Search for Events with Isolated Leptons and Large Missing Transverse Momentum in ep Collisions at HERA

Thesis submitted in accordance with the requirements
of the University of Liverpool
for the degree of Doctor in Philosophy

by

Christian Veelken

January 2005

Abstract

In this thesis, the results of a search for events with isolated electrons, muons or tau leptons and large missing transverse momentum in the HERA I e^+p and e^-p datasets are presented. The analysed data was recorded by the H1 experiment in the period 1994-2000 and corresponds to an integrated luminosity of 116.5 pb^{-1}.

In total, 19 events are observed in the data, in agreement with a Standard Model expectation of 19.9 ± 3.0, which is dominated by the contribution from the production of real W bosons with subsequent leptonic decay. In the region of high transverse momentum P_T^X of the hadronic final state, an excess over the Standard Model expectation is observed. The excess is most significant in the region $P_T^X > 40$ GeV, where 5 events are observed in the data, while only 1.4 ± 0.4 are expected from all Standard Model processes.

The observed excess may be interpreted as being either a statistical fluctuation or a signal for possible processes in theories beyond the Standard Model. The events observed at large hadronic transverse momenta are found to be compatible with the production of single top quarks by flavour-changing neutral current interactions. The origin of the excess cannot be unambigously clarified, because of the small number of events observed.

Zusammenfassung

In der vorliegenden Arbeit werden die Ergebnisse einer Suche nach Ereignissen mit isolierten Leptonen und hohem fehlenden Transversalimpuls in den e^+p und e^-p HERA I beschrieben. Die untersuchte Datenmenge wurde vom H1 Experiment in den Jahren 1994-2000 aufgezeichnet und entspricht einer integrierten Luminosität von 116.5 pb^{-1}.

Insgesamt werden 19 Ereignisse in den Daten beobachtet, in Übereinstimmung mit einer Standardmodell Vorhersage von 19.9 ± 3.0, die dominiert ist vom Beitrag der Produktion reeller W Bosonen mit anschließendem leptonischen Zerfall. Im Bereich hoher Transversalimpulse P_T^X des hadronischen Endzustands wird ein Überschuß von Ereignissen über die Vorhersage des Standardmodells beobachtet. Der Überschuß ist am signifikantesten im Bereich $P_T^X > 40$ GeV, wo 5 Ereignisse in den Daten beobachtet werden, während lediglich 1.4 ± 0.4 in der Summe aller Standardmodell Prozesse erwartet sind.

Der beobachtete Überschuß von Ereignissen kann entweder als statistische Fluktuation oder als Anzeichen für mögliche Prozesse in Theorien außerhalb des Standardmodells interpretiert werden. Die bei hohen hadronischen Transversalimpulsen beobachteten Ereignisse werden als kompatibel mit der Produktion einzelner top Quarks durch flavour verändernde Wechselwirkungen des Neutralen Stroms befunden. Aufgrund der kleinen Anzahl von beobachteten Ereignissen kann der Urprung des Überschusses jedoch nicht zweifelsfrei festgestellt werden.

Contents

Introduction

In the analysis presented in this thesis, events with isolated leptons and large missing transverse momentum are searched for in the data recorded by the H1 experiment at the HERA storage ring of the *Deutsche Elektronen Synchrotron* (DESY) in Hamburg. In the HERA storage ring, electrons and protons are accelerated to energies of 27.6 GeV and 920 GeV respectively and brought into collision at a centre-of-mass energy of $\sqrt{s} = 318$ GeV (300 GeV before 1998). In electron-proton collisions at HERA, the proton structure may be studied down to length scales of 10^{-18} m (which is about one thousandth of the proton radius) and in a kinematic regime not accessible in previous fixed-target experiments.

Additionally, the Standard Model of particle physics may the tested in the region of very high energies, currently exceeded only in proton anti-proton collisions at the TeVatron ($\sqrt{s} = 1.96$ TeV) of the *Fermi laboratory* (Fermilab) in Chicago, Illinois. The study of events with isolated leptons and large missing transverse momentum is of particular interest in this context, as previous H1 analyses reported the observation of an excess of events of this kind over the Standard Model expectation [1–5]. In a generic search for new phenomena beyond the Standard Model, the observed excess was subsequently confirmed to represent the most significant deviation from the predictions of the Standard Model in the H1 data at large transverse momenta [6]. Within the Standard Model, the dominant contribution to an event sample such as the one observed is expected to arise from the production of real W bosons with subsequent leptonic decay, a rare process with a cross-section of about 1 pb [7–9]. In contrast to the expectation for Standard Model W production, in several of the observed events hadronic final states of large transverse momentum are present, making an interpretation within the Standard Model unlikely. Within theories beyond the Standard Model, the production of single top quarks by flavour-changing neutral current interactions has been quoted as the most likely interpretation of the observed excess in previous analyses [10, 11].

In this thesis, the results of the search for events with isolated leptons and large missing transverse momentum in the e^+p and e^-p datasets recorded during the HERA I running period 1994-2000 are presented. The integrated luminosity of the analysed datasets is 116.5 pb^{-1}. In comparison to previous analyses of the same datasets, the acceptance for events with isolated leptons is increased by extending the search previously restricted to events with isolated electrons or muons to events with isolated tau leptons and searching for events with isolated muons in an extended phase-space. For the search for events with isolated tau leptons, a neural network based algorithm for the identification of hadronic tau decays is developed. The extention of the phase-space for events with isolated muons is made possible by using complementary event triggers, based on coincidences of signals in the muon system with tracks in the central tracking detectors, in addition to those based on energy deposits in the liquid argon calorimeter of the H1 detector used in previous analyses.

This thesis is organized into nine chapters. In the first chapter, the HERA storage ring and the H1 detector are described. The description of the experimental conditions is followed by a presentation of the theoretical background in the second chapter. The production of real W bosons within the Standard Model and of single top quarks by flavour-changing neutral current interactions within theories beyond the Standard Model are described in detail. The operation of neural networks is explained in chapter three, before the identification and reconstruction of electrons, muons and tau leptons is detailed in the fourth chapter. In the fifth chapter, the reconstruction of kinematic and topological quantities that are relevant for the event selection is described. The event selection is described in two chapters. In the sixth chapter, data-quality requirements that are used in similar forms in all H1 analyses are detailed. The selection criteria that are specific to the analysis presented in this thesis are described in the seventh chapter. The analysis specific selection criteria are motivated by studying the signature of background processes in a phase-space similar to that defined by the final event selection. In the eighth chapter, the estimation of systematic uncertainties is described and the results of the search for events with isolated leptons and large missing transverse momentum are presented. The results are discussed in the ninth chapter and compared with those of previous H1 analyses and similar searches performed by the ZEUS collaboration. Following these nine chapters, the results and conclusions of the analysis presented in this thesis are summarized.

Throughout this thesis, *natural units* are used in which $c = \hbar = 1$.

Chapter 1

The H1 detector at HERA

1.1 HERA Accelerator

The *"Hadron-Electron-Ring-Anlage"* [1] (*HERA*) constitutes the world's only electron [2] -proton collider. The HERA collider was designed to study proton structure in a kinematic range inaccessible to previous fixed-target experiments. As is explained in section 2.1, electron-proton scattering processes can be described by two kinematic variables; in terms of the commonly used variables x and Q^2 (described in more detail in section 2.1), HERA extends the accessible kinematic range by about two orders of magnitude. The *kinematic plane* spanned by the two variables x and Q^2 and covered by HERA is shown in figure 1.1 in comparison with the range covered by previous fixed-target experiments. In HERA, electrons are accelerated to a beam energy of $E_e^{beam} = 27.6\ GeV$ and brought into collision with a proton beam of $E_p^{beam} = 920\ GeV$ ($820\ GeV$ until 1997) energy, giving a centre-of-mass energy of $\sqrt{s} = 318\ GeV$ ($301\ GeV$ until 1997), equivalent to a $\approx 50\ TeV$ electron beam incident on a stationary target. In addition to the study of proton structure, electron-proton collisions at HERA therefore allow the Standard Model of particle physics to be tested at the highest energies available in electron-proton collisions ever.

The HERA accelerator has a circumference of $6.336\ km$ and resides in a tunnel some meters below the "Volkspark" at the *Deutsches Elektronen Synchrotron* (DESY) in Hamburg-Bahrenfeld. The accelerator consists of two storage rings, for electrons and protons, respectively. The beams are brought into collision in two general purpose detectors, *H1* and *ZEUS*, situated at the north and south experimental halls of the HERA ring. In addition to the collider type experiments H1 and ZEUS, the electrons and protons delivered by HERA are used by two fixed-target experiments, *HERMES* and *HERA-B*, that are located at the east and west sides of the HERA ring. The HERMES experiment studies collisions of the electron beam with polarized gas targets to investigate the spin-structure of the proton. In the HERA-B experiment the proton beam is incident on a wire target, allowing the copious production and study of b-quarks. A sketch of the HERA ring showing the four experiments H1, HERMES, HERA-B and ZEUS on the DESY site together with

[1] "Anlage" is the German word for facility

[2] The HERA ring can be operated with either electrons or positrons; for simplicity, both electrons and positrons will be referred to as "electrons" in this thesis.

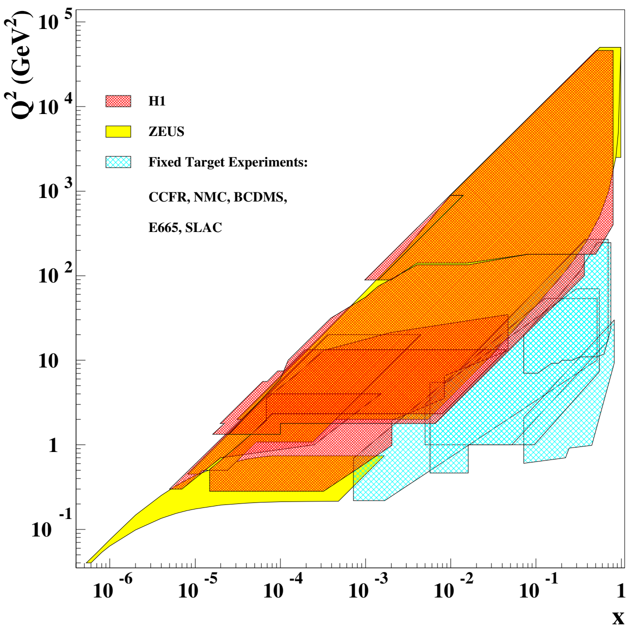


Figure 1.1: *The HERA coverage of the kinematic plane in comparison to previous fixed-target experiments.*

the *DESY* and *PETRA* pre-accelerators is displayed in figure 1.2. The construction of the HERA ring was finished in 1990, with first collisions occurring in 1992.

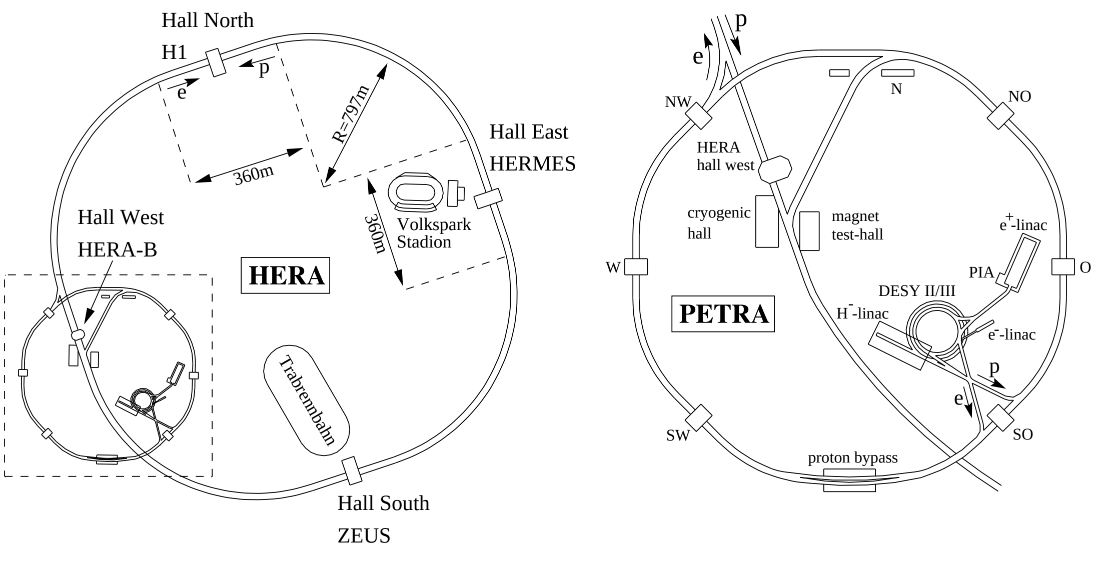


Figure 1.2: *The HERA accelerator and its pre-accelerator, PETRA.*

In operation, the HERA ring is filled with up to 220 *bunches* in each of the electron and proton beams. The bunches are brought into collision with a period of 96 *ns* between two consecutive *bunch-crossings* of the two beams. Typically only 175 of the 220 possible bunches in each beam is filled with *colliding bunches*. The gaps between these are filled with *pilot bunches* of electrons and protons that do not collide with a bunch from the other beam. Since no *ep* interactions take place when pilot bunches cross the H1 and ZEUS detectors, the pilot bunches are useful for the study of beam-related *non-ep backgrounds*, arising from the interactions of beam particles with residual gas atoms in the beam-pipe (*beam-gas*) and collisions with the beam-pipe walls (*beam-wall*). Each of the filled bunches consists of about 10^{11} particles. The particle density within the bunches is spread over a few centimeters in the longitudinal direction; it can be approximated by a Gaussian distribution of width $\sigma_z^e \sim 1\ cm$ for the electron and $\sigma_z^p \sim 10\ cm$ for the proton bunches. As a consequence of this spread, the interaction points of the electron and proton bunches are distributed in the beam, or z, direction with $\sigma_z^{vertex} \sim 10\ cm$ around the nominal interaction point.

The performance of the HERA accelerator is summarized in the produced *luminosity*, which describes the expected number of *ep* interactions in the collision regions of the H1 and ZEUS experiments. The *instantaneous* luminosity depends on the bunch-crossing frequency f, the numbers of particles N_e and N_p in the electron and proton bunches and the collimation of the beams in the transverse direction σ_{xy}^e and σ_{xy}^p,

$$\mathcal{L} = \frac{f \cdot N_e \cdot N_p}{2\pi \cdot {\sigma_{xy}^e}^2 {\sigma_{xy}^p}^2}. \tag{1.1}$$

The number of expected *ep* interactions N is proportional to the *integrated* luminosity

$$L = \int \mathcal{L}\, dt$$

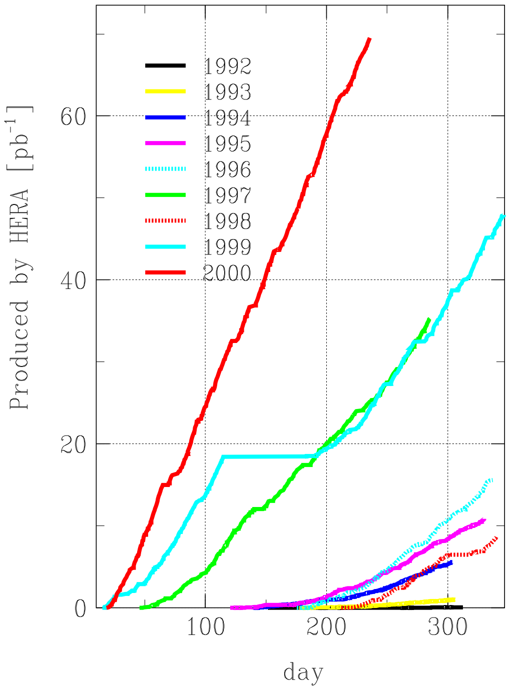
and the total *ep* cross-section σ,

$$N = L \cdot \sigma.$$

As the total *ep* cross-section is small, a large luminosity of the HERA collider is required for physics studies using the H1 and ZEUS experiments.

The luminosity produced by HERA increased steadily between the first collisions in 1992 and the end of the HERA I running period in the year 2000. Between 1994 and 1997 HERA produced e^+p data with 68.0 pb^{-1} integrated luminosity for a proton beam energy of 820 GeV, corresponding to a centre-of-mass energy of $\sqrt{s} = 301\ GeV$. In 1998, HERA switched from positrons to electrons and produced 26.4 pb^{-1} of e^-p data at $\sqrt{s} = 318\ GeV$ for a proton beam energy of 920 GeV. In 1999 and 2000, HERA was operated with positrons again and produced another 103.6 pb^{-1} of e^+p at the higher centre-of-mass energy of $\sqrt{s} = 318\ GeV$. The luminosity produced in the years 1992 to 2000 is displayed in figure 1.3. The increase in the instantaneous luminosity between 1992 and 2000 is apparent and, at the end of the HERA I running period, HERA exceeded its design luminosity of $\mathcal{L}_{nom}^{HERAI} = 1.7 \cdot 10^{31} cm^{-2} s^{-1}$. In total, HERA produced about 200 pb^{-1} integrated luminosity between the years 1994 and 2000.

After HERA had exceeded its design luminosity in the year 2000, the room for improvement was exhausted for the HERA I accelerator and, starting in September 2000, a luminosity upgrade was carried out to accelerate data-taking. For the HERA II running period, the instantaneous luminosity was increased to $\mathcal{L}_{nom}^{HERAII} = 7.4 \cdot$



Year	Luminosity
$e^+p\,(820\ GeV)$	
1994	5.9
1995	11.1
1996	14.4
1997	37.6
$e^-p\,(920\ GeV)$	
1998	8.7
1999	17.7
$e^+p\,(920\ GeV)$	
1999	33.1
2000	70.6

Figure 1.3: *Luminosity produced by HERA in individual years.*

$10^{31} cm^{-2}s^{-1}$ (design values), by increasing the number of particles N_p in the proton bunches and attaining a stronger focus of the electron and proton beams that reduces the transverse beam spreads σ^e_{xy} and σ^p_{xy} in equation 1.1. An overview of the HERA beam parameters before and after the upgrade is given in table 1.1. In order to

Parameter	HERA I		HERA II	
	e-beam	p-beam	e-beam	p-beam
E [GeV]	27.6	920	27.6	920
N_{ppb} [10^{10}]	3.5	7.3	4.0	10.3
I [mA]	50	100	58	140
σ_x [μm]	192	189	112	112
σ_y [μm]	50	50	30	30

Table 1.1: *Operating parameters of the HERA accelerator before (HERA I) and after (HERA II) the luminosity upgrade [12]; "ppb" means "particles per bunch".*

achieve the stronger focusing of the proton beam, the electron beam has to be separated from the proton beam as early as possible [12]. For this purpose, two new combined function magnets were installed inside the H1 and ZEUS detectors, 2 *m upstream* and *downstream* [3] of the nominal interaction points. The luminosity upgrade was finished as scheduled in 2002; however, the first HERA II data for

[3] In the H1 coordinate system, the z-axis is chosen to be in the direction of the proton beam with the origin at the nominal interaction point; correspondingly, "upstream" relates to the direction of negative z (before the proton beam crosses the interaction point) and "downstream" to the direction of positive z (after the proton beam crosses the interaction point).

physics analyses could only be collected in 2003, following the solution of background problems cause primarily by the increased synchrotron radiation power emitted close to the H1 and ZEUS experiments.

1.2 H1 Detector

The H1 detector is built around the northern interaction point of the HERA ring; it is a general purpose detector designed to measure the direction, energy and charge of the particles originating from the *ep* collisions in its centre. The detector is asymmetric in the z-direction, reflecting the asymmetry of the electron and proton beam energies. Due to the higher energy of the proton beam, the centre-of-mass of the colliding particles is not at rest in the laboratory frame, but *boosted* along the proton direction by a Lorentz factor of $\gamma_{cm}^{ep} = 2.86$. Correspondingly, the instrumentation of the H1 detector is predominantly concentrated in the *forward region*, defined as the direction of the incoming proton beam. The *backward* region is less densely instrumented and dedicated mainly to the detection of the scattered electron. Despite its asymmetry, the H1 detector is constructed like most general purpose detectors in a layered design. In its centre, tracking [1] detectors are installed inside a solenoidal magnetic field. These are surrounded by electromagnetic and hadronic calorimeters [2] that are enclosed by iron plates that return the magnetic flux of the solenoid magnet and are instrumented for the detection of muons. The design of the H1 detector is illustrated in figure 1.4, as is the H1 coordinate system [3] In total, the H1 detector measures approximately $12 \times 10 \times 15\ m^3$ (length $\times$ width $\times$ height) and weighs about 2800 t.

The tracking system in the innermost part of the H1 detector is subdivided into *forward*, *central* and *backward* regions. Located in the forward and central regions are, respectively. the *Forward Tracking Detector* (FTD)[3] and *Central Tracking Detector* (CTD)[2]. In the backward region, the tracking system is complemented by the *Backward Drift Chamber* (BDC), that is missing in figure 1.4, but is depicted in an enlarged view of the tracking system in figure 1.5. All tracking detectors are located in an almost homogeneous solenoidal magnetic field of $B = 1.15\ T$ provided by a superconducting coil[6], which makes possible the determination of the charge and transverse momentum of charged particles.

Around the tracking detectors, but still inside the superconducting coil, are the *Liquid Argon Calorimeter* (LAr)[4,5], which covers the forward and central region and the *"Spaghetti" Calorimeter* [4] (SpaCal)[12], which covers the backward region. Both the LAr and SpaCal calorimeters consist of an electromagnetic and a hadronic

[1] A reconstructed trajectory of a charged particle is called a *track*.

[2] The name *calorimeter* is derived from the Latin word "calor" (heat), and is used as a general term for detectors that reconstruct particle energies by measuring the development of electromagnetic or hadronic showers (see section 4.1).

[3] The origin of the right-handed H1 coordinate system is located at the nominal interaction point where the bunch-crossing of the electron and proton beam takes place. The z-axis is chosen in direction of the proton beam, with the polar angle θ defined with respect to the beam-axis in the proton direction. The x-axis points to the centre of the HERA ring and the y-axis vertically upwards. The azimuthal angle ϕ is defined with respect to the x-axis.

[4] The Spaghetti Calorimeter owes its name to the long thin scintillating fibers it is constructed of (see section 1.2.2).

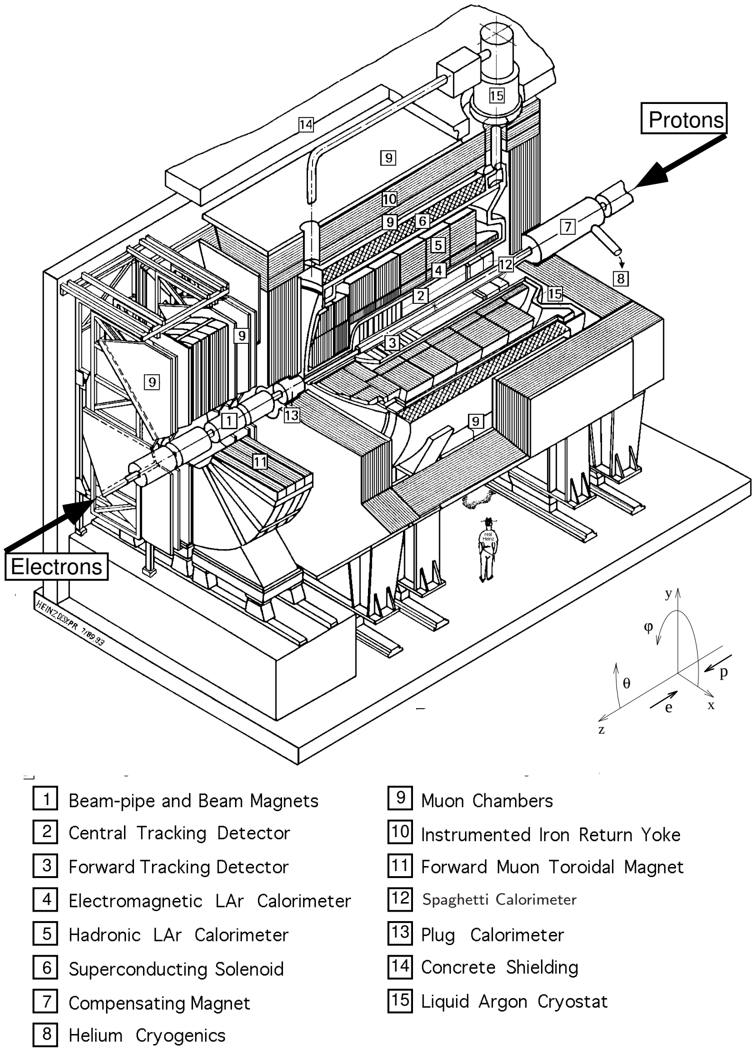


Figure 1.4: *Schematic view of the H1 detector.*

section and provide direction and energy measurements for electrons/photons, and charged/neutral hadrons. For the LAr, a compact design was chosen that fits inside the coil, in order to minimize the amount of dead material between the interaction point and the calorimeter. An additional *Plug* calorimeter[13] is installed around the beam-pipe in the very forward region, to increase the detector acceptance for particles produced at small polar angles.

In the outermost part of the H1 detector, the muon system is installed in the iron yoke[10] that encloses the tracking detectors, the calorimeters and the coil and returns the magnetic field flux. The iron yoke consists of ten iron plates, whose slits have been instrumented with *limited streamer tubes* for muon identification. The anode wires of the streamer tubes are digitally read-out to provide hits from which muon tracks may be reconstructed. In order to provide an estimate of the amount of hadronic energy "leaking" out of the LAr calorimeter, the analog signals on the cathodes are read-out in addition to the digital signals of the anode wires. For the detection of muons produced at small polar angles, the muon system is complemented in forward direction by the *Forward Muon Toroid*[11], a toroidal magnet sandwiched between three layers of drift chambers on either side of the magnet.

Altogether, the various subdetectors allow for particle reconstruction over almost the full solid angle; holes in the geometrical acceptance are mainly due to the beam-pipe, with small contributions arising from feedthroughs for cabling and cooling.

Outside the main H1 detector, additional independent detectors have been installed for the measurement of electrons and hadrons scattered by small angles. In the downstream direction, scattered hadrons are measured by the *Forward Proton Spectrometer* (FPS) [13] at $z_{FPS} = +64 - 90\ m$ [5] the *Proton Remnant Tagger* (PRT) [14] at $z_{PRT} = +24\ m$ and the *Forward Neutron Calorimeter* (FNC) [15] at $z_{FNC} = +20\ m$. For the detection of scattered electrons, three *Electron Taggers* are installed at $z_{ET_8} = -8\ m$, $z_{ET_{33}} = -33\ m$ and $z_{ET_{44}} = -44\ m$ in upstream direction. Furthest away from the interaction point, a *Photon Detector* is installed at $z_{PD} = -102\ m$; it is used in coincidence with the electron tagger located at $z_{ET_{33}} = -33\ m$ to identify Bethe-Heitler processes $ep \rightarrow ep\gamma$, the rate of which is used to measure the luminosity.

The complete *Read-Out* of the H1 detector takes more time (about 1 ms) then the HERA bunch-crossing period of 96 ns, so the entire detector cannot be read-out for each bunch-crossing. For this reason, a multi-level *Trigger-System* is installed, that initiates the read-out only when an interesting event is detected. The trigger system is aided in selecting interesting events by signals provided by two scintillator walls ("Veto-walls") located upstream of the H1 detector at $z_{Veto_8} = -8\ m$ and $z_{Veto_{11}} = -11\ m$. The scintillator walls provide accurate timing information that allows for the rejection of particles produced in collisions of stray protons with the beam-pipe walls upstream of the interaction point, as these particles hit the scintillator walls shortly before those from ep bunch crossings.

In the remainder of this chapter the individual components of the H1 detector are described. Subdetectors are described as they were in the HERA I running period in 1999/2000, as the majority of the data presented in this thesis was taken in that period; modifications of subdetectors that have been upgraded for the HERA II programme are mentioned separately where appropriate. Following the description of the subdetectors, a brief introduction to the computer simulation of the H1 detector is given: as elementary particles cannot be observed directly, the detector simulation is of vital importance for linking the experimental data with theory predictions.

1.2.1 Tracking

The tracking system provides for track triggering and reconstruction. The tracking detectors are located in the innermost part of the H1 detector, so that the momenta of particles produced in the *primary ep* interaction may be measured before the particles loose energy in *secondary* interactions with the detector material.

The H1 tracking system is based on three different detector technologies: *drift chambers*, *multi-wire proportional chambers* and *silicon detectors*. The drift chambers make possible an accurate track reconstruction, while the multi-wire proportional chambers (MWPCs) allow for a fast (but less precise) estimate of the track

[5]The FPS consists of four so-called "Roman POT" stations, installed at $z = 64\ m, 80\ m, 81\ m, and 90\ m$ [13].

parameters for trigger purposes. The silicon detectors improve the track reconstruction of the drift chambers by providing precise measurements of track parameters close to the beam pipe. The different tracking detectors of the H1 tracking system are displayed in figure 1.5. In the forward region, *Planar* and *Radial* drift chambers

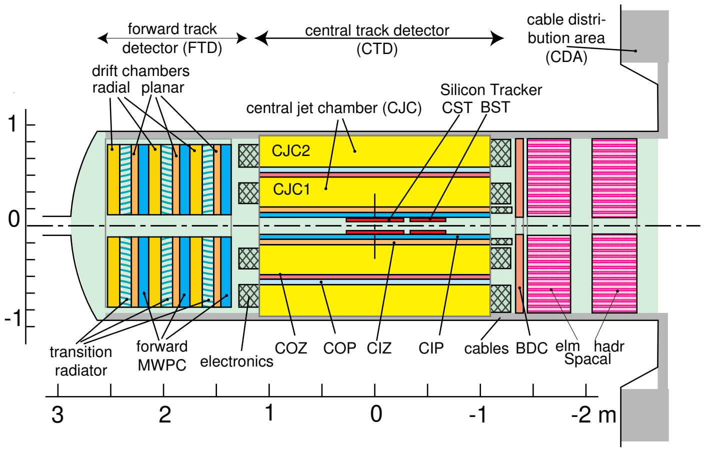


Figure 1.5: *Enlarged view of the H1 tracking system in r-z projection*

are employed for track reconstruction in the FTD. The *Forward MWPCs* provide a trigger signal for charged particles in the forward direction. In the central region, drift chambers are employed in the *Central Jet Chamber* (CJC) and the *Central Inner and Central Outer Z-Chamber* (CIZ and COZ) of the Central Tracking Detector. Trigger signals are provided in the central region by the *Central Inner and Central Outer Proportional Chambers* (CIP and COP). In the backward region, tracks are reconstructed in the BDC. Inside the FTD, CTD and BDC, directly surrounding the beam-pipe, the *Forward Silicon Tracker* (FST), *Central Silicon Tracker* (CST) and *Backward Silicon Tracker* (BST) are installed to improve the track measurements provided by the drift chambers. The FST was installed as part of the HERA II upgrade in 2001 and is therefore not shown in the illustration of the HERA I tracking system in 1.5; the CST and BST were both installed in 1998. In combination, the different tracking detectors allow the direction and momentum of charged particles to be measured in the angular range $5^\circ \lesssim \theta \lesssim 176^\circ$; the geometrical acceptances of the individual detectors are listed in table 1.2.

For drift chambers and multi-wire proportional chambers, the track reconstruction is described in more detail in section A of the appendix. The silicon detectors are not further described, as the FST, CST and BST are not used in the analysis presented in this thesis: the lifetime of tau leptons is too small in comparison to the resolution of the silicon detectors to allow for the reconstruction of the tau decay vertex (see section 4.5).

The Central Track Detector

The central track detector (CTD) shown in the centre of figure 1.5 is the principal tracking device of H1; with an angular coverage of $15^\circ \lesssim \theta \lesssim 165^\circ$, it provides track reconstruction and triggering in most of the solid angle. The CTD consists of six

Detector	θ_{min}	θ_{max}
FST	8°	16°
FTD (planar and radial chambers)	5°	25°
forward MWPCs	6°	20°
CST	$\sim 30°$	$\sim 150°$
CIP	9°	172°
CIZ	16°	169°
CJC	15°	165°
COP	25°	156°
COZ	25°	156°
BST	162°	176°
BDC	153°	176°

Table 1.2: *Geometrical acceptances of the individual tracking detectors in the forward, central and backward region of the H1 detector (values compiled from [16]); the minimum and maximum polar angles of charged particles within the detector acceptances are referred to the nominal interaction point.*

cylindrical subdetectors arranged concentrically around the beam-pipe, as illustrated in figure 1.6. Its two main components are the two central jet chambers CJC1 and

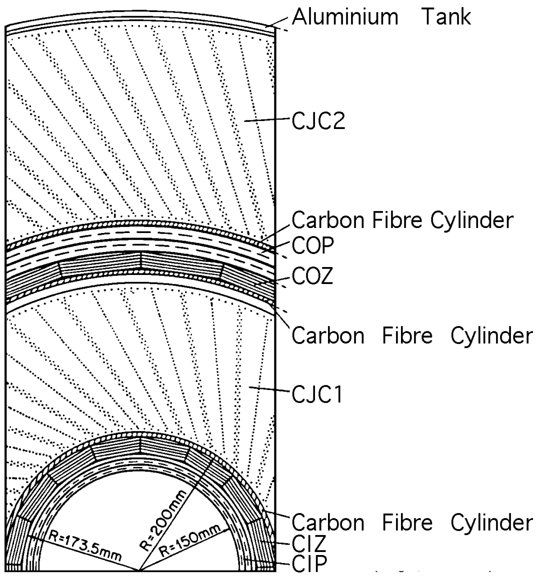


Figure 1.6: *Cross-sectional view of the central track detector.*

CJC2, which provide an accurate reconstruction of the transverse momentum and azimuthal angle of charged particles. For a precise reconstruction of the polar angle, the central jet chambers are supplemented by the CIZ and COZ, which provide a complementary measurement of the z-coordinate at two different radii. Last but not least, the CTD includes the CIP and COP, that provide a fast reconstruction of the event vertex for trigger purposes.

The CJC1 and CJC2 [17] are two large cylindrical drift chambers. Their drift

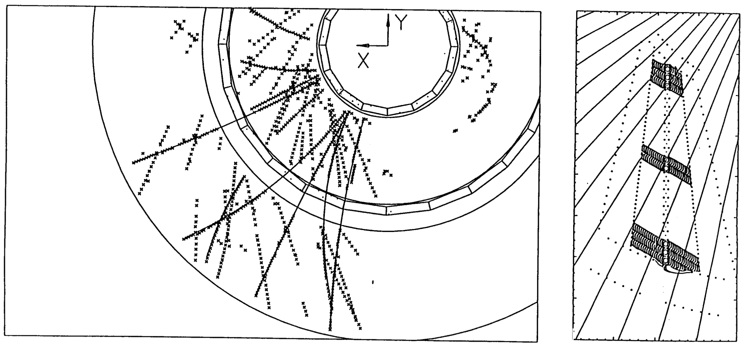


Figure 1.7: *Illustration of mirror tracks in the central track detector.*

volume is segmented in the azimuthal direction into independent trapezoidal drift cells (30 in CJC1 and 60 in CJC2). In the centre of each drift cell, a plane of anode wires is located, bounded by two planes of cathode wires on either side. In the radial direction, a total of 56 anode wires, oriented in the z-direction, is strung within each drift cell, providing accurate track reconstruction and a precise determination of the energy loss dE/dx by ionisation. From the drift-time measured by the anode wires, single hits are reconstructed with a spatial resolution of $\sigma_{r\text{-}\phi} \approx 170\ \mu m$ [16] in the r-ϕ plane. Along the anode wires a resolution of $\sigma_z \sim 30\ mm$, of the order of 1% of the wire length, is attained by charge division. By a combined measurement of all anodes wires, the energy loss of an ionizing particle is reconstructed with a relative uncertainty of about 10% [16].

The anode wire planes are inclined with respect to the radial direction by an angle of 30°, which improves the track reconstruction in three ways. Firstly, the tilt of the anode planes approximately matches the *Lorentz angle* by which electrons drifting towards the anode wires are deflected in the presence of the solenoidal magnetic field. As a result of the matching of the inclination and the Lorentz angle, the secondary electrons drift perpendicular to the anode wire plane, yielding an optimal track resolution. Secondly, the inclination helps to reject *mirror tracks* resulting from the left-right ambiguity inherent in drift chambers. As illustrated in figure 1.7, the tilted anode wire planes cause mirror track segments to neither link with track segments in neighboring drift cells nor to point to the event vertex, allowing for mirror tracks to be easily identified. Thirdly, the inclined anode wire planes are crossed at least once by all particles of sufficiently high momentum. This allows the time at which the particle passed through the chamber to be determined with an accuracy of $\sigma_t \approx 0.5\ ns$ [16] (no uncertainties due to electron drift), providing a very powerful rejection of non-ep background (see section 6.3).

The CIZ and COZ are two complementary drift chambers that improve the track reconstruction in the r-z plane. The CIZ takes the form of a regular 16-sided polygon in the transverse plane, with four layers of anode wires strung parallel to the edges of the polygon at different distances to the beam-pipe. With the four wire planes, trajectories of charged particles traversing the CIZ are measured with a spatial resolution of $\sigma_z \approx 320\ \mu m$ [16] in the z-direction. The COZ is of similar design, except that it forms a 24-sided polygon to better approximate a circular shape at its larger radius. In the COZ, the spatial resolution in the z-direction depends on the polar angle of the traversing charged particle and is between $\sigma_z \approx 200 - 500\ \mu m$.

The CIP and COP are located next to the z-chambers. Both chambers contain two layers of anode wires strung in the z-direction, interleaved with cathodes that are segmented into pads in the ϕ- and z-direction. In order to provide a fast estimation of the vertex position for trigger purposes, the cathode pads are read-out. As for trigger purposes the estimated position of the vertex along the beam-pipe is most important, the segmentation is finer in the z- than in the ϕ-direction, with 8(16) pads in ϕ- and 60(18) pads (of 36.6(120) mm length) in the z-direction for the CIP(COP).

For the HERA II running period, CIP and CIZ have been replaced by a new central inner proportional chamber with 5 cathode layers (segmented into pads of 9 mm length) and additional anode read-out [18].

The Forward Track Detector

The forward track detector (FTD) [19], shown on the left hand side of figure 1.5, covers the angular range $5° \leq \theta \leq 25°$. The detector is designed to provide track information on individual particles within jets in events with high particle multiplicities in the forward region [16]. Since charged particles produced in the forward direction no longer traverse the full chamber radius in the bending r-ϕ plane, a high radial density of instrumentation is necessary for precise track reconstruction. As a consequence, the FTD is constructed as an integrated assembly of different drift chamber types with layers of anode wires strung in the r-ϕ plane closely spaced in beam direction. In addition to *planar* and *radial* drift chambers, multi-wire proportional chambers and *transition radiators* are added for trigger purposes and particle identification, respectively. The individual chamber modules are organized in three nearly [6] identical *supermodules* that contain (in the direction of increasing z) three planar drift chambers, each rotated by 60° in azimuth with respect to the others, one *forward proportional chamber* (FPC), one transition radiator, one radial drift chamber. The assembly is illustrated in figure 1.8. The three planar drift chambers are positioned most closely to the interaction point, as their homogenous resolution in the r-ϕ plane is most suitable for linking to track segments reconstructed in the CTD. In order to maximize the geometrical acceptance of the trigger, the FPC is mounted directly behind the planars. The transition radiator is inserted between the FPC and the radial chamber to enable radial chambers to measure the transition radiation emitted. In the following, the individual chamber types are described in more detail. A brief explanation of transition radiation is also given.

The planar chambers each contain 4 planes of anode wires perpendicular to the beam direction. Each plane is divided into 32 rectangular drift cells. Coordinates perpendicular to the anode wires are measured with a spatial resolution of $150 - 170\ \mu m$ for single hits. The coordinate along the wires cannot be reconstructed as signals are read-out from one wire end only.

The forward proportional chambers contain 2 planes of anode wires interleaved by 3 cathode planes. The cathode planes are segmented in the radial and azimuthal

[6] The radial chambers in the second and third supermodule are rotated in azimuth by 3.75° and 2.5° with respect to the radial chamber in the first supermodule, in order to improve the double track resolution.

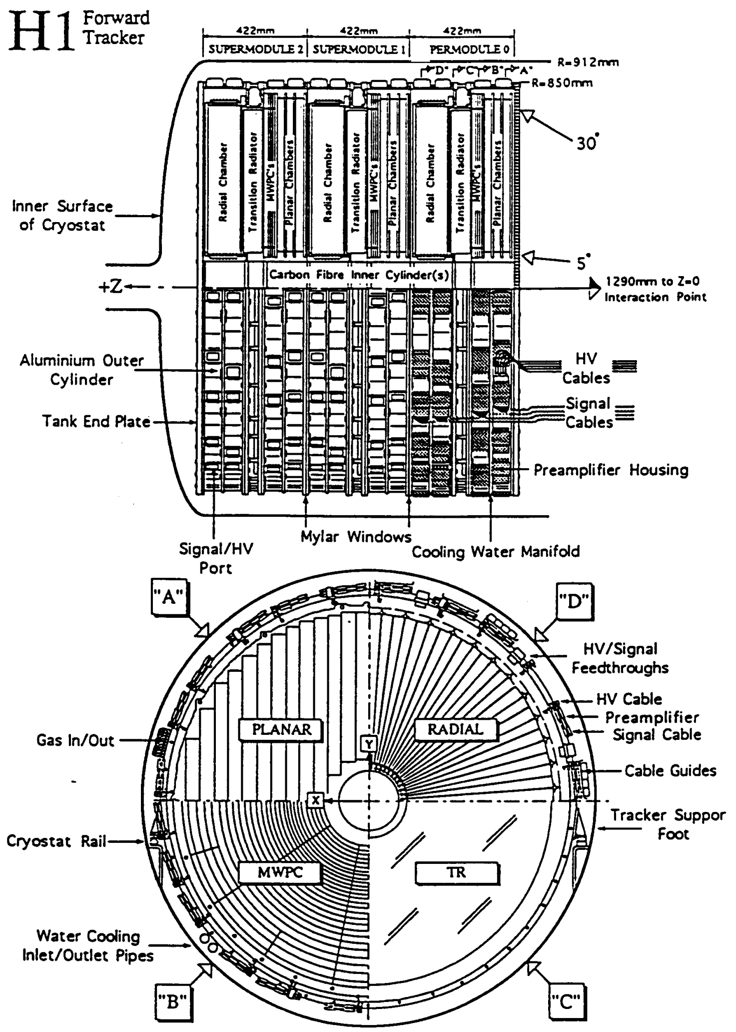


Figure 1.8: *The Forward Track Detector. Top: Longitudinal sectional view, showing the structure of supermodules. Bottom: Cross-sectional view, showing the orientation of anode wires in the planar (A) and radial (B) drift chambers and the forward proportional chambers (D).*

directions into ring-shaped cathode pads. In the radial direction, each inner plane is divided into 20 rings, each ring containing 16 pads in the azimuthal direction. The 4 outermost rings are divided into 32 pads each. The rings in two consecutive cathode planes are offset by half a ring width, to increase the effective angular resolution.

The transition radiator is a passive module (no read-out) built of 400 polypropylene foils. It allows for the separation of energetic electrons from pions by using the transition radiation emitted by charged particles crossing the boundaries between the polypropylene foils and the chamber gas [7].

[7]Transition radiation is emitted by charged particles crossing a boundary between two media of different dielectric constant. For a particle passing from vacuum to a dense medium, the transition radiation may be thought of as being emitted by an electric dipole, formed by the particle and its mirror charge in the dense medium. As the particle approaches the boundary between the two media, the electric field strength of the dipole diminishes - and finally vanishes, when the particle

The number of photons emitted in the transition radiation module is proportional to the number of foil/gas boundaries crossed, while the mean photon energy increases linearly with the *Lorentz factor* $\gamma = \frac{E}{m}$ of the charged particle (in contrast to the energy loss dE/dx by ionisation for example, that increases only logarithmically with the particle energy), which allows for particle identification at high energy. For particle energies of the order of a few GeV, the spectrum of emitted photons peaks in the X-ray domain.

The radial chambers contain 12 planes of anode wires strung radially in the transverse plane. Each wire plane is divided in the azimuthal direction into 48 wedge-shaped drift cells. Single hits are measured with a resolution of $150 - 200\ \mu m$ in azimuth, while radial coordinates may be estimated with coarse resolution by charge division. In order to measure the transition radiation produced in the adjacent transition radiator, an X-ray transparent mylar window is fitted to the chambers on the side facing the nominal interaction point.

A disadvantage of the presented design of the forward track detector is the relatively high amount of dead material that particles in the forward region have to traverse. A total of about 1.1 radiation lengths (see section 4.1.2) of dead material is located between the CTD and FTD (about 0.5 X_0 in chamber end-walls, electronics and cables) and inside the FTD (about 0.6 X_0 in the FPCs and transition radiators), resulting in a probability of $p\,(interaction) \sim e^{-1.1} \approx 70\%$ for electrons and photons to shower before or in the FTD [21]. The amount of dead material in the forward track detector has been significantly reduced (to about 0.3 X_0) for the HERA II running period, by replacing the FPCs (which were unlikely to survive the full HERA II luminosity until end of operation) and the transition radiators (whose emitted transition radiation was spoiled by showering to the extent that it was never really used in analyses) by a set of 5 new planar drift chambers (2 in supermodules 0 and 1 and 1 in SM2). Each of the new planars contain 8 planes of anode wires perpendicular to the beam direction, contributing additional space-points to the track reconstruction for increased track finding efficiency and improved track resolution.

The Backward Drift Chamber

The backward drift chamber (BDC) [22] shown in the right hand side of figure 1.5 is mounted directly in front of the SpaCal calorimeter, allowing precise reconstruction of the polar angle of the scattered electron. The detector is constructed from 8 layers of drift chambers displaced in the z-direction. The individual drift chamber layers are of octagonal shape with anode wires strung parallel to the edges of the octagon. In the radial direction, each layer contains 32 wires. The layers are paired, with the anode wires of the second layer shifted by half a wire separation with respect to the first, to resolve left-right ambiguities. With this configuration, the drift electrons move in an almost radial direction, which guarantees an optimal resolution for the reconstruction of the polar angle. To make possible the reconstruction of the azimuthal angle, the four pairs of layers are rotated by $11.25°$ around the z-axis with respect to each other.

---

crosses the boundary. In the transition, electromagnetic radiation is emitted by the particle, as a result of the changing dipole field [20].

1.2.2 Calorimetry

Calorimeters provide an energy measurement for all particles excepting neutrinos and muons of energies above a few GeV, complementing the momentum measurements provided by the tracking detectors for charged particles. In particular, calorimeters allow the reconstruction of the energy and direction of neutral hadrons and photons, which cannot be detected by tracking chambers. In addition, calorimeters improve the reconstruction of highly energetic charged particles: the relative energy resolution of calorimeters typically improves as $\frac{\sigma_E}{E} \sim \frac{1}{\sqrt{E}}$ with the energy E of the incident particles, in contrast to tracking detectors, whose momentum resolution decreases as $\frac{\sigma_{p_T}}{p_T} \sim p_T$ (*cf.* equation A.1).

The energy reconstruction in calorimeters is based on total absorption methods. Incident particles are stopped in the calorimeter volume by interactions with the atoms of the detector material (described in more detail in section 4.2), in which they finally lose all their energy. This energy is then measured - hence the term "calorimeter" [23]. The calorimeters in the H1 detector are constructed of alternating layers of *passive absorber* material and *active sampling* medium.

The function of the absorber layers is to absorb the energy of the incident particle. For this purpose the absorber layers are built of dense materials that have a small mean path length between subsequent interactions of incident particles, so as to stop the incident particle after a short distance and to contain the resulting electromagnetic and hadronic *showers* (see section 4.2) in a small volume, making possible more compact calorimeters. As the cross-section for electromagnetic interactions is proportional to the atomic charge number Z^2, while the rate of hadronic interactions is proportional to the atomic mass A, high Z (A) materials are chosen for the electromagnetic (hadronic) sections of calorimeters. In the longitudinal direction, the development of electromagnetic (hadronic) showers is described by the *radiation length* X_0 (*hadronic interaction length* λ). Lead, the material most often used as absorber in electromagnetic calorimeters, has a radiation length of $X_0 \approx 0.56\ cm$. For hadronic calorimeters often either lead or iron/steel are used, which both have a hadronic interaction length of $\lambda \approx 17\ cm$. Since λ is about one order of magnitude larger X_0 for most materials, hadronic showers are of significantly larger longitudinal extent than electromagnetic showers. As a consequence, much deeper calorimeters are necessary to contain hadronic showers, and electromagnetic calorimeters are placed before hadronic ones, closer to the interaction point.

The energy absorbed in the absorber layers is measured in the sampling layers. For the reconstruction of the absorbed energy, several different methods exist, including liquid-filled ionisation chambers and detectors based on scintillation light. The liquid-filled ionisation chambers are used to count the number of charged shower particles entering the sampling layer from the absorber layer. In comparison to the gas-filled ionisation chambers used for the construction of tracking detectors, liquid-filled detectors have the advantage of an about 10^3 higher density, leading to an about 1000 times larger energy deposit of the shower particles and a correspondingly greater signal in the sampling layers. An alternative technique is based on the registration of scintillation light. In these detectors, the shower particles are counted by measuring the light emitted by a scintillating substance, when atoms that have been excited by shower particles traversing the sampling layer revert back to their ground state. From the number of shower particles counted in the sampling layers, the energy absorbed in the absorber layers and thus the energy of the incident particle may be reconstructed. As is derived in section 4.2, the energy of the incident particle is proportional to the total path length of all shower particles in the calorimeter. The total path length s is estimated by summing the number of shower particles n_i counted in the sampling layers, and extrapolating to the combined thickness of sampling and absorber layers,

$$s \approx (d_{absorber} + d_{sampling}) \cdot \sum_i n_i = (d_{absorber} + d_{sampling}) \cdot N.$$

For the shower model described in section 4.2, the energy of the incident particle is then reconstructed as [24]

$$E \approx \frac{E_c}{2} \cdot s = \frac{E_c}{2} \cdot (d_{absorber} + d_{sampling}) \cdot N \sim N, \tag{1.2}$$

where E_c is the *critical energy*, the average energy of shower particles at which the shower development effectively stops.

In most calorimeters, the *response* to incident hadrons is lower than that to incident electrons or photons of the same energy. This occurs as about 30% of the energy of hadronic showers is on consumed in nuclear excitation and break-up and by the production of short-ranged nuclear fragments in the absorber layers that do not reach the sampling layers [24] and are thus hidden from detection. A further fraction of shower energy may escape detection if pions and kaons in the hadronic shower decay to muons and neutrinos. The invisible energy can only be corrected for on average in *compensating* calorimeters. One way of achieving this is to build the calorimeter using uranium for the absorber layers. In calorimeters containing uranium, the energy loss from nuclear excitations and break-up is compensated by a corresponding gain resulting from the spallation of the uranium atoms [8]. In *non-compensating* calorimeters, the energy of hadronic showers has to be corrected for by the reconstruction software (see the description of the Liquid Argon Calorimeter).

Uncertainties in the energy reconstruction arise from an imprecise estimation of the integrated path length s of the shower particles and from fluctuations of the energy at which individual shower particles stop producing further particles. The uncertainties in the estimation of the integrated path length are due to different effects. Above all, *sampling fluctuations* arise from variations in the energy loss of the shower particles in the absorber and sampling layers. As the energy loss in the absorber layers is not measured directly, but is inferred from the number of shower particles reaching the sampling layers, which is described by a Poisson distribution, the sampling fluctuations contribute to the energy resolution as

$$\sigma_E \sim \sigma_N = \sqrt{N} \sim E,$$

using relation 1.2. Additional uncertainties in the energy reconstruction arise from longitudinal or lateral *leakage* of energy carried by shower particles escaping detection, if showers are not fully *contained* in the calorimeter. In hadronic showers, the

[8] This technique is employed in the calorimeter of the ZEUS experiment.

uncertainties in the energy reconstruction are further increased by statistical fluctuations of the energy consumed in nuclear excitation and break-up. In electromagnetic calorimeters, the dominant uncertainty in the reconstruction of the shower energy is most often due to sampling fluctuations. Accordingly, the relative energy resolution of electromagnetic calorimeters may be parametrized to a good approximation as

$$\frac{\sigma_E}{E} = \frac{const}{\sqrt{E}},$$

with a construction dependent proportionality constant of the order of $10 - 20\%$. In hadronic calorimeters, the uncertainties introduced by nuclear excitation and break-up significantly worsen the resolution. Nevertheless, the energy resolution of hadronic calorimeters can still be parametrized by the form $\frac{\sigma_E}{E} = \frac{const}{\sqrt{E}}$, albeit with a larger construction dependent proportionality constant of the order of $50 - 100\%$.

In the H1 detector, calorimetry is provided by four individual detectors. The main component is the *Liquid Argon* calorimeter (LAr), which covers the forward and central region. The LAr is complemented in the backward direction by the *Spaghetti* calorimeter (SpaCal) and by the plug calorimeter in the very forward region. Energy leaking out of the Liquid Argon calorimeter is registered in the hadronic *Tail Catcher*, installed in the instrumented iron return yoke as part of the muon system.

The Liquid Argon Calorimeter

The *Liquid Argon* calorimeter (LAr) is the principle calorimeter in the H1 detector, covering an angular range of $4° \leq \theta \leq 153°$. It is the main detector for the reconstruction of the hadronic final state. In addition, the LAr provides excellent e/π separation at the level of 10^{-3} pion misidentification probability for an electron identification efficiency of 95% [25].

The Liquid Argon calorimeter is composed of electromagnetic and hadronic sections. The electromagnetic calorimeter is constructed from 2.4 mm thick lead plates as absorber, interspersed by gaps of width 2.35 mm filled with liquid argon as sampling medium. For the hadronic section, 16 mm thick stainless steel plates were chosen as absorber layers, alternating with layers of liquid argon filled gaps of twice 2.4 mm width. Shower particles crossing the sampling layers are detected by the ionisation of the argon atoms, read-out by rectangular cathode pads. The total depth of the absorber material varies between $20 - 30\ X_0$ in the electromagnetic section and $4.5 - 8\ \lambda$ for the combined electromagnetic and hadronic sections [26]. Both electromagnetic and hadronic sections are deeper in the forward region, where the energy of incident particles is generally higher due to the asymmetric beam energies. The electromagnetic and hadronic sections are segmented into about 45 000 read-out *cells* (displayed in figure 1.9), providing a good spatial resolution for the determination of shower shapes. In the longitudinal direction, the LAr is segmented into 3-4 layers of cells in the electromagnetic and 6-8 cell layers in the hadronic section. In the transverse direction, cell sizes vary between $5 \times 5 - 10 \times 10\ cm^2$ in the electromagnetic and $10 \times 10 - 20 \times 20\ cm^2$ in the hadronic part. The cell granularity is finer in the electromagnetic layers, so as to better resolve the more compact electromagnetic showers, and in the forward region, owing to the higher particle densities in that region.

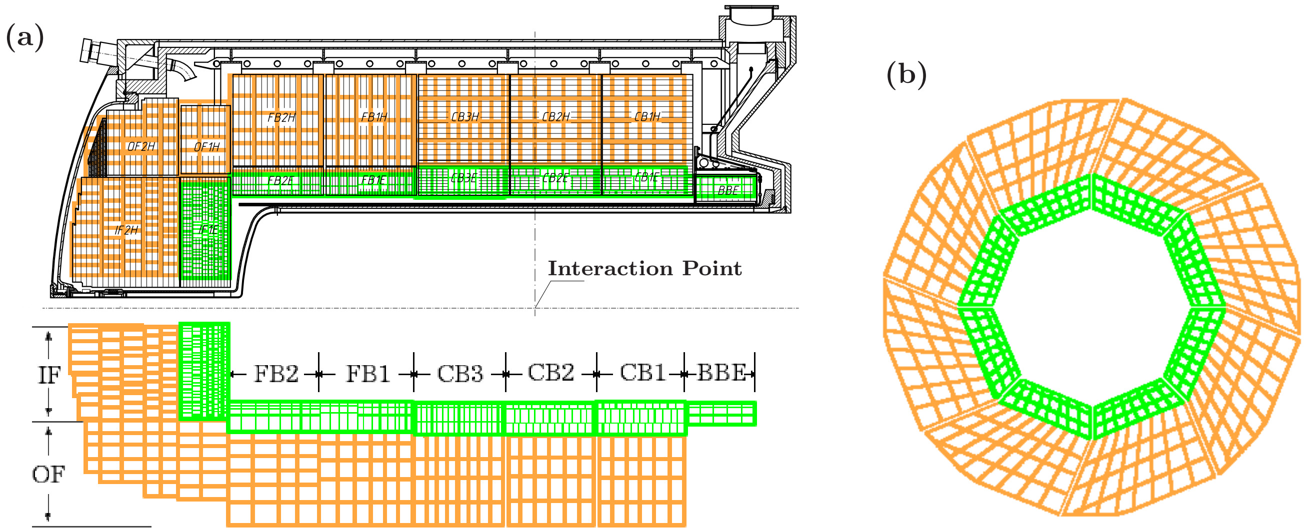


Figure 1.9: *The Liquid Argon Calorimeter. (a) Longitudinal sectional view, showing orientation of the absorber layers in the upper and the segmentation into read-out cells in the r-z plane in the lower half. (b) Cross-sectional view, showing the segmentation into read-out cells in the r-ϕ plane.*

For easier handling, the LAr calorimeter is divided into 8 *wheels* that are indicated in figure 1.9, the *Backward Barrel* (BBE), three *Central Barrels* (CB1, CB2, CB3), two *Forward Barrels* (FB1, FB2) and the *Inner* and *Outer Forward* (IF, OF) calorimeter modules. In the radial direction, all wheels except the BBE and OF are composed of two *stacks* for the electromagnetic and hadronic sections, respectively. The BBE consists of one electromagnetic stack only and the OF of two hadronic stacks. In the azimuthal direction, each wheel is divided into 8 octants. Separating the wheels and octants, as indicated by the gaps in figure 1.9, are insensitive regions, which are problematic for the event reconstruction, as there are energy losses in these regions. The gaps between the wheels are called *z-cracks*, those between the octants *ϕ-cracks*.

The energies deposited by incident particles in the electromagnetic and hadronic cells are reconstructed in several steps by the LAr reconstruction software.

An important first step is *noise suppression*. Although the noise levels of individual cells are low, between $10 - 40$ MeV [26] (increasing with the cell size), the noise in all 45 000 LAr cells typically adds up to about 40 GeV [27] in an "empty" event [9] . The basic idea of the noise suppression algorithm is to require a *localized* energy deposit several standard deviations above the noise level in at least two neighboring cells. This rejects single noisy cells as well as noise adding up from small contributions of a large number of cells. After noise suppression, the residual noise is reduced to an average of 0 GeV with a statistical variation of 0.25 GeV [27].

In the second step of the energy reconstruction, neighboring cells that have not been rejected as noise are associated to *clusters*: groups of cells that are likely to contain the shower of one and the same incident particle. The clustering algorithm works quite well for compact showers induced by electrons and photons.

[9] "Empty" events are recorded by random triggers triggering on non-colliding bunches.

It was found, that 95-97% (depending on the energy of the incident particle) of electromagnetic showers in the LAr are reconstructed as a single cluster, while simultaneously resolving pairs of electrons into two separate clusters down to opening angles of about $2°$(IF) $-$ $5°$(FB, CB) between the two electromagnetic showers [28]. The clustering algorithm does not work so well for hadronic showers, however. As a result of the broader and more fluctuating shower shape, hadronic showers induced by single hadrons are often reconstructed as several clusters.

The clusters found are then classified in the third step as either belonging to electromagnetic or hadronic showers, depending on the compactness of the cluster and on the position at which the shower started [25, 27]: an early shower start in the first electromagnetic layer of the calorimeter indicates the shower is induced by a photon or electron. Additionally, the cluster energies are corrected for energy loss in the dead-material in front of the calorimeter (between $1 - 2\ X_0$, varying with polar angle [26]).

In the fourth step, identified hadronic clusters are subject to an *energy weighting* algorithm [29, 30], that has been developed to equalize the response of the non-compensating LAr calorimeter to electromagnetic and hadronic showers. The fine granularity of the LAr makes it possible to detect the electromagnetic components of hadronic showers, which are induced by neutral pions and indicated by high local energy densities. Note that the energy contained in the electromagnetic components of showers must not need to be corrected. By applying individual energy correction factors only to those cells in the hadronic shower that are not associated with electromagnetic subshowers, the influence of the reconstructed energy on variations in the number of neutral pions in the hadronic shower is lessened.

Final corrections to the reconstructed energies are applied by a *calibration* on the particle level. The *electromagnetic energy scale* is calibrated by comparing the energy deposit of electrons reconstructed after dead-material corrections with the expectation determined by the so-called *double-angle* method [10] [31] in neutral current DIS events (see section 2.1). The *hadronic energy scale* is calibrated using neutral current DIS events as well, by requiring the transverse momentum of the hadronic final state to balance that of the calibrated scattered electron [32]. The energy resolution of the LAr calorimeter has been determined to be

$$\frac{\sigma_E^{em}}{E} = \frac{0.15}{\sqrt{E}}$$

for electrons [31] and

$$\frac{\sigma_E^{had}}{E} = \frac{0.50}{\sqrt{E}}$$

for single charged pions [33].

The "Spaghetti" Calorimeter

The *"Spaghetti"* calorimeter (SpaCal) is installed in the backward region of the H1 detector as shown in figure 1.10 and covers the angular range $153° \leq \theta \leq 177.5°$ [11].

[10] Neutral current DIS events are kinematically over-constrained, such that the event kinematics can be fully reconstructed from the polar angles of the scattered electron and the hadronic final

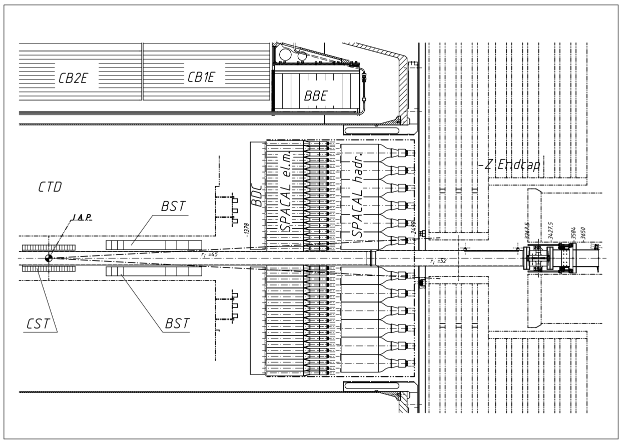


Figure 1.10: *The location of the "Spaghetti" Calorimeter in the backward region of the H1 detector.*

The primary function of the SpaCal is the detection of electrons scattered through low angles. Additionally, it allows for the measurement of backward jets.

Like the LAr calorimeter, the SpaCal is a non-compensating sampling calorimeter and consists of an electromagnetic and hadronic section. Both sections are constructed of long thin scintillating fibers (giving rise to the term "Spaghetti" calorimeter), aligned parallel to the beam direction and embedded in a lead matrix. In the lead, incident particles induce electromagnetic and hadronic showers that are detected by the scintillation light emitted by the fibers when charged shower particles excite molecules in the scintillator material. The scintillation light is collected at the backward end of the fibers and guided to *photomultiplier* tubes, that amplify the light and convert it to electric signals (see reference [24] for a more detailed description of photomultipliers). The electric signals provided by the photomultiplier are then read-out to reconstruct the energy of the incident particles.

The electromagnetic section is built of 250 mm long fibers of 0.5 mm diameter embedded in lead plates with a lead to fiber ratio of 2.3:1 (in volume), providing a depth of 28 X_0. For the read-out, the electromagnetic section is segmented into 1992 cells of size $4.05 \times 4.05\ cm^2$ in the transverse plane. The fine granularity allows for an e/π separation of better than 10^2 [34].

The hadronic section is made of 246 mm long fibers of 1.0 mm diameter embedded in a lead matrix with a lead to fiber ratio of 3.4:1 (in volume), corresponding to 1.08 λ in depth. In combination, the electromagnetic and hadronic sections of the SpaCal calorimeter provide a depth of about 2 λ. For the read-out, the hadronic section is segmented into 136 cells of size $11.93 \times 11.90\ cm^2$.

In test-beam studies, the energy resolution of the SpaCal was determined to be

$$\frac{\sigma_E^{em}}{E} = \frac{0.07}{\sqrt{E}}$$

state, yielding an expectation for the energy of the scattered electron that may be compared with the shower energy reconstructed in the LAr.

[11] The SpaCal replaced the former *Backward Electromagnetic* calorimeter (BEMC) in 1995. As only a small fraction of the data used in the analysis presented in this thesis was taken in 1994, the BEMC is not described here. A detailed description of the BEMC may be found in [16].

for electrons [35] and

$$\frac{\sigma_E^{had}}{E} = \frac{0.29}{\sqrt{E}}$$

for single charged pions [34]. The time resolution of the SpaCal has been determined with the same test-beam setup. With a resolution of 0.4 ns, the SpaCal provides a precise *time-of-flight* measurement, which may be used for the rejection of non-ep background that is not synchronized with nominal bunch crossings (see section 6.3).

For the HERA II luminosity upgrade, a new focusing magnet was installed around the beam-pipe in the backward region of the H1 detector. To make room for the magnet, some SpaCal cells had to be removed. As a consequence, the angular acceptance of the SpaCal calorimeter is reduced to about $153° \leq \theta \leq 174°$ in the HERA II running period [36].

The Plug

The *Plug* calorimeter covers the angular range $0.7° \leq \theta \leq 3.4°$, closing the gap in acceptance between the beam-pipe and the LAr calorimeter. Its main task is to minimize the undetectable transverse momentum carried away by the proton remnant.

Owing to geometrical constraints, a very compact design was chosen for the Plug calorimeter. It is constructed as a sampling-calorimeter with 9 copper plates of width 5 cm as absorber layers, providing a depth of 4.25 λ for the reconstruction of hadronic showers. The absorber layers are interleaved by 8 sampling layers, equipped with 84 mainly quadratic (5×5 cm^2) silicon detectors each. The silicon detectors of two subsequent sampling layers are read-out together, resulting in 336 cells. The energy resolution of the plug suffers from lateral and longitudinal leakage and is about

$$\frac{\sigma_E}{E} = \frac{1.50}{\sqrt{E}}.$$

The available space for the plug calorimeter is reduced in the HERA II running period by a new focusing magnet installed around the beam-pipe in the forward region of the H1 detector. A new plug calorimeter [37] of 1.5 hadronic interaction lengths in depth was built, made from 4 layers of copper absorber plates of width 6.5 cm, sandwiched between 5 sampling layers of scintillator tiles that are read-out by photomultipliers.

The Tail-Catcher

The *Tail Catcher* is located in the instrumented iron surrounding the LAr and SpaCal calorimeters. Its function is to measure the energy leaking out of the LAr and SpaCal in the tails of hadronic showers, hence the name “Tail-Catcher”.

The energy measurement of the Tail-Catcher is derived from the read-out of cathode pads on the limited streamer tubes in the instrumented iron (see section 1.2.3), that provides a depth of about 4 λ. Analog signals from five inner and six outer pad layers are combined into two *towers*, providing a two-fold longitudinal segmentation. The measured energies are corrected for the dead-material (mainly the coil) between the LAr and the Tail-Catcher. The amount of dead-material varies between $0.5 - 2.4\ \lambda$, depending on the polar angle. The energy loss in the dead-material is estimated by a linear interpolation of the energy deposits in the last layer of the LAr and the first layer of the Tail-Catcher [16]. The energy resolution of the Tail Catcher has been determined to be

$$\frac{\sigma_E}{E} = \frac{1.00}{\sqrt{E}}$$

in test-beam measurements.

1.2.3 Muon Detection

The detection of muons is based on the registration of charged particles in ionisation chambers surrounding the calorimeters. As muons do not induce particle showers when traversing the LAr and SpaCal calorimeters, and lose little energy due to ionisation (see section 4.1), muons with energies above $2-3\ GeV$ typically penetrate the H1 detector. The ionisation chambers built for the detection of muons are sandwiched between iron plates, so as to shield them against hadrons leaking out of the calorimeters [12]. In the H1 experiment, muon detection is provided by the *Central Muon Detector* and *Forward Muon Detector*.

Central Muon Detector

The *Central Muon Detector* provides muon detection in the angular range $4^\circ \leq \theta \leq 172^\circ$. It is installed in the iron return yoke of the solenoid magnet and is subdivided into three parts, the *forward* and *backward end-caps* in the forward and backward directions and the *barrel* in the central region.

The iron yoke is built of 10 iron plates of 7.5 cm thickness each. The plates are separated by slits, instrumented with layers of limited streamer tubes (LST) for the detection of penetrating particles.

