{ GeV}) \\ &= 55\text{ GeV}.\end{aligned}$$

The value measured on the particles detected in the event, δ_f, is defined as

$$\delta_f = \sum_{i=1}^{N} E_i(1-\cos\theta_i) \tag{4.1}$$

where the index i runs over the N calorimeter cells, each one characterized by a measured energy E_i and a polar angle θ_i. It is clear from Eq. 4.1 that the particles escaping from the rear beam hole, i.e. at high polar angles θ_i, can substantially reduce the measured δ_f. This is the case for electrons scattered at very high polar angles and for on-shell photons radiated before the hard interaction, whose trajectory is almost parallel to the direction of the incoming electron (see Section 4.2.4). The SLT rejects events with low values of δ_f but accepts events with initial photon radiation applying the cut:

$$\delta_f + 2E_\gamma^{LUMI} > 29\text{ GeV}, \tag{4.2}$$

where E_γ^{LUMI} is the energy of the radiated photon measured by the LUMI detector.

Additional timing cuts are applied in order to further reduce other sources of background. Timing measurement is a fast and precise ($\sigma_t/t \sim 1$ ns) way to discriminate *ep* collisions from cosmic events or beam–gas interactions. The timing of each calorimeter section is obtained by averaging the timing of the single cells with a weight equal to the deposited energy. While *ep* collision are characterized by timing close to zero, beam-gas interactions produced by the proton beam upstream of RCAL generally have a RCAL timing peaked at -10 ns [94]. A cut at -8 ns is therefore applied to reject those events.

The time distribution of cosmic events is flat. On the other hand, cosmic rays traversing the detector from the upper towards the lower part are removed applying a cut on the timing difference between upper and lower half of the BCAL.

4.1.2 Third level trigger

TLT event selection At the TLT the event kinematic is fully established and a first classification of the final state is performed. The event vertex is determined using the CTD tracks and the variable δ_f is recalculated with higher precision. The cut in Eq. 4.2 is increased to 30 GeV.

In this analysis a dedicated TLT algorithm has been used in order to select events with a muon and an electron candidate (TLT slot `DIS27`). Electron candidates are identified by four different algorithms. Events with at least one candidate with energy greater than 7 GeV are accepted. The algorithm implements a fast muon finder based on the matching of CTD tracks and BRMUON segments. It requires a reconstructed vertex, at least one segment in the internal muon chambers and at least one track in the CTD satisfying the following requirements:

- $p_{track} > 1$ GeV;
- $\theta_{track} < 20°$;
- the track must traverse at least the first three superlayers of the CTD;
- the distance of closest approach with respect to the beam line must be smaller than 10 cm;
- the distance between the point of closest approach and the primary vertex along the z-axis must be smaller than 10 cm.

The last two requirements reject cosmic muons traversing the detector far away from the reconstructed primary vertex. The TLT reconstruction software calculates the unitary vector $\vec{v}$ corresponding to the track direction at the CTD exit point. The track is extrapolated from the CTD exit point to the muon chambers segment with a straight line $\vec{v}'$. The matching between the muon segment and the CTD track is obtained by requiring the condition $||\vec{v}'_\perp|| < 200$ cm, where $\vec{v}'_\perp$ is the orthogonal component of $\vec{v}'$ respect to $\vec{v}$.

Efficiency of the TLT slot DIS27 The efficiency of the TLT slot `DIS27` was measured on data and Monte Carlo simulation using two independent DIS trigger algorithms, the `DIS1` and `DIS3` slots.

The `DIS1` trigger slot is designed to accept DIS events at low Q^2. Events are accepted by this slot if at least one electron candidate is reconstructed with energy greater than 4 GeV and $30 < \delta_f < 100$ GeV. In addition, events are accepted only if the reconstructed impact position of the scattered electron is outside the box defined by $|x_e| < 12$ cm and $|y_e| < 6$ cm. The impact position is reconstructed using the SRTD detector when available otherwise the calorimeter information is used. This cut removes electron candidates reconstructed on the inner edge of the RCAL beampipe hole.

Events at medium Q^2 are triggered by the `DIS3` slot. The cut on the electron energy and on δ_f are the same as for the slot `DIS1`. A radius cut at $(x_e^2 + y_e^2) > 35$ cm is applied on the electron impact position instead of the box shaped cut.

The event sample was selected requiring good muon and electron candidates in addition to the trigger slots. The muon candidate was reconstructed using the `BREMAT` algorithm (see Section 3.5) and requiring a matching probability $\mathcal{P}_{5d.o.f.}(\chi^2) > 1\%$. The electron candidate was reconstructed using the algorithm described in Section 3.2 and applying a cut $\mathcal{P}_{electron} > 90\%$. The efficiency ϵ was defined as

$$\epsilon = \frac{[DIS1 \vee DIS3] \wedge MUON \wedge DIS27 \wedge ELECTRON}{[DIS1 \vee DIS3] \wedge MUON \wedge ELECTRON} \tag{4.3}$$

where "MUON" and "ELECTRON" respectively stand for a good muon and electron candidate. The efficiency measured on the data sample is $\epsilon_{Data} = (93.4 \pm 0.4)\%$ while the simulated result is $\epsilon_{MC} = (96.8 \pm 0.2)\%$. The latter has been obtained using the sample B1 of open beauty events described in Section 5.2. The corresponding $\epsilon_{Data}/\epsilon_{MC}$ ratio is 1.037. A comparison of the measured and simulated trigger efficiency as a function of the muon momentum and of the photon virtuality is shown in Figure 4.1. No significant dependency of the ratio $\epsilon_{Data}/\epsilon_{MC}$ is observed as function of these variables. The measured discrepancy between data and Monte Carlo is taken in account when calculating the selection acceptance (see Section 5.4).

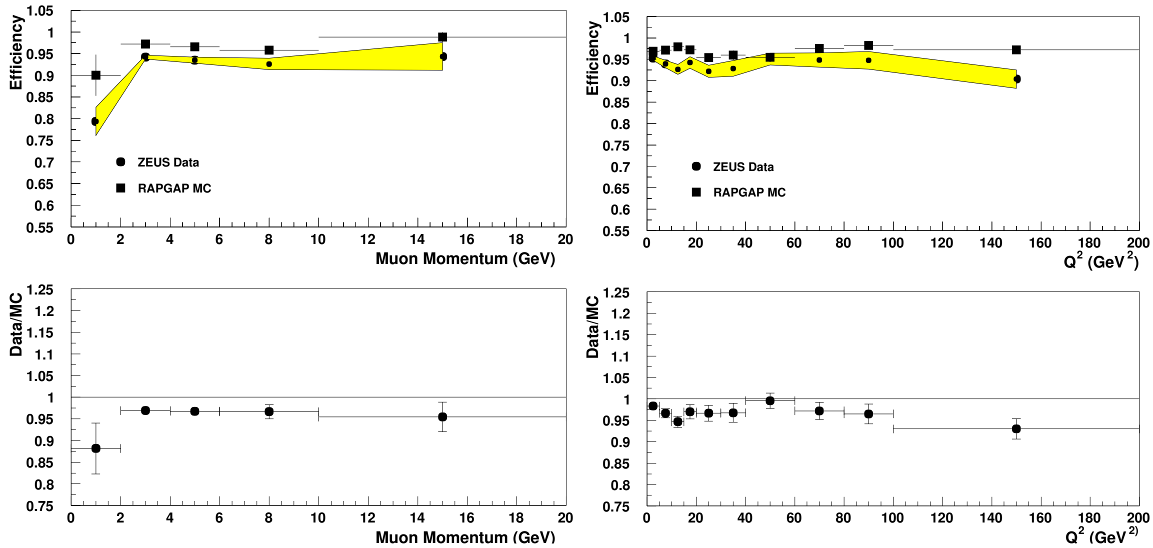

Figure 4.1 Measured and simulated efficiency of the TLT slot DIS27 as a function of the muon momentum (*upper left plot*) and of Q^2 (*upper right plot*). In the lower plots, the ratio $\epsilon_{Data}/\epsilon_{MC}$ is shown.

4.1.3 The data samples

The e^+p data used in this analysis were recorded in the years 1999 and 2000, covering the run range 33125 to 37715. Some runs were excluded either by the standard ZEUS data quality routine or by the requirement of good performances of the BRMUON detector. The e^+p data set corresponds to an integrated luminosity of 60.4 ± 1.4 pb^{-1} [95].

4.2 Offline selection

The event sample selected by the trigger still contains a fraction of background processes which are removed by applying tighter offline cuts. In addition, a precise reconstruction of the final state is only possible at the offline level. In the next section we will describe the cuts applied to select the inclusive sample of neutral current DIS events. In order to enrich the event sample with events from decays of b-quarks further cuts are applied to the muon and jet candidates.

4.2.1 Selection of a neutral current DIS event sample

The following cuts are applied to select a sample of neutral current DIS events:

- A vertex must be reconstructed using CTD tracks. At least two tracks have to be associated to the vertex and the reduced χ^2 of the fit must be smaller than 10. Such a vertex must be close to the nominal interaction point, i.e. the z coordinate $|z_{vtx}| < 50$ cm.

- Photoproduction events are a serious source of background. Due to the low virtuality of the exchanged photon, in those events the electron is scattered at large values of θ_e and it escapes the detector through the rear beam hole. As discussed in Section 4.1.1, if the scattered electron is not detected the quantity δ_f is not conserved. At the level of the offline selection a more precise estimation of δ_f is obtained using the EFOs and the electron candidate:

$$\begin{aligned}\delta &= \delta_{had} + \delta_{electron} & (4.4)\\ &= \sum_{i=1}^{N} E_i(1-\cos\theta_i) + E'_e(1-\cos\theta_e). & (4.5)\end{aligned}$$

In Eq. 4.5 the sum runs over the N reconstructed EFOs. To suppress photoproduction events δ_f is required to be

$$40 < \delta_f < 65 \text{ GeV}. \quad (4.6)$$

The upper cut is introduced to remove residual cosmic ray background.

- To ensure high purity the electron candidate is required to have an energy of at least 10 GeV. In addition a fiducial cut is applied to the electron position.
- When the scattered electron hits the RCAL surface too close to the detector edge the electromagnetic shower is not fully contained in the active volume of the calorimeter. As a consequence, the measured energy of these electrons is systematically lower than the true one. Dedicated studies have been performed to measure the energy leakage of electrons entering the RCAL near the detector edge [81]. Using kinematic peak events[1] it was shown that the response of the detector is uniform up to 2 cm from the detector edge. As discussed is Section 3.1, the energy of the scattered electron is used in the calculation of several kinematic variables. In addition, the electron energy is used to boost the hadronic final state from the laboratory frame into the Breit frame (see Section 3.4.1). Events are therefore rejected if the position of the scattered electron at the RCAL surface is contained in a box defined by $|x_e| < 12$ cm and $|y_e| < 6$ cm (see Figure 4.2).
- If the SRTD position is not available the calorimeter information must be used. In this case, if the scattered electron is reconstructed in the RCAL towers positioned over and under the beam pipe hole, hits with $|y_e| < 10$ cm are rejected. For these hits the y-position is not correctly reconstructed by the calorimeter and it is always equal to the middle point of the cell.

4.2.2 Selection of the kinematic region

After applying the preselection cuts described above the kinematic reconstruction is performed. The electron and the Jacquet-Blondel methods discussed in Section 3.1 are used. The kinematic region is defined by the following cuts:

$$\begin{aligned} Q_e^2 &> 2 \text{ GeV}^2\\ y_{JB} &> 0.05\\ y_e &< 0.7.\end{aligned}$$

Figure 4.3 shows the distribution of the selected events in the kinematic plane.

[1]In these events the energy of the scattered electron is equal to the incoming energy $E_e = 27.5$ GeV.

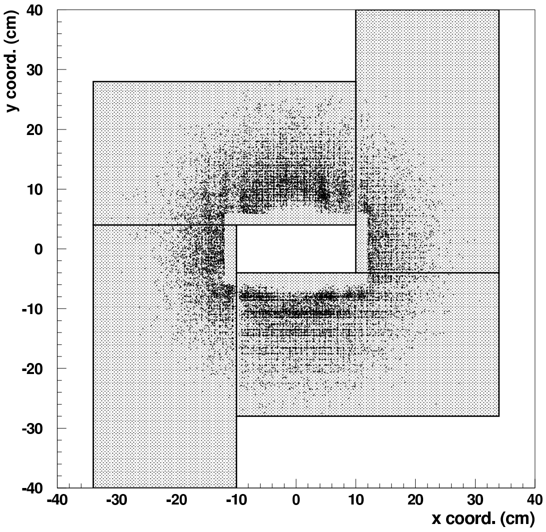

Figure 4.2 Hits of DIS scattered electrons for $Q^2 < 10$ GeV and reconstructed in the SRTD detector. The hits contained in a box of 24×12 cm around the beampipe hole are removed to allow the full containment of the electromagnetic shower in the calorimeter (*see text*).

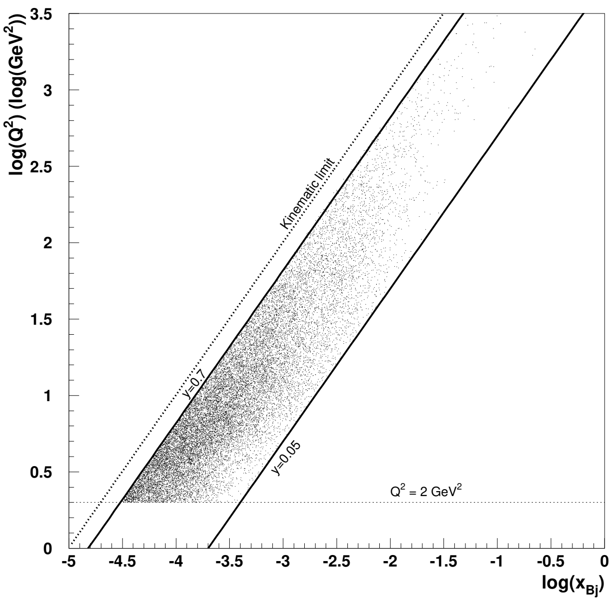

Figure 4.3 Distribution of events in the kinematic plane after the selection cuts described in the Sections 4.2.1-2.

After these cuts the sample contains 182306 events. The selection cuts defined above constrain well the physics of the data sample. To check whether unwanted sources of background are still present in the sample, a comparison with the Monte Carlo simulation is performed. Figure 4.4 shows a comparison between the measured distributions of the most relevant kinematic variables and the Monte Carlo simulation. Since no particular selection is applied at this stage to enrich the data sample with events from heavy quark decays, the comparison is performed with a Monte Carlo sample only containing u, d and s quarks in the final state. The contribution from charm and bottom at this stage of the selection is mainly due to the requirement of a muon candidate imposed by the third trigger level.

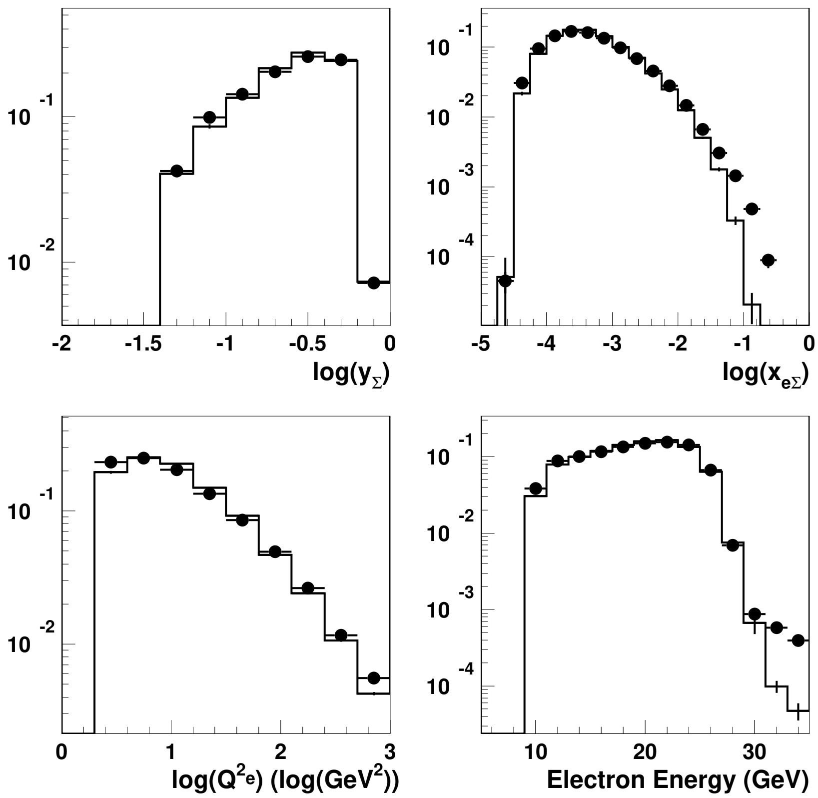

Figure 4.4 Comparison of the measured (full dots) and simulated (histograms) distributions of the DIS kinematic variables after the selection cuts described in Section 4.2.1. From the top left plot: (a) Inelasticity y_Σ, (b) Bjorken variable $x_{e\Sigma}$, (c) Virtuality of the exchanged photon Q^2_e, (d) Energy of the scattered electron.

The distributions are generally well described by the simulation, although the description at high electron energies and high values of x is not optimal.

4.2.3 Muon Selection

The muon candidates are selected according to χ^2 probability of the matching algorithm. The χ^2 probability distribution shows a peak at zero (see Figure 4.5), caused by candidates in which the measured residuals are much larger than the estimated total uncertainty. Candidates with a bad matching are therefore removed by applying the cut:

$$\mathcal{P}(\chi^2) > 1\% \quad \rightarrow \quad \chi^2 < 15. \tag{4.7}$$

The muon reconstruction procedure described in Section 3.5 implies already some kinematic cuts. For example, the requirement of a segment reconstructed in both internal and external chambers implies a minimal momentum of about 1 GeV. Since the muon reconstruction efficiency is very small for momenta below 2 GeV (see Section 3.5.3), only muon candidates with momentum greater than 2 GeV were selected.

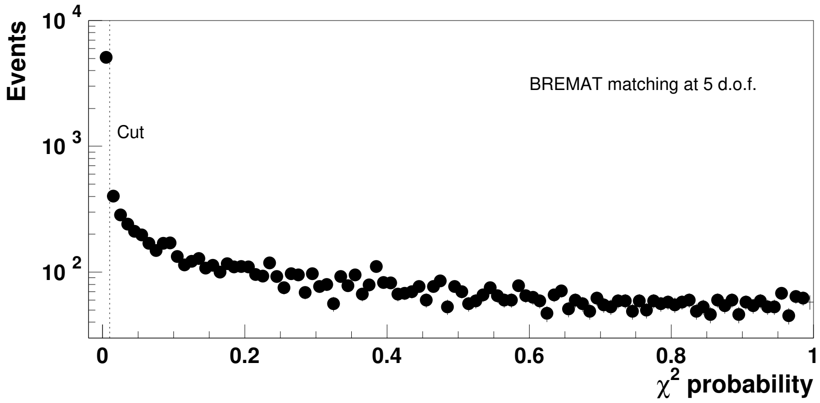

Figure 4.5 Distribution of χ^2 probability of the `BREMAT` muon finding algorithm measured on the data sample.

In order to require a good quality of the track measured in the CTD and extrapolated to the muon chambers, the measurement is restricted to the angular range

$$30^\circ < \theta < 160^\circ. \tag{4.8}$$

In this region tracks are measured by at least three superlayers of the CTD.

4.2.4 Jet Selection

As discussed in Section 1.7, the main reason to perform analyses in the Breit frame is the sensitivity to QCD processes like production of jets at high E_T with respect to the axis given by the photon-proton direction. If this axis is badly reconstructed the energy flow receives a transverse component which spoils the measurement of the jet transverse energy. The topology of different types of events in the Breit frame is sketched in Figure 4.6. The diagrams on the left side of the figure represent the QPM (*top*) and BGF processes (*bottom*). As we have seen, in the

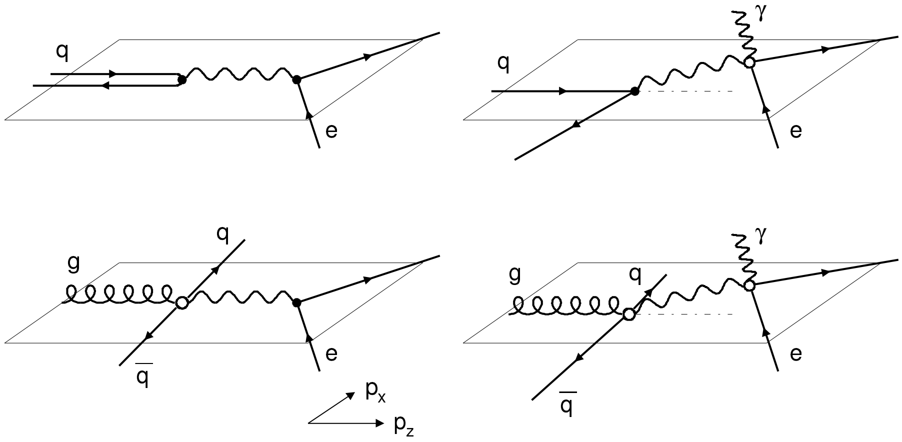

Figure 4.6 Diagrams of the QPM and BGF processes in the Breit frame (*top and bottom left*). On the right side, the effect of photon radiation on the Breit frame reconstruction is shown. If the boost vector is reconstructed using the four-momentum of the scattered positron the virtual photon does not lie on the correct axis (*dot-dashed line*).

Breit frame in a QPM process the quark is backscattered without any transverse momentum after a purely electromagnetic interaction. On the contrary, BGF processes produce a quark-antiquark pair with finite transverse momenta. If no gluon is radiated the pair is balanced in transverse momentum.

QED corrections as photon radiation may lead to a substantial change of this picture. On the right hand side of Figure 4.6 the QPM and BGF processes are displayed with an additional photon radiated collinearly to the initial state positron. If the boost vector is reconstructed from the scattered positron the virtual photon does not lie on the correct axis (represented by the dot-dashed line) and the particles of the hadronic final state are no longer balanced in transverse momentum. As a consequence, if the error in the reconstruction of the boost vector is large enough, topologies as in the upper right diagram in Figure 4.6 may be misidentified as single jets at high E_T.

In our framework, events of this kind can be rejected whenever the radiated photon is detected in the rear calorimeter. The corresponding electromagnetic cluster is most of the times reconstructed as a hard jet in the Breit frame. Therefore, events containing any jet with $E_T^{Breit} > 5$ GeV and $\eta^{Lab} < -2$ have been rejected in this analysis.

The discussion above applies also to the case in which the photon is radiated after the hard interaction. Also in those events, the photon is generally reconstructed as a hard jet in the Breit frame which is very close to the scattered positron in the $\eta - \phi$ plane. For this reason an event is rejected if it contains a jet with $E_T^{Breit} > 8$ GeV which is closer than 0.5 units in the $\eta - \phi$ plane of the laboratory frame to the electron. After the QED radiation cuts the jet energy correction discussed in Section 3.4.3 is applied.

In order to establish the optimal jet selection cuts the following arguments have been considered

- *Separation from soft QCD process:* As discussed in Section 1.7, the appearance of jets of jets with non-zero transverse energy in the Breit frame is a signature of a QCD scattering process. The transverse energy E_T^{Breit}, therefore, establishes the hardness of the process and it is used to separate it from underlying soft interactions.

- *Enhancement of the beauty/charm fraction*: Figure 4.7 shows the distribution of transverse energy in the Breit frame of the leading jet for beauty and charm quarks in the pseudorapidity range $-2 < \eta^{Lab} < 2.5$. The distribution of charm jets peaks at lower values respect to the beauty jets due to the lower charm mass. Therefore, it is possible to enhance the fraction of events from beauty decays through a sensible choice of the jet E_T cut.

- *Maximization of the statistics in the final sample*

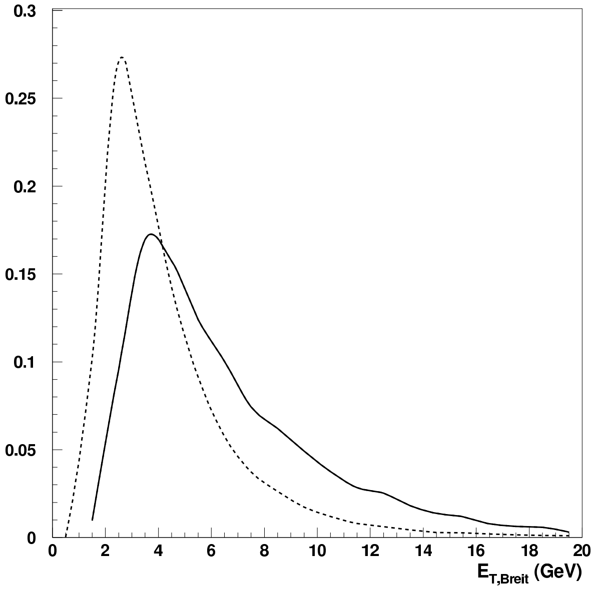

Figure 4.7 Leading jet transverse energy in the Breit frame for charm (*dashed line*) and beauty (*solid line*) quarks measured in the pseudorapidity range $-2 < \eta^{Lab} < 2.5$. The histograms are normalized to the same number of events.

On the basis of the considerations above events in the final sample were selected with **at least one jet** in the Breit frame satisfying the following requirements:

$$E_T^{Breit} > 6 \text{ GeV} \tag{4.9}$$

$$-2 < \eta^{Lab} < 2.5. \tag{4.10}$$

An alternative cut scenario consists on the requirements of two jets reconstructed in the *laboratory frame* with transverse energy above a certain value and in a defined region of the detector. This approach was applied by the H1 Collaboration in the measurement of b-quark production in DIS [54]. We have observed that, using the cuts Eqs. 4.9 and 4.10, a higher statistics can be obtained in the final sample. For instance, requiring at least two jets reconstructed in the laboratory frame with $E_T^{Lab} > 5$ GeV and $-2 < \eta^{Lab} < 2.5$ the final sample would contains $\sim 40\%$ less events and a similar fraction of events from b-decays. The reason is related to the differences in the jet selection between laboratory and Breit frame analyses. In the former case, in order to reduce the background contribution from QPM processes one usually requires at least two hard jets in the event. On the other hand, jet reconstruction can only be performed in regions of the detector which are well understood and a pseudorapity cut must be applied on both jets. As discussed in Section 1.7, QPM processes do not produce events with hard jets in the Breit frame, thus, they are already rejected requiring only one hard jet.

4.2.5 Summary of the offline selection cuts

In this Chapter we have described the selection cuts applied in this analysis. The main cuts are summarized in the Table 4.1. The selected events are characterized by the presence of the scattered positron, a muon candidate traversing the inner and outer barrel muon chambers and at least one hard jet reconstructed in the Breit frame. The event sample after all cuts contains 836 events.

In the next Chapter we will discuss how the fraction of events from b-decays is extracted and we will present the results of the cross section measurement. As we will show, the cross section is defined in a kinematic region corresponding to the measured phase space of the semi-leptonic muon and of the b-quark jet.

Sample	Cut
DIS neutral current	$\lvert z_{vtx} \rvert < 50$ cm $40 < \delta_f < 65$ GeV $E'_e > 10$ GeV box cut 24×12 cm
Kinematic region	$y_{JB} > 0.05$, $y_e < 0.7$ $Q^2_e > 2$ GeV^2
Muon cuts	$\mathcal{P}(\chi^2) > 1\%$ $p_\mu > 2$ GeV $30^\circ < \theta_\mu < 160^\circ$
Jet cuts	at least one jet with: $E_T^{Breit} > 6$ GeV $-2 < \eta^{Lab} < 2.5$

Table 4.1 Summary of the main offline selection cuts

5 Cross section measurement

In this chapter the results of the measurement of the total and differential cross sections are presented. The background sources are described and the signal extraction technique is discussed. In the last sections a detailed description of the measurement uncertainties will be given.

5.1 Background sources

The extraction of the signal from b-decays requires the correct evaluation of the background processes which survive the selection cuts. An exhaustive simulation of all those processes is crucial for the application of the statistical method which we will discuss in the following sections.

5.1.1 Fake muons

Different phenomena may produce final states with similar characteristics of the process under study. In particular, the signal reconstructed in the muon chambers can be generated from light hadrons like kaons and pions. We distinguish two kinds of processes which are usually called *sail-through* and *punch-through*:

- *Sail-through*: The process refers to hadrons traversing the material in front of the muon chambers without interacting. The probability of this process is given by

$$\mathcal{P}(d) = e^{-d/\lambda} \tag{5.1}$$

  where λ is the interaction length and d is the traversed distance. At ZEUS, the material in front of the barrel and rear muon detectors varies between 4λ (RCAL) and 5λ (BCAL). This implies a probability of $\sim 0.7 \div 1.8$ % for a hadron to reach the inner muon chambers without interacting. Muons reaching also the outer muon chambers have to traverse an additional amount of material due to the BAC detector and the iron yoke, corresponding to $\sim 5\lambda$. We should also notice that more hadrons can reach the muon chambers due to the imperfect hermeticity of the calorimeter, e.g. due to the cracks between cells.

- *Punch-through*: Hadronic particles escaping the calorimeter can traverse the inner and outer muon chambers. This contribution is partially removed by the muon reconstruction algorithm and the cut based on the quality of the matching between tracks and muon segments.

The contribution of fake muons to the muon spectrum can be factorized by the total flux f_h of hadrons and the probability of misidentification $\mathcal{P}_h^{\mu}$, representing the probability that a hadron h is detected as a muon in the muon chambers. This contribution can be estimated by generating a large amount of light hadrons using Monte Carlo simulations. The method is valid as long as the misidentification probability $\mathcal{P}_h^\mu$ and the flux f_h obtained from the MC simulation agree with those extracted from the data. In particular, some model dependence could be expected in the quantity f_h since the particle composition of all tracks is mainly determined by the hadronization model used. On the other hand, an imperfect simulation of the material between the interaction point and the muon chambers can lead to an incorrect probability $\mathcal{P}_h^\mu$. As will discuss in the next Section, the hadronization model used in this analysis to simulate the contribution of light hadrons is the Lund string model, as implemented in the program JETSET. The program has been tuned to describe the data properly and measurements at *ep* and e^+e^- collisions have shown that its description is indeed good [96].

The misidentification probability $\mathcal{P}_h^\mu$ is directly measurable on the data for pions and kaons. A direct comparison of the misidentification probability $\mathcal{P}_\pi^\mu$ between data and MC simulation has been performed in the case of pions [97]. A clean sample of pions from the decay $K_s^0 \rightarrow \pi^+\pi^-$ was selected both in the data and the MC simulation using a K_s^0 finding algorithm. Among the tracks having an invariant mass $487 < M_{\pi\pi} < 507$ MeV muon candidates were selected using the BREMAT algorithm. The matching was performed between tracks in the CTD and segments reconstructed in the inner muon chambers, applying a quality cut on the matching, $\mathcal{P}_{4d.o.f.}(\chi^2) > 1\%$. The pion misidentification probability was defined as

$$\mathcal{P}_\pi^\mu = \frac{\#\text{ tracks with } \mu \text{ identification}}{\#\text{ all tracks}} \tag{5.2}$$

The pion misidentification probability measured on data and MC simulation as function of the track momentum and polar angle is shown in Figure 5.1. The agreement between the measured and simulated probabilities is generally satisfactory. Some excess of data at low momenta and high polar angle is observed.

5.1.2 In-flight hadron decays

The detected muons can also be generated by background processes as *in-flight decays of pions and kaons*. Charged pions are the most frequent hadrons created in *ep* collisions at HERA. They can decay into muons through the reaction $\pi^\pm \rightarrow \mu^\pm\nu_\mu$ with a life-time of $\tau = 2.6\cdot 10^{-8}$ s, corresponding to a decay length of $c\tau = 7.8$ m. Charged kaons are also frequently produced at HERA. They decay mainly into $K^\pm \rightarrow \mu^\pm\nu_\mu$ with $\tau = 1.2\cdot 10^{-8}$ s corresponding to $c\tau = 3.7$ m. The probability for an unstable particle to decay within a distance d is given by

$$\mathcal{P}(d) = 1 - e^{-d/\beta\gamma c\tau} \tag{5.3}$$

where β is the particle velocity in units of c and γ is the ratio of its energy and mass. The decay muons are mostly characterized by low momenta and, therefore, partly rejected by the momentum cut, $p_\mu > 2$ GeV.

5.1.3 Decay of charmed hadrons

Muons are also produced in the semi-leptonic decay of charmed hadrons. These decays produce events topologically close to the ones under study and with a cross section almost two orders of magnitude higher. On the other hand, it is possible to define observables to distinguish

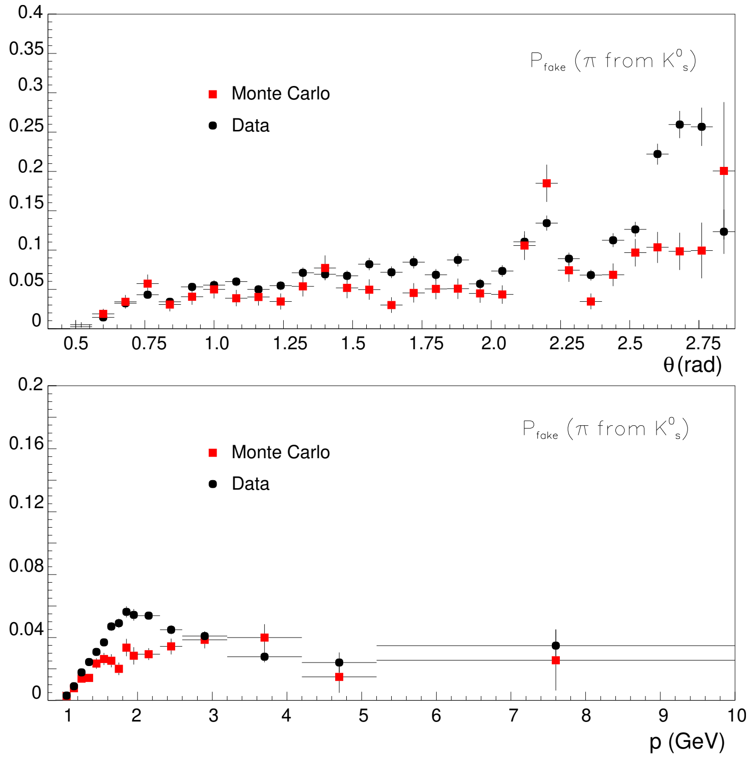

Figure 5.1 Probability $\mathcal{P}_\pi^\mu$ to misidentify a pion as muon plotted as function of the track polar angle (*top*) and momentum (*bottom*). The distribution measured on the data (*full circles*) is compared to the Monte Carlo simulation (*full squares*).

semi-leptonic decays of hadrons with charm and beauty. In Section 5.3 we will describe these observables and the methods used to measure the fraction of semi-leptonic b-decays.

