Diss. ETH No. 15969
ETHZ-IPP Internal Report 2005-02

ETH Institute for Particle Physics

Charm Production in Charged Current Interactions and Charm & Beauty Production in Neutral Current Interactions at High Q^2

A dissertation submitted to the
Swiss Federal Institute of Technology Zurich
for the degree of
Doctor of Natural Sciences

presented by
Salvatore Mangano
Dipl. Phys. ETH
born on May 28, 1974
in Zurich / Switzerland

accepted on the recommendation of
Prof. Dr. Ralph Eichler, examiner and
Prof. Dr. Günther Dissertori, co-examiner

February 2005

Abstract

The charm and beauty production cross sections in e^+p neutral current processes and the charm production cross section in e^+p charged current processes at high Q^2 are measured at HERA. Their determination relies on an inclusive method, where secondary decay vertices are explicitly reconstructed using the H1 vertex detector.

The charm and beauty cross sections in neutral current interactions are measured in the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ and the results are

$$\sigma(e^+p \rightarrow e'^+ c\bar{c}X) = 476 \pm 60 \pm 76\ pb$$

$$\sigma(e^+p \rightarrow e'^+ b\bar{b}X) = 38 \pm 6 \pm 8\ pb.$$

The results for the neutral current process are found to be compatible with the prediction of perturbative quantum chromodynamics.

In the charged current data an excess over the Standard Model prediction is seen and has been verified by different methods. The excess of events with secondary decay vertices is found in events with transverse momentum of the hadronic system above 40 GeV. At lower values the measurement is in agreement with predictions.

If the excess is interpreted as charm contribution, then the measurement of the charm cross section in charged current interactions yields in the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$

$$\sigma(e^+p \rightarrow \bar{\nu}_e\ c\ X) = 17.7 \pm 5.5 \pm 2.3\ pb.$$

Within the large statistical and systematic errors the measured results are compatible with the predictions.

Zusammenfassung

Der e^+p-Wirkungsquerschnitt für Charm und Beauty Produktion in neutralen Strom Prozessen und der e^+p-Wirkungsquerschnitt für Charm Produktion in geladenen Strom Prozessen bei hohen Q^2 wurde bei HERA gemessen. Deren Bestimmung beruht auf einer inklusiven Methode, wobei sekundäre Zerfallsvertizes unter Benutzung des H1 Vertexdetektors explizit rekonstruiert werden.

Die gemessenen Charm- und Beauty-Wirkungsquerschnitte im kinematischen Bereich $Q^2 > 150\ GeV^2$ und $0.1 < y < 0.7$ betragen im neutralen Strom

$$\sigma(e^+p \to e'^+ c\bar{c}X) = 476 \pm 60 \pm 76\ pb$$

$$\sigma(e^+p \to e'^+ b\bar{b}X) = 38 \pm 6 \pm 8\ pb.$$

Die Resultate für den neutralen Strom sind kompatibel mit den Vorhersagen der perturbativen QCD.

Ein Überschuss über der Vorhersage des Standard Modells wurde für die Daten des geladenen Stromes gesehen und mit verschiedenen Methoden bestätigt. Der Überschuss von Ereignissen mit sekundären Zerfallsvertizes wurde für Ereignisse mit Transversalimpuls des hadronischen Systems über 40 GeV gefunden. Für tiefere Werte stimmen Messung und Vorhersage überein.

Falls der Überschuss als Charm Beitrag interpretiert wird, ergibt sich für den Charm-Wirkungsquerschnitt im kinematischen Bereich $Q^2 > 150\ GeV^2$ und $0.1 < y < 0.7$ folgendes Resultat

$$\sigma(e^+p \to \bar{\nu}_e\ c\ X) = 17.7 \pm 5.5 \pm 2.3\ pb.$$

Innerhalb der grossen statistischen und systematischen Fehler stimmen gemessenes Resultat und Vorhersage überein.

Contents

Abstract **3**

Zusammenfassung **4**

Introduction **9**

1 The H1 Detector at HERA **11**
1.1 HERA 12
1.2 The H1 Detector 12
1.2.1 The Central Silicon Tracker CST 12
1.2.2 The Central Jet Chamber CJC 14
1.2.3 The Liquid Argon Calorimeter LAr 14
1.3 Track Reconstruction Issues for the z-Coordinate 14
1.4 The Trigger System 15
1.5 Monte Carlo Technique and Detector Simulation 16

2 Theoretical Framework **17**
2.1 Kinematics of Deep Inelastic Scattering 17
2.2 NC and CC Cross Sections 18
2.3 Heavy Quarks Production in NC Processes 19
2.4 Charm Cross Section in CC Processes 21
2.5 Coarse Estimation of the Charm Cross Section in CC Processes 23

3 The Inclusive Analysis Method **25**
3.1 Tagging Signature and Strategy of the Analysis 25
3.2 Measurement of the Cross Section 26
3.3 The Structure of the Final State: The Jet 27
3.4 The Observables 27
3.5 Different Inclusive Tagging Methods 29
3.5.1 Fixed Vertex Multiplicity (FVM) Method 32
3.5.2 Variable Vertex Multiplicity (VVM) Method 33
3.5.3 Pro and Con of the FVM and VVM Methods 35
3.6 Fit Procedure 35

4 The Neutral and Charged Current Selection **37**
4.1 Analyzed Data Set, Trigger and Integrated Luminosity 37
4.2 Event Simulation 37
4.3 Reconstruction of Event Kinematics 39
4.4 Control of Energy Calibration 40
4.4.1 Comparison of Generated and Reconstructed Positron Variables . . . 40
4.4.2 Electron Energy Calibration 41
4.4.3 Hadron versus Electron Energy Calibration 42
4.5 Neutral Current Measurement 43
4.6 Charged Current Measurement 45

5 Track, Primary Vertex and Jet Reconstruction **49**
5.1 Track Reconstruction 49
5.1.1 CST-improved Tracks 50
5.1.2 Strangeness Rejection 51
5.2 Primary Vertex 52
5.3 Jet Reconstruction 54

6 Charm and Beauty Production in Neutral Current Processes **57**
6.1 Signal Extraction with the FVM Method 57
6.1.1 Cut Based Heavy Quarks Cross Section with the FVM Method 59
6.1.2 Charm and Beauty Cross Section with FVM 2- and 3-tracks Methods 60
6.1.3 Charm and Beauty Cross Sections with a Fit Procedure in the FVM Method 62
6.1.4 Problems with the FVM Method 67
6.2 Signal Extraction with the VVM Method 68
6.2.1 Charm and Beauty Cross Sections with a Fit Procedure in the VVM Method 69
6.2.2 Experimental Systematic Errors 73
6.2.3 Theoretical Errors for the NLO Calculation 75
6.3 Results: Charm and Beauty Production in Neutral Current Processes 76

7 Charm Production in Charged Current Processes **79**
7.1 Tests of the Extraction Method in NC 79
7.1.1 NC-metamorphosed Sample 80
7.1.2 Results of the NC-metamorphosed Sample 80
7.1.3 Systematic Errors in the NC-metamorphosed Sample 83
7.1.4 Test of the Statistical Problem in the NC-metamorphosed Sample . . 84
7.2 Charm Extraction with the VVM Method in CC Processes 85
7.2.1 Lower Charm Efficiency in CC Processes than in NC Processes 88
7.2.2 Control of the 2 Tracks Decay Vertex Events 90
7.2.3 Pictures of the 2 Tracks Decay Vertex Events 92
7.2.4 Excess of Events with Decay Vertices at $P_{T,h} > 40\ GeV$ 94
7.2.5 Charm Extraction with the FVM Method in CC Processes 96

7.3 Results: Charm Production in Charged Current Processes 98

8 Conclusions 101

Appendix 103

A Systematic Errors in NC Processes 103

B Track, Primary Vertex and Jet Reconstruction in CC Processes 107

C Systematic Errors in CC Processes 114

D Pictures of CC Processes 116

List of Figures 130

List of Tables 132

Bibliography 134

Danksagung 139

Curriculum Vitae 140

Introduction

In the Standard Model of particle physics, matter is described by point-like matter particles (quarks and leptons), which interact via forces (the strong and electroweak interactions). The main objective of the HERA *ep* accelerator is the test of these interactions and the study of the proton structure. Whereas the strong interaction is tested in practically every measurement at HERA, the electroweak interactions are studied in the neutral current (NC) and charged current (CC) processes at high Q^2.

Until last year the heavy quark production in the high Q^2 region was not explored. No results above $Q^2 > 150\ GeV^2$ for beauty production in NC interactions existed at HERA and no results from the H1 collaboration for the charm production in NC interactions. These results have now been published [1] with a similar analysis method. The production of charmed quarks in CC events has not been measured at HERA until now. By studying the charm in CC processes the strangeness component of the proton sea could be constrained.

Indeed on one side at high Q^2 a significant fraction of the events are events containing heavy quarks, but on the other side the total cross section is small. Therefore the purpose of this work is to find, implement and test a new inclusive heavy quark tagging method. This opens the possibilities to compare the data with the theory in a kinematic region where until now the exclusive method is not applicable because of low statistics.

For the heavy quark tagging method the different lifetime signature of charm and beauty flavored hadrons compared to the light hadrons is used. The lifetime of the charm and beauty flavored hadrons is reflected in displaced tracks from the *ep*-collision point. Therefore, the tagging technique relies on the spatial separation of the production and decay vertex, exploiting the high-resolution capabilities of the H1 vertex detector.

This thesis is organized as follows. After a short overview of HERA and the H1-Detector in the first chapter, the heavy quark production in NC and CC will be illustrated in the second chapter.

In chapter three the algorithm of the two new inclusive tagging methods and the analysis strategy are presented.

The following two chapters justify the correctness of the measurement by the comparisons of data and simulation for the important distributions of the inclusive topological vertexing technique: In chapter four the NC and CC selection are described and in chapter five the important selection criteria of the lifetime tag are discussed.

In chapter six two inclusive approaches are applied to extract the charm and beauty cross sections in NC. The starting point is the measurement of the heavy quark (this means charm and beauty together) cross sections in NC with a cut based approach. The cuts in lifetime related variables provide a sufficient rejection power to reduce both the light quark production events as well as accidental combinatorial background sources. Afterwards the challenging part of the analysis is the distinction of light, charm and beauty cross sections done by statistical extraction of lifetime related distributions. Both approaches are tested, but one method proves to have smaller systematic errors and it is also taken for the final results. The measured results have been compared to theoretical models.

In chapter seven the charm production in CC is studied. To justify that this method is applicable to the CC sample, cross checks are done in a kinematically restricted NC sample which behaves similarly as the CC sample. The advantage of this restricted NC sample is that the distribution of variables can be compared between simulation and data without the problem of statistical fluctuations. This chapter concludes with the measurement of the charm in CC cross section and some control studies of candidate events for charm production in CC processes.

Chapter 1

The H1 Detector at HERA

The measurement discussed is based on positron-proton (*ep*) collision data taken by the H1 experiment. H1 is one of four detectors at HERA[1], which is the main accelerator at DESY[2] in Hamburg.

The following section gives a short overview of HERA. Afterwards the components of the H1 experiment are discussed in brief. The chapter closes with a short description of the detector simulation.

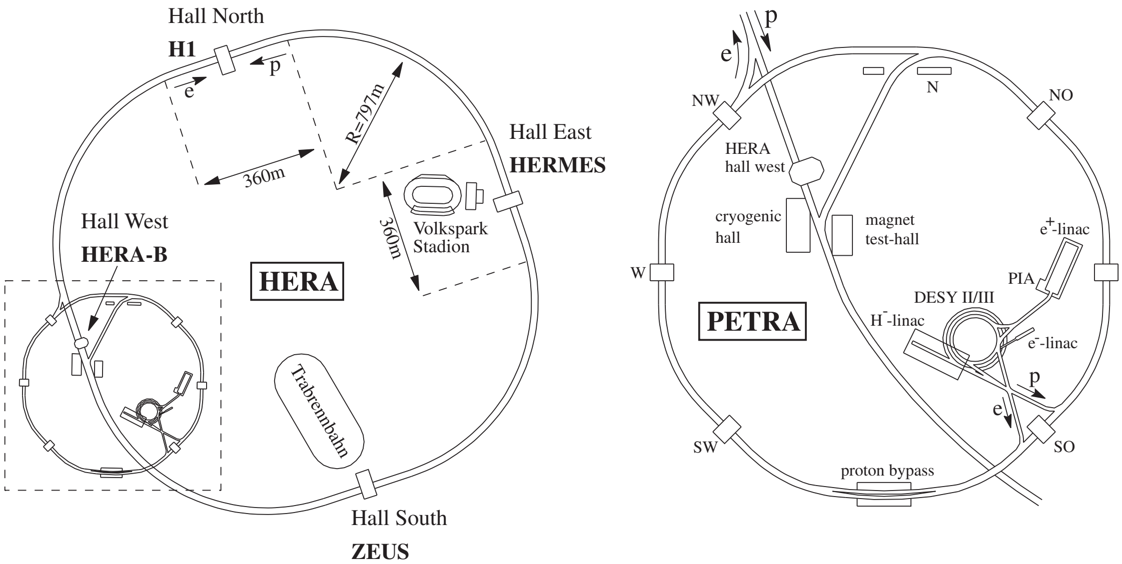


Figure 1.1: The HERA storage ring (left) and the pre-accelerators (right) at DESY.

[1]Hadron Electron Ring Accelerator
[2]Deutsches Elektronen Synchrotron

1.1 HERA

HERA is a so far unique storage ring in the study of ep collisions designed for colliding a 920 GeV proton beam with a 27.6 GeV positron beam at a center of mass energy of $\sqrt{s} = 319\ GeV$. The positron beam energy is limited by high energy losses due to synchrotron radiation and the proton beam energy is limited by the magnetic field. The main purpose of HERA are probing the structure of the proton and testing the standard model of particle physics. Figure 1.1 shows the layout of the accelerator. After passing a chain of pre-accelerators, positrons and protons are injected in opposite directions into two separate rings of 6.3 km circumference, where they are accelerated to their nominal energies. The particle beams are stored for several hours which is referred to as a luminosity fill. The particle beams are not continuous, but consist of about 200 bunches each. At two experiments, H1 and ZEUS, the bunches are brought to collision every 96 ns corresponding to a rate of 10.4 MHz. The spatial distribution of protons in a bunch is Gaussian with a width of $\sigma_x \approx 150\ \mu m$, $\sigma_y \approx 50\ \mu m$ transversal to the beam direction and of $\sigma_z \approx 11\ cm$ in beam direction.

1.2 The H1 Detector

The H1 detector displayed in figure 1.2 is built asymmetrically due to the different energies of the colliding particles. The nominal interaction point is marked in the figure with a small cross. The positive z-axis is defined by the direction of the proton beam. The x-axis points to the center of the ring and therefore the y-axis points upwards. Polar angles θ are measured with respect to the positive z-axis while azimuthal angles extend clockwise from the positive x-axis as shown in the top right corner of figure 1.2. The z projection is labeled the $r\phi$-transverse plane and the x projection the rz-plane.

The protons enter the detector in the beam pipe from the right, the positrons from the left. The interaction region is surrounded by a carbon fiber beam pipe with a radius of 45 mm. The beam pipe is surrounded by the central tracking system $\boxed{2}$ (see 1.2.1 and 1.2.2) consisting of the Central Silicon Tracker (CST), the Central Jet Chamber (CJC), the Central z-chambers and the Central Proportional Chamber. The electromagnetic $\boxed{4}$ and the hadronic $\boxed{5}$ section (see 1.2.3) of the Liquid Argon Calorimeter (LAr) encloses the tracking detectors. A superconducting coil $\boxed{6}$ that generates a homogeneous magnetic field of 1.15 T surrounds the previous parts. A complete introduction to the H1 detector at HERA can be found in [2].

1.2.1 The Central Silicon Tracker CST

The Central Silicon Tracker allows the identification of heavy flavor particles with decay lengths of a few hundred micrometers due to its close position to the interaction point and the high resolution.

The CST consists of two concentric cylindrical layers of silicon sensors which are located at radii of 57 mm to 97 mm. The position resolution is 12 μm in $r\phi$ and 22 μm in z of this detector. The response of the sensitive CST component which determines the CST

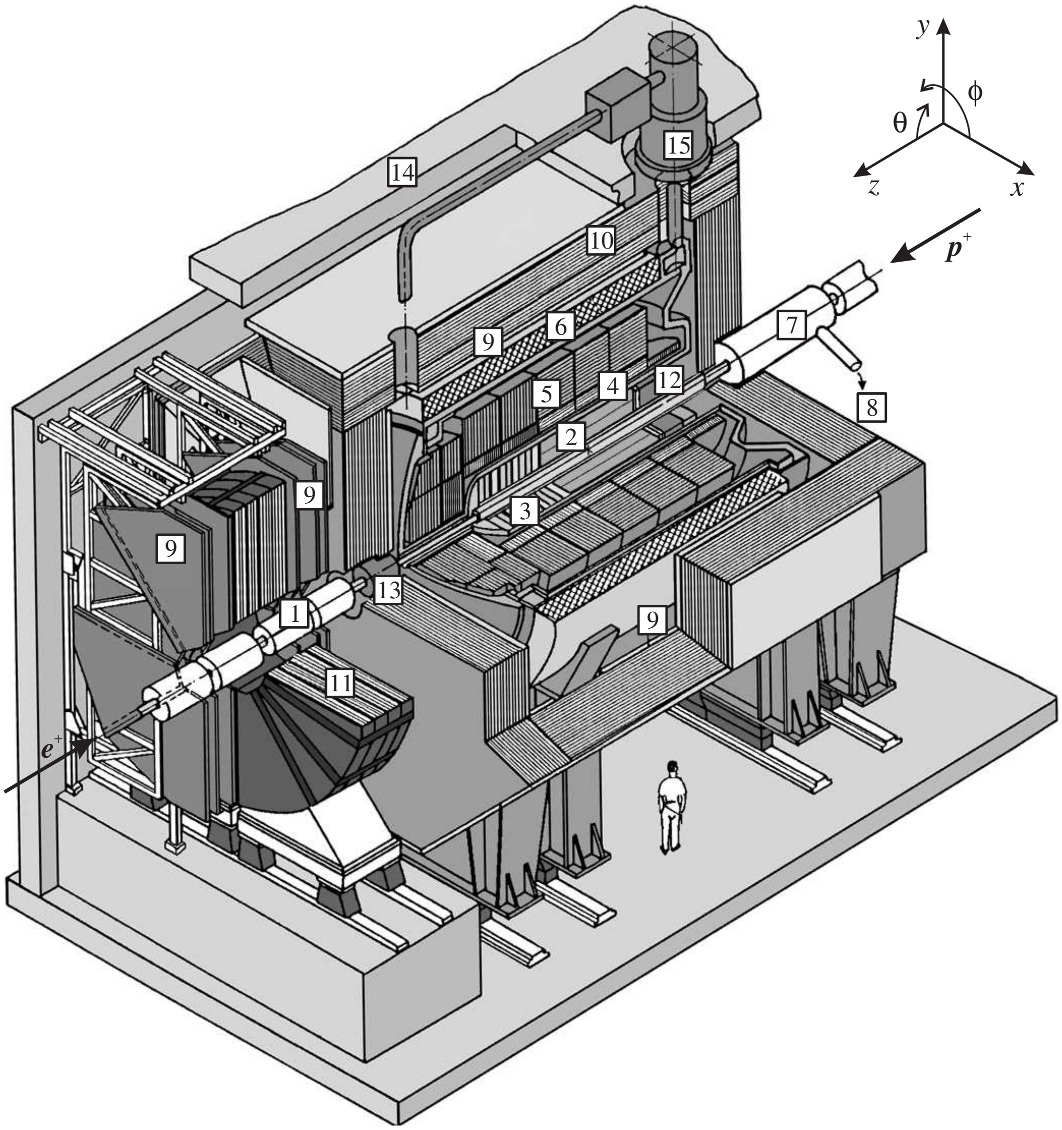


Figure 1.2: The overview of the H1 experiment shows the main detector components. The flight directions of the incoming proton and positron beams and the H1 coordinate system are indicated in the top right corner.

position point is always referred to as 'CST hit'. In the homogeneous magnetic field the charged particles are bent according to their transverse momentum and the recorded particle trajectory in the central tracking system is defined as a track. For tracks with CST hits in both layers the impact parameter in the $r\phi$-plane can be measured with a resolution of $33\ \mu m \oplus 90\ \mu m/p_T\ [GeV]$, where the first term represents the intrinsic resolution and the second term is the contribution from multiple scattering in the beam pipe; p_T is the transverse momentum of the track [1]. Detailed information concerning the CST may be found in [3].

1.2.2 The Central Jet Chamber CJC

The Central Jet Chamber is designed to measure the transverse momentum of charged particles with high precision.

The CJC consists of two cylindrical, coaxial volumes (CJC1 and CJC2) that are subdivided in 30, respectively 60 identical drift cells. Each drift cell contains 24 respectively 32 signal wires arranged parallel to the beam pipe. The intrinsic CJC resolution in $r\phi$ is 170 μm and 22 mm in z. The CJC covers the region of 25° to 155° in θ and is located at a radius of 150 mm to 850 mm.

1.2.3 The Liquid Argon Calorimeter LAr

The purpose of the Liquid Argon Calorimeter is to measure electromagnetic and hadronic energy deposits.

The LAr calorimeter covers the range in the polar angle of $4° < \theta < 154°$. It consists of an inner section for the measurement of electromagnetic showers and an outer section which is needed for the measurement of hadronic showers. The depth of the electromagnetic section is $\approx 20-30$ radiation lengths, the total depth of the calorimeter is $\approx 5-8$ interaction lengths. The calorimeter along the beam axis is segmented into 'wheels', with each wheel being further segmented into octants in ϕ. The regions between the wheels are problematic for the measurement due to energy losses. These gaps between the wheels are called 'z-cracks', between the octants 'ϕ-cracks'.

The sampling medium between the absorber plates is liquid argon. The absorber material is lead in the electromagnetic section and stainless steel in the hadronic section. Test beam measurements of LAr calorimeter modules revealed an energy resolution $\sigma_{em}(E)/E = 12\%/\sqrt{E\ [GeV]} \oplus 1\%$ for electrons and $\sigma_{had}(E)/E = 50\%/\sqrt{E\ [GeV]} \oplus 2\%$ for charged pions [2]. This allows for a very good measurement of the energy of the scattered positron as well as good resolution of the hadronic energy measurement.

1.3 Track Reconstruction Issues for the z-Coordinate

Because of the track reconstruction issues explained in the following this analysis is based on tracks with CST hits in both layers for the $r\phi$-coordinate only, ignoring the z-side by using a two-dimensional fitter.

The space points measured with the two CST layers are not sufficient for a stand alone track reconstruction. Therefore the CST information is added to CJC tracks with a combined CJC-CST track fit which determines CST-improved track parameters. The association of CST space points to vertex fitted CJC tracks is called linking and a good description is given in reference [4].

The linking of CST hits to CJC-tracks has its problems, especially for the z-coordinate. Compared to the $r\phi$-coordinate, the z-side has lower efficiency and a higher noise hit density while at the same time the precision of the tracks measured in the outer tracking is moderately low (cf. sec. 1.2.2). Depending on whether reliable Central z-chamber information is available, the uncertainty of the projection of a particle trajectory measured in the CJC onto the CST can be of the same order of magnitude as the average distance between random hits. A significant fraction of the CST-improved tracks will therefore have incorrectly assigned z-hits. This cannot be properly taken into account by the tracks covariance matrix and therefore vertex fitting with such tracks often leads to an unacceptable fit quality (χ^2), or decay length and invariant mass distributions with large tails.

One way to avoid this problem is to ignore the z-side altogether by using a two-dimensional fitter. Obviously these fitters have higher efficiency than the three dimensional fitter which needs the z-side information. Unfortunately, the in principle high precision of the CST z-measurement is lost and so is the additional constraint [5].

1.4 The Trigger System

To filter out the interesting physics events, the H1 trigger system discriminates promising ep interactions from background events. Since the positron-proton collision rate at the interaction point is approximately 10 MHz and the H1 recording bandwidth is limited to 10 Hz the rejection of uninteresting events must take place at the earliest possible stage. The trigger levels of interest are:

Level 1 (L1):

L1 combines the trigger information called *trigger elements* provided by the sub-detectors logically to *sub-triggers*. A single L1 positive sub-trigger activates the readout of an entire pipeline (1-2 ms). Frequently observed types of events or uninteresting monitor events are just stored at a fraction of there occurrence and weighted by this fraction. The L1 trigger level reduces the typical acquisition rate of 100 kHz by roughly a factor 100.

The current analysis uses a L1 sub-trigger, which triggers on the signature of high Q^2 events, i.e. highly energetic scattered positrons in the LAr calorimeter. The majority of the events in this analysis are triggered by the L1 ST67 sub-trigger. The decision of sub-trigger ST67 is based on the recognition of a local energy deposit of about 5 GeV or more in the LAr calorimeter and is combined with the timing information from the LAr calorimeter or the forward or the central proportional chambers. In addition the veto conditions are used to reject beam-gas and beam-wall background [6].

Level 2 (L2):

L2 identifies correlations among the sub-detectors which match classified event topologies. The input rate of about 1 kHz is reduced to approximately 50 Hz. The trigger level L3 was not implemented during the the data-taking period 1992-2000.

Level 4 (L4):

L4 verifies the lower level trigger decisions, performs a fast event reconstruction and determines a first set of detector calibration constants. The L4 filter farm is a software based trigger and has access to the complete event data. With the full detector information available, the selection algorithms examine each event according to the hard scale or final state finders. Events with a hard scale criterion (e.g. a track with a high transverse momentum or a high energy cluster in the calorimeter) are accepted. Depending on the reconstructed events signature, the events are classified from the final state finders into physical classes. Only if an event is recognized by at least one final state finder or one of the hard scale criteria is fulfilled, the event is accepted without a prescale factor. The L4 input rate of approximately 50 Hz is reduced to 10 Hz and stored on tape.

1.5 Monte Carlo Technique and Detector Simulation

The Monte Carlo (MC) simulation technique is indispensable for the extraction of physical quantities from the measurements. For the cross section measurement, MC programs can be used to determine corrections for acceptance, efficiencies, background contamination, and resolution effects of the detector system.

In the experiment the storage rings produce collisions. The interactions are recorded by the detectors, and the interesting ones are written to tape. Afterwards the events are reconstructed, i.e. the electronics signals (from central tracking system, calorimeters, etc.) are translated into a setup of charged tracks or neutral energy depositions. With this information the momenta and particle species may be determined and used for the physics analysis.

In the MC technique the event generator program generates the events. Every single generated long living particle propagates through active and inactive material of the detector and its interactions with the detector matter and the magnetic field is described by programs such as GEANT [7]. The energy deposited by the particles are tracked step-by-step and converted into detector signals (like in real data). The output of this simulation has exactly the same format as the real data recorded by the detector, and can be put through the same event reconstruction and physics analysis chain. The big advantage is that in the MC technique the reconstructed variables can be compared with the 'true' input (generated) variables.

Chapter 2

Theoretical Framework

2.1 Kinematics of Deep Inelastic Scattering

In a deep inelastic scattering (DIS) process the incoming positron[1] couples to an exchanged boson, which probes the structure of the proton.

In a process where the exchanged boson is a neutral particle, i.e. a photon or a neutral vector boson Z, the final state is formed by a scattered positron and a hadronic final state (X): $e^+p \rightarrow e^+X$. This process is referred to as neutral current (NC). In NC events no clear separation between the γ or Z is possible, because of the interference effects.

If the exchanged particle is a charged boson W^+, the outgoing lepton is an undetected electron anti-neutrino: $e^+p \rightarrow \bar{\nu}_e X$. This process is known as charged current (CC).

[1] or any lepton

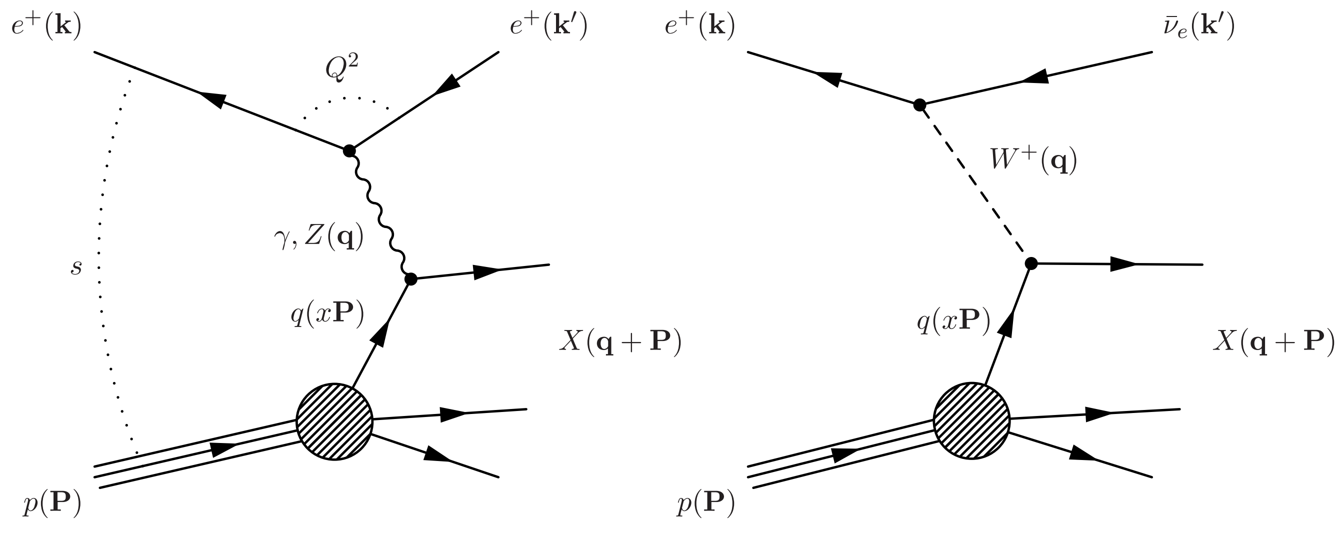


Figure 2.1: The dominant positron-proton scattering via photon and Z exchange for NC: $e^+p \rightarrow e^+X$ (left) and via W exchange for CC: $e^+p \rightarrow \bar{\nu}_e X$ (right). The four momentum vectors of the particles are given in the brackets.

Figure 2.1 shows the Feynman diagram of the dominant deep inelastic e^+p scattering process. The center-of-mass energy squared s of the reaction is given by the initial state

$$s = (\mathbf{k} + \mathbf{P})^2 \approx 4E_eE_p, \tag{2.1}$$

where $\mathbf{k} = (E_e, \vec{p}_e)$ and $\mathbf{P} = (E_p, \vec{p}_p)$ denote the four-momenta of the incoming positron and incoming proton. The second part of the equation is an approximation by neglecting the mass of the positron and proton. The variables E_e and E_p are the energies of the positron and proton in the HERA ring.

The transfer of the four-momentum squared from the positron to the proton is calculated via

$$Q^2 = -\mathbf{q}^2 = -(\mathbf{k} - \mathbf{k}')^2 \qquad Q^2 \in [0, s], \tag{2.2}$$

where $\mathbf{q}$ and $\mathbf{k'} = (E_{e'}, \vec{p}_{e'})$ denote the four-momenta of the exchanged gauge boson and the outgoing positron or the outgoing electron anti-neutrino respectively.

The Bjorken scaling variable x is given by

$$x = \frac{Q^2}{2\mathbf{P} \cdot \mathbf{q}} \qquad x \in [0, 1]. \tag{2.3}$$

The variable x can be interpreted in the Quark Parton Model as the fractional momentum of the proton carried by the struck parton if the proton and the parton masses are neglected in the infinite momentum frame.

The inelasticity y is

$$y = \frac{\mathbf{q} \cdot \mathbf{P}}{\mathbf{k} \cdot \mathbf{P}} \qquad y \in [0, 1], \tag{2.4}$$

and it can be interpreted as the relative energy transfer of a positron to the hadronic final state in the proton rest frame.

All these variables are related in the following way:

$$Q^2 = sxy. \tag{2.5}$$

In the case that the center-of-mass energy is known, two of the variables Q^2, x and y are needed to describe the kinematics of the DIS process.

2.2 NC and CC Cross Sections

If $Q^2 \gg 1\ GeV^2$ the quarks are asymptotically free and the lepton-proton DIS process is described by an incoherent superposition of elastic lepton-quark scattering processes.

The Born neutral current cross section for $e^+p \to e^+X$ scattering [8] is defined as

$$\frac{d^2\sigma_{NC}}{dxdQ^2} = \frac{2\pi\alpha^2}{x}\left(\frac{1}{Q^2}\right)^2 \phi_{NC}(x, Q^2), \tag{2.6}$$

where α denotes the electromagnetic coupling constant. In the leading order (LO) approximation of Quantum Chromodynamics (QCD), the structure function term is simply related to the sum of the quark densities, weighted with the squared quark charges, when neglecting Z exchange and beauty and top quark distributions:

$$(\phi_{NC})_{LO} = [1 + (1 - y)^2]x \left[\frac{4}{9}(u + c + \bar{u} + \bar{c}) + \frac{1}{9}(d + s + \bar{d} + \bar{s}) \right], \tag{2.7}$$

where u, d, s, c are quark density distributions and $\bar{u}, \bar{d}, \bar{s}, \bar{c}$ anti-quark density distributions. The above quark density distribution $f(x, Q^2)$ or $\bar{f}(x, Q^2)$ gives the probability to find a quark or anti-quark of a specific flavor, carrying a fraction x of the proton momentum when probing the proton with a momentum transfer Q^2.

The Born charged current cross section for $e^+p \to \bar{\nu}_e X$ is defined as

$$\frac{d^2\sigma_{CC}}{dxdQ^2} = \underbrace{\frac{G_F^2}{2\pi x}}_{coupling} \underbrace{\left[\frac{M_W^2}{Q^2 + M_W^2} \right]^2}_{propagator} \underbrace{\phi_{CC}(x, Q^2)}_{proton-structure-function}, \tag{2.8}$$

where G_F is the Fermi coupling constant and M_W the mass of the W boson. The structure function term ϕ_{CC} is given in leading order QCD by

$$(\phi_{CC})_{LO} = x([\bar{u} + \bar{c}] + (1 - y)^2[d + s]), \tag{2.9}$$

where again contributions of beauty and top quarks are neglected.

From the expressions of the NC (2.6) and CC (2.8) cross sections one sees firstly that they are proportional to a coupling (electromagnetic or Fermi respectively), a propagator (photon or W boson respectively) and a proton structure function. Secondly, one sees that the CC cross section is relatively suppressed by $\mathcal{O}(Q^4)$ at small Q^2 where the NC cross section is dominated by photon exchange, and thirdly at very high $Q^2 \gg \mathcal{O}(M_W^2)$, the charged and neutral cross sections are of the same order.

2.3 Heavy Quarks Production in NC Processes

A reasonable criterion for a quark to be called ‘heavy’ is $m_Q \gg \Lambda_{QCD} \sim 200\ MeV$, where Λ_{QCD} characterizes the energy scale above which the strong coupling constant becomes small, so that perturbative QCD (pQCD) is applicable at the scale of the heavy quark masses. The heavy flavors produced at HERA are charm and beauty, the top quark is too heavy to be created.

The cross sections of heavy quarks production can be predicted by pQCD calculations at next-to-leading order (NLO). The theoretical uncertainties come from the scheme for handling quark mass effects in QCD, the choice of the parton fit and the arbitrariness of the factorization scale [9]. Every finite order calculation depends on these scales [10, 11].

In performing calculations beyond leading order, various divergencies arise and the renormalization scale μ_r is introduced to regulate these divergencies. The factorization scale μ_f serves to define the separation of short-distance from long-distance scales. For all processes at short-distance pQCD is applicable and the long-distance processes are absorbed in the parton density functions.

In general the ep cross section for heavy quark (Q) production is the convolution of the parton density functions $f_{i/p}(x, \mu_f)$ depending on the factorization scale μ_f which describes the probability of finding partons i with a momentum fraction x, and the perturbatively calculable partonic cross section $\hat{\sigma}_{ei \to e'QX}$. This partonic cross section describes the scattering of the positron on a parton i inside the proton with

$$d\sigma(ep \to e'QX) = \sum_{partons_i} \int_0^1 f_{i/p}(x, \mu_f^2) \cdot d\hat{\sigma}_{ei \to e'QX}(x, m_Q^2, \mu_f^2) \cdot dx. \tag{2.10}$$

The partonic cross section can be calculated and depends on the masses, virtualities, coupling etc. of the particles involved. The parton densities cannot be calculated from first principles. The parton densities have to be given at some reference scale and then they can be computed for any value of the scale. The determination of the parton distributions are made with the evolution equations which fit $q, \bar{q}, g$ to data at a variety of Q^2. Two examples for the parton distributions fits from different groups are MRST [12] and CTEQ [13, 14]. The structure of an evolution equation is

$$\frac{df_j(x, \mu_f^2)}{d\ln \mu_f^2} = \frac{\alpha_s(\mu_r^2)}{2\pi} \sum_i \int_x^1 \frac{dx'}{x'} \cdot f_i(x', \mu_f^2) \cdot P_{i \to j,k}(z), \tag{2.11}$$

where f_i and f_j are the parton density functions and $P_{i \to j,k}(z)$ is the splitting function, which give the probability that the parton i is split into two partons j and k with momentum fraction $z = \frac{x}{x'}$. There exist different evolution equations evolving the density function in different variables and with different approximation [15], e.g. DGLAP [16], BFKL [17] and CCFM [18].

There are two different schemes to describe heavy flavor production, the massive and the massless scheme. In the massive scheme the heavy quark masses (charm and beauty) are taken fully into account in the calculations of the partonic cross sections, while the lighter quarks u, d and s are treated massless. In these calculations the heavy quarks contribution to the proton structure are neglected. At $Q^2 \gg m_c^2, m_b^2$ the production of heavy quarks is expected to be insensitive to the threshold effects and the heavy quarks may be treated as massless partons, like the other light quarks. This means the heavy quarks can be treated as active flavors in the proton and they are used as parton distributions in the proton [19, 20, 21, 22].

The quark-initiated NC heavy quark production is the flavor excitation ($\gamma/Z + Q \to Q$), which corresponds to the so called LO approximation. At low x the gluon density is rather large compared to the heavy quark density, so that the higher order boson gluon fusion (BGF) process ($\gamma/Z + g \to Q\bar{Q}$), which corresponds to a NLO calculation, can be numerically comparable or larger than the flavor excitation process [19]. According to [23], where the Q^2 range is slightly lower than the range defined for this analysis, the BGF process can be up to four times larger than the flavor excitation process. The two diagrams are shown in figure 2.2 and both have to be considered in order to obtain a correct prediction of observable quantities.

The main difference between the two production mechanisms is that for the BGF process two heavy hadrons are produced in the final state instead of only one as in the case of

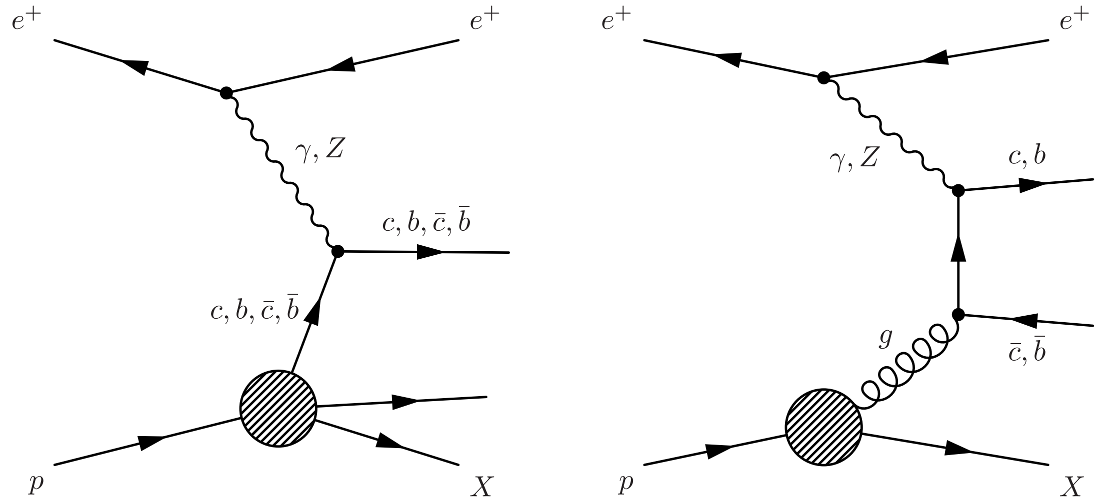


Figure 2.2: The quark-initiated or Born diagram for charm and beauty in NC (left). The gluon-initiated or BGF diagram for charm and beauty production (right). The crossed graph is not presented. The real gluon emission and virtual corrections are not shown but contribute to NLO.

the quark-initiated approach. These two production mechanisms have a different transverse momentum of the final state, which could be experimentally tested.

The produced partons then fragment into hadrons which may decay and are measured in the detector. The perturbative calculations of the hadronisation of the quarks are not converging due to confinement and thus phenomenological fragmentation models like Peterson [24] or Lund [25] are used.

A large fraction of the particles produced by the fragmentation are unstable and subsequently decay into observable stable or almost stable ones. The heavy hadrons decay via the weak interaction and this leads to finite mean lifetime expressed in units of length of $c\tau$ for charm hadrons of around 300 μm and for beauty hadrons of around 500 μm. Because of the larger beauty than charm quark mass, the average number of decay particles is larger for beauty than for charm hadrons. The charm decay charged multiplicity is 2.46 ± 0.09 [26] and the beauty decay charged multiplicity is 4.96 ± 0.06 [27]. All information about mass, branching ratio and decay properties is summarized [28].

2.4 Charm Cross Section in CC Processes

In the electroweak interaction the quark mass eigenstates are not the same as the weak eigenstates, and the matrix relating these bases is the Cabbibo-Kobayashi-Maskawa (CKM) matrix [28]. The present status of the experimental situation of the CKM matrix may be summarized as follows:

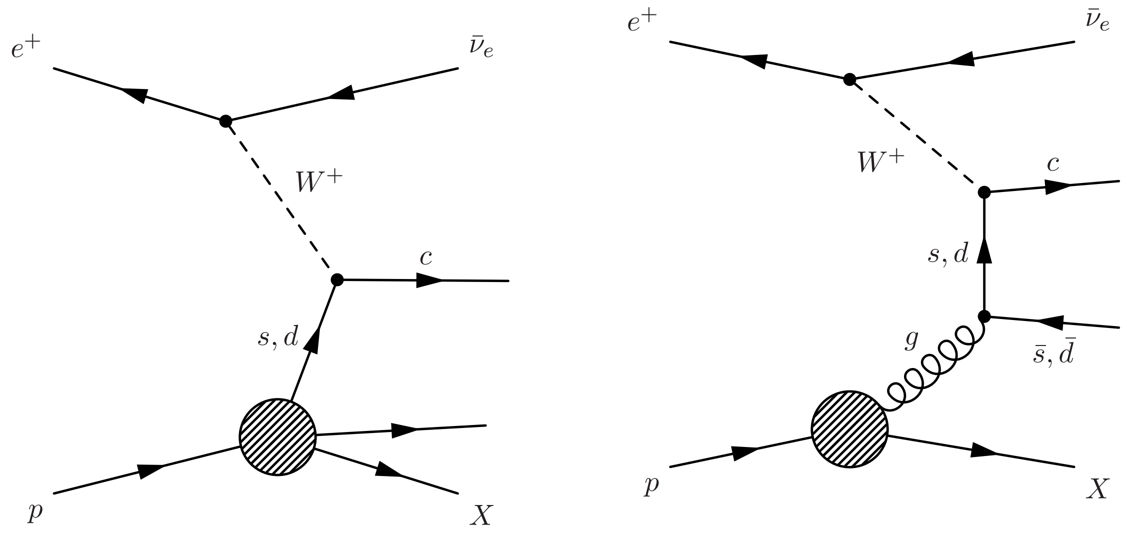


Figure 2.3: The quark-initiated or Born diagram for charm in CC (left). The Cabbibo-allowed (suppressed) contributions are obtained by s (d) quarks. The gluon-initiated or BGF diagram for charm in CC (right). The crossed diagram is not presented. The real gluon emission and virtual corrections are not shown but contribute to NLO.

$$V = \begin{pmatrix} V_{ud} \approx 0.97 & V_{us} \approx 0.22 & V_{ub} \approx 0 \\ V_{cd} \approx 0.22 & V_{cs} \approx 0.97 & V_{cb} \approx 0.04 \\ V_{td} \approx 0 & V_{ts} \approx 0.04 & V_{tb} \approx 1 \end{pmatrix}. \tag{2.12}$$

The diagonal elements V_{ud}, V_{cs}, V_{tb} are clearly dominant and are called Cabbibo-allowed transitions. The large value of V_{cs} simply reflects the experimental fact that charm particles preferentially decay into strange particles. The other elements are called Cabbibo-suppressed transitions, however the experimental observation that B mesons prefer to decay into charm particles implies $V_{cb} > V_{ub}$.

The charm cross section in charged current processes, $\sigma(e^+p \to \bar{\nu}_e cX)$, in the LO calculation obtained from equation (2.8), where the charm comes from the Cabbibo-allowed transition $s \to c$ or from the Cabbibo-suppressed transition $d \to c$, is given by

$$\frac{d^2\sigma(e^+p \to \bar{\nu}_e cX)}{dxdQ^2} = \frac{G_F^2}{2\pi} \left[\frac{M_W^2}{Q^2 + M_W^2} \right]^2 (1-y)^2 (d(x,Q^2)|V_{cd}|^2 + s(x,Q^2)|V_{cs}|^2). \tag{2.13}$$

The charm production in LO CC (see figure 2.3) is dominated by the Cabbibo-allowed $s \to c$ transition. The contribution of the Cabbibo-suppressed $d \to c$ transition is enhanced by the valence quark enhancement of d at large x. According to [29] the ratio of the Cabbibo-allowed over the Cabbibo-suppressed transition convoluted with the parton density function for the LO charm cross section is $\sigma(e^+p \to e^+s \to c\bar{\nu}_e X) : \sigma(e^+p \to e^+d \to c\bar{\nu}_e X) \approx 4 : 1$.

The beauty production is neglected. Because on the one hand $V_{ub} = 0.0037 \pm 0.0005$ [2] [28] and on the other hand the $\bar{c} \rightarrow \bar{b}$ transition is Cabbibo-suppressed and the $\bar{c}$ sea distribution is further suppressed with respect to other light quark distributions, due to the mass difference between charm and light quarks.

From equation (2.13) it is possible to extract $s(x, Q^2)$, which is not yet well known and was one motivation for this analysis. The charged conjugate process with an incident electron proceeds through an interaction with a $\bar{s}$ or $\bar{d}$ quark. Taking measurements with incident positron or electron give the opportunity to distinguish the $\bar{s}$ and s quark content in the proton.

At NLO other diagrams contribute to the charm production cross section: the real gluon emission, the virtual correction and the BGF processes. The most important contribution comes from the BGF term. The figure 2.3 shows the BGF charm production diagram for CC, where according to [29] the ratio of the allowed over the suppressed BGF charm cross section is $\sigma(e^+p \rightarrow \bar{\nu}_e c\bar{s}X) : \sigma(e^+p \rightarrow \bar{\nu}_e c\bar{d}X) \approx 20 : 1$. In the QCD framework, again the two diagrams in figure 2.3 have both to be considered in order to obtain a correct prediction of observable quantities.

Theoretical calculations [9] predict the total charm charged current cross section for the HERA kinematical range of the order of 4 *pb*. This lies in the experimentally accessible reach of H1.

Without any knowledge about QCD calculations, but only comparing the BGF diagrams from figure 2.2 and figure 2.3, it is clear that in the NC case the tagging of a charm event is more efficient than in CC, because in each event two charm quarks are produced in NC in contrast to only one charm quark in CC. Furthermore, in the case of positron scattering in the CC process the produced charm quark is only of positive charge, whereas in NC processes positive and negative charge charm quarks are produced.

2.5 Coarse Estimation of the Charm Cross Section in CC Processes

This section shows that a coarse estimation of the charm cross section in CC can be done without any complicated MC programs or extensive QCD calculations. Two assumption are made: firstly one does not care about kinematics. And secondly, beauty and top as sea quarks are neglected. The reason for this is their high mass they have compared to the other quarks.

In the case of positrons colliding with protons ($e^+p \rightarrow \bar{\nu}_e X$), and observing charge conservation, there are only eight possible transitions between quarks.

$$\mathcal{X}_{du} + \mathcal{X}_{dc} + \mathcal{X}_{sc} + \mathcal{X}_{su} + \mathcal{X}_{\bar{u}\bar{d}} + \mathcal{X}_{\bar{c}\bar{d}} + \mathcal{X}_{\bar{c}\bar{s}} + \mathcal{X}_{\bar{u}\bar{s}} = 1, \tag{2.14}$$

where $\mathcal{X}_{ab}$ is the fraction of CC events where a quark a changes into a quark b (see figure 2.4). The sum of all these transitions must be one.

Some fractions can be neglected. The charm mass is larger than the light quark mass and so charm in the sea is neglected, this means $\mathcal{X}_{\bar{c}\bar{s}} \approx \mathcal{X}_{\bar{c}\bar{d}} \approx 0$. Even $\mathcal{X}_{\bar{u}\bar{s}} \approx \mathcal{X}_{su} \approx 0$ because

[2] The $\bar{u} \rightarrow \bar{b}$ is a strongly suppressed transition, because it skips one generation.

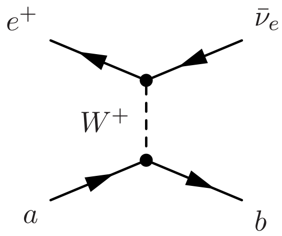


Figure 2.4: A general CC quark-initiated diagram. If beauty and top quarks are neglected, then only eight quark-initiated diagrams are possible in $e^+p \to \bar{\nu}_e X$. The value $\mathcal{X}_{ab}$ is the fraction of CC events where a (anti)quark a changes into a (anti)quark b.

this is a combination between a sea quark and a Cabbibo-suppressed transition.

