UNIVERSITÄT BONN
Physikalisches Institut

Charm Production in Charged Current Deep Inelastic e^+p Scattering at HERA

von
Meng Wang

The measurement of charm production in charged current deep inelastic positron-proton scattering is investigated with the ZEUS detector at the HERA collider. The data used has been collected from 1995 to 2000, corresponding to an integrated luminosity of $110\,\mathrm{pb}^{-1}$. Charged D^* mesons decaying in the channel $D^{*+} \to D^0\pi_s^+$ with $D^0 \to K^-\pi^+$ and the charge-conjugated channel are reconstructed to tag charm quarks. The visible cross section for D^*, $\sigma_{\mathrm{vis}}^{D^*} = 12.8 \pm 4.0(stat)^{+4.7}_{-1.5}(sys)$ pb, is measured in the kinematic range of $Q^2 > 200\,\mathrm{GeV}^2$ and $y < 0.9$, and of $p_T^{D^*} > 1.5\,\mathrm{GeV}$ and $|\eta^{D^*}| < 1.5$. The upper-limit for the charm production in the same DIS kinematic range is determined to be $\sigma^{e^+p\to\bar{\nu}_e cX} < 109\,\mathrm{pb}$ at 90% confidence level.

Post address:
Nussallee 12
53115 Bonn
Germany

BONN-IR-2006-02
Bonn University
March 2006
ISSN-0172-8741

UNIVERSITÄT BONN
Physikalisches Institut

Charm Production in Charged Current Deep Inelastic e^+p Scattering at HERA

von
Meng Wang（王 萌）
aus
Jiangsu, China

Dieser Arbeit wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen.

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online *elektronisch publiziert.*

Erscheinungsjahr: *2006*

Referent:	Prof. Dr. Erwin Hilger
Korreferent:	Prof. Dr. Norbert Wermes
Tag der Promotion:	13.02.2006

Contents

List of Figures

List of Tables

Chapter 1

Introduction

道生一	Dao gave birth to One
一生二	One to Two
二生三	Two to Three
三生万物	Three to others
万物负阴而抱阳	All objects
冲气以为和	Bear Yin
	Embrace Yang
	Harmonize with
	air of hollowness

—老子, **道德经**, 四十二章, 春秋
Laozi, ***Dao De Jing***, chapter 42, 6 B.C.
A contemporary semantic translation by E.T. Tan

What are the ultimate constituents of matter?
How do they interact with each other?

For thousands of years, humankind has pursued answers to the questions, which are provided by the present Standard Model (SM) of particle physics. The matter in the universe is made up of elementary fermions interacting through fields, which bosons are associated with. The elementary fermions are of two types: *leptons* and *quarks*. All have spin $\frac{1}{2}$, in units of $\hbar$, and are isolated into pairs of three families. Four types of interaction fields have been distinguished in Nature. The Standard Model excludes the gravitational field because of its insignificance with respect to scales of particle physics. The boson of the electromagnetic interaction field is the massless photon. The bosons of the weak interaction fields are the electrically charged $W^{\pm}$ and the neutral Z. The bosons of the strong interaction fields, the gluons, have zero mass, are electrically neutral, but charged with "color". All interaction field bosons have spin 1. The picture of the world in the Standard Model can then be summarized in Tables 1.1 and 1.2. Quarks c, b and t have much larger

mass than other quarks have, therefore the former are usually called heavy quarks or flavors, and the latter light quarks or flavors. $W^{\pm}$ and Z bosons carry mass, therefore the weak interaction is short-ranged, by the uncertainty principle, in an order of one-thousandth of the proton size.

The essence of the Standard Model is to treat elementary fermions as point-like particles and the interactions as emitting and absorbing field bosons. The underlying theory for this idea is quantum field theory and the calculation is carried out using perturbation theory. The construction has been guided by principles of symmetry and is expressed, mathematically, in terms of group theory as

$$U(1) \times SU(2) \times SU(3).$$

Theories contained in the SM are quantum electrodynamics (QED), electroweak dynamics (EW), combining the electromagnetic and weak interactions, and quantum chromodynamics (QCD). Under the EW theory, the interaction is induced by neutral current (NC), exchange of γ/Z, or charged current (CC), exchange of $W^{\pm}$. Up to now, the Standard Model has been proved successfully by tremendous experiments.

Table 1.1: *The elementary fermions in the Standard Model. Leptons and quarks are grouped in pair of families, separately. Masses are listed in parentheses, while charges in the last column in unit of the positive electron charge. All masses of neutrinos are upper limits. The masses of quarks are for comparison and the values, except that of t quark, are means of lower and higher boundaries with truncation [1].*

leptons	e (0.511 MeV)	μ (106 MeV)	τ (1.78 GeV)	-1 e
	ν_e (< 3 eV)	ν_μ (< 0.19 MeV)	ν_τ (< 18.2 MeV)	0 e
quarks	u (~ 3 MeV)	c (~ 1.25 GeV)	t (174 GeV)	$+\frac{2}{3}$ e
	d (~ 6 MeV)	s (~ 105 MeV)	b (~ 4.3 GeV)	$-\frac{1}{3}$ e

Investigating the structure of nucleons, i.e. protons and neutrons, is of fundamental importance in the development of the SM. The famous Rutherford scattering experiment represents a typical kind of experiments in this investigation. As early as in the 1950s, a *root-mean-square* charge radius of the proton was measured to be 0.74 ± 0.24 fm by McAllister and Hofstadter [2] using elastic scattering of 188 MeV electrons off a hydrogen target. In the late 1960s, Friedman, Kendall, Taylor and their collaborators carried out a similar but *inelastic* scattering experiment using electrons with energy up to 17 GeV and discovered point-like constituents within the proton [3] and the *scaling* behaviour [4]. Since then, many experiments, scattering various

Table 1.2: *The interaction field bosons in the Standard Model. The gravitation is excluded since its field quantum is still a postulate. The last column, strength, is relative to that of the strong interaction [1].*

Interaction	Field quantum	Mass (GeV)	Charge (e)	Strength
strong	g (gluon)	0	0	1
electromagnetic	γ (photon)	0	0	10^{-2}
weak	$W^{\pm}$	80	± 1	10^{-5}
	Z	91	0	

lepton beams off, usually *fixed*, nuclear targets, have been carried out and confirmed the discovery. These experiments, together with other kinds and theoretical advance, eventually led to the establishment of the QCD.

The advent of HERA in the early 1990s started a new era of scattering experiments. With the development of modern particle accelerators, at HERA, electrons and protons are accelerated simultaneously and collide at a center-of-mass energy larger than 300 GeV, probing the internal structure of the proton down to a scale of 0.001 fm. With two collider experiments, H1 and ZEUS, one of the most important discoveries is strong scaling violation at very low x.

While the scaling has perfect explanation within the naïve quark-parton model (QPM), scaling violation is anticipated from the QCD-improved parton model. In the parton model, the proton contains a sea of quarks, antiquarks as well as gluons, whereas it is made up of two u and one d quarks in the *static* quark model, which is successful in classifying hadrons, i.e. mesons and baryons. For lepton-nucleon deep inelastic scattering (DIS), the process can be regarded as an elastic scattering between the lepton and a parton within the nucleon. The probability of finding a certain parton in the nucleon is parameterized in a set of parton distribution functions (PDFs), which is usually obtained by global fits of data from DIS and other hard scattering in hadron-hadron interactions. The essence of the parton model is that the same PDFs should work for all processes, namely PDFs are universal. Therefore precise knowledge of the PDFs is a crucial input for accurate calculation and measurement of SM processes and finding new physics beyond the SM, as at the forthcoming Large Hadron Collider (LHC).

Measurements of inclusive cross sections for NC and CC DIS have been made at HERA by H1 and ZEUS. With the wide kinematic range and high statistics, these data are becoming more and more dominant in global fits for PDFs of the proton. Fits using H1 or ZEUS data alone have also been

published [5]. A very recent publication [6] of ZEUS also uses data of jet cross sections [7] to constrain the gluon distribution.

The charm quark, c, since its discovery by Ting [8] and Richter [9] in 1974, has always been of great theoretical and experimental interest. Being one of the heavy quarks, it provides a natural hard scale in perturbative QCD (pQCD) calculations. On the other hand, it is the lightest heavy quark and has comparatively high production rate in experiments. Charm quark production has been measured extensively in γp and NC DIS reactions. While the former has improved understanding of the hadronic behaviour of the photon significantly, the latter, which is directly sensitive to the gluon content in the proton, has achieved a level of precision that is highly expected to constrain the gluon distribution for global fits of PDFs in the future.

However, charm production in CC DIS has not yet been measured at HERA due to its relatively small cross section. In CC reactions, the charm quark is produced dominantly via hard scatter of $W^{\pm}$ bosons on the strange quark that is one kind of sea flavors in the proton. Even though some analyses have been performed at fixed-target experiments to constrain the strange sea, they are in a kinematic region very different from that of HERA. The behavior of the strange sea in the whole kinematic range is still uncertain. Therefore, it would be very interesting if such a measurement could be performed at HERA as well.

In this thesis, charm quark production is investigated in charged current deep inelastic scattering with the ZEUS detector at HERA. The e^+p data collected by the end of 2000, corresponding to an integrated luminosity of $110\,\mathrm{pb}^{-1}$, is used, whereas the e^-p data, of $16\,\mathrm{pb}^{-1}$, is not used, expecting a very small cross section. The charm quark is tagged in the hadronic final state via $D^{*\pm}$, undergoing the decay $D^{*+} \to D^0\pi_s^+$ with $D^0 \to K^-\pi^+$, as well as the charge conjugated channel.

A phenomenological and experimental survey on deep inelastic scattering and charm production is given in Chapter 2. The experimental setup for this analysis is then described in Chapter 3, continued in Chapter 4 with a description of the relevant event reconstruction. The succeeding chapters contain the main body of the analysis: Chapter 5 describes the Monte Carlo simulation, Chapter 6 discusses criteria for selecting CC events, Chapter 7 discusses identification of charged D^* mesons and Chapter 8 presents the results of this analysis and the discussions and the conclusions. A summary is then given in Chapter 9. In addition, a Monte Carlo study at generator level is presented in Appendix A, candidate events are visualized in Appendix B, and a description of the Bayesian approach for setting upper-limit is discussed in Appendix C.

Chapter 2

Phenomenology of charm production in DIS

Lepton-nucleon deep inelastic scattering is a class of processes where leptons, i.e. electrons, muons or neutrinos as well as their anti-particles, are scattered off nucleons, i.e. protons or neutrons, with large momentum transfer. The interaction is mediated by a virtual vector boson, which is γ or Z in electrically neutral current, and $W^{\pm}$ in charged current events. While hadron-hadron scattering suffers from large background in the hadronic final state, lepton-induced DIS offers a clear leptonic signal in the final state. Hence it is an ideal means of studying the nucleon internal structure in detail. Special kinematic variables are needed to describe the DIS process and they are defined in Section 2.1.

The parton model has been proven successful in describing DIS phenomenologically. Cross sections can be explained in terms of parton distributions in the nucleon through well defined structure functions. Therefore, the determination of parton distributions is not only a test for the Standard Model, but also, due to the universality, a basis for accurate theoretical predictions. Section 2.2 is devoted to this topic.

Heavy quark, charm and beauty, production in DIS is of particular interest because the large mass, $m_c \approx 1.5\,\text{GeV}$ or $m_b \approx 5\,\text{GeV}$, provides a hard scale in pQCD calculations. Charm or beauty production in NC DIS is directly sensitive to the gluon content in the nucleon and charm production in CC DIS is sensitive to the strange sea. While the former kind of measurement has been extensively carried out and published with precise results by the two collider experiments H1 [10, 11, 12] and ZEUS [13, 14] at HERA, the latter has not been taken on until very recently in this thesis and in another one from H1 [15], for good reasons as one will see in this work. These are discussed in Sections 2.3 and 2.4.

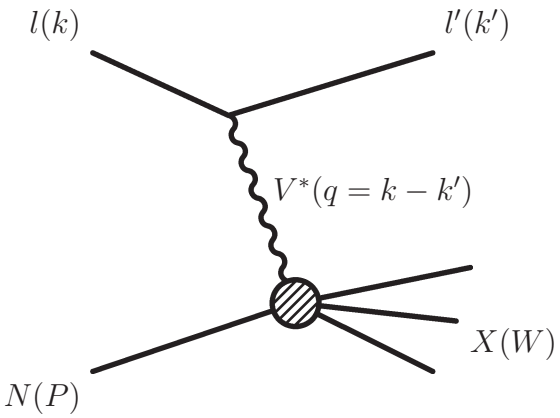


Figure 2.1: *Feynman diagram for lepton-nucleon DIS. The four momenta are given in parentheses.*

2.1 Kinematics of DIS

A deep inelastic scattering process is generally of the form

$$lN \to l'X, \tag{2.1}$$

where l and l' represent the incoming and outgoing leptons, N the nucleon and X the hadronic final state. The process, illustrated in Figure 2.1, is initiated by the exchange of a virtual vector boson V^*. Associating M for the nucleon mass and the four-momenta k, k' and P for l, l' and N respectively, the four-momentum of the virtual boson is

$$q = k - k', \tag{2.2}$$

and a set of Lorentz invariants is defined as well:

$$s = (k + P)^2, \tag{2.3}$$

the square of the center-of-mass energy,

$$Q^2 = -q^2, \tag{2.4}$$

the virtuality of the exchanged boson,

$$x = \frac{Q^2}{2q \cdot P}, \tag{2.5}$$

the *Bjorken scaling* variable, which is interpreted in the parton model as the fraction of the momentum of the incoming nucleon carried by the struck parton,

$$\nu = \frac{q \cdot P}{M}, \tag{2.6}$$

the energy transferred from the incoming lepton to the nucleon in the nucleon rest frame,

$$y = \frac{q \cdot P}{k \cdot P}, \tag{2.7}$$

the fraction of the energy of the incoming lepton transfered to the nucleon in the nucleon rest frame, therefore named *inelasticity*, and

$$W^2 = (q + P)^2, \tag{2.8}$$

the invariant mass squared of the hadronic system X.

The important relations, ignoring masses of the lepton and nucleon, are

$$Q^2 = sxy \tag{2.9}$$

and

$$W^2 = Q^2(\frac{1}{x} - 1). \tag{2.10}$$

It is obvious that at fixed center-of-mass energy, there are only two independent variables.

The square of the four-momentum transfer, or its positive value Q^2, determines the hardness, or the resolving power, of the interaction. The spatial resolution, Δb, can be approximated by

$$\Delta b \approx \frac{\hbar c}{\sqrt{Q^2}} = \frac{0.197}{\sqrt{Q^2}}\,\mathrm{GeV\,fm}. \tag{2.11}$$

For deep inelastic scattering, Q^2 is much higher than $1\,\mathrm{GeV}^2$, hence the virtual boson probes "deeply" into the nucleon, $\Delta b \ll 0.2\,\mathrm{fm}$. In the kinematic range of HERA, $\sqrt{s} = 318\,\mathrm{GeV}$, the internal structure of the proton can be resolved up to an order of $0.001\,\mathrm{fm}$, or $10^{-18}\,\mathrm{m}$.

2.2 Structure functions and parton distributions

Most formulae appearing in this section are not rigorous. Exact forms as well as other details the reader is referred to the relevant literature or some general references [16, 17, 18, 1].

2.2.1 Structure functions

At lowest order in electroweak interactions, the cross section may be split into leptonic and hadronic parts,

$$d\sigma \sim L^{\mu\nu} W_{\mu\nu}. \tag{2.12}$$

$L^{\mu\nu}$ is the leptonic tensor corresponding to the coupling of the exchange boson to the lepton, the upper vertex in Figure 2.1, and can be evaluated explicitly. The hadronic tensor $W_{\mu\nu}$, the lower vertex in Figure 2.1, describing the interaction of the electroweak current with the nucleon, is defined to all orders in the strong interaction, and may be expanded in terms of *structure functions* W_i, with $i = 1, 2$ and 3 [16],

$$\begin{aligned} W_{\mu\nu} = &- \left(g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) W_1(x, Q^2) \\ &+ \left(p_\mu - q_\mu \frac{p \cdot q}{q^2} \right) \left(p_\nu - q_\nu \frac{p \cdot q}{q^2} \right) \frac{1}{M^2} W_2(x, Q^2) \\ &- i\epsilon_{\mu\nu\lambda\kappa} q^\lambda q^\kappa \frac{1}{M^2} W_3(x, Q^2). \end{aligned} \tag{2.13}$$

Only unpolarized terms have been considered in the expansion, that is the case for all discussions in this thesis. The functions W_i are usually replaced by structure functions F_i, which will turn out to be particularly simple in the parton model,

$$\begin{aligned} F_1(x, Q^2) &= \quad W_1(x, Q^2) \\ F_2(x, Q^2) &= \frac{\nu}{M} W_2(x, Q^2) \\ F_3(x, Q^2) &= \frac{\nu}{M} W_3(x, Q^2). \end{aligned} \tag{2.14}$$

It should be mentioned that relation (2.12) is process dependent. It holds for CC, whereas for NC, the splitting sums up contributions from the exchange of γ, of Z, and their interference. Consequently, structure functions for NC are more complicated than for CC.

Explicitly, the differential cross sections for the scattering of unpolarized leptons on unpolarized nucleons is given in terms of structure functions by the generic form

$$\frac{d^2\sigma^i}{dx dQ^2} = \frac{2\pi\alpha^2}{xQ^4} \eta^i \left[Y_+ F_2^i \mp Y_- x F_3^i - y^2 F_L^i \right] \tag{2.15}$$

with i = NC or CC for neutral or charged current reactions, respectively, and $Y_\pm = 1 \pm (1-y)^2$. The $\mp$ sign corresponds to the incoming $l^\pm$, namely antileptons, e.g. e^+ or $\bar{\nu}$, and leptons, e.g. e^- or ν.

$$F_L^i = F_2^i - 2xF_1^i \tag{2.16}$$

are the longitudinal structure functions. In the naïve quark-parton model, $F_L^i = 0$, there are only two independent structure functions F_1 and F_3. For unpolarized $e^\pm$ beams, the factor

$$\eta^{\mathrm{NC}} = 1,$$

whereas

$$\eta^{\mathrm{CC}} = \frac{1}{2}\left(\frac{G_F M_W^2}{4\pi\alpha}\frac{Q^2}{Q^2+M_W^2}\right)^2, \tag{2.17}$$

with G_F the Fermi constant.

2.2.2 The parton model

Lepton-nucleon deep inelastic scattering has been phenomenologically described by the parton model since the observation of Bjorken scaling, namely structure functions are independent of Q^2 for fixed x at around $x \sim 0.25$. In the naïve quark-parton model, the parton, coined by Feynman, is identified as a quark and the inclusive DIS cross section is approximately an incoherent sum of contributions of lepton elastic scattering off nearly free partons. In the QCD-improved parton model, more subprocesses, e.g. QCD Compton scattering and boson-gluon fusion (BGF), are involved in calculations and structure functions are of logarithmic Q^2 dependence, which is confirmed experimentally by the observation of scaling violation. The scaling violation becomes particularly strong at small x, where the gluon contribution is dominant.

The validity of the parton model relies on two physical concepts: the Lorentz contraction and time dilation of internal states of the nucleon and the comparatively long time process of hadronization. The nucleon, in its own rest frame, is made up of a set of partons in virtual states with nonzero lifetime, $\tau > \tau_0$. In the center-of-mass system, the nucleon carries momentum of about $\sqrt{Q^2}/2$. Usually, $Q^2 \gg 1\,\mathrm{GeV}^2$ in DIS, therefore by the special relativity, the lifetime of the virtual state becomes $\tau/\sqrt{1-v^{*2}/c^2} \gg \tau$, with v^* the velocity. Combined with Lorentz contraction, as Q^2 goes to infinity, the time it takes the lepton to cross the nucleon vanishes, while the collection of partons is effectively "frozen" during the transit. After the collision, the

struck parton undergoes the process of hadronization, by which quarks and gluons transform into the observed particles, and this happens too late to influence the hard scattering. Therefore lepton inelastic scattering off the nucleon can be treated, in Born approximation, as an elastic scattering off the parton.

The parton model is generalized in the field theory by the *factorization theorem*, which states the DIS cross section as the convolution of non-perturbative parton distribution functions, f_i, and pQCD calculable hard-scattering cross sections, $\hat{\sigma}$,

$$\sigma(x,Q^2) \sim \sum_{i=q,\bar{q},g} \int_x^1 \frac{d\xi}{\xi} f_i(\xi,\mu_F^2,\mu^2) \times \hat{\sigma}(\frac{x}{\xi},\frac{Q^2}{\mu^2},\frac{\mu_F^2}{\mu^2},\alpha_s(\mu^2)), \tag{2.18}$$

with i running over all partons, quarks, antiquarks and gluons, and ξ the fraction of the nucleon momentum carried by the parton. μ is the usual renormalization scale. μ_F is a factorization scale: any parton that is off-shell by μ_F^2 will contribute to $\hat{\sigma}$, otherwise will be absorbed into f_i. The PDFs, therefore, depend on the choice of μ_F, a "factorization scheme". Two most popular schemes are DIS and $\overline{\text{MS}}$, which differ in the treatment of gluon contributions. Comparing with (2.12), (2.13) and (2.14), relations for the structure functions F_i can be derived in terms of parton distribution functions. Although the definition of parton distributions is scheme dependent, a set of PDFs obtained from one experiment can be used to predict measurements of another experiment, even of different type, as long as the same scheme is used.

In the naïve QPM, only quarks can couple to the electroweak current. Structure functions are expressed in terms of scale-independent quark distributions $q_i(x)$ [17],

$$\begin{aligned} xF_1 &= \frac{1}{2}\sum_i xq_i(v_i^2+a_i^2), \\ F_2 &= \sum_i xq_i(v_i^2+a_i^2), \\ xF_3 &= 2\sum_i xq_i(v_i a_i), \end{aligned} \tag{2.19}$$

where index i runs over all flavors of quarks and antiquarks allowed for the interaction. v_i and a_i are electroweak couplings of vector and axial-vector, respectively. For CC interactions, $v_i = a_i = 1$ for quarks and $v_i = -a_i = 1$ for antiquarks. For NC, v_i and a_i depend on the exchanged boson γ, Z, or

their interference, and are more complicated. A direct consequence of (2.19) is the Callan-Gross relation,

$$2xF_1 = F_2, \tag{2.20}$$

a characteristic of spin 1/2 partons. For the process $e^+p \to \bar{\nu}X$, F_2 and F_3 can be expressed as

$$\begin{aligned} F_2^{W^+} &= 2x(d + \bar{u} + s + \bar{c} + b) \\ xF_3^{W^+} &= 2x(d - \bar{u} + s - \bar{c} + b) \end{aligned} \tag{2.21}$$

at the HERA energy scale. For $e^-p \to \nu X$, the structure functions $F_{2,3}^{W^-}$ are obtained by the flavor interchanges $d \leftrightarrow u$, $s \leftrightarrow c$ etc.

When considering QCD, quarks undergo processes of emission and absorption of gluons and gluons undergo processes of creation and annihilation of $q\bar{q}$ pairs as well as self-coupling. The difference between the naïve and QCD-improved parton models is depicted in Figure 2.2. The incoherent assumption in the naïve parton model will be spoiled and the deviation rises as Q^2 increasing. Therefore, structure functions aquire a Q^2 dependence, called *scaling violation*. The Callan-Gross relation is broken as well. Parton distribution functions, as shown in (2.19), become Q^2 dependent, which is known as parton evolution.

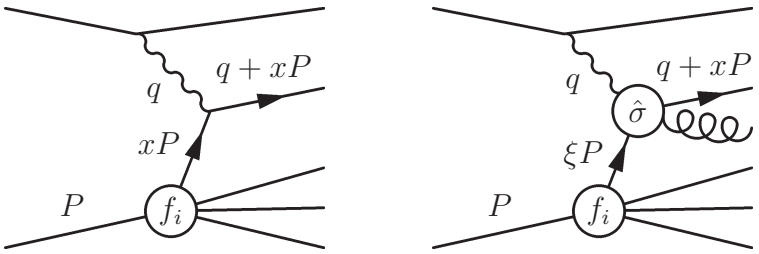


Figure 2.2: *Feynman diagrams for the naïve and the QCD-improved parton model. Relevant four-momenta are shown as well. In the QCD-improved model, right, the parton undergoes higher order QCD processes before or after hard scattering.*

The Q^2 evolution equations are known as the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [19]. With the quark and gluon distributions $q(x, Q^2)$ and $g(x, Q^2)$, the DGLAP equations are given by

$$\frac{\partial}{\partial \ln Q^2} \begin{pmatrix} q \\ g \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \begin{pmatrix} P_{qq} & P_{gq} \\ P_{gq} & P_{gg} \end{pmatrix} \otimes \begin{pmatrix} q \\ g \end{pmatrix}, \tag{2.22}$$

where P_{ij}, with $i, j = q, g$, are splitting functions. The splitting function P_{ij} provides the probability of finding parton i in parton j with a given fraction, z, of parton j's momentum, as illustrated in Figure 2.3. Given a specific renormalization and factorization scheme, the splitting functions can be expanded in pQCD as a power series in α_s,

$$P_{ij}(x, Q^2) = P_{ij}^{(0)}(x) + (\frac{\alpha_s}{2\pi})P_{ij}^{(1)}(x) + \cdots . \tag{2.23}$$

The first term, leading order (LO), in the expansion is independent of the factorization scheme and the truncation after it corresponds to the naïve parton model. Beyond LO, splitting functions depend on the factorization scale. The truncation after the first two terms defines the next-to-leading order (NLO) DGLAP evolution and gives QCD corrections to the structure functions in (2.19).

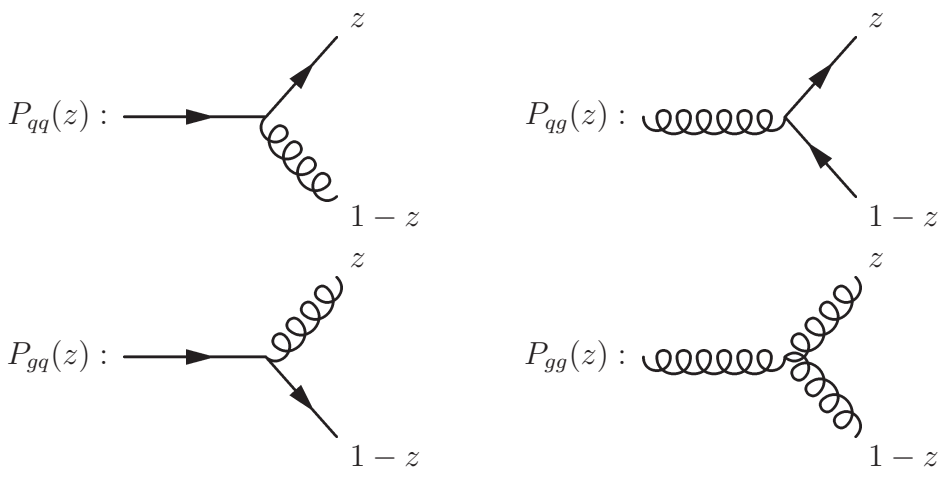


Figure 2.3: *The DGLAP splitting functions.*

In general, the series (2.23) contains both, terms proportional to $\ln Q^2$ and to $\ln(1/x)$. While the former explains successfully the logarithmic Q^2 dependence of structure functions, the latter leads to inapplicability of DGLAP equations in the very small x region. It is solved, at leading order, by the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equations [20] and, in general, by the Ciafaloni-Catani-Fiorani-Marchesini (CCFM) equations [21]. They differ in the order of summing up the series, where DGLAP is strongly ordered in Q^2, BFKL in x and CCFM in *angle*, a combination of both.

2.2.3 Determination of parton distributions

The parton distributions can not be calculated from first principles. They have to be determined by fits to data. The NLO DGLAP equations provide a reliable formalism. The parton distributions are parameterized in analytic functions of x at some start scale $Q^2 = Q_0^2$ and distributions at all values of Q^2 are then obtained by the evolution equations. The choice of Q_0^2 is somewhat arbitrary, but large enough to ensure that $\alpha_s(Q_0^2)$ is small enough for perturbative calculations. A fit is then performed to obtain parameters to describe data points in the covered x, Q^2 region.

The choice of parton distributions are usually

- $xu_v(x)$ for the u valence quark,
- $xd_v(x)$ for the d valence quark,
- $xS(x)$ for the total sea,
- $xg(x)$ for the gluon, and
- $x\Delta = x(\bar{d} - \bar{u})$ for the difference between the d and u contributions to the sea.

The sea is composed of all kinematically allowed flavors,

$$S = 2(\bar{u} + \bar{d} + \bar{s} + \bar{c} + \bar{b}),$$

with the assumption of $q_{\text{sea}} = \bar{q}$. The contributions of heavy quarks, c and b, require special theoretical treatment and will be discussed in the following section. The strange sea is not treated in that way, but suppressed relative to the u and d sea by a simple factor, $\bar{s} = (\bar{u} + \bar{d})/4$, namely 50%. This constraint comes, so far, from the CCFR NLO QCD analysis of neutrino charm production [22].

The parton distributions are traditionally determined in global fits performed by several theoretical groups, e.g. MRST [23] and CTEQ [24], using data from lepton-nucleon DIS experiments, dominantly, as well as related hadron-hadron scattering experiments. Fits made by H1 [25] and ZEUS [5, 6] also exist. The wide kinematic range covered by the HERA data has stimulated the extension of the x-Q^2 region in global fits. While the fixed-target data are important for a precise determination of the valence distributions, the HERA data is becoming dominant in determining the sea and gluon distributions. A brief description of data used in global fits is given below:

- The xF_3 data from CCFR [26] give most reliable u_v and d_v in the mid-x range, $0.1 \lesssim x \lesssim 0.65$;

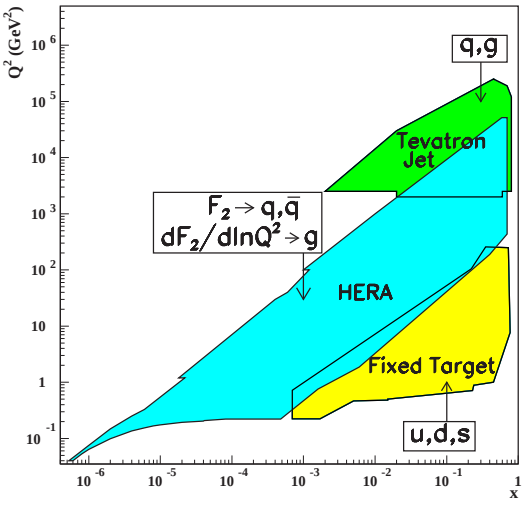


Figure 2.4: *Kinematic regions in x-Q^2 plane by fixed-target and collider experiments. The important constraints they make on the parton distribution are shown together [1].*

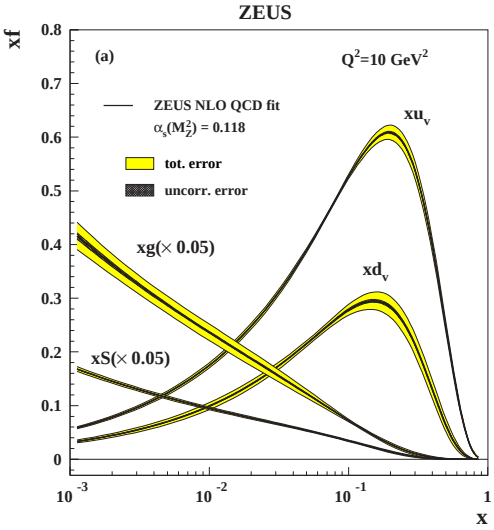


Figure 2.5: *Parton distribution functions within the proton.*

- NMC data on the ratio F_2^d/F_2^p [27], with d denoting the deuteron and p the proton, constrains the ratio d_v/u_v at high-x;
- The F_2^d and F_2^p data from BCDMS [28], E665 [29], NMC [30], SLAC [31] as well as the F_2 data from CCFR/NuTeV [32] constrain the sea distributions in the respective x-Q^2 range;
- CCFR dimuon data [22] constrains the strange sea, $\bar{s}$;
- Tevatron high-E_T jet data [33] constrains the gluon distribution;
- The F_2 data, measured in inclusive DIS cross sections, from H1 [25, 34] and ZEUS [35, 36] determine the low-x sea and gluon distributions;
- The inclusive jet cross section [7], which is directly sensitive to the gluon distribution, is also included in a recent NLO QCD analysis by ZEUS;
- The $F_2^{c\bar{c}}$ data from H1 and ZEUS, obtained in measurements of charm production in NC DIS, quite certainly, will constrain the gluon distribution in future global fits.

Kinematic regions covered by fixed-target and collider experiments are shown in Figure 2.4. An example of parton distribution functions is shown in Figure 2.5. Recently, next-to-next-to-leading order (NNLO) analyses are also performed [37]. The most up-to-date information on parton distribution functions is maintained on the Web [38].

2.3 Charm production in DIS

To discuss the charm production in DIS within the parton model, the Cabibbo-Kobayashi-Maskawa (CKM) matrix [39] has to be taken into account. In the Standard Model, the quark mass eigenstates are not the same as the weak eigenstates. The transition of the quark flavors in the charged current weak interaction is by convention parameterized in terms of a 3×3 unitary matrix V operating on the charge $-\frac{1}{3}e$ quark mass eigenstates, d, s and b,

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 0.97 & 0.22 & 0.0037 \\ 0.22 & 0.97 & 0.042 \\ 0.094 & 0.040 & 1.0 \end{pmatrix}, \tag{2.24}$$

with values from PDG [1]. For the four-quark case, the matrix reduces to a 2×2 matrix and is described by a single parameter, the Cabibbo angle. The charm quark, for instance, couples with W^+ by a rotated state s',

$$s' = |V_{cs}|^2 s + |V_{cd}|^2 d + |V_{cb}|^2 b. \tag{2.25}$$

While the last term is totally negligible, the $|V_{cd}|^2$ term, in relevant physics processes, is called Cabibbo suppression.

Within the parton model, the leading order, $\mathcal{O}(\alpha_s^0)$, QCD subprocess for charm quark production in CC DIS is quark initiated

$$W^+ s \to c \tag{2.26}$$

and

$$W^+ d \to c, \tag{2.27}$$

as shown in Figure 2.6a. The latter is Cabibbo suppressed. As d a valence quark in the proton, however, its contribution should be considered, especially in the region of high x.

The next-to-leading order, $\mathcal{O}(\alpha_s)$, QCD subprocesses are the QCD Compton (QCDC) scattering, where a real gluon is emitted, the virtual gluon corrections and the boson gluon fusion (BGF), shown in Figure 2.6b-d. Among them, the BGF

$$W^+ \mathrm{g} \to c\bar{s} \tag{2.28}$$

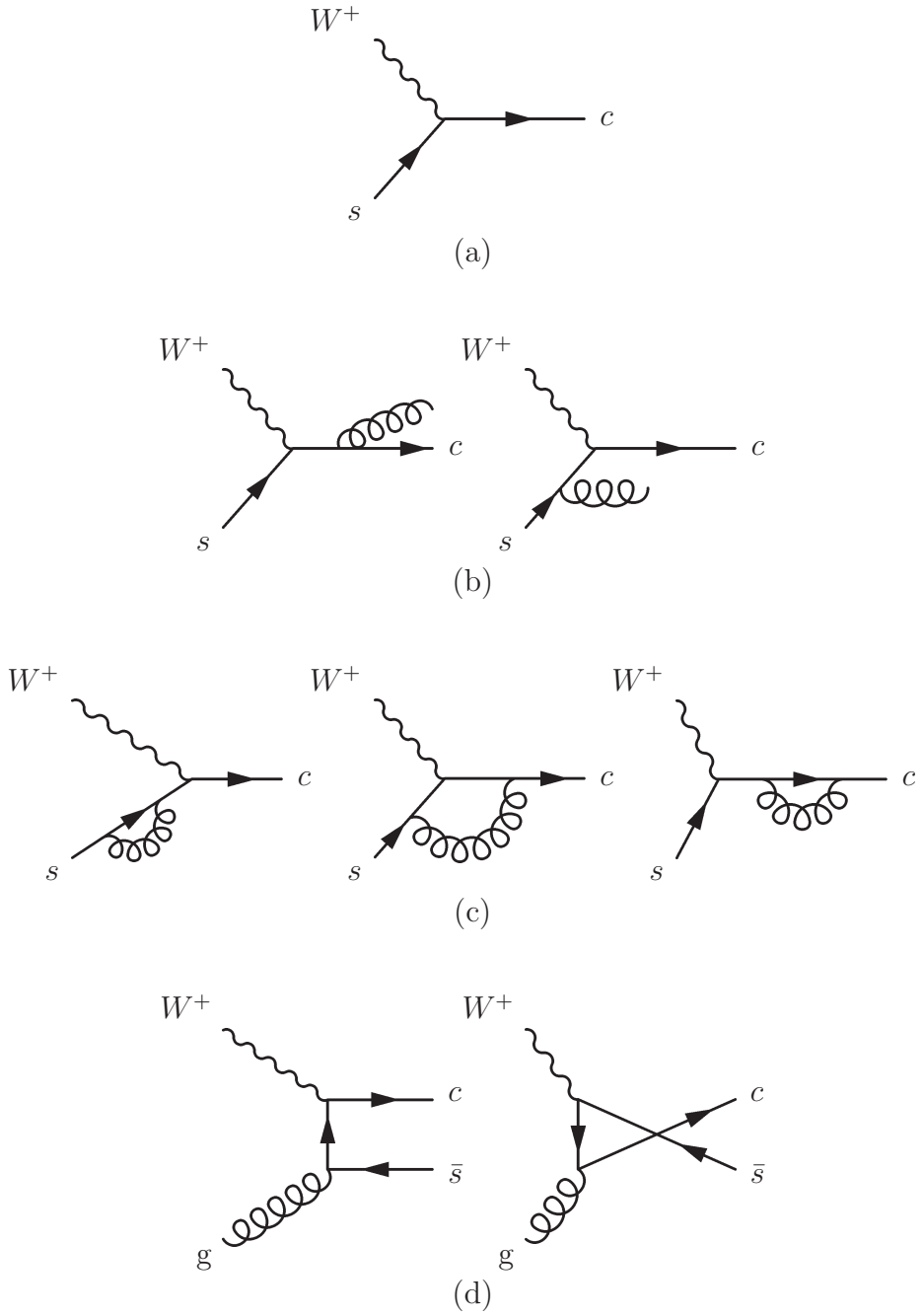


Figure 2.6: *Feynman diagrams contributing to charm production in charged current reactions up to $\mathcal{O}(\alpha_s)$. They are, separately, (a) Born term, (b) real gluon emission, or QCD Compton scattering, (c) virtual gluon corrections and (d) boson-gluon fusion. The Cabibbo suppressed contributions are obtained by substituting all s quarks by d quarks.*

is particularly interesting. As x decreases, the gluon content in the proton increases rapidly, therefore compensates the higher QCD order and the subprocess becomes dominant. For the purpose of probing the strange sea, its contribution is a kind of "background" and should be distinguished from other, LO or NLO, subprocesses.