The limited streamer tubes are built as profiles containing 8 cells of size $1 \times 1\ cm^2$. The design of the profiles is illustrated in figure 1.11. The profiles are composed of

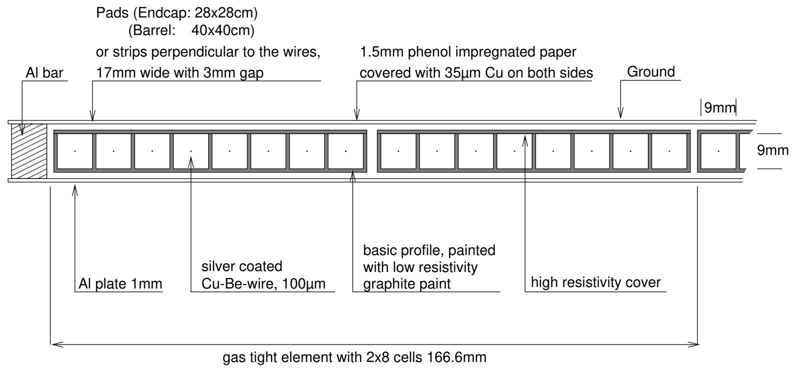


Figure 1.11: *Cross-sectional view of the structure of two profiles of limited streamer tubes.*

[12]While minimum ionizing muons loose only about 1 GeV of energy when traversing 1 m of iron, the same thickness of iron provides a shielding of more than 5 λ in depth for hadrons [24].

two parts, extruded separately from the halogen-free plastic material Luranyl [38]: an eightfold "U"-shaped base and a plane cover. The base is coated with graphite paint to enable a high voltage to be applied to the profiles. The centre of each cell contains a single anode wire. The cells are gas-filled and a negative high voltage is applied to the graphite coated base, whereas the anode wires are operated at ground potential.

In individual layers of limited streamer tubes, either *strips* or *pads* are glued as cathodes on the plane Luranyl cover. The strips are 17 mm wide and oriented perpendicular to the anode wires, with a spacing of 3 mm between neighboring strips. Together with the anode wires, the strips are read-out digitally to provide two-dimensional space points in the LST layers. For single hits, the strips provide a spatial resolution of $10-15$ mm in the direction perpendicular to their orientation, complementing the spatial resolution of the anode wires, that is $3-4$ mm in direction orthogonal to the wire direction [16]. The size of the pads varies between 30×30 cm^2 in the endcaps and 50×40 cm^2 in the barrel. The main function of the pads is to estimate the energy leaking out of the LAr and SpaCal calorimeters (see section 1.2.2). Accordingly, the pads are provided with analogue read-out. They allow a coarse spatial resolution of the order of 10 cm.

In total, the iron yoke is instrumented with 16 layers of limited streamer tubes, 11 of which are equipped with pads and 5 with strips. The orientation of the anode wires is along the z-axis in the barrel and parallel to the x-axis in the end-caps. In front of the 10 iron plates, a so-called *muon-box* is installed, containing two LST layers with strips and one with pads. Single pad layers are installed in the first three and five last slits. The fourth slit is twice as wide as the others and contains one strip and one pad layer. Behind the iron, another muon-box with two strip and one pad layer is installed, to detect particles penetrating all of the iron. The layout of the strip and pad layers is illustrated for the barrel part of the instrumented iron in figure 1.12.

Forward Muon Detector

The *Forward Muon Detector* (FMD) [39] is installed in the downstream direction outside the main H1 detector, and covers the angular range $3° \leq \theta \leq 17°$. It is built of 6 double-layers of drift chambers, mounted on either side of the *Forward Muon Toroid*.

The drift chambers provide independent track reconstruction for charged particles penetrating the forward iron end-cap, complementing the measurement of the FTD in the main H1 detector. The size of the drift-chambers increases with the distance from the interaction point from about 4 m to 6 m in diameter. The drift volume of the chambers is segmented into drift cells of 20 mm depth and 120 mm width, with a single anode wire positioned in the centre of each cell. From the measured drift times, single hits are reconstructed with a precision of about 250 μm in the direction perpendicular to the wires, while the position of hits along the wires is determined by charge-division with an accuracy of the order of 4 cm [3]. In order to resolve left-right ambiguities, the drift chambers are paired, with one drift chamber offset by half a cell width with respect to the other in each double-layer.

The double-layers are built in two different designs, as θ-layers with wires strung

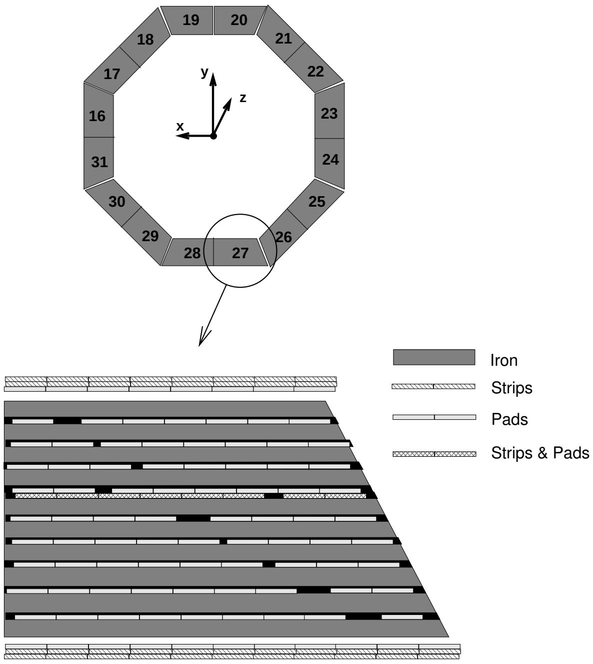


Figure 1.12: *Cross-sectional view of the instrumented iron in the barrel of the H1 detector. The figure illustrates the locations of the strip and pad layers; areas not instrumented with limited streamer tubes are marked black. The numbers numbers shown in the upper half correspond to the segmentation of the instrumented iron into independent modules for trigger purposes (see section D.2 of the appendix).*

concentrically around the beam-pipe to provide a precise measurement of the polar angle, and ϕ-layers with wires strung in the radial direction for a precise measurement of the azimuthal angle. In the θ-layers, the momentum of penetrating charged particles can also be reconstructed from the bending in the toroidal magnetic field. Of the 6 double-layers in the FMD, 4 are built as θ-layers and 2 as ϕ-layers, arranged in 2 groups of 3 double-layers on either side of the toroid magnet, as displayed in figure 1.13.

The Forward Muon Toroid sandwiched between the drift chambers is made of solid iron, 1.20 m in depth. It provides a toroidal magnetic field that varies from 1.75 T at its inner radius of 0.65 m to 1.50 T at its outer radius of 2.90 m.

As muons lose on average about 3 GeV when traversing the LAr calorimeter and forward iron end-cap, and an additional 1.5 GeV in the traversal of the toroid magnet, muon detection in the FMD is restricted to muons of energies above about 5 GeV. For muons of energies around this threshold, the momentum resolution of the FMD is limited by multiple Coulomb scattering in the iron. For muons of 5 GeV momentum, the expected resolution of the forward muon detector is

$$\frac{\sigma_p}{p}\,(p = 5\ GeV) = 24\%.$$

For muons of higher energies, the effect of multiple scattering diminishes, such that

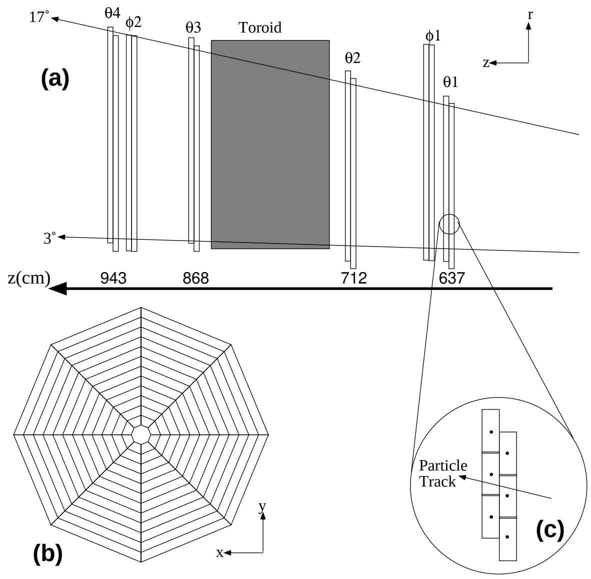


Figure 1.13: *The Forward Muon Detector. (a) Longitudinal section, showing the locations of the two ϕ- and four θ-double-layers. (b) Cross-section, showing the orientation of anode wires. (c) A magnified view of the double-layer structure that resolves the left-right ambiguity in the track reconstruction.*

the momentum resolution deteriorates only slowly to

$$\frac{\sigma_p}{p}\,(p = 200\ GeV) = 36\%$$

for very energetic muons.

1.2.4 The Luminosity System

The luminosity

$$\mathcal{L} = \frac{\dot{N}_{events}}{\sigma}$$

produced by HERA is measured by the rate of *Bethe-Heitler* [40] bremsstrahlung processes [13]

$$ep \rightarrow ep\gamma.$$

This process provides a theoretically well known cross-section (calculated with an accuracy of about 0.5%) that is large enough ($\sigma_{BH} \sim 10\ mb$) to reduce the statistical uncertainty of the luminosity measurement to a negligible level.

The bremsstrahlung photons and scattered electrons may be detected in coincidence in two *small angle* calorimeters, installed next to the beam-pipe in upstream direction at $z_{ET_{33}} = -33.4\ m$ and $z_{PD} = -102.9\ m$. An overview of the luminosity system is shown in figure 1.14. Electrons scattered by small angles are deflected out of the electron beam-pipe by a bending magnet that is installed at $z_{M_{ep}} = -23.8\ m$ to separate the electron and proton beams. Scattered electrons of energies in the

[13] See [41] for details of the luminosity measurement.

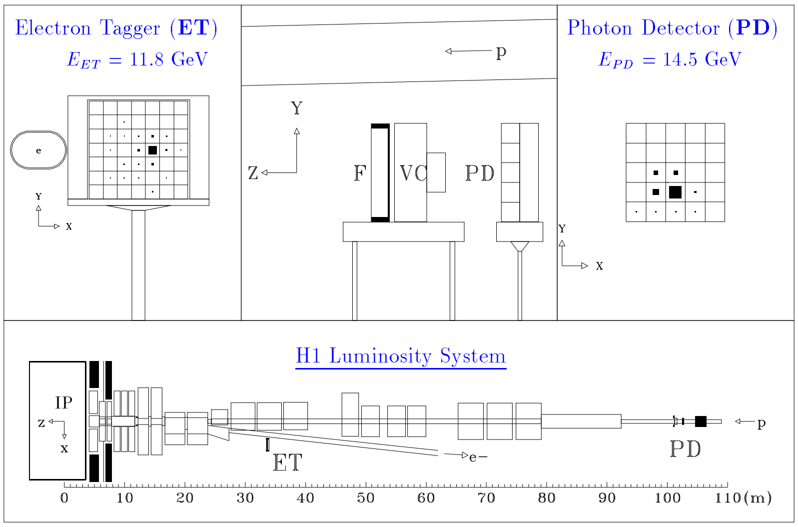


Figure 1.14: *The Luminosity System.*

interval $E'_e/E_e \epsilon [0.2, 0.8]$ pass through an exit window at $z_{W_e} = -27.3\ m$ and are detected in the *electron tagger* (ET_{33}) installed at $z_{ET_{33}} = -33.4\ m$. Bremsstrahlung photons emitted close to the electron direction leave the proton beam-pipe through an exit window at $z_{W_\gamma} = -92.3\ m$, where the proton beam-pipe bends upwards, and hit the *photon detector* (PD) installed at $z_{PD} = -102.9\ m$. The photon detector is protected from synchrotron radiation by a lead filter that absorbs the high flux of synchrotron radiation photons of energies $E_\gamma < 100\ keV$ [42]. Bremsstrahlung photons that induce an electromagnetic shower in the lead may be identified by a water-Čerenkov detector [43] installed behind the filter, that is sensitive to electrons/positrons. From the proton side, the photon detector is shielded by an iron wall of 2 m thickness.

Both the electron tagger and the photon detector are constructed from radiation-hard thalium chloride/thalium bromide (TlCl/TlBr) crystals. The crystals have a length of approximately 22 X_0 and are arranged in an array of 7×7 (5×5) units in the electron tagger (photon detector). The energy of incident electrons (photons) is reconstructed from the Čerenkov light emitted by electrons/positrons in electromagnetic showers induced in the crystals by the incident particles with a resolution of [16]

$$\frac{\sigma_E}{E} = \frac{0.10}{\sqrt{E}}.$$

The Čerenkov light is read-out from each crystal by separate photomultipliers, which provides a spatial resolution of $0.3 - 1.2\ mm$ for the reconstruction of the impact position of incident electrons (photons), as illustrated in figure 1.14.

In addition to the luminosity measurement, the electron tagger is used for the detection of scattered electrons in the analysis of photo-production processes, for which it provides an acceptance of

$$y \approx 1 - \frac{E'_e}{E_e} \epsilon [0.2, 0.8].$$

In this application, the water-Čerenkov detector installed in front of the photon detector is used as a *veto counter* (VC) to suppress the bremsstrahlung processes.

1.2.5 Trigger and Data Acquisition

At design HERA luminosity, about 1000 *ep* interactions take place in the H1 detector each second [44]. On the one hand, an event rate of 1 kHz is *small* in comparison to the HERA bunch-crossing frequency of 10.4 MHz, meaning that in only about one out of every thousand collisions an *ep* interaction happens. On the other hand, an event rate of 1 kHz means 3.6 million events each hour, a large number in comparison to the $\mathcal{O}\,(300\,\text{million})$ events recorded by H1 in seven years of running and analysed in this thesis. This rough estimate demonstrates clearly the need for selection logic that decides for which events the detector is to be read-out and the data recorded. A more detailed consideration concerning the bandwidth available for data-taking reveals that the maximum data-taking rate is limited to $\mathcal{O}\,(10\ Hz)$ for stable operation of the H1 detector [16]. The task of the trigger system is to steer the data-taking, in other words, to find the ten most interesting events each second and to initiate their read-out. As all events that are not triggered are lost for data analysis, the trigger system constitutes an extremely important component of the H1 experiment.

The H1 trigger system is composed of four levels, that are designed to successively reduce the rate from the $f_{BC} = 10.4\ MHz$ bunch-crossing frequency at its input to the $\mathcal{O}\,(10\ Hz)$ rate at its output. The layered design allows for trigger decisions of increasing complexity and precision in successive trigger levels. The H1 trigger system is illustrated in figure 1.15. In the figure, the individual trigger

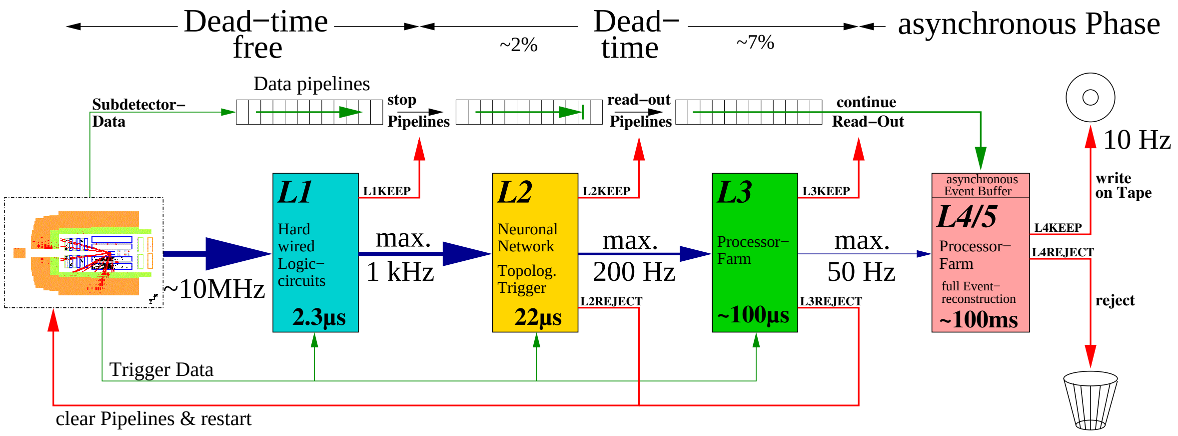


Figure 1.15: *Illustration of the four-level Trigger System of the H1 experiment. The status of the pipelines buffering the event information is shown as is the dead-time generated by the different trigger levels during nominal operation of the H1 experiment (taken from [45]).*

levels are symbolized by separate boxes. The time intervals quoted in the bottom of the boxes indicate the period scheduled for processing of the trigger signals in the corresponding trigger level. The frequencies labeling the arrows connecting the boxes represent the target output rates of the individual levels. As each trigger level can only process one (and the same) event at a time, the output rate of successive levels is inversely proportional to the increase of processing time needed for more time-consuming computations. Only limited information is available to the trigger system for the decision on whether to *keep* or *reject* an event, due to time constraints. As the complete read-out of the H1 detector by the data-acquisition system takes $1-2\ ms$, most subdetectors provide a separate data-stream optimized for trigger purposes. The information in this data-stream is restricted in quantity and to fast signals [14]. The timing of the signals provided to the trigger system is synchronized by the *central trigger logic* (CTL) which also coordinates the internal communication between the different levels of the system. The trigger decisions of the individual levels are then based on the information of the trigger data-stream in combination with the output of the preceeding trigger levels.

The first trigger level selects candidate events for interesting *ep* interactions with a trigger rate of $1\ kHz$. The processing time available to reach a trigger decision on $L1$ is 24 bunch-crossings (BC), during which time the information of all subdetectors is buffered in circular pipelines. To meet these time constraints, the first trigger level is constructed of fast hardwired logic circuits. The $L1$ trigger decision is based on 256 *trigger elements* (L1TE), that are sent from various subdetectors to the trigger system and describe the particle activity in the respective subdetectors. Since the generation of trigger elements needs a variable period of time in different subdetectors (depending on their response time, cable delays, and the time needed for processing), the trigger elements are generally not transmitted coincidently and have to be synchronized by the central trigger logic. The 256 trigger elements are logically combined to 128 *raw subtriggers* (L1RST), each of which may be gated individually with *global options* flags that describe the reconstructed vertex position and the time at which the interaction occured [15]. Most of the 128 available subtriggers are defined as *physics triggers*, which are configured to trigger efficiently the varying signatures of different *ep* processes. The physics triggers are complemented by a minority of *monitor triggers*, that are used to check the performance of individual subdetectors and to determine the efficiency of the physics triggers. As not all fulfilled physics or monitor trigger conditions are intended to initiate event read-out, each L1RST may be individually *prescaled*. A prescale factor of n means that the respective raw subtrigger is ignored for the trigger decision $n-1$ out of n times it is set, and in only 1 out of n times the corresponding *actual subtrigger* (L1AST) is set. The $L1$ trigger decision is determined by the actual subtriggers. If at least one actual subtriggers is set, an $L1keep$ signal is issued by the central trigger logic,

$$L1keep = \sum_i L1_i^{ast}.$$

The $L1keep$ signal stops the filling of the circular pipelines buffering the event information, in order to avoid them from being overwritten. As long as the pipelines

[14] To meet the timing requirements of the trigger system, subdetectors often have to provide signals faster then would be required for optimal precision, thus compromising precision for speed.

[15] The reconstructed vertex position z_{vertex} and time of interaction t_0 are of special importance at HERA, since the *ep* interaction rate of $f_{ep} \sim 1\ kHz$ is exceeded by a factor $\mathcal{O}(100)$ higher rate [46] of beam-gas, beam-wall and cosmic muon events, that can be very effectively rejected by a combination of z_{vertex} and t_0 conditions (see section 6.3).

remain stopped, the H1 detector is insensitive to further *ep* interactions. Accordingly, the *L1keep* signal starts the *dead-time* of the detector. Otherwise, in case none of the actual subtriggers is set, the event information in the pipelines is simply overwritten by the data of succeeding bunch-crossings and no dead-time is generated.

The *L1keep* decisions are validated by the second trigger level. The second trigger level is constructed of two different processor systems that confirm or falsify the decisions of the first level within $20\mu s$. The decision of both systems is in principle based on the same information as is available to the first trigger level, but is now sent by the subdetectors to the trigger system in higher granularity. For the required rate reduction, the second trigger level better exploits the correlations between the various subdetectors. The first component, the *L2 Neural Network* (L2NN) trigger, is realized in a neural network architecture. It is composed of 10 separate neural networks which are individually trained to identify a specific *ep* interaction process. Each of the neural networks may have up to 64 input neurons, 64 hidden neurons and 1 output neuron. The input to the networks are global event quantities extracted by preprocessing units in the L2NN system from the information provided by the subdetectors. Their output indicates the likelihood that the event under consideration is one of the ten processes the network is trained to identify. (for an explanation of neural networks see section 3.2.3). The second component, the *L2 Topological Trigger* (L2TT), searches for spatial correlations between the signals of different subdetectors. The correlations are searched for as coincidences in a two-dimensional histogram, in which all tracks in the drift chambers, energy deposits in the calorimeters and tracks in the muon system are represented by their θ and ϕ coordinates. From coincidences in the histogram, a rough particle identification may be derived; for instance, electrons may be identified by energy deposits localized in the electromagnetic section of the LAr that match a drift chamber track in the θ-ϕ plane [47]. The processing unit of the L2TT system is capable of computing 16 different projections of the two-dimensional histogram, each of which may define a specific event topology. The output of the 10 neural networks of L2NN and the 16 projections of L2TT provide 26 trigger elements (L2TE). In principle, these trigger elements may be logically combined into up to 48 subtriggers (L2ST); so far, however, this possibility has not been used, and the $L2$ trigger elements are in a one-to-one correspondence with the $L2$ subtriggers. The $L2$ subtriggers are used to validate the actual subtriggers of the first trigger level. The set of $L2$ subtriggers that validates a single actual $L1$ subtrigger is defined by the elements of a *validation matrix* M. The $L2$ trigger decision is determined by the actual $L1$ subtriggers that pass the validation by $L2$ subtriggers. Formally, the trigger decision of the second trigger level is evaluated as the result of the matrix multiplication

$$L2keep = \sum_{i,j} L2_i^{st} \cdot M_{ij} \cdot L1_j^{ast}.$$

An *L2keep* signal is issued by the central trigger logic if the result of the matrix multiplication is non-zero, which means that at least one actual $L1$ subtrigger passes the $L2$ validation. In order to allow for actual $L1$ subtriggers that should pass the second trigger level without $L2$ validation, the vector of $L2$ subtriggers is extended by an additional component, the *trivial $L2$ subtrigger*, that is always set [47]. The *L2keep* signal initiates the complete read-out of the H1 detector, in which the information from all subdetectors is collected by the *central event builder* (CEB) of the data-acquisition system for processing by the fourth trigger level. This process takes between $1-2\ ms$, during which the pipelines remain stopped and dead-time accumulates. In case the event is rejected by the second trigger level, the pipelines are restarted, the event is lost, and the dead-time ends.

The third trigger level may interrupt the time-consuming detector read-out and prematurely end the dead-time. Its objective is to reduce the trigger rate by about a factor four by rejecting events within a period of $\mathcal{O}\,(100\ \mu s)$ [48, 49]. In the HERA I running period the third trigger level was not used, however. Instead, the second trigger level was configured to meet the maximal $L4/5$ input rate and an automatic *L3keep* signal was generated on the third trigger level. In order to cope with the higher trigger rates in the HERA II running period, the third trigger level has been installed recently, as part of the *fast track trigger* (FTT) project [50]. It is implemented by software algorithms running on a multi-processor system [45].

The fourth trigger level performs a complete reconstruction and classification of the event. It is an asynchronous trigger level, implemented by software algorithms running on a batch system composed of off-the-shelf Personal Computers [51]. The processing of the event information on the batch system starts once the central event builder of the data-acquisition system has finished collecting the raw event data provided by the subdetectors. At this moment, the pipelines are also started again, terminating the dead-time of the detector. The raw event data is processed by the H1 reconstruction software *H1REC*, reconstructing clusters in the calorimeters and tracks in the drift chambers and muon detectors. The primary event vertex and the time of interaction are subsequently reconstructed and remaining beam-gas, beam-wall and cosmic muon events are rejected. The accepted events are classified into *event classes* such as high Q^2, high p_T, and rare exclusive final states and *event signatures* that flag, for instance, events with muons or electrons [52].

Before the event information is finally written onto tape, the events may be filtered to reduce the rate of frequent low Q^2, low p_T processes (so-called "soft physics") [16]. The events passing the filtering are written to two different data-streams, the *Production Output Tapes* (POTs) and the *Data Summary Tapes* (DSTs). On the Production Output Tapes, the complete event information (about 100 *kb* of data per event), composed of raw and reconstructed quantities, is stored. A subset of mainly reconstructed quantities (about 10 *kb* of data per event) that is sufficient for most data analyses, is duplicated on the Data Summary Tapes. In contradiction to their names, the DST files are stored on hard disks, providing a significantly higher access speed in comparison to the POT files.

1.2.6 Data Processing and Analysis

The data recorded "online" by the data acquisition system discussed in the previous subsection needs further processing to yield results such as those presented in this thesis. First, more extensive reconstruction of the information stored on the Data Summary Tapes is necessary to get access to relevant physical information such as

[16] The idea behind this technique, termed *downscaling*, is to replace a number of events by a single representative that is weighted accordingly; the resulting loss in statistics constitutes no restriction for data analysis if applied to frequent processes only.

particle identification and the kinematical variables that describe the event. After the full reconstruction, additional processing is needed to select, extract and inspect the events of interest to a particular physics analysis.

The final event reconstruction is done "offline" (asynchronous to the data-taking), on a batch system composed of off-the-shelf Personal Computers. The same batch system may also be used for individual physics analysis, but, due to the smaller amount of computing power normally needed, the final analysis programs are most often run interactively on desktop computers.

For the final event reconstruction and physics analysis, custom made software packages have been written by members of the H1 collaboration. Traditionally, the H1 software environment has been based on the *Fortran* programming language. In the Fortran based software environment, the event reconstruction is typically accomplished by a combination of functions in the common *H1PHAN* [53] library and analysis specific code, shared by a few people in a working group. The reconstructed information is written in an "N-tuple" (a set of N numbers describing each event) in *HBook* format [54] and analysed with programs written in the macro language of *PAW* [55]. Over the years, the analysis specific code used by different working groups diverged substantially, as a result of private extensions and improvements. This development made it difficult to compare and cross-check analyses done by different groups. Another disadvantage of this development is that expert knowledge is not made available to the whole of the H1 collaboration. Consequently, a new analysis environment for H1 [56] was developed, that avoids the drawbacks of the old design. The new environment is realized as an object-oriented design and implemented in the *C++* programming language. It is based on the *RooT* framework [57], which provides fundamental functionality for efficient event storage and access, graphics display facilities and basic tools for physics analysis. The new analysis environment, developed in the *H1OO* project [58], aims to provide a new unique framework for all physics analyses at H1. In the H1OO project, expert knowledge is made available in a public code reference managed by *CVS* [59], and a set of standardized algorithms are provided for event reconstruction and physics analysis.

In the object-oriented environment, the final event reconstruction is accomplished by the so-called *filling code*. The function of the filling code is to compute the information that is needed for most physics analyses and store it in an object-oriented data structure, as *objects* instantiated from *classes* [17]. The execution of the filling code proceeds on three levels. On the first level, the information written to Data Summary Tapes by the last trigger level $L4/5$ is read in and converted to object-oriented data structures. Important data structures filled on this level are those of the tracks in the drift chambers and muon detectors and those of clusters in the calorimeters. On the second level, algorithms developed for the identification and reconstruction of specific types of particles (the so-called particle-finders, see chapter 4) process the track and cluster information and fill data structures corresponding to the identified particles. By default, electrons/photons, muons and charged and neutral hadrons are searched for as *elementary* particles. The elementary particles

[17] A *class* combines data and algorithms operating on the data in a self-contained entity; for instance, a cluster class might contain uncalibrated energy values as *data members* together with *member functions* for calibration. In object-oriented designs, classes constitute the standard data structure.

are then subject to algorithms dedicated to the identification of *composed* particles such as K^0-, J/Ψ- and D^*-mesons, but also jets. On the third level, global event quantities (*cf.* section 5) are computed from the reconstructed information of the first two levels. In contrast to the data structures filled on the other two levels, the event quantities reconstructed on the third level are described by simple integer or real-valued numbers instead of objects instantiated from classes.

Corresponding to the three level structure of the filling code, the reconstructed information is stored in three *data-layers* [18]. The fundamental layer is the *"Object Data Store"* (ODS), which contains the same informations as the DST it was filled with, yet in the form of objects. The identified particles are stored in the *"micro Object Data Store"* (μODS) layer. In the *"H1 Analysis Tag"* (HAT) layer, global event quantities are stored as simple numbers. This has the technical advantage that libraries with class definitions are not necessary to read the information stored on this layer, simplifying the fast pre-selection of events. Once filled, the three data layers HAT, μODS and ODS are written as RooT *"Trees"* to separate files [19]. The filling code is executed in regular intervals (currently about every six months), to produce an updated set of HAT, μODS and ODS files for all H1 data with the most recent H1OO release. By preselecting events on the basis of information stored on the HAT, it is possible to extract a complete subset of all H1 data within a few days.

For the final physics analysis, a set of tool libraries is provided by the H1OO project, in order to simplify the development of analysis programs. Among other tools, the libraries include a complete framework providing solutions for most common analysis tasks, such as file-handling, event selection, histogramming/binning and the measurement of cross-sections [60].

The analysis presented in this thesis was done using the release series 2.5.x of the object-oriented analysis environment of H1.

1.2.7 Detector Simulation

In high energy physics experiments, the actual physics processes of interest can in general not be observed directly, but have to be reconstructed from quantities that can be directly measured, such as hits in the tracking detectors or energy deposits in the calorimeters. For the experimental data to be compared with theoretical predictions, either the underlying physics processes have to be unambiguously inferred from the measured hits and energy deposits in the detector or the hits and energy deposits have to be *simulated* for the processes predicted by theory.

As firstly, the processes particle physicists are interested in are always probabilistic in nature and secondly, theoretical predictions are often linked via empirical models (*cf.* the description of hadronisation processes in section 2.4) to the measured hits and energy deposits in the detector, the underlying physics processes can in general not be unambiguously inferred from the experimental data. In order to compare the experimental data with theoretical predictions, it is therefore necessary

[18] Additional information relevant for individual working groups may be stored in optional data-layers, supplementing the three official ones.

[19] Since H1OO release series 2.3.x large ODS trees do not have to be written to file anymore, as they can be dynamically produced in memory from the corresponding DST when needed; information in optional data-layers is written to additional files as so-called "user-Trees".

to simulate the *detector response* to the particles produced in *ep* interactions.

The probabilistic nature of particle physics experiments is accounted for by allowing for a certain “randomness” in the simulation of the interaction processes, using random numbers following specific probability distributions. Such probabilistic techniques are known as *Monte Carlo* methods [20].

The simulation of the experimental signature of a single *ep* interaction (an “event”) may be separated into two stages. In the first stage, *event generators* simulate a random *primary ep* interaction, according to the differential cross-section predicted by some theory (see chapter 2). As a result of this *generation*, a list of final state particle four-vectors is created. In the second stage, the interactions of the generated particles with the detector material and the resulting drift chamber and calorimeter signals are simulated. The generated particles are propagated through the detector in a step-by-step procedure, in each step simulating random interactions with the detector material. Simultaneously, decays of unstable particles are simulated randomly. The distance traversed before the decay and the type of the produced *secondary* particles are also determined randomly, according to probabilities corresponding to the lifetimes and branching fractions of the decayed particles. For the simulation of the interactions of propagated particles with the detector material, an accurate description of the detector geometry and its material composition is necessary [21]. The simulation of the H1 detector is implemented in the *H1SIM* software package [62], based on the *GEANT3* framework [63], that has been developed for detailed simulations of the interactions of high energetic particles with matter. In the interactions of the generated particles with the detector material further secondary particles are produced that are also propagated through the detector until they have lost all but a negligible fraction of their energy. At the end of the simulation stage, the detector response to all primary and secondary particles is simulated, creating the hits in the tracking detectors as well as the energy deposits in the calorimeters.

The simulated Monte Carlo events are then *reconstructed* in the same way as real data by the algorithms implemented in the H1REC software package and the H1OO filling code, thus ensuring a consistent treatment of simulation and reality. The event can now be compared to real experimental data.

[20] The name “Monte Carlo” was introduced by the Manhattan project at Los Alamos in the 1940’s and was suggested in reference to the gambling casinos in Monte Carlo, Monaco [61].

[21] In practice, approximations may be necessary in the description of the detector; in this case, the errors introduced by these approximations into the simulation have to be small in comparison to the experimental uncertainties, for the simulation to describe the data.

Chapter 2

Theoretical Overview

Following the description of the experiment in the last chapter, this chapter provides the theoretical background for the analysis presented in this thesis. In the first section, the variables defining the kinematics of electron-proton collisions at HERA are introduced. The kinematic variables are used for the description of the ep interactions in the second and third sections. In the second section, ep interactions are described in the limit of high magnitudes of the four-momentum transferred between the interacting electrons and protons. These *Deep-Inelastic Scattering* (DIS) processes are used to study the proton structure at HERA. The results are explained in the framework of the *Quark-Parton Model* (QPM) and the theory of *Quantum-Chromodynamics (QCD)*. In the limit of low magnitudes of the transferred four-momentum, ep interactions may be factorised into the emission of a quasi-real photon by the beam electron and a subsequent photon-proton interaction. These *Photoproduction* (γp) processes are detailed in the third section; in photoproduction processes, the electron beam may essentially be thought of as being a photon source and HERA as being a high energy photon-proton collider. In the fourth section, the transition between the quarks and gluons produced in DIS or γp processes and the elementary particles observed in nature is detailed: one of the predictions of the theory of QCD is that quarks and gluons - the constituents of the proton - cannot be observed as free particles, but only in bound states. The formation of these bound states - the *baryons* and *mesons* observed in nature - is described by *Parton Showering* and *Hadronisation* processes. Specific background processes contributing to samples of events with isolated leptons and large missing transverse momentum are discussed in the fifth section. The signal processes producing events with isolated leptons and large missing transverse momentum are detailed in the sixth and seventh sections. In the sixth section, the dominant signal process within the Standard Model, the production of real W bosons (with subsequent leptonic decay), is introduced. Possible additional signal processes within theories beyond the Standard Model are described in the seventh section, concluding the theory chapter.

2.1 Kinematics of ep-Collisions

The electron-proton interactions studied at HERA are conventionally classified into *Neutral Current (NC)* and *Charged Current (CC)* processes, depending on the charge of the exchanged gauge boson. As illustrated in figure 2.1, the term neutral current

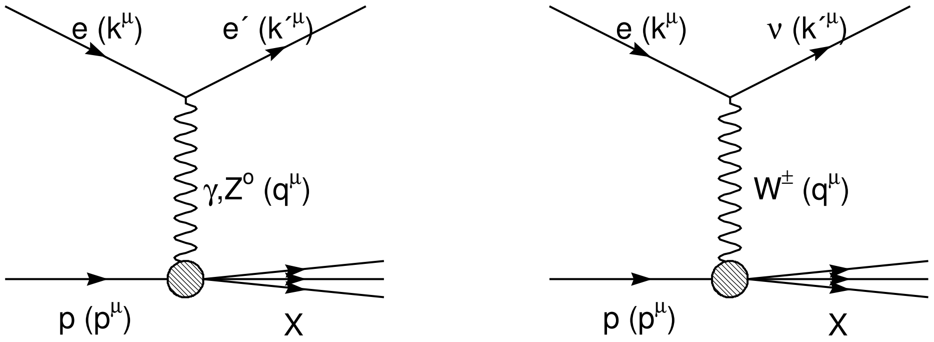


Figure 2.1: *ep scattering via virtual boson exchange in NC (left) and CC (right) interactions.*

comprises all processes $ep \to eX$ [1] in which the beam electron is scattered without changing particle type, in contrast to charged current processes $ep \to \nu_e X$ in which the beam electron changes into an electron-neutrino. As can be seen in the figure, neutral current processes are mediated by the exchange of either a photon (γ) or a Z^0 boson, whereas in charged current processes a $W^\pm$ boson is exchanged. When describing neutral and charged current processes, occasionally the adjective "inclusive" is added to emphasize that actually an entire class of processes is implied.

Following the notation used in the figure, the incident electron (e) and proton (p) have four-momenta $k^\mu = \left(E_e, \vec{k}\right)$ and $p^\mu = (E_p, \vec{p})$, respectively. The squared four-momentum transfered in the scattering process is

$$q^2 \equiv \left(k^\mu - k'^\mu\right)^2 ,$$

where k'^μ denotes the four-momentum of the scattered lepton. Since

$$q^2 = k^2 + k'^2 - 2g_{\mu\nu}k^\mu k^\nu = 2m_e^2 - 2E_eE_e'\left(1 - \cos\theta_e\right) \approx -2E_eE_e'\left(1 - \cos\theta_e\right)$$

is a negative quantity, usually the positive quantity

$$\begin{aligned} Q^2 &\equiv -q^2 \\ &\approx 2E_eE_e'\left(1 - \cos\theta_e\right) \end{aligned} \tag{2.1}$$

is defined for convenience. The variable Q^2 is a measure of the *virtuality* of the exchanged boson; for $Q^2 \ll 1\ GeV^2$ photons are considered to be quasi-real or *on mass-shell*.

In addition to the four-momentum transfer Q^2, several other Lorentz-invariant scalars are conventionally defined to describe the kinematics of the interaction, namely the *centre-of-mass energy* of the electron-proton system

$$s \equiv \left(k^\mu + p^\mu\right)^2 ,$$

[1] On the hadronic side, the label "X" includes all possible kinds of final states; the scattering process is said to be *elastic* if the proton remains intact and *inelastic* if the proton *dissociates*, typically into a single jet (formed by the *struck quark*) in the detector and a *proton-remnant* (formed by the *spectator quarks*) going down the beam-pipe.

and the two dimensionless variables

$$x \equiv \frac{Q^2}{2p \cdot q} \tag{2.2}$$

and

$$y \equiv \frac{p \cdot q}{k \cdot p}. \tag{2.3}$$

In the so-called *infinite momentum frame*, a coordinate system in which the longitudinal momentum of the proton is very large, i.e. $p^\mu \approx (E_p, 0, 0, E_p)$ and $E_p \gg m_p$ [2], the variable x describes the fraction of the proton's momentum carried by the struck quark [3]. In this frame the proton may be thought of as a system of point-like *partons*, traveling colinearly within the proton. This idea is the basis of the *Quark-Parton Model* (discussed in more detail in the next section). As a result of the Quark-Parton Model, Bjorken predicted in 1969 that for sufficiently high Q^2 (which was found to be surprisingly low, of the order of a few GeV^2 only), the inelastic scattering cross-section depends only on the *scaling variable* x [64]. Because of the great success of Bjorken's theory, the variable x is also called *Bjorken* x.

The variable y is called the *inelasticity* and may be understood, in the rest frame of the proton, as the fraction of the incident electron's momentum carried by the exchanged boson [4].

By definition, the above quantities are not independent, but are related by [5]

$$Q^2 \approx xys.$$

At HERA, the centre-of-mass energy $\sqrt{s}$ is fixed by the energies of the electron and proton beams and the scattering process can be completely described by any two of the variables x, y and Q^2 [6].

2.2 Deep-Inelastic Scattering

A large fraction of the knowledge that exists today about the structure of matter is due to scattering experiments. In these experiments, the structure of an extended object is inferred from its interactions with a point-like probe. The probably most

[2] At HERA, the laboratory system represents an infinite momentum frame to a reasonable approximation.

[3] If the momentum fraction of the struck quark is denoted as ϵ, its invariant mass after the scattering is $(\epsilon p + q)^2 = \epsilon^2 p^2 + q^2 + 2\epsilon p \cdot q = m_q^2$, from which it follows that $2\epsilon p \cdot q = Q^2$ or $\epsilon = \frac{Q^2}{2p \cdot q} = x$, if the mass terms $m_q^2 \approx 0$ and $\epsilon^2 p^2 = \epsilon^2 m_p^2 \approx 0$ are neglected.

[4] In the proton rest frame, $p = (m_p, 0, 0, 0)$ and therefore $p \cdot q = p \cdot (k - k') = m_p (E_e - E_e')$ and $p \cdot k = m_p E_e$, such that $y = \frac{E_e - E_e'}{E_e}$ is the fraction of the incident electron's momentum carried by the exchanged boson.

[5] If particles masses are neglected, then $\frac{Q^2}{xy} = \frac{Q^2}{\frac{Q^2}{2p \cdot q} \frac{p \cdot q}{k \cdot p}} = 2k \cdot p = (k+p)^2 \equiv s$.

[6] The scattering process is completely described by two variables if it is symmetric in the azimuthal direction, i.e. for unpolarized beams (as in the HERA I running period). In case of transversely polarized electrons beams (as in the HERA II running period), additionally the azimuthal angle of the scattered electron is necessary for a complete description of the scattering process.

prominent scattering experiment is the classical Rutherford experiment. In this pioneering experiment, performed by Rutherford, Geiger and Marsden [65], α particles were collided with a thin gold foil. From the deviations in the observed angular distribution of the scattered α particles with respect to the theoretical predictions for point-like particles, Rutherford inferred that in the centre of atoms there must exist a hard core, approximately $10^{-15}m$ in size, that carries about 99.9% of the atom's mass [1]. Nowadays we know that the hard core discovered by Rutherford is to be identified with the atomic nucleus, which is itself not point-like, but is composed of protons and neutrons. In scattering experiments at HERA, the substructure of the proton is probed in collisions with electrons down to length scales of the order of $10^{-18}m$ and with higher precision than any other experiment before.

The cross-section for *elastic* electron-proton scattering processes $ep \rightarrow ep$ is observed to be rapidly falling with increasing magnitude of the four-momentum transfer Q^2. Simultaneously, the cross-section for *inelastic* processes $ep \rightarrow eX$, in which the initial proton is transformed to a system X of hadrons in the final state, become more dominant. The observation of inelastic scattering processes in ep collisions is explained by the existence of substructure in the proton [2]. The proton substructure is described by two *structure functions* F_1 and F_2, that describe the distribution of electric charge and magnetic moment in the proton.

A recent measurement of the structure function F_2 of the proton by the H1 collaboration is displayed in figure 2.2. In the figure, F_2 is shown as a function of Q^2 for different values of x. The proton structure function can be seen to be independent of the four-momentum transfer Q^2 with which the proton substructure is probed over a wide range of the kinematic plane (shown in figure 1.1). This *scale invariance* of the structure function is described within the *Quark-Parton Model*. At high Q^2 and either very high or very low x, some deviations from the perfect scaling behavior are seen in the proton structure function, however: F_2 is seen to rise with Q^2 at low x and, to a lesser degree, to fall with Q^2 at high x. These deviations are termed *scaling violations* and described by *Quantum-Chromodynamics*, the gauge theory of the strong interaction within the Standard Model.

---

[1] The Rutherford experiment thus demonstrated that an atom is not really an "atom", which means "indivisible" in Greek; until the early twentieth century scientists believed the atoms listed in the periodic table to be the most elementary constituents of matter.

[2] The length scale characterising the substructure of the proton may be estimated by the magnitude of the four-momentum transfer Q^2 at which the cross-section for inelastic scattering processes becomes comparable to the elastic cross-section: in the scattering process, the proton is probed with a resolution of the order of the wavelength $\lambda_{gamma} = \frac{hc}{E_\gamma}$ of the photon emitted by the electron. The energy of the photon is determined by the kinematics of the scattering process. In the rest frame of the proton, in which $p \equiv (m_p, 0, 0, 0)$, $x \equiv \frac{Q^2}{2p \cdot q} = \frac{Q^2}{2m_p(E_e - E_e')} = \frac{Q^2}{2m_p E_\gamma}$, yielding $\lambda_{gamma} = 2x \cdot m_p \cdot \frac{hc}{Q^2}$. Experimentally, the transition between elastic and inelastic processes is observed around $Q^2 \sim 1\ GeV^2$ [66], which yields an estimate for the characteristic length scale of $\mathcal{O}(1\ fm)$ - approximately the size of the proton.

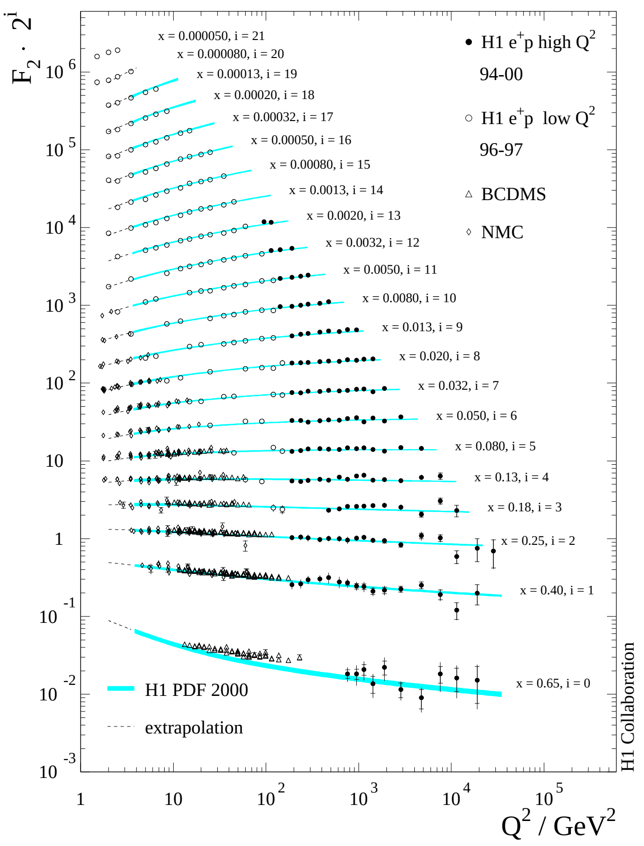


Figure 2.2: *The proton structure function* $F_2\left(x, Q^2\right)$ *as a function of* Q^2 *for a range of different* x *values. The presented H1 data (points) is compared to the H1 fit of parton density functions [67] (solid line). At low* Q^2 *the H1 data is complemented by data from the* BCDMS *[68] and* NMC *[69, 70] fixed target experiments.*

2.2.1 Scaling and the Quark-Parton Model

The scale invariance of the structure functions is explained by the presence of non-interacting point-like *partons* [3] within the proton. In the *Quark-Parton Model*, *inelastic* proton scattering is interpreted as *elastic* scattering off one of the partons that is *struck* by the photon emitted by the scattered electron. As the point-like nature of the partons does not introduce any length scale into the description of the proton structure, the observed scale invariance of the structure functions is naturally explained within this model.

An important prediction of the Quark-Parton-Model is that the two proton structure functions F_1 and F_2 are not independent. For spin 1/2 partons, the structure functions F_1 and F_2 are related to the distribution functions $f_i\left(x\right)$ of the partons

[3] The partons were later identified with quarks, the name "parton" was chosen at a time, when quarks and gluons were accepted merely as mathematical concepts, but not believed to constitute "real" particles.

within the proton by

$$
\begin{aligned}
F_1(x) &= \frac{1}{2}\sum_i Q_j^2 \cdot f_i(x) \text{ and} \\
F_2(x) &= x \cdot \sum_i Q_i^2 \cdot f_i(x),
\end{aligned}
$$

where $f_i(x)\,dx$ denotes the probability to find a parton of charge Q_i carrying a proportion between x and $x + dx$ of the proton's momentum. In particular, the two structure functions F_1 and F_2 are predicted to be not independent, but to be related by

$$F_2 = 2x \cdot F_1. \tag{2.4}$$

The relation between F_1 and F_2 in form of equation 2.4 has been unambiguously confirmed experimentally for Q^2 between $\mathcal{O}(1GeV^2)$ and $\mathcal{O}(10^4 GeV^2)$, the largest four-momentum transfers accessible at HERA.

2.2.2 Scaling Violations and Quantum-Chromodynamics

Although the Quark-Parton Model describes the observed scale invariance of the proton structure functions over a wide range of the kinematic plane, it fails to explain the deviations from perfect scaling behavior seen at high Q^2 and either very high or very low x. The scaling violations are successfully described by a phenomenologically richer theory, namely the theory of *Quantum-Chromodynamics (QCD)*, of which the Quark-Parton Model is nowadays known to be merely an approximation.

In contrast to the Quark-Parton Model, Quantum-Chromodynamics is a *dynamic* theory that includes interactions between the constituents of the proton. In QCD, the partons are identified as *quarks*, carrying an additional quantum number, the *colour charge*, and interacting by the exchange of *gluons*, which themselves carry a combination of colour and anti-colour. The gluons represent the *gauge bosons* of QCD and bind the quarks inside the proton. As a consequence of being (colour) charged, the gluons may couple to other gluons, in contrast to the photon, the (electrically) neutral gauge boson of QED [4].

The interaction of gluons with other gluons results in a specific dependence of the strong coupling constant α_s on the resolution parameter Q^2, which is in the leading-logarithm approximation described by [23]

$$\alpha_s = \frac{12\pi}{(33 - 2N_f) \cdot \ln\left(Q^2/\Lambda_{QCD}^2\right)}, \tag{2.5}$$

where $N_f = 6$ is the number of quark *flavours* and Λ_{QCD} is a scale parameter, whose value has been determined experimentally as $\Lambda_{QCD} \approx 200MeV$ [71]. In contrast to the electromagnetic coupling strength α that *increases* at short distances, the coupling constant α_s of the strong interaction *decreases* at short distances. The decrease of the coupling constant allows for the quarks to behave as nearly free particles within the proton, a property known as *asymptotic freedom*. It is in the

[4] Technically speaking, QCD is a non-Abelian ($SU(3)$) and QED an Abelian ($U(1)$) theory (the symbols in brackets denote the *gauge group* on which the theory is based).

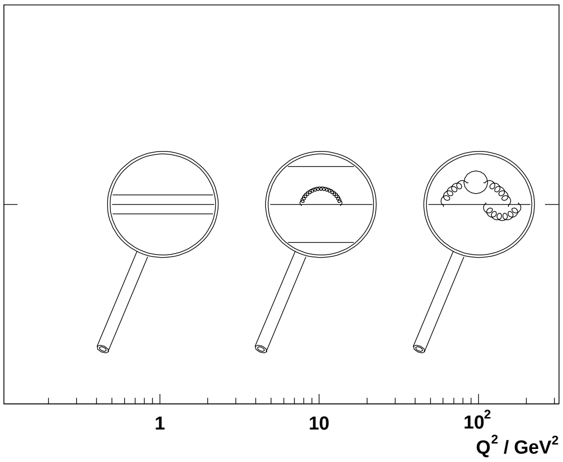


Figure 2.3: *Symbolic illustration of the proton structure observed at different Q^2.*

limit of asymptotic freedom, that the Quark-Parton Model can be derived from QCD.

The rise of the proton structure functions with increasing Q^2 at low x and the decrease at high x are a feature of the gluon interactions in Quantum-Chromodynamics: the quarks inside the proton endlessly emit and re-absorb gluons, which may then fluctuate into virtual quark anti-quark pairs. These virtual quark anti-quark pairs are termed *"sea quarks"* and distinguished from the original quark content of the proton in the static Quark-Parton Model, the *"valence quarks"*. The extent to which the virtual quarks contribute to the electron-proton scattering cross-section depends on the resolution parameter Q^2 with which the proton is probed: with increasing resolution Q^2, the photon emitted by the electron is more likely to find the proton in a state in which one of the valence quarks has radiated off one or more gluons and is surrounded by a cloud of virtual quark anti-quark pairs (see figure 2.3 for illustration). In this case, the photon may scatter off one of the sea quarks, which typically carry only a small fraction x of the proton's momentum, explaining the rise of the proton structure function at low x. If the photon however, scatters off the valence quark that has radiated off the gluons, the struck quark carries on average a smaller fraction x of the proton's momentum than it would, had it not not emitted any gluons. In other words, with increasing resolution Q^2 the proton is more likely to be "seen" by the photon in a state in which the proton's momentum is distributed over a large number of "soft" partons, such that the probability to find a large fraction of the proton's momentum concentrated in a *single* "hard" parton decreases.

2.2.3 Cross-Sections

The neutral and charged current deep-inelastic scattering cross-sections measured by H1 are shown in figure 2.4. In the graph it may be seen that the two cross-sections become comparable at large magnitudes of the four-momentum transfer Q^2. The similarity of the neutral and charged current scattering cross-sections is predicted by the unified theory of electro-weak interactions within the Standard Model: at

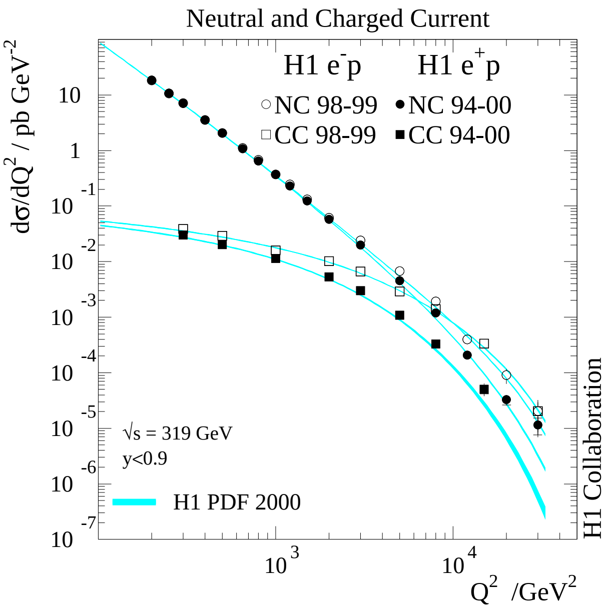


Figure 2.4: *Cross-sections for neutral (circles) and charged current (boxes) deep-inelastic scattering processes in e^+p (full symbols) and e^-p (open symbols) collisions as a function of Q^2.*

high Q^2, the mass of the exchanged gauge boson becomes less dominant in the propagator term $1/\left(Q^2 + M^2_{Z/W}\right)$ of the cross-section. Consequently, the charged current cross-section, suppressed by the high W boson mass at low and medium Q^2, becomes comparable to the neutral current cross-section at high Q^2, indicating the unification of the electromagnetic and weak forces at this scale. The difference seen in figure 2.4 between the neutral and charged current cross-sections for e^+p and e^-p collisions is a direct consequence of the quark-parton model and will be explained in more detail in the following.

Neutral Current

The full electro-weak *neutral current* cross-section is given by

$$\frac{d^2\sigma^{e^\pm p}_{NC}}{dxdQ^2} = \frac{2\pi\alpha^2}{x}\left(\frac{1}{Q^2}\right)^2 Y_+\tilde{F}_2^\pm\left(x,Q^2\right) + y^2F_L^\pm\left(x,Q^2\right) \mp Y_-x\tilde{F}_3^\pm\left(x,Q^2\right),$$

where $Y_\pm = 1 \pm (1-y)^2$ is the *helicity factor*, F_L is the *longitudinal structure function* and $\tilde{F}_2^\pm$, $\tilde{F}_3^\pm$ are generalized structure functions composed of linear combinations of five elementary structure functions that describe pure photon exchange, γZ interference and pure Z exchange [31].

For an unpolarized electron beam, the generalized structure functions may be written as

$$
\begin{aligned}
\tilde{F}_2^{\pm} &= F_2 - v_e \left(\frac{\kappa_W Q^2}{Q^2 + M_Z^2}\right) F_2^{\gamma Z} + \left(v_e^2 + a_e^2\right) \left(\frac{\kappa_W Q^2}{Q^2 + M_Z^2}\right)^2 F_2^Z \\
x\tilde{F}_3^{\pm} &= \pm a_e \left(\frac{\kappa_W Q^2}{Q^2 + M_Z^2}\right) xF_3^{\gamma Z} \mp 2a_e v_e \left(\frac{\kappa_W Q^2}{Q^2 + M_Z^2}\right)^2 xF_3^Z,
\end{aligned}
$$

where F_2 describes pure photon exchange, F_2^Z and xF_3^Z describe pure Z exchange and $F_2^{\gamma Z}$ and $xF_3^{\gamma Z}$ arise due to γZ interference. The *vector* and *axialvector* couplings of the electron (positron) are denoted by v_e and a_e and are related to the third component of *weak-isospin* by

$$
v_e = 2\sin^2\theta_W - I_3, \qquad a_e = I_3,
$$

where $I_3 = -1/2(+1/2)$ for electrons (positrons). The quantity

$$
\kappa_W = \frac{1}{4\sin^2\theta_W \cos^2\theta_W}
$$

is a function of the *Weinberg angle* θ_W that describes the *mixing* between the photon and the Z boson in the electro-weak sector of the Standard Model [5].

Within the Quark-Parton Model, the structure functions F_2, $F_2^{\gamma Z}$ and F_2^Z are expressed as the *sum* of the quark and anti-quark densities,

$$
\begin{aligned}
F_2\left(x, Q^2\right) &= x\sum_q e_q^2 \left(q\left(x, Q^2\right) + \bar{q}\left(x, Q^2\right)\right) \qquad (2.6)\\
F_2^{\gamma Z}\left(x, Q^2\right) &= x\sum_q e_q v_q \left(q\left(x, Q^2\right) + \bar{q}\left(x, Q^2\right)\right) \\
F_2^Z\left(x, Q^2\right) &= x\sum_q \left(v_q^2 + a_q^2\right) \left(q\left(x, Q^2\right) + \bar{q}\left(x, Q^2\right)\right)
\end{aligned}
$$

and the structure functions $xF_3^{\gamma Z}$ and xF_3^Z as the *difference* of the quark and anti-quark densities,

$$
\begin{aligned}
F_3^{\gamma Z}\left(x, Q^2\right) &= x\sum_q 2e_q a_q \left(q\left(x, Q^2\right) - \bar{q}\left(x, Q^2\right)\right) \\
F_3^Z\left(x, Q^2\right) &= x\sum_q 2v_q a_q \left(q\left(x, Q^2\right) - \bar{q}\left(x, Q^2\right)\right).
\end{aligned}
$$

The sums extend over all "active" quark flavours q [6]. As the neutral current cross-section falls steeply with Q^2, neutral current scattering processes occur typically at low and medium Q^2, where the contributions from Z exchange and F_L may be neglected. In this case, it may be seen from the expanded form of the proton

[5] As a consequence of the $SU(2) \otimes U(1)$ gauge group structure of the unified electro-weak interaction, a specific relation between the Weinberg angle and the masses of the W and Z bosons is predicted within the Standard Model: $M_W = \cos\theta_W \cdot M_Z$.

[6] "Active" quark flavours are the ones that are kinematically accessible. At HERA, the centre-of-mass energy is not sufficient for the production of top quark pairs, so the number of "active" quark flavours is $N_q = 5$.

structure function F_2 in equation 2.7, that neutral current scattering processes are mainly sensitive to the up and charm quark densities within the proton [7].

At HERA, deep-inelastic neutral current scattering processes are simulated by the Monte Carlo generators *Django* [72] and *Rapgap* [73] [8].

Charged Current

The *charged current* cross-section is given by

$$\frac{d^2\sigma_{CC}^{e^\pm p}}{dxdQ^2} = \frac{G_F^2 M_W^4}{2\pi x}\left(\frac{1}{Q^2+M_W^2}\right)^2 Y_+\tilde{W}_2^\pm\left(x,Q^2\right) + y^2 W_L^\pm\left(x,Q^2\right) \mp Y_- x\tilde{W}_3^\pm\left(x,Q^2\right),$$

where

$$G_F = \frac{\pi\alpha}{\sqrt{2}\sin^2\theta_W M_W^2} = \frac{g}{4\sqrt{2}M_W^2}$$

is the *Fermi constant*. The helicity factors $Y_\pm$ in the charged current cross-section are the same as in the neutral current one. The structure functions $\tilde{W}_2$, W_L and $x\tilde{W}_3$ are defined in analogy to $\tilde{F}_2$, F_L and $x\tilde{F}_3$ in the neutral current case to emphasize the structural similarity between the two cross-sections. The correspondence between the neutral and charged current cross-sections can be seen best, if both cross-sections are expressed in terms of the couplings e and g,

$$\begin{aligned} \frac{d^2\sigma_{NC}^{e^\pm p}}{dxdQ^2} &= \frac{e^4}{8\pi x}\left(\frac{1}{Q^2}\right)^2 \Phi_{NC}^\pm\left(x,Q^2\right) \\ \frac{d^2\sigma_{CC}^{e^\pm p}}{dxdQ^2} &= \frac{g^4}{64\pi x}\left(\frac{1}{Q^2+M_W^2}\right)^2 \Phi_{CC}^\pm\left(x,Q^2\right). \end{aligned}$$

In the above formula, the strong Q^2 dependence introduced by the propagator terms $1/\left(Q^2+M_{Z/W}^2\right)$ has been separated from the *reduced cross-sections* $\Phi_{NC}^\pm$ and $\Phi_{CC}^\pm$

$$\begin{aligned} \Phi_{NC}^\pm &= Y_+\tilde{F}_2^\pm + y^2 F_L^\pm \mp Y_- x\tilde{F}_3^\pm \\ \Phi_{CC}^\pm &= Y_+\tilde{W}_2^\pm + y^2 W_L^\pm \mp Y_- x\tilde{W}_3^\pm \end{aligned}$$

that are mainly functions of the quark densities within the proton.

In contrast to the neutral current cross-section, which is mainly sensitive to the up and charm quark densities in e^+p as well as in e^-p collisions, the quark densities to which the charged current cross-sections are most sensitive to depend on the charge of the beam lepton. If the reduced charged current cross-sections are expressed in terms of the quark density functions,

$$\begin{aligned} \Phi_{CC}^+\left(x,Q^2\right) &= x\left(\bar{u}(x) + (1-y)^2\left(d(x)+s(x)\right)\right) \\ \Phi_{CC}^-\left(x,Q^2\right) &= x\left(u(x) + (1-y)^2\left(\bar{d}(x)+\bar{s}(x)\right)\right) \end{aligned}$$

[7] The charge of the up and charm quarks is $e_u = e_c = \frac{2}{3}$, yielding a factor $\frac{4}{9}$ in equation 2.7 that is significantly larger than the the corresponding factor $\frac{1}{9}$ for down, strange and bottom quarks, whose charge is $e_d = e_s = e_b = -\frac{1}{3}$.

[8] The Rapgap event generator was originally developed to accurately model "Rapidity Gap" processes, that is, for diffractive analyses, but has proven useful also in the non-diffractive sector.

it can be seen that in e^+p collisions the charged current cross-section is mainly sensitive to the down and strange quark densities at low y and the anti-up quark density at high y, while in e^-p collisions charged current scattering processes mainly probe the density of up quarks within the proton. Since within the proton the up quark density is larger than the down quark density, this also explains why in figure 2.4 the charged current cross-section is significantly higher for e^-p than for e^+p collisions.

Charged current scattering processes are simulated using the Django [72] event generator.

2.3 Photoproduction

In contrast to the deep-inelastic neutral and charged current scattering processes described in the last section, the magnitude of the four-momentum transfer Q^2 between electron and proton is very small in *photoproduction* processes [1]. At low Q^2, the emission of the heavy vector bosons W and Z is heavily suppressed by the propagator terms $1/\left(Q^2 + M^2_{Z/W}\right)$ in the cross-section, and photoproduction processes are mediated by quasi-real ("on-shell") photons. As the cross-section for photon exchange is proportional to $1/Q^4$, the total ep cross-section is in fact *dominated* by photoproduction processes.

For quasi-real photons, the cross-section for ep interactions may be decomposed into two factors,

$$\sigma_{ep} = \int dx_e f_{\gamma,e}\left(x_e\right) \sigma_{\gamma p}\left(x_e\right)$$

where the first factor $f_{\gamma,e}\left(x_e\right)$ is a universal *splitting function* describing the probability for an electron to emit a photon with momentum fraction x_e (analogous to the *Altarelli-Parisi* splitting functions for quarks and gluons) and the second factor contains all the process specific details. The *photon flux* $f_{\gamma,e}\left(x_e\right)$ of the electron may be written as

$$f_{\gamma,e}\left(x_e\right) = \frac{\alpha}{2\pi}\left(\frac{1+(1-y)^2}{y}\log\left(\frac{Q^2_{max}(1-y)}{m_e^2 y^2}\right) + 2m_e^2 y\left(\frac{1}{Q^2_{max}} - \frac{1-y}{m_e^2 y}\right)\right),$$

a result known as the *Weizsäcker-Williams approximation* [74, 75] in the literature. In this ansatz, the actual interaction may be described as being between the proton and a photon in a γp process with a centre-of-mass energy of [2]

$$W_{\gamma p} = \sqrt{(q+p)^2} \approx \sqrt{ys}.$$

[1] To be precise, "true" photoproduction only occurs in the limit $Q^2 \to 0$, an idealisation which cannot be experimentally reached at HERA. Instead, $Q^2_{max} \ll 1\ GeV^2$ is conventionally chosen as criterion to define photoproduction processes.

[2] $W^2_{\gamma p} = (q+p)^2 = -Q^2 + m_p^2 + 2q \cdot p$, where $2q \cdot p = 2p \cdot ky$ from the definition of y and $2p \cdot k = (p+k)^2 - m_p^2 - m_e^2 \approx (p+k)^2 = s - m_p^2 - m_e^2$, from which it follows that $W^2_{\gamma p} = -Q^2 + m_e^2 + ys$. Then, neglecting the electron mass and assuming $Q^2 \approx 0$ in photoproduction, the γp centre-of-mass energy may be expressed in the form $W_{\gamma p} \approx \sqrt{ys}$.

One convention worth noting is that the term "photoproduction" is predominantly used to label the non-resonant production of quark anti-quark pairs, although by definition, photoproduction only depends on the magnitude of the four-momentum transfer Q^2 and in principle summarizes various different kinds of processes; the production of particles other than quark anti-quark pairs is usually denoted by the species of the produced particles, e.g. lepton pair production (discussed in more detail in section 2.5). Following this convention, the description of the photoproduction processes is restricted to the production of quark anti-quark pairs in this section.

In the description of the quark anti-quark pair production processes, the so-called *direct* and *resolved* interactions of the photon are distinguished [3]. In direct processes, the photon couples as a point-like particle to quarks, the dominant process being the *boson-gluon fusion* process $\gamma g \rightarrow q\bar{q}$ shown on the left of figure 2.5. In resolved processes, the photon fluctuates to a quark anti-quark pair, which is then said to be "resolved" by the interaction with the proton. The actual interaction in such resolved processes is then between two partons. The two dominant contributions to resolved processes are the $qg \rightarrow qg$ and $gg \rightarrow gg$ processes shown on the right of figure 2.5 (in the figure, the symbol R denotes the so-called *photon remnant*).

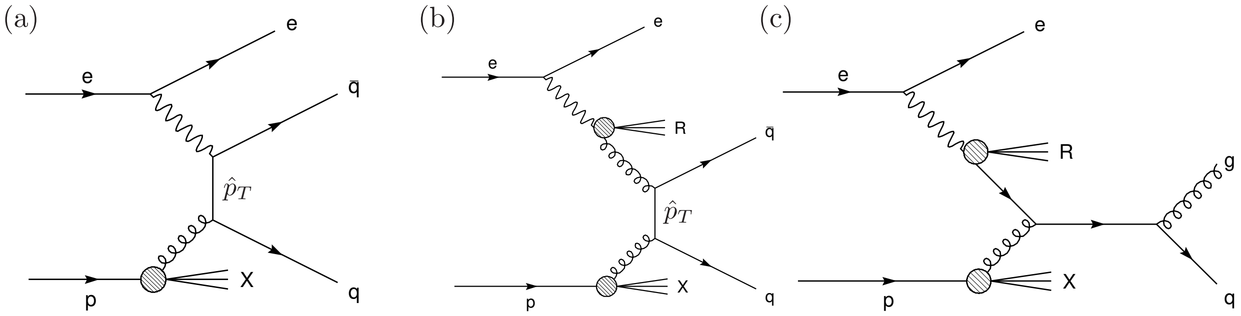


Figure 2.5: *Leading order Feynman diagrams for direct (a) and resolved (b and c) photoproduction.*

In analogy to the definition of Bjorken x, the momentum fraction of the interacting parton in the proton, resolved photoproduction processes are characterized

[3]There is actually some intrinsic ambiguity involved in the definition of direct and resolved processes, due to the quantum fluctuations inherent to both the photon and the proton. For example, the direct process depicted in figure 2.5 (left) may also be interpreted as a photon fluctuating to a quark anti-quark pair, which is then resolved by a gluon emitted by the proton. In fact, if some of the four-momenta in the process are space-like $p^2 < 0$, there is no clear time-ordering in the process and both interpretations of the interaction as direct and resolved process are possible. In practice, the ambiguity between direct and resolved processes is resolved by the introduction of an artificial parameter, the *factorisation scale*. This is a momentum transfer squared scale below which any parton activity is considered to be part of the parent structure and above which any parton activity is considered to be part of the interaction dynamics [76]. A common choice for the factorisation scale is to take the *largest* momentum transfer squared scale in any particular process (e.g. Q^2, transverse momenta or masses of particles). In the above example, if the invariant mass of the quark anti-quark pair is above the factorisation scale chosen to describe the process (e.g. Q^2), then the $q\bar{q}$ pair will be considered to be produced in the interaction, otherwise it will be ascribed to a quantum fluctuation of either the photon or the gluon (depending on which boson has the lower virtuality).

by the fraction of the photon's momentum x_γ participating in the interaction with the proton

$$x_\gamma = \frac{P \cdot p_q}{P \cdot q}.$$

Evaluation of the two four-vector products in the above definition shows that x_γ may be expressed in the two alternative forms [4]

$$x_\gamma = \frac{\sum_{jets} E - P_z}{\sum E - P_z} = \frac{1}{2E_e y}\left(E_{T_1} e^{-\eta_1} + E_{T_2} e^{-\eta_2}\right).$$

In the first form, the sum in the numerator extends over the two jets resulting from the produced quark anti-quark pair, while the sum in the denominator also includes contributions from the photon remnant. The distribution of x_γ observed in photoproduction of high p_T jets measured by the ZEUS collaboration is shown in figure 2.6. As is illustrated in the figure, the direct and resolved processes may be discriminated by a cut on the reconstructed x_γ at $x_\gamma^{cut} = 0.75$.