The ratio between charm in CC over all CC events can now be written as

$$\frac{\sigma(e^+p \to \bar{\nu}_e cX)}{\sigma(e^+p \to \bar{\nu}_e X)} = \frac{\mathcal{X}_{sc} + \mathcal{X}_{dc}}{\mathcal{X}_{du} + \mathcal{X}_{sc} + \mathcal{X}_{dc} + \mathcal{X}_{\bar{u}\bar{d}}}. \tag{2.15}$$

A further approximation is $\mathcal{X}_{\bar{u}\bar{d}} \approx \mathcal{X}_{sc}$ because both are Cabbibo-allowed transitions and a sea density distribution for $\bar{u} \approx s$ is supposed.

A little bit harder to believe is $\mathcal{X}_{dc} \approx \mathcal{X}_{sc}$: on the one hand one has a Cabbibo-suppressed transition ($d \to s$), but d is a valence quark and on the other hand it is a Cabbibo-allowed transition ($s \to c$), but s is a sea quark. According to [30] this approximation is acceptable.

After all these approximations the solution is

$$\mathcal{X}_{\bar{u}\bar{d}} \approx \mathcal{X}_{sc} \approx \mathcal{X}_{dc} = \frac{V_{cd}^2}{V_{ud}^2} \mathcal{X}_{du}. \tag{2.16}$$

The last part of the equation is exact, because it needs only the Cabbibo angles. Using equation (2.16) and equation (2.15) the result is then

$$\frac{\sigma(e^+p \to \bar{\nu}_e cX)}{\sigma(e^+p \to \bar{\nu}_e X)} \approx \frac{2\mathcal{X}_{dc}}{3\mathcal{X}_{dc} + \frac{V_{ud}^2}{V_{cd}^2}\mathcal{X}_{dc}} \approx 10\%. \tag{2.17}$$

About 10% of all CC events should contain one charm quark. With this result the next chapter shows that the charm content in CC at H1 can not be extracted with an exclusive method because of statistics problems.

Chapter 3

The Inclusive Analysis Method

The goal of this analysis is to measure the charm content in NC and CC interactions. The established method to measure the charm content is D^* production based on the Δm-tagging technique[1], where the *exclusive* decay channel $D^* \to D^0\pi \to (K\pi)\pi$ has a branching ratio of 2.6% [31]. The probability for a charm quark to fragment into a D^* is given through the fragmentation factor $f(c \to D^{*+}) = 0.235$ [31]. With the estimation done in section 2.5 the efficiency to find a charm through the Δm-tagging technique in CC is $0.1{\cdot}0.235{\cdot}0.026 \approx 0.06\%$. In the years 1999 and 2000 H1 has collected around 1500 CC events. Using the Δm-tagging technique around one event total is expected. It is clear that an *exclusive* measurement is statistically impossible.

To be independent of an exclusive final state and to obtain more statistics, an *inclusive* method should be used. An inclusive measurement of particle interaction is a partial measurement. Only a few produced particles are singled out for identification and measurement, ignoring the details of all other interaction products. The signature of interesting physics will be defined in terms of few phenomena, like high-p_T leptons or jets. In this inclusive analysis the interesting reaction can be written as

$$e^+p \to n_{measured} + (Q \to H)_{measured} + anything,$$

where n stands for the measured scattered positron e^+ in NC events or the 'unseen' $\bar{\nu}_e$ in CC events and the heavy quark Q which fragments into a bound hadron state H. The signature of the heavy quark will be the flight path of a few 100 μm of the heavy hadron.

3.1 Tagging Signature and Strategy of the Analysis

What is the signature used to tag the heavy quarks?

[1] The low transverse momentum of the pion originating from the decay $D^* \to D^0\pi$ is a characteristic property of this decay chain. The experimental value is $\Delta m = m_{D^*} - m_{D^0} = (145.42 \pm 0.01) MeV$ [28]. This is slightly more than the mass of the pion, so that the strong decay $D^* \to D^0\pi$ is just possible. The small mass difference $m_{D^*} - (m_{D^0} - m_{\pi^+}) = 5.85\ MeV$ is a big advantage for the decay reconstruction: The phase space for combinatorial random combination is very small.

The CST allows to reveal the lifetime of charm and beauty flavored hadrons of around one picosecond. The tagging technique relies on the spatial separation of the *ep*-collision point and the decay vertex, reconstructed by intersecting the tracks of the heavy hadron decay products. In the inclusive method one is not forced to fit all tracks coming from the heavy hadron to a common vertex, two (or three) of them can already determine a decay vertex! With this approach the branching ratio does not anymore limit the measurement.

How can it be confirmed that the tagging algorithms works well?

To justify the correctness of the measurement the important observables are compared between data and simulation in the NC sample, which has about 100 times more statistics than the CC sample. After cross checks in a kinematically restricted NC sample, expecting that the detector and the charm tagging behaviour for NC and CC interactions are the same, the tagging algorithm is used in the CC sample.

Two main differences between NC and CC exist. No positron is measured in CC and the heavy quarks in the NC sample are charm and beauty, where in contrast in the CC sample only charm is measured because of the low statistics collected. In the NC case the scattered positron track candidate is explicitly excluded from the *ep*-collision point fit to get the same behaviour as in CC, where no scattered positron exists.

Which strategy is chosen to extract the cross section?

The idea of the analysis can be summarized as such: To distinguish an event with heavy quarks from an event with light quarks the finite lifetime of the heavy hadrons is used, which leads to a displacement of tracks from the *ep*-collision point. Using MC program templates of a lifetime related observable for heavy and light quark distributions the light, charm and beauty fractions of the data are extracted with a fit procedure.

3.2 Measurement of the Cross Section

The measurement is performed in the visible kinematic range defined for the NC and CC[2] events by $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$. The cross section is calculated using the formula

$$\sigma(e^+p \rightarrow nQX) = \frac{f_Q \cdot N_{data}}{\varepsilon_Q \cdot \mathcal{L}}, \tag{3.1}$$

where f_Q is the charm or beauty fraction evaluated by the fit (see section 3.6) and N_{data} is the number of events selected in the final sample. In NC events n stands for the measured scattered positron e^+, whereas in CC events n stands for $\bar{\nu}_e$. The average efficiency

$$\varepsilon_Q = \frac{N^Q_{rec}}{N^Q_{gen}} \tag{3.2}$$

[2]To select CC events a cut $p_{T,miss} > 16\ GeV$ is needed, which, because of equation 4.3, is equivalent to a $Q^2 > 284\ GeV^2$ cut. The extrapolation to the used kinematic region is about 12% of the cross section.

is calculated with the MC simulation. For N^Q_{gen} one counts only the generated events which fulfill the cuts $Q^2_{gen} > 150\ GeV^2$ and $0.1 < y_{gen} < 0.7$. The number of events which satisfy all selection cuts is given by N^Q_{rec}. $\mathcal{L}$ denotes the integrated luminosity.

3.3 The Structure of the Final State: The Jet

The structure of the events suggests that the quarks in the Quark Parton Model, even when the individual quarks are not observable, give rise to jet-like configurations of hadrons. The produced or scattered quarks materialize into jets of hadrons which have momenta roughly collinear with the original quark. The jets provide a view of the hard quark interactions that occur at small distance scales.

When evaluating higher order QCD cross sections the low-momentum (soft) and small-angle (collinear) regions produce singularities. In order to handle these divergences, the observable has to be a jet quantity. This means that the observable value has to be independent of the precise number of soft and collinear particles in the final state [32]. From these properties a jet algorithm can be derived.

In this analysis the jets have a minimum transverse momentum $p_{t_{jet}}$ of 4 GeV and are reconstructed using the invariant k_t [33] algorithm in the laboratory frame. The object used for the jet algorithm is a combination of tracks and calorimeter energy deposits, the so called HFS objects used in the H1PHAN software package. The information used for this analysis from the jet algorithm are the direction of the jet axis (η_{jet}, ϕ_{jet}) and the transverse momentum $p_{t_{jet}}$.

3.4 The Observables

By reconstructing an isolated decay vertex a long-lived hadron is assumed. Therefore the ep-collision point (primary vertex) defines the heavy hadron production point and its decay point (secondary vertex) is determined by intersecting the tracks of the heavy hadron decay products. Two observables are important: the radial decay length significance and the impact parameter significance.

The separation from primary ($\vec{r}_{pv}$) to secondary ($\vec{r}_{sv}$) vertex can be described with the signed projected radial decay length l as

$$l = (\vec{r}_{sv} - \vec{r}_{pv}) \cdot \left(\frac{\vec{p}_{t_{jet}}}{|\vec{p}_{t_{jet}}|} \right), \tag{3.3}$$

where one uses $\vec{p} = (p_x, p_y)$, $|\vec{p}_t| = \sqrt{p_x^2 + p_y^2}$ and $\vec{r} = (x, y)$. Putting this into words it means: The decay length l is the projection of the distance between the primary and secondary vertex onto the jet axis in the $r\phi$-plane. A schematic picture of the decay length is shown in figure 3.1.

The decay length significance S_l is then given by

$$S_l = \frac{l}{\sigma_l}, \tag{3.4}$$

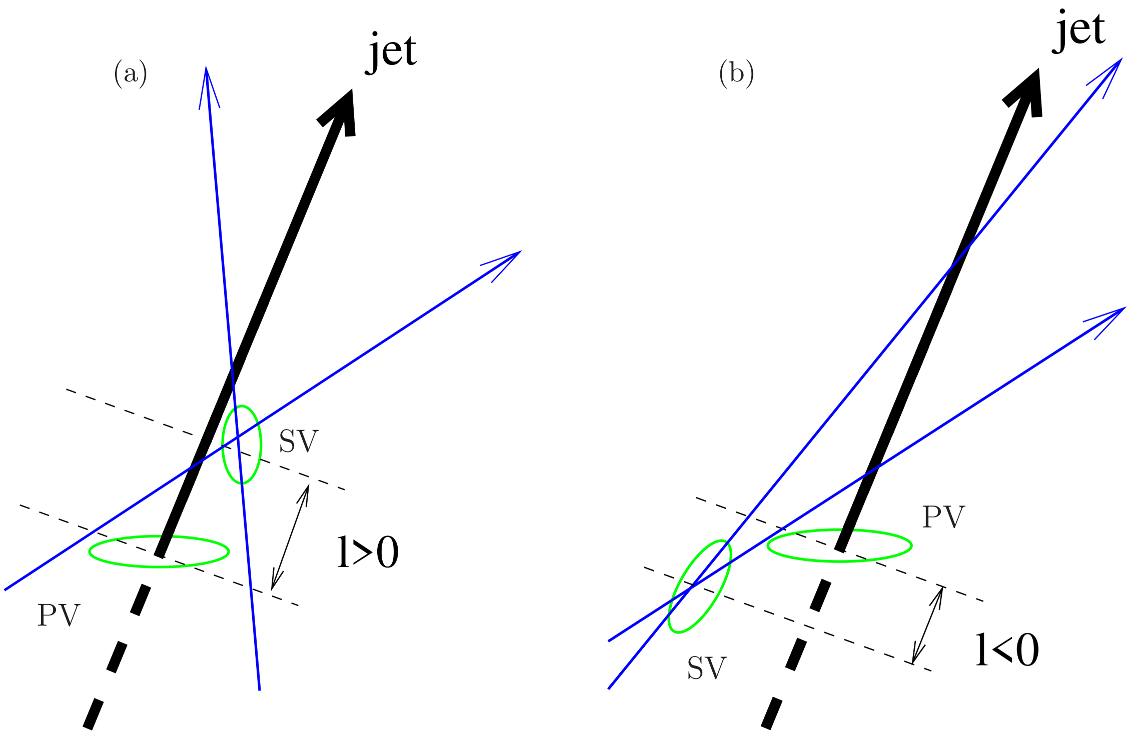


Figure 3.1: Schematic pictures of the decay length, where two tracks emerge from the secondary vertex: The two ellipses show the error of the primary (PV) respectively secondary vertex (SV). (a) Presents a positive decay length projected relative to jet axis and (b) displays a negative projected decay length. The sign is positive for decays with consistent decay and flight directions and it is negative for decays with opposite decay and flight directions which are inconsistent with a physical long-lived hadron.

where the error σ_l is composed of the primary and the secondary vertex resolution. The error σ_l is dominated by the resolution of the secondary vertex reconstruction.

Even with one track one can quantify if a track originates from a primary or a secondary vertex. This is done with the impact parameter significance S_d as

$$S_d = \frac{d}{\sigma_d}. \tag{3.5}$$

The impact parameter d denotes the minimal transverse distance of the non-vertex-fitted track to the event's primary vertex. Its sign is determined by the jet direction, namely the sign is positive if the intercept of the track with the jet axis is downstream of the primary vertex, and negative otherwise. The schematic picture of the impact parameter is presented in figure 3.2.

In figure 3.3 is manifested that heavy hadrons have lifetimes which are large enough to lead to measurable decay length or impact parameter significance within the H1 experiment. The long-lived heavy hadron decays follow an exponential decay law, so that they have most

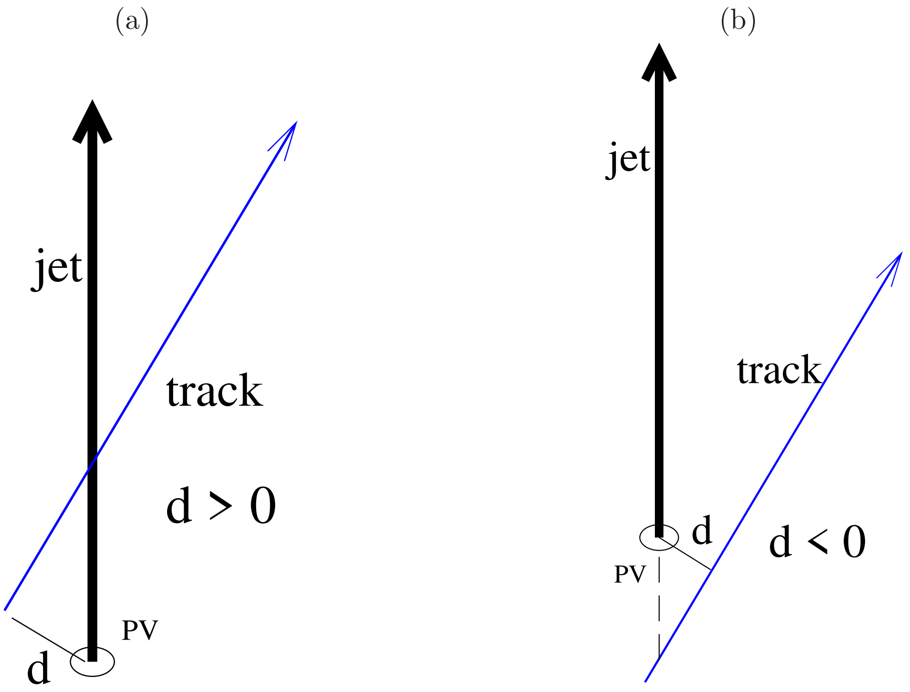


Figure 3.2: The sign of the impact parameter d: The ellipse shows the uncertainty of the primary vertex, labeled as PV. (a) The intercept of the track with the jet axis is downstream of the PV, which gives a positive sign. In (b) the negative sign is displayed, where the intercept is upstream of the PV. This definition makes sense, because decays of long-lived particles are signaled by a positive impact parameter, whereas the finite track resolution yields a symmetric distribution for positive and negative sign of the impact parameter.

a small decay length. In an idealized world tracks from the decays of long-lived particles will mainly have positive true decay length or impact parameter, whilst those produced at the primary vertex will have zero true decay length or impact parameter. Tracks reconstructed with negative decay length or impact parameter values mainly result from the finite detector resolution.

3.5 Different Inclusive Tagging Methods

An inclusive measurement which uses the good track resolution from the CST is published in [1]. In this publication the impact parameter of the tracks instead of the decay length is used. This tagging method allows also to use events with only one track, for which it is not possible to reconstruct a secondary vertex.

Two different inclusive methods will be presented here: the *fixed* vertex multiplicity (FVM) method in subsection 3.5.1 and the *variable* vertex multiplicity (VVM) method in subsection 3.5.2.

Basically both methods use the same vertex fit, the two-dimensional unconstrained track

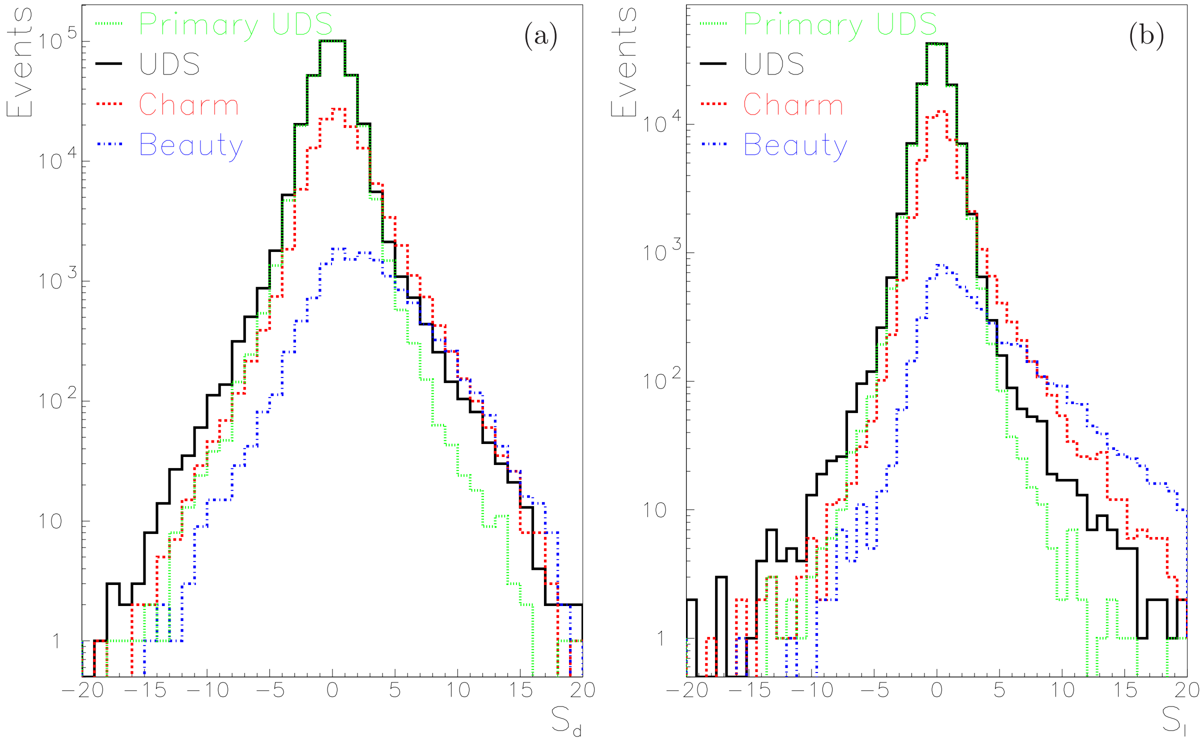


Figure 3.3: The simulated and reconstructed impact parameter and the decay length significance: (a) The reconstructed impact parameter significance from the simulation for uds, charm and beauty events. Primary tracks from uds events are especially plotted to show the resolution effects which is symmetric around zero. The difference between primary uds distribution and uds distribution is mainly due to long-lived strange particles such as K_s. For positive S_d no clear separation between the uds, charm and beauty events is seen. (b) The reconstructed decay length significance from the 3-tracks method explained in section 3.5.1 is shown. For positive S_l a better separation between the uds, charm and beauty events than for the S_d is seen. The difference in S_l and S_d for charm and beauty are due to the different lifetimes of the produced hadrons. The proportions of the distributions are the ones predicted from the Django MC simulation.

fitter (VFit2du) explained in [4]. But in the FVM method the number of tracks coming from the secondary vertex is restricted to two or three tracks[3], in contrast to the VVM method where one is working without an a priori assignment of tracks to the secondary vertex.

These methods are implemented with the help of W. Erdmann, who supports the SV fitting package [5]. The VVM method, also called *adaptive fitter*, has been developed in the CMS collaboration. W. Erdmann could simplify a version of the adaptive fitter to a version which is applicable for the H1 experiment. In chapter 6 is shown that the VVM method gives more stable results then the FVM method.

[3]Since other people use the same heavy flavor tagging method, in this analysis the one track method is not investigated any more.

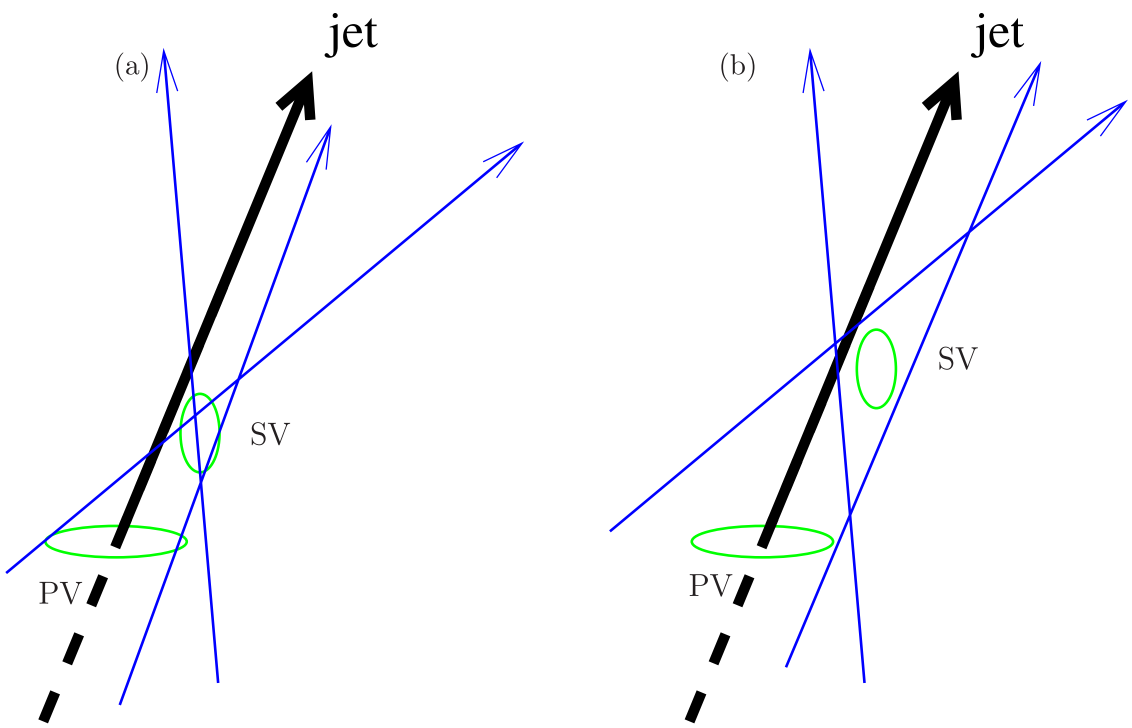


Figure 3.4: Schematic pictures of the 3-tracks method in the FVM method: (a) The candidates which are excepted because the three tracks intersect in the same point which give a high secondary vertex fit probability. (b) The discarded candidate because the three tracks gives a secondary vertex fit probability of less than 10 %.

The next two subsections show how the FVM and the VVM methods work. The input for both methods is the same, namely the following list.

In both methods the algorithm starts with the reconstruction of jets.

For both methods tracks with the measurement of two CST hits in the $r\phi$-plane, the so called CST-improved tracks, are selected, since only CST-improved tracks have the accuracy necessary to resolve the distances between the production and decay vertices of the long-lived heavy hadrons.

Afterwards the CST-improved tracks (η_{tr}, ϕ_{tr}) are selected to be inside a cone with radius one around the jet (η_{jet}, ϕ_{jet}), that is,

$$R = \sqrt{(\eta_{jet} - \eta_{tr})^2 + (\phi_{jet} - \phi_{tr})^2} < 1. \tag{3.6}$$

This produces the list of the CST-improved tracks inside the jet which of course is a subset from the list of all measured tracks.

3.5.1 Fixed Vertex Multiplicity (FVM) Method

For each event two or three CST-improved tracks inside a jet are selected and they are fitted to a common secondary vertex. These methods are referred in this analysis as the 2-tracks and 3-tracks method. A schematic picture of the 2-tracks method is displayed in figure 3.1. The other tracks, CST-improved or not, are used to improve the primary vertex. Only in the 3-tracks method the secondary vertex fit probability is a useful variable, which is used to discard fake secondary vertex candidates. A schematic picture of the 3-tracks method with a sound and a poor secondary vertex is represented in figure 3.4.

If in one jet more than two or three CST-improved tracks exist, then all possible combinations to get several candidates per event are chosen. If several candidates in the event exist, the candidate with largest $min(|S_{d1}|, |S_{d2}|, (|S_{d3}|))$ value[4] is selected. With this selection the candidates with badly reconstructed tracks are rejected. In two dimensions two tracks intersect almost every time and with one well and one badly measured track a quite large decay length value will be determined, which does not reflect the physical situation. For the candidate where one badly measured track exists and the other originates from the primary vertex, the largest $min(|S_{d1}|, |S_{d2}|)$ selection method will chose the candidate with a value of $min(|S_{d1}|, |S_{d2}|) \approx 0$. In the case where at least two CST-improved tracks are reconstructed from a long-lived hadron decay, then both decay tracks have values of $|S_d| \neq 0$. And the selected candidate has a value of $min(|S_{d1}|, |S_{d2}|) > 0$.

The starting components of the FVM algorithm are: one list of CST-improved tracks inside the jet for the secondary vertex fit, and one list with all tracks for the primary vertex fit. The 3-tracks method proceeds then as follows:

1. Fit in 2-dimension ($r\phi$-plane) three CST-improved tracks to a common vertex. This defines the secondary vertex.

2. Determine the primary vertex with the primary vertex fitter routine CSPRIM (for details see [4]). The tracks used in point 1 are explicitly excluded from the primary vertex fit[5].

3. Calculate the decay length l according to equation 3.3.

4. Calculate the secondary vertex fit probability and accept only candidates with a probability with more than 10 %.

5. Make a list of S_l and $min(|S_{d1}|, |S_{d2}|, |S_{d3}|)$.

6. Select the candidate with the largest $min(|S_{d1}|, |S_{d2}|, |S_{d3}|)$.

The result of this algorithm is one candidate per event with all possible information that can be gained with the three selected tracks (e.g. S_l, S_{di}, the transverse momentum for each track, invariant mass of all three tracks, and etc.).

[4]In the 2-tracks (3-tracks) method the candidate with the largest $min(|S_{d1}|, |S_{d2}|)$ ($min(|S_{d1}|, |S_{d2}|, |S_{d3}|)$) value is selected.

[5]The CSPRIM routine itself uses only well measured tracks. These track criteria are given in section 5.2. Additionally, in the NC case the scattered positron candidate is explicitly excluded from the primary vertex fit.

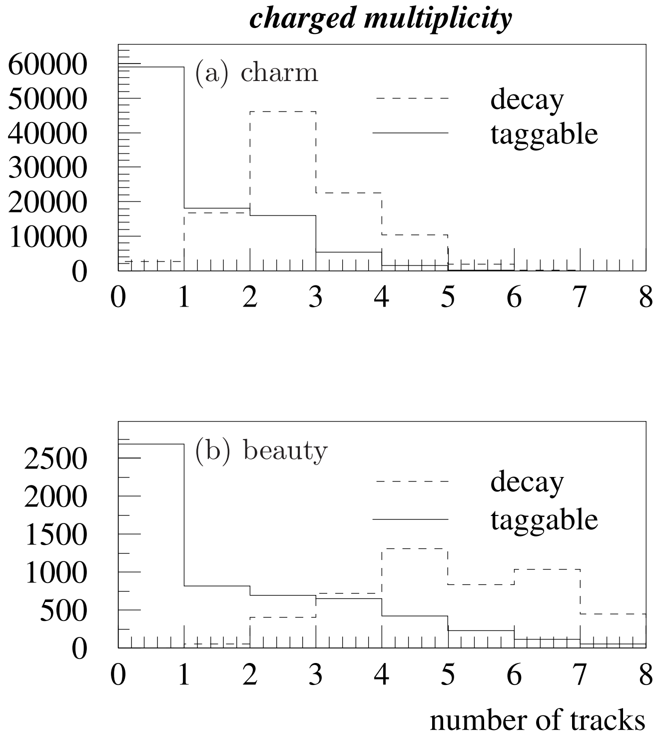


Figure 3.5: The charged decay-multiplicity for (a) charm and (b) beauty hadrons from the Django MC simulation. The dotted line shows the number of generated charged tracks from the decay. The solid line shows the charged tracks which are reconstructed after the CST-improved tracks criteria.

The 2-tracks method works in the same way, but without the secondary vertex fit probability cut, this means without point 4. This is because two tracks intersect almost always in two dimensions with the result of a secondary vertex fit probability of one. It is clear that in the 2-tracks method the selected candidate in point 6 is now the largest $min(|S_{d1}|, |S_{d2}|)$.

3.5.2 Variable Vertex Multiplicity (VVM) Method

In the FVM method the long-lived hadrons are identified only by the separation of the primary vertex from the secondary vertex. But for the long-lived heavy hadrons exist two key signatures: one is the detached vertex, which gives a measurable decay length or an impact parameter and the other is the track multiplicity of this detached vertex. As a result of the bigger mass of beauty hadrons, there are on average more tracks per decay generated, than from charm or light hadrons. The sensitivity to distinguish between light, charm and beauty signatures (e.g. decay length) can be improved, with a better use of the decay-multiplicity information. The different charged decay-multiplicities for charm and beauty (shown in figure 3.5), and the decay length provide a good light, charm and beauty separation.

Instead of the FVM method where a fixed assignment of tracks to vertices is done, a *soft assignment* of tracks to vertices is used in the VVM method. Each track is assigned a weight with a range between 0 to 1 for each vertex candidate, using a weight function. The assignment of the weight is based on the track to vertex distance. The larger the distance of the track to a vertex candidate, the smaller is the weight to this vertex.

The weight assignment of track i to vertex j is defined as

$$w_{ij}(T) = \frac{e^{-\frac{d_{ij}^2}{2T}}}{\sum_{j'} e^{-\frac{d_{ij'}^2}{2T}} + e^{-\frac{\mu^2}{2T}}}, \tag{3.7}$$

where the sum j' runs over all vertices. The weight is $0 < w_{ij} < 1$. The $d_{ij} = d/\sigma_d$ is the impact significance of track i with respect to vertex j. The track has a low weight if the track is incompatible with a vertex and has a weight $w_{ij} \approx 1$ for unambiguous assignment to the vertex. On one hand the overall weight assigned to the track is $\sum_j w_{ij} \approx 1$ if it is compatible with one of the vertices and on the other hand the overall weight assigned to the track can be $\sum_j w_{ij} \ll 1$ if it is not compatible with any vertex.

To see that the μ is a cut off which provides outlier rejection in the case of $d_{ij} > \mu$, assume for example the simplified case with only one vertex (the index j is not needed anymore). Then the weight is given by $w_i = (1 + exp[(d_i^2 - \mu^2)/2T])^{-1}$ [34] and it is used to downweight outlying tracks with $d_i > \mu$. In this analysis the parameter μ is fixed at 2.

In the VVM method a simultaneous fit for one primary and one secondary vertex is made. All tracks with some weak quality cuts of the event are considered for the primary vertex, whilst only CST-improved tracks inside the jet contribute to the secondary vertex. The vertex configuration that minimizes the global fit χ^2 is found iteratively using deterministic annealing[6] [35, 36, 37]. This fit determines the vertex position and the track assignments simultaneously! The parameter T in equation 3.7 is identified in deterministic annealing as the temperature and is sequentially reduced during the iterative fitting process.

The starting components of the VVM algorithm are: one list of the CST-improved tracks inside the jet which are considered for the primary and secondary vertex, and one list with tracks with some weak quality cuts for only the primary vertex. The determination of decay-multiplicity and decay length in the VVM method proceeds then as follows:

1. Choose two vertex prototypes.[7]

2. Set $T = T_{start}$.

3. Calculate the probability according to equation 3.7 for each track to belong to the primary or secondary vertex, respectively.

[6] The deterministic annealing technique is used for the optimization problem. Traditional optimization algorithms reduce at each iteration the cost function. For this reason they tend to get trapped in a local minimum. Deterministic annealing avoids local minima of a given cost function. The cost function is parameterized by the ‘temperature’. This cost function is optimized at each temperature sequentially, starting at high temperature and going down. The deterministic annealing method is independent of the initial choice.

[7] Take one for the primary vertex using the beam spot as constraint and one for the secondary vertex, which stays outside the beam envelope.

4. Refit both vertices individually taking into account the probability calculated in point 3.

5. With the new primary and secondary vertex reduce T.

6. If $T > T_{min}$ go to point 3 otherwise terminate.

The results of this iteration is the primary and secondary vertex position and the track assignment probability. The decay length is calculated with equation 3.3. The number of tracks contributing significantly to the secondary vertex will define a decay-multiplicity. For this analysis the tracks with a weight greater than 80% for the secondary vertex are counted for the decay-multiplicity.

3.5.3 Pro and Con of the FVM and VVM Methods

The FVM method has the advantages to be efficient, because it finds always a secondary vertex when enough CST-improved tracks exist. Therefore a good statistical precision can be achieved. The FVM method has the disadvantage, that when many candidates per jet exist the question arises which one of the many combinatorics should be selected. In the majority of cases the FVM method produces with three primary tracks a secondary vertex which is nearby the primary vertex. Sometimes the FVM method forms three track vertices even when really there are only two (or less) secondary tracks. For beauty events exist many real three track vertices but for charm events most three track candidates contain a primary track. The secondary vertex fit probability cut in the 3-tracks method does not provide sufficient fake secondary vertices rejection. The true track multiplicity for charm and beauty hadron decays enters only indirectly.

The VVM method has the advantage of having only one vertex candidate per jet. The vertex decay-multiplicity is an excellent charm and beauty discriminator, because the average multiplicity of the beauty quark mass is higher than for the charm quark mass. The prompt uds event is suppressed for high vertex multiplicity, which means the method is less sensitive to the uds simulation description. A lower efficiency for beauty and charm is the disadvantage.

3.6 Fit Procedure

The measurement of the cross section is based on the decay length or impact parameter significance distribution, to distinguish the light, charm and beauty contribution. The distinction is only done by a statistical extraction. The relative fraction of the light, charm and beauty component is estimated adjusting the proportion of all three simulated components until the best description of the data is achieved.

The figure 3.3(b) shows the measured light, charm and beauty quark distributions of the decay length significance from the MC simulation. These three different shapes are kept fixed. Only the relative normalizations are left free and are fitted to the measured decay length significance from the data. The fit yields the relative contributions of beauty (f_b), charm (f_c) and light (f_{uds}), where the fit includes the correlation.

If an analytic form or infinite statistics of the simulated and data observable exist, a standard least square fit calculates the light, charm and beauty factors correctly.

If the numbers of data points in many bins are small a maximum likelihood technique using Poisson statistics is the correct way. To account further for the limited statistics of the simulation[8] an advanced binned maximum likelihood method [38] is used, incorporating the Poisson statistics in data and MC simulation. In the Fortran HMCMLL [39] subroutine this method is implemented.

[8] The simulation is subject to statistical fluctuations, too.

Chapter 4

The Neutral and Charged Current Selection

The selection of NC and CC events is only briefly described here. The details can be found in [6, 40, 41, 42, 8, 43]. The starting point is to compare the MC program generated and reconstructed variable which are necessary to measure the NC kinematics. In addition because the CC selection relies heavily on a proper description of hadronic energy distribution, this property is studied in detail in the NC selection, where the hadron and positron energy calibration can be tested. The chapter concludes with the comparison between the data and simulation for NC and CC events.

4.1 Analyzed Data Set, Trigger and Integrated Luminosity

The data are based on e^+p collisions of the years 1999 and 2000 at the center of mass energy of 319 GeV. All important hardware components must be fully operational to select the events. These components are CST and CJC for the track reconstruction, LAr calorimeter and SpaCal for the measurement of the scattered positron, the luminosity system, the Central Proportional Chambers and ToF system. The last period of the 1999 data taking is dismissed due to the major CJC inefficiencies.

The triggering of NC events is based on the signature of high Q^2 in the detector, i.e. high energy scattered positron in the LAr calorimeter. A detailed description of the trigger elements and the selection of NC triggers is given in [44]. The majority of events is triggered by sub-trigger ST67 and some information about this trigger is given in section 1.4.

The integrated luminosity is $56.9 \pm 0.85\ pb^{-1}$.

4.2 Event Simulation

The NC and CC processes are described by electroweak matrix elements with massless heavy quarks, as implemented in the DJANGO program [45]. The event generator DJANGO simulates DIS including both QED and QCD corrections. DJANGO is based on HERACLES [46]

hadron	mass $[GeV/c^2]$	$c\tau$ $[\mu m]$
B^0	5.2794	461
B^+	5.2790	501
B_s	5.3796	438
Λ_b	5.6240	368
D^0	1.8647	124
D^+	1.8694	312
D_s	1.9681	147
Λ_c	2.2849	60

Table 4.1: The actual masses and lifetimes taken from the PDG 2003 web page for beauty and charm hadrons used in the JETSET program for the NC simulation.

for the electroweak interaction and on LEPTO [47] for the treatment of the QCD dynamics. LEPTO uses the color dipole model (CDM) [48] as implemented in the ARIADNE [49] program. The events are generated with a leading order parton density function, which is the PDF CTEQ6L [50]. The JETSET program is used for the hadron fragmentation, which takes the Lund string model [51]. The actual heavy hadron lifetimes are taken from the PDG [28] for the beauty and charm hadrons. The values are given in the table 4.1.

No background simulation is used for NC, because the background is below 1% as described in [42]. Because the analysis is strongly dependent on the heavy flavor decay distributions, one compares the heavy flavor production in NC events with RAPGAP and DJANGO. The main difference between them is that in RAPGAP the masses of the heavy quarks are taken into account with the mass of the charm quark set to $m_c = 1.5\ GeV$ and beauty quark set to $m_b = 5\ GeV$, whereas in DJANGO the heavy quarks are massless. In contrast DJANGO calculates the NC interactions with γ^* and Z^0 exchange, whereas RAPGAP uses only γ^* exchange.

The main background in the CC sample originates from direct and resolved photoproduction (γp) events and from NC events in which the scattered positron is not identified. The γp is simulated using PYTHIA [51, 52] with the GRV [53] leading order parton distribution functions for the proton and photon.

The detector response to events produced by the above programs is simulated in detail using a program based on GEANT. These simulated events are then subjected to the same reconstruction and analysis chain as the real data (cf. section 1.5).

The generated NC (CC) events correspond to a total luminosity of 519.5 pb^{-1} (2680 pb^{-1}).

Three small differences exist between the NC and CC MC simulation. Firstly the chosen PDF is CTEQ5L in CC instead of CTEQ6L in NC. Secondly the fragmentation used in LEPTO is the option MEPS (Matrix Element plus Parton Shower) [23] in CC instead of CDM in NC. Thirdly in the CC simulation the heavy hadron lifetime values are slightly different, but it has been estimated that all three differences do not make any difference in the efficiency determination.

4.3 Reconstruction of Event Kinematics

For NC DIS events there are various methods for the reconstruction of the kinematic variable since there is redundant information from both the lepton and hadron final state [54, 55].

The electron method relies only on the energy E'_e and the polar angle θ_e of the scattered positron:

$$Q^2_e = 4E_eE'_e\cos^2(\frac{\theta_e}{2}) \qquad y_e = 1 - \frac{E'_e}{E_e}\sin^2(\frac{\theta_e}{2}) \qquad x_e = \frac{Q^2_e}{sy_e}, \tag{4.1}$$

with s is the center-of-mass energy squared (see equation 2.1) and the E_e and E_p are the energies of the incoming positron and proton, respectively.

In the hadron method the hadronic final state defines the kinematics completely and must be used for charged current kinematics.

$$\Sigma_h = \sum_i(E_i - p_{z,i}) \qquad P_{T,h} = \sqrt{(\sum_i p_{x,i})^2 + (\sum_i p_{y,i})^2}, \tag{4.2}$$

where the summations are over all particles of the hadronic final state. With these variables the result is:

$$Q^2_h = \frac{P^2_{T,h}}{1 - y_h} \qquad y_h = \frac{\Sigma_h}{2E_e} \qquad x_h = \frac{Q^2_h}{sy_h}. \tag{4.3}$$

Although the particle loss in forward direction in the beam pipe is not so important for Σ_h, the hadron method is influenced by the limited energy resolution of the calorimeter and therefore has moderate precision.

The understanding of the hadronic energy scale is important for the CC measurement. In the next section the electron energy calibration is compared to the hadronic energy calibration. For this one needs the double angle (DA) method, which relies on the positron polar angle and the inclusive hadronic polar angle θ_h:

$$Q^2_{DA} = \frac{E_e}{E_p}\frac{s}{\alpha_e(\alpha_e + \alpha_h)} \qquad y_{DA} = \frac{\alpha_h}{\alpha_e + \alpha_h} \qquad x_{DA} = \frac{Q^2_{DA}}{sy_{DA}}, \tag{4.4}$$

with

$$\alpha_e = \tan\frac{\theta_e}{2} \qquad \alpha_h = \tan\frac{\theta_h}{2} = \frac{\Sigma_h}{P_{T,h}}. \tag{4.5}$$

In the quark parton model θ_h is the polar angle of the scattered quark. This method is particularly useful for the calibration of the calorimeter since it is independent of the calorimeter energy. The energy of the scattered positron determined using the DA method is given by:

$$E_{DA} = \frac{Q^2_{DA}}{4E_e\cos^2(\theta_e/2)}. \tag{4.6}$$

This form is used for the calibration between electron and hadron calorimeter.

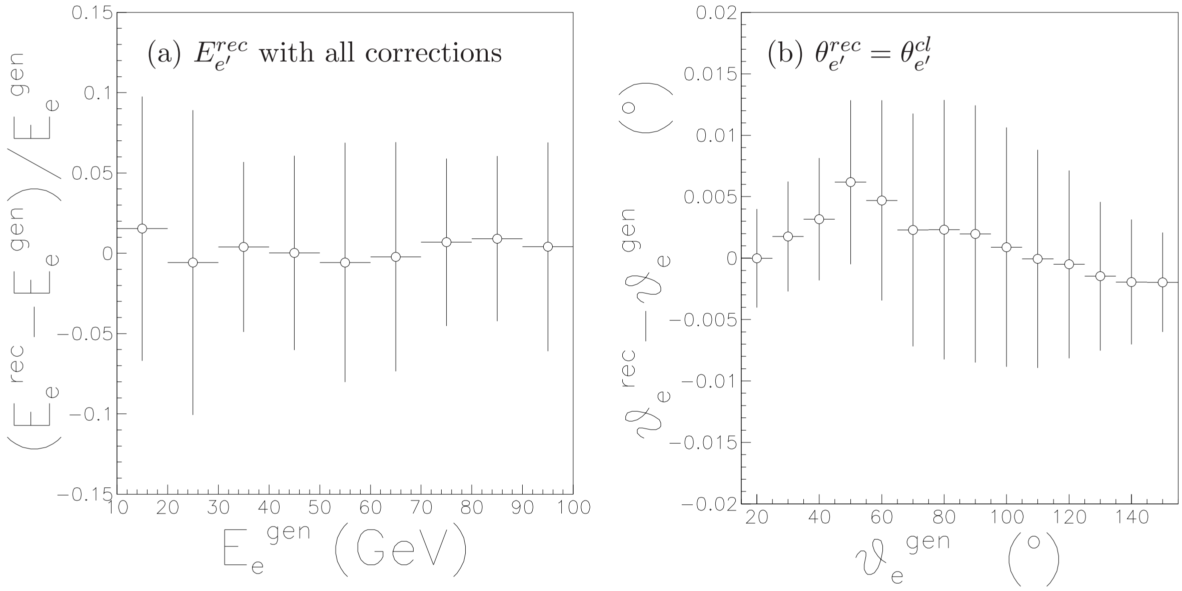


Figure 4.1: The spread between generated and reconstructed values of the scattered positron energy and the scattered positron angle: These two variables are needed to reconstruct the DIS kinematics. Figure (a) represents the variable $E_{e'}$ and figure (b) shows the variable $\theta_{e'}$. The $\theta_{e'}^{rec}$ is calculated using the position of the energy deposit in the LAr calorimeter.

4.4 Control of Energy Calibration

In three steps it will be shown that the energy calibration for the LAr calorimeter used for the NC and CC selection is sufficient. In the first step, the resolution is determined by comparing the reconstructed with the generated values, namely the scattered positron information and the kinematic variables. The second step concentrates on the positron energy calibration for the data and simulation separately. And finally the hadron versus positron energy calibration is compared for data and simulation.

4.4.1 Comparison of Generated and Reconstructed Positron Variables

In the following the generated and reconstructed positron variables are compared.

The measured variables used to reconstruct the kinematics of the DIS events are the energy deposited by the scattered positron in the LAr calorimeter and the polar angle. It is determined using the position of its energy deposit (cluster) in the LAr calorimeter, together with the interaction vertex reconstructed with tracks from charged particles in the event.

Only events with the positron scattered into the LAr calorimeter are considered. The events are simulated with the Django MC program and the selection criteria used are explained in section 4.5. The figure 4.1(a) gives the standard deviations of bin-wise Gaussian fits of the difference between the simulated and reconstructed values of positron energy. The mean value is around zero, which shows that the energy calibration in the simulation has been determined correctly and the resolution for all generated energies above 10 GeV is

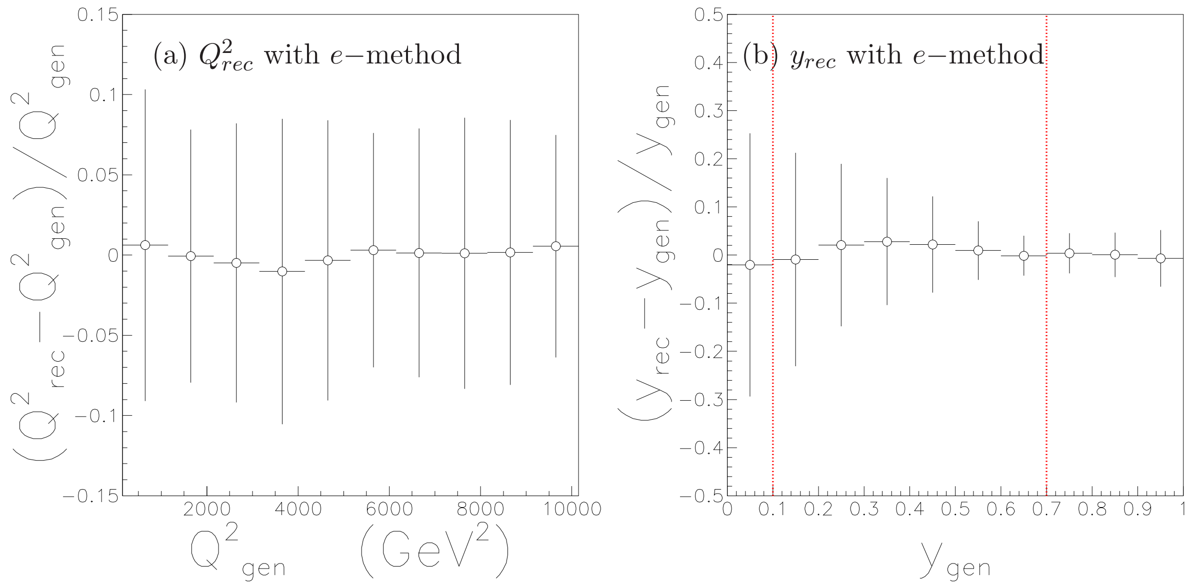


Figure 4.2: The spread of the generated and reconstructed DIS kinematics: In figure (a) Q^2_{rec} and in figure (b) y_{rec} are calculated with the electron method as used for the NC selection. The solid vertical lines in (b) indicate the cuts $0.1 < y < 0.7$.

about 10%. In the simulation (see figure 4.1(b)) only a small difference over the full range between the generated positron polar angle and the one calculated using the LAr calorimeter is observed. The resolution obtained is below $0.01°$.

The effect of this deviation in the Q^2 and y reconstruction with the positron is displayed in figure 4.2. This method gives good resolution in Q^2 and y throughout the full kinematical range.

4.4.2 Electron Energy Calibration

In NC interactions both the scattered positron and the hadronic final state are measured in the H1 detector. This means that the system is over-constrained and different methods can be used for the reconstruction of the event kinematics. The DA method does not rely on the calorimetrically measured energies. Therefore it can be used to test the calibration of the calorimeter.

For the electron energy calibration, the cluster energy in the LAr calorimeter is compared to the positron energy calculated with the DA method. The ratio E_{CL}/E_{DA} as function of ϕ, Z and time is presented in figure 4.3. No significant detector effects are seen but there is a systematic shift between data and simulation in time (for 2000 data) which has not been corrected. The maximal relative difference between data and simulation is seen in figure 4.3(b) and is less than 3%.

