Measurement of Charm Fragmentation in DIS at HERA

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von

Zuzana Rúriková

aus Bratislava

angefertigt am
Max-Planck-Institut für Physik, München
(Werner-Heisenberg-Institut)

Hamburg
2005

Gutachter der Dissertation:	Prof. Dr. B. Naroska Prof. Dr. C. Hagner
Gutachter der Disputation:	Prof. Dr. G. Heinzelmann Dr. H. Jung
Datum der Disputation:	02.03.2006
Vorsitzender des Prüfungsausschusses:	Dr. H. D. Rüter
Vorsitzender des Promotionsausschusses:	Prof. Dr. G. Huber
Dekan des Fachbereichs Physik:	Prof. Dr. G. Huber

Abstract

The charm fragmentation function has been studied in deep inelastic ep collisions with the H1 detector at HERA. The data were analyzed in the kinematic range given by the photon virtuality $2 < Q^2 < 100$ GeV^2 and the inelasticity $0.05 < y < 0.7$.

The differential cross section of D^*-meson production as a function of two scaling variables sensitive to fragmentation, $\mathrm{z}_{\mathrm{hem}}$ and $\mathrm{z}_{\mathrm{jet}}$, was measured in the D^*-meson visible range defined by cuts on pseudorapidity $|\eta(D^*)| < 1.5$ and transverse momentum $1.5 < p_{\mathrm{t}}(D^*) < 15$ GeV. In the case of the $\mathrm{z}_{\mathrm{hem}}$ observable, the momentum of the charm quark has been approximated by the momentum of a suitably defined hemisphere in the photon-proton center-of-mass frame, whereas in the case of the $\mathrm{z}_{\mathrm{jet}}$ observable, the charm quark momentum was approximated by the momentum of the reconstructed jet associated with the D^*-meson. A transverse energy $E_{\mathrm{t}}^{\gamma p}(D^*\mathrm{jet}) > 3$ GeV was required in the latter case.

Both distributions were used to extract the parameters of the non-perturbative Peterson and Kartvelishvili fragmentation functions for the Monte Carlo models RAPGAP/PYTHIA and CASCADE/PYTHIA and for the next-to-leading-order massive QCD calculation HVQDIS.

Kurzfassung

In dieser Arbeit wird das Studium der Charm-Fragmentations-Funktion anhand der H1-Daten in tief unelastischer Streuung beschrieben. Die analysierten Daten liegen in dem kinematischen Bereich, der von dem negativen Viererimpulsübertrag zum Quadrat $2 < Q^2 < 100$ GeV^2 und der Unelastizität $0.05 < y < 0.7$ festgelegt wird.
Der differentielle Wirkungsquerschnitt der D^*-Produktion wird in Abhängigkeit von zwei Skalenvariablen die sensitiv auf die Fragmentation sind, $\mathrm{z_{hem}}$ und $\mathrm{z_{jet}}$, in dem sichtbaren Bereich der D^*-Produktion gemessen. Dieser Bereich wird definiert durch die Pseudorapidität $|\eta(D^*)| < 1.5$ und dem Transversalimpuls $1.5 < p_\mathrm{t}(D^*) < 15$ GeV. Für die Observable $\mathrm{z_{hem}}$ wird der Impuls des Charm Quarks anhand einer geeignet definierten Hemisphäre im Photon-Proton Schwerpunktsystem abgeschätzt. Bei der Observablen $\mathrm{z_{jet}}$ wird der Impuls des Charm Quarks aus dem Viererimpuls des rekonstruierten Jets definiert, der das D^*-Meson beinhaltet. Für den Jet wird eine Transversalenergie von $E_\mathrm{t}^{\gamma p}(D^*\mathrm{jet}) > 3$ GeV gefordert.
Aus der Verteilung der Observablen werden die Parameter für die nicht perturbativen Fragmentations-Funktionen von Peterson und Kartvelishvili für die Monte-Carlo-Modelle RAPGAP/PYTHIA und CASCADE/PYTHIA und für eine massive QCD-Rechnung (HVQDIS) in nächst höherer Ordnung extrahiert.

Contents

Introduction

During the last decades there has been large progress achieved in the field of high energy physics, which led to the formulation of the standard model. The standard model describes all known physics phenomena except gravity. It includes the theory of electro-weak interactions and the theory of strong interactions, also called quantum chromodynamics (QCD). The theoretical predictions, for both theories, are usually based on perturbative calculations. However in QCD, due to the nature of the strong coupling, only processes with large momentum transfer can be calculated perturbatively, which limits the predictive power of this theory.

An interesting opportunity for testing QCD offer the processes which involve production of heavy quarks (charm and beauty), since the heavy quark mass already provides a hard scale. This hard scale allows to extend the range of applicability of perturbative calculations. Due to confinement, single partons cannot be directly observed, as they are bound within hadrons. This introduces a large uncertainty in the comparison of measured data on exclusive production of charmed hadrons with theoretical predictions.

The transition of a heavy quark to a heavy meson (often referred to as fragmentation) is usually described by phenomenological models, since it is not entirely perturbatively calculable. One of the major characteristics of this process is the energy fraction which is transferred from the quark to the created hadron, which is given by the non-perturbative fragmentation function.

According to the factorization theorem, the non-perturbative fragmentation function should be universal, i.e. independent of the charm production mechanism. Up to now, the parameters of phenomenological models for fragmentation have been almost exclusively determined from e^+e^- annihilation experiments, and are widely used in ep and pp collider experiments.

However, the factorization theorem has been proven only for a few processes, since the mathematical apparatus is very demanding. Apart from the question of theoretical proof, it is important to experimentally test whether universality of the fragmentation function really holds.

In this thesis the charm fragmentation function has been measured in deeply inelastic *ep* collisions at the H1 experiment. The charm events are tagged by reconstructing a $D^{*\pm}$-meson via one of its decay channels. Since the choice of a suitable observable sensitive to the fragmentation function is not so straight forward in *ep*-scattering than for example in e^+e^- annihilation, two different definitions have been investigated in parallel: the first in close analogy to e^+e^- - the hemisphere observable, the other making use of reconstructed jets - the jet observable.

The thesis is organized as follows. In the first chapter the theoretical foundations are given. In the second chapter the experimental methods for the measurement of fragmentation are introduced. A short description of the H1 detector with main focus on detector components relevant for this analysis is given in chapter 3. Then the data selection, reconstruction of event kinematics, $D^{*\pm}$-meson tagging as well as the measurement of the fragmentation spectra are discussed in chapters 4 and 5. In chapter 6 the parameters of the non-perturbative fragmentation functions are extracted for different Monte Carlo models and next-to-leading order calculations.

Chapter 1

Theoretical Framework

1.1 Quantum Chromodynamics

Quantum chromodynamics (QCD) is the theory of the strong interaction, one of the four fundamental forces in nature. It describes the interaction between quarks and gluons and at the same time gives a hint for their somewhat unconventional behavior of being bounded in the hadrons and not manifesting themselves directly as other elementary particles (e.g. leptons) do. QCD was developed in the 1970s as a field theory, where interaction between quarks is mediated by a massless spin=1 boson, the gluon. The partons (quarks and gluons) are found to have an additional degree of freedom called color charge. The color charge is analogous to the electric charge, except that it has three values rather than one (r=red, g=green and b=blue). Quarks carry only one type of color whereas the gluons are bicolored, carrying one color and one anticolor.

The major difference between QCD and quantum electrodynamics (QED) is that gluons can couple to each other since they carry color, whereas in QED the photons are chargeless. The consequence of gluon self-coupling is that the strong coupling α_s 'runs' in the opposite direction to the electromagnetic coupling and becomes very large at low momenta, ensuring that quarks and gluons are confined within color singlet hadrons.

In principle QCD can completely predict the behavior of partons on the basis of equations of motion which can be obtained from the QCD Lagrangian. But in practice this is far from being the case, as these equations cannot be solved directly. One of the frequently used techniques, which leads to testable predictions is the use of perturbation theory. For QCD processes involving exchanges of sufficiently large momenta (hard subprocesses) α_s is small enough for the perturbative approach to be valid. The problem then is to extend the domain of validity to soft processes, where the involved momenta are smaller and α_s becomes larger. To do this requires information on the dynamics of partons inside the interacting hadrons as well as how partons emerging from the hard subprocess turn into a jet of observable hadrons. Appropriate descriptions of these two soft aspects of QCD are expected to be universal according to the factorization theorem. After having been measured in simple cases, they can be applied to all other QCD processes.

1.2 Deep-Inelastic Scattering

Deep-inelastic scattering (DIS) plays an important role in our understanding of the structure of matter. The proton structure can be investigated using lepton-nucleon scattering where the boson emitted by the pointlike lepton acts as a probe of the nucleon. This provides a much simpler probing mechanism than in nucleon-nucleon collisions in which both participating particles are compound objects and thus the determination of event kinematics gets more difficult.

In DIS experiments at HERA, where electrons or positrons collide with protons at high energy, the lepton (l) interacts in leading order with one of the quarks in the proton by exchange of a virtual gauge boson (see figure 1.1). Depending on the exchanged boson, two major processes can occur. The process, $ep \to eX$, where the exchanged boson does not carry any electric charge (virtual photon or Z^0 boson), is classified as *neutral current* (NC). The second process, $ep \to \nu X$, which leads to the exchange of an electrically charged boson ($W^{\pm}$) and a different final state (a neutrino instead of a scattered electron) is classified as *charged current* (CC). In both cases X stands for the hadronic final state which includes the proton remnant, corresponding to the remainder of the proton which did not actively take part in the hard interaction.

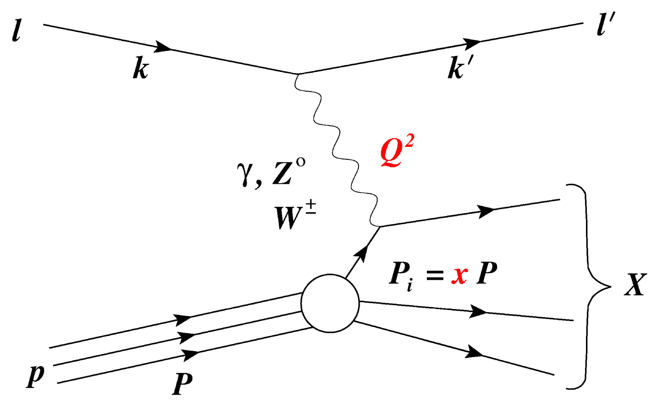

Figure 1.1: Feynman diagram of deep inelastic ep scattering.

1.3 DIS Kinematics

The kinematics of a DIS event can be parametrized in terms of Lorentz invariant quantities s, Q^2, x, y and W defined as follows:

- The center-of-mass energy squared

$$s = (k + P)^2 = 4E_e E_p + m_e^2 + m_p^2 \approx 4E_e E_p \tag{1.1}$$

  is determined by the energies of the colliding beams E_e and E_p (in this analysis it has a value of 101568 GeV^2 with $E_e = 27.6$ GeV and $E_p = 920$ GeV).

- The negative square of the four-momentum transfer or virtuality

$$Q^2 = -q^2 = -(k' - k)^2 \tag{1.2}$$

  represents the mass squared of the virtual boson. Q determines the hardness or in other words the transverse resolving power of the interaction.

- The Bjorken scaling variable

$$x = \frac{Q^2}{2Pq} \qquad (0 \leq x \leq 1) \tag{1.3}$$

can be interpreted in leading order as the fraction of the proton momentum carried by the struck parton in the infinite momentum frame of the proton.

- The inelasticity

$$y = \frac{qP}{kP} \qquad (0 \leq y \leq 1) \tag{1.4}$$

is the relative energy transfer from the lepton to the hadronic system in the proton rest frame.

- The energy of the boson-proton center-of-mass system

$$W^2 = (q + P)^2 = Q^2(\frac{1}{x} - 1) + m_{\mathrm{p}}^2 \approx ys - Q^2 \tag{1.5}$$

is equal to the invariant mass of the hadronic final state X.

Neglecting the electron and proton rest masses, the kinematic variables are related by $Q^2 = sxy$. Hence for fixed s the kinematics of the inclusive scattering process can be completely described by any set of two independent variables out of Q^2, x, y and W.

On the basis of the virtuality of the photon one distinguishes between two kinematic regimes. Interactions with small momentum transfer squared, $Q^2 \approx 0$, where the exchanged photon is almost real are called photoproduction processes. With increasing momentum transfer the wavelength of the virtual photon decreases ($\lambda \sim 1/Q$). For Q^2 values typically greater than 1 GeV^2 the wavelength becomes smaller than the proton size, making it possible to resolve its internal structure; this regime is called DIS.

1.4 Proton Structure

Neglecting the Z^0 boson exchange contribution which is insignificant at photon virtuality below 1000 GeV^2, the neutral current double differential cross section for ep scattering (see e.g. [34]) can be written as a function of two structure functions $F_1(x, Q^2)$ and $F_2(x, Q^2)$:

$$\frac{d^2\sigma_{\mathrm{NC}}}{dxdQ^2} = \frac{4\pi\alpha_{\mathrm{em}}^2}{xQ^4}\left[xy^2F_1(x, Q^2) + (1 - y)F_2(x, Q^2)\right], \tag{1.6}$$

where $\alpha_{\rm em}$ is the electromagnetic coupling, and the quantities x, y and Q^2 have been defined in the previous subsection. The structure functions are interpreted as being related to the momentum distribution of the charged partons within the proton, but they do not explicitly contain knowledge about the nature of the interaction.

Since the exchanged virtual photon is massive ($Q^2 > 0$), it can have all three helicity states, $\lambda = 0, \pm 1$. Hence the photo-absorption cross section can be written as a sum of two contributions, σ_T and σ_L, which arise from transversely ($\lambda = \pm 1$) and longitudinally ($\lambda = 0$) polarized photons respectively. It is found that $2xF_1$ is proportional to σ_T and that F_2 is proportional to $\sigma_L + \sigma_T$. Introducing the longitudinal structure function

$$F_L = F_2 - 2xF_1, \tag{1.7}$$

which is related to the exchange of longitudinally polarized photons, equation 1.6 can be rewritten as

$$\frac{d^2\sigma_{\rm NC}}{dxdQ^2} = \frac{4\pi\alpha_{\rm em}^2}{xQ^4}\left[(1 - y + \frac{y^2}{2})F_2(x,Q^2) - \frac{y^2}{2}F_{\rm L}(x,Q^2)\right] \tag{1.8}$$

The structure functions F_2 is then extracted by measuring the differential cross sections depending on x and Q^2. $F_{\rm L}$ can be determined by measuring in addition the dependence on y at fixed x and Q^2.

1.5 Quark Parton Model

The quark parton model (QPM) [30] was used to interpret the first measured data of inelastic ep scattering at the end of the sixties. In this model the proton is composed of charged pointlike spin 1/2 partons. For processes with large Q^2 where the hard interaction time scale is much shorter than the interaction between the partons, the partons can be considered as non-interacting. Therefore γ scatters incoherently, i.e. elastically on one of the partons of the proton. In QPM the internal structure of the proton can be expressed in terms of parton density functions $f_i(x)$ (PDFs), which denote the probability to find a parton of type i inside the proton, carrying fraction x of the proton's momentum. The structure functions can then be written as

$$F_2(x) = \sum_{i=1}^{n_{\rm f}} e_i^2 x f_i(x) \qquad\qquad F_{\rm L}(x) = 0 \tag{1.9}$$

The sum runs over all n_f quark flavors and e_i is the electric charge of the i-th flavor. In this 'naive' model the structure function depends only on Bjorken x and not on Q^2.[1]

Although it was found in those early days of deep-inelastic scattering experiments that the QPM described data quite well, it could not explain why the quarks were bound inside the proton. Also the assumption that the proton consisted only of charged quarks was indirectly proven to be false, since the sum over momenta of charged partons was found to be approximately 50% of the proton's total momentum. The missing momentum had to be therefore carried by neutral partons - the gluons, for which later at the e^+e^- storage ring PETRA at DESY convincing evidence was found.

By introducing gluon radiation into the QPM picture, the scaling behavior of the structure functions is violated as they now depend on both x and Q^2. The quarks in the proton can radiate gluons, and the gluons themselves can radiate gluons or split into quark-antiquark pairs as illustrated in figure 1.2. These properties lead to scaling violation, but also to divergencies in cross section calculations.

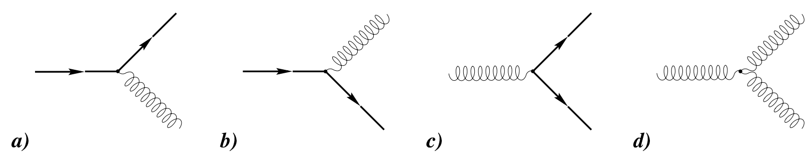

Figure 1.2: Feynman diagrams for gluon radiation processes: a) $q \to qg$, b) $q \to gq$, c) $g \to q\bar{q}$ and d) $g \to gg$.

1.6 Ultraviolet and Infrared Divergencies

One should distinguish between two types of divergencies, namely the ultraviolet and the infrared divergence. The ultra violet divergencies arising from virtual loop corrections can be absorbed into the running of the strong coupling $\alpha_s(\mu_r)$, which depends on the renormalization scale μ_r given by the 'hard' scale of the interaction. In the lowest order α_s is given by

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f)\log(Q^2/\Lambda^2_{QCD})}, \qquad \mu_r = Q \tag{1.10}$$

[1]This property is called scaling (Bjorken scaling).

where $\Lambda_{\mathrm{QCD}} \approx 0.2$ GeV is an experimentally determined parameter, and n_{f} denotes the number of quark flavors considered in virtual loop corrections to the gluon propagator. For $Q^2 \gg \Lambda^2_{\mathrm{QCD}}$ the coupling is small, and thus the interaction is asymptotically free. With decreasing hard scale $Q^2 \rightarrow \Lambda_{\mathrm{QCD}}$, the coupling becomes large (which presumably leads to confinement) and perturbative expansions cannot converge.

In case of real gluon radiation the cross section can be divergent in two regions. The case where the emitted gluon is moving in the direction of the outgoing quark is referred to as collinear divergence, while soft divergence refers to the case where the emitted gluon is soft, i.e. has low energy. It can be shown that both cases are of the infrared type which means that they involve long distances.[2] The long distance part is mostly sensitive to the partonic structure of hadrons and the hadronization mechanism. Unfortunately for infrared divergencies there does not exist such an elegant solution like the renormalization in the case of ultraviolet divergencies.

1.7 Factorization and Evolution Equations

Inspite of the plague of infrared divergencies it is possible to provide theoretical predictions within a perturbative quantum chromodynamics (pQCD) scope: the infrared divergencies at long distances can be separated from the hard interaction which is perturbatively calculable. This mechanism is called *factorization.* According to the factorization theorem the long distance part still contains infrared singularities but has the advantage of being universal (independent of the type of hard subprocess).

The 'final state' infrared divergencies[3] can be absorbed into non-calculable fragmentation functions (FFs). They are used to parametrize the transition from partons to hadrons (see section 1.8 for more details). On the other hand, 'initial state' infrared divergencies[4] are absorbed into parton density functions (section 1.5). Schematically the total cross section for a production of specific hadron h can then be written as a convolution of PDFs, hard

[2] From perturbative point of view the cross-section can be split into a short distance part (hard scale involved), which is infrared and collinear safe, and a long distance part (soft scale involved) which contains infrared and collinear singularities.

[3] This refers to gluon radiation from partons coming out of the hard subprocess.

[4] This refers to gluon radiation from partons entering the hard subprocess.

partonic cross section $\hat{\sigma}$ and fragmentation function D^h:

$$\sigma = \sum_i \sum_k f_i(x, \mu_\mathrm{f}) \otimes \hat{\sigma}_{i\gamma \to kX}(\alpha_\mathrm{s}(\mu_\mathrm{r}), \mu_\mathrm{f}) \otimes D_k^h(z, \mu_\mathrm{f}), \tag{1.11}$$

where the index i runs over all partons entering the hard subprocess, and the index k runs over partons emerging from hard subprocess. The long distance part as well as the hard subprocess cross section $\hat{\sigma}$ depend on the choice of factorization scale μ_f (similar to μ_r) and the factorization scheme. Once the PDFs and FFs are known at some starting scale μ_0, they can be extended to a different scale by the QCD evolution equations. The evolution equations are derived considering the possible gluon radiation and gluon splitting processes. Such process can occur several times which can lead to a gluon ladder in case of initial state radiation as shown on figure 1.3 (left).

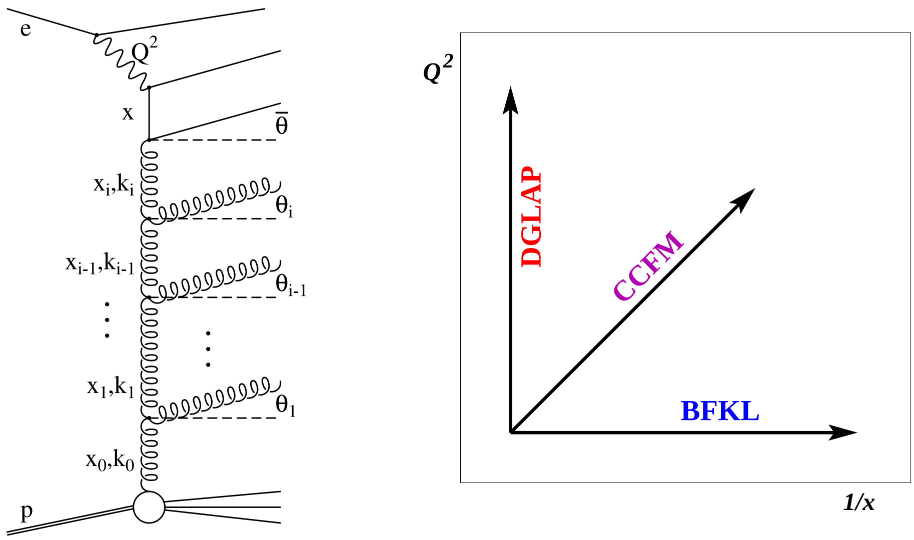

Figure 1.3: Left: gluon emissions in parton evolution, the so called gluon ladder. Right: a schematic overview of the differences in the DGLAP, BFKL and CCFM evolution schemes. The direction of particular evolutions are drawn in the $1/x - Q^2$ plane.

For the determination of PDFs there exist several approaches of summing the contributions from all ladder type diagrams. They are based on different assumptions concerning the ordering of fractional longitudinal momenta x_i, the transverse momenta $k_{\mathrm{t},i}$, which are related to the virtualities k_i^2 and the radiation angles θ_i of radiated partons. The DGLAP, BFKL and CCFM evolutions are described in the following.

DGLAP In this approach [35] the partons in the gluon ladder follow a strong ordering in transverse momenta $k_{\mathrm{t},i}^2 \gg k_{\mathrm{t},i-1}^2 \gg k_{\mathrm{t},i-2}^2 \gg ...$, while for the fractional longitudinal momenta there is only a soft ordering $x_i < x_{i-1} < x_{i-2} < ....$ The DGLAP formalism is applicable only, if the longitudinal momenta xP are larger than the transverse momenta, which is only valid in the region of large Q^2 and intermediate to large x.

BFKL In the regime of very low x and moderate Q^2 the BFKL approach [41] is more appropriate. In contrast to the DGLAP formalism, the partons in the ladder follow a strong ordering in fractional longitudinal momentum $x_i \ll x_{i-1} \ll x_{i-2} \ll ....$, while the transverse momenta k_t are not ordered.

CCFM The CCFM approach [19] attempts to unify the treatment for both the x and Q^2 dependencies of the parton distributions throughout the kinematic plane. According to CCFM, the emission of gluons is only allowed in the angular ordered region of phase space. At small x the CCFM evolution equation is 'equivalent' to BFKL and at large x to DGLAP evolution.

The difference between DGLAP, BFKL and CCFM approach is shown schematically in figure 1.3 (right). DGLAP predicts the Q^2 evolution of the structure function (parton distribution functions measured as a function of x at some starting value Q_0^2 can be evolved in Q^2), BFKL evolution in x, whereas CCFM allows to evolve in Q^2 and x at the same time.

1.8 Fragmentation

The fragmentation function formalism is to some extent similar to the treatment of the parton density functions. The fragmentation process describes in general how an energetic parton becomes a hadron (this is in fact the opposite scenario to that of PDFs). There are two independent aspects which have to be described by the process of fragmentation.

The first one involves the probabilities of a quark to hadronize into various types of hadrons. The second one deals with the probability density distribution of the energy fraction z, which is transferred in the fragmentation process from parent parton to the newly created hadron. Usually this fraction is defined as

$$z = \frac{(E + p_{\mathrm{L}})_{\mathrm{hadron}}}{(E + p)_{\mathrm{parton}}}, \tag{1.12}$$

where the energy is combined with the momentum component parallel to the parton direction. This has the advantage of being invariant under a boost along the parton direction.

As already mentioned in the previous section (see eq. 1.11), usually one assumes that the fragmentation function can be factorized. The choice of μ_{f}, which distinguishes between the long distance and short distance part of the process, is theoretically not clearly defined and is in some sense arbitrary. The precise definition of a FF for a parton i fragmenting to hadron h - $D_i^{\mathrm{h}}(z, \mu_{\mathrm{f}})$ depends on how much of the quark evolution, after its production, is absorbed into the perturbative partonic cross section $\hat{\sigma}$, and how much is assigned to the FF itself. Therefore, fragmentation functions can be further decomposed into perturbative and non-perturbative parts D_{p} and D_{np} according to

$$D_i^{\mathrm{h}}(z, \mu_{\mathrm{f}}) = D_{\mathrm{p},i}(\mu_{\mathrm{f}}) \otimes D_{\mathrm{np},i}^{\mathrm{h}} = \int_0^1 \int_0^1 D_{\mathrm{p},i}(x, \mu_{\mathrm{f}}) D_{\mathrm{np},i}^{\mathrm{h}}(y) \delta(xy - z) dx dy \tag{1.13}$$

The perturbative FF $D_{\mathrm{p},i}(z, \mu_{\mathrm{f}})$ takes into account further final state gluon emissions, and the non-perturbative FF $D_{\mathrm{np},i}^{\mathrm{h}}(z)$, which corresponds to formation of bound hadron states is then considered to be universal.[5] The non-perturbative function is usually parametrized by different phenomenological models which will be discussed later.

The perturbative fragmentation function $D_{\mathrm{p},i}(x, \mu_{\mathrm{f}})$ diverges for light quarks, but in the case of heavy quarks (charm and bottom), where the quark mass m_{Q} is significantly larger than the QCD scale Λ, better theoretical control can be achieved. In particular, the mass provides a physical infrared cutoff for collinear radiation making a part of the fragmentation process accessible to perturbative methods. Usually DGLAP inspired evolutions are used in order to resum the perturbative part, defined by the evolution in Q^2 down to the scale of the heavy quark mass m_{Q}.

[5] Universal here means depending neither on the type of hard subprocess, nor on the scale at which parton was produced.

1.8.1 Heavy Quark Fragmentation

Whereas light quark fragmentation is characterized by small energy transfers to the created hadrons, following roughly the behavior $z^{-1}(1-z)^2$, a quite different dependence is expected for the fragmentation of heavy quarks into hadrons. This can be readily seen from simple kinematic considerations as first pointed out by Bjorken and Suzuki [13].

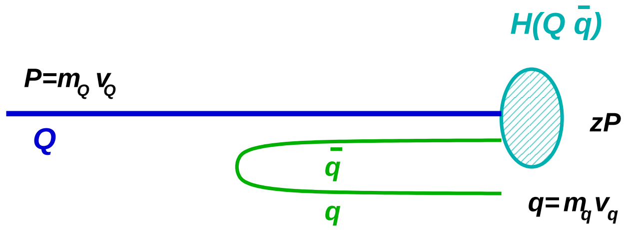

Figure 1.4: Fragmentation of heavy quark Q into heavy meson $H(Q\bar{q})$.

When a light antiquark $\bar{q}$ from the vacuum (or a diquark qq for baryon production) gets attached to a fast moving heavy quark Q, it decelerates the heavy quark in the fragmentation process only slightly, see figure 1.4. This follows from the momentum conservation law and the requirement $v_Q = v_q$ for binding of Qq in a hadron. Thus Q and the heavy hadron $H(Q\bar{q})$ (or $H(Qqq)$ respectively) should carry almost the same energy, resulting in a fragmentation function that peaks near $z = 1$.

1.8.2 Phenomenological Parametrizations

Various attempts have been made to provide firmer predictions for fragmentation functions having the behavior described above. Different functional forms (parametrizations) of $D^{\mathrm{H}}_{\mathrm{np,Q}}(z)$ were proposed.[6] Only the most frequently used ones are discussed here.

1.8.2.1 Peterson Parametrization

The Peterson fragmentation parametrization [53] is based on the assumption that the probability for heavy quark fragmentation is determined by the value

[6] In the following will be the subscript 'np' for the non-perturbative FF dropped.

of the energy transfer between the initial and final states $\Delta E = E_{\mathrm{H}}+E_{\mathrm{q}}-E_{\mathrm{Q}}$. It is given by the following formula

$$D_{\mathrm{Q}}^{\mathrm{H}}(z) = \frac{N}{z[1-1/z-\varepsilon_{\mathrm{Q}}/(1-z)]^2}, \tag{1.14}$$

The factor N normalizes the total probability for hadron formation to one, and the parameter ε_{Q} describes the hardness of the fragmentation process for quark type Q.

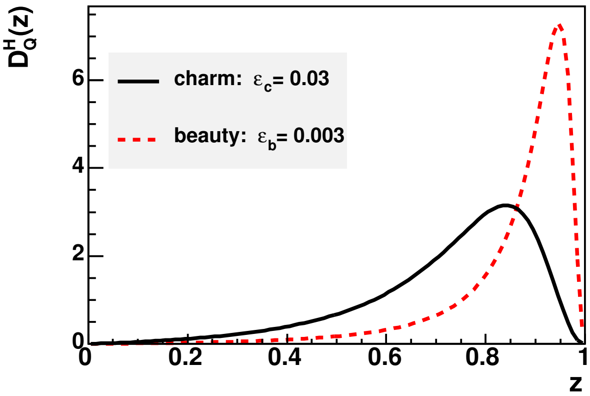

Figure 1.5: Peterson fragmentation functions for the c-quark $D_{\mathrm{c}}(z)$ and the b-quark $D_{\mathrm{b}}(z)$.

Although ε_{Q} is in principle a fixed quantity, since it is proportional to $m_{\mathrm{q}}^2/m_{\mathrm{Q}}^2$ (the ratio of the effective light and heavy quark masses), in practice it is regarded as a free parameter that can be adjusted to fit the data since m_{q} is not well defined. The Peterson model gives a fairly firm prediction for the ratio $\varepsilon_{\mathrm{b}}/\varepsilon_{\mathrm{c}} = m_{\mathrm{c}}^2/m_{\mathrm{b}}^2 \approx 0.1$. This is illustrated in figure 1.5.

1.8.2.2 Kartvelishvili Parametrization

Kartvelishvili *et al.* [40] determined the explicit form of the heavy quark fragmentation function on the basis of the reciprocity between the fragmentation function $D_Q^H(z)$ and the Q-quark density function in the heavy meson H, which can be calculated using the Kuti-Weisskopf model [22]. Assuming the validity of the 'reciprocity relation' $D_{\mathrm{Q}}^{\mathrm{H}}(z) = f_{\mathrm{H}}^{\mathrm{Q}}(z)$ at $z \sim 1$, and the universality of the quark-antiquark sea, one obtains for the FF a simple functional form

$$D_Q^H(z) = N z^{\alpha}(1-z), \tag{1.15}$$

where α is the fragmentation parameter, which should be, according to the theoretical calculations, equal to 3 for the c-quark and 9 for the b-quark.