Whereas "hard" resolved processes (photons splitting into high p_T partons) may be calculated by perturbative QCD, interactions of "soft" resolved photons (photons fluctuating into low p_T partons) cannot be calculated perturbatively and have to be treated phenomenologically. In the absence of a hard scale, strong QCD interactions become important, which distort the photon splitting. Effectively, the produced quark anti-quark pair becomes *bound* by multiple gluon emission and absorption. In this domain, the hadronic structure of the photon may be approximated by that of a *meson*. Since parity and charge-conjugation symmetry are conserved in strong interactions, the meson state representing the photon has to have the same quantum numbers $J^{PC} = 1^{--}$ as the photon, and must therefore be a *vector* meson.

---

[4] The first form follows directly from the evaluation of the vector products in the rest frame of the proton,

$$x_\gamma = \frac{P \cdot p_q}{P \cdot q} = \frac{(E_p, 0, 0, E_p) \cdot (q - R)}{(E_p, 0, 0, E_p) \cdot q} = \frac{E_p\left(E_q - P_{z_q} - (E_R - P_{z_R})\right)}{E_p\left(E_q - P_{z_q}\right)} = \frac{\sum_{jets} E - P_z}{\sum E - P_z},$$

where the sum in the numerator extends over the two highest p_T jets, excluding the photon remnant, and the sum in the denominator extends over all particles in the detector (note that in photoproduction processes the scattered beam electron typically escapes detection and does not contribute to either sum). The alternative form may be derived from the above by using the relation

$$y = \frac{p \cdot q}{k \cdot p} = \frac{E_p\left(E_q - P_{z_q}\right)}{E_p\,(2E_e)} = \frac{E_q - P_{z_q}}{2E_e} = \frac{1}{2E_e}\sum E - P_z$$

for the inelasticity y and writing the $\sum_{jet_i} E - Pz$ contribution for each jet as

$$E - P_z = \sqrt{(E + P_z)(E - P_z)\frac{E - P_z}{E + P_z}} = \sqrt{(E^2 - P_z^2)\frac{E - P_z}{E + P_z}} = E_T\left(\frac{E - P_z}{E + P_z}\right)^{\frac{1}{2}} = E_T e^{-\eta},$$

where the last steps follows directly from the definition of the *rapidity* $\eta = \frac{1}{2}\ln\frac{E+P_z}{E-P_z}$. Then, adding together the terms for the two jets in the numerator and expressing the denominator in terms of y, it follows that

$$x_\gamma = \frac{1}{2E_e y}\left(E_{T_1} e^{-\eta_1} + E_{T_2} e^{-\eta_2}\right).$$

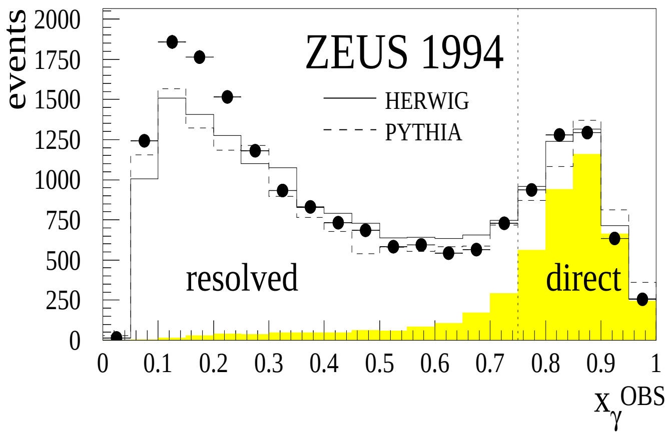


Figure 2.6: *The distribution of x_γ observed by the ZEUS collaboration in γp events with high p_T jets in comparison with the expectation for direct (hatched) and resolved (transparent) photoproduction (taken from [77]).*

In the *Vector Meson Dominance* model [78], the resolved photon is described as a superposition of $\rho(770)$, $\omega(783)$ and $\phi(1020)$ mesons [5] (numbers in brackets denote particle masses in MeV) in the mixing ratio $f_\rho : f_\omega : f_\phi = 9 : 1 : 2$, determined by the coupling of the vector mesons to the photon, which is proportional to the squared charge:

$$
\begin{aligned}
\rho &= \frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right) &\Rightarrow\quad & f_\rho \sim \frac{1}{2}\left(\frac{2}{3} - \left(-\frac{1}{3}\right)\right)^2 = \frac{1}{2} \\
\omega &= \frac{i}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right) &\Rightarrow\quad & f_\omega \sim \frac{1}{2}\left(\frac{2}{3} + \left(-\frac{1}{3}\right)\right)^2 = \frac{1}{18} \\
\phi &= s\bar{s} &\Rightarrow\quad & f_\phi \sim \left(-\frac{1}{3}\right)^2 = \frac{1}{9},
\end{aligned}
$$

assuming flavour symmetry [6]. The transition between perturbative QCD and the Vector Meson Dominance model is regulated by a *cut-off parameter* that is related to the scale of the "hard" interaction. In the calculation of photoproduction cross-sections, the hard scale is most often defined as the transverse momentum $\hat{p}_T$ of the virtual quark in figure 2.6.

The direct and resolved photoproduction processes are simulated by a variety of different Monte Carlo generators. High p_T direct and resolved production of light

[5] The original model was later extended to include heavier vector mesons (ρ' (1450), J/Ψ (3097), Υ (9460), ...) and also unbound $q\bar{q}$ pairs [79, 80].

[6] Actually, since the same couplings enter in the decay rates $vm \to e\bar{e}$, the mixing ratios f may be *measured* from the widths $\Gamma\left(vm \to e\bar{e}\right)$ [81]:

$$
f \sim \frac{\Gamma\left(vm \to e\bar{e}\right)}{m_{vm}}.
$$

Using the world averaged values for the masses and decay widths [82], the photon is found to mix in the ratio $f_\rho : f_\omega : f_\phi = 9.3 \pm 0.4 : 1 : 1.8 \pm 0.2$, in agreement with the expectation for charge squared couplings and flavour symmetry.

(u,d,s) quark anti-quark pairs is simulated by the *Pythia* [83] and *Herwig* [84] Monte Carlo generators, complemented by *PhoJet* [85] in the low p_T domain of the Vector Meson Dominance model. Heavy (c,b) quark production processes are simulated by both Pythia and *Aroma* [86] [7].

The Monte Carlo predictions for the photoproduction of heavy quark pairs seem to underestimate the cross-section measured at HERA by a factor of about two [87], in contrast to those for the photoproduction of light quark pairs, which are in good agreement with the measurements. At time of writing, the discrepancy observed in heavy flavour analyses constitutes an unresolved mystery. In the analysis presented in this thesis, the observed discrepancy is accounted for by scaling up the Monte Carlo expectation for the photoproduction of charm and bottom quark pairs by a factor of two and attributing a systematic uncertainty of 50% to it (note that the top quark is too heavy to be produced in pairs at HERA).

2.4 Parton Showering and Hadronisation

For a complete description of the electron-proton interaction, one aspect is still missing: the description presented so far ended with the production of quarks, whereas in nature, no free quarks have yet been observed. The reason for the non-observability of single quarks is *confinement*, a further important phenomenological property of Quantum-Chromodynamics.

In QCD, due to the gluon self-interactions, the field lines of the colour field are pulled together into a tube or string (see figure 2.7). As a consequence, the colour

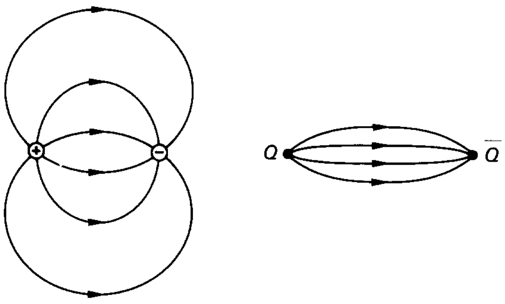


Figure 2.7: *Field lines for electromagnetic force between two electric charges (a) and strong force between two colour charges (b) (taken from [23]).*

field is concentrated in a limited spatial region and does not decrease as $1/r^2$ for increasing distance r from the (colour) charge as the electric field does [1]. The actual

[7]The Monte Carlo generators for "photoproduction" processes are used to simulate ep interactions up to $Q^2 \leq 4\ GeV^2$, so as to link up with the region $Q^2 > 4\ GeV^2$ simulated by Monte Carlo generators for DIS processes. The threshold value $Q^2 = 4\ GeV^2$ corresponds to the typical magnitude of the four-momentum transfer at which the scattered beam electron is within the geometrical acceptance of the main H1 detector (SpaCal).

[1]In classical electrodynamics, the $1/r^2$ decrease of the field strength for increasing distance r from an electric point charge is the result of Gausses law $\int_S \vec{E} \cdot \vec{n}\, dA = \int_V \nabla \cdot \vec{E}\, d^3x = \frac{q}{\epsilon_0} \equiv const$ for a surface S enclosing the volume V containing the point charge q. In the last step, Maxwell's equations have been used. For an electric field $\int_S \vec{E} \cdot \vec{n}\, dA = 4\pi r^2 E\,(r)$, from which it follows that

form of the strong interaction potential has been determined to be

$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r} + kr, \text{ with } k \approx 0.85 \text{ GeV fm}^{-1}, \tag{2.7}$$

by analysis of the linear dependence between angular momentum and squared mass for different families of baryon resonances [23].

Because of the linearly rising term in the colour potential, the energy needed to pull a single quark out of a hadron rises to infinity; at some point the creation

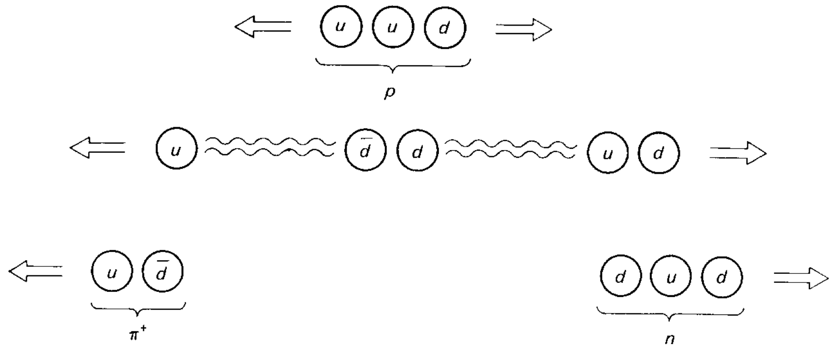


Figure 2.8: *Symbolic illustration of the hadronisation process. If a single quark is pulled out of a proton, at some separation distance the creation of a new quark anti-quark pair is energetically favoured over the further extension of the colour field, such that in the final state a pion and a neutron is observed (taken from [88]).*

of a new quark anti-quark pair is energetically favoured over a further increase in the length of the colour field (see figure 2.8). The creation of quark anti-quark pairs continues until all colour charges have formed colour-neutral bound states, the colourless baryons (qqq-states) and mesons ($q\bar{q}$ states) that are observed in nature.

In the context of electron-proton scattering, the process of baryon and meson formation may be divided into two distinct stages. In the first stage, the *primary* particles produced in the "hard" scattering have high enough energy to allow for the radiation of *secondary* quarks and gluons to be calculated in perturbative-QCD. Within time, more and more particles are produced in these *parton showering* processes, resulting in a decrease of the average particle energy. Eventually, the particle energies fall below a threshold of $\mathcal{O}(1\ GeV)$. At this moment, perturbative QCD calculations cannot be used any more, as the strong coupling α_s representing the expansion parameter becomes too large for the perturbative series to converge. The further evolution of baryon and meson formation cannot be reliably calculated in QCD [2] and has to be described by theoretically motivated empirical models. In

$E(r) = \frac{q}{4\pi\epsilon_0 r^2}$. For a colour field that is concentrated in a tube of constant diameter d, the integral $\int_S \vec{E}\cdot\vec{n}\,dA$ is independent of the distance r from the colour charge, so the strength of the colour field does *not* decrease.

[2]As described above, perturbative QCD calculations cannot be used at low particle energies. An alternative approach is to evaluate field integrals numerically on a discrete lattice instead of analytically in continuous space; the reliability of these *lattice QCD* calculations is limited so far, however, by the available computational power that does not allow for a sufficiently fine lattice granularity in the space-time volumes necessary to describe the interactions.

the "soft" *hadronisation* processes described by these empirical models, gluons split into quark anti-quark pairs and quarks combine to form colour-neutral hadrons. In order to form colour-neutral combinations, further quark anti-quark pairs may be created in this second stage.

For the simulation of parton showering and hadronisation processes, there exists a variety of Monte Carlo models. The two approaches most widely used for the simulation of parton showers are the *Parton Showers* (PS) and the *Colour Dipole Model* (CDM); for the simulation of the hadronisation process, the two most widely used approaches are *Lund String Fragmentation* and the *Cluster Decay Model*:

Parton Showers

In the PS approach, parton showers are modelled using the emission probabilities of quarks and gluons, calculated from the *Altarelli-Parisi* splitting functions [89–91]. The splitting functions $P_{\mathcal{AB}}(x, y)$ describe the probability for a parton $\mathcal{B}$ with momentum fraction y to emit a parton $\mathcal{A}$ with momentum fraction x. The radiation of quarks and gluons is described by a set of four splitting functions, which are illustrated in figure 2.9. In the parton showers resulting from these splitting functions,

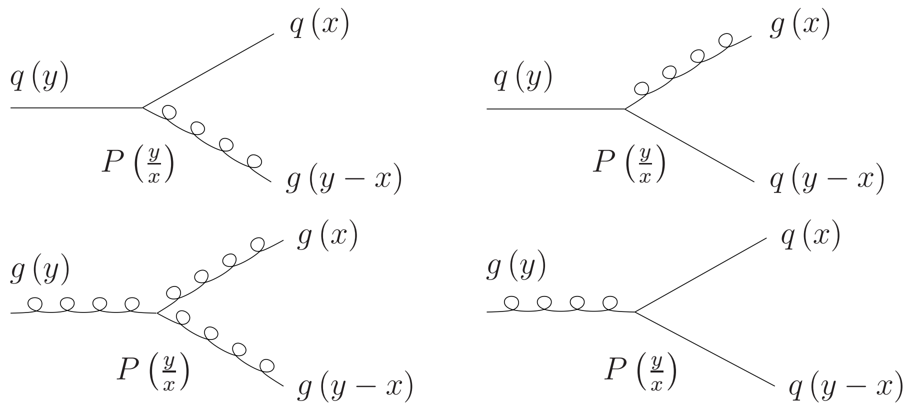


Figure 2.9: *Feynman diagrams describing the four possible Altarelli-Parisi splitting functions for the emission of quarks and gluons (taken from [31]).*

coherence effects are simulated by *angular* ordering of the radiated gluons.

Colour Dipole Model

The colour dipole model [92] represents an alternative to Altarelli-Parisi type splitting functions. In the CDM, parton showers are implemented as a *dipole cascade*, based on the interpretation of quark anti-quark systems as *colour dipoles* which act as "antenna" for gluon radiation.

Lund String Fragmentation

The Lund string fragmentation model [93] is based on the concept of string-like colour fields between colour charges (discussed in relation to the linear form of the QCD potential in equation 2.7). In this model, diverging quark anti-quark pairs produced by parton showers are separated until the string energy determined by the linear potential of the colour field exceeds twice the quark mass. At this point, the colour string between the quarks is broken into two segments by the creation of

Monte Carlo program	Algorithms used to implement Parton Showers	Hadronisation
Ariadne	Altarelli-Parisi (LO)	Lund-String
Herwig	Altarelli-Parisi (LO/NLO)	Cluster Decay
Pythia/JetSet	Dipole Cascade	Lund-String

Table 2.1: *Combinations of parton shower and hadronisation models implemented in popular Monte Carlo programs (table compiled from information in reference [96]).*

another quark anti-quark pair. The two string segments then continue to stretch as the quarks diverge. By this procedure, the string segments are consecutively split into smaller segments until eventually all quarks are close to anti-quarks, and the energy available no longer suffices for the creation of further quark anti-quark pairs. At this stage, the produced quarks and anti-quarks are combined to colour-neutral hadrons.

Cluster Decay Model

The cluster decay model [94] is based on the idea of colour *pre-confinement*, the production of colourless objects early on in the evolution of the hadronisation process. In the cluster decay model, all gluons produced in the preceeding parton shower stage are first split into quark anti-quark pairs. Neighbouring quarks and anti-quarks are then combined into colour-neutral clusters, which are finally isotropically decayed into hadrons.

A feature of QCD worth noting in relation to the empirical parton shower and hadronisation models is that the parton shower and hadronisation stages *factorize* from the details of the preceding "hard" scattering process. In other words, once the model-specific parameters of the parton shower and hadronisation models have been determined experimentally by analyzing a limited set of "hard" processes, all other "hard" processes can be completely predicted.

In commonly used Monte Carlo programs, different combinations of parton shower and hadronisation models are available, allowing for alternative models to be compared and uncertainties arising from the choice of a particular model to be estimated. An overview of the combinations of models implemented in the most widely used Monte Carlo simulations *Ariadne* [95], *Herwig* [84] and *Pythia* in combination with *JetSet* [83] is given in table 2.1.

2.5 Specific Background Processes

Following the convention that "photoproduction" processes refer to the production of quark anti-quark pairs only, the total ep cross-section is composed of the cross-sections of additional "exclusive" processes producing specific final states.

In the search for events with isolated leptons and large missing transverse momentum presented in this thesis, processes are considered as "background" if either the isolated lepton is due to a misidentified hadron or the reconstructed large missing transverse momentum is due to poor measurements in the detector. In this section, the most relevant such background processes are discussed.

2.5.1 Lepton-Pair Production

Lepton-pair production processes may contribute to samples of events with single leptons if one of the paired leptons escapes detection. At HERA, the dominant production process for lepton-pair events is the *photon-photon interaction* illustrated in figure 2.10. In the cross-section calculation of the depicted process, the Weizsäcker-

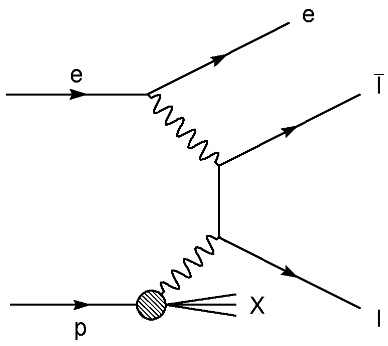


Figure 2.10: *Leading order Feynman diagram for lepton-pair production.*

Williams approximation may be applied twice to factorize out the photon-fluxes of the electron and proton, thus effectively reducing the $ep \rightarrow e\mu\bar{\mu}X$ *electroproduction* process to a $\gamma\gamma \rightarrow \mu\bar{\mu}$ interaction between two photons.

In comparison to the dominant photon-photon contribution, other electroweak processes involving the emission of virtual Z bosons (instead of photons) and the conversion of *bremsstrahlung* photons to lepton pairs (including the so called *Cabibbo-Parisi* and *Drell-Yan* processes) yield only small corrections to the cross-section in most regions of the phase-space [97, 98].

The direct production of lepton-pairs is simulated by the event generators *Lpair* [99, 100] and *Grape* [101, 102]. Grape includes the contributions from the full set of leading order electro-weak processes and approximates next-to-leading order effects by simulating initial and final state radiation, while the Lpair Monte Carlo generator only includes the photon-photon process in leading order. As the dominant contribution to the lepton-pair production cross-section is due to the photon-photon process illustrated in figure 2.10, both generators predict similar cross-sections [101, 103], despite the simplified simulation implemented in the Lpair generator.

The predictions of the Grape generator have been found to be in good agreement with recent analyses of high p_T di-electron [104] and di-muon [105] events. An analysis of di-tau events has been impeded so far by the lack of an algorithm for tau identification. First results on tau pair production based on the tau identification algorithm developed as part of this thesis are presented in section C.3 of the appendix.

2.5.2 Production of Vector Mesons

An additional contribution to events with isolated leptons can arise from the production and subsequent decay of vector meson resonances. In particular, the J/Ψ ($c\bar{c}$) and Υ ($b\bar{b}$) resonances and their excited states have sizeable leptonic decay branching fractions [1]. As shown in figure 2.11, the heavy quark vector mesons J/Ψ

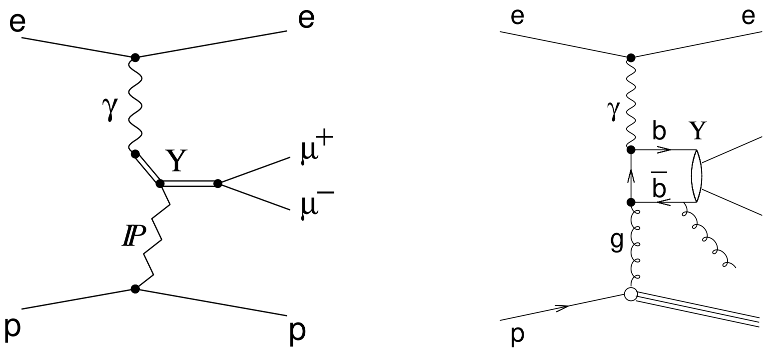


Figure 2.11: *Leading order Feynman diagrams for the elastic (left) and inelastic (right) production of vector mesons; elastic production is described by the exchange of a colourless object: the pomeron (taken from [97]).*

and Υ may be produced by either boson-gluon fusion processes with subsequent gluon radiation (necessary to make the quark anti-quark system colourless) or by *diffractive* processes, in which the photon fluctuates into a vector meson which then interacts with the proton by the exchange of a colourless object called the *pomeron*. As mentioned in section 2.4, the fluctuation of quasi-real photons into vector mesons is a "soft" process which cannot be calculated perturbatively in QCD, but has to be described phenomenologically. The diffractive production of vector mesons is described by Regge-phenomenology in combination with the Vector Meson Dominance model [2].

Production and decay of heavy quark vector mesons is implemented in the *Diffvm* [14] and *Epjpsi* [106] Monte Carlo generators.

The contribution from leptonic heavy quark vector meson decay is included in the backgrounds considered for this analysis, but is found to be negligible in the region of high lepton p_T.

[1] In comparison to other vector mesons, the branching fractions for leptonic decays are some orders of magnitude larger for the J/Ψ and Υ resonances and their excited states. The comparatively frequent leptonic decays of $c\bar{c}$ and $b\bar{b}$ resonances result from electro-weak decays; in contrast to most other vector mesons, the strong decays of the J/Ψ and Υ resonances and their excited states are suppressed, as $c\bar{c}$ and $b\bar{b}$ states are not heavy enough to decay into two D ($c\bar{q} + \bar{c}q$) respectively B ($b\bar{q} + \bar{b}q$) mesons, and because heavy flavour annihilating strong decays into pions are suppressed by the so-called "OZI"-rule [88].

[2] Although the diffractive production process cannot be *quantified* in QCD, the pomeron can of course be described *qualitatively* by the exchange of multiple quarks and gluons.

2.6 Standard Model Signal

In the search for events with isolated leptons and large missing transverse momentum, processes are considered as "signal" if the reconstructed lepton is indeed a genuine lepton (of the same species) and the reconstructed missing transverse momentum is due to a quasi-stable high p_T particle that does not interact with the detector material and thus escapes detection.

For the high p_T particle not to interact with the detector material, it has to be electrically neutral and must not initiate electromagnetic or hadronic showers. Within the Standard Model there exists only one class of such particles, namely neutrinos. The dominant Standard Model process producing isolated leptons and high p_T neutrinos in the final state is the production of real W bosons with subsequent leptonic decay of the W.

2.6.1 W Production

At HERA, real W bosons may be produced in neutral current

$$ep \rightarrow eWX$$

and charged current

$$ep \rightarrow \nu WX$$

interactions. A gauge invariant set of *leading order* (LO) Feynman diagrams describing (neutral current) $ep \rightarrow eWX$ processes are shown in figure 2.12. In addition to the diagrams shown, there exist equivalent ones with an anti-quark in the initial state.

Diagrams (a) and (b) are similar to "standard" neutral current interactions, with the additional radiation of a W boson from the *quark line*, occasionally referred to as *W-Strahlung*. The dominant contribution to the W production cross-section is due to the process shown in diagram (a), in which the W is emitted by the initial state quark. This diagram includes both a *t-channel* photon and a *u-channel* quark, that introduce two poles $1/t$ and $1/u$ into the production amplitude by the corresponding propagator terms [5].

Diagram (c) is of theoretical importance, as it includes a $WW\gamma$ triple gauge boson vertex and is thus sensitive to deviations from the triple gauge boson coupling predicted by the Standard Model. The possibility of probing the standard model gauge group structure was discussed by theorists in the early days of HERA [7,107], before LEP set quite strict limits on the *anomalous gauge couplings* [108]

$$\Delta\kappa_\gamma = -0.013 \pm 0.071 \quad \text{and} \quad \Delta\lambda_\gamma = -0.021 \pm 0.039$$

that parametrize possible deviations of the triple gauge boson coupling from the Standard Model prediction $\kappa_\gamma = 1$ and $\lambda_\gamma = 0$ [109]. In view of the limits set by LEP, the analysis of the triple gauge coupling is not so attractive anymore at HERA, as comparatively large anomalous gauge couplings are necessary to produce an observable effect in ep collisions.

Diagrams (d) and (e) are similar to the boson-gluon fusion process in direct photoproduction described in section 2.3 with the charged lepton and the neutrino

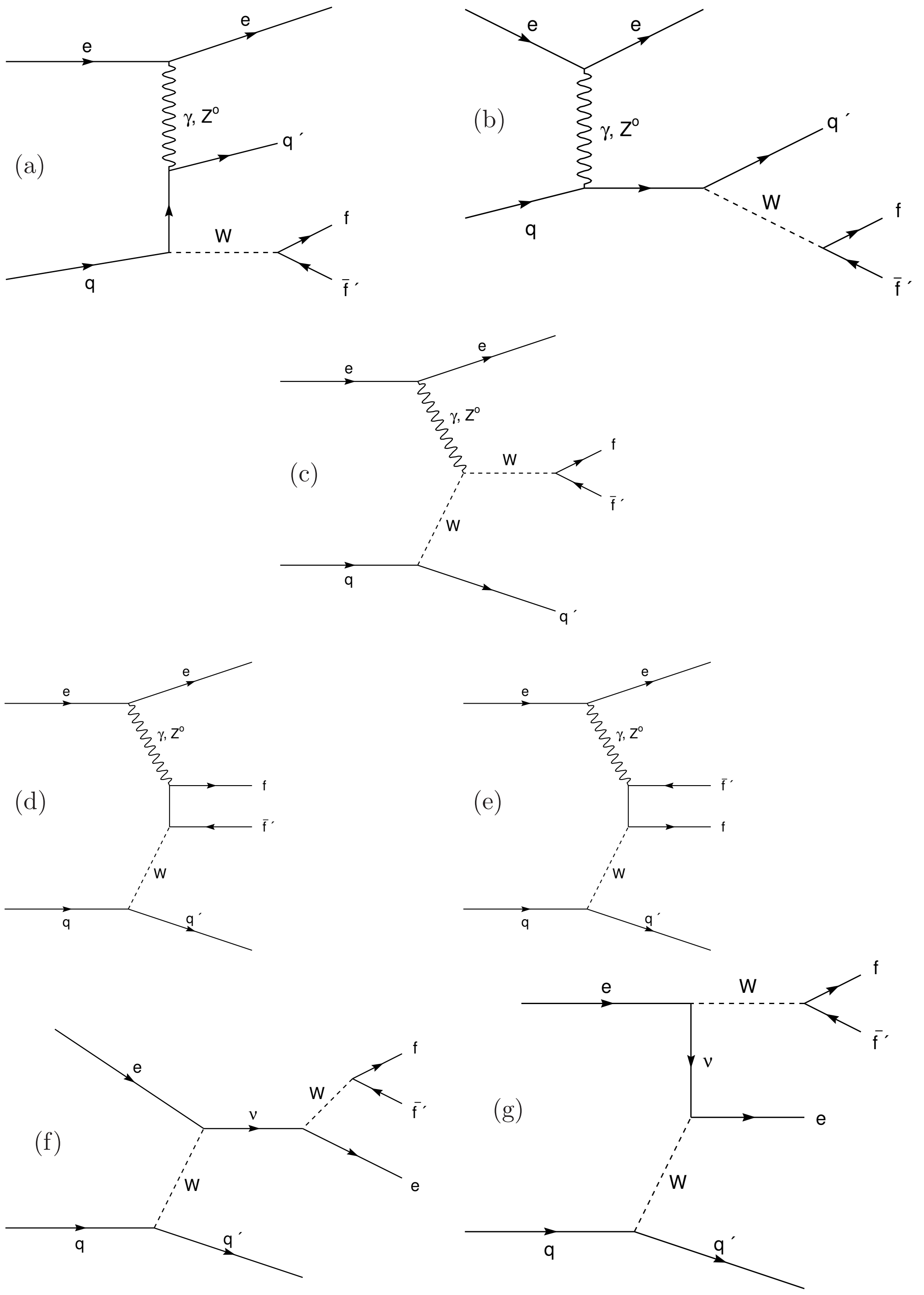


Figure 2.12: *Leading order Feynman diagrams for the production of real W bosons via neutral current $ep \to eWX$ processes at HERA and the subsequent decay $W \to f\bar{f}'$.*

taking the part of the quark anti-quark pair. Although these diagrams do not describe the production of real W bosons, but contain the W only as a virtual particle, they are required for the gauge invariant representation of the $ep \to eWX \to el\nu X$ process, which is advantageous for the numerical stability of the cross-section calculation. The contribution of diagrams (d) and (e) to the production of $el\nu X$ final states is small however, due to the propagator term associated with the virtual W.

An even smaller contribution to the cross-section is expected from the last two diagrams (f) and (g), in which the W is radiated off from the *lepton line*. These are heavily suppressed by the presence of a second W propagator term in the amplitude.

A related set of diagrams corresponding to (charged current) $ep \to \nu WX$ processes is displayed in figure 2.13. Since in these diagrams the contribution of the u- and t-channel diagram (a) dominating the $ep \to eWX$ cross-section is suppressed by the W propagator term $1/\left(Q^2 + M_W^2\right)$, $ep \to \nu WX$ processes amount to only about 10% of the total W production cross-section and will not be discussed in more detail here.

In the decay of W bosons, electrons, muons and tau leptons are produced with a *branching fraction* of about 10% each [82]. The directions of the produced leptons are only weakly correlated via the *helicity* of the W boson to the details of the ep interaction. As a consequence, the charged lepton and the neutrino produced in W decay are in general isolated from other particles in the detector. The transverse momenta of both particles are typically near the *Jacobian peak* [1] at around half the W mass:

$$p_T^{e,\mu,\tau} \sim p_T^{\nu} \sim \frac{M_W}{2} \approx 40\ GeV.$$

The distributions of transverse momenta of the produced charged leptons and neutrinos are shown in figure 2.14. As the neutrino escapes detection, its transverse momentum gets reconstructed as large missing transverse momentum

$$P_T^{miss} \approx p_T^{\nu}.$$

The transverse momenta P_T^X of the hadronic system are typically small. The expected distribution of the hadronic transverse momentum in Standard Model W production events is shown in figure 2.15. At small hadronic transverse momenta, the P_T^X distribution is steeply falling. For about two thirds of the events, the transverse momentum of the hadronic system is expected to be below $P_T^X < 12\ GeV$. At larger hadronic transverse momenta, the P_T^X distribution falls much less steeply. For about one sixth of the events (approximately half the events with $P_T^X > 12\ GeV$), hadronic transverse momenta above $P_T^X > 25\ GeV$ are expected.

The production of real W bosons in ep collisions is implemented to leading order in the *EPVEC* [7, 110] Monte Carlo generator. Recently, a *next-to-leading order (NLO)* calculation of the cross-section was performed [8,9], which has been found to reduce the theoretical uncertainty from 30% to 15%. The improved precision of the next-to-leading order calculation is taken advantage of in the analysis presented in this thesis, by reweighting the events generated by EPVEC with the ratio σ_{NLO}/σ_{LO} of the next-to-leading order to leading order cross-sections.

---

[1] The Jacobian peak may be understood as arising from a singularity of the Jacobi determinant $d\theta/dp_T$ for $\theta \to 90°$:

$$\frac{d\sigma}{dp_T} = \frac{d\sigma}{d\theta}\frac{d\theta}{dp_T} = \frac{1}{E\ \cos\theta}\frac{d\sigma}{d\theta}.$$

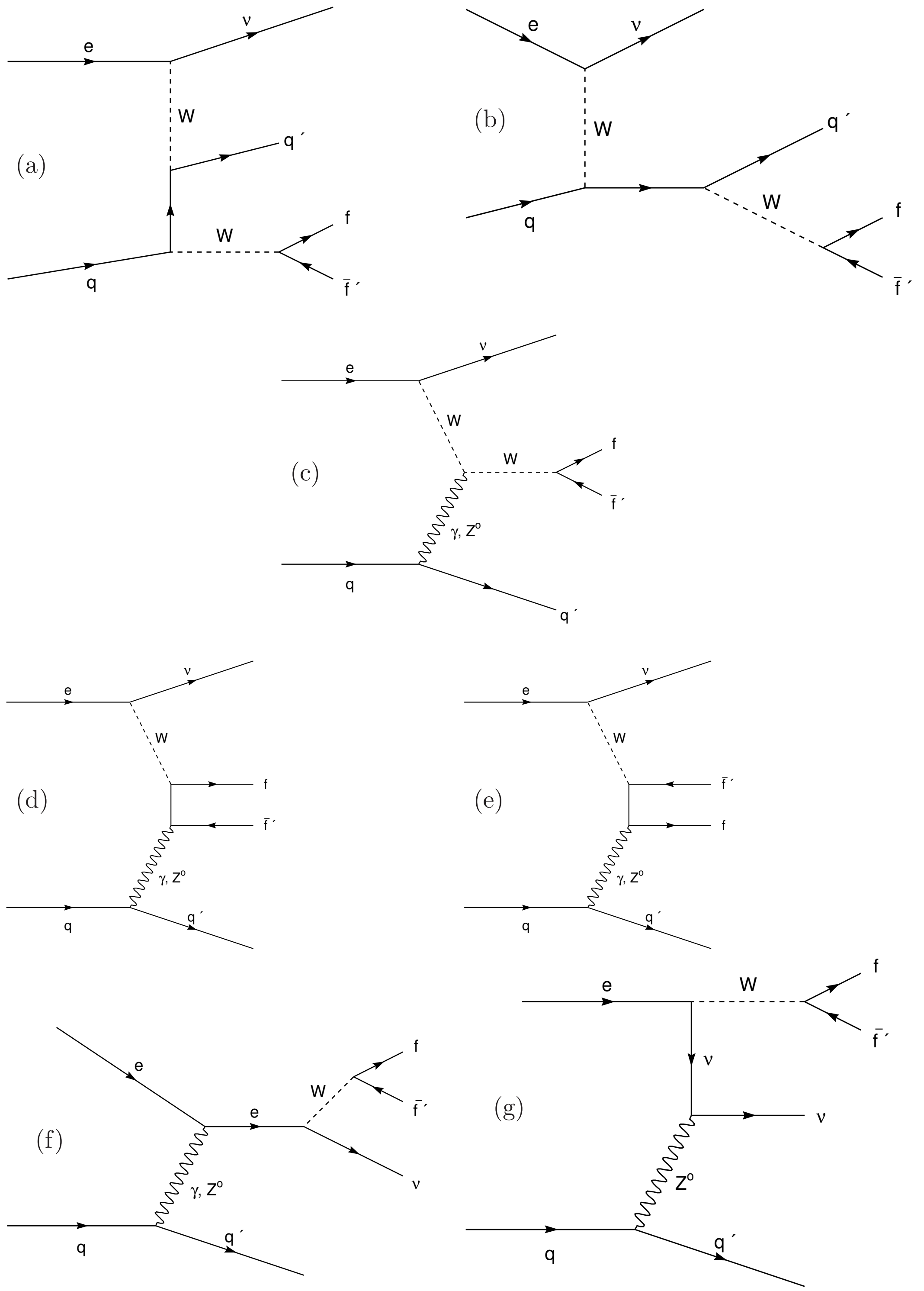


Figure 2.13: *Leading order Feynman diagrams for the production of real W bosons via charged current $ep \rightarrow \nu W X$ processes at HERA and the subsequent decay $W \rightarrow f\bar{f}'$.*

2.6.2 Z Production

The production of real Z bosons may contribute to the signal, if the Z bosons decay into two neutrinos. The branching fraction for this decay mode is about 20%

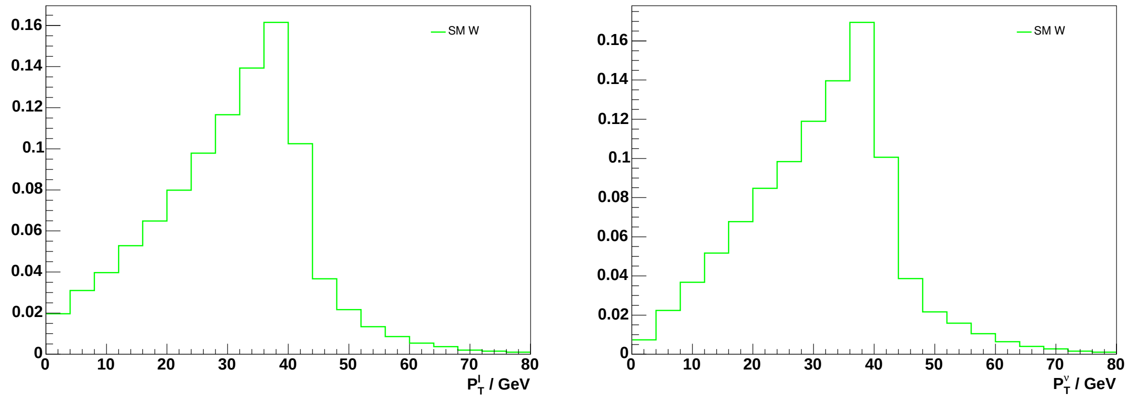


Figure 2.14: *Distribution of the transverse momenta of the charged lepton (left) and the neutrino (right) produced in the decay of real W bosons.*

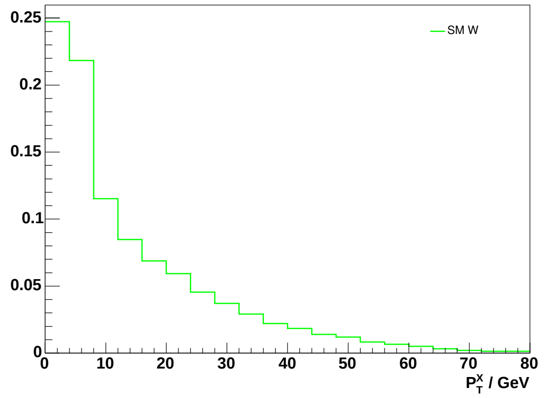


Figure 2.15: *Distribution of the hadronic transverse momentum in Standard Model W production events.*

[82]. The Feynman diagram describing the dominant contribution to the production of real Z bosons at HERA with subsequent decay into two neutrinos is shown in figure 2.16. It represents the equivalent process to the dominant contribution to the production of real W bosons shown in figure 2.12 (a), differing only in that it is a Z rather than a W that is radiated off from the quark line. In analogy to the radiation of real W bosons, this process is referred to as *Z-Strahlung*.

The events produced by Z-Strahlung may contribute to the sample of events with isolated electrons and large missing transverse momentum, if the beam electron is scattered into the detector (LAr) and identified as isolated lepton and the Z boson is produced with a large transverse momentum (if the Z boson decays into two neutrinos, its transverse momentum is reconstructed as missing transverse momentum). In the analysis presented in this thesis, the contribution from this process is suppressed, however, by selection criteria that are designed to reject the background from neutral-current DIS processes (see section 7.1.3).

In the final event sample, the dominant Z production contribution is due to the *Cabbibo-Parisi* [98] process described by the Feynman diagram displayed in figure 2.17. In this process, a photon emitted by the proton splits into an electron-positron

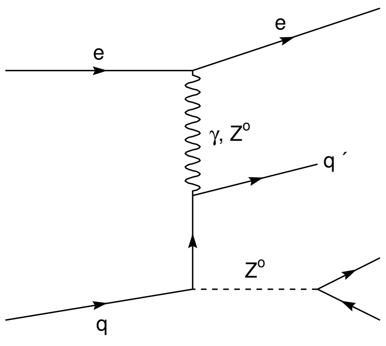


Figure 2.16: *Leading order Feynman diagram for the production of real Z bosons in the process $ep \rightarrow eZX$ and the subsequent decay $Z \rightarrow \nu\bar{\nu}$.*

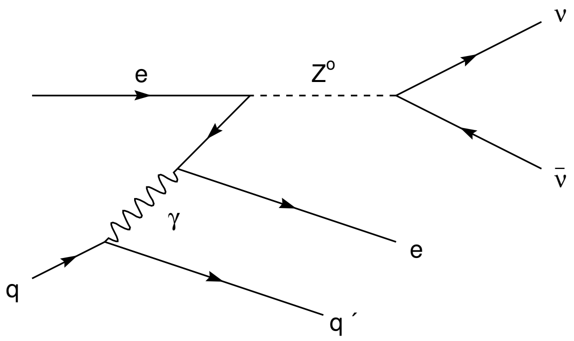


Figure 2.17: *Leading order Feynman diagram for the production of real Z bosons in the Cabbibo-Parisi process and subsequent $Z \rightarrow \nu\bar{\nu}$ decay.*

pair; one of these then annihilates with the beam lepton to produce a Z boson. The events produced by Cabbibo-Parisi type processes may contribute to the signal if the second particle of the produced e^+e^- pair is identified as isolated lepton in the detector and the Z boson is produced with a large tranverse momentum. As most Z bosons are produced with small transverse momenta, the contribution from Cabbibo-Parisi type processes to samples of events with large missing transverse momenta is suppressed, however, and the contribution from Z boson production to the signal amounts to less than 5% of the contribution from W boson production.

The production of real Z bosons is implemented in the EPVEC Monte Carlo generator.

2.7 Beyond the Standard-Model

In theories *beyond the Standard Model (BSM)*, events with isolated leptons and large missing transverse momentum may be produced by any process producing an isolated lepton in association with any particle that does not interact with the detector material and escapes detection in the final state. In addition to neutrinos, possible candidates for such undetected particles are for example *neutralinos* in supersymmetric theories.

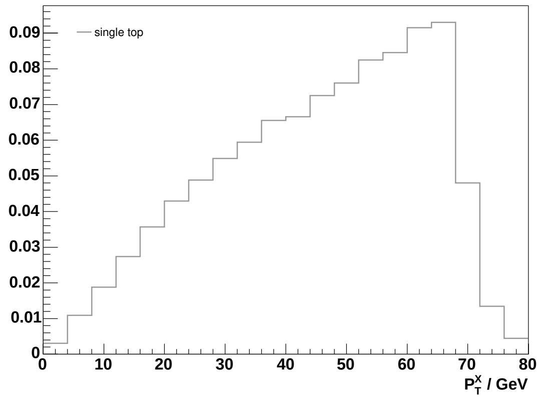


Figure 2.18: *Distribution of the hadronic transverse momentum expected in single top production events.*

The BSM theories being studied in this context are the production of *lepto-quarks* (LQ) [111–113] and *excited fermions* (f^*) [114, 115], *R-parity violating Supersymmetry* ($\not R_p$-SUSY) [116–121] and the production of single top quarks by *flavour-changing neutral current* (FCNC) interactions [11, 122]. In previous analyses, the production of single top quarks has been quoted as being the most likely of these theories to explain the excess of events with isolated leptons and large missing transverse momentum observed at large hadronic transverse momenta [3, 5] (*cf.* the introduction). In this section, only the production of single top quarks by FCNC interactions is therefore described in detail. The other theories beyond the Standard Model will be discussed in chapter 9.

2.7.1 Single top Production

In the decay of single top quarks, isolated charged leptons and neutrinos of large transverse momentum are produced in the decay $t \to bW$ (the dominant decay mode of top quarks), with subsequent leptonic decay of the produced W boson. The transverse momenta of the produced bottom quarks are typically of the order of

$$p_T^b \sim \frac{M_{top} - M_W}{2} \sim 50\ GeV.$$

The distribution of the transverse momentum of the b-jets resulting from hadronisation of the bottom quarks is shown in figure 2.18. In about three quarters of the simulated events, the transverse momentum of the hadronic system $P_T^X \approx p_T^b$ is above $P_T^X > 25\ GeV$. The production of single top quarks would therefore naturally explain the excess of events with isolated leptons and large transverse momentum observed at large hadronic transverse momenta in previous analyses.

Within the Standard Model, the cross-section for the production of single top quarks is negligible [1]. At leading order of the Standard Model, the production of single top quarks by *neutral current* interactions is excluded by the $SU(2) \otimes U(1)$ gauge group structure of the electro-weak force in combination with the unitarity of the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix that describes the probability for

[1] Note again that centre-of-mass energy is not sufficient to produce top quarks in pairs at HERA.

flavour transitions. This effect is known as the *Glashow-Iliopoulos-Maiani (GIM) mechanism* [123] in the literature and is experimentally confirmed by the rarity of the decay $K^0 \rightarrow \mu\bar{\mu}$. As a consequence of the GIM mechanism, FCNC interactions are possible only via loop diagrams and are therefore negligible within the Standard Model. The production of single top quarks by *charged current* interactions is illustrated in figure 2.19. Although this process is possible at leading order, it is heavily

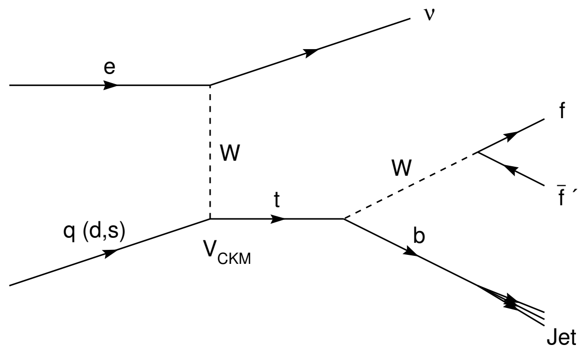


Figure 2.19: *Leading order Feynman diagram for the production of single top quarks within the Standard Model.*

suppressed by the propagator term $1/\left(Q^2 + M_W^2\right)$ in addition to the smallness of the off-diagonal elements of the CKM matrix describing the flavour transition probabilities. As a consequence, the Standard Model cross-section for the production of single top quarks is smaller than 1 fb at HERA, about three orders of magnitude too small to be observable.

In comparison to the Standard Model cross-section, the production of single top quarks may be significantly enhanced in theories beyond the Standard Model by loop contributions of new particles. The possible contributions of BSM processes to flavour-changing neutral current processes are parametrized by an effective coupling $\kappa_{tu\gamma}$ that describes the photon exchange mediated transition of an up into a top quark [2]. Within the Standard Model, this effective coupling is zero at leading order and almost zero at higher orders.

For sufficiently large effective couplings $\kappa_{tu\gamma}$, single top quarks may be produced at an observable rate in theories beyond the Standard Model. The Feynman diagram describing this process is illustrated in figure 2.20. The current experimental upper limit on the anomalous coupling $\kappa_{tu\gamma} < 0.18$ [124], is compatible with a cross-section of

$$\sigma\left(ep \rightarrow etX\right) \ < \ 0.24\ pb \tag{2.8}$$

at the 95% confidence level (CL) (and a centre-of-mass energy of $\sqrt{s} = 318GeV$). This cross-section is sufficiently large to explain the excess of events with isolated

[2]At HERA, there is practically no sensitivity to the coupling v_{tuZ} for FCNC interactions mediated by Z exchange, due to the suppression resulting from the high mass of the Z boson in the propagator term $1/\left(Q^2 + M_Z^2\right)$. An anomalous coupling of the top to the charm quark is also not observable at HERA; within the Standard Model, charm quarks exist only as "sea" quarks in the proton and have very low probability to be found at the high Bjorken $x \gtrsim 0.3$ required for the production of heavy top quarks. Even in *intrinsic charm* models, which allow charm quarks to exist at high x in the proton, the predicted cross-section for transitions of charm to top quarks is too small to be observable [122].

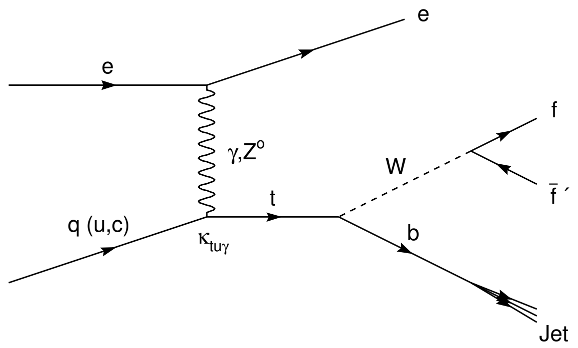


Figure 2.20: *Leading order Feynman diagram for the production of single top quarks by flavour-changing neutral currents.*

leptons and large missing transverse momentum observed in previous analyses. Incidentally, the most stringent limit on the anomalous coupling is set by searches for single top production with subsequent hadronic W decay at HERA. The current limits on the anomalous couplings $\kappa_{tu\gamma}$ set by HERA (H1 [11] and ZEUS [124]), LEP (L3 [125]) and the TeVatron (CDF [126]) are shown in figure 2.21.

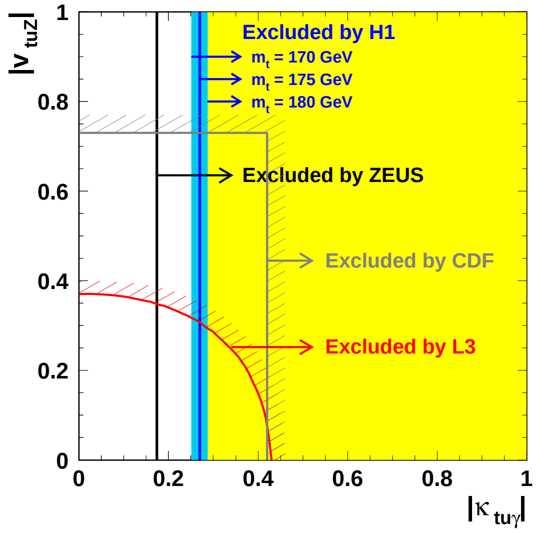


Figure 2.21: *Exclusion limits on the anomalous coupling $\kappa_{tu\gamma}$ that parametrizes transitions between up and top quarks mediated by photon exchange determined at HERA (H1, ZEUS), LEP (L3) and the TeVatron (CDF); also shown are the L3 and CDF limits on the coupling $v_{tu\gamma}$ that describes flavour-changing neutral current interactions mediated by the exchange of Z bosons (all limits at 95% CL).*

The production of single top quarks is implemented in the ANOTOP Monte Carlo generator [11]. The simulated events are weighted proportional to $\kappa^2_{tu\gamma}$, to account for the dependence of the cross-section on the square of the anomalous coupling.

Chapter 3

Pattern Classification and Neural Networks

In searches for new physics it is often the case that possible signals are hidden underneath a background that may be orders of magnitude larger. In fact, new physics is often indicated by rare processes, since otherwise the anticipated signals would already have been found in preceding experiments or previous analyses of a lower luminosity data sample. In typical search analyses, the difficulty is therefore twofold: firstly, a large fraction of the background has to be suppressed , and secondly, the efficiency for identifying the signal must be high. In short, in analyses aiming to find evidence for new physics, frequently the need arises to effectively separate rare signals from much more common backgrounds.

Generally, the problem of effectively separating signal from background may be formulated as the problem of discriminating between different categories of entities. The ability to assign individual entities to distinct categories is of manifold scientific interest also in other disciplines and has accordingly evolved into a separate research subject: *pattern classification*.

In the first section of this chapter, a brief introduction to the basics of pattern classification is given (a more detailed description of pattern classification may be found in reference [127]). With the formalism developed in the first section, neural networks will then be described in the context of pattern classification in the second section.

3.1 Pattern Classification

The central objective of pattern classification is the *design* of a suitable *classifier*, that is, an algorithm that assigns entities of unknown *categories* to one of several *classes*. The assignment of entities to classes is based on *patterns*, sets of extracted *features* of the entities to be classified. As entities belonging to different categories may feature identical patterns, classification decisions are almost never completely free of errors. In case the probabilities for the different categories to feature a certain pattern are precisely known, an optimal classification algorithm exists, that provides the lowest possible probability for misclassification: the *Bayes decision rule*. In practical pattern classification applications, such extensive knowledge about the problem domain rarely exists, however, and other algorithms have to be used. In comparison to the ideal Base case, these alternative algorithms yield higher error-rates, that is, a larger fraction of entities are assigned to a class different from their category.

In the following four subsections, the basic ideas of pattern classification will be discussed in more detail. In the first subsection, the feature space representation of the entities to be classified will be described. That the Bayes decision rule indeed defines an ideal classification algorithm will be shown in the second subsection. In the third subsection, the Bayesian decision theory will be generalized to the framework of discriminant functions, in which all algorithms used in practical pattern classification applications may be formulated. The section about pattern classification then concludes with a discussion of how to compare the performance of different classifiers in the fourth subsection.

3.1.1 Feature Space Representation

The classification decisions of classifiers are based on the *values* of different features in the patterns to be classified. The patterns may contain only one, but will typically contain more than one feature; in any case, they may be represented as vectors of real-valued numbers. The associated vector space is termed the *feature space*; technically, the feature space is simply an ordinary d-dimensional Euclidean space $\mathbb{R}^d$. Its dimensionality d is equal to the number of different features in each pattern. In the feature space, an observed pattern is represented by a *feature vector*

$$\vec{x}_{observed} \subset \mathbb{R}^d.$$

By dividing the feature space into as many regions $\mathcal{R}_i \subset \mathbb{R}^d$ as there are categories, and introducing the convention that all feature vectors in region $\mathcal{R}_i$ are assigned to class $\tilde{\omega}_i$, the classification problem for c categories may be equivalently formulated as the problem of finding the partition

$$\mathcal{R}_1, \ldots \mathcal{R}_c \subset \mathbb{R}^d$$

in feature space that minimizes the probabilities for patterns in the category ω_j to be in the (wrong) region $\mathcal{R}_j$, $i \neq j$. The classification decision is fixed once the c different regions $\mathcal{R}_i$ in the feature space are determined. Accordingly, the regions $\mathcal{R}_i$ are termed the *decision regions*. Dividing the feature space into decision regions are the *decision boundaries*, that separate the feature vectors assigned to one class from the feature vectors assigned to other classes. The concept of decision regions is illustrated in figure 3.1 for the example of classifying two different types of fish.

The abstraction provided by the feature vector representation of the patterns allows the classification problem to be treated in a largely domain-independent manner - the same formalism may be used for separating salmon from sea bass as in the above example as for discriminating between signal and background events in high energy physics applications.

A problem not addressed so far is that of the optimal choice of features. Clearly, the chosen features should have *discriminative power*, that is, the features should be differently distributed among the different categories; a good feature representation

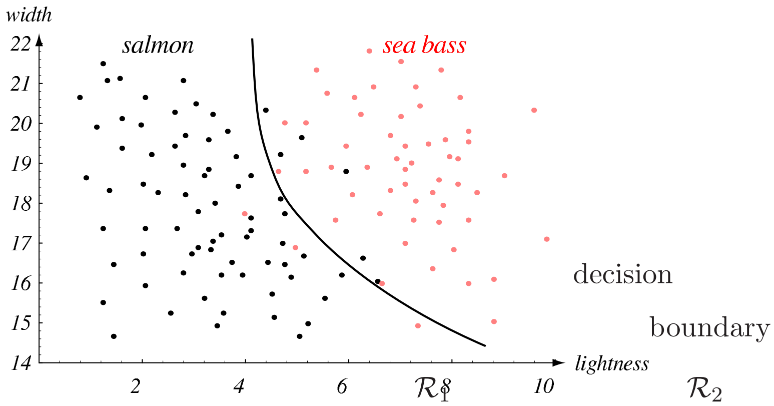


Figure 3.1: *Decision regions for a classification problem involving two types of fish based on the two features "lightness of scales" and "width" of the fish. All fish in decision region $\mathcal{R}_1$ ($\mathcal{R}_2$) are classified as salmon (sea bass). The features of salmon (sea bass) are represented by black (pink) points. The black (pink) points in decision region $\mathcal{R}_2$ ($\mathcal{R}_1$) represent salmon (sea bass) that are misclassified as sea bass (salmon) (taken from [127]).*

is one, in which the patterns belonging to the same category are close to one another, yet far from those of a different category, since this will lead to well separated decision regions in feature space [1].

3.1.2 Bayesian Decision Theory

In pattern classification applications, the true state of nature ω_i can rarely inferred with absolute certainty. The *uncertainty* that an observed pattern $\vec{x}_{observed}$ indeed belongs to category ω_i may be quantified by a *probability*, denoted as

$$P\left(\omega_i|\vec{x}\right).$$

In the ideal case that the true state of nature may be predicted to be ω_i with absolute certainty, the probability will be one; the probability will be zero if the pattern $\vec{x}$ does certainly *not* belong to category ω_i. As the probability $P\left(\omega_i|\vec{x}\right)$ may only be determined *after* all features in the pattern $\vec{x}$ have been measured, it is termed the *posterior probability*.

In Bayesian probability theory, the posterior probability is determined by the *prior probability* and the *class-conditional probability density*. For an illustration of the prior probability and class-conditional probability density, consider again the distributions of features shown in figure 3.1 for the example of two types of fish.

[1]To be precise, the notion of "close" and "far" depends on the definition of a *metric*, that is a mapping

$$D : \left(\mathbb{R}^d, \mathbb{R}^d\right) \to \mathbb{R}, \quad \left(\vec{x}, \vec{y}\right) \mapsto D\left(\vec{x}, \vec{y}\right);$$

here and in the following the *Euclidean metric*

$$D\left(\vec{x}, \vec{y}\right) = \left(\sum_{i=1}^{d} \left(x_i - y_i\right)^2\right)^{\frac{1}{2}}$$

will be assumed.

In the two dimensional feature space spanned by the "lightness of scales" and "width" features of the fish, the distribution of the salmon sample is concentrated at small lightness of scales and width, in contrast to sea bass which are generally wider and have lighter scales. Because the probability for a sea bass to be found with small lightness of scales and width is low, a thin fish with dark scales is very likely to be a salmon. This example illustrates that the posterior probability $P(\omega_i|\vec{x})$ for a certain pattern of features $\vec{x}$ to belong to category ω_i depends on the *class-conditional probability density*

$$p(\vec{x}|\omega_i)$$

for the feature vector $\vec{x}$ to be observed in a random sample of that category. More specifically, a category ω_i for which the probability density $p(\vec{x}|\omega_i)$ is large, is likely to be the right one.

The *prior probability* $P(\omega_i)$ describes some prior knowledge about the relative abundances of entities belonging to category ω_i in the set of patterns to be classified that exists *before* measuring any features of any patterns. In the above example, the prior probabilities for salmon and sea bass are about the same, indicated by an about equal number of black and red points in figure 3.1. Assume now, that the fish to be classified have been caught in some part of the ocean where sea bass are much more frequently found than salmon. In this case, the probability $P(\omega_2|\vec{x})$ for a given fish with features $\vec{x}$ to be a sea bass would increase. This consideration demonstrates that the posterior probabilities not only depend on the class-conditional probability densities, but also on the prior probabilities

$$P(\omega_i).$$

The exact relation between the prior probabilities $P(\omega_i)$, the class-conditional probability densities $p(\vec{x}|\omega_i)$ and the posterior probabilities $P(\omega_i)$ is determined by Bayesian decision theory. The *Bayes formula* states that the posterior probabilities are proportional to the product of the prior probabilities and the class-conditional probability densities. For the derivation of the Bayes formula, note first, that the *joint* probability density for finding a pattern of class ω_i with feature vector $\vec{x}$ may be written in two different ways,

$$p(\omega_i, \vec{x}) = P(\omega_i|\vec{x}) \cdot p(\vec{x}) = p(\vec{x}|\omega_i) \cdot P(\omega_i).$$

Dividing by the factor $p(\vec{x})$ then leads to the Bayes formula:

$$P(\omega_i|\vec{x}) = \frac{p(\vec{x}|\omega_i) \cdot P(\omega_i)}{p(\vec{x})}. \tag{3.1}$$

So, the probability for an observed pattern $\vec{x}$ to belong to category ω_i is proportional to the product of the prior probability of that category and the class-conditional probability density to observe the pattern of features $\vec{x}$ in a random sample of category ω_i.

If the prior probabilities and the class-conditional probability densities are precisely known for all categories ω_i, an optimal decision rule exists for classifying the feature vectors $\vec{x}$ belonging to unknown categories so as to minimize the expected misclassification rate. Formally, if a pattern $\vec{x}$ belonging to category ω_i is assigned to class $\tilde{\omega_j}$, then the decision is *correct* if $i = j$ and *in error* if $i \neq j$. If misclassification is to be avoided, the optimal decision rule is the one that minimizes the probability of error, that is, the *error rate*. Consequently, the optimal decision rule is the one with the smallest error rate. In case both the prior probabilities and the class-conditional probability densities are known for all categories ω_i, the smallest possible error rate is realized by the *Bayes decision rule* [127]:

$$\text{Decide } \tilde{\omega_i} \text{ if } p\left(\vec{x}|\omega_i\right) \cdot P\left(\omega_i\right) > p\left(\vec{x}|\omega_j\right) \cdot P\left(\omega_j\right) \;\; \forall \; i \neq j. \tag{3.2}$$

The Bayes decision rule may be rewritten in terms of the posterior probabilities $P\left(\omega_i|\vec{x}\right)$. Using relation 3.1 and omitting the factors $p\left(\vec{x}\right)$, which are common to all classes $\tilde{\omega_i}$ and thus do not influence the classification decision, the Bayes decision rule may be expressed as:

$$\text{Decide } \tilde{\omega_i} \text{ if } P\left(\omega_i|\vec{x}\right) > P\left(\omega_j|\vec{x}\right) \;\; \forall \; i \neq j.$$

This shows that the Bayes decision rule is equivalent to simply choosing the most probable hypothesis, in agreement with the common sense expectation.

That the Bayes decision rule indeed minimizes the expected misclassification rate can be of course shown formally in the multi-category case [127], but shall be demonstrated by an illustrative example here. Consider a two category problem based on a single feature x, in which the product of the prior probabilities and the class-conditional probability densities of the two categories (ω_1 and ω_2) are of the form shown in figure 3.2. Suppose there exists a non-Bayes decision rule that assigns

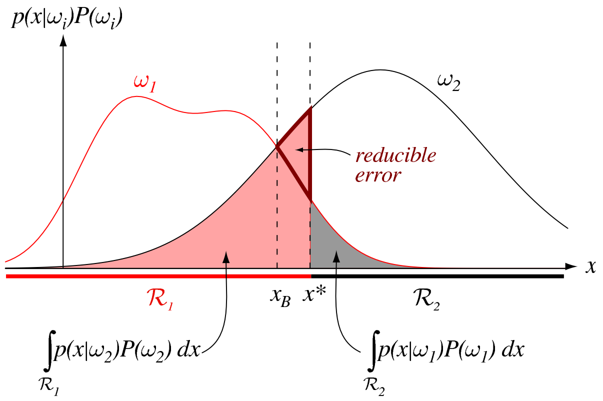


Figure 3.2: *Contributions to the expected misclassification rate in a two category classification problem based on a single feature x, resulting from the decision boundary chosen as x^*. The pink (light) area corresponds to the probability for assigning a pattern belonging to category ω_2 to class $\tilde{\omega_1}$; the grey (dark) area represents the converse, as given by equation 3.3. The decision boundary x^* is not chosen optimally. If the decision boundary were chosen as x_B instead, the probability for misclassification would be reduced by an amount indicated by the pink area framed in red; x_B corresponds to the Bayes decision rule and yields the minimal error rate (taken from [127]).*

all patterns with features x above some threshold x^* to class ω_2 and all patterns featuring x below x^* to class ω_1. Then, the probability of error expected for this decision rule is [127]

$$\begin{aligned} P\left(error\right) &= P\left(x \in \mathcal{R}_2, \omega_1\right) + P\left(x \in \mathcal{R}_1, \omega_2\right) \\ &= P\left(x \in \mathcal{R}_2 | \omega_1\right) \cdot P\left(\omega_1\right) + P\left(x \in \mathcal{R}_1 | \omega_2\right) \cdot P\left(\omega_2\right) \\ &= \int_{\mathcal{R}_2} p\left(x|\omega_1\right) \cdot P\left(\omega_1\right) dx + \int_{\mathcal{R}_1} p\left(x|\omega_2\right) \cdot P\left(\omega_2\right) dx. \end{aligned} \tag{3.3}$$

The individual contributions from the two integrals to the expected error rate are represented by the pink (light) and gray (dark) areas in the tails of the functions $p\left(x|\omega_i\right) \cdot P\left(\omega_i\right)$ in figure 3.2. The key to understanding these error rates is to note that patterns belonging to categories ω_1 and ω_2, which feature the same value x are *indistinguishable* for a classifier that is restricted to this feature (although the patterns may be distinguishable for a different choice of features). As a consequence, the classifier will assign *all* patterns featuring this value x to one and the same class $\tilde{\omega_i}$. Whenever a pattern belonging in truth to a different category ω_j, $j \neq i$ features the value x, it will be misclassified. In the classification of a particular value x, the probability of error is

$$P\left(error\right) = \begin{cases} P\left(\omega_2|x\right) & \text{if decision rule decides } \omega_1 \\ P\left(\omega_1|x\right) & \text{if decision rule decides } \omega_2 \end{cases}$$

depending on the actual decision rule of the classifier. For this value x, the misclassification probability is minimized by deciding ω_1 if $P\left(\omega_1|x\right) > P\left(\omega_2|x\right)$ and ω_2 otherwise. This is identical to the Bayes decision rule. Clearly, if the Bayes decision rule minimizes the probability of error for *any* one pattern, it likewise minimizes the error rate for *all* patterns in the integrands of equation 3.3. It has thus been shown that the Bayes decision rule indeed yields the smallest possible error rate for any given classification problem. The above example also demonstrates that the classification task becomes *simpler* the more the prior probabilities and class-conditional probability densities *differ* between categories.

3.1.3 Discriminant Functions

The Bayes decision rule is very simple and intuitive; however, in many practical pattern classification applications the prior probabilities and class-conditional probability densities of the different categories are *not* known, presenting a problem for the Bayes approach. The basic idea in these cases is to derive the decision rule from example patterns for which the true category membership is known. As the example patterns represent the prior probabilities and class-conditional probability densities of the different categories, one approach is to try to *estimate* the unknown probabilities $P\left(\omega_i\right)$ and probability densities $p\left(\vec{x}|\omega_i\right)$ from the example patterns and then use the estimated values as if they were the exact ones in the Bayes decision rule. In this approach, the estimation of the prior probabilities $P\left(\omega_i\right)$ does not present any serious difficulties; as they are pattern-independent constants that may be estimated sufficiently precisely by the fractions

$$P\left(\omega_i\right) = \frac{N_i}{\sum_j N_j}$$

of patterns belonging to category ω_i in the set of example patterns. The estimation of the class-conditional probability densities $p(\vec{x}|\omega_i)$, however, is difficult.

In pattern classification applications where the estimation of the class-conditional probability densities is impossible, the classification decision cannot be based on the Bayes decision rule. In alternative approaches, the classification problem is hence reformulated in terms of finding a set of *discriminant functions* $g(\vec{x})$, such that the i^{th} discriminant function is largest in the i^{th} decision region,

$$g_i(\vec{x}) > g_j(\vec{x}) \ \forall \ i \neq j \iff \vec{x} \in \mathcal{R}_i\,(Bayes)\,.$$

If such a set of functions can be found, the decision rule for classification may be based on assigning the feature vector $\vec{x}$ to the class $\tilde{\omega}_i$ with the largest discriminant:

$$\text{Decide } \tilde{\omega}_i \text{ if } g_i(\vec{x}) > g_j(\vec{x}) \ \forall \ j \neq i,$$

since then the resulting classification is the same as if the prior probabilities and class-conditional probability densities of all categories were known and the Bayes decision rule would have been applied, yielding the smallest possible error rate.

