

# Measurement of Charm and Beauty Dijet Cross Sections in Photoproduction

by

L. Finke



# Measurement of Charm and Beauty Dijet Cross Sections in Photoproduction

## Dissertation

zur Erlangung des Doktorgrades  
des Fachbereichs Physik  
der Universität Hamburg

vorgelegt von  
LARS FINKE  
aus Aachen

Hamburg  
2006

Gutachter der Dissertation:

Prof. Dr. B. Naroska  
Prof. Dr. R.-D. Heuer  
Prof. Dr. A. Caldwell

Gutachter der Disputation:

Prof. Dr. B. Naroska  
Dr. A. Geiser

Datum der Disputation:

15.09.2006

Vorsitzende/Vorsitzender des Prüfungsausschusses:

Dr. H.-D. Rüter

Vorsitzender des Promotionsausschusses  
und Dekan des Fachbereichs Physik:

Prof. Dr. G. Huber

## Abstract

A measurement of charm and beauty dijet photoproduction cross sections at the  $ep$  collider HERA is presented. The positron data of the years 1999 and 2000 are analysed, corresponding to an integrated luminosity of  $56.8 \text{ pb}^{-1}$ . Events are selected with two or more jets of transverse momentum  $p_t^{jet_{1(2)}} > 11(8) \text{ GeV}$  in the central range of pseudo-rapidity  $-0.9 < \eta^{jet_{1(2)}} < 1.3$ . The fractions of events containing charm and beauty quarks are determined using a method based on the impact parameter, in the transverse plane, of tracks to the primary vertex, as measured by the H1 central vertex detector. Differential dijet cross sections for charm and beauty, and their relative contributions to the flavour inclusive dijet photoproduction cross section, are measured as a function of the transverse momentum of the leading jet, the average pseudo-rapidity of the two jets, the azimuthal distance of both jets, the dijet centre-of-mass scattering angle and the observable  $x_\gamma^{obs}$ . Taking into account the theoretical uncertainties, the charm cross sections are consistent with a QCD calculation in next-to-leading order, while the predicted cross sections for beauty production are somewhat lower than the measurement.

## Kurzfassung

Ziel dieser Arbeit ist die Messung von Charm und Beauty Zweijet-Wirkungsquerschnitten in  $ep$  Kollisionen bei HERA im kinematischen Bereich der Photoproduktion. Die Positron-Daten der Jahre 1999 und 2000 mit einer integrierten Luminosität von  $56.8 \text{ pb}^{-1}$  werden analysiert. Selektiert werden Ereignisse mit mindestens zwei Jets im zentralen Rapiditätsbereich  $-0.9 < \eta^{jet_{1(2)}} < 1.3$  und Transversalimpulsen  $p_t^{jet_{1(2)}} > 11(8) \text{ GeV}$ . Zur Bestimmung der Anteile von Charm und Beauty Ereignissen am gesamten Datensatz wird eine Methode benutzt, die auf der Messung des transversalen Stossparameters von Teilchenspuren in Bezug auf den primären Vertex basiert. Diese werden von dem zentralen H1 Vertex Detektor aufgezeichnet. Differenzielle Zweijet-Wirkungsquerschnitte für Charm und Beauty und ihr relativer Anteil am flavourinklusive Photoproduktions-Wirkungsquerschnitt werden als Funktion des transversalen Impulses des führenden Jets, der gemittelten Pseudo-rapidity beider Jets, des azimuthalen Abstands beider Jets, des Zweijet-Streuwinkels sowie der Observablen  $x_\gamma^{obs}$  gemessen. Unter Berücksichtigung der theoretischen Unsicherheiten sind die gemessenen Charm Wirkungsquerschnitte konsistent mit einer QCD Rechnung in nächst führender Ordnung, während die vorhergesagten Beauty Wirkungsquerschnitte etwas niedriger als die gemessenen sind.



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Photoproduction of Heavy Quarks</b>	<b>3</b>
2.1	Quantum Chromodynamics . . . . .	3
2.2	Basics of Electron Proton Scattering . . . . .	4
2.2.1	Factorisation and Parton Evolution . . . . .	6
2.2.2	Proton Structure . . . . .	9
2.2.3	Photoproduction and the Photon Flux . . . . .	10
2.2.4	Photon Structure . . . . .	11
2.3	Heavy Quark Production . . . . .	14
2.4	Heavy Quark Fragmentation . . . . .	17
2.5	QCD Calculations . . . . .	18
2.5.1	NLO QCD Calculations . . . . .	18
2.5.2	LO Monte Carlo Event Generators . . . . .	21
2.6	Properties of Heavy Hadrons . . . . .	23
<b>3</b>	<b>The Experiment</b>	<b>26</b>
3.1	The H1 Detector at HERA . . . . .	26
3.1.1	Central Tracking Detectors . . . . .	27
3.1.2	Calorimetry . . . . .	29
3.1.3	H1 Trigger System . . . . .	30
3.1.4	Detector Simulation . . . . .	31
3.2	Experimental Review . . . . .	31
3.2.1	Heavy Quark Production at HERA . . . . .	32
3.2.2	Heavy Quark Production at other colliders . . . . .	35
<b>4</b>	<b>Strategy and Event Reconstruction</b>	<b>39</b>
4.1	Analysis Strategy . . . . .	39
4.2	Track Reconstruction . . . . .	42
4.3	CST improved tracks . . . . .	44
4.4	Primary Vertex Reconstruction . . . . .	44
4.5	Impact Parameter Resolution . . . . .	46
4.6	HFS Reconstruction . . . . .	51
4.7	Jet Algorithm . . . . .	52
4.8	Observables . . . . .	54

<b>5</b>	<b>Dijet Event Selection</b>	<b>57</b>
5.1	Preselection . . . . .	57
5.2	Photoproduction Event Selection . . . . .	57
5.3	Hadronic Final State Calibration . . . . .	58
5.4	Trigger Selection and Efficiency . . . . .	61
5.5	Background Studies . . . . .	64
5.6	Control Measurements . . . . .	65
5.7	Jet Structure . . . . .	69
5.8	Data Unfolding . . . . .	73
5.8.1	Resolution and Systematic Shifts . . . . .	73
5.8.2	Purity and Stability . . . . .	75
5.8.3	Definition of the Cross Section . . . . .	75
5.8.4	Hadronisation Corrections . . . . .	78
<b>6</b>	<b>Quark Flavour Separation</b>	<b>80</b>
6.1	Track Selection . . . . .	80
6.2	Lifetime Tag Observables . . . . .	83
6.3	Fit Procedure . . . . .	85
6.4	Multi-Impact Parameter Method . . . . .	86
6.5	Secondary Vertex Reconstruction . . . . .	88
6.6	Heavy Flavour Enriched Data Sample . . . . .	91
<b>7</b>	<b>Cross Section Measurement</b>	<b>97</b>
7.1	Systematic Uncertainties . . . . .	97
7.2	Charm and Beauty Dijet Cross Section . . . . .	103
7.2.1	Differential Cross Sections . . . . .	103
7.2.2	Double Differential Cross Sections . . . . .	108
7.3	Flavour Inclusive Dijet Cross Section . . . . .	108
7.4	Charm and Beauty Fractions . . . . .	113
7.5	Comparison with Previous Measurements at HERA . . . . .	116
<b>8</b>	<b>Conclusions and Outlook</b>	<b>118</b>
<b>A</b>	<b>Data Tables</b>	<b>121</b>
<b>B</b>	<b>Charm Cross Section with <math>D^*</math> Mesons</b>	<b>126</b>
<b>C</b>	<b>Polar angular distribution <math>\cos\theta^*</math></b>	<b>129</b>
	<b>References</b>	<b>132</b>

# CHAPTER 1

## INTRODUCTION

---

The theory of Quantum Chromodynamics (QCD) is an important part of the Standard Model of particle physics. QCD is the theory of the strong interaction, a fundamental force describing the interactions among the constituents of hadrons (quarks) via the exchange of massless gauge bosons (gluons). Both quarks and gluons carry colour and interact with each other, including gluon self-interactions. This leads to a running strong coupling constant  $\alpha_s$ . At large distances (low energies)  $\alpha_s$  is large and the principle of colour confinement prevents that quarks and gluons can be observed directly. Instead, through the mechanism of hadronisation, a collection or 'jet' of colour-neutral hadrons emerges from high energy collisions which can be investigated experimentally.

Perturbative calculations are performed by expanding scattering amplitudes in powers of  $\alpha_s$ . The use of perturbative theory is based on the fact that in the presence of a hard scale (high energy) the coupling constant is small enough to ensure the convergence of the perturbation series.

At the electron-proton collider HERA at DESY inclusive jet cross sections have been measured in a wide kinematic range and good agreement has been found between the measurements and theoretical calculations performed in the framework of perturbative Quantum Chromodynamics (pQCD) to next-to-leading order [1, 2, 3, 4]. These findings have provided increasing confidence that the production mechanisms for dijet events are well understood.

The study of heavy quark initiated jets provides a particularly interesting testing ground of pQCD, since the large heavy quark masses already provide sufficiently hard scales for the calculations. However, there are other hard scales relevant in heavy quark production processes, such as the virtuality,  $Q^2$ , of the exchanged photon and the transverse momentum of the outgoing quark. This multi-scale-problem complicates the situation for pQCD calculations and there exist many competing calculation schemes to deal with this problem.

Over the last decades charm and beauty production have been investigated intensively first in  $\gamma N$ -collisions at fixed target experiments and later in high energy particle collisions at LEP ( $p\bar{p}$ ,  $\gamma\gamma$ ), at the TEVATRON ( $p\bar{p}$ ) and at HERA ( $ep$ ,  $\gamma p$ ). While, in general, agreement of the QCD predictions with the charm data has been found, in the beauty sector some deviations have been observed with the beauty data being somewhat larger than the predictions.

The precise understanding of heavy quark production is an important aspect in current

particle physics, in particular in view of the upcoming LHC experiments, where complicated final states like charm and beauty photoproduction accompanied by jets will be the main background to the eagerly expected signals from new physics.

The goal of this thesis is to contribute to the present picture of heavy quark production at HERA. Previous measurements of charm and beauty at HERA were mostly based on the explicit reconstruction of e.g. a  $D^*$  meson in the case of charm, while for beauty usually the semi-leptonic decay into electrons or muons was used as a signature. In these measurements the branching fractions and lepton identification requirements restrict the statistical accuracy of the data. Recently, at H1 a new approach has been developed which is based on purely hadronic informations (jets) using lifetime information. In this approach events containing heavy quarks are distinguished from light quark events by the long lifetimes of  $c$  and  $b$  flavoured hadrons, which lead to displacements of tracks from the primary vertex. This technique, based on the precise spatial information from the H1 silicon vertex detector, was introduced in the measurements of the charm and beauty structure functions  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$  in deep inelastic scattering [5, 6] and is now applied to dijet events in photoproduction in this thesis. This method allows the first simultaneous measurement of charm and beauty in photoproduction and extends to larger values of transverse jet momentum than previous measurements [7, 8, 9, 10, 11, 12, 13, 14]<sup>1</sup>.

In pQCD calculations, the photoproduction of charm and beauty proceeds dominantly via the direct photon-gluon fusion process  $\gamma g \rightarrow c\bar{c}$  or  $b\bar{b}$ , where the photon interacts with a gluon from the proton to produce a pair of heavy quarks in the final state. Previous charm measurements have confirmed this prediction [7, 8, 9, 10]. In leading order QCD models a successful description of the data is obtained when additional contributions from processes involving resolved photons are taken into account [8, 9]. In such resolved photon processes the quasi-real photon fluctuates into a hadronic state before the hard interaction and thus acts as a source of partons.

In this thesis heavy quark production is investigated by measuring differential charm and beauty dijet cross sections as functions of observables which are sensitive to the underlying production mechanism. In a further step, flavour inclusive dijet cross sections are also measured and the relative contribution from charm and beauty events to dijet photoproduction is determined.

This thesis is organised as follows. The theoretical background of charm and beauty production at HERA is reviewed in chapter 2. Theoretical calculations performed in the framework of perturbative QCD are also discussed in this chapter. Chapter 3 provides a brief description of the experimental apparatus as well as a short review of recent heavy flavour measurements performed at HERA and other colliders. Chapter 4 introduces the analysis strategy and gives a detailed overview of the track and jet reconstruction, which are the central objects used within this analysis. The selection of the photoproduction dijet event sample is discussed in chapter 5. The details of the track selection are given at the beginning of chapter 6 followed by the description of the method used to determine the contributions of charm and beauty events in the dijet sample. Chapter 6 closes with the investigation of the properties of a subsample of dijet events, in which the fraction of events containing heavy quarks is enriched. The cross section measurements and their systematic uncertainties are presented in chapter 7.

---

<sup>1</sup>The results of this measurement have been published in [15].

## CHAPTER 2

# PHOTOPRODUCTION OF HEAVY QUARKS

---

This chapter gives an overview of the theoretical and phenomenological concepts relevant to describe heavy quark production in electron<sup>1</sup>-proton ( $ep$ ) collisions. In the context of this analysis heavy quark production means always charm and beauty quark production, since normal top production is not possible at HERA due to the large mass of the top quark. Starting with some basic features of Quantum Chromodynamics (QCD), the kinematics and properties of  $ep$  collisions at HERA are introduced with emphasis on the various photoproduction processes involved in the description of heavy flavour production. The discussion of heavy flavour fragmentation is followed by a description of the QCD theory calculations used in this analysis. The chapter closes with a short description of the most important properties of heavy hadrons.

### 2.1 Quantum Chromodynamics

In Quantum Electro-Dynamics (QED) the observed charge of the electron is smaller at large distances because of the screening of its electric charge by vacuum polarization. This means that the electron constantly emits virtual photons which can fluctuate into  $e^+e^-$  pairs. The electron is then effectively surrounded by a cloud of  $e^+e^-$  pairs, screening its charge at larger distances. Probing the electron at smaller scales reduces the effect and the fine structure constant,  $\alpha$ , increases.

In hadronic interactions, described by Quantum Chromodynamics, an analogous situation exists, however its explanation is not quite as intuitive as in QED. The new feature in QCD is the presence of gluons, which are massless, spin one particles that carry no electrical charge but an additional quantum number, the colour charge. Unlike photons, gluons are able to couple to one another, and therefore  $g \rightarrow gg$  pair production is possible. Due to a stronger coupling the gluon self-interaction dominates and is the reason for an anti-screening effect, causing the strong coupling constant  $\alpha_s$  to behave differently than  $\alpha$ . At small momentum scales (large distances)  $\alpha_s$  becomes large and the partons interact strongly

---

<sup>1</sup>Although HERA was operated using positrons during the data taking period considered in this thesis, both are generally referred to as electrons throughout this chapter.



**Figure 2.1:** Higher order correction diagrams with gluons. Left: gluon loop. Right: fermion loop.

being bound into colour neutral states called hadrons (*confinement*). With increasing momentum scale,  $\alpha_s$  becomes small until the parton is effectively observed in isolation, an effect referred to as *asymptotic freedom*.

In perturbative QCD (pQCD) particle scattering cross sections are given as a power series in  $\alpha_s$ . Quark and gluon loop diagrams, as for example shown in figure 2.1 start contributing beyond leading order (LO). The calculation of such loop contributions means an integration over all particle momenta  $p$  in the loop. In the limit  $p \rightarrow \infty$  ultraviolet (UV) divergences occur which can be absorbed in the running of  $\alpha_s$  and the introduction of an arbitrary dimensionless renormalisation scale  $\mu_R$ . The requirement that every physical observable has to be independent of the arbitrary choice<sup>2</sup> for  $\mu_R$ , leads to an *effective coupling constant*  $\alpha_s(\mu_R^2)$  with

$$\mu_R^2 \frac{\partial \alpha_s(\mu_R^2)}{\partial \mu_R^2} = \beta(\alpha_s). \quad (2.1)$$

Here, the  $\beta$ -function has a perturbative expansion in  $\alpha_s$  describing the dependence of the strong coupling on the renormalisation scale.

In the one loop approximation the strong coupling constant  $\alpha_s$  at the scale  $\mu_R > \Lambda_{QCD}$  can be written as

$$\alpha_s(\mu_R^2) = \frac{1}{b \cdot \ln\left(\frac{\mu_R^2}{\Lambda_{QCD}^2}\right)}, \quad b = \frac{1}{12\pi}(33 - 2n_f). \quad (2.2)$$

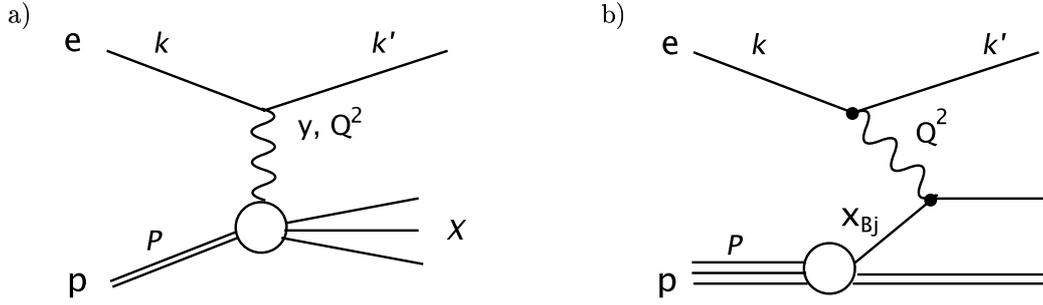
Here  $n_f$  indicates the number of quark flavours with mass below  $\mu_R$  and  $\Lambda_{QCD}$  is the scale at which  $\alpha_s$  gets large, such that the perturbative series in  $\alpha_s$  no longer converges and pQCD is not applicable anymore. If  $\alpha_s$  is known at some value  $\mu_R$ , the effective coupling can be obtained for any scale (eq. 2.2). Experimentally  $\Lambda_{QCD}$  was found to be of the order of  $\sim 200$  MeV.

The  $ep$  collisions at HERA provide a good testing ground for pQCD. QCD calculations at large momentum scales yield more reliable results as discussed above. In this analysis the large momentum scale, referred to as the *hard scale*, is achieved through the production of jets above a certain transverse energy threshold and through the study of heavy quarks.

## 2.2 Basics of Electron Proton Scattering

The generic electron-proton scattering process  $ep \rightarrow lX$  is described by the exchange of a single virtual gauge boson ( $\gamma$ ,  $Z$  or  $W$ ) as illustrated in figure 2.2 a). An incoming

<sup>2</sup>In practice,  $\mu_R$  has to be chosen close to the physical scales characterising the process under study.



**Figure 2.2:** a) Electron-Proton scattering. b) Electron-Proton scattering in lowest order perturbation theory using the parton model.

electron with 4-momentum  $k$  scatters off a proton with 4-momentum  $P$ . The final state consists of the scattered lepton with 4-momentum  $k'$  and a hadronic system  $X$ , formed by the scattered parton<sup>3</sup> and the proton remnant. Photon or  $Z$  exchange are also called *neutral currents* (NC),  $ep \rightarrow eX$ , whereas the exchange of a  $W$  boson is referred to as *charged current* (CC),  $ep \rightarrow \nu_e X$ . For fixed beam energies, two variables are needed to describe the kinematics of such a deep inelastic scattering (DIS) process. However, it is customary to define several variables and then use those most convenient for a particular problem. The main variables generally used are given below.

The squared centre-of-mass energy of the reaction is

$$s = (P + k)^2. \quad (2.3)$$

The negative squared 4-momentum transfer  $Q^2$  of the electron to the proton is given by

$$Q^2 = -q^2 = -(k - k')^2. \quad (2.4)$$

The inelasticity  $y$ , defined as

$$y = \frac{P \cdot q}{P \cdot k}, \quad (2.5)$$

describes the relative energy transfer at the electron-boson vertex in the proton rest frame. The Bjorken scaling variable is given by

$$x = \frac{Q^2}{2P \cdot q}. \quad (2.6)$$

When the struck parton is collinear to the proton (*infinite momentum frame*), the Bjorken scaling variable  $x$  represents the fraction of the 4-momentum of the proton carried by the struck parton (see figure 2.2 b)). Both  $x$  and  $y$  take values between zero and unity.

Neglecting particle masses, the above given quantities are related by

$$Q^2 = x \cdot y \cdot s. \quad (2.7)$$

In the infinite momentum frame ( $P^2 \gg m_P^2$ ) the proton can, to good approximation, be described as a parallel stream of independent partons which carry a fraction  $\xi_{p,i}$  of the longitudinal proton momentum, such that  $\sum_i \xi_{p,i} = 1$ . Deep inelastic scattering processes

<sup>3</sup>Partons are the pointlike constituents (quarks and gluons) of the proton.

are thus interpreted as elastic electron scattering from a single parton. The other partons, which form the proton remnant, do not participate in the hard interaction and are referred to as spectator partons. Due to the increasing strong force with increasing distances, the individual partons in the proton are not directly visible, and the proton content can be described by universal probabilistic parton densities. Since the partons in the proton have been identified as quarks and gluons, for each quark flavour and the gluon parton density functions (pdfs) exist which give the number density of finding a parton  $i$  with a momentum fraction  $\xi_p$ .

In figure 2.2 a photon exchange process is used to demonstrate the HERA kinematics. Neutral current  $Z$  boson exchange and charged current  $W$  boson exchange are also observed at HERA. However, due to the high mass of these bosons, the photon exchange dominates by far in the region of low  $Q^2$ . In the following only the pure photon exchange is considered. The differential NC cross section for the process  $e^\pm p \rightarrow e^\pm X$  is then given by [17]

$$\frac{d^2\sigma_{NC}^\pm}{dx dQ^2} = \frac{2\pi\alpha^2}{xQ^4} [(1 + (1 - y)^2)F_2 - y^2F_L] . \quad (2.8)$$

The fine structure constant is denoted as  $\alpha$  and  $F_2(x, Q^2)$  and  $F_L(x, Q^2)$  are called proton structure functions, which parameterize the structure of the proton as probed by the virtual photon. The contribution from  $F_L(x, Q^2)$  is kinematically suppressed compared to  $F_2(x, Q^2)$  and can lead to sizeable effects only at very large inelasticities  $y$ , as  $F_L$  is always smaller than  $F_2$ .

### 2.2.1 Factorisation and Parton Evolution

In a static model the proton is built only of the valence quarks  $u$  and  $d$ . In the naïve Quark Parton Model (QPM), taking quarks and antiquarks into account,  $F_2$  depends only on  $x$  and can be written as

$$F_2(x) = x \sum_q e_q^2 (f_q^p(x) + f_{\bar{q}}^p(x)) . \quad (2.9)$$

The sum runs over all quark flavours  $q$ ;  $e_q$  are the quark charges and the functions  $f_q^p$  ( $f_{\bar{q}}^p$ ) contain the quark (anti-quark) densities in the proton.

Early experimental results on  $F_2$  were effectively in agreement with the so-called *scaling* behaviour of  $F_2$ , i.e. no  $Q^2$ -dependence of  $F_2$  was seen first in the accessible  $x$ -range. Later, from the observation of *scaling violations* at lower  $x$  values, it was concluded that also gluons and gluon splittings into pairs of quarks and anti-quarks have to be considered for the successful description of the proton content.

The concept of proton parton densities is motivated theoretically by the *factorisation theorem*, which separates the *short distance part* (hard subprocess) and the *long distance part* (soft processes). Whereas the hard subprocess, describing the interaction of high energy partons, can be calculated within pQCD, the soft (low energy) processes are not accessible to perturbative calculations.

Within this framework, the proton structure function  $F_2$  can be written as a convolution of perturbatively calculable coefficient functions  $C^i$  and parton density functions (pdf)

$f_{i/p}(\xi)$ , which can be interpreted as the probability of finding a parton of type  $i$  carrying a fraction  $\xi$  of the proton's longitudinal momentum:

$$F_2(x, Q^2) = x \sum_{i=q, \bar{q}, g} \int_x^1 d\xi C^i \left( \frac{x}{\xi}, \frac{Q^2}{\mu_R^2}, \frac{\mu_F^2}{\mu_R^2}, \alpha_s(\mu_R^2) \right) f_{i/p}(\xi, \mu_F^2, \mu_R^2). \quad (2.10)$$

Here,  $\mu_R$  is the renormalisation scale (already discussed in section 2.1) which is used to regulate the ultraviolet divergences coming from virtual corrections. These divergences are then absorbed into the running of the strong coupling  $\alpha_s(\mu_R^2)$ . The regularisation has to be done in a well defined way. The scheme most often used is the *modified minimal subtraction* scheme ( $\overline{MS}$ -scheme).

The other scale introduced by the factorisation procedure is the factorisation scale  $\mu_F$ . The calculation of gluon radiation ( $q \rightarrow qg$ ,  $g \rightarrow gg$ ) and splittings ( $g \rightarrow q\bar{q}$ ) in pQCD gives rise to infrared divergent integrals, e.g. for collinear or soft gluon radiation. The factorisation scale  $\mu_F$  is an arbitrary cut-off parameter which separates the perturbative and non-perturbative regions and the divergences are then absorbed in the parton density functions. For instance gluon radiation at large transverse momenta is attributed to the partonic cross section while gluon radiation at small transverse momenta is part of the parton densities. If  $\mu_F \gg \Lambda_{QCD}$ , reliable pQCD calculations are possible. Several factorisation schemes exist. The most important are the  $\overline{MS}$ -scheme and the *deep inelastic scattering* scheme (DIS-scheme).

Based on the factorisation theorem, in the presence of a hard scale the  $ep$  cross section can be decomposed as

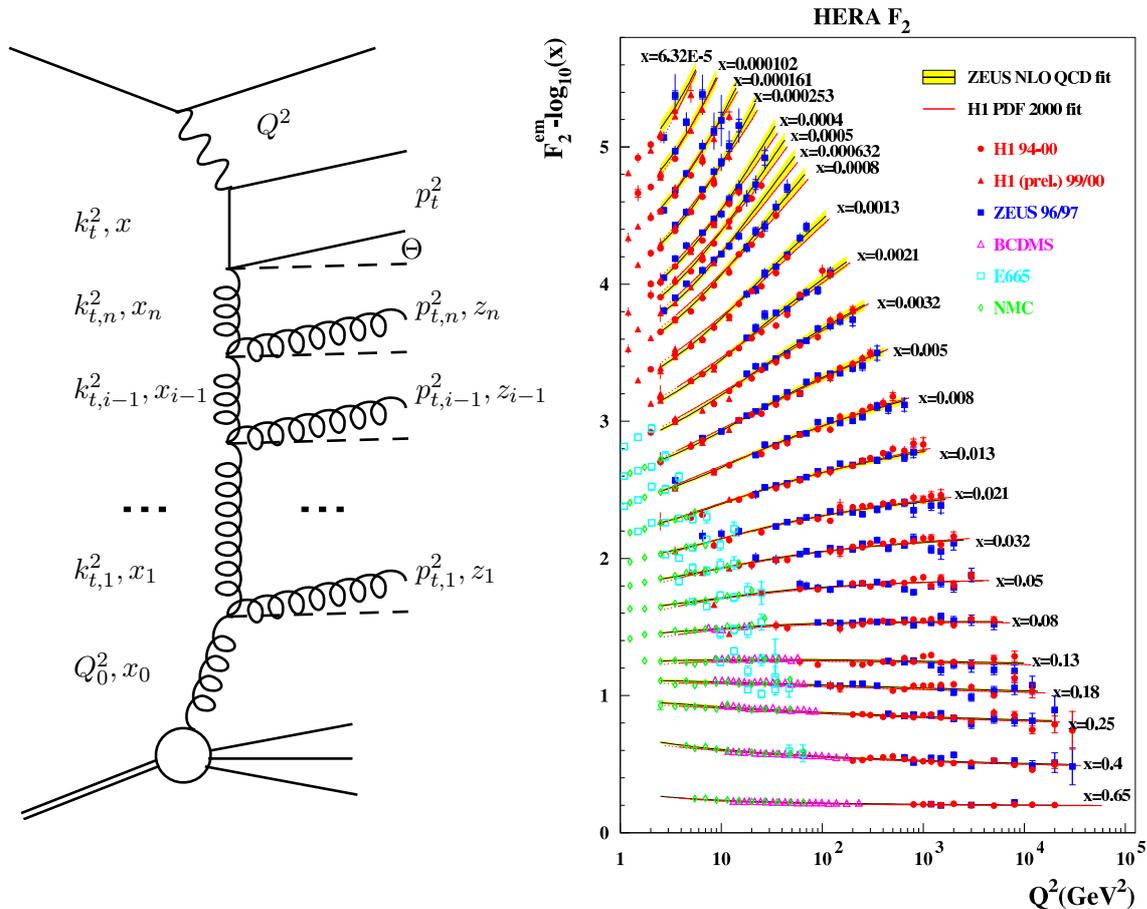
$$\sigma_{ep} \sim f_{i/p}(\xi, \mu_F^2) \otimes \hat{\sigma}_{ij}(\hat{s}, \alpha_s(\mu_R^2), \mu_R, \mu_F) \otimes f_{i/\gamma}(\xi_\gamma, \mu_F^2) \otimes D(z), \quad (2.11)$$

where  $\hat{\sigma}_{ij}$  denotes the cross section of the hard subprocess calculable in pQCD. Note, that parton density functions are no physical observables in contrast to the cross sections and structure functions. The parton density function of the photon  $f_{i/\gamma}(\xi_\gamma, \mu_F^2)$  and the fragmentation function  $D(z)$ , introduced in equation 2.11, will be discussed in sections 2.2.4 and 2.4, respectively.

## Parton Evolution Models

The parton densities discussed above depend on the scale  $\mu_F$  and cannot be calculated from first principles in pQCD. Nonetheless, since the observable cross section must be independent of the unphysical parameter  $\mu_F$  and  $\hat{\sigma}_{ij}$  is perturbatively calculable for different values of  $\mu_F$ , the dependence of the parton densities on  $\mu_F$  can be studied within pQCD. This leads to the *parton evolution equations* which predict the evolution to any scale  $\mu_F > \mu_0$  once  $f_{i/p}(\xi, \mu_0^2)$  is known at a starting scale  $\mu_0$ . In practice the pdfs are determined by a fit to the experimental  $F_2(x, Q^2)$  data from HERA and fixed target experiments. The relevant scale used in this procedure is  $\mu_F^2 = Q^2$ .

Although the concept of parton evolution is universal, in practical calculations approximations must be applied, leading to different parton evolution models, which are expected to be valid only in certain regions of phase space. To obtain the evolution equations, gluon radiation and gluon splitting processes must be considered, which can occur several times. This can result e.g. in a gluon ladder, as shown in figure 2.3, formed from n



**Figure 2.3:** Parton evolution diagram (left). Measurement of the proton structure function  $F_2(x, Q^2)$  as a function of  $Q^2$  at fixed values of  $x$  (right). Shown are the results from H1, ZEUS and various fixed target experiments. The lines correspond to the QCD fits performed by H1 and ZEUS.

gluon emissions from gluons with transverse momenta  $k_{t,i}$  and longitudinal momenta  $x_i$  with  $x_i > x_{i+1}$ . The transverse momentum of the emitted gluons is denoted as  $p_{t,i}$ , their longitudinal momentum as  $z_i$  and the radiation angle as  $\Theta_i$ .

In this analysis, two different parton evolution models are considered, which differ mainly in the kinematical ordering of the emission processes.

In the DGLAP approach [18] a strong ordering of the transverse momenta  $k_{t,i}$  is assumed, i.e.

$$Q_0^2 \ll \dots \ll k_{t,i}^2 \ll k_{t,i+1}^2 \ll \dots \ll k_{t,n}^2 \ll Q^2. \quad (2.12)$$

The DGLAP evolution equations are integro-differential equations which consist of terms where either a gluon is radiated from a parton or by considering a parton evolved from a gluon in a pair production process. The solution of the DGLAP equation results in the parton distributions as a function of  $x$  at any scale  $Q^2$ , provided that their  $x$  dependence at the starting scale  $Q_0^2$  is known. The DGLAP approximation is expected to be valid in the region where  $Q^2$  is large and  $x$  is not too small.

The CCFM ansatz [19] is a more recent parton evolution model which is based on gluon

emission only. According to the CCFM evolution equations, the gluon emission during the initial cascade is only allowed in an angular ordered region of phase space. The maximum allowed angle  $\Theta$  is defined by the hard scattering producing the quark pairs (cf. figure 2.3). The unintegrated parton density  $\mathcal{A}(x_g, k_t^2, \mu_F^2)$  describes the probability of finding a parton carrying the longitudinal momentum fraction  $x_g$  and transverse momentum fraction  $k_t$  at the factorisation scale  $\mu_F$ . Since  $\mathcal{A}(x_g, k_t^2, \mu_F^2)$  explicitly depends on  $k_t$  it is more complicated than the parton density in the DGLAP picture which depends on  $x$  and  $Q^2$  only.

As the  $k_t$  dependence of the parton densities is integrated out in the DGLAP approach it is sometimes also called *collinear factorisation* whereas the CCFM approach incorporates  *$k_t$ -factorisation*. If  $k_t$  factorisation is used, the partons entering the hard matrix element are free to be off-shell, in contrast to the collinear approach in which all partons entering the hard subprocess are on-shell.

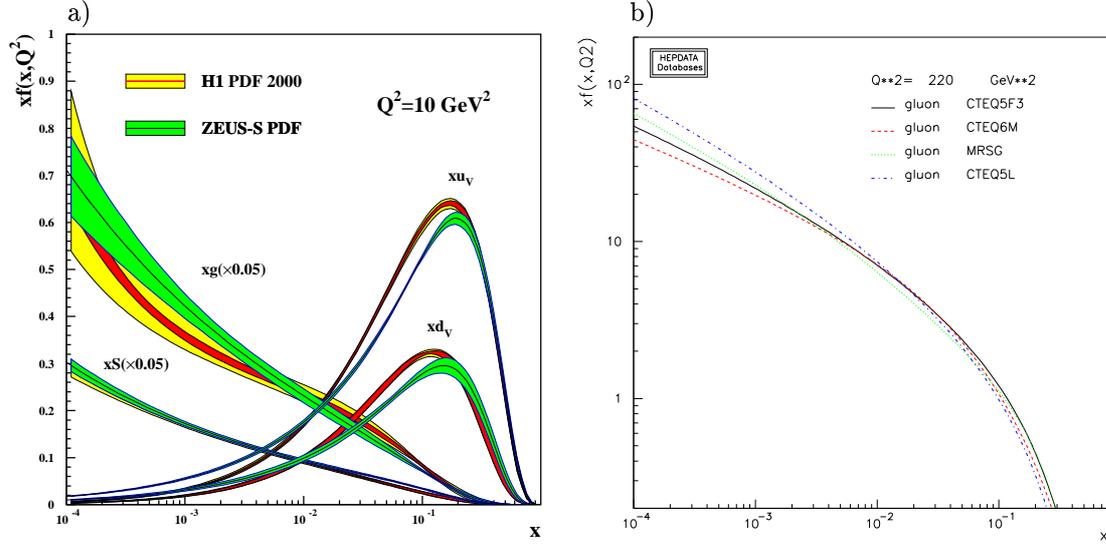
Figure 2.3 illustrates that DGLAP fits based on next-to-leading order QCD very successfully describe the HERA  $F_2(x, Q^2)$  measurements down to the smallest accessible  $x$ . However, for some exclusive processes in certain regions of phase space the need for  $k_t$ -factorisation may be necessary to fully describe the HERA data. Indications for the need of  $k_t$ -factorisation are discussed e.g. in a recent measurement of inclusive forward jet production at low  $x$  [20]. Since in heavy quark production relatively small values of  $x$  ( $10^{-3} - 10^{-4}$ ) are reached, the study of heavy quark production and the comparison to different QCD models incorporating both approaches might also give some insight into this open question.

### 2.2.2 Proton Structure

The parton density functions of the proton are not physical observables but can be determined from fits to the measured  $F_2(x, Q^2)$  data. The quark densities are well known down to very low  $x$ , while the gluon density has still uncertainties of the order 10 to 40% in this region. In figure 2.4 a) the parton density of the proton as a function of  $x$  for a fixed  $Q^2$  is shown. Recent H1 and ZEUS fits are used to extract the quark and gluon densities. It can be seen that the valence quarks  $u$  and  $d$  dominate at high  $x$ , whereas the gluons come into play at lower  $x$  exceeding the quark contribution by far in the lowest  $x$  region. This behaviour is correlated with the large scaling violations of  $F_2(x, Q^2)$  observed at low  $x$ . In the plot the gluon contribution is scaled down by a factor of 20.

A lot of pdf sets from different groups are available which are based on global fits to data (primary inclusive DIS data). Mostly relevant in this analysis are the CTEQ [21] parton distribution functions.

The CTEQ5M set is defined in the  $\overline{MS}$  scheme, matched with NLO matrix elements in the zero-quark-mass approximation for all active flavours, including charm and beauty. CTEQ5L is the corresponding LO set, appropriate for Monte Carlo programs. For the study of heavy quark processes there exist some special parton sets from the CTEQ group. The CTEQ5HQ is valid from the threshold region up to the asymptotic region ( $Q \gg m_Q$ ), but no interface for the NLO calculation program used in this analysis exists. Instead, the CTEQ5F3 parton distribution set is used, which is a fixed-3-flavour scheme treating charm and beauty quarks as heavy particles, not partons. The fixed-flavour-number scheme is valid for charm and beauty production at energy scales not far above the threshold.



**Figure 2.4:** a) Gluon and valence quark ( $u$  and  $d$ ) density of the proton as obtained from recent H1 and ZEUS fits. b) Gluon density of the proton as given by the  $CTEQ5F3$ ,  $CTEQ6M$ ,  $CTEQ5L$  and  $MRST$  pdfs. The scale relevant in this analysis is given by  $\hat{p}_t^2 + m_Q^2 = 220 \text{ GeV}^2$ .

The  $MRST$  group [22] also provide different pdf sets. Likewise they use NLO evolution equations in the  $\overline{MS}$  scheme and differ from the  $CTEQ$  fits e.g. in their experimental input. The  $MRST$  distributions are used in this work for the estimation of the systematic uncertainty of the NLO calculation which is done with the  $CTEQ5F3$  parameterisation. Figure 2.4 b) illustrates the various gluon densities of the proton as a function of  $x$ , as they are predicted by the different parametrisations for the scale relevant in this analysis.

### 2.2.3 Photoproduction and the Photon Flux

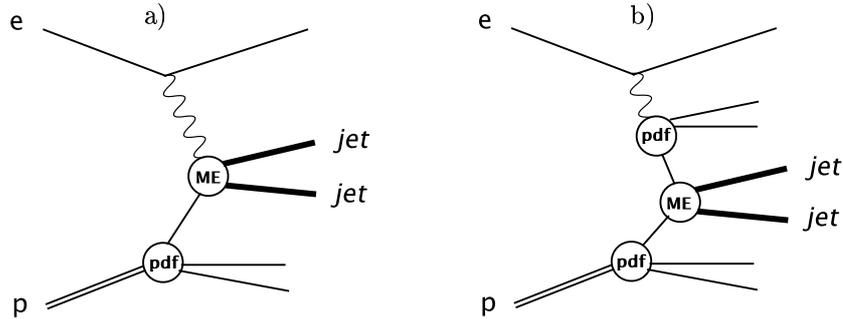
In the region of small  $Q^2$  electron-proton scattering can be simplified by factorising the radiation of a photon from the electron and the subsequent interaction of the photon with the proton ( $\gamma p$ ). This particular domain of electron-proton scattering is called photoproduction. Since the photon propagator gives rise to a factor  $1/Q^2$  in the inclusive NC DIS cross section, photoproduction dominates  $ep$ -cross sections. In  $\gamma p$  processes the electron radiates quasi-real ( $Q^2 < 1 \text{ GeV}^2$ ) photons with energy fractions  $y$  according to the *Weizsäcker-Williams approximation*. The variable  $y$  is now interpreted as the longitudinal momentum fraction of the incoming electron taken by the photon. It is directly related to the centre-of-mass energy  $W_{\gamma p}$  of the photon proton system

$$W_{\gamma p}^2 = (q + P)^2 = y \cdot s - Q^2 \approx y \cdot s, \quad (2.13)$$

where  $P$  is the proton and  $q$  the photon 4-momentum. The proton and electron mass is here and in the following expressions neglected.

The hadronic cross section, e.g. for the photoproduction of jets,  $\sigma_{ep \rightarrow eX}$ , factorises into two contributions

$$\sigma_{ep \rightarrow eX} = \int dy f_{\gamma, e}(y) \sigma_{\gamma p \rightarrow eX}(y), \quad (2.14)$$



**Figure 2.5:** a) Direct and b) resolved photoproduction of jets in leading order.

where  $\sigma_{\gamma p \rightarrow eX}$  denotes the cross section for the scattering of a real photon from a proton and  $f_{\gamma,e}$  the photon flux. Using the Weizsäcker-Williams approximation [23, 24], which neglects terms involving the longitudinal photon polarisation, the flux of photons with energy fraction  $y$  and up to a virtuality  $Q_{max}^2$  can be calculated via

$$f_{\gamma,e}(y) = \frac{\alpha}{2\pi} \left[ \frac{1 + (1-y)^2}{y} \ln \left( \frac{Q_{max}^2}{Q_{min}^2} \right) - 2m_e^2 y \left( \frac{1}{Q_{min}^2} - \frac{1}{Q_{max}^2} \right) \right]. \quad (2.15)$$

The fine structure constant is denoted as  $\alpha$  and  $Q_{min}^2 = \frac{m_e^2 y}{1-y}$  is the kinematic lower limit. The maximum photon virtuality  $Q_{max}^2$  is given by the experimental conditions and is related to the scattered electron's energy  $E_{el}$  and angle  $\theta_{el}$  via

$$Q_{max}^2 = 2E_e E_{el} (1 + \cos \theta_{el}), \quad (2.16)$$

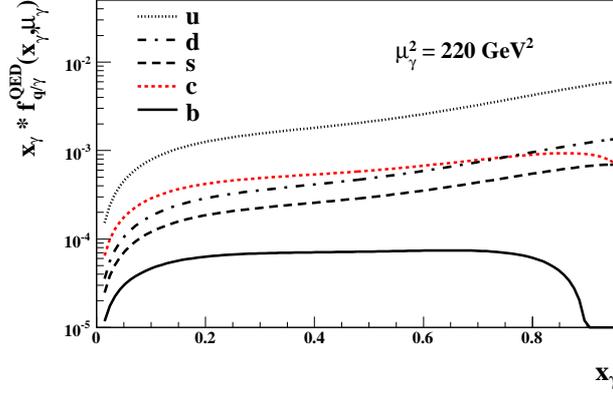
with the energy of the incoming electron beam denoted as  $E_e$ .

### 2.2.4 Photon Structure

In figure 2.5 a) the electron-proton scattering process producing jets in the photoproduction regime is illustrated. In QCD this process is understood by the hard interaction of real photons with the partons inside the proton. Of main interest for this analysis are the processes with at least two final state partons with large transverse momenta.

Two interaction classes have to be considered. If the photon interacts directly as a bare photon with a parton of the proton, it is called *direct photoproduction*. Besides, it is possible that the photon fluctuates in a  $q\bar{q}$  state before the interaction, which is then called *resolved photoproduction*. In resolved photoproduction usually two different contributions are distinguished. In the first case the  $q\bar{q}$  pair may form a vector meson which carries the quantum numbers of the photon and it then interacts with the proton like a hadron. This case is addressed by the Vector Meson Dominance Model (VDM). The VDM model gives good agreement with low energy data but its contribution to this measurement at higher energies is negligible.

The photon may also fluctuate into an unbound quark pair which is subject to gluon radiation and gluon splitting. The photon then acts as a source of partons or - in other words - the photon reveals a hadronic structure. Similar to the case of the proton (cf. figure 2.5 b)), the calculation of the latter process can be approximated by ascribing parton densities to the photon. Taking into account only the QED contributions, the probability



**Figure 2.6:** Quark density of the photon (QED result).

$f_{q/\gamma}$  of finding a quark within the photon carrying a momentum fraction  $x_\gamma$  of the initial photon momentum can be written as

$$f_{q/\gamma}(x_\gamma, \mu_\gamma^2) = e_q^2 \cdot \frac{\alpha}{2\pi} (x_\gamma^2 + (1 - x_\gamma)^2) \ln \left( \frac{\mu_\gamma^2(1 - x_\gamma)}{m_q^2 x_\gamma} \right). \quad (2.17)$$

The electric charge of the quark  $q$  is  $e_q$  and  $m_q$  is the effective mass of a free quark. The momentum fraction  $x_\gamma$  is defined as

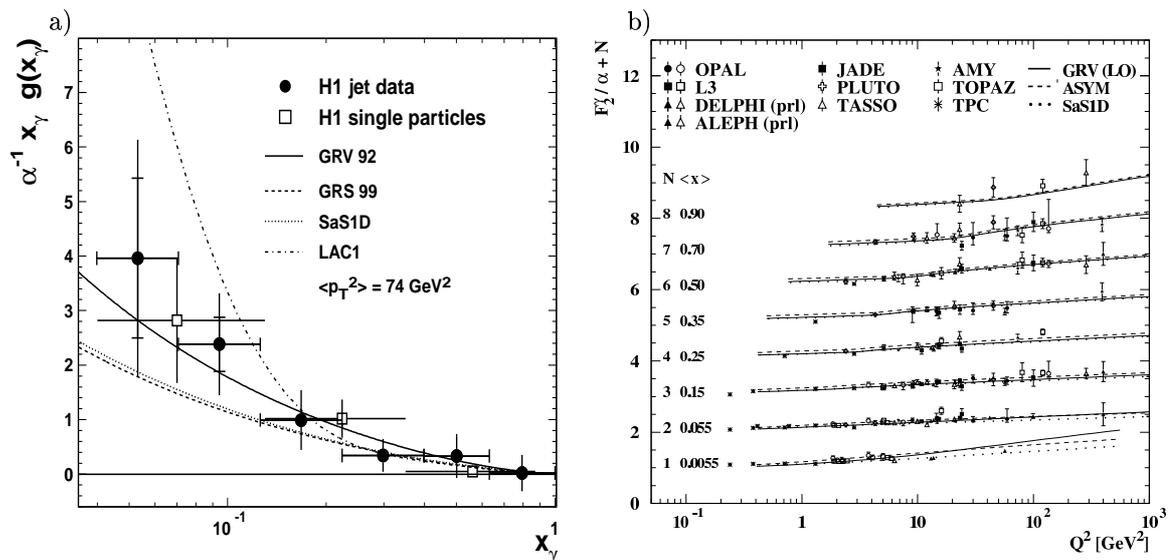
$$x_\gamma = \frac{P \cdot p_j}{P \cdot q}, \quad (2.18)$$

where  $p_j$  is the 4-momentum of the parton inside the photon. For direct processes  $f_{q/\gamma}(x_\gamma, \mu_\gamma)$  is given as the Dirac delta function  $\delta(1 - x_\gamma)$ . Summing over all colours  $n_c$  and flavours  $n_f$  results in the photon structure function  $F_2^\gamma$

$$\begin{aligned} F_2^\gamma(x_\gamma, \mu_\gamma^2) &= 2x_\gamma \sum_{n_c, n_f} e_q^2 f_{q/\gamma}(x_\gamma, \mu_\gamma^2) \\ &= 3 \sum_{n_f} e_q^4 \frac{\alpha}{2\pi} x_\gamma (x_\gamma^2 + (1 - x_\gamma)^2) \ln \left( \frac{\mu_\gamma^2(1 - x_\gamma)}{m_q^2 x_\gamma} \right). \end{aligned} \quad (2.19)$$

Unlike the proton structure function  $F_2^p$ , the quark charge contributes to the fourth power in  $F_2^\gamma$ .  $F_2^\gamma$  depends directly on the scale  $\mu_\gamma^2$  which can be interpreted as the virtuality of the particle probing the photon fluctuations. In  $F_2^p$  the scale enters only via the QCD evolution equations.

In figure 2.6 the distribution of  $x_\gamma \cdot f_{q/\gamma}$  following equation 2.17 is shown. The scale is set to  $\mu_\gamma^2 = \hat{p}_t^2 = 220 \text{ GeV}^2$  for a comparison with the present analysis. For all quark flavours  $x_\gamma \cdot f_{q/\gamma}$  increases with increasing  $x_\gamma$  and  $x_\gamma \lesssim 0.8$ . The contribution of the beauty quarks (full line) is very small due to its large mass. The probability of finding a charm quark in the photon, however, is much larger. Charm quarks even exceed the contribution of the lighter  $d$  valence quark for a wide range of  $x_\gamma$ , which is due to the larger quark charge value.



**Figure 2.7:** a) Measurement of the gluon distribution function of the photon from H1 [27] and b) summary of  $F_2^\gamma$  measurements from [26].

QCD corrections have been calculated to the QED result [25]. They preserve the logarithmic dependence on  $\mu_\gamma^2$  and thus equation 2.17 still gives a qualitative comparison of the difference of the various quark contributions to the photon structure. At very low (high) values of  $x_\gamma$  however, the QCD corrections are not negligible anymore.

Photon quark densities have been determined mainly in  $e^+e^-$  reactions. A nice review of this topic can be found in [26]. The photon structure function  $F_2^\gamma$  has been investigated up to very large scales, but the experimental uncertainties at high  $x_\gamma$  values and high scales are still large.

The knowledge about the gluon content in the photon is still less obvious. The gluon density can be calculated from the  $F_2^\gamma$  measurements only via the evolution model. As an example figure 2.7 illustrates H1 measurements of the gluon distribution function of the photon [27]. The two results, obtained from single particles and jets respectively, are consistent. In contrast to the quark density function (figure 2.6) the gluon distribution function rises towards small  $x_\gamma$  and is small at large  $x_\gamma$ . The presence of gluons at low  $x$  gives rise to scaling violations as seen in figure 2.7 b) where measurements of  $F_2^\gamma$  are summarized [26]. However, the scaling violations are much less pronounced than in the proton case, meaning that there are only few gluons in the photon and the resolved photon consists mostly of quarks.

In the figure the data are compared to different theoretical parametrisations of the photon structure. Some of them, which are important for this work, will be briefly described below. The experimental uncertainties seen in figure 2.7 a) are large and the theoretical predictions can not be distinguished. Nevertheless, the measurement from  $ep$  collisions confirms the universality of the parton distribution functions of the photon as determined in  $e^+e^-$  reactions.

Many parton distribution functions exist for photons in leading and next-to-leading order which are constructed in a way similar to the parton distribution functions of the proton. They differ in the assumptions made about the starting scale  $Q_0^2$ , the input distributions assumed at this scale and in the amount of data used in fitting their parameters.

The photon pdfs which are important for this analysis are:

**GRV** The parton distributions from Glück, Reya and Vogt [28] use a hadron like starting contribution based on VMD arguments, by applying the parton distribution functions of the pion. The point-like contribution is chosen to vanish at  $Q^2 = Q_0^2$  and for larger  $Q^2$  it is generated dynamically using the photon evolution equations. The charm and beauty quarks are included via the Bethe-Heitler formula with masses  $m_c = 1.5 \text{ GeV}$  and  $m_b = 4.5 \text{ GeV}$ . At high values of  $W$  they are treated as massless quarks in the evolution. The pdfs are available in LO starting at  $Q_0^2 = 0.30 \text{ GeV}^2$  (GRV-LO) and NLO with a starting scale of  $Q_0^2 = 0.25 \text{ GeV}^2$  (GRV-HO).

**SaS** The parton distributions from Schuler and Sjöstrand [29] use the same basic approach as the GRV distributions but in addition relate the input distribution functions to the measured photon-proton cross section. Two sets exist which differ in the starting scale, SaS1 using  $Q_0^2 = 0.36 \text{ GeV}^2$  and SaS2 using  $Q_0^2 = 4 \text{ GeV}^2$  and in the amount of the VMD contribution. For the SaS1 distributions which are used in this analysis only the shape of the VMD distribution is fitted to the data of the photon structure function. As in GRV the massless evolution equations for light quarks are used and the Bethe-Heitler formula for contributions of charm and beauty quarks, here with the masses  $m_c = 1.3 \text{ GeV}$  and  $m_b = 4.6 \text{ GeV}$ .

The parton distribution used in this work is the SaS1D parton distribution, calculated in the  $DIS_\gamma$  scheme. In the kinematic range of this measurement, the differences when using the GRV-LO parton distributions have been found to be small ( $< 3\%$ ).

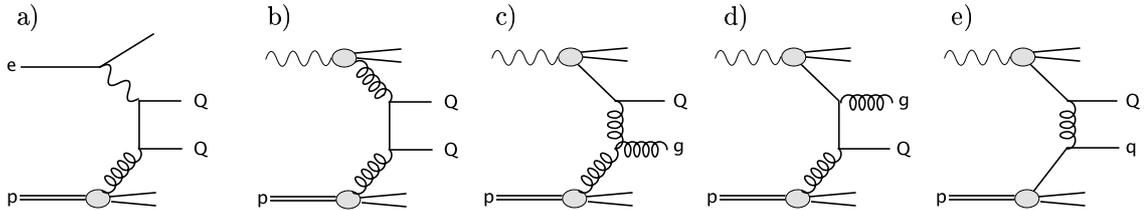
## 2.3 Heavy Quark Production

All aspects discussed up to this point are relevant for  $ep$  scattering processes irrespective of the actual final state configuration. From now on the discussion focuses on issues characteristic for the production of heavy quarks.

While the light  $u$ ,  $d$  and  $s$  quarks have masses below the  $\Lambda_{QCD}$  parameter, the masses of the heavy quarks are already high enough to serve as a hard scale for perturbative calculations. This feature sets heavy flavour production apart from other processes. At HERA heavy flavour physics means the production of  $c$  and  $b$  quarks since the  $t$  (top) quark is out of reach due to its very large mass.

The intrinsic hard scale set by the heavy quark mass can lead to complications in the theoretical calculations since in certain configurations more than one hard scale might be present. In the case of the analysis presented here the transverse momentum of the quarks is always larger than two times the beauty quark mass, thus the hard scale  $\mu$  is already given by the event selection. There might be other hard scales present, like e.g. the photon virtuality  $\mu^2 = Q^2$ . Due to the presence of multiple scales heavy quarks cannot be treated for all energies in the same way.

Different ways of organizing the perturbative calculation are necessary, depending on the relative magnitudes of  $m_Q$  and  $\mu$ . Two standard approaches exist which reduce the two scale problem to an effective and hence approximate one scale problem. In the *massless*



**Figure 2.8:** Heavy quark production processes in LO QCD: a) direct process (BGF) and b)-e) resolved processes. Heavy flavour excitation is shown in figures c)-e).

*parton model approach* all quark masses are set to zero as soon as the typical energy scale  $\mu$  is above the mass threshold  $m_Q$ . Contrary, in the *massive heavy quark approach* the quark is always treated as a heavy particle and not as a parton. The heavy quark masses are explicitly taken into account as a parameter.

In the following sections the heavy quark production mechanism is introduced within the massive scheme. More details on the calculation schemes are discussed in section 2.5.1, in which the NLO calculation program is outlined.

### Direct Photon Processes

In leading order, i.e.  $\mathcal{O}(\alpha\alpha_s)$ , the dominant contribution to the heavy quark production cross section arises from *boson-gluon fusion* (BGF), where a photon emitted from the incoming electron and a gluon coming from the proton form a heavy quark ( $Q\bar{Q}$ ) pair (figure 2.8 a)). According to [30] the cross section can be written as

$$\hat{\sigma}_{BGF}(\hat{s}, m_Q^2) = \frac{\pi e_Q^2 \alpha \alpha_s}{\hat{s}} \left[ (2 + 2\omega - \omega^2) \ln \left( \frac{1 + \chi}{1 - \chi} \right) - 2\chi(1 + \chi) \right]. \quad (2.20)$$

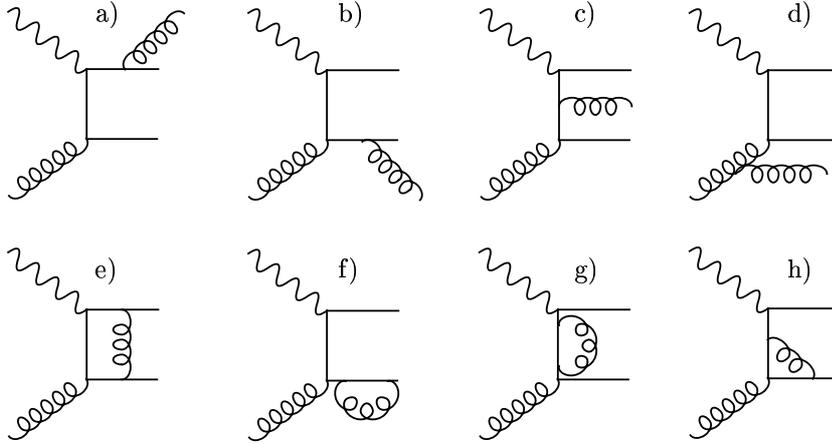
Equation 2.20 holds for the massive approach, i.e. the  $c$  and  $b$  quarks are treated as heavy particles. The centre-of-mass energy  $\hat{s}$  of the heavy quark pair is defined as  $\hat{s} = (p_Q + p_{\bar{Q}})^2$  and  $\omega$  and  $\chi$  are given by  $\omega = 4m_Q^2/\hat{s}$  and  $\chi = \sqrt{1 - \omega}$  respectively. The beauty cross section is heavily suppressed with respect to the charm cross section by the factor  $e_Q^2/\hat{s}$ , due to the relatively larger mass and the smaller charge of the  $b$  quark.

In the massless limit, i.e.  $(m_c, m_b) \rightarrow 0$ , the charm cross section is expected to be only a factor of four larger than the beauty cross section for the BGF process.

### Resolved Photon Processes

In addition to the direct component, resolved photon processes (cf. section 2.2.4) may have a sizeable contribution to the photoproduction cross section. Some of the diagrams contributing in LO QCD are shown in figures 2.8 b)-e). In the *hadron-like resolved* process a gluon or a light quark out of the photon interacts with a parton of the proton (figure 2.8 b). In other photon resolved processes the heavy quark comes directly out of the hadronic fluctuation of the photon (figures 2.8 c-e). These are summarised under the term *heavy quark excitation*.

In LO QCD models a successful description of the data is obtained when the resolved components are taken into account. For light quark photoproduction [1] the resolved component has been found to be dominant and there is also strong experimental indication



**Figure 2.9:** Some of the NLO corrections to the BGF process in the massive scheme. a)-d) Gluon radiation, e)-h) virtual corrections.

for a significant resolved contribution to charm production [9]. In the case of beauty production the situation has not been clarified yet.

Note, that the separation of the direct and resolved components is only unambiguous in the LO picture.