1.9 Charm Production in *ep* Scattering

In framework of this thesis the fragmentation of charm quarks into $D^{*\pm}$-mesons is investigated. The charm quarks can be produced in ep collisions via several processes (see figure 1.6):

- The virtual photon couples to a sea charm or anticharm quark coming from the proton.

- The virtual photon interacts with a gluon in the proton to form a charm-anticharm pair. This leading-order process in $\mathcal{O}(\alpha_s)$ is known as boson-gluon fusion (BGF).

- The virtual photon fluctuates into partons (quarks and gluons) before entering the hard interaction which then interact with a gluon from the proton side. These processes are referred to as *resolved processes* in leading-order (the previous two mechanisms are referred to as *direct processes*, since the photon acts as a point-like object interacting with a parton from the proton). Here the photon is parametrized by a structure function, similarly as is the case for the proton.

The first H1 measurement [4] of cross sections for $D^{*\pm}$ production indicates that the dominant mechanism for charm production in DIS for $10 < Q^2 < 100$ GeV^2 is the boson gluon fusion process. The contribution to the cross section coming from the charm sea in the proton is expected to be less than 5%. In this analysis this production mechanism is therefore neglected.

A recent ZEUS measurement of dijet distributions [21] in charm events indicates that at low Q^2 a sizable contribution to the cross section originates from resolved processes.

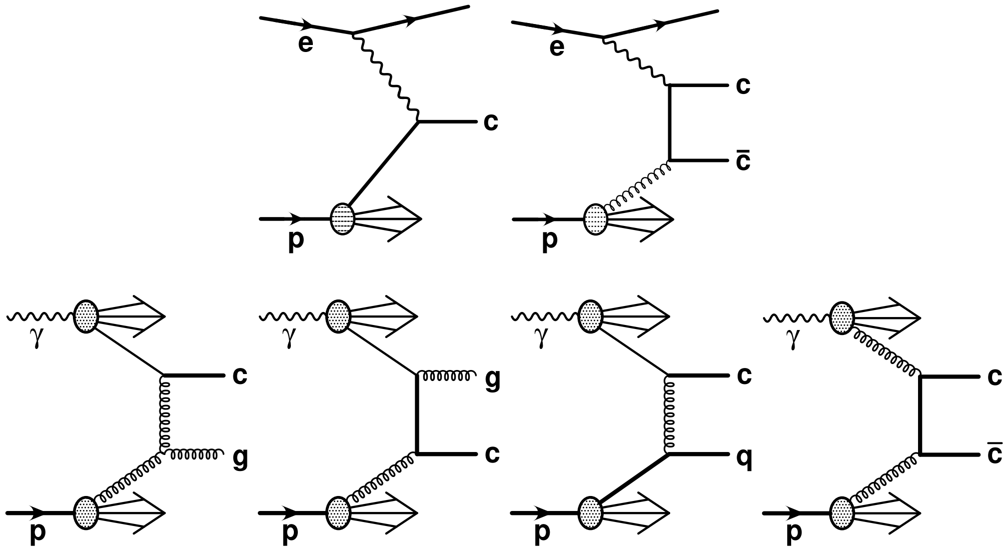

Figure 1.6: Feynman diagrams for the charm production. Direct processes (top): sea charm from proton (left), boson-gluon fusion (right). Resolved processes (bottom).

1.10 Monte Carlo Generators

Monte Carlo (MC) generators are powerful tools used to calculate the detector efficiencies and acceptance after cuts on quantities which have to be introduced in order to identify a given physics process. Only then, direct confrontation of data corrected for these effects with theory is possible.

The simulation of a physics event is done in two steps. In the first step the underlying physics is generated, i.e. all final state particles with definite four-momenta are produced. In the second step the response of the detector to the created particles is simulated. In this chapter we focus only on the first step.

Monte Carlo generators for *ep* reactions are usually based on QCD inspired models. The most wide spread types are LO+PS MC generators. They contain (for illustration see figure 1.7) the leading order (LO) matrix element (e.g. BGF), the initial parton showers (PS) which evolve the parton entering the hard subprocess from proton side, and the final parton showers which simulate QCD radiation down to a cut-off scale of typically around 1

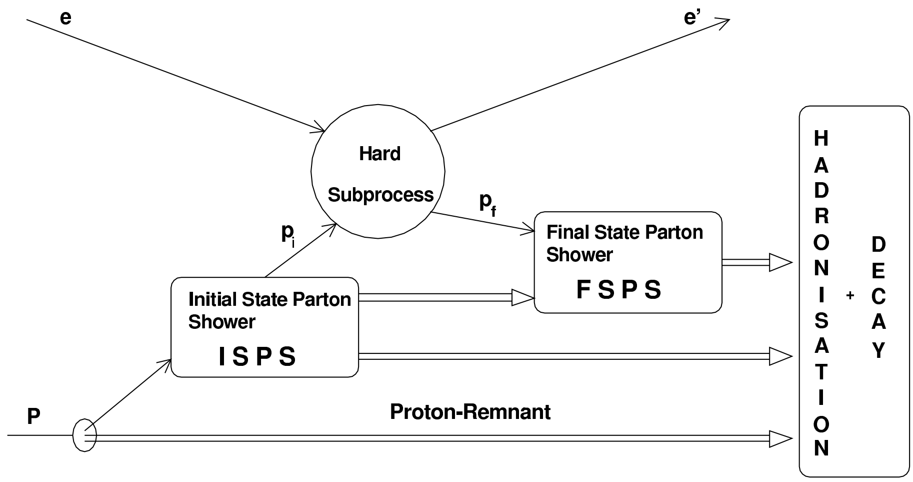

Figure 1.7: Schematic picture of the subsequent steps in the generation of an *ep* event: LO matrix element (hard subprocess), initial and final state parton showers (PS), hadronization of partons and decay of unstable particles.

GeV. Aspects of non-perturbative physics such as hadronization and initial parton density functions are included using phenomenological models and parametrizations.

In this analysis two Monte Carlo generators are used, RAPGAP version 3.1 [38] and CASCADE version 1.2 [37]. Both generators differ significantly in their approaches of modeling initial parton showers.

RAPGAP uses DGLAP evolution together with on-shell matrix element. The direct and resolved processes (figure 1.6) are generated separately and added afterwards. In the latter processes the charm quark is treated as a massive parton with $m_{\rm c} = 1.5$ GeV. The renormalization scale $\mu_{\rm r}^2 = Q^2 + p_{\rm t}^2$ has been chosen. The parton density parametrizations used are CTEQ5L [44] for the proton and SaS-G 2D [59] for the photon.

In the **CASCADE** MC the initial partons evolve according to the CCFM equation. A consequence of the angular ordering of radiated partons in the CCFM approach is that not only the photon but also the gluon entering the boson-gluon-fusion is virtual (the matrix element is off-shell). Only direct BGF processes are taken into account, but due to the CCFM approach some effects of resolved processes are already implicitly included. Thus this approach is complementary to the one used by RAPGAP, and both are used to compare the underlying physics in this thesis. For the calculations the renormalization scale $\mu_{\rm r}^2 = 4m_{\rm c}^2 + p_{\rm t}^2$ is chosen together with the unintegrated gluon density set A0 [39].

After the parton shower, the final state consists of colored partons, which are then non-perturbatively evolved into colorless hadrons. This transition is described by phenomenological models as for example the Lund-string model [8] which is incorporated in PYTHIA [62] . Both Monte Carlo programs, RAPGAP and CASCADE, are interfaced to PYTHIA version 6.2 in order to account for hadronization effects.

1.10.1 String fragmentation model

The string model is based on the idea that the chromodynamic flux between two colored partons is not spread out, like the electric field is, but takes the form of a narrow tube. This narrow tube can be well described as a string which has a constant energy density per unit length (about 1 GeV per Fermi).

As the primary partons move apart, the string stretches and its potential energy increases linearly with the distance at the expense of the parton's kinetic energy. Once the potential energy in the string is sufficient to create a quark-antiquark pair from the vacuum, the string will break into two shorter, less energetic strings. The resulting strings themselves become extended, and will break, until the original string is separated into many short pieces which do not have sufficient energy to break further: these become the conventional hadrons.

In order to generate the quark-antiquark pairs which lead to string breakups, the Lund model invokes the idea of quantum mechanical tunneling. The tunneling probability is given by the Lund symmetric fragmentation function

$$f(z) \propto z^{-1}(1-z)^a e^{-\frac{bm_\perp^2}{z}}, \tag{1.16}$$

where $m_\perp$ is the transverse mass $\sqrt{m^2 + p_\perp^2}$. The quantity z determines the $(E+p_z)$ fraction of the string's energy, which is transferred to the new hadron created by the string's breakup,[7] and a and b are free parameters related to

[7] So in contrast to the independent fragmentation, i.e. Feynman Field [31], where the energy fraction transferred to the hadron is determined from the parent quark alone, in the string fragmentation approach the hadron can gain an additional energy from the other partons bound in the string.

the behavior of $f(z)$ for $z \to 1$ and $z \to 0$. These parameters have been determined from fits to experimental data.

String fragmentation is thus able to reproduce the limited $p_\perp$ distribution of particles in the jet, and also the suppression of strange quarks, as they are heavier than u and d quarks. In case of heavy flavors, where harder fragmentation is needed, PYTHIA offers as an alternative to function 1.16 the Peterson function (equation 1.14).

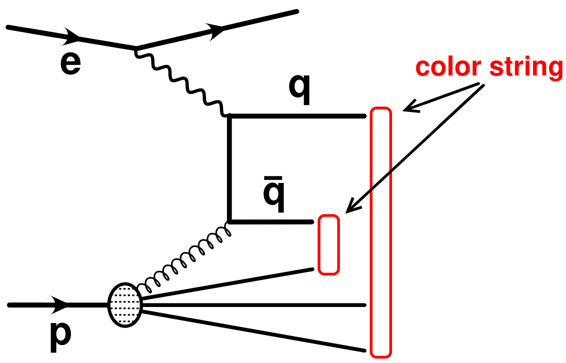

Figure 1.8: Boson-gluon fusion process with indicated color strings.

In the case of BGF, which is supposed to be the dominant process for charm production in *ep* scattering, the color strings are indicated in figure 1.8. A color octet gluon is removed from the proton leaving a quark and a diquark-quark behind (proton remnants in a color octet state). The remnants together with the colored partons from the hard interaction must form color singlet states. Therefore two strings are formed: one connecting the quark from BGF with the diquark, and one connecting the antiquark from BGF with the remnant valence quark.

When additional gluons are emitted, the strings follow the color flow of the parton shower by stretching via one or more gluons, which appear as kinks on the string.

1.11 QCD NLO Calculations

The processes which involve production of heavy quarks offer the possibility to test QCD within the perturbative scope also at small momentum transfers, as the heavy quark mass provides the essential hard scale in addition to the already involved one (e.g. Q^2, W or p_t of the heavy quark). However, the heavy quark production presents a challenge in pQCD, since the conventional pQCD was developed for hard processes depending only on one hard scale. There exist two standard schemes for the calculation of heavy quark processes which try to reduce the two-scale problem into an effective one-scale problem in diametrically opposite ways.

In the **massive** or **fixed flavor scheme** only three active flavors (u, d, s), and the gluon, are assumed to be in the proton. The scheme is called massive because the charm quark is treated as massive, however it is assumed that the charm mass and the second involved hard scale are of the same order. Hence the NLO calculations in this scheme are reliable only near the threshold.

In the **massless** or **zero mass variable-flavor scheme** the charm quark is treated as massless like the light flavors. Hence it appears as the fourth active flavor in the incoming proton parton density functions and it can initiate a hard scattering. This approach is expected to be reliable only, if the second scale involved is much larger than the heavy quark mass.

1.11.1 HVQDIS

In this analysis the HVQDIS program [36] is used for the NLO calculation ($\mathcal{O}(\alpha_s^2)$ order) of charm cross sections. It is based on the massive scheme as described above. The heavy quarks are produced at the perturbative level, where the BGF process is dominant. In addition, a small fraction of light quark induced processes with the emitted gluon splitting into a charm-anticharm pair is present. The renormalization and factorization scales have been chosen to be equal $\mu_\mathrm{r}^2 = \mu_\mathrm{f}^2 = 4m_\mathrm{c}^2 + Q^2$, the charm mass $m_\mathrm{c} = 1.5$ GeV, and the proton PDF CTEQ5F3 [44] is used.

The charm quark is fragmented independently into a $D^{*\pm}$-meson according to the Peterson or the Kartvelishvili fragmentation function (see equations 1.14 and 1.15). The fragmentation is performed in the photon-proton rest frame, which allows to mimic the behavior of the beam drag effect (energy flow in proton direction) as is observed in the laboratory frame. The $D^{*\pm}$-meson is produced on mass shell with three-momentum $\vec{p}_{\mathrm{D}^*} = z\vec{p}_\mathrm{c}$.

In order to improve the description of data the charmed hadron is given transverse momentum with respect to the charm quark direction. The p_t spectrum follows an exponential behavior $\sim \exp(-\alpha p_\mathrm{t}^2)$, with $\alpha = 6$, corresponding to mean value of $< p_\mathrm{t} > \approx 350$ MeV. The value of the α parameter was tuned to describe the average p_{t} value observed in e^+e^- experiments at low energies.

From the QCD point of view at NLO, the introduction of additional p_{t} is not apriori correct for comparison purposes, since it can mimic the effects of higher orders terms in perturbative calculation. Therefore both variants with and without additional transverse momentum are later used.

Chapter 2

Experimental Methods for the Measurement of Fragmentation

Usually the comparison of data on charm production with theory predictions in case of hadron-hadron (resp. hadron-lepton) experiments is done assuming the universality of charm fragmentation. Since fragmentation measurements at hadron colliders are rather difficult, the trend is to use the precisely measured fragmentation parametrization obtained in e^+e^- collisions for the calculation of charm production in other processes. However, this is not sufficient. In particular, one has to take care that these fragmentation functions are used consistently, i.e. they have to be used together with an appropriate QCD calculation (a fragmentation function extracted from e^+e^- data using a MC model as described in chapter 1.10 should not be used with an NLO calculation for a precision comparison). If universality should not hold then there is no guarantee to obtain a good description in any current framework.

To test this assumption of universality with special interest in heavy flavors, charm production has been studied within the framework of this thesis in order to provide exact parametrizations of some theoretically motivated functions.

2.1 Fragmentation in e^+e^- Collisions

In electron-positron collisions charm is produced predominantly via the s-channel process shown in figure 2.1. In this case the laboratory frame is at the same time also the center-of-mass frame of the $c\bar{c}$ system. In leading-order both quarks produced in the annihilation process are back to back, and they carry half of the center-of-mass energy $\sqrt{s}/2 = E_{\mathrm{beam}}$.

There are many observables sensitive to the fragmentation function. The natural choice is to measure the dependence of charm meson ($D^{*\pm}$-meson)[1] production cross section as a function of the observable z, defined as

$$\mathrm{z} = \frac{E_{\mathrm{D}^*}}{\sqrt{s}/2} = \frac{E_{\mathrm{D}^*}}{E_{\mathrm{beam}}}, \qquad \text{respectively} \qquad \mathrm{z} = \frac{|\vec{p}_{\mathrm{D}^*}|}{|\vec{p}_{\max}|}, \tag{2.1}$$

where E_{D^*} and $\vec{p}_{\mathrm{D}^*}$ are the energy and momentum of the D^*-meson, into which the charm quark fragmented, and $|\vec{p}_{\max}| = \sqrt{E^2_{\mathrm{beam}} - m^2_{\mathrm{D}^*}}$ is the maximal accessible momentum to the D^*-meson. These observables have clear advantage that in leading-order $d\sigma/dz$ is directly proportional to the non-perturbative fragmentation function.

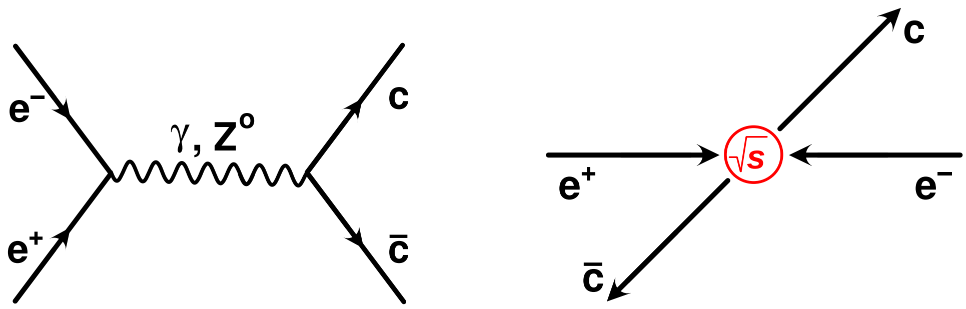

Figure 2.1: Left: Feynman diagram for charm production in e^+e^- collisions. Right: kinematic configuration of $c\bar{c}$-pair in the laboratory rest frame.

2.2 Fragmentation in *ep* Collisions

In electron-proton scattering charm is produced predominantly via boson-gluon fusion. The most appropriate frame for studying fragmentation would be the boson-gluon frame, where the kinematics looks as simple as it was in case of e^+e^- annihilation ($c\bar{c}$ is moving back to back), but unfortunately

[1]Due to the easy tagging, the $D^{*\pm}$-meson is the most frequently used charm meson. Later for simplicity it will be referred only as D^*-meson assuming always both charge conjugated states.

this frame cannot be accessed experimentally. As an alternative remains the boson-proton center-of-mass system (see figure 2.2), which is also called γp-frame, where the proton and photon collide head on along the z-axis, both of them carrying energy $W/2$.

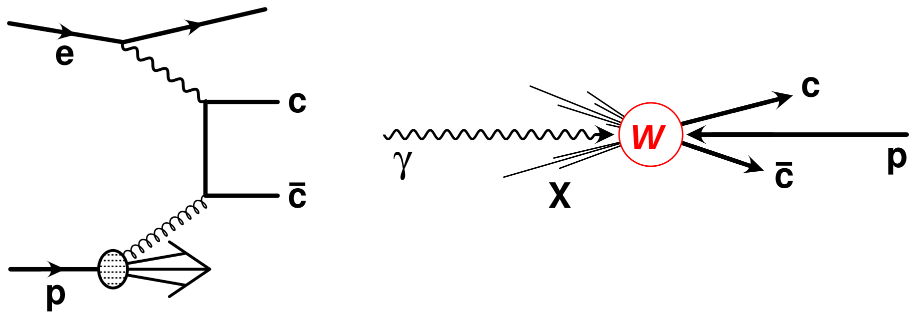

Figure 2.2: Left: Feynman diagram for charm production in ep collisions. Right: kinematic configuration of $c\bar{c}$-pair in the γp-frame.

We see that in case of ep interactions the situation is more complicated than in e^+e^- annihilation, as the energy of charm quark is not known. In this analysis we therefore study the dependence of the D^*-meson cross section on two variables, both of them defined in such a way that under assumption of no gluon radiation and an independent fragmentation of the quarks, $d\sigma/dz$ is directly proportional to the non-perturbative fragmentation function. In the definition of the first variable $\mathrm{z_{hem}}$, the $E+p$ (see equation 1.12) of the charm quark will be approximated by the $E+p$ of a suitably defined hemisphere (referred to as *hemisphere method*). For the second variable, $\mathrm{z_{jet}}$, the $E+p$ charm quark will be approximated by the $E+p$ of a suitably defined jet (referred to as *jet method*).

2.3 Hemisphere method

If one looks at the kinematics of charm production via boson-gluon fusion in the γp-frame, almost all charm and anticharm quarks are found to move in the direction of the incoming photon. This is caused by the fact that while the photon enters with its full energy into the hard subprocess, the proton interacts via its gluons, which typically carry only a very small fraction of the proton's momentum. Assuming no gluon radiation, the charm and anticharm quarks are balanced in transverse momentum.

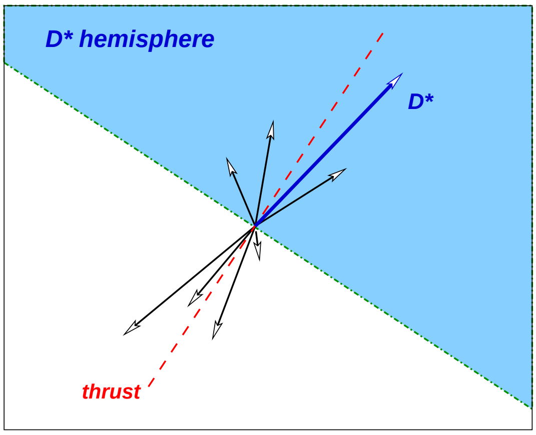

Figure 2.3: Event topology in a plane perpendicular to the photon momentum.

This topology suggests that it is possible to divide the event into two hemispheres, one containing the fragmentation products of the charm quark, the other of the anti-charm quark. To do that, first the projections of the momenta of particles onto a plane perpendicular to the γp-axis are calculated (see figure 2.3). Then the two-dimensional thrust axis [15] is found in this plane via an iterative procedure, where the absolute particle momenta parallel to the thrust axis are maximized. The value of thrust is defined as

$$T = \frac{\max(\sum_i |p_{i\parallel}|)}{\sum_i |p_i|)} \tag{2.2}$$

The value $T = 1$ corresponds to a back to back topology whereas $T = 1/2$ corresponds to an isotropic event.

With a plane perpendicular to the thrust axis the projected event is divided into two hemispheres, one of them containing the D^*-meson and usually other particles. For them we can define:

$$\mathrm{z_{hem}} = \frac{(E + p_{\mathrm{L}})_{D^*}}{\sum_{\mathrm{hem}}(E + p)}, \tag{2.3}$$

where in the denominator the energies and three-momenta of all particles with momentum projections in the D^*-hemisphere are summed, and p_L is the D^*-meson momentum component parallel to the D^*-hemisphere momentum vector. To suppress contributions from the proton remnant and initial state radiation, particles pointing in the proton direction are discarded. This method was initially proposed by [23], and has been modified for the purpose of this analysis slightly.

2.4 Jet method

In the jet method the energy and direction of the charm quark is approximated by a reconstructed jet which contains the D^*-meson. One then arrives at the following definition:

$$\mathrm{z_{jet}} = \frac{(E + p_\mathrm{L})_{D^*}}{(E + p)_\mathrm{jet}}, \tag{2.4}$$

with p_L being the momentum component of the D^*-meson parallel to the jet axis. The jets are found in the γp-frame using the inclusive $k_\perp$-cluster algorithm [28], taking the mass of the D^*-meson into account. The D^*-jet is required to have $E_\mathrm{t}(D^*\mathrm{jet})>3$ GeV. The jet algorithm is described in more detail in the next section.

The two methods differ in case of additional gluon radiation. Whereas the hemisphere method sums up contributions from gluons with low as well as high p_t with respect to the charm quark direction, in analogy with the method used in e^+e^- collisions, the jet method sums up only gluons with small p_t. The comparison of both methods thus provides a test of our understanding of parton radiation.

2.4.1 Jet Algorithm

When energetic partons from the hard subprocess hadronize, one can observe that the created hadrons remain collimated around the original parton directions (the higher the energy the parton has the more collimated the hadrons are). These bunches of hadrons are called jets. They can be interpreted as immediate link to the partons and thus can provide a deeper view of the underlying parton interactions.

In order to find jets which provide a stable basis for theoretical predictions, several jet finding algorithms have been developed [28]. A good jet algorithm should fulfill the following criteria:

- Infrared and collinear safety: the number of found jets and their properties should not change when one of the objects radiates a very soft object, or splits into two collinear objects.

- The output of the jet algorithm should be invariant under a longitudinal Lorentz boost. Since the invariance is not complete (only longitudinal) one should take care that the jet reconstruction is performed in an appropriate frame, like for example the γp or the Breit frame in case of DIS.

One of the frequently used algorithms, fulfilling these criteria, is the inclusive $k_\perp$-algorithm. This jet algorithm is also applied in this analysis using the so-called E-recombination scheme[2].

The jet algorithm starts with a list of objects (partons, hadrons or calorimeter cells, depending on which level the jet algorithm is applied) called protojets, which are characterized by their transverse energy E_t, rapidity $y = \frac{1}{2}\ln(\frac{E+p_z}{E-p_z})$ and azimuthal angle ϕ. The protojets are combined in an iterative procedure according to their distance in the $y-\phi$ plane and their transverse energies into jets in the following steps:

1. For each protojet i and each pair of protojets ij the distances

$$d_i = E_{\mathrm{t},i}^2 \quad , \qquad d_{ij} = \min(E_{\mathrm{t},i}^2, E_{\mathrm{t},j}^2)[(y_i - y_j)^2 + (\phi_i - \phi_j)^2]/R_0^2 \quad (2.5)$$

   are calculated. R_0 is an adjustable parameter related to the opening angle of the jets. In this analysis the theoretically preferred value $R_0 = 1$ has been used.

2. The minimum d_min of all d_i and d_{ij} is found.

3. If d_min is one of the d_{ij}'s, the corresponding protojets i and j are merged into a new protojet k. There exist many possibilities how their kinematic quantities can be recombined. In the E-recombination scheme, which is used in this analysis, the four-vectors of protojets i and j are added

$$(p_x, p_y, p_z, E)_k = (p_x, p_y, p_z, E)_i + (p_x, p_y, p_z, E)_j \qquad (2.6)$$

[2]In contrast to other recombination schemes the particle masses are not neglected here, the so called massive approach.

4. If d_{min} is one of the d_i's, the corresponding protojet i is considered as a final jet and is removed from the list of protojets.

The procedure is repeated until there are no protojets left. The algorithm yields a list containing typically many jets for each event. However, only the jets with large E_{t} values are of much physical interest, therefore a minimal cut on E_{t} of a jet should be demanded. Unfortunately limited statistics in the data does not allow us to cut too hard in E_{t} spectrum since it is exponentially falling. A compromise value of $E_{\mathrm{t,min}} = 3$ GeV has been chosen, which still guarantees a reasonable statistics of jet events in the data sample, as well as reasonable correlation between the quark momentum on parton level and the reconstructed momentum of the jet on hadron level.

In order to improve the correlation between the charm quark and the reconstructed D^*-jet, the D^*-meson is treated as a stable particle (which means that its decay products are removed from the protojet list and replaced by the D^*-meson). In this way it is guaranteed that none of the decay products will end up in one of the 'other' jets and so sharing of energy of the D^*-meson by two or more jets is avoided.

2.5 Fragmentation Observables on MC Generator Level

In this section the properties of both methods, described above, are studied with help of the RAPGAP Monte Carlo simulation. In particular the correlations between the parton level and hadron level are investigated. For charm quark fragmentation in the MC the Peterson parametrization with $\varepsilon = 0.04$ has been chosen, and for the steering of the production of higher excited charmed mesons such as D^1, D^2, etc. a set of parameters [55] tuned by the ALEPH collaboration to their data has been used.

In this study (and also in the whole analysis) only events lying within the kinematic range given by photon virtuality $2 < Q^2 < 100$ GeV^2 and inelasticity $0.05 < y < 0.7$ are considered. Additional cuts on the D^*-meson visibility: transverse momentum $1.5 < p_{\mathrm{t}}(D^*) < 15$ GeV and pseudorapidity[3]

[3]Pseudorapidity is defined as $\eta = -\ln(\tan\frac{\theta}{2})$. This kinematic quantity is often used instead of polar angle θ because of its invariance under the boost along the z(beam)-direction.

$|\eta(D^*)| < 1.5$, which are basically imposed by the acceptance of the central tracking system (section 3.2.1), are demanded as well.

2.5.1 Hemisphere observable

In order to study the sensitivity of the hemisphere observable to the generated fragmentation value, the correlation between the generated fragmentation value in the string breaking process z_{gen}, and the reconstructed value z_{hem} on parton level[4] and hadron level is investigated.

The obtained z_{gen} and z_{hem} distributions on parton and hadron level are shown in figure 2.4 a, b and c. A more complete picture about the situation can provide the correlation plots 2.4 d,e and f.

In spite of the expectation of hemisphere observable on parton level being close to the generated z_{gen} value, as it is not biased by hadronization effects, the z_{hem} spectrum shows strong tendency to reconstruct softer fragmentation function even on parton level. Softening effects occur in events where D^*-meson is produced in cascade decays of excited charm resonances, and thus generated z_{gen} value does not directly refer to the reconstructed D^*-meson, and in events with atypical topology caused by hard gluon radiation.

The hemisphere spectrum on hadron level has similar behavior as on parton level, but is slightly softer. The correlation between z_{hem} on both levels is reasonable.

The reason for entries at $z_{hem} > 1$ in case of the parton level distribution, which seem to be unphysical, is given by the nature of fragmentation process as implemented in the Lund-string model [8]. Since fragmentation is not an independent process, the D^*-meson can gain additional energy from nearby partons in the same string and so end up with energy slightly larger than the energy of the parent charm quark.

When studying charm fragmentation one should obviously take care that the observable to be measured is not too much affected by initial state parton showers and the beam remnants. If parts of the beam remnant and initial parton showers would enter the hemisphere calculation (they enter into the denominator), it would lead to a softer fragmentation spectrum. Therefore, in order to suppress this contribution, a simple but powerful cut $\eta > 0$ of

[4] z_{hem} on parton level means, that hemisphere momentum is calculated from partons, whereas the D^*-meson itself is taken from hadron level, as on parton level no hadrons are available.

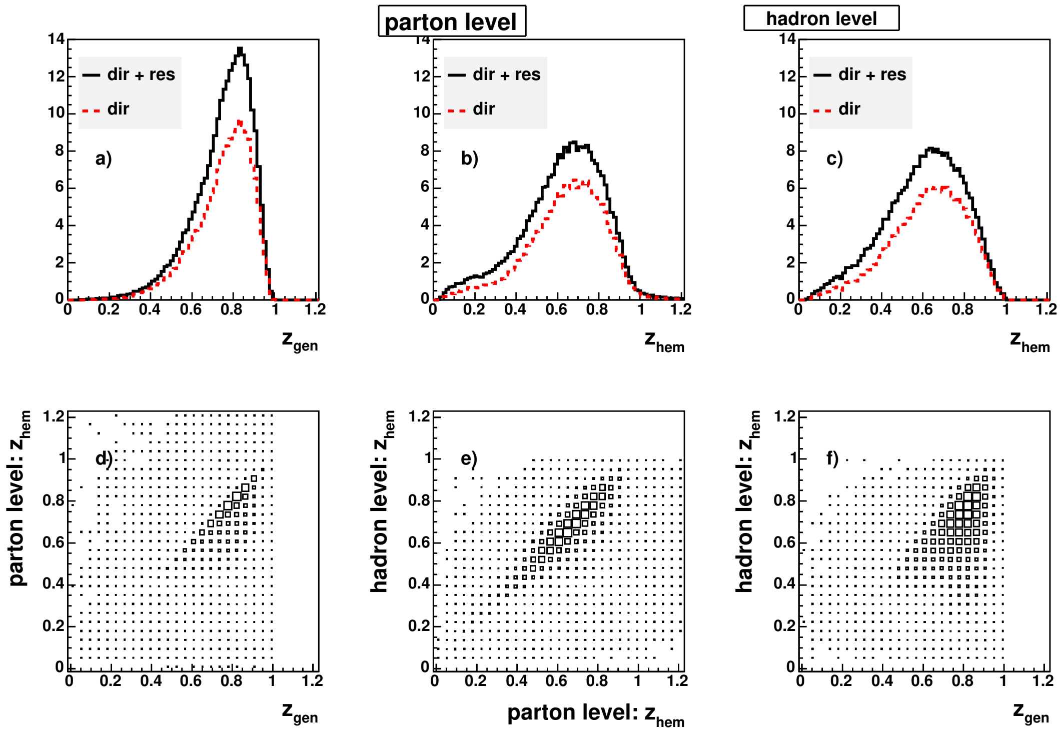

Figure 2.4: Hemisphere observable spectra : a) generated z_{gen} spectrum, b) z_{hem} spectrum on parton level, c) z_{hem} spectrum on hadron level, d) correlations between z_{gen} and z_{hem} on parton level, e) correlations between z_{hem} on parton and hadron level and f) correlations between z_{gen} and z_{hem} on hadron level.

all particles in γp-frame has been introduced (the spectra shown in figure 2.4 already contain this cut). The effect can be seen in figure 2.5, where the energy weighted η distribution of all partons is plotted. The charm quarks emerging from the hard subprocess together with final state parton showers, which are essential for the fragmentation measurement, pass the cut without being affected. On the other hand, most of the proton beam remnant and more than half of the initial state gluons are cut out.

