Università degli studi di Padova
Facoltà di Scienze MM. FF. NN.
Dipartimento di Fisica "G. Galilei"

Dottorato di Ricerca in Fisica

XV Ciclo

MEASUREMENT OF BEAUTY PHOTOPRODUCTION AT HERA

Coordinatore: *Ch.mo Prof. Attilio Stella*

Supervisore: *Dott. Riccardo Brugnera*

Dottoranda: *Monica Turcato*

31 dicembre 2002

Alla mia famiglia

Un corpo immerso in un liquido riceve una spinta verso l'alto
di una maleducazione, di una cafonaggine,
che non si può neanche fare più il bagno, in questo paese.

Brunello Robertetti – Corrado Guzzanti

Abstract

In this thesis the study of events in which beauty quarks are produced in positron–proton interactions is reported. The study is based on the analysis of the data collected by the ZEUS detector at HERA, a lepton–proton collider working at a centre–of–mass energy of $\sim$ 300 GeV. The analysed sample corresponds to the $\sim$ 100 pb^{-1} of data collected by the ZEUS experiment between 1996 and 2000.

Beauty production has been studied in the *photoproduction* regime, in which the boson exchanged between the proton and the positron is an almost real photon ($Q^2 \simeq$ 0 GeV^2). Events were selected in which a $b\bar{b}$ couple is produced and at least one quark of the pair decays semi–leptonically into a muon. The experimental signature of the process under investigation is the presence in the final state of at least two high transverse energy jets, coming from the hadronization of the b quarks, and a muon from the b semi–leptonic decay:

$$e\ p \rightarrow b\ \overline{b}\ X \rightarrow \text{dijet}\ \mu\ X.$$

In the event selection, the muon was asked to be in the acceptance region of the ZEUS muon detectors, whose performances have been investigated in detail.

The visible beauty photoproduction cross section has been measured in the three regions, forward, barrel and rear, covered by the ZEUS muon detectors, in a kinematic region where the acceptance of the detectors was well understood.

The beauty dijet cross section:

$$\sigma(e\ p \rightarrow b\ \overline{b}\ X \rightarrow \text{dijet}\ X)$$

was then determined using Monte Carlo simulations to extrapolate for beauty decay and branching ratio. In the same kinematic region also the differential cross section $d\sigma/dx_\gamma^{\text{MEAS}}$ was determined, the x_γ^{MEAS} variable representing the fraction of the photon momentum taking part to the hard interaction with the parton extracted from the proton.

The differential cross sections as a function of the muon pseudorapidity, $d\sigma/d\eta^\mu$, and transverse momentum, $d\sigma/dp_T^\mu$, have also been measured.

All the obtained results have been compared to theoretical predictions from next–to–leading order QCD calculations.

Riassunto

In questa tesi è riportato lo studio di eventi di produzione di quark beauty in interazioni positrone–protone. Lo studio si basa sull'analisi dei dati raccolti dal rivelatore ZEUS ad HERA, un collider leptone–protone con energia nel centro di massa di $\sim$ 300 GeV. Il campione analizzato corrisponde ai $\sim$ 100 pb^{-1} di dati acquisiti dal rivelatore ZEUS tra il 1996 e il 2000.

La produzione di beauty è stata studiata nel regime di *fotoproduzione*, in cui il bosone scambiato tra il protone e il positrone è un fotone quasi reale ($Q^2 \simeq 0$ GeV^2). Si sono selezionati gli eventi in cui viene prodotta una coppia $b\overline{b}$ e almeno uno dei quark della coppia decade in modo semileptonico in un muone. La firma sperimentale del processo in analisi è la presenza nello stato finale di almeno due jet ad alta energia trasversa, provenienti dall'adronizzazione dei quark b, e di almeno un muone dal decadimento semileptonico del b:

$$e\ p \rightarrow b\ \overline{b}\ X \rightarrow \text{dijet}\ \mu\ X.$$

Nella selezione degli eventi si chiede che il muone sia nella regione di accettanza dei rivelatori di muoni, le cui prestazioni sono state analizzate in dettaglio.

La sezione d'urto visibile di fotoproduzione di beauty è stata misurata nelle tre regioni angolari, in avanti, centrale e all'indietro, corrispondenti ai rivelatori di muoni di ZEUS, in una regione cinematica in cui le accettanze degli stessi rivelatori sono ben capite.

La sezione d'urto di produzione di beauty in eventi a due jet:

$$\sigma(e\ p \rightarrow b\ \overline{b}\ X \rightarrow \text{dijet}\ X)$$

è stata poi determinata usando le simulazioni Monte Carlo per estrapolare per il decadimento del beauty in muone e per il corrispondente branching ratio. Nella stessa regione cinematica si è poi calcolata la sezione d'urto differenziale $d\sigma/dx_\gamma^{\text{MEAS}}$, ove la variabile x_γ^{MEAS} rappresenta la frazione del momento del fotone che partecipa all'interazione dura con il partone estratto dal protone.

Sono anche state misurate le sezioni d'urto differenziali in funzione della pseudorapidità e del momento trasverso del muone, $d\sigma/d\eta^\mu$ e $d\sigma/dp_T^\mu$.

Tutti i risultati ottenuti sono stati confrontati con le predizioni teoriche della QCD al next–to–leading order.

Contents

Introduction **5**

1 Physics at HERA **7**
1.1 Electron–proton scattering . . . 7
1.2 Photoproduction processes . . . 8
1.2.1 Direct and resolved processes . . . 11
1.2.2 Photoproduction cross sections . . . 12
1.2.3 Heavy quark production . . . 14
1.3 Heavy quark production in DIS . . . 22
1.4 The heavy flavour production study . . . 23
1.4.1 Heavy flavour production at fixed–target experiments . . . 23
1.4.2 Heavy quark production at HERA . . . 23
1.4.3 Beauty production at the Tevatron . . . 36

2 The ZEUS detector at HERA **41**
2.1 The HERA collider . . . 41
2.2 The ZEUS detector . . . 44
2.2.1 The Central Tracking Detector (CTD) . . . 47
2.2.2 The Uranium–scintillator Calorimeter (UCAL) . . . 50
2.2.3 The Muon Detectors . . . 53
2.2.4 The Forward Muon Detector (FMUON) . . . 53
2.2.5 The Barrel and Rear Muon Detector (B/RMUON) . . . 55
2.2.6 The luminosity measurement . . . 57
2.2.7 The ZEUS trigger system . . . 58

3 Physics simulation **63**
3.1 General structure of an event generator . . . 63
3.2 The HERWIG Monte Carlo . . . 65
3.3 The PYTHIA Monte Carlo . . . 66
3.4 The detector simulation . . . 67
3.5 Monte Carlo samples used for the analysis . . . 69
3.5.1 Beauty events simulation . . . 69
3.5.2 Charm events simulation . . . 70
3.5.3 Light quark events simulation . . . 71

4 Event reconstruction and data selection **73**
4.1 Reconstruction of y and Q^2 . . . 73
4.2 Jet reconstruction . . . 75
4.2.1 Cone algorithms . . . 75
4.2.2 Clustering algorithms . . . 76
4.3 Muon reconstruction . . . 78
4.3.1 The MPMATCH matching package . . . 78
4.3.2 The BREMAT matching package . . . 79
4.4 Trigger chain for online data selection . . . 81
4.4.1 Jet First Level Trigger . . . 81
4.4.2 Jet Second Level Trigger . . . 81
4.4.3 Jet Third Level Trigger . . . 82
4.5 Offline data selection . . . 83
4.6 Rejection of DIS events . . . 83
4.7 Reconstruction of calorimetric variables . . . 84
4.7.1 Calorimeter–related corrections . . . 84
4.7.2 ZUFOs . . . 86
4.7.3 Results on y_{JB} . . . 91
4.7.4 Results on the jet variables . . . 93
4.8 Cuts on the calorimetric variables . . . 97
4.8.1 Cuts on y_{JB} . . . 97
4.8.2 Cuts on the jets . . . 98
4.9 Cuts on the muons . . . 98
4.9.1 Cuts on muons in the forward region . . . 98
4.9.2 Cuts on muons in the barrel and rear region . . . 103
4.10 Summary of the offline cuts . . . 107

5 Selection of beauty events **113**
5.1 μ–jet association . . . 113
5.2 Analysis of the background . . . 114
5.2.1 Determination of background from light quark events . . . 115
5.2.2 Determination of background from charm events . . . 116
5.3 Beauty fraction in the data: $p_T^{\rm rel}$ fit . . . 117
5.4 The x_γ distribution . . . 120
5.5 Results of the fits . . . 123
5.6 Control plots . . . 126

6 Measurement of the beauty cross sections **131**
6.1 Determination of the NLO QCD prediction . . . 131
6.2 Visible cross sections . . . 132
6.3 Dijet cross sections . . . 138
6.4 Muon differential cross sections . . . 142
6.5 Summary and discussion on the results . . . 147
6.6 Determination of the systematic uncertainties . . . 148

Conclusions **153**

A Muon chambers performances **157**
A.1 Monte Carlo simulation of the BRMUON chambers . . . 157
A.2 Method . . . 158
A.3 Selection of good runs . . . 159
A.4 Data analysis . . . 160
A.5 BREMAT performance . . . 164

B Pulls of the muon reconstruction algorithms **173**

C Values of the cross sections **179**
C.1 Cross section in the forward, barrel and rear regions . . . 179
C.2 Dijet cross sections . . . 180
C.2.1 Beauty dijet cross section . . . 180
C.2.2 $d\sigma/dx_\gamma^{\rm MEAS}$ . . . 180
C.3 Muon differential cross sections . . . 181
C.3.1 $d\sigma/d\eta^\mu$ . . . 181
C.3.2 $d\sigma/dp_T^\mu$ . . . 181

Introduction

The study of beauty production is an interesting topic in High Energy Physics since it represents an important testing ground for perturbative QCD. It is generally believed, in fact, that the mass of the beauty quark, around three times larger than that of the charm quark, can make perturbative QCD predictions more reliable than in the case of charm.

Beauty production in ep interactions has been studied for the first time at HERA (*Hadron Electron Ring Anlage*), the first and so far only electron–proton collider ever built. At the centre–of–mass energy available at HERA, $\sqrt{s} = 318$ GeV, the cross section of beauty production is much higher than at the previous fixed–target experiments, having centre–of–mass energies two orders of magnitude lower: there, beauty production could not be observed.

Beauty production in ep interactions proceeds through two main processes: deep inelastic scattering, where the virtuality of the exchanged photon goes from some GeV^2 to the kinematic limit of 10^5 GeV^2, and photoproduction processes, in which the exchanged particle is an almost real photon, whose virtuality has a mean value of $\sim 10^{-3}$ GeV^2.

The study of beauty production has been an intriguing topic in the last few years, due to the fact that in different experiments studying $p\bar{p}$ interactions, like UA1, CDF and DØ, the measured beauty cross sections have been found higher than next–to–leading–order QCD predictions. This discrepancy has been confirmed by the first HERA results, regarding beauty photoproduction with semi–leptonic decay into electrons or muons. So far, the description of beauty production cross section by NLO QCD, the so called *beauty puzzle*, is an open problem.

The work described in this thesis regards the study of beauty production in photoproduction processes. The b quark is tagged via its semi–leptonic decay into muons. The experimental signature of the process under analysis:

$$e\,p \rightarrow \quad b\,\overline{b}\,X \quad \rightarrow \mu \text{ dijet } X$$

is the presence in the final state of at least two jets coming from the hadronization of the b quarks and one muon from the semi–leptonic beauty decay. The data sample used for the analysis corresponds to the ~ 100 pb^{-1} collected by the ZEUS experiment between 1996 and 2000. The new features of this work with respect to the previous published and preliminary ZEUS results are that now almost all the available statistics have been used for the analysis, and that the angular region of the muon has been enlarged to include also the high–pseudorapidity range.

In Chapter 1 a theoretical introduction on the beauty production mechanisms is reported, together with a summary of the experimental results on beauty production obtained so far, in particular at HERA.

In Chapter 2 the HERA accelerator and the ZEUS detector are described briefly, with particular emphasis on the detector components relevant for the analysis.

In Chapter 3 a brief description of the Monte Carlo generators used in this work is reported, showing in particular the common and different aspects of the physics simulation.

In Chapter 4 the algorithms used for data selection are illustrated in some detail. The online data selection, based on a particular trigger chain, is described. The procedures used for muons and jets reconstruction at the offline level are illustrated, and the cuts to be applied to the data for the selection of a beauty enriched sample are defined.

Chapter 5 describes the extraction of beauty events from the data sample selected in Chapter 4. The discrimination between beauty and background events is performed by analysing the muon and jet dynamic and taking advantage of the large mass of the b quark.

Chapter 6 reports the results regarding the beauty visible and differential cross sections, compared to NLO QCD predictions.

At the end, the summary of the work is reported.

Appendix A describes a study on the muon reconstruction algorithm performance, for the barrel and rear regions. Differences between data and simulated events have been investigated in detail.

In Appendix B the pulls of the muon reconstruction algorithms are shown.

In Appendix C the obtained values of the different measured beauty cross sections are shown in Tables, together with the NLO QCD predictions.

Chapter 1

Physics at HERA

1.1 Electron–proton scattering

In the Standard Model an interaction between a proton and a lepton – electron or positron – occurs through the exchange of a vector boson. For neutral current (NC) processes, in which the final state lepton is an electron or positron, the vector boson can be a γ or a Z^0, and for charged current (CC) interactions, in which the final state lepton is a neutrino, the exchanged vector boson is a W. The inclusive process (the particles momenta are indicated in parentheses),

$$e(p_e)\ p(P) \rightarrow l^{\pm}(p_l)\ X(P'), \tag{1.1}$$

displayed schematically in Fig 1.1, can be described in terms of the following kinematic variables:

- the centre–of–mass energy squared, s,

$$s = (p_e + P)^2; \tag{1.2}$$

- the squared momentum transferred,

$$Q^2 = -q^2 = -(p_e - p_l)^2; \tag{1.3}$$

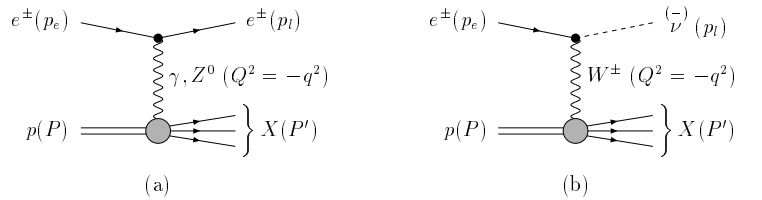

Figure 1.1: *Electron–proton scattering in Neutral Current (NC)* **(a)** *and Charged Current (CC)* **(b)** *processes.*

- the centre–of–mass energy W of the photon–proton system,

$$W^2 = (q + P)^2 \approx sy - Q^2; \tag{1.4}$$

- the Bjorken–x variable,

$$x = \frac{Q^2}{2P \cdot Q} \tag{1.5}$$

  which, in the *infinite momentum frame*[1], is the fraction of the proton momentum carried by the parton interacting with the lepton;

- the inelasticity,

$$y = \frac{P \cdot Q}{P \cdot p_e} \tag{1.6}$$

  which, in the proton rest frame, is the fraction of the lepton momentum transferred to the photon.

The kinematic variables given above are not independent. Once the centre–of–mass energy $\sqrt{s}$ is given, the kinematic of the event is completely determined by the knowledge of two of the other variables defined above. Usually the (x, y) or (x, Q^2) couple is chosen, from which the energy and direction of the outgoing lepton can be determined.

1.2 Photoproduction processes

The Q^2 variable previously defined is used to distinguish between two kinematic regimes, the Deep Inelastic Scattering (DIS) and the Photoproduction (PHP) regime. The DIS regime is defined when the squared momentum transferred, Q^2, is much larger than the proton mass. In this case the incoming lepton is deflected at some measurable angle and can be identified in the detector. On the other hand, if the squared momentum, Q^2, is of the order of 1 GeV^2, the incoming lepton is deflected at a very small angle and it cannot be observed in the detector (PHP regime). Since the study of photoproduction events is reported in this thesis, this regime is described here in some detail.

In photoproduction processes, since $Q^2 \simeq 0$, the boson exchanged between the lepton and the proton is an on–shell photon. The contribution from weak interactions is in fact negligible, since the Z^0 and W boson propagators are vanishing (they are expressed as $Q^2/(M^2_{W/Z} + Q^2)$). Therefore the lepton–proton interaction in this regime can be thought of as the scattering of a real photon on a proton.

It was first observed in fixed target photoproduction experiments, using real photon beams, [1] that the photon, besides coupling directly to one of the charged components of the proton, can also behave like a hadron. This hadronic behaviour of the photon has been interpreted in terms of the Vector Dominance Model (VDM) [2]. The photon–proton interaction is described as a two–step phenomenon: firstly the photon fluctuates

[1] The *infinite momentum frame* is defined as the reference system in which the proton is moving with very high momentum, at limit *infinite*, so that the masses are negligible. In this frame the proton can be considered as a parallel stream of particles, all moving in the same direction, since the transverse momenta are negligible. The relativistic time dilation slows down the time at which the proton constituents interact to each other, so that in the scattering with the lepton they can essentially be considered as *free* particles.

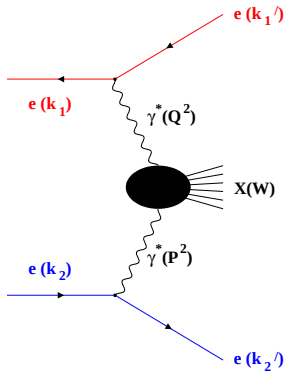

Figure 1.2: *Photon–proton scattering in a e^+e^- collision. If $P^2 = -(k_2 - k_2')^2 \simeq 0$ and $Q^2 = -(k_1 - k_1')^2 > 0$ the photon structure functions can be studied: the high–Q^2 photon is the probe, while the other photon is the target.*

into a vector meson state with the same quantum numbers, $J^{PC} = 1^{--}$ (ρ, ϕ, ω), then the nucleon interacts with the vector meson. The total cross section can be expressed as:

$$\sigma(\gamma N \to X) = \sum_{V=\rho,\phi,\omega} \frac{\pi\alpha}{\gamma_V^2} \sigma(VN \to X) \tag{1.7}$$

where γ_V is the photon–meson coupling that can be measured from the meson decay rates like $V \to e^+e^-$. The photoproduction cross section is suppressed by an $\alpha_{\rm em}$ factor with respect to the purely hadronic one, but the characteristics of the final state have been found to be the same.

The hadronic structure of the photon has been studied at e^+e^- colliders, in deep inelastic scattering $e\gamma$ events. In these processes, one highly virtual photon emitted by one of the incoming leptons interacts with an almost real photon from the other lepton (Fig. 1.2). The process $e^+e^- \to e^+e^- hadrons$ can be regarded as an interaction between a real photon and an electron, $e\gamma \to e\, hadrons$. The flux of quasi–real photons from the other lepton is factorized using the Equivalent Photon Approximation (EPA) [3]:

$$\frac{d\sigma}{dz}(ee \to eeX) = f_{\gamma/e}(z)\sigma(e\gamma \to eX), \tag{1.8}$$

where z is the fraction of the lepton energy carried by the photon, and the photon flux $f_{\gamma/e}$ is given by:

$$f_{\gamma/e}(z) = \frac{\alpha}{2\pi}\left[\frac{1+(1-z)^2}{z}\log\frac{P^2_{\rm max}}{P^2_{\rm min}} - 2m_e^2 z\left(\frac{1}{P^2_{\rm min}} - \frac{1}{P^2_{\rm max}}\right)\right]. \tag{1.9}$$

$P^2_{\rm min}$ is the kinematic limit,

$$P^2_{\rm min} = \frac{m_e^2 z^2}{1-z} \tag{1.10}$$

and $P^2_{\rm max}$ corresponds to the maximum scattering angle $\theta_{\rm max}$ for the electron not to be detected,

$$P^2_{\rm max} = (1-z)\,E^2_{\rm beam}\theta^2_{\rm max}. \tag{1.11}$$

The cross section of the deep inelastic $e\gamma$ scattering is written in analogy to the electron–nucleon DIS, introducing the two photon structure functions F_2^γ and F_2^L:

$$\frac{d\sigma(e\gamma \to eX)}{dE_{\rm tag}d\theta_{\rm tag}} = \frac{4\pi\alpha^2 E_{\rm tag}}{Q^4 y}\left\{\left[1+(1-y)^2\right]F_2^\gamma(x,Q^2) - y^2F_L^\gamma(x,Q^2)\right\} \tag{1.12}$$

where $E_{\rm tag}$ and $\theta_{\rm tag}$ are the energy and polar angle of the electron identified in the detector (the lepton that emits the quasi–real photon is deflected at a very small angle and is usually lost in the beam pipe). The y variable in $e\gamma$ scattering is defined in analogy to the ep scattering as $y = (p\cdot q)/(p\cdot k)$, where p is the momentum of the virtual photon and q that of the quasi–real γ, and k is the momentum of the incoming electron. In Eq. 1.12 y^2 is usually rather small and therefore F_L^γ can be neglected; the cross section measurement thus corresponds to the determination of the F_2^γ structure function.

The structure of the photon can be studied also in ep collisions. In the photoproduction regime, when the exchanged γ is almost real, the lepton–proton interaction can be thought of as a γp scattering, the photon direction being collinear with the lepton beam. By analogy to the e^+e^- scattering, the electroproduction cross section can be expressed in terms of the photoproduction cross section and the flux of the quasi–real photons from the lepton vertex:

$$\sigma_{ep} = \phi_{\gamma/e}\sigma_{\gamma p}. \tag{1.13}$$

The photon flux, $\phi_{\gamma/e}$, can be obtained from the expression shown above for the e^+e^- scattering (Eq. 1.9) by replacing the energy fraction z with the inelasticity y:

$$\phi_{\gamma/e} = \frac{\alpha}{2\pi}\left[\frac{1+(1-y)^2}{y}\ln\left(\frac{Q^2_{max}}{Q^2_{min}}\right) - 2m_e^2 y\left(\frac{1}{Q^2_{min}} - \frac{1}{Q^2_{max}}\right)\right] \tag{1.14}$$

where $Q^2_{\rm min}$ and $Q^2_{\rm max}$ have the same meaning of $P^2_{\rm min}$ and $P^2_{\rm max}$ defined in Eqs. 1.10 and 1.11.

The successful approach to F_2^γ, firstly suggested by Glück and Reya [4], requires the introduction of boundary conditions based on experimental data at some scale Q_0^2. The description of the Q^2 evolution is then given by perturbative QCD (DGLAP equations [5]). At the starting scale, Q_0^2, the parton densities are defined with a quark–antiquark component $q^\gamma(x,Q_0^2) = \bar{q}^\gamma(x,Q_0^2)$ and a gluon component $g^\gamma(x,Q_0^2)$. The evolution equations of the photon parton densities differ from those of the nucleon due to the anomalous coupling $\gamma \to q\bar{q}$. A peculiar characteristic of the photon structure is that there is no momentum sum rule which constrains the photon parton distributions.

The first data on F_2^γ came from the e^+e^- colliders, firstly PETRA and PEP [6], then from TRISTAN [7] and LEP [8] experiments. The photon structure functions used in this thesis have been firstly introduced on the basis of these data, and then have been further investigated and confirmed using γp data from the HERA collider.

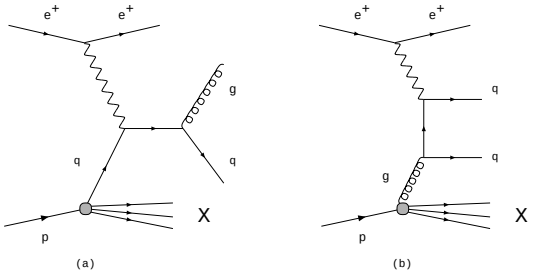

Figure 1.3: *Leading order direct photon processes:* **(a)** *QCD compton and* **(b)** *boson–gluon fusion.*

1.2.1 Direct and resolved processes

In γp interactions the photon can act in two different ways. In the *direct* photon processes (also referred to as *pointlike component*), the photon couples directly to a parton in the proton, whereas in *resolved* photon processes the photon acts as a source of partons, one of which takes part in the hard interaction. Resolved photon processes originate from the γ fluctuation into a $q\bar{q}$ or a more complex state, and depending on the relative transverse momentum between the partons this can be a bound state (as in the VDM) (*hadronic component*) or a perturbatively calculable state (*anomalous component*). Examples of direct (resolved) photon processes are shown in Fig. 1.3 (Fig. 1.4).

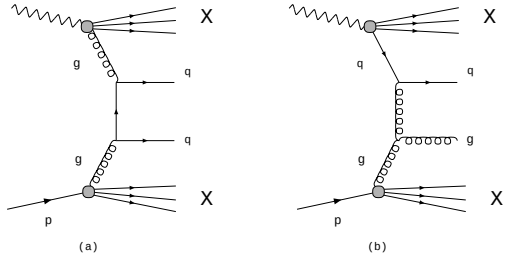

Figure 1.4: *Leading order resolved photon processes.*

In the theoretical cross section calculations, the emission of a collinear light parton from the photon before the γp interaction, as depicted in Fig. 1.4, leads to singularities in the photon–hadron cross section. These singularities are re–absorbed by introducing the concept of photon structure functions [9]. A scale μ_γ is set at which singularities coming from collinear parton emission are subtracted, and the photon structure functions depend on this scale. Neither the pointlike nor the hadronic component are separately independent of μ_γ, since the subtracted term in the pointlike component

is responsible for the redefinition of the photon structure functions in the hadronic component.

In γp interactions, the energy scale is not given by Q^2, but a *hard* interaction is defined when the produced particles have high transverse momenta. In Leading Order (LO) QCD, the expression of the strong coupling constant depends on the energy scale of the process in the following way:

$$\alpha_s(\mu^2) = \frac{12\pi}{(33-2n_f)\ln\left(\frac{\mu^2}{\Lambda^2}\right)}, \tag{1.15}$$

where Λ, the QCD scale, is of the order of 0.2 MeV. When the energy scale, μ, of the process is much higher than the QCD scale, perturbative theory can be successfully applied.

1.2.2 Photoproduction cross sections

Given the two different photon behaviours, a differential photon–hadron cross section can be written as the sum of a pointlike and a hadronic photon contribution:

$$d\sigma^{(\gamma H)}(P_\gamma, P_H) = d\sigma^{(\gamma H)}_{point}(P_\gamma, P_H) + d\sigma^{(\gamma H)}_{hadr}(P_\gamma, P_H), \tag{1.16}$$

where P_γ and P_H are the probability distributions of partons in the photon and in the hadron, respectively.

Using the factorization theorems in QCD [10], the two contributions can be written as:

$$d\sigma^{(\gamma H)}_{point}(P_\gamma P_H) = \sum_j \int dx\, f_j^{(H)}(x,\mu_f) d\hat{\sigma}_{\gamma j}(P_\gamma, xP_H, \alpha_s(\mu_R), \mu_R, \mu_F, \mu_\gamma), \tag{1.17}$$

and

$$\begin{aligned} d\sigma^{(\gamma H)}_{hadr}(P_\gamma P_H) &= \sum_{ij} \int dx\, dy\, f_i^{(\gamma)}(x,\mu_\gamma) f_j^{(H)}(y,\mu'_f) \\ &\quad \times d\hat{\sigma}_{ij}(xP_\gamma, yP_H, \alpha_s(\mu'_R), \mu'_R, \mu'_F, \mu_\gamma). \end{aligned} \tag{1.18}$$

Here the $f_i^{(\gamma)}$ are the partonic densities in the photon, and in a similar way the $f_j^{(H)}$ are the partonic densities in the hadron; μ_R and μ'_R are renormalization scales, μ_F and μ'_F are factorization scales for singularities arising from strong interactions, and μ_γ (defined in 1.2.1) is a factorization scale for singularities arising from the electromagnetic vertex. The hadronic parton densities obey the usual DGLAP evolution equations [11] in μ_F,

$$\frac{\mathrm{d}f_i(x,\mu_F)}{\mathrm{d}\log\mu_F^2} = \frac{\alpha_s}{2\pi}\int \frac{\mathrm{d}y}{y} P_{ki} f^k(y/x,\mu_f), \tag{1.19}$$

whereas the photonic parton density, not calculable in perturbation theory, obey a renormalization group equation [12] that can be obtained by slightly modifying the DGLAP equation:

$$\frac{\partial f_i^{(\gamma)}}{\partial \log\mu^2} = \frac{\alpha_{\mathrm{em}}}{2\pi} P_{i\gamma} + \frac{\alpha_s}{2\pi}\sum_j P_{ij} \otimes f_j^{(\gamma)}. \tag{1.20}$$

At lowest order

$$P_{i\gamma} = N_c e_i^2 (x^2 + (1-x)^2) \tag{1.21}$$

where N_c is the number of colours and e_i is the electric charge in units of the positron charge. Equation 1.20 differs from the evolution equation of the hadron densities for the presence of the first term, which is due to the direct coupling of the photon and the quarks.

The strict interplay between the direct and resolved components can be understood by showing the derivation of the subtracted cross sections that appears in equations 1.17 and 1.18 [13]. Firstly the partonic cross sections for the processes $\gamma+j \to Q+\bar{Q}+X$ and $i+j \to Q+\bar{Q}+X$ are evaluated. After the ultraviolet renormalization, the quantities $d\sigma_{\gamma i}$ and $d\sigma_{ij}$ have still some divergences due to the collinear emission of a massless final–state parton by one of the incoming partons. These divergences are subtracted using the factorization theorem. The partonic cross sections in next–to–leading order (NLO) QCD can be expanded in:

$$d\sigma_{\gamma i} = \alpha_{\mathrm{em}}\alpha_s d\sigma^{(0)}_{\gamma j} + \alpha_{\mathrm{em}}\alpha_s^2 d\sigma^{(1)}_{\gamma j} \tag{1.22}$$

$$d\sigma_{ij} = \alpha_s^2 d\sigma^{(0)}_{ij} + \alpha_s^3 d\sigma^{(1)}_{ij} \tag{1.23}$$

The subtracted cross sections, using dimensional regularization ($d = 4-2\epsilon$), are:

$$d\hat{\sigma}^{(0)}_{\gamma j}(p_1,p_2) = d\sigma^{(0)}_{\gamma j}(p_1,p_2), \quad d\hat{\sigma}^{(0)}_{ij}(p_1,p_2) = d\sigma^{(0)}_{ij}(p_1,p_2), \tag{1.24}$$

$$\begin{aligned} d\hat{\sigma}^{(1)}_{\gamma j}(p_1,p_2) &= d\sigma^{(1)}_{\gamma j}(p_1,p_2,\frac{1}{\bar{\epsilon}}) + \frac{1}{2\pi}\sum_k \int dx \left(\frac{1}{\bar{\epsilon}}P_{k\gamma}(x) - H_{k\gamma}(x)\right) d\sigma^{(0)}_{kj}(xp_1,p_2) \\ &\quad + \frac{1}{2\pi}\sum_k \int dx \left(\frac{1}{\bar{\epsilon}}P_{kj}(x) - K^{(H)}_{kj}(x)\right) d\sigma^{(0)}_{\gamma k}(p_1,xp_2), \end{aligned} \tag{1.25}$$

$$\begin{aligned} d\hat{\sigma}^{(1)}_{ij}(p_1,p_2) &= d\sigma^{(1)}_{ij}(p_1,p_2,\frac{1}{\bar{\epsilon}}) + \frac{1}{2\pi}\sum_k \int dx \left(\frac{1}{\bar{\epsilon}}P_{ki}(x) - K^{(\gamma)}_{ki}(x)\right) d\sigma^{(0)}_{kj}(xp_1,p_2) \\ &\quad + \frac{1}{2\pi}\sum_k \int dx \left(\frac{1}{\bar{\epsilon}}P_{kj}(x) - K^{(H)}_{kj}(x)\right) d\sigma^{(0)}_{ik}(p_1,xp_2). \end{aligned} \tag{1.26}$$

The term $1/\bar{\epsilon}$ indicates that infrared divergences are still present. The functions $K^{(H)}_{ij}$, $K^{(\gamma)}_{ij}$ and $H_{k\gamma}$, defining an extra finite part of the subtraction, are completely arbitrary; different choices correspond to different subtraction schemes. In the $\overline{MS}$ scheme, $H = K \equiv 0$.

The presence in Eq. 1.25 of the term in $P_{k\gamma}(x)$ originates from the inhomogeneous term in Eq. 1.20, and therefore from the direct coupling of the photon and the quarks. This means that the collinear divergences arising from the emission of collinear quarks from the incoming photon, in the pointlike component, are re–absorbed into the partonic densities of the photon. These appear in the hadronic component, as it was pointed out in Sect. 1.2.1.

From the discussion above it can be concluded that, since the pointlike and hadronic component are so strongly related, only their sum has a physical meaning, and that beyond the LO the discrimination between direct and resolved processes is ambiguous. As an example, the process in Fig. 1.5 is part of the $\mathcal{O}(\alpha_{\mathrm{em}}\alpha_s^2)$ cross section, and

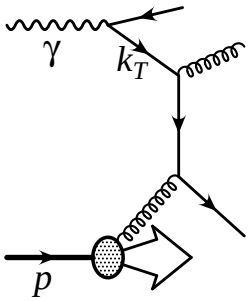

Figure 1.5: *A photon–proton interaction. If the virtuality k_T of the propagator shown in the figure is lower than the factorization scale μ_γ the process is included in the resolved calculation at leading order, otherwise it is thought of as a direct process at next–to–leading order.*

therefore can be thought of as a direct process at NLO, but if the virtuality of the propagator, k_T^2, is less than μ_γ^2, then its contribution is already included in the LO resolved process. Nevertheless, bearing in mind the previous discussion, the terms *direct* and *resolved* component will be used again, since a first–order approximation can be used in order to get a physical picture of the separation between the two components.

Since the photon parton densities are quite soft, the contribution from the hadronic component becomes important only for large centre–of–mass energy and small masses of the produced particles. In heavy quark production, this can affect charm production much more than beauty production.

1.2.3 Heavy quark production

In lepton–proton collisions, at HERA, the cross section of heavy quark production is dominated by photoproduction processes. Heavy quark photoproduction is a phenomenon of considerable interest since it is closely related to hadroproduction, but it is considerably simpler since the incoming photon is much better understood than an incoming hadron. It is a good testing ground for perturbative QCD, and helps in understanding the proton and the photon structure.

Heavy flavour photoproduction has been extensively studied in fixed–target experiments [14], at a centre–of–mass energy of ~ 30 GeV, but at the lepton–proton collider HERA, which operates at a centre–of–mass energy of ~ 300 GeV, another kinematic region can be explored, where both beauty and charm production can be observed. Since the b mass is much heavier than the charm mass, perturbative QCD calculations are expected to be more reliable.

At Leading Order, heavy flavour production proceeds via the so–called *boson–gluon fusion* process (Fig. 1.3), in which a photon from the incoming lepton interacts with a

gluon from the proton giving a $Q\overline{Q}$ couple in the final state:

$$\gamma + g \rightarrow Q + \overline{Q}. \tag{1.27}$$

At the next order in α_s the following processes have to be considered:

$$\gamma + g \rightarrow Q + \overline{Q} + g, \tag{1.28}$$
$$\gamma + q \rightarrow Q + \overline{Q} + q. \tag{1.29}$$

The *hadronic* component, in which the photon fluctuates into quarks and gluons before the hard interaction, has to be added to these *direct* photon processes.

The calculation of the NLO cross section for heavy flavour production has been performed recently. Firstly R.K. Ellis and P. Nason evaluated the QCD corrections to the direct $\gamma p \rightarrow QX$ process [15], and their calculations were then confirmed by Smith and van Neerven [16]. The NLO corrections for parton–parton scattering, needed for the hadronic part, were produced by P. Nason, S. Dawson and R.K. Ellis in [17] and by W. Beenakker et al. in [18].

S. Frixione et al. have then developed computer codes [9, 19] that integrate the NLO formulas producing cross sections useful for comparison with data.

These perturbative QCD calculations rely on the fact that the mass of the heavy quark acts as a cutoff for infrared singularities so that the splitting processes in which a heavy quark is involved are finite order by order in perturbation theory. For this reason the calculations in [15, 16, 17] have been done with a modified version [20] of the $\overline{MS}$ subtraction scheme, in which the heavy quark effects decouple in processes involving momenta much smaller than the heavy quark mass. This implies that the heavy quark does not contribute to the evolution of the running coupling constant or of the structure function. Therefore subprocesses initiated by an intrinsic heavy flavour generated by the structure function (the so–called *excitation processes*) are not present. This method is referred to as the *massive scheme*.

Nevertheless, the NLO one particle inclusive differential distributions contain terms like $\alpha_s \log(p_T^2/m^2)$ which become large in the high–p_T limit. This is reflected in a large sensitivity to the choice of the renormalization and factorization scales, and therefore in a large uncertainty in the theoretical predictions. Also terms like $\alpha_s \log(W/m)$ can become large at HERA energies.

In the *massless scheme*, the heavy quarks are treated as additional active flavours above some threshold, of the order of the quark mass. Below this threshold the heavy quarks revert to inactive flavours, the light quarks and gluons being the only active partons, as in the massive scheme. In this approach, the resummation of the $\alpha_s \log(p_T^2/m_Q^2)$ terms to all orders is performed with the so–called Perturbative Fragmentation Function (PPF) technique [21]. The PPF approach is based on the assumption that the heavy quarks act as if they were massless when their momenta are much larger than their masses. Therefore the usual $\overline{MS}$ scheme is used, assuming all the quarks to be massless, subtracting the singularities also from their lines and introducing a fragmentation function to absorb the final state divergences. This fragmentation function contains the logarithmic terms, so that their resummation is possible. However, the method is not reliable in the region $p_T \sim m_Q$, and in particular the total cross section is not calculable in the massless scheme.

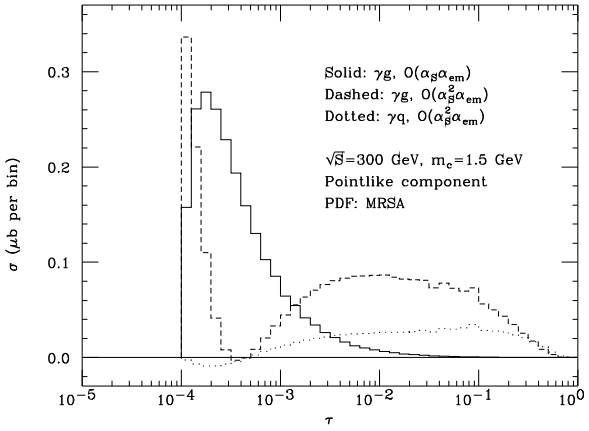

Figure 1.6: *The $\mathcal{O}(\alpha_{\rm em}\alpha_s)$ γg, $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ γg and $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ γq contributions to the total charm photoproduction cross section as a function of the variable $\tau = \hat{w}/W$. The proton renormalization and factorization scales have been taken equal to the charm mass, while the proton parton densities used are MRSA (from Ref. [22]).*

The calculations described in the following paragraphs are based on a fixed order massive approach so the logarithmic terms in $\alpha_s \log(p_T^2/m_Q^2)$ are neglected. The indetermination on the theoretical cross section calculations due to the choice of the input parameters will be investigated. The pointlike and hadronic components will be analysed separately in order to estimate the different sources of uncertainties in the two cases [22].

1.2.3.1 Charm photoproduction

In charm photoproduction at HERA, large radiative effects, which were not a problem in fixed–target experiments, are present due to the higher centre–of–mass energy. These radiative effects, proportional to $\alpha_{\rm em}\alpha_s^2 \log(W/4m_c^2)$, where W is the centre–of–mass energy of the photon–proton system and m_c the charm mass, are the first manifestation of a tower of corrections behaving as $\alpha_{\rm em}\alpha_s(\alpha_s \log(W/4m_c^2))^n$, coming from the small–$x$ region. In order to evaluate the importance of these effects, the various contributions to the charm pointlike cross sections are shown as a function of x, evaluated as the ratio of the partonic over the hadronic centre–of–mass energy $\hat{w}/W$ (Fig. 1.6). In the Figure the proton renormalization and factorization scales are set to the charm mass m_c, while the MRSA [23] parton densities are used for the proton. A hump, away from the threshold region, is clearly visible for NLO γg and γq contributions. The effect becomes even more important if, instead of choosing the MRSA parton densities for the proton, which have a relatively singular gluon distribution at small x, other parton densities, with a more regular behaviour at small x, as, for example, HMRS B [24], are used (Fig. 1.7).

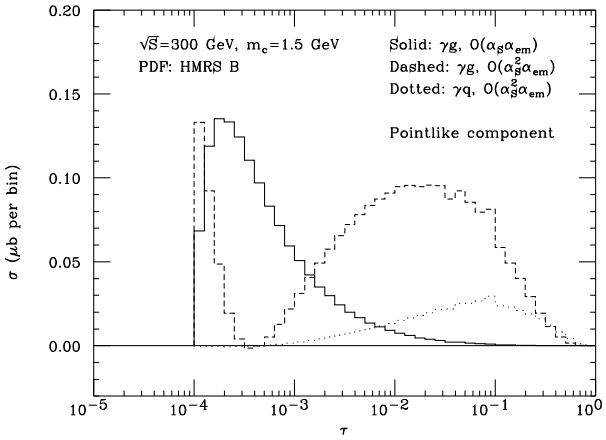

Figure 1.7: *As in Fig. 1.6, but with HMRS B proton parton densities (from Ref. [22]).*

The importance of the small–x regime can be estimated by the size of the hump relative to the total cross section. The contribution to the cross section coming from the hump is plotted and compared to the total and to the $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ cross sections (Fig. 1.8). From the Figure it can be concluded that small–x effects can have a considerable size in $c\bar{c}$ quark production at the HERA energies. A similar effect was found in the hadroproduction of b quarks [25] where the all-order resummation of small–x effects was performed. Since a similar calculation is not available for c production at HERA, the contribution of small–x effects to the cross section is assumed to be equal to the squared size of the hump relative to the total cross section as in [25].

Also in the hadronic component of the charm cross section, small–x effects are present and they can give a contribution of 45% to the total cross section at the HERA energy (20% at 50 GeV) if MRSA and GRV–HO parton densities are used for the proton and the photon, respectively. Nevertheless, the uncertainty on the final result is much smaller than other uncertainties.

Other important inputs in the computation of the total cross section are the value of the QCD scale, Λ, and the gluon density in the proton. For the calculations reported here, Λ has been chosen to vary in the interval 151 MeV $< \Lambda <$ 220 MeV in order to keep into account the values obtained from deep inelastic scattering and e^+e^- data.

The gluon density in the proton at relatively small x is constrained by deep inelastic scattering data, but at very small x the gluon density is not known with precision. For the calculations reported here, the two sets MRSD-′ and CTEQ2MF, differing strongly at small x, yet consistent with the experimental parametrizations [26], will be used as the two extreme possibilities.

The results for the pointlike component are reported in Table 1.1. The cross section is found to be very sensitive to the charm mass, whose variation changes the cross section up to a factor 4. It is also sensitive to the scale, giving a difference of a factor two for the lower choice.

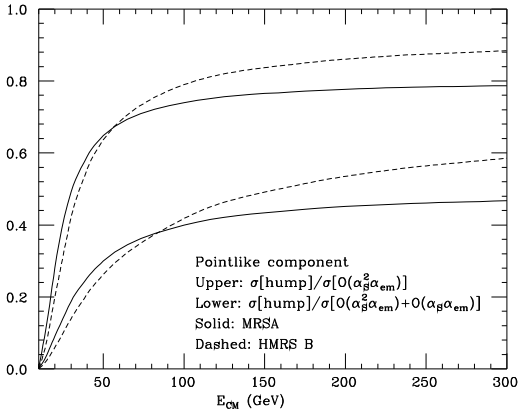

Figure 1.8: *The contribution of the hump to the total cross section relative to the* $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ *contribution* **(upper)** *and to the total cross section* **(lower)** *for the MRSA and HMRS B proton parton densities (from Ref. [22]).*

The hadronic component of the $c\bar{c}$ cross section is sensitive to the gluon distribution in the photon in the small–x region. The dominant contribution to the hadronic cross section comes from events with x_γ between 10^{-2} and 10^{-1} [22], where the x_γ variable is defined as the fraction of the photon momentum which takes part in the hard interaction. Some indications on the gluon density in the photon can come from results on dijet photoproduction at HERA [27], but the statistics are limited and do not allow to distinguish between different NLO parametrizations or different gluon densities in the photon. A reasonable way to estimate the uncertainty due to the choice of the gluon density in the photon is to take the sets GRV–HO [28] and LAC1 [29] as the two extremes. Using MRSA as the parton density in the proton, the results obtained for the hadronic component are shown in Table 1.2. At the highest photon–proton centre–of–mass energy available at HERA, the prediction for the hadronic component with LAC1 is an order of magnitude higher than the prediction with GRV–HO.

The comparison between the theoretical predictions and the experimental results [30, 31] is shown in Fig. 1.9, where only the uncertainty due to the choice of the renormalization scale is shown. Even if this is just a part of the theoretical indetermination, it should be noticed that the data are in reasonable agreement with the curves obtained by choosing $m_c = 1.5$ GeV, MRSG for the proton and GRV–HO for the photon parton densities. Also the fact that all the data can be described with a single choice of the input parameters is not trivial since the HERA centre–of–mass energy is an order of magnitude higher than that of the fixed–target experiments whose results are plotted in the Figure.

Some points have to be noticed, regarding the experimental measurement of the $c\bar{c}$ cross section. In charm analyses, the detectors are sensitive to the production in the central rapidity region and a p_T cut is usually applied to the data in order to reject

	$m_c = 1.2$ GeV			$m_c = 1.5$ GeV			$m_c = 1.8$ GeV		
$E_{\rm CM}$ μ_r	$m_c/2$	m_c	$2m_c$	$m_c/2$	m_c	$2m_c$	$m_c/2$	m_c	$2m_c$
Proton PDF set MRSA									
30 GeV	2.582	2.104	1.712	1.321	1.090	0.907	0.735	0.611	0.515
100 GeV	6.558	5.056	4.057	3.415	2.826	2.356	2.009	1.721	1.467
280 GeV	13.89	10.13	8.008	7.314	5.858	4.831	4.389	3.685	3.120
Proton PDF set CTEQ2MF									
30 GeV	2.651	2.264	1.885	1.455	1.216	1.023	0.827	0.688	0.583
100 GeV	4.371	3.856	3.236	2.699	2.425	2.088	1.761	1.593	1.393
280 GeV	7.292	5.681	4.607	4.504	3.827	3.237	3.035	2.676	2.317
Proton PDF set MRSD$-'$									
30 GeV	3.136	2.363	1.885	1.422	1.136	0.936	0.748	0.612	0.513
100 GeV	11.75	8.149	6.347	5.106	3.951	3.220	2.690	2.199	1.843
280 GeV	34.16	23.06	17.81	14.79	11.17	9.033	7.794	6.247	5.196
Proton PDF set CTEQ2M									
30 GeV	2.696	2.182	1.785	1.341	1.107	0.926	0.734	0.612	0.519
100 GeV	6.618	5.240	4.262	3.473	2.917	2.454	2.051	1.773	1.522
280 GeV	13.46	10.17	8.157	7.207	5.898	4.919	4.374	3.728	3.184
Proton PDF set CTEQ2ML									
30 GeV	3.500	2.435	1.829	1.705	1.242	0.974	0.910	0.688	0.553
100 GeV	8.818	5.625	4.171	4.270	3.169	2.497	2.451	1.942	1.583
280 GeV	18.90	10.52	7.618	8.844	6.205	4.823	5.145	3.972	3.213

Table 1.1: *Pointlike component of the total charm cross section (μb), in γp collisions, for different choices of the input parameters.*

	$m_c = 1.2$ GeV			$m_c = 1.5$ GeV			$m_c = 1.8$ GeV		
$E_{\rm CM}$ μ_r	$m_c/2$	m_c	$2m_c$	$m_c/2$	m_c	$2m_c$	$m_c/2$	m_c	$2m_c$
Photon PDF set GRV–HO									
30 GeV	0.793	0.297	0.165	0.206	0.099	0.060	0.072	0.040	0.026
100 GeV	4.572	1.680	0.924	1.438	0.673	0.405	0.589	0.314	0.200
280 GeV	15.42	5.492	2.989	5.296	2.435	1.454	2.348	1.239	0.784
Photon PDF set LAC1									
30 GeV	3.350	0.995	0.503	0.540	0.213	0.119	0.125	0.059	0.036
100 GeV	35.05	11.69	6.207	8.975	3.794	2.187	3.052	1.469	0.893
280 GeV	140.3	48.77	26.30	42.45	18.84	11.09	16.78	8.494	5.283

Table 1.2: *Hadronic component of the total charm cross section (μb), in γp collisions, for different choices of the input parameters.*

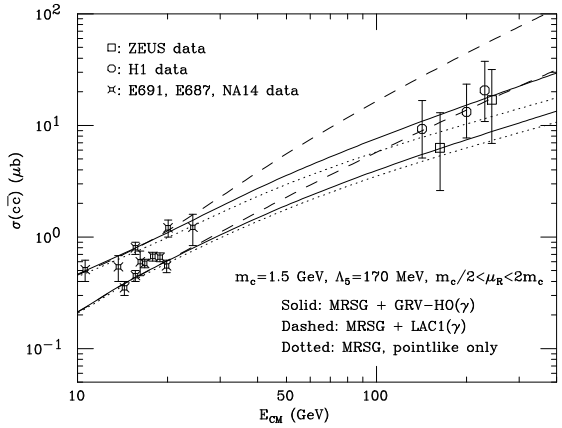

Figure 1.9: *Total cross section for $c\bar{c}$ photoproduction as a function of the γp centre–of–mass energy. Theoretical predictions and experimental data from Ref. [30] and [31] are shown.*

background events. Therefore the total cross section has to be extrapolated to the full rapidity and momentum range, and this means extrapolating to the small–x region, which is poorly known. The extrapolation is therefore subject to large uncertainties.

1.2.3.2 Beauty photoproduction

In beauty photoproduction, the uncertainties found in the case of charm are strongly reduced. The small–x problem, analysed in the same way as for charm, has been found to increase the value of the total cross section by $\sim$ 20%. The values of the cross section for the pointlike and hadronic components are shown in Tables 1.3 and 1.4, where the central value for the renormalization and factorization scales has been set to m_b = 4.75 GeV (column labelled 'DEF'). It can also be noticed in Table 1.3 that the pointlike cross section varies by a factor 2 by *simultaneously* varying all the input parameters in the direction that makes the cross section larger or smaller.

Coming to the hadronic component, the largest uncertainty comes from the choice of the parton density of the photon, as in the case of charm. However, in beauty photoproduction the small–x region is probed to a less extent than in the case of charm so the dependence on the photon density is milder. The difference on the estimated cross section is a factor 3 (to be compared to a factor 10 of the charm case) if GRV–HO and LAC1 sets are used as the two extreme possibilities. Even for beauty production, if the gluon distribution in the photon is as soft as LAC1 predicts, the hadronic component can give the dominant contribution to the cross section.

		μ_r		μ_f		m_b (GeV)	
$E_{\rm CM}$	DEF	$m_b/2$	$2m_b$	$m_b/2$	$2m_b$	4.5	5
Proton PDF set MRSA							
100 GeV	22.90	26.00	20.36	21.87	23.63	27.41	19.28
280 GeV	64.30	75.44	56.09	57.81	66.95	75.31	55.32
Proton PDF set CTEQ2MF							
100 GeV	24.91	27.93	22.33	23.80	25.73	29.70	21.02
280 GeV	62.91	73.08	55.24	56.27	65.80	72.76	54.71
Proton PDF set MRSD−′							
100 GeV	23.03	26.19	20.45	22.18	23.63	27.75	19.26
280 GeV	74.76	87.24	65.43	69.22	76.93	88.76	63.52
Proton PDF set CTEQ2M							
100 GeV	22.96	25.92	20.50	22.04	23.54	27.50	19.31
280 GeV	65.33	75.96	57.32	59.43	67.51	76.40	56.23
Proton PDF set CTEQ2ML							
100 GeV	25.07	29.53	21.76	23.38	25.92	30.03	21.07
280 GeV	71.04	86.84	60.42	61.74	74.39	82.99	61.20

Table 1.3: *Pointlike component of the total beauty cross section (nb), in γp collisions, for different choices of the input parameters.*

		μ_r		μ_f		μ_γ		m_b (GeV)	
$E_{\rm CM}$	DEF	$m_b/2$	$2m_b$	$m_b/2$	$2m_b$	$m_b/2$	$2m_b$	4.5	5
Photon PDF set GRV–HO									
100 GeV	4.244	5.589	3.290	4.246	4.258	3.883	4.525	5.444	3.347
280 GeV	27.63	38.67	20.74	25.68	29.13	25.39	28.68	34.29	22.48
Photon PDF set LAC1									
100 GeV	6.693	9.264	5.053	7.157	6.394	7.161	6.348	9.309	4.890
280 GeV	98.45	137.2	74.05	96.07	100.6	106.4	92.18	128.7	76.15

Table 1.4: *Hadronic component of the total beauty cross section (nb), in γp collisions, for different choices of the input parameters.*

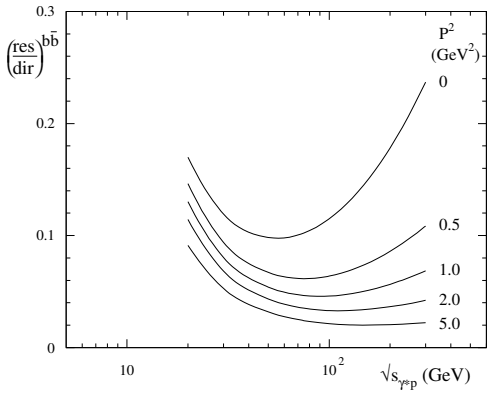

Fig. 2

Figure 1.10: *The ratio of the resolved to the direct contribution to beauty production cross section in LO QCD for various fixed values of the photon virtuality* $Q^2 \equiv P^2$ *(in* GeV^2 *units).*

1.3 Heavy quark production in DIS

In the deep inelastic scattering regime, the processes contributing to heavy flavour production are expected to be different than in the photoproduction regime. The parton densities in the photon, in fact, fall rapidly to zero when the photon virtuality approaches the hard scale of the process, which is of the order of the quark mass, as is pointed out in Ref. [32, 33, 34].

In Ref. [35] it is shown that as long as the virtuality Q^2 of the exchanged photon is much lower than the factorization scale, $Q^2/\mu_F^2 \lesssim 10^{-1}$, the partonic content of the photon is applicable to investigations concerning the *resolved photon* contributions. The resolved parton content of the photon is discernible provided it contributes significantly more than 10% to the involved production rate. These conditions are almost never satisfied for heavy quark (c, b) electroproduction as is shown in Fig. 1.10 in the case of beauty quarks. In these calculations, the parton density used are GRV 94 [36] for the proton, while those used for the photon are described in [37]. The b–quark mass is set to 4.75 GeV, and the factorization scale to $\mu_F = 4m_b^2$.

The smallness of the resolved contribution of virtual photons ($Q^2 > 0$) to the heavy quark electroproduction cross section implies that ep scattering at HERA, in the high Q^2 regime, is dominated by the direct boson–gluon fusion process, whereas the resolved contributions are negligible. Therefore, in the study of heavy flavour production at HERA, a comparison between DIS and PHP results is useful to the understanding of the two different photon behaviours.

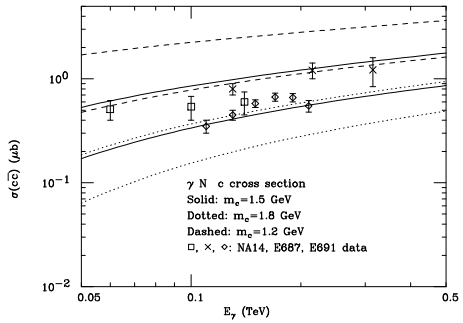

Figure 1.11: *Total cross section for* $c\bar{c}$ *production in* γN *collisions, compared to NLO QCD predictions.*

1.4 The heavy flavour production study

Heavy quark production has been studied intensively, first at fixed–target experiments, and then at the colliders like HERA and the Tevatron. In the following sections the results obtained by the various experiments will be summarized, with particular emphasis on beauty production at HERA.

1.4.1 Heavy flavour production at fixed–target experiments

Heavy quark production has been studied extensively at fixed target experiments. Since the typical centre–of–mass energy is between 10 and 40 GeV, where the beauty cross section is quite small, most of the available data concern charm production.

Because of the small value of the charm–quark mass, the perturbative expansion may not be reliable due to higher–order terms and to the non–perturbative contributions. This fact has to be kept in mind while comparing experimental results and theoretical predictions.

The production of $c\bar{c}$ pairs in γN collisions has been studied in various experiments [38, 39, 40]. The experimental results in comparison with NLO QCD predictions are shown in Fig. 1.11. The theoretical uncertainties found in photoproduction, shown here, are smaller than in hadroproduction [13]. In the Figure, the bands are obtained by varying the renormalization scale between $m_c/2$ and $2m_c$, while the factorization scale has been kept fixed to $2m_c$. The data are consistent with a charm mass of 1.5 GeV, but some of the experimental results are incompatible with one another so that no final response could be obtained.

1.4.2 Heavy quark production at HERA

At the lepton–proton collider HERA, working at a centre–of–mass energy of ~ 300 GeV, the study of heavy quarks is a subject of great interest. At that energy, as was shown in Sects. 1.2.3.1 and 1.2.3.2, the $c\bar{c}$ production cross section is of the order of 1 μb,

while the beauty cross section is $\mathcal{O}(10)$ nb. With the integrated luminosity collected until 2000, very large statistics are available for the study of charm production, and also beauty analyses can be carried on and give significant results. In the following we will illustrate some of the most recent results obtained by the ZEUS experiment on charm production, and some of the published and preliminary results of the ZEUS and H1 Collaborations on beauty production.

1.4.2.1 Charm photoproduction at HERA

At HERA, large statistics are available for the analysis of charm events. Here we want to focus on the study of open charm production, that proceeds via the search for D^* mesons, whose experimental signature is quite clear. In Ref. [41] the production of D^* mesons and of the associated jets is studied in detail, and the measured cross sections are compared to NLO QCD predictions using the massive or massless approach.

In the analysis in [41] charm was tagged by identifying $D^{*\pm}(2010)$ mesons in the final state via their charged decay products:

$$D^{*+} \to \quad D^0\pi_s^+ \quad \to (K^-\pi^+)\pi_s^+ \tag{1.30}$$
$$D^{*+} \to \quad D^0\pi_s^+ \quad \to (K^-\pi^+\pi^+\pi^-)\pi_s^+ \tag{1.31}$$

and the charge conjugated processes. The small mass difference $\Delta M = M(D^*) - M(D^0) = 145.42 \pm 0.02$ MeV [42] yields a low momentum pion, referred to as the *soft pion*, π_s, from the D^* decay, and a large signal just above the threshold of the $M(K\pi\pi_s) - M(K\pi)$ and $M(K\pi\pi\pi\pi_s) - M(K\pi\pi\pi)$ distributions, where the phase space contribution is highly suppressed. These features of the events are used to discriminate the signal from the background.

The differential cross sections as a function of the D^* transverse momentum, $d\sigma/dp_T^{D^*}$, and pseudorapidity[2], $d\sigma/d\eta^{D^*}$, were measured in the kinematic region defined by $p_T^{D^*} > 2$ GeV, $-1.5 < \eta^{D^*} < 1.5$, for $Q^2 < 1$ (photoproduction regime) and $130 < W < 280$ GeV.

The obtained results were compared to NLO QCD predictions, in the massive charm scheme [22, 43], performed using the MRSG [44] and GRV–G HO [28] parton densities for the proton and the photon, respectively. The renormalization scale chosen was $\mu_R = m_T = \sqrt{m_c^2 + p_T^2}$ ($m_c = 1.5$ GeV) and the factorization scale was set to $2m_T$. The charm fragmentation into a D^* was performed using the Peterson fragmentation function [45]. The comparison between the data and the NLO predictions is shown in Figs. 1.12 and 1.13. The predicted cross sections are well below the measured values, the values obtained with the extreme choice of the theoretical parameters being still below the data. Also the shapes of both the differential cross sections, $d\sigma/dp_T^{D^*}$ and $d\sigma/d\eta^{D^*}$, are not well described by the theory.

In the same Figures also the predictions using the massless scheme [46, 47, 48] are reported. The parton densities used were CTEQ4M [49] for the proton and GRV–G HO [28] for the photon; the renormalization and factorization scales were the same used for the calculations in the massive scheme. In the Figures the results obtained using two different massless approaches [46, 48] are reported: the perturbative and non–perturbative parts of the fragmentation function are factorised differently in the two

[2] The *pseudorapidity*, η, is defined as $\eta = -\ln(\tan\frac{\theta}{2})$, where θ is the polar angle.