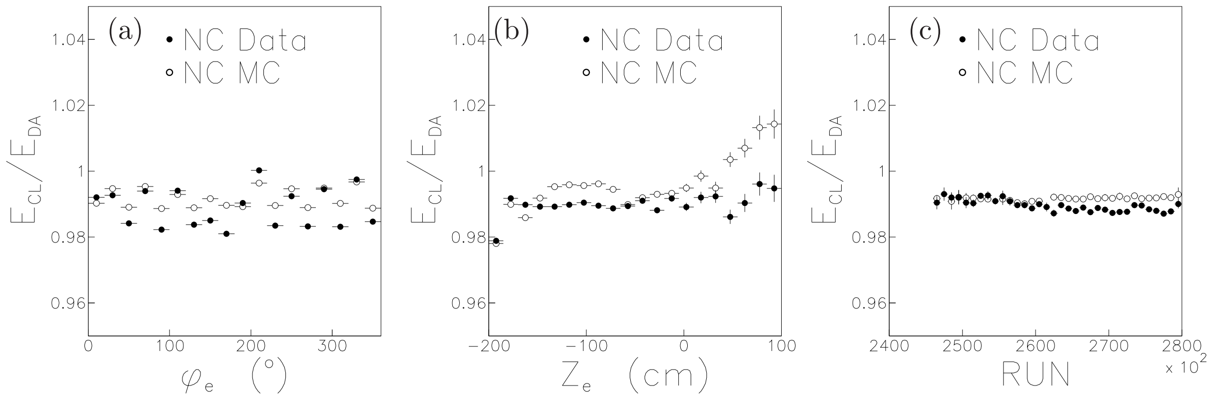


Figure 4.3: The electron energy calibration: For the data and simulation separately, the ratio of the scattered positron energy measured in the LAr calorimeter E_{CL} and the positron energy calculated using the DA method E_{DA} is shown as function of: (a) the azimuth angle of the positron, (b) the Z position of the positron impact point and (c) the run numbers.

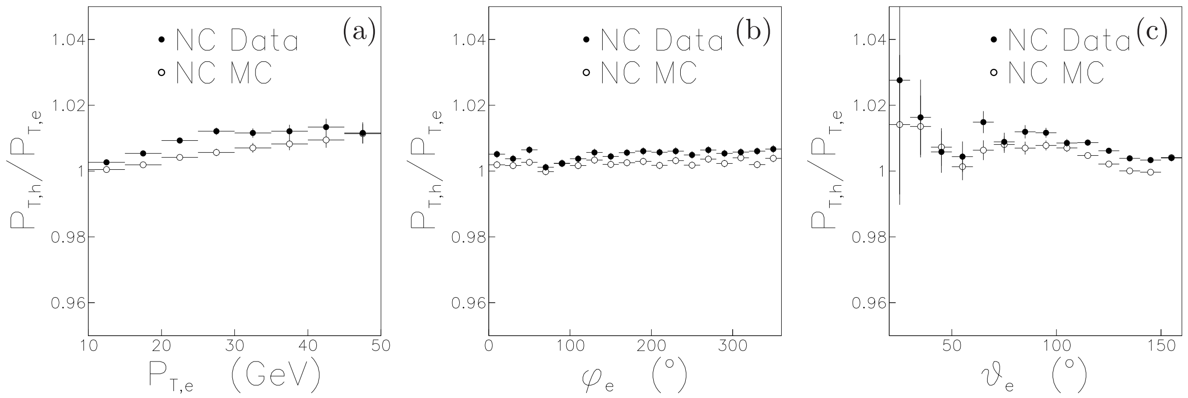


Figure 4.4: The hadron and electron energy calibration: The mean values of transverse momentum of the hadronic final state $P_{T,h}$ divided by the mean values of the scattered positron $P_{T,e}$ is displayed as a function of: (a) $P_{T,e}$, (b) the azimuth angle ϕ_e and (c) the polar angle θ_e, for the data and simulation separately.

4.4.3 Hadron versus Electron Energy Calibration

A precise hadronic energy measurement is crucial for all measurements relying on the hadronic final state. The test of the hadronic and electronic energy calibration is based on the transverse energy conservation, namely the transverse momentum of the hadronic final state $P_{T,h}$ should be equal to that of the scattered positron $P_{T,e}$. These two variables are on one side only hadronic energy measurements and in the other side only positronic energy measurements.

In figure 4.4 the mean value of the ratio $P_{T,h}/P_{T,e}$ is shown. It demonstrates the quality of the hadronic calibration in the region $10\ GeV < P_{T,e} < 50\ GeV$, over the full azimuth angle and over the full polar angle. No detector effects are seen. A good agreement is found to within 3%.

4.5 Neutral Current Measurement

The criteria used for the NC events do not differ in essence from the one used in other NC analyses [6, 40, 41, 42]. The characteristic feature of DIS is an identified scattered positron in the LAr calorimeter and the hadronic final state which is mainly measured in the LAr calorimeter as well.

DIS events are selected by requiring a compact and isolated energy deposit in the electromagnetic part of the LAr calorimeter. The scattered positron is identified as the cluster of highest transverse momentum. The cluster has to be associated with a track measured in the inner tracking chambers. This combination of an electromagnetic cluster finder and a track cluster link is done by a software package XASELE.

The NC events are required to have an interaction vertex within ± 20 cm of the nominal z position to suppress beam induced background and also to be in the acceptance of the CST. NC events are triggered mainly using information from the LAr calorimeter. The sub-triggers used are ST67, ST71, ST75 and ST77[1]. If the scattered positron energy $E_{e'}$ measured in the LAr calorimeter is greater than 11 GeV, the trigger was fully efficient, except in a small fraction of inefficient regions. In an other analysis [6] these regions are excluded by applying fiducial cuts, which are not used in this analysis. The influence of these inefficient regions have not been observed in this analysis and therefore they are retained.

Because of LAr calorimeter geometry, the experimental control of efficiencies and calibration becomes difficult for some Z-crack and ϕ-crack. The discarded region are $Z_e < -190$ cm, 15 $cm < Z_e < 25$ cm and $|\phi_e - n \cdot 45^o| < 2^o, n = 0, 1, ..., 7$.

The CJC-T0 timing is used to suppress background events originating from beam-halo events or cosmic muons. In addition the background events are rejected by a few topological background filters [56, 57], which are implemented in the software package QBGFMAR.

The energy and momentum conservation[2] requires $(E - P_z)_{total} \equiv (E'_e - P_{z,e}) + \Sigma_h = 2E_e$, where a smaller value is expected for the γp events as the contribution from the scattered positron is not included in the sum. Restricting the measured $(E - P_z)_{total}$ to be greater than 35 GeV reduces the γp events.

The figure 4.7(a) shows the measurable kinematic range for NC events with jets in the center of the detector. For this analysis the tracks need to pass through the CST detector; this means the hadronic final state going forward is outside the CST acceptance and low values of y are discarded. Thus the analysis is restricted to $0.1 < y < 0.7$. The upper y limit ensures good trigger acceptance for the scattered positron. Additionally, $Q^2 > 150$ GeV^2 was imposed, because on one hand the positron variables have to be measured in the LAr calorimeter and on the other hand the same kinematical region as used in the CC measurement is chosen.

[1]The sub-trigger information is not used for the simulation because it is not well implemented.

[2]The situation before the scattering is $(E - P_z)_{total,initial} = (E_e - P_{e,z})_{initial} + (E_p - P_{p,z})_{initial} = 2E_e$ by neglecting the positron and proton mass.

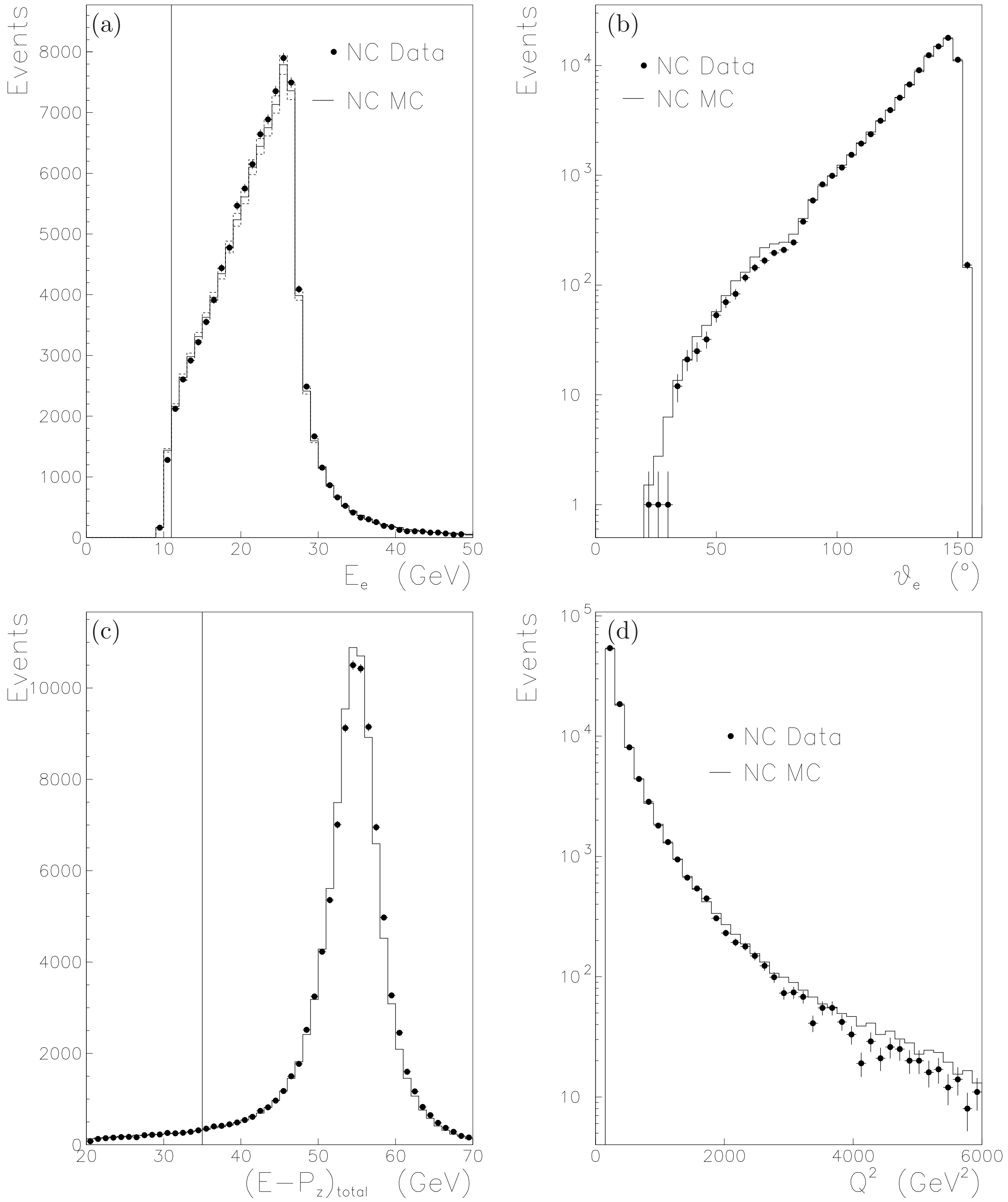


Figure 4.5: Comparison of the data description by the simulation for the selected 99-00 e^+p NC events: Distributions of (a) the scattered positron energy E_e', (b) the polar angle θ_e, (c) the longitudinal momentum balance $(E - P_z)_{total}$ and (d) the distribution of the kinematic variable Q^2 are presented. The simulation is normalized to the experimental luminosity. The vertical line in (a) indicates the cut $E_e \geq 11\ GeV$ and the dotted histograms are $\pm 2\%$ around the simulation prediction. The vertical line in (c) indicates the cut $(E - P_z)_{total} \geq 35\ GeV$.

Kinematic	Background	Quality
$Q^2 > 150\ GeV^2$	QBGFMAR	Run selection
$0.1 < y < 0.7$	non-ep background cuts	ST67, ST71, ST75, ST77
$(E - P_z)_{total} > 35\ GeV$	$\lvert Z_{vtx}\rvert < 20\ cm$	Z-crack and ϕ-crack rejection
$E_{e'} > 11\ GeV$	CJC-T0 timing	

Table 4.2: The final neutral current selection criteria.

Due to the large number of inefficient CJC1 cells, the run range from 257637 to 262144 is excluded. This is the reason for the around 12% lower luminosity between this analysis and the other NC analyses [42, 6].

A luminosity of $\mathcal{L} = 56.9 \pm 0.85\ pb^{-1}$ is collected. After applying the NC selection criteria summarized in table 4.2 the NC data sample contains about 95000 events. The comparison of the data and the simulation is presented in figure 4.5 for the scattered positron energy distribution, the polar angle variable and the distribution $E - P_z$ and Q^2. All distributions are well described by the simulation, which is normalized to the luminosity of the data.

4.6 Charged Current Measurement

The criteria used for the CC events are described in more detail in [41, 40]. The CC events are characterized by a missing transverse momentum due to the undetected neutrino in the final state. For CC events P_T^{miss} is equivalent to $P_{T,h}$.

The main selection cut to suppress background is $P_{T,h} > 16\ GeV$. This background is split in ep and non-ep background. The non-ep background like halo muons, cosmic rays and beam-gas and beam-wall interactions is rejected by the vertex requirement, by the timing information from the CJC and by the topological filters implemented in QBGFMAR. The ep background originates from γp and NC events.

Mismeasurement of energies and limited geometrical acceptance can in both cases lead to events which are not balanced in transverse momentum. To suppress these events the ratio V_{ap}/V_p is used. The quantities V_p and V_{ap} are the transverse energy flow parallel and anti-parallel to $\vec{P}_{T,h}$ respectively. They are determined from the transverse momentum vectors $\vec{P}_{T,i}$ of all particles i in the hadronic final state according to

$$V_{ap} = -\sum_i \frac{\vec{P}_{T,h} \cdot \vec{P}_{T,i}}{P_{T,h}} \quad for \quad \vec{P}_{T,h} \cdot \vec{P}_{T,i} < 0$$

$$V_p = \sum_i \frac{\vec{P}_{T,h} \cdot \vec{P}_{T,i}}{P_{T,h}} \quad for \quad \vec{P}_{T,h} \cdot \vec{P}_{T,i} > 0.$$

For well measured CC events with only one jet, the transverse momentum component opposite to the jet (V_{ap}) is expected to be much smaller than that in the direction of outgoing jet (V_p) thereby leading to $V_{ap}/V_p \approx 0$. For NC and γp events the energy is expected to be distributed isotropically, resulting in higher values of V_{ap}/V_p. The cut defined is $V_{ap}/V_p < 0.1$ for all

Kinematic	Non-CC rejection	Quality
$P_{T,h} > 16\ GeV$	QBGFMAR	Run selection
$0.1 < y_{had} < 0.7$	non-ep background cuts	ST66, ST67, ST71, ST77
$Q^2_{had} > 300\ GeV^2$	$V_{ap}/V_p < 0.1$	$\lvert Z_{vtx}\rvert < 20\ cm$
	$\Delta\phi_{h,PLUG} - V_{ap}/V_p$ plane cut	CJC-T0 timing

Table 4.3: The final charged current selection criteria.

events with $P_{T,h} > 16\ GeV$, which significantly suppresses the background. For events with $P_{T,h} < 25\ GeV$ a further $P_{T,h}$ dependent cut is used in the $\Delta\phi_{h,PLUG}$ and V_{ap}/V_p plane, which is described in [41]. This cut rejects about 20 γp background events and retains most of the CC events.

The sub-triggers used are ST66, ST67, ST71 and ST77. The CC event selection is based mainly on the missing transverse energy provided by the LAr calorimeter, in coincidence with timing information from either the proportional chambers or the LAr calorimeter. From the detector point of view, the hadronic final state of a CC event is the same as that of a NC event. The trigger efficiencies are again not used in the MC program. The 99-00 CC trigger efficiency is taken from a $P_{T,h}$ and γ_h parametrisation which can be found in [58]. The CC trigger efficiency is taken from the high statistics NC events from the real data when all information associated to the scattered positron is removed. A short explanation of this so called pseudo-CC sample is given in [40]. The average weight of the selected CC data events caused by the trigger efficiency is about 1.05.

The run selection, the vertex requirement, the CJC-T0 timing and the inelasticity range for the CC analysis is equivalent to the NC analysis[3]. Again the same run range as in NC is excluded because of the large number of inefficient CJC1 cells. Figure 4.7 shows that the measured Q^2 range in CC is higher than in NC. The $P_{T,h} > 16\ GeV$ cut is, because of equation 4.3, equivalent to a $Q^2 > 284\ GeV^2$ cut. The extrapolation to the same kinematic range as in NC is about 12%.

Further rejection is done by using the topological filter QBGFMAR, which is based on the characteristic signature of cosmic or halo muon events and the accidental overlap of cosmic or halo muons with γp events. QBGFMAR is optimized for the inclusive CC analysis at $P_{T,h} > 12\ GeV$.

The luminosity is the same as in NC. After applying the CC selection criteria summarized in table 4.3 the final CC data sample contains about 800 events. The CC data and simulation are compared in figure 4.6 for the $P_{T,h}, E_h - P_{z,h}, Q^2_{had}$ and y_{had} distributions. The simulated backgrounds for CC are the resolved γp, the direct γp and NC. The main background shown in figure 4.6 is due to the resolved γp simulation. The simulation gives a rather good description of the selected data.

[3]The upper inelasticity cut could be extended to $y < 0.85$ as in the nominal CC selection. The number of selected CC data events would increase from about 800 to 850 events. All tests for CC are always done with the NC sample restricted in $0.1 < y < 0.7$, so that the same restricted y range for CC as for NC is used.

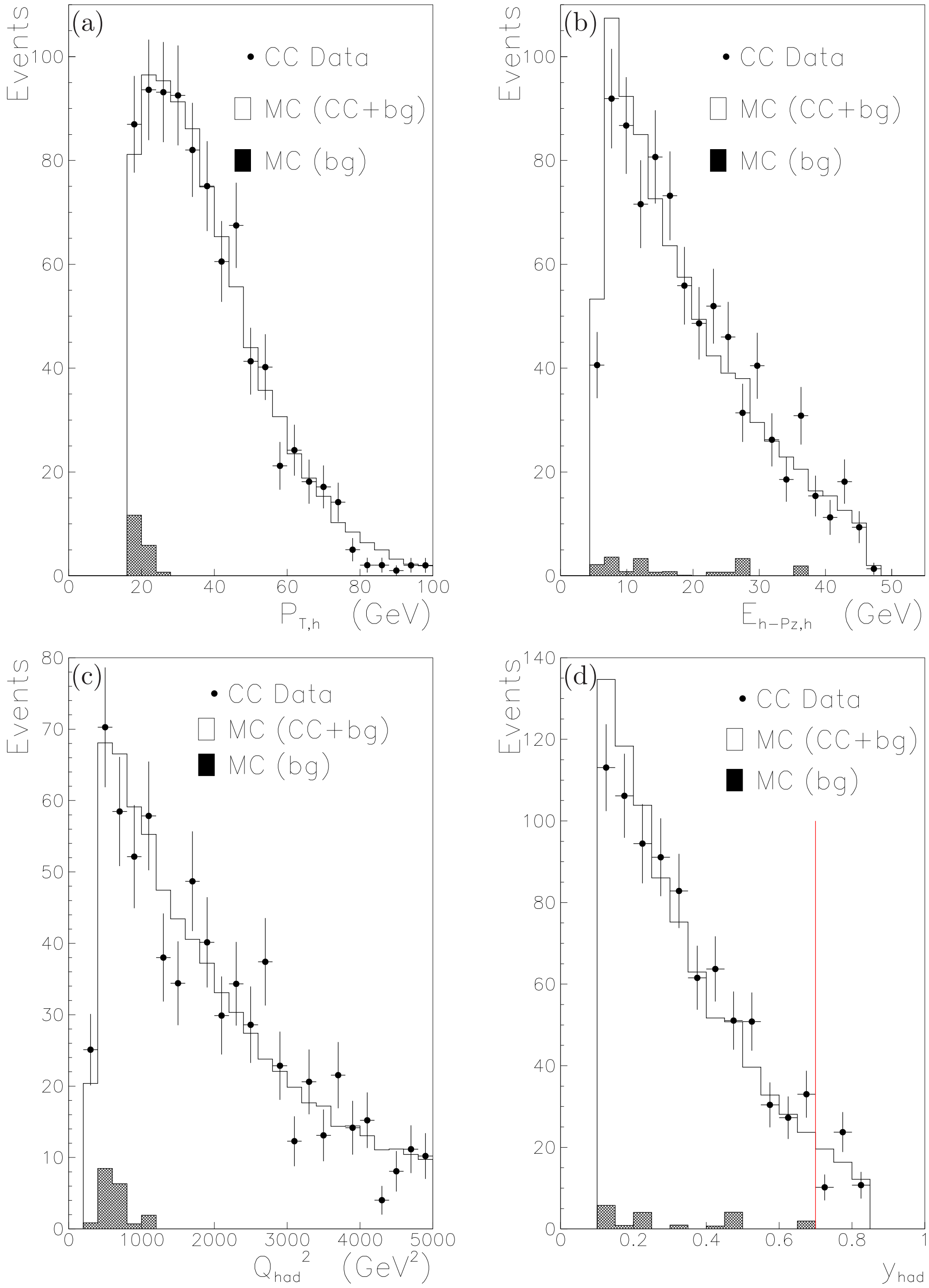


Figure 4.6: Comparison of the data description by the simulation for the selected 99-00 e^+p CC events: Distributions of (a) the missing transverse momentum $P_{T,h}$, (b) $E_h - P_{z,h}$, (c) Q^2_{had} and (d) y_{had} for CC data (points) and simulation (solid line) are compared. The filled histograms show the γp and NC contribution. All simulations are normalized to the experimental luminosity. In (d) the upper inelasticity cut at $y = 0.7$ is shown.

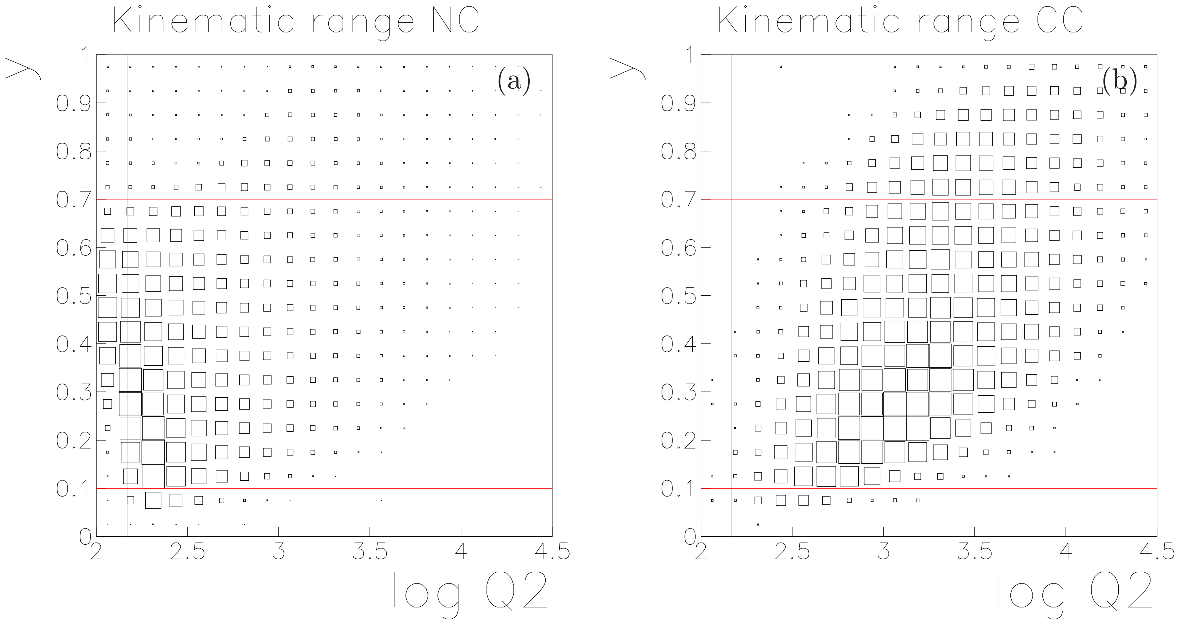


Figure 4.7: The measurable kinematic range for NC and CC events with jets in the center of the detector: Figure (a) shows NC events with tracks passing through the CST detector. Figure (b) represents CC events with tracks passing through the CST detector and the $P_{T,h} > 16\ GeV$ cut. The solid horizontal line indicate the cuts $0.1 < y < 0.7$ and the solid vertical lines indicates the cut $Q^2 > 150\ GeV$. These cuts define the kinematic range for the measurements. The extrapolation to this kinematic range defined for the CC measurement is about 12%.

Chapter 5

Track, Primary Vertex and Jet Reconstruction

The lifetime of around 10^{-12} seconds of the charm and beauty hadrons leads to a decay length of a few hundred μm. To distinguish signal and background events on the basis of the lifetime information the decay topology is reconstructed. The primary event vertex is used to define the heavy hadron production point, and its decay point is determined by intersecting some of the hadron decay tracks, which is referred to as the secondary vertex. Because of the low-z-resolution described in section 1.3, the measurement relies on the $r\phi$-projection alone. The separation between primary and secondary vertex is expressed also in the radial decay length.

To reconstruct the decay topology the precise knowledge of the primary vertex and tracks is required. The jets are only needed to give the flight direction of the produced or scattered quark. To persuade the reader that the heavy quark tagging algorithm works, an understanding of the tracks, the primary vertex and jets is important.

5.1 Track Reconstruction

The projection of a charged particle track in a homogeneous magnetic field is a circle with a radius R. The transverse momentum is proportional to the charge of the particle, the strength of the magnetic field and the radius ($p_t = c \cdot qBR$). In three dimensions the flight path is described by a helix. This is insofar an approximation as the flight direction is distorted by multiple scattering and energy loss in matter and in addition the magnetic field has inhomogeneities. These effects are corrected by the routine CSTCOR [4] and are important for tracks which are reconstructed in the CJC and then extrapolated outside of the measured volume. This is the case in the reconstruction chain for the tracks used. The track first reconstructed in the CJC is then extrapolated into the silicon layers and the CST hit closest to these CJC crossings is associated with the track. The track is refitted using the additional information of the CST hits in order to improve the precision of the track parameters. The analysis is based on tracks which are linked to hits in both planes of the CST, which will be called CST-improved tracks.

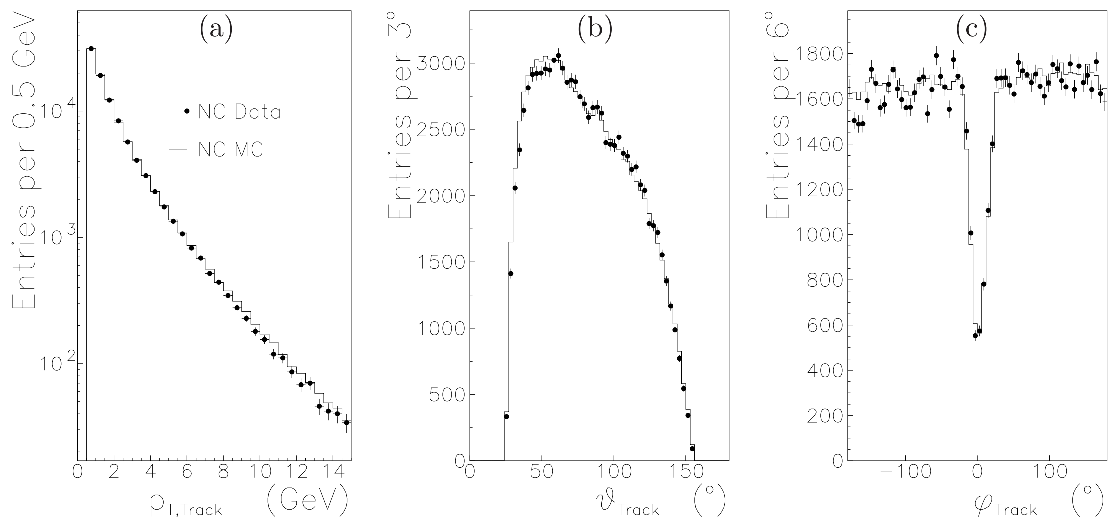


Figure 5.1: (a) The transverse momentum distribution, (b) the polar angle distribution and (c) the azimuth angle distribution of CST-improved tracks which are selected for the secondary vertexing in the 2-tracks method in the NC sample. Data (dots) with statistic errors are well described by the simulation (histogram). The simulation is normalized to the data luminosity.

5.1.1 CST-improved Tracks

In the following section the track selection criteria for the secondary vertex fit are described.

Since the CJC covers the full CST acceptance, only tracks measured in the CJC are used. Selected are the charged tracks in the central detector region, i.e. $|\eta| < 1.5$.

The hit measurements of the CJC are the basis for the reconstruction of tracks. But only the position measurement of the CST hits have the accuracy necessary to resolve the rather small distances separating the production and decay vertices of the selected heavy hadrons. So the tracks are first reconstructed with the CJC and are then extrapolated into the CST, where the closest $r\phi$ point measured in each CST layer is used to get the CST-improved tracks. Only the transverse CST information is used.

The resolution of CST-improved tracks arises from multiple scattering in the material in front of the detector, which are the beam pipe and the silicon layers. Wrongly linked CST hits are another source of error.

For a reliable determination of the track parameters, a minimum track length in the $r\phi$-plane of 22 cm is required with the trajectory starting in the volume of the CJC1 at $R_{start} < 50\ cm$. A minimum transverse momentum of $p_t > 0.5\ GeV$ is required for all tracks, to reduce the effect of multiple scattering. All tracks must come from the region around the primary vertex, so that the impact parameter is $|d| < 0.2\ cm$. This cut reduces the effects of long lived strange particles such as K_s and Λ, cosmic ray shower particles and secondary interactions like photon conversions at the silicon layers. Only those tracks which are linked to hits in both planes of the CST with a probability of greater than 10% from the CJC-CST

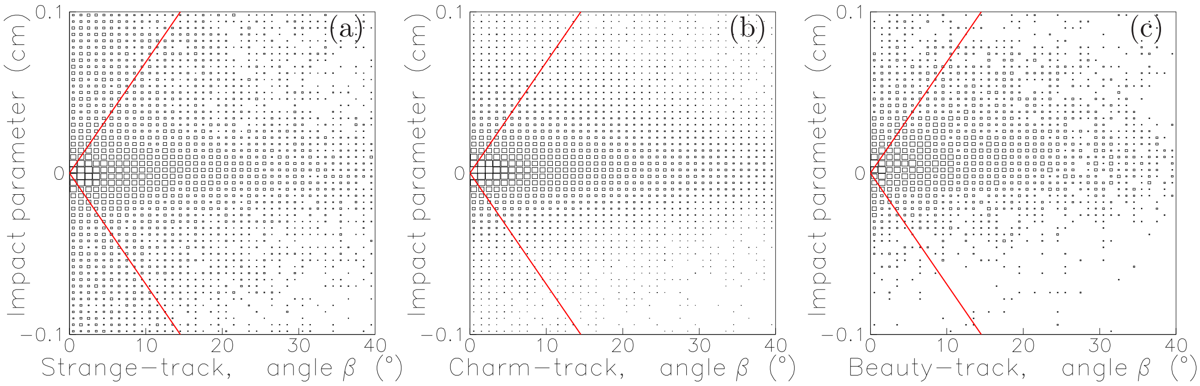


Figure 5.2: The simulated impact parameter versus the absolute value of the angle β from jet and track for decaying tracks produced for different mother tracks are shown: In (a) the decaying tracks produced from K_s and Λ, in (b) produced from $D^0, D^{\pm}$ and D_s and in (c) produced from $B^0, B^{\pm}$ and B_s are displayed. The lines show the strangeness reduction cut, which rejects tracks near the jet axis with $d < 0.4 \sin\beta$.

$r\phi$-track fit are selected. This probability cut is a quality criterion for the hit assignment and is a method to reject bad or wrongly assigned tracks to hits, which will fake large impact parameters with small errors due to the good CST resolution. All these track criteria are recapitulated in the right column of the table 5.1. At the end the tracks are ordered in p_t. According to [1] the efficiency for a single charged particle with $p_T > 0.5\ GeV$ to produce a CST-improved track varies between 40% and 70% depending on the polar angle of the track. In figure 5.1 the CST-improved tracks selected for the 2-tracks method are compared to the MC simulation. The typical falling transverse momentum distribution (a) for charged tracks with $p_T > 0.5\ GeV$ is shown. The polar angle distribution (b) shows more tracks for low values, due to the different energies of the colliding particles. The hole in (c) is due to the inefficient CJC cells. The simulation gives a good description of these distributions.

5.1.2 Strangeness Rejection

The long lived strange particles such as K_s and Λ give long lifetime tails for signal and background. These tails are similar for signal and background and not useful information when one is using the fit procedure to disentangle signal and background. One way to reduce the contribution from strangeness is a cut in the impact parameter. The other way is to exploit the $d = c \sin\beta$ for decays near the jet axis, where β is defined as $\beta = \phi_{track} - \phi_{jet}$ and c is a constant. Only tracks with $d > 0.4 \sin\beta$ are selected for the secondary vertex fit. The figures 5.2 shows the simulated impact parameter versus the absolute value of the angle from the jet and track for (a) strange decay track like decays from K_s and Λ, (b) charm decay track like decays from $D^0, D^{\pm}$ and D_s and (c) beauty decay track like decays from $B^0, B^{\pm}$ and B_s. The track from decays of a strange hadron has in average larger impact parameter and simultaneous smaller aperture angle of jet and track than the heavy decay track. This is caused by the lower mass of the strange hadrons than the heavy hadrons. The heavy hadron decay tracks have larger momentum along and more transverse momentum with respect to the jet direction, resulting in a larger aperture angle of jet and track than the strange hadron decay tracks.

5.2 Primary Vertex

To measure the radial decay length not only a good knowledge of tracks is needed, but even a good understanding of the primary vertex is essential.

The tracks used in the secondary vertex (SV) fit are excluded from the primary vertex (PV) fit. The PV is determined through the routine CSPRIM, independent if the tracks are inside the jet or not.

The primary event vertex is reconstructed from charged tracks with the information of the average beam position. The mean position of the interaction region has movements of 0.1 cm around $(x_{beam}, y_{beam}) \approx (-0.2\ cm, 0.2\ cm)$ which are made to optimize the luminosity. The movement of the beam in the x direction in function of the run number is represented in figure 5.3(a) for data. Only the candidates selected for the 2-tracks method are presented. In (b) the number of tracks used for the PV xy-fit for data and simulation are compared. This is important because the PV resolution is a combination of spread of the beam interaction region and the PV tracks. The beam position is determined with good precision by CST tracks with high momenta of some hundred events and therefore the difference between run vertex and PV is rather independent of the beam movements. In the simulation the beam spot position is generated according to two Gaussian distributions with fixed average position ($x_{beam} \approx -0.2\ cm, y_{beam} \approx 0.2\ cm$) with the same widths as the data distribution. The movements of the beam are well corrected in data for the x and y direction as one sees in 5.3(c) and (d), where the difference of the PV calculated for each event in the 2-tracks method and the beam position is shown.

The measured transverse profile of the spread of the beam interaction region is measured to be in (e) $\sigma(x_{pv}) = 145\ \mu m$ and in (f) $\sigma(y_{pv}) = 30\ \mu m$. The maximal size of the errors are limited by the size of the elliptic beam spot. The error of the beam position is irrelevant. The peak at $145 \mu m$ represents the events with no track which are selected from CSPRIM for the PV. In NC events at least one track has to exist, the scattered positron, which comes from the PV. But in this analysis the track with the positron hypothesis is rejected for the primary event vertex calculation, to get the same behavior of the PV as in CC, where no positron exists.

In the variable vertex multiplicity (VVM) method the selection of tracks used for the PV are more stringent than in the fixed vertex multiplicity (FVM) method. The FVM method utilizes for the calculation of the PV the track selection criteria contained in CSPRIM[1], whereas the VVM method takes the track selection criteria listed in the left side of the

[1] The start radius R_{Start} and the impact parameter d are the same for CSPRIM and the PV track selection of this analysis. In addition CSPRIM selects tracks with track length $L_{CJC} > 10\ cm$ and a track impact significance $|S_d| \leq 2$, whereas the impact parameter and its significance is calculated with respect to the mean beam position.

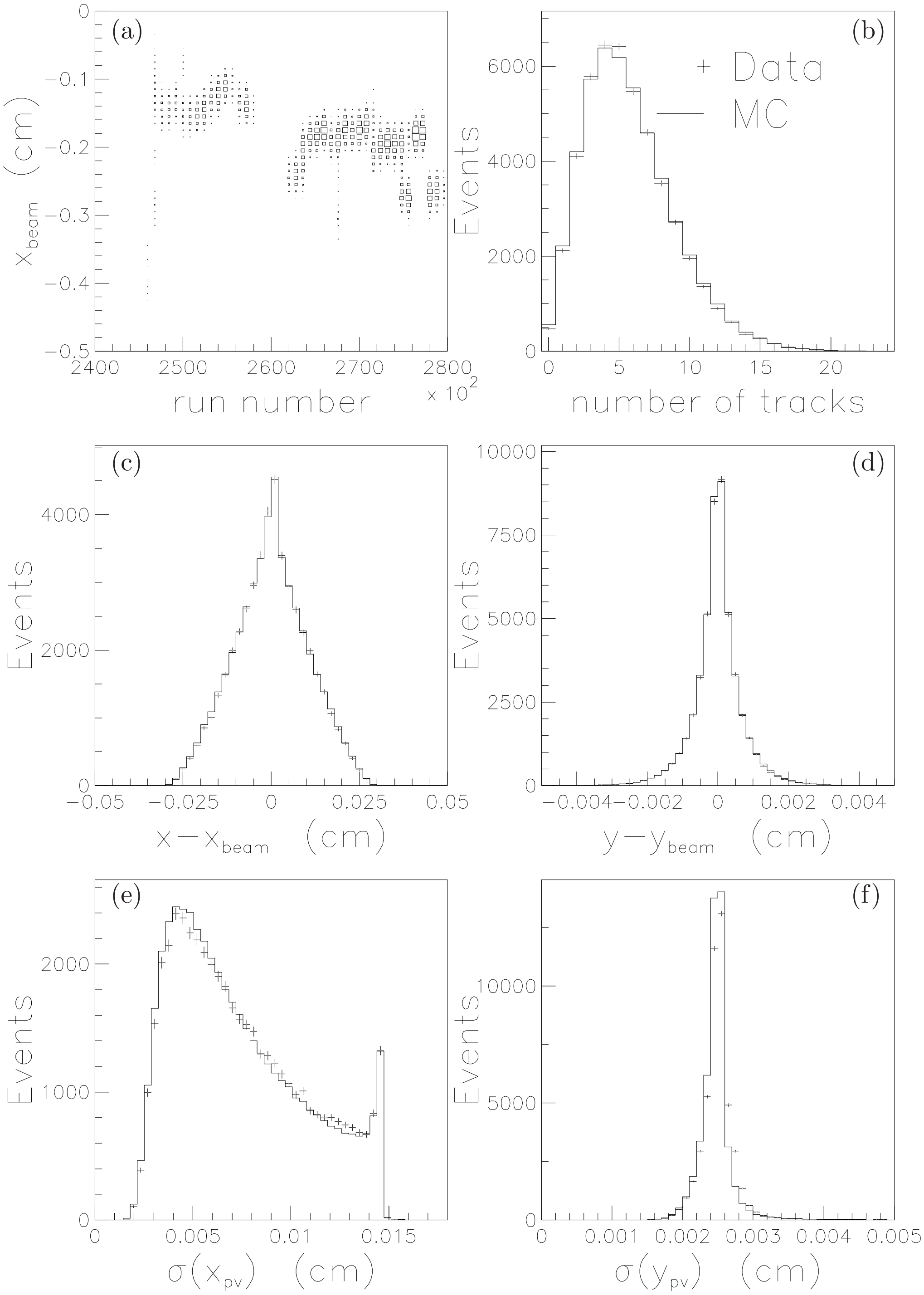


Figure 5.3: Different primary vertex informations for the selected candidates with the 2-tracks method are presented. Figure (a) shows for data only the movement of the beam in the x-direction as a function of the run number. (b) compares for simulation and data the number of tracks used for the two dimensional PV fit. In (c) and (d) the difference of the calculated PV and the position of the beam spot measured as the average over some hundred events is displayed. The calculated error distributions for the PV coordinates in the (e) x and the (f) y directions are compared for simulation and data.

Criteria		VMM PV	(VMM & FMM) SV
transverse momentum p_t	[MeV]	> 400	> 500
impact parameter $\|d\|$	[cm]	< 2	< 0.2
CST hit		≥ 1	≥ 2
CJC-CST track fit probability	[%]	≥ 10	≥ 10
track length L_{CJC}	[cm]	> 22	> 22
start radius R_{start}	[cm]	< 35	< 35
rapidity track $\|\eta\|$		-	< 1.5

Table 5.1: The track selection criteria utilized, for the variable vertex multiplicity (VVM) and the fixed vertex multiplicity (FVM) method, for primary or secondary vertex fit respectively. Additionally to the track criteria for secondary vertex (SV) fit the strangeness rejection cut is used.

table 5.1.

5.3 Jet Reconstruction

The jets serve to define the direction of flight of the light or heavy quark.

In this analysis only the jet axis is important and not the jet energy itself. Once a jet is defined, then all CST-improved tracks in the cone with radius one around the jet are collected, and afterwards a search of a secondary vertex with these track candidates starts. The jet is a physically motivated construct to reduce the number of tracks and consequently the combinatorics which can come from a decay.

Jets are reconstructed using the inclusive k_T algorithm with radius $R = 1$ in the $\eta\phi$ plane. The k_T algorithm is applied in the laboratory frame over all reconstructed HFS particles. HFS particles are reconstructed using a combination of tracks and calorimeter energy deposits. The P_T-recombination scheme is applied giving massless jets. The selection requires at least one jet with transverse momentum of $> 4\ GeV$.

Because of the high Q^2 defined, most events have a jet with $p_T > 4\ GeV$. In figure 5.4(a) only less than 1000 events have no jet. This means, that in about 98% of the events a jet has been found and the algorithm can start to search tracks and secondary vertices inside the jet cone. In (b) the number of CST-improved tracks in each jet which are in the angular range $10^o < \vartheta_{Jet} < 170^o$ for data and simulation is compared. The simulation gives a rather poor description of these distributions. Not so important for this analysis is the number of jets per event, but the number of CST-improved tracks per jet. It has an impact on the combinatorics when selecting two or three tracks for the 2- or 3-tracks method. One will see later that this incomplete description of the simulation seems to have negligible effects. Only jets with at least two CST-improved tracks for the 2-tracks method and at least three CST-improved tracks for the 3-tracks method are used. The figure 5.4 shows (c) the polar angle ϑ_{Jet} and (d) the transverse momentum $p_{T,Jet}$ distribution of the jets which contain at least two CST-improved tracks selected referring to equation 3.6. The jet with the candidate

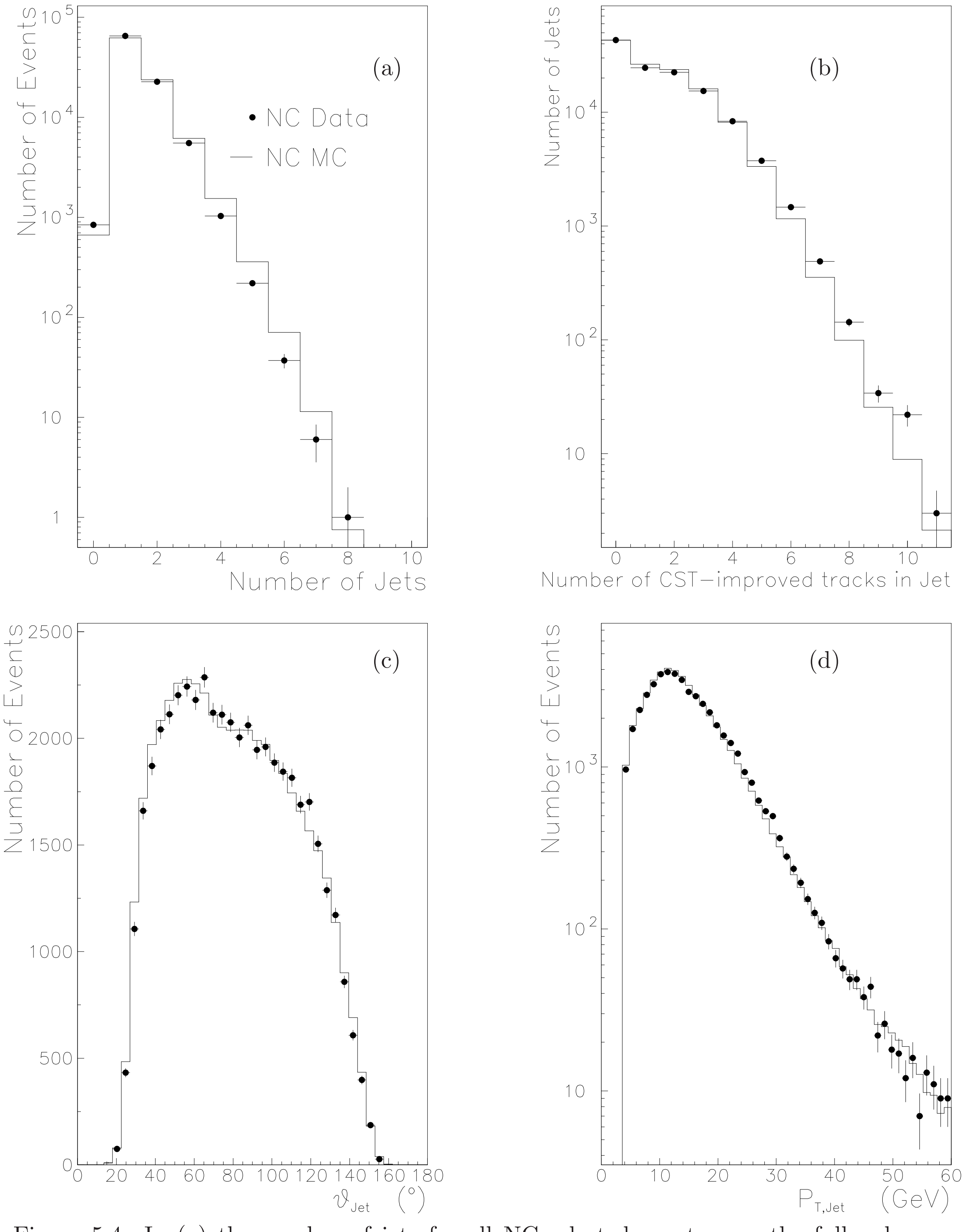


Figure 5.4: In (a) the number of jets for all NC selected events over the full polar range and $p_{T,Jet} > 4\ GeV$ are compared between data and simulation. In (b) the number of CST-improved tracks inside the jet in the angular range $10^o < \vartheta_{Jet} < 170^o$ are shown. In (c) the polar angle ϑ_{Jet} and in (d) the transverse momentum $p_{T,Jet}$ distribution of the jets which contain at least two CST-improved tracks within a cone of radius 1 in the $\eta - \phi$ plane are displayed.

as in the 2-tracks method selected is taken. The simulation describes the jet variable well if the jets are in the CST acceptance and one candidate per event is selected. A harder cut in the transverse momentum of the jet will not cause much loss in efficiency, because the $p_{T,Jet}$ spectrum increases until about 14 GeV.

Recapitulating: the CST-improved tracks, the primary vertex, the jet and the NC selection are well described by the simulation. In the next chapter the signal extraction is discussed.

Chapter 6

Charm and Beauty Production in Neutral Current Processes

In this chapter the results of the measurement of the charm and beauty production cross sections in neutral current processes (NC) are presented. The signal extraction technique is discussed. The analysis of the background and the extraction of the signal is developed using large samples of MC simulation events.

Two different approaches are taken to extract the signal. The fixed vertex multiplicity (FVM) method (cf. sec 3.5.1) and the variable vertex multiplicity (VVM) method (cf. sec. 3.5.2). Different ways to extract the signal are presented. Both methods use the information of the displaced secondary vertex. When testing systematic effects the FVM method gives unstable results, whereas the VVM method is robust. The VVM method is taken to extract the final heavy quarks content in the NC and CC selection.

The cross sections given in the next pages are determined for the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$. In the last section a description of the measurement uncertainties is given and a comparison with the NLO calculation.

6.1 Signal Extraction with the FVM Method

In the FVM method two or three CST-improved tracks inside a jet are fitted with a two dimensional unconstrained track fit (VFit2du) with two or three CST improved tracks to a common vertex. In the 3-tracks method only the candidates with a secondary vertex fit probability with more than 10% are selected. For each event only the candidate with the largest $min(|S_{d1}|, |S_{d2}|, |S_{d3}|)$ value is chosen. The motivation therefore and the detailed explanation of the FVM method are given in section 3.5.1.

