The Internal Structure of Charmed Jets in Photoproduction at HERA

and

Tests of the ZEUS Microvertex Silicon Sensors

by

Margherita Milite

The Internal Structure of Charmed Jets in Photoproduction at HERA

and

Tests of the ZEUS Microvertex Silicon Sensors

DISSERTATION

zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von

Margherita Milite
aus Salerno, Italy

Hamburg
Mai 2001

Gutachter der Dissertation:	Prof. Dr. R. Klanner
	Prof. Dr. B. Naroska
	Prof. Dr. N. Wermes
Gutachter der Disputation:	Prof. Dr. Fr.-W. Büßer
	Prof. Dr. G. Heinzelmann
Datum der Disputation:	04.09.2001
Dekan des Fachbereichs Physik und Vorsitzender des Promotionsausschusses:	Prof. Dr. Fr.-W. Büßer

Abstract

In the first part of this thesis, the internal structure of charm initiated jets in high energy γp interactions has been studied at HERA for the first time. The mean subjet multiplicity and the differential and integrated jet shape, describing the distribution of the transverse energy inside the jet, have been measured for jets with $E_T^{\rm jet} \geq 6$ GeV and $-1 \leq \eta^{\rm jet} \leq 2$. The data collected with the ZEUS detector during the 1996-2000 e^+p running periods, corresponding to an integrated luminosity of ~ 100 pb^{-1}, has been used. A charm-enriched jet sample has been obtained by selecting dijet events where at least one of the two jets contains a $D^*(2010)^\pm$ meson with $p_T(D^*) \geq 3.0$ GeV and $|\eta(D^*)| \leq 1.5$. The daughter particles of the decay mode $D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+$ (and the charge-conjugate channel) have been measured with the ZEUS central tracking detector (CTD). The closest jet in azimuth to the D^* meson is tagged and the internal structure of the second jet, which in most of the events also originates from a charm (or anti-charm) quark, is investigated. The $E_T^{\rm jet}$ and $\eta^{\rm jet}$ dependence of the measured jet substructure has been studied and compared with the predictions from leading-logarithm parton-shower Monte Carlo programs. In addition, a measurement of the internal jet structure for inclusive (i.e. without flavour tagging) dijets in photoproduction, at much lower $E_T^{\rm jet}$ compared with previous analyses, has been performed. In particular, it has been observed that the jets in the inclusive sample become broader as $\eta^{\rm jet}$ increases whereas the charm initiated jets do not. This is consistent with the predicted fractions of quark and gluon initiated jets in photoproduction processes. The internal structure of gluon jets has been extracted with a statistical method using the measurements for the inclusive dijet and the charm-enriched jet samples for $E_T^{\rm jet} \geq 15$ GeV. It is observed that gluon jets have a higher mean subjet multiplicity and are broader than quark jets of the same $E_T^{\rm jet}$ which is in agreement with pQCD predictions. The second part of this thesis describes the new silicon microvertex (MVD), which has been installed in the ZEUS detector during the HERA luminosity upgrade, and the main results of the beam test program performed at DESY to determine the characteristic performances of the MVD silicon sensors.

Zusammenfassung

Im ersten Teil dieser Arbeit wurde die innere Struktur von Charm-Quark induzierten Jets in hochenergetischen γp Wechselwirkungen am HERA-Speicherring untersucht. Es wurden die mittlere Subjet-Multiplizität und die differentiellen und integrierten Jet-Formen (engl.: *jet shapes*), die die Verteilung der transversalen Energie innerhalb des Jets beschreiben, im kinematischen Bereich von $E_T^{\rm jet} \geq 6$ GeV und $-1 \leq \eta^{\rm jet} \leq 2$ gemessen. Dafür wurden die mit dem ZEUS Detektor während der e^+p Runperiode 1996-2000 aufgezeichneten Daten verwendet, die einer integrierten Luminosität von ~ 100 pb^{-1} entsprechen. Eine Ereignismenge von Charm-Quark induzierten Jets konnte durch Auswahl von 2-Jet-Ereignissen, in denen mindestens einer der Jets ein $D^*(2010)^\pm$ Meson (mit $p_T(D^*) \geq 3.0$ GeV und $|\eta(D^*)| \leq 1.5$) enthält, erzeugt werden. Die Zerfallsprodukte des Zerfallskanals $D^{*+} \to D^0(\to K^-\pi^+)\pi_s^+$ (und dem dazu ladungskonjugierten Kanal) wurden mit der zentralen Spurkammer (CTD) des ZEUS Detektors gemessen. Der Jet, der den kleinsten azimuth Winkel zum D^* Meson hat, wurde identifiziert und die innere Struktur des zweiten Jets, der in den meisten Fällen auch durch ein Charm (oder Anti-Charm) Quark entstanden ist, wurde analysiert. Die $E_T^{\rm jet}$ und $\eta^{\rm jet}$ Abhängigkeit der gemessenen Unterstruktur dieser Jets wurde untersucht und mit Vorhersagen von Partonschauer Monte Carlo Modellen verglichen. Zusätzlich wurde eine Messung der inneren Jetstruktur für inklusive (d.h. ohne Flavour-Identifizierung) 2-Jet-Ereignisse in Photoproduktion, bei viel kleinerem $E_T^{\rm jet}$ als in vorhergehenden Analysen, durchgeführt. Die gemessene innere Struktur von charmangereicherten Jets wurde mit der von inklusiven Jets verglichen. Insbesondere wurde festgestellt das in der inclusive Messung Jets in vorwärts Richtung breiter werden, die Quark induzierten Jets hingegen nicht. Dies ist konsistent mit dem erwarteten Quark- und Gluonanteil aus Photoproduktionsprozessen. Mit Hilfe einer statistischen Methode und durch Kombination der Messungen von inklusiven und charmangereicherten 2-Jet-Ereignissen wurde die innere Struktur von Gluonjets für $E_T^{\rm jet} \geq 15$ GeV extrahiert. Dabei wurde festgestellt, daß Gluonjets eine höhere mittlere Subjet-Multiplizität haben und breiter sind als Quarkjets bei vergleichbarem $E_T^{\rm jet}$, was in guter Übereinstimmung mit pQCD Vorhersagen ist. Im zweiten Teil dieser Arbeit wurde der neue Silizium Mikrovertex Detektor (MVD) vorgestellt, der während des HERA Luminositäts-Upgrades in das ZEUS Experiment eingebaut worden ist. Die wichtigsten Ergebnisse des am DESY durchgeführten Teststrahlprogramms zur Untersuchung der charakteristischen Eigenschaften der MVD Siliziumsensoren wurden präsentiert.

Contents

List of Figures

List of Tables

Introduction

The dynamics of strong interaction phenomena is currently best understood in the framework of Quantum ChromoDynamics (QCD), a non-Abelian gauge theory which describes the interaction among the constituents of baryons and mesons (quarks) via the exchange of massless gauge bosons (gluons), which themselves carry colour and interact with each other. Because of colour confinement, quarks and gluons are not directly observable as free particles but they give rise to narrow sprays of colourless hadrons called 'jets'. In our present understanding, these jets result from bremsstrahlung processes, i.e. chains of elementary splitting processes off the initial quarks or gluons. The relative strengths of the different splittings, i.e. $g \to gg$, $q \to q'g$ and $g \to q\bar{q}$, are determined by the QCD colour factors which directly follow from the $SU(3)$ group structure of QCD: $C_A = 3$ and $C_F = 4/3$ and $T_R = 1/2$. As a consequence of the bremsstrahlung process, hadronic jets develop an internal structure which is expected to depend mainly on the type of primary parton from which they originated. Asymptotically (at large energy scale), in leading order calculations at parton level, $C_A/C_F = 9/4$ times more splittings are predicted in gluon initiated jets than in quark initiated jets [1]. Although non-perturbative effects are expected to have an influence on the physical observables used to define the internal properties of hadronic jets (in particular at low energy), the basic characteristics of the perturbative evolution should still be visible after hadronization. The investigation of the nature of quark and gluon initiated jets thus constitutes an intuitive test of perturbative QCD predictions and might allow to gain insight into the dynamical processes responsible for the transition from partons to hadrons.

The topic of this thesis is the measurement of the internal jet structure of charm (i.e. quark) initiated jets in high energy γp interactions at HERA. A sample of D^*-tagged dijet events is selected, one jet is associated with the charmed meson and the internal structure of the second jet in the event is investigated. In most of the D^*-tagged dijet events, the second jet also originates from a charm (or anti-charm) quark and therefore by looking at the untagged jet (referred to as 'other' jet), an unbiased measurements, i.e. independent of the selection criteria applied to find the charmed meson, for a charm-enriched sample of jets can be performed. The data taken by the ZEUS experiment during the pre-upgrade e^+p running periods of the 1996-2000 years, corresponding to an integrated luminosity of ~ 100 pb^{-1}, is used for the analysis.

Sofar the identification of the charmed mesons with the ZEUS experiment (as well as the estimation of the background contribution to the signal) is based on detection methods using tracks reconstructed by the central tracking component. The purity and efficiency of the various charm tagging methods (e.g. D^* reconstruction, semi-leptonic charm decays, etc.) would of course significantly benefit from a direct tagging of secondary vertices.

During the shutdown in 2000-2001, HERA has undergone a significant luminosity upgrade which aims for an increase of the e$^\pm$p peak luminosity from 1.5 to 7.0$\times$ 10^{31} cm^{-2}s^{-1}, providing new sensitivity to low cross-section physics and in particular a higher event rate for heavy quark production. To take full advantage of the post-upgrade increased statistics, events containing heavy quark hadrons have to be directly tagged. A silicon microvertex detector (MVD) has been installed in the ZEUS experiment, aiming to significantly improve the overall tracking capabilities and to provide the ability to identify events containing secondary vertices. This will represent a great step forward for many physics topics made accessible by the luminosity upgrade. For example, charm initiated jets in photoproduction will give new insight into the parton content of the photon, a very precise measurement of the structure function $F_2^{c\bar{c}}$ will be performed, extending the accessible kinematic range beyond the present limits and a precise measurement of the beauty production cross section will be possible.

This thesis is divided in two parts: the first part is dedicated to the physics analysis performed using the data collected by the ZEUS experiment in the pre-upgrade phase. The second part describes the silicon microvertex detector and the main results of the beam test program performed at DESY to determine the characteristic performances of the MVD silicon sensors.

The first part of the thesis is organized as follows. In chapter 1 an overview of the theoretical background relevant for the investigations made in the analysis is presented. The theory of strong interactions (QCD) and the perturbative regime in which calculations can be performed are introduced. The different properties of quark and gluon initiated jets are discussed and the jet observables under study are defined. In chapter 2 the HERA collider and the ZEUS detector are described, putting emphasis on those detector components important for the analysis. Chapter 3 gives a brief overview of the Monte Carlo generators used, focusing on the different aspects of the physics simulation process. The reconstruction and selection of the dijet photoproduction data sample is explained in chapter 4. The procedure used for the charm identification with the ZEUS detector, which is based on the reconstruction of D^* mesons in the central tracking detector, the method used to estimate the background contribution to the charm signal and the tagging of the charm initiated jets are discussed in chapter 5. The measurement of the internal jet structure at detector level for the dijet photoproduction sample and the charm-enriched jet sample is presented in chapter 6. The final results corrected for detector effects are presented in chapter 7 and compared with the available theoretical predictions. A summary and discussion of the results conclude the first part of this thesis.

Chapters 8 and 9 are devoted to the ZEUS MVD project. In chapter 8 the layout of the ZEUS MVD is presented. The main measurements performed with the MVD prototype sensors at the DESY testbeam are explained in detail in chapter 9. The conclusions on the MVD prototype performances put the final endpoint to this thesis.

Part I

Analysis of 1996-2000 Data

Chapter 1

Theoretical Overview

This chapter provides a brief overview of the physics background relevant for the investigations made in this thesis and introduces the physical observables measured in the analysis. In the first two sections (1.1, 1.2) the theory of strong interactions, the perturbative regime in which calculations can be performed and the different properties of quark and gluon jets are introduced. The following sections (1.3, 1.4) illustrate the criteria to define a jet of hadrons and introduce the observables used to measure the internal jet structure. The main characteristics of lepton-hadron scattering and particularly of the kinematic regime (photoproduction) in which the measurements of this analysis are performed, are summarized in sections 1.5.1 and 1.5.2. Finally, the goal and strategy of the analysis are presented in section 1.5.3.

1.1 Quantum Chromodynamics (QCD) and the Perturbative Regime

Quantum Chromodynamics (QCD) is the theory of the strong interaction, one of the four known fundamental forces in nature. It describes the interaction among the constituents of baryons and mesons and represents one of the cornerstones of the 'Standard Model' of elementary particles.

The fundamental axiom of QCD is that matter is made of pointlike half-integer spin constituents, called 'quarks' which carry a fractional electromagnetic charge (+2/3 or -1/3) and one of the three possible colour charges (red, green, blue). The colour degree of freedom was introduced to solve the apparent violation of the Pauli's exclusion principle in the Δ^{++} baryons. The number of colours, N_C, must be at least three in order to construct a totally asymmetric wave function for the Δ^{++} baryon, which consists of three u quarks.

QCD is a non-abelian gauge theory based on the SU(3) symmetry group. Colour is exchanged between quarks via gauge bosons, called 'gluons'. As a consequence of the non-abelian structure of QCD, the gluons themselves carry colour and therefore not only interact with quarks and antiquarks but also with each other[1]. This is shown in Fig. 1.1 where the fundamental interaction vertices of the QCD Lagrangian are illustrated. The symbol α_S, associated with the three different interaction vertices, represents the strong coupling constant (defined in analogy with

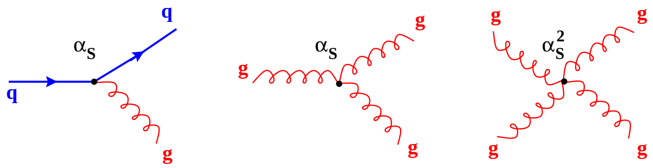

Figure 1.1: Diagrams of the three fundamental interaction vertices in the QCD Lagrangian: (a) quark-gluon vertex (order α_S) , (b) triple-gluon vertex (order α_S), (c) four-gluon vertex (order α_S^2). The symbol α_S represents the strong coupling constant.

the fine structure constant of QED). Despite its name, the coupling constant of the strong force shows an energy scale dependence. The 'running' of the coupling constant is a characteristic feature of renormalizable gauge theories: when a dimensionless physical observable $R(Q)$ (where Q is the energy scale on which the observable depends) is calculated as a perturbation series in the coupling α_S:

$$R(Q) = R_0(Q) + \alpha_S \cdot R_1(Q) + \alpha_S^2 \cdot R_2(Q) + ...$$

ultraviolet (UV) divergences arise, i.e. the observable $R(Q)$ becomes infinity. In order to remove the UV divergences, a **renormalization** procedure is employed. This procedure introduces a second energy scale, μ_R, which represents the point at which the subtraction to remove the UV divergences is performed. As a consequence the observable R will depend in general on the ratio Q^2/μ_R^2. However the scale μ_R is an arbitrary parameter and therefore any physical observable cannot depend on the choice made for it. Mathematically this is expressed as:

$$\mu_R^2 \frac{d}{d\mu_R^2} R(Q^2/\mu_R^2, \alpha_S) = 0 \tag{1.1}$$

and the running of the renormalized coupling constant $\alpha_S^R = \alpha_S^R(Q^2/\mu_R^2)$ is determined by the renormalization group equation[3]:

$$\mu_R^2 \frac{d}{d\mu_R^2}\alpha_S = \beta(\alpha_S) = -\alpha_S \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_S}{4\pi}\right)^{n+1} \tag{1.2}$$

with[2]

$$\beta_0 = 4\pi b = 11 - \frac{2}{3}N_f, \quad \beta_1 = 16\pi^2 bb' = 102 - \frac{38}{3}N_f, \quad ...$$

where N_f is the number of active quark flavours with a mass smaller than the energy scale μ_R. The lowest order solution for the Eq. 1.2 is given by:

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu_R^2)}{1 + \alpha_S(\mu_R^2)(\beta_0/4\pi)\ln(Q^2/\mu_R^2)} \xrightarrow{Q^2\to\infty} 0 \tag{1.3}$$

which gives the relation between $\alpha_S(Q^2)$ and $\alpha_S(\mu_R^2)$. From Eq. 1.3 one can see that for large values of the energy scale Q (or equivalently for short distances $\lambda \sim 1/Q$) the running coupling constant $\alpha_S(Q^2)$ vanishes logarithmically. This property is known as '**asymptotic freedom**' and allows the application of perturbative techniques to calculate short distance interactions.

[1] The self-coupling of the gauge bosons in QCD represents the main difference to the quantum electrodynamics (QED), which is the theory of the electromagnetic interactions.

[2] The values of the coefficients $\beta_i = 0, 1, 2, ..$ depend on the renormalization scheme used. Here the values obtained in the $\overline{MS}$ scheme are reported.

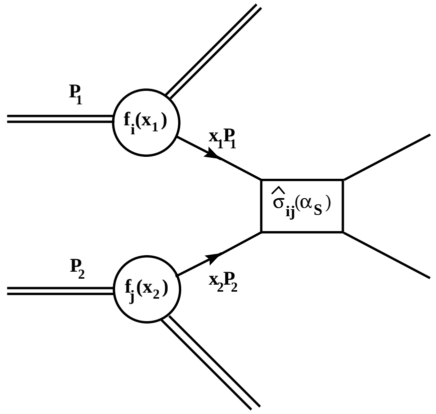

Figure 1.2: Schematic diagram of a hard scattering process.

For small values of the energy scale Q, i.e. $Q^2 \lesssim \mathcal{O}(\Lambda_{\rm QCD}^2)$ (or equivalently for long distances λ) the running coupling constant $\alpha_S(Q^2)$ increases and perturbation theory is no longer applicable. The constant $\Lambda_{\rm QCD}$, which has the dimension of an energy, is by convention used in the definition of $\alpha_S(Q^2)$, and is defined as the constant of integration over the β function:

$$\ln\frac{Q^2}{\Lambda_{\rm QCD}^2} = -\int_{\alpha_S(Q^2)}^{\infty}\frac{dx}{\beta(x)} = \int_{\alpha_S(Q^2)}^{\infty}\frac{dx}{bx^2(1+b'+...)}$$

$\Lambda_{\rm QCD}$ is regarded as the scale at which the coupling $\alpha_S(Q^2)$ diverges and therefore represents the limit of the perturbative domain. The value of $\Lambda_{\rm QCD}$ (of the order of $\sim$ 200 MeV) is not predicted by the theory and has to be determined from fits to experimental data.

The fact that the strong force becomes larger and larger as the energy scale decreases is an argument used to introduce another fundamental property of QCD known as '**confinement**': quarks and gluons[3] cannot be observed as free particles but only as bound states in colourless hadrons. The non-perturbative process in which the coloured partons are transformed into colourless hadrons is called **hadronization** and will be discussed in more detail in chapter 3. This chapter is instead focused on the large energy scale regime in which perturbative QCD (**pQCD**) is able to provide precise predictions.

In pQCD a scattering process is viewed as the result of a hard interaction between quarks and gluons (in hadron-hadron scattering) or of the leptons with the quarks (lepton-hadron scattering). The incoming hadrons can be seen as providing 'broad-band' beams of partons which carry a fraction of the momenta of their parent hadrons. The cross section for a hard scattering initiated by two hadrons with 4-momenta $\rm P_1$ and $\rm P_2$, illustrated in Fig. 1.2, can be written as:

$$\sigma(\mathrm{P}_1, \mathrm{P}_2) = \sum_{\mathrm{i,j}} \int \mathrm{dx}_1\,\mathrm{dx}_2\,\mathrm{f_i}(\mathrm{x}_1, \mu_\mathrm{F}^2)\,\mathrm{f_j}(\mathrm{x}_2, \mu_\mathrm{F}^2)\hat{\sigma}_{\mathrm{ij}}(\mathrm{x}_1\mathrm{P}_1, \mathrm{x}_2\mathrm{P}_2, \alpha_\mathrm{S}(\mu_\mathrm{R}^2), \mu_\mathrm{F}^2, \mu_\mathrm{R}^2) \tag{1.4}$$

The momenta of the partons which participate in the hard scattering are $\rm p_1 = x_1P_1$ and $\rm p_2 = x_2P_2$ and the characteristic scale of the hard scattering is denoted by $\rm Q^2$. The short-distance

[3] The general term 'parton' will be used in the following when referring either to a quark or to a gluon.

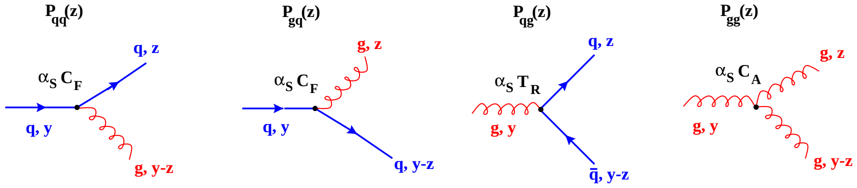

Figure 1.3: Diagrams of the fundamental QCD couplings. The two diagrams on the left show the gluon radiation from a quark which is proportional to $\alpha_S \cdot C_F$, whereas the two diagrams on the right show the gluon splittings which are proportional to $\alpha_S \cdot T_R$ and $\alpha_S \cdot C_A$, respectively.

cross section, $\hat{\sigma}_{ij}(p_1, p_2, \alpha_S(\mu_R^2), \mu_F^2, \mu_R^2)$, for the hard scattering of partons of type i and j involves only high momentum transfer and is independent of the type of the incoming hadrons.

The parton distribution functions $f_i(x, \mu_F^2)$ (**PDFs**) have to be determined experimentally. They describe the distribution of longitudinal momentum fractions x of parton type i at an energy scale μ_F^2. The PDFs depend on the type of the hadron, but not on the particular scattering process considered and therefore are said to be 'universal'. Although the initial PDFs cannot be calculated perturbatively, pQCD predicts their evolution with the energy scale Q^2, if their value at a certain scale Q_0^2 is given. The evolution is described by a set of integro-differential equations known as the **DGLAP**[4] evolution equations [4].

The **factorization scale**, μ_F, is an arbitrary parameter and can be considered as the scale which separates the long and the short-distance physics. Typically the renormalization and the factorization scales are set equal, $\mu_F = \mu_R = \mu$, and μ is chosen to be of the same order as Q. Any fixed order perturbative QCD calculation depends on the choice of the factorization and renormalization scales and the sensitivity of the theoretical predictions to μ is often taken as a measure of the uncertainty from the contributions of higher order terms.

The final state partons manifest themselves as collimated streams of hadrons called **jets** which represent one of the most striking phenomena in high-energy physics. In the present understanding, the jets result from a chain of elementary splitting processes (**parton branchings**) of the quarks and gluons emerging from the hard scattering. At each parton branching, the virtuality of the parton Q decreases, thus leading to a corresponding increase in the strong coupling constant $\alpha_S(Q^2)$ and therefore the process can be described only up to a certain cut-off value $Q^2_{\mathrm{cut-off}}$. The latter is chosen in such a way that the best overall description of the experimental data is achieved. Typical values are in the range $Q_{\mathrm{cut-off}} \simeq$ 0.6-1.6 GeV. The probability for a parton i to emit a parton j with momentum fraction z of the parent parton is given by one of the four different splitting functions (see Fig. 1.3), which at first order in α_S have the following expression:

$$P^0_{qq}(z) = \mathbf{C_F} \cdot \left[\frac{1+z^2}{1-z} + \frac{3}{2}\delta(1-z)\right]$$
$$P^0_{gq}(z) = \mathbf{C_F} \cdot \left[\frac{1+(1-z)^2}{z}\right]$$
$$P^0_{qg}(z) = \mathbf{T_R} \cdot \left[z^2 + (1-z)^2\right]$$

[4]Dokshitzer, Gribov, Lipatov, Altarelli and Parisi.

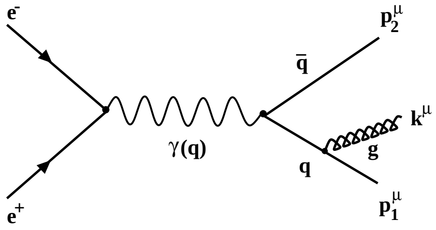

Figure 1.4: Soft gluon emission from a $q\bar{q}$ pair.

$$P^0_{gg}(z) = 2\,\mathbf{C_A} \cdot \left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right] + \delta(1-z)\frac{11C_A - 4N_f T_R}{6} \tag{1.5}$$

where C_F, C_A and T_R are the QCD **colour factors** (also called Casimir factors) given by:

$$C_A = N_C = 3, \quad C_F = \frac{N_C^2 - 1}{2N_C} = 4/3 \quad \text{and} \quad T_R = 1/2 \tag{1.6}$$

The values of the colour factors follow directly from the SU(3) group structure of QCD and determine the relative strengths of the parton branchings. Since $C_A/C_F = 9/4$, more branchings are expected for jets originating from gluons than for jets originating from quarks[5]. The comparison of gluon and quark jets thus allows for basic test of QCD and offers a direct handle on the transition of colour charged partons into colourless hadrons.

An important key idea in the pQCD parton branching formalism, essential for the modelling of the internal jet structure, is the so-called **angular ordering property** [5] of soft gluon emission (which is an example of a coherence effect common to all gauge theories): each subsequent branching has a smaller opening angle than the previous one. The simplest example in order to understand the angular ordering property is given by the emission of a soft gluon from a $q\bar{q}$ pair (see Fig. 1.4). In the soft limit (with the assumption of massless partons) the gluon emission amplitude factors out from the lowest order $q\bar{q}$ pair production cross-section $\sigma_0(e^+e^- \to q\bar{q})$:

$$\begin{aligned} d\sigma(e^+e^- \to q\bar{q}g) &= \frac{4\pi\alpha_S C_F d^3k}{(2\pi)^3 2k}\frac{2p_1 \cdot p_2}{(k\cdot p_1)(k\cdot p_2)}\sigma_0(e^+e^- \to q\bar{q}) = \\ &= \frac{\alpha_S C_F dk d\Omega}{2\pi^2 k}\frac{1-\cos\theta_{p_1p_2}}{(1-\cos\theta_{kp_1})(1-\cos\theta_{kp_2})}\sigma_0(e^+e^- \to q\bar{q}) \end{aligned} \tag{1.7}$$

where p_1, p_2 and k are the momenta of the quark, anti-quark and the gluon, respectively. θ_{kp_1} (θ_{kp_2}) is the angle between the direction of motion of the gluon and the quark (anti-quark), $\theta_{p_1p_2}$ is the angle between the quark and the anti-quark and Ω is the solid angle of $\vec{k}$. The angular factor in Eq. 1.7 is positive definite and singular both for $\theta_{kp_1} \to 0$ and $\theta_{kp_2} \to 0$. It can be re-written as the sum of two terms each singular in one of these limits:

$$\frac{1-\cos\theta_{p_1p_2}}{(1-\cos\theta_{kp_1})(1-\cos\theta_{kp_2})} \equiv (1|2) = \frac{1}{2}\left\{(1|2) + \frac{1}{1-\cos\theta_{kp_1}} - \frac{1}{1-\cos\theta_{kp_2}}\right\} + \{1 \to 2\}$$

[5]In the following jets originating from quarks (gluons) will be referred to as quark (gluon) jets.

By integrating over the azimuthal angle ϕ (around the direction of $\vec{p_1}$), making use of the relation:

$$\cos\theta_{kp_2} = \cos\theta_{kp_1}\cos\theta_{p_1p_2} + \sin\theta_{kp_1}\sin\theta_{p_1p_2}\cos\phi$$

one finds that the angular factor vanishes outside a cone with $\theta_{kp_1} > \theta_{p_1p_2}$:

$$\frac{1}{2\pi}\int_0^{2\pi} d\phi \frac{1}{2}\left\{(1|2) + \frac{1}{1-\cos\theta_{kp_1}} - \frac{1}{1-\cos\theta_{kp_2}}\right\} = \frac{1}{1-\cos\theta_{kp_1}}\Theta(\theta_{p_1p_2} - \theta_{kp_1})$$

where $\Theta(x)$ is the step function[6]. So, with the angular restriction $\theta_{kp_1} > \theta_{p_1p_2}$ one can use the term $1/(1-\cos\theta_{kp_1})$ for the emission amplitude. Although the amplitude, before azimuthal integration, is a more complicated quantity and only identical to the quantity $1/(1-\cos\theta_{kp_1})$ at very small angles, the approximation is exact in evaluating angle-averaged quantities.

The cross-section in Eq. 1.7 is of the form:

$$d\sigma_{n+1} = d\sigma_n \frac{d\omega}{\omega}\frac{d\Omega}{2\pi}\frac{\alpha_S}{2\pi}\sum_{i,j} C_{ij}W_{ij}$$

where the sum runs over all pairs of external lines $\{i,j\}$, ω is the gluon energy, $d\Omega$ is the element of solid angle for the emitted gluon, C_{ij} is a colour factor which can be calculated and W_{ij} is the so-called radiation function. For the case of one gluon emission in Eq. 1.7 W_{ij} is given by:

$$W_{ij} = \frac{1-\cos\theta_{ij}}{(1-\cos\theta_{iq})(1-\cos\theta_{jq})} \quad \text{with} \quad i = p_1,\ j = p_2,\ q = k$$

and was written as the sum of two terms:

$$W_{ij} = W_{ij}^{[i]} + W_{ij}^{[j]}$$

If one considers a soft gluon emission from the $q\bar{q}g$ system (see Fig. 1.5), it can be shown that the soft gluon radiation pattern is determined by the expression [3]:

$$W = -\mathbf{Q_i}\cdot\mathbf{Q_j}W_{ij} - \mathbf{Q_j}\cdot\mathbf{Q_k}W_{jk} - \mathbf{Q_i}\cdot\mathbf{Q_k}W_{ik}$$

where the products $\mathbf{Q_a}\cdot\mathbf{Q_b}$ represent the colour factors C_{ab} with the condition $\mathbf{Q_a^2} = C_F$ for a quark and $\mathbf{Q_a^2} = C_A$ for a gluon. For the configuration in which i and j are close in angle, forming a system l with colour charge $\mathbf{Q}_l = \mathbf{Q}_i + \mathbf{Q}_j = -\mathbf{Q}_k$ and approximating the direction of $\vec{i}$ and $\vec{j}$ by the direction $\vec{l}$, one can deduce the following expression:

$$W \simeq \mathbf{Q}_i^2 W_{ij}^{[i]} + \mathbf{Q}_j^2 W_{ij}^{[j]} + \mathbf{Q}_k^2 W_{lk}^{[k]} + \mathbf{Q}_l^2 \tilde{W}_{lk}^{[ij]}$$

where:

$$\tilde{W}_{lk}^{[ij]} = \begin{cases} W_{lk}^{[l]} & \text{for } \theta_{lq} > \theta_{ij} \\ 0 & \text{for } \theta_{lq} < \theta_{ij} \end{cases}$$

So, each parton i, j and k radiates in proportion to its colour charge, but when i and j are close in angle, their incoherent contributions are limited (after azimuthal averaging) to a cone of half-angle θ_{ij}. At larger angles, out of the direction of $\vec{k}$, they give a coherent contribution in proportion to their combined colour charge $\mathbf{Q}_l^2$, which can be computed as if it came from an internal line of momentum $\vec{p}_l = \vec{p}_i + \vec{p}_j$. By generalizing this argument to multiple soft gluon emissions, one can show that the emission angles along each branch can be ordered according to decreasing values, i.e. $\theta_{p_1p_2} > \theta_{k_1p_1} > \theta_{k_2k_1} > \theta_{k_3k_2} > ....$ This property is the basis for the so-called 'parton shower' algorithms which are used to describe the perturbative phase of the jet evolution.

[6] The step function $\Theta(x)$ is defined as $\Theta(x) = 0$ for $x \leq 0$ and $\Theta(x) = 1$ for $x > 0$.

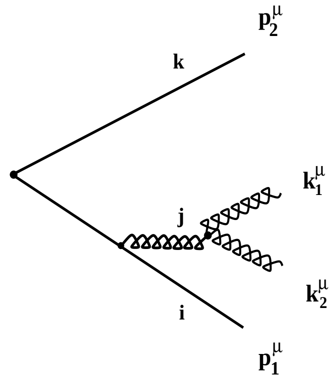

Figure 1.5: Soft gluon emission in a three parton system $q\bar{q}g$.

1.2 Quark and Gluon Jet Differences

In the frame of pQCD, the average number of hadrons of type h in a jet initiated by a parton i at a scale t can be expressed as:

$$\langle n_h(t)\rangle_i = \int_0^1 dx D_i^h(x,t) \tag{1.8}$$

where the **fragmentation function** $D_i^h(x,t)$ describes the single particle distribution in the final state[7]. The variable x is usually referred to as the scaled hadron energy. A possible definition of x, commonly used in both e^+e^- and hadron(lepton)-hadron scattering, is $x = E_h/E_{jet}$. The fragmentation functions satisfy the DGLAP evolution equations:

$$t\frac{\partial}{\partial t}D_i^h(x,t) = \sum_j \int_x^1 \frac{dz}{z}\frac{\alpha_S}{2\pi}P_{ji}(z,\alpha_S(t))D_j^h(x/z,t) \tag{1.9}$$

thus, as in the case of the parton distribution functions, it is possible to predict their value at an arbitrary energy scale once the functions have been parameterized at some fixed scale t_0.

The relation between the asymptotic behaviour (i.e. at very large energy scale $t \gg t_0$) for the averaged multiplicity of any type of object in a quark and gluon jet can be expressed as [3]:

$$\langle n(t)\rangle_q \sim \frac{C_F}{C_A}\langle n(t)\rangle_g \propto \exp\left[\frac{1}{b}\sqrt{\frac{2C_A}{\pi\alpha_S(t)}}\right] \tag{1.10}$$

In addition the asymptotic behaviour for the multiplicity distribution widths, $M_{q,g}(t)$ is given by:

$$M_q(t) \sim (1 + \frac{C_F}{3C_A})\langle n(t)\rangle_q^2 \quad \text{and} \quad M_g(t) \sim \frac{4}{3}\langle n(t)\rangle_g^2 \tag{1.11}$$

This means that the overall particle multiplicity in a gluon jet should become $C_A/C_F = 9/4$ times greater than that in a quark jet for $t \gg t_0$. The multiplicity distribution should also become wider for a gluon jet, by the factor $\sqrt{C_A/C_F}$. As a consequence of the higher multiplicity,

[7] More explicitly, the function $D_i^h(x,t)$ represents the probability for a parton i to form a jet that contains the hadron h, carrying a fraction x of the longitudinal momentum of the parton i.

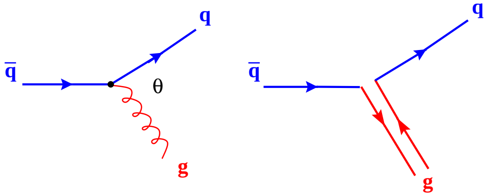

Figure 1.6: A three-parton (jet) event and the colour flow to $\mathcal{O}(\alpha_S \mathrm{N}_\mathrm{C}^2)$.

the fragmentation function of a gluon jet is expected to be softer, i.e. more concentrated at low values of the energy fraction x, than that of a quark jet.

Figure 1.7 shows a measurement of the charged hadron multiplicities for gluon and quark jets as a function of the jet scale Q_jet in 3-jets events from $\mathrm{e}^+\mathrm{e}^-$ scattering at LEP [11] where the characteristic difference predicted by pQCD is clearly seen[8]. The definition of the jet scale $Q_\mathrm{jet} = E_\mathrm{jet}\,\sin(\theta/2)$ involves the angle θ between the quark (gluon) jet and the closest jet to it in order to account for coherence effects in the parton branchings [12]. The way the scale Q_jet incorporates the topological dependence can be seen, for example, at leading order in α_S for the three parton $q\bar{q}g$ vertex (see Fig. 1.6): since the colour charge is conserved by QCD, the gluon can be represented as a $q'\bar{q}'$ pair that compensates exactly the total colour charge of the initial $q\bar{q}$ pair. As already discussed in Sec. 1.1 one expects that as the quark and gluon, for example, come close to each other their colour charges mutually shield one another and the amount of subsequent gluon radiation is reduced, i.e. the quark and gluon are not independent but form a coherent system. The colour shielding is of course minimal for back-to-back partons, and the scale Q_jet becomes equal to the parton energy in that case.

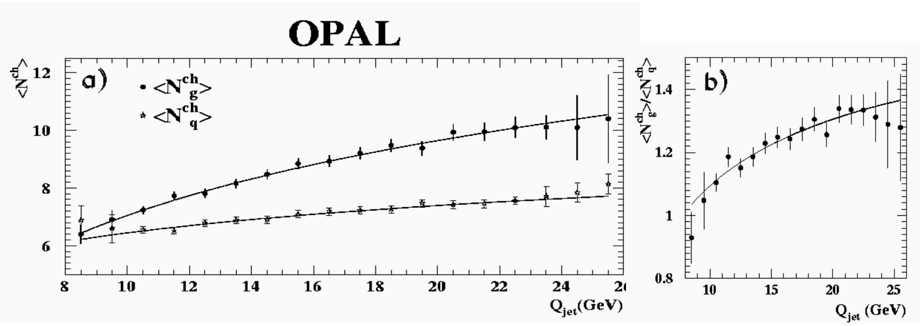

Figure 1.7: Average multiplicity of charged hadrons for pure gluon and quark jets (a) and the ratio of the average charged hadron multiplicities in gluon and quark jets as a function of the scale Q_jet (b). The continuous curves are from a fit to a phenomenological formula. Note that the ratio $\left\langle n_g^{ch}\right\rangle / \left\langle n_q^{ch}\right\rangle$ is smaller than $C_A/C_F = 9/4$ as expected, since the latter represents the value for the ratio at infinite energy ($t \to \infty$).

In Fig. 1.8 the fragmentation function for both gluon and light quark (u, d, s) jets is shown [13]. A roughly exponential fall-off is observed which is more pronounced for gluon jets than

[8] The jet definition will be discussed in the next section.

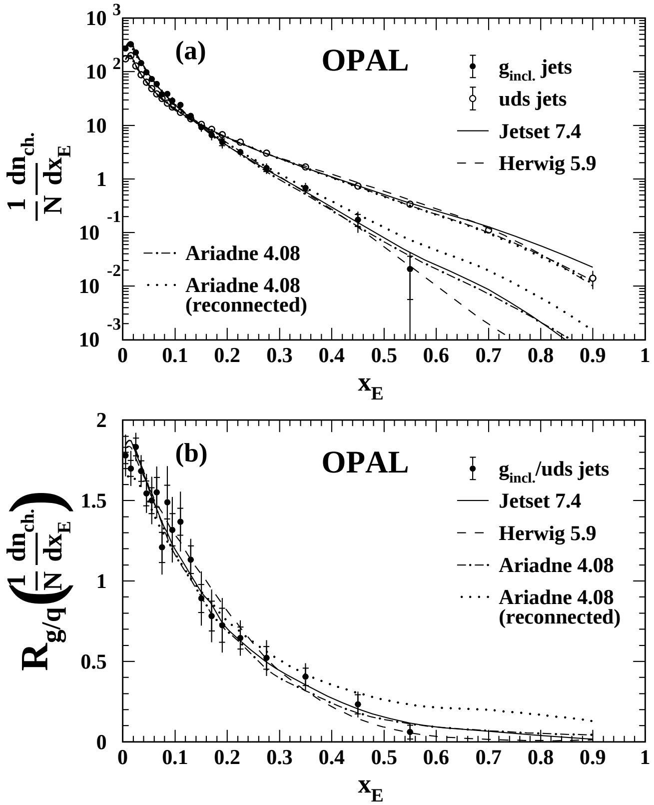

Figure 1.8: Gluon and quark fragmentation functions to charged hadrons as a function of the scaled hadron energy x_E (upper plot). The lower plot shows the ratio of the x_E distribution obtained from quark and gluon jets. The average jet energy is $\simeq 40.1$ GeV for gluon jets and $\simeq 45.6$ GeV for quark jets. The predictions of various parton shower Monte Carlo event generators (see chapter 3) are also shown.

for quark jets. At large x_E the gluon distribution is suppressed with respect to the quark one by more than one order of magnitude, i.e. the production of fast, so called 'leading' particles is strongly suppressed in gluon jets. At small x_E values, however, the gluon to quark ratio reaches values up to 1.8.

Another characteristic difference between quark and gluon jets is that gluon jets are expected to be 'broader' than quark jets. This property could be seen even without a precise formulation of the jet definition. In the first jet cross section definition employed in $\mathrm{e}^+\mathrm{e}^-$ annihilation by Sterman and Weinberg [7], the final state is classified as two-jet like if all but a fraction ϵ of the total available energy is contained in a pair of cones of half-angle δ (see Fig. 1.9). By means of this relatively elementary definition, the angular size δ_q of a quark jet in lowest order can be

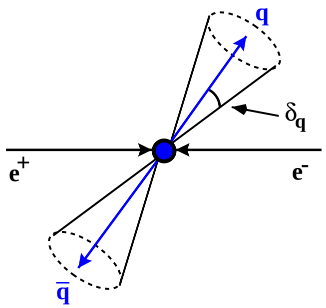

Figure 1.9: Schematic representation of a two-jet final state in e^+e^- annihilation.

shown to be:

$$\delta_q \sim \exp\left\{-\frac{\pi(1-f_2)}{4\,\mathbf{C_F}\,\alpha_\mathbf{S}(\mathbf{s})\,|\ln\epsilon|}\right\}$$

where the two-jet fraction is defined as the ratio of the two-jet and the total cross section ($f_2 = \sigma_2/\sigma_{tot}$) and s is the centre of mass energy. Since $1/\alpha_S(s)$ increases logarithmically with the energy s, the angular jet width, for a fixed fraction f_2, decreases like a negative power of s. For a final state consisting of two gluon jets, the analogous calculation gives a factor C_A in place of C_F and therefore:

$$\delta_g \sim \delta_q^{\mathbf{C_F/C_A}}$$

Since the ratio $C_F/C_A < 1$ the angular size for a gluon jet $\delta_g > \delta_q$ and moreover its power decrease as a function of s is less rapid (see Fig. 1.10).

1.3 Jet Definition

The first evidence for jet structure in hadron production by e^+e^- annihilation was reported in 1975 [6] using an intuitive and very qualitative jet definition (i.e. 'large amount of hadronic energy in a small angular region'). However, in order to perform a quantitative analysis, one needs a jet algorithm which is able to specify a jet configuration unambiguously both in

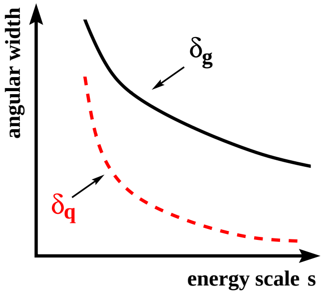

Figure 1.10: Schematic representation of the power decrease of the angular width for quark and gluon jets as a function of the energy scale s.

theoretical calculations and in experiments (starting from the hadrons detected in the final state). Therefore, it is necessary to explain the way a jet is defined before introducing the physical observables used in this analysis to investigate the differences between quark and gluon jets.

A clear difficulty in choosing a jet definition is that the latter most likely depends on the kind of high energy scattering under consideration. Although the basic hard scattering processes studied in hadron(lepton)-hadron and in electron-positron collisions can be described within the same theoretical framework, the overall event structure is quite different. In e^+e^- annihilation the initial state is purely electromagnetic and the entire final state can be thought of as arising from the virtual photon (or Z^0) which creates a $q\bar{q}$ pair. Therefore, all the hadrons in the final state are associated with the hard scattering process. In contrast, in hadron(lepton)-hadron collisions, there are different contributions to the overall final state: only a fraction of the hadrons are associated with the hard scattering process, and the remainder form part of the beam remnants or come from eventual soft (i.e. small transverse momentum) interactions of the remaining partons in the incident hadrons ('**underlying event**'). Moreover the partons participating in the hard scattering process produce additional initial QCD radiation (**initial state branchings**) that is not present in e^+e^- annihilation. These differences in the event structure have led naturally to differences in the jet definition and the corresponding algorithms.

However, as a general rule, any jet definition should satisfy the conditions of being:

(i) collinear and infrared safe;

(ii) easy to measure in experimental analyses from the hadronic final state and to calculate order-by-order in perturbation theory from the partonic final state;

(iii) less sensitive to hadronization corrections. Having small hadronization corrections entails a closer correspondence between distributions of the final state partons and those of the hadrons.

In the case of hadron-hadron and lepton-hadron collisions, the jet definition should also fulfill the additional requirements of being:

(iv) able to factorize the collinear singularities arising from the initial state branchings into universal distributions;

(v) not strongly affected by contamination from hadron remnants and underlying soft event.

The requirement (i) is related to the fact that in pQCD, divergences arise when two massless partons are emitted collinearly or when a parton with very low energy is emitted. To cancel these divergences in the calculation, it is necessary that a collinear pair of particles is treated identically to a single particle with their combined momentum and that the emission of soft partons does not affect at all the observable $R(p_1, p_2, ..., p_n)$, where $p_{i=1,..,n}$ are the four-momenta of the final state particles. Mathematically this can be expressed as:

$$R(p_1, p_2, ..., p_n) = R(p_1 + p_2, ..., p_n) \quad \text{if} \quad \vec{p}_1 \parallel \vec{p}_2 \qquad \textbf{collinear safe}$$

and

$$R(p_1, p_2, ..., p_n, p_{n+1}) = R(p_1, p_2, ..., p_n) \quad \text{if} \quad p^0_{n+1} \to 0 \quad \textbf{infrared safe}$$

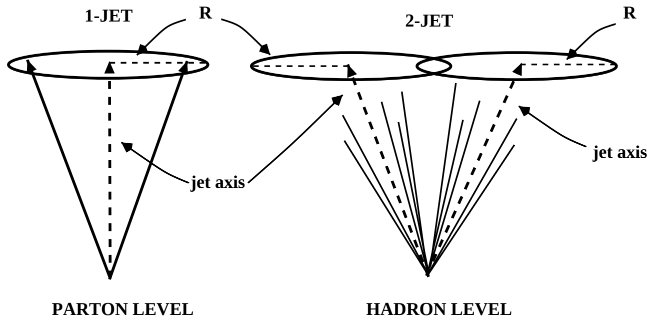

Figure 1.11: Schematic diagram of a jet configuration in which cone-type algorithms give a different result at parton and hadron level. The left drawing shows a jet of radius R consisting of two partons at a distance $\sim 2R$ from each other. The right drawing shows a 2-jet configuration originating from the same two parton system: after the parton shower and the hadronization process the cone-type algorithm reconstructs two different jets sharing a fraction of the total energy.

At present there are essentially two classes of jet algorithms in use: cone-type algorithms 'à la Sterman-Weinberg' [7] and cluster-type algorithms of the type first introduced by the JADE collaboration [8]. In cone-type algorithms jets are defined by finding a direction (jet axis) which maximizes the amount of energy in a cone of fixed radius R drawn around the chosen direction. Although conceptually simple to define, the cone-type algorithms suffer from ambiguities related to the overlapping of two cones. In the right drawing of Fig. 1.11 one can see two jets which are separated by less than 2R in the $(\eta,\, \phi)$ space. In a final state with many hadrons the two hadronic sprays are recognized as separate jets. In a final state consisting of only few partons (as in the left drawing of Fig. 1.11) the situation is unclear. Although it is clearly natural to consider the two partons as separate jets, both partons fit into a single cone of radius R and are therefore combined into one jet.

Cluster-type algorithms [9] follow a different approach than cone-type ones: instead of globally finding a jet axis direction, they start by finding pairs of particles that are 'nearby' in phase-space and merging them together to form new 'pseudoparticles'. The procedure continues iteratively until the event consists of a few well-separated pseudoparticles, which are the output jets. There is however no unique definition of 'closeness' in phase-space and several algorithms using different definitions have been developed. Nowadays, cluster-type algorithms are preferred because they allow for an unambiguous and exhaustive assignment of hadrons to jets using the same procedure both in theoretical calculations and in experimental practice.

A cluster-jet algorithm which fulfills the requirements (i)-(v) can be implemented by defining:

1. a separation variable (energy-angle resolution) d_{kl}
2. a recombination procedure
3. a hard scattering scale

The test variable d_{kl} is needed to specify whether or not two hadrons $(h_k,\, h_l)$ belong to the same cluster, whilst the recombination procedure defines how the cluster variables are related

to those of the hadrons belonging to it. The introduction of a hard scattering scale is necessary only in the case of hadron(lepton)-hadron collisions since for e^+e^- annihilation, no factorization of hard and soft sub-processes has to be performed and the hard scattering scale coincides with the centre-of-mass energy.

1.3.1 The Longitudinally Invariant k_T Cluster Algorithm

The longitudinally invariant k_T cluster algorithm [10], which has been used to define the jets in the analysis presented in this thesis, starts with a list of 'particles' (e.g. partons, hadrons or calorimeter cells), each of which is characterized by its transverse energy $E_{T,i} = |\vec{p_{T,i}}|$, pseudorapidity $\eta_i = -\ln\tan(\theta_i/2)$ and azimuthal angle ϕ_i. The algorithm proceeds according to the following steps:

1. for every pair of particles, define a closeness d_{ij}:

$$d_{ij} = \min(E_{T,i}\,,\, E_{T,j})^2 \cdot \Delta R_{ij}^2 \quad \text{where} \quad \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$$

   Note that for small opening angles, the quantity $\Delta R_{ij} \ll 1$ and therefore:

$$\min(E_{T,i}\,,\, E_{T,j})^2 \cdot \Delta R_{ij}^2 \sim \min(E_i\,,\, E_j)^2 \cdot \Delta\theta_{ij}^2 \sim k_{T,ij}^2$$

   where $k_{T,ij}$ represents the transverse momentum of particle i relative to particle j

2. for every particle, define a closeness to the beam direction:

$$d_{i,\mathrm{beam}} = E_{T,i}^2 \cdot R_0^2$$

   where R_0 is an adjustable parameter of the algorithm which plays the rôle of a jet radius in the $(\eta,\, \phi)$ plane, assumed to be of the order of 1[9]. In the limit of small opening angles, the variable $d_{i,\mathrm{beam}}$ reduces to the square of the transverse momentum $k_{T,i,\mathrm{beam}}$ of the particle i with respect to the beam direction

3. find the smallest element $d_{\mathrm{min}} = \min\{d_{ij},\, d_{i,\mathrm{beam}}\}$

4. if $d_{\mathrm{min}} = d_{ij}$, particles i and j are merged together into a new pseudoparticle k with:

$$E_{T,k} = E_{T,i} + E_{T,j}$$

   and

$$\eta_k = (E_{T,i} \cdot \eta_i + E_{T,j} \cdot \eta_j)\,/\,E_{T,k} \quad , \quad \phi_k = (E_{T,i} \cdot \phi_i + E_{T,j} \cdot \phi_j)\,/\,E_{T,k}$$

5. if $d_{\mathrm{min}} = d_{i,\mathrm{beam}}$, the 'proto-jet' i is complete and is added to the output list

The algorithm continues from step (1) considering pseudoparticles on an equal footing with particles. These steps are iterated until all proto-jets are complete and put into a list with successively larger values of transverse energy E_T. Only the proto-jets with a transverse energy larger than a specified threshold value, $E_T \geq E_T^{\mathrm{min}}$, are defined as jets. It should be noticed that a particle or pseudoparticle is merged with its nearest neighbour if and only if their invariant

[9]The value $R_0 = 1$ is strongly preferred theoretically as it treats initial and final-state radiation on equal footing.

opening angle is smaller than R_0, i.e. all opening angles within each jet are $< R_0$ and all opening angles between jets are $> R_0$. Thus the jets produced are phenomenologically not very dissimilar to those defined by a cone-type algorithm. However, the invariant k$_T$ cluster algorithm is less influenced by soft particles than the cone-type algorithms, which results in smaller hadronization and detector corrections. Qualitatively one can say that the invariant k$_T$ cluster algorithm pays more attention to the core of the jet and only merges neighbouring particles if they are close enough, whereas the cone algorithm, in order to maximize the jet energy, pulls in as much neighbouring energy as possible.

1.4 The Internal Jet Structure

The internal structure of quark and gluon jets, which in the frame of pQCD is ascribed to the emission of individual partons according to the SU(3) colour factors (as discussed in sections 1.1, 1.2), is an exclusive event property which allows for a direct test of pQCD. In this section the physical observables, which have been measured in the analysis presented in this thesis, will be defined.

The most common way of resolving the internal jet structure, inspired by the cone-type algorithms, is to measure the energy distribution around the jet axis direction. The physical observable used is the so-called **jet shape** which will be defined in Sec. 1.4.1. Close to the jet axis, the jet shape is dominated by collinear gluon emission, whereas at large distances (angles) from the jet axis, the jet shape reflects large angle gluon emission, which can be calculated perturbatively. However, the cluster-type algorithms naturally suggest an alternative, additional way of analyzing the internal structure of a jet that is much closer to the partonic picture of how that structure arises, i.e. by resolving 'sub-objects' within a jet making use of a resolution variable. The physical observable used in this case is the so-called **mean subjet multiplicity** which will be defined in Sec. 1.4.2.

1.4.1 The Jet Shape

The **integrated jet shape** $\psi(r,R)$ is defined as the average fraction of the jet's transverse energy that lies inside an inner cone of radius r, in the (η, ϕ) space, concentric with the jet axis, as shown in the left drawing of Fig. 1.12:

$$\psi(r,R) = \frac{\int_0^r dr' dE_T(r')/dr'}{\int_0^R dr' dE_T(r')/dr'} = \frac{1}{N_{\text{jets}}} \sum_{\text{jets}} \frac{E_T^{\text{jet}}(r)}{E_T^{\text{jet}}(r=R)} \qquad (1.12)$$

where $E_T^{\text{jet}}(r)$ is the transverse energy within the inner cone of radius r, N_{jets} is the total number of jets in the sample and the jet radius $R = 1$ in the case of the invariant k$_T$ cluster algorithm[10]. Narrower jets are characterized by larger values of $\psi(r,R)$. Thus, from what was already discussed in Sec. 1.2, one expects:

$$\psi_{\text{quark}}(r,R) > \psi_{\text{gluon}}(r,R) \quad \forall\, r < R$$

The jet shapes can be also analyzed in terms of the **differential jet shape**, i.e. the average fraction of the jet's transverse energy that lies in an angular annulus concentric with the jet

[10] Note that by definition, $\psi(R,R) = 1$.

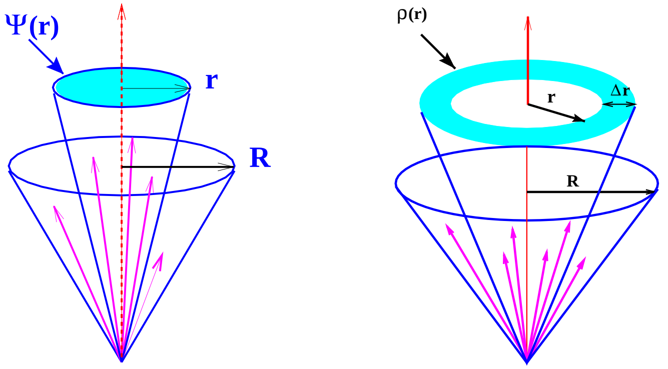

Figure 1.12: Schematic representation of the integrated (left) and differential (right) jet shape.

axis (as shown in the right drawing of Fig. 1.12):

$$\rho(r) \equiv d\psi/dr = \frac{1}{N_{\text{jets}}} \sum_{\text{jets}} \frac{E_T^{\text{jet}}(r, r+\Delta r)}{E_T^{\text{jet}}(r=R)} \qquad (1.13)$$

where $E_T^{\text{jet}}(r, r+\Delta r)$ is the transverse energy of the jet which lies in the region between the two cones of radius r and $r+\Delta r$, respectively. Narrower jets are characterized by larger values of $\rho(r)$ at small values of the radius r and by smaller values of $\rho(r)$ at large values of the radius r. Figure 1.13 shows the jet shape measured in e^+e^- annihilation [14] for light quark and gluon jets. Large differences are observed between the two samples, the gluon jets being much broader than the quark ones, as predicted by pQCD.