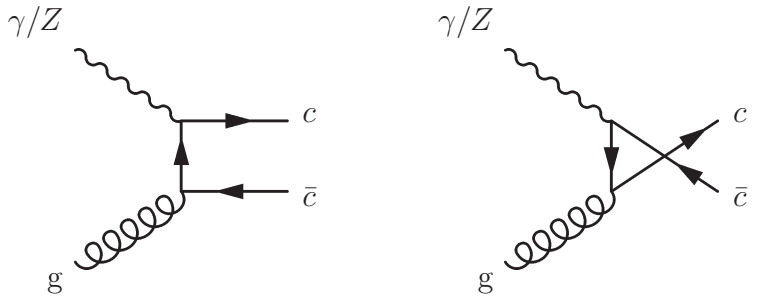


Figure 2.7: *Feynman diagrams contributing to charm production in neutral current reactions at the order $\mathcal{O}(\alpha_s)$. They are the boson-gluon fusion in t-channel, left, and u-channel, right.*

In the NC DIS case, there is no corresponding quark initiated subprocess if there is no intrinsic charm in the nucleon. The leading order QCD subprocess is at the order $\mathcal{O}(\alpha_s)$, i.e. the BGF, as shown in Figure 2.7. In contrast to CC, pairs of $c\bar{c}$ appear in the final state. The cross section can be expressed in terms of the charm contribution, $F_{2,L}^{c\bar{c}}$, to the NC DIS structure functions $F_{2,L}^{\mathrm{NC}}$,

$$\frac{d^2\sigma^{c\bar{c}}(x,Q^2)}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left\{\left[1+(1-y^2)\right]F_2^{c\bar{c}}(x,Q^2) - y^2F_L^{c\bar{c}}(x,Q^2)\right\}. \quad (2.29)$$

In the region $Q^2 \ll M_Z^2$, $F_L^{c\bar{c}}$ is negligible, therefore, $F_2^{c\bar{c}}$ can be extracted directly from the measurement of differential cross sections, and in turn, will constrain the gluon distribution.

As one of the heavy flavors, the charm quark provides a naturally hard scale for pQCD calculations. There are several schemes, however, in the treatment of heavy quark masses, $m_{c,b}$:

- The fixed-flavor-number scheme (FFNS) or "massive" scheme [40], in which heavy quarks are non-active flavors in the nucleon and are produced through hard scatter such as BGF;

- The zero-mass variable-flavor-number scheme (ZM-VFNS) or "massless" scheme [41], in which heavy quarks are just the contents of the nucleon and can enter the reaction directly;

- The general-mass variable-flavor-number scheme (GM-VFNS), with two prescriptions from Aivazis-Collins-Olness-Tung (ACOT) [42] and Thorne-Roberts [43], in which heavy flavors are treated as massive quarks when $Q^2 \sim m_{c,b}^2$ and are treated as massless partons when $Q^2 \gg m_{c,b}^2$.

While the massive scheme works well near the threshold of heavy quark production and the massless scheme works better in the higher region, the GM-VFNS provides pQCD predictions over the whole kinematic range in Q^2.

Many calculations have been presented for charm production in DIS, particularly in the case of NC, as the data from HERA keeps improving our understanding of the proton structure. Several predictions for the CC case are reviewed below since they are relevant to this thesis.

Overall cross sections for heavy quark production at HERA were predicted by Ingelman and Schuler [44]. Charm production in CC DIS is estimated as

$$\sigma^{e^+p}(d \to c) = 0.75\,\text{pb},$$
$$\sigma^{e^+p}(s \to c) = 3.3\ \ \text{pb},$$

for leading order QPM, and

$$\sigma^{e^+p}(\text{g} \to c\bar{d}) = 0.46\,\text{pb},$$
$$\sigma^{e^+p}(\text{g} \to c\bar{s}) = 8.2\ \ \text{pb},$$

for BGF. Although the predictions are quite rough, it is clear that the contribution of BGF is larger than that of LO flavor mixing. It is because of the dominance of the gluon distribution at small x covered by HERA. The largest uncertainties in the BGF cross sections arise from the choice of different parameterizations for the gluon distribution, less than 40%, and from varying the charm quark mass from the center value 1.5 GeV, 68% and 36% with a decrease and an increase by 0.3 GeV respectively.

A study made by Barone, D'Alesio and Genovese [45] concentrates on the theoretical uncertainties arising from the choice of factorization scale, heavy flavor mass scheme and parton distribution fit. It is found that the main uncertainty, $\sim 40\%$, comes from the choice of mass scheme. While only Cabibbo unsuppressed processes are taken into account up to NLO, the total cross section is given as

$$\sigma(e^+p \to \bar{\nu}_e cX) \sim 5\,\text{pb}$$

in the kinematic region $Q^2 > 200\,\text{GeV}^2$, $x > 0.006$ and $\sqrt{s} \approx 300\,\text{GeV}$.

Complete NLO QCD corrections to charm production in CC DIS have been studied by Krezter and Stratmann [46]. Differential cross sections for

D-meson production, using the Peterson fragmentation function [47], are calculated and a cut on a final-state scaling variable, $z = p_D \cdot P/q \cdot P$, with p_D the four-momentum of D-meson, is proposed to suppress the BGF contribution. The cross sections for $e^- p \to \nu_e \bar{D} X$ are found "all in the ballpark" of about 50 pb for $x \approx 0.05$ in the kinematic region $Q^2 > 500\,\mathrm{GeV}^2$, $0.01 \leq y \leq 0.9$ and $\sqrt{s} = 300\,\mathrm{GeV}$. In the case of the charge-conjugated process, $e^+ p \to \bar{\nu}_e D X$, although the d valence quark contribution should be taken into account, it only becomes significant at large x. Therefore, the prediction can be translated directly to D meson production, and in turn by the fragmentation fraction, $f(c \to D)$, to charm production, in an order of 10 pb.

In conclusion, the calculations are consistent within large theoretical uncertainties and all predict roughly 10 pb total cross section for charm production in CC DIS in the HERA kinematic region.

2.4 Experimental situation

For the time being, the most sensitive constraints on the strange sea are made by the CCFR dimuon data, with neutrino or antineutrino scattering off an iron target. In the NLO QCD analysis [22], it is found that the quark sea is not SU(3)-symmetric and the strange sea has an overall suppression of the momentum distribution, $\int x\bar{s}dx$, of about 50%. No shape difference is found between the strange and non-strange components of the sea, which is not consistent with their LO analysis [48] but the discrepancy is understood. Also no shape difference is found between the s and $\bar{s}$ distributions. The analysis is performed in a mid-x range, $0.01 \lesssim x \lesssim 0.20$. The entire x dependence of the flavor-decomposed light quark sea is, however, still rather unclear. Although the strange sea can also be inferred from a comparison of charged lepton and neutrino structure functions by the "$\frac{5}{18}$ rule", $F_2^{lN}/F_2^{\nu N} = \frac{5}{18}\{1 - \frac{3}{5}[(s+\bar{s})/(q+\bar{q})]\}$, it suffers from the fact that the strange sea appears only as a small residual of two large quantities, $x\bar{s} \ll F_2^{lN}, F_2^{\nu N}$.

At HERA, measurements of structure functions for both NC and CC DIS have obtained highly accurate results, and become more and more dominant in global fits for the parton distribution functions. The charm contribution, $F_2^{c\bar{c}}$, to the NC DIS structure function has also been measured extensively, as shown in Figure 2.8. With increasing accuracy, it is highly expected that the $F_2^{c\bar{c}}$ data will be used in future global fits to constrain the gluon distribution.

However, the lack of a corresponding measurement in CC DIS has been existing for long. The limitation is mainly due to the comparatively small cross section and low charm detection efficiency. The charm quark is tradi-

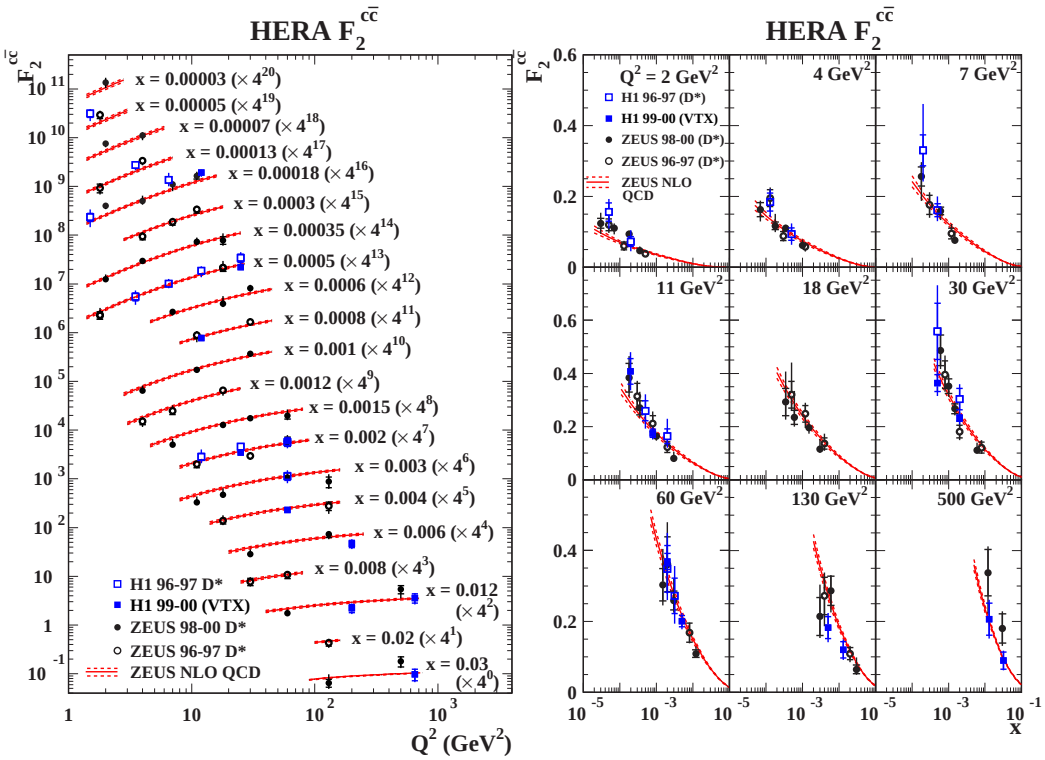


Figure 2.8: *The charm contribution, $F_2^{c\bar{c}}$, to the proton structure function F_2 for NC DIS.*

tionally identified in the hadronic final state by the charged D^* through so called "golden" channel,

$$D^{*+} \to D^0 \pi_s^+ \text{ with } D^0 \to K^- \pi^+, + c.c. \tag{2.30}$$

where π_s represents a "slow" pion, to be explained in Section 7.1. Taking account of the fragmentation fraction of the charm quark to D^{*+} and branching ratios of the decay modes involved in the golden channel, the charm detection efficiency would be

$$\epsilon = f(c \to D^{*+}) \times \mathcal{B}(D^{*+} \to D^0 \pi_s^+) \times \mathcal{B}(D^0 \to K^- \pi^+) \times \mathcal{A} \approx 0.006 \mathcal{A},$$

with $\mathcal{A}$ denoting acceptance and using values of the fragmentation fraction $f(c \to D^{*+}) \approx 0.22$ [49], branching ratios $\mathcal{B}(D^* \to D^0 \pi_s^+) \approx 68\%$ and $\mathcal{B}(D^0 \to K^- \pi^+) \approx 3.8\%$ [1]. A typical value for $\mathcal{A}$ in ZEUS kinematic region is about 30% [13]. Using the HERA I integrated luminosity for e^+p data, $\mathscr{L} \approx 110\,\text{pb}^{-1}$ and the estimated charm cross section in the preceding section, $\sigma(e^+p \to \bar{\nu}_e cX) \approx 10\,\text{pb}$, the number of D^* observed in the golden channel would be expected as

$$N_{\text{obs}}^{D^*} \approx 2, \tag{2.31}$$

which would be a quite difficult measurement, if not impossible.

In the year 2000, HERA started a major upgrade with an expectation to achieve an integrated luminosity, in a short period, several times as much as that having been accumulated by then. That arose the interest, both theoretically and experimentally, in charm production in CC DIS at HERA. The upgrade suffered from unexpected high background, however. Even though the problem was eventually solved and HERA II has been providing luminosity since 2003, the performance is not as good as the early expectation. Nevertheless, it is still tempting to take such an analysis in the HERA I data and it would be a worthwhile investigation if enough integrated luminosity would be obtained in the future, or any deviation from theoretical expectation would be found.

Chapter 3

ZEUS and HERA

The experimental setup of the analysis are described in this chapter. A brief description of the HERA collider as well as the relevant data taking periods are give in Section 3.1. The ZEUS detector is described in Section 3.2, with some detailed discussion for the components tightly related to this analysis. The online, i.e. trigger and data acquisition, system and the offline system of ZEUS are explained in Sections 3.3 and 3.4.

3.1 The HERA collider

The HERA collider (**H**adron **E**lektron **R**ing**a**nlage) is the first and so far unique lepton-neutron collider on earth. It is located at DESY (**D**eutsches **E**lektronen **Sy**nchrotron) in Hamburg, Germany. The machine was proposed in 1981, approved in 1984, then built on schedule and within budget. The first electron-proton collisions were achieved in October 1991, the first luminosity was provided to the colliding beam detectors H1 and ZEUS in the summer of 1992. Since then HERA has been running successfully, with a major break started in September, 2000, for upgrading the machine to provide significantly higher instantaneous luminosity and longitudinal polarization to the collider detectors. For distinction, we speak of HERA I before the upgrade and HERA II after.

HERA has a circumference of 6.3 km and is located in a tunnel 10-15 m underground. HERA consists of two storage rings, for electrons[1] and protons separately. The two beam pipes run on top of each other and merge into one at two areas along the circumference, where the beams are made to collide at

[1] At HERA, electrons represent negatively charged electrons or positively charged positrons without special distinction. The data sample used in this analysis, however, is of the positron-proton colliding.

zero crossing angle to provide ep interactions for the experiments of H1 and ZEUS. In addition, there are two other fixed target experiments HERMES and HERA-B. The former makes use of the electron beam by colliding it with a polarized proton gas jet to study the spin structure of the proton. The latter used wire targets in the proton beam to study CP-violation in the B sector physics, and ceased to run in 2004.

While the electron ring magnets operate at ambient temperatures, the proton ring magnets are superconducting because the proton energy is 30 times higher than the electron energy. On the other hand, the mass of electron is much less than that of the proton, by a factor of about two thousands, and the synchrotron radiation energy loss is inversely proportional to the particle's mass to the fourth power, hence superconducting cavities are used for electrons to compensate their enormous energy losses. But ambient temperature cavities suffice for protons.

The electrons and protons are bunched, with bunches in one bunch train separated by 96 ns. Some small number of bunches are unpaired, namely the corresponding bunch in the other beam is empty, for background studies.

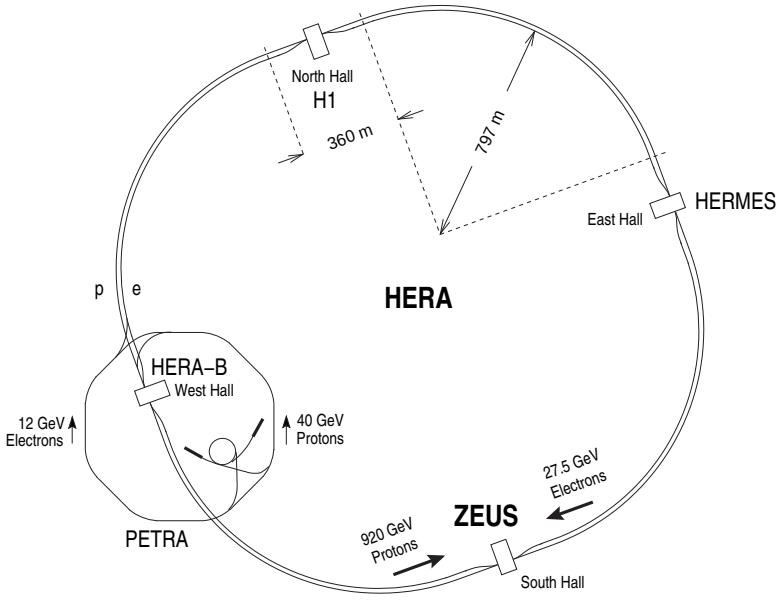


Figure 3.1: *The schematic drawing of HERA.*

Figure 3.1 is a schematic layout of HERA as well as the relevant pre-accelerator chain. The big circle-like rings are of HERA and the small one of PETRA (**P**ositron **E**lektron **T**andem **R**inga**a**nlage), from which 12 GeV electrons and 40 GeV protons are injected separately into HERA for further acceleration. The ZEUS detector is located in the south hall.

HERA initially began as an electron-proton collider, but switched to positron-proton collisions in 1994 because the electron beam severely suffered a short lifetime at high currents. When the problem was finally solved in 1998, HERA returned to operate with electrons for nine months. However, for maximizing statistics to further investigate the potential physics beyond the Standard Model, it was decided to switch again to positron-proton collisions in the spring 1999 and this configuration was kept till the end of HERA I.

The data used in this analysis are the e^+p collisions collected from 1995 to 2000, corresponding to an integrated luminosity of $111.6\,\mathrm{pb}^{-1}$. Table 3.1 lists the integrated luminosities collected in all HERA I data-taking periods, and curves in Figure 3.2 illustrate the chronological luminosity accumulation.

Table 3.1: *HERA-delivered and ZEUS-gated physics luminosities in the years 1994-2000. Relative uncertainties for ZEUS physics luminosities are listed in parentheses. Luminosities in bold font are used in this analysis.*

Year	Beam energy (GeV)		Integrated luminosity (pb^{-1})			
			HERA delivered		ZEUS physics	
	E_e	E_p	e^-p	e^+p	e^-p (%)	e^+p (%)
1994	27.5	820.0	1.1	5.1	0.3 (1.5)	3.0 (1.5)
1995	27.5	820.0	–	12.3	–	**6.6** (1.1)
1996	27.5	820.0	–	17.2	–	**10.8** (1.1)
1997	27.5	820.0	–	36.4	–	**27.9** (1.8)
1998	27.5	920.0	8.1	–	4.6 (1.8)	–
1999	27.5	920.0	17.1	28.5	12.1 (1.8)	**19.7** (2.3)
2000	27.5	920.0	–	66.4	–	**46.6** (2.3)
Total integrated luminosity for e^+p of 1995-2000 :						**111.6** (2.0)

3.2 The ZEUS detector

About 450 physicists coming from 52 different institutes of 15 countries form the ZEUS collaboration. The ZEUS detector is a general purpose detector with nearly hermetic calorimetry coverage. Figure 3.3 shows cross sectional views of the detector. The coordinate system [50] is chosen such that the proton beam points along the z-axis, the x-axis points to the center of the HERA ring, and the y-axis completes a right-handed coordinate system,

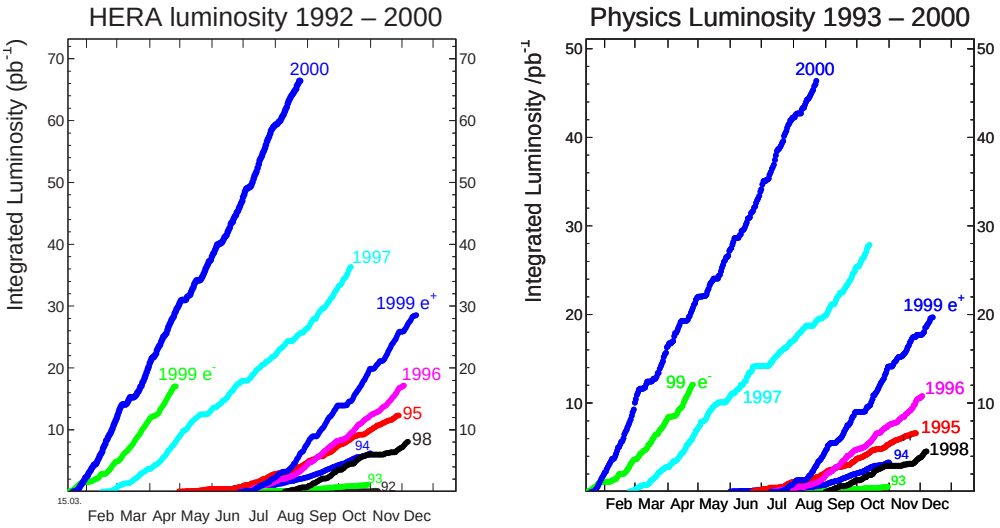


Figure 3.2: *Luminosities versus date in running periods of HERA I . The left one is of HERA delivered and the right used for ZEUS physics.*

hence upwards, as shown in Figure 3.4. The nominal *ep* interaction point (IP) is at $x = y = z = 0$.

The design of the ZEUS detector emphasizes optimum calorimetry for hadron energy measurements, hence a uranium-scintillator sampling calorimeter (CAL) with equal response to electrons and hadrons is implemented. Additional calorimetric detector components such as HES (Hadron Electron Separator) and BAC (Backing Calorimeter) improve the calorimetry as well as particle identification. The tracking detectors comprise a central tracking detector (CTD), a forward detector (FDET), which straw tube trackers were added into after the upgrade, and a rear tracking detector (RTD). There was a vertex detector (VXD) at the beginning, but not really used and taken out very soon; a micro vertex detector (MVD) was installed however during the upgrade. The CTD lives in a magnetic field of 1.43 T which is generated by a superconducting solenoid mounted between CTD and calorimeter. Muon detectors have inner and outer chambers as well as forward, barrel and rear parts for each (F/B/RMUI and F/B/RMUO).

The ZEUS detector has undergone continuous changes, like upgrades, or some components being removed or simply not used. The detector components related to this thesis are reviewed in this section. Table 3.2 summarizes the principle parameters of them. For more details and descriptions of other components, the reader is referred to the status report[51].

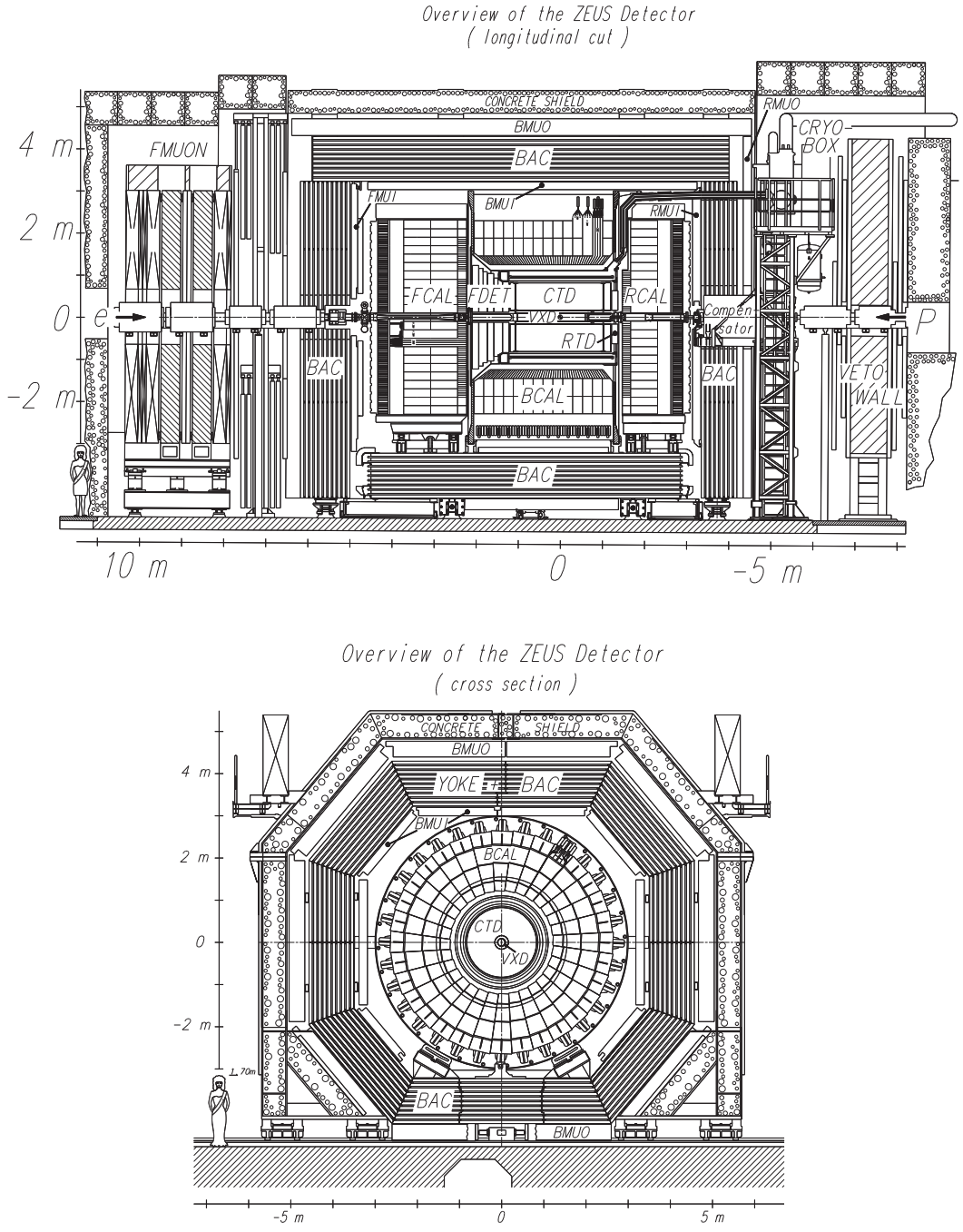


Figure 3.3: *Cross sectional views of the ZEUS detector. The top one is the Z-Y view and the bottom one X-Y view.*

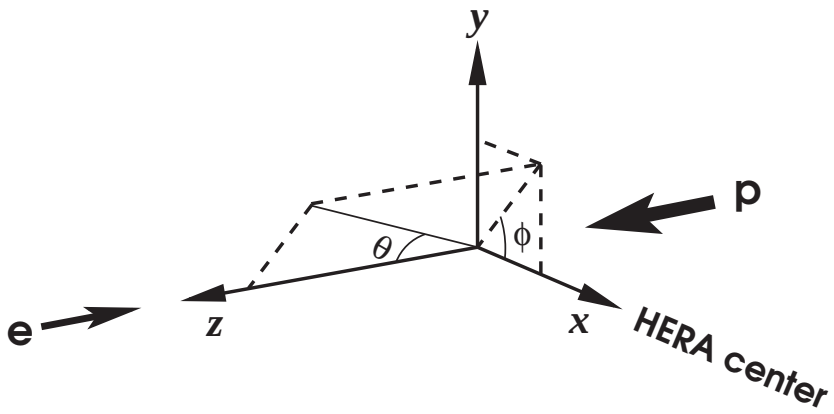


Figure 3.4: *The ZEUS coordinate system. Arrows with e and p indicate the directions of electron and proton beams, respectively. Two angles, θ for the polar angle and ϕ for the azimuthal angle, are also shown.*

Table 3.2: *The principle parameters of the CAL, the CTD and the luminosity detector of ZEUS The calorimeter performances were measured in test beams. CTD resolutions of transverse momentum and vertex are for full length tracks. $\oplus$ denotes quadratic summation.*

Component	Parameter	
CAL	Angular coverage	$2.6^\circ < \theta < 178.4^\circ$
	σ_E/E(EM shower)	$0.18/\sqrt{E(\mathrm{GeV})} \oplus 0.02$
	EM E scale uncertainty	1-3%
	σ_E/E(hadronic shower)	$0.35/\sqrt{E(\mathrm{GeV})} \oplus 0.03$
	hadronic E scale uncertainty	3%
	position resolution	$\sim 1\,\mathrm{cm}$
	time resolution	$< 1\,\mathrm{ns}$
CTD	magnetic field	1.43 T
	angular coverage	$11^\circ < \theta < 168^\circ$
	σ_{p_T}/p_T	$0.005 p_T(\mathrm{GeV}) \oplus 0.016$
	z-vertex resolution	0.4 cm
	R-ϕ vertex resolution	0.1 cm
Luminosity	normalization uncertainty	2.0%

3.2.1 The calorimeter

The calorimeter (CAL) is the key component of the ZEUS detector. It is a high resolution compensating calorimeter, and is nearly hermetic, covering a solid angle of 99.8% in the forward hemisphere and 99.5% in the backward hemisphere.

The calorimeter consists of depleted uranium plates as absorber layers, interleaved with plastic scintillator tiles as sampling layers. The relative thickness of the absorber layer and the sampling layer has been tuned to achieve $e/h = 1.00 \pm 0.05$, where e/h is the ratio of responses to electrons and hadrons, hence a compensating calorimeter. It provides linearity of hadronic response as well as optimum hadronic energy resolution. Test beam measurements gave the resolution for hadronic and electromagnetic (EM) showers separately:

$$\begin{aligned} \sigma_E/E &= 0.35/\sqrt{E(\text{GeV})} \oplus 0.03 \text{ for hadronic showers,} \\ \sigma_E/E &= 0.18/\sqrt{E(\text{GeV})} \oplus 0.02 \text{ for EM showers.} \end{aligned} \tag{3.1}$$

The CAL surrounds the solenoid and the tracking detectors, and mechanically divides into a forward part (FCAL), a barrel part (BCAL), and a rear part (RCAL). All three parts are in modular structure and segmented longitudinally into the electromagnetic sections (EMC) and the hadronic sections (HAC). The EMCs have nearly the same depth of $\sim 1\lambda (\sim 25X_0)$. The HACs vary in depth from $\sim 6\lambda$ in the very forward direction to $\sim 3\lambda$ in the rear. In FCAL and BCAL, where the HAC is deeper than 3λ, it is divided further into two subsections.

FCAL and RCAL have the same construction, differing mainly in depth. Each of them is made up of 24 modules, with a common width of 20 cm, and the heights arranged to form approximately a circle of 230 cm radius. The placement of modules is parallel to the y-axis. Figure 3.5 shows a three-dimensional view of an FCAL module. A module consists of non-projective towers. An EMC tower has a front face of about $5 \times 20\,\text{cm}^2$ in FCAL and $10 \times 20\,\text{cm}^2$ in RCAL. The HAC tower in both is $20 \times 20\,\text{cm}^2$. Each tower subsection is a readout unit, called a cell.

The BCAL is made up of 32 identical wedge-like modules, which are located coaxially with the beam, and each covers 11.25° in azimuthal angle. The axis of each module is tilted by 2.5° relative to a radius vector to prevent projective module boundaries. The EMC towers are projective, with $5 \times 24\,\text{cm}^2$ at the front face. The HAC towers are non-projective in θ. Except at the ends, four EMC towers are followed by one single hadronic tower.

On both sides of a module, there are separate wave length shifters for EMC and HACs. They read out the light generated in the scintillator towers

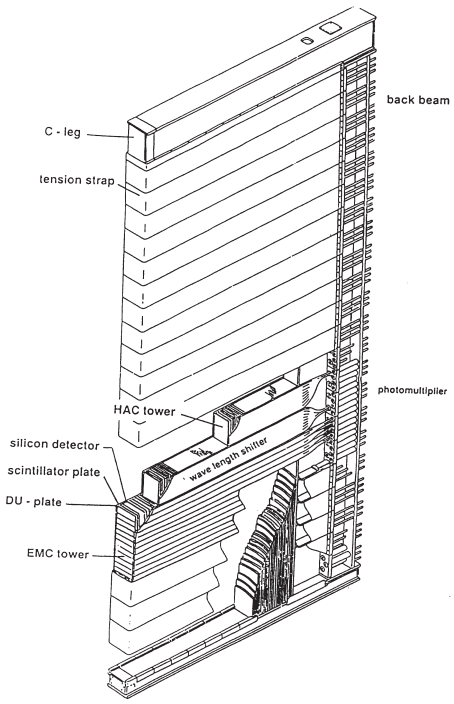


Figure 3.5: *Isometric view of the largest FCAL module.*

and transport it via light guides to photomultiplier tubes (PMTs) to form electronic pulses. The coordinate measurement can be achieved based on knowledge of the attenuation length in the scintillator.

The choice of scintillator as readout medium avoids pile up effects by keeping the pulses shorter than the bunch crossing time of 96 ns. And, due to their fast rising time, the timing of the pulses can be determined to within one nanosecond, which is important for suppressing the cosmic ray background.

The uranium radioactivity provides an extremely valuable calibration and monitoring tool. Besides the uranium activity, the calorimeter has very low noise, typically 10 MeV for an EMC PMT and 20 MeV for a HAC PMT.

The performance of the calorimeter has been measured in detail in test beams, and some results are summarized in Table 3.2. In addition, Table 3.3 summarizes some geometric parameters and the total depth for normal incidence.

3.2.2 The central tracking detector

The central tracking detector, CTD, is the most important tracking component in the ZEUS detector. It is located between the beam pipe and the

Table 3.3: *Some geometric parameters of the calorimeter.*

	FCAL	BCAL	RCAL
Polar angle $\theta(^\circ)$	2.2 - 39.9	36.7 - 129.1	128.1 - 176.5
z coordinate (cm)	221 - 451		-307 - -146
Total depth (λ) for normal incidence			
EMC	0.96 ($25.9X_0$)	1.10 ($24.6X_0$)	0.95 ($25.9X_0$)
HAC1	3.09	2.11	2.32-3.09
HAC2	1.54-3.09	2.11	–

superconducting solenoid. The axis of the chamber is parallel to the z-axis and the center is at $z = +2.5\,\text{cm}$. Its active axial coverage is 202.4 cm between end-plates ($-98.7 < z < 103.7\,\text{cm}$), with radial coverage between $R = 19.0\,\text{cm}$ of the innermost layer and $R = 78.5\,\text{cm}$ of the outermost layer[52].

The CTD is a cylindrical drift chamber composed of 9 super-layers (SLs), each with 8 layers of sense wires. Among the super-layers, the five axial super-layers, of which the sense wires are parallel to the beam axis, are interleaved with four stereo super-layers, the wires of which are strung left or right at an angle with respect to the z-axis. The stereo angle of a wire is defined to be the angle between the wire and the plane formed by the z-axis and a radial vector pointing to the center of the wire. The stereo angle of about 5° is chosen to achieve roughly the same resolutions in the polar and azimuthal angle measurements. Within each super-layer, sense wires as well as other wires are arranged in cells. The cells are tilted nearly 45° with respect to the radial vector, which powerfully breaks the left-right ambiguity. Figure 3.6 shows the layout of wires of a 45° sector of the CTD.

The CTD lives in a nearly uniform solenoidal magnetic field of $B \approx 1.43\,\text{T}$. While passing the CTD, a charged particle will leave out approximately helix-like trajectory. In the approximation,

$$p_T = qB\varrho,$$

where p_T is the transverse momentum of the particle, q the charge and ϱ the curvature of the projection of the helix in the R-ϕ plane. The axial super-layers measure the positions of hits in R-ϕ plane with a resolution $\sim 200\,\mu\text{m}$ (θ dependent), which in turn provide the measurement of the curvature ϱ, hence of the transverse momentum. The best resolution is obtained with a trajectory passing all nine super-layers:

$$\sigma_{p_T}/p_T = 0.0058 p_T \oplus 0.0065 \oplus 0.0014/p_T \quad (p_T \text{ in GeV}), \tag{3.2}$$

where the first term is due to the hit position, the second and the third due to multiple scattering, in the CTD and between the interaction point and the CTD[53].

The z-coordinate can be measured by the stereo super-layers (*z-by-stereo*), or by the time difference (*z-by-timing*). The former has the z-resolution $\sim 1.5\,\mathrm{mm}$ and the latter $\sim 5\,\mathrm{cm}$. With z-coordinates, the three-dimensional helix can be reconstructed and the momentum of a charged particle is obtained by

$$|\mathbf{p}| = \frac{p_T}{\sin\theta},$$

where θ is the polar angle determined from z-coordinates.

In addition to the momentum measurement, the CTD also provides a dE/dx measurement. Although further elaborate correction is still needed [54], this information is very useful in particle identification.

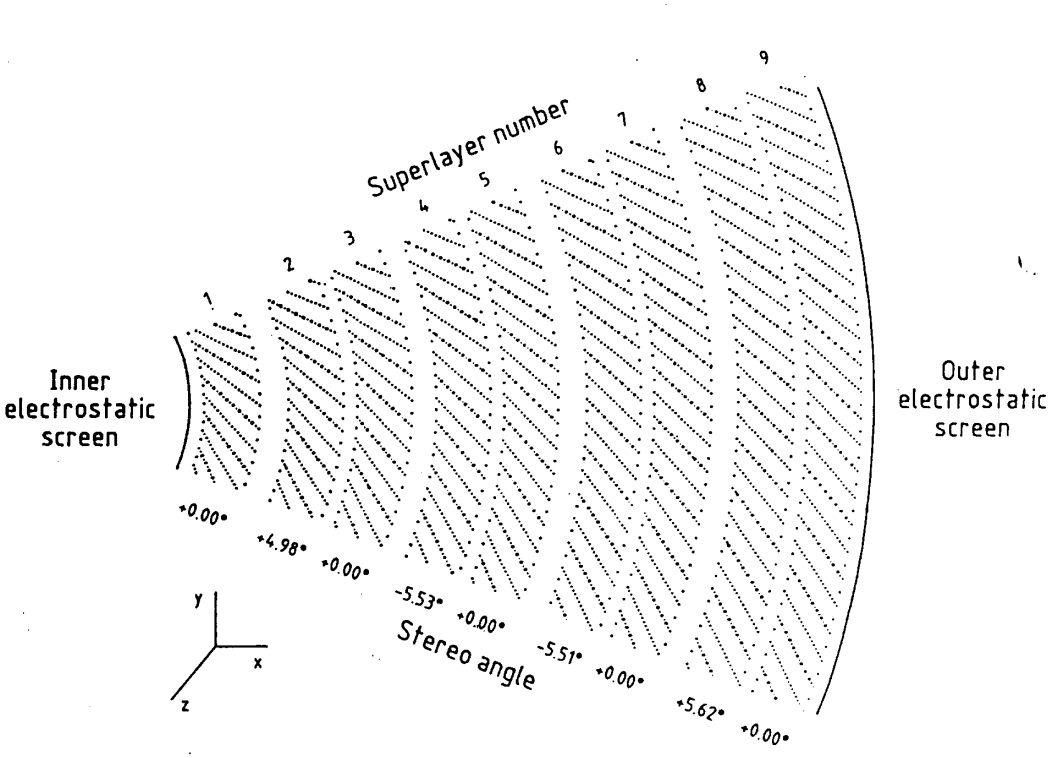


Figure 3.6: *Layout of wires in a* 45° *sector of the CTD.*

3.2.3 The luminosity monitor

The luminosity is measured through the QED bremsstrahlung or Bethe-Heitler [55] process

$$ep \to ep\gamma. \tag{3.3}$$

The luminosity monitor consists of a photon detector, Lumi-γ, and an electron detector, Lumi-e, as shown in Figure 3.7. The photon detector, measuring the photon energy as well as the position, consists of a 1.5 mm thick copper-beryllium exit window from the proton beam pipe at $z = -92.5$ m, a filter to absorb a large flux of direct synchrotron radiation and a $24X_0$ deep lead-scintillator electromagnetic calorimeter at $z = -106$ m. The electron detector, measuring the energy of the electron in coincidence with the photon, serves for consistency check but is not used for the luminosity measurement directly. The systematic uncertainty is achieved 1% for 1996 data and around 2% for other data taking periods [56], seeing Table 3.1 for details.

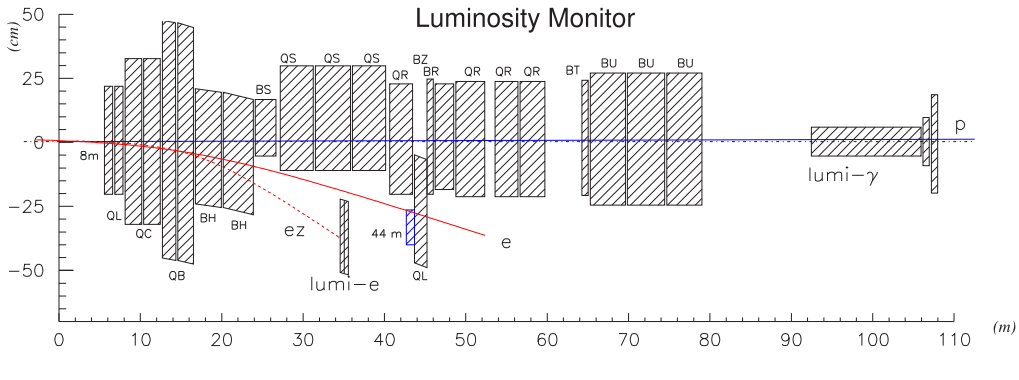


Figure 3.7: *The luminosity monitor.*

3.3 The trigger and data acquisition system

The ZEUS online system is a very sophisticated three-level trigger and data acquisition system.