### NLO contributions

Perturbative calculations of heavy quark production are available up to NLO. NLO corrections account for the presence of hard parton emissions and the contribution to the total rates is expected to be significant. Some of the contributing graphs within the massive scheme are shown in figure 2.9. Due to gluon emissions and virtual corrections, ultraviolet, infrared and collinear singularities appear. Renormalisation of charge and mass remove the ultraviolet divergences. The infrared divergences cancel with loop contributions and the collinear divergences are absorbed in fragmentation functions and pdfs [31]. As a consequence the cross section depends directly on the renormalisation and factorisation scale  $\mu_R$  and  $\mu_F$ , respectively. At NLO there is no unique distinction between direct and resolved processes, only their sum has a meaningful physical interpretation. Besides, there are ambiguities between diagrams in LO and NLO. The diagrams in figure 2.9 d) and figure 2.8 c) describe effectively the same process, meaning that heavy flavour excitation is in principle a contribution at NLO.

Finally a word on the possible production of charm and beauty in the fragmentation process via gluon splitting  $g \rightarrow Q\bar{Q}$ . At HERA this process is very rare since it is of the order  $\mathcal{O}(\alpha\alpha_s^2)$  and thus heavily suppressed compared to the BGF process. Its contribution to charm quark production has been estimated by H1 and is approximately 2% [32]. Due to the larger mass of the  $b$  quark compared to the  $c$  quark, the contribution to beauty production should be even smaller. But e.g. for the LHC the gluon splitting into a heavy quark pair ( $gg \rightarrow gQ\bar{Q}$ ) is an important contribution to heavy flavour production since the cross section of  $gg \rightarrow gg$  is much larger than for the process  $gg \rightarrow Q\bar{Q}$  (see also section 3.2.2).

## 2.4 Heavy Quark Fragmentation

Due to colour confinement, a measurement of free  $c$  and  $b$  quarks can not be performed directly. Instead, the strongly interacting outgoing partons are visible as colour neutral hadrons resulting from parton *hadronisation*, also referred to as *fragmentation*. Since at large distances the strong coupling constant rises strongly spoiling the perturbative calculations, phenomenological models are used to describe the fragmentation process.

In the factorisation ansatz (cf. eq. 2.11), the fragmentation process is treated to be independent of the hard scattering process. For the charm case this assumption has been tested by a recent H1 measurement [33] in which the fragmentation function of charm quarks into  $D^{*\pm}$ -mesons has been measured and has been shown to agree with results coming from  $e^+e^-$  experiments, thus proving the universality of the fragmentation process. For heavy quarks a perturbative treatment of the fragmentation is partially possible and calculations are available in leading and next-to-leading order [34]. The probability  $D_Q^h$  to produce a hadron  $h$  from the original heavy quark  $Q$  can be split into short and long distance parts, as

$$\mathcal{D}_Q^h(z, \mu_F) = \int_z^1 \mathcal{D}_Q(x, \mu_F) \cdot \mathcal{D}_i^h(z/x) \cdot dx. \quad (2.21)$$

Starting with a heavy quark produced at the scale  $\mu_F$  the perturbative fragmentation function  $\mathcal{D}_Q(x, \mu_F)$  describes its evolution to a quark on mass shell via subsequent gluon emissions. The dependence of  $\mathcal{D}_Q(x, \mu_F)$  on the scale  $\mu_F$  can be described by the DGLAP evolution equations. This has inspired the model of *parton showers* which is used to simulate the parton evolution on a statistical basis in Monte Carlo programs. A parton shower develops through successive parton splittings until the partons reach a scale where the perturbative treatment becomes unreliable ( $\approx 1$  GeV).

The transition from the quarks on mass shell to hadrons,  $\mathcal{D}_i^h(z/x)$ , do not follow from first principles and must be modelled phenomenologically. Two approaches, the *independent* and the *string* fragmentation are relevant for the analysis presented here.

As the name suggests, in the model of independent fragmentation all outgoing partons hadronise independently. The transition probability from a parton  $i$  to a hadron  $h$  carrying a fraction  $z$  of the parton's longitudinal momentum is given by fragmentation functions  $D_i^h(z)$ . To create e.g. a bound meson  $Q\bar{q}$ , the light quark is taken from a  $q\bar{q}$  pair created out of the vacuum with the remaining quark continuing the hadronisation process until the available energy is used up. The hadron's transverse momentum distribution with respect to the original quark is assumed to be Gaussian. The *Peterson fragmentation function* [35]

$$\mathcal{D}_Q^h(z) \sim \frac{1/z}{(1 - 1/z - \epsilon_Q/(1 - z))^2} \quad (2.22)$$

realises independent fragmentation and is widely used in HERA heavy flavour analyses. The free parameter  $\epsilon_Q$  has to be determined experimentally and its value differs according to the perturbative treatment applied. At NLO additional gluon radiation has to be considered and thus  $\epsilon_Q^{NLO}$  is effectively smaller than  $\epsilon_Q^{LO}$ . Comparing charm and beauty, the experimental results suggest significantly lower  $\epsilon_Q$  values for  $b$  quarks. The fragmentation is harder for  $b$  quarks, meaning that  $D_b^h(z)$  peaks at higher  $z$  values than  $D_c^h(z)$ . This can

be understood by the larger mass of the  $b$  quark which make it easier for the  $b$  quark to carry the light quarks away, leading to larger momentum fractions  $z$  of the hadron.

In an alternative hadronisation model, the string fragmentation, confinement is implemented by stretching colour flux tubes (strings) between  $q\bar{q}$  pairs. The strings represent a linear confinement potential,  $E_{string} = \kappa \cdot \Delta x$ , with string tension constant  $\kappa$  and  $\Delta x$  being the distance between the color charges.  $\kappa$  is of the order of 1GeV/fm. When the quarks move apart, the potential energy in the flux tube increases until at a specific threshold is reached. The string breaks up producing a new  $q\bar{q}$  pair out of the vacuum. Since the creation follows the selection rule  $u\bar{u} : d\bar{d} : s\bar{s} : c\bar{c} \approx 1 : 1 : 0.3 : 10^{-11}$ , the creation of heavy quarks through fragmentation is highly suppressed. The gluons are involved in this model as *kinks* in the colour string. The string splitting processes proceed until the quarks are on their mass shells. Adjacent quarks are then combined to colorless mesons. These are the constituents of the hadronic final state.

Alternatively to the Peterson fragmentation the Lund string fragmentation which incorporates string fragmentation with the Lund fragmentation function [36] is used in this analysis within the framework of Monte Carlo event generators.

But while the Lund string model provides a description of the full hadronic final state, the Peterson model is only applicable for independent heavy quark fragmentation. Thus, in Monte Carlo generator programs Peterson fragmentation for heavy quarks can be optionally added to the Lund string fragmentation which then deals with the light quarks only.

## 2.5 QCD Calculations

Perturbative QCD calculations of charm and beauty production are available up to NLO. The QCD predictions of these calculations will be compared to the measured photoproduction cross sections. In the following, two different types of calculations will be discussed: Monte Carlo event generators and cross section integration programs. The Monte Carlo programs provide an event-by-event prediction for the complete hadronic final state. This is generally needed to study details of the event topology and detector effects which is beyond the scope of the pure cross section integration programs. The latter have the advantage that they are available in NLO, whilst in Monte Carlo programs the hard matrix element is usually calculated in LO which is then supplemented with parton showers to approximate higher order contributions. At NLO the different diagrams can produce negative contributions to the cross section which makes it difficult to implement NLO calculations in full hadron level Monte Carlo event generators. Since recently there exists also a Monte Carlo program which matches NLO matrix elements with parton showers (MC@NLO [37]) but this is presently only available for comparisons with  $pp/p\bar{p}$  data (cf. section 3.2.2).

### 2.5.1 NLO QCD Calculations

In the literature different approaches of perturbative QCD calculations of heavy quark production at NLO are presently discussed. They are motivated by the more complicated situation introduced by the hard scale set by the heavy quark mass in addition to the already present scales like  $p_t$  or  $Q^2$ . In the perturbative calculation the multi-scale problem is treated in different ways, according to the relative magnitudes of the present scales.

In the *massive* or *fixed-flavour-number-scheme* (FFNS) only light quarks and gluons are considered as active partons in the proton or the (resolved) photon. Heavy quarks are produced via the BGF mechanism. In the perturbative expansion of the hard scattering cross section all terms up to  $\mathcal{O}(\alpha_s^2)$  are taken into account. Some of the diagrams corresponding to heavy flavour production in the massive scheme are shown in figure 2.9. The massive approach is reliable near threshold, i.e. if the transverse momentum of the heavy quark is less than or similar to its mass. In the other case, if  $p_t \gg m_Q$ , the FFNS approach becomes unreliable because terms of the form  $\log^n(p_t^2/m_Q^2)$  occur at any order  $n$ , accounting for collinear gluon emissions and gluon or photon splittings into a heavy quark pair and spoiling the convergence of the perturbative expansion. The uncertainty of the 3-flavour calculation is expected to grow as  $p_t/m_Q$  becomes large.

In *resummed calculations*, the contributions from collinear gluon emissions and gluon or photon splittings are included also beyond  $\mathcal{O}(\alpha_s^2)$ . This is achieved by absorbing the singularities into fragmentation functions and pdfs. As this requires setting the quark masses to zero, this approach is called the *massless* or *zero-mass-variable-flavour-number-scheme* (ZM-VFNS). Thus in the ZM-VFNS heavy quarks are considered as active partons in the proton and the photon.

For the sake of a comparison of the best theory knowledge with the experimental data a combination of both schemes is desirable. FONLL [38] is a full massive NLO calculation matched with an additional resummation of the logarithms of  $(p_t^2/m_Q^2)$  with next-to-leading logarithmic accuracy (NLL). Thus, in this approach the calculation at low momenta is effectively performed in the massive scheme and at large momenta in the massless scheme. However, the theory predictions based on the FONLL approach are relatively new and presently unavailable for H1 data.

## FMNR

In the present analysis QCD calculations to next-to-leading order are performed using the program FMNR [39]. A comprehensive discussion of the FMNR program and its application to  $ep$  processes can be found in [40]. FMNR implements the calculation at fixed order in the massive scheme. The predictions from FMNR are therefore expected to be valid in the lower  $p_t$  region. FMNR provides weighted parton level events with two or three outgoing partons, i.e. the heavy quark antiquark pair and possibly a third parton. Values of 1.5 and 4.75 GeV are chosen for the  $c$  and  $b$  quark masses respectively. The renormalisation and factorisation scales are set to the transverse masses

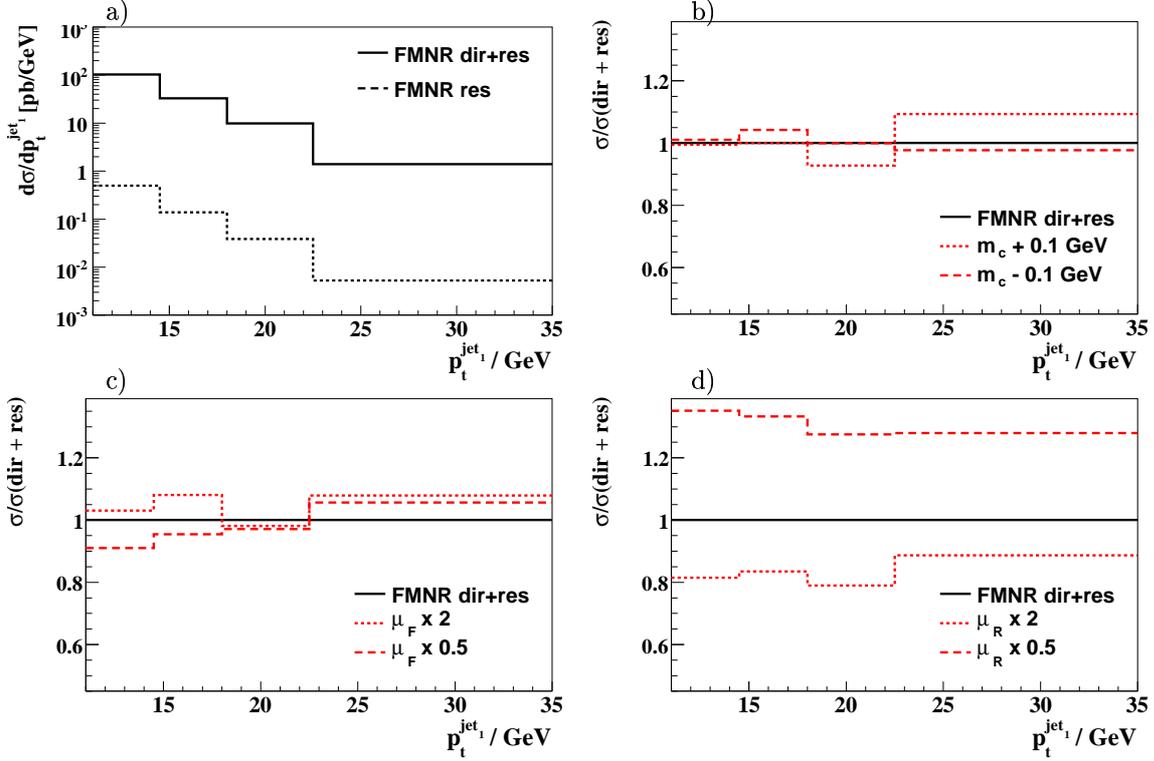
$$m_t = \sqrt{m_q^2 + p_{t,q\bar{q}}^2}, \quad (2.23)$$

where  $p_{t,q\bar{q}}^2$  is the average of the squared transverse momenta of the heavy quark and anti-quark<sup>4</sup>.

The *CTEQ5F3* parameterisation [21] is chosen for the parton distribution functions in the proton. Contributions from processes with resolved photons are calculated using the

---

<sup>4</sup>Recent H1 charm measurements at lower transverse momenta use  $\mu_F = 2m_t$  to ensure a large enough scale. In the transverse momentum range considered in this analysis the difference is small as figures 2.10 c) and 2.11 c) suggest.



**Figure 2.10:** Results of the NLO calculations using the FMNR program for **charm** quarks as a function of  $p_t^{\text{jet}1}$ . a) Differential charm dijet cross section and the relative effect of varying b) the charm quark mass  $m_c$ , c) the factorisation scale  $\mu_F$  and d) the renormalisation scale  $\mu_R$ .

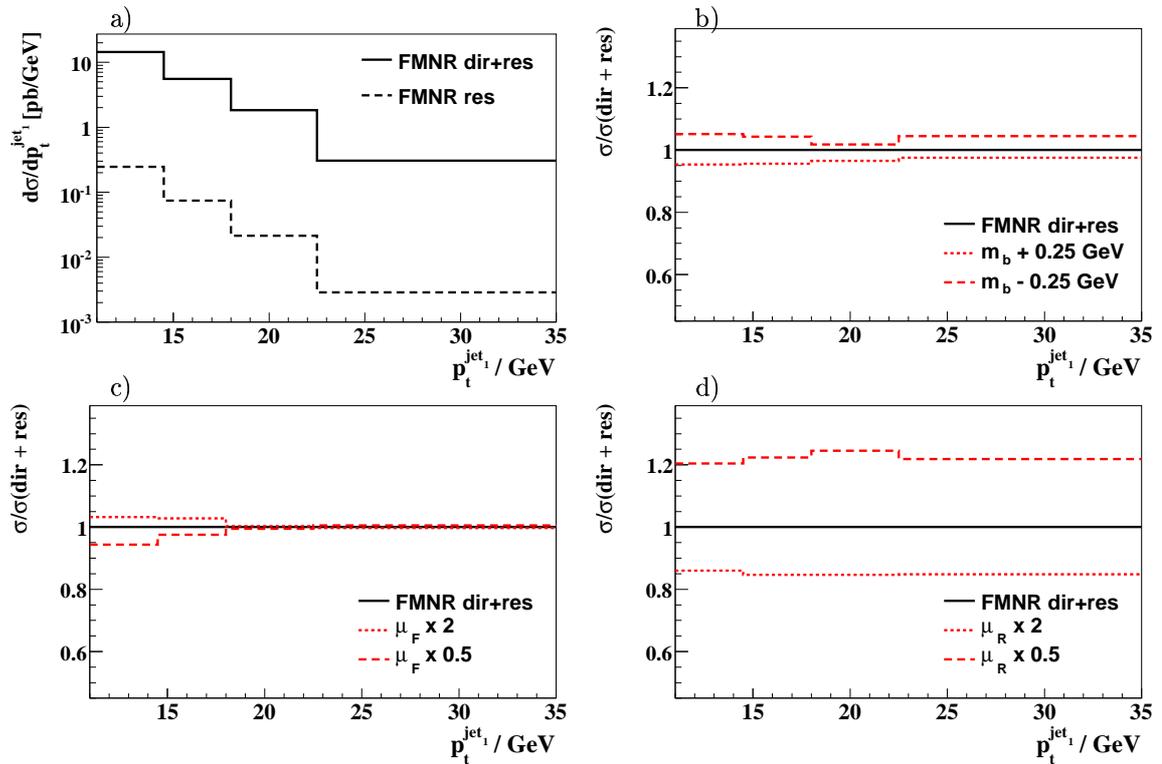
GRV-G  $HO$  distributions of partons in the photon [41]. Note, that for the resolved calculation in FMNR the photon density has to be convoluted with the photon flux from the electron for the specific range of the inelasticity  $y$  [42]<sup>5</sup>.

The results of the NLO calculation as a function of the transverse momentum of the leading jet<sup>6</sup> are shown in figures 2.10 and 2.11. The calculations are performed in the visible range of this measurement (cf. equation 5.7). For both charm and beauty the contributions from resolved processes in the NLO predictions are found to be small ( $< 3\%$ ). In addition the effect of the variations of the heavy quark masses and of the renormalisation and factorisation scales by factors of one half and two are visualised. The uncertainty applied to the heavy quark masses has only a minor effect on the cross section in the kinematic region of the measurement ( $p_t \gg m_b$ ) leading to cross section changes of maximal  $\pm 9\%$  for charm and  $\pm 4\%$  for beauty.

A larger factorisation scale  $\mu_F$  leads to increasing parton densities and therefore generally increases the cross section. The effect of this variation however is also seen to be small, i.e. of the same order of magnitude as the quark mass variations. The variation of the

<sup>5</sup>For the resolved calculation of e.g.  $d\sigma/dx_\gamma$  the information about  $y$  is needed and thus its calculation must be performed in  $\gamma p$  mode at a photon-proton centre-of-mass energy corresponding to a fixed value of  $y = W_{\gamma p}^2/s$ . The result is then multiplied with the integrated photon flux of a small range around the fixed  $y$ .

<sup>6</sup>The jet is reconstructed using the  $k_t$  jet algorithm (discussed in section 4.7) to the two or three outgoing partons.



**Figure 2.11:** Results of the NLO calculations using the FMNR program for **beauty** quarks as a function of  $p_t^{jet1}$ . a) Differential beauty dijet cross section and the relative effect of varying b) the beauty quark mass  $m_b$ , c) the factorisation scale  $\mu_F$  and d) the renormalisation scale  $\mu_R$ .

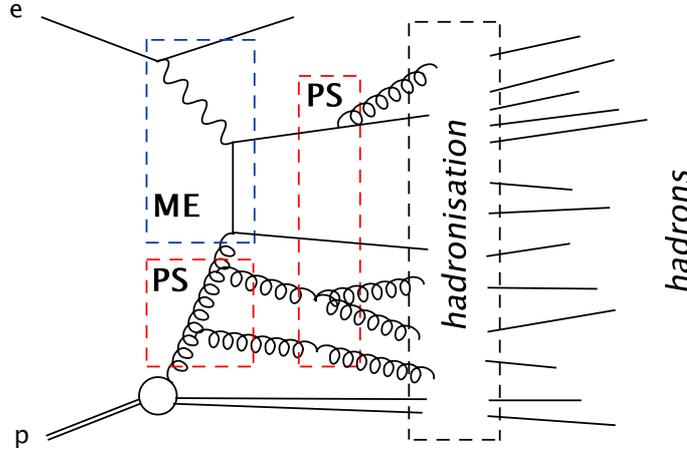
renormalisation scale  $\mu_R$  has the largest influence on the result. Differences of up to 30–35% for charm and 20–30% for beauty are observed in certain bins of the measurement. This indicates the need for higher order contributions. With a larger  $\mu_R$ ,  $\alpha_s(\mu_R)$  gets smaller, leading to a decreased cross section.

Theoretical uncertainties of the NLO calculation are estimated by independent variations of  $\mu_R$ ,  $\mu_F$  and the quark masses and the maximal changes to the cross section predictions are adopted as systematic errors. The cross section variations when using other proton structure functions such as CTEQ6M [43], MRSG or MRST1 [22] are less than 8% in all regions of the measurement. The latter uncertainty is added in quadrature to the uncertainties from the scales and the quark mass.

## 2.5.2 LO Monte Carlo Event Generators

Taking the BGF process as an example the elements of a Monte Carlo event generator are sketched in figure 2.12. The starting point is the hard partonic subprocess described by a matrix element (ME), which is calculated using perturbation theory. Initial and final state parton showers (PS) simulate QCD radiation down to a cut-off scale. After the cut-off the hadronisation process starts, in which the coloured partons coalesce into colourless hadrons. Within the Monte Carlo programs used in this analysis, the last two steps are implemented via the JETSET program [44].

Beauty and charm production events are simulated for this analysis using the PYTHIA



**Figure 2.12:** Schematic view of the subsequent steps of an event generator: matrix element (ME), parton shower (PS) and hadronisation.

[44] and CASCADE [45] Monte Carlo generators. The cross section predictions of PYTHIA version 6.1 and CASCADE version 1.2010 will be compared to the measurement. The two Monte Carlo generators differ significantly in the underlying QCD treatment. Whereas PYTHIA computes the on-shell matrix element and the initial state parton shower in collinear factorisation, i.e. using the DGLAP evolution equations, CASCADE contains an implementation of the CCFM evolution equation based on  $k_t$  factorisation using off-shell matrix elements (see also section 2.2.1).

PYTHIA is run in an inclusive mode (MSTP(14)=30) in which direct and resolved photon processes, including heavy quark excitation, are generated using massless matrix elements for all quark flavours. In the visible range of this measurement the cross section of the resolved process with a gluon from the photon is less than 2%. In contrast, the excitation processes contribute up to  $\approx 52\%$  and  $\approx 43\%$  to the visible photoproduction cross section for charm and beauty, respectively. The CTEQ5L [21] parton densities are used for the proton and those of SaS1D [46] for the photon. The charm and beauty quark masses are set to 1.5 and 4.75 GeV respectively, and the fragmentation is modelled by the Lund string model, using the Peterson function for the longitudinal fragmentation of beauty and charm quarks. In order to study the systematic uncertainty resulting from the chosen fragmentation model, additional PYTHIA Monte Carlo samples are generated using the Lund fragmentation also for heavy quarks.

Multiple interactions between the proton and the resolved photon are incorporated in PYTHIA by adding additional interactions between spectator partons within the same event. The perturbative parton-parton scattering must be extended to the low transverse momentum region to calculate these processes.

In CASCADE the process  $\gamma g \rightarrow c\bar{c}$  or  $b\bar{b}$  is implemented using off-shell matrix elements convoluted with  $k_t$  unintegrated parton distributions in the proton. In contrast to the standard parton density the unintegrated gluon density explicitly depends on the transverse momentum  $k_t$  of the parton. Although only the direct BGF process is taken into account, some effects of heavy quark excitation processes are indirectly included by this procedure. In this analysis the parametrisation A0 [47] is used for the parton distributions and the

	PYTHIA (LO)	CASCADE (LO)	FMNR (NLO)	
			central	variations
Proton pdf	CTEQ5L	A0	CTEQ5F3	CTEQ6M, MRSG
Photon pdf	SaS1D	-	GRV-G HO	
$m_c$ / GeV	1.5	1.5	1.5	1.4 - 1.6
$m_b$ / GeV	4.75	4.75	4.75	4.5 - 5.0
$\mu_R$	$m_t$	$4m_q^2 + p_t^2(q\bar{q})$	$m_t$	$(1/2 - 2)m_t$
$\mu_F$	$m_t$	$\hat{s} + Q_t^2$	$m_t$	$(1/2 - 2)m_t$
Peterson $\epsilon_c$	0.058	-		-
$\epsilon_b$	0.0069	-		-

**Table 2.1:** Overview of settings used in the QCD calculations. The heavy quark mass is denoted as  $m_q$ ,  $p_t^2(q\bar{q})$  is the average of the transverse momenta of the two heavy quarks,  $\hat{s}$  is the centre-of-mass energy squared and  $Q_t$  is the transverse momentum of the heavy quark.

same mass parameters as used in PYTHIA. The final state parton showers, the treatment of the proton remnant and the hadronisation are adopted from JETSET being part of the PYTHIA program. Multiple interactions are not treated in CASCADE.

An overview of the settings used in the QCD calculations is given in table 2.1.

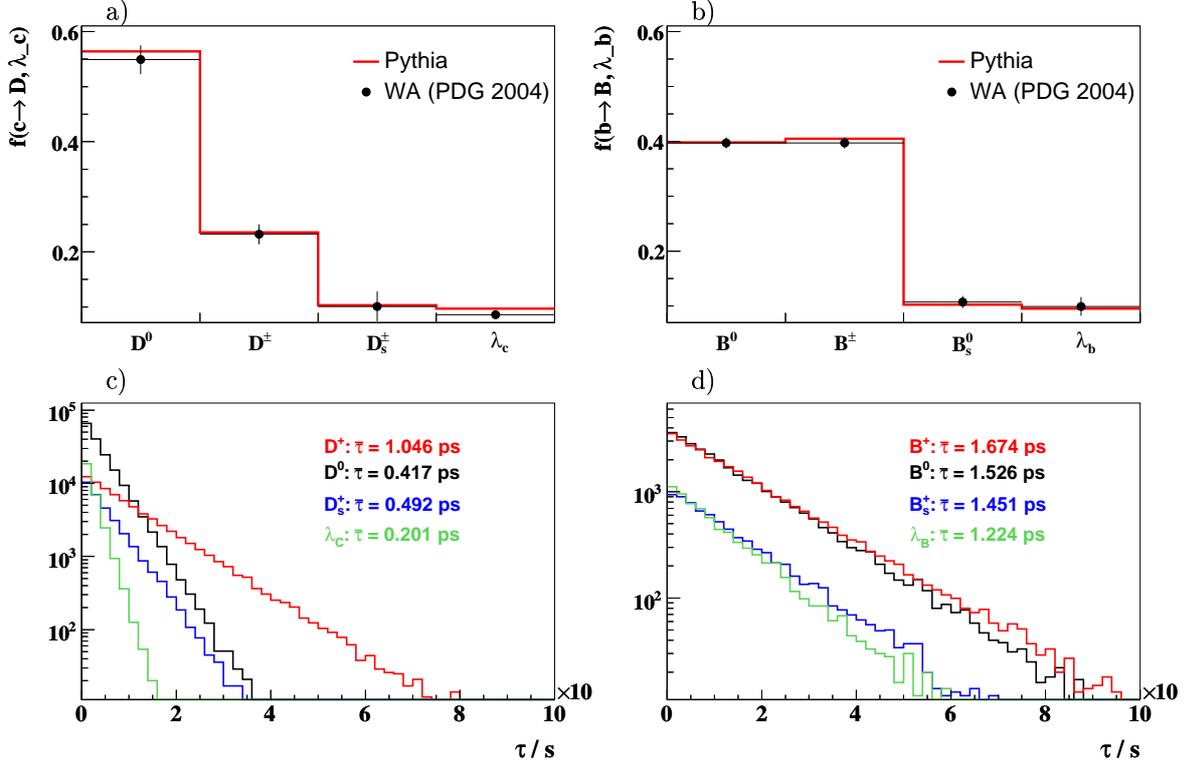
A DJANGO 6.2 [48] Monte Carlo sample is used in this analysis in order to estimate the NC DIS background for the dijet measurement. This sample covers the low  $Q^2$  region and is also exploited to study the trigger efficiency. In addition a high  $Q^2$  DJANGO Monte Carlo sample ( $Q^2 > 100 \text{ GeV}^2$ ) is used to cross check the hadronic calibration.

## 2.6 Properties of Heavy Hadrons

Due to the differences in quark masses, charm- and beauty-flavoured hadrons are considerably heavier than hadrons composed only of the light quarks  $u$ ,  $d$  and  $s$ . The light meson  $\pi^0$  with a mass of  $m_{\pi^0} = 0.135 \text{ GeV}$  has to be compared with e.g. the  $D^0$  charm meson with a mass of  $m_{D^0} = 1.865 \text{ GeV}$  and the  $B^0$  beauty meson with a mass of  $m_{B^0} = 5.28 \text{ GeV}$ .

	quark content	mass $m_Q$ / MeV	lifetime $\tau$ / ps	decay length $c\tau$ / $\mu\text{m}$	branching fraction $f[c(b) \rightarrow D(B), \Lambda_{c(b)}]$
$D^0$	$cu$	$1864.6 \pm 1.0$	$0.4103 \pm 0.0015$	123	$0.549 \pm 0.026$
$D^+$	$cd$	$1869.4 \pm 0.5$	$1.040 \pm 0.007$	312	$0.232 \pm 0.018$
$D_s^+$	$cs$	$1968.3 \pm 0.5$	$0.490 \pm 0.009$	147	$0.101 \pm 0.027$
$\Lambda_c^+$	$cud$	$2284.9 \pm 0.6$	$0.200 \pm 0.006$	60	0.086
$B^0$	$bd$	$5279.4 \pm 0.5$	$1.536 \pm 0.014$	460	$0.397 \pm 0.010$
$B^+$	$bu$	$5279.0 \pm 0.5$	$1.671 \pm 0.018$	501	$0.397 \pm 0.010$
$B_s^0$	$bs$	$5369.6 \pm 2.4$	$1.461 \pm 0.057$	438	$0.107 \pm 0.011$
$\Lambda_b^0$	$bdu$	$5624.0 \pm 9.0$	$1.229 \pm 0.080$	368	$0.099 \pm 0.017$

**Table 2.2:** Properties of D- and B-hadrons [49].



**Figure 2.13:** Branching ratios and lifetimes of charm and beauty hadrons. The properties have been calculated using the PYTHIA Monte Carlo simulation.

Charmed hadrons are the  $D^+$ ,  $D^0$  and  $D_s^+$  mesons and the  $\Lambda_c$  baryon. The corresponding beauty hadrons are  $B^+$ ,  $B^0$ ,  $B_s^+$  and  $\Lambda_b$ . The spin excited states are written as e.g.  $D^*$ . The second characteristic property of heavy hadrons is their long lifetime. While charm mesons have lifetimes of (0.4-1) ps, the lifetimes of beauty mesons are of the order of 1.6 ps. The masses, mean lifetimes and branching ratios of charm and beauty hadrons can not be predicted accurately from theory and are therefore measured experimentally (table 2.2).

The different lifetimes of the weak decays are driven by the different sizes of the Cabbibo-Kobayashi-Maskawa (CKM) matrix elements  $|V_{ij}|$  corresponding to the dominant decays  $b \rightarrow cW^-$  and  $c \rightarrow sW^+$  ( $|V_{cb}| = 0.0412 \pm 0.0020$ ,  $|V_{cs}| = 0.224 \pm 0.016$ ).

The decay length  $\ell$  is related to the lifetime  $c\tau$  via

$$\ell = \beta\gamma \cdot c\tau = \frac{p_Q}{m_Q} \cdot c\tau, \quad (2.24)$$

where  $p_Q$ ,  $m_Q$  and  $\beta\gamma$  are the heavy particle's momentum, mass and boost factor respectively. The probability for an unstable particle to decay at some distance  $L$  between the  $Q\bar{Q}$  production vertex and the point of the subsequent hadron's decay is given by

$$\mathcal{P}(L) = 1 - \exp\left(-\frac{L}{\ell}\right). \quad (2.25)$$

Taking into account the production branching fractions  $f(c \rightarrow C, \Lambda_c)$  and  $f(b \rightarrow B, \Lambda_b)$  the mean weighted decay lengths of charm and beauty hadrons is  $155 \mu\text{m}$  and  $466 \mu\text{m}$

respectively. Since the average boost  $\beta\gamma$  at HERA is about 0.7 at low  $Q^2$ , the average observed decay length is only  $108\ \mu\text{m}$  for charm and  $326\ \mu\text{m}$  for beauty.

In consequence, a lifetime based tag of  $c$ - or  $b$ -flavoured hadrons and their separation from the light quark background should be feasible if the production and decay vertices are known with a precision  $\sim 100\ \mu\text{m}$ .

Figure 2.13 shows the branching ratios and lifetimes of charm and beauty hadrons as obtained with the PYTHIA Monte Carlo program. The simulated branching ratios of the Monte Carlo program are in good agreement with the world average values taken from [49]. Similarly, the calculated mean lifetimes computed from the exponential lifetime spectra reproduce nicely the world average values.

## CHAPTER 3

# THE EXPERIMENT

---

The analysis presented in this thesis is based on positron-proton ( $ep$ ) collision data recorded by the H1 experiment. H1 is one of the general purpose detectors which study such  $ep$  collisions at the *Hadron Electron Ring Accelerator* (HERA) at the *Deutsches Elektronen Synchrotron* laboratory (DESY), located in Hamburg, Germany.

The H1 detector is described in detail in [50]. In the first part of this chapter only its components of main interest for the present measurement are introduced. In the second part the experimental status of heavy quark production at HERA and other colliders is briefly reviewed.

### 3.1 The H1 Detector at HERA

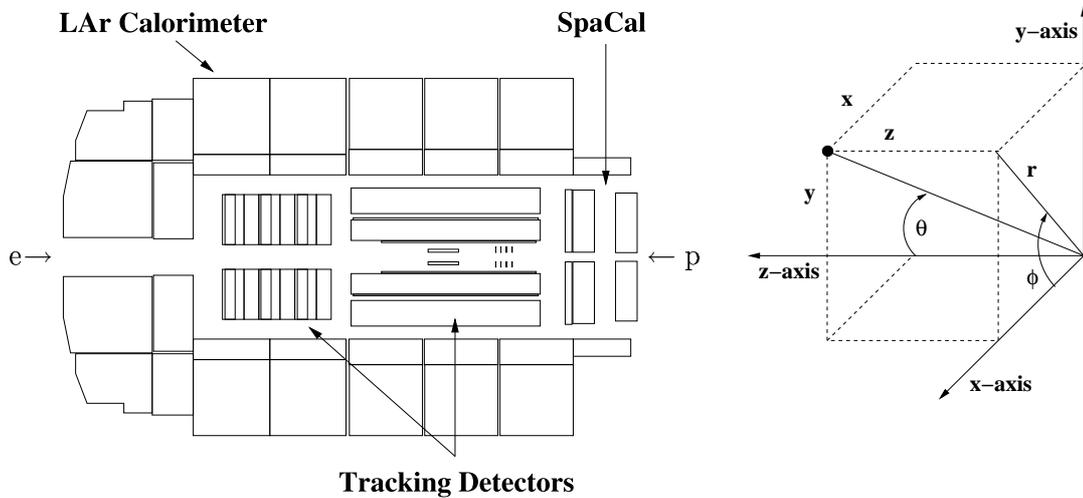
At HERA positrons (or electrons) collide with protons at energies of 27.6 GeV and 920 GeV respectively, yielding a centre-of-mass energy of  $\sqrt{s} \approx 319$  GeV. The particle beams are stored in bunches with a time distance of 96 ns, corresponding to a 10.4 MHz collision rate.

Note, that in a lively and creative high energy physics experiment the detector can undergo several steps of modifications with respect to its original design. In the following the discussion of the detector components refers to its status in the years 1999-2000.

The H1 detector measures roughly 12m x 10m x 15m and weighs approximately 2800 tons. A schematic side-view of the H1 detector is given in figure 3.1. The asymmetric arrangement of the components along the beam axis is due to the large difference of the HERA beam energies resulting in a large boost of the final state particles along the direction of the incoming proton. Charged particles emerging from the interaction region are measured by the tracking detectors. The calorimeters surround the tracking chambers and measure energy deposits of charged and neutral particles. The instrumented iron return yoke of the solenoidal magnet encloses the calorimeters.

#### Coordinate System

Points within the detector are usually defined using the Cartesian coordinates  $(x, y, z)$  with respect to the origin, i.e. the nominal interaction point. The direction of the proton beam, also referred to as the forward direction, defines the  $z$ -axis, the  $y$  direction is vertically



**Figure 3.1:** Schematic longitudinal section of the H1 detector. Only the most important components used in this analysis are shown. The H1 coordinate system is shown on the right.

upwards and  $x$  points horizontally to the centre of the storage rings. Angles are described with a spherical coordinate system  $(r, \theta, \phi)$ , such that the azimuthal angle  $\phi$  lies in the  $xy$ -plane and the polar angle  $\theta$  is measured from the  $z$ -axis as depicted in figure 3.1.

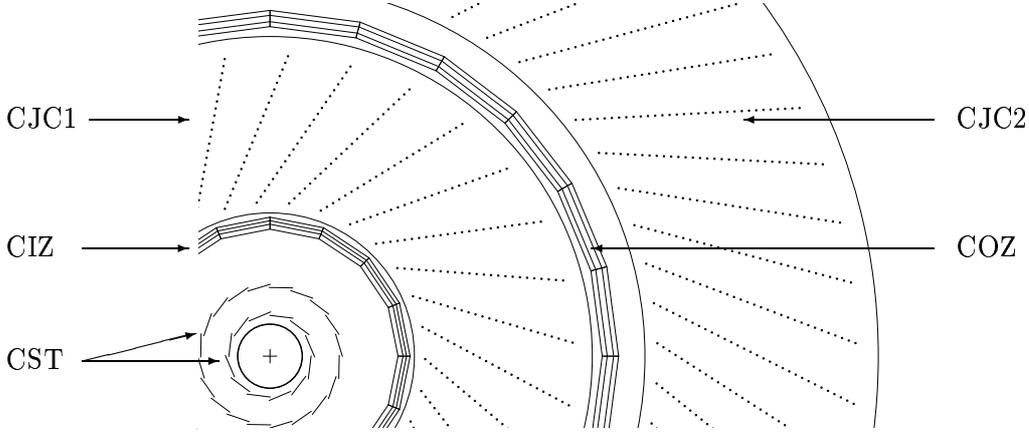
## Luminosity

The accurate determination of the luminosity is essential for accurate cross section measurements. The luminosity is measured using the Bethe-Heitler bremsstrahlung process ( $ep \rightarrow ep\gamma$ ), for which the cross section is precisely known from QED. The scattered positron as well as the outgoing photon are measured with special detectors situated in the backward direction: the Electron Tagger at  $z = -33.4$  m and the Photon Detector (PD) at  $z = -103$  m respectively. The integrated luminosities are measured with a precision of better than 1.5%.

### 3.1.1 Central Tracking Detectors

Charged particles emerging from the interaction region are measured by the central tracking detector (CTD) in the pseudo-rapidity range  $-1.74 < \eta < 1.74$ <sup>1</sup>. The CTD consists of two large cylindrical central jet drift chambers (CJCs), two  $z$ -chambers (CIZ/COZ) and two multi-wire proportional chambers (CIP/COP) arranged concentrically around the beam-line within a solenoidal magnetic field of 1.15 T. A radial view of the CTD without CIP and COP is shown in figure 3.2. The CTD provides triggering information based on track segments in the  $r\phi$  plane from the CJC and the  $z$ -position of the vertex from the multi-wire proportional chambers.

<sup>1</sup>The pseudo-rapidity  $\eta$  of an object with polar angle  $\theta$  is given by  $\eta = -\ln \tan(\theta/2)$ .



**Figure 3.2:** Radial view of the central tracking system (section).

### Central Jet Chamber

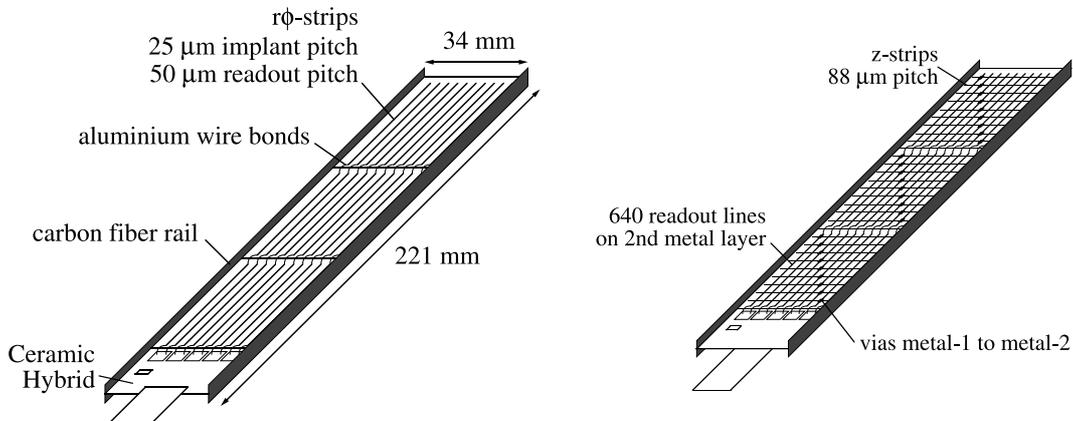
The CJCs are the backbone of the track reconstruction in the central region. Two cylindrical, coaxial volumes (CJC1 and CJC2), with sense wires strung parallel to the  $z$ -axis, measure charged particles with a single hit resolution in the  $r\phi$ -plane of  $\sigma_{r\phi}^{CJC} \approx 140 \mu\text{m}$  and a transverse momentum resolution of  $\sigma^{CJC}(p_t)/p_t = 0.01 \cdot p_t/\text{GeV}$ . The intrinsic hit resolution in  $z$  is significantly worse ( $\geq 22 \text{ mm}$ ) than in the transverse plane. An improved  $z$  measurement is provided by including information from the two thin drift chambers (CIZ and COZ). More information on the CJC, including details of the track reconstruction procedure can be found e.g. in [51].

### Central Silicon Tracker

Tracks reconstructed with the CJC are linked with hits in the Central Silicon Tracking detector (CST) [52], which consists of two cylindrical layers of silicon strip sensors, surrounding the beam pipe at radii of  $R = 57.5 \text{ mm}$  and  $R = 97 \text{ mm}$  from the beam axis. Its active length of  $356 \text{ mm}$  in the  $z$ -direction covers polar angles in the range  $30^\circ < \theta < 150^\circ$ .

The two layers are composed of 12 (inner layer) and 20 (outer layer) slightly overlapping *ladders*, each containing two identical *half ladders* of  $221 \text{ mm}$  length and  $34 \text{ mm}$  width. Each half ladder consists of three silicon sensors and a hybrid structure carrying the readout electronics (see figure 3.3).

The silicon sensors are double-sided having on the outer face ( $p$ -side) 1280  $p^+$ -acceptor strip implants positioned parallel to the  $z$ -axis with a pitch of  $25 \mu\text{m}$ . Only every second strip is readout, providing a single hit resolution in the  $r\phi$ -plane of  $\sigma_{r\phi}^{CST} = 12 \mu\text{m}$ . On the opposite side ( $n$ -side)  $n^+$ -donator strip implants of  $88 \mu\text{m}$  pitch are oriented perpendicular to the  $z$ -axis. Every  $n$ -side strip is read out determining the  $z$  coordinate with a single hit resolution of  $\sigma_z^{CST} = 25 \mu\text{m}$ .



**Figure 3.3:** CST half ladders:  $p$ -side (left) and  $n$ -side (right).

### Hit Finding

The hit finding algorithm is described in [53]. Charged particles crossing the CST produce pairs of positive and negative charge carriers leading to signals which are read out by the silicon strips. In a first step neighboring strips with a signal above noise threshold define a *cluster*. The centre-of-gravity of the cluster then determines the hit position. This is done independently for the  $p$ - and  $n$ -side. In a second step the association of  $p$ - and  $n$ -side clusters results in three dimensional *space points*.

The total signal-to-noise ratio of a cluster must exceed five (four) on the  $p$ -side ( $n$ -side). On the  $n$ -side an additional metal layer is needed for the readout which deteriorates the signal-to-noise ratio by a factor of two compared to the  $p$ -side (cf. figure 3.4).

The average hit efficiencies are 97 (92)% in  $r\phi$  ( $z$ ).

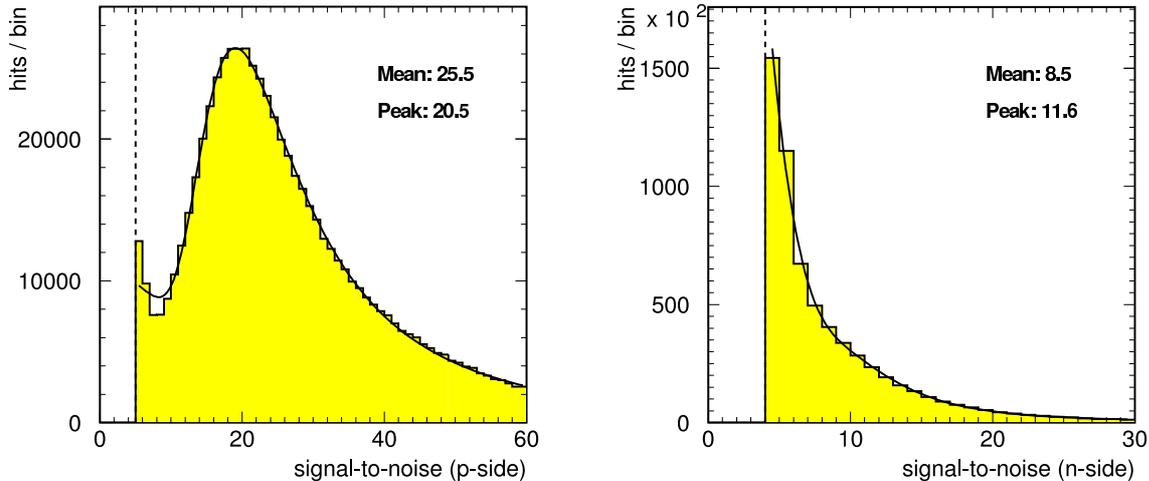
More details on the CST detector design, readout and reconstruction can be found in [52, 53].

### 3.1.2 Calorimetry

The H1 detector contains four separate calorimetric units with different angular acceptances. Important for this analysis are the Liquid Argon (LAr) Calorimeter and the Spaghetti Calorimeter (SpaCal).

#### The Liquid Argon Calorimeter

The LAr calorimeter surrounds the tracking chambers and covers the range  $-1.5 < \eta < 3.4$ . It consists of two sections, both contained in a single liquid argon cryostat. The inner electromagnetic (EM) region with high granularity and lead absorber plates is optimised for the detection of electrons and photons. In the outer part, the hadronic (HAD) section, the cells are larger and the absorber material is stainless steel. It is designed for the measurements of hadronic energy deposits. The LAr calorimeter is a non-compensating calorimeter, i.e. its response to electrons and hadrons is not equal: the average energy deposit of hadrons is  $\approx 30\%$  smaller than in the case of electrons with the same initial



**Figure 3.4:** Signal-to-noise ratio for the  $p$ -side (left) and  $n$ -side (right). The dashed line marks the cut made in the hit finding algorithm.

energy. In the energy reconstruction this behaviour is considered by a suitable signal weighting procedure. The achieved energy resolution of the LAr calorimeter for electrons is  $\sigma_{EM}(E)/E \approx 0.12/\sqrt{E/\text{GeV}} \oplus 0.01$  and for pions  $\sigma_{HAD}(E)/E \approx 0.5/\sqrt{E/\text{GeV}} \oplus 0.02$ . More information about the LAr calorimeter can be found in [54].

### The Spaghetti Calorimeter

The backward region ( $-4.0 < \eta < -1.4$ ) is covered by a lead-scintillating fibre calorimeter (SpaCal [55]), also with electromagnetic and hadronic sections. Energies are measured with  $\sigma_{EM}(E)/E \approx 0.08/\sqrt{E/\text{GeV}} \oplus 0.01$  and  $\sigma_{HAD}(E)/E \approx 0.3/\sqrt{E/\text{GeV}} \oplus 0.07$  respectively.

The electromagnetic section of the LAr calorimeter and the SpaCal is used in this analysis as a veto against DIS events. As described in section 5.2 events are selected with no scattered electron reconstructed in the LAr calorimeter or SpaCal. Thus the upper bound on  $Q^2$  is defined by the lower acceptance limit ( $\theta < 177.8^\circ$ ) of the SpaCal to be  $Q^2 < 1 \text{ GeV}^2$ .

### 3.1.3 H1 Trigger System

The main objective of the trigger system is to give a fast decision for the acquisition of interesting  $ep$  events while sorting out background events. Background sources are mainly due to reactions of the protons, e.g. interaction of the proton beam with gas in the beam pipe (beam-gas) or with material of the beam pipe (beam-wall) or beam halo muons and muons emerging from cosmic radiation. Typical background rates are of the order of 1 kHz, while the electron and proton bunches collide at the interaction region with a rate of 10.4 MHz. With the help of the multi-level H1 Trigger System this input rate is reduced to about 10 Hz that then is written to tape for the offline analysis.

The H1 Trigger System is divided into four levels (L1-L4) of which only L1-L2 and L4 are operational in the data taken period considered here. L1 and L2 are phase locked to the HERA accelerator clock of 10.4 MHz.

The L1 trigger system provides a trigger decision after  $2 \mu\text{s}$  without causing dead time.

The trigger decision is based on information provided by various subsystems (*trigger elements*, TEs). The LAr calorimeter provides TEs which are used to trigger NC, CC and photoproduction jet events. The signals are derived from the total energy measured in the LAr calorimeter or from summed energies in subsections (Big-Towers) of it. A full description of the LAr trigger can be found in [56, 57].

The central trigger logic combines several TEs to 128 so-called *subtriggers*. An event is kept at L1 if at least one of the subtriggers gives a positive decision. While the trigger decision is made, the readout is stored in pipelines. If a specific subtrigger has a too large rate, it is *prescaled*, i.e. scaled down by a factor  $n$ . Thus only every  $n$ -th positive decision of this subtrigger is taken into account, effectively scaling down the integrated luminosity collected by this subtrigger. The L1 output rate is of the order of a few kHz.

The L2 system decision is derived within  $20\mu\text{s}$  from one of two independent hardware systems, a topological trigger (L2TT) and a neural network (L2NN). A positive decision on L2 stops the pipelines and the full event is read out. This might cause some dead time in which no data can be collected. The L2 output rate is of the order of 50 kHz.

The L4 trigger is based on a full but simplified reconstruction of the event and decides in approximately 100 ms. The selected events are grouped into L4 classes. The physics event classes are classified by the presence of a hard scale, e.g. a sufficiently large  $Q^2$  or transverse momentum or by specific final state finders like e.g. 'open charm' or 'open beauty'. Events which cannot be assigned to a hard scale physics class are downscaled (L4 weight) and collected in a special class, called 'soft physics'.

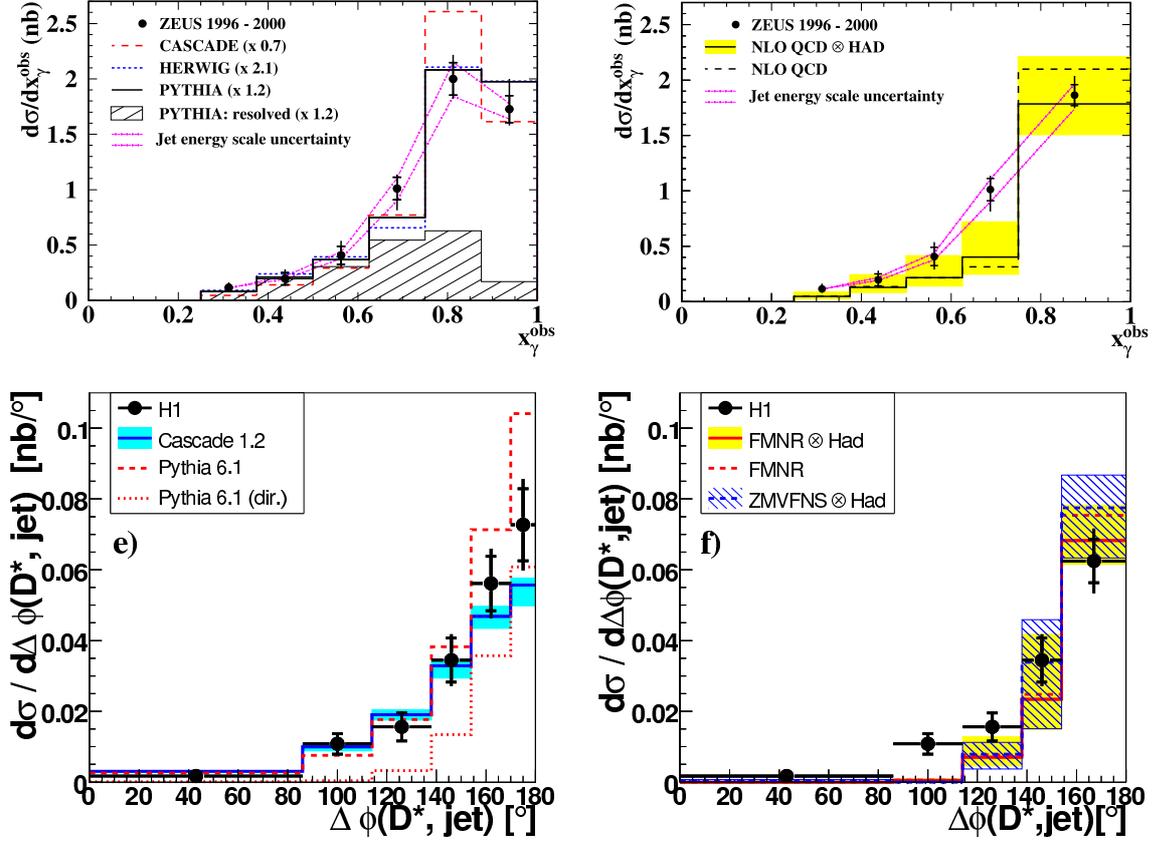
The events accepted on L4 are then fully reconstructed and written in a compressed format to data summary tapes (DST).

### 3.1.4 Detector Simulation

As discussed in section 2.5 the theoretical models to be tested by the experiment are often implemented in Monte Carlo event generators, which provide event-by-event topologies as four-vectors of the final state particles. For these particles the complete detector response is simulated by the standard H1 detector simulation which is based on GEANT 3 [58]. The trigger response is also simulated. The simulated four-vectors are mapped on detector signals, which then undergo the same reconstruction procedure as the measured data. In this way the data and the Monte Carlo simulation can directly be compared. If the physics model is correct and the data are well described by the reconstructed quantities of the simulation, efficiencies and detector acceptances can be extracted from the Monte Carlo simulation.

## 3.2 Experimental Review

Heavy quark production has been studied intensively since the discovery of the charm quark in 1974 [59] and the beauty quark in 1977 [60]. In the following sections some of the most interesting results obtained by the HERA experiments H1 and ZEUS as well as from other colliders like the TEVATRON will be summarized. In the last section the impact of these measurements is discussed.



**Figure 3.5:** Charm dijet measurements at HERA. *Top:* Differential  $D^*$  dijet photoproduction cross section as a function of  $x_\gamma^{obs}$  from ZEUS [9]. *Bottom:* Differential charm cross section as a function of  $\Delta\phi$  in a  $D^*$ -jet sample from H1 [16].

### 3.2.1 Heavy Quark Production at HERA

A complete and very nicely readable overview of the measurements of heavy quark production at HERA can be found in [61, 62]. Here, only the measurements of dijet photoproduction are outlined.

#### Charm Photoproduction

In the measurement presented in the ZEUS publication of charm dijets in photoproduction [9] at least two jets are required with transverse energy  $E_t > 5$  GeV in the pseudorapidity range  $|\eta^{jet}| < 2.4$ . Charm events are tagged in this measurement by the presence of  $D^*$  mesons. In figure 3.5 the measured cross section as a function of the variable  $x_\gamma^{obs}$  is shown and compared with various LO Monte Carlo predictions (left) as well as a fixed order massive calculation in NLO QCD (right). The variable  $x_\gamma^{obs}$  is a measure of the contributions from direct and resolved processes (for a discussion of dijet observables see section 4.8). A significant part of the cross section is situated at low values of  $x_\gamma^{obs}$ , indicating the presence of resolved charm contributions.

The recent H1 analysis [16] requires a reconstructed  $D^*$  and in addition the presence of at least one jet in the final state, which is not associated to the  $D^*$ . This *other* jet is initiated

by another hard parton which can be either the other charm quark, a gluon or a light quark. The investigation of correlations of the tagged charm jet with the *other* jet gives access to the underlying charm production process. In a similar analysis [63] the ZEUS experiment uses a dijet sample to measure the difference  $\Delta\phi$  of the two jets. In LO the quantity  $\Delta\phi$  is expected to have a sharp peak at  $180^\circ$  while higher order processes as well as fragmentation and detector resolution effects give rise to contributions at smaller values of  $\Delta\phi$ . In figure 3.5 the measurement from H1 is shown and compared to LO Monte Carlo programs (left) and NLO calculations in the massive and massless scheme (right). While the description by PYTHIA and CASCADE is good, the NLO calculations fail at lower  $\Delta\phi$ , indicating the need for higher order contributions or different hadronisation effects.