1.4.2 The Subjet Multiplicity

After identifying a jet of a given transverse energy E_T, the invariant k_T cluster algorithm is rerun only on those particles that were assigned to the jet. The clustering procedure is stopped when all values of d_{ij} satisfy the condition:

$$d_{ij} > y_{\text{cut}} \cdot E_T^2$$

i.e. when all the internal relative transverse momenta are greater than $\sqrt{y_{\text{cut}}} \cdot E_T$ and the number of sub-structures inside the jet is counted. For a resolution scale of $y_{\text{cut}} \sim 1$ the jet always consists of only one subjet (the jet itself), i.e. the subjet multiplicity $n_{\text{sbj}} \to 1$, whereas for $y_{\text{cut}} \to 0$ every hadron is eventually considered as a separate subjet, i.e. the subjets multiplicity $n_{\text{sbj}} \to N_{\text{hadrons}}$, where N_{hadrons} represents the number of hadrons within the jet. This is qualitatively illustrated in Fig. 1.14. Thus, varying the resolution scale y_{cut} allows to study in some detail the transition of the colour charged fields (partons) into colourless hadrons. The mean subjet multiplicity is defined as:

$$\langle n_{\text{sbj}} \rangle (y_{\text{cut}}) = \frac{1}{N_{\text{jet}}} \sum_{i=1}^{N_{\text{jets}}} n_{\text{sbj}}^i(y_{\text{cut}}) \qquad (1.14)$$

From what was already discussed in Sec. 1.2, one expects that:

$$\langle n_{\text{sbj}} \rangle_{\text{gluon}} > \langle n_{\text{sbj}} \rangle_{\text{quark}} \quad \forall\, y_{\text{cut}} \ll 1$$

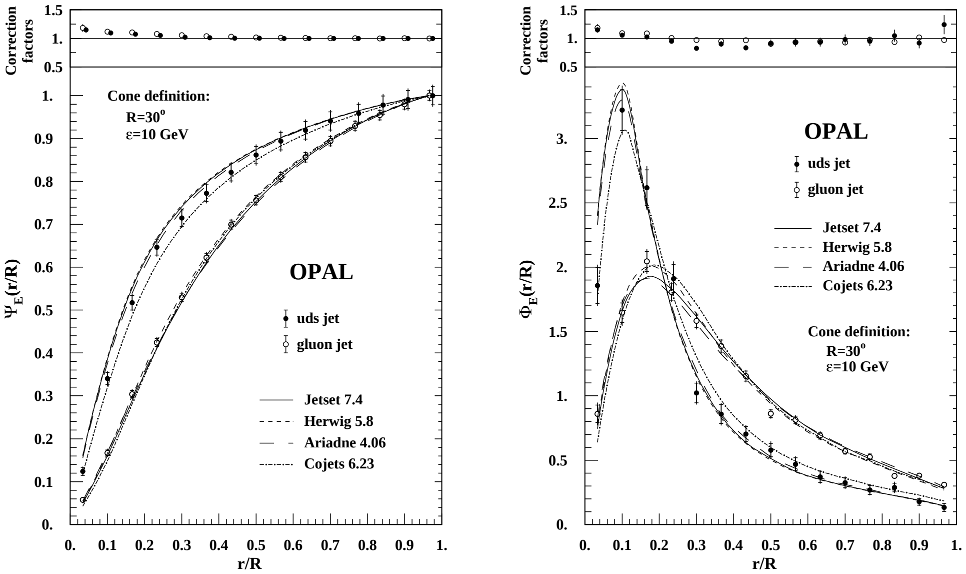

Figure 1.13: Integrated (left) and differential (right) jet shape for gluon and light quarks measured in e^+e^- annihilation [14]. The jets are defined using a cone-type algorithm which employs two resolution parameters: the cone size R and the minimum particle energy within the cone ϵ. Comparison with different theoretical predictions (LO Monte Carlo simulations) are shown. The upper plots show the experimental correction factors from detector to hadron level.

1.5 Definition of the Measurement

1.5.1 Lepton Proton Scattering

The scattering of a high energy lepton off a hadron target (e.g. a proton) is mediated by the exchange of a virtual gauge boson. In a neutral current interaction (NC), e(k)+p(P) $\rightarrow$ e(k$'$)+X(P$'$), the exchanged boson is either a virtual photon γ^* or a Z^0 and the final state lepton is a charged electron (or positron), whereas in a charged current interaction (CC), e(k)+p(P) $\rightarrow$ ν_e(k$'$)+X(P$'$), the exchanged boson is a $W^\pm$ and the final state lepton is a neutrino (or anti-neutrino) which escapes the detector undetected. Figure 1.15 shows a schematic diagram of a lepton proton scattering event.

For unpolarized beams, at a given centre-of-mass energy, $\sqrt{s} = \sqrt{(k+P)^2}$, only two variables are needed to completely describe the kinematics of the lepton proton scattering. Usually two of the following Lorentz invariant quantities are used:

$$Q^2 = -q^2 = -(k-k')^2$$
$$y = \frac{P \cdot q}{P \cdot k}$$

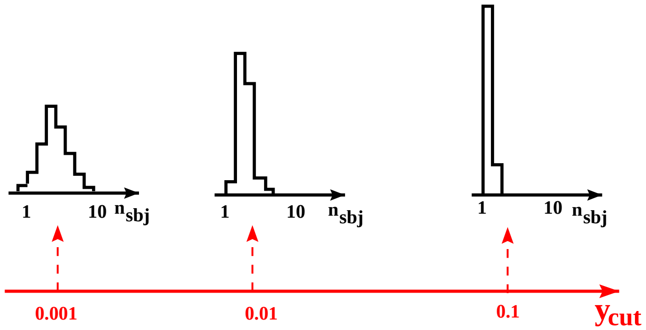

Figure 1.14: Schematic representation of the subjet multiplicity distribution for increasing values of the resolution scale $y_{\rm cut}$. For small values of $y_{\rm cut}$ the mean value ($\langle n_{\rm sbj} \rangle$) and the spread ($\sigma_{n_{\rm sbj}}$) of the distribution are relatively large, whereas for larger values of $y_{\rm cut}$ they become smaller until a resolution scale is reached at which it is no longer possible to resolve sub-structures inside the jet.

$$x = \frac{Q^2}{2\,P \cdot q} \qquad (1.15)$$

where Q^2 is the negative squared four momentum transfer from the incoming lepton to the proton and can be thought of as a measure of the resolving power of the exchanged boson. In the proton rest frame, y represents the fractional energy transferred from the lepton to the proton. The Bjorken scaling variable x, in the quark parton model [2], can be interpreted as the momentum fraction of the proton carried by the struck parton. Neglecting the lepton and proton masses, the three variables Q^2, x and y are related to each other by the following equation:

$$Q^2 \simeq x \cdot y \cdot s \qquad (1.16)$$

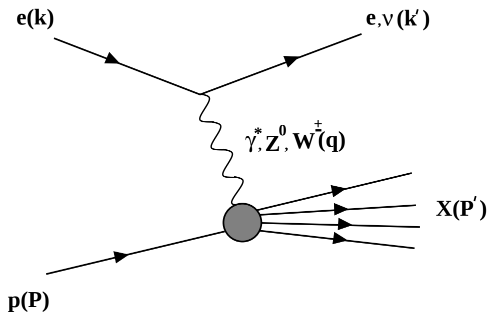

Figure 1.15: Schematic diagram of a lepton proton scattering event. The four vectors of the particles are given in parentheses.

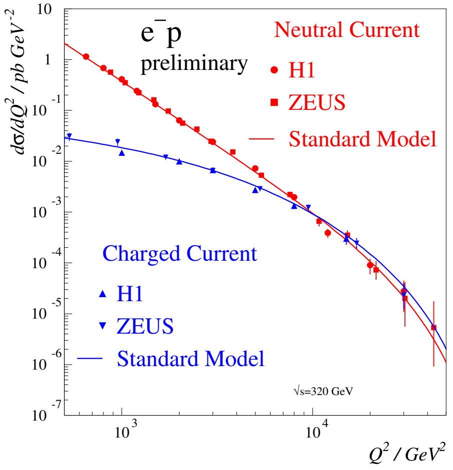

Figure 1.16: Measurements of the NC and CC cross-sections at HERA.

and the square of the invariant mass of the final hadronic state, W^2, is given by:

$$W^2 = (P+q)^2 \simeq Q^2 \cdot \left(\frac{1}{x} - 1\right) \tag{1.17}$$

The lepton-proton cross sections (NC and CC) fall rapidly as a function of Q^2: at low Q^2, the photon propagator dominates the NC cross section and only when Q^2 is sufficiently large is the contribution from Z^0 significant. This can be seen in Fig. 1.16, where the NC and CC cross section measured by the HERA experiments (ZEUS and H1) are shown. At low Q^2 values, the NC cross section is about two orders of magnitude larger than the CC cross section and only at large Q^2 values (where the strengths of the weak and electromagnetic forces become equal), the two cross sections become of the same order.

1.5.2 Photoproduction Processes

Photoproduction at HERA refers to processes where the electron is scattered under a very small angle emitting a quasi-real photon ($Q^2 \approx 0$ GeV^2), which than interacts with the proton. Already the first γp scattering experiments in the 1960's showed striking similarities to hadron-hadron interactions which can be explained by assuming that the photon can exhibit an internal structure similar to that of hadrons, except that it occurs with a relative small probability of the order of $\alpha_{\text{em}} \simeq 1/137$. The electromagnetic coupling of the photon to charged particles allows it to fluctuate into a quark-antiquark pair ($q\bar{q}$) pair. The physical photon can thus be thought of as a superposition of two type of states [17], a bare photon $|\gamma_b\rangle$ and a hadronic component $|h\rangle$:

$$|\gamma\rangle = \sqrt{Z_3}|\gamma_b\rangle + \sqrt{\alpha_{\text{em}}}|h\rangle\ , \tag{1.18}$$

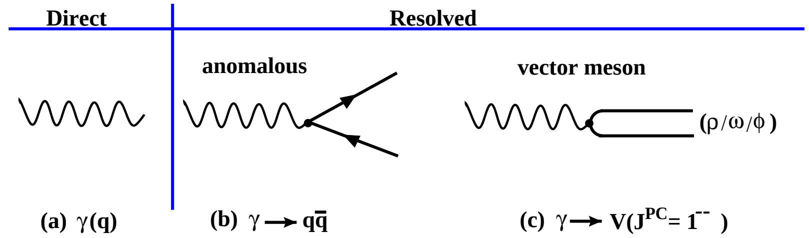

Figure 1.17: Different components of the photon: (a) point-like (b) anomalous and (c) hadron-like.

where Z_3 assures the correct normalization of $|\gamma\rangle$. Because of conservation laws, $|h\rangle$ must have the same quantum numbers as the photon. In the Vector Meson Dominance (VMD) model the hadronic component of the photon is assumed to be:

$$|h\rangle = a|\rho^0\rangle + b|\omega\rangle + c|\phi\rangle$$

where $|\rho\rangle$, $|\omega\rangle$ and $|\phi\rangle$ are the three lightest vector mesons. The fluctuation time can be estimated by using the uncertainty principle to be:

$$t_f \approx \frac{2E_\gamma}{Q^2 + m_V^2}$$

In the proton rest frame the interaction time t_i is of the order of the proton radius $r_p \simeq 1$ fm, and therefore for small values of Q^2 and m_V^2 the fluctuation time $t_f \gg t_i$ and the interaction occurs between the $q\bar{q}$ pair and the proton. The coupling of the photon to an unbound $q\bar{q}$ adds an extra component to the partonic structure of the photon, which is called **anomalous** component and is, in principle, perturbatively calculable (see Fig. 1.17).

For very low virtualities Q^2, the time fluctuation of the exchanged photon can be considered large with respect to the characteristic time of the hard sub-process, the electron beam can be regarded as a source of approximately massless (quasi-real), collinear photons. In the limit of low Q^2, the ep cross-section mediated by photon exchange can be expressed in terms of the photon-proton cross-section as follows:

$$\frac{d\sigma^{\text{ep}}}{dy dQ^2} = f_\gamma^e(y,Q^2) \cdot \sigma_{\text{tot}}^{\gamma p}(W(y), Q^2)$$

where the total γp cross section reduces to the transversely polarized γp component[11], $\sigma_{\text{tot}}^{\gamma p} = \sigma_{\text{T}}^{\gamma p}$. The function $f_\gamma^e(y,Q^2)$ represents the photon flux and can be considered as the probability that an electron radiates a photon with a given virtuality Q^2 and fractional energy y. In the Equivalent Photon Approximation (EPA) [26] the function $f_\gamma^e(y,Q^2)$ is given by:

$$f_\gamma^e(y,Q^2) = \frac{\alpha_{\text{em}}}{2\pi Q^2}\left[\frac{1+(1-y)^2}{y} - 2\frac{1-y}{y}\frac{Q^2_{\text{min}}}{Q^2}\right] \tag{1.19}$$

where Q^2_{min} is the lowest virtuality that can be obtained:

$$Q^2_{\text{min}} = m_e^2 \frac{y^2}{1-y}$$

[11] Real photons ($Q^2 \to 0$ GeV^2) can only be transversely polarized.

By integrating the EPA equation from the minimum Q^2_{min} to some maximum, experimentally fixed virtuality Q^2_{max}, one obtains:

$$f^e_\gamma(y) = \frac{\alpha_{\text{em}}}{2\pi}\left[\frac{1+(1-y)^2}{y}\ln\frac{Q^2_{\text{max}}}{Q^2_{\text{min}}} - 2\frac{1-y}{y}\left(1-\frac{Q^2_{\text{min}}}{Q^2_{\text{max}}}\right)\right] \tag{1.20}$$

which is known as the Weizsäcker Williams Approximation (WWA).

In hard photoproduction processes, the hard scale of the interaction is given by the transverse energy of the outgoing partons that give rise to high E_T jets of hadrons. In leading order perturbation theory, two processes contribute to the photoproduction of jets: the **direct** process in which the photon couples directly to a parton in the proton, and the **resolved** process in which the photon acts as a source of partons, one of which, carrying only a fraction x_γ of the total photon momentum, interacts with a parton in the proton. Figure 1.18 shows the LO direct photoproduction diagrams ($2 \to 2$ parton scattering processes):

- in the **QCD Compton process** (left diagram in Fig. 1.18) the photon couples to a quark within the proton which then radiates a gluon. Thus a quark and a gluon jet with high transverse energy are produced in the final state
- in the **Boson Gluon Fusion (BGF) process** (right diagram in Fig. 1.18) the photon couples to a gluon in the proton which splits into a quark-antiquark pair. In this case the final state consists of two high transverse energy quark jets.

In hard photoproduction, the proton momentum fraction x, carried by the parton participating in the interaction, can be expressed as:

$$x \simeq M^2_{JJ}/W^2$$

where M_{JJ} represents the invariant mass of the two outgoing partons and W is the γp centre-of-mass energy. In the kinematic range in which the analysis presented in this thesis is performed, i.e. for $W \sim \mathcal{O}(10^2)$ GeV and with hadronic jets having a transverse energy of $E_T^{\text{jet}} \sim \mathcal{O}(10^1)$ GeV, the variable x is of the order of $\sim 10^{-2}$. For these x values, the gluon distribution function inside the proton is much higher than the valence and sea quark ones, as shown in Fig. 1.19 [16]. Thus, although the matrix elements for the two LO direct processes are of the same order of magnitude [15], BGF results to be the dominant process in direct photoproduction.

In Fig. 1.20 examples of resolved photoproduction diagrams are shown:

- the left diagram (a) in Fig. 1.20 shows the interaction of a gluon from the photon with a gluon from the proton, giving rise to a quark-antiquark pair in the final state
- the right diagram (b) in Fig. 1.20 shows, instead, the interaction of a quark from the photon with a gluon from the proton. In this case the final state consists of a quark and a gluon jet

In resolved photoproduction, processes with initial gluons are dominant[12]. Therefore a higher rate of gluon jets in the final state with respect to the direct processes is expected. At a fixed value of the photon energy y, because only a fraction x_γ of the photon momentum enters the hard sub-process, resolved photoproduction events are also expected to be 'softer' than direct processes at the same y, resulting in a lower transverse energy and a different kinematics

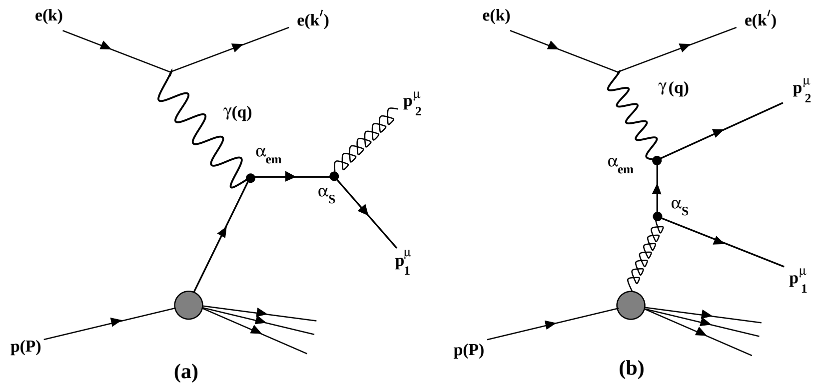

Figure 1.18: Leading order direct processes in photoproduction.

for the outgoing partons. The cross section for photoproduction is sensitive to the hadronic structures of both the proton and the photon. In an analogous way as for the proton, the concept of the parton densities within the photon together with a corresponding factorization scale μ_γ can be introduced. This is illustrated in Fig. 1.21. As in the case of hadron-hadron scattering, the cross-section $\gamma p \to cd$ is factorized into a convolution of a hard sub-process and non-perturbative proton and photon parton densities:

$$d\sigma_{\gamma p\to cd} = \sum_{ab}\int_{x_p} dx_p \int_{x_\gamma} dx_\gamma f_{p\to b}(x_p,\, \mu_p^2) f_{\gamma\to a}(x_\gamma,\, \mu_\gamma^2)\, d\hat{\sigma}_{ab\to cd} \tag{1.21}$$

[12]By calculating the invariant matrix elements for the various two-to-two parton subprocesses, it can be shown that the $gg \to gg$ and $gq \to gq$ subprocesses give the largest contribution to the parton cross section [3].

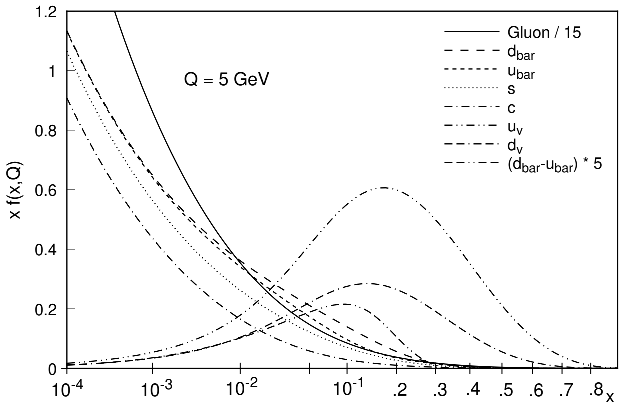

Figure 1.19: Quark and gluon distribution functions in the proton at μ =5 GeV, according to the CTEQ Collaboration (version CTEQ5M). The gluon distribution is scaled down by a factor of 15, and the ($\bar{d}$-$\bar{u}$) distribution is scaled up by a factor of 5.

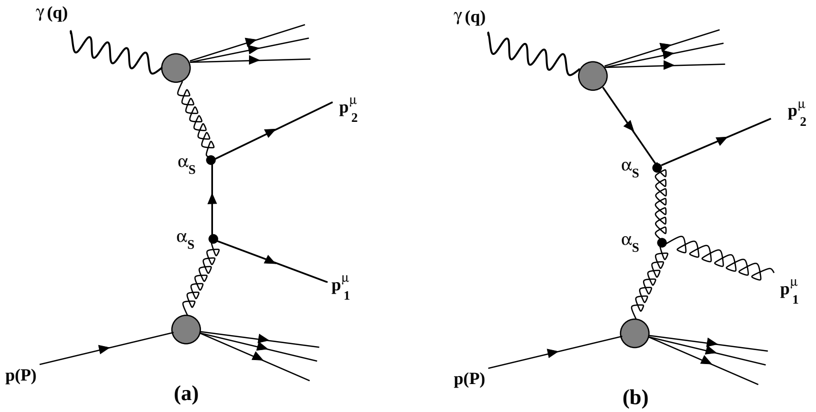

Figure 1.20: Examples of leading order resolved processes in photoproduction.

For the case of direct photoproduction, the parton from the photon is the photon itself and the distribution function is a delta function at $x_\gamma = 1$.

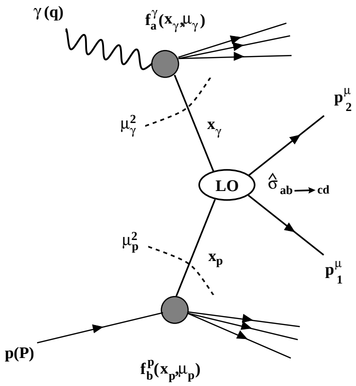

Figure 1.21: General leading order diagram for resolved photoproduction processes.

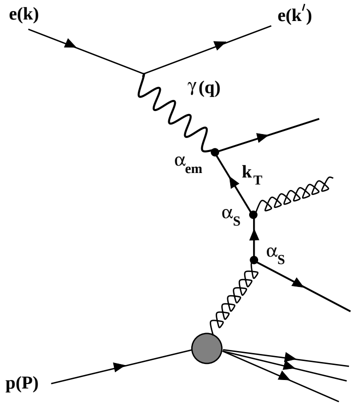

Figure 1.22: A NLO diagram for direct photoproduction process.

Note however that an unambiguous distinction between direct and resolved processes exists only at leading order. At next-to-leading (NLO) order there is an ambiguity in the classification of direct and resolved diagrams and only their sum is physically meaningful. This is illustrated in Fig 1.22 where an NLO direct process is shown. Although the diagram contributes nominally at $\mathcal{O}(\alpha_{\rm em}\alpha_S^2)$, if the virtuality k_T^2 of the propagator labeled in figure is smaller than the factorization scale μ_γ, the diagram could as well be included in the LO resolved contribution.

1.5.3 Measurement Strategy

In this thesis the measurements of the internal jet structure for:

- exclusive dijet events
- identified charm (i.e. quark) initiated jets

in photoproduction are presented. The goal of the analysis is to measure the properties of identified quark jets and by comparing with the inclusive measurement (i.e. mixture of gluon and quark jets) to be able to infer the properties of gluon jets.
The jet shape and the mean subjet multiplicity observables, defined in Sec. 1.4, are measured as a function of the jet transverse energy $E_T^{\rm jet}$ and of its pseudorapidity $\eta^{\rm jet}$ and the results are compared with the available pQCD predictions. Moreover, a comparison of the measurement for quark (charm) initiated jets in photoproduction with similar measurements performed in deep inelastic scattering ($Q^2 \gg 1$ GeV^2) allows to test the universality of the QCD radiation processes. Universality means that the evolution of the outgoing partons, which determines the internal structure of the final state hadronic jets, is independent of the kind of hard sub-process from which the outgoing partons originate.

Chapter 2

The ZEUS Detector at HERA

This chapter gives a brief overview of the HERA accelerator complex and the ZEUS detector. Emphasis is put on those detector components important for the analysis. For a detailed description of the ZEUS detector refer to [18].

2.1 The HERA Collider

Figure 2.1: Aerial view of the DESY site showing the locations of the HERA and PETRA rings and the ZEUS and H1 experiments (marked with S and N respectively).

HERA (Hadron Elektron Ring Anlage) [19] is the first and so far only lepton-proton collider in the world. It is located at the DESY (Deutsches Elektronen SYnchrotron) laboratory in

Hamburg, Germany (see Fig. 2.1) and has been operating since autumn 1991. HERA is designed to collide electrons or positrons, accelerated up to an energy of 30 GeV, with protons of 820 (920) GeV energy, yielding a centre of mass energy of $\sqrt{s} \simeq \sqrt{4E_eE_p} \sim 314$ GeV. Compared to fixed target experiments the centre of mass energy at HERA is an order of magnitude higher and therefore a new kinematic region is accessible[1].

The HERA tunnel is 6.3 km long and it is situated 15-25 m under ground level. It consists of four straight segments, each 360 m long, joined by four arcs with a radius of 779 m. Electrons (positrons) and protons are accelerated in two different rings, using conventional and superconducting magnets, respectively.

Four experiments have been installed along the HERA ring: ZEUS and H1 (located in the South and North Hall respectively, where the beams are brought into collision every 96 ns with zero crossing angle) measure $e^{\pm}p$ interactions. The HERMES experiment (located in the East Hall) uses polarized electrons in collision with an internal polarized gas target (hydrogen, deuterium or He^3) to investigate the nucleon spin structure, whereas the HERA-B experiment (located in the West Hall) was designed to use collisions of the proton beam halo with a wire target to produce B-mesons for the study of $\mathcal{CP}$-violation in the $B\overline{B}$ system. Figure 2.2 shows a

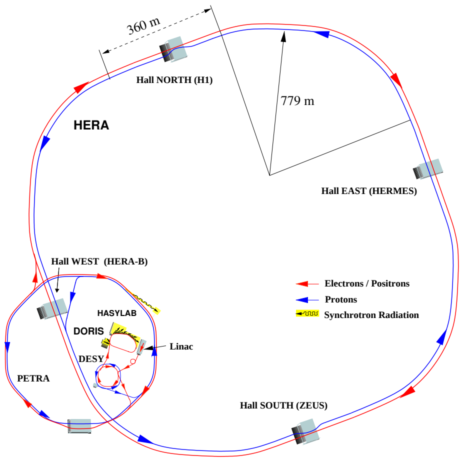

Figure 2.2: The HERA storage ring and its pre-accelerator complex. Four experiments are located in the experimental halls South (ZEUS), West (HERA-B), North (H1), and East (HERMES).

schematic layout of the HERA facility and its pre-accelerator system. The proton acceleration chain starts with negative hydrogen ions (H^-) which are accelerated to 50 MeV in a LINAC. The electrons are then stripped off the H^- ions to yield protons which are injected into the proton synchrotron DESY III where they are accelerated up to 7.5 GeV. They are further accelerated in PETRA up to 40 GeV and then injected into the HERA proton storage ring, where they reach their final energy of 820 GeV. The electron pre-acceleration chain starts in the LINAC II, where the lepton beam is accelerated up to 450 MeV. The electrons are then injected

[1]To reach the same centre of mass energy, in a fixed target experiment, where $\sqrt{s} \simeq \sqrt{2E_em_p}$, one would need to accelerate the incident electron beam up to ~48 TeV.

into DESY II and, once accelerated up to 7.5 GeV, into PETRA II, reaching an energy of 14 GeV. They are transferred then into the HERA lepton ring and further accelerated to their final energy of 27.5 GeV. Electrons and protons are grouped in bunches of $\mathcal{O}(10^{10})$ particles each. 210 bunches of each electrons and protons spaced by 96 ns can be filled into HERA. In normal running, some of the 210 positions are left empty ('pilot bunches') in order to study the background conditions. Non-colliding bunches (where either the lepton or the proton bunch is empty) enable the measurement of beam related backgrounds, whereas empty pilot bunches (where neither of the two is filled) allow the study of cosmic ray background rates. Since

Year	HERA Luminosity (pb^{-1})
e^- 92-94	2.19
e^+ 94-97	70.92
e^- 98-99	25.2
e^+ 99-00	94.95

Table 2.1: Overview of the luminosity delivered by HERA from 1992 to 2000.

HERA parameters	Design	1996/1997	Upgrade
proton beam energy (GeV)	820	820[1]	820/920
electron beam energy (GeV)	30	27.52	30
number of bunches(proton/electron)	210	174+6[2]/174+15[2]	174+6[2]
maximum proton current (mA)	160	76	140
maximum electron current (mA)	60	36	58
proton transverse beam size at IP[3] $\sigma_x \times \sigma_y$ (mm)	0.29×0.07	0.2×0.054	0.118×0.032
electron transverse beam size at IP[3] $\sigma_x \times \sigma_y$ (mm)	0.26×0.02	0.2×0.054	0.118×0.032
proton longitudinal beam size σ_z (mm)	110	170	110
electron longitudinal beam size σ_z (mm)	8.0	8.0	8.0
max. instantaneous luminosity ($cm^{-2}s^{-1}$)	1.5×10^{31}	1.4×10^{31}	7.0×10^{31}

Table 2.2: HERA parameters for the 1996-1997 running period compared to the original design values and to the upgrade ones.

its first year of running, the performance of HERA has steadily improved and the integrated luminosity gathered has continuously increased. In the second half of the 1994 running period, it was decided to operate HERA with positrons instead of electrons in order to achieve a longer and stable lepton beam lifetime (about 8 hours). For electrons above a current of about 20 mA lifetime problems were experienced, which were attributed to capturing of positively charged dust. During the 1997/1998 shutdown period, new pumps were installed in the lepton ring to improve the electron beam lifetime and therefore during 1998 and part of 1999 HERA was operating again with electrons. In 1998 the proton energy was also raised to 920 GeV.
Figure 2.3 shows the luminosity delivered by HERA for the different years versus the days of running. A summary for the data taking periods can be found in table 2.1. With the HERA luminosity upgrade[20], an increase in the luminosity by a factor five is expected (see table 2.2 for the most relevant HERA parameters).

[1]During 1998-2000 running period the proton beam energy was 920 GeV
[2]Non colliding bunches
[3]IP = interaction point

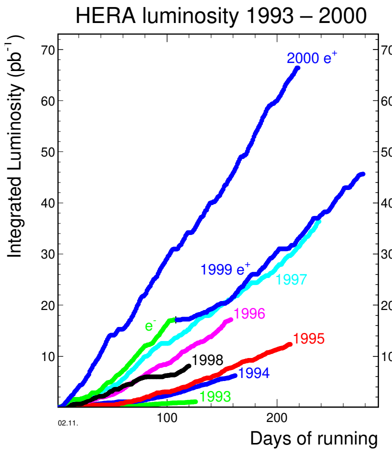

Figure 2.3: Integrated luminosity delivered by HERA in the 1993-2000 running periods.

2.2 The ZEUS Detector

The ZEUS detector is a multipurpose detector designed to study ep scattering at HERA. It covers most of the 4π solid angle, except for small regions around the beampipe. ZEUS was commissioned in spring 1992 and since then it has undergone several detector upgrades (adding essentially new components) driven by the physics and technical understanding gained during the first years of data taking.

The ZEUS coordinate system (see Fig. 2.4) is a righthanded, orthogonal system, with the origin at the nominal interaction point (IP), the Z axis pointing in the proton beam direction (also referred to as the forward direction), the X axis pointing towards the centre of HERA and the Y axis pointing upwards. Figures 2.5 and 2.6 show a longitudinal and transverse

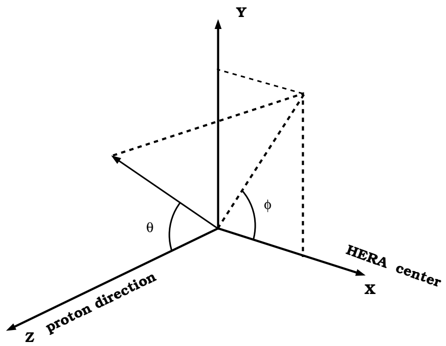

Figure 2.4: The ZEUS coordinate system.

cross sectional view (with respect to the beam direction) of the main components of the ZEUS

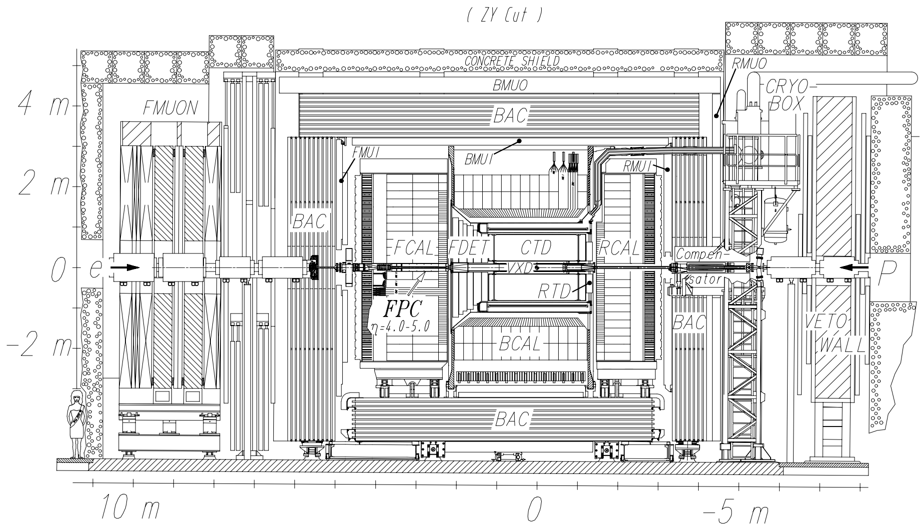

Figure 2.5: Cross section of the ZEUS detector along the beam direction.

detector. Because of the large momentum imbalance between the electron and proton beam, most of the final state particles are boosted to the forward direction and therefore the detector layout is asymmetric with respect to the IP. A brief outline of the major detector components is given below. The principal components used in the analysis will be described in some detail in the following sections. Starting from the interaction point and moving radially outwards, in Fig.2.6 one finds that the vertex detector (VXD) is the innermost component in the ZEUS experiment. However, the VXD was removed during the 1995/1996 shutdown. At present a new silicon microvertex detector (MVD) [21] is under construction and it will be installed in the ZEUS experiment in spring 2001 (see chapter 8). Therefore, in the 1996-2000 configuration, the central tracking detector (CTD) is the nearest component to the IP. It is complemented by forward and rear tracking detectors (FTD, RTD). For charge and momentum determination, the tracking system is surrounded by a super-conducting solenoid providing a magnetic field of 1.43 T.

Outside the superconducting magnet lies the ZEUS calorimeter. It is a compensating high resolution uranium-scintillator calorimeter and it is divided into forward, barrel and rear sections (FCAL, BCAL and RCAL). To improve the discrimination between electromagnetic and hadronic showers for low energy particles (< 5 GeV), silicon diodes have been added in the FCAL and RCAL (hadron-electron separator, HES). The whole uranium calorimeter is enclosed by an iron yoke which provides the return path for the solenoid magnetic field flux and serves as absorber for the backing calorimeter (BAC) which measures energy leakage from the main calorimeter. Limited streamer tube chambers are located inside (FMUI, BMUI, RMUI) and outside (FMUON, BMUON, RMUON) the yoke.

A small angle rear tracking detector (SRTD) is situated between the RTD and the RCAL, covering a radius of ~ 34 cm around the beam pipe. To measure electrons with an even smaller

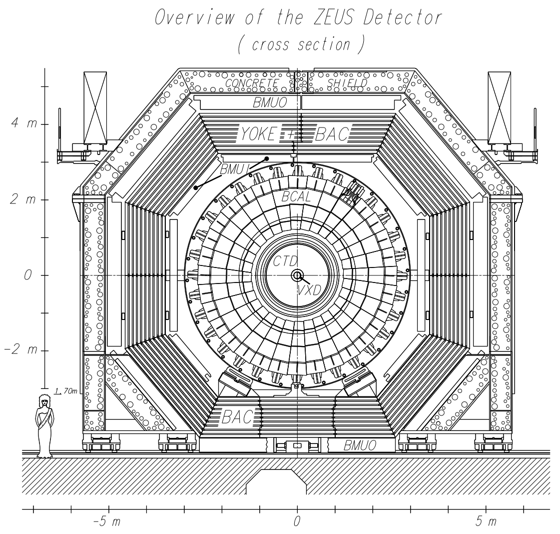

Figure 2.6: Cross section of the ZEUS detector perpendicular to the beam direction.

scattering angle, a small electromagnetic beam pipe calorimeter (BPC) was installed in 1995 in the beam hole of the RCAL. In 1997 the position resolution of the BPC was improved by the installation of a silicon tracker in front of the BPC, the beam pipe tracker (BPT). In 1998 a small forward plug calorimeter (FPC) was installed to extend the calorimetric coverage by one unit in pseudorapidity η[2].

Additional detectors are located outside the main detector along the beam pipe: at distances of about 24 - 90 m from the interaction point, the leading proton spectrometer (LPS) is installed inside the proton beam pipe. It consists of 6 silicon strip detector stations which measure protons scattered at small angles. At Z = 105.6 m, a lead/scintillator calorimeter (FNC), is installed to measure forward neutrons. In the rear direction, at Z = -7.3 m a scintillator hodoscope with iron wall (VETO) is used to reject proton beam-related background. The LUMI detectors, consisting of two small lead/scintillator calorimeters (LUMIE, LUMIG), installed at Z = -34 m and Z = -104 m, detect bremsstrahlung events for the luminosity measurement. The LUMI detectors are also used to identify the scattered electron in photoproduction events and to identify the photon in radiative DIS events. The time interval of 96 ns between bunch crossings in HERA leads to a nominal rate of 10 MHz. In ZEUS a three-level trigger reduces the rate to a few Hz before writing the data to tape.

2.2.1 The Central Tracking Detector (CTD)

The CTD is a cylindrical wire drift chamber which measures the direction and momentum of charged particles and estimates the energy loss dE/dx used for particle identification. The chamber has an overall length of 241 cm and an outer radius of 85 cm. Its active volume ranges

[2]The pseudorapitidy is defined as $\eta = -\ln(\tan\frac{\theta}{2})$

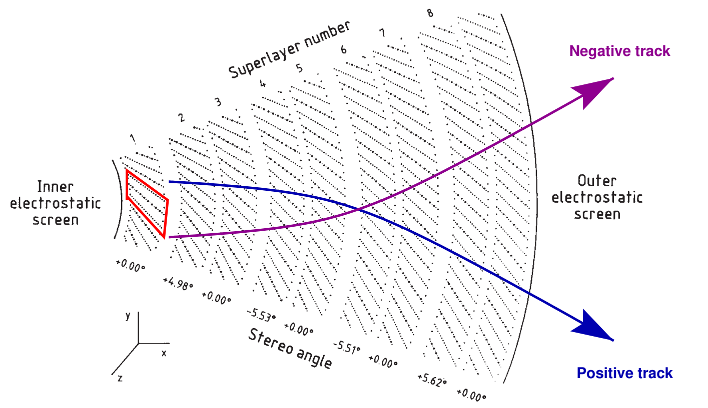

Figure 2.7: X-Y cross section through one octant of the CTD. The larger dots indicate the sense wires. In superlayer 1, a cell (8 sense wires plus associated field-shaping wires) is also indicated.

from Z = -100 cm to Z = 104 cm with outer and inner radii of 18.2 cm and 79.4 cm, resulting in a polar angular coverage of $15° < \theta < 164°$. The CTD consists of 72 radial layers of sense wires which are arranged into nine superlayers (SL). A group of eight wires in the r-ϕ plane of each superlayer defines a cell. One octant of the CTD is shown in figure 2.7. The CTD has a total of 4608 sense wires (30 μm thick) and 19584 field wires. The drift chamber is filled with a gas mixture of argon (Ar), carbon dioxide (CO_2) and ethane (C_2H_6) in the ratio 83:5:12, which is bubbled through ethanol[3]. A charged particle, traversing the CTD, ionizes the gas, creating electron-ion pairs along its trajectory. Under the action of the electric field, the freed electrons drift towards the positive sense wires (with an approximately constant velocity of 50 μm/ns), whereas the positive ions are accelerated towards the negative field wires. In the field of the sense wires, avalanche-like multiplication of the electrons occurs, where the amplification factor is about 10^4. The produced sizable pulse is then read out and digitised by 8-bit flash ADCs. Figure 2.8 shows the ionization path in a CTD cell. The wire orientation is tilted by 45° with respect to the radial direction. The tilt compensates the Lorentz angle of 45° from the combined electric and magnetic field such that the electrons drift azimuthally towards the sense wires, thus ensuring optimal resolution. The position resolution achieved for single hits in (r,ϕ) is approximately 200 μm.
The superlayers are numbered from 1 (innermost) to 9 (outermost). The odd numbered superlayers (SL1, SL3, SL5, SL7, SL9) have wires parallel to the chamber axis and are called axial superlayers. The even numbered superlayers (SL2, SL4, SL6, SL8) are at a small stereo angle ($\pm 5°$) to aid a good Z position reconstruction and are called stereo layers. Combining information from the axial and stereo layers, a resolution in the Z hit position of $\sigma_z \sim 2$ mm can be achieved. The three inner superlayers (SL1, SL3, SL5) are instrumented with a z-by-timing

[3]A pure mixture of argon and ethane (in the ratio 50:50) would have provided a better resolution but, on grounds of safety precautions and detector lifetime, it was decided to avoid it.

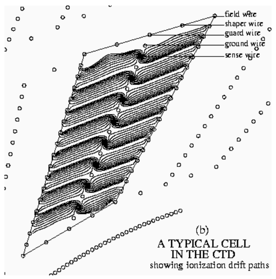

Figure 2.8: Representation of a single cell and its ionization drift paths.

system in which the arrival times (t_1, t_2) of a pulse at the two opposite ends of a wire are measured and the Z position is determined from the time difference $|t_1 - t_2|$. This provides a fast information about the Z position of a track which can be used for trigger purposes. The resolution in Z obtained by the z-by-timing method is $\sigma_z \sim 4$ cm. For tracks fitted to the interaction vertex, passing at least three superlayers and with $p_t > 0.150$ GeV, the resolution in transverse momentum is given by (p_t in GeV)[4]:

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

where the first term corresponds to the hit position resolution whereas the second and third terms to the multiple scattering effects within and before the CTD, respectively.

I. Track Reconstruction

The spatial information necessary for the track reconstruction is obtained by measuring the drift time t_d of the electrons generated in the ionization process. In the case of an almost constant drift velocity, u_d, the relation between t_d and the distance from the sense wire, d_s, is given by:

$$d_s \simeq u_d \cdot t_d \qquad t_d = (t_f - t_i) \tag{2.2}$$

where t_i is the time of passage of the ionizing particle[5] and t_f is the time at which the signal pulse appears at the sense wire. The time information is, however, not enough to determine from which side of the wire the hit originated, since the drift direction is unknown. To solve this left-right ambiguity, the CTD was designed in such a way that each cell is tilted by 45° with respect to the radial direction and therefore straight tracks coming from the interaction point tend to pass the boundary of adjacent cells within a superlayer. The pattern recognition can consequently break the left-right ambiguity by searching for track segments within one superlayer, as illustrated in Fig. 2.9.

In the track finding algorithm each track candidate begins as a 'seed' of 3 hits in an outer axial superlayer which is extrapolated towards the vertex. To facilitate the inward reconstruction of the trajectory, a virtual hit at X=Y=0 is added, which is retained throughout the whole procedure. A circle arc in the (r,ϕ) plane is formed and followed inward, gathering additional

[4]Here $\oplus$ stands for addition in quadrature.
[5]Each sense wire uses its own calibrated t_i.

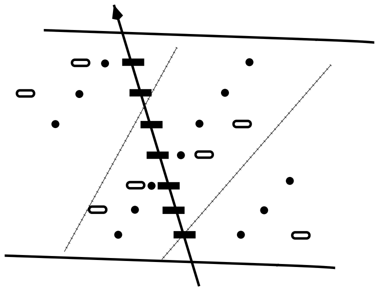

Figure 2.9: Left-right ambiguity resolution: the hits that line up to form a track segment crossing the cell boundaries are genuine hits (solid rectangles), whereas their mirrored counterparts (open ellipses) are ghost hits.

axial hits on the way and updating the trajectory parameters accordingly. Once the trajectory spans several axial superlayers, the arc in the XY plane is used for the stereo pattern recognition, i.e. stereo hits that match with the arc after being rotated are selected. The Z position is given by the amount of rotation needed to superimpose the stereo track segment on the existing track candidate. The pattern recognition algorithm first reconstructs the longest tracks, which are successfully continued from SL9 all the way down to SL1, then shorter ones, whose track seed was found in SL7, and so on down to those which already ended in SL1. Tracks with too many shared hits are removed. Normally, 85% of a candidate's hits must be unique to it[6]. A second pass is made to pick up those tracks which failed to reach down to SL1, but that span at least two axial superlayers and an intermediate stereo one. These tracks may in fact arise from decays of long-lived particles or from secondary interactions.

Each track candidate is then fitted to a 5 parameters helix model, by evolving the trajectory through the magnetic field, starting from the vicinity of the innermost used hit and proceeding outward. The helix parameters, specifying the trajectory at the closest approach to the reference point in the XY plane $(X_{ref}, Y_{ref}) = (0,0)$ (see Fig. 2.10), are:

1. $a_1 = \phi_H$, the angle tangent to the helix in the XY plane
2. $a_2 = Q/R$, Q is the charge and R the local radius
3. $a_3 = QD_H$, D_H connects the helix to the reference point (X_{ref}, Y_{ref})
4. $a_4 = Z_H$, Z position of the point of closest approach
5. $a_5 = \cot\theta$

The trajectory fit can incorporate a 6th parameter, δ_S, which is the projected scattering angle in the XY plane at the component boundaries. When using only hits inside the CTD, as in this analysis, this parameter is set to zero. The coordinates of each point on the helix are a

[6]This rule is somewhat relaxed to enhance the acceptance for photon conversions, when a track candidate spans at least two axial superlayers and shared hits occur only in the innermost superlayer.

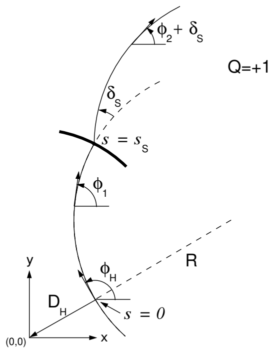

Figure 2.10: Track helix in the (X,Y) plane, without vertex constraint.

function of the azimuthal angle:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}(\phi) = \begin{pmatrix} -(\sin\phi)/a_2 + (1/a_2 + a_3)\sin a_1 \\ (\cos\phi)/a_2 - (1/a_2 + a_3)\cos a_1 \\ a_4 + a_5(a_1 - \phi)/a_2 \end{pmatrix} \tag{2.3}$$

and the momentum components are given by:

$$(p_x, p_y, p_z) = (p\cos\phi\sin\theta, p\sin\phi\sin\theta, p\cos\theta) \tag{2.4}$$

where p is the estimated track momentum from fit.

The reconstructed tracks are then used for the determination of the event vertex, which is done in the following three steps:

1. a preliminary filter removes trajectories whose extrapolation is not close enough to the beam-line
2. a simple linear vertex fit calculates a weighted centre of gravity $(X,Y,Z)^{\mathrm{simple}}$ using the surviving tracks; any tracks which contribute too much to the χ^2_{simple} are discarded until the fit quality is acceptable
3. a non-linear fit determines, starting from $(X,Y,Z)^{\mathrm{simple}}$, the final vertex position $(X,Y,Z)^{\mathrm{final}}$; the participating tracks are then constrained to it and their helix parameters are recalculated.

2.2.2 The Uranium Calorimeter (UCAL)

The ZEUS calorimeter (UCAL) is a high resolution calorimeter. It completely surrounds the solenoid and the tracking detectors, as shown schematically in Fig. 2.11. It is essential for the complete measurement of jets and therefore plays a crucial role in the present analysis.

The UCAL is a sandwich uranium-scintillator calorimeter. It consists of alternating layers of 3.3 mm thick depleted uranium (98.1% U^{238}, 1.7% Nb, 0.2% U^{235}) plates as absorber and 2.6 mm thick organic scintillators (SCSN-38 polystyrene) as active material for readout purposes.

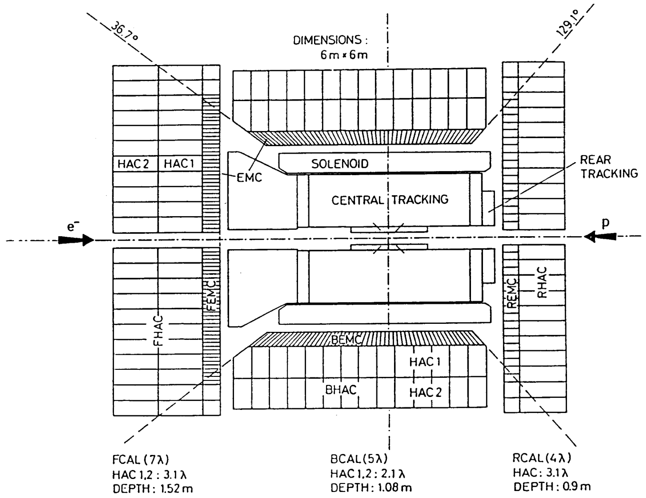

Figure 2.11: View of the UCAL geometry.

The thicknesses of the plates have been chosen such that the calorimeter is compensating, i.e. its response to electrons and hadrons of equal energy is the same (e/h= 1.00±0.02). In this way optimum accuracy for the absolute value and the resolution of hadron energies is achieved.

The energy resolution of the calorimeter, measured under test-beam conditions, is [22]:

$$\frac{\sigma_E}{E} = \frac{18\%}{\sqrt{E/\mathrm{GeV}}} \oplus 2\% \tag{2.5}$$

for electrons and

$$\frac{\sigma_E}{E} = \frac{35\%}{\sqrt{E/\mathrm{GeV}}} \oplus 1\% \tag{2.6}$$

for hadrons.

The UCAL is mechanically divided into three components: the forward (FCAL), barrel (BCAL) and rear (RCAL) calorimeters, covering together 99.7% of the 4π solid angle around the nominal interaction point. The planes of division are at constant Z values and thus non-projective. Hence, there are overlap regions between the FCAL and the BCAL ($36.7° < \theta < 48.2$ °) and between the BCAL and the RCAL ($118.6° < \theta < 129.1°$).

	FCAL	BCAL	RCAL
η-range	3.8 - 1.0	1.1 - -0.74	-0.72 - -4.2
θ-range	2.5° - 39.9°	36.7° - 129.1°	128.1° - 178.4°
EMC radiation length X_0	25.9	22.7	25.9
total absorbtion length λ	7.14	4.92	3.99

Table 2.3: Angular acceptance and longitudinal depth of the UCAL.

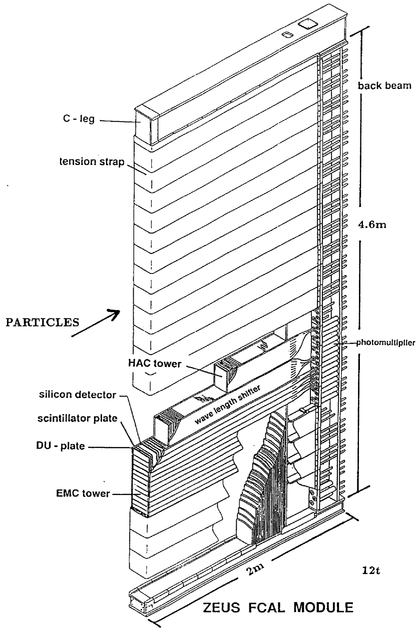

Figure 2.12: Layout of a FCAL module.

Each of the calorimeter parts is subdivided into modules, which in turn are segmented into towers. Each tower is longitudinally divided into an inner, electromagnetic (EMC) and two (one in RCAL) outer hadronic (HAC) sections. The EMC sections consist of four (two in RCAL) cells with transverse dimensions of 5×20 cm^2 (10×20 cm^2 in RCAL), whereas each HAC section consists of one 20×20 cm^2 cell. As an example a module of the FCAL is shown in figure 2.12, where also the readout mechanism is illustrated: the scintillator light of each cell is read out on 2 opposite sides via a coupling to wavelength shifters guiding the signal to photomultiplier tubes (PMTs). Summing the signals of both PMTs, the response becomes independent of the particles point of incidence. Comparison of the two PMTs signals allows the determination of the horizontal impact position.

The calorimeter also provides accurate timing information with a time resolution better than 1 ns for energy deposits larger than 4.5 GeV. This has proven to be a powerful tool to reject beam-gas background efficiently at the trigger level.

The radioactive decay of the U^{238} produces a constant dark current[7] which is used for the channel-by-channel calorimeter calibration and allows to achieve an accuracy of better than 1%. The PMTs and the readout electronics are additionally calibrated using LED, laser and test pulses.

[7]The depleted uranium plates are wrapped in stainless steel foils in order to reduce the signal from the natural radioactivity of the uranium to a level which does not interfere with the measurements of real physics particles but it is still high enough for calibration purposes.

2.2.3 The Luminosity Monitor

The ep luminosity at ZEUS is determined by measuring the bremsstrahlung process [23] ep → eγp. The cross section for this process is high, which is important for a fast and continuous monitoring of the event rate and is known to an accuracy of ~ 0.5 %, thus allowing a very precise determination of the ep luminosity.

The luminosity monitor consists of a photon and an electron calorimeter positioned downstream of the electron beam. Although initially intended as a coincidence measurement, due to the sensitivity of the acceptance of the electron calorimeter on beam position and angle, the bremsstrahlung event rate is currently determined by counting the rate of photons above a certain energy threshold, E_γ^{th}, in the photon calorimeter and the luminosity is defined as:

$$\mathcal{L} = \frac{\dot{N}_{e\gamma p}}{\sigma^{acc}_{e\gamma p}} \tag{2.7}$$

where $\dot{N}_{e\gamma p}$ is the photon rate and $\sigma^{acc}_{e\gamma p}$ is the calculated cross section (corrected for detector acceptance).

Photons, radiated at angles $\theta_\gamma < 0.5$ mrad, exit the beampipe through a Cu-Be window at Z = -92 m and are detected in a lead-scintillator calorimeter positioned at Z = -106 m. Under test beam conditions the energy resolution of the calorimeter was found to be $18\%/\sqrt{E(GeV)}$. However, for protection against synchrotron radiation, the calorimeter has been shielded by a carbon-lead filter, effectively degrading the resolution to $23\%/\sqrt{E(GeV)}$. The impact position of the photons is reconstructed with an accuracy of ~ 2 mm using scintillator fingers embedded in the calorimeter at a depth of 7 X_0. Bremsstrahlung of electrons on the rest gas molecules represents a severe background for the luminosity measurement as its experimental signature is exactly the same as that of ep bremsstrahlung. Its contribution is measured using 'unpaired' electron bunches (i.e. without colliding p-bunches) and is subtracted:

$$\dot{N}_{e\gamma p} = \dot{N}_{tot} - k \cdot \dot{N}_{pilot} \tag{2.8}$$

where $\dot{N}_{e\gamma p}$ and $\dot{N}_{pilot}$ are the total event rate and the one measured in the pilot bunches respectively, and k is the corresponding current ratio. The achieved accuracy on the luminosity measurement, which depends crucially on the quality of the vacuum in the interaction region, is of the order of 1.5-2.0%.

2.2.4 The ZEUS Trigger and Data Acquisition System

The ZEUS trigger system has the task of selecting interesting ep physics events amongst many background events. The total event rate at HERA is dominated by interactions of the proton beam with residual gas in the beam pipe, which is of the order of 10-100 kHz, whereas the rate of ep physics events (if excluding the very low Q^2 regime) is only of the order of a few Hz.

ZEUS uses a three-level trigger system in order to select interesting physics events. A schematic diagram of the trigger and data acquisition chain is shown in figure 2.13.

The first level trigger (FLT), which has to deal with the HERA bunch crossing rate of 10 MHz, is a hardware trigger, designed to reduce the input rate below 1 kHz. Data from every bunch crossing are stored into pipelines. These pipelines are 46 bunch crossing deep and allow for a

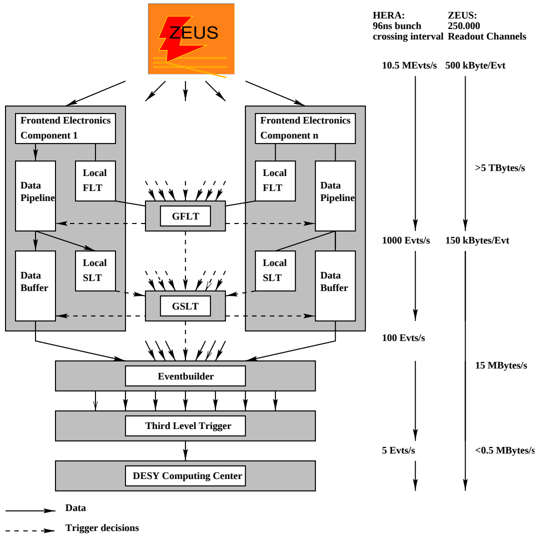

Figure 2.13: Schematic diagram of the ZEUS trigger and data acquisition system.