In figure 6.1 the properties of the tagging tracks from the 3-tracks method between data and simulation are compared. Because the tracks are sorted in transverse momentum, the first row shows the fastest track of the three selected tracks, whereas for the following rows the transverse momentum decreases. Because of multiple scattering effects the implication is that the error of the impact parameter increases if the transverse momentum decreases. The hole around zero in the impact parameters comes from the selection of the largest

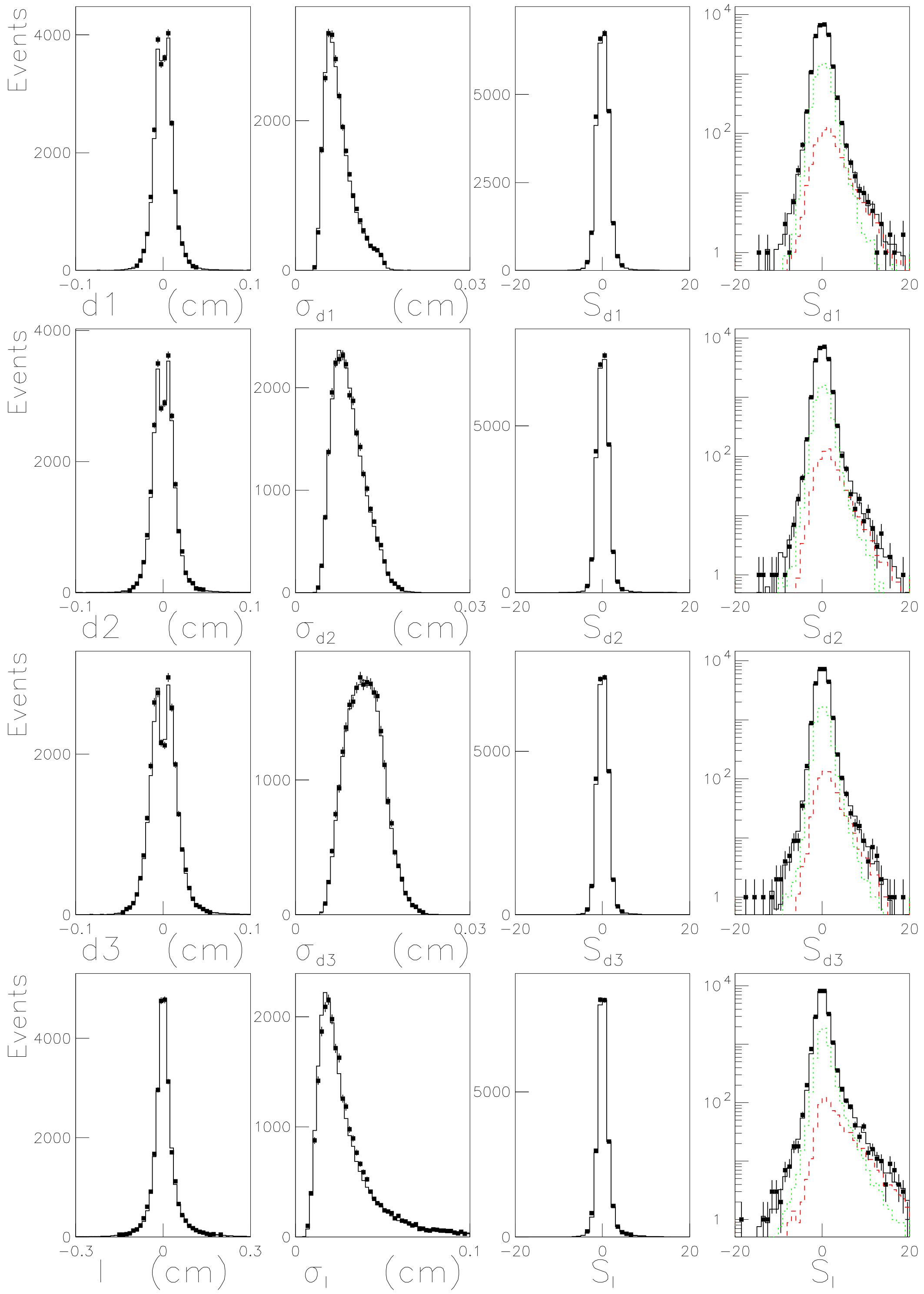


Figure 6.1: The selected events with the three tracks used for the 3-tracks method. The simulation (solid histogram) is normalized to data luminosity (points with statistical errors). In the first (second and third) row the impact parameter, the error of the impact parameter and its significance in linear and logarithmic scale of the first (second and third) track is displayed. The tracks are sorted in transverse momentum, where the fastest is upmost. In the last row the decay length, error of the decay length and decay length significance in linear and logarithmic scale is presented. In the last column the beauty (red, dashed) and charm (green, fine dashed) components as predicted by the simulation is shown. For high values of the significance the beauty contributions dominate.

$min(|S_{d1}|, |S_{d2}|, |S_{d3}|)$ value. The significances are displayed in linear and logarithmic scale. The linear scale shows the good description of the track resolution and the logarithmic scale demonstrates the good agreement of the tails of the distribution between the simulation and data.

Two different methods are used to extract the heavy quarks cross section: one is the *cut* based approach and the other is the *fit* procedure.

In the cut based approach, the idea is to cut on some variable increasing the signal over background ratio. The events which survive the cut are counted in the data and the cross section is calculated with the efficiency estimated in the MC simulation. This method is less dependent on the light quark content described by the simulation, since the selected cuts should throw away as many as possible from the light quark events and keep remaining signal events.

In the fit procedure (cf. sec. 3.6) the shapes of lifetime-related variables from the simulation are fitted to the data. If the shapes from light, charm and beauty components are too similar, the fit has problems to discriminate between the three components. Because the distribution of the observable and the efficiency is taken from the simulation, it is clear that this method is more dependent on the simulation than the cut based approach.

The observables used for the cut and the fit procedure are the decay length significance and the impact parameter significance.

6.1.1 Cut Based Heavy Quarks Cross Section with the FVM Method

The lifetime tag distinguishes signal and background events on the basis of their decay topology. The heavy hadrons produce secondary vertices separated from the ep interaction point of a few 100 μm, while the combinatorial background is symmetrically distributed around the ep interaction point, because of the track resolution. On the one hand cutting on lifetime related variables increases the signal over background ratio, but on the other hand makes the signal sample smaller. Thus choosing a large S_l region will allow the direct measurement of the heavy quarks production cross section in NC.

The idea used in the cut based approach is simple. First a cut is applied, increasing the signal over background ratio, and then one counts the events left. The goal is to get an enriched heavy hadron sample. As can be seen in the last column of figure 6.1, with a simultaneous cut in the variables S_{d1}, S_{d2}, S_{d3} and S_l an enriched sample of heavy quarks can be produced. The simultaneous cut in the impact and decay length significance is needed, because the three track vertices are sometimes formed with one primary track and two secondary tracks. The first bin in figure 6.2(a) shows the sum of the events in the 3-tracks method which fulfill the cuts $0 < S_m < 20$ with $m = d1, d2, d3, l$. The next bins give the sum of events when the lower bound of the cuts is increased. The x-coordinate of the figure 6.2(a) gives the value of the lower bound of the cuts. For the 2-tracks method the cuts for bin i in figure 6.2(c) are $i * 1. < S_m < 20$ with $m = d1, d2, l$ and $i = 0, 1...8$.

In figure 6.2(b) and (d) the calculated heavy quarks cross section is shown, which means the beauty and charm contributions are not distinguished. The cross section is calculated with the cuts explained above. The number of heavy quarks N_H is calculated for each cut from the difference of the data events N_{Data} and the uds-events $N_{MC,uds}$ as predicted by the simulation, $N_H = N_{Data} - N_{MC,uds}$. The efficiencies for charm and beauty together are taken from the MC simulation. The heavy quarks cross section σ_H is calculated as

$$\sigma_H = \frac{N_H}{\varepsilon_H \cdot \mathcal{L}}$$

in which the average efficiency $\varepsilon_H = \frac{N^H_{rec}}{N^H_{gen}}$ is defined in equation 3.2 and $\mathcal{L}$ denotes the luminosity.

The results indicated by the full circles have been obtained differently. The number of heavy quarks is now the difference between the data events in the positive side as illustrated in figure 6.2(a) and the sum of data events when all cuts used have a negative sign (e.g. $-20 < S_m < -i * 1.$). This means the events where the negative bins are subtracted from the positive $N'_H = N_{Data,positive} - N_{Data,negative}$ are counted. The idea is to use less information from the MC simulation and indeed with this method only the efficiency for heavy quarks calculated before is taken and no uds-events from MC simulation is any more needed. These two methods (N_H and N'_H) differ only for the first few bins and then give the same result. That is, the result is the same when the uds contribution is negligible. For the first few bins the method with the negative side subtraction gives a too low result because of the resolution effects of the heavy and light quark distributions, which means too many heavy quark events are subtracted for the first few bins. This could be corrected, but then more information from the simulation is needed. For high cut values the two ways give the same results, which means that the uds contribution behaves like the negative tails. This tells us that the uds contribution is less affected by the long lived strange particles like K_s and Λ, which will give a positive tail, than by random badly measured tracks, which are symmetric around zero.

The error bars shown are only statistical. An ad hoc systematic error like 10% uncertainty of the background will produce large error bars for the first few bins and then, because the background and statistics of the signal decrease, the statistical errors will dominate this ad hoc systematic error in the last bins. The heavy quarks cross section seems to be dependent on the cut values which shows that there exist systematics effects in this method that are not fully understood. However the heavy quarks cross section in NC is somewhere between $400 - 550\ pb$.

6.1.2 Charm and Beauty Cross Section with FVM 2- and 3-tracks Methods

A nice way to extract the beauty and charm cross section is to use simultaneously the 2- and 3-tracks methods and some information from the MC simulation. A cut value somewhere around two in figure 6.2(a) and (c) will give a rather clean heavy hadrons content for the 2- and 3-tracks sample. A harder cut will give on the one hand a cleaner heavy hadrons sample, but on the other hand the tagging efficiencies for charm and beauty will decrease and also the statistical errors will increase. Because the charm hadrons have a shorter lifetime than beauty hadrons the efficiency will decrease for charm events even faster than for beauty events. After taking a cut to enrich the heavy hadrons, the data events are subtracted from the corresponding uds-events simulated by the MC program for the 2- respectively 3-tracks method. The number of heavy quarks is then $N_i^H = N_i^{Data} - N_i^{MC,uds}$, where $i = 2, 3$ stands for 2- or 3-tracks method respectively.

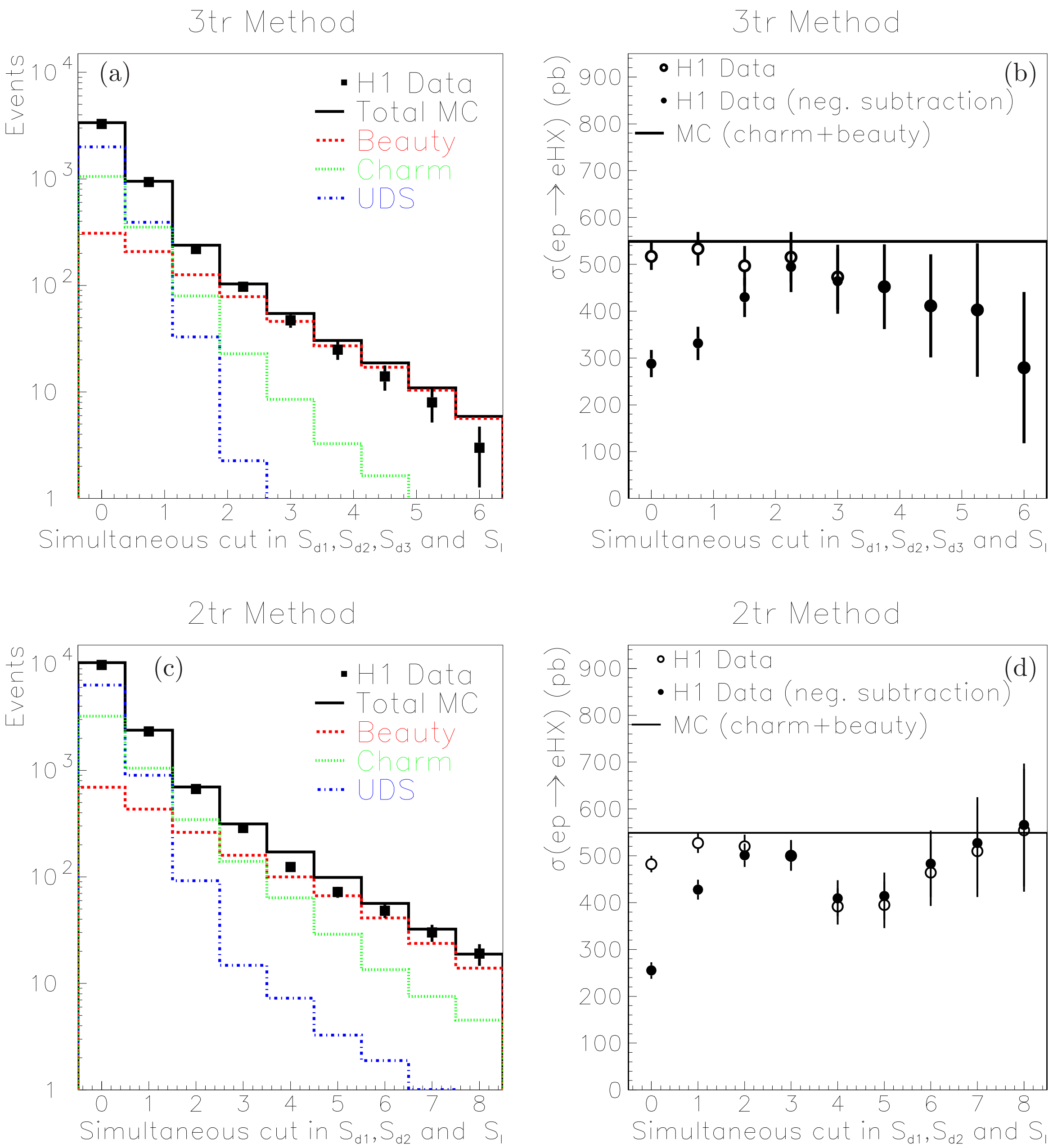


Figure 6.2: The composition of light, charm and beauty in the NC events selected with the 3- respectively 2-tracks method with a cut based approach explained in the text are plotted in (a) respectively (c). A simultaneous cut of 3 in the 3-tracks method will give a rather clean beauty sample of about 70 events (60 beauty, 10 charm and 0 uds, as predicted by the Django MC program). In figures on the right (b) and (d) the cross section for heavy quarks (charm and beauty together) $\sigma(ep \rightarrow e'HX)$ is calculated. The line is the expected cross section predicted from the Django MC program. The empty circles gives the dependence of the measured heavy quarks cross section in function of the cut values. A slightly different calculation is used for the full circles, which is explained in the text. Only statistical errors are shown for the data points.

The charm and beauty cross sections are then calculated with the equations:

$$N_2^H = (\epsilon_{2b} \cdot \sigma_b + \epsilon_{2c} \cdot \sigma_c) \cdot \mathcal{L} \quad (6.1)$$
$$N_3^H = (\epsilon_{3b} \cdot \sigma_b + \epsilon_{3c} \cdot \sigma_c) \cdot \mathcal{L} \quad (6.2)$$

where the unknowns are the cross sections for charm and beauty (σ_c, σ_b). The efficiencies are taken from the simulation with the same definition as given in equation 3.2 for charm and beauty for the 2- ($\epsilon_{2c}, \epsilon_{2b}$) respectively 3-tracks ($\epsilon_{3c}, \epsilon_{3b}$) method. $\mathcal{L}$ is the luminosity. All this is only correct if the 2- and 3-tracks sample are independent, otherwise the equations are linearly dependent. The result is only reasonable, where the heavy hadrons content is large enough and is not dominated by the uds-events. Changing the cut from the second bin until the fourth bin in figure 6.2(a) and (c) gives results around $\sigma_c \approx (500 \pm 100)$ *pb* and $\sigma_b \approx (50 \pm 10)$ *pb*. These errors mean only the fluctuations resulting from the different cuts. A harder cut gives significantly different results for the charm cross section. The explanation is that in the 3-tracks method the hard cut gives a very low charm efficiency and so the statistical errors are dominant.

6.1.3 Charm and Beauty Cross Sections with a Fit Procedure in the FVM Method

The measurement of the cross section is based on a observable by which it is possible to distinguish the signal from the background. Such a distinction is not performed on an event by event basis but it consists in the study of the observable spectrum over a large sample of events. The relative fraction of signal events is estimated adjusting the proportion of all simulated components until the best description of the measured observable is achieved. As a consequence, the method relies on the correct simulation of the observable distribution.

The observable used are S_l or $min(S_{di})$. The measured S_l or $min(S_{di})$ distributions from the data are fitted with the shapes of light, charm and beauty contributions, which are obtained from the Django MC simulation. The simulation templates used for the 3-tracks method of the S_l distributions are presented in figure 3.3(b). The relative fractions of all three components are adjusted such that the likelihood is maximized. The normalization of the summed contributions is adjusted to match the data. The fit results are used to extract the number of events attributed to light, charm and beauty hadron decays and allows to determinate the cross sections (cf. sec. 3.2).

The three parameters fit with the advanced binned maximum likelihood method (cf. sec. 3.6) for figure 6.3(a) obtained with the 3-tracks method yields a sample composition of $f_{uds} = 73.5 \pm 1.9\%$, $f_c = 23.8 \pm 2.2\%$ and $f_b = 2.6 \pm 0.4\%$ with a $\chi^2/n.d.o.f = 35.8/20$. The statistical errors of the fit include contributions from the limited statistics of the data and simulated samples representing the different components. The cross sections displayed in the figures 6.3 are calculated using the fit results with the equation 3.1. The correlation coefficients are $\rho_{b,c} = -0.724$, $\rho_{b,uds} = 0.605$ and $\rho_{c,uds} = -0.935$. With a Kolmogorov test the similarity between the uds, charm and beauty distributions is checked. The Kolmogorov test of this distributions gives the probability of less than 10^{-5}, which means that the distributions are very unlikely that they arose from the same parent distribution.

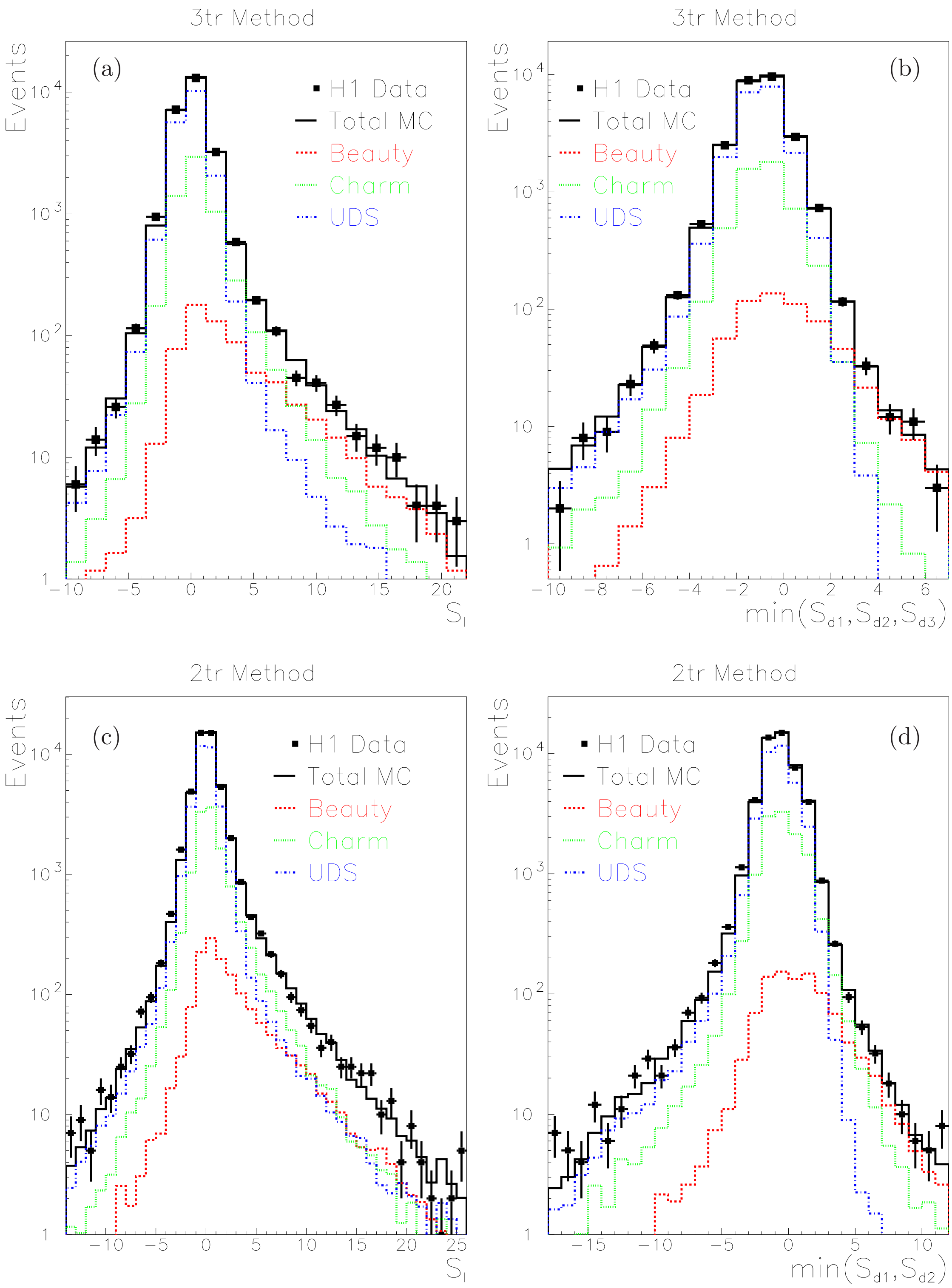


Figure 6.3: Measured decay length significance (left) and the minimal impact parameter significance (right) for the data and Django MC simulation with the 3-tracks (upper) and 2-tracks (lower) method. The fitted contribution from beauty, charm and light hadrons are displayed separately. The solid line corresponds to the sum of all MC program contributions. The data points are plotted with the statistical uncertainties only. Also included is the result from the fit to the data of the simulation distributions extrapolated to the defined cross section. A clear beauty enriched sample is visible at high positive values for the 3-tracks method in figures (a) and (b). Unacceptable for the 2-tracks method is the worse χ^2 resulting from the fit. Nevertheless even in figure (d) a clear enrichment of beauty at high values is manifested. In figure (c) no clear separation between the three components is seen.

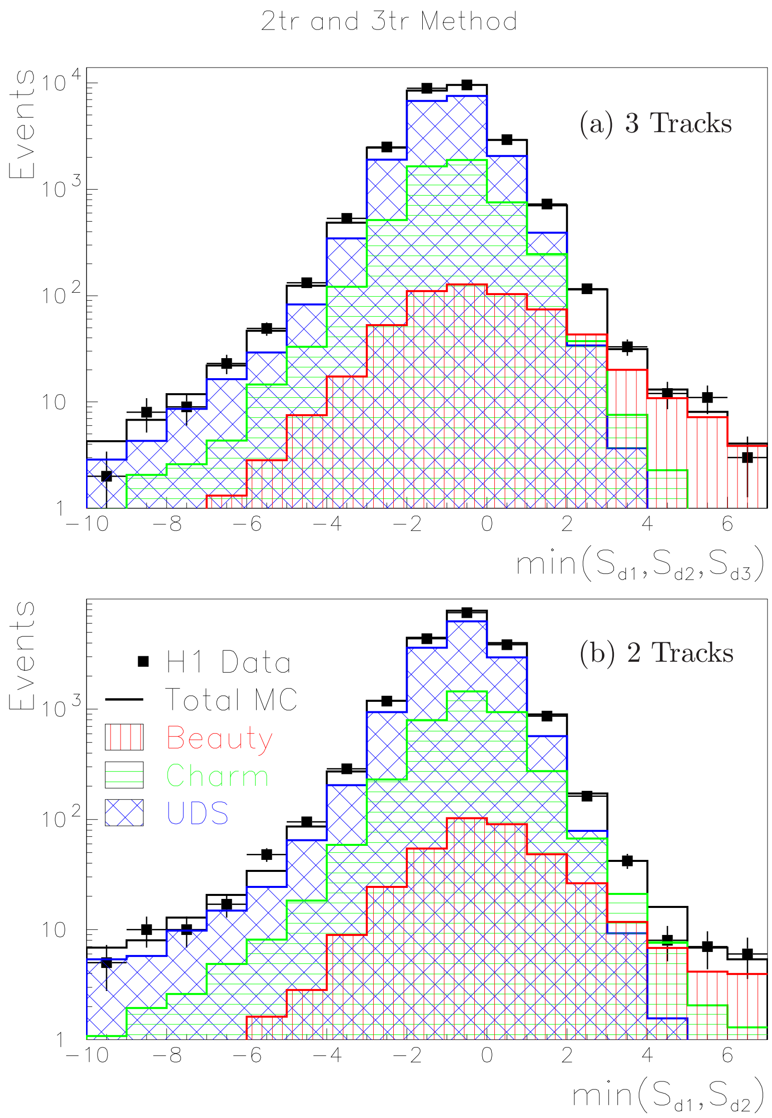


Figure 6.4: A simultaneous fit of the combination from the 2- and 3-tracks method, where no overlap of events is possible. The figure (a) is the same as figure 6.3(b). For figure (b) only events with exact two CST-improved tracks inside a jet are selected for the 2-tracks method. It is also a subsample of the figure 6.3(d). The result is consistent with the other results and is given in table 6.1(e). Even a fit of only figure (b), where events with exactly two CST-improved measured tracks per event are selected, gives reasonable results.

A slightly different observable is used as a cross check, namely the minimal impact parameter significance of the selected candidate, which is shown in figure 6.3(b). The minimal impact parameter significance gives similar results as the variable S_l and these results are summarized in table 6.1(a).

In the 3-tracks method around 25% of the selected NC events are used to extract the

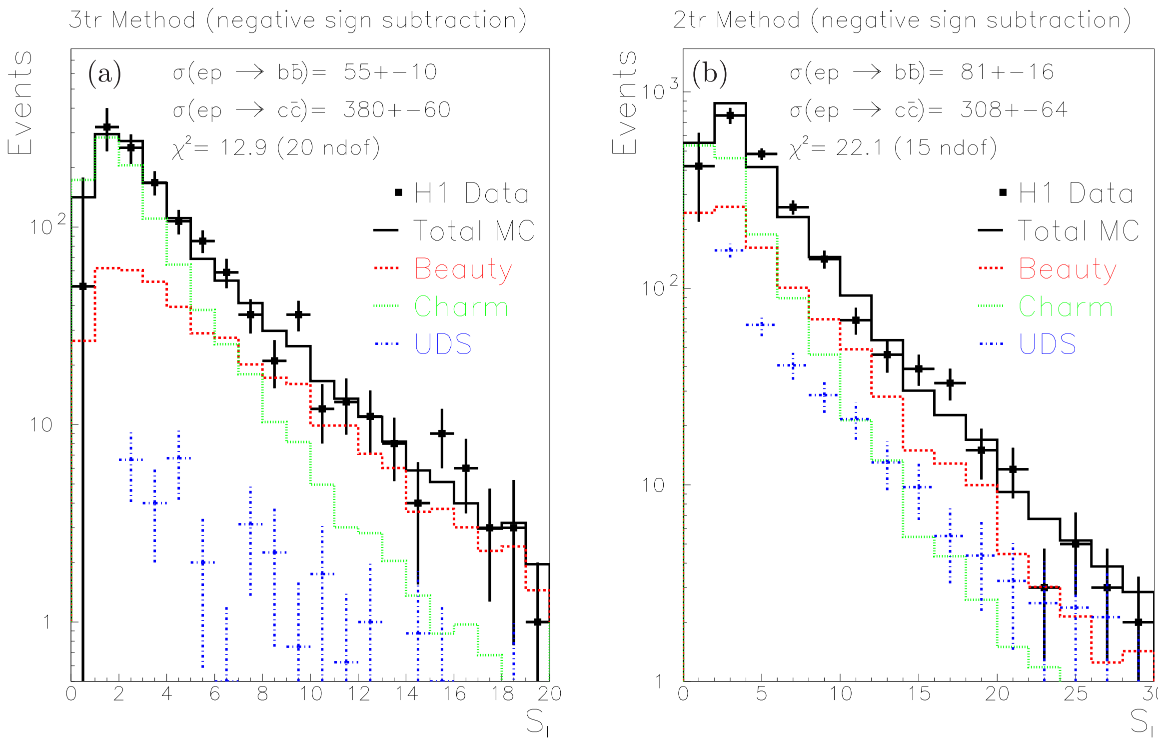


Figure 6.5: Fit of the decay length significance after negative sign subtraction: The S_l distribution with the 3-tracks method in (a) and with the 2-tracks method in (b), after subtracting the bins with equal magnitude but negative sign, from the positive. The light quark distribution is fixed and only the result from the fit for charm and beauty are shown.

cross section and one question that one should pose is: Have events without a secondary vertex hypothesis, like events with less than three CST-improved tracks inside a jet, a bias for heavy quark events? If the multiplicity of the charged tracks for light and heavy quark events is not well simulated by the MC program, then the extrapolation for the cross section could be wrong. That can be checked with a fit with all NC events. Namely all NC events except the NC events with a secondary vertex candidate in the 3-tracks method are put in one bin, the so called normalization bin. Then a simultaneous fit of the normalization bin and the distribution of secondary vertex candidates used above in figure 6.3(a) or (b) is done. The results are compatible with the results without the normalization bin and are listed in table 6.1(b).

The normalization bin works well, but what happens if the sample is augmented with events containing only two CST-improved tracks inside a jet? The 2-tracks sample has around two times more events than the 3-tracks sample. The results of the fit from the 2-tracks method for the decay length significance and minimal impact parameter significance are shown in figure 6.3(c) and (d). For example, the fractions yield from the minimal impact parameter significance are $f_{uds} = 72.6 \pm 2.0\%$, $f_c = 25.2 \pm 2.3\%$ and $f_b = 2.2 \pm 0.4\%$, but the χ^2 is bad. The results are consistent with the 3-tracks method. The decay length

	FVM Method	$\sigma(ep \to bb)$	$\sigma(ep \to c\bar{c})$	$\sigma(ep \to uds)$	$\chi^2/ndof$	$\rho_{b,c}$	$\rho_{b,uds}$	$\rho_{c,uds}$
	3-tracks							
(a)	S_l	50 ± 8	462 ± 42	1788 ± 47	35/20	-0.72	0.61	-0.94
	$min(S_{d1}, S_{d2}, S_{d3})$	46 ± 7	382 ± 55	1892 ± 64	16/17	-0.57	0.46	-0.96
(b)	norm. & S_l	44 ± 7	498 ± 32	1727 ± 30	38/21	-0.65	0.48	-0.95
	norm. & $min(S_{d1}, S_{d2}, S_{d3})$	39 ± 6	486 ± 35	1746 ± 35	20/18	-0.45	0.27	-0.96
	2-tracks							
(c)	S_l	58 ± 10	486 ± 38	1684 ± 38	108/40	-0.82	0.68	-0.94
	$min(S_{d1}, S_{d2})$	40 ± 6	469 ± 43	1730 ± 48	77/30	-0.76	0.68	-0.97
(d)	norm. & S_l	85 ± 10	366 ± 31	1828 ± 26	129/41	-0.80	0.60	-0.93
	norm. & $min(S_{d1}, S_{d2})$	49 ± 5	389 ± 26	1841 ± 25	89/31	-0.64	0.48	-0.94
	Combined 2tr & 3tr							
(e)	$min(S_{d1}, S_{d2}, (S_{d3}))$	44 ± 6	391 ± 40	1825 ± 46	73/34	-0.59	0.47	-0.96
	subsample 2 tracks	43 ± 20	345 ± 90	1800 ± 91	22/17	-0.85	0.77	-0.98
	S_l	50 ± 7	470 ± 32	1719 ± 35	140/39	-0.74	0.60	-0.93
	subsample 2 tracks	72 ± 18	403 ± 60	1701 ± 56	70/19	-0.83	0.69	-0.94
	Negative subtraction							
(f)	3-tracks S_l	55 ± 10	380 ± 60		13/20			
	2-tracks S_l	81 ± 16	308 ± 64		22/15			

Table 6.1: The integrated light, charm and beauty cross sections given in pico barn for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ extracted with the different FVM methods. Only statistical errors are given. Different results found when different variables and combinations of histograms are fitted. The different methods are explained in the text. The 3-tracks method gives for all ways consistent results, whereas the 2-tracks method gives great variations of the results with bad χ^2.

significance of the 2-tracks method has less separation power between light and heavy quark as in the 3-tracks method (compare figure 6.3(a) and (c)). After all the components are different enough to use the fitting procedure. The Kolmogorov test gives a probability of less than 10^{-5} between these distributions. The worse separation between heavy and light quark distributions is caused firstly because the K_s are favored to be reconstructed in the 2-tracks method than in the 3-tracks method. Secondly, for light quarks the chance of three tracks which intersect in one point with a high significance due to the resolution effects is smaller than for two tracks.

The fit of the variable S_l with the normalization bin in the 2-tracks method gives significantly different result for the beauty cross section, caused by the worse separation between light, charm and beauty components than in the 3-tracks method. The results with the normalization bin for the 2-tracks method are listed in table 6.1(d).

The 2-tracks method is not fully independent from the 3-tracks method. The two CST-improved tracks per event of course are not used in the 3-tracks method and are surely independent. But for the three and more CST-improved tracks per event it could be that the tracks selected for the secondary vertex fit in the 2- or 3-tracks method are the same. The simultaneous fit of the 2- and 3-tracks method together, is surely independent if in the 2-tracks method only events with two CST-improved tracks per event are taken. The minimal impact significance distribution for the disjoint 2- and 3-tracks method is presented in figure 6.4. Again the results are consistent and are given in table 6.1. Even a fit of the two CST-improved track per event subsample from the 2-tracks method alone gives reasonable results, listed as subsample 2 tracks in table 6.1(e); for the sake of completeness the results of S_l are added in the table.

One method to reduce the sensitivity to the resolution effects and light quark contribution is the negative bins subtraction in the S_l distributions from the positive bins. This method can not be applied for the $min(S_{di})$ distribution, because it is not symmetric.

In figure 6.5 the negative bins subtraction method is applied for the 2- and 3-tracks method, where the results of the fit are displayed. In contrast to the advanced binned maximum likelihood (Barlow) fits used above, a standard least squares fit is used. Only two free fit parameters for charm and beauty are taken, whereas the light quark distribution is taken from the simulation and is a fixed parameter. The least squares fit has not the problem, in contrast to the Barlow fit, with the light quark histogram which has negative number of entries for the first two bins. This is caused by the selection of the largest $min(S_{di})$ and the succeeding negative bins subtractions. The fit gives acceptable χ^2. Consistent results are also found in the 3-tracks method when a fit without any contribution of the light quark or when all three parameters are free. This shows that the light quarks contribution is negligible for the 3-tracks method. For the 2-tracks method a large fixed light contribution has to be included and the results are not completely satisfactory. Again the figure 6.5 (b) shows a worse separation between the different components than figure 6.5(a). The results are summarized in table 6.1(f).

6.1.4 Problems with the FVM Method

The fixed vertex multiplicity method fails in three crucial tests, two of them appear when testing the systematic effects explained in section 6.2.2. Firstly the 2-tracks method has an unacceptable χ^2. Secondly because the beam spot is elliptical, it is a good test to split the selected events into a horizontal and a vertical sample. The charm and beauty distribution is very sensitive to this division. The charm (beauty) cross section variation of this test is about 20 % (30 %). Thirdly the track parameters and errors in the MC program can be tuned to data by two fudge factors, one is the resolution and one the multiple scattering term. Two different sets of these fudge factors are used. The motivation is that the default fudge factors are not consistent with the fudge factors determined for this analysis. Taking the difference between the two sets of fudge factors (1.35(resolution)/1.22(multiple scattering) $\rightarrow$ 1.35/1.09) leads to large systematic errors. Because of this effect one gets for the 3-tracks (2-tracks) method a 36 % (40 %) variation of the cross section for charm and a 18 % (30 %) variation of the cross section for beauty. The charm cross section is even more sensitive to the resolution than the beauty cross section. This is because the shape of the reconstructed fit variable is dependent on the resolution and a large fraction of charm decays have a decay length comparable to the resolution of the detector.

One of the problems is that in the 3-tracks sample many candidates of the charm events have one primary track. This track has with great probability the smallest impact parameter. In the minimum method only this track is considered and so a great part of the charm signal looks like the uds contribution. The lifetime part with three real charm tracks is relatively small and looks like the beauty events. This is because many of the charm decays are D^+,

FVM Method	$\sigma(ep \to b\bar{b})$ $[pb]$	$\sigma(ep \to c\bar{c})$ $[pb]$
3-tracks		
S_l	$50 \pm 8 \pm 15$	$462 \pm 42 \pm 162$
$min(S_{d1}, S_{d2}, S_{d3})$	$46 \pm 7 \pm 9$	$382 \pm 55 \pm 153$
2-tracks		
S_l	$61 \pm 10 \pm 20$	$470 \pm 38 \pm 197$
$min(S_{d1}, S_{d2})$	$40 \pm 6 \pm 9$	$469 \pm 43 \pm 188$

Table 6.2: The integrated beauty and charm cross sections with systematic and statistical errors for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ extracted with the FVM methods. The systematic error is dominated by the track resolution.

which often decay in three charged tracks, but the D^+ has a decay length like beauty.

The systematic errors for the variable S_l and $min(S_{di})$ used in the 2-tracks respectively 3-tracks method are listed in the tables A.1, A.2, A.3 and A.4 in the appendix A, where the errors are dominated by the track resolution. The result is that the FVM method is very sensitive to small variations of the input variables (mainly due to the track resolution), which gives big variations on the beauty cross section and huge variation on the charm cross section. The results of the systematic and statistical errors are summarized in table 6.2.

The next section shows that the variable vertex multiplicity method is less sensitive to these problems and is less affected from the systematic studies.

6.2 Signal Extraction with the VVM Method

In the VVM method the secondary vertex is reconstructed explicitly like in the FVM but with variable track number. The use of a variable track multiplicity gives a better separation of the different track multiplicity of charm and beauty decays. The algorithm of the adaptive fitter starts only if there exist at least two CST improved tracks inside a jet. That is equivalent to the sample of the 2-tracks method. Now the adaptive fitter has always a secondary vertex but in the end it will not assign any tracks to the secondary vertex. This because, after the last annealing step, the vertex multiplicity is determined by tracks contributing with a weight of greater than 80% to the secondary vertex. That is the big difference to the FVM method! The selected candidates in the VVM method have really tracks which are not compatible to the primary vertex, whereas in the FVM method most selected candidates have tracks which are compatible with the primary vertex. The detailed explanation of the VVM method is given in section 3.5.2.

In figure 6.6 the track multiplicity of the secondary vertices found by the adaptive fitter are shown. All events in the bin -2 or -1 respectively are events with zero or one CST improved track inside a jet, where the adaptive fitter is not used. As mentioned above the adaptive fitter starts only for events with at least two CST improved tracks inside a jet. In most of these events the adaptive fitter does not assign any tracks to the secondary vertex, which are put in the zero bin. About 10% of the NC events have one track assigned to the

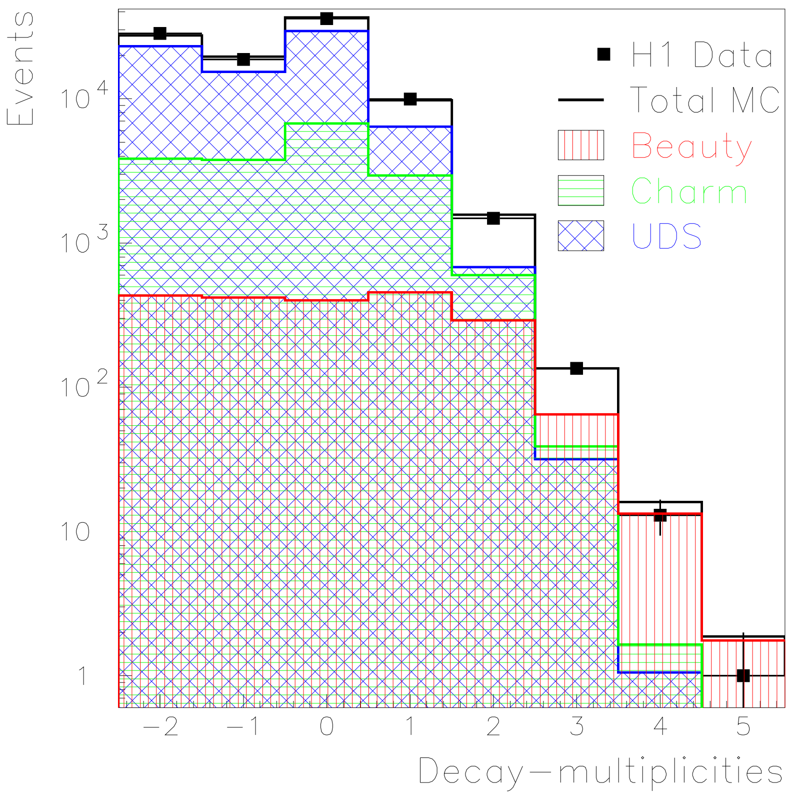


Figure 6.6: The number of tracks contributing with a weight greater than 80 % to the secondary vertex, is used as a measure of the decay-multiplicity. Beauty dominates ≥ 3 track vertices. The adaptive fitter is not used for events in the negative bins, see text.

secondary vertex and in only about 1.5% (0.15%) of the NC events the adaptive fitter ends up with two (three and more) tracks assigned to the secondary vertex.

Some events have in different jets two candidates with at least one track compatible with a secondary vertex. Only one candidate per event is considered, namely the candidate with the higher track multiplicity and if this is not sufficient the candidate with the higher decay length significance.

6.2.1 Charm and Beauty Cross Sections with a Fit Procedure in the VVM Method

The charm and beauty cross sections are extracted again with the fit procedure. Different variables and different combinations of the decay-multiplicity are tested. The variables are again the S_l and $min(S_{di})$. The variables are plotted in different histograms depending if the selected candidate has less than 1, 1, 2, 3 or 4 tracks which are significantly compatible with the secondary vertex. The histograms with the same variable but with different multiplicities are fitted simultaneously using MC program templates for the corresponding distribution of

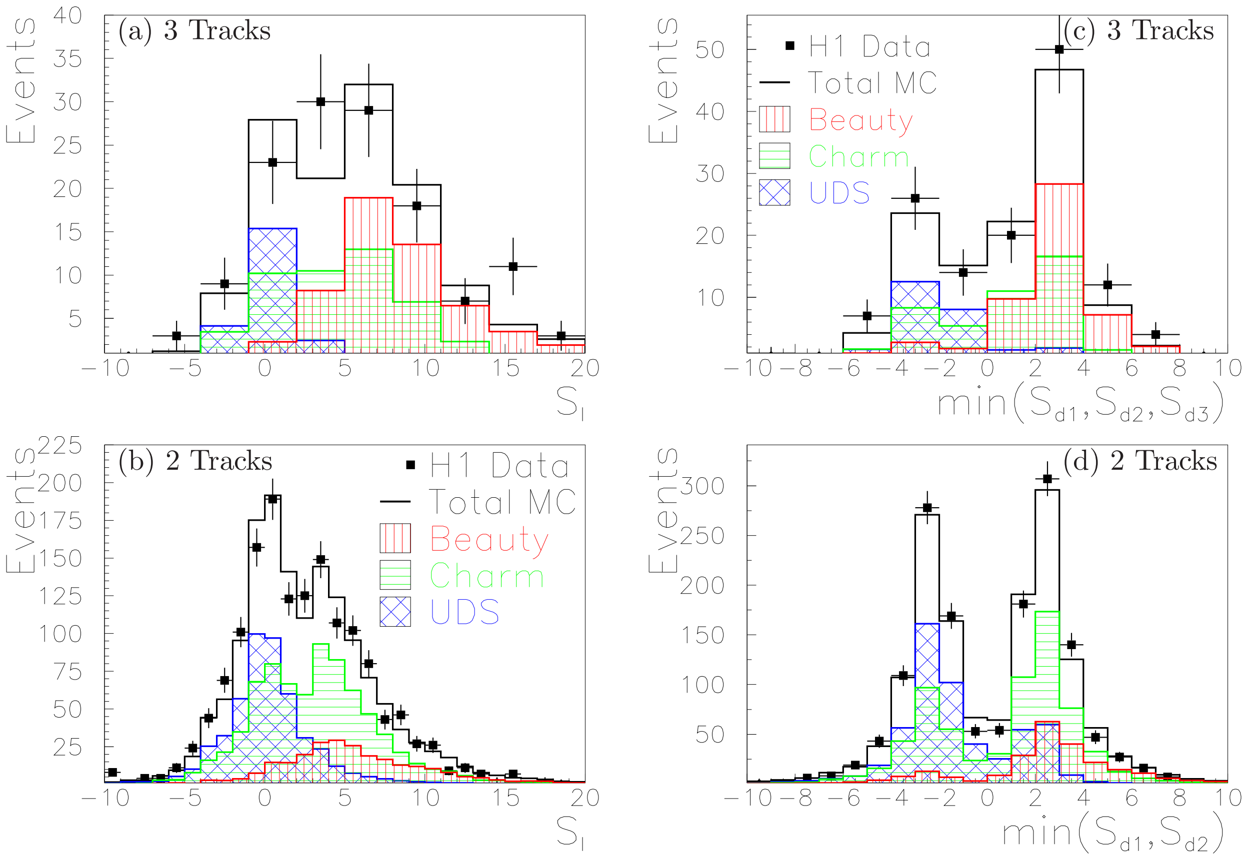


Figure 6.7: The VVM method for 2 & $\geq$ 3 tracks: The figure shows the variable S_l (left) and $min(S_{di})$ (right) for the same events with $\geq$ 3 tracks (a, c) and 2 tracks (b, d) associated with the secondary vertex. The simulation shows a good description of the data. Superimposed on the data points with statistical error only are the light, charm and beauty quark contributions as obtained from the fit with the Django MC program templates.

light, charm and beauty quark respectively.

In the two histograms on the left (right) side of the figure 6.7 the variable S_l ($min(S_{di})$) of the 2 tracks candidates are put in one histogram and $\geq$ 3 tracks candidates in the other, then the simulated shapes are fitted simultaneously to both data histograms respectively

For the 2 tracks decay-multiplicity in figure 6.7(b) charm and beauty quarks are well separated from the light quarks. The light quarks are peaked around zero, whereas the heavy quarks are shifted to higher values. In figure 6.7(a) where $\geq$ 3 tracks contribute to the secondary vertex, the beauty quarks are well separated.

For the same events, the variable $min(S_{di})$ is plotted in figure 6.7(c) and (d). The hole near zero in (d) comes from the cut that the tracks have to contribute at least with 80 % to the secondary vertex, which removes tracks which are compatible with the primary vertex. For the 2 tracks sample most light quark events have one track with negative impact

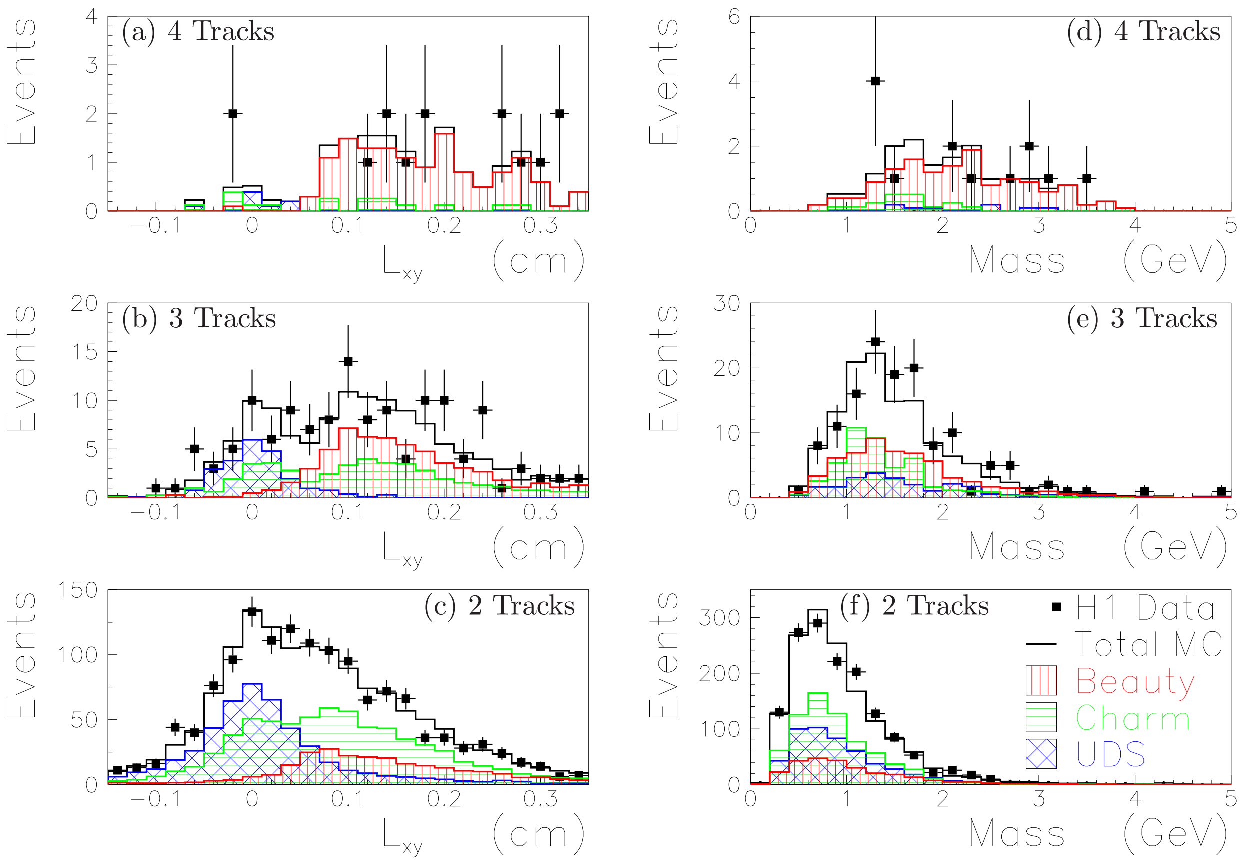


Figure 6.8: The VVM method for 2 & 3 & 4 tracks: On the left side the radial decay length distribution (L_{xy}) for events with (a) 4, (b) 3 and (c) 2 tracks associated with a weight greater than 80% to the secondary vertex is shown. The difference between (c) and the figure 6.7(b) is that the decay length error calculation is not used here. On the right side for the same events the invariant mass is calculated. The charm masses are below 2 GeV whereas the beauty masses extend to higher values. Superimposed on the data points are charm, beauty and light quark contributions as obtained from a simultaneous fit of the three histograms with Django MC program templates of the (L_{xy}) variable.

parameter, whereas the heavy quark events have both tracks with positive impact parameter. The simulation gives a good description of the data for both variables.