The framework of discriminant functions is general enough to solve arbitrary classification problems. In particular, the Bayes decision rule itself may be expressed equivalently by discriminant functions

$$g_i(\vec{x}) = P(\omega_i|\vec{x})\,, \tag{3.4}$$

if the posterior probabilities are chosen as discriminant functions.

3.1.4 Performance of Classifiers

The performance of classifiers is benchmarked by the probability for misclassification, the *classification error rate*, which may be decomposed into three different contributions. First and foremost, the Bayes or *indistinguishability error* represents the irreducible component of the classification error rate that results from the overlapping probability densities $p(\vec{x}|\omega_i)$ of different categories ω_i in feature space. For a fixed set of features, the indistinguishability error is an inherent property of the classification problem; it cannot be eliminated and is the same for all possible classification algorithms. The second contribution to the classification error rate is that due to *estimation errors*, which arise from statistical uncertainties if the decision regions are estimated from a finite set of example patterns. Accordingly, the estimation error may be reduced by increasing the number of example patterns. The third and last contribution may be due to *model errors* that arise if the algorithm used for classification is not flexible enough to accurately model the Bayes decision regions. The model errors may be reduced by using a more flexible classification algorithm; however, if the decision boundaries are estimated by example patterns, a more flexible classifier typically requires more example patterns for the determination of its decision regions, otherwise the estimation error increases.

An equivalent measure of the classification performance often used in practical pattern classification applications is the *separation power*

$$s \equiv \frac{\varepsilon_{sig}}{\varepsilon_{bgr}} = \frac{N_{sig,selected}/N_{sig}}{N_{bgr,selected}/N_{bgr}} \tag{3.5}$$

for a given *signal efficiency* $\varepsilon_{sig} = N_{sig,selected}/N_{sig}$. The separation power corresponds to the improvement in the signal-to-background ratio provided by the classifier; consequently, a higher separation power (for a given signal efficiency) means better performance.

3.2 Artificial Neural Networks

Inspired by the amazing performance of the human brain, artificial neural networks are an attempt at modelling the information processing capabilities of nervous systems. Although they are relatively simple devices in comparison to their biological models, artificial neural networks represent a powerful tool for pattern classification. Like their biological models, artificial neural networks are built of comparatively simple computing units, the *neurons*. The information processing capabilities of neural networks mostly arise from the intercommunication of the neurons, established by the connections between them. The structure of these connections is represented by the *topology* of the network, which may be quite complex. The structure of the network is only partially predetermined by the network topology, however; free parameters in the network may be adjusted so as to optimally solve the classification problem. The adjustment of these parameters is done by a *learning algorithm*, i.e. not through explicit programming, but through an automatic adaptive method.

In the following six subsections the principles of artificial neural networks is described in more detail. The biological model on which they are based is presented in the first subsection. In the second subsection, single artificial neurons are introduced. The interconnection of individual neurons to artificial neural networks is described in subsection three. In the context of pattern classification, a particular type of network is most relevant, that of *multi-layer* neural networks. These may be understood in terms of a *network function* that provides a mapping between the input and the output of the neural networks. The computation of the network functions proceeds via the so-called *"feed-forward" operation* of the network. As the output may be interpreted as discriminant functions, feed-forward operation represents the basis for pattern classification, as is described in more detail in the fourth subsection. In subsection five the training of the free parameters in the network by the *"error back-propagation"* learning algorithm will be described. The aim of the training is to modify the free parameters in such a way that the network function approximates the optimal discriminant functions corresponding to the Bayes decision rule in equation 3.4 [1]. The error back-propagation algorithm used for training and the feed-forward operation used for classification constitute the two primary modes of operation of multi-layered neural networks. Last but not least, the capability of artificial neural networks to solve arbitrary classification problems will be discussed and geometrically visualized in the sixth subsection.

[1] Under certain conditions conditions, the network output may even be interpreted as posterior probabilities [127]. This property of neural networks is not essential for pattern classification, however.

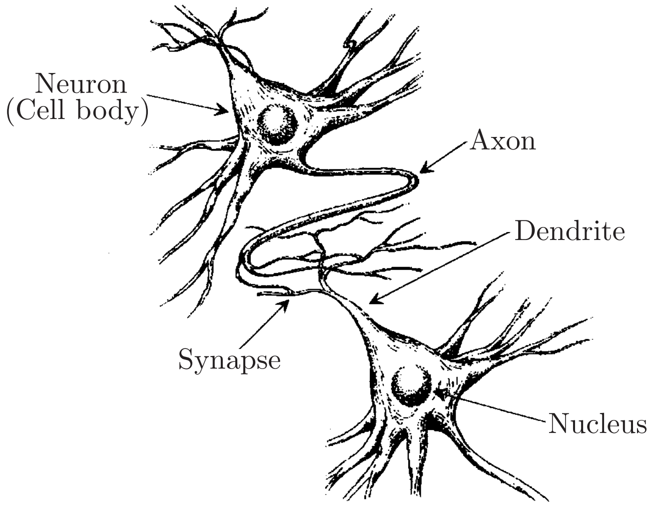


Figure 3.3: *Two biological neurons (taken from [128]).*

3.2.1 Biological Model

Artificial neural networks are simplified copies of the nervous system of humans and animals. The most complex of these, the human brain, contains about 10^9 neural cells or *neurons* in its outer layer, the *cortex*. For the processing of information, each of the 10^9 neurons is connected with up to 10^4 other neurons, forming a widely branched network of neural cells. In this network, information is processed by neurons receiving signals from other neurons and sending a response. Through this intercommunication, networks of neurons can develop complex behavior, as the emergence of consciousness in the human brain shows.

The complexity of neural networks is in contrast to the simplicity of individual neurons. As far as the processing of information is concerned, a biologocal neuron is composed of only four distinct functional units, a *dendrite*, a *soma*, an *axon*, and a *synapse*. An association of two such biological neurons is illustrated in figure 3.3. In the figure, signals are propagated from the upper left to the lower right; signals are sent by the *axon* of the sending neuron to the *dendrite* of the receiving neuron via the *synapse* that establishes the connection between the neurons. The information is transmitted by electrical signals that travel along the bio-membranes of the neural cells. When the so-called *activation potential* reaches the end of an axon, transmitter molecules are emitted from the *synaptic bouton*, a thickening at the end of the axon, into the *synaptic gap* between the axon of the sending neuron and the dendrite of the receiving one. The emitted *neurotransmitter* molecules then bind to receptors located at the surface of the dendrite, modifying the configuration of tunnel proteins so as to allow ions to flow into the dendrite. Depending on the type of neurotransmitter, the *post-synaptic potential* induced by the ion flow is either *exhibitory* or *inhibitory*, that is, it either increases or decreases the electric potential at the dendrite. The induced change in potential is subsequently propagated along the dendrite and eventually reaches the *soma*, the cell body of the receiving neuron. Depending on whether more exhibitory or inhibitory post-synaptic potentials reach the receiving neuron from different sending neurons and the *threshold* of the receiving neuron to "fire", the soma may then emit an action potential itself, transmitting the activation to the next neuron.

A biological neuron, whose activation never or only rarely reaches threshold gradually loses its ability to "fire". This mechanism provides a means for neural networks to adapt to changing environments, to "learn".

3.2.2 Artificial Neurons

The four structures, dendrite, soma, axon, and synapse of biological neurons are mimiced in the artificial copy. Artificial neurons are equipped with input and output channels for connections to other neurons and a body for the computation of an activation. Synapses are simulated by contact points between the input and output channels of different neurons; a *weight* is associated to these points in order to model the efficacy of the synaptic connections.

The transmission of information between artificial neurons proceeds in two stages: when calculating the output sent to the next neuron, an artificial neuron first computes its activation as the sum of the weighted input signals and then emits an output signal as a function of the activation. Writing the input signals as an input vector $\vec{x}$ and the weights assigned to the synaptic connections as a weight vector $\vec{w}$, the activity of a neuron may be expressed by the scalar product

$$\Sigma \equiv \vec{w} \cdot \vec{x}.$$

The output signal emitted by the neuron may be written in the form

$$f\left(\Sigma - \theta\right) = f\left(\vec{w} \cdot \vec{x} - \theta\right),$$

where θ denotes the *activation threshold* and f the *activation function* of the neuron. As the activation function f transfers the activity from the input to the output channels of the neuron, it is sometimes also termed the *transfer function* in the literature. In principle, *any* function may be used as the activation function; most often, either the hyperbolic tangent $\tanh\left(x\right) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ or the *Fermi function* $F\left(x\right) = \frac{1}{1+e^{-x}}$ are used, however [2].

In order to simplify the classification and training of neural networks, the activation thresholds θ of artificial neurons may be substituted by an extra weight, which is connected to an imaginary *threshold neuron* whose output is "clamped" to one. Formally, the input vector $\vec{x}$ of each neuron is augmented by the constant output of this threshold neuron and the weight vector $\vec{w}$ is augmented by the activation

[2] In the analysis presented in this thesis, the hyperbolic tangent is used as activation function for hidden and output neurons; for input neurons the identity $f\left(x\right) = x$ is used, to feed-in the features undistorted into the network. In comparison to the Fermi function, the hyperbolic tangent has the advantage of being antisymmetric, $\tanh\left(-x\right) = -\tanh\left(x\right)$, which allows for a symmetric treatment of exhibitory and inhibitory activations. Expressed in more technical language, the hyperbolic tangent $\tanh : \mathbb{R} \rightarrow [-1, 1]$ implements *bipolar logic*, while the Fermi function $F : \mathbb{R} \rightarrow [0, 1]$ implements *binary logic*. In neural network applications, bipolar logic is found to be more powerful [129]; as a result, the classification performance of neural networks using the Fermi function is generally slightly lower than that of corresponding networks using the hyperbolic tangent as activation function.

threshold θ that is to be substituted. The activation of the neurons may then be written in the form

$$f(\Sigma') = f(\vec{w}' \cdot \vec{x}'), \quad \text{with } \vec{x}' \equiv \begin{pmatrix} 1 \\ \vec{x} \end{pmatrix}, \ \vec{w}' \equiv \begin{pmatrix} \theta \\ \vec{w} \end{pmatrix}. \tag{3.6}$$

In this formalism, the activation thresholds of artificial neurons may be learned as weights, which allows for a common training algorithm to be used for both.

3.2.3 Multi-layer Neural Networks

Although artificial neurons are relatively simple devices of limited computational power, they may provide powerful tools when combined in a neural network. Artificial neural networks typically consist of between 10 and 100 neurons, a small number in comparison to the 10^9 neurons in the human brain. Decisive for the performance of artificial neural networks is not so much the number of neurons, but instead the organisation of the connections between them, the *topology* of the network. In the most simple types of networks, neurons are organized in *layers*, which means that all neurons in any one layer receive all their input from neurons in one preceding layer and likewise send all their output to neurons in one succeeding layer. Such simple *multi-layer neural networks* are found to perform surprisingly well in solving general pattern classification problems; hence, there is no need for more complicated topologies and the class of multi-layer neural networks is the most popular one in practical pattern classification applications.

In multi-layer neural networks, a distinction is mades between *input neurons*, *hidden neurons*, and *output* neurons. Input neurons are directly connected to the external input and constitute the *input layer*, which has the purpose of feeding the features of the patterns that are to be classified into the network. Hidden neurons are connected to other neurons only; they are "hidden" from the external environment. The hidden neurons are the ones that do the principal share of computing the classification decision. They may be organized in one or several *hidden layers*. Output neurons are located in the final layer of the network, where the classification decision is read-out. The output neurons constitute the *output layer*. For the substitution of the activation thresholds of neurons in the hidden and output layers, one threshold neuron is added to the preceding layer that is connected to all neurons in the layer of neurons whose activation thresholds are to be substituted. For input neurons, no activation thresholds are used, so no threshold neurons are necessary. An example of a multi-layer neural network is shown in figure 3.4. This shows a topology with an input layer, one hidden layer, and an output layer. The neurons in this example are *fully connected*; that is, all neurons in any one layer are connected to *all* neurons in the preceding and succeeding layers.

Classification

In multi-layer neural networks, classification proceeds by propagating the activation of neurons through the network layer by layer. For every pattern to be classified, the d-dimensional feature vector $\vec{x}$ is presented to the input layer. The input layer then distributes each component of the input vector to all neurons in the first hidden

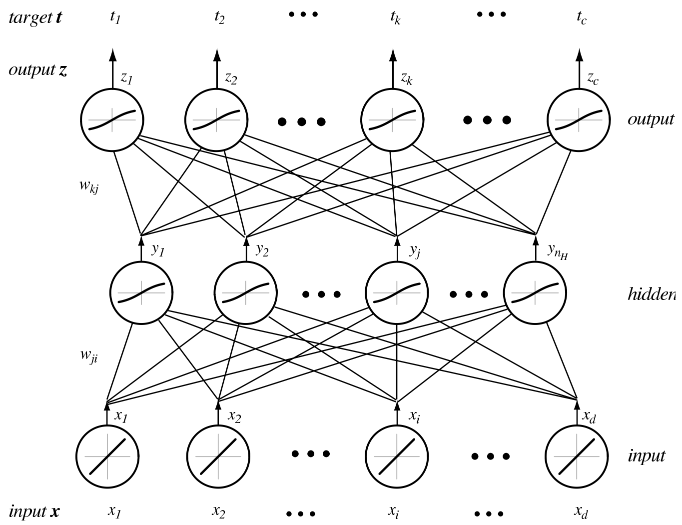


Figure 3.4: *A fully connected three-layer neural network with d input neurons, n_H neurons in the hidden layer and c output neurons. The neurons are symbolized as circles, their connections as lines. The activation functions of the neurons are indicated in the circles. Hyperbolic tangents are used for hidden and output neurons; for input neurons the identity $f(x) = x$ is used to feed-in the undistorted features (taken from [127]).*

layer. Each neuron in the first hidden layer then computes its *net activation* as the weighted sum of the inputs [3],

$$\Sigma_j^{(1)} = \sum_i w_{ij}^{(1)} x_i \equiv \vec{w}_j^{(1)} \cdot \vec{x},$$

where the weight vector $\vec{w}_j^{(1)}$ corresponds to the weights between input neurons i and neuron j in the first hidden layer. The activation of all neurons in the first hidden layer may be summarized in matrix notation as

$$\vec{\Sigma}^{(1)} \equiv W^{(1)} \cdot \vec{x},$$

where $W^{(1)}$ denotes the *weight matrix* between the input layer and the first hidden layer. The neurons in the first hidden layer then emit the output

$$y_j^{(1)} = f\left(\Sigma_j^{(1)}\right),$$

or in matrix notation

$$\vec{y}^{(1)} \equiv f\left(\vec{\Sigma}^{(1)}\right),$$

where the *same* activation function was assumed for all hidden neurons, but this is not obligatory. Each neuron in the second hidden layer similarly computes its activation based on the signals received from the first hidden layer,

$$\vec{\Sigma}_{(2)} = W^{(2)} \cdot \vec{y}^{(1)}$$

[3] In the following, the convention 3.6 is used to substitute the activation thresholds of neurons by an extra weight. In order to simplify the notation, the dashes denoting the augmented weight and input vectors are omitted in the following two section.

and emits the output

$$\vec{y}^{(2)} = f\left(\vec{\Sigma}^{(2)}\right).$$

In this way, the activation is sequentially propagated through all n hidden layers until it finally reaches the output layer where

$$\vec{\Sigma}^{(o)} = W^{(o)} \cdot \vec{y}^{(n)}$$

is computed and the output

$$\vec{z} = f\left(\vec{\Sigma}^{(o)}\right)$$

is emitted. This sequential layer-wise propagation of the activation from the input to the output layer is called the *"feed-forward"* operation of the network.

Training

When the network is set-up, the feed-forward operation provides a simple means of classifying patterns whose category is unknown. A remaining problem is, however, *how* to set-up the neural network for optimal classification. One of the simplest and most general methods widely used for the *training* of multi-layer neural networks is the *"error back-propagation"* algorithm. In the back-propagation algorithm, the weights of the connections between neurons are optimized in an interactive procedure so as to best classify a set of example patterns for which the true category is exactly known [4].

Starting with an untrained network of randomly initialized weights, the training patterns are presented to the input layer one by one. The activations of the input neurons are then propagated through the network, and the value of the discriminant is computed at the output layer. The network output is compared with the true category of the pattern, which represents the *target* for the classification:

$$t_i\left(\vec{x}\right) = \begin{cases} +1 & \text{if } x \,\epsilon\, \omega_i \\ -1 & \text{otherwise.} \end{cases}$$

Any difference between actual output and target represents an *error*. The weights of connections between the neurons in the network are then adjusted to reduce the error and better match the output to the target. As a measure of the error, usually the squared difference between the output and target vector

$$E = \frac{1}{2}|\vec{t} - \vec{z}|^2 = \frac{1}{2}\sum_{k=1}^{c}\left(t_k - z_k\right)^2 \tag{3.7}$$

is defined. The key idea of the back-propagation algorithm is, that the network output $\vec{z}$ is a *function* of the weights and hence function minimisation procedures

[4] As a rule of thumb, the number of patterns used for training should be at least one order of magnitude larger then the number of parameters

$$n_{weights} \sim n_{hidden}^2$$

to be trained [127].

can be applied to reduce the error $E \equiv E(w)$. By considering the neural network as representing a function

$$F : \mathbb{R}^d \to \mathbb{R}^c \subset [-1, +1]^c$$

mapping the input vector $\vec{x}$ to the output vector $\vec{z}$, the effect of each individual weight on the output error can be determined by ordinary partial derivatives.

Using the usual rules for function composition to express the propagation of the activation through the different layers, the function F mapping the input to the output vector may be written as

$$\begin{aligned} \vec{z} &= f\left(\vec{\Sigma}^{(o)}\right) = f\left(W^{(o)} \cdot \vec{y}^{(n)}\right) = f\left(W^{(n)} \cdot f\left(\vec{\Sigma}^{(n)}\right)\right) \\ &= \cdots = f\left(W^{(o)} \cdot f\left(W^{(n-1)} \cdot f\left(\ldots W^{(1)} \cdot \vec{x}\right)\right)\right). \end{aligned}$$

The function F is called the *network function*. In particular, if the activation functions f of the neurons are continuous and differentiable, the network function F is continuous and differentiable, allowing gradient descent procedures to be applied for the minimisation of the error function E. In the back-propagation learning rule, gradient descent is used in an iterative procedure, sequentially looping over the training patterns: for each training pattern presented to the network, the error made in the classification of that pattern is calculated according to equation 3.7; the weights are then modified by some fraction of the gradient of the error function

$$w_{ij}(n+1) = w(n) + \Delta w_{ij}(n+1), \quad \Delta w_{ij}(n+1) = -\eta \frac{\partial E}{\partial w_{ij}}, \tag{3.8}$$

such as to reduce the error. In the above *weight update rule*, w_{ij} describes the weight connecting neuron i with neuron j, and η is the *learning rate*, that defines the rapidity with which the weights change in the direction of minimal error.

The weight update process is started at the output layer, where the target activation is directly known. Each weight w_{ij} connecting neurons in the last hidden layer with those in the output layer is modified according to

$$\frac{\partial E}{\partial w_{ij}^{(o)}} = \frac{\partial E}{\partial z_j} \cdot \frac{\partial z_j}{\partial \Sigma_j^{(o)}} \cdot \frac{\partial \Sigma_j^{(o)}}{\partial w_{ij}^{(n)}} = (t_j - z_j) \cdot f'\left(\Sigma_j^{(o)}\right) \cdot y_i^{(n)},$$

where i indicates the neuron in the last hidden layer and j that in the output layer. The modifications for weights connecting other layers may then be expressed in terms of the error at the output layer by a recursive relationship. Applying the chain rule for differentiation, the modifications for weights connecting neurons i in any layer with neurons j in the succeeding hidden layer h are determined by

$$\frac{\partial E}{\partial w_{ij}^{(h)}} = \frac{\partial E}{\partial y_j^{(h)}} \cdot \frac{\partial y_j^{(h)}}{\partial \Sigma_j^{(h)}} \cdot \frac{\partial \Sigma_j^{(h)}}{\partial w_{ij}^{(h)}} = \frac{\partial E}{\partial y_j^{(h)}} \cdot f'\left(\Sigma_j^{(h)}\right) \cdot y_i^{(h-1)}.$$

Since the error function E does not depend explicitely on y_j, the chain rule has to be used again to evaluate the derivative

$$\frac{\partial E}{\partial y_j^{(h)}} = \sum_k \frac{\partial E}{\partial y_k^{(h+1)}} \cdot \frac{\partial y_k^{(h+1)}}{\partial \Sigma_k^{(h+1)}} \cdot \frac{\partial \Sigma_k^{(h+1)}}{\partial y_j^{(h)}} = \sum_k \frac{\partial E}{\partial y_k^{(h+1)}} \cdot \frac{\partial y_k^{(h+1)}}{\partial \Sigma_k^{(h+1)}} \cdot w_{jk}^{(h+1)},$$

where the sum extends over all neurons k in the layer succeeding the layer containing neuron j. By substituting this expression in the previous equation, the derivative of the error function with respect to the weight $w_{ij}^{(h)}$ may be written as

$$\frac{\partial E}{\partial w_{ij}^{(h)}} = \sum_k \frac{\partial E}{\partial y_k^{(h+1)}} \cdot \frac{\partial y_k^{(h+1)}}{\partial \Sigma_k^{(h+1)}} \cdot w_{jk}^{(h+1)} \cdot f'\left(\Sigma_j^{(h)}\right) \cdot y_i^{(h-1)}.$$

The recursive relation shows how the update of weights connecting neurons in the input and hidden layers (according to equation 3.8) depends on the weights and derivatives evaluated in the succeeding layer. It constitutes the basis of the back-propagation algorithm.

With the definition of the *sensitivity*

$$\delta_j \equiv \frac{\partial E}{\partial \Sigma_j} = \begin{cases} \frac{\partial E}{\partial z_j} \cdot \frac{\partial z_j}{\partial \Sigma_j} = (t_j - z_j) \cdot f'(\Sigma_j) & \text{, for neurons } j \text{ in the output layer} \\ \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{\partial y_k}{\partial \Sigma_k} \cdot w_{jk} \cdot f'(\Sigma_j) = f'(\Sigma_j) \cdot \sum_k w_{jk} delta_k & \text{, for neurons } j \text{ in any hidden layer } h \end{cases}$$

of the output error to the activation of individual neurons, the weight update rule may be written in a unified form,

$$w_{ij}(n+1) = w(n) - \eta \cdot y_i \cdot \delta_j, \tag{3.9}$$

for all weights in the network. The back-propagation algorithm defined by equation 3.9 is illustrated in figure 3.5. The graphic shows how, during training of a neural

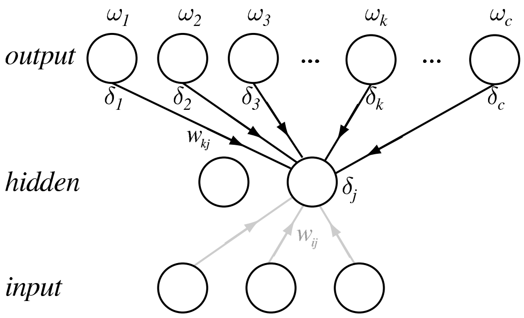


Figure 3.5: *The sensitivity of hidden neurons is proportional to the weighted sum of the sensitivities of the output neurons:* $\delta_j = f'(net_j) \sum_{k=1}^{c} w_{kj}\delta_k$. *The sensitivities of the output neurons are thus propagated "back" to the hidden neurons (taken from [127]).*

network with one hidden layer, the error is propagated from the output layer back to the hidden layer in order to perform the learning of weights connecting the input layer with the hidden layer.

By repeated application of the back-propagation learning rule, the training error decreases and finally converges to (at least a local) minimum of the error function. In practical neural network applications, the network may possibly learn the training patterns perfectly, but this does generally *not* mean that it will also classify novel patterns not used for training without error. To avoid *over-training*, an independent validation set is used to determine when to stop the training; the learning process is stopped, when the validation error reaches a minimum (see figure 3.6). Provided the error is satisfactorily low, learning may also be stopped after a fixed number of training cycles (before the minimal error is reached), in order to save computing time.

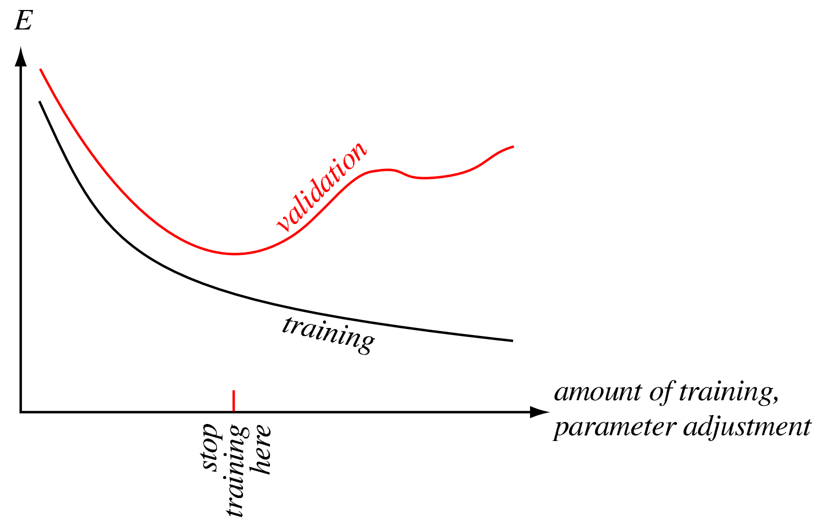


Figure 3.6: *Typical development of the classification error for patterns in the training and validation sets during training of neural networks. While the training error decreases monotonically, the validation error starts to increase after some time, indicating that the network gets tuned too much to the particular training sample; this is when training should be stopped (taken from [127]).*

3.2.4 Expressive Power and Geometric Visualisation

For optimal performance in arbitrary classification problems, neural networks have to approximate closely the decision regions defined by the Bayes decision rule. For this to be the case, on the one hand the free parameters of the networks need to be trained (as described in the last subsection). On the other, the topologies of the networks need to be flexible enough for the neural network to be trainable to closely approximate the optimal decision regions. The capability of neural networks to approximate arbitrary decision regions is termed the *expressive power* of the network. In this subsection it will be shown that the expressive power of neural networks crucially depends on the number of hidden layers in the network; in addition, the expressive power of neural networks with zero, one and two hidden layers will be visualized by geometric considerations of the approximizable decision regions in feature space.

Neural networks with no hidden layers are restricted to approximate hyperplane decision boundaries. As the input neurons merely distribute the components of the feature vector to the neurons in the succeeding layer, only the output neurons have an influence on the classification decision in neural networks without hidden layers. The classification decision of the network is determined by the output neuron with the largest output signal, that is, the maximum

$$f\left(net_i'\right) = f\left(net_i - \theta\right) = f\left(\vec{w}_i \cdot \vec{x} - \theta_i\right),$$

for all output neurons i. For feature vectors on decision boundaries, the output signals of (at least) two output neurons are equal,

$$f\left(net_i - \theta\right) = f\left(net_j - \theta\right) \iff \vec{w}_i \cdot \vec{x} - \theta_i = \vec{w}_j \cdot \vec{x} - \theta_j.$$

So,

$$\left(\vec{w}_i - \vec{w}_j\right) \cdot \vec{x} = \theta_i - \theta_j$$

for all feature vectors $\vec{x}$ on the decision boundary. This is the equation for points on a hyperplane normal to $\vec{n} = \vec{w_i} - \vec{w_j}$ and with distance $d = \theta_i - \theta_j$ to the origin, showing that the decision boundaries of neural networks without hidden layers are indeed hyperplanes in feature space. As a consequence, classification problems that are not *linearly separable* cannot be optimally solved with neural networks with no hidden layers.

The flexibility of neural networks to approximate non-linear decision boundaries is significantly increased by adding an extra hidden layer. In network architectures with one hidden layer, the hidden neurons may now approximate the hyperplane decision boundaries. By forming logical combinations of the output of hidden neurons, the hidden neurons may then approximate decision regions defined by intersections of hyperplanes [5]. As a result, neural networks with one hidden layer may approximate convex polytopes in feature space [130, 131].

By adding another hidden layer, the expressive power of neural networks may be again increased significantly. In network architectures with two hidden layers, the neurons in the second hidden layer may now approximate the convex polytopes. The output neurons may then approximate intersections and unions of the polytopes, by forming logical combination of the output of neurons in the second hidden layer. Since arbitrary shapes can be represented by unions and intersections of convex polytopes [130], neural networks with two hidden layers (and a sufficient number of neurons) are flexible enough to approximate decision regions of any form in feature space and hence optimally solve arbitrary classification problems [6].

As an example for the expressive power of neural networks, the decision regions that neural networks with zero, one and two hidden layers are able to approximate are illustrated in figure 3.7 for a two-category classification problem based on two different features x_1 and x_2.

---

[5] For example, a logical "AND" may be realized by choosing the activation θ_o of an output neuron slightly smaller then the sum of weights $\sum_j w_{ho,j}$ of its connections to neurons in the hidden layer.

[6] That neural network topologies with two hidden layers are indeed sufficient in general pattern classification applications is underpinned theoretically by a theorem due to the Russian mathematician Kolmogorov, which states that *any* continuous function $g(x)$ defined on the unit hypercube (this condition of Kolmogorov's theorem does not limit generality, as the feature vectors may always be mapped inside the unit hypercube by an appropriate scale transformation of the feature space) can be represented in the form

$$g(x) = \sum_{j=1}^{2d+1} \Phi_j \left(\sum_{i=1}^{d} \phi_i(x) \right)$$

for a suitably chosen set of functions Φ_j and ϕ_i [132]. In the context of neural networks, Kolmogorov's theorem may be interpreted as stating that *every* classification problem may in principle be optimally solved by neural networks with two hidden layers [127, 129]; the optimal solution may be realized by chosing the set of functions Φ_j and ϕ_i for the activation functions of neurons in the hidden layers, such that the resulting network function equals the discriminant functions corresponding to the Bayes decision regions in equation 3.4. Surprisingly, if the right set of activation functions is chosen for the neurons in the hidden layers, the theorem also shows that a moderate number of hidden neurons $n_{hidden} \sim d$ is sufficient for optimal performance in arbitrary classification problems.

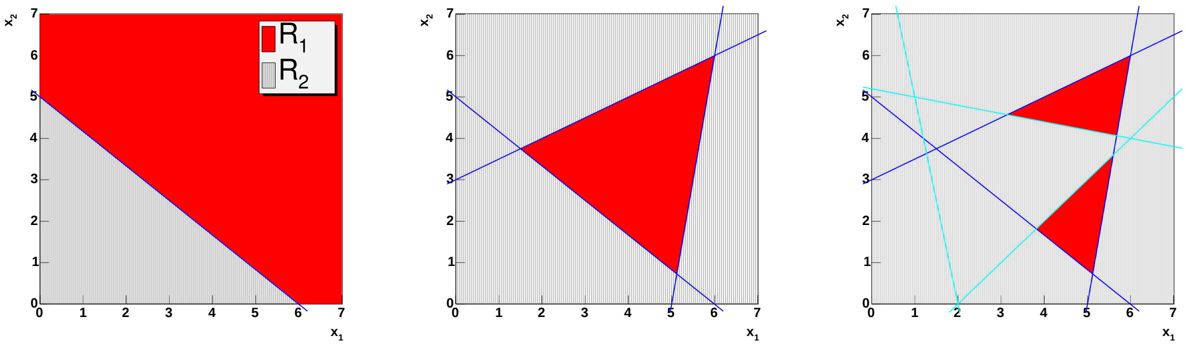


Figure 3.7: *Decision regions $\mathcal{R}_1$ (white) and $\mathcal{R}_2$ (grey) for a two-category classification problem realizable by neural networks with zero (left), one (centre) and two hidden layers (right).*

Chapter 4

Particle Reconstruction

In this chapter, the identification and reconstruction of individual particles in the H1 detector will be described. The identification of electrons, photons and muons is based on the detection of characteristic signatures, arising from their distinctive interactions with the detector material. In the H1 detector, tau leptons may be identified by their characteristic decay products only. All particles not identified as either electrons, photons, muons or tau leptons are considered to be hadrons.

In the first section of the following, the different kinds of interactions of particles with matter are described. As a result of these interactions, electrons, photons and hadrons may induce particle showers when traversing the detector. The development of these particle showers is described in the second section. In sections three, four and five, the identification and reconstruction of electrons, muons and tau leptons is described. The chapter concludes with a description of the reconstruction of hadrons.

4.1 Interactions of Particles with Matter

4.1.1 Ionisation

When traversing matter, energetic charged particles lose energy due to excitation and ionisation of the atoms in the traversed material. In a single collision process, a maximum energy of

$$\Delta E_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma \frac{m_e}{m} + \left(\frac{m_e}{m}\right)^2}$$

may be transferred to an atomic electron, depending on the mass m and momentum (via β and γ) of the traversing particle. The energetic particle suffers many interactions with atomic electrons, typically losing a few eV of energy in each collision. While *individual* collision processes can only be described statistically, the *average* energy loss arising from those collisions has been calculated within electrodynamics [133]. The result is the *Bethe-Bloch* formula [1]

$$-\frac{dE}{dx} = 2\xi\rho\left(\ln\left(\frac{\Delta E_{max}}{I}\right) - \beta^2 - \frac{\delta}{2}\right), \quad \xi = \frac{2\pi e^4 z^2 N_A}{m_e c^2 \beta^2}\left(\frac{Z}{A}\right),$$

which describes the average energy loss as function of the density ρ, ionisation potential I, charge Z and mass A (in atomic units) of the traversed material. The parameter δ is introduced to correct for saturation effects at high particle momenta; for highly energetic particles it is given by $\delta = 2\ln\gamma + \zeta$, with ζ being a material dependent constant.

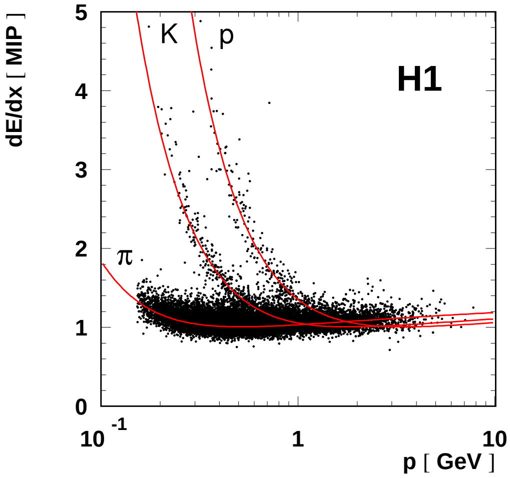


Figure 4.1: *Energy loss due to ionisation for different particle types. The curves describe the average energy losses expected for different particle types according to the Bethe-Bloch formula, while the points represent actually measured energy losses of charged particles traversing the central jet chamber of the H1 detector (taken from [134]).*

The energy loss due to ionisation is displayed in figure 4.1 for different particle types. At low momenta, the energy loss falls rapidly with the particle momentum as the time for interaction decreases. The energy loss becomes minimal around $\beta\gamma \approx 4$, giving rise to the term *minimum ionizing particle* (MIP) for particles with an energy loss near this minimum. For higher particle momenta, the energy loss rises logarithmically because of the relativistic increase of the electric field of the traversing particle in the transverse direction. At very high particle momenta, the transverse electric field of the traversing particle becomes screened by the polarisation of the atoms in the traversed material, thus limiting the logarithmic rise of the energy loss. As the polarisation of the traversed material depends on its density, the saturation of the energy loss is known as the *density effect*. While being negligible in gases, the density effect introduces an important correction to the energy loss in dense materials like lead or iron. The value of the energy loss in the limit of high particle momenta is termed the *Fermi-plateau*.

[1] The Bethe-Bloch formula describes the average energy loss due to excitation and ionisation for all traversing particles except electrons; for the latter, the Bethe-Bloch formula has to be adjusted to account for spin effects and modifications arising from the indistinguishability of the passing and atomic electrons [24].

Sizeable statistical fluctuations may arise around the average energy loss described by the Bethe-Bloch formula particularly in thin material layers, or where the maximum transferable energies are large. As a result of rare collision processes with a high energy transfer, the energy loss distribution is observed to be asymmetric, with a tail which extends to energy losses significantly above the Bethe-Bloch expectation. The relative size of the statistical fluctuations is characterized by a parameter that depends on the thickness of the traversed material Δx and the maximal transferable energy ΔE_{max},

$$\kappa = \frac{\xi \rho \Delta x}{\Delta E_{max}}.$$

Different mathematical models are suitable for the description of the energly loss fluctuations, depending on the value of κ. For $\kappa \geq 10$, the influence of single collision processes with a high energy transfer may be neglected and the energy loss distribution may be approximated by a symmetric *Gaussian* distribution. For $10 > \kappa \geq 10^{-3}$ the statistical fluctuations of the energy loss are best described by a *Vavilov* distribution [63]. For $\kappa < 10^{-3}$, the contribution of rare collision processes with high energy transfers become more dominant and the energy loss is best described by a *Landau* distribution.

4.1.2 Bremsstrahlung

In addition to the energy loss due to ionisation, fast charged particles loose energy due to *bremsstrahlung* processes. Bremsstrahlung photons are emitted whenever charged particles traversing material pass close to an atomic nucleus and are decelerated in its electromagnetic field. As the deceleration of the traversing charged particle decreases inversely proportional to its mass, the energy loss by Bremsstrahlung is, however, mainly relevant for electrons, for which it is the dominant energy loss mechanism at high energies.

For energetic electrons, the energy loss due to Bremsstrahlung may be approximated by [24]

$$-\frac{dE}{dx} = 4\alpha N_A \frac{Z^2}{A} \rho \left(\frac{1}{4\pi\epsilon_0} \frac{e^2}{m_e c^2} \right) \cdot E \cdot ln \left(\frac{183}{Z^{\frac{1}{3}}} \right).$$

As the energy loss is proportional to the energy of the traversing electron, the electron's energy decreases exponentially during its traversal of matter. The exponential decrease of the energy due to Bremsstrahlung may be more clearly seen if the material dependent parameters are combined to form the single quantity

$$X_0 \equiv \frac{A}{\frac{1}{4\pi\epsilon_0} \frac{e^2}{m_e c^2} 4\alpha N_A Z^2 \rho \ln \left(\frac{183}{Z^{\frac{1}{3}}} \right)}, \tag{4.1}$$

and the energy loss due to Bremsstrahlung is written in the form

$$-\frac{dE}{dx} = \frac{E}{X_0}.$$

The solution for the energy of the traversing electron is

$$E(x) = E_0 \cdot e^{-\frac{x}{X_0}}$$

as a function of the distance traversed x: $E_0 = E(x = 0)$ is the initial energy of the electron.

The quantity X_0 is a characteristic of the traversed material and describes the distance at which the energy of an incident electron has decreased to $1/e$ of its initial value due to the emission of bremsstrahlung photons. Accordingly, X_0 is termed the *radiation length*.

4.1.3 Pair Production

Energetic photons traversing matter may produce electron-positron pairs via the *photon conversion* process $\gamma \to e^+e^-$ in the electromagnetic field of an atomic nucleus. The probability for electron-positron pair production in a material layer of thickness x is approximately described by

$$P_{conv}(x) = 1 - e^{-\frac{x}{\mu}},$$

with the characteristic path length [24]

$$\mu = \frac{A}{\frac{1}{4\pi\epsilon_0}\frac{e}{m_e c^2} 4\alpha N_A Z^2 \rho \left(\frac{7}{9}\ln\left(\frac{183}{Z^{\frac{1}{3}}}\right) - \frac{1}{54}\right)}. \tag{4.2}$$

If the small term $\frac{1}{54}$ in the bracket of the denominator is neglected, the characteristic path length μ for electron-positron pair production by highly energetic photons may be expressed in terms of the radiation length X_0. More precisely, as follows by comparison of equations 4.2 and 4.1, the characteristic path length for photon-conversion is about 30% larger than a radiation length,

$$\mu = \frac{9}{7} X_0.$$

4.1.4 Hadronic Interactions

In addition to their electromagnetic interactions, hadrons traversing matter may interact strongly with the atoms in the traversed material. Typically, hadronic interactions are inelastic processes in which further strongly interacting particles are produced.

In analogy to the radiation length X_0 for electromagnetic interactions, the *hadronic interaction length*

$$\lambda = \frac{A}{N_A \rho \sigma}$$

is defined as the characteristic length scale for the interaction of hadrons in matter. The probability for an incident hadron to interact strongly in a material layer of thickness x is given by

$$P(x) = 1 - e^{-\frac{x}{\lambda}}.$$

In contrast to the radiation length X_0, λ cannot be determined from first principles, as the hadronic cross-sections cannot be calculated. Experimentally measured values must be used in calculating λ and these are usually considered to be independent of the energy of the incident hadron and of the hadron species, so that the hadronic interaction length depends only on the properties of the traversed material. For materials with $Z \geq 6$, the hadronic interaction lengths λ are much larger then the radiation lengths X_0. In particular, for lead, iron and copper, which are used as absorber materials in the H1 calorimeters, the mean free path length of hadrons exceeds that of electrons and photons by about one order of magnitude [24].

4.2 Development of Particle Showers

4.2.1 Electromagnetic showers

Through alternating bremsstrahlung and pair production processes, energetic electrons and photons traversing matter may produce an exponentially increasing number of further electrons and photons, an *electromagnetic shower*. Particle showers induced by electrons and photons become very similar after a short distance, as is shown by the comparison of the cascades

$$e \to \gamma e \to eee\gamma \to e\gamma e\gamma e\gamma ee \quad \text{and} \quad \gamma \to ee \to e\gamma e\gamma \to e\gamma eee\gamma ee.$$

The essential characteristics of the shower may be illustrated by a simple model: A photon of energy E_0 enters a material and converts to an e^+e^--pair after travelling a distance X_0. The produced electron and positron emit bremstrahlung photons after traversing another radiation length. In this way, the number of particles doubles within each radiation length, producing an exponentially increasing particle shower, as is illustrated in figure 4.2. After a distance of t radiation lengths, the number of

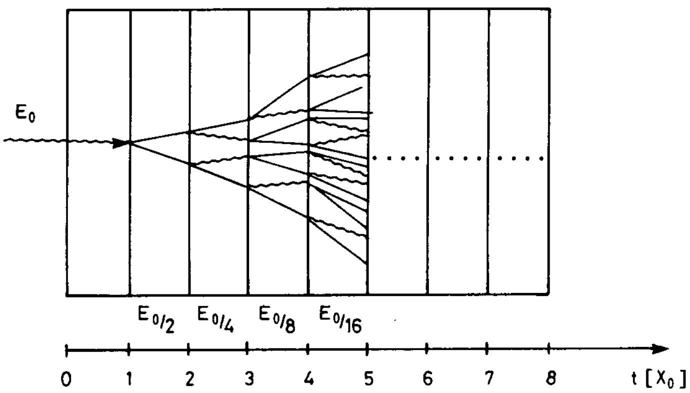


Figure 4.2: *Illustration of a simple model for the development of particle showers (taken from [24]).*

shower particles is

$$N(t) = 2^t.$$

If the energy is assumed to be distributed evenly, the energy of the shower particles at a distance t is then

$$E(t) = \frac{E_0}{N(t)} = E_0 \cdot 2^{-t}.$$

The particle multiplication continues as long as the energy of the shower particles is sufficient for further bremsstrahlung and pair production processes. Once the energy of the electrons and photons in the shower drops below a *critical energy* E_c, energy losses due to ionisation and Compton scattering become dominant and the shower particles are decelerated and finally absorbed by the traversed material. The shower dies out.

The critical energy E_c is determined by the energy dependence of the bremsstrahlung and pair production processes on the one hand and that of the competing ionisation and Compton scattering energy loss mechanisms on the other. The energy loss of electrons and photons by means of ionisation and Compton scattering becomes dominant below a material dependent threshold of the order of $10 - 100\ MeV$. For materials with atomic charge number $Z \geq 13$, the critical energy may be approximated by [20]

$$E_c \approx \frac{500\ MeV}{Z}.$$

The above simple shower model is sufficient to describe the longitudinal development of particle showers. The maximum number of shower particles is reached after a distance of about

$$t_{max} \sim \ln\left(\frac{E_0}{E_c}\right)$$

and is

$$N\left(t_{max}\right) \approx \frac{E_0}{E_c}.$$

Experimentally, around 98% of the shower energy is contained within a depth of $l = 2.5\ t_{max}$ on average [24].

The development of the shower in the transverse direction is not described by the simple shower model. In nature, the particle shower broadens with increasing depth due to multiple scattering. The extent of the electromagnetic shower in the transverse direction is characterized by the *Molière radius*

$$R_m = \frac{21\ MeV}{E_c} \cdot X_0.$$

In spite of the broadening due to multiple scattering, the majority of the energy is contained within a compact core. Experimentally, on average about 95% of the shower energy is found within a cylinder of radius $r = 2\ R_m$ [24].

The statistical nature of the bremsstrahlung, pair production and scattering processes implies fluctuations in the shower development, that have been neglected in the simple shower model. In fact, these fluctuations are insignificant for electromagnetic showers.

4.2.2 Hadronic Showers

The main difference between hadronic and electromagnetic showers is that in hadronic showers the particles are produced by hadronic interactions instead of bremsstrahlung and pair production processes. Correspondingly, hadronic showers develop over length scales increased by the ratio of the hadronic interaction length λ to the radiation lenth X_0. In the transverse direction, the shower particles are more widespread in hadronic than in electromagnetic showers, because of the high transverse momenta that may be transferred in nuclear interactions.

Inelastic nuclear interactions result in the production of large numbers of charged and neutral pions, as well as smaller numbers of kaons, nucleons and other hadrons. On average, neutral pions amount to about one third of the total number of pions produced in inelastic hadronic interactions. They decay rapidly, in about 10^{-16} s, into two energetic photons, which induce electromagnetic subshowers in the hadronic showers. The fraction of neutral pions produced in nuclear interactions generally increases with the energy of the interacting hadron; the fraction is subject to sizeable statistical fluctuations in individual hadronic showers, however.

Because of this, and because the development of hadronic showers depends decisively on the nature of the first inelastic interaction of the incident hadron [24], the properties of hadronic showers fluctuate significantly between individual showers.

4.3 Electron Identification

The identification of electrons is based on their characteristic shower-shape in calorimeter detectors. In the H1 experiment, electrons are detected in the LAr and SpaCal calorimeters, and in dedicated small angle detectors ("electron taggers") installed close to the beam-pipe in upstream direction. As the signatures of electromagnetic showers differ significantly in these detectors, separate algorithms ("electron-finders") have been developed for electron identification and reconstruction in each calorimeter. The principal LAr [135] and SpaCal [136] electron finders are optimized to efficiently identify electrons that are of a comparatively high energy. For the identification of low energy electrons, produced for instance in the decays of heavy quark vector mesons, a complementary finder algorithm has been developed [137]. In the analysis presented in this thesis, the identification of "isolated electrons" is performed using the high energy LAr electron finder, since it provides the most pure electron sample in most of the solid angle. Electron candidates found in the SpaCal and the small angle detectors are considered to be the scattered beam electrons and analysed solely to improve the event reconstruction (*cf.* section 5). Hence, only the algorithm for the identification of high energy electrons in the LAr calorimeter will be discussed in more detail in the following.

The LAr electron finder identifies electron candidates as compact energy clusters in the electromagnetic part of the LAr calorimeter. Owing to the fine granularity of the calorimeter, detailed information describing the shape and size of the shower profile is available to discriminate between electromagnetic and hadronic showers. The shower-shape analysis is performed in a cone of 7.5 degrees opening angle around an axis defined by the position of the shower centre relative to the reconstructed primary vertex and having its apex 100 cm before the inner edge of the LAr surface (see figure 4.3). This cone is termed the *electron shower envelope*; its size has been optimized to contain electromagnetic showers of all energies [25]. Surrounding the electron shower envelope is an *isolation cone* of opening angle 0.25 (defined in the η-ϕ metric), having its apex at the reconstructed primary vertex. Energy leaking out of the electron shower envelope in the transverse direction is detected in the

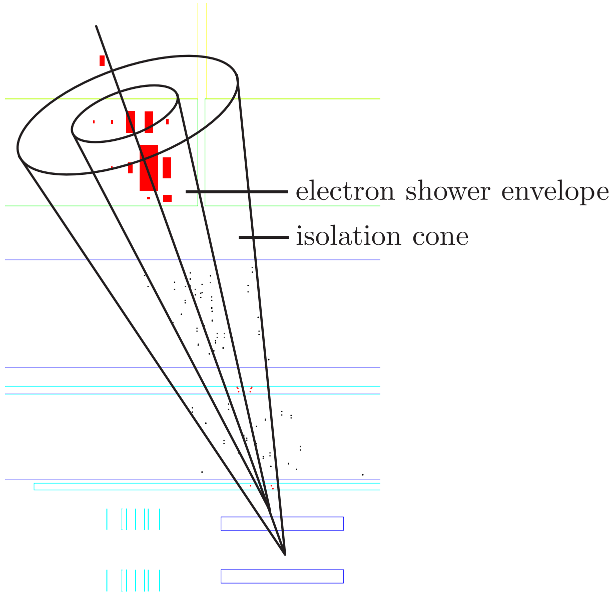


Figure 4.3: *Illustration of the electron shower envelope and the isolation cone around an electron cluster in the H1 detector.*

isolation cone. Electromagnetic showers are generally contained within the electron shower envelope, while hadronic showers are not; a sizeable fraction of their energy may be detected in the isolation cone. The different longitudinal development of electromagnetic and hadronic showers in the electron shower envelope allows for further separation of electrons and photons from hadrons: the fraction of enery contained in the electromagnetic part of the LAr calorimeter is generally larger for electromagnetic than for hadronic showers. In addition to the smaller transverse and longitudinal extent, a compact *"hot core"* is expected for electromagnetic showers. The fraction of energy contained in the "hot core" is estimated by summing the energy in a group of N contiguous cells in the first two (three in the IF wheel) electromagnetic layers of the LAr calorimeter including the "hottest" (most energetic) cell. The number of cells summed is adapted to the varying cells sizes in the different regions of the LAr calorimeter ($N = 4$ in the BBE, CB1 and CB2 wheels; $N = 8$ in CB3, FB1 and FB2; $N = 12$ in the IF wheel), ensuring the core contains, on average, a similar energy fraction for electromagnetic showers. A further estimator of the compactness of the shower is the cluster radius, calculated as

$$R_{cluster} = \frac{\sum_i \rho_i r_i^2 - \left(\sum_i \rho_i r_i\right)^2}{\sum_i \rho_i},$$

summed over all cells i in the electron shower envelope. The distances r_i of the cells from the cone axis are weighted with the energy density $\rho_i = E_i/V_i$ in the cells, to increase the weight of (smaller) cells in the first electromagnetic layer. On the basis of these variables, electron candidates are selected using the criteria listed in table 4.1. The thresholds of the cuts on some estimators are adapted to the cell granularity of the LAr calorimeter, that varies between individual wheels; the dependence on the polar angle is shown in figure 4.4.

Property	Requirement
E_e	$> 5\ GeV$
P_T^e	$> 3\ GeV$
N_{cell}	≥ 4
E_{em}/E_e	$> f_{em}(\theta_e)$
E_{core}/E_e	$> f_{core}(\theta_e)$
$R_{cluster}$	$< f_{radius}(\theta_e)$
$E_e/(E_e + E_{iso})$	< 0.98
E_{iso}^{had}	$< 300\ MeV \Vert E_e/(E_e + E_{iso}) < 0.95$

Table 4.1: *Selection of electron candidates in the LAr calorimeter. E_e and P_T^e denote the total and transverse energy reconstructed in the electron shower envelope. The estimators E_{em} and E_{core} describe the energy detected in the electromagnetic section of the LAr and in the "hot core" of the electron cluster, respectively. The cluster radius $R_{cluster}$ characterizes the spread of the shower energy in the transverse direction. The energy leaking out of the electron shower envelope is estimated by the variables E_{iso} and E_{iso}^{had}, which describe the energy reconstructed in the isolation cone and the hadronic section of the isolation cone, respectively. All energies are calculated for the electron hypothesis, that is, using the electromagnetic energy scale, which is corrected for energy losses in the dead material between the interaction point and the LAr calorimeter (see section 1.2.2).*

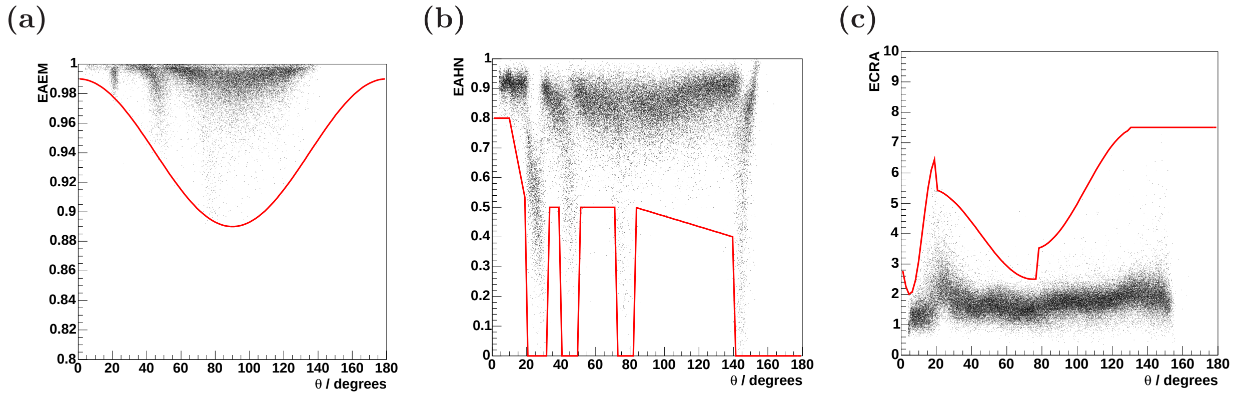


Figure 4.4: *Dependence of the cut thresholds (curves) on (a) the electromagnetic energy fraction E_{em}/E_e, (b) the "hot-core" fraction E_{core}/E_e, and (c) the cluster radius $R_{cluster}$ on the polar angle of the electron candidates, shown with the values of these estimator variables reconstructed for simulated electrons (dots).*

The identification efficiencies for electrons and misidentification probabilities for hadrons (determined here for charged pions) resulting from this candidate selection are displayed in figures 4.5 and 4.6, respectively. As can be seen in the figures, the electron identification efficiency is above 95% for electrons of energies above $p^{track} > 6\ GeV$ in the central region ($35° \leq \theta^{track} \leq 130°$). In the forward region, the electron identification efficiency is seen to decrease to about 70% at small polar angles. The probability for the misidentification of hadrons as electrons decreases for hadrons of higher energies, as the particle showers they induce become more extended and thus easier to distinguish from compact electromagnetic showers. As a result of the varying cell granularity of the LAr and the polar angle dependent cut thresholds shown in figure 4.4, the probability for hadrons to be misidentified as electrons strongly depends on the polar angle; it is smaller in the forward region, in which direction the highest number of hadrons is produced.

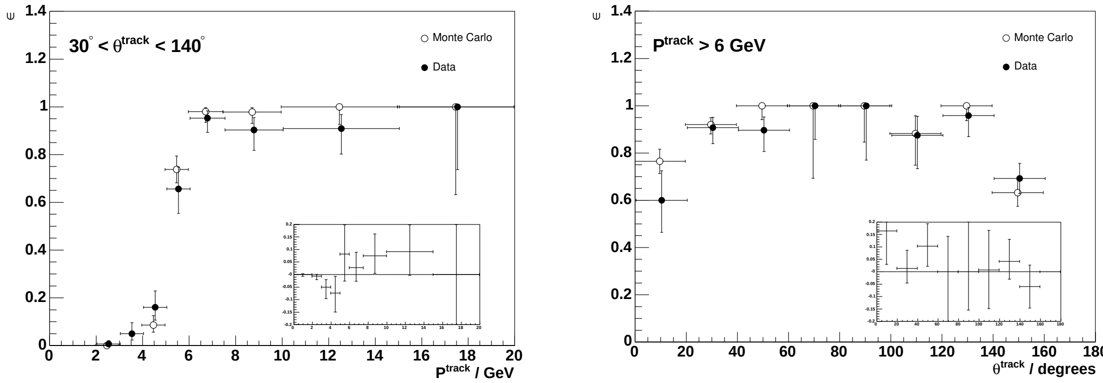


Figure 4.5: *Electron identification efficiency resulting from the candidate selection summarized in table 4.1. The identification efficiencies are estimated separately for the data (solid circles) and for the Monte Carlo simulation (open circles) using samples of elastic di-electron events (see section C.1 of the appendix). The difference between data and Monte Carlo is shown in magnified view in the insets.*

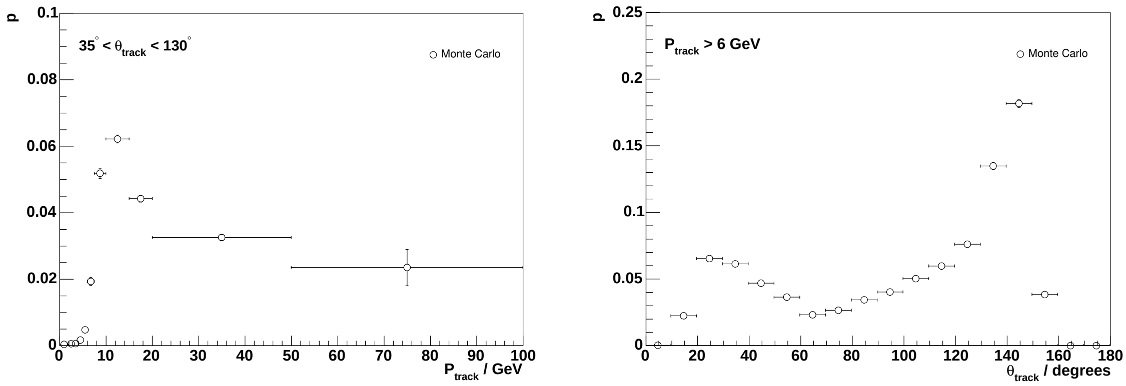


Figure 4.6: *Misidentification probabilities for hadrons, resulting from the electron selection summarized in table 4.1. The probabilities for hadrons to be misidentified as electrons are estimated using a sample of simulated single charged pion events.*

To separate the electromagnetic showers induced by electrons from those induced by photons, the energy clusters in the LAr are required to be linked either to a track in the central or forward drift chambers CTD and FTD or to a hit in the central inner and outer proportional chambers CIP and COP. In particular, the identification of electrons produced in the forward direction is improved by the CIP, as it allows for charged particles to be tagged before they traverse the dead material between the CTD and FTD, which is problematic for the electron/photon separation, because of the high probability for photons to convert into electron-positron pairs.

The electron four-vector is determined from the combined cluster and track information. Optimal precision is attained by reconstructing the energy and polar angle from the calorimeter energy and cluster position relative to the event vertex and the azimuthal angle from the slope of the linked drift chamber track at the vertex. The reconstructed energy is calibrated using the method described in references [31, 138]. Insensitive regions of the LAr calorimeter are excluded from the analysis; electron candidates are removed in the region within $\pm 2°$ around the ϕ-cracks between different octants and within 15..30 cm of the z-crack separating the CB2 and CB3 wheels.

The isolation of electrons is estimated from the energy deposits not attributed to the electron contained in a cone of opening angle 0.5 (in the η-ϕ metric). Electron candidates are flagged as "isolated leptons" if the energy deposits in this cone amount to less than 5% of the electron energy.

4.4 Muon Identification

The identification of muons is based on their characteristically low energy loss when traversing matter. Unlike electrons, muons of energies in the kinematic range of HERA do not lose a significant fraction of their energy due to bremsstrahlung, because of their higher mass. At HERA, only the energy loss due to the ionisation of the traversed material is relevant for muons. As this energy loss increases logarithmically with the particle energy, it deviates only little from its minimum value (see section 4.1.1), so muons of energies accessible at HERA may be characterized as minimum ionizing particles. In particular, unlike electrons, photons and hadrons, muons do not induce particle showers when traversing the detector material. As the energy loss due to ionisation is approximately constant, the energy deposits of muons are evenly distributed in the calorimeter and contained in a narrow cylinder around the muon track. In iron, the energy loss of muons is of the order of 10 $MeV cm^{-1}$. As a consequence, muons with energies above a few GeV penetrate the calorimeter and the surrounding muon systems and then leave the detector.

In the H1 detector, muons are identified by their characteristic signatures in the LAr calorimeter, the instrumented iron return yoke surrounding the LAr calorimeter and the forward muon detector. The momenta of identified muons are measured in the forward muon detector as well, and also in the central and forward tracking detectors.

In the analysis presented in this thesis, the selection of muon candidates is based on tracks reconstructed in either the instrumented iron or the forward muon detector ("iron" muons). In order to compensate for inefficiencies of the track reconstruction and improve the reconstruction efficiency for muons in the instrumented iron, especially in the transition regions between the barrel and the endcaps [97], muons in the instrumented iron may also be identified by a combination of the pattern of energy deposits in the tail-catcher and in the liquid argon calorimeter. As the energetic muons searched for in this analysis are expected to penetrate the calorimeter, muons identified solely in the LAr calorimeter ("calorimeter" muons) [137] are not considered. The tracks and tail-catcher clusters in the instrumented iron are required to match a track in either the central or forward tracking detectors within a distance of 0.5 (in the η-ϕ metric) [1]. As the forward muon detector provides a

[1] No *track-link* between the tracks in the instrumented iron and the drift chambers is required

precise measurement of the muon momentum on its own, no matching track in the forward tracking detector is required for tracks in the FMD.

The reliability of the muon signature in the instrumented iron, the forward muon detector and the LAr calorimeter are estimated by various variables calculated by the muon idenification algorithm ("muon-finder"). The quality of tracks reconstructed in the instrumented iron is assessed by the number of associated hits in the limited streamer tubes and the depth of those hits in the iron (see reference [141] for details of the track reconstruction). Only streamer tube layers that are enclosed by iron plates are taken into account, the philosophy being that layers in the inner and outer muon-boxes do not improve the muon signature, as no material needs to be traversed. The quality of tracks reconstructed in the forward muon detector is determined by the number of hits in the drift chambers on either side of the toroid included in the track fit. Tracks containing track segments in front of and behind the toroid are assigned track qualities 1 and 2, respectively. The reliability of the muon signature in the liquid argon calorimeter is estimated from the sum of energy deposits contained in a cylinder around the extrapolated muon track of radius 25 cm in the electromagnetic and 50 cm in the hadronic section of the LAr and the depth of the energy deposits in the calorimeter [2] On the basis of these estimators, muon candidates are selected using the criteria listed in table 4.2. For muons produced in the forward region, more stringent selection criteria are applied in the instrumented iron, to avoid the misidentification of hadrons penetrating the LAr calorimeter in this region, in which the number of hadrons is highest.

The identification efficiencies for muons and misidentification probabilities for hadrons (here determined for charged pions) resulting from the muon selection listed in table 4.2, are displayed in figures 4.7 and 4.8, respectively. As can be seen in the figures, the muon identification efficiency is about 90% for muons of momenta above about 3 GeV; muons of lower momenta are less likely to be identified in the iron, as they are more likely to stop in the LAr calorimeter. The probability for hadrons to be misidentified as muons is between 1 and 2%. The misidentification is mainly due to hadrons traversing the LAr calorimeter and reaching the muon system without undergoing a strong interaction. The probability for hadrons to penetrate the calorimeter without interacting strongly (termed *"sail-through"* [142]) is (*cf.* section 4.1.4)

$$P(\Delta x) = e^{-\frac{\Delta x}{\lambda}},$$

where Δx is the depth of the LAr calorimeter which amounts between 4.5 and 8 hadronic interaction lengths λ, depending on the polar angle.

For the reconstruction of the muon four-vector, preference is given to tracks in the forward muon detector as it provides the best resolution for energetic muons.

---

in the muon identification, as it would reduce the identification efficiency by about 10%. Moreover, the probability for linking extrapolated drift chamber tracks with tracks reconstructed in the instrumented iron is not precisely modelled in the Monte Carlo simulation [139], due to the necessary simplifications in the track extrapolation routines that are used for simulation and reconstruction [140]. For isolated muons, the track-link is not absolutely necessary, as there is no ambiguity in matching the tracks reconstructed in the instrumented iron with those in the drift chambers.

[2] All energies for the muon hypothesis are calculated using the electromagnetic energy scale. No corrections for energy loss in dead material are applied (as it is small for minimum ionizing particles).

Iron Tracks			
Forward endcap		**Backward endcap**	
ρ_0	$< 100\ cm$	ρ_0	$< 100\ cm$
z_0	$< 100\ cm$	z_0	$< 100\ cm$
no. layers	≥ 6	no. layers	≥ 3
first layer	≤ 5	first layer	≤ 8
last layer	≥ 6	last layer	≥ 3
matching Drift Chamber track		matching Drift Chamber track	
Barrel		**Forward muon toroid**	
ρ_0	$< 100\ cm$	track quality	$= 1\|\|2$
z_0	$< 100\ cm$	z_0	$> -400\ cm$
no. layers	≥ 2	z_0	$< 300\ cm$
first layer	≤ 5		
last layer	≥ 2		
matching Drift Chamber track			

Tail-Catcher Cluster and MIP Pattern in LAr	
no. Tail Catcher clusters	≥ 1
no. LAr clusters	≥ 1
E_{cone}^{LAr}	$\leq 5\ GeV$
$< l >$	$\geq 40\ cm$
matching Drift Chamber track	

Table 4.2: *Selection criteria for "iron" muons. The muons may be identified either as tracks reconstructed in the instrumented iron or the forward muon detector (above) or by the combination of energy deposits in the instrumented iron (tail-catcher) and the LAr calorimeter (below). For muons whose momenta are not measured in the FMD, a track in either the central or forward tracking detectors is required, matching the muon signal in the iron within a distance of* 0.5 *in* $\eta - \phi$. *The variables* ρ_0 *and* z_0 *describe the distance of the extrapolated iron tracks to the nominal interaction point and are used to reject cosmic muons and muons produced outside the ep interaction region in interactions of the proton beam with residual gas atoms in the beam-pipe or with the beam-pipe wall.*

In the case that muon candidates are associated with a track in the forward muon detector, their momentum is determined by the curvature, and their polar angle by the first hit, of the track. If this muon candidate is also associated with a track in the forward tracking detector, the azimuthal angle is determined by the slope of this track at the vertex; otherwise the azimuthal angle is determined using the first hit of the track in the forward muon detector. If muon candidates are not associated to a track in the FMD, their four-vector is determined by the momentum, polar and azimuthal angles of the associated track in the central or forward tracking detectors.

The isolation of muons is estimated from the energy in the LAr calorimeter contained in a cylinder around the extrapolated muon track, the radius of the cylinder being 35 cm in the electromagnetic and 75 cm in the hadronic section. Muon candidates are flagged as "isolated leptons" if the energy in the cylinder is less than 5 GeV (on the electromagnetic energy scale) and the muon is separated from other tracks in the central and forward tracking detectors by a distance of at least 0.5 (in the η-ϕ metric).