4.4 μs latency per event. The FLT level operates on a subset of the full data, calculating crude event observables such as regional energy sums, number of tracks and timing information. Each component completes its internal trigger calculations and passes the information for a particular bunch crossing on to the global first level trigger (GFLT) between 1.0 and 2.5 μs after the crossing occurred. The GFLT is issued exactly after 46 bunch crossings ($\sim$4.4μs); if the GFLT decides to keep the event the data is passed on to the second level trigger (SLT).

The SLT is software-based and runs on a network of transputers. It has access to a large fraction of the full data for the event. The SLT is designed to reduce the rate to below 100 Hz. Typically the SLT decision takes 30 μs introducing a $\sim$3% dead time (at a GFLT rate of $\sim$ 1 kHz). At the SLT level, objects like track momenta, the event vertex and calorimeter clusters are reconstructed, permitting a more restrictive trigger decision. As in the case of the FLT, each component has its own local SLT process, passing information to the global SLT (GSLT) which then takes the decision to reject or accept the event. If the GSLT decision is positive, all components send their data to the event builder, which combines their data, writes it into the final data format and makes it accessible to the third level trigger (TLT).

The TLT runs part of the offline reconstruction on a computer farm of Silicon Graphics (SGI) CPUs. On the TLT level, detailed tracking as well as jet and electron finding are performed. After the final TLT decision, the rate is reduced to a few Hz. The accepted events have a typical data size of $\sim$ 100 kB and are written to disk at the DESY computing centre via a fiber-link (FLINK) connection. From here on events are available for the full offline reconstruction and data analysis.

Chapter 3

Physics Simulation

3.1 Introduction

The simulation of physics events in the ZEUS experiment is done in two main steps. In the first step, the ep scattering process is simulated by means of a so called 'event generator' which provides the detailed list of the 4-momenta of all final state particles. In the second step, a simulation of the detector and trigger response to the collection of outgoing particles is performed. The output of this simulation has the same format as the real data recorded by the detector and can therefore be put through the same event reconstruction and physics analysis chain. The basis for these simulations are Monte Carlo (MC) techniques which, due to the complexity of high energy processes and particle detectors, turn out to be an essential tool.

The event generators used for this analysis, HERWIG (version 5.9) [25] and PYTHIA (version 6.1) [30]) will be discussed in the first section whereas the second section shortly describes the detector simulation.

3.2 Event Generators

The main theoretical justification for QCD Monte Carlo event generators lies in the factorization theorem for hard processes [31]. The event generators utilize the fact that a scattering process with a hard scale can be factorized into separate stages (see Fig.3.1):

1. **hard sub-process**: it involves a hard scale (μ) and can therefore be calculated in a fixed order perturbative expansion. The nature of this interaction determines the main characteristics of the event.

2. **initial and final-state radiation**: in every process that contains coloured and/or charged objects in the initial or final state, gluon and/or photon radiation may give large corrections to the overall topology of the event. A very common approach to the modelling of these perturbative corrections is the so called parton shower method, in which the radiation is simulated by an arbitrary number of branchings of one parton into two partons (a $\rightarrow$ bc), i.e. $e \rightarrow e\gamma$, $q \rightarrow qg$, $q \rightarrow q\gamma$, $g \rightarrow gg$, $g \rightarrow q\bar{q}$. The kernel $P_{a\rightarrow bc}(z)$ of a branching gives the probability distribution of the energy sharing, with daughter b taking a fraction z and daughter c the remaining $1-z$ of the initial energy E_a. Once

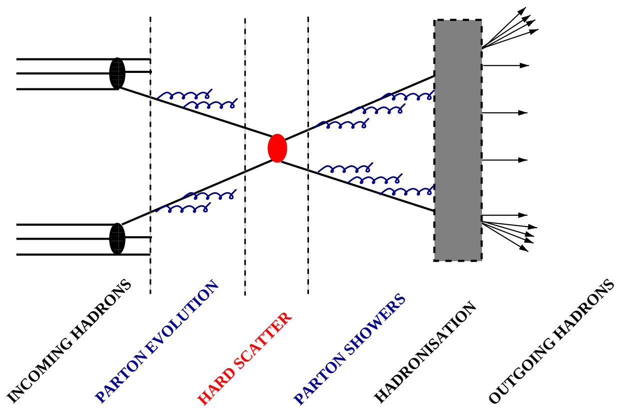

Figure 3.1: Schematic representation of the processes involved in a typical hard QCD interaction.

formed, the daughters b and c may branch in their turn, and so on. The shower evolution is cut off at some scale μ_0, typically around 1 GeV for QCD branchings and around m_e for initial state QED ones.

Note that the initial state radiation leads to the evolution of the parton density functions $f_i(x, \mu)$[1] of the incident particles.

3. **hadronization**: it is the process in which the coloured partons are transformed into colourless hadrons. This non-perturbative process has yet to be understood from first principles and at present only phenomenological models exist. The main current models are the cluster and string hadronization. Their versions, implemented in the HERWIG and PYTHIA event generator programs, respectively, will be briefly discussed in the subsections 3.2.1 and 3.2.2.

4. **beam remnant fragmentation**: in the scattering process, the algorithm of the initial-state radiation reconstructs in each beam particle the shower initiator by backwards evolution from the hard sub-process. This shower initiator takes only some fraction of the total beam energy, leaving behind a beam remnant that takes the rest. If the shower initiator is coloured, so is the remnant. Being colour-connected to the hard interaction, the beam remnant forms part of the same fragmenting system and needs to be reconstructed and connected to the rest of the event. In addition, in collisions where the incoming beam particles have a composite nature (e.g. hadron-hadron interaction, resolved photoproduction) there is the additional possibility that several parton pairs undergo separate hard or semihard scattering, usually referred to as 'multiple interactions' (see Fig.3.2). At present there is very little understanding of these mechanisms and the event generators make use of phenomenological models and parameterization from existing data to simulate the effects on the hadronic final state.

[1] $f_i(x, \mu)$ describes the probability of finding a parton i inside a beam particle, carrying a fraction x of the total momentum, when the beam particle is probed at a given scale μ.

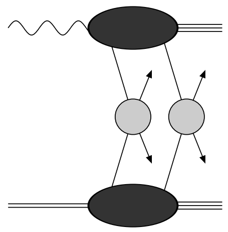

Figure 3.2: Schematic representation of a multiple interaction in a γp collision.

3.2.1 HERWIG Monte Carlo

The philosophy of the HERWIG Monte Carlo [25] is to describe the perturbative stages 1. and 2. of the previous section in as much detail as possible. Therefore meaningful results can only be expected for physical observables which are little affected by the non-perturbative processes. Such observables are then defined in close connection to the topology of the hard subprocess. With regard to the description of the non-perturbative hadronization and beam remnant components simple models are assumed. As a consequence of this approach the HERWIG Monte Carlo contains very few tunable parameters and has a high predictive power.

Spacelike parton showers from the hadronic initial states are generated by backwards evolution from the hard sub-process and matched to the parton density parameterizations for the initial state hadrons at some infrared cut-off. The HERWIG Monte Carlo uses the equivalent photon approximation (EPA) to generate the spectrum of photons radiated from the incoming positron. For resolved photoproduction processes, HERWIG also allows at each branching the possibility for a dynamical $\gamma \to q\bar{q}$ anomalous splitting.

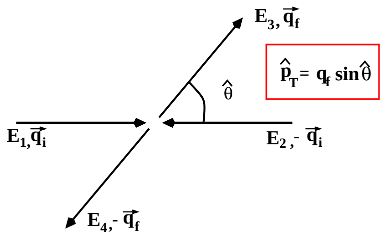

Figure 3.3: Schematic representation of a $2 \to 2$ process in the centre of mass.

The scale for the hard sub-process, μ^2, is given by:

$$\mu^2 = \frac{2\hat{s}\hat{t}\hat{u}}{\hat{s}^2 + \hat{t}^2 + \hat{u}^2} \simeq \hat{p}_T^2 \cdot \frac{4}{3 + \cos^2\hat{\theta}} \tag{3.1}$$

where $\hat{s}$, $\hat{t}$, $\hat{u}$ are the Mandelstam variables and $\hat{p}_T$ is the transverse momentum of the outgoing partons defined in the rest frame of the hard sub-process (see Fig. 3.3). In the assumption of massless partons, the resulting cross-section is divergent in the limit $\hat{p}_T^2 \to 0$, and therefore a minimum transverse momentum $\hat{p}_T^{min}$ is introduced as regularization cut-off.

The outgoing virtual partons (with large time-like mass) generate a shower of partons with decreasing virtuality. In the branching process, the energy fractions are distributed according to the Altarelli-Parisi splitting functions [27] and the available phase space is restricted to an angular-ordered region, i.e. the angle between two emitted partons is smaller than that of the previous branching. Such angular restriction is the result of gluon interference and takes leading infrared singularities correctly into account. The parton emissions continue until the squared transverse momentum of the radiated partons becomes too small and perturbative QCD is no longer applicable (around ~ 1 GeV^2). It is worth noting that in the case of heavy quarks, the collinear divergences are regularized by the non-zero quark mass and that the gluon radiation from a heavy quark is characterized by an angular screening, i.e. the soft collinear gluon emission is reduced within the angle $\Theta_0 = m_q/E_q$, where m_q and E_q are the quark mass and energy, respectively. This screening of the collinear singularity is known as the 'dead-cone' effect [28].

The non-perturbative hadronization in the final state is governed by the phenomenological cluster fragmentation model. First, all the outgoing gluons are split into quark-antiquark or diquark anti-diquark pairs. Then, quarks are combined with their nearest neighbouring (in the colour field) anti-quark or diquark, to form colour singlet clusters. These clusters have a distribution of mass and spatial size peaked at relatively low values, falling rapidly for large cluster masses and sizes and are asymptotically independent of the hard sub-process scale. The so formed clusters are fragmented into hadrons (see Fig. 3.4, left plot). If a cluster is too light to decay into two hadrons, it is allowed to become the lightest hadron of the relevant flavour, adjusting its mass to the appropriate value by momentum exchange with a neighbouring cluster. Massive clusters, below a certain fission threshold M_f, decay isotropically into pairs of hadrons. Unstable hadrons formed in this way are then allowed to decay. Some clusters are too heavy for isotropic two-body decay and therefore are first allowed to fragment into lighter clusters using an iterative fission model, until the masses of the fission products fall below M_f and subsequently decay into hadrons. The fission threshold M_f is defined according to the formula [29]:

$$M_f^{\mathrm{par1}} = (\mathrm{par2})^{\mathrm{par1}} + (m_1 + m_2)^{\mathrm{par1}} \tag{3.2}$$

where m_1 and m_2 are the masses of the quarks forming a cluster, and par1, par2 are input parameters in the Monte Carlo program, tuned in order to reproduce experimental results.

3.2.2 PYTHIA Monte Carlo

The PYTHIA Monte Carlo [30] is also a general purpose event generator, sharing many common features with the HERWIG Monte Carlo, but also with significant differences since its philosophy is rather to describe the hadronization stage in as much detail as possible.

In the case of photoproduction, the Weizsäcker Williams Approximation is used to generate the spectrum of photons radiated from the incoming positron.

The scale for the hard sub-process, μ^2, is given by the transverse mass of the two outgoing partons, m_T^2:

$$\mu^2 = \frac{1}{2}(m_3^2 + p_{T_3}^2 + m_4^2 + p_{T_4}^2) = \frac{1}{2}(m_3^2 + m_4^2) + \frac{\hat{t}\hat{u} - m_3^2 m_4^2}{\hat{s}} \tag{3.3}$$

The parton shower is similar to the one developed in the HERWIG Monte Carlo, but in the case of PYTHIA the parton emissions are ordered only according to the virtuality of the radiated partons. As in HERWIG, the perturbative parton radiation stage is stopped when the squared transverse momentum of the radiation becomes of the order of 1 GeV^2. In the PYTHIA Monte Carlo, the coherence in the final state radiation is incorporated by explicitly disallowing parton emissions with disordered opening angles.

The non-perturbative hadronization in the final state is described by the string fragmentation or Lund model [32]. In this model, as a quark and an anti-quark (q, $\bar{q}$) move apart from their common production vertex, a colour string is stretched between the two partons. As the length of this string increases, the potential energy stored in it linearly increases and the string may break by the production of a new quark anti-quark pair (q', $\bar{q}'$) so that the system splits into two colour-singlet systems, $q\bar{q}'$ and $q'\bar{q}$. If the invariant mass of either of these string pieces is large enough, further break-ups may occur. For a (q, $\bar{q}$, g) configuration, a string is stretched from the quark end via the gluon to the anti-quark, i.e. the gluon being a kink on the string, carrying localized energy and momentum. The string then fragments into two $q\bar{q}$ string segments which are boosted with respect to the overall centre-of-mass frame of the initial (q, $\bar{q}$, g) system. In the case of n partons (after a parton shower), the produced partons are arranged in a planar configuration, where each parton has an equal and opposite colour to that of its neighbour (both neighbours in the case of a gluon). The string is stretched between these colour-connected neighbours, from a quark end via a number of intermediate gluons to an anti-quark end (see Fig 3.4, right plot). At each stage, the string iteratively fragments into smaller segments and the transverse momentum, p_T, for each new $q\bar{q}$ pair created is generated using a Gaussian distribution in p_x and p_y separately[2]. When the energy of individual strings is too small to enable the partons to separate further, hadrons are formed. Their longitudinal momentum is determined with the help of a phenomenological fragmentation function, $f(z)$, which expresses the probability that a given fraction z of the available $E + p_z$ quantity is taken by a hadron:

$$f(z) \propto z^{-1}(1-z)^{\alpha} \exp\left(-\frac{\beta \cdot m_T^2}{z}\right) \tag{3.4}$$

where α and β are free parameters which are determined from fits to experimental data.
For strings where one end consists of a heavy quark, the form of the function $f(z)$ can be modified to account for the harder fragmentation distribution measured in heavy quark (i.e. charm and bottom) production.

3.3 Detector Simulation

The collection of particles produced by an event generator constitutes the input to the ZEUS detector and trigger simulation program MOZART. MOZART is based on the GEANT [33] package. Its kernel is a description of the detector components, including the materials they are made of, their exact geometry and position. The program tracks the particles through the whole detector, taking into account physics processes such as energy loss, multiple scattering and particle decays in flight.

MOZART contains subprograms for the simulation of the trigger (ZGANA), based on simulated signals from the different detector components, and for the offline reconstruction (ZEPHYR). Figure 3.5 shows a schematic diagram of the ZEUS reconstruction scheme for data and MC.

[2] The rms of the p_T distribution is a tunable parameter in the program.

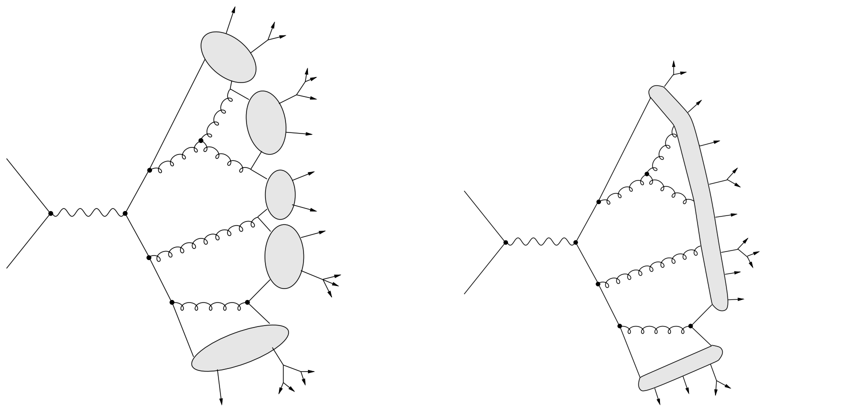

Figure 3.4: Schematic picture showing a parton shower with cluster hadronization model (left) or string hadronization model (right) for $e^+e^- \rightarrow$ hadrons.

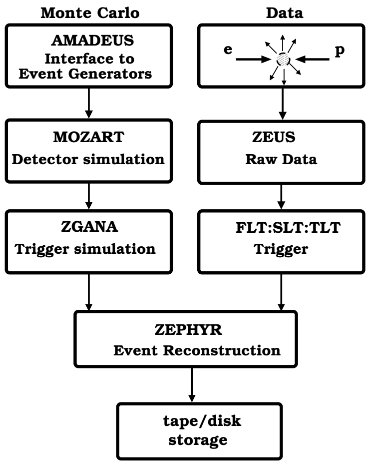

Figure 3.5: Diagram of the ZEUS reconstruction scheme.

Chapter 4

Event Reconstruction and Selection of Dijets Photoproduction Events

In this chapter the selection of the dijet photoproduction sample is discussed. First the reconstruction of the event kinematics is explained. Then the trigger selection and the offline cuts used for this analysis are discussed and the comparison between the selected data and Monte Carlo simulation is shown.

4.1 Reconstruction of the Kinematic Variables

Fig. 4.1 shows a schematic diagram of a hard photoproduction process at HERA. The four

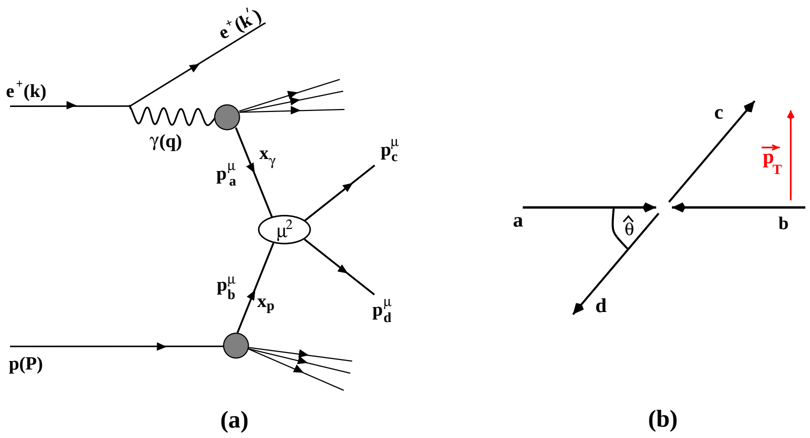

Figure 4.1: Schematic diagram for a photoproduction process.

vectors of the positron in the initial and final state are labelled by k and k', respectively. The negative square of the invariant mass of the virtual photon is denoted by Q^2. The momentum fractions of the partons which participate in the hard scattering are labelled with x_γ and x_p and the four vectors of the incoming and outgoing partons are $(p_a^\mu,\, p_b^\mu)$ and $(p_c^\mu,\, p_d^\mu)$, respectively.

In Fig. 4.1(b) the centre of mass view of the colliding partons is shown. The scattering angle between the two outgoing partons is labelled $\hat{\theta}$ and the relative transverse momentum with respect to the incoming partons is labelled p_T.

4.1.1 Reconstruction of y and Q^2

The negative square of the invariant mass of the photon (defined in Eq. 1.15) is given by:

$$Q^2 = -(k-k')^2 = 2E_e E'_e(1+\cos\theta'_e)$$

where E_e and E'_e are the energy of the incoming and scattered positron, respectively, and θ'_e is the positron scattered angle. In the ZEUS experiment, photoproduction events are defined by the requirement that the scattered positron is not detected in the uranium calorimeter, which corresponds approximately to a scattering angle $\theta'_e > 177°$ and therefore to a cut of $Q^2 \lesssim 3-4$ GeV^2.

The dimensionless variable y defined in Eq. 1.15 as:

$$y = \frac{p\cdot q}{p\cdot k} = \frac{E_p E_\gamma - \vec{p}_p\cdot\vec{p}_\gamma}{E_p E_e - \vec{p}_p\cdot\vec{p}_e}$$

can be determined from the lepton vertex to be:

$$y_{\mathrm{elec}} = 1 - \frac{E'_e}{2E_e}(1-\cos\theta'_e) \tag{4.1}$$

In the photoproduction regime ($Q^2 \to 0$), neglecting the proton mass, y can also be written as:

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

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

$$y \simeq \frac{E_\gamma}{E_e}$$

The reconstruction of the y variable is based on the so-called Jacquet-Blondel method [34]. In this method the variable y is reconstructed as :

$$y = \frac{p\cdot q}{p\cdot k} = \frac{p\cdot(p_X - p)}{p\cdot k} = \frac{E_p E_X - E_p p^z_X}{2E_p E_e} = \frac{E_X - p^z_X}{2E_e} = \frac{\sum_h (E_h - p^z_h)}{2E_e} \tag{4.2}$$

where the momentum conservation at the hadronic vertex has been used:

$$q = p_X - p$$

where $p_X = (E_X, \vec{p}_X)$ denotes the four momentum of the complete hadronic final state X and the sum in the last equation extends ideally over all final state hadrons and in practice over all calorimeter cells. The equation above is exact with the assumption that all hadrons in the final state are measured. Most of the hadrons, which are not detected, have a small p_T and escape

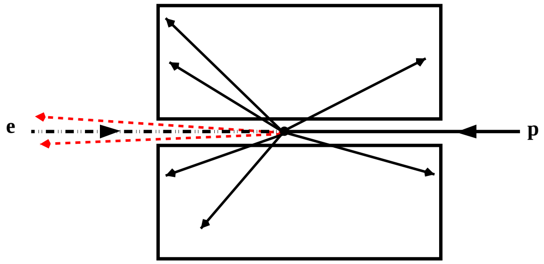

Figure 4.2: Sketch for the understanding of the Jaquet-Blondel method.

in the forward direction (see Fig. 4.2) hence their contribution to the quantity $E_X - p^z_X$ can be considered negligible and one thus obtains a good estimate of y using the formula in Eq. 4.2.

For the reconstruction of Q^2 one can also write:

$$Q^2 = \frac{(E'_e \sin\theta)^2}{1-y} = \frac{(p_T)^2_e}{1-y} = \frac{(p_T)^2_X}{1-y} = \frac{\sum_h (p_T)^2_h}{1-y}$$

where the momentum conservation, $(p_T)_e = (p_T)_X$, had been used.

So, with the Jacquet-Blondel method one obtains:

$$y_{\mathrm{JB}} = \frac{\sum_h (E_h - p^z_h)}{2E_e} \quad \text{and} \quad Q^2_{\mathrm{JB}} = \frac{\sum_h (p^x_h)^2 + (p^y_h)^2}{1-y_{\mathrm{JB}}} \tag{4.3}$$

The γp centre-of-mass energy W which is equal to $\sqrt{ys}$ where s is the ep centre-of-mass energy can also be measured from the hadronic energy deposits in the calorimeter.

4.1.2 Reconstruction of x_γ and x_p

At leading order, the fractional momentum of the incoming partons from the proton and the photon are defined by:

$$x_p \equiv \frac{p_\gamma\cdot p_b}{p_\gamma\cdot p_p} \quad \text{and} \quad x_\gamma \equiv \frac{p_p\cdot p_a}{p_p\cdot p_\gamma} \tag{4.4}$$

In the collinear approximation, one can write:

$$x_p = \frac{E_b}{E_p} \quad \text{and} \quad x_\gamma = \frac{E_a}{E_\gamma}$$

where E_a and E_b are the energies of the partons emitted by the photon and the proton, respectively. The momentum fractions may be determined from the energies and angles of the outgoing partons according to the following formulas:

$$x_p^{\mathrm{LO}} = \frac{\sum_{j=c,d} E^j_T e^{\eta_j}}{2E_p} \quad \text{and} \quad x_\gamma^{\mathrm{LO}} = \frac{\sum_{j=c,d} E^j_T e^{-\eta_j}}{2E_\gamma} \tag{4.5}$$

These relations can be understood by considering that for a massless parton j:

$$E^j_T\, e^{-\eta_j} = E_j - p^z_j \quad \text{and} \quad E^j_T\, e^{\eta_j} = E_j + p^z_j$$

However, partons are not directly measurable, and one needs to translate the formulas in Eq. 4.5 into experimentally measurable quantities, by replacing the sum over partons with a sum over hadronic jets. The variables x_p^{OBS} and x_γ^{OBS} are therefore defined as:

$$x_p^{\mathrm{OBS}} = \frac{\sum_{\mathrm{jets}} E_T^{\mathrm{jet}} e^{\eta^{\mathrm{jet}}}}{2E_p} \quad \text{and} \quad x_\gamma^{\mathrm{OBS}} = \frac{\sum_{\mathrm{jets}} E_T^{\mathrm{jet}} e^{-\eta^{\mathrm{jet}}}}{2yE_e}$$

where the sum runs over the two highest transverse energy jets in the event. Thus x_p^{OBS} and x_γ^{OBS} represent the fraction of the incoming proton and photon energy which participates in the production of the two jets with the highest transverse energy. Figures 4.3 and 4.4 show

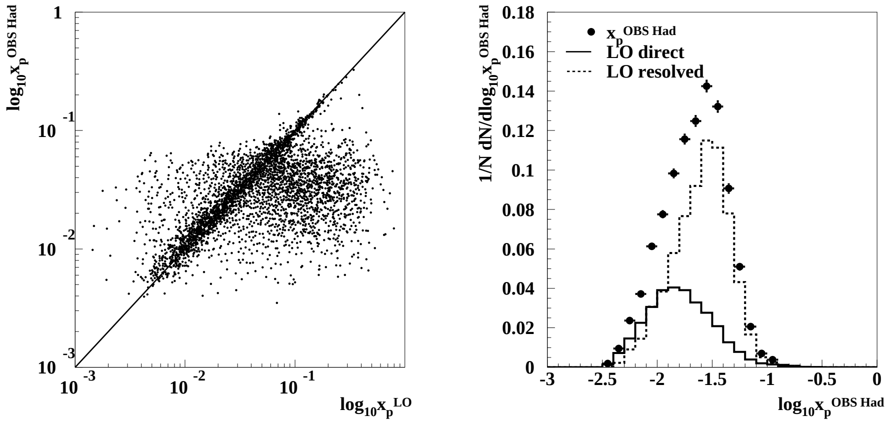

Figure 4.3: Correlation between x_p^{OBS} (at hadron level) and x_p^{LO} and the x_p^{OBS} distribution in PYTHIA.

the correlation between $x_{\gamma,p}^{\mathrm{OBS}}$ calculated at the parton level with the corresponding leading order definition, and the distributions of the $x_{\gamma,p}^{\mathrm{OBS}}$ variables calculated at hadron level. The events are from the PYTHIA Monte Carlo with **Multiple Interaction** (MI) and the jets are found using the KTCLUS algorithm applied to the final state hadrons before the detector simulation stage. From the left plot in Fig. 4.3 one can see that the correlation between x_p^{OBS} and x_p^{LO} is relatively good. The difference between x_p^{OBS} and x_p^{LO} is mostly due to parton shower and hadronization effects. From the right plot in Fig. 4.3 one can see that the direct photoproduction events lie towards smaller values of $\log_{10} x_p^{\mathrm{OBS}}$. This is reasonable since when the maximum photon momentum is involved in the hard scattering, it is possible for partons, carrying smaller fractions of the proton momentum, to participate in the hard interaction.

From the left plot in Fig. 4.4 one can see that the correlation between x_γ^{OBS} and x_γ^{LO} is also relatively good. The direct and resolved processes tend to populate different regions in the x_γ^{OBS} distribution, but a clear overlap between the two contributions is present (as shown in the right plot of Fig. 4.4). The x_γ^{OBS} is also subject to parton shower and hadronization effects which mostly determine the smearing of the direct peak at $x_\gamma^{\mathrm{OBS}} = 1$ to lower values of x_γ^{OBS}.

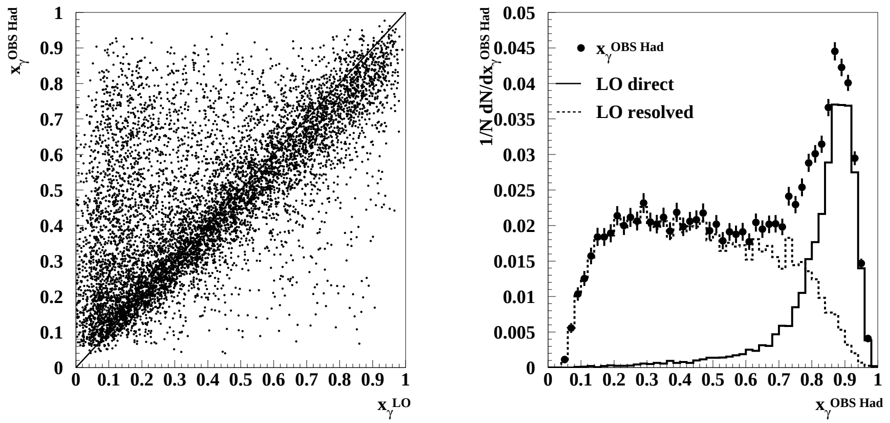

Figure 4.4: Correlation between x_γ^{OBS} (at hadron level) and x_γ^{LO} for resolved photoproduction events and the x_γ^{OBS} distribution in PYTHIA.

4.2 Data Selection

For the analysis presented in this thesis data collected with the ZEUS detector during the 1996-97 and 1999-2000 e^+p HERA running periods is used. The data samples correspond to an integrated luminosity of 106 pb^{-1}. Note that for the exclusive photoproduction dijet analysis without any flavour tagging only the data from 1996-97 (corresponding to a luminosity of 38.4 pb^{-1}) has been used.

The data selection can be divided into two separate steps:

1. online trigger selection
2. offline selection cuts

The online and offline selection criteria are designed to select dijet events and to reject at the same time as many background events as possible. There are two different classes of potential background:

- non ep-physics background processes:
  - proton beam gas interactions occurring upstream in the proton beam
  - positron beam gas interactions occurring upstream in the positron beam
  - beam halo muons travelling collinear to the proton beam
  - cosmic muons

These background processes are rejected efficiently when a reconstructed interaction vertex is required, near the nominal interaction point and when the timing information of the event is required to be consistent with an ep interaction.

- background originating from ep interactions:
  - neutral current deep inelastic scattering (NC DIS)
  - charged current deep inelastic scattering (CC DIS)

The rejection criteria for the ep-physics background will be illustrated in Sec. 4.2.2 where the offline selection cuts are discussed.

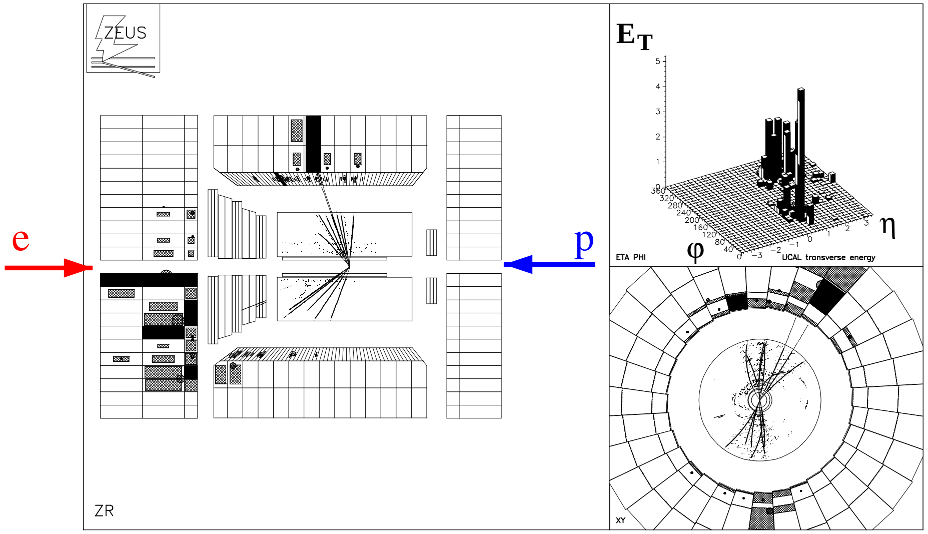

Figure 4.5: A dijet photoproduction event as seen by the ZEUS detector.

Figure 4.5 shows an example of a dijet photoproduction event registered with the ZEUS detector. The energy and position measurements of the uranium calorimeter and the central tracking chamber for the event are displayed in three different views. The picture on the left side shows the rZ view of the event, where a projection of the UCAL and the tracking detectors is shown. The shaded area in each calorimeter cell is proportional to the energy deposited in the cell. One can see large energy deposits in the forward and barrel calorimeter sections which are associated with the jets. There is a well defined vertex defined from tracks close to the nominal interaction point, and a large energy deposit around the beampipe, associated with the proton remnant. In the XY view of the event, shown in the lower right picture, one can see the back-to-back spray of charged particles whose tracks (associated with the two jets) have been reconstructed by the CTD and point to the primary interaction vertex. The upper right picture

shows the transverse energy deposits in the calorimeter, according to their position in the η-ϕ space. One can distinguish two hadronic jets each with $E_{T_{1,2}}^{\rm jet} \sim 19$ GeV and pseudorapidity $\eta_1^{\rm jet} \sim 0.4$, $\eta_2^{\rm jet} \sim 1.4$.

4.2.1 The Online Event Selection

In the online selection of dijet events, based on the ZEUS three-level trigger system, the properties of high transverse energy in the calorimeter and a well defined vertex from tracks reconstructed by the CTD are used.

I. First Level Trigger

The FLT selection criteria are based on global and regional energy sums in the calorimeter[1] together with simple tracking requirements from the CTD FLT and vetoes from additional components. The FLT requires:

- a large energy deposit in the calorimeter:

$$\mathrm{E}^{\rm FLT} \gtrsim 15 \text{ GeV} \quad \text{or} \quad \mathrm{E}^{\rm FLT}_{\rm EMC} \gtrsim 10 \text{ GeV} \quad \text{or} \quad \mathrm{E}^{\rm FLT}_{\rm BEMC} \gtrsim 3.4 \text{ GeV} \quad \text{or} \quad \mathrm{E}^{\rm FLT}_{\rm REMC} \gtrsim 2 \text{ GeV}$$

- at least one good CTD FLT track, where a good track is defined as a track for which the Z position in the first superlayer of the CTD is between -50 cm and 80 cm
- timing information from the vetowall, the C5 counter and the SRTD to veto events having a beam gas timing and no physics timing

II. Second Level Trigger

At the SLT the following cuts are required to select high transverse energy events:

- a reconstructed vertex with $-60 < Z_{\rm vertex} < 60$ cm
- $E - p_z > 8.0$ GeV, where E and p_z are the energy and the longitudinal momentum of the event, determined from the energy deposits in the calorimeter
- $E_T^{\rm cone} > 8.0$ GeV where $E_T^{\rm cone}$ is the sum of transverse energy in all calorimeter cells outside a cone of $10°$ around the forward beampipe
- $E - p_z > 12$ GeV or $p_z/E < 0.95$[2]

Furthermore events are vetoed at the SLT when one of the following conditions holds:

- only one photomultiplier tube in the calorimeter gives a signal

[1] For the determination of calorimeter energy sums at the FLT, 3 inner rings around the FCAL and the inner ring around the RCAL beampipe are excluded. Furthermore, the energy resolution is rather poor, since a coarse digitization of the signals is performed.

[2] Proton beam gas events are enormously boosted in the proton direction and are therefore characterized by $p_z \sim E$.

- $E - p_z > 75$ GeV

Events are also vetoed according to the timing information of different calorimeter sections. The rejection happens when:

- $t_{\mathrm{DOWN}} - t_{\mathrm{UP}} > 10$ ns, where t_{DOWN} and t_{UP} are the timing of the lower and upper half of the BCAL
- $t_{\mathrm{FCAL}} - t_{\mathrm{RCAL}} > 8$ ns
- $|t_{\mathrm{RCAL}}| > 8$ ns
- $|t_{\mathrm{FCAL}}| > 8$ ns

III. Third Level Trigger

At the TLT, more time is available and a software trigger is used. A cone-type jet algorithm (EUCELL) is applied to the energy deposits in the calorimeter and events are accepted when one or more of these algorithm find at least two jets with:

- $\eta^{\mathrm{jet}} < 2.5$
- $E_T^{\mathrm{jet}} > 4.0$ GeV[3]

It is furthermore required that:

- a reconstructed vertex is found with $-60 < Z_{\mathrm{vertex}} < 60$ cm
- the number of bad tracks $N_{\mathrm{tracks}}^{\mathrm{bad}} < 6$, where a bad track is defined as a track that has traversed at least 3 superlayers of the CTD, but does not point towards the interaction vertex

In addition, the characteristic pattern of hit cells for cosmic muons and for muons which travel in the proton beam halo is used at the TLT to reject these backgrounds.

It can be noticed that the TLT trigger used in the analysis is only minimally dependent on the tracking information and is predominantly a calorimeter trigger.

4.2.2 The Offline Event Selection

The event sample obtained with the online trigger selection still contains some contamination from non ep-physics interactions and from non-photoproduction e^+p interactions which must be rejected. Before cuts on quantities based on calorimeter energy values can be made, corrections for noise, originating from the electronics and the radioactive decay of the uranium, are implemented.

[3] In the 1999-2000 trigger configuration, the E_T^{jet} threshold was raised to 4.5 GeV in order to reduce the event rate.

I. Calorimeter noise suppression

The noise per calorimeter cell due to the radioactivity of the uranium has approximately a Gaussian shape, centered around zero, with a standard deviation of ~ 18 MeV in the EMC and ~ 27 MeV in the HAC section. A noise suppression cut of 60 MeV for EMC cells and 110 MeV for HAC cells is therefore applied to each calorimeter cell at the offline level. A cut on the relative cell imbalance, $Imb = (E_l - E_r)/E_{\mathrm{cell}}$ (where E_l, E_r are the signals measured by the two PMTs on the opposite sides of a cell) is also applied:

$$|Imb| > 0.49 + \frac{0.03\mathrm{GeV}}{E_{\mathrm{cell}}} \quad \text{where} \quad E_{\mathrm{cell}} = (E_r + E_l) \tag{4.6}$$

Some calorimeter cells are identified as 'noisy' during certain running periods, i.e. they have a significantly higher mean energy or 'firing' frequency compared to all other cells. A list of noisy cells is made for each year and cells belonging to the list are not included in the offline reconstruction, if their energy is less than three sigma above the noise level:

$$E_{\mathrm{cell}} < \langle E_{\mathrm{cell}} \rangle + 3 \cdot \sigma_{E_{\mathrm{cell}}}$$

To further reduce the noise level a cut of 100 (150) MeV is applied to isolated EMC (HAC) cells. The imbalance and isolation cuts are predominantly used to suppress the contribution of cells characterized by 'sparks' in the PMTs whereas the noisy cells are mainly due to electronic malfunctions.

II. Selection cuts

The following 'cleaning cuts', motivated by comparing data and Monte Carlo, are used to select the final sample of dijet photoproduced events[4]:

1. $-45 \leq Z_{\mathrm{vertex}} \leq 45$ cm, $N_{\mathrm{track}}^{\mathrm{good}} \geq 2$ and $N_{\mathrm{track}}^{\mathrm{bad}} < 5$

   These cuts are used to remove the eventual background due to proton beam gas interactions or cosmic showers still contained in the online selected data sample. The variable $N_{\mathrm{track}}^{\mathrm{good}}$ represents the number of tracks fitted to the primary vertex traversing at least three superlayers and having $p_T^{\mathrm{track}} > 0.12$ GeV. The variable $N_{\mathrm{track}}^{\mathrm{bad}}$ represents instead the number of forward tracks ($5° < \theta_{\mathrm{track}} < 70°$) with $p_T^{\mathrm{track}} > 0.2$ GeV and intercepting the Z axis at a distance from the z_{vertex} larger than 75 cm in the rear direction. The left plot in figure 4.6 shows the distribution of the Z position of the vertex compared to the Monte Carlo, which describes well the data. In the right plots of Fig. 4.6 the distributions for the two variables $N_{\mathrm{track}}^{\mathrm{good}}$ and $N_{\mathrm{track}}^{\mathrm{bad}}$ are presented.

2. $p_T/\sqrt{E_T} \leq 2$ GeV$^{1/2}$

   This cut is applied to remove the small contamination from CC DIS events. The total vectorial transverse momentum has been scaled by the inverse square root of the deposited energy to account for the calorimeter resolution. The distribution of $p_T/\sqrt{E_T}$ is shown in the left plot of Fig. 4.7. One can see that data and Monte Carlo are in good agreement.

[4] For each distribution all selection criteria described in the following are applied, excluding the cut on the observable which is plotted.

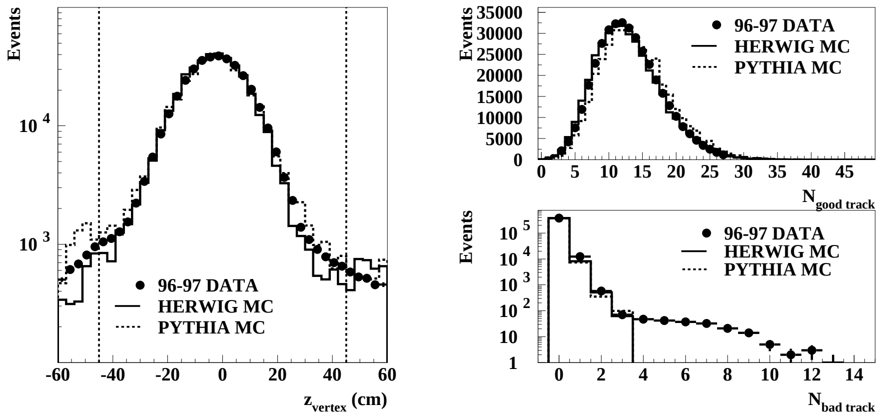

Figure 4.6: Reconstructed Z position of the vertex (left) and number of 'good' and 'bad' tracks (right) for data and for the Monte Carlo models.

3. $0.16 \leq y_{\mathrm{JB}} \leq 0.7$

   The lower cut is used to eliminate possible proton beam gas interactions which are mostly boosted in the proton direction and are therefore characterized by $p_z \sim E$. Thus for proton beam gas background events $y_{\mathrm{JB}} \to 0$. The upper cut is instead used to remove NC DIS background. For NC DIS events the variable y_{JB} peaks around 1, because the energy of the scattered positron is taken as part of the hadronic final state in the calculation of y_{JB}. In the right plot of Fig. 4.7 the y_{JB} distribution is compared to the two Monte Carlo models. One can see that for low values of y_{JB} ($y_{\mathrm{JB}} < 0.35$) the Monte Carlo description is rather poor.

4. the event is rejected if a positron candidate is found by the electron finder SINISTRA [35] with $y_{\mathrm{electron}} \leq 0.7$.

   When in a photoproduction event, an electromagnetic cluster (due, for example, to electromagnetic showers of π^0 and η mesons) is identified in the hadronic final state, the corresponding y_{elec} value is in general higher than that of a genuine NC DIS scattered positron. This can be seen in the right plot of Fig. 4.8 where the y_{elec} distribution is compared to the HERWIG and PYTHIA photoproduction Monte Carlo models and good agreement is found between data and Monte Carlo at higher y_{elec} values. At low values of y_{elec} a deficit in the Monte Carlo is seen, since this region is expected to be populated by NC DIS scattered positrons.

The cuts on y_{JB} and on y_{elec} effectively restricts the range of the virtuality for the exchanged photon to $Q^2 \lesssim 1$ GeV^2, with a median of 10^{-3} GeV^2. This can be seen in Fig. 4.9 where the effect of the two cuts on the Q^2 distribution is shown.

To obtain the final sample of dijet candidate events the k_T cluster algorithm is applied to the calorimeter cells and events are selected in which two calorimeter jets are found with:

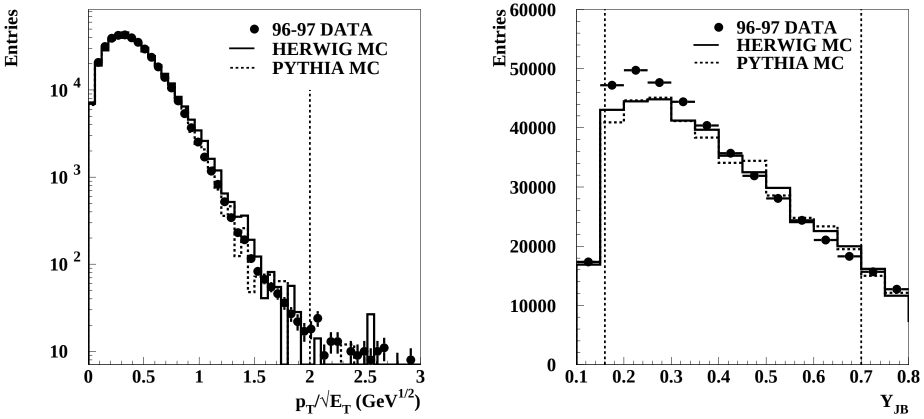

Figure 4.7: Total p_T distribution scaled by $1/\sqrt{E_T}$ (left plot) and the y_{JB} spectrum (right plot) compared with the Monte Carlo predictions.

- $-1 \leq \eta^{\mathrm{jet}} \leq 2$
- $E_T^{\mathrm{jet}_1} \geq 7.0$ GeV and $E_T^{\mathrm{jet}_2} \geq 6.0$ GeV, where $E_T^{\mathrm{jet}_{1,2}}$ is the jet transverse energy corrected for energy losses in inactive material[5] of the leading and second-leading jets

Figure 4.10 shows the x_γ^{OBS} and the x_p^{OBS} distributions (at calorimeter level) for the final sample of dijet candidate events. The data in the x_γ^{OBS} spectrum show two clear peaks: the first one is close to $x_\gamma^{\mathrm{OBS}} \sim 1$ and it is attributed at leading order to a predominance of direct processes, whereas the second one at low x_γ^{OBS} values is attributed to a predominance of resolved processes. The direct component from the Monte Carlo samples is shown separately. A cut at $x_\gamma^{\mathrm{OBS}} \geq 0.75$ is usually used to select a sample of events strongly enriched in direct photoproduction. The resolved component of both Monte Carlo samples includes the simulation of multi-parton interaction (MI). The Monte Carlo prediction without MI underestimates significantly the cross section for $x_\gamma^{\mathrm{OBS}} < 0.3$. Only the inclusion of a model simulating MI effects on the hadronic final state can improve the description of data in analyses where jets are studied starting from low values of the transverse energy E_T^{jet}.

Since the photoproduction Monte Carlo models contain only the leading order contributions to the hard scattering, they underestimate the absolute jet cross sections. To obtain a good description of the data, the resolved and direct contributions in the Monte Carlo models are scaled. The scale factor is determined from a fit to the measured x_γ^{OBS}. For a given bin i in x_γ^{OBS} containing n_i^{dir} entries from the direct processes and n_i^{res} entries from the resolved processes, the combined number of Monte Carlo events in the bin, n_i^{comb} is given by:

$$n_i^{\mathrm{comb}} = \alpha \cdot n_i^{\mathrm{dir}} + (1-\alpha) \cdot n_i^{\mathrm{res}}$$

[5] The correction procedure for the jet transverse energy is explained in App. A.

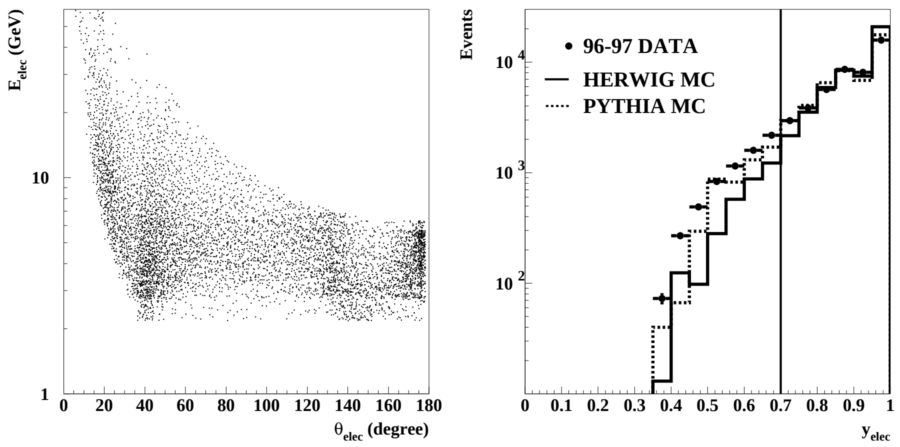

Figure 4.8: Scatter plot of the electron energy, $E_{\rm elec}$, as a function of the electron scattering angle, $\theta_{\rm elec}$, (left plot) and the $y_{\rm elec}$ spectrum compared with the Monte Carlo predictions for the events with a positron candidate in the dijet photoproduction sample (right plot).

with the condition that:

$$n_{\rm MC}^{\rm tot} = \sum_{i=1}^{N_{\rm bins}} n_i^{\rm comb} = N_{\rm data}^{\rm tot}$$

where $N_{\rm data}^{\rm tot}$ is the total number of data events. The parameter α gives the relative proportion between the direct and the resolved components to be combined and it is determined by performing a χ^2 minimization. The value $\alpha = 1$ corresponds to a sample of entirely direct events, whereas the value $\alpha = 0$ corresponds to a sample of entirely resolved events. The result of the fit for the HERWIG and PYTHIA Monte Carlo samples used in this analysis is:

$$\alpha_{\rm HERWIG} = 31.7\% \quad \text{and} \quad \alpha_{\rm PYTHIA} = 30.5\%$$

All the predictions from the HERWIG and PYTHIA Monte Carlo which are shown in this thesis always include this scale factor. Figure 4.11 shows the $x_\gamma^{\rm OBS}$ distribution for different thresholds on the jet with the highest E_T, ranging from $E_T \geq 6$ GeV up to $E_T \geq 29$ GeV. The data are compared to the HERWIG Monte Carlo and the normalization of the direct and resolved components in the Monte Carlo is determined by a fit to each of the $x_\gamma^{\rm OBS}$ spectra. The distributions show that the relative contribution of the resolved processes, as defined in a leading order Monte Carlo model, decreases as the transverse energy of the jets increases.

Figure 4.12 shows the transverse energy and the pseudorapidity distributions for the selected dijet sample. The transverse energy distribution falls rapidly with increasing $E_T^{\rm jet}$ and is well described by the Monte Carlo models which also give a reasonable description of the $\eta^{\rm jet}$ distribution, although neither of the two Monte Carlo models is really able to completely follow the shape of the data distribution.

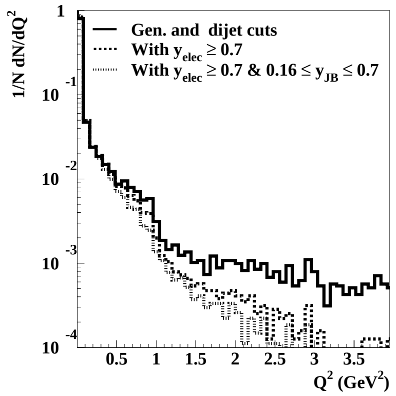

Figure 4.9: Effect of the cuts used to remove NC DIS candidates on the generated Q^2 distribution for the HERWIG Monte Carlo.

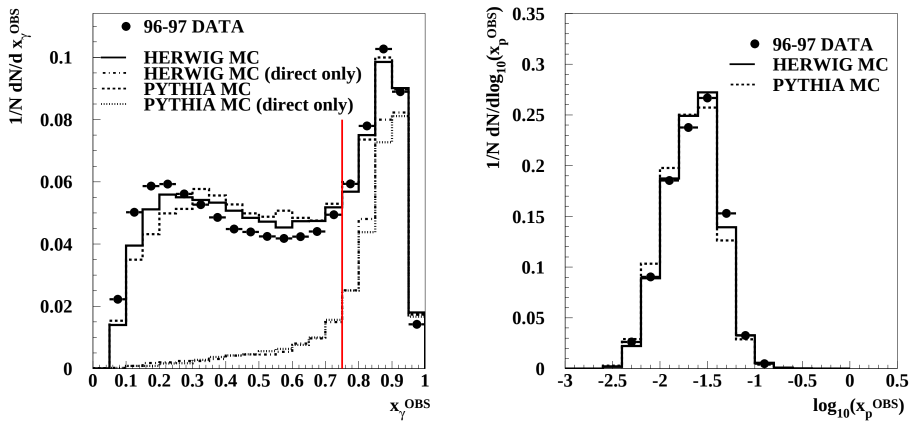

Figure 4.10: The $x_\gamma^{\rm OBS}$ and the $x_p^{\rm OBS}$ spectrum of the selected dijet sample, compared to the HERWIG and PYTHIA Monte Carlo predictions. The direct component in the $x_\gamma^{\rm OBS}$ distribution from the Monte Carlo models is shown separately. Only statistical errors are plotted.

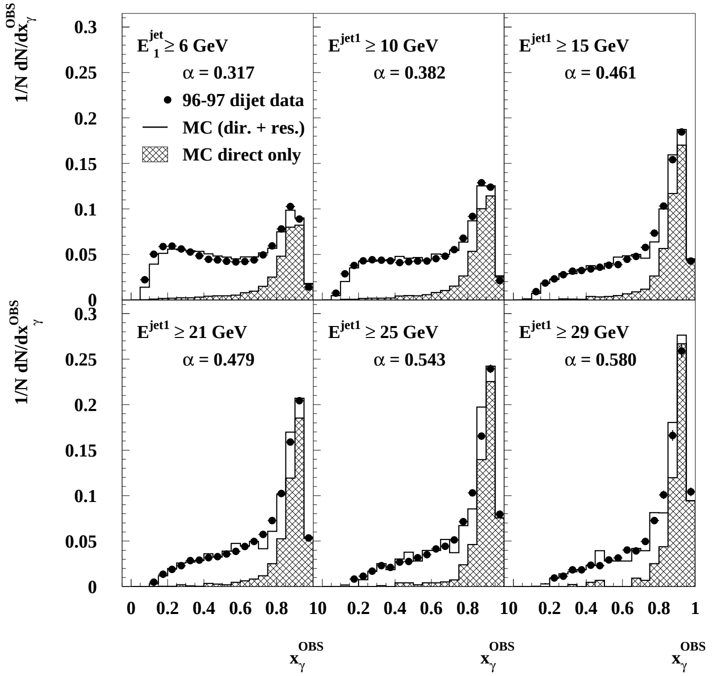

Figure 4.11: x_γ^{OBS} distribution for different thresholds on the highest transverse energy jet. The second jet is always required to have $E_T^{\mathrm{jet2}} \geq 6$ GeV. The open histogram is the prediction of the HERWIG Monte Carlo and the shaded histogram is the direct component of the same Monte Carlo. The data points (dot) are shown with statistical errors only.

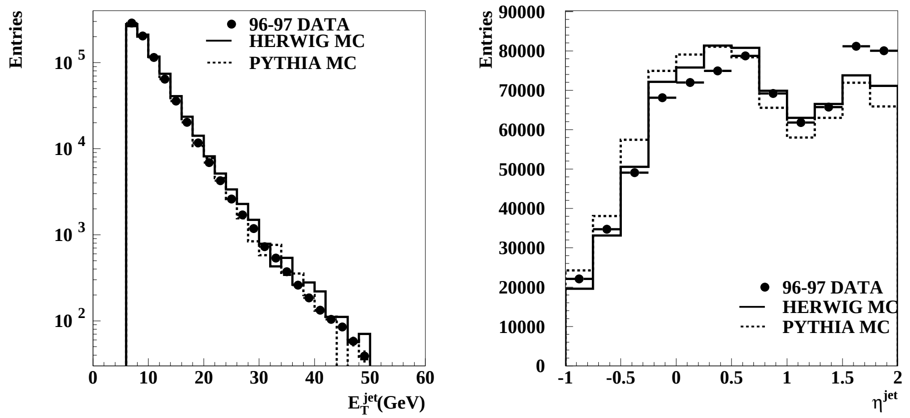

Figure 4.12: The distributions of the transverse energy and pseudorapidity for the dijet sample.