5.2 Signal and background Monte Carlo samples

The analysis of the background sources and the extraction of the signal is performed using large samples of Monte Carlo events. In order to simulate the contribution of fake muons from punch-through and sail-through as well as muons from pion or kaon decays, samples of events with light quarks (u, d and s) have been produced. Separate samples of heavy quarks have been generated in addition.

Light quark sample Due to the efficient background suppression of the selection procedure a very high statistics is required for these processes. In addition, the DIS cross section falls steeply with the photon virtuality Q^2 and the inelasticity y. Therefore, the Monte Carlo samples generated according to the respective cross sections are generally dominated by low Q^2 and low y events such that it is difficult to obtain sufficient statistics at high Q^2 and y. A solution to this problem is given by the *weighted Monte Carlo generation* [98]: events can be generated according to a weight depending on the variables for which the DIS cross section is singular. Since the DIS cross section is falling like $1/(Q^2y)$ such a function is a natural choice.

In addition, events with QCD radiation, for which hard jets can be measured, can also be enhanced. To this purpose, the 'hardness' of the partons of the QCD subprocess has to be quantified in a suitable variable. The transverse momentum $p_\perp$ of the partons from the leading order matrix element in the photon-parton frame can be used in case of a MEPS Monte Carlo program like `RAPGAP` (see Section 1.6) .

In our sample of light quarks for each event a weight $w(Q^2, y, p_\perp)$ is assigned:

$$w(Q^2, y, p_\perp) = \max\left(1, \frac{E_{cut}^2}{1 + Q^2 y + p_\perp^2}\right) \tag{5.4}$$

where E_{cut}^2 can be interpreted as the value above which all events are kept. E_{cut}^2 has been fixed to 100 in our case.

The light quark event sample has been generated with the program `RAPGAP` . The selected kinematic region is $Q^2 > 1$ GeV^2, $0.01 < y < 0.85$ and $p_\perp > 3$ GeV. The `CTEQ5L` [99] set of parton densities was used with $\Lambda_{QCD} = 250$ MeV. Using this weight definition Eq. 5.4 and the cuts described above a *reduction factor*[1] of 25 is achieved. This means that for a given computing time a 25 times larger date sample can be simulated.

Open charm sample Events with semi-leptonic decays of charmed hadrons into muons have been produced generating a sample of *open charm* with the program `RAPGAP`. The `CTEQ5L` set of parton densities was used with $\Lambda_{QCD} = 250$ MeV; the mass of the charm quark was set to 1.5 GeV. To enhance the number of decays into muons the following selection cuts were applied during the generation:

- Kinematic region defined by $Q^2 > 1$ GeV^2 and $0.01 < y < 0.85$
- At least one muon at hadron level with momentum $p_\mu > 1$ GeV and in the angular region $5° < \theta_\mu < 175°$.

The cuts correspond to a reduction factor of about 4.

Open beauty samples Several samples of open beauty events have been generated for this analysis using the Monte Carlo programs `RAPGAP` and `CASCADE`. The latter implements the CCFM evolution equations and the k_T-dependent gluon density. The samples include different QCD evolution equations, models of multi-parton production (see Section 1.6) and fragmentation functions of the b-quark into hadrons. In all samples the hadronisation is obtained using the Lund model as implemented in `JETSET` and the b-quark mass was set to 5 GeV. All samples have been produced with colliding e^+p beams; a single sample with colliding e^-p beams has been produced for checking purposes. For the `RAPGAP` samples the `CTEQ5L` set of parton densities was used with $\Lambda_{QCD} = 250$ MeV. The default fragmentation function is given by the Bowler distribution (see Section 1.8) with parameters set to $a = 0.3$ and $b = 0.58$. Here follows a detailed description of all samples:

- `RAPGAP` sample with parton shower (B1): This sample was generated using the parton shower approach to multi-parton production. The selected kinematic region is: $Q^2 > 1$ GeV and $0.01 < y < 0.85$. A muon requirement at the hadron level is imposed as for the charm sample.

[1]The reduction factor is defined as the ratio between the number of generated events and the number of events satisfying the imposed cuts.

- RAPGAP sample with Colour Dipole Model (B2): This sample was generated using the CDM approach to multi-parton production. It has been used to evaluate systematic uncertainties on the detector acceptance. The selected kinematic region is: $Q^2 > 2$ GeV and $0.01 < y < 0.85$. A muon requirement at the hadron level is imposed as for the charm sample.
- RAPGAP sample with colliding e^-p beams (B3): The sample has been generated in the kinematic region: $1 < Q^2 < 5000$ GeV^2 and $0.01 < y < 0.85$. A muon requirement at generator level as for the charm sample.
- CASCADE sample (B4): The sample was generated using the gluon density extracted by Jung and Salam [31] from the HERA measurements of the structure function F_2. The mass of the b-quark was fixed to 4.75 GeV. The sample was generated in the kinematic region defined by $1 < Q^2 < 5000$ GeV^2 and $0.01 < y < 0.85$. No additional filters were applied.
- RAPGAP sample with Peterson fragmentation function (B5-B7): three event samples were generated using the Peterson function to simulated the fragmentation of b-quarks into hadrons. The ϵ parameter was set to the values of 0.0055 (B5), 0.0041 (B6) and 0.0031(B7) as recently measured respectively by the SLD, OPAL and ALEPH collaborations [100, 101, 44].

A summary of the characteristics of all samples is given in Table 5.1 and 5.2. In the latter table the *equivalent luminosity* is calculated dividing the generated number of events by the cross section relative to the process.

Sample	Multi-Parton Prod Model	QCD Evolution	Fragmentation function
Light	P.S.	DGLAP	-
Charm	P.S.	DGLAP	-
B1,B3	P.S.	DGLAP	Bowler: $a = 0.3$, $b = 0.58$
B2	C.D.M.	DGLAP	Bowler: $a = 0.3$, $b = 0.58$
B4	P.S.	CCFM	Bowler: $a = 0.3$, $b = 0.58$
B5	P.S.	DGLAP	Peterson: $\epsilon = 0.0055$ [100]
B6	P.S.	DGLAP	Peterson: $\epsilon = 0.0041$ [101]
B7	P.S.	DGLAP	Peterson: $\epsilon = 0.0031$ [44]

Table 5.1 Summary of Monte Carlo samples generated for this analysis.

5.3 Extraction of events from b decays

The measurement of the cross section is based on a observable by which it is possible to distinguish the signal from the background. Such a distinction is not performed on a event-by-event basis but it consists in the study of the observable spectrum over a large sample of events. In practice, the relative fraction of signal events is estimated adjusting the proportion

Sample	Generated events (Selected×Red.factor)	Eq. luminosity (pb^{-1})
Light	$\sim 120 \cdot 10^6$	200
Charm	$\sim 7.5 \cdot 10^6$	94
B1	$\sim 430 \cdot 10^3$	540
B2	$\sim 860 \cdot 10^3$	$1 \cdot 10^3$
B3	$\sim 830 \cdot 10^3$	573

Table 5.2 Summary of Monte Carlo samples generated for this analysis.

of all simulated components until the best description of the measured observable is achieved. As a consequence, the method relies on the correct simulation of the observable distribution. In what follows we will refer to this approach as *statistical signal extraction.*

In the following sections we will first discuss the observable and we will then describe the fit technique that implements the statistical signal extraction.

5.3.1 Observable definition

In order to separate the semi-leptonic decays of beauty hadrons from charm decays and processes with light quarks we introduce the transverse momentum of the muon relative to its jet axis

$$p_T^{rel} = |\vec{p}_\mu| \sin\left(\mathrm{acos}\left(\frac{\vec{p}_\mu \cdot \vec{p}_{Jet}}{|\vec{p}_\mu||\vec{p}_{Jet}|}\right)\right) \tag{5.5}$$

where $\vec{p}_\mu$ and $\vec{p}_{Jet}$ are the three-momenta of the semi-leptonic muon and jet respectively. The association between the muon and the jet is performed using the k_T-clustering algorithm[2]. Figure 5.2-b shows a comparison of the MC distributions of the variable p_T^{rel} for different processes after all selection cuts. The value of p_T^{rel} of b-decays is, on average, higher than for charm and light quarks. The kinematic effect is due to the higher mass of the b-quark ($m_b \sim 3m_c$). In Figure 5.2-a the p_T^{rel} of b-decays is shown separating the contribution from direct semi-leptonic decays ($b \to c\mu X$) and indirect decays through a charmed hadron ($b \to c \to s\mu X$). The direct decays constitute the contribution at high values of p_T^{rel} while indirect decays have a shape similar to the charm decays. In addition, the p_T^{rel} distribution for charm decays is very similar to the light quarks.

5.3.2 Statistical signal extraction

Analysis of the results from HEP experiments often involves the estimation of the relative magnitude of different processes composing the data. After a brief introduction to the most common fit techniques, as the χ^2 minimization and the Maximum Likelihood Method, we will describe how the fluctuations due to the finite Monte Carlo statistics are taken in account.

[2]The k_T clustering algorithm returns which reconstructed jet contains the EFO associated to the muon candidate.

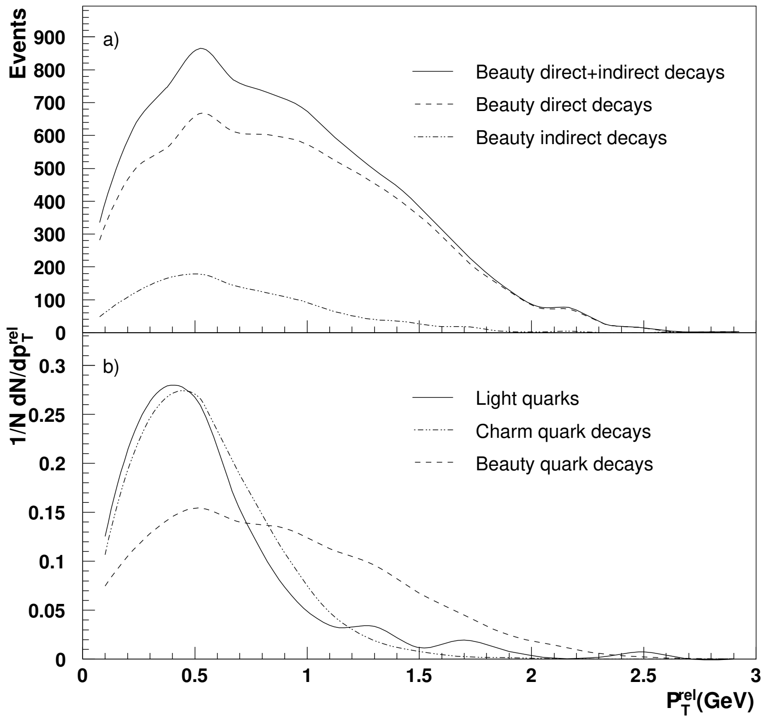

Figure 5.2 Distribution of p_T^{rel} after all selection cuts described in Chapter 4. a) Distribution of p_T^{rel} from events with direct ($b \to c\mu X$) and indirect ($b \to c \to s\mu X$) decays as well as their sum. b) Distribution of p_T^{rel} for light quarks, charm and beauty decays. The distributions are normalized to the same number of events.

χ^2 method and maximum likelihood

Let us suppose to have m different Monte Carlo samples used to describe the distribution observed in the data as a function of a certain variable x. If no analytic form for the function is available the range of the x variable has to be splitted into n bins. We denote with d_i the number of events in the data that fall into bin i and with a_{ji} the numbers of Monte Carlo events from the source j in bin i. The total number of events in the data N_D and the total number in the MC sample for source j are:

$$N_D = \sum_{i=1}^{n} d_i, \quad N_j = \sum_{i=1}^{n} a_{ji}. \tag{5.6}$$

The number of predicted events in bin i can be expressed in terms of fractions P_j of the m MC sources as

$$f_i(P_1, ..., P_m) = N_D \sum_{j=1}^{m} \frac{P_j a_{ji}}{N_j} \quad \text{with} \quad \sum_j P_j = 1. \tag{5.7}$$

It is convenient to incorporate the normalization factors into the fraction p_j, writing $p_j = N_D P_j / N_j$. In this way Eq. 5.7 can be rewritten as

$$f_i = \sum_{j=1}^{m} p_j a_{ji}. \tag{5.8}$$

A first approach to the problem is to estimate the fractions p_j by adjusting them to minimize the χ^2

$$\chi^2 = \sum_{i=1}^{n} \frac{(d_i - f_i)^2}{d_i}. \tag{5.9}$$

In this χ^2 definition it is assumed that the distribution for each d_i is Gaussian. The Gaussian distribution is a good approximation of the Poisson for large number of events in the bin i. If many bins have small values of d_i it is more appropriate to use the Poisson distribution, for which the probability to observe a value d_i in bin i is given by

$$e^{-f_i} \frac{f_i^{d_i}}{d_i!}. \tag{5.10}$$

In this case the solution is given by the fractions f_i which maximize the total likelihood $\mathcal{L}$, defined as the product over all bins of the n probabilities defined in Eq. 5.10. It is convenient to calculate the fractions which maximize the logarithm of $\mathcal{L}$, given by:

$$\ln \mathcal{L} = \sum_{i=1}^{n} (d_i \ln f_i - f_i) \tag{5.11}$$

where the constants have been omitted.

An advanced Binned Maximum likelihood method

The methods described so far do not consider the statistical error associated to the quantities a_{ji}, which is due to the finite statistics in each Monte Carlo sample. The procedure that we will now describe does not include this approximation [102].

For each Monte Carlo source and for each bin we define the expected number of events A_{ji}, which can differ from a_{ji} because of the limited statistics available in the MC. The number of predicted events, Eq. 5.7, as function of the new variables is

$$f_i = \sum_{j=1}^{m} p_j A_{ji}. \tag{5.12}$$

The quantities a_{ji} have a binomial distribution, but they can be approximated with a Poisson distribution if $A_{ji} \ll N_j$. This condition is true in our case since even large Monte Carlo samples can lead to a small number of events in each bin. The likelihood which is to be maximized is now the combined probability of the observed d_i and the observed a_{ji}:

$$\ln \mathcal{L} = \sum_{i=1}^{n} (d_i \ln f_i - f_i) + \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ji} \ln A_{ji} - A_{ji}). \tag{5.13}$$

In order to calculate the values of P_j and A_{ji} which maximize the likelihood we have to differentiate Eq. 5.13 and set the derivatives to zero. This gives two sets of equations

$$\sum_{i=1}^{n} \left(\frac{d_i A_{ji}}{f_i} - A_{ji} \right) = 0 \quad \forall\, j \tag{5.14}$$

$$\frac{d_i A_{ji}}{f_i} - p_j + \frac{a_{ji}}{A_{ji}} - 1 = 0 \quad \forall\, i, j. \tag{5.15}$$

These equations are not linear and coupled, however, they can be remarkably simplified. Eq. 5.15 can be rewritten as

$$1 - \frac{d_i}{f_i} = \frac{1}{p_j}\left(\frac{a_{ji}}{A_{ji}} - 1\right). \tag{5.16}$$

The left hand side depends only on i, we can therefore substitute $t_i = 1 - d_i/f_i$ and rewrite Eq. 5.16 as

$$A_{ji} = \frac{a_{ji}}{1 + p_j t_i}. \tag{5.17}$$

For a given set of p_j the $n \times m$ unknown quantities A_{ji} are given by the n unknown quantities t_i. Using the definition of t_i, if $d_i = 0$ then $t_i = 1$, otherwise

$$\frac{d_i}{1 - t_i} = f_i = \sum_j p_j A_{ji} = \sum_j \frac{p_j a_{ji}}{1 + p_j t_i}. \tag{5.18}$$

If these n equations are satisfied then all the other $m \times n$ equations are also satisfied. The Fortran `HMCMLL` subroutine [102] solves the n decoupled equations Eq. 5.18 obtaining the quantities t_i. The respective quantities A_{ji} are then obtained through Eq. 5.17. Using the definition Eq. 5.13 the logarithm of the likelihood is calculated. Finally, the maximum of $\ln \mathcal{L}$ is obtained numerically using the Newton algorithm.

5.3.3 Fit results

In order to perform the fit the distributions of light quark and charm have been combined. The combined distribution is obtained by assigning to each MC source a weight inversely proportional to the equivalent luminosity of the sample. Indicating with $N(i)$ the content of bin i of the MC mixture we have

$$N(i) = \frac{\sum_j n_j(i)/\mathcal{L}_j}{\sum_j 1/\mathcal{L}_j \sum_i n_j(i)} \tag{5.19}$$

where $\mathcal{L}_j$ is the equivalent luminosity of the j-th MC source and $n_j(i)$ is the content of the j-th MC source in bin i.

Applying the fit procedure described in the last Section the fraction f_b of events from b decays is obtained such that the total MC content in the bin i is

$$N(i; f_b) = f_b N_b(i) + (1 - f_b) N_{c+l}(i) \tag{5.20}$$

where $N_b(i)$ and $N_{c+l}(i)$ are the content of bin i respectively of the beauty MC and of the charm and light quark mixture. The fit result is

$$\boxed{f_b = (25.3 \pm 4.8)\%}. \tag{5.21}$$

The log-likelihood Eq. 5.13 as a function of the parameter f_b as a clear minimum, $f_{b,min}$. The asymmetric errors on $f_{b,min}$ are calculated following the log-likelihood function from the minimum and finding the values of f_b where it crosses the point $\ln \mathcal{L} = \ln \mathcal{L}(f_{b,min}) + \delta$ where δ was set to 0.5. Following this procedure we find $f_b = \left(25.3(0)^{+4.8(0)}_{-4.7(9)}\right)\%$. These values are in agreement with the symmetric error found assuming a parabolic shape of the log-likelihood function.

Figure 5.3 shows the distribution of the p_T^{rel} variable for data and Monte Carlo using the fraction f_b obtained from the fit.

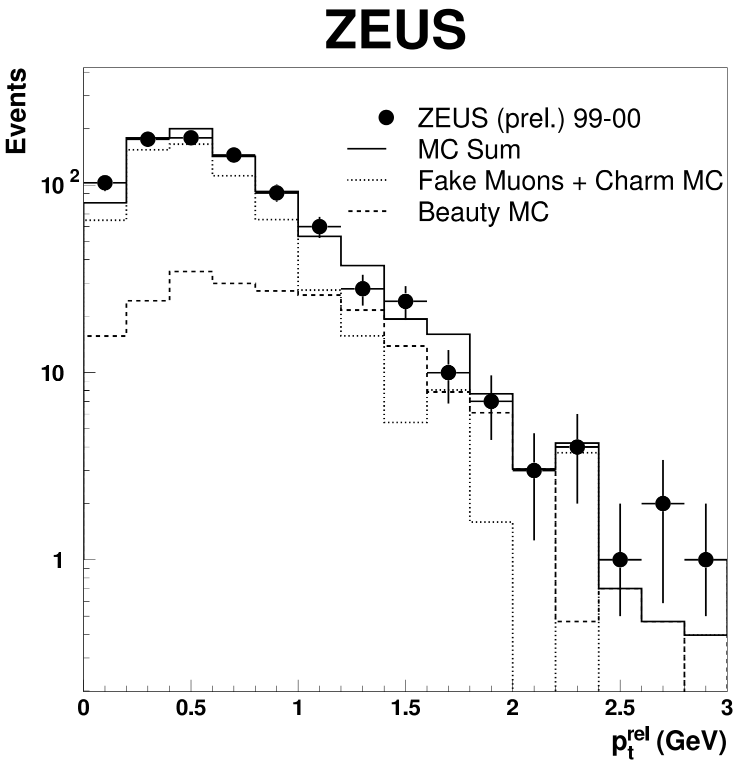

Figure 5.3 Measured p_T^{rel} distribution for the data and Monte Carlo simulations. The fitted contribution from beauty quarks (*dashed histogram*) and from light-plus-charm quarks (*dotted histogram*) are shown separately. The solid line corresponds to the sum of all Monte Carlo contributions. The data points are plotted with the statistical uncertainties only.

5.3.4 Comparison between data and Monte Carlo simulation

The extracted fraction of events from b-decays can be cross checked by comparing the measured kinematic distributions with the combination of all Monte Carlo predictions which has been determined by the fit. Since the separate fractions of events with charm and light quarks are not determined in the fit the two samples are summed according to the respective equivalent luminosities, as described in Section 5.3.3. Figure 5.4 shows a comparison of the measured and simulated distributions of several DIS kinematic variables. A comparison of the variables relative to the jet associated to the muon is shown in Figure 5.5; the respective muon variables are in Figure 5.6. The dip in the distribution of the muon azimuthal angle ϕ_μ corresponds to the octant of the barrel muon detector in which a smaller chamber is installed. The agreement between data and MC is satisfactory in all distributions, showing that the simulated processes in the fractions extracted with the fit can describe all the relevant kinematic variables of the data sample.

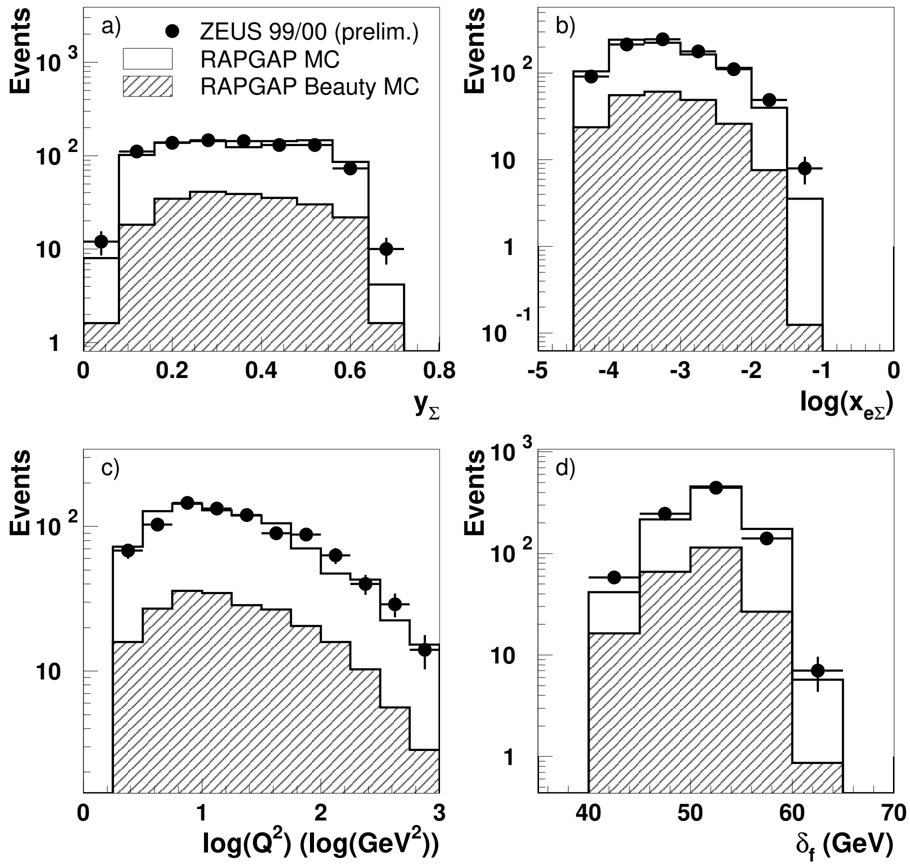

Figure 5.4 Data (*dots*) and RAPGAP MC (*solid line*) DIS distributions after the final event selection for a) inelasticity y_Σ, b) Bjorken variable $x_{e\Sigma}$, c) virtuality of the exchanged photon Q^2 and d) variable δ_f. The solid line represents all MC contributions according to the percentage given by the fit, while the hatched histograms show the contribution from b quarks. The error bars are statistical only.

5.4 Measurement of the total cross section

Using the data samples selected according to the criteria described in the previous chapter a total visible cross section is measured for the process

$$ep \rightarrow e\ b\ \bar{b}\ X \rightarrow e\ \mu\ Jet\ X. \tag{5.22}$$

The measurement has been performed in the DIS kinematic region

$$Q^2 > 2\ \mathrm{GeV}^2 \tag{5.23}$$

$$0.05 < y < 0.7 \tag{5.24}$$

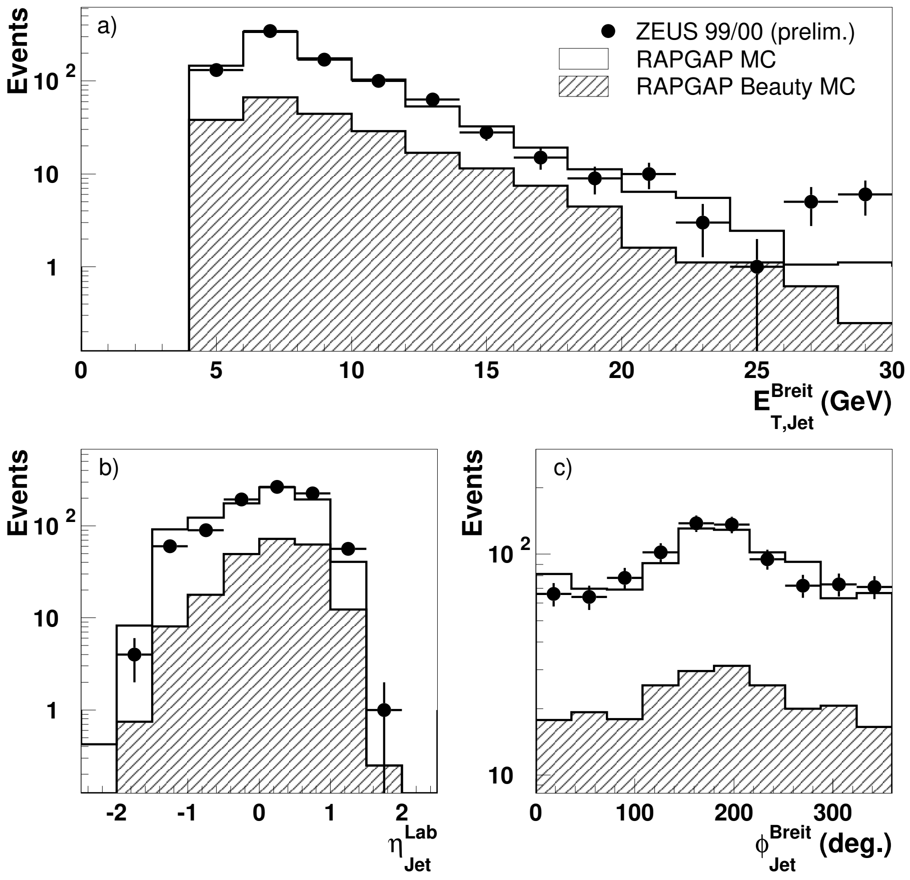

Figure 5.5 Data (*dots*) and RAPGAP MC (*solid line*) distributions after the final event selection of the jet associated to the muon, as described in Section 5.3.1. a) Jet transverse energy in the Breit frame, b) jet pseudorapidity in the laboratory frame, c) Jet azimuthal angle in the Breit frame. The solid line represents all MC contributions according to the percentage given by the fit, while the hatched histograms show the contribution from b quarks. The error bars are statistical only.

with the final state containing at least one jet in the Breit frame with

$$E_{T,Jet}^{Breit} > 6 \text{ GeV} \quad \text{and} \quad -2 < \eta_{Jet}^{Lab} < 2.5 \tag{5.25}$$

and at least one muon satisfying the cuts

$$p_\mu > 2 \text{ GeV} \quad \text{and} \quad 30^\circ < \theta_\mu < 160^\circ. \tag{5.26}$$

The total cross section σ_{tot} is calculated using the formula

$$\sigma_{tot} = \frac{N_{tot} \cdot f_b \cdot C}{\delta_\mu \cdot \mathcal{L}} \tag{5.27}$$

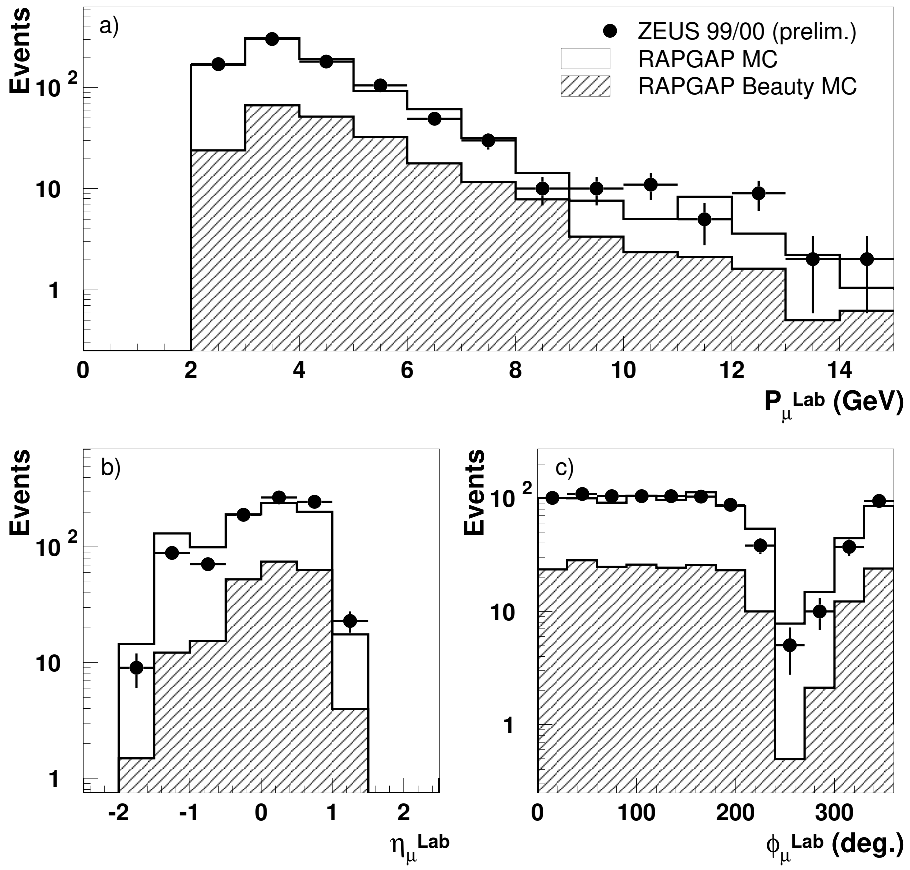

Figure 5.6 Data (*dots*) and RAPGAP MC (*solid line*) distributions after the final event selection of the muon associated to the jet, as described in Section 5.3.1. a) Muon momentum, b) muon pseudorapidity, c) muon azimuthal angle. The solid line represents all MC contributions according to the percentage given by the fit, while the hatched histograms show the contribution from b quarks. The error bars are statistical only.

where N_{tot} is the number of events selected in the final sample, f_b is the fraction of events from b-decays, C is the correction factor due to the selection acceptance, $\mathcal{L}$ the total luminosity of the data sample and δ_μ is the correction factor for the muon reconstruction efficiency discussed in Section 3.5.3.

The correction factor of the event selection is defined as the ratio $C = N_{gen}/N_{rec}$ between the number of events generated in the kinematic region defined by Eqs. 5.23-5.26 and the number of events reconstructed in the same kinematic region. Both quantities are estimated using the sample (B1-B4) of the signal Monte Carlo simulation (see Section 5.2). The value of N_{rec} is given by the number of events satisfying all selection cuts described in the previous chapter including the trigger preselection. In order to measure the number of events N_{gen} the cuts Eqs. 5.23-5.26 are applied to the *hadron level* of the simulation[3].

The values of the correction factors estimated with different Monte Carlo simulations are summarized in Table 5.3, where the relative uncertainty is computed using the binomial formula for error propagation. In the table the values of the selection efficiency, ϵ, and purity, p are given as well. These quantities are defined as follows

$$\epsilon = \frac{N_{gen} \cap N_{rec}}{N_{gen}}, \qquad p = \frac{N_{gen} \cap N_{rec}}{N_{rec}}. \tag{5.28}$$

A more detailed discussion of these variables is given in the next Section. The measured efficiency is $\sim 8\%$ whereas high values of purity ($\sim 90\%$) are observed.

Sample	N_{gen}	N_{rec}	C	Efficiency	Purity
B1	12499	1140	10.96±0.25	$(0.81 \pm 0.02) \cdot 10^{-1}$	$(89.3 \pm 0.9) \cdot 10^{-2}$
B2	11329	1004	11.29±0.27	$(0.79 \pm 0.03) \cdot 10^{-1}$	$(89.3 \pm 1.0) \cdot 10^{-2}$
B3	12288	1073	11.45± 0.26	$(0.78 \pm 0.02) \cdot 10^{-1}$	$(90.1 \pm 0.9) \cdot 10^{-2}$
B4	11580	1024	11.31±0.27	$(0.81 \pm 0.03) \cdot 10^{-1}$	$(92.1 \pm 0.8) \cdot 10^{-2}$

Table 5.3 Selection acceptance from different samples of signal Monte Carlo simulation. A detailed description of all samples is given in Section 5.2

The values of the correction factor C extracted from the MC samples B1, B2 and B4 agree within the statistical uncertainties. This implies that no dependence is introduced by different QCD evolution equations or by different models of the multi-parton production. No dependence on the lepton charge is observed (sample B3).

In the cross section calculation we use the value obtained from the Monte Carlo sample B1. Using Eq. 5.27 the measured cross section is

$$\boxed{\sigma_{tot}(ep \rightarrow e\ b\ \bar{b}\ X \rightarrow e\ \mu\ Jet\ X) = 38.7 \pm 7.7\ (\text{stat.})^{+6.1}_{-5.0}\ (\text{syst.})\ \text{pb}} \tag{5.29}$$

A detailed discussion of the statistical and systematic uncertainties is given in the Section 5.6.