An attempt to increase the η cut did not bring any improvement, since the p_t-balance of remaining particles got worse, which negatively influenced the determination of the charm quark hemisphere.

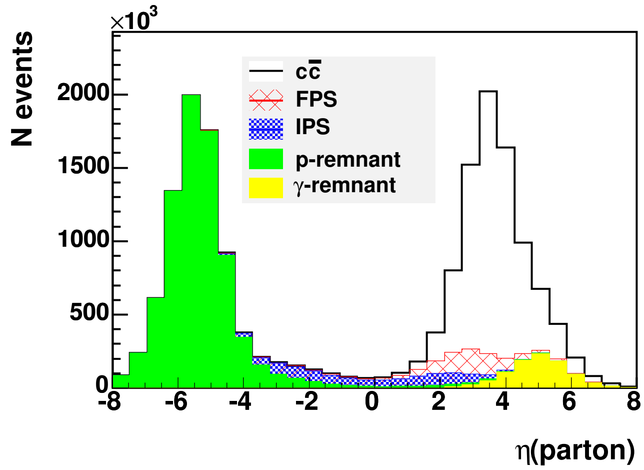

Figure 2.5: Energy weighted η distribution of all generated partons in the γp-frame.

In case of resolved processes the situation is more complicated due to the presence of the photon remnant. In contrast to the proton beam remnant, which can be completely excluded, the photon remnant populates the same kinematic phase space as partons coming from the hard subprocess, and it often contains a charm quark. According to the MC simulation in roughly 7% of all cases the tagged D^*-meson originates from the charm quark of the photon remnant. Especially for these events the hemisphere method fails to reconstruct the generated z value (in most cases yielding values for $z_{\mathrm{hem}} \simeq 0$).

In spite of the subtleties and deficiencies mentioned, the correlation of z_{hem} to the generated fragmentation value is still reasonable. Thus the hemisphere method can be used for the extraction of the fragmentation parameter.

2.5.2 Jet Observable

Also for the jet method the correlation between the generated and reconstructed fragmentation value on parton level and hadron level has been studied. The z_{jet} distributions obtained on parton and hadron level, as well as their correlation matrix are shown in figure 2.6.

The spectra show several differences with respect to the distributions obtained previously by the hemisphere method (figure 2.4). In case of the jet

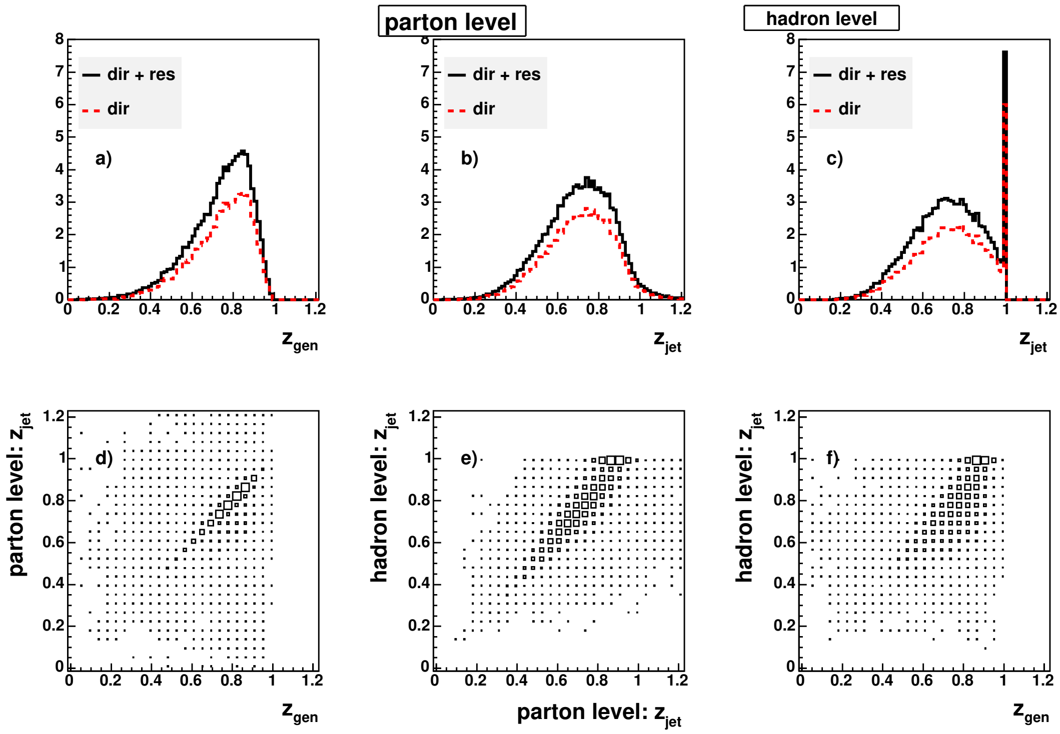

Figure 2.6: Jet observable spectra: a) z_{gen} spectrum, b) z_{jet} spectrum on parton level, c) z_{jet} spectrum on hadron level, d) correlation between z_{gen} and z_{jet} on parton level, e) correlation z_{jet} on parton and hadron level and f) correlation between z_{gen} and z_{jet} on hadron level.

method the fragmentation spectra are harder than those from the hemisphere method, which is caused by the fact that in the jet less radiated gluons are summed than in the D^*-hemisphere.

The correlation between z_{gen} and z_{jet} observable is very good. However, on the hadron level a new characteristic feature appears - a large peak exactly at $z = 1$. These entries correspond to jets, which consist only of the D^*-meson itself. In general, since charm is at HERA mainly produced close to the kinematic threshold, the charm jets are rather broad and not well collimated. Therefore, the jet algorithm sometimes fails to find soft hadrons belonging to the charm quark fragments. By increasing the cut on the jet transverse momentum the fraction of these entries drops, but unfortunately the limited statistics in data does not allow to increase the cut on $E_{\mathrm{t}}(D^{*}\mathrm{jet})$.

This peak is just an artefact and must not be taken as evidence for an extremely hard fragmentation as can be seen from the correlation plot 2.6 e (the horizontal line at $\mathrm{z}_{\mathrm{jet}}^{\mathrm{hadron}} = 1$ covers the interval $0.7 < \mathrm{z}_{\mathrm{jet}}^{\mathrm{parton}} < 1$). Nevertheless, since the $\mathrm{z}_{\mathrm{jet}}^{\mathrm{parton}}$ for this subset of events is significantly harder than the rest of the sample, these events should not be excluded, as they may contain interesting information about the fragmentation process.

Chapter 3

H1 Experiment at HERA

3.1 HERA Accelerator

The HERA machine was build as the first lepton-proton accelerator with two independent storage rings placed close to each other; one for protons and a second for leptons. It is located at the DESY research center in Hamburg, Germany. Since 1999 the HERA accelerator provides proton and lepton beams with energies of 920 GeV and 27.6 GeV respectively, yielding in center-of-mass energy $\sqrt{s} \approx 320$ GeV. The leptons and protons are stored in bunches which collide in 96 ns intervals in two interaction regions where the H1 and ZEUS experiments are located.

3.2 H1 Detector

The H1 detector is one of two multi purpose detectors designed to study *ep* collisions at HERA collider. It consists of several layers of detectors symmetrically placed around the beam pipe, allowing to measure energy and momentum of charged as well as neutral particles produced in collisions. Since the beam energies are highly asymmetric, most of the particles are boosted along the proton direction. Due to this feature the detector instrumentation is enhanced in the proton (forward) direction.

At H1 a right handed coordinate system (figure 3.2) is used. The positive z-axis is defined parallel to the proton beam direction (often referred as forward direction) and the x-axis points towards the center of HERA ring.

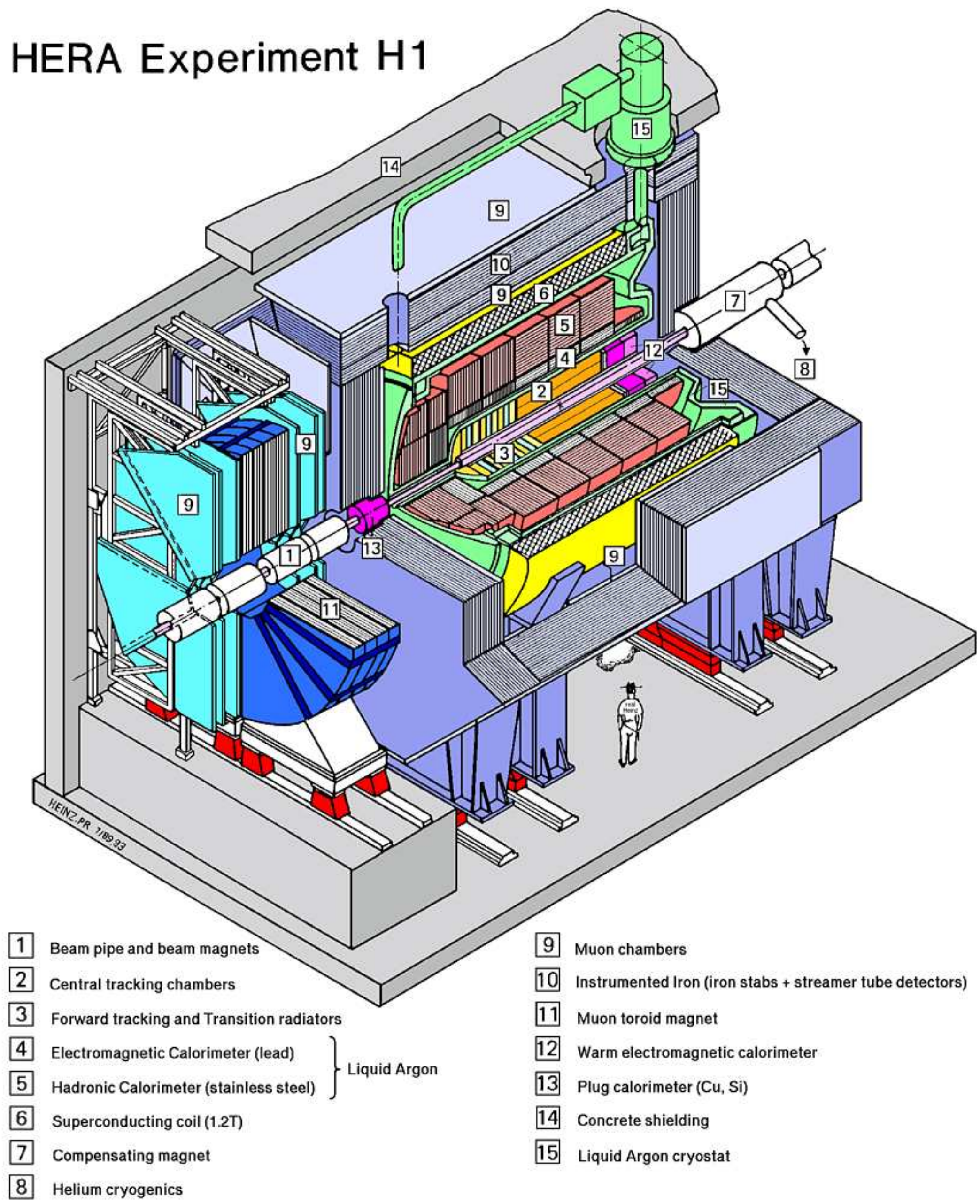

Figure 3.1: A schematic view of the H1 detector.

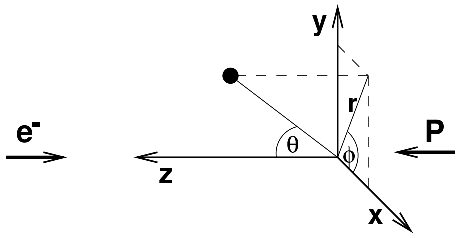

Figure 3.2: The H1 coordinate system.

The main components of the H1 detector are shown on figure 3.1. The beam pipe around the interaction point is first surrounded by silicon trackers (located in central and backward direction), then by forward and central tracking chambers followed by a liquid argon calorimeter and muon chambers. Between the muon chambers and the calorimeter is positioned the superconducting solenoid which provides a magnetic field of 1.15 T. The presence of magnetic field is essential for transverse momenta measurement of charged particles in the tracking chambers.

In addition several small detectors are located along the beam pipe to provide a luminosity measurement, the tagging of photoproduction via detecting the electron scattered under very small angles in the backward direction, and a measurement of the scattered proton or proton remnant in the forward direction. More detailed descriptions can be found in references [1] [2].

In the following only detector components related to this analysis, including the trigger system, will be described.

3.2.1 Tracking System

The H1 tracking system (see figure 3.3) is designed to measure three-dimensional particle trajectories within a large acceptance range. A part of it is used also for triggering purposes. From a geometric point view it can be divided into three parts: the forward tracking detector (FTD) covering angular range $7^o < \theta < 25^o$, central tracking detector (CTD) covering range $15^o < \theta < 165^o$ and the backward drift chamber (BDC) covering range $153^o < \theta < 177^o$.

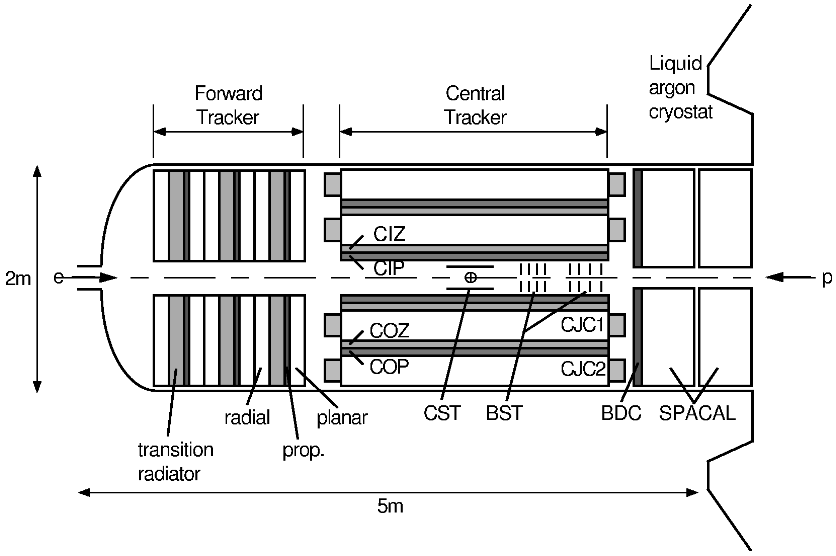

Figure 3.3: The H1 tracking system ($r - z$ view): forward, central and backward trackers.

3.2.1.1 Central Tracking Detectors

The track reconstruction in the central region is performed with help of four concentric cylindric drift chambers (the central jet chambers CJC1,CJC2 [16] and the central z-chambers CIZ,COZ [27]) and a central silicon tracker (CST) [54] which serves mainly for improving the event vertex reconstruction. Both central jet chambers provide besides the trajectory measurement also the measurement of specific energy losses dE/dx, which are in this analysis used for particle identification.

The CJC1 consists of 30 cells each containing 24 sense wires, whereas the outer CJC2 chamber consists of 60 cells with 32 sense wires running parallel to the beam axis providing a good resolution in $r - \phi$ plane. The cells are separated with planes of field shaping wires, which provide a homogeneous electric drift field. The charged particles moving through the jet chambers in the homogeneous magnetic field follow a helical trajectory, with radius proportional to the transverse momenta p_t with respect to the magnetic field lines. The particles ionize the gas in the chamber and the free electrons produced in ionization process drift towards to sense wires, driven by the electric field formed with field-shaping wires. Due to the gas amplification the drifting electrons close to the sense wire initiate an ionization avalanche generating a signal. The drift velocity and the time at which the signal is registered at the sense wire determine the hit position in $r-\phi$ plane. In this way a spacial resolution of about 170 μm can be achieved.

The z trajectory component can be obtained on the basis of charge division between the ends of active sense wires. This method leads to a resolution of 2.2 cm, which is much worse than the resolution in $r-\phi$ plane. Nevertheless it can be improved by including the hit information from inner and outer z-chambers (CIZ and COZ), as the sense wires in both z-drift chambers are stretched perpendicular to $r-\phi$ plane. This topology provides excellent resolution of $\sigma_z = 350$ μm.

Close to the z-chambers are located multiwire proportional chambers CIP and COP [45] which are used for triggering on events with significant number of tracks pointing to the nominal vertex region in z plane.

3.2.1.2 Forward Tracking Detector

The FTD [17] is build up from three identical supermodules arranged along the z-axis. Each of them consists of radial and planar drift chambers, proportional chamber and a transition radiator. Till now the forward trackers are not satisfactory understood within the Monte Carlo simulation. Therefore, pure forward tracks which are not linked to any central tracks are not used in this analysis.

3.2.1.3 Backward Drift Chamber

The backward drift chamber (BDC) [18] is mounted in front of the SPACAL calorimeter. It is mainly used to improve the precision of the polar angle measurement of the scattered electron in the SPACAL, and to distinguish electrons from photons. The BDC consists of four double-layers drift chambers with wires arranged in octagonal form. These layers are rotated by 11.5^{o} to each other, which enables to measure the ϕ coordinate. The resolution of polar angle θ is better than 1 mrad with a systematic uncertainty of 0.5 mrad.

3.2.2 Calorimeters

Calorimeters offer a complementary measurement to the tracker measurement, allowing to detect the neutral particles which are for trackers 'invisible' and to discriminate between leptons and hadrons. In sampling calorimeters the energy measurement is achieved by putting layers of passive dense material (called absorber), where the incident particles loose energy via production of secondary particle showers, and layers of sampling material where the deposited energy can be measured. The two most important calorimeters in the H1 experiment are the liquid argon calorimeter (LAr), covering the central and forward region, and the backward scintillating fiber calorimeter SPACAL.

3.2.2.1 Liquid Argon Calorimeter

The LAr calorimeter [9] (see figure 3.4) is divided into inner electromagnetic section with lead absorber plates and outer hadronic section with stainless steel absorber plates in order to optimize the electron (photon) and hadron energy measurement. In both sections liquid argon is used as an active medium. The choice of argon brings several benefits: it is a noble gas with high atomic density resulting in large ionization and guaranteeing a stable performance of the calorimeter over years.

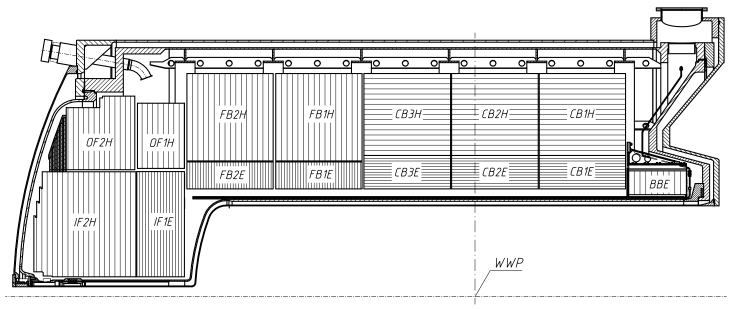

Figure 3.4: LAr calorimeter side view.

The charge produced during the ionization of argon atoms by secondary particle showers is collected on cathode pads and the induced signal is readout. In total there are 44352 cells with 63488 electronic read-out channels.

Granularity of the cells increases towards the forward direction, in order to provide homogeneous coverage in pseudorapidity.

The LAr is a non-compensating type of calorimeter which means that hadronic particles cause in average around 1/3-rd smaller energy deposit than electromagnetic particles of the same energy. In order to account for this behavior a weighting technique is applied during the reconstruction of hadronic energy measurement.

The energy resolution of LAr calorimeter for electromagnetic showers is $\sigma_{\mathrm{em}}(E)/E = 11.5\%/\sqrt{\mathrm{E[GeV]}} \oplus 1\%$ and for hadronic shower of $\sigma_{\mathrm{had}}(E)/E = 50\%/\sqrt{E[\mathrm{GeV}]} \oplus 2\%$.

3.2.2.2 SPACAL

The primary function of the SPACAL calorimeter [47] is to trigger and precisely measure the electrons scattered under small angles, and thus allow to study the low Q^2 DIS events. It is a non-compensating scintillation fiber calorimeter with an electromagnetic and hadronic part covering angular range from 153^o to 175.5^o. Both parts are divided into cells, the electromagnetic cells being smaller than the hadronic cells.

Scintillating fibers, which are aligned along the beam axis, are embedded in a lead matrix. When shower particles pass the scintillation material, scintillation light is emitted which is then collected at the end of the fibers and guided to photomultipliers. The light is in photomultipliers converted to electronic signals which are then amplified and read-out.

The achieved resolution for electrons is $\sigma_{\mathrm{em}}(E)/E = 7\%/\sqrt{\mathrm{E[GeV]}}$ and for hadrons $\sigma_{\mathrm{had}}(E)/E = 56\%/\sqrt{E[\mathrm{GeV}]}$.

3.2.3 Trigger System

At designed HERA luminosity around 1000 ep interactions are expected per second, whereas DIS events occur only with couple of Hz frequency. However the total interaction rate is completely dominated by non ep background events ($\sim$ 100 kHz) coming from beam-gas, beam-wall collisions, etc. ...

This constant data stream cannot be handled since the read-out rate as well as the data transfer and storage capability is limited. Therefore a sophisticated trigger system which discriminates events of physics interest from background events is needed. The H1 trigger system consists of 4 levels [1] which are designed to successively reduce the rate from 10 MHz down to hundreds of Hz (see figure 3.5). The reduction rate of the levels is inversely proportional to the processing time needed for more and more complex calculations.

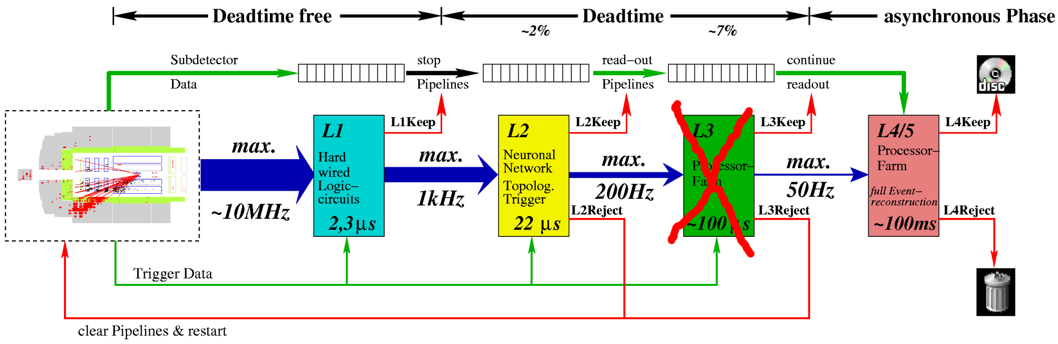

Figure 3.5: The H1 multilevel trigger system.

3.2.3.1 Trigger Level 1

Due to the high bunch crossing frequency at HERA each 96 ns an interesting interaction can possibly occur. Therefore it is essential that the level 1 trigger (L1) provides a decision for each bunch crossing (BC) without causing a dead time. This is achieved by feeding detector signals into pipelines whose length varies form 27 to 35 bunch crossing intervals depending on the subdetector system. In this way time is gained to obtain general event characteristics on which hand a decision can be made. The information from different subdetectors are in form of trigger elements delivered into central trigger logic (CTL) where they are logically combined to various so-called subtriggers. The subtriggers are available 24 BC after the real interaction time and the L1 decision can be made. If any of the subtrigger conditions are fulfilled an L1KEEP signal is sent to stop the pipelines, to protect them from being overwritten, and the L2 trigger is initiated. From this moment the detector is insensitive to further interactions and the dead-time starts.

[1] During the 2000 running period, which is analyzed here, the level 3 trigger system was not yet implemented.

3.2.3.2 Trigger Level 2

On the second level two independent parallel running systems have nearly ten times more time (20 μs) to verify or falsify the L1 decision. The first system is based on a neural network architecture L2NN and the second on event topologies L2TT. They work with more detailed information delivered by most of the trigger sub-systems. If the event is not rejected the L2KEEP signal initiates the read-out of the whole detector (complete read-out takes between 1-2 ms).

3.2.3.3 Trigger Level 4

On the fourth level full detector information is available, hence the online event reconstruction can be performed. The detector information is filled into a buffer and then is asynchronously processed on a PC-farm, so it does not further contribute to the detector dead time.

If the decision of L1 and L2 triggers is confirmed, a L4 classification of the event is done. If the event is classified as a interesting physics channel event ('high Q^2', 'open charm', 'diffraction', ...) then it is being kept, otherwise it is downscaled according to its Q^2. The smaller the Q^2 value is the larger the downscaling factor is attributed. In this way the L4 output rate is reduced to 10 Hz and the selected events are written on tapes.

Chapter 4

Selection of D^*-Mesons

This chapter describes the selection of the analyzed data sets. First an overview of the online selection, which is performed using a dedicated subtrigger, is given. Then the reconstruction of the event kinematics used for the definition of the DIS phase space of this measurement and the reconstruction of D^*-mesons is discussed.

4.1 Online Selection

The starting point of the analysis is the selection of the events of interest from a huge amount of background events. This is done in two steps: First a preselection is done online during the data taking. It is based on a trigger decision which selects a certain class of physics events. Secondly, the final selection is then done offline, where analysis specific cuts are applied, and fine tuning is possible.

As has already been mentioned in chapter 3, the data are selected using the 4-level trigger system. This D^* analysis makes use of one single dedicated L1 subtrigger ST61. This subtrigger combines information from several detector components, demanding a signal from the SPACAL trigger, in coincidence with a charged track signal from the z-vertex and the DCRPhi triggers. The basic definition of subtrigger ST61 is as follows:

$$\mathrm{ST61} = (\mathrm{SPCLe\_IET} > 2 || \mathrm{SPCLe\_Cen3}) \& \mathrm{DCRPhi\_THigh} \& \mathrm{zVtx\_sig} \quad (4.1)$$

- The **SPACAL IET**(inclusive electron trigger) subsystem is used to detect events with a scattered electron in the backward region. In order to achieve high efficiency a so-called sliding window technique is used: from 4x4 neighboring electromagnetic cells trigger towers are built which overlap with each other in both x and y directions. If the energy deposit in one of the trigger towers exceeds a threshold of 6 GeV, the trigger element is set to on. (More precisely, the condition that is demanded consists of a logical or between the energy deposit in the non-central and central region of SPACAL: SCPLe_IET>2 ||SPCLe_IET_Cent_3.)

- The **DCRPhi Trigger** uses as input digitized signals from 10 wire layers of the CJC. They are compared with predefined track masks corresponding to charged particle trajectories. The subtrigger ST61 requires the trigger element DCRPh_THigh to be on, which means that at least one track mask above a threshold of $p_{\mathrm{t}} > 800$ MeV is found.

- The **z-Vertex Trigger** decision is based on the CIP, COP and FPC. The signals from the proportional chambers are divided into z and ϕ sectors which are then used to obtain rough information about the event z-vertex position. All possible combinations of particle trajectories are constructed, and their intersections with the z-axis are filled into a histogram (see figure 4.1). Only particles originating from an interaction point cause an excess in the corresponding histogram bin, whereas the combinatorial background is randomly distributed. In this analysis the zVtx_sig trigger element is required, which implies a significant peak in the z-vertex histogram.

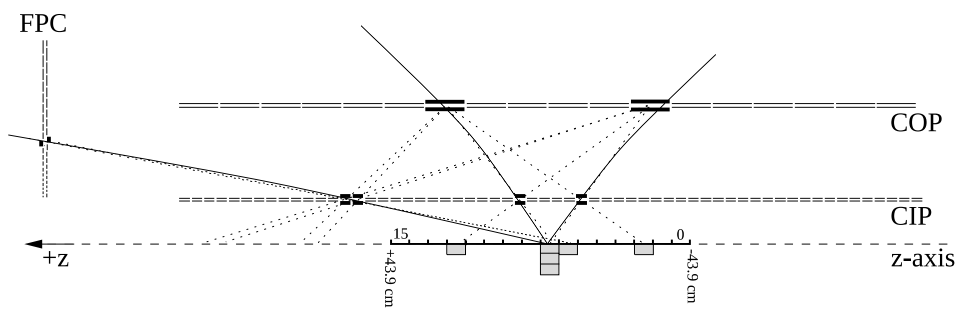

Figure 4.1: The principle of the z-vertex trigger histogram procedure.

In addition, veto conditions from time-of-flight systems are applied to suppress out of time background coming from beam-gas and beam-wall interactions.

In order to keep L1 rates at a tolerable level prescale factors may be applied on subtriggers. Prescale factor n means that only each n-th time, when the subtrigger fires, is an event accepted and processed further. In the analyzed 2000 data taking period the average prescale factor of ST61 is 1.2.

Since there is no additional L2 requirement for the ST61 subtrigger, the events which passed L1 are then directed to level 4, where they can be downscaled once again according to their Q^2 (an L4-weight is attributed to each event, which is then later used in the off-line analysis). A photoproduction or low Q^2 event can still be saved, if one of the hadronic final state finders, such as HQSEL [32] for open charm production, considers the event to be interesting. HQSEL searches for various charmed hadrons in different decay channels, for example the $D^* \to K\pi\pi_s$ channel which is used in this analysis. The D^*-candidates are identified by using the tracks reconstructed in the central tracking detector with a rather loose set of cuts, which will be discussed later.

4.2 Data Quality Cuts

4.2.1 Run Selection

The data used in this analysis were collected during the 2000 running period when HERA collided 27.6 GeV positrons with 920 GeV protons (the run range with shifted vertex position is not included). From the overall data sample, only runs which fulfill certain quality criteria are selected.

- The runs flagged as 'poor' are excluded. This flag is set, if either calorimeters or trackers are not operational, or other serious problems have been found such as readout inconsistencies of some of the detector components, not properly loaded system, etc. ...

- Only data, collected during the time when all detector systems essential for this analysis had fully operational read-out and high voltage on, are taken. The relevant detector components are the CJC1 and CJC2, the central and forward proportional chambers (CIP, COP and FPC), the

LAr and SPACAL calorimeters, the BDC, the time-of-flight scintillators and luminosity system.

- The integrated luminosity $\mathcal{L}^{\text{int}}$ per run, after high voltage and satellite bunch corrections has to be larger than 0.1 nb^{-1}.

With these requirements the total integrated luminosity of the analyzed data sample is 43.35 pb^{-1}, and after correction for the prescale factor of subtrigger ST61 $\mathcal{L} = 36.23\ \text{pb}^{-1}$.