Chapter 5

Charm Selection

In this chapter the selection of charm candidate events from the dijet photoproduction sample is discussed. Sections 5.1 and 5.2 give a general overview about the standard method used for the charm identification at ZEUS, based on the reconstruction of the D^* mesons. In Sec. 5.3 the method to estimate the background contribution to the signal is discussed and the signal distributions are shown. The final section 5.4 explains the tagging of the charm initiated jets.

5.1 Charm Identification

The standard procedure, used in the ZEUS experiment, for the identification of charm production is the measurement of the decay products of charmed hadrons. In the case of open charm production the predominantly produced hadrons are D mesons. Only a small fraction of the produced charm quarks fragment into baryons (cqq), e.g. Λ_c, Σ_c, Ω_c or Ξ_c. To create a D meson, in the fragmentation process the charm quark 'captures' a u or a d quark forming a $D^0(\bar{D}^0)$ or a $D^{\pm}$ either in the ground state or in the short-lived exited states such as D^{*0} or $D^{*\pm}$. A smaller fraction of the charm quarks fragment into a D_s by picking up an s quark. Figure 5.1 shows the mass spectrum of the known D mesons. The heavier states decay immediately (width $\Gamma \sim 0.5 - 25$ MeV), through strong interactions, into lighter D mesons plus another hadron or a photon. The ground state D mesons decay weakly into light particles with lifetime $\tau \sim 10^{-12}$ s. Charmed hadrons which have charged decay products can be observed in the central tracking chamber (CTD). Although providing clean signals, the branching ratios, $\mathcal{B}$s, for specific exclusive decay channels are very small and therefore a high statistics sample (i.e. high luminosity) is needed in order to perform a measurement. The statistics for the charm signal can be improved using inclusive channels such as the semileptonic decay of charm quarks, requiring either a muon or an electron in the final state (see Fig. 5.2). In fact the branching ratio $\mathcal{B}(c \to e\nu X(\mu\nu X)) \sim 10\%$ to be compared, for example, with $\mathcal{B}(c \to D^* \to D^0\pi_s \to K\pi\pi_s) \simeq 0.6\%$ [37]. However, since it is an inclusive channel, the background of electrons or muons that do not stem from semileptonic charm decays is relatively large. Therefore this channel cannot be used for the jet analysis presented in this thesis.

As the charmed particles of interest are selected as 'leading' particles within a jet, the charged hadrons arising from charm decays are to be found within narrow bunches of several other

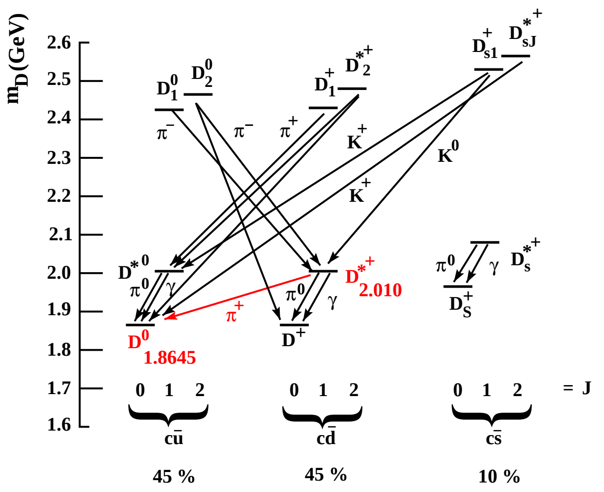

Figure 5.1: Mass spectrum of the D mesons and their transitions. Given are the quark contents, the total angular momentum $\vec{J} = \vec{S} + \vec{l}$ and the mass for the mesons of interest in this analysis.

particles. For this reason, the decay chain[1] (also shown in Fig. 5.3):

$$\begin{aligned} D^{*+} \longrightarrow\ & D^0 + \pi_s^+ \\ & \hookrightarrow K^- + \pi^+ \end{aligned} \tag{5.1}$$

with the lowest multiplicity decay of the D^0 meson into charged particles has been chosen. The branching ratio for this decay chain are [37]:

$$\mathcal{B}(D^* \to D^0\pi_s^+) = 67.7 \pm 0.5\% \quad \text{and} \quad \mathcal{B}(D^0 \to K^-\pi^+) = 3.83 \pm 0.09\%$$

Taking into account the charm fragmentation probability into D^{*+} mesons [38] (obtained combining together four different measurements from Z^0 decays performed in e^+e^- annihilation at LEP [39][2]):

$$f(c \to D^{*+}X) = (23.5 \pm 0.7 \pm 0.7)\,\%$$

the total fraction of charm quarks that follow the described decay chain is:

$$f(c \to D^{*+} \to [D^0 \to K^-\pi^+]\pi_s^+) = 0.62 \pm 0.02\%$$

The mass difference between the D^{*+} and the D^0 mesons is only slightly above the threshold of the pion mass, m_π:

$$m_{D^{*+}} = 2010.0 \pm 0.5\,\text{MeV} \quad \text{and} \quad m_{D^0} = 1864.5 \pm 0.5\,\text{MeV}$$

[1] Here and in the following the charge conjugation will be always implied.

[2] The OPAL and DELPHI collaborations have measured the fragmentation probability times the $D^{*+} \to D^0\pi_s^+$ branching ratio, $f(c \to D^{*+}) \cdot \mathcal{B}(D^{*+} \to D^0\pi^+)$, using a double tag method based on the detection of a slow pion opposite to the full reconstructed D^* meson and the ALEPH and OPAL collaborations have performed $D^{*\pm}$ rate measurements.

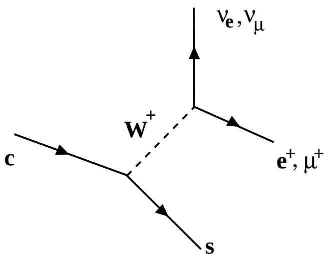

Figure 5.2: Schematic diagram of the semileptonic charm quark decay.

This means that the D^* meson decay products have very little kinetic energy, $Q \equiv m_{D^{*+}} - (m_{D^0} + m_{\pi^+}) = 5.8$ MeV, in the D^{*+} rest frame. Therefore the pion has a very low momentum and is referred to as the 'slow pion' (hence the subscript in the symbol, π_s^+). The tight kinematical constraints of this decay imply that the mass difference $\Delta m \equiv (m_{D^{*+}} - m_{D^0})$ can be measured more accurately than the D^* meson mass itself:

$$\Delta m = 145.436 \pm 0.016 \ \text{MeV}$$

Plotting the Δm distribution yields a prominent signal over a combinatorial background which is heavily phase-space suppressed [41].

In summary, the decay chain $D^{*+} \to D^0(\to K^-\pi^+) + \pi_s^+$ has a small branching ratio but a very clean signal. The D^{*+} meson is reconstructed by measuring tracks which are then associated to the charged decay hadrons, as explained in detail in the next section. Since no particle identification is used, a hypothesis on the mass of each charged track has to be made. This produces a combinatorial background contribution which is however highly suppressed by measuring the mass difference Δm.

The lifetime of a D^0 meson is $\tau = 0.4126 \pm 0.0028 \cdot 10^{-12}$ s, therefore if the D^0 meson is not produced at rest, its decay vertex is displaced from the production vertex by $\beta\gamma c\tau$ with $c\tau \sim 120\mu$m. The resolution of the ZEUS detector is currently not adequate to resolve secondary vertices from decays of heavy quark hadrons and only the installation of the microvertex detector (described in chapter 8) will significantly improve the quality of the charm tagging by allowing the identification of secondary vertices.

5.2 D* Mesons Reconstruction

Since the D^0 meson decay length is, in general, smaller than the vertex resolution of the CTD, the efficiency for tracks from D^* decays to be associated with the vertex is approximately the same as for the other tracks. Therefore all candidate tracks used in the D^* meson reconstruction are required to be associated with the primary vertex. The D^* meson reconstruction is performed by combining all permutations of three tracks, (K, π, π_s) candidates, fulfilling certain conditions which will be detailed in the following. In a first step, pairs of oppositely charged CTD tracks are combined to form a D^0 meson candidate. Each of the two tracks is alternatively assumed to be a kaon or a pion, i.e. with no particle identification being employed. Then an additional slow track with opposite 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 candidate to form a D^* meson candidate.

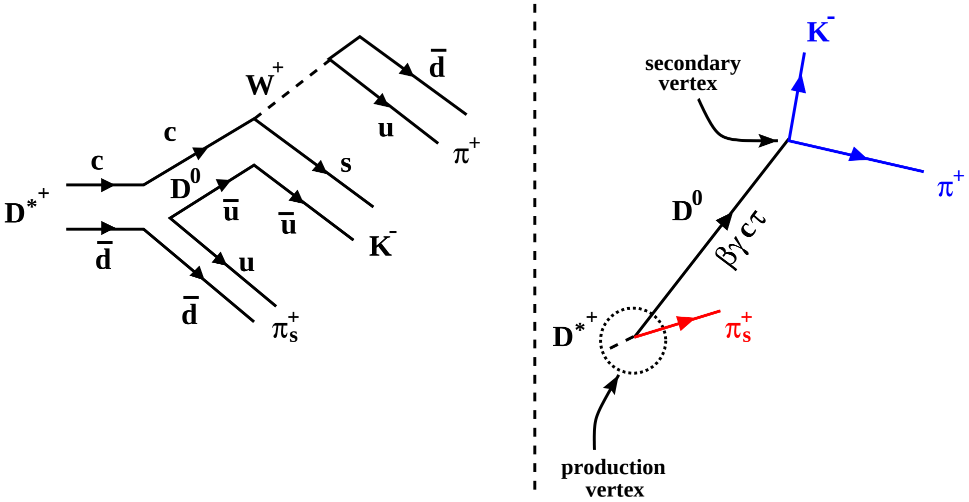

Figure 5.3: Schematic diagram of the D^{*+} meson decay. The left drawing shows the flavour content of the D^{*+} meson decay products, whereas the right drawing is a spatial representation of the decay chain in the laboratory frame.

Only those combinations of three tracks are kept which fall into the D^* meson signal region defined in this analysis as:

$$0.144 \leq \Delta m \leq 0.147 \text{ GeV} \quad \text{and} \quad 1.82 \leq M(K\pi) \leq 1.90 \text{ GeV}$$

The mass window chosen are narrower than the those used in previous ZEUS analyses in both photoproduction and deep-inelastic scattering [42]. The reason for that is related to the background contribution and will be discussed in detail in chapter 6.

The kinematic region of the measured D^* meson candidate is restricted to:

$$p_T(D^*) \geq 3.0\,\text{GeV} \quad \text{and} \quad |\eta(D^*)| \leq 1.5$$

Since the event selection requires two jets with high transverse energy E_T, the tight cut on the transverse momentum of the D^* meson, applied to reduce the combinatorial background, still allows the detection of a significant fraction of D^* candidates.

The tracks used in the D^* meson reconstruction are required to be well contained in the CTD, i.e. only tracks that reach at least the third superlayer (SL3) of the CTD are considered. This implicitly restricts their minimum transverse momentum p_T^{track} and pseudorapidity η^{track}:

$$p_T^{\text{track}} \geq 0.12 \text{ GeV} \quad \text{and} \quad |\eta^{\text{track}}| \leq 1.75$$

Figure 5.4 shows the correlation between the D^* meson transverse momentum and those for its decay products (π_s, K, π) from the HERWIG Monte Carlo. One can see that the D^* meson and the π_s transverse momenta are highly correlated and the transverse momentum threshold for the slow pion track candidate limits the acceptance for low-p_T D^* mesons. A cut of

$$p_T(\pi_s) \geq 0.2 \text{ GeV}$$

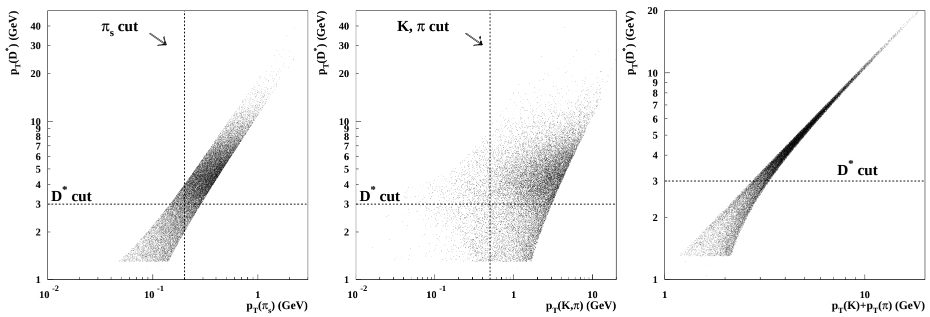

Figure 5.4: Correlations between the D^*, π_s, π and K meson transverse momenta from the HERWIG Monte Carlo.

is sufficient to ensure a high acceptance in the kinematic region chosen for the D^* mesons. Similarly the transverse momenta of the D^* meson and the D^0 meson are also strongly correlated. However, since the masses of the pion and kaon are both much smaller than the D^0 mass, the momenta of the kaon and the pion from the two body decay of the D^0 are individually not so strongly correlated with that of the D^* meson when boosted in the laboratory frame. This is illustrated in the middle plot of Fig. 5.4. Each of the transverse momenta ($p_T(\pi)$, $p_T(K)$) is free to extend to very small values, with the only constraint that:

$$p_T(K) + p_T(\pi) \sim p_T(D^*)$$

in the laboratory frame (see right plot in Fig. 5.4). Figure 5.5 shows the p_T distribution of tracks in the CTD, where generally good agreement between data and Monte Carlo can be seen. The MC spectra are however slightly harder than the data one. The p_T distribution of tracks in the CTD rises steeply with falling p_T and therefore one expects a high multiplicity of (π, K, π_s) candidates for low p_T values and consequently a much higher combinatorial background for the D^* reconstruction. To keep the background at a reasonable level, the following tighter cut for the kaon and pion track candidates is applied in the analysis:

$$p_T(K,\ \pi) \geq 0.5 \text{ GeV}$$

5.3 The D* Signal and the Background Subtraction

The sample of D^* meson candidates obtained using the reconstruction method described in the previous section, is not a pure sample of D^* mesons. A significant amount of combinatorial background is still present and has to be estimated and statistically subtracted. For the estimation of the combinatorial background contribution associated with the signal, two different methods have been employed, the so-called 'wrong charge' (WC) and 'control region' (CR) combinations, which provide independent measurements. In the 'wrong-charge' method, track combinations, which have no possibility of stemming from a genuine D^* meson, are selected in the signal region. This is achieved by requiring that the D^0 meson candidate is composed of like-charge tracks. The slow pion has instead the right charge sign, i.e. the same as the D^*

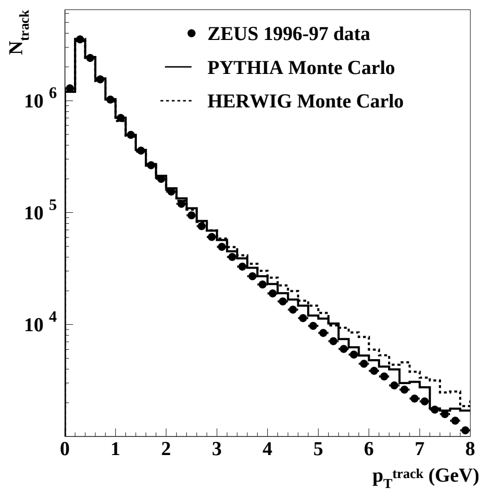

Figure 5.5: Transverse momentum distribution of CTD tracks, assigned to the primary vertex and traversing at least three superlayers, for dijet events in photoproduction.

meson candidate: (K^-, π^-, π_s^+) for the D^{*+} and (K^+, π^+, π_s^-) for the D^{*-}. The phase-space of the 'wrong-charge' combinations is only slightly smaller than that for the right-charge combinations. In the correct charge case, there are N^+N^- possible combinations for the production of a D^0 meson whereas in the wrong charge case, these are $N^+(N^+ - 1)$.

In addition to the wrong charge method, one can estimate the background combinations by means of a control region. This means that events which lie outside, but close to the signal region, in the $(\Delta m,\ M(D^0))$ mass window, are taken. In this case the combinatorial phase-space is approximately equal, but the momentum phase-space is no longer the same. When using the control region, a much larger sample of combinations is available for the estimation of the background contribution, thus allowing the statistical uncertainty in the measurements to be significantly reduced. Moreover, in the present analysis the eventual differences caused by the different momentum phase-space do not constitute any bias for the jet distributions, as will be shown in detail in chapter 6. Therefore the control region background with the highest statistics will be used to estimate the background for all the measurements of the charm jets sample. The wrong charge background will be used instead for the estimation of the systematic uncertainty. Figure 5.6 shows the $(\Delta m,\ M(D^0))$ mass window with the different regions that have been used for the background estimation.

The mass difference Δm and the $M(K\pi)$ invariant mass distributions (from the dijet photoproduction sample, whose selection has been explained in chapter 4) after all selection cuts described above are shown in Fig. 5.7. Clear signals can be seen. The width of the Δm and $M(K\pi)$ peaks can be attributed to the experimental resolution and their Gaussian shape is reasonably well reproduced by the Monte Carlo simulation. Since a combinatorial background also passes the applied cuts, the overall shape is described by a Gaussian distribution for the

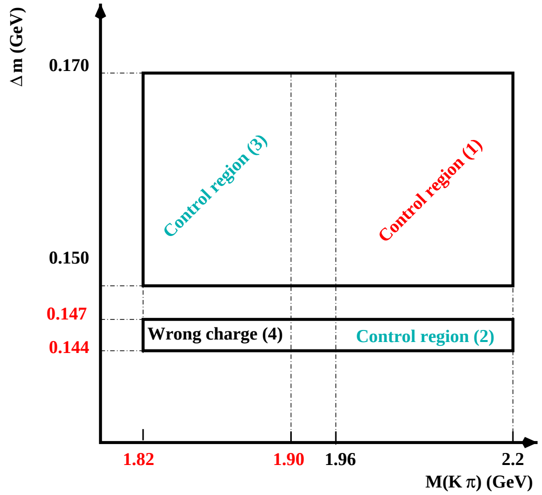

Figure 5.6: $(\Delta m,\ M(D^0))$ mass window with signal and different background regions.

signal superimposed on a background of the form:

$$dN/d\Delta m = a(\Delta m - m_\pi)^b$$

for the Δm distribution, where the parameters a and b are to be determined from a fit and m_π is the pion mass, and of the form:

$$dN/dM(K\pi) = Ae^{-BM(K\pi)}$$

for the $M(K\pi)$ invariant mass distribution, where again the parameters A and B are determined from a fit. The number of D^* mesons reconstructed in the signal region, $N_{\mathrm{D}^*_{\mathrm{rec}}}$, can be obtained from the fit to the Δm distribution or by subtracting the background estimated with either the control region or the wrong charge combinations[3]. The results obtained for $N_{\mathrm{D}^*\mathrm{rec}}$ in the exclusive dijet event sample are:

$$N^{\mathrm{fit}}_{\mathrm{D}^*_{\mathrm{rec}}} = 941 \pm 41\ , \quad N^{\mathrm{CR1}}_{\mathrm{D}^*_{\mathrm{rec}}} = 925 \pm 40\ , \quad N^{\mathrm{WC}}_{\mathrm{D}^*_{\mathrm{rec}}} = 943 \pm 40$$

The fit to the Δm distribution already provides the necessary information, i.e. $N_{\mathrm{D}^*_{\mathrm{rec}}}$, and the eventual fit to the $M(K\pi)$ invariant mass distribution is only used as a cross-check. The value for the D^0 mass obtained from the fit to the $M(K\pi)$ distribution is:

$$M(D^0) = 1861.5 \pm 1.1 \text{ MeV}$$

which is in good agreement with the Particle Data Group (PDG) value. Note, that in addition to the background from random combinations, the $M(K\pi)$ distribution contains also signals from decay modes other than $D^0 \to K\pi$, which are only partially reconstructed and therefore yield a lower mass value $M(K\pi) < M(D^0)$ by at least the π^0 mass. This explains the excess of right charge combinations over the wrong charge (control region) ones in the region $M(K\pi) < 1.7$ GeV.

[3] The wrong charge and control region distributions are normalized to the right charge distribution in a region where no signal is expected: $0.150 < \Delta m < 0.170$ GeV for the Δm distribution and $2.0 < M(K\pi) < 2.2$ for the $M(K\pi)$ distribution.

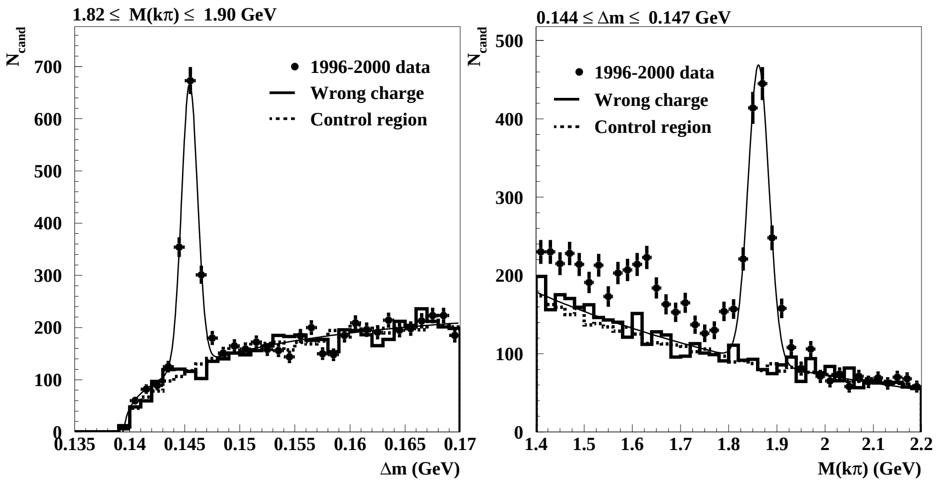

Figure 5.7: Δm (left) and $M(K\pi)$ (right) distributions of the measured D^* candidates (dots) after all selection cuts. Each histogram contains the combinations that fall into the signal region of the distribution shown in the other histogram. The curves show the fits to the data points, as described in the text. The background contribution estimated using the wrong charge (solid line) and control region (dashed line) combinations is also shown.

5.4 Tagging of Charm Initiated Jets

From the final sample of dijet events, selected as explained in the previous chapter, a subsample is extracted, consisting of all dijet events where a D^* meson has been reconstructed, using the method described in the previous sections. A matching criteria is applied in order to decide

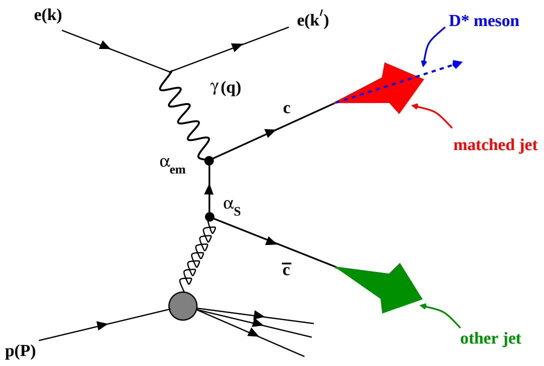

Figure 5.8: Schematic diagram of a direct charm photoproduction process.

which of the two high energy jets contains the D^* meson:

$$\Delta\phi \equiv |\phi_{D^*} - \phi_{\text{jet}}| \leq \phi_{\text{cut}} \quad \text{with} \quad \phi_{\text{cut}} = 45^\circ$$

where ϕ_{D^*} and ϕ_{jet} are the azimuthal angle of the reconstructed D^* meson and of the jet, respectively. The matching criteria and consequently the cut value ϕ_{cut} has been decided by looking at the $\Delta\phi$ distribution for the HERWIG and PYTHIA Monte Carlo samples [4] at both hadron and detector level. Figure 5.9 shows the $\Delta\phi$ and the Δr (i.e. the distance in the (ϕ, η) plane between the D^* meson and the jet, $\Delta r \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$) distributions for the D^*-tagged dijet events in Monte Carlo. From the left plot one can see that in almost all events the D^* meson has a very small azimuthal angle separation from one jet and the other jet is very often contained in the opposite azimuthal hemisphere. In the Δr distribution (right plot in Fig. 5.9) one can distinguish two peaks for $\Delta r \sim 0$ and $\Delta r \sim \pi$. The D^* meson is nearly always very close to one of the two jets in the event. In Fig. 5.10, one can see that after the matching cut has been applied, only a very small fraction of all events has the second jet (other jet) in the same azimuthal hemisphere containing the D^* meson and that only for a negligible fraction of all the events $|\phi_{D^*} - \phi_{\text{jet out}}| < 45^\circ$. Since the η^{jet} distribution for the total dijet photoproduction sample is not perfectly described by the Monte Carlo simulation (see Fig. 4.12), the matching criteria is based only on the azimuthal distance and not on the distance in the (ϕ, η) plane and the few events where both jets have $\Delta\phi < 45^\circ$ are rejected.

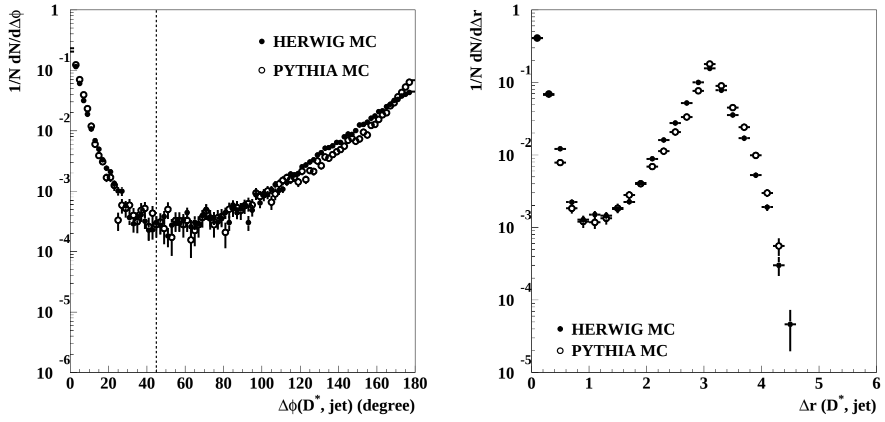

Figure 5.9: $\Delta\phi$ (left) and Δr (right) distributions for the D^*-tagged dijet events in Monte Carlo. The dashed line in the left plot indicates the value used for ϕ_{cut} in the analysis.

The jet not associated to the charm meson ('other' jet) in each event represents the unbiased, i.e. not influenced by the selection cuts on the D^* meson, jet candidate whose internal properties are investigated in this analysis (see Fig. 5.8). Figure 5.11 shows the $p_T(D^*)$ and the $\eta(D^*)$ distributions for all the D^* meson candidates after background subtraction, whereas in Fig. 5.12 the E_T^{jet} and the η^{jet} distributions for the jets in the D^*-tagged sample are shown. One can see that the agreement between data and Monte Carlo predictions is very good. The left

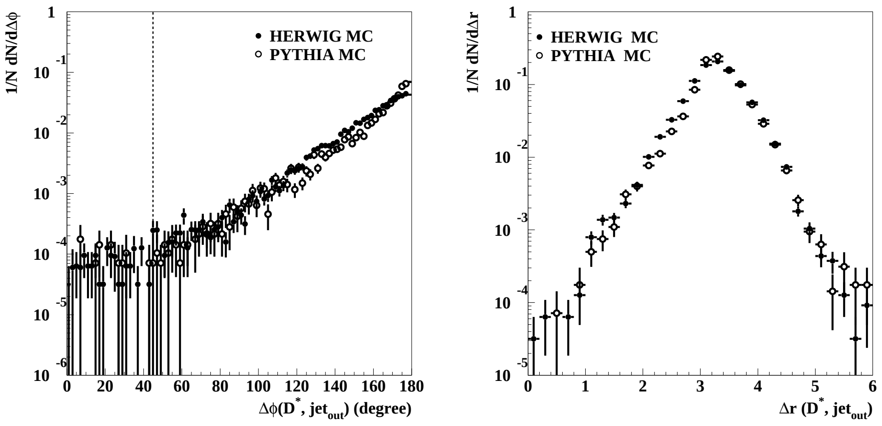

Figure 5.10: $\Delta\phi$ and Δr distributions for the other jet, after the matching criteria has been applied. The distance in the (η, ϕ) plane (right plot) is very large, $\Delta r \gtrsim 1$.

plot in Fig. 5.13 shows the $x_\gamma^{\rm OBS}$ distribution for the D^*-tagged dijet events selected in this analysis. As for the total sample of dijet events, the relative proportion between direct and resolved components of the D^* Monte Carlo samples has been determined from a fit to the $x_\gamma^{\rm OBS}$ distribution. Since variations in the different background distributions result in slightly different $x_\gamma^{\rm OBS}$ spectra for the signal (as shown in the right plot of Fig. 5.13), the result of the fit for the D^* HERWIG and PYTHIA Monte Carlo samples may depend in this case on the region in the $(\Delta m, M(D^0))$ mass window taken to estimate the background contribution. The results of the fit for the scale factor α are reported, separately for the 1996-1997 and 1999-2000 data samples, in table 5.1. One can notice that, for the 1996-1997 data sample, the value of α

	$\alpha_{\rm fit}$ (%)			
	D^* **HERWIG MC**		D^* **PYTHIA MC**	
	1996-1997	**1999-2000**	**1996-1997**	**1999-2000**
Data corr. with CR1	$75\,^{+6}_{-7}$	$70\,^{+5}_{-5}$	$65\,^{+8}_{-8}$	$64\,^{+6}_{-6}$
Data corr. with CR2	$71\,^{+7}_{-8}$	$68\,^{+5}_{-5}$	$60\,^{+8}_{-7}$	$62\,^{+6}_{-6}$
Data corr. with CR3	$75\,^{+7}_{-8}$	$71\,^{+5}_{-5}$	$65\,^{+8}_{-8}$	$65\,^{+6}_{-6}$
Data corr. with WC	$64\,^{+7}_{-7}$	$69\,^{+5}_{-6}$	$52\,^{+8}_{-7}$	$63\,^{+6}_{-6}$

Table 5.1: Values for the parameter α obtained using different background estimations.

changes by about 15-20 %, for both Monte Carlo samples, when using different estimation for the background contribution to the D^* signal. The scale factor α when the data distribution is corrected with the wrong-charge background (WC) is systematically lower than in all other cases. However, since for the 1999-2000 data sample, which corresponds to ~ 65 % of the total luminosity used, the results from the fit using the different background samples are consistent

[4] Monte Carlo samples for dedicated studies of D^* mesons have been generated using a filter which requires a true D^* meson in the event but no cuts on jets are made.

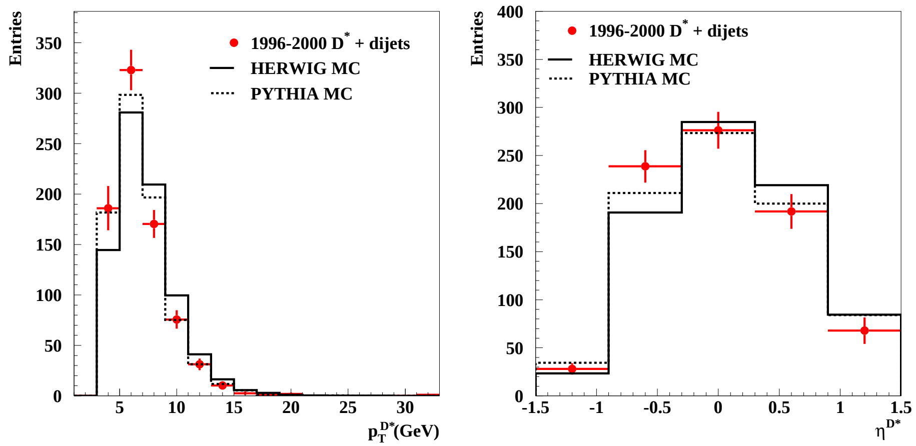

Figure 5.11: $p_T^{D^*}$ and η^{D^*} distributions for the D^* candidates after background subtraction.

with each other, the discrepancy can be attributed to a statistical fluctuation. Thus the result for the scale factor α obtained when correcting the $x_\gamma^{\rm OBS}$ distribution for 1996-1997 data sample with the wrong-charge background will not be used in the following. The value of α turns out to be slightly different for the two Monte Carlo models, whereas for the dijet samples the agreement is very good. This effect is included in the estimation of the systematic uncertainties of the finals results (see chapter 6). Independent of the particular value chosen for α, one can notice that the contribution of the resolved component is not negligible. According to the Monte Carlo predictions, a fraction between $\sim 25-40\%$ of the total D^*-tagged dijet events originates from resolved processes where the photon acts as a source of partons. One possibility of resolved charm production is the gluon-gluon fusion (shown in the left diagram of Fig. 5.15) where one gluon comes from the proton and the other from the photon. This process is very similar to the direct BGF process, producing two high $E_T^{\rm jet}$ jets each 'containing' a charmed hadron, with the difference that in this case only a fraction of the photon's momentum enters into the jet production, the rest being carried off in a photon remnant. However, if a charm contribution is included in the evolution of the parton density functions of either the proton or the photon, one can obtain another class of resolved processes, e.g. $c^\gamma g^p \to cg$ (shown in the right diagram of Fig. 5.15), which are referred to as 'charm excitation'. These processes also lead to high $E_T^{\rm jet}$ jets in the final state, but only one of them contains a charmed hadron, whereas the second charmed hadron is carried off in the photon (or proton) remnant. The measurement of the cross section $d\sigma/dx_\gamma^{\rm OBS}$ for $ep \to D^*X$ + dijets [36] indicated that the contribution at low $x_\gamma^{\rm OBS}$ is significant and both direct and resolved processes are needed to obtain a good description of the charm production data. Moreover, according to the HERWIG LO Monte Carlo the resolved contribution is almost entirely due to charm excitation in the photon. However, the resolved contribution is suppressed relative to the direct one in comparison to the non charm-tagged case (compare Fig. 5.13 with the left plot in Fig. 4.10). In the kinematic region selected in this analysis, the resolved contribution is additionally reduced if compared with the measurement presented in [36], but, as already pointed out, it is not negligible. This means that the other jet in a relatively small fraction of the D^*-tagged dijet events is expected to originate from

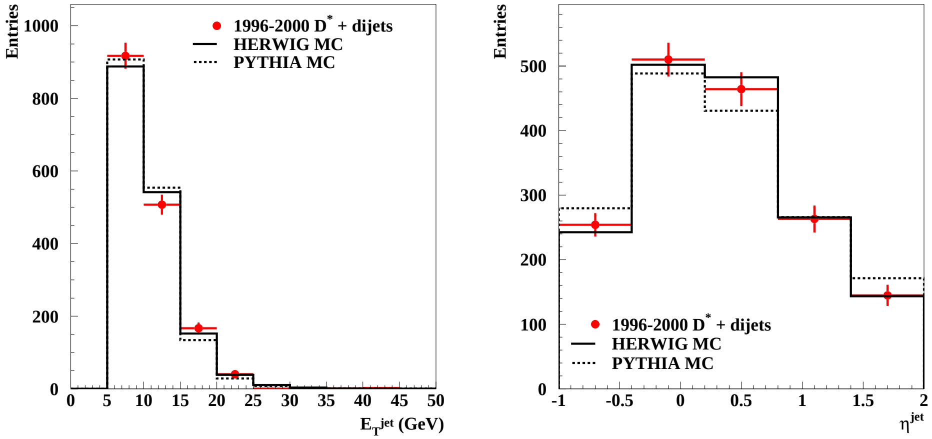

Figure 5.12: E_T^{jet} and η^{jet} distributions for the two jets in the D^* tagged events, compared with the HERWIG and PYTHIA Monte Carlo models.

a light quark or a gluon. The gluon 'contamination' to the sample of quark jets selected in the analysis has been estimated using both HERWIG and PYTHIA Monte Carlo models: for each D^*-tagged dijet event passing all selection cuts described in this chapter, the other jet is matched with the closest outgoing parton in the (η, ϕ) space and the flavour[5] of the outgoing parton is assigned to the reconstructed jet. The results are shown in Fig. 5.16. The estimation of the gluon contamination depends of course on the value of α chosen to combine the direct and resolved components in the Monte Carlo samples. In table 5.2 the estimations obtained using the two most different values for the scale factor α are reported. One can see that, even in

	f_{gluon} (%)			
	D^* **HERWIG MC**		D^* **PYTHIA MC**	
	$\alpha = 75\%$	$\alpha = 69\%$	$\alpha = 65\%$	$\alpha = 62\%$
no x_γ^{OBS} **cut**	15.8	19.5	23.6	25.6
$x_\gamma^{\mathrm{OBS}} \geq 0.75$	4.4	5.7	11.2	12.5
$x_\gamma^{\mathrm{OBS}} < 0.75$	40.0	44.5	50.2	51.8

Table 5.2: Values for the fraction f_{gluon} obtained using the two most different values for the scale factor α for the two Monte Carlo models.

the 'worst' hypothesis (smallest value for α) both Monte Carlo predict that $\lesssim 25\%$ of the total jet sample comes from gluon initiated jets, when no cut on the variable x_γ^{OBS} is applied. The requirement $x_\gamma^{\mathrm{OBS}} \geq 0.75$ yields a significant smaller gluon contamination, but also a reduction in the available statistics. Therefore the analysis has been performed using the total D^*-tagged dijet sample.

[5]The term flavour is here used not only to differentiate between the various quarks but also to indicate the type of the parton, i.e. quark or gluon.

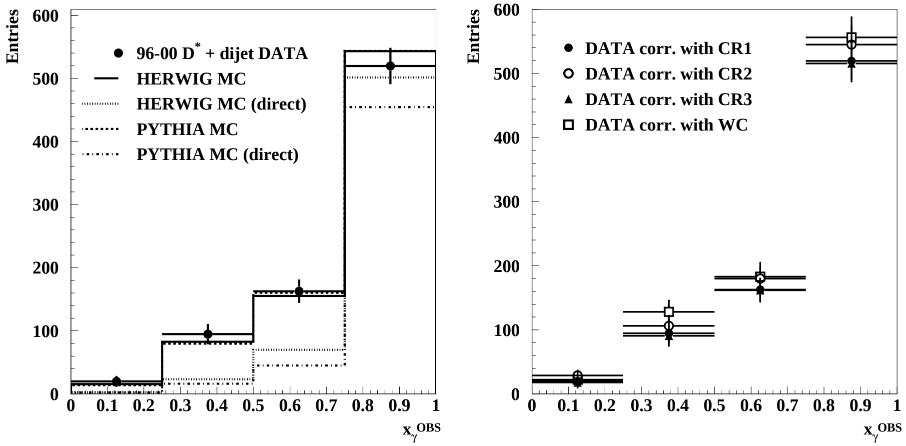

Figure 5.13: The x_γ^{OBS} spectrum (left) for the D^*-tagged dijet subsample after background subtraction, compared with the HERWIG and PYTHIA Monte Carlo predictions. The direct component from both Monte Carlo models is shown separately. Only statistical errors on the data points are shown. The right plot shows the x_γ^{OBS} spectrum obtained using different estimations for the background contribution (as explained in Sec. 5.3).

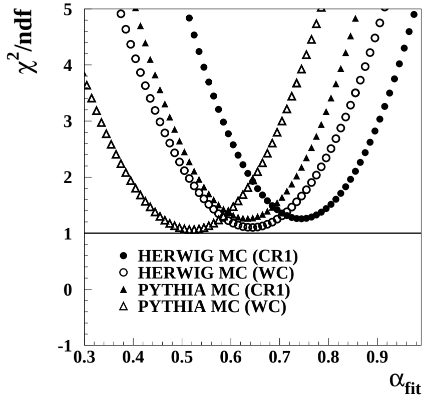

Figure 5.14: χ^2/ndf of the fit to the x_γ^{OBS} distribution (from the 1996-1997 data sample) for the two most different backgrounds (CR1, WC) and for the two Monte Carlo models (HERWIG, PYTHIA).

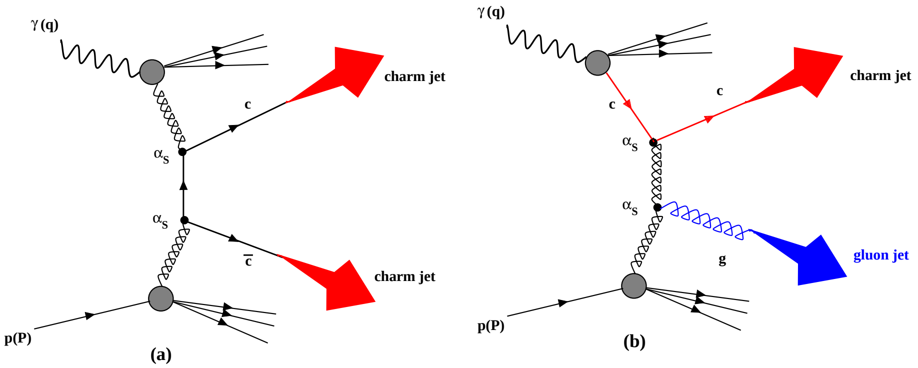

Figure 5.15: Schematic diagrams for charm production in (a) gluon-gluon fusion (b) charm excitation in the photon.

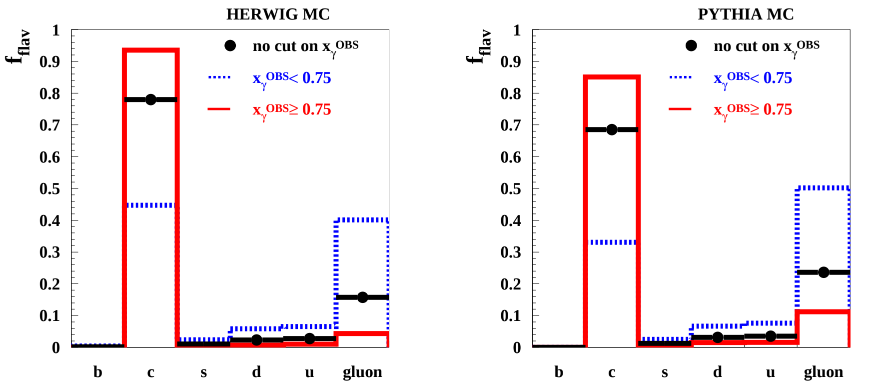

Figure 5.16: Fraction, $f_{\rm flav}$ of the different flavours assigned to the other jet in D^*-tagged dijet events sample. The solid points represent the MC prediction when no cut on $x_\gamma^{\rm OBS}$ is applied. The solid histogram shows the $f_{\rm flav}$ for the sample enriched in direct photoproduction processes ($x_\gamma^{\rm OBS} \geq 0.75$), whereas the dashed histogram shows the result for the sample enriched in resolved photoproduction processes.

Chapter 6

Measurements of the Internal Jet Structure

In this chapter the internal jet structure for the dijet photoproduction sample and the D^*-tagged subsample are presented and compared with the Monte Carlo predictions in different bins of $\eta^{\rm jet}$ and $E_T^{\rm jet}$ at detector level. The method applied to correct for detector effects is explained and the measurements for the observables corrected for detector effects, together with the expectations from the Monte Carlo models, are discussed. The last section is dedicated to the study of the systematic uncertainties.

6.1 Measurements at Detector Level for Inclusive and Charmed Jets

The input for the jet algorithm at detector level consists of all the UCAL cells which survive the noise suppression cuts. The same reconstruction procedure is applied to both real and simulated events.

6.1.1 Jet Shapes and Mean Subjet Multiplicity for the Dijet Photoproduction Sample

In Fig. 6.1 the integrated and differential jet shapes as a function of r/R for the dijet photoproduction sample, selected as explained in chapter 4, are compared with the expectations of the HERWIG and PYTHIA Monte Carlo in four different $\eta^{\rm jet}$ bins. The measured observables are very well described by both MC models. In Fig. 6.2 the integrated and differential jet shapes as a function of r/R for the dijet photoproduction sample are compared with the expectations of the HERWIG and PYTHIA Monte Carlos in three different $E_T^{\rm jet}$ bins. Also in this case the agreement between data and Monte Carlo is very good. Figure 6.3 shows the mean subjet multiplicity as a function of the resolution scale $y_{\rm cut}$ for the dijet photoproduction sample, again compared with the expectations of HERWIG and PYTHIA Monte Carlos in four $\eta^{\rm jet}$ bins and in three different $E_T^{\rm jet}$ bins, respectively.

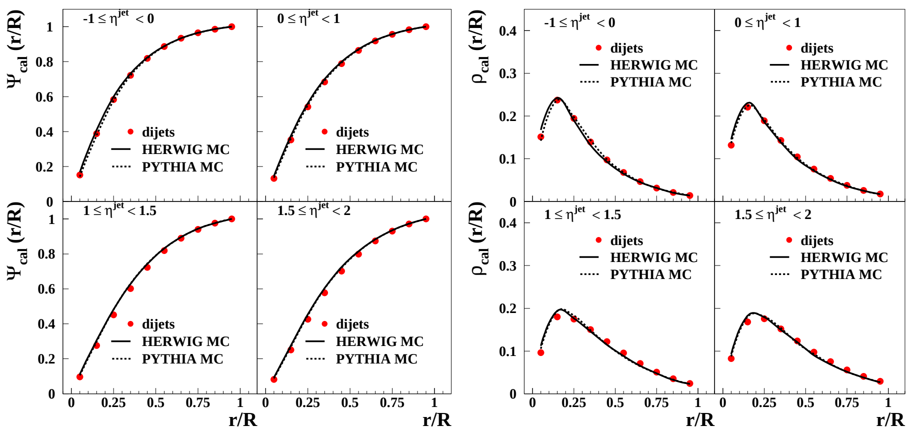

Figure 6.1: Measured integrated jet shape, $\Psi_{\text{cal}}(r/R)$, and differential jet shape, $\rho_{\text{cal}}(r/R)$, for the dijet photoproduction sample in four regions of η^{jet}, integrated over $E_T^{\text{jet1,2}} \geq 7, 6$ GeV. The small statistical errors ($< 1\%$) cannot be seen in the plot. For comparison, the HERWIG and PYTHIA Monte Carlo predictions are shown as lines.

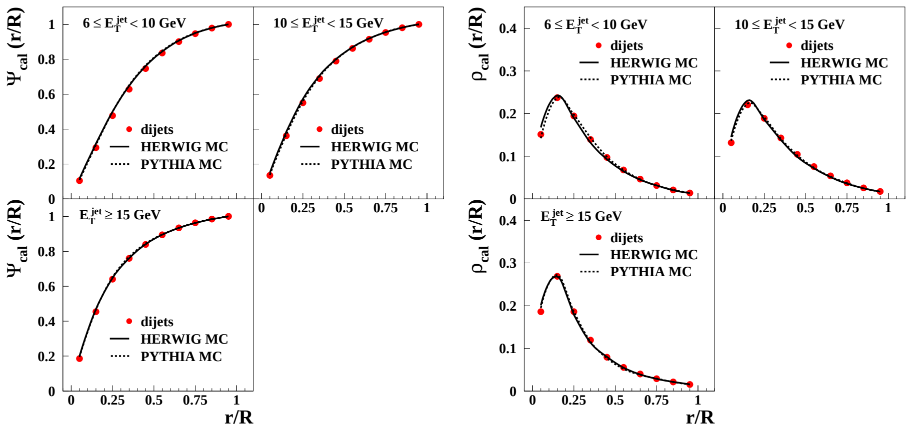

Figure 6.2: Measured integrated jet shape, $\Psi_{\text{cal}}(r/R)$, and differential jet shape, $\rho_{\text{cal}}(r/R)$, for the dijet photoproduction sample in three regions of E_T^{jet}, integrated over $-1 \leq \eta^{\text{jet}} \leq 2$. For comparison, the HERWIG and PYTHIA Monte Carlo predictions are shown as lines.

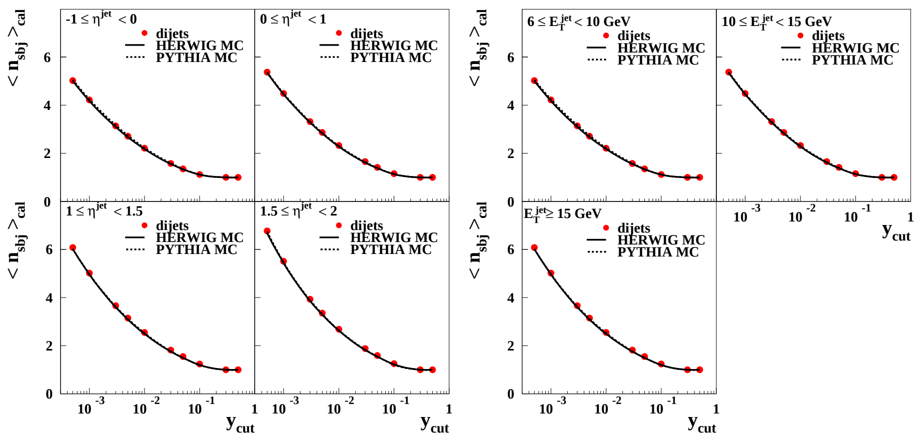

Figure 6.3: Mean subjet multiplicity $\langle n_{\text{sbj}} \rangle_{\text{cal}}$ as a function of the resolution scale y_{cut} for the dijet photoproduction sample in four regions of η^{jet} (left), integrated over $E_T^{\text{jet1,2}} \geq 7, 6$ GeV, and in three regions of E_T^{jet} (right), integrated over $-1 \leq \eta^{\text{jet}} \leq 2$. For comparison, the HERWIG and PYTHIA Monte Carlo predictions are also shown.

6.1.2 Jet Shapes and Mean Subjet Multiplicity for the D^*-tagged Photoproduction Sample

I. Background Subtraction Method

As already discussed in chapter 5, the sample of D^* meson candidates obtained using the reconstruction method described in Sec. 5.2, is not a pure sample of charm hadrons and thus the jet sample, consisting of the 'other' jet in the D^*-tagged dijet events, contains a background contribution which has to be statistically subtracted. The observables measured in this analysis are average values obtained from the total jet sample:

$$\mathcal{O}^{\text{meas}} = \frac{1}{N_{\text{jet}}^{\text{other}}} \sum_{i=1}^{N_{\text{jet}}^{\text{other}}} O_i$$

and can therefore be written as:

$$\mathcal{O}^{\text{meas}} = (1 - f^{\text{back}}) \cdot \mathcal{O}^{\text{sig}} + f^{\text{back}} \cdot \mathcal{O}^{\text{back}} \quad \text{with} \quad f^{\text{back}} = N^{\text{back}}/N^{\text{tot}} \tag{6.1}$$

where f^{back} is the fraction of background jets in the total jet sample and $\mathcal{O}^{\text{sig}}$ and $\mathcal{O}^{\text{back}}$ are the average values for the signal and the background jet sample, respectively. The value for $\mathcal{O}^{\text{sig}}$ is then obtained from the measured observable $\mathcal{O}^{\text{meas}}$ by subtracting the background contribution:

$$\mathcal{O}^{\text{sig}} = \frac{\mathcal{O}^{\text{meas}} - f^{\text{back}} \cdot \mathcal{O}^{\text{back}}}{1 - f^{\text{back}}} \tag{6.2}$$

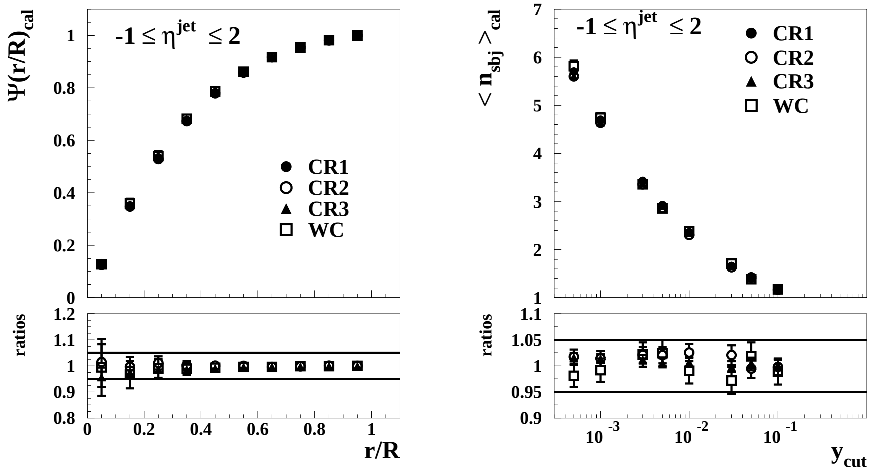

Figure 6.4: Integrated jet shape $\Psi(r/R)_{\rm cal}$ (left plot) and mean subjet multiplicity $\langle n_{\rm sbj}\rangle_{\rm cal}$ (right plot) calculated for the different background regions (CR1, CR2, CR3, WC). The bottom plots show the ratio $\mathcal{O}^{\rm CR1}/\mathcal{O}^{j}$, where $j =$ CR2, CR3, WC. The error bars represent the statistical errors on the ratio $\mathcal{O}^{\rm CR1}/\mathcal{O}^{j}$.

and its statistical error $\sigma_{\rm stat}(\mathcal{O}^{\rm sig})$ is given by:

$$\sigma_{\rm stat}(\mathcal{O}^{\rm sig}) = \frac{1}{(1-f^{\rm back})} \cdot \sqrt{\sigma^2_{\rm stat}(\mathcal{O}^{\rm meas}) + (f^{\rm back}\sigma_{\rm stat}(\mathcal{O}^{\rm back}))^2 + \left(\frac{\mathcal{O}^{\rm meas}-\mathcal{O}^{\rm back}}{1-f^{\rm back}}\right)^2 \sigma^2_{\rm stat}(f^{\rm back})} \tag{6.3}$$

From Eq. 6.3, one can see that the statistical error on the observable $\mathcal{O}^{\rm sig}$ depends strongly on the background contribution to the total jet sample. The tighter cuts in the (Δm, $M(K\pi)$) mass plane applied in the D^* reconstruction are used to reduce the fraction $f^{\rm back}$ ($\simeq 0.3$) and therefore the statistical uncertainty on the final measurements. The estimation of $\mathcal{O}^{\rm back}$ in each bin of $\eta^{\rm jet}$ and $E_T^{\rm jet}$ has been performed using both the wrong charge (WC) and control region combinations (CR1, CR2, CR3). Figure 6.4 shows the integrated jet shape (left plot) and the mean subjet multiplicity (right plot) for the four different background samples, integrated over $-1 \le \eta^{\rm jet} \le 2$. No significant systematic differences are seen and the ratios $\mathcal{O}^{\rm CR1}/\mathcal{O}^{j}$, where $j =$ CR2, CR3, WC, (bottom plots in Fig. 6.4) show that the relative variation are smaller than the statistical errors associated to the measurements. The largest variation is seen when comparing with the WC background, which is however the data sample with the fewest events and with the biggest statistical uncertainty. The background sample with the smallest statistical uncertainty (CR1) is used to extract $\mathcal{O}^{\rm sig}$ whereas the other three samples (CR2, CR3, WC) are used for the estimation of the corresponding systematic error. To elucidate the effect of the background subtraction, the ratios $\mathcal{O}^{j}/\mathcal{O}^{\rm dijet}$, where j indicates the measured (meas), the background (back) and the background corrected (sig) values for the observable $\mathcal{O}^{\rm det}$, are shown in Fig. 6.5. The data points for the background observable $\mathcal{O}^{\rm back}$ are much closer to the ones measured from the dijet sample than the data points measured from the

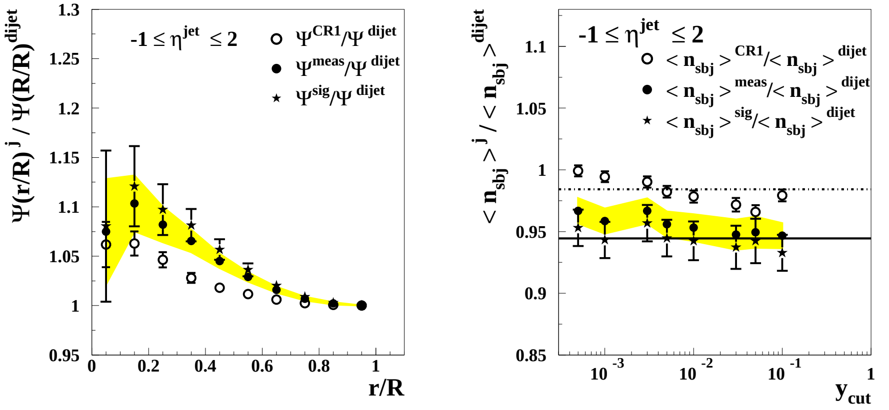

Figure 6.5: Ratios $\Psi^{j}/\Psi^{\rm dijet}$ (left plot) and $\langle n_{\rm sbj}\rangle^{j} / \langle n_{\rm sbj}\rangle^{\rm dijet}$, where j indicates the measured (meas), the background (back, e.g. CR1) and the background corrected (sig) values for the observables $\Psi(r/R)$ and $\langle n_{\rm sbj}\rangle$. The shaded area indicates the statistical errors on the ratio $\mathcal{O}^{\rm meas}/\mathcal{O}^{\rm dijet}$.