At the HERA accelerator, the time between beam crossings is 96 ns, or the crossing frequency is 10 MHz. The highest activity in the detector arises from beam-losses, along the beam pipes, with a rate of order of 100 kHz/m, beam-gas interactions, with a rate of 10 kHz/m, and cosmic rays, with a rate of 5 kHz. The rate of beam-beam, namely ep, interactions is comparatively low, among which the quasi-real photoproduction is dominant, and other physics processes only contribute with 1 Hz or less. Therefore, a highly efficient

trigger system is desired. The ZEUS three-level trigger system as well as the event builder are described briefly below. A schematic flow chart is shown in Figure 3.8.

The first level trigger (FLT) reduces the rate of background to below to 1 kHz while maintaining good physics acceptance. The time for making a trigger decision is within 52 beam crossings, corresponding to 5 μs. In order to make a decision on every crossing, FLT implements pipelines to buffer the data before a decision can be made. Local decisions are first made based on energy deposits and tracks found in individual detector components and then a global decision is formed. Detector components taking part in the FLT decision are the CAL, the CTD and the muon detectors.

The second level trigger (SLT) reduces the rate by a factor of 10, to come down from a FLT trigger rate of 1 kHz to a trigger rate of 100 Hz. It is built of a transputer network. Instead of using only a subset of data at the FLT, all the data from participating detectors are available at this level. More precise quantities as well as some new ones, e.g. calorimeter timing, are used to reduce background. Results from local SLT components, like FLT, are collected to the global second level trigger (GSLT) to make a trigger decision. Physics filters have also been defined at this level [57].

The event builder (EVB) is a powerful processor farm. When a positive decision is made at SLT, it assembles the data from all detector components to form an event and passes the event to the next trigger level for the final decision.

The third level trigger (TLT) reduces the rate from SLT by a factor of 20 to about 5 Hz. Each event is fully reconstructed with a software similar to the one for the offline reconstruction. Physical quantities of the event, such as kinematic variables, output of electron finders, topology of the hadronic final state, and so on, are used to decide upon recording the event.

In addition to physics related trigger configuration at each level, there are special ones to check for environmental and test triggers, to pass a fraction of events (passthrough) as well as to run veto algorithms. Raw data of events, together with electronic calibration constants, monitoring data of running conditions and all trigger data used in local processing, are recorded permanently. They will be used in the full offline reconstruction for physics analyses.

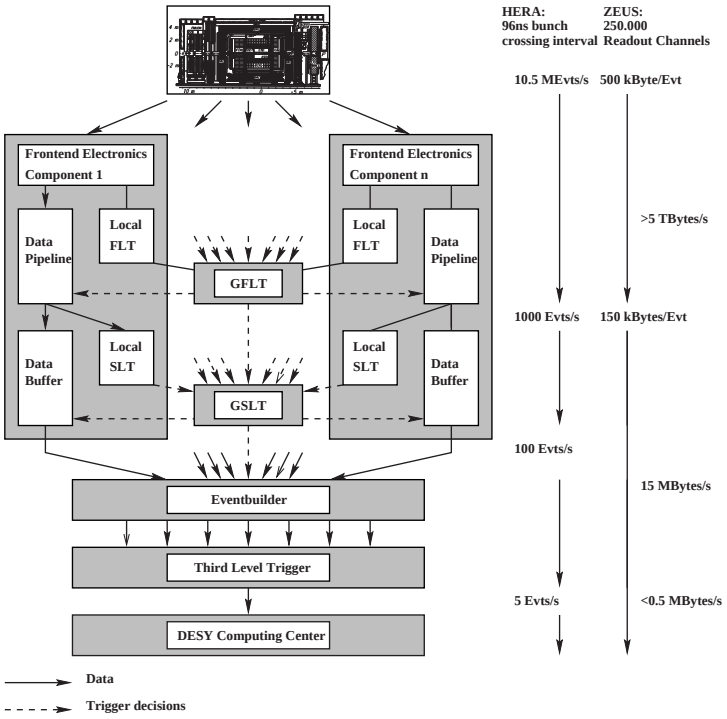


Figure 3.8: *Flow chart of the ZEUS trigger and data acquisition system.*

3.4 The offline system

The main tasks of the offline system are the event reconstruction and the improvement of the measurement by a variety of corrections. The offline reconstruction, of which the program package is called ZEPHYR [58], is processed on a PC farm running Linux as operating system. During reconstruction, events are also classified by assigning DST (data summary tape) bits according to certain physics characteristics. The DST bits are mainly based on a logical combination of trigger bits, extended by some offline requirements. These DST bits provide offline analyses with the advantage of fast access to a given class of events.

Chapter 4

Event reconstruction

Aspects of offline reconstruction tightly related to the analysis are discussed in this chapter. It is particularly crucial for charged current events to reconstruct accurately the hadronic final state, which is described in Section 4.1. As for D^* identification, good tracks as well as the primary vertex are demanded. The relevant description is given in Section 4.2. Reconstruction of DIS kinematics is important for the cross section measurement. Various methods are discussed in Section 4.3.

4.1 The hadronic final state

Within the ZEUS detector, the scattered neutrino of the charged current reaction can not be detected. Therefore, only the hadronic system will be present in the final state. The CAL is the key component for reconstructing the hadronic final state. Many efforts have been devoted to improve its energy measurement, from which important event variables are evaluated.

4.1.1 Global event variables

Although CAL is nearly hermetic, there are unavoidable holes for passing through the beam pipe. The proton remnant, concentrating mostly in the very forward direction, will cause a large uncertainty in the measurement of the total energy. Contrarily, the energy measurement in the transverse plane can be achieved to high accuracy. Therefore, the energy measurement with ZEUS is generally considered in two, longitudinal and transverse, directions.

For a CC event, the most important variable is the missing transverse momentum, $\not{p}_T$, caused by the escaping energetic final-state neutrino. $\not{p}_T$ is balanced by the net transverse momentum, p_T, of the final state observed in

the detector. They are calculated as

$$\not{p}_T{}^2 = p_T^2 = p_x^2 + p_y^2 = \left(\sum_i E_i \sin\theta_i \cos\phi_i\right)^2 + \left(\sum_i E_i \sin\theta_i \sin\phi_i\right)^2. \quad (4.1)$$

The sum runs over all calorimeter energy deposits E_i, where θ_i and ϕ_i are corresponding polar and azimuthal angles measured from the interaction vertex. The energy deposits are uncorrected in the trigger, but corrected in the offline analysis. They can be in objects of the calorimeter cells, clusters or islands, as discussed later. The hadronic polar angle is defined by

$$\cos\gamma_h = \frac{p_T^2 - \delta^2}{p_T^2 + \delta^2}, \quad (4.2)$$

where

$$\delta = \sum_i (E_i - E_i \cos\theta_i) = \sum_i (E - p_z)_i. \quad (4.3)$$

In general, p_T and δ in (4.2) should also take a subscript h to denote the hadronic final state, but the subscript can be left out in the case of CC. In the naïve quark-parton model, γ_h gives the angle of the struck quark. The variable δ gives a better longitudinal energy measurement than the total energy E or z-component momentum p_z does individually. With ideal detection, δ would be twice the positron beam energy, $\delta = 55\,\text{GeV}$, following from energy-momentum conservation. Undetected particles escaping through the forward beam hole, e.g. the proton remnant, change δ negligibly while particle loss through the rear beam hole, e.g. the outgoing positron in photoproduction, can lead to a substantial reduction of δ. Another variable used in the event selection is the total transverse energy, E_T, given by

$$E_T = \sum_i E_i \sin\theta_i. \quad (4.4)$$

p_T and δ are used for reconstructing DIS kinematic variables, as discussed later.

4.1.2 Energy corrections

Before proceeding with the determination of p_T, δ and E_T, a number of corrections are applied to CAL's energy measurement. Methods developed and tested for the NC cross section determination [35] have been used for the case of CC [36]. A brief description is given here.

Noise suppression: The natural radioactivity of the depleted uranium produces a constant level of noise in the CAL cells. While energy cuts, 60 MeV for EMC and 110 MeV for HAC cells, have been applied online, stricter cuts, 100 MeV and 150 MeV for EMC and HAC cells, are applied offline.

Dead material: Energy losses in the inactive material between the interaction point and the calorimeter are corrected by using dead material maps, which contain a detailed description of the detector layouts for the different data taking periods.

Energy scale: Detailed studies of energy measurements and simulated detector responses disclose the under-measurement of energy. Absolute energy scales are applied to different parts of the calorimeter as well as for different type of particles, i.e. electrons and hadrons.

Non-uniformity: There are dips and peaks in the energy responses at the boundaries of the calorimeter cells and modules. The effects are different in data and Monte Carlo simulation and are corrected accordingly.

Detector alignment: The positions of the calorimeter parts changed slightly for different data taking periods. The effects have been taken into account as well.

4.1.3 Calorimeter islands

The resolution of energy measurement is inversely proportional to the square root of the energy, see (3.1). Therefore, individual energy deposits in cells are clustered into relatively large objects to improve the energy measurement as well as the angular information of the calorimeter. For CC analyses, the objects are usually *islands* with the method developed for ZUFOs (ZEUS unidentified flow objects) [59]. The algorithm works in two steps, as shown in Figure 4.1. First, adjacent cells in EMC, HAC1 and HAC2 sections are clustered into two-dimensional objects, *cell islands*, based on local maxima of energy deposits. The resulting cell islands are then joined into three-dimensional objects, *cone islands*, using a probability function depending on the angular separation. The position of a cone island is determined by the logarithmic center-of-gravity of the shower.

The resulting *islands* undergo further corrections for *backsplash* as well as redirected hadronic energy from interactions in material between the primary vertex and the calorimeter. High energy hadrons interacting in the calorimeter with a non-negligible probability can produce particles at large angles

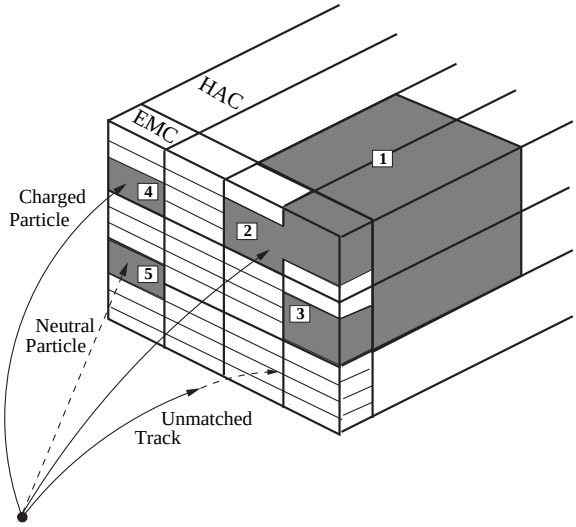


Figure 4.1: *Schematic plot of the island clustering algorithm. Neighbouring calorimeter cells are clustered into cell islands. One HAC (1) and four EMC (2-5) cell islands are shown in the plot. Cell islands, island 1, 2 and 3, are joined into cone islands. Refer to the text for details.*

with respect to the direction of the main shower. Some of these particles travel backwards and generate energy deposits far away from their original source in the calorimeter which is referred to as backsplash. MC studies indicate that γ_h or δ is biased by such kind of energy deposits. The bias can be minimized by removing islands with energy below 3 GeV and with a polar angle larger than $\gamma_{\max}$, which is a function of γ_h derived from a MC sample. After the island removal, γ_h as well as $\gamma_{\max}$ are recalculated, and the procedure is repeated until the relative change of γ_h is below 1%.

This correction as well as some energy corrections described previously are implemented in a specific software package, `CorAndCut` [60]. DIS kinematic variables are then calculated with the Jacquet-Blondel method as described in Section 4.3. Furthermore, the Jacquet-Blondel variables, y_{JB} and Q^2_{JB}, are regularized by taking account of two constraints:

$$\begin{aligned} y_{\mathrm{JB}} &= \frac{\delta}{2E_e} < 1, \\ Q^2_{\mathrm{JB}} &= \frac{p_T^2}{1 - y_{\mathrm{JB}}} < s = 4E_eE_p, \end{aligned} \tag{4.5}$$

where E_e and E_p are the energy of positron and proton beams. The package is used for the reconstruction of the hadronic final state and DIS kinematics in this analysis.

4.1.4 Verification

The improvement in the reconstruction of the hadronic final state variables by applying the correction can be verified with Monte Carlo. The comparison of the relative resolution in p_T with and without the correction is shown in Figure 4.2. Although the whole resolution with the correction is slightly worsened, the bias is significantly reduced.

4.2 Tracking and vertexing

As for HERA I data, the offline reconstruction performs track finding twice per event. The first pass, *CTD-only*, uses solely information from CTD, and the second pass, *regular*, uses information from other tracking detectors, such as FTD and RTD, as well. Which detectors are used in *regular* mode depends on when the data was taken, and on the status of the reconstruction program. To reduce uncertainties in corrections for efficiency and acceptance, the *CTD-only* mode is used for this analysis. Besides this consideration, there are several other advantages:

- CTD participated in the online trigger in contrast to other tracking detectors;
- Tracking in the CTD is relatively clean;
- There is no bias due to relative alignment with other detectors;
- CTD pattern recognition is not confused by large scatters or secondary charged particle production in the material between the CTD and the other tracking detectors.

A sophisticated algorithm is used to find tracks in the CTD [61]. For a brief description, the track reconstruction is completed in three steps.

1. The first step is *pattern recognition* by starting each track candidate from a "seed" in an outer part of the CTD and extrapolating inward. A very broad "virtual hit" is added at the beam line to guide the trajectory inward.
2. The second step is CTD *hit corrections* for the Planar Drift Approximation[1] in the position measurements of hits. Many effects, e.g. angle of track from drift plane (ψ'), are taken into account. The effect on the

[1] The plane, centered on a sense wire, is by definition perpendicular to the line connecting $x = y = 0$ to the cell's central ground wire.

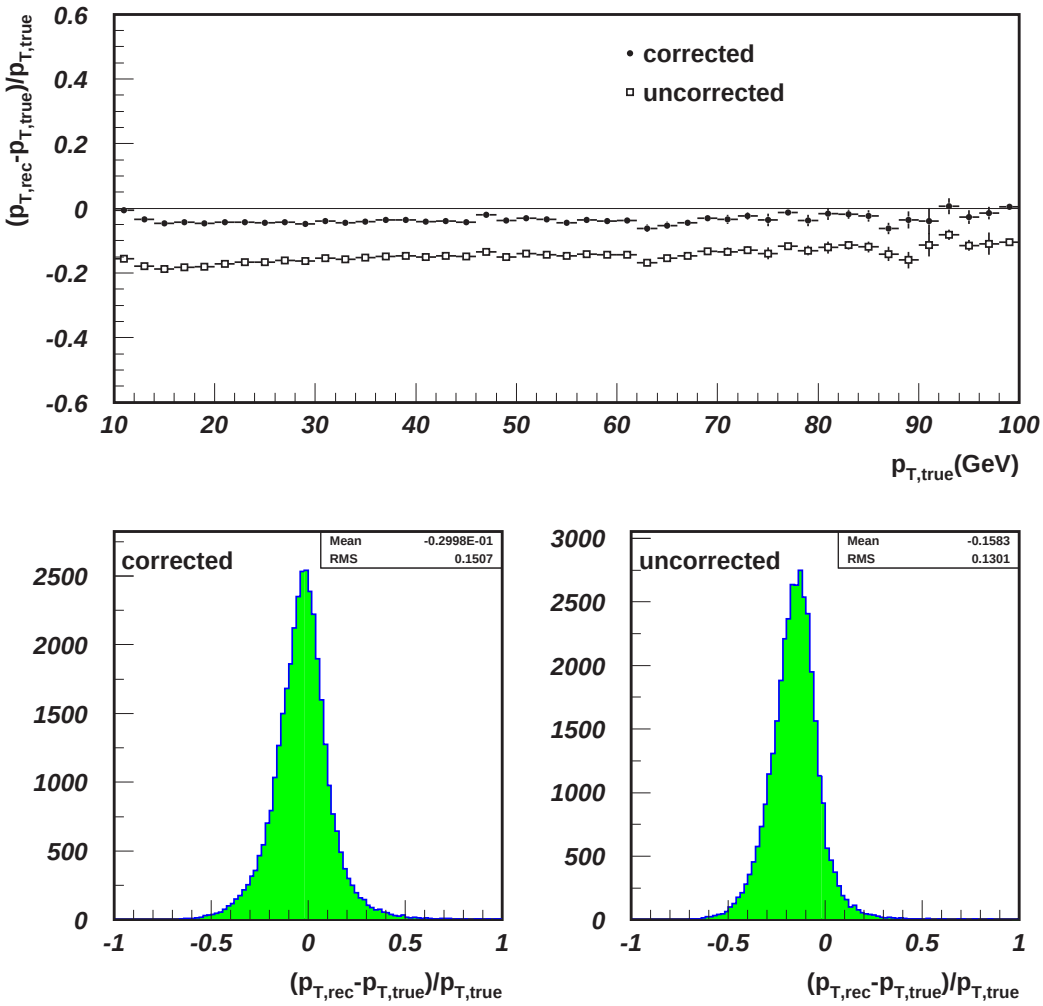


Figure 4.2: *Comparison of the relative resolution in p_T with and without energy correction. In the upper plot, the resolution is shown as a function of true transverse momentum, $p_{T,true}$, for corrected, solid dots, and uncorrected, open squares, p_T. In the lower part, the whole resolution with the correction, left, and without the correction, right, are shown. The signal Monte Carlo sample is used for the comparison.*

final track parameters is usually relatively small, but a large effect happens for positive tracks with transverse momentum less than 200 MeV, where the track becomes parallel to the plane of sense wires.

3. The third step is to *fit* each track candidate to a 5-parameter helix model. The fit procedure is crude but fast, and one iteration suffices since the *pattern recognition* is fairly accurate.

An example of a helix in the X-Y plane is schematically shown in Figure 4.3.

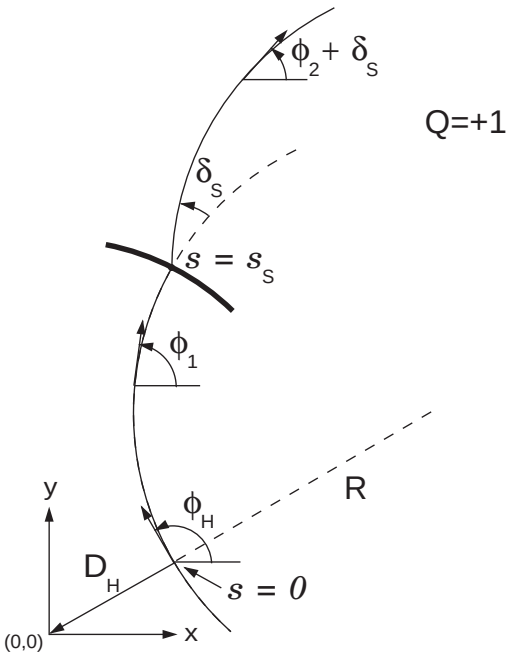


Figure 4.3: *A track helix in x-y plane. Definitions of the variables can be found in the reference [61].*

Primary and secondary event vertices are reconstructed by fitting tracks based on the simple and full fit perigee techniques [62]. An imagined proton, centered at the beam spot in the X-Y plane with $\sigma_x = \sigma_y = 0.7\,\text{cm}$, is used to constrain the primary vertex. Technical details are described in the references [61]. Primary-vertex tracks then take advantage of simultaneously refitting both direction and curvature.

4.3 DIS kinematics

As discussed in Section 2.1, DIS kinematics are specified with two variables at fixed center-of-mass energy

$$\sqrt{s} \approx 2\sqrt{E_e E_p}, \tag{4.6}$$

where E_e and E_p are the energy of positron and proton beams, and masses of the positron and the proton are ignored. Many methods have been developed to calculate DIS kinematics by reconstructing the leptonic final state, or the hadronic final state, or both. Several methods are reviewed here.

Electron method (EL): using solely the leptonic final state,

$$Q^2_{\mathrm{EL}} = 2E_e E'_e(1+\cos\theta'_e), \tag{4.7}$$

$$y_{\mathrm{EL}} = 1 - \frac{E'_e}{2E_e}(1-\cos\theta'_e), \tag{4.8}$$

with E'_e and θ'_e the energy and the polar angle of the scattered positron. This is the simplest method and has traditionally been used in fixed-target experiments. It is, however, only applicable for NC events. The method suffers from large radiative corrections and a seriously degraded x resolution at small y. The resolution is, however, very good at large y.

Double angle method (DA): using the measured angles of both leptonic and hadronic final states,

$$Q^2_{\mathrm{DA}} = 4E_e^2 \frac{\sin\gamma_h(1+\cos\theta'_e)}{\sin\gamma_h + \sin\theta'_e - \sin(\theta'_e+\gamma_h)}, \tag{4.9}$$

$$x_{\mathrm{DA}} = \frac{E_e \sin\gamma_h + \sin\theta'_e - \sin(\theta'_e+\gamma_h)}{E_p \sin\gamma_h + \sin\theta'_e - \sin(\theta'_e+\gamma_h)}, \tag{4.10}$$

with γ_h defined in (4.2). This method does not require precise knowledge of energy scales, and results in small radiative corrections. However, the resolution is poor at very small y. Again, the method is applicable only for NC events.

Hadron method: also known as **Jacquet-Blondel (JB)** method [63], using the measurement of the hadronic final state,

$$y_{\mathrm{JB}} = \frac{\delta_h}{2E_e}, \tag{4.11}$$

$$Q^2_{\mathrm{JB}} = \frac{p^2_{T,h}}{1-y_{\mathrm{JB}}}, \tag{4.12}$$

with δ_h and $p_{T,h}$ defined in (4.3) and (4.1) for the hadronic final state. This is the only possible method for CC events. The method is stable against energy losses down the forward beam pipe since they contribute very little to δ_h and $p_{T,h}$. However, it is sensitive to calorimeter noise at very small y, and is sensitive to energy losses in the rear direction at high values of y. It also requires a good understanding of energy scales and energy losses in the inactive material. On the other hand, it is rather insensitive to radiative corrections.

The Jacquet-Blondel method implemented in `CorAndCut` is used for this analysis. It is clear from (4.12) that Q^2_{JB} is directly correlated with p_T. Any cut on p_T in the event selection will automatically restrict the kinematic region of events of interest. A kinematic plane is shown in Figure 4.4 with lines of iso-y and iso-p_T.

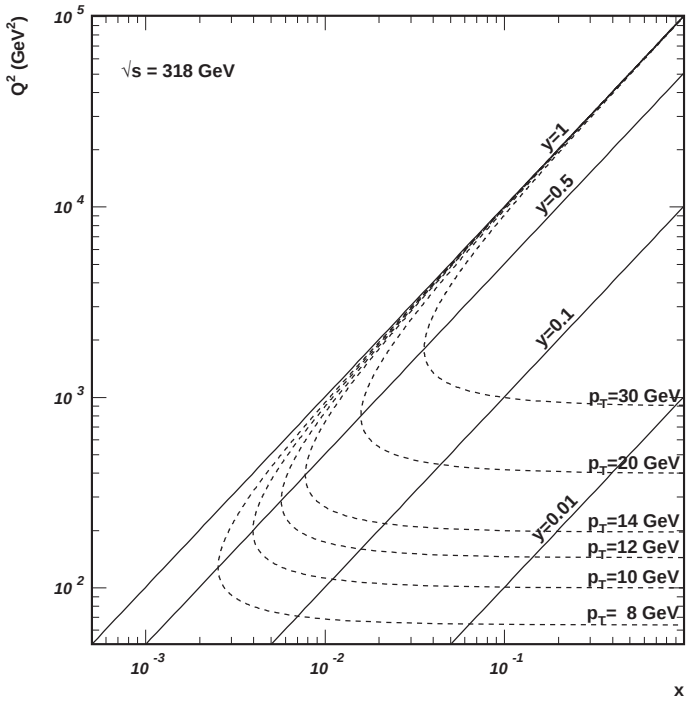


Figure 4.4: *DIS kinematic plane in x and Q^2 with lines of iso-y and iso-p_T. Contour lines are calculated at $\sqrt{s} = 318$ GeV and those of p_T are obtained according to the Jacquet-Blondel method, as discussed in the text.*

Chapter 5

Monte Carlo simulation

Monte Carlo (MC) simulation is used to determine the efficiency for selecting events and the accuracy of kinematic reconstruction, to estimate the *ep* background rates and to extract cross sections for the full kinematic region. MC samples generated for these purposes as well as some techniques in generation are discussed in Section 5.1. The MC samples are normalized to the total integrated luminosity of the data that is discussed in Section 5.2. As a direct application, the resolution of DIS kinematic reconstruction is checked in Section 5.3.

5.1 Generation

The common interface for all MC generators used at ZEUS is AMADEUS. Generated events are submitted to a facility, called FUNNEL, for full detector simulation. The ZEUS detector response is simulated with a program, called MOZART, based on GEANT 3.13 [64]. The simulated events are subjected to the same trigger requirements as the data, and processed by the same reconstruction programs. The ZEUS MC chain is shown in Figure 5.1. More details can be found on the Web [65].

Various MC samples, as listed in Table 5.1, have been generated for different purpose. Inclusive CC and NC DIS as well as photoproduction MC samples are used to study the separation of CC events from NC and photoproduction background. To increase statistics at high Q^2, CC and NC events were simulated in sets of samples with different Q^2 thresholds. Both direct and resolved photoproduction samples were generated. A signal MC sample is used to determine the acceptance as well as to study D^* finding. It has been generated in the same configuration as the inclusive CC MC samples with the additional requirement that at least one charged D^* exists within a

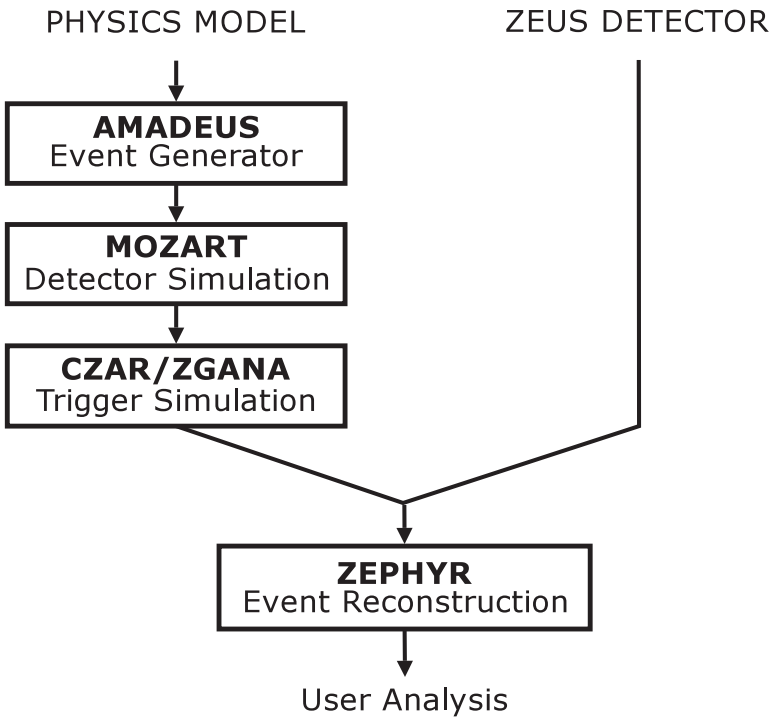


Figure 5.1: *The ZEUS Monte Carlo simulation chain.*

certain kinematic region in the final state. Another similar sample, differing only in the simulation of high order QCD effects, is used to estimate the uncertainty of MC modeling. A sufficient number of events were generated for different samples separately to ensure that the statistical uncertainties arising from the MC simulation are negligible compared to those of the data. Some details are discussed in the following.

5.1.1 Inclusive CC and NC DIS MC samples

DJANGOH 1.1 [66] is used to simulate charged current and neutral current deep inelastic *ep* scattering including QED and QCD radiative effects at the parton level. It is an interface of the MC generators HERACLES 4.6.1 [67] and LEPTO 6.5 [68]. The Lund string model implemented in JETSET 7.4 and PYTHIA 5.7 [69] is used for a complete event simulation including fragmentation and hadronization of the scattered quark and the proton remnant. In HERACLES, appropriate radiative corrections, initial-state radiation for CC and initial- and final-state radiation and Compton scattering for NC, and weak virtual corrections are included. The QCD corrections for the naïve parton model are simulated using the color-dipole model as implemented in

Table 5.1: *Monte Carlo samples. Details about generator are discussed in the text. Only difference in generators for D^* CC MC samples are listed. Integrated luminosities are used for MC normalization.*

Sample	Generator	Threshold	N_{event}	$\mathscr{L}(\,\text{pb}^{-1})$
D^* CC	ARIADNE	$Q^2 > 100\,\text{GeV}^2$	50000	1.469×10^4
	MEPS	$Q^2 > 100\,\text{GeV}^2$	50000	1.216×10^4
inclusive CC	DJANGOH	$Q^2 > 10\,\text{GeV}^2$	25000	553.1
		$Q^2 > 100\,\text{GeV}^2$	25000	628.5
		$Q^2 > 5000\,\text{GeV}^2$	15000	4.688×10^3
		$Q^2 > 10000\,\text{GeV}^2$	5000	7.322×10^3
		$Q^2 > 20000\,\text{GeV}^2$	5000	8.078×10^4
inclusive NC	DJANGOH	$Q^2 > 100\,\text{GeV}^2$	940k	115.7
		$Q^2 > 400\,\text{GeV}^2$	120k	102.8
		$Q^2 > 1250\,\text{GeV}^2$	50000	253.1
		$Q^2 > 2500\,\text{GeV}^2$	24000	407.4
		$Q^2 > 5000\,\text{GeV}^2$	24000	1.617×10^3
		$Q^2 > 10000\,\text{GeV}^2$	24000	8.591×10^3
		$Q^2 > 20000\,\text{GeV}^2$	24000	7.742×10^4
		$Q^2 > 30000\,\text{GeV}^2$	12000	2.204×10^5
		$Q^2 > 40000\,\text{GeV}^2$	12000	1.104×10^6
		$Q^2 > 50000\,\text{GeV}^2$	12000	5.664×10^6
direct PHP	HERWIG	$E_T > 40\,\text{GeV}$	100k	633.7
resolved PHP		$p_T > 10\,\text{GeV}$	100k	577.0

ARIADNE 4.10 [70]. The CTEQ5D [24] parton distribution functions, accessed from the PDFLIB 8.04 [71], are used for the proton. Values from the PDG [1] are used for the Fermi constant and the masses of the W and Z bosons as well as of the top quark. The Higgs-boson mass is set to 100 GeV.

As an alternative to ARIADNE, the QCD effects can also be simulated using the matrix elements and parton shower (MEPS) model of LEPTO. Actually, this was the first choice when starting this analysis and the simulation has been studied at the generation level, see Appendix A. Both models can describe data well in most aspects of the final state, e.g. p_T and $\delta = E - p_z$ distributions as shown in Figure 5.2. However, as shown in Figure 5.3, ARIADNE agrees well with data in the distribution of number of good tracks, $N_{\text{track}}^{\text{good}}$, whereas MEPS does not. Since good tracks are important in D^* finding, ARIADNE is chosen as the primary simulation for the QCD effects in the signal MC sample, and MEPS as a systematic check.

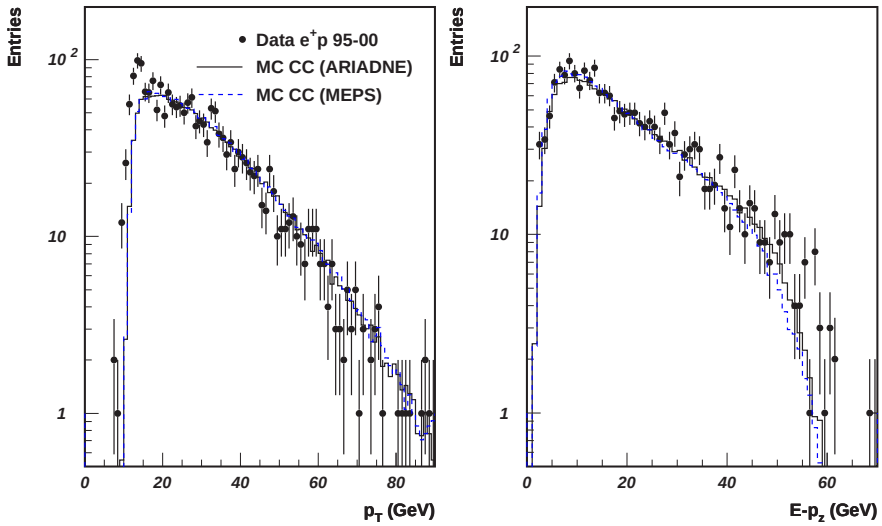


Figure 5.2: *Comparison of ARIADNE and MEPS on the distributions of p_T and $\delta = E - p_z$. All event criteria except those on p_T, y_{JB} and Q^2_{JB} are performed. The dots represent data and histograms the MC simulation for CC DIS with ARIADNE (solid) or MEPS (dashed). MC samples are normalized to the data's total integrated luminosity.*

5.1.2 Photoproduction MC samples

The dominant reaction in ep scattering is neutral current, exchanging mainly virtual photons. When the virtuality Q^2 of the exchanged boson is very small,

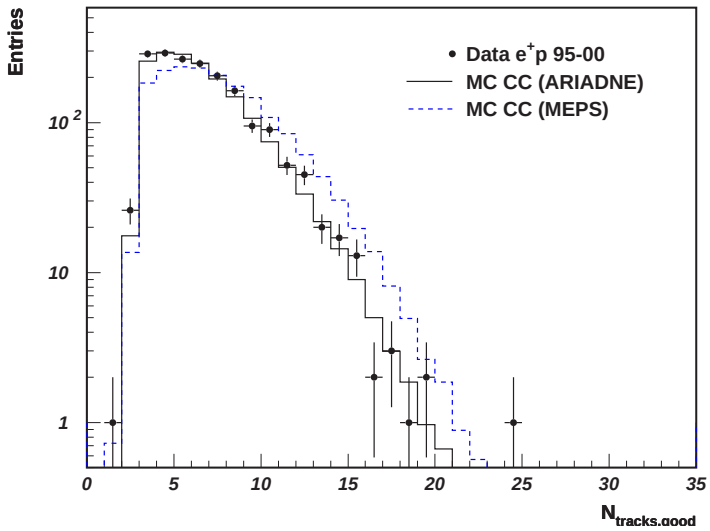


Figure 5.3: *Comparison of ARIADNE and MEPS on the distribution of number of good tracks, N_{track}^{good}. All event criteria except that on N_{track}^{good} is performed. The dots represent data and histograms the MC simulation for CC DIS with ARIADNE (solid) or MEPS (dashed). MC samples are normalized to the data's total integrated luminosity.*

$Q^2 \to 0\,\mathrm{GeV}^2$, the virtual photon resembles a real one and the ep collision is referred to as the γp process, or photoproduction (PHP). While most of the *quasi-real* photons will react on quarks directly, a certain fraction of them can be resolved with parton contents. For distinction, they are called direct and resolved photoproduction, respectively. Feynman diagrams in Figure 5.4 illustrate such processes in boson-gluon fusion.

Both direct and resolved photoproduction samples are simulated with the MC generator HERWIG 5.9 [72]. The PDFs of GRV-G [73] are used for the photon and CTEQ4L for the proton. Generator level cuts, $E_T > 40\,\mathrm{GeV}$ to all final state particles with polar angles in the range $0.0384 < \theta < 3.08$ and $p_T > 10\,\mathrm{GeV}$ to the hard process, are applied. Photoproduction contamination from lower thresholds are found to be negligible [74].

5.1.3 D^* signal MC samples

Based on the simulation of CC DIS, signal MC samples are generated by requiring the existence of at least one D^*, positively or negatively charged, in the hadronic final state. As discussed previously, ARIADNE is chosen for higher order QCD simulation, while MEPS is used as a systematic check.

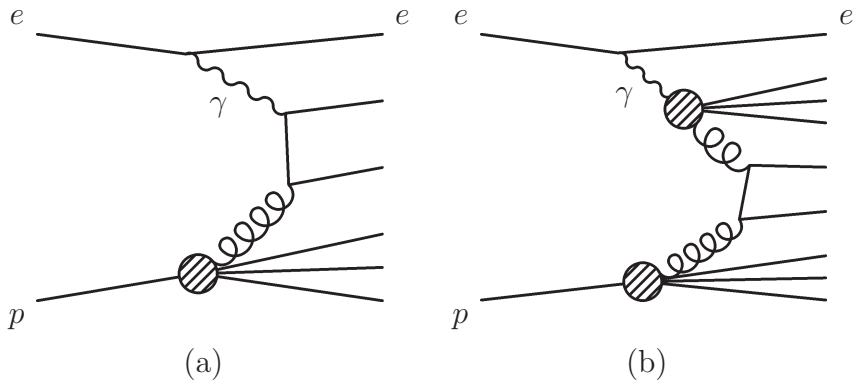


Figure 5.4: *Feynman diagrams for photoproduction in boson-gluon fusion. In direct PHP (a), the photon interacts with the gluon directly, whereas in resolved PHP (b), the photon resolves into partons before entering the interaction.*

The signal MC sample is used to determine the acceptance and to extrapolate the measurement of D^* cross section to the full kinematic region.

Distributions of true D^* kinematics, the transverse momentum $p_{T,\text{true}}^{D^*}$ and the pseudorapidity[1] $\eta_{\text{true}}^{D^*}$ are shown in Figure 5.5. In D^* reconstruction, as discussed in Section 7.1, the transverse momentum and the pseudorapidity will be restricted in a region defined by cutdstarkine. The full region distributions here will be used to obtain the extrapolation factor defined in (8.5). In the distribution of $\eta_{\text{true}}^{D^*}$, the peak at around zero corresponds to D^*s coming from hard scattering, while the plateau in the very forward direction corresponds mainly to the D^* production from fragmentation. Models of ARIADNE and MEPS, however, show very similar behavior here.

5.2 Normalization

In general, a MC sample is normalized to data according to the calculation for the cross section

$$\sigma = \frac{N}{\mathscr{L}} = \frac{N^{Data}}{\mathscr{L}^{Data}} = \frac{N^{MC}}{\mathscr{L}^{MC}}, \tag{5.1}$$

where N represents the number of events, $\mathscr{L}$ the total integrated luminosity. Superscripts denote data and MC, respectively. The normalization weight,

[1]The pseudorapidity is defined as $\eta = -\ln(\tan\theta/2)$, where the polar angle, θ, is measured with respect to the proton beam direction. The larger the pseudorapidity, the more forward it is.