4.2.2 Event Vertex

To reduce the number of background events coming from interactions of the proton beam with the residual gas or with the beam pipe walls, which can happen anywhere along the whole beam pipe, only events with a reconstructed z-vertex in region close to the nominal z-vertex position are taken into account ($|z_{\text{vtx}}| < 35$ cm). The resulting z_{vtx} spectrum is shown in figure 4.2.

Unfortunately the z-vertex position is not easy to simulate within the Monte Carlo framework, because it depends a lot from the properties of injected bunches during the HERA run operation. Hence good agreement between the data and MC simulation should not be expected. However, it is important to have a good description of the z-vertex position of the data by Monte Carlo, since it can influence the reconstruction of the scattered electron.Therefore the z-vertex position of Monte Carlo simulation have been reweighted to describe the data.

4.3 Selection of DIS Events

This section describes the criteria applied to identify the scattered electron, the reconstruction of kinematic variables as well as the final cuts for the DIS event selection.

4.3.1 Scattered Electron

In low Q^2 DIS events the incident electron is scattered at small angles into the backward SPACAL calorimeter. Its energy is reconstructed from the

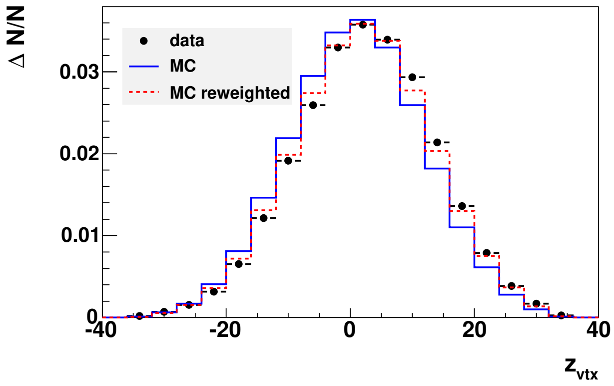

Figure 4.2: Description of the z-vertex distribution by the Monte Carlo simulation before and after a reweighting procedure.

deposited energy in the form of a cluster in the electromagnetic section of the SPACAL.

The largest source of background clusters, which fakes the electron signature, comes from hadronic final state particles moving in the backward direction from the interaction point into the SPACAL. One source of misidentified electrons is due to photons from π^0 decays, $\pi^0 \rightarrow \gamma\gamma$. Also hadrons, due to fluctuations in the shower development, are sometimes misidentified as electrons. Therefore, a small fraction of photoproduction events, where the scattered electron goes unseen down the beam pipe, can be misidentified as a DIS event (after applying all analysis cuts the fraction is below 1% [60]).

In order to reject background, the cluster with highest p_{t} which passes the following criteria, is selected as the scattered electron:

- **Cluster energy**
  To suppress the contribution from backward moving hadrons, the center of gravity of a reconstructed cluster and 90% of its energy has to lie in the electromagnetic part of the SPACAL calorimeter. The cluster energy must be larger than 8 GeV.

- **Cluster radius**
  Since hadronic showers tend to be broader than electromagnetic ones, laterally as well as longitudinally, a small cluster radius of 3.5 cm is

required.

- **BDC hit association**
  The BDC is used to reject background from high energetic photons from π^0 decays. Photons, unless converted in material, do not produce any signal in the drift chamber in contrast to electrons, hence a track segment in the BDC with a small projected distance from the center of the SPACAL cluster is required ($\Delta R_{\mathrm{BDC}} < 1.5$ cm).

- **Radial cluster position**
  To guarantee a fully contained shower in the calorimeter, electron candidates in the innermost SPACAL region $r_\theta = (z_{\mathrm{clus}} - z_{\mathrm{vtx}}) \tan\theta < 9.1$ cm are rejected. In contrast to a simple geometrical cut on $\rho_{\mathrm{clus}} = \sqrt{x^2_{\mathrm{clus}} + y^2_{\mathrm{clus}}}$ this cut allows a beam displacement between data and MC simulation. In addition, events with energy deposits in Veto layers and $\rho_{\mathrm{clus}} > 74$ cm are rejected.

Because of better resolution the polar angle θ'_{e} as well as the azimuthal angle ϕ'_{e} are calculated from the track segment in the BDC instead of from the SPACAL cluster[1].

To account for trigger cells which did not work efficiently for long run-periods, fiducial regions for the electron in SPACAL have been cut out [60]. The distributions characterizing the scattered electron, in events containing a D^*-meson, are shown in figure 4.3 a-c). The MC simulation has been corrected for the BDC inefficiency of 10% in the transition region between the small and large BDC cells at cluster radius $\rho_{\mathrm{clus}} \approx 27$ cm, which is not implemented in the standard detector simulation.

4.3.2 The Energy Balance

$E - p_z$ is a quantity characterizing the energy and longitudinal momentum balance of an event. From energy momentum conservation before and after interaction one obtains following relation

$$E - p_z = E_{\mathrm{p}} - p_{\mathrm{p},z} + E_{\mathrm{e}} - p_{\mathrm{e},z} = \sum_{i=\mathrm{hadrons}} (E_i - p_{i,z}) + (E'_{\mathrm{e}} - p'_{\mathrm{e},z}), \tag{4.2}$$

[1] Only the polar angle resolution of the BDC is better, but since the azimuthal resolution is comparable with the one from SPACAL, for simplicity both quantities are taken from the BDC.

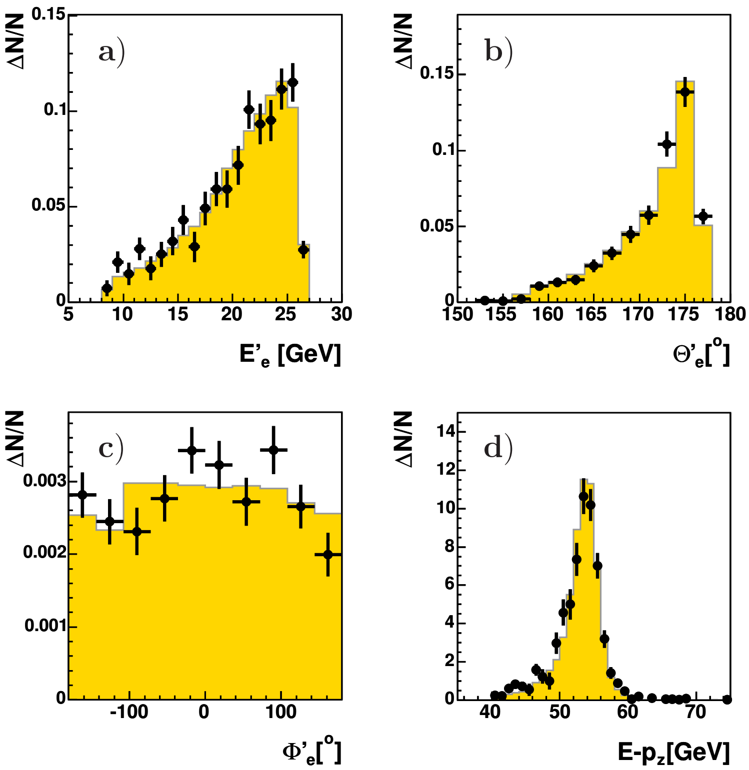

Figure 4.3: Scattered electron quantities for events containing a D^*-meson: a) energy E'_e, b) polar angle θ'_e, c) azimuthal angle ϕ'_e and e) $E - p_z$ control plot. Shown are data (full points) together with RAPGAP MC predictions after $z_{\rm vtx}$ reweighting (yellow filled area).

where the sum runs over all measured hadronic objects. The observed distribution is shown in figure 4.3 d. In the ideal case $E - p_z$ should be equal to $2E_e \approx 55$ GeV.

This quantity is very sensitive to particle losses in the backward direction and therefore can be used to reduce the contamination from photoproduction events, where the electron escapes down the beam pipe, by imposing a lower limit. By applying an upper limit poorly reconstructed events are rejected. In this analysis the range 40 GeV $< E - p_z <$ 75 GeV is required.

4.3.3 Reconstruction of the Kinematics

Various methods are used for the reconstruction of the event kinematics [10], which differ in their sensitivity to QED radiation, to the accuracy of the energy scales and polar angle measurements. For this analysis the so-called electron method has been chosen.

The **electron method** is based only on measurements of the energy E'_{e} and the polar angle θ'_{e} of the scattered electron. The kinematic variables defined in section 1.3 can be calculated using the following formulae:

$$Q^2_{\mathrm{e}} = 4E'_{\mathrm{e}}E_{\mathrm{e}}\cos^2\frac{\theta'_{\mathrm{e}}}{2} \qquad\qquad y_{\mathrm{e}} = 1 - \frac{E'_{\mathrm{e}}}{E_{\mathrm{e}}}\sin^2\frac{\theta'_{\mathrm{e}}}{2} \tag{4.3}$$

This method provides very good Q^2 and y resolution, which then degrades towards lower y values, where accurate energy calibration becomes essential. The measured inelasticity y and photon virtuality Q^2 are shown in figure 4.4.

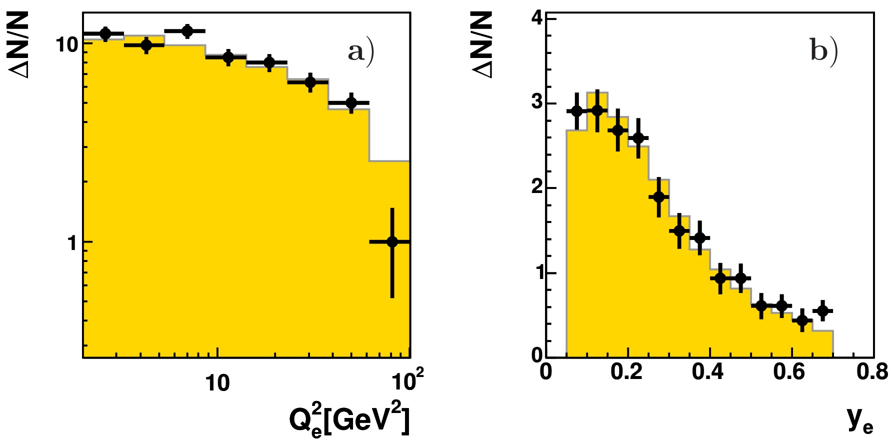

Figure 4.4: Reconstructed kinematic variables for events containing a D^*-meson: a) Q^2, b) y. Shown are data (full points) together with RAPGAP MC predictions after z_{vtx} reweighting (yellow filled area).

4.3.4 DIS Kinematic Range

The DIS kinematic range covered by this analysis is defined in terms of cuts on Q^2 and y:

$$2 < Q^2 < 100 \text{ GeV}^2 \qquad\qquad 0.05 < y < 0.7 \tag{4.4}$$

The cuts on Q^2 are determined by the geometrical acceptance of the SPACAL calorimeter. The lower cut on inelasticity is applied to avoid the region of poor resolution in the event kinematics by the electron method, and the upper cut suppresses events with large QED radiative effects.

4.4 Reconstruction of D^*-mesons

The commonly used method for reconstruction of D^*-mesons (or any kind of unstable hadrons) is to plot the invariant mass of the tracks identified with a given decay channel. The tracks originating from the meson decay have to obey energy conservation laws, and hence their invariant mass is equal to the meson mass (due to detector resolution the reconstructed mass is smeared around the central meson mass value according to a more or less Gaussian distribution). On the other hand randomly combined tracks, not related with a D^*-meson, contribute to the background which can be approximated by a polynomial shape. Assuming a good enough invariant mass resolution leading to a reasonable ratio of signal to background events, this simple method allows to extract events with particular charm decays.

4.4.1 The D^* Decay Mode

The D^{*+} (D^{*-}) meson is an excited $c\bar{d}$ ($\bar{c}d$) state with mass $m(D^*) = 2.010$ GeV [24]. In the framework of this analysis it is reconstructed in the exclusive decay channel

$$D^{*\pm} \to D^0\pi_s^\pm \to K^\mp\pi^\pm\pi_s^\pm \tag{4.5}$$

The mass difference between the D^* and D^0 of 145.4 MeV is only slightly above the pion mass $m_\pi = 139.6$ MeV. Therefore, the available kinetic energy of the decay products is very small, and the D^0 and $\pi_s^\pm$ are produced almost at rest in the D^* rest frame. When boosting to the laboratory frame most of the D^* momentum will be transferred to the D^0-meson because of its large mass compared to that of the pion. Hence, the pion coming directly from the D^*-meson decay has the smallest momentum among the three decay particles ($K\pi\pi_s$) and is being referred to as the 'slow' pion (π_s). Although the branching ratio of this decay channel is small

$$\begin{aligned} \mathcal{BR}(D^{*\pm} \to K^{\mp}\pi^{\pm}\pi_s^{\pm}) &= \mathcal{BR}(D^{*\pm} \to D^0\pi_s^{\pm})\mathcal{BR}(D^0 \to K^{\mp}\pi^{\pm}) \\ &= (67.7 \pm 0.5)\%.(3.8 \pm 0.09)\% \\ &= (2.57 \pm 0.06)\%, \end{aligned} \tag{4.6}$$

the advantages which it offers compensate for it.

Due to the tight kinematic constraint one can achieve a clear signal with very good resolution, basically dominated by the $p_t(\pi_s)$ resolution, when plotting the invariant mass difference

$$\Delta M = M(K^{\mp}\pi^{\pm}\pi_s^{\pm}) - M(K^{\mp}\pi^{\pm}) \tag{4.7}$$

as shown in figure 4.5 (left). Using this ΔM tagging technique [29], instead of a simple invariant mass $M(K\pi\pi_s)$ calculation, the resolution is significantly improved, since several measurement errors cancel at least partially. In addition, the phase space for the combinatorial background, which grows with increasing ΔM, is significantly suppressed.

4.4.2 D^* Finder Algorithm

The D^* decay products are measured in the central tracking detectors (CJC, CIZ/COZ and CST). In order to reach a good signal to background ratio, only well measured tracks fitted to the primary vertex are considered as K, π or π_s candidates.[2] These tracks are used as input to the so-called D^* finder algorithm which works as follows.

D^0	$\lvert M(K\pi) - M(D^0)\rvert < 70$ MeV $p_t(K) + p_t(\pi) > 2$ GeV
D^*	$\Delta M = M(K\pi\pi_s) - M(K\pi) < 170$ MeV $1.5 < p_t(D^*) < 15$ GeV $\lvert\eta(D^*)\rvert < 1.5$

Table 4.1: D^*-candidate kinematic cuts.

[2] Although the D^0-meson is decaying weakly, due to very short mean decay distance of $c\tau = 123\ \mu m$, the daughter tracks are fitted to the primary vertex of the event.

In each event oppositely charged tracks are combined in pairs, assigning them the pion and kaon mass hypothesis. If their invariant mass lies inside the $M(D^0)$ mass window (see table 4.1) the search for a corresponding slow pion continues. The remaining tracks with opposite charge to that taken as the kaon are then added one by one, assuming the pion mass hypothesis, to form the D^*-meson candidate.

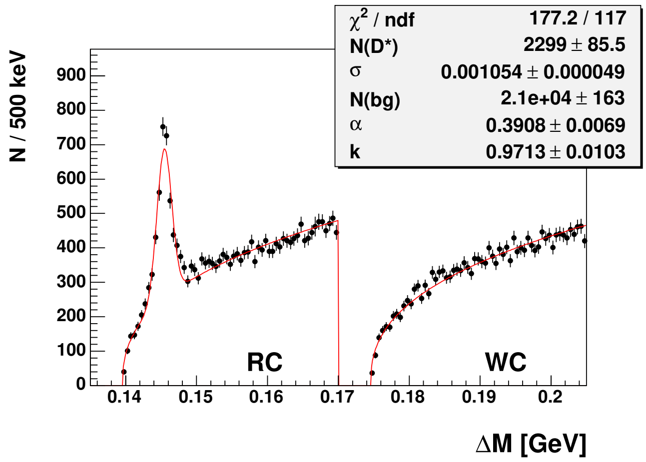

Figure 4.5: The ΔM invariant mass spectrum of selected 'right charged' (RC) and 'wrong charged' (WC) D^*-candidates for the 2000 data sample, together with the results of the fit. WC D^*-candidates are artificially shifted by a constant value to the right.

Due to limited geometrical acceptance and reconstruction efficiency of the central trackers, the kinematic range of a D^*-candidate is constrained by the cuts: $|\eta(D^*)| < 1.5$ and $1.5 < p_t(D^*) < 15$ GeV.[3] The invariant mass spectrum with resonant peak of correctly identified D^*-mesons is shown on figure 4.5 (left). Table 4.2 summarizes all kinematic cuts, together with the track quality cuts used for each of the decay particles separately.

4.4.3 D^* Signal Extraction

The number of D^*-mesons is obtained from a fit to the ΔM distribution. Unfortunately the signal peak sits exactly on top of a steeply rising background,

[3]These two cuts, together with the DIS cuts on Q^2 and y, define the D^* visibility range.

	Kaon	Pion	Slow Pion
p_{t}	> 0.25 GeV	> 0.25 GeV	> 0.12 GeV
Track length	> 18.9 cm	> 18.9 cm	> 10 cm
R_{start}	< 30 cm	< 30 cm	< 30 cm
N_{CJC} hits	> 10	> 10	> 10
$\lvert d'_{\mathrm{ca}}\rvert$	< 1 cm	< 1 cm	< 1 cm
$\Delta z = \lvert z'_{\mathrm{ca}} - z_{\mathrm{vtx}}\rvert$	< 20 cm	< 20 cm	< 20 cm
$\lvert d'_{\mathrm{ca}} \sin\theta\rvert$	< 0.5 cm	< 0.5 cm	< 0.7 cm
$\lvert \Delta z \sin\theta\rvert$	< 18 cm	< 18 cm	< 18 cm

Table 4.2: Track selection cuts for $K\pi\pi_s$ candidates.

therefore, the so-called 'wrong charged' background is used to stabilize the background fit in the signal region.

The wrong charged background consist of wrong charged 'D^*'-candidates, which are selected by the same algorithm and the same cuts as described above, only instead of combining a positively and a negatively charged track to reconstruct a D^0-candidate, two like sign tracks are taken. This 'D^0'-candidate is then combined with an oppositely charged track, the 'π_s', to form a wrong charged 'D^*', a $K^{\mp}\pi^{\mp}\pi_{\mathrm{s}}^{\pm}$ combination. The shape of the ΔM distribution for wrong charged (WC) combinations agrees well with the shape of the D^*-meson background in the right charged (RC) ΔM distribution as can be seen in figure 4.5 (in order to be able to plot both RC and WC distributions into one histogram the WC distribution is artificially shifted by the constant value of $m_0 = 0.035$ GeV to the right).

The ΔM distribution is fitted with a Gaussian for the signal and a power law function for the background:

$$\begin{aligned}
\mathrm{RC:} \quad & x \in (m_\pi, 0.17), & f(x) &= N_{D^*}\frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-\mu)^2}{2\sigma^2}) \\
& & &+ N_{\mathrm{bg}}N_{\mathrm{n}}(x-m_\pi)^\alpha \\
\mathrm{WC:} \quad & x \in (m_\pi + m_0, 0.17 + m_0), & f(x) &= kN_{\mathrm{bg}}N_{\mathrm{n}}(x-m_\pi-m_0)^\alpha
\end{aligned} \tag{4.8}$$

where N_{D^*} refers to the total number of D^*-mesons, σ and μ are the width and mean of the signal peak, N_{bg} is the background normalization, α is a parameter characterizing the steepness of the rising background, $N_{\mathrm{n}} = (\alpha+1)/(0.17-m_\pi)^{\alpha+1}$ is the normalization factor for background function and k is the relative normalization between RC and WC background. All these variables are used in the fit as free parameters. The results of the fit on full 2000 data statistics are listed in figure 4.5. Overall the studied data sample contains almost 2300 D^*-mesons.

For the ΔM fits in bins of differential distributions the mean position of the D^*-peak is kept fixed to the μ value obtained from the global fit (see figure 4.5).

4.4.4 Particle Identification

Further suppression of combinatorial background due to tracks with wrong mass hypothesis assignment can be achieved by particle identification via ionization losses.

The mean energy loss of a charged particle (except of electron) passing matter is described by the Bethe-Bloch formula [25]:

$$-\frac{dE}{dx} = 4\pi N_{\mathrm{A}} r_{\mathrm{e}}^2 m_{\mathrm{e}} c^2 z^2 \frac{Z}{A}\frac{1}{\beta^2}[\frac{1}{2}\ln\frac{2m_{\mathrm{e}}c^2\beta^2\gamma^2 T_{\mathrm{max}}}{I^2} - \beta^2 - \frac{\delta}{2}] \tag{4.9}$$

N_{A} denotes Avogadro's number, r_{e} the classical electron radius, m_{e} the electron mass, z the charge of the incident particle, Z the atomic number, A the atomic mass and I the effective ionization potential of the traversed matter, and δ a correction for the density effect. The variable T_{max} refers to the maximum kinetic energy which can be transferred to a free electron in a single collision. The Lorentz variables $\beta = p/E$ and $\gamma = 1/\sqrt{1-\beta^2}$ are defined as usual.

The ionization loss dE/dx depends only on the particle's velocity (or $\beta\gamma$). At small $\beta\gamma$ values dE/dx decreases quadratically towards the $\beta\gamma \approx 4$, where a minimum is reached. This is followed by a slow logarithmic rise, which at large $\beta\gamma$, due to density effects, ends in a plateau. Particle identification can be done at H1 via measurement of ionization losses in the region of steep fall $\sim 1/\beta^2$ using the CJC detector. Due to detector and track reconstruction effects and threshold requirements, the measured dE/dx values differ from the theoretical expectations (equation 4.9). Therefore the measured values need to be empirically parametrized

$$-\frac{dE}{dx} = a_1 z^2 \beta^{-a_2}(1 + a_3 e^{-a_4 \log(0.25+\beta\gamma)}) \tag{4.10}$$

as given in [63]. At the same time run dependent corrections are applied in order to achieve a better dE/dx description.

For identification purposes it is convenient to plot the dependence of dE/dx on the logarithm of the measured particle momenta, as the curves for different particles (e.g. p, K, π) are then shifted with respect to each other by $\log(m)$ as shown in figure 4.6 (left).

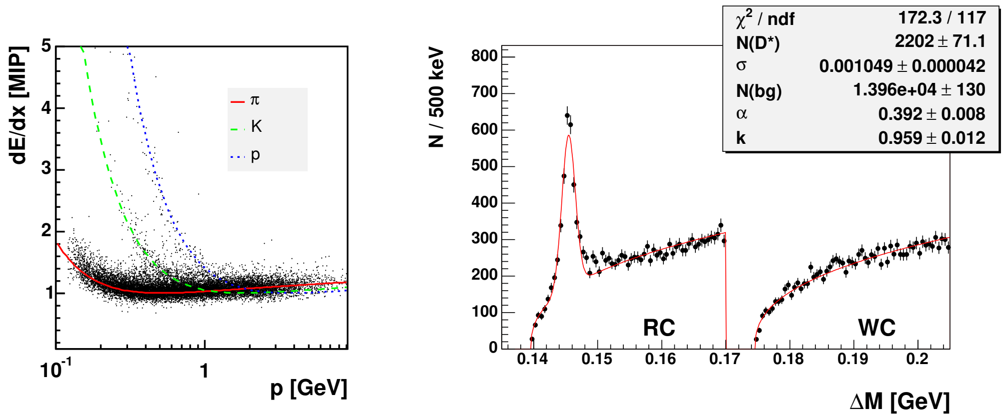

Figure 4.6: Left: ionization loss dE/dx as a function of momenta for central tracks in events with a D^*-candidate. The drawn curves show expected $< dE/dx >$ values for pions, kaons and protons. Right: ΔM invariant mass distribution of the D^*-candidates after applying cuts on normalized dE/dx likelihoods. The full curve denotes the fitted function as defined in section 4.4.3.

With the dE/dx information and the measurement uncertainties $\sigma(dE/dx)$ and $\sigma(p)$, the likelihood of a track to be due to a K, π or p can be calculated. The sum of these likelihoods is normalized to unity. To suppress the combinatorial background from tracks with obviously wrong mass assignment, momentum dependent cuts on normalized likelihoods are applied on D^* daughter particles. The cut values listed in table 4.3 are adjusted in a way that the combinatorial background is effectively suppressed and at the same time the signal efficiency remains high .

In case of $K\pi$ combinations from decays of D^0-candidates only tracks with reliable dE/dx measurements, i.e. with more than 10 dE/dx hits, are accepted.

The signal to background ratio after particle identification cuts improves from 0.73 to 1.05 (see figure 4.6 (right)).

Momentum range	Kaon $LH_n(K)$	Pion $LH_n(\pi)$	Slow Pion $LH_n(\pi)$
$p < 0.7$ GeV	5%	20%	5%
$0.7 < p < 1.2$ GeV	5%	5%	5%
$p > 1.2$ GeV	–	–	–

Table 4.3: Lower cuts on the dE/dx normalized likelihoods applied on K, π and π_s candidates.

4.4.5 Signal Extraction in MC

In contrast to the data, with its large combinatorial background, the situation with MC events is different, since the events are generated such that at least one generated D^*-meson decaying into $K\pi\pi_s$ is present. But this does not necessarily mean that background subtracted data can be directly compared with Monte Carlo prediction.

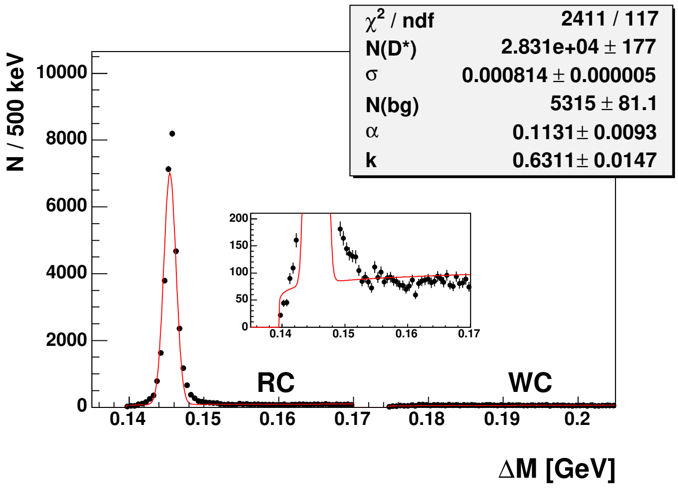

Figure 4.7: ΔM invariant mass distribution for Monte Carlo sample, with the zoomed background region. The full curve denotes the fitted function as defined in section 4.4.3.

As can be seen in figure 4.7 the ΔM distribution still contains a very small fraction of background (roughly 2.6% in signal region) from random track combinations which cannot be neglected. Performing the same kind of fit as in data, to obtain the number of reconstructed D^*-mesons is not the right approach, since the signal shape is not any longer Gaussian[4], which leads to a large χ^2.

Therefore, instead of fitting the signal, only simple statistical WC background subtraction is used. In the non-resonant region, $\Delta M > 0.155$ GeV, the relative right charge to wrong charge background normalization k is calculated, which is then used in the signal region $|\Delta M - \mu| < 3\sigma$ to subtract the wrong charge background:

$$\text{Signal} = \text{RC} - \frac{1}{k}\text{WC} \tag{4.11}$$

[4]Even in the data the signal is not of a Gaussian shape (a small tail on right side of the peak is present), but due to the large statistical errors the χ^2 is acceptable. Besides this feature, one can see that the resolution of track momenta is overestimated in MC simulation, since the peak width is significantly smaller than in data.

Chapter 5

Fragmentation Measurement

In this chapter the experimental results obtained by the hemisphere and jet methods, using so-called hadronic final state objects, are presented. In order to be able to compare them with measurements done at other experiments and to extract the fragmentation parameters, the measured distributions are corrected for detector and QED radiative effects.

5.1 Hadronic Final State

For the hemisphere and jet methods, which have been introduced in chapter 2, it is essential to have events with a properly reconstructed hadronic final state. In this analysis the information from the tracking system and the calorimeter are combined to create hadronic final state (HFS) objects. These objects are then used as input for the jet finder or for the thrust calculation in case of the hemisphere method.

The HFS objects are built on the basis of the H1 Hadroo2 algorithm [52]. The algorithm benefits from the advantages of track and cluster measurements in different regions of phase-space. In general charged particles at low transverse momenta can be measured rather precisely in tracking chambers. This precision decreases with growing p_{t}. The calorimeter can measure charged as well as neutral particles, which are quite frequently produced via decays of unstable resonances (on average every third produced particle is neutral). In contrast to the behavior of the tracking system, the uncertainty of calorimeter cluster energy measurement decreases with increasing energy.

The Hadroo2 algorithm starts with a list of good quality tracks, fulfilling the so-called Lee-West criteria [43], and a list of LAr and SPACAL clusters.[1] Various noise finders are applied to identify and remove noise related clusters in addition to the noise suppression already done during the reconstruction of clusters [2]. A neighboring track and cluster(s), which most likely belong to the same particle, are merged to form an HFS object. Three different configurations can occur: an HFS object is built out of a track and one or several clusters, from a single track or from one or several clusters only.

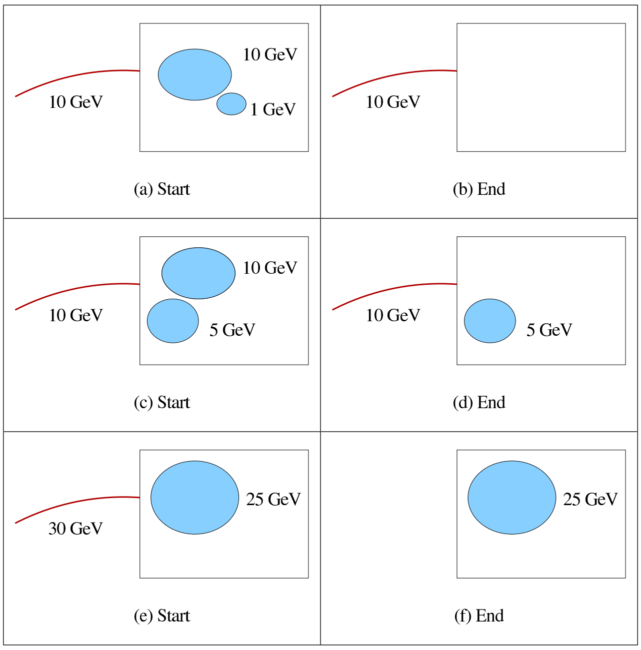

Figure 5.1: Sketch of Hadroo2 algorithm for the construction of HFS objects: a-b): the precisely measured track is kept and all calorimetric information is removed. c-d): in addition to a track a non-track related energetic cluster is found, hence two objects are created. e-f): the more precise calorimetric cluster energy is used.

For each HFS object the algorithm compares the relative energy resolution obtained from the tracker measurement (each track is assumed to originate from a pion $E_{\text{track}} = \sqrt{m_\pi^2 + p_{\text{track}}^2}$) with the expected resolution of

[1] Isolated leptons like the scattered electron, which have already been identified, are kept unchanged, and the corresponding tracks and clusters are excluded from any additional treatment.

$E_{\rm track}$, if measured by the calorimeter, taking possible fluctuations of each measurement within its standard error into account.

In the case that the track resolution is better and its energy is found to be compatible within 1.96 σ with the calorimetric energy deposition, the track information is used (see figure 5.1 a,b).

If the track resolution is better, but the energy deposition in the calorimeter is much larger than expected from the track measurement, the track energy $E_{\rm track}$ is subtracted from the matched cluster's energy to avoid double counting, and two HFS objects are created (figure 5.1 c,d). One object containing the track information and a remaining cluster object with energy equal $E_{\rm clus} - E_{\rm track}$, which is assumed to come from one or more neutral particles.