D^*-tagged dijet sample. From both distributions of Fig. 6.5 one can clearly see a systematic difference between $\mathcal{O}^{\rm dijet}$ and $\mathcal{O}^{\rm sig}$. The measurements of the internal structure of the other jet in the D^*-tagged dijet sample, before correcting for any detector effects, already shows that the other jet in the D^*-tagged dijet sample is narrower and has a smaller mean subjet multiplicity, as expected for quark initiated jets. Figures 6.6, 6.7 and 6.8 show the integrated and differential jet shapes and the mean subjet multiplicity for the other jet in the D^*-tagged dijet sample after background subtraction. From the comparison with the D^* HERWIG and PYTHIA Monte Carlo expectations a reasonable agreement between data and Monte Carlo can be seen.

6.2 Correction Method from Detector to Hadron Level

Before the measured observables can be compared with the theoretical predictions or with results from other measurements, they must be first corrected ('unfolded') for various detector related effects, such as geometrical acceptance, detector efficiency and resolution, particle decays, interactions with inactive material and also for effects related to the trigger and event selection. In this analysis, multiplicative correction factors C, relating the measured value for an observable $\mathcal{O}^{\rm det}$ at detector level (e.g. $\Psi_{\rm cal}(r/R)$, $\rho_{\rm cal}(r/R)$ or $\langle n_{\rm sbj}\rangle_{\rm cal}$) to the corrected value[1], are used:

$$\mathcal{O}^{\rm corr} = \mathcal{O}^{\rm det} \cdot C \quad \text{with} \quad C = \left.\frac{\mathcal{O}^{\rm had}}{\mathcal{O}^{\rm det}}\right|_{\rm MC} \tag{6.4}$$

[1]In the following, the corrected values for the measured observables will be referred to also as the value corrected at 'hadron level'.

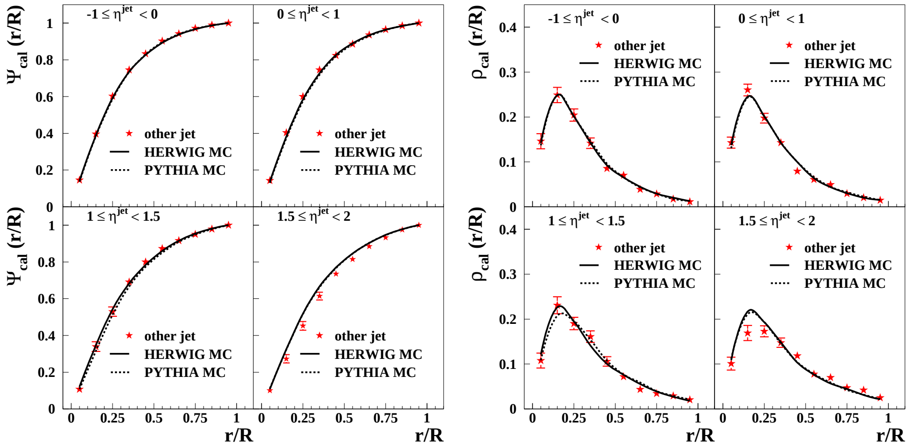

Figure 6.6: Integrated jet shape, $\Psi_{\rm cal}(r/R)$, and differential jet shape, $\rho_{\rm cal}(r/R)$, for the other jet in the D^*-tagged dijet photoproduction sample (after background correction) in four regions of $\eta^{\rm jet}$, integrated over $E_T^{\rm jet1,2} \geq 7,6$ GeV. For comparison the HERWIG and PYTHIA Monte Carlo predictions are shown as lines.

where $\mathcal{O}^{\rm had}$ indicates the value for the observable $\mathcal{O}$, computed using the Monte Carlo simulated hadronic final state before the detector simulation stage. The correction factors are calculated

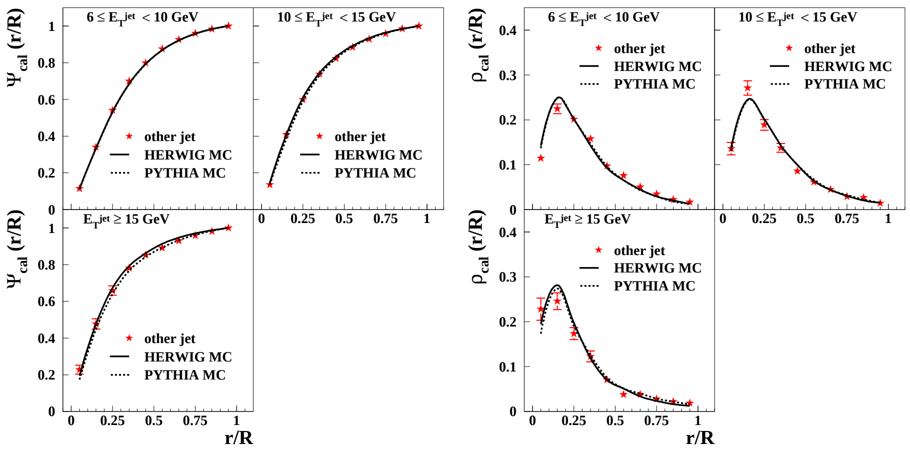

Figure 6.7: Integrated jet shape, $\Psi_{\rm cal}(r/R)$, and differential jet shape, $\rho_{\rm cal}(r/R)$, for the other jet in the D^*-tagged dijet photoproduction sample (after background correction) in three regions of $E_T^{\rm jet}$, integrated over $-1 \leq \eta^{\rm jet} \leq 2$. For comparison, the HERWIG and PYTHIA Monte Carlo predictions are shown as lines.

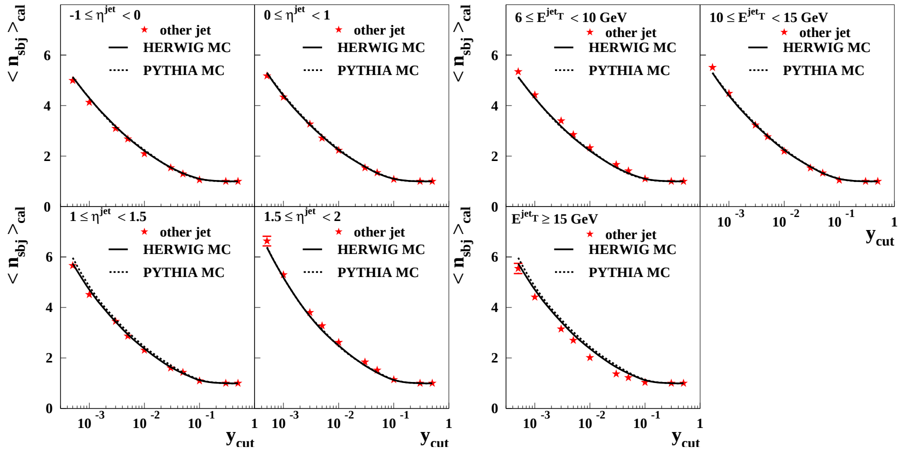

Figure 6.8: Mean subjet multiplicity $\langle n_{\rm sbj}\rangle_{\rm cal}$ as a function of the resolution scale $y_{\rm cut}$ for the other jet in the D^*-tagged dijet photoproduction sample (after background correction) in four regions of $\eta^{\rm jet}$ (left), integrated over $E_T^{\rm jet1,2} \geq 7,6$ GeV, and in three regions of $E_T^{\rm jet}$ (right), integrated over $-1 \leq \eta^{\rm jet} \leq 2$. For comparison, the HERWIG and PYTHIA Monte Carlo predictions are shown as lines.

individually for each bin. Although they are to a good approximation independent of the event generator used, a residual influence remains and must be taken into account in the estimation of the systematic errors. Therefore the two different Monte Carlo generators, HERWIG and PYTHIA, are used to compute the correction factors ($C_{\rm HERWIG}$, $C_{\rm PYTHIA}$) and the difference between the corrected values $\mathcal{O}^{\rm corr}_{\rm HERWIG}$ and $\mathcal{O}^{\rm corr}_{\rm PYTHIA}$ is included into the estimation of the total systematic error on the final measurements.

6.2.1 Definition of the Hadronic Final State

The definition of the hadronic final state is not unique and must be unambiguously specified. The standard criteria to define which particles have to be considered as part of the hadronic final state, is to look at their lifetime and fix a cut value $\tau_{\rm cut}$. All known particles can be divided, according to their lifetime, in four different groups, as schematically shown in Fig 6.9. In this analysis, all the hadrons with lifetime $\tau > 10^{-11}$ s are considered as 'stable' particles whereas all other particles (having $\tau < 10^{-11}$ s) are defined as unstable and their decay products are then included in the hadronic final state. As an example, if a charmed meson D^0 is produced and decays as follows:

$$D^0 \to K^- + \pi^+ + \pi^0(\to \gamma\,\gamma)$$

the following particles are included in the reconstruction of the hadronic final state:

$$K^-,\ \pi^+,\ \gamma,\ \gamma$$

The definition of the hadronic final state is a crucial point when investigating exclusive prop-

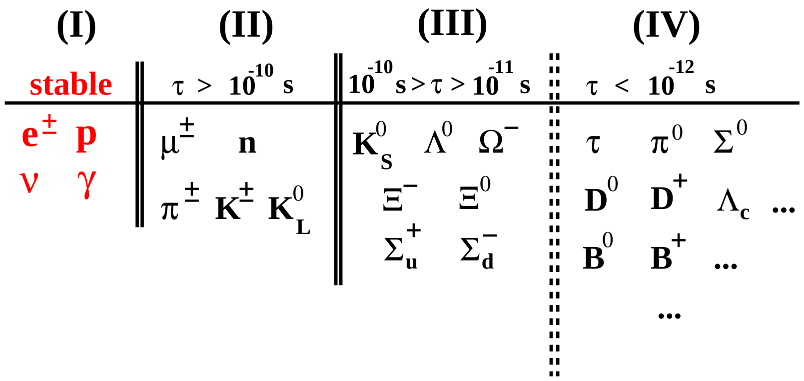

Figure 6.9: Classification of the final state particles in four different groups according to their lifetime.

erties, as in the case of the internal structure of jets, particularly when aiming at comparison between measurements performed in different kinematic regimes in the same experiment (e.g. photoproduction versus deep inelastic scattering at HERA) or in different experiments (e.g. lepton-proton scattering versus electron-positron annihilation). Depending on the definition adopted, the correction factors C may vary significantly. A striking example is shown in Fig. 6.10, where the mean subjet multiplicity for quark and gluon jets, separately, from the HERWIG Monte Carlo simulation is shown for two different definition of the hadronic final state:

$$\text{(a)} \quad \tau > 10^{-13}\,s \qquad \text{(b)} \quad \tau > 10^{-11}\,s$$

In the case (a) charmed mesons such as $D^{\pm}$, D^0, etc. are considered as stable particles[2]. The observable $\langle n_{\mathrm{sbj}} \rangle$ for the quark jets (left plot in Fig. 6.10) shows a significant dependence on the convention used to define the hadronic final state, whereas for the gluon jets (right plot in Fig. 6.10) no significant difference is seen. Figure 6.11 shows the observable $\langle n_{\mathrm{sbj}} \rangle$ at hadron level for quark jets separated according to the flavour of the parton from which they originate. One can clearly see that the large difference observed in the quark jets originated from the charm contribution[3]. According to definition (b) the charmed mesons are considered unstable, and their decay products are included in hadronic final state. This of course leads to a higher particle multiplicity than in the case when definition (a) is used. More importantly, since the charm is a heavy quark, the decay products of charmed particles can have a high relative transverse momentum k_T and can therefore be reconstructed as separated sub-objects within a jet if the resolution scale y_{cut} is small enough, i.e. $y_{\mathrm{cut}} \lesssim (m_c/E_T^{\mathrm{jet}})^2$. One can also notice that in both cases (a) and (b), charm jets are predicted to have a different internal structure when compared with jets originating from light quarks at hadron level. In the case of the definition (b), this effect is more pronounced for the HERWIG predictions (as can be seen by comparing them with the PYTHIA predictions in Fig. 6.11 and 6.12). Moreover, in the

[2] This definition has been used by the ZEUS collaboration during various years.

[3] The contribution of beauty jets to the dijet event sample was estimated using the LO Monte Carlo models and it was found to be negligible: the HERWIG (PYTHIA) Monte Carlo predicts (before re-weighting) a beauty contribution to the dijet sample, in the kinematic region chosen for this analysis, of $\sim 4.6(5.7)\%$ for the direct processes and of $\sim 0.4(0.8)\%$ for resolved processes. Therefore in Fig. 6.11 the internal structure for beauty jets is not shown.

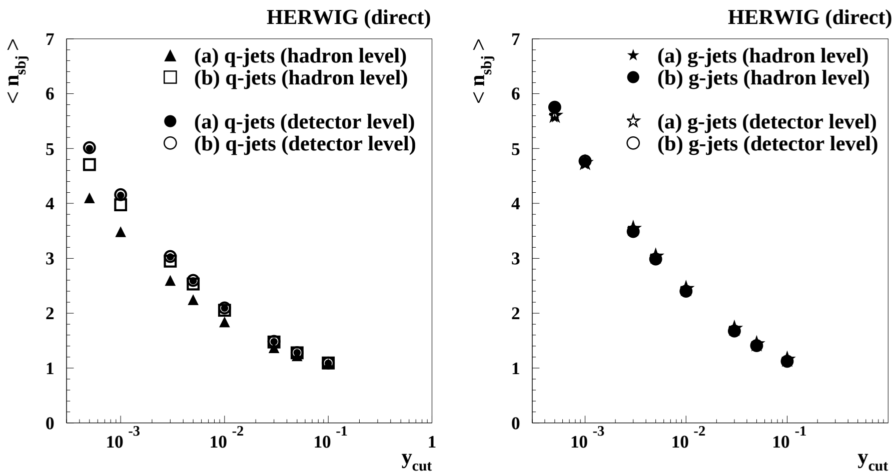

Figure 6.10: $\langle n_{\mathrm{sbj}} \rangle$ for quark jets (left plot) and gluon jets (right plot) for two different definitions of the hadronic final state, calculated using the HERWIG Monte Carlo.

resolved processes, where 'extra' activity in the final state is present (due to the implementation of the multiple interaction in the event simulation), the effect becomes smaller: in the case of PYTHIA, when using the definition (b), the difference between the charm and the light quark jets almost vanishes.

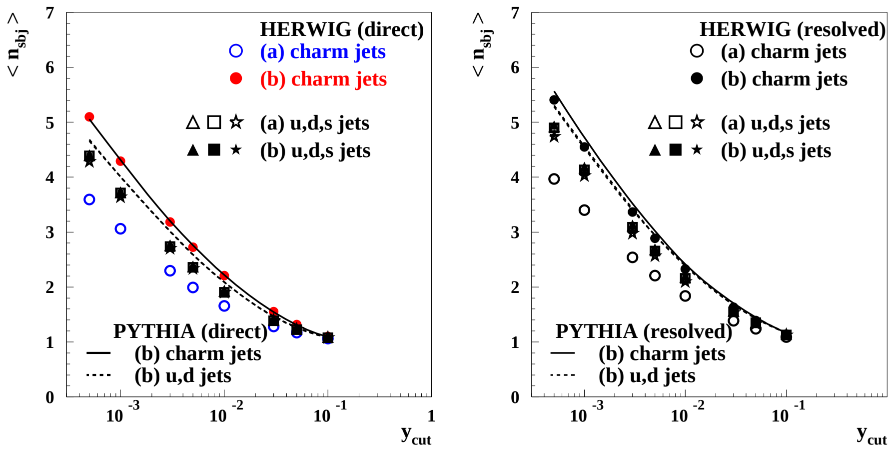

Figure 6.11: $\langle n_{\mathrm{sbj}} \rangle (y_{\mathrm{cut}})$ for light (u, $\bar{u}$, d, $\bar{d}$, s, $\bar{s}$,) and heavy quark (c, $\bar{c}$) jets for two different definition of the hadronic final state, calculated using the HERWIG and PYTHIA Monte Carlos.

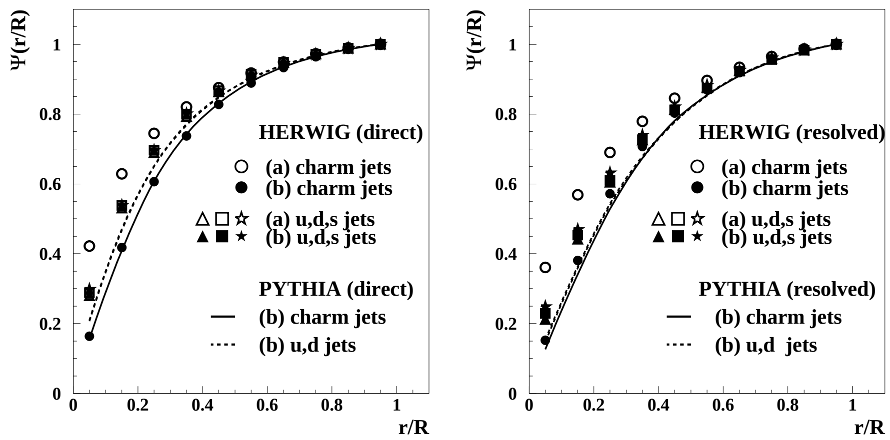

Figure 6.12: $\Psi(r/R)$ for light (u, $\bar{u}$, d, $\bar{d}$, s, $\bar{s}$,) and heavy quark (c, $\bar{c}$) jets for two different definition of the hadronic final state, calculated using the HERWIG and PYTHIA Monte Carlos.

Since the measured observables at detector level are independent from the hadron level definition (as shown in Figure 6.10), the correction factors calculated in the two cases (a), (b) are clearly different and lead thus to a different value for the corrected observable $\mathcal{O}^{\rm corr}$. Therefore only measurements using the same definition of the final state can be directly compared.

6.2.2 Difference between Light and Heavy Quark Jets

The difference in the internal structure at hadron level between light and heavy quark jets (shown in figures 6.11 and 6.12) can be understood, in the parton shower formalism, in terms of the difference between the quark masses. As already mentioned in Sec. 3.2.1, soft gluon emission around the heavy quark direction is suppressed within an angle of the order m_q/E_q, where m_q and E_q are the mass and the energy of the heavy quark, respectively. Moreover, the parton shower evolution is terminated at a cut-off value of the form:

$$Q_i = m_i + Q_0 \quad i = d,\, u,\, s,\, c,\, b$$

where m_i is the quark mass and Q_0 is a parameter of the order of $\sim$ 1 GeV. So, parton branchings from heavy quarks stop 'earlier', i.e. at a higher virtuality Q_i. Fewer hadrons are then formed in the hadronization process, which subsequently decay. If the charm mesons are stable (definition (a)), the heavy quark jets at hadron level are then narrower and have fewer sub-objects when compared with light quark jets of the same tranverse energy. When using the definition (b) for the hadronic final state, because of the higher particle multiplicity in the decay of the heavy quark hadrons, the opposite effects are seen in the internal jet structure, as long as the energy scale involved ($E_T^{\rm jet}$) is comparable with the heavy quark mass (m_Q), $E_T^{\rm jet}/m_Q \lesssim \mathcal{O}(10)$. Figure 6.13 shows the integrated charged multiplicity and the integrated energy fraction for quark and gluon jets measured in three-jet events at LEP [43], where the effects due to the heavy quark (beauty) mass are clearly seen.

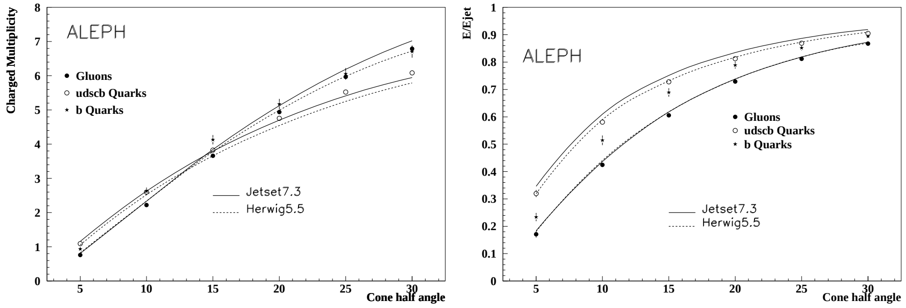

Figure 6.13: Integrated charged multiplicity (left plot) and integrated energy fraction (right plot) within successive cones for quark and gluon jets measured in three-jet events at LEP [43]. The average jet energy is $\left\langle E_T^{\rm jet}\right\rangle \simeq 24$ GeV. One can see that beauty jets have a higher multiplicity and are broader than light quark jets.

6.3 Measurements Corrected for Detector Effects

The corrected measurements are given for a hadronic final state defined as follows:

- $Q^2 \leq 1$ GeV^2
- $0.2 \leq y \leq 0.85$ corresponding to $134 \leq W \leq 277$ GeV
- $E_T^{\rm jet1,2} \geq 7,6$ GeV and $-1 \leq \eta^{\rm jet} \leq 2$
- definition (b) for the hadrons in the final state (see Sec. 6.2.1)

Note that the limits on the y range are $\sim$ 20% higher than the selection cuts applied on the variable $y_{\rm JB}$ (see Sec. 4.2). This takes into account the systematic deviation of the measured $y_{\rm JB}$ from the true value y. The correlation between the two quantities is shown in Fig. 6.14. The observed deviation is essentially caused by the energy loss and the smearing of particles traversing the inactive material in front of the calorimeter and by the loss of particles through the rear beam-pipe hole.

The correction factors C for all the $\eta^{\rm jet}$ and $E_T^{\rm jet}$ bins used in this analysis can be found in App. B. Figures 6.16, 6.17, and 6.18, 6.19 show the measurements for the observables corrected at hadron level:

$$\langle n_{\rm sbj}\rangle_{\rm corr} = \langle n_{\rm sbj}\rangle_{\rm cal}\cdot C(\langle n_{\rm sbj}\rangle), \quad \Psi_{\rm corr} = \Psi_{\rm cal}\cdot C(\Psi), \quad \text{and} \quad \rho_{\rm corr} = \rho_{\rm cal}\cdot C(\rho)$$

for the dijet sample and the other jet in the D^*-tagged dijet sample, respectively, integrated over $-1 \leq \eta \leq 2$.

From the comparison of the mean subjet multiplicity and the jet shapes for the dijet photoproduction sample with the HERWIG and PYTHIA predictions (Fig. 6.16, 6.17), one can see that the leading-logarithm parton shower approximation (implemented in both Monte Carlo programs to model the internal jet structure) provides a good description of the data. The mean subjet multiplicity and the jet shapes, as predicted by PYTHIA, for quark and gluon jets are shown separately. The predictions from the Monte Carlo are separately calculated for direct

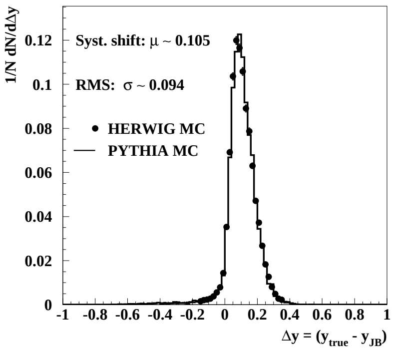

Figure 6.14: Correlation between the measured quantity y_{JB} and the generated y.

and resolved processes and then averaged according to the normalization determined by the fit to the x_γ^{OBS} distribution. One can see that the predicted mean subjet multiplicity is larger for gluon jets than for quark jets and that the data points are located between the two curves, showing that the dijet sample selected is a mixture of quark and gluon jets, as predicted in the frame of pQCD. The same conclusion can be drawn by looking at the jet shapes. The quark jets are predicted to be narrower then the gluon jets, and also in this case, the data points are located in between the two curves.

The procedure used, for the simulated events, to classify a jet (at hadron level) as a quark or a gluon jet is the following:

- the distances ΔR_{had} in the $(\eta,\ \phi)$ plane between the jet and each of the two outgoing

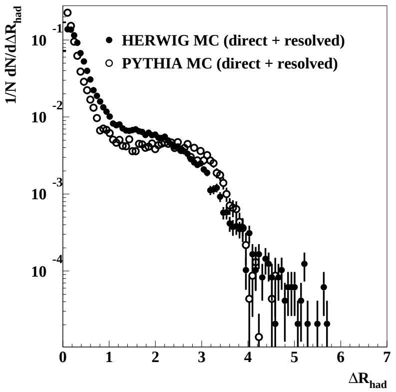

Figure 6.15: ΔR_{had} distribution between the hadronic jet and its closest parton.

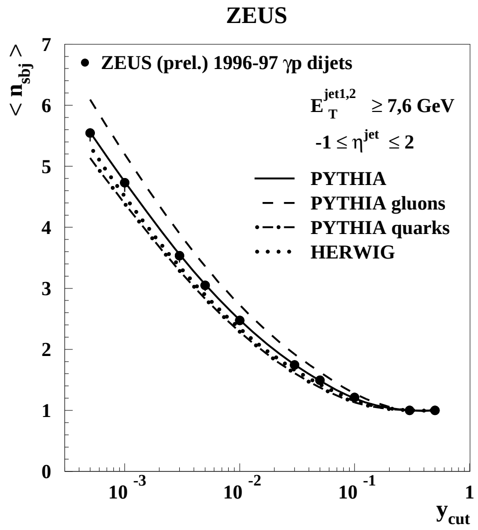

Figure 6.16: Mean subjet multiplicity $\langle n_{\mathrm{sbj}}\rangle\,(y_{\mathrm{cut}})$ corrected to hadron level for $E_T^{\mathrm{jet1,2}} \geq 7,6$ GeV, integrated over $-1 \leq \eta^{\mathrm{jet}} \leq 2$, for the dijet photoproduction sample. The predictions of the PYTHIA (solid line) and HERWIG (dotted line) Monte Carlos including direct plus resolved processes are shown. For the PYTHIA Monte Carlo, the expectations for quark jets (dash-dotted line) and gluon jets (dashed line) are shown.

partons from the two-to-two subprocess are calculated

- the parton which is closer to the jet is associated to it and its flavour (i.e. quark or gluon) is used to define the jet type

Figure 6.15 shows the ΔR_{had} distribution between the hadronic jet and its closest parton, calculated using both the HERWIG and PYTHIA Monte Carlos.

In Fig.6.18 and 6.19 the mean subjet multiplicity and the jet shapes for the other jet in the D^*-tagged dijet sample are compared with the HERWIG and PYTHIA Monte Carlo predictions[4]. The agreement between data and Monte Carlo is very good and it can be clearly seen that the data points are consistent with the predictions for a pure sample of quark jets. The prediction for a pure sample of charm jets is also shown and proven to be in perfect agreement with the measured distributions.

6.4 Study of the Systematic Uncertainties

In the following the systematic checks which were performed for this analysis are listed. They are grouped in four classes which are assumed to be approximately independent:

[4]Also in this case, the predictions from the Monte Carlo are calculated separately for direct and resolved processes and then averaged according to the normalization determined by the fit to the x_γ^{OBS} distribution.

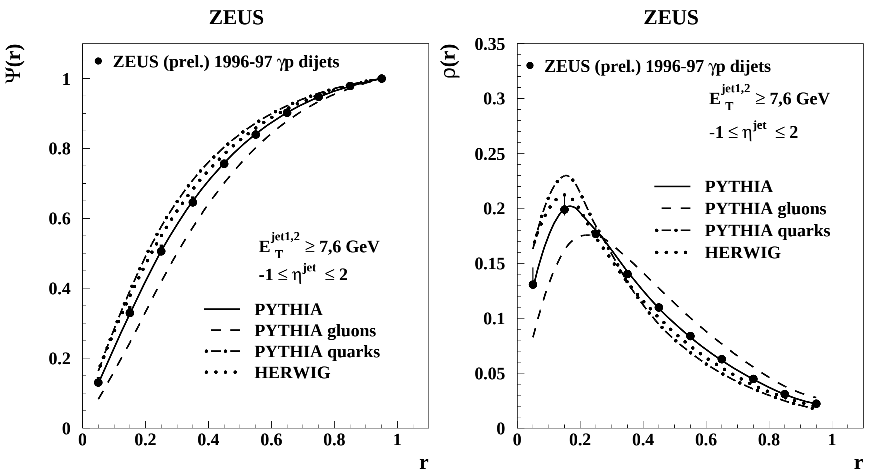

Figure 6.17: Integrated jet shape, $\Psi(r/R)$, and differential jet shape, $\rho(r/R)$, corrected to hadron level for $E_T^{\rm jet1,2} \geq 7,6$ GeV, integrated over $-1 \leq \eta^{\rm jet} \leq 2$, for the dijet photoproduction sample. The predictions of the PYTHIA (solid line) and HERWIG (dotted line) Monte Carlos including direct plus resolved processes are shown. For the PYTHIA Monte Carlo, the expectations for quark jets (dash-dotted line) and gluon jets (dashed line) are also shown.

1. checks performed on the kinematic selection cuts:
   - the cut on the reconstructed vertex range is varied between ± 35 cm and ± 55 cm
   - the cut on the total vectorial transverse momentum p_T divided by the total transverse energy E_T is varied between $p_T/\sqrt{E_T} \leq 1.5$ GeV$^{1/2}$ and $p_T/\sqrt{E_T} \leq 2.5$ GeV$^{1/2}$
   - the cut on $y_{\rm elec}$ is varied between $y_{\rm elec} > 0.8$ and $y_{\rm elec} > 0.6$
   - the variable $y_{\rm JB}$ is varied by $\pm 10\%$. This provides an estimation of the effect due to the eventual energy scale difference between data and Monte Carlo and, since the $y_{\rm JB}$ distribution is poorly described by the Monte Carlo at low values of $y_{\rm JB}$, it is also used to check the effect of the kinematic selection cut
2. checks performed on the calorimeter energy scale:
   - the reconstructed transverse jet energy is varied (only in the data samples) by $\pm 5\%$ according to the estimated uncertainty on the calorimeter energy scale
   - the noise cut for isolated EMC (HAC) cells is raised to 140 (180) MeV
3. checks performed on the Monte Carlo simulation:
   - the HERWIG Monte Carlo is used for the unfolding procedure
   - the scale factor α used to combine direct and resolved processes is varied by $\pm 15\%$
4. checks performed on the D^* selection and background subtraction method:

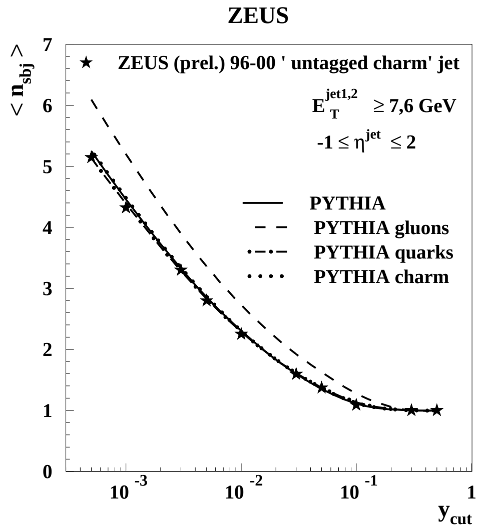

Figure 6.18: Mean subjet multiplicity $\langle n_{\rm sbj}\rangle\,(y_{\rm cut})$ corrected to hadron level for $E_T^{\rm jet1,2} \geq 7,6$ GeV, integrated over $-1 \leq \eta^{\rm jet} \leq 2$, for the other jet in the D^*-tagged dijet photoproduction sample. The predictions of the PYTHIA Monte Carlo (solid line) including direct plus resolved processes is shown. The expectations for quark jets (dash-dotted line), gluon jets (dashed line) and for only charm jets (dotted line) are also shown. Note that in the plot the other jet is labelled as "untagged charm" jet. This convention will be used in many of the plots to follow.

   - the cut on $p_T(\pi_s)$ is varied between $p_T(\pi_s) > 0.175$ GeV and $p_T(\pi_s) > 0.225$ GeV
   - the cut on $p_T(K,\pi)$ is varied between $p_T(K,\pi) > 0.450$ GeV and $p_T(K,\pi) > 0.550$ GeV
   - the cut on the $\Delta\phi \equiv |\phi_{D^*} - \phi_{\rm jet}|$ is varied between $\Delta\phi \equiv |\phi_{D^*} - \phi_{\rm jet}| > 40°$ and $\Delta\phi \equiv |\phi_{D^*} - \phi_{\rm jet}| > 50°$
   - the wrong charge (WC) and the control region track combinations (CR2, CR3) are used for the estimation of the background contribution to the D^* signal and to the measured jet observables

For each change the corrected observable $\mathcal{O}^{\rm corr}$ has been recalculated and compared to the **central values** in each region of $\eta^{\rm jet}$ and $E_T^{\rm jet}$. The central values ($\mathcal{O}^{\rm corr}_{\rm central}$) have been obtained with the following specifications:

- Monte Carlo generator: **PYTHIA** for direct and resolved processes (the latter including MI)[5]
- Monte Carlo normalization according to the fit to the $x_\gamma^{\rm OBS}$ distribution

[5]The proton and photon parton distributions used are **MRSA** and **GRV-LO**, respectively.

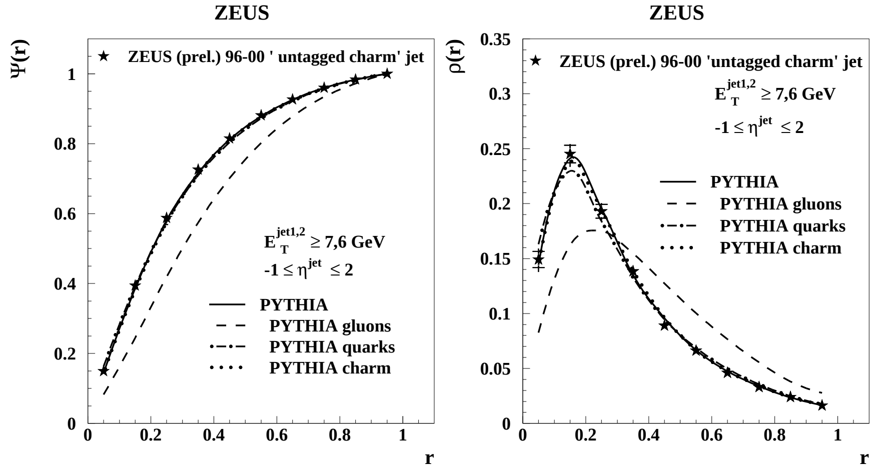

Figure 6.19: Integrated jet shape, $\Psi(r/R)$, and differential jet shape, $\rho(r/R)$, corrected to hadron level for $E_T^{\mathrm{jet1,2}} \geq 7,6$ GeV, integrated over $-1 \leq \eta^{\mathrm{jet}} \leq 2$, for the other jet in the D^*-tagged dijet photoproduction sample. The predictions of the PYTHIA Monte Carlo (solid line) including direct plus resolved processes is shown. The expectations for quark jets (dash-dotted line), gluon jets (dashed line) and only charm jets (dotted line) are also shown.

- reconstruction of the jet observables using calorimeter cells

and with the following additional specification for the D^*-tagged dijet sample:

- background subtraction using the control region CR1

The systematic uncertainty is defined as:

$$\sigma_{\mathrm{sys}} = \frac{\mathcal{O}^{\mathrm{corr}}_{\mathrm{new}} - \mathcal{O}^{\mathrm{corr}}_{\mathrm{central}}}{\mathcal{O}^{\mathrm{corr}}_{\mathrm{central}}}$$

where $\mathcal{O}^{\mathrm{corr}}_{\mathrm{new}}$ is the value obtained when one of the cuts is changed. For each bin, the positive and negative contributions from the different systematic uncertainties are separately added in quadrature to determine the total (positive and negative) systematic uncertainty.

For the dijet photoproduction sample, the largest systematic uncertainty (up to $\sim 3\%$ for the $\langle n_{\mathrm{sbj}} \rangle$ and up to $\sim 15\%$ for $\Psi(r/R)$ and $\rho(r/R)$ at small radius $r/R \leq 0.2$) comes from the use of the HERWIG Monte Carlo in the unfolding procedure. For all the other checks performed, variations of less than 2% are observed.

For the D^*-tagged photoproduction sample, the largest systematic uncertainty (up to $\sim 3\%$ for the $\langle n_{\mathrm{sbj}} \rangle$ and up to $\sim 10\%$ for $\Psi(r/R)$ and $\rho(r/R)$) comes from the use of the HERWIG Monte Carlo in the unfolding procedure and from the use of the wrong-charge combinations for the background estimation.

Chapter 7

Results

In this chapter the distributions of the mean subjet multiplicity and the jet shape corrected at hadron level are presented as a function of the jet pseudo-rapidity η^{jet} and the transverse jet energy E_T^{jet} and compared with the available pQCD predictions from the LO Monte Carlo models. The comparison between the observables measured for the exclusive dijet photoproduction and the charm-enriched jet sample is presented and discussed. A simple statistical method to extract the measurement of the internal structure for a pure sample of gluon jets is explained and the results for the mean subjet multiplicity and the integrated jet shape are shown.

7.1 Comparison of the Internal Jet Structure for Inclusive and Charmed Jets in Photoproduction

7.1.1 Dependence on the Jet Pseudorapidity

The mean subjet multiplicity, $\langle n_{\mathrm{sbj}} \rangle$, for the dijet photoproduction sample is presented in Fig. 7.1 as a function of the resolution scale y_{cut}, in four different η^{jet} regions. One can see that for smaller values of y_{cut}, an increasing number of sub-structures with smaller relative transverse momenta is resolved and that the Monte Carlo predictions (including direct and resolved processes) are in good agreement with the data. The mean subjet multiplicity for quark and gluon jets are shown separately in each η^{jet} bin. The predicted mean subjet multiplicity is larger for gluon jets than for quark jets and exhibits a small dependence with η^{jet}. At small η^{jet} values ($\eta^{\mathrm{jet}} \leq 0$), the data points are very close to the prediction for a pure sample of quark jets whereas at large η^{jet} values ($\eta^{\mathrm{jet}} \geq 1.5$) the data points are very close to the predictions for a pure sample of gluon jets. In the central bins ($0 \leq \eta^{\mathrm{jet}} \leq 1.5$) the data points are located in between the two curves. Figure 7.2 shows the integrated jet shape as a function of the radius r/R in four different η^{jet} regions. One can see that more than 60% of the transverse jet energy is already contained in cone of radius $r/R = 0.4$. Also in this case, the Monte Carlo predictions are in good agreement with the data and the data points are again very close to the prediction for a pure sample of quark jets, at small η^{jet} values ($\eta^{\mathrm{jet}} \leq 0$), and very close to the prediction for a pure sample of gluon jets at large η^{jet} values ($\eta^{\mathrm{jet}} \geq 1.5$). To investigate this effect, the kinematics for the different two-body subprocesses contributing, at leading order, to the dijet photoproduction, has been studied. In the case of direct processes, where the dominant subprocess in the kinematic regime studied is photon-gluon fusion $\gamma g \rightarrow q\bar{q}$,

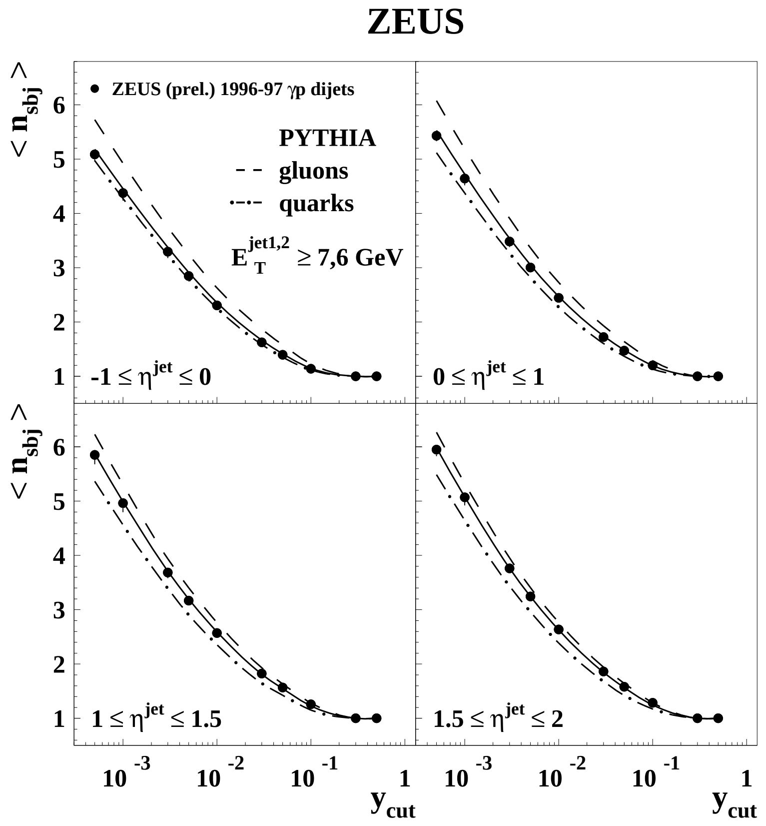

Figure 7.1: Mean subjet multiplicity $\langle n_{\rm sbj}\rangle$ for the dijet photoproduction sample with $E_T^{\rm jet} \geq 7,6$ GeV in different $\eta^{\rm jet}$ bins. The error bars (statistical and systematic errors added in quadrature) are small and not always visible in the plot. The prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown.

PYTHIA (HERWIG) predicts the fraction of quark jets to be 93(94)% for $-1 \leq \eta^{\rm jet} \leq 0$ and 82(87)% for $1.5 \leq \eta^{\rm jet} \leq 2$. In the case of resolved processes, the dominant subprocesses in the kinematic regime studied are the gluon-gluon fusion and $q_\gamma g_p \to qg$, and the fraction of quark jets has a stronger $\eta^{\rm jet}$ dependence: from $\sim 57(63)\%$ for $-1 \leq \eta^{\rm jet} \leq 0$ to $\sim 33(32)\%$ for $1.5 \leq \eta^{\rm jet} \leq 2$. So, the larger mean subjets multiplicity seen in the dijet photoproduction data as $\eta^{\rm jet}$ increases is consistent with an increased fraction of gluon jets in the forward direction.

The D^*-tagged photoproduction sample offers the possibility to disentangle the quark and gluon contribution in the different $\eta^{\rm jet}$ regions by comparing the measurements for the dijet sample with the measurements for the charm-enriched sample (other jet in D^*-tagged events). Figures 7.3 and 7.4 show the mean subjet multiplicity and the integrated jet shape for the other jet in

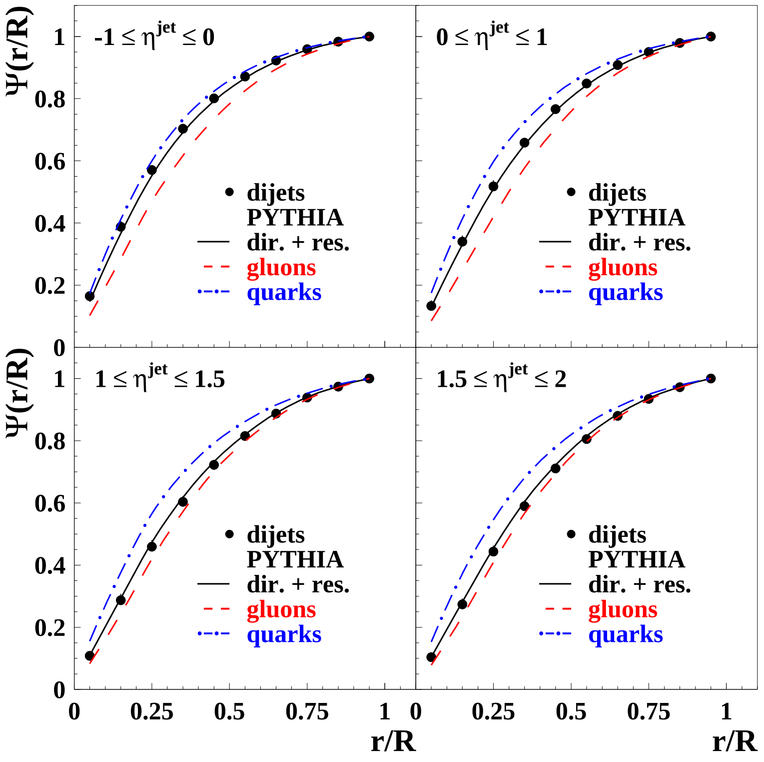

Figure 7.2: Integrated jet shape $\Psi(r/R)$ for the dijet photoproduction sample with $E_T^{\rm jet} \geq 7,6$ GeV in different $\eta^{\rm jet}$ bins. The prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown.

the D^*-tagged photoproduction sample in four different $\eta^{\rm jet}$ regions. It can be seen that the variable $\langle n_{\rm sbj}\rangle$ for the other jet is consistent with the prediction for a pure sample of quark jets for $-1 \leq \eta^{\rm jet} \leq 1.5$. The same agreement is observed for the integrated jet shape, $\Psi(r/R)$. In the highest $\eta^{\rm jet}$ bin, for both observables, the data points show a deviation from the prediction of a pure sample of quark jets. Since the estimated gluon contamination to the charm-jet sample (mainly due to the photon excitation process, as discussed in Sec. 5.4) has its largest contribution in the forward region, this deviation could be explained by the increase of the gluon jet fraction in the charm-enriched sample. However , the Monte Carlo prediction in the highest $\eta^{\rm jet}$ bin does not describe the data as well as for the inclusive dijets. The D^* Monte Carlo samples were generated without implementing the effect of multiple interaction, thus resulting in the lack of 'extra' activity, particularly in the forward direction. This could partially explain the worse agreement between data and Monte Carlo for $\eta^{\rm jet} > 1.5$. To elucidate the effect of

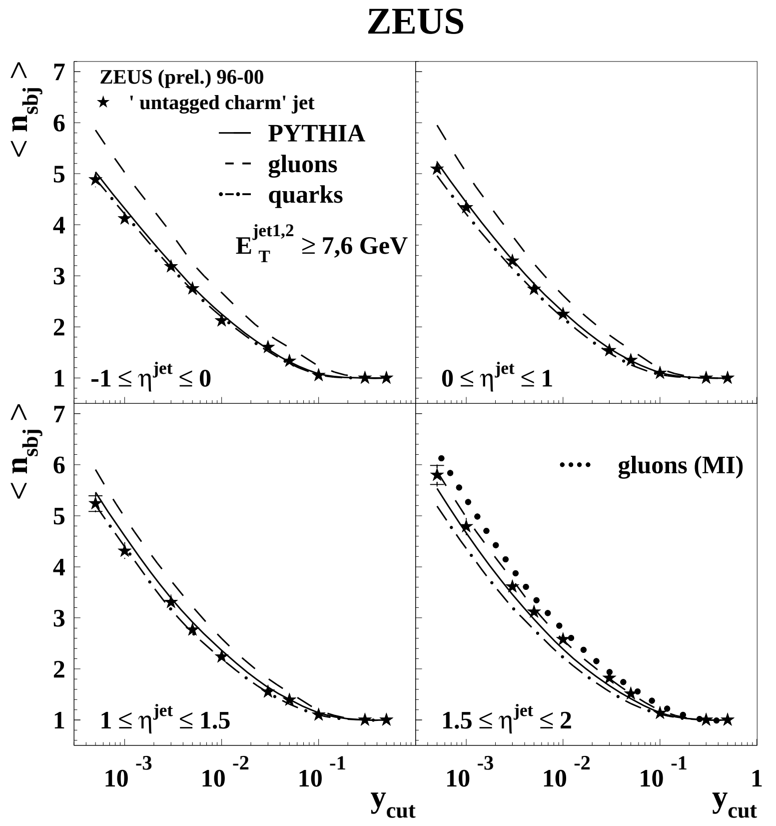

Figure 7.3: Mean subjet multiplicity $\langle n_{\rm sbj} \rangle$ for the other jet in the D^*-tagged dijet photoproduction sample with $E_T^{\rm jet} \geq 7,6$ GeV in different $\eta^{\rm jet}$ bins. The inner and outer error bars denote the statistical errors and the statistical and systematic errors added in quadrature, respectively. For comparison, the prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown. For $1.5 \leq \eta^{\rm jet} \leq 2.0$, the prediction of PYTHIA for gluon jets with multiple interaction (MI) is also shown.

the multiple interaction, the difference between the prediction for a pure sample of gluon jets when implementing or not the multiple interaction in the Monte Carlo simulation is shown in Fig. 7.3 and 7.4 for the most forward bin in $\eta^{\rm jet}$.

To check whether the effect seen in the data is due to the resolved contribution, the direct-enriched D^*-tagged sample (i.e. $x_\gamma^{\rm OBS} \geq 0.75$) has been separately analysed. Figure 7.5 shows the comparison between the measurements for the direct-enriched D^*-tagged sample and the total D^*-tagged sample. Although the errors on the measurements are relatively large, a clear difference between the total and the direct-enriched sample, in the highest $\eta^{\rm jet}$ bin, is seen for

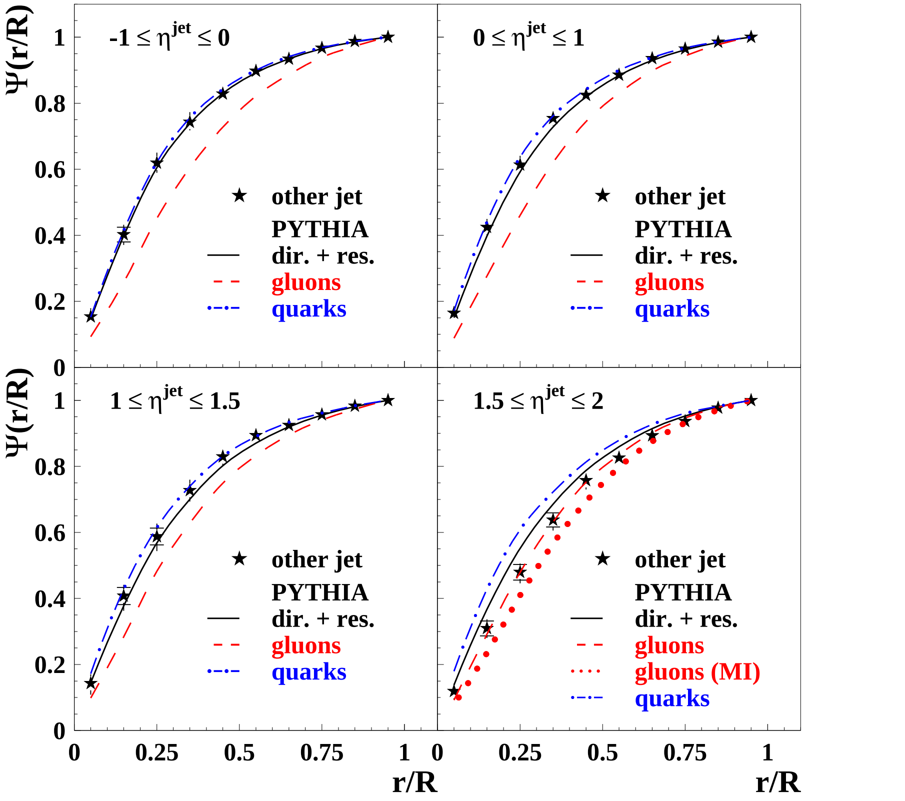

Figure 7.4: Integrated jet shape $\Psi(r/R)$ for the other jet in the D^*-tagged photoproduction sample with $E_T^{\rm jet} \geq 7,6$ GeV in different $\eta^{\rm jet}$ bins. The prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown. For $1.5 \leq \eta^{\rm jet} \leq 2.0$, the prediction of PYTHIA for gluon jets with multiple interaction (MI) is also shown.

both the mean subjet multiplicity (at a fixed value of the resolution scale $y_{\rm cut} = 0.01$) and the integrated jet shape (at a fixed value of the radius $r/R = 0.3$). The Monte Carlo predictions for $x_\gamma^{\rm OBS} \geq 0.75$ are in very good agreement with the data points. For the direct-enriched sample, no $\eta^{\rm jet}$ dependence is seen (as predicted by the Monte Carlo models). So, the measurements in the highest $\eta^{\rm jet}$ bin are consistent with the prediction of charm excitation processes whose kinematics is such that a large fraction of gluon jets is expected in the forward direction. Figure 7.6 shows the mean subjet multiplicity (at a fixed value of $y_{\rm cut} = 0.01$) and the integrated jet shape (at a fixed value of $r/R = 0.3$) for the dijets and the other jet in the D^*-tagged events. The difference between the two sets of data points can be clearly seen. The results for the other jet in the D^*-tagged photoproduction sample are consistent with the prediction for a

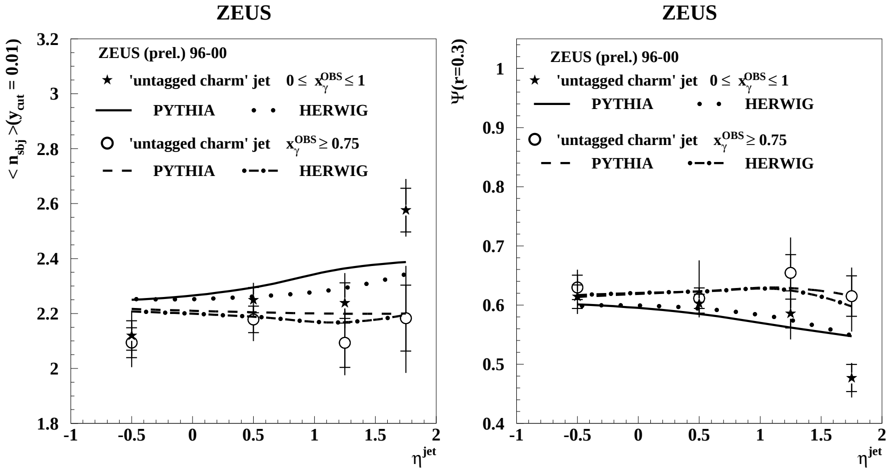

Figure 7.5: Mean subjet multiplicity (left plot) at a fixed value of $y_{\rm cut} = 0.01$ and integrated jet shape (right plot) at a fixed value of $r/R = 0.3$ as a function of $\eta^{\rm jet}$ for the other jet in the total and the direct-enriched D^*-tagged sample ($x_\gamma^{\rm OBS} \geq 0.75$), compared with the predictions from the PYTHIA and HERWIG Monte Carlos.

pure sample of quark jets (in the $\eta^{\rm jet}$ regions where the gluon contamination to the sample is negligible).

7.1.2 Evolution with the Jet Transverse Energy

The mean subjet multiplicity $\langle n_{\rm sbj}\rangle$ for the dijet photoproduction sample is presented in Fig. 7.7 in three different $E_T^{\rm jet}$ bins. In the lowest $E_T^{\rm jet}$ bin, the data points are closer to the prediction for a pure sample of gluon jets, whereas in the highest $E_T^{\rm jet}$ bin, they tend to be much closer to the prediction for a pure sample of quark jets. The same tendency is also observed in the integrated jet shape (see Fig. 7.8): the data points become more 'quark-like' as the transverse jet energy increases. This is consistent with the smaller contribution from resolved processes at higher $E_T^{\rm jet}$, as already pointed out in Sec. 4.2.2, which leads to a smaller fraction of gluon jets in the final state at higher transverse jet energies. For the other jet in the D^*-tagged photoproduction sample, the data points are always consistent with the prediction for a pure sample of quark jets, as can be seen from the plots in Fig. 7.9 and 7.10. In the lowest $E_T^{\rm jet}$ bin, the difference between the prediction for a pure sample of gluon jets when implementing or not the multiple interaction in the Monte Carlo simulation is also shown.