### Beauty Photoproduction

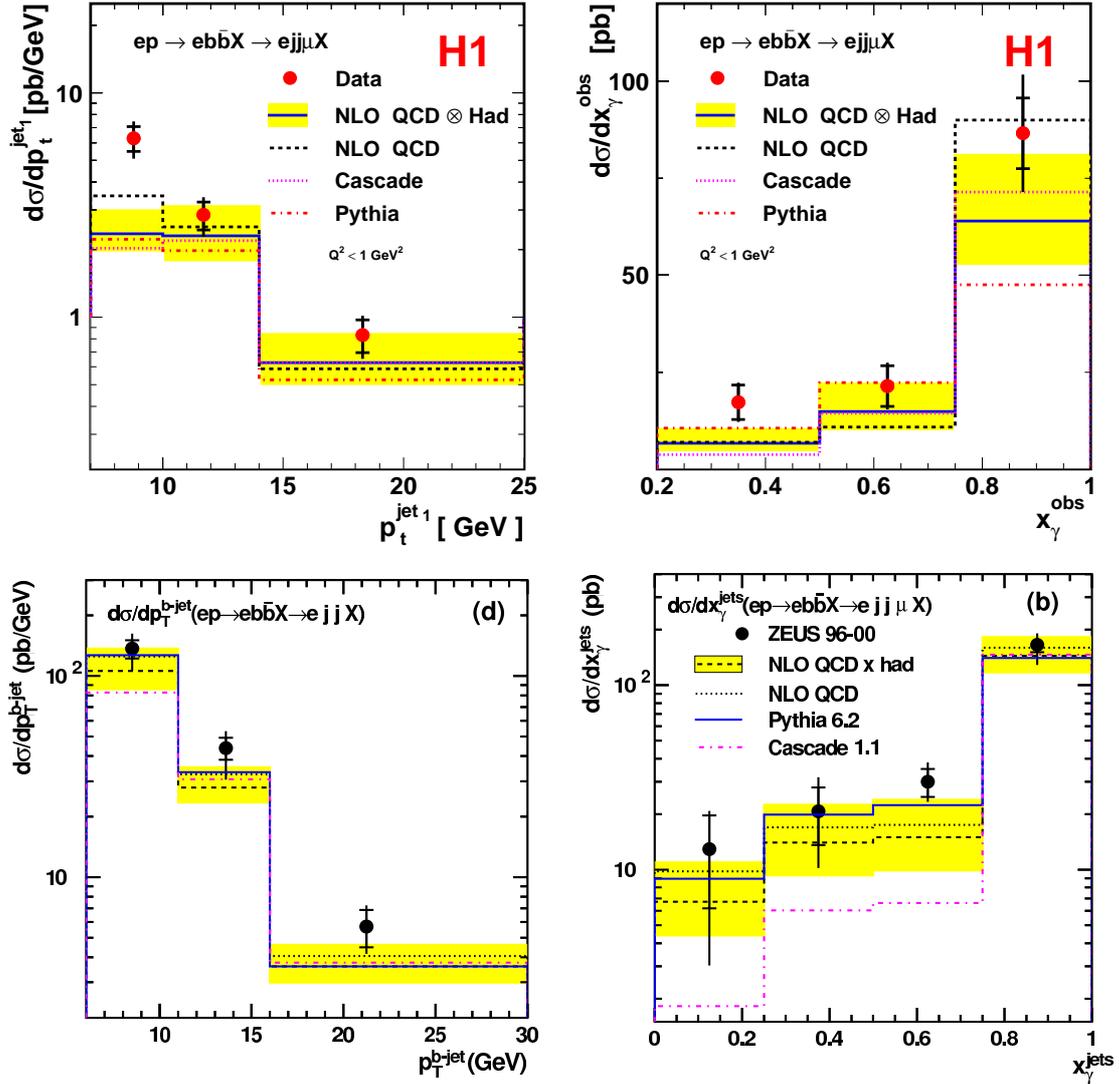
In contrast to the charm case no explicit reconstruction of hadronic resonances containing beauty is made at HERA due to the comparably small beauty production rates and the little branching ratios for suitable decay channels. The measurements of beauty cross sections performed at HERA usually make use of semi-leptonic decays of beauty hadrons. The beauty signal can be enriched by selecting events containing high transverse momentum muons or electrons. The techniques used to distinguish the beauty signal from the charm and light quark background are based on observables which exploit either the large mass ( $p_t^{rel}$ ) or the long lifetime (*impact parameter*) of the beauty quarks or a combination of both.

The first measurements of beauty production cross sections at HERA [11, 12] were higher than calculations in NLO QCD.

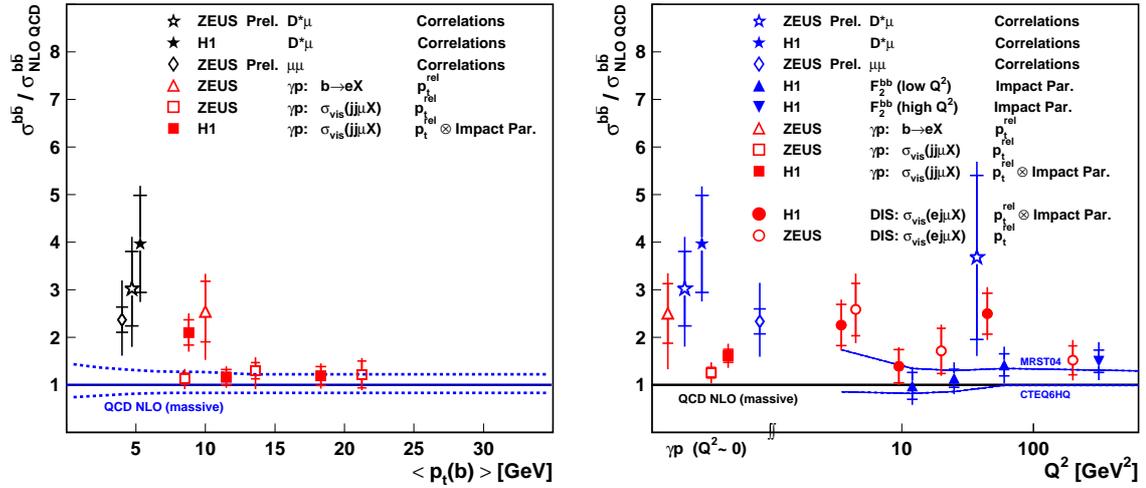
A recent H1 measurement [14] requires at least two jets with a transverse momentum  $p_t > 7(6)$  GeV in the pseudo-rapidity range  $|\eta^{jet}| < 2.5$  accompanied by a muon associated to one of the two jets. A similar analysis has been performed by ZEUS in a comparable phase space region [13]. In figure 3.6 the results of both measurements as a function of  $p_t^{jet}$  and of  $x_\gamma^{obs}$  are presented. The data is compared to predictions from PYTHIA, CASCADE and a fixed order massive NLO calculation (FMNR).

In both measurements the data show reasonable agreement with the NLO prediction with the data lying on the upper bound of the theory uncertainty band, except for the lowest  $p_t^{jet}$  bin of the H1 measurement, where the data is significantly above NLO. The LO Monte Carlo models provide in general a reasonable description, except for the lower  $x_\gamma^{obs}$  region where CASCADE tends to underestimate the data.

There are plenty of other results of beauty production at HERA available which can not be covered here, such as measurements based on  $\mu\mu$  correlations or analyses in the DIS regime. Figure 3.7 summarises the wealth of results achieved at HERA in recent years. In this figure the ratios between the measured cross sections and the predictions from the fixed order massive NLO prediction is shown as a function of the mean transverse momentum of the  $b$ -quark and as a function of  $Q^2$ . Taking into account both, the theoretical and experimental errors, all measurements are in reasonable agreement with the theory. However, in most cases the central values of the measurements tend to be somewhat higher than NLO QCD. Note that the phase space of the various measurements can differ drastically.



**Figure 3.6:** Differential beauty dijet photoproduction cross section as a function of  $p_t^{jet}$  and  $x_\gamma^{obs}$  from H1 [14] (top) and ZEUS [13] (bottom).



**Figure 3.7:** Summary of beauty measurements at HERA. The left plot shows the ratio of the cross sections to the NLO predictions as a function of the mean transverse momentum of the  $b$ -quark. On the right the dependence on  $Q^2$  is shown.

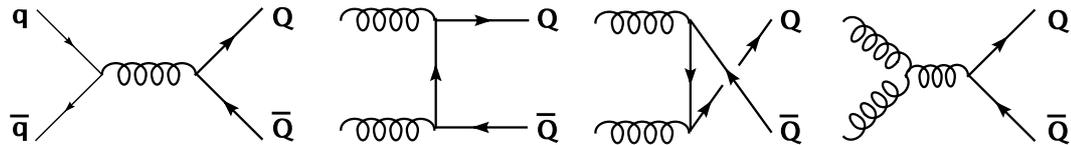
### 3.2.2 Heavy Quark Production at other colliders

It is also interesting to compare the HERA heavy quark production results with that of other colliders.

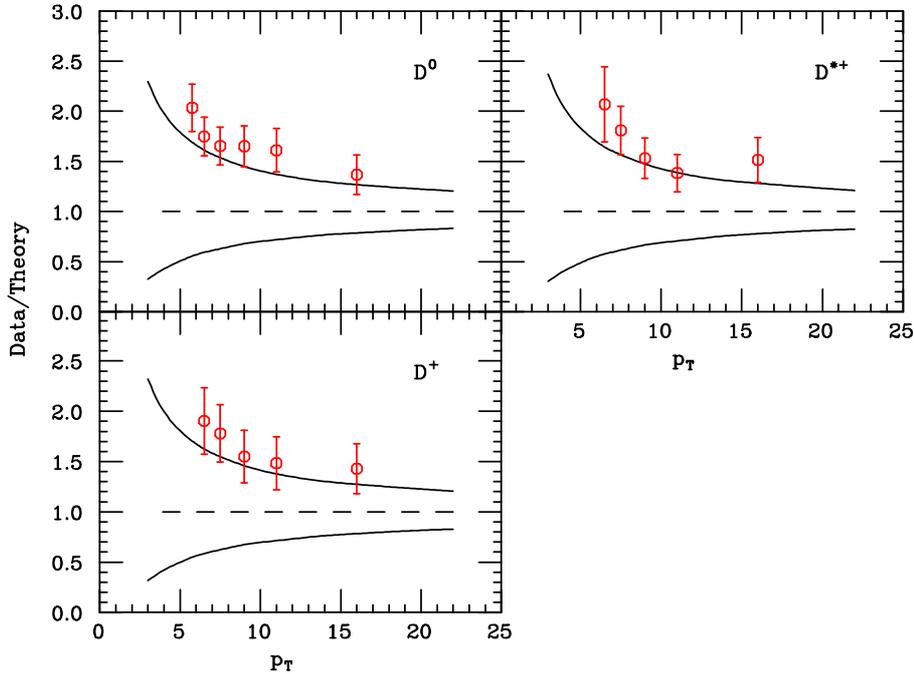
At fixed-target experiments most of the available data concerns charm quark production, since the typical centre-of-mass energy reaches at maximum 40 GeV, where the beauty cross section is quite small. In [64] the measurements of the total cross section of  $c\bar{c}$  pairs in  $\gamma N$  collisions collected by various fixed target experiments are reviewed. The experimental results are compared to NLO predictions and are found to be consistent with a charm quark mass of 1.5 GeV. However, the statistical accuracy of the data as well as the precision of the theory calculations are limited so that no conclusive picture could be found.

At the TEVATRON the experiments D0 and CDF can measure heavy quark production in  $p\bar{p}$  collisions at a centre-of-mass energy of 1.8 TeV (1.96 TeV) in Run I (Run II), respectively. In  $p\bar{p}$  collisions charm and beauty quark production is dominated by the processes shown in figure 3.8.

The results from a recent CDF measurement [65] of charm quark production using D-mesons is presented in figure 3.9. The measured ratios between the experimental results and the theoretical predictions from FONLL as a function of the transverse momentum of the D-meson show that the overall agreement is reasonable, with the data lying on the upper limit of the theoretical uncertainty band. The theory predictions based on the improved



**Figure 3.8:** Leading order processes for charm and beauty quark production in  $p\bar{p}$  collisions.



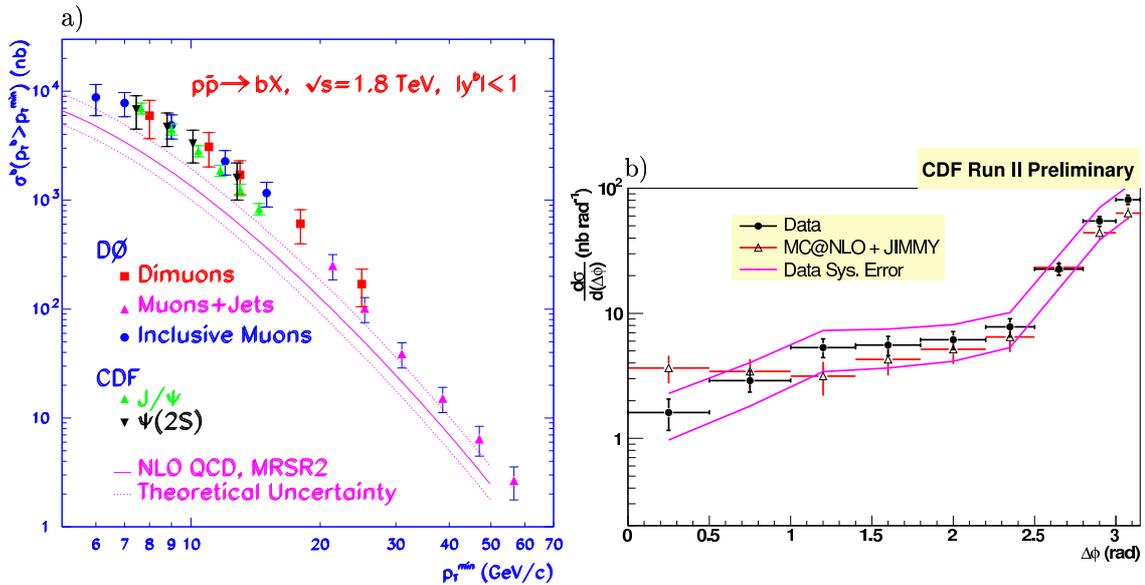
**Figure 3.9:** CDF Run II results for charm quark production as a function of the transverse momentum of the D-meson. Shown are ratios between the experimental results and the central theoretical predictions obtained with the FONLL approach.

FONLL approach are relatively new and presently not available for comparisons with the H1 data (cf. section 2.5.1).

The beauty quark production as measured at the TEVATRON in Run I reveals a widely discussed significant excess of data over NLO predictions, as for example illustrated in figure 3.10 a).

In a recent article [66] all available beauty measurements at the TEVATRON are reviewed and compared to perturbative QCD predictions (NLO and FONLL). For the single  $b$ -quark production cross section the derived average ratio of the data to the standard theory (NLO) is  $2.39 \pm 0.17$ . The standard NLO prediction is based on old (pre-HERA) parton distribution functions (MRSD and MRSA). The improved QCD calculation (FONLL) is approximately 60% higher than the standard NLO prediction, yielding a ratio of 1.5 with a relatively large uncertainty of the FONLL prediction of  $\sim 40\%$ . The increase of the FONLL prediction with respect to the standard NLO calculation is mostly due to the usage of a harder fragmentation function which is based on LEP and SLC data instead of the Peterson fragmentation function which was used for the standard theory. The parton level cross sections are the same in both predictions. In addition the latest parton distribution functions, including HERA data, have been considered leading to an increase of the prediction of 20% alone. In conclusion, the average single  $b$  cross section measured at the TEVATRON is within the range of values predicted by the FONLL calculation.

A corresponding ratio is derived for the  $b\bar{b}$  correlation, i.e. the production of both  $b$  and  $\bar{b}$  quark centrally and above a given  $p_t$  threshold. As stated in [66] the status of the  $\sigma_{b\bar{b}}$  measurements at the TEVATRON is quite disconcerting. On average they find a ratio



**Figure 3.10:** TEVATRON Run I and Run II results for beauty quark production. a) A collection of D0 and CDF beauty measurements. b) A recent  $b\bar{b}$  jets measurement performed by CDF. The data is compared to the prediction from the MC@NLO generator.

data over improved theory of 1.8 with a 0.8 RMS deviation. Such a high RMS deviation indicates that the experimental results are inconsistent among themselves.

A recent measurement of CDF [67], studying events containing two high  $p_t$  central jets, where beauty jets are identified by requiring the presence of displaced secondary vertices, finds a ratio of data over theory of  $1.0 \pm 32\%$ . In this case however, the NLO prediction has been evaluated using the MC@NLO Monte Carlo generator [37]. As an example the results as a function of the azimuthal difference of both jets is shown in figure 3.10 b).

For some time it has been speculated that new physics beyond the Standard Model may be needed in order to explain the measured excess mainly found in the beauty production cross section [68]. Given the fact that both, charm and beauty cross sections are found to be in agreement at least with the upper limit of the improved NLO predictions, indicates that the remaining discrepancies may be due to uncalculated QCD effects, such as higher order contributions, or insufficient tuning of scales and/or quark masses.

### Prospects for the LHC

At the LHC complicated final states like charm and beauty photoproduction accompanied by jets will be very copious. Although interesting for its own sake, a deep understanding of heavy quark production is therefore needed in particular for the LHC since heavy quark events are background to signals from new physics. As an example the ATLAS collaboration studied the production of a  $b\bar{b}$  pair along with a supersymmetric Higgs particle which subsequently decays to a  $b\bar{b}$  pair [69]. Even requiring four tagged beauty jets a signal-to-background ratio of only a few percent would be achieved for this channel. The main background arises from QCD processes with a gluon splitting to a  $b\bar{b}$  pair. Thus, the verification of theory with HERA data will aid the understanding of these high rate QCD background events.

### Summary

Whereas the charm data seem to be already well understood, there is still room for improvement in the QCD description of the beauty data. However, the situation for the description of beauty production has recently changed significantly. Generally QCD provides an acceptable description of the data with some hints at differences in specific regions. Both HERA experiments will produce several new measurements in the next years with higher precision and covering larger regions of phase space, e.g. at both low and high  $p_t$  and forward pseudo-rapidities. Together with expected improvements of the calculations it is desirable that the picture of heavy quark production can be completed by the turn-on of the LHC.

## CHAPTER 4

# STRATEGY AND EVENT RECONSTRUCTION

---

The distributions used for the heavy flavour signal extraction exploit the long lifetime of heavy hadrons and are derived from the reconstructed impact parameter or decay length spectra. In this chapter the reconstruction chain of the objects which enter in these distributions are explained. In section 4.1 introductory remarks on the general analysis strategy are given.

The two tracking detectors measuring charged particles' flight trajectories in the central region of the detector have been described in section 3.1.1. The CJC is the backbone of the track measurement here, but the desired accurate track resolution for an efficient lifetime tagging can only be achieved in combination with the information of the central silicon detector CST. The association of CST hits and the combined CJC-CST track fit determine the CST improved track parameters which enter the lifetime distributions.

The CJC and CST track reconstruction and their linking will be outlined in sections 4.2 and 4.3. The impact parameter is measured with respect to the primary event vertex. Section 4.4 covers this aspect followed by the discussion of the impact parameter resolution in section 4.5.

Tracks are associated to jets for which the lifetime content is then studied separately. The reconstruction of jets is based on the definition of the hadronic final state objects, called HFS, which are discussed in section 4.6. Finally, the algorithm which builds the jets from the HFS objects as input is described in section 4.7. A brief overview of the observables which are measured in the present analysis closes this chapter.

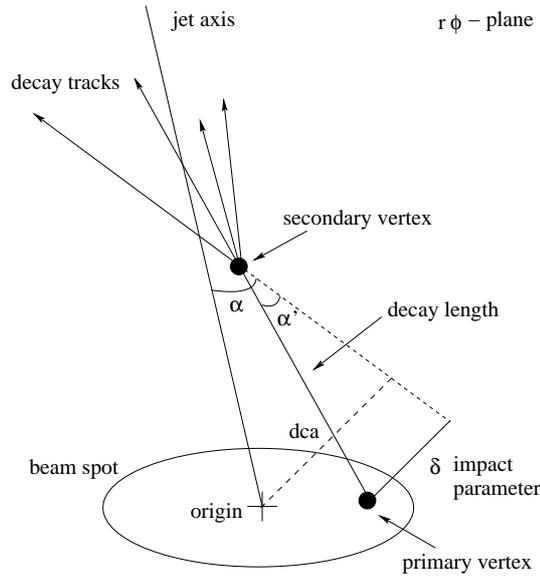
### 4.1 Analysis Strategy

Before entering the details of the analysis, some key aspects of the measurement are reviewed in this section.

Most measurements in the heavy flavour sector at H1 and ZEUS are based on the explicit reconstruction of  $D$ -Mesons (e.g. [16]) or a muon produced in the semi-leptonic decay mode (e.g. [14]). Although the branching ratio of about 10% for the decay to muons is small<sup>1</sup>, it

---

<sup>1</sup>  $BR(c \rightarrow \mu X) = (9.6 \pm 0.4)\%$  and  $BR(b \rightarrow \mu X) = (10.95^{+0.29}_{-0.25})\%$  [49].



**Figure 4.1:** Schematic illustration of the impact parameter in the  $r\phi$ -plane.

provides a clean experimental signature. Also combinations of both strategies are possible (e.g. [70]). However, the branching fractions and the lepton identification requirements restrict the statistical accuracy of the data.

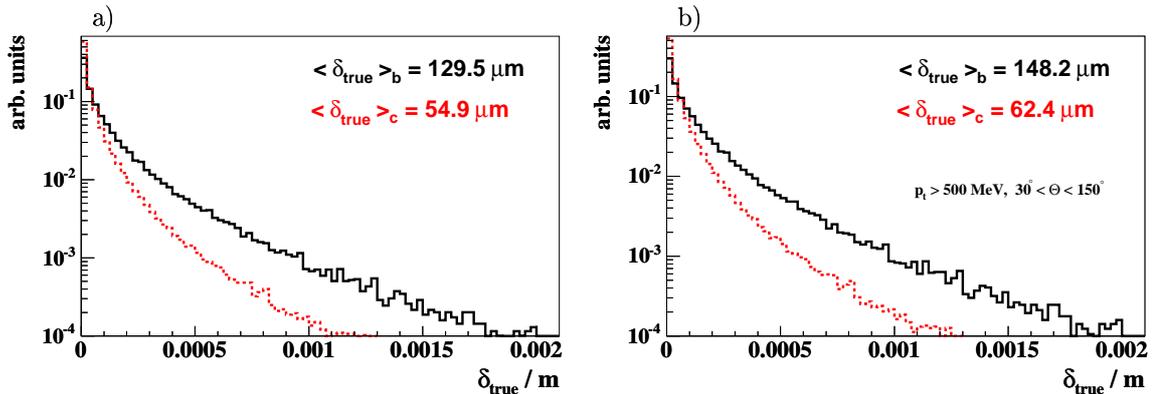
In the present analysis an inclusive ansatz using the displacement of tracks from the primary vertex is chosen to measure the fractions of charm and beauty events in a dijet photoproduction event sample. Dedicated triggers are operational at H1 which are sensitive to events with hard jets above a certain minimum transverse momentum. This analysis is based on a high statistics data sample which covers higher transverse jet momenta than previously possible.

### Impact Parameter

Events containing heavy quarks are distinguished from light quark events by the long lifetimes of charm and beauty flavoured hadrons, which lead to displacements of tracks from the primary vertex. These displacements can be quantified by the *impact parameter*. The definition of the impact parameter is illustrated in figure 4.1. Due to the limited  $z$ -resolution of the detector, the measurement and all following considerations are made for the  $r\phi$ -plane only.

A heavy hadron is produced at the primary event vertex inside the beam spot which is the transverse profile of the interaction region. Due to its long lifetime, the heavy hadron travels typically for a few hundred micrometers (cf. table 2.2) through the detector and decays at the secondary vertex, also referred to as the decay vertex. The transverse distance between the primary and the decay vertex is denoted as the *decay length*. The impact parameter  $\delta$  of a particle produced at the secondary vertex is the closest distance in the  $r\phi$ -plane between the particle's trajectory and the primary vertex.

In section 2.6 the exponential dependence of a particle's characteristic decay length  $\ell$  and the particle's real traversed distance  $L$  in the laboratory frame between production and decay vertex has been discussed. Since the analysis is carried out in the  $r\phi$ -plane, only the



**Figure 4.2:** Normalised distributions of the true track impact parameter  $\delta_{true}$  for charm (dotted) and beauty (full) events calculated with PYTHIA; a) without cuts and b) for tracks with  $p_t > 500 \text{ MeV}$  in the angular acceptance range of the CST.

transverse part of  $L$

$$L_T = L \cdot \sin \theta \quad (4.1)$$

is relevant here with  $\theta$  being the polar angle of the decaying particle. The real impact parameter  $\delta_{true}$  of a particle originating from a hadron decay is then defined as

$$\delta_{true} = L_T \cdot \sin \alpha, \quad (4.2)$$

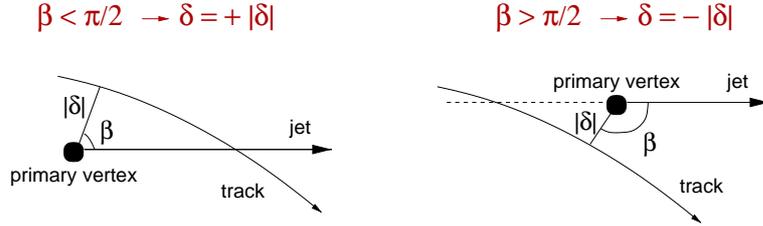
where  $\alpha$  is the angle between the hadron direction and the direction of the decay particle. The boost dependence of  $L_T$  is approximately compensated by the factor  $\sin \alpha$ .

The lifetime difference between charm and beauty flavoured hadrons leads to significantly different decay spectra  $\delta_{true}$ , as illustrated in figure 4.2. For central tracks with  $p_t > 500 \text{ MeV}$  the mean of the true impact parameter distribution is predicted to be about  $62 \mu\text{m}$  and  $148 \mu\text{m}$  for charm and beauty decays, respectively.

The flight direction of the decaying hadron is approximated by a jet-based reference axis. Thus a sign is ascribed to the impact parameter  $\delta$ , which reflects the relative orientation of the single track with respect to the jet axis. The sign is positive if the angle  $\beta$  between the axis of the associated jet and the line joining the primary vertex to the point of closest approach of the track is less than  $90^\circ$ , and is defined as negative otherwise. Figure 4.3 gives a schematic illustration of the sign convention.

Events without lifetime information have a true impact parameter of  $\delta_{true} = 0$ . Their reconstructed spectrum of  $\delta$  will therefore be symmetric around zero, the width of the distribution reflecting the finite track and vertex reconstruction resolutions. Events with decays of long-lived particles are expected to have an excess at positive  $\delta$  values, resulting in an asymmetric  $\delta$  distribution.

In order to use an impact parameter based observable to determine the charm and beauty content of a selected dijet event sample, all quantities entering the calculation of  $\delta$  need to be reconstructed with sufficient precision. According to the expected  $\delta_{true}$  values an impact parameter resolution of the order of  $100 \mu\text{m}$  is needed. Also a good knowledge of the jet reference axis is essential. Otherwise, an increased number of signal events with large  $\delta_{true}$  will be reconstructed with negative sign, making long lifetime effects indistinguishable from insufficiently modelled resolution.



**Figure 4.3:** Sign convention of the impact parameter  $\delta$ .

With the impact parameter method an event-by-event separation of beauty, charm and light quark events is not possible. Instead, the contribution from the various quark flavours are determined on a statistical basis. The Monte Carlo simulations are used to model the contributing processes and the relative fractions of beauty, charm and light quark events are then obtained from a least squares fit to observables based on the definition of  $\delta$ . The variables and the fit procedure are introduced and discussed in sections 6.2 and 6.3.

### Decay Length

A complementary ansatz to distinguish events with long lived hadrons from events with zero lifetime is to reconstruct explicitly a secondary vertex. In contrast to the impact parameter based method, the reconstruction of the hadron decay vertex requires at least two well measured tracks and is therefore less efficient by construction. Also, the impact parameter depends only weakly on the boost of the decaying particle, whereas the decay length directly depends on it (cf. equation 2.24).

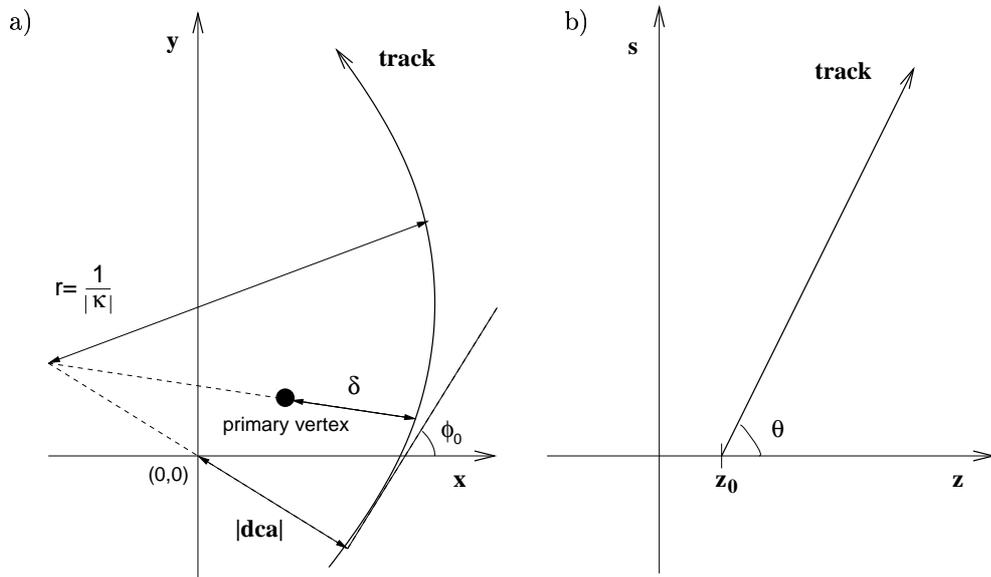
However, it turns out that the decay length reconstruction sometimes can be favorable, e.g. it shows to have a better separation power of light to heavy quark events than the impact parameter method. The explicit reconstruction of a secondary vertex is addressed in section 6.5.

## 4.2 Track Reconstruction

The reconstruction of the track parameters is based on hits in the central tracking detectors. Due to the presence of a homogeneous magnetic field parallel to the  $z$ -direction, charged particles are bent in the  $r\phi$ -plane. The bending radius  $r$  depends on the transverse momentum  $p_t$  of the particle and the strength of the magnetic field. The flight path can be described by a helix and its parameterization in H1 coordinates as a function of the arclength  $s$  is given by

$$\begin{aligned}
 x(s) &= +\left(dca - \frac{1}{\kappa} \cdot \sin(\phi_0) + \frac{1}{\kappa} \cdot \sin(\phi_0 + \kappa s)\right), \\
 y(s) &= -\left(dca - \frac{1}{\kappa} \cdot \cos(\phi_0) - \frac{1}{\kappa} \cdot \cos(\phi_0 + \kappa s)\right), \quad s \geq 0, \\
 z(s) &= z_0 + s \cdot \cot(\theta).
 \end{aligned} \tag{4.3}$$

The five track parameters  $(\kappa, \phi_0, \theta, dca, z_0)$  unambiguously describe the helix with respect to the origin  $(0,0,0)$ . These quantities are displayed in figure 4.4.



**Figure 4.4:** The track parameters  $(\kappa, \phi_0, \theta, dca, z_0)$  describe the helix in a) the  $r\phi$  and b) the  $zs$  plane with respect to the origin of the H1 coordinate system.

$\kappa = \frac{1}{r}$  is the curvature or inverse bending radius of the trajectory with positive sign for negative charged particles and vice versa. The flight direction in the transverse plane is given by the azimuthal angle  $\phi_0$ . The angle  $\phi_0$  is measured at the point of closest approach to the  $z$  axis which is seen as the starting point of the helix ( $s = 0$ ).

The distance of closest approach  $dca$  denotes the signed minimal radial distance of the track to the origin  $(0,0)$  in the  $r\phi$ -plane. The sign of  $dca$  is chosen equal to the sign of  $\kappa$ , if the origin is inside of the circle which describes the track in the  $r\phi$  plane, otherwise it is chosen opposite to it. In the  $zs$  plane the track is described by a straight line which intercepts the  $z$ -axis at the point  $z_0$ . It has a slope  $\cot(\theta)$ , where the polar angle  $\theta$  gives the flight direction with respect to the positive  $z$ -axis.

The reconstruction of the flight trajectories in the Central Jet Chambers and their corresponding track parameters are determined by a fit to the measured hits in the CJC. The pattern recognition algorithm fits the track parameters in the transverse plane first because the hit resolution is with  $\mathcal{O}(200 \mu\text{m})$  superior to the  $z$  resolution by two orders of magnitude. For the  $r\phi$  plane a circular trajectory is assumed. The parameters  $\theta$  and  $z_0$  are determined by a straight line fit in the  $zs$  plane.

If a real particle travels through the tracking devices its flight direction is distorted mainly by multiple scattering in the material of the detector. When extrapolating a track measured in the CJC to the CST or still further to the particles' production point inside the beam pipe this effect has to be taken into account. The track extrapolation routine CSTCOR handles the distortion coming from multiple scattering, energy loss and inhomogeneities of the magnetic field. A detailed description of CSTCOR can be found for instance in [71].

### 4.3 CST improved tracks

The position measurements of the CST hits provide the accuracy necessary to resolve the rather small distances separating the production and decay vertices of long lived heavy hadrons. However, the two CST layers alone are not sufficient to reconstruct a track. Therefore the identified CJC tracks are extrapolated to both layers of the CST using the CSTCOR routine.

The association of CST hits and CJC tracks is done using the linking routine CSTLIN [72]. In this algorithm non-vertex fitted tracks (DTNV) are chosen as input and the linking takes place separately in the  $r\phi$ - and  $zs$ -plane. Thus CST improved track parameters can be obtained using p-side information alone without being affected by possible n-side inefficiencies which can occur due to a worse S/N ratio and hit ambiguities on the n-side (cf. section 3.1.1). If CST hits are found in both the inner and outer layers, they are linked simultaneously, i.e. the hit combination of inner and outer layers is chosen which maximises the total link probability. This is superior to a separate linking, since obviously the positions of all hits emerging from one particle are correlated.

In particular, the CST improved  $r\phi$  track parameters are obtained according to the following procedure: For each CJC non vertex fitted track and all possible combinations of p-side hits from the inner and outer CST layers, a circle track fit is applied minimizing the following  $\chi^2$  function:

$$\chi^2 = (\vec{T} - \vec{T}_{DTNV})^t V_{DTNV}^{-1} (\vec{T} - \vec{T}_{DTNV}) + \sum_{\text{layer } i} \sum_{\text{hits } j} \frac{d_{ij\vec{T}}^2}{\sigma^2(d_{ij\vec{T}})}. \quad (4.4)$$

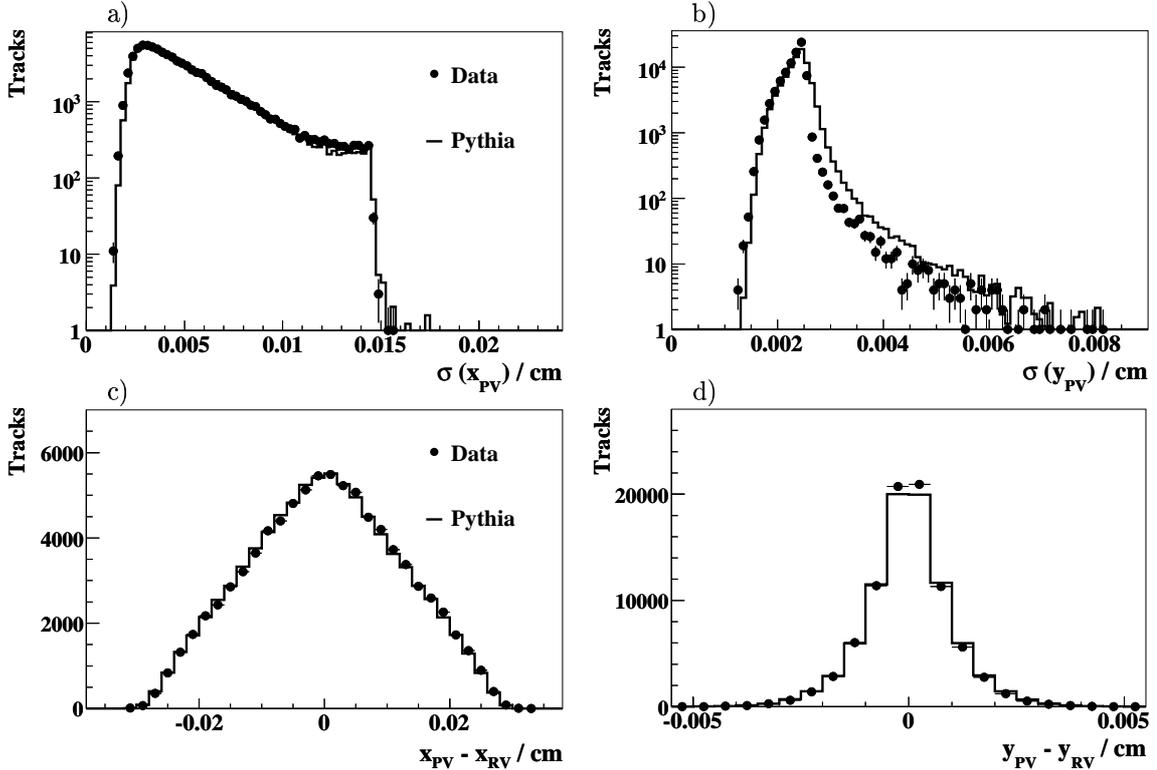
The first sum runs over the inner and the outer CST layer and the second sum runs over their corresponding hits.  $\vec{T} = (\kappa, \phi_0, dca)$  denote the  $r\phi$  track fit parameters,  $\vec{T}_{DTNV}$  the CJC track parameters and  $V_{DTNV}$  the covariance matrix. The separation between the track and the hit in  $r\phi$  is given by  $d_{ij\vec{T}}$  and its corresponding error by  $\sigma(d_{ij\vec{T}})$ . Solutions with a maximal number of CST hits are preferred but must have a reasonable  $\chi^2$ . If more than one CJC track can be associated to one or more CST hits, the CJC track with the smallest extrapolation error is chosen.

### 4.4 Primary Vertex Reconstruction

The transverse profile of the interaction region at HERA, the *beam spot*, extends over a few hundred micrometers with a rather stable mean position for a sequence of runs. For the determination of the so called *run vertex* the average coordinates  $x_{beam}$  and  $y_{beam}$  (both defined at  $z = 0$ ) of the *ep* interaction point as well as the *beam tilts*<sup>2</sup>  $\alpha_x$  and  $\alpha_y$  are determined by collecting information from many events. The run vertex coordinates are obtained by a least-squares fit minimising the overall distances of closest approach using only well measured non-vertex fitted tracks with high transverse momentum.

The standard procedure at H1 for the reconstruction of the *primary vertex* which is also referred to as the *event vertex* is called CSPRIM and is described in [73]. To ensure a high reconstruction quality, it is based on the selection of CST improved tracks and makes use

<sup>2</sup>The incoming electron and proton beams at H1 are slightly tilted with respect to the  $z$ -axis.



**Figure 4.5:** The calculated errors on the primary vertex positions (top) and the difference of the reconstructed primary vertex and the run vertex (bottom). H1 data are shown and compared with the PYTHIA simulation.

of the CST improved run vertex as a constraint. The track selection requires all tracks to be compatible with the run vertex within two standard deviations. Note, that the tracks used later in the impact parameter measurement are not excluded explicitly from the primary vertex fit. This leads to a small bias in the corresponding distributions of the impact parameter significance as discussed in section 6.2.

In this analysis the primary vertex fit is repeated based on CST improved non-vertex fitted tracks after the treatment of data and simulation as discussed in the following section. Furthermore, in contrast to the standard procedure, the uncertainty of the beam tilt in the error of the run vertex is taken into account in the simulation. The beam spot position is generated according to two Gaussian distributions with fixed average positions and with the width taken from data.

### Primary Vertex Resolution

In figure 4.5 a) and b) the calculated errors on the radial position of the event vertex are shown. Since the beam position is precisely known, the maximal size of the primary vertex errors is given by the size of the elliptical beam spot. The upper edge of the distribution of  $\sigma(x_{PV})$  corresponds to good approximation to the beam spot extent in  $x$  of  $\sigma_x = 145 \mu\text{m}$ . The average resolution however, is much better and measured to be  $\langle\sigma(x_{PV}^{Data})\rangle = 50.9 \mu\text{m}$  in data and  $\langle\sigma(x_{PV}^{MC})\rangle = 51.1 \mu\text{m}$  in the Monte Carlo simulation.

The already small beam spot width in  $y$  of  $\sigma_y = 25 \mu\text{m}$  can hardly be resolved and an average resolution of  $\langle\sigma(y_{PV}^{Data})\rangle = 23.0 \mu\text{m}$  in data and  $\langle\sigma(y_{PV}^{MC})\rangle = 23.7 \mu\text{m}$  in the Monte

Carlo simulation is achieved. The tails towards large values of  $\sigma(y_{PV})$  are due to short luminosity runs in which an accurate determination of the run vertex is difficult. Less than 2% of the selected data events have an error in  $y$  which is larger than  $25 \mu\text{m}$ .

Whereas the error on the primary vertex position in  $x$  is well modelled by the simulation, a small shift towards larger errors is observed in the simulation of  $\sigma(y_{PV})$ . This effect can be understood by the non perfect simulation of the beam tilt in the Monte Carlo. The beam tilt contributes to the error on the run vertex and thus enters also the calculation of the error on the reconstructed primary vertex. Since in  $y$  the tilt is one order of magnitude larger than in  $x$ , its accurate description in the simulation has a more pronounced effect on the error calculation in  $y$ . However, due to the small width of the beam spot in  $y$ ,  $\sigma(y_{PV})$  is smaller than  $\sigma(x_{PV})$  and the non perfect description of  $\sigma(y_{PV})$  by the simulation has a negligible effect.

The difference of the reconstructed primary vertex and the run vertex is displayed in figure 4.5 c) and d). In the  $x$  ( $y$ ) projection the distribution has a RMS of  $120 \mu\text{m}$  ( $10 \mu\text{m}$ ) which has to be compared with the beam spot size. Good agreement of the data and the Monte Carlo simulation is observed.

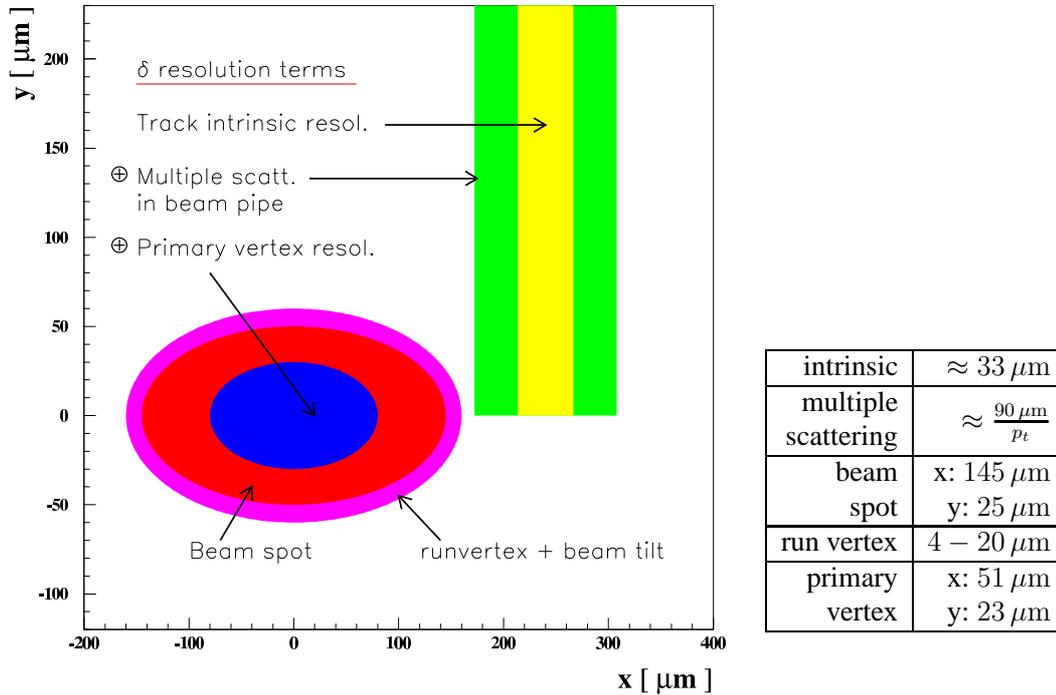
## 4.5 Impact Parameter Resolution

The understanding and description of the impact parameter resolution is crucial for this analysis. On the one hand the resolution has to be proven small enough to disentangle tracks coming from long lived hadrons from tracks of the zero lifetime combinatorial background. On the other hand the description of the impact parameter resolution must be well understood in order to keep the systematic error small. For a discussion of the systematic error emerging from the uncertainty of the impact parameter resolution see section 7.1.

Figure 4.6 illustrates the various sources which add up to the total impact parameter resolution. For tracks with hits in both layers of the CST the *intrinsic* resolution due to uncertainties on the internal alignment of the CST with respect to the other detector components is measured to be  $33 \mu\text{m}$ . Effects from *multiple scattering* with the beam pipe and the first layer of the CST are strongly  $p_t$  dependent and contribute by  $\approx 90 \mu\text{m}/p_t [\text{GeV}]$ . Further contributions are all dependent on the reference point to which the impact parameter is measured. When using the run vertex (RV) as reference the impact parameter is denoted as  $dca_{RV}$ . Otherwise, when the impact parameter is measured with respect to the real, reconstructed primary event vertex (PV), it is called  $dca_{PV}$  or simply  $\delta$ . Accordingly, the primary vertex resolution, the beam spot and uncertainties on the run vertex and the beam tilt can contribute to the total impact parameter resolution.

### Track Resolution

Having identified the different sources contributing to the impact parameter resolution, these must be discussed in more detail. In order to preserve a high statistics data sample, in the final selection tracks down to a transverse momentum of  $500 \text{ MeV}$  are analysed. In the low  $p_t$  region the impact of multiple scattering effects is large and must be well controlled. For a reasonable description of the data by the simulation adjustments are necessary.



**Figure 4.6:** Schematic view of the various contributions to the impact parameter resolution taken from [74]. The numbers in the table correspond to the 1999-2000 data taking phase.

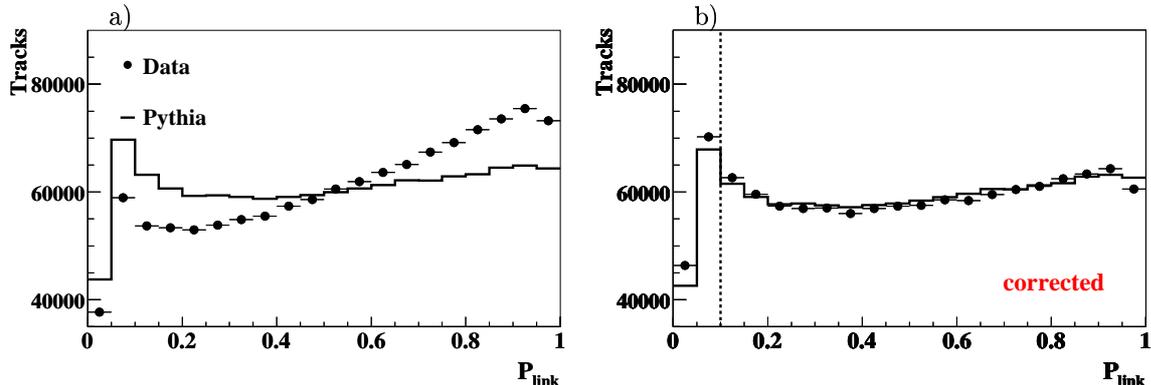
The description of the track resolution can be investigated by looking at the CJC-CST track linking probability,  $\mathcal{P}_{link}$ , which is a direct measure of the accuracy of the covariance matrix of the track parameters. It is defined as

$$\mathcal{P}_{link}(\chi^2, N) = \frac{1}{\sqrt{2^N} \Gamma(N/2)} \int_{\chi^2}^{\infty} e^{-\frac{1}{2}t} t^{-\frac{1}{2}N-1} dt. \quad (4.5)$$

$\mathcal{P}_{link}$  reflects the probability of a successful matching of the CST hits and the CJC tracks. Strictly speaking,  $\mathcal{P}_{link}$  is the probability of having a larger  $\chi^2$  value as achieved by the minimisation algorithm and is based on the  $\chi^2$  distribution following equation 4.4 and the available number degrees of freedom,  $N$ , which are the linked CST hits in this case. The probability distribution should be flat between 0 and 1 if the covariance matrix is accurately described.

Extensive investigations demonstrating a good understanding of the track linking probability after adjusting the covariances of the CST track parameters have been performed in [75]. For tracks with  $p_t > 2.5 \text{ GeV}$  correction factors of  $f_{r\phi} = 1.35$  in the  $r\phi$ -plane and  $f_{rz} = 1.0$  in the  $rz$ -plane have been determined for the intrinsic part of the covariances. In addition in the Monte Carlo simulation the intrinsic track resolution was smeared by the same factors to adjust the track resolution to that of the data.

The contribution from multiple scattering was studied separately by selecting events with a relaxed cut on the transverse momentum ( $p_t > 0.3 \text{ GeV}$ ). It was found that in the simulation the amount of dead material is underestimated giving rise to an additional scaling factor,  $f_{MS} = 1.22$ . This factor corresponds effectively to a beam pipe, thicker by 22%.



**Figure 4.7:** Distribution of the CJC-CST link probability  $\mathcal{P}_{Link}$  before (left) and after correction (right). The vertical dashed line indicates the cut applied in the track selection. See text for details.

Based on the conclusions given in [75] the reconstructed data sample and all Monte Carlo samples used in this analysis have been treated with the above discussed scale factors. However, in order not to overestimate the influence of the multiple scattering part, a reduced factor of  $f_{MS} = 1.09$  is chosen, while keeping the factors  $f_{r\phi} = 1.35$  and  $f_{rz} = 1.0$  for the intrinsic part. The initial choice of a smaller factor  $f_{MS} = 1.09$  allows to vary the  $p_t$  track cut in later stages of the analysis but makes a further  $p_t$  track dependent correction necessary in order to describe the linking probability for the track samples used in this analysis.

In figure 4.7 the CJC-CST link probability for tracks with  $p_t > 0.5$  GeV (as in the final track selection) is shown. On the left side the data distribution of  $\mathcal{P}_{Link}$  is compared with the simulation, both without further corrections. The Monte Carlo distribution is flat for values of  $\mathcal{P}_{Link} > 0.1$  while the data distribution is not. This indicates that the real track resolution in the data is still worse than described in the covariance matrix of the data and as assumed in the simulation. Therefore, a further  $p_t$  dependent correction is applied to the initial  $\chi_0^2$  of the data, following

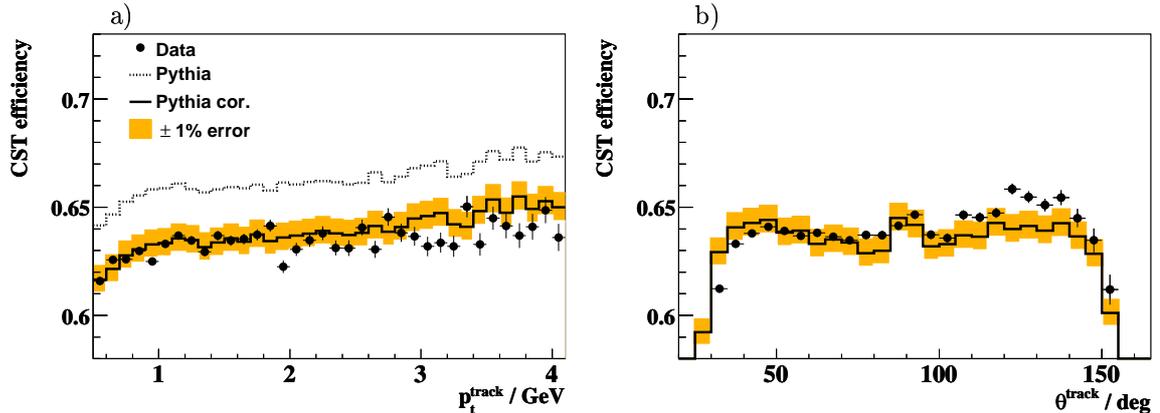
$$\chi_{cor}^2 = \chi_0^2 + \chi_0^2 \cdot \exp(f_0 + f_1 \cdot p_t). \quad (4.6)$$

For the parameters  $f_0$  and  $f_1$  values of -0.87 and -0.47 have been found, respectively. The distribution of  $\mathcal{P}_{Link}$  according to the modified  $\chi_{cor}^2$  is displayed in figure 4.7 b). The data distribution is almost flat for  $\mathcal{P}_{Link} > 0.1$  and is now well modelled by the simulation.

### Track Finding Efficiency

Based on the corrected CJC-CST link probability the track finding efficiency can be investigated. For this purpose tracks with at least 2 CST hits and  $\mathcal{P}_{Link} > 0.1$  are selected. The CST efficiency, defined as the ratio of the number of tracks fulfilling these CST cuts over the number of CJC tracks, is shown in figure 4.8 as functions of the track transverse momentum and polar angle. The total CST track finding efficiency includes the CST hit efficiency, the CJC-CST linking efficiency and losses due to inactive CST regions.

The efficiency is seen to be nearly flat but slowly rising with increasing  $p_t^{track}$  and amounts to  $\approx 63\%$  on average in data. The default Monte Carlo efficiency is a bit higher (dotted



**Figure 4.8:** CST track finding efficiency for tracks with at least 2 CST hits and  $\mathcal{P}_{\text{Link}} > 0.1$  as functions of  $p_t^{\text{track}}$  (a) and  $\theta^{\text{track}}$  (b). The Monte Carlo efficiency is shown separately for the default (dashed line) and modified (full line) version. The error band corresponds to a remaining uncertainty of  $\pm 1\%$ .

line) which is due to the still not perfect description of the CJC-CST link probability (cf. figure 4.7).

In order to correct for this small disagreement, in the simulation the track finding efficiency is downgraded by  $\approx 3\%$  with a small  $p_t$  dependence. The improved Monte Carlo simulation is shown as a full line. The 1% error band reflects the remaining uncertainty on the CST track finding efficiency. The agreement between data and simulation is good; this holds also for the efficiency as a function of  $\theta^{\text{track}}$ , displayed in figure 4.8 b).

### Impact Parameter Resolution

Having discussed the track resolution, the beam spot and the primary vertex resolution, the impact parameter resolution can now be investigated in detail.

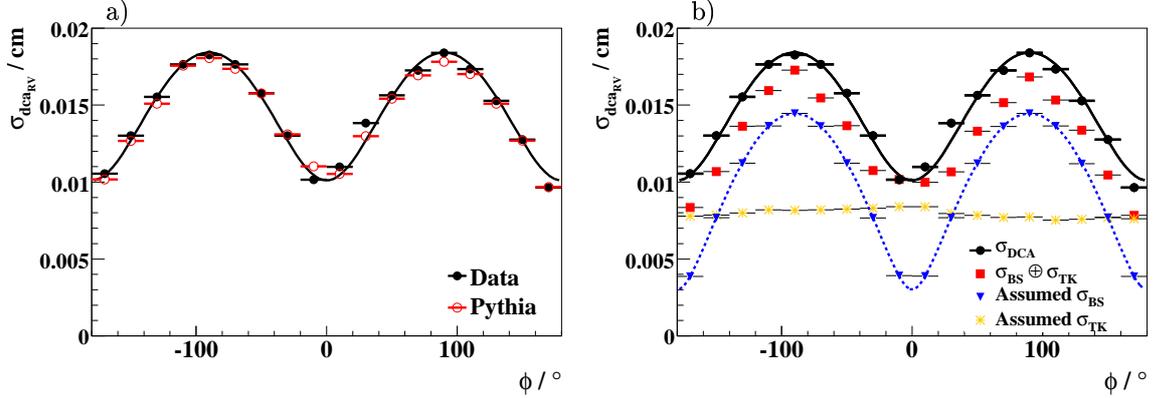
First, in order to separate contributions from the primary vertex fit, the run vertex is used as reference in the impact parameter definition. For tracks with at least two CST hits and a minimum transverse momentum of  $p_t > 0.5 \text{ GeV}$  the Gaussian width of the  $dca_{RV}$  distribution as a function of the track azimuthal angle  $\phi$  is shown in figure 4.9 a). Neglecting the error on the run vertex, the width of  $dca_{RV}$  depends on the beam spot size and the track resolution. The  $\phi$  dependence of the  $dca_{RV}$  resolution is expected to follow a function of the form

$$\sigma_{dca_{RV}}^2(\phi) = \sigma_{Tk}^2 + \sigma_x^2 \sin^2(\phi) + \sigma_y^2 \cos^2(\phi), \quad (4.7)$$

where  $\sigma_{Tk}$  accounts for the track resolution and  $\sigma_x$  and  $\sigma_y$  reflect the transverse size of the beam spot.

The parameters of the beam spot size have been independently measured to be  $\sigma_x = 145 \mu\text{m}$  and  $\sigma_y = 25 \mu\text{m}$  [75]. These parameters are also used in the simulation and in the fit (black line) following equation 4.7. Both give a good description of the  $\phi$  dependence seen in the data.

In figure 4.9 b) the size and structure of the  $dca_{RV}$  resolution is investigated in more detail. The full dots and the full curve are the same as in figure 4.9 a) representing the *measured* impact parameter resolution of the data, i.e. the Gaussian width of the measured



**Figure 4.9:** Gaussian width of the  $dca_{RV}$  distribution as a function of the track direction around the beam. Inclusive tracks with at least two CST hits and  $p_t > 0.5 \text{ GeV}$  are shown for data and simulation. The curves corresponds to a fit  $\sim \sigma_x^2 \sin^2(\phi) + \sigma_y^2 \cos^2(\phi)$  with the parameters  $\sigma_x = 145 \mu\text{m}$  and  $\sigma_y = 25 \mu\text{m}$ . On the right side the contributions from the beam spot  $\sigma_{BS}$  and the tracking resolution  $\sigma_{TK}$  are shown separately.

$dca_{RV}$  distribution. The *assumed* contribution coming from the track resolution,  $\sigma_{TK}$ , and from the beam spot,  $\sigma_{BS}$ , can be separated by exploiting their different  $\phi$  dependence. The assumed track resolution  $\sigma_{TK}$  is *calculated* from the assumed contributions of the intrinsic track resolution plus multiple scattering effects plus tracking fit errors using error propagation. Making use of the such extracted values of  $\sigma_{TK}$ , the assumed beam spot contribution  $\sigma_{BS}$  can then be calculated from the difference of the assumed, i.e. calculated  $\sigma_{dca_{RV}}$  values and  $\sigma_{TK}$ .

Whereas the contribution coming from the tracks is constant at  $\approx 90 \mu\text{m}$ , the extension of the beam spot can be read from the dotted line. As expected, for  $\phi = \pm 90^\circ$  essentially the large dimension of the beam ellipse in  $x$  direction is seen, resulting in a resolution of  $\approx 145 \mu\text{m}$ , whereas towards  $\phi = 0^\circ$  ( $\phi = 180^\circ$ ) the resolution reaches  $\approx 30 \mu\text{m}$ , corresponding to the beam extension in  $y$ .