5.5 Measurement of the differential cross sections

Differential cross sections of beauty production in DIS have been measured in the kinematic region defined by the cuts Eqs. 5.23-5.26 as function of the virtuality of the exchanged photon Q^2 and of the Bjorken scaling variable x_{Bj}. As we will discuss in Section 7.1.1 the ratio of the cross section of beauty to charm quark production depends on both the Q^2 and x_{Bj} variables. Therefore the fraction of b-decays was measured in each bin of the differential cross sections

[3]In practice, while the event kinematic variables Q^2 and y are given directly by the event generator, the muon and jet requirements must be verified on the table of particles generated for each event. To reconstruct the jets, the k_T-clustering algorithm is applied to all the particles of the hadronic final state excluding the scattered positron. The muon requirement is satisfied if at least one muon in the table survives the cuts Eq. 5.26.

by repeating the fit. The differential cross section in a bin of width $d\xi_i$ is calculated using the formula

$$\frac{d\sigma}{d\xi_i} = \frac{N_i \cdot f_{b,i} \cdot C_i}{\delta_\mu \cdot \mathcal{L}} \cdot \frac{1}{d\xi_i} \tag{5.30}$$

where N_i is the number of events in bin i, $f_{b,i}$ is the fraction of events from beauty decays in bin i and $C_i = N_{gen,i}/N_{rec,i}$ is the respective correction factor. In order to calculate the statistical uncertainty associated to the factor C_i we introduce the vector $N_{gen,i} \cap N_{rec,i}$, corresponding to the number of events which are reconstructed *and* generated in bin i. Taking into account the correlation between $N_{gen,i}$ and $N_{rec,i}$ the uncertainty on C_i is

$$\delta C_i = \sqrt{\frac{N_{gen,i}}{N_{rec,i}}(N_{gen,i} + N_{rec,i} - 2 \cdot N_{gen,i} \cap N_{rec,i})}. \tag{5.31}$$

For each bin we also define the *efficiency* ϵ_i and the *purity* p_i, given by

$$\epsilon_i = \frac{N_{gen,i} \cap N_{rec,i}}{N_{gen,i}}, \qquad p_i = \frac{N_{gen,i} \cap N_{rec,i}}{N_{rec,i}}. \tag{5.32}$$

One can think of the efficiency as the fraction of true events in a given bin which are *also* reconstructed in that bin and of the purity as the fraction of events reconstructed in a given bin which are *also* generated in that bin. The errors associated with the purity, δp_i, and efficiency, $\delta\epsilon_i$, are given by the binomial expressions

$$\delta p_i = \sqrt{\frac{(1-p_i) \cdot p_i}{N_{rec,i}}}, \qquad \delta\epsilon_i = \sqrt{\frac{(1-\epsilon_i) \cdot \epsilon_i}{N_{gen,i}}}. \tag{5.33}$$

5.5.1 Measurement of $d\sigma/dQ^2$

The differential cross section $d\sigma/dQ^2$ was measured in the range $2 < Q^2 < 1000$ GeV^2. To achieve comparable statistical uncertainties in each bin the final data sample was divided in three bins with approximately the same number of events. Figure 5.7 shows the distribution of the p_T^{rel} variable for data and Monte Carlo in each bin of Q^2 using the fractions $f_{b,i}$ extracted by the fit. The measured distributions are well described by the simulation in all bins. In each fit the measured log-likelihood function has a clear minimum. The values of f_b extracted in each Q^2 bin are shown in Table 5.4. The uncertainties on the b-fraction are calculated assuming a parabolic shape of the log-likelihood function as discussed in Section 5.3.3.

The correction factors, efficiencies and purities for each Q^2 bin are shown in Figure 5.8. The measured efficiency lies within 7% and 9% while a very high purity of about 85% is observed in all bins. The differential cross section is calculated applying Eq. 5.30. The values used in the calculation and the final results are summarised in Table 5.4. Figures 5.9 shows the measured differential cross section as functions of Q^2.

5.5.2 Measurement of $d\sigma/dx_{Bj}$

The differential cross section $d\sigma/dx_{Bj}$ was measured in the range $-4.5 < \log(x_{Bj}) < -1$ and in the kinematic region defined by the cuts Eqs. 5.23-5.26. As for the measurement of $d\sigma/dQ^2$, the final data sample was divided into three bins with approximately the same number of events. The fraction of events from beauty events was determined in each bin (see Figure

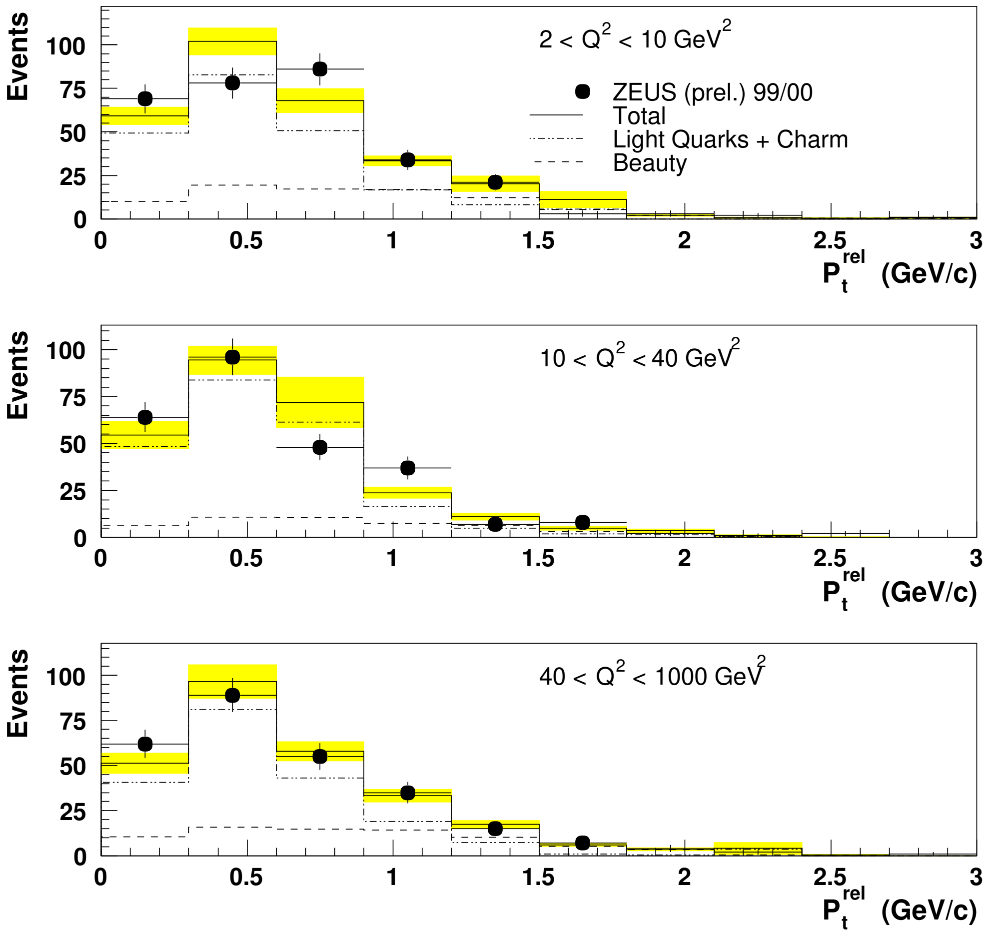

Figure 5.7 Measured p_T^{rel} distribution for the data and MC in bins of Q^2. The fitted contribution from beauty quarks (*dashed histogram*) and from light-plus-charm quarks (*dotted histogram*) are shown separately. The solid line corresponds to the sum of all MC contributions and the shaded band its statistical uncertainty.

Bin	Q^2 range (GeV^2)	Events	f_b (%)	Correction factor	$d\sigma/dQ^2$ (pb/GeV^2)
1	$[2, 10]$	298	28.5±8.2	12.5±0.5	2.2±0.7
2	$[10, 40]$	265	17.7±8.8	9.9±0.4	0.3±0.1
3	$[40, 10^3]$	271	27.7±8.3	10.4±0.5	$(13 \pm 4) \cdot 10^{-3}$

Table 5.4 Summary of the values used in the cross section measurement. All the errors are statistical only.

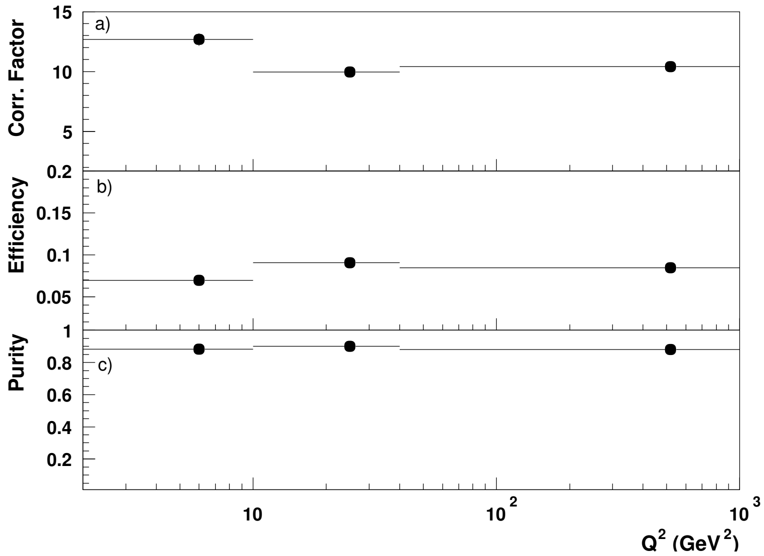

Figure 5.8 a) Correction factor, b) efficiency and c) purity as function of Q^2. All the variables are measured using the MC sample B1.

5.10). The measured b-fraction lies between $\sim 25 \div 33\%$ whereas the efficiency is $\sim 8\%$. A very high purity ($\sim 90\%$) is observed also in this case. The correction factors, the efficiency and the purity as a function of the Bjorken variable x_{Bj} are shown in Figure 5.11

The differential cross section was calculated using Eq. 5.30 and the values summarised in Table 5.5. The measured differential cross section as a function of $\log(x_{Bj})$ is shown in Figure 5.12.

Bin	$\log(x_{Bj})$ range	Events	f_b (%)	Correction factor	$d\sigma/dx_{Bj}$ (pb)
1	$[-4.5, -3.5]$	289	25.8±8.2	13.1±0.5	16.1±5.3
2	$[-3.5, -2.9]$	268	24.7±9.2	9.7±0.4	17.6±6.7
3	$[-2.9, -1.0]$	279	33.0±7.9	10.2±0.5	8.2±2.0

Table 5.5 Summary of the values used in the cross section measurement. All the errors are statistical only.

5.6 Experimental uncertainties

Various sources contribute to the total uncertainty on the measured beauty cross section. In particular, systematic errors are introduced by the selection procedure, by the fit technique, by the uncertainty on the detector calibration. In what follows we discuss how the different sources were evaluated and how the total error is obtained.

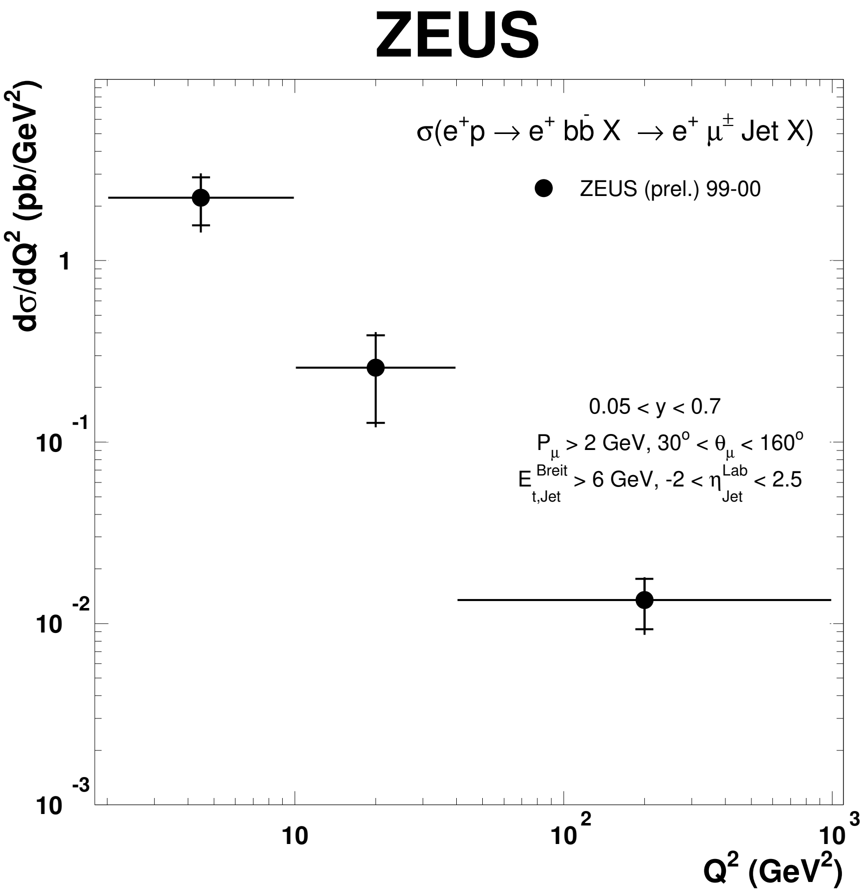

Figure 5.9 Differential beauty cross section as a function of Q^2. The error bars on the data points correspond to the statistical uncertainty (*inner error bars*) and to the statistical and systematic uncertainty added in quadrature (*outer error bars*).

5.6.1 Statistical uncertainties

Statistical uncertainties arise from the limited number of events in the data set and in the Monte Carlo samples used for the statistical signal extraction. The finite statistics of the simulated samples implies also an uncertainty on the selection acceptance discussed in the previous section. The statistical error on the cross section is therefore calculated propagating the single contributions of the terms appearing in Eq. 5.27.

5.6.2 Systematic uncertainties

In the following sections we will describe the contributions to the total systematic uncertainty. The systematic checks can be of different nature. A first group concerns the systematic effects introduced by the selection cuts. In these checks the analysis is repeated changing one of the selection cuts both on the data and Monte Carlo samples. A second group includes all systematic errors depending on the calibration scales and constants. In addition, we have considered the systematic uncertainties related to the efficiency of the muon detectors (see Section 3.5.3) and to the statistical extraction method. Several Monte Carlo programs were used to obtain the correction factors, assuming different models for the multi parton production

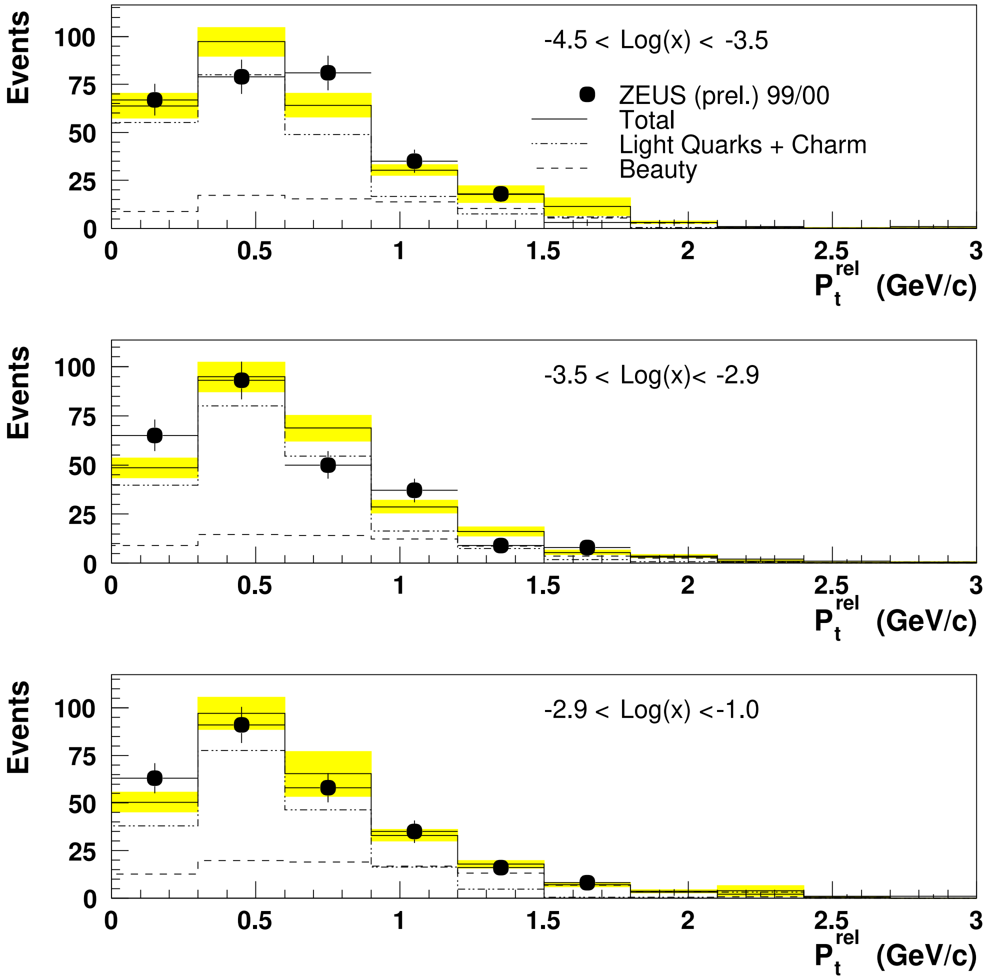

Figure 5.10 Measured p_T^{rel} distribution for the data and MC in bins of x_{Bj}. The fitted contribution from beauty quarks (*dashed histogram*) and from light-plus-charm quarks (*dot-dashed histogram*) are shown separately. The solid line corresponds to the sum of all MC contributions and the band to its statistic uncertainty.

and for the fragmentation of the b-quark. These model dependences has been included in the systematic checks.

The contributions to the systematic error of the total cross section Eq. 5.29 are summarized in Table 5.6. The relative difference between the cross section resulting from the systematic check and the nominal one is considered the systematic uncertainty. The total systematic error is obtained is obtained summing in quadrature the single contributions of the same sign.

For the differential cross sections, the systematic errors were evaluated in each bin. The results are shown in Figure 5.13 and 5.14 in bins of Q^2 and x_{Bj} respectively.

Uncertainties in the DIS selection The following studies were performed to evaluate the systematic uncertainties associated to the selection of the DIS sample (see Section 4.2.1):

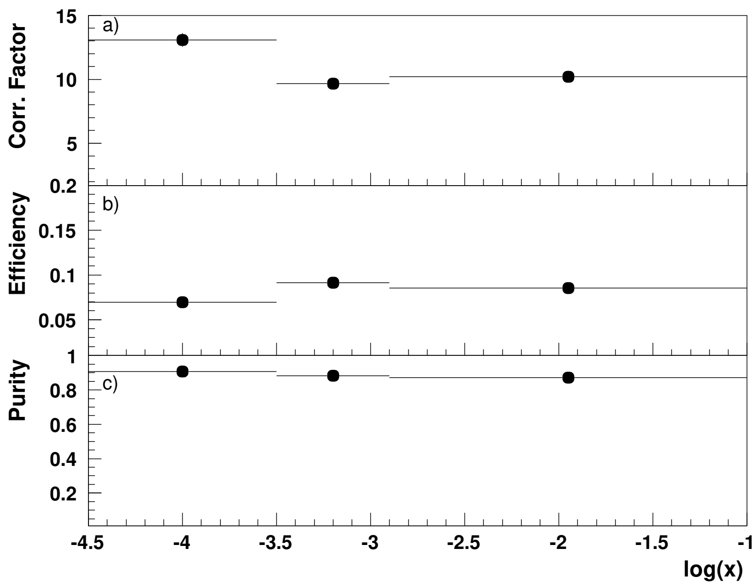

Figure 5.11 a) Correction factor, b) efficiency and c) purity as function of x_{Bj}. All the variables are measured using the MC sample B1.

- *Variation of the electron energy cut:* The cut on the electron energy was varied by ± 1 GeV
- *Variation of the cut on δ:* The lower and upper cut on the variable δ (see Eq. 4.5) were varied by ±2 GeV.
- *Uncertainty on the SRTD-RCAL alignment:* The relative alignment between the RCAL and the SRTD detector was estimated with a precision of ±2 mm [103]. This implies a systematic uncertainty on the measurement of the deflection angle of the scattered positron. The related systematic error has been estimated shifting the reconstructed SRTD hit position by ±2 mm in both coordinates.
- *Variation of the box cut:* The cut was extended by +3 cm on all sides.
- *Uncertainty on the electromagnetic energy scale:* The electron energy calibration between data and Monte Carlo was changed by ±1%.
- *Variation of the kinematic reconstruction method:* In order to evaluate the uncertainty introduced by the chosen method of kinematic reconstruction the analysis was repeated using the *Double angle* method.

Uncertainties in the jet selection The following studies were performed to evaluate the systematic uncertainties associated to the selection of jets in the Breit frame (see Section 4.2.4).

- *Uncertainty on the jet energy scale:* The uncertainty on the jet energy scale has been discussed in Section 3.4.4. According to this study an uncertainty of ±2% must be taken

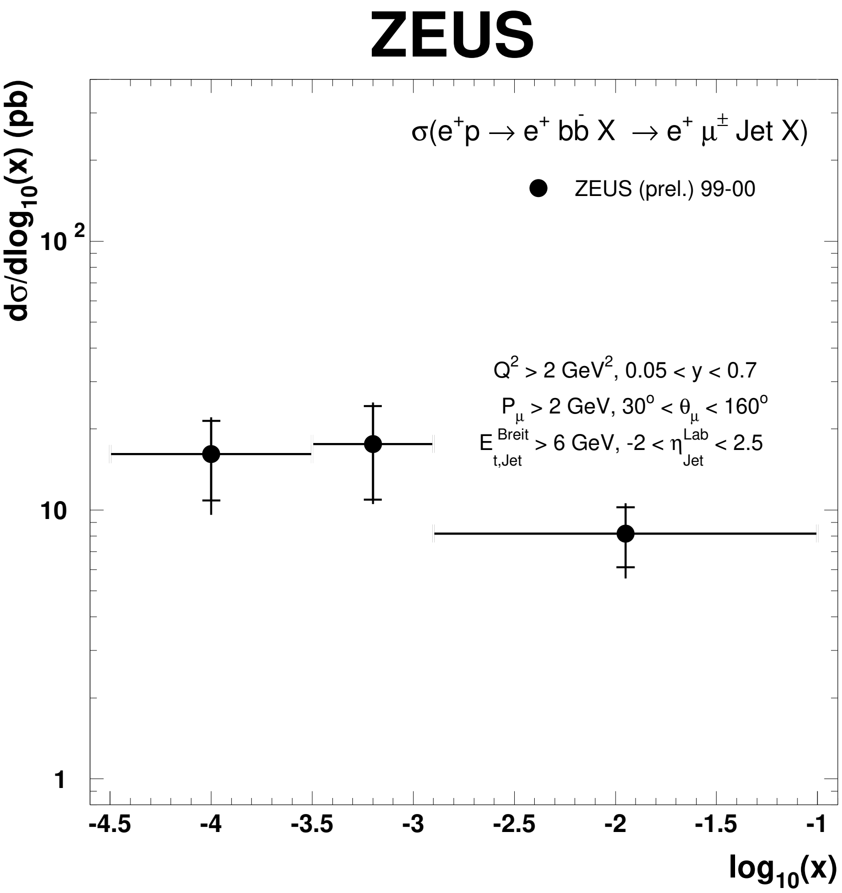

Figure 5.12 Differential beauty cross section as a function of $\log(x_{Bj})$. The error bars on the data points correspond to the statistical uncertainty (*inner error bars*) and to the statistical and systematic uncertainty added in quadrature (*outer error bars*).

in account. The corresponding error was estimated by varying the jet transverse energy of $\pm 2\%$ on the Monte Carlo sample and recalculating the selection acceptance.

- *Variation of the jet E_T cut:* The cut on E_T^{Breit} has been varied by ± 1 GeV.
- *Variation of the jet η cut:* The jet pseudorapidity region has been restricted to $-2 < \eta^{Breit} < 2.3$ to remove jets reconstructed in the very forward region.

Uncertainties of the statistical signal extraction The systematic uncertainties related to the statistical signal extraction were estimated by repeating the fit of the p_T^{rel} distribution with a different number of bins. The p_T^{rel} distribution was divided in 10, 15, 20 and 25 bins. The uncertainty was set to the maximum observed deviation from the nominal fit result.

Uncertainty on the muon reconstruction efficiency The ratios between the measured and simulated muon reconstruction efficiencies have been discussed in Section 3.5.3. The correction factors summarized in Table 3.2 have an uncertainty mostly due to the limited statistics of the used data sample. As a systematic check, the correction factors have been varied independently within one standard deviation and the selection acceptance was recalculated accordingly.

Uncertainty on the selection acceptance The selection acceptance has been estimated using several Monte Carlo samples, obtained using different assumptions on the QCD evolution equations and on the multi-parton production. We define the associated systematic uncertainty as the maximal deviation respect to acceptance of the reference sample.

Uncertainty of the luminosity measurement The uncertainty associated with the luminosity measurement for the 1999-00 data taking periods used in this analysis was estimated to be ±2.5%. This introduces an overall normalization uncertainty on each measured cross section, which is correlated between all data points.

Check	Description	$\Delta\sigma_{tot}/\sigma_{tot}$ (%)
1	$E'_e > 10 \pm 1$ GeV	$^{+0.4}_{-3.2}$
2	42< $E - P_z$ <67 GeV	-1.4
3	38< $E - P_z$ <63 GeV	+0.9
4	Jet cut: $E_T^{Breit} > 5 \pm 1$ GeV	$^{-2.4}_{-3.3}$
5	Jet cut: $\eta^{Lab} < 2.3$	-0.9
6	SRTD box cut + 3 cm for each side	-3.9
7	Hadronic energy scale ±2%	$^{+1.3}_{-0.6}$
8	E.M. energy scale ±1%	$^{-1.0}_{+1.1}$
9	SRTD Align.: $x' = x_{SRTD} + 2$ mm	-2.0
10	SRTD Align.: $y' = y_{SRTD} + 2$ mm	-0.9
11	SRTD Align.: $x' = x_{SRTD} - 2$ mm	-1.6
12	SRTD Align.: $y' = y_{SRTD} - 2$ mm	+0.6
13	BMUON $\delta'_\mu = \delta_\mu$+0.08	-10.2
14	BMUON $\delta'_\mu = \delta_\mu$-0.08	13.4
15	RMUON $\delta'_\mu = \delta_\mu$+0.05	-1.2
16	RMUON $\delta'_\mu = \delta_\mu$-0.05	1.1
17	Double Angle	1.3
18	Fit	$^{+8.0}_{-4.6}$

Table 5.6 Systematic uncertainties of the total cross section Eq. 5.27

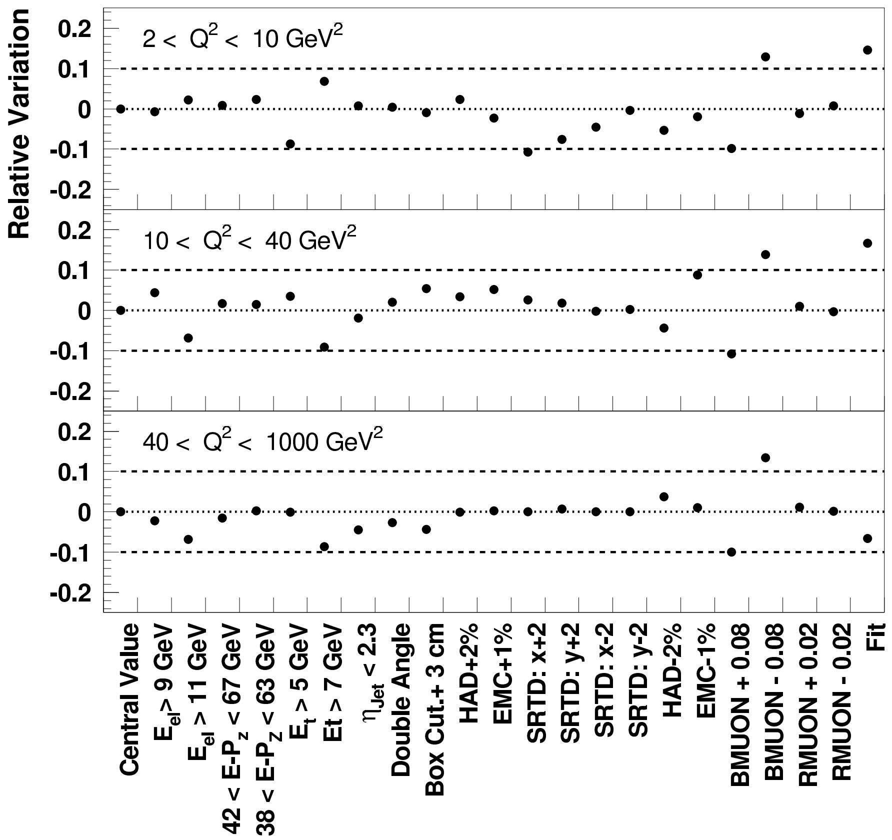

Figure 5.13 Systematic uncertainties of the differential cross section $d\sigma/dQ^2$ in bins of Q^2 (*see text*).

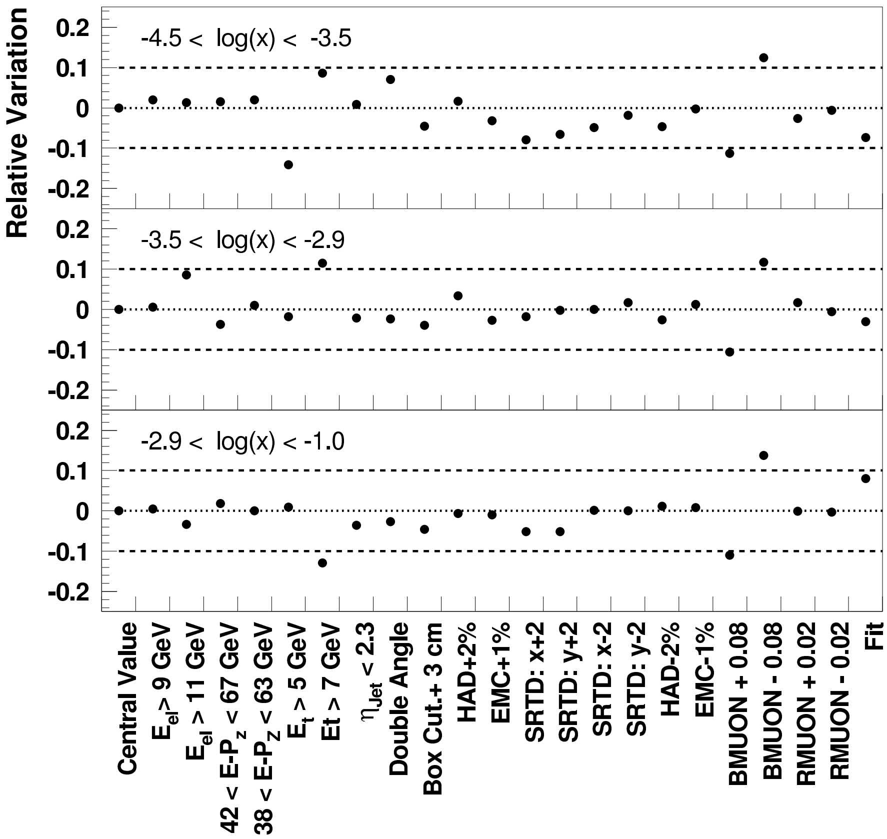

Figure 5.14 Systematic uncertainties of the differential cross section $d\sigma/dx_{Bj}$ in bins of $\log(x_{Bj})$ (*see text*).

6 Results

In this section we present a comparison of the measured cross sections with QCD Predictions. The next-to-leading order QCD prediction and the relative uncertainties will be discussed. The experimental results will also be compared to the predictions obtained with Monte Carlo programs based on leading-order matrix elements and implementing different QCD evolution equations.

6.1 Next-to-leading order QCD prediction

The comparison of the measured b cross section with QCD predictions needs a three-step calculation:

1. The calculation of the partonic processes which generate the final state with b quarks and gluons. This is performed to next-to-leading accuracy using the calculation by Harris and Smith implemented in the program `HVQDIS` [12]. The calculation was performed using the subtraction scheme and it is discussed in more details in Section 1.4.

2. The fragmentation of the b-quark to a B-meson. This is parameterized by the fragmentation functions discussed in Section 1.8.

3. The decay of the B-meson into a muon and hadrons. The spectrum of the B-meson is folded with the momentum spectrum of the semi-leptonic muons extracted from Monte Carlo programs.

The calculation scheme is illustrated in Figure 6.1. Since the first two steps of the calculation have been previously discussed we will here cover the third step.

To model the kinematics of the muon decay we use the muon momentum spectrum in the center-of-mass frame of the B-hadron extracted from `JETSET`. As discussed, `JETSET` is generally interfaced with the most common Monte Carlo event generators and it simulates the hadronization of the partonic final state to hadrons and their successive decay. Starting with a B-meson a muon can be produced in the final state with a branching probability $\mathcal{B}_\mu(B \to \mu X) = 0.22$ [1]. The muon can either be produced through a direct semi-leptonic decay ($b \to \mu$), or by indirect decay through a charmed hadron ($b \to c \to \mu$) or, in small fractions, through indirect decays containing $\tau^\pm$ or a J/ψ meson. In Table 6.1 are summarized the values of the branching ratios $B \to \mu X$ implemented in `JETSET` and compared with the measurements collected in Ref. [104].

In the calculation of the cross-section $ep \to b\bar{b} \to \mu X$ it has to be taken into account that the muon which has been measured in the detector can stem from the b-quark or the $\bar{b}$-quark.

[1] This value of the branching is implemented in the program `JETSET`

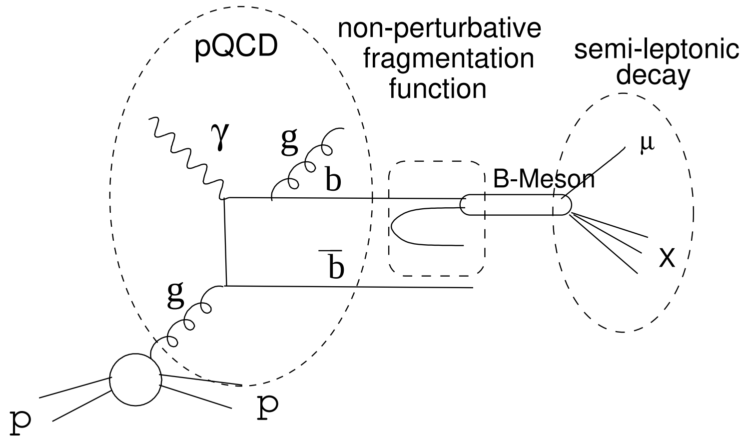

Figure 6.1 Schematic representation of the NLO QCD calculation. The partonic process is calculated to the NLO accuracy using the program `HVQDIS`. The kinematics of the B-meson is obtained folding the b-quark with a fragmentation function. The B-meson then decays semi-leptonically into muons.