The fit method to extract the quark content is the same as used in figure 6.4. The difference to figure 6.4 is that in the VVM method the tracks are really compatible with the secondary vertex, whereas in the FVM method most of the selected candidates (e.g. from the light quark events) have tracks from the primary vertex.

The three parameters fitted simultaneously to the two histograms in figure 6.7(a) and (b) yield a sample composition of $f_{uds} = 31.3 \pm 3.4\%$, $f_c = 49.5 \pm 5.9\%$ and $f_b = 19.1 \pm 3.4\%$ with

	VVM Method	$\sigma(ep \to bb)$	$\sigma(ep \to c\bar{c})$	$\sigma(ep \to uds)$	$\chi^2/ndof$	$\rho_{b,c}$	$\rho_{b,uds}$	$\rho_{c,uds}$
	Combined 2tr & $\geq$ 3tr							
(a)	S_l	41 ± 7	510 ± 60	1316 ± 144	46/37	-0.79	0.47	-0.75
	$min(S_{di})$	38 ± 6	476 ± 60	1439 ± 157	29/29	-0.77	0.47	-0.78
(b)	norm. & S_l	47 ± 6	404 ± 35	1828 ± 34	64/38	-0.7	0.58	-0.97
	norm. & $min(S_{di})$	42 ± 5	395 ± 34	1842 ± 33	36/30	-0.67	0.56	-0.96
	Combined 2tr & 3tr & 4tr							
(c)	L_{xy}	42 ± 8	474 ± 67	1386 ± 150	46/57	-0.83	0.51	-0.75
	S_l	39 ± 7	522 ± 60	1294 ± 144	49/44	-0.79	0.47	-0.75
	$min(S_{di})$	37 ± 6	489 ± 59	1465 ± 154	31/33	-0.75	0.45	-0.77
	Combined 1tr & 2tr & $\geq$3tr							
(d)	S_l	47 ± 6	435 ± 32	1785 ± 58	174/66	-0.73	0.49	-0.85
(e)	norm. & S_l	47 ± 5	429 ± 23	1800 ± 22	175/67	-0.7	0.53	-0.92
	Combined 1tr & $\geq$ 2tr							
(f)	S_l	47 ± 9	432 ± 39	1790 ± 63	163/57	-0.87	0.6	-0.83

Table 6.3: The integrated light, charm and beauty cross sections given in pico barn for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ extracted with the variable vertex multiplicity methods. Only statistical errors are given. Different combinations of decay-multiplicities are fitted simultaneously (see text).

a $\chi^2/n.d.o.f = 45.9/37$. Again the statistical errors of the fit include contributions from the limited statistics of the data and simulated samples representing the different components. The correlation coefficient are $\rho_{b,c} = -0.79$, $\rho_{b,uds} = 0.47$ and $\rho_{c,uds} = -0.75$. Similar results are obtained with the variable $min(S_{di})$ and the results are summarized in table 6.3(a), where the cross sections are calculated with the equation 3.1.

In the VVM method the statistics of the fitted histogram is lower than in the FVM method (up to a factor 30!), but with a better signal sensitivity the loss of efficiency is compensated and similar statistical errors as in the FVM method are achieved.

Again, as in the FVM method, the normalization bin is used as a cross check. The normalization bin contains all NC events except the ones with ≥ 2 tracks associated with the secondary vertex. Then a simultaneous fit of the normalization bin and the distributions of histogram in figure 6.7(a) and (b) respectively (c) and (d) is done. The results are compatible with the results without the normalization bin and are listed in table 6.3(b).

If the uds contribution is correctly described in the simulation, but because of the small efficiency it has a big statistical error, then the normalization bin can improve the uds result considerably. One sees this comparing the results in table 6.3(a) and (b). If the result in table 6.3(a) has a significant wrong uds quark content, then a fit with normalization bin will falsify the charm and beauty contents. Therefore the result without normalization bin is preferable. Indeed the question remains, why one has to believe into the charm and beauty contents and accept that the uds content is significantly wrong. One argument is that for the heavy hadrons the real lifetime information is used, whereas for the light hadrons one looks only into the tails of the distributions.

A further test is to fit the radial decay length L_{xy} proper. The fit results are similar as for the decay length significance. On the left side of the figure 6.8 the radial decay length for events with 2, 3 and 4 tracks associated with the secondary vertex are presented. The vertices found in the light quark events peak at decay length zero, while the vertices found in heavy quark events are significantly displaced to positive values. Both charm and beauty decays contribute to the 2 tracks vertices, whereas beauty dominates the 3 and 4 tracks vertices. The right side of the figure 6.8 shows the calculated invariant mass of the same events and same tracks used on the left side. For each track associated with the secondary vertex a mass assumption of a pion is used. These are only control plots, because the fit do not converge for these distributions. The charm masses are below 2 GeV, whereas the beauty masses extend to higher values. The results are in table 6.3(c). No difference for the fit results exist if the histogram with $\geq$ 3 tracks vertices is split into two histograms of 3 and 4 tracks vertices, which is listed in table 6.3(c).

It is possible to include vertices with 1 track to the 2 tracks and $\geq$ 3 tracks vertices. Again a simultaneous fit of these three histograms gives consistent results given in table 6.3(d), but now the χ^2 is bad. The normalization bin (this means all NC events except the ones with $\geq$ 1 track vertices) and these three histograms give together the results shown in table 6.3(e).

In the next chapter for charm in CC only a simultaneous fit of 1 track and 2 tracks is possible, because of the low statistics. The results are again consistent and given in table 6.3(f), but the $\chi^2/ndof$ is high as soon as the 1 track events are taken.

6.2.2 Experimental Systematic Errors

Various sources contribute to the total uncertainty of the measured heavy quark cross sections. In particular, systematic errors are introduced by the selection procedure, by the fit technique and by the uncertainty of the detector calibration. The other error is the statistical uncertainty. The total error results from the quadratic sum of the systematic and statistical errors.

The sources of the systematic uncertainty tested for this analysis are as follows:

- The track parameters and errors in the simulation can be tuned to data by two fudge factors, one is the resolution and one the multiple scattering term. The uncertainty on the track resolution is tested by changing the multiple scattering term from the actual value 1.09 of the H1SIMREC version to the previous value of 1.22 (cf. sec. 6.1.4).
- A track reconstruction inefficiency of -3% per track is taken for the simulation for all reconstructed tracks.
- Two different generators (Django and Rapgap, cf. sec. 4.2) for the heavy hadron decay variables are used as templates for the fit.
- The uncertainty on the fragmentation function of the heavy quarks used in the simulation is estimated by repeating the fits with MC program templates generated using the Peterson fragmentation functions instead of the Lund fragmentation. The different fragmentations are done with the Rapgap MC program, which have (beauty/charm) cross sections for the Lund fragmentation of (48 pb/457 pb) and for the Peterson fragmentation of (41 pb/519 pb). This difference is taken as fragmentation uncertainty, even if the Rapgap Peterson fragmentation gives similar results as the Django Lund fragmentation.

S_l with 2 & ≥ 3 tracks	Default	41 ± 7 pb	510 ± 61 pb	15
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$	fit prob. (%)
Resolution	$1.09 \rightarrow 1.22$	10	-8	91
Efficiency	-3% per track	15	1	42
Model	Django $\rightarrow$Rapgap	15	-11	89
Fragmentation	Lund $\rightarrow$ Peterson	14	13	87
B multiplicity	$5.2 \rightarrow 5.0$	8	1	60
D^+ fraction	$0.197 \rightarrow 0.232$	3	1	43
E(jet) scale	10 %	-2	-1	18
E(e) scale	5 %	-5	2	16
Luminosity		1.5	1.5	
Total systematic error		29	20	
Statistical error		17	12	

Table 6.4: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_b (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable S_l in the VVM method with the simultaneous fit of the 2 & $\geq$ 3 tracks histograms. In the last column the fit probability is given.

- The average charged decay multiplicity of the beauty hadrons is changed from the simulation mean value of 5.2 to the world average measurements value of 5.0.
- The uncertainty on the decay branching fractions of charmed hadrons (e.g. D^+, where the simulation input value of 0.197 is changed to world average value of 0.232) is tested.
- A 10% uncertainty on the hadronic energy scale is taken into account by varying the jet cut in the data by this value.
- The NC selection uncertainty are considered with: The luminosity is known to a value 1.5%. An uncertainty on the scattered positron energy of 5% is tested by varying the scattered positron energy cut in the data by this value.

The systematic errors for the variable S_l fitting the 2 & $\geq$ 3 tracks histograms simultaneously are listed in table 6.4. Similar results for the systematic errors are obtained for the variable $min(S_{di})$ and are given in the appendix A in the table A.5.

The additional cross check with the splitting of the selected events into a horizontal and a vertical sample gives variations for charm or beauty cross sections of 18% or 20% respectively. This separation into two samples is not included into the systematics, because it reduces by a factor of two the statistics and the results are more sensitive to statistical fluctuations. As it should be, the results do not depend on the fitting method used, which are the Barlow fit or least squares fit. The systematic error on the heavy quark cross sections for the VVM method are dominated by the track resolution, the track inefficiency, the generator models and the fragmentations. A much better stability for the systematic tests is seen for the VVM method

VVM Method	$\sigma(ep \to b\bar{b})$ $[pb]$	$\sigma(ep \to c\bar{c})$ $[pb]$
Combined 2tr & $\geq$ 3tr		
S_l	$41 \pm 7 \pm 12$	$510 \pm 61 \pm 102$
$min(S_{di})$	$38 \pm 6 \pm 8$	$476 \pm 60 \pm 76$

Table 6.5: The integrated beauty and charm cross section with statistical and systematic errors for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ extracted with the VVM methods with the simultaneous fit of the 2 & $\geq$ 3 tracks histograms.

Default		$32\ pb$	$343\ pb$
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$
Mass (c resp. b)	$\pm 0.2\ GeV$ resp. $\pm 0.25\ GeV$	8	8
Scale (lower resp. upper)	$\frac{1}{4} \cdot \mu$ resp. $4 \cdot \mu$	8	6
PDF	CTEQ6M $\to$ CTEQ4F3	6	2
Total error		13	10

Table 6.6: The theoretical uncertainties for the NLO calculations based on the program HVQDIS.

than for the FVM method. The results for the variable S_l and $min(S_{di})$ of the systematic and statistical errors are summarized in table 6.5.

6.2.3 Theoretical Errors for the NLO Calculation

The results can not only be compared with the LO calculation and parton shower prediction of the generator MC program, but they can be compared with a NLO calculation, too. The integrated charm and beauty cross sections for the measured kinematical region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ is calculated using the NLO pQCD program HVQDIS [59, 21] with the parton density function CTEQ6M. The program calculates the parton level cross section $\sigma(e^+p \to c\bar{c}X, b\bar{b}X)$ without the use of fragmentation. The theoretical uncertainties for the NLO calculation is estimated with the variation of the heavy quark masses, the variation of the factorization respectively renormalization scale and the use of an older parton density function CTEQ4F3. The charm mass is varied by $\pm 0.2\ GeV$ around the default value of $m_c = 1.5\ GeV$, whereas the beauty mass is varied by $\pm 0.25\ GeV$ around the default value of $m_b = 4.75\ GeV$. The factorization and renormalization scales are set to the default value of $\mu_f = \mu_r = \mu = (Q^2 + 4M_q^2)$, where M_q is the mass of the heavy quark and the scale is varied by a factor 4, this means the lower scale value is one fourth and the upper scale value four times the default value. The total theoretical error results from the quadratic sum of the individual errors. The total relative theoretical uncertainty of the NLO calculations for the beauty (charm) cross section is 13% (10 %) and is given in table 6.6.

The NLO prediction with the CTEQ6M parton density function based on the program HVQDIS gives therefore $\sigma_{b\bar{b}} = 32 \pm 4\ pb$ and $\sigma_{c\bar{c}} = 343 \pm 34\ pb$.

6.3 Results: Charm and Beauty Production in Neutral Current Processes

The charm and beauty production cross sections are measured in the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$. The following results are summarized in figure 6.9. The results extracted with the VVM method with the variable S_l are:

$$\sigma(e^+p \to e'^+c\bar{c}X) = 510 \pm 61 \pm 102\ pb$$

$$\sigma(e^+p \to e'^+b\bar{b}X) = 41 \pm 7 \pm 12\ pb$$

and with the variable $min(S_{di})$ are:

$$\sigma(e^+p \to e'^+c\bar{c}X) = 476 \pm 60 \pm 76\ pb$$

$$\sigma(e^+p \to e'^+b\bar{b}X) = 38 \pm 6 \pm 8\ pb.$$

The FVM 2-tracks method has an unacceptable χ^2. The FVM 3-tracks method has huge systematic errors for the charm cross section, caused mainly by the track resolution, whereas the beauty cross section has comparable errors like the VVM method and the results are $\sigma_{b\bar{b}} = 50 \pm 8 \pm 15\ pb$ for the variable S_l and $\sigma_{b\bar{b}} = 46 \pm 7 \pm 9\ pb$ for the variable $min(S_{di})$.

The results agree with the published cross sections of the impact parameter method [1], which are $\sigma_{c\bar{c}} = 373 \pm 39 \pm 47\ pb$ and $\sigma_{b\bar{b}} = 55.4 \pm 8.7 \pm 12.0\ pb$ and which have better systematic uncertainties for the charm cross section than this analysis.

The NLO prediction with a CTEQ6M parton density function gives $\sigma_{b\bar{b}} = 32 \pm 4\ pb$ and $\sigma_{c\bar{c}} = 343 \pm 34\ pb$, which agree with the data.

The Django MC generator cross sections prediction with the leading order parton density function CTEQ6L are $\sigma_{c\bar{c}} = 492\ pb$ and $\sigma_{b\bar{b}} = 57\ pb$, whereas the Rapgap MC generator prediction with the MRSH leading order parton density function are $\sigma_{c\bar{c}} = 461\ pb$ and $\sigma_{b\bar{b}} = 36\ pb$. The predictions of the two MC generators are compatible with the data.

Recapitulating: The VVM method has smaller systematic and similar statistical errors when compared to the FVM method. The VVM method results agree within errors with the other H1 measurement, with QCD NLO predictions and with the MC prediction of the generator Django and Rapgap.

Finally, it shall be mentioned here that D. Meer has used the VVM method to extract the charm and beauty cross sections in photoproduction [60].

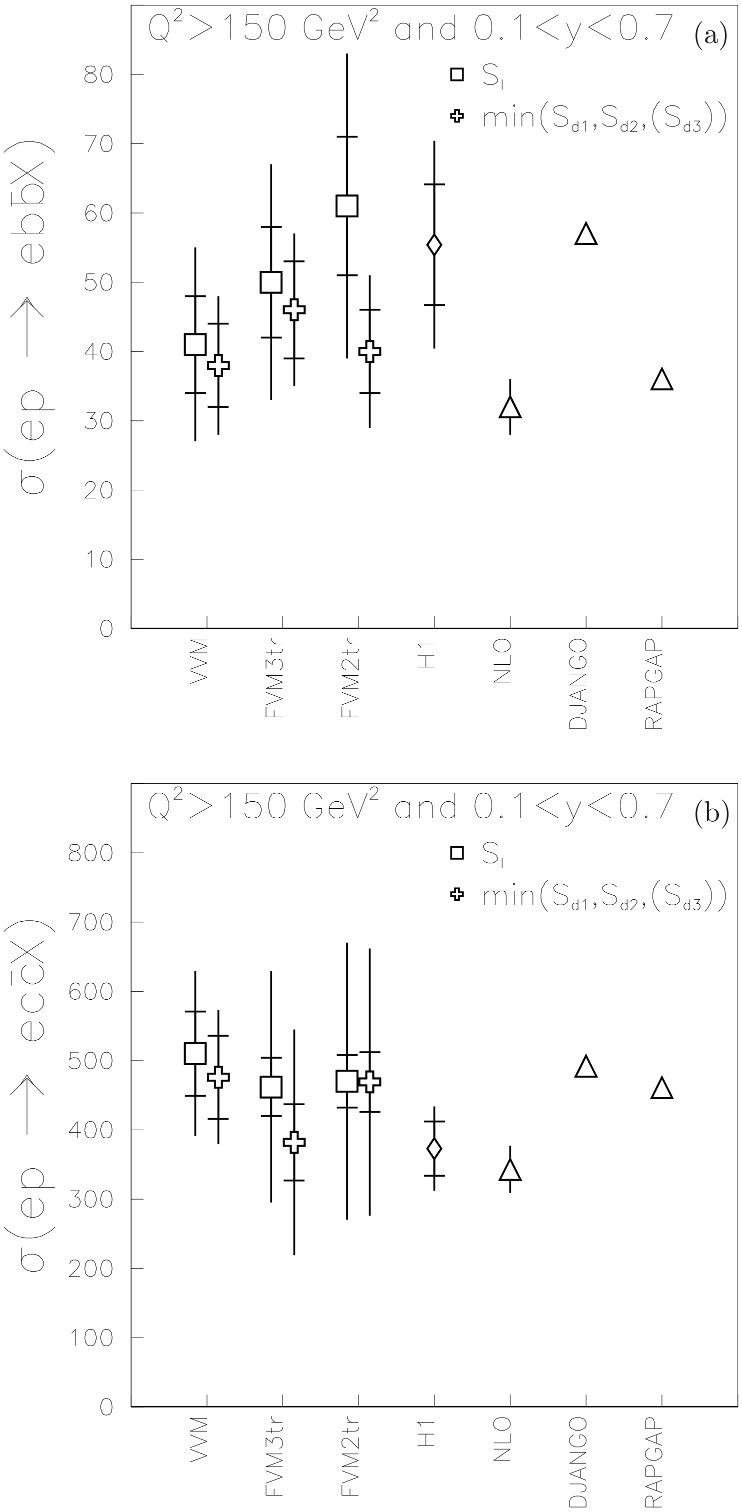


Figure 6.9: The measured beauty (a) and charm (b) cross sections given in pico barn in neutral current processes in the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ extracted with the VVM, FVM 3-tracks and FVM 2-tracks method. The results are displayed for the variables S_l and $min(S_{di})$. Also shown are the measurement from H1 [1], the prediction of the NLO QCD calculation and the prediction of the Django and Rapgap MC generators. The inner error bars show the statistical error only, whereas the outer error bars indicate the quadratically added statistical and systematic errors.

Chapter 7

Charm Production in Charged Current Processes

In order to determine the charm cross section in charged current processes (CC), a reliable extraction method has to be found. The VVM method, which was presented in section 6.2, has to be found to be the appropriate one. This fact is established by using a NC higher statistics subsample which is preselected to resemble as much as possible the CC kinematics. This particular preselection is based on a $P_{T,h} > 16\ GeV$ cut and it is referred to as the 'NC-metamorphosed' sample. Furthermore, the positron in the NC sample is not used for the primary vertex fit (see section 5.2) to get the same behavior of the primary vertex fit as in the CC sample. The full statistics NC sample without a $P_{T,h}$ cut is referred to as the 'standard NC' sample.

In the CC sample the beauty contribution is neglected, because the beauty cross section is expected to be approximately 100 times smaller than the charm cross section[1]. According to the MC program prediction, the about 800 selected CC events contain less than two beauty events.

To extract the charm content in the CC sample with the VVM method only two free fit parameters are necessary, the uds and charm contributions. The beauty contribution is set to zero. The cross sections discussed below are determined for the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$, which is extrapolated from the 'visible' kinematic region $Q^2 > 284\ GeV^2$ and $0.1 < y < 0.7$ (see section 4.6). The extrapolation is about 12%.

7.1 Tests of the Extraction Method in NC

The confirmation that the VVM method is applicable in CC is divided into four steps:

1. The NC and CC kinematics is made similar by using the same $P_{T,h}$ cut in the NC sample.

[1] The prediction from the Django MC program in the kinematic range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ is for the anti-beauty cross section $\sigma(e^+p \rightarrow \bar{\nu}_e \bar{b} X) = 0.056\ pb$ and for the charm cross section $\sigma(e^+p \rightarrow \bar{\nu}_e c X) = 6.05\ pb$.

2. The fit procedure is checked in the NC-metamorphosed sample.
3. The systematic errors for the determination of the cross sections in CC are determined in the NC-metamorphosed sample.
4. A low statistics NC-metamorphosed sample is produced (see table 7.3).

The following subsections will discuss these four points.

7.1.1 NC-metamorphosed Sample

Important for the verification of the extraction method is that the NC events have a similar kinematic behaviour as the CC events. The most prominent way to make the NC sample resemble the CC kinematics is a cut on $P_{T,h}$. Its effect can be readily seen by comparing the distributions ϑ_h with and without $P_{T,h}$ cut.

In the figure 7.1(b) the ϑ_h distribution of the NC events (open dots) and the ϑ_h distribution of the CC events in the figure 7.1(a) look very different. Only after a cut with $P_{T,h} > 16\ GeV$ the NC ϑ_h distribution (solid dots) in figure 7.1(b) is similar to the CC ϑ_h distribution. The $P_{T,h} > 16\ GeV$ cut does not change much the other NC distributions. As one example, in the lower two figures 7.1(c) and (d), the CC $E - P_{z,h}$ distribution is compared to the NC $E - P_{z,h}$ distribution before and after the $P_{T,h}$ cut.

The jet related variables, namely the $p_{T,Jet}$ and ϑ_{Jet} distributions, are compared for the CC sample and the NC sample before and after the $P_{T,h}$ cut in the appendix B in figure B.5. The $P_{T,h}$ cut makes again the $p_{T,Jet}$ and ϑ_{Jet} distributions between NC-metamorphosed and CC sample similar. In addition, it is important to confirm that the decay multiplicity and S_l distributions between NC-metamorphosed and CC sample are similar. These distributions are presented in appendix B in figure B.6. All NC-metamorphosed data distributions are rather well described by the simulation.

7.1.2 Results of the NC-metamorphosed Sample

The tests of the charm extraction method are done in the NC-metamorphosed sample with a method which is applicable in the CC sample. The VVM method applied to the CC sample yields zero data events with a decay-multiplicity of three. The only method which is applicable in the CC sample is therefore to use events with 1 & 2 tracks decay vertex.

In the standard NC sample extraction the light, charm and beauty cross sections with the simultaneous fit of the 1 & ≥ 2 decay-multiplicities is consistent with the simultaneous fit of the 2 & ≥ 3 decay-multiplicities, which can be seen in table 6.3(f) and (a).

One weak point in this extraction method is that the $\chi^2/ndof$ in table 6.3 of the 1 & ≥ 2 decay-multiplicities is high. Additionally, the $P_{T,h} > 16\ GeV$ cut could produce unexpected effects. E.g. the higher track density in the jet could enhance the accidental linking of CST hits to CJC-tracks.

To confirm the correctness of the charm extraction method the NC-metamorphosed sample with the 1 & ≥ 2 decay-multiplicities is used.

A simultaneous fit of the distributions S_l or $min(S_{di})$, of the 1 & ≥ 2 decay-multiplicities in the NC-metamorphosed sample is done using MC templates, where only the charm and

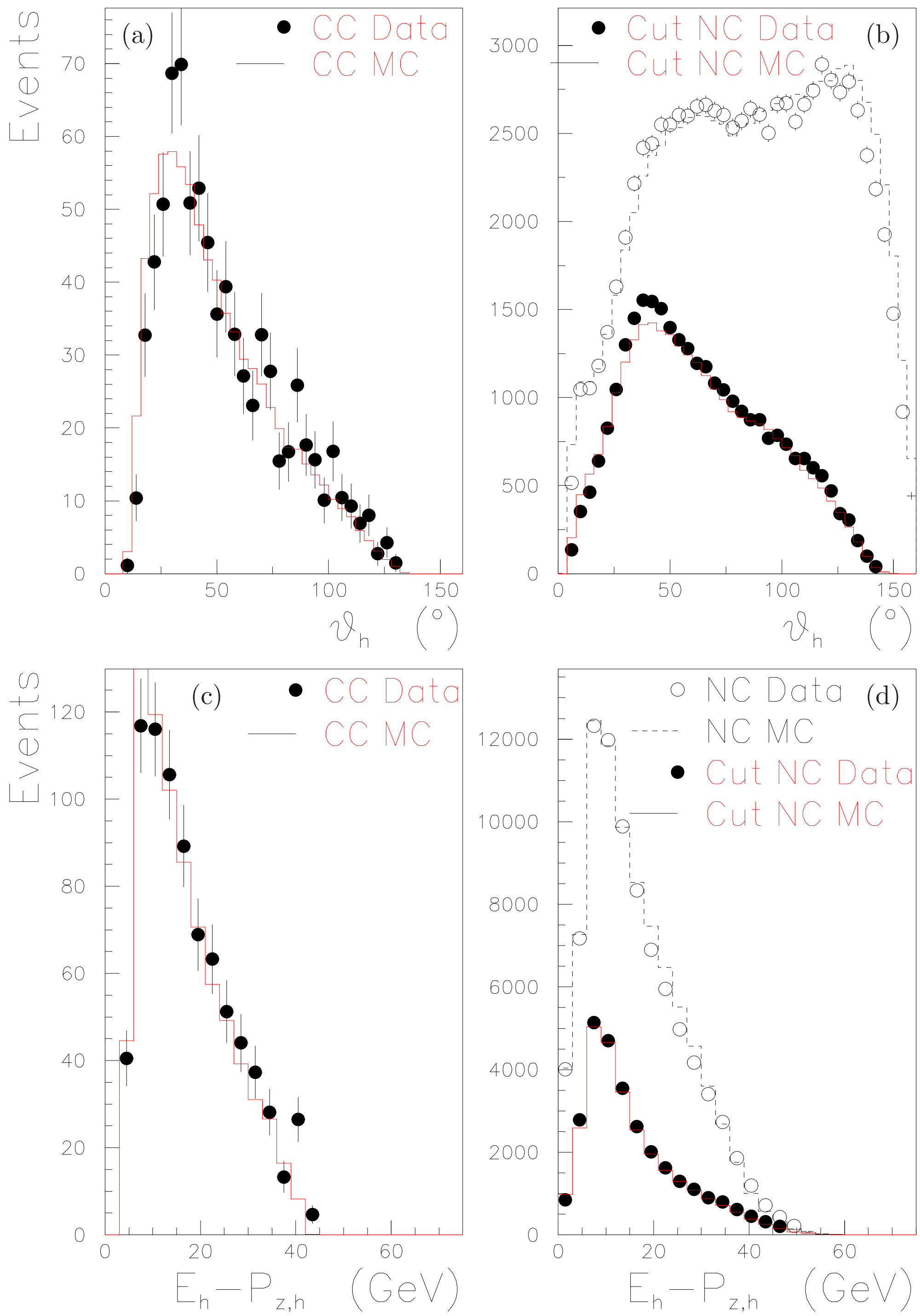


Figure 7.1: Comparison of CC (a,c) and NC (b, d) kinematics: In figure (b) the hadronic polar angle ϑ_h for the standard NC events (open dots) and for the NC-metamorphosed sample (solid dots) are displayed. The $P_{T,h} > 16\ GeV$ cut produces a similar distribution as for the CC events, shown in figure (a). The $E - P_{z,h}$ distribution does not change much for the NC-metamorphosed sample evidenced in figure (d) and looks similar as in the CC which is represented in figure (c).

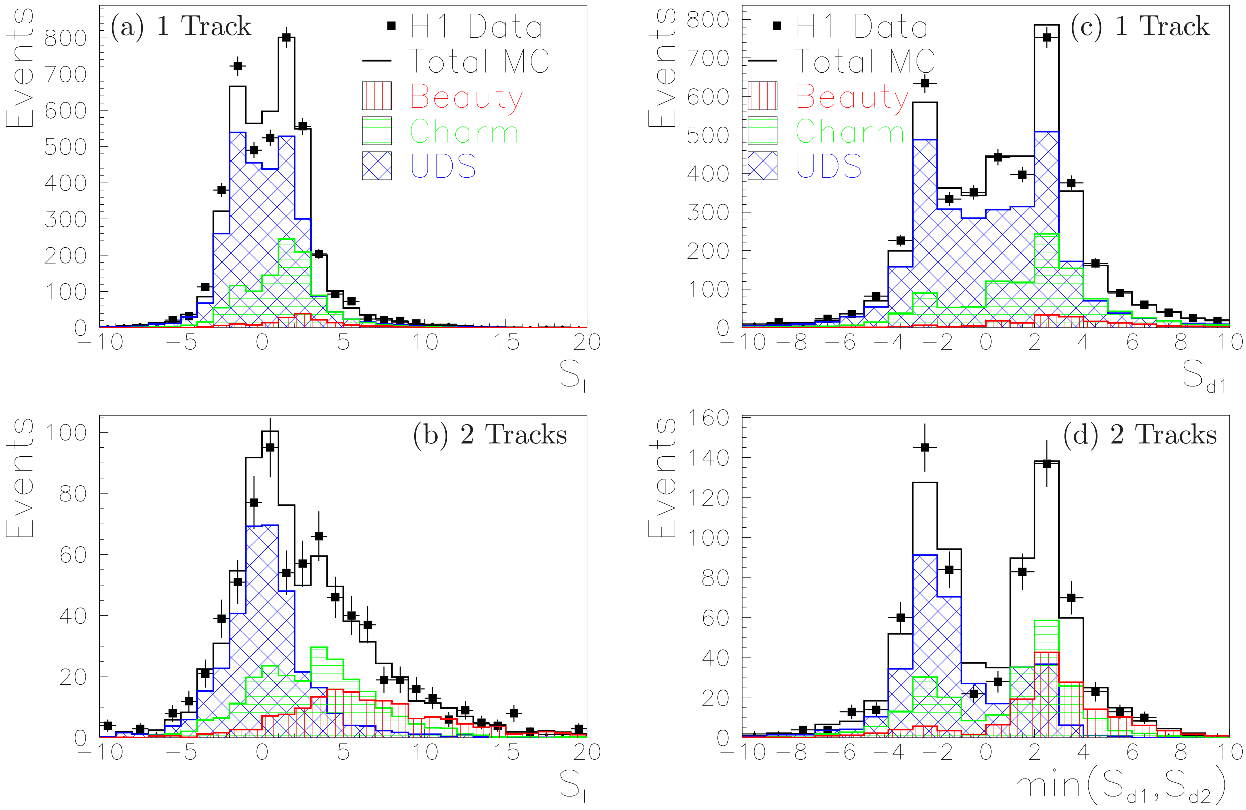


Figure 7.2: The distributions S_l and $min(S_{di})$ are shown for the VVM method for 1 & ≥ 2 tracks in the NC-metamorphosed sample with fixed beauty parameter. The figure displays the variable S_l (left) and $min(S_{di})$ (right) for the same events with ≥ 2 tracks (b, d) and 1 track (a, c) associated with the secondary vertex. Superimposed on the data points (with statistical error only) are the light, charm and beauty quark contributions as obtained from the fit with the Django MC program templates.

uds normalizations are free fit parameters. In figure 7.2 these distributions of the NC-metamorphosed sample are displayed. The beauty fraction is taken from the simulation and is fixed, in analogy to the CC sample where the beauty distribution is fixed and set to zero.

The results, summarized in table 7.1(a), are consistent with standard NC selection and with the fit procedure described in subsection 6.2.1. Again the normalization bin is used. It contains all NC-metamorphosed events except the ones with ≥ 1 tracks associated with the secondary vertex. Then a simultaneous fit of the normalization bin and the distributions of the histogram in figure 7.2(a) and (b), as well as the normalization bin and figure 7.2(c) and (d) is done. The results are compatible with the results extracted without the normalization bin. They are listed in table 7.1(b). All fit probabilities in the table 7.1 are rather low,

	VVM Method in NC-metamorphosed	$\sigma(ep \to c\bar{c})$ $[pb]$	$\sigma(ep \to uds)$ $[pb]$	$\chi^2/ndof$	$\rho_{c,uds}$
	Combined 1tr & $\geq$ 2tr				
(a)	S_l	466 ± 43	2093 ± 84	94/57	-0.83
	$min(S_{di})$	447 ± 42	2126 ± 81	92/40	-0.81
(b)	norm. & S_l	577 ± 33	1820 ± 29	100/58	-0.90
	norm. & $min(S_{di})$	559 ± 32	1825 ± 27	67/43	-0.89

Table 7.1: The fit results for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ with the VVM method in the NC-metamorphosed sample: The beauty fraction is taken from the Django MC program and is fixed. Only a two parameters fit for charm and light quarks is used. The integrated light and charm cross sections are listed. Only statistical errors are given. These results are compatible with the results of the standard NC sample listed in table 6.3(a) and (b).

S_l with 1 & $\geq$ 2 tracks	Default	2093 ± 84 pb	$466 \pm 43pb$	94/57
Systematic		$\delta(uds)(\%)$	$\delta(c)(\%)$	$\chi^2/ndof$
Resolution	$1.09 \to 1.22$	10	4	68/55
Efficiency	-3% per track	15	5	63/57
Model	Django $\to$Rapgap	1	3	82/57
Fragmentation	Lund $\to$ Peterson	6	11	82/57
Total systematic error		19	13	

Table 7.2: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_{uds} (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable S_l using the VVM method with the simultaneous fit of the 1 & $\geq$ 2 tracks histograms in the NC-metamorphosed sample. This total systematic error is used for the CC result in table 7.6.

around 1%. This indicates that the MC program templates do not describe well the data.

7.1.3 Systematic Errors in the NC-metamorphosed Sample

To avoid statistical fluctuations in the CC sample with less than 1000 events, the systematic studies are done in the NC-metamorphosed sample with around 40000 data events. The stability of the algorithm is tested only with the dominant contribution of the systematic uncertainties seen in NC (see section 6.2.2).

The systematic errors for the variable S_l fitting the 1 & $\geq$ 2 tracks histograms simultaneously are listed in table 7.2. A similar result for the systematic error is obtained for the variable $min(S_{di})$ and is given in appendix C in the table C.1. If the normalization bin is taken the statistical error is halved as one sees when comparing the results for the CC sample in table 7.5(a) with (b). However, the systematic errors with the normalization bin are bigger than without normalization bin and are given in appendix C in the tables C.2 for

S_l	1 & ≥ 2 tracks		1 & ≥ 2 tracks including normalization bin	
	$\sigma(ep \to c\bar{c})$ $[pb]$	$\sigma(ep \to uds)$ $[pb]$	$\sigma(ep \to c\bar{c})$ $[pb]$	$\sigma(ep \to uds)$ $[pb]$
Default	466 ± 43	2093 ± 84	577 ± 33	1820 ± 29
Experiment				
1	349 ± 282	2055 ± 555	439 ± 178	1831 ± 154
2	270 ± 257	2940 ± 498	700 ± 194	1766 ± 167
3	214 ± 243	2658 ± 485	521 ± 194	1842 ± 164
4	488 ± 273	2360 ± 534	765 ± 200	1663 ± 168
5	78 ± 218	2694 ± 459	381 ± 165	1839 ± 146
6	560 ± 266	2139 ± 495	696 ± 200	1787 ± 168
7	531 ± 262	2141 ± 494	750 ± 198	1583 ± 163
8	815 ± 273	1536 ± 487	794 ± 197	1615 ± 165
9	92 ± 210	2470 ± 438	122 ± 160	2321 ± 143
10	239 ± 226	2574 ± 456	448 ± 171	1956 ± 150
mean	364 ± 251	2357 ± 490	562 ± 186	1820 ± 159

Table 7.3: The charm and uds cross sections for the low statistics NC-metamorphosed samples with a simultaneous fit of the 1 & ≥ 2 tracks histograms: In the third row the cross sections with the full statistics of the NC-metamorphosed sample without (2^{nd} and 3^{rd} col.) and with normalization bin (4^{th} and 5^{th} col.) are given. In the following rows the results from the test sample are listed, as explained in the text. In the 2^{nd} and 3^{rd} (4^{th} and 5^{th}) column the charm and uds cross sections without (with) the normalization bin are shown.

the distribution S_l and in table C.3 for the distribution $min(S_{di})$. The explanation of the enlargement of the systematic errors is given on page 72.

7.1.4 Test of the Statistical Problem in the NC-metamorphosed Sample

One question arises: Is it possible to extract a cross section with the low statistics accumulated in the CC sample?

To test this, ten test samples, each with the same number of events as the CC sample, are generated by randomly selecting events from the NC-metamorphosed sample. The cross sections are then measured from the distribution S_l of these low statistics NC-metamorphosed samples with the same fit procedure as used in section 7.1.2. In appendix B the last of these tests is demonstrated in figure B.7. In table 7.3 the cross sections extracted from the ten test samples are reported and they are compared to the results of the full statistic of the NC-metamorphosed sample. Variations up to three sigma are observed, but otherwise the tests are compatible with the default result.

Recapitulating: The VVM method gives consistent results for the NC-metamorphosed sample even when a two parameters fit with a fixed beauty contribution for the simultaneous fit of the 1 & ≥ 2 tracks candidate is done. The systematic errors are calculated and will be used for the CC results. The statistics should be large enough to give a reliable results for

	All CC	1 track	2 tracks	3 tracks
Data	805.4 ± 28.4	70.3 ± 8.4	14.4 ± 3.7	0
MC all	789.7 ± 4.3	61.2 ± 1.2	7.4 ± 0.4	0.74 ± 0.13
uds	613.4 ± 3.8	40.9 ± 1.0	4.2 ± 0.3	0.25 ± 0.08
charm	175.4 ± 2.0	20.2 ± 0.7	3.2 ± 0.3	0.49 ± 0.11
beauty	0.9 ± 0.1	0.1 ± 0.1	0.04 ± 0.04	0

Table 7.4: Comparison of the number of CC data events with the MC program prediction for the VVM method, where the simulation is normalized to the experimental luminosity: The data agree with the predictions within the statistical errors. For the 2 tracks decay vertex events a two sigma deviation between simulation and data is seen. The 2 tracks vertex events have a ratio of charm events to uds events of about 1:1. The beauty contribution can be neglected. Only statistical errors are given.

the cross section.

The next step is to use the algorithm for the real CC data.

7.2 Charm Extraction with the VVM Method in CC Processes

In the following section the charm cross section in CC with the VVM method is determined.

The CC selection is illustrated in section 4.6, where the simulation describes the kinematic variables of the selected CC data rather well. The same selection criteria for CC events as for the NC events are used for the CST improved tracks, the primary vertex and the jets (see chapter 5). The important distributions are presented in the appendix B in the figures B.1, B.2 and B.3[2]. All distributions are well described by the simulation within the statistical errors.

The about 800 CC data events are analyzed with the VVM method exactly in the same way as for the NC sample. From row two in table 7.4 it can be seen that in the CC data around 70 events have a decay-multiplicity of one, 14 events have decay-multiplicity of two and zero events exist with a three track vertex. The MC program prediction shown in the second row is similar apart from the 2 tracks decay vertex events, where a two sigma deviation between simulation and data exists. The MC program predictions for the number of the light, charm and beauty quark events are given in the following rows of the table 7.4.

The same information is displayed in the figure 7.3, where the track multiplicity of the secondary vertices found by the VVM method are presented[3]. The relative uds and charm fractions indicated in this figure are those obtained from the fit of the histograms in the figure 7.4.

[2]These are the analog distributions shown for the NC events in figures 5.1, 5.3 and 5.4.

[3]The analog plot is shown for the NC sample in figure 6.6 and the corresponding explanations are given in section 6.2.

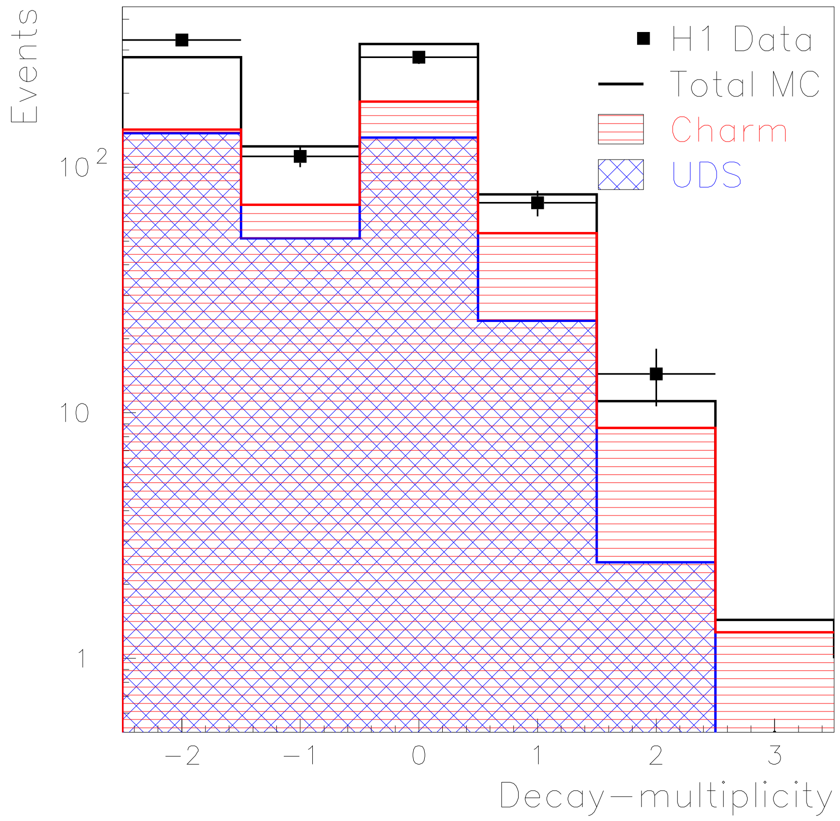


Figure 7.3: The number of tracks contributing with a weight greater than 80% to the secondary vertex for the CC sample, is used as a measure of the decay-multiplicity. All events in the bin -2 or -1 respectively are events with zero or one CST improved track inside a jet, where the adaptive fitter is not used. No data event is found with a 3 tracks vertex. Superimposed on the data points with statistical errors only are the light and charm quark contributions as obtained from the fit of the figures 7.4.

The measured S_l and $min(S_{di})$ distributions from the data are fitted with the shapes of the light and charm contributions, which are obtained from the simulation. The two histograms on the left (right) side of the figure 7.4 present the variable S_l ($min(S_{di})$) for the 1 track & ≥ 2 tracks candidates. The two parameters fit of the 1 track & ≥ 2 tracks for the figure 7.4(a) and (b) simultaneously yields a sample composition of $f_{uds} = 18.5 \pm 24.2\%$ and $f_c = 81.5 \pm 25.2\%$ with a $\chi^2/ndof = 9/20$ and the correlation coefficient $\rho_{c,uds} = -0.91$. Similar results are obtained when using the variable $min(S_{di})$ and those results are summarized in table 7.5(a), where the cross sections are calculated with the help of equation 3.1.

Again as a cross check the normalization bin is used. It contains all CC events except the ones with ≥ 1 tracks associated with the secondary vertex. A simultaneous fit of the normalization bin and the distributions of the histograms in figure 7.4(a) and (b) as well as the normalization bin and figure 7.4(c) and (d) is done. This fit yields for the variable S_l a

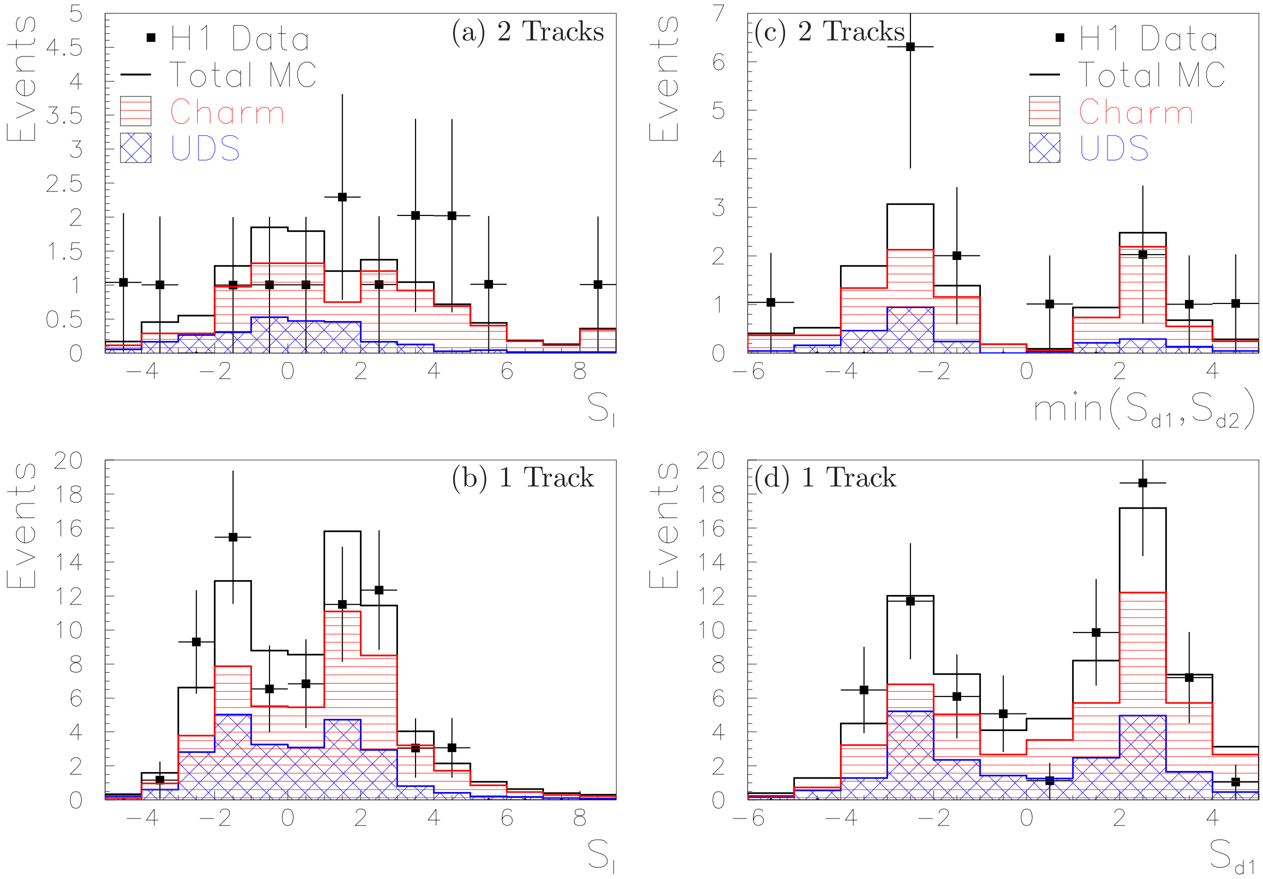


Figure 7.4: The VVM method for 1 & ≥ 2 tracks in the CC sample: The figure shows the variable S_l (left) and $min(S_{di})$ (right) for the same events with 2 tracks (a, c) and 1 track (b, d) associated with the secondary vertex. Superimposed on the data points with statistical error only are the light and charm quark contributions as obtained from the fit with the Django MC program templates.

sample composition of $f_{uds} = 45.9 \pm 10.9\%$ and $f_c = 54.1 \pm 11.0\%$ with a $\chi^2/ndof = 10/21$ and the correlation coefficient $\rho_{c,uds} = -0.95$. The results are compatible with the results without normalization bin and are listed in table 7.5(b). The difference between the results with and without normalization bin in table 7.5(a) and (b), is that the statistical error is halved for the result with normalization bin. In contrast the systematic errors are larger for the fit with normalization bin than without as is demonstrated by the systematic studies in section 7.1.3.

The final results for charm and uds cross sections in CC, with the systematic errors obtained in section 7.1.3, are listed in the table 7.6.

	VVM Method in CC	$\sigma(ep \to c)$ $[pb]$	$\sigma(ep \to uds)$ $[pb]$	$\chi^2/ndof$	$\rho_{c,uds}$
	Combined 1tr & $\geq$ 2tr				
(a)	S_l	17.7 ± 5.5	5.8 ± 7.8	9/20	-0.91
	$min(S_{di})$	17.2 ± 6.5	8.1 ± 9.2	19/16	-0.93
(b)	norm. & S_l	15.3 ± 3.1	10.5 ± 2.5	10/21	-0.95
	norm. & $min(S_{di})$	16.2 ± 3.3	9.9 ± 2.7	19/18	-0.95

Table 7.5: The integrated charm and light cross sections in CC for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ with the VVM method are summarized. Only statistical errors are given.

VVM Method in CC	$\sigma(ep \to c)$ $[pb]$	$\sigma(ep \to uds)$ $[pb]$
Combined 1tr & $\geq$ 2tr		
S_l	$17.7 \pm 5.5 \pm 2.3$	$5.8 \pm 7.8 \pm 1.1$
$min(S_{di})$	$17.2 \pm 6.5 \pm 1.9$	$8.1 \pm 9.2 \pm 1.5$
norm. & S_l	$15.3 \pm 3.1 \pm 5.8$	$10.5 \pm 2.5 \pm 1.2$
norm. & $min(S_{di})$	$16.2 \pm 3.3 \pm 5.8$	$9.9 \pm 2.7 \pm 0.9$

Table 7.6: The integrated charm and light cross sections with statistical and systematic errors for the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ with the VVM methods with the simultaneous fit of the 1 & ≥ 2 tracks histograms are summarized. The systematic errors are measured with the NC-metamorphosed sample.

7.2.1 Lower Charm Efficiency in CC Processes than in NC Processes

By comparing the 2 tracks decay vertex events from the uds and charm distributions in the CC sample with these two distributions in the NC sample one recognizes the following: The charm distribution in the figure 6.7(d) from the NC sample has a pronounced excess for positive impact parameters, which is not seen for the charm distribution in the CC sample represented in figure 7.4(c). This means an evidently better separation between the uds and charm components is recognized in the NC sample than in the CC sample.

The explanation, that the different kinematics between NC and CC can make the separation between uds and charm worse, can be partially discarded. The reason is, that if one compares these distributions in the NC-metamorphosed sample displayed in figure 7.2(d) with these distributions in the NC sample indicated in figure 6.7(d), the separation between uds and charm is good in both samples.