Figure 7.11 shows the integrated jet shape (at a fixed value of $r/R = 0.3$) and the mean subjet multiplicity (at a fixed value of $y_{\rm cut} = 0.01$) as a function of $E_T^{\rm jet}$. One can clearly notice that, as expected from pQCD predictions, both observables show a dependence with $E_T^{\rm jet}$. The integrated jet shape $\Psi(r/R)$ increases faster with r for jets with a larger transverse energy, indicating that these jets are more collimated, and the mean subjet multiplicity $\langle n_{\rm sbj}\rangle$ is smaller for jets with a larger transverse energy. Since the number of subjets, at a given value of the resolution scale $y_{\rm cut}$, reflects the amount of relative transverse momentum with respect to

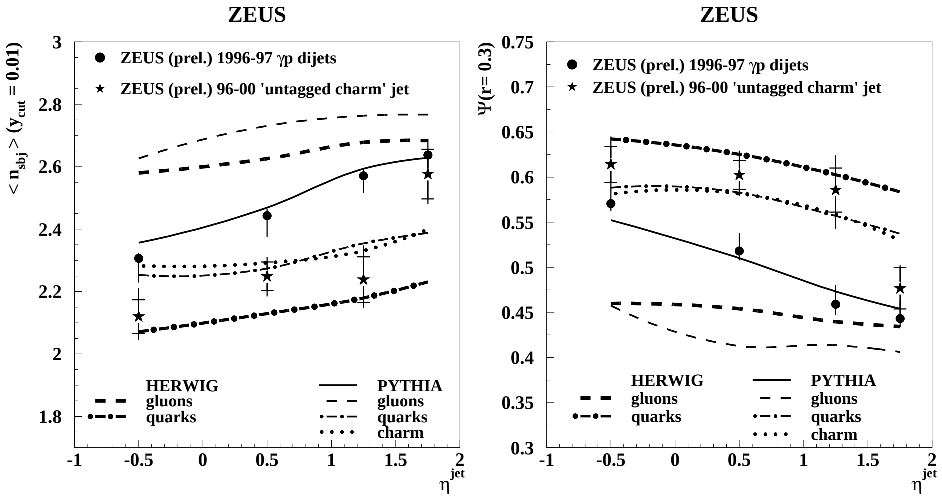

Figure 7.6: Mean subjet multiplicity (left plot) at a fixed value of $y_{\rm cut} = 0.01$ and integrated jet shape (right plot) at a fixed value of $r/R = 0.3$ as a function of $\eta^{\rm jet}$. The inner and outer error bars indicate the statistical errors and the statistical and systematic errors added in quadrature, respectively. For comparison, the prediction of the PYTHIA Monte Carlo including direct plus resolved processes are shown: all jets (solid line), quark jets (dash-dotted line), gluon jets (dashed line) and charm jets (dotted line). The predictions of the HERWIG Monte Carlo for quark and gluon jets are also shown (using thicker lines).

the jet axis, the observed $E_T^{\rm jet}$ dependence is consistent with the measurement of the jet shape.

7.2 Comparison with DIS

The investigation of the internal structure of jets has been performed at HERA also in deep-inelastic-scattering (DIS) at moderate and large squared four momentum transfers Q^2 [44]. The differential and integrated jet shapes have been measured for jets in the NC and CC reactions [45]:

$$e^+p \to e^+(\bar{\nu}) + {\rm jet} + X$$

with $Q^2 \geq 100$ GeV^2 by the ZEUS collaboration. The analysis has been carried out in the laboratory frame, where jets were required to have $E_T^{\rm jet} > 14$ GeV and $-1 < \eta^{\rm jet} < 2$. This means that, at the relatively large Q^2 scale chosen, mostly events (originating from the scattering process $eq \to eq$) with only one jet contribute to the DIS sample. The left plot in Fig. 7.12 shows the differential jet shape $\rho(r/R)$ for the other jet in the D^*-tagged photoproduction sample (measured for $E_T \geq 15$ GeV, with $\langle E_T\rangle \simeq 18$ GeV) compared with the differential jet shape for the NC DIS jets (measured for $14 \leq E_T \leq 21$ GeV, with $\langle E_T\rangle \simeq 17$ GeV). One can see that the two measurements are in agreement with each other. This is consistent with the expectation that at high Q^2 the DIS sample is dominated by quark initiated jets and

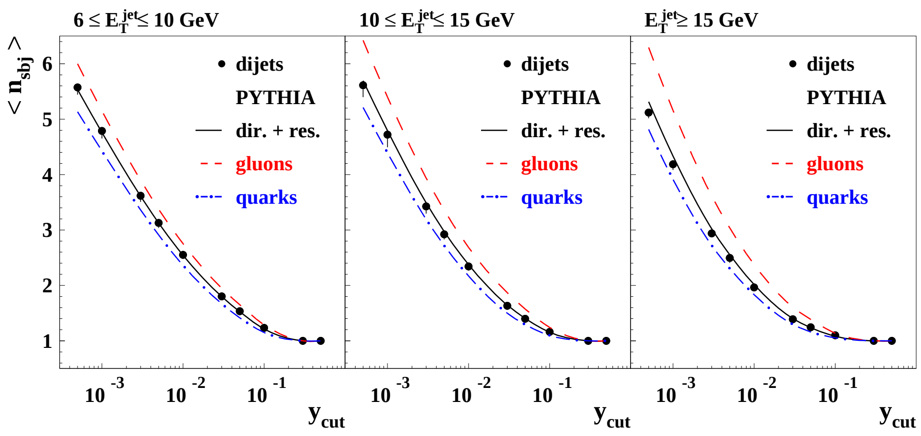

Figure 7.7: Mean subjet multiplicity $\langle n_{\rm sbj} \rangle$ for the dijet photoproduction sample with $-1 \leq \eta^{\rm jet} \leq 2$ in different $E_T^{\rm jet}$ bins. For comparison, the prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown.

that the outgoing parton evolution, and hence the resulting internal jet structure, is to a large extent independent of the particular hard scattering process in consideration. The comparison is however only qualitative since a different definition of the hadronic final state (definition (a) in Sec. 6.2.1) and a different jet algorithm (cone-type) were used in the analysis of the DIS

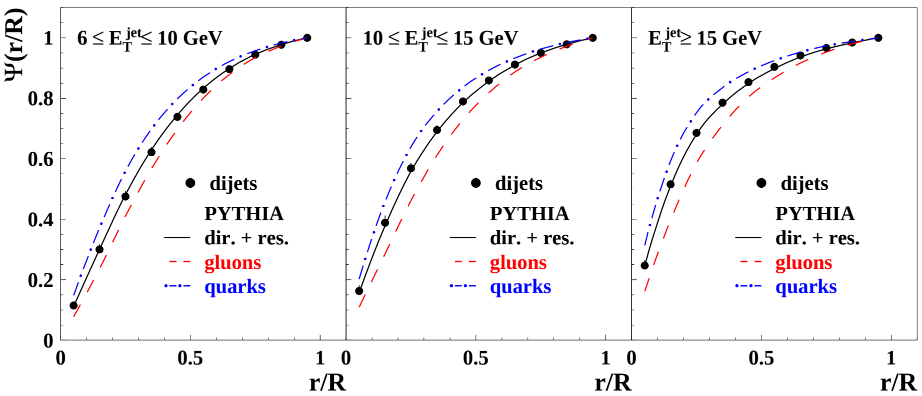

Figure 7.8: Integrated jet shape $\Psi(r/R)$ for the dijet photoproduction sample with $-1 \leq \eta^{\rm jet} \leq 2$ in different $E_T^{\rm jet}$ bins. For comparison, the prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown.

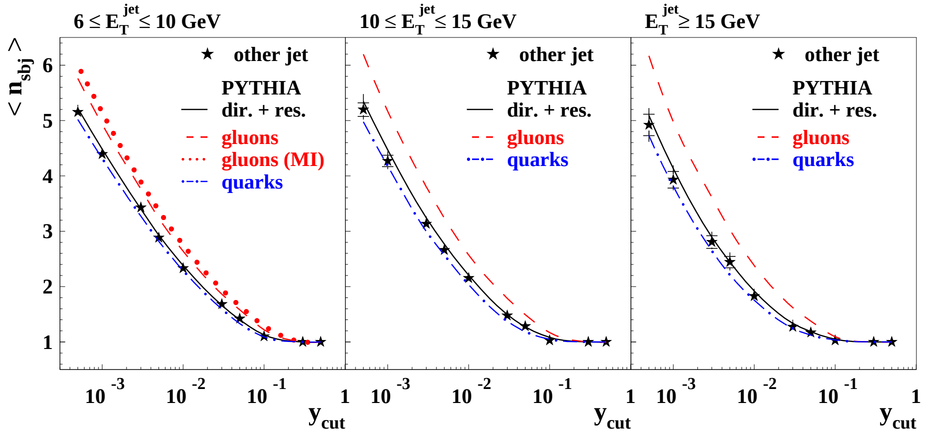

Figure 7.9: Mean subjet multiplicity $\langle n_{\rm sbj} \rangle$ for the other jet in the D^*-tagged dijet photoproduction sample with $-1 \leq \eta^{\rm jet} \leq 2$ in three different $E_T^{\rm jet}$ bins. The inner and outer error bars denote the statistical errors and the statistical and systematic errors added in quadrature, respectively. The prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown. For $6 \leq E_T^{\rm jet} \leq 10$ GeV, the prediction of PYTHIA for gluon jets with multiple interaction (MI) is also shown.

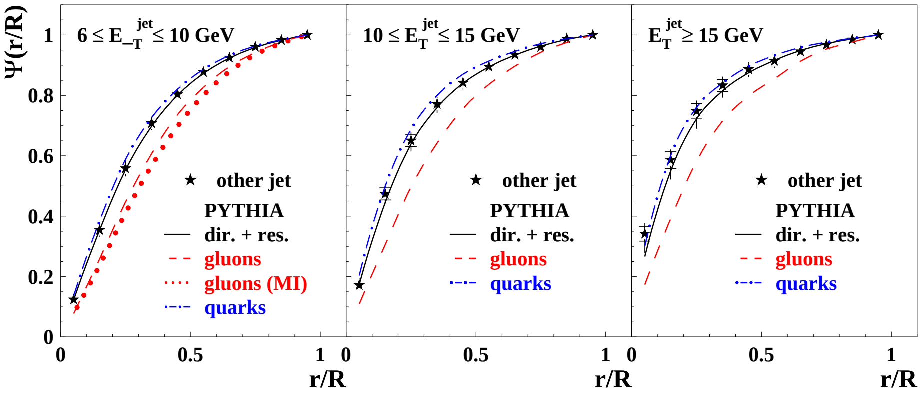

Figure 7.10: Integrated jet shape $\Psi(r/R)$ for the other jet in the D^*-tagged dijet photoproduction sample with $-1 \leq \eta^{\rm jet} \leq 2$ in three different $E_T^{\rm jet}$ bins. The inner and outer error bars denote the statistical errors and the statistical and systematic errors added in quadrature, respectively. The prediction of PYTHIA for resolved plus direct processes (solid line), for quark jets (dash-dotted line) and for gluon jets (dashed line) are shown. For $6 \leq E_T^{\rm jet} \leq 10$ GeV, the prediction of PYTHIA for gluon jets with multiple interaction (MI) is also shown.

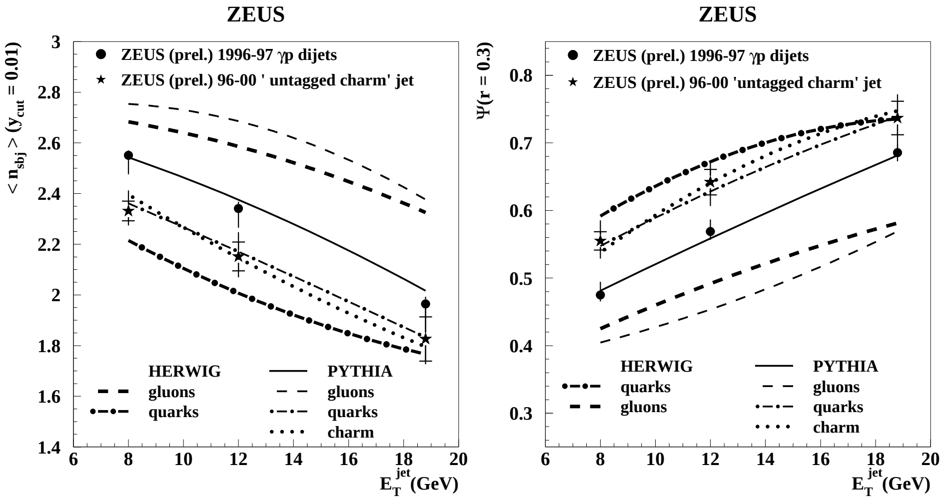

Figure 7.11: Mean subjet multiplicity (left plot) at a fixed value of $y_{\rm cut} = 0.01$ and integrated jet shape (right plot) at a fixed value of $r/R = 0.3$ as a function of $E_T^{\rm jet}$. The data points are located at the weighted mean in each $E_T^{\rm jet}$ region. The inner and outer error bars indicate the statistical errors and the statistical and systematic errors added in quadrature, respectively. For comparison, the prediction of the PYTHIA Monte Carlo including direct plus resolved processes are shown: all jets (solid line), quark jets (dash-dotted line), gluon jets (dashed line) and charm jets only (dotted line). The predictions of the HERWIG Monte Carlo for quark and gluon jets are also shown (using thicker lines).

jet sample. The right plot in Fig. 7.12 shows the mean subjet multiplicity for the other jet in the D^*-tagged photoproduction sample (measured for $E_T \geq 15$ GeV, with $\langle E_T \rangle \simeq 18$ GeV) compared with the mean subjet multiplicity for the NC DIS jets[1] (measured for $15 \leq E_T \leq 21$ GeV, with $\langle E_T \rangle \simeq 17$ GeV) [46]. In this case, the same definition of the hadronic final state and the same jet algorithm used in this analysis are applied. From the comparison, one can see that the data points are in very good agreement and differ by less than one standard deviation:

$$\Delta \equiv \frac{\langle n_{\rm sbj}\rangle^{\rm other\ jet} - \langle n_{\rm sbj}\rangle^{\rm DIS}}{\langle n_{\rm sbj}\rangle^{\rm DIS}} < 1 \cdot \sigma$$

Moreover, for resolution scales $y_{\rm cut} \geq 0.01$, where one expects the effect of the charm mass to be negligible, the data points show an even better agreement ($\Delta < 0.5\,\sigma$).

7.3 Extraction of the Jet Shape and Mean Subjet Multiplicity for Gluon Jets

In this section a simple method to extract the measurement of the internal structure of gluon jets, on a statistical basis, is presented.

[1]The cut on the squared four momentum transfers applied in this analysis is $Q^2 \geq 125$ GeV^2.

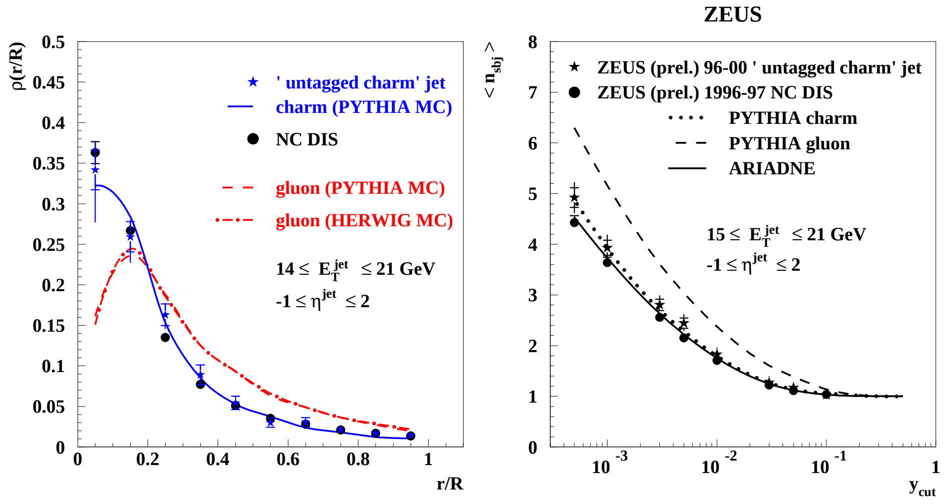

Figure 7.12: Differential jet shape $\rho(r/R)$ (left plot) and mean subjet multiplicity $\langle n_{\rm sbj}\rangle$ for the other jet in the D^*-tagged photoproduction sample compared with the same observables measured in NC DIS. In the left plot the DIS MC prediction is not shown.

Since the dijet photoproduction sample consists of a mixture of quark and gluon jets, the measured observables can be written as:

$$\mathcal{O}_{\rm dijet} = f_q \cdot \mathcal{O}_{\rm quark} + f_g \cdot \mathcal{O}_{\rm gluon} \tag{7.1}$$

where f_q and $f_g = (1 - f_q)$ are the fractions of quark and gluon jets in the sample and $\mathcal{O}_{\rm quark}$ and $\mathcal{O}_{\rm gluon}$ are the observables for a pure sample of quark and gluon jets, respectively. From what has been discussed in the previous sections, the measurements for the charm-enriched jet sample at high transverse energy ($E_T^{\rm jet} \geq 15$ GeV) can be considered as 'gauge' measurements for a pure sample of quark jets[2], i.e. $\mathcal{O}_{\rm charm} = \mathcal{O}_{\rm quark}$. Therefore by inverting the formula in Eq. 7.1 one can extract the values for a pure sample of gluon jets. The fractions f_q and f_g for the dijet photoproduction sample are taken from the re-weighted LO Monte Carlo. The fractions f_q and f_g predicted by the PYTHIA and HERWIG Monte Carlo models have been compared and found to be very similar:

$$f_q^{\rm PYTHIA} = 0.628 \quad \text{and} \quad f_q^{\rm HERWIG} = 0.633 \quad \text{for} \quad E_T^{\rm jet} \geq 15 \text{ GeV}$$

Since the method relies completely on the correct prediction for the fraction f_q, its value has been varied by $\pm 10\,\%$ and the variation on the extracted $\mathcal{O}_{\rm gluon}$ has been included as a systematic error on the measurements. For the mean subjet multiplicity the corresponding systematic error is always smaller than 5 %, whereas for the integrated jet shape it is smaller than 10% for $r/R \geq 0.2$. For $r/R = 0.1$ the variation of $\pm 10\%$ in the fraction of quark and gluon jets, results in a systematic error of the order of 40%. Figure 7.13 shows the extracted mean subjet multiplicity and integrated jet shape for the gluon jets, compared with the measurements

[2]In the case of the mean subjet multiplicity this is true only for values of the resolution scale at which the charm mass has no effect, i.e. $y_{\rm cut} \gtrsim (m_c/E_T^{\rm jet})^2$.

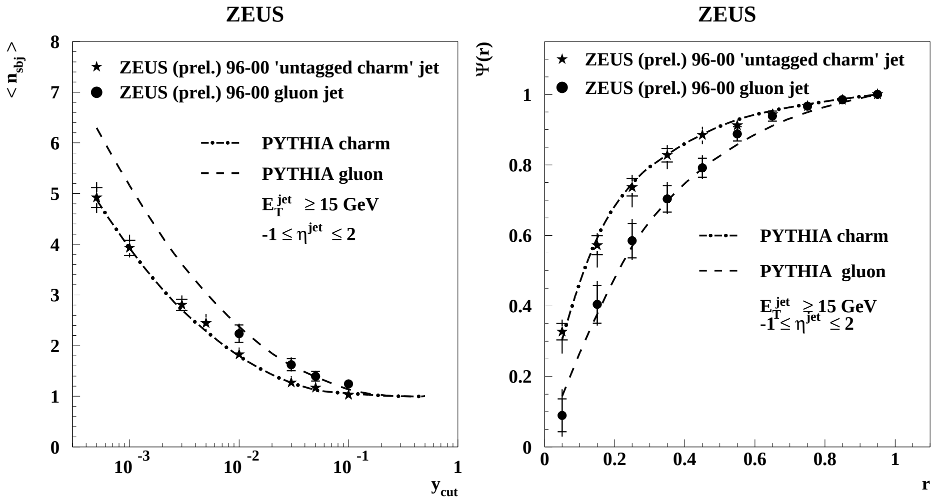

Figure 7.13: Mean subjet multiplicity $\langle n_{\mathrm{sbj}} \rangle (y_{\mathrm{cut}})$ (left plot) and integrated jet shape $\Psi(r/R)$ (right plot) for gluon jets with $E_T^{\mathrm{jet}} \geq 15$ GeV compared with the observables measured for the charm-enriched jet sample. The dotted line in the left plot indicates the 'threshold' value of the resolution scale y_{cut} from which the effect of the charm mass is estimated to be negligible.

for the charm-enriched jet sample and with the LO Monte Carlo predictions. One can see that the measurements are in very good agreement with the pQCD predictions. On a statistical level, the internal structure of gluon jets is quite different from that of quark jets: gluon jets have more subjets and are broader than quarks jets and the measured properties are consistent with the pQCD predictions.

Summary and Conclusions

In the first part of this thesis a measurement of the internal structure for charm (i.e. quark) initiated jets and for exclusive dijets in high energy γp interactions at HERA has been presented. The measurements are performed at a γp centre-of-mass energy of $134 \leq W \leq 277$ GeV in the photoproduction regime which corresponds to the kinematic region defined by $Q^2 \lesssim 1$ GeV^2.

The mean subjet multiplicity, $\langle n_{\mathrm{sbj}} \rangle (y_{\mathrm{cut}})$, and the differential and integrated jet shape, $\rho(r/R)$ and $\Psi(r/R)$, have been measured using the data collected with the ZEUS detector during the 1996-2000 e^+p running periods, corresponding to an integrated luminosity of 106 pb^{-1}.

Dijets with transverse energy $E_T^{\mathrm{jet}} \geq 7,6$ GeV and pseudorapidity $-1 \leq \eta^{\mathrm{jet}} \leq 2$ are defined in the laboratory frame using the longitudinally invariant k_T cluster algorithm.

A charm-enriched jet sample has been obtained by selecting dijet events where at least one of the two jets is initiated by a charm quark. The latter is identified via the measurement of a D^* meson in the final state in the following decay channel:

$$\begin{aligned} D^{*+} \longrightarrow\ & D^0 + \pi_s^+ \\ & \hookrightarrow K^- + \pi^+ \end{aligned} \tag{7.2}$$

and its charge-conjugate channel. The tight kinematical constraints of this decay imply that the mass difference $\Delta m \equiv (m_{D^{*+}} - m_{D^0})$ can be measured more accurately than the mass of the D^* meson itself. Plotting the Δm distribution yields a prominent signal over a relatively small combinatorial background. The D^0 meson decay channel with the lowest multiplicity has been chosen in order to minimize the contribution from combinatorial background. The kinematic region of the measured D^* meson candidate is restricted to:

$$p_T(D^*) \geq 3.0\,\mathrm{GeV} \quad \text{and} \quad |\eta(D^*)| \leq 1.5$$

The limits on the selected η^{jet} region are dictated by the requirement to have well reconstructed tracks in the ZEUS central tracking detector. The D^* meson is associated with the closest jet in the ϕ space. The internal structure of the second jet (referred to as the 'other jet'), which in most of the events also originates from a charm (or anti-charm) quark, is investigated and the combinatorial background associated with the signal is statistically subtracted.

The dependence of the measured jet substructure on E_T^{jet} and η^{jet} has been studied and leading-logarithm parton-shower Monte Carlo calculations for direct and resolved photoproduction processes, including multi parton interactions, have been compared to the measured observables. Good agreement with the data has been observed.

For both the dijet photoproduction and the charm-enriched jet sample, the mean subjet multiplicity decreases and the jet shape narrows as E_T^{jet} increases, indicating that jets with larger E_T^{jet} become more collimated. This result is in agreement with the pQCD predictions, as implemented in the MC programs.

For the dijet photoproduction sample, it is also observed that the mean subjet multiplicity increases and the jet shape broadens as $\eta^{\rm jet}$ increases. For $\eta^{\rm jet} \leq 0$, the data points are very close to the prediction for a pure sample of quark jets whereas for $\eta^{\rm jet} \geq 1.5$ the data points are very close to the predictions for a pure sample of gluon jets. This effect has already been observed in previous analyses [47], [48] where it was pointed out that the observed broadening of the jet shape and the increase in the mean subjet multiplicity as $\eta^{\rm jet}$ increases are consistent with the predicted increase in the fraction of gluon initiated jets with $\eta^{\rm jet}$. Since the $\eta^{\rm jet}$ dependence of the internal jet structure for either quark or gluon initiated jets, as predicted by LO Monte Carlo models, is very small, the measurement of the internal structure of the charm-enriched jet sample can be used to disentangle the quark and gluon contribution in the photoproduction processes.

The measurements performed for the charm-enriched jet sample are consistent with the Monte Carlo prediction for a pure sample of charm jets in all $E_T^{\rm jet}$ regions (integrated over $\eta^{\rm jet}$) and for $-1 \leq \eta^{\rm jet} \leq 1.5$ (integrated over $E_T^{\rm jet} \geq 7, 6$ GeV). In the highest $\eta^{\rm jet}$ bin, however, the data points show a deviation from the prediction of a pure sample of charm jets. According to LO Monte Carlo predictions, part of the charm production in resolved processes originates from so called 'charm excitation' processes in which only one of the high $E_T^{\rm jet}$ jets in the final state contains a charmed hadron, whereas the second charmed hadron is carried off in the photon (or proton) remnant. Since the contributions from charm excitation is expected to be larger in the forward direction (large values of $\eta^{\rm jet}$), the deviation seen in the charm-enriched jet sample could be explained by a non-negligible gluon fraction in the charm-enriched jet sample.

A direct-enriched D^*-tagged sample, i.e. $x_\gamma^{\rm OBS} \geq 0.75$ where the boson gluon fusion $\gamma g \rightarrow c\bar{c}$ results to be the dominant process and thus both jets are charm initiated, has been separately analysed. The jet substructure for this subsample is found to be in agreement with the Monte Carlo prediction for a highly charm-enriched jet sample in the full $\eta^{\rm jet}$ range. Thus the comparison between the measurements for the direct-enriched D^*-tagged sample and the total D^*-tagged sample supports the interpretation that the deviation from the prediction for a pure sample of charm jets in the highest $\eta^{\rm jet}$ bin could be due to the higher gluon contamination to the charm-enriched jet sample.

The internal structure of gluon jets has been extracted with a simple statistical method using the measurements for the inclusive (i.e. without flavour tagging) dijet and the charm-enriched jet samples. For jets of the same transverse energy, it is observed that gluon jets have a higher mean subjet multiplicity and are broader than quark jets and the results agree within errors with the theoretical pQCD predictions.

In conclusions, for the first time at HERA, the internal structure of charm initiated jets has been investigated and an indirect measurement for a pure sample of gluon initiated jets has been performed.

Part II

Outlook

Chapter 8

The ZEUS Microvertex Detector (MVD)

The HERA ring is currently undergoing a significant upgrade, which aims for an increase of the $e^{\pm}p$ peak luminosity from 1.5 to 7.0×10^{31} cm^{-2}s^{-1} and longitudinal $e^{\pm}$ polarization for the two collider experiments (H1 and ZEUS) as well as for the fixed target experiment HERMES [20]. At the same time, a silicon microvertex detector (MVD) and, in the forward direction, a straw-tube-tracker (STT) will be installed in the ZEUS experiment, which will significantly improve the tracking capabilities. The coverage of the existing central tracking detector will be substantially increased in the forward direction, thus enhancing the acceptance for high-mass and high Q^2 physics. Moreover, the improved tracking will allow the reconstruction of secondary vertices, leading to the measurement of heavy quark production with reduced background. In this chapter the layout of the ZEUS microvertex detector will be presented.

8.1 The ZEUS MVD Design

The MVD design is based on the following requirements:

- polar angular coverage between $\sim 10°$-$150°$
- measurements of three points per track, in two projections each
- hit efficiency $> 97\%$
- $< 20\mu$m intrinsic hit resolution for normal incident tracks
- two track separation down to 240 μm
- noise occupancy $< 10^{-3}$

It has to fit into the 324 mm diameter hole left by the existing central tracking detector (CTD). Figure 8.1 shows the schematic layout of the MVD. For the fast handling of the signals from the silicon sensors, the readout electronics, and therefore also the cooling pipes and electrical cables, are located near the sensors.

The detector is divided into two sections:

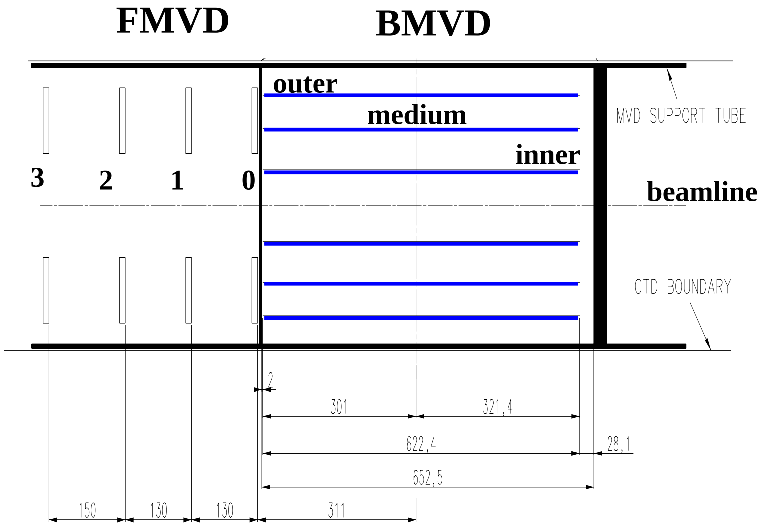

Figure 8.1: Schematic layout of the MVD (along the beam direction). The horizontal and vertical scales are different. Shown are the four wheels (0, 1, 2, 3) of the FMVD and the three layers (outer, medium, inner) parallel to the beamline of the BMVD.

- a $\sim$ 60 cm long barrel part (BMVD), which consists of three double layers of silicon strip sensors, arranged in concentric planes around the interaction point, with a polar coverage between 30° and 150°. A cross section of the BMVD in the r-ϕ plane is shown in Fig. 8.2, where the elliptical beam pipe is also visible. The shape of the beam pipe takes into account the background present in the interaction region. The greatest contribution comes in fact from direct and back-scattered synchrotron radiation emitted mainly in the horizontal plane by the electron beam. The first BMVD layer has been placed as close as possible to the beam pipe, but due to the limited space it cannot surround it completely[1]. Therefore $\sim$ 25% of the azimuthal angle is covered by only two cylinders. The longitudinal dimension of the BMVD is constrained by the spread of the interaction region, which is determined by the proton bunch length ($\sim$ 10 cm). The BMVD sensors are square shaped (64.2×64.2 mm^2) with a constant strip length of 62.4 mm. Two silicon sensors, with strips perpendicular to each other, are glued together using epoxy-resin and polyimide spacers, and are electrically interconnected via copper traces etched on a 35 μm thick kapton foil, as shown in Fig. 8.3. The structure obtained is referred to as a 'half-module'. Two half modules (one the mirror image of the other) are mounted on top of each other to form a module or readout cell (see Fig. 8.3). Five modules are fixed side by side, with 1 mm gap in between, on a support frame ('ladder'). The ladder provides also the support for readout electronics, cabling and cooling for the front-end electronics

[1]In order to optimize the impact parameter resolution, the first hit measurement has to be performed as close as possible to the primary vertex.

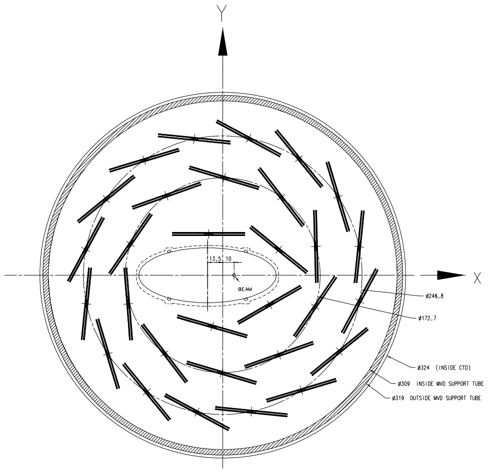

Figure 8.2: Cross section of the BMVD in the r-ϕ plane. The three silicon sensors layers around the elliptical beampipe can be seen.

(see Fig. 8.4). Table 8.1 summarizes the BMVD geometrical details.

- a forward part (FMVD), which consists of four double layers of silicon strip sensors, arranged around the beampipe as shown in Fig 8.5, extending the polar coverage down to 7°. The two layers of each wheel are mounted back to back on a support structure with $\sim$ 8 mm separation in the Z direction and the forward wheels are positioned at Z $\simeq$ 31, 44, 57, 70 cm from the nominal interaction point. The FMVD sensors are trapezoidal shaped with two sided tilted by 13° in opposite direction (see Fig. 8.6) and have a variable strip length (from 0.6 to 73.3 mm). The shape and positioning of the beampipe requires the use of shorter sensors (in the radial direction) in some locations. Table 8.2 summarizes the FMVD geometrical details.

	inner layer	**medium layer**	**outer layer**	**total**
# of ladders	4	10	16	30
# of silicon sensors	80	200	320	600
# of readout cells	40	100	160	300
# of readout channels	20480	51200	81920	153600
# of readout chips	160	400	640	1200

Table 8.1: Statistics on the components of the BMVD

8.2 The MVD Silicon Sensor Layout

The MVD detectors are made of high resistivity ($\sim$ 5 kΩcm) n-type silicon and have a thickness of $\sim$ 330 μm. The backside of the detector is n$^+$ doped and aluminised whereas the front side

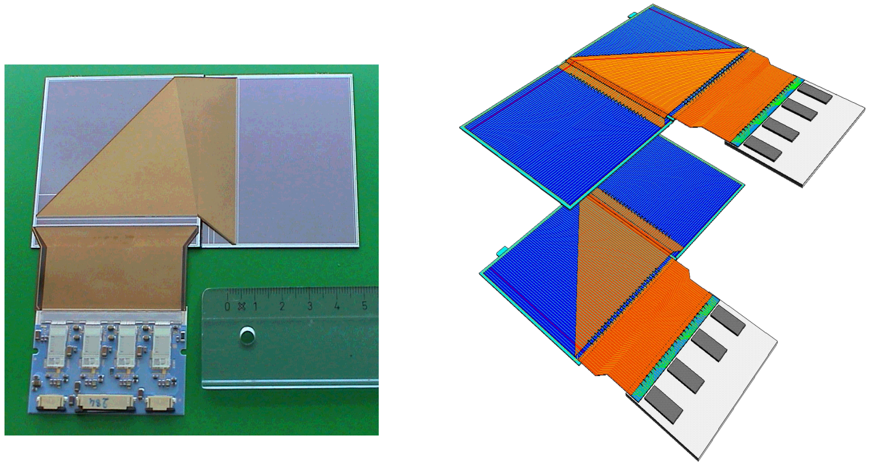

Figure 8.3: Photo of a half-module (left), where the kapton interconnection between the two sensors and the connection to the hybrid (housing 4 HELIX3.0 readout chips) can be seen. On the right a schematic drawing shows the geometry of a full readout cell.

is covered with p^+ doped strips. Every sixth strip (14 μm wide) is AC-coupled by a SiO_2-Si_3N_4 dielectric ($\sim$ 250 nm thick) to an aluminium readout strip, resulting in a capacitive charge division [49] with five intermediate strips (each 12 μm wide). Capacitive charge division between readout strips is used in order to limit the number of readout channels still allowing to achieve a spatial hit resolution of $\sim$ 10-15 μm for particles traversing the detector at small

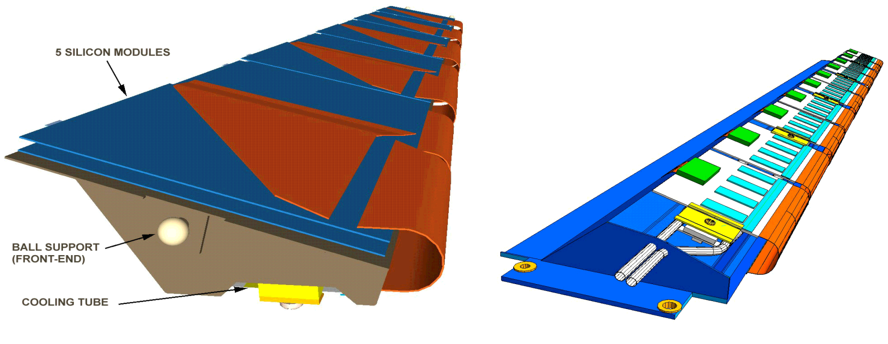

Figure 8.4: Front view of a BMVD ladder with the five silicon modules mounted on it (left) and back view of a BMVD ladder (right) where the hybrids and the cooling pipes can be seen.

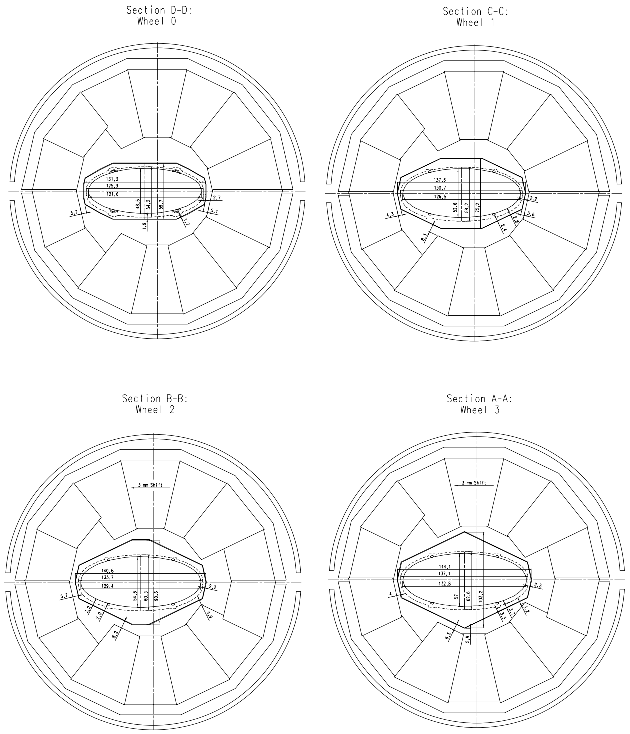

Figure 8.5: Cross section of the FMVD in the r-ϕ plane showing the arrangements of the silicon sensors in the forward wheels. Wheel #0 is nearest to the BMVD. The shielding of the beampipe is also shown.

incidence angles ($\lesssim 30^\circ$). The strip and readout pitches are 20 μm and 120 μm respectively.

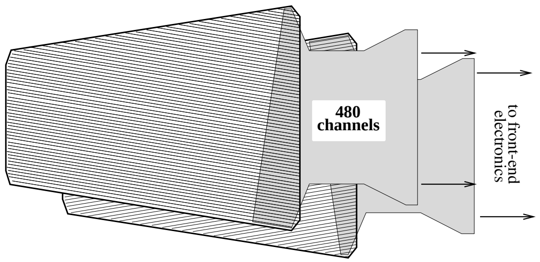

Figure 8.6: Schematic drawing of two FMVD silicon sensors.

	FMVD
outer radius	∼ 132 mm
inner radius	∼ 60 mm
wedge angle	∼ 13°
# of wheels	4
# of sensors/wheel	2×14
# of readout strips/sensor	480
# of readout channels	53760
# of readout chips	448

Table 8.2: Statistics on the components of the FMVD

Figure 8.7 shows a cross section (not to scale) of the MVD sensor, where the different layers can be seen. Polysilicon resistors are used to connect all the p^+ strips to the common voltage supply line. They alternately connect even and odd p^+ strips to either of the two bias lines, located perpendicular to the strips close to the detector edges. The first and last readout strip are directly connected to the bias line, completing the bias ring around the sensitive area (62.2×61.4 mm^2 for the BMVD sensors) which is surrounded by a triple guard-ring structure (see Fig. 8.8). An additional n^+ doped band is placed beyond the last guard ring. It allows to bias the backplane with a contact from the top. The detectors have been produced by Hamamatsu Photonics (Japan). In November 1998 the first MVD prototypes have been delivered and extensively tested to verify that the detector specifications were fulfilled. In the next chapter the results of the extensive testbeam program on the MVD prototypes are presented whereas the electrical measurements performed on the detectors themselves and on test-structures are described elsewhere [50].

For the understanding of the measurements described in the next chapter it is helpful to have in mind the following concepts. A silicon strip sensor is essentially obtained by dividing a p-n junction into small segments that act as independent electrodes[2]. Its **depletion region** is

[2] A very short review of the basic principles of a p-n junction is reported in App.C. For a complete treatment refer to [51]

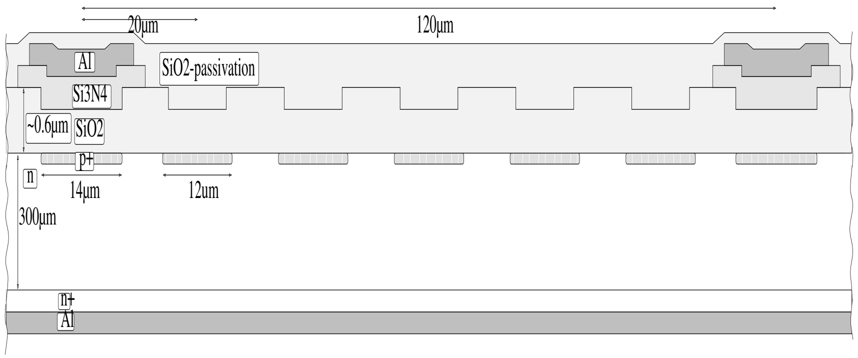

Figure 8.7: Cross section of the MVD silicon sensor (not to scale) showing the region between two readout strips.

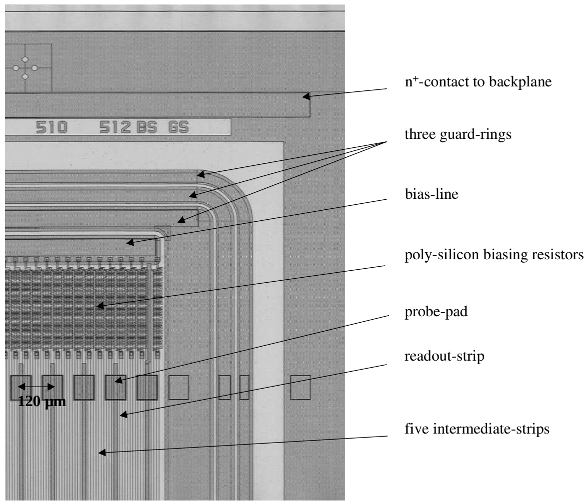

Figure 8.8: Mask of one corner of the BMVD silicon sensor.

the sensitive volume from which electron and hole pairs from ionizing radiation are collected, since only there an electric field is present. When a charged particle traverses the depletion region in the silicon bulk, electron-hole pairs are generated along its ionization path in a very narrow tube. In the electrical field, the holes and electrons drift towards the p^+ and n^+ sides, respectively, and cause a change in the induced surface charges, thus creating a short current pulse on the electrodes. For a 300 μm thick fully depleted silicon sensor, the most probable energy loss is $\Delta E \sim 88$ keV, corresponding to a signal charge $Q_{\rm coll}$ of $\sim$ 24600 electron-hole pairs.
The reverse bias voltage, $V_{\rm dep}$, necessary to extend the depletion region over the entire detector bulk is approximately (ignoring effects of the segmentation of the p^+-electrode) given by[3]:

$$V_{\rm dep} = \frac{t^2}{2\,\epsilon_{\rm Si}\,\rho\,\mu_e} \tag{8.1}$$

where t is the detector thickness, $\epsilon_{\rm Si}$ is the silicon dielectric constant, ρ is the resistivity of the doped silicon and μ_e is the electron mobility. Because of its electrical configuration, the depletion region has also a certain dynamic capacitance $C_{\rm junc} = dQ/dV$, given by [51]:

$$C_{\rm junc} = \epsilon_{\rm Si} \cdot \frac{A}{W(V_{\rm bias})} \tag{8.2}$$

where A is the area of the depletion layer, $W(V_{\rm bias})$ is the width of the depletion region and $V_{\rm bias}$ is the applied reverse bias voltage. This means that the detector behaves like a parallel plate capacitor with the two electrodes at a distance W(V). Taking into account the voltage

[3] This formula is only valid for n-type detectors.

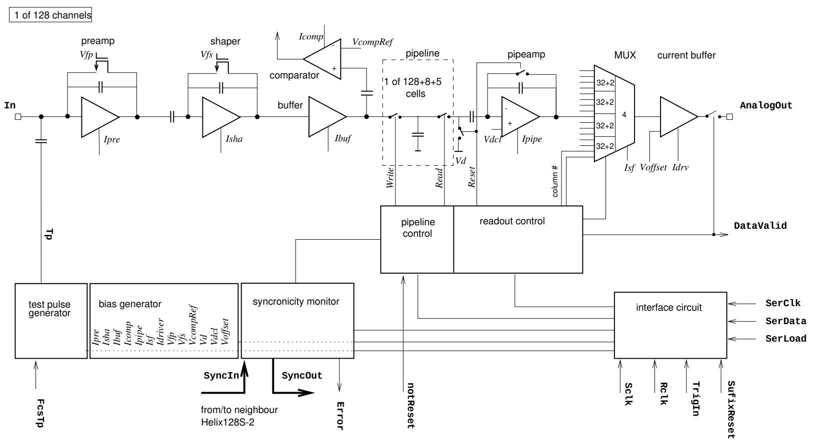

Figure 8.9: Schematic layout of the HELIX analog readout chip

dependence of the depletion depth (see Eq. C.13), one can write[4]

$$C_{junc}(V_{bias}) = A \cdot \sqrt{\frac{\epsilon_{Si} q N_D}{2 \cdot (V_{bias} + V_{bi})}} \quad \text{for} \quad V_{bias} \le V_{dep} \tag{8.3}$$

$$C = C_{geom} = \epsilon_{Si} \cdot \frac{A}{t} \quad \text{for} \quad V_{bias} > V_{dep}$$

8.3 Readout Electronics for the MVD Detector

The MVD silicon sensors are readout by HELIX3.0[5] chip [52], a successor of an analog chip designed by the ASIC laboratory at Heidelberg (modified for ZEUS in collaboration with NIKHEF) and manufactured in the 0.8 μm CMOS process by AMS [53]. Figure 8.9 shows a schematic of the HELIX chip. The chip integrates 128 channels with a charge sensitive pre-amplifier followed by a shaper forming a semi Gaussian pulse with a peak time of ~50-70 ns. The signals are sampled in an analog pipeline capable of storing up to 136 event measurements to await the GFL trigger decision (see Sec. 2.2.4). The 128 channels of a pipeline column, selected by a trigger signal, are transferred via a multiplexer to a serial output readout line. The bias voltages and current for the different amplifier stages can be adjusted via programming of digital-to-analog (DAC) registers in order to select the most appropriate operation point and to compensate for radiation damages. An important feature of the HELIX is the capability to synchronize the data transfer of several chips on the serial output line using a daisy-chain mechanism (**token ring**). The major modification introduced in the HELIX3.0 version of the chip is the implementation of the failsafe token which allows to exclude a bad chip from the readout chain, without hindering the functioning of the remaining ones in the chain. Four HELIX chips are mounted on a ceramic hybrid and two hybrids, equipping one BMVD module and two wheel sensors, are readout as a unit, having in common all the power lines, control signals and output analog line.

[4] In the first formula, N_D indicates the 'donor' concentration and V_{bi} is the 'built-in' potential difference established in equilibrium between the n-type and the p-type material.

[5] The HELIX2.2 version has been used for the beam tests on the MVD prototypes (see chapter 9).

Chapter 9

MVD Prototypes and Beam Test Results

The production of the MVD has been accompanied by an extensive test program, including testbeam measurements. The aim of the testbeam program was to measure the general performance of the prototype detectors and its readout electronics:

- detection and charge collection efficiency
- noise level and signal over noise ratio (S/N)
- position resolution as a function of the angle of the incident particle θ (across the detector bulk) and of the pulse height
- performances after photon and neutron irradiation
- performances of the half-modules and eventually full readout modules

In this chapter measurements performed with single BMVD prototype detectors readout by HELIX2.2 chips (with a clock rate of 1 MHz) are presented.

9.1 Testbeam Setup at DESY

The test measurements were performed with the DESY-II testbeam, which is a parasitic electron beam obtained after two conversions. A 10 μm thick carbonfiber target in the machine intercepts the beam and produces a bremsstrahlung beam which, by means of a 0.1 X_0[1] thick copper target, is converted into electron-positron pairs. A bending magnet together with a collimator slit, used to select the particle momentum, delivers the beam into the experimental hall. Depending on the primary use of DESY-II the maximal momentum varies between 4.3 and 7.5 GeV. Most measurements were done at 3 and 6 GeV resulting in a trigger rate of ~ 10 Hz and ~ 2 Hz, respectively, limited by the data acquisition system.

In order to investigate the performances of the device under test (DUT), the position information of the traversing particle has to be known with a resolution better than that expected

[1] X_0 is the radiation length.

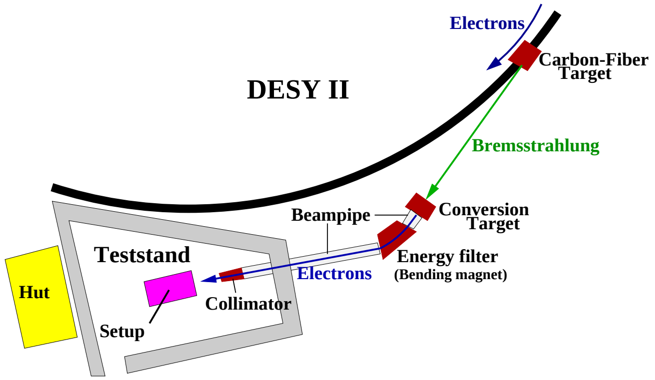

Figure 9.1: Schematic drawing of the test-beam area.

from the tested detector. A reference telescope has been assembled to allow a precise determination of the particle impact point on the MVD prototypes. The silicon telescope consists of three modules which are mounted on an optical bench, placed on a metal table (see figures 9.2, 9.3). The modules can be mounted at different positions along the beam axis (Z), by means of screws and fixation holes. The DUT is mounted between two telescope modules on linear and rotational positioners which allow to investigate the performances in different areas of the detector and for different incidence angle of the beam particles. Two small plastic scintillators, defining a surface in the (X,Y) plane of 9×9 mm^2, are located at each end of the telescope and their coincidence signals are used for triggering purpose.

The data acquisition system (DAQ) is handled by a Sun workstation mounted on a VME module which controls the initialization and operation of the readout chips, the pattern generators (sequencers), the analog-to-digital converters (ADCs), the GPIB-interface for the positioners and the data taking (see Fig. 9.4). The DAQ runs under LabVIEW [54], a software application which makes use of a graphical, object-oriented programming language.

9.1.1 Silicon Telescope

The telescope modules are a recent version of a CERN development [55]. Each of them consists of two high resolution single-sided strip sensors, mounted close to each other with perpendicular strip orientation, and measures two orthogonal coordinates (X,Y). Both detectors are mounted in an electrical shielded metal box with thin aluminium windows. The AC coupled sensors are square shaped (32×32 mm^2) and have a thickness of $\sim$ 300 μm. Their strip pitch is 25 μm while the readout pitch is 50 μm. The sensors are fully depleted at a bias voltage of $\sim$ 50 V and the average leakage current is of the order of 1.5-2.0 nA per strip.

The 640 readout strips on each sensor are read out by five VA2 chips [56]. The VA2 chips have been built in the 1.2 μm CMOS technology and contain 128 channels. Each channel has a charge sensitive pre-amplifier followed by a CR-RC shaper (set at 2 μs peaking time) and a sample-and-hold circuitry. Data from each channel are put serially in an output buffer through an on-chip multiplexer (see Fig. 9.5). The analog signals are digitized by CAEN V550 flash ADCs (with 10 bits resolution and a maximum conversion rate of 5 MHz) which are

Figure 9.2: Photo of the test-beam setup at DESY. The electron beam comes from the right. The DUT is contained inside the cylindrical box, which is placed on linear and rotational positioners and mounted on the optical bench between two of the three telescope modules.

controlled by the CAEN V551 sequencer [57]. Figure 9.6 shows the energy loss distribution for

	VA2
Equivalent Noise Charge @ 2 μs	135 e$^-$+12.3 e$^-$/pF
Power Consumption	1.2 mW/channel
Peaking Time	0.5-3.0 μs
Dynamic Range	$\pm$ 4 MIPs
Typical Applications	medium sized detectors (cap. load $\sim$ 10 pF)

Table 9.1: Technical details of telescope readout VA2 chips.

the telescope detectors[2].

9.2 MVD Data Analysis

For each event, the voltage outputs from all the readout channels of the DUT as well as the relevant information from the silicon telescope are stored into files in the form of digitized ADC counts.

The measured ADC value is a convolution of several contributions which have to be estimated in order to extract the signal caused by a traversing particle. The raw signal PH$_i^k$, where i and k denote the event and the strip number respectively, can be expressed as:

$$PH_i^k = S_i^k + P^k + N_i^k + C_i \tag{9.1}$$

[2]The analysis procedure for the telescope detectors is very similar to the one used for the MVD sensors and will be discussed in Sec. 9.2 and further.

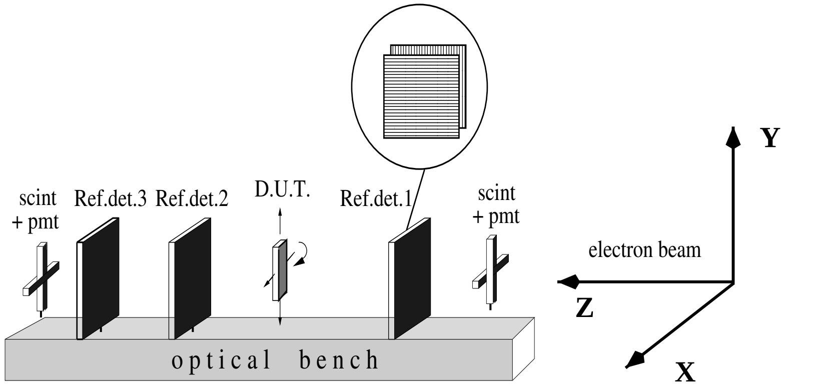

Figure 9.3: Schematic layout of the test-beam setup.

where:

- S_i^k is the Landau distributed signal due to the energy loss of the traversing particle

- P^k is a DC offset on the voltage output, usually referred to as pedestal level; it has to be calculated for each channel separately because large variations, due to inhomogeneities in the front-end electronics, may occur. The pedestal level may also slowly vary with time ($\gtrsim \mathcal{O}(10^1)$ hour), for example, because of temperature changes, requiring a dedicated monitoring during data taking. The pedestal levels for all the channels of a DUT prototype are shown in Fig. 9.7

- N_i^k is the individual random noise level of each channel. It is Gaussian distributed with mean zero, roughly independent of the neighbouring channel noise, uncorrelated between events and cannot be deconvoluted from the particle signal S_i^k. Nevertheless, the spread (rms) of its distribution, N^k_{noise}, can be calculated from the raw data. N^k_{noise} is in general rather uniform within one chip, but because of detector and/or front-end chip faults, some channels may have a much higher noise than the average chip level. The noise level N^k_{noise} for all the channels of a DUT prototype is shown in Fig. 9.7

- C_i is a positive or negative voltage offset common to all channels of one readout chip, usually referred to as common mode noise (CMN). Its variations from event to event are mainly caused by low frequency pick-up on the power supply lines. C_i is uncorrelated between events and obeys in good approximation a Gaussian distribution.

9.2.1 Pedestal and Noise Calculations

To obtain the pedestal and noise levels, the first N events (typically between 100 and 200) at the beginning of the data-taking are selected with the electron beam shutter closed using a random trigger (i.e. S_i^k=0).