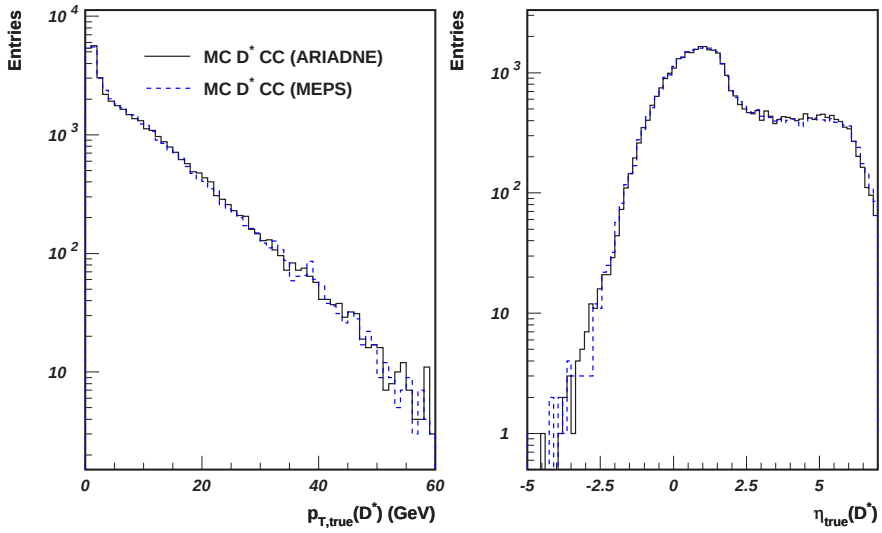


Figure 5.5: *Distributions of true D^* kinematics, $p^{D^*}_{T,true}$ and $\eta^{D^*}_{true}$. The solid histogram is for the signal MC simulation with ARIADNE and the dashed for MEPS.*

w, is then given by

$$w = \frac{\mathscr{L}^{Data}}{\mathscr{L}^{MC}}. \tag{5.2}$$

For inclusive CC or NC samples with different Q^2 thresholds, however, the overlap has to be taken into account. Given a MC event with $Q^2 = Q^2_0$, its weight is

$$w = \frac{\mathscr{L}^{Data}}{\sum_i \mathscr{L}^{MC}_i}, \tag{5.3}$$

where the denominator sums over all samples with $Q^2_{\min_i} < Q^2_0$. Therefore, in this situation, the normalization weight is Q^2 dependent and must be evaluated event by event. In Figure 5.6, the upper plot is the Q^2 spectrum without normalization for all CC MC samples in use, and the zigzagged shape can be seen with clear Q^2 thresholds; the lower plot is the one with normalization, and the spectrum becomes smooth. Indeed, it has been checked that the distribution in the lower plot has the same shape as that of the part with the lowest Q^2 threshold in the upper plot.

There is a difficulty in normalizing photoproduction samples. What really happens for photoproduction at HERA is so far not very clear. Photoproduction contamination is dependent on the process under study. Therefore, the general normalization (5.2) is not sufficient. A further normalization is

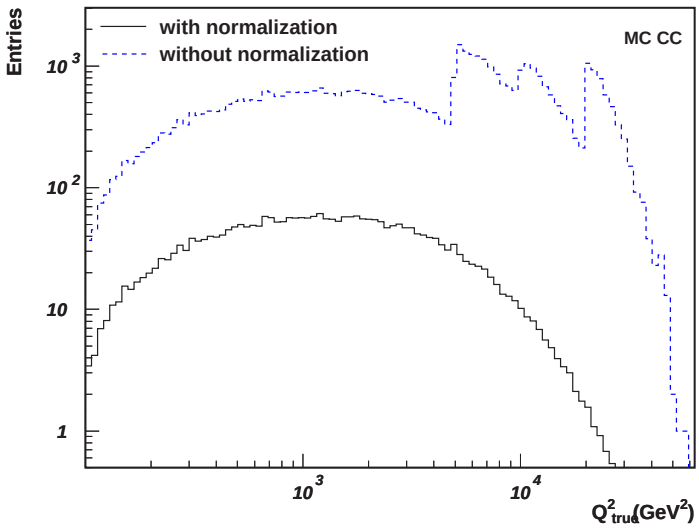


Figure 5.6: *Normalization for inclusive CC MC samples. The solid histogram is obtained with an appropriate normalization, whereas the dashed histogram is without normalization.*

needed. This is done by performing a fit,

$$H_{\text{data}} = H_{\text{CC}} + H_{\text{NC}} + w_{\text{dir}} H_{\text{dir.PHP}} + w_{\text{res}} H_{\text{res.PHP}}, \tag{5.4}$$

where H's represent distributions of some variable for data and different MC samples, indicated by subscripts, and w_{dir} and w_{res} are two free parameters. The distribution is chosen so that the shape of it in photoproduction MC samples is significantly different from those in other MC samples. The distribution of p_T/E_T is a good choice and is used for this purpose here. The fit is performed when all cuts except the one on p_T/E_T are made and the result is shown in Figure 5.7. The results for w_{dir} and w_{res} are

$$\begin{aligned} w_{\text{dir}} &= 1.3, \\ w_{\text{res}} &= 0.5, \end{aligned} \tag{5.5}$$

with a fit quality $\chi^2/\text{ndf} = 70.7/47 \approx 1.5$. The two weights are then fixed for normalization of direct and resolved photoproduction through the whole analysis.

5.3 Resolution of DIS kinematics

Using MC samples, the resolution in the reconstruction of the DIS kinematics can be obtained by comparing measured quantities with the *true* values,

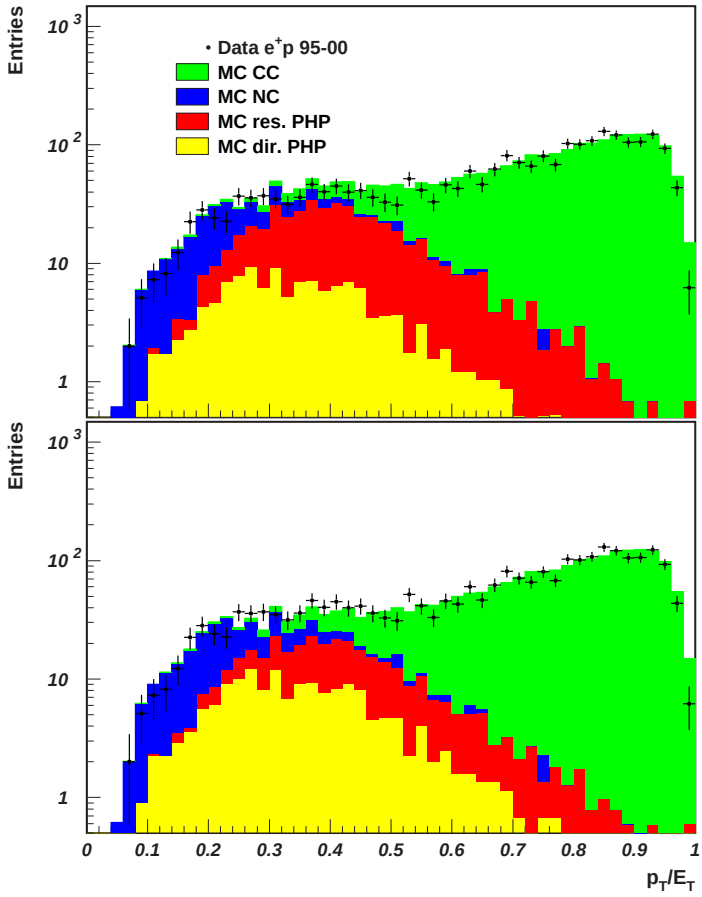


Figure 5.7: *Normalization for photoproduction MC samples. The upper plot is before performing a fit as discussed in the text, and the lower one after a fit. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

which are accessible within the MC sample. The DIS kinematics are reconstructed with the Jacquet-Blondel method, see Section 4.3.

Relative resolutions of Q^2_{JB}, y_{JB} and x_{JB} are shown in Figures 5.8, 5.9 and 5.10. Each is shown as a function of corresponding true values as well as in a overall distribution. Except in extreme regions, $Q^2 \gtrsim 10^4\,\mathrm{GeV}^2$, $x \lesssim 0.01$ and $x \gtrsim 0.3$, the JB method gives very good resolutions, indicated by error bars in y-direction. There are slight biases in Q^2_{JB} and y_{JB} that is consistent with the bias in p_T, shown in Figure 4.2, whereas the distribution of x_{JB} is centered at zero perfectly. The overall resolutions, RMS' as read from each histograms, are separately

$$\sigma_{Q^2}/Q^2 \approx 30\%, \tag{5.6}$$
$$\sigma_y/y \approx 19\%, \tag{5.7}$$
$$\sigma_x/x \approx 26\%. \tag{5.8}$$

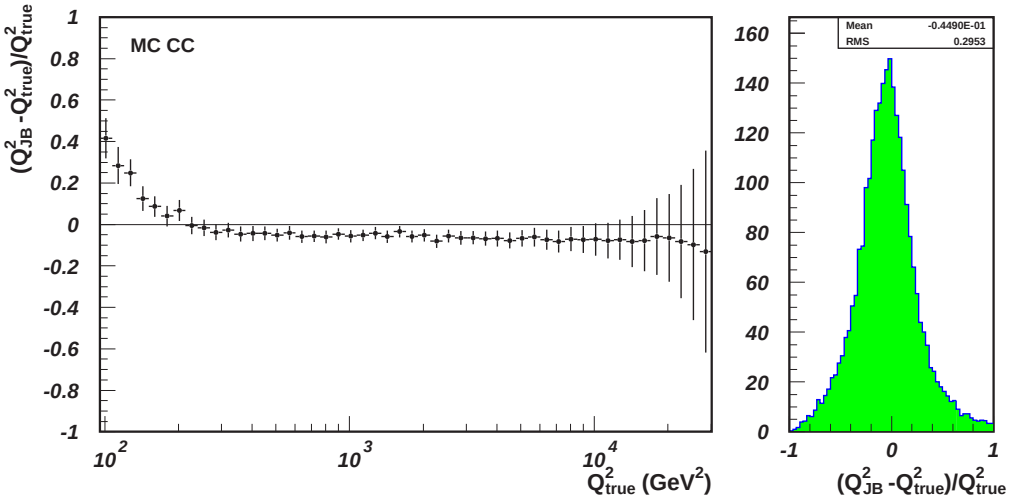


Figure 5.8: *Resolution of reconstructed Q^2. In the left, the relative difference of reconstructed Q^2 to the true value is shown as a function of the true values, where the dots are for mean differences and the vertical error bars (RMS) for the relative resolutions. In the right, the distribution of overall relative differences is shown, where the RMS corresponds to an overall resolution.*

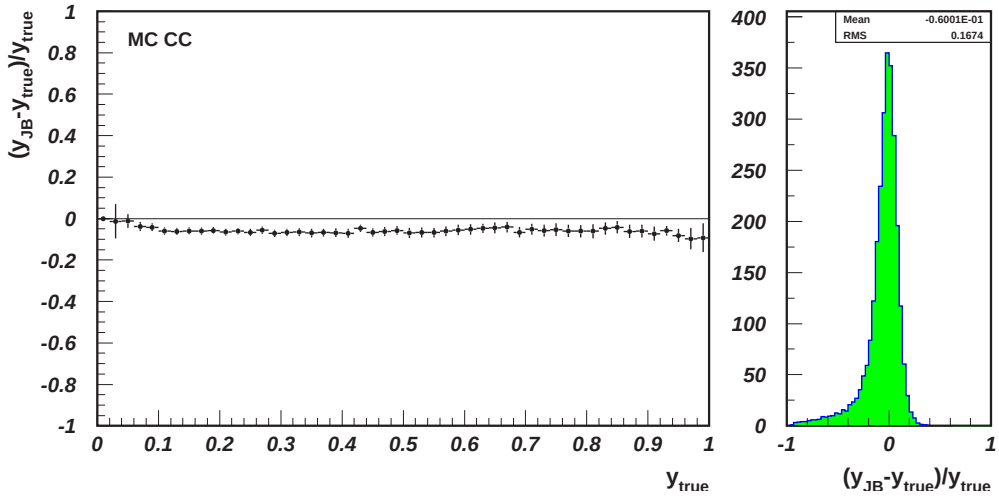


Figure 5.9: *Resolution of reconstructed y. In the left, the relative difference of reconstructed y to the true value is shown as a function of the true values, where the dots are for mean differences and the vertical error bars (RMS) for the relative resolutions. In the right, the distribution of overall relative differences is shown, where the RMS corresponds to an overall resolution.*

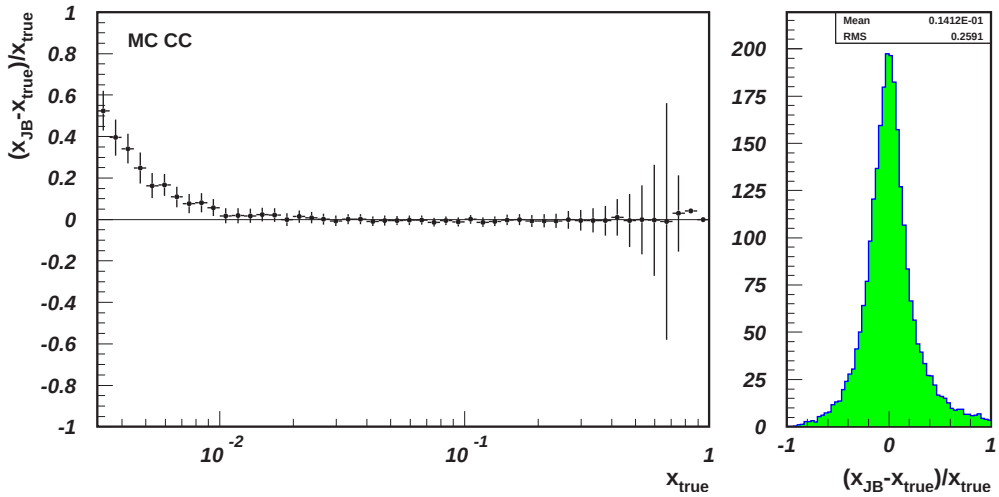


Figure 5.10: *Resolution of reconstructed x. In the left, the relative difference of reconstructed x to the true value is shown as a function of the true values, where the dots are for mean differences and the vertical error bars (RMS) for the relative resolutions. In the right, the distribution of overall relative differences is shown, where the RMS corresponds to an overall resolution.*

Chapter 6

Charged current event selection

The physics process under study in this thesis is charm production, indicated by D^* production, in CC reactions, hence the two phases—selecting CC events and identifying D^* in the final states. Naturally, there are two approaches to accomplish the mission: filtering CC events first, and then finding the D^* signal in them; or the other way around. The former approach is taken for this analysis. It is due to the choice of the trigger configuration, which is usually the first step for any offline analysis. While the CC trigger configuration had been relatively stable for nearly all HERA I running periods, the open charm trigger configuration had been designed mainly for photoproduction as well as the NC process, and that for CC is still under study. This and the next chapters discuss details in selecting CC events and identifying D^*'s, respectively.

The chapter starts in Section 6.1 with discussion of data and MC samples. It is followed in Section 6.2 by a description of online trigger related to CC events. Data quality and preselection for the data sample are then discussed in Section 6.3. The primary CC event selection is discussed in Section 6.4. Discussion of vertex and tracking is in Section 6.5. The two following Sections 6.6 and 6.7 concentrate on suppressing non-ep collisions and non-CC ep collsion background. Then, constraints on DIS kinematics are expressed in Section 6.8. In the end, Section 6.9 is discussion on the final CC sample.

6.1 Data and MC samples

To maximize statistics, all e^+p collision data collected with the ZEUS detector at HERA I from 1995 to 2000 are used for this work. The MC samples used for optimizing CC event selection are listed in Table 5.1. They are simulated according to the online trigger configuration for the running period

of 1999. Trigger configuration related to CC events for 2000 is almost the same as that for 1999, and is slightly different from those for other years. To minimize systematic uncertainties, MC simulations should be made in sets different for each running period. However, this analysis is dominantly statistics limited, hence one set of simulations is sufficient.

Another issue is that the center-of-mass energy for running periods 1995-1997, $\sqrt{s} = 300\,\text{GeV}$, is different from that for 1999 and 2000, $\sqrt{s} = 318\,\text{GeV}$. Owing to the same consideration, events of the data with the center-of-mass energy of 300 GeV are scaled by a factor, w_{scale}, that is determined with MC simulations:

$$w_{\text{scale}} = \frac{\sigma_{\text{CC}}^{\text{MC}}(\sqrt{s} = 318\,\text{GeV})}{\sigma_{\text{CC}}^{\text{MC}}(\sqrt{s} = 300\,\text{GeV})} = 1.1, \tag{6.1}$$

where the total cross sections, $\sigma_{\text{CC}}^{\text{MC}}$, are evaluated within a kinematic region of $Q^2 > 200\,\text{GeV}^2$.

As discussed in Section 5.2, MC samples are normalized to the total integrated luminosity of data, to be given in Section 6.3. Factors obtained in (5.5) are used for additional PHP normalization.

6.2 Online trigger

The ZEUS online system contains three trigger levels, FLT, SLT and TLT, introduced in Section 3.3. At each level, trigger conditions are organized in slots, or bits. CC related slots are summarized in Table 6.1 and are briefly described below.

- FLT—using coarse energy measurement within CAL and track information within CTD. There are six slots directly or indirectly related to CC events:

  FLT-41: high transverse energy, $\texttt{E}_{\texttt{T}}^{\texttt{-2ir}}$, without track requirement. Two innermost rings of CAL around the beam pipe are excluded for the measurement, hence the superscript $-$`2ir`.

  FLT-42: for photoproduction. It requires sufficient energy deposit in total or in the EMC section of some sub-CALs and a good track, `TRKb`, that is defined as a track pointing to the nominal interaction point.

  FLT-43: similar to `FLT-41` with a lower threshold but requiring a good track.

FLT-44: for DIS jets. It requires energy deposits in the EMC section of BCAL or RCAL combined with high track multiplicity, `TRKa`, or a good track.

FLT-60: high missing p_T. This is the main slot dedicated to CC events.

FLT-61: dedicated to diffractive CC events. It requires a moderate missing p_T with a good track, and, specifically, small energy deposits around the beam pipe (bp) in FCAL, E^{bp}_{FCAL}.

- SLT—defining several major physics filters, among which CC events are classified into exotic (Exo) filters. At this trigger level, CAL timing is available and is used to suppress non-*ep* collision background. A coarse CTD tracking vertex is used to suppress beam-gas events as well. Energy measurements $E - p_z$ and p_T^{-ir}, the transverse momentum excluding the innermost CAL ring around the beam pipe, are also available to improve the trigger condition. Only one slot is CC relevant:

  ExoSLT-4: dedicated to CC events. It contains four branches, `CC1-CC4`, by combining different p_T cuts with other energy cuts as well as good track requirement.

  Definitions of some trigger conditions appearing in Table 6.1 for this level are listed below, while details can be found found on the Web [75].

  `CCglobTime` : $|T_{global}| < 7$ ns and $N_{PMT,\ global} > 1$
  `CTD_OK` : NtrkSLT ≥ 1 and not CTDBeamGas
  `SLTGoodTrack` : FLF-49 or NtrkSLT ≥ 1
  `NoOffBeamProton`: $|p_y| > 3$ GeV or $p_T > 15$ GeV or $p_T^{-ir} > 6$ GeV or $p_T > 0.06\, p_z$

- TLT—improving trigger decision with fully reconstructing events. It contains two slots for CC events:

  ExoTLT-2: compensation for `ExoTLT-6`. It has a relatively low threshold for transverse momentum, but requires a good vertex with at least one good track.

  ExoTLT-6: the main CC channel. It suppresses further non-*ep* background from events selected by `ExoSLT-4`.

  Definitions of some trigger conditions appearing in Table 6.1 for this level are listed below, while details can be found found on the Web [75].

GoodVertex	:	vertex with $\|z\| < 60\,cm$ and $r < 10\,cm$
GoodTrack	:	$p_T > 0.2\,GeV$ and $DCA < 1.5\,cm$ and $\|ZCA\| < 60\,cm$
UDTimeOK	:	against cosmic rays
BeamGasVertex	:	vertex with $z < -80\,cm$ and $N_{track} \geq 5$
TooManyBeamGasTracks	:	$N_{BeamGasTrack} \geq 5$ and $N_{track} \geq 25$
TwoTrackGoodVertex	:	GoodVertex and $N_{track} \geq 2$

For different running periods, trigger definitions were usually adjusted according to the beam condition at HERA. Even in the same data-taking period, there are also slight changes in thresholds. It happened to the first level trigger more frequently than to the others. Details can be found on the Web [75].

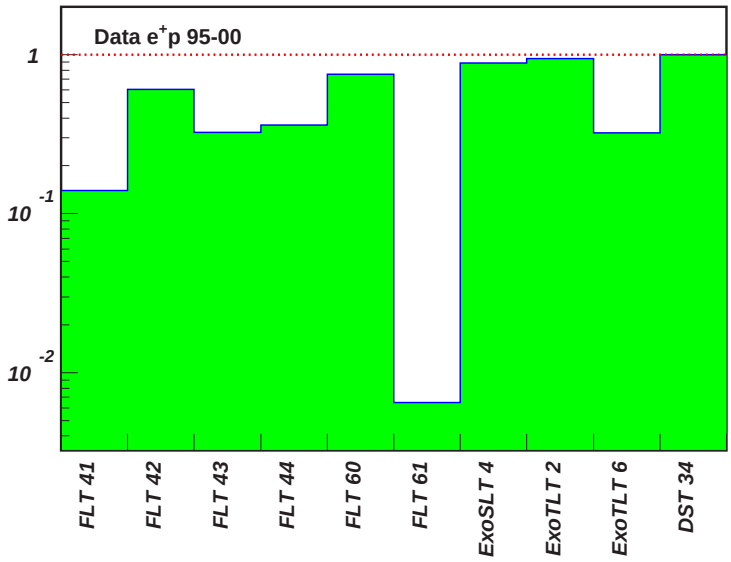


Figure 6.1: *Trigger frequencies of CC relevant bits. It is obtained for the final data sample. Frequencies are normalized to that of DST-34.*

The online trigger bits are not directly used in this analysis. Instead, an offline DST bit is in use. Trigger frequencies of those bits listed in Table 6.1 are shown in Figure 6.1 for the final data sample. It is obvious on the plot that the most "fired" bit at the first trigger level is the $\not\!\!p_T$ related bit, FLT-60.

The online trigger is vital because not only is it the first step for any offline analysis, but also the loss of events by the trigger decision would be unrecoverable afterwards. The effect of some trigger slot can be verified by turn-on plots of the data sample selected with another orthogonal trigger slot. In the case of inspecting FLT-60, a choice of orthogonal slot is FLT-42,

Table 6.1: *Trigger bits relevant to CC event selection. Definitions are based on the data-taking period of year 2000 and main changes are indicated for other data-taking periods. Only key features are given in the definition for brevity.*

Trigger	Definition
FLT-41	$E_T^{-2ir} > 30\,GeV$ (21 GeV in 1999, 15 GeV in 1995)
FLT-42	$(E_{BEMC} \geq 3.4\,GeV$ or $E_{REMC} \geq 2\,GeV)$ and TRKb **96-99** : with (or $E_{CAL} \geq 15\,GeV$ or $E_{EMC} \geq 10\,GeV$)
FLT-43	$E_T^{-2ir} > 11.6\,GeV$ and TRKb
FLT-44	$(E_{BEMC} \geq 4.8\,GeV$ and TRKab) or $E_{REMC} \geq 3.4\,GeV$
FLT-60	$(p_T \geq 5\,GeV$ and $E_T \geq 5\,GeV$ and TRKb) or $(p_T \geq 8\,GeV$ and TRKab) or $(p_T \geq 11\,GeV$ and $E_{FCAL}^{-ir} \geq 10\,GeV)$ **96-99** : $(p_T \geq 8\,GeV$ and (TRKab or $E_{FCAL}^{-ir} \geq 10\,GeV)$ **96** : $5\,GeV \rightarrow 6\,GeV$, $8\,GeV \rightarrow 9\,GeV$
FLT-61	$p_T \geq 5\,GeV$ and TRKb and $(E_{FCAL}^{bp} < 12.5\,GeV)$ **96-99** : $5\,GeV \rightarrow 3\,GeV$ **96-97** : $\rightarrow E_{FCAL}^{bp} < 12.5\,GeV$ or $(E_{FCAL}^{bp} < 7.5\,GeV$ and no PRT hit)
ExoSLT-4	(CCglobTime or CTD_OK) and NoOffBeamProton and (CC1 or CC2 or CC3 or CC4) CC1 = $p_T > 6\,GeV$ and $E_T^{-2ir} > 6\,GeV$ and SLTGoodTrack CC2 = $p_T > 9\,GeV$ and $E_T^{-1ir} > 8\,GeV$ and $E_{FCAL} > 20\,GeV$ CC3 = $p_T > 9\,GeV$ and $p_T^2 > 2.31E_T$ and $E_{FCAL} > 80\,GeV$ CC4 = $E - p_z > 6\,GeV$ and $p_T^2 > 2.25E_T$ and SLTGoodTrack
ExoTLT-2	$p_T > 6\,GeV$ and GoodVertex and $N_{GoodTrack} > 0$ and UDTimeOK **95** : $6\,GeV \rightarrow 7\,GeV$, without UDTimeOK
ExoTLT-6	$p_T > 8\,GeV$ and ExoSLT-4 and $(E_{FCAL} > 10$ or FLF-49) and (not OffBeamProton) and (not BeamGasVertex) and (not TooManyBeamGasTracks or TwoTrackGoodVertex) and $(p_T^{-ir} > 10\,GeV$ or $N_{CTD\,Hits} < 2500)$ and UDTimeOK **95-97** : without (not OffBeamProton) **95** : $p_T > 9\,GeV$ without UDTimeOK
DST-34	$(ExoTLT\text{-}2$ or ExoTLT-6 or $p_T^{-ir} > 6\,GeV)$ and $p_T > 7\,GeV$ **96-97** : $p_T^{-ir} > 7\,GeV$, without cut on p_T **95** : $(p_T > 8\,GeV$ or ExoTLT-2) and $p_T^{-ir} > 8\,GeV$

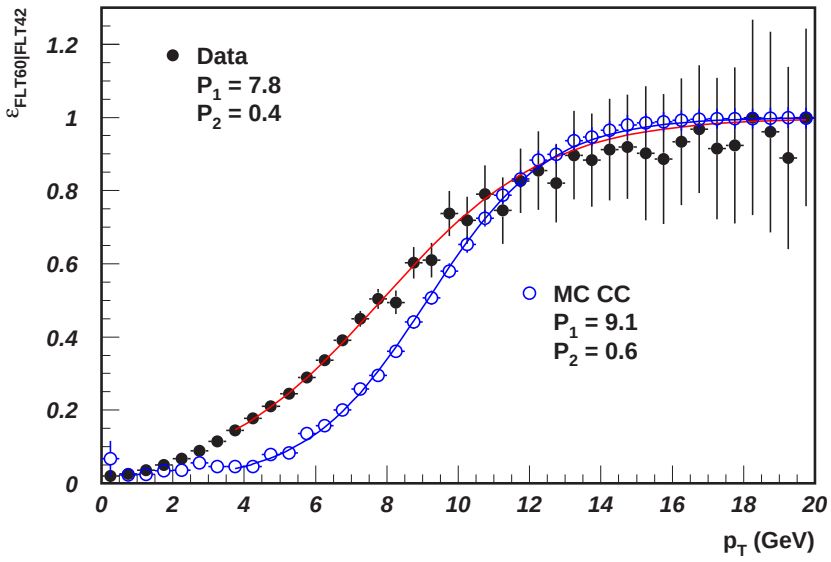


Figure 6.2: *Efficiencies of the trigger slot FLT-60 as a function of p_T for data (solid dots) and CC MC (open dots) samples. FLT-42 is chosen as an orthogonal slot and no other cuts are used. Fit curves are represented by solid lines. P_1 and P_2 are fit parameters. Explanation can be refered in the text.*

and the turn-on plot is the distribution of trigger efficiencies as a function of transverse momentum, p_T. The trigger efficiency in the ith bin of p_T is defined by

$$\epsilon_i = \frac{N_i^{\mathrm{FLT60\,|\,FLT42}}}{N_i^{\mathrm{FLT42}}}, \tag{6.2}$$

with N_i^{FLT42} denoting number of events selected by FLT-42 and $N_i^{\mathrm{FLT60\,|\,FLT42}}$ by both FLT-42 and FLT-60. Examples are shown in Figure 6.2 for data and CC MC samples. Curves on the plot are obtained by performing fits with the form

$$\epsilon(p_T) = 1 - \frac{1}{1 + e^{P_2(p_T - P_1)}}, \tag{6.3}$$

with P_1 representing the turn-on point and P_2 the slope of the curve at turn-on point. The plot is made without any other event cuts except the two trigger slots, therefore data and MC give very different fit results. While the turn-on point reflects the threshold used in trigger, the slope corresponds to resolution of measurement. An early study [76] shows uncertainty caused by the resolution difference between online trigger and MC simulation is negligible.

6.3 Data sample preselection

There are in total 3904 physics runs of e^+p data corresponding to an integrated luminosity of 111.1 pb^{-1} collected with the ZEUS detector during the HERA running periods of from 1995 to 2000. Among them, there are some runs with shifted interaction point, 24 runs in 1995 and 33 runs in 2000, corresponding to an integrated luminosity of 1.1 pb^{-1}. These runs are for special use and excluded from this analysis. Data taking is a very complicated process, and the quality of data can be affected by many things. For instance, some key detector component might fail to work properly and was switched off temporarily. A lot of effort has been paid for Data Quality Monitoring (DQM) and results are summed in a special offline software package, EVTAKE, that picks out problematic runs and events. With EVTAKE, the data sample is reduced to 3665 runs, corresponding to an integrated luminosity of 110 pb^{-1}, that is used for this analysis. The distribution of integrated luminosities of individual runs is shown in Figure 6.3.

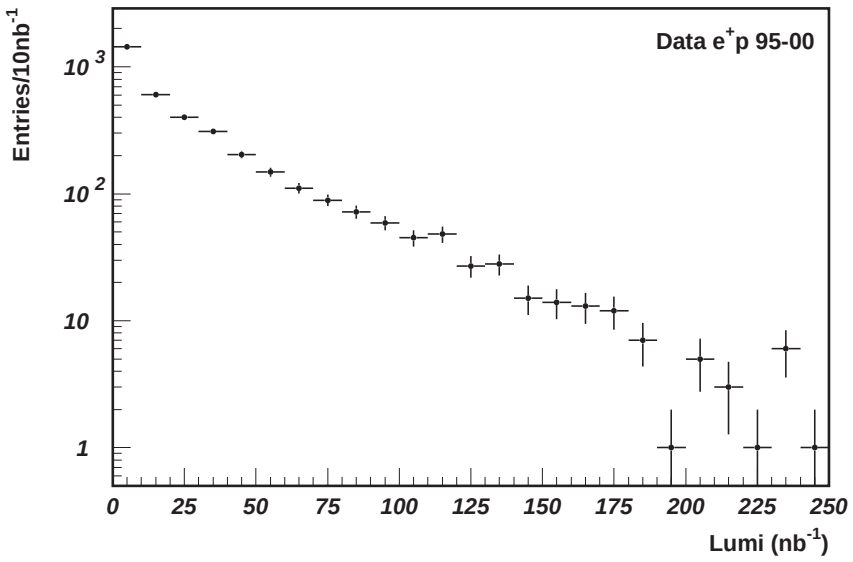


Figure 6.3: *Distribution of integrated luminosities of individual runs for the data used in this analysis.*

The trigger-like offline DST bits, see Section 3.4, are used for the purpose of selecting a certain class of physics events. The DST bit-34 is designed for CC events and is in use as physics preselection. Its definition is in Table 6.1 as well. It is a combination of the CC related TLT trigger bits with offline improved measurements of p_T and p_T^{-ir}.

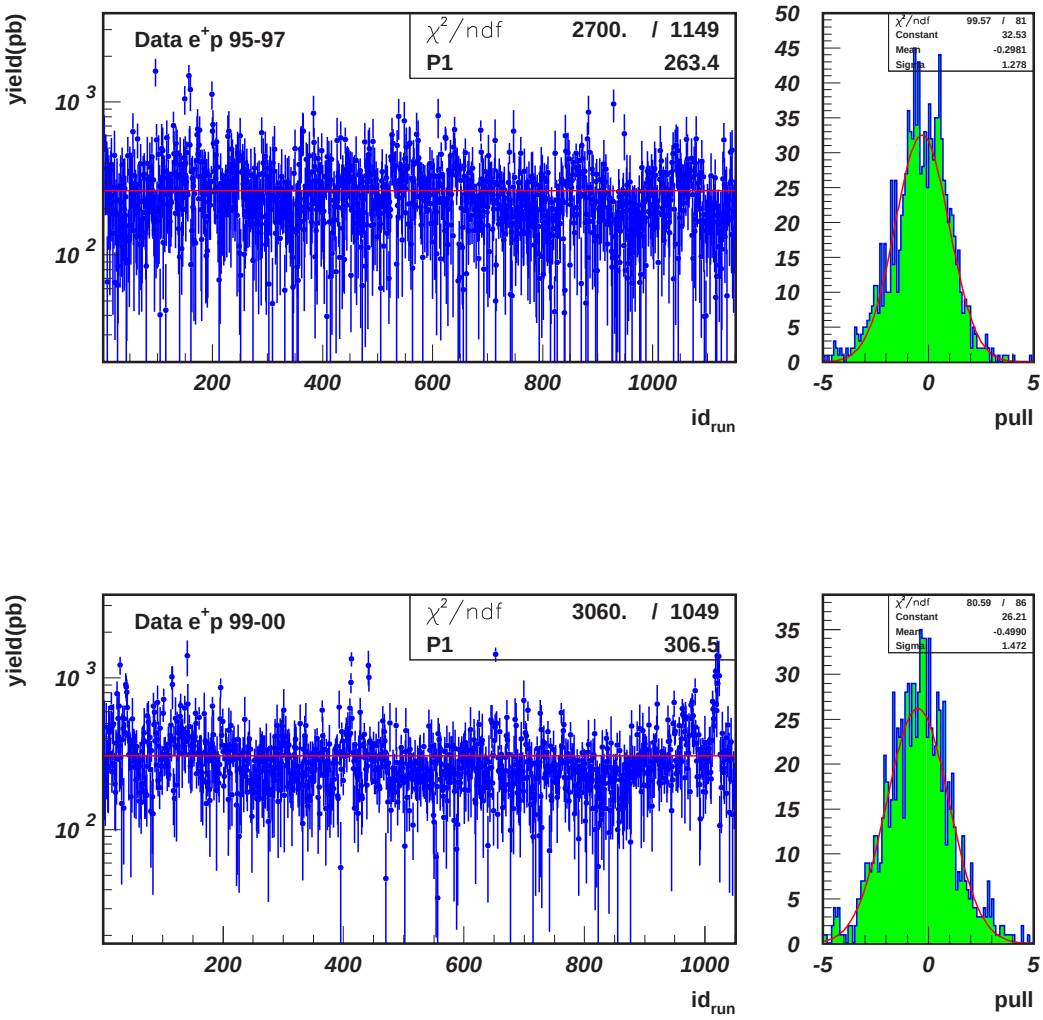


Figure 6.4: *Distributions of yield-of-run with respect to run indices (left) and corresponding pull plots (right) for preselected data samples of different center-of-mass energy. Short runs with integrated luminosities less than* $10\,nb^{-1}$ *are collected as one bin. Gaussian fit is performed for pull distribution, whereas a zero-order polynomial fit for yield distribution, shown as solid lines on the plots.*

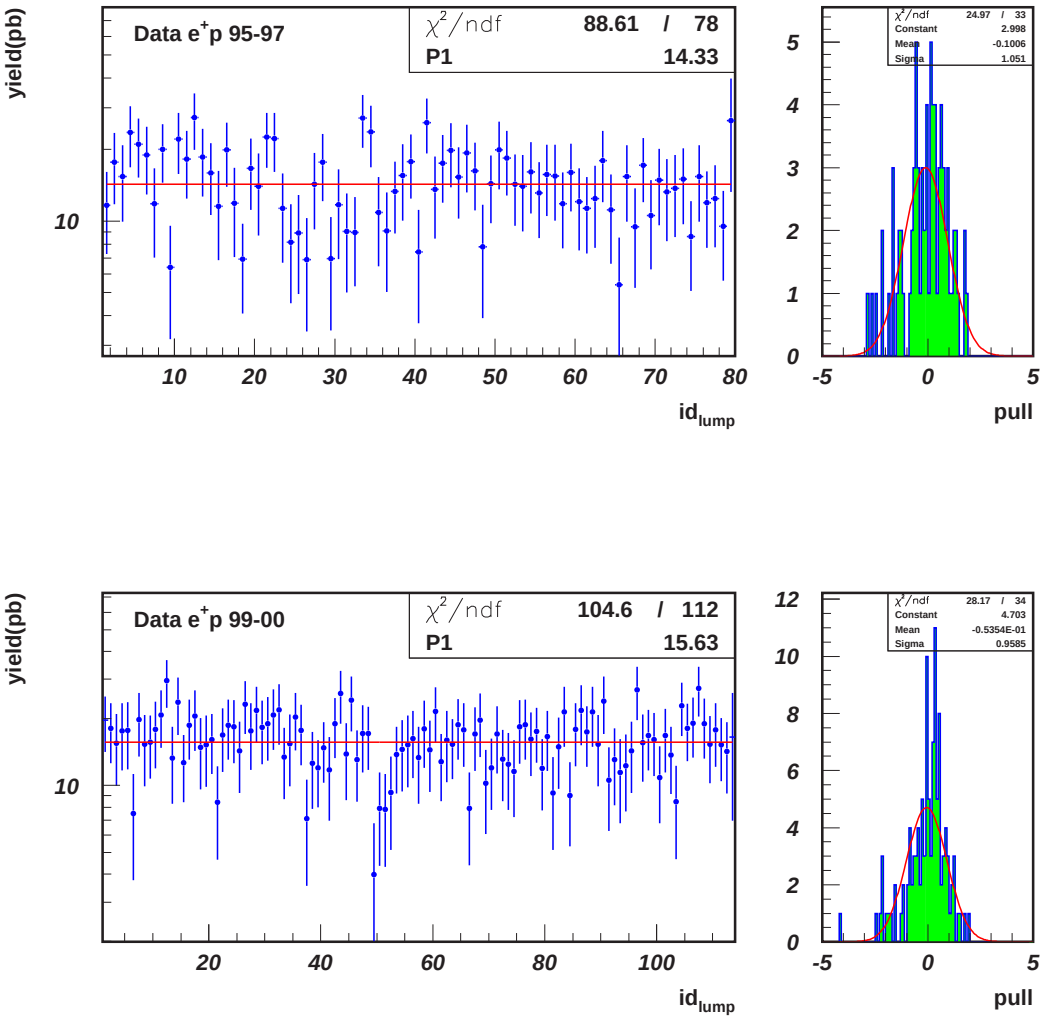


Figure 6.5: *Yield distributions of run-lumps (left) and corresponding pull plots (right) for final data samples of different center-of-mass energy. Run-lumps are formed chronologically with integrated luminosity not less than* $0.5\,pb^{-1}$. *Gaussian fit is performed for pull distribution, whereas a zero-order polynomial fit for yield distribution, shown as solid lines on the plots.*

The quality of runs can be inspected by means of yield-of-run defined by

$$\text{yield}_i = \frac{N_i}{\mathscr{L}_i}, \tag{6.4}$$

where $\mathscr{L}_i$ is the integrated luminosity of some run i and N_i is the corresponding number of events after some event selection. For good runs, the distribution of yields as a function of runs will be roughly constant within statistical fluctuations and its pull plot will exhibit a standard normal distribution with $\mu = 0$ and $\sigma^2 = 1$. Large deviations from these expectation, however, will indicate potential problems in selected runs. Yield-of-run distributions for running periods with different center-of-mass energy and their pull plots are shown in Figure 6.4 after applying the offline preselection as well as cuts discussed in Sections 6.4 and 6.5. To avoid large statistical fluctuations, those short runs with integrated luminosity less than $10\,\text{nb}^{-1}$ are added together to form one bin in each yield-of-run plot. No large deviation is observed in these plots. This technique is very useful, hence through the whole procedure of analysis, the yield distribution is inspected after every cut. However, the cross section of charged current is relatively small, and the yield-of-run will become of little statistics while proceeding with more cuts. As an alternative, consecutive runs are put together until the lumped integrated luminosity is larger than $0.5\,\text{pb}^{-1}$, and yields of these run-lumps may be inspected. Distributions for final data sample are shown in Figure 6.5. Pull plots exhibit very well Gaussian distributions.