Decay type	`JETSET`	PDG 2002 [104]
$b \to \mu^-$	0.10475	0.106 ± 0.002
$b \to c \to \mu^+$	0.08244	0.080 ± 0.004
$b \to \bar{c} \to \mu^-$	0.02428	0.016 ± 0.005
$b \to \tau^- \to \mu^-$	0.00636	0.0063 ± 0.0008
$b \to J/\Psi \to \mu^\pm$	0.00145	0.00068 ± 0.00006
$b \to \tau^+ \to \mu^+$	0.00106	0.0063 ± 0.0008
$b \to$ other $\to \mu^\pm$	0.00015	-

Table 6.1 b-quark decay modes and branching ratio as implemented in `JETSET` and measured in e^+e^- collisions.

The b-quark and the $\bar{b}$-quark are both independently decayed into muons according to the branching ratio $\mathcal{B}_\mu$. If one of the muons is in the detector acceptance the event is counted in the calculation and the cross-section as calculated by NLO QCD is multiplied with the probability that at least one of the quarks decays into muons $1-(1-\mathcal{B}_\mu)(1-\mathcal{B}_\mu) = 2\mathcal{B}_\mu - \mathcal{B}_\mu^2$. To improve the efficiency of the calculation, one can also decay one of the b-quark with probability $\mathcal{B}_\mu$ and then decay the other one according to the probability that exactly two muons are in the event under the condition that there is at least one muon, i.e. $\mathcal{B}_\mu^2(2\mathcal{B}_\mu - \mathcal{B}_\mu^2) = \mathcal{B}_\mu/(2-\mathcal{B}_\mu)$.

Muons are generated isotropically in the rest frame of the B-hadron. The momentum is assigned through a probability function extracted from `JETSET` muon momentum spectrum. Figure 6.2 shows the momentum distribution of final state muon in the rest frame of the B-hadron, dN/dp. The distributions corresponding to the direct and indirect b-decays are shown separately. From this distribution, the probability function $\mathcal{P}(p)$ is obtained through the relation

$$\mathcal{P}(p) = \frac{\int_0^p dp' \, \frac{dN}{dp'}}{\int_0^{p_{max}} dp' \, \frac{dN}{dp'}}, \tag{6.1}$$

where p_{max} was set to 2.5 GeV.

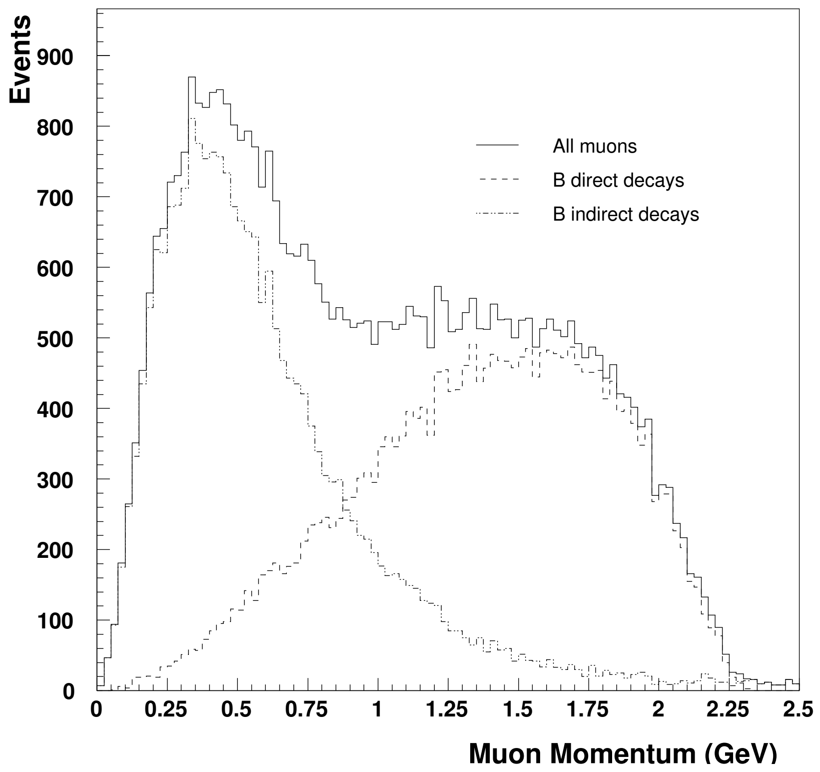

Figure 6.2 Momentum distribution of the decay muons in the rest frame of the parent B-hadron. The contributions from direct ($b \to \mu$) and indirect ($b \to c \to \mu$) decays are shown separately.

6.1.1 `HVQDIS` predictions at leading order

A comparison of the LO `HVQDIS` prediction with the `RAPGAP` generator provides a check of the kinematics of the leading order matrix elements through the shape and overall normalization of the distributions. For this purpose, both codes were run with the CTEQ5L proton parton distributions. The set has $\Lambda^{(5)} = 146$ MeV and $n_f = 5$ in the one-loop strong coupling α_s. The renormalization and factorization scales were both set to Q^2 and the beauty quark mass to $m_b = 5$ GeV. In the `RAPGAP` program the parton showering was turned off.

In what follows we define the *DIS cross section* as the one obtained applying only the DIS cuts Eqs. 5.23-5.24 and no further requirement. The *jet cross section* includes the DIS cuts and the jet cuts Eq. 5.25 whereas for the jet-muon cross section both the jet and muon requirements Eq. 5.26 are applied.

The inclusive DIS b-quark cross-section as calculated by LO `HVQDIS` is $\sigma_b = 530$ pb and agrees within 2% with the prediction from `RAPGAP`. For the jet b-quark cross-section $\sigma_{b,jet} = 148$ pb is found. The agreement with RAPGAP is within 4%. This is an indication that the LO matrix elements used in the two programs are the same. However, difference up to 5% are found for large transverse energies of the b-quark $E_{T,b}$ and of the jet $E_{T,jet}$. This can be seen in Figure 6.3 where the inclusive DIS b-quark cross section as a function of the transverse energy of the b-quark $E_{T,b}$ and as a function of the transverse jet energy $E_{T,jet}$ is shown. No strong dependence on the pseudo-rapidities of the b-quark η_b or of the jet η_{Jet} is found (see Figure 6.3-b,d). For the jet-muon cross section $\sigma_{b,jet-\mu} = 19$ pb is found with LO HVQDIS, while the one from RAPGAP is 10% larger. The observed difference in $\sigma_{b,jet-\mu}$ is probably due to an incomplete modeling of the muon fragmentation in the HVQDIS program.

6.1.2 Renormalization scale and b-quark mass uncertainties

The NLO QCD calculation of b-quark production depends on two scales: the renormalization (μ_R) and factorization (μ_F) scales. The former arises in any fixed-order calculation, in which the expansion of the observable in power of the strong coupling constant is truncated at a certain order. The latter is a direct consequence of the factorization theorem and it defines the separation between short-distance QCD processes and long-distance, non-perturbative processes, which are absorbed in the parton distributions. A comprehensive discussion of the theoretical uncertainties for this process can be found in Ref. [105]. In what follows we set the renormalization and factorization scales to

$$\mu_R^2 = \mu_F^2 = k_\mu^2 \cdot \bar{\mu}, \quad \bar{\mu} = f(m_b^2, Q^2, p_{T,b}^2) \tag{6.2}$$

where f is a function of the b-quark mass, m_b, of the photon virtuality Q^2 and of the transverse momentum of the b-quark, $p_{T,b}$. The renormalization scale uncertainties are defined as the relative variation of the cross section when the factor k_μ is varied in the range $0.5 < k_\mu < 2$. An additional source of uncertainty is the mass of the b-quark, which is varied between 4.5 and 5 GeV.

Figure 6.4 shows the scale and mass uncertainties as a function of Q^2 for the cross section defined in the kinematic region Eqs. 5.23-5.26 and for different combinations of the possible scales. The inner bands represent the uncertainty due to the variation of the renormalization scale while the outer band represents the uncertainty due to the variation of the b-quark mass. In all distributions we observe a decreasing uncertainty for increasing values of the photon virtuality. The total uncertainty is in all cases between 10% and 20%. The smallest uncertainties are observed with the function $f = p_{T,b}^2 + 4m_b^2$.

The predicted b-quark cross section and the relative uncertainty, as a function of the jet transverse energy in the Breit frame and of the pseudorapidity of the jet associated to the muon in the laboratory frame are shown in Figure 6.5 for the scale $\mu^2 = p_{T,b}^2 + 4m_b^2$. The scale uncertainties are large for values of the jet transverse energy $E_{T,jet}^{Breit} \sim 0$ and decrease for increasing values of $E_{T,jet}^{Breit}$ reaching a minimum around the value of the b-quark mass. Then, the uncertainties increase for increasing values of $E_{T,jet}^{Breit}$ reaching a plateau of about 20%. The uncertainty related to the b-quark mass dominate the total error for $E_{T,jet}^{Breit} \sim m_b$. As a function of the jet pseudorapidity, we observe increasing uncertainties of the calculation as we move towards the forward region. The respective distributions for the scale $\mu^2 = Q^2 + 4m_b^2$ are shown in Figure 6.6.

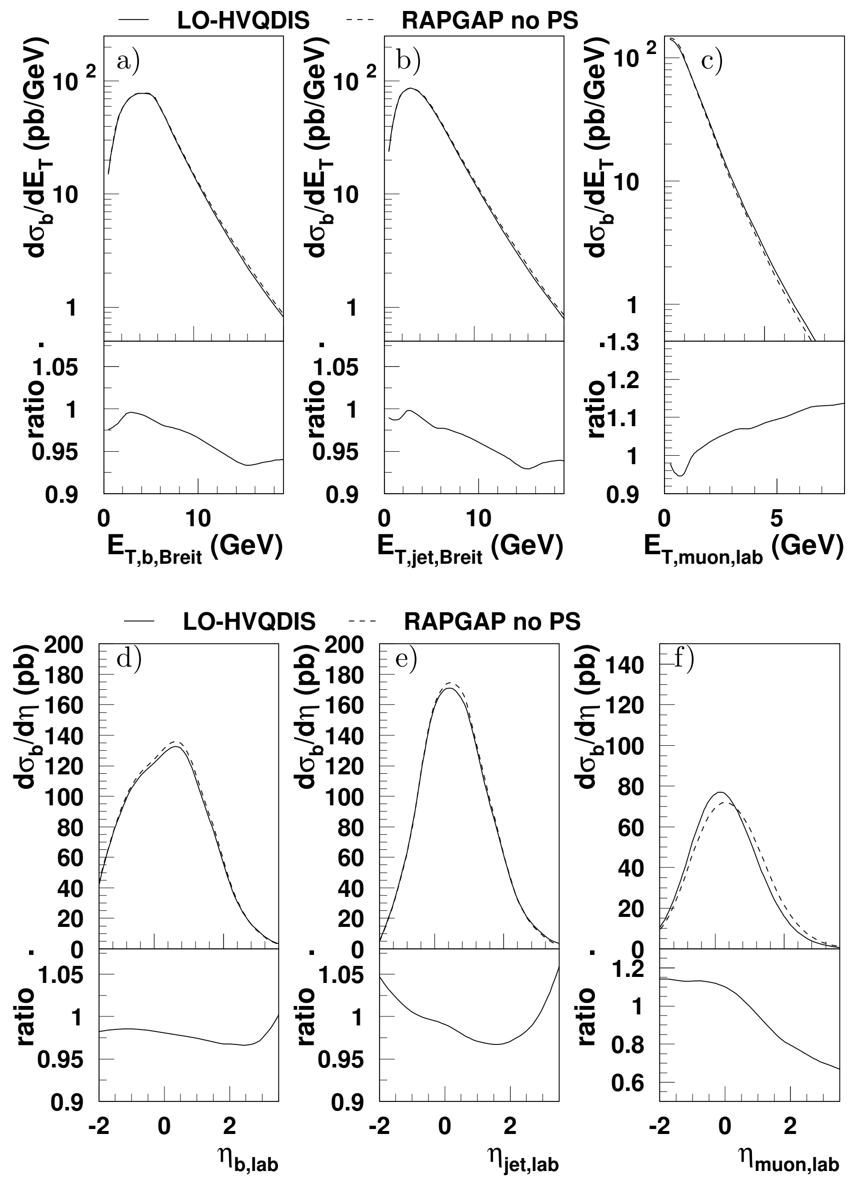

Figure 6.3 Comparison of the LO HVQDIS prediction with the RAPGAP Monte Carlo generator. *Top:* inclusive DIS b-quark cross section as a function of the transverse energy of the b-quark in the Breit frame (a), of the b-quark jet in the Breit frame (b) and of the semi-leptonic muon in the laboratory frame (c). *Bottom:* b-quark cross section as a function of the pseudorapidity of the b-quark (d), of the b-quark jet (e) and of the semi-leptonic muon (f). For each distribution the ratio between the HVQDIS and RAPGAP prediction is also shown.

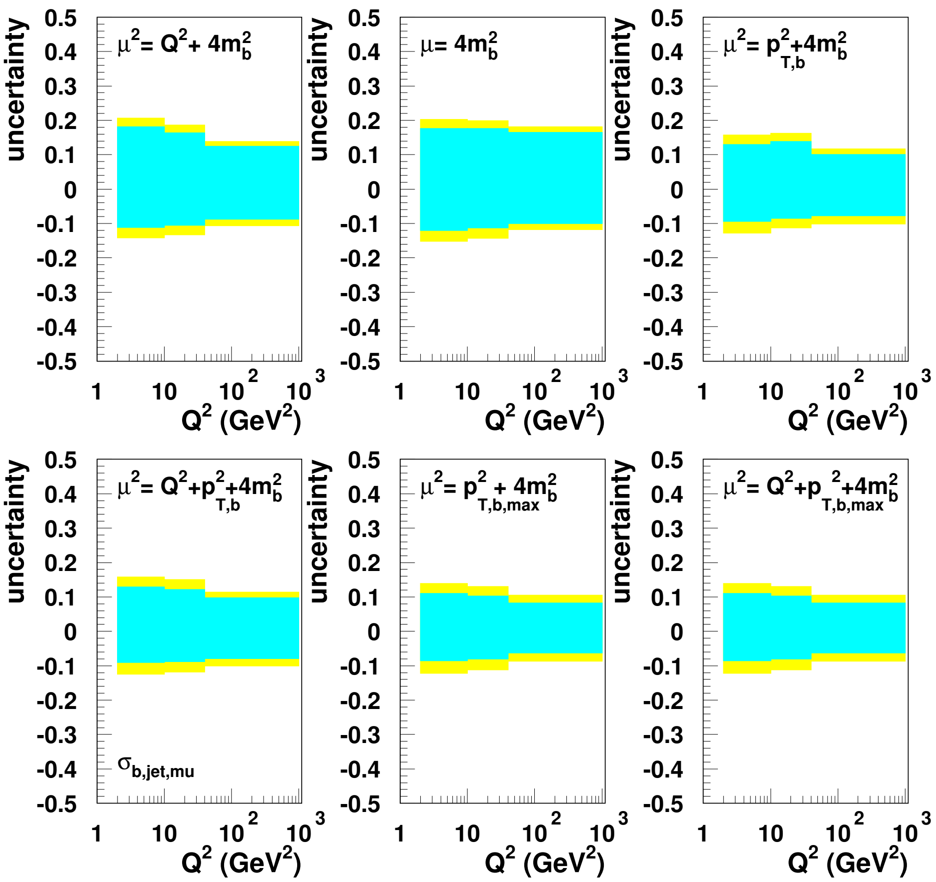

Figure 6.4 Uncertainty of the beauty quark cross section as a function of Q^2 for different choices of the renormalization and factorization scales. The inner band represents the uncertainty related to the variation of the renormalization scale, μ while the outer band represents the uncertainty due to the variation of the b-quark mass, m_b.

6.2 Comparison with NLO QCD predictions

Using the data samples collected by the ZEUS Collaboration in the years 1999-2000 total and single differential cross sections were measured for the process:

$$ep \rightarrow e\ b\ \bar{b}\ X \rightarrow e\ \mu\ Jet\ X.$$

The integrated luminosity of the data sample is 60 pb^{-1}. The measurement was performed in the DIS kinematic region defined by

$$Q^2 > 2\ \text{GeV}^2$$

$$0.05 < y < 0.7$$

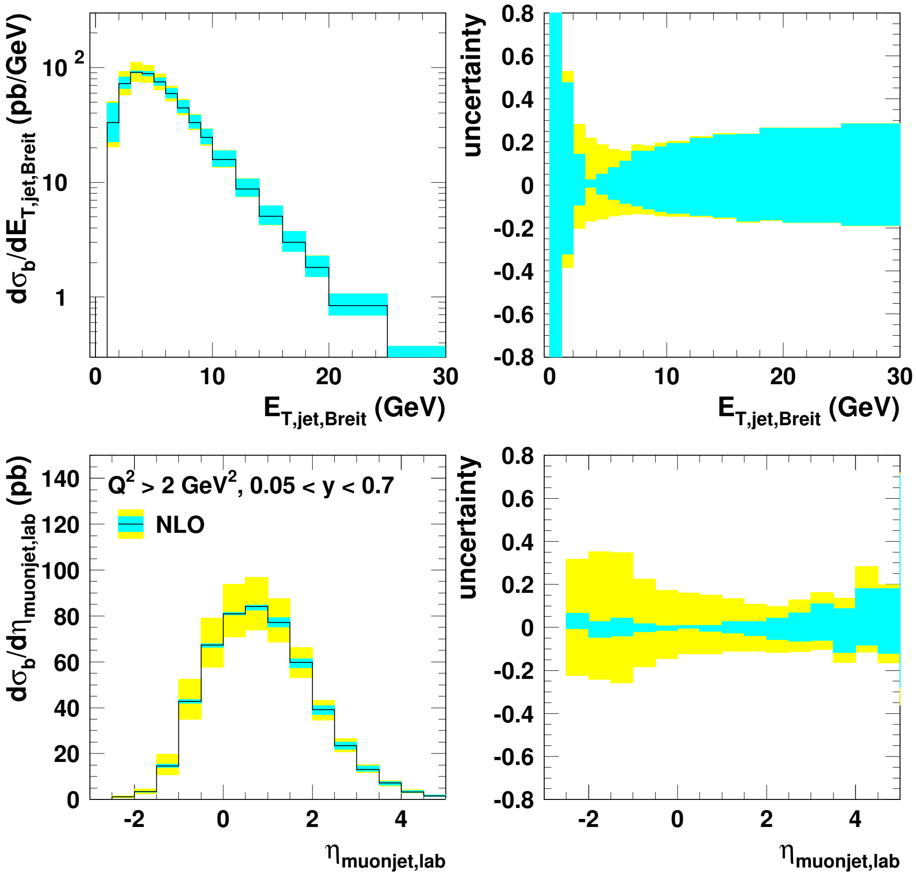

Figure 6.5 Scale and mass uncertainties of the b-quark cross section for the scale choice $\mu^2 = p_{T,b}^2 + 4m_b^2$. *Top:* b-quark cross section as a function of the jet transverse energy in the Breit frame (*left*) and relative uncertainties (*right*). The inner band represents the uncertainty related to the variation of the renormalization scale, μ while the outer band represents the uncertainty due to the variation of the b-quark mass, m_b. *Bottom:* b-quark cross section as a function of the jet pseudorapidity in the laboratory frame (*left*) and relative uncertainties (*right*).

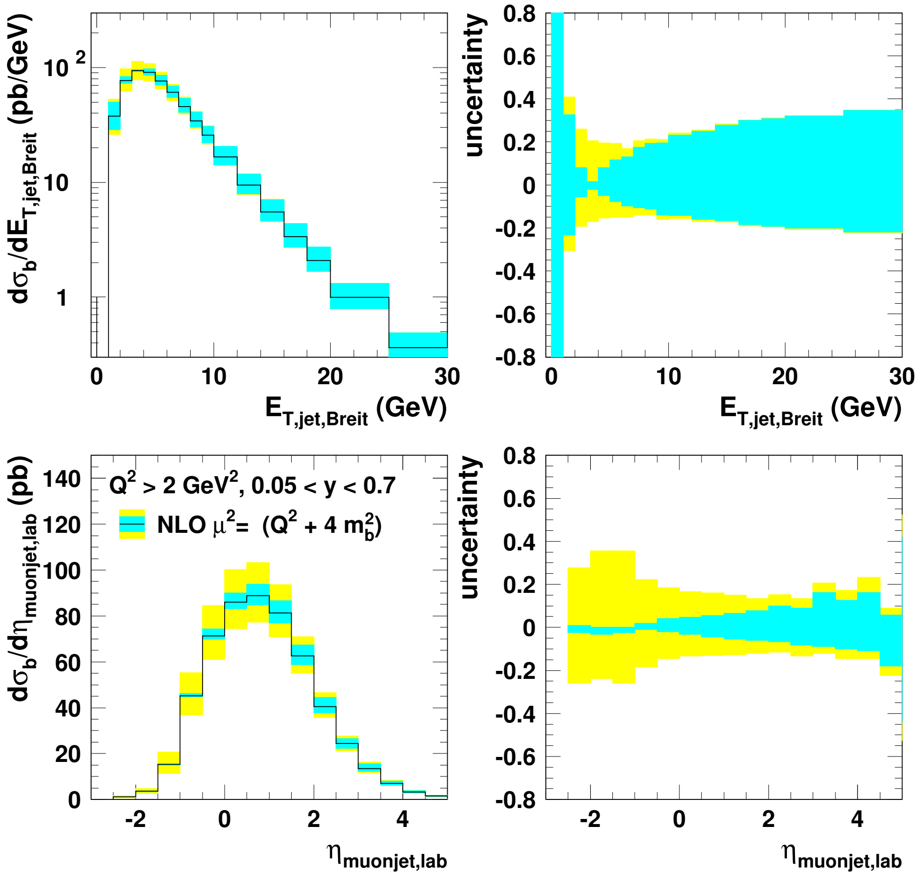

Figure 6.6 Scale and mass uncertainties of the b-quark cross section for the scale choice $\mu^2 = Q^2 + 4m_b^2$. *Top:* b-quark cross section as a function of the jet transverse energy in the Breit frame (*left*) and relative uncertainties(*right*). The inner band represents the uncertainty related to the variation of the renormalization scale, μ while the outer band represents the uncertainty due to the variation of the b-quark mass, m_b. *Bottom:* b-quark cross section as a function of the jet pseudorapidity in the laboratory frame (*left*) and relative uncertainties (*right*).

with the final state containing at least one jet in the Breit frame with

$$E_T^{Breit} > 6 \text{ GeV} \quad \text{and} \quad -2 < \eta^{Lab} < 2.5$$

and one muon satisfying the cuts

$$p_\mu > 2 \text{ GeV} \quad \text{and} \quad 30^\circ < \theta_\mu < 160^\circ.$$

The measured cross section is

$$\boxed{\sigma_{tot}(ep \to e\ b\ \bar{b}\ X \to e\ \mu\ Jet\ X) = 38.7 \pm 7.7\ (\text{stat.})^{+6.1}_{-5.0}\ (\text{syst.})\ \text{pb}}$$

The NLO QCD prediction was evaluated using the `HVQDIS` program and the procedure described in Section 6.1. The fragmentation of b-quarks was performed using the Kartvelishvili function with $\alpha_b = 27.5$ as proposed in Ref. [43] (see Section 1.8.1). The b-quark mass was set to $m_b = 4.75$ and the renormalization and factorization scales to $\mu^2_{R,F} = p_T^2 + 4m_b^2$. The CTEQF4 proton parton distributions were used. The NLO QCD prediction is

$$\boxed{\sigma_{tot}(ep \to e\ b\ \bar{b}\ X \to e\ \mu\ Jet\ X) = 26.7^{+3.3}_{-3.0}\ \text{pb}}$$

The error corresponds to the variation of the scale μ by a factor of 2 and of the mass m_b between 4.5 and 5.0 GeV. The respective variations were added in quadrature. The measured total cross section is a factor ~ 1.4 above the NLO QCD calculation but agrees with it within the experimental and theoretical uncertainties.

Single differential distributions were measured in the same kinematic range as the total cross section. Figure 6.7 shows the differential cross section as a function of Q^2 compared to the NLO calculation. Figure 6.8 shows the differential cross section as a function of the Bjorken variable x, compared to the NLO calculation. In both figures, the shaded bands show the uncertainty of the theoretical prediction coming from the variation of the renormalization and factorization scale, μ, and the b-quark mass, m_b. In both distributions the measured b-quark cross section agrees with the NLO QCD prediction.

6.3 Comparison with Monte Carlo programs

Predictions for the measured cross section have been calculated using Monte Carlo programs which implement different QCD evolution equations. A theoretical introduction to the QCD evolution equations and to their implementation has been given in Sections 1.5 and 1.6.

The results obtained with `HVQDIS` and different Monte Carlo programs are summarized in Table 6.2. The definitions of the b-quark cross section are the same as described in Section 6.1.1. In what follow we will describe how the single predictions are calculated and we will compare them to the measured cross section.

Predictions for a calculation based on LO matrix elements and DGLAP parton shower have been obtained using the `RAPGAP` program. The CTEQ5L proton parton distribution was used and the b-quark mass was fixed to 5 GeV. The renormalization and factorization scales were set to $\mu^2 = p_T^2 + m_b^2$ and the fragmentation of the b-quark into B-hadrons was performed using the Bowler parameterization:

$$f(z) = z^{-(1+bm_T^2)}(1-z)^a \exp(-bm_T^2/z) \tag{6.3}$$

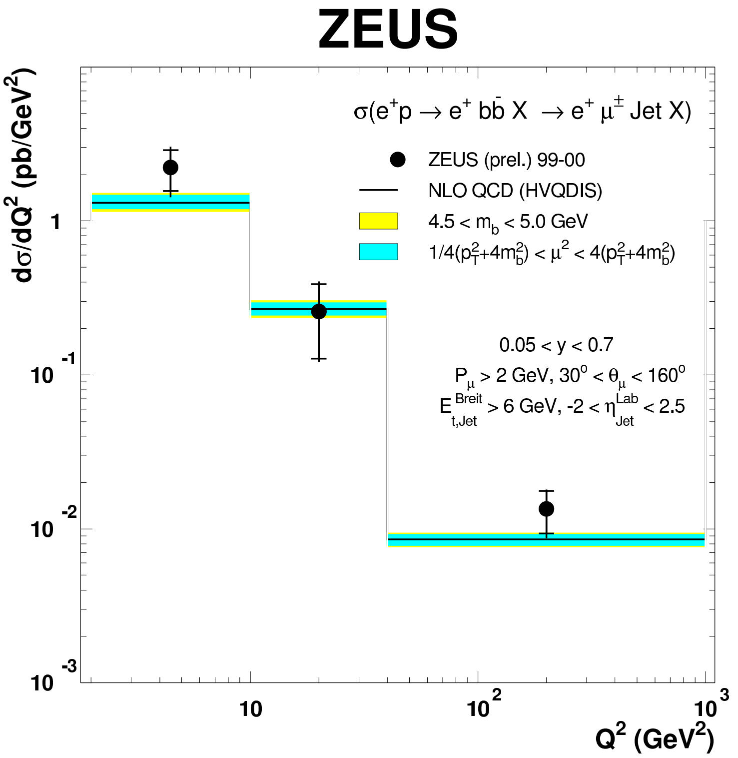

Figure 6.7 Differential beauty quark cross section as a function of Q^2 compared to the NLO QCD calculations. The error bars on the data points correspond to the statistical uncertainty (*inner error bars*) and to the statistical and systematic uncertainty added in quadrature (*outer error bars*). The shaded bands show the uncertainty of the theoretical prediction coming from variation of the renormalization and factorization scale, μ, and the b-quark mass, m_b.

Model	σ_b (pb)	$\sigma_{b,jet}$ (pb)	$\sigma_{b,jet-\mu}$ (pb)
NLO	597.6	230.4	26.1
LO	449.4	113.6	14.4
ME + PS	475.2	164.7	17.1
ME only	475.2	123.3	13.8
CDM	475.2	163.5	16.3
ME* + CCFM	822.0	364.4	36.8

Table 6.2 b-quark cross section predicted by different QCD calculation for several cross section definitions (see Section 6.1.1). The cross sections are calculated using the scale $\mu^2 = p_T^2 + 4m_b^2$ with $m_b = 5$ GeV. The first two rows show `HVQDIS` prediction at NLO and LO. The prediction based on LO matrix element (ME) `RAPGAP` prediction, with or without parton shower (PS), are shown in the third and fourth row. CDM stands for the prediction corresponding to the colour dipole model. The last row shows the `CASCADE` prediction, based on LO ME with CCFM parton shower. ME^* denotes the matrix element where the incoming gluon can be off-shell.

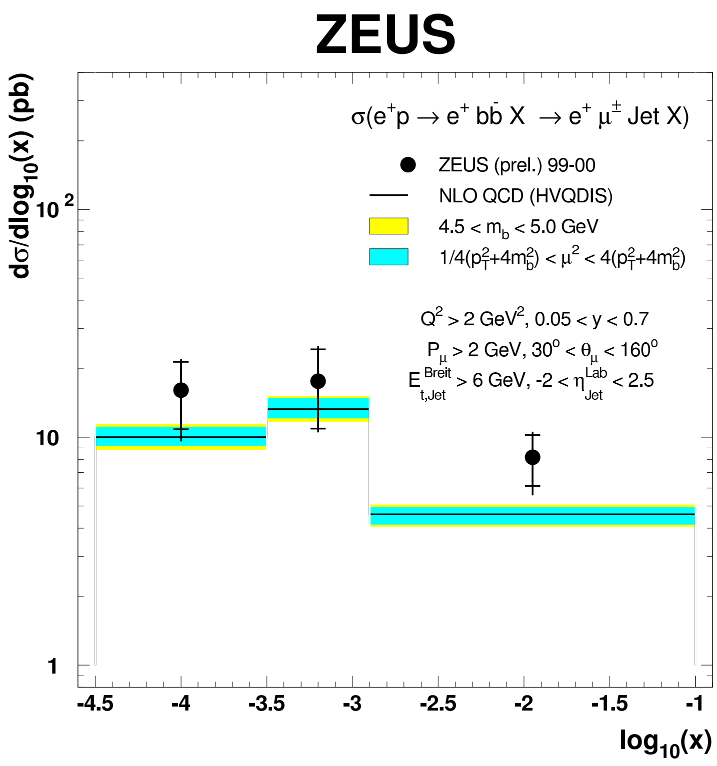

Figure 6.8 Differential beauty quark cross section as a function of the Bjorken variable x (*see the description of Figure 6.7*).

where m_T is the transverse mass[2] of the B-hadron and the parameters have been fixed respectively to $a = 0.3$ and $b = 0.58$. Figure 6.9 and 6.10 show the comparison of the measured differential distribution, respectively as functions of Q^2 and x, compared with the RAPGAP prediction.

A prediction based on the CCFM evolution equation and the k_T-factorization was obtained using the CASCADE program. The program implements only gluon initiated processes. The gluon distribution in the proton were extracted with a CCFM fit to the measurements of the structure function F_2 measured by the H1 Collaboration (see Section 1.6.1). The b-quark mass was fixed to 5 GeV and the renormalization and factorization scales were set to $\mu^2 = p_T^2 + m_b^2$. The Bowler fragmentation was used, with the same parameters used for RAPGAP. The CASCADE prediction for the total cross section is 35 pb, which is in good agreement with the measurement presented here. The comparison of the measured differential cross sections with the CASCADE prediction is shown in Figure 6.9 and 6.10. In both distributions the CASCADE prediction is in good agreement with the measurement. The RAPGAP prediction is always below the data.

[2] The transverse mass of the heavy hadron is defined in the center-of-mass frame of the $q\bar{q}$ pair as $m_T^2 = m^2 + p_x^2 + p_y^2$, where the quark moves in the positive z direction and the antiquark in the $-z$ one.

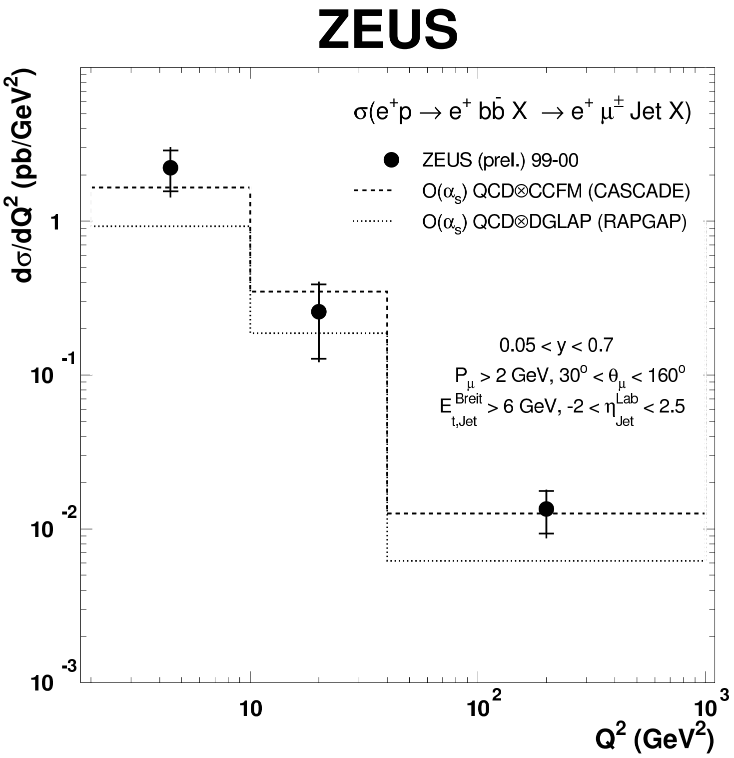

Figure 6.9 Differential beauty quark cross section as a function of Q^2 compare to the LO QCD calculations folded with DGLAP evolution (*dotted line*) and to the LO `CASCADE` prediction with CCFM evolution (*dashed line*). The error bars on the data points correspond to the statistical uncertainty (*inner error bars*) and to the statistical and systematic uncertainty added in quadrature (*outer error bars*).

6.4 Conclusions

The QCD predictions for the measured cross sections were calculated using the NLO program `HVQDIS` and Monte Carlo programs based on LO matrix elements and implementing different QCD evolution equations. The measured cross section was defined by requiring a final state containing the scattered lepton, a muon traversing the inner and outer muon chambers and a hard jet reconstructed in the Breit frame. Single differential distributions of b-quark production in DIS were measured at HERA for the first time.

The measured visible cross section is a a factor ~ 1.4 above the NLO QCD predictions but in agreement within the theoretical and experimental uncertainties. The former, which include the variation of the renormalization and factorization scales and of the b-quark mass, are below 15%.

The prediction of the `CASCADE` Monte Carlo generator, implementing the CCFM QCD evolution equations is in good agreement with the measurement. The prediction of the `RAPGAP` Monte Carlo program is below the data. A detailed comparison of this result with the existing measurements of b-quark production at HERA is given in Section 1.9.