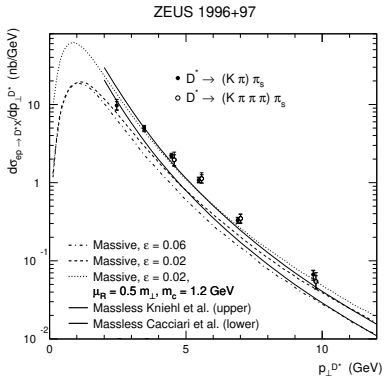

Figure 1.12: *Differential cross section $d\sigma/dp_T^{D^*}$ for D^* photoproduction for the $(K\pi)\pi_s$ (full dots) and $(K\pi\pi\pi)\pi_s$ (open dots) channels. The NLO QCD predictions in the massive and massless schemes are compared to the data (from Ref. [41]).*

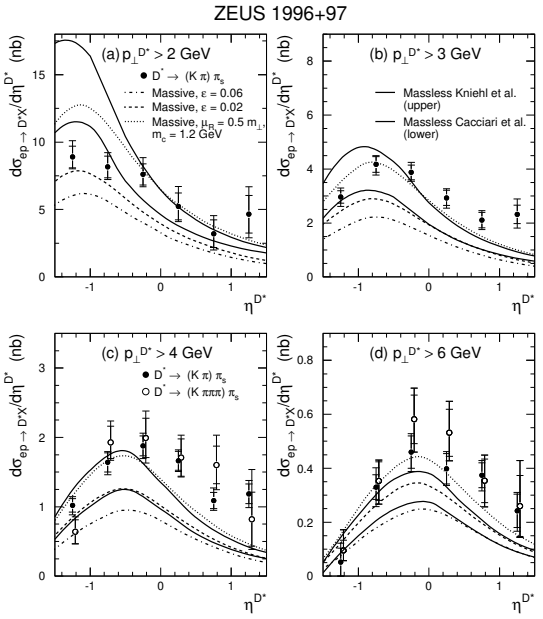

Figure 1.13: *Differential cross section $d\sigma/d\eta^{D^*}$ for D^* photoproduction, for the $(K\pi)\pi_s$ (full dots) and $(K\pi\pi\pi)\pi_s$ (open dots) channels. The NLO QCD predictions in the massive and massless schemes are compared to the data (from Ref. [41]).*

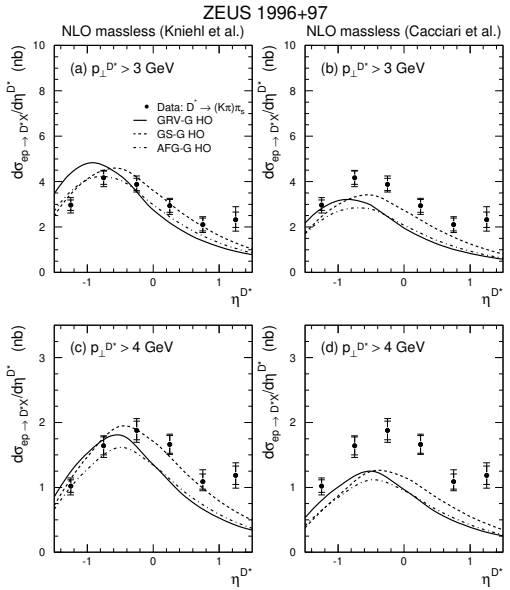

Figure 1.14: *Differential cross section $d\sigma/d\eta^{D^*}$, for the two highest–p_T bins, in the $(K\pi)\pi_s$ channel. The curves show the predictions of the massless charm NLO of Ref. [46]* **(a,c)** *and Ref. [48]* **(b,d)**, *with various photon parton density parametrizations (from Ref. [41]).*

calculations. The non–perturbative part is then fitted to the Peterson fragmentation [45], using recent e^+e^- data on D^* production to extract the ϵ parameter.

As it can be seen in Figs. 1.12 and 1.13, where the results of the two different massless approaches are shown as the full lines, the two models give similar shapes of the distributions but differ with each other in absolute magnitude by $\sim 40\%$. The predicted cross sections are mostly below the data, in the forward region in particular. The use of the MRSG [44] parton density for the proton does not change the results significantly. A bigger effect is obtained by changing the parton density of the photon, as can be seen in Fig. 1.14. The differential cross section $d\sigma/d\eta^{D^*}$ is reported in the two highest–p_T bins, where the massless predictions are expected to be more reliable, and compared to NLO massless charm predictions obtained with the photon parton densities GRV–G HO [28], GS–G HO [50] and AFG [51]. The best description of the data is given by the massless scheme of Ref. [46] with the GS–G HO [50] parton density for the photon.

From the measurement of the final state jets the kinematic of the hard process can be reconstructed. The $x_\gamma^{\rm OBS}$ variable, defined in terms of the two highest transverse

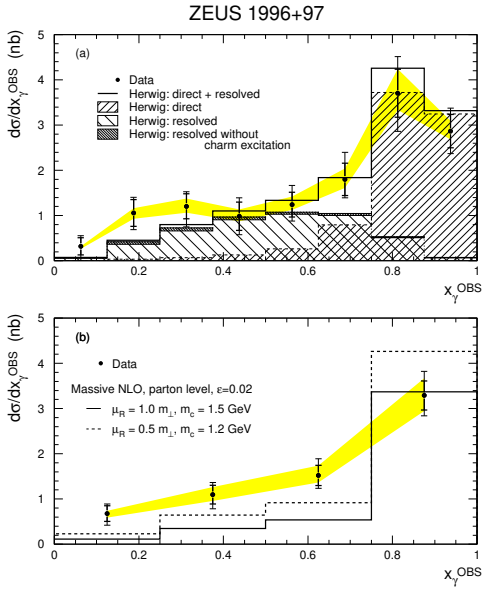

Figure 1.15: *The differential cross section $d\sigma/dx_\gamma^{\rm OBS}$ for dijet plus D^* meson events. The shaded band represents the uncertainty due to the indetermination of the calorimeter energy scale. In* **(a)** *the experimental data are compared to the expectations of the HERWIG Monte Carlo simulations, normalised to the data. In* **(b)** *the data are compared to a parton level NLO massive charm calculation [22, 43] (from Ref. [41]).*

energy jets in the event

$$x_\gamma^{\rm OBS} = \frac{E_T^{\rm jet1} e^{-\eta^{\rm jet1}} + E_T^{\rm jet2} e^{-\eta^{\rm jet2}}}{2yE_e}, \qquad (1.32)$$

where E_e is the incoming electron energy and y is the inelasticity, has the meaning of the fraction of the photon momentum which takes part to the hard interaction. Therefore an experimental separation of direct and resolved processes can be done using a cut on $x_\gamma^{\rm OBS}$: *direct* (*resolved*) processes are defined for $x_\gamma^{\rm OBS} > 0.75$ ($x_\gamma^{\rm OBS} < 0.75$).

The differential cross section $d\sigma/dx_\gamma^{\rm OBS}$ was calculated for $Q^2 < 1$ GeV^2, $130 < W < 280$ GeV, requiring two jets with $E_T^{\rm jet1} > 7$ GeV, $E_T^{\rm jet2} > 6$ GeV, and at least one D^* with $p_T^{D^*} > 3$ GeV in the angular range $-1.5 < \eta^{D^*} < 1.5$.

The $d\sigma/dx_\gamma^{\rm OBS}$ cross section is shown in Fig. 1.15. In the upper part, the cross section is compared to the HERWIG [52] Monte Carlo simulation, normalized to the data. Leading Order direct and resolved components are included in the simulation, in the fraction predicted by the generator. A peak for high $x_\gamma^{\rm OBS}$ values can be seen in the data distribution since the contribution from direct processes dominates, but also a tail at low $x_\gamma^{\rm OBS}$ values is present, and this tail needs the contribution from

the resolved HERWIG component to be described. In the LO–resolved Monte Carlo histogram, the dominant contribution is found to come from charm excitation processes (lightly hatched histogram) rather than from other resolved processes (densely hatched histogram).

In the lower part of Fig. 1.15 the data are compared to the NLO predictions using the massive charm approach [22, 43]. Since this is a massive approach no explicit charm excitation component is included. The tail for low $x_\gamma^{\rm OBS}$ values can not be described by the NLO calculation, even with extreme choices of the input parameters. The extent to which a NLO calculation for charm with dijet may describe the $d\sigma/dx_\gamma^{\rm OBS}$ distribution must await further theoretical developments, in particular massless charm NLO predictions are needed.

Recently the ZEUS experiment has also published the measurement of the inclusive $D_s^\pm$ cross section in the photoproduction regime [53], which represents another test of perturbative QCD calculations for charm production, independent of the $D^{*\pm}$ measurement reported above.

The $D_s^\pm$ mesons were reconstructed through the decay mode $D_s^\pm \to \phi\pi^\pm$, with the ϕ identified via the decay $\phi \to K^+K^-$. The D_s was required to have $3 < p_T^{D_s} < 12$ GeV and to be in the angular range $-1.5 < \eta^{D_s} < 1.5$. The discrimination between signal and background was done using the distribution of $\cos\theta_K^*$, the angle between one of the kaons and the pion in the ϕ rest frame. The $\cos^3\theta_K^*$ distribution for signal events is expected to be flat, whereas the background distribution is peaked at zero. A partial K and π separation was done using the dE/dx information from the Central Tracking Detector (see Sect. 2.2.1).

The differential cross sections for D_s photoproduction, $d\sigma/dp_T^{D_s}$ and $d\sigma/d\eta^{D_s}$, were calculated and compared to two different NLO QCD predictions; one is the same as for the D^* [22, 43], the second is described in Ref. [54]. In the first approach, a massive calculation based on the program by S. Frixione et al., the parton densities used were MRSG [44] for the proton and GRV–G HO [28] for the photon, and, as usual, the renormalization scale was set to the transverse mass, $m_T = \sqrt{p_T^2 + m_c^2}$ ($m_c = 1.5$ GeV), and the factorization scale to $2m_T$. The Peterson [45] fragmentation function was used to describe charm fragmentation into hadrons.

The differential cross sections compared to NLO predictions are shown in Fig. 1.16. In the Figures, NLO predictions are shown for two sets of the input parameters: the thick curves correspond to the standard choice described above, whereas the thin lines correspond to the values obtained with $m_c = 1.2$ GeV and $\mu_r = m_T/2$. The theory reproduces within the errors the shape of the $p_T^{D_s}$ distribution, but underestimates the data for the nominal parameter set. Regarding the η^{D_s} distribution, the NLO predictions are below the data in the central and forward regions.

The approach to the data of the second model, proposed by Berezhnoy, Kiselev and Likhodel, (BKL) [54] is different since no specific fragmentation function is used. In this calculation, the $(\bar{c}, q)$ state produced in perturbative QCD is hadronized, taking into account singlet and octet contributions. The ratio of the colour–singlet and colour–octet components was tuned to describe the ZEUS data on D^* photoproduction, and used to describe D_s photoproduction. The agreement with the data is better than that of the NLO calculations of Refs. [22, 43], but the simulation of the shape of the η^{D_s} distribution is not very good [53].

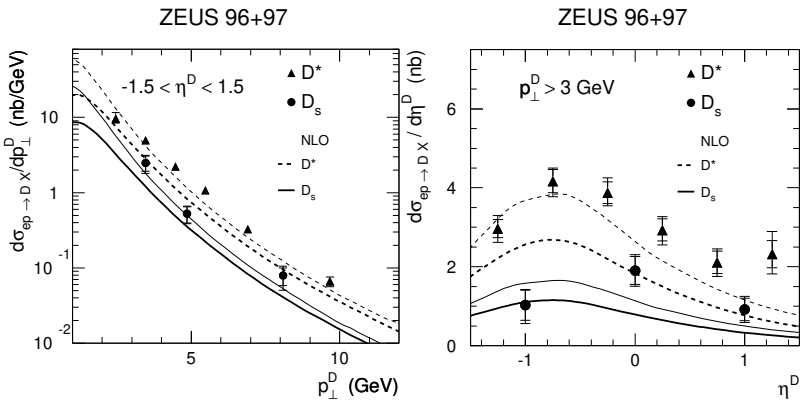

Figure 1.16: *Differential cross sections $d\sigma/dp_T^D$ and $d\sigma/d\eta^D$ where D stands for D_s (dots) or D^* (triangles). The data are compared to NLO predictions for D_s (full curves) and D^* (dashed curves) with two parameter settings: $m_c = 1.5$ GeV, $\mu_r = m_T$ (thick curves) and $m_c = 1.2$ GeV, $\mu_r = m_T/2$ (thin curves) (from Ref. [53]).*

1.4.2.2 Beauty photoproduction at HERA

The study of beauty production at HERA has been carried on by both the experiments ZEUS and H1, which have published results on beauty photoproduction and presented preliminary results on b production in the deep inelastic scattering regime.

The H1 Collaboration has studied the production of b quarks in the photoproduction regime [55] using a sample of 6.6 pb^{-1} collected during 1996 data taking. The production of beauty was tagged looking for muons coming from the semi–leptonic decay of the b quark. Events were selected by requiring the presence of at least two jets having $E_T > 6$ GeV, identified using a cone algorithm [56] with radius 1. The muon, included in one of the jets, had to be observed in the central region of the detector ($35^o < \theta^\mu < 130^o$) and to have a transverse momentum $p_T^\mu > 2$ GeV. DIS production was rejected by suppressing events in which an electron of more than 4 GeV was found in the final state. The inelasticity, y, calculated using the Jacquet–Blondel [57] method, was asked to be in the interval $0.1 < y < 0.8$. These conditions limit the data sample to the PHP regime, $Q^2 < 1$ GeV^2.

Beauty events were separated from charm and other background on a statistical basis, using the transverse momentum of the muon relative to the thrust axis of the jet which contains it, $p_T^{\rm rel}$. The thrust axis was defined as the direction such that

$$T = \max \frac{\sum |p_i^L|}{\sum |p_i|} \tag{1.33}$$

where p_i^L are the longitudinal momenta (with respect to the thrust axis) and p_i the momenta of all the charged and neutral particles in the jet, except for the muon. Since the mass of the b quark is much larger than that of the charm and lighter quarks,

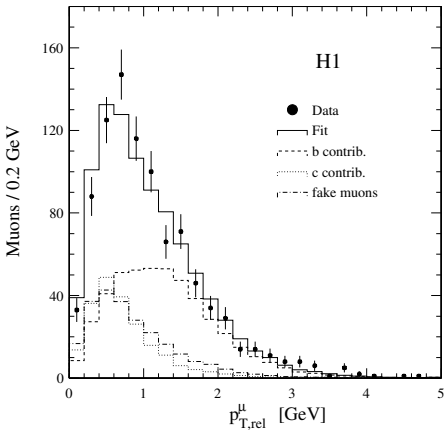

Figure 1.17: *The measured $p_T^{\rm rel}$ distribution in the data and the fitted sum (solid line) of the contributions of beauty (dashed line), charm (dotted line) and the fixed fake muon background (dashed–dotted line).*

the $p_T^{\rm rel}$ distribution of muons coming from beauty decay is expected to be harder than for lighter quark decays. The $p_T^{\rm rel}$ spectrum was used for a combined fit of the different contributions to the selected data sample, in order to extract the fraction of beauty events in the data. The shapes of the beauty and charm $p_T^{\rm rel}$ distributions were taken from the AROMA [58] Monte Carlo event generator, based on LO QCD matrix elements and implementing only direct photon processes. The background from light quark events was modelled by studying the probability of misidentification of π, K and p as muons in the H1 detector. To this extent, large samples of single pions, kaons and protons were passed to the full detector simulation, and the fraction measured as muons determined. The probabilities obtained in this way were then verified on the data, using K_S^0 and ϕ decays as a source of pions and kaons. The agreement between the predicted number of fake muons and those really observed in the data was found to be very good, and also the fake muon distributions in p_T^μ and θ^μ well agreed between data and Monte Carlo.

In Fig. 1.17 the observed $p_T^{\rm rel}$ distribution for the data is shown, together with the contributions from beauty, charm and the background from light quarks obtained as explained above. The relative composition of the data sample amounts to $f_b = 50.8 \pm 4.9\%$ (beauty), $f_c = 22.4 \pm 5.0\%$ (charm) and $f_{fake} = 25.9\%$.

From the fraction of b quarks in the data the visible beauty electroproduction cross section was measured to be:

$$\sigma_{vis}(ep \to b\bar{b}X \to \mu X') = 0.176 \pm 0.016\ (stat.)\ ^{+0.026}_{-0.017}\ (syst.)\ \text{nb} \qquad (1.34)$$

in the kinematic region defined by $Q^2 < 1$ GeV^2, $0.1 < y < 0.8$, $p_T^\mu > 2$ GeV and $35^o < \theta^\mu < 130^o$.

To estimate the theoretically expected cross section, the NLO QCD calculations by S. Frixione et al. [22] were used. The b–quark mass was set to 4.75 GeV, and the parton densities used for the proton and the photon were MRSG [44] and GRV–HO [28], respectively. The renormalization and factorization scales were set to the transverse mass, $m_T = \sqrt{m_b^2 + p_T^2}$. The b quark fragmentation was done using the Peterson [45] parametrisation. The value found for the theoretical prediction was 0.054 ± 0.09 nb, more than a factor three lower than the measured value.

The ZEUS Collaboration has published the results of an analysis on beauty photoproduction with semi–leptonic decay into electrons [59]. The data sample used corresponds to $\sim$ 38 pb^{-1} collected during 1996–97 data–taking. Events were selected by requiring the presence of at least two jets in the final state, reconstructed by the KTCLUS [60] algorithm, having transverse energy $E_T^{\rm jet1(2)} > 7(6)$ GeV, in the angular region defined by $|\eta^{\rm jet1,2}| < 2.4$. DIS background was rejected by removing events with a well reconstructed positron in the CAL and by cutting on the inelasticity, y_{JB}, measured using the Jacquet–Blondel method [57]. The inelasticity was asked to be in the range $0.2 < y_{JB} < 0.8$. These cuts limit the data sample to $Q^2 < 1$ GeV^2. Electrons were identified in the Central Tracking Detector (see Sect. 2.2.1) from the dE/dx of the charged tracks, using a method of statistical subtraction. The identified electrons were asked to have $p_T^{e^-} > 1.6$ GeV and to be in the angular region defined by $|\eta^{e^-}| < 1.1$.

The differential cross sections $d\sigma/dp_T^{\rm rel}$ and $d\sigma/dx_\gamma^{\rm OBS}$ for events with a semi-leptonic decay of a heavy quark into an electron were calculated in the kinematic region defined by $0.2 < y < 0.8$, $Q^2 < 1$ GeV^2, $E_T^{\rm jet1(2)} > 7(6)$ GeV, $|\eta^{\rm jet1,2}| < 2.4$, $p_T^{e^-} > 1.6$ GeV and $|\eta^{e^-}| < 1.1$. The obtained distributions are shown in Fig. 1.18. The data are compared to the expectations of the HERWIG [52] Monte Carlo simulation, which has been area normalized to the data for a comparison of shape; the scaling factor was 3.8.

In the $d\sigma/dp_T^{\rm rel}$ differential cross section, the peak at low values is consistent with predominant semi–leptonic decays of charm quark, whereas the tail at high $p_T^{\rm rel}$ values is consistent with a significant contribution from b decays. The contribution from light flavours, essentially coming from τ and η decays, was estimated to be less than 3%.

In the $d\sigma/dx_\gamma^{\rm OBS}$ distribution a peak at high values is observed, consistent with the dominance of direct photon processes, but the tail at low–$x_\gamma^{\rm OBS}$ values cannot be explained by direct processes alone. HERWIG predicts a contribution of 35% from resolved processes to the total cross section, giving a distribution in good agreement with the data. From the fit of the shape of the direct and resolved components to the data distribution, the resolved photon component is estimated to give a $28 \pm 5\ (stat.)\%$ contribution to the total cross section. Therefore, LO Monte Carlo models require resolved photon processes to describe heavy quark production, as was already observed for charm production in [41].

The $p_T^{\rm rel}$ distribution of the data was fitted to the sum of the contributions from beauty and charm, including the cascade decay $b \to c \to e^-$ in the background expected for charm. The fitted beauty percentage, $f_b = 14.7 \pm 3.8\ (stat.)$, was used to determine the cross section for beauty production in the kinematic region defined above. The result was:

$$\sigma^{b\to e^-}_{e^+p\to e^+dijet\ e^-X} = 24.9 \pm 6.4^{+4.2}_{-7.3}\ \text{pb}. \qquad (1.35)$$

This cross section was extrapolated to the parton level in a restricted range of the transverse momentum and pseudorapidity of the b quark, defined as $p_T^b > 5$ GeV, $|\eta^b| < 2$. The kinematic region was also defined by $Q^2 < 1$ GeV^2 and $0.2 < y < 0.8$.

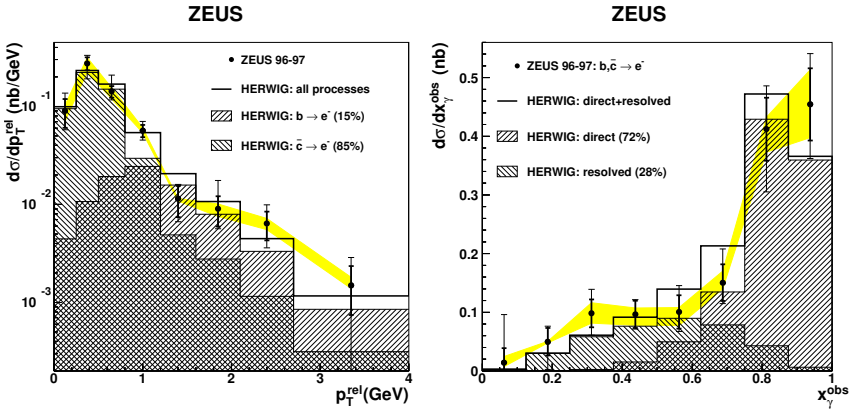

Figure 1.18: **Left**: *differential cross section $d\sigma/dp_T^{\rm rel}$ for heavy flavour production. The shaded band represents the effect of the uncertainty on the calorimeter energy scale. The experimental data are compared to the* Herwig *Monte Carlo prediction (solid line) for all quark flavours, fitted to the data and scaled up by a factor 3.8. The contributions from beauty (forward–diagonally hatched histogram) and from charm processes (backward–diagonally hatched histogram) predicted by the Monte Carlo model are shown separately.* **Right**: *differential cross section $d\sigma/dx_\gamma^{\rm OBS}$ for heavy quark decays. The direct (forward–diagonally hatched histogram) and resolved photon components (backward–diagonally hatched histogram) are shown separately. From Ref. [59].*

The extrapolation factor evaluated by the Herwig Monte Carlo was 6.8, with a large uncertainty due to the hadronization model, quantified in -26% by using the Pythia [61] Monte Carlo to extrapolate (Pythia uses the Lund model while Herwig uses the cluster model for the hadronization). To correct to the full beauty cross section the branching ratio for the process $b \to e^-$ $(10.73 \pm 0.35)\%$ [62] was used. The extrapolated cross section was found to be:

$$\sigma^{\rm ext}_{ep\to e^+bX} = 1.6 \pm 0.4(stat.)^{+0.3}_{-0.5}(syst.)^{+0.2}_{-0.4}(ext.)\ \text{nb}. \qquad (1.36)$$

The cross section extrapolated in this way was compared to the NLO QCD predictions by Frixione et al. [22, 43]. The parton densities used for the proton and the photon were MRST99 [63] and GRV–G HO [28], respectively; the renormalization and factorization scales were set to the transverse mass, $\mu_R = \mu_F = m_T = \sqrt{m_b^2 + p_T^2}$, using for the mass of the b quark the value $m_b = 4.75$ GeV. The predicted beauty cross section with the above settings was 0.64 nb. To estimate the uncertainty on the theoretical calculation the b–quark mass and the scales were changed simultaneously in order to obtain the maximum upper and lower variation of the predicted cross section. The obtained results are shown in Fig. 1.19: the extrapolated cross section lies somewhat above the central NLO prediction.

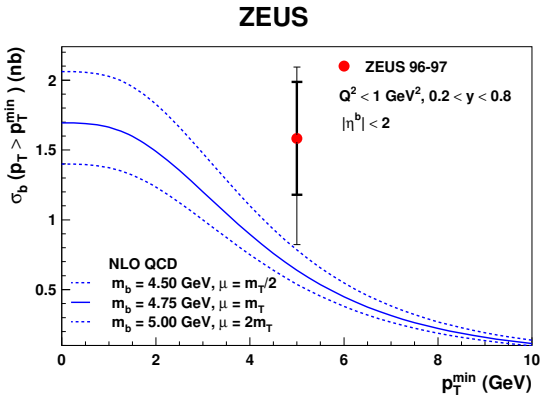

Figure 1.19: *The extrapolated beauty cross section at a fixed $p_T^{\rm min}$ value compared with theoretical predictions plotted as a function of $p_T^{\rm min}$. The curves show the predictions from NLO QCD for varying the b–quark mass and the factorization and renormalization scales (from Ref. [59]).*

The H1 Collaboration has presented preliminary results [64] on beauty production in the PHP regime, measured using information on the particle lifetime, obtained with the H1 vertex detector. The analysed sample corresponds to 14.7 pb^{-1} of data collected during the 1997 data taking, when the Central Silicon Tracker (CST) [65] of the H1 detector was fully commissioned. The long lifetime of the b hadrons, that can be measured by the vertex detector, provides a signature of the production of b quarks which is independent of that obtained with the $p_T^{\rm rel}$ distribution.

Photoproduction events were selected by requiring no electromagnetic cluster with an energy above 8 GeV in the backward (lepton) direction. In this way the data sample was limited to the kinematic region $Q^2 < 1$ GeV^2 (photoproduction regime). A cut on the inelasticity, y, calculated with the Jacquet–Blondel method, was applied, $0.1 < y < 0.8$, in order to further reduce the background from DIS events. Jets were reconstructed by the k_T algorithm [60] and asked to have transverse energy $E_T > 5$ GeV. The muon, which had to be identified in the barrel part of the detector ($35^o < \theta^\mu < 130^o$) and to have a transverse momentum $p_T^\mu > 2$ GeV, had to be included in one of the jets. At least two hits in the CST had to be found, associated with the muon candidate track.

The Aroma [58] Monte Carlo program was used to generate the signal events ($ep \to b\bar{b}X$) and the background from charm events ($ep \to c\bar{c}X$). The background from fake muons was determined using tracks from a tagged photoproduction event sample which fulfilled the same selection criteria as the signal sample, except that no muon identification was required.

The impact parameter δ distribution in the data, together with the contributions from beauty and background events, is shown in Fig. 1.20 (left). The decomposition is obtained from a likelihood fit to the δ spectrum [66], using for b and c components

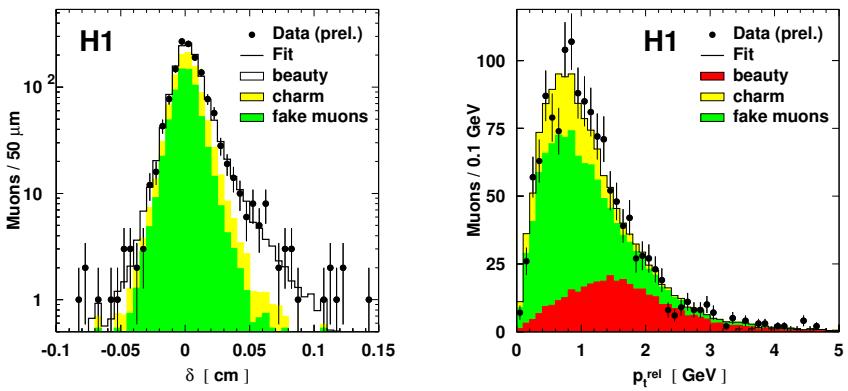

Figure 1.20: **Left:** *Impact parameter distribution in the data (dots), compared to the sum of beauty, charm and fake muons contributions, mixed according to the fractions given by the likelihood fit.* **Right:** $p_T^{\rm rel}$ *distribution in the data (dots), compared to the sum of beauty, charm and fake muons contributions, in the fractions obtained by fitting the beauty and charm components, keeping the fake muons fraction fixed. From Ref. [64].*

the shapes given by the Monte Carlo, while for the fake muons the distribution from real data was used. From the fit, the contributions of the three different components were found to be $f_b = 26 \pm 5\%$ (beauty), $f_c = 24 \pm 12$ (charm) and $f_{fake} = 50 \pm 10\%$ (fake muons).

The $p_T^{\rm rel}$ distribution of the data was used to obtain an independent determination of the fraction of b quarks. The data distribution is shown in Fig. 1.20 (right), compared to the sum of b, c and fake muons contributions. In this case, since the $p_T^{\rm rel}$ distributions of charm and light quarks cannot be well distinguished, the fraction of fake muon events was kept fixed at $f_{fake} = 56\%$, a value obtained using the light hadrons misidentification probabilities, as in [55]. From the likelihood fit, the fraction of b quark events was found to be $f_b = (27 \pm 3)\%$, in good agreement with the result of the δ fit.

A fit of the two–dimensional $(\delta, p_T^{\rm rel})$ distribution of the data was performed and gave for the beauty cross section the value:

$$\sigma_{vis}(ep \to b\bar{b}X \to \mu X) = 160 \pm 16\ (stat.) \pm 29\ (syst.)\ \text{pb}, \qquad (1.37)$$

in good agreement with the published one (Eq. 1.34). The two values were combined to give

$$\sigma_{vis}(ep \to b\bar{b}X \to \mu X) = 170 \pm 25\ \text{pb}. \qquad (1.38)$$

The NLO QCD prediction by Frixione et al. [22] gives a value of (54 ± 9) pb, as seen above [55], thus still underestimating the measured value.

It is interesting to see if the discrepancy between data and theoretical predictions is observed also in the deep inelastic scattering regime, since in this case the production mechanism of heavy quarks is quite different than in the photoproduction case (see Sect. 1.3).

The H1 Collaboration has presented preliminary results [67] on beauty production in the high–Q^2 region. The analysed sample corresponds to 10.5 pb^{-1} of data collected during 1997. The analysis is very similar to that of the PHP regime, reported in [64], but in this case the scattered positron had to be observed in the detector, having a polar angle $\theta^e < 177^o$ and an energy $E^e > 9$ 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 and $35^o < \theta^\mu < 130^o$. The beauty fraction in the data was determined by a fit to the two dimensional $(\delta, p_T^{\rm rel})$ distribution, as for the photoproduction analysis. The resulting beauty fraction in the data was $f_b = 43 \pm 8\%$. From this value the beauty production cross section in the DIS regime could be extracted:

$$\sigma_{vis}(ep \to bX \to \mu X) = 39 \pm 8\ (stat.) \pm 10\ (syst.)\ \text{pb}. \qquad (1.39)$$

The NLO QCD prediction was obtained using the HVQDIS program [68]. The Peterson function was used to describe the b fragmentation. The proton parton densities chosen were GRV98 [69]; the renormalization and factorization scales were $\mu_R = \mu_F = \mu = \sqrt{Q^2 + 4m_b^2}$, with the b–quark mass set to $m_b = 4.75$ GeV. The value found was (11 ± 2) pb, well below the experimental result.

The ZEUS experiment also has presented preliminary results [70] on beauty production in the DIS regime. The analysed sample corresponds to 60 pb^{-1} of data collected during 1999–2000 running period. DIS events were selected by requiring a well reconstructed positron with $E^e > 10$ GeV, $Q^2 > 2$ GeV^2 and $0.05 < y < 0.7$. At least one jet reconstructed in the Breit frame was required, having transverse energy $E_T^{\rm jet} > 6$ GeV, in the pseudorapidity range $-2 < \eta^{\rm lab} < 2.5$, where $\eta^{\rm lab}$ is the pseudorapidity measured in the laboratory frame. The muon, reconstructed in the region $30^o < \theta^\mu < 160^o$, and with $p_T^\mu > 2$ GeV, had to be included in one jet having $E_T > 4$ GeV in the Breit frame. The Monte Carlo used to simulate signal and background events was RAPGAP [71]. The fraction of beauty in the data was determined by a fit to the $p_T^{\rm rel}$ variable, mixing the charm and light flavour components according to their relative luminosities, since the shape of their $p_T^{\rm rel}$ distribution was very similar and cannot be distinguished by the fit. The fraction of beauty events in the data was $f_b = 25 \pm 5\%$ $(stat.)$, and the corresponding beauty cross section was:

$$\sigma_{b\bar{b}}(e^+p \to e^+ b\bar{b}X \to e^+\text{jet}\ \mu^\pm X) = 38.7 \pm 7.7\ (stat.)^{+6.1}_{-5.0}\ (syst.)\ \text{pb}. \qquad (1.40)$$

The NLO QCD predictions were determined using the HVQDIS [68] program. The Peterson function was used to describe the b fragmentation into hadrons, and the semi–leptonic decay of b hadrons was modelled using a parametrisation of the muon momentum spectrum extracted from the Monte Carlo. The proton parton densities used were CTEQ5F4 [72], the b–quark mass was set to $m_b = 4.75$ GeV, and the renormalization and factorization scales to $\mu = \sqrt{Q^2 + 4m_b^2}$. The resulting value was $28^{+5.3}_{-3.5}$ pb, which is somewhat below but agrees within the experimental and theoretical uncertainties with the measured value.

1.4.3 Beauty production at the Tevatron

The analyses on beauty production in $p\bar{p}$ collisions at the Tevatron [73] have shown that the b quark production cross section is higher than the theoretical predictions (obtained with the standard choice of the input parameters) by a factor of 2 to 3. One of the most recent publications by the CDF Collaboration [74] is of particular interest, since it has been investigated by M. Cacciari and P. Nason in order to make clearer some aspects of the NLO calculations.

In Ref. [74] the B^+ total cross section and the differential cross section $d\sigma/dp_T^{B^+}$ were measured, using fully reconstructed $B^\pm$ mesons decaying into the exclusive final state $J/\psi\ K^\pm$. The measurement used a data sample of 98 pb^{-1} collected from 1992 to 1995 by the CDF experiment from $p\bar{p}$ collisions at a centre–of–mass energy of 1.8 TeV. The whole data sample is divided into two sub–samples, treated in different ways in the analysis: the run 1A sample consists of 19.3 pb^{-1} of data used for a previous analysis on B^+ production [75], while the rest of the sample is referred to as run 1B.

The J/ψ coming from the B^+ decay was reconstructed by measuring both muons coming from its decay with the silicon vertex detector. For run 1A data, the transverse momentum of each muon was asked to be $p_T^\mu > 1.8$ GeV, with one of the muons having $p_T^\mu > 2.8$ GeV, while for run 1B data both muons were asked to have $p_T^\mu > 2$ GeV. The two muons had to have opposite sign, and to be less than 5.0 cm far to each other at the point of closest approach to the vertex. They were then constrained to come from a common point in space. The width of the reconstructed J/ψ peak was 16 MeV, the signal region being defined as that with dimuon candidates with reconstructed mass within 3.3 σ of the known J/ψ mass.

To reconstruct the B candidate, each charged particle in an event was combined with the two muons if the z coordinate of the track at the point of closest approach to the beam line was within 5 cm of the z position of the J/ψ candidate decay vertex. Since kaons from B decays have much harder transverse momentum spectrum than K coming from background events, a cut was imposed on p_T^K, requiring $p_T^K > 1.25$ GeV. The p_T of each B candidate was required to be $p_T > 6.0$ GeV, and the proper decay length to be greater than 100 μm, to suppress background from prompt J/ψ mesons.

The differential cross section $d\sigma/dp_T^{B^+}$ was determined by dividing the full B candidate sample into four p_T ranges, and in each range the invariant mass distribution was fitted using an unbinned maximum likelihood fit, to determine the number of B candidates. The value of the cross section in each $p_T^{B^+}$ bin was then given by:

$$\frac{d\sigma(B^+)}{dp_T^{B^+}} = \frac{N_{\rm sig}/2}{\Delta p_T^B \mathcal{L} A \mathcal{B}} \tag{1.41}$$

where $N_{\rm sig}$ is the number of B in each p_T^B bin as given by the fit, Δp_T^B is the width of the bin, A is the acceptance, which also includes trigger and reconstruction efficiency, $\mathcal{B}$ is the product of the branching ratios $BR(B^\pm \to J/\psi\ K^\pm) \cdot BR(J/\psi \to \mu^+\mu^-)$. The factor 1/2 keeps into account the fact that both B^+ and B^- mesons are detected, while the cross section is reported for B^+ mesons, assuming charge invariance in the production process. The differential cross section $d\sigma/dp_T^{B^+}$ is shown in Fig. 1.21, compared to NLO QCD predictions. In the theoretical calculation the proton parton densities used were MRST [63] and the renormalization and factorization scales were set to $\mu_R = \mu_F = \mu = m_T = \sqrt{m_b^2 + p_T^2}$. The fragmentation of b quarks into hadrons

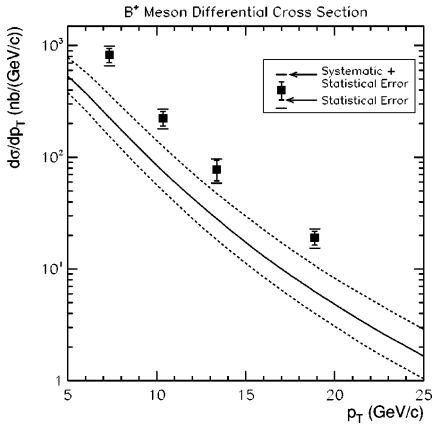

Figure 1.21: *Differential cross section $d\sigma/dp_T^{B^+}$ compared to NLO QCD prediction. The solid curve represents the theoretical prediction for $m_b = 4.75$ GeV, $\mu_R = \mu_F = \mu = \sqrt{m_b^2 + p_T^2}$, while the dashed curves are obtained by setting $m_b = 4.5$ GeV, $\mu_R = \mu_F = \mu/2$, and $m_b = 5.0$ GeV, $\mu_R = \mu_F = 2\mu$ (from Ref. [74]).*

was modelled using the Peterson fragmentation function, with $\epsilon = 0.006$, as obtained from a fit to data collected at e^+e^- colliders [76]. The uncertainties on the NLO predictions were estimated by varying the mass of the b–quark from 4.5 to 5.0 GeV, and the renormalization and factorization scales from $m_T/2$ to $2m_T$. The ratio between the measured values and the NLO QCD predictions was calculated in the four p_T bins, and then fitted by a line. The value found by the fit, corresponding to the mean data to theory ratio, was 2.9.

Recently, the results obtained by the CDF Collaboration in Ref. [74] have been analysed by M. Cacciari and P. Nason [77] to see how the big discrepancy between the data and the NLO predictions could be explained. They focused their attention on the implementation of the hadronization effects in the NLO QCD calculations, which is a non–perturbative issue.

The non–perturbative hadron–formation effect is usually introduced by writing the hadron–level cross section for B mesons as:

$$\frac{d\sigma^B}{dp_T} = \int d\hat{p}_T dz \frac{d\sigma^b}{d\hat{p}_T} D(z) \delta(p_T - z\hat{p}_T) \tag{1.42}$$

where the function $D(z)$ is a phenomenological parametrization of hadronization effects. Usually, the Peterson [45] form of the $D(z;\epsilon)$ function is used, and the quark cross section is given by a shower Monte Carlo program; the ϵ parameter is taken from e^+e^- data. The effect of the fragmentation is to reduce the momentum of the B meson with respect to that of the b quark of roughly 10%. Nevertheless, the value of the cross

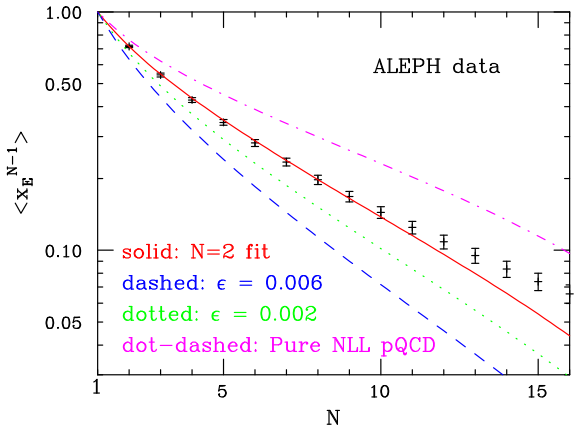

Figure 1.22: *Moments of the measured B meson fragmentation function compared with the perturbative NLL calculation supplemented with different non–perturbative fragmentation forms $D(z)$. The solid line is obtained using a one–parameter form fitted to the second moment.*

section is influenced by this effect since the transverse momentum spectrum of the b quark is steeply falling.

In Ref. [77] the NLO QCD predictions for the $d\sigma/dp_T^{B^+}$ were recalculated by correctly implementing the effect of heavy quark fragmentation. To this aim, a complete resummation of large transverse momentum logarithms at the next–to–leading level (NLL) [78] was needed, and a formalism for merging the NLL resummed results to the fixed order (FO) NLO calculation had to be used (FONLL) [79].

In Ref. [80] the fact is pointed out that the value $\epsilon = 0.006$ is appropriate only when a leading–log (LL) calculation of the spectrum is used, while when NNL calculations are used, smaller values of ϵ are needed to fit the data. Another fact has to be noticed, that, rather than the knowledge of the whole spectrum of $D(z)$ for $z \in [0,1]$, the knowledge of some specific moment of the fragmentation function

$$D_N \equiv \int D(z) z^N \frac{dz}{z} \tag{1.43}$$

is sufficient to calculate the hadronic cross section, assuming the power law behaviour of the differential cross section as a function of the transverse momentum of the B meson, $d\sigma/dp_T \propto 1/p_T^n$. It is, in fact:

$$\frac{d\sigma}{dp_T} = \int dz d\hat{p}_T \; D(z) \frac{A}{\hat{p}_T^N} \delta(p_T - z\hat{p}_T) = \frac{A}{p_T^N} D_N. \tag{1.44}$$

Therefore, the hadronic cross section is the product of the partonic cross section and the n^{th} moment of the fragmentation function. The value of n for the p_T spectrum in the region of interest ranges from 3 to 5. It is thus clear that having a good determination

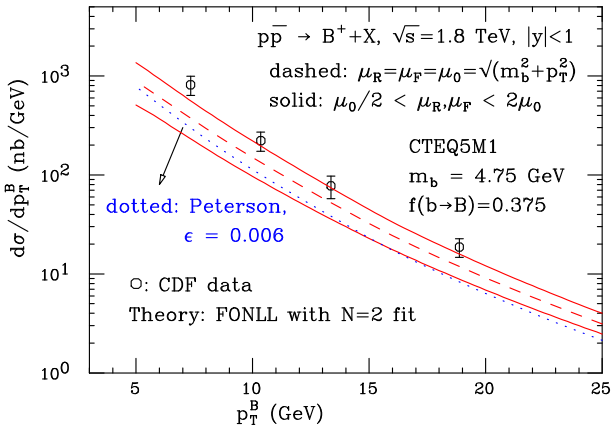

Figure 1.23: *Prediction for the B cross section obtained using the calculation in [79], supplemented with the $N = 2$ fit of the non–perturbative fragmentation function, compared to the CDF data in Ref. [74].*

of the moments of the fragmentation function, when fitting to e^+e^- data, is more important than describing the whole $D(z)$ spectrum.

In Fig. 1.22 the moments of the fragmentation function are shown, calculated from the distribution of the B meson energy fraction with respect to the beam energy, x_E, published by the ALEPH collaboration [81]. The dot–dashed line is the pure perturbative part, while the dotted and the dashed lines represent the convolution of the perturbative part with a Peterson function with $\epsilon = 0.002$ and $\epsilon = 0.006$, respectively. It is clear from the Figure that the parametrization with $\epsilon = 0.006$ does not describe the data.

In order to give a good description of the low moments, a one–parameter form of the non–perturbative fragmentation function was used, its free parameter being fixed by fitting the $N = 2$ point in moments space. The results are shown by the solid line in Fig. 1.22. The description of the data is good up to $N \simeq 10$. Therefore, this "$N = 2$ fit" is a good candidate for the calculations of the hadronic cross sections according to Eq. 1.44.

The theoretical prediction of the differential cross section $d\sigma/dp_T^{B^+}$ was re–calculated using this form of the fragmentation function, and then compared to the results obtained by the CDF Collaboration (Fig. 1.23). In the Figure, also the NLO QCD prediction obtained using the Peterson fragmentation with $\epsilon = 0.006$ is shown, and it can be clearly seen that this choice underestimates the B cross section at large p_T values. The ratio between the measured values and the theoretical predictions has in this case a mean value of 1.7, to be compared to the 2.9 obtained with the previous choice of the fragmentation function.

It is also interesting to see a comparison of the different theoretical predictions as a function of the p_T of the B meson. This is reported in Fig. 1.24, where the

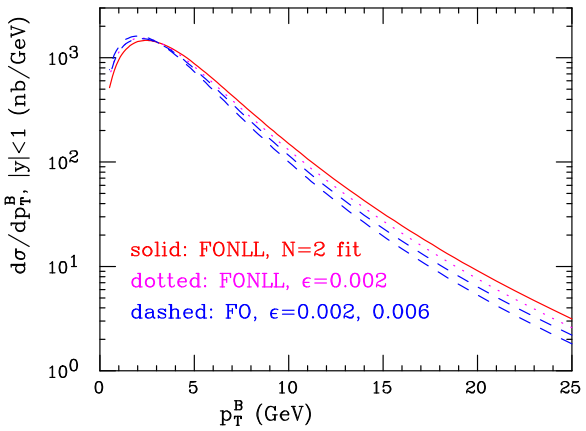

Figure 1.24: *Theoretical prediction of the differential cross sections $d\sigma/dp_T^B$, calculated with different schemes and forms of the fragmentation function.*

results obtained using a fixed–order NLO QCD calculation are shown, together with the FONLL calculation with $\epsilon = 0.002$ or with the $N = 2$ fit. It is clear from the Figure that the difference is more evident for high–p_T values, whereas it is much smaller at low p_T.

Chapter 2

The ZEUS detector at HERA

In this Chapter the HERA accelerator and the ZEUS detector are described briefly, giving particular emphasis on the parts of the detector used in the analysis described in this thesis. A detailed description of the ZEUS detector can be found in [82].

2.1 The HERA collider

HERA (Hadron Electron Ring Anlage) [83] is the first and so far the only lepton–proton collider in the world. It is located at the DESY (Deutsches Electronen SYnchrotron) laboratory, in Hamburg, northern Germany, and operates since autumn 1991. 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 centre–of–mass energy is 300(318) GeV, more than an order of magnitude higher than the previous fixed–target experiments. As a consequence a new and wider kinematic region is accessible at HERA.

The HERA tunnel is 6.3 km long and it is located 15–30 m under the ground level. Electrons (positrons) and protons are accelerated in two different rings. The magnetic system 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.

Four experiments are located in four experimental halls along the HERA ring. *ep* 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 for the study of $\mathcal{CP}$ violation in the $B - \bar{B}$ system.

Fig. 2.1 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 then stripped off the H^- ions to obtain 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.

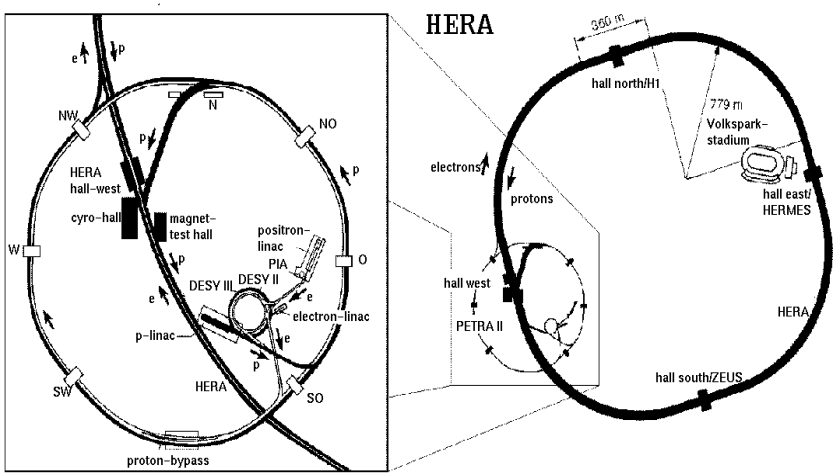

Figure 2.1: *The HERA storage ring with its pre–accelerators system.*

The electron (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 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 the bunches are empty the non–beam–related background can be estimated, as the rates of cosmic rays.

The main design parameters of HERA are reported in Table 2.1.

In Fig. 2.2 and in Table 2.2 the performance of HERA in the last years is shown. 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. It was also in 1998 that the energy of the proton beam was raised from 820 to 920 GeV.

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**
Centre–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 HERA [83].*

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: *Overview of the luminosity delivered by HERA from 1992 to 2000.*

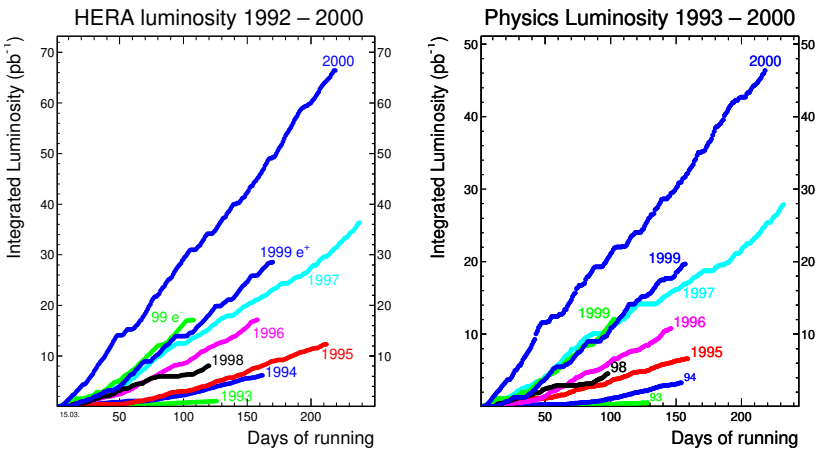

Figure 2.2: *Integrated luminosity delivered by HERA* **(left)** *and usable for ZEUS physics analyses* **(right)** *in the 1993–2000 running period.*

Although a lot of interesting measurements have already been performed at HERA, the desire was expressed by the experiments for an increase in the luminosity. The motivations for this increase were studied in a one–year workshop held between 1995 and 1996, when it was concluded that having 1 fb^{-1} of integrated luminosity would have opened the possibility for new interesting measurements. The luminosity upgrade [84] of the machine, done between the end of 2000 and 2001 shutdown period, should bring a significant increase in the luminosity (around a factor 5).

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 keeping into account the needs of the physics and the technical understanding gained during the running period.

The original form of the detector was driven by the different processes that can be observed at HERA. The detector can measure energies from few MeV to hundreds of GeV in the forward region. For low–momentum particles the tracking on a magnetic field is very precise (the resolution behaves as $\sigma(p_T)/p_T \sim p_T$), while high energy particles are well measured by the calorimetric system (resolution goes as $\sigma(E)/E \sim \sqrt{E}/E$).

Particle identification is needed in a wide momentum range: in Neutral Current (NC) DIS events the scattered lepton has to be identified and measured with high

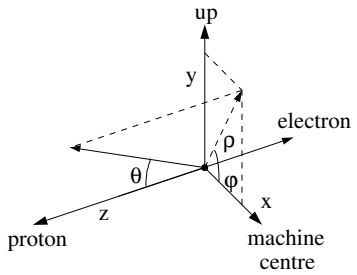

Figure 2.3: *The ZEUS coordinate system.*

precision; the identification of electrons, positrons and muons is needed in order to study the semi–leptonic decay of heavy quarks and exotic processes involving leptons.

In Charged Current (CC) DIS processes a hermetic detector is needed in order to reconstruct the missing transverse momentum carried by the outgoing neutrino. In this kind of events and also in untagged photoproduction events the precise reconstruction of the final state energy is important in order to determine the event kinematic.

The ZEUS coordinate system (Fig. 2.3) 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 centre 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 sub–detectors that build up ZEUS are coaxial but asymmetric with respect to the IP. The detector layout is shown in Fig. 2.4 (longitudinal view) and 2.5 (transverse view). The main detector is approximately 20 m long, 12 m large and 11 m high, and it weights around 3600 tonnes.

A brief outline of the various detector components is given below; a more detailed description of the sub–detectors of particular interest for the analysis reported here will be given later in this Chapter.

The innermost detector that can be seen in Figs. 2.4 and 2.5 is the Vertex Detector (VXD), that was removed during the 1995–96 shutdown. Therefore during 1996–2000 data–taking the closest detector 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 (FCAL, BCAL, 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 (inner muon chambers, FMUI, BMUI and RMUI) and outside (FMUO, BMUO, RMUO) the yoke.

The Small angle Rear Tracking Detector (SRTD) is a scintillator strip detector

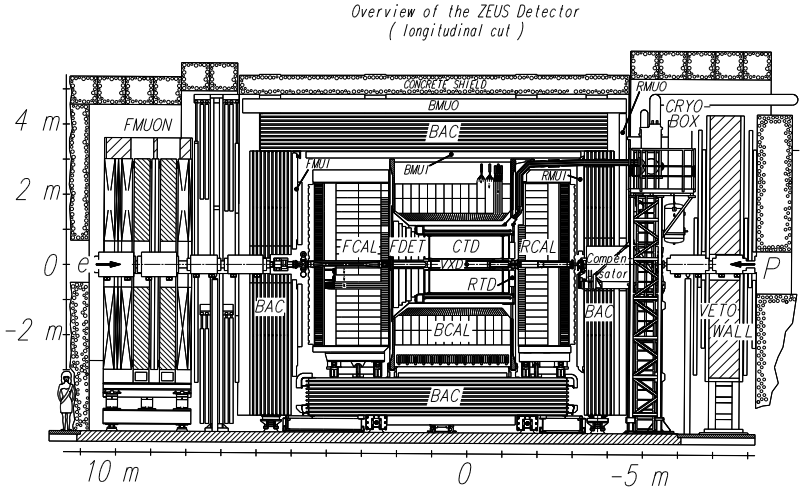

Figure 2.4: *Cross section of the ZEUS detector along the beam axis.*

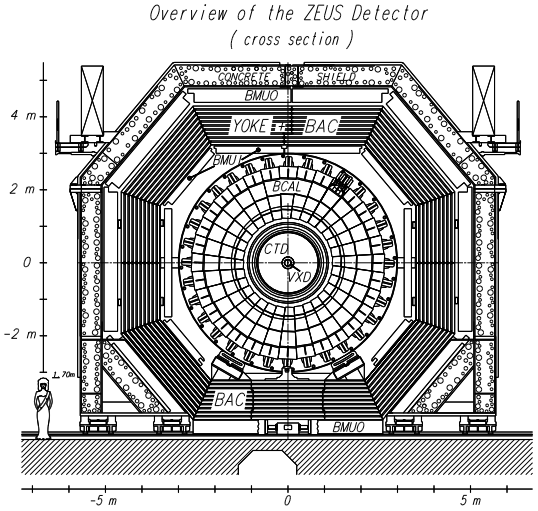

Figure 2.5: *Cross section of the ZEUS detector orthogonal to the beam axis.*

Figure 2.6: *$x-y$ cross section of one octant of the CTD. The sense wires are indicated with dots.*

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. The C5 counter, located at the rear end of the calorimeter ($z = -314$ cm), consists of two planes of scintillators, one above and one below the beam pipe. The timing information given by these two sub–detectors are used to reject proton–beam gas events (see Sect. 4.4.1).

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

2.2.1 The Central Tracking Detector (CTD)

The Central Tracking Detector (CTD) [85] 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 $15^o < \theta < 164^o$. 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. The CTD 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 makes up the CTD. The field wires are tilted of 45^o with respect to the radial direction, in order to compensate the Lorentz angle of 45^o due to the electric and magnetic field. One octant of the CTD is shown in Fig. 2.6.

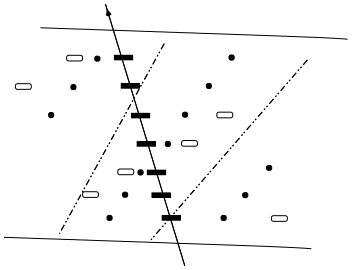

Figure 2.7: *Hits coming from a genuine track (solid rectangles) tend to cross the cell boundaries within a superlayer. The open rectangles are ghost hits.*

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; during the drift an avalanche effect occurs, giving an amplification factor on the electrons of $\sim 10^4$, so that a readable 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 $\pm 5^o$ (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 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.

2.2.1.1 Track reconstruction

The information used to determine the spatial position of a particle, needed for the track reconstruction, come from the time of arrival of the ionization electrons on the sense wire. If the drift velocity is approximately constant and equal to u_d, the relation between the drift time, t_d, and the distance from the sense wire, d_s, is given by:

$$d_s \simeq u_d \cdot t_d \tag{2.2}$$

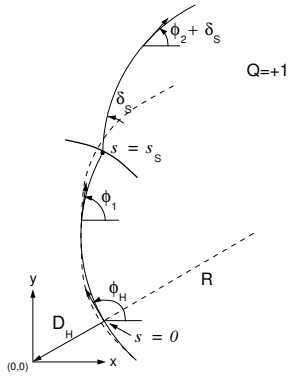

Figure 2.8: *The track helix in the xy plane.*

where t_d is defined as the difference between the time t_f at which the pulse appears on the sense wire, and the time t_i of passage of the charged particle, calibrated for every wire, $t_d = t_f - t_i$. However, this kind of information is not sufficient to determine from which side of the wire the particle comes, therefore a left–right ambiguity is still present. The 45^o tilt of the sense wires can solve this ambiguity, since tracks coming from the interaction point tend to pass the boundary of adjacent cells within a superlayer, as shown in Fig. 2.7.

The pattern recognition of the track begins looking for a *seed*, a group of hits in the outermost superlayer (SL9). To these hits a virtual hit at $x = y = 0$ is added, keeping into account the transverse dimension of the beam as an error on this hit. The two hits are enveloped with a circle arc, and inner hits on the axial superlayers are added on the way, updating the circle parameters and refining the trajectory determination. Once the trajectory spans several axial superlayers, the arc in the XY plane is used for the stereo pattern recognition: stereo hits are selected that match with the arc after being rotated. The pattern recognition begins with the longest tracks, those going from SL9 to SL1, then continues with shorter tracks, reaching inner superlayers (SL7, SL5...), and finally includes tracks with no hits in the innermost superlayer, that can come from the decay of long lived particles.

When the pattern recognition is done, all the candidate tracks are fitted with a helix, starting with the innermost superlayer and adding the outer ones on the way. The five helix parameters are (Fig.2.8):

- $a_1 = \phi_H$, the angle tangent to the helix in the xy plane;
- $a_2 = Q/R$, where Q is the charge and R the radius of the helix;
- $a_3 = QD_H$, where D_H is the distance of the helix from the reference point;
- $a_4 = z_H$, the z position of the point of closest approach;
- $a_5 = \cot\theta$, where θ is the polar angle.

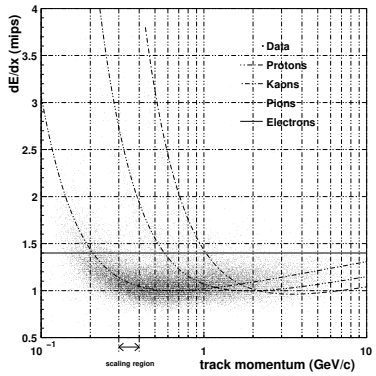

Figure 2.9: *The energy lost by different particles, as reconstructed by the CTD, as a function of the particle momentum. The curves indicate the particle type.*

The event vertex is then reconstructed from the information on the fitted tracks. Tracks too far from the determined vertex are discarded, the surviving tracks are constrained to the vertex and the fit parameters recalculated.

2.2.1.2 dE/dx measurement

The CTD is also used to estimate the ionization energy loss of a particle in the gas, in order to have information on the particle identification.

The ionization energy loss is parametrized by the Bethe–Bloch equation:

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 z^2 \frac{Z}{A}\frac{1}{\beta^2}\left[\ln\left(\frac{2m_e c^2\gamma^2\beta^2}{I}\right) - \beta^2 - \frac{\delta}{2}\right] \tag{2.3}$$

for a particle with charge ze passing through a medium with atomic number Z and mass number A. The energy loss is a function of the velocity, $c\beta$, of the particle: for low β the ionization loss decreases steeply, down to a minimum for $\gamma = (1-\beta^2)^{-0.5} \simeq 3$, and then rises as a logarithm up to a plateau at very high γ. At low momenta, $p = \beta\gamma m$, the energy loss is different for different masses of the particles. The behaviour as observed by the CTD is shown in Fig. 2.9.

2.2.2 The Uranium–scintillator Calorimeter (UCAL)

The ZEUS calorimeter (UCAL) [86] is a high–resolution compensating calorimeter. It completely surrounds the tracking devices and the solenoid, and covers the 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.