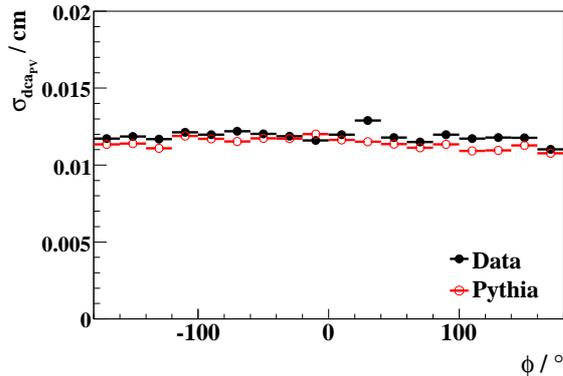
Adding up  $\sigma_{TK}$  and  $\sigma_{BS}$  quadratically results in the distribution which is represented by the full boxes. The remaining difference to the measured impact parameter resolution is due to missing contributions of multiple scattering effects which would increase the track contribution independent of  $\phi$ . This coincides with the observation made in the previous paragraphs that the real track resolution in the data is worse than assumed in their covariance matrix.

In summary, using the run vertex as an anchor for the impact parameter measurement results for tracks with a transverse momentum larger than  $500 \text{ MeV}$  in a resolution of  $\approx 100 - 180 \mu\text{m}$ , depending on the track direction in  $\phi$ .

The resolution can be improved by replacing the run vertex with the primary event vertex. In figure 4.10 the  $\phi$ -dependence of the impact parameter resolution with respect to the primary vertex  $\sigma_{dca_{PV}}$  is shown.

The distribution is dominated by the (flat) contribution of the track resolution (cf. figure 4.9 b)) such that the  $\phi$ -dependence of the vertex resolution is seen to be suppressed. A mean impact parameter resolution of  $\approx 115 \mu\text{m}$  is observed in both, data and simulation.

In conclusion, the reconstruction of the run vertex and the primary vertex as well as the



**Figure 4.10:** Gaussian width of the  $dca_{PV}$  distribution as a function of the track direction around the primary vertex. Inclusive tracks with at least two CST hits and  $p_t > 0.5 \text{ GeV}$  are shown for data and simulation.

track resolution are well under control. The best impact parameter resolution achieved with tracks down to 500 MeV is about  $115 \mu\text{m}$ . Comparing this to a typical  $\delta_{true}$  value of  $148 \mu\text{m}$  and  $62 \mu\text{m}$  for beauty and charm hadron decays respectively (cf. section 4.1), a good understanding of the  $\delta$ -resolution is clearly a key issue of this analysis. In addition further discriminators might be needed in order to separate beauty, charm and light quark events based on lifetime effects.

As discussed in chapter 6 this additional discriminator is the track multiplicity.

## 4.6 HFS Reconstruction

In this section the reconstruction of the *hadronic final state* (HFS) will be outlined. The HADROO2 algorithm [76] was used in this analysis to reconstruct the hadronic energy of the events. Only a brief description of the algorithm is given here.

The basic idea behind the HADROO2 algorithm is to use a combined information from the track measurement and the calorimeter, depending on the uncertainty of the measurements in the respective devices.

### Track and Cluster selection

The algorithm starts by selecting good quality tracks as defined by the heavy flavour group. These tracks, measured with the central and forward tracking detectors, are classified in three categories, Central, Combined and Forward and are called *Lee West tracks*. For details on the specification see [76].

Calorimetric clusters are made out of LAr and SpaCal clusters only. The main requirement when selecting the clusters is to suppress the noise coming from the electronics or non  $ep$ -physics events.

Any identified electrons and muons which are not isolated are considered as being part of the HFS but their associated tracks and clusters are removed from the input list.

### The algorithm

The HADROO2 algorithm is based on the combination of the selected tracks and clusters. The track and cluster information is matched avoiding double counting of energies. The decision whether to take the track or the cluster information to construct the HFS object is based on a comparison on their relative resolutions. Due to possible contributions of neutral particles the exact precision of the calorimeter is unknown. Therefore the average relative error expected for the calorimeter measurement is calculated as

$$\left(\frac{\sigma_E}{E}\right)_{LAr}^{expected} = \frac{\sigma_{E_{LAr}}^{expected}}{E_{track}} = \frac{50\%}{\sqrt{E_{track}}}. \quad (4.8)$$

The track measurement is considered as being superior if

$$\left(\frac{\sigma_E}{E}\right)_{track} < \left(\frac{\sigma_E}{E}\right)_{LAr}^{expected}. \quad (4.9)$$

Generally the tracker measurement is better up to 25 GeV for central tracks ( $20^\circ < \theta < 160^\circ$ ). If equation 4.9 is true, the track information is preferred and an HFS object is created based on the track measurement. In order to avoid double counting of energies, electromagnetic (hadronic) energy in clusters within a cylinder of 25cm (50cm) around the extrapolated track is removed until the discarded energy is approximately equal to the track energy (details in [76]). Possible fluctuations of both measurements are taken into account in this procedure. Potentially remaining cluster energy is due to neutral particles or belongs to another track extrapolated in the same region of the calorimeter.

If equation 4.9 is false and if  $E_{track}$  is within  $2\sigma$  of  $E_{cylinder}$  (with  $\sigma_{E_{cylinder}} = 0.5\sqrt{E_{cylinder}}$ ), the track and the calorimetric energy measurements are considered to be consistent and the calorimetric measurement is used to create the HFS object.

If the track energy is below the measured cluster energy ( $E_{track} < E_{cylinder} - 1.96\sigma_{E_{cylinder}}$ ), the track measurement is used and the calorimetric energy is subtracted in the same way as described before when the track precision is better than the expected calorimeter precision. Otherwise, if  $E_{track} > E_{cylinder} + 1.96\sigma_{E_{cylinder}}$ , the cluster energy defines the HFS object and the track measurement is suppressed.

Once all tracks have been treated, HFS objects are created for the remaining clusters. These particles correspond to neutral hadrons with no associated track or to charged particles with a badly measured track.

## 4.7 Jet Algorithm

The best way to fully analyze  $\gamma p$  interactions would be to measure the final state quarks and gluons. However, due to the fact that the partons carry colour they must combine to a group of colourless hadrons, which are the measurable final state particles of the reaction. This group has a typical intrinsic transverse momentum of the hadronisation scale, i.e. several GeV. When the group has a transverse momentum relative to the beam axis much greater than this scale, a collimated formation of hadrons, called a jet, results.

One of the most important requirements of a jet algorithm is that it should minimise the effect of the long distance hadronisation. It turned out that the inclusive  $k_t$  algorithm

is best suited here. The  $k_t$  algorithm has other advantages like the fact that the cone algorithm problem with overlapping jets disappears (cf. e.g. [77]).

Another aspect of great importance is that the jet algorithm should be collinear and infrared safe. The reason is twofold. In perturbative QCD, divergencies occur whenever two massless partons are parallel (collinear) or one massless parton has a vanishingly small (infrared) energy. Both divergencies are cancelled in the total cross section by virtual contributions. However, for this cancellation to take place also in the jet calculation, the treatment of two parallel particles must be identical to the treatment of a single particle with their combined momentum. Similarly, the jet must not be affected by the addition of a soft particle.

On the other hand, from the experimental point of view, the jet algorithm must not depend on the resolution of the detector, e.g. when two parallel particles go into the same calorimeter cell, or on e.g. trigger response when additional low energy particles are emitted.

In the present analysis jets are always defined using the inclusive  $k_t$  jet algorithm [78] which has become standard in jet analyses at HERA. The inclusive  $k_t$  jet algorithm is a clustering algorithm which starts by finding pairs of particles nearby in phase-space and merging them together to form new pseudo-particles. This procedure continues iteratively until the event consists of a few well separated pseudo-particles, which are the output jets. In the present analysis the algorithm is applied in the laboratory frame using a  $p_t$  weighted recombination scheme with distance parameter  $R=1$  in the  $\eta\phi$ -plane. A more detailed description of the iterative procedure follows:

- For each object  $i$  a separation to the beam axis is defined

$$d_i = p_{t,i}^2 \quad (4.10)$$

and for each pair of objects a separation to each other is defined as

$$d_{i,j} = \min(p_{t,i}^2, p_{t,j}^2) R_{i,j}^2 / R, \quad (4.11)$$

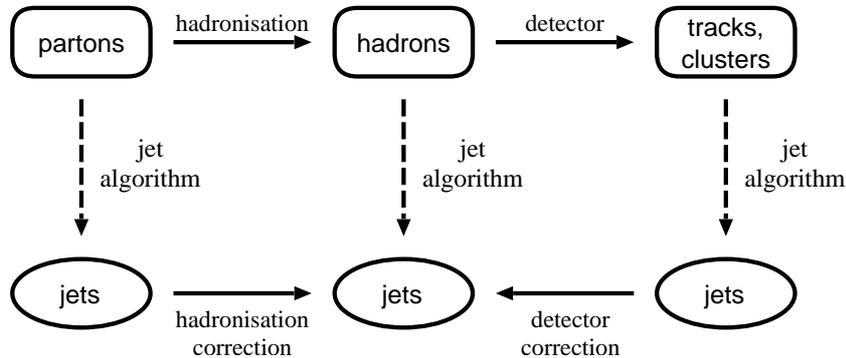
where  $R_{i,j}$  is the distance of two objects in the  $\eta\phi$  plane

$$R_{i,j} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \quad (4.12)$$

and  $R$  is the distance parameter.

- The smallest distance of a single object  $d_i$  or pair  $d_{i,j}$  is labeled  $d_{min}$ .
- If  $d_{min}$  is a single particle distance, remove object  $i$  and add it to the list of jets.
- If  $d_{min}$  is a pair distance then the two objects  $i$  and  $j$  are merged to a new object  $k$  with:

$$\begin{aligned} p_{t,k} &= p_{t,i} + p_{t,j} \\ \eta_k &= (p_{t,i}\eta_i + p_{t,j}\eta_j) / p_{t,k} \\ \phi_k &= (p_{t,i}\phi_i + p_{t,j}\phi_j) / p_{t,k}. \end{aligned} \quad (4.13)$$



**Figure 4.11:** Schematic diagram of the parton, hadron and detector level jets. At each different stage the jet finding algorithm can be applied which is used to compare detected data with parton level calculations.

This procedure is continued until the list of objects is empty. The resulting jets are massless and ordered ascendingly in  $p_t$ .

In figure 4.11 the various levels of the analysis chain are illustrated. In this analysis the same jet algorithm is applied to the *detector*, *hadron* and *parton* levels. In the Monte Carlo simulations jets can be reconstructed on the parton level using the partons produced from the hard scattering process after the parton shower. Similarly, one can run the jet algorithm on hadron level using the decayed hadrons of the Monte Carlo events. On the detector level jets are found based on the final state particles in both, data and the reconstructed Monte Carlo simulation.

In order to compare experimental data and theory, both need to be given at the same level. Usually the hadron level is chosen. Thus, the data are corrected for detector properties such as acceptance and efficiency. Next-to-leading order calculations are only available on parton level. The effects of the correction to the hadron level are discussed in section 5.8.4.

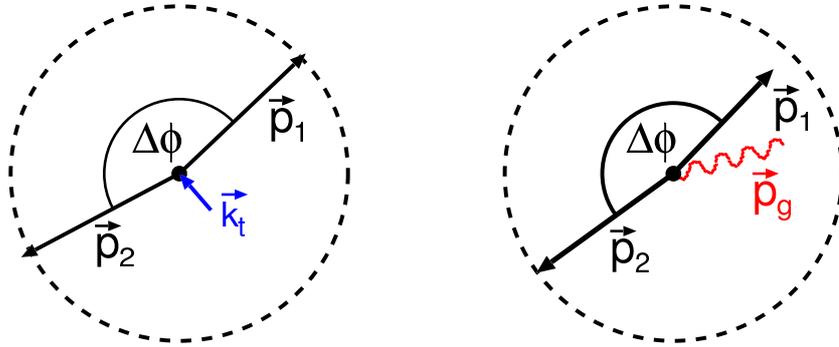
## 4.8 Observables

The measurement of dijet events gives the possibility to test several quantities which are related to the heavy quark production mechanism. The following considerations are based on the assumption that the two highest transverse momentum jets approximate the kinematics of the outgoing heavy quark pair in case of the boson gluon fusion process or one heavy quark and another parton, i.e. a light quark or a gluon in resolved processes.

A natural choice of observables to be considered in dijet photoproduction physics are the jet transverse momentum  $p_t$ , the jet pseudo-rapidity  $\eta$  and the azimuthal angles  $\phi$ . In order to avoid divergencies in the next-to-leading order calculations all observables must be infra-red safe [79]. Since the jets are ordered according to decreasing  $p_t$ , the pseudo-rapidities of the leading or second jet are both not infra-red safe observables [80]. However, the mean pseudo-rapidity

$$\bar{\eta} = \frac{1}{2}(\eta^{jet_1} + \eta^{jet_2}), \quad (4.14)$$

is infra-red safe and thus converges in the perturbative calculations to next-to-leading



**Figure 4.12:** Two heavy quarks in the  $r\phi$ -transverse plane. An initial transverse momentum  $k_t \neq 0$  of the incoming partons in the hard scattering (left) or gluon radiation (right) can lead to a non back-to-back topology of the outgoing partons. The picture is taken from [42].

order. The same is true for the absolute value of the difference in pseudo-rapidity of the two jets,  $|\eta^{jet_1} - \eta^{jet_2}|$ , and the absolute value of the difference in azimuthal angle of the two jets

$$(\Delta\phi)^{jj} = \begin{cases} |\phi^{jet_1} - \phi^{jet_2}| & \text{for } |\phi^{jet_1} - \phi^{jet_2}| < 180^\circ \\ 360^\circ - |\phi^{jet_1} - \phi^{jet_2}| & \text{for } |\phi^{jet_1} - \phi^{jet_2}| > 180^\circ. \end{cases} \quad (4.15)$$

The scattering angle  $\theta^*$  of the jets in their centre-of-mass system is accessible via the measurement of  $|\eta^{jet_1} - \eta^{jet_2}|$ :

$$\cos(\theta^*) = |\tanh((\eta^{jet_1} - \eta^{jet_2})/2)|. \quad (4.16)$$

The angle  $\theta^*$  is sensitive to the dynamics of the jet production, i.e. to the nature of the propagator of the involved matrix element. The observable is expected to show different distributions for direct and resolved events. In direct events the quark (spin 1/2) propagator yields a  $(1 - |\cos \theta^*|)^{-1}$  behaviour while in resolved events processes with a gluon (spin 1) propagator contribute, yielding a stronger rise like  $(1 - |\cos \theta^*|)^{-2}$ . It has been shown in a recently published analysis of charmed dijet events [9] that the data are consistent with the behaviour expected from a gluon propagator, i.e. it is due to charm excitation processes.

Cuts in jet transverse momentum and pseudo-rapidity lead to a distortion of the expected  $\cos \theta^*$  distribution. Differential cross sections measured as a function of  $\cos(\theta^*)$  are therefore evaluated with additional restrictions on the invariant dijet mass  $M_{jj}$  and on the mean pseudo-rapidity  $\bar{\eta}$  as explained in appendix C.

By measuring  $(\Delta\phi)^{jj}$  the azimuthal correlation of the two hard partons is investigated. At leading order the direct photon gluon process leads to production of the heavy quark pairs back-to-back in their centre-of-mass system, i.e. with  $(\Delta\phi)^{jj} = 180^\circ$ . This remains true in the laboratory frame, since the observable  $(\Delta\phi)^{jj}$  is invariant under a longitudinal boost along the  $z$ -axis.

Smaller values of  $(\Delta\phi)^{jj}$  are due to higher order processes such as gluon radiation or due to a non zero initial transverse momentum of the incoming partons. Both situations are illustrated in figure 4.12.

The observable  $x_\gamma^{obs}$  represents the fraction of the photon momentum entering the hard interaction. For the reconstruction of  $x_\gamma^{obs}$  the definition

$$x_\gamma^{obs} = \frac{\sum_{h \in \text{jet}_1} (E - p_z) + \sum_{h \in \text{jet}_2} (E - p_z)}{\sum_h (E - p_z)} \quad (4.17)$$

is used, where both sums in the numerator run over the particles associated with the two jets, respectively, and the sum in the denominator runs over all detected hadronic final state particles. For  $2 \rightarrow 2$  processes with massless partons  $x_\gamma^{obs}$  is equal to the true momentum fraction  $x_\gamma$  (eq. 2.18). Relation 4.17 can easily be derived from equation 2.18 for leading order processes (cf. e.g. [40]). With the above definition  $x_\gamma^{obs}$  is restricted to the range  $0 < x_\gamma^{obs} < 1$ . For direct photon gluon processes  $x_\gamma^{obs} \sim 1$ , as the hadronic final state consists only of the two hard jets and the proton remnant which is forced to the very forward region and contributes only little to  $\sum_h (E - p_z)$ . However, due to radiated partons and subsequent hadronisation processes the peak at 1 is smeared to lower values of  $x_\gamma^{obs}$ . In resolved photon processes  $x_\gamma^{obs}$  can be small due to the presence of the photon remnant in the hadronic final state. For next-to-leading order processes  $x_\gamma^{obs}$  is also smaller than unity, since the additional hard outgoing parton produces a third jet. In general, the observable  $x_\gamma^{obs}$  is sensitive to the resolved photon structure and higher order processes.

In this analysis differential charm and beauty cross section measurements will be presented as a function of the transverse momentum of the leading jet  $p_t^{jet_1}$ , the mean pseudo-rapidity  $\bar{\eta}$  of the two jets, the absolute difference of the azimuthal angles  $(\Delta\phi)^{jj}$  of both jets, the variable  $\cos\theta^*$  and the variable  $x_\gamma^{obs}$ . Furthermore, differential cross sections will be given for large values of  $x_\gamma^{obs} > 0.85$  where direct production processes are expected to dominate. The investigation of differential cross sections in the region of small  $x_\gamma^{obs}$  is not possible. The reason is twofold. On the one hand the data sample is statistically limited in the region of small  $x_\gamma^{obs}$ . On the other hand hard gluon jets are expected to play an important role in that region of phase space. Since gluon jets carry no lifetime information the fit used to distinguish light, charm and beauty events does not converge (cf. section 6.3).

## CHAPTER 5

# DIJET EVENT SELECTION

---

In this chapter the inclusive photoproduction dijet sample is presented which is the basis for the measurement of the charm and beauty cross sections. The chapter starts with the description of the selection criteria and continues with a data to Monte Carlo comparison including the discussion of some aspects related to the topic of jet structure. At the end the data unfolding mechanism is introduced and the definition of the cross section is given.

### 5.1 Preselection

The data sample of this analysis was collected with the H1 detector in the years 1999-2000, when protons of 920 GeV energy were collided with positrons of 27.6 GeV energy, yielding a centre-of-mass energy of  $\approx 319$  GeV.

Various stages of cuts have been used to reduce the enormous number of events to a reasonable sized sample of candidate events. In a first step *good* and *medium* runs were selected. This means that only data taking periods with the main detector components operational and with stable running conditions were kept for which the integrated luminosity was measured. In a second step it was required that for each event the high voltage of the following subdetectors important for this measurement were at nominal settings: *LAr*, *SpaCal*, *CJC1*, *CJC2*, *CST*, *CIP*, *COP*, *Lumi*, *ToF*. In a next step certain run ranges have been excluded from the analysis. These are the runs where the CJC had a broken wire and the so-called 'shifted vertex runs'<sup>1</sup> at the end of the data taking period in the year 2000.

On the basis of the above conditions together with the trigger requirements discussed in section 5.4, the total integrated luminosity of the analysed event sample amounts to  $56.8 \text{ pb}^{-1}$ .

### 5.2 Photoproduction Event Selection

In this section the main cuts, which also define the accessible phase space of the measurement, will be discussed. Some of the selection cuts are motivated by CST detector acceptance requirements, which are needed to perform the lifetime analysis.

---

<sup>1</sup>Shifted vertex runs are runs where the bunches are not colliding at the nominal interaction point.

Non  $ep$ -background is rejected to a large extent by requiring the primary vertex of the event to be reconstructed within a certain range of the  $z$  position of the nominal vertex. This condition excludes events coming from collisions of the satellite bunches with the nominal bunches. Due to the limited CST acceptance, the  $z$ -distance of the vertex from the nominal interaction point, the  $z$ -Vertex, is restricted further to the range  $|z_{\text{Vertex}}| < 20$  cm.

### Kinematic selection

In order to select photoproduction events, it is necessary to anti-tag electrons in the main calorimeters. The scattered electron is searched for by electron finders in the LAr and SpaCal calorimeters. For the so-called untagged photoproduction events under study the photon virtuality is restricted by the SpaCal acceptance to  $Q^2 < 1$  GeV<sup>2</sup>.

The kinematic variables must be calculated using the hadrons since the scattered electron is not measured. The inelasticity  $y_{\text{JB}}$ , calculated with the Jacquet-Blondel method [81] is restricted to the range  $0.15 < y_{\text{JB}} < 0.8$ . The upper cut removes background from deep inelastic scattering (DIS) events, which might be still in the sample because of inefficiencies of the electron finder (cf. also section 5.5). DIS background events are characterised by values of  $E - P_z$  of typically twice the beam energy, i.e. 55.2 GeV. Since  $y_{\text{JB}} = (E - P_z)/(2E_e)$  such events are found at high inelasticities  $y_{\text{JB}}$ . The lower cut rejects events in the extreme forward region of the detector, thus ensuring a complete reconstruction of the final state.

### Jet Selection

The definition of the hadronic final state and the jet reconstruction were discussed in sections 4.6 and 4.7. Due to trigger conditions (cf. section 5.4) the jet selection criteria require a  $p_t$  of the highest transverse momentum jet of  $p_t^{\text{jet1}} > 11$  GeV, and the transverse momentum of the second highest transverse momentum jet to be  $p_t^{\text{jet2}} > 8$  GeV.

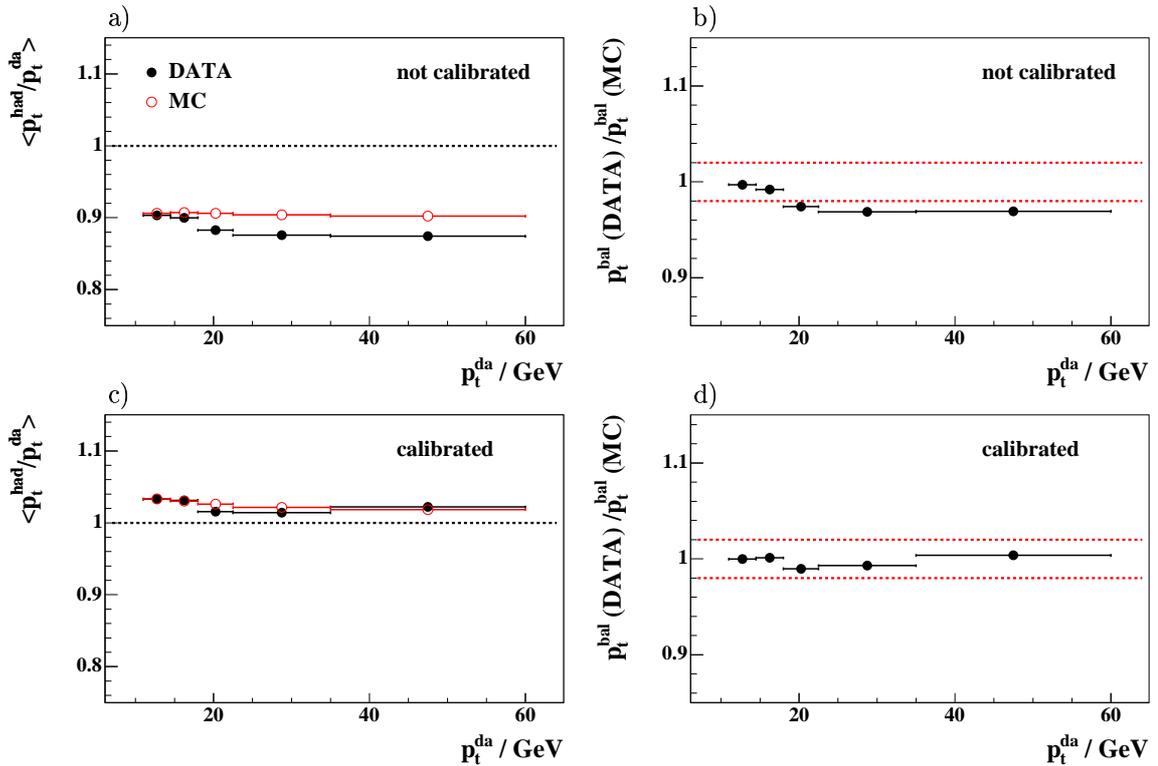
The need for an asymmetric cut on the transverse momentum of the two leading jets is motivated by theory arguments [79]. In this way regions of phase space affected by large uncertainties in the NLO calculation are avoided.

Since well measured CST tracks (cf. section 4.5) play a significant role for the measurement of charm and beauty jets, only jets in the central region of the detector are considered. For both jets the pseudo-rapidity is thus restricted to the range  $-0.9 < \eta < 1.3$ . The upper cut coincides exactly with the CST acceptance whereas the lower cut is chosen a bit tighter to exclude remaining background from deep inelastic scattering events.

## 5.3 Hadronic Final State Calibration

The algorithm to reconstruct the hadronic final state was already specified in section 4.6. An additional jet calibration is applied to the individual jets to further improve the reconstruction of the jet quantities. The principle of this calibration, a cross-check performed with a DIS neutral current dijet sample and the consequences for the systematic error on the hadronic energy scale, will be discussed in this section.

Since the tracks are already calibrated the aim of the calibration procedure is to change the energy of the calorimeter clusters only. As a basic principle it makes use of the fact



**Figure 5.1:** The mean  $p_t$  balance distributions as a function of  $p_t^{da}$  before (top) and after (bottom) final calibration.

that for neutral current DIS events the transverse momentum of the hadronic system  $p_t^{had}$  is balanced by the transverse momentum of the electron  $p_t^e$ .

The hadronic calibration is studied in [76] and the constants obtained there are applied in this analysis. The calibration procedure described in [76] is performed in two steps. First, a relative correction of the  $p_t$  balance, i.e. the ratio of the  $p_t$  of the hadronic system and the  $p_t$  of the electron, is applied separately for all wheels of the LAr and thus depends on the pseudo-rapidity of the jet. This correction is made for data relative to Monte Carlo. In a second step a  $p_t$  dependent absolute calibration of the HFS is performed and applied to both data and Monte Carlo.

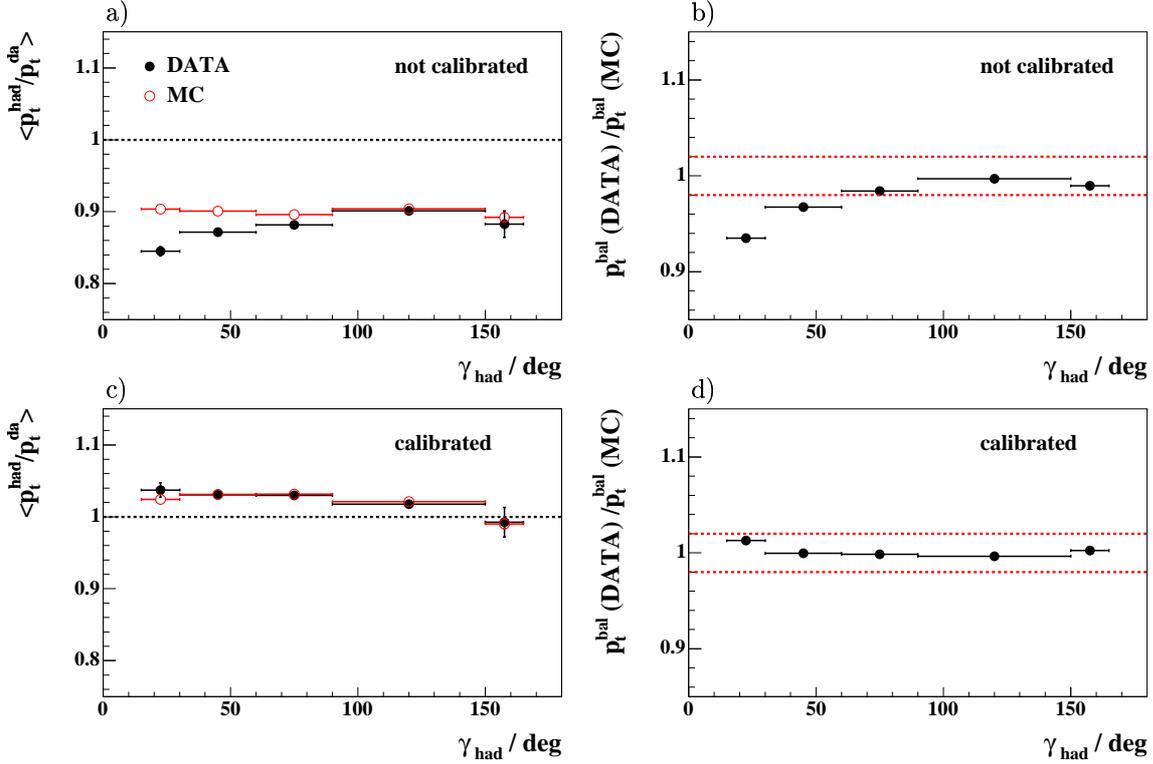
It was found in [76] that the transverse momentum balance after the recalibration agrees within 2% between data and Monte Carlo simulation.

In this analysis it is checked how well the transverse momentum balance is described by the simulation before and after recalibration. For this purpose a neutral current DIS dijet sample is compared to the DJANGO Monte Carlo simulation. The selection cuts are summarised in table 5.1.

Figures 5.1 and 5.2 show the evolution of the mean values of the  $p_t$  balance

$$p_t^{bal} = \frac{p_t^{had}}{p_t^{da}}. \quad (5.1)$$

Here  $p_t^{had}$  is the total hadronic transverse momentum calculated from all jets in the event and  $p_t^{da}$  is the transverse momentum obtained with the double angle method. The double angle method exploits the scattering angle of the outgoing lepton and the effective angle



**Figure 5.2:** The mean  $p_t$  balance between electron and hadrons as a function of  $\theta_{had}^e$  before (top) and after (bottom) calibration.

of the hadronic system. Its major advantage is that it is almost insensitive to the energy scale itself and hence ideal to test the calibration of the calorimeter.

The  $p_t$  balance as a function of  $p_t^{da}$  is presented in figure 5.1 before and after the recalibration is applied to the data and the simulation. In addition the ratio  $p_t^{bal}(\text{DATA})/p_t^{bal}(\text{MC})$  is shown. After applying the final jet calibration the  $p_t$  balance becomes flatter for the data and it is shifted from  $\sim 0.9$  to values close to unity. Moreover, the agreement between data and Monte Carlo is improved. The difference is smaller than 2% over the whole  $p_t$  range.

In figure 5.2 the same quantities are displayed as a function of the hadronic angle  $\gamma_{had}$ . In the central and backward region agreement between data and simulation before calibration is good although the response for hadrons is low. In the more forward region for angles  $\gamma_{had} < 30^\circ$  the jet energy is typically reconstructed too low. With the final calibration the data is well described by the Monte Carlo simulation even in the forward region and an agreement within  $\pm 2\%$  is found.

In general the recalibration of the jets improves the description of the data by the simulation. The disagreement is found to be smaller than 2% in all regions of  $p_t$  and  $\gamma_{had}$ . An uncertainty of the hadronic energy scale of  $\pm 2\%$  is used as estimate of the systematic error of the absolute hadronic energy scale.

NC DIS Dijet Sample
$p_t^{el} > 10 \text{ GeV}$
$Q_{el}^2 > 100 \text{ GeV}^2$
$y_{el} > 0.15$
$40 \text{ GeV} < E - P_z < 70 \text{ GeV}$
$ z_{vertex}  < 35 \text{ cm}$
Number of jets $\geq 2$

**Table 5.1:** The selection criteria for the NC DIS dijet event sample.

## 5.4 Trigger Selection and Efficiency

The H1 trigger system was introduced in section 3.1.3. The subtriggers chosen for this analysis are S67, S71, S75 and S77. S67 and S75 are based on energy thresholds in the electromagnetic and hadronic part of the LAr calorimeter. Although these triggers are optimised for electrons they also respond to high energetic jets. S77 is a missing transverse energy trigger.

S71 was originally designed as a charged current trigger [56]. On L1 it is based on a coincidence of an energy deposit in one of the LAr *big towers* and a *big ray* (CIP/COP track candidate). The track candidate is verified by the LAr big tower if the energy exceeds 1 GeV. On L2 a validation of this subtrigger is performed via the selection of *big towers* unbalanced in  $\phi$  in the central part of the calorimeter ( $25^\circ \lesssim \theta \lesssim 145^\circ$ ).

In addition all these triggers require vertex, timing and veto conditions.

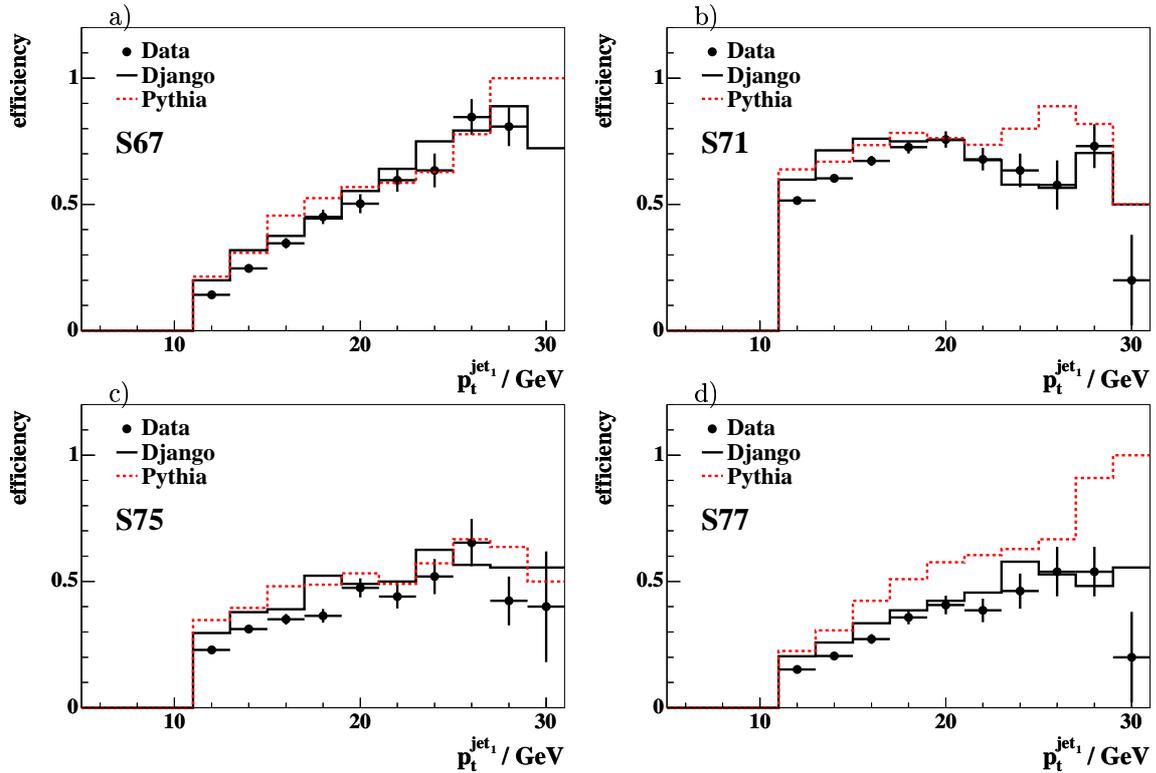
The efficiency of the above subtriggers is determined by selecting events which were triggered by an independent subtrigger, the so-called monitor trigger. The monitor trigger must not contain any of the trigger elements that the analysis subtriggers are built of. Since no such subtrigger for high transverse momentum jets in photoproduction exists in the data-taking period considered in this analysis, a trigger for DIS events is used. Based on a data to Monte Carlo simulation comparison for the DIS events, correction factors are determined which are also applied to the photoproduction sample.

The subtrigger S3 requires only a certain energy deposit in the electromagnetic SpaCal. In order to use it, a data sample without the cut against electrons in the SpaCal is selected. The trigger test sample thus contains mainly events with an electron identified in the SpaCal and two high transverse momentum jets. Note that in order to make a comparison with an available DJANGO Monte Carlo simulation, an additional cut  $Q^2 > 10 \text{ GeV}^2$  is applied to the trigger test sample.

The trigger efficiency is now given by the fraction of events that were triggered by both the signal and the monitor trigger ( $N_{signal \wedge monitor}$ ) relative to the events that were triggered by the monitor trigger alone  $N_{monitor}$ :

$$\varepsilon^{trigger} = \frac{N_{signal \wedge monitor}}{N_{monitor}} \quad (5.2)$$

The efficiency of the various subtriggers as a function of the transverse momentum of the



**Figure 5.3:** The trigger efficiency for the individual signal triggers as a function of the transverse momentum of the leading jet before correction.

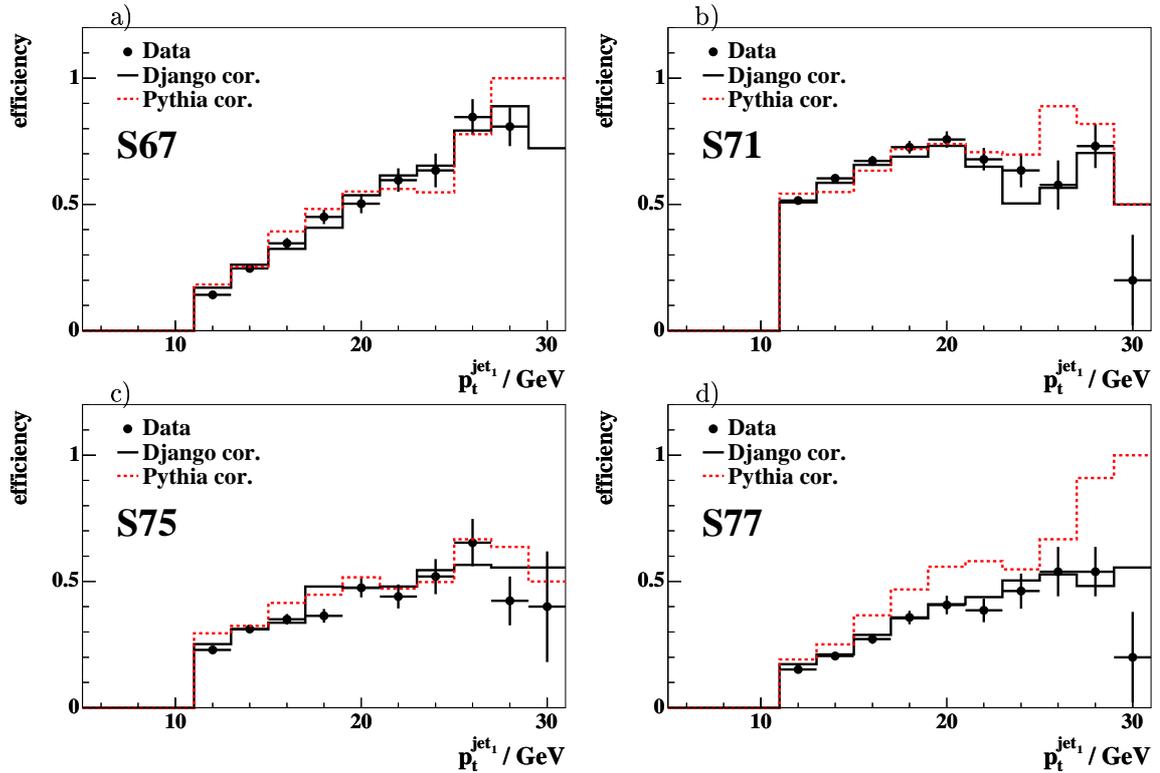
leading jet is shown in figure 5.3. The data is compared to two different simulations. While the efficiency calculated by DJANGO is in agreement for all subtriggers with the data for values of  $p_t \geq 20 \text{ GeV}$ , PYTHIA is generally worse in describing the data. This might be expected since PYTHIA is generated to describe photoproduction events and is not able to describe the kinematics of the trigger test sample. For  $p_t < 20 \text{ GeV}$  the trigger efficiency found in the data is generally lower than in the simulations.

In order to keep these events even at low transverse momenta, a correction to the Monte Carlo simulation is applied. More than 85% of the events that pass the dijet selection cuts (cf. section 5.2) are triggered by subtrigger S71. Weight factors for the Monte Carlo simulation are obtained by adjusting DJANGO to the data for  $p_t < 20 \text{ GeV}$  based on S71 alone. In addition, the trigger efficiency is lowered by 4% in the Monte Carlo simulation to account for inefficiencies of the other subtriggers (S67, S75, S77).

The effect of this procedure can be seen in figures 5.4 and 5.5. The corrected Monte Carlo trigger efficiencies shown in figure 5.4 as a function of the transverse momentum of the leading jet agree well with the data for all subtriggers apart from S77, where PYTHIA is still off. S77 alone triggers only 2% of the selected dijet events.

The trigger efficiency of the combined subtriggers is presented in figure 5.5 as a function of  $p_t^{\text{jet}_1}$  and  $\theta^{\text{jet}_1}$ . It reaches 100% for events with transverse momentum larger than 20 GeV and the dependence on  $\theta^{\text{jet}_1}$  is small. A similarly good description of the trigger efficiency is found for the second leading jet in the event (not shown).

Although the kinematics of the trigger test sample and the analysis sample might be different, the weight factors obtained with DJANGO are also applied to PYTHIA in the



**Figure 5.4:** The trigger efficiency for the individual signal triggers as a function of the transverse momentum of the leading jet after correction.

analysis of dijet photoproduction events. This procedure yields satisfactory results as can be seen by the good description of the overall trigger efficiency by PYTHIA in figure 5.5. A systematic error of 5% is attributed to the remaining differences between data and the simulation and for possible effects in the photoproduction region.

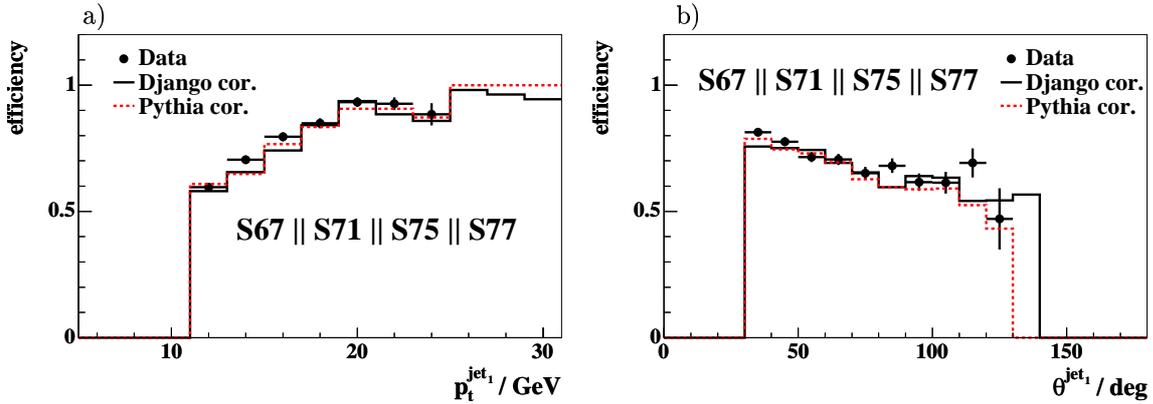
#### L4 Trigger Selection

The objects used in the offline event selection differ slightly from the objects the event reconstruction and classification performed on the fourth trigger level L4 is based on. Furthermore, the calorimeter clusters and drift chamber tracks used in this analysis are improved by offline calibration and by including CST hit information in the track fit.

At L4, events are classified into one of 15 event selection classes. A classification and subsequent selection of the events requires that the events either have a sufficient hard scale (e.g.  $E_t$ ,  $Q^2$ ,  $E_t^{\text{miss}}$ ) or are recognized by one of the exclusive final state finders. Otherwise the events are assigned to class 3 ("soft physics") in which only a fraction of classified events is accepted.

The majority of the selected dijet events (92%) is found in class 7 ("high  $E_T$  jets") which requires a minimum transverse energy of 10 GeV for at least one jet reconstructed in the LAr calorimeter. A cone jet algorithm is applied to the energy deposits in the calorimeter clusters.

Most of the events which are not found by class 7 are assigned to class 4 ("high  $Q^2$ ") or class 8 ("high  $E_t$  central"). Only a very small fraction of 0.08% of the dijet events



**Figure 5.5:** The trigger efficiency for the combination of all subtriggers as a function of the transverse momentum (a) and the polar angle (b) of the leading jet after correction.

are classified as soft physics, receiving large weight factors between 10 and 60. The mean weight for these events is 32 corresponding to  $\sim 3\%$  of the total event yield.

It has been checked that the Monte Carlo simulation gives a reasonable description of the L4 classification. Hence, in the final selection all class 3 exclusive events are removed in both data and Monte Carlo simulation such that no weighted events are left in the data sample.

## 5.5 Background Studies

At this stage of the event selection procedure events from non  $ep$ -interactions as well as non-photoproduction events could still spoil the event sample. Both background sources are investigated in this section.

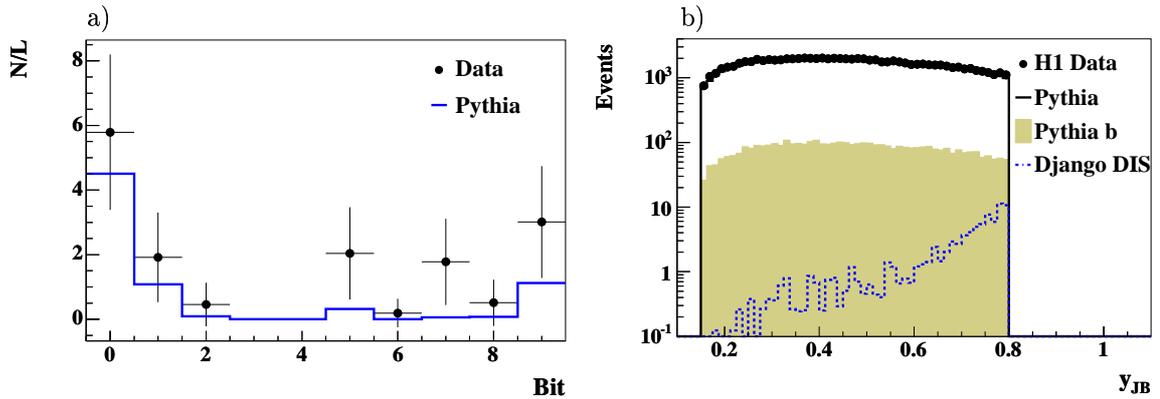
### Non $ep$ -background finders

The so-called QBGF<sub>MAR</sub> background finders [82, 83] are designed to search for signatures from events with muons which do not originate from  $ep$ -collisions. The package consists of several independent algorithms that look for different signatures of *cosmic ray muon*<sup>2</sup> and *halo muon*<sup>3</sup> events. In figure 5.6 a) the luminosity normalised number of events is shown which have been identified as background by the QBGF<sub>MAR</sub> finders as a function of the bit number corresponding to the individual finder. All cuts previously described in this chapter have been applied before. Only the first 10 algorithms of the QBGF<sub>MAR</sub> package are considered and displayed. These are reported to be safe in the sense that they ensure minimal inefficiency for  $ep$ -physics [83].

As can be seen in figure 5.6 a) only a few events per inverse picobarn are removed as background in both data and Monte Carlo. From this we conclude that any remaining non  $ep$ -background in the data sample is negligibly small compared with a total yield of  $\approx 1525$  events/ $\text{pb}^{-1}$  in the dijet photoproduction sample.

<sup>2</sup>Cosmic ray muons are produced in the earth's atmosphere by highly energetic cosmic ray particles.

<sup>3</sup>Halo muons originate from interactions of the proton beam with the beam pipe wall or remaining gas molecules inside the beam pipe.



**Figure 5.6:** a) First 10 bits of the QBGMAR background finder for the final dijet photoproduction selection. The luminosity normalised number of events is shown, which have been identified as halo muon background (bits 0-4) and cosmic muon background (bits 5-9). b) Distribution of the inelasticity  $y_{JB}$  for the final dijet photoproduction selection. The predicted contribution from remaining DIS events in the dijet photoproduction sample is shown as a dashed line.

## DIS background

As pointed out in section 5.2 the main handle against DIS background is the veto against electrons in the LAr and SpaCal calorimeters and the upper cut on  $y_{JB}$ . However, due to inefficiencies of the electron finder algorithm, the scattered electron can remain unidentified. In this case it is treated as part of the HFS and it may fake a jet or could be part of it.

In previous dijet analyses in photoproduction at H1 it turned out that the DIS background coming from events where the scattered electron has been misidentified as a high transverse momentum jet is not negligible [80, 84]. Recently a new track based electron finder algorithm has become available [85] which improves the electron finding efficiency substantially. At the same time the fake rate is reduced significantly. Figure 5.6 b) illustrates that according to the simulation the remaining background from DIS events in the final dijet sample is negligible. The distribution of the inelasticity  $y_{JB}$  is displayed and the data is compared to the PYTHIA and DJANGO simulations, the latter describing the DIS background. In this figure the predictions from PYTHIA are normalised with the scale factors obtained by the fit described in section 6.3. The DJANGO Monte Carlo simulation is normalised according to the generated luminosity.

It can be seen that the contamination with DIS events becomes larger for large  $y_{JB}$ , but even at the highest inelasticities the contribution from DIS background is less than 1% for the total dijet sample. Comparing with the measured beauty fraction of the total sample alone which is indicated by the shaded area, the predicted contribution of flavour inclusive DIS events by DJANGO is much lower than the beauty photoproduction fraction everywhere in  $y_{JB}$ .

## 5.6 Control Measurements

In this section the properties of the selected dijet photoproduction sample are studied and compared to Monte Carlo predictions using the PYTHIA and CASCADE generators. Flavour inclusive distributions are shown. The decomposition into contributions from

Dijet Event Selection
Good, medium runs. LAr, SpaCal, CST,CJC1, CJC2, CIP, COP, Lumi, ToF HV on
Subtrigger: S67, S71, S75, S77
$ z_{\text{Vertex}}  < 20$ cm non <i>ep</i> -background finder (bits 0-9 not fired)
no identified scattered electron in LAr or SpaCal $0.15 < y_{\text{JB}} < 0.8$
Number of jets $\geq 2$ $p_t^{\text{jet1}} > 11$ GeV $p_t^{\text{jet2}} > 8$ GeV $-0.9 < \eta^{\text{jet}} < 1.3$

**Table 5.2:** The final dijet selection criteria.

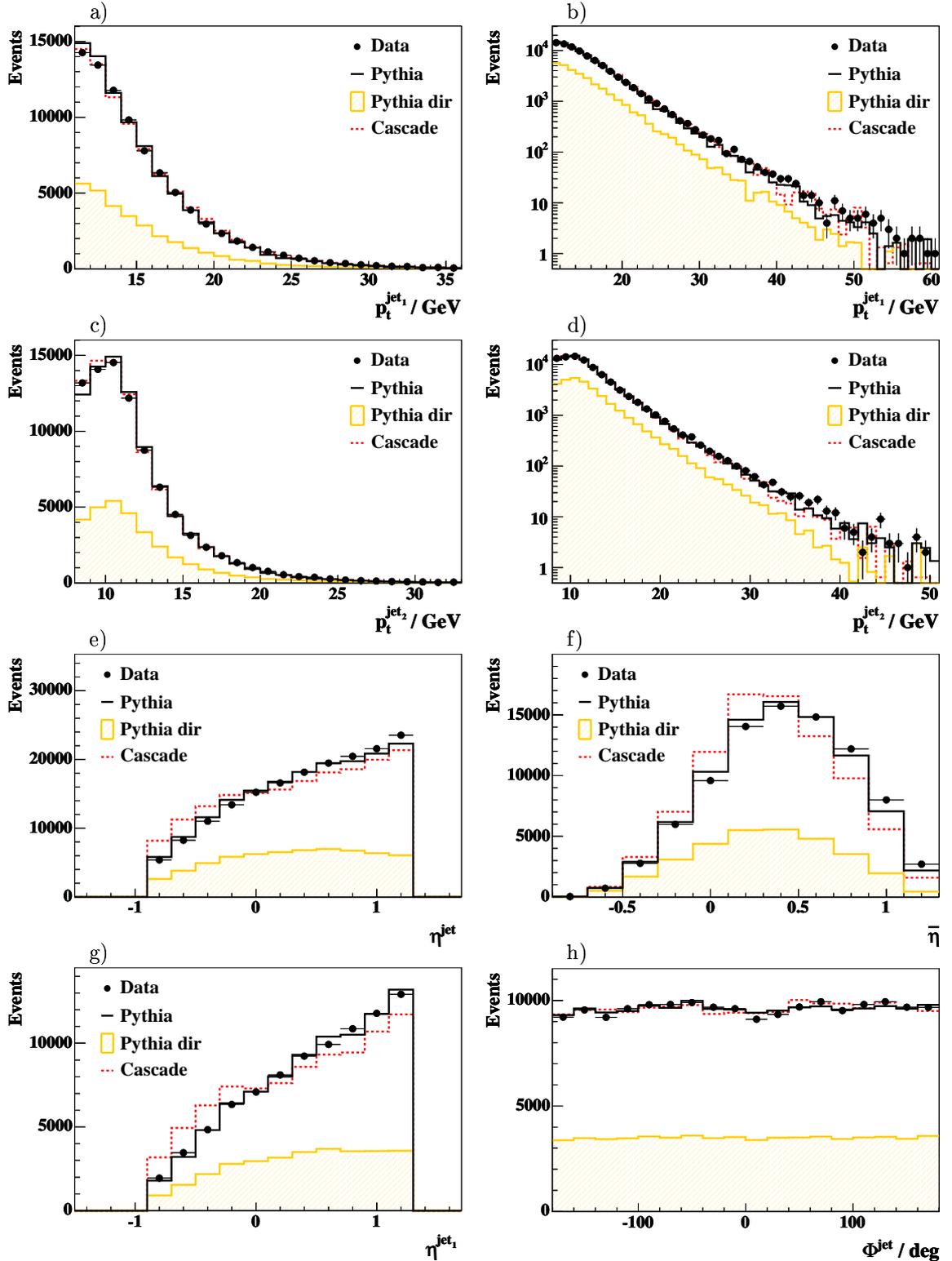
light quark, charm quark and beauty quark events will be discussed later in chapter 6. The normalisation of both Monte Carlo simulations is chosen to match the data in this section.

The full set of selection criteria is summarised in table 5.2. After applying these cuts based on the final hadronic calibration described in section 5.3 the total number of selected events is 86356.

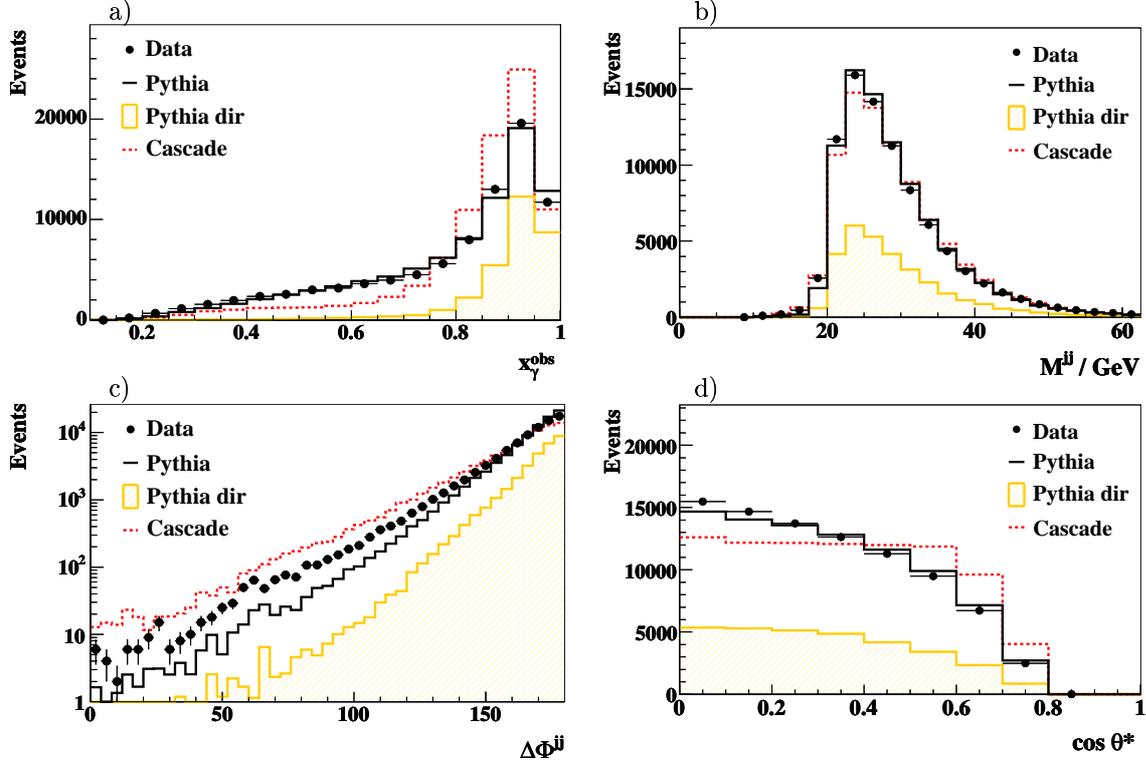
In figures 5.7 a)-d) the transverse momentum distributions of the individual jets are presented. The data are well described by both Monte Carlo simulations over 4 orders of magnitude. Jets are found with transverse momenta up to 60 GeV.

Figures 5.7 e)-h) show the distributions of the jet angular variables. The pseudo-rapidity of the two jets is presented in figure e). Most of the jets are found in the forward region, i.e. at positive  $\eta$ . PYTHIA slightly underestimates the number of jets at high  $\eta^{\text{jet}}$  and CASCADE is seen to be flatter than the data over the whole range of  $\eta$ . According to PYTHIA, events originating from resolved interactions are predominantly found in the forward region as illustrated in figures 5.7 f) and g) which show the mean pseudo-rapidity and the pseudo-rapidity of the leading jet, respectively. The azimuthal angle distribution shown in figure 5.7 h) is almost flat indicating full detector efficiency in azimuth.