If the expected calorimeter energy resolution is more accurate than the measured track resolution, the calorimeter information is used and the track is discarded (figure 5.1 e,f).

The ratio of the reconstructed hadronic final state energy to the generated energy, obtained by RAPGAP MC simulation (figure 5.2 a), shows that the Hadroo2 algorithm performs well, since the ratio $E_{\rm HFS}/E_{\rm gen}$ peaks close to one. For events to be considered the generated particles were required to fulfill minimal detection criteria: either their energy was above noise threshold, or for charged particles the transverse momenta passed the track selection criteria.

The quality of hadronic final state reconstruction as well as the description of data by Monte Carlo simulation can be further checked by looking at the $p_{\rm t}$ balance of the hadronic final state and the scattered electron. In the ideal case, when all produced particles are fully detected, the ratio $p_{\rm t,had}/p_{\rm t,e}$ equals 1. In reality, the measured distribution is smeared around the expected value of 1, due to the detector resolution as shown in figure 5.2 b. The description of the data by Monte Carlo simulation could be partially improved by an energy calibration of hadronic clusters. Especially at low Q^2, this appears difficult and is not yet available.

5.2 Trigger Efficiency

The detector simulation used in this analysis already includes a simulation of the trigger. But before drawing any conclusion from the comparison of measured distributions with MC predictions, it is important to cross-check

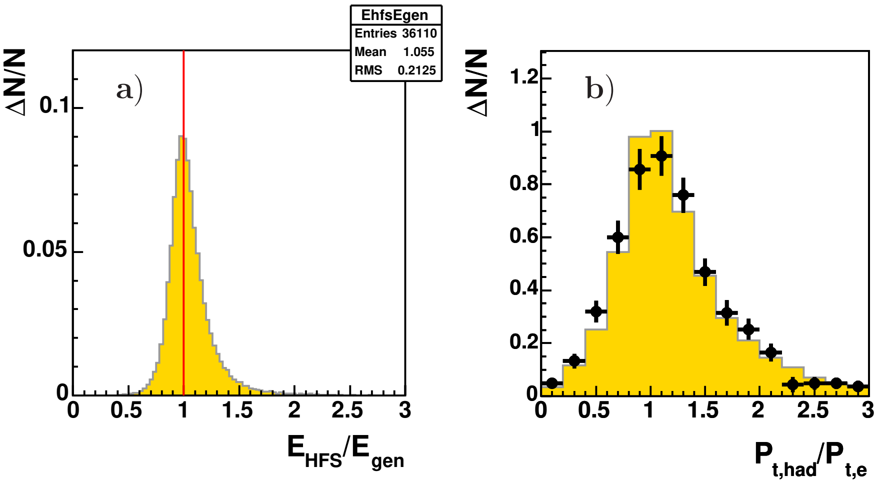

Figure 5.2: a) The ratio of the total energy of the HFS objects to sum of the energy of the generated particles per event as obtained by Monte Carlo simulation. b) The distribution of the measured ratio $p_{\mathrm{t,had}}/p_{\mathrm{t,e}}$ in data compared to Monte Carlo simulation (yellow filled area).

whether MC describes the trigger efficiencies correctly. If the description turns out not to be satisfactory, correction factors need to be applied to MC simulation.

The trigger efficiency in data can be in general calculated by analyzing events triggered by an independent subtrigger with sufficiently loose requirements. The efficiency can be expressed as the ratio of events passing all analysis cuts, which were fired by the subtrigger STa to be studied and an independent subtrigger STb as well, to the total number of events which were fired by subtrigger STb:

$$\varepsilon_{\mathrm{trig}} = \frac{N_{\mathrm{STa\&STb}}}{N_{\mathrm{STb}}} \tag{5.1}$$

The subtrigger ST61 (see section 4.1) relevant for this analysis consists of three independent parts: an energy deposition in the backward SPACAL calorimeter and the track related DCRPhi and z-Vertex trigger elements (TE). Hence they can be evaluated separately. The time of flight trigger elements, used as a non-ep background veto in ST61, are assumed to be 100% efficient.

After cutting out SPACAL regions where photomultipliers with small gain caused smaller trigger efficiencies, the efficiency for electron identification is nicely described by MC simulation and reaches almost 100%. This has already been shown in several analysis, i.e. [61] which is based on the same electron and D^* selection cuts as this analysis.

The efficiency of the track related trigger elements for the data is determined with a sample of events which contain besides a scattered electron also a D^*-meson candidate, and which were triggered with an independent subtrigger. As independent subtriggers the inclusive electron monitor triggers ST0, ST1, ST3, ST4 and ST9, for convenience referred to as STall, were chosen. For calculation of the trigger efficiency error the formula $\sigma(\varepsilon) = \varepsilon\sqrt{(\frac{\sigma(N_{\mathrm{STb}})}{N_{\mathrm{STb}}})^2 + (1-2\varepsilon)(\frac{\sigma(N_{\mathrm{STa\&STb}})}{N_{\mathrm{STa\&STb}}})^2}$ was used [33], since $N_{\mathrm{STa\&STb}}$ and N_{STb} in equation 5.1 are highly correlated as $N_{\mathrm{STa\&STb}}$ is a subset of N_{STb}.

The overall ST61 trigger efficiencies for data and MC are summarized in table 5.1. The trigger efficiency of the MC simulation was calculated as the ratio of events with a positive trigger decision to the number of all events with a reconstructed D^*-meson.

TE condition	data	RAPGAP
DCRPhi_THigh	98%	98%
zVtx_sig	93%	94%
ST61	91%	92%

Table 5.1: Overall trigger efficiencies for data and Monte Carlo simulation.

The efficiencies of the DCRPhi and z-Vertex trigger elements for data and Monte Carlo are shown in figure 5.3 as a function of the D^* fragmentation observables, $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$, and p_{t} and η of the D^*-meson. The DCRPhi efficiency is rather high as expected and is well described by MC simulation. For the z-Vertex significance efficiency the distributions show small deviations, which are, when taking into account statistical errors, of little significance.

Some of the SPACAL subtriggers used for the determination of the trigger efficiencies contain besides the SPACAL trigger elements and time of flight veto conditions also track based veto conditions, which can in principle introduce a bias; in other words events triggered by the affected subtriggers can be correlated to some extent with events containing the DCRPhi or z-Vertex trigger elements, whose efficiencies are to be measured, and thus the assumption of independence of equation 5.1 may not be fulfilled. Hence, an

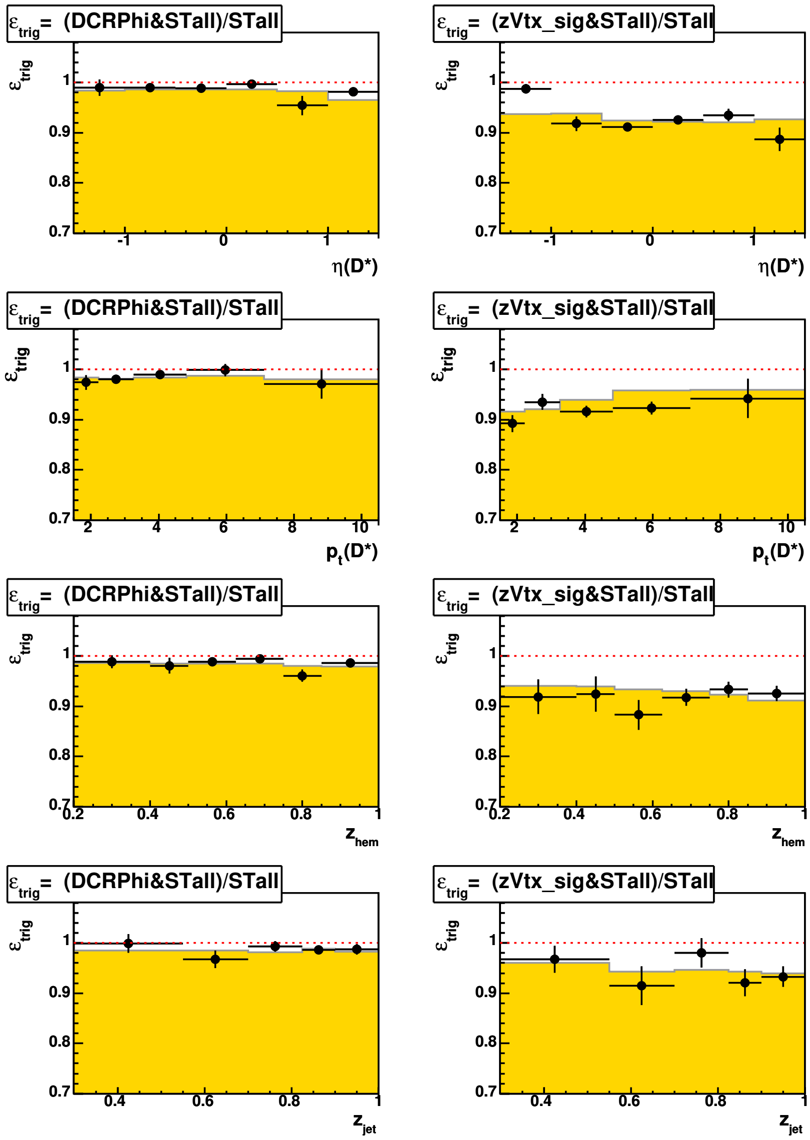

Figure 5.3: The measured efficiency of DCRPhi_THigh (left column) and zVtx_sig (right column) trigger elements based on subtriggers ST0, ST1, ST3, ST4, ST9 as a function of $p_t(D^*)$, $\eta(D^*)$, z_{hem} and z_{jet} compared with detector simulation (yellow filled area).

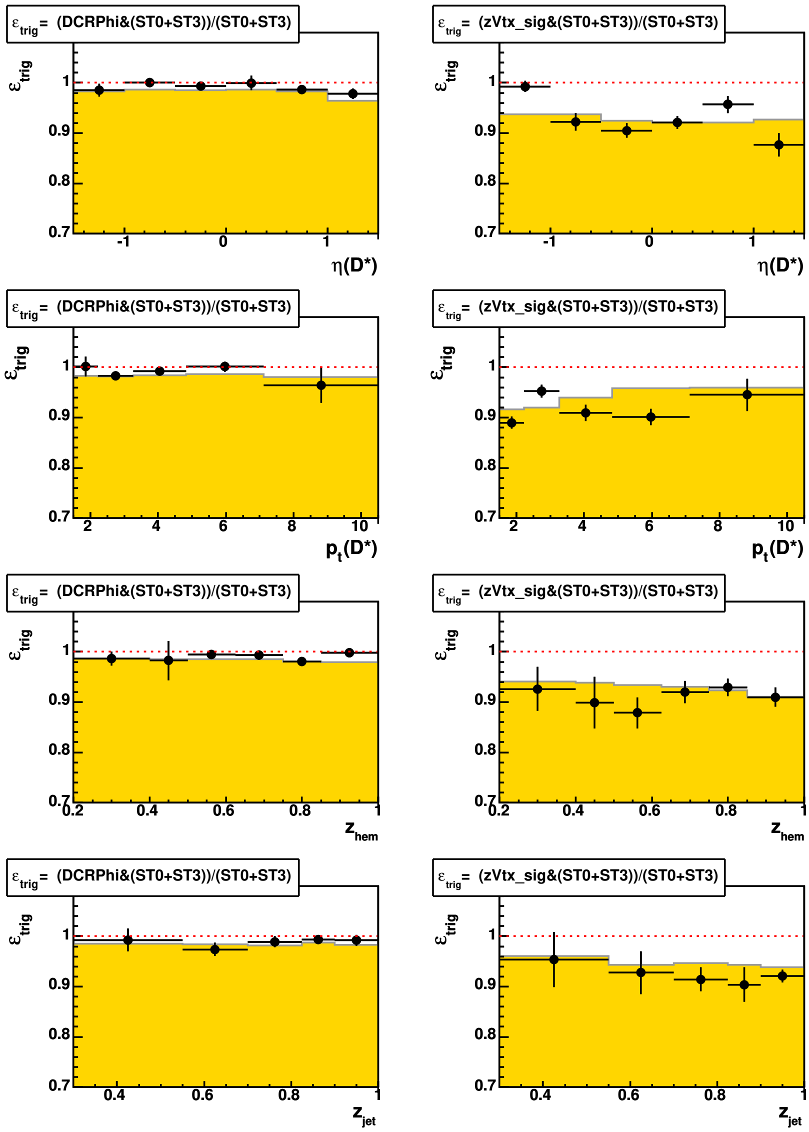

Figure 5.4: The measured efficiency of DCRPhi_THigh (left column) and zVtx_sig (right column) trigger elements based on subtriggers ST0 and ST3 as a function of $p_{\mathrm{t}}(D^{*})$, $\eta(D^{*})$, $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ compared with detector simulation (yellow filled area).

extensive study was performed only with a subset of events triggered by SPACAL subtriggers without any track veto condition requirement. These are the subtriggers ST0 and ST3. Unfortunately the statistics for these events is quite limited as they are heavily prescaled. Moreover, the ST0 and ST3 subtriggers do not include the inner part of the SPACAL, where approximately one half of the events triggered by ST61 have their scattered electron located. Therefore this study is considered only as a rough check.

The obtained DCRPhi_THigh and zVtx_sig trigger elements efficiencies are shown in figure 5.4 for the same set of variables as before. Except for a few bins there is very good agreement with the efficiencies determined using the full subtrigger sample. Thus, one can conclude that the ST61 trigger efficiency is sufficiently well described by Monte Carlo simulation.

5.3 Hemisphere Observable: $\mathrm{z}_{\mathrm{hem}}$

Using HFS objects the hemisphere observable $\mathrm{z}_{\mathrm{hem}}$ can be calculated for events containing a D^*-meson candidate. The HFS objects are boosted into the γp-frame, where the hemisphere containing the fragmentation products of the D^* parent charm quark is determined on the basis of the two dimensional thrust-axis, which is determined in the plane perpendicular to the virtual photon direction (see section 2.5.1).

For events with two or more reconstructed D^*-candidates a sort of equal treatment has been introduced, since in principle it is not possible to distinguish which of the D^*-candidates are real and which are due to the combinatorial background. For each D^*-candidate a corresponding $\mathrm{z}_{\mathrm{hem}}$ value is calculated, so that in the end such kind of an event is equivalent to n events (n being the number of D^*-candidates contained in the event) with the same event kinematics but different D^*-candidates.[2] In data the fraction of events with two (three or more) D^*-candidates, due to the large combinatorial background, is at the 7.4% (0.8%) level, whereas in the charm signal Monte Carlo simulation this fraction is negligible.

Since the Monte Carlo simulations (mainly RAPGAP) with default parameter settings failed to describe basic distributions such as $p_{\mathrm{t}}(D^*)$ or $\eta(D^*)$, the MC events had to be reweighted in order to correct for these deficiencies. The reweighting procedure is described in appendix A.

[2] All D^* related distributions are treated similarly.

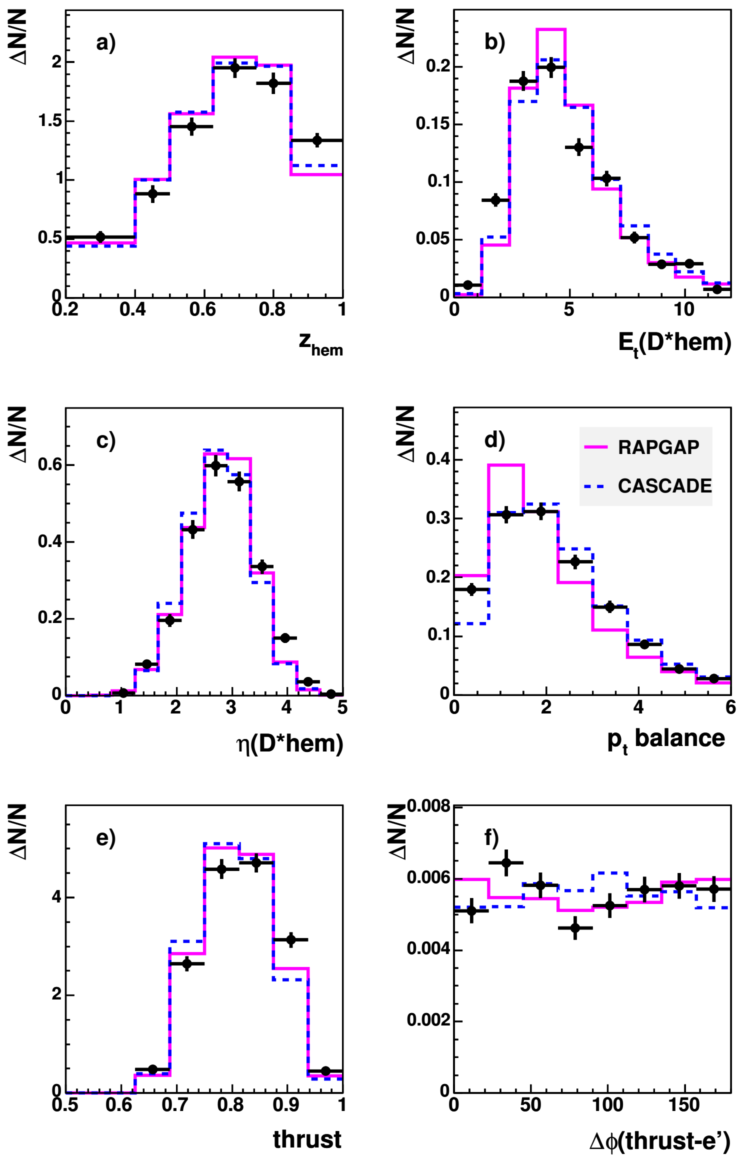

Figure 5.5: Description of hemisphere related observables in data by RAPGAP and CASCADE Monte Carlo models: a) z_{hem}, b) the transverse energy $E_{\mathrm{t}}(D^{*}\mathrm{hem})$, c) the pseudorapidity $\eta(D^{*}\mathrm{hem})$, d) the p_{t} balance in the current hemisphere, e) the thrust value and f) the angle between constructed thrust-axis and scattered electron vector in a plane perpendicular to the virtual photon direction in γp-frame.

Figure 5.5 a) shows the shape normalized z_{hem} distribution after background subtraction for data overlayed with RAPGAP and CASCADE MC expectation. Except for small discrepancies in the first and last z_{hem} bins the distribution is well described. The very low z_{hem} bin in the interval $z_{\mathrm{hem}} \in (0., 0.2)$ is dominated by background events such that it was not possible to extract the small signal expected in data within reasonable error. Hence, this bin was dropped. Towards higher z_{hem} values the background significantly decreases so that the highest bin is almost background free.

Figure 5.5 b-f) shows other interesting observables characterizing the D^*-hemisphere: the transverse energy and the pseudorapidity of the D^*-hemisphere in the γp-frame, as well as the p_t balance of particles in the current hemisphere, the thrust magnitude and the angle between the reconstructed thrust-axis and the scattered electron measured in the plane perpendicular to the virtual photon direction, which are important for the hemisphere determination.

5.4 Jet Observable: z_{jet}

Jets are reconstructed in the γp-frame using the inclusive $k_\perp$-algorithm with the E-recombination scheme (see section 2.5.2). As input hadronic final state objects were used as for the hemisphere method.

The measured observable z_{jet} as well as distributions of $E_t(D^*\mathrm{jet})$, $\eta(D^*\mathrm{jet})$ and of the distance between D^*-meson and D^*-jet directions, $\Delta R(D^*\mathrm{jet} - D^*) = \sqrt{(\eta_{D^*\mathrm{jet}} - \eta_{D^*})^2 + (\phi_{D^*\mathrm{jet}} - \phi_{D^*})^2}$, are in general very well described by the MC simulations[3] (see figure 5.6). Also the fraction of jets which consist only of the D^*-meson itself (roughly 10%) is within the statistical errors in very good agreement with MC prediction.

The region at small z_{jet} values is, as it was the case for small z_{hem} values, background dominated. Hence z_{jet} has been measured only within the interval $z_{\mathrm{jet}} \in (0.3, 1.)$.

5.5 Correcting Data to Hadron Level

For further analysis it is convenient and useful to correct the measured distributions to hadron level. This includes detector as well as QED-radiative

[3]The Monte Carlo simulations are reweighted in observables $p_t(D^*)$, $\eta(D^*)$ and z_{vtx}.

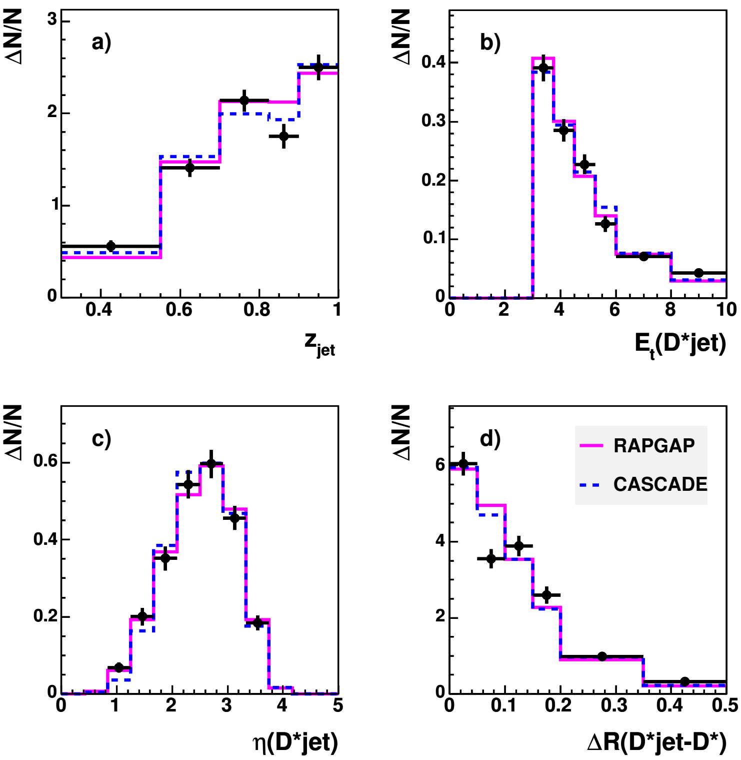

Figure 5.6: Description of jet related observables in data by RAPGAP and CASCADE Monte Carlo models: a) $z_{\rm jet}$, b) the transverse energy $E_{\rm t}(D^*{\rm jet})$, c) the pseudorapidity $\eta(D^*{\rm jet})$ and d) the distance between the D^*-jet and D^*-meson vectors measured in $\eta - \phi$ plane.

corrections. Having the fragmentation spectra corrected to hadron level one can easily compare them with predictions of various phenomenological models, fragmentation parametrizations and one can extract the fragmentation parameters for Monte Carlo models and even NLO calculations.

The following procedure was carried out to correct data to hadron level:

- **Beauty subtraction**
  In a first step the beauty contribution has been subtracted from the already background subtracted data. For this purpose the RAPGAP Monte Carlo prediction for beauty production including detector simulation has been used.

- **Hadron level corrections**
  The beauty subtracted data are then corrected for detector effects and the $D^* \rightarrow D^0\pi_\mathrm{s} \rightarrow K\pi\pi_\mathrm{s}$ reconstruction efficiency using the charm Monte Carlo simulation reweighted in $p_\mathrm{t}(D^*)$, $\eta(D^*)$ and z_vtx as described in appendix A. For this purpose a simple minded bin by bin unfolding method have been used.

- **QED corrections**
  Since theoretical models usually do not include higher orders of electroweak interactions, like real photon emission by the incoming respectively outgoing electron and the corresponding virtual corrections, the measurement is corrected for these effects.

5.5.1 Subtraction of Beauty Contribution

The D^*-mesons coming from weak decays of beauty hadrons have a significantly softer fragmentation spectrum than the D^*-mesons produced directly in charm events. The RAPGAP Monte Carlo prediction for beauty production has been used in this analysis to estimate this contribution. Figure 5.7 shows both measured fragmentation spectra normalized to luminosity together with the beauty MC expectation. According to the most recent H1 F_2^c and F_2^b measurements [7] carried out in a similar Q^2 range as used in this analysis, the RAPGAP beauty production cross section is in more or less good agreement with data. In case of $\mathrm{z_{hem}}$ the fraction of beauty events contributing is around 1.7% and for $\mathrm{z_{jet}}$ it is 2.2%.

While for the hemisphere distribution the beauty component is as expected clearly peaking at low $\mathrm{z_{hem}}$ values, for the $\mathrm{z_{jet}}$ distribution it is of rather flat shape. Since the beauty jets are according to the MC simulation broader than the charm jets, in some cases only part of the b quark fragmentation products are collected within a D^*-jet, which leads to overestimated $\mathrm{z_{jet}}$ values. The more robust hemisphere method is not much affected by this jet property.

5.5.2 Unfolding Method

The bin by bin unfolding method completely relies on a correct description of the underlying physics processes and a correct simulation of the detector response by the Monte Carlo. The correction factor for a certain bin i of the

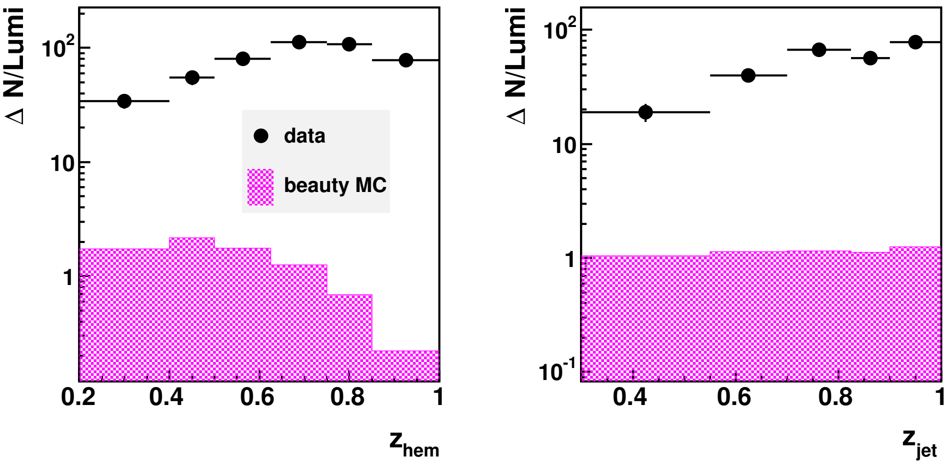

Figure 5.7: Luminosity normalized fragmentation spectra with indicated beauty contribution as predicted by the RAPGAP MC: for the hemisphere method (left), for the jet method (right).

measured data distribution can be expressed as a ratio of the bin content on hadron level to the bin content on detector level

$$C_i = \frac{N_i^{\text{had}}}{N_i^{\text{det}}} \tag{5.2}$$

This procedure works only, if the migrations of events from one bin into other bins are reasonably small. Therefore, it is necessary to choose proper binning for the observed quantities in order to reduce possible migrations. To quantify the amount of migrations and to make sure that usage of bin based corrections is appropriate, the following quantities are introduced:

$$\text{purity}: \quad P_i = \frac{N_i^{\text{det\&had}}}{N_i^{\text{det}}} \tag{5.3}$$

$$\text{stability}: \quad S_i = \frac{N_i^{\text{det\&had}}}{N_i^{\text{had}}}, \tag{5.4}$$

where N_i^{had} is the number of generated D^*-mesons (D^*-jets) within the visible range in bin i, N_i^{det} is the number of reconstructed D^*-mesons (D^*-jets) and $N_i^{\text{det\&had}}$ is the number of D^*-mesons which have been reconstructed and generated in the same bin.

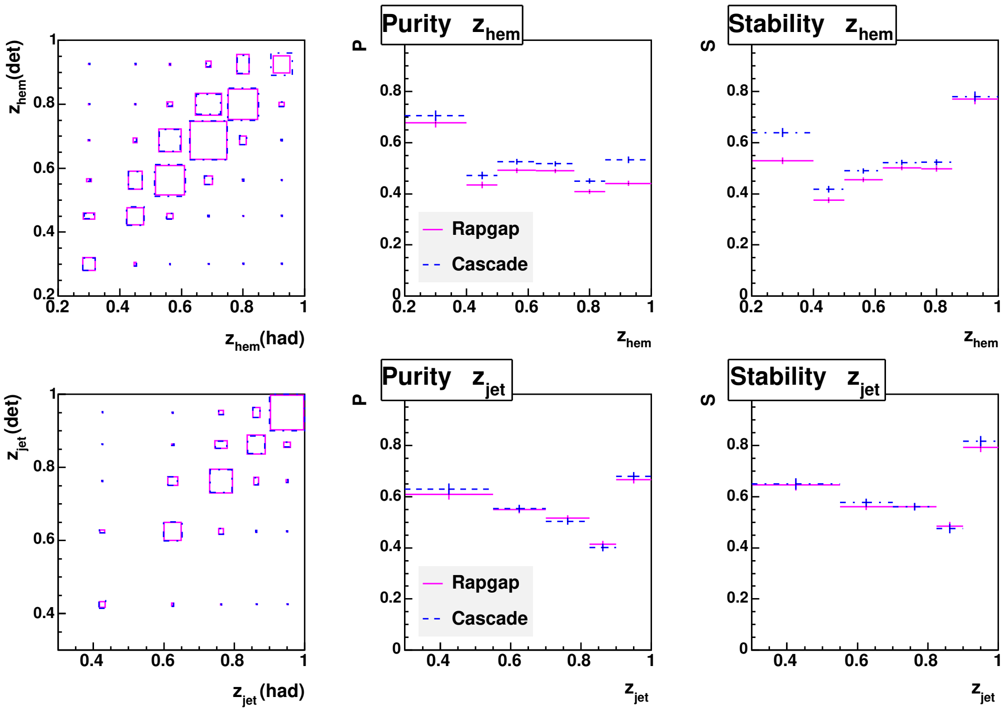

Figure 5.8: The correlation matrix, purity and stability as a function of z_{hem} (top) and z_{jet} (bottom) as obtained from the RAPGAP and CASCADE MC simulations.

The bin widths have been adjusted in such a way that the purity is always above 40%. This is the lower tolerable limit for bin by bin unfolding. The purity, stability and correlation matrix between the hadron and detector level are shown in figure 5.8.