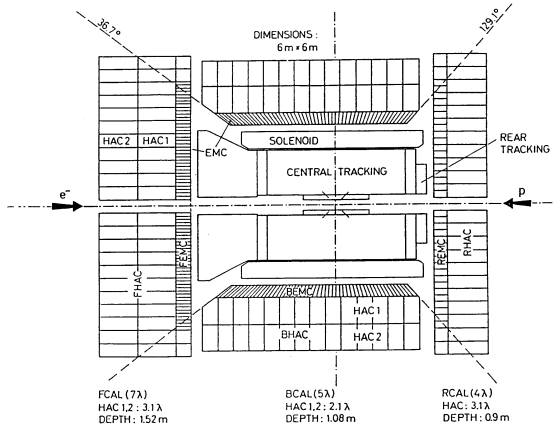

Figure 2.10: *Schematic view of the UCAL along the beam axis.*

UCAL Part	Angular coverage	EMC $x \times y$	HAC $x \times y$
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.*

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

$$\frac{\sigma(E)}{E} = \frac{18\%}{\sqrt{E}} \oplus 2\%, \tag{2.4}$$

while the hadronic resolution is

$$\frac{\sigma(E)}{E} = \frac{35\%}{\sqrt{E}} \oplus 1\% \tag{2.5}$$

where E is the particle energy, measured in GeV.

The UCAL is divided into three parts: the forward (FCAL), barrel (BCAL) and rear (RCAL) calorimeter (Fig. 2.10). 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 part of the calorimeter 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.

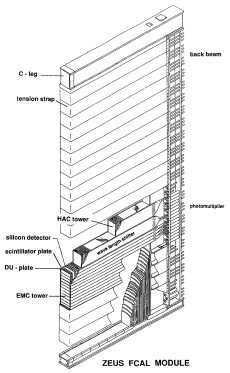

Figure 2.11: *A FCAL module.*

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 26 X_0, while each hadronic section is 3.1 λ deep. The EMC and HAC cells are superimposed to form a rectangular module, one of which is shown in Fig. 2.11. 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 (WLS) 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 tracks with 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 it will be illustrated later, in the trigger Section.

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%.

2.2.3 The Muon Detectors

The main aim of these detectors is the measurement of the tracks coming from the interaction region (*pointing tracks*) and that are able to cross the whole calorimeter and the iron yoke (*penetrating tracks*). This behaviour is characteristic of muons, that can cross large amount of material without being absorbed since, being much heavier than the electrons and not interacting strongly, they lose their energy just by ionization.

The muon detection system, as the other detectors in ZEUS, has to keep into account 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 found, whereas 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 detectors, designed and realized in different ways: the Forward Muon Detector (FMUON) and the Barrel and Rear Muon Detector (BMUON and RMUON).

2.2.4 The Forward Muon Detector (FMUON)

The Forward Muon Detector is made of two modules (Fig. 2.12): one is located between the FCAL and the BAC (inner detector, FMUI), the other is positioned outside the BAC (outer detector, FMUO). The FMUON detector consists of:

- a system of four limited streamer tubes [87] 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 $5^o < \theta < 16^o$.

The first LST plane and the first drift chamber make up the FMUI detector, while the FMUO detector consists of the rest of the system.

2.2.4.1 The Limited Streamer Tubes (LST) 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 sheat. 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 are induced on copper strips with polar geometry, glued on the outer side of the plastic sheat. The number of radial ρ strips, 1.9 cm wide, is 132, and 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 last 64 strips far away from the beam line are OR'ed two by two since as θ becomes larger a coarser resolution is needed.

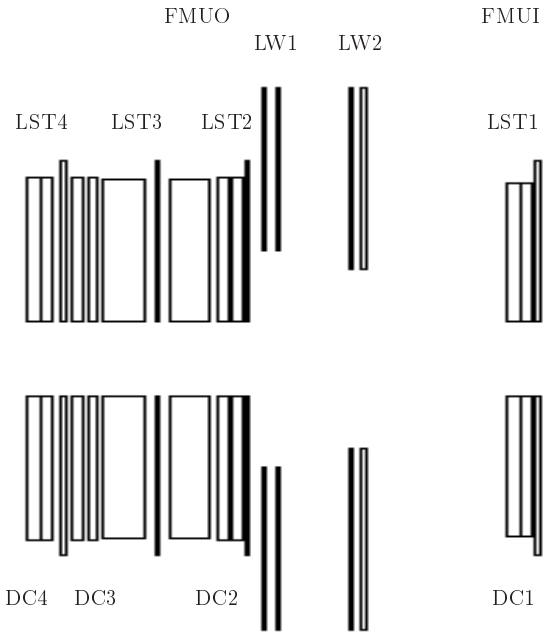

Figure 2.12: *Schematic view of the forward muon detector along the beam axis.*

Therefore the number of channels needed for each octant is 96. The number of ϕ strips is 32 per octant: each covers a 1.4^o interval in azimuth.

2.2.4.2 The drift chambers (DC)

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.

2.2.4.3 The Large Angle Coverage Planes (LW)

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 ($16^o < \theta < 32^o$).

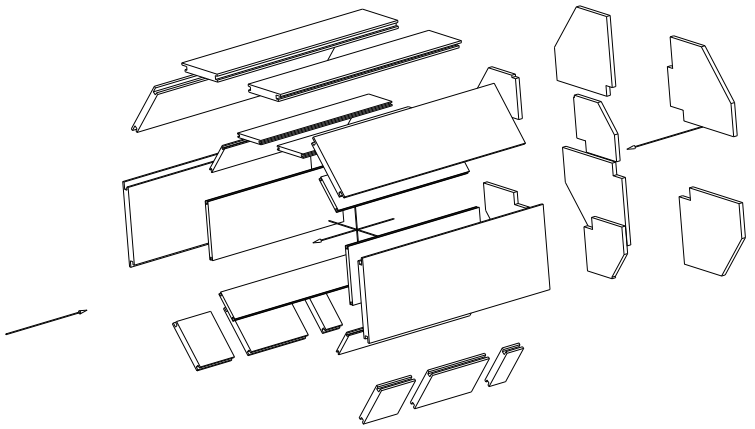

Figure 2.13: *An exploded view of the Barrel and Rear Muon Detector, showing the positioning of the chambers.*

Each plane consists of eigth aluminum tight wrappings that contain a LST layer. The LST signal is induced on copper strips with radial geometry, spaced of 0.7^o 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 charge barycentre method, is $\sim$ 1 mm.

2.2.5 The Barrel and Rear Muon Detector (B/RMUON)

The Barrel and Rear Muon Detector [88] has to cover a very large area, of the order of two thousands squared meters, so a modular structure was chosen. The basic element is the *chamber*. The chambers covering the inner barrel part, between the CAL and the iron yoke, are called BMUI, whereas the chambers situated in the outer barrel part, ouside the yoke, are denoted as BMUO. In a similar way, in the rear region the detector is divided into RMUI and RMUO chambers. In Fig. 2.13 the layout of B/RMUON is schematically shown. The chambers have different shapes and dimensions depending on where they are located, but their internal structure is kept the same: the element bearing the weight of the chamber is an aluminum honeycomb structure, 20 cm thick in the rear chambers, 40 cm in the barrel ones. On both sides of the honeycomb a couple of planes of limited streamer tubes (LST) [87] is placed (see Fig. 2.14). The choice of LST planes was due to the large area to be covered, joined with the necessity of a good space resolution, of the order of $\sim$ 1 mm. Each tube contains eight cells, each with one sense wire; the distance between two sense wires is 1 cm. During the data–taking the sense wires are brought to $\sim$ 4500 V so that they behave as anodes, while the inner cell walls, covered by graphite, act as cathodes. The cells are filled with a mixture of carbon dioxide (CO_2), argon and isobuthane. On one of the outer wall of the LST plane conductive strips orthogonal to the wires are placed. The distance between two adjacent strips is 1.5 cm. The signal read by the wires is also induced

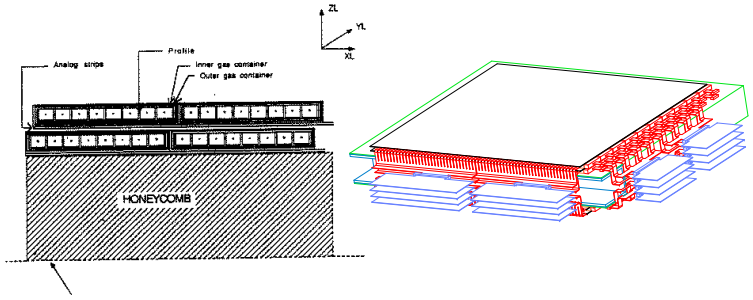

Figure 2.14: *Transverse section of a BMUON chamber, with the honeycomb structure visible* **(left)**. *Scheme of a BMUO chamber* **(right)**.

on the strips, so that a single plane of wires and strips is sufficient to determine the particle position. For redundancy each chamber has two LST planes on each side. The cells in one plane are displaced with respect to those in the other plane by half cell, in order to have good acceptance also for the particles passing near the boundary of two cells in a plane. In the BMUI and BMUO chambers the LST are parallel to the beam direction, whereas in RMUI and RMUO they are horizontal (parallel to the ZEUS x direction).

The determination of the position of a particle passing through the muon chambers proceeds by collecting the information coming from the wires, read by electronic cards called SGS, and from the strips, with readout cards named STAR. The reading of the cards depends on the presence of a trigger signal in the event. There are two types of trigger signals in an event: when the muon chambers identify a *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 specified by the inner tracking detectors (*pointing tracks*). In the case of positive response by the matrices, a strong trigger signal is present.

When there is 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 (analog readout). With the analog readout the achievable spatial resolution on the coordinate orthogonal to the wires is 200 μm, while it is 700 μm for the coordinate parallel to the wires.

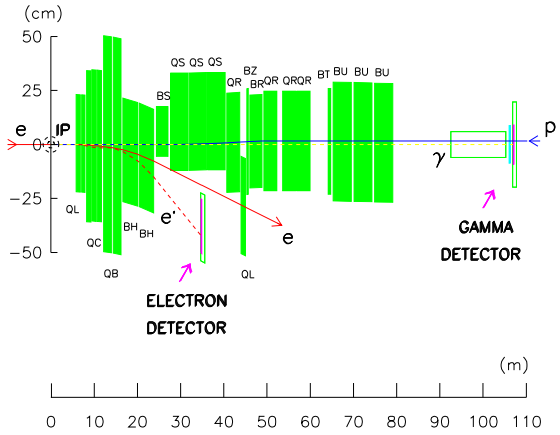

Figure 2.15: *Schematic view of the luminosity monitor.*

If there is no trigger signal in the event, the recorded information are the addresses of the hit wires and of the OR of four adjacent strips, called *quadruplets* (digital readout). 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.

2.2.6 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 [89] and is known with an accuracy of $\sim 0.5\%$, allowing a precise determination of the ep 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. 2.15). To protect the photon calorimeter against synchrotron radiation, it has been 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 extracted 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 paired 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%.

2.2.7 The ZEUS trigger system

The ZEUS trigger [90] is a three–level system, each level being differently sophisticated: the first level, having to deal with very big event rates, handles simpler information from the events, while the second can use more refined objects, and in the third even part of the offline reconstruction software can be run, in order to decide if an event has to be written to disk or not. The aim of the ZEUS trigger system is to reject background events, coming mainly from beam–gas interactions and from cosmics, and at the same time to keep events coming from ep interactions, having much smaller rate ($\mathcal{O}(1)$ Hz) than the background.

2.2.7.1 The First Level Trigger (FLT)

The First Level Trigger has to strongly suppress the following background events, in order to pass a cleaner sample to the other trigger components:

- 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 with a vacuum of 10^{-9} Torr, is 50 kHz;
- 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; these events, however, have a typical topology (the muons are typically parallel to the proton beam direction) and usually can easily be distinguished from ep events;
- cosmic ray muons, with a rate of the order of 1 kHz.

The First Level Trigger has to deal with the HERA bunch crossing so it has to handle events at a rate of 10 MHz, giving as output events at a rate of the order of 1 kHz, the design rate of the Second Level Trigger. The FLT is a hardware trigger, designed to analyse one event for every bunch crossing. The data of each bunch crossing are stored into pipelines, that are 46 bunch crossing deep and allows the FLT a time of 4.4 μs to accept or discard an event. The FLT operates on a subset of the full data coming from an event, basing essentialy on the calculation of crude event observables (regional energy sums, number of tracks, timing information...).

Each detector component has its own first level trigger processor, and the 4.4 μs interval has to be shared between the components trigger and the Global First Level Trigger (GFLT). After 26 bunch crossing times (2.5 μs) all the components send their FLT signal to the GFLT, which uses the remaining 20 crossing times (1.9 μs) to take the final decision. The components data are processed and combined in parallel in eight Trigger Logic Modules of the GFLT, and 64 individual sub–triggers (*slots*) are generated. The GFLT accepts or refuses the event looking at the OR of these 64 sub–triggers.

If the event is accepted, all the components have to digitize their data in order to send them to a system of digital CPUs for the next analyses. This operation takes $\sim$ 10 μs after the GFLT decision, and during this time no event acquisition is possible. This is the only dead–time of the GFLT chain, and is of the order of 1%.

2.2.7.2 The Second Level Trigger (SLT)

The Second Level Trigger (SLT) [91] has the aim to further reduce the background events with respect to the ep events. The SLT, which receives events from the FLT with a rate of 1000 Hz, has an output rate of 100 Hz. The SLT is software–based and runs on a network of transputers [92]. 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. As in the FLT, each detector component has its own SLT processor, and all the information from the single components are sent to the Global Second Level Trigger (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 was shown in section 2.2.2, 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. Therefore, a proton–gas event, originating upstream of the detector, will produce particles arriving to the RCAL before than to the FCAL, so with time in FCAL and RCAL differing by a $\sim$ 10 ns. In the same way, events coming from lepton–gas interactions downstream the detector will produce particles arriving to the FCAL before than to 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. In more detail, the filters applied to separate ep and background events are:

- *RCAL timing:* used to reject events coming from proton beam interactions with the residual gas in the beam pipe. Events are rejected if $|T_{RCAL}| > 8$ ns.
- *FCAL timing:* used to reject events coming from interactions of the lepton beam with residual gas in the beam pipe. Events are rejected if $|T_{FCAL}| > 8$ ns.
- *F–RCAL timing:* events coming from real ep interactions have particles with the same time in FCAL and in RCAL, so that $T_{FCAL} - T_{RCAL} \simeq 0$, while beam–gas events upstream the interaction point have times in FCAL and RCAL that differs of $\sim$ 10 ns. The events are rejected if $T_{FCAL} - T_{RCAL} > 8$ ns.
- *Up–down timing:* used to reject cosmic muon events, that reach the upper part of the BCAL before than the lower. The event is rejected if both the halves of the BCAL have valid timing, if there is any activity nor in the FCAL neither in the RCAL, and if $T_{up} - T_{down} < -10$ ns.
- *E and p_z*: used to reject beam–gas interactions occurring near the interaction region. Since these events can essentially be thought as proton collisions on a fixed target, the produced particles have a small polar angle so that

$$\frac{\sum_i E_i \cos\theta_i}{\sum_i E_i} \sim 1 \tag{2.6}$$

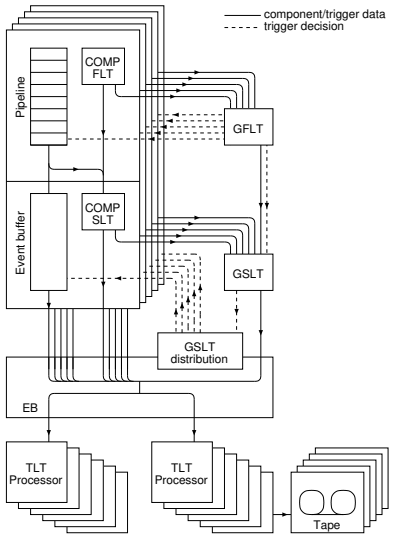

Figure 2.16: *Schematic diagram of the ZEUS trigger and data acquisition system.*

where E_i is the energy of the ith cell of the calorimeter and θ_i is its polar angle. The event is rejected if

$$\frac{\sum_i E_i \cos\theta_i}{\sum_i E_i} > 0.96. \qquad (2.7)$$

- *Global timing:* if one of the calorimeter times is valid and greater than 10 ns, the event is rejected.

2.2.7.3 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 of 100 Hz, while the output rate is $3-5$ Hz, similar to the rate of the ep interactions. Therefore the aim of the TLT is not only to reject background events but also to select the particular classes of ep interactions under investigation.

The TLT consists of a serie of algorithms (those used for the analysis reported here will be described in detail later) written in FORTRAN, running on a farm of SGI CPUs, where a partial event reconstruction is done. The analysis of the events is not parallel, but the events are distributed to all the CPUs of the farm. In addition to the reconstruction program, also some selection programs are run on the data, in order to select good ep events. The main information that the reconstruction program sends to the selection algorithms are the addresses and the energies of the calorimeter cells above

threshold, and the parameters (momentum and position) of the tracks reconstructed by the CTD. The resolution on the energy and on the track parameters is lower than that obtained with the final reconstruction program.

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

A schematic view of the ZEUS trigger and data acquisition system is shown in Fig. 2.16.

Chapter 3

Physics simulation

The use of simulation programs is essential in physics analyses. In fact, a better understanding of the data and of the detector behaviour can be achieved by simulating the detector response to physics events. Moreover, the theoretical models implemented in the simulations can be tested by comparisons to real data.

The simulation of physics events at HERA is done in two main steps. First, the ep scattering process is simulated using an *event generator*. This program, following the prescriptions of the theoretical models implemented in it, provides the complete list of the four–momenta of the final state particles. In the second step, all the detector and the trigger system are simulated, in order to determine their response to the particles produced in the physical process.

These simulations are based on Monte Carlo (MC) techniques, which turn out to be an essential tool to understand the complexity of high energy physics processes and of particle detectors.

3.1 General structure of an event generator

The factorization theorem for hard processes [93] is the main theoretical justification for QCD Monte Carlo event generators. Following the prescriptions of the theorem, an ep scattering process characterized by a hard scale can be factorized into the following separate stages (Fig. 3.1):

- *hard sub–process:* it is the main feature of the event, the interaction between a parton, extracted from the proton, and the photon, or a photon constituent in resolved photon events. This process can be calculated in a fixed order perturbative expansion since it involves a hard scale μ;

- *initial and final state radiation:* in processes involving charged and coloured objects the topology of an event can be strongly influenced by the emission of gluons and photons in the initial or final state. These perturbative corrections are usually modelled by the so called *parton shower method*: the radiation is simulated by an arbitrary number of branchings of one parton into two, like $e \to e\gamma$, $q \to qg$, $q \to q\gamma$, $g \to q\bar{q}$. The kernel $P_{a\to bc}(z)$ of a branching gives the probability distribution of the energy sharing, with daughter b taking a fraction z and daughter c the remaining $(1-z)$ of the initial energy E_a. The two daughters may branch

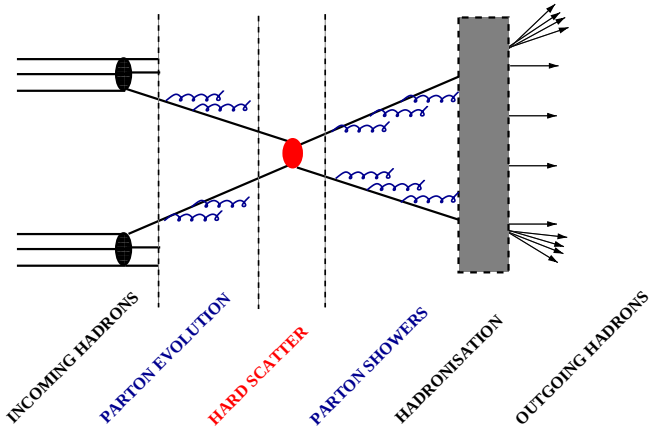

Figure 3.1: *Schematic representation of the processes involved in a typical hard QCD interaction.*

in turn, producing other partons, and so on. Via the *initial state radiation* a parton, constituent of the incident hadron (photon), having low space–like virtuality, radiates time–like partons, increasing its space–like virtual mass. On the other hand, via the *final state radiation* an outgoing virtual parton with large time–like mass generates a shower of partons having lower virtuality. The shower evolution stops at some fixed scale μ_0, typically of the order of 1 GeV;

- *hadronization:* it is the process in which colourless hadrons are formed starting from coloured partons. It is a non–perturbative phenomenon and is still not completely understood. Therefore simulation programs model the hadronization process using phenomenological inputs. The main hadronization models now available are the cluster model, implemented in the HERWIG Monte Carlo (see Sect. 3.2) and the string model, implemented in the PYTHIA Monte Carlo program (see Sect. 3.3);

- *beam remnant:* in ep interactions, the simulation programs reconstruct the particles that have gone through the initial state radiation and have then generated the partons involved in the hard sub–process. These initial particles, one from each beam, are a parton extracted from the proton and a photon or, in resolved photon processes, a parton from the photon. They carry only a fraction of the initial beam energy, the rest is taken by the beam remnant. If the shower initiator is coloured, also is the beam remnant, that is therefore connected to the rest of the event and has to be fragmented and reconstructed coherently.

3.2 The HERWIG Monte Carlo

The HERWIG (Hadron Emission Reactions With Interfering Gluons) [52] Monte Carlo program is a general–purpose particle physics event generator which can simulate lepton–lepton, lepton–hadron and hadron–hadron interactions. It has a relatively small number of free parameters, a large predictive power, and is as independent as possible of non–perturbative parameters.

The initial and final state radiation are factorized as successive branching processes in which the energy fractions are distributed according to the Altarelli–Parisi splitting functions. The full available phase–space is restricted to an angular ordered region, as a result of interference, and leading infrared singularities are correctly kept into account.

In the case of final state radiation, at each branching the angle between the two emitted partons is smaller than that of the previous branching. In the case of initial state radiation, the angular ordering applies to the angle θ_i between the incoming hadron and the emitted partons i. The scale of α_s at each branching is set at the relative transverse momentum of the two emitted partons. In the case of heavy flavour production, the mass of the quark modifies the angular–ordered phase–space so that the soft radiation in the direction of the heavy quark is depleted. The emission within a cone with an opening angle m_Q/E_Q, where m_Q and E_Q are the mass and the energy of the heavy quark, vanishes. This angular screening determines the shape of the heavy flavour jet.

A large variety of electroweak and QCD processes can be simulated using HERWIG. The factorization scale μ^2 of the hard subprocess is given by:

$$\mu^2 = \frac{2\hat{s}\hat{t}\hat{u}}{\hat{s}^2 + \hat{t}^2 + \hat{u}^2} \tag{3.1}$$

where $\hat{s}$, $\hat{t}$ and $\hat{u}$ are the Mandelstamm variables. The resulting cross section for light flavour production diverges in the limit $p_T \to 0$, therefore a cut on the minimum transverse momentum of the produced partons has to be applied.

HERWIG uses the Equivalent Photon Approximation (EPA) [3] to generate the spectrum of the photons radiated from the incoming lepton in an ep scattering. The boson–gluon fusion process in heavy flavour production uses the exact cross section of the process $ep \to Q\bar{Q}X$ as fully computed by Schuler [94]. In the case of resolved photon processes, the program allows at each branching the possibility for a dynamical $\gamma \to q\bar{q}$ anomalous splitting.

In the hadronization process, three types of non–perturbative contributions have to be considered: the representation of the incoming partons as constituents of the incoming particles, the conversion of the outgoing partons into hadrons, and the description of the soft underlying event due to the presence of spectator partons.

The incoming parton treatment is related to the factorization theorem of collinear singularities. A phenomenological input structure function describes the distribution of the partons longitudinal momentum fractions, while the transverse momentum distribution is characteristic of the size of the hadron.

Regarding the formation of hadrons from the outgoing partons below the time–like cutoff scale Q_0, perturbative QCD predicts [95] that in hard processes the confinement of partons is local in colour and independent of the hard scale μ. The cluster

hadronization model used in HERWIG is assumed to satisfy these requirements. After the perturbative parton branching process, all the outgoing gluons are splitted non–perturbatively into light (u or d) quark–antiquark or diquark–antidiquark pairs, the splitting into diquarks being suppressed with respect to that into quarks. Quarks are then combined with their nearest neighbours to form colour singlet clusters, whose mass distribution is peaked at low values and falls rapidly for large masses. The clusters formed in this way are fragmented into hadrons following some prescriptions. If a cluster is too light to decay into two hadrons, it is taken to represent the lightest single hadron of its flavour, and its mass is shifted to the right value by an exchange of momentum with a neighbouring cluster. The clusters which are massive enough decay isotropically into pairs of hadrons, following some rules that bring to the formation of an unbiased selection of decay products conserving flavour. The small fraction of clusters having too high masses for isotropic two–body decay are fragmented using an iterative fission model, until the masses of the products are below the fission threshold. The mechanism is not unlike string fragmentation (see Sect. 3.3).

The spectators extracted from the incoming particles in hadron–hadron and lepton–hadron collisions are included into beam clusters. In the formation of those clusters, the colour connection between the spectators and the initial state parton shower is cut by the forced emission of a quark–antiquark pair. In a lepton–hadron collision, the soft underlying event is represented by a soft collision between the beam cluster and the adjacent cluster, produced in the forced emission mentioned above.

3.3 The PYTHIA Monte Carlo

The PYTHIA [96] Monte Carlo is a general–purpose event generator, sharing many common features with HERWIG, but having also some significant differerences, in particular in the treatment of the non–perturbative processes. The PYTHIA philosophy, in fact, is to describe also the hadronization processes in as much detail as possible.

PYTHIA and HERWIG are slightly different also in the description of the hard subprocess. In particular, in photoproduction events PYTHIA uses the Weiszäcker and Williams [97] approximation to generate the spectrum of the photons radiated from the incoming lepton, whereas HERWIG uses the equivalent photon approximation (see Sect. 3.2). Another difference is the scale of the hard scattering, μ^2; PYTHIA sets μ^2 to the transverse mass of the two outgoing partons, m_T:

$$\mu^2 = \frac{1}{2}(m_3^2 + p_{T_3}^2 + m_4^2 + p_{T_4}^2) = \frac{1}{2}(m_3^2 + m_4^2) + \frac{\hat{t}\hat{u} - m_3^2 m_4^2}{\hat{s}}, \qquad (3.2)$$

whereas the scale used by HERWIG is given by the Formula 3.1.

The parton shower processes implementation is similar to that of the HERWIG Monte Carlo, but in PYTHIA the parton emissions are ordered only according to the virtuality of the radiated partons. The perturbative parton radiation is stopped at some cutoff scale Q_0^2, of the order of 1 GeV^2.

The main difference between PYTHIA and HERWIG is the hadronization model. In PYTHIA fragmentation is performed by the JETSET program, which implements the *string fragmentation* or *Lund model* [98]. The string model is based on the starting assumption of a linear confinement picture. According to this model, the energy stored in the colour dipole field between a charge and an anticharge increases linearly with the separation between the charges. This assumption is supported by lattice QCD studies performed on one of the simplest possible systems, the colour–singlet $q\overline{q}$ 2–jet event.

As q and $\overline{q}$ move apart from their common production vertex, the physical picture is that of a colour flux tube, with transverse dimensions of the typical hadronic sizes (1 fm), being stretched between the q and the $\overline{q}$. If the tube is uniform along its length, this leads to a confinement picture with a linearly rising potential. The simplest way to obtain a Lorentz covariant and causal description of the energy flow due to this confinement is the use of the dynamics of the massless relativistic strings with no transverse degrees of freedom.

Coming back to the fragmentation process, as q and $\overline{q}$ move apart the potential energy stored in the string increases, and the string may break producing a new $q'\overline{q}'$ pair. The original system therefore splits into two colour–singlet systems $q\overline{q}'$ and $\overline{q}q'$. Further breakings may occur if the invariant mass of either of these string pieces is large enough. The string break–up process proceeds until only on–mass–shell hadrons remain.

In the generation of quark–antiquark pairs from string break–ups quantum mechanical tunneling is used, leading to a flavour–independent Gaussian spectrum for the transverse momentum p_T of the $q'\overline{q}'$ pair. Since the string has no transverse excitation this p_T is locally compensated between the quark and the antiquark of the pair. The p_T of a hadron is made up of the p_T of the quark and antiquark forming it. The tunneling picture also implies a suppression of heavy quark production so that charm and heavier quarks are not expected to be produced in the soft fragmentation, but only in perturbative parton shower branchings $g \to q\overline{q}$. A tunneling mechanism can also be used to explain the production of baryons.

The different string breakings are causally disconnected. This means that, for example, starting from a system made up of a quark moving in the $+z$ direction and an antiquark moving in the $-z$ direction, the formation of mesons can be described in the following way: by a string break–up, the production of a $q_1\overline{q}_1$ pair brings to the formation of a $q\overline{q}_1$ meson, leaving behind an unpaired quark q_1. A second pair $q_2\overline{q}_2$ may now be produced, to give a new meson with flavours $q_1\overline{q}_2$, etc. This process can be iterated until all energy is used, with some modifications near the $\overline{q}$ end to make the total energy and momentum consistent.
The case of several partons moving apart from a common origin is more complicated. For a $\overline{q}qg$ event, a string is stretched from the q to the $\bar{q}$ end via the gluon, so that the gluon has two string pieces attached. The string fragments into two $\overline{q}q$ pairs, boosted with respect to the centre–of–mass frame of the original $q\bar{q}g$ system, and into one hadron straddling both string pieces.

3.4 The detector simulation

All the event generators supported in ZEUS, like PYTHIA and HERWIG, are gathered in a software program called AMADEUS. The user can choose the event generator, which gives as output all the four–momenta of the particles produced in the hard scattering process and all the relevant kinematic variables.

The data produced by an event generator are the input to the ZEUS detector

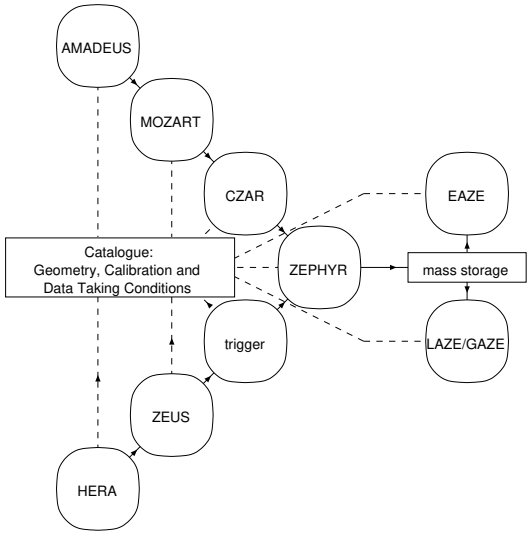

Figure 3.2: *A schematic diagram of ZEUS data and Monte Carlo reconstruction chain.*

and trigger simulation program, MOZART[1]. MOZART is based on the GEANT [99] package, whose kernel contains a description of all the detector components, including the material they are made of, their shapes and positions. The program traces the particles through the whole detector, simulating its response and keeping into account physics processes such as energy loss, multiple scattering and particle decays in flight.

The events pass then through the CZAR[2] [100] package, that simulates the trigger logic as implemented in the data taking.

As a final step, the generated sample is processed by the ZEUS reconstruction program, ZEPHYR[3]. This program reconstructs the event variables, like particle momenta and energies, treating the data and the Monte Carlo in the same way. All the information coming from the different detectors making up ZEUS are taken as inputs by ZEPHYR.

All the ZEUS data are organized using the ADAMO[4] [101] data management system, used for the data storage in memory or on external media and for their documentation.

The access to the data by the users is done with the EAZE program; the LAZE program generates bi–dimensional graphical representation of the real or simulated events. A diagram of the ZEUS reconstruction scheme for data and Monte Carlo is shown in Fig. 3.2.

[1] MOnte Carlo for Zeus Analysis Reconstruction and Triggering.
[2] Complete Zgana Analysis Routine.
[3] ZEus PHYsics Reconstruction.
[4] Aleph DAta MOdel.

3.5 Monte Carlo samples used for the analysis

In this Section, a brief description of all the Monte Carlo samples generated for the simulation of the data used in this work will be given. Due to the fact that the HERA centre–of–mass energy was changed in 1998 from ~ 300 to ~ 320 GeV, the data samples corresponding to the two different data–taking periods have been analysed separately. Two complete Monte Carlo samples with different centre–of–mass energy have been generated for the two periods. The characteristics of the different samples will be illustrated in the following paragraphs.

3.5.1 Beauty events simulation

The production of beauty quarks in an event proceeds through different sub–processes. The main contribution to the cross section is given by the *direct photon* process, which at LO reduces to the boson–gluon fusion:

$$\gamma g \to Q\overline{Q}. \tag{3.3}$$

Other lower but still significant contributions to the cross section come from *resolved photon* processes. Other processes leading to the production of beauty quarks are the *beauty excitation* phenomena, in which a beauty quark is extracted from the photon or from the proton.

Monte Carlo samples simulating all these sub–process have been generated, the respective cross sections and the integrated luminosities generated are reported in Table 3.1. Beauty events generated by PYTHIA have been used to determine the central values of the cross sections, whereas beauty samples generated by HERWIG have been used to determine the systematic uncertainties, mainly related to the different hadronization models used by the two generators. Four different samples were generated using the PYTHIA Monte Carlo, simulating direct, resolved and excitation processes. The sub–processes simulated in each sample were:

- b **in direct** γ**:**

$$\begin{aligned} \gamma\, g &\to b\, \overline{b}; \end{aligned}$$

- b **in resolved** γ**:**

$$\begin{aligned} q\, \overline{q} &\to b\, \overline{b}; \\ g\, g &\to b\, \overline{b}; \end{aligned}$$

- b **excitation in** γ**:**

$$\begin{aligned} b\, q &\to b\, q; \\ b\, g &\to b\, g; \end{aligned}$$

- b **excitation in** p**:**

$$\begin{aligned} b\, \gamma &\to b\, \gamma; \\ b\, g &\to b\, g; \\ b\, q &\to b\, q; \\ b\, \overline{b} &\to b\, \overline{b}. \end{aligned}$$

In the case of beauty excitation in the proton, both direct and resolved processes are involved.

Two samples were generated using the HERWIG Monte Carlo, one simulating all the direct γ processes (boson–gluon fusion, direct processes of beauty excitation in the proton) the other including all the resolved photon processes ($q\,\overline{q}$ and $g\,g$ interactions, and beauty excitation in the photon). Therefore, in Table 3.1, referring to HERWIG, the direct and resolved phenomena have to be interpreted in the sense defined above.

In the event generation, the b–quark mass was set to $m_b = 4.5$ GeV. The parton densities used for the proton and the photon were CTEQ4L [49] and GRVG–LO [28], respectively. The events were generated with no requirements on the presence of muons or jets. The beauty quarks were not forced to decay into muons: as a consequence, in the beauty Monte Carlo samples the fraction of beauty quarks decaying into muons is given by the branching fraction implemented in the simulation. This choice was made in order to generate a Monte Carlo sample suitable for the evaluation of the beauty photoproduction cross section without requests on the semi–leptonic decay into muons.

3.5.2 Charm events simulation

The production mechanisms of $c\overline{c}$ quark pairs are identical to those of $b\overline{b}$. Charm can be produced in direct photon processes, like the boson gluon fusion, and in resolved processes. Also charm excitation in the photon and in the proton gives a sizable contribution to the cross section. The charm photoproduction cross section is expected to be two orders of magnitude larger than that of beauty; the resolved processes are expected to give a larger contribution than in the case of beauty. Resolved processes are in fact more important for lower masses of the produced quarks.

Monte Carlo samples simulating charm photoproduction have been generated using PYTHIA. The cross sections and the integrated luminosities generated are shown in Table 3.1. As in the case of beauty, four samples were generated, simulating the different sub–processes. For each sample, the simulated interactions are the same described above in the case of beauty, with b quarks replaced by charm quarks.

In the event generation the mass of the c quark was set to $m_c = 1.35$ GeV. Charm quarks were forced to decay into muons and jets had to be produced from the charm hadronization and decay. An event was included in the simulation if at least one muon and two jets were found in the final state satisfying the following requirements:

- **muons:** the muon was asked to be in the acceptance region of the muon chambers:

  $$p_T^{\mu} > 2 \text{ GeV} \quad \text{or} \quad p_z^{\mu} > 3 \text{ GeV} \quad \text{or} \quad p_z^{\mu} < -1.5 \text{ GeV};$$

  and

  $$10^o < \theta^{\mu} < 170^o \quad ;$$

- **jets:** due to the different trigger requirements (see Sect. 4.4.3) the cuts are different for the two Monte Carlo samples simulating 1996–97 and 1999–2000 data:
  - **1996–97:** $E_T^{\text{jet}1,2} > 4$ GeV, $-3 < \eta^{\text{jet}1,2} < 3$;
  - **1999–2000:** $E_T^{\text{jet}1} > 4.5$ GeV, $E_T^{\text{jet}2} > 3.5$ GeV, $-3 < \eta^{\text{jet}1,2} < 3$.

3.5.3 Light quark events simulation

In light quark photoproduction, due to the low mass of the produced quarks, resolved photon processes are expected to be dominant. The cross section is much larger than that of charm photoproduction, and therefore also of beauty.

Monte Carlo samples simulating light flavour photoproduction in direct and resolved photon processes have been generated. The cross sections and the integrated luminosities generated are reported in Table 3.1, where the light flavours are indicated by l. The sub–processes simulated in each sample were:

- l **in direct** γ**:**

  $$f\,\gamma \to f\,g;$$
  $$g\,\gamma \to f\,\bar{f};$$

- l **in resolved** γ**:**

  $$\begin{aligned} f\,f' &\to f\,f' \text{ (QCD)};\\ f\,\bar{f} &\to f'\,\bar{f}';\\ f\,\bar{f} &\to g\,g;\\ f\,g &\to f\,g;\\ g\,g &\to f\,\bar{f};\\ g\,g &\to g\,g. \end{aligned}$$

As in the case of charm, cuts have been applied on the generated events in order to reduce the size of the data to be analysed, keeping only those of interest for the analysis. As it will be shown in detail in Chapter 5, a possible source of background for the analysis reported here is given by two–jet events, in which a light particle (usually a pion, kaon or proton) is misidentified as a muon by the muon chambers. The probability of such a particle to reach the muon detectors is very low, but its production cross section is much higher than that of beauty quarks, and therefore it is necessary to estimate the probability of these particles to give a fake muon signal. Hence, the generated light quark events were asked to have at least two jets in the final state, and a relative high momentum particle in the acceptance region of the muon chambers. More precisely, the imposed cuts were:

- at least one track in the angular region defined by:

  $$10^o < \quad \theta \quad < 170^o$$

  and satisfying the following momentum cut:

  $$p_T > 1.35 \text{ GeV} \quad \text{or} \quad p > 2.5 \text{ GeV};$$

- **jets:** at least two jets, reconstructed by the KTCLUS algorithm (see Sect. 4.2.2) and satisfying the following requirements:

  $$E_T^{\text{jet}1,2} > 5.25 \text{ GeV} \quad \text{and} \quad -3 < \eta^{\text{jet}1,2} < 2.7.$$

The production of heavy quarks was forbidden in the event generation. Strong cuts were imposed in the generation in order to reduce the number of events to store: nevertheless, million of events had to be generated in order to have a significant statistics for this kind of processes (see Table 3.1).

PYTHIA				
	1996–97 sample		**1999–2000 sample**	
Process	σ (pb)	$\mathcal{L}$ (pb^{-1})	σ (pb)	$\mathcal{L}$ (pb^{-1})
b in direct γ	6716	119	7305	320
b in resolved γ	1147	174	1282	335
b exc. in γ	1542	176	1728	275
b exc. in p	95	150	109	325
c in direct γ	135200	253	125500	325
c in resolved γ	12220	273	12020	448
c exc. in γ	96380	269	87580	302
c exc. in p	9420	272	49540	302
l in direct γ	87420	137	93360	100
l in resolved γ	443100	130	491900	107
HERWIG				
	1996–97 sample		**1999–2000 sample**	
Process	σ (pb)	$\mathcal{L}$ (pb^{-1})	σ (pb)	$\mathcal{L}$ (pb^{-1})
b in direct γ	4600	195	4940	324
b in resolved γ	1040	192	1170	342

Table 3.1: *Cross section and integrated luminosity generated for the various Monte Carlo samples, beauty, b, charm, c, and light flavours, l. The statistics are reported separately for MC samples simulating 1996–97 and 1999–2000 events.*

Chapter 4

Event reconstruction and data selection

In this Chapter the selection of events of interest for the work reported here is described. An event is defined *interesting* if it is possible that a $b\bar{b}$ couple has been produced. The experimental signature of this kind of events is the presence of at least two jets in the final state, coming from the hadronization of the b quarks, and of at least one high transverse momentum muon, used to tag beauty via its semi–leptonic decay.

First, the reconstruction method for the kinematic variables of interest, like jets and muons parameters, and kinematic variables used to distinguish photoproduction and deep inelastic scattering regimes, is illustrated. Then, the trigger chain used for the online selection of the event sample is described. Finally, the cuts applied to the data to obtain the final photoproduction sample are illustrated, together with the comparison between the data and the Monte Carlo.

4.1 Reconstruction of y and Q^2

In a lepton–proton collision, the inelasticity, y, and the squared four–momentum transferred, Q^2, defined in Sect. 1.1, are related to the kinematic of the outgoing lepton:

$$Q^2 = -(p_e - p'_e)^2 \tag{4.1}$$

$$y = \frac{P \cdot Q}{P \cdot p_e} \tag{4.2}$$

where p_e and p'_e are the four–momenta of the incoming and outgoing positron, respectively, and P is the four–momentum of the incoming proton. These two variables can be evaluated in many ways; here two methods are described in detail: the *electron method* and the *Jacquet–Blondel method* [57].

The *electron method* can be used when the kinematic of the outgoing lepton is known with precision. Naming E_e and E'_e the energies of the incoming and outgoing lepton, respectively, and θ_e the scattering angle of the outgoing lepton, it follows from simple considerations of relativistic dynamics that y and Q^2 can be expressed as follows:

$$y_e = 1 - \frac{E'_e}{2E_e}(1 - \cos\theta_e) \tag{4.3}$$

$$Q^2_e = 2E_e E'_e(1 + \cos\theta_e) \tag{4.4}$$

The electron method cannot be used when, as in the photoproduction case, the outgoing lepton cannot be detected in the calorimeter since it is scattered at very low angles. With the ZEUS acceptance the maximum polar angle at which a positron can be identified in the detector is $\theta_e \sim 177^o$, corresponding to $Q^2 \sim 1$ GeV^2.

In the photoproduction case, neglecting the proton mass, the inelasticity can also be expressed as:

$$y_e = \frac{E_p E_\gamma (1 - \cos\theta_{p\gamma})}{E_p E_e (1 - \cos\theta_{pe})} \tag{4.5}$$

where $\theta_{p\gamma}$ is the opening angle between the proton and the photon and θ_{pe} is the opening angle between the proton and the positron. In the collinear approximation the photon is emitted collinearly with the incoming positron, so that $\theta_{p\gamma} = \theta_{pe}$, and therefore:

$$y_e \simeq \frac{E_e - E'_e}{E_e} = \frac{E_\gamma}{E_e}. \tag{4.6}$$

The Formula 4.6 can be obtained also from Eq. 4.3, in the limit $\cos\theta_e \to -1$, i.e. when the scattered lepton is deflected at a very small angle (PHP). Looking at Eq. 4.6 it is clear why y_e can be interpreted as the fraction of the positron momentum transferred to the photon.

When the information on the outgoing positron are not available, the reconstruction of the y variable is performed using the *Jacquet–Blondel* method. The quantity:

$$\delta = \sum_f (E_f - p_{z,f}) \tag{4.7}$$

where f runs over all the final state particles, and E_f and $p_{z,f}$ are the energy and momentum along the beam axis (z) of each particle, is not very sensitive to the fact that the detector is not completely hermetic. This is because by subtracting p_z the contribution to the sum coming from all the particles escaping detection along the forward beam pipe is made negligible. These particles have in fact typical high momenta along the z axis, but very low transverse momenta. If the detector measures correctly all the final state particles the quantity δ is conserved, within a good approximation:

$$\sum_f (E_f - p_{z,f}) = \sum_f E_f (1 - \cos\theta_f) \approx \sum_{i=e,p} E_i (1 - \cos\theta_i) = 2E_e \tag{4.8}$$

where i runs on the initial state particles, the proton and the positron. The y_{JB} variable is then defined as:

$$y_{JB} = \frac{\sum_f (E_f - p_{z,f})}{2E_e}. \tag{4.9}$$

If the scattered positron is identified in the detector, then $y_{JB} \sim 1$, whereas if the outgoing positron escapes the detection y_{JB} is an estimation of the inelasticity y of the event.

In an analogous way, the Q^2_{JB} variable is expressed as:

$$Q^2_{JB} = \frac{(\sum_f p_{x,f})^2 + (\sum_f p_{y,f})^2}{1 - y_{JB}}. \tag{4.10}$$

Clearly, on real data the complete information on the kinematic of all the final state particles are not available. Therefore the y_{JB} variable is reconstructed by combining information coming from the tracking system (see Sect. 2.2.1) and the calorimeter (see Sect. 2.2.2) as will be shown later in this Chapter.

4.2 Jet reconstruction

The features of the jets in a hadronic final state are strictly related to those of the partons which gave origin to them. However, jets are complex objects, and they are not uniquely defined in QCD: their definition relies on the algorithms used to reconstruct them. In the Snowmass workshop in 1990 [102] some criteria were fixed to be satisfied by every jet reconstruction algorithm. In particular, two requirements have to be fulfilled, in order to define an *infrared* and *collinear safe* algorithm, reconstructing jets in the proper way:

- the results must be independent of the fact that one parton can split into two partons moving collinearly, or, from the experimental point of view, that a particle can release energy in two adjacent calorimeter cells. This dependence in fact causes collinear divergences in the theoretical calculations, which disappear if no distinction is made between two particles having energy $E_1 + E_2 = E$ and one single particle of energy E, moving in the same direction. From the experimental point of view, this means that the results are independent of the detector granularity;

- the results must be independent of the emission of very low energy particles; this fact causes infrared divergences in the theoretical calculations, removed by integration; in experiments these small energy deposits are related to the noise of the detector, removed by using thresholds or corrected by suitable algorithms.

Jet algorithms to be used for the analysis of photoproduction processes have to fulfill two further requirements: they have to treat the proton and, if present, the photon remnant in a proper way, that means, the remnants have to be separated from the jets and not to influence their search. Moreover, in photoproduction processes in general the laboratory frame is different from the frames of physical interest (i.e. the γp centre–of–mass frame). Nevertheless, the interesting frames are usually Lorentz boosted along the beam direction with respect to the laboratory frame. Therefore the jet algorithm has to be independent of this kind of transformations: this can be done by reconstructing jets using their transverse energy in a pseudorapidity–azimuth plane $(\eta - \phi)$.

4.2.1 Cone algorithms

The standard definition of a cone algorithm was given by the Snowmass Convention [102]: a jet is a cone of radius R in the $\eta - \phi$ plane, containing a maximum in the transverse energy of the particles (or the calorimeter cells) composing it. The jet consists of the particles or calorimeter cells i which are at a distance R_i from the centre of the jet, satisfying:

$$R_i = \sqrt{(\eta_i - \eta^{\rm jet})^2 + (\phi_i - \phi^{\rm jet})^2} \leq R. \tag{4.11}$$

where η_i and ϕ_i are the pseudorapidity and azimuth of the ith particle. The transverse energy and the coordinates of the jet are defined as:

$$E_T^{\rm jet} = \sum_i E_{t,i} \tag{4.12}$$

$$\eta^{\rm jet} = \frac{1}{E_T^{\rm jet}} \sum_i E_{t,i}\eta_i \tag{4.13}$$

$$\phi^{\rm jet} = \frac{1}{E_T^{\rm jet}} \sum_i E_{t,i}\phi_i. \tag{4.14}$$

In ZEUS, two variants of the cone algorithm are used, EUCELL [103] and PUCELL [104], both based on the pre–clustering technique, but differing in the recombination method. In PUCELL, a pre–cluster is formed by combining the highest E_T object with all the objects within a radius R in the $\eta - \phi$ plane, until all objects are included into a pre–cluster. The centre of the cluster is then recalculated and the process is repeated until a stable situation is reached. The clusters with transverse energy above some threshold are considered jets.

Also EUCELL uses the pre–clustering concept, but clusters are identified using a grid in the $\eta - \phi$ plane, with cells having dimensions $\Delta\eta^{\rm cell} \approx \Delta\phi^{\rm cell} \approx R/2$. A pre–cluster is a 3×3 cell window over the grid, above some fixed transverse energy threshold. A cone of radius R is placed around the pre–cluster and an iterative process is performed. The first jet is defined as the one with highest transverse energy, the following jets are determined using the same procedure until no more cones are found above some fixed transverse energy threshold.

The use of cone algorithms can lead to ambiguities: the Snowmass Convention does not give limits on the radius of the cone, or rules on how to treat overlapping jets. Moreover, the process is not infrared safe at NNLO without modifications [105].

4.2.2 Clustering algorithms

The ambiguities found in the use of cone algorithms can be solved by using clustering algorithms, in which the treatment of overlapping jets is clear, and the assignment of hadrons to jets can be done using the same procedure both in theoretical calculations and in experiments. This is the reason why in this work a clustering algorithm has been used to reconstruct jets, the k_T algorithm [106], implemented in the KTCLUS [60] library of routines.

The resolution variable used by the k_T algorithm to identify the jets is the relative transverse momentum k_T between the particles. The use of this variable follows from the fact that, with the present understanding of perturbative QCD, the jets are not sprays of hadrons confined in cones of fixed angle. Soft hadrons produced coherently by the fragmentation of hard partons should be assigned to the jet of the hard parton nearest in angle, independently of the actual value of its angular distance [106, 107]. This means that the jets have an effective radius depending on the hardness of the jet itself and on the color flow of the hard subprocess.

The clustering procedure of the k_T algorithm is performed in two steps:

- **Pre–clustering of hadrons:** for every final–state hadron h_i and for every hadron pair (h_i, h_j) the resolution variables d_i and d_{ij} are evaluated. These

quantities have, respectively, in the small angle limit, the meaning of the squares of the transverse momentum of the hadron i with respect to the beam axis, $k_{T,i}^2$ and of the relative transverse momentum of the hadron i relative to j, $k_{T,ij}^2$:

$$\begin{aligned} d_i &\simeq E_i^2\theta_i^2 \simeq k_{T,i}^2 && \text{for } \theta_i \to 0 \\ d_{ij} &\simeq \min\{E_i^2, E_j^2\}\theta_{ij}^2 \simeq k_{T,ij}^2 && \text{for } \theta_{ij} \to 0 \end{aligned} \tag{4.15}$$

where θ_i is the angle between the particle direction and the beam, and θ_{ij} is the angle between the two hadrons. Then, the minimum between (d_i, d_{ij}) is considered. If d_{ij} is the minimum, the partons i and j are combined into a single cluster (*pseudoparticle*), while if d_i is the minimum, the hadron i is included in the *beam jets*. The procedure is repeated for all particles and pseudoparticles not included in the beam jets, until all the objects have d_i, d_{ij} larger than some parameter $d_{\rm cut}$, which defines the hard scale of the process. At the end of the procedure, one ends up with beam jets, which include also the beam remnants, having vanishing transverse momentum relative to the beam axis, and with other objects called *hard final–state jets*.

- **Resolving event structure into sub–jets:** after having factorized the hard scattering process, jet structure can be resolved in the same way as in e^+e^- processes. First, a resolution parameter $y_{\rm cut}$ is defined, and for any pair of hadrons in a hard final–state jet the rescaled resolution variable y_{ij} is considered:

$$y_{ij} = \frac{d_{ij}}{d_{\rm cut}}. \tag{4.16}$$

The smallest y_{kl} of the y_{ij} is taken, and if $y_{kl} < y_{\rm cut}$ the corresponding hadron pair (h_k, h_l) is combined into a single cluster (pseudoparticle) $h_{(kl)}$. The procedure is repeated until all pairs of objects have $y_{ij} > y_{\rm cut}$. The resulting objects are called *sub–jets*.

The small angle form of the resolution variables, d_i and d_{ij}, ensures that the main features of jets in the soft and collinear limits are correctly described in terms of QCD coherence. Away from that limit, it can be shown [106] that the variables d_i and d_{ij} admit a *longitudinal–boost–invariant* extrapolation of the form:

$$d_i = p_{t,i}^2, \qquad d_{ij} = \min\{p_{t,i}^2, p_{t,j}^2\} R_{ij}^2 \tag{4.17}$$

where $p_{t,i}$ is the transverse momentum of the hadron i with respect to the beam axis, and the simplest form of the radius R_{ij} is

$$R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2. \tag{4.18}$$

The recombination scheme, i.e. the way in which the resolution variables of the pseudoparticles are defined, has also to be taken into account in the definition of the clustering algorithm. It can happen, in fact, that the jet properties, and in particular the hadronization corrections, are influenced by the recombination scheme. Here three different recombination schemes are briefly illustrated:

- **covariant E–scheme:** the pseudoparticle (h_i, h_j) is considered as a real particle with momentum $p_{(ij)} = p_i + p_j$, or, more precisely, once the hadrons h_i and h_j have been merged into a cluster, the corresponding resolution variables are:

$$d_{(ij)} = (\vec{p}_{ti} + \vec{p}_{tj})^2 \qquad d_{(ij)k} = \min\left[(\vec{p}_{ti} + \vec{p}_{tj})^2, \vec{p}_{tk}^{\,2}\right] R^2_{(ij)k}. \tag{4.19}$$

  In this case, the pseudoparticle is not massless.

- **p_t–weighted scheme:** the form 4.17 of the resolution variables is used both for particles and pseudoparticles, and the pseudoparticle kinematic variables are obtained by the p_t–average of the corresponding particle variables:

$$\begin{aligned} p_{t(ij)} &= p_{ti} + p_{tj} \\ \eta_{(ij)} = \frac{p_{ti}\eta_i + p_{tj}\eta_j}{p_{t(ij)}} &\qquad \phi_{(ij)} = \frac{p_{ti}\phi_i + p_{tj}\phi_j}{p_{t(ij)}}. \end{aligned} \tag{4.20}$$

- **monotonic p_t–weighted scheme:** in this case the resolution variables are:

$$d_{(ij)} = (p_{ti} + p_{tj})^2 \qquad d_{(ij)k} = \min\left[(p_{ti} + p_{tj})^2, p_{tk}^2\right] R^2_{(ij)k} \tag{4.21}$$

  and the radius is defined as:

$$R^2_{(ij)k} = \frac{p_{ti}^2 R_{ik}^2 + p_{tj}^2 R_{jk}^2}{p_{ti}^2 + p_{tj}^2}. \tag{4.22}$$

The KTCLUS algorithm satisfies all the requirements of a good jet algorithm: it can separate true jets from the beam remnants and it is infrared and collinear safe. It is invariant for longitudinal Lorentz boosts, and it can be used in the same way in experiments and in theoretical calculations. Then, since every particle (or calorimetric cell) is without ambiguity associated to a unique jet, the problem of overlapping jets does not exist.

4.3 Muon reconstruction

Due to the different structure of the BRMUON and FMUON detectors (see Sect. 2.2.4 and 2.2.5) two different muon reconstruction algorithms exist, one for muon identification in the forward region, MPMATCH [108], and one for muons in the barrel and rear region, BREMAT [109]. These two algorithms, used in the data analysis presented here, are briefly described below. More information can be found in Refs. [108, 109].

4.3.1 The MPMATCH matching package

The MPMATCH package has the main purpose to match segments reconstructed by the FMUON detector with tracks reconstructed by the inner tracking devices, mainly the CTD (see Sect. 2.2.1). It has been developed for the selection and the reconstruction of a clean muon sample in inelastic events with non–isolated muons, and therefore the angular coverage is limited to the overlapping region between the FMUON and the CTD.

A FMUON track is defined by five parameters in the ZEUS reference frame: $(x, y, dx/dz, dy/dz, Q/p)$. The search for a match starts from a FMUON reconstructed track. A θ and ϕ road is opened around the track, and tracks from the CTD are searched inside the road. The FMUON track, with its parameters and its covariance matrix, is then extrapolated backward to the z coordinate measured by CTD, using the GEANE [110] package. Then, a fit between the extrapolated FMUON track and the CTD track parameters is performed using Kalman filtering [111]. If in the CTD there is more than one track inside the road, a choice based on the χ^2 probability is done so that just one track is retained. The procedure is then repeated for all available FMUON tracks. Once that a match is found, a vertex fit is done, extrapolating the FMUON track toward the z of the reconstructed vertex, and performing a fit between the parameters of the extrapolated track and those of the CTD track, the x and y coordinates of the reconstructed vertex and the direction and momentum of the track.

4.3.2 The BREMAT matching package

The *Barrel and Rear Extrapolation MATching* package, BREMAT [109], is used to match segments reconstructed in the barrel and rear muon detectors to tracks measured in the inner tracking detectors, mainly the CTD. The limited streamer tubes, and the associated strips, used in the BRMUON detector have a resolution of the order of $\sim$ 1 mm on both the coordinates they measure, x and y (see Sect. 2.2.5) but the momentum resolution for most of the muons is dominated by multiple scattering in the iron yoke placed between the inner and the outer chambers. The most powerful way to reconstruct muons in this context requires that measurement errors, multiple scattering and energy losses are correctly taken into account, and this is the technique used by BREMAT.

The main purpose of the algorithm is to find out candidate muons, associating inner detector tracks to segments in the muon chambers and providing the resulting matching χ^2. The main input to the algorithm is the MBXYSG table [112], containing the information on the reconstructed segments in the BRMUON chambers, and the VCTRHL table [113], containing the parameters of the tracks reconstructed by the inner detectors. When an entry in MBXYSG is found, i.e. a segment in the muon chambers, BREMAT looks for candidates to be associated to it in the VCTRHL table, i.e. between all the tracks reconstructed by the inner tracking devices, mainly the CTD. A loose preselection is done on VCTRHL tracks to be associated to MBXYSG segments:

- the momentum p of the track has to be $p > 1$ GeV; this is a minimal request for tracks that have to cross all the calorimeter before being identified by the chambers;
- the polar angle θ of the track has to be $\theta > 20^o$ (better acceptance of the tracking detectors, mainly CTD; rejection of forward tracks);
- the track has to start from CTD superlayer 1 and to extend at least to superlayer 3 (see Sect. 2.2.1);
- the impact parameter of the track, D_H, has to be $|D_H| < 10$ cm;

- the z coordinate at the point of closest approach has to be $|z_H| < 75$ cm;
- $\chi^2_{\rm track}/n.d.f. < 5$;
- $\Delta \leq 150$ cm, Δ being the distance between a central point on the BRMUON segment and the straight line obtained by extrapolating the CTD track at the calorimeter entrance.

Tracks passing the preselection are extrapolated through the calorimeter using the GEANE [110] package. The extrapolation of the parameters and error matrix of the candidate track proceeds from the outer surface of the inner tracking devices, through the calorimeter up to a reference surface on the inner muon chambers, where matching is done. This surface is shaped as a prism, having one base on the rear chambers plane at $z = -310.53$ cm, the other base lying on the forward end of the barrel chambers, at $z = +450.0$ cm, and the side faces corresponding to the barrel sectors (see Sect. A.5). The track extrapolation is done outward, starting from the inner region, for the best treatment of the low–momentum muons, which suffer big energy losses compared to their initial energy. With the GEANE package also tracks that completely loose their energy before reaching the chambers are simulated.

The output track parameters and error matrix at the reference surface are usually referred to as *predictions*. The predicted variables and error matrix are obtained in a convenient Cartesian parametrization, related to the local coordinate frame of each sector. The slopes of the track x' and y' are also given in order to determine the predicted direction of the particle. The fifth coordinate is the momentum of the particle Q/p, that can be measured by the BRMUON detector just in the case the muon reaches the outer chambers.

Naming ξ_{M_i} and ξ_{P_i}, respectively, the measured and the predicted track parameters at the reference surface, the residuals are defined as $\delta_i = \xi_{M_i} - \xi_{P_i}$. If the coordinate of the Cartesian parametrization are written as $(x, y, x', y', Q/p)$ the matching χ^2 is given by:

$$\chi^2 = \sum_{i,j} S^{-1}_{ij} \delta_i \delta_j \tag{4.23}$$

where the sum goes from 1 to $n = 4$ or 5, depending whether the momentum is included or not in the χ^2, and S^{-1} is the inverse of the covariance matrix of the residuals:

$$S_{ij} = \sum_{k,l} T_{ik} (\sigma^2_{\rm track})_{kl} T^T_{lj} + (\sigma^2_{\rm random})_{ij} + (\sigma^2_{BRMU})_{ij} \tag{4.24}$$

with $k, l = 1, ..., n$; T is the transport matrix, between the start and the stop of the swim, $\sigma^2_{\rm track}$ is the error matrix of the inner detector track, $\sigma^2_{\rm random}$ the error matrix due to multiple scattering and energy loss, and σ^2_{BRMU} the error matrix of the BRMUON segment from the MBXYSG table. The first two terms are calculated by steps during the GEANE extrapolation using the detailed geometry of the ZEUS detector and the magnetic field map. BREMAT makes a loose cut on the matched tracks, considering as matching the tracks having $\chi^2 < 100$.