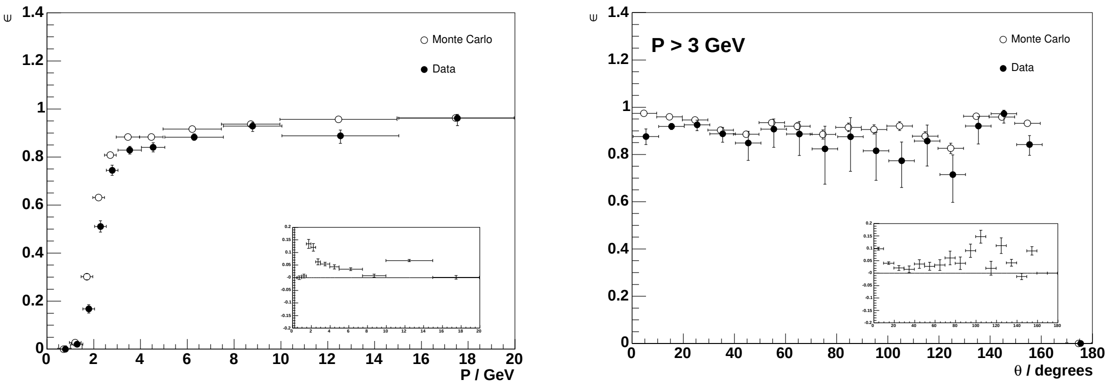


Figure 4.7: *Muon identification efficiency resulting from the candidate selection summarized in table 4.2. The identification efficiencies are estimated separately for the data (solid circles) and for the Monte Carlo simulation (open circles) using a sample of elastic di-muon events (see section C.2 of the appendix). The difference between data and Monte Carlo is shown in magnified view in the insets.*

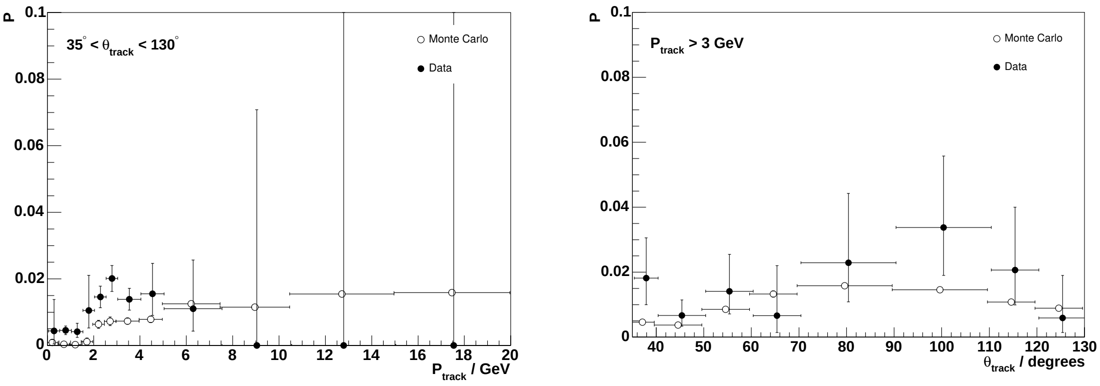


Figure 4.8: *Misidentification probabilities for hadrons resulting from the muon selection summarized in table 4.2. The probabilities for hadrons to be misidentified are estimated separately for the data (solid circles) and for the Monte Carlo simulation (open circles). In the data, the estimated misidentification probabilities are determined using a sample of identified K^0 mesons decaying to two charged pions (as described in reference [142]); in the Monte Carlo simulation, the misidentification probabilities are determined using a sample of single charged pion events.*

4.5 Tau Identification

The identification of tau leptons is substantially different from the identification of electrons and muons. With a mass of $m_\tau \approx 1.78\ GeV$ and a lifetime of $\tau \approx 290\ fs$ [82], the tau decays inside the beam-pipe within $c\tau \approx 87\ \mu m$. Therefore, only the tau decay products are observed in the detector.

The most frequent decay channels of tau leptons are listed in table 4.3, along with their branching fraction. About 35% of taus decay *leptonically* into electrons and muons, the remaining 65% decay *hadronically*. Unlike the lighter electrons and muons, which have masses below the pion mass, tau leptons are heavy enough for the

Decay mode	Branching fraction
leptonic Decay modes	
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$	$\approx 17\%$
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$	$\approx 18\%$
hadronic Decay modes ("one-prong")	
$\tau^- \to \pi^- \nu_\tau$	$\approx 11\%$
$\tau^- \to \rho^- \nu_\tau \to \pi^- \pi^0 \nu_\tau$	$\approx 25\%$
$\tau^- \to a_1^- \nu_\tau \to \pi^- \pi^0 \pi^0 \nu_\tau$	$\approx 9\%$
other	$\approx 5\%$
hadronic Decay modes ("three-prong")	
$\tau^- \to a_1^- \nu_\tau \to \pi^- \pi^- \pi^+ \nu_\tau$	$\approx 10\%$
$\tau^- \to \pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	$\approx 4\%$
other	$\approx 1\%$

Table 4.3: *The main decay modes of tau leptons and their branching fractions. The decay modes are classified into leptonic and hadronic decays; the hadronic decays are further subclassified into decays into one charged hadron ("one-prong") and three charged hadrons ("three-prong").*

decay into hadrons to be kinematically possible. In nearly all hadronic tau decays either one or three charged particles, mainly pions, are produced, together with a few neutral pions [1]. As listed in table 4.3, some of the hadronic tau decay modes proceed via intermediate vector meson resonances.

The identification of tau leptons is based on the identification of their decay products in the detector. Due to the variety of decay modes, there is no unique tau signature, but several different signatures that must be combined for maximal reconstruction efficiency.

The most difficult decay modes of tau leptons to reconstruct are the decays into electrons and muons. The only means of distinguishing between leptonic tau decays and electrons and muons produced in the primary *ep* interaction is through observables that are sensitive to the displaced decay vertex and the neutrinos produced in the decay. As is illustrated in figure 4.9, the displaced event vertex may give rise to a measureable *impact parameter* of the electron or muon track with respect to the event vertex. Unfortunately, the lifetime of the tau lepton is too short for significant impact parameters to be observed in the H1 detector, given the resolution of the tracking detectors. [2]. The possibility of identifying leptonic tau decays by the missing momentum arising from the neutrinos produced in the tau decay is analysis dependent. In general, missing momentum reconstruction is complicated by the presence of more than one neutrino in events with tau leptons [3]. In the analysis

[1] The charged particles produced in the decay of tau leptons are conventionally referred to as "prongs".

[2] The feasibility of using lifetime information to identify tau decays has been studied by the *CDF* collaboration [143]. The impact parameter resolution of the CDF detector is about three times better than that of H1 [144], giving a separation power (defined by equation 3.5) of about 2-4 between particles coming from tau decays and those produced at the event vertex.

[3] This is strictly true in the Standard Model and remains so in processes beyond the Standard Model, as long as lepton flavour is conserved, as each ν_τ then has to be associated with a $\bar{\nu}_\tau$ in the

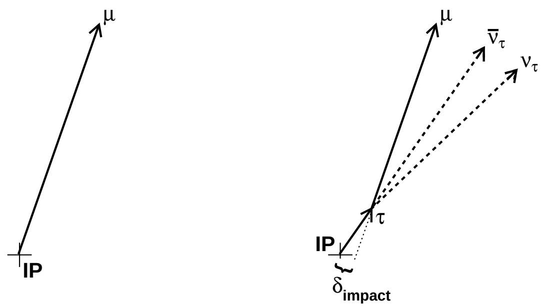


Figure 4.9: *Production of muons in primary ep interactions (left) and through the decays of tau leptons (right). In leptonic tau decays, a measureable impact parameter of the muon track with respect to the event vertex may be observed or may be missing momentum arising from the muon- and tau neutrinos.*

of events with isolated leptons and missing transverse momentum presented in this thesis, electrons and muons from leptonic tau decays cannot be distinguished from those produced in the primary *ep* interaction. Consequently, leptonic tau decays may contribute to the event samples with isolated electrons or muons, but not to the event sample with isolated tau leptons.

The experimentally accessible decay channels are the hadronic decays of the tau. The jets resulting from hadronic tau decays ("tau jets") can be distinguished from quark- and gluon-induced jets ("QCD jets"), by their characteristic low particle multiplicity. In comparison to the number of hadrons produced in "one-prong" and "three-prong" decays of tau leptons, the typical particle multiplicities in jets produced by quarks and gluons are significantly higher, due to parton showering before hadronisation. Additionally, the hadrons produced in decays of high momentum tau leptons are boosted in the direction of the original tau lepton. Therefore, high momentum tau jets are very collimated. In comparison, QCD jets are significantly less collimated, since the colour-connection to the rest of the hadronic final state enhances the spread of particles emerging from the hadronisation.

The identification of hadronic tau decays is based on the low particle multiplicity and radius of tau jets. These two characteristic features are illustrated in figure 4.10 together with the same quantities for a sample of simulated QCD jets. In the H1 detector, the particle multiplicity and collimation of jets is reconstructed from information provided by the CTD, the FTD and the LAr calorimeter. The "one-prong" and "three-prong" decay modes of the tau are expected to give rise to either one or three tracks in the tracking detectors, matching narrow and deep hadronic showers in the calorimeter. Electromagnetic showers initiated by photons resulting from neutral pion decay may also be found in the calorimeter. All tracks and particle showers are expected to be collimated, due to the boost of the charged and neutral pions in the original tau direction. A typical experimental signature of hadronic tau decays is shown in figure 4.11, in comparison to that of a typical QCD-jet.

The algorithm for the identification of hadronic tau decays (the "tau finder") processes the list of jets that result from the hadronic reconstruction described in section 4.6. Each jet is tested individually for its compatibility with the signature

final state.

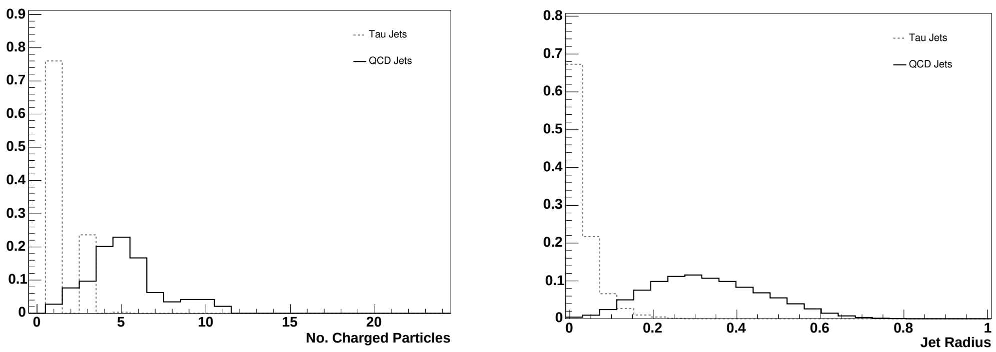


Figure 4.10: *Distributions of the expected charged particle multiplicities and jet radii of tau jets in comparison to QCD jets.*

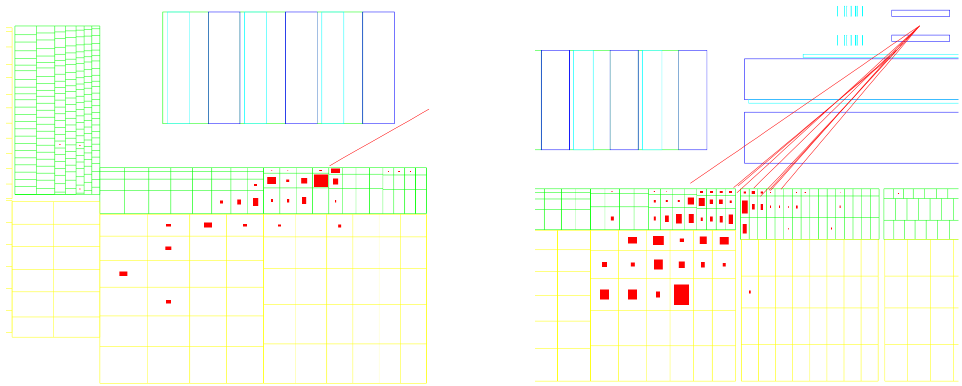


Figure 4.11: *Event displays of a typical tau jet (left) in comparison to a QCD-jet (right) in the H1 detector. The tau jet is associated to fewer tracks in the drift chambers and is more collimated in the calorimeter than the QCD-jet.*

expected for a hadronic tau decay, based on the information provided by the CTD, the FTD and the LAr calorimeter. The efficienct separation of tau jets from similar low multiplicity QCD jets requires a detailed analysis at the track and cluster level. At the track level, the multiplicity and collimation of charged particles in the tau jet is reconstructed from the tracks in the drift chambers. At the cluster level, the multiplicity and collimation of all particles in the tau jet, including the photons produced by neutral pion decays, is reconstructed by analysing the shape of the particle showers in the calorimeter. The shower-shape analysis is composed of two independent parts: the reconstruction of the neutral particle multiplicity in the tau jet and the computation of jet shape observables that distinguish tau jets from QCD jets. At low and medium tau momenta, the fine granularity of the LAr calorimeter makes it possible to separate the electromagnetic and hadronic showers induced by the neutral and charged decay particles and reconstruct the individual photons and neutral pions that are produced in varying numbers in the different decay modes of tau leptons. At high tau momenta, the multiplicity and collimation of the particles in the jet is estimated using jet shape observables; as the distances between the decay particles diminish due to the Lorentz boost in the tau direction, the electromagnetic and hadronic showers overlap too much for individual particles to be reconstructed.

In the following subsections, the reconstruction of the charged and neutral particle multiplicities is described in more detail. In the first subsection, the definition of the *tau cone* is given within which the track and cluster level analyses are performed. The size of the tau cone is such that it contains the charged and neutral particles produced in tau decays. It is surrounded by an *isolation cone* of fixed size (opening angle of 1.0 in the η-ϕ metric), that is expected to be empty for tau jets, but will contain the particles outside the tau cone for QCD jets. The track level analysis is described in the second subsection. The tau jet candidates are preselected on the basis of the charged particle multiplicity, estimated from the number of tracks reconstructed in the CTD and FTD tracking detectors: in order to minimize the overall run-time of the tau identification algorithm, the shower-shape analysis on the cluster level is only started for jets that are compatible with a tau signature on the track level. The two independent parts of the cluster level analysis, the reconstruction of the neutral particle multiplicity and the computation of jet shape observables, are discussed separately in subsections three and four. In the fifth subsection, the identification of hadronic tau decays with neural networks is described. In subsections six and seven, the rejection of unidentified electrons, which may fake hadronic tau decays if they become misidentified as hadrons, is discussed. At HERA, the rejection of unidentified electrons is especially important, as the cross-section for beam electrons to be scattered into the liquid argon calorimeter exceeds the cross-section for the production of tau leptons by several orders of magnitude.

4.5.1 Definition of the tau cone

The particle multiplicity and collimation of tau jets is analysed in a *tau cone*, whose size is adapted to contain the charged and neutral decay particles produced in the tau decay. The expected angle between the decay particles in the detector depends on the Lorentz boost in the tau direction. In order to quantify the effect of the Lorentz boost, the tau lepton is considered to decay to two massless particles, as illustrated in figure 4.12 [4]. With this simplification, the angle α between the decay

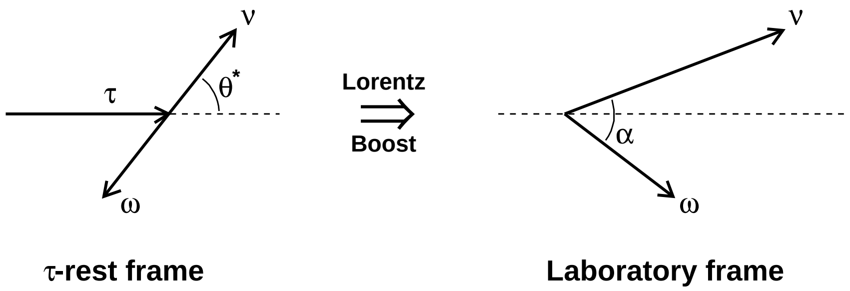


Figure 4.12: *Illustration of the idealized two-body decay of tau leptons.*

particles in the detector may be written as a function of the tau momentum p_τ and

[4] A detailed discussion of hadronic tau decays may be found in reference [145].

the decay angle θ^* in the rest frame of the tau lepton as

$$\alpha = \arccos\left(\frac{p_\tau^2 \sin^2\theta^* - m_\tau^2}{p_\tau^2 \sin^2\theta^* + m_\tau^2}\right).$$

The dependence of α on the decay angle θ^* and the momentum p_τ of the tau lepton is shown in figure 4.13. As expected, the angle between the decay particles in the detector is seen to decrease with increasing tau momentum.

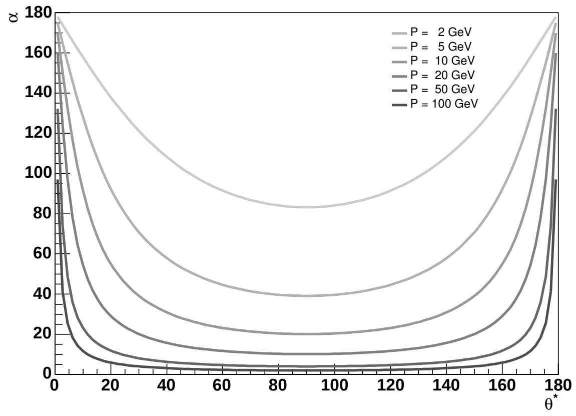


Figure 4.13: *Angle α between the decay particles in the laboratory frame as function of the decay angle θ^* in the rest frame of the tau lepton for tau momenta p_τ of* $2, 5, 10, 20, 50$ *and* 100 *GeV.*

The dependence of α on the decay angle θ^* and the tau momentum p_τ constitutes a practical problem for defining the size of the tau cone. In general, θ^* and p_τ cannot be reconstructed, due to the undetected momentum carried away by the tau neutrino. As a consequence, the size of the tau cone can only be chosen such, that the decay particles are contained in the tau cone for most - but not all - tau decays. In the tau identification algorithm developed for the analysis presented in this thesis, the opening angle of the tau cone is chosen to be

$$\alpha_{cone} = \min\left(\arccos\left(\frac{p_\tau^2 - m_\tau^2}{p_\tau^2 + m_\tau^2}\right), 30^\circ\right). \tag{4.3}$$

Note that this opening angle corresponds to a maximum angle of $2\alpha_{cone}$ between the decay particles in the detector.

The momentum p_τ of the tau lepton that defines the size of the tau cone is estimated from the detected momentum p_{jet} of the tau jet corrected for the average momentum fraction that is carried away by the tau neutrino. The average fraction of the tau momentum that is detected in hadronic "one-prong" and "three-prong" tau decays is shown in figure 4.14 for different Monte Carlo simulations. In the figure, it can be seen that this fraction is compatible within statistical uncertainties in the different Monte Carlo simulations. The correction factors that need to be applied to correct the detected momentum p_{jet} of the tau jet to the momentum p_τ of the original tau lepton depend on the momentum spectrum of the tau leptons. This dependence is visible in figure 4.14. It may be seen, for instance, that tau

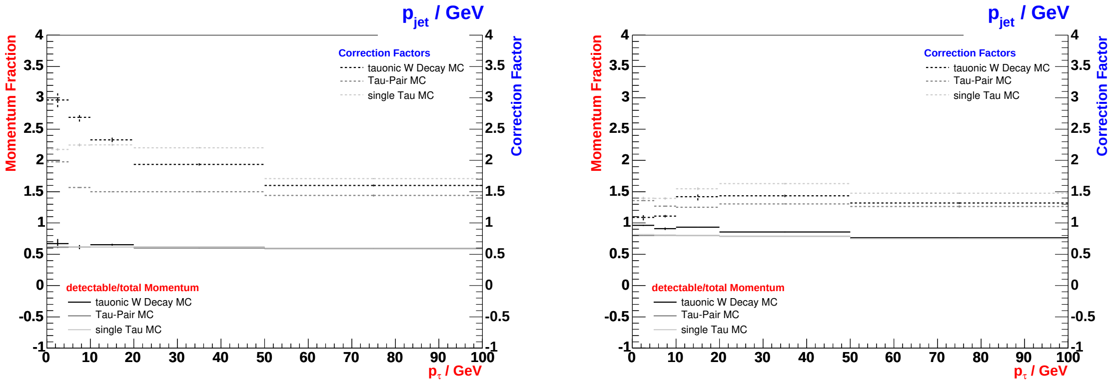


Figure 4.14: *Average fractions of tau momentum p_τ that is detectable as jet momentum p_{jet} (solid markers) for hadronic "one-prong" tau decays (left) and "three-prong" tau decays (right) and corresponding correction factors for the jet momentum (dashed markers); the undetectable momentum is carried away by the tau neutrino.*

jets resulting from W decays need to be corrected by a larger factor than tau jets resulting from tau pair production processes, as tau leptons resulting from W decays are produced with tau momenta of typically half the W mass (*cf.* section 2.6.1), whereas most tau leptons resulting from tau pair production processes are produced with low tau momenta. The correction factors that need to be applied to correct the momentum of tau leptons with a uniform momentum distribution are also shown in the figure ("single Tau"). For these, the correction factors are almost independent of the detected tau momentum and are of the order of 1.5. In order for the tau identification algorithm to be universal, this constant factor is used for estimating the momentum p_τ of the original tau lepton from the detected momentum p_{jet} of the tau jet in the determination of the size of the tau cone. In figure 4.15 it is shown that a cone defined by $p_{tau} \equiv 1.5 p_{jet}$ and equation 4.3 contains most of the decay products of tau leptons resulting from W decays. Note the band at $\alpha = 0$ that represents tau decays into a single detected particle via $\tau^- \rightarrow \pi^- \nu_\tau$.

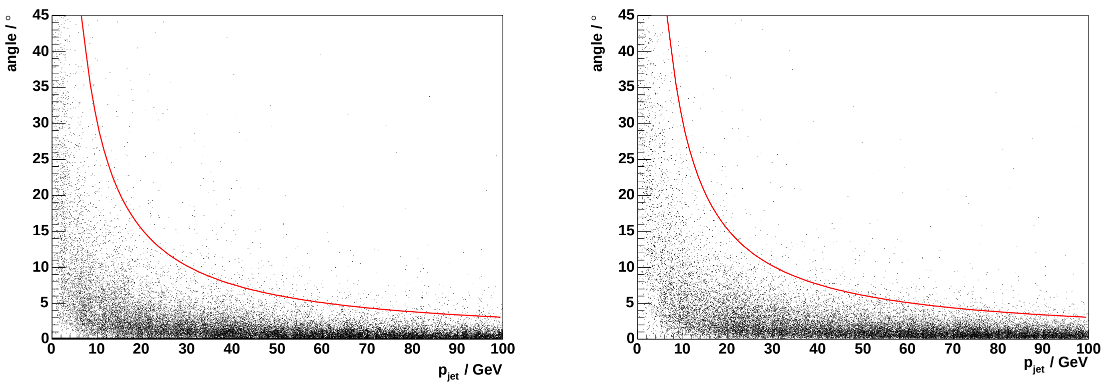


Figure 4.15: *Angle between the decay particles and the axis of the tau jet (dots) in comparison to the opening angle of the tau cone (solid line) for hadronic "one-prong" tau decays (left) and "three-prong" tau decays (right).*

The size of the tau cone still needs to be corrected for the spatial extent of the electromagnetic and hadronic showers induced by the decay particles. In order to account for the effects of electromagnetic showers (*cf.* section 4.3), the opening angle of the tau cone is enlarged by 7.5 degrees. Including this correction, the final size of the tau cone is

$$\alpha_{cone} = \min\left(\arccos\left(\frac{(1.5p_{jet})^2 - m_\tau^2}{(1.5p_{jet})^2 + m_\tau^2}\right), 30^\circ\right) + 7.5^\circ,$$

as a function of the detected momentum p_{jet} of the tau jet. As in hadronic tau decays, all particles inducing hadronic showers are charged (*cf.* the list of decay modes in table 4.3), the opening angle of the tau cone does not need to be enlarged further to account for the greater spatial extent of hadronic showers unless tracks are reconstructed at the edge of the tau cone. In this case, energy deposits that are within cones of opening angle 12.5 degrees around the extrapolated tracks are also considered to be within the tau cone.

4.5.2 Reconstruction of charged particle multiplicity

The multiplicity of charged particles produced in the tau decay is reconstructed as the number of tracks in the tau cone. At least one (three) vertex-fitted tracks in the CTD and FTD drift chambers are required within the tau cone for "one-prong" ("three-prong") tau jet candidates. Additional tracks are considered to be potential products of photon conversions. They are flagged as resulting from photon-conversion, if they can either be fitted to *secondary vertices* displaced from the primary event vertex or start behind dead material in the tracking detectors. The detailed conditions for tracks to be flagged as candidates for photon-conversion are listed in table 4.4. For "one-prong" ("three-prong") tau jet candidates, the number

Tracking detector	Conditions
CTD	two collimated opposite charged tracks fitted to a secondary vertex, flagged by the identification algorithm for converted photons
FTD	expected radial track length in CTD > 10 cm and no CTD track $\|\|$ ($z_{starthit} > 145$ cm && $r_{starthit} > 25$ cm)

Table 4.4: *Selection criteria for tracks to be flagged as possible products of photon-conversion. For the identification of electron/positron tracks produced by converted photons in the CTD, a dedicated algorithm ("converted photon finder") exists in the H1OO framework.*

of vertex-fitted tracks of transverse momentum above $p_T > 150$ MeV that cannot be interpreted as being due to photon-conversion is required to be not more than one (three). Jets with a higher number of tracks in the tau cone or tracks in the isolation cone surrounding the tau cone are rejected and excluded from further analysis, in order to minimize the overall run-time of the tau identification algorithm. For the same reason, jets are also excluded from further analysis, if the sum of track momenta is below 2 GeV. Note that ambiguities may arise in the interpretation

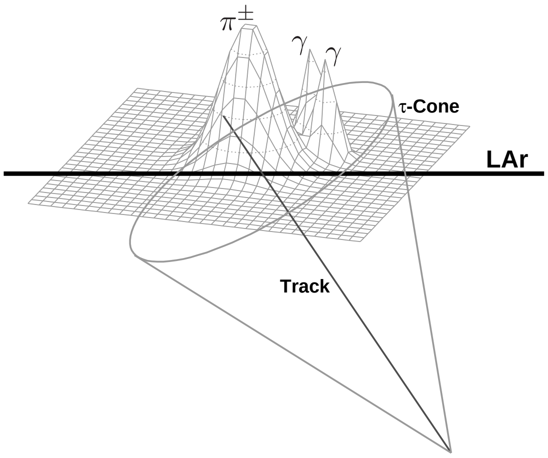


Figure 4.16: *Illustration of the energy distribution in the plane perpendicular to the jet axis for a tau jet with one charged and one neutral pion that decays into two photons.*

of tau jets as either "one-prong" or "three-prong" in some cases, for instance, if there are three vertex-fitted tracks of which two are compatible with being due to photon-conversion. In such cases, the one-prong hypothesis is given preference.

4.5.3 Reconstruction of the neutral particle multiplicity

The multiplicity of neutral particles is reconstructed using a shower-shape analysis of the energy deposits within the tau cone in the LAr calorimeter [5]. In the context of tau identification, the reconstruction of the neutral particle multiplicity is considerably eased by the fact that the decay modes of tau leptons are known. In hadronic tau decays, all charged particles are either charged pions or kaons, while all neutral particles are photons, produced in the decays of neutral pions. The electromagnetic showers are reconstructed from the distribution of energy deposits in the electromagnetic section of the LAr calorimeter in a three stage procedure. In the first stage, the energy deposits that are due to hadronic showers are removed. In the second stage, the distribution of energy deposits not attributed to charged hadrons is searched for local maxima, indicating the presence of electromagnetic showers induced by single photons (see figure 4.16 for illustration). Finally, an attempt is made to combine the reconstructed photons to form neutral pions, based on a mass-window criterion.

The distribution of energy deposits is analysed in two-dimensional histograms, which represent the distribution of energy deposits in the LAr calorimeter projected onto a plane perpendicular to the jet axis and which are filled before the three stage shower-shape analysis procedure is started. The energy deposits in the electromag-

[5]The reconstruction of the neutral particle multiplicity is performed by the tau identification algorithm, as the more general H1 algorithms for reconstruction of the hadronic final state (described in section 4.6) are designed to reconstruct the hadronic energy flow (sum of particle momenta), but not to reconstruct individual neutral particles. For example, no attempt is made to compensate for the splitting of hadronic showers into several clusters in the reconstruction software of the LAr calorimeter (*cf.* section 1.2.2) and to merge the clusters that result from one and the same hadronic shower. As a result, the multiplicity of neutral particles is often overestimated by this algorithm.

netic and hadronic sections are entered into separate histograms (see illustration in figure 4.17). The sizes of the histogram bins are $7 \times 7\ cm^2$, roughly matching the

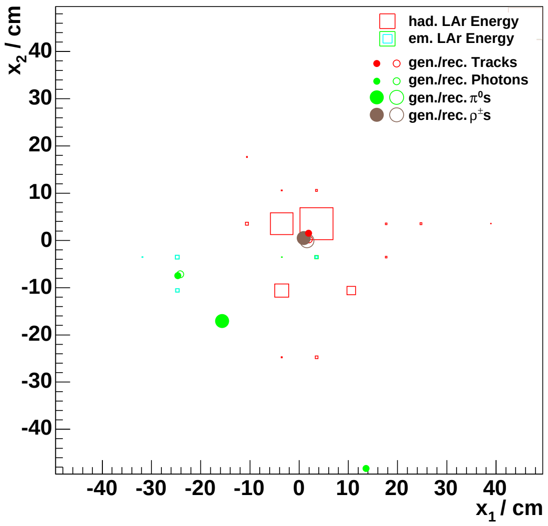


Figure 4.17: *Reconstruction of the neutral particle multiplicity in the LAr calorimeter. For an example of a simulated tau decay into one charged and one neutral pion, the reconstructed positions of the hadronic (open red points) and electromagnetic showers (open small green points) are shown together with the true impact positions of the charged pion (full red point), the neutral pion (full large green point) and the two photons (full small green points). The boxes indicate the energy deposits in the electromagnetic (green) and hadronic (red) sections of the LAr; the energy deposits in the first electromagnetic layer are also indicated (cyan).*

average size of the cells in the electromagnetic section of the LAr calorimeter. The energy deposited in cells that overlap with more than one bin is split between those bins according to the overlap of the projected cross-section of the cell with the area of the bin [6].

After the histograms have been filled, the first stage of the shower-shape analysis is the removal of the contributions of charged hadrons to the energy distribution in the electromagnetic section. These contributions are estimated from the difference between the track momenta and the energy deposits in the hadronic section of the LAr within a cylinder of radius 25 cm around the extrapolated tracks. In the case of ambiguities, that is, if the energy deposits are within the cylinders around more than one track, the fraction of energy deposited in cell i of the hadronic section attributed to track j is determined by the ratio

$$f_{ij} = \frac{\delta_j}{\sum_k \delta_k}, \qquad \delta_j \equiv P_j - \sum_l E_l^{had},$$

so as to best match all track momenta P_j and energy deposits E_l^{had} in the hadronic section. In the above formula, $\sum_l E_l^{had}$ denotes the sum over all energy deposits in

[6]To calculate the overlap, the cells are approximated by a cube of edge length $l = V^{1/3}$, which is a necessary simplification as only the position of the centre of cells and their volume V is available in the data-base that stores the calorimeter geometry.

the hadronic section that are unambigously attributed to track j and the sum over k extends over all tracks near cell i. The difference

$$\Delta E_j = \left(P_j - \left(\sum_k E_k + \sum_i f_{ij} E_i^{had} \right) \right)$$

between the momentum of track j and the sum of the energy deposits in the hadronic section attributed to that track, is considered to be the contribution of the charged hadron to the energy deposits in the electromagnetic section of the LAr and is removed before the reconstruction of electromagnetic showers. The contribution of charged hadrons from the distribution of energy deposits in the electromagnetic section is removed by assuming the energy distribution of hadronic showers in the transverse plane is Gaussian [7]. The energy attributed to charged hadron j is then determined by

$$E_{ij}^{had} = \Delta E_j \cdot \int_A Gauss\,(x, y)\, dx\ dy$$

and is substracted from the energy distribution in the electromagnetic section. The integral extends over the area A of the bins i within the cylinder. The integrand is a normalized Gaussian distribution centered at the centre of gravity of the energy depositions attributed to track j in the hadronic section (if there are none, the impact position of the extrapolated track is used); the Gaussian is assigned a width of one bin width, so as to attribute most of the energy close to the extrapolated track to the corresponding charged hadron.

After the contributions of charged hadrons have been removed from the energy deposits in the electromagnetic section, electromagnetic showers are reconstructed in the second stage of the shower-shape analysis. The reconstruction of electromagnetic showers is based on finding local maxima in this energy distribution. These are processed independently in order of decreasing energy. The local maxima are found by identifying the calorimeter cell that maximizes the sum of energy deposits within a narrow cylinder of radius 15 cm in the energy deposits not attributed to previously reconstructed hadronic or electromagnetic showers. The position of this cell defines the position of a two-dimensional Gaussian function that is fitted to the energy distribution around the maximum. As most of the energy of electromagnetic showers is expected to be contained in a compact "hot core", only the energy deposits within the narrow cylinder are considered for the fit, in order to make the fit result less sensitive to possible other nearby electromagnetic showers. In the fit, the amplitude and width of the Gaussian function and its position are allowed to vary. After the fit has converged, the energy deposit attributed to electromagnetic shower j in bins i within the cylinder is determined by the integral of the fitted Gaussian function over the area A of the bins,

$$E_{ij}^{em} = E \cdot \int_A Gauss\,(x, y)\, dx\ dy,$$

where E denotes the amplitude of the Gaussian function. The energy of the elec-

[7] It has been checked that this approximation does not introduce any unphysical behaviour into the reconstruction of the neutral particle multiplicity.

tromagnetic shower is then reconstructed as the sum

$$E_{shower} = \sum_i E_{ij}^{em}$$

over all bins i within the cylinder. The reconstructed electromagnetic shower candidate is subject to additional selection criteria, before it is considered to have been induced by a photon. In order to reject showers induced by neutral hadrons and noise, at least one cell in the first electromagnetic layer of the LAr and a total of at least two LAr cells are required [8]. If the electromagnetic shower candidate fullfills these criteria, it is considered to have been induced by a single photon, the four-vector of which is determined by the fitted shower energy and the position of the shower centre relative to the event vertex. Otherwise, the reconstructed shower is not considered to be electromagnetic and no photon is reconstructed. In both cases, the energies E_{ij}^{em} attributed to the shower are substracted from the energy deposits of all bins within the cylinder. The energy in bins i is set to zero, if the energy remaining in that bin after the substraction is compatible with a fluctuation of the energy deposits attributed to previously reconstructed showers,

$$E_i^{em} - \sum_j E_{ij}^{em} < 2 \cdot 0.12 \sqrt{\sum_j E_{ij}^{em}},$$

on a two sigma confidence level. The energy resolution for the previously reconstructed showers is estimated to be $0.12\sqrt{\sum_j E_{ij}^{em}}$ [9]. After the contribution of shower j has been removed from the energy distribution in the electromagnetic layer, the next local maximum (of lower energy) is searched for in the remaining distribution of energy deposits and the next electromagnetic shower is reconstructed. The reconstruction of electromagnetic shower candidates continues until no further maxima can be found.

After the reconstruction of electromagnetic showers is complete, an attempt is made to combine pairs of identified photons to form neutral pions in the third stage of the shower-shape analysis. A π^0 candidate is formed, if the invariant mass of the photon pair is compatible with the neutral pion mass at the five sigma level,

$$M_{\gamma\gamma} - m_{\pi^0} < 5 \cdot \sigma_{M_{\gamma\gamma}},$$

taking into account the expected energy resolution of the LAr calorimeter for electromagnetic showers and the uncertainty of the reconstructed position of the shower centers. If one photon may be associated with more than one π^0 candidate, the combination that is closer to the π^0 mass in terms of the estimated uncertainties is given preference.

[8] Depending on the polar angle, the depth of the first electromagnetic layer varies from $\Delta x = 2.2\ X_0$ to $3.6\ X_0$; given the $\Delta x' = 1\ X_0$ to $2\ X_0$ of dead material between the interaction point and the first electromagnetic layer [25,26], the probability for electrons and photons to start showering in the first electromagnetic layer $p\,(\Delta x + \Delta x') = 1 - e^{-(\Delta x + \Delta x')}$ varies between 95% and 99%.

[9] The constant 0.12 corresponds to the energy resolution of the LAr calorimeter for electromagnetic showers (see section 1.2.2).

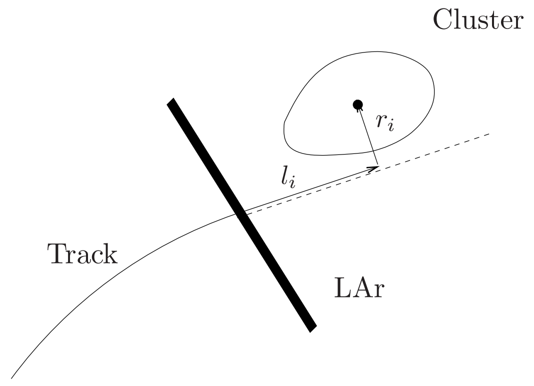


Figure 4.18: *Illustration of the variables r_i and l_i defining the radial and longitudinal moments of the energy deposits within the tau cone in the calorimeter.*

4.5.4 Reconstruction of estimator variables that distinguish tau jets from QCD jets

In order to identify hadronic tau decays at high momenta, a set of estimators that are sensitive to the multiplicity and collimation of the particles in the jet is defined. These observables may be used to separate tau jets from QCD jets if the electromagnetic and hadronic showers overlap too much for individual particles to be reconstructed.

The most important observables for the separation of tau jets and QCD jets are the *radial moments* of the jet. The first (second) radial moment describes the mean (mean squared) distance of the energy deposits in the tau cone from the jet axis, the *jet radius*. The i^{th} radial moment is defined as

$$< r^i >= \frac{\sum_j E_j r_j^i}{\sum_j E_j},$$

where the sum extends over all LAr cells j in the tau cone and r_j is the distance of the j^{th} cell to the jet axis, as is illustrated in figure 4.18. The lower response of the LAr calorimeter to hadronic showers is corrected for statistically by the energy weighting procedure described in section 1.2.2. The radial moments are sensitive to both the collimation and (to a lesser degree) the particle multiplicity of the jet and therefore provide a high separation power between tau jets and QCD jets. The longitudinal moments

$$< l^i >= \frac{\sum_j E_j l_j^i}{\sum_j E_j},$$

also provide some sensitivity to the particle multiplicity, as for a fixed jet energy, higher particle multiplicities imply lower particle energies and therefore less deep particle showers. However, in comparison to the radial moments, the separation provided by the longitudinal moments is rather limited, since the longitudinal extent of the shower depends only logarithmically on the energy of the incident particle (*cf.* section 4.2). In addition to the radial and longitudinal moments, a few complementary jet shape observables of lesser importance are defined, as is described in section B of the appendix.

4.5.5 Rejection of QCD jets

The final selection of tau jets is based on the reconstructed neutral and charged particle multiplicities in combination with the jet shape observables and some estimator variables that are specific to certain decay modes, such as the estimated probability for the formation of an intermediate ρ meson resonance. The set of observables reconstructed for each tau jet candidate may be considered as a feature vector and the separation of tau jets from QCD jets as a pattern classification problem. In the analysis presented in this thesis, this problem is solved using neural networks.

As is detailed in chapter 3, there exists some freedom in choosing the set of features used for classification. Clearly, the chosen features should have discriminative power. Equally important for tau identification, however, is that the observables are well modelled by the Monte Carlo simulation, as the neural networks used to classify the data will be trained with simulated patterns. The full set of observables used for the identification of hadronic tau decays is presented in section B of the appendix. As can be seen in the figures shown in the same section, the observables used do have discriminative power and are well modelled by the Monte Carlo simulation for QCD jets in deep-inelastic neutral and charged current processes. For tau jets, the agreement between the Monte Carlo simulation and the data is checked using a sample of tau pair production events (see section C.3 of the appendix). The significance of the cross-check provided by the tau pair events is limited, however. Firstly, the cross-section for tau pair production is small and hence the event statistics of the data sample restricted. Secondly, the tau jets found in the tau pair events are of significantly lower momenta than those expected in events with isolated tau leptons and large missing transverse momentum, which introduces some systematic uncertainties in the cross-check due to the extrapolation to higher momenta. Note, however, that the uncertainties in the simulation of tau jets are small in comparison to those of QCD jets, as the particle multiplicities in tau jets are precisely known (see table 4.3). In contrast to QCD jets, in the modelling of which significant uncertainties arise from parton showering and hadronisation, uncertainties in the modelling of tau jets arise solely from the detector simulation. These uncertainties have been estimated by comparing different simulations and found to be small (*cf.* section 8.1).

The distributions of observables shown in section B of the appendix have been corrected for their dependence on the energy and polar angle of tau jets (and QCD jets). The reconstructed neutral particle multiplicity and jet shape observables depend on the energy of the tau jet, as tau jets of higher momenta are more collimated by the Lorentz boost in tau direction. Further, the electromagnetic and hadronic showers induced by more energetic particles have a more compact "hot core" [24] and extend deeper in the calorimeter. The observed dependence on the polar angle is due to the varying granularity of the LAr calorimeter. In the forward region, the calorimeter cells are of smaller size, so the shower-shape may be more precisely reconstructed. As an example, the energy and polar angle dependence of the average jet radii reconstructed for tau jets are shown in figure 4.19. If the observables used for tau identification were not corrected for their energy and polar angle dependence, the identification of hadronic tau decays would depend on the different energy and polar angle distributions of tau jets and QCD jets, that is, on the *event kinematics*.

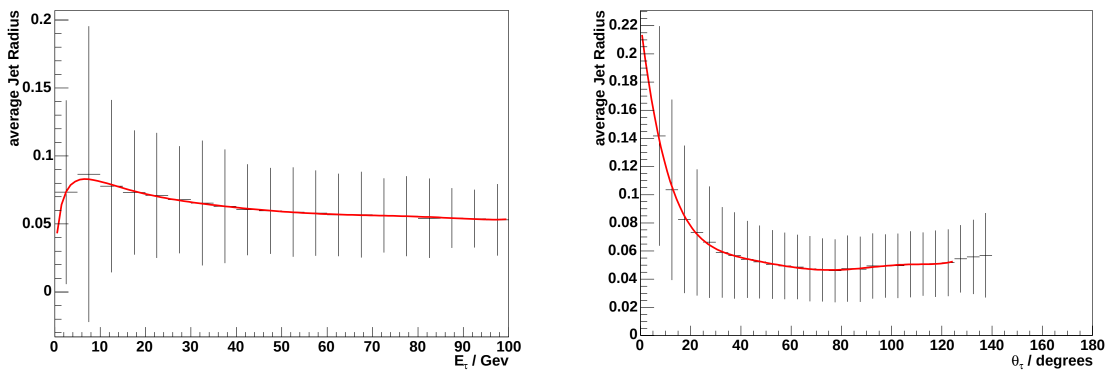


Figure 4.19: *Energy and polar angle dependence of the average jet radii reconstructed for simulated tau jets. The average jet radii are represented by crosses. The "error bars" indicate the spread of individual tau jets around the average. The solid line represents a smooth interpolation that may be used to correct the jet radii for their energy and polar angle dependence.*

In order to obtain a universal tau identification algorithm that is independent of the event kinematics, either the energy and polar angle dependence of the observables used for tau identification has to be eliminated, or the samples of tau jets and QCD jets that are used for the training of the tau identification algorithm have to be reweighted to a common energy and polar angle distribution. The method of correcting the observables for their energy and polar angle dependence is described in reference [146]. In the analysis presented in this thesis, the reweighting technique is used, as it is better suited to the neural network approach. In the reweighting technique, the feature vectors of tau jets and QCD jets are binned in energy and polar angle. They are then assigned the weights

$$w = \begin{cases} \max\left(1, \frac{N_{QCD-jet}(E,\theta)}{N_{tau-jet}(E,\theta)}\right) & \text{if the jet is a tau jet} \\ \max\left(1, \frac{N_{tau-jet}(E,\theta)}{N_{QCD-jet}(E,\theta)}\right) & \text{if the jet is a QCD jet,} \end{cases}$$

where $N_{tau-jet}(E,\theta)$ and $N_{QCD-jet}(E,\theta)$ are the number of tau jets and QCD jets in the same energy and polar angle bin as the feature vector that is to be assigned the weight, respectively. After these weights have been applied to all feature vectors, the energy and polar angle distributions of the tau jets and QCD jets are equal. The energies and polar angles of the jets are appended to the weighted feature vectors. Given this reweighting, the neural network may for example find that a jet of radius 0.15 (in the η-ϕ metric) and an energy of 5 GeV is likely to be a tau jet, while a jet of the same radius, but an energy of 50 GeV is likely to be a QCD-jet (*cf.* the energy dependence of the average jet radii of tau jets displayed in figure 4.19).

Separate neural networks are trained to identify each decay mode of the tau leptons. The classification outputs of neural networks that have been trained to identify the "one-prong" and "three-prong" decay modes are displayed in figure 4.20. In the figure, the classification output for jets in deep-inelastic scattering events selected in the 1999/2000 e^+p data is compared with the Monte Carlo expectation.

The identification efficiencies for tau jets resulting from a cut of $f_{NN} > 0.75$ on the neural network output are shown in figure 4.21. As can be seen in the figure, the

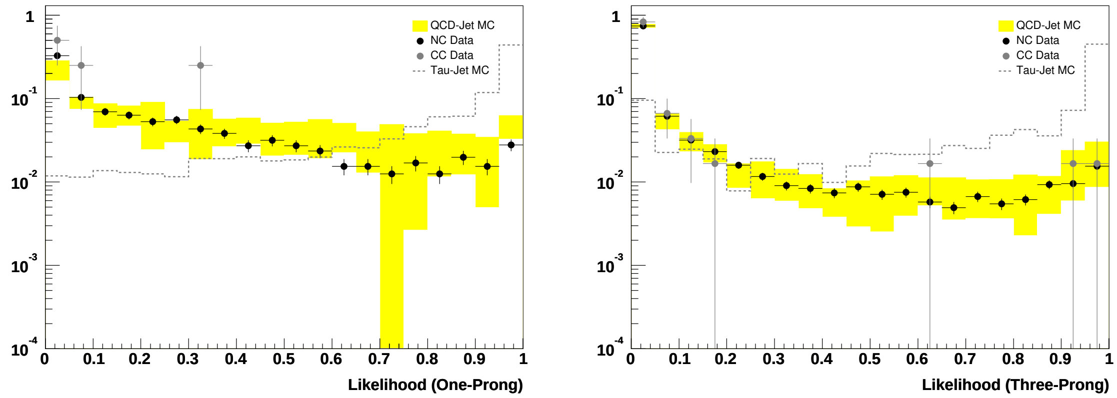


Figure 4.20: *Distributions of the classification output of neural networks that have been trained to identify the "one-prong" and "three-prong" decay modes of tau leptons for a sample of QCD jets in NC and CC DIS events selected in the 1999/2000 e^+p data (see references [31, 138] for details of the event selection) in comparison to simulated QCD jets and tau jets produced in decays of W bosons. The difference between the upper and lower limits of the shaded band represent the uncertainties of the Monte Carlo simulation that has been estimated by comparing different hadronisation models and detector simulations (see section B of the appendix for further details).*

efficiency for the identification of hadronic tau decays increases monotonically with the energy of the tau lepton. The efficiency to identify tau leptons of energies less than $E_\tau < 5\ GeV$ is small, due to the minimum transverse momentum of $2\ GeV$ required for the summed track momenta in the preselection of tau jet candidates (*cf.* section 4.5.2). For a fixed tau energy, "three-prong" decays are identified with higher efficiency, as the fraction of transverse momentum on average carried away by the tau neutrino is smaller than in "one-prong" decays. The identification efficiency for both "one-prong" and "three-prong" decays is also seen to depend strongly on the polar angle of the tau lepton. In the forward region, the identification efficiencies are smaller than in the central region, due to the lower efficiency for track reconstruction in the FTD and the high probability for photon-conversion in the dead material between the central and forward tracking detectors which may lead to an increase in the reconstructed charged particle multiplicity. As a consequence of the low track reconstruction efficiency and the high QCD-jet background in the forward region, jets with polar angles below $\theta_{jet} < 20°$ are not considered as "three-prong" candidates (the non-vanishing probability for "three-prong" decays to be identified in the forward region results from decays identified as "one-prongs" and decays in which the detected particles are detected at larger polar angles than the tau lepton was actually produced). The decrease seen in the identification efficiency for "one-prong" decays at polar angles $\theta_\tau \sim 80°$ is due to the *fiducial cuts* necessary to reject unidentified electrons (see section 4.5.6). The increase that may be seen in the identification efficiency for "three-prong" decays at polar angles around 40 and 120 degrees is an artefact of the $p_T^\tau > 10\ GeV$ requirement: this transverse momentum threshold implies higher energies and hence a higher collimation of the tau jet for tau leptons of polar angles at the extremes of the allowed polar angle range. All in all, the efficiency for identifying hadronic "one-prong" and "three-prong" decays

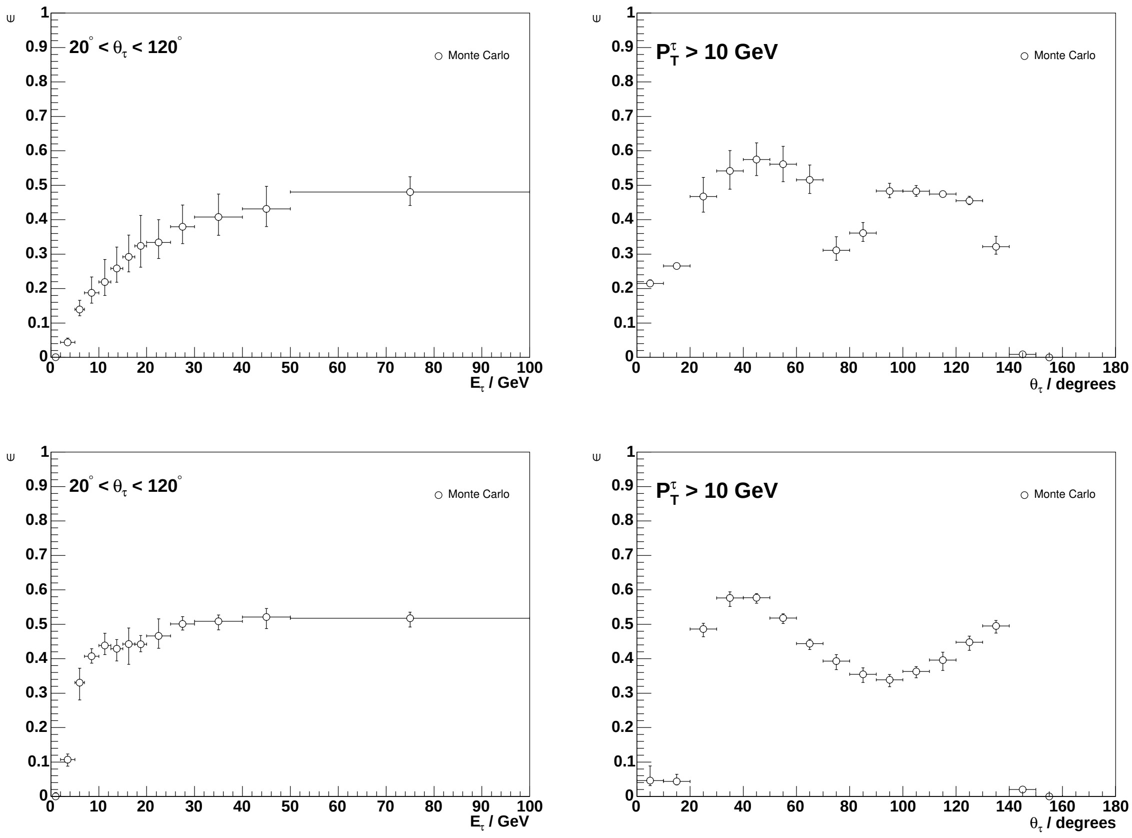


Figure 4.21: *Identification efficiency for tau jets resulting from "one-prong"-tau decays (top) and "three-prong" tau decays (bottom). The identification efficiencies are estimated for a cut of $f_{NN} > 0.75$ on the neural network output shown in figure 4.20 by a simulated sample of Monte Carlo events with single tau leptons.*

is about 50% for tau leptons of transverse momentum above $p_T^\tau > 10\ GeV$ in the central region.

The misidentification probabilities for QCD jets resulting from a cut of $f_{NN} > 0.75$ on the neural network output are shown in figure 4.22. As expected, due to their high particle multiplicity, QCD jets are more likely to be misidentified as "three-prong" than as "one-prong" tau jet candidates: The probability for QCD jets to be misidentified as "three-prong" ("one-prong") tau jet candidates is about 4% (0.5%). The increase seen in the probability for QCD jets to be misidentified as "one-prong" tau decays in the forward region is due to the looser track isolation criteria applied in the preselection of tau jet candidates in that direction (*cf.* section 4.5.2). No corresponding increase in the probability for QCD jets to be misidentified as "three-prong" tau decays is seen, as jets of polar angles below $\theta_{jet} < 20°$ are not considered as "three-prong" tau jet candidates.

Note that the misidentification probabilities for tau leptons displayed in figure 4.22 cannot be directly compared with those for electrons and muons displayed in figures 4.6 and 4.8, respectively: the tau misidentification probabilities refer to jets, whereas those for electrons and muons refer to single charged hadrons. As in most processes which form the background to events with isolated leptons, jets are pro-

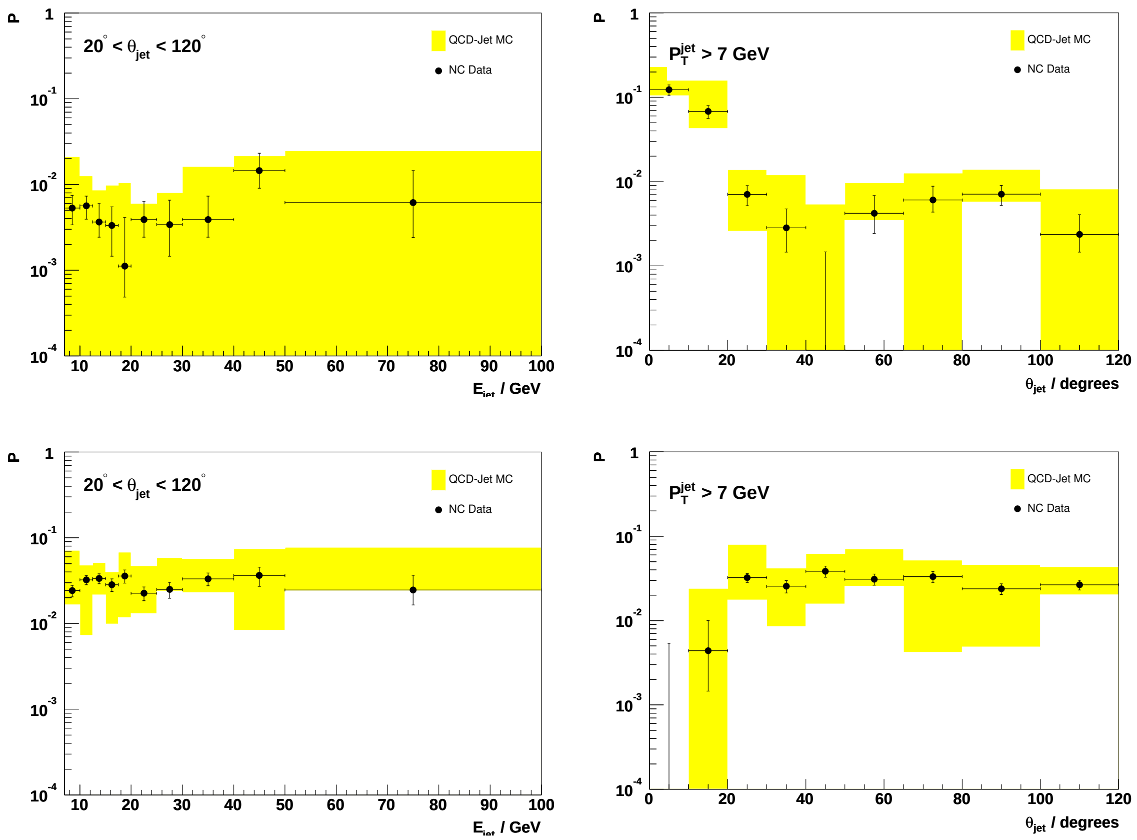


Figure 4.22: *Misidentification probabilities for QCD jets as "one-prong" (top) and "three-prong" (bottom) tau decays. The misidentification probabilities for QCD jets are estimated for a cut of $f_{NN} > 0.75$ on the neural network output shown in figure 4.20 for a sample of QCD jets in NC DIS events selected in the 1999/2000 e^+p data (see references [31, 138] for details of the event selection) and simulated QCD jets. The difference between the upper and lower limits of the shaded band represent the uncertainties of the Monte Carlo simulation that has been estimated by comparing different hadronisation models and detector simulations (see section B of the appendix for further details). The uncertainties on the probability for QCD jets to be misidentified as "one-prong" tau decays are overestimated as statistical uncertainties of the Monte Carlo samples used for this comparison become relevant because of the small misidentification probabilities.*

duced with significantly higher rate than isolated charged hadrons, the background arising from lepton misidentification in event samples with isolated electrons and muons is significantly smaller than in event samples with isolated tau leptons, although the misidentification probabilities for tau leptons displayed in figure 4.22 are roughly comparable with those for electrons and muons displayed in figures 4.6 and 4.8.

4.5.6 Electron rejection

In the identification of hadronic tau decays, special attention needs to be given to the rejection of electrons, firstly because the cross-section for beam electrons to be

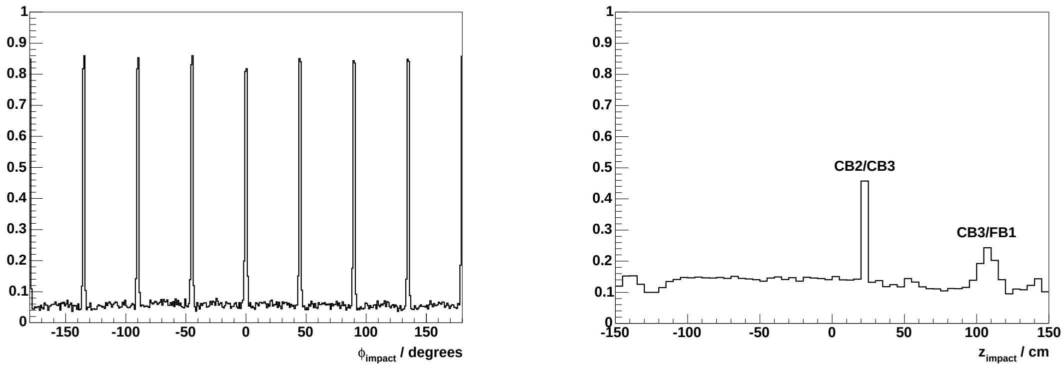


Figure 4.23: *Dependence of the probability for electrons to fail the electron identification criteria described in section 4.3 on the azimuthal angle (left) and the z-coordinate of the impact position in the LAr calorimeter (right).*

scattered in the LAr calorimeter is very large, and secondly because the experimental signatures of electrons may be similar to those of tau leptons decaying into one charged pion and one or more neutral pions of high energy.

In order to avoid the misidentification of electrons as "one-prong" tau decays, tau jet candidates are rejected if they contain any (isolated) electromagnetic showers passing the selection criteria described in section 4.3.

After identified (isolated) electromagnetic showers have been rejected, the remaining unidentified electrons are concentrated in the crack regions of the LAr calorimeter (see figure 4.23). In these, the electron identification is inefficient as interactions of the shower particles with the dead material can cause distortions of the shower-shape that lead to the estimator variables to take on atypical values. The electron identification is particular inefficient in the ϕ-cracks, due to their projective geometry (*cf.* the geometry of the LAr calorimeter in figure 1.9): in the gaps between the different octants, incident electrons may reach the first hadronic layer before they start showering, causing the electron to fail the identification due to the low fraction of energy deposited in the electromagnetic section of the LAr. To suppress the unidentified electrons in the crack regions - and also because the crack regions are difficult to describe in the Monte Carlo simulation - "one-prong" tau jet candidates are subject to *fiducial cuts*. "One-prong" candidates are rejected if the impact position of the extrapolated track in the LAr calorimeter is within $\Delta\phi = \pm 2^\circ$ of a ϕ-crack, or the centre-of-gravity of the energy deposits in the calorimeter is close to the z-crack between the CB2 and CB3 wheels at $z \ \epsilon \ [15, 30] \ cm$ [10].

After the exclusion of the crack regions, a few unidentified electrons remain due to statistical fluctuations in the shower development. Typically, the remaining unidentified electromagnetic showers fail the electron selection criteria, because they are not quite compact enough or start a little too deep in the calorimeter. In order to remove these remaining unidentified electrons, the identification algorithm for hadronic tau decays reconstructs the same estimator variables that are used in the electron identification algorithm. By applying looser conditions on these estimator

[10] The z-crack between the CB3 and FB1 wheels does not need to be cut, as the cross-section for beam electrons to be scattered so forward is small.

variables, the majority of unidentified electrons may be rejected.

4.6 Hadronic Reconstruction

The identification of hadrons is implicit and is based on ruling out other particle hypotheses. In the *HadrooII* [32] reconstruction algorithm for the *hadronic final state* ("hfs-finder"), all particles that have not been identified as leptons or photons are assumed to be charged pions. The algorithm performs two important tasks, namely, associating tracks with clusters and determining the best estimate of the particle four-vector for these associated tracks and clusters. As the momentum resolution of tracking detectors deteriorates with increasing momentum, while the relative energy resolution of calorimeters improves, an optimal reconstruction of the particle four-vector is attained by a momentum dependent combination of the track and cluster information.

The inputs to the algorithm are vertex-fitted tracks in the central and forward tracking detectors passing the quality criteria given in reference [147] and clusters in the LAr and SpaCal calorimeters. The clusters in the LAr calorimeter are subject to filter algorithms [148], mainly to further suppress electronic noise that may remain after the standard noise-suppression algorithms implemented in the reconstruction software of the LAr calorimeter described in section 1.2.2.

The association of tracks and clusters is performed by extrapolating tracks to the first electromagnetic layer of the LAr and SpaCal calorimeters; matching clusters are searched for inside a cylinder of radius 25 *cm* in the electromagnetic and 50 *cm* in the hadronic section. All energy deposits in the cylinder are corrected for the lower calorimeter response to hadronic showers by the energy weighting procedure described in section 1.2.2. The measured track momentum is then compared to the sum of the energy deposited within the cylinder, to determine if this energy is compatible with that of the charged track or if it is likely to contain contributions from other nearby hadrons. The track and calorimeter signatures are considered to be comptible with a single charged hadron if the energy contained within the cylinder does not exceed the reconstructed track momentum by more than 1.96 standard deviations (i.e. is compatible at the 5% confidence level) in terms of the estimated uncertainty of the reconstructed track momentum and the expected energy resolution of the LAr and SpaCal calorimeters for hadronic showers:

$$E_{cylinder} < E_{track} \; + \; 1.96 \cdot \sqrt{\sigma_E^{track\,2} + \sigma_E^{cylinder\,2}}.$$

The expected energy resolution of the calorimeters for a hadron of energy $E_{track} = \sqrt{\left(\frac{p_T^{track}}{\sin\theta_{track}}\right)^2 + m_{\pi^\pm}^2}$, determined by the reconstructed transverse momentum and polar angle of the track, is (*cf.* section 1.2.2)

$$\sigma_E^{LAr} = \frac{0.50}{\sqrt{E_{track}}}, \qquad \sigma_E^{SpaCal} = \frac{0.29}{\sqrt{E_{track}}}.$$

If the track and calorimeter signatures are compatible with arising from a single charged hadron, all clusters in the cylinder are associated with the track. Otherwise, only a fraction of the energy deposits within the cylinder corresponding to E_{track} is associated with the track, and the rest is attributed to other hadrons. In this case, the clusters that are closest to the extrapolated track are associated first, possibly splitting the last associated cluster to exactly match the energy E_{track}.

The best estimate for the particle four-vector depends on whether the track momentum and the energy deposits within the cylinder are compatible. If this is the case, or if the energy contained in the cylinder is lower than the track momentum, perhaps due to energy loss in the dead material in front of the calorimeter, the four-vector of the particle is determined as follows, depending on the estimated uncertainty of the track momentum and the expected resolution of the calorimeter for a particle of energy E_{track}. If

$$\sigma_E^{track} < \sigma_E^{calorimeter},$$

the particle four-vector is set to E_{track}; otherwise it is set to $E_{cylinder}$. If the energy deposits within the cylinder are incompatible with arising from a single charged hadron, the calorimeter information cannot be used for the determination of the particle four-vector and the four-vector of the particle is set to E_{track}.

After tracks and clusters have been associated and the particle four-vector has been determined, the particle is completely reconstructed. The track and the associated clusters are removed from the input list, and the algorithm continues to the next track.

Clusters remaining in the input list after all tracks have been associated are interpreted as energy deposits resulting from neutral particles. Individual neutral particles are associated to each cluster; no attempt is made to merge clusters that may be due to the same hadronic shower. The four-vector of neutral particles is set to the cluster energy, calculated at either the electromagnetic energy scale or including the corrections for the lower response of the calorimeters to hadronic showers. The electromagnetic energy scale is used for clusters with a high electromagnetic energy fraction, which are likely to be due to neutral pions or photons. The hadronic energy scale is used for all other clusters that are likely to be due to neutral hadrons, such as neutrons or long-lived neutral Kaons.

The charged and neutral hadrons reconstructed by the *HadrooII* algorithm are then assigned to jets. The aim is to associate all particles that originate from parton showering and hadronisation of the same high energy quark or gluon to the same jet; accordingly, non-isolated leptons and photons are included in the jets, for they are probably produced in the decays of heavy quarks, charged or neutral pions, and kaons. The jets are reconstructed using the longitudinally invariant k_T *algorithm* [149, 150] with a separation parameter $R_0 = 1.0$, and a minimal transverse momentum of $p_T = 2.5$ GeV is required for the jets. The jet four-vector is determined by a p_T-weighted *recombination scheme* [58],

$$p_T^{jet} = \sum_i p_{T_i}, \qquad \eta^{jet} = \frac{\sum_i p_{T_i}\eta_i}{\sum_i p_{T_i}}, \qquad \phi^{jet} = \frac{\sum_i p_{T_i}\phi_i}{\sum_i p_{T_i}},$$

that produces massless jets.

The energy of reconstructed jets is recalibrated as detailed in reference [32]. If no jets are found, the four-vector sum of all hadrons and non-isolated leptons is subjected to recalibration.

The hadronic final state is reconstructed from the calibrated jets and individual hadrons that are not assigned to jets.

Chapter 5

Reconstruction of global Event Quantities

The dominant Standard Model contribution to events with isolated leptons and large missing transverse momentum at HERA is the production and subsequent decay of real W bosons, a rare process with a cross-section of $\mathcal{O}\,(1\ pb)$, as detailed in section 2.6. Some of the background processes described in section 2.5 have a cross-section that is a few orders of magnitude larger than this. Efficient selection of the signal and rejection of the background requires careful definition of the observables that are employed for the event selection.

In this chapter, the global event quantities used in the the selection and analysis of events with isolated leptons and large missing transverse momentum are presented. Following the definition of each variable, its main characteristics and discriminatory properties are briefly described.

5.1 Kinematic Variables

5.1.1 P_T^{calo}

The quantity P_T^{calo} is defined as the transverse component of the vectorial sum of the energy deposits in the calorimeters,

$$P_T^{calo} \equiv \sqrt{V_x^2 + V_y^2}, \quad \text{with} \quad V_x \equiv \sum_i E_i \sin\theta_i \cos\phi_i, \quad V_y \equiv \sum_i E_i \sin\theta_i \sin\phi_i.$$

The index i runs over all cells above noise level in the LAr and SpaCal calorimeters; energy deposits in the tail-catcher and the plug are not included in the sums. E_i, θ_i and ϕ_i are the energy, polar and azimuthal angles of the i^{th} calorimeter cell, respectively. The angles θ_i and ϕ_i are determined by the positition of the cell centre relative to the nominal interaction point. The energies E_i are corrected for energy losses in the inactive material between the interaction point and the calorimeter and for nuclear excitation and break-up in hadronic showers using the energy-weighting algorithm described in section 1.2.2.

In events without muons, P_T^{calo} is similar to P_T^{miss}, the missing transverse momentum. As the relative energy resolution of the calorimeters improves for particles of higher energies, whereas the relative momentum resolution of the tracking detectors degrades, energetic particles are predominantly measured in the calorimeters and the contribution of the tracking detectors to the event reconstruction is small in high p_T analyses. The relation between P_T^{calo} and P_T^{miss} is different for events with muons, however. As high p_T muons deposit only a fraction of their energy in the calorimeter, they contribute significantly to P_T^{miss}, but influence P_T^{calo} hardly at all. In events with muons (and no detected scattered electron of high p_T), P_T^{calo} is approximately equal to P_T^X, the transverse momentum of the hadronic final state.

5.1.2 P_T^{miss}

The *missing transverse momentum*, P_T^{miss}, is defined as the transverse component of the sum of the four-vectors of all detected particles in the event,

$$P_T^{miss} \equiv \sqrt{P_x^{miss^2} + P_y^{miss^2}}, \quad \text{with} \quad P_x^{miss} \equiv -\sum_i p_{x_i}, \quad P_y^{miss} \equiv -\sum_i p_{y_i}. \tag{5.1}$$

As the initial state particles in the electron and proton beams have no transverse momentum, any imbalance in the sum of transverse momenta of the detected particles has to be balanced by undetected particles (such as neutrinos) in the final state. Note that undetected particles in the proton remnant and electrons scattered through small angles going down the beam-pipe in the forward or backward directions, do not contribute significantly to the undetected transverse momentum.

5.1.3 P_T^X

The transverse momentum of the hadronic final state, P_T^X, is defined as the transverse component of the sum of the four-vectors of all hadrons, photons and non-isolated leptons in the event.

5.1.4 γ

The *inclusive hadronic angle*, γ, is defined as

$$\gamma \equiv 2 \arctan \left(\frac{E_X - P_z^X}{P_T^X} \right),$$

where E_X, P_T^X and P_z^X are the energy, transverse and longitudinal momentum of the hadronic final state, respectively.