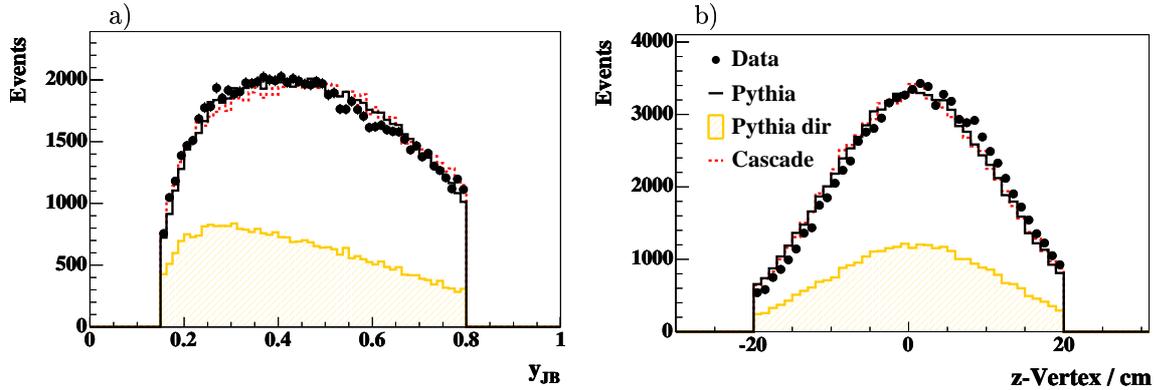
Further dijet distributions, calculated from both jets, are presented in figure 5.8. The longitudinal photon momentum fraction  $x_\gamma^{\text{obs}}$  is displayed in figure 5.8 a). The data covers  $x_\gamma^{\text{obs}}$  values between 0.1 and 1. PYTHIA can reproduce the data very well. Only at very high  $x_\gamma^{\text{obs}}$  it has slight difficulties in predicting the trend of the data. According to the PYTHIA simulation resolved events are mandatory to describe the shape of the data in particular in the region  $x_\gamma^{\text{obs}} < 0.85$ . In CASCADE resolved photon processes are not simulated separately but a hadronic photon component is partially included indirectly by



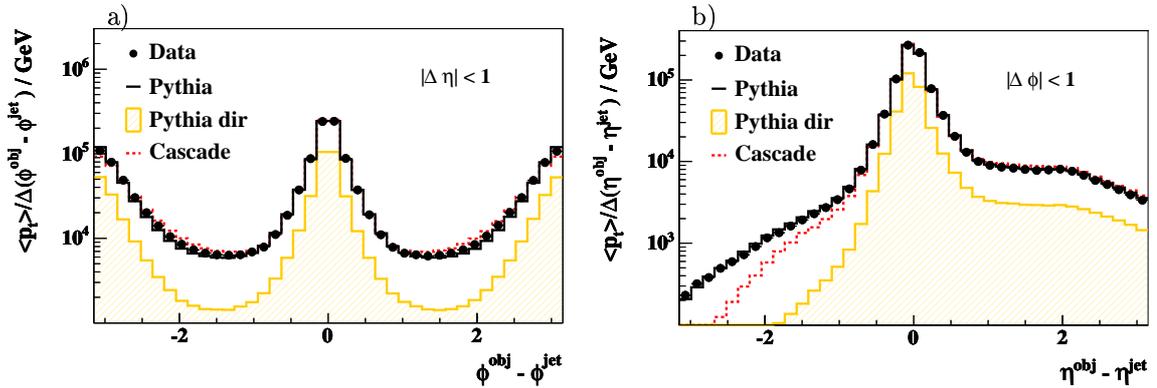
**Figure 5.7:** Distributions of the two leading jets in the selected dijet sample. The data are compared to the PYTHIA (full line) and CASCADE (dotted line) simulation for the distributions of a-b) the transverse momentum  $p_t^{jet1}$  of the leading jet, c-d) the transverse momentum  $p_t^{jet2}$  of the second jet, e) the pseudo-rapidity of both jets  $\eta^{jet}$ , f) the mean pseudo-rapidity  $\bar{\eta}$  of the two jets, g) the pseudo-rapidity of the leading jet  $\eta^{jet1}$  and h) the azimuthal angle of both jets  $\Phi^{jet}$ . The Monte Carlo distributions are normalised to the number of events in the data.



**Figure 5.8:** Distributions of the two leading jets in the selected dijet sample. The data are compared to the PYTHIA (full line) and CASCADE (dotted line) simulation for the distributions of a) the observable  $x_{\gamma}^{obs}$ , b) the invariant dijet mass  $M^{jj}$ , c) the azimuthal difference  $(\Delta\phi)^{jj}$  of the two jets and d) the observable  $\cos\theta^*$ . The Monte Carlo distributions are normalised to the number of events in the data.



**Figure 5.9:** Comparison of the dijet data with PYTHIA and CASCADE. Shown are event-normalised distributions of a) the inelasticity  $y_{JB}$  and b) the  $z$  position of the vertex.



**Figure 5.10:** The transverse energy flow around the jet axis as a function of the distance to the jet axis in azimuthal angles  $\phi^{obj} - \phi^{jet}$  a) and as a function of the distance in pseudo-rapidity  $\eta^{obj} - \eta^{jet}$  (b).

the concept of  $k_t$  factorisation (cf. section 2.2.1). However, CASCADe fails to describe the shape of the  $x_\gamma^{obs}$  distribution in the kinematic range of this analysis.

The distribution of the invariant dijet mass, shown in figure 5.8 b), is well reproduced by PYTHIA. CASCADe underestimates the data mainly at small values of  $M^{jj}$ .

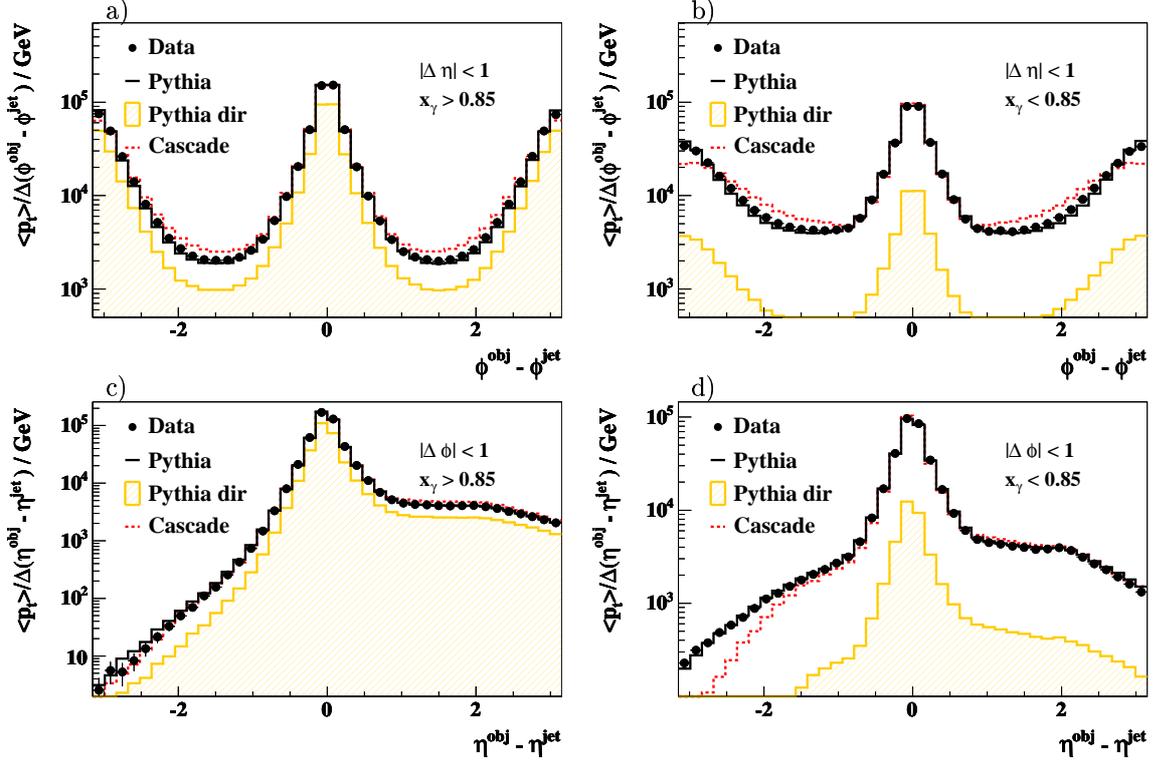
The distributions of the azimuthal difference  $(\Delta\phi)^{jj}$  is presented in figure 5.8 c). Both Monte Carlo simulations are not able to reproduce the characteristics of the data. Whereas for high values of  $(\Delta\phi)^{jj} > 150^\circ$  the description is still satisfactory, it becomes insufficient for lower values.

The observable  $\cos\theta^*$  is shown in figure 5.8 d). The data shows a slightly steeper behaviour than the PYTHIA Monte Carlo simulation. The CASCADe simulation fails to describe the shape of the distribution over the whole range of  $\cos\theta^*$ . Note, that the observed decrease with increasing  $\cos\theta^*$  is not an effect of the QCD matrix element, it is due to the jet transverse momentum and pseudo-rapidity cuts (cf. discussion in appendix C).

Figure 5.9 shows quantities which are not calculated from jets. The inelasticity  $y_{JB}$  is sufficiently modelled by both generators with a pronounced fraction of resolved events in the region of high  $y$  according to PYTHIA. The z-Vertex distribution shown in figure 5.9 b) reveals a slight shift of the data with respect to both simulations. Since the uncertainties on the acceptance corrections introduced by the cut on the vertex position are found to be small, no correction of the Monte Carlo simulations to the data is performed. The slightly unsatisfactory description of the z position of the vertex has no influence on the measurement because the impact parameter and the decay length are measured in the transverse plane only.

## 5.7 Jet Structure

The measurement of the transverse energy flow around the jet axis gives additional insight into the properties of dijet events. The energy flow is illustrated by so-called jet profiles where the distance in  $\eta$  and  $\phi$  between the objects of the hadronic final state and the jet axis weighted by the transverse energy of each object is considered. With this definition the energy flow is sensitive to effects from possible soft underlying events [86]. In PYTHIA, multiple interactions between the proton and the resolved photon remnants are used as a



**Figure 5.11:** The transverse energy flow around the jet axis as a function of the distance to the jet axis in azimuthal angles  $\phi^{obj} - \phi^{jet}$  (a-b) and as a function of the distance in pseudo-rapidity  $\eta^{obj} - \eta^{jet}$  (c-d) for different  $x_\gamma^{obs}$  regions.

way to include effects of underlying events.

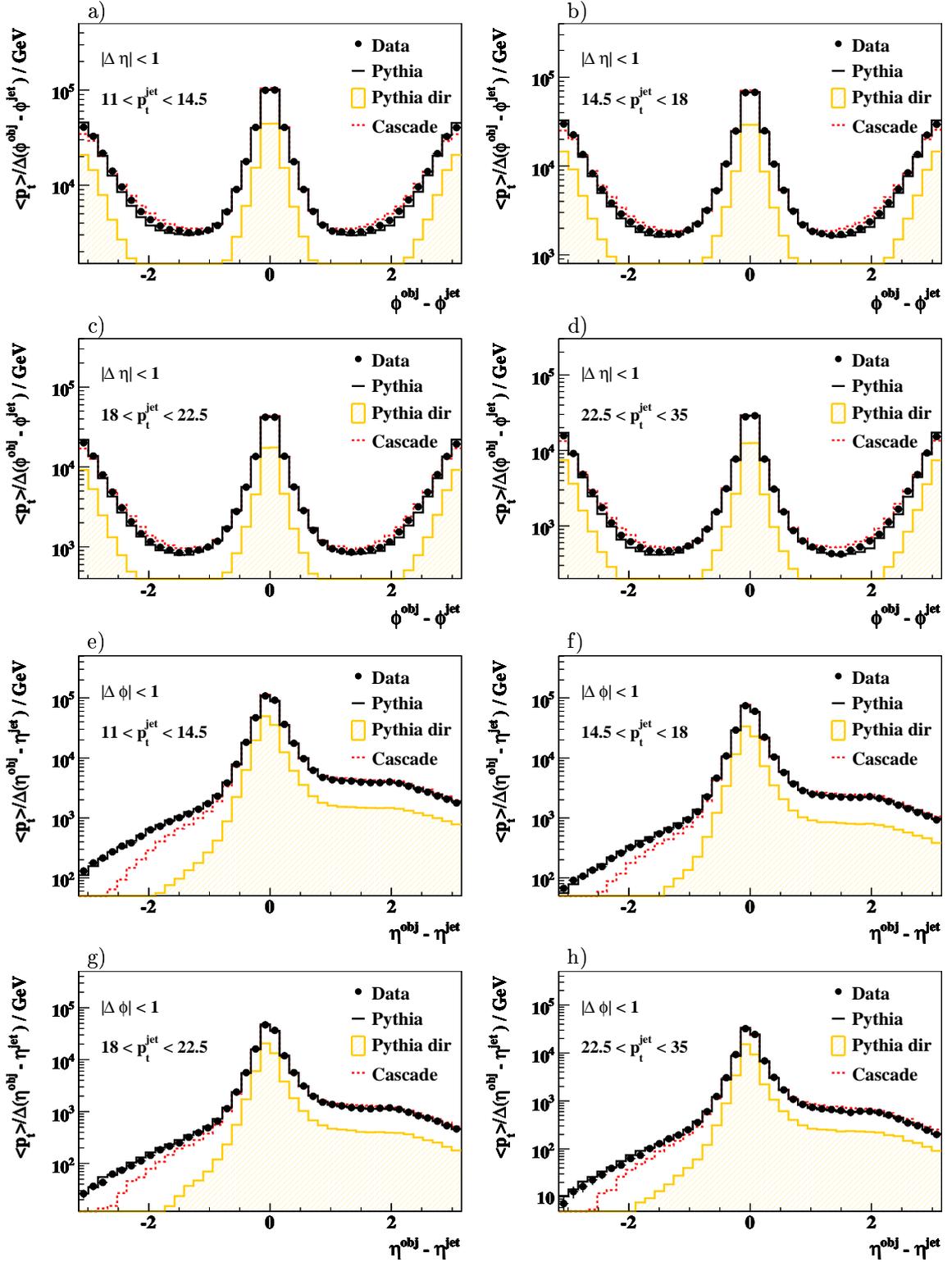
In figure 5.10 a) the transverse energy flow as a function of  $\phi^{obj} - \phi^{jet}$  between the  $\phi$  positions of the objects which contribute to the jet and the  $\phi$  position of the jet axis is shown. The objects are selected in a region  $|\Delta\eta| = |\eta^{obj} - \eta^{jet}| < 1$ . The peak close to zero originates from the particles of the jet itself. Due to the presence of the second jet the energy flow increases towards  $\pm\pi$ .

In a similar way figure 5.10 b) shows the transverse energy flow with respect to the  $\eta$  position of the jet axis. In this case the objects are restricted to the region  $|\Delta\phi| = |\phi^{obj} - \phi^{jet}| < 1$ . The asymmetry in the distribution, showing an enhancement for positive  $\Delta\eta$ , reflects the higher energy flow in the direction of the proton remnant. This may be attributed to multiple interactions, e.g. interactions between the proton remnant and the jets originating from the hard subprocess.

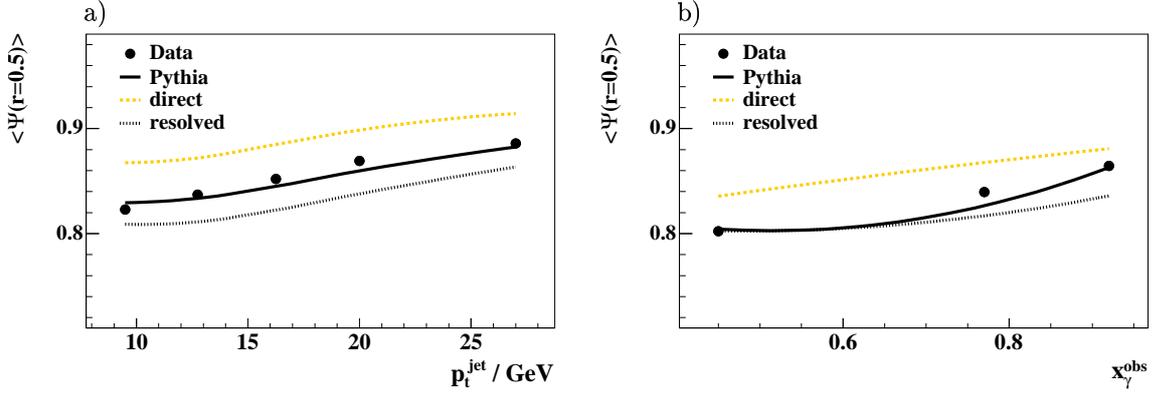
In more detail, figures 5.11 and 5.12 show the energy flow as a function of  $\Delta\phi$  and  $\Delta\eta$  for different regions of  $x_\gamma^{obs}$  and  $p_t^{jet1}$ .

CASCADE shows a clear disagreement with the data which is most prominent in the transverse region of  $|\phi^{obj} - \phi^{jet}| \sim \pi/2$  (figure 5.11 a and b) and in the region  $\eta^{obj} - \eta^{jet} < 0$  (figure 5.11 d). In the latter region, PYTHIA enhances the activity of multiple interactions through interactions with the photon remnant.

For PYTHIA generally a good agreement is observed between the measured and simulated jet profiles meaning that phenomena related to multiple interactions and soft underlying events are well treated within the PYTHIA Monte Carlo simulation.



**Figure 5.12:** The transverse energy flow around the jet axis as a function of the distance to the jet axis in azimuthal angles  $\phi^{\text{obj}} - \phi^{\text{jet}}$  (a-d) and as a function of the distance in pseudo-rapidity  $\eta^{\text{obj}} - \eta^{\text{jet}}$  (e-h) for different  $p_t^{\text{jet}}$  regions.



**Figure 5.13:** Mean integrated jet shape at fixed radius  $r = 0.5$  for the flavour inclusive dijet sample as functions of a)  $p_t^{\text{jet}}$  and b) the observable  $x_\gamma^{\text{obs}}$ . The statistical errors in both data and Monte Carlo simulation are negligible. Smooth curves connect the Monte Carlo predictions which are measured at the same points as the data.

These findings are confirmed by the measurement of the jet shape of the two leading jets. The jet shape  $\Psi(r)$  is defined as the fraction of the jet transverse momentum deposited in a cone of radius  $r$  around the jet axis relative to the transverse momentum of the jet deposited in a bigger cone with radius  $R = 1$ . This reads as

$$\Psi(r) = \frac{\sum_{i, r_i < r} p_{t,i}}{\sum_{i, r_i < R} p_{t,i}}, \quad (5.3)$$

where the sum runs over the final state objects of a jet. The mean integrated jet shape is then defined as

$$\langle \Psi(r) \rangle = \frac{1}{N_i} \sum_i \Psi(r), \quad (5.4)$$

with  $N_i$  being the total number of events and the mean being calculated for both investigated jets. The radius  $R = 1$  is chosen in analogy to the resolution parameter of the  $k_t$  clustering algorithm (cf. section 4.7). Some ambiguity arises from particles which belong to the jet, but have a distance to the jet axis  $r_i > R$ . It has been checked in [40] that the effect on the measured jet shapes from these particles is small.

In figure 5.13 the mean integrated jet shapes at fixed radius  $r = 0.5$  are shown as functions of  $p_t^{\text{jet}}$  and  $x_\gamma^{\text{obs}}$ . The flavour inclusive dijet sample is well described by the PYTHIA simulation. Distributions for direct and resolved processes are shown separately.

The result that multiple interaction models are needed to describe flavour inclusive jet shapes in photoproduction is demonstrated e.g. in [80, 87]. The whole concept of multiple interactions is somewhat controversial, especially its relevance for the heavy flavour sector [44]. In a recent study, jet shapes in charm photoproduction events are investigated at H1 [40] and compared to a flavour inclusive sample. While the latter sample can be described fairly well by the PYTHIA simulation including multiple interactions, there are indications that in the charm sample the impact from multiple interactions may be overestimated.

The discussion of this topic is continued in section 6.6 where jet shapes of charm and beauty enriched data samples are presented.

## 5.8 Data Unfolding

In the last sections it has been shown that the Monte Carlo simulations including detector effects give a reasonable description of the data. Next-to-leading order calculations are only available on parton level and have to be corrected to the hadron level. For a meaningful comparison of the data to theoretical calculations, detector effects, such as limited resolution and efficiencies, are removed from the measurement.

The technique used to correct for detector and hadronisation effects is the so-called bin-by-bin method, which examines the distributions of a given observable at two different levels. The technique is valid if the correlations between bins of a given distribution are small. To ensure this, the bin sizes should be larger or of the same order as the resolution of the observables, such that migrations between bins are small. Resolution and migration aspects will be discussed in the following two sections.

The definition of the cross section and the corrections from parton to hadron level for the theory calculations are then explained in sections 5.8.3 and 5.8.4.

The bin-by-bin method is only usable if the simulation widely agrees with the data. Since PYTHIA is found to give a somewhat more consistent description of the dijet data than CASCADE, only investigations performed with the PYTHIA simulation are presented in the next sections.

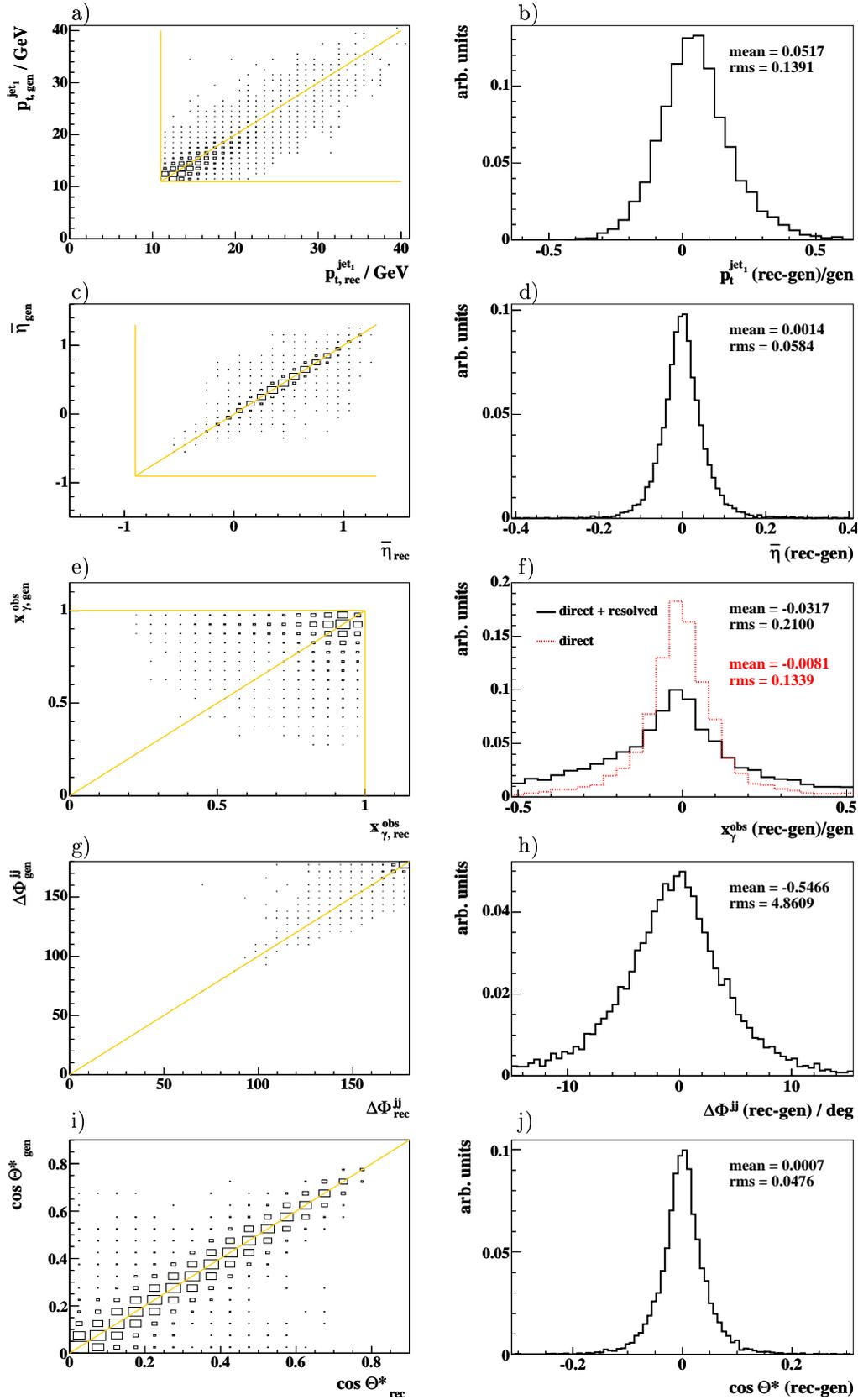
### 5.8.1 Resolution and Systematic Shifts

In this section the relationship between reconstructed and generated variables is studied. For the comparisons only events are used which fulfill all cuts on both detector and hadron level.

The correlations between detector and hadron level for the variables  $p_t^{jet_1}$ ,  $\bar{\eta}$ ,  $x_\gamma^{obs}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  are shown in figure 5.14 for the PYTHIA Monte Carlo simulation of beauty events. The correlations are found to be similar for the charm and light quark samples. The matching between detector and hadron level is not always perfect as can be seen e.g. for the observable  $x_\gamma^{obs}$ . This is mainly due to events with more than two jets where the two jets on detector level are not the same as on hadron level. The best correlation is seen for the mean pseudo-rapidity of the two jets  $\bar{\eta}$ . However, the other observables are generally also reasonably correlated.

The resolution is a measure of the relative deviation of the reconstructed variable from the generated variable. Figure 5.14 b) illustrates the resolution  $(p_{t,rec}^{jet_1} - p_{t,gen}^{jet_1})/p_{t,gen}^{jet_1}$  of the transverse momentum of the leading jet again only for the beauty PYTHIA Monte Carlo simulation. On average it is reconstructed  $\sim 5\%$  too high and has a resolution of  $\sim 14\%$ . In figure 5.14 d) the resolution of  $\bar{\eta}$  is shown. The resolution of the mean pseudo-rapidity is 0.06 units in  $\bar{\eta}$ .

The ability to reconstruct the variable  $x_\gamma^{obs}$  is illustrated in 5.14 f). Here the resolution is investigated separately for all production processes (full line) and direct processes only (dotted line). Both distributions peak slightly below zero with a mean value close to zero. However, a difference in resolution is observed. While for the full simulation including direct and resolved processes the resolution is found to be 21%, the direct processes only have a resolution of 13%. This is due to the more complex event structure of the resolved processes. For the other variables the resolution for direct and resolved events is comparable.



**Figure 5.14:** Correlations and resolutions of the observables  $p_t^{\text{jet}_1}$ ,  $\bar{\eta}$ ,  $x_\gamma^{\text{obs}}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  determined with the beauty PYTHIA Monte Carlo simulation. The *left column* shows the reconstructed versus the generated variables. Horizontal and vertical lines indicate the cuts applied. In the *right column* the resolutions are investigated. In f) the resolution obtained with beauty PYTHIA direct only is shown separately (dotted line). The distributions are normalised to one.

Finally the resolutions of  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  are shown in figures 5.14 h) and j). The difference of the jet's azimuthal angle can be measured with a resolution of  $\sim 5$  degrees. The resolution of the the centre-of-mass scattering angle is 0.05 units in  $\cos\theta^*$ . No substantial shifts are observed in the angular variables.

### 5.8.2 Purity and Stability

As a measure of the migrations between bins the quantities purity and stability are determined. They are defined as the ratio of the number of events generated and reconstructed in one bin and the number of events generated (stability) or reconstructed (purity) in this bin:

$$Purity(i) = \frac{N_{rec\&gen}(i)}{N_{rec}(i)}, \quad (5.5)$$

$$Stability(i) = \frac{N_{rec\&gen}(i)}{N_{gen}(i)}. \quad (5.6)$$

Following this definition the value of  $(1 - Purity(i))$  indicates the fraction of events migrating into the bin  $i$  while  $(1 - Stability(i))$  indicates how many events migrate out of bin  $i$ . In the ideal case of perfect resolution both purity and stability are equal to one.

The choice of the binning should ensure that all bins contain a reasonable number of events and that the bin sizes are larger than the obtained resolutions. Both, purity and stability are required to be larger than  $\sim 30\%$ .

Figures 5.15 and 5.16 show the purities and the stabilities for the variables  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $x_\gamma^{obs}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  for the chosen bins. They are calculated from the PYTHIA Monte Carlo simulation using events in the visible region of the measurement, which is defined by the cuts summarised in equation 5.7 for the generated level and in table 5.2 for the reconstructed level. The purities and stabilities are found to be similar for beauty (full points) and charm (open points) in all variables and are always 40% or larger.

### 5.8.3 Definition of the Cross Section

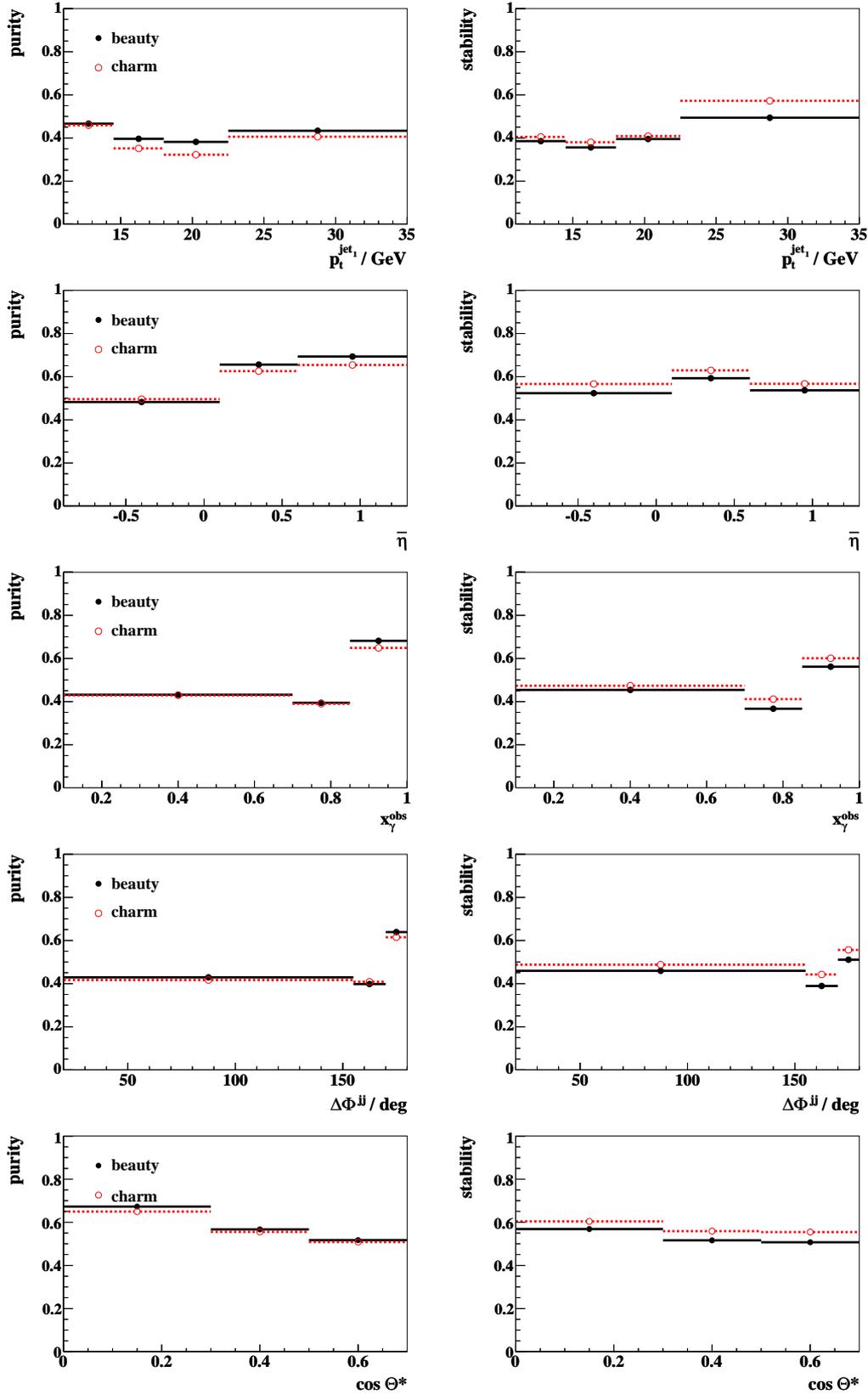
The cross sections is measured in the region

$$\begin{aligned} Q^2 < 1 \text{ GeV}^2 & & 0.15 < y < 0.8 \\ p_t^{jet1(2)} > 11(8) \text{ GeV} & & -0.9 < \eta^{jet1(2)} < 1.3 \end{aligned} \quad (5.7)$$

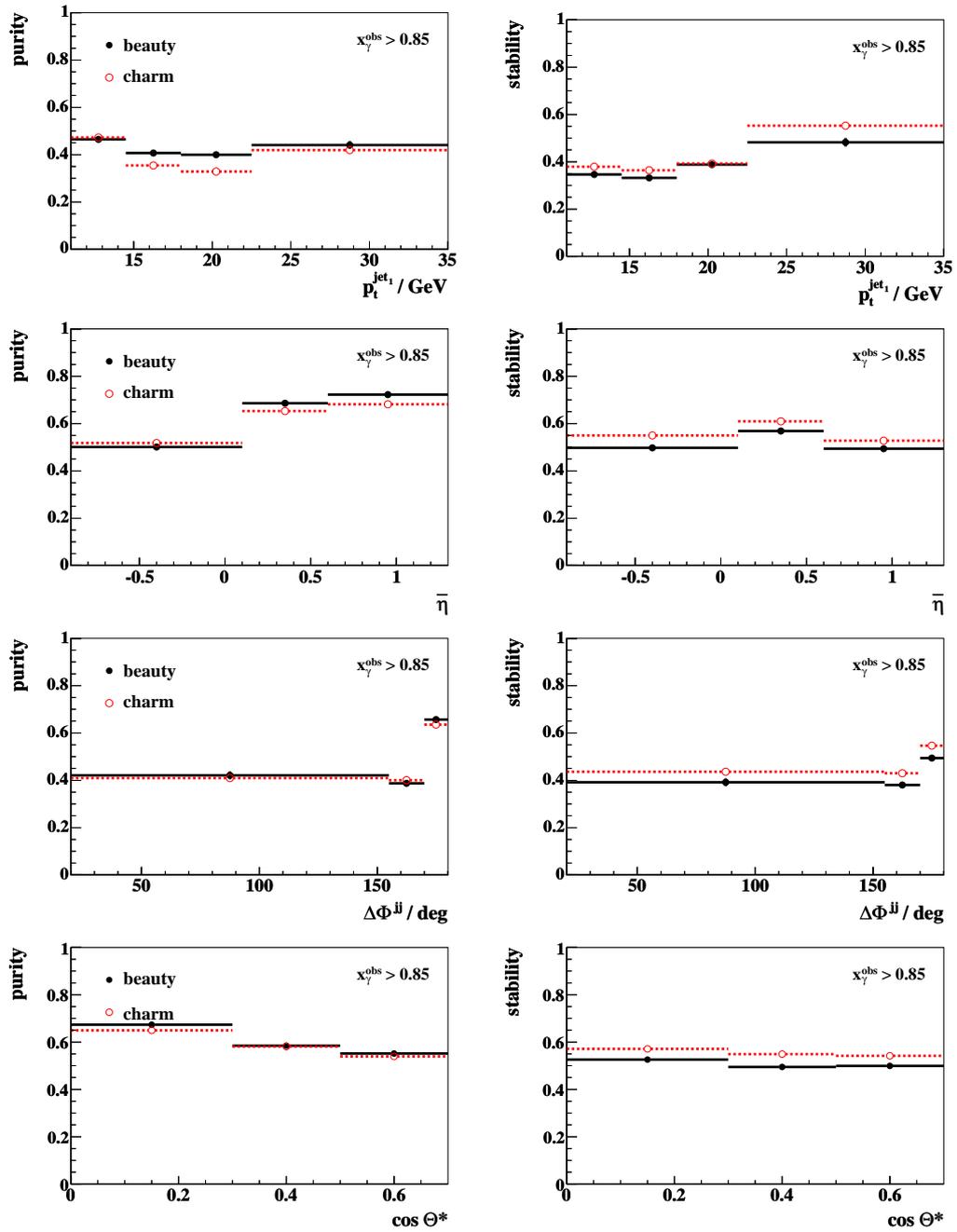
Most cuts are due to experimental reasons as discussed in this chapter. The dijet production cross section can be measured by counting the number of measured dijet events ( $N^{Data}$ ) produced in the visible range correcting for detector efficiencies and normalising to the corresponding integrated luminosity ( $\mathcal{L}_{Data}$ ) collected by the experiment

$$\sigma_{Data}^{meas}(ep \rightarrow ejjX) = \frac{N^{Data}}{\mathcal{L}_{Data} \cdot \varepsilon A}. \quad (5.8)$$

The factor  $\varepsilon A$  represents the correction for detector efficiencies and acceptances. This includes contributions depending on geometry and limited resolution of the detector as well



**Figure 5.15:** Purities and Stabilities for the chosen bins in  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $x_\gamma^{obs}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$ . The full (open) points show the values obtained with the beauty (charm) PYTHIA Monte Carlo simulation.



**Figure 5.16:** Purities and Stabilities for the chosen bins in  $p_t^{\text{jet}_1}$ ,  $\bar{\eta}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  in the region  $x_\gamma^{\text{obs}} > 0.85$ . The full (open) points show the values obtained with the beauty (charm) PYTHIA Monte Carlo simulation.

as on the selection cuts and the performance of the algorithms used for the reconstruction. The factor  $\varepsilon A$  is determined from the Monte Carlo simulation by dividing the number of reconstructed and generated events:

$$\varepsilon A = N^{rec}/N^{gen}. \quad (5.9)$$

For the Monte Carlo simulation the visible cross section is defined as

$$\sigma_{MC}^{vis}(ep \rightarrow ejjX) = \frac{N^{gen}}{\mathcal{L}_{MC}}, \quad (5.10)$$

with  $\mathcal{L}_{MC}$  being the generated Monte Carlo luminosity. Combining 5.8 with equations 5.9 and 5.10 one obtains

$$\sigma_{Data}^{meas}(ep \rightarrow ejjX) = \frac{N^{Data} \mathcal{L}_{MC}}{N^{rec} \mathcal{L}_{Data}} \cdot \sigma_{MC}^{vis}. \quad (5.11)$$

For the measurement of the charm and beauty cross sections the first factor is identified with the scale factors  $P_c$  and  $P_b$  which are obtained from a fit to the data as explained in the next chapter.

#### 5.8.4 Hadronisation Corrections

Theory calculations performed to next-to-leading order are available on parton level (cf. section 2.5.1). In order to compare the theory calculations with the data, correction factors from the parton to the hadron level are applied for each bin of the measurement. The correction factors are determined using the PYTHIA Monte Carlo event generator.

The hadronisation correction factors are given by the ratio of the cross sections obtained with jets as reconstructed from hadrons after the full event simulation and from jets as reconstructed from partons after parton showers but before hadronisation. The hadron level includes also effects from underlying events which is not the case on parton level. The jets at both the parton and the hadron level are reconstructed using the inclusive  $k_t$  jet-algorithm in the  $p_t$  recombination scheme.

For flavour inclusive jets the effects of underlying events and of the fragmentation are investigated separately in [88]. The total hadronisation corrections are assumed to be the sum of both contributions. Low momentum hadrons from the underlying event tend to increase the jet transverse momentum and thus the hadron level cross section at a given jet transverse momentum. The correction due to the underlying event is therefore always positive and for jet transverse momenta between 5 and 12 GeV it is found to be between +30% at lower pseudo-rapidities and up to +100% in the forward region. For larger transverse momenta ( $> 20$  GeV) the effect is reduced to about +10%. The correction factor due to the effect of the underlying event is found to be partially compensated by fragmentation effects, which generally reduce the cross section by  $-30\%$  in the lower transverse momentum region and by around 5% for larger transverse momenta.

For this analysis the overall hadronisation correction factors for the phase space given in 5.7 are  $-7\%$  and  $+1\%$  for charm and beauty, respectively. The bin-by-bin correction factors for the observables considered in this analysis are found to be less than  $\pm 10\%$  everywhere except in the bins  $0.7 < x_\gamma^{obs} < 0.85$  and  $0.85 < x_\gamma^{obs} < 1$  where the corrections are about

---

+35% and  $-15\%$  respectively. This is due to hadrons from the hadronisation process which are wrongly assigned to the two leading jets at hadron level. This hadronisation effect can reduce  $x_\gamma^{obs}$  drastically, leading to large migrations.

## CHAPTER 6

# QUARK FLAVOUR SEPARATION

---

The different quark flavours that contribute to the flavour inclusive cross section can be distinguished on the basis of differences in the lifetime spectra of the produced hadrons. As discussed in chapter 4 the uncertainty of the measurement of the impact parameter receives contributions from the position and resolution of the primary vertex, the intrinsic track resolution as well as distortions due to multiple scattering in the beam pipe and surrounding material. For a successful description of the data the Monte Carlo parameters for the beam spot size, the tracking resolution and detector material are adjusted to those observed in the data (cf. section 4.5).

In this analysis the fractions of events containing charm and beauty quarks are determined using the same method as in previous H1 measurements [5, 6]. The impact parameter of selected tracks is used to construct the significance distributions  $S_1$  and  $S_2$ . In section 6.2 this method is described in detail. In section 6.3 the fit procedure is outlined.

Alternative methods are used as a cross check. The Multi-Impact Parameter approach is presented in section 6.4, followed by a discussion of the explicit secondary vertex reconstruction using deterministic annealing in section 6.5. Based on the latter approach a heavy flavour enriched subsample is selected and described in the last section of this chapter.

### 6.1 Track Selection

For the track selection only events which have passed the dijet selection criteria (cf. table 5.2) are considered. The non vertex fitted CST tracks as described in section 4.2 are the basis of the track selection. Only those tracks are selected which can be associated to one of the two leading jets in the event. For technical reasons the direct association of the tracks to the jets via the jet algorithm is not easily possible, since updated track parameters as discussed in section 4.5 are used and these are not available for the jet algorithm. Therefore the track-jet matching is done using a cone condition, requiring selected tracks to be inside a cone of radius  $R = 1$  centred about the jet axis. The cone radius  $R$  is defined as

$$R = \sqrt{(\eta_{jet} - \eta_{track})^2 + (\phi_{jet} - \phi_{track})^2}, \quad (6.1)$$

with the pseudo-rapidities and azimuthal angles of the jet and track, respectively.

Track Selection	
Transverse momentum	$p_t > 0.5 \text{ GeV}$
Angular region	$30^\circ < \theta_{track} < 150^\circ$
CJC radial track length	$l_{track} > 10 \text{ cm}$
CJC track start radius	$R_{start} < 50 \text{ cm}$
CST hits	$N_{CST} \geq 2$
CJC-CST Link probability	$\mathcal{P}_{link} > 0.1$

**Table 6.1:** Track selection criteria.

Table 6.1 summarises the track selection criteria which have been developed to ensure a track reconstruction with a reliable precision while keeping enough statistics. Since at lower transverse momentum the impact parameter resolution is spoiled by multiple scattering effects, the transverse momentum of the tracks must exceed 500 MeV to ensure a reasonable resolution. The angular region corresponds to the CST acceptance. Although the jet axes are already required to lie in this central region, the additional track angular cut guarantees that each associated and selected track is well measured inside the CST.

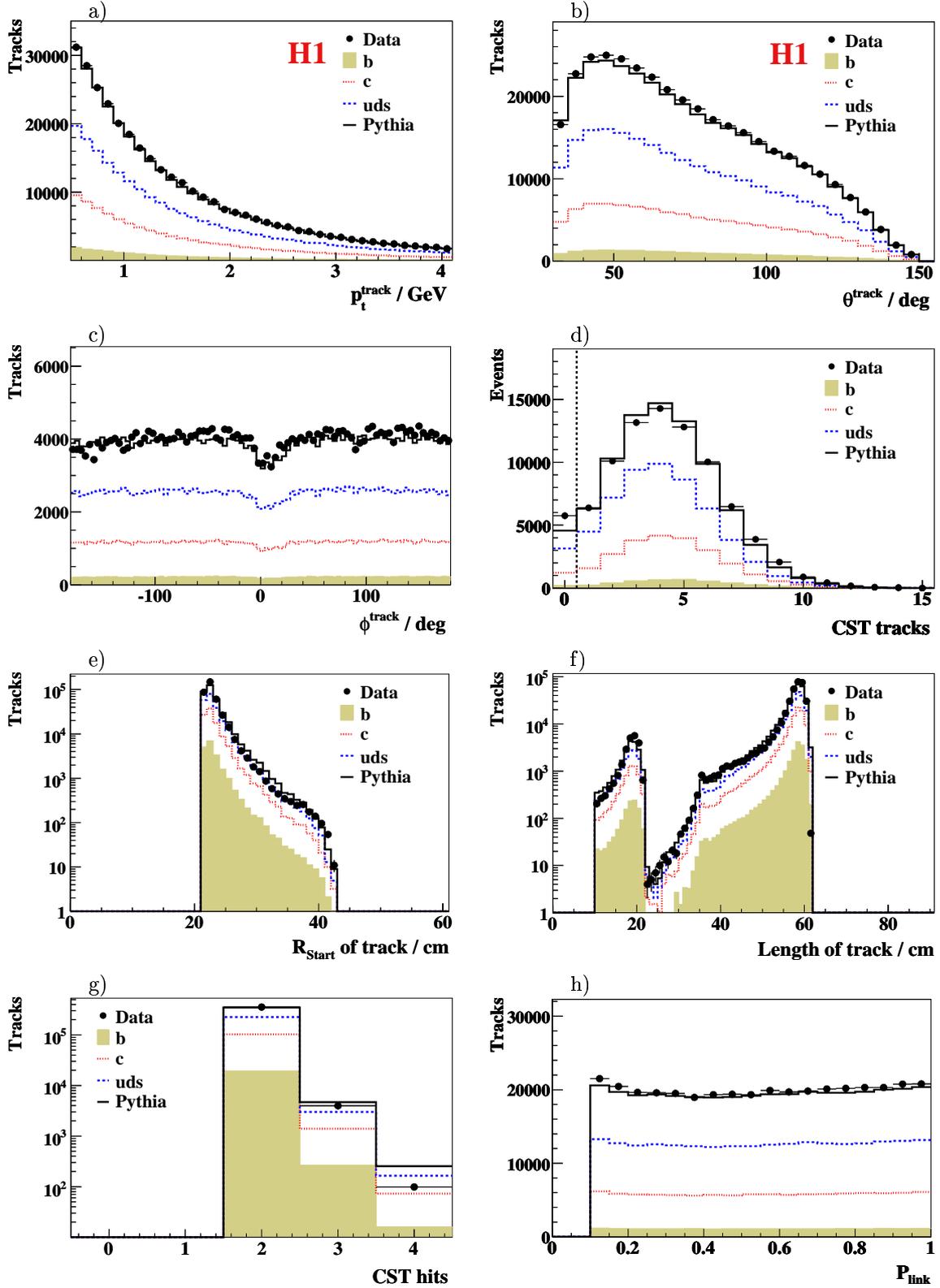
The radial track length is determined by the design of the jet chambers with a maximum track length of 64.1 cm given by the difference of the outer radius of CJC2 and the inner radius of CJC1. For a reasonable determination of the CJC track parameters the minimum radial track length is required to be 10 cm. To ensure a precise extrapolation of the CJC track to the CST, the start radius of the CJC track must be inside the CJC1, i.e.  $R_{start} < 50 \text{ cm}$ .

The cuts on the CST information provide the necessary quality of the track reconstruction close to the interaction region. At least two hits in the two layers of the CST are requested with a minimum CJC-CST track linking probability of 10%.

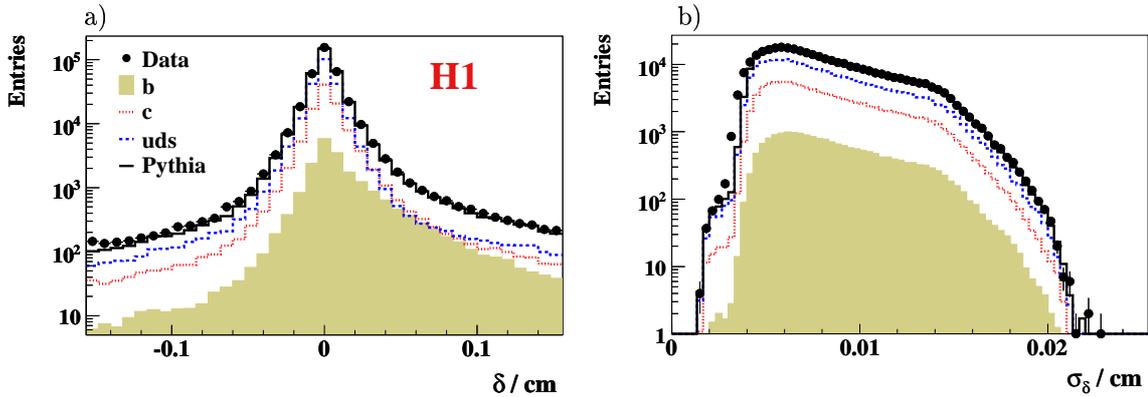
Figure 6.1 shows control distributions of the above discussed track parameters. All tracks which belong to one of the two leading jets, and fulfilling the track cuts according to table 6.1, enter these distributions. The data is compared to the expectation from the PYTHIA Monte Carlo simulation. The contributions from the various quark flavours are shown separately with relative normalisations according to the individual scale factors for each quark flavour as described in section 6.3.

The overall agreement between data and simulation is good, especially the track transverse momentum (figure 6.1 a) and the angular (figure 6.1 b) spectra are well reproduced by the Monte Carlo simulation. The small disagreement seen at small track angles  $\theta$  is due to the non-perfect description of the track efficiency in that region (cf. figure 4.8). The uncertainties arising from these differences are covered by the variations of the track efficiency as discussed in section 7.1. The small dip at  $\phi_{track} \approx 0^\circ$  (figure 6.1 c) is due to inefficient cells in the CJC in the data taking period considered here.

In the distribution of the number of selected CST tracks (figure 6.1d) a small shift between data and simulation is observed. The track multiplicity is seen to be somewhat higher in the data. However the agreement is still reasonable. The small deviations are mainly due to a non-perfect modeling of the multiplicities in light quark jets and have a negligible



**Figure 6.1:** Control distributions of the selected tracks. The expectation from the PYTHIA Monte Carlo simulation is included in the figure, showing the contributions from the various quark flavours after applying the scale factors obtained from the fit to the subtracted significance distributions of the data (see section 6.3).



**Figure 6.2:** Distributions of a) the reconstructed impact parameter  $\delta$  and b) its error  $\sigma_\delta$ . The expectation from the PYTHIA Monte Carlo simulation is included in the figure, showing the contributions from the various quark flavours after applying the scale factors obtained from the fit to the subtracted significance distributions of the data (see section 6.3).

effect on the measurement. The systematic error arising from uncertainties on the heavy quark multiplicities are discussed in section 7.1.

The number of events rejected by the above requirements is 5757. This corresponds to 7% of the dijet events, resulting in 80769 events remaining after the final track selection. The loss is dominantly due to the CST efficiency and the uncertainty is covered by the corresponding systematic error (see section 7.1).

## 6.2 Lifetime Tag Observables

The different quark flavours that contribute to the cross section can be distinguished on the basis of differences in the reconstructed impact parameter spectrum of selected CST tracks. The distribution of the reconstructed signed impact parameter  $\delta$  and its error  $\sigma_\delta$  is shown in figure 6.2.

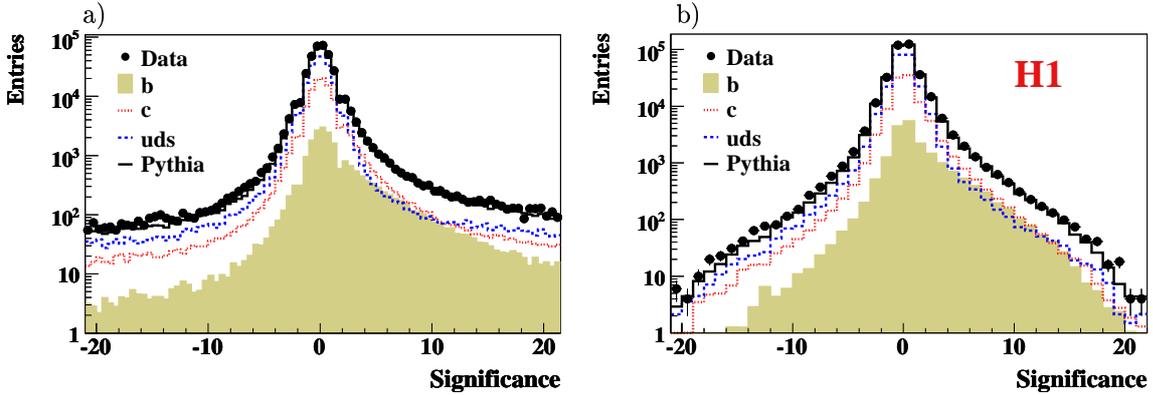
Due to the long lifetimes of charm and beauty flavoured hadrons the  $\delta$  distribution is asymmetric, the number of tracks with positive values exceeding the number of tracks with negative values. The simulation gives a reasonable description of the data. While the component that arises from light quarks is almost symmetric at low  $\delta$ , the  $c$  component has a moderate asymmetry and the  $b$  component shows a marked asymmetry. The differences are due to the different lifetimes of the produced hadrons. The asymmetry seen at  $|\delta| > 0.1$  cm is mainly due to decays of long lived strange particles such as  $K_S^0$ . In order to reduce the effects of the strange component, a cut of  $|\delta| < 0.1$  cm is imposed on all tracks used in the following. The systematic error attributed to the uncertainty of the strange component in the sample is discussed in section 7.1.

The distribution of  $\sigma_\delta$  receives contributions from the vertex and track errors and reaches values of  $220 \mu\text{m}$  at maximum. Its description by the simulation is reasonable.

The significance, defined as the ratio of the impact parameter to its error,

$$S_i = \frac{\delta_i}{\sigma_{\delta_i}}, \quad (6.2)$$

is shown in figure 6.3 for all selected tracks  $i$  without (a) and with the cut  $|\delta| < 0.1$  cm (b).



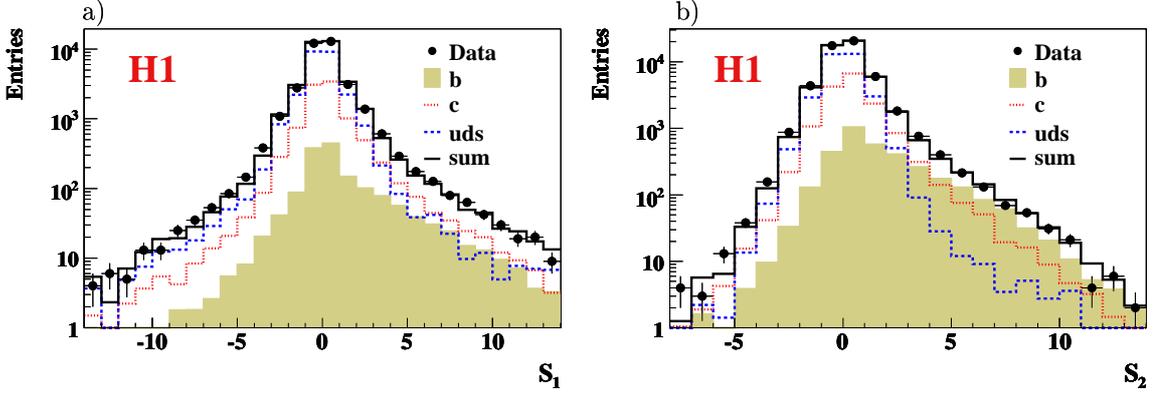
**Figure 6.3:** Distributions of a) the impact parameter significance for tracks and b) the impact parameter significance for tracks with  $|\delta| < 0.1$  cm. The expectation from the PYTHIA Monte Carlo simulation is included in the figure, showing the contributions from the various quark flavours after applying the scale factors obtained from the fit to the subtracted significance distributions of the data (see section 6.3).

The construction of the significance optimises the separation of the quark flavours, because the error reflects the influence of the specific decay topology and of multiple scattering on the impact parameter measurement. The simulation is observed to describe the data very well, apart from the tails of the distribution where a non perfect modelling of the resolution leads to some differences. The deviations in resolution between data and Monte Carlo are treated as a systematic error (cf. section 7.1).

In the distribution of the significance (figure 6.3 a), a small kink is observed at a significance of about two. This is an effect of the primary vertex reconstruction where all tracks compatible within two standard deviations with the run vertex are used. The tracks corresponding to the impact parameter significances shown in this figure are not excluded from the vertex fit. Thus, for tracks with small impact parameter significances, the track is included in the vertex fit and the primary vertex is reconstructed closer to the track resulting in artificially lower measured impact parameter significances. This effect is well modelled in the simulation and hence does not spoil the measurement.

A further gain in optimising the separation of the quark flavours is made by using different significance distributions for events with different track multiplicities. The first significance distribution  $S_1$  is defined for jets where exactly one selected track is linked to the jet, and is simply the significance of this track. The second significance distribution  $S_2$  is defined for jets with two or more associated tracks and is the significance of the track with the second highest absolute significance. For jets contributing to the distribution of  $S_2$  it is required that the track with the first and the track with the second highest absolute significance in the jet have the same sign of  $\delta$ . The choice of the second highest significance is made because in light quark events it is very unlikely that two tracks are produced at large significance due to resolution effects. The  $S_1$  and  $S_2$  distributions are shown in figures 6.4 a and b. While the distribution of  $S_1$  shows only small differences for the various quark flavours, the distribution of  $S_2$  gives a good separation power of light to heavy quarks. At moderate and large values of  $S_2$  the beauty contribution exceeds that from charm.

The fit procedure for the determination of the quark fractions is described in the following section. Although the  $S_2$  distribution provides better separation between charm and



**Figure 6.4:** Distributions of a) the significance  $S_1$  of tracks in jets with exactly one selected track and b) the significance  $S_2$  of the track with the second highest significance in jets with two or more selected tracks. Only tracks with an impact parameter  $|\delta| < 0.1$  cm are considered. The expectation from the PYTHIA Monte Carlo simulation is included in the figure, showing the contributions from the various quark flavours after applying the scale factors obtained from the fit to the subtracted significance distributions of the data (see section 6.3).

beauty, events with only one selected CST track ( $S_1$ ) are retained to improve the statistical precision of the fit.