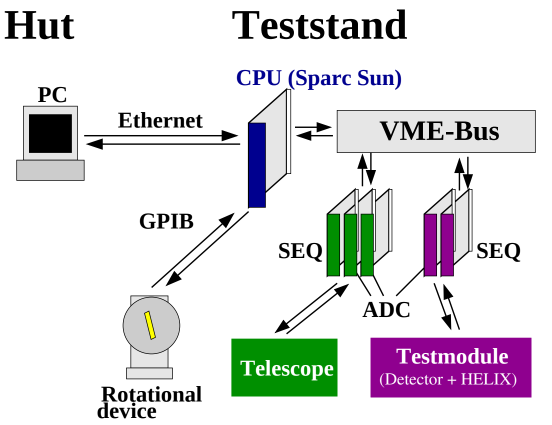

Figure 9.4: Schematic representation of the acquisition system.

First the CMN contribution for each readout chip separately is calculated:

$$C_i = \frac{1}{N_{\text{channel}}} \cdot \sum_{k=1}^{N_{\text{channel}}} PH_i^k \tag{9.2}$$

where N_{channel} is the number of readout channels per chip[3].
The pedestal and noise levels are then given by:

$$P^k = \frac{1}{N_{\text{event}}} \sum_{i=1}^{N_{\text{event}}} (PH_i^k - C_i) \tag{9.3}$$

$$N^k_{\text{noise}} = \frac{1}{N_{\text{event}} - 1} \cdot \sqrt{\sum_{i=1}^{N_{\text{event}}} (P^k - PH_i^k - C_i)^2} \tag{9.4}$$

[3]For both the reference detectors and the MVD prototypes $N_{\text{channel}} = 128$.

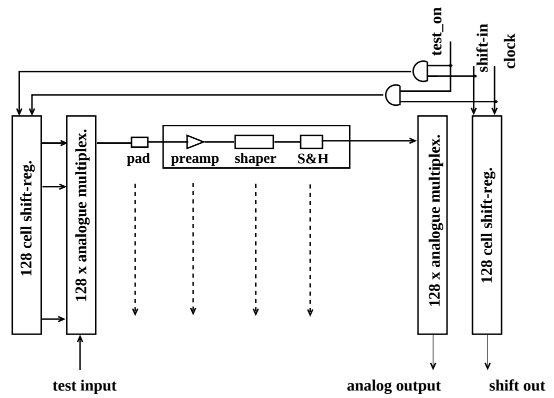

Figure 9.5: Schematic block diagram of a VA circuit.

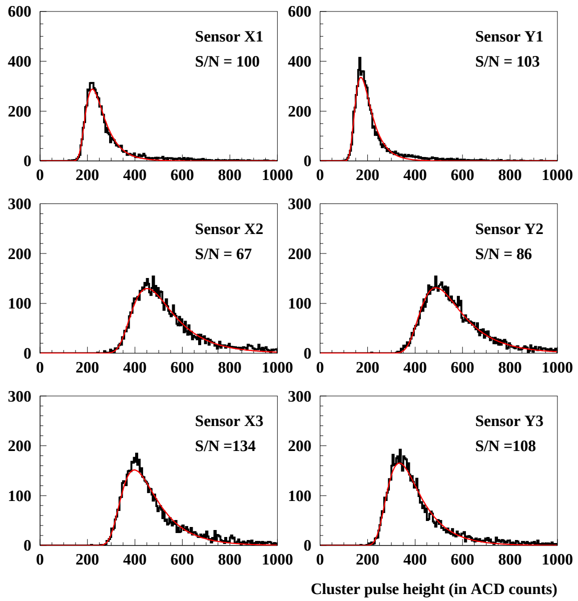

Figure 9.6: Pulse height distribution for the telescope detectors fitted by a Landau curve. The signal over noise ratio S/N for the six sensors is also indicated. Note that for technical reasons, the ADC full scale voltage range for the first two sensors (X1,Y1) is smaller (300 mV) than that used for the other four sensors (750 mV).

For the telescope detectors, pedestal and noise levels are calculated online in the LabVIEW DAQ program, neglecting the contribution of the CMN, and their values are stored into an initialization file. Before starting the beam data taking, these pedestal and noise levels are read-in from the file, a threshold level T^k for each channel is calculated:

$$T^k = P^k + 3 \cdot N^k_{\text{noise}} \tag{9.5}$$

and pedestal and threshold levels for the telescope detectors are stored into memories on the CAEN V550 ADC modules. These values are then used for the online pedestal subtraction and eventual zero suppression[4] performed directly by the CAEN ADC modules, in order to reduce the size of the output data files. For the DUT, instead, the data from all the readout channels is recorded with no online preselection and CMN, pedestal, noise and threshold levels are calculated offline.

During beam data-taking, for the CMN determination, an iterative procedure is used to exclude channels with a real signal ($S^k_i \neq 0$). First, for all the channels of one readout chip the following

[4]Zero suppression means that, once the analog signal from a readout channel had been sampled by the ADC, its digital value is compared with the threshold level for that specific channel. If the channel value is above threshold, then it is stored into the local FIFO memory and afterwards recorded into file.

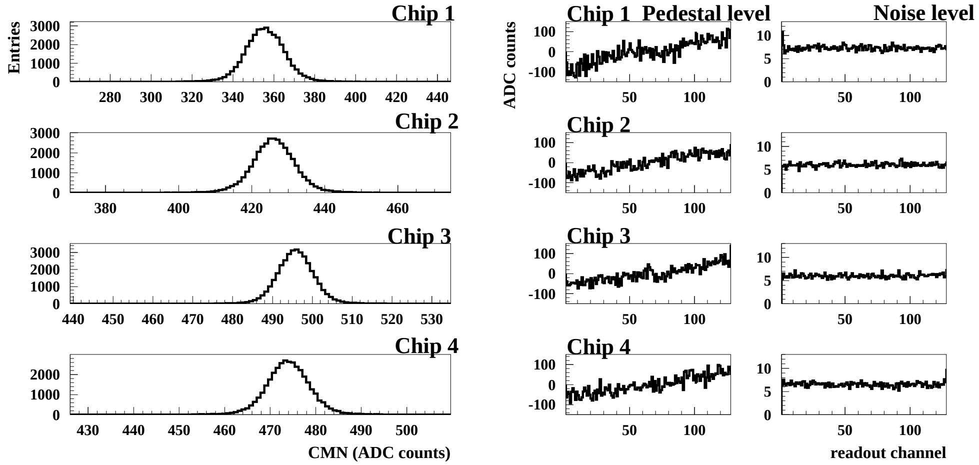

Figure 9.7: Common mode noise distribution (left) and pedestal and noise levels (right) for the DUT.

quantities are calculated:

$$M_i = \frac{1}{N_{\text{channel}}} \cdot \sum_{k=1}^{N_{\text{channel}}} A^k_i \quad \text{and} \quad V_i = \frac{1}{N_{\text{channel}} - 1} \cdot \sqrt{\sum_{k=1}^{N_{\text{channel}}} (A^k_i - M_i)^2} \quad \text{where} \quad A^k_i = PH^k_i - P^k \tag{9.6}$$

Then, readout channels with $(A^k_i - M_i) > 2 \cdot V_i$ are excluded and the others are used to determine C_i for each chip. Figure 9.7 shows the CMN distribution and the pedestal and noise levels separately for the 4 readout chips connected to a DUT. The common mode noise is not too high and the noise level is found to be constant throughout the whole detector.

9.3 Energy Loss Distribution at 0° Incidence Angle

The signal S^k_i (including any noise fluctuation N^k_i) for each strip is calculated from the raw data as:

$$S^k_i = PH^k_i - P^k - C_i \tag{9.7}$$

Figure 9.8 shows the charge distribution measured by a BMVD prototype detector at perpendicular angle of incidence for the 4 readout chips separately. For each event, the two neighbouring strips with the highest signals are selected and the cluster pulse height, S_{evt}, is defined as the sum of their signals. The most probable energy deposition, estimated with a fit to the distribution by a Landau curve[5] and the strip noise level (averaged over the readout chip) are listed

[5]The function used for the fit is:

$$f(x, P_1, P_2, P_3) = P_1 \cdot \exp\left[-\frac{1}{2} \cdot \lambda + \exp(-\lambda)\right] \quad \text{where} \quad \lambda = P_3 \cdot (x - P_2) \tag{9.8}$$

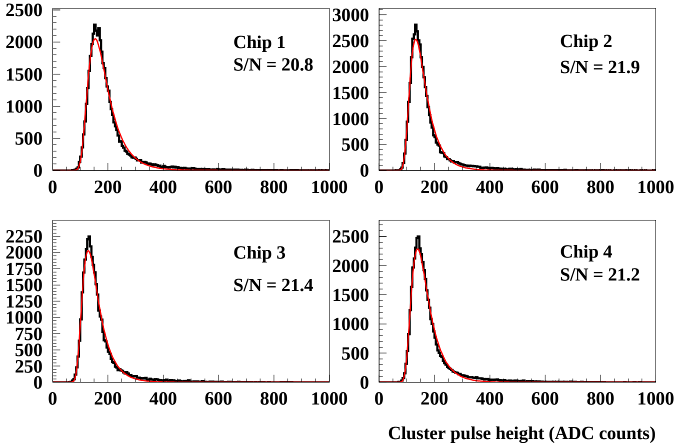

Figure 9.8: Pulse height distribution for the DUT fitted by a Landau curve for the four different readout chips.

in table 9.2, together with the resulting signal over noise ratio S/N.

Chip #	Signal S (ADC counts)	Average Chip Noise N (ADC counts)	S/N
1	153.9	7.30	20.8
2	132.3	6.04	21.9
3	128.6	6.02	21.4
4	139.0	6.55	21.2

Table 9.2: Most probable energy deposition (S), average chip noise (N) and S/N ratio of the DUT for the 4 readout chips separately.

9.4 MVD Prototype Detector Efficiency

The detector efficiency $\epsilon_{\rm det}$ is defined as:

$$\epsilon_{\rm det} = \frac{N_{\rm good\ cluster}}{N_{\rm tot}} \tag{9.9}$$

where $N_{\rm good\ cluster}$ is the total number of events for which a signal $S_{\rm evt} > 8 \cdot N_{\rm noise}$[6] has been measured and $N_{\rm tot}$ is the total number of events used. Different high statistics data runs ($\mathcal{O}(10^4)$ events) have been used to estimate the efficiency for BMVD prototype detectors and $\epsilon_{\rm det}$ values between 0.9996 and 0.9999 have been found. A similar analysis for the FMVD prototype detectors has been performed [58] and the results are in agreement with the one here reported.

[6] $N_{\rm noise}$ is the average strip noise.

9.5 MVD Prototype Depletion Voltage

During operation, all MVD silicon detectors have to be fully depleted in order to achieve a high charge collection efficiency and the best position resolution. The equivalent noise charge (ENC) of a charge sensitive amplifier is to a good approximation given by [59]:

$$Q_{\rm noise} \simeq \alpha_{\rm ENC} + \beta_{\rm ENC} \cdot C_{\rm input} \tag{9.10}$$

where $C_{\rm input}$ is the capacitive load at the input of the amplifier and α and β are parameters describing the properties of the chosen amplifier. From the formula in Eq. 8.3, one can see that if the applied bias voltage is small the dynamic capacitance $C_{\rm junc}$ becomes larger, implying that the equivalent capacitive load at the input of the amplifier increases and the noise level becomes larger. This effect can be clearly seen in the right plot of Fig. 9.9, in which the average chip noise level as a function of the bias voltage is shown. One can notice that the noise level reaches a plateu already for voltage values of the order of 20 V. From the left plot of Fig. 9.9, showing the most probable value for the energy deposition as a function of the applied bias voltage, one can conclude that the maximum charge collection is achieved for $V_{\rm bias} \gtrsim 80$ V. Two prototype detectors (det#10, det#16) were irradiated with photons (using a ^{60}Co source, $E_\gamma \sim 1.2$ MeV). During irradiation, det#10 was taken under a bias voltage of 100 V, whereas det#16 was unbiased. No effect is observed for both irradiated detectors when compared with the non-irradiated one (det#6), as expected since the photon irradiation does not produce damages in the detector bulk.

The depletion voltage $V_{\rm dep}$ of the MVD prototypes was estimated using a standard capacitance measurement method described in [50]. For all the 20 detectors belonging to the first production batch, the estimated depletion voltage was:

$$V_{\rm dep} \sim 65 \ \ \text{Volts}$$

A bias voltage scan was performed in the testbeam for different BMVD and one FMVD prototype detectors. By fitting the (V, $S^{-2}_{\rm evt}$) distribution one obtains:

$$V_{\rm dep} \sim 66 \ \ \text{Volts}$$

in agreement with the capacitance measurements.

A bias voltage of $V_{\rm dep} \simeq 100$ V has been applied to the MVD prototype detectors used for all the testbeam measurements described in this chapter.

9.6 MVD Prototype Detector Calibration

Due to possible gain variations in the readout electronics, for a fixed value of the charge induced on a readout strip, $Q_{\rm coll}$, the corresponding output signal, $S^k(Q_{\rm coll})$, may vary significantly in absolute value from strip to strip, thus introducing a systematic bias in any position reconstruction algorithm. It is therefore essential to quantify the strip-to-strip gain variation using a calibration procedure and to correct the data accordingly.

The calibration procedure is based on a matrix calculation. For all the events, $N_{\rm sel}$, the strip k with the highest signal, and its left and right neighbours (k-1, k+1) are selected (see Fig.

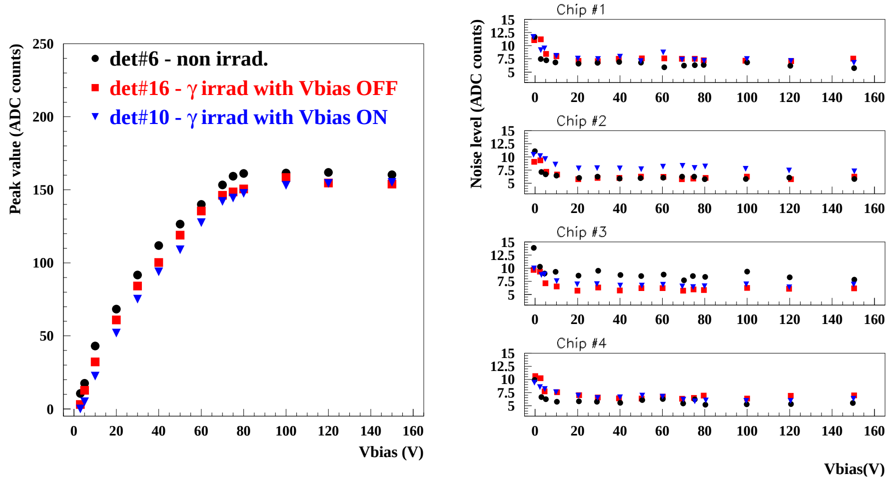

Figure 9.9: Signal peak as a function of the bias voltage, V_{bias}, applied to deplete the detector bulk, for three different prototype detectors (left plot) and average chip noise level as a function of V_{bias} (right plot).

9.10) and the respective signals $(h_i^k, l_i^{k-1}, r_i^{k+1})_{i=1,N_{sel}}$ are used to calculate the following average values:

$$h^k = \frac{1}{N_{\mathrm{sel}}} \cdot \sum_i^{N_{sel}} h_i^k, \quad l^k = \frac{1}{N_{\mathrm{sel}}} \cdot \sum_i^{N_{sel}} l_i^{k-1}, \quad r^k = \frac{1}{N_{\mathrm{sel}}} \cdot \sum_i^{N_{sel}} r_i^{k+1} \tag{9.11}$$

A $N_{\mathrm{strip}} \times N_{\mathrm{strip}}$ matrix is defined and the matrix elements (k,k), (k,k-1), (k,k+1) are filled with the corresponding signals h^k, l^k and r^k.

$$\begin{pmatrix} h^1 & r^1 & & & & \\ l^2 & h^2 & r^2 & & & \\ & l^3 & h^3 & r^3 & & \\ & & \cdots & \cdots & \cdots & \\ & & & & l^{N_{strip}} & h^{N_{strip}} \end{pmatrix}_{N_{strip} \times N_{strip}}$$

The three elements of each matrix column are summed up:

$$s^k = r^{k-1} + h^k + l^{k+1}$$

and then divided by the averaged value over the columns, s, to obtain the strip calibration constant c_k:

$$c^k = \frac{s^k}{s} \quad \text{where} \quad s = \frac{1}{N_{\mathrm{strip}}} \sum_{k=1}^{N_{strip}} s^k \tag{9.12}$$

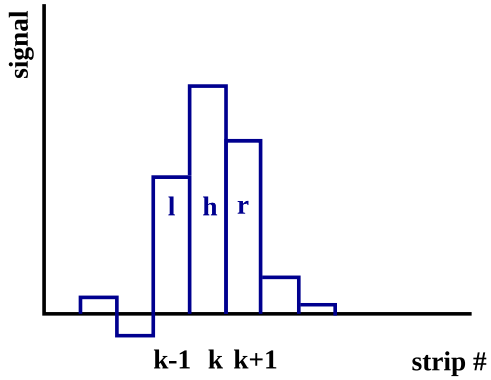

Figure 9.10: Schematic picture of a cluster from the DUT.

Figure 9.11 shows the strip calibration constants for a BMVD prototype detector[7]. In the calibration procedure, the first and last readout strips are always excluded (because either the left or the right neighbour is missing) as well as all dead channels plus their closest neighbours. For all excluded channels the calibration constant in the plot of Fig. 9.11 is set to zero. The variations within a readout chip are found to be small ($\lesssim$2%) and almost of the same order of the statistical uncertainty on the calibration measurement itself ($\sim$1%). The chip to chip variations are rather large and cannot be neglected. Gain variations $\gtrsim$ 20% cause observable systematic bias in the position reconstruction and have therefore to be taken into account. As a result, all the 128 readout channels of the same chip are calibrated using a single constant and only 4 different values for each BMVD detector are needed.

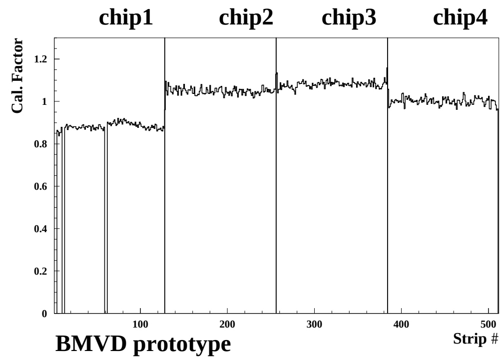

Figure 9.11: Result of the calibration procedure for one BMVD prototype detector. The vertical lines indicate the chip boundaries.

The same procedure has been also used for the FMVD prototype detectors[8] [58]. Figure 9.12

[7] The BMVD prototype detector used has two dead (i.e. inactive) readout channels and less than 1% noisy readout channels

[8] The FMVD prototype detector used has 3 dead readout channels.

shows the strip calibration constants for a FMVD prototype detector. No correlation between

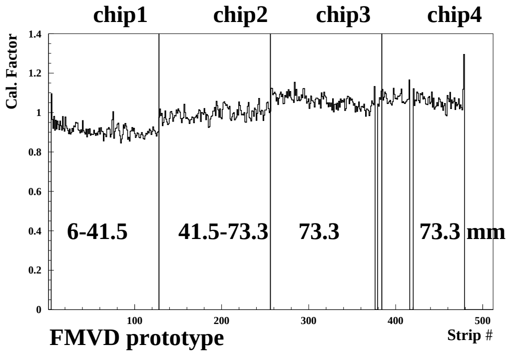

Figure 9.12: Result of the calibration procedure for a FMVD prototype detector. The vertical lines indicate the chip boundaries and the strip length range within each readout chip is also shown.

the gain fluctuations within one readout chip and the variation of the strip length is observed and therefore also for the FMVD detectors only one calibration value for each readout chip is needed.

BMVD	Chip1	Chip2	Chip3	Chip4
Mean value	0.884	1.050	1.081	1.001
RMS value	0.002	0.014	0.017	0.010
FMVD	**Chip1**	**Chip2**	**Chip3**	**Chip4**
Mean value	0.909	0.990	1.062	1.070
RMS value	0.036	0.024	0.029	0.039

Table 9.3: Average chip calibration constants and their rms for a BMVD and a FMVD prototype detector.

9.7 Alignment Procedure

To reconstruct the particle trajectory and calculate the impact point on the DUT, the telescope detectors and the DUT need to be aligned with respect to each other in a common reference system.

In the alignment procedure, three independent movements per each sensor are allowed, which correspond to three free parameters (see Fig. 9.13):

- a global offset perpendicular to the strip orientation ($\mathrm{off}_{x,y}$)
- a rotation around the strip axis ($\theta_{x,y}$)

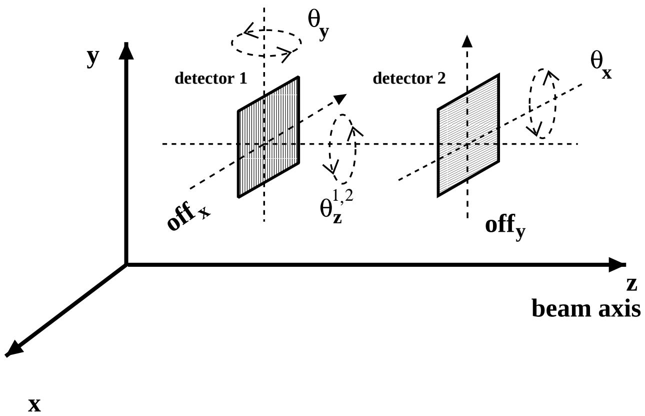

Figure 9.13: Drawing showing the alignment parameters for the testbeam setup.

- a rotation around the beam axis ($\theta_z^{(1,2)}$)[9]

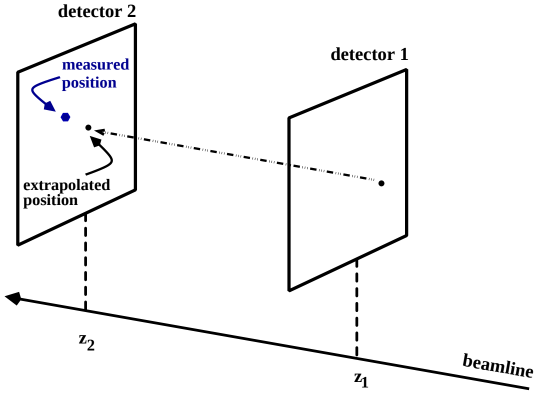

Figure 9.14: Representation of the position extrapolation in the alignment procedure.

The alignment constants are determined from a χ^2 fit which minimizes the distances between the measured and predicted position:

$$\chi^2 = \frac{(x - x_{\mathrm{pred}})^2 + (y - y_{\mathrm{pred}})^2}{\sigma^2_{\mathrm{tot}}} \tag{9.13}$$

where the predicted coordinated ($\mathrm{x_{pred}}$, $\mathrm{y_{pred}}$) are defined as the straight line extrapolation (parallel to the beam line) on the detector #2 of the position reconstructed on detector #1 (as shown in Fig. 9.14) and:

$$x = \left(\cos\theta_y \cdot \cos\theta_z^{(1)}\right) \cdot x_{\mathrm{meas}} + \left(-\sin\theta_x \cdot \sin\theta_y \cdot \cos\theta_z^{(1)} + \cos\theta_x \cdot \sin\theta_z^{(1)}\right) \cdot y_{\mathrm{meas}} + \mathrm{off_x}$$

[9] The upper numerical index refers to the two sensors of each telescope modules. For the DUT, which measures one coordinate, only one angle is of course needed.

$$y = \left(-\cos\theta_y \cdot \cos\theta_z^{(2)}\right) \cdot x_{\text{meas}} + \left(\sin\theta_x \cdot \sin\theta_y \cdot \cos\theta_z^{(2)} + \cos\theta_x \cdot \sin\theta_z^{(2)}\right) \cdot y_{\text{meas}} + \text{off}_y \ .$$

σ_{tot} is the total error on the reconstructed position given by the quadratic sum of the intrinsic resolution of the telescope modules ($\sigma_{\text{intr}}^{\text{ref}}$) and of the extrapolation error due to the multiple Coulomb scattering effect ($\sigma_{\text{mult.scatt.}}$)[10]:

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{intr}}^2 + \sigma_{\text{mult.scatt.}}^2}$$

The standard procedure starts by aligning the first telescope module (refer to Fig. 9.3 for the geometrical setup) with respect to the second one and afterwards the third telescope module and the DUT are aligned, using the extrapolated position determined with the first two reference modules. Since the DUT can only measure one coordinate, the other coordinate is replaced in the alignment fit by the predicted one. Alignment errors can be made visible by looking at the correlations between the residuals, ($y_{\text{meas}} - y_{\text{pred}}$) or ($x_{\text{meas}} - x_{\text{pred}}$), and the absolute predicted coordinate positions, x or y. Figure 9.15 shows an example of such correlation plots for the DUT, from which it is clear that the alignment procedure works correctly. The profile histograms are fitted by a polynomial of grade 0 with a resulting $\chi^2_{\text{ndf}} \sim 1.2$ and a shift from the zero of less than 0.3 μm.

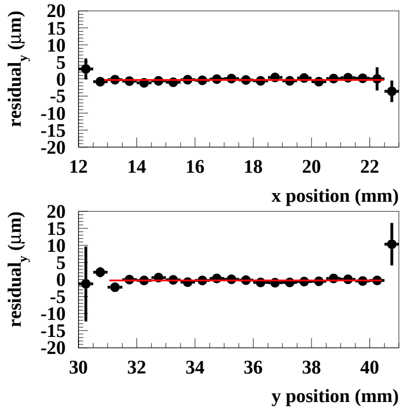

Figure 9.15: Mean value of the residual versus the y/x coordinate position.

9.8 Cross Talk Effects

The intermediate strips are coupled to the readout strips by means of the intrinsic capacitance between the implants (C_i). In the simple model where the capacitance of an intermediate strip towards next-to-closest readout strips is neglected, the charge collected on one intermediate

[10] A more detailed discussion can be found in Sec. 9.9.1

strip is then essentially transferred to the two closest readout channels (**principle of charge division**). However, in a more realistic model, also the couplings to the next-to-closest readout strips play a role, meaning that a non-negligible fraction of the total charge, which should be collected on the two closest readout strips, may be symmetrically spread on more readout strips (see Fig. 9.16). The term **symmetric cross talk** is here used to indicate this charge flowing from the readout strips closest to the point where the particle traversed the silicon detector to their neighbouring readout strips.

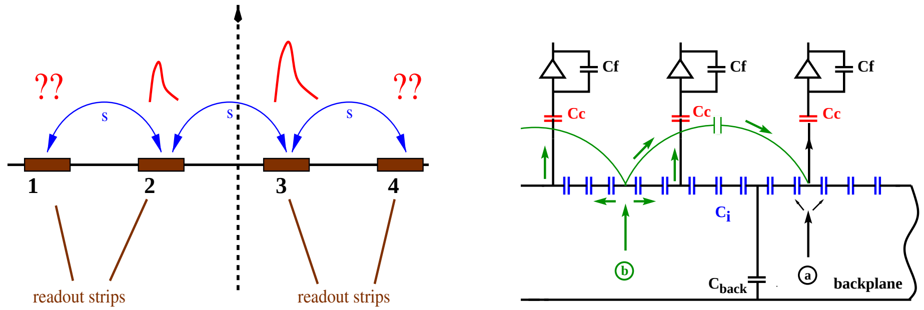

Figure 9.16: Schematic representation of the cross talk effect (left) and of the various coupling capacitances (right) in a detector making use of the capacitive charge division mechanism.

To ensure that most of the signal is transferred to the preamplifier of the two readout strips closest to the impact point, a high coupling capacitance (C_c) between the strip implant and the amplifier input is required, as well as a comparably low capacitance to neighbouring intermediate strips (C_i) and to the backplane (C_{back}). Typical values of:

$$C_c \sim 27pF/cm \quad , \quad C_i \sim 1.5pF/cm \quad \text{and} \quad C_{\text{back}} \sim 0.07pF/cm$$

have been measured for the MVD prototype detectors. In addition to the symmetric cross talk occurring on the detector itself, an additional symmetric charge flow may occur in the readout chip. This second contribution cannot be separated from the detector one and only the overall combined effect can be measured from the testbeam data.

Measurements performed on the readout chips (HELIX2.2 and successive versions), by means of external test pulses, have revealed the presence of an **asymmetric cross talk** among readout channels at both 10 MHz and 1 MHz clock rate. When pulsing a channel n, a fraction of the charge is measured alternatively on the channel (n-1) and (n+1). For the odd (even) channels[11], the previous (following) channel has essentially zero cross talk while the following (previous) one has a cross talk varying between 6-8 % (see Fig. 9.17). This effect presumably originates from an asymmetry in the chip pipeline layout.

To study the cross talk effects high statistics runs, at normal incidence angle and at different position along the DUT (in order to scan the whole detector) have been taken. The events are divided into two samples, depending on whether the beam particle[12] crossed the detector

[11] The readout channels on the HELIX chip are numbered from 0 to 127.
[12] The silicon telescope is used to calculate the particle impact position on the DUT.

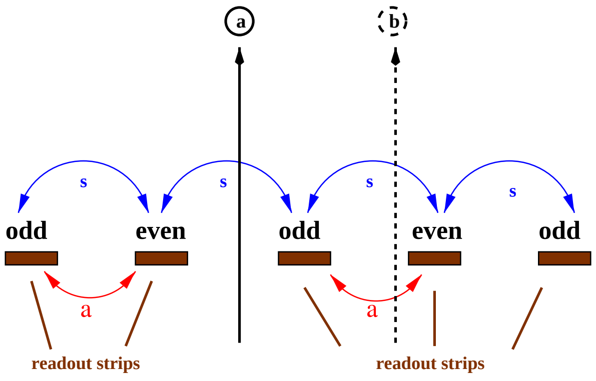

Figure 9.17: Asymmetric cross talk effect.

between an even and an odd readout strip (case a in Fig. 9.17) or between an odd and an even readout strip (case b in Fig. 9.17). Five neighbouring readout strips on both sides with respect to the particle position are selected and the total signal, $S_{\rm tot}$, is calculated as the sum of the signal S_i from the selected readout strips. Since the fraction of the total signal spread on to the next-to-next closest readout strips is very small, using two neighbouring readout strips on each side with respect to the particle position, two linear systems can be written and solved:

- **case a (even-odd)**

$$\begin{aligned} S_1 &= (a+s)\cdot T_2 \\ S_2 &= (1-2s-a)\cdot T_2 + s\cdot T_3 \\ S_3 &= s\cdot T_2 + (1-2s-a)\cdot T_3 \\ S_4 &= (a+s)\cdot T_3 \end{aligned} \tag{9.14}$$

- **case b (odd-even)**

$$\begin{aligned} S_1 &= s\cdot T_2 \\ S_2 &= (1-2s-a)\cdot T_2 + (a+s)\cdot T_3 \\ S_3 &= (a+s)\cdot T_2 + (1-2s-a)\cdot T_3 \\ S_4 &= s\cdot T_3 \end{aligned} \tag{9.15}$$

where $S_{1,2}$ and $S_{3,4}$ are the signals measured by the closest readout strips to the particle impact position (two on each side), $T_{2,3}$ are the 'true' signals which would have been measured by the two closest readout strips (one on each side), if there were no cross talk effects, and (s, a) are the symmetric and asymmetric cross talk coefficients respectively. By adding the first and last equations of the system in Eq. 9.14, one obtains:

$$S_1 + S_4 = (a+s)\cdot(T_2+T_3)$$

and assuming that $(T_2+T_3) = \sum_{i=1}^{4} S_i$ one can write:

$$a + s = \frac{S_1+S_4}{\sum_{i=1}^{4} S_i} \tag{9.16}$$

The same calculation for the system in Eq. 9.15 gives the following result:

$$s = \frac{S_1+S_4}{\sum_{i=1}^{4} S_i} \tag{9.17}$$

From Eqs. 9.16, 9.17 it is thus possible to estimate both coefficients $(a,\ s)$. The left plot in figure 9.18 shows the symmetric and asymmetric cross talk coefficients as a function of the particle position between the two closest readout strips. The symmetric coefficient, s, shows a clear dependence on the particle position whereas the asymmetric one, a, does not. The right plot in figure 9.18 shows instead their average values in different areas within one readout chip: both coefficients are constant within the chip. The average value for the asymmetric coefficient is $\sim 4.9\%$ and was found to be essentially the same for various HELIX chips, mounted on different hybrids. Table 9.4 shows the cross talk coefficients calculated for a single detector and a half-module. The symmetric coefficient for a half-module is higher than the one calculated for a single detector whereas the asymmetric coefficient shows only a $\lesssim 2\cdot\sigma$ variation.

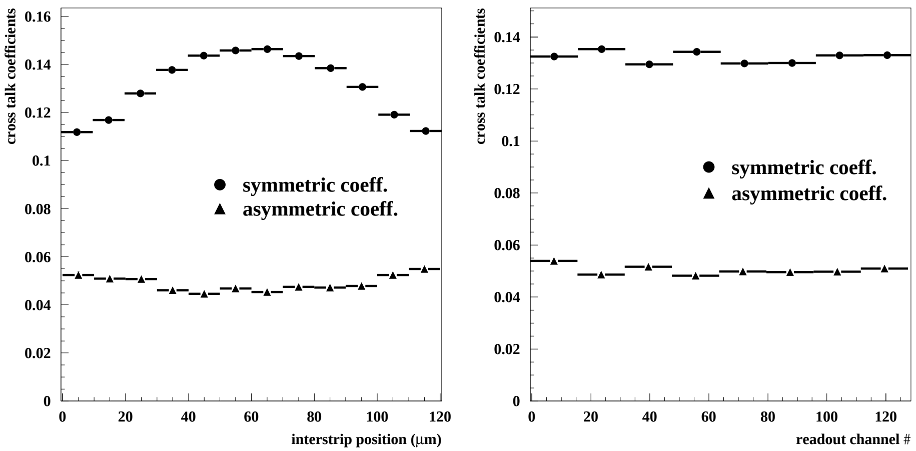

Figure 9.18: Cross talk coefficients as a function of the interstrip position (left) and of the readout channel # (right).

In the testbeam data analysis, the asymmetric cross talk contribution is corrected for, in order to avoid a systematic bias. The symmetric cross talk is not corrected for since it does not introduce any systematic bias in the position reconstruction algorithm and it can actually be used to improve the position resolution as will be shown in Sec. 9.9.

	half-module (%)		single detector (%)	
	symm. coeff. s	asymm. coeff. a	symm. coeff. s	asymm. coeff. a
chip 1	17.7 ± 0.4	4.2 ± 0.3	–	–
chip 2	17.2 ± 0.4	4.4 ± 0.5	13.1 ± 0.1	4.9 ± 0.3
chip 3	16.3 ± 0.4	4.2 ± 0.5	12.4 ± 0.1	5.0 ± 0.3
chip 4	15.9 ± 0.4	4.5 ± 0.3	12.7 ± 0.1	4.8 ± 0.3

Table 9.4: Average values for the cross talk coefficients for a single detector and a half-module, for the different readout chips.

9.9 Position Reconstruction at Small Incidence Angle

The algorithms used to estimate the impact position on the DUT, as well as on the telescope modules, are based on the fact that the signal on the readout strips depends on the track position between the strips. A standard method, proven to work at small incidence angle, is the so called η algorithm [60]. It consists of a non-linear interpolation between the two neighbouring strips of the cluster which have collected the highest signals (indicated in the following as S_{right} and S_{left}, respectively). A cluster, in this case, is defined by taking all the strips whose signal is above a certain threshold level T (usually $T = 3\times\sigma_{chip}$, where σ_{chip} is the average chip noise[13]). For each event, the quantity η:

$$\eta = \frac{S_{\text{right}}}{S_{\text{right}} + S_{\text{left}}} \tag{9.18}$$

is calculated. In case of a linear interpolation, η represents the particle position relative to the readout strips. Figure 9.19 shows an example of the $dN/d\eta$ distribution for the DUT obtained

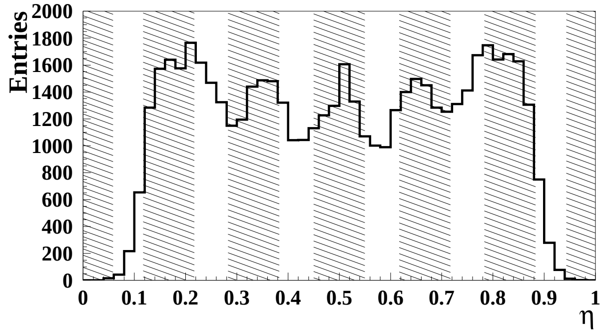

Figure 9.19: Example of $dN/d\eta$ distribution for the DUT.

from the testbeam data: the region close to the readout strips is almost desert, and clear peaks can be seen around the interstrip position regions. The fact that the $dN/d\eta$ is not uniform, although the beam profile is uniform over the detector area, indicates that the capacitive charge

[13] All the readout channels with $N^k_{noise} > 3 \cdot \sigma_{chip}$ are masked and not used in the analysis.

division mechanism does not depend linearly on the hit position between readout strips. To correct for this non-linearities a probability density function is introduced:

$$f(\eta_0) = \frac{1}{N_0} \cdot \int_0^{\eta_0} \frac{dN}{d\eta} d\eta \tag{9.19}$$

where N_0 is the total number of entries in the $dN/d\eta$ distribution and η_0 is the η value for the considered event. The corrected impact position is then given by:

$$y_{\text{rec}} = \text{p} \cdot f(\eta_0) + y_{\text{left}} \tag{9.20}$$

where p is the readout pitch of 120 μm and y_{left} denotes the position of the left strip. Figure 9.20 shows the function used to reconstruct the track position[14].

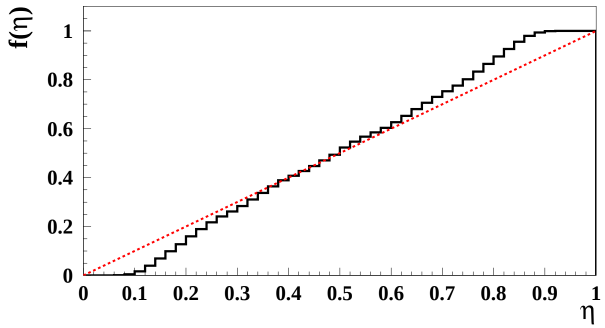

Figure 9.20: Example of the probability density function f(η) (solid line). The dashed line is plotted to indicate the deviation from the linear interpolation. The major effect occurs in the area close to the readout strips.

Figure 9.21 shows the $dN/d\eta$ distribution and the probability density function $f(\eta)$ for the three telescope modules. In this case the region close to the readout strips are populated (indicating a much smaller cross talk among readout strips than the one observed on the DUT) and a peak in the region around the intermediate strip can also be seen.

Using the η algorithm to reconstruct the hit position on the DUT as well as on the telescope modules, the residual distribution for the DUT has been obtained (see Fig. 9.22). The residual ($y_{recon} - y_{pred}$) is defined as the difference between the reconstructed and the predicted position coordinate. The standard deviation of the Gaussian fit to the residual distribution represents the overall detector resolution:

$$\sigma = 7.7 \pm 0.1 \ \mu\text{m}$$

The different contributions to the overall position resolution will be discussed in Sec. 9.9.1. The value for the overall position resolution obtained from the testbeam measurements represents a striking achievement for a silicon sensor having a readout pitch of 120 μm.

[14] The correction procedure is also referred to as 'bias correction'.

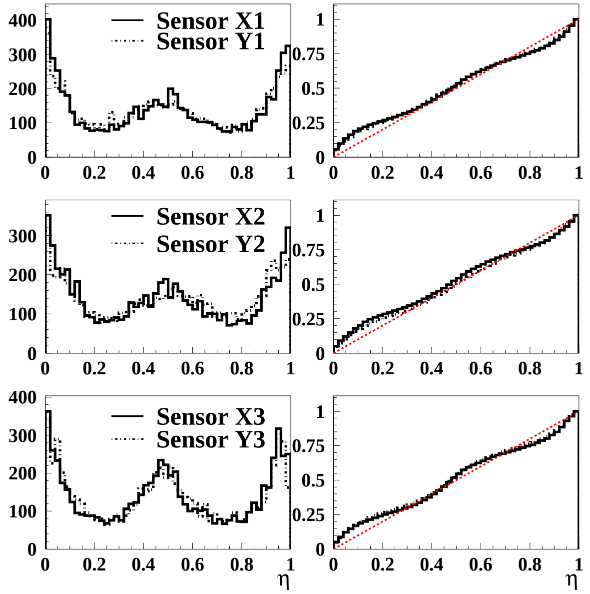

Figure 9.21: Example of $dN/d\eta$ distribution and the probability density function $f(\eta)$ for the telescope modules.

The left plot in Fig. 9.23 shows the position resolution as a function of the interstrip position. One can notice the presence of an alternate systematic pattern when moving from a p^+ implant to the next one. The variations observed are in general very small ($\lesssim 1\mu$m) and do not have any relevant effect on the measurements presented in the following. In the region close to the readout strips, one can notice that the position resolution becomes slightly worse. This effect is a consequence of the use of only two readout strips for the position reconstruction: when a particle traverses the silicon sensor very close to a readout strip, a high signal is induced on that readout strip and only relatively small signals on the neighbouring ones. Therefore the η algorithm becomes more sensitive to noise fluctuations. The right plot in Fig. 9.23 shows the mean value of the Gaussian fit to the residual distribution (corresponding to a systematic shift from the origin) as a function of the interstrip position. Close to the readout strips the systematic shift in the position reconstruction is larger ($\sim$ 1.5-2 μm) than in the central area between readout strips ($\lesssim 1$ μm).

Figure 9.24 shows the most probable value for the energy deposition (corresponding to the peak in the energy loss distribution) as a function of the interstrip position. It can be seen that the amount of charge, collected by the two closest readout strips, strongly depends on the particle impact position. If the particle traverses the detector in the central region between two readout strips, $\sim$ 17% of the signal is lost due to the capacitive coupling to the detector backplane. This effect has been cross-checked using a detector simulation [61] and the result from the simulation is found to be in fair agreement with the testbeam measurement.

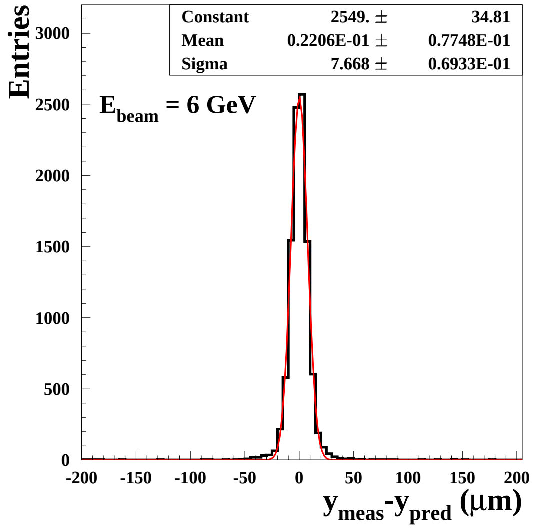

Figure 9.22: Residual distribution for the DUT, at $E_{beam} = 6$ GeV, fitted by a Gaussian function. A cut to remove δ-electrons, which can spoil the detector resolution, has been applied: $(S_{\mathrm{left}} + S_{\mathrm{right}}) < 1.7$ peak signal. The standard deviation of the Gaussian fit represents the overall detector resolution: $\sigma = 7.7 \pm 0.1\,\mu$m. Note that the shift from the origin is $\sim 0.02\mu$m.

9.9.1 Intrinsic Position Resolution

The intrinsic resolution is defined as the spatial precision of the DUT independent from environmental conditions, e.g. test setup geometry, beam energy and so on.

The measured position resolution σ_{res}, shown in Fig. 9.22, includes several contributions:

$$\sigma_{\mathrm{res}} = \sqrt{(\sigma_{\mathrm{DUT}}^{\mathrm{intr}})^2 + k \cdot (\sigma_{\mathrm{tele}}^{\mathrm{intr}})^2 + \sum_i k_i \cdot \Delta\theta_{\mathrm{ms}}^2} \qquad (9.21)$$

where $\sigma_{\mathrm{DUT}}^{\mathrm{intr}}$ is the intrinsic resolution of the DUT, $\sigma_{\mathrm{tele}}^{\mathrm{intr}}$ is the intrinsic resolution of the telescope sensors, (k, k_i) are geometrical factors related to the relative distances between the telescope modules, the DUT and also including the thickness of the aluminium window foils, and $\sum_i k_i \cdot \Delta\theta_{\mathrm{ms}}$ is the extrapolation error due to the multiple Coulomb scattering along the particle direction [37]:

$$\Delta\theta \simeq \frac{13.6\mathrm{MeV}}{\beta c p_{\mathrm{beam}}[GeV/c]} \cdot \sqrt{\frac{\Delta x}{X_0}} \cdot [1 + 0.038\ln(\Delta x/X_0)] \qquad (9.22)$$

In Eq. 9.22, p_{beam} and $c\beta$ are the momentum and velocity of the incident particle, and $\Delta x/X_0$ is the thickness of the scattering medium in radiation lengths.

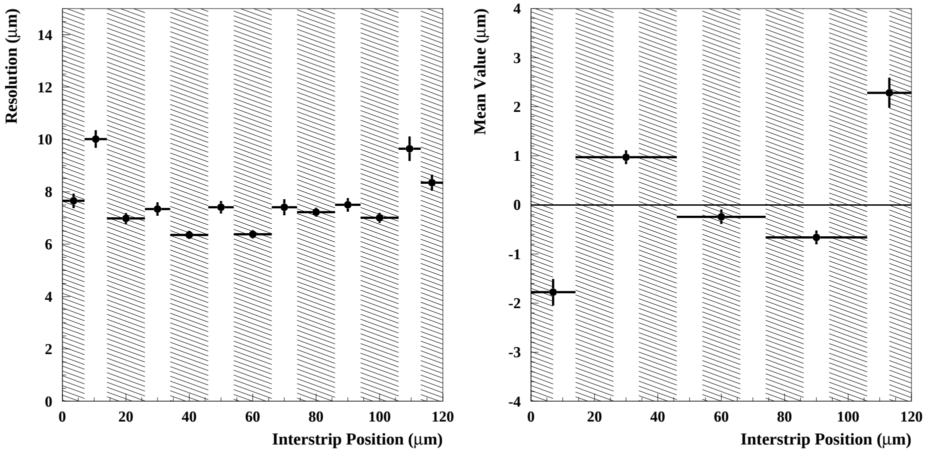

Figure 9.23: Position resolution and average residual shift from the origin as a function of the interstrip position. Note that in this case no cut to reject δ-electrons has been applied. The hatched bands indicate the position of the readout and intermediate strips.

To extract the intrinsic position resolution of the DUT ($\sigma^{\mathrm{intr}}_{\mathrm{DUT}}$), the contribution of the second and third term in Eq. 9.21 has been evaluated. The detailed analysis is described elsewhere [62]. The effect of the multiple Coulomb scattering at low beam energy (2-6 GeV) cannot be neglected; the residual distribution has therefore been measured, using the η algorithm, at several beam energies and by fitting the data points (shown in Fig. 9.25) to the formula in Eq. 9.21 the intrinsic resolution at 0° has been extracted. The same procedure has been performed first for the telescope modules only (taking apart the DUT) in order to extract the intrinsic resolution $\sigma^{\mathrm{intr}}_{\mathrm{tele}}$ of the telescope sensors[62], $\sigma^{\mathrm{intr}}_{\mathrm{tele}} \simeq 2.8\mu$m.

The intrinsic position resolution obtained for the DUT at $\theta = 0°$ incidence angle is:

$$\sigma^{\mathrm{intr}}_{\mathrm{DUT}} = 7.2 \pm 0.2\ \mu\mathrm{m}$$

Since the production of δ-rays can spoil the position resolution (see Fig. 9.26), a selection cut to reject the events in the tail of the energy loss distribution has been applied in the previous calculation:

$$S_{\mathrm{cluster}} \leq 1.7 \cdot S_{\mathrm{peak}}$$

where S_{cluster} is the total cluster charge and S_{peak} is the most probable energy deposition.

9.9.2 The 3-Strips Position Reconstruction Algorithm

The use of only two readout strips to extract position information may not be the most appropriate choice for the MVD detectors, since the symmetric cross talk to the next-to-closest

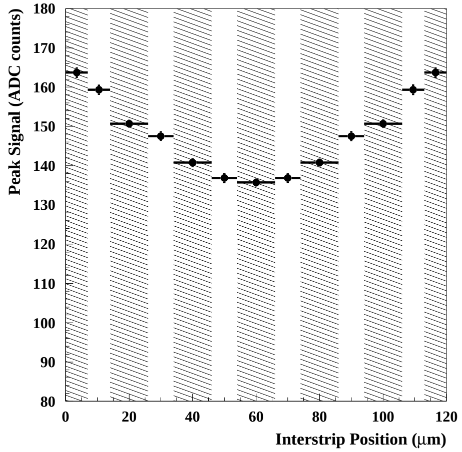

Figure 9.24: Most probable value for the energy deposition in the DUT as a function of the interstrip position. The hatched bands indicate the position of the readout and intermediate strips.

readout strips is not negligible. Figure 9.27 shows the cluster size[15] distribution at four dif-

[15] The cluster size is defined as the total number of strips forming a cluster.

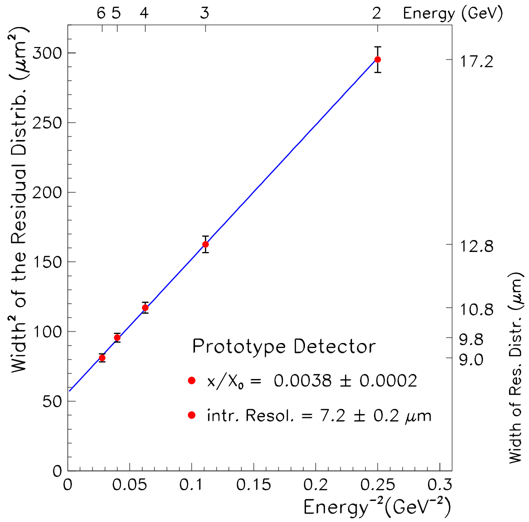

Figure 9.25: Position resolution as a function of the beam energy [62].

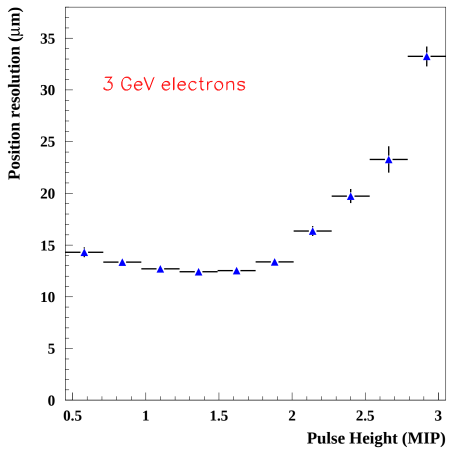

Figure 9.26: Position resolution as a function of the pulse height, normalized to the most probable energy deposition from a minimum ionizing particle (MIP).

ferent incidence angles (0°, 10°, 20°, 30°). At $\theta = 0°$, already $\sim 28\%$ of all cluster consists of more than two strips and this percentage becomes much larger ($\gtrsim 45\%$) for $\theta \gtrsim 20°$. An alter-

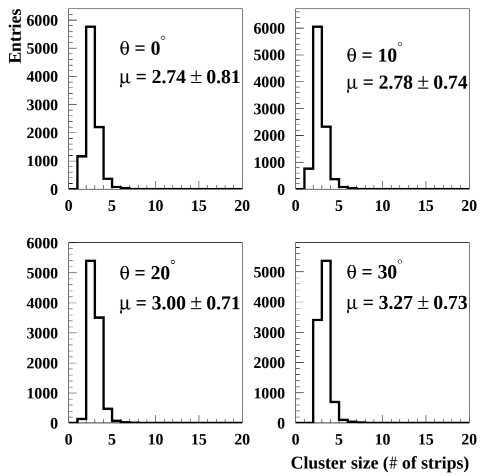

Figure 9.27: Cluster size distribution for small incidence angles (0°, 10°, 20°, 30°). The mean value for the cluster size is also indicated with the rms as error.

native position reconstruction algorithm, which makes use of three strips, has been therefore developed. From each cluster, the strip with the highest signal, k, and its closest neighbours (k-1,k+1) are selected and the following quantities are calculated:

$$\mathrm{p_{left}} = \frac{S_k \cdot k + S_{k-1} \cdot (k-1)}{S_k + S_{k-1}} \quad \text{and} \quad \mathrm{p_{right}} = \frac{S_k \cdot k + S_{k+1} \cdot (k+1)}{S_k + S_{k+1}} \tag{9.23}$$

The uncorrected reconstructed position $\mathrm{p_{rec}}$ (in analogy to the linear η interpolation) is then defined as the weighted mean of $\mathrm{p_{left}}$ and $\mathrm{p_{right}}$:

$$\mathrm{p_{rec}} = \frac{\mathrm{p_{left}} \cdot w^{2n} + \mathrm{p_{right}}}{1 + w^{2n}} \quad \text{where} \quad w = S_{k-1}/S_{k+1} \quad \text{and} \quad n = 1 \tag{9.24}$$

and the corresponding interstrip position $\tilde{p}$ is given by:

$$\tilde{p} = \mathrm{mod}(\mathrm{p_{rec}}, 1.)$$

Figure 9.28 shows the dN/dη and dN/d$\tilde{p}$ distributions for small incidence angles (0°, 10°, 20°, 30°). By means of the 3-strips algorithm a more uniform distribution can be obtained. In analogy with the η algorithm a probability density function can be defined:

$$f(\tilde{p}_0) = \frac{1}{N_0} \cdot \int_0^{\tilde{p}_0} \frac{dN}{d\tilde{p}} d\tilde{p} \tag{9.25}$$

and the corrected impact position is given by:

$$y_{\mathrm{rec}} = \mathrm{p} \cdot f(\tilde{p}_0) + p_{\mathrm{left}} \tag{9.26}$$

where p is the readout pitch and p_{left} denotes the position reconstructed using the left and the central strips. Figure 9.29 shows the comparison between $f(\tilde{p})$ and $f(\eta)$ for $\theta = 0°$. The position resolution obtained using the η algorithm and the 3-strips algorithm for the four different incidence angles (0°, 10°, 20°, 30°) are shown in Fig. 9.30.

The resolution at 0° calculated with the 3-strips algorithm is slightly worse than the one obtained with the η algorithm since for perpendicular tracks the use of a third strip causes a higher sensitivity to noise fluctuation. Nevertheless the bias correction for the 3-strips algorithm is much smaller than the one needed for η algorithm up to 30°, as shown in the right plot of Fig. 9.30. So the 3-strips algorithm is much less sensitive to possible variations in the probability density function due to eventual non-uniformities on the MVD detectors (e.g. bond pads regions, overlap region for the half-module, etc.)[16].