4.4 Trigger chain for online data selection

The first data selection is done by requiring in the events the presence of at least two jets, coming from the hadronization of the heavy quark pair. The trigger chain used to select this kind of events is described in detail in this Section, at all the three levels of the ZEUS trigger system (see Sect. 2.2.7). Some of the trigger requests were changed in 1998, due to the fact that the proton–beam energy was raised from 820 to 920 GeV.

4.4.1 Jet First Level Trigger

At the first level trigger, events are asked to fire at least one of the following slots:

- **FLT 40:** the energy reconstructed in the electromagnetic part of the calorimeter has to be $E_{\rm emc} > 14968$ MeV;
- **FLT 41:** the transverse energy reconstructed by the UCAL has to be $E_T > 20982$ MeV;
- **FLT 42:** this trigger slots requires the *or* of four calorimetric requirements:
  - the total calorimetric energy has to be $E > 14968$ MeV;
  - the total electromagnetic energy has to be $E_{\rm emc} > 10068$ MeV;
  - the electromagnetic energy collected by the barrel calorimeter has to be $E^{\rm BCAL}_{\rm emc} > 3404$ MeV;
  - the electromagnetic energy collected by the rear calorimeter has to be $E^{\rm RCAL}_{\rm emc} > 2032$ MeV;

  Moreover, at least one track coming from the nominal interaction region, $-80\,{\rm cm} < z_{\rm vtx} < 80$ cm, where $z_{\rm vtx}$ is the z coordinate of the vertex, as given by the CTD, has to be present in the event.
- **FLT 43:** the transverse energy has to be $E_T > 11574$ MeV, and a track coming from the nominal interaction point has to be found in the event.

All these trigger slots require some vetoes criteria to be fulfilled: the timing observed in the C5 counters and in the SRTD have to be consistent with physics timing and not with that of the beam gas, and no coincidence of the inner and outer veto wall has to be found.

In the determination of the energies measured by the calorimeter, the three inner rings around the FCAL beam pipe and the innermost around the RCAL beam pipe are excluded.

The track requests, when present, also ask the ratio of vertex to non–vertex tracks to be greater than $25 - 30\%$, depending on the track multiplicity in the event.

4.4.2 Jet Second Level Trigger

At the second level trigger, events are selected having large transverse energy. The requirements the events have to fulfill are the following:

- -60 cm $< z_{\rm vtx} < 60$ cm, where $z_{\rm vtx}$ is the z coordinate of the vertex as reconstructed by the CTD;
- at least one vertex track has to be present in the event;
- $p_z/E < 0.95$ if $8 < E - p_z < 12$ GeV; E is the energy of all the particles in the event and p_z their momentum along the z axis, all reconstructed by the UCAL; the event is always accepted if $E - p_z > 12$ GeV;
- $E_T({\rm box}) > 8$ GeV, $E_T({\rm box})$ being the sum of the transverse energy collected in all the UCAL cells, excluding the first ring around the FCAL beam pipe.

4.4.3 Jet Third Level Trigger

Due to the enlarged proton–beam energy, changed in 1998 from 820 to 920 GeV, and to the consequent increase of the centre–of–mass energy, the requests of the jet Third Level Trigger (*dijet* trigger bit) have slightly changed in the two data–taking periods used in this analysis. The following requests were kept the same in both 1996–97 and 1999–2000 data–taking:

- the event must have a vertex;
- -60 cm $< z_{\rm vtx} < 60$ cm, $z_{\rm vtx}$ being the z coordinate of the vertex;
- $E - p_z < 75$ GeV
- $p_z/E < 1.0$;
- the number of bad tracks has to be lower than six; a good track is defined when:
  - number of d.o.f. > 20;
  - $P_T > 0.2$ GeV
  - $0.35 < \eta < 3.13$
  - number of hits in the CTD axial superlayers > 5;
  - number of hits in the CTD stereo superlayers > 5;
  - z value at the distance of closest approach ≤ 75 cm.

These two different requirements on the jet transverse energy were applied in 1996–97 and 1999-2000 data taking:

- **1996–97**: at least two jets with transverse energy $E_T > 4$ GeV and pseudorapidity $\eta < 2.5$;
- **1999–2000**: at least two jets with transverse energy $E_T > 4.5$ GeV and pseudorapidity $\eta < 2.5$;

In both the data–taking periods, the jets at the TLT level were reconstructed using a simplified version of the EUCELL cone algorithm (see Sect. 4.2.1).

After the trigger selection, the data sample to be used in the rest of the analysis amounts to about ten million of events.

4.5 Offline data selection

Having selected events using the trigger chain described above, the final data sample will be chosen by applying proper cuts on the variables of interest for the analysis. Some of these, for example the calorimetric variables, need to be corrected before cutting on them, in order to make them more similar to the *true* ones, that can be obtained from the Monte Carlo. The motivations and details of the corrections will be described below. Our aim is to obtain, after having applied all the cuts, a clean dijet sample enriched in the $b\bar{b}$ component.

4.6 Rejection of DIS events

In order to select a photoproduction event sample, deep inelastic scattering events have to be rejected, in which the incoming positron is scattered at some measurable angle and can be identified in the detector. The rejection of this kind of events is performed using an electron finder, i.e. an algorithm having the purpose to analyse the energy deposits in the calorimeter, in both EMC and HAC cells, and to distinguish between electromagnetic and hadronic clusters. The electron finder used in this analysis is SINISTRA [114]. The algorithm proceeds in two steps: first, SINISTRA takes as input the energies of the calorimeter cells, and, using a neural network, described in detail in [114], gives as output the probability for each cluster to be electromagnetic or hadronic. Then, the second part of the algorithm selects the scattered positron from the list of the candidates.

The algorithm can be used in different ways; for the analysis reported here, the choice has been done to require the candidate positron to be reconstructed only by the UCAL, with no request of an associated track in the CTD. The motivation behind this choice is the fact that the DIS cross section falls rapidly as the Q^2 of the event increases. Therefore, in most of our DIS background events the outgoing lepton is scattered at a very small angle, and is hence detected by the UCAL but not by the CTD, since it falls outside its coverage region.

SINISTRA gives as output the lepton candidate having the highest probability to be the scattered positron, providing its probability is greater than 0.9. For this candidate also the energy deposited in the calorimeter, E_{el}, and the inelasticity, y_{el}, are given as output by the algorithm. The event is rejected if:

$$E_{el} > 5 \text{ GeV} \quad \text{and} \quad y_{el} < 0.9. \tag{4.25}$$

The request on y_{el} is done to reduce the risk to reject photoproduction events wrongly identified as DIS. In fact, final state pions, electrons or photons, present in a photoproduction event, can be misidentified as the scattered lepton. In this case, however, the reconstructed y_{el}, that clearly is not the inelasticity of the event, is usually large. Therefore, events with an electromagnetic deposit of considerable energy but high y_{el} are included in the analysis as photoproduction events.

4.7 Reconstruction of calorimetric variables

As it was pointed out above, the inelasticity and the jet variables (transverse energy, pseudorapidity and azimuth) in an event can be evaluated using the energy released by the particles and measured by the calorimeter. In Monte Carlo events the evaluation of the *true* variables is possible, using as input the four–momenta of the final–state particles. The values measured on real data, i.e. using the calorimeter, differ from the true ones. Some corrections have to be applied to the data in order to make the measured values more similar to the true ones. In the first part of this Section, all the corrections to be applied to the energy measured by the calorimeter will be illustrated:

- rejection of noisy cells and suppression of noise–related effects;
- comparison of the calorimeter response to real data and to Monte Carlo events;
- simulation of effects related to the UCAL geometry (cracks, dead material, etc.).

Then, in the second part, we will focus on the possibility of merging calorimeter and tracking system information in order to better reconstruct the calorimetric variables. The concept of ZUFO will be introduced.

After all the corrections have been applied, the results will be investigated by comparing the true variables to the reconstructed ones, before and after the corrections. Results regarding y_{JB} and the jets variables will be shown.

4.7.1 Calorimeter–related corrections

The first correction applied to the calorimeter measurements is the suppression of badly–working calorimeter cells and of noise–related effects. The list of noisy cells for each year is reported in standard routines, used to correct for these effects. The correction falls into distinct categories: the removal of noisy cells according to the list, a cut on the relative cell imbalance[1], and the removal of isolated EMC and HAC cells with energies below 80 and 140 MeV, respectively. The noise associated with the imbalance and isolation cuts is mainly due to sparks in the PMTs, while the noisy cells come mainly from electronic malfunctioning.

A second correction regards the calorimeter energy scale. Studies on the calorimeter response to data and Monte Carlo [115] have shown that there is a multiplicative factor between the actual calorimeter energy scale and the one simulated in MOZART. This discrepancy can easily be corrected by multiplying the energy as measured on real data by a correction factor, dependent on the sector of the calorimeter (FCAL, BCAL, or RCAL) and on the type of the cell (EMC or HAC) (see Sect. 2.2.2). These correction factors are summarized in Table 4.1. No correction is applied to the Monte Carlo.

In addition to these two effects, that can easily be corrected, systematic discrepancies between the variables measured by the calorimeter and the true ones arise from the calorimeter structure and the way it measures particles: the presence of dead material in front of the UCAL, cracks between the calorimeter sectors, and the presence of particles like muons and neutrinos which do not release all their energy in the calorimeter are the causes of a systematic shift of the calorimeter measurements toward lower

[1] As described in Sect. 2.2.2, each cell is read on both sides by a photomultiplier. The imbalance is defined as the fractional difference of the two independent measurements.

Sector	Cell	Correction on DATA	Correction on Monte Carlo
FCAL	EMC	1.04	1
	HAC	0.95	1
BCAL	EMC	1.04	1
	HAC	1.08	1
RCAL	EMC	1.025	1
	HAC	1.025	1

Table 4.1: *Correction factors to be applied to data and Monte Carlo to keep into account the difference in the calorimeter energy scale. Correction factors are 1 for the Monte Carlo, while for data they depend on the detector sector and on the type of the cell.*

values than the true ones. All these effects have to be taken into account in order to correct the calorimetric variables up to the true values. An estimation of the goodness of the corrections can be obtained by comparing the true variables and the corrected ones on a selected Monte Carlo sample.

Suitable routines have been developed in order to correct for the effects reported above. In particular:

- the map of the *dead material* is available and has been used to write corrections algorithms. The materials constituting the beam pipe, the tracking devices and the solenoid correspond to a number of radiation lengths varying from 1 to 3 (see Fig. 4.1) and the energy loss of the particles, especially of those having low momenta, can be relevant in such a thickness. Since energy losses due to the presence of the dead material are difficult to be included with sufficient precision in the detector simulation, the correction is done offline, and is parametrized as a function of the energy and of the polar angle of the particles;
- the zones of the cracks between the calorimeter sectors are not well simulated by the Monte Carlo so corrections have been introduced offline;
- pions, kaons and protons having energies lower than 1 GeV loose all their energy by ionization before any hadronic interaction. In this case the calorimeter is no more compensated ($e/h \sim 0.6$) and this causes an overestimation of the energy of the low–energy hadrons. Algorithms correcting for these effects have been used in this thesis.

In reconstructing the calorimetric variables, like y and the jets parameters, another point has to be considered, that, for low energy charged particles, the reconstruction of the particle energy is done better by the tracking devices than by the calorimeter. In ZEUS the technique of merging information from the CTD and the UCAL for the best estimation of the inelasticity and jet variables has been introduced. The resulting objects are called ZUFOs (Zeus Unidentified Flow Objects).

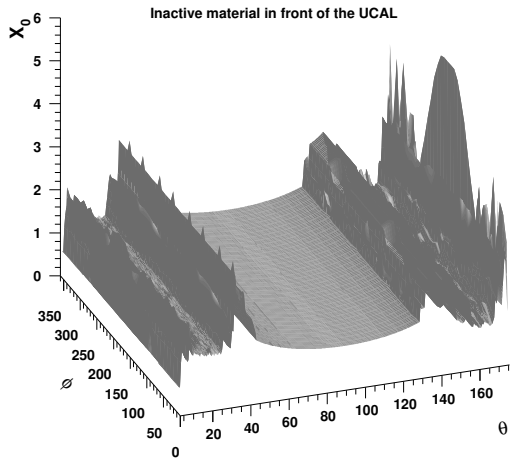

Figure 4.1: *Distribution of inactive material in front of the UCAL in units of the radiation length, X_0, in the $\theta - \phi$ plane, as implemented in the simulation of the detector.*

4.7.2 ZUFOs

The use of the ZUFOs is based on the fact that the hadronic energy has both a charged particle and a neutral particle component. Both are measured by the calorimeter, but a large fraction of the charged particles is also measured by the tracking detectors. In some cases, especially when the charged particles have low energies, or when they cross a large thickness of dead material before being detected, the resolution of the tracking devices is better than that of the calorimeter.

The use of the ZUFOs rather than that of the usual energy deposits in the calorimeter is also driven by the design of the ZEUS calorimeter (see Sect. 2.2.2), which is divided in three parts, the forward (FCAL), barrel (BCAL) and rear (RCAL) calorimeter. This spatial separation is a serious complication for a local clustering algorithm in handling the energy deposits of a single particle which is not confined in a single part of the UCAL, since the energy will be split in two or more clusters.

Because of these complications, the clustering algorithm for ZUFOs is done in two steps [116]: one local, driving to the formation of *cell islands*, and one based on global clustering and used to create the *cone islands*. The cell island algorithm starts considering each cell and its local neighbourhood. If the cell has sufficient energy, it becomes a candidate to be connected with one of its neighbours, chosen between the highest energy nearest neighbour and the highest energy next to nearest neighbour. The procedure is repeated for each cell so that at the end each cell is uniquely assigned to an island. The algorithm looks for cell islands in each layer of the calorimeter separately (FEMC, FHAC1, FHAC2, BEMC, BHAC1, BHAC2, REMC, RHAC). In this algorithm the detector dependence is hidden in the definition of the local neighbourhood.

The cone island algorithm has the purpose to collect the cell islands which belong

to a shower of a single particle or to a jet of particles. This algorithm has as input the cell islands and performs a clustering in the $\theta - \phi$ plane. The matching starts from cell islands in the HAC2 layer and goes inward. For each HAC2 cell island, the angular separation from all HAC1 cell islands is determined and translated to a probability according to a distribution obtained from a single pion Monte Carlo. The cell with highest probability is taken and matched to the HAC2 cell island if its probability is higher than a fixed cut. If no valid HAC1 cell island is found, an attempt is done of a matching with an EMC cell island.

In a second step, a similar procedure is done for the HAC1 cell island, in order to find a match with an EMC cell. As third and final step, EMC cell islands can be connected to each other. After this linking, HAC cell islands matched to the same EMC cell island are joined to give a cone island. The position of the cluster is not reconstructed as the usual centre of gravity of the shower, calculated using as weights the energies of the cells, because this method yields systematic biases due to varying cells projectivity as seen by the vertex. Instead, the geometric centre of the cell is corrected using the imbalance energy information (each cell is read out on opposite sides by two PMTs) and choosing weights which are logarithmic rather than linear in the cell energy [116].

Once that the cone islands are reconstructed, the next step [117] for the ZUFOs identification is the search for good charged tracks fitted to the vertex, traversing at least 4 superlayers of the CTD and having transverse momentum p_T in the range $0.1 < p_T < 20$ GeV (the maximum p_T is raised to 25 GeV for tracks passing more than 7 superlayers). These good tracks are extrapolated to the inner surface of the calorimeter and associated to cone islands.

The tracks are considered matched to the cone island if the distance of closest approach between the extrapolated track and the island is less than 20 cm, or if it is less than the maximum radius of the island on a plane perpendicular to a ray drawn from the vertex to the island. A schematic view of all the procedure can be seen in Fig. 4.2.

For each of the resulting groups of matched tracks and islands, the decision has to be taken on which energy flow information to use. Three cases are clearer than the others:

- good tracks not associated with any calorimetric object are counted as charged particles, and the CTD information are used. The assumption is done that the particle is a pion;
- calorimeter objects not associated with any track are counted as neutral particles and the calorimeter information are chosen;
- for calorimeter objects associated with more than three tracks the calorimetric information are chosen.

In the case of a 1–to–1 track–island matching, the track information are used instead of those from the UCAL if these two requirements are fulfilled: the energy deposit in the calorimeter has to be due to the associated track alone, and the momentum resolution of the track has to be better than the energy resolution of the corresponding calorimetric object (see Fig. 4.3), i.e. if both the following requests are satisfied:

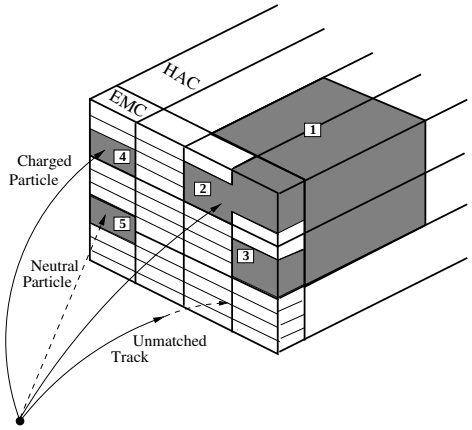

Figure 4.2: *Neighbouring calorimeter cells are clustered into cell islands. This schematic picture shows four EMC cell islands and one HAC cell island. EMC cell islands 2 and 3 are joined with HAC cell island 1 to form a cone island. Then the cone islands are matched to tracks.*

- $E_{\rm cal}/p < 1.0 + 1.2 \cdot \sigma(E_{\rm cal}/p)$: the track momentum is favoured if it exceeds the energy measured in the calorimeter, within the resolution on the measured ratio $E_{\rm cal}/p$. The uncertainty is given by $\sigma(E_{\rm cal}/p) = (E_{\rm cal}/p^2)\ \sigma(p) \oplus (1/p)\ \sigma(E_{\rm cal})$, where $\sigma(p)$ and $\sigma(E_{\rm cal})$ are the resolutions on the momentum and the energy, respectively;
- $\sigma(p)/p < \sigma(E_{\rm cal})/E_{\rm cal}$: the relative uncertainty on the momentum determination has to be lower than that on the measured energy to favour the track momentum over the calorimetric energy. In Fig. 4.3 the transverse momentum resolution of the CTD is compared to the energy resolution of the CAL using a single particle Monte Carlo. It can be clearly seen that low–energy particles are better measured by the CTD than by the UCAL.

Objects having the properties of muons are treated separately; the tracking information are taken instead of the calorimetric ones if:

- $E_{\rm cal} < 5$ GeV;
- $E_{\rm cal}/p < 0.25$;
- $p_T < 30$ GeV.

The more complicated 1–to–2, 1–to–3 2–to–1 and 2–to–2 track–island matches are treated similarly to the 1–to–1 match, substituting the UCAL energy and the CTD momentum with the sum of the energies of the islands and the sum of the momenta of the tracks, respectively.

Finally, in the case a single track is matched to two or more islands, and the energy of the UCAL is favoured, the more precise angular information of the track is used.

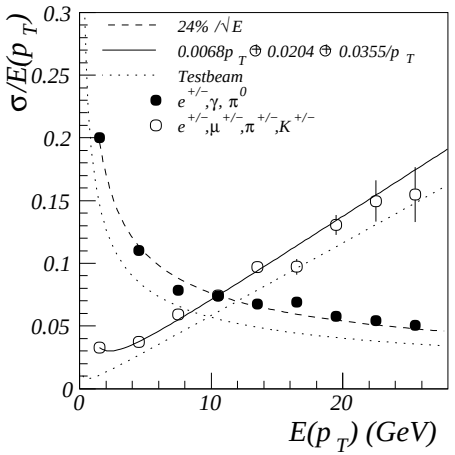

Figure 4.3: *The resolution on the measurement of the electromagnetic energy, E, in the EMC calorimeter and of the transverse momentum, p_T, in the CTD, obtained from single particle Monte Carlo, $(e^{\pm}, \pi^0, \gamma)$ and $(e^{\pm}, \pi^{\pm}, \mu^{\pm}, K^{\pm})$, respectively.*

4.7.2.1 Corrections for the presence of the muon

In the data analysis reported in this thesis, a search is performed of events in which jets are present, and at least one of the jets has to contain a muon. It is therefore of particular importance that the muon is identified properly, and that its momentum is correctly kept into account in the jet reconstruction. Muons are, in fact, particles that do not release all their energy in calorimeters (*minimum ionizing particles, mips*). Hence, if the energy as reconstructed by the UCAL is used to include a muon in a jet, rather than the information from the CTD, the obtained energy of the jet will be systematically lower than the true one.

As it was pointed out above, the ZUFOs algorithm should in principle distinguish the muons from the other particles, and use for them, in the ZUFOs reconstruction, the tracking variables rather than the calorimetric ones. Nevertheless, since this step is of particular interest, another correction has been introduced in order to modify the ZUFOs parameters when the main algorithm does not succeed in the correct identification of the muon. This can be done taking advantage of a particular feature of the ZUFO algorithm: for each ZUFO, the list of the particles belonging to it is available. Therefore, once that the muon algorithms have found the candidate muon between all the CTD tracks, the ZUFO in which the muon is included can be identified. The information are also available on the ZUFO type, i.e. if it has been identified as a muon or not, and if the calorimetric or tracking variables have been used for the ZUFO reconstruction.

This additional correction for the muon presence proceeds as follows: first, the muon reconstruction algorithms, MPMATCH and BREMAT, find the candidate muons and

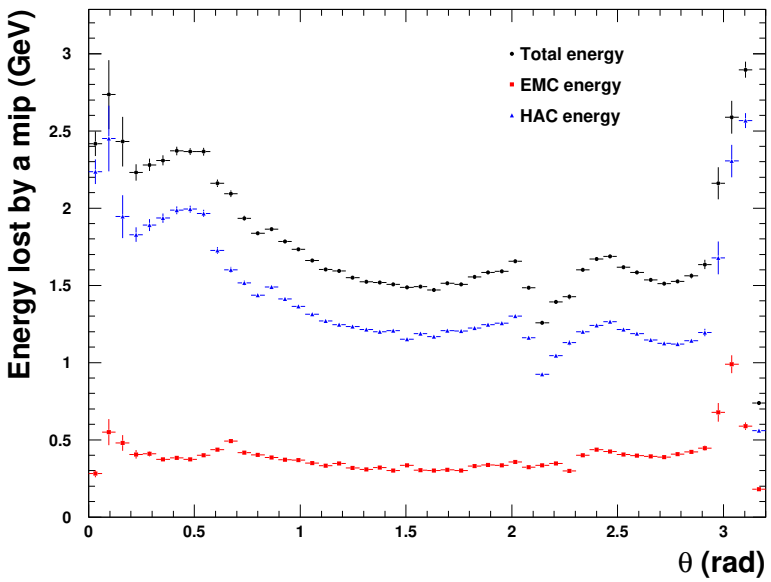

Figure 4.4: *Parametrization of the energy lost in the calorimeter by a minimum ionizing particle, shown as a function of the polar angle θ. The dots are the total energy released, while the squares and the triangles represent the electromagnetic and hadronic contributions, respectively. The plot is done using a data sample of dimuon events, collected by the ZEUS detector during 2000.*

perform the matching between the segments in the muon chambers and the tracks in the inner tracking devices, including also the momentum of the particle. Then, some minimal cuts on the candidate quality are imposed: the muons are asked to have a momentum greater than 2 GeV, as measured by the CTD (a particle with lower momentum cannot reach the outer muon chambers) and a matching probability in momentum greater than 0.01 (as it will be shown in the following, the cut of the analysis will be set to 0.05). If a candidate muon satisfying these requests is found, the algorithm proceeds in the following way: after the ZUFO containing the muon has been identified, its energy is analysed. The ZUFO energy is considered compatible to that of a mip if it is between the 50% and the 150% of the energy predicted for a mip in the same angular range (in Fig. 4.4 the energy released by a mip in the UCAL is shown as a function of the polar angle θ). Then, one of the following can happen:

- if the ZUFO already uses the tracking information, nothing is done;
- if the ZUFO uses the calorimetric information and its energy is compatible to that of a mip at the same polar angle, the ZUFO variables are replaced by those of the muon, as given by the CTD;

- if the muon is included in a jet, usually the ZUFO energy is larger than that of a mip. In this case the ZUFO is splitted into two parts, one having as parameters the muon variables as reconstructed by the CTD, the other being equal to the initial ZUFO, to which the energy of a mip is subtracted; in this way the problem of the underestimation of the jet energy is solved;
- if the muon is matched to a ZUFO having too low energy to be a mip, or does not point to any calorimetric deposit, the CTD information are used instead of that of the UCAL.

Having built and corrected the ZUFOs in this way, the jet parameters and the inelasticity can be reconstructed, using as input the ZUFOs rather than the calorimeter cells. Comparisons to the true variables will confirm if all these corrections are sufficient for our purposes or not.

4.7.3 Results on y_{JB}

In this thesis, since we deal with the photoproduction regime, the Jacquet–Blondel [57] method is used to reconstruct the inelasticity y_{JB} of the event (see Sect. 4.1). From what was pointed out above, the inelasticity can be expressed by the Formula 4.9:

$$y_{JB} = \frac{\sum_f (E_f - p_{z,f})}{2E_e},$$

taking as input the energy E_f measured by each calorimeter cell, or the energy E_f of each ZUFO before any correction, or the same ZUFO energy after all the corrections, and so on. The effects of the use of the ZUFOs rather than the calorimeter cells can therefore be evaluated by analysing the y_{JB} variable reconstructed using the two different inputs. Moreover, the effects of the corrections applied to the ZUFOs can be investigated in this way.

In Fig. 4.5, in the plots on the left, the inelasticity resolution, $(y_{JB}^{\rm rec} - y^{\rm true})/y^{\rm true}$, is evaluated for different calculations of $y_{JB}^{\rm rec}$, and in the scatter plots on the right the reconstructed $y_{JB}^{\rm rec}$ is plotted versus the true variable, evaluated using the electron method (see Sect. 4.1). The event sample, here and in the following, is composed by a mixture of the four beauty Monte Carlo samples (direct, resolved, excitation in the photon and in the proton, see Sect. 3.5) weighted proportionally to their inverse luminosities. No cuts were applied, except for the trigger selection and the cuts for DIS rejection. In the first row of the figure, denoted by **CAL**, the inelasticity is evaluated by looping on the calorimeter cells, and taking as input for the calculation their energy without any correction. It is clearly visible in the Figure that in this case the centre of the ratio distribution $(y_{JB}^{\rm rec} - y^{\rm true})/y^{\rm true}$ is shifted toward negative values (around -0.12), that means that the reconstructed $y_{JB}^{\rm rec}$ is systematically $\sim 12\%$ lower than the true one. The situation is better if the standard ZUFOs (i.e. without any correction) are used to reconstruct $y_{JB}^{\rm rec}$ (second row of the plot, indicated by **Unc. ZUFOs**): the shift toward negative values almost disappears. In the third row, denoted by **Corr. ZUFOs**, ZUFOs with all the corrections applied (i.e. dead material, low energy hadrons and muons) are used for the determination of $y_{JB}^{\rm rec}$. In this case the mean value of the distribution is positive, and the resolution is slightly better than in the case of the

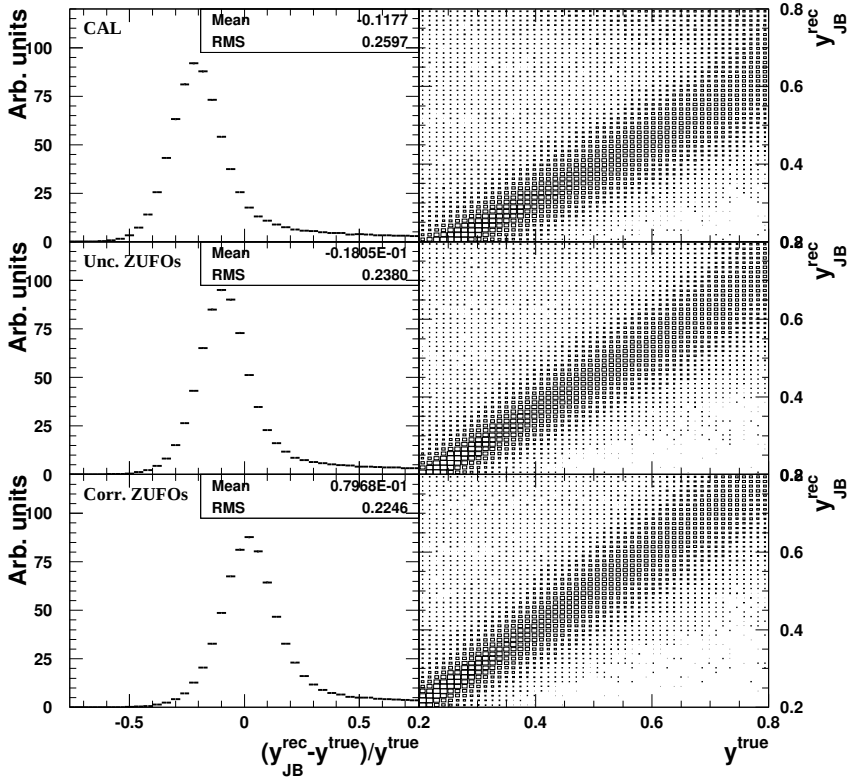

Figure 4.5: $(y_{JB}^{\rm rec} - y^{\rm true})/y^{\rm true}$ **(left)** *and* $y_{JB}^{\rm rec}$ *vs* $y^{\rm true}$ **(right)**, *plotted for various estimations of the inelasticity of the event. In the first row, the resolution* **(left)** *and the comparison between the reconstructed and the true variable* **(right)** *are done calculating* $y_{JB}^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* y_{JB} *is calculated taking as inputs the standard ZUFOs, without any correction; in the third row, all the corrections described in the text have been applied to the ZUFOs. The sample is obtained by mixing the beauty Monte Carlo samples (direct, resolved and excitations) in fractions given by their inverse luminosities so that they sum up to 1.*

standard ZUFOs. Nevertheless, we have to notice that a long tail is present at high values of the ratio, $(y_{JB}^{\rm rec} - y^{\rm true})/y^{\rm true}$. The fact has to be kept into account that no cuts have been applied to this sample, and that therefore most of these events do not contain a muon. All the corrections have to be verified by analysing what happens when all the cuts of the analysis are applied to the Monte Carlo sample, i.e. when *all* the events contain at least one good muon. We will come back to this point after having described all the cuts of the analysis, at the end of this Chapter.

4.7.4 Results on the jet variables

The same analysis done on y_{JB} can be performed on the jet variables. However, before some points have to be clarified. As was shown in Sect. 4.2.2, even if the k_T algorithm is chosen to reconstruct jets, several possibilities are available for the choice of the recombination scheme. In this thesis, having to deal with jets coming from the decay of heavy particles, the E recombination scheme was chosen, in order to have massive objects as output of our algorithm. Other justifications for this choice, mainly related to the correction from parton to hadron jets, will be clearer in the following.

Another important point is the level at which we want to reconstruct jets. On real data, all the information available are the energy and momenta of the final state particles reconstructed by the CTD and the UCAL: this means, of all the particles, except neutrinos, not escaping detection, produced after the quark hadronization and the decay of short–living particles. Hence, the only possibility to reconstruct jets is to use as input the momenta and energies of these particles, measured as well as possible.

On the other hand, in the Monte Carlo all the particle momenta are available. The event can be reconstructed from the beginning: the kinematic of the outgoing partons, after the hard scattering process, is available. Also the momenta of the hadrons produced in the non–perturbative hadronization process can be reconstructed, and, in particular, the variables of the B particles coming from the hadronization of the b quarks, that cannot be measured by the detector due to their short lifetime. All these information can be used to reconstruct jets in the Monte Carlo, at various levels, some closer to theory, other to the jet definition in the data:

- **parton level jets:** are obtained by running on all the partons, as given by the Monte Carlo, before the final state parton shower. This kind of jets comes directly out of the matrix element inserted in the Monte Carlo to describe the hard scattering process. Therefore, they can be easily compared to theoretical calculations, that describe only the perturbative process. Nevertheless, jets reconstructed in this way are difficultly comparable to real jets since no non–perturbative process is taken into account, or the hadron decay. No easy matching is therefore possible between these jets and real data;

- **hadron level jets:** in this case, the inputs to the algorithm are the momenta of all the final state particles. However, a definition of *final state particles* is needed. Usually, the particles considered as *stable* are taken, i.e. those that do not decay or that reach the detector before decaying, usually electrons, muons, neutrinos, pions, protons, strange hadrons, but not charmed or beauty hadrons. Using this definition, the reconstructed jets should be very close to data, the difference being mainly due to the presence of neutrinos and to the non–perfect determination

of the particle energies by the detector. A correction procedure should bring data jet parameters very close to the hadronic jet parameters. Nevertheless, a further step is possible, i.e. the choice of charm and beauty hadrons as final state particles. In this way jets are closer to the theoretical calculations, and also not so far from the jets reconstructed on the data; they do not depend on the decay of the b and c hadrons simulated in the Monte Carlo, but they differ from theory only because they include the quark fragmentation into hadrons. The particles included in the jets are heavy, their masses being compatible to their momenta, and this fact drives to the choice of the E recombination scheme for the k_T algorithm, which gives as output massive jets. In this thesis, Monte Carlo jets are reconstructed at this level.

In Fig. 4.6 the true transverse momentum of the jet, $p_T^{\rm jet}$, (we do not refer anymore to the jet 'transverse energy' since in our definition jets are massive objects) in the sense defined above, is compared to the reconstructed one. First, a matching between the jets at the different levels has to be performed. For each of the three highest transverse momentum jets reconstructed in the event a loop is done over the true jets, reconstructed starting from the B hadrons, and the distance between the two jets in the $\eta - \phi$ plane is evaluated:

$$d_{\eta\phi} = \sqrt{(\eta_{\rm jet}^{\rm rec} - \eta_{\rm jet}^{\rm true})^2 + (\phi_{\rm jet}^{\rm rec} - \phi_{\rm jet}^{\rm true})^2}. \qquad (4.26)$$

The reconstructed jet is associated to the nearest true jet in the $\eta - \phi$ plane, and if their distance is lower than 1 unit the two jets are considered matched. In this case the ratio reported in the plots, $(p_T^{\rm rec} - p_T^{\rm true})/p_T^{\rm true}$, is evaluated. The Monte Carlo sample under analysis has been defined above, and is the same used to investigate the y_{JB} behaviour. Also here no cuts have been applied, except for the trigger requirements and the DIS cuts by SINISTRA.

The calculation is done for two types of reconstructed jets: jets obtained by running on the calorimeter cells, without any correction, shown in the first row of Fig. 4.6, and jets obtained using ZUFOs with all the corrections (in the following, we will refer to such jets with the name 'ZUFO jets', for simplicity) shown in the second row of Fig. 4.6. In the first case, the shift toward negative values of the $(p_T^{\rm rec} - p_T^{\rm true})/p_T^{\rm true}$ ratio is evident. The shift almost disappears when the corrected ZUFOs are used. Also in this case, as it was for the y_{JB} variable, a long tail is visible for high values of the ratio. In the scatter plots, the improvement obtained using the corrected ZUFOs is also evident.

The analysis done on the transverse jet momentum can be repeated for the pseudorapidity, $\eta^{\rm jet}$, of the jets. The procedure is the same, the matched jets being exactly the same used for the transverse momentum analysis. The results are reported in Fig. 4.7, the meaning of the plots is analogous to that of Fig. 4.6. The reconstructed and the true variable are nicely correlated, as can be noticed also in the scatter plot. A tail is present in this case in the low η region, even using the corrected ZUFOs.

The plots will be shown again after all the cuts have been applied to the sample, in order to see what happens to a reduced sample enriched with muon events.

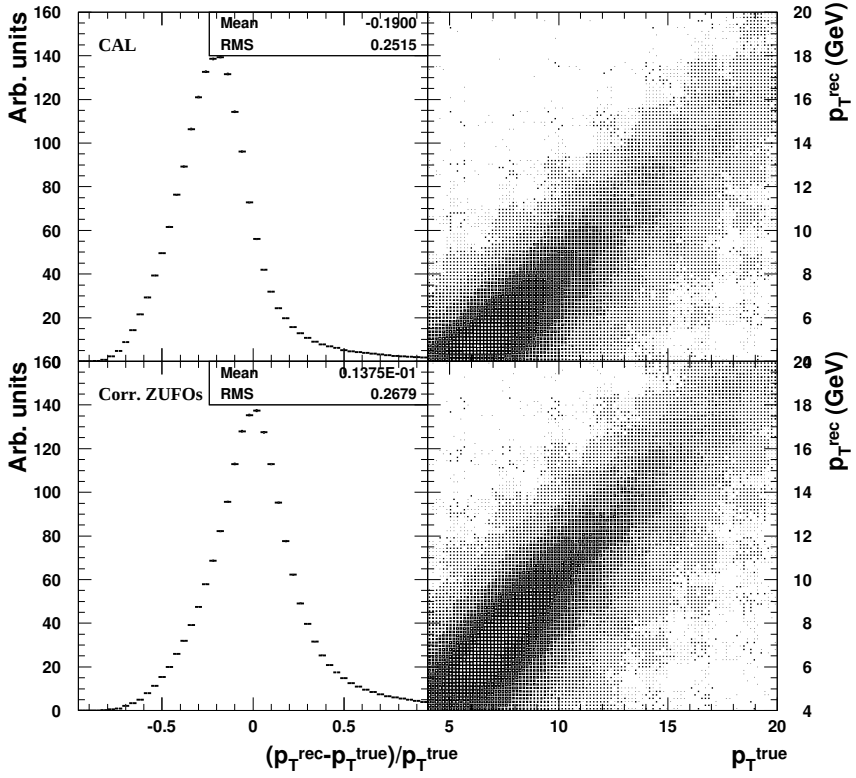

Figure 4.6: $(p_T^{\rm rec} - p_T^{\rm true})/p_T^{\rm true}$ **(left)** *and* $p_T^{\rm rec}$ *vs* $p_T^{\rm true}$ **(right)**, *plotted for two different reconstructions of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $p_T^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* $p_T^{\rm rec}$ *is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1.*

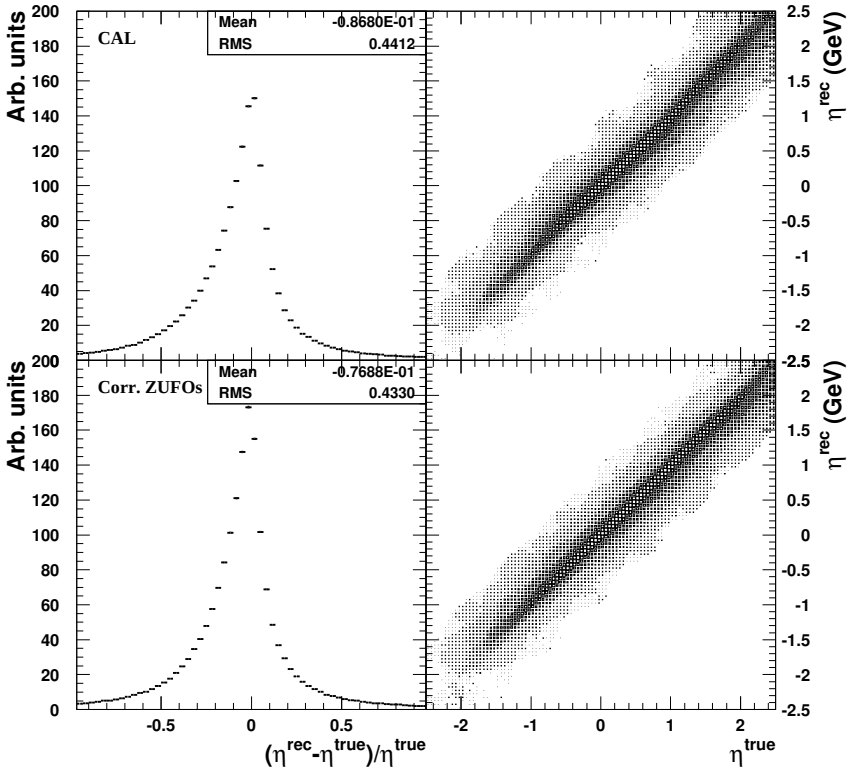

Figure 4.7: $(\eta^{\rm rec} - \eta^{\rm true})/\eta^{\rm true}$ **(left)** *and* $\eta^{\rm rec}$ *vs* $\eta^{\rm true}$ **(right)**, *plotted for two different reconstructions of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $\eta^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* $\eta^{\rm rec}$ *is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1.*

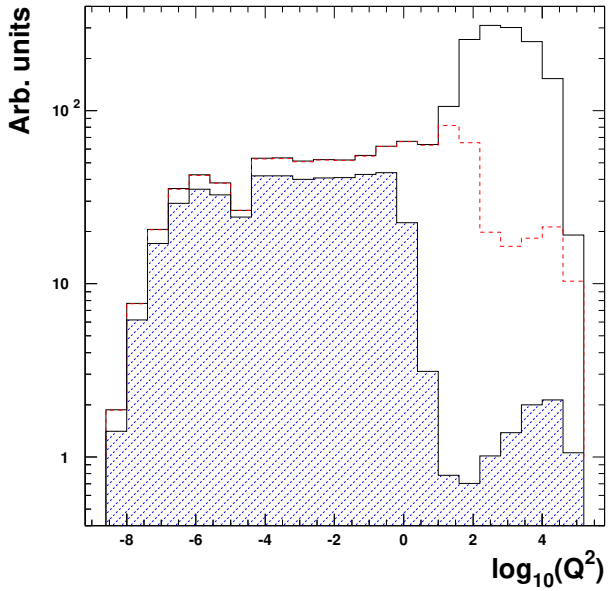

Figure 4.8: *The* Q^2 *distribution of a beauty Monte Carlo sample, obtained mixing the direct, resolved and excitation components in fractions given by their respective inverse luminosities. The empty histogram with full line is the distribution as obtained by applying only the trigger requests; the dashed histogram is obtained by implementing the cuts given by the SINISTRA algorithm, and the diagonally–hatched histogram is the distribution after the cut* $0.2 < y_{JB} < 0.8$ *has been applied.*

4.8 Cuts on the calorimetric variables

4.8.1 Cuts on y_{JB}

After having verified the goodness of the correction to the calorimetric variables, cuts can be applied on them, in order to discriminate signal from background and to select a sample enriched in beauty events.

To this aim, the inelasticity y_{JB} is asked to be in the range:

$$0.2 < y_{JB} < 0.8. \tag{4.27}$$

The effect of this request on y_{JB} is twofold. The rejection of events having low inelasticity values removes the background from proton–beam gas events (see Sect. 2.2.7.2). The cut at high y_{JB} values has the purpose to exclude residual deep inelastic scattering events, not removed by SINISTRA. The effects of the two different cuts can be seen in Fig. 4.8, in which the Q^2 distribution is reported, using as Monte Carlo sample the one defined above for the analysis of the corrections to the calorimetric variables. In

the Figure the Q^2 distribution is reported after the trigger request (empty histogram with full line), after the positron rejection by SINISTRA has been applied (dashed histogram) and finally after the cut on y_{JB} has been implemented (diagonally–hatched histogram). It can be clearly seen in the Figure that the big peak at high Q^2 values almost disappears after the cuts so that the fraction of events having $Q^2 > 1$ GeV^2 is highly suppressed. The residual contamination from events with $Q^2 > 1$ GeV^2 amounts to $\sim 2\%$.

4.8.2 Cuts on the jets

Events are requested to have at least two jets reconstructed using the k_T algorithm in the longitudinal–boost–invariant mode, with the E recombination scheme, and fulfilling the following requirements:

$$p_T^{\rm jet1} > 7 \text{ GeV} \quad ; \quad p_T^{\rm jet2} > 6 \text{ GeV} \tag{4.28}$$

$$|\eta^{\rm jet1,2}| \; < \; 2.5. \tag{4.29}$$

The values of the jet cuts are chosen to ensure that the jets are well reconstructed and understood at both calorimetric and hadronic level. They depend mainly on the trigger cuts (see Sect. 4.4.3) applied on the reconstructed transverse energy. In Fig. 4.9 the efficiency of the EUCELL and KTCLUS algorithm in identifying jets is reported, as a function of the E_T cut applied at the third level trigger. It can be seen in the Figure that for energy greater than 6 GeV the KTCLUS algorithm is perfectly efficient if the TLT cut is $E_T^{\rm TLT} > 4$ GeV, as it was in 1996–97, and also for $E_T^{\rm TLT} > 4.5$ GeV, as it was in 1999–2000 [118]. Since the cone radius of the EUCELL algorithm, $R = 1$, is larger than the effective cone radius of the KTCLUS algorithm, jets reconstructed by EUCELL have, on average, higher transverse energy than those reconstructed by k_T. A cut on the same values is hence harder for the jets found by KTCLUS.

4.9 Cuts on the muons

The cuts on the muons are driven by the attempt to enrich the data sample with beauty events. The b quark, having higher mass than that of the charm and the other quarks, produces in its decay muons having larger momenta than those coming from the decay of lighter quarks. This fact can be observed by analysing the momentum distribution of muons coming from beauty, charm and lighter quarks using Monte Carlo. The obtained distributions of the muon momentum, p^μ, are shown in Fig. 4.10: the spectrum of the muons coming from beauty decay has a long tail toward large p^μ values, whereas muons from charm and light quarks have softer spectra, peaked at smaller values. Light quark events, however, show a tail at high momenta, even if this tail is less pronounced than in the case of beauty quarks. These features of the muon spectra will be used in the following to suppress background events with respect to signal events.

4.9.1 Cuts on muons in the forward region

The reconstruction of muons in the forward region uses the MPMATCH algorithm [108], described in some detail in Sect. 4.3.1. The muon is asked to reach the outer

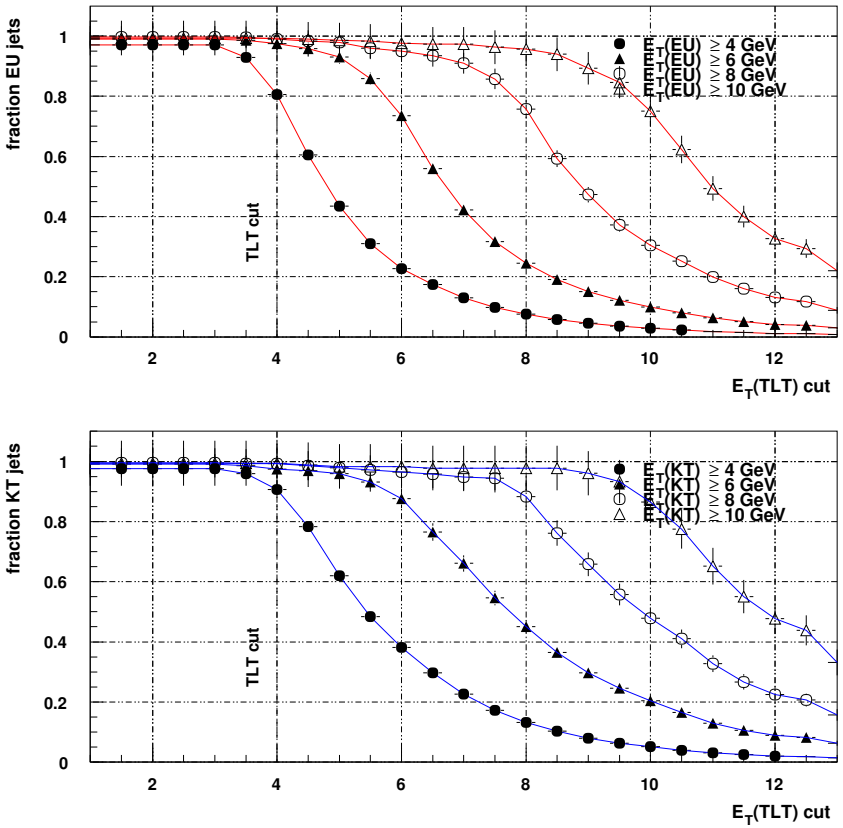

Figure 4.9: *The efficiency of the dijet TLT as a function of the TLT jet E_T threshold, for offline thresholds of 4,6,8 and* 10 GeV. *The dashed line shows the 1996–97 TLT threshold of* 4 *GeV. In the top plot, jets are reconstructed offline using the EUCELL algorithm; in the bottom plot, jets are reconstructed using KTCLUS. From ref. [118].*

FMUON chambers, and the muon momentum to be reconstructed by both the CTD and the muon detector, so that also the momentum can be included in the fitting parameters, when matching the CTD track to the FMUON track.

The muon track is asked to have hits in at least four planes of the forward muon detector, with the first hit in the innermost plane of the detector. The reason to ask at least four hits is that in the FMUON detector a lot of background hits are present, due to the fact that the apparatus is very close to the beam, and therefore photons coming from lepton *bremsstrahlung* and secondary particles emitted by the proton can hit the chambers. Requiring four hits for the FMUON track reduces the probability

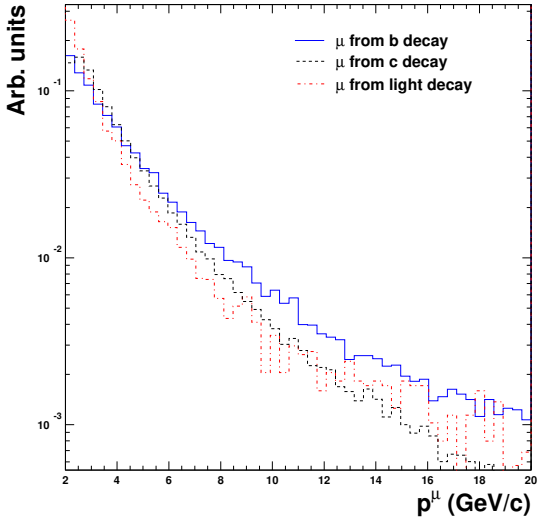

Figure 4.10: *Distribution of the muon momentum, p^μ, for muons coming from beauty (solid line), charm (dashed line) and light quark (dot–dashed line) decays.*

that spurious hits are wrongly included in the track fit. The reason to ask the first hit in the first FMUON plane is similar: the innermost plane is the nearest to the central tracking detector, and therefore the presence of a hit in that plane can increase significantly the quality of the fit.

The resolutions in the muon momentum, p^μ, and pseudorapidity, η^μ, are shown in Fig. 4.11. The Monte Carlo sample has been defined above, and is the same used for the previous studies on the calorimetric variables. The obtained resolution on the pseudorapidity is very good, while the momentum resolution is worse, due to the fact that tracks reconstructed by the Forward Muon Detector have low polar angles, and hence they cross a small number of CTD superlayers: the momentum reconstruction by the drift chamber is therefore worse.

The cuts on the muon momentum are essentially driven by the muon detector performance. Studies performed on data and Monte Carlo have suggested as safe angular region the range [119]:

$$1.48 < \eta^\mu < 2.3 \qquad (11^o < \theta^\mu < 26^o). \tag{4.30}$$

Outside that region, the detector response is not well simulated by the Monte Carlo. Performance considerations drive also the choice of the momentum cut:

$$p^\mu > 4 \text{ GeV} \quad \text{and} \quad p_T^\mu > 1 \text{ GeV} \tag{4.31}$$

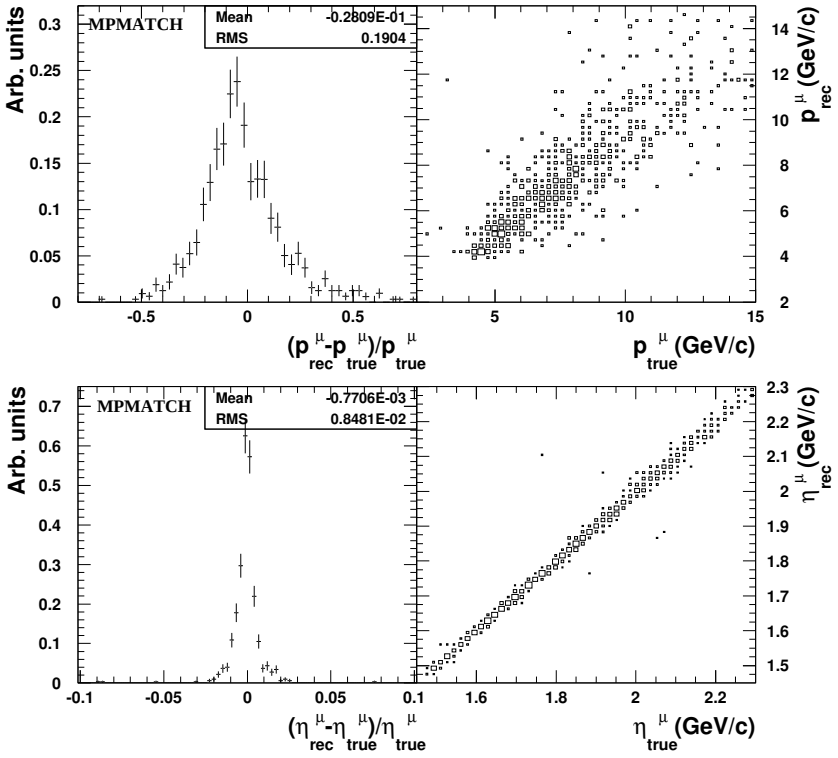

Figure 4.11: *Comparison between the muon variables as reconstructed by the MPMATCH algorithm and the true muon variables. In the first row, left plot, the ratio $(p^\mu_{\rm rec} - p^\mu_{\rm true})/p^\mu_{\rm true}$ is shown, while the right plot shows the correlation between the reconstructed and true muon momentum, p^μ. In the second row, the same is done using the muon pseudorapidity, η^μ: the left plot shows the resolution $(\eta^\mu_{\rm rec} - \eta^\mu_{\rm true})/\eta^\mu_{\rm true}$, and the right plot illustrates the comparison between the reconstructed and the true muon pseudorapidity.*

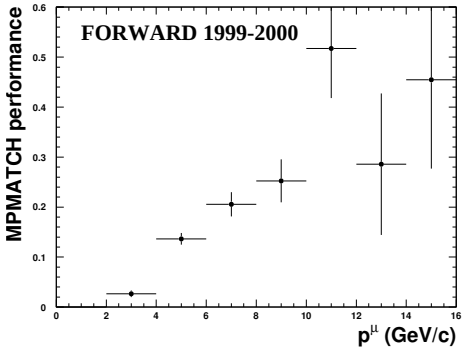

Figure 4.12: *Performance of the MPMATCH algorithm as a function of the muon momentum, p^μ, evaluated on 1999-2000 data sample.*

where the cut on p^μ is due to the fact that the detector starts to see muons at those momenta (see Fig. 4.12), and the cut on p_T is done to avoid the region nearest to the beam pipe, badly simulated by the Monte Carlo. The cut on the probability of the position and momentum (5 d.o.f.) matching is done to reject background from fake muon events:

$$P_{\rm match} > 0.05. \tag{4.32}$$

The performance of the muon matching algorithm for the forward region was investigated in detail [119], as was done for the BREMAT algorithm (see Appendix A). The procedure is rather similar to that used for the barrel and rear muon chambers, described in detail in Sect. A.2, therefore just a few details will be given here.

The performance of the matching algorithm in reconstructing muon tracks has been studied using a sample of two–track events, mainly elastic J/ψ production with subsequent decay into two muons, and dimuon production in $\gamma\gamma$ events. One of the two tracks was asked to trigger the event, then the second was used for the performance determination. The performance is defined as:

$$\text{performance} = \frac{\text{number of reconstructed tracks}}{\text{total number of tracks}} \tag{4.33}$$

where the total number of tracks is given by all the tracks pointing to the FMUON fiducial region defined above (see 4.30), and the numerator is given by the number of the tracks reconstructed by the MPMATCH algorithm, where *reconstructed* means:

- the track has hits in at least four planes of the FMUON detector, the first in the innermost plane;
- the probability of the 5 d.o.f. matching, $P_{\rm match}$, is greater than 0.05.

The track triggering the event is asked to be in the barrel region: the forward region is rejected to avoid biases, the rear is not included to avoid background events, cosmics

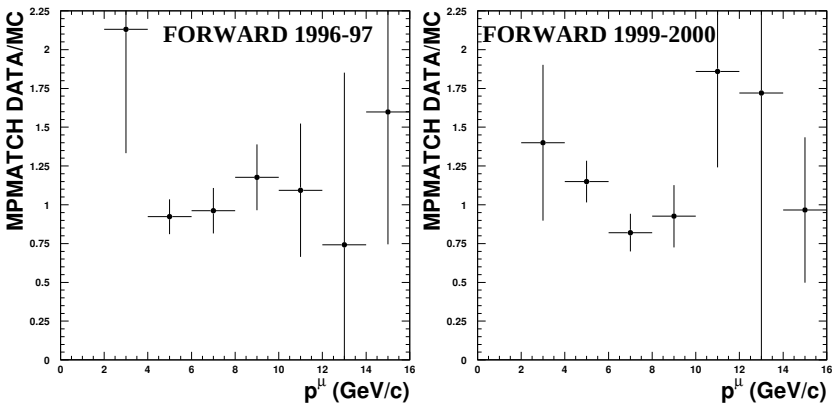

Figure 4.13: *Ratio between the MPMATCH performance measured on the data and on the Monte Carlo, plotted as a function of the muon momentum, p^μ, for 1996–97* **(left)** *and 1999–2000* **(right)** *data–taking. Inefficiencies found in the data are simulated in the Monte Carlo so that the ratio should be nearly equal to 1, as it is.*

or halo–muons. The MPMATCH performance has been evaluated in the same way on the Monte Carlo.

The ratio between the MPMATCH algorithm performance measured on real data and simulated by the Monte Carlo is plotted in Fig. 4.13, as a function of the muon momentum, p^μ. In this case, following a different strategy than that used for the Barrel and Rear Muon Detector (see Appendix A and Sect. 4.9.2), the data were studied in detail before the Monte Carlo was generated, and the inefficiency found in the data are simulated in the Monte Carlo version used for the analysis. Therefore, in the plots shown above, the data/MC ratio should be almost equal to 1. Nevertheless, this number has to be determined with its error, looking at real data, to be used in the following to correct the cross sections obtained using the MPMATCH algorithm. The weighted mean of the data/MC ratios of Fig. 4.13, for $p^\mu > 4$ GeV, gives:

$$\begin{aligned} 1996-1997 \text{ data}: &\quad 0.98 \pm 0.08 \\ 1999-2000 \text{ data}: &\quad 0.98 \pm 0.08 \end{aligned} \tag{4.34}$$

in perfect agreement with each other.

4.9.2 Cuts on muons in the barrel and rear region

The request for a hard muon in the barrel and rear region is done by asking the muon to reach the outer chambers of the detector. In this configuration the particle momentum can also be measured, taking advantage of the magnetic field present in the yoke, positioned between the inner and outer muon chambers (see Sect. 2.2.5). The BREMAT algorithm [109], described in Sect. 4.3.2, is used to reconstruct muons in the barrel and rear region.