### 6.3 Fit Procedure

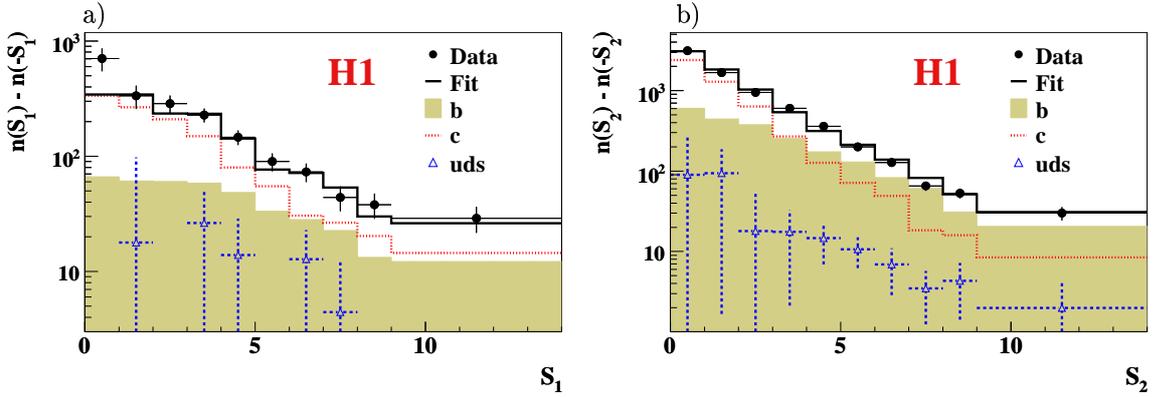
In order to substantially reduce the uncertainty due to the resolution of  $\delta$  and the light quark normalisation the negative bins in the  $S_1$  and  $S_2$  distributions are subtracted from the positive ones. In the tails of the  $S_1$  and  $S_2$  distributions the number of events in data and Monte Carlo simulation is small. Therefore, to retain the Gaussian limit in the fit, the bins with insufficient statistics ( $< 10$  events) are combined into one large bin in the subtracted significance distributions (cf. figure 6.5).

The  $b$ ,  $c$  and light quark fractions in the data are then extracted by simultaneously fitting the subtracted  $S_1$  and  $S_2$  distributions  $N_i^{data}$  and the total number of events  $N_{tot}^{data}$  before any CST track selection. The Monte Carlo  $b$ ,  $c$  and light quark distributions are used as shape templates and are allowed to be modified by the scale factors  $P_b$ ,  $P_c$  and  $P_l$ , respectively, such that

$$\chi^2 = \sum_i \frac{(N_i^{data} - P_b N_{bi}^{MC} - P_c N_{ci}^{MC} - P_l N_{li}^{MC})^2}{\sigma^2(N_i^{data}) + (P_b \sigma(N_{bi}^{MC}))^2 + (P_c \sigma(N_{ci}^{MC}))^2 + (P_l \sigma(N_{li}^{MC}))^2} + \frac{(N_{tot}^{data} - P_b N_{totb}^{MC} - P_c N_{totc}^{MC} - P_l N_{totl}^{MC})^2}{\sigma^2(N_{tot}^{data}) + (P_b \sigma(N_{totb}^{MC}))^2 + (P_c \sigma(N_{totc}^{MC}))^2 + (P_l \sigma(N_{totl}^{MC}))^2}$$

is minimized, with  $i$  running over all bins of the subtracted  $S_1$  and  $S_2$  distributions.

Only the statistical errors of the data and the Monte Carlo simulation are taken into account. The fit to the  $S_1$  and  $S_2$  distributions mainly constrains  $P_c$  and  $P_b$ , where the  $c$  and  $b$  quark fractions are distinguished by their different shapes. In contrast, the light quark normalisation is constrained by the total number of dijet events without any track association.



**Figure 6.5:** Distributions of a) the subtracted signed significance for the sample with exactly one selected track and b) the subtracted signed significance for the sample with two or more selected tracks. Only tracks with an impact parameter  $|\delta| < 0.1$  cm are considered. The expectation from the PYTHIA Monte Carlo simulation is included in the figure, showing the contributions from the various quark flavours after applying the scale factors  $P_b = 1.98$ ,  $P_c = 1.45$  and  $P_l = 1.44$  as obtained from the fit.

The fit to the complete data sample gives scale factors

$$\begin{aligned} P_b &= 1.98 \pm 0.22 \\ P_c &= 1.45 \pm 0.14 \\ P_l &= 1.44 \pm 0.05 \end{aligned} \tag{6.3}$$

for PYTHIA and has a  $\chi^2/n.d.f.$  of 13.1/18. The number of degrees of freedom (*n.d.f.*) of the fit are the number of bins of the subtracted  $S_1$  and  $S_2$  distributions, plus one (total number of events), minus three for the free parameters of the fit ( $P_l$ ,  $P_c$  and  $P_b$ ). Consistent results have been found when varying the number of bins in the fit.

The subtracted  $S_1$  and  $S_2$  distributions are shown together with the fit result in figure 6.5. The distributions are dominated by charm quark events, with an increasing fraction of beauty quark events towards larger values of significance. The contribution from light quarks is seen to be small.

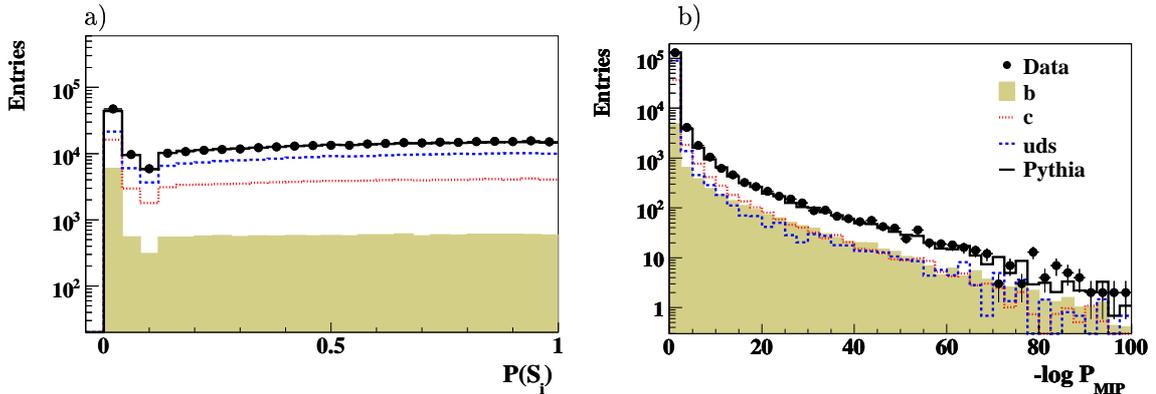
## 6.4 Multi-Impact Parameter Method

The results obtained with the  $\chi^2$ -fit to the subtracted  $S_1$  and  $S_2$  distributions are cross checked using an alternative method to separate the quark flavours. The multi-impact parameter (MIP) method is also based on the use of the significance distributions of the selected tracks  $S_i$  and was employed by the ALEPH collaboration [89]. It was adopted to H1 and first investigated in [90].

The quantity

$$P(S_i) = \frac{1}{\sqrt{2\pi}} \int_{\chi_i^2}^{\infty} e^{-t^2} dt, \tag{6.4}$$

with  $\chi_i^2 = S_i^2$ , can be interpreted as the probability that a track originates at the primary vertex. More precisely,  $P(S_i)$  is the probability of having a larger significance value  $S_i$  than



**Figure 6.6:** Distributions of a) the track probability  $P(S_i)$  and b) the negative logarithm of the multi impact parameter probability  $-\log(P_{MIP})$  for both jets. Only tracks with an impact parameter  $|\delta| < 0.1$  cm are used. The expectation from the PYTHIA Monte Carlo simulation is included in the figure, showing the contributions from the various quark flavours after applying the scale factors as obtained from the fit to the subtracted significance distributions.

the measured one. For a Gaussian distribution of the significances, the track probability distribution should be flat between 0 and 1, meaning that the tracks are compatible with the primary vertex. In contrast, tracks with large significances accumulate at very low values of  $P(S_i)$ .

A multi impact parameter probability  $P_{MIP}$  is then constructed by combining the probabilities of the  $N$  selected CST tracks within each jet:

$$P_{MIP} = \Pi \sum_{j=0}^{N-1} (-\ln \Pi)^j / j!, \quad (6.5)$$

where  $j$  runs over all selected CST tracks and

$$\Pi = \prod_{i=1}^N P(S_i). \quad (6.6)$$

The construction of  $P_{MIP}$  assumes the  $N$  track probabilities  $P(S_i)$  to be independent and  $P(S_i)$  is set to unity for tracks with negative significance. The heavy quark jet characteristic of having a large track multiplicity and producing on average more tracks with large impact parameter significance than light quark jets is combined in the variable  $P_{MIP}$ .

The distributions of the track probability  $P(S_i)$  and the negative logarithm  $-\log(P_{MIP})$  of the MIP probability for both jets is shown in figure 6.6. When performing a  $\chi^2$ -fit of the fractions of  $b$ ,  $c$  and light quarks to the measured  $-\log(P_{MIP})$  distribution of the data using the Monte Carlo expectations for the shapes of each of these quark flavours, the results are found to agree well with the  $(S_1, S_2)$ -method. To illustrate this, the Monte Carlo contributions in figure 6.6 are scaled by the factors obtained using the  $(S_1, S_2)$ -method.

The  $\chi^2$ -fit to the  $-\log(P_{MIP})$  distribution gives scale factors  $P_b = 2.25 \pm 0.38$ ,  $P_c = 1.69 \pm 0.21$  and  $P_l = 1.31 \pm 0.07$  for PYTHIA and has a  $\chi^2/n.d.f.$  of 23.3/36. Thus the statistical precision of the multi-impact parameter method is worse by a factor 1.5 compared to the  $(S_1, S_2)$ -method.

The important drawback of the multi-impact parameter method however is that it relies on the impact parameter resolution description by the Monte Carlo simulation. As discussed in section 7.1 the impact parameter resolution is one of the dominating systematic errors. It was found that this uncertainty is a factor 1.5-2 larger for the MIP-method than for the  $(S_1, S_2)$ -method, where the resolution dependence is highly reduced by the subtraction of the negative side.

## 6.5 Secondary Vertex Reconstruction

A further check is performed using an independent method to separate the quark flavours based on the explicit reconstruction of decay vertices in the transverse plane. In this approach, events with at least two selected CST tracks in a given jet are used and the primary and secondary vertex are fitted simultaneously. The selected CST tracks are used for both vertices whereas a relaxed track selection, corresponding to the standard primary vertex fit CSPRIM [73], is applied for the refit of the primary vertex.

The method was first presented at H1 in [5] and a more detailed description can be found e.g. in [91]. Here, only the main aspects of this approach are outlined.

Usually, when trying to fit two or more vertices simultaneously, the problem of the correct assignment of tracks to vertices arises. The basic idea of this approach is, that there is no definite mapping of tracks to vertices done, but each track is assigned a weight in the range 0 to 1 for each vertex candidate. This soft assignment is based on the track to vertex distance and follows the weight function of [92]

$$w_{ij}(T) = \frac{e^{-\frac{-S_{ij}^2}{2T}}}{\sum_{j'} e^{-\frac{-S_{ij'}^2}{2T}} + e^{-\frac{-\mu^2}{2T}}}, \quad (6.7)$$

with the sum running over all vertices  $j'$ . The  $S_{ij}$  are the impact parameter significances of track  $i$  with respect to vertex  $j$ . The smaller the distance of the track to the vertex candidate, the larger the weight. The parameter  $\mu$  is fixed to  $\mu = 2$  and it is used as a cut-off rejecting outliers<sup>1</sup>.

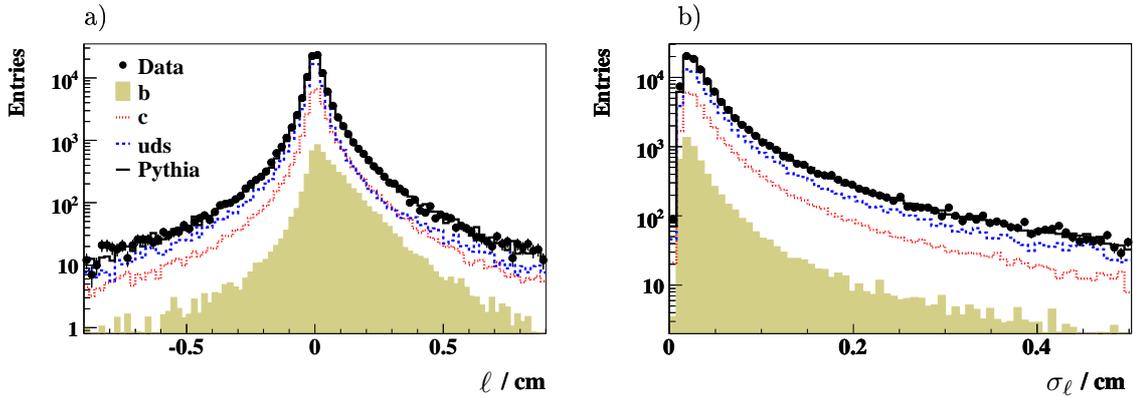
For each jet the vertex configuration that minimises the global fit  $\chi^2$  is found iteratively using deterministic annealing [93].

The concept of deterministic annealing optimises minimisation algorithms in which many local minima may occur. In this case, for e.g. three selected CST tracks up to four different secondary vertices corresponding to four local minima in the fit algorithm can be found. Deterministic annealing avoids local minima by introducing a 'temperature' function, denoted with  $T$  in equation 6.7. The global  $\chi^2$  of the fit is then optimised at each temperature iteratively, starting at a high temperature and then decreasing it slowly. This procedure is independent of the initial choice of  $T$ .

As a result of the iterative fit, the primary vertex ( $\vec{r}_{PV}$ ) and the secondary vertex position ( $\vec{r}_{SV}$ ) and their multiplicity is obtained. The number of tracks which contribute with a weight greater than 0.8 to the secondary vertex (SV) is used as a measure of the decay vertex track multiplicity.

---

<sup>1</sup>For  $S_{ij'} > \mu$   $w_{ij} \rightarrow 0$ .



**Figure 6.7:** Distributions of a) the decay length  $\ell$  and b) the decay length error  $\sigma_\ell$  for both jets. For the secondary vertex only tracks with an impact parameter  $|\delta| < 0.1$  cm are chosen. PYTHIA is scaled with the scale factors as obtained from the fit to the subtracted ( $S_1$ ,  $S_2$ )-distributions.

The measured decay length  $\ell$  is the transverse distance between both vertices according to

$$\ell = (\vec{r}_{SV} - \vec{r}_{PV}) \cdot \left( \frac{\vec{p}_t^{jet}}{|\vec{p}_t^{jet}|} \right), \quad (6.8)$$

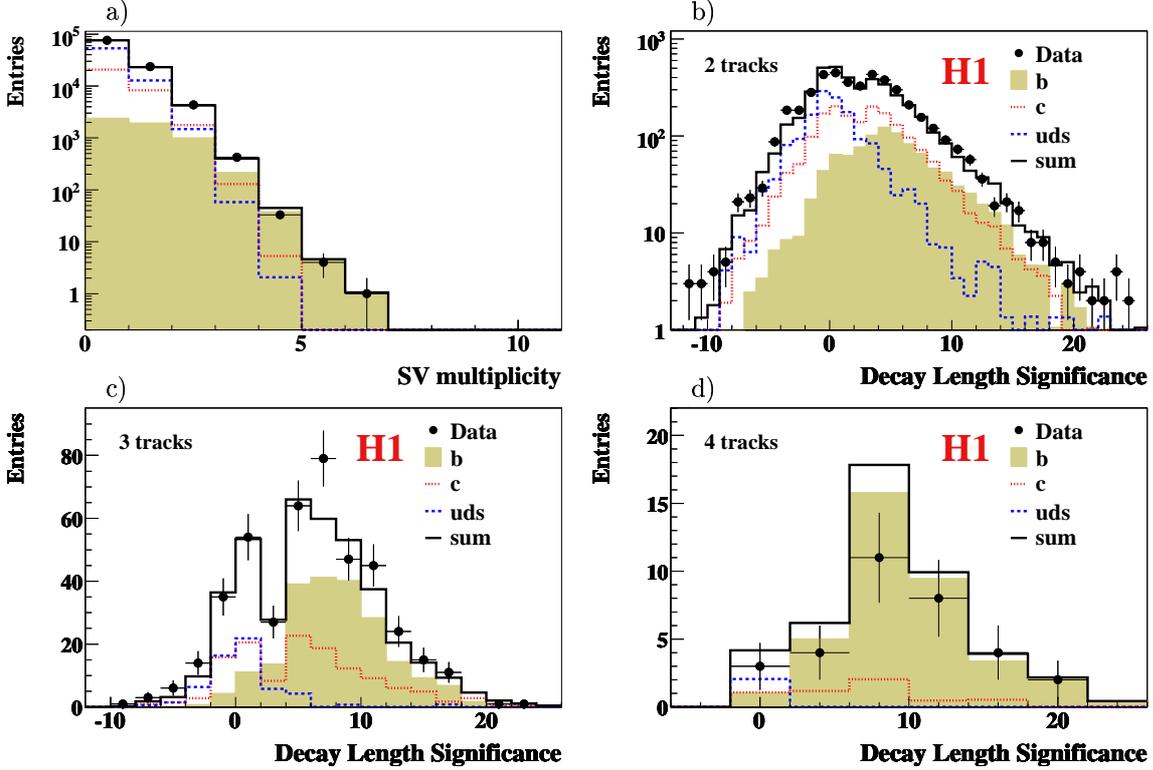
with  $\vec{p}_t = (p_x, p_y)$ ,  $|\vec{p}_t| = \sqrt{p_x^2 + p_y^2}$  and  $\vec{r} = (x, y)$ . Thus, a sign is assigned to the decay length using the projection of the corresponding vector onto the jet direction.

Figure 6.7 shows the decay length  $\ell$  and its error  $\sigma_\ell$  for both jets. Good agreement is seen between the data and the distributions from the PYTHIA Monte Carlo simulation which are scaled by the factors as obtained from the fit to the subtracted significance distributions  $S_1$  and  $S_2$ .

The decay length spectrum has similar features as the impact parameter distribution (figure 6.2) showing a marked (moderate) asymmetry of the  $b$  ( $c$ ) component with an almost symmetric contribution from light quarks. The decay length error  $\sigma_\ell$  is driven by the accuracy in the primary and decay vertex reconstruction. Due to the directional constraint these positions are correlated which is taken into account in the error calculation. The mean decay length error is about  $65 \mu\text{m}$  with a long tail towards larger errors.

The decay length significance is defined as the ratio of the signed decay length over its error. The secondary vertex multiplicity and the decay length significance for various SV-multiplicities is presented in figure 6.8. Apparently, in heavy quark events secondary vertices are significantly displaced in the direction of the related jet axis. Contributions from light quarks are mainly observed in the two-track sample at small decay length significances. In the two-track sample the contribution from charm and beauty is similar to that of light quark events while the three and four-track samples are dominated by the beauty component, as expected from the mean charged particle multiplicity in heavy hadron decays (cf. figure 6.8 a).

The PYTHIA Monte Carlo simulation is again scaled in the figure by the factors obtained with the ( $S_1$ ,  $S_2$ )-method, demonstrating the good agreement of both independent methods. Conversely, a simultaneous  $\chi^2$ -fit of the three shown decay length significance distributions for the different track multiplicities gives results consistent with those from the subtracted significance distributions.



**Figure 6.8:** Distributions of a) the secondary vertex (SV) multiplicity and the decay length significance ( $\ell/\sigma_\ell$ ) with multiplicities SV=2 (b), SV=3 (c) and SV=4 (d). PYTHIA is scaled with the scale factors as obtained from the fit to the subtracted ( $S_1, S_2$ )-distributions.

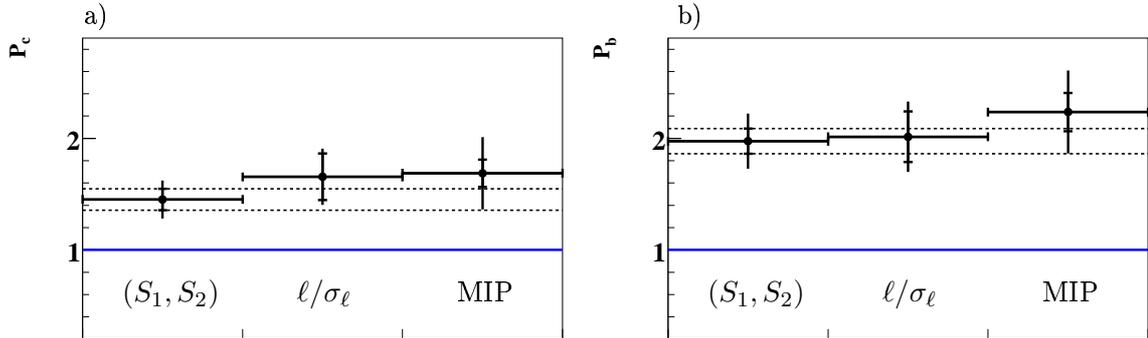
Scale factors of  $P_b = 2.0 \pm 0.5$ ,  $P_c = 1.66 \pm 0.35$  and  $P_l = 1.35 \pm 0.13$  with a  $\chi^2/n.d.f.$  of 43.0/51 are obtained by the fit.

Compared to the ( $S_1, S_2$ )-method the fitted scale factors are statistically less precise, which is due to the higher track multiplicity needed to perform the fit. However, the purity of a heavy quark sample selected by subsequent cuts on the secondary vertex track multiplicity and the decay length significance, is superior to that when using the impact parameter significances. An investigation of a heavy quark enriched subsample is presented in the next section.

### Summary of the fit methods

Figure 6.9 summarises the results of the various quark flavour separation methods presented in this chapter. The scale factors  $P_c$  and  $P_b$  are shown with the inner error bars indicating the statistical uncertainty and the outer error bars showing the statistical and systematic errors added in quadrature.

Consistent results are found for all fits with the smallest statistical error seen using the ( $S_1, S_2$ )-method. The dashed lines indicate the size of the statistical error of this method. The systematic error, which is discussed in detail in section 7.1, is largest in the MIP-approach with the largest contribution coming from the impact parameter resolution uncertainty. The systematic error of the decay length method is compatible with that of the ( $S_1, S_2$ )-approach. But since the latter has a more powerful statistical precision, the



**Figure 6.9:** Summary of the results from the independent quark flavour separation methods. Shown are the scale factors  $P_c$  (a) and  $P_b$  (b) as obtained from the fit to the subtracted  $(S_1, S_2)$ -distributions, the fit to the decay length significance ( $\ell/\sigma_\ell$ ) distributions and the fit to the multi-impact parameter (MIP) distribution. The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic error added in quadrature.

$(S_1, S_2)$ -method is chosen as default in this analysis for the measurement of charm and beauty dijet cross sections presented in the next chapter.

## 6.6 Heavy Flavour Enriched Data Sample

A subsample of dijet events with an enriched fraction of heavy quarks is investigated in this section.

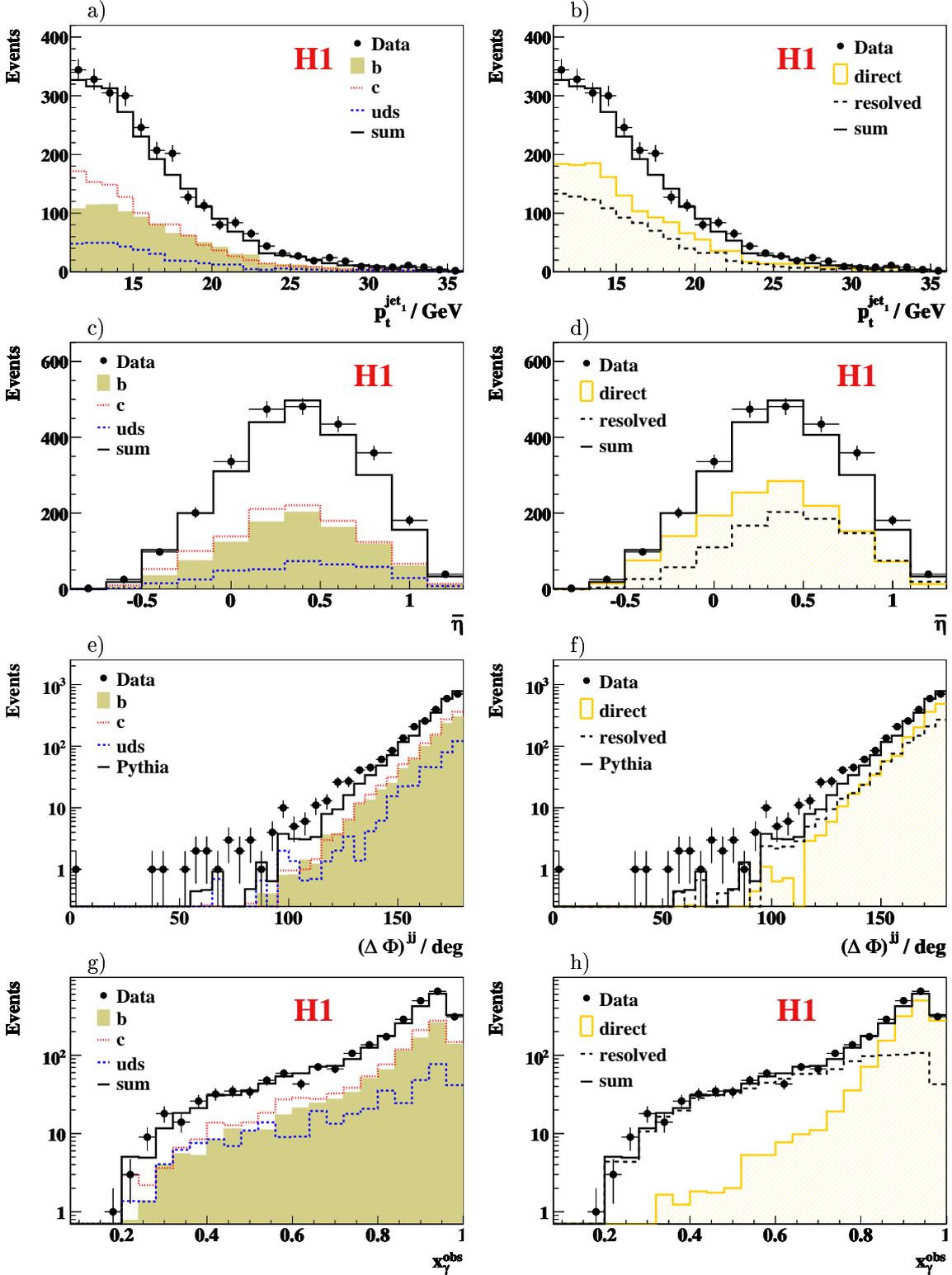
The relative fraction of heavy quarks in the event sample is enhanced by requiring a secondary decay vertex with two or more tracks and a decay length significance larger than 2.0 (cf. figure 6.8). Applying these cuts, the distributions for  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $(\Delta\phi)^{jj}$  and for  $x_\gamma^{obs}$  are shown in figure 6.10. The data are compared with the PYTHIA Monte Carlo simulations. In figures 6.10 a, c, e and g the decomposition into light, charm and beauty quark contributions is indicated. These are determined using the scale factors as obtained from the fit to the subtracted significance distributions of the full sample. The contributions from charm and beauty events are 44% and 41% respectively, while 15% of the events contain only light quarks.

Apart from the tail of the  $(\Delta\phi)^{jj}$  distribution, good agreement between the PYTHIA prediction and the data is seen. The description by PYTHIA for the  $(\Delta\phi)^{jj}$  distribution is considerably better for the heavy quark dominated sample than for the flavour inclusive sample (figure 5.8 c).

In figures 6.10 b, d, f and h the same data are shown together with the contributions from direct and resolved processes, as predicted by the PYTHIA simulation. According to the leading order QCD calculation, as implemented in PYTHIA, the fraction of events that arise from processes with resolved photons is about 40% for both, charm and beauty. In the region  $x_\gamma^{obs} > 0.85$  the fraction of direct processes dominates ( $\gtrsim 80\%$ ).

In the region  $x_\gamma^{obs} < 0.85$  the contribution from resolved processes is enhanced and amounts to about 85% (73%) for the selected charm (beauty) events. According to PYTHIA the dominant contribution results from processes with heavy quark excitation, in which one of the two jets originates from a light parton, i.e. a light quark or a gluon.

In conclusion, the selected data sample contains 85% of heavy quark events and is well



**Figure 6.10:** Distributions of the heavy quark enriched sample with two or more tracks originating from a secondary vertex and a decay length significance larger than 2.0. The left column shows the decomposition of the distribution into beauty, charm and light quarks after scaling the PYTHIA predictions by the scale factors obtained from the fit to the subtracted significance distributions of the full sample. In the right column the contributions from direct and resolved processes in PYTHIA are indicated as shaded histogram and dashed line respectively.

described by the leading order Monte Carlo simulation PYTHIA in which significant contributions from heavy quark excitation processes are predicted.

### Heavy Quark Jet Structure

The internal jet structure and the mean integrated jet shapes for flavour inclusive dijet events were investigated in section 5.7. In this paragraph the same quantities are discussed for the heavy quark enriched dijet sample.

In figure 6.11 the jet profiles are shown as functions of the difference in azimuthal angle and pseudo-rapidity between the jet and the objects contributing to the jet. The agreement with PYTHIA is similarly good as for the flavour inclusive sample (figure 5.10). In particular, PYTHIA gives a reasonable description of the energy flow in the directions of both the proton remnant ( $\eta^{obj} - \eta^{jet} > 0$ ) and the photon remnant ( $\eta^{obj} - \eta^{jet} < 0$ ). In the latter region CASCADE significantly underestimates the data.

Figure 6.12 shows the energy flow for different regions of  $x_\gamma^{obs}$ . While PYTHIA and partly also CASCADE are able to reproduce the energy profiles for the region of large values of  $x_\gamma^{obs}$ , the description is worse for events with  $x_\gamma^{obs} < 0.85$ . For the transverse region, i.e.  $|\phi^{obj} - \phi^{jet}| \sim \pi/2$  (figure 6.12b), differences between the data and the models become apparent. In section 5.7 it was found that this region, which is very sensitive to multiple parton interactions, is well described by PYTHIA for the flavour inclusive data sample (cf. figure 5.11 b).

These findings support the hypothesis discussed in [40] that multiple parton interactions, which are needed to describe light quark dominated data samples, behave differently for heavy quark events.

In order to investigate the mean integrated jet shape  $\langle \Psi(r) \rangle$  for heavy quark events, a background subtraction has to be performed. The mean integrated jet shape measured in the data can be written as

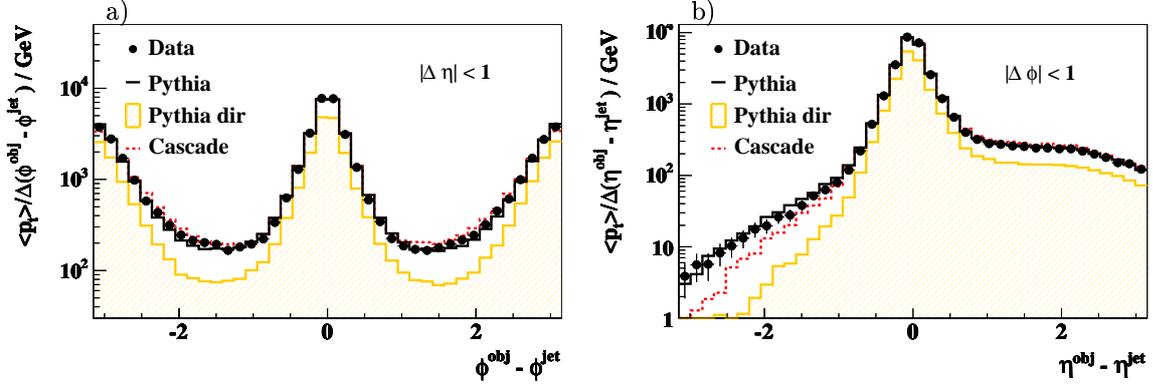
$$\langle \Psi(r) \rangle^{data} = f_b \langle \Psi(r) \rangle^b + f_c \langle \Psi(r) \rangle^c + f_{uds} \langle \Psi(r) \rangle^{uds}, \quad (6.9)$$

with  $f_b$ ,  $f_c$  and  $f_{uds}$  being the fractions of beauty, charm and light quark events, respectively. For the selected heavy quark enriched sample these fractions are determined to be  $f_b = 0.41$ ,  $f_c = 0.45$  and  $f_{uds} = 0.15$ . Using the same method as described in [40], a background subtraction procedure can be performed to extract both,  $\langle \Psi(r) \rangle^c$  and  $\langle \Psi(r) \rangle^b$ . Hence, subtracting the background from beauty and light quark events, the measured mean integrated charm jet shape is defined as

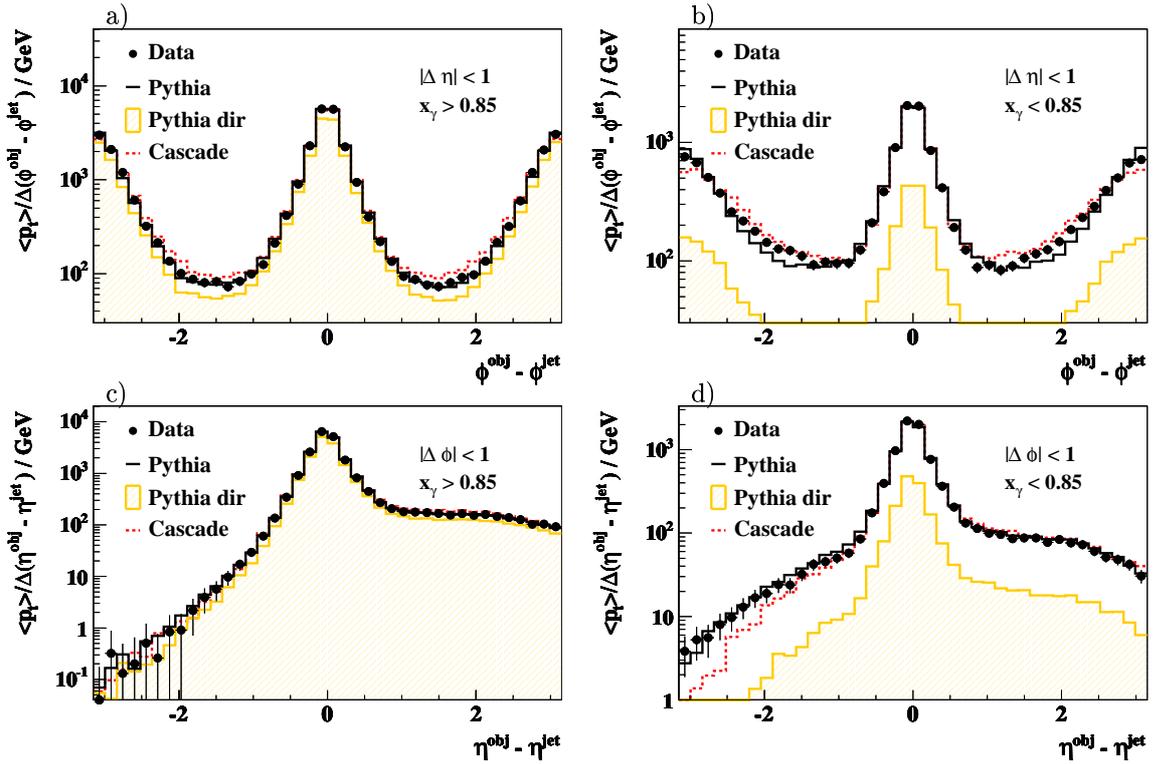
$$\langle \Psi(r) \rangle^c = \frac{\langle \Psi(r) \rangle^{data} - f_b \langle \Psi(r) \rangle^b - f_{uds} \langle \Psi(r) \rangle^{uds}}{1 - f_b - f_{uds}}. \quad (6.10)$$

Accordingly,  $\langle \Psi(r) \rangle^b$  is obtained by replacing the charm related factors in equation 6.10 with the corresponding beauty factors and vice versa.

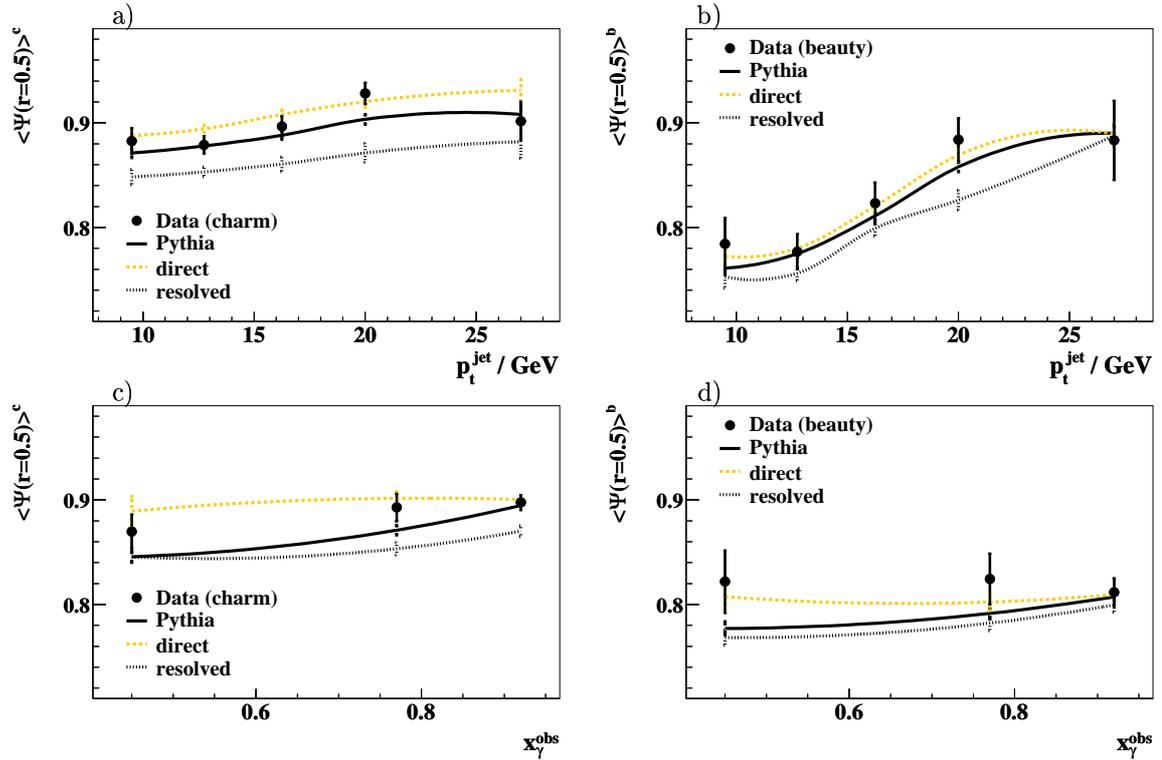
In figure 6.13 the mean integrated charm and beauty jet shapes  $\langle \Psi(r = 0.5) \rangle^c$  and  $\langle \Psi(r = 0.5) \rangle^b$  are presented as functions of the variables  $p_t^{jet}$  and  $x_\gamma^{obs}$ . The data is compared to the total charm (beauty) PYTHIA predictions. The contributions of direct and resolved events are shown separately.



**Figure 6.11:** The transverse energy flow around the jet axis for the heavy quark enriched sample as a function of a) the distance to the jet axis in azimuthal angles  $\phi^{obj} - \phi^{jet}$  and b) as a function of the distance in pseudo-rapidity  $\eta^{obj} - \eta^{jet}$ . The predictions from PYTHIA and CASCADE are normalised to the data and indicated as full line and dotted line respectively.



**Figure 6.12:** The transverse energy flow around the jet axis for the heavy quark enriched sample as a function of a), b) the distance to the jet axis in azimuthal angles  $\phi^{obj} - \phi^{jet}$  and c), d) as a function of the distance in pseudo-rapidity  $\eta^{obj} - \eta^{jet}$  for different  $x_\gamma^{obs}$  regions. The predictions from PYTHIA and CASCADE are normalised to the data and indicated as full line and dotted line respectively.



**Figure 6.13:** Mean integrated jet shape at fixed radius  $r = 0.5$  for the heavy quark enriched dijet sample as functions of a), b)  $p_t^{\text{jet}}$  and c), d) the observable  $x_\gamma^{\text{obs}}$ . The left (right) column shows the charm (beauty) data after background subtraction (see text). The data is compared to the total charm and beauty PYTHIA predictions (full line). Direct (dashed line) and resolved (dotted line) contributions are shown separately. Only statistical errors are shown.

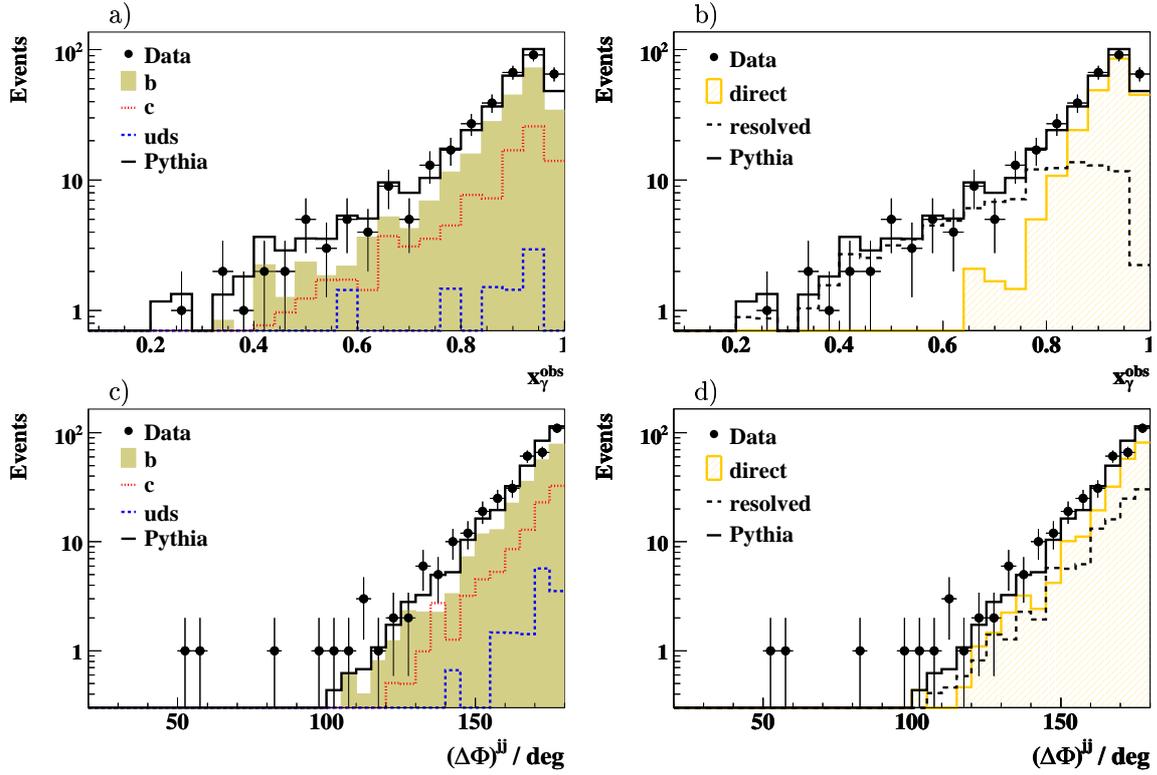
Figure 6.13 illustrates the differences in the internal jet structure of charm and beauty events. Narrower jets are characterised by larger values of  $\langle \Psi(r) \rangle$ , and beauty jets are seen to be broader than charm jets. This may be understood from the larger beauty quark mass giving rise to a larger relative transverse momentum of the decaying particles with respect to the originating parton, and thus leading to broader jets.

Both, charm and beauty data, are generally in reasonable agreement with the total PYTHIA prediction (full line). At low values of  $x_\gamma^{\text{obs}}$  the data suggest a trend to larger  $\langle \Psi(r) \rangle$  as expected for direct processes. Although the statistical precision of the data is limited, this observation is similar to that reported in [40] for charm jets.

### High Purity Beauty Sample

A high purity beauty event sample is selected by requiring a secondary vertex with *three* or more tracks and a decay length significance larger than 2.0 (cf. figure 6.8). In figure 6.14 the distributions of  $x_\gamma^{\text{obs}}$  and  $(\Delta\phi)^{jj}$  are shown, together with the PYTHIA predictions. In figures 6.14 a) and c), the decomposition into light, charm and beauty quark contributions is indicated using the scale factors as obtained from the fit to the subtracted significance distributions of the full sample. In this subsample 68% (28%) of the events originate from beauty (charm) quark events while only 4% of the events contain only light quarks.

The data is fairly well described by the PYTHIA Monte Carlo simulation. In particular



**Figure 6.14:** Distributions of the beauty quark enriched sample with three or more tracks originating from a secondary vertex and a decay length significance larger than 2.0. The left column shows the decomposition of the distribution into beauty, charm and light quarks after scaling the PYTHIA predictions by the scale factors obtained from the fit to the subtracted significance distributions of the full sample. In the right column the contributions from direct and resolved processes in PYTHIA are indicated as shaded histogram and dashed line respectively.

for the distribution of  $(\Delta\phi)^{jj}$  the agreement is better than for the charm quark event dominated heavy quark sample which is presented in figure 6.10.

Figures 6.14 b) and d) illustrate the need of resolved processes for a successful description of data samples in which most of the events originate from beauty quark decays.

## CHAPTER 7

# CROSS SECTION MEASUREMENT

---

In this chapter the measured cross sections for inclusive, charm and beauty dijet production in the reaction  $ep \rightarrow e(q\bar{q})X \rightarrow ejjX$  are presented. The data are corrected for detector effects and are presented at the level of stable hadrons for the phase space region defined in equation 5.7. The results are compared to various QCD models including NLO predictions for charm and beauty obtained with the settings described in section 2.5.

In section 7.2 differential dijet charm and beauty cross sections are measured as a function of the transverse momentum of the leading jet,  $p_t^{jet1}$ , the average pseudo-rapidity of the two jets,  $\bar{\eta}$ , the observable  $x_\gamma^{obs}$ , the azimuthal separation  $(\Delta\phi)^{jj}$  of the two jets and the centre-of-mass scattering angle  $\cos\theta^*$ .

Measuring the corresponding flavour inclusive dijet cross sections (section 7.3) gives access to the relative charm and beauty fractions in the visible range. These are presented in section 7.4.

Finally, the results are compared to previous charm and beauty measurements at HERA in section 7.5. The chapter starts with the discussion of the systematic error.

### 7.1 Systematic Uncertainties

Systematic uncertainties of the cross section measurement are evaluated by variations of the Monte Carlo simulations. The resulting systematic uncertainties of the total charm and beauty dijet cross sections are listed in table 7.1 and are detailed below.

#### Impact Parameter Resolution

The uncertainty arising from the knowledge of the impact parameter resolution of the tracks is estimated by varying the resolution by an amount that encompasses the differences between the data and the simulation. This is achieved by applying an additional Gaussian smearing in the Monte Carlo simulation of 200  $\mu\text{m}$  to 5% of randomly selected tracks and of 25  $\mu\text{m}$  to the rest. In figure 7.1 the distributions of  $S_1$  and  $S_2$  are shown on a linear and logarithmic scale. In addition to the default PYTHIA simulation (full line), the modified (smeared) simulation is shown as dashed line. The smearing of 200  $\mu\text{m}$  which affects mostly the tails of the distributions, is motivated by the presence of non-gaussian tails in a small

Source	Variation	Uncertainty [%]	
		Charm	Beauty
Impact parameter resolution	$\oplus 25\mu\text{m}$ $\oplus 200\mu\text{m}$	7	10
Jet axis $\phi$ direction	$1^\circ$ shift in $\phi$	3	2
Track finding efficiency	$2\% \oplus 1\%$	3	8
$uds$ asymmetry	$\pm 50\%$	1	6
HQ production model	resolved $\gamma$ , $p_t$ dependence	7	14
Fragmentation model	Peterson / Lund	1	2
Fragmentation fractions	PDG	0.5	1.6
Hadron lifetimes	PDG	0.1	0.3
Charged track multiplicities	MARK-III, LEP, SLD	1.5	4
Jet energy scale	2%	6	5
Trigger efficiency		5	5
Luminosity measurement		1.5	1.5
Total		14	22

**Table 7.1:** Systematic uncertainties of the measured total charm and beauty dijet cross sections.

fraction of the data which are not fully incorporated in the simulation. These effects are visible in single bins of the measurement (not shown).

### Jet Axis

The uncertainty of the jet axis reconstruction which can affect the determination of the correct sign of the impact parameter is estimated by shifting the jet axis in azimuth  $\phi$  by  $\pm 1^\circ$ , corresponding to the resolution.

### Track Finding Efficiency

The reconstruction efficiency of CJC tracks is uncertain to the level of 2% [94].

In section 4.5 the efficiency to find two CST hits on a given drift-chamber track is analysed. As documented in figure 4.8, the uncertainty of the CST hit linking efficiency is known to 1%. Thus, the total uncertainty on the track finding efficiency amounts to 2.24%.

### Light Quark Asymmetry

The uncertainty resulting from the shape of the subtracted significance distributions  $S_1$  and  $S_2$  for light quarks (cf. figure 6.5) is estimated by repeating the fits with the light quark  $S_1$  and  $S_2$  distributions varied by  $\pm 50\%$  of the default value. The motivation for this variation is illustrated in figure 7.2 where the 2 track mass for a strangeness enhanced data sample is shown. The distribution is shown after background subtraction, i.e. events with negative or positive significance are displayed with a weight  $w = -1$  or  $w = +1$ , respectively. Strangeness is enriched by requiring only events in the range  $0.1 \text{ cm} < \delta < 0.5 \text{ cm}$ .

In figure 7.2 a clear peak at  $\approx 500 \text{ MeV}$  signals the presence of  $K^0$  particles in the sample. The data is reasonably well described by the simulation. The error band of  $\pm 50\%$  (figure 7.2 b) encompasses the uncertainties of the light quark description in the simulation.

### Heavy Quark Production Model

Figure 7.3 a) shows the distribution of  $x_\gamma^{obs}$  for the heavy flavour enriched sample for data and the PYTHIA Monte Carlo simulation. The simulation is normalised to match the number of entries in the data and the contributions from the direct and resolved processes are indicated separately. Figure 7.3 b) shows similar plots for the CASCADE Monte Carlo. It can be seen that the data is reasonably well described by the PYTHIA model. However the CASCADE model fails to describe the data.

In the CASCADE simulation the amount of events with low values of  $x_\gamma^{obs}$  appears to be underestimated. This amounts to a larger relative fraction of  $Q\bar{Q}$ -events compared to  $gQ$ -processes. As gluon jets carry no lifetime information the lifetime spectrum of the CASCADE model is too broad. However in the PYTHIA model most of the resolved photon events have one jet produced by a heavy quark while the other originates from a gluon, resulting in a more favorable lifetime spectrum.

Thus, the choice of the production model (CASCADE or PYTHIA) has a large effect on the lifetime distributions and hence on the resulting measurements. The measured charm (beauty) cross sections differ by  $-15\%$  ( $-18\%$ ) when using CASCADE instead of PYTHIA for the fit.

A contribution to the systematic error due to the uncertainty of the underlying physics model is therefore made by varying the relative amounts of the resolved photon and direct processes in PYTHIA by  $\pm 50\%$ . The resulting change in the distribution of  $x_\gamma^{obs}$  due to this variation is shown as the shaded region in figure 7.3 b). As can be seen the data is well contained inside the error band. The average error of the charm and beauty cross sections due to this variation is  $7\%$  and  $11\%$ , respectively.

A further contribution to the systematic error was added after considering the effect of the modified PYTHIA simulation on the distributions of  $p_t^{jet1}$  and  $\eta^{jet1}$  for the heavy quark enriched sample (figure 7.4). The  $p_t$  spectrum is seen to be fairly independent of a variation of the direct and resolved contributions. The prediction of CASCADE is also shown, revealing a  $p_t$  spectrum less steep than that seen in PYTHIA and the data. In order to determine a model error for the acceptance corrections in  $p_t$ , the PYTHIA prediction is reweighted to that of CASCADE resulting in a systematic error of  $1\%$  and  $8\%$  for charm and beauty, respectively.

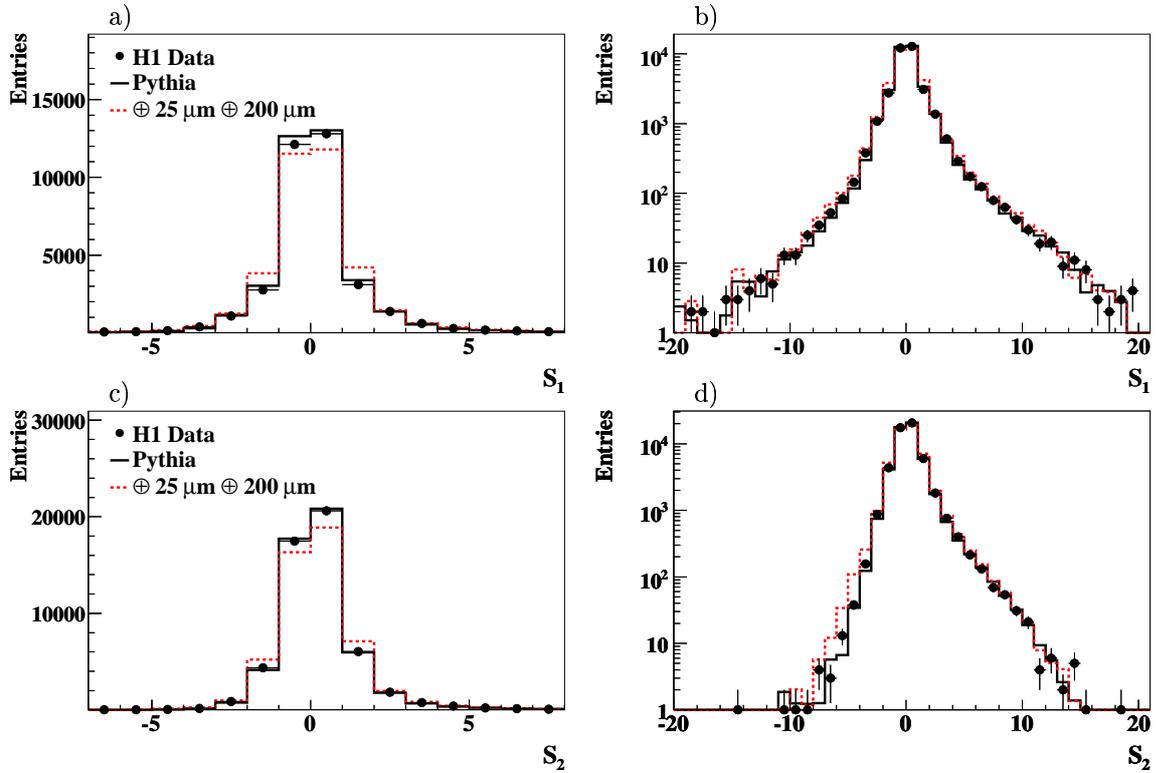
Thus the total systematic error arising from the uncertainty of the heavy quark production model is obtained by adding the contribution from the modified PYTHIA simulation to the contribution from the  $p_t$  reweighting in quadrature. These variations lead to systematic uncertainties of  $\pm 7\%$  for charm and  $\pm 14\%$  for beauty.

### Fragmentation Model

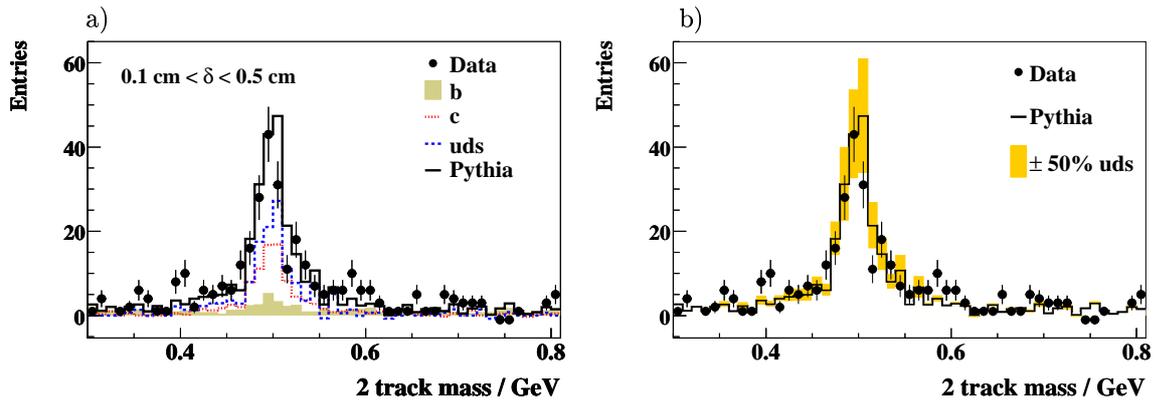
The uncertainties in the description of the heavy quark fragmentation are estimated by repeating the fits with a PYTHIA simulation in which the Lund Bowler function [36] is used for the longitudinal fragmentation instead of the Peterson function.