Inclusion of QED radiative corrections implies modification of the equation 5.2. In general, two Monte Carlo simulations are used, one with QED radiative effects implemented and a second one without. The total bin correction factor can be written as:

$$C_i = \frac{N_i^{\mathrm{had}}}{N_i^{\mathrm{det,rad}}} \tag{5.5}$$

where subscript 'rad' denotes the inclusion of QED radiative effects. Assuming that the QED and detector corrections can be factorized one obtains:

$$C_i = \frac{N_i^{\mathrm{had}}}{N_i^{\mathrm{had,rad}}} \frac{N_i^{\mathrm{had,rad}}}{N_i^{\mathrm{det,rad}}} = C_i^{\mathrm{QED}} C_i^{\mathrm{det}} \tag{5.6}$$

The radiative corrections are calculated using the HERACLES 4.63 [42] program. Unfortunately this program can be interfaced only with the RAPGAP MC for the generation of direct processes but not for resolved photon processes. Therefore the C_i^{QED} radiative correction factors for RAPGAP resolved and also for CASCADE are taken to be the same as for RAPGAP direct. Hence, the final formulas for RAPGAP and CASCADE correction factors are as follows:

$$\text{RAPGAP}: \qquad C_{i,\mathrm{Rap}} = C_{i,\mathrm{RapDir}}^{\mathrm{QED,had}} \frac{N_{i,\mathrm{RapDir}}^{\mathrm{had,rad}} + \frac{1}{C_{i,\mathrm{RapDir}}^{\mathrm{QED,had}}} N_{i,\mathrm{RapRes}}^{\mathrm{had}}}{N_{i,\mathrm{RapDir}}^{\mathrm{det,rad}} + \frac{1}{C_{i,\mathrm{RapDir}}^{\mathrm{QED,det}}} N_{i,\mathrm{RapRes}}^{\mathrm{det}}} \tag{5.7}$$

$$\text{CASCADE}: \qquad C_{i,\mathrm{Cas}} = C_{i,\mathrm{RapDir}}^{\mathrm{QED,had}} \frac{\frac{1}{C_{i,\mathrm{RapDir}}^{\mathrm{QED,had}}} N_{i,\mathrm{Cas}}^{\mathrm{had}}}{\frac{1}{C_{i,\mathrm{RapDir}}^{\mathrm{QED,det}}} N_{i,\mathrm{Cas}}^{\mathrm{det}}} = \frac{N_{i,\mathrm{Cas}}^{\mathrm{had}}}{\frac{1}{C_{i,\mathrm{RapDir}}^{\mathrm{QED,det}}} N_{i,\mathrm{Cas}}^{\mathrm{det}}} \tag{5.8}$$

where $C_{i,\mathrm{RapDir}}^{\mathrm{QED,had}} = N_{i,\mathrm{RapDir}}^{\mathrm{had}} / N_{i,\mathrm{RapDir}}^{\mathrm{had,rad}}$ is the radiative correction factor on hadron level and $C_{i,\mathrm{RapDir}}^{\mathrm{QED,det}} = N_{i,\mathrm{RapDir}}^{\mathrm{det}} / N_{i,\mathrm{RapDir}}^{\mathrm{det,rad}}$ the radiative correction factor on detector level. In general the expression $(1/C^{\mathrm{QED}})N_{\mathrm{Cas}}$ (respectively $(1/C^{\mathrm{QED}})N_{\mathrm{Rap,Res}}$) either on hadron or on detector level stands for CASCADE (RAPGAP resolved) MC corrected for QED effects.

In contrast to the jet method, where radiative corrections play just a minor role ($\sim 6\%$), the sensitivity of the hemisphere method to radiative effects is significantly larger especially at low $\mathrm{z_{hem}}$ values ($\sim 3-8\%$ and in lowest $\mathrm{z_{hem}}$ bin even 15%).

The fragmentation spectra corrected to hadron level are shown in figure 5.9. The data points contain besides the statistical error also the systematic error which will be discussed in the next section.

5.6 Systematic Uncertainties

In order to study the sensitivity of the fragmentation spectra to systematic uncertainties, many checks have been performed. Some of the uncertainties arise from incomplete knowledge of detector effects, some from incomplete

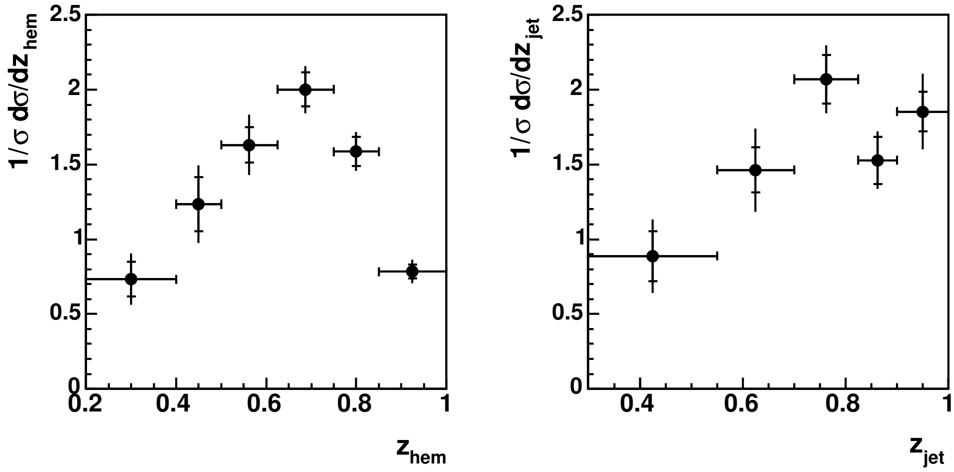

Figure 5.9: The fully corrected and shape normalized fragmentation spectra. The hemisphere method observable z_{hem} (left), the jet observable z_{jet} (right).

knowledge of the underlying physics which is incorporated in the Monte Carlo model, and some from the cuts and extraction methods used in this analysis.

To estimate the effects of a given systematic uncertainty, the Monte Carlo simulation, with a quantity under investigation modified according to its systematic uncertainty, is used to correct the measured data to hadron level. The difference between the spectrum obtained in this way and the spectrum obtained with the 'nominal' MC simulation, is a measure for the systematic error of a given source. Those systematic uncertainties which are considered to be most important are discussed below.

5.6.1 Experimental Systematic Uncertainties

5.6.1.1 Hadronic energy scale

The uncertainty of the calorimetric hadronic energy scale propagates to the reconstruction of HFS objects. Hence the energy of all hadronic HFS clusters in the liquid argon calorimeter was scaled up and down by 5% [57] and the energy of SPACAL hadronic HFS clusters by 7% [5]. This variation covers the differences between data and Monte Carlo simulation. The obtained systematic error for the z_{hem} and z_{jet} distributions is less than 5% and 4%

respectively, but is strongly correlated between the bins.

5.6.1.2 Energy and polar angle of the scattered electron

The energy and polar angle of the scattered electron is not only used to determine the Q^2 and y kinematic region of an event, but also to perform the boost of the particle momenta from the laboratory frame to the γp-frame, and hence can indirectly affect the fragmentation spectra.

To study the systematic effects the energy of the scattered electron has been varied by $\pm 3\%$ for E'_e of 8 GeV and by $\pm 1\%$ for 27.6 GeV [56][57]. For intermediate energies the size of uncertainty is assumed to be linearly dependent on E'_e. The resulting systematic error has been found to be less than 1%.

Changing the polar angle θ'_e by ± 1 mrad, corresponding to BDC resolution, had only a small impact on the fragmentation spectra of roughly 0.5%.

5.6.1.3 Tracking efficiency

The uncertainty of the tracking efficiency can mainly affect the D^* tag. Since in this analysis the differential spectra are shape normalized, precise knowledge of the overall uncertainty is not as important as the proper description of the tracking efficiency in the $\eta - p_{\mathrm{t}}$ phase space by the Monte Carlo simulation. An uncertainty depending on η and p_{t} could distort the corrected z spectra.

The tracking efficiencies as predicted by MC have been varied for the D^* daughter particles according to two extreme scenarios. First, the tracking efficiency for low p_{t} and large $|\eta|$ was decreased and for central η and moderate and high p_{t} increased. In the second scenario the inverse variation was performed. The magnitude of the variation was chosen to be larger in corners of the phase space (small p_{t} and forward-backward η region of the CJC chamber) $\sim 4\%$, and smaller 2-3% in the central region which is expected to be better simulated. These limits on the variation of the uncertainty are based on an average track reconstruction inefficiency for $D^* \rightarrow K\pi\pi_s$ of $\sim 2\%$ [23] [60].

The resulting systematic error due to the uncertainty of tracking efficiency was found to be for both $\mathrm{z}_{\mathrm{hem}}$ and $\mathrm{z}_{\mathrm{jet}}$ below 1% except for the first and last bin of the $\mathrm{z}_{\mathrm{hem}}$ distribution where it reaches 2%.

5.6.1.4 dE/dx cuts

Since the description of the normalized dE/dx likelihoods by the Monte Carlo simulation is not satisfactory at larger track momenta, the dE/dx cuts for suppression of combinatorial D^* background have been switched-off in the data and in MC as well. This resulted in $\sim 3-6\%$ systematic uncertainty.

5.6.1.5 D^* signal extraction

The systematic error due to the signal extraction in the data has been estimated by a different extraction procedure. The procedure is based on the same fit of the ΔM invariant mass spectrum as has been described in section 4.4.3, but instead of relying on the number of D^*-mesons coming from the Gaussian normalization, the number is calculated by subtracting the integral of the fitted background function from the number of histogram entries in the signal region ($\mu \pm 3\sigma$). In this way one is less sensitive to the not entirely proper description of the signal shape by a Gaussian. On average the error has been found to be about 4%.

5.6.2 Theoretical Systematic Uncertainties

5.6.2.1 Resolved fraction

To determine the systematic error due to the uncertainty of the size of the resolved component, the nominal resolved photon fraction of 33% predicted by the RAPGAP Monte Carlo was changed to 10 and 50%. The hemisphere observable turns out to be much more sensitive to the resolved fraction than the jet observable. It has been found that a part of the photon remnant is contaminating the D^*-hemisphere, which has a direct impact on the hardness of the measured fragmentation spectrum. The effect of varying the resolved component was found to be below 6% for the $\mathrm{z}_{\mathrm{hem}}$ and 2% for the $\mathrm{z}_{\mathrm{jet}}$ spectrum.

5.6.2.2 Beauty contribution

The nominal beauty contribution, which was taken according to the RAPGAP prediction, was increased by a factor of two. This leads to an effect on the fragmentation spectra below 2%.

5.6.2.3 Model dependence

Since the hemisphere related observables, as for example the p_{t} balance of the current hemisphere, are better described by the CASCADE MC, while the description of the jet observables is as good as the one by RAPGAP, the model dependence of the correction factors was investigated.[4] Besides the errors due to the dE/dx and the D^* signal extraction the model dependence is one of the most dominant sources of systematic uncertainty (3-5% effect).

sys. source	$\mathbf{z_{hem}}$ **error**	$\mathbf{z_{jet}}$ **error**
E'_{e}	0.5%	0.8%
θ'_{e}	0.5%	0.6%
E_{HFSclus}	2.3%	2.2%
Track. eff.	0.7%	0.5%
dE/dx cut	4.0%	5.0%
Signal extraction	4.8%	4.6%
Res. fraction	3.0%	1.0%
Beauty fraction	1.6%	1.0%
Model dependence	3.4%	5.3%
<Total sys.err.>	8.3%	9.1%

Table 5.2: The list of the relative systematic errors averaged over the bins of the $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ distributions.

The average systematic errors obtained are summarized in table 5.2. The total systematic error σ_{tot} for a given bin is calculated as follows

$$\sigma_{\mathrm{tot}} = \sqrt{\sigma^2_{\mathrm{stat}} + \sum_i \sigma^2_{i,\mathrm{sys}}} \quad , \tag{5.9}$$

where the index i runs over all sources of systematic uncertainties. As can be seen from figure 5.9, the statistical error σ_{stat} dominates over the systematic uncertainties, except for the last $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ bins.

[4] The dependence on the fragmentation model has been briefly investigated using the HERWIG MC [46], which in contrast to RAPGAP and CASCADE uses an alternative model for fragmentation - the cluster model. Unfortunately this Monte Carlo failed to describe the fragmentation spectra and hence was abandoned.

It should be mentioned that some of the systematic errors are strongly correlated between bins. Those are: the uncertainty of the hadronic energy scale, the resolved fraction and the beauty component. They are shown in figure 5.10. The remaining systematic errors are uncorrelated. The treatment of the correlated errors will be discussed in more detail in chapter 6.

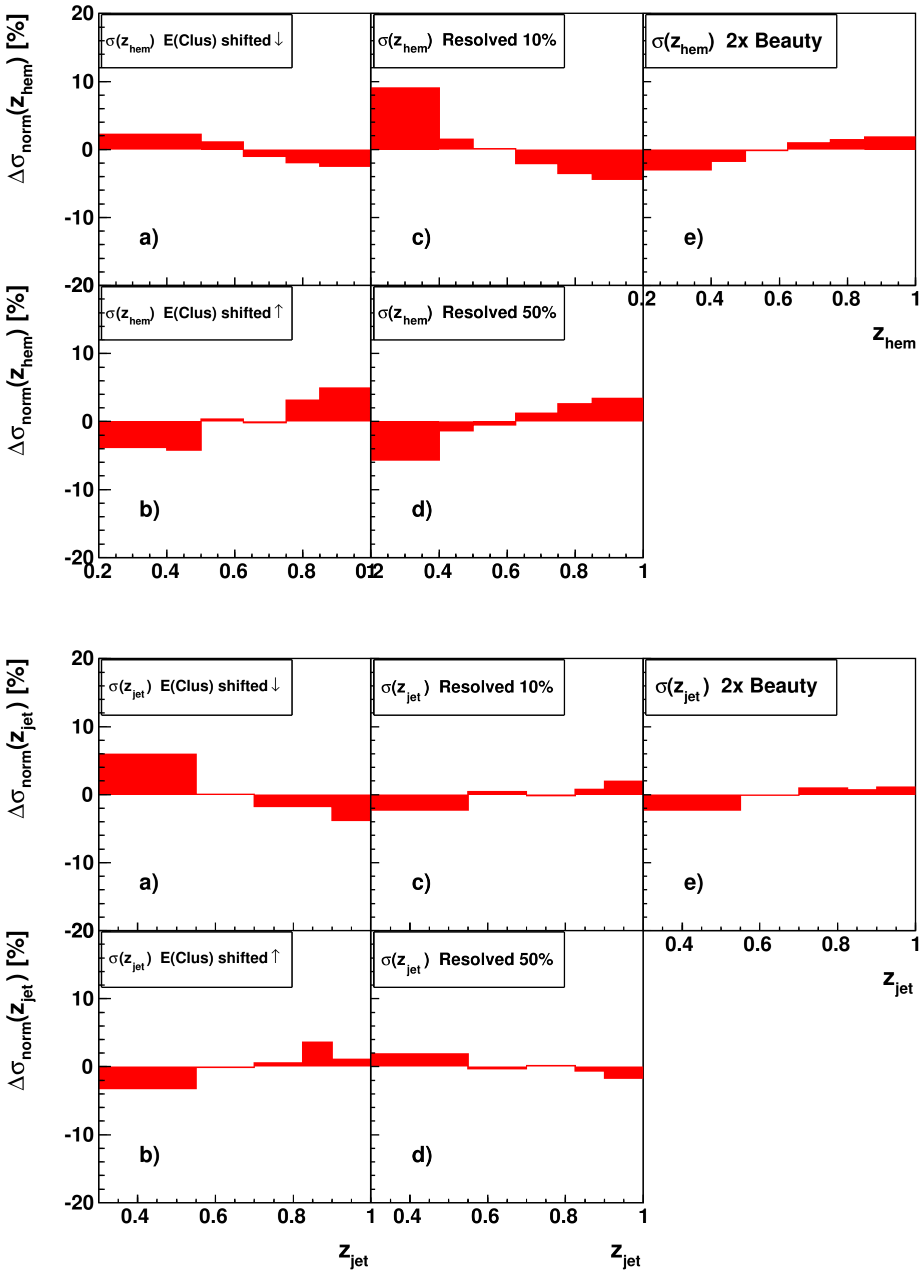

Figure 5.10: The relative correlated systematic errors as a function of z_{hem} (the upper set of plots) and z_{jet} (the lower set of plots). a-b) the influence of the hadronic energy scale variation, c-d) the variation of resolved fraction and e) the variation of the beauty contribution.

Chapter 6

Results

Using fragmentation spectra corrected to hadron level, the charm fragmentation parameters of the Peterson and Kartvelishvili non-perturbative fragmentation functions were extracted for the MC models RAPGAP and CASCADE and for a NLO (massive) calculation as well. In the end the fragmentation spectra were compared with e^+e^- measurements.

6.1 Extraction Procedure and Treatment of Correlated Systematic Errors

For the extraction of fragmentation parameters, cross sections as a function of $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ on hadron level were generated for various values of the respective fragmentation parameters. For the Peterson parametrization (equation 1.14) the value of ε was varied over a wide range from 0.01 up to 0.1, and for the Kartvelishvili parametrization (equation 1.15) in the range of 2 to 7. To determine the most optimal parameter value a method of least squares [26] (often referred to as χ^2 method) have been used. The χ^2 between data and model prediction was calculated using the full covariance matrix, taking statistical and uncorrelated systematic errors as well as correlations into account:

$$\chi^2(\varepsilon) = \sum_{jk} (\mathrm{z}_j - \mathrm{z}_j^{\mathrm{MC}}(\varepsilon)) S^{-1}_{jk} (\mathrm{z}_k - \mathrm{z}_k^{\mathrm{MC}}(\varepsilon)) \tag{6.1}$$

The sum runs over all z bins, ε is the tunable fragmentation parameter and S the covariance matrix with the following form

$$\mathbf{S} = \begin{pmatrix} \sigma^2_{\mathrm{stat},1} + \sum \sigma^2_{\mathrm{sys},1} + \sum_{\mathrm{cor}} \sigma'^2_{\mathrm{sys},1} & \sum_{\mathrm{cor}} \sigma'_{\mathrm{sys},1}\sigma'_{\mathrm{sys},2} & \cdots \\ \sum_{\mathrm{cor}} \sigma'_{\mathrm{sys},2}\sigma'_{\mathrm{sys},1} & \sigma^2_{\mathrm{stat},2} + \sum \sigma^2_{\mathrm{sys},2} + \sum_{\mathrm{cor}} \sigma'^2_{\mathrm{sys},2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$\sigma_{\mathrm{stat},i}$ denotes the statistical error, $\sigma_{\mathrm{sys},i}$ the uncorrelated systematic error and $\sigma'_{\mathrm{sys},i}$ the correlated systematic error for i-th bin of the z spectrum. The sources of systematic errors which were considered to be correlated are: the hadronic energy scale uncertainty, the fraction of the resolved component and the beauty component (see chapter 5.6).

The $\pm 1\sigma$ error of the fitted parameter ε' is then determined from the shape of the χ^2 distribution according to the rule $\chi^2(\varepsilon' \pm 1\sigma) = \chi^2_{\mathrm{min}} + 1$.

The fragmentation parameters were fitted for RAPGAP and CASCADE Monte Carlo models which use the same Lund-string fragmentation model as implemented in PYTHIA 6.2. For extraction of the fragmentation parameters the $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ distributions were shape normalized, in order to minimize any sensitivity to the total D^*-meson production cross section, which is not necessarily correctly predicted by Monte Carlo models.

6.2 Parameter Extraction for RAPGAP

First the parameter ε of the Peterson parametrization has been extracted. The dependence of χ^2 on the ε parameter for the $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ observables is shown in figure 6.1 (each point corresponds to a fully generated Monte Carlo set).

Close to the minimum, χ^2 is expected to depend quadratically on the parameter value, however over a larger range this is true only for a linear function of the parameter. In case of the Peterson parametrization the dependence is slightly asymmetric.[1] Therefore a superposition of two exponential functions was fitted through the points lying close to the minimum to obtain the best fit value and the 1σ confidence limits. In this way the sensitivity to fluctuations as well as to the limited number of Monte Carlo sets is being reduced. The results of the fits for the $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ methods are $\varepsilon = 0.022^{+0.007}_{-0.004}$ and $\varepsilon = 0.040^{+0.013}_{-0.009}$ respectively. The comparison of the

[1] In case of the Kartvelishvili parametrization this asymmetry gets even smaller.

measured fragmentation spectra with the model predictions for the fitted ε values is shown in figure 6.1 as well.

The Kartvelishvili fits have been performed in analogy to the Peterson fits. The dependence of χ^2 on α together with the comparison of the measured fragmentation spectra with the RAPGAP predictions are shown in figure 6.2.

Similarly as in the case of Peterson parametrization, the parameter values obtained for both methods do not agree within their errors: $\alpha = 5.1^{+0.8}_{-0.7}$ for the hemisphere and $\alpha = 3.8^{+0.6}_{-0.5}$ for the jet method. This discrepancy between the two methods, at a level of $\lesssim 2\sigma$, could be due to a statistical fluctuation, however, it could also be a hint that the RAPGAP/PYTHIA model or the simple one parameter fragmentation functions by Peterson and Kartvelishvili are not able to describe the underlying physics sufficiently well for the application of the two methods. The reasons for these not fully consistent results are investigated in chapter 6.4.

Both parametrizations seem to describe the data equally good, as can be seen from the value of χ^2/N_{df} which is close to one (see the summary of the results in table 6.1).

Parametrization		**Hem. method**		**Jet method**	
		par. value	$\chi^2_{\mathrm{min}}/N_{\mathrm{df}}$	par. value	$\chi^2_{\mathrm{min}}/N_{\mathrm{df}}$
Peterson	ε	$0.022^{+0.007}_{-0.004}$	5.3/5	$0.040^{+0.013}_{-0.009}$	3.8/4
Kartvelishvili	α	$5.1^{+0.8}_{-0.7}$	4.2/5	$3.8^{+0.6}_{-0.5}$	4.4/4

Table 6.1: Extracted fragmentation parameters for the RAPGAP/PYTHIA MC.

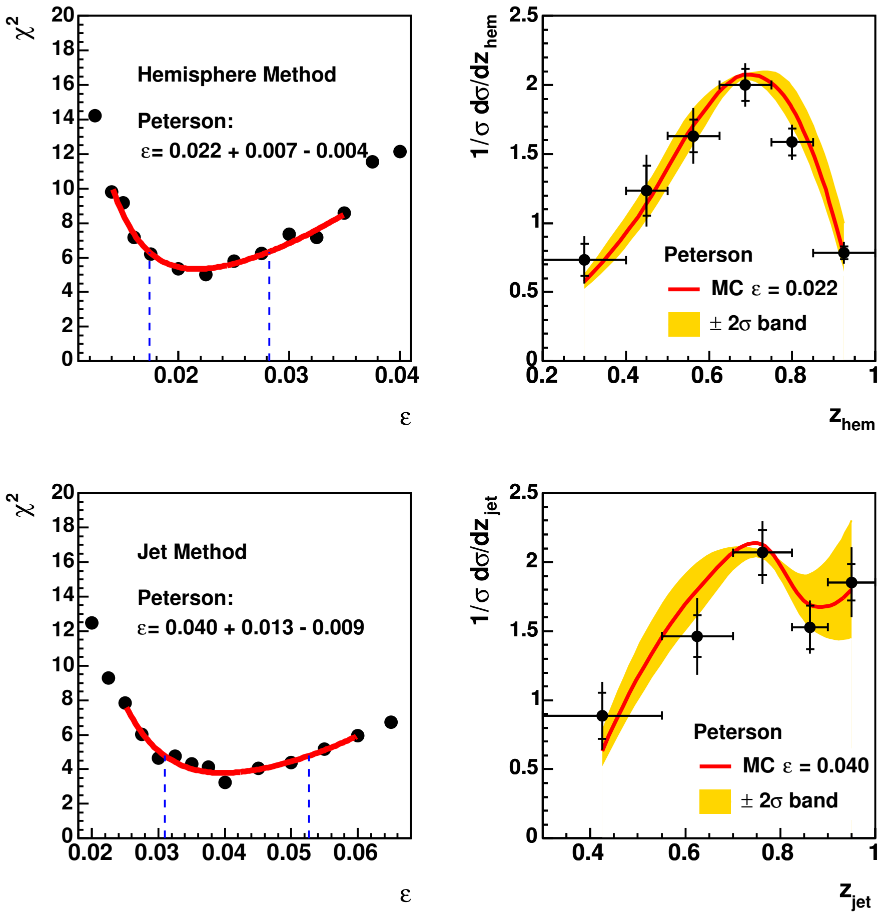

Figure 6.1: The χ^2 dependence and the measured z-distributions for the hemisphere (upper) and jet method (lower) compared to the RAPGAP/PYTHIA model predictions using the Peterson parametrization for the non-perturbative fragmentation function.

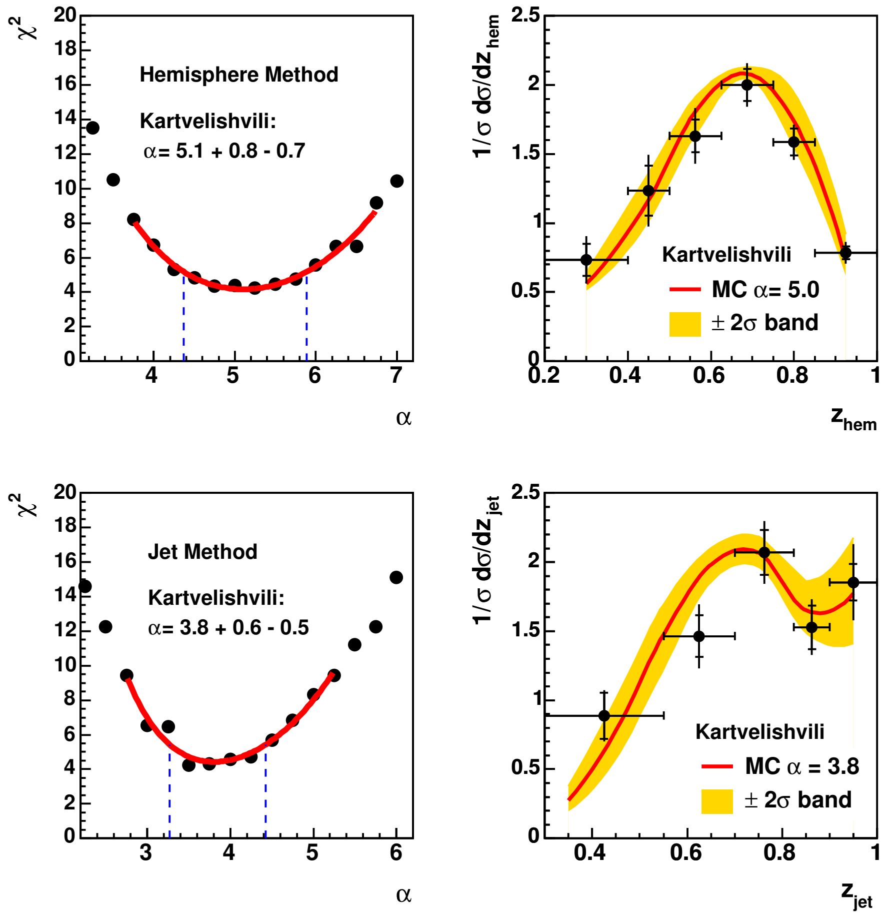

Figure 6.2: The χ^2 dependence and the measured z-distributions for the hemisphere (upper) and jet method (lower) compared to the RAPGAP/PYTHIA model predictions using the Kartvelishvili parametrization for the non-perturbative fragmentation function.

6.3 Parameter Extraction for CASCADE

The χ^2 fits of fragmentation parameters for the CASCADE/PYTHIA Monte Carlo model can be found in figure 6.3 for the Peterson and in figure 6.4 for the Kartvelishvili parametrization respectively. The extracted parameter values are summarized in table 6.2.

The CASCADE model with its CCFM parton evolution prefers, according to the χ^2 fit, the same values of fragmentation parameters like the RAPGAP model. Here again a discrepancy of $\lesssim 2\sigma$ between the extracted values for the hemisphere and jet method is observed.

Parametrization		**Hem. method**		**Jet method**	
		par. value	$\chi^2_{\mathrm{min}}/N_{\mathrm{df}}$	par. value	$\chi^2_{\mathrm{min}}/N_{\mathrm{df}}$
Peterson	ε	$0.022^{+0.006}_{-0.004}$	2.7/5	$0.041^{+0.014}_{-0.010}$	2.8/4
Kartvelishvili	α	$5.3^{+0.7}_{-0.6}$	2.2/5	$4.0^{+0.6}_{-0.6}$	2.6/4

Table 6.2: Extracted fragmentation parameters for the CASCADE/PYTHIA MC.

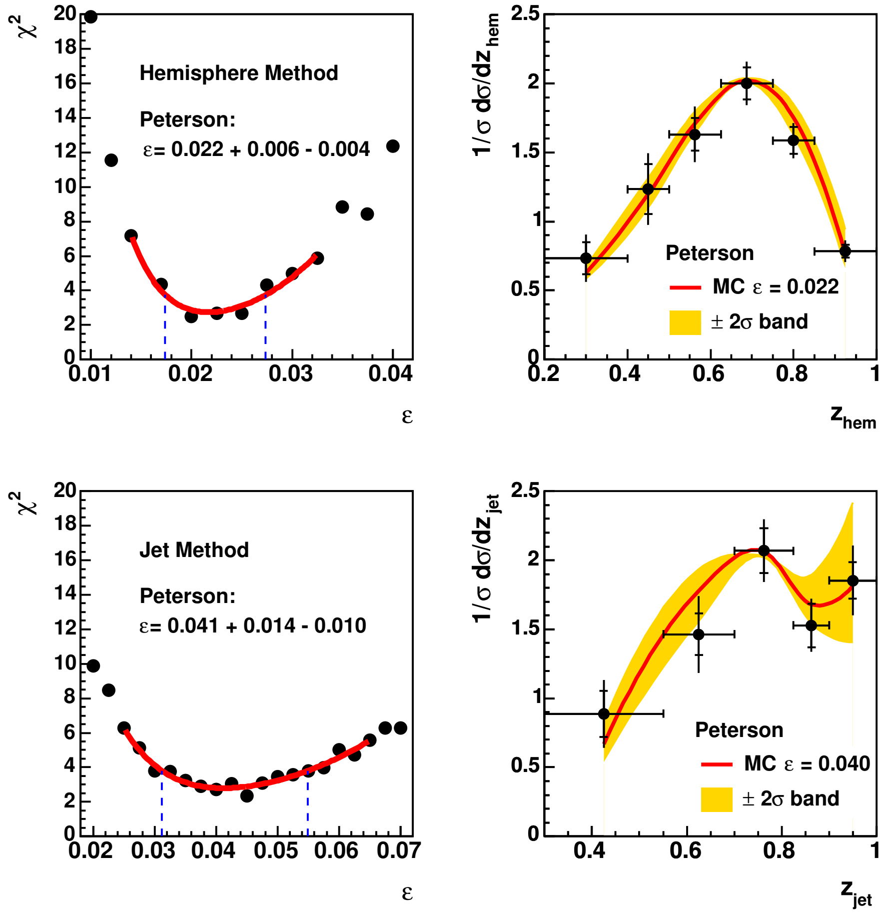

Figure 6.3: The χ^2 dependence and the measured z-distributions for the hemisphere (upper) and jet method (lower) compared to the CASCADE/PYTHIA model predictions using the Peterson parametrization for the non-perturbative fragmentation function.