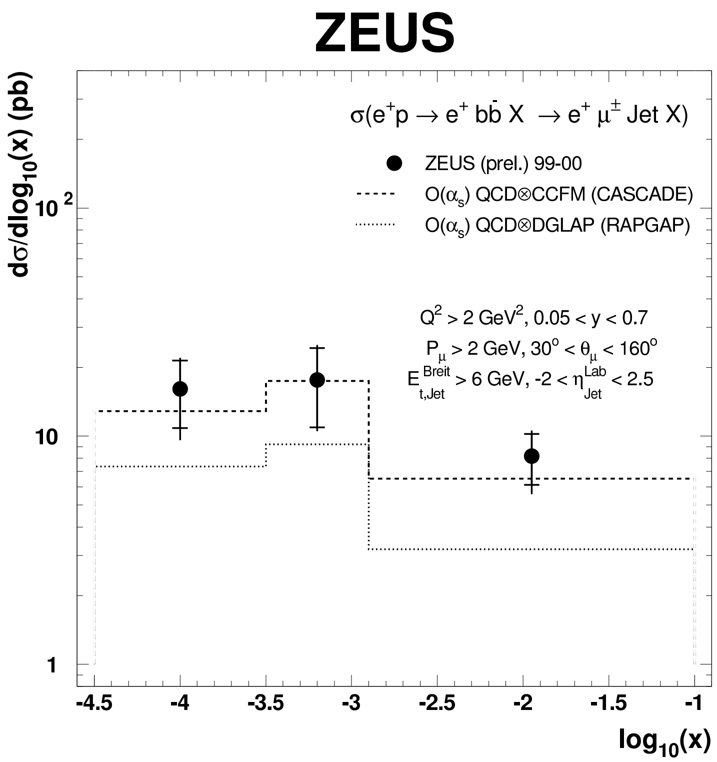

Figure 6.10 Differential beauty quark cross section as a function of the Bjorken variable x (*see the description of Figure 6.9*).

7 HERA II and the ZEUS Micro Vertex Detector

In this Chapter we present the luminosity upgrade of the HERA collider and of the ZEUS detector. The heavy flavour physics potential of the upgraded accelerator is discussed together with a brief description of the HERA machine. The upgraded ZEUS detector includes a new silicon Micro Vertex Detector (MVD) which will enhance the ability to identify displaced vertices. In the second part of this Chapter the MVD is described and the test beam measurements performed with the MVD half modules are presented. We conclude describing the measurements performed with a MVD full module.

7.1 The HERA luminosity upgrade

The data accumulated by the HERA experiments to date have mainly been used to explore the structure of the proton and to make new and precise tests of QCD. During a long shutdown started in May 2000 the HERA collider has been modified in order to deliver more luminosity for the ZEUS and H1 experiments. This is achieved by substantial modifications of the interaction regions, in order to focus the beams more strongly. Beam polarization and higher luminosity will give more inside on electroweak physics. Furthermore, new important tests of strong interactions will be possible.

A review of the physics goals of the next HERA run is collected in Ref. [106]. In this section we will discuss some selected topics of heavy flavour production in the light of a large integrated luminosity and describe the main features of the new HERA interaction regions.

7.1.1 Prospects of heavy flavour physics at HERA II

In Section 1.3.2 the limit of the heavy quark structure function at high virtualities of the exchanged photon was discussed. In the asymptotic coefficient functions logarithmic terms $\ln(Q^2/m^2)$ appear, where m is the mass of the heavy quark. The terms at high values Q^2 could possibly spoil the convergence of the perturbative expansion. This problem lead to the introduction of the variable flavour number scheme (VFNS) [107] in which at high Q^2 the charm quark is described as a light quark, i.e. as a constituent parton of the proton, whereas at small Q^2 (of the order m^2) the boson-gluon fusion mechanism, in which the charm quark can only be excited by a hard scattering, describes the production process.

This picture is in apparent contradiction with the findings. Both ZEUS and H1 have measured the inclusive production of D^* mesons and have translated this into a measurement of the charm contribution to the proton structure function, F_2^c. The data are well described by the fixed order (FO) NLO QCD calculation, in which the charm quark is not an active parton in the proton.

Recently the VFNS calculations have been carried up to the next-to-next to-leading order (NNLO) and have been compared to the exact NLO coefficient functions in Ref. [108]. The result is that the differences between the two VFNS calculation schemes on one hand and between the VFNS and the FO NLO approach on the other hand are both hardly noticeable (see Figure 7.1). The main conclusion is that the resummation effect is very small which means that the logarithms of the type $\ln(Q^2/m^2)$ do not vitiate the perturbative series [7]. In the future, the high luminosity run will allow measurements of charm production at very high Q^2 and more precise comparison will be performed.

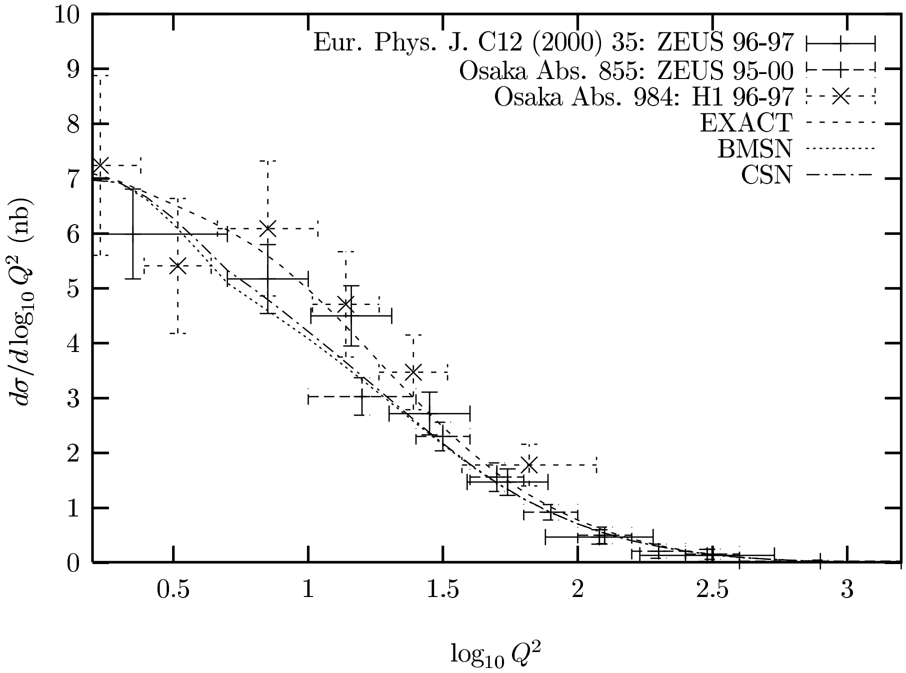

Figure 7.1 Differential cross section of D^* meson production as a function of Q^2 measured by the ZEUS and H1 Collaborations. The data points are compared with a FO QCD NLO calculation (`HVQDIS`, *dashed line*) and with two different calculation schemes based on the VFNS approach (BMSN-scheme [109], *dotted line* and CSN-scheme [110], *dot-dashed line.*).

On the basis that charm production is dominated by boson-gluon-fusion the measurement of the structure function F_2^c is also an excellent probe to extract the gluon density in the proton at small x values. In Figure 7.2 we show the gluon density extracted from the inclusive D^* production by the H1 Collaboration [111]. Data from photoproduction at a scale $\mu = 25\ \text{GeV}^2$ are plotted together with those from DIS at $Q^2 = 25\ \text{GeV}^2$ and are compared to a QCD fit to F_2. In view of a theoretically well-behaved observable and these first experimental results a more precise measurement of the gluon density can be expected.

While the production mechanism of charm quarks seems to be well described by FO NLO QCD the situation for b-quarks is still unclear. In Section 1.9 we have reviewed the actual status of the experimental results and their comparison to the theoretical predictions. In some cases, sizeable discrepancies are observed between the measured cross sections and the NLO QCD predictions. On the other hand, for the cross sections defined without extrapolations to

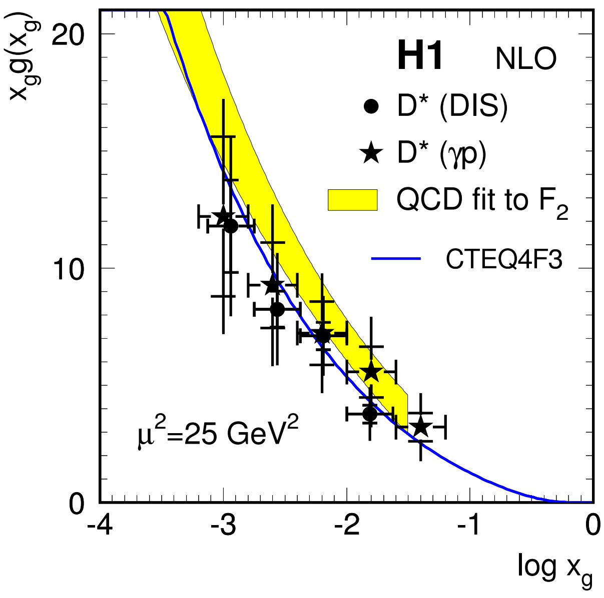

Figure 7.2 Gluon density extracted in DIS (*full circles*) and in photoproduction (*full starts*) by the H1 Collaboration. Both results are compared to a H1 QCD analysis of the inclusive F_2 measurement at a scale $\mu = 25\ \mathrm{GeV}^2$ (*light dashed band*) and to the `CTEQ4F3` parameterization [111].

unmeasured regions of the phase space, the excess is reduced. The actual ZEUS measurements, both in the photoproduction and DIS regions, rely on the statistical extraction of the signal using the semi-leptonic b-decays into muons. In the future, tagging of heavy mesons will be possible thanks to the newly installed silicon Micro Vertex Detector. This will provide an additional and independent selection tool for heavy flavour analyses reducing the systematic uncertainties. The increased tagging capabilities and the high luminosity will also reduce the statistical errors.

Figure 7.3 shows the anticipated result of a measurement of the ratio of the beauty contribution to the proton structure function, F_2^b over F_2^c assuming a luminosity of 500 pb^{-1} [112]. The error bars show the statistical and the total experimental uncertainties of the anticipated measurement. For a given Q^2 or x the ratio is expected to raise with decreasing x or increasing Q^2, respectively. Integrated over the kinematic range a mean value of $F_2^b/F_2^c \sim 0.02$ is predicted.

7.1.2 HERA luminosity constraints and the new interaction region

The luminosity of the HERA ep collider can be written as

$$L = \frac{N_e N_p N_B f_{rev}}{2\pi(\sqrt{\sigma_{xp}^2 + \sigma_{xe}^2} + \sqrt{\sigma_{yp}^2 + \sigma_{ye}^2})} \tag{7.1}$$

where N_p is the number of protons per bunch, N_e is the number of leptons per bunch, N_B is the number of bunches, f_{rev} is the revolution frequency, $\sigma_{x,y,e,p}$ are the beam sizes at the

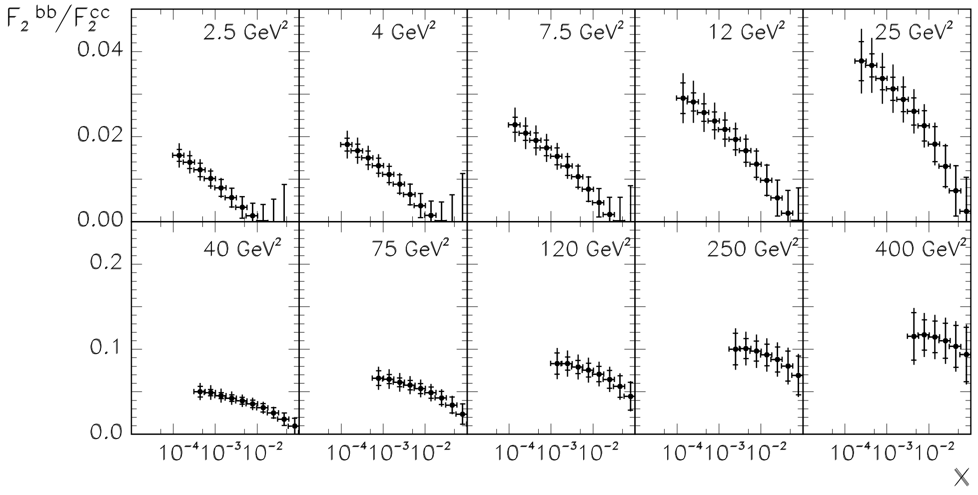

Figure 7.3 Expected ratio of F_2^b to F_2^c for a luminosity of 500 pb^{-1}. The inner(outer) error bars represent the statistical(total) experimental uncertainty of the anticipated measurement [112].

interaction point (IP) for protons and leptons respectively. The following constraints must be taken in account:

- The beam cross section of protons and leptons have to match to reduce the non-linear effects of the beam-beam interaction;
- The total current of the lepton beam is limited by the available RF power. This problem cannot be overcome by reducing the beam energy, since the necessity of longitudinally polarized beams sets a lower limit on the beam energy;
- The number of protons per bunch is limited by space charge effects in the injector chain;
- Since the two beams have to be separated after collisions before focusing elements for the proton beam can be placed, the beam size at the IP is expected to be limited by the maximum β-function of the protons at the low β quadrupole around the interaction region (IR).

These consideration lead to the conclusion that the most promising way to increase the HERA luminosity is to reduce the β-functions at the IP by a new IR geometry with the proton and lepton low-β quadrupoles closer to the IP [113].

The accelerator redesign incorporates machine magnets inside the detectors and new types of septum magnets to decrease the proton β-functions to $\beta_x^* = 2.45$ m and $\beta_y^* = 0.18$ m at the IP. Improvements factors of up to five compared to the HERA-I design value and up to three compared to the 1997 running conditions can be reached with this configuration . The expected instantaneous luminosity is $7 \cdot 10^{31}$ cm^{-2}s^{-1}.

A schematic top view of the interaction region is shown in Figure 7.4. The IP in the middle of the drawing is surrounded by superconducting combined function magnets which separate the two beams by an 8 mrad kick and which provide vertical focusing to the lepton beam. The strong synchrotron radiation generated by these magnets pass through the detector region and the downstream magnets and is absorbed at 11 m, 19 m and 25 m downstream the IP. The next elements are two low-β quadrupoles which are aligned on the design orbit of the lepton beam. Behind these magnets, at 11 m from the IP, the two beams are sufficiently separated, to place strong septum quadrupoles which provides focusing in the vertical plane for the proton beam. Current septum quadrupoles complete the vertical focusing of the proton beam. Horizontal focusing starts at 30 m from the IP using conventional quadrupoles.

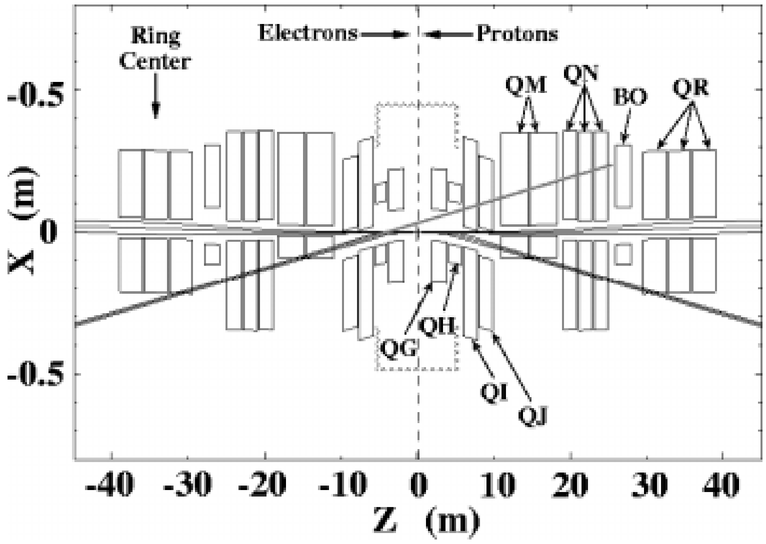

Figure 7.4 Schematic top view of the new HERA interaction region

7.2 The ZEUS Micro Vertex Detector

Together with the upgrade of the HERA collider, new devices have been realized for the ZEUS and H1 detectors. In particular, the ZEUS experiment has been equipped with a new silicon Micro Vertex Dectector (MVD) and a Straw Tube Tracker (STT) in the forward region. The latter replaces the TRD detector and it is conceived to improve the pattern recognition and the track reconstruction. The information of both detectors will be combined already at the Second Level Trigger. In this Section, after a brief introduction to the experimental techniques of heavy meson tagging, the Micro Vertex Detector will be described in details.

7.2.1 Tagging of heavy meson decays

In order to measure the decay of heavy mesons their production and decay vertices have to be determined with an accuracy sufficient to resolve the decay length. The decay lengths of some heavy mesons is shown in Table 7.1. For lifetime measurements of heavy quarks two methods are commonly applied:

- *Reconstruction of the secondary vertex:* Both the production and decay vertex are reconstructed in the events using charged tracks. The distance between the two vertices d is related to the particle lifetime τ by the expression

$$d = \beta\gamma c\tau. \tag{7.2}$$

Meson	Mass (MeV)	$c\tau$ (μm)
$D^{\pm}$	1869.3±0.5	315
D^{0}	1864.5±0.5	123.7
$D_s^{\pm}$	1968.5±0.6	148.6
$B^{\pm}$	5279.0±0.5	496
B^{0}	5279.4±0.5	464
B_s^{0}	5369.6±2.4	448

Table 7.1 Masses and decay lengths of some heavy mesons

The method requires at least two charged particles created at both the primary and secondary vertex. In addition the method depends strongly on the momentum of the decaying particle.

- *Impact parameter method:* the impact parameter is defined as the orthogonally closest distance between a particle track and the primary vertex in the $r-\phi$ plane. For instance, the impact parameter δ of a muon from a heavy quark decay is related to the decay length by the expression

$$\delta = d \sin(\theta) \sin(\alpha) \tag{7.3}$$

where θ is the polar angle of the decaying hadron and α is the angle between the direction of the hadron and the direction of the muon in the $r-\phi$ plane. This method has already been used in *ep* collisions by the H1 Collaboration, as discussed in Section 1.9.

An advantage of the impact parameter method is the weak dependence of δ on the momentum of the decaying hadron. Since $\sin(\alpha) \sim 1/\gamma \sim 1/p$ the dependence cancels with the decay length d in Eq. 7.2.

7.2.2 Detector specifications

Taking into account the constraints of available space and costs the ZEUS Collaboration has designed and realized a Micro Vertex Detector (MVD) on the basis of the following requirements [114]:

- polar angular coverage between 10° and 150°;
- measurement of three points per track, in two projections each;
- at least 20 μm hit resolution for normal incident tracks;
- two-tracks separation of 200 μm.
- hit efficiency better than 95%.
- noise occupancy better than 10^{-3}.

The proposed layout is a 3-layer device surrounding the beampipe. The choice of three, rather than two, layers is dictated by considerations of efficiency, pattern recognition, and some expectation of using a momentum estimate in the trigger.

The expected resolution on the impact parameter was extracted from a Monte Carlo simulation of the MVD [115]. Figure 7.5 shows the resolution on the impact parameters σ_r and σ_z as a function of the track momentum for tracks perpendicular to the beamline ($\eta = 0$). The impact parameter σ_r is the distance of closest approach of the reconstructed track at the IP in the projection onto the $r - \phi$ plane whereas σ_z is the distance along the z-axis between the IP and the intersection point of the track and the beamline in the projection onto the track plane. The points in the plot are obtained requiring tracks traversing all three layers of the detector. The simulation is based on the distribution of material corresponding to the proposed MVD layout.

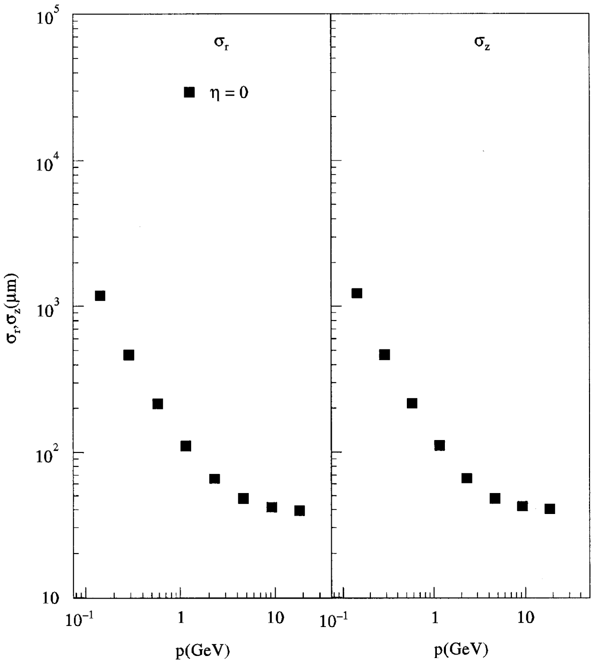

Figure 7.5 Resolution on the impact parameter σ_r and σ_z as a function of the track momentum, for tracks perpendicular to the beam line ($\eta = 0$).

The forward wheels (see Section 7.2.3) are constrained mainly by the length of the interaction region, which sets the closest position at $z \sim 40$ cm. This implies that the wheels are not well suited to impact parameter determination at the 100 μm level. On the other hand, the wheels will provide essential tracking and vertexing information in the regions not well covered by existing tracking detectors.

The expected MVD tagging efficiency for charm quarks is estimated between 10% and 30% with a purity greater than 30%[114]. Better performance on beauty quarks can be expected considering the larger decay length of the b-mesons.

7.2.3 Mechanical design

The Micro Vertex Detector is divided in two regions: a *barrel region* (BMVD) and a *forward region* (FMVD).

The length of the barrel region, 640 mm, is constrained by the longitudinal distribution of the interaction region which at HERA is dominated by the length of the proton bunch ($\sim$ 10 cm). The BMVD consists of three double layers of silicon strip sensors, arranged in concentric planes around the interaction point, with a polar coverage between 30° and 150°. Cross sections of the BMVD across the beam line and across the xy plane are shown in Figure 7.6. The first layer of the detectors, placed at $3.6 \div 4.5$ cm from the detector axis, follows the elliptical shape of the beam pipe. The second and third layers are placed at 8.6 and 12.3 cm from the detector axis.

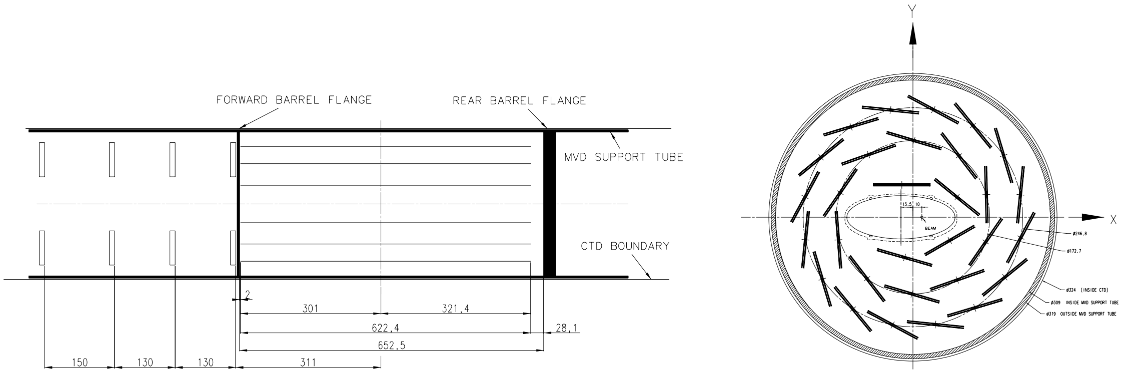

Figure 7.6 *On the left:* Schematic layout of the MVD along the beam direction. *On the right:* Cross section of the BMVD across the xy plane. In the drawing only the detector layers are shown, without the ladder supporting structure.

The BMVD silicon sensors are square shaped ($64.2 \times 64,2$ mm^2) and 320 μm thick. A more detailed description of the sensor layout follows in Section 7.2.4. In the barrel region two sensors are glued together so that the strips are perpendicular to each other. One sensor is electrically connected to the other one via copper trace etched on 50 μm thick Upilex[1] foil (see Fig. 7.7). The connection of the sensor assembly with the front end readout is also made via a Upilex foil, glued at one side on the sensor and at the other side on the hybrid. The resulting surface of 123.68×64.24 mm^2 forms one readout cell of 512 channels and is called a *half module* [116].

Two readout cells are placed on top of each other forming a *full module* of 1024 channels and providing two coordinates for a track traversing it. Five full modules are fixed side by side, with 1 mm gap in between, on a support frame and form a *ladder* where the silicon covers an area of 64.2×622 mm^2. The ladder is made of a pitch based carbon-fiber and polycarbonate resin, with fibers of 60 μm in diameter [117]. This material ensures a strong and lightweight construction. The ladder has also to provide the support for the hybrids, cabling and cooling for the front-end electronics.

The forward MVD extends the detector coverage to a pseudorapidity of $\eta = 2.6$. It consists of four planes (*wheels*) perpendicular to the beam axis each of them made of two layers of 14 silicon sensors of the same type as in the barrel section but with a trapezoidal shape. The wheels are positioned at $z \sim 31$, 44, 57 and 70 cm from the nominal interaction point. In each wheel the two layers are parallel but strips are tilted by 180°/14 in opposite directions.

[1]Polyimid film produced by UBE Technologies Ltd, Japan

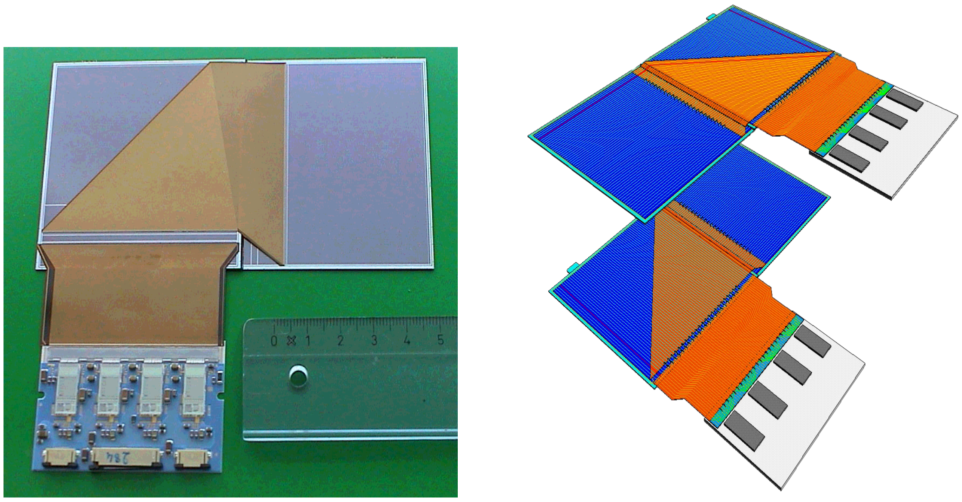

Figure 7.7 *On the left:* photo of a BMVD half module, showing the two single sensors electrically connected via the Upilex foil. The half module is equipped with a hybrid circuit on which the four read-out chips are mounted. *On the right:* Two half modules are placed on top of each other forming a full module.

The support tube is an assembly of light weight cylinders connected to each other via flanges: two flanges at the rear and forward end of the support tube and other two flanges in the barrel section supporting both ends of the ladders.

7.2.4 Layout of the silicon sensors

The sensors used for the MVD are single-sided, AC coupled, silicon strip detectors[2] with p^+ strips implanted into a n-type bulk. The front side is covered with 3082 p^+ strips. Every sixth strip is AC coupled by a SiO_2-Si_3N_4 dielectric to an aluminum read-out strip (see Figure 7.8). The width of the intermediate strips is 12 μm whereas the read-out strips are 14 μm wide. The strip pitch is 20 μm corresponding to a read-out pitch of 120 μm.

Capacitive charge division between read-out strips limits the number of read-out channels and yields a spatial hit resolution in the range $10 \div 15$ μm for tracks perpendicular to the detector surface.

The biasing of the strips is implemented using polysilicon resistors ($\sim 1.5 - 2.5$ MΩ) connected to the sensor ground line (called *biasing ring*), located alternatively on both ends of the strips. The first and last strips close the biasing ring, being directly connected to it. Three p^+ guard rings, designed to adjust the potential towards the detector edges, surround the sensitive area (see Figure 7.9). An additional n^+ doped implant beyond the last guard ring allows to bias the backplane with a contact from the top.

A schematic layout of a BMVD sensor is shown in Figure 7.10. A probe pad, with an area of 75×100 μm^2 is located close to each polysilicon resistor, at both ends of each read-out strip. The BMVD sensors have also three rows of bond pads, with a pad area of 75×250 μm^2, located at 8, 10 and 19 mm from one sensor edge. The bond pads allow the electrical connection to

[2]produced by Hamamatsu Photonics.

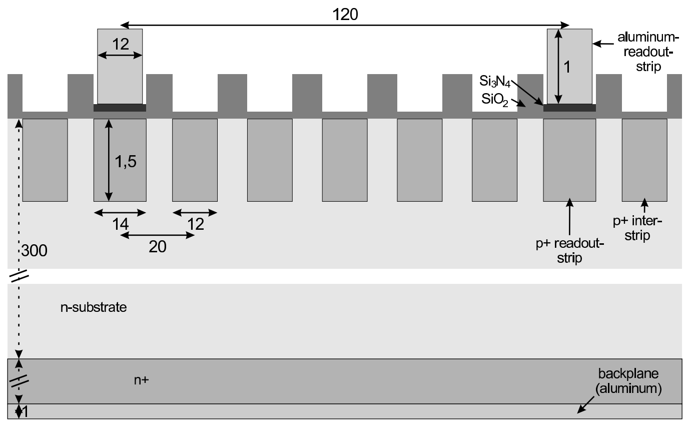

Figure 7.8 Cross section of a MVD silicon sensor (not to scale, all dimensions are in μm [118]).

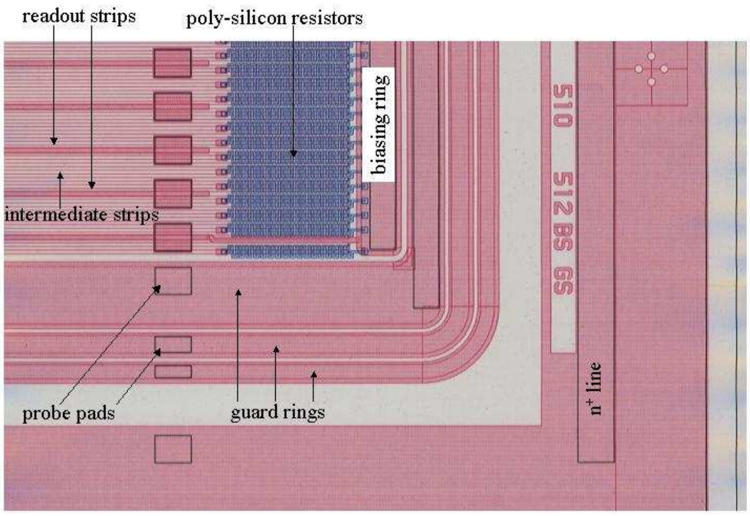

Figure 7.9 Layout of one corner of the BMVD sensors (courtesy of Hamamatsu Photonics K.K.).

the other sensor of a half module or to the front-end electronics via the Upilex foils.

The FMVD sensors have a trapezoidal shape with the tilted edges inclined by $\sim 13°$ with respect to the height. The p^+ strips run parallel to one tilted side and have a varying strip length. The FMVD sensors have only one row of bond pads perpendicular to the read-out strips and therefore to one of the tilted edges of the sensor.

A more detailed description of the MVD sensors and their electrical properties before and after irradiation can be found in Ref. [118].

7.2.5 Readout electronics and data acquisition

The MVD silicon sensors are readout by the `HELIX 3.0` chip [119], a successor of the `HELIX 2.2` version designed by ASIC laboratories in Heidelberg and manufactured in the 0.8 μm

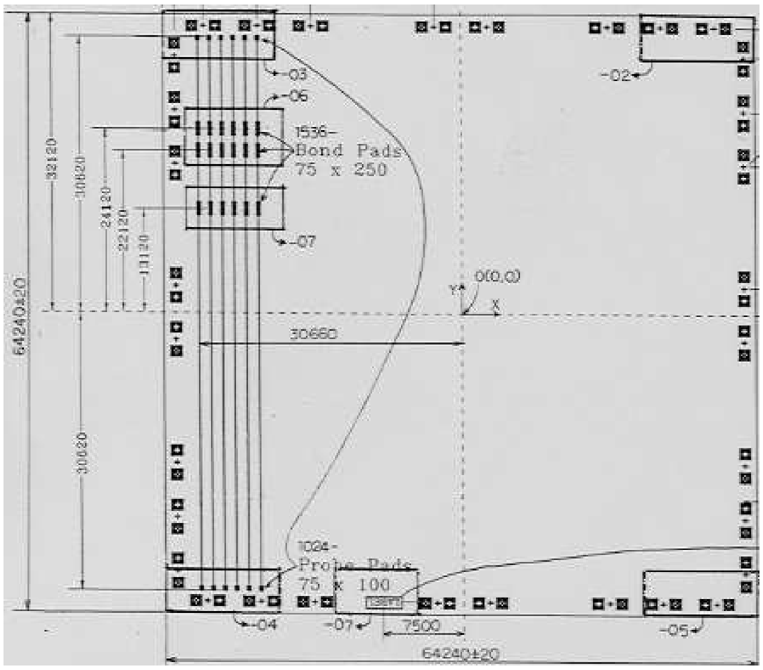

Figure 7.10 Layout of a BMVD sensor. All dimensions are in micron (courtesy of Hamamatsu Photonics K.K.).

CMOS process by AMS. The chip was modified for ZEUS in collaboration with NIKHEF. It contains 128 channels, each one consisting of

- a low noise charge sensitive preamplifier yielding a gain of 11.2 mV/MIP (1 MIP$=24\cdot 10^3$ electrons);
- a shaper forming a semi-Gaussian with a peak time of $50 \div 70$ ns;
- a buffer amplifier driving the internal pipeline write line and the pipeline cell;
- a comparator circuit indicating hit channels;
- an analog pipeline consisting of 128+8+5 capacitors resulting in a maximum latency of 128 events and a multi-event buffer (FIFO) capable of storing up to eight triggered events. The signals are sampled in the analog pipeline waiting for the GFLT trigger decision (see Section 2.8.1);
- a switched charge sensitive pipeline amplifier.
- a cascaded 128+8+1 channel multiplexer and a current output buffer close the chain. The latter is a differential voltage input/single current output amplifier that converts the voltage signals delivered by the multiplexer.