### Fragmentation Fractions, Lifetimes and Multiplicities

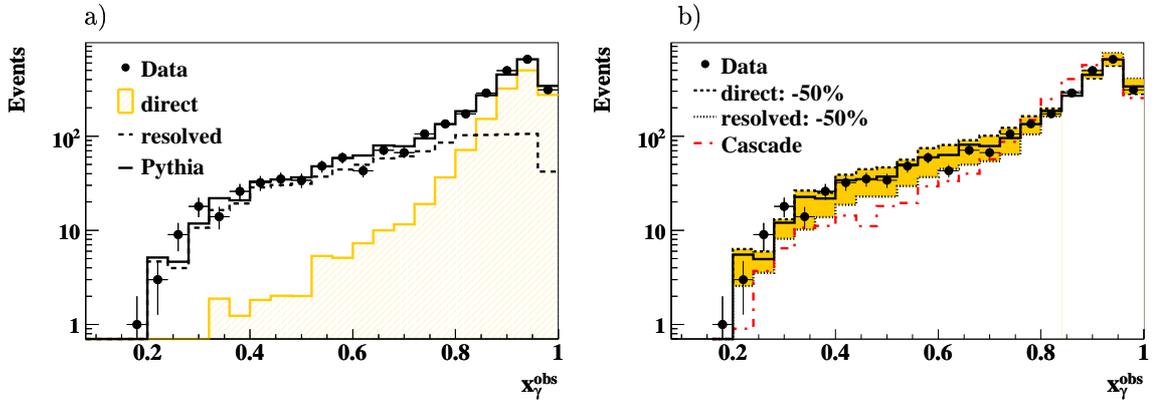
The uncertainties arising from the various  $D$  and  $B$  hadron lifetimes, fragmentation fractions and mean charged track multiplicities are estimated by varying the input values of



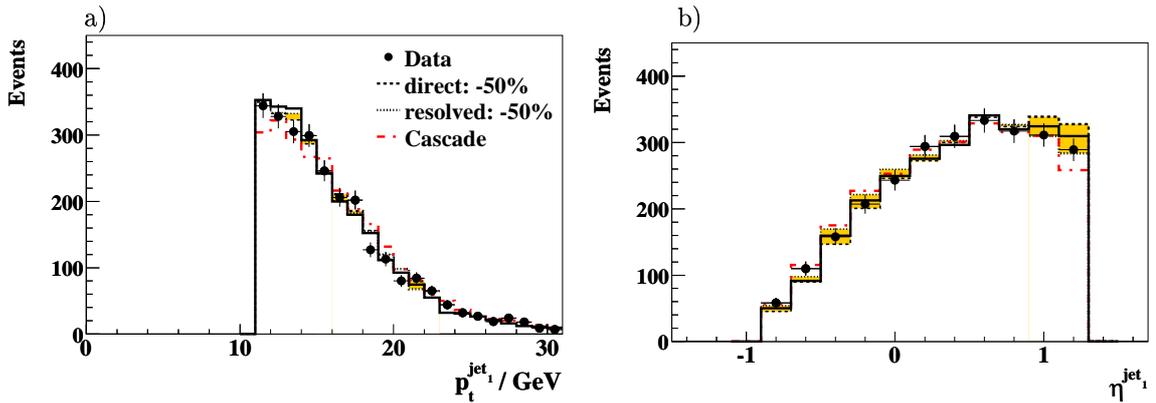
**Figure 7.1:** Distributions of the impact parameter significances  $S_1$  and  $S_2$  in linear (a and c) and semi-logarithmic scale (b and d). In addition to the default PYTHIA simulation (full line) a modified simulation is shown in which a Gaussian smearing of the impact parameter of  $200 \mu\text{m}$  to 5% of randomly selected tracks and of  $25 \mu\text{m}$  to the rest is applied (dashed line).



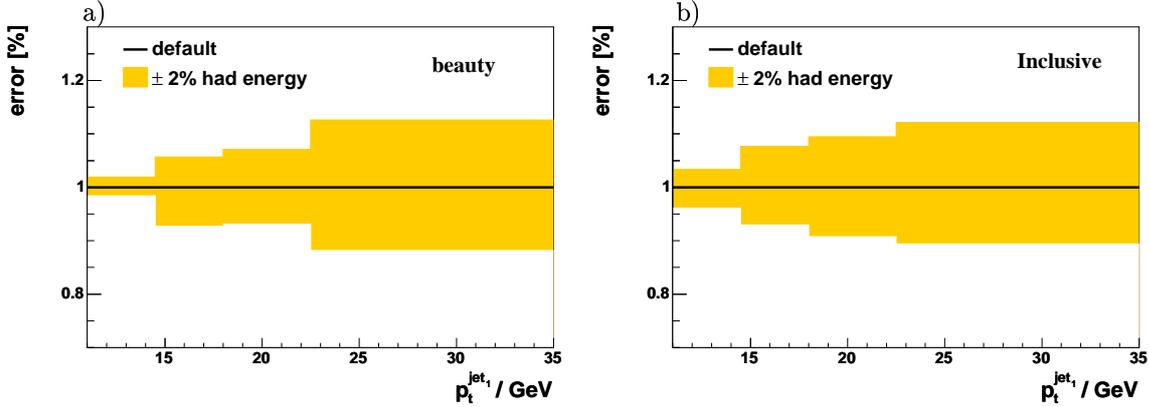
**Figure 7.2:** Distribution of the invariant 2 track mass in a strangeness enhanced data sample after background subtraction. Events with negative or positive impact parameter significance enter with a weight  $w = -1$  or  $w = +1$ , respectively. In figure a) the data is compared to the PYTHIA simulation. The decomposition into beauty, charm and light quarks is shown separately. In figure b) the same data is shown and compared to modified PYTHIA predictions in which the relative fraction of  $uds$  events is varied. The error band of  $\pm 50\%$  accounts for the uncertainty in the light quark description of the simulation.



**Figure 7.3:** Distributions of the variable  $x_\gamma^{obs}$  for the heavy flavour enriched sample (see section 6.6). The Monte Carlo distributions are normalised to match the data. Contributions from direct and resolved processes are shown separately for the default PYTHIA in figure a). Variations of the relative fractions of direct and resolved processes lead to the uncertainty band (shaded region) in figure b). Cascade (dashed-dotted line) is shown separately.



**Figure 7.4:** Distributions of  $p_t^{jet_1}$  and  $\eta^{jet_1}$  for the heavy flavour enriched sample (see section 6.6). The Monte Carlo distributions are normalised to match the data. Variations of the relative fractions of direct and resolved processes lead to the uncertainty band (shaded region). Cascade (dashed-dotted line) is shown separately.



**Figure 7.5:** Effect of the LAr energy scale variation on  $p_t^{\text{jet}_1}$  for the beauty (left) and flavour inclusive (right) Monte Carlo simulation. Shown is the estimated error of the cross section.

the Monte Carlo simulation by the experimental errors of the corresponding measurements or world averages. For the fragmentation fractions of  $c$  and  $b$  quarks to hadrons and for the lifetimes of these hadrons the central values and errors on the world averages are taken from [49]. For the mean charged track multiplicities the values and uncertainties for  $c$  hadrons are taken from Mark-III [95] and for  $b$  hadrons from LEP and SLD measurements [96].

### Jet Energy Scale

The LAr calorimeter energy scale has an uncertainty of 2% (4%) for the electromagnetic (hadronic) part. Since the jet energy measurement is predominantly track based ( $\geq 60\%$ ), a shift of the jet energy scale of  $\pm 2\%$  is considered here (cf. also section 5.3). As pointed out in [84], this procedure is similar to a variation of the individual LAr cell energies by  $\pm 2\%$ .

The uncertainty of the jet energy scale leads to cross section uncertainties from 3% at small  $p_t^{\text{jet}}$  to 12% at large  $p_t^{\text{jet}}$  and 6% on average, largely independent of the quark flavour. As an example in figure 7.5 the effect of the energy scaling as a function of  $p_t^{\text{jet}_1}$  can be seen for the beauty and flavour inclusive PYTHIA simulation.

### Trigger Efficiency and Luminosity

The trigger efficiency is studied using monitoring events from neutral current processes in deep inelastic scattering in which the scattered positron triggers the events independently of the triggers under study. The uncertainty is determined to be 5% (cf. section 5.4).

The luminosity at H1 is known to an accuracy of 1.5%.

### Total Systematic Uncertainties

Total systematic uncertainties of 14% and 22% are obtained for the measurement of the charm and the beauty production cross sections respectively. The total systematic error for the flavour inclusive dijet cross section, in which track selection and resolution effects do not arise, is 8% resulting from the uncertainty of the hadronic energy scale (6%), the trigger efficiency uncertainty (5%) and the uncertainty of the luminosity measurement (1.5%).

For the relative contributions of charm and beauty production to the flavour inclusive dijet cross section, the statistical errors are added in quadrature and the systematic errors include those sources that are specific to the charm and beauty cross section measurement. The same uncertainties are equally attributed to all bins of the measurement except for the uncertainty of the hadronic energy scale for which the uncertainties are determined and applied individually in each bin of the measurement.

## 7.2 Charm and Beauty Dijet Cross Section

For the measurement of the charm and beauty cross sections, the scale factors  $P_c$  and  $P_b$ , which are determined from fits of the subtracted impact parameter significance distributions to the data, are multiplied with the cross section predictions of the PYTHIA Monte Carlo simulation (cf. equations 5.11 and 6.3). The total dijet charm photoproduction cross section in the range  $Q^2 < 1 \text{ GeV}^2$ ,  $0.15 < y < 0.8$ ,  $Q^2 < 1 \text{ GeV}^2$ ,  $p_t^{\text{jet}_{1(2)}} > 11(8) \text{ GeV}$  and  $-0.9 < \eta^{\text{jet}_{1(2)}} < 1.3$  is measured to be

$$\sigma(ep \rightarrow ec\bar{c}X \rightarrow ejjX) = 702 \pm 67(\text{stat.}) \pm 95(\text{syst.}) \text{ pb.} \quad (7.1)$$

For the same kinematic range the measured beauty cross section is

$$\sigma(ep \rightarrow eb\bar{b}X \rightarrow ejjX) = 150 \pm 17(\text{stat.}) \pm 33(\text{syst.}) \text{ pb.} \quad (7.2)$$

A cross check of the charm cross section result using reconstructed  $D^*$  events is performed and described in appendix B.

The predictions from the theoretical calculations are detailed in table 7.2. The results of the massive NLO calculations are stated before and after correction to the hadron level using correction factors as determined from PYTHIA (see section 5.8.4). The uncertainties on the NLO calculation are determined as described in section 2.5.1. In the kinematic range of the measurement the variation of the renormalisation scale has the largest contribution to the theory error indicating the importance of higher order terms. The average total theory error accumulates to +35% and -20% for charm and +23% and -17% for beauty, respectively.

For charm, the central FMNR prediction after applying the hadronisation correction is somewhat lower than the measurement but still in reasonable agreement within the large theoretical errors.

For beauty, FMNR is lower than the data by a factor 1.8, corresponding to 1.6 standard deviations, taking both experimental uncertainties and uncertainties in the theory into account.

For both, charm and beauty, the LO plus parton shower calculations of PYTHIA and CASCADE predict values similar to that of FMNR.

### 7.2.1 Differential Cross Sections

For the measurement of the differential charm and beauty cross sections the fit is performed separately in each bin  $i$ . The resulting scale factors  $P_{c,i}$  and  $P_{b,i}$  are then multiplied with the bin-integrated cross section predictions of the PYTHIA Monte Carlo simulation, divided

	Charm [pb]	Beauty [pb]
Data	$702 \pm 67(stat.) \pm 95(syst.)$	$150 \pm 17(stat.) \pm 33(syst.)$
FMNR	$538^{+186}_{-107}$	$82^{+19}_{-14}$
FMNR $\otimes$ had	$500^{+173}_{-99}$	$83^{+19}_{-14}$
PYTHIA	484	76
CASCADE	438	80

**Table 7.2:** The measured charm and beauty photoproduction dijet cross sections in the kinematic range  $Q^2 < 1 \text{ GeV}^2$ ,  $0.15 < y < 0.8$ ,  $p_t^{jet1(2)} > 11(8) \text{ GeV}$  and  $-0.9 < \eta^{jet1(2)} < 1.3$  in comparison to predictions in massive NLO QCD (FMNR) and from the LO Monte Carlo programs PYTHIA and CASCADE.

by the respective bin size. PYTHIA is used to determine the point in the bin at which the bin-averaged cross section equals the differential cross section.

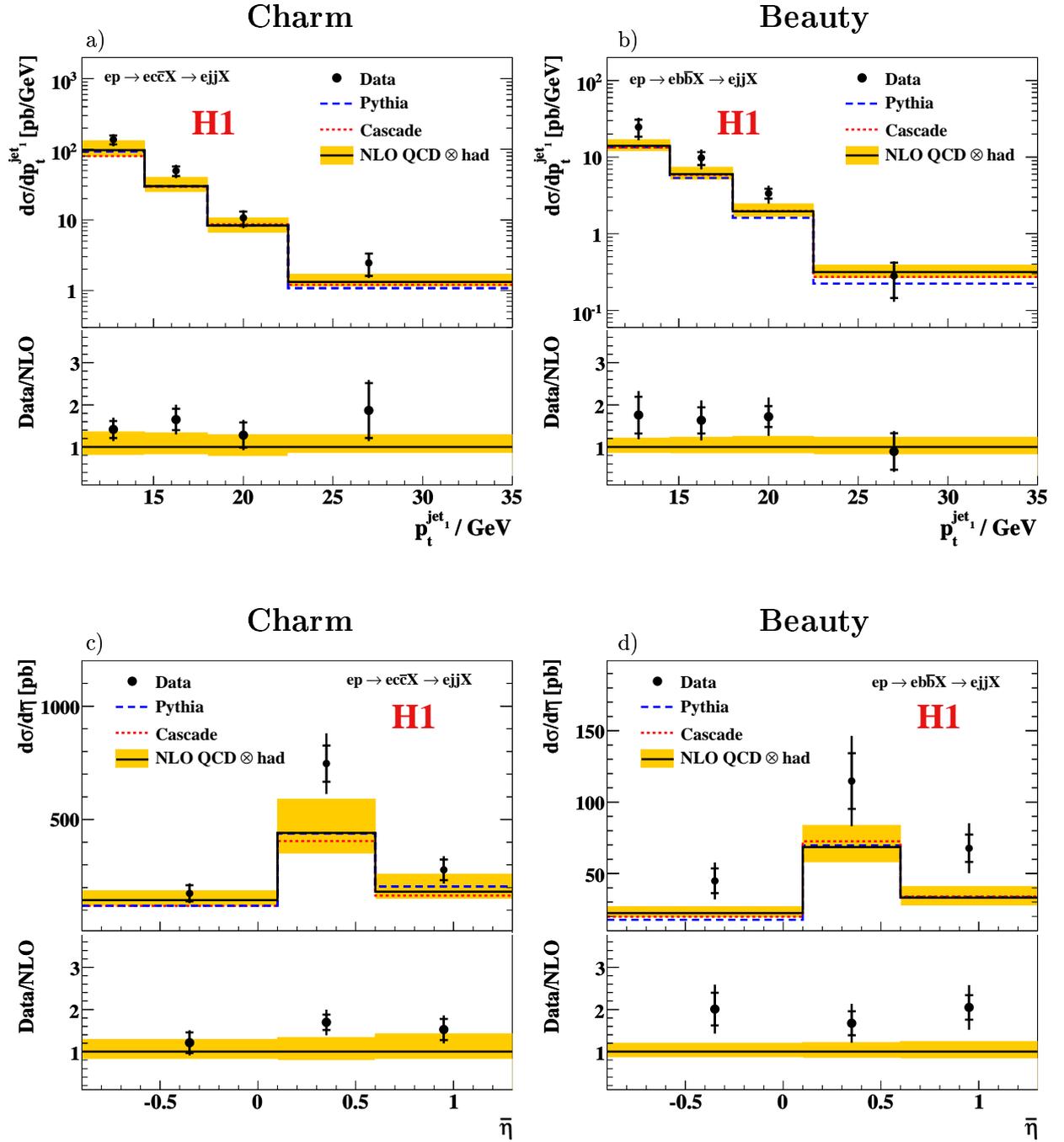
The data are presented at the level of stable hadrons after decay for the phase space region defined in equation 5.7. The inner error bars in the figures indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. All results are compared to FMNR predictions corrected to hadron level (NLO QCD  $\otimes$  had) as well as to LO Monte Carlo predictions from PYTHIA and CASCADE. For a better comparison of data and NLO predictions, the ratio between data and NLO calculation is shown on a linear scale below each histogram.

The measured differential cross sections as functions of  $p_t^{jet1}$  and of  $\bar{\eta}$  are shown in figure 7.6 and are listed in tables A.1 and A.2.

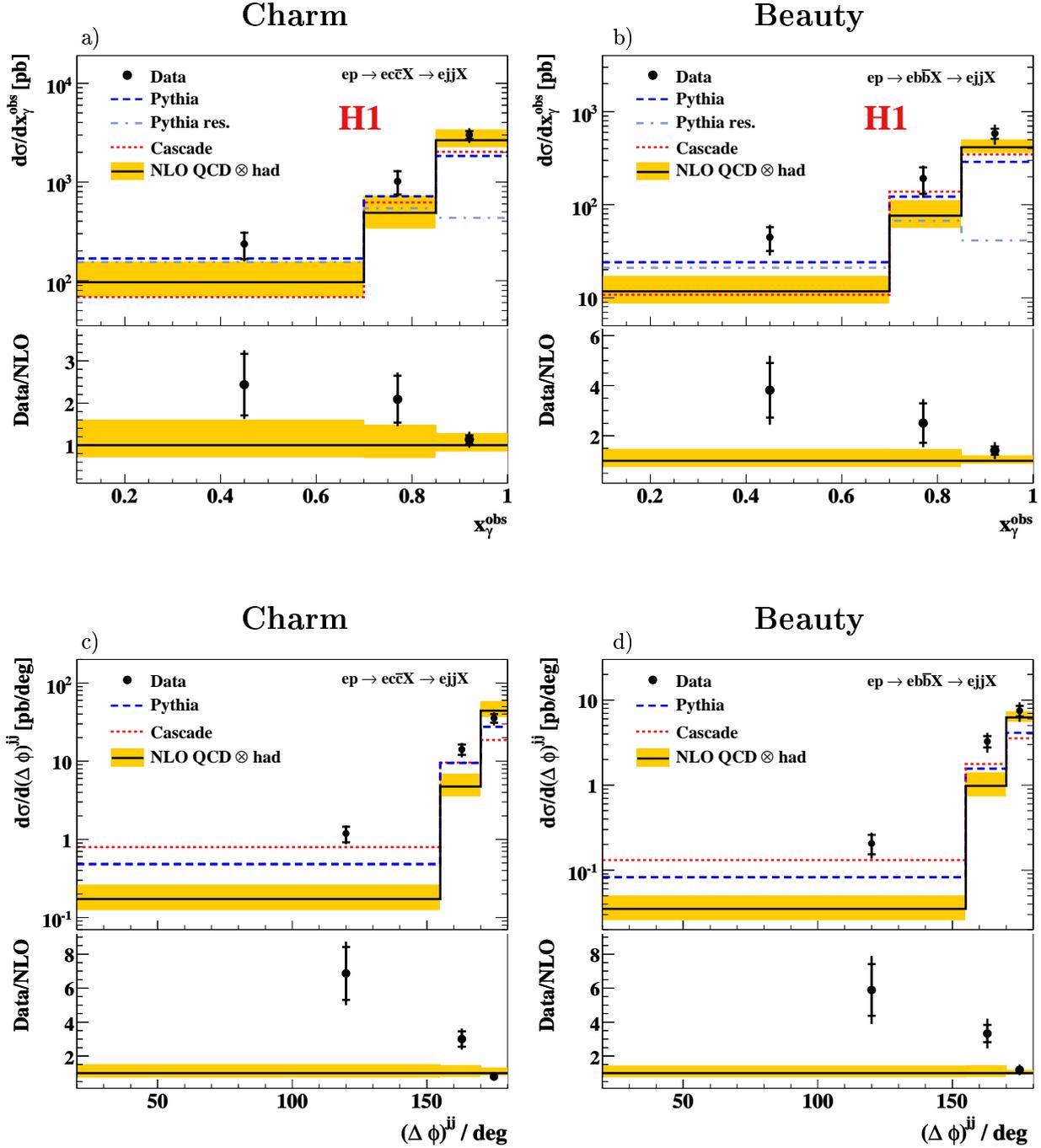
The  $p_t$  distribution is steeply falling over two orders of magnitude and is measured up to a central value of 27 GeV. The shape is similar for charm and beauty and is reasonably well described by all theories. Comparing the ratio of the measured data over NLO theory, the charm data is compatible with the NLO calculation but lying at the upper edge of the theory uncertainty. The theory uncertainty is not reduced significantly with increasing  $p_t^{jet1}$ . For beauty the difference between data and NLO theory is more pronounced with the data reaching the level of the NLO calculation only at largest  $p_t^{jet1}$ .

The cross section as a function of the mean pseudo-rapidity  $\bar{\eta}$  is sensitive to the parton distributions in both the photon and the proton. The shape is similar for charm and beauty and reproduced reasonably well by the theories. For charm, a reasonable agreement between data and NLO QCD is observed taking into account the uncertainties in the calculations and in the data points. The beauty data, again, is seen to be somewhat underestimated by the NLO QCD prediction over the whole range of  $\bar{\eta}$ .

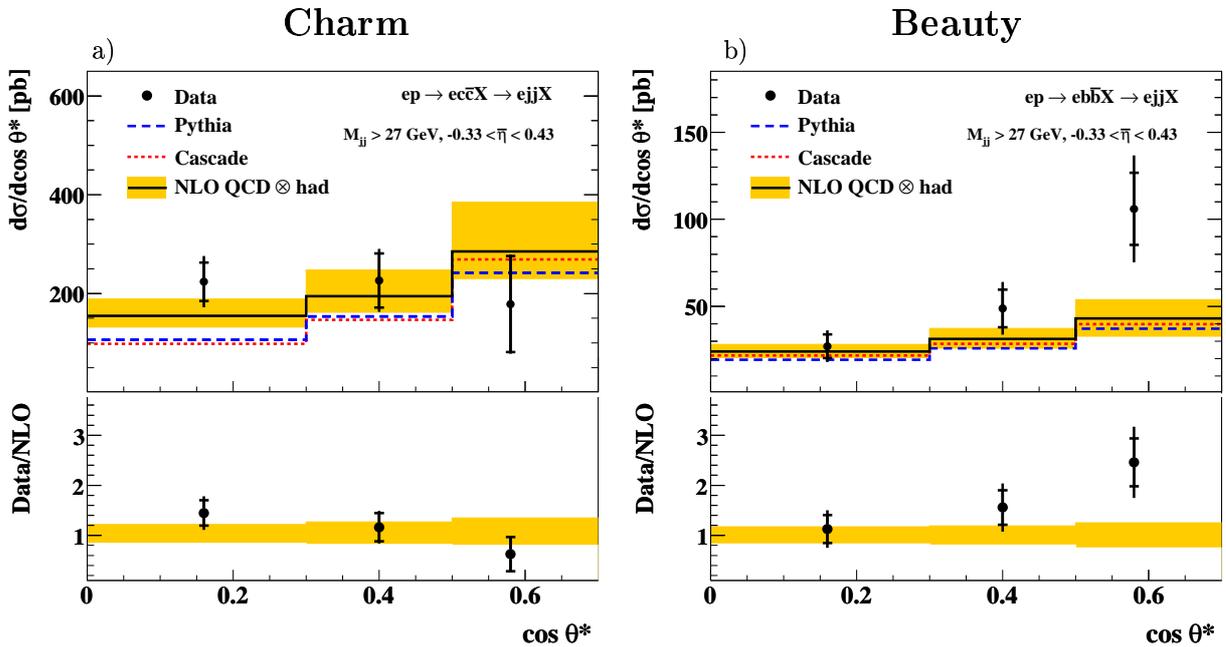
Figures 7.7 a) and b) display the measured differential cross sections as a function of  $x_\gamma^{obs}$ . For resolved processes, the observable  $x_\gamma^{obs}$  is related to the momentum fraction of the photon, which according to equation 4.17 is defined as the fraction of the photon's energy participating in the production of the two highest  $p_t$  jets. The charm and beauty data have a similar shape with a substantial tail towards low values of  $x_\gamma^{obs}$ . In this region, PYTHIA predicts a large contribution from events with resolved photons, as indicated in the figures by the dashed-dotted lines. The contribution of resolved events which are not due to 'heavy quark excitation' in the photon is found to be less than 2% for both charm and beauty.



**Figure 7.6:** Differential charm and beauty photoproduction cross sections a, b)  $d\sigma/dp_t^{jet1}$  and c, d)  $d\sigma/d\eta$  for the process  $ep \rightarrow e(c\bar{c} \text{ or } b\bar{b})X \rightarrow ejjX$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the prediction from a massive NLO QCD calculation, corrected for hadronisation effects, and the shaded band shows the estimated uncertainty. The absolute predictions from PYTHIA (dashed lines) and CASCADE (dotted lines) are also shown.



**Figure 7.7:** Differential charm and beauty photoproduction cross sections a, b)  $d\sigma/dx_\gamma^{obs}$  and c, d)  $d\sigma/d(\Delta\phi)^{jj}$  for the process  $ep \rightarrow e(c\bar{c} \text{ or } b\bar{b})X \rightarrow ejjX$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the prediction from a massive NLO QCD calculation, corrected for hadronisation effects, and the shaded band shows the estimated uncertainty. The absolute predictions from PYTHIA (dashed lines) and CASCADE (dotted lines) are also shown. The contribution from resolved processes in PYTHIA (dash-dotted lines) is depicted separately in figures a) and b).



**Figure 7.8:** Differential charm and beauty photoproduction cross sections a, b)  $d\sigma/d\cos\theta^*$  for the process  $ep \rightarrow e(c\bar{c} \text{ or } b\bar{b})X \rightarrow ejjX$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the prediction from a massive NLO QCD calculation, corrected for hadronisation effects, and the shaded band shows the estimated uncertainty. The absolute predictions from PYTHIA (dashed lines) and CASCADE (dotted lines) are also shown.

PYTHIA describes nicely the shapes of the charm and beauty data distributions, while the normalisations are low. The CASCADE prediction is too small in the region of small  $x_\gamma^{obs}$ , but approaches the measurement in the region  $x_\gamma^{obs} > 0.85$ . The NLO calculation is in good agreement with the charm and beauty data in the highest bin but towards small values of  $x_\gamma^{obs}$  the NLO prediction tends to be below the data. For beauty, the ratio between data and NLO theory is 3.8 in the lowest bin, corresponding to  $1.9\sigma$ .

The measured differential cross section as a function of  $(\Delta\phi)^{jj}$  is shown in figures 7.7 c) and d). The cross section rises towards its maximum close to  $180^\circ$  where both jets are oriented exactly back-to-back in  $r\phi$ . Both, charm and beauty data show a significant contribution at lower values of  $(\Delta\phi)^{jj}$  which can not be accounted for by any of the theories. The shape is again nicely reproduced by PYTHIA, whereas CASCADE is too flat but reaching the measurement at low values of  $(\Delta\phi)^{jj}$ . FMNR is in good agreement with the charm and beauty data at large values of  $(\Delta\phi)^{jj}$  but undershoots the data at low values by far. For charm (beauty) the ratio data over theory in the lowest bin is 6.9 (5.9), corresponding to  $3.1\sigma$  ( $2.4\sigma$ ).

Finally, the charm and beauty dijet cross sections  $d\sigma/d\cos\theta^*$  are shown in figures 7.8 a) and b). The measurement is presented in a phase space region with additional cuts on the dijet mass,  $M_{jj} > 27 \text{ GeV}$ , and the mean pseudo-rapidity,  $-0.33 < \bar{\eta} < 0.43$  (cf. discussion in appendix C). These cuts remove phase space effects due to the jet transverse momentum and pseudo-rapidity requirements and modify the shape of the  $\cos\theta^*$

distribution towards that expected from the QCD matrix elements. Production processes with a gluon propagator are expected to show a steeper rise of the cross section towards large values of  $\cos\theta^*$  than processes with a quark propagator.

Due to the additional cuts the statistical errors of  $d\sigma/d\cos\theta^*$  increases, in particular at large values of  $\cos\theta^*$ . Within the errors the charm cross section agrees with the NLO predictions. However, the measured shape is somewhat flatter than expected from all QCD models. For beauty, the measured cross section is in good agreement with the theories for low values of  $\cos\theta^*$  but suggests a steeper rise towards large values of  $\cos\theta^*$  than predicted by the QCD models. The ratio data over NLO in the last bin yields a factor 2.5, corresponding to  $1.5\sigma$ .

### 7.2.2 Double Differential Cross Sections

Differential cross sections are also measured separately for the region  $x_\gamma^{obs} > 0.85$  and the results as functions of  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  are shown in figures 7.9 and 7.10 and are listed in tables A.3 and A.4.

In the region of large  $x_\gamma^{obs}$  the contribution from resolved photon processes is suppressed and, referring to figure 7.7 a) and b), in this region of phase space the agreement between charm and beauty data with the NLO QCD calculation is expected to improve.

In all measured distributions the charm cross sections are found to be nicely described by the NLO QCD calculation both in normalisation and shape. The beauty cross sections are also reasonably well described, the agreement being clearly better than for the whole range of  $x_\gamma^{obs}$ . A difference between data and NLO prediction appears only at small angles  $(\Delta\phi)^{jj}$  (figure 7.10 b).

In the region of  $x_\gamma^{obs} > 0.85$  differences between PYTHIA and CASCADE become more apparent. As seen in figures 7.9 a) and b) CASCADE predicts a softer spectrum for the transverse momentum distribution than PYTHIA. Also the shapes of the  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  distributions (figures 7.10 a-d) are different for PYTHIA and CASCADE. The statistical precision of the data however is not sufficient to distinguish between both models.

Differential measurements in the region  $x_\gamma^{obs} < 0.85$ , in which the largest deviations and differences of the models are observed, are not feasible due to the dominance of gluon jets with no lifetime information and the reduced statistics in that region of phase space.

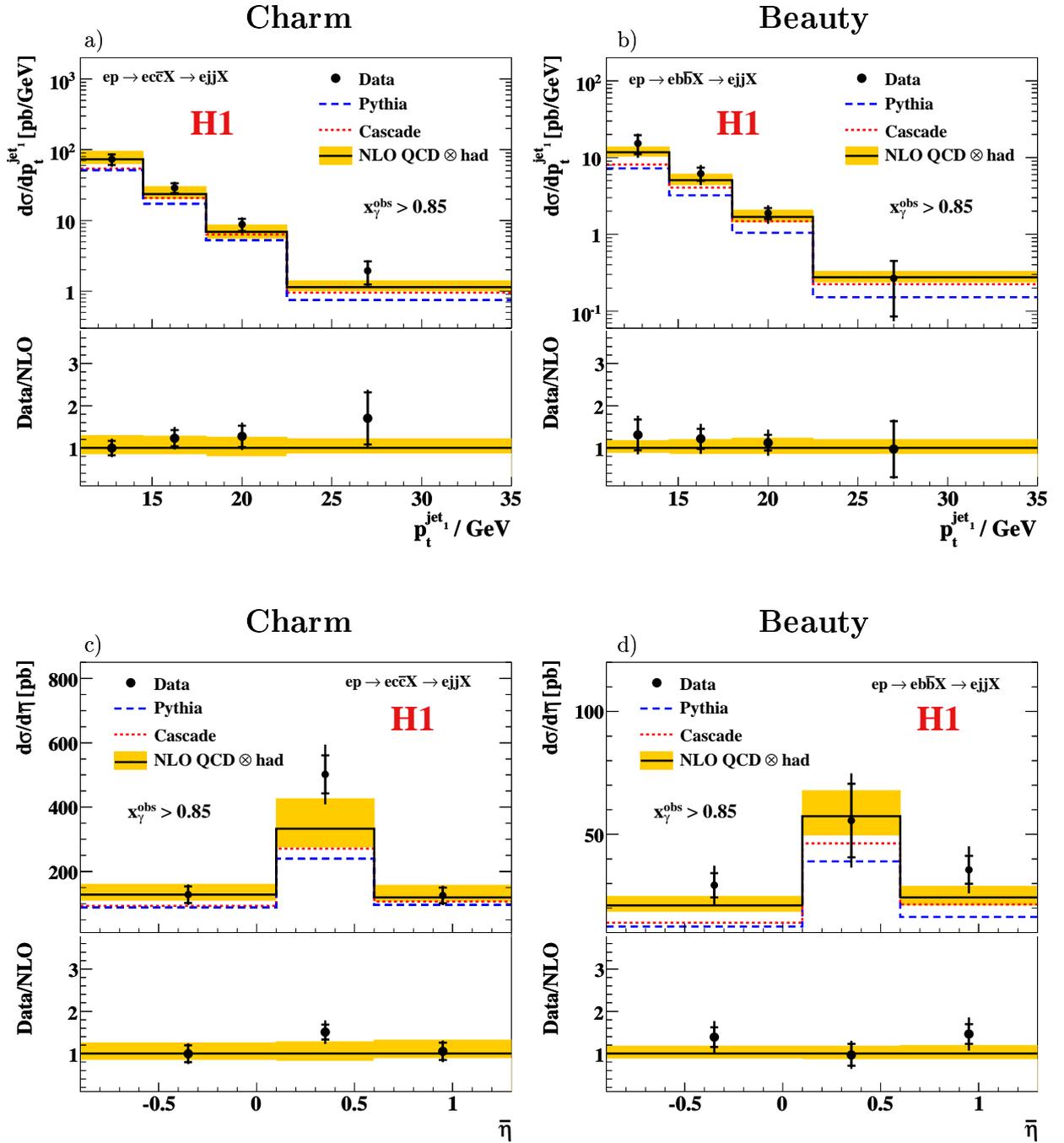
## 7.3 Flavour Inclusive Dijet Cross Section

The flavour inclusive dijet cross section is measured by correcting the observed number of events before track selections for detector efficiencies and acceptances and dividing by the integrated luminosity (equation 5.11).

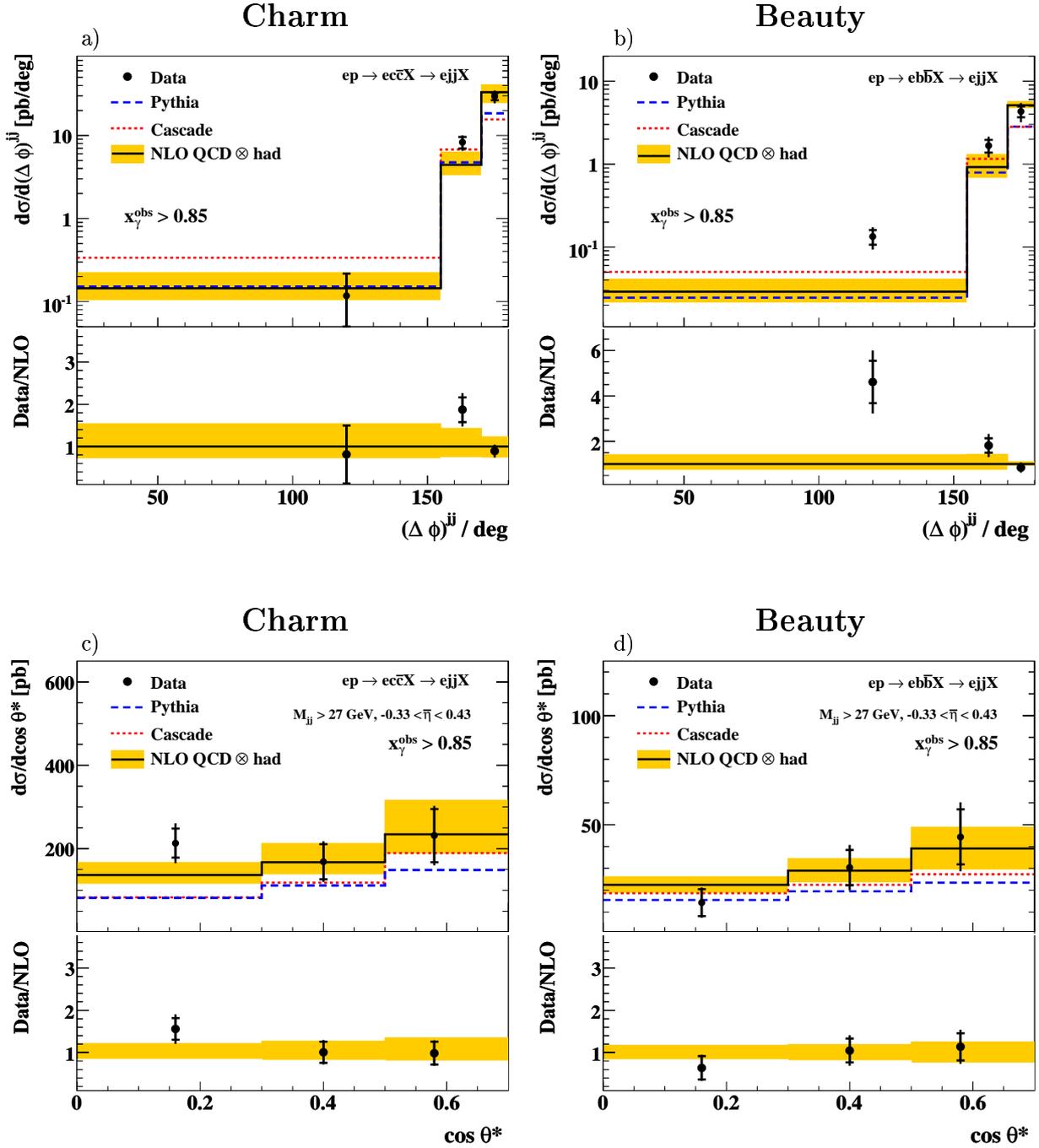
The total flavour inclusive dijet photoproduction cross section in the range  $Q^2 < 1 \text{ GeV}^2$ ,  $0.15 < y < 0.8$ ,  $p_t^{jet1(2)} > 11(8) \text{ GeV}$  and  $-0.9 < \eta^{jet1(2)} < 1.3$  is measured to be

$$\sigma(ep \rightarrow eq\bar{q}X \rightarrow ejjX) = 2261 \pm 8(stat.) \pm 176(syst.) \text{ pb}. \quad (7.3)$$

This result agrees within 2% with the sum of the measured total charm, beauty and light cross sections obtained by multiplying the scale factors  $P_c$ ,  $P_b$  and  $P_l$  with the corresponding cross section predictions of the PYTHIA Monte Carlo simulation. Note, that the latter

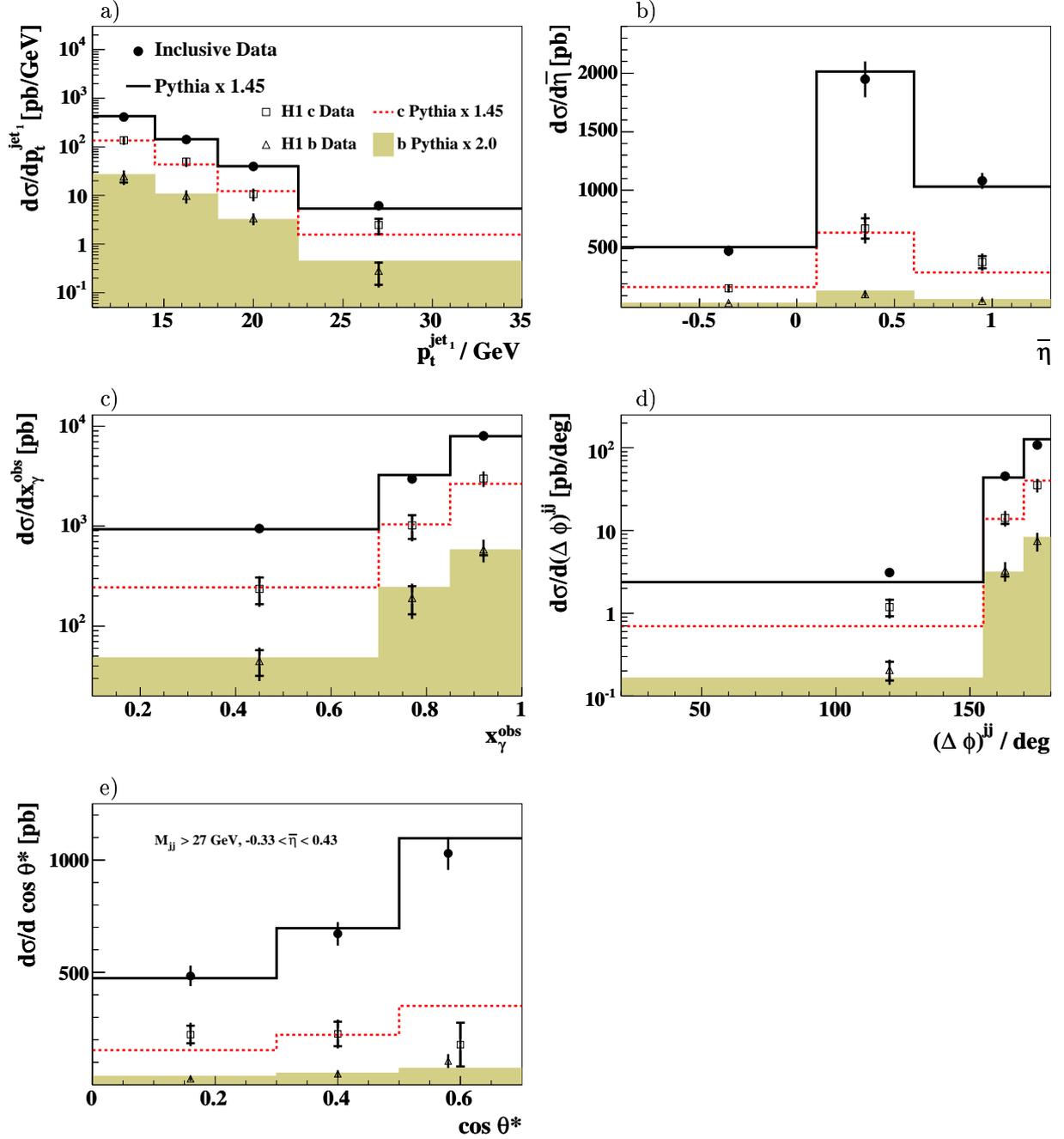


**Figure 7.9:** Differential charm and beauty photoproduction cross sections a, b)  $d\sigma/dp_t^{jet1}$  and c, d)  $d\sigma/d\eta$  for the process  $ep \rightarrow e(c\bar{c} \text{ or } b\bar{b})X \rightarrow ejjX$  in the region  $x_\gamma^{obs} > 0.85$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the prediction from a massive NLO QCD calculation, corrected for hadronisation effects, and the shaded band shows the estimated uncertainty. The absolute predictions from PYTHIA (dashed lines) and CASCADE (dotted lines) are also shown.



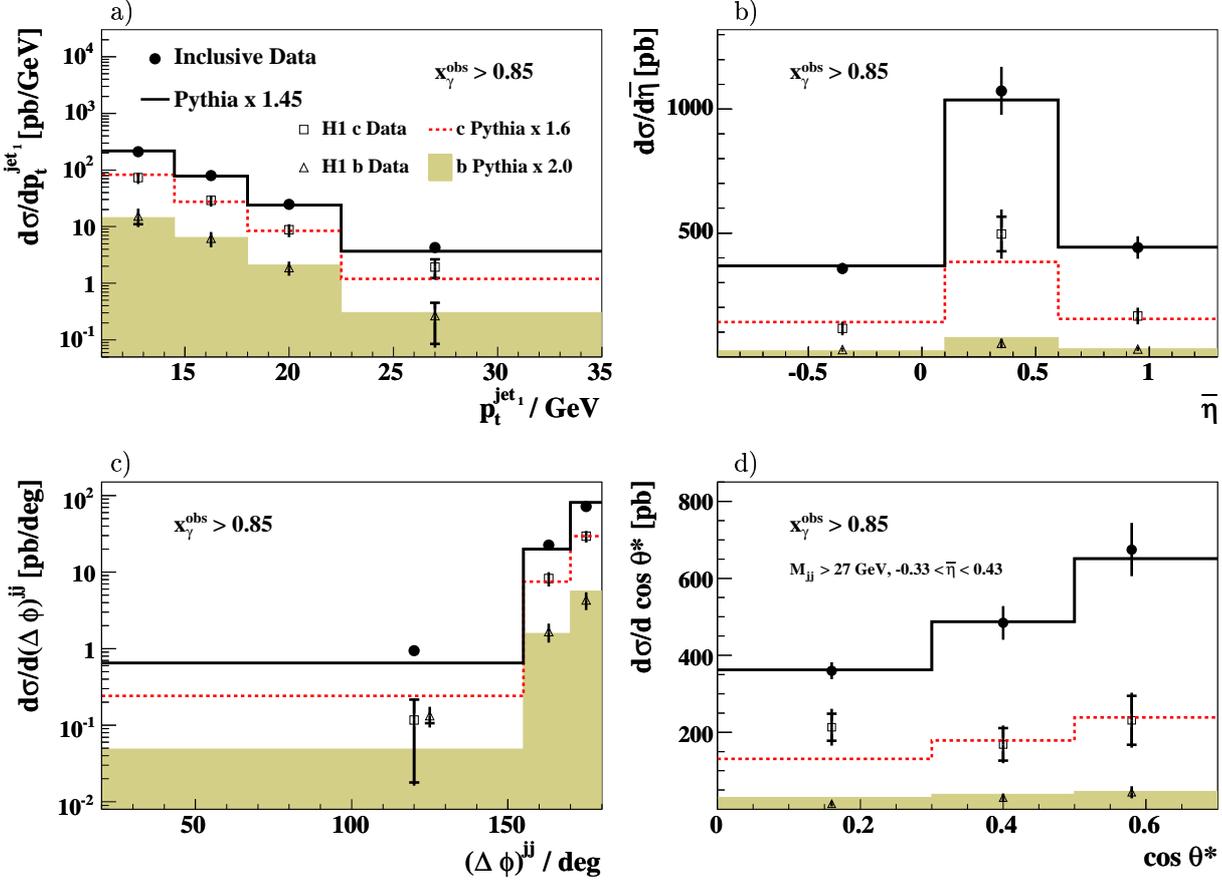
**Figure 7.10:** Differential charm and beauty photoproduction cross sections a, b)  $d\sigma/d(\Delta\phi)^{jj}$  and c, d)  $d\sigma/d\cos\theta^*$  for the process  $ep \rightarrow e(c\bar{c} \text{ or } b\bar{b})X \rightarrow ejjX$  in the region  $x_\gamma^{obs} > 0.85$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the prediction from a massive NLO QCD calculation, corrected for hadronisation effects, and the shaded band shows the estimated uncertainty. The absolute predictions from PYTHIA (dashed lines) and CASCADE (dotted lines) are also shown.

## Inclusive



**Figure 7.11:** Differential flavour inclusive photoproduction cross sections a)  $d\sigma/dp_t^{\text{jet}_1}$ , b)  $d\sigma/d\bar{\eta}$ , c)  $d\sigma/dx_\gamma^{\text{obs}}$ , d)  $d\sigma/d(\Delta\phi)^{jj}$  and e)  $d\sigma/d \cos \theta^*$  for the process  $ep \rightarrow eq\bar{q}X \rightarrow ejjX$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the absolute prediction from PYTHIA scaled with a factor 1.45. The charm (boxes) and beauty (triangles) data are also shown together with the charm and beauty PYTHIA prediction scaled with the factors  $P_c = 1.45$  and  $P_b = 2.0$ .

## Inclusive



**Figure 7.12:** Differential flavour inclusive photoproduction cross sections a)  $d\sigma/dp_t^{jet1}$ , b)  $d\sigma/d\bar{\eta}$ , c)  $d\sigma/d(\Delta\phi)^{jj}$  and d)  $d\sigma/d \cos \theta^*$  for the process  $ep \rightarrow eq\bar{q}X \rightarrow ejjX$  in the region  $x_\gamma^{obs} > 0.85$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic errors added in quadrature. The solid lines indicate the absolute prediction from PYTHIA scaled with a factor 1.45. The charm (boxes) and beauty (triangles) data are also shown together with the charm and beauty PYTHIA prediction scaled with the factors  $P_c(x_\gamma^{obs} > 0.85) = 1.6$  and  $P_b(x_\gamma^{obs} > 0.85) = 2.0$ .

result makes use of the fit of the individual quark contributions and hence depends on the description of the tracking efficiencies and uncertainties whereas the result stated in equation 7.3 is independent of all tracking conditions.

The measured differential flavour inclusive cross sections are presented in figures 7.11 - 7.12 and compared to LO QCD predictions from PYTHIA. In the figures PYTHIA (full line) is scaled by a factor 1.45 which corresponds to the ratio of the measured total cross section and the prediction from PYTHIA<sup>1</sup>. The shape of the data distributions (full dots) is well described by the theory with some discrepancies seen only in the distributions of  $(\Delta\phi)^{jj}$  (figure 7.11 d) and 7.12 c)) where the predicted cross sections are generally steeper than the measured ones.

<sup>1</sup>The original predictions from the LO plus parton shower calculations of PYTHIA and CASCADE are 1559 pb and 1207 pb, respectively.

The charm (open boxes) and beauty (open triangles) cross sections discussed before are also included in the figures together with the scaled predictions from PYTHIA (dotted line for charm and shaded areas for beauty). Here the same scale factors  $P_c$  and  $P_b$  are applied which have been used to extract the differential cross sections. The shapes of the differential charm and beauty dijet cross sections are all reasonably described by the PYTHIA model. Similar to the observations made for the flavour inclusive cross sections the agreement for the observable  $(\Delta\phi)^{jj}$  is worse, showing a somewhat steeper distribution for the theory. While the measured flavour inclusive cross section as a function of  $\cos\theta^*$  is well described by the theory, the measured charm cross section shows a tendency to be flatter than the prediction (figures 7.11 e and 7.12 d).

The relative fractions of charm and beauty events with respect to the flavour inclusive data sample are investigated in the following section.

## 7.4 Charm and Beauty Fractions

To obtain the fractional contributions of events containing charm and beauty quarks, the measured differential cross sections for charm and beauty dijet production are divided by the corresponding flavour inclusive cross sections.

The relative contributions from charm and beauty to the inclusive dijet cross sections are presented in figures 7.13 - 7.14 and listed in tables A.1 - A.4. The charm and beauty fractions are shown with statistical errors including the (small) statistical error of the inclusive cross sections. The systematic errors are calculated based only on those sources that are specific to the charm and beauty cross section measurements.

In the kinematic range of this analysis the total charm fraction is measured to be

$$f^{c\bar{c}} = (31.0 \pm 3.0(\text{stat.}) \pm 3.6(\text{sys.}))\%. \quad (7.4)$$

In the same region of phase space the beauty fraction is measured to be

$$f^{b\bar{b}} = (6.6 \pm 0.8(\text{stat.}) \pm 1.4(\text{sys.}))\%. \quad (7.5)$$

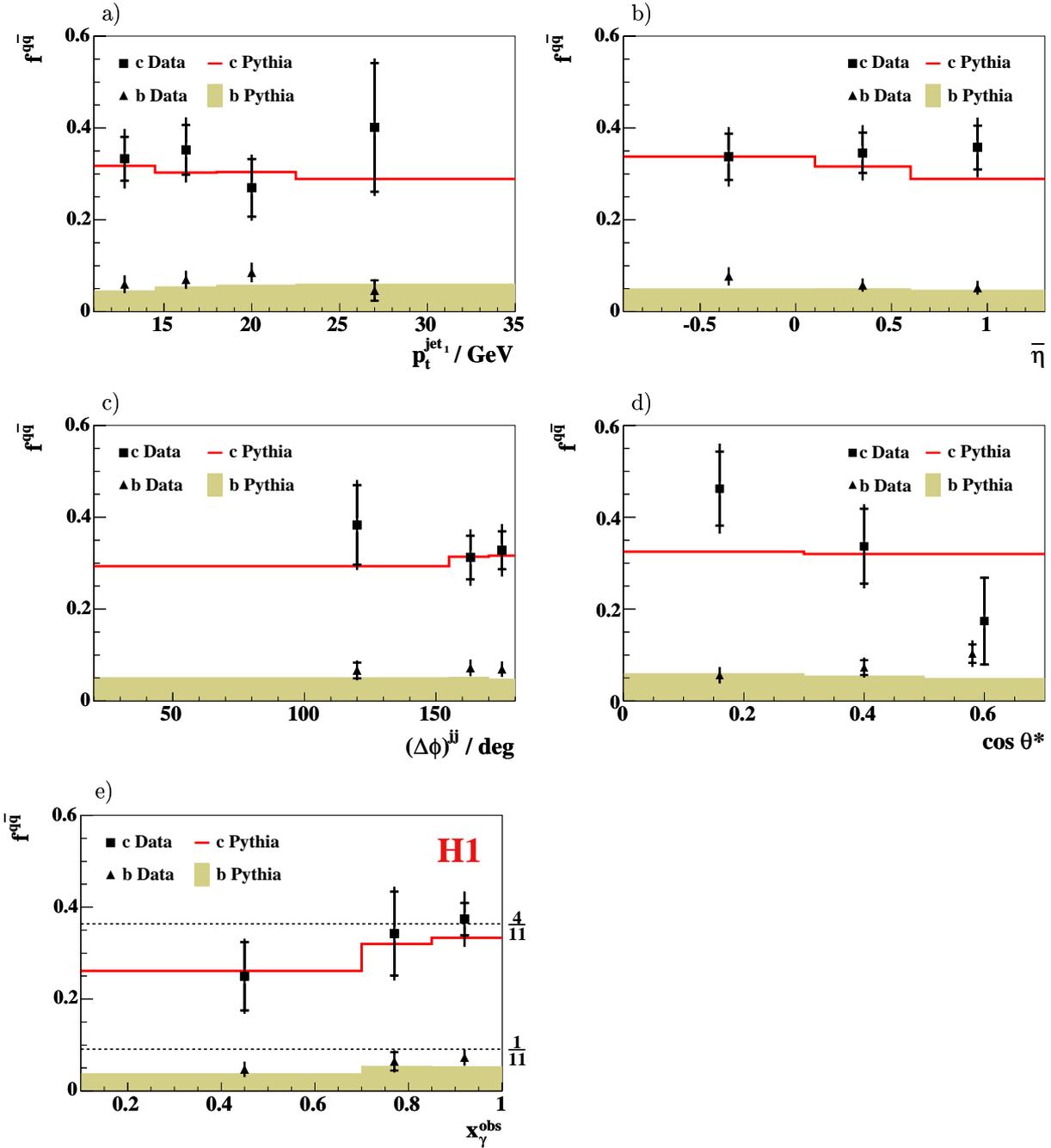
For charm the measured fraction is in very good agreement with the prediction of 31.1% of the PYTHIA Monte Carlo simulation. The predicted beauty fraction of 4.9% however is lower than the measurement.

In figure 7.13 the measured relative contributions as functions of  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $(\Delta\phi)^{jj}$ ,  $\cos\theta^*$  and  $x_\gamma^{obs}$  are compared with the PYTHIA Monte Carlo simulation.

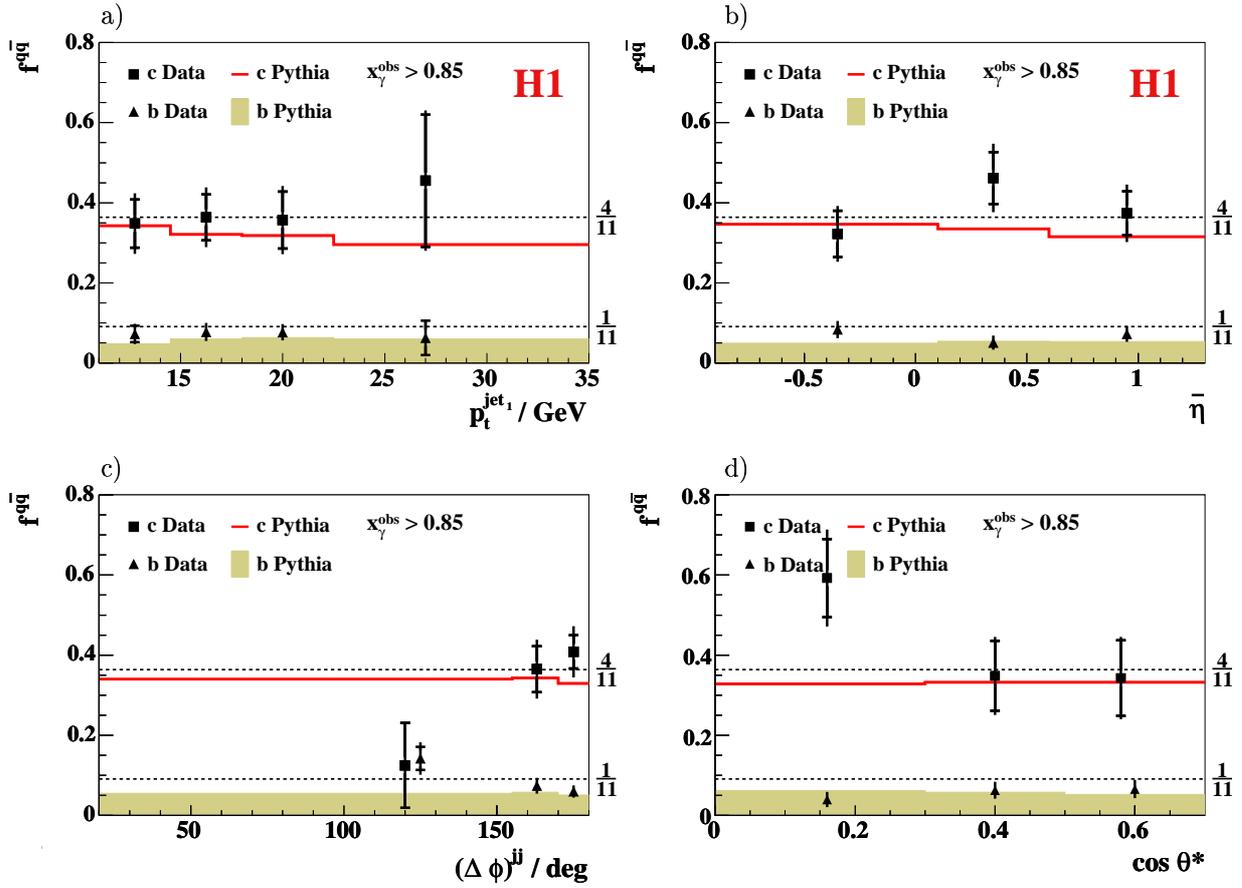
The charm and beauty fractions as functions of  $p_t^{jet1}$ ,  $\bar{\eta}$  and  $(\Delta\phi)^{jj}$  (figure 7.13 a)-c)) show a similar behaviour and the ratios are constant within their uncertainties.

In figure 7.13 d) the charm and beauty fractions are shown as a function of the observable  $\cos\theta^*$ . While the PYTHIA prediction is flat for both quark flavours, the data distributions tend to have a slope with opposite signs for charm and beauty. The measured contribution from charm events at large values of  $\cos\theta^*$  is lower than predicted by PYTHIA. On the other hand the beauty fraction increases for smaller scattering angles. If confirmed this might point to the relative contributions of the various production mechanisms to be different for charm and beauty.