9.10 Position Reconstruction at Large Incidence Angle

At large incidence angles ($\theta > 30°$) the charge is spread over several strips and the total cluster signal becomes larger since the particle path length increases (see Fig. 9.31)[17]:

$$S(\theta) \propto \frac{S(\theta = 0°)}{\cos\theta} \tag{9.27}$$

Since the central strips have on average the same signal, the information on the impact position is essentially contained only in the positions and signals of the cluster edge strips. Therefore

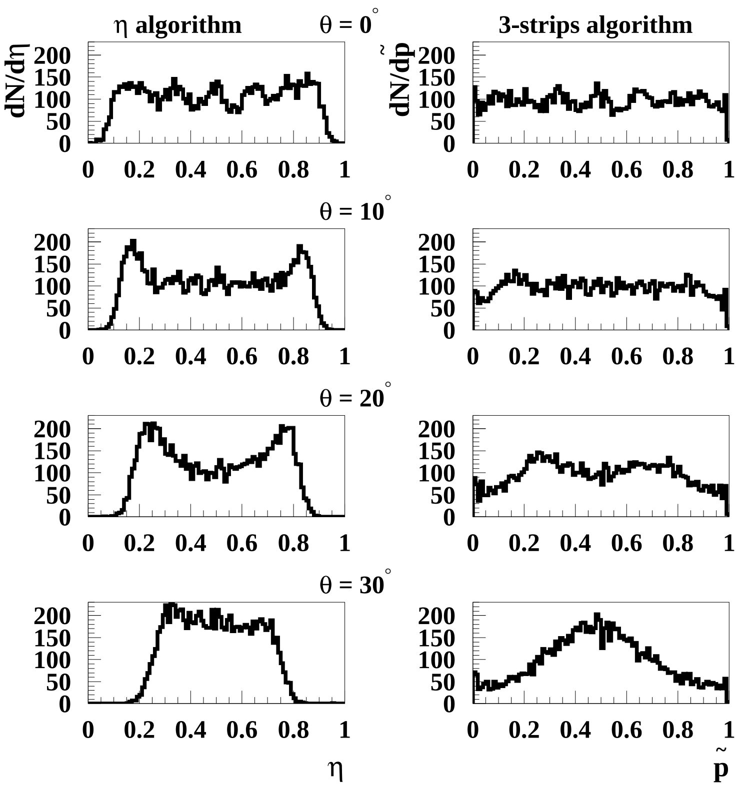

Figure 9.28: dN/dη and dN/d$\tilde{p}$ distributions for small incidence angles.

reconstruction algorithms such as the η and the 3-strips ones are inadequate to calculate the particle impact point on the detector. Figure 9.32 shows the cluster size distributions for large incidence angles (40°, 50°, 60°, 70°). The left plot in Fig. 9.33 shows the energy loss distribution for various incidence angles (0°, 30°, 50°, 60°, 70°), whereas in the right plot the most probable value for the energy deposition as a function of the incidence angle is shown. In this case the energy deposition is defined by summing up the signals measured on ± 5 strips around the predicted position:

$$S_{\mathrm{tot}} = \sum_{i=-5}^{+5} S_i$$

[16] In the testbeam data, it was observed that the probability density function showed large variations in correlation with the particular detector area irradiated by the electron beam.

[17] The probability of δ-ray emissions also increases according to the path length.

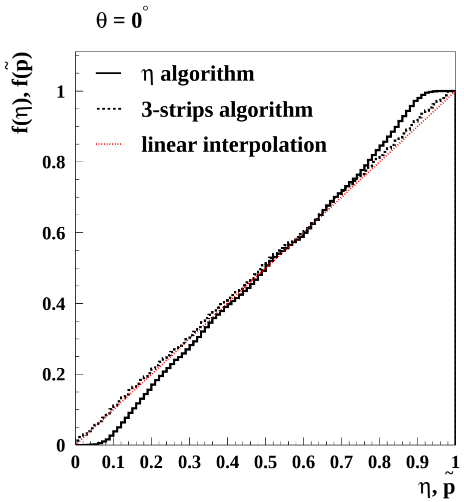

Figure 9.29: Probability density function for the η and the 3-strips algorithm. Note that at $\theta = 0°$ the result for f($\tilde{p}$) is very close to the linear interpolation.

The result of a fit to the function:

$$f(\theta, P_1, P_2) = \frac{P_1}{\cos\theta^{P_2}} \tag{9.28}$$

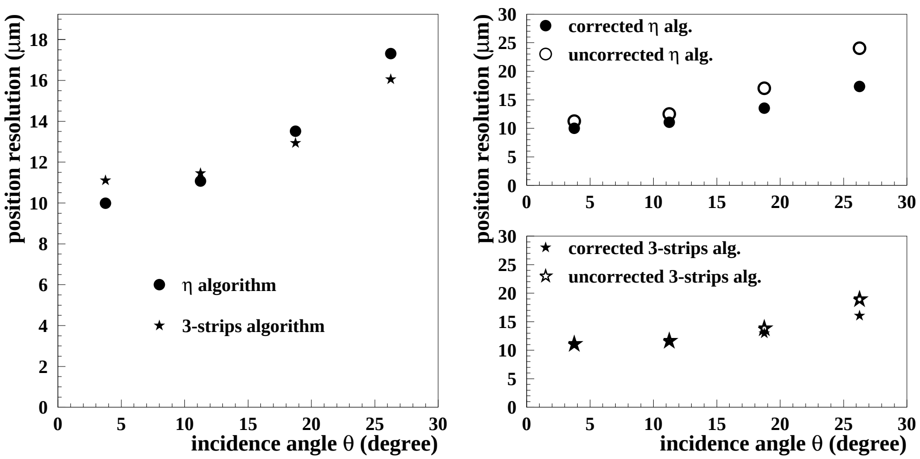

Figure 9.30: Position resolution for small incidence angles (0°, 10°, 20°, 30°). The left plot shows the comparison between the corrected position resolution calculated using the η and the 3-strips algorithm. The right plot shows the comparison between the uncorrected and corrected position resolution for both algorithms.

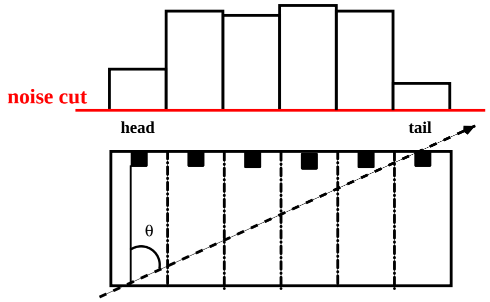

Figure 9.31: Geometrical representation of the detector geometry in the case of large incidence angles, where also the expected cluster charge distribution is shown.

is also presented. The value $P_2 = 1.08 \pm 0.04$ is in agreement with the expectation in Eq. 9.27 (i.e. $P_2 = 1.0$). The standard procedure of applying a threshold cut to select the strips, $\mathrm{N_{strips}}$, belonging to a cluster and then to define the energy deposition as :

$$S_{\mathrm{cluster}} = \sum_{i=1}^{\mathrm{N_{strip}}} S_i$$

causes at low angle of incidence a systematic shift of the peak signal to lower values (see Fig. 9.34), thus resulting in a deviation from the expected behaviour.

A standard position reconstruction algorithm, proven to work at large incidence angle, is the

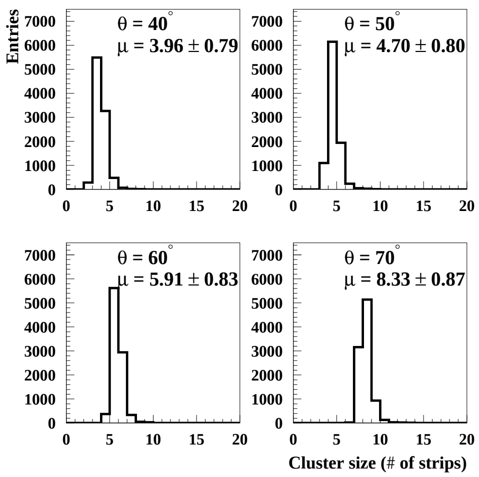

Figure 9.32: Cluster size distributions for large incidence angles (40°, 50°, 60°, 70°)

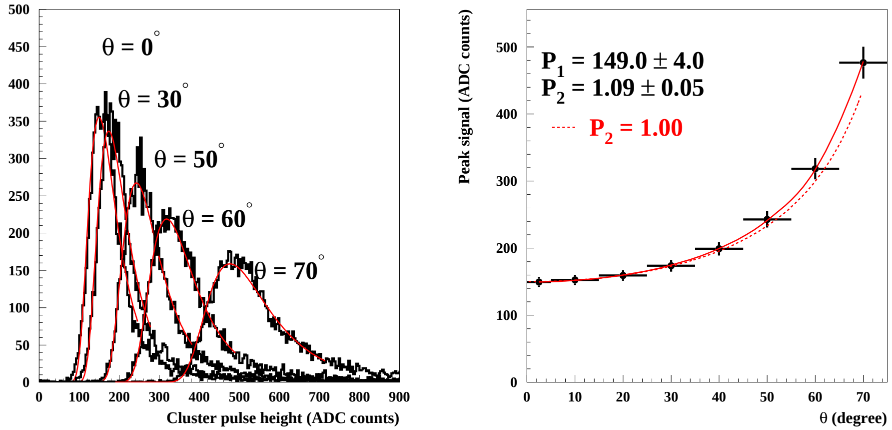

Figure 9.33: Energy loss distribution at various incidence angles (left) and most probable value for the energy deposition as a function of the incidence angle (right).

so called 'head-tail' algorithm [63]. All the strips with:

$$S_{\mathrm{strip}} > 3 \cdot \sigma_{\mathrm{chip}}$$

where S_{strip} is the strip signal and σ_{chip} is the average chip noise, are considered. The first (head) and the last (tail) strips belonging to a cluster are selected (see Fig. 9.31) and the

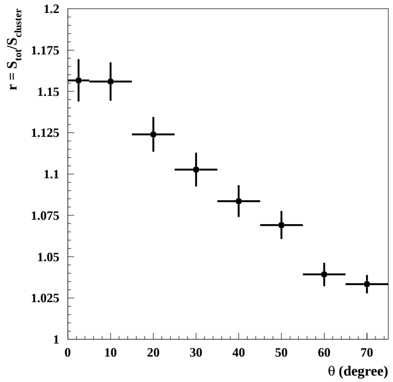

Figure 9.34: Ratio of the most probable value for the energy deposition calculated with the two different definitions ($\mathrm{S_{tot}}$, $\mathrm{S_{cluster}}$) as a function of the incidence angle.

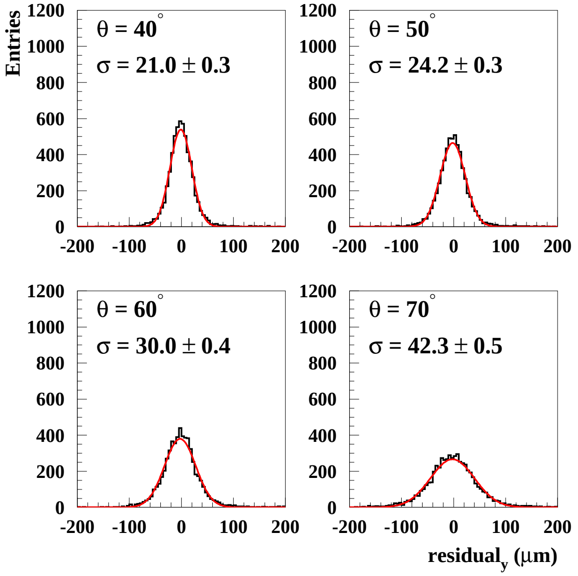

Figure 9.35: Residual distributions for large incidence angles (40°, 50°, 60°, 70°).

impact position is defined as:

$$y_{\rm rec} = \frac{y_{\rm head} + y_{\rm tail}}{2} + {\rm ht}_{\rm corr} \quad \text{with} \quad {\rm ht}_{\rm corr} = \frac{S_{\rm tail} - S_{\rm head}}{2 \cdot < S >} \cdot p \tag{9.29}$$

where $y_{\rm head}$ ($y_{\rm tail}$) is the position of the head (tail) strip, $S_{\rm head}$ ($S_{\rm tail}$) is the corresponding signal, p is the strip pitch and $< S >$ is the average strip signal defined as :

$$< S > = \frac{1}{N_{\rm cluster}} \sum_{i={\rm head}}^{\rm tail} S_i$$

where $N_{\rm cluster}$ is the number of strips belonging to a cluster.

The difference $(S_{\rm tail} - S_{\rm head})$ in the correction factor ${\rm ht}_{\rm corr}$ is used to shift the average position $(y_{\rm head} + y_{\rm tail})/2$ towards the tail $(S_{\rm tail} - S_{\rm head} > 0)$ or head $(S_{\rm tail} - S_{\rm head} < 0)$ strip of the cluster, taking into account the rough proportionality of the energy loss to the particle path in the detector. Figure 9.35 shows the residual distributions obtained using the 'head-tail' algorithm for four different incidence angles (40°, 50°, 60°, 70°).

9.11 Optimization of Reconstruction Algorithms

The reconstruction algorithms described in sections 9.9 and 9.10 are based on the knowledge of the incidence angle θ of the particle, used to select the most effective algorithm to be applied.

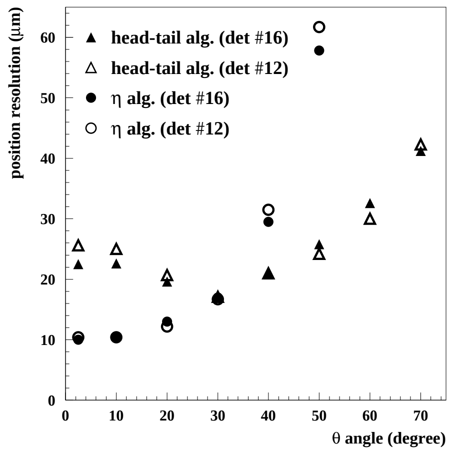

Figure 9.36: Position resolution calculated with the η and head-tail algorithms as a function of the angle of incidence θ. At angles $\lesssim 30°$ the η algorithm as well as the 3-strips algorithm (not shown here) give a much better resolution than the head-tail algorithm. However for angles $\gtrsim 30°$, the latter proves to work much better. The results using two different MVD barrel prototypes are presented.

For small incidence angles, the algorithms depend also on the bias correction which has been seen to depend on the position of the hit in the detector. The sharp switch at some value $\theta = \overline{\theta}$ between different algorithms clearly implies complications in the track reconstruction procedure. Ideally an optimal algorithm should calculate impact positions with a single calculation procedure, using as little additional external information as possible to minimize systematic effects. A first easy attempt for a more general algorithm could be based only on the rough knowledge of the angle θ of the incident particle ($\theta_{\rm cut}$ algorithm):

- if $\theta \lesssim \theta_{\rm cut}$ (with $\theta_{\rm cut} \simeq$ 30-35°) the uncorrected 3-strips algorithm is used (see Sec. 9.9)
- if $\theta > \theta_{\rm cut}$ a modified version of the head-tail algorithm is used, such that similar formulas, compared to the ones used for small angles, are employed. This should reduce as much as possible any systematic effect introduced by the change in the calculation of position reconstruction. The head strip (k) and tail strip (m) belonging to a cluster[18] are selected and the impact position is defined as:

$$y_{\rm rec} = \frac{\tilde{y}_{\rm head} + \tilde{y}_{\rm tail}}{2} \cdot p + \tilde{\rm ht}_{\rm corr} \quad \text{with} \quad \tilde{\rm ht}_{\rm corr} = \frac{S^2_{\rm tail} - S^2_{\rm head}}{(2 \cdot < S >)^2} \cdot p \tag{9.30}$$

where:

$$\tilde{y}_{\rm head} = \frac{S_{\rm k} \cdot k + S_{k+1} \cdot (k+1)}{S_k + S_{k+1}} \quad \text{and} \quad \tilde{y}_{\rm tail} = \frac{S_{\rm m} \cdot m + S_{m-1} \cdot (m-1)}{S_m + S_{m-1}}$$

[18] A cluster is defined, as usual, by taking all the strips whose signal is above a certain threshold level T_0 (usually $T_0 = 3 \times \sigma_{\rm chip}$), where $\sigma_{\rm chip}$ is the average chip noise

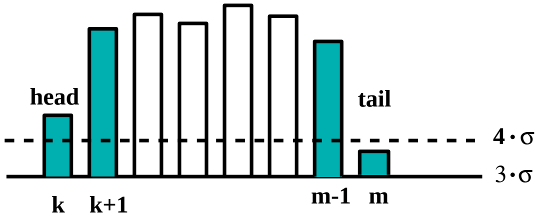

Figure 9.37: Schematic representation of a cluster at large incidence angles.

The correction $\tilde{\mathrm{ht}}_{\mathrm{corr}}$ is applied only in the case either the head or the tail strip signal is below a fixed threshold T_1 (here set to $T_1 = 4 \cdot \sigma_{\mathrm{chip}}$), as shown in Fig. 9.37. If both edge strips signals are above the threshold T_1, then the impact position is given by:

$$y_{\mathrm{rec}} = \frac{\tilde{y}_{\mathrm{head}} + \tilde{y}_{\mathrm{tail}}}{2} \cdot p \tag{9.31}$$

Note that in this case no bias correction is applied with the disadvantage of having a slightly worse position resolution at small angles. Figure 9.38 shows the position resolution obtained with this algorithm as function of the incidence angle θ for two different prototype detectors.

A more general approach is based on the use of the cluster size, N_{cluster}, as switch parameter ('cluster cut' algorithm). However, it turned out that by simply replacing the switch parameter, the position resolution obtained is significantly worse and thus an additional parameter, R_{out} (see Fig. 9.40), related to the cluster 'shape' has been introduced:

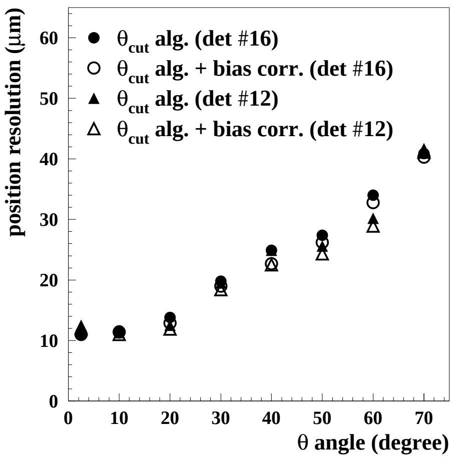

Figure 9.38: Position resolution as a function of the angle of incidence θ calculated using the θ_{cut} reconstruction algorithm.

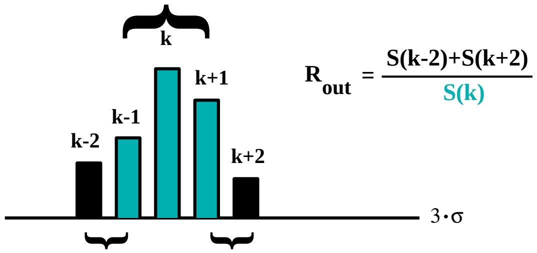

Figure 9.39: Schematic representation of a small cluster: depending on the value of the parameter R_{out}, the 'cluster cut' algorithm decides whether to use the 3-strips algorithm with the central strips or the head-tail algorithm with the edge strips.

- if $N_{\mathrm{cluster}} < N_{\mathrm{cut}}$ (with $N_{\mathrm{cut}} = 5$ or 6) and $R_{\mathrm{out}} < R_{\mathrm{cut}}$, where $R_{\mathrm{out}} = (S_{k-2} + S_{k+2})/S_k$[19], the 3-strips algorithm is used (see Sec. 9.9).
- if $N_{\mathrm{cluster}} > N_{\mathrm{cut}}$ or $R_{\mathrm{out}} \geq R_{\mathrm{cut}}$ the same modified version of the head-tail algorithm used in the previous case is employed.

The value for R_{cut} is 0.3 and was chosen by comparing the distribution for different incidence angles shown in Fig. 9.40. From the plot in Fig. 9.40 it is clear that the cluster shape depends on the angle of incidence: for larger angles, the outer strips of small clusters have a sizable signal compared to the highest signal in the cluster.

[19] The notation is the same used in Sec. 9.9, i.e. k is the strip with the highest signal in the cluster.

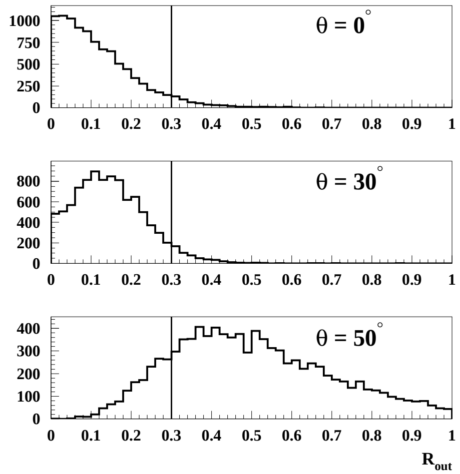

Figure 9.40: R_{out} distribution at $\theta = 0°$, $30°$, $50°$

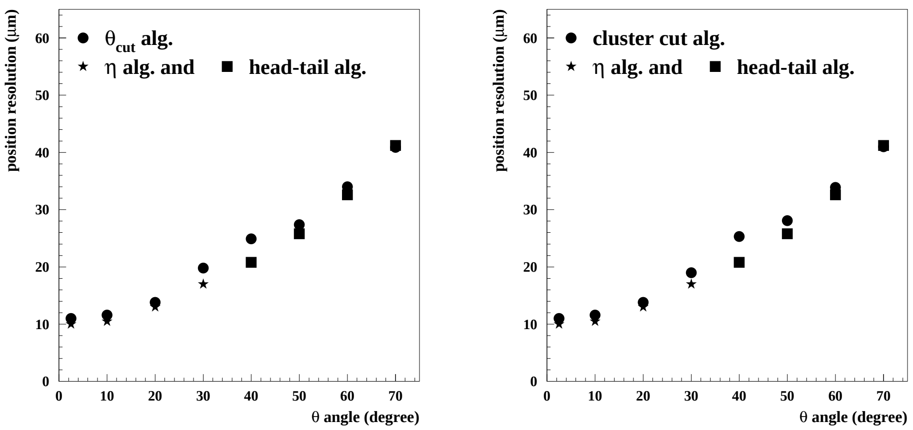

Figure 9.41: Position resolution as a function of the angle of incidence θ obtained with the θ_{cut} (left) and the 'cluster cut' (right) algorithms. The comparison with the η algorithm (at small incidence angles) and head-tail algorithm (at large incidence angles) is also shown.

Figure 9.41 shows the comparison between the position resolution obtained with the two different switch parameters and the position resolution obtained with the η and the head-tail algorithms. Although the position resolution obtained with the uncorrected 'cluster-cut' algorithm is a little worse than the one calculated with the standard reconstruction methods (i.e. η and the 'head-tail'), the 'cluster-cut' algorithm does not need any angular information and could therefore be a valuable choice for a first position reconstruction in a general track reconstruction procedure.

9.12 Conclusions

The testbeam setup at DESY has been used to characterize the performance of BMVD and FMVD prototype detectors plus their readout electronics (HELIX chips). The main results from the testbeam measurements presented in this chapter can be summerized as follows:

- the detector efficiency ϵ_{det} is very high ($> 99.96\%$)
- a signal over noise ratio S/N $\simeq 21$ has been achieved for the readout electronics (HELIX2.2) coupled to BMVD prototype detectors and the noise level is distributed uniformly over the strips.
- gain variations within single HELIX2.2 readout chip are in the order of 2% and do not influence the position reconstruction algorithm
- the intrinsic position resolution at normal angle of incidence $\sigma_{\text{DUT}}^{\text{intr}} = 7.2 \pm 0.2\ \mu$m is highly satisfactory if compared to the value $p/\sqrt{12} \simeq 35\,\mu$m which represents the limit for a digital system with a readout pitch p = 120 μm

- the position resolution at relatively large angle of incidence θ is still very satisfactory ($\lesssim 40\mu$m up to $\theta = 70°$)
- a position reconstruction algorithm which uses the rough knowledge of the angle of incidence θ to choose the most effective reconstruction procedure has been developed. For small incidence angles a 3-strips algorithm is applied whereas for large incidence angles a modified version of the head-tail algorithm is used
- a general position reconstruction algorithm, not relying on the knowledge of the angle of incidence, has proven to work well up to $\theta = 70°$ and could therefore be a valuable choice for a first position reconstruction in the track reconstruction procedure

Appendix A

Jet Energy Correction

The comparison of the reconstructed jet variables for jets defined at hadron level (referred to as **hadron jets**) and jets defined at detector level (referred to as **cal jets**) reveals that the transverse energy of the cal jets underestimates that of the hadron jets by an average amount of $\sim 17.7\%$ with an r.m.s. of $\sim 12.5\%$, as shown in the left plot of Fig. A.1[1]. The difference

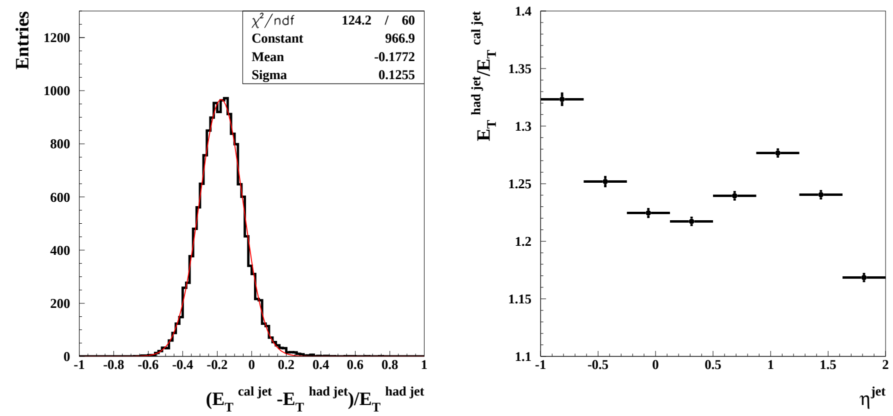

Figure A.1: Correlation between the transverse jet energy at hadron and detector level (left plot) and dependence of the ratio $E^{\text{ had jet}}/E^{\text{cal jet}}$ from the jet direction η^{jet} (right plot).

in the transverse jet energy is mainly due to the energy loss in the inactive material placed in front of the calorimeter UCAL. In the right plot of Fig. A.1 one can see that the difference in the transverse energy between hadron and cal jets depends on the jet direction. Two peaks in the regions $\eta^{\text{jet}} \sim -1$ and $\eta^{\text{jet}} \sim 1$ are present, which are related to the increase of the inactive material in front of the UCAL in those regions. Since no significant systematic shift in the angular variable η^{jet} and ϕ^{jet} is observed (see Fig. A.2), the correction procedure, applied in

[1]Each cal jet is associated to a had jet if their relative distance in the (η, ϕ) plane is smaller than 1.

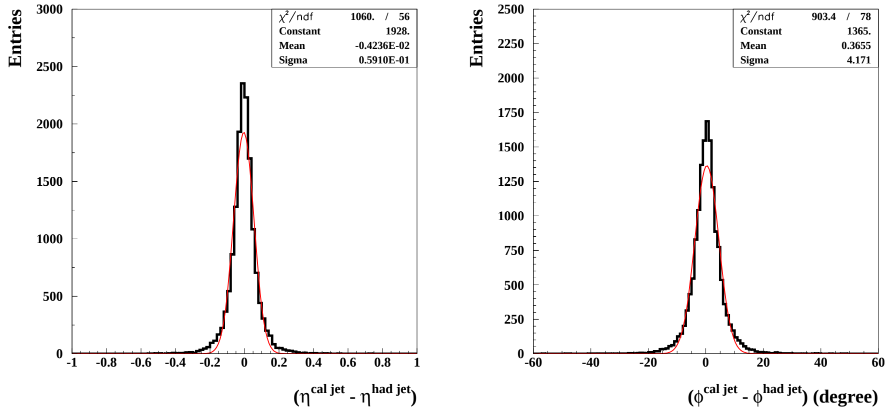

Figure A.2: Correlation between the angular jet variables ($\eta^{\rm jet}$, $\phi^{\rm jet}$) at hadron and detector level.

this analysis, is only needed to 'calibrate' the transverse jet energy measured at detector level. The corrected transverse jet energy is given by:

$$E_T^{\rm corr\ jet} = C(E_T^{\rm cal\ jet}, \eta^{\rm cal\ jet}) \cdot E_T^{\rm cal\ jet}$$

where the correction factors $C(E_T^{\rm cal\ jet}, \eta^{\rm cal\ jet})$ are obtained from a linear fit to the ($E_T^{\rm cal\ jet}$, $E_T^{\rm had\ jet}$) profile distributions in 12 different bins of the $\eta^{\rm cal\ jet}$ variable between $-1 \leq \eta^{\rm cal\ jet} \leq 2$ (as shown in Fig. A.3).

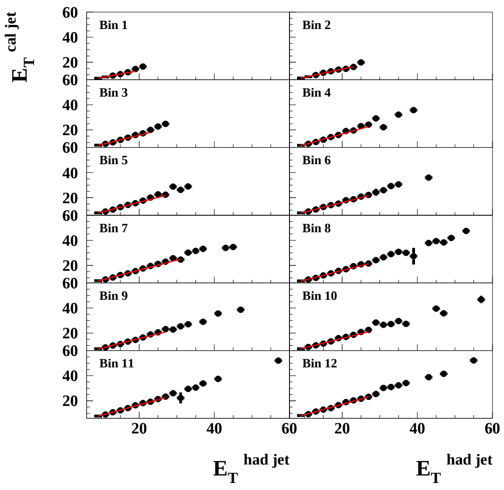

Figure A.3: ($E_T^{\rm cal\ jet}$, $E_T^{\rm had\ jet}$) profile distributions in 12 bins of the $\eta^{\rm cal\ jet}$ variable.

Appendix B

Correction Factors from Detector to Hadron Level

Figures B.1, B.2 and B.3 show the correction factors for the dijet sample and figures B.4, B.5 and B.6 the correction factors for the other jet in the D^*-tagged dijet sample, applied to the observables measured in this analysis. The deviations from unity for the observable $\langle n_{\rm sbj}\rangle$ are found to be small. The largest deviation for both the dijet sample and the other jet in D^*-tagged dijet sample is of the order of $\sim 12\%$. The correction factors for the observable $\Psi(r/R)$ differ from unity by less than 5% for subcone radii $r \geq 0.3$. For the measurement at $r = 0.1$ the deviation from unity is found to be large, especially at high $E_T^{\rm jet}$: $C(\Psi(0.1)) \simeq 63\%$ for the dijet sample and $C(\Psi(0.1)) \simeq 50\%$ for the other jet in D^*-tagged dijet sample, for $E_T^{\rm jet} \geq 15$ GeV. For the observable $\rho(r/R)$, the correction factors differ from unity typically by $\sim 10\%$ for subcone radii $r \geq 0.2$. At high $E_T^{\rm jet}$ the deviations become larger but on average always smaller than 20%.

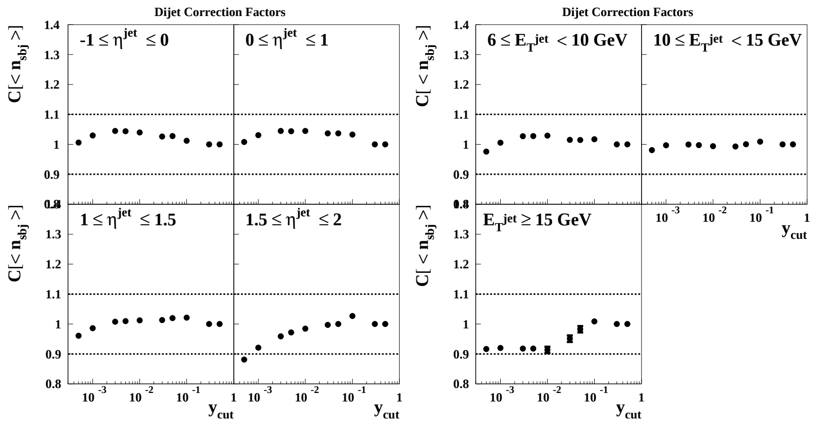

Figure B.1: Correction factors $C(\langle n_{\rm sbj}\rangle)(y_{\rm cut})$ in the four $\eta^{\rm jet}$ bins (left plot) and the three $E_T^{\rm jet}$ bins (right plot) for the dijet photoproduction sample, computed using the PYTHIA Monte Carlo.

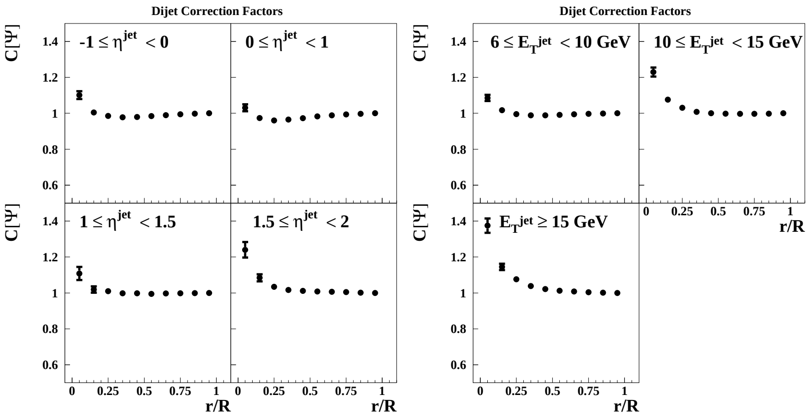

Figure B.2: Correction factors $C(\Psi)(r/R)$ in the four $\eta^{\rm jet}$ bins (left plot) and the three $E_T^{\rm jet}$ bins for the dijet photoproduction sample, computed using the PYTHIA Monte Carlo.

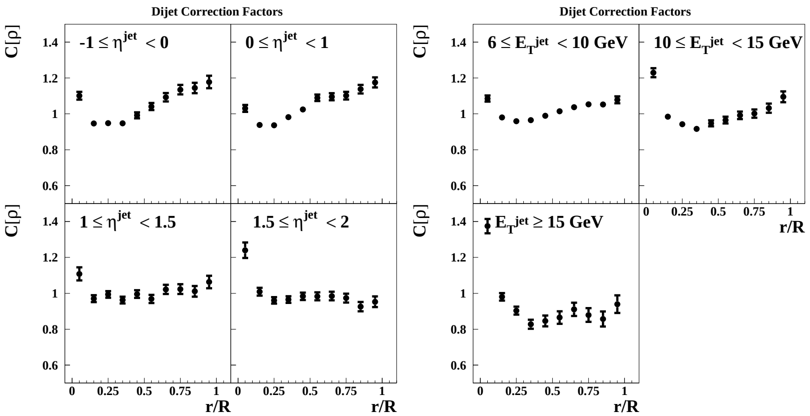

Figure B.3: Correction factors $C(\rho)(r/R)$ in the four $\eta^{\rm jet}$ bins (left plot) and the three $E_T^{\rm jet}$ bins for the dijet photoproduction sample, computed using the PYTHIA Monte Carlo.

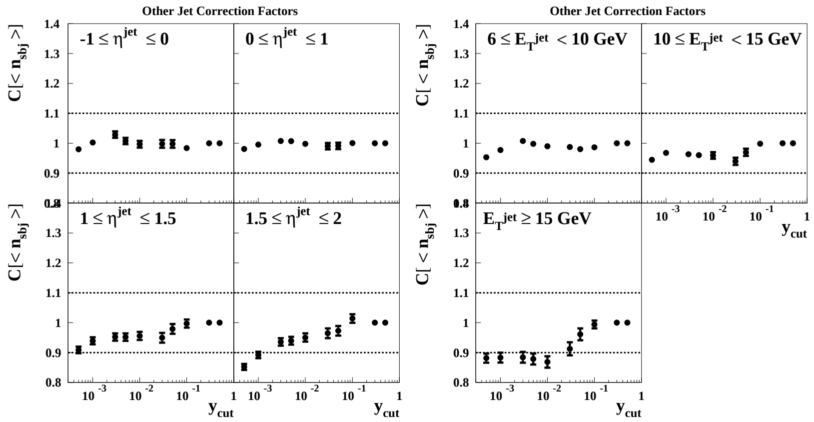

Figure B.4: Correction factors $C(\langle n_{\rm sbj}\rangle)(y_{\rm cut})$ in the four $\eta^{\rm jet}$ bins (left plot) and the three $E_T^{\rm jet}$ bins for the other jet in the D^*-tagged dijet photoproduction sample, computed using the PYTHIA Monte Carlo.

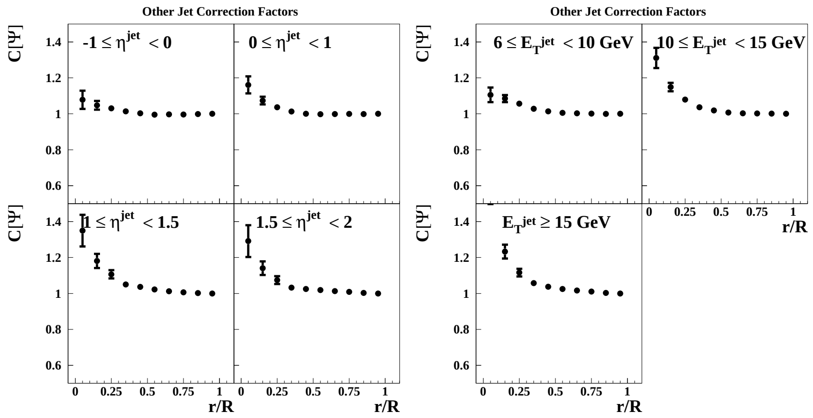

Figure B.5: Correction factors $C(\Psi)(r/R)$ in the four $\eta^{\rm jet}$ bins (left plot) and the three $E_T^{\rm jet}$ bins for the other jet in the D^*-tagged dijet photoproduction sample, computed using the PYTHIA Monte Carlo.

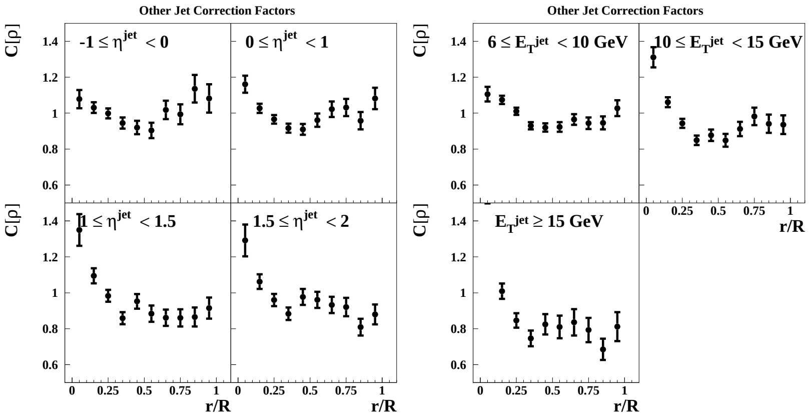

Figure B.6: Correction factors $C(\rho)(r/R)$ in the four $\eta^{\rm jet}$ bins (left plot) and the three $E_T^{\rm jet}$ bins for the other jet in the D^*-tagged dijet photoproduction sample, computed using the PYTHIA Monte Carlo.

Appendix C

Silicon Strip Detector Principles

Silicon atoms are arranged in a diamond-like face centred cubic lattice. Each silicon atom is surrounded by four other atoms in a tetrahedral configuration and shares its four valence electrons to form covalent bonds. At T = 0K, all valence electrons participate in the covalent bonding and no electrons are available for electrical conduction. Already at room temperature (kT $\simeq$ 25 meV), however, the action of the thermal vibration can break some bonds and electrons enter the conduction band, leaving a 'hole' in its original position. An electric current is thus generated due to the movement of free electrons in the conduction band and of holes in the valence band. In intrinsic silicon, the concentrations of free electrons ($\mathrm{n_e}$) and holes ($\mathrm{n_h}$), described by the Fermi-Dirac statistics, are equal:

$$\mathrm{n_e = n_h = n_i = \sqrt{N_c N_v}\exp\left(-\frac{E_g}{2kT}\right) = AT^{3/2}\exp\left(-\frac{E_g}{2kT}\right)} \tag{C.1}$$

where $\mathrm{N_c(T)}$ and $\mathrm{N_v(T)}$ are the effective densities of occupied states at the conduction and valence band edges, $\mathrm{E_g}$ is the energy gap between the two bands[1] at T = 0K and k is the Boltzmann constant.

The equilibrium density $\mathrm{n_i}$ can be changed by doping the intrinsic silicon with a small amount of impurity atoms having one more or less valence electron in their outer atomic shell (pentavalent or trivalent elements, respectively). In the case of a pentavalent dopant, only four electrons are needed for the covalent bonding and the fifth one resides in a discrete energy level created by the presence of the impurity atoms in the energy gap, very close to the conduction band. At normal temperature, these extra electrons can be easily excited into the conduction band or they can fill up holes, thereby decreasing the normal hole concentration. The current is then mainly due to the movement of electrons, the impurity material is called 'donor' (N_D being its concentration) and the doped silicon is referred to as **n-type** material. If the impurity is trivalent, there are not enough electrons to fill the valence band and thus an excess of holes in the crystal is created. Additional energy levels are created very close to the valence band, where valence electrons can be easily excited into, leaving extra holes behind. The normal electron concentration decreases, and the holes become the majority charge carriers. In this case the impurity material is called 'acceptor' (N_A being its concentration) and the doped silicon is referred to as **p-type** material[2].

[1] For silicon, $\mathrm{E_g(T = 0K)} \sim 1.21$ eV.
[2] Highly acceptor or donor doped materials ($\sim 10^{20}$ cm^{-3}) are usually indicated by $\mathrm{p^+}$ and $\mathrm{n^+}$.

Regardless of the type of dopant, in the case of thermal equilibrium, the mass action law applies:

$$n_e \cdot n_h = n_i^2 = AT^{3/2} \exp\left(-\frac{E_g}{2kT}\right) \tag{C.2}$$

The resistivity ρ of the doped silicon is given by:

$$\rho = \frac{1}{q \cdot (\mu_e n_e + \mu_h n_h)} \tag{C.3}$$

where μ_e and μ_h are the electron and hole mobility, respectively, and q the electric charge[3]. From the formula in Eq. C.3 it is clear that to achieve a very high resistivity, the doping concentration must be very low ($\rho > 2k\Omega$cm $\rightarrow N_{eff} < 10^{12}$ cm^{-3}, where $N_{eff} = N_D - N_A$ is the effective concentration).

In order to use silicon as a particle detector, nearly all free charge carriers have to be removed from the silicon bulk. This is achieved by creating so called p-n diode junctions. A p-n junction is a juxtaposition of a p-type silicon with an n-type silicon. The different free charge carrier concentrations lead to a diffusion current across the junction. Electrons diffuse to the p-side and recombine with holes, while the diffusing holes capture electrons on the n-side. Since the p-region is injected with extra electrons, it becomes negative while the n-region becomes positive. The static donor (positive) and acceptor (negative) ions produces an electric field gradient across the junction which counteracts the diffusion current. In equilibrium, diffusion currents and field induced current are equal and a 'built-in' potential difference, V_{bi}, establishes between the n-type and the p-type material. As a consequence a space charge region depleted of mobile charge carriers (**depletion region**) is formed. Note that the depletion region is the only ionization sensitive volume of a silicon detector, since only there an electric field is present. The width of the depletion region is generally small and depends on the concentration of donor and acceptor impurities. It can be determined from the Poisson's equation (if the charge density $\rho_c(x)$ is known):

$$-\frac{d^2V}{dx^2} = \frac{dE}{dx} = \frac{\rho_c(x)}{\epsilon_{Si}} \quad \text{with} \quad \rho_c = q \cdot [n_h(x) - n_e(x) + N_D(x) - N_A(x)] \tag{C.4}$$

where ϵ_{Si} is the dielectric constant. For an abrupt junction, one can write:

$$\frac{dE}{dx} = \frac{q}{\epsilon_{Si}} \cdot N_D \quad \text{for} \quad 0 < x \le x_n$$

$$\frac{dE}{dx} = -\frac{q}{\epsilon_{Si}} \cdot N_A \quad \text{for} \quad -x_p \le x < 0 \tag{C.5}$$

(C.6)

and by integrating:

$$E(x) = \frac{qN_D}{\epsilon_{Si}} \cdot (x - x_n) \quad \text{for} \quad 0 < x \le x_n$$

$$E(x) = -\frac{qN_A}{\epsilon_{Si}} \cdot (x_p + x) \quad \text{for} \quad -x_p \le x < 0 \tag{C.7}$$

(C.8)

[3] At T = 300K, $\mu_e = 1500$ cm^2/Vs and $\mu_h = 450$ cm^2/Vs, yielding a resistivity of $\rho = 219$ kΩcm

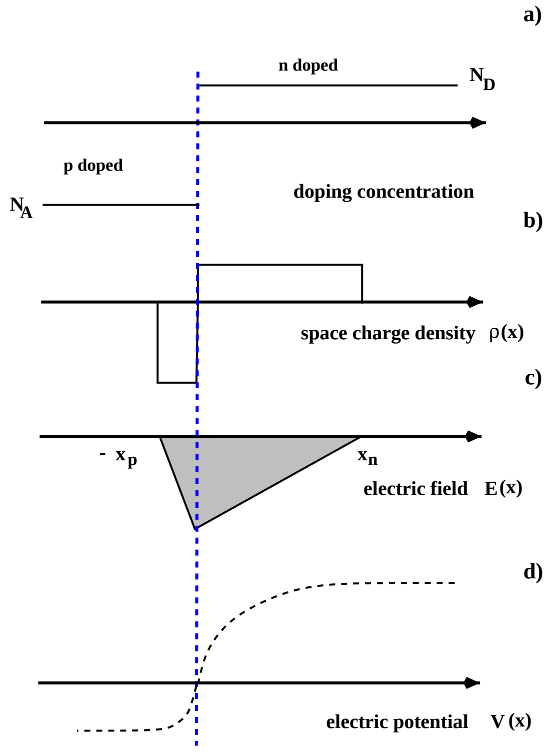

Figure C.1: Abrupt p-n junction in thermal equilibrium: a) doping profile, b) space charge density, c) electric field, d) electric potential.

where x_n and x_p are the extent of the depletion zone on the n-side and p-side respectively (see Fig. C.1). The built-in potential V_{bi} is given by:

$$V_{bi} = -\int_{-x_p}^{x_n} E(x)dx = \frac{1}{2}E(x=0) \cdot W \quad \text{where} \quad W = (x_n + x_p)$$
$$\text{with} \quad E(x=0) = \frac{qN_D x_n}{\epsilon_{Si}} = \frac{qN_A x_p}{\epsilon_{Si}} \tag{C.9}$$

Combining the last two equations, one obtains:

$$x_p = \sqrt{\frac{2\epsilon_{Si}V_{bi}}{qN_A(1 + N_A/N_D)}} \quad \text{and} \quad x_n = \sqrt{\frac{2\epsilon_{Si}V_{bi}}{qN_D(1 + N_D/N_A)}} \tag{C.10}$$

(C.11)

and thus the total depth of the depletion region is given by:

$$W = x_p + x_n = \sqrt{\frac{2\epsilon_{Si}V_{bi}}{q}\frac{(1 + N_D/N_A)}{N_D}} \tag{C.12}$$

For a p^+-n junction ($N_A \gg N_D$), this leads to a deep depletion region on the n-side, which is the active particle detection volume, and a very shallow one on the p-side. The shallow p^+ implant represent the measuring electrode and is required for the depletion of the n-side material.
By applying a reverse bias voltage, V_{bias}, to the diode (positive voltage to the n-side and negative voltage to the p-side) the depletion region can be enlarged. Its width, W, can be calculated from the formula in Eq. C.12 by replacing V_{bi} with (V_{bi}+V_{bias}):

$$W = \sqrt{2\epsilon_{Si}\rho\mu_e(V_{bias} + V_{bi})} \quad \text{with} \quad \rho = \frac{1}{q\mu_e N_D} \tag{C.13}$$

The voltage, V_{dep}, necessary to extend the depletion region over the entire detector bulk is thus:

$$V_{dep} = \frac{t^2}{2\,\epsilon_{Si}\mathrm{Si}\,\rho\,\mu_e} \tag{C.14}$$

where t is the detector thickness.

Bibliography

[1] S.J. Brodsky and J.F. Gunion, *Phys. Rev. Lett.* 37 (1976) 402;
K. Konishi, M. Massoth, A. Ukawa and G. Veneziano, *Phys. Rev. Lett.* 78 (1978) 243

[2] R.P. Feynman, *Phys. Rev. Lett.* 23 (1969) 1415

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

[4] V.N. Gribov and L.N. Lipatov, *Sov. J. Nucl. Phys.* 15 (1972) 438;
G. Altarelli and G. Parisi, *Nucl. Phys.* B126 (1977) 298;
Y.L. Dokshitzer, *Sov. Phys. JETP* 46 (1977) 641

[5] A.H. Mueller, *Phys. Lett.* 104B (1981) 161;
B.I. Ermolayer and V.S. Fadin, *JETP Lett.* 33 (1981) 269;
A. Bassetto, M. Ciafaloni, G. Marchesini and A.H. Mueller, *Nucl. Phys.* B207 (1982) 189

[6] G. Hanson et al., *Phys. Rev. Lett.* 35 (1975) 1609

[7] G. Sterman and S. Weinberg, *Phys. Rev. Lett.* 39 (1997)

[8] JADE Collaboration, W. Bartel et al., *Z. Phys.* C33 (1986) 23;
S. Bethke et al., *Phys. Lett.* B213 (1988) 235

[9] S. Catani, Yu.L. Dokshitzer and B.R. Webber, *Phys. Lett.* 285 (1992) 291;
S. Catani, Yu.L. Dokshitzer M.H. Seymour and B.R. Webber, *Nucl. Phys.* B406 (1993) 187

[10] S.D. Ellis and D.E. Soper, hep-ph/9305266

[11] OPAL Collaboration, G. Abbiendi et al., *Eur. Phys. J.* C17 (2000) 373;
ALEPH Collaboration, D. Buskulic et al., *Z. Phys.* C 76 (1997) 191

[12] Yu.L. Dokshitzer, V.A. Khoze, A.H. Mueller and S.L. Troyan, *Basics of Perturbative QCD* (Editions Frontieres, Gif-sur-Yvette) 1991

[13] OPAL Collaboration, G. Abbiendi et al., *Eur. Phys. J.* C11 (1999) 217

[14] OPAL Collaboration, G. Alexander et al., *Z. Phys.* C69 (1996) 543

[15] F. Halzen and A.D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, John Weley and Sons, New York 1984

[16] H.L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J.F. Owens, J. Pumplin and W.K. Tung, *Eur.Phys.J.* C12 (2000) 375

[17] T.H. Bauer, R.D. Spital and D.R. Yennie, *Rev. Mod Phys.* 50 (1978) 261

[18] ZEUS Collaboration, *The ZEUS detector, Status Report 1993*, U. Holm ed. (1993)

[19] *HERA: A Proposal for a Large Electron Proton Colliding Beam Facility at DESY* (Hamburg, 1981), DESY-HERA-81/10

[20] *The HERA Luminosity Upgrade* DESY-HERA 98-05 (1998)

[21] ZEUS Collaboration, J. Breitweg et al., DESY-PRC 97/01 (1997)

[22] A. Andresen et al. (ZEUS Calorimeter group), *Nucl. Instr. and Meth.* A309 (1991) 101;
M. Derrick et al., *Nucl. Instr. and Meth.* A309 (1991) 77;
A. Bernstein et al., *Nucl. Instr. and Meth.* A336 (1993) 23

[23] H. Bethe and W. Heitler, *Proc. Roy. Soc.* A146 (1934) 83

[24] ZEUS Collaboration, M. Derrick et al, *Phys. Lett.* B316 (1993) 412;
H1 Collaboration, I. Abt et al, *Nucl. Phys.* B407 (1993) 515;
ZEUS Collaboration, M. Derrick et al, *Z. Phys.* C65 (1995) 379

[25] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, *Comput. Phys. Commun.* 67, 465 (1992);
G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, hep-ph/9607393

[26] C.F. von Weizsacker, *Z. Phys.* 88 (1934) 612;
E.J. Williams, *Phys. Rev.* 45 (1975) 729;
S. Frixione,M.L. Mangano,P. Nason and G. Ridolfi, *Phys. Lett.* B 319 (1993) 339

[27] G. Altarelli and G. Parisi, *Nucl. Phys.* B126 (1997) 298

[28] Yu.L. Dokshitzer, V.A. Khoze, S.I. Troyan, *J. Phys.* G17 (1991) 1481

[29] A. Kupco, hep-ph/9906412

[30] H.U. Bengtsson and T. Sjöstrand, *Comput. Phys. Commun.* 46 (1987) 43;
T. Sjöstrand, *Comput. Phys. Commun.* 82 (1994) 74

[31] J.C. Collins, D.E. Soper and G. Sterman, *Nucl. Phys.* B308 (1988) 833

[32] T. Sjöstrand, *Nucl. Phys.* B248 (1984) 469

[33] GEANT 3.13: R. Brun et al., CERN DD/EE/84-1 (1987)

[34] F. Jaquet and A. Blondel, *Proceedings of the study of an ep facility for Europe* DESY-79-48 (1979) 391

[35] H. Abramowicz, A. Caldwell, R. Sinkus, *Neural Network Based Electron Identification in the ZEUS Calorimeter*, DESY 95-054;
R. Sinkus, *Nucl. Instr. Meth.* A361 (1995) 290

[36] ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* C6 (1999) 67

[37] C. Caso et al., (Particle Data Group). Review of particle physics, *Eur. Phys. J.* C3 (1998) and 1999 off-year partial update for the 2000 edition available on the PDG WWW pages (URL: http://pdg.lbl.gov/)

[38] L. Gladilin, *Charm Hadron Production Fractions*, hep-ph/9912064

[39] OPAL Collaboration, G. Alexander et al., *Z. Phys.* C72 (1996) 1;
ALEPH Collaboration, R. Barate et al., CERN/EP 99-094, hep-ex/9909032, submitted to *Eur. Phys. J.* C;
DELPHI Collaboration, P. Abreu et al., *Eur. Phys. J.* C12 (2000) 209

[40] G.J. Feldman et al., *Phys. Rev. Lett.* 38 (1977) 1313

[41] S. Nussinov, *Phys. Rev. Lett.* 35 (1977) 1672

[42] ZEUS Collaboration, M. Derrick et al., *Phys. Lett.* B349 (1995) 225;
ZEUS Collaboration, J. Breitweg et al., *Phys. Lett.* B407 (1997) 402;
ZEUS Collaboration, J. Breitweg et al., *Phys. Lett.* B401 (1997) 192;
ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* C6 (1999) 67;
ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* C12 (2000) 35

[43] ALEPH Collaboration, R. Barate et al., *Phys. Rep.* 294 (1998) 1

[44] H1 Collaboration, C. Adloff et al., *Nucl.Phys.* B545 (1999) 3;
ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* C2 (1998) 61

[45] ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* C2 (1998) 61

[46] ZEUS Collaboration, *Determination of α_S using subjet multiplicities in NC-DIS*, paper submitted to the 9th International Workshop on Deep Inelastic Scattering (2001)

[47] ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* C2 (1998) 61

[48] ZEUS Collaboration, *Measurements of the Jet Substructure in Photoproduction at HERA*, paper 530 submitted to the International Europhysics Conference on HEP, Tampere (1999)

[49] U. Koetz et al., *Nucl. Instrum. and Methods* A235 (1985) 481

[50] D. Dannheim, *The silicon strip detectors of the ZEUS microvertex detector*, Diplomaarbeit, University Hamburg;
A. Garfagnini, *Nucl. Instrum. Methods* A435 (1999) 34;
C. Coldewey, *Nucl. Instrum. Methods* A447 (2000) 44

[51] S.M. Sze, *Physics of semiconductor devices*, John Wiley & Sons, 1981

[52] M. Feuerstack-Raible, *Overview ot microstrip readout chips*, *Nucl. Instr. and Meth.* A447 (2000), Proceedings VERTEX99;
W. Fallot-Burghardt et al., HELIX128.2x User Manual V2.1,03/02/1999, HD-ASIC-33-0697.

[53] Austria Mikrosysteme International GmbH

[54] LabView, product of National Instruments

[55] C. Colledani et al., *Nucl. Instr. and Meth.* A372 (1996) 379

[56] VA2 chip product of the VA readout chip series of IDE AS, Norway

[57] products of CAEN, Italy

[58] I. Redondo, *Charm Electroproduction at HERA*, PhD thesis in preparation

[59] A.S. Schwarz, *Phys.Rep* 238 (1994) 1

[60] E. Belau et al., *Nucl. Instr. and Meth.* 214 (1983) 253;
R.Turchetta, *Nucl. Instr. and Meth.* A335 (1993) 44

[61] J. Martens, *Simulationen und Qualitätssicherung der Siliziumstreifendetektoren des ZEUS-Mikrovertexdetektors*, Diplomaarbeit, University Hamburg

[62] M. Moritz, PhD thesis in preparation

[63] R. Turchetta, *Nucl. Instr. and Meth.* A335 (1993) 44;
KEK Preprint 96-172, *Measurement of the Spatial Resolution of Wide-pitch Silicon Strip Detectors with Large Incident Angle*