To avoid as much as possible the different kinematics between NC and CC the NC-metamorphosed sample is compared to the CC sample in the figures 7.5. For both samples the figures 7.5 (a) and (b) show the variable $min(S_{d1}, S_{d2})$ for the uds and the charm distributions with the 2 tracks decay vertex events. All histograms in figure 7.5 are normalized to the same number of entries. The figure (a) with the uds distribution evidences no large difference between the NC-metamorphosed sample and the CC sample. In contrast, the charm distribution (figure (b)) for the CC sample does not have the pronounced excess for positive

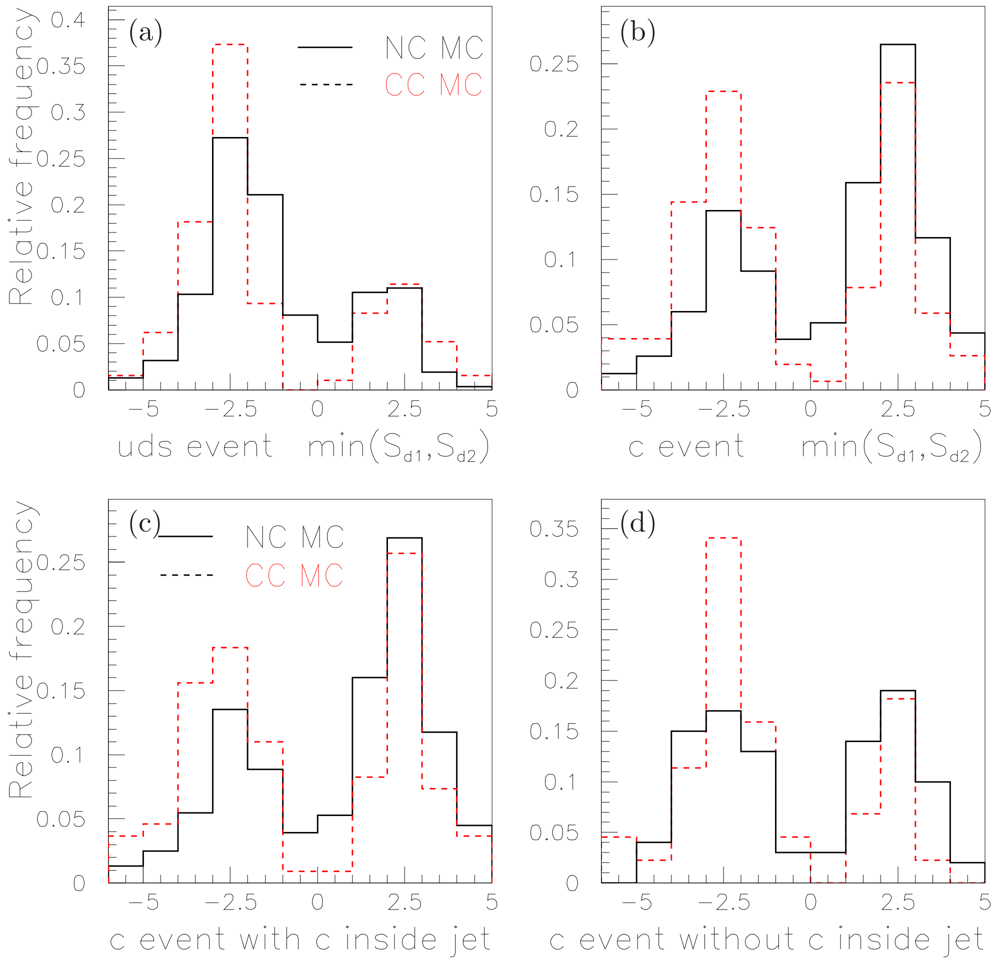


Figure 7.5: Comparison of the simulated 2 tracks decay vertices from the CC sample (dashed line) and from the NC-metamorphosed sample (full line): In figure (a) respectively (b) the relative frequency of the reconstructed 2 tracks vertices for the uds respectively charm events in CC and NC-metamorphosed are displayed. For charm events a better separation is evidenced for the NC-metamorphosed sample than for the CC sample. The charm events are distinguished on the one hand on a secondary vertex inside a jet which really originates from a charm hadron put in figure (c) and on the other hand on a secondary vertex inside a jet which has no charm hadron put in figure (d). The histogram from the CC sample in figure (d) looks similar to the histogram in (a) for the uds event. For further explanations see the text.

impact parameters, which is seen in the NC-metamorphosed sample.

As noticed in section 2.4 the BGF Feynman diagrams from NC (figure 2.2) and CC (figure 2.3) are different. In a CC event with a BGF diagram only one charm quark is produced, the other is a anti-strange or anti-down quark, whereas in the NC event with a BGF diagram two charm quarks are produced.

To check if the BGF diagrams can degrade the charm signature, one looks inside a cone around the jet of each 2 track vertices with a tagged charm event, which means one looks inside the jet of the candidate selected for figure 7.5(b). If the selected jet candidate has a charm hadron, then the candidate really originates from a charm quark and is put in figure 7.5(c), whereas the candidate without a charm hadron inside the jet is put in figure 7.5(d). These vertex tagged non-charm jets originate from light quarks or gluons and for them one introduces the term 'mistag'. The ratio in the NC-metamorphosed sample between vertex tagged charm jets and mistag jets is 17:1, whereas in CC the ratio is 2.5:1. The histogram from the CC sample in figure (d) with mistag events looks similar to the figure (a) from the CC uds events, which can be understood with the tagging of the anti-strange or anti-down quark events. On the other hand the histogram from the NC-metamorphosed sample in figure (d) with mistag events looks symmetric and can be understood for a random selection of candidates.

The two histograms from the CC and NC-metamorphosed sample in figure (c) with correctly tagged charm events looks similar, but the separation between positive and negative value entries for the $min(S_{d1}, S_{d2})$ variable is still slightly better for the NC-metamorphosed sample than for the CC sample.

7.2.2 Control of the 2 Tracks Decay Vertex Events

The excess of the 2 tracks decay vertex events should be checked for consistency. Are these 14 data events really CC events or is something else wrong with these 2 tracks decay vertex events?

The non-CC background is around 2.5% of the total CC selected sample, where the largest contribution is located at $P_{T,h} < 25\ GeV$, as seen in figure 4.6. In appendix D the table D.1 shows that 12 of the 14 selected 2 tracks candidates have a higher $P_{T,h}$. According to [41] the background of this CC selection, where the $P_{T,h} > 12\ GeV$, is around 2% and is distributed at low Q^2, at low $P_{T,h}$ and high y.

In figure 7.6 the distributions of these 14 events in the variables the decay length, the impact parameter, the total charges and the invariant mass are presented. Because the charm quark is positively charged the hope is that the total charges of both tracks have some discriminating power between light and charm hadrons. Unfortunately in the figure (c), with the total charges variable prediction from the simulation, no clear separation between these two components is seen. The invariant mass distribution in figure (d) shows that the selected candidates do not have a peak for the invariant K_s mass, which indicates that at least not only strange quarks are accumulated. Additionally, it is unlikely that the candidate with the invariant mass of 2.8 GeV is a charm hadron event. The decay length distribution in figure 7.6(a) has no events with suspicious too large decay length values. Furthermore, one sees that the events are distributed in a manner that they are compatible with an exponential decay. In the figure 7.6(b) both impact parameters are put in the same histogram and are compared to the simulation.

One further test is to check how the (symmetric) distributions (e.g. S_l and S_{di}) behave, if one looks at only the positive or negative values. This makes sense, because decays of long-lived particles are signaled by positive S_l and S_{di}, whereas the finite track resolution yields a symmetric distribution for positive and negative signs of S_l and S_{di}. In table 7.7 the

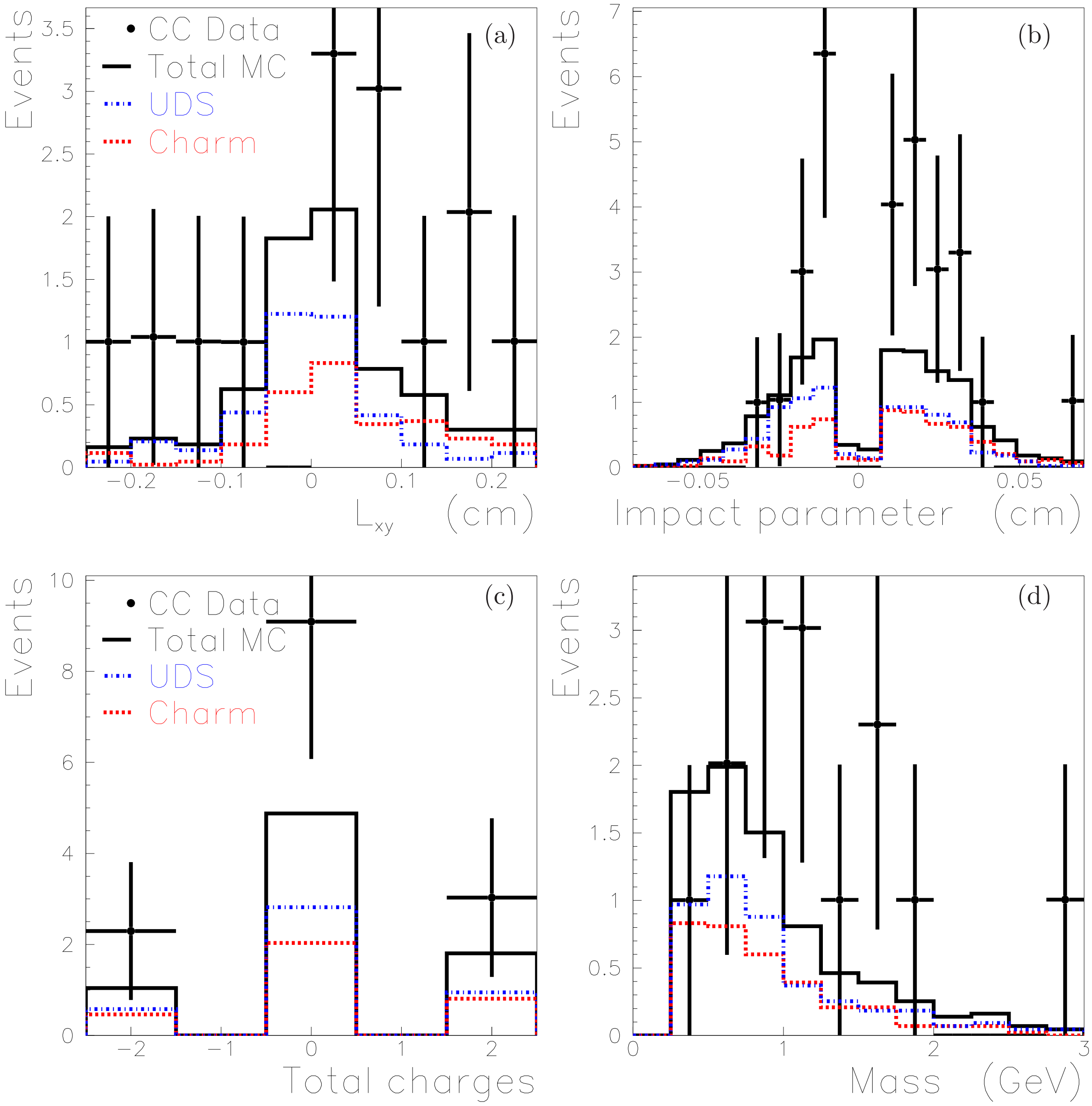


Figure 7.6: The 14 CC events with a 2 tracks decay vertex measured with the VVM method: In (a) the decay length, (b) both impact parameters, (c) total charges and (d) the invariant mass distributions are shown. The relative uds and charm fractions are that predicted from the Django MC program.

14 events are separated into events with a negative decay length and in events with a positive decay length. Four data events have negative decay length where the MC program predicts 3.1 events. Ten data events are found with positive decay length, which is to be compared with the MC program prediction of 4.2 events[4]. Furthermore these 14 events are divided

[4]One can check this with figure 7.6(a).

Test	Data	MC All	Charm	UDS
All	14.4	7.4 ± 0.4	3.2 ± 0.3	4.2 ± 0.3
$S_l < 0$	4.0	3.1 ± 0.3	1.0 ± 0.2	2.1 ± 0.2
$S_l > 0$	10.4	4.2 ± 0.3	2.1 ± 0.2	2.1 ± 0.2
$S_{d1} < 0$ && $S_{d2} < 0$	2.0	1.7 ± 0.2	0.5 ± 0.1	1.3 ± 0.2
$S_{d1} > 0$ && $S_{d2} > 0$	5.1	2.6 ± 0.3	1.4 ± 0.2	1.2 ± 0.2

Table 7.7: The measured and expected number of the 2 tracks decay vertex: The negative distributions are well described by the simulation, whereas for the positive distributions an excess exist in the data compared to the MC program prediction. The statistical errors for the MC prediction are given.

into events with positive impact parameters for both tracks and in events with negative impact parameters for both tracks. These results are listed in table 7.7. For negative impact parameters the data is reproduced by the MC simulation, whereas a cut for positive impact parameters for both tracks, gives again more events measured in the data than predicted from the simulation.

The concluding remark is that nothing evidently wrong is seen in these tests of the 2 tracks decay vertex events and the next step is to visualize these events.

7.2.3 Pictures of the 2 Tracks Decay Vertex Events

The event display of one CC event with a 2 tracks decay vertex is presented in figure 7.7, the other 13 events are shown in the appendix D in the figures from D.1 to D.13.

The H1 detector with calorimeter information is visualized in figure 7.7(a) in the side view and in figure (b) in a front view. One recognizes in the front view that this event has a large missing transverse momentum, assumed to be carried by an unseen neutrino.

In figure 7.7(c) the schematic $r\phi$ view of the CST is displayed with around 25-times magnified of the center of figure (b). The lines are the trajectory of all reconstructed charged tracks from the event. This event has enough tracks to determine well the primary vertex.

The figure (d) shows the CST including the hits, where now only the two selected CST-improved tracks which produce the 2 tracks decay vertex are displayed.

In figure (e) in a closer view the decay vertex of the two intersecting tracks are evidenced. The reconstructed secondary vertex position and the reconstructed primary vertex position with a one sigma error are displayed. The arrow gives the direction of the reconstructed transverse momentum of the decaying objects. This event has a positive radial decay length with a decay length significance of 5.3 and a missing transversal momentum of $P_{T,h} = 56.7\ GeV$.

This event has been presented in the H1 publication [61] as an isolated muon event with missing transverse momentum. In figure 7.7(b) two jets can be distinguished: In one isolated jet the muon candidate is found and in the other jet a 2 tracks decay vertex is measured. This publication notes that there is an excess of observed muon and electron numbers of events over the Standard Model predicted for $P_{T,h} > 40\ GeV$. In the muon channel for all e^+p data the observed numbers of events is three whereas the Standard Model expectation

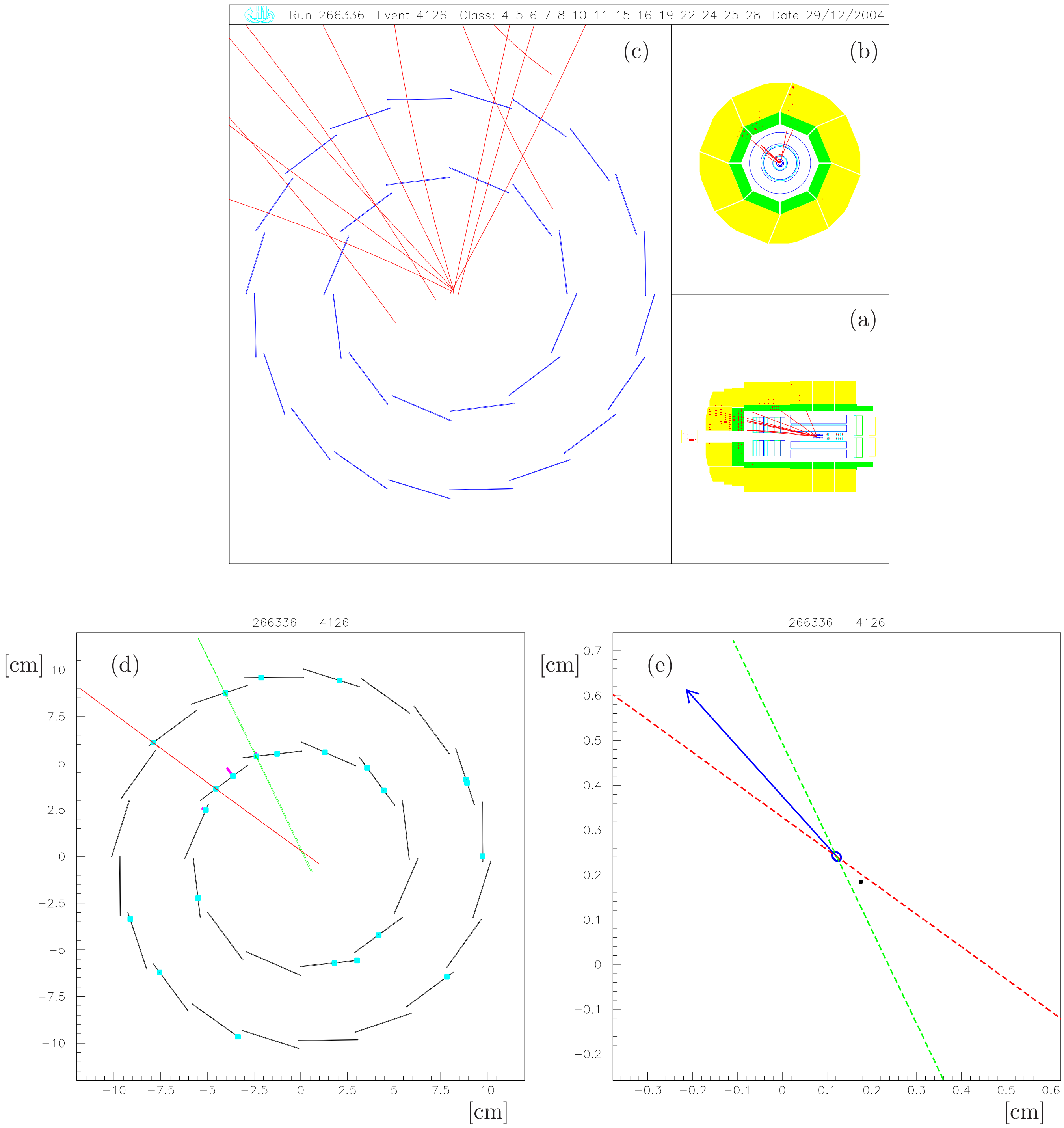


Figure 7.7: A CC event with a 2 tracks decay vertex selected with the VVM method: A CC event measured by the H1 detector in the side view (a) and in the $r\phi$ view (b) is displayed. With around 25-times magnified of the center of figure (b) is presented in (c) a schematic $r\phi$ view of the CST with all reconstructed charged tracks from the event. In (d) the CST including hits (blue points) and the two selected CST-improved tracks (lines) which produce the 2 tracks decay vertex is visualized. Again in a closer view, allowing the different scale on the axis, one sees in (e) the decay vertex of the two intersecting tracks, where the one sigma error of the secondary vertex is the ellipse. The arrow represents the direction of the reconstructed transverse momentum of the decaying objects. In opposite direction to the arrow the primary vertex position with a one sigma error (black ellipse) is displayed. The measured radial decay length is $L_{xy} = 0.078\ cm$ and the reconstructed invariant mass is $1.54\ GeV$.

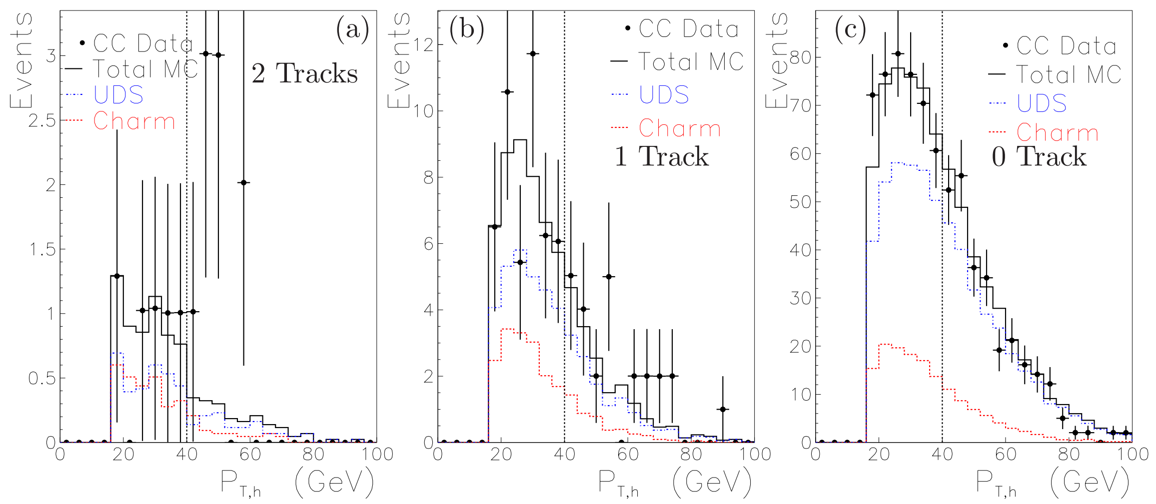


Figure 7.8: The transverse momentum $P_{T,h}$ distribution for events with (a) 2 tracks, (b) 1 track and (c) less or equal 0 track associated with the secondary vertex. The vertical dotted line displays the $P_{T,h} = 40\ GeV$ cut. For (c) the data (dots) with statistical errors are well described by the simulation (histogram). A significant excess in the data is seen in (a) for the 2 tracks decay vertex at $P_{T,h} > 40\ GeV$. The simulation is normalized to the data luminosity.

is (0.55 ± 0.12) events.

All 13 selected events in the appendix D look well, except the event in run 268001 with event-number 89571. This event has two collinear CST-improved tracks. This is of course not the nice long-lived signature that one tries to find. But both tracks have been chosen by the adaptive fitter to contribute with a weight greater than 80% to the secondary vertex. The huge one sigma error of the secondary vertex manifests that a clear separation between primary and secondary vertex is impossible. Nevertheless this event is kept.

Recapitulating: A visual scanning of these 14 events demonstrates no evidently strange signature.

7.2.4 Excess of Events with Decay Vertices at $P_{T,h} > 40\ GeV$

The figure 7.8 shows the missing transverse momentum $P_{T,h}$ distribution for the events with 2 tracks, 1 track and less or equal 0 track associated with the secondary vertex. A clear discrepancy between simulation and data is seen for the 2 tracks decay vertex at $P_{T,h} > 40\ GeV$.

The table 7.8 compares the number of all events, the 1 track decay vertex events and 2 tracks decay vertex events between simulation and data at $P_{T,h} > 16\ GeV$, for the region $16 < P_{T,h} < 40\ GeV$ and at $P_{T,h} > 40\ GeV$. The observed and predicted number of events (listed in first column of table 7.8) agree within each region. This indicates that the normalization is correct. Furthermore, the data are well described by the simulation for the region $16 < P_{T,h} < 40\ GeV$ for the 1 track and 2 tracks decay vertex events. At

	All CC	1 track	2 tracks
$P_{T,h} > 16\ GeV$			
Data	805.4	70.3	14.4
MC all	789.7 ± 4.3	61.2 ± 1.2	7.4 ± 0.4
$16 < P_{T,h} < 40\ GeV$			
Data	488.9	45.2	5.4
MC all	474.4 ± 3.4	43.9 ± 1.0	5.6 ± 0.4
$P_{T,h} > 40\ GeV$			
Data	316.6	25.1	9.1
MC all	315.3 ± 2.8	17.3 ± 0.7	1.9 ± 0.2

Table 7.8: Comparison of the number of events for data with the simulation for the regions $P_{T,h} > 16\ GeV$, $16 < P_{T,h} < 40\ GeV$ and $P_{T,h} > 40\ GeV$. The simulation describes the data for the region $16 < P_{T,h} < 40\ GeV$, whereas the excess for the 1 track and 2 tracks decay vertex is present at $P_{T,h} > 40\ GeV$. The statistical errors for the MC prediction are given.

$P_{T,h} > 40\ GeV$ one sees an excess for the 1 track and 2 tracks decay vertex events in the data over the simulation prediction. The excess seen in table 7.4 is therefore only located at $P_{T,h} > 40\ GeV$.

Again (as in table 7.7) one checks how the distributions behave for positive or negative values for the two regions separately. The table 7.9 lists the results for the region $16 < P_{T,h} < 40\ GeV$. The negative and positive distributions for the data are well described by the simulation. For the region $16 < P_{T,h} < 40\ GeV$ the charm cross section in CC is similar to the MC program prediction.

The table 7.10 lists the results at $P_{T,h} > 40\ GeV$, where the excess is dominated by positive distributions. For example, at $P_{T,h} > 40\ GeV$ and positive decay length, seven events are seen, where the expectation is (1.1 ± 0.2) events. The probability for the MC program expectation to fluctuate to the observed number of events or more in this region is $1.5 \cdot 10^{-4}$. This is calculated with the equation

$$P(n \geq n_{obs}|\nu_b) = 1 - \sum_{n=0}^{n_{obs}-1} \frac{\nu_b^n}{n!} e^{-\nu_b},$$

where $\nu_b = 1.1$ and $n_{obs} = 7$ is used. If additionally one takes the statistical error $\delta\nu_b = 0.2$ into account the probability is $6.3 \cdot 10^{-4}$. The statistical error $\delta\nu_b$ on the mean number of expected events ν_b is taken into account by using the convolution:

$$P(n \geq n_{obs}|\nu_b, \delta\nu_b) = \int_0^{+\infty} dx\ G(x, \nu_b, \delta\nu_b)\ (1 - \sum_{n=0}^{n_{obs}-1} \frac{x^n}{n!} e^{-x}),$$

where $G(x; \nu_b, \delta\nu_b)$ is the probability density function for the MC expectation x, namely a Gaussian of mean value ν_b and width $\delta\nu_b$ [62].

Therefore, only the region $P_{T,h} > 40\ GeV$ is responsible that the measured charm cross section determined in section 7.2 is higher than predicted!

Test	Data	MC All	Charm	UDS
All	5.4	5.5 ± 0.4	2.5 ± 0.3	3.0 ± 0.3
$S_l < 0$	2.0	2.4 ± 0.2	0.8 ± 0.1	1.6 ± 0.2
$S_l > 0$	3.3	3.2 ± 0.3	1.8 ± 0.2	1.4 ± 0.2
$S_{d1} < 0$ && $S_{d2} < 0$	1.0	1.3 ± 0.2	0.3 ± 0.1	0.9 ± 0.2
$S_{d1} > 0$ && $S_{d2} > 0$	2.0	1.9 ± 0.2	1.2 ± 0.2	0.7 ± 0.1

Table 7.9: The measured and expected number of 2 tracks decay vertex events for the region $16 < P_{T,h} < 40\ GeV$: The negative and positive distributions are well described by the simulation.

Test	Data	MC All	Charm	UDS
All	9.1	1.8 ± 0.2	0.6 ± 0.1	1.2 ± 0.2
$S_l < 0$	2.0	0.8 ± 0.1	0.3 ± 0.1	0.5 ± 0.1
$S_l > 0$	7.0	1.1 ± 0.2	0.4 ± 0.1	0.7 ± 0.1
$S_{d1} < 0$ && $S_{d2} < 0$	1.0	0.5 ± 0.1	0.1 ± 0.1	0.3 ± 0.1
$S_{d1} > 0$ && $S_{d2} > 0$	3.0	0.7 ± 0.1	0.3 ± 0.1	0.4 ± 0.1

Table 7.10: The measured and expected number of 2 tracks decay vertex events at $P_{T,h} > 40\ GeV$: An excess exists in the data compared to the MC program prediction mainly for the positive distributions. E.g. for positive decay length seven events are observed, where the MC program prediction is 1.1 ± 0.2 events. Only statistical errors for the MC prediction are given.

7.2.5 Charm Extraction with the FVM Method in CC Processes

In this section the FVM method is briefly mentioned. Although the FVM method has its problems, especially for the determination of charm, as seen in the NC case, it will be used as a cross check of the VVM method. In the following, the cut based extraction and the fit procedure is described.

Cut based charm cross section with the FVM method in CC processes

In the cut based approach the goal is to get a charm enriched sample and this is done in an analogous manner as described in section 6.1.1. The advantage of the CC sample with respect to the NC sample is that the beauty contribution is negligible and so the problem to distinguish the heavy quarks is inexistent.

The selected events with the two tracks used in the 2-tracks method are compared between simulation and data in the appendix B in figure B.4. Within the errors the data distributions are described well by the simulation. In the 2-tracks method a simultaneous cut S_{d1}, S_{d2} and S_l increases the charm over light quark ratio. The first bin in figure 7.9(a) shows the sum of the events in the 2-tracks method which fulfill the cuts $0 < S_m < 10$ with $m = d1, d2, l$. The next bins give the sum of events when the lower bound of the cuts is increased. The

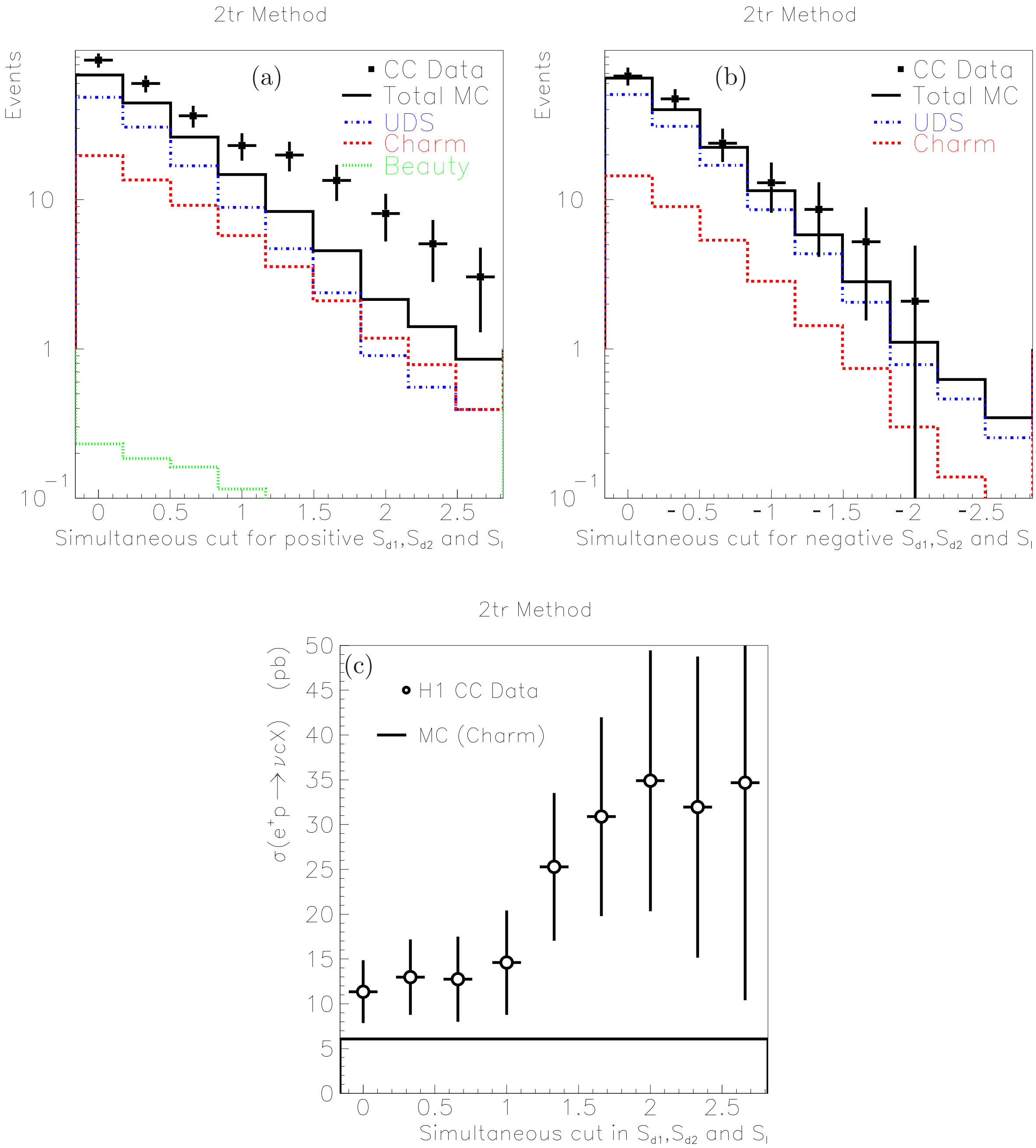


Figure 7.9: Composition of the light, charm and beauty contribution in the CC events predicted from the MC program selected with the FVM 2-tracks method: With a cut based approach in figure (a) and (b) the CC data are compared to the simulation. A simultaneous cut for positive S_{d1}, S_{d2} and S_l gives for the data events an excess over the simulation. No such excess is seen on the negative lifetime side. This evidences that the normalization and the track resolution between data and simulation are correct and suggests that in the MC program something which produces positive decay length is missing. In the figure (c) the charm cross section $\sigma(e^+p \rightarrow \bar{\nu}_e cX)$ in function of the cut values (open dots) is calculated. The solid line is the expected cross section predicted from the Django MC program. Only statistical errors are shown for the data points.

x-coordinate of the figure 7.9(a) gives the value of the lower bound of the cuts. A clear excess for the data compared to the simulation is seen. From the simulation one sees that only with cut values around two the charm exceeds the uds. However, only few events survive. The excess is not a consequence of the wrong normalization between data and simulation, as one sees in figure 7.9(b), where now all cuts used have a negative sign. E.g. the first bin represents the sum of the events which fulfill the cuts $-10 < S_m < 0$ with $m = d1, d2, l$ and the next bins give the sum of events when the upper bound of the cuts is decreased. In figure 7.9(b) the data points are consistent with the MC program prediction. However, even in this figure each data bin is systematically higher than the MC program prediction, but less pronounced as seen in the figure 7.9(a). Furthermore, the cuts with negative sign does not enrich charm and the charm contribution never exceeds the uds contribution.

The calculated charm cross section in CC is displayed in figure 7.9(c). The number of charm quarks N_c is calculated for each cut from the difference of the data events N_{Data} and the uds-events $N_{MC,uds}$ as predicted by the simulation, $N_c = N_{Data} - N_{MC,uds}$. The efficiencies for charm are taken from the simulation. The charm quark cross section σ_c is calculated as

$$\sigma_c = \frac{N_c}{\varepsilon_c \cdot \mathcal{L}}$$

in which the average efficiency $\varepsilon_c = \frac{N^c_{rec}}{N^c_{gen}}$ is defined in equation 3.2 and $\mathcal{L}$ denotes the luminosity. The measured charm cross section is for the first few bins two times higher than predicted by the Django MC program. One problem is that the charm cross section increases if the cut values increase. When the charm contribution is a significant fraction of all events (with cut values around 2) then the charm cross section is larger than the cross section predicted for all CC events which is $\sigma_{CC}(e^+p \to \bar{\nu}_e X) = 23.2\ pb$. It is clear that the efficiency determination from the simulation is wrong for charm and there must exist components in data which are not included in the MC program.

Charm cross section with a fit procedure in the FVM method

As seen in the NC sample the fit procedure in the FVM method has its problems especially for the charm cross section.

The fit procedure for the FVM method gives arbitrary results (see table 7.11), because the S_l and $min(S_{di})$ distributions between charm and uds are similar. Only the tails of these distributions are sensitive to discriminate the charm and uds, whereas the rest of the distribution is more or less identical. But in the tails there exist only few events, hence no relative ratio can be obtained with reasonable χ^2. This method is therefore not used for the determination of the cross section.

7.3 Results: Charm Production in Charged Current Processes

An excess of data over the MC program prediction is seen on the positive lifetime side and has been verified by different methods. Furthermore the excess is located at $P_{T,h} > 40\ GeV$, whereas the region $16 < P_{T,h} < 40\ GeV$ is well described by the simulation.

	FVM Method in CC	$\sigma(ep \to c)$ [pb]	$\sigma(ep \to uds)$ [pb]	$\chi^2/ndof$	$\rho_{c,uds}$
	3-tracks				
(a)	S_l	17.8 ± 5.4	5.6 ± 5.1	7/10	-0.96
	$min(S_{d1}, S_{d2}, S_{d3})$	5.6 ± 5.8	17.3 ± 5.6	15/15	-0.97
(b)	norm. & S_l	8.3 ± 2.9	14.7 ± 2.8	9/11	-0.96
	norm. & $min(S_{d1}, S_{d2}, S_{d3})$	-	-	-	-
	2-tracks				
(a)	S_l	16.2 ± 6.4	7.1 ± 3.9	7/11	-0.96
	$min(S_{d1}, S_{d2})$	16.7 ± 6.4	6.8 ± 6.1	5/10	-0.98
(b)	norm. & S_l	7.0 ± 2.5	15.9 ± 2.4	12/12	-0.94
	norm. & $min(S_{d1}, S_{d2})$	5.6 ± 2.6	17.3 ± 2.5	8/11	-0.95

Table 7.11: The integrated light quark and charm quark cross sections for the kinematic range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ with the FVM method. Only statistical errors are given. The fit does not converge for the 3-tracks method of the variable $min(S_{d1}, S_{d2}, S_{d3})$ with the normalization bin. This method is insensitive to charm and uds and is therefore not used for the determination of the cross sections.

If the excess is interpreted as charm contribution, then the measured charm production cross section in CC processes, integrated over the range $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$, is found to be:

$$\sigma(e^+p \to \bar{\nu}_e\ c\ X) = 17.7 \pm 5.5 \pm 2.3\ pb.$$

Indeed it means that the uds quark cross section in CC is:

$$\sigma(e^+p \to \bar{\nu}_e\ uds\ X) = 5.8 \pm 7.8 \pm 1.1\ pb.$$

The Django MC generator predicts a cross sections, of $\sigma_c = 6.1\ pb$ and $\sigma_{uds} = 17.1\ pb$, using leading order parton density function CTEQ5L.

Although there is an indication of an excess in the data compared to the MC program prediction, within the large statistical and systematic errors the predictions are still compatible with the measurements.

Chapter 8

Conclusions

The production of charm quarks and beauty quarks in neutral current processes and the production of charm quarks in charged current processes at high Q^2 is measured.

This work provided an independent cross-check of another H1 analysis in the neutral current processes and as such served as supporting evidence for the publication [1]. The techniques applied rely on the precise tracking information from the vertex detector of H1 at HERA. The same technique, now proven to be applicable, has been used to determine for the first time the production cross section of charm in charged current processes at HERA. An integrated luminosity of $\mathcal{L} = (56.9 \pm 0.85)\ pb^{-1}$ for e^+p scattering has been used.

Due to the long lifetime of the heavy hadrons, the inclusive heavy quark cross sections are measured with a lifetime tagging technique. The technique relies on the spatial separation of the primary and the secondary vertex, exploiting the two dimensional vertexing capability of the Central Silicon detector. Using Monte Carlo (MC) templates of lifetime related variables for the heavy and the light quark distributions, the light, charm and beauty fractions of the data are extracted with a fit procedure.

Two new inclusive approaches are tested, namely, the fixed vertex multiplicity (FVM) method and the variable vertex multiplicity (VVM) method, which give both consistent results. The VVM method has smaller systematic and similar statistical errors when compared to the FVM method, and thus the VVM method is chosen to extract the final results.

The charm and beauty cross sections in neutral current are measured for the kinematic region $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$. The results are

$$\sigma(e^+p \to e'^+c\bar{c}X) = 476 \pm 60 \pm 76\ pb$$

$$\sigma(e^+p \to e'^+b\bar{b}X) = 38 \pm 6 \pm 8\ pb.$$

The prediction from perturbative quantum chromodynamics at next-to-leading order, the MC prediction of the generator Django (heavy quarks are massless) and Rapgap (heavy quark masses are taken into account) agree with the measured integrated cross sections within errors.

The VVM method is used in the CC sample only after the confirmation, that it produces reliable results in the NC-metamorphosed sample. This reduced NC sample has the same kinematic behaviour as the CC sample, however, the results are known and the statistics are higher than in the CC sample.

In the CC data an excess over the MC prediction of the Django generator is seen on the positive lifetime side and has been verified with different methods. Furthermore the excess is located at $P_{T,h} > 40\ GeV$, whereas the region $16 < P_{T,h} < 40\ GeV$ is well described by the simulation.

If the excess is interpreted as being due to charm, then the charm production cross section can be measured in charged current, within the kinematic region $Q^2 > 284\ GeV^2$ and $0.1 < y < 0.7$. Using the Django MC program to extrapolate to $Q^2 > 150\ GeV^2$ and $0.1 < y < 0.7$ (a correction of order 10%), then the cross section is found to be

$$\sigma(e^+p \to \bar{\nu}_e\ c\ X) = 17.7 \pm 5.5 \pm 2.3\ pb.$$

This means that the uds quark cross section in CC is:

$$\sigma(e^+p \to \bar{\nu}_e\ uds\ X) = 5.8 \pm 7.8 \pm 1.1\ pb.$$

Within the large statistical and systematic errors the predictions are compatible with the measured results.

This is the first observation of charm production in charged current scattering at HERA. Only a substantial increase of statistics will permit a measurement of the strange and anti-strange sea in the proton.

Appendix A

Systematic Errors in NC Processes

The procedure to evaluate the systematic errors for the integrated charm and beauty cross sections in NC processes is explained in section 6.2.2. The systematic errors are evaluated for:

- the FVM 3-tracks method for the variable S_l in table A.1 and for the variable $min(S_{d1}, S_{d2}, S_{d3})$ in table A.2.
- the FVM 2-tracks method for the variable S_l in table A.3 and for the variable $min(S_{d1}, S_{d2})$ in table A.4.
- the VVM method for the variable $min(S_{di})$ in table A.5.

The systematic error results from the FVM method are used in the table 6.2 and the systematic error results from the VVM method are used in the table 6.5.

S_l 3-tracks	Default	50 ± 8 pb	462 ± 42 pb	2
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$	fit prob. (%)
Resolution	1.09 → 1.22	18	33	67
Efficiency	−3% per track	6	11	0
Model	Django →Rapgap	18	1	2
Fragmentation	Lund → Peterson	10	5	7
B multiplicity	5.2 → 5.0	3	3	2
D^+ fraction	0.197 → 0.232	5	1	2
E(jet) scale	10 %	1	1	2
E(e) scale	5 %	0	1	1
Luminosity		1.5	1.5	
Total systematic error		29	35	
Statistic error		16	9	

Table A.1: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_b (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable S_l in the FVM method with the 3-tracks method. In the last column the fit probability is given.

$min(S_{d1}, S_{d2}, S_{d3})$ 3-tracks	Default	46 ± 7 pb	382 ± 55 pb	52
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$	fit prob. (%)
Resolution	1.09 → 1.22	2	36	0
Efficiency	−3% per track	13	18	1
Model	Django →Rapgap	7	1	41
Fragmentation	Lund → Peterson	9	2	77
B multiplicity	5.2 → 5.0	5	2	52
D^+ fraction	0.197 → 0.232	2	2	42
E(jet) scale	10 %	1	1	42
E(e) scale	5 %	4	2	46
Luminosity		1.5	1.5	
Total systematic error		19	40	
Statistic error		15	14	

Table A.2: The summary of the experimental systematic: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_b (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable $min(S_{d1}, S_{d2}, S_{d3})$ in the FVM method with the 3-tracks method. In the last column the fit probability is given.

S_l 2-tracks	Default	$61 \pm 10\ pb$	$470 \pm 38\ pb$	0
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$	fit prob. (%)
Resolution	1.09 $\rightarrow$ 1.22	30	40	3
Efficiency	-3% per track	3	11	0
Model	Django $\rightarrow$Rapgap	8	2	0
Fragmentation	Lund $\rightarrow$ Peterson	9	1	0
B multiplicity	5.2 $\rightarrow$ 5.0	2	1	0
D^+ fraction	0.197 $\rightarrow$ 0.232	5	2	0
E(jet) scale	10 %	2	2	0
E(e) scale	5 %	3	3	0
Luminosity		1.5	1.5	
Total systematic error		33	42	
Statistic error		16	8	

Table A.3: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_b (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable S_l in the FVM method with the 2-tracks method. In the last column the fit probability is given.

$min(S_{d1}, S_{d2})$ 2-tracks	Default	$40 \pm 6\ pb$	$469 \pm 43\ pb$	0
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$	fit prob. (%)
Resolution	1.09 $\rightarrow$ 1.22	20	35	0
Efficiency	-3% per track	5	19	0
Model	Django $\rightarrow$Rapgap	5	4	0
Fragmentation	Lund $\rightarrow$ Peterson	2	1	0
B multiplicity	5.2 $\rightarrow$ 5.0	8	1	0
D^+ fraction	0.197 $\rightarrow$ 0.232	5	1	0
E(jet) scale	10 %	0	0	0
E(e) scale	5 %	3	1	0
Luminosity		1.5	1.5	
Total systematic error		23	40	
Statistic error		15	9	

Table A.4: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_b (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable $min(S_{d1}, S_{d2})$ in the FVM method with the 2-tracks method. In the last column the fit probability is given.

$min(S_{di})$ with 2 & ≥ 3 tracks	Default	38 ± 6 pb	476 ± 60 pb	45
Systematic		$\delta(b)(\%)$	$\delta(c)(\%)$	fit prob. (%)
Resolution	$1.09 \rightarrow 1.22$	5	4	33
Efficiency	-3% per track	16	2	76
Model	Django $\rightarrow$Rapgap	8	12	43
Fragmentation	Lund $\rightarrow$ Peterson	12	9	70
B multiplicity	$5.2 \rightarrow 5.0$	1	3	32
D^+ fraction	$0.197 \rightarrow 0.232$	2	1	43
E(jet) scale	10 %	0	2	49
E(e) scale	5 %	2	1	47
Luminosity		1.5	1.5	
Total systematic error		22	16	
Statistic error		16	13	

Table A.5: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_b (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable $min(S_{di})$ in the VVM method with the simultaneous fit of the 2 & ≥ 3 tracks histograms. In the last column the fit probability is given.

Appendix B

Track, Primary Vertex and Jet Reconstruction in CC Processes

In the following the CST-improved tracks, the primary vertex and the jets control-plots as used for the NC events are shown for the CC events. In figure B.4 the properties of the tagging tracks from the FVM 2-tracks method between CC data and simulation are compared. Then the jet (see figure B.5) and tagging related variables (see figure B.6) are compared for the CC sample and the NC sample before and after the $P_{T,h} > 16\ GeV$ cut. The figure B.7 presents the last of the ten low statistical NC-metamorphosed samples obtained as explained in section 7.1.4.

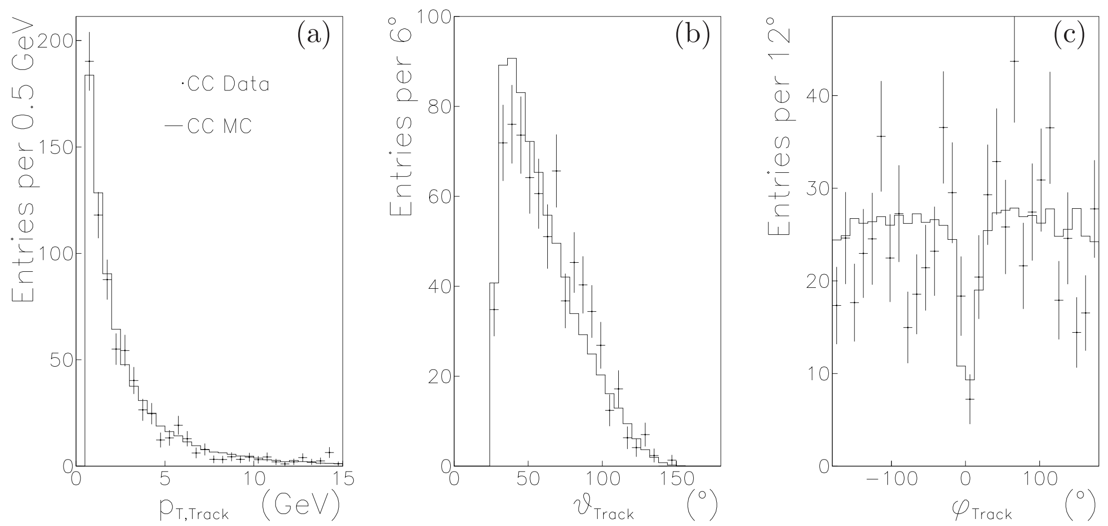


Figure B.1: (a) The transverse momentum distribution, (b) the polar angle distribution and (c) the azimuth angle distribution of CST-improved tracks which are selected for the secondary vertexing in the 2-tracks method in the CC sample. Data (dots) with statistic error are well described by simulation (histogram). The MC is normalized to the data luminosity.