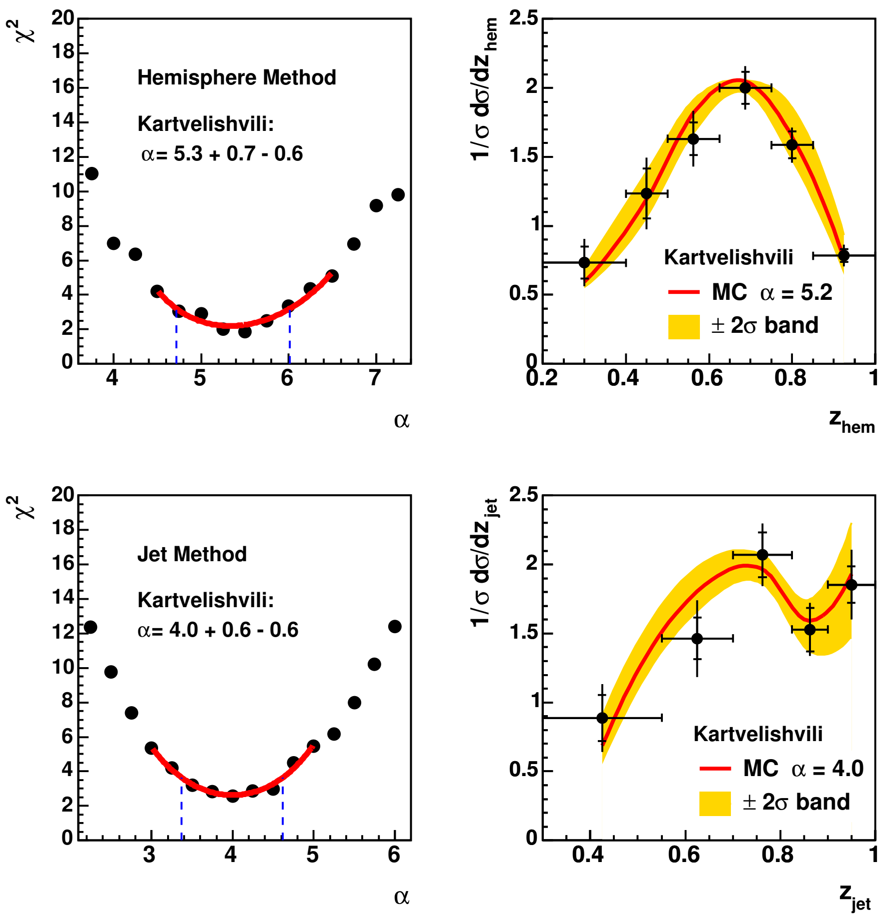

Figure 6.4: The χ^2 dependence and the measured z-distributions for the hemisphere (upper) and jet method (lower) compared to the CASCADE/PYTHIA model predictions using the Kartvelishvili parametrization for the non-perturbative fragmentation function.

6.4 Hemisphere Method Revisited

If the QCD based Monte Carlo models provide a correct description of all aspects of D^*-production[2], then the jet and hemisphere methods should both lead to the same extracted fragmentation parameter. In appendix B it is shown that if MC event sample ('pseudo-data') is used instead of real data, then the values extracted with both methods are compatible. This test confirms the self-consistency of the hemisphere and jet method. Hence the discrepancy of $\lesssim 2\sigma$ between the extracted parameters (see table 6.1 and 6.2) may signal inadequacies in one or more aspects of the model.

Since the hemisphere method does not explicitly demand any additional cut, besides the usual D^*-meson visibility requirements, it contains in contrast to the jet method a big fraction of events produced close to the kinematic threshold. On one hand this yields twice the statistics of the jet method, but on the other hand this makes the hemisphere method very sensitive to a good description of the charm production close to threshold by the MC models, which may not be the case.

In order to figure out whether this could be the reason for the observed discrepancy, the hemisphere method events have been divided into two subsamples: one subsample containing events with $E_{\mathrm{t}}(D^*\mathrm{jet}) > 3$ GeV and a second subsample containing the remaining soft QCD events. According to the above mentioned hypothesis the value of the fragmentation parameter obtained by the hemisphere method, applied to events containing a D^*-jet, should be in agreement with the parameter value obtained by the jet method.

The description of the measured $\mathrm{z_{hem}}$ distributions on detector level by the RAPGAP and CASCADE Monte Carlo models is shown in figure 6.5. For the sample without a D^*-jet a large excess of data over MC expectation is observed in the last bin, while for the sample of events containing a D^*-jet the excess of data is in the first bin. In general the $\mathrm{z_{hem}}$ spectrum for events with a D^*-jet is significantly softer than for the events without D^*-jet. This behavior is mainly caused by the minimal transverse momentum requirement $p_{\mathrm{t}}(D^*) > 1.5$ GeV. Most affected are the events produced close to the kinematic threshold, where basically only D^*-mesons carrying a large fraction of the c-quark's energy pass the transverse momentum requirement, as well as the events where due to the soft charm quark fragmentation the

[2]This includes sufficient flexibility of the fragmentation function over the range in z probed by the measurement.

D^*-meson does not fulfill the former cut. In principle the more energetic the parent charm quark is, the less the lower $z_{\rm hem}$ bins are affected by the $p_t(D^*)$ cut and the $z_{\rm hem}$ spectrum gets softer, which is in agreement with the observation.

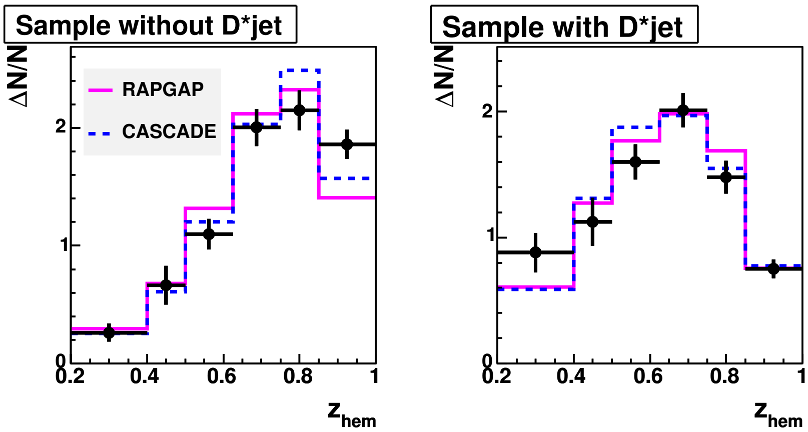

Figure 6.5: $z_{\rm hem}$ distributions on reconstructed level for subsample of events which do not contain D^*-jet (left) and which do contain a D^*-jet (right).

Description of the other hemisphere related distributions for two former subsamples of events by MC models can be seen in figure 6.6. Events with a D^*-jet have a harder $E_t(D^*\text{hem})$ spectrum, are less p_t balanced in the current hemisphere, and according to the thrust distribution they have a more pronounced back to back topology in a plane perpendicular to the γ direction.

After performing the χ^2 fits (see figure 6.7) the following fragmentation parameters have been extracted for the RAPGAP/PYTHIA MC:

$$\begin{array}{lll} \text{Sample without } D^*\text{-jet:} & \varepsilon = 0.014^{+0.006}_{-0.004}, & \chi^2_{\rm min}/N_{\rm df} = 1.0/5 \\ \text{Sample with } D^*\text{-jet:} & \varepsilon = 0.040^{+0.010}_{-0.007}, & \chi^2_{\rm min}/N_{\rm df} = 3.3/5 \end{array}$$

The extracted ε value using the sample of events containing a D^*-jet is in very good agreement with the value obtained from the jet method, whereas the ε value obtained from events which do not contain any D^*-jet is significantly harder ($\sim 2\sigma$ discrepancy).

This would indicate that either events produced close to the kinematic threshold or events where the c-quark radiates a hard gluon and thus does

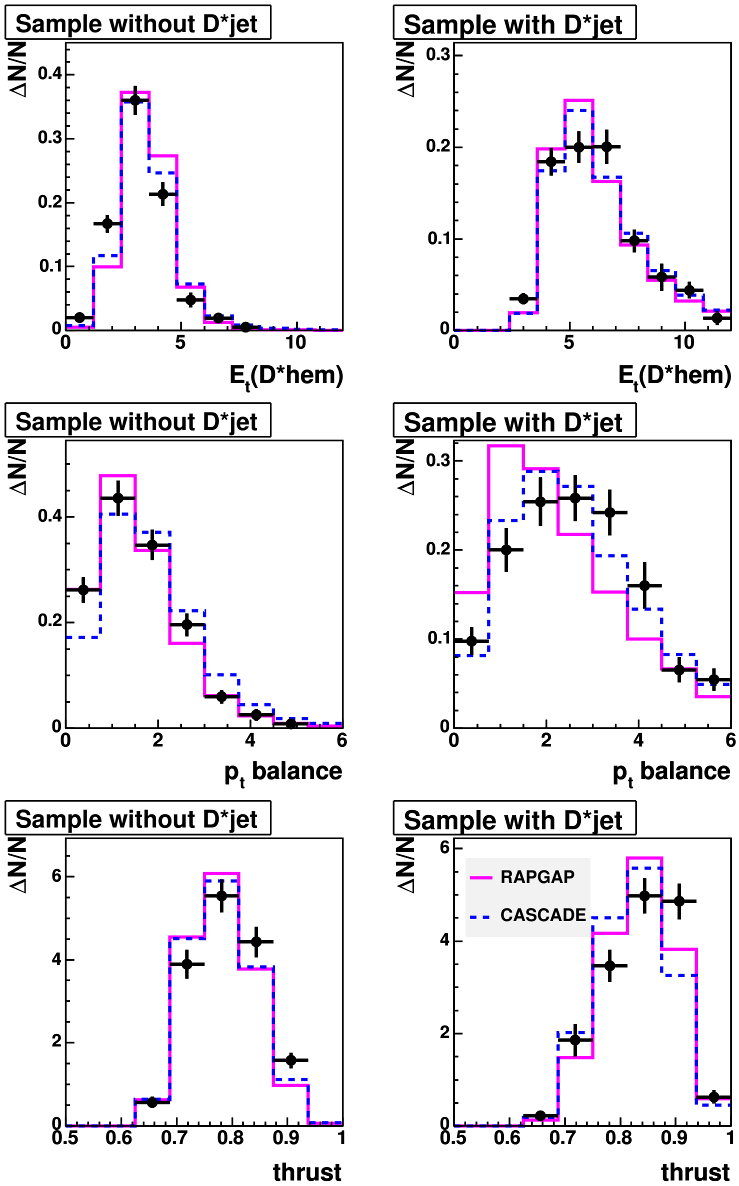

Figure 6.6: Description of hemisphere related observables in data by RAPGAP and CASCADE Monte Carlo models on reconstructed level, shown for subsample of events which do not contain D^*-jet (left column) and which do contain a D^*-jet (right column): $E_t(D^*\text{hem})$, the p_t balance in the current hemisphere and the thrust value.

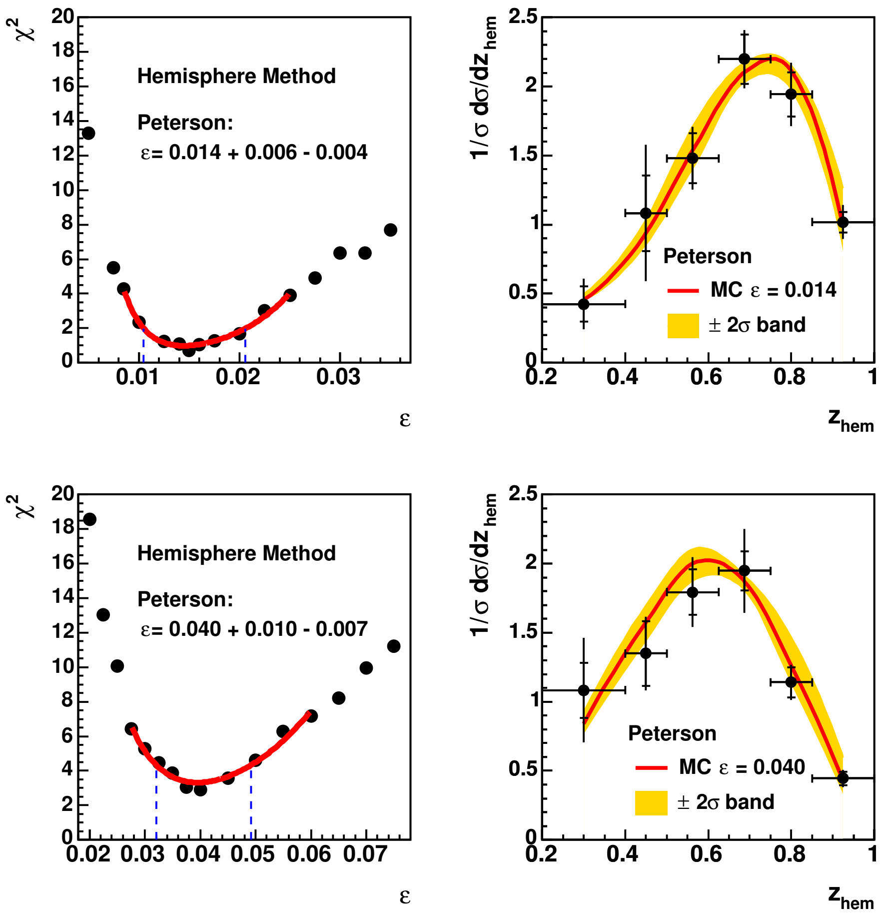

Figure 6.7: The χ^2 dependence and the measured z_{hem} distributions for two samples of events, one containing no D^*-jet (upper) and the other containing at least one D^*-jet (lower), compared to the RAPGAP/PYTHIA model predictions using the Peterson parametrization for the non-perturbative fragmentation function.

not fulfill the $E_\mathrm{t}(D^*\mathrm{jet})$ cut are not properly described by Monte Carlo simulation. This information is from the physics point of view important and exciting as well, since agreement with the jet method at larger E_t suggests that the hemisphere method works properly.

6.5 Parameter Extraction for NLO Calculations

NLO QCD programs, like for example, HVQDIS provide a theoretical prediction only on parton level in fixed $\mathcal{O}(\alpha_\mathrm{s}^2)$ without additional parton showers. Therefore the fragmentation functions obtained for MC models cannot be used, they need to be determined within the framework of the NLO calculations itself. The fragmentation of a charm quark into a D^*-meson is implemented as described in chapter 1.11.1. But for distributions like the momentum of the D^*-hemisphere or D^*-jet additional hadronization corrections need to be applied. There are two equivalent ways how this can be done. First, one can correct the NLO calculations for hadronization effects, or the second possibility is to correct the data to parton level.

Since the hadronization corrections can in principle depend on the hardness of the charm fragmentation function, it would mean that for each NLO calculation set with different fragmentation parameter they would have to be determined separately. Therefore it is more convenient to choose the second option where the hadronization corrections for measured data need to be determined only once.

The hadronization corrections were calculated using the RAPGAP Monte Carlo model , with incorporated Lund-string fragmentation model.[3] For both major $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ distributions bin based correction factors were determined as a ratio of entries before and after the hadronization in a given z bin: $C_i = \mathrm{z}_i^{\mathrm{parton}}/\mathrm{z}_\mathrm{i}^{\mathrm{hadron}}$.[4] In contrast to the calculations done on hadron level, where the D^*-meson is used to determine the hemisphere (jet) belonging to the parent charm quark, on parton level the hemisphere (jet) containing the D^* parent quark is directly chosen.

[3] For the purpose of this analysis the Monte Carlo simulation with a central fit value of the fragmentation parameter $\varepsilon = 0.04$, which has been extracted in the previous chapter, was used.

[4] This approach is valid only under assumption that hadronization corrections for LO+PS Monte Carlo model are the same as for NLO theory calculations.

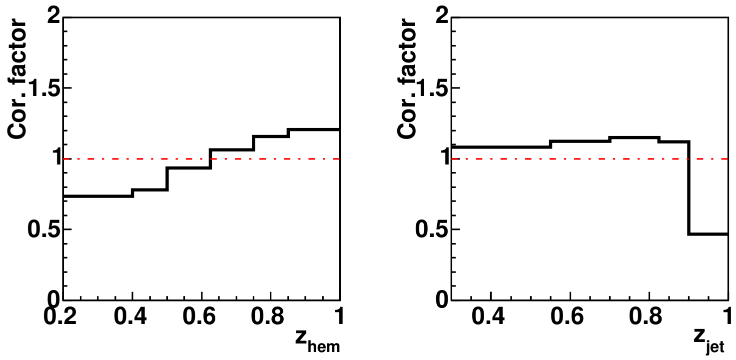

Figure 6.8: Hadronization correction factors for z_{hem} (left) and for z_{jet} (right) obtained by the RAPGAP/PYTHIA Monte Carlo model.

The hadronization corrections obtained are shown in figure 6.8. In case of the hemisphere method the hadronization corrections are below 25%, but depend on z_{hem}. For the jet method the hadronization corrections are flatter and less than 15% except for the last z_{jet} bin, where they reach 50%. These large hadronization corrections in the last bin are due to the jets consisting only of a D^*-meson itself ($z_{\mathrm{jet}}^{\mathrm{hadron}} = 1$), which has no analogy at parton level.

The data corrected to parton level are then compared with HVQDIS predictions for different fragmentation parameter values and a χ^2 is calculated. As has already been mentioned in section 1.11.1 two variants of fragmentation procedure have been investigated. In the first variant an additional p_{t} momentum component is attributed to the D^*-meson, besides a conventional longitudinal momentum parallel to the charm quark direction, whereas in the second variant this option is disabled, i.e. the D^*-meson is moving parallel to the c-quark.

The fitted values with and without additional p_{t} generation for the hemisphere and the jet observables are summarized in table 6.3. The shape of both fragmentation spectra turns out not to be sensitive to the additional transverse momentum of the D^*-meson.

An additional systematic error due to uncertainty of the hadronization corrections has been considered for the extracted fragmentation parameters listed in table 6.3. Since both fragmentation spectra on MC parton level have entries at values $z > 1$ (see figures 2.4 and 2.6), which has no anal-

Parametrization		**Hem. method**		**Jet method**	
		par. value	$\chi^2_{\min}/N_{\rm df}$	par. value	$\chi^2_{\min}/N_{\rm df}$
Peterson ($p_{\rm t}$)	ε	$0.040^{+0.010}_{-0.007}$	45.4/5	$0.057^{+0.014}_{-0.009}$	20.2/4
Peterson	ε	$0.042^{+0.012}_{-0.008}$	47.6/5	$0.058^{+0.014}_{-0.009}$	20.6/4
Kartvelishvili ($p_{\rm t}$)	α	$3.6^{+0.4}_{-0.4}$	9.5/5	$3.1^{+0.4}_{-0.4}$	6.1/4
Kartvelishvili	α	$3.5^{+0.4}_{-0.4}$	11.1/5	$3.1^{+0.4}_{-0.4}$	6.4/4

Table 6.3: Extracted fragmentation parameters for HVQDIS NLO-calculations. The quoted errors include the systematic error due to the different treatment of hadronization corrections.

ogy in HVQDIS where the charm quarks are fragmented independently[5], the data have been corrected with alternative bin by bin corrections, where all entries above 1 were put to the last $z_{\rm hem}$ or $z_{\rm jet}$ bin. The difference between the extracted parameter values with the first and second set of hadronization corrections gives the size of systematic uncertainty, which is then added quadratically to the original error obtained by the χ^2 fit. This uncertainty turned out to be smaller than 1%, only for Peterson fit to the $z_{\rm jet}$ distribution does it reach 5%.

In order to check whether the last $z_{\rm jet}$ bin, with the largest hadronization corrections, does not push the fit result away from the parameter value preferred by the rest of the $z_{\rm jet}$ spectrum, the fit has been repeated only with the first 4 points. The extracted parameter values are softer than before, $\varepsilon = 0.072^{+0.012}_{-0.011}$ for Peterson and $\alpha = 2.8^{+0.4}_{-0.4}$ for Kartvelishvili parametrization, but still compatible within the quoted errors. So apparently the last point does not seem to spoil the fit result.

The extracted parameter values for Peterson (figure 6.9) and Kartvelishvili parametrization (figure 6.10) are softer than in the MC simulations. This is in agreement with expectation, since the charm quark fragmentation as implemented in HVQDIS does not take into account higher excited charm resonant states, as is the case for MC simulation. The effect of the inclusion (exclusion) of higher charm resonant states on the value of the extracted fragmentation parameter has been investigated with the Monte Carlo simulation

[5]Some entries at $z_{\rm hem} > 1$ can occur even in HVQDIS, but only if additional $p_{\rm t}(D^*)$ is generated in the fragmentation process. This effect is only on a permille level, which is one order of magnitude smaller than in RAPGAP MC, and hence negligible.

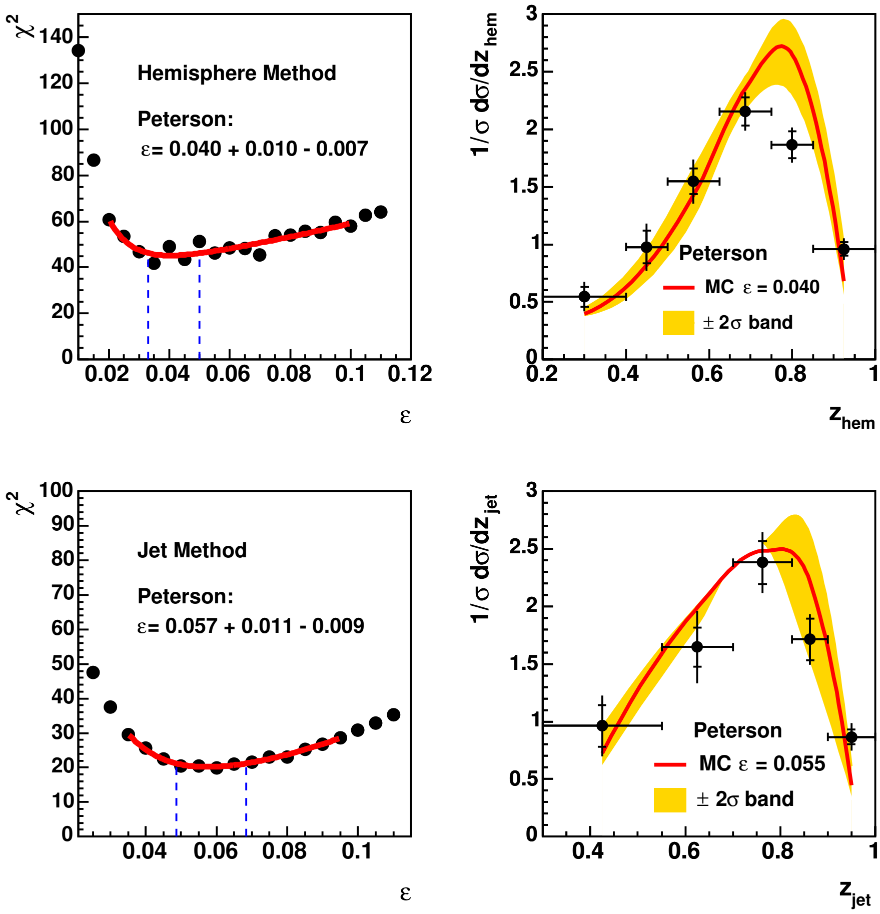

Figure 6.9: The χ^2 dependence and the measured z-distributions corrected on parton level for the hemisphere (upper) and jet method (lower) compared with HVQDIS prediction using the Peterson parametrization for the non-perturbative fragmentation function and additional $p_t(D^*)$ with respect to the c-quark direction.

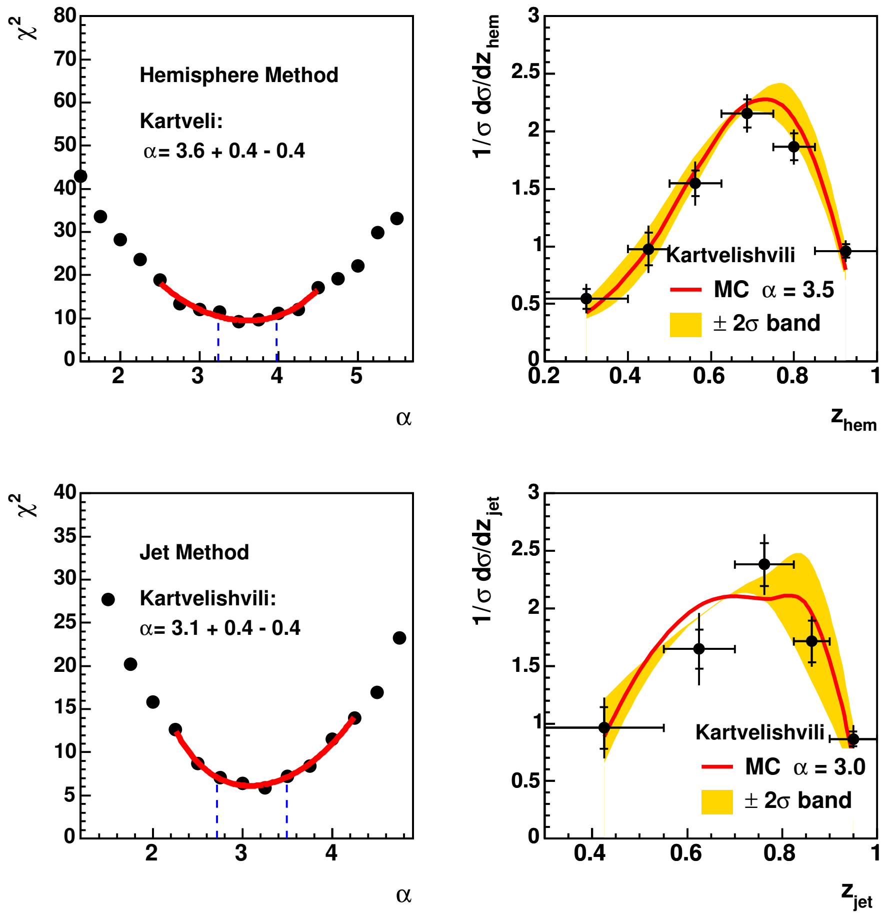

Figure 6.10: The χ^2 dependence and the measured z-distributions corrected on parton level for the hemisphere (upper) and jet method (lower) compared with HVQDIS prediction using the Kartvelishvili parametrization for the non-perturbative fragmentation function and additional $p_t(D^*)$ with respect to the c-quark direction.

(see appendix C). A general feature of the D^*-mesons produced directly in the fragmentation process is that they carry on average larger a fraction of the charm quark momentum than those D^*-mesons which are produced in decays of higher resonances. Hence ignorance of higher resonances results in a softer extracted non-perturbative fragmentation function.

When studying fragmentation at NLO level new features appear. The Peterson parametrization is strongly disfavored by the χ^2 of the fit. Moreover, the discrepancy between the extracted parameter values by $\mathrm{z_{hem}}$ and $\mathrm{z_{jet}}$ methods, which was in LO+PS Monte Carlo models around $2-3\sigma$, decreases to $1-2\sigma$.

6.6 Comparison with e^+e^- Experiments

In order to draw conclusions about the universality of the charm fragmentation function, the data and fitted fragmentation parameters have been compared with measurements done at e^+e^- collider experiments.

Since the $\mathrm{z_{hem}}$ observable has been designed in a way that includes gluon radiation from the charm quark, it is comparable to the observable definitions used at e^+e^- annihilation experiments (section 2.1 and 2.3). Therefore the measured $\mathrm{z_{hem}}$ distribution is compared with CLEO [11], OPAL [6] and ALEPH [12] measurements (see figure 6.11).

One can see that despite the different definitions of z observables, the shape of the measured spectra is quite similar.

Since $c\bar{c}$-pairs at the CLEO experiment are produced at similar center of mass energy ($\sqrt{s} = 10.6$ GeV) as the events covered by kinematic range of $\mathrm{z_{hem}}$ analysis, the comparison with this measurement is considered to be the most relevant. At the OPAL and ALEPH experiments charm is produced in the Z^0 decays ($\sqrt{s} = 91.2$ GeV), so in principle a larger phase space for gluon radiation is available which can have an impact on the shape of the z spectrum, since the z observable is sensitive to perturbative effects.[6]

Except for a small shift towards larger z values, the shape of the z spectrum is in rough agreement with the data from CLEO. This shift could be due to the visible range cuts applied on η and p_{t} of the D^*-meson. In general the increase of $p_{\mathrm{t}}(D^*)$ cut implies harder z spectrum.

[6] Indeed excess at low z values from OPAL arises from gluon splitting into $c\bar{c}$. In the ALEPH spectrum the gluon splitting component is already subtracted.

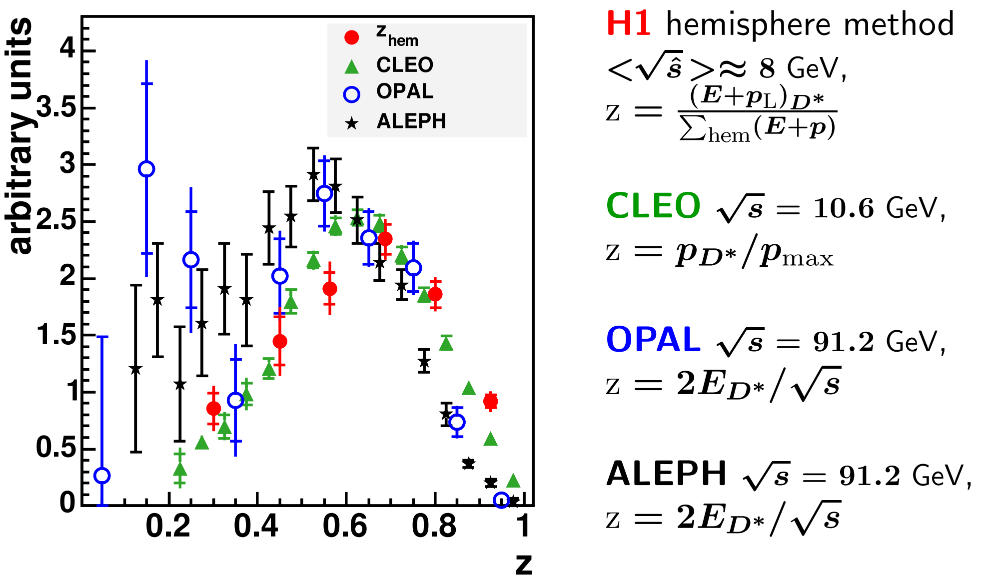

Figure 6.11: Comparison of the fragmentation function measurement from CLEO, OPAL and ALEPH with the hemisphere method measurement.

Whereas in e^+e^- measurements the spectra are extrapolated to the full $\eta(D^*)$ and $p_t(D^*)$ range, in ep collisions extrapolation would introduce too large uncertainty (due to much harder cuts which have to be applied in contrast to e^+e^- collisions) and hence is not performed. On the other hand even after extrapolation both spectra would be most likely not equivalent, since without $p_t(D^*)$ cut the mean $\hat{s}$ value of $c\bar{c}$ pair drops, and thus measurements would not cover the same kinematic region.

A better agreement between the CLEO data and z_{hem} distribution is observed when a sample of events with D^*-jet is considered (see figure 6.12). For this subsample of events the mean center-of-mass energy of the $c\bar{c}$-system is higher, $\sqrt{\hat{s}} \approx 10$ GeV, and thus closer to the $\sqrt{s}$ of CLEO experiment.[7] Furthermore, the size of corrections which have to be applied in order to extrapolate the spectrum to the full $\eta(D^*)$ and $p_t(D^*)$ range, is smaller than for the complete sample of D^*-mesons. For the extrapolation the CASCADE Monte Carlo model has been used. As can be seen from figure 6.12, agreement with CLEO data is anyway not achieved since the z_{hem} spectrum gets after

[7] Ideal for comparison would be to restrict the data sample only to events from an immediate vicinity of $\sqrt{\hat{s}} = 10.6$ GeV, but due to the lack of statistics this is not possible.

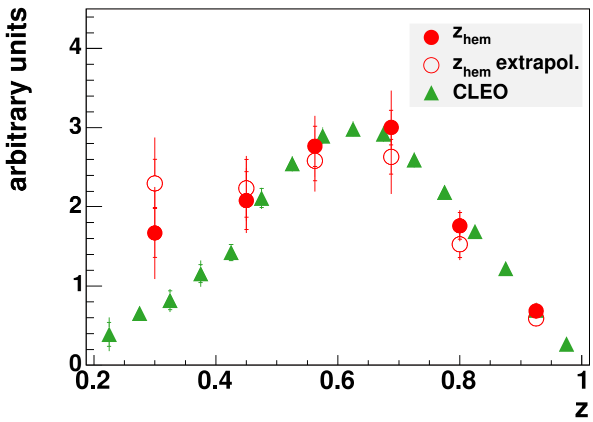

Figure 6.12: Comparison of the fragmentation function measurement from CLEO with the hemisphere method measurement obtained for the subset of events with a D^*-jet. The empty circles correspond to an extrapolation of the measured z_{hem} to the full $\eta(D^*)$, $p_{\mathrm{t}}(D^*)$ phase space, analogous to that of the CLEO analysis.

the extrapolation much softer. However, this does not necessarily mean that universality does not hold, since the definitions of observables as well as the basic hard subprocesses are different and the extrapolation procedure may have large uncertainties.