6.4 Charged current selection

Primary signature of a CC event is a high missing transverse momentum, $\not{p}_T$, or equivalently the event's transverse momentum, p_T. Distributions of p_T for data as well as for MC samples after the offline preselection are shown in Figure 6.6. Peaks around 8 GeV are due to the cuts on p_T in the online trigger slots and the offline DST bit discussed before. Large discrepancies between data and MC samples in the range of $p_T < 25\,\text{GeV}$ are caused by non-simulated background that are dominantly non-ep collision events. In the range of $p_T > 25\,\text{GeV}$, data points are very well described by the MC simulation. It is clear that the higher the cut on p_T, the purer the CC sample will be. However, it is also clear that contributions of NC and PHP background, simulated in MC samples, fade out only at about 40 GeV, at which a cut on p_T will get rid of most CC events as well. Therefore, other means than only a p_T cut are needed to suppress background events and keep a relatively low p_T cut.

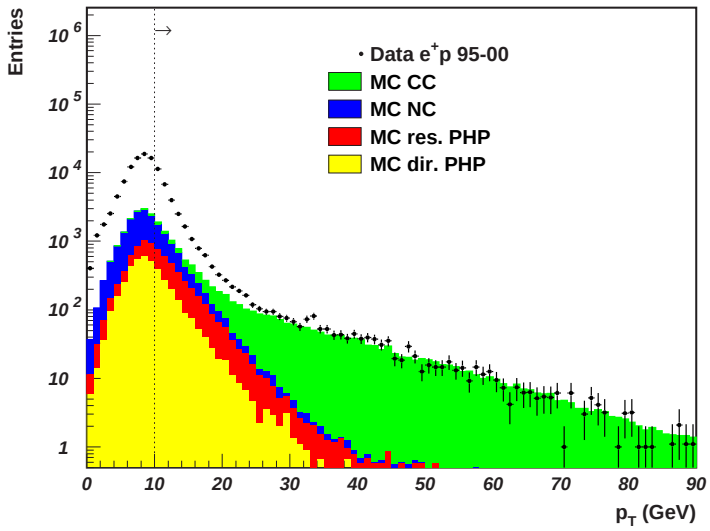


Figure 6.6: *Distributions of p_T for preselected data and MC samples. The soft p_T cut is indicaded by the line with an arrow. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

The discussion can be verified quantitatively by defining cut efficiency, $\epsilon_{\rm cut}$, and purity, $\mathcal{P}_{\rm cut}$ for MC samples,

$$\begin{aligned} \epsilon_{\rm cut} &= \frac{N^{\rm CC}_{\rm cut}}{N^{\rm CC}}, \\ \mathcal{P}_{\rm cut} &= \frac{N^{\rm CC}_{\rm cut}}{N^{\rm CC}_{\rm cut} + N^{\rm non\text{-}CC}_{\rm cut}}, \end{aligned} \tag{6.5}$$

where N denotes the number of events before a cut, and $N_{\rm cut}$ after a cut. Superscripts *"CC"* and *"non-CC"* represent MC samples. For the data sample, $\epsilon^{\rm Data}_{\rm cut}$ can be defined similarly. In Addition, the fraction of non-simulated background after a cut, $f^{\rm non\text{-}sim}_{\rm cut}$, can be estimated as

$$f^{\rm non\text{-}sim}_{\rm cut} = \frac{N^{\rm Data}_{\rm cut} - (N^{\rm CC}_{\rm cut} + N^{\rm non\text{-}CC}_{\rm cut})}{N^{\rm Data}_{\rm cut}}, \tag{6.6}$$

with $N^{\rm Data}_{\rm cut}$ denoting the number of events in the data sample after a cut. Variation of $\epsilon^{\rm CC}_{\rm cut}$, $\mathcal{P}_{\rm cut}$, $\epsilon^{\rm Data}_{\rm cut}$ and $f^{\rm non\text{-}sim}_{\rm cut}$ for different cut positions in p_T are shown in Figure 6.7. A curve for the products of $\epsilon^{\rm CC}_{\rm cut}$ and $\mathcal{P}_{\rm cut}$ is also shown in the same plot to indicate the trend of balance between cut efficiency and purity.

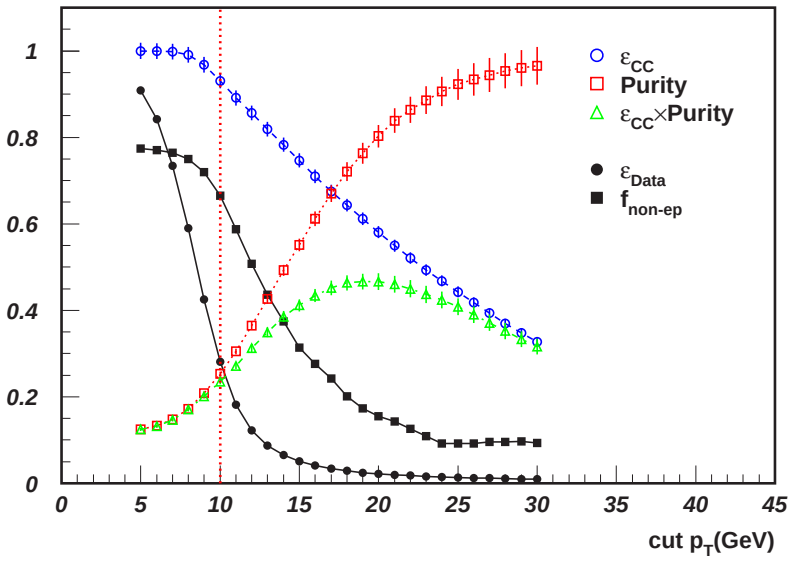


Figure 6.7: *Cut efficiencies and purities as functions of p_T cut positions for preselected samples. Explanation for variables shownq on the plot can be found in the text. The initial p_T cut-position is indicated by the vertical line.*

After applying event criteria discussed in the following sections, a hard cut,

$$p_T > 12\,\text{GeV}, \tag{6.7}$$

will be used as the primary condition for selecting CC events. However, a soft cut,

$$p_T > 10\,\text{GeV}, \tag{6.8}$$

is used as a preselection as well in order to study efficient means for background suppression and to reduce the volume of data.

6.5 Primary vertex and good tracks

Primary vertices of events are important for reconstructing transverse momenta and identifying D^*s. While the reason for the former is clear in (4.1), the reason for the latter is that D^*s decay immediately after being produced, namely at primary vertices. Resolution of primary vertices in the X-Y plane, as shown in Figure 6.8, is much worse than that of measured beam spots, $0.3 \times 0.1\,\text{mm}^2$. Therefore, X and Y coordinates of primary vertices are always set to the nominal interaction point: $x_{\text{vertex}} = y_{\text{vertex}} = 0\,\text{cm}$. As for z_{vertex}, an early study [77] showed most of them lie within $\pm 30\,\text{cm}$ around the

nominal interaction point, with secondary peaks centering at about ±70 cm. A conservative cut,

$$|z_{\text{vertex}}| < 50\,\text{cm}, \tag{6.9}$$

is chosen for this analysis. A distribution for data is shown in Figure 6.9.

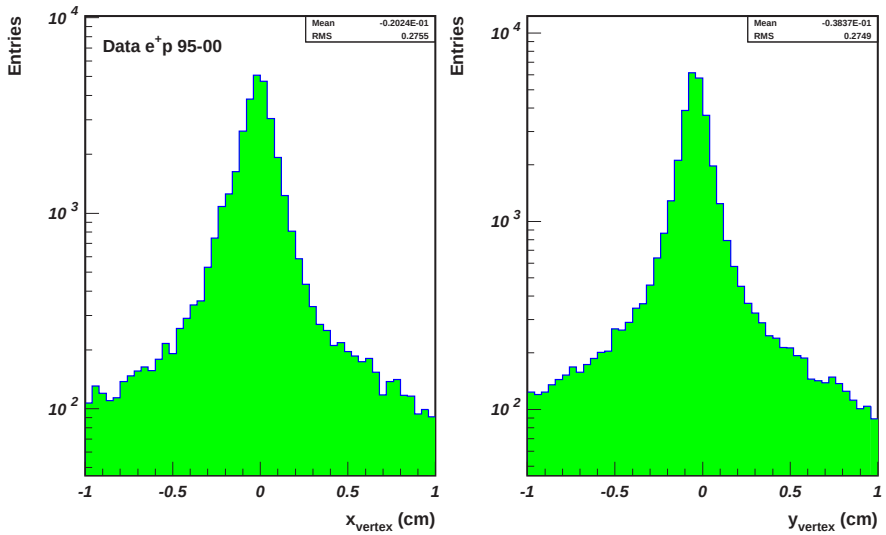


Figure 6.8: *Distributions of x_{vertex} and y_{vertex} for data sample preselected with the soft p_T cut and the cut on number of good tracks, N^{good}_{track}.*

At least three *good tracks* are needed for reconstructing one D^* with the method used for this work, to be discussed in Section 7.1. Good tracks are subject to the following criteria:

$$\begin{aligned} &\text{associated with the primary vertex,} \\ &\text{reaching at least the third super-layer,} \\ &p_T^{\text{track}} > 0.2\,\text{GeV}. \end{aligned} \tag{6.10}$$

The last condition takes the effect of CTD hit corrections into account, see Section 4.2. Characteristic distributions of good tracks are shown in Figure 6.10. The pattern in the distribution of the outermost super-layer of good tracks is because the tracking algorithm uses mainly hits in the axial, i.e. odd-numbered, super-layers.

Since the requirement of good tracks is mandatory for D^* identification, a cut on the number of good tracks,

$$N^{\text{good}}_{\text{track}} \geq 3, \tag{6.11}$$

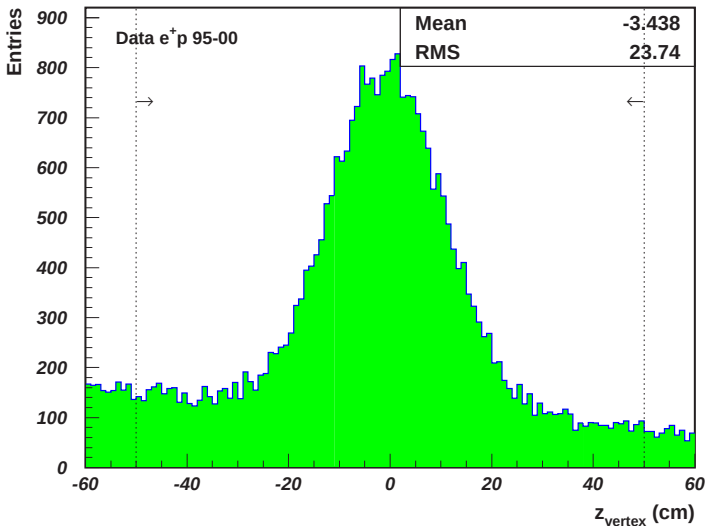


Figure 6.9: *Dsitribution of z_{vertex} for the same data sample as for Figure 6.8. The signal region is indicaded by the lines with arrows.*

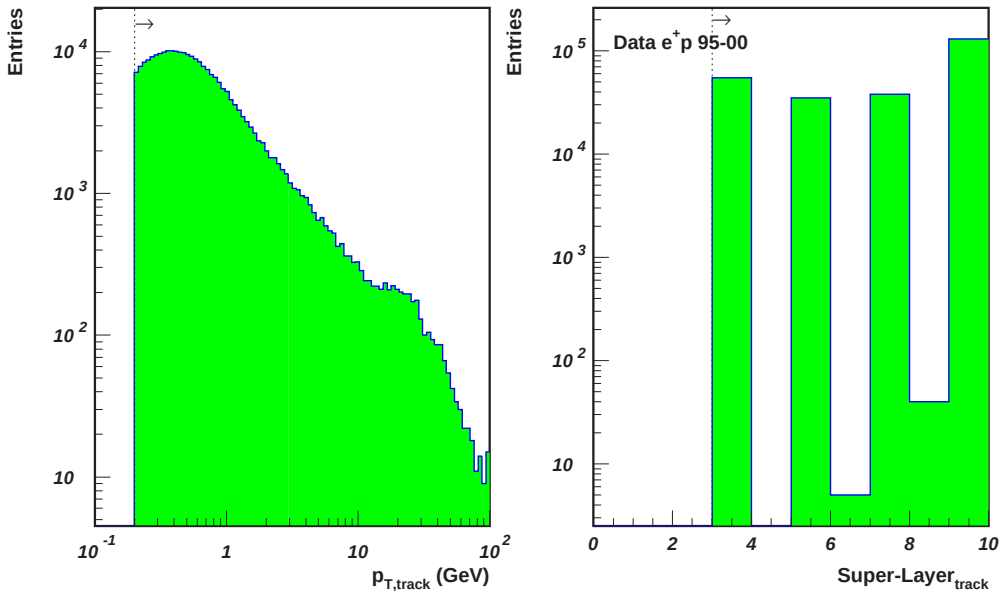


Figure 6.10: *Distributions of transverse momenta and outermost super-layer of good tracks. Signal regions are indicaded by the lines with arrows, respectively.*

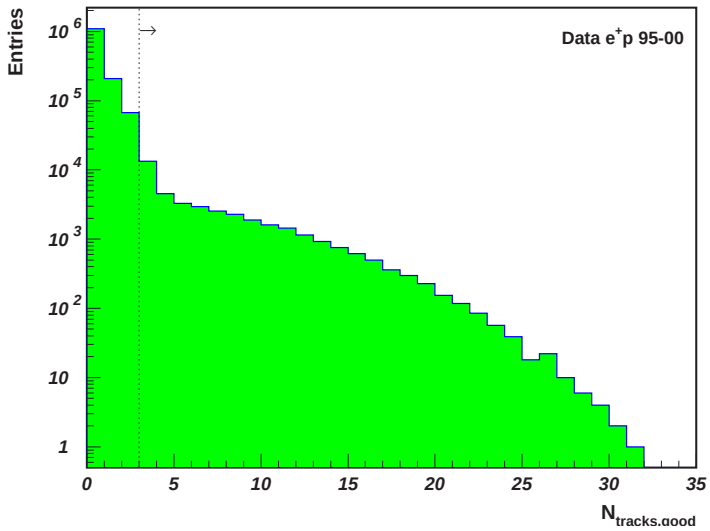


Figure 6.11: *Distribution of number of good tracks, N_{track}^{good}, for data sample preselected with the same condition as for Figure 6.8 except the cut on N_{track}^{good}. The cut is indicaded by the line with an arrow.*

is demanded. As shown in Figure 6.11, this cut reduces the data volume tremendously.

Technically, cuts on the primary vertex and number of good tracks are classified into preselection as well. The following discussions are based on data samples preselected with conditions introduced so far.

6.6 Non-*ep* collision background

Non-*ep* collision background are events not being produced by *ep* collisions. They are typically cosmic rays, beam-gas interactions, and halo-μ's. These events are produced and recorded randomly during data taking. Although most of them have been filtered out by the online trigger, a large amount of them still remaines in the data sample. However, they cannot be modelled via MC simulation. It is therefore important to suppress these kinds of background as much as possible.

Cosmic rays have been heavily suppressed by the online calorimeter timing requirement. A certain amount of them leaked into the data sample, as shown in Figure 6.12. The requirement is tightened on the more precisely

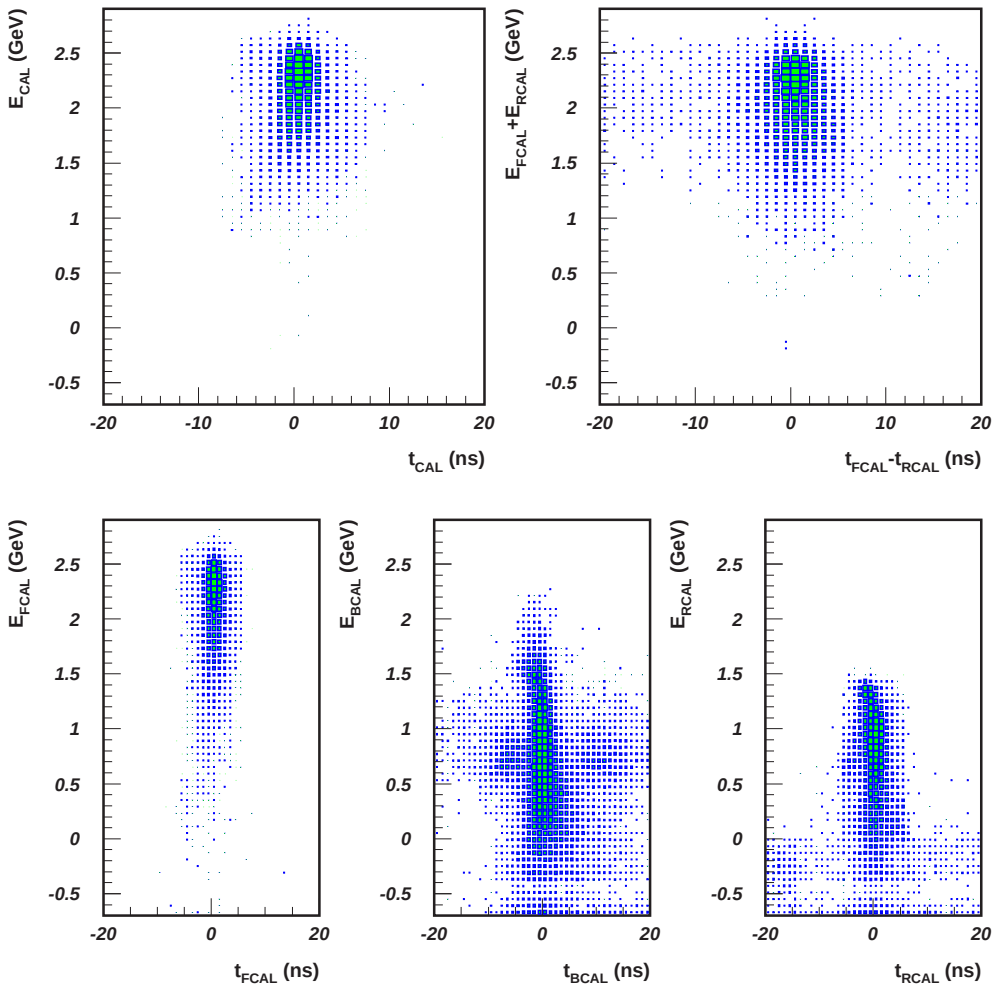


Figure 6.12: *Scatter plots of calorimeter energy deposits versus timing for data sample selected with already discussed criteria.*

reconstructed offline quantities:

$$\begin{aligned}
|t_{\mathrm{FCAL}}| < 6\,\mathrm{ns} \text{ or } N_{\mathrm{FCAL}} < 2 \text{ or } E_{\mathrm{FCAL}} < 1\,\mathrm{GeV},\\
|t_{\mathrm{BCAL}}| < 6\,\mathrm{ns} \text{ or } N_{\mathrm{BCAL}} < 2 \text{ or } E_{\mathrm{BCAL}} < 1\,\mathrm{GeV},\\
|t_{\mathrm{RCAL}}| < 6\,\mathrm{ns} \text{ or } N_{\mathrm{RCAL}} < 2 \text{ or } E_{\mathrm{RCAL}} < 1\,\mathrm{GeV},\\
|t_{\mathrm{CAL}}| < 6\,\mathrm{ns} \text{ or } N_{\mathrm{CAL}} < 2 \text{ or } E_{\mathrm{CAL}} < 2\,\mathrm{GeV},\\
|t_{\mathrm{FCAL}} - t_{\mathrm{RCAL}}| < 6\,\mathrm{ns} \text{ or } \min(N_{\mathrm{FCAL}}, N_{\mathrm{RCAL}}) < 2\\
\text{or } \min(E_{\mathrm{FCAL}}, E_{\mathrm{RCAL}}) < 1\,\mathrm{GeV},
\end{aligned} \tag{6.12}$$

with t denoting calorimeter timing, N number of PMTs used for timing and E the corresponding energy deposit. Subscripts represent different calorimeter parts, FCAL, BCAL and RCAL, or the whole, CAL. The requirements on E and N ensure good resolution of the calorimeter timing. These cuts are only applied to real data because calorimeter timing is not simulated in MC properly.

The beam-gas interaction is caused by the proton or positron beam scattered on the atoms of the residual gas in the beam pipe. Most of the particles produced are concentrated in a narrow cone around the beam axis and have a small transverse energy. Therefore, they will deposit energy mainly in the inner part of FCAL. This characteristic can be visualized in scatter plots of the net transverse momentum excluding the innermost part of FCAL, p_T^{-ir}, versus event's p_T, shown in Figure 6.13. Comparing the distributions of the MC samples, there is a large excess for data in the region of p_T less than about 25 GeV. A cut on the linear relation, $p_T^{-ir} > k \cdot p_T + b$, is expected to remove the excess efficiently with little loss in CC signal. To find appropriate parameters for the linear function, a two-dimensional fit is performed, shown in Figure 6.14, and the difference between p_T^{-ir} and p_T is checked, shown in Figure 6.15. Resulting slopes of the fits show very good agreement between the data and the CC MC sample. Taking into account RMS in Figure 6.15(b), a cut is formed as

$$p_T^{-ir} > 0.9 p_T - 6. \tag{6.13}$$

Beam-gas events are also characterized by high multiplicities of non-primary-vertex associated tracks, shown in Figure 6.16. A cut [76],

$$N_{\mathrm{track}}^{\mathrm{good}} > \frac{1}{4}(N_{\mathrm{track}}^{\mathrm{all}} - 20), \tag{6.14}$$

can remove this kind of beam-gas background efficiently.

Halo-μ events are actually a kind of proton-beam-gas events which however happen upstream far away from the detector. While most of other final

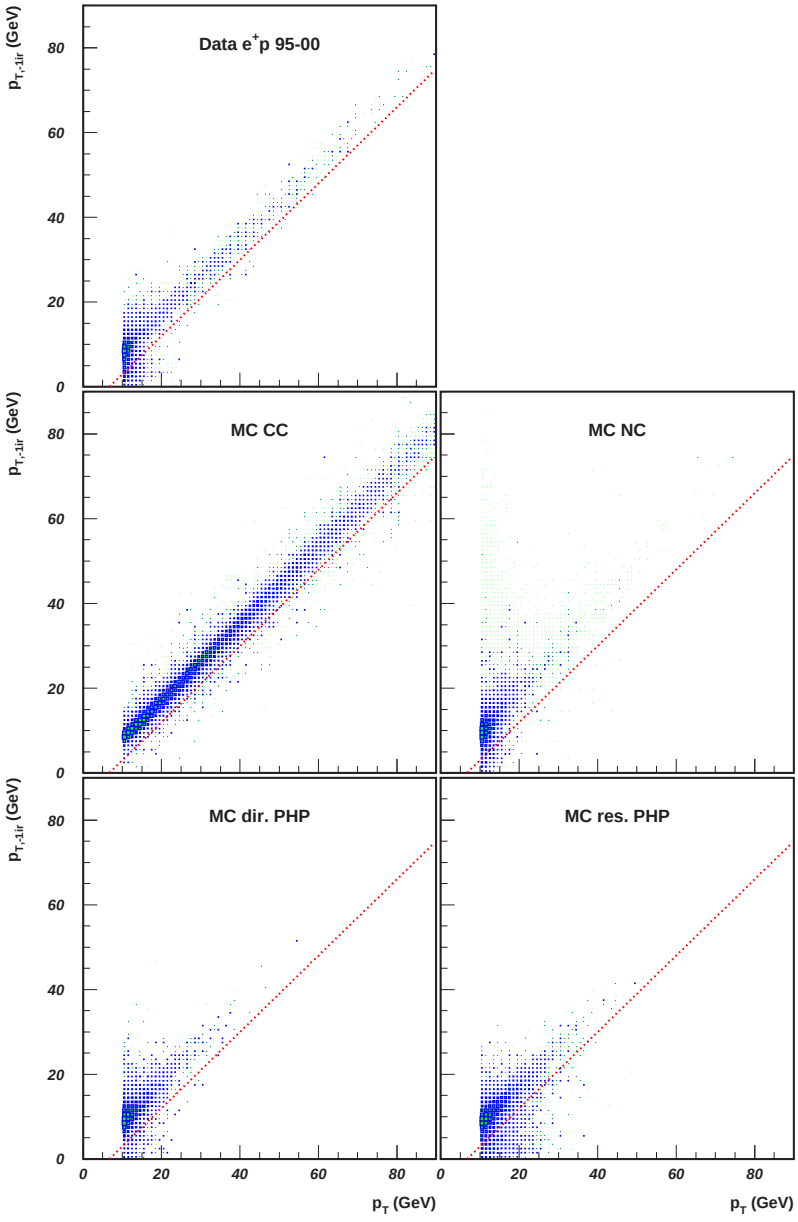


Figure 6.13: *Scatter plots of p_T^{-ir} versus p_T for data and MC samples selected with already discussed criteria. The cut is indicated by the line.*

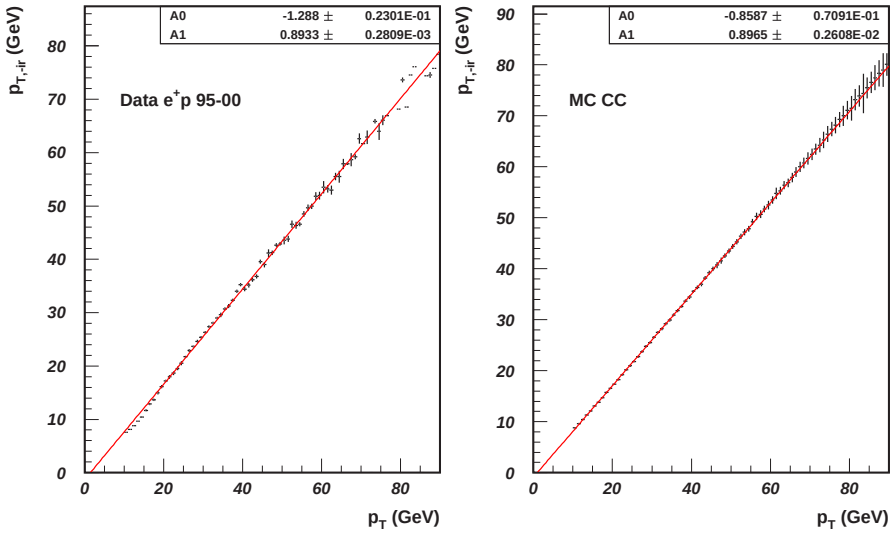
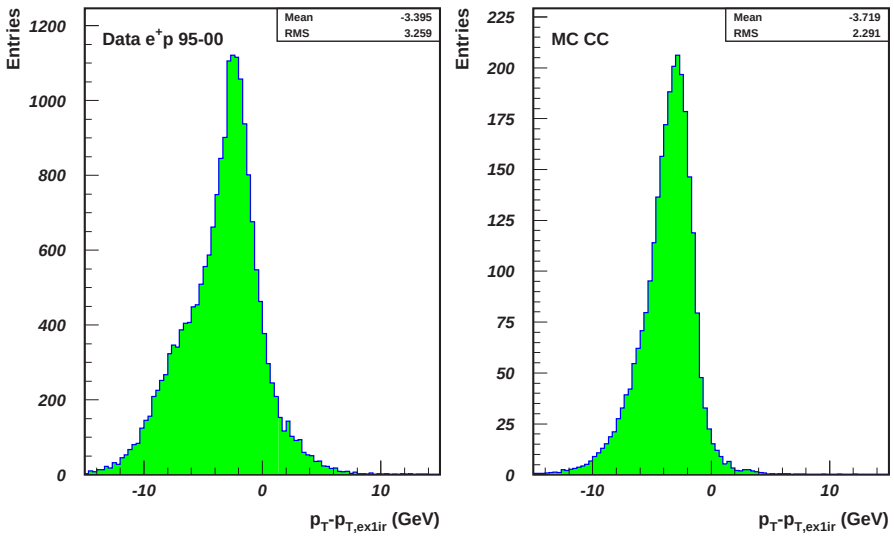


Figure 6.14: *Two-dimensional fit for p_T^{-ir} as a function of p_T for data and CC MC samples. The line shows the result of the fit described in the text.*



(a) (b)

Figure 6.15: *Distributions of $p_T - p_T^{-ir}$ for data and CC MC samples.*

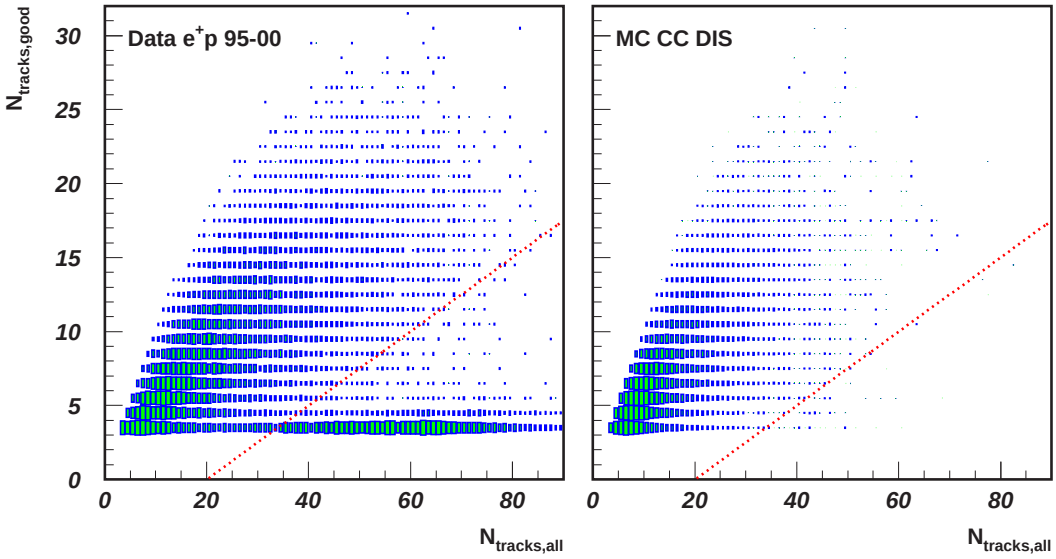


Figure 6.16: *Scatter plots of N^{good}_{track} versus N^{all}_{track} for data and CC MC samples selected with already discussed criteria. The cut is indicated by the line.*

state particles are prevented from entering the detector by the veto wall outside of the detector, a μ will penetrate into the detector because of its minimum ionization. Therefore, the halo-μ's are parallel to the beam axis and have very special patterns of energy deposits in the calorimeter. Most of halo-μ events have been gotten rid of by the online calorimeter timing requirement. The ones left are further removed by the good tracking requirement specialized for this analysis because a pure halo-μ event normally lacks good tracks. Although, overlay of an halo-μ and an ep collision event has so far no clear ways to deal with, it happens rarely and has little effect in D^* identification.

Besides these kinds of background events, there is another contamination owing to a technical reason [78]. As described in Section 3.2.1, energy deposit in one calorimeter cell is read out by two PMTs. If one of the two PMTs is dead, its readout is set to that of the other one. This setup usually has little influence in p_T measurement. On the other hand, spontaneously discharging PMTs will fake energy deposits in cells. They can be identified usually by checking the imbalance of readout of two PMTs for one cell, $A_{\rm imb} = (E_L - E_R)/(E_L + E_R)$, with E_L and E_R denoting energy measurements of left and right PMTs. However, there is a probability for a PMT discharging in a cell with one dead PMT. In this case, $A_{\rm imb}$ is artificially zero, and the faked energy deposit may be large enough to affect the p_T measurement. The effect is not simulated in MC and can be disclosed by inspecting cells with

highest energy deposit (hc), as shown in Figure 6.17. There is a clear spike at zero in the distribution of $A_{\rm imb}$, and excess of data in the distribution of $p_T^{\rm hc}/p_T$ can be observed in the region of $p_T^{\rm hc}/p_T > 0.5$. The condition against this problem is

$$A_{\rm imb}^{\rm hc} \neq 0 \text{ or } \frac{p_T^{\rm hc}}{p_T} < 0.5. \tag{6.15}$$

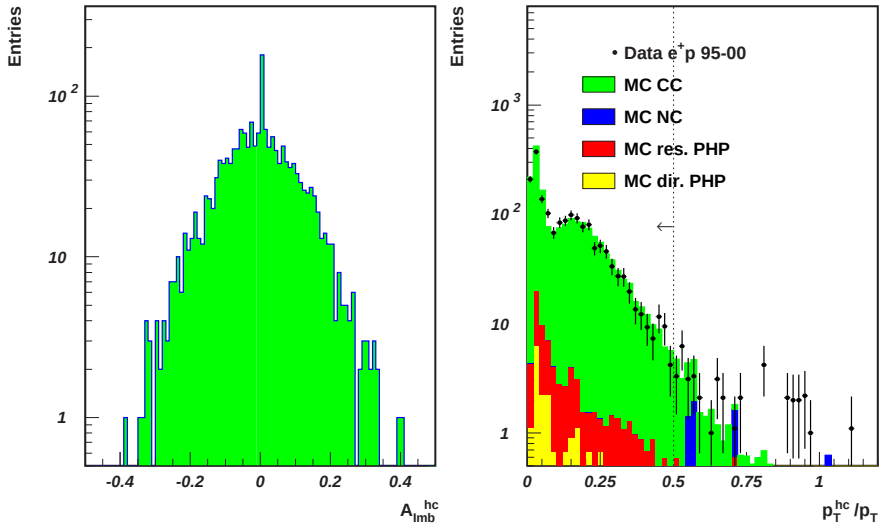


Figure 6.17: *Distributions of A_{imb}^{hc} and p_T^{hc}/p_T for data and MC samples selected with already discussed criteria. Explanation can be found in the text. The cut on p_T^{hc}/p_T is indicaded by the line with an arrow. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

6.7 Non-CC *ep* collision background

In *ep* collision, processes other than the CC reaction, namely NC and PHP reactions, are the main type of background to be suppressed in this analysis. The most efficient means to suppress this background is by the cut on p_T, i.e. (6.7). However, they still populate the final data sample considerably, especially the PHP background, due to their much larger cross-sections in comparison to CC.

An efficient way is to consider energy flow in the transverse event plane. The non-isotropic distribution can be quantified as the quantity p_T/E_T. For CC events, this quantity tends to be large, while for NC and PHP, it will be

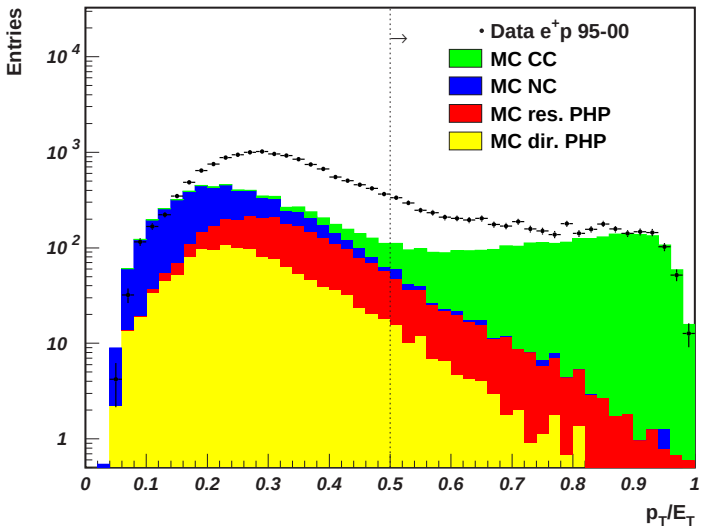


Figure 6.18: *Distributions of p_T/E_T for data and MC samples selected with already discussed criteria. The cut is indicaded by the line with an arrow. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

small, as shown in Figure 6.18. A cut,

$$\frac{p_T}{E_T} > 0.5, \tag{6.16}$$

is able to get rid of a large fraction of NC and PHP events.

Another tool is to calculate the vector sum of transverse momenta of good tracks, $p_T^{\rm CTD}$. Although this can not be used as a measurement of event's transverse momentum due to neutral particles undetected in CTD, its direction, namely azimuthal angle, is not affected much, hence is a measurement of p_T direction. As for resolved PHP, which is the dominant background after applying previous cuts, large p_T is mainly caused by calorimetry mismeasurement. The azimuthal angular difference between $p_T^{\rm CTD}$ and the event p_T, $\Delta\phi(p_T, p_T^{\rm CTD})$, should distribute isotropically for resolved PHP, whereas it tends to be small for the CC signal. A comparison is made in Figure 6.19 and results in a criterion,

$$|\Delta\phi(p_T, p_T^{\rm CTD})| < 30^\circ. \tag{6.17}$$

Asymmetric distribution of good tracks in the transverse event plane can also help separate the PHP background from the CC signal. For a CC event, tracks are mainly in the forward region of the direction of p_T and rarely in

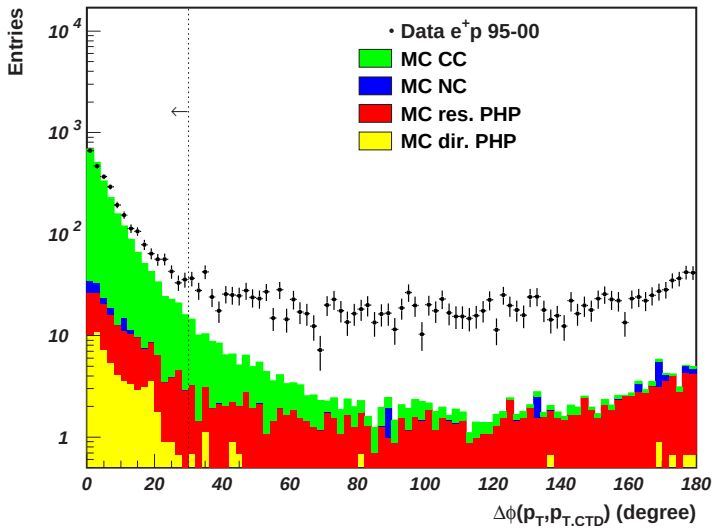


Figure 6.19: *Distributions of* $\Delta\phi(p_T, p_T^{CTD})$ *for data and MC samples selected with already discussed criteria. The cut is indicaded by the line with an arrow. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

the backward region, whereas for a PHP event, whenever it has a high p_T, there is also a high probability for a small amount of tracks leaking into the backward region. With azimuthal angular difference, $\Delta\phi(p_T, p_T^{\rm track})$, between event's p_T and individual track's transverse momentum, $p_T^{\rm track}$, two variables are defined:

$$N^+_{\rm track} \equiv \text{number of tracks in region of } |\Delta\phi(p_T, p_T^{\rm track})| < 30^\circ,$$
$$N^-_{\rm track} \equiv \text{number of tracks in region of } |\Delta\phi(p_T, p_T^{\rm track})| > 150^\circ.$$

Scatter plots of $N^-_{\rm track}$ versus $N^+_{\rm track}$ for data and MC samples are shown in Figure 6.21. A condition for selecting CC events is

$$\begin{aligned} &N^-_{\rm track} = 0, \text{ or} \\ &N^-_{\rm track} = 1 \text{ and } N^+_{\rm track} \geq 7. \end{aligned} \tag{6.18}$$

There are no further cuts against NC background, since only little are left and they will be removed by a cut on $y_{\rm JB}$ as discussed next.