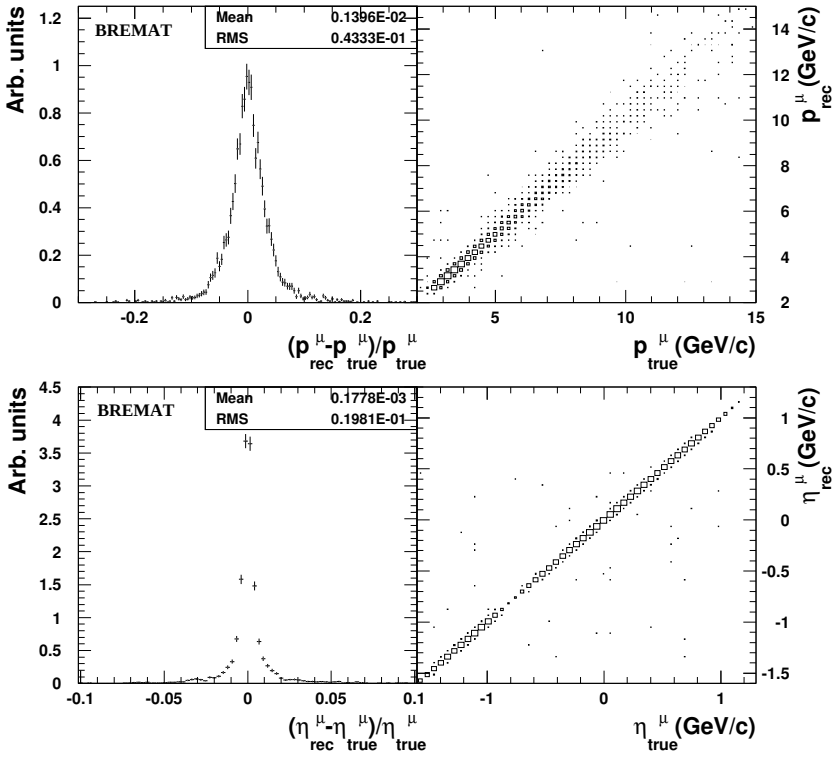

Figure 4.14: *Comparison between the muon variables as reconstructed by the BREMAT algorithm and the true muon variables. In the first row, left plot, the ratio* $(p^{\mu}_{\rm rec} - p^{\mu}_{\rm true})/p^{\mu}_{\rm true}$ *is shown, while the right plot shows the correlation between the reconstructed and true muon momentum,* p^{μ}*. In the second row, the same is done using the muon pseudorapidity,* η^{μ}*: the left plot shows the resolution* $(\eta^{\mu}_{\rm rec} - \eta^{\mu}_{\rm true})/\eta^{\mu}_{\rm true}$*, and the right plot illustrates the comparison between the reconstructed and the true muon pseudorapidity.*

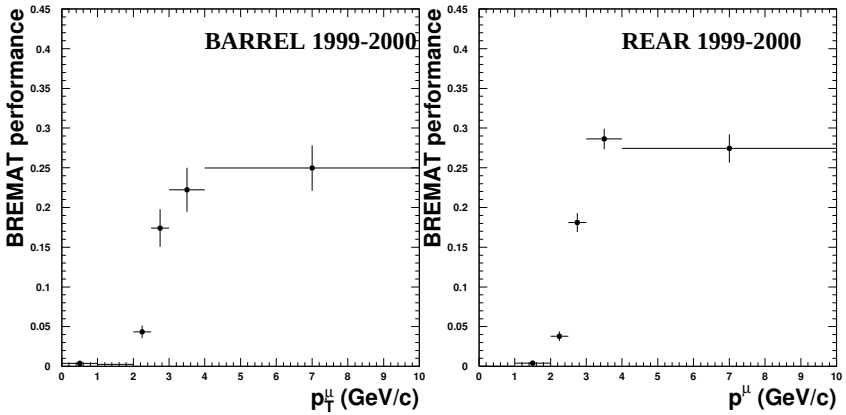

Figure 4.15: *Performance of the barrel* (**left**) *and rear* (**right**) *muon detectors, plotted as a function of the muon transverse momentum,* p^{μ}_{T}*, and of the muon momentum,* p^{μ}*, respectively. The sample used corresponds to the data collected between 1999 and 2000 (see also Appendix A).*

The resolution of the BREMAT algorithm, together with a scatter plot showing the correlation between the true and the reconstructed muon variables, is shown in Fig. 4.14. The Monte Carlo used is the sample described above, used also for the analysis of the resolution of the MPMATCH algorithm. In the Figure it can be noticed that the resolution in momentum and pseudorapidity are very good.

As in the case of the FMUON detector, the cuts on the momentum and angular region of the particle are essentially driven by the detector performance and by the necessity of working in a region well understood and simulated by the Monte Carlo.

The BREMAT performances in finding muon tracks have been extensively investigated in this work (see appendix A); here we will briefly summarize the results. The performances of the algorithm have been evaluated on real data and on Monte Carlo events. The Monte Carlo used is a mixture of elastic J/ψ and $\gamma\gamma$ events, with weights chosen in order to simulate the invariant mass distribution of the data.

The method is similar to that used for the forward region: two–track events are selected, one track is asked to trigger the event and the second is used for the determination of the algorithm performance. The BREMAT performance is defined as in Eq. 4.33, where in this case a *reconstructed* track is a track fulfilling the following requirement:

- the BREMAT algorithm finds a matching by fitting the position and momentum variables, and gives as matching probability a value $P_{\rm match} > 0.05$.

The BREMAT performances are shown in Fig. 4.15 (built as Fig. A.7). As can be seen in the Figure, the barrel detector starts to see particles for transverse momenta $p^{\mu}_{T} > 2.5$ GeV, and the rear detector identifies muons with a sizable efficiency for

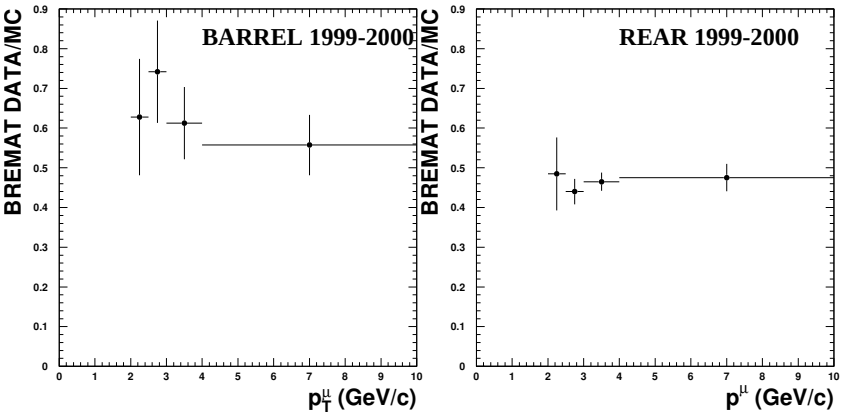

Figure 4.16: *Ratio between the BREMAT performance as measured on real data and on Monte Carlo events, as a function of p_T^μ in the barrel region* **(left)** *and of p^μ in the rear region* **(right)**. *The sample used corresponds to the data collected between 1999 and 2000 (see also Appendix A).*

momenta $p^\mu > 2.5$ GeV. Therefore, the cuts on the muon momenta are defined as follows:

- **BARREL region:** $p_T^\mu > 2.5$ GeV;
- **REAR region:** $p^\mu > 2.5$ GeV.

The cuts on the pseudorapidity are chosen to select a region well described by the Monte Carlo and where the performances of the algorithm are well understood:

- **BARREL region:** $-0.9 < \eta^\mu < 1.3$, $(30^o < \theta^\mu < 135^o)$;
- **REAR region:** $-1.6 < \eta^\mu < -0.9$, $(135^o < \theta^\mu < 157^o)$.

In the case of the Barrel and Rear Muon Detector the inefficiencies are not simulated in the Monte Carlo, but correction factors have been evaluated to be applied to the cross sections to keep into account the detector inefficiencies.

The ratio between the BREMAT performance as measured on real data and on Monte Carlo events is shown in Fig. 4.16 for the 1999-2000 sample, and in Sect. A.5. It is clear from the Figures that, in the regions selected by our cuts, the data/MC ratio is rather flat (the situation is less clear for the barrel region since the errors are larger). Therefore the correction factors to be applied to the data are evaluated as the weighted mean of the ratios in the region of interest, $p_T^\mu > 2.5$ GeV in the barrel region and $p^\mu > 2.5$ GeV in the rear region. The obtained results are shown in Table 4.2; the errors on the correction factors are given by the quadratic sum of the statistical uncertainties and of systematic effects illustrated in detail in Appendix A.

BARREL REGION		REAR REGION	
1996–97	1999–2000	1996–97	1999–2000
0.71 ± 0.11	0.61 ± 0.09	$0.54^{+0.03}_{-0.07}$	0.46 ± 0.05

Table 4.2: *Summary of the correction factors given by the ratio of the BREMAT performance evaluated on the data and on the Monte Carlo. The factors are evaluated in the momentum region of interest, i.e. for $p_T^\mu > 2.5$ GeV in the barrel region and for $p^\mu > 2.5$ GeV in the rear region.*

4.10 Summary of the offline cuts

The cuts applied to the data in order to select a clean photoproduction event sample, with at least two jets and a muon in the final state, coming from the process

$$e^+ p \to e^+ b \bar{b} X \to e^+ \mu^\pm \text{ dijet } X \tag{4.35}$$

are summarized in the following.
First of all, deep inelastic scattering background has been rejected by removing events in which a candidate positron was found, having energy E_e and inelasticity y_e:

$$E_e > 5 \text{ GeV} \qquad y_e < 0.9. \tag{4.36}$$

The following cut has been applied to the inelasticity, y_{JB}, of the event, in order to further reduce DIS and proton–beam gas background:

$$0.2 < \quad y \quad < 0.8. \tag{4.37}$$

In this way the data sample has been limited to the region:

$$Q^2 \quad < \quad 1 \text{ GeV}^2. \tag{4.38}$$

Then, cuts have been applied to the jets to select a region where they are well understood and reconstructed, and to enrich the data sample in beauty events:

$$\begin{aligned} p_T^{\text{jet1}} > 7 \text{ GeV} \quad ; \quad & p_T^{\text{jet2}} > 6 \text{ GeV} \\ |\eta_T^{\text{jet1,2}}| \quad < \quad & 2.5 \text{ GeV}. \end{aligned} \tag{4.39}$$

The cuts on the muons were also driven by the necessity to enrich the data sample in signal events, and to select a region where the muon detectors have good performance and are well understood:

- **FORWARD region:** the muon has to have hits in at least four planes of the FMUON detector, the first in the innermost plane, and to fulfill the following requirements:

$$\begin{aligned} p^\mu > 4 \text{ GeV} \quad ; \quad & p_T^\mu > 1 \text{ GeV} \\ 1.48 < \eta^\mu < 2.3 \quad & (11^o < \theta^\mu < 26^o) \\ P_{\text{match}}^{\text{MPMATCH}} \quad > \quad & 0.05; \end{aligned} \tag{4.40}$$

- **BARREL region:**

$$\begin{aligned} p_T^{\mu} &> 2.5 \text{ GeV} \\ -0.9 < \eta^{\mu} < 1.3 & \quad (30^o < \theta^{\mu} < 135^o) \\ P_{\text{match}}^{\text{BREMAT}} &> 0.05; \end{aligned} \tag{4.41}$$

- **REAR region:**

$$\begin{aligned} p^{\mu} &> 2.5 \text{ GeV} \\ -1.6 < \eta^{\mu} < -0.9 & \quad (135^o < \theta^{\mu} < 157^o) \\ P_{\text{match}}^{\text{BREMAT}} &> 0.05. \end{aligned} \tag{4.42}$$

This completes our data selection. After all the cuts, the data sample to be analysed amounts to 3334 events, 1941 collected during 1999–2000 data–taking, 1393 in 1996–97.

The plots showing the resolution on the inelasticity and on the jets variables and the comparison between reconstructed and true variables, after all the cuts of the analysis have been applied, are shown in Figs. 4.17, 4.18 and 4.19. The meaning of the Figures is analogous to that of Figs. 4.5, 4.6 and 4.7, the only difference is that now, having applied all the cuts described above, the sample is enriched in muon events.

The effects of the correction to the y_{JB}^{rec} variable remain evident, the shift toward negative values of the ratio $(y_{JB}^{\text{rec}} - y^{\text{true}})/y^{\text{true}}$ being strongly reduced by using standard ZUFOs with respect to calorimeter cells, and corrected ZUFOs with respect to standard ZUFOs. The tail at large values of the ratio is less pronounced. For what regards the jet transverse momentum, p_T^{jet}, the results are highly improved by using corrected ZUFOs rather than calorimeter cells. Also here the tail at high values of the ratio $(p_T^{\text{rec}} - p_T^{\text{true}})/p_T^{\text{true}}$ is reduced. Regarding the jet pseudorapidity, η^{jet}, the reconstructed and the true variable are still well correlated and the tail at low values is slighlty reduced.

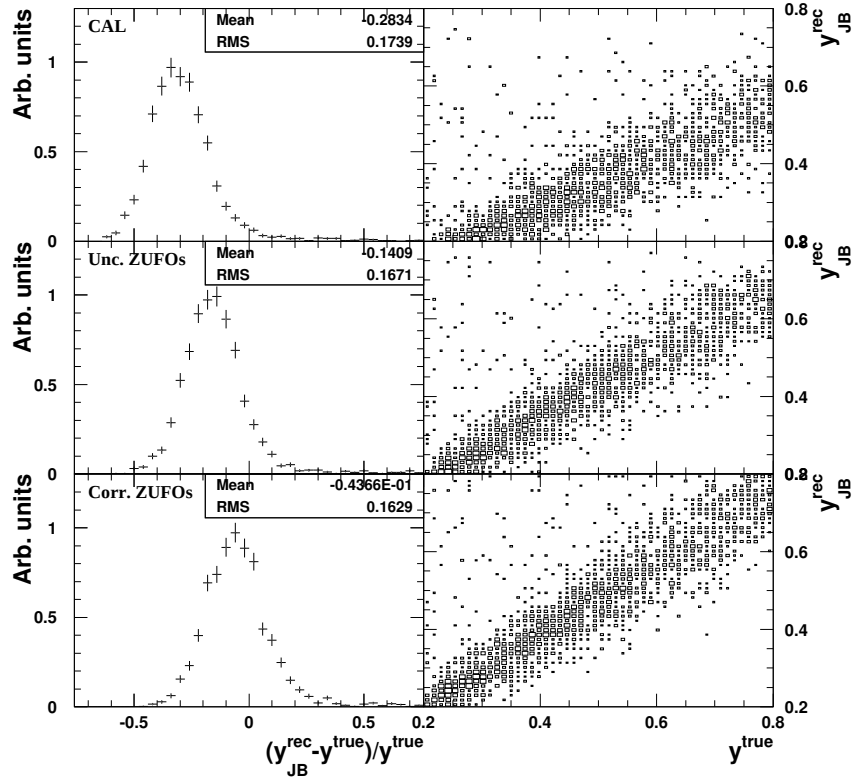

Figure 4.17: $(y_{JB}^{\text{rec}} - y^{\text{true}})/y^{\text{true}}$ **(left)** *and* y_{JB}^{rec} *vs* y^{true} **(right)**, *plotted for various estimations of the inelasticity of the event. In the first row, the resolution* **(left)** *and the comparison of the reconstructed to the true variable* **(right)** *are done calculating* y_{JB}^{rec} *by looping on the calorimeter cells; in the second row,* y_{JB} *is calculated taking as inputs the standard ZUFOs, without any correction; in the third row, all the corrections described in the text have been applied to the ZUFOs. The sample is obtained by mixing the beauty Monte Carlo samples (direct, resolved and excitations) in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo, in order to obtain a muon–enriched sample.*

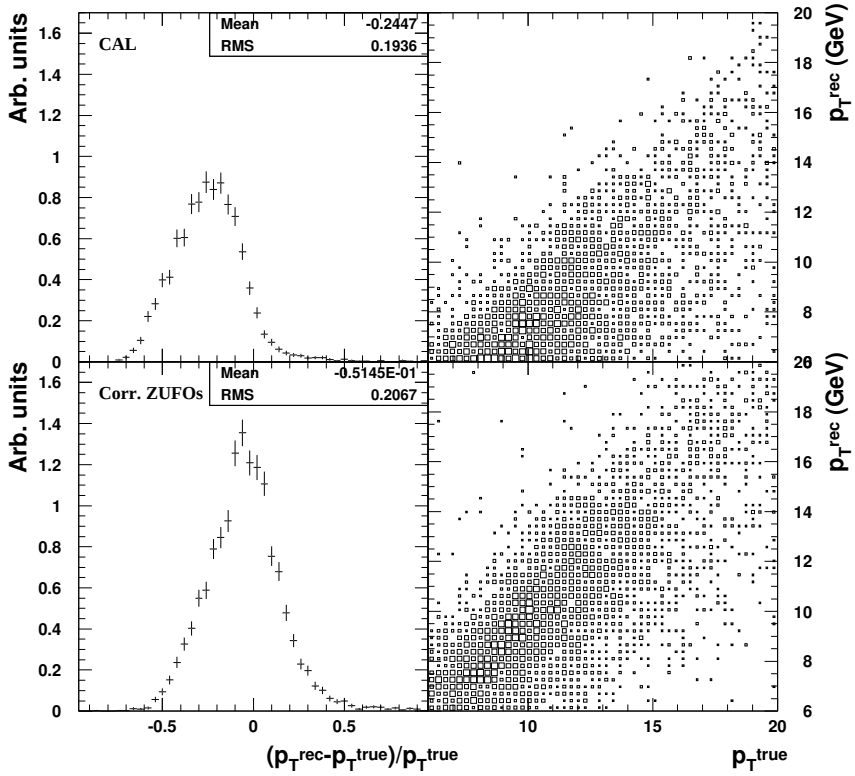

Figure 4.18: $(p_T^{\rm rec} - p_T^{\rm true})/p_T^{\rm true}$ **(left)** *and* $p_T^{\rm rec}$ *vs* $p_T^{\rm true}$ **(right)**, *plotted for two different reconstruction of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $p_T^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* $p_T^{\rm rec}$ *is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo, in order to obtain a muon–enriched sample.*

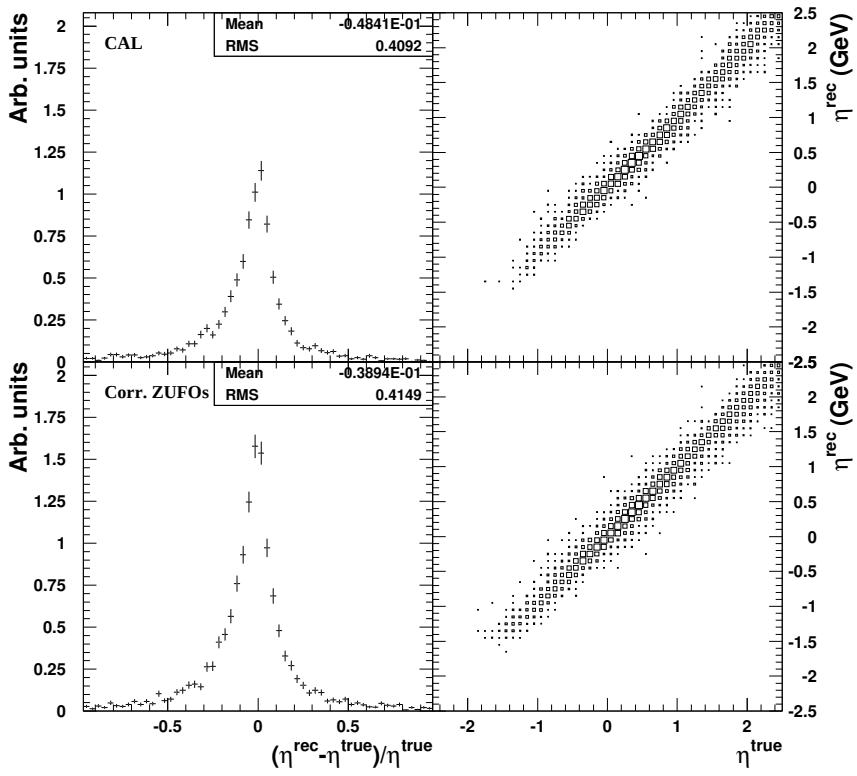

Figure 4.19: $(\eta^{\rm rec} - \eta^{\rm true})/\eta^{\rm true}$ **(left)** *and* $\eta^{\rm rec}$ *vs* $\eta^{\rm true}$ **(right)**, *plotted for two different reconstruction of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $\eta^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* $\eta^{\rm rec}$ *is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo, in order to obtain a muon–enriched sample.*

Chapter 5

Selection of beauty events

At this stage of the analysis a data sample has been selected in which at least two jets and a muon are present in the final state. This sample contains both signal (beauty) and background events. In this Chapter the various sources of background will be described, and the method used to recognize the signal will be illustrated in detail. At the end of the Chapter the fraction of beauty events present in the data will be determined.

5.1 μ–jet association

A further cut has to be applied on the selected muon plus dijet data sample in order to choose the events in which the muon is included in one of the jets. Muons coming from the decay of beauty quarks, in fact, are part of the jet resulting from the hadronization and the subsequent decay of the b quark.

The multiplicity of muons in events passing all the cuts is shown in Fig. 5.1 (left plot). It is clear that the number of events in which two good muons are found is very small. Fig. 5.1 (right plot) shows the multiplicity of good jets. The cuts to be applied to the first two jets have already been described (see Eqs. 4.28 and 4.29). If more than two jets are found in an event, the third and the fourth jet are considered good for the matching if they fulfill the same requirements imposed to the second jet:

$$p_T^{\rm jet} > 6 \text{ GeV} \quad ; \quad |\eta^{\rm jet}| < 2.5. \tag{5.1}$$

No more than four jets in an event are selected for the matching with the muon. In the plot it can be noticed that in most of the cases the number of good jets in an event is two; the number of events having more than four jets passing all the cuts is negligible (less than 0.1%).

To perform the muon–jet association a particular feature of the k_T algorithm is used. The algorithm, in fact, gives as output a pointer connecting each particle to the jet in which it is included. Therefore, the reconstruction of muons and jets and the subsequent association proceed as follows:

- **muons** are reconstructed by the muon algorithms, MPMATCH and BREMAT, so that a track in the CTD is matched to a track in the muon detectors: the track corresponding to a muon is therefore clearly identified in the CTD;

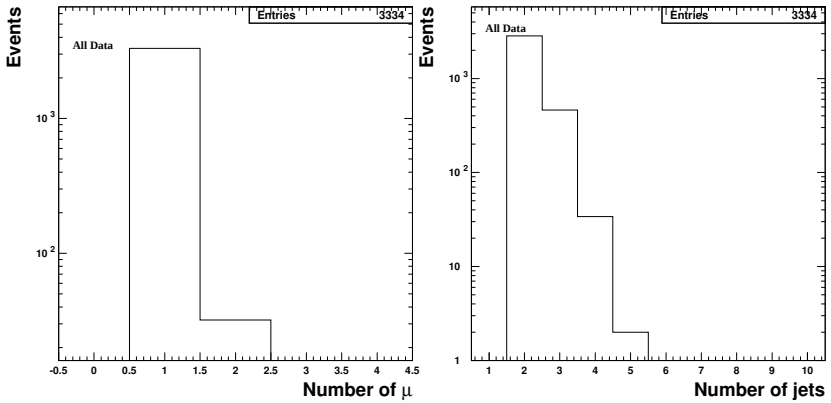

Figure 5.1: *Muon (***left***) and jet (***right***) multiplicity in 1996–2000 data sample, for events passing all the cuts, before the muon–jet association. The cuts on the third and fourth jets are the same imposed on the second (see text). No more than four jets are selected for the matching with the muon.*

- **ZUFOs** are reconstructed starting from energy deposits in the calorimeter and from tracks in the CTD. For each identified particle it is possible to find the ZUFO containing it;
- **jets** are reconstructed by running the k_T algorithm on the ZUFOs. For each jet, the list of the ZUFOs belonging to it is available.

Therefore, for each muon we know the ZUFO in which it is contained, and for each ZUFO we know the jet to which it belongs. A muon–jet association is done in these terms, asking the muon to be included in the jet by the k_T algorithm. The k_T algorithm reconstructs jets using criteria of relative transverse momenta between the particles, and this should be a safer way than a criterion of angular distance to include the right particles in the jets.

The number of events passing also this additional request regarding the muon–jet association is reported in Table 5.1. The total number of events available, summing 1996–97 and 1999–2000 data–taking periods, amounts to 2845, to be compared to the 3334 events selected without this additional request (see Sect. 4.10). The final event sample consists of the 2845 events having the muon associated to the jet.

5.2 Analysis of the background

The selected data sample contains signal events, in which a $b\bar{b}$ couple is produced, together with background events. The way in which background is rejected depends strongly on the kind of processes involved. A complete inclusion and understanding

	Barrel–Rear	Forward	Total
1996–1997	891	304	1195
1999–2000	1194	456	1650

Table 5.1: *Number of events passing all the cuts, including the muon–jet association, shown separately in the barrel–rear and forward region, for the two data–taking periods.*

of all the background sources are of particular importance for the correct background estimation, given the method that will be used for its subtraction.

A possible source of background is the misidentification of light hadrons as muons by the muon chambers. Such hadrons, mainly kaons and pions, have a small probability to reach the muon chambers and to be reconstructed as muons. Nevertheless, since they are produced in large number in ZEUS events, some of them can actually give a wrong signal in the muon chambers. In this case we talk about *fake muons* or *mistags*.

Another important contribution to the background is given by events in which a quark pair is produced, other than a $b\bar{b}$ couple, i.e. a $c\bar{c}$ pair, with the subsequent semi–leptonic decay of one quark in the muon channel. The experimental signature of these events is very similar to that of beauty events: at least two jets are present in the final state, and a muon is included in one of the jets. In this case, the discrimination between signal and background will be based on the large mass difference between the beauty and the other quarks, as it will be shown later in this Chapter.

5.2.1 Determination of background from light quark events

In this Section the determination and treatment of background coming from light quark events will be described in some detail. This kind of background comes from dijet events with a signal in the muon chambers given by a particle other than a muon coming from beauty or charm decay. The particle hitting the muon chambers can be a hadron (in this case we talk about fake muons or mistags) or a secondary muon, and can come from different processes:

- *punch–through:* hadrons passing through the calorimeter can interact and give way to secondary particles that hit the muon detectors; this source of background is reduced by requiring a good position and momentum matching between the segment in the muon detector and the CTD track;
- *sail–through:* in some cases, hadrons can pass through the whole detector without being stopped, and then reach the muon chambers being mistagged as muons.
- *decay in flight:* muons can come from particle decays, like $\pi^{\pm} \to \mu^{\pm}\nu$, or $K^{\pm} \to \mu^{\pm}\nu$. Such muons typically have low momenta and therefore the request for a hard muon reduces this kind of background. Moreover, these events are rejected also by requiring the muon to come from the interaction vertex.

The effect of this background has been quantified by using a large sample of Monte Carlo events, described in detail in Sect. 3.5.3. Each event of this photoproduction sample was asked to have at least two jets in the final state and one track in the acceptance region of the muon chambers. The production of heavy flavours was forbidden.

The sample was treated in the same way as the beauty Monte Carlo and as the data in order to determine the number of light quark events passing all the cuts. In this way, also the distributions of the kinematic variables of interest could be determined for background events included in the final data sample. From the analysis of these distributions some indications on how to distinguish between signal and background can be obtained, as it will be shown in Sect. 5.3.

To quantify the fraction of light quark events in the final data sample the same procedure was used as for the charm–induced background. This procedure will be illustrated in detail in Sect. 5.3.

5.2.2 Determination of background from charm events

The characteristics of events in which a $c\bar{c}$ pair is produced, with a subsequent decay of one of the quarks in the muon channel, are very similar to those of $b\bar{b}$ events. The cross section of $c\bar{c}$ photoproduction is expected to be two orders of magnitude larger than that of beauty. Therefore, even if muons coming from charm decay have typically lower momenta than those coming from beauty decay (see Fig. 4.10), and are thus rejected by the request of a hard muon, they represent a significant source of background.

The main difference between the two processes, beauty and charm photoproduction, is represented by the mass of the produced quarks. The mass of the b quark is in fact around three times larger than the mass of the charm quark, and therefore the dynamics of the decay products will be different in the two processes. In particular, we expect the relative transverse momenta of the particles coming from beauty decay to be larger than those of the charm decay products. Referring to muons and jets, we can consider the transverse momentum of the muon relative to the axis of the jet in which it is included, $p_T^{\rm rel}$ (see also Fig. 5.2):

$$p_T^{\rm rel} = \frac{\vec{p}_\mu \cdot \vec{p}_{\rm jet}}{|\vec{p}_\mu|\ |\vec{p}_{\rm jet}|}. \qquad (5.2)$$

From what was pointed out above, the $p_T^{\rm rel}$ distribution is expected to be harder in $b\bar{b}$ events than in events in which a $c\bar{c}$ or lighter quark pair is produced. This fact can be verified by looking at the $p_T^{\rm rel}$ distributions in the Monte Carlo. A large sample of Monte Carlo events in which a $c\bar{c}$ couple is produced was generated (see Sect. 3.5.2). The $p_T^{\rm rel}$ distribution was plotted using charm, light and beauty MC samples, selecting events passing all the cuts, in order to determine the shape of the $p_T^{\rm rel}$ distribution contributing to the final data sample. The results are reported in Fig. 5.3.

In the Figure it can be noticed that the $p_T^{\rm rel}$ distribution is harder for beauty events. Charm events have a much steeper falling $p_T^{\rm rel}$ shape; the same is true for light quark events. Looking at the data distribution, it is clear that its shape cannot be described by summing only the charm and light–flavour components, but a contribution from beauty events is needed. The relative contributions to the data coming from the different MC components can be obtained by fitting the shape of the $p_T^{\rm rel}$ distribution of the data with the beauty, charm and light quark components provided by the Monte Carlo.

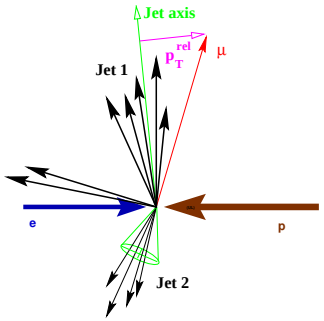

Figure 5.2: *Graphical representation of the meaning of the $p_T^{\rm rel}$ variable.*

5.3 Beauty fraction in the data: $p_T^{\rm rel}$ fit

The determination of the beauty fraction in the data relies on the difference between the shapes of the $p_T^{\rm rel}$ distributions of beauty, charm and light quark components. The $p_T^{\rm rel}$ shape of the data results from the sum of various contributions, simulated by the different Monte Carlo components. A fit on the $p_T^{\rm rel}$ variable can distinguish between these contributions.

First of all, the $p_T^{\rm rel}$ distributions of beauty, charm and light flavours have to be built. Each of these components (b, c and light quarks, l), in fact, are the sum of different sub–processes, simulated using Monte Carlo. As an example (see Sect. 3.5.1) the beauty contribution to the data comes from a mixture of four different sub–processes in which b is produced: direct photoproduction, resolved photoproduction, beauty excitation in the photon and in the proton. These sub–processes have to be mixed in a proper way in order to obtain a single $p_T^{\rm rel}$ distribution for the beauty component. The same is true for charm, whose production mechanisms are identical to those of b (direct and resolved PHP, charm excitation in p or γ). Finally, the light quark component is the sum of two sub–processes, direct and resolved photoproduction.

For each component the various contributing sub–processes have been summed up, using as weights the inverse of the integrated luminosity generated for each sub–process:

$$N_i^f = \frac{\sum_{j\in\{f\ \rm proc\}} \mathcal{L}_j^{-1} N_i^j}{\sum_{j\in\{f\ \rm proc\}} \mathcal{L}_j^{-1}} \qquad (5.3)$$

where N_i^f is the number of events in the ith bin of the $p_T^{\rm rel}$ distribution of flavour f, where f can refer to beauty, charm or light flavours. $\mathcal{L}_j$ is the integrated luminosity generated for each sub–process, N_i^j is the ith bin content of the $p_T^{\rm rel}$ distribution relative to the sub–process j. In this way, each sub–process is taken into account considering both its cross section and the statistics available. The resulting plots, with the binning used for the fit but normalized to one, are shown in Fig. 5.3.

Once the three $p_T^{\rm rel}$ distributions are ready, the fit to the data can be performed. As

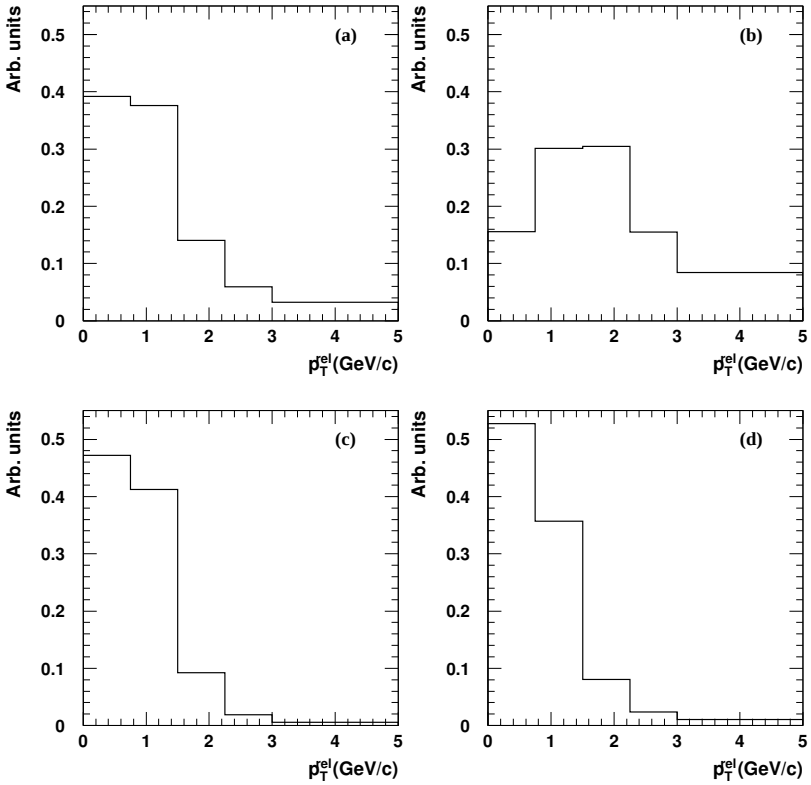

Figure 5.3: $p_T^{\rm rel}$ *distributions for data* **(a)**, *beauty* **(b)**, *charm* **(c)** *and light flavour* **(d)** *components.*

a first attempt, the $p_T^{\rm rel}$ distribution of the data was fitted using the following function:

$$N_i^{\rm fit} = \frac{f_b N_i^b + f_c N_i^c + f_l N_i^l}{f_b + f_c + f_l} \tag{5.4}$$

where N_i^b, N_i^c, N_i^l and $N_i^{\rm fit}$ are, respectively, the contents of the ith bin in the beauty, charm, light flavour and fitted $p_T^{\rm rel}$ distributions. f_b, f_c and f_l, the free parameters of the fit, can be interpreted as the relative fractions (percentages) of beauty, charm and light flavour events in the data, provided they sum up to 1 (100). A χ^2 function was built, using a FORTRAN program, to perform the fit:

$$\chi^2 = \frac{\frac{\sum_i (N_i^{\rm data} - N_i^{\rm fit})^2}{\sigma_i^2}}{d.o.f.}, \tag{5.5}$$

where $N_i^{\rm data}$ is the ith bin content of the $p_T^{\rm rel}$ distribution of the data; σ_i is the error on the ith bin content of the $p_T^{\rm rel}$ distributions. σ_i is the result of different contributions, added in quadrature:

- the error on the data, given by the square root of the bin content, if the bin content is greater than 5, otherwise by the value given by the Poisson statistics;
- the error on the Monte Carlo, given by the square root of the bin content, taking properly into account the applied weights. This error is of particular importance for the light flavour component, the one having the lowest statistics.

The FORTRAN program varies the free parameters of the fit, f_b, f_c and f_l, in order to find the values giving the lowest χ^2. For each choice of the triplet (f_b, f_c, f_l) χ^2 is evaluated, and the triplet corresponding to the minimum value is chosen. In a first loop, the percentages are varied in steps of 1 unit, until a minimum χ^2 value is found. In a second step, the percentages are varied in steps of 0.1 units in a restricted interval around the previously determined minimum. No more loops are performed, considering the error affecting the determination of the beauty percentage coming from the fit.

The fit fails in distinguishing between the charm and light flavour component. As it was shown above (Fig. 5.3) the two $p_T^{\rm rel}$ distributions are rather similar to each other and therefore cannot be distinguished by a fit. Hence, the choice was made to use as input to the fit two MC distributions instead of three. The charm and light flavour components were summed up using the technique of Eq. 5.3. A single charm plus light flavour $p_T^{\rm rel}$ distribution was obtained, its bin content being:

$$N_i^{cl} = \frac{\sum_{j\in\{c,l\ \rm procs\}} \mathcal{L}_j^{-1} N_i^j}{\sum_{j\in\{c,l\ \rm procs\}} \mathcal{L}_j^{-1}} \tag{5.6}$$

where j now runs over all the six charm+light flavour sub–processes, and N_i^j is the number of events in the ith bin of the jth $p_T^{\rm rel}$ distributions contributing to the final $c+l$ component.

The formulas reported above were modified keeping into account the reduced number of free parameters of the fit. The χ^2 definition in Eq. 5.5 is still valid, but now the content of the ith bin of the fitted distribution is given by:

$$N_i^{\rm fit} = \frac{f_b N_i^b + f_{cl} N_i^{cl}}{f_b + f_{cl}}. \tag{5.7}$$

5.4 The x_γ distribution

A quantity of particular interest for our analysis is the fraction of the photon momentum taking part to the hard interaction, referred to as x_γ. From the experimental point of view, x_γ can be determined using the variables of the two highest transverse energy jets in the event. The $x_\gamma^{\rm MEAS}$ variable can be defined:

$$x_\gamma^{\rm MEAS} = \frac{(E-p_z)^{\rm jet1} + (E-p_z)^{\rm jet2}}{E-p_z}, \tag{5.8}$$

this definition being valid for massive and massless jets. When the jets are reconstructed as massless objects, i.e. when $(E-p_z)^{\rm jet} \simeq p(1-\cos\theta^{\rm jet})$, Eq. 5.8 is equivalent to:

$$x_\gamma^{\rm OBS} = \frac{E_T^{\rm jet1} e^{-\eta^{\rm jet1}} + E_T^{\rm jet2} e^{-\eta^{\rm jet2}}}{2yE_e}, \tag{5.9}$$

which we have already defined in Eq. 1.32. In our case, since the decision has been taken to use massive jets the definition of x_γ in Eq. 5.8 has to be used.

Particular emphasis is given to the x_γ variable since it is used to distinguish between direct and resolved photon processes. We talk about *direct* photon processes if $x_\gamma > 0.75$, whereas the *resolved* photon processes are defined for $x_\gamma < 0.75$.

A correct description of the $x_\gamma^{\rm MEAS}$ variable by the Monte Carlo means that the relative fractions of direct and resolved processes are well implemented in the simulation. A good description of the $x_\gamma^{\rm MEAS}$ distribution can be obtained by reweighting the different MC components by changing the ratio between the direct and resolved cross sections of the various contributing processes. In order to determine how these cross sections have to be modified, the $p_T^{\rm rel}$ distribution of the signal and background components were fitted to the data in the two $x_\gamma^{\rm MEAS}$ regions, $x_\gamma^{\rm MEAS} < 0.75$ and $x_\gamma^{\rm MEAS} > 0.75$, using the method described in Sect. 5.3. The obtained percentages are reported in Table 5.2.

Years	b fractions	
	$x_\gamma^{\rm MEAS} < 0.75$	$x_\gamma^{\rm MEAS} > 0.75$
1996–97	$(24.3 \pm 4.7)\%$	$(22.1 \pm 3.8)\%$
1999–2000	$(25.8 \pm 4.2)\%$	$(31.3 \pm 3.4)\%$

Table 5.2: *Beauty percentages in the data, for $x_\gamma^{\rm MEAS} < 0.75$ (resolved photon processes) and $x_\gamma^{\rm MEAS} > 0.75$ (direct γ processes), extracted by fitting the $p_T^{\rm rel}$ distributions in the two regions.*

The beauty cross sections have then been reweighted with respect to those of the background, in order to better simulate the relative fractions of b and $c+l$ components in the two $x_\gamma^{\rm MEAS}$ regions. The reweighting has been done for the direct and resolved cross sections separately. Since the different Monte Carlo contributions are weighted with their inverse luminosities, a x factor has been applied to the beauty cross sections so that:

$$\frac{x^{\rm dir(res)} \sum_{i\in b\ {\rm dir(res)\ procs}} \mathcal{L}_i^{-1}}{\sum_{i\in cl\ {\rm dir(res)\ procs}} \mathcal{L}_i^{-1}} = \frac{f_b^{\rm dir(res)}}{f_{cl}^{\rm dir(res)}} \tag{5.10}$$

where $f_b^{\rm dir(res)}$ and $f_{cl}^{\rm dir(res)}$ are the fractions of beauty and background events in the $x_\gamma^{\rm MEAS} > 0.75$ ($x_\gamma^{\rm MEAS} < 0.75$) region, reported in Table 5.2; i runs over the direct (resolved) processes of each of the beauty and background components. By multiplying the beauty direct (resolved) cross sections by $x^{\rm dir}$ ($x^{\rm res}$) the ratio between signal and background events in the direct (resolved) photon region is simulated in the correct way.

The obtained results have then been used to perform a fit to the $x_\gamma^{\rm MEAS}$ variable. A $x_\gamma^{\rm MEAS}$ distribution of resolved photon processes has been obtained in the following way: first, two $x_\gamma^{\rm MEAS}$ distributions have been built, one summing up the contributions of all the beauty resolved sub–processes, and the other by summing all the charm and light flavours resolved sub–processes. The usual formula was used:

$$N_i^{b,cl\ {\rm res}} = \frac{\sum_{j\in b,cl\ {\rm res\ procs}} \mathcal{L}_j^{-1} N_i^j}{\sum_{j\in b,cl\ {\rm res\ procs}} \mathcal{L}_j^{-1}} \tag{5.11}$$

where now i runs over all the resolved photon sub–processes, for the beauty or for the charm and light flavours components. Then, the contributions from resolved beauty and background components have been mixed according to the $x^{\rm res}$ factor obtained above. In this way a single distribution from all the beauty, charm and light flavour resolved sub–process has been obtained. With a similar procedure, a single $x_\gamma^{\rm MEAS}$ distribution including all the contributions from beauty and background direct sub–process has been built.

A fit has then been performed of the direct and resolved photon Monte Carlo contributions to the data $x_\gamma^{\rm MEAS}$ shape, in order to estimate the relative fractions of direct and resolved processes in the data. The results are reported in Fig. 5.4, and the obtained fractions of direct processes are:

- **1996–97:** $f_d = (0.45 \pm 0.03)\%$;
- **1999–2000:** $f_d = (0.51 \pm 0.03)\%$.

Once the relative fractions of direct and resolved processes in the data are known, the weights to be applied to all the direct and resolved processes can be recalculated in order to give:

$$\frac{w_d}{w_r} = \frac{f_d}{f_r} \tag{5.12}$$

where $f_r = (1-f_d)$ is the fraction of resolved γ events in the data. We decide to apply a multiplicative factor, y, only to the resolved part in order to reproduce the f_d/f_r ratio. Eq. 5.12 can be written as:

$$\frac{x^{\rm dir} \sum_{i\in b\ {\rm dir}} \mathcal{L}_i^{-1} + \sum_{i\in cl\ {\rm dir}} \mathcal{L}_i^{-1}}{y \cdot (x^{\rm res} \sum_{i\in b\ {\rm res}} \mathcal{L}_i^{-1} + \sum_{i\in cl\ {\rm res}} \mathcal{L}_i^{-1})} = \frac{f_d}{f_r} \tag{5.13}$$

from which the y factor can be determined, to be applied to the cross sections of the resolved processes of each component.

At the end, four weights are found, to be applied to the signal and background direct and resolved components. Since only the ratio of the weights is important for our aims, one of them can be set to one: in our procedure, the cross section of the

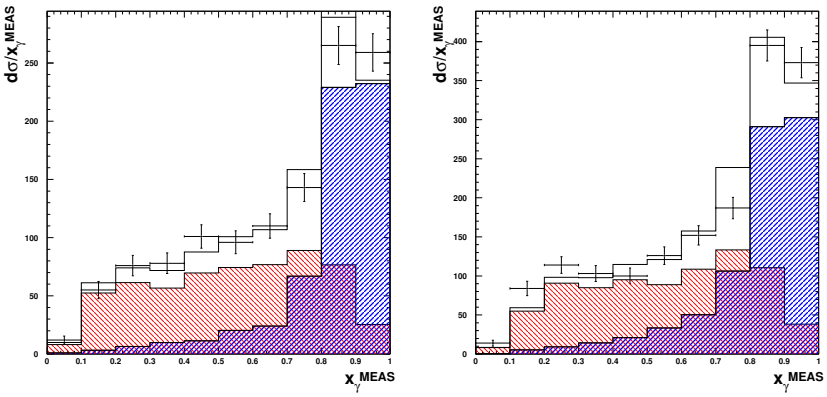

Figure 5.4: *Results of the fit of the direct and resolved Monte Carlo components to the* $x_\gamma^{\rm MEAS}$ *distribution of the data, for 1996–1997 (***left plot***) and 1999–2000 (***right plot***) data samples. The dots represent the data, the fitted Monte Carlo distribution is drawn as the empty histogram and the contributions coming from direct and resolved photon processes are shown as the forward– and backward–diagonally hatched histograms, respectively.*

background coming from direct photon processes does not change. All the resolved cross section are multiplied by y, and in addition the beauty direct and resolved cross sections are multiplied by $x^{\rm dir}$ and $x^{\rm res}$, respectively. The obtained weights are shown in Table 5.3, for the 1996–97 and 1999–2000 Monte Carlo samples.

In the Table it can be noticed that the weights to be applied to the beauty cross sections are much lower in 1996–97 than in 1999–2000. This is due to the different Monte Carlo statistics available in the two periods: in 1996–97 the ratio between the beauty and the background statistics is lower than in 1999–2000. As a consequence, in reweighting the beauty luminosities with respect to the background, a lower factor $x^{\rm dir,res}$ is needed.

Years	b direct	b resolved	cl direct	cl resolved
1996–97	0.281	0.495	1	1.165
1999–2000	1.101	0.990	1	1.155

Table 5.3: *Weights to be applied to the cross sections of the different Monte Carlo contributions in order to correctly take into account the fractions of direct and resolved events in the data, and the fractions of beauty and background events in the regions of direct and resolved photon processes. All the weights are referred to the one applied to the background direct processes, that is therefore equal to 1.*

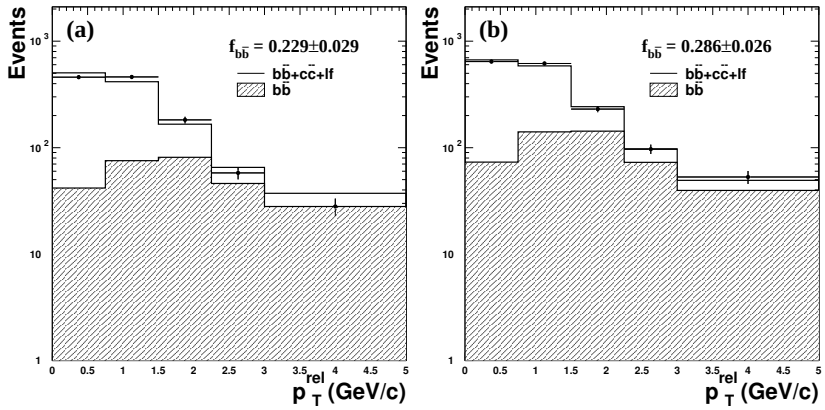

Figure 5.5: *The* $p_T^{\rm rel}$ *distributions for* **(a)** *1996–97 and* **(b)** *1999–2000 data. The data are compared to the* PYTHIA *Monte Carlo with the mixture of beauty and charm+light flavour given by the fit. The beauty contribution is represented by the hatched histogram.*

5.5 Results of the fits

After the MC cross sections had been reweighted in the way described above, the $p_T^{\rm rel}$ fit was redone, in order to determine the beauty fraction in the overall data sample. The fit was performed separately for 1996–97 and 1999–2000 data samples. The results are shown in Fig. 5.5: the dots represent the data, the empty histogram is the fit, and the hatched histogram is the beauty contribution. The fitted distribution has been obtained by mixing the beauty and charm+light flavour contributions in the fractions given by the $p_T^{\rm rel}$ fit.

As can be seen in the plots, the fit describes the 1999–2000 data distribution very well, while the shape of the 1996–97 data distribution is reasonably but not very well described by the fit.

The beauty fractions found in the data are summarized below:

- **1996–97:** $f_b = 0.23 \pm 0.03$ $(stat.)$;
- **1999–2000:** $f_b = 0.29 \pm 0.03$ $(stat.)$.

The $x_\gamma^{\rm MEAS}$ distribution was then plotted using the reweighted cross sections to obtain the beauty and charm+light flavour components and then mixing signal and background in the fractions given by the $p_T^{\rm rel}$ fit. The results are shown in Fig. 5.6, where the data (dots) are compared to the sum of b and $c+l$ components. The fitted distributions are area normalized to the data, and the hatched histograms represent the contribution coming from beauty events. The description of the data by the Monte Carlo is reasonably good, for both 1996–97 and 1999–2000 data.

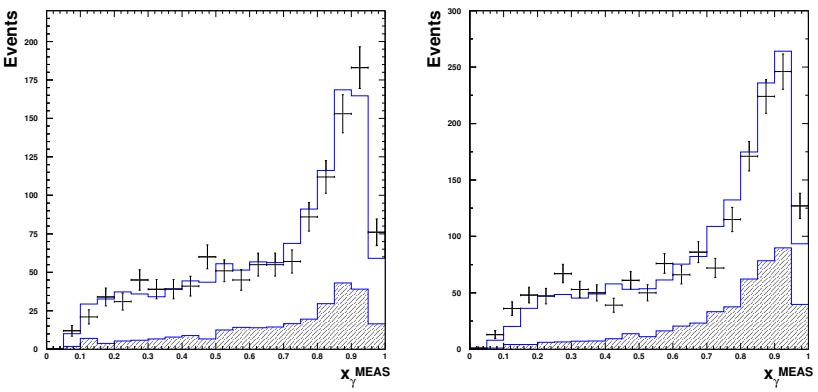

Figure 5.6: *The $x_\gamma^{\rm MEAS}$ distributions plotted using 1996–97* **(left plot)** *and 1999–2000 data* **(right plot)** *compared to the sum of beauty and charm plus light flavour Monte Carlo, mixed according to the b and background fractions given by the new $p_T^{\rm rel}$ fit, performed after the cross sections of direct and resolved b and $c+l$ sub–processes have been reweighted as explained in the text. The fitted distribution is area–normalized to the data. The hatched histogram represents the beauty contribution.*

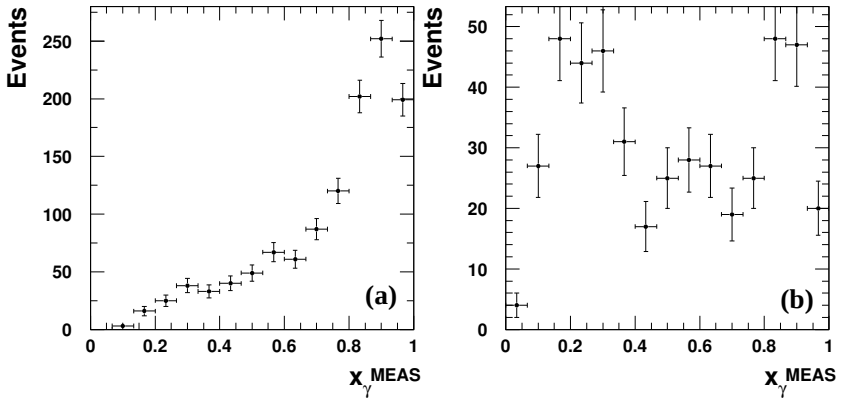

Figure 5.7: *The $x_\gamma^{\rm MEAS}$ distribution for data having good muons in the barrel/rear* **(a)** *and forward region* **(b)**. *The data sample refers to the 1999–2000 period.*

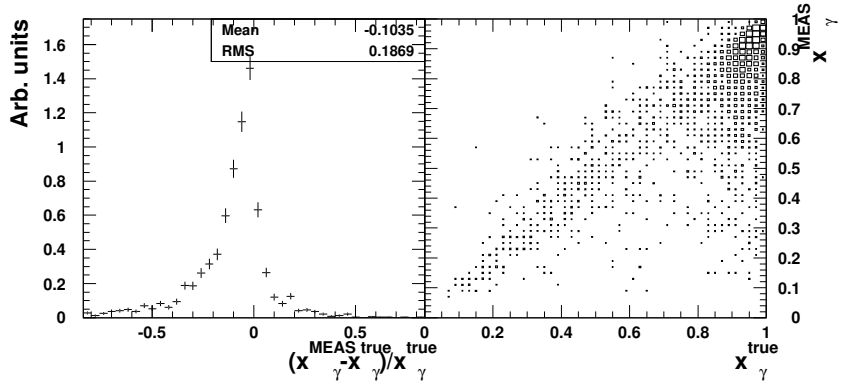

Figure 5.8: *$(x_\gamma^{\rm MEAS} - x_\gamma^{\rm TRUE})/x_\gamma^{\rm TRUE}$ and $x_\gamma^{\rm MEAS}$ vs $x_\gamma^{\rm TRUE}$, where $x_\gamma^{\rm TRUE}$ is evaluated using massive jets before the beauty decay, as illustrated in Sect. 4.7.4. The sample is obtained by mixing the beauty Monte Carlo samples (direct, resolved and excitations) in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo. The observed resolution on the $x_\gamma^{\rm MEAS}$ variable is $\sim 20\%$.*

In the plots reported in Fig. 5.6 the fact can be observed that in our data the direct photon processes are dominant, since a strong peak is present in the high–$x_\gamma^{\rm MEAS}$ region. Nevertheless, also a significant tail is observed at low $x_\gamma^{\rm MEAS}$, indicating the presence of resolved photon processes. Some interesting indications can be obtained by looking at the two $x_\gamma^{\rm MEAS}$ distributions in the barrel/rear and forward region, reported in Fig. 5.7. It can be noticed that the $x_\gamma^{\rm MEAS}$ distribution in the barrel and rear region is strongly peaked at 1, indicating the dominance of direct photon processes. The distribution in the forward region is quite flat, indicating that resolved photon processes are as important as the direct in this region. Therefore, the mechanisms leading to the production of muon+dijet events are different in the two regions.

The resolution on the $x_\gamma^{\rm MEAS}$ variable is shown in Fig. 5.8 (left plot), and is $\sim 20\%$. In Fig. 5.8 (right plot) a comparison between the reconstructed and the true $x_\gamma^{\rm MEAS}$ is shown. In Table 5.4 the resolutions on the different variables of interest for our analysis are summarized.

Variable	Resolution	Variable	Resolution
y_{JB}	16%	p^μ (MPMATCH)	19%
$x_\gamma^{\rm MEAS}$	19%	η^μ (MPMATCH)	8%
$p_T^{\rm jet}$	21%	p^μ (BREMAT)	4%
$\eta^{\rm jet}$	41%	η^μ (BREMAT)	2%

Table 5.4: *Summary of the resolutions on the variables of interest for the analysis.*

5.6 Control plots

Using our technique, we have succeeded in describing reasonably well the $p_T^{\rm rel}$ and $x_\gamma^{\rm MEAS}$ distributions of the data. We have used these two distributions in order to determine the fractions of beauty and background events and of direct and resolved photon processes in our selected data sample. To better describe the $x_\gamma^{\rm MEAS}$ distribution we have reweighted the Monte Carlo cross sections. To validate our procedure now other kinematic variables of the event, like muon and jets variables, will be plotted and compared to the Monte Carlo distributions obtained with the cross sections reweighted as explained above, and inserting a beauty contribution equal to the fraction given by the $p_T^{\rm rel}$ fit. A good description of the data by the Monte Carlo is a confirmation that the procedure followed in the previous Sections is correct.

As in Fig. 5.5, in the plots the beauty and charm+light flavour contributions are mixed according to the percentages given by the $p_T^{\rm rel}$ fit and the fitted distribution is area–normalized to the data. The various distributions are plotted separately for 1996–97 and 1999–2000 data sample.

In Figs. 5.9 (for 1996–97 data–sample) and 5.10 (for 1999–2000) the muon variables are plotted. In the upper plots, the muon momentum, p^μ, and transverse momentum, p_T^μ, are shown, and the data description by the Monte Carlo is reasonably good. In the middle plot the pseudorapidity, η^μ, can be seen. A better description of the region $-0.7 < \eta^\mu < -0.2$ and of the very high pseudorapidity region can be obtained by further reweighting the Monte Carlo, but this has not been done in the plots shown. In the lower plots the matching probability given by the two algorithms, MPMATCH and BREMAT, is shown. The description of the data seems to be better for the MPMATCH algorithm, whereas for BREMAT a high peak in the low probability region is present, which is not simulated by the Monte Carlo. This behaviour is observed in both 1996–97 and 1999–2000 data samples. From previous studies it is known that this peak comes mainly from light flavour events. Studies on the simulation of background events, especially in the high $p_T^{\rm rel}$ region, are now undergoing.

In Figs. 5.11 and 5.12 the jets variables are shown, relative to the highest and second highest transverse momentum jets in the event ($p_T^{\rm jet1}$, $\eta^{\rm jet1}$, $p_T^{\rm jet2}$ and $\eta^{\rm jet2}$) and to the jet associated to the muon ($p_T^{\rm sel\ jet}$ and $\eta^{\rm sel\ jet}$). The data description by the Monte Carlo simulation is reasonably good, in particular for the jet associated with the muon, which is the most interesting for the analysis since it is used for the $p_T^{\rm rel}$ calculation.