Nevertheless the same fragmentation parameter should describe the data at different center-of-mass energies, since the evolution of the perturbative component of the fragmentation function is in LO+PS MC simulations already accounted for (at least the dominant contribution). The extracted parameter values for the JETSET package, which is often used in Monte Carlo simulations for the treatment of hadronization effects, are listed in table 6.4. The extracted values do not match, even not between e^+e^- experiments themselves. However this should be not taken too seriously since the fragmentation parameter depends a lot on the choice of other MC steering parameters as well as on the inclusion of higher charm excited states (see appendix C).

At this level of understanding it is too early to draw any conclusions about the universality of fragmentation functions. A comparative analysis of

type	exp.	Parametrization		fit. value	$\chi^2_{\rm min}/N_{\rm df}$
e^+e^-	OPAL	Peterson	ε	$0.035 \pm 0.007 \pm 0.006$	5.2/6
		Kartvelishvili	α	$4.2 \pm 0.5 \;\; 0.4$	11.5/6
		Lund	a	$1.95^{+0.78}_{-0.53} \pm 0.08$	
			b	$1.58^{+0.64}_{-0.42} \pm 0.06$	3.4/5
e^+e^-	ALEPH	Peterson	ε	0.034	—
e^+e^-	CLEO	Lund	a	0.18	—
			b	0.40	—
e^+e^-	BELLE	Peterson	ε	0.054	3003/54
		Kartvelishvili	α	5.6	1271/54
		Lund	a	0.58	
			b	default	965/55
ep	ZEUS	Peterson	ε	$0.064 \pm 0.006^{+0.011}_{-0.008}$	—
ep	H1 $z_{\rm hem}$	Peterson	ε	$0.022^{+0.007}_{-0.004}$	5.3/5
		Kartvelishvili	α	$5.1^{+0.8}_{-0.7}$	4.2/5
ep	H1 $z_{\rm jet}$	Peterson	ε	$0.040^{+0.013}_{-0.009}$	3.8/4
		Kartvelishvili	α	$3.8^{+0.6}_{-0.5}$	4.4/4

Table 6.4: Extracted fragmentation parameters for JETSET, from e^+e^- (ALEPH [12], OPAL [6], CLEO [11] and BELLE [58]) and ep (ZEUS [48], this measurement) data.

the e^+e^- and ep results, which is beyond the scope of this thesis, is essential. As the definitions of z observables used in e^+e^- and ep experiments are different, the most straight forward way would be to repeat an e^+e^- analysis with a Monte Carlo simulation tuned to our data and then to compare the obtained fragmentation spectrum with the e^+e^- data if it is able to describe them. However this would require a detailed knowledge of the particular e^+e^- analysis.

Conclusions and Outlook

The fragmentation function of charm quarks fragmenting into a D^*-meson has been studied in ep collisions using two different methods: the hemisphere and the jet method. Observables of both methods are reasonably well described by the Lund string fragmentation model, and both methods show a strong sensitivity to the studied non-perturbative QCD aspects.

Using the method of least squares, the fragmentation parameters have been extracted for the most frequently used Peterson and Kartvelishvili parametrizations of non-perturbative fragmentation functions, taking statistical and systematic uncorrelated and correlated errors into account. The parameters were extracted for the QCD inspired RAPGAP and CASCADE Monte Carlo models, which include leading-order massive matrix elements for charm production matched with collinear (RAPGAP) or k_{t}-ordered (CASCADE) parton showers, and for an analytical next-to-leading-order QCD calculation (HVQDIS).

The results obtained with the hemisphere and jet method for the Peterson and Kartvelishvili parametrizations, which are summarized in tables 6.1, 6.2 and 6.3, show discrepancies of approximately 2σ. This may signal inadequacies in one or more aspects of the Monte Carlo models, but the possibility of a statistical fluctuation cannot be excluded either.

Data corrected for detector acceptance have been compared to e^+e^- annihilation measurements. Rough agreement with recent data from the CLEO collaboration, collected at similar center-of-mass energies of the $c\bar{c}$ system, is observed. Nevertheless, before making a final statement about the universality of the fragmentation function, a consistent phenomenological analysis of e^+e^- and ep data must be performed, as different definitions of observables are used.

The results of this analysis provide an important test of our understanding of heavy quark production. They can help to reduce the uncertainty due to c-quark fragmentation in other D^* based charm analyses at ep experiments. The extracted non-perturbative parameters should be used only with corresponding perturbative QCD description and generator tuned quantities.

Inclusion of data from the HERA II running period would allow to reduce the statistical and perhaps also the systematic uncertainty, and thus lead to an increase in the precision of this type of measurement.

Appendix A

Reweighting Procedure

For the data unfolding to hadron level it is important to have a Monte Carlo simulation which describes the data distributions reasonably well. In case the deficiencies are not too large, a satisfactory description can be achieved by applying a reweighting procedure to the MC simulation.

By means of the of measured distributions in data and in Monte Carlo simulation an event weight is determined. First distributions of a given observable x in data and MC simulation are fitted with a function $\mathcal{F}(x)$ of an appropriate shape, and then the event weight is calculated as the ratio of $\mathcal{F}^{\mathrm{data}}(x)/\mathcal{F}^{\mathrm{MC}}(x)$ according to the generated x value in the MC event.

In case of the RAPGAP and CASCADE Monte Carlo simulations reweighting in one observable was not sufficient. Altogether three observables were reweighted: z_{vtx}, $p_{\mathrm{t}}(D^*)$ and $\eta(D^*)$. Whereas the first observable is related to the detector simulation (section 4.2.2), the remaining two are directly sensitive to the underlying physics implemented in the MC model. Under the assumption that these quantities are uncorrelated, the total event weight is given by the product of single weights

$$W_{\mathrm{tot}}(z_{\mathrm{vtx}}, p_{\mathrm{t}}(D^*), \eta(D^*)) = W(z_{\mathrm{vtx}})W(p_{\mathrm{t}}(D^*))W(\eta(D^*)) \qquad \text{(A.1)}$$

This assumption seems to be approximately correct, since already after the first iteration a quite reasonable data description is achieved. For the z_{vtx} and $\eta(D^*)$[1] distributions a Gauss function and for $p_{\mathrm{t}}(D^*)$ a steeply falling

[1]This variable is reweighted in the γp-frame, where the distribution has a simple Gaussian shape.

function have been fitted:

$$\begin{aligned} z_{\rm vtx}: \quad & \mathcal{F}(x) = \mathcal{G}(x, N, \sigma, \mu) \\ p_{\rm t}(D^*): \quad & \mathcal{F}(x) = N x^{\alpha} exp(-\beta x) \\ \eta(D^*): \quad & \mathcal{F}(x) = \mathcal{G}(x, N, \sigma, \mu) \end{aligned} \tag{A.2}$$

In figure A.1 the $p_{\rm t}(D^*)$ and $\eta(D^*)$ distributions are shown before and after reweighting. Besides an improvement of the data description by MC in the reweighted distributions, which is of course expected, also distributions related to the reweighted ones, as e.g. $E_{\rm t}(D^*\text{jet})$, $\eta(D^*\text{jet})$ and to some extent also $E_{\rm t}(D^*\text{hem})$, $\eta(D^*\text{hem})$ are positively affected. A spoiling of the description of other distributions has not been observed.

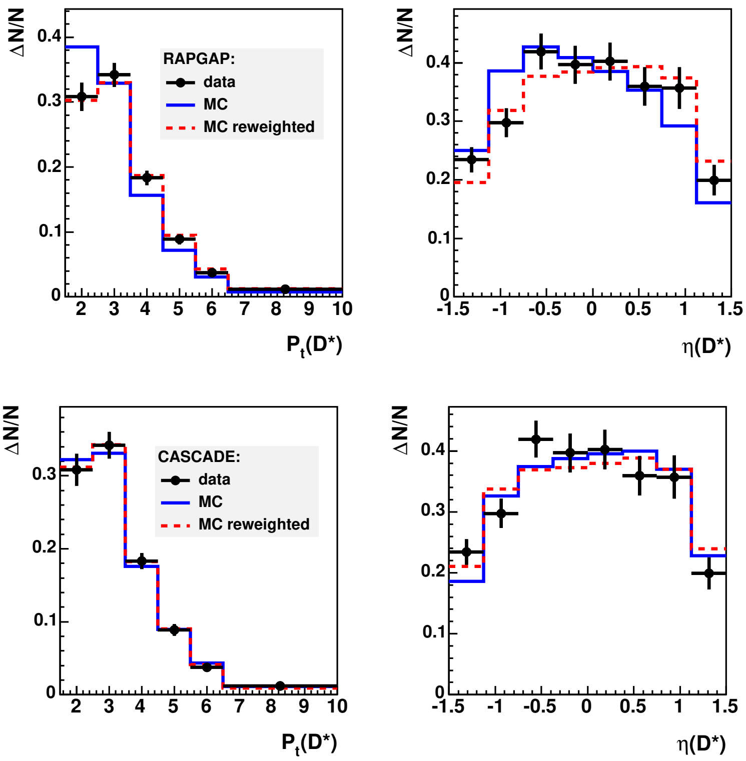

Figure A.1: The description of $p_{\rm t}(D^*)$ and $\eta(D^*)$ distributions before and after Monte Carlo reweighting. Upper plots RAPGAP, lower plots CASCADE Monte Carlo simulation.

Appendix B

Consistency Check of Experimental Methods

To prove the internal consistency of hemisphere and jet method within a Monte Carlo simulation, the following test has been performed.

The value of fragmentation parameter was extracted by minimizing the χ^2 as described in section 6.1. But this time instead of the data the CASCADE Monte Carlo with Kartvelishvili parametrization $\alpha = 5.0$ was used with statistics equivalent to the data statistics (this MC sample will be further referred as pseudo-data). For simplicity only statistical errors were taken into account during the χ^2 calculation.

The χ^2 dependence on fragmentation parameter obtained for the CASCADE MC model is shown on figure B.1. Both methods lead to the same α value which has been chosen also for pseudo-data. Hence a conclusion can be drawn that both experimental methods are consistent within the Monte Carlo simulation.

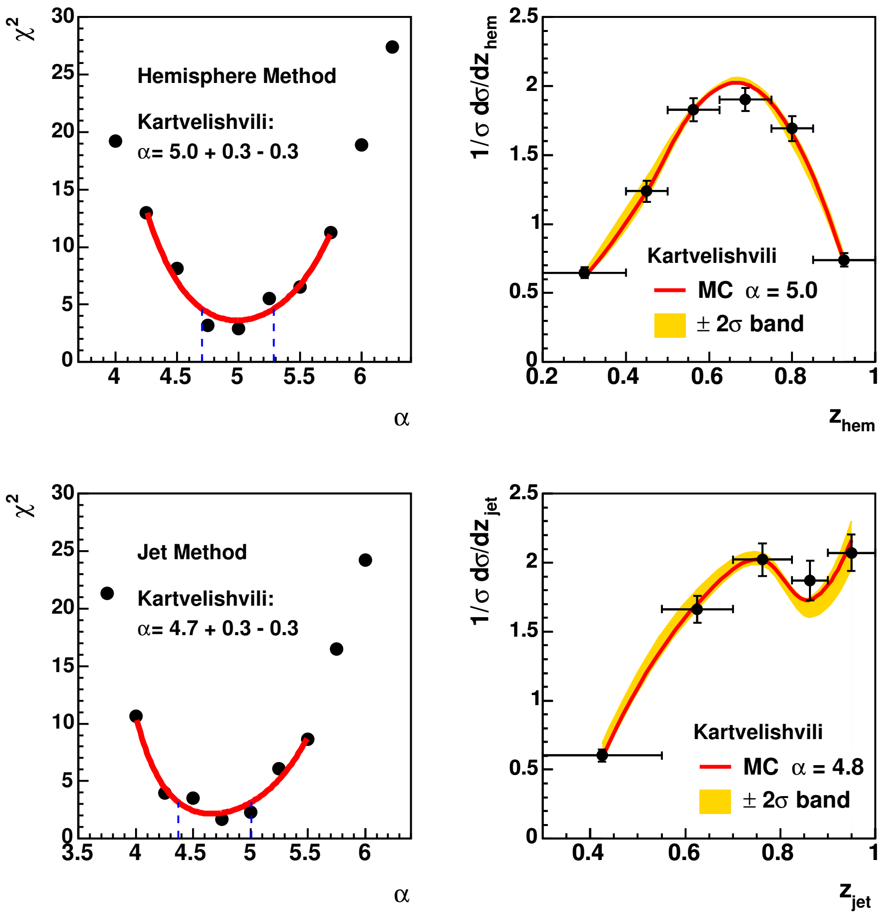

Figure B.1: The χ^2 dependence on fragmentation parameter α and the corresponding z-distributions for the hemisphere (upper) and jet method (lower) compared to the CASCADE/PYTHIA model prediction, using pseudo-data instead of the real data.

Appendix C

Effect of Higher Charmed Resonances

In many older charmed analyses done at HERA the effect of higher charmed resonances used to be underestimated. Either too few higher resonances were assumed, or they have been completely neglected, i.e D^*-meson was assumed to come directly from charm quark fragmentation. However the recent measurements, as e.g [49][50][51], show that the fraction of D^*-mesons originating from decays of higher charmed resonances is quite sizeable and hence should not be ignored.

In this work for steering the production of excited charmed states and other hadronization parameters a set tuned by ALEPH collaboration [55] to their data has been used. Overall about 27% of the D^*-mesons in the kinematic range of this analysis come from decays of excited states. The list of included excited states as well as the fraction of D^*-mesons resulting from their decays is shown in figure C.1 (left). Some of them are seen also in H1 data [23] (see figure C.1 (right)) and the fitted fraction of the D^*-mesons coming from $D_1(2420)^0$ and $D_2^*(2460)^0$ resonance decays of around 8% is in agreement with the Monte Carlo expectation.

Taking into account higher excited states the non-perturbative fragmentation function $FF(c \to D^*)$ can be written as a superposition of a direct fragmentation of the c-quark into D^*-meson $FF_p(c \to D^*)$, and of a fragmentation of the c-quark into excited charm meson (let us call it D^{**}-meson) which then subsequently decays to D^*-meson:

$$FF(c \to D^*) = FF_{\mathrm{p}}(c \to D^*) + \sum_i FF_{\mathrm{p}}(c \to D_i^{**}) \otimes FF(D_i^{**} \to D^*) \quad \text{(C.1)}$$

where the sum runs over all excited mesons decaying to D^*-meson. Usually one assumes that all directly produced charm mesons are fragmented according to the same fragmentation function FF_{p}, independently from the meson type.

The latter subset of D^*-mesons have significantly softer fragmentation spectrum. The fragmentation spectra $FF(c \to D^*)$ and $FF_{\mathrm{p}}(c \to D^*)$ are compared in figure C.2 in terms of $\mathrm{z}_{\mathrm{hem}}$ and $\mathrm{z}_{\mathrm{jet}}$ observables.

The inclusion of higher excited states therefore leads to the softer fragmentation spectra. Due to this feature the extracted non-perturbative fragmentation function is significantly harder for the Monte Carlo models with incorporated higher resonances (see figure C.3). The extracted ε parameter value moves from 0.022 to 0.038 for the hemisphere method and from 0.040 to 0.75 for the jet method.

This could possibly explain the difference between the Peterson parameter value of $\varepsilon = 0.061 \pm 0.0073(stat)^{+0.012}_{-0.008}(syst.)$ extracted by the ZEUS collaboration [48] with the jet method at higher jet energies $E_{\mathrm{t}}(D^*\mathrm{jet}) > 9$ GeV, and the parameter value extracted within framework of this analysis by the jet method $\varepsilon = 0.040 + 0.013 - 0.009$.

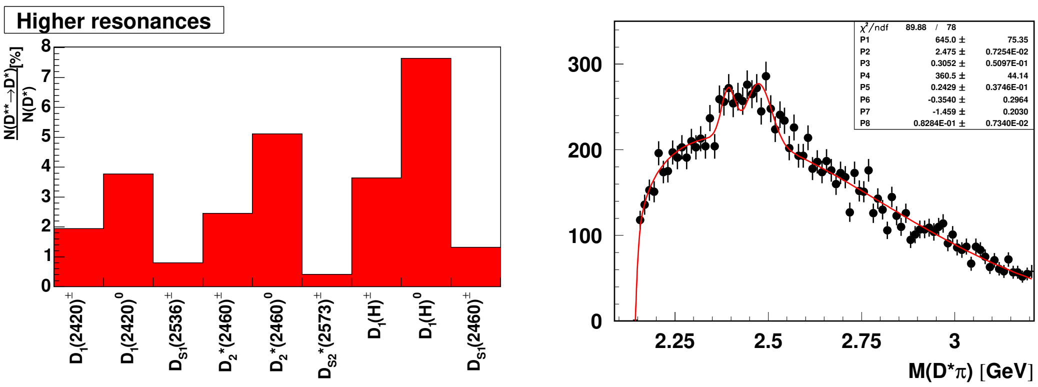

Figure C.1: Left: the list of excited charmed states together with the fraction of D^*-mesons resulting from their decays as incorporated in the Monte Carlo simulation. Right: the fitted invariant mass spectrum of $D^*\pi$ combinations with visible D_1^0 and D_2^{*0} signals in the H1 data.

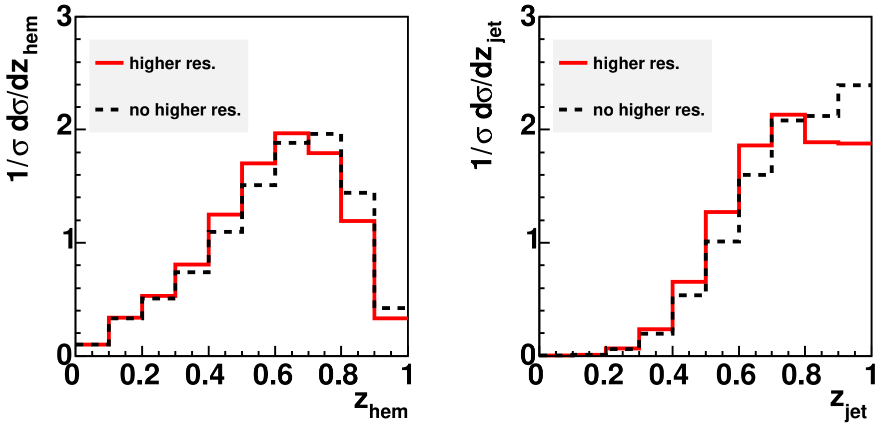

Figure C.2: The effect of higher charmed resonances on hardness of the D^* fragmentation spectra. Left for z_{hem} and right for z_{jet} observable on hadron level. The red line corresponds to $FF(c \to D^*)$, with the folded effects of higher resonances, and the black line to $FF_{\text{p}}(c \to D^*)$.

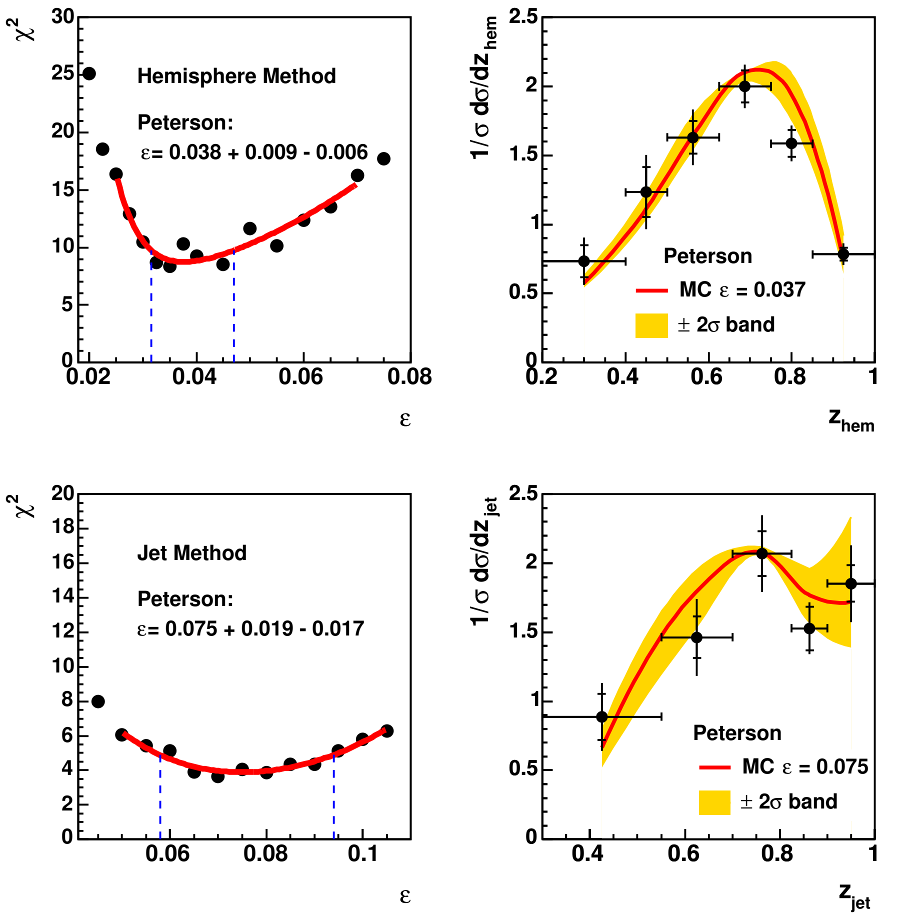

Figure C.3: The χ^2 dependence and the measured z-distributions for the hemisphere (upper) and jet method (lower) compared to the RAPGAP/PYTHIA model prediction without incorporated higher resonances using the Peterson parametrization for the non-perturbative fragmentation function.

Bibliography

[1] I. Abt *et al.* [H1 Collaboration], Nucl. Instrum. Meth. A **386**, (1997) 310.

[2] I. Abt *et al.* [H1 Collaboration], Nucl. Instrum. Meth. A **386**, (1997) 348.

[3] I.Abt *et al.* [H1 Collaboration], *The H1 detector at HERA.* , Nucl. Instrum. Meth. A **386** (1997) 310.

[4] C. Adloff *et al.* [H1 Collaboration], Z. Phys. J. C **72**, (1996) 593 [arXiv:hep-ex/9607012].

[5] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **13**, (2000) 609.

[6] R. Akers *et al.* [OPAL Collaboration], Z. Phys. C **67** (1995) 27.

[7] A. Aktas *et al.* [H1 Collaboration], Eur. Phys. J. C **40**, (2005) 349 [hep-ex/0411046].

[8] B. Andersson *et al.* [H1 Collaboration], Phys. Rep. **97**, (1983) 31;
B. Andersson, G. Gustafson, and B. Soderberg, Z. Phys. J. C **20**, (1983) 317.

[9] B. Andrieu *et al.* [H1 Calorimer Group], Nucl. Instrum.Meth. A **336**, (1993) 460.

[10] A. Arbuzov *et al.*, Comput. Phys. Commun. **94** (1996) 128, [hep-ph/9511434].

[11] M. Artuso *et al.* [CLEO Collaboration], Phys. Rev. D **70** (2004) 112001 [arXiv:hep-ex/0402040].

[12] R. Barate *et al.* [ALEPH Collaboration], Eur. Phys. J. C **16** (2000) 597, [arXiv:hep-ex/9909032].

[13] J. D. Bjorken, Phys. Re. D 17, 171 (1978);
M. Suzuki, Phys. Lett. 71B, 189 (1977).

[14] Stan Bentvelsen *et al.*, *Reconstruction of (x,Q^2) and extraction of structure functions in neutral current scattering at HERA*, Proceedings of the workshop: Physics at HERA, volume 1 (2002),pages 23-41 DESY.

[15] S. Brandt, Ch. Peyrou, R. Sosnowski and A. Wroblewski Phys. Lett. A **12**, (1964) 57.

[16] J. Bürger *et al.*, Nucl. Instrum. Meth. A **279** (1989) 217.

[17] S. Burke *et al.*, Nucl. Instrum. Meth. A **373** (1996) 227.

[18] H1 Collaboration, *Technical proposal for the upgrade of the backward region of the H1 detector*, DESY internal report PRC-93/02.

[19] M. Ciafaloni, Nucl. Phys. B **296**, (1988) 49;
S. Catani, F. Fiorani and G. Marchesini, Phys. Lett. B **234**, (1990) 339;
S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. B **336**, (1990) 18;
G. Marchesini, Nucl. Phys. B **445**, (1995) 49 [hep-ph/9412327].

[20] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant K(t) clustering algorithms for hadron hadron* , Nucl. Phys. B **406**, 187 (1993).

[21] S. Chekanov *et al.* (ZEUS), Phys. Lett., B **565**, 87 (2003) [arXiv:hep-ph/0302025].

[22] J. Cuti and V. Weisskopf, Phys. Rev., D **4**, 3418 (1971).

[23] K. Daum [H1 collaboration], private communication.

[24] S. Eidelman *et al.*, (Particle Data Group), Phys. Lett. B **592**, (2004) 50.

[25] S. Eidelman *et al.*, (Particle Data Group), Phys. Lett. B **592**, (2004) 242.

[26] S. Eidelman *et al.*, (Particle Data Group), Phys. Lett. B **592**, (2004) 280.

[27] S. Egli *et al.*, Nucl. Instrum. Meth. A **283**, (1989) 487.

[28] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions* , Phys. Rev. D **48** (1993) 3160 [arXiv:hep-ph/9305266].

[29] G.J. Feldman *et al.*, Phys. Rev. Lett. **38** (1977) 1313.

[30] R. P. Feynman, Phys. Rev. Lett. **23** (1969) 1415.

[31] R.P. Feynman and R.D. Field, Nucl. Phys. B **136** (1978) 1.

[32] Heavy Flavour Group [H1 Collaboration], *HQSEL45* (1999), URL https://www-h1.desy.de/h1/iww/iwork/ihq/sw-doc/hqsel99.html.

[33] J. Gassner, *A Measurement of D-Meson Production by Decay Vertex Identification*, PhD Thesis, Zürich (2002), 175.

[34] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press (1996).
F. Halzen and A Martin, *Quarks and Leptons*, John Wiley (1984).

[35] V.N. Gribov and L.N. Lipatov, Yad. Fiz.**15** (1972) 781 [Sov.J. Nucl. Phys. **15** (1972) 438];
V.N. Gribov and L.N. Lipatov, Yad. Fiz.**15** (1972) 1218 [Sov.J. Nucl. Phys. **15** (1972) 675];
L.N. Lipatov, Yad. Fiz.**20** (1975) 181 [Sov.J. Nucl. Phys. **20** (1975) 94];
G. Altareli and G. Parisi, Nucl. Phys. B **126** (1977) 298;
Y. L. Dokshitzer, Sov Phys. JETP **46** (1977) 641 [Zh. Eksp. Teor. Fiz. **73** (1977) 1216].

[36] B.W. Harris and J. Smith, Phys. Rev. D **57** (1998) 2806.

[37] H. Jung, *The CCFM Monte Carlo Generator CASCADE Version 1.2007 (2004).*
URL http://www.desy.de/ jung/cascade/cascade12007.ps.gz.

[38] H. Jung, Comput. Phys. Commun., **86**, (1995) 147, version 3.1 [version 3.1].

[39] H. Jung, *Un-integrated Parton Density Functions in CCFM*, DIS 2004, Štrbské Pleso, Slovakia, hep-ph/0010017.

[40] V.G. Kartvelishvili, A.K. Likhoded, and V.A. Petrov, Phys. Lett. B 78, 615 (1978).

[41] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Sov. Phys. JETP **45**, (1977) 199;
Y.Y. Balitski and N.L. Lipatov, Sov. J. Nucl. Phys. **28**, (1978) 822.

[42] A. Kwiatkowski, H. Spiesberger and H.J. Mohring, Comput. Phys. Commun. **69**, (1992) 155, *HERACLES - An event generator for ep interactions at HERA including radiative processes*, version 4.6, 1996.

[43] Lee West, *Heavy flavor working group track selection code by Lee West*, Available from H1 Collaboration.

[44] H.L. Lai, J. Fuston, S. Kuhlmann, J. Morfin, F. Olness, J.F. Owens, J. Pumplin, W.K. Tung, Eur. Phys. J. C **12** (2000) 375.

[45] P. Marage *et al.*, Nucl. Phys. B **16** (1990) 518;
K. Müller *et al.*, Nucl. Instrum. Meth. A **312** (1992) 457.

[46] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L.Stanco, Comp. Phys. Comm. **67** (1992) 465.

[47] T. Nichols *et al.*, [H1 SPACAL Group], Nucl. Instrum. Meth. A **386** (1997) 397.

[48] S. Chekanov, [ZEUS Collaboration], *Measurement of charm fragmentation function in D^* photoproduction at HERA*, Abstract: 778, submitted to XXXIst Int. Conf. on High Energy Physics, 24-31 July 2002, Amsterdam, Netherlands.

[49] [ZEUS Collaboration], *Production of P-wave charm mesons at HERA*, Abstract: 854, submitted to XXXth Int. Conf. on High Energy Physics, July 27 - August 2, 2002, Osaka, Japan.

[50] [ALEPH Collaboration], *Production of D_1 and D_2^* mesons in hadronic Z decays*, Abstract: 5_411, submitted to HEP99, July 15-21, 2002, Tampere.

[51] [ALEPH Collaboration], Phys. Lett. B **526**, (2002) 34-49, [hep-ex/0112010].

[52] Matti Peez, Benjamin Portheault, Emmanuel Sauvan, *An energy flow algorithm for hadronic reconstruction in OO: Hadroo2*, H1-Note H1-01/05-616,DESY, 2005.

[53] C. Peterson *et al.*, Phys. Rev. D27 (1983) 105.

[54] D. Pitzl *et. al.*, ETHZ-IPP PR-2000-1.

[55] G. Rudolph [ALEPH collaboration], private communication.

[56] E. Sauvan, talk on H1OO meeting Jun (2005).

[57] E. Sauvan, private communication.

[58] R. Seuster *et al.* [BELLE Collaboration], Belle Preprint 2005-23, arXiv:hep-ex/0506068].

[59] G. Schuler, T. Sjöstrand, Phys. Lett. B **376** (1996) 193.

[60] S. Schmidt [H1 collaboration], private communication.

[61] S. Schmidt, *Messung charminduzierter Zweijetereignisse in tief inelastischer ep-Streuung mit dem H1-Detektor bei HERA*, Ph.D. thesis, Technische Universität München.

[62] T. Sjöstrand, L. Lönnblad, S. Mrenna and P. Skands, *PYTHIA 6.2 - Physics and Mannual*, [arXiv:hep-ph/0108264] (2002);
T. Sjöstrand, P. Eden, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun., **135**, (2001) 238, [arXiv:hep-ph/0010017] [version 6.1].

[63] J. Steinhart, *Die Messung des totalen $c\bar{c}$-Photoproductions-Wirkungsquerschnittes von λ_c-Baryonen unter Verwendung der verbesserten dE/dx-Teilchenidentifikation am H1-Experiment bei HERA*, Ph.D. thesis, Universität Hamburg (1999).