In figure 7.13 e) the relative charm and beauty fractions are shown as a function of the observable  $x_\gamma^{obs}$ . PYTHIA predicts an increase of the relative charm and beauty contributions



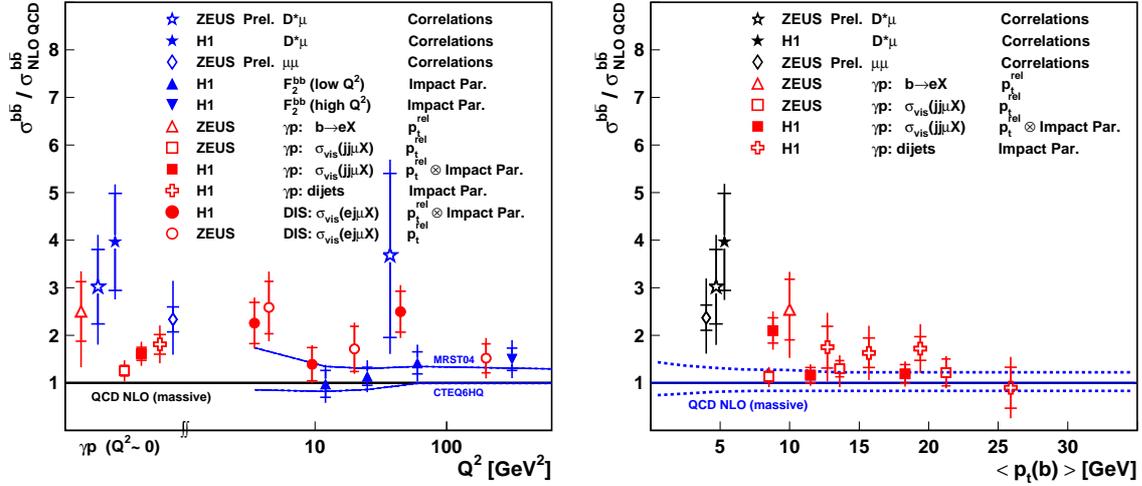
**Figure 7.13:** Relative contributions from charm (squares) and beauty events (triangles) as a function of a) the transverse momentum of the leading jet  $p_t^{\text{jet}_1}$ , b) the mean pseudo-rapidity  $\bar{\eta}$ , c) the azimuthal distance of the two jets  $(\Delta\phi)^{jj}$ , d) the observable  $\cos\theta^*$  and e) the observable  $x_\gamma^{\text{obs}}$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic error added in quadrature. The solid line (shaded area) indicates the absolute prediction from PYTHIA for charm (beauty). The dashed lines in figure e) correspond to the absolute predictions from naïve quark charge counting.



**Figure 7.14:** Relative contributions from charm (squares) and beauty events (triangles) as a function of a) the transverse momentum of the leading jet  $p_t^{jet_1}$ , b) the mean pseudo-rapidity  $\bar{\eta}$ , c) the azimuthal distance of the two jets  $(\Delta\phi)^{jj}$  and d) the variable  $\cos\theta^*$  for the region  $x_\gamma^{obs} > 0.85$ . The inner error bars indicate the statistical uncertainty and the outer error bars show the statistical and systematic error added in quadrature. The solid line (shaded area) indicates the absolute prediction from PYTHIA for charm (beauty). The dashed lines correspond to the absolute predictions from naïve quark charge counting.

towards large values of  $x_\gamma^{obs}$  where direct photon-gluon fusion processes dominate. Both, the charm and the beauty data follow this trend of the theory.

According to equation 2.20 the cross section for the direct photon-gluon fusion process depends strongly on the quark mass. In the kinematic region of this measurement the transverse momentum of the quarks, which is approximated by the transverse momentum of the selected jets, is much larger than the quark masses. The only difference between charm and beauty quarks in equation 2.20 is then their charge. Hence, assuming the charm and beauty quarks to be light, naïve quark charge counting predicts a value of four for the relative production rates of charm to beauty dijets in direct photon-gluon fusion processes. In comparison, the measurement in the region  $x_\gamma^{obs} > 0.85$  yields a ratio of  $5.1 \pm 1.1$  (stat.) taking into account the statistical correlation of both measurements. Normalised to all five possible quark flavours and their corresponding charges the absolute prediction of naïve quark charge counting for the charm and beauty fractions is  $4/11$  and  $1/11$ , respectively. This is shown in with the dashed line in figure 7.13 e).



**Figure 7.15:** Summary of beauty measurements at HERA. The left plot shows the ratio of the cross sections to the (massive) NLO predictions as a function of  $Q^2$ . On the right the dependence on the mean transverse momentum of the  $b$ -quark is shown. The results of this analysis are represented by the open crosses.

In figure 7.14 the relative contributions to the dijet cross section are shown for the region  $x_\gamma^{obs} > 0.85$  as functions of  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$ . Compared to figure 7.13, both charm and beauty fractions are somewhat larger in this kinematic range where contributions from resolved photon processes are suppressed.

Apart from the lowest bin in  $(\Delta\phi)^{jj}$  and in  $\cos\theta^*$ , which have large errors, the ratios are constant within their uncertainties. The measured fractions are consistent with the PYTHIA prediction and the expectation of naïve quark charge counting, which predicts a somewhat larger heavy quark fraction than implemented in PYTHIA.

## 7.5 Comparison with Previous Measurements at HERA

The general observation at HERA that the charm cross sections are reasonably well described by NLO calculations while the beauty cross sections are found to be somewhat higher than NLO QCD is confirmed by this measurement.

For charm, some discrepancies between data and NLO predictions occur only in the regions of small  $x_\gamma^{obs}$  and  $(\Delta\phi)^{jj}$ . Similar observations have been reported in previous analyses on independent data sets from H1 and ZEUS [16, 10].

The measured beauty cross sections are on average a factor 1.8 higher than the NLO predictions with more pronounced differences seen in the regions of low  $x_\gamma^{obs}$ , low  $(\Delta\phi)^{jj}$  and large values of  $\cos\theta^*$ . Figure 7.15 summarises the present status of all beauty measurements performed at HERA. In figures 7.15 a) and b) the ratio of data to NLO theory (massive) is shown as a function of  $Q^2$  and the mean transverse momentum of the beauty quark  $\langle p_T(b) \rangle$ , respectively. The results of the present analysis are shown as open crosses<sup>2</sup>. While the ratio shows no dependence on  $Q^2$ , some trends are visible that the agreement between data and theory improves with larger transverse momentum of the  $b$  quark.

<sup>2</sup>The measured jet cross sections as a function of  $p_t^{jet1}$  are corrected to the mean transverse momentum of the beauty quark according to the PYTHIA Monte Carlo simulation.

---

In recent publications of H1 [5, 6] the relative charm and beauty fractions  $f^{c\bar{c}}$  and  $f^{b\bar{b}}$  have been measured for the regions of low and high  $Q^2$ . The experimental techniques applied in those analyses are similar to those used here. While in the present measurement in photoproduction  $f^{c\bar{c}}$  and  $f^{b\bar{b}}$  are seen to be almost constant over the entire region of phase space, the charm and beauty fractions presented in [5, 6] show a  $Q^2$ -dependence, which is particularly strong for beauty. Due to quark mass effects  $f^{q\bar{q}}$  rises with increasing  $Q^2$ . Such quark mass effects are not visible in the case of the present photoproduction analysis which requires the transverse momentum of the quarks to exceed the beauty quark mass by a factor of two.

## CHAPTER 8

# CONCLUSIONS AND OUTLOOK

---

Charm and beauty dijet events in photoproduction at HERA are investigated using a technique based on the lifetime of the heavy quark hadrons. The contributions of charm and beauty quark events to the flavour inclusive cross section are determined using the subtracted impact parameter significance distributions of tracks in dijet events. For the first time at HERA differential dijet cross sections in photoproduction are measured simultaneously for charm and beauty.

The cross sections are measured as functions of the transverse momentum  $p_t^{jet_1}$  of the leading jet, of the mean pseudo-rapidity  $\bar{\eta}$  of the two jets, of the azimuthal separation  $(\Delta\phi)^{jj}$  of both jets, the dijet centre-of-mass scattering angle  $\cos\theta^*$  and of the observable  $x_\gamma^{obs}$ . In general, the measured charm and beauty distributions show a similar shape in all variables indicating that kinematic effects due to different quark masses are small in the kinematic domain considered. Some smaller differences in shape are observed for the distribution of  $\cos\theta^*$  which shows for beauty a steep rise with increasing values of  $\cos\theta^*$ , while it is flat for charm.

The results are compared with the QCD models as implemented in the leading order plus parton shower Monte Carlo simulation programs PYTHIA and CASCADE and the massive next-to-leading order (NLO) QCD calculation FMNR which is presently the only NLO prediction available for the production processes under study. Theoretical uncertainties arising from contributions beyond NLO are estimated by scale variations and found to be particularly large for charm. The total theory error amounts to 35% (23%) for charm (beauty) compared to a total experimental error of 17% (25%).

Taking into account the theoretical uncertainties, the charm cross sections are found to be consistent both in normalisation and shape with the perturbative QCD next-to-leading order predictions. The beauty cross sections tend to be higher than NLO by a factor of 1.8 ( $1.6\sigma$ ) for the total cross section, with an increased difference observed in the region of small  $x_\gamma^{obs}$ , small  $(\Delta\phi)^{jj}$  and large  $\cos\theta^*$ , where processes involving resolved photons or higher order contributions are expected to be enhanced. Significant differences between the measured charm and beauty cross sections and the NLO predictions are found in the region of small values of  $(\Delta\phi)^{jj}$  ( $3.1\sigma$  and  $2.4\sigma$  for charm and beauty, respectively).

The LO plus parton shower calculations of PYTHIA and CASCADE predict total charm and beauty cross sections similar to that of the NLO calculation. For charm, PYTHIA and CASCADE fall below the data in normalisation by factors 1.45 and 1.6, respectively. The

corresponding ratios for beauty are 2.0 for PYTHIA and 1.9 for CASCADE.

Differential cross sections as functions of  $p_t^{jet1}$ ,  $\bar{\eta}$ ,  $(\Delta\phi)^{jj}$  and  $\cos\theta^*$  are also measured separately for the region  $x_\gamma^{obs} > 0.85$  in which the contribution of direct photon-gluon fusion processes dominate. The charm cross sections are in good agreement with the NLO QCD predictions. The beauty cross sections are also reasonably well described, the level of agreement being considerably better than for the whole range of  $x_\gamma^{obs}$ .

The Monte Carlo generator PYTHIA gives a good description of the shape of both the charm and the beauty data. The data do confirm the assumption used in this model that a significant contribution to the charm and beauty quark dijet cross section comes from processes with resolved photons. These findings are supported by a detailed study of a heavy quark enriched data sample, which was selected based on the explicit reconstruction of secondary vertices. A good description of this subsample can be achieved by including the contributions from resolved photons. According to PYTHIA, in the region  $x_\gamma^{obs} < 0.85$  the contribution from resolved photons is enhanced to about 80%, with the dominant contribution resulting from processes with charm and beauty quark excitation.

The CASCADE model includes  $k_t$  unordered gluon radiation in the parton evolution. This leads to radiative processes with similar final state configurations as the excitation processes implemented in the PYTHIA model. The CASCADE model is disfavored by the shape of the measured charm and beauty differential cross sections as a function of  $x_\gamma^{obs}$ , but gives the best prediction of all models considered for the region of small values of  $(\Delta\phi)^{jj}$ , which is sensitive to the transverse momentum  $k_t$  of the partons before the hard interaction.

The transverse energy flow relative to the jet axis as a function of the azimuth  $\phi$  and the pseudo-rapidity has been studied for different regions of  $p_t^{jet1}$  and  $x_\gamma^{obs}$ . While the flavour inclusive data can be described nicely by PYTHIA including multiple parton interactions, the heavy quark data show some deviations in particular in the region of small  $x_\gamma^{obs}$  in which resolved photon processes are important. The measurement of the mean integrated jet shapes  $\langle\Psi(r)\rangle$  support these findings. The jet shapes are generally well described by the PYTHIA simulation for both charm and beauty, with some deviations only in the region of small  $x_\gamma^{obs}$  where the data tend to have larger values of  $\langle\Psi(r)\rangle$  than predicted. In this region the PYTHIA model predicts a significantly increased number of gluon jets resulting from heavy quark excitation processes which lead to smaller values of  $\langle\Psi(r)\rangle$  than for quark initiated jets.

The flavour inclusive dijet cross sections are also measured in the same phase space. In the region  $x_\gamma^{obs} > 0.85$ , the relative charm and beauty contributions to the inclusive dijet cross sections are found to be consistent with values of 4/11 and 1/11, respectively, the naïve expectation for the direct photon-gluon fusion process assuming massless quarks.

## Outlook

Further studies of the heavy quark production mechanism and a more precise test of the various QCD models require a reduction of the theoretical scale uncertainties, as well as the statistical and systematic uncertainties in the data. This would be particularly useful in view of the upcoming experiments at the LHC, which are currently in their final construction phase.

The HERA II data sample, which already exceeds the data sample studied here, will allow a reduction of the data uncertainties and the measurement of more differential cross

sections. Furthermore, the installation of a dedicated photoproduction jet trigger will allow the study of heavy quark jets in a sizeable wider transverse energy range. In order to get more insight into resolved photon processes, which are expected to be enhanced at large pseudo-rapidities, the inclusion of the *Forward Silicon Tracker* data would be desirable. Next-to-leading order calculations performed in the appropriate schemes and with the appropriate scales would significantly improve the interpretation of the results. A promising step in this direction is the MC@NLO generator which matches parton showers to NLO matrix elements. These calculations are presently not yet available for HERA.

## APPENDIX A

### DATA TABLES

---

		Charm						
$p_t^{jet1}$ range [GeV]	$\langle p_t^{jet1} \rangle$ [GeV]	$d\sigma/dp_t^{jet1}$ [pb/GeV]	stat.	syst.		$f^{c\bar{c}}$	stat.	syst.
11.0 14.5	12.75	137	20	20		0.333	0.048	0.044
14.5 18.0	16.25	49.5	7.6	7.9		0.352	0.054	0.046
18.0 22.5	20.0	10.6	2.5	1.8		0.270	0.062	0.035
22.5 35.0	27.0	2.46	0.86	0.45		0.402	0.140	0.052
$\bar{\eta}$ range	$\langle \bar{\eta} \rangle$	$d\sigma/d\bar{\eta}$ [pb]	stat.	syst.		$f^{c\bar{c}}$	stat.	syst.
-0.90 0.10	-0.35	162	24	22		0.337	0.051	0.045
0.10 0.60	0.35	674	86	101		0.346	0.044	0.045
0.60 1.30	0.95	386	52	61		0.358	0.049	0.047
$x_\gamma^{obs}$ range	$\langle x_\gamma^{obs} \rangle$	$d\sigma/dx_\gamma^{obs}$ [pb]	stat.	syst.		$f^{c\bar{c}}$	stat.	syst.
0.10 0.70	0.45	236	70	34		0.250	0.074	0.033
0.70 0.85	0.77	1017	271	162		0.342	0.091	0.047
0.85 1.00	0.92	2994	280	498		0.374	0.035	0.049
$(\Delta\phi)^{jj}$ range [deg]	$\langle (\Delta\phi)^{jj} \rangle$ [deg]	$d\sigma/d(\Delta\phi)^{jj}$ [pb/deg]	stat.	syst.		$f^{c\bar{c}}$	stat.	syst.
20 155	120	1.19	0.27	0.16		0.383	0.087	0.046
155 170	163	14.17	2.14	2.13		0.312	0.047	0.039
170 180	175	35.46	4.50	5.61		0.328	0.041	0.039
$\cos\theta^*$ range	$\langle \cos\theta^* \rangle$	$d\sigma/d\cos\theta^*$ [pb]	stat.	syst.		$f^{c\bar{c}}$	stat.	syst.
0.0 0.3	0.16	224	39	30		0.462	0.081	0.056
0.3 0.5	0.4	226	55	34		0.337	0.081	0.042
0.5 0.7	0.58	179	97	28		0.174	0.094	0.021

**Table A.1:** The measured charm dijet photoproduction cross sections and the relative contributions  $f^{c\bar{c}}$  to the inclusive dijet photoproduction cross section with statistical and systematic errors.

		Beauty						
$p_t^{jet1}$ range [GeV]	$\langle p_t^{jet1} \rangle$ [GeV]	$d\sigma/dp_t^{jet1}$ [pb/GeV]	stat.	syst.	$f^{b\bar{b}}$	stat.	syst.	
11.0 14.5	12.75	24.7	6.1	5.5	0.060	0.015	0.013	
14.5 18.0	16.25	9.79	1.87	2.15	0.070	0.013	0.015	
18.0 22.5	20.0	3.37	0.48	0.78	0.085	0.012	0.018	
22.5 35.0	27.0	0.28	0.14	0.07	0.046	0.022	0.010	
$\bar{\eta}$ range	$\langle \bar{\eta} \rangle$	$d\sigma/d\bar{\eta}$ [pb]	stat.	syst.	$f^{b\bar{b}}$	stat.	syst.	
-0.90 0.10	-0.35	37.1	5.9	7.8	0.077	0.012	0.016	
0.10 0.60	0.35	112	17	25	0.057	0.009	0.012	
0.60 1.30	0.95	56.2	11.3	12.4	0.052	0.010	0.011	
$x_\gamma^{obs}$ range	$\langle x_\gamma^{obs} \rangle$	$d\sigma/dx_\gamma^{obs}$ [pb]	stat.	syst.	$f^{b\bar{b}}$	stat.	syst.	
0.10 0.70	0.45	44.6	12.8	9.9	0.047	0.014	0.010	
0.70 0.85	0.77	191	60	44	0.064	0.020	0.014	
0.85 1.00	0.92	584	72	135	0.073	0.009	0.015	
$(\Delta\phi)^{jj}$ range [deg]	$\langle (\Delta\phi)^{jj} \rangle$ [deg]	$d\sigma/d(\Delta\phi)^{jj}$ [pb/deg]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.	
20 155	120	0.206	0.053	0.043	0.067	0.017	0.013	
155 170	163	3.27	0.50	0.72	0.072	0.011	0.015	
170 180	175	7.48	1.07	1.66	0.069	0.010	0.014	
$\cos\theta^*$ range	$\langle \cos\theta^* \rangle$	$d\sigma/d\cos\theta^*$ [pb]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.	
0.0 0.3	0.16	27.2	6.8	5.7	0.056	0.014	0.011	
0.3 0.5	0.4	48.9	10.8	10.7	0.073	0.016	0.015	
0.5 0.7	0.58	106	20.6	23.5	0.103	0.020	0.021	

**Table A.2:** The measured beauty dijet photoproduction cross sections and the relative contributions  $f^{b\bar{b}}$  to the inclusive dijet photoproduction cross section with statistical and systematic errors.

$x_\gamma^{obs} > 0.85$			Charm					
$p_t^{jet1}$ range [GeV]	$\langle p_t^{jet1} \rangle$ [GeV]		$d\sigma/dp_t^{jet1}$ [pb/GeV]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.
11.0 14.5	12.75		73.3	12.7	10.7	0.348	0.060	0.046
14.5 18.0	16.25		29.0	4.6	4.6	0.364	0.057	0.048
18.0 22.5	20.0		8.79	1.75	1.46	0.357	0.071	0.047
22.5 35.0	27.0		1.94	0.70	0.35	0.455	0.165	0.060
$\bar{\eta}$ range	$\langle \bar{\eta} \rangle$		$d\sigma/d\bar{\eta}$ [pb]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.
-0.90 0.10	-0.35		115	21	16	0.322	0.058	0.043
0.10 0.60	0.35		496	70	74	0.462	0.065	0.060
0.60 1.30	0.95		165	24	26	0.374	0.055	0.049
$(\Delta\phi)^{jj}$ range [deg]	$\langle (\Delta\phi)^{jj} \rangle$ [deg]		$d\sigma/d(\Delta\phi)^{jj}$ [pb/deg]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.
20 155	120		0.118	0.099	0.160	0.125	0.106	0.015
155 170	163		8.26	1.28	1.24	0.365	0.057	0.046
170 180	175		29.5	2.99	4.66	0.408	0.041	0.049
$\cos\theta^*$ range	$\langle \cos\theta^* \rangle$		$d\sigma/d\cos\theta^*$ [pb]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.
0.0 0.3	0.16		213	35	29	0.592	0.097	0.072
0.3 0.5	0.4		169	42	25	0.348	0.087	0.044
0.5 0.7	0.58		231	8	37	0.343	0.094	0.041

**Table A.3:** The measured charm dijet photoproduction cross sections and the relative contributions  $f^{c\bar{c}}$  to the inclusive dijet photoproduction cross section for the region  $x_\gamma^{obs} > 0.85$  with statistical and systematic errors.

$x_\gamma^{obs} > 0.85$			Beauty					
$p_t^{jet1}$ range [GeV]	$\langle p_t^{jet1} \rangle$ [GeV]		$d\sigma/dp_t^{jet1}$ [pb/GeV]	stat.	syst.	$f^{b\bar{b}}$	stat.	syst.
11.0 14.5	12.75		15.3	4.3	3.4	0.073	0.020	0.015
14.5 18.0	16.25		6.19	1.20	1.42	0.078	0.015	0.016
18.0 22.5	20.0		1.89	0.31	0.44	0.076	0.013	0.016
22.5 35.0	27.0		0.27	0.18	0.07	0.063	0.043	0.014
$\bar{\eta}$ range	$\langle \bar{\eta} \rangle$		$d\sigma/d\bar{\eta}$ [pb]	stat.	syst.	$f^{b\bar{b}}$	stat.	syst.
-0.90 0.10	-0.35		29.8	4.7	6.3	0.084	0.013	0.018
0.10 0.60	0.35		54.4	15.5	12.0	0.051	0.014	0.011
0.60 1.30	0.95		31.9	5.7	7.1	0.072	0.013	0.016
$(\Delta\phi)^{jj}$ range [deg]	$\langle (\Delta\phi)^{jj} \rangle$ [deg]		$d\sigma/d(\Delta\phi)^{jj}$ [pb/deg]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.
20 155	120		0.134	0.027	0.028	0.142	0.029	0.028
155 170	163		1.67	0.29	0.37	0.074	0.013	0.015
170 180	175		4.31	0.64	0.95	0.060	0.009	0.012
$\cos\theta^*$ range	$\langle \cos\theta^* \rangle$		$d\sigma/d\cos\theta^*$ [pb]	stat.	syst.	$f^{c\bar{c}}$	stat.	syst.
0.0 0.3	0.16		14.3	6.2	3.0	0.040	0.018	0.008
0.3 0.5	0.4		30.4	8.1	6.7	0.063	0.017	0.013
0.5 0.7	0.58		44.4	12.6	9.8	0.065	0.019	0.013

**Table A.4:** The measured beauty dijet photoproduction cross sections and the relative contributions  $f^{b\bar{b}}$  to the inclusive photoproduction dijet cross section for the region  $x_\gamma^{obs} > 0.85$  with statistical and systematic errors.

## APPENDIX B

# CHARM CROSS SECTION WITH $D^*$ MESONS

---

The result of the measured total charm cross section using the  $(S_1, S_2)$ -method is cross checked with events in which a  $D^*$  meson is explicitly reconstructed. The standard mass difference technique is applied in which the  $D^*$  is detected via its golden decay channel  $D^{*+} \rightarrow D^0 \pi_S^+ \rightarrow K^- \pi^+ \pi_S^+$ . Only a brief description of this procedure is given here.

### Decay Mode

The probability of the hadronisation of a charm quark into a  $D^{*+}$  meson is experimentally determined to be  $f(c \rightarrow D^{*+}) = 0.235 \pm 0.007 \pm 0.007$  [49]. The  $D^{*+}$  meson is an excited  $c\bar{d}$  state decaying mainly into a  $D^0$  meson and a charged pion via the strong interaction, i.e. with a very short lifetime. Since the mass difference between the  $D^*$  and the  $D^0$  is only slightly above the pion mass, the decay pion has only a very small momentum. This is also true in the laboratory frame. For this reason the 'slow' pion of the  $D^*$  decay is usually written as  $\pi_S^+$ .

The  $D^0$  meson then decays weakly with several possible decay modes. In this case only the decay  $D^0 \rightarrow K^- \pi^+$  is considered. The total branching ratio for the  $D^*$  mesons in the so called 'golden decay channel' amounts to

$$BR(D^{*+} \rightarrow K^- \pi^+ \pi_S^+) = (2.57 \pm 0.06)\%. \quad (\text{B.1})$$

### Reconstruction and Selection

The  $D^*$  mesons are reconstructed via the decay chain  $D^{*+} \rightarrow K^- \pi^+ \pi_S^+$  and the charge conjugate process. Note, that only primary vertex fitted tracks are taken into account since this cross check is based on a standard H1 reconstruction technique. Tracks in the central tracking devices with opposite charges are combined in pairs to form  $D^0$  candidates. Each of the two tracks is alternatively assumed to be a kaon or pion, i.e. without employing any particle identification. Then an additional slow track with opposite track charge to that of the kaon is assumed to be the slow pion and is combined with the two tracks of the  $D^0$  meson to form a  $D^*$  candidate. Events with a mass difference  $\Delta m = m_{K\pi\pi_S} - m_{K\pi}$  in the

$D^*$ track selection
$p_t^{D^*} > 3 \text{ GeV}$ $ \eta^{D^*}  < 1.5$
$p_t^\pi > 0.3 \text{ GeV}$ $p_t^K > 0.5 \text{ GeV}$ $ m_{K\pi} - m_{D^0}  < 0.08 \text{ GeV}$ $\frac{p_t^{D^*}}{\sum_{HFS}^{\Theta > 10^\circ} E_i \sin \Theta_i} > 0.1$

**Table B.1:**  $D^*$  selection criteria applied to the dijet event sample.

range  $0.143 < \Delta m < 0.148 \text{ GeV}$  and the range  $1.82 < m_{K\pi} < 1.90 \text{ GeV}$  around the  $D^0$  mass are called  $D^*$  candidates.

The  $D^*$  candidate is required to have a transverse momentum of  $p_t^{D^*} > 3.0 \text{ GeV}$  in the central pseudo-rapidity range  $|\eta^{D^*}| < 1.5$ . Table B.1 summarises all cuts applied to the  $(K, \pi, \pi_S)$  system.

In addition to the cuts on the transverse momentum of the  $D^*$  meson and its decay tracks, the cut  $p_t^{D^*} / \sum_{HFS}^{\Theta > 10^\circ} E_i \sin \Theta_i > 0.1$  suppresses further the combinatorial background. The denominator contains the transverse energy of all HFS objects outside a cone of  $\Theta = 10^\circ$  in the forward direction.

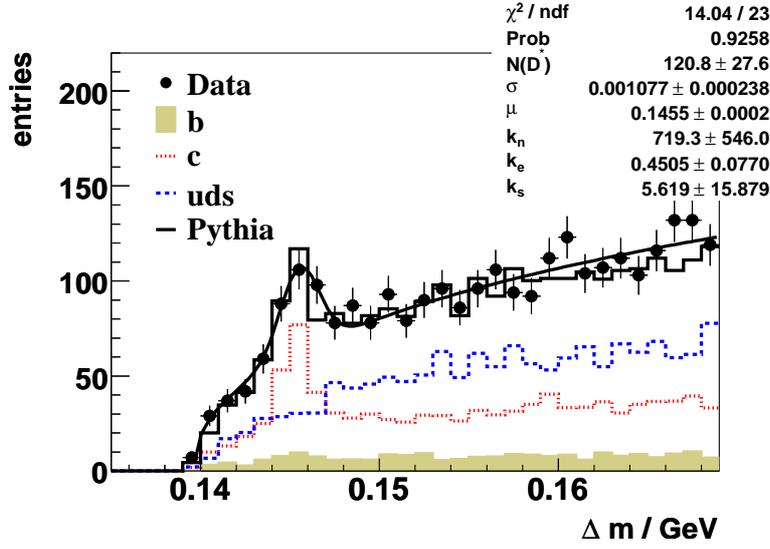
### $D^*$ Mesons in the Dijet Photoproduction Sample

The  $D^*$  meson sample obtained when applying the selection criteria mentioned above to the dijet photoproduction sample as defined in table 5.2 is visualised in figure B.1. The data is compared to the PYTHIA Monte Carlo simulation with the contributions from charm, beauty and light quarks shown separately. The Monte Carlo prediction is scaled by the individual scale factors from the fit to the subtracted impact parameter significance distributions. The resonance in the  $\Delta m$  distribution at  $\sim 0.145 \text{ GeV}$  indicates the presence of successfully reconstructed  $D^*$  mesons. The signal sits on top of a smaller, steeply rising background. According to the Monte Carlo simulation only a small fraction of beauty events contribute to the signal.

Since the combinatorial background also passes the applied cuts, the fit function is a sum of a Gaussian for the signal and a background term. The number of  $D^*$  mesons  $N^{D^*}$  is determined by fitting the function

$$f(\Delta m) = \frac{N^{D^*}}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\Delta m - \mu)^2}{2\sigma^2}\right) + k_n(\Delta m - m_\pi)^{k_e} \cdot (1 - k_s(\Delta m)^2) \quad (\text{B.2})$$

to the measured  $\Delta m$  distribution. Here,  $\sigma$  and  $\mu$  are the width and the mean of the signal and the parameters  $k_n, k_e, k_s$  are determined from the fit. The fit result is shown in figure B.1 as the full smooth line. The number of  $D^*$  mesons reconstructed in the signal region from the fit is found to be  $121 \pm 28$  at the mass difference  $m_{D^*} - m_{D^0} = 145.5 \pm 0.2 \text{ MeV}$ .



**Figure B.1:** Distribution of the mass difference of the  $D^*$  an  $D^0$  candidates  $\Delta m = m_{K\pi\pi_S} - m_{K\pi}$ . The number of  $D^*$  events in the untagged photoproduction dijet sample determined by the fit is  $N^{D^*} = 121 \pm 28$ .

Taking into account the total branching ratio (equation B.1) the charm cross section can be extrapolated from the number of observed  $D^*$  mesons

$$\sigma(ep \rightarrow ec\bar{c}X \rightarrow ejjX) = \frac{N^{D^*}}{BR(D^{*+} \rightarrow K^- \pi^+ \pi_S^+) \cdot \mathcal{L} \cdot \varepsilon A}. \quad (\text{B.3})$$

The efficiency and acceptance  $\varepsilon A$  is calculated from the Monte Carlo simulation with the number of reconstructed  $D^*$  mesons determined from the same fit (equation B.2) as used for the data. This is possible because the combinatorial background is well modelled by the simulation as can be seen in figure B.1.

Applying these numbers to equation B.3, the total charm dijet photoproduction cross section in the visible range of this measurement (cf. equation 5.7) amounts to

$$\sigma(ep \rightarrow ec\bar{c}X \rightarrow ejjX) = (692 \pm 158) \text{ pb}, \quad (\text{B.4})$$

with the error denoting the statistical uncertainty of the fit.

These findings confirm nicely the result obtained from the fit to the subtracted impact parameter significance distributions (cf. equation 7.1).

## APPENDIX C

# POLAR ANGULAR DISTRIBUTION $\cos \theta^*$

In this section the  $\cos \theta^*$  distribution and its dependence on kinematic constraints, such as the jet transverse momenta and pseudo-rapidities, is investigated.

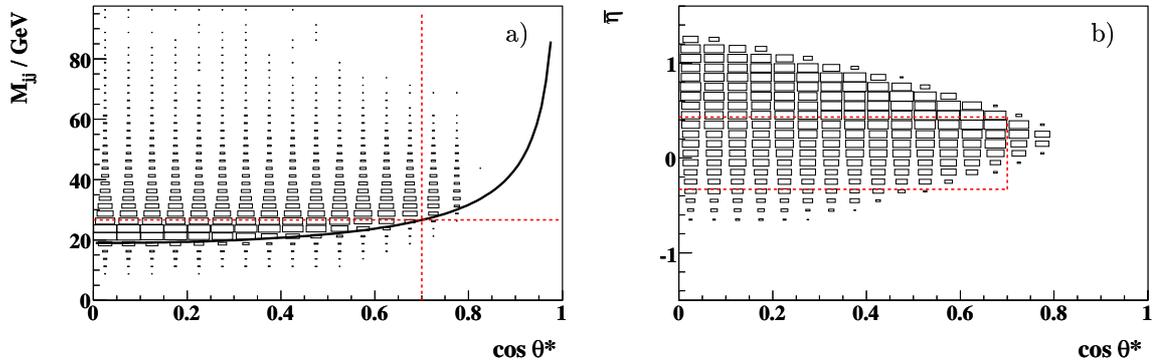
As defined in equation 4.16 the dijet centre-of-mass scattering angle  $\theta^*$  is reconstructed using

$$\cos(\theta^*) = |\tanh((\eta^{jet_1} - \eta^{jet_2})/2)|. \quad (\text{C.1})$$

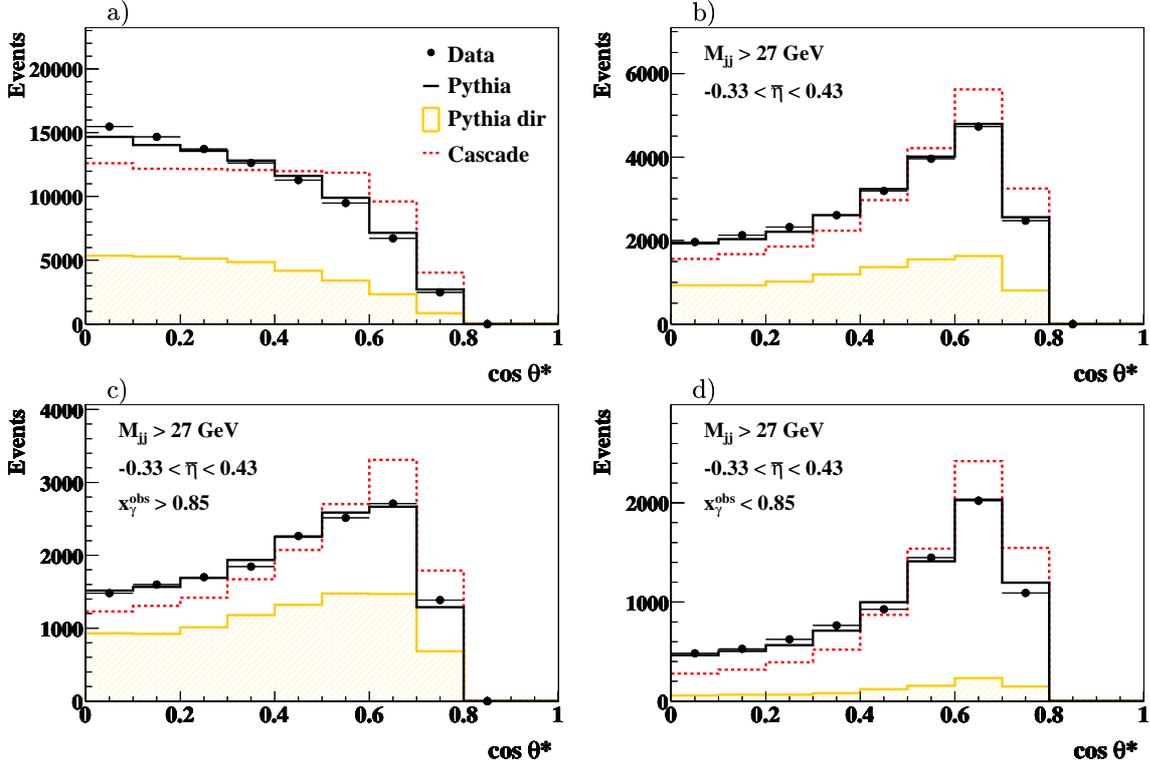
In the simple case in which two jets with equal transverse momenta  $p_t^{jet}$  are back-to-back in the transverse plane, the following equation holds for the relation between  $\theta^*$ ,  $p_t^{jet}$  and the invariant mass  $M_{jj}$  of the dijet system:

$$M_{jj} = \frac{2 \cdot p_t^{jet}}{\sqrt{1 - |\cos \theta^*|^2}}. \quad (\text{C.2})$$

Hence, for a given  $M_{jj}$  events with high values of  $\cos \theta^*$  have a lower  $p_t^{jet}$ . In figure C.1 a) the correlation between  $\cos \theta^*$  and  $M_{jj}$  is shown for the selected dijet events (cf. table 5.2). The solid curve corresponds to equation C.2 for jet transverse momenta  $p_t^{jet} > 11$  GeV. Most of the events are found close to this curve. In these events the jets have lower



**Figure C.1:** Correlations of the scattering angle  $\theta^*$  and a) the invariant mass  $M_{jj}$ , b) the mean pseudo-rapidity  $\bar{\eta}$ . The horizontal and vertical dashed lines show the cuts, while the solid curve in a) corresponds to equation C.2 for jet transverse momenta  $p_t^{jet} > 11$  GeV.



**Figure C.2:** Distributions of the scattering angle  $\theta^*$  for the flavour inclusive dijet sample. The data are compared to the PYTHIA (full line) and CASCADE (dotted line) simulation for a)  $\cos\theta^*$  b)  $\cos\theta^*$  with additional cuts on  $M_{jj}$  and  $\bar{\eta}$  and for c) direct ( $x_\gamma^{obs} > 0.85$ ) and d) resolved ( $x_\gamma^{obs} < 0.85$ ) enriched samples. The Monte Carlo distributions are normalised to the number of events in the data.

transverse momenta and show a tendency to produce high scattering angles  $\cos\theta^*$ . For values of  $M_{jj} > 27$  GeV the bias due to the  $p_t^{jet}$  cut is removed and the  $\cos\theta^*$  distribution can be investigated up to values  $\cos\theta^* < 0.7$ .

Figure C.1 b) shows the  $\bar{\eta} - \cos\theta^*$  plane for the selected dijet events. In order to study dijet events with  $p_t^{jet} > 11$  GeV,  $-0.9 < \eta^{jet} < 1.3$ ,  $M_{jj} > 27$  GeV and scattering angles  $\cos\theta^* < 0.7$  without bias from the  $\eta^{jet}$  cut, the average pseudo-rapidity

$$\bar{\eta} = \frac{1}{2}(\eta^{jet_1} + \eta^{jet_2}) \quad (\text{C.3})$$

is chosen to be in the range  $-0.33 < \bar{\eta} < 0.43$ . This cut on  $\bar{\eta}$  is based on the combination of equations C.1 and C.3 within the boundary conditions of the two jet pseudo-rapidities. In figure C.1 b) the region of the selected events is indicated by the dashed lines.

In summary, the additional cuts on the dijet invariant mass,  $M_{jj} > 27$  GeV, and the average pseudo-rapidity,  $-0.33 < \bar{\eta} < 0.43$ , select an  $M_{jj}$  and  $\bar{\eta}$  region where the biases from the jet kinematic cuts are minimised.

Figure C.2 illustrates the effect of the additional cuts on the  $\cos\theta^*$  distribution. The distributions are shown for the flavour inclusive dijet sample. The shape of the biased  $\cos\theta^*$  distribution (figure C.2 a) is shifted towards that expected from the QCD matrix elements in figure C.2 b). Figures C.2 c) and d) illustrate the shape differences of the  $\cos\theta^*$  distribution for direct ( $x_\gamma^{obs} > 0.85$ ) and resolved ( $x_\gamma^{obs} < 0.85$ ) enhanced event samples.

The distribution rises faster with increasing  $\cos\theta^*$  for resolved photoproduction, where processes involving gluon (spin 1) exchange dominate, than for direct photoproduction, where processes involving quark (spin 1/2) exchange dominate (cf. also figure 2.8).

# References

- [1] A. Aktas *et al.* [H1 Collaboration], “Photoproduction of dijets with high transverse momenta at HERA,” [[hep-ex/0603014](#)].
- [2] A. Aktas *et al.* [H1 Collaboration], “Measurement of dijet production at low  $Q^2$  at HERA,” *Eur. Phys. J. C* **37** (2004) 141 [[hep-ex/0401010](#)].
- [3] S. Chekanov *et al.* [ZEUS Collaboration], “The dependence of dijet production on photon virtuality in e p collisions at HERA,” *Eur. Phys. J. C* **35** (2004) 487 [[hep-ex/0404033](#)].
- [4] S. Chekanov *et al.* [ZEUS Collaboration], “An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment,” *Eur. Phys. J. C* **42** (2005) 1 [[hep-ph/0503274](#)].
- [5] A. Aktas *et al.* [H1 Collaboration], “Measurement of F2(c anti-c) and F2(b anti-b) at high  $Q^2$  using the H1 vertex detector at HERA,” *Eur. Phys. J. C* **40** (2005) 349 [[hep-ex/0411046](#)].
- [6] A. Aktas *et al.* [H1 Collaboration], “Measurement of F2(c anti-c) and F2(b anti-b) at low  $Q^2$  and x using the H1 vertex detector at HERA,” *Eur. Phys. J. C* **45** (2006) 23 [[hep-ex/0507081](#)].
- [7] S. Aid *et al.* [H1 Collaboration], “Photoproduction of  $D^{*\pm}$  Mesons in  $ep$  Collisions at HERA,” *Nucl. Phys. B* **472** (1996) 32 [[hep-ex/9604005](#)].
- [8] J. Breitweg *et al.* [ZEUS Collaboration], “Measurement of inclusive  $D^{*\pm}$  and associated dijet cross sections in photoproduction at HERA,” *Eur. Phys. J. C* **6** (1999) 67 [[hep-ex/9807008](#)].
- [9] S. Chekanov *et al.* [ZEUS Collaboration], “Dijet angular distributions in photoproduction of charm at HERA,” *Phys. Lett. B* **565** (2003) 87 [[hep-ex/0302025](#)].
- [10] S. Chekanov *et al.* [ZEUS Collaboration], “Inclusive jet cross sections and dijet correlations in  $D^{*\pm}$  photoproduction at HERA,” *Nucl. Phys. B* **729** (2005) 492 [[hep-ex/0507089](#)].
- [11] C. Adloff *et al.* [H1 Collaboration], “Measurement of open beauty production at HERA,” *Phys. Lett. B* **467** (1999) 156 [Erratum-*ibid.* B **518** (2001) 331] [[hep-ex/9909029](#)].

- [12] J. Breitweg *et al.* [ZEUS Collaboration], “Measurement of open beauty production in photoproduction at HERA,” *Eur. Phys. J. C* **18** (2001) 625 [[hep-ex/0011081](#)].
- [13] S. Chekanov *et al.* [ZEUS Collaboration], “Bottom photoproduction measured using decays into muons in dijet events in  $e p$  collisions at  $s^{*(1/2)} = 318\text{-GeV}$ ” *Phys. Rev. D* **70** (2004) 012008 [[hep-ex/0312057](#)].
- [14] A. Aktas *et al.* [H1 Collaboration], “Measurement of beauty production at HERA using events with muons and jets,” *Eur. Phys. J. C* **41** (2005) 453 [[hep-ex/0502010](#)].
- [15] A. Aktas *et al.* [H1 Collaboration], “Measurement of charm and beauty dijet cross sections in photoproduction at HERA using the H1 vertex detector,” to appear in *Eur. Phys. J. C.*, [[hep-ex/0605016](#)].
- [16] A. Aktas *et al.* [H1 Collaboration], “Photoproduction of  $D^{*\pm}$  Mesons Associated with a Jet at HERA”, in preparation for publication.
- [17] C. Adloff *et al.* [H1 Collaboration], “Measurement and QCD analysis of neutral and charged current cross sections at HERA,” *Eur. Phys. J. C* **30** (2003) 1 [[hep-ex/0304003](#)].
- [18] V.N. Gribov and L.N. Lipatov, *Yad. Fiz.* **15** (1972) 781 [*Sov. J. Nucl. Phys.* **15** (1972) 438]; G. Altarelli and G. Parisi, *Nucl. Phys. B* **126** (1977) 298; Y.L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641 [*Zh. Eksp. Teor. Fiz.* **73** (1977) 1216].
- [19] M. Ciafaloni, *Nucl. Phys. B* **296** (1988) 49; S. Catani, F. Fiorani and G. Marchesini, *Phys. Lett. B* **234** (1990) 339; *idem*, *Nucl. Phys. B* **336** (1990) 18; G. Marchesini, *Nucl. Phys. B* **445** (1995) 49.
- [20] A. Aktas *et al.* [H1 Collaboration], “Inclusive dijet production at low Bjorken- $x$  in deep inelastic scattering,” *Eur. Phys. J. C* **33** (2004) 477 [[hep-ex/0310019](#)].
- [21] H. L. Lai *et al.*, *Eur. Phys. J. C* **12** (2000) 375 [[hep-ph/9903282](#)].
- [22] A.D. Martin, R.G. Roberts and W.J. Stirling, *Phys. Lett. B* **354** 155 (1995); A.D. Martin, R.G. Roberts, W.J. Stirling and R.S. Thorne, *Eur. Phys. J. C* **4** (1998) 463.
- [23] C.F. von Weizsäcker, “Ausstrahlung bei Stößen sehr schneller Elektronen,” *Z. Phys.* **88** (1934) 612.
- [24] E.J. Williams, “Nature of the High Energy Particles of Penetrating Radiation and Status of Ionisation and Radiation Formulae,” *Phys. Rev* **45** (1934) 729.
- [25] E. Witten, “Anomalous Cross-Section For Photon - Photon Scattering In Gauge Theories,” *Nucl. Phys. B* **120** (1977) 189.
- [26] R. Nisius, “The photon structure from deep inelastic electron photon scattering,” *Phys. Rept.* **332** (2000) 165 [[hep-ex/9912049](#)].
- [27] C. Adloff *et al.* [H1 Collaboration], “Measurement of di-jet cross-sections in photoproduction and photon structure,” *Phys. Lett. B* **483** (2000) 36 [[hep-ex/0003011](#)].

- [28] M. Gluck, E. Reya and A. Vogt, "Parton structure of the photon beyond the leading order," *Phys. Rev. D* **45** (1992) 3986; M. Gluck, E. Reya and A. Vogt, "Photonic parton distributions," *Phys. Rev. D* **46** (1992) 1973.
- [29] G. A. Schuler and T. Sjostrand, "Low and high mass components of the photon distribution functions," *Z. Phys. C* **68** (1995) 607 [[hep-ph/9503384](#)].
- [30] J. Smith and W. K. Tung, "Heavy flavor production," NIKHEF-H-93-20.
- [31] J. Smith and W. L. van Neerven, "QCD corrections to heavy flavor photoproduction and electroproduction," *Nucl. Phys. B* **374** (1992) 36.
- [32] C. Adloff *et al.* [H1 Collaboration], "Inclusive  $D^0$  and  $D^{*+}$  production in neutral current deep inelastic  $e p$  scattering at HERA," *Z. Phys. C* **72** (1996) 593 [[hep-ex/9607012](#)].
- [33] A. Aktas *et al.* [H1 Collaboration], "The Charm Fragmentation Function in DIS," in preparation for publication.
- [34] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, "Heavy-quark production," *Adv. Ser. Direct. High Energy Phys.* **15** (1998) 609 [[hep-ph/9702287](#)].
- [35] C. Peterson, D. Schlatter, I. Schmitt and P.M. Zerwas, *Phys. Rev. D* **27** (1983) 105.
- [36] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, "Parton Fragmentation And String Dynamics," *Phys. Rept.* **97** (1983) 31.
- [37] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP* **0206** (2002) 029 [[hep-ph/0204244](#)];  
S. Frixione, P. Nason and B. R. Webber, "Matching NLO QCD and parton showers in heavy flavour production," *JHEP* **0308** (2003) 007 [[hep-ph/0305252](#)];  
M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, "QCD analysis of first b cross section data at 1.96-TeV," *JHEP* **0407** (2004) 033 [[hep-ph/0312132](#)].
- [38] M. Cacciari, M. Greco and P. Nason, "The  $p(T)$  spectrum in heavy-flavour hadroproduction," *JHEP* **9805** (1998) 007 [[hep-ph/9803400](#)].
- [39] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, "Total cross sections for heavy flavor production at HERA," *Phys. Lett. B* **348** (1995) 633 [[hep-ph/9412348](#)].
- [40] M. Martisikova, "Jet Shapes in Charm Photoproduction at HERA," PhD thesis, Univ. Hamburg, 2005, DESY-THESIS-2005-047.
- [41] M. Glück, E. Reya and A. Vogt, *Phys. Rev. D* **45** (1992) 3986; *idem*, *Phys. Rev. D* **46** (1992) 1973 .
- [42] G. Flucke, "Photoproduction of  $D^*$  Mesons and  $D^*$  Mesons Associated with Jets at HERA," Dissertation, Univ. Hamburg, 2005, DESY-THESIS-2005/06.
- [43] S. Kretzer, H. L. Lai, F. I. Olness and W. K. Tung, "CTEQ6 parton distributions with heavy quark mass effects," *Phys. Rev. D* **69** (2004) 114005 [[hep-ph/0307022](#)].

- 
- [44] T. Sjöstrand, “High-Energy Physics Event Generation with PYTHIA 5.7 and JETSET 7.4,” *Comput. Phys. Commun.* **82** (1994) 74.
- [45] H. Jung and G. P. Salam, “Hadronic final state predictions from CCFM: The hadron-level Monte Carlo generator CASCADE,” *Eur. Phys. J. C* **19** (2001) 351 [[hep-ph/0012143](#)];  
H. Jung, “The CCFM Monte Carlo generator CASCADE,” *Comput. Phys. Commun.* **143** (2002) 100 [[hep-ph/0109102](#)].
- [46] G.A. Schuler and T. Sjöstrand, “Parton Distributions of the Virtual Photon,” *Phys. Lett. B* **376** (1996) 193 [[hep-ph/9601282](#)].
- [47] H. Jung, “Un-integrated uPDFs in CCFM,” [[hep-ph/0411287](#)].
- [48] K. Charchula, G. A. Schuler and H. Spiesberger, *Comput. Phys. Commun.* **81** (1994) 381.
- [49] S. Eidelman et al. [Particle Data Group Collaboration], “Review of particle physics,” *Phys. Lett. B* **592** (2004) 1.
- [50] I. Abt *et al.* [H1 Collaboration], “The Tracking, calorimeter and muon detectors of the H1 experiment at HERA ,” *Nucl. Instrum. Meth. A* **386** (1997) 310 and 348.
- [51] W. Erdmann, “Untersuchung der Photoproduktion von  $D^*$ -Mesonen am ep Speicher-ring HERA,” PhD thesis, ETH Zürich, 1996, ETH No. 11441.
- [52] D. Pitzl *et al.*, “The H1 silicon vertex detector,” *Nucl. Instrum. Meth. A* **454**, 334 (2000) [[hep-ex/0002044](#)].
- [53] M. Kausch “The Silicon Microvertex Detector of the H1 Experiment: Readout, Event Reconstruction, and Studies on Heavy Quark Decays,” PhD thesis, Univ. Hamburg, 1998, DESY-THESIS-1998-0333.
- [54] B. Andrieu *et al.* [H1 Calorimeter Group], “The H1 liquid argon calorimeter system,” *Nucl. Instrum. Meth. A* **336** (1993) 460.
- [55] T. Nicholls *et al.* [H1 SPACAL Group Collaboration], “Performance of an electromagnetic lead / scintillating fiber calorimeter for the H1 detector,” *Nucl. Instrum. Meth. A* **374** (1996) 149.
- [56] J.C. Bizot *et al.*, “Pure LAr calorimeter L2TT trigger elements,” H1 internal note [[H1-07/99-574](#)].
- [57] T. Carli *et al.*, “Performance of the H1 LAr Trigger in 1994,” H1 internal note [[H1-07/95-445](#)].
- [58] R. Brun *et al.*, “Geant 3,” (1987) CERN-DD/EE/84-1.
- [59] J. E. Augustin *et al.*, “Discovery of a Narrow Resonance in  $e^+e^-$  Annihilation,” *Phys. Rev. Lett.* **33** (1974) 1406.
- [60] S. W. Herb *et al.*, *Phys. Rev. Lett.* **39** (1977) 252.

- [61] A.B. Meyer, "Heavy Quark Production at HERA," Habilitation, Univ. Hamburg (2005).
- [62] O. Behnke, "Production of Charm and Beauty Quarks at HERA," Habilitationsschrift, Ruprecht-Karls Universität Heidelberg (2005).
- [63] S. Chekanov *et al.* [ZEUS Collaboration], "Inclusive Jet Cross Sections and Dijet Correlations in D\* Photoproduction at HERA," Nuclear Physics B **729** (2005) 492-525. [[hep-ex/0507089](#)].
- [64] J. C. Anjos *et al.* [Tagged Photon Spectrometer Collaboration], "Photon - Gluon Fusion Analysis Of Charm Photoproduction," Phys. Rev. Lett. **65** (1990) 2503.
- [65] M. Cacciari and P. Nason, "Charm cross sections for the Tevatron Run II," JHEP **0309** (2003) 006 [[hep-ph/0306212](#)].
- [66] F. Happacher, P. Giromini and F. Ptohos, "Status of the observed and predicted b anti-b production at the Tevatron," Phys. Rev. D **73** (2006) 014026 [[hep-ph/0509348](#)].
- [67] Tara Shears, "Charm and Beauty production at the Tevatron," CDF note CDF7905.
- [68] E. L. Berger, B. W. Harris, D. E. Kaplan, Z. Sullivan, T. M. P. Tait and C. E. M. Wagner, "Low energy supersymmetry and the Tevatron bottom-quark cross section," Phys. Rev. Lett. **86** (2001) 4231 [[hep-ph/0012001](#)].
- [69] ATLAS Collaboration, Technical Report CERN/LHCC 99-14, 1999.
- [70] A. Aktas *et al.* [H1 Collaboration], "Measurement of charm and beauty photoproduction at HERA using D\* mu correlations," Phys. Lett. B **621** (2005) 56 [[hep-ex/0503038](#)].
- [71] J. Gassner, "A Measurement of D-Meson Production at HERA by Decay Vertex Identification," PhD thesis, Swiss Federal Inst. of Technology Zürich, 2002, ETHZ No. 14774.
- [72] O. Behnke, J. Kroseberg, "CSTLIN: Combined CJC-CST Track Fit," H1 internal note, in preparation.
- [73] T. Kuhr and M. Kausch-Blecken von Schmeling, "CS PRIM: Primary Vertex Determination using CST Measurements," H1 internal document (2000).
- [74] O. Behnke, H1 internal document.
- [75] C. Gerlich, "Messung der Produktion schwerer Quarks in inklusiven Zweijet-Ereignissen mit einem Myon unter Anwendung einer Multi-Impaktparameter Methode am H1 Experiment bei HERA," Dissertation, Univ. Heidelberg, 2005.
- [76] M. Peez, B. Portheault, E. Sauvan, "An energy flow algorithm for Hadronic Reconstruction in OO: Hadroo2," H1 internal note [[H1-01/05-616](#)].
- [77] M.H. Seymour, "Jets in QCD", CERN-TH/95-176, talk given at the 10th Topical Workshop on Proton-Antiproton Collider Physics, Batavia, IL, May 9-13, 1995.

- [78] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, “Longitudinally invariant  $K(t)$  clustering algorithms for hadron hadron collisions,” Nucl. Phys. B **406** (1993) 187.
- [79] S. Frixione and G. Ridolfi, “Jet photoproduction at HERA,” Nucl. Phys. B **507** (1997) 315 [[hep-ph/9707345](#)].
- [80] S. Caron, “Jets in Photoproduction at HERA,” PhD thesis, RWTH Aachen, 2002, DESY-THESIS-2002-035.
- [81] F. Jacquet and A. Blondel, “Report from the Study Group on Detectors for Charged Current Events,” in Proc. of the Study for an ep Facility for Europe, edited by U. Arnaldi, (1979) 377, DESY-79-048.
- [82] E. Chabert et al., “Qbgfmar: An updated phan package for cosmic and halo muon topological rejection in high  $p_t$  physics analysis,” H1 internal note [[H1-01/98-556](#)].
- [83] Ch. Veelken, “H1nonepbgfinder - rejection of cosmic muon and beam-halo events in the h1oo framework,” H1 internal note [[H1-01/02-603](#)].
- [84] I. Strauch, “Jets with High Transverse Momenta in Photoproduction at HERA,” PhD thesis, Univ. Hamburg, 2004, DESY-THESIS-2004/047.
- [85] A. Mehta, “New track based electron finder,” H1 internal document.
- [86] T. Affolder *et al.* [CDF Collaboration], “Charged jet evolution and the underlying event in proton anti-proton collisions at 1.8 TeV,” Phys. Rev. D **65** (2002) 092002.
- [87] J. Breitweg *et al.* [ZEUS Collaboration], “Measurement of jet shapes in photoproduction at HERA,” Eur. Phys. J. C **2** (1998) 61 [[hep-ex/9710002](#)].
- [88] C. Adloff *et al.* [H1 Collaboration], “Measurement of inclusive jet cross sections in photoproduction at HERA,” Eur. Phys. J. C **29** (2003) 497 [[hep-ex/0302034](#)].
- [89] D. Buskulic *et al.* [ALEPH Collaboration], “A Precise measurement of Gamma ( $Z \rightarrow b$  anti- $b$ ) / Gamma ( $Z \rightarrow$  hadrons),” Phys. Lett. B **313** (1993) 535.
- [90] L. Finke, “Untersuchungen zum Erkennen von B-Mesonen im H1-Detektor,” Diploma thesis, RWTH Aachen, 2003.
- [91] S. Mangano, “Charm Production in Charged Current and Charm & Beauty Production in Neutral Current Interactions at High  $Q^2$ ,” PhD thesis, ETH Zuerich, 2005.
- [92] R. Frühwirth *et al.*, “New developments in vertex reconstruction for CMS,” Nucl. Instrum. Meth. A **502** (2003) 446.
- [93] R. Frühwirth and A. Strandlie, “Adaptive multi-track fitting,” Comput. Phys. Commun. **120** (1999) 197.
- [94] H1 Tracking Group, <http://www-h1.desy.de/idet/itracker/TrackingGroup/home.html>, H1 internal document.

- [95] D. Coffman *et al.* [MARK-III Collaboration], “Measurement of the inclusive decay properties of charmed mesons,” *Phys. Lett. B* **263** (1991) 135.
- [96] D. Abbaneo *et al.* [LEP/SLD Heavy Flavour Working Group], LEPHF 2001-01, available from <http://lepewwg.web.cern.ch/LEPEWWG/heavy>.

# Acknowledgment

It is a pleasure for me to express my gratitude to everyone who supported me during the course of my stay at H1.

Special thanks go to my supervisor Prof. Dr. Beate Naroska for her excellent support, advice and confidence in my work throughout my research. This was truly appreciated. I am thankful to Prof. Dr. Rolf-Dieter Heuer for kindly accepting to be my second referee. I wish to thank Andreas Meyer for all the help and countless discussions we had about details of the analysis. It would not have been possible for me to reach so far were it not for his experience, advice and motivation.

Many thanks go to the whole H1 collaboration, especially the members of the Heavy Flavour Working Group, for the friendly and professional atmosphere I enjoyed. Here my special thanks go to Olaf Behnke for his many fruitful ideas and the stimulating discussions we had.

Special mention to those who participated in the proofreading of the manuscript.

Further thanks go to the H1 Hamburg group, who are too numerous to mention, for their good team work and the enjoyable atmosphere we shared. I would like to thank especially Michael Steder and Shiraz Habib for their company, bad coffee and many meaningless discussions. I thank Maria Martisikova, with whom I shared my office space, for the encouraging times and good working environment.

My stay at DESY would not have been so pleasant were it not for my colleagues and best friends Johannes Haller, Linus Lindfeld and Carsten Schmitz.

Lovely thanks go to charming Mona for being the beauty in my life.

Finally I would like to thank my family for their invaluable support during the past 29 years. This work would not be possible without all of them.