5.1.5 δ^{miss}

The *missing longitudinal momentum*, δ_{miss}, is defined to be the difference in $E - P_z$ between the initial state and all detected particles in the final state of an event,

$$\delta_{miss} \equiv 2E_e^{beam} - \sum_i (E - P_z)_i \,.$$

Here, the summation index i runs over all detected particles, E_e^{beam} is the energy of the electron beam and $2E_e^{beam}$ corresponds to the energy minus longitudinal momentum of the initial electron and proton.

From energy-momentum conservation it follows that δ_{miss} vanishes for events in which only particles going down the beam-pipe in the proton direction are undetected. Note that if the scattered electron escapes detection by going down the beam-pipe in backward direction, it contributes twice its energy to δ_{miss}. As a result, the quantity δ_{miss} is most useful in the analysis of events with a detected electron. In these events, the missing longitudinal momentum may be identified with the contribution $(E - P_z)_\nu \equiv \delta_{miss}$ of a neutrino (assuming there is only one such particle in the final state).

5.1.6 ζ^2

The quantity ζ^2 is defined as

$$\zeta^2 \equiv 2E_e^{beam} E_e \,(1 + \cos\theta_e)$$

for the highest p_T electron in the event, where E_e and θ_e are the energy and polar angle of this electron, respectively. The variable ζ^2 describes the four-momentum transfer Q^2 at the electron vertex under the assumption that the highest p_T electron in the event is the scattered beam electron.

For scattered beam electrons, the ζ^2 spectrum falls as ζ^{-4} (*cf.* the discussion of the neutral-current cross-section in section 2.2.3). In contrast to this, the ζ^2 spectrum of electrons produced in the ep interaction, and in particular of electrons resulting from the decays of heavy particles such as W bosons, falls much less steeply.

5.2 Topological Variables

5.2.1 V_{ap} and V_p

The quantities V_{ap} and V_p are defined by the projections of the energy deposits in the two hemispheres that are anti-parallel and parallel to the vectorial sum of all energy deposits in the calorimeters,

$$V_{ap} \equiv |\sum_{\substack{i \\ \vec{v}_i \cdot \vec{V} < 0}} \vec{v}_i|, \quad V_p \equiv |\sum_{\substack{i \\ \vec{v}_i \cdot \vec{V} > 0}} \vec{v}_i|, \quad \text{with}$$

$$\vec{v}_i \equiv \begin{pmatrix} V_x \\ V_y \end{pmatrix} \quad \text{and} \quad \vec{V} \equiv \sum_i \vec{v}_i .$$

The index i runs over all cells above noise level in the LAr and SpaCal calorimeters; the quantities V_x and V_y are defined in section 5.1.1.

The ratio V_{ap}/V_p allows discrimination between events which are intrinsically balanced in p_T, and where the reconstructed missing p_T is due to fluctuations in the measurement, and events in which the reconstructed missing p_T results from the presence of undetected particles (such as neutrinos) in the final state: In intrinsically balanced events (NC, γp) the ratio V_{ap}/V_p is typically close to one, whereas in events with undetected particles (neutrinos) in the final state (CC, W production), it is close to zero.

5.2.2 $\Delta\phi_{\ell-X}$

The *acoplanarity* $\Delta\phi_{\ell-X}$ is defined by the difference in azimuthal angle between the isolated lepton and the hadronic final state,

$$\Delta\phi_{\ell-X} \equiv |\phi_\ell - \phi_X| \qquad \left(P_T^X > 5\ GeV\right).$$

In order that this quantity be well defined, the transverse momentum of the hadronic final state is required to be above $P_T^X > 5\ GeV$. For $P_T^X < 5\ GeV$, the energy deposits in the plug calorimeter (if any) are taken as reference instead of the hadronic final state,

$$\Delta\phi_{\ell-X} \equiv |\phi_\ell - \phi_{plug}| \qquad \left(P_T^X < 5\ GeV\right).$$

The definition of the acoplanarity angle is illustrated in figure 5.1.

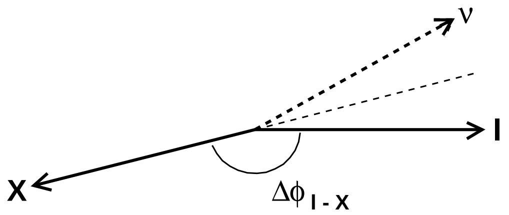


Figure 5.1: *Definition of the acoplanarity angle.*

Like the ratio V_{ap}/V_p, the acoplanarity angle allows discrimination between events which are intrinsically balanced in p_T, and where the reconstructed missing p_T is due to a mismeasurement, and events in which the reconstructed missing p_T results from the presence of undetected particles (such as neutrinos) in the final state: In intrinsically balanced events (NC, γp, lepton pair production) the hadronic final state is typically opposite to the isolated lepton in azimuth. This "back-to-back" configuration corresponds to acoplanarity angles near 180 degrees. In contrast, in events with high p_T undetected particles (such as neutrinos) in the final state (CC, W production), the isolated lepton and hadronic final state are often found with a significant acoplanarity angle, as the undetected particles are seldom produced exactly parallel to either the hadronic system or the isolated lepton.

5.3 Isolation Criteria

5.3.1 D_{jet} and D_{track}

The separation of isolated leptons to the nearest charged particle, D_{track}, is defined by

$$D_{track} \equiv \min_{\text{tracks}_i} \sqrt{\Delta\eta^2_{\ell-\text{track}_i} + \Delta\phi^2_{\ell-\text{track}_i}}.$$

Only tracks that pass the quality criteria defined in [147] are considered and the track associated with the isolated lepton is excluded in the calculation. Additionally, tracks within a cone of radius $r_e = 0.1$ (in the η-ϕ metric) around the electron direction are excluded in the calculation of D_{track} with respect to isolated electrons, in order to ensure that tracks due to photon-conversion of bremsstrahlung photons emitted by the electron do not disturb the result. In the calculation of D_{track} with respect to isolated tau leptons, all tracks within a cone of radius $r_\tau = 1.0$ around the jet axis are excluded.

The distance of isolated leptons to the nearest jet, D_{jet}, is similarly defined by

$$D_{jet} \equiv \min_{\mathrm{jet}_i} \sqrt{\Delta\eta^2_{\ell-\mathrm{jet}_i} + \Delta\phi^2_{\ell-\mathrm{jet}_i}}, \qquad \text{if number of jets} > 0.$$

As isolated leptons are excluded from the jet-finder (see section 4.6), no extra care has to be taken to avoid the D_{jet} distance from refering to the isolated lepton. In case no jet is found in the event, the entire hadronic final state X is taken as reference,

$$D_{jet} \equiv \sqrt{\Delta\eta^2_{\ell-X} + \Delta\phi^2_{\ell-X}}, \qquad \text{if number of jets} = 0,$$

with η_X being reconstructed as $\eta_X \equiv -\log\left(\tan\frac{\gamma}{2}\right)$. The isolation criteria are sensitive to the production mechanism of leptons. In contrast to leptons produced in the primary *ep* interaction, which are uncorrelated with other particles, leptons resulting from pion or kaon decays - and "fake" leptons - are typically near other hadrons, as hadrons resulting from the fragmentation of quarks and gluons are rarely isolated.

5.4 Mass Reconstruction

5.4.1 $M_T^{\ell\nu}$ and $M^{\ell\nu}$

Assuming that the missing momentum vector in an event is due to a single neutrino in the final state [1], the mass of the lepton-neutrino system may be reconstructed and used as an indication of whether or not events observed in the data are compatible with the production and decay of W bosons.

For events with detected scattered electrons ("tagged" events), the (massless) neutrino four-vector, and hence the *invariant mass* of the lepton neutrino system, may be reconstructed by employing the conservation laws for energy and momentum. The invariant mass, $M^{\ell\nu}$, is defined as

$$M^{\ell\nu} = \sqrt{(E_\ell + E_\nu)^2 - \left((P_x^\ell + P_x^\nu)^2 + \left(P_y^\ell + P_y^\nu\right)^2 + (P_z^\ell + P_z^\nu)^2\right)},$$

with the identification $P_x^\nu \equiv P_x^{miss}$ and $P_y^\nu \equiv P_y^{miss}$ (see equation 5.1). The energy and longitudinal momentum of the neutrino, determined by energy-momentum conservation, are:

$$\begin{aligned} E_\nu &= \frac{{P_T^{miss}}^2 + \delta^2_{miss}}{2\delta_{miss}}, \\ P_z^\nu &= \frac{{P_T^{miss}}^2 - \delta^2_{miss}}{2\delta_{miss}}. \end{aligned}$$

[1] No masses are reconstructed in events with isolated tau leptons, as there are at least two neutrinos (tau and anti-tau) in the final state of these events.

For events in which the scattered electron is not detected ("untagged" events) and carries away an undetermined contribution to $E - P_z$, the reconstruction of the neutrino vector is restricted to the transverse plane; in this case only the *transverse mass* of the lepton-neutrino system may be recontructed. The transverse mass, $M_T^{\ell\nu}$, is defined as

$$M_T^{\ell\nu} = \sqrt{\left(E_T^\ell + E_T^\nu\right)^2 - \left((P_x^\ell + P_x^\nu)^2 + \left(P_y^\ell + P_y^\nu\right)^2\right)},$$

where $E_T^\ell \equiv P_T^\ell$ and with the identification $E_T^\nu \equiv P_T^{miss}$.

5.4.2 $M^{\ell\nu X}$

For testing the hypothesis that the excess of isolated lepton events observed at high P_T^X may be due to the decay of a massive particle (like the top quark), a significant quantity is the invariant mass of the lepton-neutrino system plus the hadronic final state. If some of the events observed in the data result from decays of a heavy particle, a peak is expected in the three-body invariant mass.

The invariant mass of the lepton-neutrino-hadron system, $M^{\ell\nu X}$, is defined as

$$\begin{aligned} M^{\ell\nu X} = \big((E_\ell + E_\nu + E_X)^2 \\ - \left(\left(P_x^\ell + P_x^\nu + P_x^X\right)^2 + \left(P_y^\ell + P_y^\nu + P_y^X\right)^2 + \left(P_z^\ell + P_z^\nu + P_z^X\right)^2\right)\big)^{\frac{1}{2}}. \end{aligned}$$

For events with detected scattered electrons, the neutrino four-vector is reconstructed as for the invariant mass of the lepton-neutrino system, $M^{\ell\nu}$.

For events with undetected scattered electrons, the neutrino four-vector may be reconstructed by constraining the invariant mass of the lepton-neutrino system. As events with isolated leptons and missing transverse momentum are produced by decays of (intermediate) W bosons in many BSM processes (single top production and R-parity violating SUSY models, *cf.* section 2.7), the invariant mass of the lepton-neutrino system is constrained to the W mass in events with undetected scattered electrons. The solutions for the energy and longitudinal momentum of the neutrino determined by the W mass constraint are not unique, however. There exists one "forward" and one "backward" solution [10], defined by

$$\begin{aligned} E^\nu_{forward/backward} &= \frac{a \cdot E_\ell}{2{P_T^\ell}^2} \pm \sqrt{\left(\frac{a \cdot P_z^\ell}{2{P_T^\ell}^2}\right)^2 - \left(\frac{P_T^{miss} \cdot P_z^\ell}{P_T^\ell}\right)^2}, \\ P^\nu_{z\ forward/backward} &= \frac{a \cdot P_z^\ell}{2{P_T^\ell}^2} \pm \sqrt{\left(\frac{a \cdot E_\ell}{2{P_T^\ell}^2}\right)^2 - \left(\frac{P_T^{miss} \cdot E_\ell}{P_T^\ell}\right)^2}, \end{aligned}$$

where $a \equiv 2\left(P_x^\ell P_x^{miss} + P_y^\ell P_y^{miss}\right) + M_W^2$. Due to the limited experimental resolution, the numerical solutions for the neutrino energy and longitudinal momentum may be complex or correspond to negative $(E - P_z)_\nu$.

The forward and backward solutions are either both complex or both real, as complex solutions arise from a negative argument within the square-root. In the case of two complex solutions, the argument of the square-root is assumed to be negligibly small and its value is set to zero. In this case, the forward and backward solutions become real and coincide. In case of two real solutions, both are considered equally likely.

Not all solutions are physically possible, however. Solutions are rejected if they violate any of the following four conservation laws:

- $M^{\ell\nu X} < \sqrt{s}$
- $0 < (E - P_z)_\nu < 2E_e^{beam}$
- $(E - P_z)_\nu + \sum_i (E - P_z)_i < 75\ GeV$
- $E_\nu + \sum_i E_i < E_p^{beam} + E_e^{beam}$,

where the index i runs over all detected particles.

Each of the physical solutions for the neutrino four-vector then defines a different invariant mass $M^{\ell\nu X}$. Note that as a result of this procedure either zero, one or two physical solutions may be obtained for the invariant mass $M^{\ell\nu X}$ of the lepton-neutrino-hadron system in events with undetected electrons.

Chapter 6

Basic Event selection

6.1 Run and High Voltage Selection

The times in which data are taken at HERA are divided into *luminosity fills*, which are defined by one filling of electron and proton beams in the HERA ring. The luminosity fills are further subdivided into *luminosity runs*, shorter periods in which the experimental conditions at H1 do not vary significantly. The run and high voltage selection aims at selecting data-taking periods in which the experimental conditions are well understood and stable.

The experimental conditions are determined mainly by the working condition of the H1 experiment. To ensure a well understood detector response, the subdetectors that are essential to this analysis have to be operational, that is they have to be powered with the nominal high voltage and must be in read-out mode. The subdetectors that are considered essential comprise the central jet chambers CJC1/CJC2, the central proportional chambers CIP/COP, the LAr and SpaCal calorimeters, and the ToF and Luminosity systems. Occasionally the high voltage of the drift chambers may be reduced for part of a run, to protect the chamber from being damaged due to temporary high rates of beam-gas/beam-halo or synchrotron radiation backgrounds. In such runs, the events which are recorded at the reduced high voltage are rejected. In order to exclude data-taking periods with unstable experimental conditions, very short runs with luminosities $L_{run} < 0.1\ nb^{-1}$ are rejected. Additionally, a few runs corresponding to data-taking periods with noisy cells in the LAr or SpaCal calorimeters are also rejected.

The integrated luminosities resulting from this run selection for the different data-taking periods are shown in table 6.1. The quoted luminosities are calculated as described in reference [58]. In total, the analysis presented in this thesis is based on an integrated luminosity of 116.5 pb^{-1}. The stability of the experimental conditions over the whole HERA I data-taking period has been verified by an analysis of the *event yield*, the number of events observed per unit of integrated luminosity after application of the basic selection criteria described in this chapter and an additional $P_T^{calo} > 12\ GeV$ condition.

Run-Period	Luminosity
1994 - 97 e^+p $\left(E_p^{beam} = 820\ GeV\right)$	35.1 pb^{-1}
1998 - 99 e^-p $\left(E_p^{beam} = 920\ GeV\right)$	13.2 pb^{-1}
1999 - 2000 e^+p $\left(E_p^{beam} = 920\ GeV\right)$	68.3 pb^{-1}
1994 - 2000	116.5 pb^{-1}

Table 6.1: *Integrated luminosities for e^+p collisions at $\sqrt{s} = 301\ GeV$ and 318 GeV during the years 1994-97 and 1999/2000 and for e^-p collisions at $\sqrt{s} = 318\ GeV$ in the years 1998/99.*

6.2 Vertex Position

As a consequence of the spatial extension of the electron and proton bunches (see section 1.1), the interaction points of colliding electrons and protons are distributed around the nominal vertex position. In the beam direction and near the nominal interaction point, the distribution of interaction points is approximately Gaussian, with a spread of $\sigma_z \sim 10\ cm$ corresponding to the distribution of the particle density in the proton bunches [1]. In addition to the approximately Gaussian distribution near the nominal interaction point, that results from collisions of primary electron and proton bunches, the distribution of interactions points features two further maxima that result from the collisions of satellite bunches. The submaxima are displaced in the beam direction by about $\Delta z \sim \pm 150\ cm$ with respect to the nominal interaction point [47]. As the detector acceptance and resolution are optimized for *ep* collisions at the nominal interaction point, for an optimal measurement (and to reject non-*ep* background), an event vertex reconstructed near the nominal interaction point is required in most physics analyses.

In the analysis presented in this thesis, an event vertex determined by tracks in the central tracking detector ("central vertex") reconstructed within $-40\ cm < z_{vertex} < 100\ cm$ around the nominal position is required [2]. As in previous analyses [4,5], an assymmetric region enlarged in the forward direction is chosen for the z_{vertex} requirement. The choice of this assymmetric region is a consequence of studies on the vertex reconstruction [151]. These found that, in events with few central tracks, there was a shift between the reconstructed event vertex and the true interaction point. In particular in low multiplicity events, this produced a tail in the z_{vertex} distribution extending into the forward region [151].

6.3 Rejection of non-*ep* Background

In addition to the events produced by interactions of the electron and proton beams, events containing cosmic muons or originating from collisions of the proton beam with the beam-pipe wall or residual gas molecules in the beam-pipe are recorded by the H1 experiment. These *cosmic muon*, *beam-halo* and *beam-gas* events constitute

[1] The actual position of the Gaussian peak may be shifted by a few mm with respect to the nominal interaction point, due to variations in the HERA beam optics.

[2] The nominal position is defined as the run-period averaged position of the Gaussian peak in the z_{vertex} distribution and is determined from a sample of NC DIS events selected in the data.

the non-ep background to physics analyses at HERA and must be rejected. The three different types of non-ep background events are rejected on the basis of their characteristic signatures in the detector:

- Cosmic muons are produced in collisions of energetic cosmic particles with gas molecules in the earth's atmosphere. Correspondingly, cosmic muons penetrate the H1 detector typically with incident angles near $\theta \approx 90^\circ$, $|\phi| \approx 90^\circ$ and are in general not in time with ep collisions. The experimental signature of cosmic muon events is two isolated muon tracks in the instrumented iron, the liquid argon calorimeter and the central tracking detector "back-to-back" in polar and azimuthal angle.

- Beam-halo particles are produced in collisions of stray protons in the tails of the transverse beam profile [47] with the beam-pipe walls. The hadronic component of the resulting showers of particles is quickly absorbed, so that mainly muons are observed in beam-halo events in the H1 detector. The experimental signature of beam-halo events is a muon track travelling parallel to the beam direction, seen usually in the backward iron endcap, the liquid argon calorimeter and the forward iron endcap.

- Proton Beam-gas events originate from collisions of the proton beam with residual gas molecules in the beam-pipe. As a result of the high proton beam energy, the particles produced in beam-gas interactions are strongly boosted in the forward direction. The produced particles may be scattered out of the beam-pipe when they hit collimators that shield the detector from synchrotron radiation emitted by the electron beam [152]. The experimental signatures of proton beam-gas events are energy depositions mostly in the forward direction, many low p_T tracks isotropically distributed in azimuth and often an event vertex reconstructed near the collimators in the backward region of the H1 detector.

The three different types of non-ep background events may coincide or overlap with an *underlying ep* interaction (typically soft photoproduction) event. An example for a coincidence of a beam-halo and a beam-gas event is shown in figure 6.1 (top). In the same figure (bottom), an example of an overlayed *cosmic muon* and a soft photoproduction event is shown.

At HERA, the rate of non-ep background events is typically higher than the rate of ep interactions. In table 6.3, the rate of comic muon, beam-halo and beam-gas events is compared with the rate of different types of ep processes. As is shown in the table, the rates of rare high p_T processes (deep-inelastic NC and CC scattering and W production) are some orders of magnitude smaller then the non-ep background rates. In comparison to other collider experiments, the ratio of the beam-beam interaction rate to the rate of beam-induced backgrounds is especially unfavourable at HERA: in comparison to proton-proton colliders, the electron-proton scattering cross-section is small, due to the electro-weak interaction of the electron beam, while the beam-induced background is large in comparison to electron-electron colliders, due to the strong interactions of the proton beam with the beam-pipe wall and residual gas molecules in the beam-pipe [47]. As a consequence, special attention needs to be given to the suppression of this type of background at HERA.

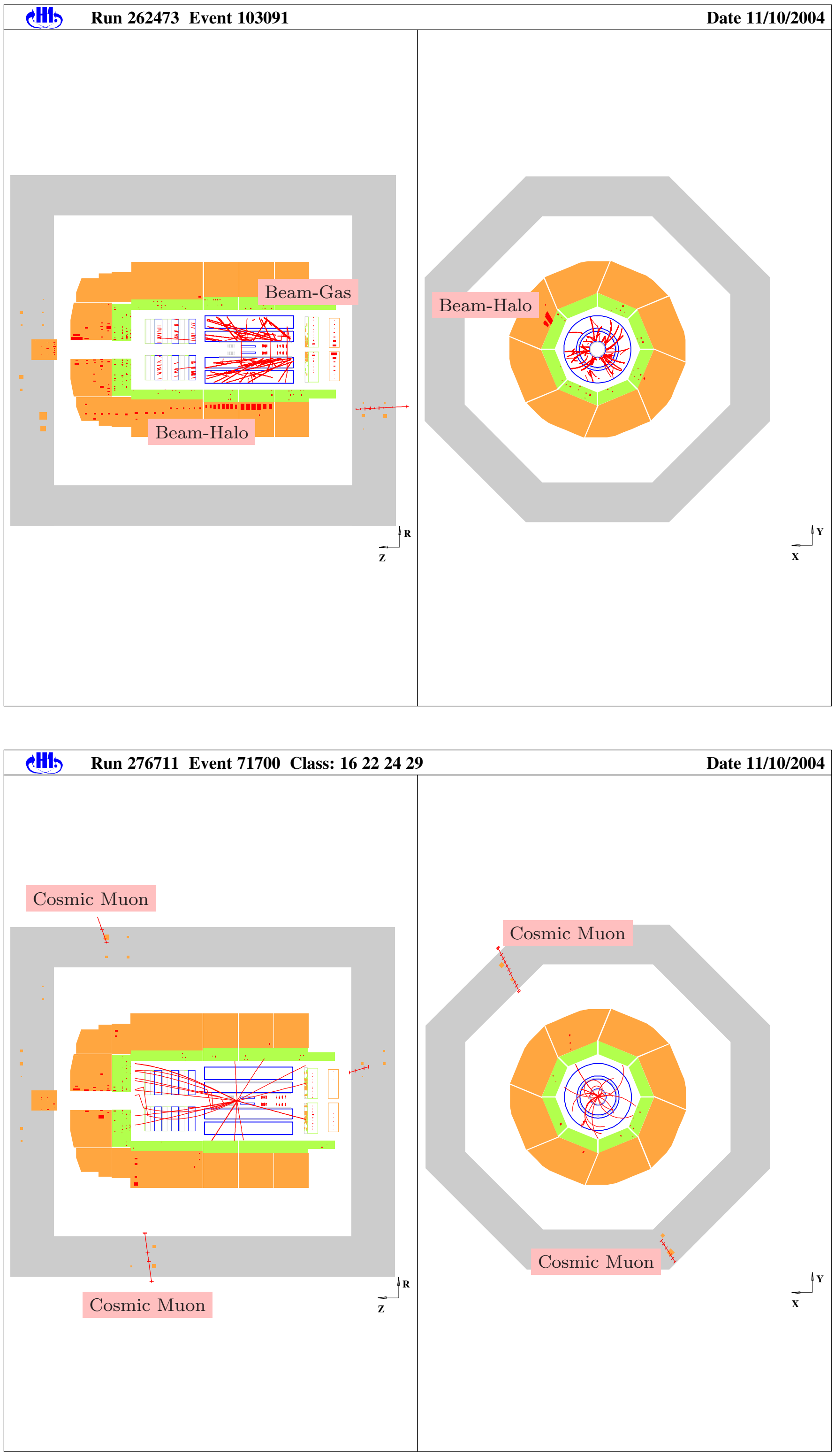


Figure 6.1: *Displays of non-ep background events. Top: Coincidence of a beam-halo and a beam-gas event. Bottom: Overlay event of a cosmic muon and a soft photoproduction event.*

Event Type	Rate
Cosmic Muons	$\mathcal{O}\left(10^3\ Hz\right)$
Beam-Halo	$\mathcal{O}\left(10^5\ Hz\right)$
Beam-Gas	$\mathcal{O}\left(10^3\ Hz\right)$
Photoproduction	$\mathcal{O}\left(10^3\ Hz\right)$
DIS (NC)	$\mathcal{O}\left(1\ Hz\right)$
DIS (CC)	$\mathcal{O}\left(10^{-3}\ Hz\right)$
W Production	$\mathcal{O}\left(10^{-5}\ Hz\right)$

Table 6.2: *Event rates expected at nominal HERA I running conditions for cosmic muon, beam-halo and beam-gas backgrounds in comparison to different ep processes (instantenous luminosity $\mathcal{L} = 2 \cdot 10^{31}\ cm^{-2}s^{-1}$, lifetime of the proton beam $\tau_p = 24\ h$, quality of the vacuum in the beam-pipe $p = 10^{-10}\ Torr$, and proton beam current $I_p = 100\ mA$) [47].*

In the H1 experiment, the non-*ep* background is suppressed by various measures. The vast majority of the cosmic muon, beam-halo and beam-gas background is already rejected "online" during data-taking; on the first trigger level by trigger elements provided by the ToF system and on the fourth trigger level by dedicated filter algorithms [51]. In "offline" data analyses, the non-*ep* background may be suppressed further by timing information provided by the central jet chambers and the liquid argon calorimeter [153] and by employing more complex filter algorithms, the *topological non-ep background finders* [154–156].

In the analysis presented here, the majority of cosmic muon, beam-halo and beam-gas induced background is rejected by a combination of the timing information provided by the CJC and the topological non-*ep* background finders, which are described in more detail in the following.

6.3.1 CJC Timing

After the track reconstruction, the time of traversal of charged particles in the central jet chamber may be reconstructed from the measured drift times to the anode wires. By comparison with the time of traversal expected for particles produced in *ep* collisions, an event timing relative to the nominal *ep* bunch crossing may be derived. The resolution of this t_0^{CJC} is approximately $\sigma_{t_0^{CJC}} \approx 2\ ns$ [31], but has significant non-Gaussian tails.

In the analysis presented in this thesis, events are rejected if their event timing is not within $\pm\Delta t_0^{CJC} = 9.6\ ns$ around the nominal bunch crossing time. The size of this time window is chosen so as to reject only a very small fraction of *ep* events. Most cosmic muon events are rejected by this requirement, as they traverse the CJC at times uncorrelated with the electron and proton beam crossings. The beam-halo and beam-gas background is not significantly reduced by the t_0^{CJC} requirement, as beam-halo muons traverse the detector nearly parallel to the anode wires of the central jet chamber, making the reconstruction of t_0^{CJC} difficult, while particles produced in beam-gas interactions traverse the CJC at similar times to those produced in *ep* interactions.

Flag	Bits
Ibg	$0, 1, 2, 3, 4, 5, 7, 8$
Ibgfm	$0, 1, 2, 4, 9, 10, 11, 16$
Ibgam	$0, 1, 2, 5$

Table 6.3: *Bits of the Ibg, Ibgfm and Ibgam flags that are considered for the rejection of cosmic muon, beam-halo and beam-gas background; the bit-coding of the flags is defined in reference [156].*

6.3.2 Non-*ep* Background Finders

The majority of the beam-halo, beam-gas and cosmic muon background may be suppressed by algorithms that reject the non-*ep* background on the basis of topological criteria ("non-*ep* background finders"), that is, associations of tracks and clusters in different subdetectors that are characteristic for beam-halo, beam-gas and cosmic muon events.

In the H1OO implementation, the results of individual finder algorithms are bit-coded in three different flags, *Ibg*, *Ibgfm* and *Ibgam* [156]. The non-*ep* background finders are applied to both data and simulated events. For simplicity, a common subset of finders is applied to events with isolated electrons, muons and tau leptons; those finders that misidentify more than about one per cent of simulated electronic, muonic or tauonic W decays as non-*ep* background are not applied. The bits of the *Ibg*, *Ibgfm* and *Ibgam* flags that are employed for non-*ep* background rejection in the analysis presented in this thesis are listed in table 6.3.

In addition to the flags listed in table 6.3, one extra finder algorithm is used, that has been developed to improve the suppression of the beam-gas background. The definition of this additional beam-gas finder is detailed in section E of the appendix, together with a display of an example event exclusively rejected by this finder. As can be seen in figure 6.2, the additional beam-gas finder significantly improves the suppression of the beam-gas background. In the figure, the number of events selected by the condition $P_T^{calo} > 12\ GeV$ in the 1999/2000 e^+p dataset per unit of integrated luminosity is shown after application of the basic selection criteria described in this chapter with and without employing the additional beam-gas finder.

6.4 Trigger selection

The triggering of events with isolated leptons and large missing transverse momentum is based on subtriggers derived from energy deposits in the LAr calorimeter. In order to increase the trigger efficiency for events with isolated muons (that do not deposit much energy in the calorimeter), the LAr triggers are complemented by muon triggers derived from hits in the instrumented iron.

The trigger efficiency is defined as

$$\varepsilon \equiv \frac{\substack{\text{Number of (weighted) events in sample} \\ \text{triggered by at least one of the subtriggers}}}{\text{Number of (weighted) events in sample}}$$

for a given event sample. Note that in the case of weighted events, the events in

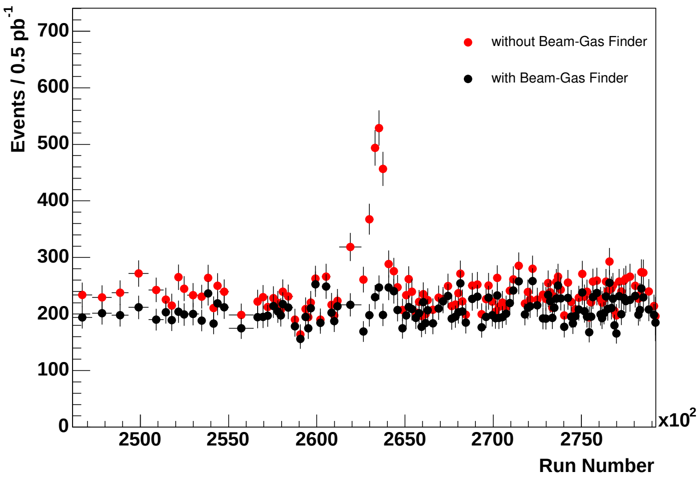


Figure 6.2: *Number of selected events per unit of integrated luminosity after application of the basic selection criteria described in this chapter and an additional $P_T^{calo} > 12$ GeV condition with and without using the extra finder algorithm for the suppression of beam-gas background. Before applying the additional beam-gas finder, an increased event yield is seen in the run range* 262000 - 265000*. The increase is interpreted as contribution of beam-gas background resulting from bad vacuum conditions in that run range. No increase in the event yield is seen anymore after applying the additional beam-gas finder, indicating that the majority of beam-gas background has been rejected by the algorithm.*

the sample may contribute to the numerator and denominator with fractional event counts, as is often the case for simulated Monte Carlo events.

In the following two sections, the LAr triggers and muon triggers employed for the triggering of events with isolated leptons and large missing transverse momentum in the HERA I data-taking period are described [1].

6.4.1 LAr Triggers

Most of the events considered in the analysis presented in this thesis are triggered by subtriggers based on the liquid argon calorimeter. The LAr triggers used are composed of two sets of subtriggers, the subtriggers $ST67$ and $ST75$, designed to trigger on the scattered electron in NC DIS events and the subtriggers $ST66$, $ST71$ and $ST77$, designed to trigger CC DIS events with large missing transverse momentum. The subtriggers $ST67$ and $ST75$ are sensitive to compact energy deposits of energy in the electromagnetic section of the LAr, as expected for the electromagnetic showers induced by the scattered beam electron. These triggers provide a very high efficiency for the triggering of events with isolated electrons. Additionally, they may trigger events with isolated muons or tau leptons and high energetic jets. The subtriggers $ST66$, $ST71$ and $ST77$ are sensitive to an imbalance of the energy deposits in the LAr calorimeter and efficiently trigger events with isolated electrons, muons or tau leptons and high P_T^{calo}.

The trigger efficiency of the LAr triggers is determined from the data, using a

[1] For future HERA II data-taking, a new subtrigger for high p_T muons has been developed. The new subtrigger is described in section D of the appendix.

sample of NC DIS events. For events with electrons in the LAr, the trigger efficiency of the subtriggers $ST67$ and $ST75$ is almost 100% [31]. The trigger efficiency of the combined set of subtriggers $ST66$, $ST67$, $ST71$, $ST75$ and $ST77$ for events without electrons in the LAr is determined as a function of P_T^{calo} and γ using the same sample of NC DIS events after removing the scattered electron that triggered the event: the *pseudo-charged current* (PsCC) method.

In the pseudo-charged current method, a sample of selected neutral current events is made equivalent to a charged current event sample, in order to allow the trigger and selection efficiencies for charged current events to be studied with the higher statistical precision provided by the neutral current sample. This is done by removing all evidence of the scattered electron from the NC data and reweighting the events by the ratio

$$w = \frac{d\sigma_{CC}\left(x, Q^2\right)/dx\, dQ^2}{d\sigma_{NC}\left(x, Q^2\right)/dx\, dQ^2}$$

so they give the same kinematic distributions as the charged current DIS events. The PsCC method relies on the similarity of the hadronic final state in neutral and charged current events; an assumption, that has been verified by independent studies [157]. The main advantage of this method is that there are about 100 times more events in the PsCC sample than in the CC sample.

The trigger efficiencies determined by the pseudo-charged current method are illustrated in figure 6.3. as a function of P_T^{calo} and γ. For all γ, the trigger efficiency rises with increasing P_T^{calo}, as expected. The decrease of the trigger efficiency seen at small γ mainly results from calorimeter cells near the beam-pipe in the forward direction that are not included in the energy sums, due to the high particle activity in that region. The decrease of the trigger efficiency seen at large γ results from the higher energy thresholds of the LAr triggers in this region, that are due to the higher noise levels of the calorimeter cells [31, 128]. Comparison of the trigger efficiencies determined by the PsCC method from the NC data with the simulated trigger efficiencies for CC Monte Carlo events shows that the trigger efficiencies of the combined set of subtriggers $ST66$, $ST67$, $ST71$, $ST75$ and $ST77$ are inaccurately modelled by the Monte Carlo. As a consequence, the trigger conditions cannot be applied to Monte Carlo events in the same way as they are applied to the data. Instead, the simulated events are weighted with the trigger efficiencies determined by the PsCC method from the NC data.

The weighting factors applied to the simulated events are determined by interpolating the P_T^{calo} and γ dependence of the trigger efficiency determined from the data. The P_T^{calo} dependence is parametrized by the functional form [3]

$$f_\gamma\left(P_T^{calo}\right) = A_\gamma - B_\gamma \cdot e^{C_\gamma \cdot P_T^{calo}}$$

and the coefficients A_γ, B_γ and C_γ are fitted independently in the different γ ranges shown in figure 6.3. In order to take into account minor changes in the trigger definitions in different data-taking periods, the trigger efficiency of the combined set of subtriggers $ST66$, $ST67$, $ST71$, $ST75$ and $ST77$ is fitted separately for the data-taking periods $1994-96e^+p$, $1997e^+p$, $1998/99e^-p$ and $1999/2000e^+p$. In the $1994-97e^+p$ data-taking period, the trigger efficiency is fitted separately for $1994-96$ and 1997, because the subtrigger $ST71$ was not implemented until the end

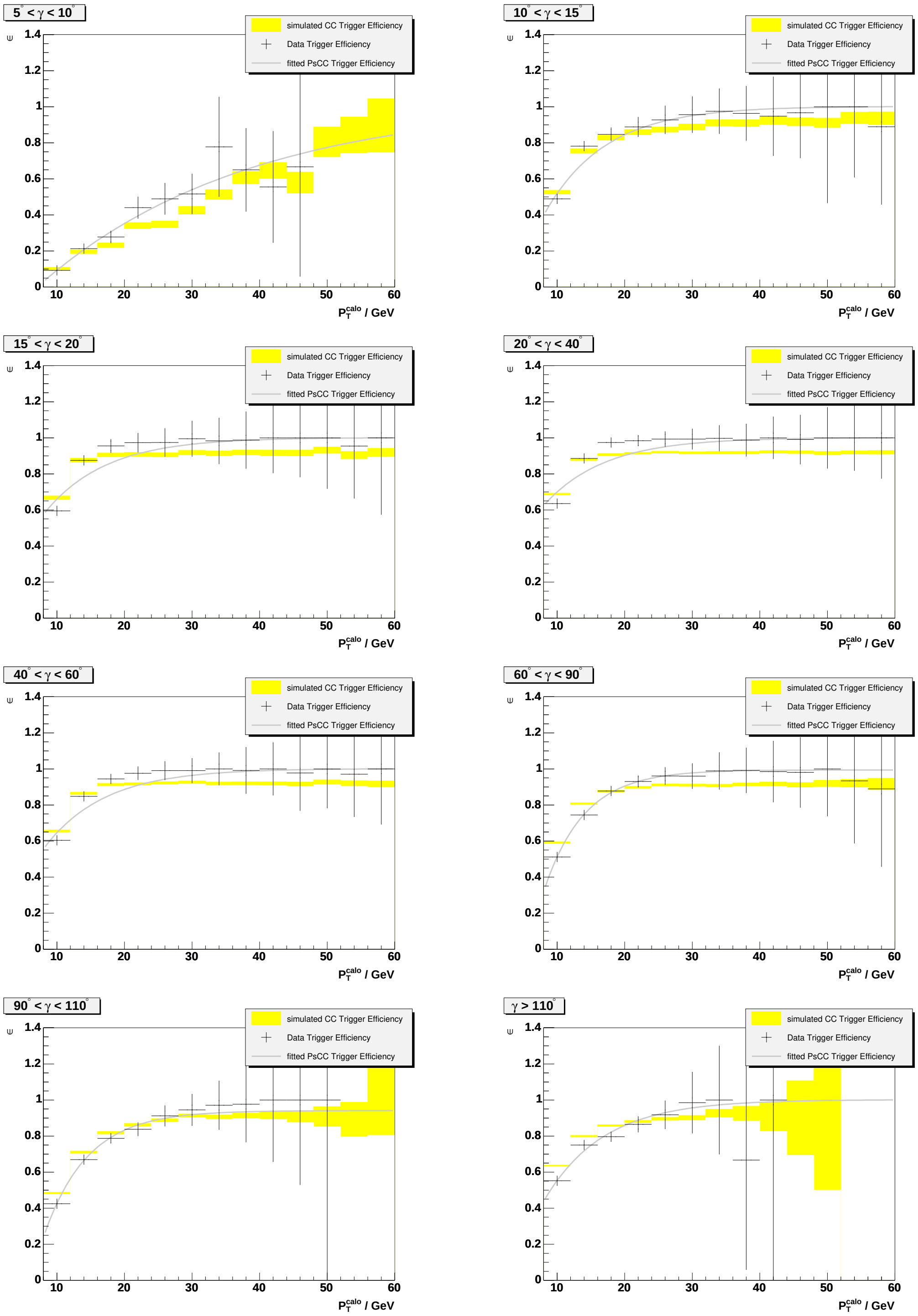


Figure 6.3: *The trigger efficiency of the combined subtriggers ST66, ST67, ST71, ST75 and ST77 for events with large missing transverse momentum. The trigger efficiencies indicated by points have been determined with the PsCC method from NC DIS events selected in the 1999/2000 e^+p dataset. The solid curves represent an interpolation of the trigger efficiencies as function of P_T^{calo} for different ranges of γ. For comparison, the trigger efficiencies predicted by the Monte Carlo simulation are indicated by the shaded areas; the differences between the upper and lower limits represent the statistical uncertainties.*

of 1996. The trigger efficiencies determined from the data are well approximated by the interpolation functions, as may be seen in figure 6.3.

In order to correctly model the combined efficiency of the LAr and muon triggers in the Monte Carlo, the weighting factors are only applied to simulated events that are neither triggered by electrons in the LAr nor by muon triggers.

6.4.2 Muon Triggers

In addition to the LAr triggers described in the last section, a set of muon triggers is employed to improve the trigger efficiency for events with isolated muons, especially at low P_T^X [2]. In the analysis presented in this thesis, the muon triggers used are the subtriggers $ST15$, $ST18$, $ST19$, $ST22$, $ST34$ and $ST56$, that are designed to trigger muonic J/Ψ decays and muon pair production events. With the exception of $ST15$, the muon triggers are described in detail in [97]. A detailed description of $ST15$ may be found in reference [158]. In contrast to the LAr triggers, the trigger efficiency of the muon triggers is found to be well described by the Monte Carlo simulation [97], so that the muon trigger conditions are applied in the same way to both data and simulated events.

The trigger efficiency of the muon triggers, estimated using an event sample of simulated muonic W decays passing the selection criteria listed in table 7.4, varies between $\varepsilon_{forward} \approx 13\%$ in the forward region and $\varepsilon_{central} \approx 52\%$ in the central region; averaged over the solid angle, the trigger efficiency of the muon triggers is estimated to be $\varepsilon \approx 31\%$. In some data-taking periods with high trigger rates, some of the muon triggers have been prescaled. The prescales that were applied to the muon triggers during data-taking are modelled for simulated events by appropriate weighting factors, determined as described in references [58, 159].

[2]As has been described in section 5.1.1, $P_T^{calo} \sim P_T^X$ in most events with isolated muons, so the LAr triggers are very inefficient at low transverse momenta of the hadronic final state.

Chapter 7

Final Event Selection

The sections in this chapter describe the criteria for the final selection of events with isolated electrons, muons and tau leptons. In each section, the selection criteria used to suppress the background are motivated by studying the dominant background contributions to event samples in an enlarged phase-space. The selection criteria defining the final event selection are then described in detail. These need to be carefully defined - as the cross-sections of some background processes exceed that of the expected signal by a few orders of magnitude: the selection criteria need to reject a very large fraction of the background and simultaneously maintain a high selection efficiency for the rare signal. The selection criteria are chosen as generically as possible. In this way, it is guaranteed that the search for events with isolated leptons and large missing transverse momentum is model independent and sensitive not only to the Standard Model processes producing such events, but to possible signals of beyond the Standard Model processes, too. At the end of each section, the efficiency of the chosen selection criteria for signal processes is estimated representatively using simulated samples of leptonic decays of real W bosons, the dominant contribution to events with isolated leptons and large missing transverse momentum within the Standard Model. The selection efficiency for signal processes other than Standard Model W production may be approximated by this estimate.

7.1 Events with isolated Electrons

7.1.1 Definition of the Phase-Space

The enlarged phase-space for events with isolated electrons is defined by requiring an electron candidate of transverse momentum $p_T^e > 10\ GeV$ passing the selection criteria listed in table 4.1 in the polar angle range $5^\circ < \theta_e < 140^\circ$ and a reconstructed imbalance $P_T^{calo} > 12\ GeV$ of the energy deposits in the calorimeter. In addition, the events are required to pass the basic selection criteria discussed in chapter 6. In the enlarged phase-space selection, the electron candidates are not required to be isolated, so as to to increase the statistics of the background samples and show the effect of the isolation criteria on the background suppression. The selection criteria defining the enlarged phase-space for events with isolated electrons are listed in table 7.1.

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-*ep* background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 7, 8$ *Ibgfm* bits $0, 1, 2, 4, 9, 10, 11, 16$ *Ibgam* bits $0, 1, 2, 3, 4, 5, 7, 8$
Electron	$p_T^e > 10$ GeV $5° < \theta_e < 140°$ "LAr electron" linked to track in CTD $\parallel$ FTD
Additional requirements	$P_T^{calo} > 12$ GeV

Table 7.1: *Selection criteria defining the enlarged phase-space for events with isolated electrons. "LAr electrons" are defined as passing the identification criteria listed in table 4.1.*

7.1.2 Background Studies

In the phase-space defined by the selection criteria listed in table 7.1, 3318 events are selected in the 1994-2000 e^+p and e^-p datasets, which - within the systematic uncertainties discussed in section 8.1 - compares well with an expectation of 2818 events from the Monte Carlo simulation. For the selected event sample, distributions of the quantities defined in chapter 5 are shown in figures 7.1, 7.2, 7.3, and 7.4.

The selected event sample is dominated by the contribution of NC DIS processes. In NC DIS events, the electron identified in the detector is the scattered beam electron and the reconstructed imbalance of transverse energy in the calorimeter is due to a mismeasurement of an intrinsically balanced event in the detector. As expected for scattered beam electrons, the polar angle distribution of the electrons in NC DIS events is seen to peak in the backward direction. The different kinematic distributions of the electrons resulting from decays of heavy particles and the scattered beam electrons can be seen most clearly in the distribution of the quantity ζ^2: for NC DIS events, ζ^2 corresponds to the magnitude of the four-momentum transfer Q^2 and as the neutral current cross-section decreases with $1/Q^4$ (*cf.* section 2.2.3), the ζ^2 spectrum falls steeply. In contrast, the ζ^2 spectrum reconstructed for electrons resulting from decays of heavy particles is approximately constant. Another characteristic of NC DIS events is that only particles going down the beam-pipe in the proton direction are undetected. This allows for the reconstruction of the missing longitudinal momentum δ_{miss}, the distribution of which is seen to peak at zero for NC DIS events. As the scattered beam electron balances the hadronic final state in the transverse plane, the distribution of the acoplanarity angle $\Delta\phi_{e-X}$ is very strongly peaked in the "back-to-back" configuration. The balance between the scattered beam electron and the hadronic final state can be seen also in the distribution of the topological variable V_{ap}/V_p that describes the angular distribution of the energy deposits in the calorimeter: in contrast to events with undetected high

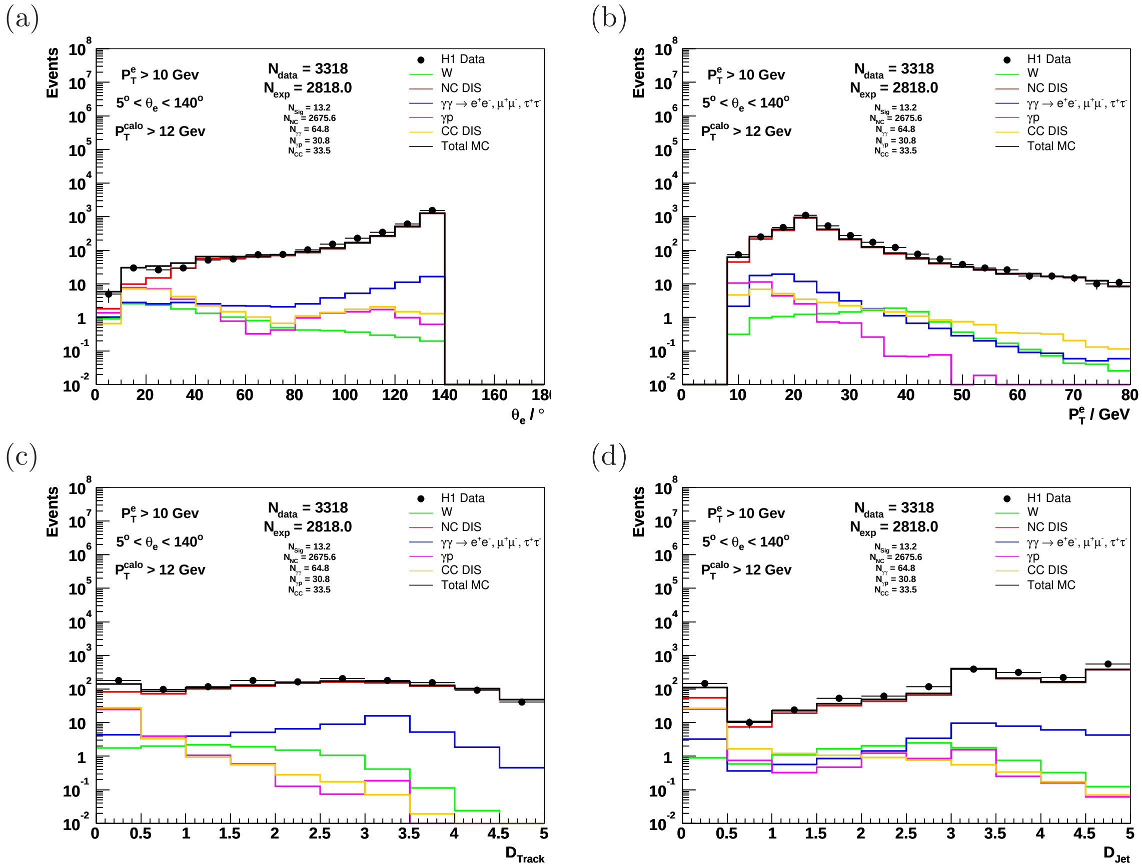


Figure 7.1: *Distributions of the polar angle (a), the transverse momentum (b), the distance to the nearest track (c) and the distance to the nearest jet (d) of the electron candidates in events in the enlarged phase-space defined by the criteria listed in table 7.1.*

p_T particles in the final state, the distribution of V_{ap}/V_p does not peak at low ratios for NC DIS events.

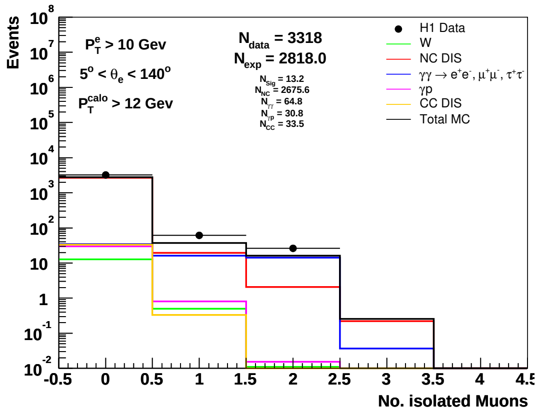


Figure 7.2: *Distribution of the number of isolated muons in events in the enlarged phase-space defined by the criteria listed in table 7.1. The criteria for muons to be considered as "isolated" are defined in section 4.4.*

The contribution of lepton pair production processes is much smaller in size than that of NC DIS processes. The events contributing to the selected event sample are similar in lepton pair production and NC DIS processes, however: in most of the selected lepton pair production events, the electron identified in the detector is the scattered beam electron. Perhaps surprisingly, the contribution from muon pair production processes is actually larger than that from electron pair production processes. This may be seen in the distribution of the number of isolated muons in the selected events. The relatively large contribution of muon pair production processes results from events, in which the scattered electron is balanced by muons of which at least one is produced with high transverse momentum. As muons deposit only little energy in the calorimeter, a significant imbalance of the energy deposits in the calorimeter may be reconstructed in this case. The similarity between the events produced by lepton pair production and NC DIS processes can be seen in the distribution of the electron polar angle, the quantity ζ^2, the missing longitudinal momentum and the acoplanarity angle. The distribution of the topological variable V_{ap}/V_p is shifted towards lower ratios for lepton pair production events. The difference results from events in which the scattered electron is balanced by muons of high transverse momentum in the final state. In this case, the vectorial sum of energy deposits in the calorimeter typically points to the hemisphere with the scattered electron and V_p is of the order of p_T^e, while V_{ap} is of the order of the typical energy deposit of muons in the calorimeter, i.e. $\mathcal{O}\,(3\ GeV)$.

The contribution of photoproduction processes arises from events in which the electron candidate is either a misidentified hadron or produced in a heavy quark decay and the reconstructed imbalance of energy deposits in the calorimeter is due to mismeasurement. In typical photoproduction events, two jets are produced "back-to-back" in the transverse plane. The electron candidate is typically within one of these jets and seldom isolated, as parton showering and hadronisation processes typically produce a couple of hadrons near the electron candidate in each jet. This can be seen in the distributions of the distances D_{track} and D_{jet} of the electron candidate to the nearest charged particle and the nearest jet, both of which peak at small values. The closeness of the electron candidate to one of the two jets also explains the peak of the acoplanarity angle distribution near 180 degrees. Overall, the contribution of photoproduction processes to the selected event sample is small, as photoproduction events are suppressed both by the electron selection requirements and by the demanded imbalance of energy deposits in the calorimeter.

The contribution of CC DIS processes is similar in size to that of photoproduction processes, although the charged current cross-section is much smaller. In the selected CC DIS events, the electron candidate is either due to a misidentified hadron or the decay of a heavy quark as in photoproduction events, but the reconstructed imbalance of the energy deposits in the calorimeter is due to an undetected neutrino in the final state. The neutrino balances the hadronic final state in azimuth, so that in charged current events all detectable particle are typically concentrated in one hemisphere with the detector, a characteristic which can best be seen in the peak of the V_{ap}/V_p distribution at very low values. The electron candidate is typically within the hadronic final state, as may be seen directly in the distribution of the azimuthal distance $\Delta\phi_{e-X}$ between the electron candidate and the hadronic final state, that peaks at small values. As expected for electron candidates within the

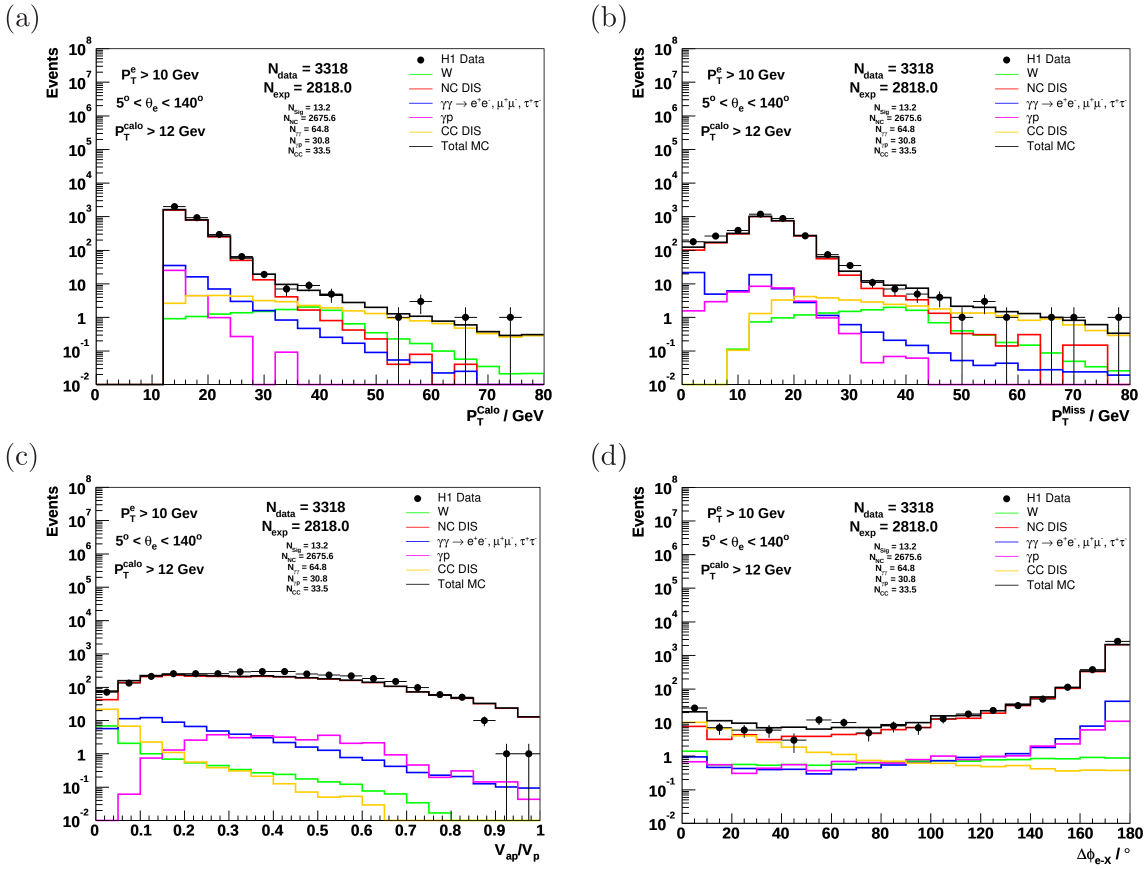


Figure 7.3: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the ratio V_{ap}/V_p (c) and the acoplanarity angle between the electron candidate and the hadronic final state (d) in events in the enlarged phase-space defined by the criteria listed in table 7.1.*

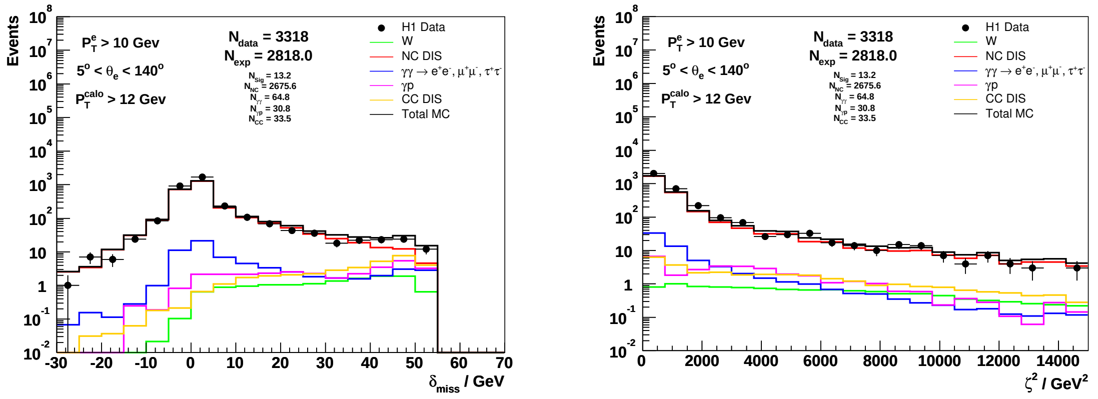


Figure 7.4: *Distributions of the missing longitudinal momentum (left) and the quantity ζ^2 (right) in events in the enlarged phase-space defined by the criteria listed in table 7.1.*

hadronic final state, the distributions of D_{track} and D_{jet} also peak both at small values for CC DIS events.

7.1.3 Definition of the Final Selection Criteria

The final selection of events with isolated electrons is based on the enlarged phase-space defined by the selection criteria listed in table 7.1. In addition to passing the selection criteria for the enlarged phase-space, events in the final sample are required to contain an isolated electron, have large missing transverse momentum, and pass some additional criteria designed to suppress the dominating NC DIS background [1].

The requirement that the electron candidates be isolated effectively suppresses the background contribution from events in which the electron candidate is due to a misidentified hadron or the decay of a heavy quark. For electron candidates to be considered as isolated, they are required to be flagged as "isolated lepton" by the electron identification algorithm (see section 4.3) and be separated from other charged particles and jets by at least $D_{track} > 0.5$ and $D_{jet} > 1.0$, respectively [2]. The isolation requirements effectively suppress the background contribution from photoproduction and CC DIS processes.

The large missing transverse momentum requirement does not suppress much background, as, in events with isolated electrons, the missing transverse momentum P_T^{miss} is similar to the imbalance P_T^{calo} of the energy deposits in the calorimeter. For events to be considered as having a large missing transverse momentum, $P_T^{miss} > 12\ GeV$ is required in the final selection. In order to suppress intrinsically balanced events in which the reconstructed missing transverse momentum is due to a mismeasurement, a significant acoplanarity angle $\Delta\phi_{e-X} < 160°$ is required between the isolated electron and the hadronic final state. The background of intrinsically balanced events is further suppressed by a requirement on the topological variable V_{ap}/V_p. In order not to reject too much signal, the V_{ap}/V_p requirement is formulated as a two-dimensional cut. The correlation between V_{ap}/V_p and the transverse momentum of the isolated electron is shown in figure 7.5, separately for signal and background processes. As can be seen in the figure, the electrons (mainly scattered beam electrons) in NC DIS and lepton pair production processes are typically of significantly lower transverse momentum than the electrons produced in the decays of heavy particles (represented by leptonic W decays [3]). A large fraction of the events in which the reconstructed missing transverse momentum is due to a mismeasure-

[1] As in the enlarged phase-space selection, the electron candidates are required to have transverse momenta above $p_T^e > 10\ GeV$ and be in the polar angle range $5° < \theta_e < 140°$ in the final selection. The backward region is excluded, because of the increase of the NC DIS cross-section at small scattering angles and also because the BBE wheel in the backward region is less deep than the other wheels of the LAr calorimeter (*cf.* figure 1.9). This increases the probability for electromagnetic showers to leak out out of the calorimeter, in which case the energy of the electron inducing the shower is likely to be underestimated and missing transverse momentum to be reconstructed in intrinsically balanced events. Neither the required transverse momentum nor the exclusion of the backward region reduces the acceptance for events with electrons resulting from decays of heavy particles: the transverse momenta of the electrons produced in the decays are typically of the order of half the mass m of the decayed heavy particle (see section 2.6.1), and as Bjorken $x > m^2/s$ needs to be large in order to produce such heavy particles, $x \cdot E_p^{beam} \gg E_e^{beam}$, so that the produced heavy particle (and its decay products) are significantly Lorentz boosted in the forward direction.

[2] The $D_{track} > 0.5$ requirement is not applied for electron candidates at small polar angles $\theta_e < 45°$, due to the problem with showering in the dead material in the forward region.

[3] The small contribution from decays of Z bosons into neutrinos (*cf.* section 2.6.2) is included in the "W" expectation.

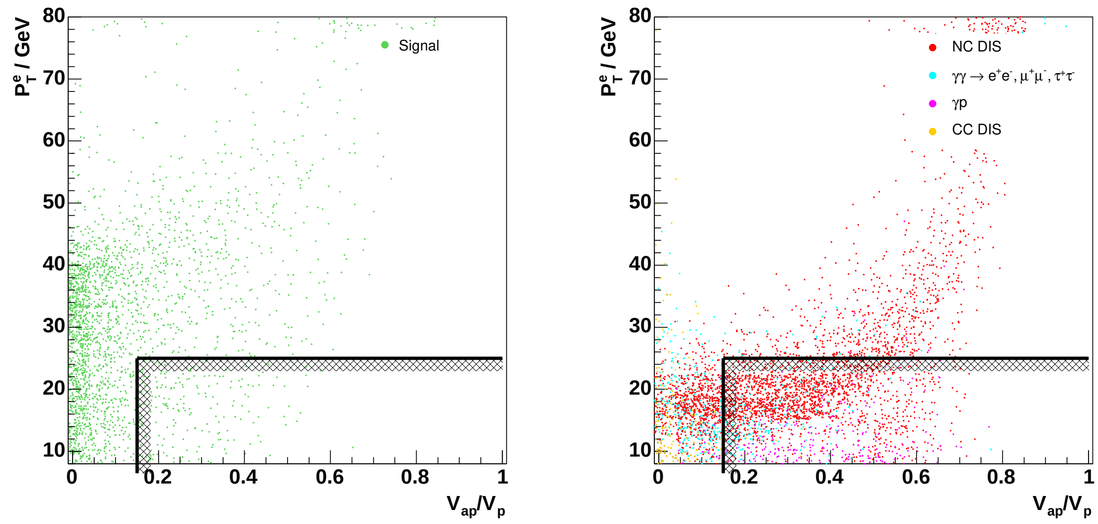


Figure 7.5: *Correlation between the ratio V_{ap}/V_p and the transverse momentum p_T^e of the isolated electron for simulated W production events (left) and events produced by background processes (right). The two-dimensional cut applied in the final selection of events with isolated electrons is depicted by the hatched line in the lower right corner.*

ment (NC, γp, lepton pair production) is rejected by requiring $V_{ap}/V_p < 0.15$ for $P_T^e < 25\ GeV$ and $V_{ap}/V_p < 0.5$ for $P_T^e > 25\ GeV$, while most events with undetected high p_T particles in the final state (W production, but also CC) pass these requirements. The combination of the requirements on P_T^{miss}, $\Delta\phi_{e-X}$ and V_{ap}/V_p effectively suppresses the "back-to-back" topologies characteristic of the NC DIS, photoproduction and lepton pair production backgrounds.

The large background contribution from NC DIS processes needs to be further suppressed. The majority of NC DIS events can be rejected by requirements on the quantity ζ^2, that corresponds to the magnitude of the four-momentum transfer Q^2 in NC DIS events, and the missing longitudinal momentum δ_{miss}. In order not to reject too much signal, the requirement on ζ^2 is formulated as a two-dimensional cut in combination with the imbalance of energy deposits in the calorimeter. The correlation between ζ^2 and P_T^{calo} is shown in figure 7.6, separately for signal and background processes. In the figure, NC DIS events are seen to dominate at low values of ζ^2 and P_T^{calo}. The signal events (represented by leptonic W decays) are seen to be evenly distributed in ζ^2 and concentrated around significantly higher values of P_T^{calo}. A large fraction of NC DIS events is rejected by requiring $\zeta^2 > 5000\ GeV^2$ ($\zeta^2 > 500\ GeV^2$) for $P_T^{calo} < 25\ GeV$ ($P_T^{calo} > 25\ GeV$). The NC DIS background is suppressed further by requiring $\delta_{miss} > 5\ GeV$ for events in which only one electron candidate is detected that has the same charge as the beam electron. The requirements on ζ^2 and δ_{miss} not only suppress the dominant NC DIS background, but also lepton pair production processes in which the isolated electron identified in the detector is the scattered beam electron.

The final selection criteria for events with isolated electrons and large missing transverse momentum are listed in table 7.2.

In addition to passing the selection criteria described so far, it is required that events in the final sample do not contain isolated muons. This requirement is intro-

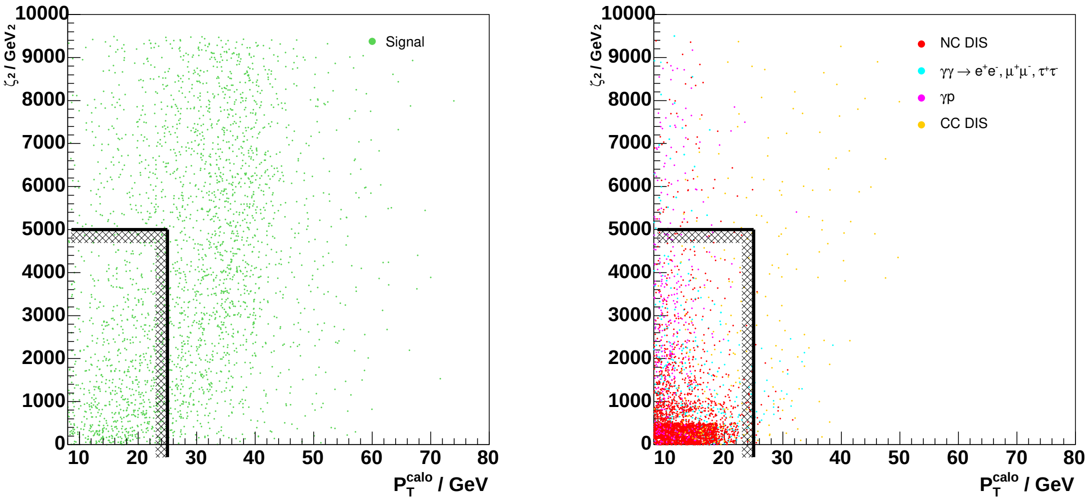


Figure 7.6: *Correlation between the quantity ζ^2, that corresponds to the magnitude of the four-momentum transfer Q^2 in NC DIS events, and the imbalance of energy deposits in the calorimeter, P_T^{calo}, for simulated W production events (left) and events produced by background processes (right). The two-dimensional cut applied in the final selection of events with isolated electrons is depicted by the hatched line in the lower left corner.*

duced to ensure events contribute to only one of the isolated electron, muon or tau samples [3]. It also reduces the background contribution from muon pair production processes that is difficult to suppress by other means because of its high P_T^{calo} and low V_{ap}/V_p values.

7.1.4 Selection Efficiency

The efficiency ε of the selection criteria listed in table 7.2 for events with isolated electrons and large missing transverse momentum is estimated as a function of the generated hadronic transverse momentum P_T^X. For each bin of the P_T^X distribution it is defined by the ratio

$$\varepsilon = \frac{\text{Number of (weighted) events in bin passing selection}}{\text{Number of (weighted) events in bin}}.$$

The estimated selection efficiency for Standard Model W production is shown in figure 7.7. Also shown in the figure is the estimated efficiency for decays of single top quarks produced by flavour-changing neutral current interactions, as an example for possible signal processes beyond the Standard Model. The selection efficiency for Standard Model W production is about 40% and is almost independent of P_T^X. For single top production, an efficiency of about 50% is estimated. The higher efficiency is mainly due to the smaller polar angles of the electrons (and hence higher ζ^2) and larger transverse momenta of the neutrinos produced in top quark decays, causing fewer single top events to fail the two-dimensional $P_T^{calo} > 25 || \zeta^2 > 5000\ GeV^2$ requirement than Standard Model W production events.