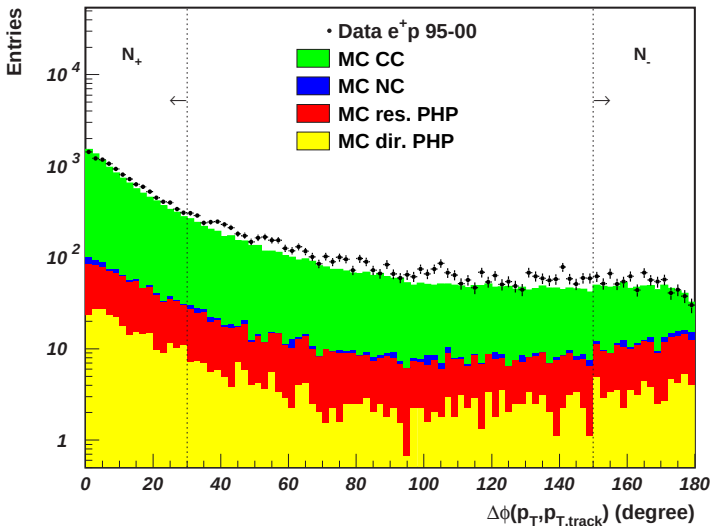


Figure 6.20: *Distributions of $\Delta\phi(p_T, p_T^{track})$ for data and MC samples selected with already discussed criteria. Regions for defining N^+_{track} and N^-_{track} are indicaded by the lines with arrows. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

6.8 DIS kinematic region

As discussed in Section 4.3, the reconstructed DIS kinematics are strongly dependent on p_T as well as on $E - p_z$. A cut on p_T is also a direct cut on Q^2_{JB}. For comparing to theoretical predictions, it is necessary to put a constraint on Q^2_{JB}. On the other hand, the equation (4.12) implies that the Q^2_{JB} resolution degenerates when y_{JB} is close to 1, hence a constraint as well. DIS kinematic cuts are set to be

$$\begin{aligned} Q^2_{\mathrm{JB}} &> 200\,\mathrm{GeV}^2, \\ y_{\mathrm{JB}} &< 0.9. \end{aligned} \tag{6.19}$$

Distributions of Q^2_{JB} and y_{JB} are shown in Figure 6.22 and Figure 6.23, respectively. The large excess of data in the low Q^2_{JB} region will be removed after applying the hard p_T cut. The distribution of y_{JB} shows most of remaining NC background to be cleaned up by this cut.

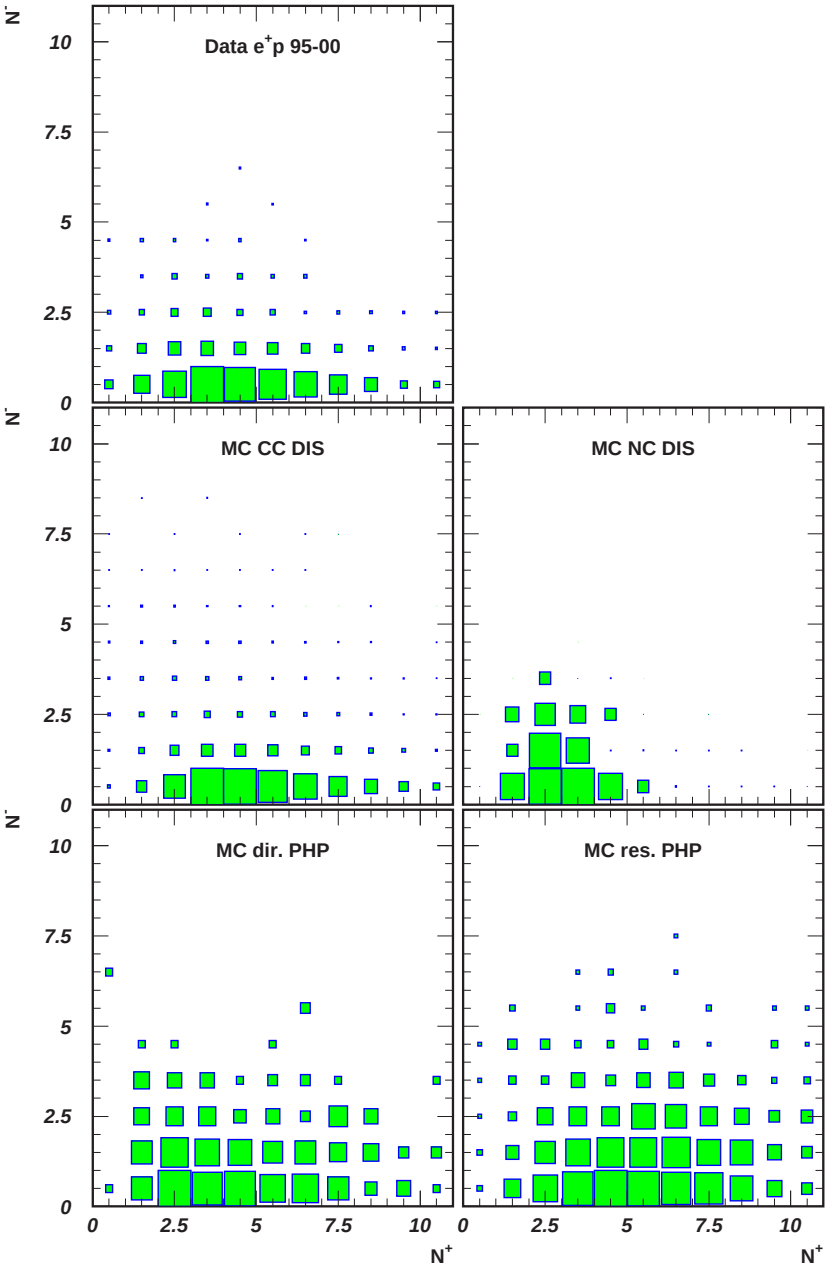


Figure 6.21: *Scatter plots of N^{+}_{track} versus N^{-}_{track} for data and MC samples selected with already discussed criteria.*

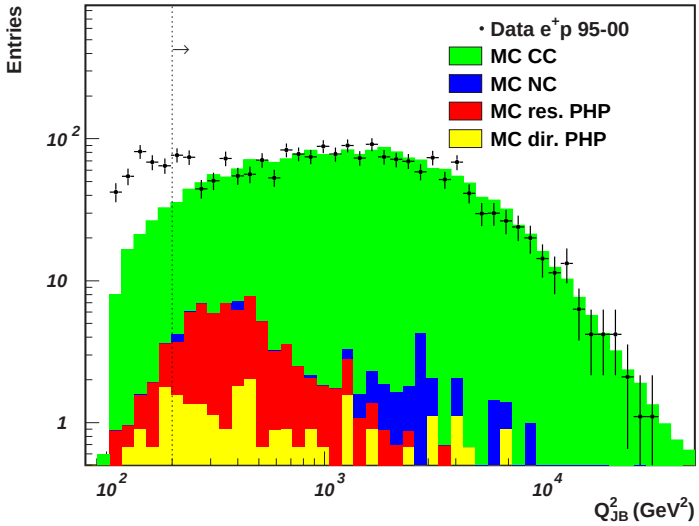


Figure 6.22: *Distributions of Q^2_{JB} for data and MC samples selected with already discussed criteria. The cut is indicaded by the line with an arrow. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

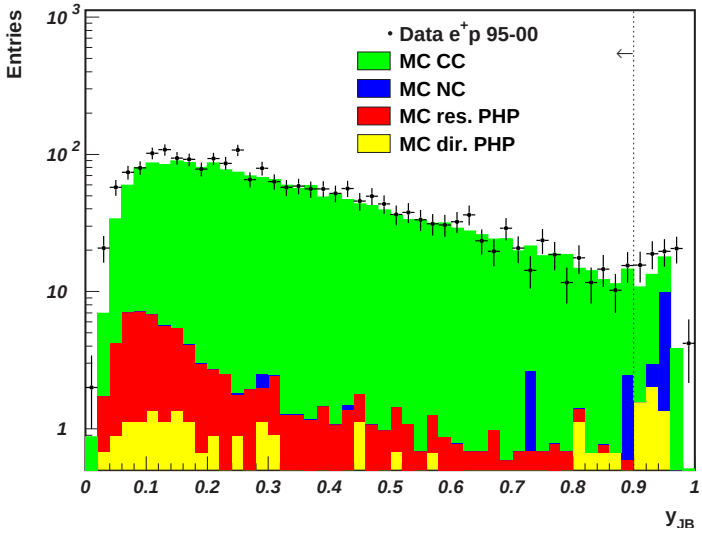


Figure 6.23: *Distributions of y_{JB} for data and MC samples selected with already discussed criteria. The cut is indicaded by the line with an arrow. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

6.9 Final CC sample

After proceeding with all discussed event criteria, the hard p_T cut (6.7) is applied to achieve final CC samples. The effect of the hard p_T cut can be verified in Figure 6.24. Comparing Figure 6.7, the cut efficiencies for the data and the CC MC samples now behave very similarly and the cut at 12 GeV achieves a high efficiency for both.

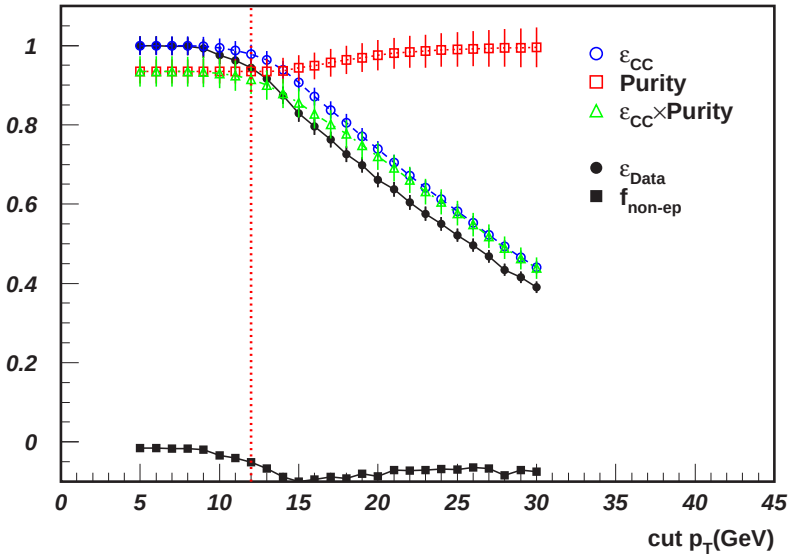


Figure 6.24: *Cut efficiencies and purities as functions of p_T cut positions for final samples. Explanation for variables shown on the plot can be found in Section 6.4. The final p_T cut-position is indicated by the line.*

The numbers of events in final data and MC samples are listed in Table 6.2. There are 1763 events left in the final data sample, in agreement with the sum of the MC samples. With the MC simulation, the remaining NC background is very little, less than 1%, whereas the PHP background is relatively large, about 4%. Comparing to the large statistical uncertainty discussed later, both contamination are negligible.

All the discussed criteria are summarized in Table 6.3. For each cut except those classified into preselection, number of selected events and efficiencies for data and MC are obtained from preselected samples. The number of events cut exclusively by one criterion is obtained by comparing the final data sample with the sample using all criteria except that one or directly correlated ones. These exclusively selected samples are also used to get the final control plots, shown in Figure 6.25. Very good agreement between data and MC simulations can be observed in these control plots.

Table 6.2: *Number of events in final data and MC samples.*

Sample		N_{final}
Data		1763
MC	CC DIS	1733.8
	NC DIS	3.6
	resolved PHP	20.3
	direct PHP	50.7
	Sum	1808.4

Table 6.3: *Summary of CC event criteria. Number of selected events, N_{sel}, selecting efficiencies for data, ϵ_{Data}, and for CC MC, ϵ_{MC}, are obtained with respect to samples selected by EVTAKE and DST bit for cuts of preselection, and with respect to preselected samples for other cuts. Number of exclusively cut events, N_{cut}, is obtained by excluding a cut and comparing the resulted data sample with the final data sample.*

Criterion		N_{sel}	N_{cut}	ϵ_{Data}	ϵ_{MC}
DST 34 and EVTAKE	Preselection	2579719	–	–	–
$p_T > 10\,\text{GeV}$		1470649	–	57.0%	93.2%
$\lvert z_{\text{vertex}}\rvert < 50\,\text{cm}$		2044364	32	79.2%	95.4%
$N_{\text{track}}^{\text{good}} \geq 3$		161060	22	6.2%	82.7%
Spark and dead PM		31102	8	97.4%	100%
CAL timing		26092	140	81.7%	100%
$p_T^{-ir} > 0.9 p_T - 6\,\text{GeV}$		26320	30	82.4%	98.9%
$N_{\text{track}}^{\text{good}} > (N_{\text{track}}^{\text{all}} - 20)/4$		26052	4	81.5%	99.8%
$p_T/E_T > 0.5$		8163	442	25.6%	85.9%
$\Delta\phi(p_T, p_T^{\text{CTD}}) < 30^\circ$		11052	358	34.6%	92.0%
N_{track}^{-} vs N_{track}^{+}		13887	341	43.5%	82.6%
$y_{\text{JB}} < 0.9$		28092	50	87.9%	97.7%
$Q_{\text{JB}}^2 > 200\,\text{GeV}^2$		18731	61	58.6%	93.5%
$p_T > 12\,\text{GeV}$		13866	104	43.4%	92.0%

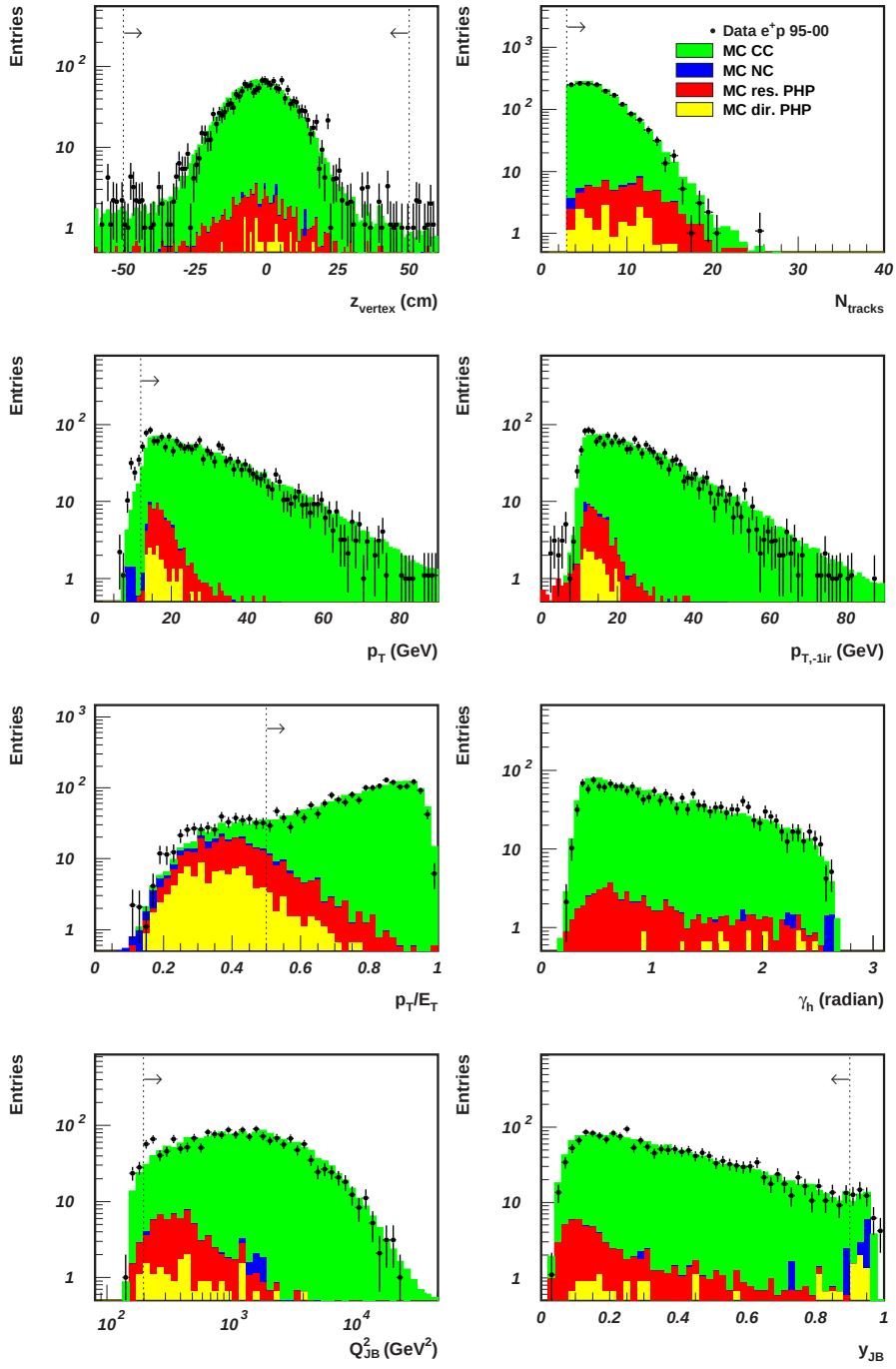


Figure 6.25: *Control plots for final samples. Each plot is obtained by using all cuts except the one related to the plot. The corresponding cuts are indicaded by the lines with arrows. The dots are for data, while histograms are for different MC samples, distinguished by different degrees of gray as shown, and ploted on top of each other.*

Chapter 7

D^* identification

The charm quark is tagged by the existence of charged D^* mesons in the final CC data sample. The D^* CC MC sample with ARIADNE simulating the high order QCD process is used to guide the tagging process and obtain the acceptance, while the alternative sample using MEPS is for systematic check. Because of the statistical limit in the data sample, only the MC results are shown in this chapter.

The charged D^* meson is reconstructed with a combinatorial method that is discussed in Section 7.1. Different methods for extracting the number of D^* signals are studied in Section 7.2. Separation of positively and negatively charged D^*s is discussed in Section 7.3.

7.1 Combinatorial method

The charged D^* is reconstructed through the so-called "golden channel",

$$\begin{aligned} D^{*+} &\to D^0 \pi_s^+ \\ &\hookrightarrow K^- \pi^+, \ c.c., \end{aligned} \tag{7.1}$$

where π_s^+ denotes the *slow* pion. The mass difference of charged D^* and D^0, $m_{D^{*+}} - m_{D^0} \approx 145\,\text{MeV}$, is only a little higher than the mass of charged pion, $m_{\pi^+} \approx 140\,\text{MeV}$, that gives very limited phase space for the pion decayed directly from D^*. In the rest frame of D^*, the momentum of the directly decayed pion is only about 40 MeV, hence it is called the "slow" pion. The choice of the D^0 decay has an advantage of high acceptance, because only two charged particles are required for reconstructing one D^0. Another decay mode, $D^0 \to K^-\pi^+\pi^+\pi^-$, has also been studied. Its branching ratio is nearly twice as large as that of the mode in the golden channel, but it needs four charged particles for D^0 reconstruction, hence has a relatively small

acceptance and eliminates the benefit of high branching ration. On the other hand, this channel suffers from much high combinatorial background, hence is not used.

Table 7.1: *Charged-track combinations for D^* reconstruction of the decay chain $D^{*+} \to D^0\pi_s^+$ with $D^0 \to K^-\pi^+$, and the charge-conjugated channel. All possible combinations of three charged tracks are listed. Wrong-charge combinations are for estimating the combinatorial background, while other kinds of combinations are not used.*

Combination	K	π	π_s
$D^{*\pm}$	$\mp$	$\pm$	$\pm$
wrong-charge	$\mp$	$\mp$	$\pm$
non-in-use	$\pm$	$\mp$	$\pm$
	$\pm$	$\pm$	$\pm$

With the ZEUS detector, charged tracks can be reconstructed very well in the CTD. A charged-track combinatorial method is therefore used for distinguishing D^*s in the golden channel. Pairs of oppositely-charged tracks are first combined to form a D^0 candidate. The tracks are alternatively assigned the masses of a charged kaon and a charged pion. An additional slow track, with charge opposite to that of the "kaon" track, is assigned the pion mass and combined with the D^0 candidate to form a D^* candidate. The combinations are summarized in Table 7.1 along with other possible ones. Given the reconstructed invariant masses of D^*, $M_{K\pi\pi_s}$, and D^0, $M_{K\pi}$, the spectrum of mass difference, $\Delta M = M_{K\pi\pi_s} - M_{K\pi}$, can be inspected for the D^* signal. The signal regions for $M_{K\pi}$ and ΔM are defined around their nominal values [1]:

$$1.80 \quad < M_{K\pi} < \quad 1.92 \ \text{GeV}, \tag{7.2}$$

$$0.143 \quad < \Delta M < \quad 0.148 \,\text{GeV}. \tag{7.3}$$

The number of D^*s can be extracted from the signal regions with an appropriate method to estimate the combinatorial background. A scatter plot of $M_{K\pi}$ versus ΔM for the D^* MC sample is shown in Figure 7.1. Clear signals can be observed in the mass signal regions.

Only good tracks defined in (6.10) are used for the combinations. The cut on transverse momenta of these tracks is therefore a cut on transverse momenta of kaons and pions,

$$p_T^{K,\pi,\pi_s} > 0.2\,\text{GeV}. \tag{7.4}$$

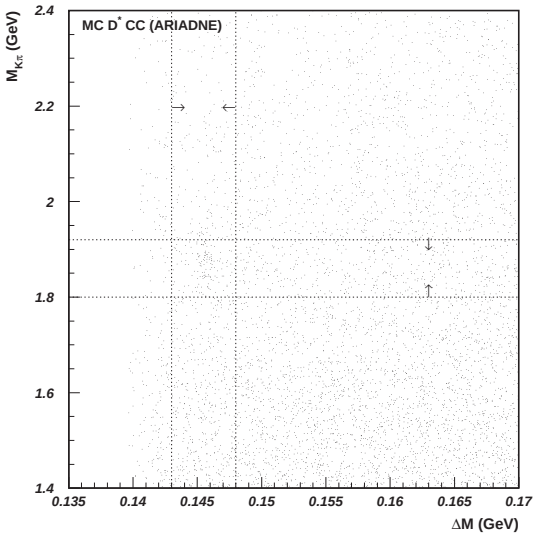


Figure 7.1: *Scatter plot of $M_{K\pi}$ versus ΔM for the D^* CC MC sample. The mass signal regions are indicaded by the lines with arrows.*

The higher threshold for fast kaons and pions, e.g. $p_T^{K,\pi} > 0.4\,\text{GeV}$, implemented in analyses of D^* production in PHP and NC DIS, is mainly for reducing combinatorial background, which is not the case for this analysis, therefore is not used. The spectra of transverse momenta of kaons and pions associated with D^*s in the signal regions are shown in Figure 7.2. The cut (7.4) affects mainly the slow pions, but does little on the fast kaons and pions.

As discussed in Section 5.1.3, D^*s coming from direct ep scattering favor a phase space differing from that of D^* produced via other processes, e.g. fragmentation. Cuts on D^*s' kinematics,

$$\begin{aligned} p_T^{D^*} &> 1.5\,\text{GeV}, \\ |\eta^{D^*}| &< 1.5, \end{aligned} \tag{7.5}$$

are hence required, as shown in Figure 7.3. The kinematic cut mainly affects the high η^{D^*} region, namely the forward region, where D^*s from fragmentation are dominant.

The spectrum of ΔM is obtained by restricting $M_{K\pi}$ in its signal region, as shown in Figure 7.4. The background is mainly due to the arbitrary assignment of masses to tracks used for the combinations, hence called *combinatorial background*. MC studies have shown that the contribution of other D^0 decay modes to the ΔM peak is small and can be neglected. A binned

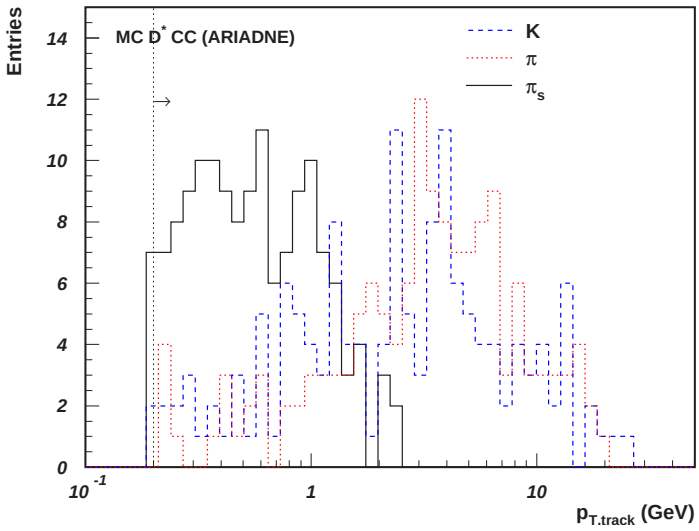


Figure 7.2: *Transverse momenta of K, π and π_s for D^* candidates in the mass signal regions. The cut is indicaded by the line with an arrow. The solid-line histogram is for π_s, the dashed for K and the dotted for π.*

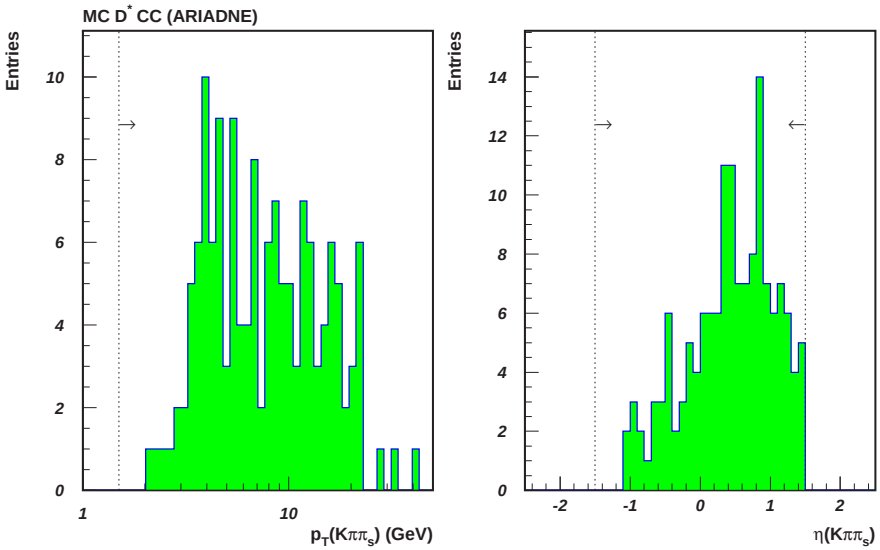


Figure 7.3: *Distributions of D^* kinematics $p_T^{D^*}$ (left) and η^{D^*} (right) for the D^* CC MC sample. The signal regions are indicaded by the lines with arrows.*

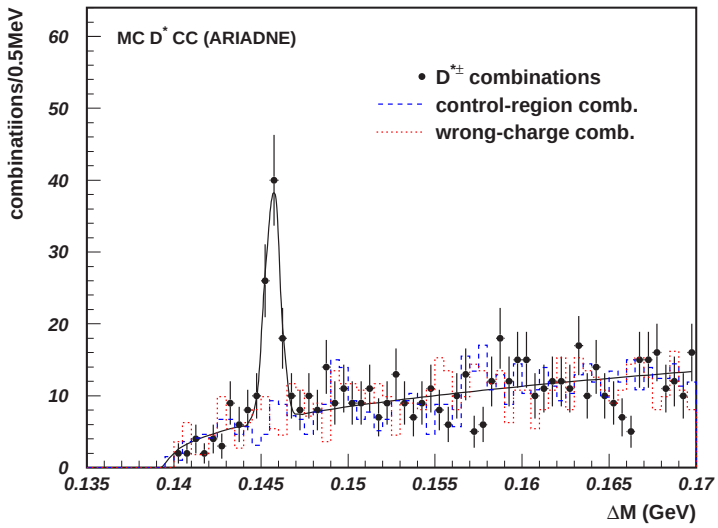


Figure 7.4: *Distribution of ΔM for D^* candidates of the D^* CC MC sample (solid dots). The ΔM distributions from control-region (dashed) and from wrong-charge (dotted) combinations, normalized in the region* $0.150 < \Delta M < 0.170$ *GeV, are shown in histograms. The solid line shows the result of the fit described in the text.*

maximum-likelihood fit is performed for the ΔM distribution. The fit function has the form

$$F(\Delta M) = P_1 e^{-\frac{(\Delta M - P_2)^2}{2P_3^2}} + P_4(\Delta M - m_\pi)^{P_5}, \tag{7.6}$$

with P_1 to P_5 being free parameters, and m_π the charged pion mass. The Gaussian term is used to fit the mass peak, and the other term to fit the combinatorial background. The fitted mass-difference, 145.66 ± 0.09 MeV, is consistent with the PDG value [1], 145.421 ± 0.010 MeV, and the fitted width, 0.43 ± 0.09 MeV, is consistent with the experimental resolution. The corresponding $M_{K\pi}$ spectrum can be used as a check, as shown in Figure 7.5. The bump below the D^0 peak mainly originates from the decay $D^0 \to K^-\pi^+\pi^0$ in which the neutral pion is not reconstructed. A fit with the form,

$$F(M_{K\pi}) = P_1 e^{-\frac{(M_{K\pi} - P_2)^2}{2P_3^2}} + e^{P_4 + P_5 M_{K\pi}}, \tag{7.7}$$

is performed in the limited region, $M_{K\pi} > 1.7$ GeV, and similar agreements are observed in the fit results.

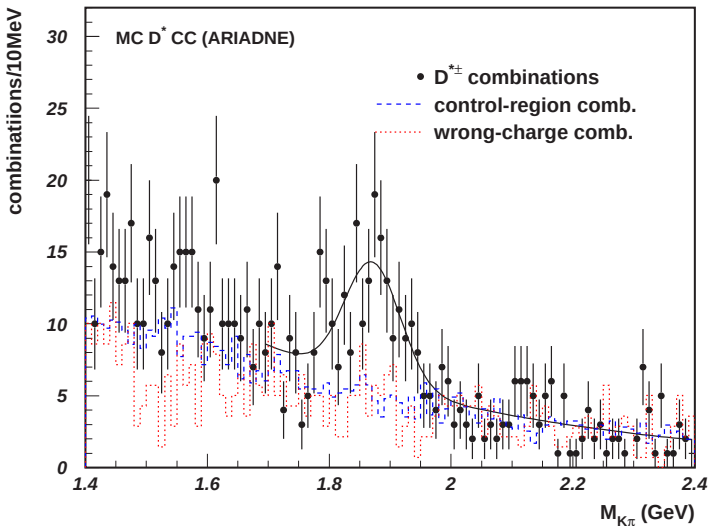


Figure 7.5: *Distribution of $M_{K\pi}$ for D^* candidates of the D^* CC MC sample (solid dots). The $M_{K\pi}$ distributions from the control-region (dashed) and from wrong-charge (dotted) combinations, normalized in the region $2.0 < M_{K\pi} < 2.4$ GeV, are shown in histograms. The solid line shows the result of the fit described in the text.*

7.2 Signal extraction

There are two general ways to determine the number of D^*s under the signal peak in Figure 7.4: *fit* and *background subtraction*. While the former method is clear, the latter subtracts the background from the signal region by estimating the shape of the combinatorial background properly.

With the fit function (7.6) and binned fitting, the number of signals, N_{signal}, under the Gaussian peak is given by

$$N_{\text{signal}} = \frac{\sqrt{2\pi} P_1 P_3}{w_{\text{bin}}}, \tag{7.8}$$

with w_{bin} denoting the bin width.

For background subtraction, several methods exist to estimate the combinatorial background shape. In Figure 7.4, results of two methods, wrong-charge combinations and $M_{K\pi}$ control-region combinations, are shown. In the former method, as listed in Table 7.1, two same-charged tracks are combined for one D^0 and an additional track with the opposite charge of the "kaon" is added to form a D^*. The background shape is then obtained by the faked ΔM distribution. With the latter method, the combinatorial background is

obtained by restricting $M_{K\pi}$ in the *control region*:

$$2.0 < M_{K\pi} < 2.4\,\text{GeV}. \tag{7.9}$$

In Figure 7.4, the background shapes obtained with the two methods are normalized to the signal distribution in the *normalization region*,

$$0.150 < \Delta M < 0.170\,\text{GeV}, \tag{7.10}$$

and can describe the background very well. Defining the number of combinations in the signal region, N_{sr}, and the number of combinations in the normalization region, N_{nr}, for the signal distribution, and correspondingly, N'_{sr} and N'_{nr} for the background distribution, the number of D^* signals can be given by

$$N_{\text{signal}} = N_{\text{sr}} - N'_{\text{sr}} \cdot \frac{N_{\text{nr}}}{N'_{\text{nr}}}, \tag{7.11}$$

with the squared statistical error

$$\sigma^2_{N_{\text{signal}}} = N_{\text{sr}} + \left(N'_{\text{sr}} \cdot \frac{N_{\text{nr}}}{N'_{\text{nr}}}\right)^2 \left(\frac{1}{N'_{\text{sr}}} + \frac{1}{N_{\text{nr}}} + \frac{1}{N'_{\text{nr}}}\right), \tag{7.12}$$

while the estimation of the combinatorial background is

$$N_{\text{background}} = N'_{\text{sr}} \cdot \frac{N_{\text{nr}}}{N'_{\text{nr}}}, \tag{7.13}$$

with the squared error

$$\sigma^2_{N_{\text{background}}} = \left(N'_{\text{sr}} \cdot \frac{N_{\text{nr}}}{N'_{\text{nr}}}\right)^2 \left(\frac{1}{N'_{\text{sr}}} + \frac{1}{N_{\text{nr}}} + \frac{1}{N'_{\text{nr}}}\right). \tag{7.14}$$

Similarly, all methods can be used for the $M_{K\pi}$ signal distribution in Figure 7.5 with (7.9) being the normalization region and (7.10) for the ΔM control-region combinations. When applicable, the fit method usually gives better statistical errors than the subtraction method. In the case of small statistics, however, the subtraction method is prefered, if not the only choice.

The choice of the normalization region (7.10) has been studied with the D^* MC sample. As shown in Figure 7.6, the upper edge of the normalization region, $\Delta M_{\text{nr,upper}}$, corresponding to the choice of the ΔM window range, is varied and the numbers of D^* signals are plotted with respect to the variation for the two subtraction methods respectively. Also shown in the plot as a verification is the signal numbers obtained with the fit method, which is not sensitive to the choice of the ΔM window. The results of

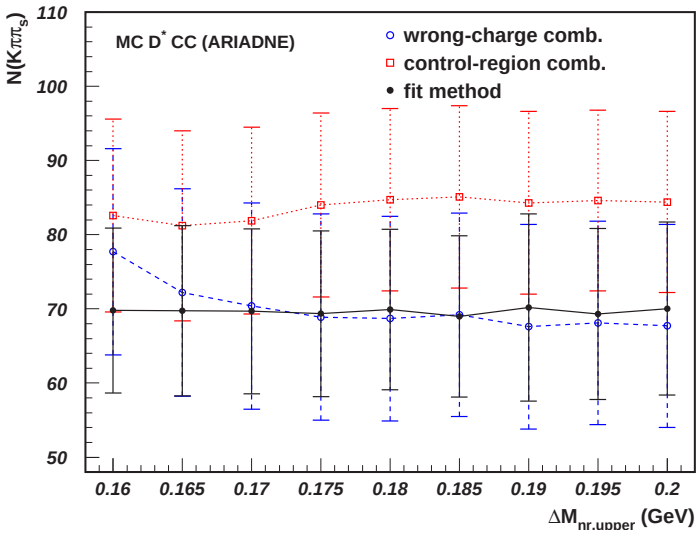


Figure 7.6: *Numbers of D^* candidates as a function the upper edge of the ΔM normalization region, $\Delta M_{nr,upper}$, with methods of wrong-charge (open circles) and control-region (open squares) combinations as well as the fit method (solid dots). Vertical lines with bars represent statistical errors. The D^* CC MC sample is used.*

all methods are in good agreement within errors. The $M_{K\pi}$ control-region method gives results systematically higher than other methods, because the $M_{K\pi}$ control region is chosen only above the signal region that systematically selects tracks with high momenta for D^0 reconstruction and underestimates the number of combinations in the ΔM signal region. With the wrong-charge combinations, the results become much more stable as $\Delta M_{\mathrm{nr,upper}}$ gets larger than 0.170 GeV, which is hence chosen. Similar studies have also been made for the $M_{K\pi}$ case, hence the region (7.9) chosen.

7.3 D^* charge separation

Unlike charm production in NC or PHP, charm production in CC is charge asymmetric. In the e^+p CC reaction, only the charm quark, c, can be produced by the hard scatter directly. The anti-charm quark, $\bar{c}$, can only be produced through the processes of fragmentation, or decay of a beauty quark. These processes, however, will produce both c and $\bar{c}$ with equal probability. Therefore, number of hard-scatter charm quark can be obtained statistically by subtracting number of anti-charm quarks from total number of charm quarks. The conclusion can be translated directly in terms of charged D^*.

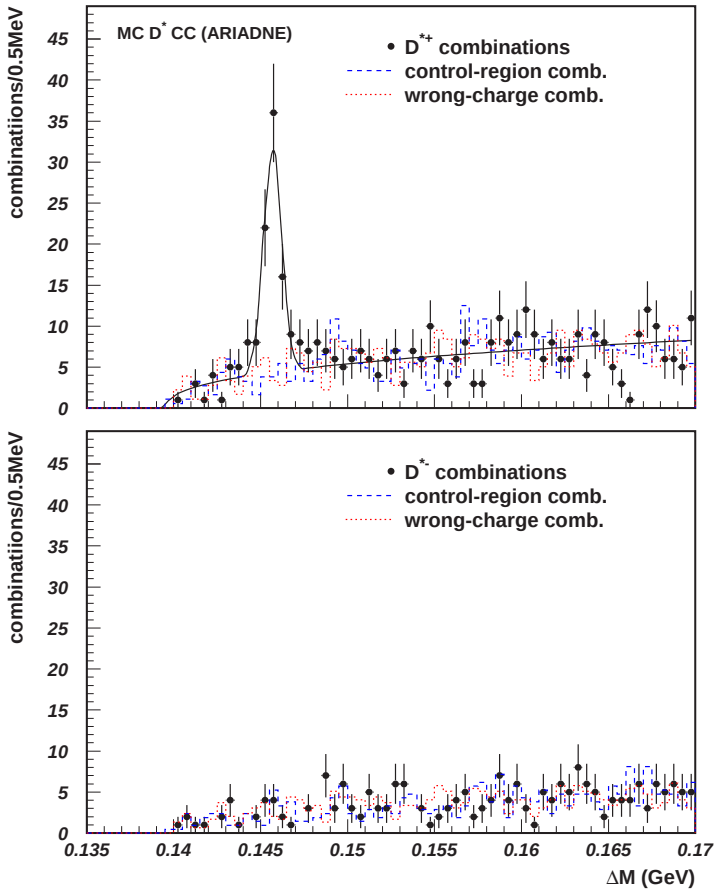


Figure 7.7: *Charge-separated ΔM distributions for D^{*+} (up) and D^{*-} (low) candidates in the D^* CC MC sample. The solid line in the D^{*+} distribution shows the result of the fit described in the text. The dots are for right-charge combinations and the histograms for control-region (dashed) and wrong-charge (dotted) combinations.*

With the combinatorial method, reconstructed D^* samples may be separated in positively and negatively charged samples distinguished by the charge of the slow pions. The ΔM distributions of D^{*+} and D^{*-} for the D^* CC MC sample are shown in Figure 7.7, together with the normalized control-region and wrong-charge combinations. The spectrum of D^{*-} is flat and no peak can be observed in the signal region, whereas a sharp peak is observed for D^{*+}. The observation implies that the contribution of the anti-charm quark in the MC simulation is negligible and the corresponding ΔM distribution can be used as an estimation for that of the charm quark.

Chapter 8

Results and discussions

The charm production can be derived from the measurement of the D^* cross section. The procedure to determine the cross section is discussed in Section 8.1. The numeric results of the cross sections are given in Section 8.2, while the systematic uncertainties are given in Section 8.3. The results are discussed at length and conclusions are addressed in Section 8.4.