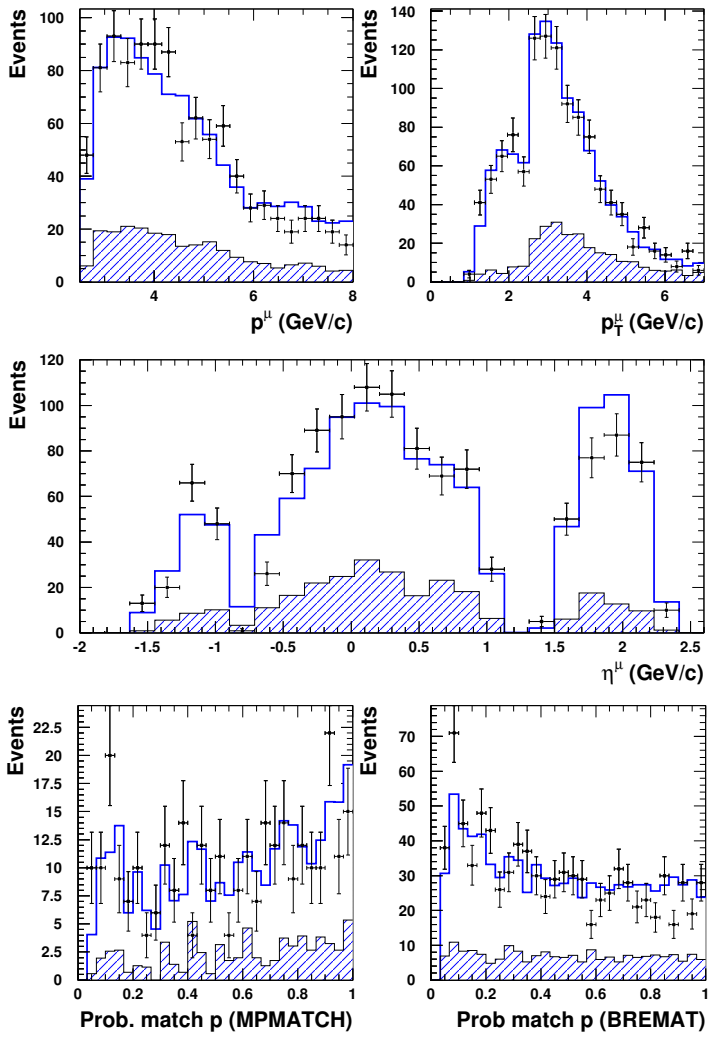

Figure 5.9: *Control plots for the muon variables, relative to the* **1996–97** *data sample. In the upper plots, the muon momentum,* p^μ, **(left)** *and transverse momentum,* p_T^μ, **(right)** *are shown, in the middle plot the muon pseudorapidity,* η^μ, *is displayed. The lower plots show the probability of the 5 d.o.f. matching,* Prob. match p, *of the two muon matching algorithms, MPMATCH and BREMAT . In all the plots the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the* $p_T^{\rm rel}$ *fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.*

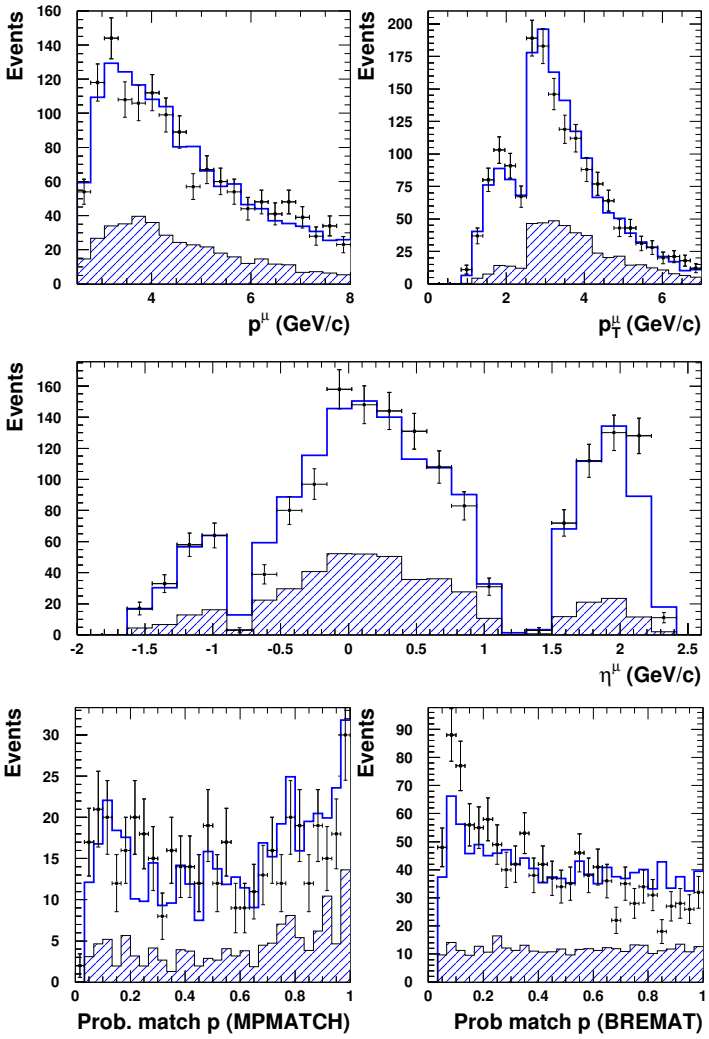

Figure 5.10: *Control plots for the muon variables, relative to the* **1999–2000** *data sample. In the upper plots, the muon momentum, p^{μ},* (**left**) *and transverse momentum, p_T^{μ},* (**right**) *are shown, in the middle plot the muon pseudorapidity, η^{μ}, is displayed. The lower plots show the probability of the 5 d.o.f. matching,* Prob. Match p, *of the two muon matching algorithms, MPMATCH and BREMAT . In all the plots the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the p_T^{rel} fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.*

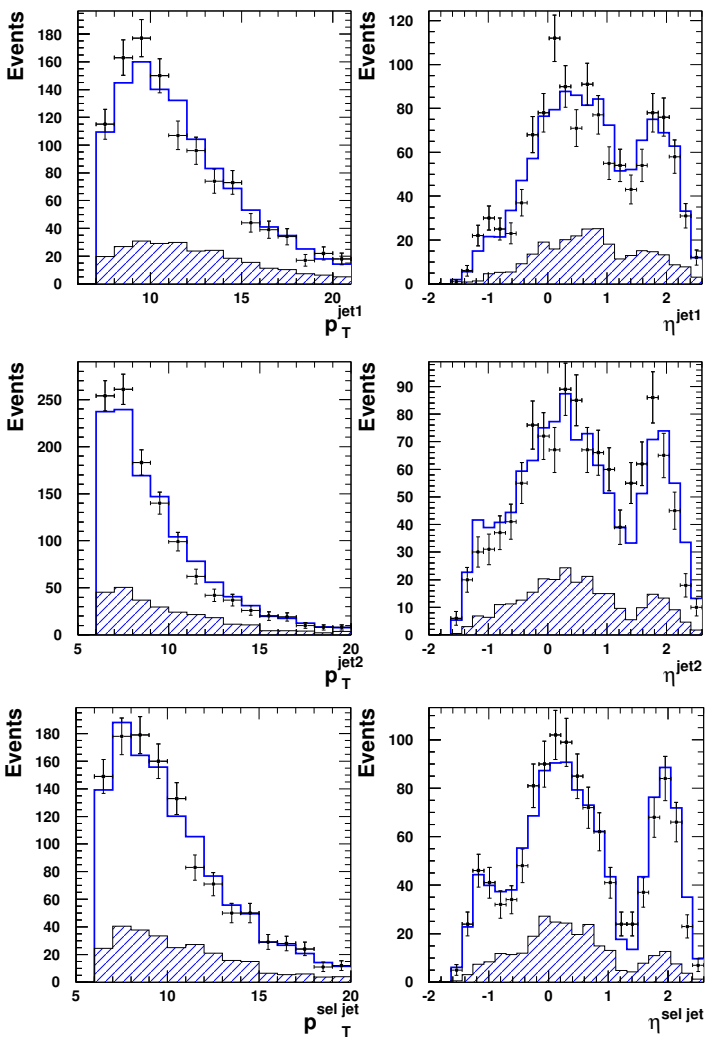

Figure 5.11: *Control plots for the jets variables, relative to the* **1996–97** *data sample. In the upper plots, the transverse momentum, p_T^{jet1},* (**left**) *and pseudorapidity, η^{jet1},* (**right**) *of the highest transverse momentum jet in the event are shown, and the same variables, referred to the second–highest transverse momentum jet, are shown in the middle plots. In the lower plots the transverse momentum, $p_T^{\mathrm{sel\ jet}}$,* (**left**) *and pseudorapidity, $\eta^{\mathrm{sel\ jet}}$,* (**right**) *of the jet associated to the muon are displayed. In all the plots, the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the p_T^{rel} fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.*

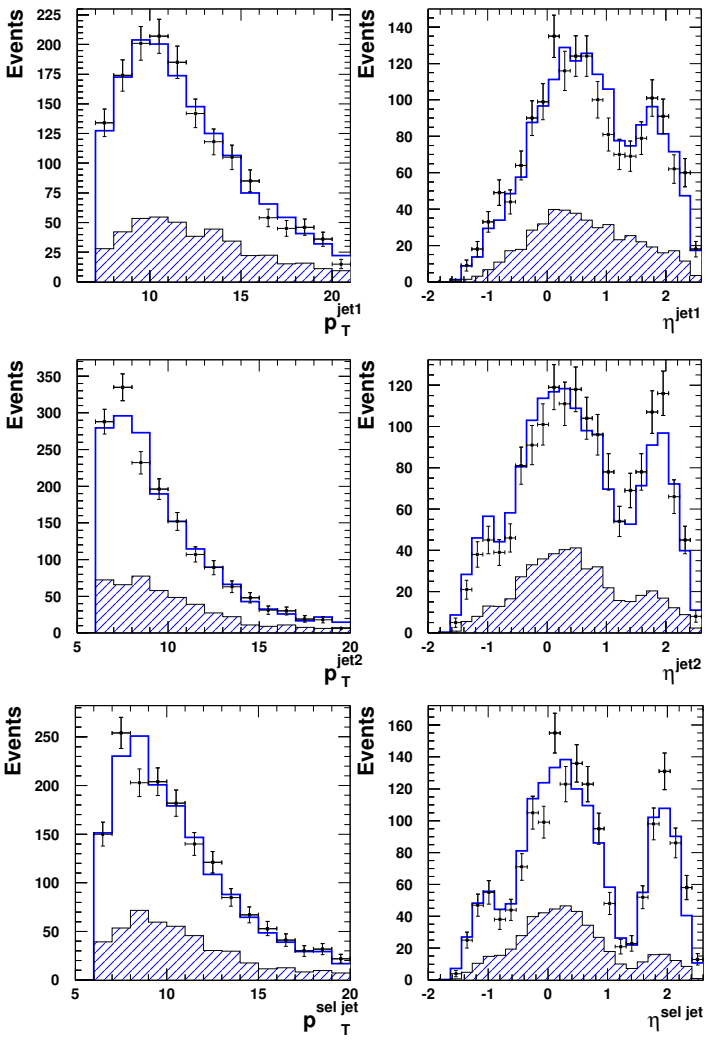

Figure 5.12: *Control plots for the jets variables, relative to the* **1999–2000** *data sample. In the upper plots, the transverse momentum, $p_T^{\rm jet1}$,* **(left)** *and pseudorapidity, $\eta^{\rm jet1}$,* **(right)** *of the highest transverse momentum jet in the event are shown, and the same variables, referred to the second–highest transverse momentum jet, are shown in the middle plots. In the lower plots the transverse momentum, $p_T^{\rm sel\ jet}$,* **(left)** *and pseudorapidity, $\eta^{\rm sel\ jet}$,* **(right)** *of the jet associated to the muon are displayed. In all the plots, the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the $p_T^{\rm rel}$ fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.*

Chapter 6

Measurement of the beauty cross sections

In this Chapter the results concerning the measured beauty cross sections will be shown. The cross sections will be calculated separately for the two data–taking periods, and then the obtained values will be merged. The systematic uncertainties affecting the cross section determination will be analyzed. The obtained results will be compared to NLO QCD predictions.

6.1 Determination of the NLO QCD prediction

In the following Sections, the measured beauty photoproduction cross sections will be compared to the predictions of NLO QCD, based on the program by Frixione et al. [22]. The program gives as output the NLO cross sections for $b\bar{b}$ photoproduction at the parton level, without hadronization and decay of the produced particles. Dealing with data, we can measure beauty cross sections using events in which the heavy quarks produce jets and decay into muons. Therefore, in order to compare the measurements to the NLO predicted cross sections, the NLO needs the implementation of fragmentation and hadronization, and in some cases the measurements have to be extrapolated.

In the determination of the NLO QCD predictions [120] the parton distribution functions used for the proton and the photon were CTEQ5M [121] and GRVG–HO [28], respectively. The b–quark mass was set to 4.75 GeV, and the renormalization and factorization scales to the transverse mass, $\mu_R = \mu_F = \mu = m_T = \sqrt{m_b^2 + p_T^2}$, where m_b is the b–quark mass and p_T its transverse momentum. Jets were reconstructed running the k_T algorithm on partons, and then correcting from parton to hadron level using the PYTHIA Monte Carlo.

In Fig. 6.1 the parton to hadron correction factors (see Eq. 6.1) are shown, extracted using different reconstructed jets, to compare the results obtained using the various methods. The parton to hadron corrections have been evaluated in the three regions, forward, barrel and rear, in which the visible cross sections will be measured (see Sect. 6.2), in a unique bin, for the beauty dijet cross section (Sect. 6.3), and in all the bins of the measured differential cross sections (Sects. 6.3 and 6.4) in order to compare them to NLO. In Fig. 6.1, for simplicity, also the corrections to be applied in the three angular regions of the visible cross section are shown as an histogram.

The correction factor in the ith bin is given by the ratio:

$$ph = \frac{N_i^{\rm hadr}}{N_i^{\rm part}} \tag{6.1}$$

where $N_i^{\rm hadr}$ ($N_i^{\rm part}$) is the number of events with $Q^2 < 1$ GeV^2, $0.2 < y < 0.8$ having two good hadronic (partonic) jets in the final state, i.e. two jets satisfying the cuts in Eqs. 4.28 and 4.29.

The ph factor has been evaluated comparing massive (massless) hadronic jets to massive (massless) partonic jets. The hadronic jets have been reconstructed using the final state stable particles, including the b decay products (jets *after beauty decay*) or all the particles before the beauty decay, including the initial B meson (jets *before beauty decay*).

It can be noticed that the choice of jets before the beauty decay makes the parton to hadron corrections much lower since in this case the reconstructed jets are more similar to the partonic ones. Moreover, in the case of massive jets before the b decay, looking at the $x_\gamma^{\rm MEAS}$ correction factors we notice that the corrections in each bin are very similar. This is not true if massless jets are used: a large bump can be seen in the third bin, probably due to bin migrations of the events and to still unclear effects. This is the motivation why in this analysis the choice has been done to use massive jets before the beauty decay in the evaluation of the cross sections.

In the calculation of NLO–predicted muon cross sections, the b fragmentation into hadrons was described using the Peterson [45] fragmentation function, with $\epsilon = 0.0035$ [122]. The spectrum of muons from B–hadron decay was taken from PYTHIA.

To estimate the uncertainty on the NLO predictions, the b–quark mass and the renormalization and factorization scales were varied simultaneously in order to obtain the extreme values of the predicted cross sections. The minimum value was obtained for $m_b = 5.0$ GeV, $\mu_R = \mu_F = 2 \cdot m_T$, the maximum for $m_b = 4.5$ GeV, $\mu_R = \mu_F = m_T/2$. Also the variation of the predicted values with the ϵ parameter was investigated, varying ϵ from 0.002 to 0.005: the obtained effect was found negligible. The parton to hadron corrections were recalculated using HERWIG and the values were found in agreement with those from PYTHIA.

The NLO hadronic cross sections have upward uncertainties of the order of 30–50% and downward uncertainties slightly lower, of the order of 20–40%, due to the change in the theoretical input parameters.

6.2 Visible cross sections

Once that the beauty percentage in the data is determined, the cross section of the process

$$e\ p \to b\ \overline{b}\ X \to \text{dijet}\ \mu\ X \tag{6.2}$$

can be calculated in the region of interest. First, we want to evaluate the visible cross sections in the three angular regions covered by the muon detectors, forward, barrel and rear. The beauty cross section is evaluated in the **photoproduction regime**:

$$Q^2 < 1\ \text{GeV}^2 \quad ; \quad 0.2 < y < 0.8, \tag{6.3}$$

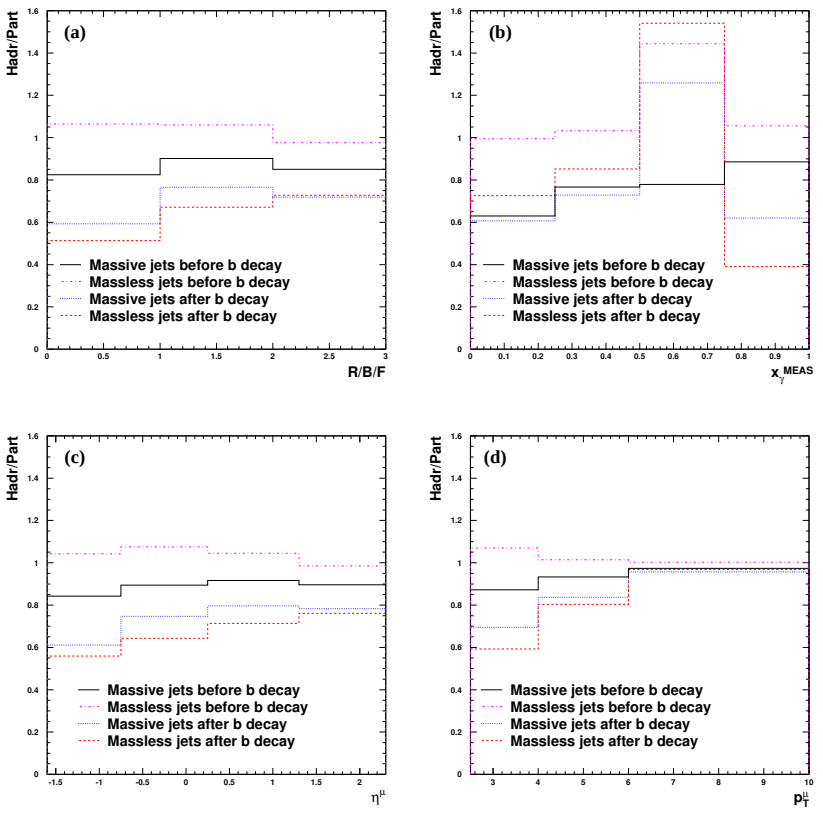

Figure 6.1: *Parton to hadron corrections, evaluated using massive jets before the beauty decay (full line), massless jets before the beauty decay (dot–dashed line) massive (dotted line) and massless (dashed line) jets after the beauty decay. The Monte Carlo used is* PYTHIA *and the different beauty sub–processes have been mixed according to their inverse integrated luminosities. The correction factors are shown in the three angular regions of the visible cross sections* **(a)** *and in all the bins of the differential distributions* $d\sigma/dx_\gamma^{\rm MEAS}$ **(b)**, $d\sigma/d\eta^\mu$ **(c)** *and* $d\sigma/dp_T^\mu$ **(d)**.

for events in which at least two jets are found in the final state, coming from the hadronization of the b quarks and satisfying:

$$p_T^{\rm jet1,2} > 7(6)\ \text{GeV} \quad ; \quad |\eta^{\rm jet1,2}| < 2.5. \tag{6.4}$$

Years	Beauty percentages		
	Forward	Barrel	Rear
1996–97	14.0 ± 7.7	24.1 ± 3.9	21.5 ± 6.9
1999–2000	31.1 ± 8.1	31.7 ± 3.4	29.0 ± 7.2

Table 6.1: *Beauty percentages in the three angular regions defined for the visible cross section calculation, obtained by fitting to the data the $p_T^{\rm rel}$ distributions of the beauty and charm plus light flavour components given by the Monte Carlo. The results have been evaluated separately in the two data–taking periods, 1996–97 and 1999–2000.*

At least one muon coming from the beauty semi–leptonic decay has to be present in the final state. The kinematic region of the muon depends on its angular position and is essentially defined by the acceptance of the muon chambers:

- **forward region:**

$$p^\mu > 4 \text{ GeV}; \quad p_T^\mu > 1 \text{ GeV}; \quad 1.5 < \eta^\mu < 2.3; \tag{6.5}$$

- **barrel region:**

$$p_T^\mu > 2.5 \text{ GeV} \quad ; \quad -0.9 < \eta^\mu < 1.3; \tag{6.6}$$

- **rear region:**

$$p^\mu > 2.5 \text{ GeV} \quad ; \quad -1.6 < \eta^\mu < -0.9. \tag{6.7}$$

In the evaluation of the cross section in the three regions, the beauty percentage in the data has been determined in each angular range. The method is exactly the one described in the previous Chapter. The obtained beauty fractions are shown in Table 6.1. It can be noticed that the fractions in 1999–2000 are always higher than those obtained in 1996–97, even if, keeping the errors into account, the largest discrepancy is of 1.5 σ. An increase in the beauty fraction is expected since the centre–of–mass energy is higher in 1999–2000, and the beauty cross section is therefore larger.

The visible cross section in the defined kinematic region is derived using the following formula:

$$\sigma_{\rm vis} = \frac{f_b \; N^{\rm evt}_{\rm F,B,R}}{\mathcal{L} \; acc \; M}, \tag{6.8}$$

where f_b is the fraction of beauty events in the data and $N^{\rm evt}_{\rm F,B,R}$ is the total number of data events in the forward (F), barrel (B) or rear (R) region (see Table 6.2). $\mathcal{L}$ is the integrated luminosity of the data, and *acc*, whose meaning will be explained in more detail below, is the acceptance of the detector. M is the correction factor due to the muon chambers performances, defined in Eq. 4.34 for the forward region and in Table 4.2 for the barrel and rear region; the different correction factors to be applied in the three regions are summarized in Table 6.3.

The acceptance, *acc*, in Eq. 6.8, is evaluated using beauty Monte Carlo events and is defined as:

$$acc = \frac{N^{\rm rec}}{N^{\rm true}}. \tag{6.9}$$

Here, $N^{\rm rec}$ and $N^{\rm true}$ have the following meaning:

	Cuts in Eq. 6.5-6.7		
Years	Forward	Barrel	Rear
1996–97	302	741	146
1999–2000	453	1015	172
	Cuts in Eq. 6.30		
Years	Forward	Barrel	Rear
1996–97	103	741	54
1999–2000	187	1015	54

Table 6.2: *Comparison between the number of events selected in the forward, barrel and rear regions with two different sets of cuts: the first set, defined in Eqs. 6.5, 6.6 and 6.7, has been used for the evaluation of the visible cross sections in the three angular regions, for the dijet cross section and for $d\sigma/dx_\gamma^{\rm MEAS}$; the second set, defined in 6.30, has been used for the determination of the muon differential cross sections, $d\sigma/d\eta^\mu$ and $d\sigma/dp_T^\mu$.*

- $N^{\rm rec}$ is the total number of reconstructed Monte Carlo events that satisfy all the cuts applied to the data: i.e. they have at least one muon in the final state reconstructed by the BREMAT or MPMATCH algorithm, satisfying the muon cuts, and two jets reconstructed using the KTCLUS algorithm running on the corrected ZUFOs, etc. This means that in $N^{\rm rec}$ all the sources of inefficiency in the data selection are taken into account, except for that of the muon reconstruction algorithm, already included in the M factor;

- $N^{\rm true}$ is the number of events generated in the kinematic region of interest. These are events in which a $b\bar{b}$ couple is produced. Two jets have to be present in the final state, coming from the hadronization of the b quarks and fulfilling the requirements in Eq. 6.4, and a muon coming from a b decay has to be found in the previously defined kinematic region: in the simulation all these information are available. The variables considered to determine if an event is or is not in the interesting kinematic region are the true ones, as given by the Monte Carlo.

In our case, since four different beauty sub–processes have been simulated using Monte Carlo (see Sect. 3.5.1) the contribution of each sub–process has been taken into

Region	1996–97	1999–2000
Forward	0.98 ± 0.08	0.98 ± 0.08
Barrel	0.71 ± 0.11	0.61 ± 0.09
Rear	$0.54^{+0.03}_{-0.07}$	0.46 ± 0.05

Table 6.3: *Summary of the correction factors to be applied to the data in the three angular regions, due to the muon detectors performances. The correction factors are given by the ratio of the chambers performances evaluated on the data and on the Monte Carlo. The values relative to the 1996–97 and 1999–2000 samples have been evaluated separately.*

Years	Forward	Barrel	Rear
1996–97	0.13	0.20	0.19
1999–2000	0.11	0.19	0.18

Table 6.4: *Measured values of the acceptance in the three angular regions in the two data–taking periods.*

account in the acceptance evaluation according to the inverse luminosity of the sample:

$$acc = \frac{\sum_{i \in b \text{ procs}} \mathcal{L}_i^{-1} N_i^{\text{rec}}}{\sum_{i \in b \text{ procs}} \mathcal{L}_i^{-1} N_i^{\text{true}}}, \tag{6.10}$$

where i runs over the four beauty sub–processes, direct and resolved photon and beauty excitation in the proton and in the photon. The measured acceptances are shown in Table 6.4.

The cross sections measured using the **1996–97** data are:

- **forward:**
$$\sigma_{\text{vis}}^F = 16.0 \pm 5.1 \text{ (stat)}^{+2.7}_{-3.0} \text{ (syst.) pb;} \tag{6.11}$$
- **barrel:**
$$\sigma_{\text{vis}}^B = 32.8 \pm 5.3 \text{ (stat)}^{+6.4}_{-6.1} \text{(syst.) pb;} \tag{6.12}$$
- **rear:**
$$\sigma_{\text{vis}}^R = 5.1 \pm 2.9 \text{ (stat)}^{+2.4}_{-0.7} \text{ (syst.) pb.} \tag{6.13}$$

For the **1999–2000** data–taking period, the obtained results are:

- **forward:**
$$\sigma_{\text{vis}}^F = 20.5 \pm 5.1 \text{ (stat)}^{+2.6}_{-6.7} \text{ (syst.) pb;} \tag{6.14}$$
- **barrel:**
$$\sigma_{\text{vis}}^B = 46.3 \pm 4.9 \text{ (stat)}^{+9.3}_{-8.4} \text{ (syst.) pb;} \tag{6.15}$$
- **rear:**
$$\sigma_{\text{vis}}^R = 10.7 \pm 2.8 \text{ (stat)}^{+2.6}_{-1.4} \text{ (syst.) pb.} \tag{6.16}$$

The way in which the systematic uncertainties have been evaluated is described in detail at the end of the Chapter (see Sect. 6.6).

The measured values in 1996–97 are lower than those obtained in 1999–2000. This is due to the increase of the cross section with the centre–of–mass energy. If the statistical errors are kept into account, considering the systematics as totally correlated, the largest discrepancy between the two measured cross sections is $\sim 1.9\ \sigma$.

The results have been merged in a simple way, in order to have a unique value with a reduced statistical error. The raise of the cross section with the increasing centre–of–mass energy has been assumed linear: this hypothesis has its justification in the fact that the difference between the two centre–of–mass energies is very small (300 GeV in 1996–97 and 318 GeV in 1999–2000). After the two values of the cross section have been measured the straight line interpolating them is determined. The cross section is then evaluated at an average centre–of–mass energy given by:

$$E_{CM}^{\text{mean}} = \frac{\mathcal{L}_{9697} \cdot E_{CM}^{9697} + \mathcal{L}_{9900} \cdot E_{CM}^{9900}}{\mathcal{L}_{9697} + \mathcal{L}_{9900}} \simeq 311 \text{ GeV} \tag{6.17}$$

where $\mathcal{L}_{9697}$ and $\mathcal{L}_{9900}$ are the integrated luminosities of the data collected in the two periods, and E_{CM}^{9697}, E_{CM}^{9900} the corresponding centre–of–mass energies. The statistical error on the cross section calculated at E_{CM}^{mean} has been obtained by propagating the statistical errors on the two cross sections in input.

The obtained results for the merged cross sections in the forward, barrel and rear regions are reported below:

- **forward:**
$$\sigma_{\text{vis}}^F = 18.7 \pm 3.7 \text{ (stat)}\ ^{+2.4}_{-5.1}\text{(syst.) pb;} \tag{6.18}$$
- **barrel:**
$$\sigma_{\text{vis}}^B = 41.0 \pm 3.6 \text{ (stat)}\ ^{+8.0}_{-7.5}\text{(syst.) pb;} \tag{6.19}$$
- **rear:**
$$\sigma_{\text{vis}}^R = 8.5 \pm 2.0 \text{ (stat)}\ ^{+2.4}_{-1.1}\text{(syst.) pb.} \tag{6.20}$$

The systematic error on the merged cross sections has been determined in the following way: for each effect causing a systematic uncertainty the resulting cross sections have been evaluated, on the 1996–97 and on the 1999–2000 data samples. We refer to these two values as $\sigma_{\text{syst}_i}^{9697}$ and $\sigma_{\text{syst}_i}^{9900}$, where i runs on the sources of systematic indetermination. As before, from these values the cross section at $\sqrt{s} = 311$ GeV can be determined, assuming the input cross sections to lay on a straight line. The obtained $\sigma_{\text{syst}_i}^{\text{mean}}$ is then compared to the previously evaluated central value of the cross section at $\sqrt{s} = 311$ GeV, and the systematic uncertainty relative to the source i is determined from the difference between these two cross sections. The obtained effects are then summed as usually to give a unique value: the quadratic sum of the positive and negative effects are performed separately.

The obtained merged results have then been compared to the NLO QCD predicted cross sections in the three regions of interest:

- **forward:**
$$\sigma_{\text{NLO}}^F = 6.1\ ^{+2.8}_{-1.5} \text{ pb;} \tag{6.21}$$
- **barrel:**
$$\sigma_{\text{NLO}}^B = 33.1\ ^{+9.1}_{-7.4} \text{ pb;} \tag{6.22}$$
- **rear:**
$$\sigma_{\text{NLO}}^R = 4.2\ ^{+1.2}_{-1.0} \text{ pb.} \tag{6.23}$$

As it can be noticed from the numbers, the NLO QCD predictions agree within errors with the measured values in the barrel and rear region, even if both the predictions are below the measurements. The discrepancy between data and NLO seems to be slightly more pronounced in the forward region (around 1.8 σ). This can be due to the fact that the processes contributing to beauty production in the three regions are

Years	Acceptance
1996–97	0.0158
1999–2000	0.0154

Table 6.5: *Measured values of the acceptance for the dijet beauty cross section determination in the two data–taking periods.*

different. From the $x_\gamma^{\rm MEAS}$ distribution of the data (see Fig. 5.7) it has been noticed that, while in the barrel and rear region the direct photon processes are dominant, in the forward region resolved photon processes become important. Resolved photon processes involve also beauty excitation phenomena not included in the NLO calculations, since the massive approach is used and heavy quarks are not contained in the photon and proton structure functions.

6.3 Dijet cross sections

Starting from events in which beauty is produced and decays semi–leptonically into muons, the cross section of the process

$$e\ p \to b\ \overline{b}\ X \to \text{dijet}\ X \tag{6.24}$$

can be extrapolated using a suitable Monte Carlo sample.

As we pointed out in Sect. 3.5.1, the beauty Monte Carlo used in this analysis has been generated without any request on the muons or on the jets, and the b quark has not been forced to decay into muons. Therefore, the beauty dijet cross section can be evaluated simply by recalculating the acceptance to be inserted in Eq. 6.8. This can be done using the Formula in 6.10 and defining $N_i^{\rm true}$ as the total number of Monte Carlo events generated in the kinematic region of interest, for each of the beauty sub–processes. This kinematic region will not contain any request on the muon presence or momentum.

The cross section of the process 6.24 has been evaluated in the kinematic region defined by:

$$\begin{aligned} Q^2 < 1\ \text{GeV}^2 \quad &; \quad 0.2 < y < 0.8 \\ p_T^{\rm jet1(2)} > 7(6)\ \text{GeV} \quad &; \quad |\eta^{\rm jet1(2)}| < 2.5. \end{aligned} \tag{6.25}$$

The measured acceptances are reported in Table 6.5. The obtained values are one order of magnitude lower than the previous ones, due to the extrapolation for the muon decay and branching ratio ($\mathcal{B}(b \to \mu) = 0.220 \pm 0.011$) [96].

The beauty dijet cross section obtained using **1996–97** data is:

$$\sigma(ep \to b\bar{b}X \to \text{dijet}\ X) = 631 \pm 80\ (\text{stat.})^{+111}_{-95}\ (\text{syst.})\ \text{pb}, \tag{6.26}$$

while in the **1999–2000** data–taking period the measured cross section is:

$$\sigma(ep \to b\bar{b}X \to \text{dijet}\ X) = 809 \pm 74\ (\text{stat.})^{+148}_{-108}\ (\text{syst.})\ \text{pb}. \tag{6.27}$$

The merged value is:

$$\sigma(ep \to b\bar{b}X \to \text{dijet}\ X) = 740 \pm 55\ (\text{stat.})^{+130}_{-98}\ (\text{syst.})\ \text{pb}. \tag{6.28}$$

The NLO–predicted cross section is:

$$\sigma_{\rm NLO} = 381\ ^{+117}_{-78}\ \text{pb}. \tag{6.29}$$

In this case the discrepancy between data and NLO predictions is $\sim 2.2\ \sigma$. The fact has to be kept into account, however, that this cross section has been extrapolated using the Monte Carlo, and that therefore a dependence on the model can be present, even if it was taken into account in the systematic error. The slight discrepancy between data and NLO predictions found in the determination of the visible cross sections reflects in the disagreement observed here.

In order to better investigate where this discrepancy comes from, the differential cross section $d\sigma/dx_\gamma^{\rm MEAS}$ has been determined and compared to NLO QCD predictions. In the evaluation of this differential cross section, the beauty fraction in the data has been extracted by fitting in each $x_\gamma^{\rm MEAS}$ bin the $p_T^{\rm rel}$ shape of the signal and background components to the data. The value of the cross section in each bin is then obtained by the usual formula in Eq. 6.8. This method has been used also for the muon differential cross sections shown in the following Section.

The results are shown in Figs. 6.2 and 6.3. In Fig. 6.2 the results obtained using 1996–97 and 1999–2000 data samples are compared: in the first three bins the measured values are compatible, while in the last bin, corresponding to the direct photon processes, the result obtained with 1996–97 data is lower than that extracted using 1999–2000 data, but the errors are large. The solid histogram represents the cross sections predicted by the PYTHIA Monte Carlo at $\sqrt{s} = 318$ GeV, while the dashed histogram is the prediction at $\sqrt{s} = 300$ GeV: the difference in the predictions, due to the increase in the centre–of–mass energy, is around 10%.

In Fig. 6.3 the merged values are reported and compared to NLO QCD predictions. The NLO predictions underestimate the data in both the direct ($x_\gamma^{\rm MEAS} > 0.75$) and the resolved ($x_\gamma^{\rm MEAS} < 0.75$) part.

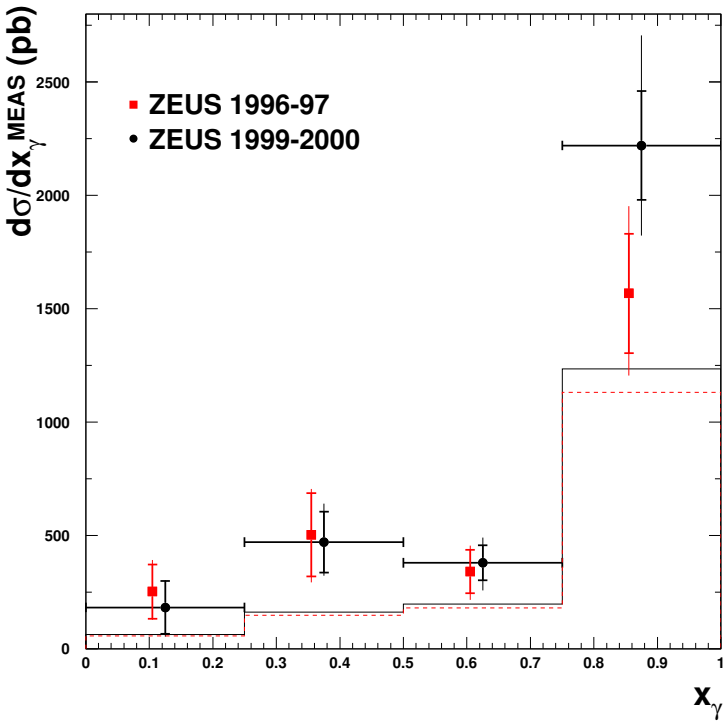

Figure 6.2: *Differential cross section* $d\sigma/dx_\gamma^{\rm MEAS}$ *, evaluated using the two data samples available, relative to the 1996–97 (squares) and 1999–2000 (circles) data–taking periods. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The two histograms represent the* PYTHIA *predictions: the solid histogram is the Monte Carlo cross section at* $\sqrt{s} = 318$ GeV*, the dashed one is the prediction at* $\sqrt{s} = 300$ GeV*.*

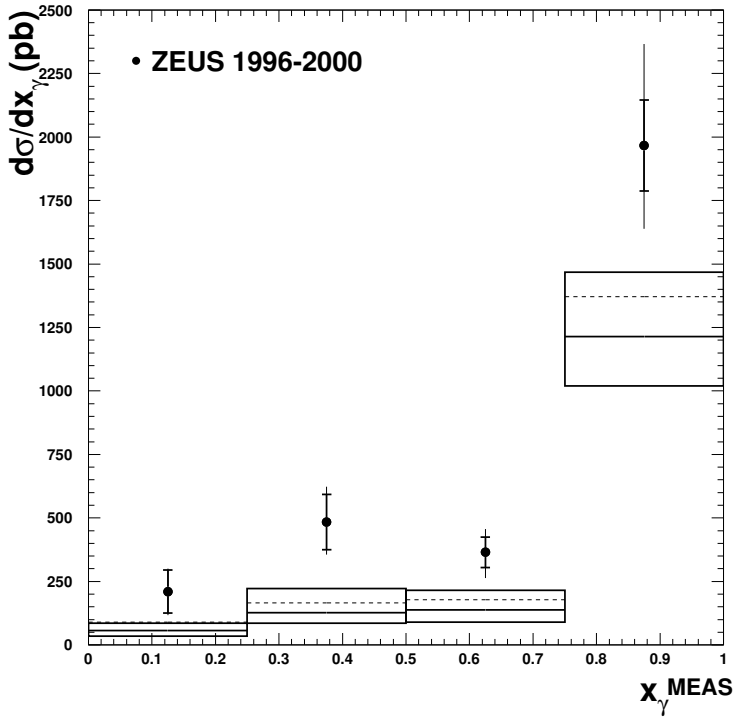

Figure 6.3: *Differential beauty cross section as a function of* $x_\gamma^{\rm MEAS}$ *for two-jets events, compared to the predictions of NLO QCD (dashed line: parton–level jets, full line: corrected to hadron–level jets). The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The results of the two data samples have been merged as explained in the text. The band around the NLO prediction shows the range obtained by varying the b–quark mass as well as* μ_r *and* μ_f*, as explained in the text.*

6.4 Muon differential cross sections

The differential cross sections as a function of the muon pseudorapidity, $d\sigma/d\eta^{\mu}$, and transverse momentum, $d\sigma/dp_T^{\mu}$, have been determined in a different kinematic region with respect to that defined by the cuts in Eqs. 6.5, 6.6 and 6.7. The kinematic region is defined as usual as:

$$\begin{aligned} Q^2 < 1 \text{ GeV}^2 \quad ; \quad 0.2 < y < 0.8 \\ p_T^{\text{jet1(2)}} > 7(6) \text{ GeV} \quad ; \quad |\eta^{\text{jet1(2)}}| < 2.5, \end{aligned}$$

but the muon region is now:

$$\begin{aligned} p_T^{\mu} > 2.5 \text{ GeV} \\ -1.6 < \eta^{\mu} < 2.3. \end{aligned} \tag{6.30}$$

This means that the momentum range has been restricted: now most of the events come from the barrel region. Events are selected in the angular region defined by $-1.6 < \eta^{\mu} < 1.3$ and $1.5 < \eta^{\mu} < 2.3$: the gap between the barrel and the forward muon detectors ($1.3 < \eta^{\mu} < 1.5$) is included using the Monte Carlo for the extrapolation. The number of events selected in each angular region with these new momentum cuts is reported in Table 6.2. In this way, all the data can be used for the determination of the differential cross sections, since a region has been chosen where the acceptance is known to be good for the rear, barrel and forward muon chambers.

The differential cross section as a function of the muon pseudorapidity, $d\sigma/d\eta^{\mu}$, is shown in Fig. 6.4. The results obtained with the two different data samples are compared and the Pythia Monte Carlo predictions are displayed for the two different centre–of–mass energies. It can be noticed that the 1996–97 cross section is systematically lower than the 1999–2000, even if the difference is always within the statistical error. Part of the difference can be explained by the Monte Carlo, that also in this case predicts a raise in the cross section of $\sim 10\%$.

In Fig. 6.5 the results obtained merging 1996–97 and 1999–2000 data are shown and compared to NLO QCD predictions. In this case, the NLO calculations describe the data reasonably well, keeping into account the experimental and theoretical uncertainties.

The differential cross section as a function of the muon transverse momentum, $d\sigma/dp_T^{\mu}$, has been determined in the same kinematic region as $d\sigma/d\eta^{\mu}$. The comparison between 1996–97 and 1999–2000 results is shown in Fig. 6.6.

The merged results are plotted in Fig. 6.7 and compared to NLO QCD predictions. As in the case of $d\sigma/d\eta^{\mu}$, the differential cross section is reasonably well described by the theoretical calculations, keeping into account the experimental and theoretical uncertainties.

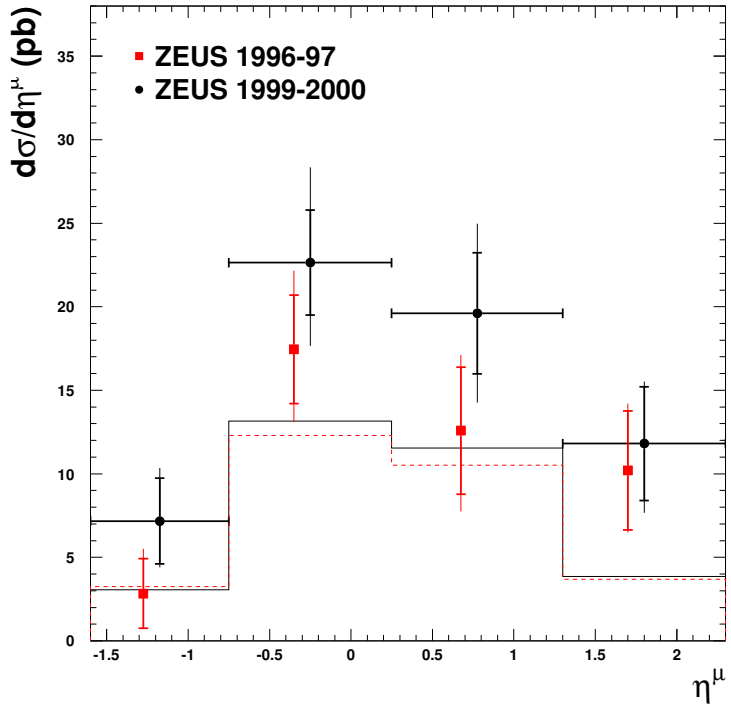

Figure 6.4: *Differential cross section $d\sigma/d\eta^{\mu}$, evaluated using the two data samples available, relative to the 1996–97 (squares) and 1999–2000 (circles) data–taking periods. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The two histograms represent the* Pythia *predictions: the solid histogram is the Monte Carlo cross section at $\sqrt{s} = 318$ GeV, the dashed one is the prediction at $\sqrt{s} = 300$ GeV.*

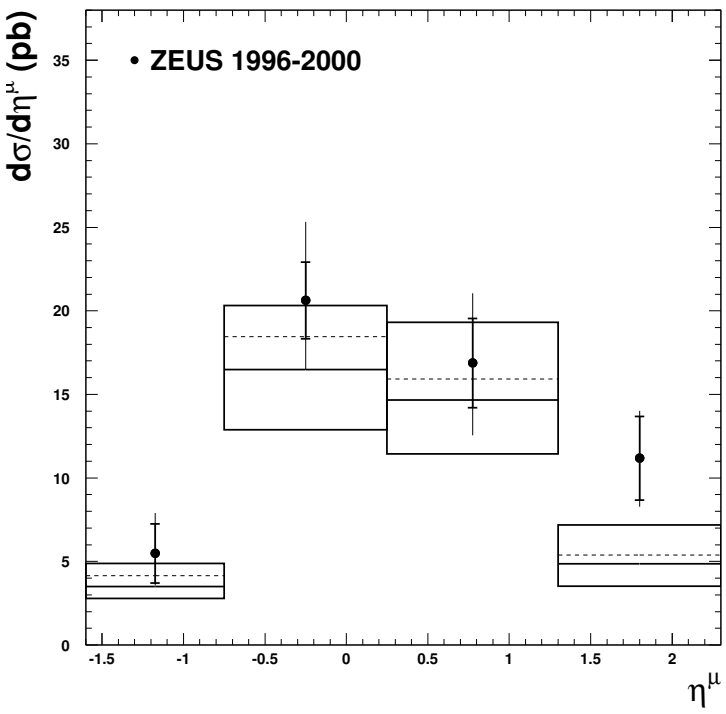

Figure 6.5: *Differential beauty cross section as a function of the muon pseudorapidity, η^{μ}, for events with two jets and a muon, compared to the predictions of NLO QCD (dashed line: parton–level jets, full line: corrected to hadron–level jets). The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The results of the two data samples have been merged as explained in the text. The band around the NLO prediction shows the result obtained by varying the b–quark mass as well as μ_r and μ_f, as explained in the text.*

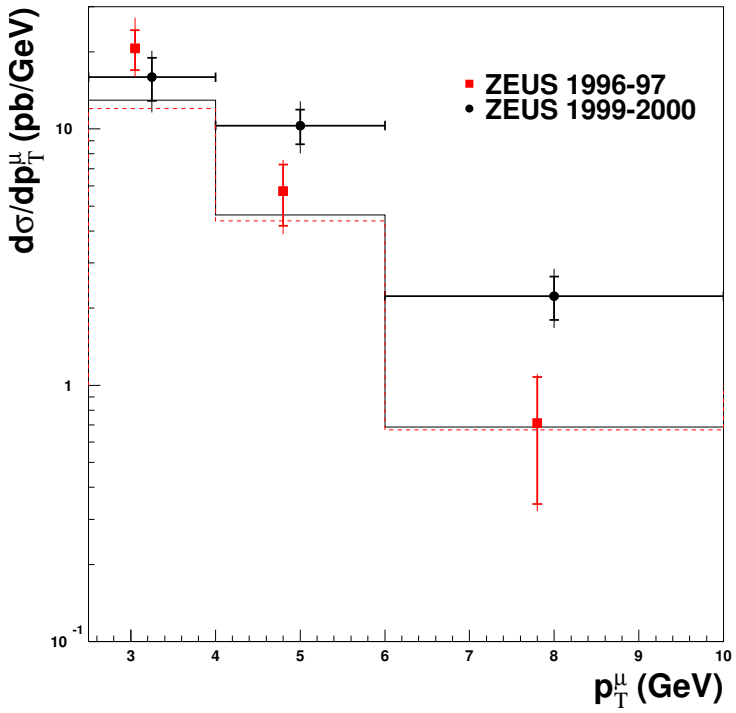

Figure 6.6: *Differential cross section $d\sigma/dp_T^{\mu}$, evaluated using the two data samples available, relative to the 1996–97 and 1999–2000 data–taking periods. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The two histograms represent the* Pythia *predictions: the solid histogram is the Monte Carlo cross section at $\sqrt{s}$ = 318 GeV, the dashed one is the prediction at $\sqrt{s}$ = 300 GeV.*

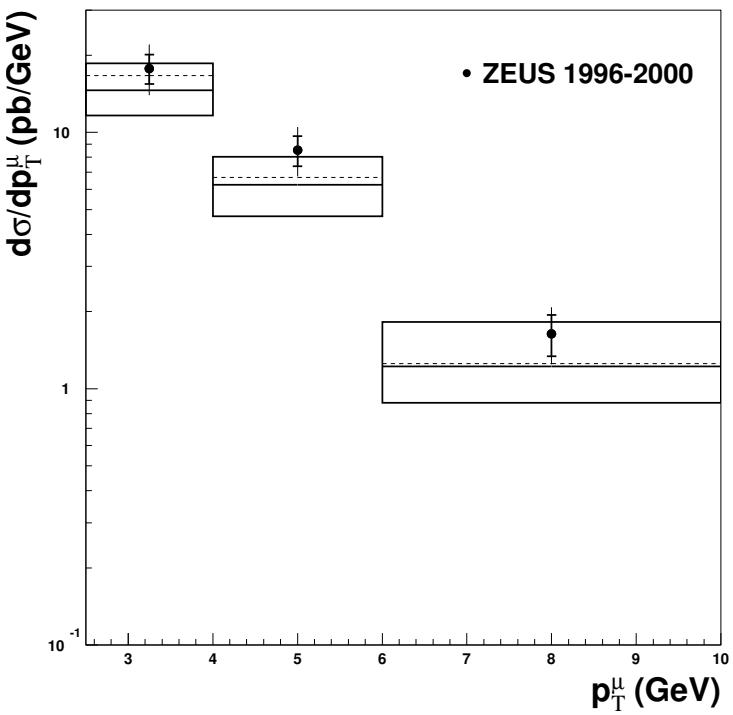

Figure 6.7: *Differential beauty cross section as a function of the muon transverse momentum, p_T^μ, for events with two jets and a muon, compared to the predictions of NLO QCD (dashed line: parton–level jets, full line: corrected to hadron–level jets). The results of the two data samples have been merged as explained in the text. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The band around the NLO prediction shows the range obtained by varying the b–quark mass as well as μ_r and μ_f, as explained in the text.*

Cross sections with muons defined by cuts in 6.5-6.7							
Forward	Barrel	Rear	Dijet	$d\sigma/dx_\gamma^{\rm MEAS}$			
				0.:0.25	0.25:0.5	0.5:0.75	0.75:1.
3.1	1.2	2.0	2.0	3.7	3.8	2.6	1.6

Cross sections with muons defined by cuts in 6.30						
$d\sigma/d\eta^\mu$				$d\sigma/dp_T^\mu$		
-1.75:-0.75	-0.75:0.25	0.25:1.3	1.3:2.3	2.5:4	4:6	6:10
1.6	1.2	1.1	2.3	1.2	1.4	1.3

Table 6.6: *Ratio between the cross sections measured on the data and the predictions of NLO QCD, reported for all the measured cross sections and for every bin of the differential cross sections.*

6.5 Summary and discussion on the results

The fact that the muon differential cross sections reported above are reasonably well described by NLO QCD while the dijet cross sections are not has been investigated recently, and some studies are still undergoing. The ratios between the measured cross sections and the theoretical predictions are summarized in Table 6.6. The measured cross sections have been divided into two groups, depending on the selected data sample used for the calculation. It can be noticed that the discrepancy between data and NLO is higher for the cross sections calculated starting from the largest sample, defined by the cuts in Eqs. 6.5, 6.6 and 6.7. It is interesting to compare the cross sections in the forward, barrel and rear region to the differential cross section $d\sigma/d\eta^\mu$. In this differential distribution, the first bin $(-1.75 : -0.75)$ roughly corresponds to the rear region, the second and the third to the barrel, and the fourth $(1.3 : 2.3)$ to the forward. The data/NLO ratio remains the same in the barrel region since the data sample does not change (cuts in Eq. 6.6 are equal to those in Eq. 6.30). However, in the forward and rear regions the data/NLO ratio reduces by $\sim 25\%$ by increasing the requests on the muon transverse momentum.

Coming to the beauty dijet cross section and to $d\sigma/dx_\gamma^{\rm MEAS}$, if we lower the data/NLO ratios of $\sim 25\%$ to keep into account the effect explained above, the discrepancy between data and NLO does not disappear, in particular in the first three $x_\gamma^{\rm MEAS}$ bins. To understand the residual differences we have to analyse how the cross sections and the NLO predictions have been evaluated.

In the determination of the muon differential cross sections, no Monte Carlo extrapolation has been done in order to compare data and theoretical calculations. The NLO predictions have been determined for events with two jets and a muon in the defined kinematic range, using Peterson fragmentation function and the spectrum of the muon momentum given by Pythia. On the other hand, in the determination of the beauty dijet cross section and of $d\sigma/dx_\gamma^{\rm MEAS}$, the Monte Carlo has been used to extrapolate the measured cross section to the kinematic region defined in Eq. 6.25. The NLO calculations and the Pythia Monte Carlo do not use the same fragmentation model: Pythia uses the parton shower to describe the perturbative part of the b–quark fragmentation, and at the end of the parton shower the string containing the b–quark is

fragmented producing a B–hadron and a left–over string. The energy sharing between the string and the hadron is extracted from a fragmentation function $f(z)$. The default function used by PYTHIA is the Bowler function:

$$f(z) = z^{-(1+bm_T^2)}(1-z)^a e^{-bm_T^2/z} \tag{6.31}$$

and not the Peterson. Here, a and b are parameters to be extracted from fits to real data.

Possible explanations of the different observed data/NLO ratios are therefore the different fragmentation models or different jets distributions between PYTHIA and NLO, since the $B \to \mu$ decay is in both cases described by PYTHIA.

From recent preliminary studies [123] it has been noticed that the fragmentation does not seem to produce the main differences between PYTHIA and NLO. A significant discrepancy has been found in the simulation of excitation processes: in PYTHIA there is an excess of low p_T B–hadrons with respect to NLO, due to flavour excitation processes in which just one jet is produced, the other being a light–flavoured jet. The significant presence of these hadrons traduces in an excess of low p_T muons, mainly in the low $x_\gamma^{\rm MEAS}$ region.

6.6 Determination of the systematic uncertainties

The systematic uncertainties affecting the determination of the cross sections that have been taken into account in this work are the following (the effects of the various contributions to the systematic error are shown in Fig. 6.8 and 6.9 for the different measured cross sections):

- **calorimeter energy scale:** as we pointed out in Sect. 4.7.1, there is an uncertainty on the absolute calorimeter energy scale, in the sense that the calorimeter response to data is not perfectly reproduced in Monte Carlo events. This effect has been corrected by applying a multiplicative factor to the energy measured in the data. The effect of the uncertainty on the calorimeter energy scale has been evaluated by changing the energy measured in the Monte Carlo by $\pm 5\%$. The acceptance has then been re–evaluated in each bin of the measured cross sections, and the cross sections recalculated without redoing the fit. When it is possible, in fact, we will avoid to redo the fit for the estimation of the systematic uncertainties: this is because the statistical error of the fit is large and in some cases the number of events is small. Therefore, a small variation in the number of events in the input distributions can cause a large variation in the fitted beauty fraction. Nevetheless, this variation is strictly related to the statistics available, and already estimated by the statistical uncertainty of the fit. We will therefore avoid, when possible, to mix it with the systematic uncertainty, whose origin is different.

  The systematic error on the measured cross sections caused by the variation of the calorimeter energy scale has been found of the order of 5–10%;

- **uncertainty on the estimation of the muon chambers performances:** as it was pointed out in Sects. 4.9.1 and 4.9.2, the performances of the muon detectors on data and Monte Carlo are different, and the ratio between the two has been evaluated with its error. The summary of the correction factors applied to the Monte Carlo in order to simulate the real muon detectors performances are reported in Table 6.3, for the different angular regions and for the two data–taking periods. The error on the measured data/MC ratio has been included as a systematic uncertainty on the cross sections determination. The acceptance has been re–evaluated using the extreme values of the ratios, and the cross sections recalculated with these new acceptances, without redoing the fit. The effects found are of the order of 15%;

- **model dependence: HERWIG:** to estimate the dependence of the cross sections on the theoretical model used to evaluate them, the HERWIG Monte Carlo has been used to recalculate the acceptances. The cross sections have then been re–evaluated with the new acceptances, without performing again the fit. The difference between the obtained cross sections and the central values has been considered as a systematic uncertainty. The obtained effects vary between a few percent and 25% depending on the considered cross section or bin, but are usually between 5 and 10%. Different behaviours in some bins in 1996–97 with respect to 1999–2000 are mainly due to a lack of Monte Carlo statistics since the bins giving problems are those with the lowest number of events;

- **model dependence: charm/light reweighting:** in order to keep into account the uncertainty on the ratio between the charm and light flavour photoproduction cross sections, the charm cross section has been varied by a factor four, from half to twice the nominal value given by the Monte Carlo and reweighted as explained in Sect. 5.4. In this way the $p_T^{\rm rel}$ distribution of the background has changed in shape, and consequently the beauty and background fractions found by fitting the data change. The effect given by this reweight is usually lower than 5%. The different behaviour in 1996–97 with respect to 1999–2000 in the last bin of the differential cross section $d\sigma/dp_T^\mu$ is due to the small statistics of that bin, and to the consequent unstability of the fit;

- **model dependence: use of the nominal Monte Carlo cross sections:** the difference between the central value of the measured cross sections, obtained by reweighting the nominal Monte Carlo cross sections as explained in Sect. 5.4, and the values obtained by using the nominal cross sections given by PYTHIA has been included as a systematic uncertainty. The obtained effect is relevant only for the extrapolated cross sections, $\sigma(e\,p \to \text{dijet}\ X)$, where it can be of the order of 5–15%; it is much lower for the muon cross sections;

- **cut on the matching probability:** the cut on the matching probability for both the muon reconstruction algorithms has been slightly varied (from 0.05 to 0.04 and 0.06) so that the number of selected events changes a bit. The $p_T^{\rm rel}$ fit has then been redone and the cross sections recalculated. The differences found with respect to the central values are small, usually of the order of 5%, larger in the bins with low statistics where the fitted beauty fraction has a large error;

- **fit method:** all the fits have been redone using a CERN routine, HMCMLL [124], using a binned maximum likelihood fit which includes the effect of both data and

Monte Carlo statistics. The obtained results are usually in nice agreement with those from the χ^2 fit, the difference being lower than a few percent and much lower than the statistical error; some larger discrepancies are found for fits having low statistics (rear region, first bin of $d\sigma/d\eta^{\mu}$). The fit in the first bin of $d\sigma/dp_T^{\mu}$ for 1996–97 is unstable and the values given by the two fits seem both reasonable, even if they differ by $\sim 20\%$ (the statistical error from the fit is $\sim 18\%$);

- **error on the measured luminosity:** the integrated luminosity collected by the ZEUS experiment is known with a certain error, which is of 2.25% on the 1999–2000 data sample and 1.8% on the 1996–97 data sample. The uncertainty on the measured cross sections arising from this error has been included in the systematic indetermination;
- **branching ratio $\mathcal{B}(b \to \mu)$:** the uncertainty on the branching ratio of beauty semi–leptonic decay into muons, $\mathcal{B}(b \to \mu) = (0.220 \pm 0.011)$ [96], has been included as a systematic indetermination on the measured cross sections.

The main contribution to the systematic uncertainty has been found to come from the uncertainties on the muon chambers efficiency and from the model used, PYTHIA or HERWIG.

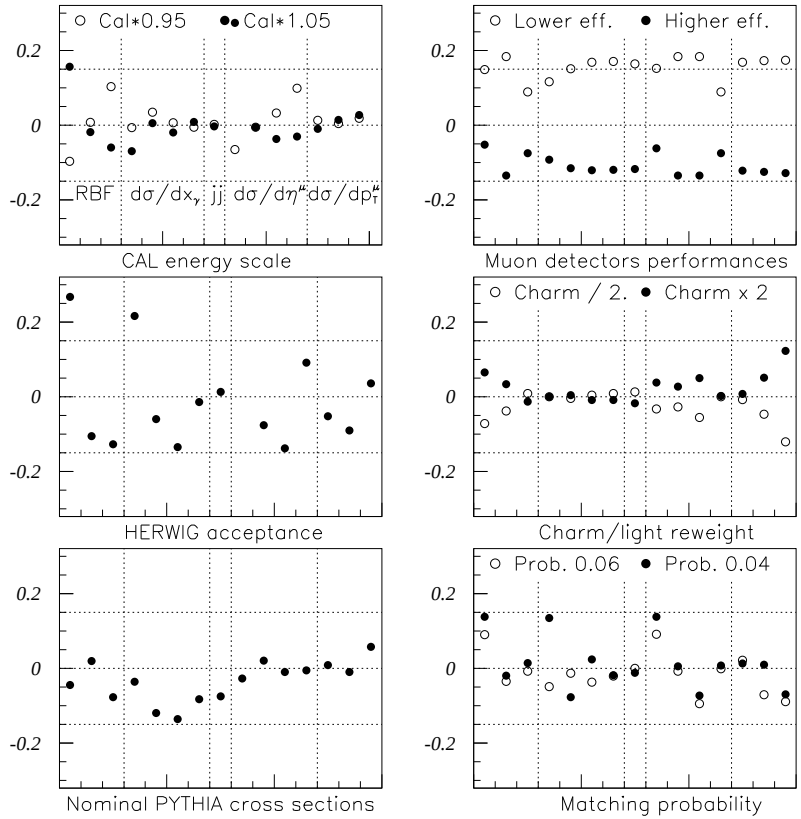

Figure 6.8: *Systematic uncertainties on the measured values of the cross sections, shown for all the measured cross sections and for some of the different considered contributions. In each histogram the systematic uncertainty is shown for the cross sections in three bins (RBF, Rear, Barrel, Forward), for the differential cross section $d\sigma/dx_{\gamma}^{\mathrm{MEAS}}$, for the beauty dijet cross section (jj), and for the two differential muon cross sections, $d\sigma/d\eta^{\mu}$ and $d\sigma/dp_T^{\mu}$. The considered contributions are described in detail in the text. The evaluation is referred to the* **1996–97** *cross sections.*

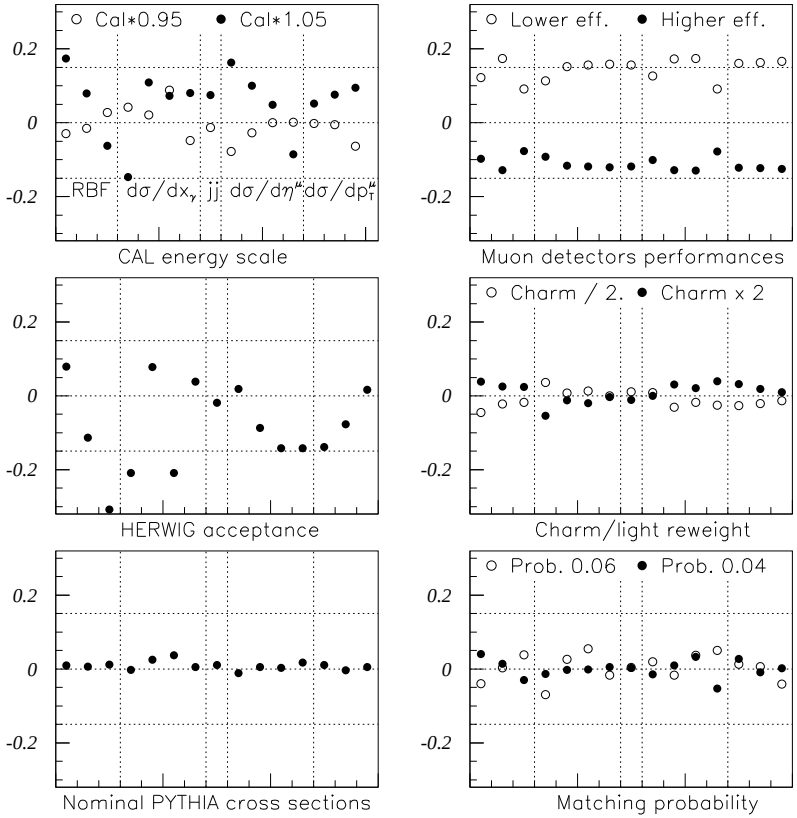

Figure 6.9: *Systematic uncertainties on the measured values of the cross sections, shown for all the measured cross sections and for the different considered contributions. In each histogram the systematic uncertainty is shown for the cross sections in three bins (RBF, Rear, Barrel, Forward), for some of the differential cross section* $d\sigma/dx_\gamma^{\rm MEAS}$ *, for the beauty dijet cross section* (jj)*, and for the two differential muon cross sections,* $d\sigma/d\eta^\mu$ *and* $d\sigma/dp_T^\mu$*. The considered contributions are described in detail in the text. The evaluation is referred to the* **1999–2000** *cross sections.*

Conclusions

In this thesis a study of beauty photoproduction with semi–leptonic decay into muons has been reported, using the data collected by the ZEUS detector at HERA. The analysed sample corresponds to the ~ 100 pb^{-1} of data collected between 1996 and 2000. Beauty visible and differential cross sections have been measured and compared to NLO QCD predictions.