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-ep background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 7, 8$ *Ibgfm* bits $0, 1, 2, 4, 9, 10, 11, 16$ *Ibgam* bits $0, 1, 2, 3, 4, 5, 7, 8$
Electron	$p_T^e > 10\ GeV$ $5° < \theta_e < 140°$ "LAr electron" flagged as "isolated lepton" and linked to track in CTD $\parallel$ FTD $D_{track} > 0.5$ for $\theta_e > 45°$ $D_{jet} > 1.0$
Missing momentum	$P_T^{miss} > 12\ GeV$ $\Delta\phi_{e-X} < 160°$ $V_{ap}/V_p < 0.5$ (< 0.15 for $P_T^e < 25\ GeV$)
NC DIS suppression	$\zeta^2 > 500\ GeV^2$ $\big(> 5000\ GeV^2$ for $P_T^{calo} < 25\ GeV\big)$ $\delta_{miss} > 5\ GeV$[1]
Additional requirements	$P_T^{calo} > 12\ GeV$ No isolated Muon

[1] if only one electron candidate is detected, which has the same charge as the beam electron

Table 7.2: *Selection criteria for the final sample of events with isolated electrons. "LAr electrons" are defined as passing the identification criteria listed in table 4.1. The "isolated lepton" flag is defined in section 4.3.*

7.2 Events with Isolated Muons

The selection of events with isolated muons is based on two separate sets of selection criteria. The two sets are adapted to specific regions of the phase-space: the events with high hadronic transverse momenta that are triggered by the LAr triggers described in section 6.4.1 and the events with isolated muons in the central region that are triggered by the muon triggers described in section 6.4.2. In the following two subsections, the selection of events with isolated muons is described separately for each of the two phase-space regions.

7.2.1 LAr triggered Phase-Space

Definition of the Phase-Space

The enlarged LAr triggered phase-space for events with isolated muons is defined by requiring a muon candidate of transverse momentum $p_T^{\mu} > 10\ GeV$ passing the selection criteria listed in table 4.2 in the polar angle range $5° < \theta_{\mu} < 140°$ and a

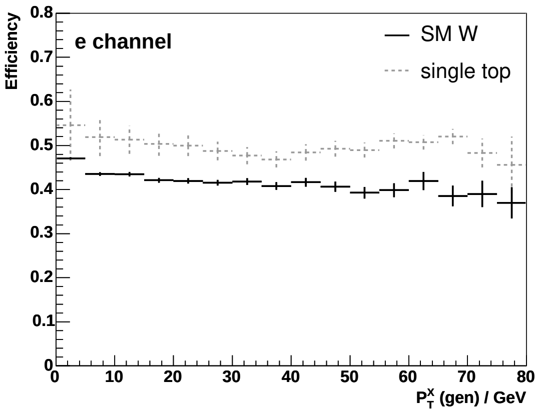


Figure 7.7: *Efficiency of the selection criteria listed in table 7.2 for events with isolated electrons and large missing transverse momentum. The selection efficiency for Standard Model signal processes is represented by the solid line; the dashed line represents the efficiency for single top production as an example process beyond the Standard Model.*

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST15 \parallel ST18 \parallel ST19 \parallel ST22 \parallel ST34 \parallel ST56 \parallel$ $ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-*ep* background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 7, 8$ *Ibgfm* bits $0, 1, 2, 4, 9, 10, 11, 16$ *Ibgam* bits $0, 1, 2, 3, 4, 5, 7, 8$
Muon	$p_T^{\mu} > 10\ GeV$ $5° < \theta_{\mu} < 140°$ "Iron muon"
Additional Requirements	$P_T^{calo} > 12\ GeV$

Table 7.3: *Selection criteria defining the enlarged LAr triggered phase-space for events with isolated muons. "Iron muons" are defined as passing the identification criteria listed in table 4.2.*

reconstructed imbalance $P_T^{calo} > 12\ GeV$ of the energy deposits in the calorimeter. Additionally, the events are required to pass the basic selection criteria discussed in chapter 6. In the enlarged phase-space selection, the muon candidates are not required to be isolated, in order to increase the statistics of the background samples and show the effect of the isolation criteria on the background suppression. The selection criteria defining the enlarged LAr triggered phase-space for events with isolated muons are listed in table 7.3.

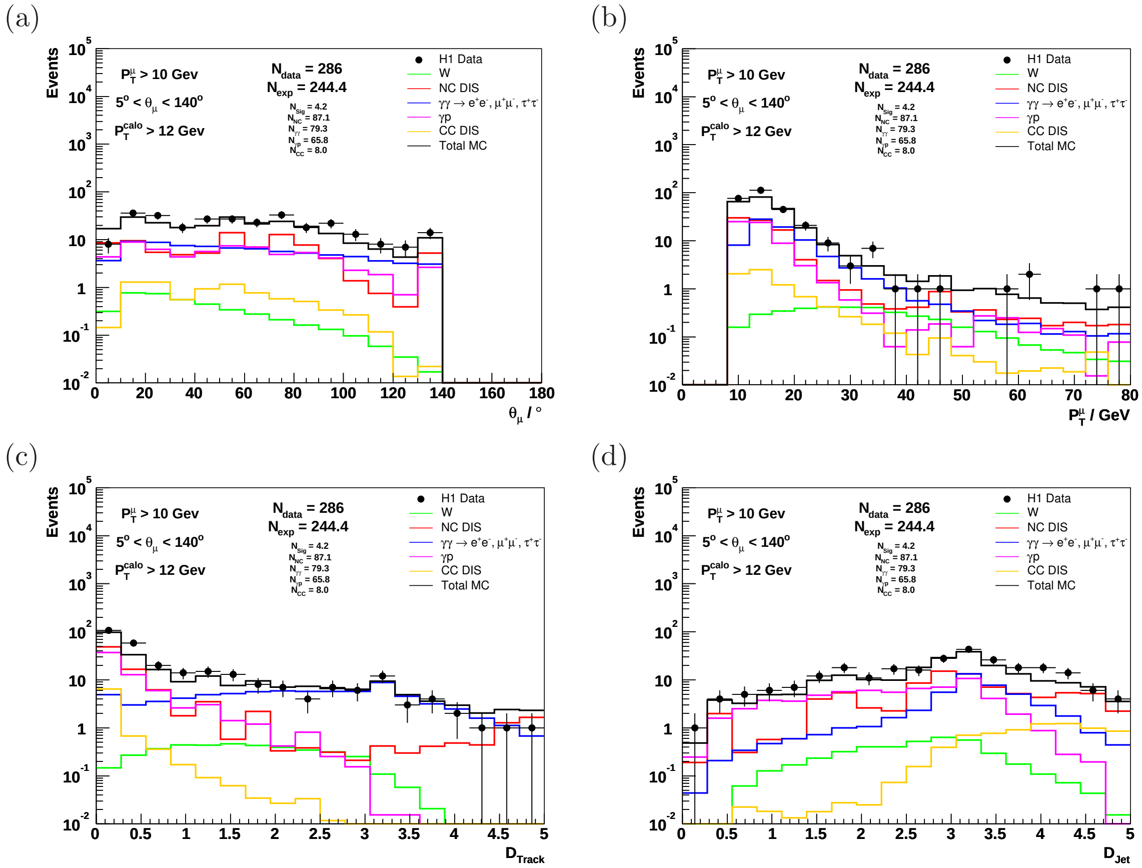


Figure 7.8: *Distributions of the polar angle (a), the transverse momentum (b), the distance to the nearest track (c) and the distance to the nearest jet (d) of the muon candidates in events in the enlarged LAr triggered phase-space defined by the criteria listed in table 7.3.*

Background Studies

In the phase-space defined by the selection criteria listed in table 7.3, 286 events are selected in the 1994-2000 e^+p and e^-p datasets, which - within the systematic uncertainties discussed in section 8.1 - compares well with an expectation of 244 events from the Monte Carlo simulation. In the selected event sample, the contributions from NC DIS, lepton pair production and photoproduction processes are of approximately equal size. Distributions of the quantities defined in chapter 5 are shown in figures 7.8, 7.9, and 7.10. Overall, the contribution of background processes to events with isolated muons is significantly smaller than to events with isolated electrons, as there is no contribution equivalent to that of the scattered beam electrons in events with isolated muons.

The contributions from NC DIS and photoproduction processes are characterised by very similar event topologies. In both cases, the muon candidate identified in the detector is due to a misidentified hadron or produced in the decay of a heavy quark and the reconstructed imbalance of energy deposits in the calorimeter is due to a mismeasurement of an intrinsically balanced event in the detector: in the distribution of the acoplanarity angle $\Delta\phi_{\mu-X}$ and of the topological variable V_{ap}/V_p it does not make any difference whether the transverse momentum balance is between the scattered beam electron and the hadronic final state as in NC DIS events or between two jets as in typical photoproduction events. For both NC DIS and photoproduction events, the distribution of the acoplanarity angle $\Delta\phi_{\mu-X}$ peaks at the "back-to-back" configuration and the distribution of the ratio V_{ap}/V_p at high values, in contrast to events with undetected particles of high transverse momentum in the final state. The muon candidates in NC DIS and photoproduction events are typically within jets and seldom isolated, as parton showering and hadronisation processes typically produce a couple of hadrons near the muon candidate in each jet. This can be seen in the distributions of the distances D_{track} and D_{jet} of the muon candidate to the nearest charged particle and the nearest jet, both of which peak at small values.

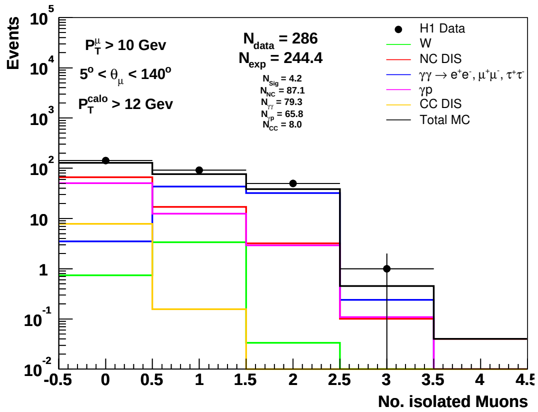


Figure 7.9: *Distribution of the number of isolated muons in events in the enlarged LAr triggered phase-space defined by the criteria listed in table 7.3. The criteria for muons to be considered as "isolated" are defined in section 4.4.*

Perhaps surprisingly, the contribution from photoproduction processes to events with muons is larger than to events with electrons, although the probability for hadrons to be misidentified as muons is smaller than the probability for them to be misidentified as electrons (*cf.* sections 4.3 and 4.4). The reason for the higher probability for photoproduction processes to contribute to event samples with muons is a correlation between the misidentification of a hadron as a muon and the reconstruction of an imbalance of transverse energy in the calorimeter: if a highly energetic hadron does not lose all its energy in the calorimeter and reaches the muon system, the difference between the energy of the hadron and its energy loss in the calorimeter is reconstructed as an imbalance of the energy deposits in the calorimeter. This effect can be seen in the distribution of the reconstructed missing transverse momentum P_T^{miss}: as the four-vector of muon candidates is reconstructed from their track in the drift chambers, the energy of misidentified hadrons is likely to be reconstructed correctly and the reconstructed missing transverse momentum likely to be small.

The contribution from lepton pair production processes arises mostly from muon pair production events in which the scattered beam electron is balanced by muons of high transverse momentum in the final state. In these events, the reconstructed imbalance of the energy deposits in the calorimeter is due to the electron and the

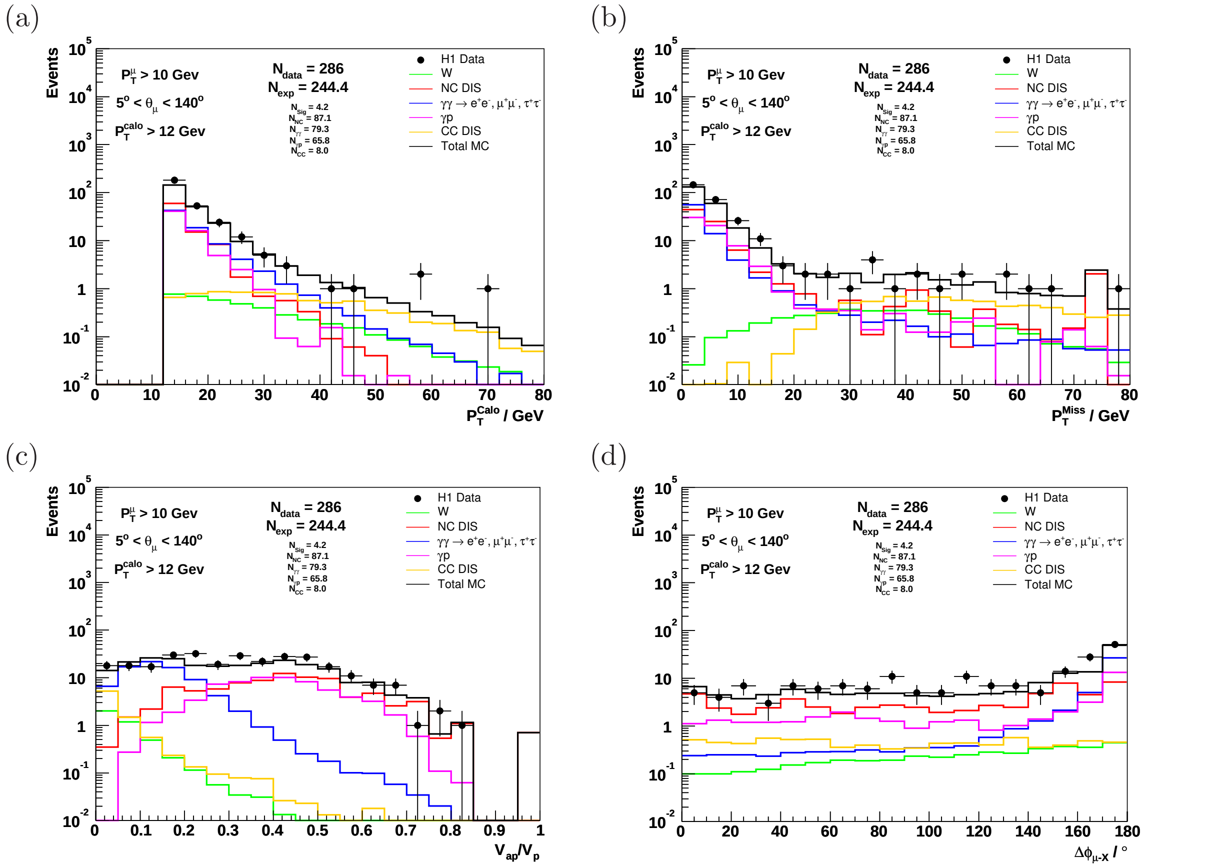


Figure 7.10: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the ratio V_{ap}/V_p (c) and the acoplanarity angle between the muon candidate and the hadronic final state (d) in events in the enlarged LAr triggered phase-space defined by the criteria listed in table 7.3.*

hadronic transverse momenta are typically small. The distribution of the topological variable V_{ap}/V_p is shifted towards low values, due to the typically small energy deposits of the muons in the calorimeter, while the distribution of the acoplanarity angle is very strongly peaked in the "back-to-back" configuration. Perhaps surprisingly, the distribution of the number of isolated muons in the selected events does not peak at two. The reasons for the lower number of isolated muons reconstructed in many muon pair production events are muons produced at either very small or very large polar angles, outside the geometrical acceptance of the detector, and muons that fail the isolation criteria.

The contribution from CC DIS processes is smaller by about a factor four than the contribution of such events to the enlarged phase-space electron sample [1]. This factor is explained by the smaller probability for hadrons to be misidentified as muons than as electrons and shows that in CC DIS events most electron and muon candidates are due to misidentified hadrons and only a few due to heavy quark

[1] Note that the correlation between the misidentification of a hadron as a muon and the reconstruction of an imbalance of energy deposits in the calorimeter that increased the contribution from photoproduction processes is not relevant for CC DIS processes, as CC DIS events are intrinsically imbalanced due to the presence of an undetected neutrino in the final state.

decays. (Heavy quarks decay with equal probability into electrons and muons.) As expected for misidentified hadrons, the distributions of the distances D_{track} and D_{jet} of the muon candidate to the nearest charged particle and the nearest jet both peak at small values.

Definition of the Final Selection Criteria

The final selection of events with isolated muons in the LAr triggered phase-space is based on the enlarged phase-space defined by the selection criteria listed in table 7.3. In addition to passing the selection criteria for the enlarged phase-space, events in the final sample are required to contain an isolated muon and large missing transverse momentum [2].

The requirement for the muon candidates to be isolated effectively suppresses the background contribution of events in which the muon candidate is due to a misidentified hadron or the decay of a heavy quark. For muon candidates to be considered as isolated, they are required to be flagged as "isolated lepton" by the muon identification algorithm (see section 4.4) and be separated from other charged particles and jets by at least $D_{track} > 0.5$ and $D_{jet} > 1.0$, respectively. As in the case of events with isolated electrons, the isolation requirements effectively suppress the background contribution from NC DIS, photoproduction and CC DIS processes.

The large missing transverse momentum requirement effectively suppresses the background of intrinsically balanced events with muons of high transverse momentum or hadrons reaching the muon system. For events to be considered as having a large missing transverse momentum, $P_T^{miss} > 12\ GeV$ is required in the final selection. In order to further suppress the background of intrinsically balanced events in which the reconstructed missing transverse momentum is due to a mismeasurement, a low value of the topological variable V_{ap}/V_p is required. In order not to reject too much signal, the V_{ap}/V_p requirement is formulated as a two-dimensional cut. The correlation between V_{ap}/V_p and the imbalance of the energy deposits in the calorimeter is shown in figure 7.11, separately for signal and background processes. In the figure, it can be seen that P_T^{calo} is typically small in intrinsically balanced events in which the reconstructed imbalance of energy deposits is due to a mismeasurement in the calorimeter. In signal events, P_T^{calo} is typically of the order of the hadronic transverse momentum P_T^X and comparatively high. A large fraction of the NC DIS and photoproduction backgrounds is rejected by requiring $V_{ap}/V_p < 0.15$ for $P_T^{calo} < 25\ GeV$ and $V_{ap}/V_p < 0.5$ for $P_T^{calo} > 25\ GeV$, while most signal events pass these requirements.

The background from lepton pair production processes is suppressed further by requiring a significant acoplanarity angle $\Delta\phi_{\mu-X} < 170°$ between the isolated

[2] As in the enlarged phase-space selection, the muon candidates are required to have transverse momenta above $p_T^{\mu} > 10\ GeV$ and be in the polar angle range $5° < \theta_{\mu} < 140°$ in the final selection. The transverse momentum and polar angle requirements for the muon candidates are the same as for the electron candidates in events with isolated electrons discussed in the last section. In events with isolated muons, the backward region is excluded in order to avoid misidentification of hadrons traversing the electromagnetic section of the BBE and reaching the iron (the BBE has no hadronic section, *cf.* figure 1.9). As in the case of events with isolated electrons, the requirements on the transverse momentum and polar angle do not significantly reduce the acceptance for events in which the isolated muon results from the decay of a heavy particle.

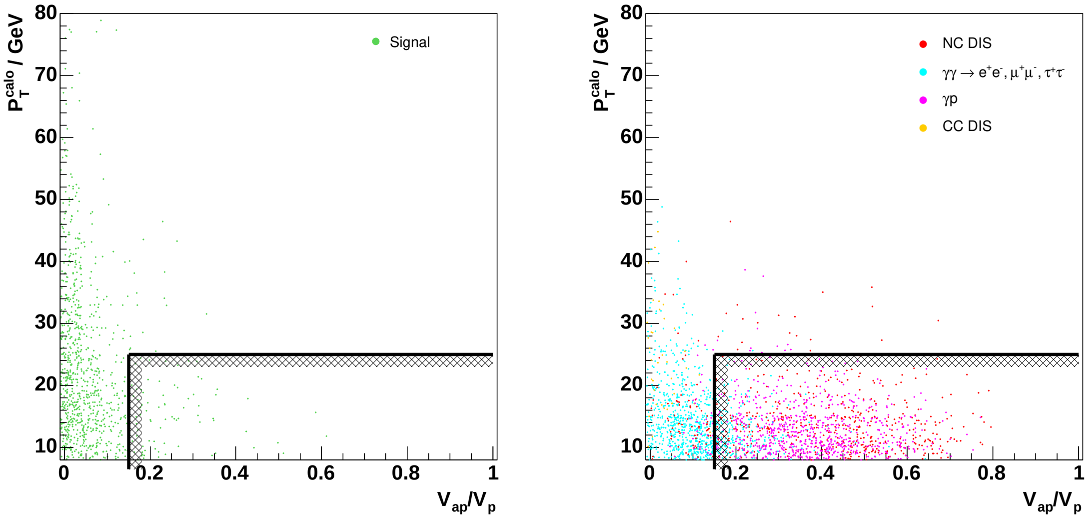


Figure 7.11: *Correlation between the ratio V_{ap}/V_p and the imbalance of energy deposits in the calorimeter P_T^{calo} for simulated W production events (left) and events produced by background processes (right). The two-dimensional cut applied in the final selection of events with isolated muons in the LAr triggered phase-space is depicted by the hatched line in the lower right corner.*

muon and the hadronic final state in combination with a large hadronic transverse momentum $P_T^X > 12\ GeV$. The requirement that $P_T^X > 12\ GeV$ effectively halves the contribution from lepton pair production processes by rejecting elastic muon pair production events in which the reconstructed imbalance of energy deposits in the calorimeter is due to the scattered beam electron. For the signal, the $P_T^X > 12\ GeV$ requirement rejects virtually no events, because of its equivalence to the requirement of $P_T^{calo} > 12\ GeV$ in the LAr triggered phase-space. In addition to significantly reducing the contribution from lepton pair production processes, the $P_T^X > 12\ GeV$ requirement also guarantees that the acoplanarity angle is well defined.

The final selection criteria for events with isolated muons and large missing transverse momentum in the LAr triggered phase-space are listed in table 7.4.

Selection Efficiency

The efficiency of the selection criteria listed in table 7.4 for events with isolated muons and large missing transverse momentum in the LAr triggered phase space is estimated as a function of the generated hadronic transverse momentum P_T^X. The estimated efficiency for Standard Model W production is shown in figure 7.12. Also shown in the figure is the estimated efficiency for decays of single top quarks produced by flavour-changing neutral current interactions, as an example for possible signal processes beyond the Standard Model. The selection efficiencies for both processes are seen to agree within a few percent. Due to the $P_T^{calo} > 12\ GeV$ and $P_T^X > 12\ GeV$ requirements in the event selection, the selection efficiency vanishes for events with generated hadronic transverse momenta below $\mathcal{O}\,(10\ GeV)$, but reaches up to 50% at high P_T^X. For hadronic transverse momenta above $\mathcal{O}\,(20\ GeV)$, the efficiency of the selection criteria listed in table 7.4 for events with isolated muons

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST15 \parallel ST18 \parallel ST19 \parallel ST22 \parallel ST34 \parallel ST56 \parallel$ $ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-ep background finder	Ibg bits $0, 1, 2, 3, 4, 5, 7, 8$ $Ibgfm$ bits $0, 1, 2, 4, 9, 10, 11, 16$ $Ibgam$ bits $0, 1, 2, 3, 4, 5, 7, 8$
Muon	$p_T^\mu > 10$ GeV $5° < \theta_\mu < 140°$ "Iron muon" flagged as "isolated lepton" $D_{track} > 0.5$ $D_{jet} > 1.0$
Missing momentum	$P_T^{miss} > 12$ GeV $\Delta\phi_{\mu-X} < 170°$ $V_{ap}/V_p < 0.5$ $\left(< 0.15 \text{ for } P_T^{calo} < 25\ GeV\right)$
Additional Requirements	$P_T^{calo} > 12$ GeV $P_T^X > 12$ GeV Not two isolated Muons

Table 7.4: *Selection criteria for the final sample of events with isolated muons in the LAr triggered phase-space. "Iron muons" are defined as passing the identification criteria listed in table 4.2. The "isolated lepton" flag is defined in section 4.2.*

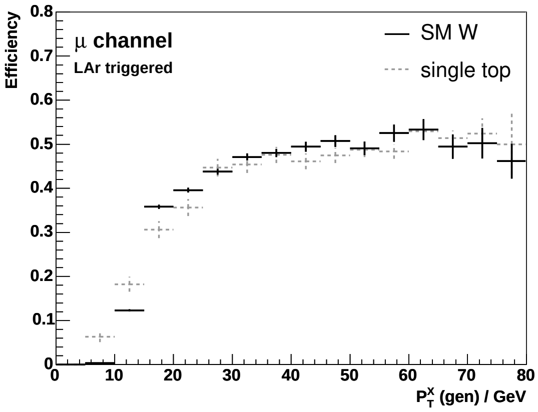


Figure 7.12: *Efficiency of the selection criteria listed in table 7.4 for events with isolated muons and large missing transverse momentum. The selection efficiency for Standard Model signal processes is represented by the solid line; the dashed line represents the efficiency for single top production as an example process beyond the Standard Model.*

becomes comparable to that of the selection criteria listed in table 7.2 for events with isolated electrons.

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST15 \parallel ST18 \parallel ST19 \parallel ST22 \parallel ST34 \parallel ST56 \parallel$ $ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-*ep* background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 7, 8$ *Ibgfm* bits $0, 1, 2, 4, 9, 10, 11, 16$ *Ibgam* bits $0, 1, 2, 3, 4, 5, 7, 8$
Muon	$p_T^{\mu} > 10\ GeV$ $5^\circ < \theta_\mu < 140^\circ$ Radial track length in CTD $> 30\ cm$ "Iron muon" with track-link probability $> 10^{-4}$

Table 7.5: *Selection criteria defining the enlarged muon triggered phase-space. "Iron muons" are defined as passing the identification criteria listed in table 4.2. The track-link probability is described in reference [137].*

7.2.2 Muon Triggered Phase-Space

Definition of the Phase-Space

The enlarged muon triggered phase-space is defined by requiring a muon candidate of transverse momentum $p_T^{\mu} > 10\ GeV$ passing the selection criteria listed in table 4.2 in the polar angle range $20^\circ < \theta_\mu < 140^\circ$. The polar angle range is limited to the geometric acceptance of the central tracking detector as this is the region covered by the trigger. For the muon candidate, a minimal radial track length (distance between the first and last hit associated to the track in the transverse plane) of 30 cm in the central tracking detector and an estimated probability for linking the extrapolated central track with a track reconstructed in the instrumented iron of at least 10^{-4} is required [3]. These requirements guarantee that the transverse momentum of the muon candidate is well measured and reduce the contribution from background processes at lower P_T^X [4]. The muon candidates are not required to be isolated in the enlarged phase-space, in order to increase the statistics of the background samples, and show the effect of the isolation criteria on the background suppression. The selection criteria defining the enlarged muon triggered phase-space are listed in table 7.5 [5].

[3] See reference [137] for a description of the track-linking.

[4] In events with low hadronic transverse momentum, the requirement that there be a well measured muon track is also necessary for technical reasons: if the transverse momentum of the muon candidate were not required to be well measured, low p_T background events, in which the reconstructed high transverse momentum of the muon track is due to a mismeasurement, would contribute significantly to the selected event sample. As it is very difficult to simulate Monte Carlo samples of sufficient size to model this contribution, it must be strongly suppressed.

[5] Note that the selection criteria in the enlarged LAr and muon triggered phase-space are not exclusive, so that any given event may contribute to both samples.

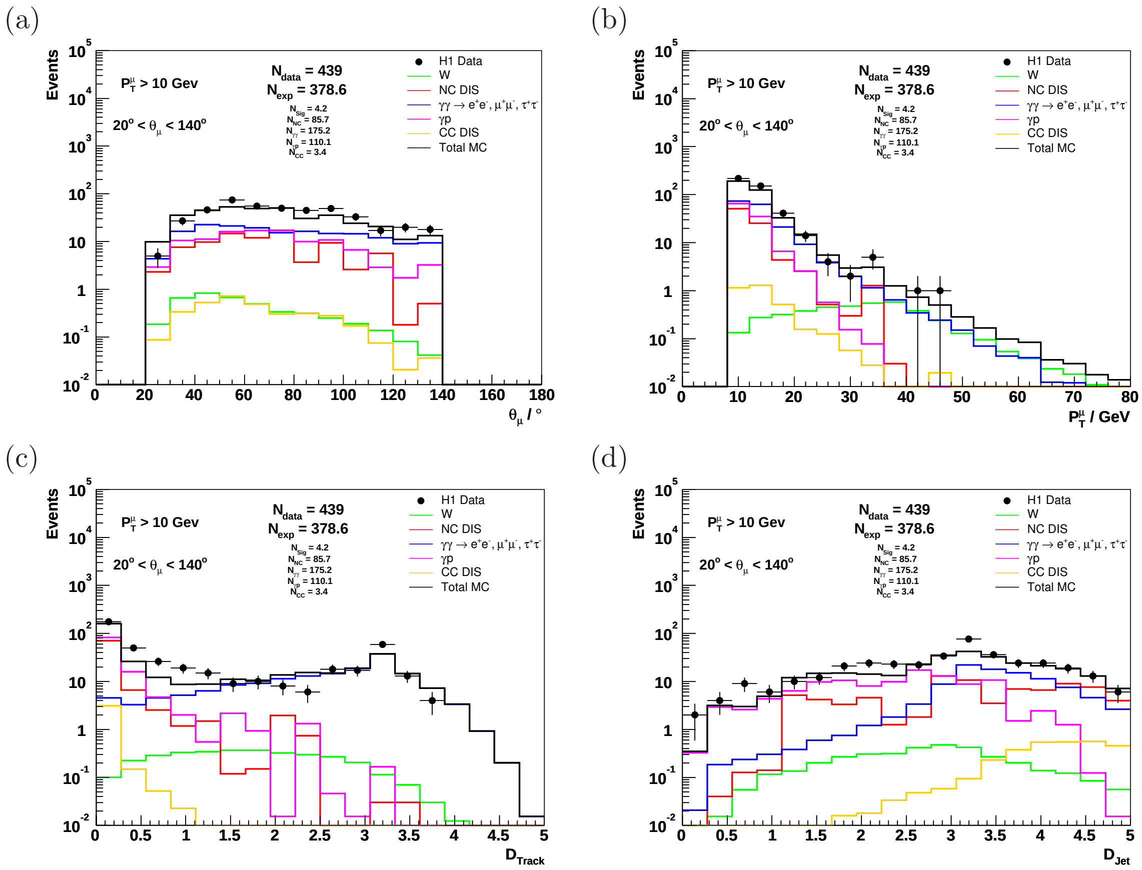


Figure 7.13: *Distributions of the polar angle (a), the transverse momentum (b), the distance to the nearest track (c) and the distance to the nearest jet (d) of the muon candidates in events in the enlarged muon triggered phase-space defined by the criteria listed in table 7.5.*

Background Studies

In the phase-space defined by the selection criteria listed in table 7.5, 439 events are selected in the 1994-2000 e^+p and e^-p datasets, which - within the systematic uncertainties discussed in section 8.1 - compares well with an expectation of 379 events from the Monte Carlo simulation. For the selected event sample, distributions of the quantities defined in chapter 5 are shown in figures 7.13, 7.14, and 7.15.

The selected event sample is dominated by the contribution of lepton pair production processes, in particular by elastic muon pair production events. In typical elastic muon pair production processes, two muons of equal transverse momentum are produced "back-to-back" in the transverse plane, and neither the hadronic system nor the scattered beam electron have significant transverse momentum. In fact, elastic processes form the largest fraction of the total lepton pair production cross-section, but are suppressed by the $P_T^{calo} > 12\ GeV$ requirement in the LAr triggered phase-space.

As in the LAr triggered phase-space, the contributions of NC DIS and photoproduction processes are characterised by very similar event topologies. In both cases, the muon candidate identified in the detector is due to a misidentified hadron

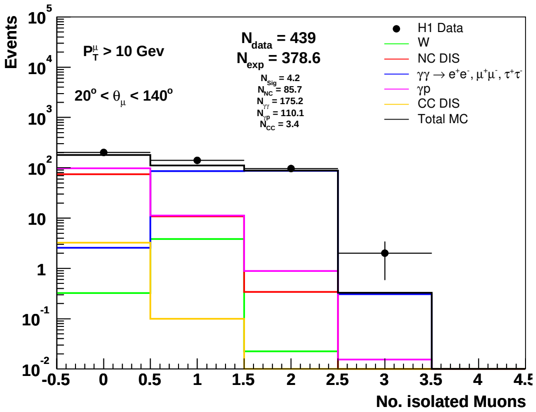


Figure 7.14: *Distribution of the number of isolated muons in events in the enlarged muon triggered phase-space defined by the criteria listed in table 7.5. The criteria for muons to be considered as "isolated" are defined in section 4.4.*

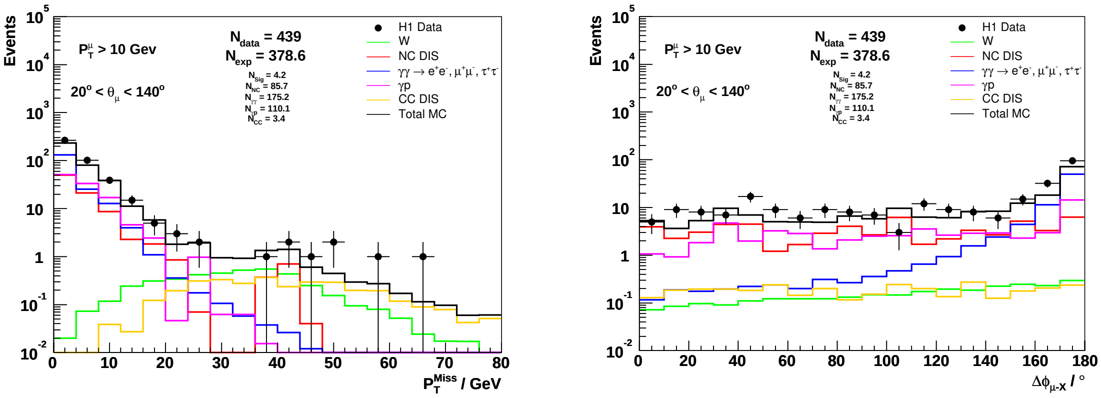


Figure 7.15: *Distributions of the reconstructed missing transverse momentum (left) and the acoplanarity angle between the muon candidate and the hadronic final state (right) in events in the enlarged muon triggered phase-space defined by the criteria listed in table 7.5.*

or produced in the decay of a heavy quark. The muon candidates in NC DIS and photoproduction events are typically within jets and seldom isolated, as can be seen in the distributions of the distances D_{track} and D_{jet} of the muon candidate to the nearest charged particle and the nearest jet, both of which peak at small values.

The contribution from CC DIS processes is small. In comparison to the LAr triggered phase-space, the contribution from CC DIS processes is suppressed by the more stringent quality criteria on the muon track in the muon triggered phase-space: the requirement of a track-link between an extrapolated central track and a track reconstructed in the instrumented iron reduces the probability for hadrons to be misidentified as muons by about a factor two to three. As the $P_T^{calo} > 12\ GeV$ requirement rejects only very few CC DIS events, almost all of the events selected in the muon triggered phase-space are also selected in the LAr triggered phase-space. The distributions of the distances D_{track} and D_{jet} of the muon candidate to the nearest charged particle and the nearest jet are of similar shape to those in the LAr triggered phase-space [6] and both peak at small values.

Definition of the Final Selection Criteria

The final selection of events with isolated muons in the muon triggered phase-space is based on the enlarged phase-space defined by the selection criteria listed in table 7.5. In addition to passing the selection criteria for the enlarged phase-space, events in the final sample are required to contain an isolated muon and have large missing transverse momentum [7].

For muon candidates to be considered as isolated, they are required to be flagged as "isolated lepton" by the muon identification algorithm and be separated from other charged particles and jets by at least $D_{track} > 0.5$ and $D_{jet} > 1.0$. The isolation requirements effectively suppress the background contributions from NC DIS, photoproduction and CC DIS processes.

The requirement of large missing transverse momentum effectively suppresses the background of intrinsically balanced events. For events to be considered as having a large missing transverse momentum, $P_T^{miss} > 12\ GeV$ is required in the final selection.

As in the LAr triggered phase-space, the background from lepton pair production processes is further suppressed by rejecting events with two isolated muons in the final state.

The final selection criteria for events with isolated muons and large missing transverse momentum in the muon triggered phase-space are listed in table 7.6.

The final event samples in the LAr and muon triggered phase-space are made exclusive by requiring that the events in the final sample of the muon triggered phase-space are not selected in the LAr triggered phase-space. This avoids double-counting when both event samples are combined.

Selection Efficiency

The efficiency of the selection criteria listed in table 7.6 for events with isolated muons and large missing transverse momentum in the muon triggered phase-space is estimated as a function of the generated hadronic transverse momentum P_T^X in the same way as for the LAr triggered phase-space. The estimated efficiencies for Standard Model W production and the decays of single top quarks produced by flavour-changing neutral current interactions, as an example for possible signal processes beyond the Standard Model, are shown in figure 7.16. For comparison with the efficiencies for events with isolated electrons, the efficiencies for events with isolated muons in the combined LAr and muon triggered phase-space are also shown. The muon triggered phase-space is seen to increase the efficiency for events with isolated muons especially at low P_T^X. In the region $P_T^X < 12\ GeV$, the efficiency is about 20%.

[6] This shows that the track-link probability is not correlated with the isolation of the muon candidates.

[7] As in the enlarged phase-space selection, the muon candidates are required to have transverse momenta above $p_T^{\mu} > 10\ GeV$ and to be in the polar angle range $20° < \theta_{\mu} < 140°$ in the final selection. See footnote 2 on page 152 for further remarks.

Quantity	Requirement
Vertex	Reconstructed using CTD tracks
	z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST15 \parallel ST18 \parallel ST19 \parallel ST22 \parallel ST34 \parallel ST56 \parallel$
	$ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system,
	CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-*ep* background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 7, 8$
	Ibgfm bits $0, 1, 2, 4, 9, 10, 11, 16$
	Ibgam bits $0, 1, 2, 3, 4, 5, 7, 8$
Muon	$p_T^{\mu} > 10$ GeV
	$5° < \theta_{\mu} < 140°$
	Radial track length in CTD > 30 cm
	"Iron muon" with track-link probability $> 10^{-4}$
	flagged as "isolated lepton"
	$D_{track} > 0.5$
	$D_{jet} > 1.0$
Missing momentum	$P_T^{miss} > 12$ GeV
Additional requirements	Not two isolated Muons
	Not selected in LAr triggered phase-space

Table 7.6: *Selection criteria for the final sample of events with isolated muons in the muon triggered phase-space. "Iron muons" are defined as passing the identification criteria listed in table 4.2; the "isolated lepton" flag is defined in section 4.2, and the track-link probability is described in reference [137].*

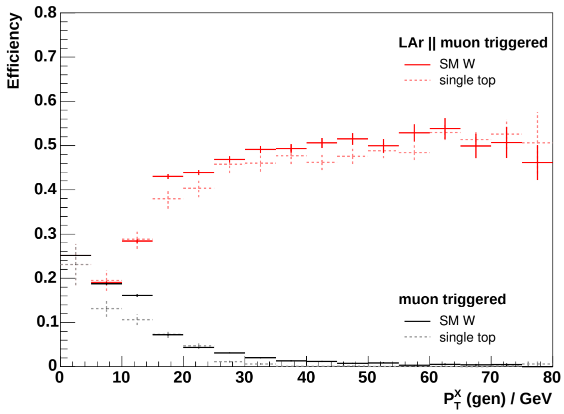


Figure 7.16: *Efficiency of the selection criteria listed in table 7.6 for events with isolated muons and large missing transverse momentum. The selection efficiency in the muon triggered phase-space alone is indicated in black; additionally, the combined efficiency of the LAr triggered and muon triggered phase-space regions is depicted in red. The solid line represents the expected selection efficiency for Standard Model processes, the dashed line an estimate of the efficiency for processes beyond the Standard Model.*

At low P_T^X, the efficiency of the combined LAr and muon triggered phase-space is still almost a factor of two lower than the efficiency for events with isolated electrons. The lower efficiency for events with isolated muons is mainly due to the limitation of the muon triggered phase-space to the geometric acceptance of the central tracking detector, the lower efficiency for muons to be triggered and identified in the instrumented iron than for electrons to be triggered and identified in the LAr, and the additional requirements on the radial track length in the CTD and the track-link probability for muon candidates in the muon triggered phase-space.

7.3 Events with Isolated Tau Leptons

7.3.1 Definition of the Phase-Space

The enlarged phase-space for events with isolated tau leptons is defined by requiring a jet compatible with being a hadronic "one-prong" or "three-prong" tau decay of transverse momentum $p_T^{jet} > 7\ GeV$ in the polar angle range $20° < \theta_\mu < 120°$ and a reconstructed imbalance $P_T^{calo} > 12\ GeV$ of the energy deposits in the calorimeter. The transverse momentum required for tau jet candidates is lower than that for electrons and muons, as on average only about 70% (80%) of the momentum of tau leptons is detected in hadronic "one-prong" ("three-prong") tau decays. (The rest is carried away by the tau neutrino, *cf.* section 4.5.1.) In comparison to electrons and muons, the tau jet candidates are required to be in the more central region of the H1 detector. In the forward region, the identification of hadronic tau decays is difficult, due to the dead material in front of and within the forward tracking detector that causes an increase of the track multiplicity through photon conversions. As a consequence, the identification efficiency for hadronic tau decays is low at polar angles below 20° (*cf.* figure 4.21). For this reason, and also because the QCD jet background is very high at small polar angles, the forward region is excluded from the identification of hadronic tau decays. The backward region is excluded mainly because of the high background of unidentified electrons. As also the depth of the LAr calorimeter decreases at polar angles above about 120° (*cf.* figure 1.9), the identification of hadronic tau decays is restricted to polar angles below 120°. In order to increase the size of the background samples and show the effect of the neural network based tau identification on the background suppression, the tau jet candidates are not required to pass any requirements on the output of the neural networks trained to identify hadronic tau decays as described in section 4.5.5; in the enlarged phase-space, jets are considered to be compatible with being hadronic "one-prong" or "three-prong" tau decays if they pass the preselection criteria described in section 4.5.2 on the number of tracks (not compatible with photon-conversion) within the jet. The selection criteria defining the enlarged phase-space for events with isolated tau leptons are listed in table 7.7.

7.3.2 Background Studies

In the phase-space defined by the selection criteria listed in table 7.7, 402 events are selected in the 1994-2000 e^+p and e^-p datasets, which - within the systematic

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-ep background finder	Ibg bits $0, 1, 2, 3, 4, 5, 7, 8$ $Ibgfm$ bits $0, 1, 2, 4, 9, 10, 11, 16$ $Ibgam$ bits $0, 1, 2, 3, 4, 5, 7, 8$
Tau lepton	$p_T^{jet} > 7\ GeV$ $20^\circ < \theta_{jet} < 120^\circ$ Compatible with hadronic "one-prong" $\parallel$ "three-prong" tau decay
Additional requirements	$P_T^{calo} > 12\ GeV$

Table 7.7: *Selection criteria defining the enlarged phase-space for events with isolated tau leptons. Jets are considered as being compatible with hadronic tau decays if they pass the preselection criteria detailed in section 4.5.2; they are not required to pass any selection criteria on the neural network output described in section 4.5.5.*

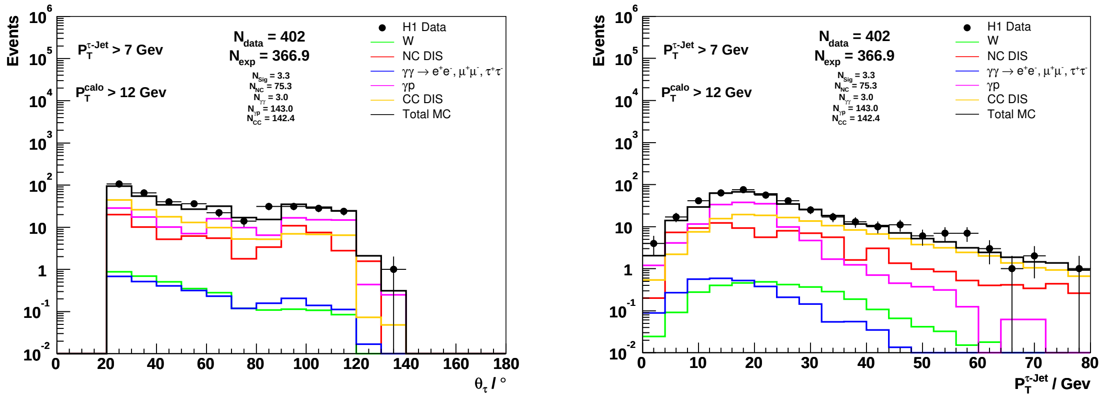


Figure 7.17: *Distributions of the polar angle (left) and the transverse momentum (right) of the tau jet candidates in events in the enlarged phase-space defined by the criteria listed in table 7.7.*

uncertainties discussed in section 8.1 - compares well with an expectation of 367 events from the Monte Carlo simulation. For the selected event sample, distributions of the quantities defined in chapter 5 are shown in figures 7.17, 7.18, 7.19 and 7.20.

The selected event sample is dominated by the contribution from photoproduction and CC DIS processes. In the selected CC DIS events, the tau jet candidate is due to a misidentified QCD-jet and the reconstructed imbalance of the energy deposits in the calorimeter is due to an undetected neutrino in the final state. Due to the presence of the neutrino, the distributions of P_T^{calo}, V_{ap}/V_p and $\Delta\phi_{\tau-X}$ are very similar for CC DIS and signal events (represented by leptonic W decays). The main difference between signal and CC DIS events is the presence (absence) of a tau

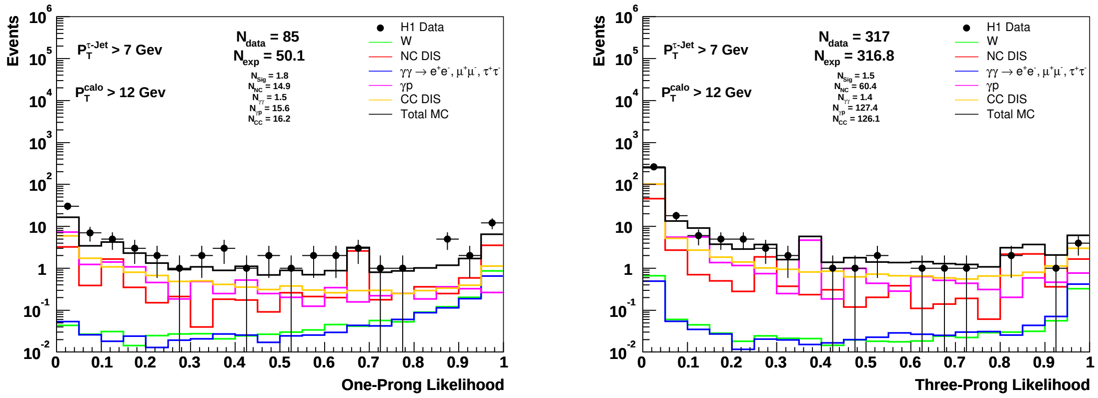


Figure 7.18: *Distributions of the classification output of the neural networks described in section 4.5.5 for "one-prong" (left) and "three-prong" (right) tau jet candidates in events in the enlarged phase-space defined by the criteria listed in table 7.7. Note the different number of entries in the two distributions: for any given tau jet candidate to enter the "one-prong" ("three-prong") likelihood distribution, it is required to pass the precriteria detailed in section 4.5.2 for hadronic "one-prong" ("three-prong") tau decays.*

jet resulting from a hadronic tau decay in signal (CC DIS) events. The presence (absence) of a tau jet in signal (CC DIS) events can be seen in the distributions of the likelihood for the tau jet candidate to be compatible with either hadronic "one-prong" or hadronic "three-prong" tau decays: in most signal events, the tau jet candidates are classified as being likely to be a tau jet ($f_{NN} > 0.75$), while the tau jet candidates in most CC DIS events are classified as being likely to be a QCD-jet ($f_{NN} < 0.75$). In the likelihood distributions, it can also be seen that the tau jet candidates in most signal (CC DIS) events are candidates for hadronic "one-prong" ("three-prong") tau decays (in agreement with the branching fractions in tau decays detailed in section 4.5 and the charged particle multiplicity in QCD jets discussed in sections 4.10 and 4.5.5).

The contribution from photoproduction processes arises from events in which the tau jet candidate is due to a misidentified QCD-jet and the reconstructed imbalance of the energy deposits in the calorimeter is due to the mismeasurement of an intrinsically balanced event in the detector. In typical photoproduction events, two jets of equal transverse momentum are produced "back-to-back" in the transverse plane, of which one is misidentified as a tau jet candidate. As expected for intrinsically balanced events, the distribution of the ratio V_{ap}/V_p peaks at high values, and the distribution of the acoplanarity angle $\Delta\phi_{\tau-X}$ at the "back-to-back" configuration (but not very strongly).

The contribution from NC DIS processes arises from events in which the tau jet candidate is either due to a misidentified QCD-jet or to the (unidentified) scattered beam electron and the reconstructed imbalance of the energy deposits in the calorimeter is due to the mismeasurement of an intrinsically balanced event in the detector. When comparing the polar angle distributions of the tau jet candidates in NC DIS events with those in CC DIS events, it may be seen that the contribution of scattered beam electrons becomes relevant at polar angles above about 90°, but that

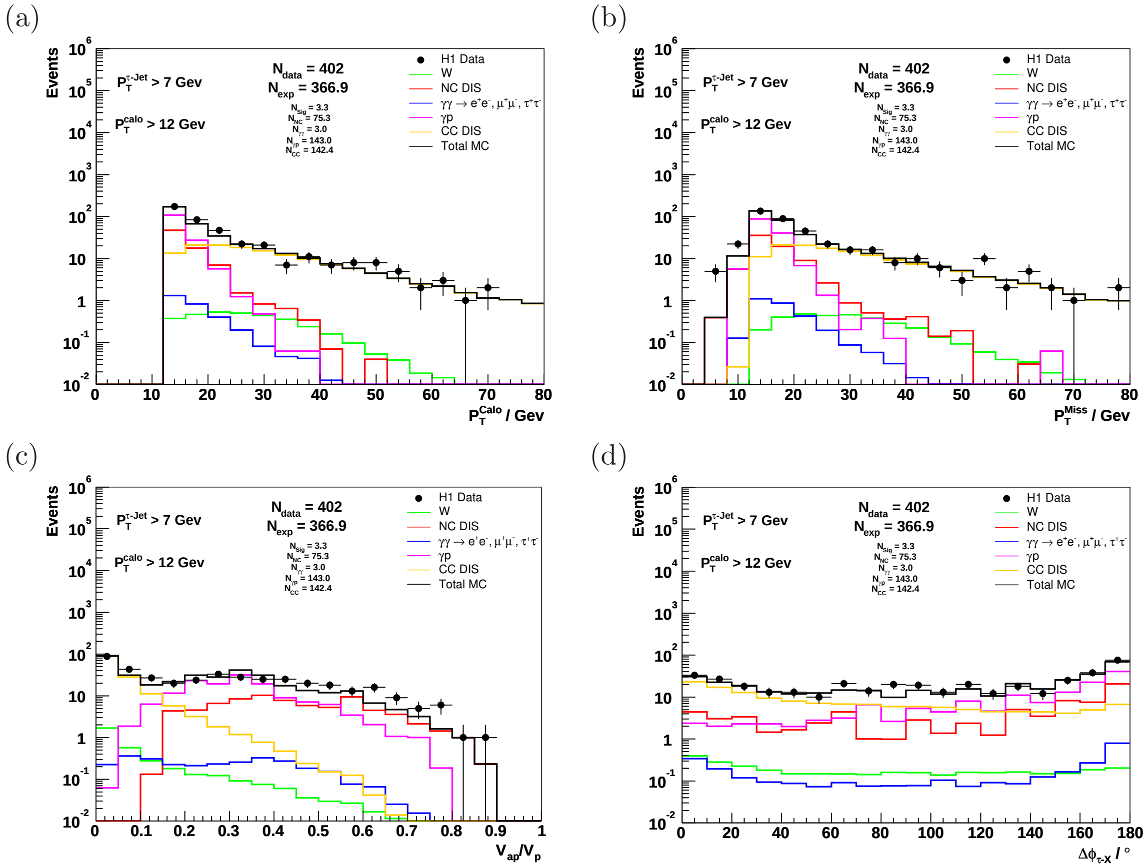


Figure 7.19: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the ratio V_{ap}/V_p (c) and the acoplanarity angle between the tau jet candidate and the hadronic final state (d) in events in the enlarged phase-space defined by the criteria listed in table 7.7.*

most tau jet candidates are due to misidentified QCD jets in NC DIS events. The comparatively small contribution of unidentified electrons is a direct consequence of the fiducial cuts discussed in section 4.5.6, in particular the exclusion of the ϕ-cracks in the LAr calorimeter from the identification of hadronic "one-prong" tau decays. Regardless of whether the tau jet candidate is due to a misidentified QCD-jet or the scattered beam electron, only particles going down the beam-pipe in the proton direction are undetected in NC DIS events: as expected, the distribution of the missing longitudinal momentum δ_{miss} is seen to peak at zero.

The contribution from lepton pair production processes is relatively small. The majority of the selected events are due to elastic tau pair production in which the tau jet candidate is due to the hadronic decay of one of the two tau leptons and the reconstructed imbalance of the energy deposits in the calorimeter is due to the neutrinos produced in the tau decays [1]. In typical elastic tau pair production processes, neither the hadronic system nor the scattered beam electron have significant

[1] The contribution of tau pair production processes is considered as background in the analysis of events with isolated leptons and large missing transverse momentum presented in this thesis, as in tau pair production processes no undetected particles are produced *together* with the tau leptons - the neutrinos are produced *subsequently* in the tau decays only.

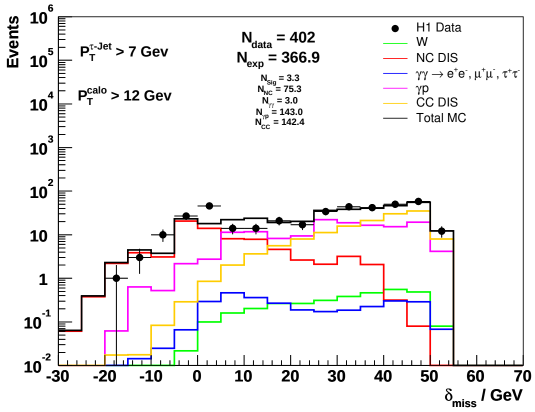


Figure 7.20: *Distribution of the missing longitudinal momentum in events in the enlarged phase-space defined by the criteria listed in table 7.7.*

transverse momentum and the two tau leptons are produced "back-to-back" in the transverse plane and with equal transverse momenta. As a consequence of the relatively small mass of tau leptons and the Lorentz boost in the tau direction, the detected particles as well as the neutrinos resulting from the tau decays are likely to be produced collinear to the directions of the original tau leptons (*cf.* section 4.5). The detected particles are therefore likely to be "back-to-back" in elastic tau pair production events and the contributions of the neutrinos produced in the tau decays to the transverse momentum balance of the events are likely to cancel. This effect can be seen in the distribution of the missing transverse momentum P_T^{miss} which falls almost as steeply as in events without undetected neutrinos in the final state (NC DIS and γp), and in the distribution of the topological variable V_{ap}/V_p which is shifted towards higher ratios in comparison to the contribution from lepton pair production processes to the enlarged phase-space samples of events with isolated electrons or muons.

7.3.3 Definition of the Final Selection Criteria

The final selection of events with isolated tau leptons is based on the enlarged phase-space defined by the selection criteria listed in table 7.7. In addition to passing the selection criteria for the enlarged phase-space, events in the final sample are required to contain a jet compatible with being a hadronic "one-prong" tau decay, have large missing transverse momentum, and pass some additional criteria designed to suppress the background of NC DIS events in which the tau jet candidate is due to the misidentified scattered electron [2].

The background arising from the misidentification of QCD jets is suppressed by more stringent requirements on the signature of the tau jet candidates. For tau jet candidates to be considered as compatible with hadronic "one-prong" tau

[2] As in the enlarged phase-space selection, the tau jet candidates are required to have transverse momenta above $p_T^{jet} > 7\ GeV$ and be in the polar angle range $20^\circ < \theta_{jet} < 120^\circ$ in the final selection. The "three-prong" tau jet candidates are excluded from the final selection, because of the high background due to misidentified QCD jets.

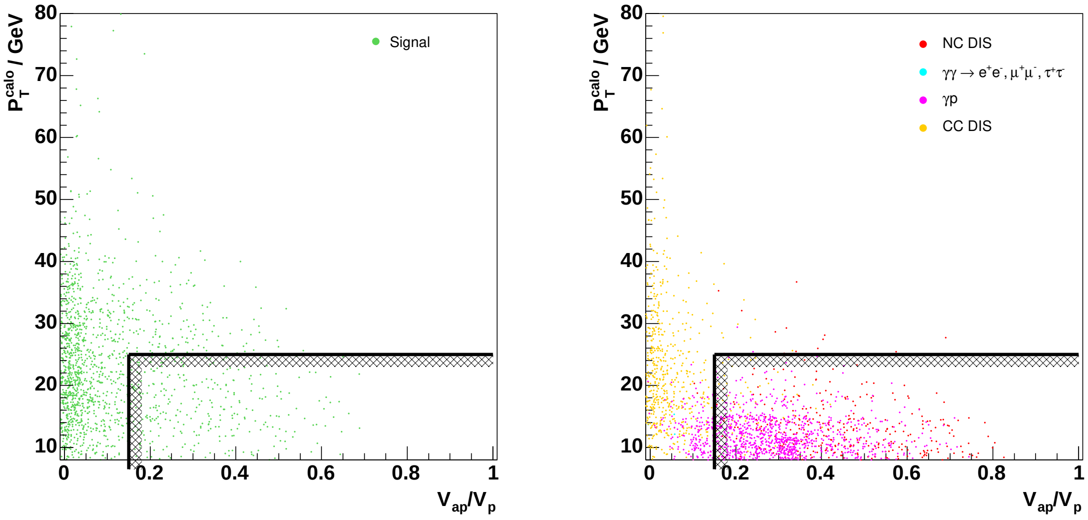


Figure 7.21: *Correlation between the ratio V_{ap}/V_p and the imbalance of energy deposits in the calorimeter P_T^{calo} for simulated W production events (left) and events produced by background processes (right). The two-dimensional cut applied in the final selection of events with isolated tau leptons is depicted by the hatched line in the lower right corner.*

decays, a classification output of the neural networks described in section 4.5.5 of at least $f_{NN} > 0.75$ is required in the final selection. The additional requirement on the neural network output suppresses the background contribution from NC DIS, photoproduction and CC DIS processes [3].

The large missing transverse momentum requirement does not suppress much background, as, in events with isolated tau leptons, the missing transverse momentum P_T^{miss} is similar to the imbalance P_T^{calo} of the energy deposits in the calorimeter. For events to be considered as having a large missing transverse momentum, $P_T^{miss} > 12\ GeV$ is required in the final selection. In order to further suppress the background of intrinsically balanced events in which the reconstructed missing transverse momentum is due to a mismeasurement, a low value of the topological variable V_{ap}/V_p is required. In order not to reject too much signal, the V_{ap}/V_p requirement is formulated as a two-dimensional cut. The correlation between V_{ap}/V_p and the imbalance of the energy deposits in the calorimeter is shown in figure 7.21, separately for signal and background processes. In the figure, it can be seen that P_T^{calo} is typically small in intrinsically balanced events in which the reconstructed imbalance of energy deposits is due to a mismeasurement in the calorimeter. In signal events, P_T^{calo} is comparatively high. A large fraction of the NC DIS, photoproduction and lepton pair production backgrounds is rejected by requiring $V_{ap}/V_p < 0.15$ for $P_T^{calo} < 25\ GeV$ and $V_{ap}/V_p < 0.5$ for $P_T^{calo} > 25\ GeV$, while most events with undetected high p_T particles in the final state (W production, but also CC) pass these requirements. The background of intrinsically balanced events is suppressed further by requiring a significant acoplanarity angle $\Delta\phi_{\tau-X} < 170°$. The combination of the requirements on P_T^{miss}, V_{ap}/V_p and $\Delta\phi_{\tau-X}$ effectively suppresses the "back-to-back" topologies characteristic of the NC DIS, photoproduction and lepton

[3]CC DIS processes are particularly difficult to suppress by other means.

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within $^{+100}_{-40}$ cm around nominal position
Trigger	$ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-ep background finder	Ibg bits $0, 1, 2, 3, 4, 5, 7, 8$ $Ibgfm$ bits $0, 1, 2, 4, 9, 10, 11, 16$ $Ibgam$ bits $0, 1, 2, 3, 4, 5, 7, 8$
Tau lepton	$p_T^{jet} > 7\ GeV$ $20^\circ < \theta_{jet} < 120^\circ$ Compatible with hadronic "one-prong" tau decay Neural network output $f_{NN} > 0.75$
Missing momentum	$P_T^{miss} > 12\ GeV$ $\Delta\phi_{\tau-X} < 170^\circ$ $V_{ap}/V_p < 0.5$ $\left(< 0.15 \text{ for } P_T^{calo} < 25\ GeV\right)$
NC DIS suppression	$\delta_{miss} > 5\ GeV$[1]
Additional requirements	$P_T^{calo} > 12\ GeV$ No isolated Muon

[1] if no electron and a "one-prong" tau jet candidate, which has the same charge as the beam electron, is dete

Table 7.8: *Selection criteria for the final sample of events with isolated tau leptons. Jets are considered as being compatible with hadronic "one-prong" tau decays if they pass the preselection criteria detailed in section 4.5.2 and the requirement on the classification output f_{NN} of the neural networks described in section 4.5.5.*

pair production backgrounds.

The background from NC DIS processes arising from events in which the tau jet candidate is due to the misidentified scattered beam electron is further suppressed by requiring a missing longitudinal momentum of $\delta_{miss} > 5\ GeV$ for events in which no electron and a "one-prong" tau jet candidate that has the same charge as the beam electron is detected.

The background from lepton pair production processes is suppressed further by rejecting events with isolated muons. This requirement not only suppresses the contribution of muon pair production events in which an unidentified muon is misidentified as "one-prong" tau jet, but also of tau pair production events in which one of the tau leptons decays hadronically and the other into a muon.

The final selection criteria for events with isolated tau leptons and large missing transverse momentum are listed in table 7.8.

7.3.4 Selection Efficiency

The efficiency of the selection criteria listed in table 7.8 for events with isolated tau leptons and large missing transverse momentum is estimated as a function of the gen-

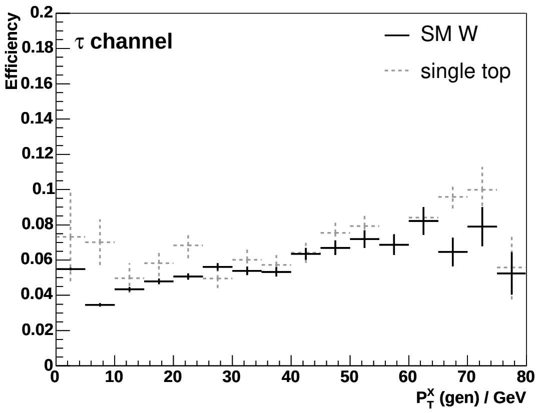


Figure 7.22: *Efficiency of the selection criteria listed in table 7.8 for events with isolated tau leptons and large missing transverse momentum. The selection efficiency for Standard Model signal processes is represented by the solid line; the dashed line represents the efficiency for single top production as an example of a beyond the Standard Model process.*

erated hadronic transverse momentum P_T^X. The estimated efficiencies for Standard Model W production and decays of single top quarks produced by flavour-changing neutral current interactions, as an example for possible signal processes beyond the Standard Model, are shown in figure 7.22. The selection efficiencies for both processes are seen to agree and are about 6%. The selection efficiency for events with isolated tau leptons is significantly lower than that for events with isolated electrons or muons. The lower efficiency is mainly due to the efficiency with which hadronic tau decays can be identified. In comparison to the identification efficiencies of about 90% for electrons and muons (see sections 4.3 and 4.4, respectively), the efficiency for the identification of hadronic tau decays is limited firstly by the branching fraction for tau leptons to decay hadronically of about 70% and secondly by the efficiency of about 50% with which the hadronic tau decays are identified by the algorithm described in section 4.5. The selection efficiency for events with isolated tau leptons is reduced further by the necessary exclusion of the forward region from the identification of hadronic tau decays, the undetected momentum carried away by the neutrino produced in the tau decay (whose transverse momentum is likely to cancel that of the neutrino produced in the W decay), and the lower trigger efficiency for events with isolated tau leptons in comparison to events with isolated electrons or muons. Overall, the selection efficiency for events with isolated tau leptons is only about one seventh of that for events with isolated electrons or muons.

Chapter 8

Results

In this chapter, the results of the search for events with isolated leptons and large missing transverse momentum are presented. For an assessment of the results, the uncertainty on the Standard Model expectation needs to be estimated. The total uncertainty is composed of statistical and *systematic* uncertainties. The systematic uncertainties are described in more detail in the first section of this chapter, before the events with isolated electrons, muons or tau leptons and large missing transverse momentum selected in the data are presented and compared to the Standard Model expectation. The results for events with isolated electrons, muons and tau leptons are first presented separately. Within the Standard Model, *lepton universality* predicts that an equal number of events with electrons, muons and tau leptons is produced. In theories beyond the Standard Model, this is not necessarily the case, however. For this reason, the compatibility of the results for events with isolated electrons, muons and tau leptons is explicitly checked before the individual channels are combined and the combined results are presented. At the end of this chapter, cross-sections for the production of events with isolated leptons and large missing transverse momentum in *ep* collisions are extracted.

8.1 Systematic Uncertainties

The systematic uncertainties summarize the uncertainty on the Standard Model expectation arising from a possibly imprecise description of the data by the Monte Carlo simulation. Corresponding to the two separate stages in the Monte Carlo production, the event generation and detector simulation (*cf.* section 1.2.7), there exist two different kinds of systematic uncertainties: theoretical uncertainties that represent the effect of uncertainties in the modeling of the *ep* interaction process, and experimental uncertainties that represent the effect of uncertainties in the simulation of the detector response to the particles produced in the *ep* interaction. In the analysis presented in this thesis, the systematic uncertainties on the signal and background expectations are estimated differently.

The systematic uncertainty on the signal expectation is determined by separately estimating the theoretical and experimental uncertainties and adding them in quadrature. The theoretical uncertainty on the signal expectation is taken as the uncertainty estimated by theorists on the cross-section for signal processes. The experimental uncertainties are estimated by varying experimental quantities and recalculating the signal expectation after each variation [1]. The magnitude of the variation is determined by the level of agreement between data and Monte Carlo simulation in independent representative processes with sufficiently large event numbers [2].

The uncertainty on the background expectation is estimated directly from the data. It is estimated separately for individual background processes. For each background process, the systematic uncertainty is estimated by the level of agreement between the data and the Monte Carlo simulation for a background enriched sample in a phase-space similar to that defined by the final selection [3]. The selection criteria defining the enriched control samples are chosen such as to maximize the contribution of the background process under study and minimize that of other processes [4].

The estimation of the systematic uncertainties on the signal and background expectations is detailed separately in the following two subsections.

8.1.1 Signal Processes

The theoretical uncertainty on the signal expectation is taken to be the uncertainty on the cross-section for the production of real W bosons, within the Standard Model the dominant contribution to events with isolated leptons and large missing transverse momentum at HERA. The uncertainty on the cross-section has been estimated by theorists to be of the order of 15% (for the NLO calculation, *cf.* section 2.6.1) [5].

The experimental uncertainty on the signal expectation is estimated by variation of the quantities listed in the following:

Electron Identification and Reconstruction

The uncertainty on the electron identification efficiency is estimated from the study of electron pair production events presented in section C.1 of the appendix. The uncertainty attributed to the electron identification efficiency is 5%.

The uncertainty on the reconstructed electron four-vector is estimated using studies of NC DIS events [31]. The uncertainty attributed to the electron energy is determined by the uncertainty on the electromagnetic energy scale of the LAr calorimeter and varies between 0.7% in the backward region and 3% in the forward region, as shown in figure 8.1. The uncertainty is larger in the forward region, as a

[1] The recalculation of the signal expectation includes the redetermination of all quantities that are reconstructed from the varied experimental quantities; for example, the redetermination of $\zeta^2 = 2E_e^{beam} E_e\,(1 + \cos\theta_e)$ if either the electron energy E_e or polar angle θ_e is varied.

[2] The systematic uncertainties cannot be verified directly for signal processes, as their cross-sections are too small for sufficiently large event numbers to be obtained.

[3] The resulting estimates may exceed the sum of the uncertainties estimated by theorists on the cross-section and the experimental uncertainties: as only a very small fraction of background events contributed to the phase-phase in which the signal is searched for, the theoretical uncertainties on the background expectation in that "exotic" phase-space may well be significantly larger than the uncertainty on the cross-section.

[4] The cross-sections for most background processes are more than large enough for sufficiently large event numbers to be obtained in the enriched control samples.