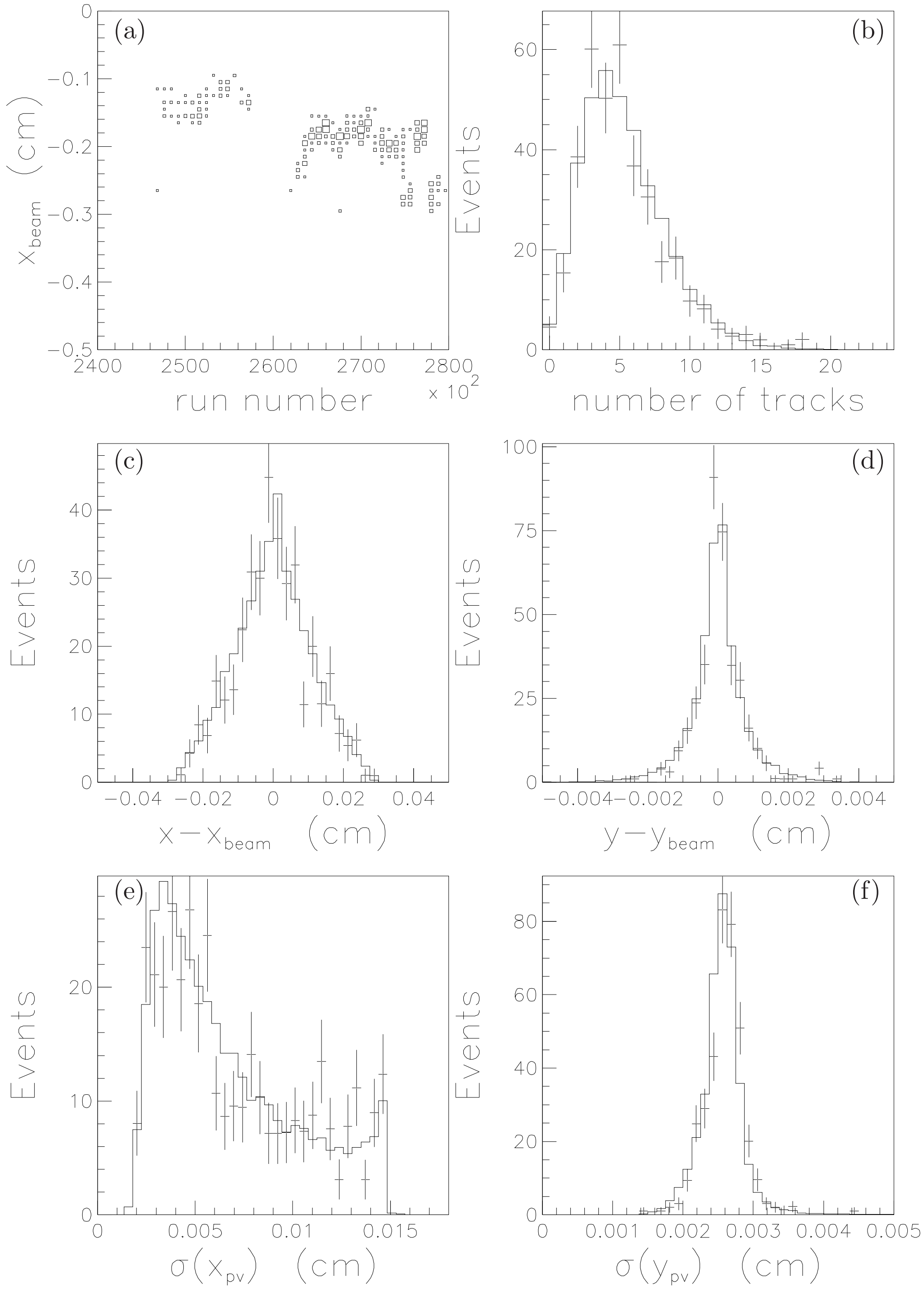


Figure B.2: Different primary vertex informations for the selected CC candidates with the 2-tracks method are presented. Figure (a) shows for data only the movement of the beam in the x-direction as a function of the run number. (b) compares for simulation and data the number of tracks used for the two dimensional PV fit. In (c) and (d) the difference of the calculated PV and the position of the beam spot measured as the average over some hundred events is displayed. The calculated error distributions for the PV coordinates in the (e) x and the (f) y directions are compared for simulation and data.

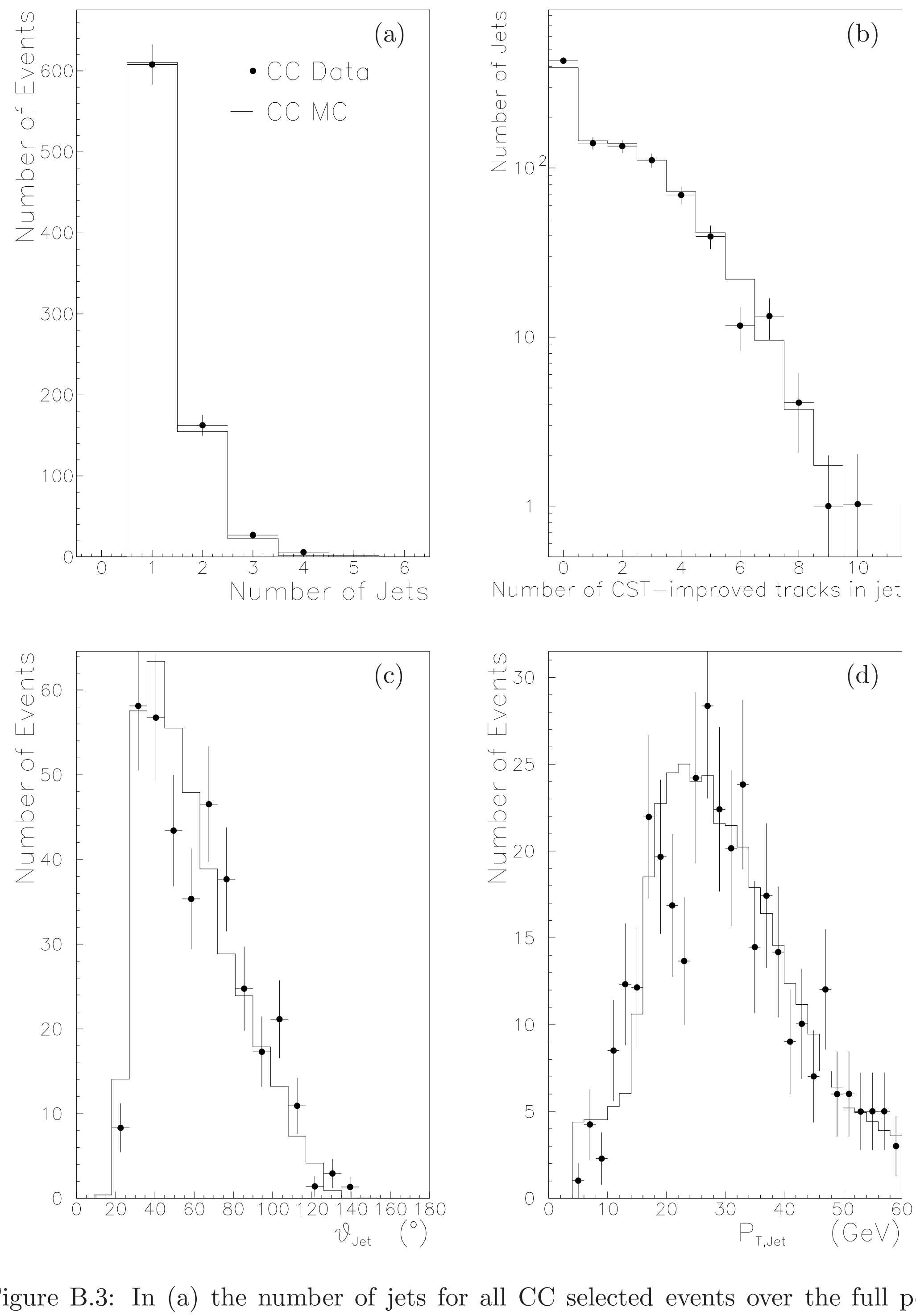


Figure B.3: In (a) the number of jets for all CC selected events over the full polar range and $p_{T,Jet} > 4\ GeV$ are compared between data and simulation. In (b) the number of CST-improved tracks inside the jet in the angular range $10^o < \vartheta_{Jet} < 170^o$ are shown. In (c) the polar angle ϑ_{Jet} and in (d) the transverse momentum $p_{T,Jet}$ distribution of the jets which contain at least two CST-improved tracks within a cone of radius 1 in the $\eta - \phi$ plane are displayed.

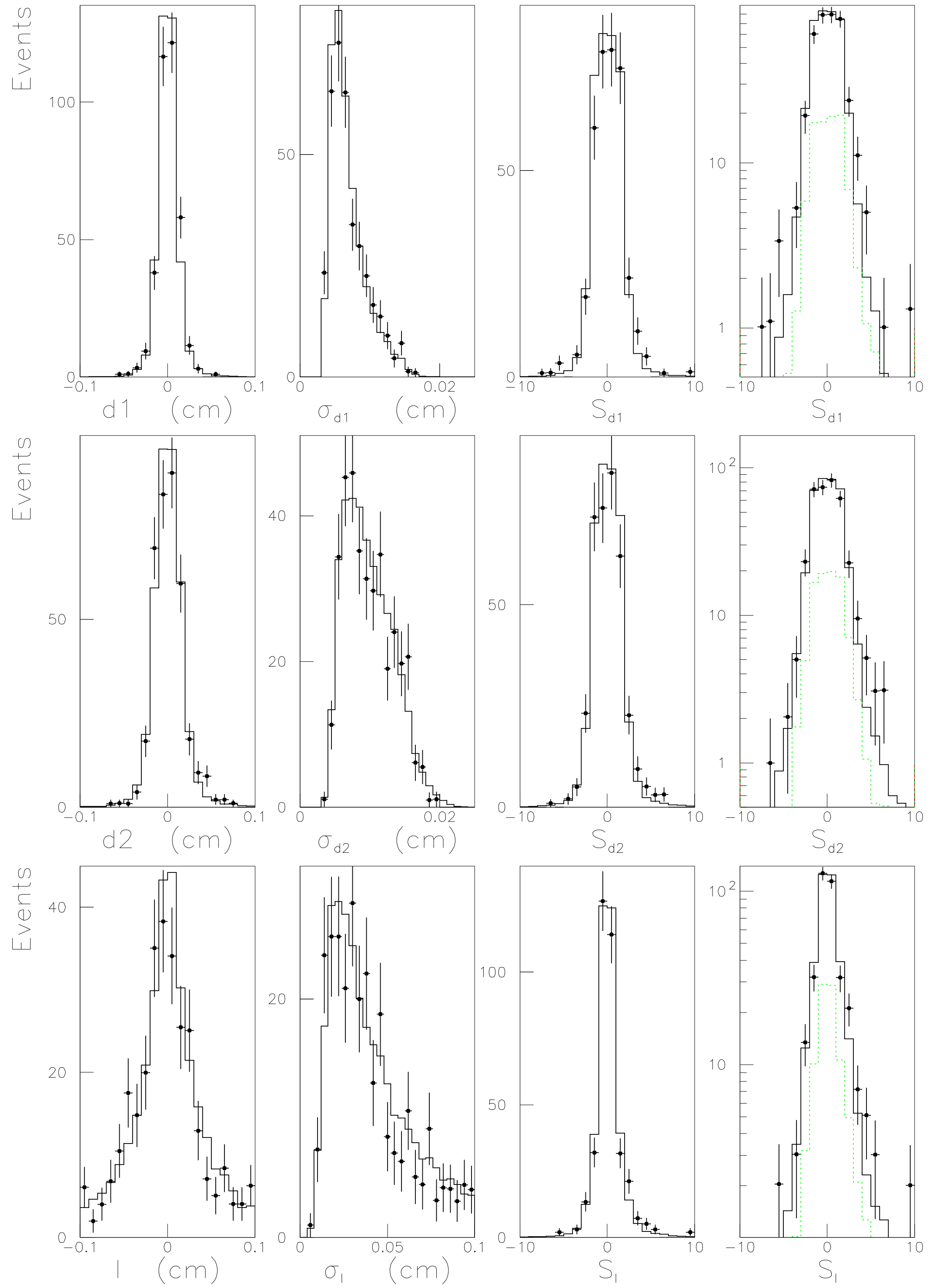


Figure B.4: The selected events with the two tracks used for the 2-tracks method. The simulation (solid histogram) is normalized to data luminosity (points with statistical errors). In the first (second) row the impact parameter, the error of the impact parameter and its significance in linear and logarithmic scale of the two tracks is displayed. The tracks are sorted in transverse momentum, where the fastest is upmost. In the last row the decay length, error of the decay length and decay length significance in linear and logarithmic scale is shown. In the last column the charm (green, fine dashed) component as predicted by the simulation is shown.

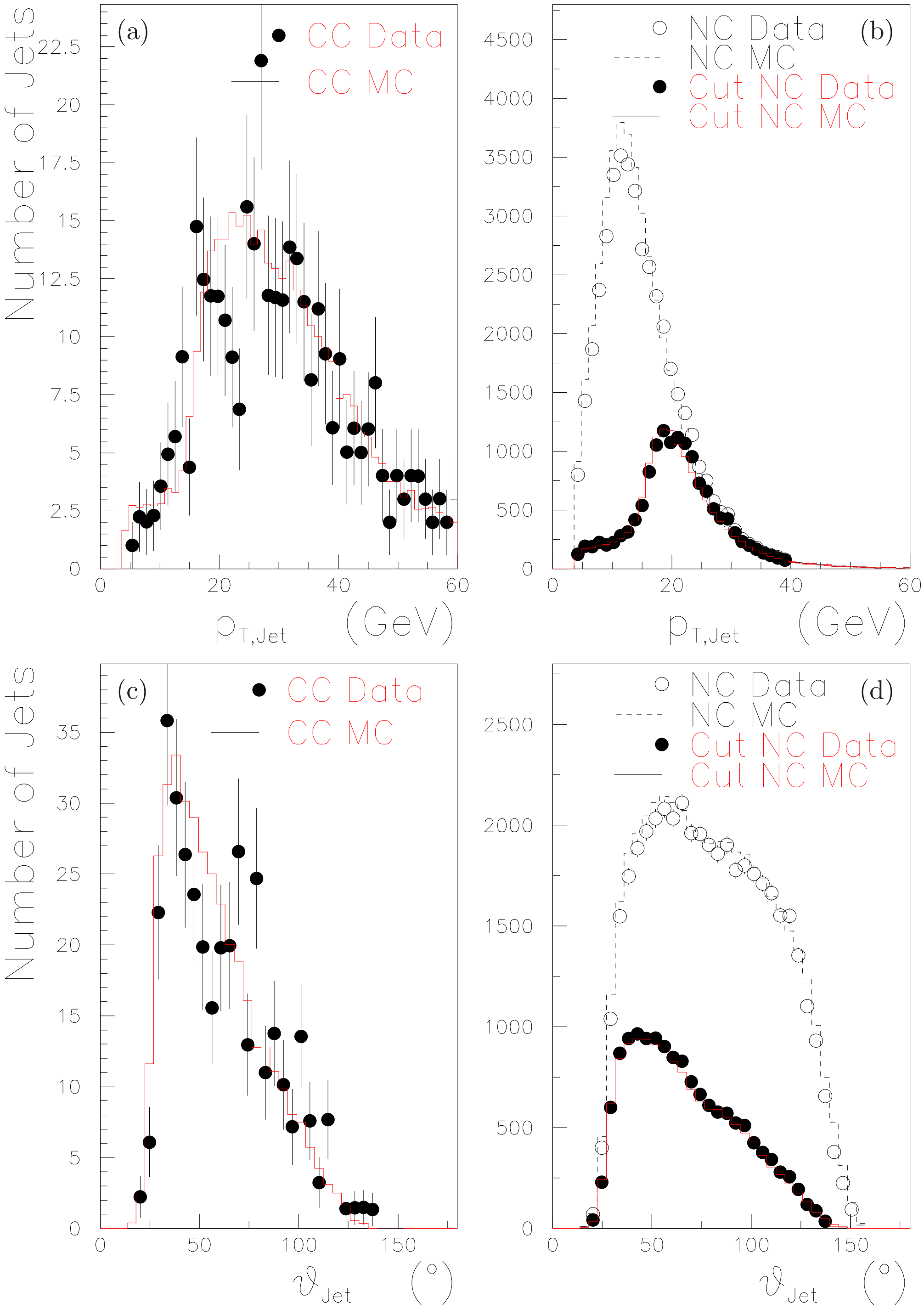


Figure B.5: Comparison of CC and NC jet information: On the left (right) side the CC (NC) events for data and simulation are presented. In figure (b) the $p_{T,Jet}$ for the 'standard NC' events (empty circle) and for the NC events after a $P_{T,h} > 16\ GeV$ (full circle) cut are displayed. This cut produce a similar distribution as for the CC events, which is shown in figure (a). Similar behaviour is seen for the ϑ_{Jet} distribution in the two lower figures.

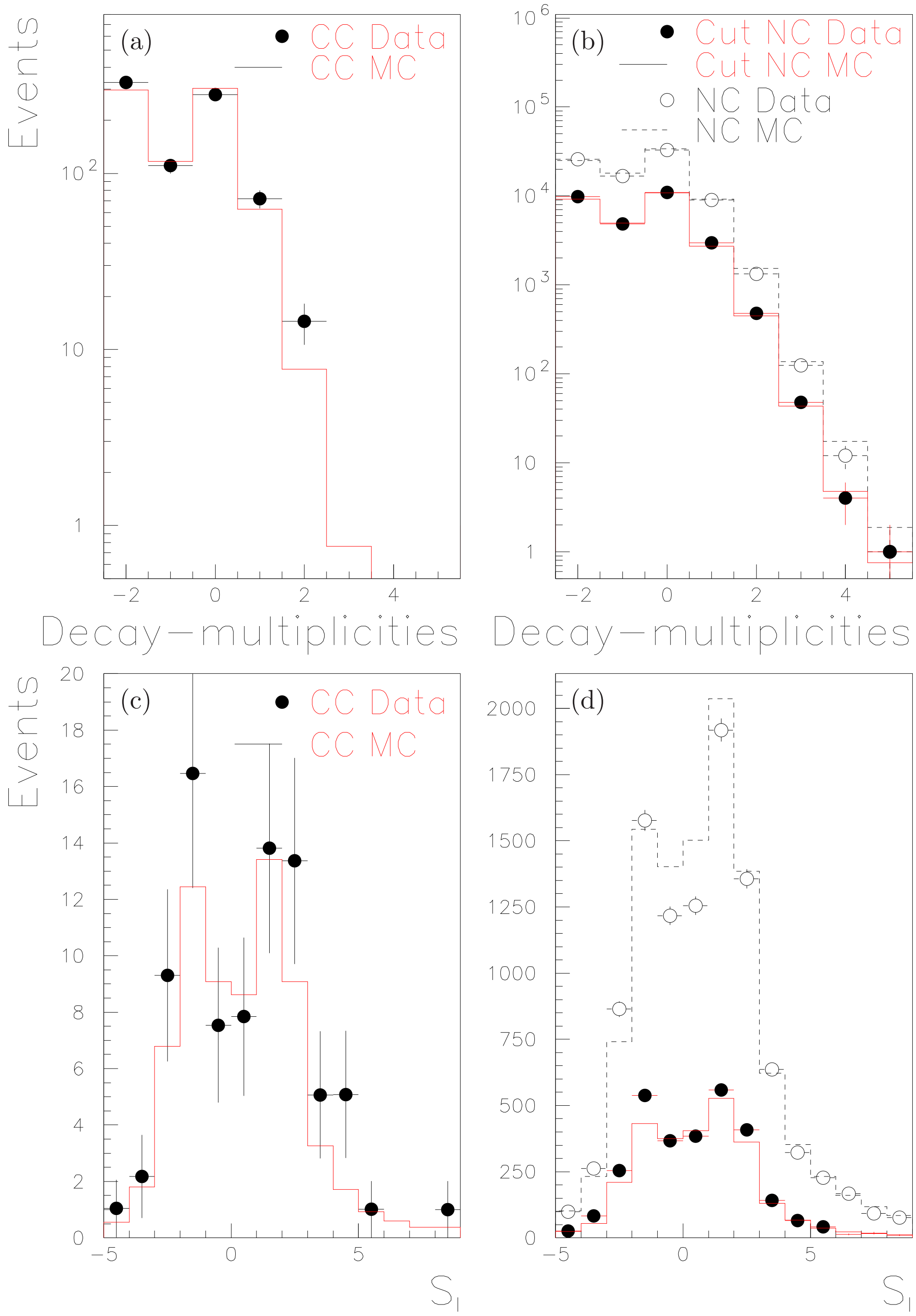


Figure B.6: Comparison of CC and NC heavy quark tagging information: On the left (right) side the CC (NC) events for data and simulation are presented. In the upper (lower) two figures the decay-multiplicity (S_l for more than one track) for the 'standard NC' events (empty circle) and for the NC events after a $P_{T,h} > 16\ GeV$ (full circle) cut are displayed. These distributions do not change much for the NC-metamorphosed and they look similar to the CC sample.

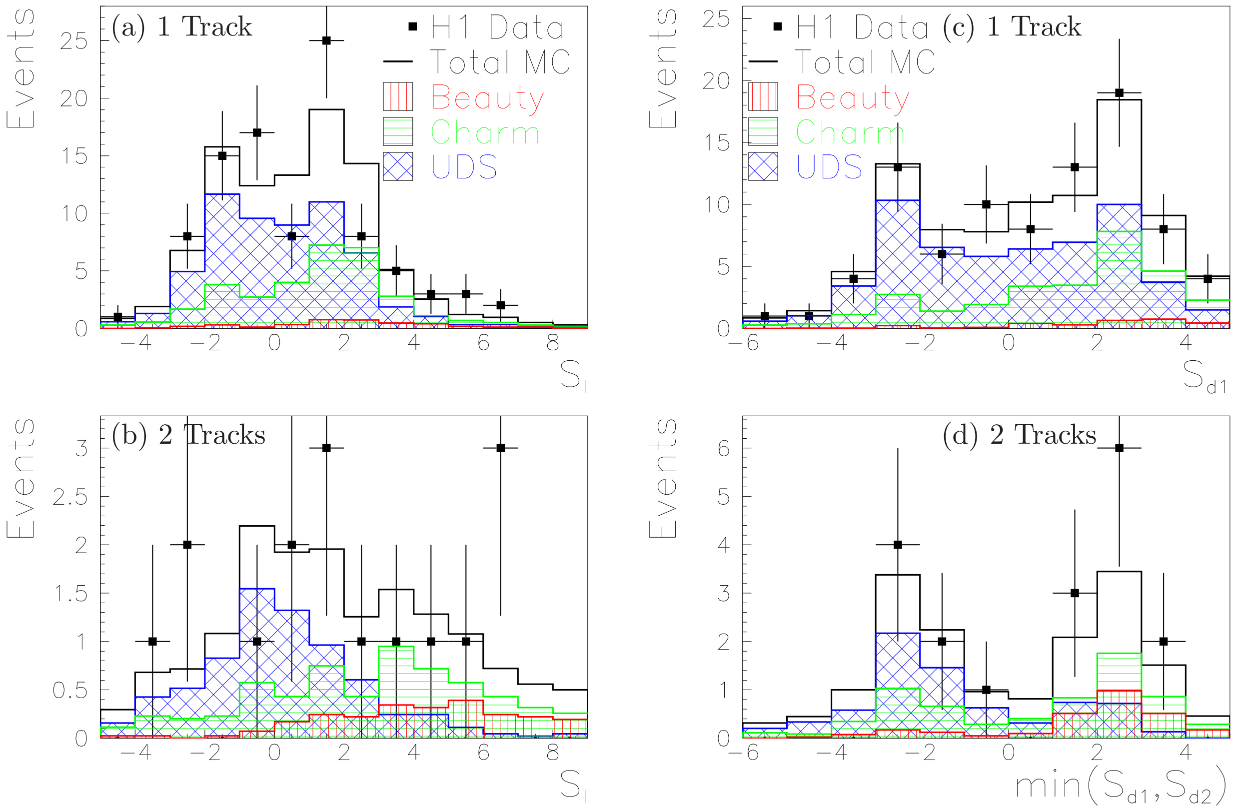


Figure B.7: The VVM method for 1 & $\geq$ 2 tracks in the NC-metamorphosed sample with fixed beauty parameter and low statistic: The figure shows the variable S_l (left) and $min(S_{di})$ (right) for the same events with 2 or more tracks (b, d) and 1 track (a, c) associated with the secondary vertex. So many candidates are discarded to get a similar statistic as in the CC sample. The last of the ten tests is shown. Superimposed on the data points with statistical error only are the light, charm and beauty quark contributions as obtained from the fit with the Django MC templates.

Appendix C

Systematic Errors in CC Processes

The systematic errors for the CC processes are determined in the NC-metamorphosed sample to avoid statistical fluctuations. The stability of the algorithm is tested only with the dominant contribution of the systematic uncertainties seen in NC (see section 6.2.2). The systematic errors are evaluated for the VVM method:

- without normalization bin for the variable $min(S_{di})$ in table C.1.
- with normalization bin for the variable S_l in table C.2.
- with normalization bin for the variable $min(S_{di})$ in the table C.3.

These results are used for the systematic errors of the integrated charm and light cross sections in the table 7.6.

$min(S_{di})$ with 1 & ≥ 2 tracks	Default	$2126 \pm 81\ pb$	$447 \pm 42 pb$	92/40
Systematic		$\delta(uds)(\%)$	$\delta(c)(\%)$	$\chi^2/ndof$
Resolution	$1.09 \rightarrow 1.22$	11	1	57/39
Efficiency	-3% per track	14	2	39/39
Model	Django $\rightarrow$Rapgap	2	9	63/39
Fragmentation	Lund $\rightarrow$ Peterson	3	5	63/39
Total systematic error		18	11	

Table C.1: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_{uds} (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable $min(S_{di})$ with the VVM method with the simultaneous fit of the 1 & ≥ 2 tracks histograms in the NC-metamorphosed sample. This total systematic error is used for the CC result in table 7.6.

norm. & S_l with 1 & ≥ 2 tracks	Default	1810 ± 29 pb	$577 \pm 33pb$	100/58
Systematic		$\delta(uds)(\%)$	$\delta(c)(\%)$	$\chi^2/ndof$
Resolution	$1.09 \rightarrow 1.22$	7	26	69/56
Efficiency	-3% per track	7	27	114/58
Model	Django $\rightarrow$Rapgap	3	4	95/58
Fragmentation	Lund $\rightarrow$ Peterson	2	5	87/58
Total systematic error		11	38	

Table C.2: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_{uds} (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable S_l with the VVM method with the simultaneous fit of the normalization bin, the 1 & ≥ 2 tracks histograms in the NC-metamorphosed sample. This total systematic error is used for the CC result in table 7.6.

norm. & $min(S_{di})$ with 1 & ≥ 2 tracks	Default	1825 ± 27 pb	$559 \pm 32pb$	67/43
Systematic		$\delta(uds)(\%)$	$\delta(c)(\%)$	$\chi^2/ndof$
Resolution	$1.09 \rightarrow 1.22$	6	23	40/40
Efficiency	-3% per track	6	27	115/40
Model	Django $\rightarrow$Rapgap	2	8	84/40
Fragmentation	Lund $\rightarrow$ Peterson	1	2	66/40
Total systematic error		9	36	

Table C.3: The summary of the experimental systematic uncertainty: Shown are the variations of parameters (2^{nd} col.) and the effects on σ_{uds} (3^{rd} col.) and σ_c (4^{th} col.). All tests are done for the variable $min(S_{di})$ with the VVM method with the simultaneous fit of the normalization bin, the 1 & ≥ 2 tracks histograms in the NC-metamorphosed sample. This total systematic error is used for the CC result in table 7.6.

Appendix D

Pictures of CC Processes

The following figures visualize the remaining 13 candidates with a 2 tracks decay vertex found in the CC sample. One candidate is presented in figure 7.7.

For all events the two selected intersecting tracks and the primary and secondary vertex with a one sigma error ellipses are displayed. The arrow represents the direction of the reconstructed transverse momentum of the decaying objects. Furthermore the H1 detector with the calorimeter information is shown in the side and radial view.

In the table D.1 all useful information for the CC events with 2 tracks decay vertex are listed.

Run	Event	l $[cm]$	Mass $[GeV]$	Charge	$d1$ $[cm]$	$d2$ $[cm]$	$P_{T,miss}$ $[GeV]$	x_h
262397	254359	0.035	1.35	0	-0.0075	0.0164	59.3	0.158
263337	17795	-0.081	0.79	0	0.0175	-0.0315	49.2	0.096
266336	4126	0.078	1.54	0	0.0135	0.0250	56.7	0.177
268001	8957	-0.211	0.32	2	-0.0144	-0.0123	45.6	0.082
268336	36823	-0.143	1.03	-2	0.0072	-0.0202	33.8	0.055
272280	52700	0.117	2.83	2	0.0199	0.0417	50.5	0.133
274922	125441	0.058	1.20	0	-0.0156	0.0322	49.6	0.100
276165	214817	0.029	1.86	0	-0.0101	0.0192	45.7	0.115
276429	13318	0.231	0.66	0	0.0316	0.0075	37.3	0.058
277355	64022	0.052	0.61	0	-0.0132	0.0258	45.4	0.090
277958	125695	0.180	1.08	0	0.0125	0.0160	40.2	0.064
278082	102056	0.027	1.63	-2	-0.0129	0.0290	17.7	0.017
278375	14245	0.161	0.89	2	0.0224	0.0693	24.1	0.050
279139	12574	-0.153	0.78	0	-0.0251	-0.0137	29.2	0.037

Table D.1: The 14 CC events with 2 tracks decay vertex are introduced: The run, event, decay length, invariant mass, total charges, impact parameter one, impact parameter two, missing transversal momentum and the variable x_h are listed. All events have a weight of around one, apart from run 278082 with the event-number 102056 has a weight of 1.3.

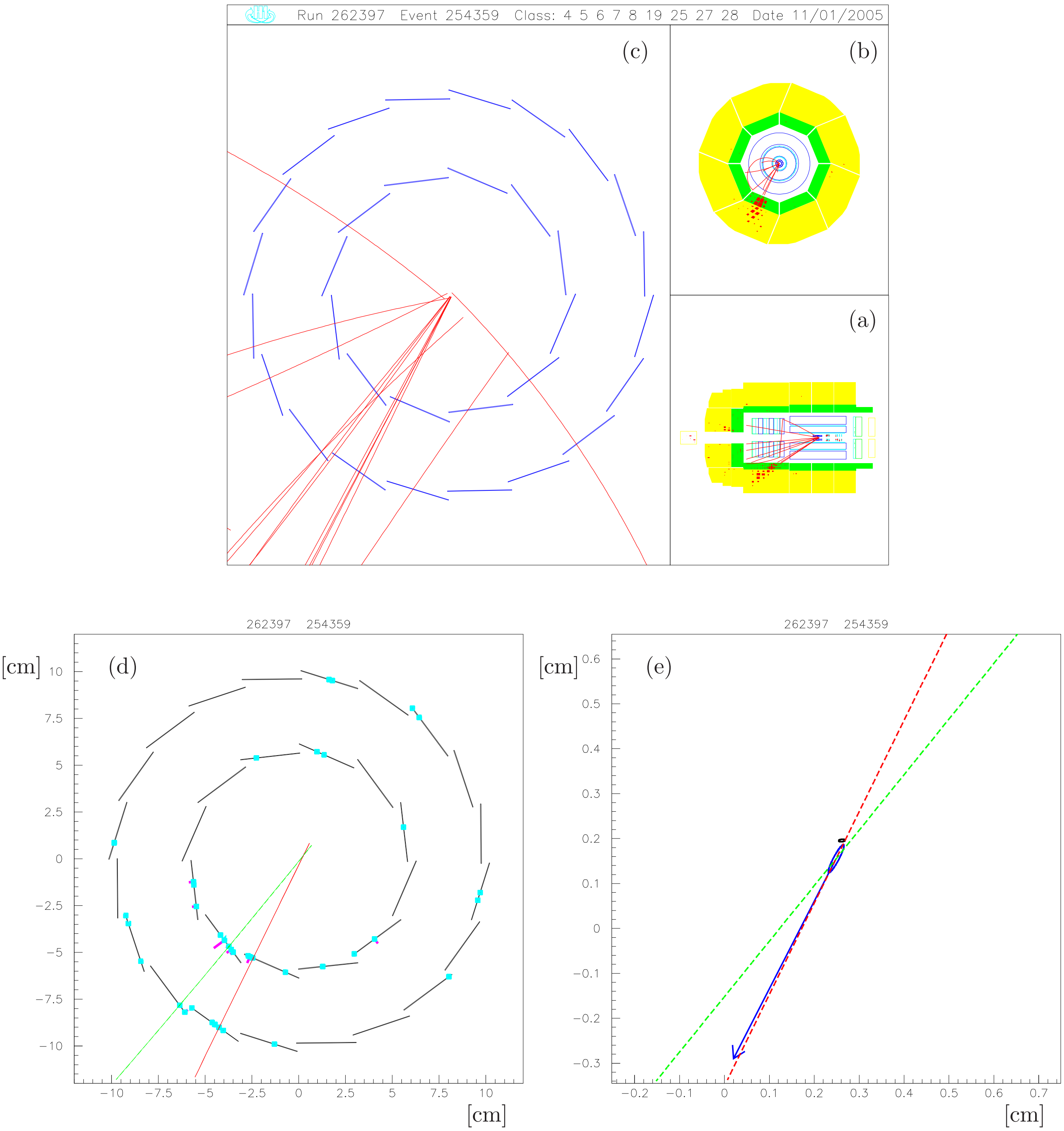
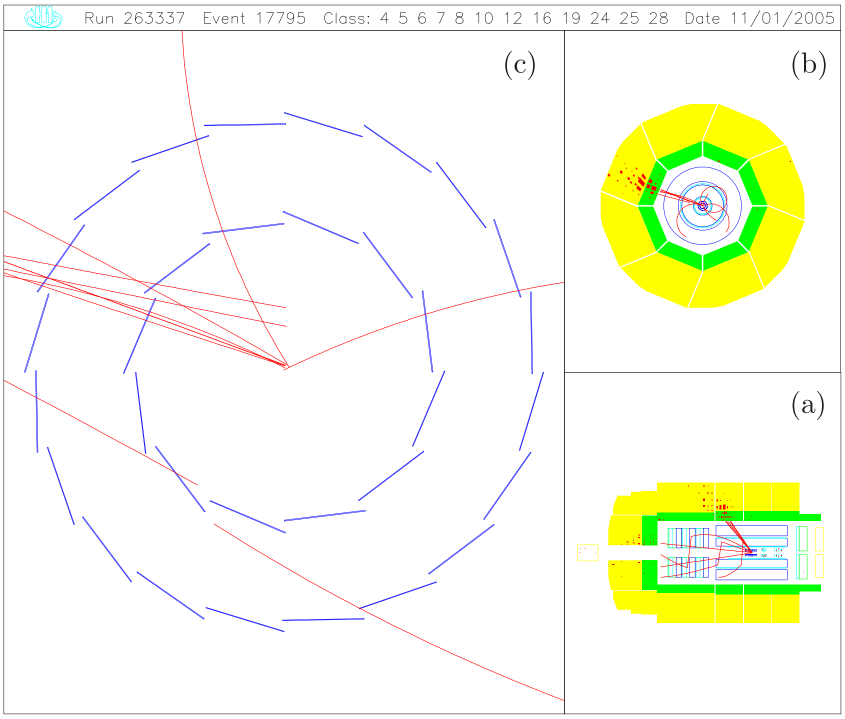


Figure D.1: A CC event with a 2 tracks decay vertex selected with the VVM method: A CC event measured by the H1 detector in the side view (a) and in the $r\phi$ view (b) is displayed. With around 25-times magnified of the center of figure (b) is presented in (c) a schematic $r\phi$ view of the CST with all reconstructed charged tracks from the event. In (d) the CST including hits (blue points) and the two selected CST-improved tracks (lines) which produce the 2 tracks decay vertex is visualized. Again in a closer view, allowing the different scale on the axis, one sees in (e) the decay vertex of the two intersecting tracks, where the one sigma error of the secondary vertex is the ellipse. The arrow represents the direction of the reconstructed transverse momentum of the decaying objects. In opposite direction to the arrow the primary vertex position with a one sigma error (black ellipse) is displayed. The measured radial decay length is $L_{xy} = 0.035\ cm$ and the reconstructed invariant mass is $1.35\ GeV$.



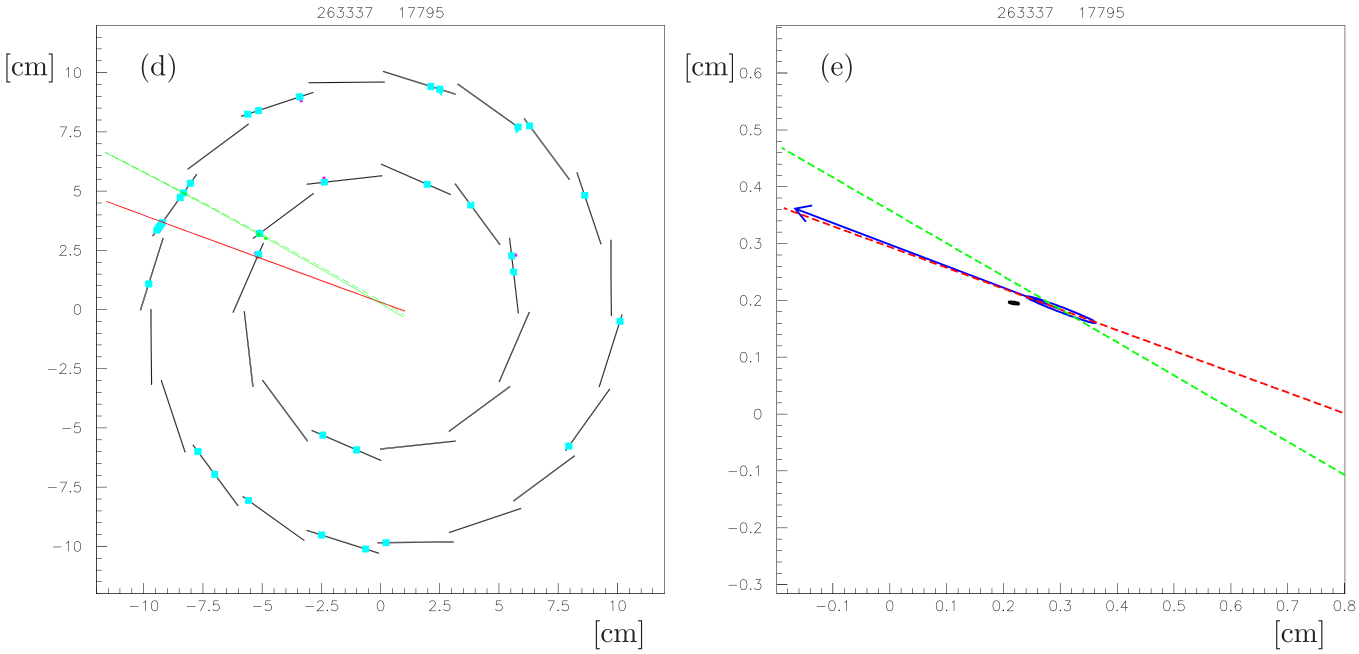


Figure D.2: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = -0.081\ cm$ and the reconstructed invariant mass is $0.79\ GeV$. For further explanations see figure D.1.

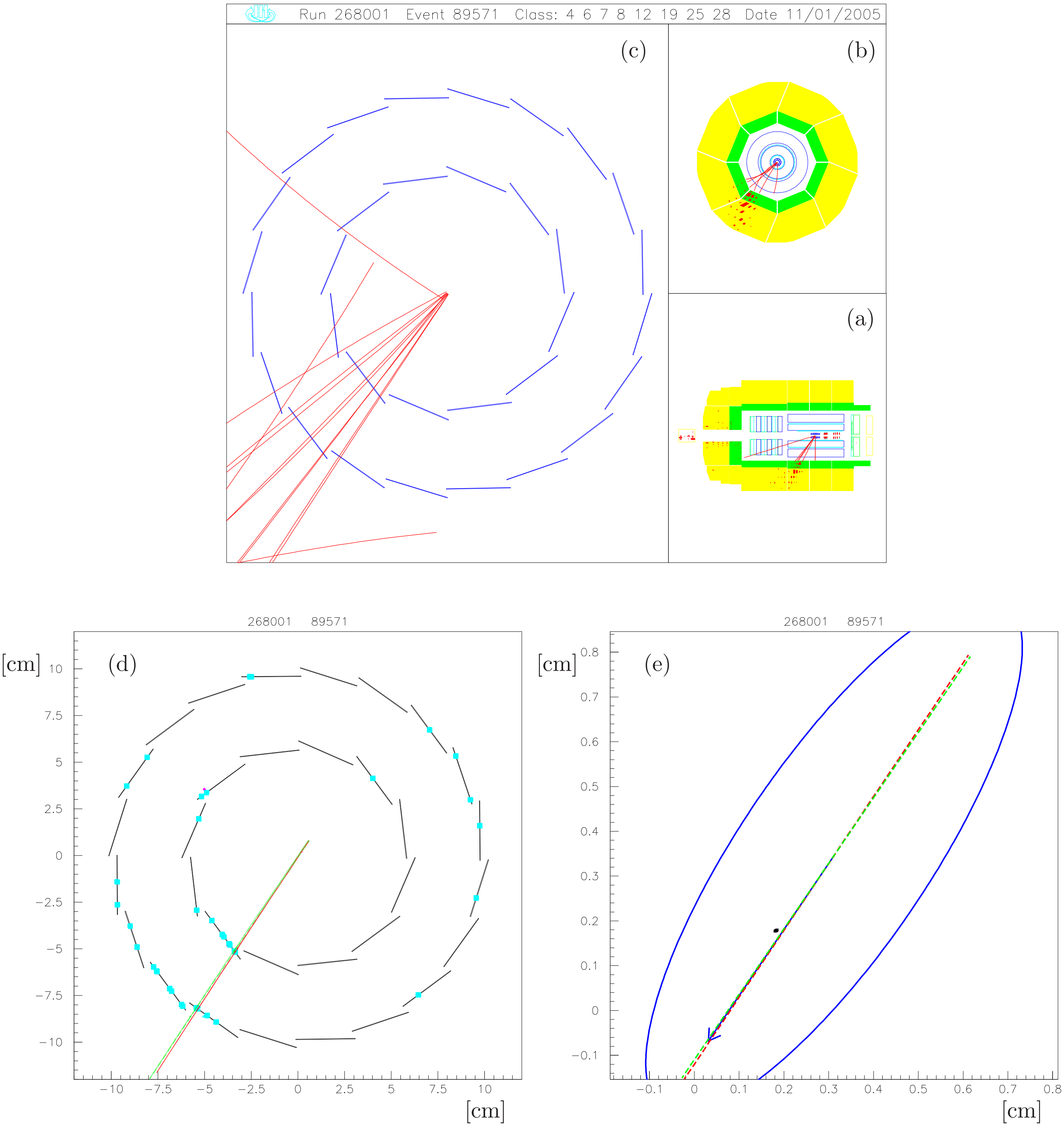


Figure D.3: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = -0.211\ cm$ and the reconstructed invariant mass is $0.32\ GeV$. For further explanations see figure D.1.

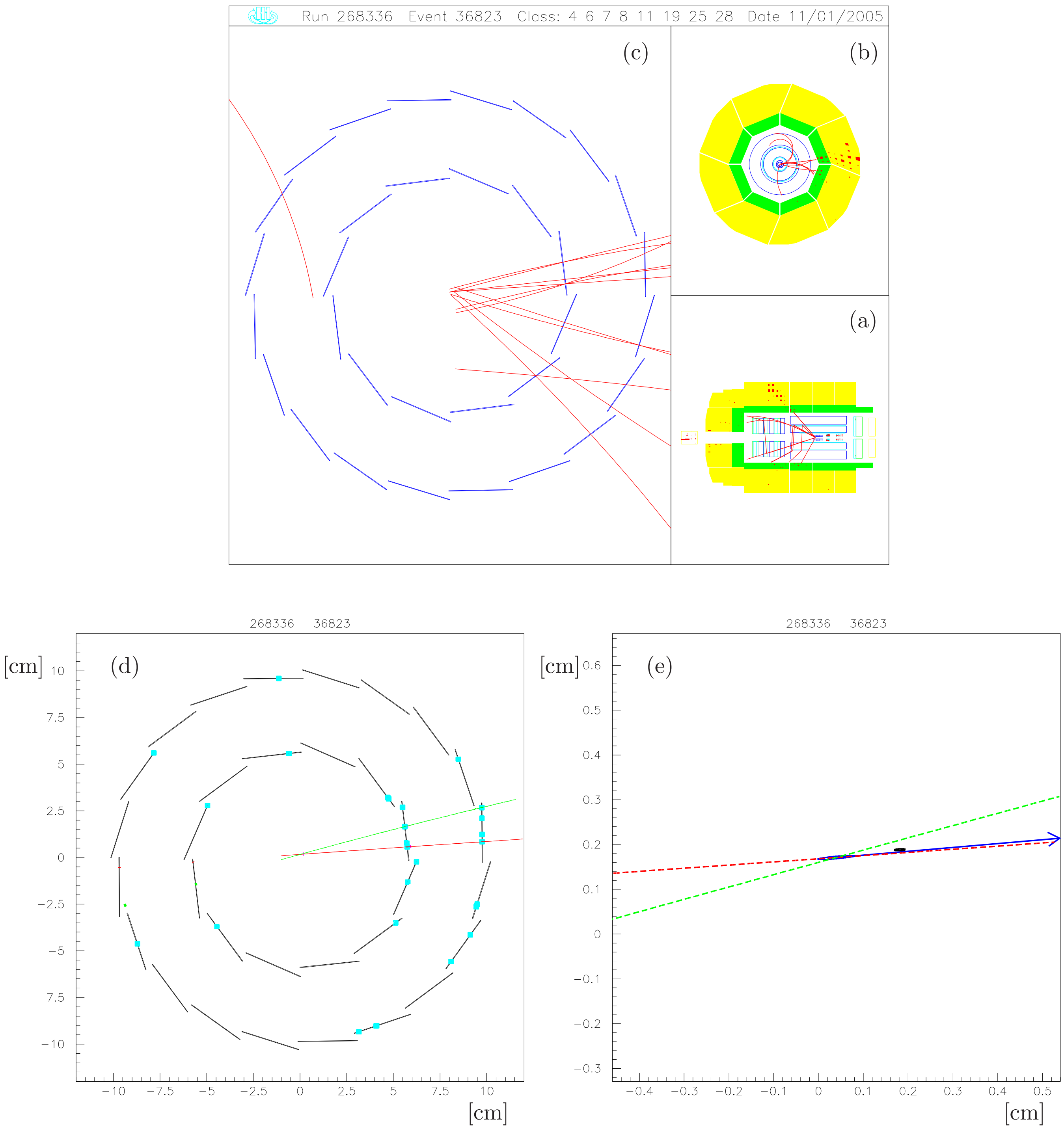


Figure D.4: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = -0.143\ cm$ and the reconstructed invariant mass is $1.03\ GeV$. For further explanations see figure D.1.

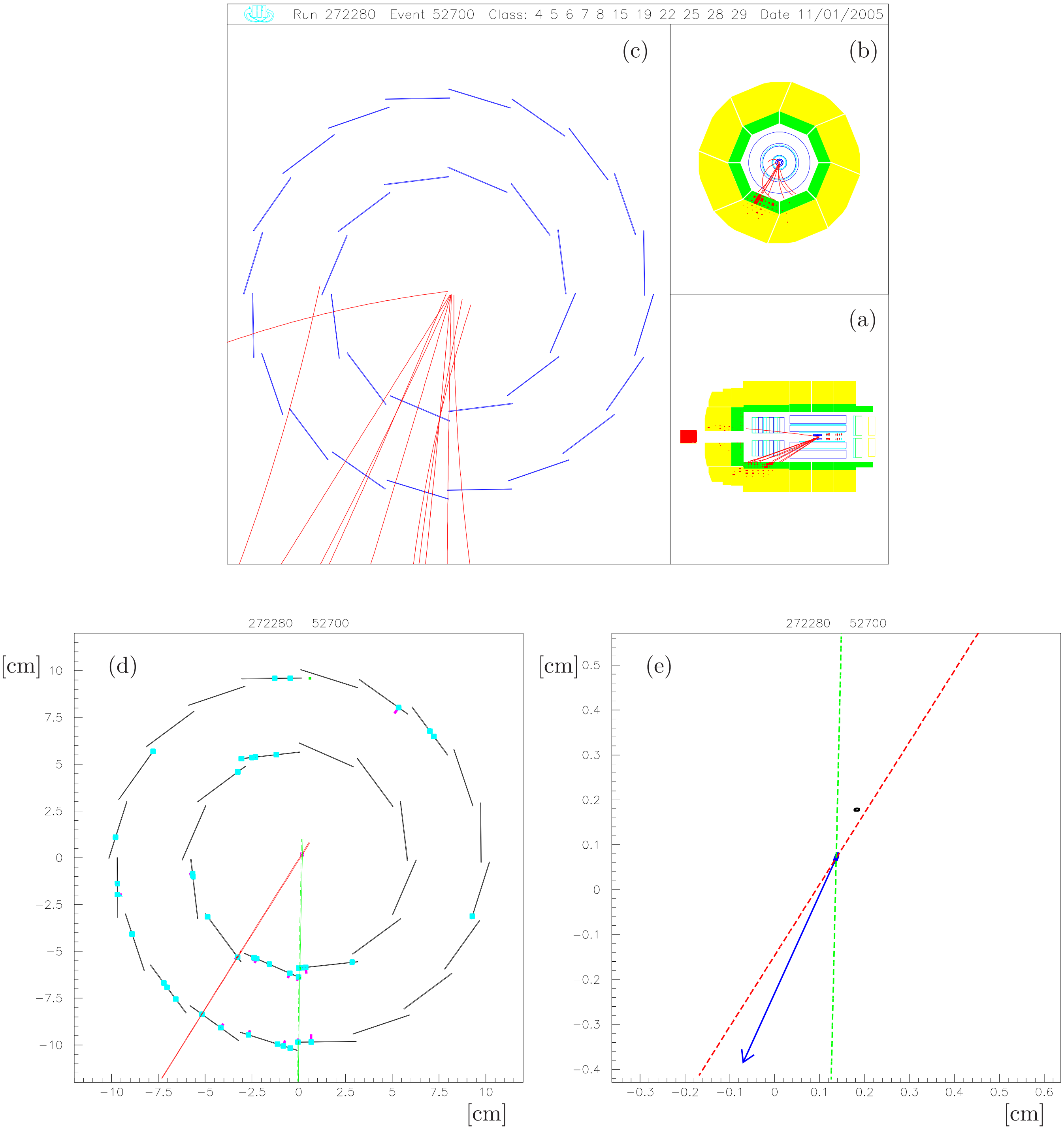


Figure D.5: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.117\ cm$ and the reconstructed invariant mass is $2.83\ GeV$. For further explanations see figure D.1.