All the measured cross sections have been evaluated in the kinematic region defined by:

$$\begin{aligned} Q^2 &< 1 \text{ GeV}^2 & 0.2 < y &< 0.8 \\ p_T^{\text{jet1(2)}} &> 7(6) \text{ GeV} & |\eta^{\text{jet1(2)}}| &< 2.5. \end{aligned} \tag{7.1}$$

Therefore, the photoproduction regime has been investigated, in the kinematic region characterised by the presence of two high–transverse momentum jets.

The visible cross section of the process:

$$e\,p \to \;\; b\,\bar{b}\,X \;\; \to \mu \text{ dijet } X$$

has been evaluated in the three angular regions covered by the ZEUS muon detectors, defined by:

- **forward region:**

$$p^\mu > 4 \text{ GeV}; \;\; p_T^\mu > 1 \text{ GeV}; \;\; 1.5 < \eta^\mu < 2.3;$$

- **barrel region:**

$$p_T^\mu > 2.5 \text{ GeV} \;\; ; \;\; -0.9 < \eta^\mu < 1.3;$$

- **rear region:**

$$p^\mu > 2.5 \text{ GeV} \;\; ; \;\; -1.6 < \eta^\mu < -0.9.$$

The obtained results have been found higher but in agreement within errors with NLO QCD predictions. Some slightly higher discrepancy has been observed in the forward region, where resolved photon processes, mainly flavour excitation phenomena, are expected to dominate. The difference between data and NLO in the forward region is around 1.8 σ.

The cross section of the process:

$$e\,p \to \;\; b\,\bar{b}\,X \;\; \to \text{dijet } X$$

has been measured in the kinematic region defined by the cuts in Eq. 7.1, starting from the muon sample used for the visible cross section determination, and using the Monte Carlo to extrapolate for muon decay and branching ratio. The obtained result is:

$$\sigma(ep \to b\bar{b}X \to \text{dijet } X) = 740 \pm 55 \text{ (stat.)}^{+130}_{-98} \text{ (syst.) pb},$$

to be compared to the NLO prediction of:

$$\sigma_{\text{NLO}} = 381\ ^{+117}_{-78} \text{ pb}.$$

In this case the discrepancy between data and NLO is about 2.2 σ.

The differential cross section $d\sigma/dx_\gamma^{\text{MEAS}}$ has been evaluated in the same kinematic range as the beauty dijet cross section, to see where the discrepancy between data and NLO comes from. The NLO predictions have been found to underestimate the measured differential cross section in both the resolved ($x_\gamma^{\text{MEAS}} < 0.75$) and the direct ($x_\gamma^{\text{MEAS}} > 0.75$) part, even if the discrepancy in the resolved part is more pronounced.

Differential beauty cross sections as a function of the muon pseudorapidity, η^μ, and transverse momentum, p_T^μ, have been evaluated in the kinematic region defined in Eq. 7.1, plus the following cuts on the muon:

$$p_T^\mu > 2.5 \text{ GeV} \qquad -1.6 < \eta^\mu < 2.3.$$

The obtained differential cross sections agree within errors with NLO QCD predictions.

Studies on why some distributions are described by NLO and other are not are still undergoing. Differences in the obtained results are due to the different starting sample, defined by the two sets of cuts on the muons, and to the theoretical model used to describe the fragmentation in the PYTHIA Monte Carlo and in NLO QCD predictions. Inputs are needed from theoreticians in order to decide which is the best fragmentation model for our data. Nevertheless, having discovered some possible reasons why in some cases NLO QCD does not succeed in explaining the measured beauty cross section is a step forward in the understanding of the beauty puzzle.

In Fig. 7.1 the ratios between the measured beauty cross sections at HERA and the NLO QCD predictions are reported. The result of this thesis, indicated as the full square in the low–Q^2 region, has been obtained using the measured beauty dijet cross section. The other points have been obtained from measurements in different kinematic regions and with different extrapolation techniques. A comparison of all these results, including the analysis of the applied cuts and of the theoretical models used in the extrapolations, can help in understanding which are the points in which the theory fails, and, consequently, can bring to the solution of the beauty puzzle at HERA.

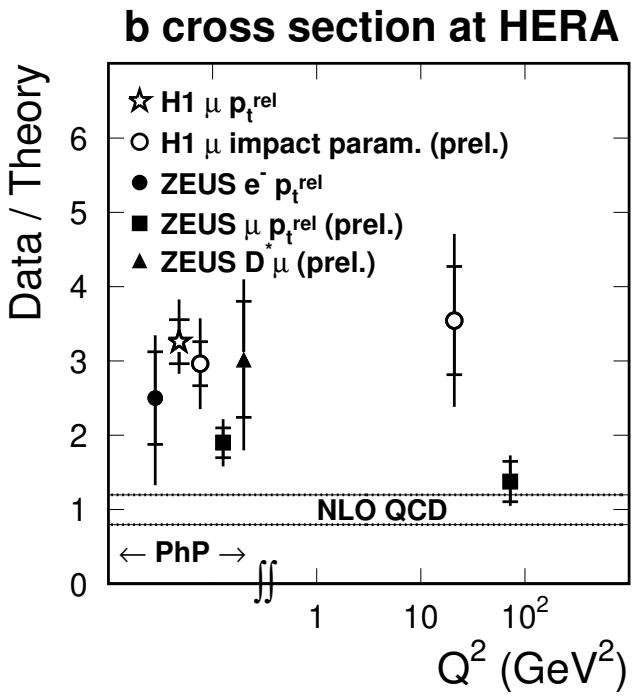

Figure 7.1: *Ratio of measured b–production cross sections at HERA and the theoretical expectation from NLO QCD, as a function of Q^2. The full error bars are the quadratic sum of the statistical (inner part) and systematic uncertainties. The full square in the low–Q^2 region corresponds to the value obtained from the beauty dijet cross section reported in this thesis.*

Appendix A

Muon chambers performances

In the analysis of physics data, the correct understanding of the sub–detectors used for data reconstruction is essential. In particular, in the determination of the correction factors for the cross sections calculations, any existing difference between the events simulated with Monte Carlo programs and the real data has to be analysed and understood completely. In this Chapter we will focus on the Barrel and Rear muon chambers, whose performances will be investigated in detail.

A.1 Monte Carlo simulation of the BRMUON chambers

The Barrel and Rear Muon Detector is simulated in the ZEUS detector and trigger simulation program, MOZART (see Sect. 3.4) which includes information on the detector geometry, shape and constituent material. Therefore, the inefficiency of the detector due to its geometrical acceptance (angular coverage, cracks between the chambers) is correctly implemented in the simulation programs. Nevertheless, there are other sources of inefficiencies, that would be absent in a perfect detector:

- during data taking, the high–voltage of the detector is not kept at the values at which the particle detection efficiency is maximal, but at slightly lower values. This is because the lifetime of the detector, if the voltage is kept at this slightly lower value, is expected to be much longer, and therefore a compromise between efficiency and durability was chosen;
- during the years, some of the limited streamer tubes can break, giving an inefficiency in the region they previously covered;
- the general performances of the electronics can deteriorate in the years.

There are two possible choices for the simulation of the performance of a detector: these sources of inefficiencies can be studied in detail on real data, and then implemented in a Monte Carlo version suitable for physics analyses, or they can be studied *a posteriori*, giving a correction factor to be applied to the measured cross sections. Both the solutions have negative and positive aspects: a correct mapping of all the dead tubes of the detector, joined with an estimation of the decrease in efficiency due to the lower value of the high–voltage chosen and to the changes in the electronics

performances, all implemented in an event simulation, would be probably a more precise solution, but more time consuming, since the generation of Monte Carlo events is in this case postponed after the complete analysis of the detector performance. The estimation of the correction factors on real data is less time consuming and gives a sufficient precise determination of the detector performances. This is the motivation why, for the results reported in this work, the second method was chosen, bearing in mind that a complete mapping of the dead tubes can be used as an independent method (and also as a cross check) for the performance determination.

A.2 Method

In this Sections the results are described concerning the performance of the muon reconstruction algorithm for the barrel and rear regions, BREMAT [109], used in the data analysis reported in this thesis (see Sect. 4.3.2). The algorithm uses information of both the inner and outer muon chambers, therefore the performances of the reconstruction package are strongly related to those of the Barrel and Rear Muon Detector.

For the determination of the performances of the Barrel and Rear Extrapolation MATching package, BREMAT, the idea is quite simple: having chosen a sample of very clean events (two–track events) we select in each event one track that gives the trigger signal, and use the other to analyse the muon reconstruction algorithm performance. This second track is asked to be in an angular region different from that of the first, in order to choose an unbiased sample for the performance determination.

The appropriate trigger slot to be used to select such events is a third level trigger bit used for the selection of isolated muon events:

- **GLOMU TLT bit:** the GLOMU trigger bit is defined by a quite complex algorithm, available at the Third Level Trigger of the ZEUS experiment, that combines the information coming from the Central Tracking Detector (see Sect. 2.2.1), the Calorimeter (Sect. 2.2.2) and the Barrel and Rear Muon Detector (2.2.5). A good track is required to be reconstructed by the CTD, an energy deposit has to be found in the calorimeter, consistent with the passage of one single minimum ionizing particle, and an activity has to be present in the inner muon chambers. A fit is performed, using the direction of the track as reconstructed by the CTD, the energy deposit in the calorimeter and the hits in the muon chambers, and good candidates, having the χ^2 of the fit lower than a certain (fixed) value, are selected as muon candidates.

On the event sample selected requiring the GLOMU TLT bit to be fired, a screening is needed in order to determine which runs are good for the analysis and which are not. The selection of good runs is based on two main criteria, explained in more detail in the following: the calculation of the GLOMU trigger rate and the determination of the runs in which the detector or part of it were not working properly. These two criteria are only partially correlated, as it will be shown later. After this screening, all the runs defined as *good* will be used for the performance determination.

The BREMAT performance is defined in the following way:

$$\text{performance} = \frac{\text{number of reconstructed tracks}}{\text{total number of tracks}} \qquad \text{(A.1)}$$

where *reconstructed* means identified as good by the BREMAT algorithm, in the sense that will be clarified in the following Sections. The *total number of tracks* is given by all the tracks pointing toward the angular region under investigation, as stated by the CTD. In this way, the performance represents a convolution between the geometrical acceptance of the detector (as an example, inefficiencies due to the cracks between the chambers are included in the quantity defined above) and the proper detector and algorithm performance. The proper detector performance is determined, as was shown above, by the detector type, its high–voltage, the number of dead channels, the thresholds, and so on. The proper algorithm performance is determined by the goodness of the software itself, the cuts applied, the correct implementation of the detector geometry and alignment, etc.

A.3 Selection of good runs

In this Section, the data under investigation are those collected during 1998–2000 data–taking period. The selection of runs suitable for physics analyses in the previous period had already been done [125]. As was pointed out above, the main criteria for rejecting a run are a bad rate of events triggered by the GLOMU TLT bit, or a total or partial malfunction of the apparatus. The two circumstances are not fully correlated, since the number of GLOMU events, and hence the rate, is not so large to be sensitive to small variations due to rather small parts (for example, a chamber) of the detector not working properly. Nevertheless, runs in which small parts of the detector were known to have not worked properly were rejected, in order to have the cleanest possible event sample for the evaluation of the algorithm performance. A selection of good runs based mainly on the trigger rate, on the other hand, has to be done with extreme caution, because if runs having bad rates just for statistical fluctuations were rejected, the resulting sample could be biased.

The main reasons driving to the rejection of runs, besides the bad GLOMU rate, were problems occurred to the apparatus after power failures, one or more chambers not working properly because of problems on the high voltage, impossibility to read one or more planes of the chambers because of broken readout cards.

The list of the selected runs, good for physics analyses, relative to 1998–2000 data taking and chosen using the criteria defined above, can be found in [126]. As a check, the rates of the GLOMU trigger bit, after the subtraction of bad runs, were calculated and are shown in Fig. A.1.

In the Figure, the data are divided into the different data–taking periods, 1998 and the first half of 1999, in which HERA was running with electrons, the second half of 1999 and 2000, during which HERA was running with positrons (see Sect. 2.1). The number of bins in each histogram depends on the integrated luminosity collected in the period: each bin corresponds to about $\sim$ 700 nb^{-1} of data. The dots represent the number of collected events divided by the corresponding luminosity. In the Figure it can be noticed that the rate is reasonably flat, and constant during one year, as it should. However, the number of collected events divided by the luminosity lowers during the years: it is ~ 0.9 in 1998 and it goes down to ~ 0.8 in 2000. This is because the use of better triggers has suppressed in the years the number of background events; only a small part of the decrease in the rate depends on the worse performance of the

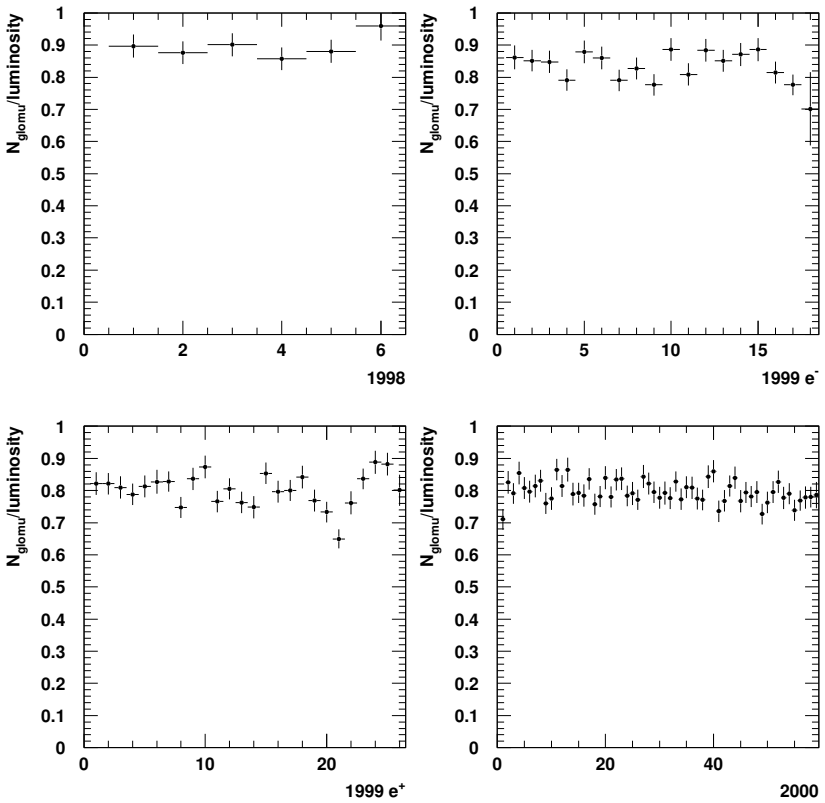

Figure A.1: *Rates of the GLOMU trigger bit, for 1998* **(upper left)**, *1999 electron* **(upper right)**, *1999 positron* **(lower left)** *and 2000* **(lower right)** *data. In each histogram every bin corresponds to a luminosity interval of* ~ 700 nb^{-1}. *The rate is obtained dividing the number of the GLOMU events in each bin by the corresponding luminosity. Bad runs have been rejected.*

muon detector. If tighter cuts are applied to the data before evaluating the rate, so that more background events are rejected, the observed decrease in the rate during the years is much lower: this confirms the fact that the lower rates do not depend very much on the muon detector performances.

A.4 Data analysis

Once the good runs have been chosen, the real data analysis can start. From the selected data a subsample of clean two–track events can be extracted, coming from the elastic production of J/ψ mesons (fig. A.2, left) and from the two–photon process

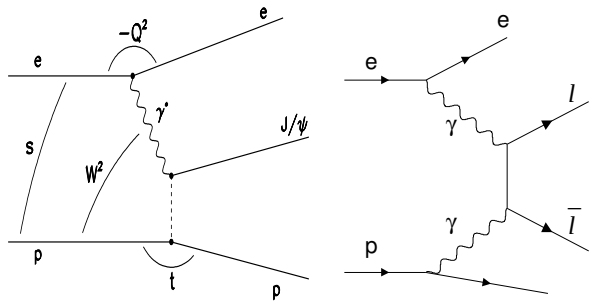

Figure A.2: **Left***: Schematic diagram for the elastic* J/ψ *production process,* $ep \to e\ J/\psi\ p$. **Right:** *Two–photon process* $ep \to \gamma\gamma \to l^+l^-$.

$\gamma\gamma \to \mu^+\mu^-$ (fig. A.2, right). Moreover, a sizable part of the selected data comes from J/ψ and $\gamma\gamma$ events in which the incoming proton dissociates into a different state.

The following cuts were applied to the data, in order to clean the sample and reject the main background, coming from cosmic muon events triggered by GLOMU:

- a muon trigger was asked at all the three trigger levels;
- two and only two tracks fitted to the vertex had to be present in the event;
- a cut on the acollinearity was imposed, $\theta_{\rm acoll} > -0.95$, sufficient to reject most of the cosmics;
- the invariant mass of the dimuon system was asked to be $m_{\mu^+\mu^-} > 2.5$ GeV.

In the analysis, the *barrel* part is defined as the polar angle interval $35^o < \theta < 135^o$ ($-0.9 < \eta < 1.15$) while the *rear* part is defined as $135^o < \theta < 157^o$ ($-1.6 < \eta < -0.9$).

Our data sample was simulated by mixing two different Monte Carlo contributions: the elastic J/ψ production, modelled using the DIPSI [127] program, and the two–photon process, modelled by ZLPAIR [128]. The relative weight of the two contributions was extracted by requiring the Monte Carlo mixture to correctly describe the invariant mass distribution of the dimuon system. In the determination of the BREMAT performances, the fact that the non–elastic component of the data is not simulated does not bias the analysis or the results.

In Fig. A.3 the mass of the $\mu^+\mu^-$ pair is shown using 2000 data, the sample with largest statistics, and is compared to the Monte Carlo distribution. Dimuon events were asked to have an invariant mass greater that 2.5 GeV since the region below that value could not be well simulated by the Monte Carlo. In addition, that region is not very interesting for our aims since it involves very low–momentum muons that can difficultly reach the chambers. The big peak at ~ 3 GeV comes from events in which a J/ψ meson is produced, while the tail at large mass values is well described by the two–photon contribution. The agreement between the data and the Monte Carlo weighted in this way is good.

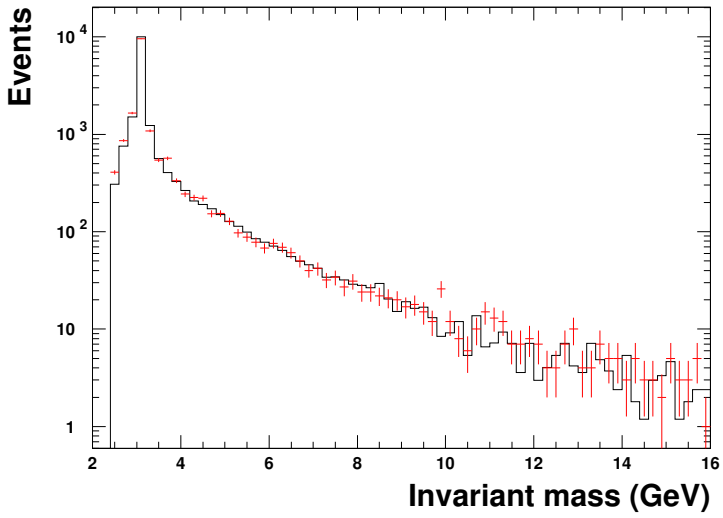

Figure A.3: *Invariant mass of the $\mu^+\mu^-$ system, for 2000 data, compared to the Monte Carlo mixture described in the text, area–normalised to the data. The elastic J/ψ and $\gamma\gamma$ contributions have been weighted in order to reproduce the shape of the distribution.*

Some kinematic variables of the tracks used for the analysis have been plotted and compared to the Monte Carlo, independent of the fact that they are reconstructed by the muon reconstruction algorithm or not. In Fig. A.4 the polar angle distribution of the data is shown, in the barrel and rear regions; the distribution in the barrel region, **(a)**, is plotted for tracks having $p_T > 1.5$ GeV, while in the rear region, **(b)**, no cuts on p have been applied. The Monte Carlo, represented by the shaded band, gives a good description of the data distributions.

In Fig. A.5 the azimuthal angle distribution of the tracks under analysis is shown in the barrel and rear regions and compared to the distribution given by the Monte Carlo. Also in this case in the barrel region the data are plotted for $p_T > 1.5$ GeV, while no cuts have been applied in the rear region. The data description by the Monte Carlo is reasonably good.

The data–Monte Carlo comparisons in Figs. A.4 and A.5, together with the invariant mass distribution in Fig. A.3, make us confident that our description and understanding of the data is good for our aims.

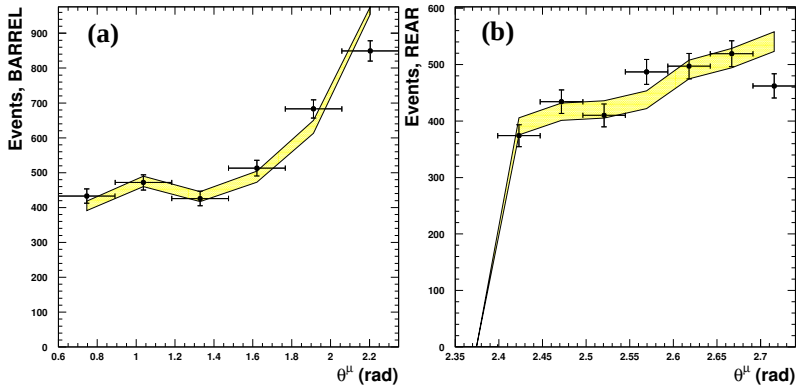

Figure A.4: **(a):** *Polar angle distribution of the track under analysis, for tracks in the barrel region ($35^o < \theta^\mu < 135^o$) having transverse momentum $p_T > 1.5$ GeV. The dots represent the data corresponding to the 2000 data taking, while the shaded band is the Monte Carlo with its error.* **(b):** *Polar angle distribution for the track under analysis, for tracks in the rear region ($135^o < \theta^\mu < 157^o$) and without cuts on the momentum. The shaded band is the Monte Carlo with its errors.*

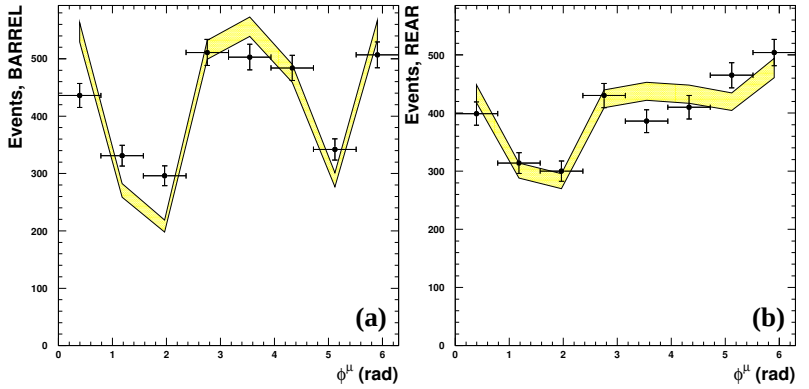

Figure A.5: **(a):** *Azimuthal angle distribution of the track under analysis, for tracks in the barrel region ($35^o < \theta^\mu < 135^o$) having transverse momentum $p_T > 1.5$ GeV. The dots represent the data corresponding to the 2000 data taking, while the shaded band is the Monte Carlo with its error.* **(b):** *Azimuthal angle distribution for the track under analysis, for tracks in the rear region ($135^o < \theta^\mu < 157^o$) and without cuts on the momentum. The shaded band is the Monte Carlo with its errors.*

A.5 BREMAT performance

The data analysis presented in this thesis makes strong use of the BREMAT algorithm, since the data selection is done by requiring the presence of a well reconstructed muon in the forward or in the barrel–rear region. The fact that some malfunctions of the muon detector are not simulated by the Monte Carlo affects the response of the reconstruction package, that consequently has better performances in muon identification on the Monte Carlo than on the data. A complete analysis of the differences between data and Monte Carlo is therefore needed for the full understanding of the data and consequently of the results reported in this thesis.

As it was pointed out in Sect. A.2, since one of the two tracks in the event has to give the muon trigger signal, while the other is used for the determination of BREMAT performances, the second track is asked to be in an angular region different from that of the first. Therefore, some details on the regions that can be distinguished by the trigger have to be illustrated, in order to describe the method used.

One of the tracks in the event is asked to give a trigger signal in the inner muon chambers. This trigger can distinguish between eight sectors in the barrel region and two in the rear. The division of the detector into sectors follows its natural hardware structure: the whole BRMUON is divided into two halves, having as axis the direction of the beam, and being positioned one at negative x values, in the ZEUS coordinate system (see Sect. 2.2) and the other at positive x values. These two parts are called *north* and *south balcony*. Each balcony is composed by twelve chambers, eight in the barrel part and four in the rear part. Each inner chamber has its outer counterpart, covering the same angular region but being positioned outside the iron yoke, except for an outer chamber positioned right below the detector, that has no inner counterpart. Each balcony is divided into four trigger sectors in the barrel region, each corresponding to an inner chamber with its outer counterpart, while the whole rear balcony corresponds to one trigger sector.

In our analysis events having one track in the barrel and one in the rear region can be used, together with events having both tracks in the barrel or in the rear region, providing they hit different trigger sectors.

The sample used for the determination of BREMAT performances is the same used in the data analysis presented in this thesis: it corresponds to the ~ 38 pb^{-1} of data collected during 1996–97 period, and to the ~ 60.5 pb^{-1} of data collected during the positron run in 1999 and the whole 2000. The 1996–97 data sample has been analysed separately from that of 1999–2000 for two main reasons: first, to investigate differences in the behaviour of the detector due to aging or other causes, second, because the Monte Carlo versions to be compared to the data are different, due to slight modifications of the detector and of its simulation.

Events were selected satisfying the cuts defined in Sect. A.4 and having two tracks fulfilling the following requirements:

- *if the first track points to the barrel region and the second to the rear:* if the first track gives a trigger signal in the barrel region, the second is used to analyse the BREMAT performance in the rear region; on the other hand, if the second track gives a trigger signal in the rear region, the first is used for the analysis of the BREMAT performance in the barrel region;

- *both tracks in the barrel region:* the two tracks are asked to point one to the north balcony, the second to the south. If one track gives a trigger signal on a sector of the balcony it points to, the second track is used for the analysis of BREMAT performance;

- *both tracks in the rear region:* the situation is analogous to the previous one; the same method is used, analysing the rear balcony triggers instead of the barrel ones, and checking if BREMAT succeeds in reconstructing the track.

In order to evaluate the BREMAT performances in the same conditions as in the analysis, the track reconstructed by BREMAT is considered good if it fulfills the following requirement:

- the muon reconstructed by BREMAT is asked to have a probability of the 5 d.o.f. matching greater than 0.05. This cut was applied in the analysis in order to reduce the background from fake muon events (K and π coming from particle decays or crossing all the calorimeter and reaching the muon chambers, see Sect. 5.2.1).

The fact that the muon momentum is reconstructed by the muon algorithm means, in practice, that the particle has hits in both the inner and the outer chambers, and that its deflection in the yoke has been determined for the momentum evaluation (see Sect. 2.2.5).

The BREMAT performance is defined, according to Eq. A.1, as the number of good tracks reconstructed by the algorithm and fulfilling the cut on the matching probability, divided by the total number of tracks under analysis.

The BREMAT performance is evaluated as a function of the muon transverse momentum, p_T, in the barrel region and of the momentum, p, in the rear region. The obtained results are shown in Figs. A.6 for the 1996–97 and A.7 for the 1999–2000 data sample, and are summarized in Table A.1. It can be noticed in the Figures that the algorithm starts to reconstruct particles with a significant efficiency for $p_T > 2.5$ GeV in the barrel region and $p > 2.5$ GeV in the rear region. The performance reaches its plateau value for $p_T > 3$ GeV in the barrel and for $p > 3$ GeV in the rear region.

For the analysis purposes, the absolute BREMAT performance is of modest interest. More interesting is the ratio between the algorithm performances as determined on the data and on the Monte Carlo. The plots of the data/MC ratios are shown in Fig. A.8, for the 1996–97, and A.9, for the 1999–2000 data sample, and the numbers are summarized in Table A.2. In the last row of the Table the weighted mean of the data/MC ratios in the last three bins of the plotted distribution is reported. This number is of particular importance since it is used as a correction factor in the evaluation of the cross sections reported in this thesis.

The errors on the correction factors shown in the last row of Table A.2 are only statistical. Systematic checks have been performed in order to understand the stability of the obtained results and to have, at the end, the correct determination of the size of the error:

- the data sample has been limited to events having one track in the barrel and one in the rear region: in this way the statistics slightly reduce;

BREMAT PERFORMANCE						
GeV/c	1996	1997	Mean	1999 e^+	2000	Mean
p_T	**Barrel region**					
0-1	/	0.4 ± 0.2	0.4 ± 0.2	0.4 ± 0.3	0.4 ± 0.2	0.4 ± 0.1
1-2	0.2 ± 0.1	0.6 ± 0.1	0.4 ± 0.1	0.3 ± 0.1	0.3 ± 0.1	0.3 ± 0.1
2-2.5	5.6 ± 2.1	4.8 ± 1.3	5.0 ± 1.1	7.3 ± 1.9	3.8 ± 0.9	4.3 ± 0.8
2.5-3	15.4 ± 5.0	16.8 ± 3.7	16.3 ± 3.0	21.3 ± 4.5	16.0 ± 2.8	17.4 ± 2.4
3-4	34.1 ± 7.1	26.8 ± 5.2	29.3 ± 4.2	24.2 ± 5.2	21.4 ± 3.3	22.2 ± 2.8
4-10	17.9 ± 7.1	35.6 ± 5.1	29.6 ± 4.1	23.5 ± 5.1	25.6 ± 3.4	25.0 ± 2.9
p	**Rear region**					
0-1	/	/	/	/	/	/
1-2	2.0 ± 1.4	/	2.0 ± 1.4	/	0.4 ± 0.3	0.4 ± 0.3
2-2.5	4.9 ± 1.6	5.1 ± 1.2	5.0 ± 1.0	4.6 ± 1.3	3.5 ± 0.7	3.8 ± 0.6
2.5-3	16.6 ± 2.7	23.0 ± 2.1	20.5 ± 1.7	22.3 ± 2.4	16.8 ± 1.4	18.1 ± 1.2
3-4	32.1 ± 3.2	35.0 ± 2.2	34.1 ± 1.8	31.4 ± 2.4	27.5 ± 1.5	28.6 ± 1.3
4-10	21.4 ± 3.8	34.3 ± 3.1	29.1 ± 2.4	27.9 ± 3.3	27.3 ± 2.1	27.4 ± 1.8

Table A.1: *Performance of the reconstruction algorithm for muons in the barrel and rear region, BREMAT [109]. The two different data sets used in the analysis, 96–97 and 99–00, have been analysed separately in order to investigate effects due to detector aging and different running conditions. The weighted mean for each data taking period is reported.*

BARREL			REAR		
	DATA/MC			DATA/MC	
p_T (GeV/c)	1996–97	1999–2000	p (GeV/c)	1996–97	1999-2000
0–1	/	/	0–1	/	/
1–2	/	/	1–2	/	/
2–2.5	0.70 ± 0.20	0.63 ± 0.15	2–2.5	0.71 ± 0.18	0.485 ± 0.092
2.5–3	0.73 ± 0.16	0.74 ± 0.13	2.5–3	0.507 ± 0.051	0.440 ± 0.032
3–4	0.76 ± 0.13	0.612 ± 0.091	3–4	0.567 ± 0.048	0.465 ± 0.023
4–10	0.66 ± 0.11	0.557 ± 0.076	4–10	0.517 ± 0.035	0.475 ± 0.034
2.5–10	0.706 ± 0.074	0.608 ± 0.053	2.5–10	0.539 ± 0.025	0.461 ± 0.016

Table A.2: *Ratio between the BREMAT performance evaluated on the data and on the Monte Carlo, for the two different data–taking periods. The values are reported in bins of the muon transverse momentum, p_T, in the barrel region, and of the momentum, p, in the rear region. In the last row, the weighted mean of the last three bins is reported, to be used as a correction factor for the evaluated cross sections.*

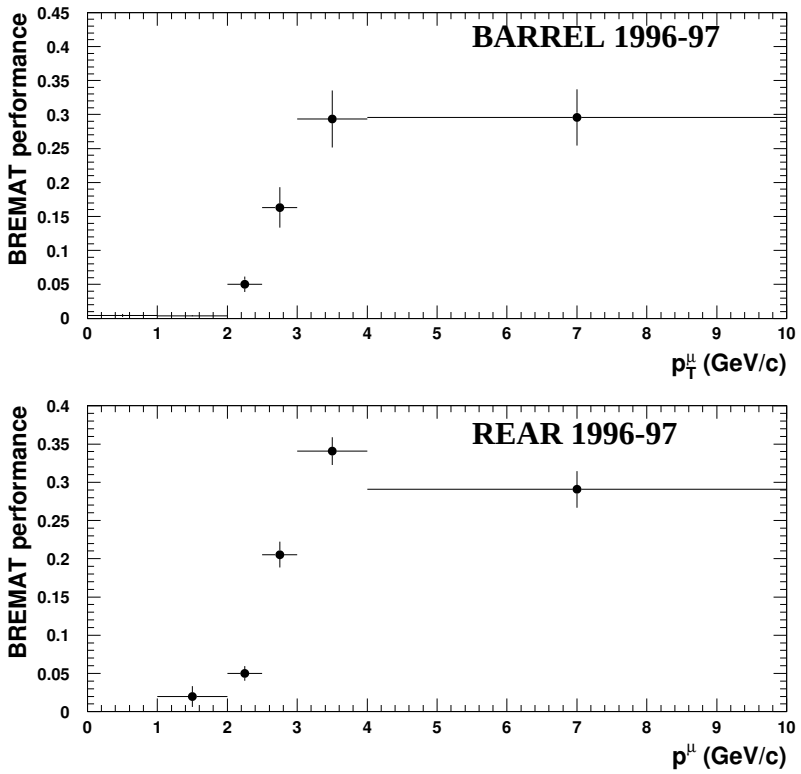

Figure A.6: *BREMAT performance in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1996–97** *data sample.*

- the cut on the acollinearity angle has been changed from -0.95 to -0.995 (this last value is still safe for cosmics rejection);
- the angular range has been restricted, defining the barrel region as $37^o < \theta^\mu < 126^o$ and the rear as $140^o < \theta^\mu < 155^o$;
- the angular range has been enlarged, defining the barrel region as $32^o < \theta^\mu < 135^o$ and the rear as $135^o < \theta^\mu < 160^o$;
- the correction factors have been evaluated also calculating the weighted mean of the data/MC ratios as a function of the polar angle, θ^μ, for $p_T > 2.5$ GeV in the barrel region and $p > 2.5$ GeV in the rear region.

The obtained systematic error has been added in quadrature to the statistical for

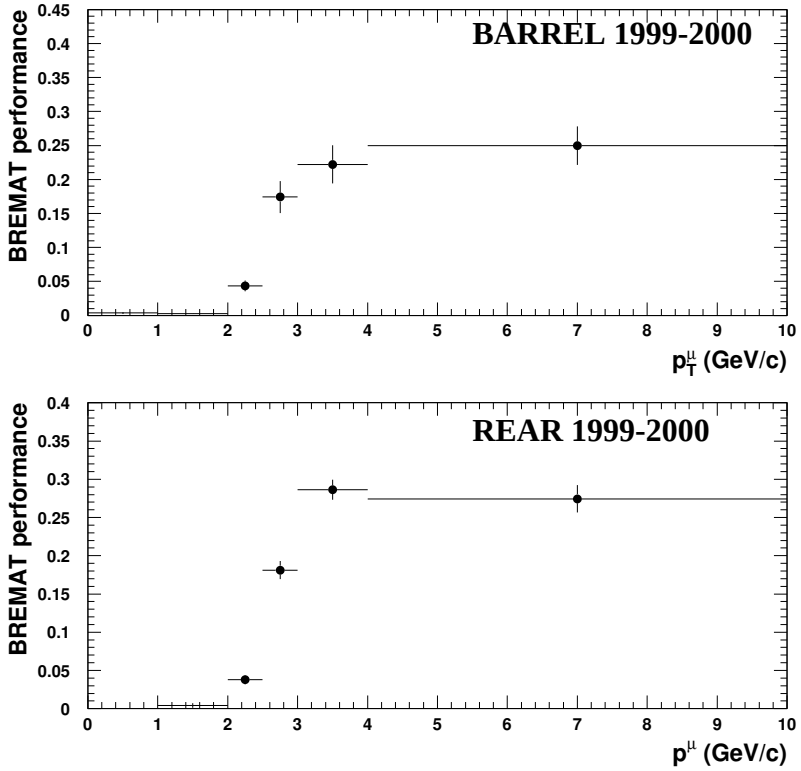

Figure A.7: *BREMAT performance in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1999–2000** *data sample.*

the determination of the total error on the correction factors. The results are:

- **1996–97**:

$$\text{Barrel region}: \quad \frac{Data}{MC} = 0.71 \pm 0.11; \tag{A.2}$$

$$\text{Rear region}: \quad \frac{Data}{MC} = 0.54^{+0.03}_{-0.07}; \tag{A.3}$$

- **1999–2000**:

$$\text{Barrel region}: \quad \frac{Data}{MC} = 0.61 \pm 0.09; \tag{A.4}$$

$$\text{Rear region}: \quad \frac{Data}{MC} = 0.46 \pm 0.05. \tag{A.5}$$

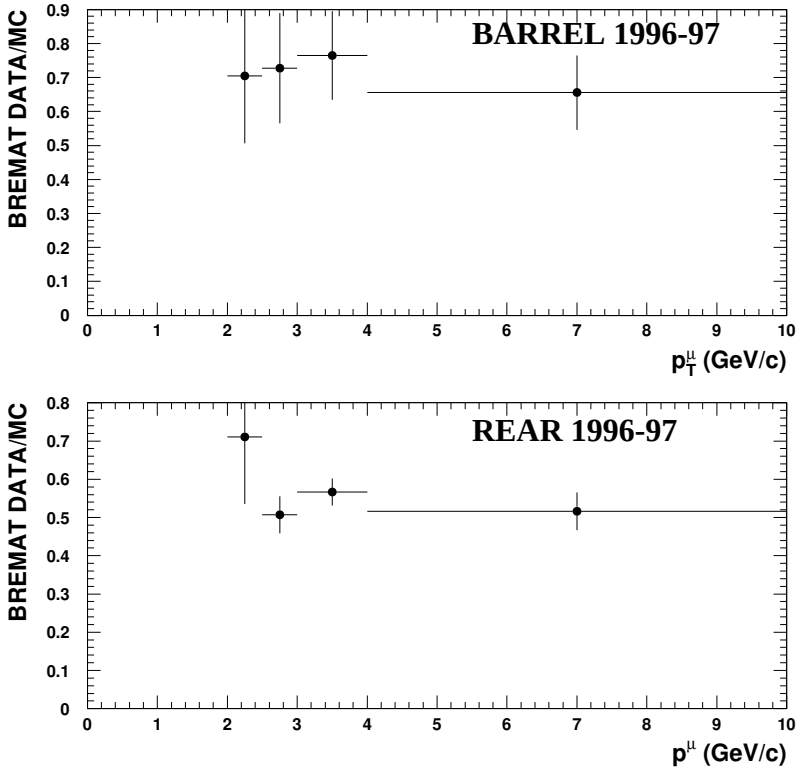

Figure A.8: *Ratio between the performances evaluated on the data and on the Monte Carlo, in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1996–97** *data sample.*

Looking at the last row of Table A.2, it can be noticed that the performances of both the barrel and rear muon detectors are lower in 1999–2000 than in 1996–97. In order to investigate if this was a real detector problem, i.e. if it really depends on an increase in the number of dead tubes and in the electronics malfunctions, the efficiencies of wire and strip planes (see Sect. 2.2.5) were estimated.

Each chamber of the BRMUON detector is made up of four planes, each including a wire plane and a strip plane. The probability to have hits in four, three or less planes of a chamber clearly depends on the efficiency of the planes themselves. Events need to have at least two hits in both the wire and strip planes to be triggered, with a hit in each of the two different plane doublets. Therefore, naming m the efficiency of a wire or strip plane in a chamber, the probability to have four (P_4), three (P_3) or less hits

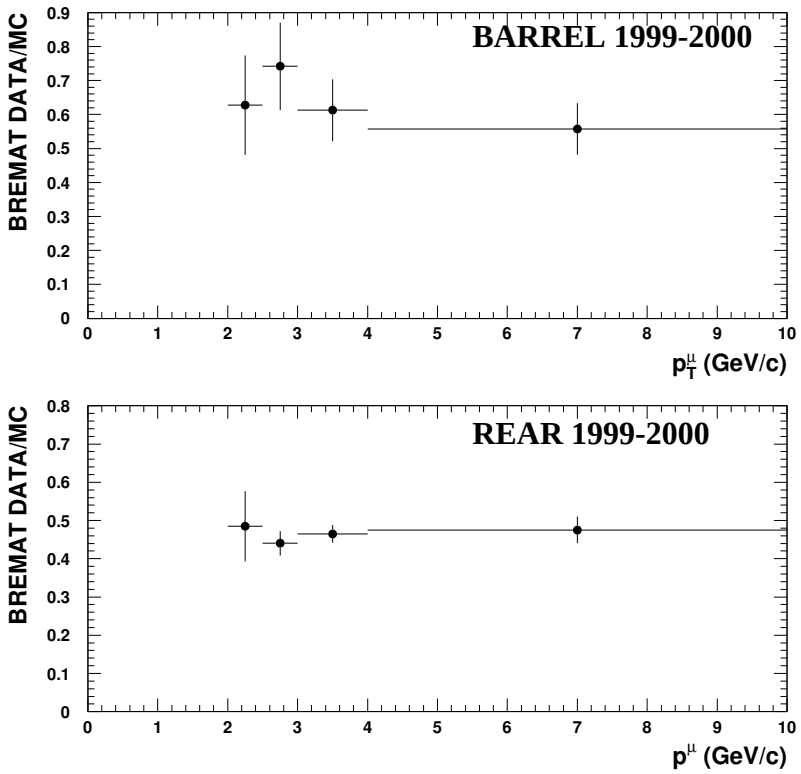

Figure A.9: *Ratio between BREMAT performances evaluated on the data and on the Monte Carlo, in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1999–2000** *data sample.*

(P_2, P_1, P_0) can be expressed as:

- $P_4 = m^4$;
- $P_3 = 4m^3(1-m)$;
- $P_2 = 4m^2(1-m)^2$;
- $P_1 = 0$;
- $P_0 = 0$.

Years	Barrel region		Rear region	
	Effic. wires	Effic. strips	Effic. wires	Effic. strips
1996–97	79.2 ± 0.5	74.8 ± 0.6	75.1 ± 0.6	71.0 ± 0.8
1999–2000	72.3 ± 0.6	65.9 ± 0.8	73.5 ± 0.5	66.1 ± 0.7
Difference	8.7 ± 0.8	11.9 ± 1.0	2.1 ± 0.8	6.9 ± 1.1

Table A.3: *Comparison between the efficiencies of the strip and wire planes in 1997 and 2000. In the last row the relative difference with respect to the 1996–97 value is shown,* (effic 9697 − effic 9900)/effic 9697, *expressed in percent.*

The efficiency m can therefore be extracted from the ratio between the number of events with four hits and the number of events with three hits:

$$\frac{m}{4(1-m)} = \frac{\text{number of events with 4 hits}}{\text{number of events with 3 hits}}. \tag{A.6}$$

The efficiency of the wire and strip planes has been evaluated in the different years, for the inner and the outer muon chambers. The results are shown in Fig. A.10: it can clearly be seen that the tendency is for a decrease in efficiency during the years of both the wire and strip planes, in both the inner and outer chambers. The strips efficiency is lower than that of the wires due to the lower voltage of the induced signal.

In Table A.3 the efficiencies of the wire and strip planes in 1997 and 2000, the samples with highest statistics, are shown. In the last row the relative difference with respect to the 1996–97 value is reported, expressed in percent.

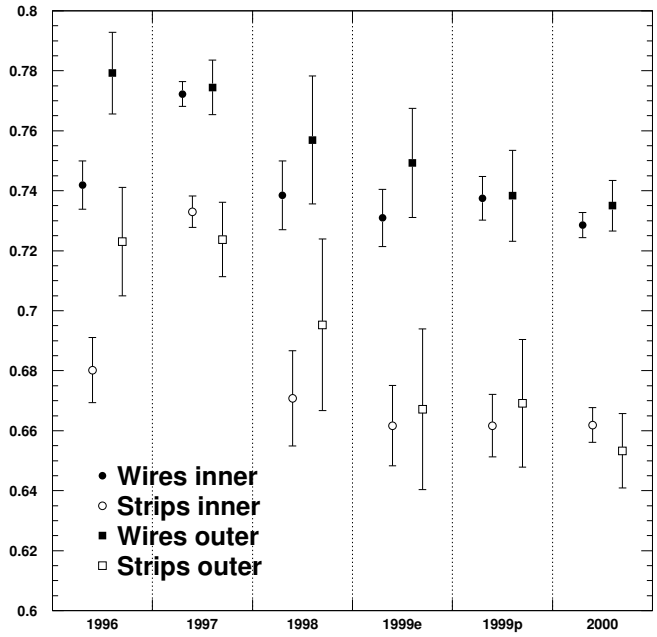

Figure A.10: *Efficiency of the wire (full markers) and strip planes (empty markers) for the inner (circles) and outer (squares) muon chambers, evaluated in the different years of data–taking.*

Appendix B

Pulls of the muon reconstruction algorithms

For the two muon reconstruction algorithms, MPMATCH and BREMAT, the pulls have been evaluated of the variables used for the track fit, i.e. the x and y coordinates in the ZEUS reference frame, the slope in these two coordinates, x' and y', and the inverse of the momentum, Q/p. The pulls are defined as:

$$\text{pull} = \frac{x^{\text{meas}} - x^{\text{pred}}}{cov(x^{\text{meas}}, x^{\text{pred}})} \tag{B.1}$$

where x^{meas} is the measured coordinate of the particle, x^{pred} the one predicted by the fit, and cov is the covariance matrix of the fit, which keeps into account the errors on the measured and the predicted variables and all the correlations. Eq. B.1 is used for the evaluation of the pulls of all the variables used for the fit.

If the algorithms are working properly, the mean of each pull is ~ 0 and its root mean square is ~ 1. This means that there are no systematic shifts between the measured and the predicted variables, and that the errors are evaluated in the proper way.

In Figs. B.1 and B.2 the pulls of MPMATCH and BREMAT are shown, respectively, evaluated on the 1996–97 data sample, after all the cuts of the analysis have been applied. In Figs. B.3 and B.4 the same pulls are shown, evaluated on the 1999–2000 data sample.

The obtained results are quite good. The means are always close to zero, values around 0.1 are found for the momentum distribution, for both MPMATCH and BREMAT, but the means for the coordinates and the slopes are usually lower. The BREMAT pulls seem slightly better than those of MPMATCH. In particular, for BREMAT the root mean squares of the distributions are closer to 1, while for MPMATCH they are around 0.8.

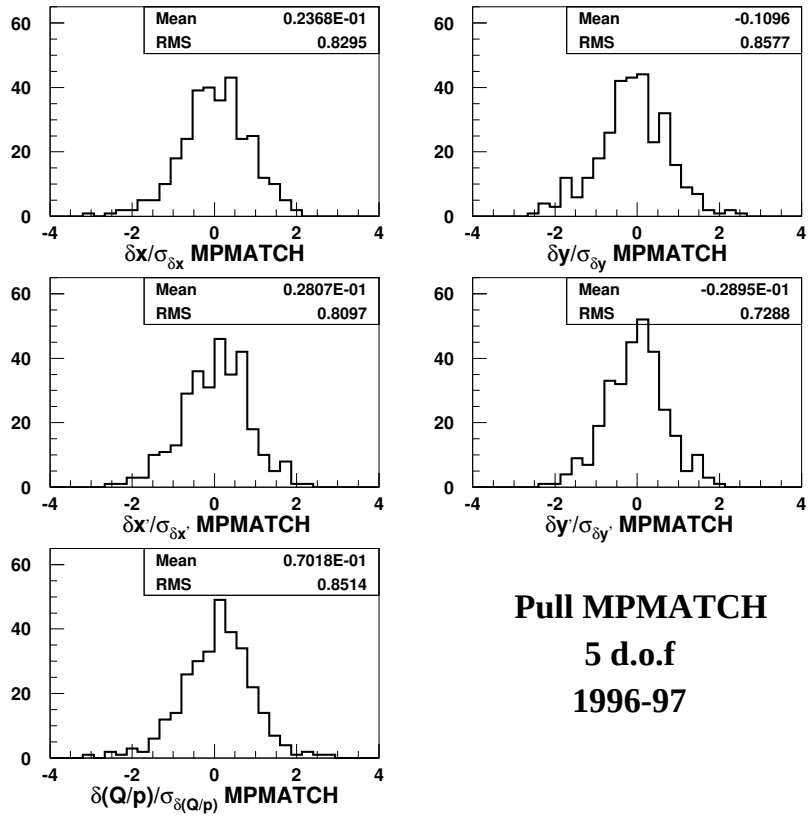

Figure B.1: *Pulls of the variables used by the MPMATCH algorithm to fit the muon track: x and y coordinates, slopes x' and y' and inverse of the momentum Q/p. The pulls have been evaluated on 1996–97 data, after all the cuts of the analysis have been applied.*

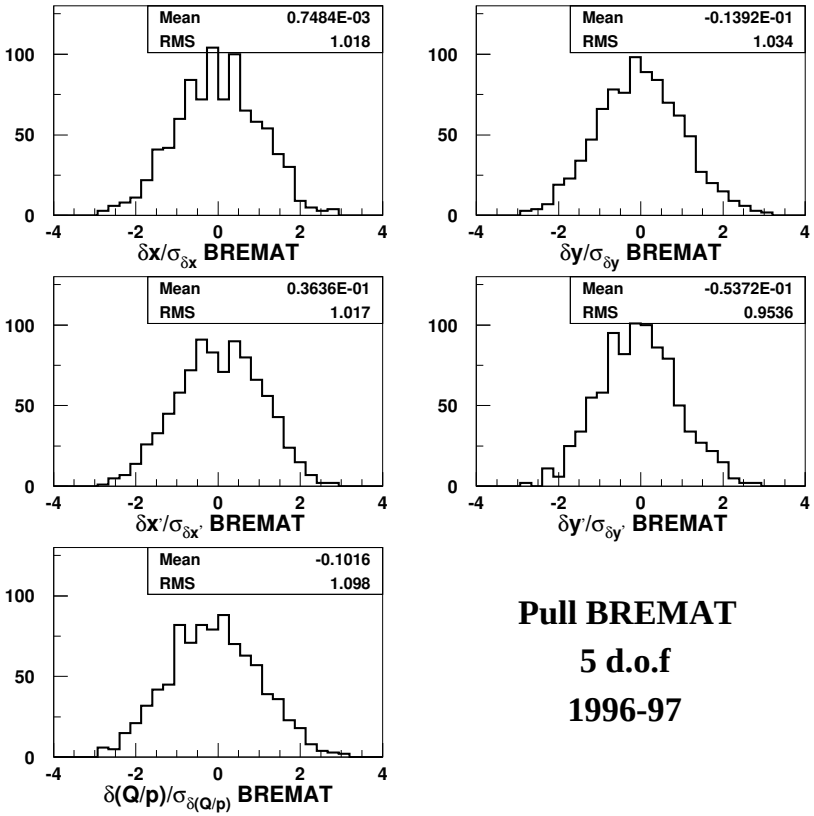

Figure B.2: *Pulls of the variables used by the BREMAT algorithm to fit the muon track: x and y coordinates, slopes x' and y' and inverse of the momentum Q/p. The pulls have been evaluated on 1996–97 data, after all the cuts of the analysis have been applied.*

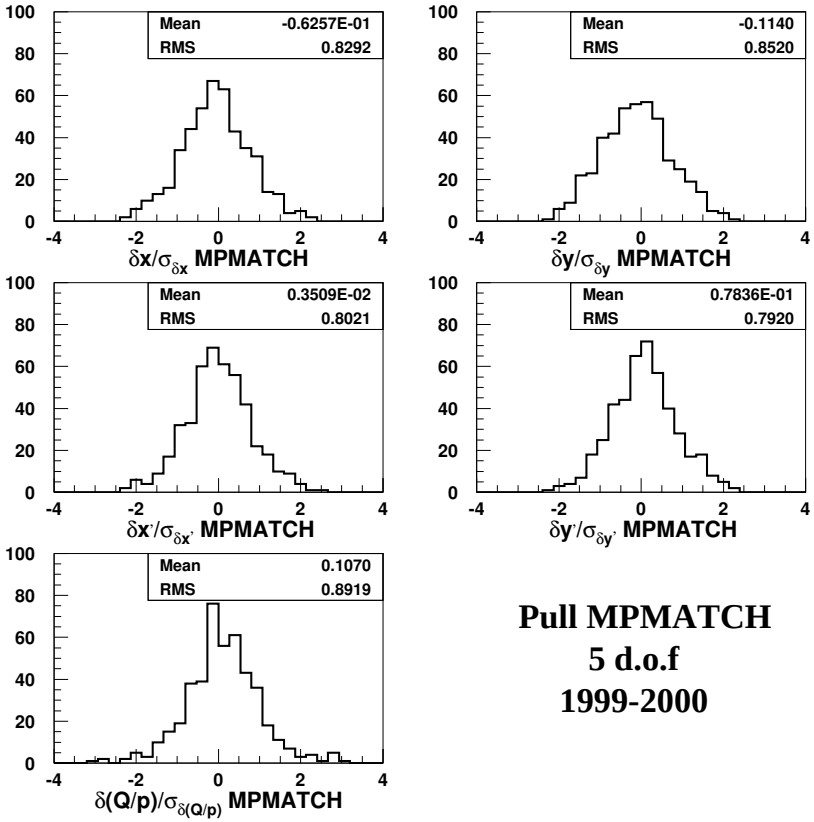

Figure B.3: *Pulls of the variables used by the MPMATCH algorithm to fit the muon track: x and y coordinates, slopes x' and y' and inverse of the momentum Q/p. The pulls have been evaluated on 1999–2000 data, after all the cuts of the analysis have been applied.*

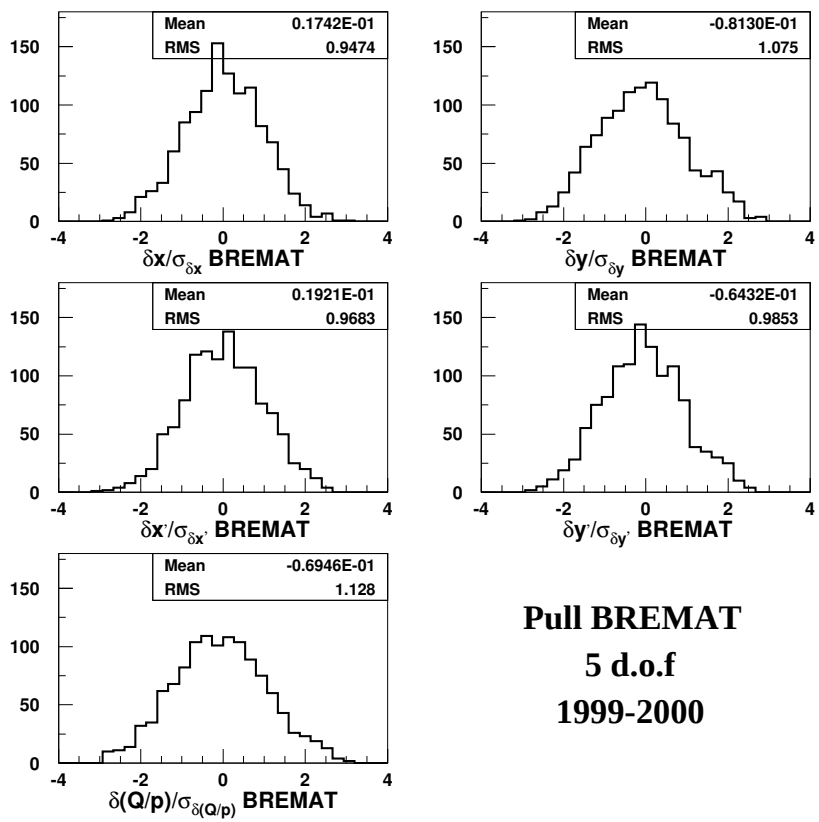

Figure B.4: *Pulls of the variables used by the BREMAT algorithm to fit the muon track: x and y coordinates, slopes x' and y' and inverse of the momentum Q/p. The pulls have been evaluated on 1999–2000 data, after all the cuts of the analysis have been applied.*

Appendix C

Values of the cross sections

In this Appendix, the values of the measured beauty cross sections are reported in Tables together with the NLO QCD predictions. All the cross sections have been measured in the kinematic region defined by:

$$\begin{aligned} Q^2 < 1 \text{ GeV}^2 \quad &; \quad 0.2 < y < 0.8, \\ p_T^{\text{jet}1,2} > 7(6) \text{ GeV} \quad &; \quad |\eta^{\text{jet}1,2}| < 2.5. \end{aligned}$$

The requirements on the muons, when present, will be specified in each case.

C.1 Cross section in the forward, barrel and rear regions

The muon kinematic region is defined by:

- **forward region:** $p^\mu > 4$ GeV; $\quad p_T^\mu > 1$ GeV; $\quad 1.5 < \eta^\mu < 2.3$;
- **barrel region:** $p_T^\mu > 2.5$ GeV; $\quad -0.9 < \eta^\mu < 1.3$;
- **rear region:** $p^\mu > 2.5$ GeV; $\quad -1.6 < \eta^\mu < -0.9$.

	1996–97	1999–2000	Merged	NLO QCD
Forward	$16.0 \pm 5.1\ ^{+2.7}_{-3.0}$	$20.5 \pm 5.1\ ^{+2.6}_{-6.7}$	$18.7 \pm 3.7\ ^{+2.4}_{-5.1}$	$6.1\ ^{+2.8}_{-1.5}$
Barrel	$32.8 \pm 5.3\ ^{+6.4}_{-6.1}$	$46.3 \pm 4.9\ ^{+9.3}_{-8.4}$	$41.0 \pm 3.6\ ^{+8.0}_{-7.5}$	$33.1\ ^{+9.1}_{-7.4}$
Rear	$5.1 \pm 2.9\ ^{+2.4}_{-0.7}$	$10.7 \pm 2.8\ ^{+2.6}_{-1.4}$	$8.5 \pm 2.0\ ^{+2.4}_{-1.1}$	$4.2\ ^{+1.2}_{-1.0}$

Table C.1: *Beauty visible cross sections in the three angular regions of the ZEUS detector. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.*

C.2 Dijet cross sections

The beauty dijet cross section and the differential cross section $d\sigma/dx_\gamma^{\rm MEAS}$ have been measured using the Monte Carlo to extrapolate for muon decay and branching ratio.