The chip data output is shown in Figure 7.11. The 128 amplifier channels are output channel 0 up to channel 127. The analog data are followed immediately by a 8 bit trailer showing the pipeline column number the event has been stored in.

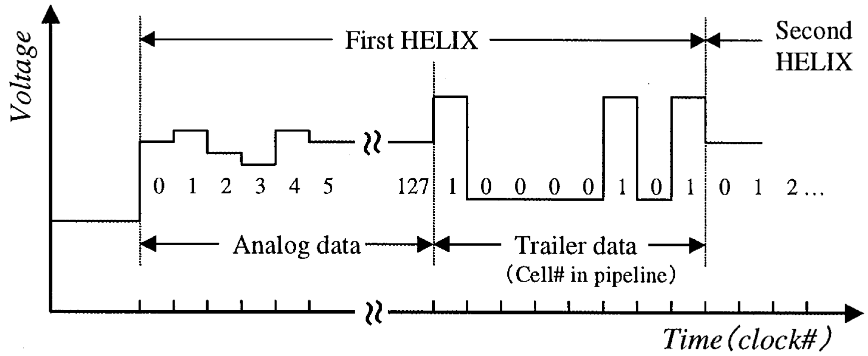

Figure 7.11 Output signal of the `HELIX` chip.

The noise performance of the `HELIX` depends on the input capacitance, C, and is $ENC = 400 + 40 \cdot C$ pF equivalent noise charge (ENC). Although the `HELIX` chip is not realized in a radiation hard technology, the layout has been optimized to be radiation tolerant. Results of irradiation tests on `HELIX` chips can be found in Ref. [120].

The bias voltages and current for the different amplifier stages can be adjusted via programming of digital-to-analog register in order to select the most appropriate operation point and to compensate for radiation damages. The `HELIX` chips can synchronize the data transfer with other chips using a daisy-chain mechanism. In addition, the `HELIX 3.0` version implements a failsafe version of the desy chain which allows to exclude a bad chip from the readout chain, without interrupting the functioning of the whole chain.

The output of the `HELIX` daisy chain is amplified and transferred by an analog link board, and fed to the 10 bit ADC module [121]. At the input of the ADC module, the signal range is 0-2 V, which corresponds to 0-10 MIP. The total number of read-out strips amounts to about 200 000. The ADCs perform the common noise and pedestal subtraction. Then, strips with charge greater than threshold values are kept for cluster finding. A further threshold can be set on the total charge of the cluster for further data reduction. The optimal values of the ADC thresholds were investigated using a MVD full module; the results are reported in Section 7.4.1.

The ADC modules can provide three output data streams: the *raw data*, the *strip data*, which consist of strip I.D. and the ADC values of the strips above threshold, and the *cluster data*, which contain the total charge, the location of the cluster and the charge collected on the first and last two strips in the cluster. The read-out is performed via VME Power PC boards running Lynx OS.

As a trigger component, the MVD has stimulated the realization of a new Global Tracking Trigger (GTT), based on the combined information of the Micro Vertex Detector with the existing CTD and the newly installed forward Straw Tube Tracker. The GTT will provide efficient trigger and reconstruction capability already at the Second Level Trigger [122].

7.3 Beam test measurements on MVD half modules

A test beam program has been set up to understand the performance of our silicon detectors and front-end electronics. Results on single sensors have been presented in Ref. [123]. Half modules of the final production chain have been exposed to an electron beam of 2-6 GeV at the DESY-II accelerator.

A drawing of the testbeam setup is shown in Figure 7.12. The charged tracks are measured by three pairs of silicon strip detectors with 50 μm readout pitch and one intermediate strip, providing a position accuracy of 5.4 μm for a 6 GeV beam. The detector to be studied is mounted between the telescope modules on linear and rotational positioners which allow to investigate the performance in different areas of the detector and for different angles of incidence. The rotations can be around the strip axis (θ angle) or around the axis perpendicular to the strip in the detector plane (ϕ angle) (see Figure 7.12). A trigger was generated by coincidence of the signals from scintillator fingers located at both end sides of the optical bench.

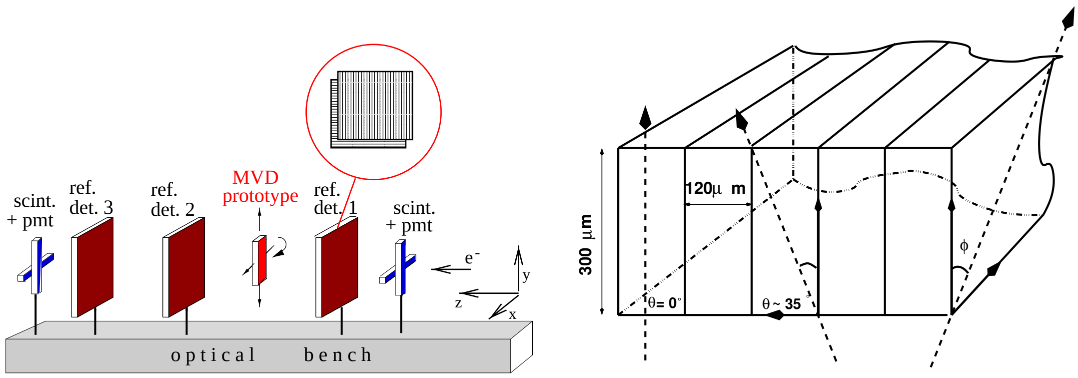

Figure 7.12 *On the left:* layout of the test beam setup. *On the right:* Definition of the impact angles θ and ϕ.

The digitized ADC output coming from the MVD detector and the telescope detectors is stored in files during the data taking. The channel noise and pedestal levels are measured using special random trigger runs of 100-200 events taken without beam. During the offline analysis, the common mode noise and the pedestal levels are subtracted from the data [124]. The pedestal, determined once per day, has shown negligible variation over time. The variation of the pedestals within one readout chip, from the first to the last readout channel, has been found to be of the order of the cluster pulse height. The strip noise was stable and showed uniform behaviour; its variation within regions read by the same chip are much smaller than those observed between chips ($\sim$ 20%). The common mode noise is Gaussian distributed with a r.m.s. comparable to the single strip noise level.

This section is structured as follows:

- In Section 7.3.1 we show the measurement of charge collection in a half module as a function of the bias voltage;
- In Section 7.3.2 the algorithms used for the reconstruction of the hit position are presented;

- In Section 7.3.3 and 7.3.4 we show respectively the results of the measurements performed on the overlap between single sensors and a detailed study of the effects related to the probe and bond pads. In all the measurements the *eta algorithm* is used to reconstruct the hit position. As we will show in Section 7.3.5, this method gives the best performance for tracks perpendicular to the sensor. In this section we emphasize the dependence of this reconstruction method on the charge sharing and discuss the related systematic errors.

- An overview of the performance of the algorithms described in Section 7.3.2 is given is Section 7.3.5 as a function of the incidence angles. The eta algorithm is compared with other methods which use the charge collected on three consecutive strips (instead of two). In each case the hit resolution and the dependence of the reconstruction on the inter-strip position is discussed.

7.3.1 Performance as a function of the bias voltage

The dependence of the collected charge as a function of the bias potential, V_{bias}, was investigated with the half module. Figure 7.13 shows the most probable value of the energy loss distribution as a function of the bias voltage and measured on the two sensors of an half module. We call x-detector the sensor directly connected to the front-end electronics whereas the y-detector is the other sensor of the half module. The y-detector is glued to the x-detector and electrically connected to it via a Upilex foil. The signal was defined as the sum of the ADC counts of the strips with charge greater than 3σ, where σ is the r.m.s. noise of each strip. The most probable value of the energy loss distribution was defined as the peak value from a fit by a Landau function.

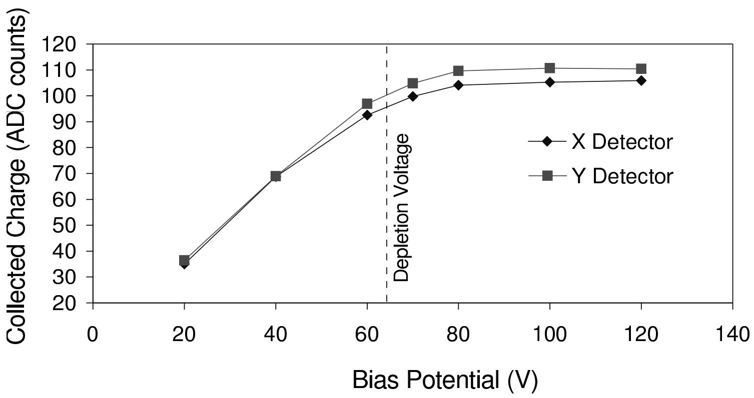

Figure 7.13 Most probable value of the charge loss distribution as a function of the bias voltage measured on the two sensors of a half module.

For values of V_{bias} above the depletion voltage ($V_{dep} \sim 65$ V) a plateau is reached. The signal collection degrades with decreasing V_{bias}. The signal collected on the x-detector is $\sim 5\%$ higher than on the y-detector. The effect is related to the additional capacity introduced by the Upilex foil and seen by the y-detector.

Figure 7.14 the shows the signal over noise ratio (S/N) as a function of the bias voltage. Values of $S/N \sim 15$ are observed when the detector is overdepleted by a factor ~ 1.5.

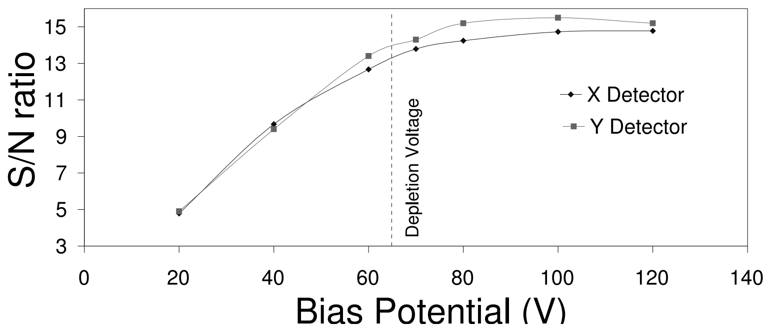

Figure 7.14 Signal over noise ratio as a function of the bias voltage measured on the two sensors of a half module.

7.3.2 Reconstruction of the hit position

The charge created by a traversing particle is collected by several consecutive strips, which form a *cluster*. In order to find the cluster we first look for the strip whose charge is higher then a given threshold. As we will discuss in Section 7.3.5 the cluster size for perpendicular tracks is between 2 and 3 strips. The threshold was defined as 5σ, where σ is the r.m.s. noise of each strip. The reconstruction of the impact position has been performed with several algorithms [125], i.e. (1) the Eta algorithm [126, 127, 128], (2) the Double Centroid algorithm, (3) the Gaussian algorithm, (4) the Head–Tail algorithm and (5) the Double Centroid algorithm with four strips.

1. *Eta algorithm*: is based on the assumption that most of the charge is collected by the two readout strips closest to the impact position. If $Q_{left}(Q_{right})$ is the charge collected on the left(right) side of the impact point, P the readout pitch and x_{left} the position of the left strip, the hit position x_{Eta} can be reconstructed by

$$x_{Eta} = P \cdot f(\eta) + x_{left} \tag{7.4}$$

with:

$$\eta = \frac{Q_{right}}{Q_{right} + Q_{left}}, \quad f(\eta) = \frac{1}{N}\int_0^{\eta} \frac{dN}{d\eta'}\, d\eta'. \tag{7.5}$$

In Eq. 7.5 η is calculated for each hit, while the probability function $f(\eta)$ is evaluated using all hits of each data sample.

2. *Double Centroid algorithm (DC)*: reconstructs the impact position using the strip with the highest collected charge and its two neighbouring strips. Having computed the center of gravity $C_{left}(C_{right})$ between the central strip and the left(right) neighbouring strip, the hit position is given by

$$x_{DC} = \frac{C_{left}/dr + C_{right}/dl}{dr + dl}, \quad dl = \frac{Q_{left}}{Q_{right}} = 1/dr. \tag{7.6}$$

3. *Gaussian algorithm (G)*: the impact position is calculated using the strip with the highest collected charge and its two neighbouring strips. A Gaussian interpolation is performed using the three strips and the center of the Gaussian distribution is defined as the impact position.

4. *Head–Tail algorithm (HT)*: for large incident angles, the charge generated by a particle spreads among many strips. In this case, the charges collected by central strips in the cluster do not contain position information. The Head–Tail algorithm uses the information from strips on both edges to solve this problem. We define the head(tail) strip as a strip which has the largest(smallest) strip number and collected charge higher than three times the noise level (see Figure 7.15-a). The hit position x_{HT} is then given by
$$x_{HT} = \frac{x_h + x_t}{2} + \frac{Q_t + Q_h}{2Q_{AV}} \cdot P \tag{7.7}$$
where $x_h(x_t)$ is the position of the head(tail) strip and $Q_h(Q_t)$ its charge. The average pulse height per strip within the cluster is Q_{AV}.

5. *Double Centroid algorithm with four strips*: this algorithm was newly developed. It uses the information from strips on both edges as the HT algorithm. In this case, the charge collected on the last two strips on both sides is used. For each couple of strips the center of gravity, $pos_{l(r)}$ is calculated with the expression
$$pos_{l(r)} = \frac{Q_{l(r)} \cdot strip_{l(r)} + Q_{l+1(r-1)} \cdot strip_{l+1(r-1)}}{Q_{l(r)} + Q_{l+1(r-1)}}, \tag{7.8}$$
where $Q_{l(r)}$ is the charge collected on the out most left(right) strip and $strip_{l(r)}$ the respective position (see Figure 7.15-b). The hit position is given by
$$x_{DC4S} = \frac{pos_l/dr + pos_r/dl}{dr + dl}, \tag{7.9}$$
where
$$dl = \frac{Q_l + Q_{l+1}}{Q_{r-1} + Q_r} = \frac{1}{dr}. \tag{7.10}$$

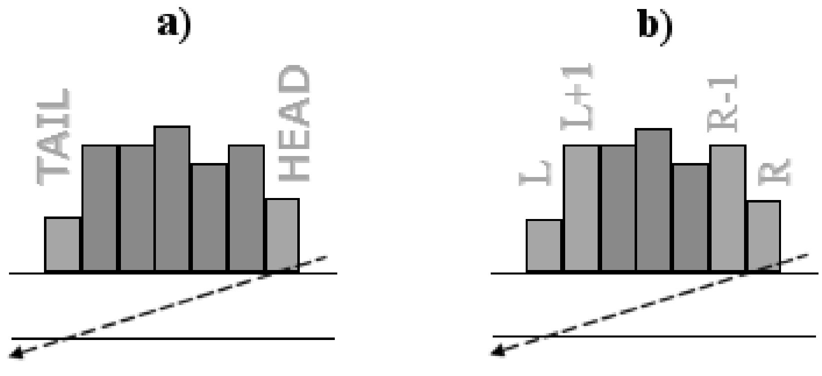

Figure 7.15 Hit reconstruction for a track traversing the detector at large values of the impact angle θ. The Head–Tail algorithm a) and the Double Centroid with four strips b) are shown.

For each hit on the detector under test (DUT) a predicted position is given by the reference telescopes. The *residual* is defined as

$$\delta = x_{pred} - x_{meas}, \tag{7.11}$$

where x_{pred} and x_{meas} are respectively the predicted and reconstructed position. The measured resolution, σ_{res}, is defined as the width of the residual distribution and is given by

$$\sigma_{res}^2 = \left(\sigma_{DUT}^{intr}\right)^2 + k \cdot \left(\sigma_T^{intr}\right)^2 + \sum_i k_i \cdot \Delta\theta_{m.s.}^2, \tag{7.12}$$

where σ_{DUT}^{intr} is the *intrinsic resolution* of the DUT, σ_T^{intr} is the intrinsic resolution of the telescopes, k and k_i are geometrical factors related to the relative distances between the telescope modules and the last term is the extrapolation error due to the multiple scattering along the particle direction. The calculation of the second and third terms on the r.h.s. of Eq. 7.12 for our setup is discussed in Ref. [129]. The *intrinsic resolution* of a reconstruction algorithm is thus defined as the spatial precision independent from environmental conditions, e.g. the setup geometry, the beam energy and so on.

7.3.3 Performance in the overlap between sensors

The two sensors forming a half module are glued together as shown in Figure 7.16. A Cirlex bar (5.8×65.44, 0.4 mm thick) glued at the edge in between the sensors, forms the mechanical connection between the two. The use of this material assures a good isolation of the HV side of one sensor from the ground plane of the other one [117]. While the sensors overlap by 4.8 mm, the sensitive areas only overlap by 2.76 mm.

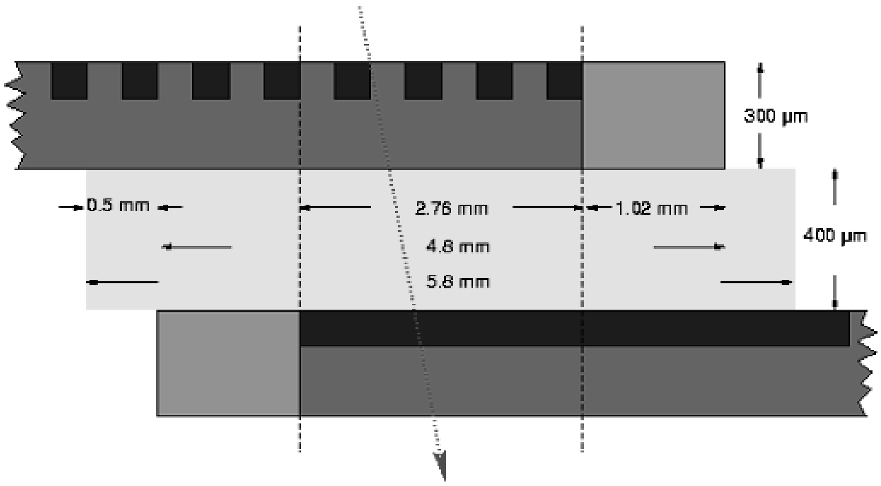

Figure 7.16 Cross section of the overlap region between two sensors in a half module. A Cirlex bar is glued at the edge in between the sensors.

On the upper sensor, the sensitive area in the overlap is covered by 23 readout strips. The strips are parallel to the long side of the Cirlex bar. On the lower sensor the readout strips are orthogonal to the upper ones. The strip implantations are covered with the polysilicon resistors and with a row of probe pads (see Figure 7.9).

The detector performances in the overlap region were measured and compared to other areas of the half module. Several runs were recorded using a narrow beamspot ($\sim 3 \times 9$ mm^2) hiting the overlap region and a wider beamspot ($\sim 9 \times 9$ mm^2) hiting the regions far from the overlap and from the bond or probe pads.

Figure 7.17 shows the hit pattern on the overlap region of the upper sensor. The black dots represent the hits reconstructed by the upper sensor whereas the grey dots are the respective hit positions predicted by the telescopes. Since the upper sensor measures only the x-coordinate, the y-coordinate of the black dots is the one predicted by the telescopes. The overlap region of the detector sensitive areas is delimited by the two solid lines. The dashed line shows the position of the lower sensor edge. A similar pattern is reconstructed on the lower sensor, where the y-coordinate of each hit is measured.

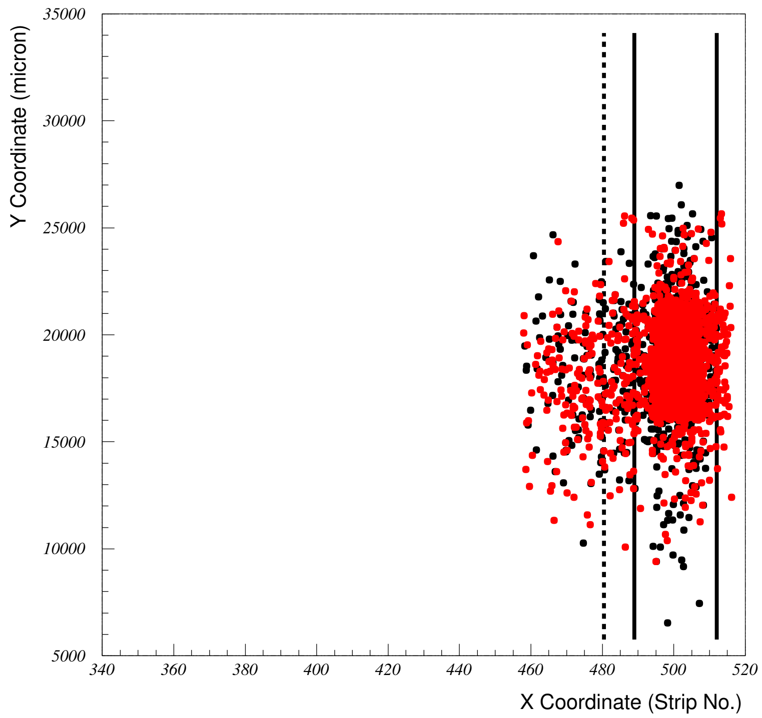

Figure 7.17 Hit pattern reconstructed in the overlap region. The black dots represent the hits reconstructed on the half module whereas the grey dots are the positions predicted by the telescopes. The overlap region of the active sensors is delimited by solid lines. The dashed line shows the position of the lower detector edge.

The hit position was reconstructed using the eta algorithm. Figure 7.18 shows the η-distribution and probability function $f(\eta)$ (see Eq. 7.5) measured on the upper and lower sensors of the half module. The shape of the η-distribution is different for the two sensors. In the upper sensor the distribution is almost flat for $0.2 < \eta < 0.8$ whereas in the lower sensor two peaks at $\eta \sim 0.2$ and $\eta \sim 0.8$ are observed. The different shape of the distribution is due to the presence of the probe pads on the lower sensor. The effect has been clarified with a set of dedicated measurements which are discussed in Section 7.3.4.

In principle, the characteristic shape of η-distribution in the lower sensor could also be related to the gluing of the Cirlex bar on the active surface of the sensor. This possibility is excluded by our observations. Figure 7.19 shows the η-distribution measured on the lower sensor in the region covered by the Cirlex bar (between 0 and 2 mm from the active edge of the sensor) and compared with the distribution measured outside the overlap (between 4 and 6 mm from the active edge of the sensor). A good agreement between the two distribution is observed.

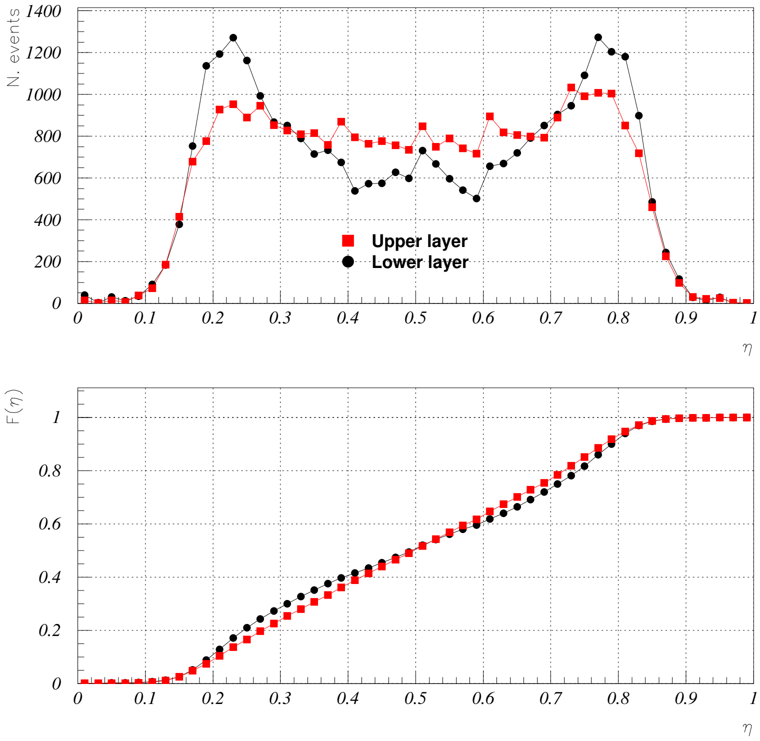

Figure 7.18 η-distribution of the hits reconstructed in the overlap region (*upper plot*). The distribution measured on the upper (*full squares*) and lower (*full circles*) sensors are shown separately. The respective probability functions are shown in the lower plot.

The hit positions were calculated for each sensor using the respective probability functions. The residual widths measured on the upper and lower sensors are respectively (14.6 ± 0.3) μm and (15.1 ± 0.3) μm whereas the value measured outside the overlap region is ~ 12 μm. The values of the residual widths measured in the overlap region are in agreement with the expectation obtained using Eq. 7.12 and including the additional multiple scattering contribution introduced by the Cirlex bar and the second silicon sensor.

For the lower sensor, the width of the residual distribution was also measured as a function of the distance from the end of the active area. The measurement was performed using a wide beamspot ($\sim 9 \times 9$ mm^2) centered on the overlap region. The beamspot size allows to analyze in the same experimental conditions areas in and outside the overlap region. A high statistics run was recorded to achieve an acceptable number of events in all areas. The results are summarized in Table 7.2. All distribution have a Gaussian shape and are well centered at

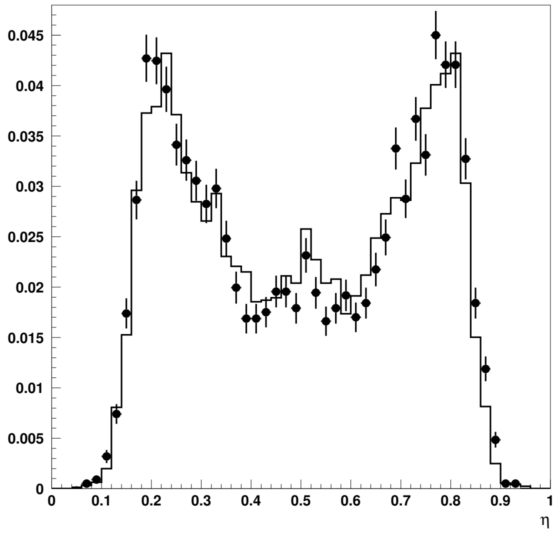

Figure 7.19 η-distribution of the hits reconstructed in a region covered (*solid line*) and uncovered (*full circles*) by the Cirlex bar. The distribution are normalized to the number of events.

Range (mm)	Residuals width
$[0-2]$	(15.2 ± 0.2) μm
$[2-4]$	(14.5 ± 0.2) μm
$[4-6]$	(13.8 ± 0.2) μm

Table 7.2 Widths of the residual distributions as a function of the distance from the edge of the sensor active area.

zero. The residual widths decrease with increasing distance from the edge of the active area. This is due to the decreasing multiple scattering contribution as we move out of the overlap region.

Using the same data sample the dependence of the charge collection was measured as a function of the distance from the edge of the active area. In Table 7.3 the most probable value of the charge loss distribution is shown as a function of the distance from the active edge. The total charge was defined as the sum of the ADC counts of the two highest strips and the most probable value as the peak value from a fit by a Landau function. We observe a decrease of the collected charge towards the detector edge. The overall difference is of the order of 5%.

7.3.4 Effects related to probe and bond pads

Due to the effects in the charge division observed during the measurement in the overlap region, detailed studies were performed in the areas close to the probe and bond pads. The

Range (mm)	Collected charge (ADC counts)
$[0-1]$	114.8 ± 0.4
$[1-2]$	115.9 ± 0.4
$[2-3]$	117.0 ± 0.4
$[3-4]$	118.1 ± 0.4
$[4-6]$	119.5 ± 0.4

Table 7.3 Most probable value of the charge loss distribution as a function of the distance from the edge of the sensor active area.

cross section of a BMVD sensor in a region with no pads is shown in Figure 7.20-a. This layout is described in details in Section 7.2.4. In Figure 7.20-b the correspondent electric diagram is shown. Each read-out strips is coupled to the front-end electronics through a *coupling capacitance*, $C_c \simeq 26$ pF/cm, and to its neighboring strips through a *inter-strip capacitance*, $C_c \simeq 1$ pF/cm. An additional capacitance, not shown in the diagram, couples each strip to the backplane.

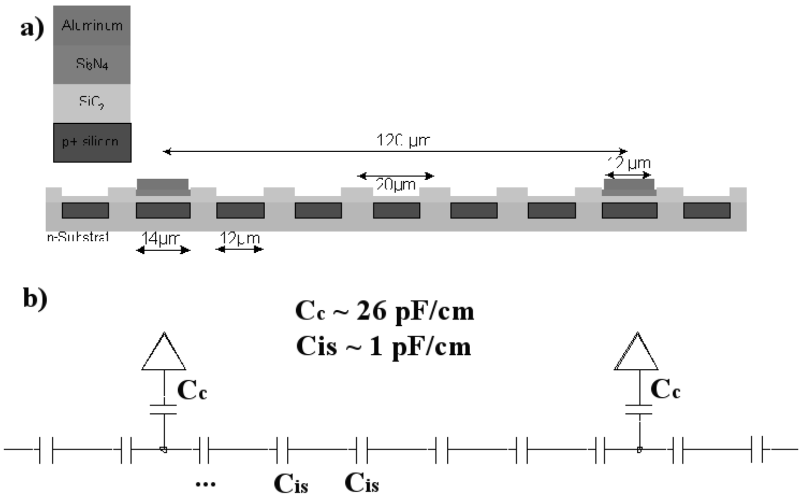

Figure 7.20 Cross section of a BMVD sensor. The schematic layout (a) and the correspondent electric diagram (b) are shown.

The cross section of a BMVD sensor across a bond (or probe) pad is shown in Figure 7.21-a. The pad is 75 μm wide and it covers the read-out strip, the first neighboring strip and part of the second strip. The correspondent electrical diagram is shown in Figure 7.21-b. In this case, two more coupling capacitances are placed in parallel to the usual one and couple the first neighboring strips to the aluminum pad. The second neighboring strip is only partially covered by the pad. The estimated coupling capacitance is $C_c' \simeq 1/3 \cdot C_c$.

Since the coupling capacitance is different in the two regions effects on the charge division between strips might be observed. A set of runs recorded using the half module and moving the

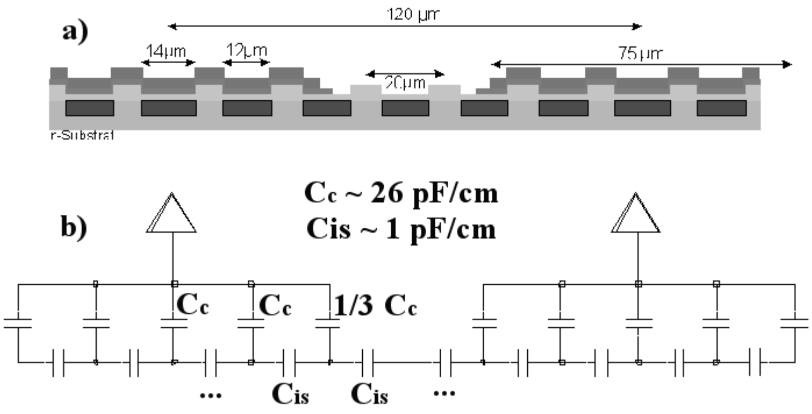

Figure 7.21 Cross section of a BMVD sensor across the bond pad. The schematic layout (a) and the correspondent electric diagram (b) are shown.

Run	Range (μm)	Events	Note
3095	$[0, 7144]$	10000	Probe pads
3116	$[3144, 12133]$	5000	-
3114	$[12144, 21144]$	10000	-
3090	$[15144, 24144]$	5000	-
3098	$[26144, 35144]$	5000	-
3109	$[30144, 39144]$	7000	-
3103	$[39144, 48144]$	15000	Bond pads (1^{st} row)
3106	$[48144, 57144]$	30000	Bond pads (2^{nd} and 3^{rd} rows)
3115	$[53144, 61240]$	10000	Bond pads (3^{rd} row) and Probe pads

Table 7.4 Summary of the runs recorded moving the beam spot along the read-out strips. The coordinate refers to the sensitive area of the sensor and measures the distance along the strips moving upwards in Figure 7.10.

6 GeV beam spot along the read-out strips of a sensor. Data were recorded on the y-detector of the half module using a 9×9 mm^2 beam spot. The analyzed runs are summarized in Table 7.4. The coordinate which measures the position along the strip has its origin at the beginning of the strip implantation at the bottom side of the sensor as in Figure 7.10. In this frame the position of the bond pad rows are respectively at 43.74 mm (*first row*), 52.74 mm (*second row*) and 54.74 mm (*third row*).

The η-distribution measured in different regions of the sensor is shown in Figure 7.22. The distributions in Figure 7.22-a,b are measured in a region in which the strips are not covered by probe or bond pads. The distribution in Figure 7.22-c corresponds to the first row of bond pads whereas Figure 7.22-d corresponds to the second and third rows. The distribution a) and b) have a similar shape while the distribution c) and d) show two peaks at $\eta = 0.2$ and $\eta = 0.8$. The corresponding probability function is shown in Figure 7.23. The largest differences between the curves correspond to the peaks in the η-distribution.

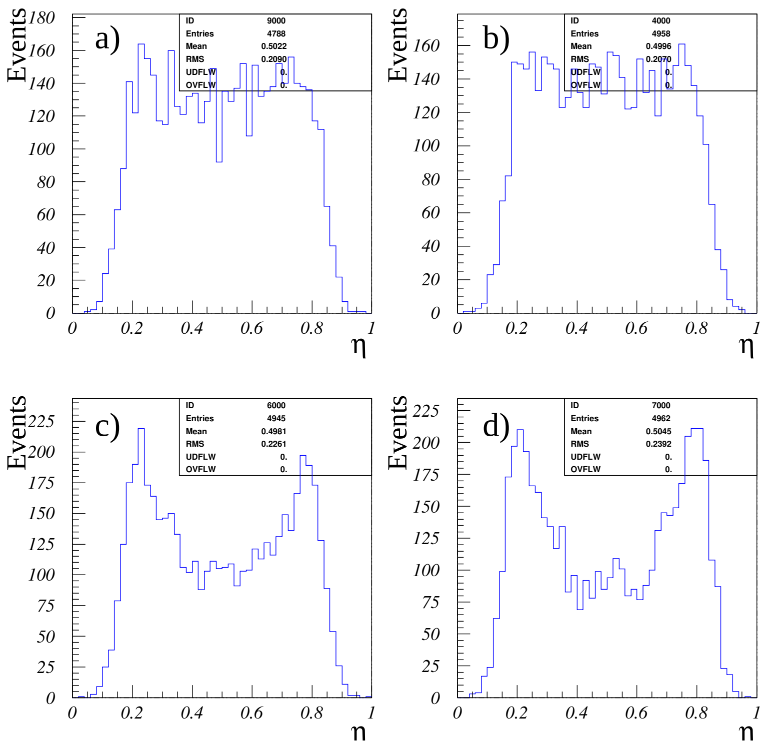

Figure 7.22 η-distribution measured at different positions along the read-out strip. a) 20 mm, b) 31 mm, c) 44 mm and d) 53 mm. The coordinate refers to the sensitive area of the sensor and measures the distance along the strips moving upwards in Figure 7.10. The positions a) and b) are not covered by probe or bond pads. The position c) corresponds to the first row of bond pads whereas the position d) corresponds to the second and third row of bond pads.