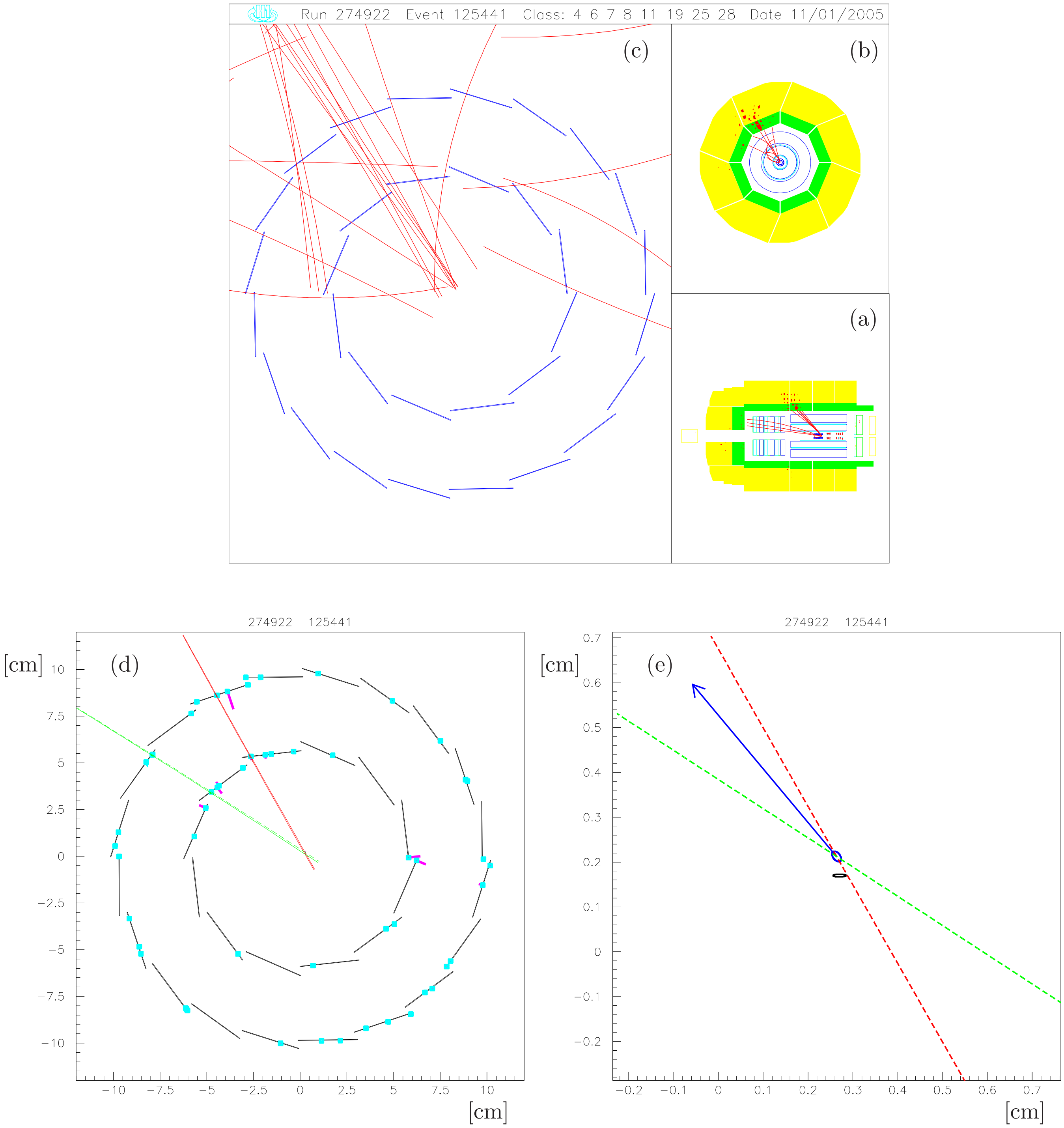


Figure D.6: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.058\ cm$ and the reconstructed invariant mass is $1.20\ GeV$. For further explanations see figure D.1.

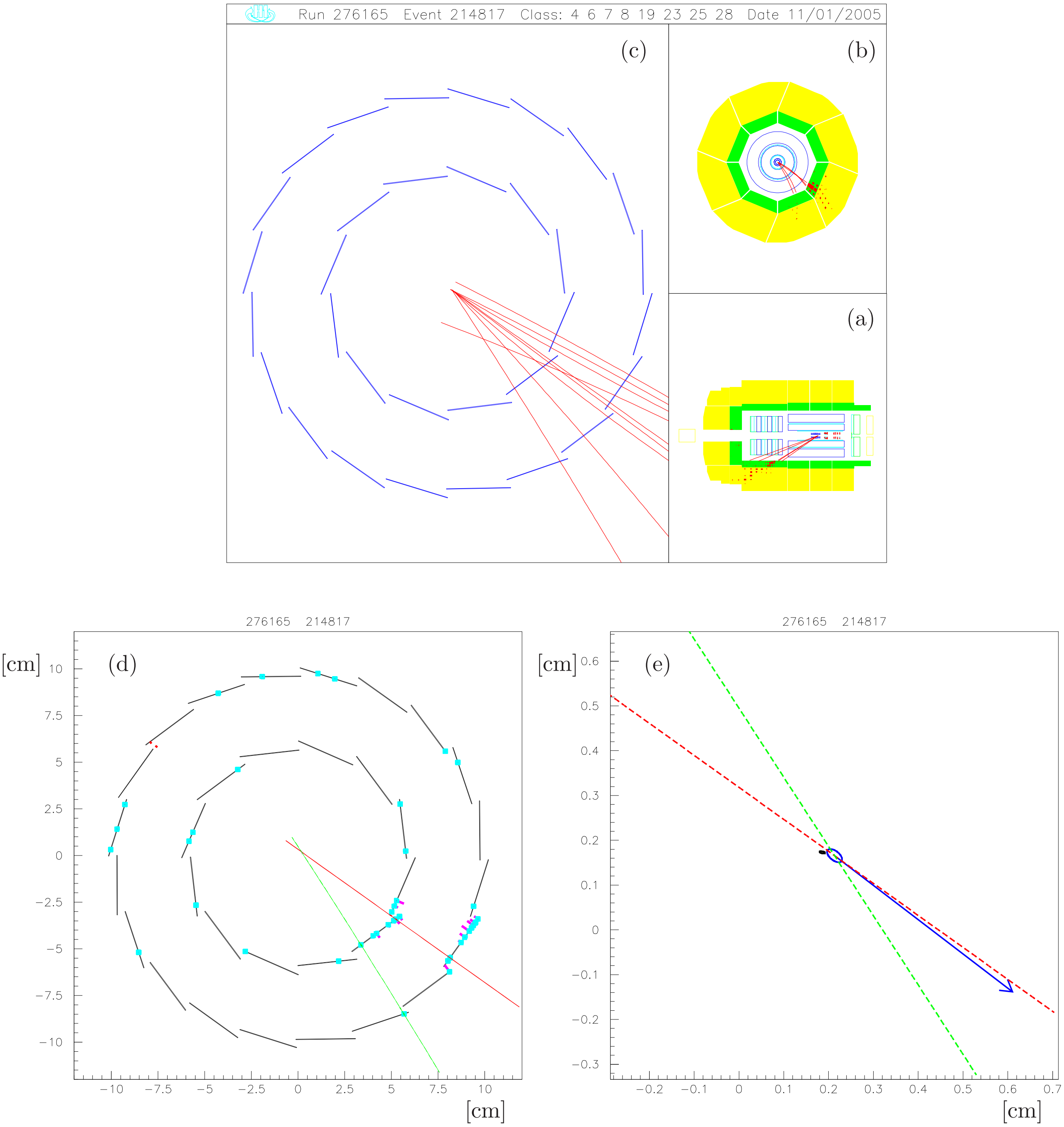


Figure D.7: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.029\ cm$ and the reconstructed invariant mass is $1.86\ GeV$. For further explanations see figure D.1.

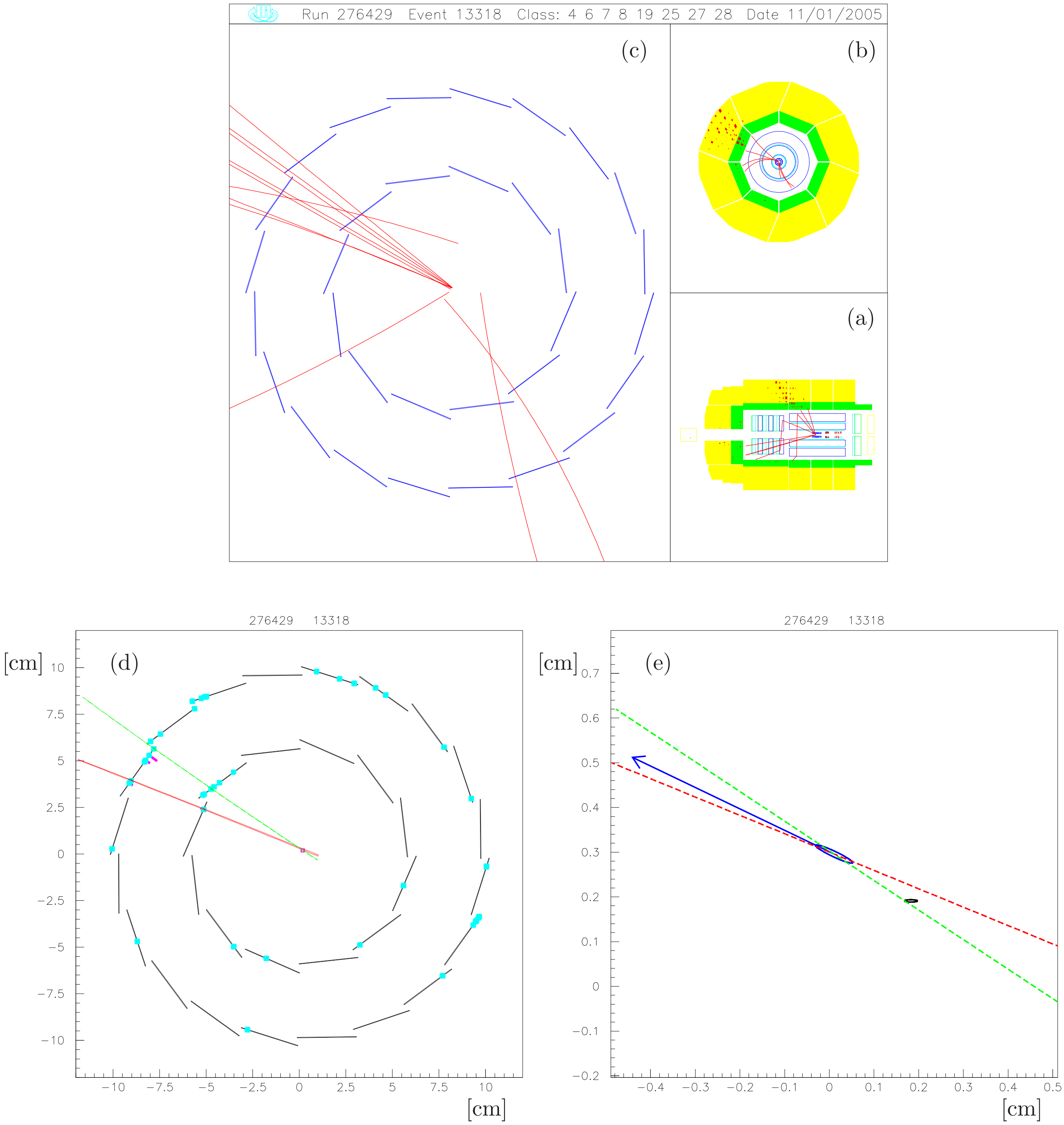


Figure D.8: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.231\ cm$ and the reconstructed invariant mass is $0.66\ GeV$. For further explanations see figure D.1.

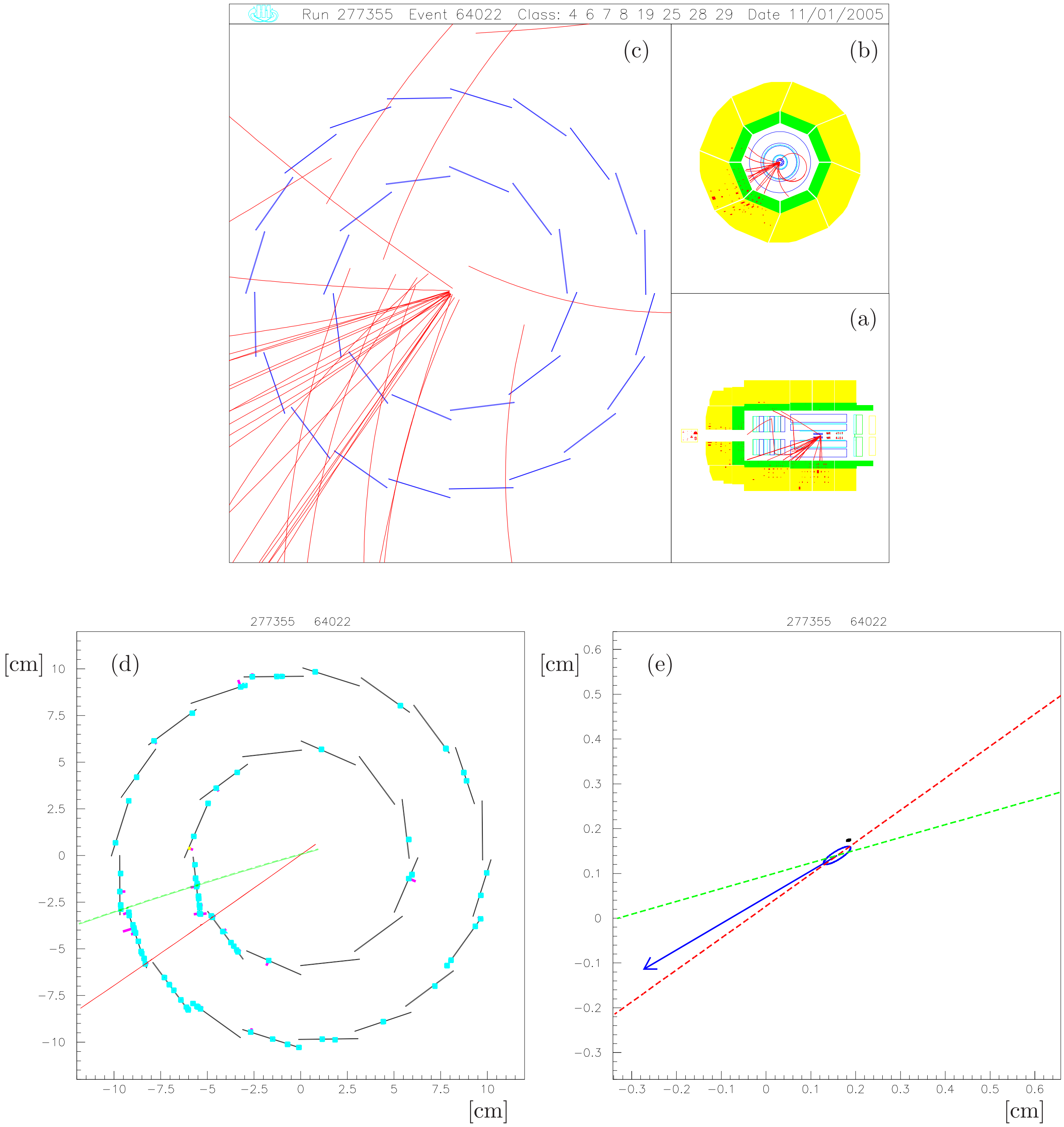


Figure D.9: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.052\ cm$ and the reconstructed invariant mass is $0.61\ GeV$. For further explanations see figure D.1.

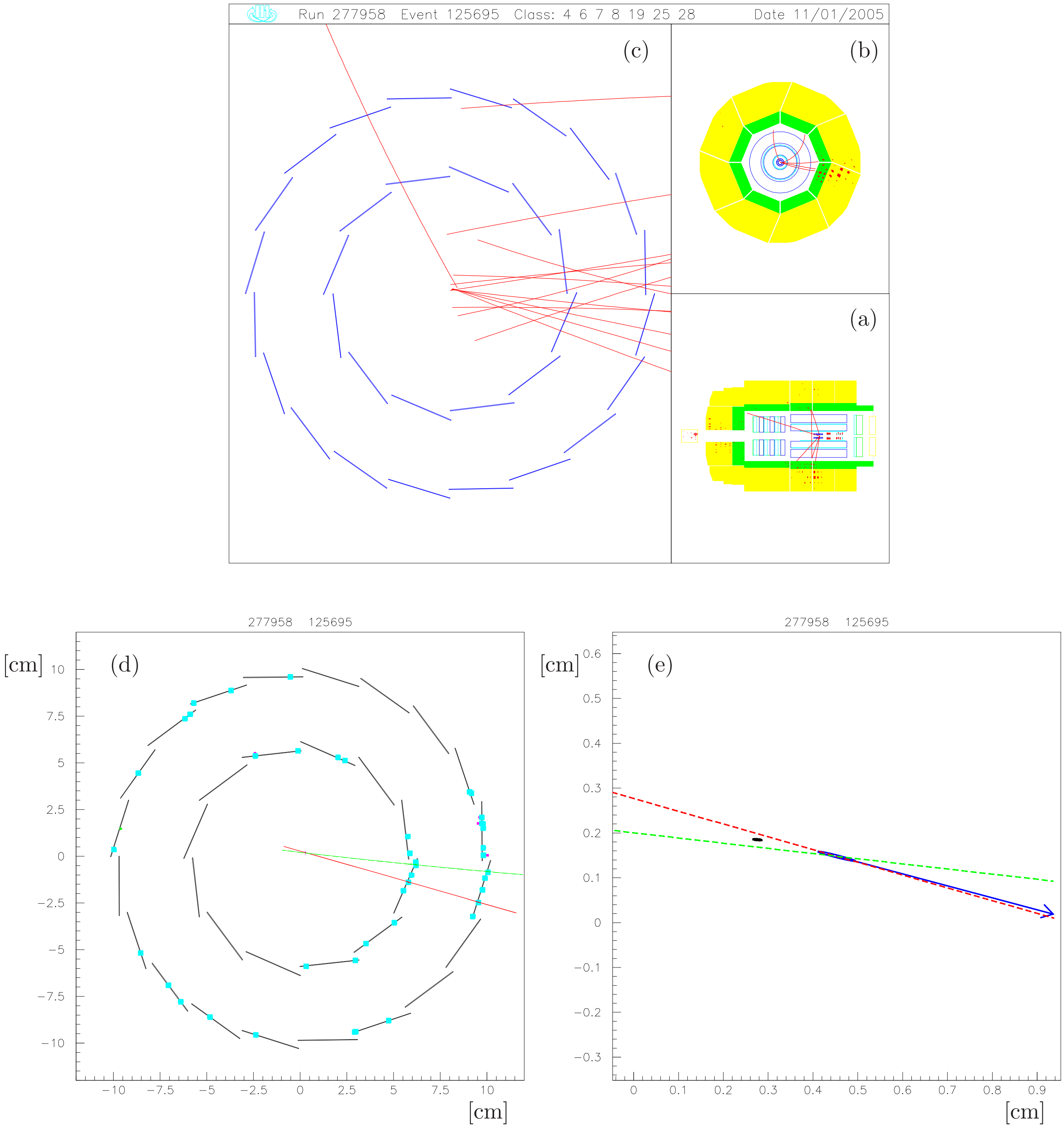


Figure D.10: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.180\ cm$ and the reconstructed invariant mass is $1.08\ GeV$. For further explanations see figure D.1.

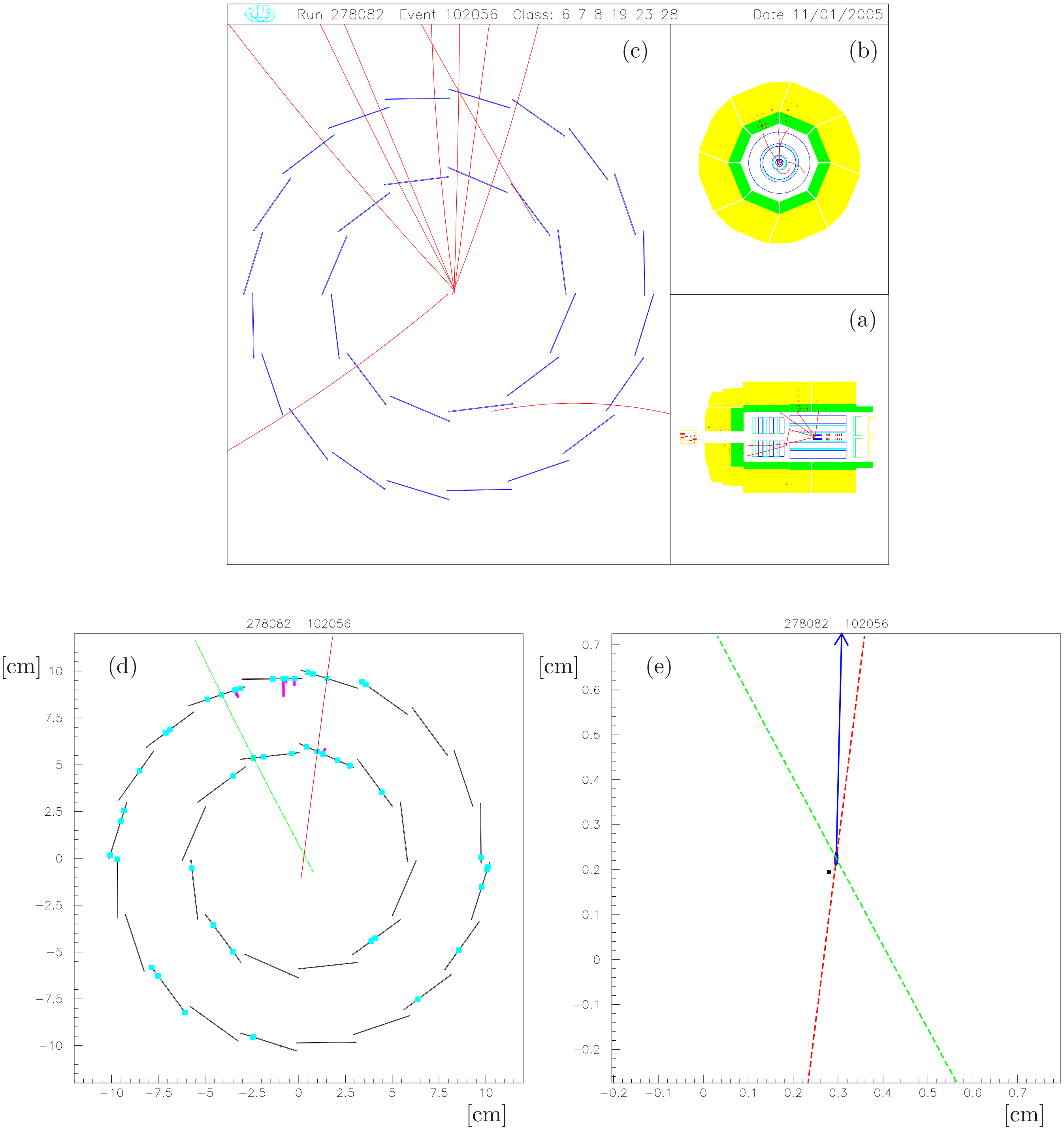


Figure D.11: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.027\ cm$ and the reconstructed invariant mass is $1.63\ GeV$. For further explanations see figure D.1.

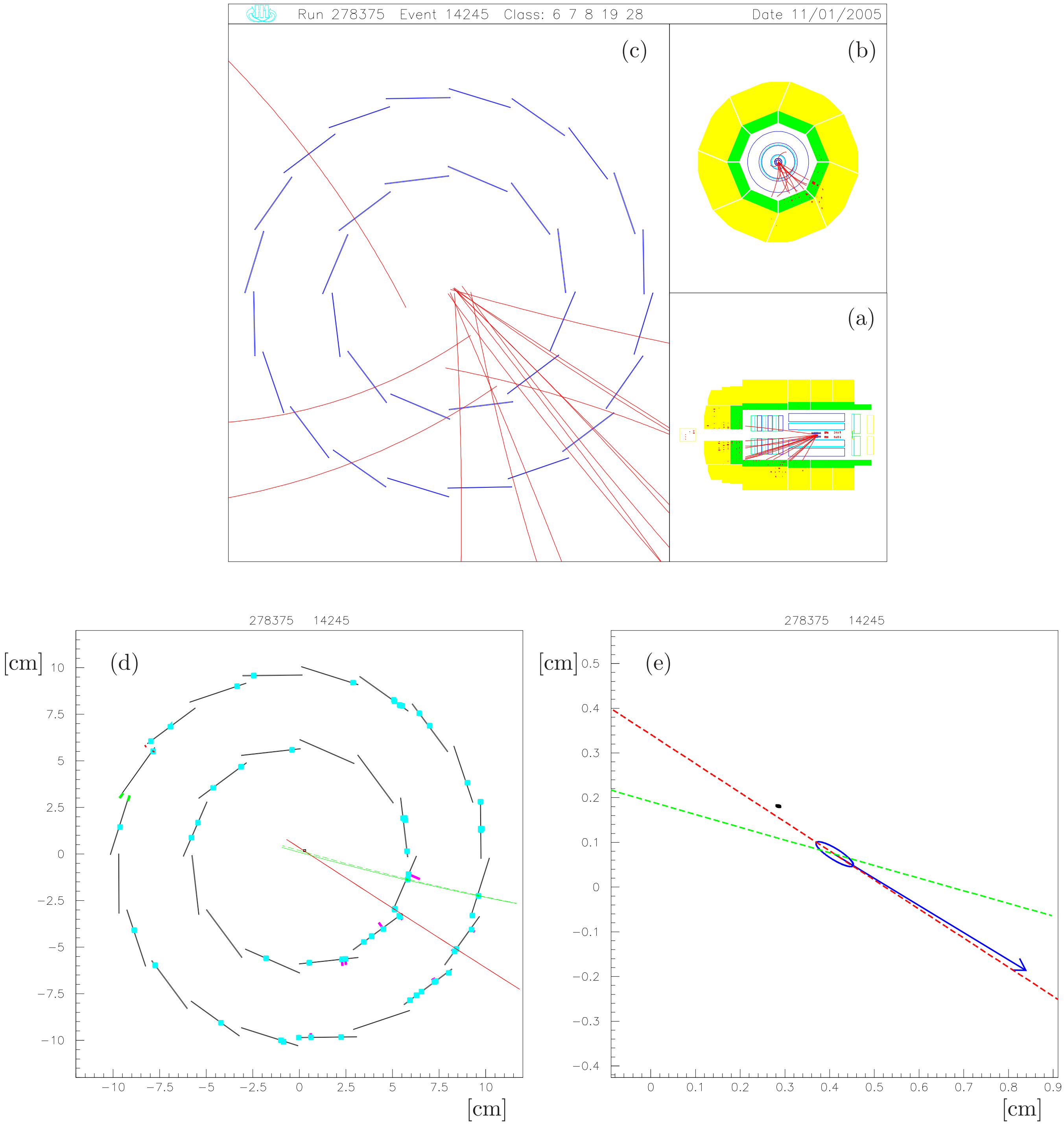


Figure D.12: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = 0.161\ cm$ and the reconstructed invariant mass is $0.89\ GeV$. For further explanations see figure D.1.

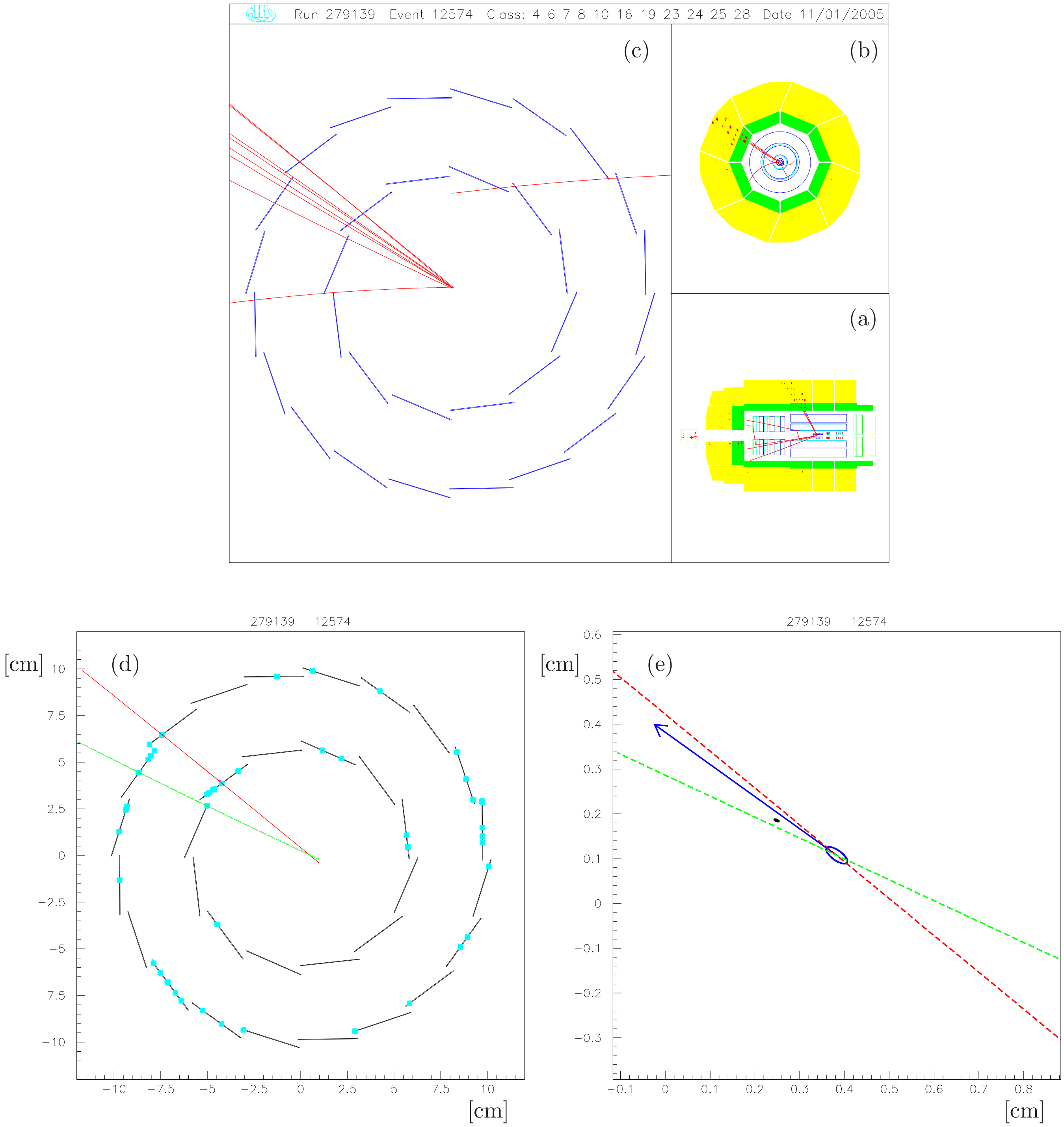


Figure D.13: A CC event with a 2 tracks decay vertex selected with the VVM method: The measured radial decay length is $L_{xy} = -0.153\ cm$ and the reconstructed invariant mass is $0.78\ GeV$. For further explanations see figure D.1.

List of Figures

1.1 The HERA storage ring and the pre-accelerators at DESY 11
1.2 Overview of the H1 detector and the definition of the coordinate system . . . 13

2.1 Neutral current and charged current process in e^+p scattering 17
2.2 Feynman diagram for charm and beauty production in NC 21
2.3 Feynman diagram for charm production in CC 22
2.4 A general CC Feynman diagram 24

3.1 Schematic pictures of the decay length and the 2-tracks method 28
3.2 The impact parameter 29
3.3 The reconstructed impact parameter and decay length significance from MC . 30
3.4 The 3-tracks method 31
3.5 The charged decay-multiplicity for charm and beauty hadrons 33

4.1 The spread between generated and reconstructed $E_{e'}$ and $\theta_{e'}$ values 40
4.2 The spread between generated and reconstructed Q^2_{rec} and y_{rec} kinematics . . 41
4.3 Electron energy calibration 42
4.4 Hadron versus electron energy calibration 42
4.5 Comparison of the data description by the simulation for NC events 44
4.6 Comparison of the data description by the simulation for CC events 47
4.7 The measured kinematic range for NC and CC 48

5.1 Comparison of data by the simulation for track information in NC 50
5.2 Strangeness reduction cut 51
5.3 Comparison of data by the simulation for primary vertex information in NC . 53
5.4 Comparison of data by the simulation for jet information in NC 55

6.1 Comparison of data by the simulation for the 3-tracks method in NC 58
6.2 Cut based heavy quarks cross section with FVM method in NC 61
6.3 Charm and beauty cross sections with fit procedure in FVM method 63
6.4 Fit combination of the 2- and 3-tracks method 64
6.5 Fit decay length significance after negative sign subtraction 65
6.6 Vertex decay-multiplicity in the VVM method in NC 69
6.7 Variable S_l and $min(S_{di})$ for the 2 and 3 tracks decay vertex in NC 70

6.8 Variable L_{xy} and mass for the 2, 3 and 4 tracks decay vertex 71
6.9 The beauty and charm cross sections in NC 77

7.1 Comparison of CC and NC-metamorphosed kinematics 81
7.2 Variable S_l and $min(S_{di})$ for 1 and 2 tracks decay vertex in NC-metamorphosed 82
7.3 Vertex decay-multiplicity in the VVM method in CC 86
7.4 Variable S_l and $min(S_{di})$ for the 1 and 2 tracks decay vertex in CC 87
7.5 Better charm efficiency in NC than in CC 89
7.6 Control of the 14 CC events with a 2 tracks decay vertex 91
7.7 CC event measured by the H1 detector with a 2 tracks decay vertex 93
7.8 Comparison of data by the simulation for the variable $P_{T,h}$ 94
7.9 Cut based charm cross section with FVM method in CC 97

B.1 Comparison of data by the simulation for track information in CC 107
B.2 Comparison of data by the simulation for primary vertex information in CC . 108
B.3 Comparison of data by the simulation for jet information in CC 109
B.4 Comparison of data by the simulation for the 2-tracks method in CC 110
B.5 Comparison of CC and NC-metamorphosed jet information 111
B.6 Comparison of CC and NC-metamorphosed heavy quark tagging information 112
B.7 Low statistic NC-metamorphosed fit 113

D.1 2 tracks decay vertex in CC: Run=262397 and Event=254359 117
D.2 2 tracks decay vertex in CC: Run=263337 and Event=17795 118
D.3 2 tracks decay vertex in CC: Run=268001 and Event=8957 119
D.4 2 tracks decay vertex in CC: Run=268336 and Event=36823 120
D.5 2 tracks decay vertex in CC: Run=272280 and Event=52700 121
D.6 2 tracks decay vertex in CC: Run=274922 and Event=125441 122
D.7 2 tracks decay vertex in CC: Run=276165 and Event=214817 123
D.8 2 tracks decay vertex in CC: Run=276429 and Event=13318 124
D.9 2 tracks decay vertex in CC: Run=277355 and Event=64022 125
D.10 2 tracks decay vertex in CC: Run=277958 and Event=125695 126
D.11 2 tracks decay vertex in CC: Run=278082 and Event=102056 127
D.12 2 tracks decay vertex in CC: Run=278375 and Event=14245 128
D.13 2 tracks decay vertex in CC: Run=279139 and Event=12574 129

List of Tables

4.1 Masses and lifetimes for beauty and charm hadrons 38
4.2 Final NC cuts 45
4.3 Final CC cuts 46

5.1 Track selection criteria for primary and secondary vertex fit 54

6.1 Light, charm and beauty cross sections with FVM method in NC 66
6.2 Results with FVM method in NC 68
6.3 Light, charm and beauty cross sections with VVM method in NC 72
6.4 Systematic errors for variable S_l in VVM method in NC 74
6.5 Results with VVM method in NC 75
6.6 Theoretical errors for NLO calculations 75

7.1 Light and charm cross sections in NC-metamorphosed 83
7.2 Systematic errors for variable S_l in NC-metamorphosed 83
7.3 Results of ten low statistic NC-metamorphosed cross sections 84
7.4 Comparison of CC data events with MC prediction 85
7.5 Light and charm cross sections with VVM method in CC 88
7.6 Results in CC 88
7.7 Control of positive and negative distributions at $P_{t,h} > 16\ GeV$ 92
7.8 Comparison of data events with MC program prediction for different $P_{T,h}$ regions 95
7.9 Control of positive and negative distributions for the region $16 < P_{T,h} < 40\ GeV$ 96
7.10 Control of positive and negative distributions at $P_{T,h} > 40\ GeV$ 96
7.11 Light and charm cross sections with FVM method in CC 99

A.1 Systematic errors for variable S_l in FVM 3-tracks method in NC 104
A.2 Systematic errors for variable $min(S_{di})$ in FVM 3-tracks method in NC . . . 104
A.3 Systematic errors for variable S_l in FVM 2-tracks method in NC 105
A.4 Systematic errors for variable $min(S_{di})$ in FVM 2-tracks method in NC . . . 105
A.5 Systematic errors for variable $min(S_{di})$ in VVM method in NC 106

C.1 Systematic errors for variable $min(S_{di})$ in NC-metamorphosed 114
C.2 Systematic errors for variable norm. & S_l in NC-metamorphosed 115
C.3 Systematic errors for variable norm. & $min(S_{di})$ in NC-metamorphosed . . . 115

D.1 Information to the 14 CC events with 2 tracks decay vertex 116

Bibliography

[1] 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, (2004), hep-ex/0411046.

[2] I. Abt et al., The H1 detector at HERA, Nucl. Instrum. Meth. **A386**, 310–347 (1997).

[3] D. Pitzl et al., The H1 silicon vertex detector, Nucl. Instrum. Meth. **A454**, 334–349 (2000), hep-ex/0002044.

[4] J. Gassner, *A Measurement of D-Meson Production at HERA by Decay Vertex Identification*, PhD thesis, Swiss Federal Institute of Technology Zurich, 2001.

[5] W. Erdmann, SV: A package for secondary vertex, http://www-h1.desy.de/iwork/ibtag/workpages/werdmann/sv/sv.html.

[6] A. Dubak, *Measurement of the e^+p Neutral Current DIS Cross Section and the F_2, F_L, xF_3 Structure Functions in the H1 Experiment at HERA*, PhD thesis, Technische Universität München, 2003.

[7] R. Brun, F. Bruyant, M. Maire, A. C. McPherson, and P. Zanarini, GEANT3, CERN-DD/EE/84-1.

[8] C. Adloff et al., Measurement of neutral and charged current cross-sections in positron proton collisions at large momentum transfer, Eur. Phys. J. **C13**, 609–639 (2000), hep-ex/9908059.

[9] V. Barone, U. D'Alesio, and M. Genovese, The charm-strange contribution to charged-current DIS structure functions, (1996), hep-ph/9610211.

[10] R. K. Ellis, W. J. Stirling, and B. R. Webber, QCD and collider physics, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8**, 1–435 (1996).

[11] M. E. Peskin and D. V. Schroeder, An Introduction to quantum field theory, Reading, USA: Addison-Wesley (1995) 842 p.

[12] 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**, 73–87 (2002), hep-ph/0110215.

[13] H. L. Lai et al., Global QCD analysis and the CTEQ parton distributions, Phys. Rev. **D51**, 4763–4782 (1995), hep-ph/9410404.

[14] J. Pumplin et al., New generation of parton distributions with uncertainties from global QCD analysis, JHEP **07**, 012 (2002), hep-ph/0201195.

[15] M. Hansson, H. Jung, and L. Jonsson, The unintegrated gluon density in the photon and heavy quark production, (2004), hep-ph/0402019.

[16] V. N. Gribov and L. N. Lipatov, Deep inelastic *ep* scattering in perturbation theory, Yad. Fiz. **15**, 781–807 (1972).

[17] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Multi - Reggeon Processes in the Yang-Mills Theory, Sov. Phys. JETP **44**, 443–450 (1976).

[18] M. Ciafaloni, Coherence effects in initial jets at small Q^2/s, Nucl. Phys. **B296**, 49 (1988).

[19] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, Leptoproduction of heavy quarks. 2. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies, Phys. Rev. **D50**, 3102–3118 (1994), hep-ph/9312319.

[20] F. I. Olness and S. T. Riemersma, Leptoproduction of heavy quarks in the fixed and variable flavor schemes, Phys. Rev. **D51**, 4746–4755 (1995), hep-ph/9409208.

[21] B. W. Harris and J. Smith, Heavy quark correlations in deep inelastic electroproduction, Nucl. Phys. **B452**, 109–160 (1995), hep-ph/9503484.

[22] E. Laenen, S. Riemersma, J. Smith, and W. van Neerven, Complete O(α_s) corrections to heavy-flavour structure functions in electroproduction, Nucl. Phys. **B392**, 162–228 (1993).

[23] M. Kuhlen, Hadronic final states in deeply inelastic scattering, (1995), hep-ex/9508014.

[24] D. Peterson, C. Schlatter, I. Schmitt, and P. Zerwas, Scaling violations in inclusive e^+e^- annihilation spectra, Phys. Rev. D **27**, 105–111 (1983).

[25] B. Andersson, G. Gustafson, G. Ingelman, and Sjostrand, Parton fragmentation and string dynamics, Phys. Reports. **97**, 31–145 (1983).

[26] D. Coffman et al., Measurement of the inclusive decay properties of charmed mesons, Phys. Lett. B **263**, 135–140 (1991).

[27] D. Abbaneo et al., Combined results on b-hadron production rates and decay properties, Technical Report CERN-EP/2001-050, ALEPH, CDF, DELPHI, L3, OPAL, SLD, 2001, hep-ex/0112028.

[28] S. Eidelman et al., Review of particle physics, Phys. Lett. **B592**, 1 (2004).

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

[30] 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**, 189–198 (1995), hep-ex/9406007.

[31] L. Gladilin, Charm hadron production fractions, (1999), hep-ex/9912064.

[32] S. Catani and M. H. Seymour, A general algorithm for calculating jet cross sections in NLO QCD, Nucl. Phys. **B485**, 291–419 (1997), hep-ph/9605323.

[33] S. D. Ellis and D. E. Soper, Successive combination jet algorithm for hadron collisions, Phys. Rev. **D48**, 3160–3166 (1993), hep-ph/9305266.

[34] R. Frühwirth, W. Waltenberger, K. Prokofiev, T. Speer, and P. Vanlaer, New developments in vertex reconstruction for CMS, Nucl. Instrum. Meth. **A502**, 699–701 (2003).

[35] K. Rose, E. Gurewitz, and G. Fox, Statistical Mechanics and Phase Transitions in Clustering, Physical Review Letters **65**, 945 (1990).

[36] K. Rose, E. Gurewitz, and G. Fox, A Deterministic Annealing Approach to Clustering, Pattern Recognition Letters **11**(9), 589 (1990).

[37] R. Frühwirth and A. Strandlie, Track fitting with ambiguities and noise: a study of elastic tracking and nonlinear filters, Comput. Phys. Commun. **120**, 197–214 (1999).

[38] R. Barlow and C. Beeston, Fitting using finite MonteCarlo samples, Computer Physics Communications **77**, 219–228 (1993).

[39] M. Goossens, HBOOK routines: HMCMLL, (1997), http://wwwasdoc.web.cern.ch/wwwasdoc/hbook_html3/node138.html.

[40] B. Heinemann, *Measurement of Charged Current and Neutral Current Cross Sections in Positron-Proton Collisions at* $\sqrt{s} \simeq 300$ *GeV*, PhD thesis, Universität Hamburg, 1999.

[41] Z. Zhang, *New Insights into the Proton Structure with ep Collider HERA*, 2000, Université de Paris-Sud, Habilitation.

[42] C. Adloff et al., Measurement and QCD analysis of neutral and charged current cross sections at HERA, Eur. Phys. J. **C30**, 1–32 (2003).

[43] S. Aid et al., Measurement of the e^+ and e^- induced charged current cross-sections at HERA, Z. Phys. **C67**, 565–576 (1995), hep-ex/9506002.

[44] B. Reisert, *Elektron-Proton-Streuunung bei hohen Impulsüberträgen am H1-Experiment bei HERA*, PhD thesis, Ludwig-Maximilians-Universität München, 2000.

[45] G. A. Schuler and H. Spiesberger, DJANGO: The Interface for the event generators HERACLES and LEPTO, In Hamburg 1991, Proceedings, Physics at HERA, vol. 3 1419-1432. (see high energy physics index 30 (1992) No. 12988).

[46] 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**, 155–172 (1992).

[47] G. Ingelman, A. Edin, and J. Rathsman, LEPTO 6.5 - A Monte Carlo Generator for Deep Inelastic Lepton-Nucleon Scattering, Comput. Phys. Commun. **101**, 108–134 (1997), hep-ph/9605286.

[48] L. Lonnblad, The colour-dipole model and the ARIADNE program at high Q^2, (1999), hep-ph/9908368.

[49] L. Lonnblad, ARIADNE version 4: A Program for simulation of QCD cascades implementing the color dipole model, Comput. Phys. Commun. **71**, 15–31 (1992).

[50] J. Pumplin et al., New generation of parton distributions with uncertainties from global QCD analysis, JHEP **07**, 012 (2002), hep-ph/0201195.

[51] T. Sjostrand, PYTHIA 5.7 and JETSET 7.4: Physics and manual, (1995), hep-ph/9508391.

[52] T. Sjostrand, High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4, Comput. Phys. Commun. **82**, 74–90 (1994).

[53] H. Plothow-Besch, PDFLIB: A Library of all available parton density functions of the nucleon, the pion and the photon and the corresponding alpha-s calculations, Comput. Phys. Commun. **75**, 396–416 (1993).

[54] U. Bassler and G. Bernardi, On the kinematic reconstruction of deep inelastic scattering at HERA: The Sigma method, Nucl. Instrum. Meth. **A361**, 197–208 (1995), hep-ex/9412004.

[55] F. Jacquet and A. Blondel, Detection and Study of the Charged Current Event, Study of an ep-Facility for Europe, DESY , 391–396 (1979).

[56] I. Négri, *Recherche de fermions excités dans l'expérience H1 auprès du collisionneur positron-proton HERA*, PhD thesis, Université de la méditerranée Marseille II, 1998.

[57] E. Chabert, C. Diacnou, S. Kermiche, I. Négri, and C. Vallée, QBGFMAR: An updated phan package for cosmic and halo muon topological rejection in high P_T physics analysis, H1-11/98-556 .

[58] Fast parametrisation of CC trigger efficiency 1994-2000, http://www-h1.desy.de/h1/iww/ipwg/iinc/analysis/cctrigger.html .

[59] B. W. Harris and J. Smith, Invariant mass distributions for heavy quark - anti-quark pairs in deep inelastic electroproduction, Phys. Lett. **B353**, 535–540 (1995), hep-ph/9502312.

[60] D. Meer, *Heavy quarks in photoproduction at HERA and a linking algorithm for the Fast Track Trigger*, PhD thesis, Swiss Federal Institute of Technology Zurich, 2005.

[61] V. Andreev et al., Isolated electrons and muons in events with missing transverse momentum at HERA, Phys. Lett. **B561**, 241–257 (2003), hep-ex/0301030.

[62] C. Adloff et al., Observation of events at very high Q^2 in *ep* collisions at HERA, Z. Phys. **C74**, 191–206 (1997), hep-ex/9702012.

Danksagung

Ich danke Herrn Prof. Eichler für die Aufnahme in seine Arbeitsgruppe und seine interessierte Anteilnahme am Fortgang der Arbeit.

Bei Wolfram Erdmann bedanke ich mich für die gute Betreuung der Arbeit, die immerwährende Bereitstellung neuer Ideen und die pragmatischen Fragestellungen. Ohne ihn und seine Erfahrung im Vertexing wäre diese Arbeit um einiges bescheidener ausgefallen.

Die unendliche Anzahl von Computer- und Physikproblemen konnten nur durch die Hilfe von Dotorandenkollegen Simon Baumgartner, Markus Bischofberger und David Meer gelöst werden. Schön war auch die Zeit ausserhalb der Physik, die wir miteinander hatten.

Zu Beginn der Arbeit waren die vielen kleinen Tipps hilfreich, die von Benno List, Johannes Gassner und Michael Hilgers kamen.

Diese Arbeit war natürlich nur mit der Unterstützung der H1-Kollaboration möglich.

Für das Korrekturlesen der Arbeit danke ich Wolfram Erdmann, Christoph Grab, Prof. Eichler und meinem Koreferenten Prof. Dissertori.

Marta Sekolec und Rosa Bächli danke ich für die kompetente Hilfe in allen administrativen Belangen.

Infine ringrazio mia madre e mio padre di avermi sempre dato la possibilità di andare la via che potevo.

Curriculum Vitae

Personal Data

Full Name	Salvatore Mangano
Date of Birth	May 28, 1974
Place of Birth	Zürich (Switzerland)
Nationality	Swiss

Education

1981 – 1995	Public Schools of Zurich
1995 – 2000	Studies of physics at the Swiss Federal Institute of Technology (ETH) in Zurich
2000 – 2001	Diploma thesis (4 months) at INFN Napoli (Italy). Title of the thesis: "Misura della risoluzione angolare in film di emulsioni nucleari per l'esperimento OPERA" Scientific assistant (4 months)
2001 – 2005	Graduate education in experimental particle physics at the ETH Zurich