8.1 Procedure to determine the cross section

The D^* is reconstructed in the kinematic regions (7.5) and (6.19). Therefore, only the *visible* cross section for the D^* production, $\sigma_{\text{vis}}^{D^*}$, is measured. This integrated cross section is given by the form

$$\sigma_{\text{vis}}^{D^*} = \frac{N_{\text{meas}}^{D^*}}{\mathcal{A} \times \mathscr{L} \times \mathcal{B}}, \tag{8.1}$$

where $N_{\text{meas}}^{D^*}$ is the number of D^* candidates measured, $\mathcal{A}$ the acceptance and $\mathcal{B}$ the product of the appropriate branching ratios for D^* and D^0 [1],

$$\mathcal{B} = \mathcal{B}(D^* \to D^0\pi_s^+) \times \mathcal{B}(D^0 \to K^-\pi^+) = (2.57 \pm 0.06)\%. \tag{8.2}$$

As discussed in Section 6.3, the e^+p data sample used in this analysis corresponds to an integrated luminosity

$$\mathscr{L} = 110\,\text{pb}^{-1}. \tag{8.3}$$

The overall uncertainty on the luminosity measurement is 2.0%, seeing Table 3.1, and is taken into account for calculating the systematic uncertainty.

The acceptance is determined from the D^* CC MC sample. Given the number of D^*'s generated in the true kinematic regions, $N_{\text{gen}}^{D^*}$, and the number

of D^*'s measured in the reconstructed kinematic regions, $N^{D^*}_{\text{meas}}$, the acceptance is defined as

$$\mathcal{A} = \frac{N^{D^*}_{\text{meas}}}{N^{D^*}_{\text{gen}}}. \tag{8.4}$$

In general, the acceptance is not necessarily less than one, since reconstructed kinematics are usually not the same as generated ones.

Since no anti-charm quarks are expected to appear in the hard scatter of CC e^+p interactions, the charm production can be derived from the cross section for the hard-scatter D^{*+}. For this purpose, D^*-charge separation should be taken into account in the previous procedure. In addition, the visible cross section should be extrapolated to the full D^* phase space, which can be done with MC. The extrapolation factor is defined by

$$\mathcal{C}_{\text{ext}} = \frac{N^{\text{full}}_{\text{gen}}}{N^{\text{kine}}_{\text{gen}}}, \tag{8.5}$$

where the denominator is the number of D^*'s generated in the restricted kinematic region, and the numerator that in the full region.

Using the hadronization fraction of the charm quark to D^{*+} from the e^+e^- experiment [49],

$$f(c \to D^{*+}) = 0.222 \pm 0.014, \tag{8.6}$$

the charm cross section in the study is then given by

$$\sigma^{e^+p\to\bar{\nu}_e cX} = \frac{\sigma^{D^{*+}}_{\text{vis}} \times \mathcal{C}_{\text{ext}}}{f(c \to D^{*+})}. \tag{8.7}$$

8.2 The results

Results of the D^* CC MC samples for determining the acceptance as well as the extrapolation factor are given in Table 8.1. Since the D^{*-} is not of interest in this study and has only a small fraction in the MC simulation, its acceptance is not considered. Only the measurement of D^{*+} production will be converted into charm production, hence the extrapolation factor is calculated just for this. The number of hard-scatter D^{*+}s generated can be estimated by subtracting the number of D^{*-}s from the number of all D^{*+}s, assuming D^{*+}s and D^{*-}s are produced equally in processes other than the hard scatter. The numbers of measured D^*s are obtained with the fit method. For D^*s without charge separation, the acceptance is

$$\mathcal{A}^{D^*} = (15.5 \pm 2.4)\%, \tag{8.8}$$

while for D^{*+} only, the acceptance is

$$\mathcal{A}^{D^{*+}} = (16.7 \pm 2.7)\% \tag{8.9}$$

and the extrapolation factor is

$$\mathcal{C}_{\rm ext}^{D^{*+}} = 2.10, \tag{8.10}$$

with a negligible, hence not shown, statistical error.

Table 8.1: *MC results for determining the acceptance and extrapolation factor. The numbers of generated D^*s in the full D^* kinematic region, N_{gen}^{full}, and in the restricted region, N_{gen}^{kine}, are counted for all decay modes, while $N_{gen}^{K\pi\pi_s}$ is for the golden decay mode in the restricted region. The number of measured D^*, $N_{meas}^{K\pi\pi_s}$, is obtained by the fit method. Calculation for the $\mathcal{A}$ and $\mathcal{C}_{ext}$ is explained in the text.*

D^* CC MC		$N_{\rm gen}^{\rm full}$	$N_{\rm gen}^{\rm kine}$	$N_{\rm gen}^{K\pi\pi_s}$	$N_{\rm meas}^{K\pi\pi_s}$	$\mathcal{A}\,(\%)$	$\mathcal{C}_{\rm ext}$
ARIADNE	$D^{*\pm}$	–	–	449	69.5 ± 10.7	15.5 ± 2.4	–
	D^{*+}	40437	18910	438	71.4 ± 11.5	16.7 ± 2.7	2.10
	D^{*-}	1696	473	11	–	–	–
MEPS	$D^{*\pm}$	–	–	436	89.0 ± 11.7	20.4 ± 2.7	–
	D^{*+}	40339	18918	430	90.1 ± 11.1	21.3 ± 2.6	2.09
	D^{*-}	1782	435	6	–	–	–

Table 8.2: *The numbers of combinations for extracting the number D^* candidates with different background-subtraction methods. The D^* CC MC samples are not normalized to the data. Details are explained in the text.*

Combination		Data		MC (ARIADNE)		MC (MEPS)	
		$N_{\rm signal}$	$N_{\rm norm}$	$N_{\rm signal}$	$N_{\rm norm}$	$N_{\rm signal}$	$N_{\rm norm}$
right-charge	D^{*+}	3	13	124	272	128	277
	D^{*-}	4	15	21	165	13	169
control-region	D^{*+}	4	30	73	500	102	536
	D^{*-}	5	32	46	346	30	365
wrong-charge		2	32	80	486	70	502
wrong-charge ($M_{K\pi}$)		2	11	80	166	70	132

Table 8.3: *The numbers of D^* candidates determined with the different background-subtraction methods. The D^* CC MC samples are not normalized to the data. Details are explained in the text.*

Method		Data	MC (ARIADNE)	MC (MEPS)
control-region	$N^{D^{*+}}$	1.3 ± 2.1	84.3 ± 12.4	75.3 ± 13.1
	$N^{D^{*-}}$	1.9 ± 2.4	-0.9 ± 6.0	-0.9 ± 4.6
wrong-charge	$N^{D^{*+}}$	2.3 ± 1.9	79.2 ± 12.7	89.4 ± 12.6
	$N^{D^{*-}}$	3.3 ± 2.2	-6.2 ± 6.0	-10.6 ± 5.0
wrong-charge	N^{D^0}	2.3 ± 1.9	88.8 ± 12.8	73.9 ± 14.9
	$N^{\bar{D}^0}$	3.2 ± 2.2	-1.2 ± 6.4	-2.9 ± 5.2

Charge-separated distributions of ΔM and $M_{K\pi}$ for data are shown in Figure 8.1 and Figure 8.2. The events in the mass signal region are shown with the ZEUS event display in Appendix B. The statistics are very low, and no peaks can be observed in the signal regions. For estimating the number of D^* candidates, only the background subtraction methods are applicable.

Table 8.2 lists the numbers of charge-separated combinations in signal and normalization regions of ΔM distribution for combinations of right-charge, control-region and wrong-charge, defined in Section 7.2. Since the ΔM normalization region is also for getting the $M_{K\pi}$ distribution with the control-region method, and vice verse, the only different numbers for the $M_{K\pi}$ distribution are of wrong-charge combinations, which are listed in Table 8.2 as well. Also listed are the corresponding numbers for the two D^* CC MC samples.

As discussed in Section 7.2, the wrong-charge-combination method has better performance than the control-region-combination method. Although D^*-charge separation makes no sense for wrong-charge combinations, it can still be used by assuming the same behavior of positive and negative tracks in the detector[1]. The control-region method as well as the wrong-charge method for $M_{K\pi}$ distribution are used as cross checks. The numerical results are listed in Table 8.3. For the MC samples, the results of $D^{*\pm}$ and D^{*+} are in agreement with those obtained by the fit method, see Table 8.1, while the results of D^{*-} are consistent with zero. All results obtained with the two background-subtraction methods and the wrong-charge method for $M_{K\pi}$

[1]This is not really true. With the arrangement of cells, the CTD has higher efficiency for positive tracks than for negative tracks, and the difference becomes large for tracks with low transverse momenta, which implies a higher efficiency for D^{*+} than for D^{*-} since the D^* reconstruction is sensitive to the slow pion.

distribution agree within the error.

Without charge separation, the number of D^* candidates in the e^+p is

$$N^{D^*}_{\text{meas}} = 5.6 \pm 3.0. \tag{8.11}$$

Combining with the acceptance (8.8), the visible cross section for D^* is

$$\sigma^{D^*}_{\text{vis}} = 12.8 \pm 7.1\,\text{pb}, \tag{8.12}$$

with statistical error. Considering the charge separation, the numbers of D^{*+}s and D^{*-}s are

$$N^{D^{*+}}_{\text{meas}} = 2.3 \pm 1.9 \tag{8.13}$$

and

$$N^{D^{*-}}_{\text{meas}} = 3.3 \pm 2.2, \tag{8.14}$$

with the statistical errors. Neglecting the D^{*-}, the result (8.13) can be combined with the acceptance (8.9) and the visible cross section for D^{*+} is

$$\sigma^{D^{*+}}_{\text{vis}} = 4.9 \pm 4.0\,\text{pb}. \tag{8.15}$$

Combining with the extrapolation factor (8.10), the cross section for charm quark production is derived as

$$\sigma^{e^+p\to\bar{\nu}_e cX} = 46.3 \pm 38.2\,\text{pb}, \tag{8.16}$$

with a large statistical error of about 83%. Taking D^{*-} into account, its distribution consists of contributions from physics processes other than the e^+p hard scatter as well as from the combinatorial background. Using the same technique of background subtraction, the number of hard-scatter D^{*+}s can be obtained,

$$N^{D^{*+}}_{\text{hard}} = -0.6 \pm 2.9. \tag{8.17}$$

It is consistent with zero with a very large statistical error. Therefore, it is more appropriate to set a limit than to determine a value for the cross section. The *Bayesian* approach is used for this purpose, referring to Appendix C for technical details. The observed D^{*+} combinations in the ΔM signal region is

$$n = 3, \tag{8.18}$$

and the background estimated from the D^{*-} distribution is

$$b = 3.7 \pm 2.3, \tag{8.19}$$

calculated according to (7.13) and (7.14). The upper-limit ($s_{\rm ul}$) for the number of D^{*+} candidates (s) at 90% confidence level (CL) is

$$s < s_{\rm ul} = 5.4, \qquad \text{at } 90\% \text{ CL}. \tag{8.20}$$

With the Bayesian approach, conversion of the upper-limit on the number of events to that on the cross section is straightforward. The upper-limit on the cross section for the D^{*+} is

$$\sigma_{\rm vis}^{D^{*+}} < \sigma_{\rm ul}^{D^{*+}} = 11.5\,{\rm pb}, \qquad \text{at } 90\% \text{ CL}, \tag{8.21}$$

and for charm production,

$$\sigma^{e^+p\to\bar{\nu}_e cX} < \sigma_{\rm ul}^{e^+p\to\bar{\nu}_e cX} = 109\,{\rm pb}, \qquad \text{at } 90\% \text{ CL}. \tag{8.22}$$

As a cross check, the cross section for the CC reaction in the kinematic region (6.19) is

$$\sigma^{\rm CC} = 32.9 \pm 0.8\,{\rm pb}, \tag{8.23}$$

with the acceptance

$$\mathcal{A}^{\rm CC} = 46.8\%. \tag{8.24}$$

The error on the cross section is statistical, while the statistical error of the acceptance is neglected.

8.3 Systematic uncertainties

The systematic error is estimated according to the visible cross section for D^* without charge separation since its statistical error is better than that of the D^{*+} result. The major uncertainty comes from the MC model for determining the acceptance, as given in Table 8.1. The relative error on the acceptance by using the alternative D^* MC sample is 32%.

The systematic uncertainties caused by the event selection can be obtained by varying cut values up and down by a small factor, e.g. 10%, repeating the whole analyzing procedure and comparing results with the nominal one. The resulting uncertainties are summarized in Table 8.4. The overall systematic uncertainty is obtained by adding individual ones in quadrature, also shown in the table.

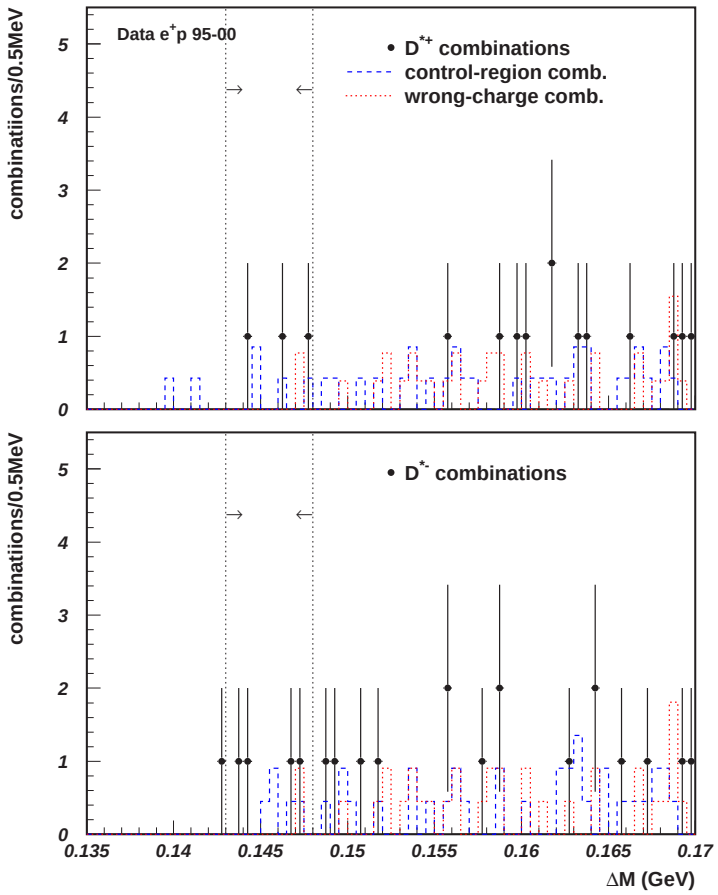


Figure 8.1: *Charge-separated ΔM distributions for data. The dots are for right-charge combinations and the histograms for control-region (dashed) and wrong-charge (dotted) combinations. The mass signal region is indicaded by the lines with arrows.*

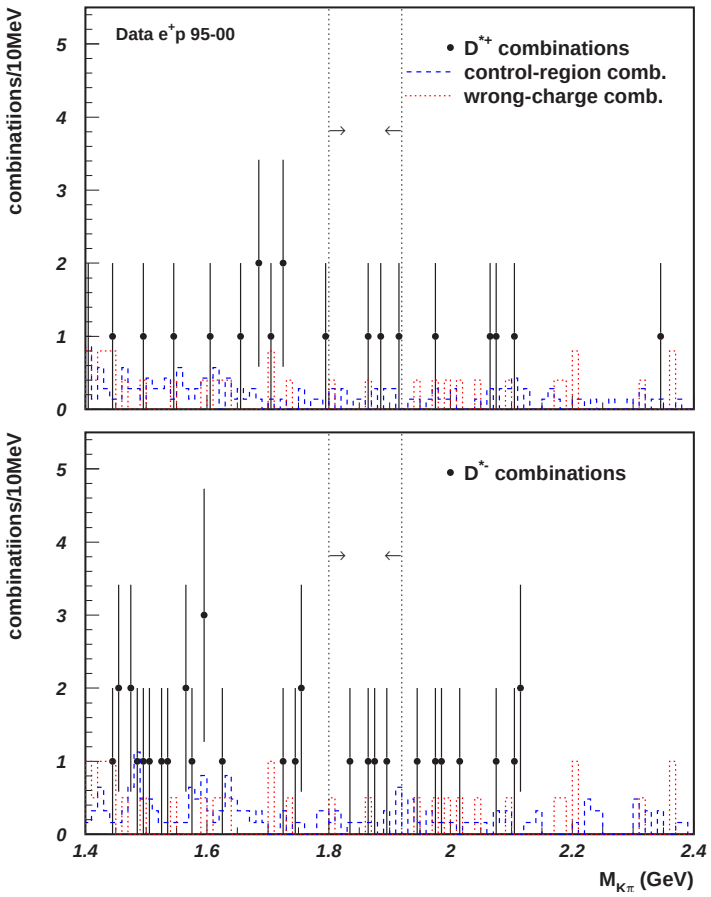


Figure 8.2: *Charge-separated $M_{K\pi}$ distributions for data. The dots are for right-charge combinations and the histograms for control-region (dashed) and wrong-charge (dotted) combinations. The mass signal region is indicaded by the lines with arrows.*

Table 8.4: *Systematic uncertainties of most important event criteria for the cross sections of D^* (σ^{D^*}) as well as CC (σ^{CC}). The uncertainty from the MC simulation is also listed for D^*. The total is the square root of the sum in quadrature.*

Criterion	Variation	$\Delta_{\sigma^{D^*}}/\sigma^{D^*}(\%)$	$\Delta_{\sigma^{\mathrm{CC}}}/\sigma^{\mathrm{CC}}(\%)$
$p_T > 12\,\mathrm{GeV}$	13 GeV	+2.3	−1.5
	11 GeV	+0.8	+0.9
$\lvert z_{\mathrm{vertex}}\rvert < 50\,\mathrm{cm}$	55 cm	−0.8	+0.3
	45 cm	0.0	−0.6
$N_{\mathrm{track}} \geq 3$	4	+3.1	−0.9
	2	0.0	+0.3
$p_T/E_T > 0.5$	0.55	+13	+0.6
	0.45	+11	−0.3
$\lvert\Delta\phi(p_T, p_T^{\mathrm{CTD}})\rvert < 30°$	35°	−2.3	+0.6
	25°	−2.3	−0.6
$p_T^{\pi_s} > 0.20\,\mathrm{GeV}$	0.30 GeV	−6.2	−1.5
	0.15 GeV	−9.4	+0.3
MC CC D^* with MEPS		+32	–
Total		$^{+37}_{-12}$	$^{+1.2}_{-2.6}$

8.4 Discussions and conclusions

A sizable D^* cross section with large statistical error, roughly 50%, has been measured in the CC process of e^+p DIS. The contribution of D^{*+} and D^{*-} are almost the same within the statistical error, while the latter is slightly higher. The number of hard-scatter D^{*+}s is obtained by subtracting the D^{*-} distribution under the assumption that the physics processes other than the hard scatter will produce charm and anti-charm quarks equally, and is consistent with zero within the statistical error. This observation is consistent with the expectation given in Section 2.4.

With the Standard Model, the charm production in CC e^+p DIS is charge asymmetric, namely only the charm and no anti-charm quark is produced in the hard process. The anti-charm quark can be produced in the soft process, i.e. fragmentation, in which the charm quark is usually produced with the same probability, too. The mechanism of fragmentation is not completely understood so far. However, it is commonly believed that heavy quark production is suppressed in the fragmentation process. With the Lund model, used for the MC simulation, the suppression is $u : d : s : c \approx 1 : 1 : 0.3 : 10^{-11}$ [69]. Even this is not true in reality, the phase space of charm quarks of fragmentation would be quite different from that of hard-scatter ones. The requirement on D^* kinematics (7.5) favors the charm production of interest. Therefore, the observed D^* cross section cannot be explained with the fragmentation process.

Another important source of open-charm production, at HERA's center-of-mass energy, is the decay of B-mesons, namely the beauty production. In the e^+p CC reaction, the dominant process for beauty production is the BGF process,

$$W^+g \to c\bar{b}, \tag{8.25}$$

whereas the process $W^+g \to t\bar{b}$ is eliminated by the HERA energy scale and the process $W^+g \to u\bar{b}$ is suppressed by the CKM matrix element $|V_{ub}|^2$. With $\bar{b} \to \bar{c}$, this process will produce $c\bar{c}$ pairs. The contribution of the BGF depends on the gluon's PDF in the proton. As discussed in Section 2.3, the BGF contribution (2.28) to charm production is comparable to the contribution of the leading-order process (2.26) in the high Q^2 region. On the other hand, a very recent measurement of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at high Q^2 has been performed by H1 [11] and the ratio of the $c\bar{c}$ and $b\bar{b}$ cross sections can be derived as:

$$\frac{\sigma^{c\bar{c}}}{\sigma^{b\bar{b}}} = 6.7 \pm 2.1, \tag{8.26}$$

in the kinematic range $Q^2 > 150\,\mathrm{GeV}^2$ and $0.1 < y < 0.7$, which is close to that used in this thesis. Therefore, the beauty contribution to the charm

production is considerable, and underestimated in the MC simulation. This might be a reason for the sizable D^* cross section, but cannot fully explain the equality of the observed numbers of D^{*+} and D^{*-} candidates.

There are also other potential processes, e.g. diffractive heavy flavor production and intrinsic charm in the proton, for the charm DIS production. They either have been measured to have small cross sections relative to the BGF, or are not expected to contribute in the relevant kinematic regions [17].

The charge dependence of the cross section for D^* production had been observed in early studies of electroproduction by ZEUS. An excess of about 18% was observed for D^{*-} candidates [79], and the reason is still unclear. A similar observation was also made at Fermilab [80].

In conclusion, the observed D^{*+}s and D^{*-}s are mainly due to the statistical fluctuation. A constraint on the strange sea in the proton from the measurement of the charm production in CC e^+p DIS cannot be made with the HERA I data. An integrated luminosity larger by several times than that available in this study is needed to perform such a measurement, hence the high luminosity HERA II data will have to be awaited for to yield a proper measurement.

Chapter 9

Summary

For the first time, the measurement of charm production in charged current deep inelastic positron-proton scattering has been investigated with the ZEUS detector at HERA in the kinematic region $Q^2 > 200\,\mathrm{GeV}^2$ and $y < 0.9$. The e^+p data used for this study were collected from 1995 to 2000, corresponding to an integrated luminosity of $110\,\mathrm{pb}^{-1}$. The charm quark is tagged by requiring the existence of at least one charged D^* in the kinematic range $p_T^{D^*} > 1.5\,\mathrm{GeV}$ and $|\eta^{D^*}| < 1.5$. D^* mesons are identified in the decay channel $D^{*+} \to D^0\pi_s^+$ with $D^0 \to K^-\pi^+$ and corresponding antiparticle decay, which are reconstructed with the charged-track combinatorial method.

The number of D^*s observed in the final data sample is

$$N^{D^*} = 5.6 \pm 3.0,$$

with the acceptance $\mathcal{A}^{D^*} = (15.5 \pm 2.4)\%$, where both errors are statistical. The observation corresponds to a visible cross section of

$$\sigma_{\mathrm{vis}}^{D^*} = 12.8 \pm 4.0^{+4.7}_{-1.5}\,\mathrm{pb},$$

where the first error is statistical and the second is systematic. The contribution of the D^{*+} is slightly smaller than that of D^{*-}, but both are almost the same within the large statistical errors. The number of D^{*+}s in the hard process is estimated by assuming that both D^{*+} and D^{*-} are produced with the same probability in other processes. The direct subtraction gives a result consistent with zero with very large statistical uncertainty. An upper-limit is hence determined with the Bayesian approach,

$$N_{\mathrm{hard}}^{D^{*+}} < 5.4, \qquad \text{at } 90\%\ \mathrm{CL}.$$

The upper-limit is converted to the charm cross section in the DIS kinematic region

$$\sigma^{e^+p\to\bar{\nu}_e cX} < 109\,\mathrm{pb}, \qquad \text{at } 90\%\ \mathrm{CL}.$$

The results have been verified from different aspects and have been discussed in possible physics processes.

To achieve these results, the event selecting for the CC reaction and for the D^* reconstruction with the combinatorial method have been studied carefully, and the procedure has been documented in detail. The MC simulation is crucial for this study, hence has been discussed at length. Although it is impossible to fully explain the ZEUS experiment and all aspects of offline event reconstruction and various corrections in a thesis, a brief description of them is given. Besides these, the phenomenology of the charm production in DIS as well as the experimental situation are also introduced in this thesis.

Appendix A

MC study at generator level

The MC simulation with DJANGOH has been studied carefully at the generator level. The QCD processes simulated within LEPTO are the leading order (LO) process

$$V^*q \to q \tag{A.1}$$

and the first order processes

$$V^*q \to qg, \tag{A.2}$$

the gluon radiaton or QCD Compton (QCDC) process, and

$$V^*g \to q\bar{q}, \tag{A.3}$$

the boson-gluon fusion (BGF), where V^* is the exchanged virtual boson γ/Z or W. Quark masses are not included in the simulation, but a threshold factor is applied for boson-gluon fusion into heavy quarks. The contribution of these processes is shown in Figure A.1, and the corresponding kinematic distributions in Figure A.2.

There is a special parameter, `LST(19)`, to control the evaluation of probabilities for the three processes. By default, `LST(19)=-10`, a pre-calculated grid is used. This configuration has an advantage in time consuming, but causes the problem as shown in Figure A.3, when using MEPS. However, this problem is absent in the simulation with ARIADNE, and will disappear for MEPS when the configuration is changed to some other way, for instance, `LST(19)=-1`, to evaluate the probabilities event by event.

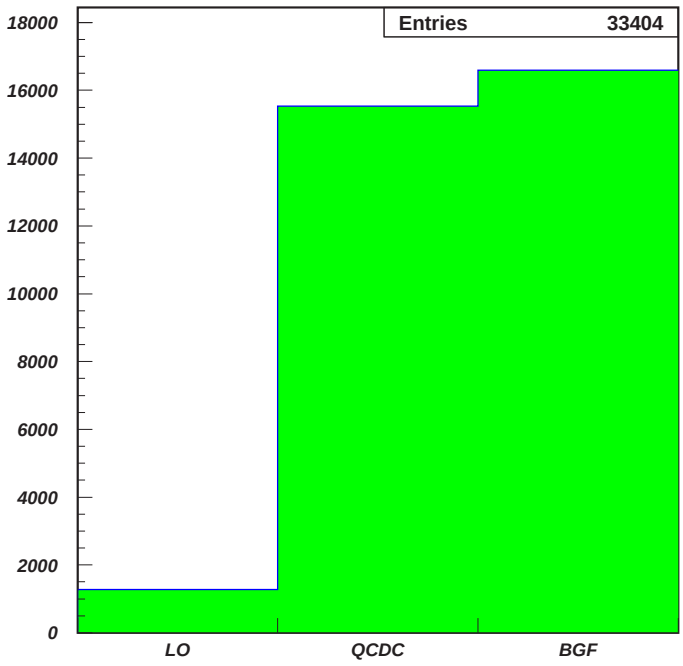


Figure A.1: *Distribution of QCD processes simulated with LEPTO. D^* enriched CC MC sample simulated with ARIADNE is used for illustration. Explanation can be found in the text.*

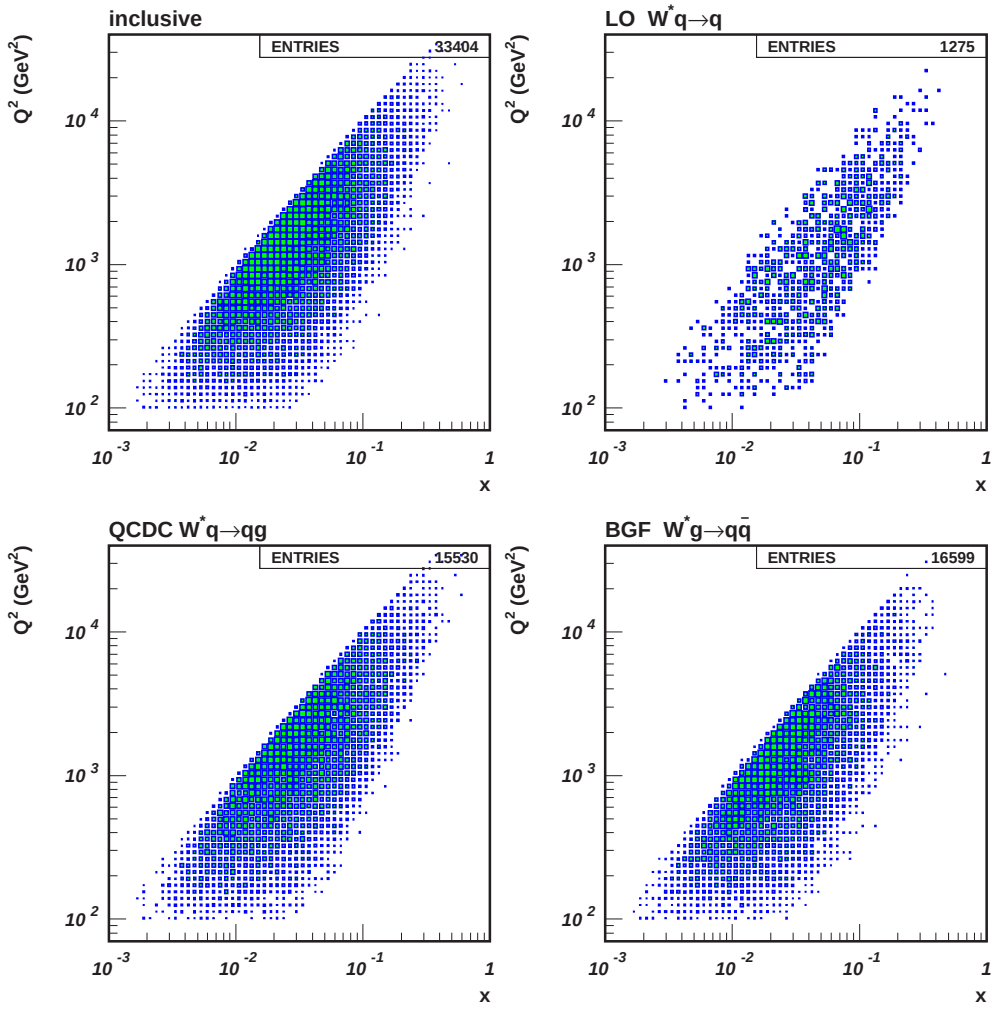


Figure A.2: *Simulation of events in the DIS plane for all and different QCD processes. D^* enriched CC MC sample simulated with ARIADNE is used for illustration. Details can be found in the text.*

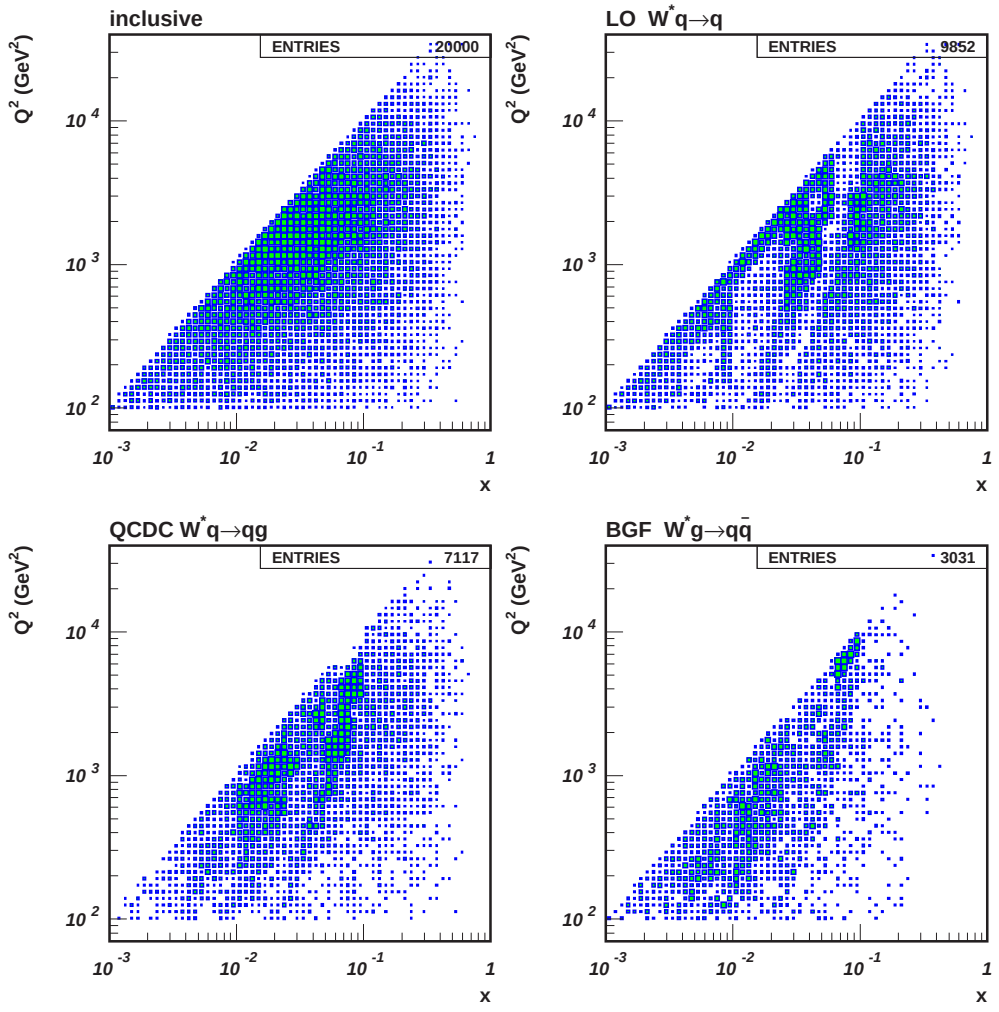


Figure A.3: *Problematic simulation of events in the DIS plane for different QCD processes. Inclusive CC MC sample simulated with MEPS and* `LST(19)=-10` *is used for illustration. Details can be found in the text.*

Appendix B

D^* candidates

There are in total three D^{*+} and four D^{*-} candidates in the charged current e^+p DIS in the HERA I data. They are shown with the ZEUS event visualization (ZeVis) [81]. Values of important global event variables as well as of the DIS kinematics shown in the plots are slightly different from those used in the analysis because only simple offline corrections have been implemented in the event display.