C.2.1 Beauty dijet cross section

	1996–97	1999–2000	Merged	NLO QCD
b in dijet	$631 \pm 80\ ^{+111}_{-95}$	$809 \pm 74\ ^{+148}_{-108}$	$740 \pm 55\ ^{+130}_{-98}$	$381\ ^{+117}_{-78}$

Table C.2: *Cross section of the process $ep \to b\bar{b}\ X \to$ dijet X. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.*

C.2.2 $d\sigma/dx_\gamma^{\rm MEAS}$

$d\sigma/dx_\gamma^{\rm MEAS}$	1996–97	1999–2000	Merged	NLO QCD
0 : 0.25	$253 \pm 119\ ^{+73}_{-38}$	$183 \pm 117\ ^{+25}_{-53}$	$210 \pm 85\ ^{+30}_{-36}$	$57\ ^{+28}_{-22}$
0.25 : 0.5	$503 \pm 184\ ^{+85}_{-101}$	$471 \pm 135\ ^{+103}_{-61}$	$483 \pm 109\ ^{+88}_{-66}$	$127\ ^{+94}_{-41}$
0.5 : 0.75	$341 \pm 96\ ^{+63}_{-81}$	$380 \pm 77\ ^{+81}_{-94}$	$365 \pm 60\ ^{+69}_{-82}$	$139\ ^{+75}_{-49}$
0.75 : 1	$1568 \pm 264\ ^{+280}_{-248}$	$2220 \pm 240\ ^{+423}_{-316}$	$1966 \pm 179\ ^{+358}_{-274}$	$1214\ ^{+254}_{-194}$

Table C.3: *Differential cross section $d\sigma/dx_\gamma^{\rm MEAS}$ for two–jet events. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.*

C.3 Muon differential cross sections

The muon kinematic region is defined by:

$$p_T^\mu > 2.5 \text{ GeV}$$
$$-1.6 < \eta^\mu < 2.3.$$

C.3.1 $d\sigma/d\eta^\mu$

$d\sigma/d\eta^\mu$	1996–97	1999–2000	Merged	NLO QCD
-1.6 : -0.75	$2.8 \pm 2.1\ ^{+1.7}_{-0.3}$	$7.2 \pm 2.6\ ^{+1.9}_{-1.0}$	$5.5 \pm 1.8\ ^{+1.6}_{-0.7}$	$3.5\ ^{+1.4}_{-0.7}$
-0.75 : 0.25	$17.5 \pm 3.2\ ^{+3.4}_{-2.9}$	$22.6 \pm 3.1\ ^{+4.8}_{-3.9}$	$20.6 \pm 2.3\ ^{+4.1}_{-3.5}$	$16.5\ ^{+3.8}_{-3.6}$
0.25 : 1.3	$12.6 \pm 3.8\ ^{+2.5}_{-3.0}$	$19.6 \pm 3.6\ ^{+3.9}_{-3.9}$	$16.9 \pm 2.7\ ^{+3.2}_{-3.4}$	$14.7\ ^{+4.6}_{-3.3}$
1.3 : 2.3	$10.2 \pm 3.6\ ^{+1.8}_{-1.0}$	$11.8 \pm 3.4\ ^{+1.5}_{-2.4}$	$11.2 \pm 2.5\ ^{+1.3}_{-1.5}$	$4.9\ ^{+2.3}_{-1.4}$

Table C.4: *Differential cross section $d\sigma/d\eta^\mu$ for two–jet plus muon events. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.*

C.3.2 $d\sigma/dp_T^\mu$

$d\sigma/dp_T^\mu$	1996–97	1999–2000	Merged	NLO QCD
2.5 : 4	$20.6 \pm 3.7\ ^{+5.5}_{-3.0}$	$15.9 \pm 3.0\ ^{+2.9}_{-3.1}$	$17.8 \pm 2.3\ ^{+3.6}_{-3.0}$	$14.6\ ^{+4.0}_{-2.9}$
4 : 6	$5.7 \pm 1.5\ ^{+1.0}_{-1.0}$	$10.3 \pm 1.6\ ^{+2.0}_{-1.6}$	$8.5 \pm 1.1\ ^{+1.6}_{-1.4}$	$6.2\ ^{+1.8}_{-1.5}$
6 : 10	$0.7 \pm 0.4\ ^{+0.1}_{-0.1}$	$2.2 \pm 0.4\ ^{+0.4}_{-0.3}$	$1.6 \pm 0.3\ ^{+0.3}_{-0.3}$	$1.2\ ^{+0.6}_{-0.3}$

Table C.5: *Differential cross section $d\sigma/dp_T^\mu$ for two–jet plus muon events. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.*

Bibliography

[1] T.H. Bauer et al., Rev. Mod. Phys., **50**, 261 (1978); **51**, 407 (1979).

[2] J.J. Sakurai, Ann. Phys., **11**, 1 (1960).

[3] V.M. Budnev et al., Phys. Rep., **15 C**, 181 (1975).

[4] M. Glück and E. Reya, Phys. Rev. **D 28**, 2749 (1983).

[5] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438, 675 (1972);
L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 95 (1975);
Yu.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977);
G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977).

[6] D. Morgan, M.R. Pennington and M.R. Whalley, J. Phys. **G 20** Suppl. 8A, A1-A147 (1994).

[7] K. Muramatsu et al, TOPAZ Collaboration, Phys. Lett. **B 332**, 477 (1994);
S.K. Sahu et al., Phys. Lett. **B 346**, 208 (1995).

[8] R. Akers et al., OPAL Collaboration, Z. Phys. **C 61**, 199 (1994);
K. Ackerstaff et al., OPAL Collaboration, Z. Phys. **C 74**, 33(1997);
P. Abreu et al., DELPHI Collaboration, Z. Phys. **C 69**, 223 (1996).

[9] S. Frixione et al., Nucl. Phys. **B 412** 225 (1994), (`hep-ph/9306337`).

[10] R.K. Ellis et al., Nucl. Phys. **B 152** 285 (1989);
J.C. Collins, D.E. Soper and G. Sterman, in *Perturbative Quantum Chromodynamics*, ed. A.H. Mueller (World Scientific, Singapore, 1989) and references therein.

[11] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977).

[12] E. Witten, Nucl. Phys. **B 120**, 189 (1977);
C.H. Llewellyn Smith, Phys. Lett. **B 79**, 83 (1978);
W.A Bardeen and A.J. Buras, Phys. Rev. **D 20**, 166 (1979);
R.J. De Witt et al., Phys. Rev. **D 19**, 2046 (1979).

[13] S. Frixione et al., Adv. Ser. Direct. High Energy Phys. **15**, 609 (1998) (`hep-ph/9702287`).

[14] G. Bellini, proceedings of *Les Rencontres de Physique de la Vallee d'Aoste*, La Thuile, Aosta Valley, March 6–12, 1994;
J.C. Anjos et al., E691 Collaboration, Phys. Rev. Lett. **62**, 513 (1989);
M.P. Alvarez et al., NA14/2 Collaboration, Z. Phys. **C 60**, 53 (1993).

[15] R.K. Ellis and P. Nason, Nucl. Phys. **B 312**, 551 (1989).

[16] J. Smith and W.L. van Neerven, Nucl. Phys. **B374**, 36 (1992).

[17] P. Nason, S. Dawson and R.K. Ellis, Nucl. Phys. **B 303**, 607 (1988);
P. Nason, S. Dawson and R.K. Ellis, Nucl. Phys. **B 327**, 49 (1989).

[18] W. Beenakker et al., Phys. Rev. **D 40**, 54 (1989);
W. Beenakker et al., Nucl. Phys. **B 351**, 507 (1991).

[19] M.L. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. **B 373**, 295 (1992).

[20] J. Collins, F. Wilczek and A. Zee, Phys. Rev. **D 18**, 242 (1978);
W.J. Marciano, Phys. Rev. **D 29**, 580 (1984).

[21] M. Cacciari and M. Greco, Z. Phys. **C 69** 459 (1996), (hep-ph/9505419).

[22] S. Frixione et al., Phys. Lett. **B 348**, 633 (1995), (hep-ph/9412348).

[23] A.D. Martin, R.G. Roberts and J. Stirling, RAL–94–055, DTP/94/34.

[24] P. Harriman et al., Phys. Rev. **D 37** 798 (1990).

[25] J.C. Collins and R.K. Ellis, Nucl. Phys. **B 360**, 3 (1990).

[26] M. Derrick et al., ZEUS Collaboration, Phys. Lett. **B 345**, 576 (1995).

[27] T. Ahmed et al., H1 Collaboration, Nucl. Phys. **B 445**, 195 (1995), (hep-ex/9504004).

[28] M. Glück, E. Reya and A. Vogt, Phys. Rev. **D 46**, 1973 (1992).

[29] H. Abramowicz, K. Charchula and A. Levy, Phys. Lett. **B 269**, 458 (1991).

[30] M. Derrick et al., ZEUS Collaboration, Phys. Lett. **B 349**, 225 (1995), (hep-ex/9502002).

[31] S. Aid et al., H1 Collaboration, Nucl. Phys. **B 472**, 32 (1996), (hep-ex/9604005).

[32] T. Uematsu and I.F. Walsh, Nucl. Phys. **B 199**, 93 (1982).

[33] F. Borzumati and G.A. Schuler, Z. Phys. **C 58**, 139 (1993).

[34] M. Drees and R.M. Godbole, Phys. Rev. **D 50**, 3124 (1994).

[35] M. Glück, E. Reya and M. Stratmann, Phys. Rev. **D 54**, 5515 (1996), (hep-ph/9605297).

[36] M. Glück, E. Reya and A. Vogt, Z. Phys. **C 67**, 433 (1995).

[37] M. Glück, E. Reya and M. Stratmann, Phys. Rev. **D 51**, 3220 (1995).

[38] M.P. Alvarez et al., NA14/2 Collaboration, Z. Phys. **C60**, 53 (1993).

[39] G. Bellini, Proc. of Les Rencontres de Physique de la Vallée d'Aoste, La Thuile, 1994.

[40] J.C. Anjos et al., E691 Collaboration, Phys. Rev. Lett. **62**, 513 (1989);
J.C. Anjos et al., E691 Collaboration, Phys. Rev. Lett. **65**, 2503 (1990).

[41] J. Breitweg et al., ZEUS Collaboration, Eur. Phys. J. **C6**, 67 (1999).

[42] R.M. Barnett et al., Particle Data Group, Phys. Rev. **D 54**, 1 (1996).

[43] S. Frixione et al., Nucl. Phys. **B 454**, 3 (1995).

[44] A.D. Martin, W.J. Stirling and R.G. Roberts, Phys. Lett. **B 354**, 155 (1995).

[45] C. Peterson et al., Phys. Rev. **D 27**, 105 (1983).

[46] B.A. Kniel et al., Z. Phys. **C 76**, 689 (1997).

[47] J. Binnewies et al., DESY 97–241, Phys. Rev. **D 58** 014014 (1998), (hep-ph/9712482).

[48] M. Cacciari et al., Phys. Rev. **D 55**, 2736 (1997), ibid. 7134.

[49] H.L. Lai et al., Phys. Rev. **D 55**, 1280 (1997).

[50] L.E. Gordon and J.K. Storrow, Z. Phys. **C 56**, 307 (1992).

[51] P. Aurenche et al., Z. Phys. **C 64**, 621 (1994).

[52] G. Marchesini et al., Comp. Phys. Comm. **67**, 465 (1992).

[53] J. Breitweg et al., ZEUS Collaboration, Phys. Lett. **B 481**, 213 (2000), (hep-ex/0003018).

[54] A.V. Berezhnoy, V.V. Kiselev and A.K. Likhoded, Phys. Atom. Nucl. **63** 1595 (2000), Yad. Fiz. **63**, 1682 (2000), (hep-ph/9905555).

[55] C. Adloff et al., H1 Collaboration, Phys. Lett. **B 467**, 156 (1999).

[56] M.H. Seymour, Z. Phys. **C 62**, 127 (1994).

[57] F. Jacquet and A. Blondel, *Proceedings of the Study of an* ep *facility for Europe*, Ed. U. Amamldi, DESY 79/48, 391.

[58] G. Ingelman, J. Rathsman and G.A. Schuler, Comp. Phys. Comm., **101**, 135 (1997).

[59] J. Breitweg et al., ZEUS Collaboration, Eur. Phys. J. **C 18**, 625 (2001).

[60] S. Catani et al., Nucl. Phys. **B 406**, 186 (1993).

[61] H.U. Bengtsson and T. Sjöstrand, Comp. Phys. Comm. **46**, 43 (1987).

[62] D.E. Groom et al., Eur. Phys. J. **C 15**, 1 (2000).

[63] A.D. Martin et al., Eur. Phys. J. **C 4**, 463 (1998);
A.D. Martin et al., Eur. Phys. J. **C 14**, 133 (2000).

[64] H1 Collaboration, *Measurement of the Beauty Production Cross Section at HERA Using Lifetime Information*, abstract 979,982, submitted to the 30th International Conference on High Energy Physics ICHEP 2000, Osaka, Japan, July 2000.

[65] D. Pitzl et al., *The H1 Silicon Vertex Detector*, `hep-ex/0002044`.

[66] R. Barlow and C. Beeston, Comp. Phys. Comm. **77**, 219 (1993).

[67] H1 Collaboration, *Beauty Production in Deep Inelastic ep Scattering*, abstract 807, submitted to the International Europhysics Conference on High Energy Physics EPS 2001, Budapest, Hungary, July 2001.

[68] B.W. Harris and J. Smith, Nucl. Phys. **B 452**, 109 (1995); Phys. Lett. **B 353**, 535 (1995); Phys. Rev. **D 57**, 2806 (1998).

[69] M. Glück, E. Reya and A. Vogt, Eur. Phys. J. **C 5**, 461 (1998), (`hep-ph/9806404`).

[70] ZEUS Collaboration, *Measurement of beauty production in deep inelastic scattering at HERA*, abstract 783, submitted to the 31st International Conference on High Energy Physics ICHEP 2002, Amsterdam, The Netherlands, July 2002.

[71] H. Jung, Comp. Phys. Comm. **86**, 147 (1994).

[72] H.L. Lai et al., CTEQ Collaboration, Eur. Phys. J. **C 12**, 375 (2000).

[73] F. Abe et al., CDF Collaboration, Phys. Rev. Lett. **71**, 500 (1993);
F. Abe et al., CDF Collaboration, Phys. Rev. Lett. **79**, 572 (1997);
F. Abe et al., CDF Collaboration, Phys. Rev. Lett. **75**, 1451 (1995);
B. Abbott et al., D0 Collaboration, Phys. Lett. **B 487**, 264 (2000);
B. Abbott et al., D0 Collaboration, Phys. Rev. Lett. **85**, 5068 (2000).

[74] D. Acosta et al., CDF Collaboration, Phys. Rev. **D 65**, 052005 (2002).

[75] F. Abe et al., CDF Collaboration, Phys. Rev. Lett. **75**, 1451 (1995).

[76] J. Chrin, Z. Phys. **C 36**, 163 (1987).

[77] M. Cacciari and P. Nason, `hep-ph/0204025`.

[78] M. Cacciari and M. Greco, Nucl. Phys. **B 421**, 530 (1994).

[79] M. Cacciari, M. Greco and P. Nason, JHEP **05**, 007 (1998).

[80] M. Cacciari and M. Greco, Phys. Rev. **D 55**, 7134 (1997);
P. Nason and C. Oleari, Nucl. Phys. **B 565**, 245 (2000).

[81] A. Heister et al., Phys. Lett. **B 512**, 30 (2001).

[82] ZEUS Collaboration, *The ZEUS Detector, Status Report 1993*, U. Holm ed. (1993).

[83] *HERA: A Proposal for a Large Electron Proton Beam Facility at DESY* (Hamburg, 1981), DESY–HERA–81/10.

[84] ZEUS Collaboration, *HERA Luminosity Upgrade Report for Funding Agencies*, October 28, 1997.

[85] N. Harnew et al., Nucl. Inst. Meth. **A 279**, 290 (1989),
B. Foster et al., Nucl. Phys. Proc. Suppl. **B 32**, 181 (1993);
B. Foster et al., Nucl. Inst. Meth. **A 338**, 254 (1994).

[86] M. Derrick et al., Nucl. Inst. Meth. **A 309**, 77 (1991);
A. Andresen et al., Nucl. Inst. Meth. **A 309**, 101 (1991);
A. Caldwell et al., Nucl. Inst. Meth. **A 321**, 356 (1992);
A. Bernstein et al., Nucl. Inst. Meth. **A 336**, 23 (1993).

[87] E. Iarocci et al., Nucl. Inst. Meth. **A 217**, 30, (1983).

[88] G. Abbiendi et al., Nucl. Inst. Meth. **A 333**, 342, (1993).

[89] H. Bethe and W. Heitler, Proc. Roy. Soc., **A 146**, 83 (1934);
H. Bethe, Phys. Rev. **89**, 1256 (1953).

[90] W.H. Smith et al., *The ZEUS trigger system*, Zeus-Note 89–084 (1989).

[91] F. Chlebana, Zeus–Note 94–102 (1995).

[92] INMOS Limited, *The Transputer Book*, 2nd edition (1989).

[93] J.C. Collins, D.E. Soper and G. Sterman, Nucl. Phys. **B 308**, 833 (1988).

[94] G.A. Schuler, Nucl. Phys. **B 299**, 21 (1988).

[95] D. Amati and G. Veneziano, Phys. Lett. **B 83**, 87 (1979);
Yu.L. Dokshitzer and S.I. Troian, Leningrad Nuclear Physics Institute preprint N922 (1984);
Ya.I. Azimov et al., Phys. Lett. **B 165**, 147 (1985), Z. Phys. **C 27**, 65 (1985).

[96] T. Sjöstrand et al., Comp. Phys. Comm. **135** 238, (2001).

[97] C.F. Weizsäcker, Z. Phys., **88**, 612 (1934);
E.J. Williams, Phys. Rev., **45**, 729 (1934).

[98] T. Sjöstrand, Nucl. Phys. **248**, 469 (1984).

[99] GEANT 3.13, R. Brun, CERN DD/EE 84-1 (1984).

[100] ZGANA, ZEUS trigger simulation library, Els de Wolf (ed.) et al.

[101] S.M. Fisher and A. Palazzi, *The ADAMO Data System Programmmer's Manual.*

[102] J. Huth et al., Proceedings of the DPF Summer Study on High Energy Physics, 1990, ed. L. Berger, 134 (1992).

[103] P.J. Bussey, EUCELL, ZEUS *Phantom* Users Library.

[104] P. de Jong, *Status of the Uranium Calorimeter Reconstruction Software*, ZEUS–Note 92-019.

[105] M. Seymour, proceedings of *Les Rencontres de Physique de la Vallee d'Aoste*, La Thuile, Aosta Valley, March 2–8, 1997, `(hep-ph/9707349)` .

[106] S. Catani, Yu.L. Dokshitzer and M.H. Seymour, Phys. Lett. **B 285**, 291 (1992).

[107] A. Bassetto, M. Ciafaloni and G. Marchesini, Phys. Rep. **100**, 202 (1983);
Yu.L. Dokshitzer et al., *Basics of perturbative QCD*, Ed. Frontiéres, Gif–Sur–Yvette, 1991.

[108] L. Bellagamba, *MVMATCH: A package to match FMUON tracks with Central Detectors*, ZEUS–Note 96-051; M. Corradi, private communication, and
`http://www-zeus.desy.de/~corradi/ZEUS_ONLY/mpmatch/mpmatch2.html`.

[109] G. Abbiendi, *Global Tracking of muons in the Barrel and Rear region*, ZEUS–Note 99–063.

[110] V. Innocente, M. Maire and E. Nagy, *GEANE: average tracking and error propagation package*, proceedings of MC91: Detector and event simulation in High Energy Physics, Amsterdam, 1991.

[111] R. Fruhwirth, Nucl. Instr. Meth. **A 262**, 444 (1987).

[112] S. Limentani, M. Posocco and L. Stanco, *Muon Barrel Reconstruction*, ZEUS–Note 98-058.

[113] G.F. Hartner et al., *VCTRACK(3.07/04): Offline Output Information*, ZEUS–Note 97–064;
G.F. Hartner, *VCTRAK Briefing: Program and Math*, ZEUS–Note 98–058.

[114] R. Sinkus, T. Voss, Nucl. Instr. Meth. **A 391**, 360 (1997).

[115] J. Repond, *Jet Energy Corrections*, ZEUS–Note 96–133.

[116] G. Briskin, *Diffractive dissociation in ep deep inelastic scattering*, PhD Thesis, DESY–THESIS–1998–036, 1998.

[117] N. Tuning, *ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction*, ZEUS–Note 01–021.

[118] S. Stonjek et al., *D^* and associated dijet photoproduction at HERA*, ZEUS–Note 98–047.

[119] M. Corradi, private communication.

[120] M. Corradi, private communication.

[121] H.L. Lai et al., Eur. Phys. J. **C 12**, 375 (2000).

[122] L. Gladilin, private communication.

[123] M. Corradi, private communication.

[124] R. Barlow, C. Beeston, *Fitting using finite Monte Carlo samples* Comp. Phys. Comm. **77**, 219 (1993).

[125] A. Bertolin, R. Brugnera, private communication. For a list, see:
`http://www-zeus.desy.de/~mturcato/ZEUS_ONLY/BMUON/mbtake.fpp`.

[126] M. Turcato, R. Brugnera, for the list see:
`http://www-zeus.desy.de/~mturcato/ZEUS_ONLY/BMUON/mbtake9800.fpp`.

[127] M. Arneodo, L. Lamberti and M.G. Ryskin, Comp. Phys. Comm. **100**, 195 (1997).

[128] J.A.M. Vermaseren, Nucl. Phys. **B229**, 347 (1983).

List of Figures

1.1 *Electron–proton scattering in Neutral Current (NC)* **(a)** *and Charged Current (CC)* **(b)** *processes.* 7

1.2 *Photon–proton scattering in a e^+e^- collision. If $P^2 = -(k_2 - k_2')^2 \simeq 0$ and $Q^2 = -(k_1 - k_1')^2 > 0$ the photon structure functions can be studied: the high–Q^2 photon is the probe, while the other photon is the target.* . . 9

1.3 *Leading order direct photon processes:* **(a)** *QCD compton and* **(b)** *boson–gluon fusion.* 11

1.4 *Leading order resolved photon processes.* 11

1.5 *A photon–proton interaction. If the virtuality k_T of the propagator shown in the figure is lower than the factorization scale μ_γ the process is included in the resolved calculation at leading order, otherwise it is thought of as a direct process at next–to–leading order.* 14

1.6 *The $\mathcal{O}(\alpha_{\rm em}\alpha_s)$ γg, $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ γg and $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ γq contributions to the total charm photoproduction cross section as a function of the variable $\tau = \hat{w}/W$. The proton renormalization and factorization scales have been taken equal to the charm mass, while the proton parton densities used are MRSA (from Ref. [22]).* 16

1.7 *As in Fig. 1.6, but with HMRS B proton parton densities (from Ref. [22]).* 17

1.8 *The contribution of the hump to the total cross section relative to the $\mathcal{O}(\alpha_{\rm em}\alpha_s^2)$ contribution* **(upper)** *and to the total cross section* **(lower)** *for the MRSA and HMRS B proton parton densities (from Ref. [22]).* . 18

1.9 *Total cross section for $c\bar{c}$ photoproduction as a function of the γp centre–of–mass energy. Theoretical predictions and experimental data from Ref. [30] and [31] are shown.* 20

1.10 *The ratio of the resolved to the direct contribution to beauty production cross section in LO QCD for various fixed values of the photon virtuality $Q^2 \equiv P^2$ (in GeV^2 units).* 22

1.11 *Total cross section for $c\bar{c}$ production in γN collisions, compared to NLO QCD predictions.* 23

1.12 *Differential cross section $d\sigma/dp_T^{D^*}$ for D^* photoproduction for the $(K\pi)\pi_s$ (full dots) and $(K\pi\pi\pi)\pi_s$ (open dots) channels. The NLO QCD predictions in the massive and massless schemes are compared to the data (from Ref. [41]).* 25

1.13 *Differential cross section $d\sigma/d\eta^{D^*}$ for D^* photoproduction, for the $(K\pi)\pi_s$ (full dots) and $(K\pi\pi\pi)\pi_s$ (open dots) channels. The NLO QCD predictions in the massive and massless schemes are compared to the data (from Ref. [41]).* 25

1.14 *Differential cross section $d\sigma/d\eta^{D^*}$, for the two highest–p_T bins, in the $(K\pi)\pi_s$ channel. The curves show the predictions of the massless charm NLO of Ref. [46]* **(a,c)** *and Ref. [48]* **(b,d)**, *with various photon parton density parametrizations (from Ref. [41]).* 26

1.15 *The differential cross section $d\sigma/dx_\gamma^{\rm OBS}$ for dijet plus D^* meson events. The shaded band represents the uncertainty due to the indetermination of the calorimeter energy scale. In* **(a)** *the experimental data are compared to the expectations of the HERWIG Monte Carlo simulations, normalised to the data. In* **(b)** *the data are compared to a parton level NLO massive charm calculation [22, 43] (from Ref. [41]).* 27

1.16 *Differential cross sections $d\sigma/dp_T^D$ and $d\sigma/d\eta^D$ where D stands for D_s (dots) or D^* (triangles). The data are compared to NLO predictions for D_s (full curves) and D^* (dashed curves) with two parameter settings: $m_c = 1.5$ GeV, $\mu_r = m_T$ (thick curves) and $m_c = 1.2$ GeV, $\mu_r = m_T/2$ (thin curves) (from Ref. [53]).* 29

1.17 *The measured $p_T^{\rm rel}$ distribution in the data and the fitted sum (solid line) of the contributions of beauty (dashed line), charm (dotted line) and the fixed fake muon background (dashed–dotted line).* 30

1.18 **Left**: *differential cross section $d\sigma/dp_T^{\rm rel}$ for heavy flavour production. The shaded band represents the effect of the uncertainty on the calorimeter energy scale. The experimental data are compared to the* HERWIG *Monte Carlo prediction (solid line) for all quark flavours, fitted to the data and scaled up by a factor 3.8. The contributions from beauty (forward–diagonally hatched histogram) and from charm processes (backward–diagonally hatched histogram) predicted by the Monte Carlo model are shown separately.* **Right**: *differential cross section $d\sigma/dx_\gamma^{\rm OBS}$ for heavy quark decays. The direct (forward–diagonally hatched histogram) and resolved photon components (backward–diagonally hatched histogram) are shown separately. From Ref. [59].* 32

1.19 *The extrapolated beauty cross section at a fixed $p_T^{\rm min}$ value compared with theoretical predictions plotted as a function of $p_T^{\rm min}$. The curves show the predictions from NLO QCD for varying the b–quark mass and the factorization and renormalization scales (from Ref. [59]).* 33

1.20 **Left:** *Impact parameter distribution in the data (dots), compared to the sum of beauty, charm and fake muons contributions, mixed according to the fractions given by the likelihood fit.* **Right:** *$p_T^{\rm rel}$ distribution in the data (dots), compared to the sum of beauty, charm and fake muons contributions, in the fractions obtained by fitting the beauty and charm components, keeping the fake muons fraction fixed. From Ref. [64].* . . 34

1.21 *Differential cross section $d\sigma/dp_T^{B^+}$ compared to NLO QCD prediction. The solid curve represents the theoretical prediction for $m_b = 4.75$ GeV, $\mu_R = \mu_F = \mu = \sqrt{m_b^2 + p_T^2}$, while the dashed curves are obtained by setting $m_b = 4.5$ GeV, $\mu_R = \mu_F = \mu/2$, and $m_b = 5.0$ GeV, $\mu_R = \mu_F = 2\mu$ (from Ref. [74]).* 37

1.22 *Moments of the measured B meson fragmentation function compared with the perturbative NLL calculation supplemented with different non–perturbative fragmentation forms $D(z)$. The solid line is obtained using a one–parameter form fitted to the second moment.* 38

1.23 *Prediction for the B cross section obtained using the calculation in [79], supplemented with the $N = 2$ fit of the non–perturbative fragmentation function, compared to the CDF data in Ref. [74].* 39

1.24 *Theoretical prediction of the differential cross sections $d\sigma/dp_T^B$, calculated with different schemes and forms of the fragmentation function.* . 40

2.1 *The HERA storage ring with its pre–accelerators system.* 42

2.2 *Integrated luminosity delivered by HERA* **(left)** *and usable for ZEUS physics analyses* **(right)** *in the 1993–2000 running period.* 44

2.3 *The ZEUS coordinate system.* 45

2.4 *Cross section of the ZEUS detector along the beam axis.* 46

2.5 *Cross section of the ZEUS detector orthogonal to the beam axis.* 46

2.6 *$x - y$ cross section of one octant of the CTD. The sense wires are indicated with dots.* 47

2.7 *Hits coming from a genuine track (solid rectangles) tend to cross the cell boundaries within a superlayer. The open rectangles are ghost hits.* . . . 48

2.8 *The track helix in the xy plane.* 49

2.9 *The energy lost by different particles, as reconstructed by the CTD, as a function of the particle momentum. The curves indicate the particle type.* 50

2.10 *Schematic view of the UCAL along the beam axis.* 51

2.11 *A FCAL module.* 52

2.12 *Schematic view of the forward muon detector along the beam axis.* . . . 54

2.13 *An exploded view of the Barrel and Rear Muon Detector, showing the positioning of the chambers.* 55

2.14 *Transverse section of a BMUON chamber, with the honeycomb structure visible* **(left)**. *Scheme of a BMUO chamber* **(right)**. 56

2.15 *Schematic view of the luminosity monitor.* 57

2.16 *Schematic diagram of the ZEUS trigger and data acquisition system.* . . 60

3.1 *Schematic representation of the processes involved in a typical hard QCD interaction.* 64

3.2 *A schematic diagram of ZEUS data and Monte Carlo reconstruction chain.* 68

4.1 *Distribution of inactive material in front of the UCAL in units of the radiation length, X_0, in the $\theta - \phi$ plane, as implemented in the simulation of the detector.* 86

4.2 *Neighbouring calorimeter cells are clustered into cell islands. This schematic picture shows four EMC cell islands and one HAC cell island. EMC cell islands 2 and 3 are joined with HAC cell island 1 to form a cone island. Then the cone islands are matched to tracks.* 88

4.3 *The resolution on the measurement of the electromagnetic energy, E, in the EMC calorimeter and of the transverse momentum, p_T, in the CTD, obtained from single particle Monte Carlo, $(e^{\pm}, \pi^0, \gamma)$ and $(e^{\pm}, \pi^{\pm}, \mu^{\pm}, K^{\pm})$, respectively.* 89

4.4 *Parametrization of the energy lost in the calorimeter by a minimum ionizing particle, shown as a function of the polar angle θ. The dots are the total energy released, while the squares and the triangles represent the electromagnetic and hadronic contributions, respectively. The plot is done using a data sample of dimuon events, collected by the ZEUS detector during 2000.* 90

4.5 $(y_{JB}^{\rm rec} - y^{\rm true})/y^{\rm true}$ **(left)** *and* $y_{JB}^{\rm rec}$ *vs* $y^{\rm true}$ **(right)**, *plotted for various estimations of the inelasticity of the event. In the first row, the resolution* **(left)** *and the comparison between the reconstructed and the true variable* **(right)** *are done calculating* $y_{JB}^{\rm rec}$ *by looping on the calorimeter cells; in the second row, y_{JB} is calculated taking as inputs the standard ZUFOs, without any correction; in the third row, all the corrections described in the text have been applied to the ZUFOs. The sample is obtained by mixing the beauty Monte Carlo samples (direct, resolved and excitations) in fractions given by their inverse luminosities so that they sum up to 1.* 92

4.6 $(p_T^{\rm rec} - p_T^{\rm true})/p_T^{\rm true}$ **(left)** *and* $p_T^{\rm rec}$ *vs* $p_T^{\rm true}$ **(right)**, *plotted for two different reconstructions of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $p_T^{\rm rec}$ *by looping on the calorimeter cells; in the second row, $p_T^{\rm rec}$ is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1.* 95

4.7 $(\eta^{\rm rec} - \eta^{\rm true})/\eta^{\rm true}$ **(left)** *and* $\eta^{\rm rec}$ *vs* $\eta^{\rm true}$ **(right)**, *plotted for two different reconstructions of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $\eta^{\rm rec}$ *by looping on the calorimeter cells; in the second row, $\eta^{\rm rec}$ is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1.* 96

4.8 *The Q^2 distribution of a beauty Monte Carlo sample, obtained mixing the direct, resolved and excitation components in fractions given by their respective inverse luminosities. The empty histogram with full line is the distribution as obtained by applying only the trigger requests; the dashed histogram is obtained by implementing the cuts given by the SINISTRA algorithm, and the diagonally–hatched histogram is the distribution after the cut $0.2 < y_{JB} < 0.8$ has been applied.* 97

4.9 *The efficiency of the dijet TLT as a function of the TLT jet E_T threshold, for offline thresholds of 4,6,8 and* 10 GeV. *The dashed line shows the 1996–97 TLT threshold of* 4 *GeV. In the top plot, jets are reconstructed offline using the EUCELL algorithm; in the bottom plot, jets are reconstructed using KTCLUS. From ref. [118].* 99

4.10 *Distribution of the muon momentum, p^{μ}, for muons coming from beauty (solid line), charm (dashed line) and light quark (dot–dashed line) decays.* 100

4.11 *Comparison between the muon variables as reconstructed by the MPMATCH algorithm and the true muon variables. In the first row, left plot, the ratio $(p^{\mu}_{\rm rec} - p^{\mu}_{\rm true})/p^{\mu}_{\rm true}$ is shown, while the right plot shows the correlation between the reconstructed and true muon momentum, p^{μ}. In the second row, the same is done using the muon pseudorapidity, η^{μ}: the left plot shows the resolution $(\eta^{\mu}_{\rm rec} - \eta^{\mu}_{\rm true})/\eta^{\mu}_{\rm true}$, and the right plot illustrates the comparison between the reconstructed and the true muon pseudorapidity.* 101

4.12 *Performance of the MPMATCH algorithm as a function of the muon momentum, p^{μ}, evaluated on 1999-2000 data sample.* 102

4.13 *Ratio between the MPMATCH performance measured on the data and on the Monte Carlo, plotted as a function of the muon momentum, p^{μ}, for 1996–97* **(left)** *and 1999–2000* **(right)** *data–taking. Inefficiencies found in the data are simulated in the Monte Carlo so that the ratio should be nearly equal to 1, as it is.* 103

4.14 *Comparison between the muon variables as reconstructed by the BREMAT algorithm and the true muon variables. In the first row, left plot, the ratio $(p^{\mu}_{\rm rec} - p^{\mu}_{\rm true})/p^{\mu}_{\rm true}$ is shown, while the right plot shows the correlation between the reconstructed and true muon momentum, p^{μ}. In the second row, the same is done using the muon pseudorapidity, η^{μ}: the left plot shows the resolution $(\eta^{\mu}_{\rm rec} - \eta^{\mu}_{\rm true})/\eta^{\mu}_{\rm true}$, and the right plot illustrates the comparison between the reconstructed and the true muon pseudorapidity.* 104

4.15 *Performance of the barrel* **(left)** *and rear* **(right)** *muon detectors, plotted as a function of the muon transverse momentum, p^{μ}_T, and of the muon momentum, p^{μ}, respectively. The sample used corresponds to the data collected between 1999 and 2000 (see also Appendix A).* 105

4.16 *Ratio between the BREMAT performance as measured on real data and on Monte Carlo events, as a function of p^{μ}_T in the barrel region* **(left)** *and of p^{μ} in the rear region* **(right)**. *The sample used corresponds to the data collected between 1999 and 2000 (see also Appendix A).* 106

4.17 $(y_{JB}^{\rm rec} - y^{\rm true})/y^{\rm true}$ **(left)** *and* $y_{JB}^{\rm rec}$ *vs* $y^{\rm true}$ **(right)**, *plotted for various estimations of the inelasticity of the event. In the first row, the resolution* **(left)** *and the comparison of the reconstructed to the true variable* **(right)** *are done calculating* $y_{JB}^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* y_{JB} *is calculated taking as inputs the standard ZUFOs, without any correction; in the third row, all the corrections described in the text have been applied to the ZUFOs. The sample is obtained by mixing the beauty Monte Carlo samples (direct, resolved and excitations) in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo, in order to obtain a muon–enriched sample.* 109

4.18 $(p_T^{\rm rec} - p_T^{\rm true})/p_T^{\rm true}$ **(left)** *and* $p_T^{\rm rec}$ *vs* $p_T^{\rm true}$ **(right)**, *plotted for two different reconstruction of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $p_T^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* $p_T^{\rm rec}$ *is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo, in order to obtain a muon–enriched sample.* 110

4.19 $(\eta^{\rm rec} - \eta^{\rm true})/\eta^{\rm true}$ **(left)** *and* $\eta^{\rm rec}$ *vs* $\eta^{\rm true}$ **(right)**, *plotted for two different reconstruction of the jet variables of the event. In the first row, the resolution* **(right)** *and the comparison of the reconstructed to the true variable* **(left)** *are done calculating* $\eta^{\rm rec}$ *by looping on the calorimeter cells; in the second row,* $\eta^{\rm rec}$ *is calculated taking as inputs the ZUFOs, with all the corrections applied. The sample is obtained by mixing the beauty Monte Carlo samples in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo, in order to obtain a muon–enriched sample.* 111

5.1 *Muon (***left***) and jet (***right***) multiplicity in 1996–2000 data sample, for events passing all the cuts, before the muon–jet association. The cuts on the third and fourth jets are the same imposed on the second (see text). No more than four jets are selected for the matching with the muon.* . . 114

5.2 *Graphical representation of the meaning of the* $p_T^{\rm rel}$ *variable.* 117

5.3 $p_T^{\rm rel}$ *distributions for data* **(a)**, *beauty* **(b)**, *charm* **(c)** *and light flavour* **(d)** *components.* 118

5.4 *Results of the fit of the direct and resolved Monte Carlo components to the* $x_\gamma^{\rm MEAS}$ *distribution of the data, for 1996–1997 (***left plot***) and 1999–2000 (***right plot***) data samples. The dots represent the data, the fitted Monte Carlo distribution is drawn as the empty histogram and the contributions coming from direct and resolved photon processes are shown as the forward– and backward–diagonally hatched histograms, respectively.* 122

5.5 *The* $p_T^{\rm rel}$ *distributions for* **(a)** *1996–97 and* **(b)** *1999–2000 data. The data are compared to the* PYTHIA *Monte Carlo with the mixture of beauty and charm+light flavour given by the fit. The beauty contribution is represented by the hatched histogram.* 123

5.6 *The* $x_\gamma^{\rm MEAS}$ *distributions plotted using 1996–97* **(left plot)** *and 1999–2000 data* **(right plot)** *compared to the sum of beauty and charm plus light flavour Monte Carlo, mixed according to the b and background fractions given by the new* $p_T^{\rm rel}$ *fit, performed after the cross sections of direct and resolved b and c + l sub–processes have been reweighted as explained in the text. The fitted distribution is area–normalized to the data. The hatched histogram represents the beauty contribution.* 124

5.7 *The* $x_\gamma^{\rm MEAS}$ *distribution for data having good muons in the barrel/rear* **(a)** *and forward region* **(b)**. *The data sample refers to the 1999–2000 period.* 124

5.8 $(x_\gamma^{\rm MEAS} - x_\gamma^{\rm TRUE})/x_\gamma^{\rm TRUE}$ *and* $x_\gamma^{\rm MEAS}$ *vs* $x_\gamma^{\rm TRUE}$, *where* $x_\gamma^{\rm TRUE}$ *is evaluated using massive jets before the beauty decay, as illustrated in Sect. 4.7.4. The sample is obtained by mixing the beauty Monte Carlo samples (direct, resolved and excitations) in fractions given by their inverse luminosities so that they sum up to 1. All the cuts of the analysis have been applied to the Monte Carlo. The observed resolution on the* $x_\gamma^{\rm MEAS}$ *variable is* $\sim 20\%$. 125

5.9 *Control plots for the muon variables, relative to the* **1996–97** *data sample. In the upper plots, the muon momentum,* p^μ, **(left)** *and transverse momentum,* p_T^μ, **(right)** *are shown, in the middle plot the muon pseudorapidity,* η^μ, *is displayed. The lower plots show the probability of the 5 d.o.f. matching,* Prob. match p, *of the two muon matching algorithms, MPMATCH and BREMAT . In all the plots the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the* $p_T^{\rm rel}$ *fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.* 127

5.10 *Control plots for the muon variables, relative to the* **1999–2000** *data sample. In the upper plots, the muon momentum,* p^μ, **(left)** *and transverse momentum,* p_T^μ, **(right)** *are shown, in the middle plot the muon pseudorapidity,* η^μ, *is displayed. The lower plots show the probability of the 5 d.o.f. matching,* Prob. Match p, *of the two muon matching algorithms, MPMATCH and BREMAT . In all the plots the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the* $p_T^{\rm rel}$ *fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.* 128

5.11 *Control plots for the jets variables, relative to the* **1996–97** *data sample. In the upper plots, the transverse momentum,* p_T^{jet1}, **(left)** *and pseudorapidity,* η^{jet1}, **(right)** *of the highest transverse momentum jet in the event are shown, and the same variables, referred to the second–highest transverse momentum jet, are shown in the middle plots. In the lower plots the transverse momentum,* $p_T^{\text{sel jet}}$, **(left)** *and pseudorapidity,* $\eta^{\text{sel jet}}$, **(right)** *of the jet associated to the muon are displayed. In all the plots, the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the* p_T^{rel} *fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.* 129

5.12 *Control plots for the jets variables, relative to the* **1999–2000** *data sample. In the upper plots, the transverse momentum,* p_T^{jet1}, **(left)** *and pseudorapidity,* η^{jet1}, **(right)** *of the highest transverse momentum jet in the event are shown, and the same variables, referred to the second–highest transverse momentum jet, are shown in the middle plots. In the lower plots the transverse momentum,* $p_T^{\text{sel jet}}$, **(left)** *and pseudorapidity,* $\eta^{\text{sel jet}}$, **(right)** *of the jet associated to the muon are displayed. In all the plots, the data (dots) are compared to the fitted distribution (empty histogram) obtained summing the beauty and charm plus light flavour distributions in the fractions given by the* p_T^{rel} *fit. The fitted distribution is area normalized to the data. The hatched histogram represents the contribution coming from beauty events.* 130

6.1 *Parton to hadron corrections, evaluated using massive jets before the beauty decay (full line), massless jets before the beauty decay (dot–dashed line) massive (dotted line) and massless (dashed line) jets after the beauty decay. The Monte Carlo used is* PYTHIA *and the different beauty sub–processes have been mixed according to their inverse integrated luminosities. The correction factors are shown in the three angular regions of the visible cross sections* **(a)** *and in all the bins of the differential distributions* $d\sigma/dx_\gamma^{\text{MEAS}}$ **(b)**, $d\sigma/d\eta^\mu$ **(c)** *and* $d\sigma/dp_T^\mu$ **(d)**. 133

6.2 *Differential cross section* $d\sigma/dx_\gamma^{\text{MEAS}}$ *, evaluated using the two data samples available, relative to the 1996–97 (squares) and 1999–2000 (circles) data–taking periods. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The two histograms represent the* PYTHIA *predictions: the solid histogram is the Monte Carlo cross section at* $\sqrt{s} = 318$ GeV*, the dashed one is the prediction at* $\sqrt{s} = 300$ GeV. 140

6.3 *Differential beauty cross section as a function of* x_γ^{MEAS} *for two-jets events, compared to the predictions of NLO QCD (dashed line: parton–level jets, full line: corrected to hadron–level jets). The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The results of the two data samples have been merged as explained in the text. The band around the NLO prediction shows the range obtained by varying the b–quark mass as well as* μ_r *and* μ_f*, as explained in the text.* 141

6.4 *Differential cross section* $d\sigma/d\eta^\mu$*, evaluated using the two data samples available, relative to the 1996–97 (squares) and 1999–2000 (circles) data–taking periods. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The two histograms represent the* PYTHIA *predictions: the solid histogram is the Monte Carlo cross section at* $\sqrt{s} = 318$ GeV*, the dashed one is the prediction at* $\sqrt{s} = 300$ GeV. 143

6.5 *Differential beauty cross section as a function of the muon pseudorapidity,* η^μ*, for events with two jets and a muon, compared to the predictions of NLO QCD (dashed line: parton–level jets, full line: corrected to hadron–level jets). The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The results of the two data samples have been merged as explained in the text. The band around the NLO prediction shows the result obtained by varying the b–quark mass as well as* μ_r *and* μ_f*, as explained in the text.* 144

6.6 *Differential cross section* $d\sigma/dp_T^\mu$*, evaluated using the two data samples available, relative to the 1996–97 and 1999–2000 data–taking periods. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The two histograms represent the* PYTHIA *predictions: the solid histogram is the Monte Carlo cross section at* $\sqrt{s} = 318$ GeV*, the dashed one is the prediction at* $\sqrt{s} = 300$ GeV. 145

6.7 *Differential beauty cross section as a function of the muon transverse momentum,* p_T^μ*, for events with two jets and a muon, compared to the predictions of NLO QCD (dashed line: parton–level jets, full line: corrected to hadron–level jets). The results of the two data samples have been merged as explained in the text. The full error bars are the quadratic sum of statistical (inner part) and systematic uncertainties. The band around the NLO prediction shows the range obtained by varying the b–quark mass as well as* μ_r *and* μ_f*, as explained in the text.* 146

6.8 *Systematic uncertainties on the measured values of the cross sections, shown for all the measured cross sections and for some of the different considered contributions. In each histogram the systematic uncertainty is shown for the cross sections in three bins (RBF, Rear, Barrel, Forward), for the differential cross section* $d\sigma/dx_\gamma^{\text{MEAS}}$ *, for the beauty dijet cross section* (jj)*, and for the two differential muon cross sections,* $d\sigma/d\eta^\mu$ *and* $d\sigma/dp_T^\mu$*. The considered contributions are described in detail in the text. The evaluation is referred to the* **1996–97** *cross sections.* 151

6.9 *Systematic uncertainties on the measured values of the cross sections, shown for all the measured cross sections and for the different considered contributions. In each histogram the systematic uncertainty is shown for the cross sections in three bins (RBF, Rear, Barrel, Forward), for some of the differential cross section* $d\sigma/dx_\gamma^{\rm MEAS}$ *, for the beauty dijet cross section (jj), and for the two differential muon cross sections,* $d\sigma/d\eta^\mu$ *and* $d\sigma/dp_T^\mu$*. The considered contributions are described in detail in the text. The evaluation is referred to the* **1999–2000** *cross sections.* . . . 152

7.1 *Ratio of measured b–production cross sections at HERA and the theoretical expectation from NLO QCD, as a function of* Q^2*. The full error bars are the quadratic sum of the statistical (inner part) and systematic uncertainties. The full square in the low–*Q^2 *region corresponds to the value obtained from the beauty dijet cross section reported in this thesis.* 155

A.1 *Rates of the GLOMU trigger bit, for 1998* **(upper left)**, *1999 electron* **(upper right)**, *1999 positron* **(lower left)** *and 2000* **(lower right)** *data. In each histogram every bin corresponds to a luminosity interval of* ~ 700 nb^{-1}*. The rate is obtained dividing the number of the GLOMU events in each bin by the corresponding luminosity. Bad runs have been rejected.* . . . 160

A.2 **Left***: Schematic diagram for the elastic* J/ψ *production process,* $ep \to e\ J/\psi\ p$*.* **Right:** *Two–photon process* $ep \to \gamma\gamma \to l^+l^-$. . . . 161

A.3 *Invariant mass of the* $\mu^+\mu^-$ *system, for 2000 data, compared to the Monte Carlo mixture described in the text, area–normalised to the data. The elastic* J/ψ *and* $\gamma\gamma$ *contributions have been weighted in order to reproduce the shape of the distribution.* . . . 162

A.4 **(a):** *Polar angle distribution of the track under analysis, for tracks in the barrel region (*$35^o < \theta^\mu < 135^o$*) having transverse momentum* $p_T >$ 1.5 GeV*. The dots represent the data corresponding to the 2000 data taking, while the shaded band is the Monte Carlo with its error.* **(b):** *Polar angle distribution for the track under analysis, for tracks in the rear region (*$135^o < \theta^\mu < 157^o$*) and without cuts on the momentum. The shaded band is the Monte Carlo with its errors.* . . . 163

A.5 **(a):** *Azimuthal angle distribution of the track under analysis, for tracks in the barrel region (*$35^o < \theta^\mu < 135^o$*) having transverse momentum* $p_T > 1.5$ GeV*. The dots represent the data corresponding to the 2000 data taking, while the shaded band is the Monte Carlo with its error.* **(b):** *Azimuthal angle distribution for the track under analysis, for tracks in the rear region (*$135^o < \theta^\mu < 157^o$*) and without cuts on the momentum. The shaded band is the Monte Carlo with its errors.* . . . 163

A.6 *BREMAT performance in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1996–97** *data sample.* . . 167

A.7 *BREMAT performance in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1999–2000** *data sample.* . 168

A.8 *Ratio between the performances evaluated on the data and on the Monte Carlo, in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1996–97** *data sample.* . . . 169

A.9 *Ratio between BREMAT performances evaluated on the data and on the Monte Carlo, in the barrel* **(up)** *and rear* **(down)** *region of the BRMUON detector, calculated using the* **1999–2000** *data sample.* . . . 170

A.10 *Efficiency of the wire (full markers) and strip planes (empty markers) for the inner (circles) and outer (squares) muon chambers, evaluated in the different years of data–taking.* . . . 172

B.1 *Pulls of the variables used by the MPMATCH algorithm to fit the muon track: x and y coordinates, slopes* x' *and* y' *and inverse of the momentum* Q/p*. The pulls have been evaluated on 1996–97 data, after all the cuts of the analysis have been applied.* . . . 174

B.2 *Pulls of the variables used by the BREMAT algorithm to fit the muon track: x and y coordinates, slopes* x' *and* y' *and inverse of the momentum* Q/p*. The pulls have been evaluated on 1996–97 data, after all the cuts of the analysis have been applied.* . . . 175

B.3 *Pulls of the variables used by the MPMATCH algorithm to fit the muon track: x and y coordinates, slopes* x' *and* y' *and inverse of the momentum* Q/p*. The pulls have been evaluated on 1999–2000 data, after all the cuts of the analysis have been applied.* . . . 176

B.4 *Pulls of the variables used by the BREMAT algorithm to fit the muon track: x and y coordinates, slopes* x' *and* y' *and inverse of the momentum* Q/p*. The pulls have been evaluated on 1999–2000 data, after all the cuts of the analysis have been applied.* . . . 177

List of Tables

1.1 *Pointlike component of the total charm cross section (μb), in γp collisions, for different choices of the input parameters.* 19
1.2 *Hadronic component of the total charm cross section (μb), in γp collisions, for different choices of the input parameters.* 19
1.3 *Pointlike component of the total beauty cross section (nb), in γp collisions, for different choices of the input parameters.* 21
1.4 *Hadronic component of the total beauty cross section (nb), in γp collisions, for different choices of the input parameters.* 21

2.1 *Main design parameters of HERA [83].* 43
2.2 *Overview of the luminosity delivered by HERA from 1992 to 2000.* . . . 43
2.3 *Angular coverage of the UCAL parts and dimensions of the cells.* 51

3.1 *Cross section and integrated luminosity generated for the various Monte Carlo samples, beauty, b, charm, c, and light flavours, l. The statistics are reported separately for MC samples simulating 1996–97 and 1999–2000 events.* 72

4.1 *Correction factors to be applied to data and Monte Carlo to keep into account the difference in the calorimeter energy scale. Correction factors are 1 for the Monte Carlo, while for data they depend on the detector sector and on the type of the cell.* 85
4.2 *Summary of the correction factors given by the ratio of the BREMAT performance evaluated on the data and on the Monte Carlo. The factors are evaluated in the momentum region of interest, i.e. for $p_T^\mu > 2.5$ GeV in the barrel region and for $p^\mu > 2.5$ GeV in the rear region.* 107

5.1 *Number of events passing all the cuts, including the muon–jet association, shown separately in the barrel–rear and forward region, for the two data–taking periods.* 115
5.2 *Beauty percentages in the data, for $x_\gamma^{\mathrm{MEAS}} < 0.75$ (resolved photon processes) and $x_\gamma^{\mathrm{MEAS}} > 0.75$ (direct γ processes), extracted by fitting the p_T^{rel} distributions in the two regions.* 120
5.3 *Weights to be applied to the cross sections of the different Monte Carlo contributions in order to correctly take into account the fractions of direct and resolved events in the data, and the fractions of beauty and background events in the regions of direct and resolved photon processes. All the weights are referred to the one applied to the background direct processes, that is therefore equal to 1.* 122

5.4 *Summary of the resolutions on the variables of interest for the analysis.* 125

6.1 *Beauty percentages in the three angular regions defined for the visible cross section calculation, obtained by fitting to the data the $p_T^{\rm rel}$ distributions of the beauty and charm plus light flavour components given by the Monte Carlo. The results have been evaluated separately in the two data–taking periods, 1996–97 and 1999–2000.* 134

6.2 *Comparison between the number of events selected in the forward, barrel and rear regions with two different sets of cuts: the first set, defined in Eqs. 6.5, 6.6 and 6.7, has been used for the evaluation of the visible cross sections in the three angular regions, for the dijet cross section and for $d\sigma/dx_\gamma^{\rm MEAS}$; the second set, defined in 6.30, has been used for the determination of the muon differential cross sections, $d\sigma/d\eta^\mu$ and $d\sigma/dp_T^\mu$.* 135

6.3 *Summary of the correction factors to be applied to the data in the three angular regions, due to the muon detectors performances. The correction factors are given by the ratio of the chambers performances evaluated on the data and on the Monte Carlo. The values relative to the 1996–97 and 1999–2000 samples have been evaluated separately.* 135

6.4 *Measured values of the acceptance in the three angular regions in the two data–taking periods.* 136

6.5 *Measured values of the acceptance for the dijet beauty cross section determination in the two data–taking periods.* 138

6.6 *Ratio between the cross sections measured on the data and the predictions of NLO QCD, reported for all the measured cross sections and for every bin of the differential cross sections.* 147

A.1 *Performance of the reconstruction algorithm for muons in the barrel and rear region, BREMAT [109]. The two different data sets used in the analysis, 96–97 and 99–00, have been analysed separately in order to investigate effects due to detector aging and different running conditions. The weighted mean for each data taking period is reported.* 166

A.2 *Ratio between the BREMAT performance evaluated on the data and on the Monte Carlo, for the two different data–taking periods. The values are reported in bins of the muon transverse momentum, p_T, in the barrel region, and of the momentum, p, in the rear region. In the last row, the weighted mean of the last three bins is reported, to be used as a correction factor for the evaluated cross sections.* 166

A.3 *Comparison between the efficiencies of the strip and wire planes in 1997 and 2000. In the last row the relative difference with respect to the 1996–97 value is shown,* (effic 9697 − effic 9900)/effic 9697, *expressed in percent.* 171

C.1 *Beauty visible cross sections in the three angular regions of the ZEUS detector. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.* . 179

C.2 *Cross section of the process $ep \to b\bar{b}\ X \to$ dijet X. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.* 180

C.3 *Differential cross section $d\sigma/dx_\gamma^{\rm MEAS}$ for two–jet events. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.* 180

C.4 *Differential cross section $d\sigma/d\eta^\mu$ for two–jet plus muon events. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.* 181

C.5 *Differential cross section $d\sigma/dp_T^\mu$ for two–jet plus muon events. The values found in the two data–taking periods are shown, together with the merged values and the NLO–predicted cross sections. For the measured value, the first error is statistical, while the second is systematic; for the theoretical predictions, the indetermination derives from the change in the input parameters, the mass of the b quark and the renormalization and factorization scales, as explained in Sect. 6.1.* 181

Acknowledgments

Una ovvia e già sentita premessa: mi scuso con le persone che dovessi eventualmente dimenticare, ma due anni e mezzo sono lunghi e ci sono state tante persone che mi sono state vicino, per poco o tanto tempo... Quindi, chiunque dimentichi, mi avvisi e avrà le mie scuse pubbliche!

Partiamo dal lavoro, visto che è quello l'argomento riportato nella Tesi. Vorrei in primo luogo ringraziare tutto il gruppo di ZEUS–Padova, per avermi aiutata anche dal punto di vista umano, da quando ero laureanda, ai periodi bui del pre–dottorato, quando non sapevo dove sarei finita, al dottorato stesso. Ringrazio prima di tutto il mio tutore, Riccardo Brugnera, per essere sempre stato disponibile quando avevo bisogno di aiuto o di un semplice consiglio, per aver letto con attenzione la mia tesi, per avermi insegnato una buona parte delle cose che so. Poi, ringrazio Luca Stanco, per il suo ruolo di infaticabile organizzatore, anche in situazioni estreme (!) e, anche lui, per essere stato sempre disponibile nonostante i numerosi impegni. Ringrazio Roberto Carlin per essere stato anche lui sempre disponibile per aiuti o chiarimenti, per avermi sempre supportata e fatto coraggio nei momenti difficili... Ringrazio la professoressa Silvia Limentani per essersi interessata al mio lavoro e per avermi fornito interessanti opinioni. E poi, Enrico Borsato per aver più volte rallegrato l'atmosfera dell'ufficio, anche se spesso parlando male delle donne! Da non dimenticare anche Flavio Dal Corso, che mi ha più volte tolta da guai BMUON–eschi che non avrei saputo altrimenti come risolvere! Ringrazio tutti anche per le loro qualità umane, che mi hanno fatto trovare nel gruppo, oltre che un valido supporto per la mia attività di ricerca, anche delle persone con cui parlare serenamente anche nei momenti difficili, e, perché no, con cui condividere quelli felici. Poi, nuova–vecchia entrata nel gruppo di Padova, ringrazio anche Alessandro Bertolin, per i consigli di lavoro, per il supporto morale e non solo, per i pranzi e le cene condivisi, per la compagnia che mi ha fatto ad Amburgo, per la fiducia che mi ha sempre dato... Insomma, per un sacco di cose!

Poi, a Bologna... Un doveroso e sentito ringraziamento va a Massimo Corradi, che ha fatto la mia stessa analisi in questi anni, e con il quale speriamo di arrivare presto a un articolo vero. Lo ringrazio, tra l'altro, per avermi spinto ad arrivare al risultato preliminare per ICHEP2002, grazie a questo risultato poi la strada per la Tesi è stata più in discesa! Non posso poi dimenticarmi di Graziano Bruni, per essere stato parte fondamentale nel preliminare di EPS2001, e per averci anche aiutato con il più recente risultato. Non ultimo, per aver generato tutti quegli eventi di light flavours!

Un ringraziamento va poi a tutto il gruppo degli Heavy Flavours, a DESY, e in generale a tutta la Collaborazione ZEUS nell'ambito della quale ho condotto il mio lavoro.

Chiudo col lavoro e passo alla famiglia. Voglio ringraziare infinitamente mia madre,

alla quale va parte del merito del fatto che sono arrivata alla laurea, bene e in tempi brevi, e se non mi fossi laureata non avrei nemmeno mai scritto questa Tesi. Grazie mamma. Poi, mio padre, per essersi sempre informato sul mio lavoro, per essermi stato vicino e per avermi capita, nel lavoro come nella vita. La mia sorellina, anche se l'ho vista poco in questi anni a causa della mia presenza molto effimera a casa, spero non si sia dimenticata che le voglio bene e che desidero moltissimo che sia felice! Ha delle buone possibilità, spero che le sfrutti appieno! Vai Tania!

Dopo i parenti strettissimi, quelli un po' meno stretti... Voglio ricordare in primo luogo la mia nonna che se ne è andata proprio nel periodo del mio Dottorato, mi spiace che non possa essere felice con me a festeggiarne la conclusione, assieme al marito e alle figlie, tutte le mie zie, che ricordo caramente, assieme alle loro famiglie. Infine, ricordo la mia nonna paterna, spero che in questa occasione sia felice con me!

Il capitolo degli amici è complicato, viste le varie persone che mi sono state vicine in questi anni, che lo sono ancora o che non lo sono più per motivi logistici o di tempo... Partiamo da Padova, e dagli amici del Copernico, la casa dello studente che ora non frequento più ma che per quattro anni è stata il mio punto di riferimento a Padova. Prima di tutto, Daniele, che adesso vedo così poco ma che è sempre nel mio cuore, ci vuole altro per farmi dimenticare quanto è stato importante per me, quanto mi è stato vicino e mi ha aiutata ogni volta che poteva. Poi, Anna e Ilaria, le mie care amiche Copernicane, su cui so che posso sempre contare, e loro spero sappiano che possono contare su di me. Infine, in particolare per una sera in cui ha raccolto i miei cocci sparsi un po' ovunque, voglio ringraziare Isabella. Mi ha aiutato tanto, anche se ci siamo viste così poco. Infine, un'amica non del Copernico ma sempre di casa, Silvia: di lei che dire? La conosco da tanto tempo e ci sono poche cose che lei non sa di me... Grazie anche a lei.

Poi, i fisici, dell'ufficio e non: Giorgio, che finalmente si è comprato il portatile!: lo voglio ringraziare soprattutto per avermi portato le brioches ogni martedì mattina, puntualissimo! A parte tutto, insieme abbiamo fatto tante belle pause caffè, e lui sa tante cose di me... Lo ringrazio per essermi sempre stato ad ascoltare quando ne avevo bisogno, anche se era difficile! Quindi, gli amici dell'ufficio: Andrea Longhin, con cui abbiamo discusso tante volte di bellezza e di camere (!), Moreno e Fulvio, per avere installato Linux sul mio portatile nel tempo record di mezz'ora, Marcello con cui tante volte abbiamo fatto le ore piccole in ufficio, in periodi impegnativi per entrambi, Enrico Feltresi, salutoni a Dresda!, Alvise, Gionni, il Professor Camerata, alias Federico, per essere sempre stato in minoranza e non essersi eccessivamente lamentato, Melisa, per tante cose, tra cui il sempre proficuo shopping, Enrico Carrara e tutti gli altri che gironzolano lì attorno e con cui si beve il caffè. Quindi, gli amici di qualche vacanza e/o delle pizze di laboratorio, Baz, Flora, Sara, Fabiana e gli altri.

Qualche altro: Emanuele, Danijel, anche lui che mi ha aiutato molto in un momento difficile, Chiara, per rapprensentare la mia coscienza pacifista, Ruggero, Amita, Vincenzo, Sabine, Margherita, Nicola, Alessandro, Barbara, Michela, e tutti quelli che ho dimenticato.

Infine, per essermi stato vicino in questo ultimo anno e più, per avermi aiutato tante volte, per aver avuto pazienza nei momenti difficili, un grazie enorme va ad Andrea, per aver realizzato qualcosa che non credevo più mi potesse accadere. Grazie.