The measured distribution were compared with a simulation of the BMVD sensors described in Ref. [130]. A comparison of the measured and simulated η-distribution is shown in Figure 7.24 for two different regions of the sensor. The simulation reproduces the observed enhancement of the peaks in the region covered by bond pads although the simulated peak position does not correspond to the measured one. This is probably due to additional contributions to the cross-talk between read-out strips introduced by the electrical connections and the front-end electronics, which are not present in the simulation.

The variation of the η-distribution must be taken in account when reconstructing the hit position with the η algorithm. The reconstructed position, x_{Eta}, is directly proportional to the value of the probability density $f(\eta)$ (see Eq. 7.4) which can vary in different regions of the detector, depending on the value of η. Here we should stress that η is an observable measured for each hit using the charge collected on the read-out strips. On the other hand, the probability function $f(\eta)$ is calculated integrating the η-distribution over many events and it is, a priori, unknown. Therefore, a probability function measured in a certain detector region should be used only to reconstruct hits in that region.

The effect has been quantified by measuring the hit position at a fixed value of $\eta = 0.29$ as a function of the longitudinal strip position (see Figure 7.25). For each bin in the longitudinal

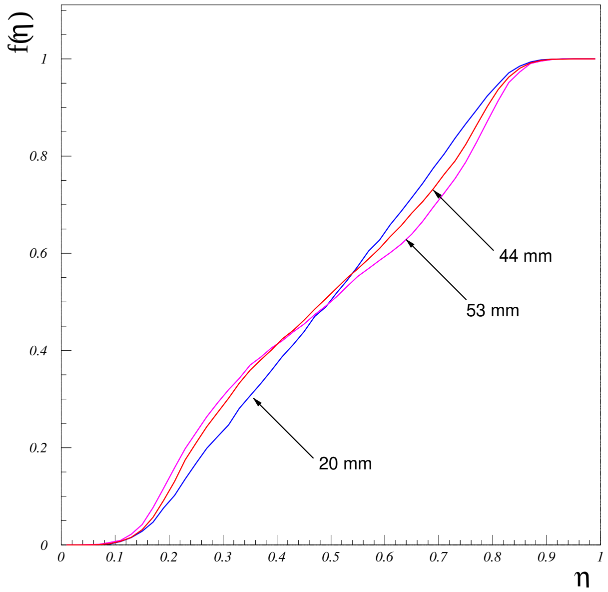

Figure 7.23 Probability function $f(\eta)$ measured at different position along the read-out strips.

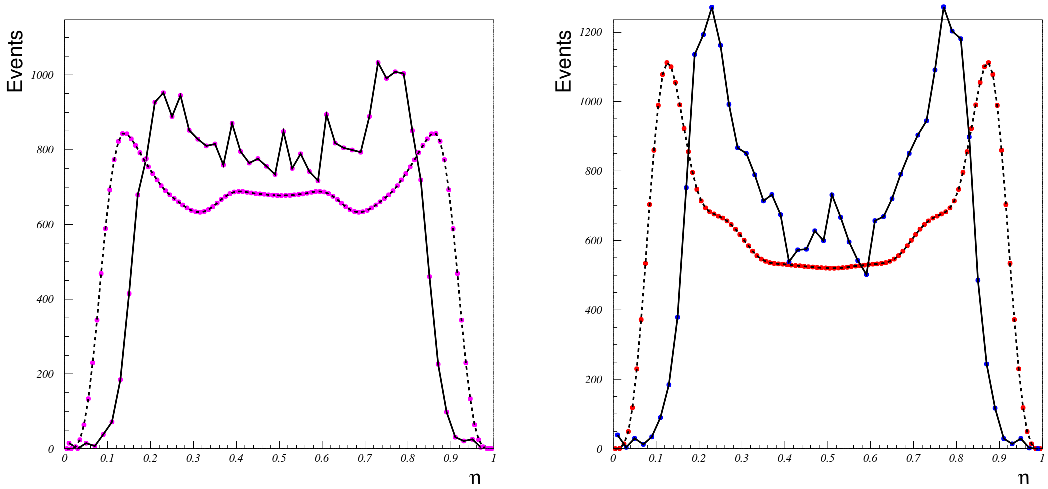

Figure 7.24 η-distribution measured on two different regions of the BMVD sensors. *On the left:* η-distribution measured in a region withoud pads (*solid line*) and compared with a simulation of the BMVD sensors (*dashed line*). *On the right:* η-distribution measured in a region with bond pads (*solid line*) and compared with a simulation of the BMVD sensors (*dashed line*). The distributions are normalized to the number of events.

strip position the probability function was extracted and the hit position calculated using Eq. 7.4 with $x_{left} = 0$. The reconstructed position varies between 27 and 36 μm depending on the longitudinal hit position. Furthermore, it is almost flat between 6 and 40 mm were the read-out strips are not covered with probe or bond pads. For longitudinal positions greater than 40 mm (or lower than 6 mm) the reconstructed position changes in relation to the presence of bond pad rows (or probe pad rows). The position of the probe/bond pad rows are indicated by vertical lines. The measured points are compared with the result of the PSPICE simulation of the BMVD sensors. The data points generally agree with the simulation. Some discrepancy between the measurement and the simulation is observed at the edges of the sensor.

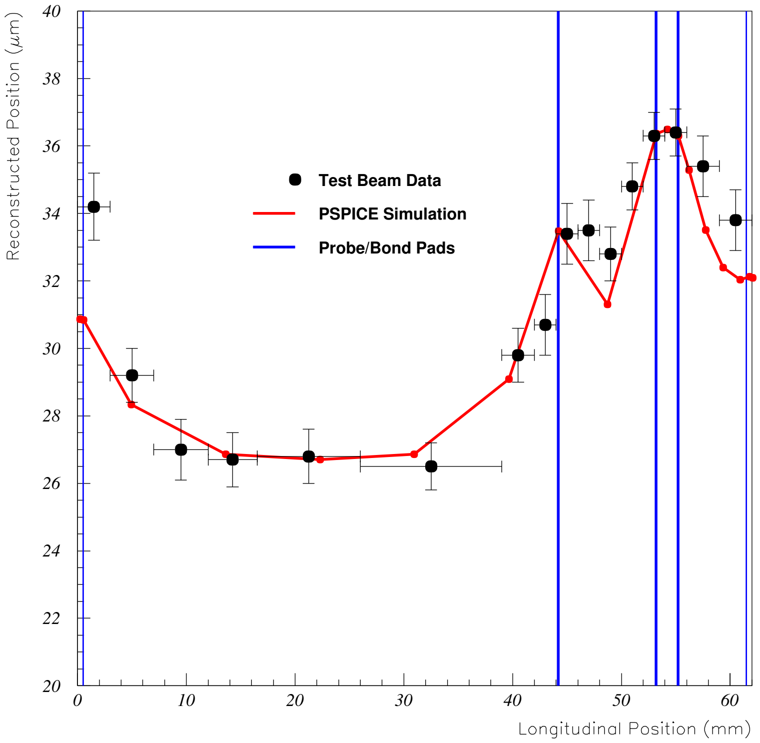

Figure 7.25 Reconstructed hit position as a function of the longitudinal strip position for hits with $\eta = 0.29$ (*full circles*). The measurement is compared with the result of a PSPICE simulation of the BMVD sensor (*solid line*). The position of the probe and bond pads is indicated by the vertical lines in the plot.

From this measurements we can conclude that a maximum systematic errors of the order of 10 μm can be made if the hit position is reconstructed with the Eta algorithm and the wrong probability function is used. This is the case for hits in which $20 \div 30$ % of the deposited charge is collected by one of the read-out strips (corresponding to $\eta \sim 0.2 \div 0.3$). The position reconstruction for hits in which the charge is equally shared between two strips ($\eta \simeq 0.5$) or in which the whole charge is collected in one strip ($\eta \simeq 0, 1$) is less dependent on the detector region.

Run	θ (degrees)	Events	Region
3354	0	5000	No pads
3360	5	10000	No pads
3355	15	10000	No pads
3358	25	10000	No pads
3359	35	10000	No pads
3370	45	10000	No pads
3371	55	10000	No pads
3363	0	5000	Bond pad region
3367	5	10000	Bond pad region
3364	15	10000	Bond pad region
3365	25	10000	Bond pad region
3366	35	10000	Bond pad region
3374	45	10000	Bond pad region
3372	55	10000	Bond pad region

Table 7.5 Summary of the runs recorded at different values of the impact angle θ. Runs marked with *"No pads"* are taken in a region of the detector centered between the probe pads and the first row of bond pads. Runs marked with *"Bond pad region"* are taken over the second and third bond pad rows. The energy of the beam was set to 6 GeV in all runs.

Run	ϕ (degrees)	Events
3450	0	5000
3458	15	5000
3451	30	10000
3459	40	5000
3456	50	5000

Table 7.6 Summary of the runs recorded at different values of the impact angle ϕ. The energy of the beam was set to 6 GeV in all runs.

7.3.5 Performance as a function of the impact angle

Several runs were recorded with the half module at different values of the impact angles θ and ϕ (see Figure 7.12) and in different regions of the sensor. The beam energy was set to 6 GeV in most of the runs and the beam spot size to 9×9 mm^2. The runs at different values of θ were taken placing the beam spot on the y-detector of the half module whereas the runs at different values of ϕ were taken on the x-detector. The analyzed runs are summarized in Table 7.5 and 7.6.

For increasing values of the impact angle ϕ the deposited charge increases as $1/\cos(\phi)$ and the cluster size is constant. The signal over noise ratio measured at different impact angles ϕ is shown in Figure 7.26 for clusters of two and three strips. The S/N for perpendicular tracks is ~ 13.5 and ~ 15 respectively for two and three strip clusters.

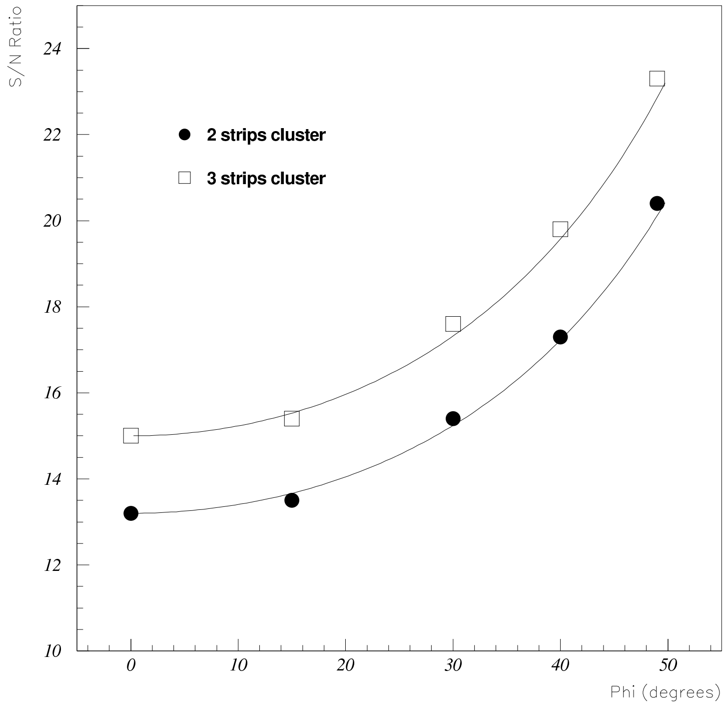

Figure 7.26 Signal over noise ratio as a function of the impact angle ϕ for cluster of two strips (*full circles*) and three strips (*open squares*). The curve represents the function $S(\phi) = S_0/\cos(\phi)$, where S_0 is the S/N ratio measured at $\phi = 0$.

As the impact angle θ increases, the deposited charge is collected by more strips. In Figure 7.27 the distribution of the cluster size is shown for $\theta = 15°$, $25°$ and $35°$ degrees. For small impact angle, say $\theta \leq 30°$, a cluster is typically composed by two or three strips and algorithms as the Eta, Double Centroid and Gaussian can be used to reconstruct the impact position. The Eta algorithm uses the two highest strips whereas both the Double Centroid and the Gaussian algorithms also use the third highest strip in the cluster.

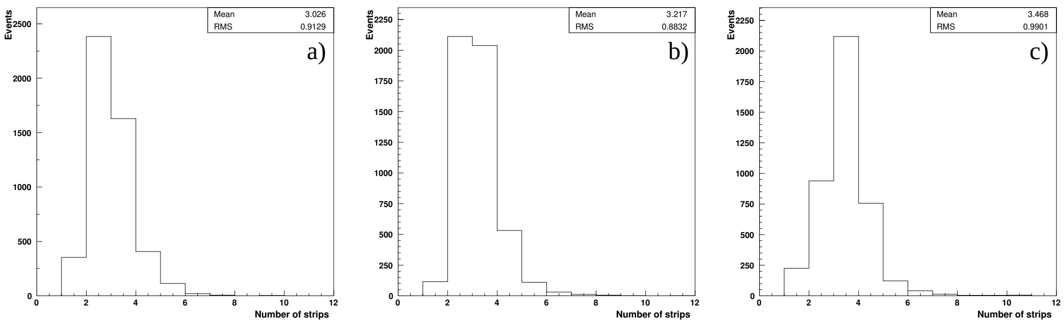

Figure 7.27 Distribution of the cluster size for $\theta = 25°$ (a), $\theta = 25°$ (b) and $\theta = 35°$ (c)

The distribution of the charge collected on the first, second and third highest strip in a cluster is shown in Figure 7.28 as a function of the hit distance from the highest strip. The distributions are measured with tracks normal to the sensor surface. As the modulus of the distance from the highest strip increases the charge collected on the first strip decreases and more charge is collected on the second strip. If the hit position is right in-between two readout strips the first and second strips collect the same amount of charge. The charge collected on the third strip is also clearly correlated to the hit position. If a readout strip is hit, i.e. when the distance is zero in the plot, the first strip collects most of the charge whereas the second and third equally share the remaining deposited charge. If the modulus of the distance increases the charge collected on the third strip will decrease.

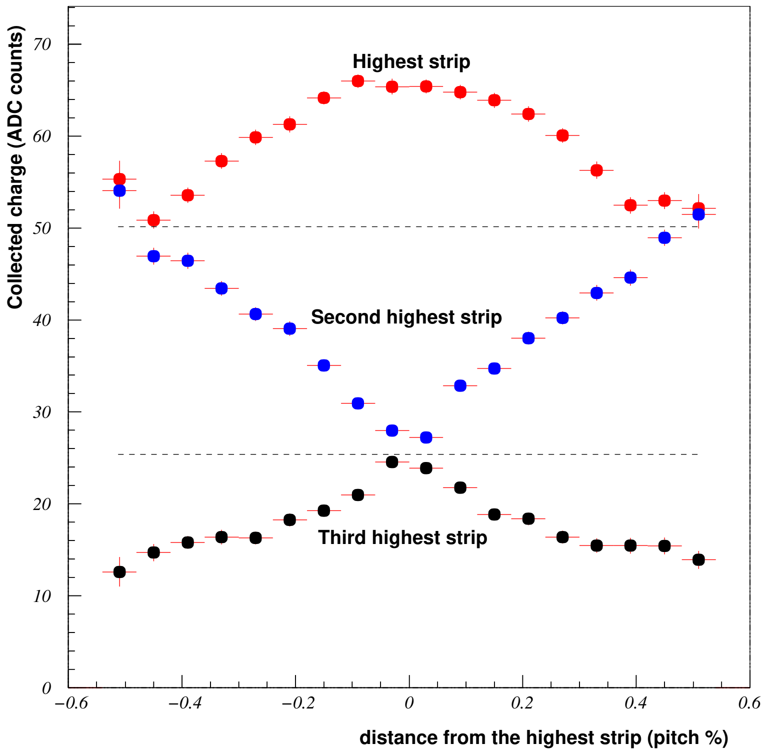

Figure 7.28 Signal collected on the first, second and third highest strip as a function of the distance from the first highest strip (in percentage of the read-out pitch).

In this section we will briefly discuss some features of the Gaussian and DC algorithms and we will review the performance of the reconstruction algorithms as a function of the impact angle θ.

Gaussian vs Double Centroid

Figure 7.29-a shows the average value of the residual, $\delta = x_{pred} - x_{meas}$, as a function of the predicted inter-strip position, x_{pred}, measured in two different regions of the sensor using the Gaussian algorithm and perpendicular tracks. The vertical error bars in the plot represent the spread of the residual distribution in each bin. In both regions the distribution of the residuals is not centered at zero but shows systematic deviations towards positive or negative values. The maximum observed deviation is of the order of ~ 10 μm and ~ 14 μm respectively in the

regions uncovered and covered with bond pads. Thus, the Gaussian algorithm systematically pulls the reconstructed position towards one of the read-out strips. The distribution of events reconstructed with the Gaussian algorithm as a function of the distance of the hit from the highest strip is shown in Figure 7.29-b. The distribution has a maximum in correspondence of the read-out strip and decreases towards the middle position in-between two read-out strips.

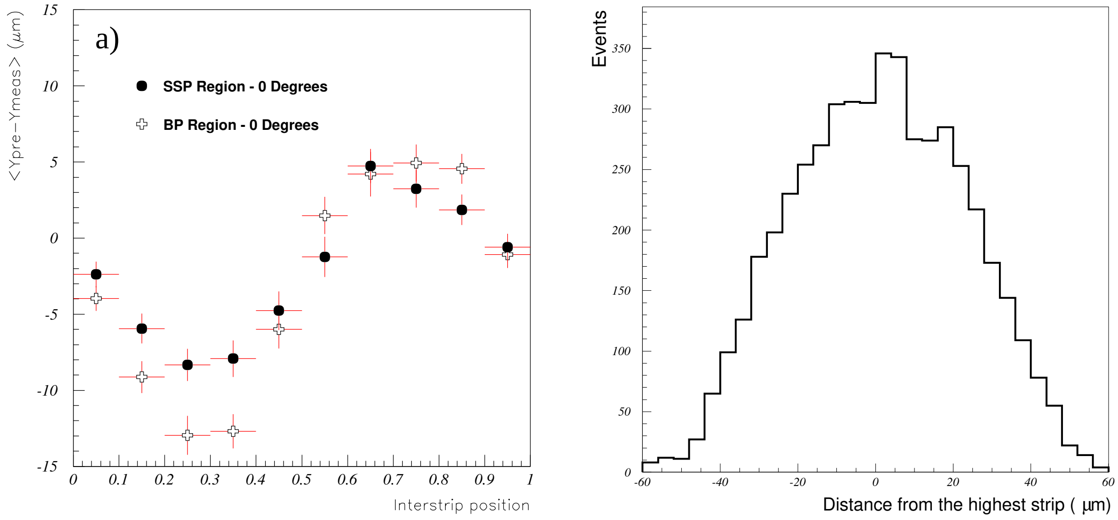

Figure 7.29 *On the left:* Average value of the residual distribution δ as a function of the inter-strip position for hits reconstructed with the Gaussian algorithm. The distribution measured in the bond pad region (*open crosses*) and in a region not covered by pads (*full circles*) are shown separately. The vertical error bars correspond to the width of the residual distribution in each bin. The inter-strip position is presented as fraction of the read-out pitch. *On the right:* Distribution of events as a function of the distance from the highest strip reconstructed with the Gaussian algorithm.

Figure 7.30 shows the distribution of events reconstructed using the Double Centroid algorithm as a function of the reconstructed position. The distribution is shown for three different values of the parameter α, where

$$x_{DC} = \frac{C_{left}/dr^{\alpha} + C_{right}/dl^{\alpha}}{dr^{\alpha} + dl^{\alpha}}$$

and the weights dl and dr are defined in Eq. 7.6. The parameter α can be tuned in order to obtain an unbiased distribution of events as a function of the inter-strip position. Since the beam produces a uniform irradiation on the sensor the optimal event distribution should be flat. From Figure 7.30 we notice that with $\alpha = 1$ a sensible reconstruction is achieved. The average value of the residual distribution as a function of the predicted inter-strip position is shown in Figure 7.31 for hits reconstructed with the Double Centroid algorithm and $\alpha = 1$. In this case the distribution is almost constant for both regions of the detector.

In conclusion, the Gaussian algorithm introduces sizeable systematic shifts in the reconstructed position, depending on the inter-strip position. The effect is also observed when using the Double Centroid algorithm with $\alpha \neq 1$. Nevertheless, in the latter case an unbiased reconstruction can be achieved with a proper choice of the parameter α.

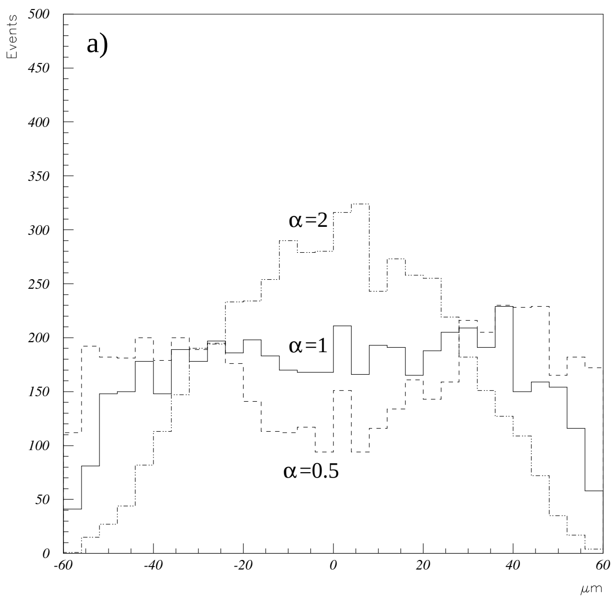

Figure 7.30 Distribution of events as a function of the distance from the highest strip reconstructed with the Double Centroid. The distribution is shown for different values of the parameter α (*see text*).

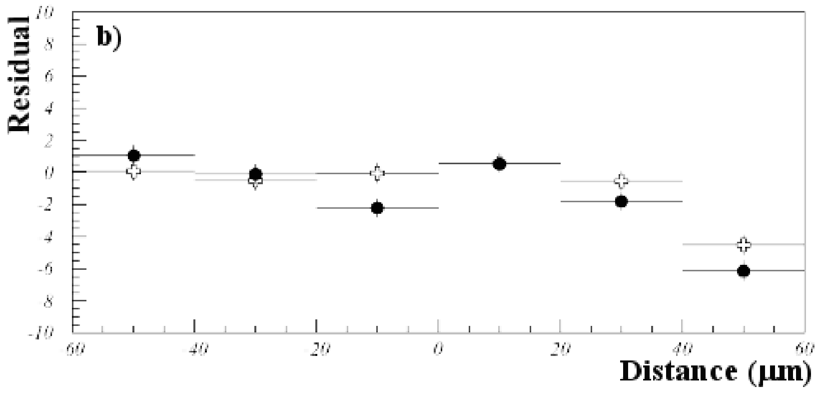

Figure 7.31 Average value of the residual distribution δ as a function of the inter-strip position for hits reconstructed with the Double Centroid algorithm. The distribution measured in the bond pad region (*open crosses*) and in a region not covered by pads (*full circles*) are shown separately.

Position resolution as a function of the impact angle

The intrinsic hit resolution of the BMVD half module was measured as a function of the impact angle θ and using the reconstruction algorithms discussed above.

Figure 7.32 shows the intrinsic resolution as a function of θ measured in the region of the sensor not covered with bond or probe pads. The best achieved intrinsic resolution for perpendicular tracks is 13 μm with the Eta algorithm as compared with 9 μm obtained with the single sensor [123]. The deterioration is expected as a result of the additional coupling capacitance. The DC algorithm has a slightly worse resolution (17 μm). For incident angles greater than 40 degrees the HT algorithm gives the best performance. The DC algorithm with four strips gives a resolution similar to the DC algorithm with three strips at small angles and to the HT algorithm at larger angles. The intrinsic resolution as a function of θ measured in the bond pad region is shown in Figure 7.33. For perpendicular tracks we observe a deterioration of the resolution of about $2 \div 3$ μm for almost all the reconstruction algorithms. Also in this case, the Eta algorithm gives the best performance at small impact angles and the HT at large angles.

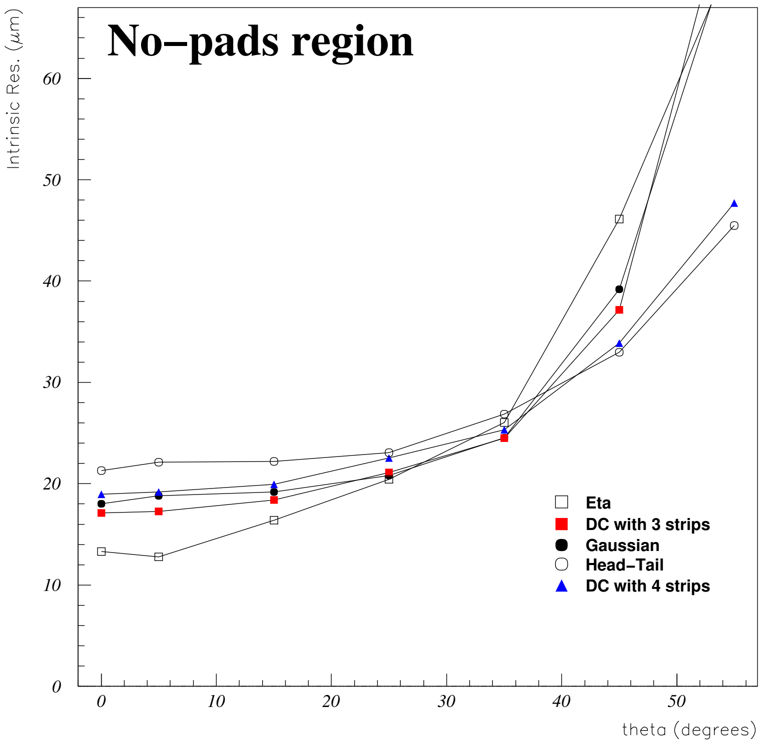

Figure 7.32 Intrinsic resolution of a BMVD half module as a function of the impact angle θ

A direct comparison between the resolution measured in the bond pad region and the standard region is shown in Figure 7.34 for the Eta and DC algorithms. For impact angles below 5° the performance of the Eta algorithm are independent on the detector region. This algorithm takes into account the difference in terms of charge division through the probability function $f(\eta)$. For the DC algorithm the differences are contained within ~ 3 μm. Furthermore, this algorithm does not require the parameterization of the probability function and it gives an un-

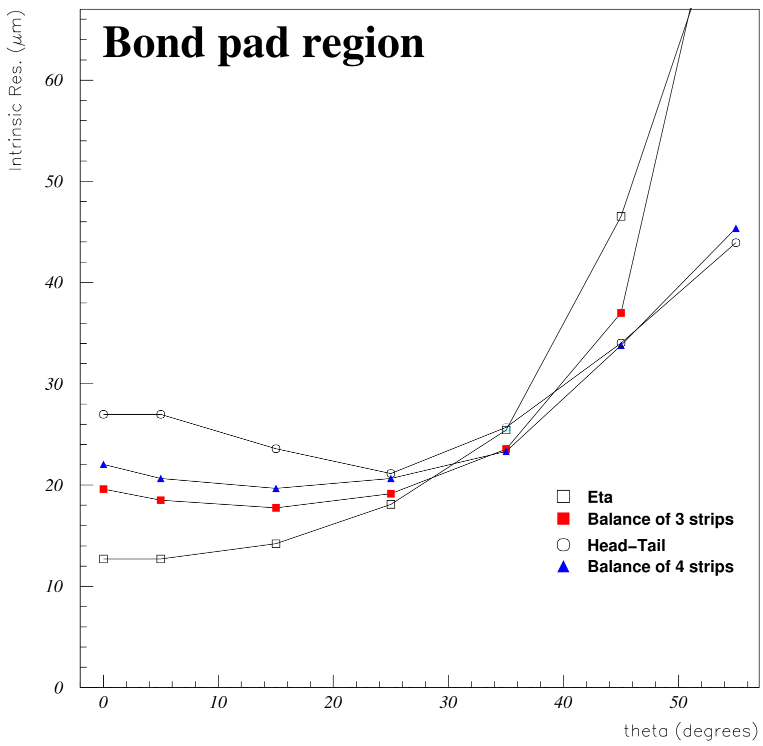

Figure 7.33 Intrinsic resolution of a BMVD half module as a function of the impact angle θ, measured in the bond pad region.

biased reconstruction, as discussed in the last Section. For angles larger than 30° no difference is observed between the two regions for both algorithms.

7.4 Test of a full module

As discussed in Section 7.2.3 a full module is obtained by placing two half modules on top of each other. For each half module a hybrid circuit containing four front-end chips is connected to the sensors. The two hybrid circuits are connected through a multi-layer PCB with a flexible layer. The full module provides two coordinates for a track traversing it and has a total of 1024 channels.

The purpose of the tests described in the following sections is to test the detector read-out with an electronic chain similar to the final one. In what follows we describe the experimental setup and the software used to reconstruct tracks traversing the full module. In Section 7.4.1 we present a study of the MVD ADC clustering thresholds performed using the data collected with the full module.

A full module was connected to a read-out chain as sketched in Figure 7.35. A Ru^{106} source was placed over the sensor with a dedicated support. The trigger signal was generated with a scintillator detector placed under the sensor. Prototype units of the Clock and Control, of the Helix driver and of the Analog link were used. The final ADC module was connected to a PC for the data acquisition. Prototype units of the low voltage supplies were used to power

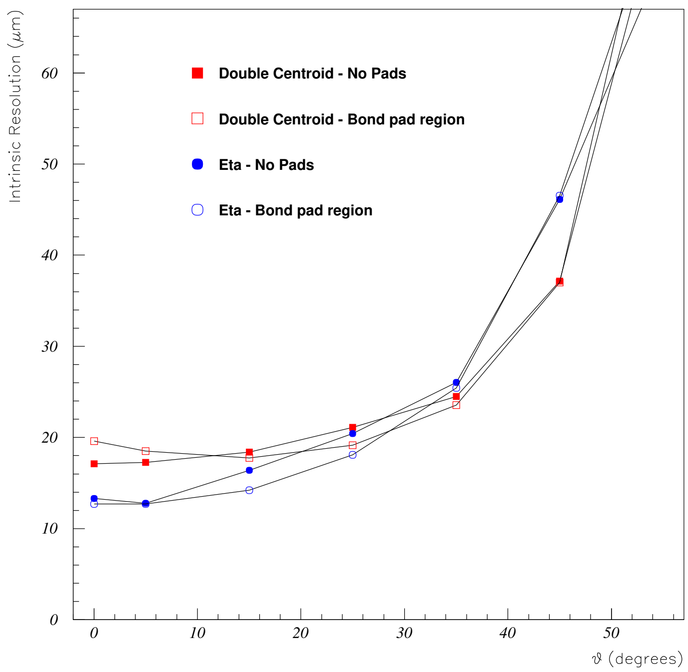

Figure 7.34 Intrinsic resolution of a BMVD half module as a function of the impact angle θ, measured in two regions of the sensor.

the front-end electronics whereas a commercial high voltage power supply was used to bias the sensors.

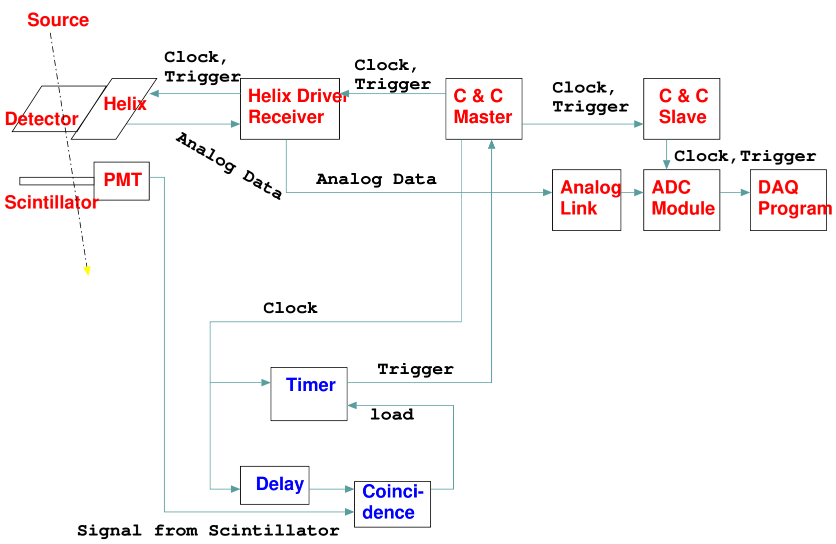

Figure 7.35 Setup for the test of a full MVD module with a radioactive source.

The ADC read-out mode was set to *raw data*, such that the signal from all strips was recorded on disk. The common mode and pedestal subtraction was performed offline. The hit position was reconstructed with a center of gravity of the two highest strips in each half module. Figure 7.36-a,b show the hit distribution reconstructed respectively in the upper and lower half module. The corresponding hit pattern in two dimensions is shown in Figure 7.36-c. The two spatial measurements give the two dimensional position of the traversing particle. The distribution of the signal collected in the two half modules is shown in Figure 7.37. The threshold was defined as 5σ, were σ is the r.m.s. noise of each strip. For each hit, the signal was defined as the sum of the two highest strips.