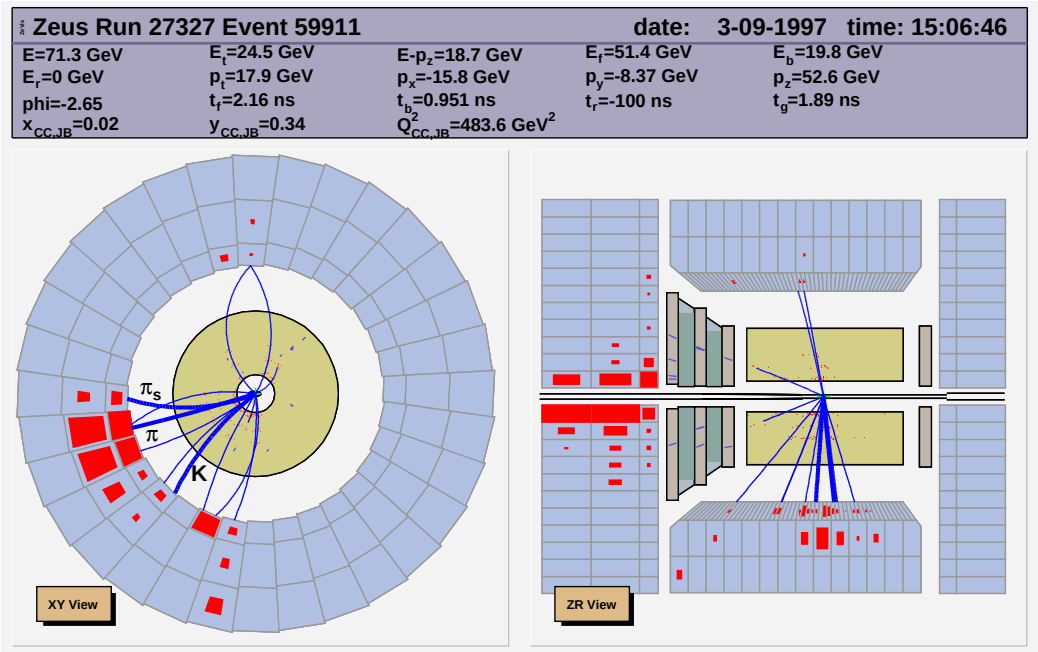


Figure B.1: *D^{*+} in CC DIS candidate: Run 27327 Event 59911. K^-, π^+ and π_s^+ are indicated by bold tracks.*

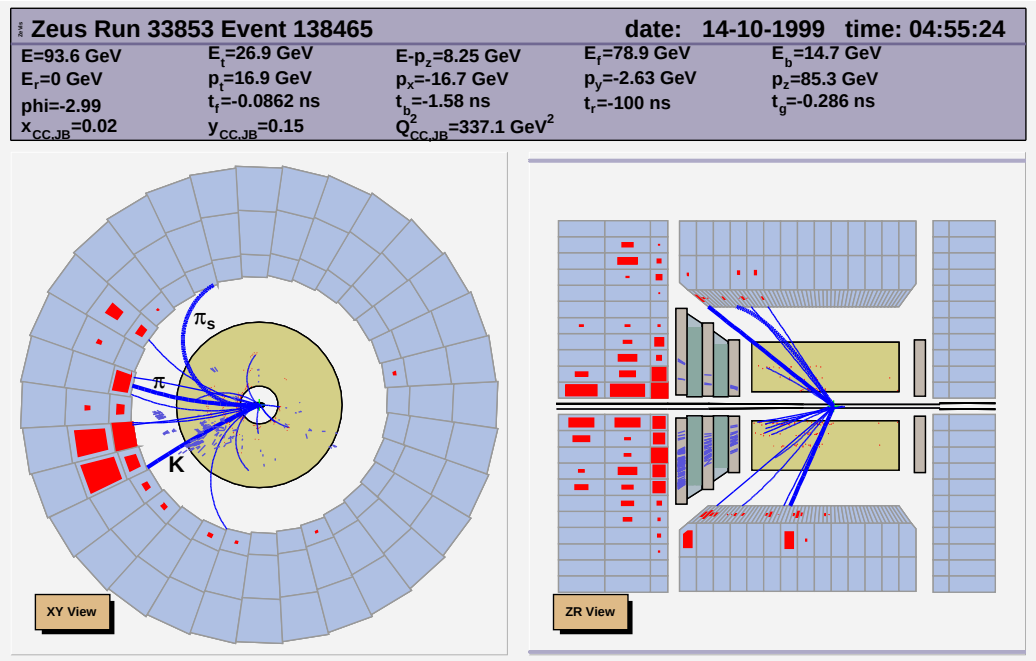


Figure B.2: *D^{*+} in CC DIS candidate: Run 33853 Event 138465. K^-, π^+ and π_s^+ are indicated by bold tracks.*

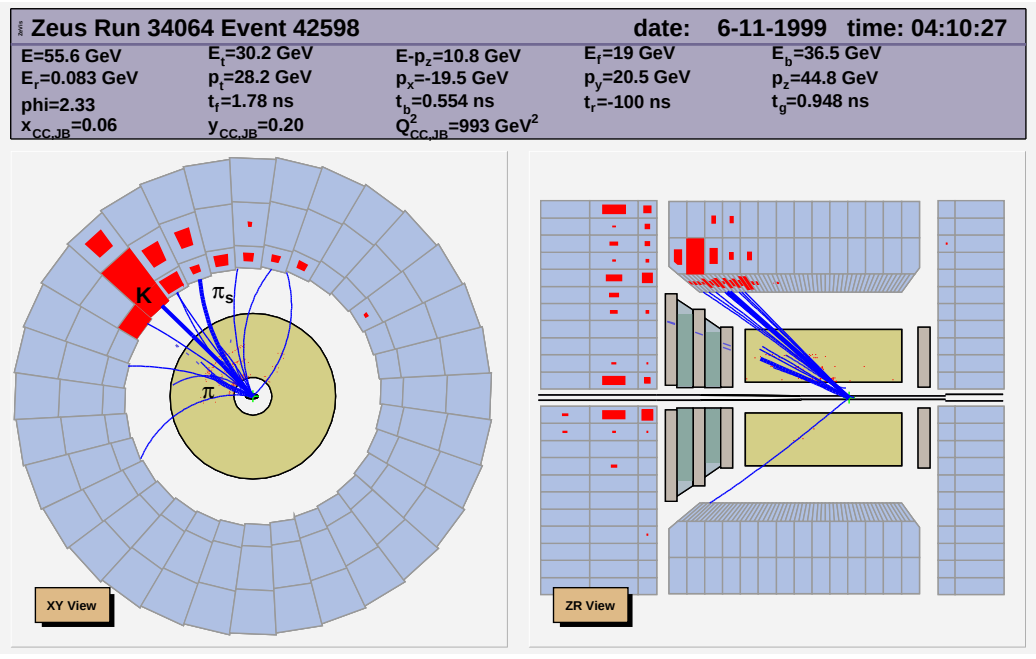


Figure B.3: *D^{*+} in CC DIS candidate: Run 34064 Event 42598. K^-, π^+ and π_s^+ are indicated by bold tracks.*

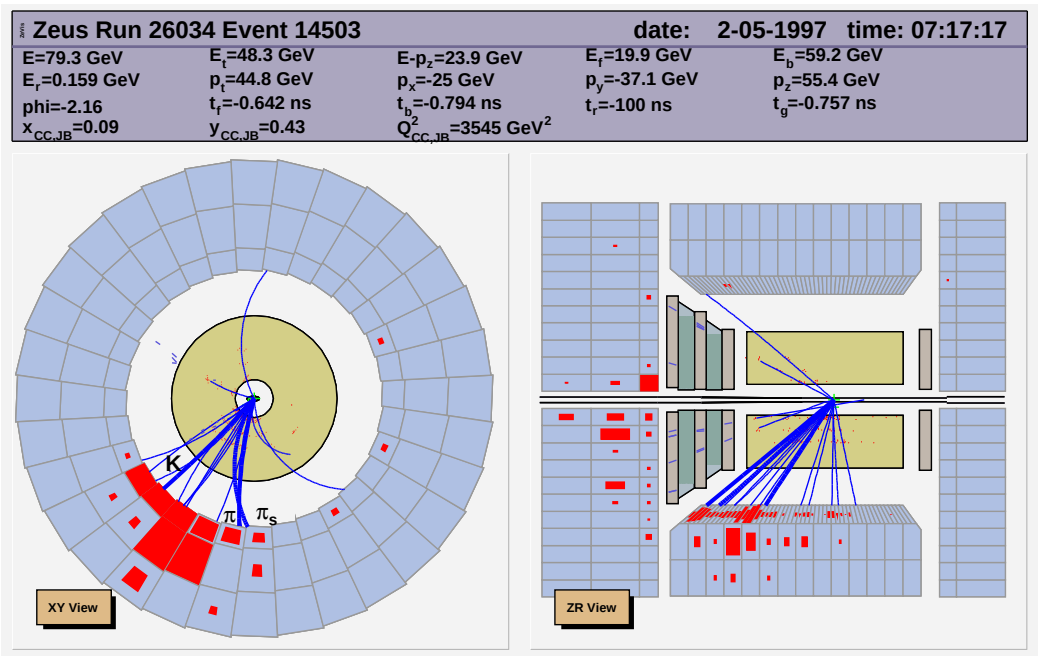


Figure B.4: *D^{*-} in CC DIS candidate: Run 26034 Event 14503. K^+, π^- and π_s^- are indicated by bold tracks.*

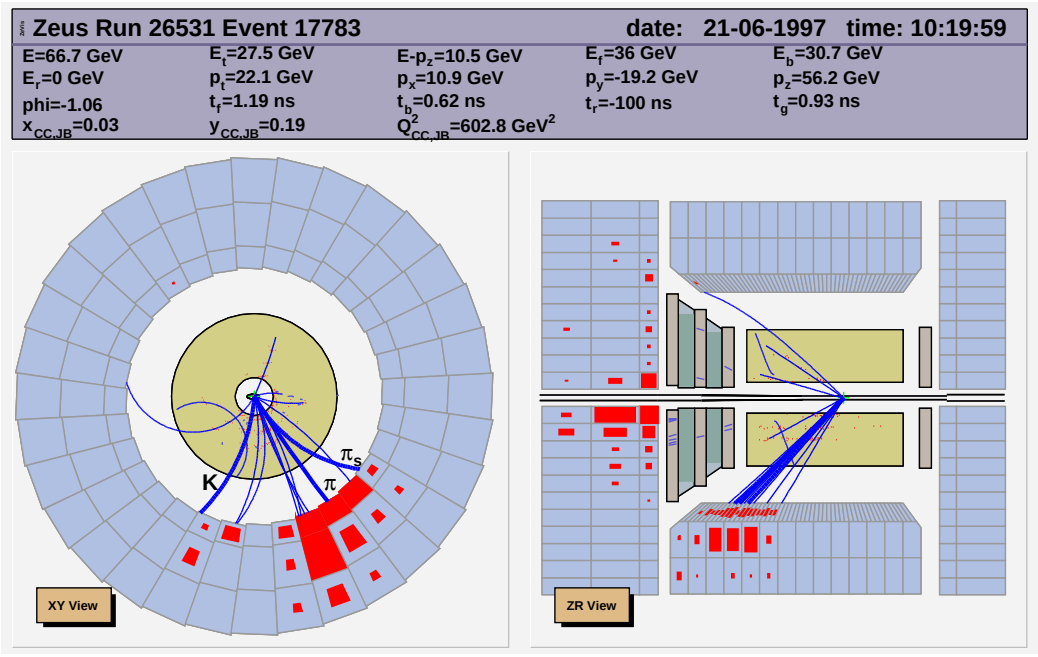


Figure B.5: *D^{*-} in CC DIS candidate: Run 26531 Event 17783. K^+, π^- and π_s^- are indicated by bold tracks.*

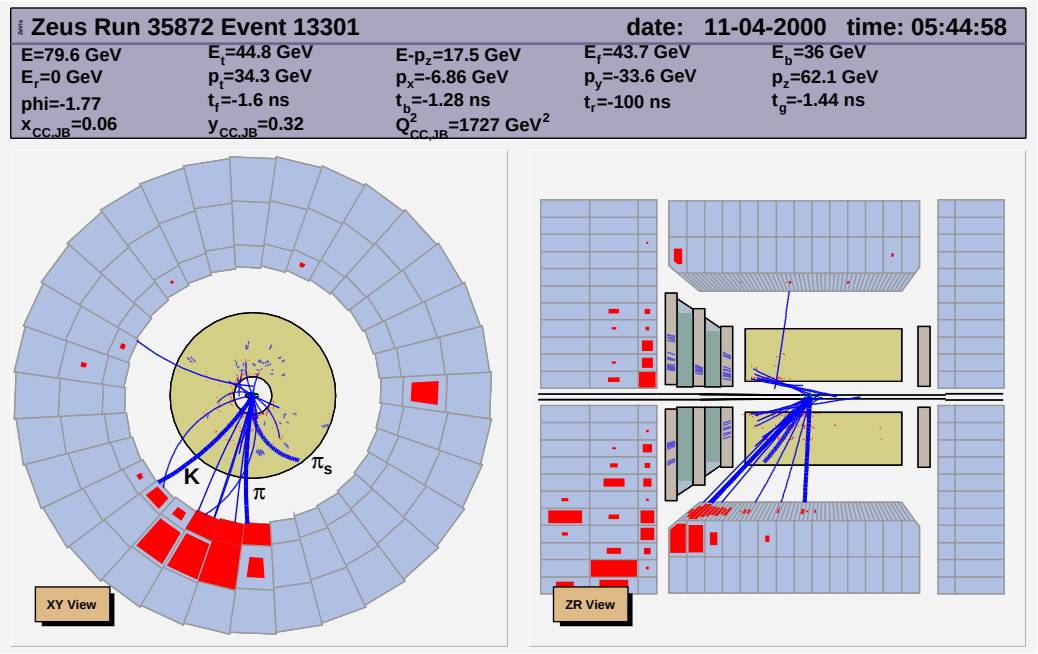


Figure B.6: *D^{*-} in CC DIS candidate: Run 35872 Event 13301. K^+, π^- and π_s^- are indicated by bold tracks.*

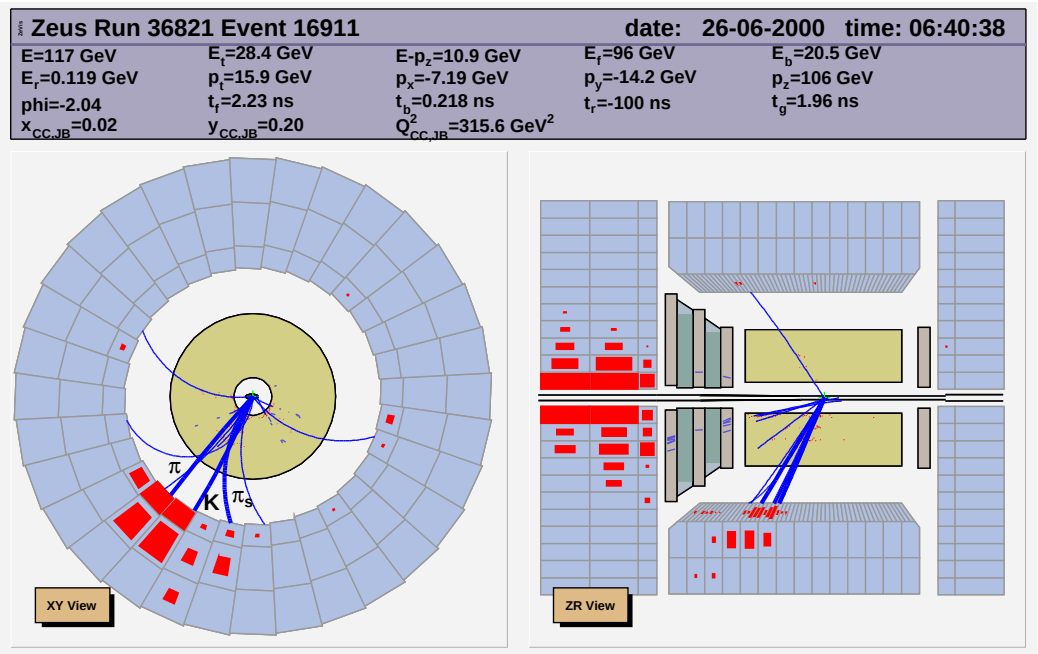


Figure B.7: *D^{*-} in CC DIS candidate: Run 36821 Event 16911. K^+, π^- and π_s^- are indicated by bold tracks.*

Appendix C

Bayesian approach for determining confidence intervals

The main reason to choose the Bayesian approach [82, 1] is its probabilistic meaning in the result, while another reason is that the inclusion of uncertainties is relatively easy in this approach.

Given the number of candidates (n) that contains the number of background (b) with some uncertainty (σ_b), the general form of the Bayesian approach to determine the probability density function (p.d.f.) of the number of signal (s) is

$$\begin{aligned} p(s|n,b,\sigma_b) &\propto L'(s;n,b,\sigma_b)\pi(s) \\ &= \left(\int L(s;n,b+\nu)\pi(\nu|\sigma_b)d\nu\right)\pi(s), \end{aligned} \tag{C.1}$$

with ν an *nuisance parameter*. Using the likelihood function for Poisson distributed n,

$$L(s;n,b+\nu) = \frac{(s+b+\nu)^n}{n!}e^{-(s+b+\nu)}, \tag{C.2}$$

the prior

$$\pi(s) = \begin{cases} 0 & s<0 \\ 1 & s\geq 0 \end{cases}, \tag{C.3}$$

and a Gaussian p.d.f. for the nuissance parameter

$$\pi(\nu|\sigma_b) = \frac{1}{\sqrt{2\pi}\sigma_b}e^{\frac{\nu^2}{2\sigma_b^2}}, \tag{C.4}$$

the upper-limit ($s_{\rm ul}$) at the confidence level (CL) of $1-\alpha$ can be obtained by integrating the posterior

$$1-\alpha = \int_{-\infty}^{s_{\rm ul}} p(s|n,b,\sigma_b)ds. \tag{C.5}$$

The result for determining the upper limit on the number of hard-scatter D^{*+}s is shown in Figure C.1. The p.d.f. for s with $\sigma_b = 0$ is also shown for comparison.

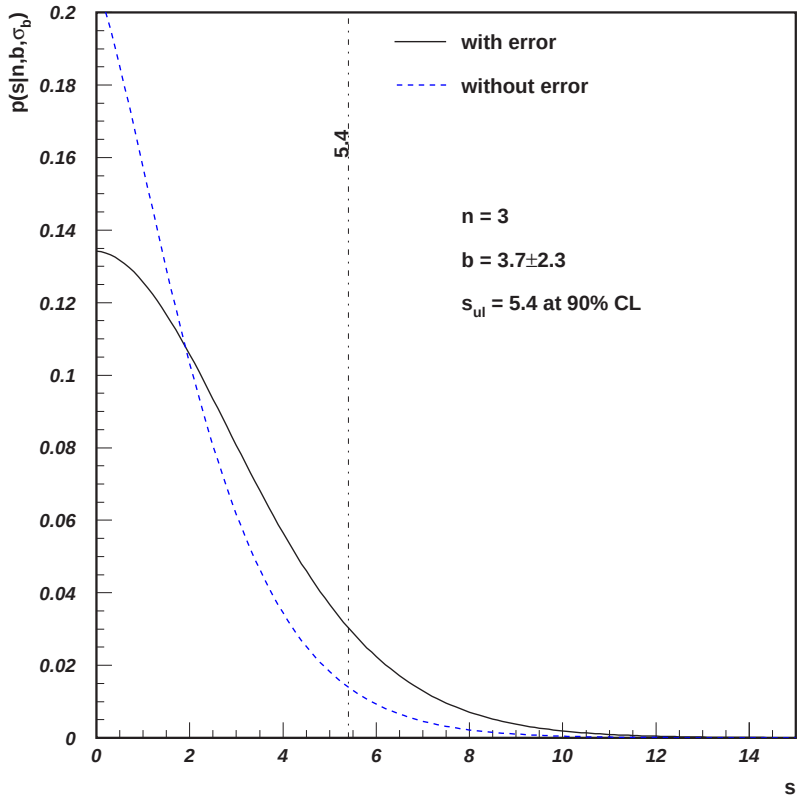


Figure C.1: *Bayesian approach for determing the upper limit on the number of hard-scatter D^{*+}s. The solid curve is the p.d.f with the error on b and the dashed curve without the error. The upper limit is indicated by the vertical line.*

Bibliography

[1] **Particle Data Group** Collaboration, K. Hagiwara *et al.*, "Review of particle physics", *Phys. Rev.* **D66** (2002) 010001;
Particle Data Group Collaboration, S. Eidelman *et al.*, "Review of particle physics", *Phys. Lett.* **B592** (2004) 1.

[2] R. W. Mcallister and R. Hofstadter, "Elastic scattering of 188-MeV electrons from the proton and the alpha particle", *Phys. Rev.* **102** (1956) 851–856.

[3] **SLAC-MIT** Collaboration, E. D. Bloom *et al.*, "High-energy inelastic *ep* scattering at 6° and 10°", *Phys. Rev. Lett.* **23** (1969) 930–934.

[4] **SLAC-MIT** Collaboration, M. Breidenbach *et al.*, "Observed behavior of highly inelastic electron-proton scattering", *Phys. Rev. Lett.* **23** (1969) 935–939.

[5] **ZEUS** Collaboration, S. Chekanov *et al.*, "A ZEUS next-to-leading-order QCD analysis of data on deep inelastic scattering", *Phys. Rev.* **D67** (2003) 012007, hep-ex/0208023.

[6] **ZEUS** Collaboration, S. Chekanov *et al.*, "An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment", *Eur. Phys. J.* **C42** (2005) 1–16, hep-ph/0503274.

[7] **ZEUS** Collaboration, S. Chekanov *et al.*, "Dijet photoproduction at HERA and the structure of the photon", *Eur. Phys. J.* **C23** (2002) 615–631, hep-ex/0112029;
ZEUS Collaboration, S. Chekanov *et al.*, "Inclusive jet cross sections in the Breit frame in neutral current deep inelastic scattering at HERA and determination of α_s", *Phys. Lett.* **B547** (2002) 164–180, hep-ex/0208037.

[8] **E598** Collaboration, J. J. Aubert *et al.*, "Experimental observation of a heavy particle J", *Phys. Rev. Lett.* **33** (1974) 1404–1406.

[9] **SLAC-SP-017** Collaboration, J. E. Augustin *et al.*, "Discovery of a narrow resonance in e^+e^- annihilation", *Phys. Rev. Lett.* **33** (1974) 1406–1408.

[10] **H1** Collaboration, C. Adloff *et al.*, "Measurement of $D^{*\pm}$ meson production and F_2^c in deep inelastic scattering at HERA", *Phys. Lett.* **B528** (2002) 199–214, hep-ex/0108039.

[11] **H1** Collaboration, A. Aktas *et al.*, "Measurement of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at high Q^2 using the H1 vertex detector at HERA", *Eur. Phys. J.* **C40** (2005) 349–359, hep-ex/0411046;
H1 Collaboration, A. Aktas *et al.*, "Measurement of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at low Q^2 and x using the H1 vertex detector at HERA", *Eur. Phys. J.* **C45** (2006) 23–33, hep-ex/0507081.

[12] **H1** Collaboration, A. Aktas *et al.*, "Measurement of beauty production at HERA using events with muons and jets", *Eur. Phys. J.* **C41** (2005) 453–467, hep-ex/0502010.

[13] **ZEUS** Collaboration, J. Breitweg *et al.*, "Measurement of $D^{*\pm}$ production and the charm contribution to F_2 in deep inelastic scattering at HERA", *Eur. Phys. J.* **C12** (2000) 35–52, hep-ex/9908012;
ZEUS Collaboration, S. Chekanov *et al.*, "Measurement of $D^{*\pm}$ production in deep inelastic $e^{\pm}p$ scattering at HERA", *Phys. Rev.* **D69** (2004) 012004, hep-ex/0308068.

[14] **ZEUS** Collaboration, S. Chekanov *et al.*, "Measurement of beauty production in deep inelastic scattering at HERA", *Phys. Lett.* **B599** (2004) 173–189, hep-ex/0405069.

[15] S. Mangano, *Charm production in charged current interactions and charm and beauty production in neutral current interactions at high Q^2*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2005.

[16] **CTEQ** Collaboration, R. Brock *et al.*, "Handbook of perturbative QCD", *Rev. Mod. Phys.* **67** (1995) 157–248.

[17] H. Abramowicz and A. Caldwell, "HERA collider physics", *Rev. Mod. Phys.* **71** (1999) 1275–1410, hep-ex/9903037.

[18] A. M. Cooper-Sarkar, R. C. E. Devenish, and A. De Roeck, "Structure functions of the nucleon and their interpretation", *Int. J. Mod. Phys.* **A13** (1998) 3385–3586, hep-ph/9712301;
R. Devenish and A. Cooper-Sarkar, *Deep Inelastic Scattering*. Oxford University Press, UK, 2004.

[19] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438–450;
L. N. Lipatov, "The parton model and perturbation theory", *Sov. J. Nucl. Phys.* **20** (1975) 94–102;
G. Altarelli and G. Parisi, "Asymptotic freedom in parton language", *Nucl. Phys.* **B126** (1977) 298;
Y. L. Dokshitzer, "Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics.", *Sov. Phys. JETP* **46** (1977) 641–653.

[20] V. S. Fadin, E. A. Kuraev, and L. N. Lipatov, "On the pomeranchuk singularity in asymptotically free theories", *Phys. Lett.* **B60** (1975) 50–52;
E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, "Multi-reggeon processes in the Yang-Mills theory", *Sov. Phys. JETP* **44** (1976) 443–450;
E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, "The pomeranchuk singularity in nonabelian gauge theories", *Sov. Phys. JETP* **45** (1977) 199–204;
I. I. Balitsky and L. N. Lipatov, "The pomeranchuk singularity in quantum chromodynamics", *Sov. J. Nucl. Phys.* **28** (1978) 822–829.

[21] M. Ciafaloni, "Coherence effects in initial jets at small Q^2/s", *Nucl. Phys.* **B296** (1988) 49;
S. Catani, F. Fiorani, and G. Marchesini, "QCD coherence in initial state radiation", *Phys. Lett.* **B234** (1990) 339;
S. Catani, F. Fiorani, and G. Marchesini, "Small-x behavior of initial state radiation in perturbative QCD", *Nucl. Phys.* **B336** (1990) 18.

[22] **CCFR** Collaboration, A. O. Bazarko *et al.*, "Determination of the strange quark content of the nucleon from a next-to-leading order QCD analysis of neutrino charm production", *Z. Phys.* **C65** (1995) 189–198, hep-ex/9406007.

[23] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Parton distributions: A new global analysis", *Eur. Phys. J.* **C4**

(1998) 463–496, hep-ph/9803445;
A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "MRST2001: Partons and alpha(s) from precise deep inelastic scattering and Tevatron jet data", *Eur. Phys. J.* **C23** (2002) 73–87, hep-ph/0110215;
A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Uncertainties of predictions from parton distributions. I: Experimental errors. ((T))", *Eur. Phys. J.* **C28** (2003) 455–473, hep-ph/0211080.

[24] **CTEQ** Collaboration, J. Botts *et al.*, "CTEQ parton distributions and flavor dependence of sea quarks", *Phys. Lett.* **B304** (1993) 159–166, hep-ph/9303255;
H. L. Lai *et al.*, "Global QCD analysis and the CTEQ parton distributions", *Phys. Rev.* **D51** (1995) 4763–4782, hep-ph/9410404;
H. L. Lai *et al.*, "Improved parton distributions from global analysis of recent deep inelastic scattering and inclusive jet data", *Phys. Rev.* **D55** (1997) 1280–1296, hep-ph/9606399;
CTEQ Collaboration, H. L. Lai *et al.*, "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions", *Eur. Phys. J.* **C12** (2000) 375–392, hep-ph/9903282;
CTEQ Collaboration, J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **07** (2002) 012, hep-ph/0201195.

[25] **H1** Collaboration, C. Adloff *et al.*, "Deep-inelastic inclusive ep scattering at low x and a determination of α_s", *Eur. Phys. J.* **C21** (2001) 33–61, hep-ex/0012053.

[26] W. G. Seligman *et al.*, "Improved determination of α_s from neutrino nucleon scattering", *Phys. Rev. Lett.* **79** (1997) 1213–1216.

[27] **New Muon** Collaboration, M. Arneodo *et al.*, "Accurate measurement of F_2^d/F_2^p and $R^d - R^p$", *Nucl. Phys.* **B487** (1997) 3–26, hep-ex/9611022.

[28] **BCDMS** Collaboration, A. C. Benvenuti *et al.*, "A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2", *Phys. Lett.* **B223** (1989) 485;
BCDMS Collaboration, A. C. Benvenuti *et al.*, "A high statistics measurement of the deuteron structure functions $F_2(x, Q^2)$ and R

from deep inelastic muon scattering at high Q^2", *Phys. Lett.* **B237** (1990) 592.

[29] **E665** Collaboration, M. R. Adams *et al.*, "Proton and deuteron structure functions in muon scattering at 470-GeV", *Phys. Rev.* **D54** (1996) 3006–3056.

[30] **New Muon** Collaboration, M. Arneodo *et al.*, "Measurement of the proton and deuteron structure functions, F_2^p and F_2^d, and of the ratio σ_L/σ_T", *Nucl. Phys.* **B483** (1997) 3–43, hep-ph/9610231.

[31] L. W. Whitlow, E. M. Riordan, S. Dasu, S. Rock, and A. Bodek, "Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections", *Phys. Lett.* **B282** (1992) 475–482.

[32] **CCFR/NuTeV** Collaboration, B. T. Fleming *et al.*, "A first measurement of low x low Q^2 structure functions in neutrino scattering", *Phys. Rev. Lett.* **86** (2001) 5430–5433, hep-ex/0011094.

[33] **DØ** Collaboration, B. Abbott *et al.*, "The inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8\,\text{TeV}$", *Phys. Rev. Lett.* **82** (1999) 2451–2456, hep-ex/9807018;
CDF Collaboration, T. Affolder *et al.*, "Measurement of the inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8\,\text{TeV}$", *Phys. Rev.* **D64** (2001) 032001, hep-ph/0102074.

[34] **H1** Collaboration, C. Adloff *et al.*, "Measurement and QCD analysis of neutral and charged current cross sections at HERA", *Eur. Phys. J.* **C30** (2003) 1–32, hep-ex/0304003.

[35] **ZEUS** Collaboration, J. Breitweg *et al.*, "Measurement of high-Q^2 neutral-current e^+p deep inelastic scattering cross-sections at HERA", *Eur. Phys. J.* **C11** (1999) 427–445, hep-ex/9905032;
ZEUS Collaboration, S. Chekanov *et al.*, "Measurement of the neutral current cross section and F_2 structure function for deep inelastic e^+p scattering at HERA", *Eur. Phys. J.* **C21** (2001) 443–471, hep-ex/0105090;
ZEUS Collaboration, S. Chekanov *et al.*, "Measurement of high-Q^2 e^-p neutral current cross sections at HERA and the extraction of xF_3", *Eur. Phys. J.* **C28** (2003) 175, hep-ex/0208040;
ZEUS Collaboration, S. Chekanov *et al.*, "High-Q^2 neutral current

cross sections in e^+p deep inelastic scattering at $\sqrt{s} = 318\,\text{GeV}$", *Phys. Rev.* **D70** (2004) 052001, hep-ex/0401003.

[36] **ZEUS** Collaboration, J. Breitweg *et al.*, "Measurement of high-Q^2 charged-current e^+p deep inelastic scattering cross sections at HERA", *Eur. Phys. J.* **C12** (2000) 411–428, hep-ex/9907010;
ZEUS Collaboration, S. Chekanov *et al.*, "Measurement of high-Q^2 charged current cross sections in e^-p deep inelastic scattering at HERA", *Phys. Lett.* **B539** (2002) 197–217, hep-ex/0205091;
ZEUS Collaboration, S. Chekanov *et al.*, "Measurement of high-Q^2 charged current cross sections in e^+p deep inelastic scattering at HERA", *Eur. Phys. J.* **C32** (2003) 1–16, hep-ex/0307043.

[37] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "NNLO global parton analysis", *Phys. Lett.* **B531** (2002) 216–224, hep-ph/0201127;
A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Uncertainties of predictions from parton distributions. II: Theoretical errors", *Eur. Phys. J.* **C35** (2004) 325–348, hep-ph/0308087.

[38] Webpage, "HEPDATA—The Durham HEP Databases". http://durpdg.dur.ac.uk/HEPDATA/.

[39] N. Cabibbo, "Unitary symmetry and leptonic decays", *Phys. Rev. Lett.* **10** (1963) 531–532;
M. Kobayashi and T. Maskawa, "CP violation in the renormalizable theory of weak interaction", *Prog. Theor. Phys.* **49** (1973) 652–657.

[40] S. Frixione, P. Nason, and G. Ridolfi, "Differential distributions for heavy flavor production at HERA", *Nucl. Phys.* **B454** (1995) 3–24, hep-ph/9506226;
B. W. Harris and J. Smith, "Charm quark and $D^{*\pm}$ cross sections in deeply inelastic scattering at DESY HERA", *Phys. Rev.* **D57** (1998) 2806–2812, hep-ph/9706334.

[41] M. Cacciari and M. Greco, "D^* production from e^+e^- to ep collisions in NLO QCD", *Phys. Rev.* **D55** (1997) 7134–7143, hep-ph/9702389;
G. Heinrich and B. A. Kniehl, "Next-to-leading-order predictions for $D^{*\pm}$ plus jet photoproduction at DESY HERA", *Phys. Rev.* **D70** (2004) 094035, hep-ph/0409303.

[42] F. I. Olness and W.-K. Tung, "When is a heavy quark not a parton? Charged higgs production and heavy quark mass effects in the QCD

based parton model", *Nucl. Phys.* **B308** (1988) 813;
M. A. G. Aivazis, F. I. Olness, and W.-K. Tung, "QCD formulation of charm production in deep inelastic scattering and the sea quark-gluon dichotomy", *Phys. Rev. Lett.* **65** (1990) 2339–2342;
M. A. G. Aivazis, F. I. Olness, and W.-K. Tung, "Leptoproduction of heavy quarks. I. General formalism and kinematics of charged current and neutral current production processes", *Phys. Rev.* **D50** (1994) 3085–3101, hep-ph/9312318;
M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, "Leptoproduction of heavy quarks. II. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies", *Phys. Rev.* **D50** (1994) 3102–3118, hep-ph/9312319.

[43] R. S. Thorne and R. G. Roberts, "An ordered analysis of heavy flavour production in deep inelastic scattering", *Phys. Rev.* **D57** (1998) 6871–6898, hep-ph/9709442;
R. S. Thorne and R. G. Roberts, "A practical procedure for evolving heavy flavour structure functions", *Phys. Lett.* **B421** (1998) 303–311, hep-ph/9711223;
R. S. Thorne and R. G. Roberts, "A variable number flavour scheme for charged current heavy flavour structure functions", *Eur. Phys. J.* **C19** (2001) 339–349, hep-ph/0010344.

[44] G. Ingelman and G. A. Schuler, "Characteristics of heavy flavor production in *ep* collisions", *Z. Phys.* **C40** (1988) 299.

[45] V. Barone, U. D'Alesio, and M. Genovese, "The charm-strange contribution to charged-current DIS structure functions", hep-ph/9610211.

[46] S. Kretzer and M. Stratmann, "Next-to-leading order QCD corrections to charged current charm production and the unpolarized and polarized strange sea at HERA", *Eur. Phys. J.* **C10** (1999) 107–119, hep-ph/9902426.

[47] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, "Scaling violations in inclusive e^+e^- annihilation spectra", *Phys. Rev.* **D27** (1983) 105.

[48] **CCFR** Collaboration, S. A. Rabinowitz *et al.*, "Measurement of the strange sea distribution using neutrino charm production", *Phys. Rev. Lett.* **70** (1993) 134–137.

[49] **OPAL** Collaboration, K. Ackerstaff *et al.*, "Measurement of $f(c \to D^{*+}X)$, $f(b \to D^{*+}X)$ and $\gamma_{c\bar{c}}/\gamma_{\rm had}$ using $D^{*\pm}$ mesons", *Eur. Phys. J.* **C1** (1998) 439–459, hep-ex/9708021.

[50] E. Hilger, "ZEUS coordinate system". ZEUS-86-17 (unpublished), 1986.

[51] U. Holm *et al.*, "The ZEUS Detector". Status Report, 1993. Available from http://www-zeus.desy.de/bluebook/bluebook.html.

[52] K. Long, D. H. Saxon, and B. T. Payne, "ZEUS CTD parameters issue 5—revised". ZEUS-89-23 (unpublished), 1989.

[53] R. Hall-Wilton, N. McCubbin, P. Nylander, M. Sutton, and M. Wing, "The CTD tracking resolution". ZEUS-99-024 (unpublished), 1999.

[54] D. Bartsch, "Corrections for dE/dx measurements with the ZEUS CTD". ZEUS note in preparation, 2006;
O. M. Kind, *Open Beauty Production at HERA*. PhD thesis, Universität Bonn, 2006. In preparation;
S. Marawske, "Improvement of dE/dx measurements in the central tracking detector of ZEUS at HERA", Master's thesis, Universität Bonn, 2005.

[55] H. Bethe and W. Heitler, "On the Stopping of fast particles and on the creation of positive electrons", *Proc. Roy. Soc. Lond.* **A146** (1934) 83–112.

[56] **ZEUS Luminosity Group** Collaboration, J. Andruszkow *et al.*, "Luminosity measurement in the ZEUS experiment", *Acta Phys. Polon.* **B32** (2001) 2025–2058;
K. Olkiewicz and A. Eskreys, "Off-line luminosity calculation in the ZEUS experiment in 1997, 1998 and 1999". ZEUS-99-044 (unpublished), 1999.

[57] F. Chlebana, "Description of the ZEUS global second level trigger in 1994". ZEUS-94-102 (unpublished). Available from http://www-zeus.desy.de/components/gslt/old/GSLT/gslt_1994/gslt_doc/gslt_doc.html, 1994.

[58] Webpage, "ZEUS physics reconstruction program packages". http://www-zeus.desy.de/~ifhuta/ZEUS_ONLY/zephyr/.

[59] N. Tuning, "ZUFOs: Hadronic final state reconstruction with calorimeter, tracking and backsplash correction". ZEUS-01-021 (unpublished), 2001.

[60] J. Grosse-Knetter, "Corrections for the hadronic final state". ZEUS-98-031 (unpublished). Available from http://www-zeus.desy.de/~grosse/hqex/newcor/, 1998.

[61] G. F. Hartner, Y. Iga, J. B. Lane, N. A. McCubbin, and M. Wing, "VCTRAK (3.07/04): Offline output information". ZEUS-97-064 (unpublished), 1997;
G. F. Hartner, "VCTRAK briefing: Program and math". ZEUS-98-058 (unpublished), 1998.

[62] P. Billoir and S. Qian, "Fast vertex fitting with a local parametrization of tracks", *Nucl. Instrum. Meth.* **A311** (1992) 139–150.

[63] F. Jacquet and A. Blondel, "Report from the study group on detectors for charged current events", in *Study of an ep facility for Europe*, p. 391. Hamburg, Germany, 1979.

[64] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, "GEANT3", tech. rep., CERN, 1987.

[65] Webpage, "The ZEUS Monte Carlo Chain". http://www-zeus.desy.de/components/funnel/MONTE_CARLO/MONTE_CARLO.html.

[66] H. Spiesberger, *HERACLES and DJANGOH: Event generation for ep interactions at HERA including radiative processes*, 1998.

[67] A. Kwiatkowski, H. Spiesberger, and H. J. Mohring, "HERACLES: An event generator for *ep* interactions at HERA energies including radiative processes: version 1.0", *Comp. Phys. Commun.* **69** (1992) 155–172;
H. Spiesberger, "An event generator for *ep* interactions at HERA including radiative processes (version 4.6)", 1996.

[68] G. Ingelman, A. Edin, and J. Rathsman, "LEPTO 6.5 — A Monte Carlo generator for deep inelastic lepton-nucleon scattering", *Comput. Phys. Commun.* **101** (1997) 108–134, hep-ph/9605286. Available from http://www3.tsl.uu.se/thep/lepto/.

[69] T. Sjostrand, "The Lund Monte Carlo for jet fragmentation and e^+e^- physics: JETSET version 6.2", *Comput. Phys. Commun.* **39** (1986) 347–407;
T. Sjostrand and M. Bengtsson, "The Lund Monte Carlo for jet fragmentation and e^+e^- physics: JETSET version 6.3: an update", *Comput. Phys. Commun.* **43** (1987) 367;
T. Sjostrand, "High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4", *Comput. Phys. Commun.* **82** (1994) 74–90;
T. Sjostrand, "PYTHIA 5.7 and JETSET 7.4: physics and manual", hep-ph/9508391.

[70] L. Lonnblad, "ARIADNE version 4: A program for simulation of QCD cascades implementing the color dipole model", *Comput. Phys. Commun.* **71** (1992) 15–31. Available from http://www.thep.lu.se/~leif/ariadne/.

[71] H. Plothow-Besch, *PDFLIB—Proton, pion and photon parton density functions, parton density functions of the nucleus, and α_s calculations, Version 8.04*. CERN, 2000. Available from http://wwwasdoc.web.cern.ch/.

[72] G. Marchesini *et al.*, "HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. version 5.1", *Comput. Phys. Commun.* **67** (1992) 465–508.

[73] M. Gluck, E. Reya, and A. Vogt, "Photonic parton distributions", *Phys. Rev.* **D46** (1992) 1973–1979.

[74] A. D. Tapper, *Measurement of charged current deep inelastic scattering cross sections using the ZEUS detector at HERA*. PhD thesis, University of London, 2001.

[75] Webpage, "Exotic DST Bits". http://www-zeus.desy.de/physics/exo/ZEUS_ONLY/trigger/.

[76] K. Nagano, *Measurement of charged-current e^+p deep inelastic scattering cross sections at $\sqrt{s} = 300$ GeV*. PhD thesis, University of Tokyo, 1998.

[77] M. Moritz, *Measurement of the high Q^2 neutral current DIS cross section at HERA*. PhD thesis, Universität Hamburg, 2002.

[78] J. Rautenberg, *Measurement of high-Q^2 charged current cross sections in e^+p deep inelastic scattering at HERA*. PhD thesis, Universität Bonn, 2004.

[79] O. Deppe, *Measurement of $D^{*\pm}$ electroproduction at HERA*. PhD thesis, Universität Hamburg, 1999.

[80] **E687** Collaboration, P. L. Frabetti *et al.*, "Charm-anticharm asymmetries in high energy photoproduction", *Phys. Lett.* **B370** (1996) 222–232.

[81] Webpage, "ZeVis—ZEUS Event Visualisation". http://www-zeus.desy.de/~zevis/.

[82] G. D'Agostini, "Probability and measurement uncertainty in physics—a Bayesian primer", hep-ph/9512295.

Acknowledgments（致谢）

The success of this doctoral work is due to many people. I would like to thank them for various reasons from giving physics discussion and guidance to making life in Bonn and Hamburg enjoyable.

My thanks first go to Prof. Erwin Hilger for providing me an opportunity to study in Germany, giving me invaluable guidance in research and paying fatherly attention to my life. What I learned with him is not only knowledge, in physics and in all respects of life, but also a spirit of research. I am also pleased to thank Prof. Ian Brock and Prof. Ewald Paul for all their disinterested help of kindness.

I really rejoice to be a member of the ZEUS-Bonn group. I am grateful to all dear fellows for helping me to make life as a foreigner much easier. I would like to thank Julian Rautenberg and Detlef Bartsch for fruitful discussions, keeping our local computer cluster always state-of-the-art and, in particular, criticizing my thesis carefully. I am grateful to Alexander Kappes and Ainas Weber for helping me to get familiar with ZEUS physics analyses fast. I want to thank young colleagues, Stefan Marawske, Stefan Hoppe, Felix Fehr and Michał Własenko, for bringing so much fun into "dull" life. I really appreciate all the help from the FDET folks, especially, from Oliver Kind, Roger Renner and Ursula Samson.

I am grateful to Massimo Corradi, Matthew Wing, Richard Hall-Wilton and Achim Geiser for their excellent coordination of the heavy flavor group and valuable advice to my research. Specifically, I would like to thank Achim Geiser for plenty of discussions and strong support. I am also grateful to Leonid Gladilin for his maintaining the common-interest analysis software.

Any high energy research is impossible without collaborative effort of a huge number of people. I really appreciate all contributions made by DESY people to the HERA collider, the ZEUS detector, the computing environment, and so on.

Last but not least, I sincerely thank my wife, Qi Jingtao, my parents, Wang Yifu and Chen Longzhu, as well as younger sister, Wang Lei, for giving me so much love and support. In addition, special thanks are reserved for my little daughter, Shuying, for bringing me so great joy and happiness.

最后，我由衷地感谢我的妻子亓靖涛，我的父母王益赋和陈龙珠，以及妹妹王蕾，为了他们给予我那么多的爱和支持。另外，一份特别的感谢留给我的小女儿舒颖，为了她带给我如此巨大的愉快和幸福。