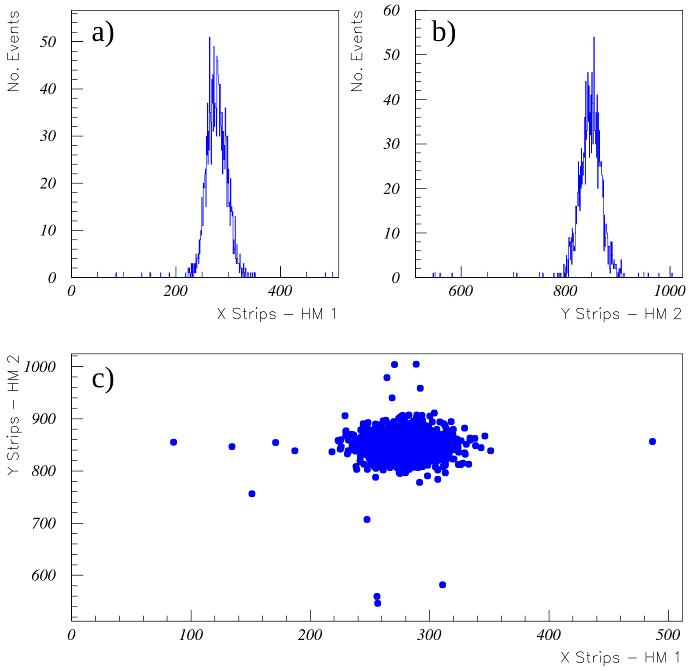

Figure 7.36 Hit reconstruction with a BMVD full module. The distribution of the hits reconstructed in the upper (a) and lower (b) half module are shown separately. The corresponding two dimensional hit pattern is shown in c).

7.4.1 Optimization of the ADC clustering thresholds

The data collected with the full module setup were used to study the optimization of the ADC thresholds for strip clustering. In the MVD ADC modules two kinds of thresholds can be set: A *strip threshold*, variable strip by strip and representing the minimum charge above which the strip is included in the clustering procedure, and a *cluster threshold* that is the minimum total charge in a cluster. The selected strips are processed by the clustering algorithm, described in Ref. [131].

The ADC algorithm was implemented in a `FORTRAN` code and the clustering was performed offline on the common mode and pedestal subtracted data. Figure 7.38 shows the distribution of the cluster charge for different values of the strip threshold and no cluster threshold. For each strip, the threshold was set to multiples of the r.m.s. of the pedestal. Each distribution shows a peak at low values produced by noise clusters and a tail corresponding to the real hits. As the value of the strip threshold increases the number of noise clusters, represented by the

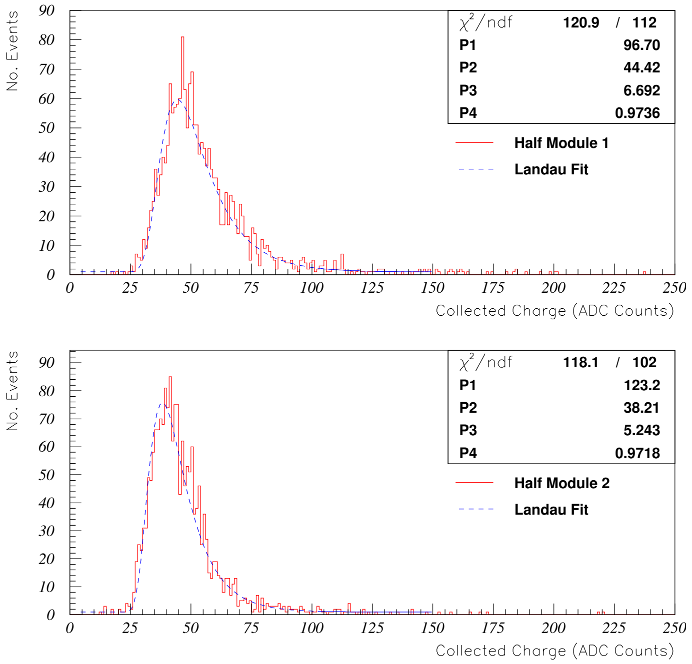

Figure 7.37 Distribution of the signal collected in the two half modules of a full module (*solid line*). A fit with a Landau function is shown (*dashed line).*

area under the first peak, decreases.

The clustering thresholds were studied in terms of two estimators, the *purity* and the *efficiency*. For each cluster charge distribution a Landau fit was performed in the region corresponding to the signal. The purity, $p(s,c)$, for a certain value of the strip threshold s and of the cluster threshold c was defined as

$$p(s,c) = \frac{\text{Number of signal clusters}}{\text{Total number of clusters}}, \tag{7.13}$$

where the numerator corresponds to the integral of the function extracted with a Landau fit. The purity is the fraction of events coming from real hits in the module over the the total number of reconstructed hits. On the other hand, the inefficiency is the fraction of signal events lost due to the threshold cuts. The efficiency, $\epsilon(s,c)$ was defined as

$$\epsilon(s,c) = \frac{\text{Number of signal clusters above } c}{\text{Total number of signal clusters}}, \tag{7.14}$$

where the numerator is the integral of the Landau function for signal above the threshold c.

The purity is shown in Figure 7.39 as a function of the cluster threshold. For a fixed value of the strip threshold the purity increases for increasing values of the cluster threshold. Therefore, for increasing values of both thresholds more and more clusters are real hits in the detector. On the other hand, if we increase the thresholds to a too high value we also lose hits. This is shown in Figure 7.40, in which the efficiency is plotted as a function of the cluster thresholds for different values of the strip threshold. We observe that, for cluster thresholds smaller than ~ 35 ADC counts the efficiency is greater than 99%. For larger values of the cluster threshold the efficiency decreases.

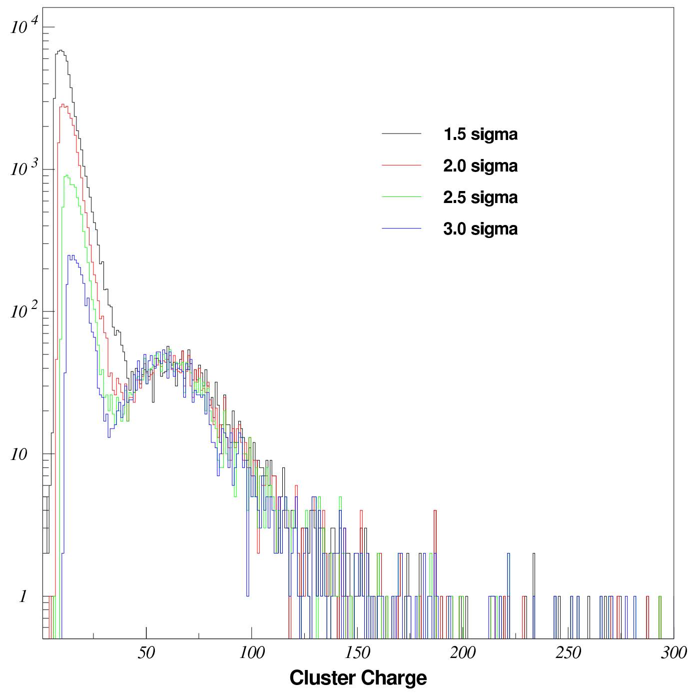

Figure 7.38 Distribution of the cluster charge measured on a MVD full module. The distribution corresponding to different values of the strip threshold are shown.

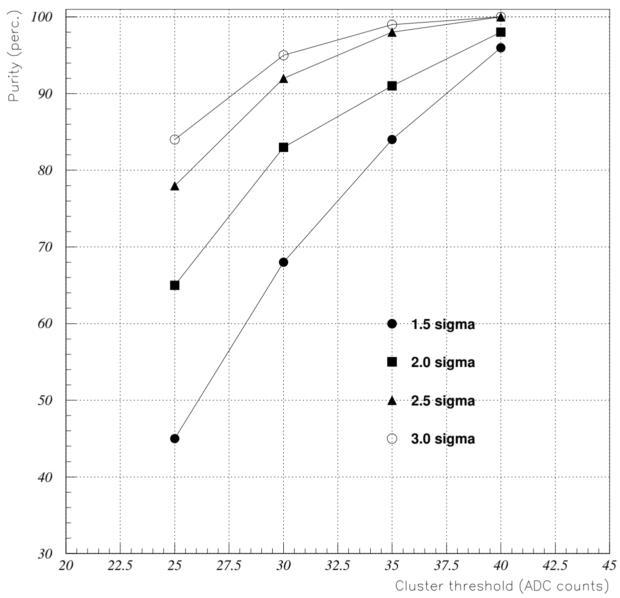

Figure 7.39 Purity of the ADC clustering as a function of the cluster threshold measured for different values of the strip threshold.

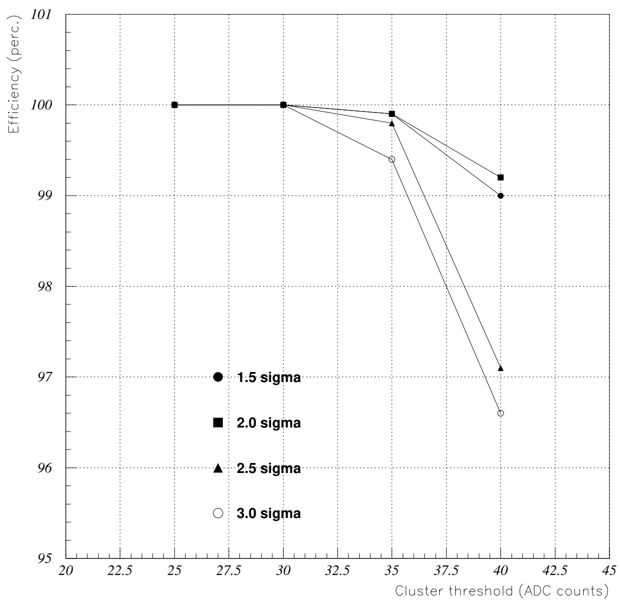

Figure 7.40 Efficiency of the ADC clustering as a function of the cluster threshold measured for different values of the strip threshold.

We conclude that sensible values for the ADC clustering are a cluster threshold of about 35 ADC count and a strip threshold between 2.5 and 3 times the strip noise. With these values we observe an efficiency $\epsilon > 99\%$ and a purity $p > 98\%$.

7.5 Conclusions

During the HERA luminosity upgrade the ZEUS detector was upgraded with a silicon Micro Vertex Detector. In this Chapter we have presented the results of the beam test measurements performed on the MVD half modules. The results can be summarized as follows:

1. The signal to noise ratio at a bias voltage $V_{bias} = 100$ V is around 15. The difference in the collected charge between the two single sensors in a half module is of the order of $\sim 5\%$;

2. The effect related to the presence of probe and bond pads were investigated using the η-algorithm for the hit reconstruction. Sizable dependence of the charge sharing between read-out strips on the distance from the pads was observed. The effect is reflected in the η-distribution used to reconstruct the impact position. From theses measurements we conclude that a maximum systematic error of the order of 10 μm can be made if the hit position is reconstructed with the η-algorithm and the wrong probability function is used. The measurement are confirmed by a numerical simulation of the BMVD sensors performed with the `PSPICE` software package.

3. Existing and newly developed hit reconstruction algorithms were studied. In particular, due to the observation of point 2) the performance of alternative algorithms were investigated. We summarize the observations as follows:

    a) For *small incidence angles* ($\theta < 30°$) the best intrinsic resolution is achieved with the η-algorithm ($\sigma_{intr} \sim 13$ μm). Nevertheless, the algorithm requires the parameterization of the probability function $f(\eta)$ in different regions of the sensor to avoid systematic errors. The DC algorithm is a good alternative to the η-algorithm. The intrinsic resolution of the DC algorithm for tracks perpendicular to the sensor is $\sigma_{intr} \sim 17$ μm. Furthermore, the algorithm can be tuned to obtain an unbiased reconstruction in all regions of the sensor. In both cases a very precise reconstruction can be achieved despite the read-out pitch of 120 μm, fulfilling the detector specification of $\sigma_{intr} < 20$ μm for perpendicular tracks.

    b) For *large incidence angles* ($\theta > 30°$) the most precise reconstruction is achieved using the HT algorithm. On the other hand, it was shown that using the DC algorithm of 4 strips acceptable precision of the spatial measurement can be achieved for all values of the incidence angle.

In addition to the beam tests performed on half modules, full modules were connected to prototype units of the final read-out chain. The functionality of the MVD readout electronics was successfully tested. The data collected with a radioactive source were used to establish the optimal values of the cluster thresholds of the MVD ADC units. It was shown that through a sensible choice of the ADC thresholds excellent values of the clustering purity and efficiency can be achieved.

Summary and outlook

In this dissertation we have presented the measurements of total and single differential cross sections of b-quark production in neutral current deep inelastic e^+p scattering. The measurements were performed with the ZEUS detector at HERA using the data sample collected in the years 1999-2000 and corresponding to an integrated luminosity of 60 pb^{-1}. The single differential cross sections represent the first measurement of this kind at HERA.

Events were selected containing the scattered positron, a hard jet reconstructed in the Breit frame and a semi-leptonic muon traversing the barrel and rear muon chambers. The fraction of events from b-decays was extracted on a statistical basis using the distribution of the muon transverse momentum relative to the closest jet. The phase space region was carefully selected in order to minimize the experimental and theoretical uncertainties as well as the model dependency of the result.

QCD predictions to the NLO accuracy were obtained using the calculation by Harris and Smith implemented in the `HVQDIS` program. A detailed analysis of the NLO QCD uncertainties was carried out and these predictions were compared to Monte Carlo programs based on LO matrix elements and implementing different QCD evolution equations.

Both the total and single differential cross sections are almost a factor 1.4 above the NLO QCD predictions but compatible within the experimental and theoretical uncertainties. The result is also compatible with the prediction of the `CASCADE` Monte Carlo program, based on k_T-factorization and CCFM QCD evolution equation. Other Monte Carlo predictions lie below the data.

In the total cross section obtained by the H1 Collaboration an excess of a factor three is observed between the data and the NLO QCD predictions. However, this result is defined in a different kinematic region and relies on additional model assumptions and it is, therefore, not directly comparable with this result.

A more precise measurement will be possible after the upgrade of the HERA collider. The new layout of the interaction region includes low-β quadrupoles close to the interaction point by which an instantaneous luminosity of $7 \cdot 10^{-31}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ can be expected. This value corresponds to an improvement of a factor five with respect to the HERA design specifications and up to three compared to the 1997 running conditions.

To fully benefit from the new experimental conditions, the ZEUS detector was upgraded by an improved tracking system consisting of a Straw-Tube-Tracker and a Micro Vertex Detector. The latter will allow the measurement of the impact parameter of a track produced in a heavy meson decay with a precision better than 100 μm for tracks perpendicular to the beam line.

In this dissertation we have presented the results of beam test measurements performed on the MVD half modules. The charge collection, the signal/noise ratio and the hit resolution have been measured for different incidence angles using existing and newly developed hit reconstruction algorithms. Detailed studies of the charge sharing mechanism have been performed on the regions of the MVD sensors corresponding to the bonding pads and on the overlap between single sensors.

The data analysis has highlighted new unexpected phenomena, i.e. the sizable dependence of the charge sharing on the presence of the aluminum contacts on the silicon surface. These results have been confirmed by numerical simulations.

In addition, MVD full modules were connected to prototype units of the final read-out chain. Measurements of signal/noise ratio were performed using radioactive sources. The data were also used to establish the optimal values of the cluster thresholds of the MVD ADC units.

The MVD detector was installed in the ZEUS experiment during April 2001. After the completion of the HERA upgrade, it was realized in the commissioning stage of the machine that the beam-related background was unacceptable from the point of view of operation of the wire chambers of ZEUS and H1. Therefore, the latest period of the HERA post-upgrade run was mainly devoted to the understanding of the backgrounds and to the comparison of the observed rates to the simulation. The solution proposed by the ZEUS Collaboration for the reduction of the background rates to acceptable conditions are reported in Ref. [132]. According to the guidelines suggested by these reports, modifications to the elements in the interaction region will be applied during a shutdown period of 19 weeks starting in March 2003. The HERA physics program will then continue until year 2006.

Bibliography

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

[2] S. J. Brodsky, P. Hoyer, A. H. Mueller, and W.-K. Tang, Nucl. Phys. **B369**, 519 (1992).

[3] R. Vogt and S. J. Brodsky, Nucl. Phys. **B438**, 261 (1995), hep-ph/9405236.

[4] S. Schilling, *Implementation of BGF-processes in Monte Carlo generators for electron proton scattering*, Master's thesis, University of Heidelberg, 2000, DESY-THESIS-2000-040.

[5] R. Field, *Applications of Perturbative QCD* (Addison-Wesley, 1995).

[6] M. Glueck, E. Reya, and A. Vogt, Z. Phys. **C67**, 433 (1995).

[7] W. L. van Neerven, hep-ph/0107193, Proceedings of the 23rd Annual MRST Conference on High Energy Physics (MRST 2001), London, Ontario (Canada), 2001.

[8] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, 1995).

[9] J. C. Collins, D. E. Soper, and G. Sterman, Nucl. Phys. **B263**, 37 (1986).

[10] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, Nucl. Phys. **B392**, 229 (1993).

[11] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, Nucl. Phys. **B392**, 162 (1993).

[12] B. W. Harris and J. Smith, Nucl. Phys. **B452**, 109 (1995), hep-ph/9503484.

[13] Z. Kunszt and D. E. Soper, Phys. Rev. **D46**, 192 (1992).

[14] V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 1218 (1972).

[15] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[16] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[17] A. Martin, Lectures given at the XXI Int. Meeting on Fundamental Physics, Miraflores de la Sierra, Madrid (Spain), 1993.

[18] G. P. Salam, Acta Phys. Polon. **B30**, 3679 (1999), hep-ph/9910492.

[19] R. Ellis *et al.*, *QCD and Collider Physics* (Cambridge Univ. Press, 1996).

[20] Small-x Collab., B. Andersson *et al.*, Eur. Phys. J. **C25**, 77 (2002), hep-ph/0204115.

[21] T. Carli, Int. J. Mod. Phys. **A17**, 3185 (2002), hep-ph/0111232.

[22] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **44**, 443 (1976).

[23] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977).

[24] I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[25] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B366**, 135 (1991).

[26] M. Ciafaloni, Nucl. Phys. **B296**, 49 (1988).

[27] S. Catani, F. Fiorani, and G. Marchesini, Phys. Lett. **B234**, 339 (1990).

[28] S. Catani, F. Fiorani, and G. Marchesini, Nucl. Phys. **B336**, 18 (1990).

[29] G. Marchesini, Nucl. Phys. **B445**, 49 (1995), hep-ph/9412327.

[30] M. Bengtsson and T. Sjostrand, Z. Phys. **C37**, 465 (1988).

[31] H. Jung and G. P. Salam, Eur. Phys. J. **C19**, 351 (2001), hep-ph/0012143.

[32] G. Marchesini and B. R. Webber, Nucl. Phys. **B349**, 617 (1991).

[33] G. Marchesini and B. R. Webber, Nucl. Phys. **B386**, 215 (1992).

[34] H1 Collab., S. Aid *et al.*, Nucl. Phys. **B470**, 3 (1996), hep-ex/9603004.

[35] B. R. Webber, J. Phys. **G19**, 1567 (1993).

[36] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, Phys. Rev. **D27**, 105 (1983).

[37] R. L. Jaffe and L. Randall, Nucl. Phys. **B412**, 79 (1994), hep-ph/9306201.

[38] V. G. Kartvelishvili, A. K. Likhoded, and V. A. Petrov, Phys. Lett. **B78**, 615 (1978).

[39] M. G. Bowler, Zeit. Phys. **C11**, 169 (1981).

[40] B. Andersson, G. Gustafson, and B. Soderberg, Z. Phys. **C20**, 317 (1983).

[41] P. D. B. Collins and T. P. Spiller, J. Phys. **G11**, 1289 (1985).

[42] P. Nason and B. R. Webber, Phys. Lett. **B395**, 355 (1997), hep-ph/9612353.

[43] M. Cacciari and P. Nason, Phys. Rev. Lett. **89**, 122003 (2002), hep-ph/0204025.

[44] ALEPH Collab., A. Heister *et al.*, Phys. Lett. **B512**, 30 (2001), hep-ex/0106051.

[45] B. Mele and P. Nason, Nucl. Phys. **B361**, 626 (1991).

[46] M. Cacciari and S. Catani, Nucl. Phys. **B617**, 253 (2001), hep-ph/0107138.

[47] H1 Collab., C. Adloff *et al.*, Phys. Lett. **B467**, 156 (1999), hep-ex/9909029.

[48] H1 Collab., C. Adloff *et al.*, Contributed paper no. 979/982 to ICHEP 2000, Osaka.

[49] ZEUS Collab., J. Breitweg *et al.*, Eur. Phys. J. **C18**, 625 (2001), hep-ex/0011081.

[50] S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, Phys. Lett. **B348**, 633 (1995), hep-ph/9412348.

[51] ZEUS Collab., S. Chekanov *et al.*, Contributed paper no. 785 to ICHEP 2002, Amsterdam.

[52] T. Sjostrand, Comput. Phys. Commun. **82**, 74 (1994).

[53] E. Norrbin and T. Sjostrand, Eur. Phys. J. **C17**, 137 (2000), hep-ph/0005110.

[54] H1 Collab., C. Adloff *et al.*, Contributed paper no. 1013 to ICHEP 2002, Amsterdam.

[55] ZEUS Collab., S. Chekanov *et al.*, Contributed paper no. 783 to ICHEP 2002, Amsterdam.

[56] H1 Collab., C. Adlof *et al.*, Contributed paper no. 1016 to ICHEP 2002, Amsterdam.

[57] ZEUS Collab., S. Chekanov *et al.*, Contributed paper no. 784 to ICHEP 2002, Amsterdam.

[58] ZEUS Collab., The ZEUS detector: Status report 1993, ZEUS-STATUS-REPT-1993.

[59] ZEUS Collab., The ZEUS Detector: Status report 1989, ZEUS-1989.

[60] DESY HERA Report No. 81-10, 1981 (unpublished).

[61] N. Harnew *et al.*, Nucl. Instrum. Meth. **A279**, 290 (1989).

[62] ZEUS Collab., B. Foster *et al.*, Nucl. Instrum. Meth. **A338**, 254 (1994).

[63] M. Derrick *et al.*, Nucl. Instrum. Meth. **A309**, 77 (1991).

[64] ZEUS Calorimeter Group, A. Andresen *et al.*, Nucl. Instrum. Meth. **A309**, 101 (1991).

[65] A. Caldwell *et al.*, Nucl. Instrum. Meth. **A321**, 356 (1992).

[66] ZEUS Barrel Calorimeter Group, A. Bernstein *et al.*, Nucl. Instrum. Meth. **A336**, 23 (1993).

[67] G. Battistoni *et al.*, Nucl. Instr. Meth. **217**, 429 (1983).

[68] G. Abbiendi *et al.*, Nucl. Instrum. Meth. **A333**, 342 (1993).

[69] M. De Giorgi *et al.*, Nucl. Instrum. Meth. **A378**, 472 (1996).

[70] ZEUS Collab., A. Bamberger *et al.*, Nucl. Instrum. Meth. **A401**, 63 (1997).

[71] W. Smith *et al.*, The ZEUS trigger system, ZEUS-Note 89-084.

[72] M. Klein, in *Physics at HERA*, edited by W. Buchmüller and G. Ingelman Vol. 1, p. 71, Hamburg, 1991, DESY.

[73] A. Blondel and F. Jacquet, in *Study for an ep facility for Europe*, p. 391, 1979.

[74] S. Bentvelsen *et al.*, Reconstruction of (x, Q^2) and extraction of structure functions in neutral current scattering at HERA, in *Physics at HERA*, edited by W. Buchmüller and G. Ingelman Vol. 1, p. 23, Hamburg, 1991, DESY.

[75] U. Bassler and G. Bernardi, Nucl. Instrum. Meth. **A426**, 583 (1999), hep-ex/9801017.

[76] U. Bassler and G. Bernardi, Nucl. Instrum. Meth. **A361**, 197 (1995), hep-ex/9412004.

[77] H. Abramowicz, A. Caldwell, and R. Sinkus, Nucl. Instrum. Meth. **A365**, 508 (1995), hep-ex/9505004.

[78] R. Sinkus and T. Voss, Nucl. Instrum. Meth. **A391**, 360 (1997).

[79] A. López-Durán Viani and S. Schlensted, Electron finder efficiencies and impurities. a comparison between `SINISTRA95`, `EM` and `EMNET`, ZEUS-Note 99-077.

[80] T. Doeker, A. Frey, and M. Nakao, Electron position reconstruction. Update of the `ELECPO` routines, ZEUS-Note 94-123.

[81] J. Ng and W. Verkerke, An overview of SRTD analysis, ZEUS-Note 95-037.

[82] N. Tuning, ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction, ZEUS-Note 99-021.

[83] G. M. Briskin, *Diffractive dissociation in ep Deep Inelastic Scattering*, PhD thesis, Tel Aviv University, School of Physics and Astronomy, 1998, DESY-THESIS 1998-036.

[84] G. Hartner *et al.*, VCTRAK (3.06/20): Offline output information, ZEUS-Note 96-013.

[85] A. Longhin, *Misura delle sezioni d'urto di fotoproduzione di Open Beauty ad HERA*. (In Italian), Master's thesis, University of Padua, 2000.

[86] ZEUS Collab., S. Chekanov *et al.*, Phys. Lett. **B531**, 9 (2002), hep-ex/0112030.

[87] ZEUS Collab., S. Chekanov *et al.*, Eur. Phys. J. **C23**, 615 (2002), hep-ex/0112029.

[88] M. Wing, hep-ex/0206036, Proceedings from the 10th International Conference on Calorimetry in High Energy Physics, CALTECH, Pasadena, 25-29 March 2002.

[89] G. Abbiendi, Global tracking of muons in the barrel and rear region, ZEUS-Note 99-063.

[90] V. Innocente, M. Maire, and E. Nagy, In *Amsterdam 1991, Proceedings, MC91: Detector and event simulation in high energy physics* 58-78. (see HIGH ENERGY PHYSICS INDEX 30 (1992) No. 3237).

[91] S. Limentani, M. Posocco, and L. Stanco, Muon barrel reconstruction, ZEUS-Note 92-027.

[92] M. Arneodo, L. Lamberti, and M. Ryskin, Comput. Phys. Commun. **100**, 195 (1997), hep-ph/9610286.

[93] S. P. Baranov, O. Duenger, H. Shooshtari, and J. A. M. Vermaseren, In *Hamburg 1991, Proceedings, Physics at HERA, vol. 3* 1478-1482. (see HIGH ENERGY PHYSICS INDEX 30 (1992) No. 12988).

[94] M. Wodarczyk, *Measurement of the F_2 structure function of the proton at HERA from 1996 and 1997 ZEUS data*, PhD thesis, University of Wisconsin, 1999.

[95] L. Adamczyk *et al.*, Luminosity measurement in the ZEUS experiment, ZEUS-NOTE 01-004.

[96] SLD Collab., K. Abe *et al.*, Contributed to International Europhysics Conference on High- Energy Physics (HEP 2001), Budapest, Hungary, 12-18 Jul 2001 and Contributed to 20th International Symposium on Lepton and Photon Interactions at High Energies (LP 01), Rome, Italy, 23-28 Jul 2001.

[97] A. Longhin, private communication.
Available on: `http://www-zeus.desy.de/~longhin/mudstar/mudstar.html`.

[98] T. Carli, H. Kuester, and M. Lindstroem, Weighted Monte Carlo Fact ot Fiction ?, H1 Software Note 58-5/98.

[99] CTEQ Collab., H. L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (2000), hep-ph/9903282.

[100] SLD Collab., K. Abe *et al.*, Phys. Rev. **D65**, 092006 (2002), hep-ex/0202031.

[101] OPAL Collab., G. Abbiendi *et al.*, (2002), hep-ex/0210031.

[102] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219 (1993).

[103] F. Goebel, *Measurement of the diffractive contribution to the DIS cross section using the ZEUS Forward Plug Calorimeter*, PhD thesis, Hamburg University, 2001.

[104] Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. **D66**, 010001 (2002).

[105] T. Carli, V. Chiochia, and K. Klimek, (2003), hep-ph/0305103, DESY-03-063, to be published on JHEP.

[106] G. Ingelman, A. De Roeck, and R. Klanner, DESY-96-235.

[107] M. Aivazis, J. C. Collins, F. I. Olness, and W. Tung, Phys. Rev. **D50**, 3102 (1994), hep-ph/9312319.

[108] A. Chuvakin, J. Smith, and B. W. Harris, Eur. Phys. J. **C18**, 547 (2001), hep-ph/0010350.

[109] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, Phys. Lett. **B411**, 211 (1997), hep-ph/9707263.

[110] A. Chuvakin, J. Smith, and W. L. van Neerven, Phys. Rev. **D62**, 036004 (2000), hep-ph/0002011.

[111] H1 Collab., C. Adloff *et al.*, Nucl. Phys. **B545**, 21 (1999), hep-ex/9812023.

[112] K. Daum, S. Riemersma, B. W. Harris, E. Laenen, and J. Smith, (1996), hep-ph/9609478, Proceedings of Workshop on Future Physics at HERA, Hamburg, Germany, 25-26 Sep 1995.

[113] U. Schneekloth, DESY-HERA-98-05.

[114] ZEUS, A Microvertex Detector for ZEUS, ZEUS-Note 97-006.

[115] A. Garfagnini, Nucl. Instrum. Meth. **A435**, 34 (1999).

[116] A. Garfagnini and U. Kotz, Nucl. Instrum. Meth. **A461**, 158 (2001).

[117] E. Koffeman and H. Tiecke, Mechanical design of the zeus micro vertex detector, ZEUS-Note 00-028.

[118] D. Dannheim *et al.*, Nucl. Instrum. Meth. **A505**, 663 (2003), hep-ex/0212026.

[119] W. Fallot-Burghardt *et al.*, *Helix128-2.x User Manual*, ASIC Labor Heidelberg, HD-ASIC-33-0697. The modifications for ZEUS are described in `http://wwwasic.unihep.uni-heidelberg.de/~feuersta/Helix/index.html.`

[120] J. Velthuis, Irradiation test of the ZEUS vertex detector frontend chips, the HELIX 128-3.0, 2001, proceedings of the 7th International Conference on Advanced Technology and Particle Physics, Villa Olmo, Como.

[121] T. Fusayasu and K. Tokushuku, Nucl. Instrum. Meth. **A436**, 281 (1999).

[122] A. Polini, The ZEUS Micro Vertex Detector, 2001, proceedings of the 7th International Conference on Advanced Technology and Particle Physics, Villa Olmo, Como.

[123] L. A. T. Bauerdick *et al.*, Nucl. Instrum. Meth. **A501**, 340 (2003), hep-ex/0212037.

[124] M.Milite, PhD thesis, University of Hamburg, 2001, DESY-THESIS-2001-050. Available at `http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis-01-050`.

[125] V. Chiochia, Nucl. Instrum. Meth. **A501**, 60 (2003), hep-ex/0111061.

[126] E. Belau *et al.*, Nucl. Instr. Meth. **214**, 253 (1983).

[127] U. Koetz *et al.*, Nucl. Instr. Meth. **235**, 481 (1985).

[128] R. Turchetta, Nucl. Instrum. Meth. **A335**, 44 (1993).

[129] M.Moritz, *Measurement of the high Q^2 neutral current DIS cross section at HERA*, PhD thesis, University of Hamburg, 2002, DESY-THESIS-2002-009.

[130] J. Martens, *Simulations and quality control of the silicon strip detectors of the ZEUS Microvertex Detector.* (In German), Master's thesis, University of Hamburg, 1999, DESY-THESIS-1999-044.

[131] T. Fusayasu, MVD ADC module cluster algorithm
, Available at `http://www-zeus.desy.de/~toku/mvd/adcm-argo.htm`.

[132] ZEUS Background working group, A. Antonov *et al.*, Study of beam-induced backgrounds in the ZEUS detector from 2002 HERA running, ZEUS Note 02-18. Addenda in ZEUS Notes 02-020 and 02-027.

List of Figures

List of Tables

Acknowledgments

This thesis owes a large intellectual debt to the colleagues I have had the privilege of working with. I am thankful to Prof. Robert Klanner who gave me the opportunity to work on my Ph.D. within the ZEUS Collaboration at DESY. I am also gratefully indebted to Tancredi Carli for his unstinting commitment to helping see this thesis through its final completion, and his equally generous and wise guidance during its development. A special thank goes also to Prof. Roberto Carlin for his encouragement and guidance during my studies. To him and to Prof. Peter Schleper goes all my gratitude for correcting the drafts of this thesis.

The measurement of b-quark production in deep inelastic scattering was possible thanks to the help and expertise of many people. In particular, I would like to thank Massimo Corradi, Andrea Longhin, Monica Turcato and Nicola Coppola for developing parts of analysis code used in this thesis. Their contribution was also essential to achieve a satisfactory understanding of the muon chamber performances and of the background processes. I am also grateful to Florian Goebel for his suggestions on the electron reconstruction. Under the coordination of Alessandro Bertolin, Leonid Gladilin and Matthew Wing and thanks to their constant support it was possible the release our results to the summer conferences in 2002. The results were validated by an independent analysis performed by Katarzyna Klimek, whom I kindly thank for the frenetic work and her constant optimism.

In various stages of the work I profited from the discussion with several theoretical physicists. In particular, I would like thank Matteo Cacciari, Brian Harris, Hannes Jung and Gavin Salam for their suggestions. All my gratitude goes also to Sabine Schilling and Enrico Lunghi, who helped me to disentangle the QCD calculations at all orders of perturbation theory and beyond.

Working at the ZEUS Micro Vertex Detector project was indeed a great scientific opportunity as well as a nice experience of team work. A very special thank goes to Uli Kötz and Roberto Carlin for their guidance and help in the acquisition and analysis of the MVD test beam data. The first measurements on MVD half modules were performed by Ainas Weber whose analysis framework helped me to profitably achieve the results detailed in this thesis. I would also like to thank Maria Carmela Petrucci and Ignacio Redondo with whom I shared the pleasure to unveil some properties of the MVD sensors.

The MVD was successfully installed in the ZEUS detector thanks to the efforts of many excellent physicists. Here I would like to thank all the colleagues with whom I shared the exhausting period of the MVD system test. Many thanks in particular to Alessandro Polini, Chris Youngman, Roberto Carlin, Uli Kötz, Bern Surrow, Els Koffeman and Henk Tiecke.

The time spent in Hamburg was unforgettable thanks to the colleagues and friends in and outside DESY, whatever their nationality was, is and will be. Many thanks the entire crew of

DESY students and post-docs of Hamburg University and in particular to Volker Adler, Ingo Bloch, Devis Contarato, Dominik Dannheim, Oliver Gutsche, Genosse Frank Hönniger, Kostya Holubyev, Benjamin Kahle, Gregor Kramberger, Damir Lelas, Chi-Nhan Nguyen, Ioana Pintilie, Amita Raval, Andreas "Mokka-Fix" Schramm and Jörg Stahl. Each of them (and many others) contributed to the beautifully deconstructed picture I have of this country. On the other hand, the Italians always reminded me where I come from and where we are going to. My gratitude also goes to the everyday colleagues and friends of the *Arbeitsgruppe Nukleare Meßtechnik* for not being able to find a better name to our group.

The last "thanks" is for my parents to whom this work his dedicated. Grazie.