[5] The higher uncertainty on the cross-section for the production of real Z bosons (that is calculated at LO only) may be neglected, as the contribution from Z boson production amounts to less than 5% of the contribution from W boson production (*cf.* section 2.6.2).

consequence of the small number of scattered electrons with which the agreement between data and Monte Carlo simulation can be verified in that region. The un-

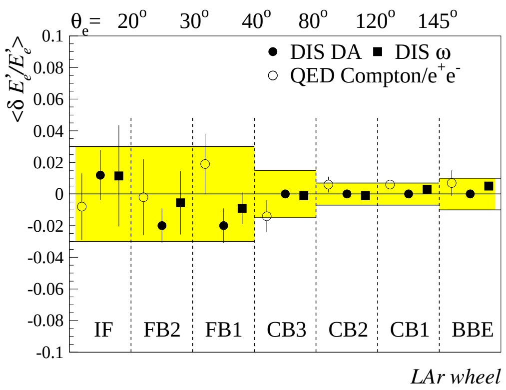


Figure 8.1: *Uncertainty on the electromagnetic energy scale in the different wheels of the LAr calorimeter. The IF is the most forward and the BBE the most backward wheel. The agreement between data and Monte Carlo is studied in samples of NC DIS and QED Compton events. In both processes, the kinematics are overconstrained, such that the electron energy can be reconstructed with sufficient precision independently of the energy of the electromagnetic shower measured in the LAr. The points represent the difference between the kinematically reconstructed electron energy and the energy of the electromagnetic shower measured in the LAr (the kinematics of NC DIS events are reconstructed using two different methods, the so called Double-Angle and ω methods; for more details see reference [31]); the uncertainty attributed to the electromagnetic energy scale is indicated by the grey band (taken from [31]).*

certainty attributed to the polar angle is 1 $mrad$, 2 $mrad$ and 3 $mrad$ for electrons of polar angles $\theta_e > 135°$, $120° < \theta_e < 135°$ and $\theta_e < 120°$, respectively. The uncertainty attributed to the azimuthal angle is 1 $mrad$, independent of the polar angle.

An additional uncertainty of 3% is attributed to the cluster-track link requirement for the electron [5].

Muon Identification and Reconstruction

The uncertainty on the muon identification efficiency is estimated from the studies of muon pair production events presented in section C.2 of the appendix. The uncertainty attributed to the muon identification efficiency is 6%.

The uncertainty on the reconstructed muon momentum is estimated using studies of the transverse momentum balance in muon pair production events and is taken to be 5% [160]. The uncertainties on the polar and azimuthal angles of muons are assumed to be similar to those of electrons. An uncertainty of 3 $mrad$ is attributed to the polar angle and 1 $mrad$ to the azimuthal angle of muons.

An additional uncertainty of 5% is attributed to the track-link requirement in the muon triggered phase-space. This uncertainty is estimated by comparing the link probability for isolated muons in the Monte Carlo simulation with that for cosmic muons in the data [6]. The results are shown in figure 8.2.

[6] The number of muon pair production events is not sufficient to allow these to be used for this

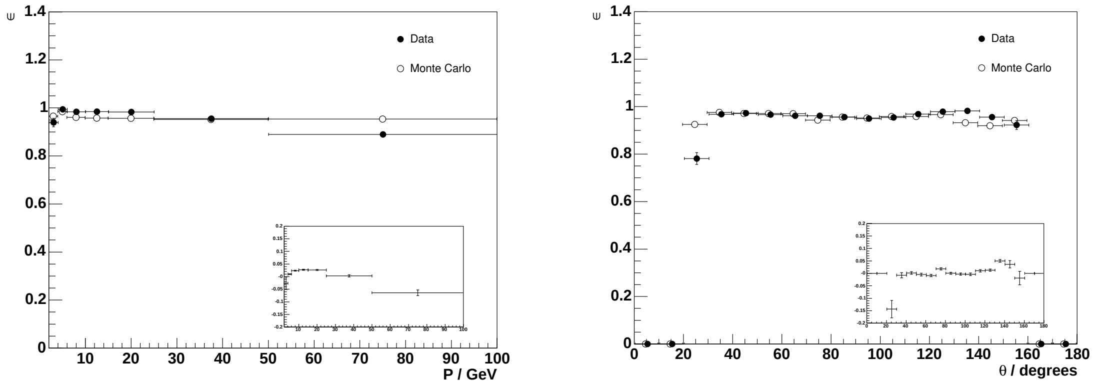


Figure 8.2: *Probability to link an extrapolated central track with a track reconstructed in the instrumented iron for muons in the geometric acceptance of the central tracking detector. The link probability is estimated separately for the data (solid circles) and the Monte Carlo simulation (open circles). In the Monte Carlo simulation, the link probability is determined using generated events with single muons (the H1SIM software package allows for the "inline" generation of single particles for detector studies) and is parametrized as a function of the momentum and polar angle of the generated muon. In the data, the link probability is determined using cosmic muon events. The events are selected by requiring two central tracks with a small distance of closest approach to the nominal interaction point and two tracks "back-to-back" in polar and azimuthal angle reconstructed in the instrumented iron. In addition, a track-link is required for the central track in the hemisphere with positive azimuthal angle, so as to guarantee that the momentum is well measured. The link probability is then estimated as the fraction of central tracks in the hemisphere with negative azimuthal angle that are linked to an iron track and is parametrized as a function of the momentum and polar angle of the central track in the hemisphere with positive azimuthal angle.*

Hadronic Reconstruction

The uncertainty on the energy of hadrons is estimated using the transverse momentum balance between the scattered electron and the hadronic final state in NC DIS events [32] as shown in figure 8.3 [7]. In the figure, the data are seen to agree with the Monte Carlo at the 2% level. In order to take the extrapolation into the different phase-space region in which the signal is searched for into account, the difference between the data and the Monte Carlo simulation seen in the NC DIS events is doubled, and an uncertainty of 4% is attributed to the energy of hadrons.

An uncertainty of 20 *mrad* is attributed to the polar and azimuthal angles of the hadronic system respectively [5].

Tau Identification and Reconstruction

The uncertainty on the identification efficiency for hadronic tau decays is determined by the uncertainties on the track requirements described in section 4.5.2

study at high momentum.

[7] The uncertainty on the energy of hadrons is determined by the electromagnetic and hadronic energy scales of the LAr calorimeter and by the track reconstruction in the central and forward tracking detectors (*cf.* the description of the hadronic reconstruction in section 4.6).

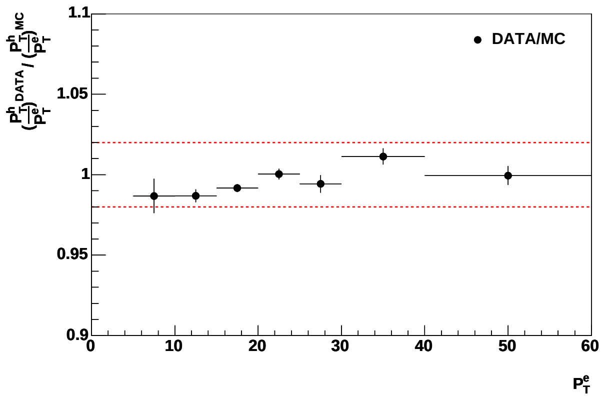


Figure 8.3: *Uncertainty on the energy of hadrons. The agreement between data and Monte Carlo is studied with the transverse momentum balance between the scattered electron and the hadronic system in NC DIS events.*

and the requirements on the classification output of the neural networks described in section 4.5.5. The uncertainty attributed to the efficiency of the track requirement is 4%. The uncertainty on the efficiency of the requirement $f_{NN} > 0.75$ on the neural network output is estimated as the difference between the percentages of tau jets simulated with the fast and detailed detector simulations (*cf.* section B of the appendix) that pass this requirement [8]. The difference in efficiency is mainly due to differences in the electromagnetic shower development (that results in more or less collimated showers) implemented in the fast and detailed simulations and is estimated to be 5% for hadronic "one-prong" tau decays.

The uncertainties on the energy and the polar and azimuthal angles of tau jets are assumed to be similar to those of QCD jets. The uncertainty attributed to the energy of tau jets is 4%. The uncertainties attributed to the polar and azimuthal angles are 5 $mrad$ for $\theta_\tau < 30°$ and 10 $mrad$ for $\theta_\tau > 30°$ [6].

Non-*ep* Background Rejection

An uncertainty of 2% is attributed to the efficiency of the requirements on the event timing and topological non-*ep* background finders described in sections 6.3.1 and 6.3.2. The value of 2% corresponds to about the inefficiency of these requirements in the Monte Carlo simulation and is assumed to be a conservative estimate.

Trigger Efficiency

The uncertainties on the trigger efficiency are different for events with isolated electrons, muons and tau leptons, as different subtriggers are used to trigger these events (*cf.* section 6.4).

The events with isolated electrons are triggered by the presence of compact energy deposits in the electromagnetic section of the LAr calorimeter. The efficiency of these LAr electron triggers is very high and the uncertainty small. The uncertainty on the trigger efficiency for events with an electron in the LAr is estimated to be 0.5% from studies of NC DIS events [31].

[8]The tau jets are simulated in events with single hadronic tau decays by using the "inline" generator of the H1SIM software package.

P_T^X $GeV]$	δ_{ε_e} [%]	δ_{E_e} [%]	δ_{θ_e} [%]	δ_{ϕ_e} [%]	$\delta_{\varepsilon_{tr}}$ [%]	δ_{ε_μ} [%]	δ_{P_μ} [%]	δ_{θ_μ} [%]	δ_{ϕ_μ} [%]	$\delta_{\varepsilon_\tau}$ [%]	δ_{E_τ} [%]	δ_{θ_τ} [%]	δ_{ϕ_τ} [%]
< 25	3.2	0.4	0.1	0.0	1.9	1.8	0.1	0.0	0.0	0.3	0.1	0.0	0.0
> 25	2.3	0.3	0.2	0.0	1.4	2.9	0.3	0.0	0.0	0.3	0.1	0.0	0.0
All P_T^X	3.0	0.4	0.1	0.0	1.8	2.1	0.1	0.0	0.0	0.3	0.1	0.0	0.0

P_T^X $GeV]$	δ_{E_h} [%]	δ_{γ_h} [%]	δ_{ϕ_h} [%]	δ_{bg} [%]	δ_{trig} [%]	δ_{lumi} [%]	δ_{sys} [%]
< 25	0.9	2.2	0.1	2.0	2.3	1.5	7.6
> 25	5.6	10.5	0.2	2.0	1.5	1.5	14.1
All P_T^X	2.1	4.2	0.1	2.0	2.1	1.5	9.1

Table 8.1: *Contributions of individual quantities to the systematic uncertainty δ_{sys} on the signal expectation. The individual quantities are: the electron identification efficiency ε_e, the energy E_e, the polar and azimuthal angles θ_e and ϕ_e, and the cluster-track requirement of electrons; the muon identification efficiency ε_μ, the momentum P_μ, and the polar and azimuthal angles θ_μ and ϕ_μ of muons; the tau identification efficiency ε_τ, the energy E_τ, and the polar and azimuthal angles θ_τ and ϕ_τ of tau jets; the energy E_h and the polar and azimuthal angles θ_h and ϕ_h of the hadronic system; the efficiency ε_{bg} of the requirements for non-ep background rejection; the trigger efficiency ε_{tr}; and the luminosity measurement.*

The events with isolated muons are triggered by the presence of either hits in the instrumented iron or an imbalance of the energy deposits in the LAr calorimeter. The uncertainty increases for events with small imbalances P_T^{calo}, for which the trigger efficiency varies significantly (see figure 6.3). The uncertainty on the efficiency ε of the combination of muon and LAr triggers is parametrized as $2\% \oplus (1-\varepsilon) \cdot 30\%$ [9] [5].

The events with isolated tau leptons are triggered by the presence of an imbalance of the energy deposits in the LAr calorimeter. The uncertainty on the efficiency of the LAr triggers is estimated by the "pseudo-charged current" technique in samples of NC DIS events (*cf.* section 6.4.1) and is parametrized as $2\% \oplus (1-\varepsilon) \cdot 15\%$ [31].

Luminosity

An uncertainty of 1.5% is attributed to the luminosity measurement [67].

The effect on the signal expectation of varying the individual experimental quantities by their estimated uncertainties are summarized in table 8.1. In comparison to the theoretical uncertainty of 15% on the cross-section, the contribution of the experimental uncertainties to the systematic uncertainty of the signal expectation is relatively small.

8.1.2 Background Processes

The systematic uncertainty on the background expectation is estimated separately for events with electrons, muons and tau leptons and for the individual background processes. For each background process, the level of agreement between the data and the Monte Carlo simulation is studied using background enriched control samples

[9] The symbol $\oplus$ means "add in quadrature".

in the enlarged phase-spaces described in sections 7.1.1, 7.2.1, 7.2.2, and 7.3.1 [10]. The selection criteria defining the background enriched control samples are listed in table 8.2.

The numbers of events selected in the background enriched control samples in the 1994-2000 e^+p and e^-p datasets are compared with the expectations for individual background processes in tables 8.3 and 8.4. In all background enriched control samples (except the photoproduction enriched control sample in the enlarged phase-space for events with isolated tau leptons), the enriched process dominates the expectation. The level of agreement between the data and the Monte Carlo simulation is of the order of 15%. The difference is mainly in the overall scale: for all background enriched control samples, the expected distributions of the quantities defined in chapter 5 have been compared with the data and no significant deviations other than an overall scale factor have been found [11].

A systematic uncertainty of 30% is attributed to the background expectation in the final samples of events with isolated electrons, muons and tau leptons [12]. This is twice the uncertainty associated with the level of agreement observed between the data and the Monte Carlo simulation in the background enriched control samples. The uncertainty is doubled in order to account for the additional errors arising from the extrapolation into the more "exotic" phase-space of the final event sample.

8.2 Events with Isolated Electrons

Using the selection criteria listed in table 7.2, 11 events with isolated electrons and large missing transverse momentum are selected in the 1994-2000 e^+p and e^-p datasets. The number of events selected in the data is compared with the Standard Model expectation in four bins of the hadronic transverse momentum P_T^X in table 8.5. In total, 11.5 events with isolated electrons and large missing transverse momentum are expected from all Standard Model processes. The dominant contribution is expected from signal processes: of the 11.5 events, 7.5 events are expected from W production and 4.0 events from other Standard Model processes. The background contribution is mainly due to NC and CC DIS processes and is concentrated

[10] The background contributions from CC DIS processes to the enlarged phase-space of events with muons and from lepton pair production processes to the enlarged phase-space of events with isolated tau leptons are too small to be directly verified with sufficiently large numbers of events. In the enlarged phase-space of events with electrons, the background contributions from photoproduction and lepton pair production processes are too small in comparison to the contribution from NC DIS processes to be separately verified.

[11] The distributions are shown in section F of the appendix.

[12] The background contribution to the final sample of events with isolated tau leptons is attributed an additional uncertainty of 20% arising from the uncertainty on the percentage of QCD jets which pass the requirement $f_{NN} > 0.75$ on the neural network output in the final event selection. This corresponds to the difference between the percentages of QCD jets simulated with the fast and detailed detector simulations and different hadronisation models that pass the $f_{NN} > 0.75$ requirement (*cf.* section B of the appendix). The additional uncertainty of 20% is added in quadrature to the systematic uncertainty estimated using the background enriched control samples, as in the background enriched control samples, the tau jet candidates are not required to pass any requirements on the neural network output (*cf.* section 7.3.1). The effect of this additional uncertainty is small, however.

Enriched Selection	Requirements
Events with Isolated Electrons[1]	
NC	Electron flagged as "isolated lepton" $D_{track} > 0.5$ for $\theta_e > 45°$ $D_{jet} > 1.0$
CC	$\Delta\phi_{e-X} < 160°$ $V_{ap}/V_p < 0.15$ $\zeta^2 > 2500\ GeV^2$ $\delta_{miss} > 5\ GeV$
Events with Isolated Muons (in LAr triggered phase-space[2])	
NC	One electron
$\ell\bar{\ell}$	Muon flagged as "isolated lepton" $D_{track} > 0.5$ $D_{jet} > 1.0$ $V_{ap}/V_p < 0.2$
γp	No electron
Events with Isolated Muons (in muon triggered phase-space[3])	
NC	One electron Does not pass NC enriched selection in LAr triggered phase-space
$\ell\bar{\ell}$	Muon flagged as "isolated lepton" $D_{track} > 0.5$ $D_{jet} > 1.0$ $V_{ap}/V_p < 0.4$ Does not pass $\ell\bar{\ell}$ enriched selection in LAr triggered phase-space
γp	No electron Does not pass γp enriched selection in LAr triggered phase-space
Events with Isolated Tau Leptons[4]	
NC	One electron
γp	No electron $\delta_{miss} > 5\ GeV$
CC	$V_{ap}/V_p < 0.15$ $\delta_{miss} > 5\ GeV$

[1] In the enlarged phase-space defined by the selection criteria listed in table 7.1
[2] Defined by the selection criteria listed in table 7.3
[3] Defined by the selection criteria listed in table 7.5
[4] In the enlarged phase-space defined by the selection criteria listed in table 7.7

Table 8.2: *Selection criteria defining the background enriched control samples. The listed criteria are required in addition to those given in tables 7.1, 7.3, 7.5, and 7.7.*

at low P_T^X [1]. The signal contribution becomes more dominant at high P_T^X. In total,

[1] The contributions from individual background processes to the final sample of events with isolated electrons are tabulated in table H.1 in section H.1 of the appendix.

Enriched Selection	Data	W	NC	$\ell\bar{\ell}$	γp	CC
Events with Isolated Electrons						
NC	3082	16.7 ± 0.1	2516.1 ± 11.4	59.2 ± 0.5	3.3 ± 0.5	3.5 ± 0.1
CC	27	7.2 ± 0.1	0.4 ± 0.1	0.7 ± 0.1	0.3 ± 0.1	19.4 ± 0.3
Events with Isolated Muons in LAr Triggered Phase-Space						
NC	72	1.3 ± 0.0	46.5 ± 7.0	16.6 ± 0.2	3.8 ± 0.4	0.1 ± 0.0
$\ell\bar{\ell}$	62	3.1 ± 0.0	5.0 ± 1.6	56.7 ± 0.4	2.0 ± 0.5	0.0 ± 0.0
γp	66	2.4 ± 0.0	7.4 ± 2.7	12.6 ± 0.2	38.2 ± 1.5	6.2 ± 0.2
Events with Isolated Muons in Muon Triggered Phase-Space						
NC	52	0.4 ± 0.0	40.1 ± 5.6	19.2 ± 0.4	9.1 ± 2.4	0.0 ± 0.0
$\ell\bar{\ell}$	90	1.8 ± 0.0	2.0 ± 0.8	92.8 ± 0.9	3.3 ± 1.0	0.0 ± 0.0
γp	72	1.8 ± 0.0	3.5 ± 1.7	22.9 ± 0.4	41.4 ± 4.9	0.3 ± 0.0
Events with Isolated Tau Leptons						
NC	70	0.9 ± 0.0	55.7 ± 6.5	1.3 ± 0.0	7.6 ± 0.9	8.5 ± 0.2
γp	276	2.1 ± 0.0	8.6 ± 1.9	1.1 ± 0.0	122.2 ± 12.0	131.9 ± 0.8
CC	150	2.3 ± 0.0	0.1 ± 0.1	0.7 ± 0.0	8.2 ± 0.9	127.1 ± 0.8

Table 8.3: *Number of events in the 1994-2000 e^+p and e^-p datasets in the background enriched control samples defined by the criteria listed in table 8.2 compared with the Monte Carlo expectations for the individual processes. The quoted uncertainties represent the statistical uncertainties of the Monte Carlo samples.*

the number of events observed in the data is in agreement with the Standard Model expectation. In the region of large hadronic transverse momenta, an excess of events is observed, however: in the data, 2 events are observed with $P_T^X > 40\ GeV$, where only 0.7 are expected from all Standard Model processes.

For the selected events, distributions of the polar angle and transverse momentum of the isolated electron, the missing transverse momentum, the acoplanarity angle, and the transverse momentum of the hadronic system are shown in figure 8.4. Also shown in the figure is the distribution of the transverse mass of the hypothetical electron-neutrino system reconstructed in the selected events.

8.3 Events with Isolated Muons

The results of the search for events with isolated muons and large missing transverse momentum are presented in three subsections. In the first two subsections, the results for the LAr triggered and the muon triggered phase-spaces are presented separately, so as to facilitate the comparison with previous analyses that are restricted to the LAr triggered phase-space. The combined results for the LAr and muon triggered phase-space are presented in the third subsection.

8.3.1 LAr Triggered Phase-Space

Using the selection criteria listed in table 7.4, 8 events with isolated muons and large missing transverse momentum are selected in the LAr triggered phase-space in

Enriched Selection	Data	SM Expectation	$\frac{\text{Data - SM Expectation}}{\text{SM Expectation}}$
Events with Isolated Electrons			
NC	3082	2593.7 ± 11.4	$18.8 \pm 2.2\%$
CC	27	27.9 ± 0.4	$-3.2 \pm 18.7\%$
Events with Isolated Muons (in LAr triggered phase-space)			
NC	72	68.3 ± 7.0	$5.4 \pm 16.5\%$
$\ell\bar{\ell}$	62	66.8 ± 1.7	$-7.2 \pm 12.0\%$
γp	66	66.7 ± 3.1	$-1.0 \pm 13.0\%$
Events with Isolated Muons (in muon triggered phase-space)			
NC	52	68.9 ± 6.1	$-24.5 \pm 12.4\%$
$\ell\bar{\ell}$	90	99.9 ± 1.6	$-9.9 \pm 9.6\%$
γp	72	69.8 ± 5.3	$3.2 \pm 14.5\%$
Events with Isolated Tau Leptons			
NC	70	74.1 ± 6.5	$-5.5 \pm 14.0\%$
γp	276	266.0 ± 12.2	$3.8 \pm 7.9\%$
CC	150	138.5 ± 1.2	$8.3 \pm 8.9\%$

Table 8.4: *Number of events in the 1994-2000 e^+p and e^-p datasets in the background enriched control samples compared with the Monte Carlo expectation for all Standard Model processes. The uncertainties quoted on the difference between the data and the Monte Carlo expectation represent the statistical uncertainties on the number of events selected in the data and Monte Carlo samples.*

the 1994-2000 e^+p and e^-p datasets. The number of events selected in the data is compared with the Standard Model expectation in four bins of the hadronic transverse momentum P_T^X in table 8.6. In the LAr triggered phase-space, 2.8 events with isolated muons and large missing transverse momentum are expected from all Standard Model processes. The dominant contribution is expected from signal processes: of the 2.8 events, 2.4 events are expected from W production and 0.4 events from other Standard Model processes. The background contribution is mainly due to lepton pair production processes [1]. The contributions from signal and background processes are restricted to the region $P_T^X > 12\ GeV$ by the selection criteria (*cf.* table 7.4).

8.3.2 Muon Triggered Phase-Space

Using the selection criteria listed in table 7.6, 0 events with isolated muons and large missing transverse momentum are selected in the muon triggered phase-space in the 1994-2000 e^+p and e^-p datasets. The number of events selected in the data is compared with the Standard Model expectation in four bins of the hadronic

[1]The contributions from individual background processes to the final sample of events with isolated muons in the LAr triggered phase-space are tabulated in table H.2 in section H.2.1 of the appendix.

P_T^X [GeV]	Data	SM Expectation	SM Signal	Other SM Processes
		e^+p collisions		
< 12	6	6.66 ± 0.85	4.12 ± 0.69	2.54 ± 0.50
12 - 25	1	1.61 ± 0.23	1.31 ± 0.21	0.30 ± 0.09
25 - 40	1	1.03 ± 0.17	0.75 ± 0.14	0.29 ± 0.09
> 40	2	0.57 ± 0.15	0.47 ± 0.14	0.10 ± 0.04
All P_T^X	10	9.87 ± 1.36	6.65 ± 1.18	3.23 ± 0.65
		e^-p collisions		
< 12	0	1.13 ± 0.17	0.55 ± 0.09	0.58 ± 0.14
12 - 25	1	0.29 ± 0.05	0.18 ± 0.03	0.11 ± 0.04
25 - 40	0	0.16 ± 0.03	0.10 ± 0.02	0.05 ± 0.02
> 40	0	0.09 ± 0.03	0.07 ± 0.02	0.03 ± 0.02
All P_T^X	1	1.66 ± 0.25	0.89 ± 0.16	0.77 ± 0.19
		ep collisions		
< 12	6	7.79 ± 1.01	4.67 ± 0.78	3.12 ± 0.62
12 - 25	2	1.90 ± 0.27	1.49 ± 0.24	0.41 ± 0.11
25 - 40	1	1.19 ± 0.19	0.85 ± 0.16	0.34 ± 0.10
> 40	2	0.66 ± 0.17	0.53 ± 0.16	0.13 ± 0.05
All P_T^X	11	11.53 ± 1.60	7.54 ± 1.34	3.99 ± 0.82

Table 8.5: *Number of events with isolated electrons and large missing transverse momentum selected using the criteria listed in table 7.2 in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background contributions. The uncertainties quoted on the expectation are calculated by adding the statistical and systematic uncertainties in quadrature.*

transverse momentum P_T^X in table 8.7. In the muon triggered phase-space, 3.4 events with isolated muons and large missing transverse momentum are expected from all Standard Model processes. The dominant contribution is expected from signal processes: of the 3.4 events, 1.9 events are expected from W production and 1.5 events from other Standard Model processes. The background contribution is mainly due to lepton pair production processes [2]. The contributions from signal and background processes are essentially restricted to the region $P_T^X < 12$ GeV by the requirement that the events not be found in the LAr triggered phase-space (*cf.* table 7.6).

8.3.3 Combined Phase-Space

In the combined LAr and muon triggered phase-space, 8 events with isolated muons and large missing transverse momentum are selected in the 1994-2000 e^+p and e^-p datasets. The number of events selected in the data is compared with the Standard Model expectation in four bins of the hadronic transverse momentum P_T^X in table 8.8. In total, 6.1 events with isolated muons and large missing transverse momentum

[2]The contributions from individual background processes to the final sample of events with isolated muons in the muon triggered phase-space are tabulated in table H.3 in section H.2.2 of the appendix.

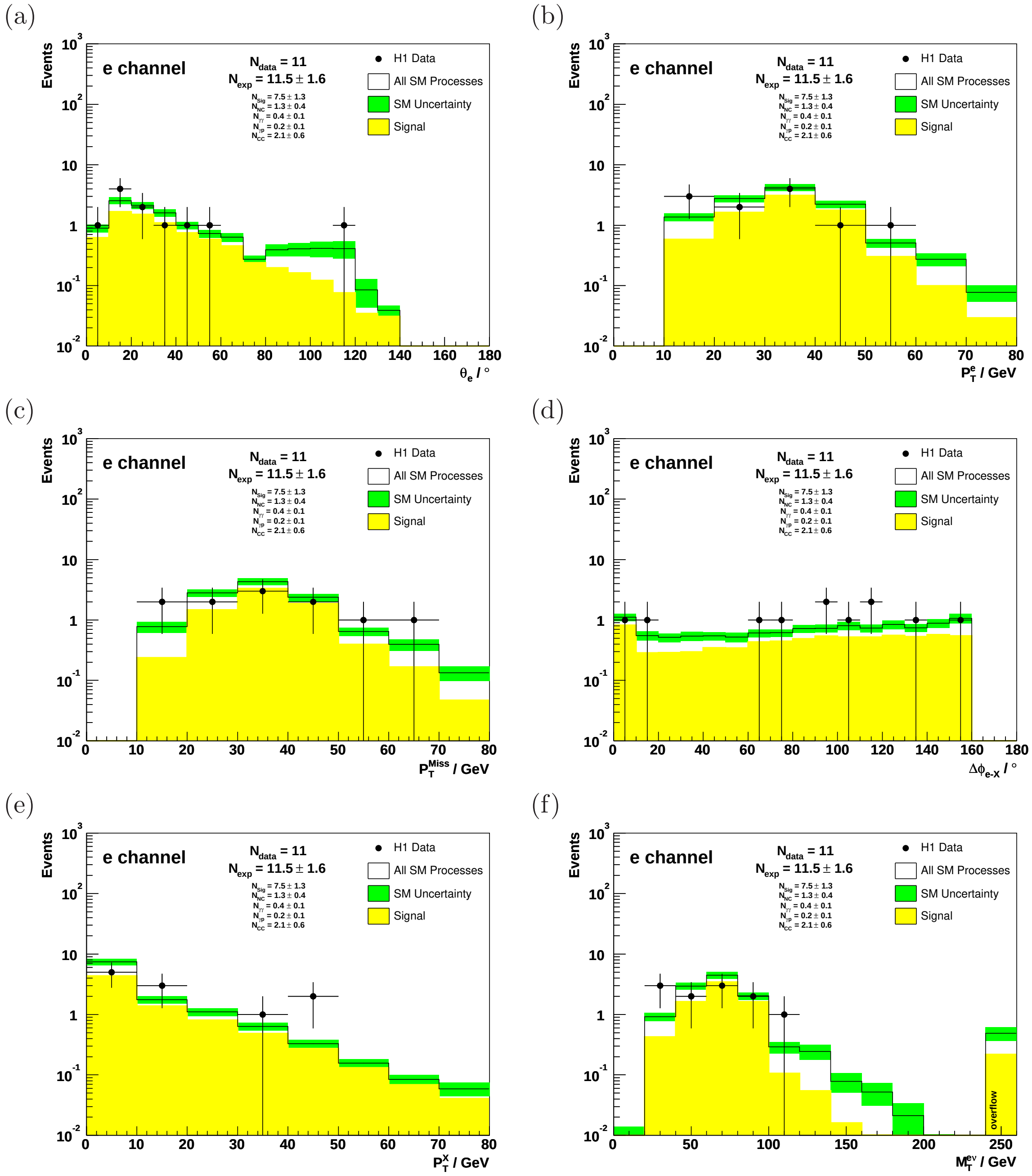


Figure 8.4: *Distributions of the polar angle (a) and the transverse momentum (b) of the isolated electron, the missing transverse momentum (c), the acoplanarity angle (d), the transverse momentum of the hadronic system (e) and the transverse mass of the hypothetical electron-neutrino system (f) in the final sample of events with isolated electrons and large missing transverse momentum.*

are expected in the combined LAr and muon triggered phase-space from all Standard Model processes. The dominant contribution is expected from signal processes: of the 6.1 events, 4.3 events are expected from W production and 1.8 events from other Standard Model processes. The background contribution is mainly due to lepton

P_T^X $[GeV]$	Data	SM Expectation	SM Signal	Other SM Processes
		e^+p collisions		
< 12	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
12 - 25	3	1.02 ± 0.14	0.88 ± 0.13	0.14 ± 0.04
25 - 40	2	0.84 ± 0.12	0.72 ± 0.11	0.12 ± 0.04
> 40	3	0.58 ± 0.08	0.52 ± 0.08	0.06 ± 0.02
All P_T^X	8	2.44 ± 0.33	2.12 ± 0.32	0.32 ± 0.09
		e^-p collisions		
< 12	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
12 - 25	0	0.12 ± 0.02	0.11 ± 0.02	0.01 ± 0.01
25 - 40	0	0.11 ± 0.02	0.09 ± 0.01	0.01 ± 0.00
> 40	0	0.08 ± 0.01	0.07 ± 0.01	0.01 ± 0.00
All P_T^X	0	0.31 ± 0.04	0.27 ± 0.04	0.03 ± 0.01
		ep collisions		
< 12	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
12 - 25	3	1.14 ± 0.16	0.99 ± 0.15	0.15 ± 0.04
25 - 40	2	0.95 ± 0.13	0.81 ± 0.12	0.13 ± 0.04
> 40	3	0.66 ± 0.09	0.59 ± 0.09	0.07 ± 0.02
All P_T^X	8	2.75 ± 0.38	2.40 ± 0.36	0.35 ± 0.10

Table 8.6: *Number of events with isolated muons and large missing transverse momentum in the LAr triggered phase-space in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background components. The uncertainties quoted on the expectation are calculated by adding the statistical and systematic uncertainties in quadrature.*

pair production processes and is concentrated at low P_T^X [3]. The signal contribution becomes more dominant at high P_T^X. In total, the number of events observed in the data is compatible with the Standard Model expectation. However, in the region of large hadronic transverse momenta, an excess of events is observed: all 8 events are observed with $P_T^X > 12\ GeV$, where only 3.0 are expected from all Standard Model processes. The observation of no events with $P_T^X < 12\ GeV$ is compatible with the expectation within the estimated statistical and systematic uncertainties.

For the selected events, distributions of the polar angle and transverse momentum of the isolated muon, the missing transverse momentum, the acoplanarity angle, and the transverse momentum of the hadronic system are shown in figure 8.5. Also shown in the figure is the distribution of the transverse mass of the hypothetical muon-neutrino system reconstructed in the selected events.

[3]The contributions from individual background processes to the final sample of events with isolated muons in the combined LAr and muon triggered phase-space are tabulated in table H.4 in section H.2.3 of the appendix.

P_T^X $[GeV]$	Data	SM Expectation	SM Signal	Other SM Processes
		e^+p collisions		
< 12	0	2.78 ± 0.43	1.51 ± 0.23	1.27 ± 0.36
12 - 25	0	0.17 ± 0.03	0.15 ± 0.03	0.01 ± 0.01
25 - 40	0	0.03 ± 0.01	0.03 ± 0.01	0.01 ± 0.00
> 40	0	0.01 ± 0.00	0.01 ± 0.00	0.00 ± 0.00
All P_T^X	0	2.98 ± 0.46	1.70 ± 0.26	1.29 ± 0.37
		e^-p collisions		
< 12	0	0.38 ± 0.06	0.20 ± 0.03	0.18 ± 0.05
12 - 25	0	0.02 ± 0.00	0.02 ± 0.00	0.00 ± 0.00
25 - 40	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
All P_T^X	0	0.40 ± 0.06	0.22 ± 0.03	0.18 ± 0.05
		ep collisions		
< 12	0	3.15 ± 0.49	1.71 ± 0.26	1.44 ± 0.41
12 - 25	0	0.18 ± 0.03	0.17 ± 0.03	0.02 ± 0.01
25 - 40	0	0.04 ± 0.01	0.03 ± 0.01	0.01 ± 0.00
> 40	0	0.01 ± 0.00	0.01 ± 0.00	0.00 ± 0.00
All P_T^X	0	3.38 ± 0.52	1.92 ± 0.29	1.47 ± 0.42

Table 8.7: *Number of events with isolated muons and large missing transverse momentum in the muon triggered phase-space in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background components. The uncertainties quoted on the expectation are calculated by adding the statistical and systematic uncertainties in quadrature.*

8.4 Events with isolated Tau Leptons

Using the selection criteria listed in table 7.8, 0 events with isolated tau leptons and large missing transverse momentum are selected in the 1994-2000 e^+p and e^-p datasets. The number of events selected in the data is compared with the Standard Model expectation in four bins of the hadronic transverse momentum P_T^X in table 8.9. In total, 2.3 events with isolated tau leptons and large missing transverse momentum are expected from all Standard Model processes. Of the 2.3 events, 0.8 events are expected from W production and 1.5 events from other Standard Model processes. The background contribution is mainly due to CC DIS processes and is concentrated at low P_T^X [1]. The signal contribution becomes dominant at high P_T^X. The observation of no events in the data is compatible with the expectation within the estimated statistical and systematic uncertainties. The contributions from signal and background processes to the Standard Model expectation are summarized in table 8.9.

Distributions of the polar angle and transverse momentum of the tau jet, the missing transverse momentum, the acoplanarity angle, and the transverse momentum of the hadronic system are shown in figure 8.6. Also shown in the figure is the

[1]The contributions from individual background processes to the final sample of events with isolated tau leptons are tabulated in table H.5 in section H.3 of the appendix.

P_T^X $[GeV]$	Data	SM Expectation	SM Signal	Other SM Processes
		e^+p collisions		
< 12	0	2.78 ± 0.63	1.51 ± 0.43	1.27 ± 0.46
12 - 25	3	1.18 ± 0.24	1.04 ± 0.23	0.15 ± 0.05
25 - 40	2	0.87 ± 0.17	0.75 ± 0.16	0.13 ± 0.04
> 40	3	0.59 ± 0.17	0.52 ± 0.17	0.07 ± 0.02
All P_T^X	8	5.42 ± 1.20	3.82 ± 0.99	1.60 ± 0.56
		e^-p collisions		
< 12	0	0.38 ± 0.09	0.20 ± 0.06	0.18 ± 0.06
12 - 25	0	0.14 ± 0.03	0.13 ± 0.03	0.02 ± 0.01
25 - 40	0	0.11 ± 0.02	0.10 ± 0.02	0.02 ± 0.01
> 40	0	0.08 ± 0.02	0.07 ± 0.02	0.01 ± 0.00
All P_T^X	0	0.71 ± 0.16	0.49 ± 0.14	0.22 ± 0.08
		ep collisions		
< 12	0	3.15 ± 0.71	1.71 ± 0.48	1.44 ± 0.52
12 - 25	3	1.33 ± 0.27	1.16 ± 0.27	0.16 ± 0.05
25 - 40	2	0.98 ± 0.19	0.84 ± 0.18	0.14 ± 0.04
> 40	3	0.67 ± 0.19	0.60 ± 0.19	0.07 ± 0.02
All P_T^X	8	6.13 ± 1.36	4.31 ± 1.12	1.82 ± 0.64

Table 8.8: *Number of events with isolated muons and large missing transverse momentum in the combined LAr and muon triggered phase-space in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background components. The uncertainties quoted on the expectation are calculated by adding the statistical and systematic uncertainties in quadrature.*

distribution of the transverse mass reconstructed for the tau jet and the missing momentum vector in the selected events. Note that the distribution of the transverse mass reconstructed for signal events is shifted towards lower values in comparison to the transverse mass distributions of the hypothetical electron-neutrino and muon-neutrino systems shown in figures 8.4 and 8.5, due to the presence of an additional neutrino produced in the tau decay.

8.5 Compatibility of Different Channels

The compatibility of the isolated electron, muon and tau lepton results with lepton universality is checked by estimating the probability for the differences in the number of events observed in the individual channels to be due to statistical fluctuations, taking into account the different selection efficiencies, background contributions and systematic uncertainties. The probability of the observation in all channels is given by the product

$$\mathcal{P} \equiv \prod_i P\left(N^i_{obs} | N^i_{exp}, \sigma^i_{exp}\right)$$

over the i individual channels, where P denotes the probability to observe N^i_{obs} events in channel i when N^i_{exp} are expected. In the case of non-zero uncertainties

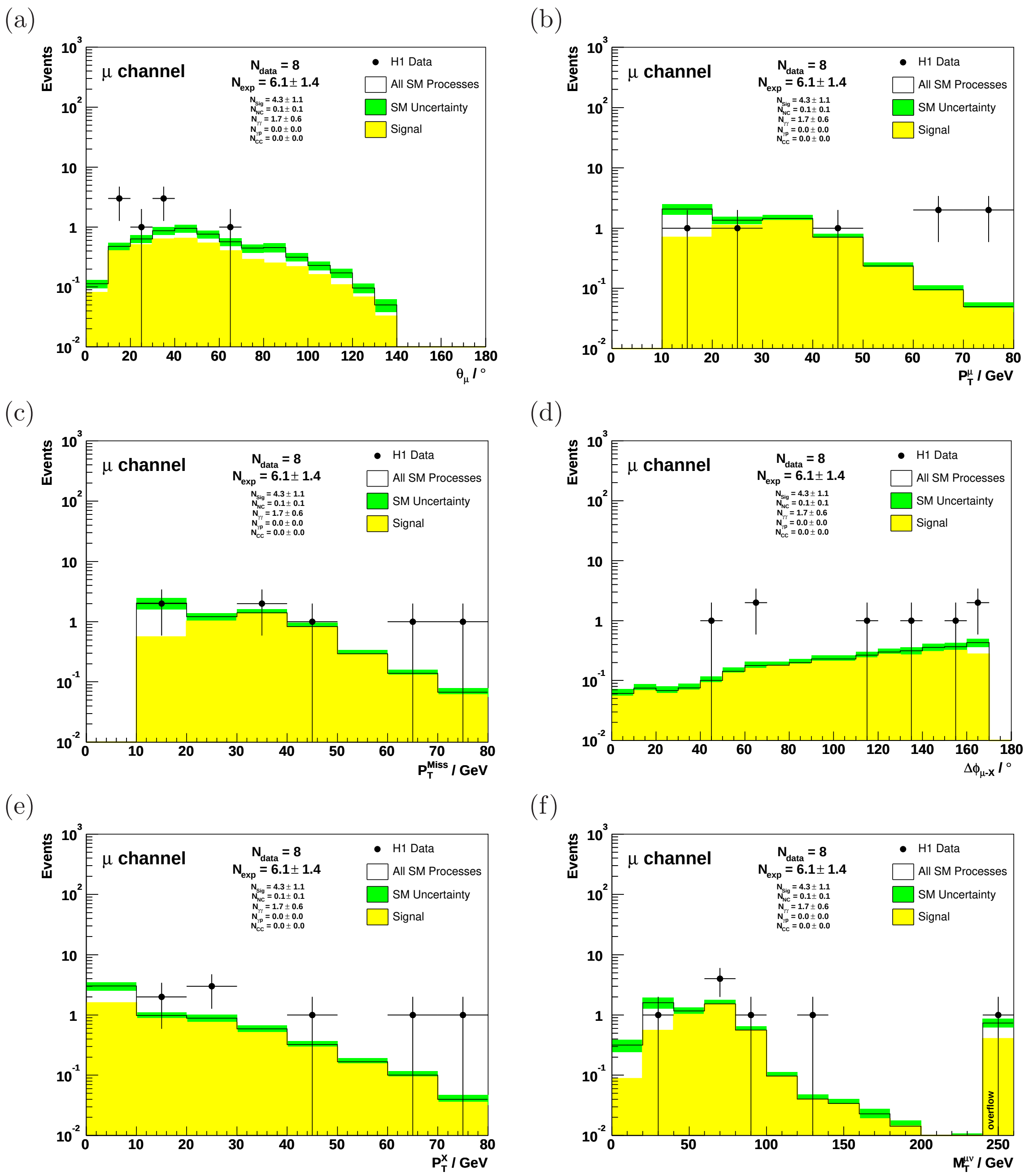


Figure 8.5: *Distributions of the polar angle (a) and the transverse momentum (b) of the isolated muon, the missing transverse momentum (c), the acoplanarity angle (d), the transverse momentum of the hadronic system (e) and the transverse mass of the hypothetical muon-neutrino system (f) in the final sample of events with isolated muons and large missing transverse momentum in the combined LAr and muon triggered phase-space.*

σ^i_{exp} on the expectation N^i_{exp} [1], P is not simply the Poisson probability

$$P_{poisson}\left(N_{obs}|N_{exp}\right) = e^{-N_{exp}} \frac{(N_{exp})^{N_{obs}}}{N_{obs}!},$$

[1]The uncertainties σ^i_{exp} represent the combined statistical and systematic uncertainties on the number N^i_{exp} of expected events.

P_T^X $[GeV]$	Data	SM Expectation	SM Signal	Other SM Processes
		***e^+p* collisions**		
< 12	0	1.54 ± 0.46	0.41 ± 0.14	1.13 ± 0.44
12 - 25	0	0.16 ± 0.05	0.12 ± 0.05	0.04 ± 0.02
25 - 40	0	0.12 ± 0.04	0.10 ± 0.04	0.02 ± 0.01
> 40	0	0.08 ± 0.05	0.07 ± 0.05	0.01 ± 0.01
All P_T^X	0	1.90 ± 0.60	0.70 ± 0.28	1.20 ± 0.46
		***e^-p* collisions**		
< 12	0	0.35 ± 0.13	0.05 ± 0.02	0.30 ± 0.13
12 - 25	0	0.02 ± 0.01	0.02 ± 0.01	0.00 ± 0.00
25 - 40	0	0.01 ± 0.01	0.01 ± 0.01	0.00 ± 0.00
> 40	0	0.02 ± 0.01	0.01 ± 0.01	0.01 ± 0.01
All P_T^X	0	0.40 ± 0.15	0.09 ± 0.03	0.31 ± 0.13
		***ep* collisions**		
< 12	0	1.89 ± 0.58	0.46 ± 0.16	1.43 ± 0.56
12 - 25	0	0.18 ± 0.06	0.14 ± 0.05	0.04 ± 0.02
25 - 40	0	0.13 ± 0.05	0.11 ± 0.05	0.02 ± 0.01
> 40	0	0.10 ± 0.06	0.08 ± 0.06	0.02 ± 0.01
All P_T^X	0	2.30 ± 0.74	0.79 ± 0.31	1.51 ± 0.59

Table 8.9: *Number of events with isolated tau leptons and large missing transverse momentum selected using the criteria listed in table 7.8 in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background contributions. The uncertainties quoted on the expectation are calculated by adding the statistical and systematic uncertainties in quadrature.*

but its convolution with a Gaussian [2],

$$P\left(N_{obs}|N_{exp},\sigma_{exp}\right) = \int_0^\infty P_{poisson}\left(N_{obs}|x\right) \frac{1}{\sqrt{2\pi}\sigma_{exp}} e^{\frac{-\left(N_{exp}-x\right)^2}{2\left(\sigma_{exp}\right)^2}}. \tag{8.1}$$

In this ansatz, possible contributions from (non-lepton universality violating) BSM signal processes to the final event sample are accounted for by introducing a common scale factor s by which the Standard Model signal expectations $N_{exp}, signal^i$ in the electron, muon and tau channels are allowed to vary. The BSM contribution

[2]The integral may be evaluated analytically by using the identity [161] (integral 3.462)

$$\int_0^\infty x^n e^{-px^2-qx}\,dx = (2p)^{-\frac{n+1}{2}}\, n! e^{\frac{q^2}{8p}} D_{-n-1}\left(\frac{q}{\sqrt{2p}}\right) \qquad (p>0).$$

Its value is

$$P\left(N_{obs}|N_{exp},\sigma_{exp}\right) = \frac{1}{\sqrt{2\pi}\sigma_{exp}} e^{-pN_{exp}^2}\left(\sigma_{exp}\right)^{N_{obs}+1} e^{\frac{q^2}{2p}} D_{-N_{obs}-1}\left(\frac{q}{\sqrt{2p}}\right),$$

with the abbreviations $p \equiv 1/\left(2\left(\sigma_{exp}\right)^2\right)$ and $q \equiv p \cdot N_{exp} - 1/2$. The function D_{-n-1} denotes a *parabolic cylinder function*, whose value is computed numerically [162].

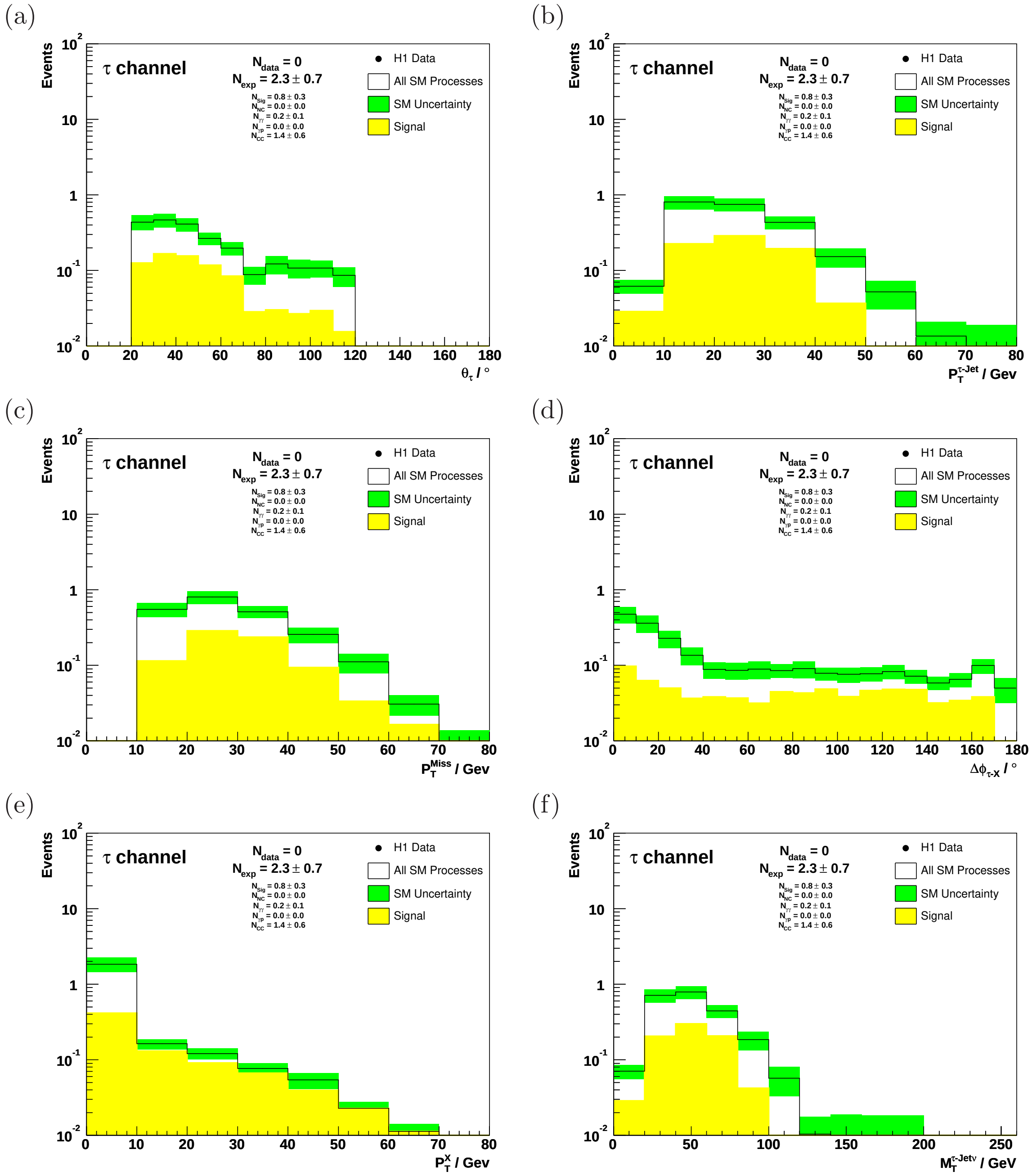


Figure 8.6: *Distributions of the polar angle (a) and the transverse momentum (b) of the tau jet, the missing transverse momentum (c), the acoplanarity angle (d), the transverse momentum of the hadronic system (e) and the transverse mass reconstructed for the tau jet and the missing momentum vector (f) in the final sample of events with isolated tau leptons and large missing transverse momentum.*

is then parametrized as deviation of this scale factor from unity [3]. The common

[3]Note that the introduction of a common scale factor requires the selection efficiencies in the electron, muon and tau lepton channels to be equal for Standard Model signal processes and possible BSM contributions. In figures 7.7, 7.16 and 7.22 it may be seen that the selection efficiencies are approximately equal in the muon and tau channels, but not in the electron channel. The different selection efficiencies for events with isolated electrons produced by Standard Model signal processes

scale factor s is chosen such as to maximize the probability

$$\mathcal{P}_s \equiv \prod_i P\left(N^i_{obs} | N^i_{exp,background} + s \cdot N^i_{exp,signal}, s \cdot \sigma^i_{exp}\right) \tag{8.2}$$

of the observation as a function of the varied expectations

$$N^i_{exp} = N^i_{exp,background} + s \cdot N^i_{exp,signal}$$

in the individual channels.

The probabilities $\mathcal{P}_s$ estimated by equation 8.2 still have to be normalized. The normalisation is done by Monte Carlo techniques [163]. For each individual channel i, random event numbers are generated in a two stage procedure. In the first stage, random expectations $N^{i,MC}_{exp,signal}$ and $N^{i,MC}_{exp,background}$ for the number of signal and background events are generated, according to Gaussian distributions with means $N^i_{exp,signal}$ and $N^i_{exp,background}$ and variances $\left(\sigma^i_{exp,signal}\right)^2$ and $\left(\sigma^i_{exp,background}\right)^2$. In the second stage, a random number of events $N^{i,MC}_{obs}$ is generated according to a Poisson distribution, the mean of which is $N^{i,MC}_{exp} = N^{i,MC}_{exp,signal} + N^{i,MC}_{exp,background}$. For the event numbers generated in the individual channels, an unnormalized "Monte Carlo probability" $\mathcal{P}^{MC}_s$ is calculated according to

$$\mathcal{P}^{MC}_s = \prod_i P\left(N^{i,MC}_{obs} | N^{i,MC}_{exp,background} + s^{MC} \cdot N^{i,MC}_{exp,signal}, s^{MC} \cdot \sigma^{i,MC}_{exp}\right).$$

The normalized probability $\hat{\mathcal{P}}_s$ of the observation is then estimated as the percentage of generated event numbers with a "Monte Carlo probability" $\mathcal{P}^{MC}_s \leq \mathcal{P}_s$.

The probability $\hat{\mathcal{P}}_s$ for the number of events observed in the electron, muon and tau lepton channels to be compatible with lepton universality is estimated separately in the region of small ($P^X_T < 25\ GeV$) and large ($P^X_T > 25\ GeV$) hadronic transverse momenta. In the combined 1994-2000 e^+p and e^-p datasets, the estimated levels of compatibility are 61% in the region $P^X_T < 25\ GeV$ and 66% for $P^X_T > 25\ GeV$. These probabilities are interpreted as confirming that the results for events with isolated electrons, muons and tau leptons are compatible with lepton universality within the estimated statistical and systematic uncertainties.

Now that the compatibility of the electron, muon and tau lepton channels is confirmed, the final samples of events with isolated electrons, muons and tau leptons may be combined and the combined results presented.

8.6 Combined Results

In the combined electron, muon and tau lepton channels, 19 events with isolated leptons and large missing transverse momentum are selected in the 1994-2000 e^+p and e^-p datasets. The number of events selected in the data is in agreement with the Standard Model expectation: in total 19.9 events are expected from all Standard Model processes in the combined electron, muon and tau lepton channels. Of

and possible BSM contributions are accounted for by introducing a *model uncertainty* of $10\% \cdot s \cdot N_{exp}$ on the selection efficiency that is added in quadrature to the uncertainty σ_{exp}.

P_T^X $[GeV]$	Data	SM Expectation	SM Signal	Other SM Processes
		e^+p collisions		
< 12	6	10.97 ± 1.47	6.04 ± 1.12	4.93 ± 0.95
12 - 25	4	2.95 ± 0.46	2.46 ± 0.44	0.49 ± 0.12
25 - 40	3	2.02 ± 0.34	1.59 ± 0.32	0.43 ± 0.11
> 40	5	1.23 ± 0.36	1.06 ± 0.36	0.17 ± 0.05
All P_T^X	18	17.17 ± 2.61	11.15 ± 2.24	6.02 ± 1.16
		e^-p collisions		
< 12	0	1.85 ± 0.29	0.80 ± 0.15	1.05 ± 0.24
12 - 25	1	0.44 ± 0.07	0.32 ± 0.06	0.13 ± 0.04
25 - 40	0	0.28 ± 0.05	0.21 ± 0.04	0.07 ± 0.03
> 40	0	0.19 ± 0.06	0.15 ± 0.05	0.05 ± 0.02
All P_T^X	1	2.76 ± 0.44	1.47 ± 0.30	1.29 ± 0.30
		ep collisions		
< 12	6	12.82 ± 1.75	6.83 ± 1.27	5.99 ± 1.18
12 - 25	5	3.39 ± 0.52	2.78 ± 0.50	0.61 ± 0.14
25 - 40	3	2.30 ± 0.38	1.80 ± 0.36	0.50 ± 0.12
> 40	5	1.42 ± 0.42	1.21 ± 0.41	0.22 ± 0.06
All P_T^X	19	19.93 ± 3.04	12.62 ± 2.54	7.32 ± 1.44

Table 8.10: *Number of events with isolated electrons, muons or tau leptons and large missing transverse momentum in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background components. The uncertainties quoted on the expectation are calculated by adding the statistical and systematic uncertainties in quadrature.*

the 19.9 events, the dominant contribution is expected from signal processes: 12.6 events are expected from W production and 7.3 from other Standard Model processes. The number of events selected in the data is compared with the Standard Model expectation and its signal and background contributions in four bins of the hadronic transverse momentum P_T^X in table 8.10. In the table, it can be seen that the background contribution is concentrated at low P_T^X and the signal contribution becomes more dominant at high P_T^X. An excess of events is observed in the signal dominated region of large hadronic transverse momenta. The excess is most significant for $P_T^X > 40\ GeV$, where only 1.4 events are expected from all Standard Model processes, but 5 are observed.

The main kinematic and topological properties of the events selected in the data are listed in table 8.11 (event displays are shown in section G of the appendix). The distributions of the kinematic and topological quantities of the selected events are compared with the Standard Model expectations in figure 8.7. In figure 8.7 (a), the distribution of the polar angles of isolated electrons and muons and of tau jets is shown. A slight excess of events over the Standard Model expectation can be seen at small polar angles. Such an excess is expected for BSM processes in which the isolated lepton results from the decay of a heavy particle (such as a top quark). The distribution of the transverse momenta of isolated electrons and muons

Run Event	Lepton	P_T^{ℓ} $[GeV]$	P_T^{miss} $[GeV]$	P_T^X $[GeV]$	$M_T^{\ell\nu}$ $[GeV]$	$M^{\ell\nu}$ $[GeV]$
90264 313	e^- (4.9σ)	$39.0^{+0.6}_{-0.5}$	$32.4^{+1.0}_{-1.0}$	$7.6^{+1.2}_{-1.2}$	$71.1^{+1.4}_{-1.4}$	-
186729 702	μ	$78.6^{+387.8}_{-35.9}$	$31.7^{+365.5}_{-12.5}$	$77.7^{+5.9}_{-5.8}$	$78.3^{+782.1}_{-44.4}$	-
188108 5066	μ^- (8.3σ)	$40.9^{+5.5}_{-4.4}$	$43.8^{+4.6}_{-3.5}$	$29.8^{+4.0}_{-3.9}$	$79.2^{+10.5}_{-8.2}$	$85.1^{+8.6}_{-6.7}$
192227 6208	μ^- (7.0σ)	$73.3^{+12.1}_{-9.2}$	$20.0^{+9.0}_{-3.4}$	$64.2^{+3.6}_{-3.7}$	$67.7^{+27.0}_{-19.1}$	-
195308 16793	μ^+ (4.3σ)	$60.1^{+18.3}_{-11.4}$	$35.3^{+18.0}_{-11.0}$	$27.4^{+1.9}_{-1.8}$	$91.4^{+37.4}_{-23.8}$	-
196406 38438	e	$14.0^{+0.4}_{-0.4}$	$19.7^{+1.3}_{-1.3}$	$10.8^{+1.4}_{-1.4}$	$32.2^{+1.2}_{-1.2}$	$76.4^{+1.8}_{-1.8}$
236176 3849	e	$10.2^{+0.2}_{-0.2}$	$21.1^{+1.6}_{-1.5}$	$19.8^{+1.8}_{-1.7}$	$24.2^{+0.7}_{-0.7}$	-
248207 32134	e^+ (14.5σ)	$32.7^{+0.5}_{-0.5}$	$44.1^{+2.4}_{-2.2}$	$43.0^{+3.1}_{-3.1}$	$63.6^{+1.3}_{-1.3}$	-
251415 43944	μ^- (8.8σ)	$22.7^{+2.8}_{-2.3}$	$14.7^{+2.0}_{-1.3}$	$17.8^{+1.8}_{-1.8}$	$32.9^{+5.5}_{-4.3}$	$76.1^{+6.3}_{-5.7}$
252020 30485	e^+ (40.0σ)	$25.6^{+0.6}_{-0.6}$	$38.0^{+3.1}_{-2.9}$	$39.6^{+3.8}_{-3.7}$	$49.7^{+1.4}_{-1.4}$	-
253700 90241	μ^- (11.1σ)	$168.1^{+216.8}_{-61.0}$	$177.0^{+216.4}_{-60.6}$	$17.9^{+1.8}_{-1.8}$	$344.6^{+433.4}_{-121.7}$	-
266336 4126	μ^+ (26.0σ)	$19.7^{+0.8}_{-0.7}$	$64.7^{+3.5}_{-3.4}$	$49.5^{+3.7}_{-3.6}$	$68.3^{+2.4}_{-2.3}$	$68.9^{+2.8}_{-2.6}$
268338 70014	e^+ (5.0σ)	$32.0^{+0.5}_{-0.5}$	$66.3^{+4.8}_{-4.6}$	$45.6^{+5.3}_{-5.2}$	$87.0^{+2.8}_{-2.8}$	-
269672 66918	e	$17.4^{+0.3}_{-0.3}$	$19.3^{+0.6}_{-0.5}$	$3.0^{+0.6}_{-0.6}$	$36.6^{+0.8}_{-0.8}$	-
270132 73115	μ	$63.9^{+\infty}_{-40.2}$	$76.1^{+\infty}_{-37.0}$	$22.7^{+2.2}_{-2.2}$	$138.1^{+\infty}_{-79.7}$	-
274357 6157	e	$40.2^{+0.6}_{-0.6}$	$40.3^{+0.7}_{-0.7}$	$4.4^{+0.8}_{-0.7}$	$80.3^{+1.2}_{-1.2}$	$121.5^{+4.8}_{-5.0}$
275991 29613	e^+ (36.5σ)	$37.5^{+0.6}_{-0.6}$	$37.5^{+1.6}_{-1.1}$	$19.9^{+4.6}_{-4.6}$	$72.3^{+1.4}_{-1.3}$	$75.6^{+1.2}_{-1.1}$
276220 76295	e^- (42.0σ)	$52.2^{+0.8}_{-0.8}$	$52.2^{+0.8}_{-0.8}$	-	$104.5^{+1.6}_{-1.6}$	-
277699 91265	e	$28.2^{+0.4}_{-0.4}$	$25.4^{+0.9}_{-0.9}$	$3.2^{+0.8}_{-0.8}$	$53.5^{+1.2}_{-1.2}$	-

Table 8.11: *Kinematic and topological properties of the events with isolated leptons and large missing transverse momentum in the 1994-2000 e^+p and e^-p datasets. As explained in section 5.4.1, the invariant mass $M^{\ell\nu}$ of the lepton-neutrino system can only be reconstructed in events in which the scattered electron is detected. The number in brackets indicates the significance of the charge measurement. In case the significance of the charge measurement amounts to less than two Gaussian standard deviations in terms of its estimated uncertainty, the charge of the lepton is considered as unmeasured.*

is shown in figure 8.7 (b) [1]. A slight excess may be seen at large transverse momenta. The excess is due to events with isolated muons, the transverse momenta of which are inaccurately measured, and is not significant. The same slight excess may be seen in the distribution of the missing transverse momentum shown in figure 8.7 (c). In figure 8.7 (d), the acoplanarity angle distribution is shown. In agreement with the expectation, the events selected in the data follow an approximately flat distribution. In particular, no concentration of events is seen at either small or large acoplanarity angles, the regions in which the background contributions become more dominant. For events with isolated electrons or muons, the distribution of the transverse mass of the hypothetical lepton-neutrino system is shown in figure 8.7 (f) [2]. The transverse mass distribution is in agreement with the Jacobian peak expected for the leptonic decays of real W bosons. In the distributions (a)-(d) and (f), a good agreement between the events selected in the data and the Standard Model expectation for signal and background processes is seen. The only distribution, in

[1] The transverse momenta of tau jets are not included in the distribution shown in the figure, as they are shifted towards lower values due to the undetected momentum carried away by the neutrino produced in the tau decay.

[2] The transverse masses reconstructed in events with isolated tau leptons are not included in the distribution shown in the figure, as they are shifted towards lower values (*cf.* section 8.4).

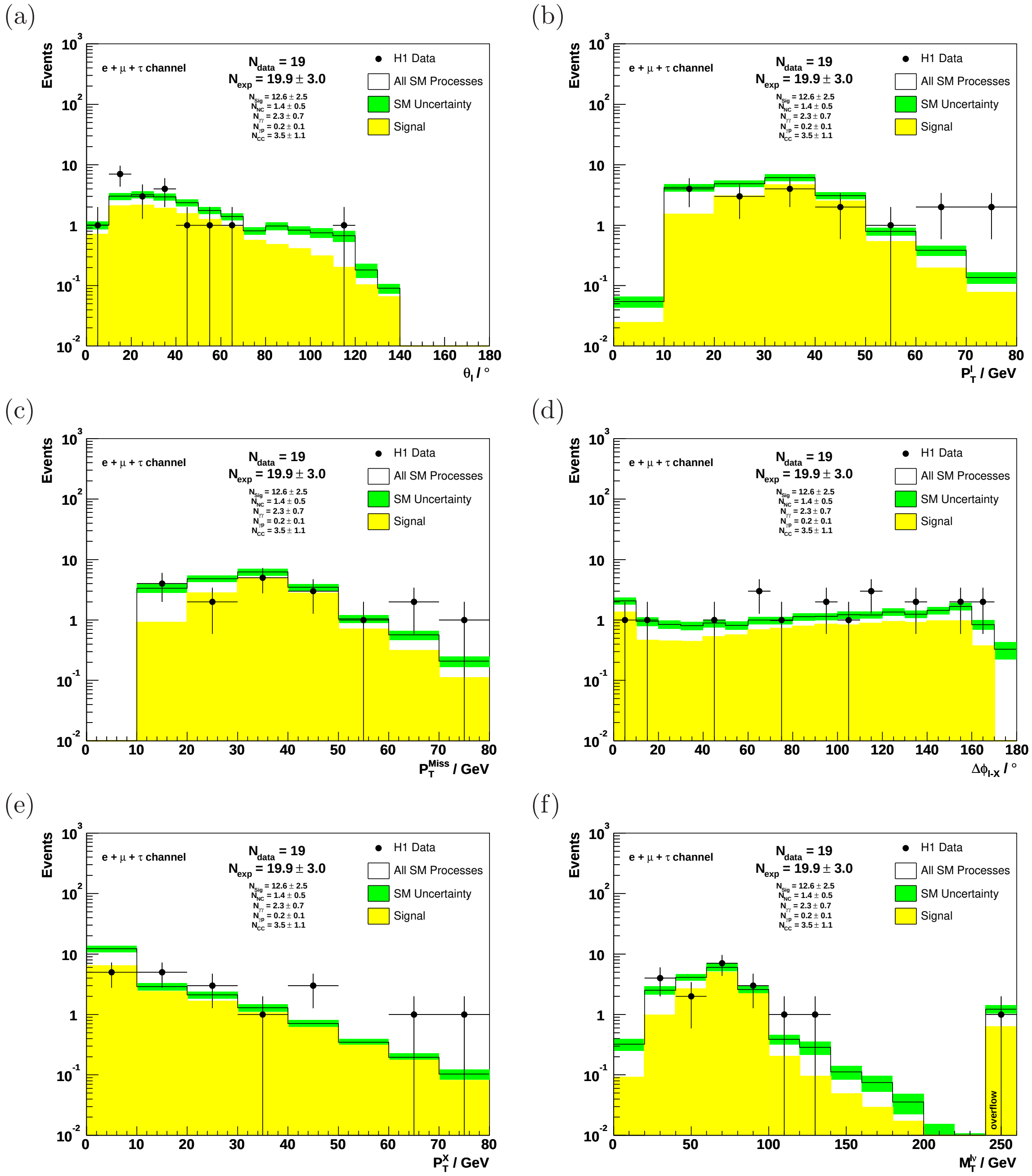


Figure 8.7: *Distributions of the polar angles (a) and the transverse momenta (b) of electrons, muons and tau jets, the missing transverse momentum (c), the acoplanarity angle (d), the transverse momentum of the hadronic system (e) and the transverse mass of the hypothetical lepton-neutrino system (f) in the combined final samples of events with isolated electrons, muons and tau leptons (in distributions (b) and (f) only events with isolated electrons or muons are included).*

which a significant deviation from the Standard Model expectation is seen is that of the hadronic transverse momentum shown in figure 8.7 (e). In figure 8.7 (e), the events selected in the data are seen to significantly exceed the expectation at very large hadronic transverse momenta.

All in all, the distributions shown support the interpretation of the excess observed in the data in the region $P_T^X > 40\ GeV$ as being produced by signal processes. In particular, the agreement seen between the data and the Standard Model expectation in the distribution of the transverse mass of the hypothetical lepton-neutrino system support the interpretation of the selected events as resulting from leptonic W decays. In this interpretation, the atypically large hadronic transverse momenta of the selected events may be explained by decays of heavy particles produced by BSM processes (such as the production of single top quarks) into W bosons. This interpretation is discussed in more detail in chapter 9.

8.7 Extraction of Cross-Sections

In order to quantify the excess of events observed at large hadronic transverse momenta in a model independent way, cross-sections for the production of events with isolated leptons and large missing transverse momentum are extracted. In this thesis, separate cross-sections are extracted for the production of events with isolated electrons, muons and tau leptons, in order to facilitate comparison with previous analyses restricted to the electron and muon channels. For comparison with previous analyses restricted to the 1994-2000 e^+p dataset, the cross-sections are extracted separately for e^+p and e^-p collisions. As in previous analyses, the cross-sections are measured in two bins, at small ($P_T^X < 25\ GeV$) and large ($P_T^X > 25\ GeV$) hadronic transverse momenta.

For each channel and P_T^X bin, a cross-section is extracted by correcting the number of events selected in the data for the expected background contribution and the estimated *acceptance* of the event selection for signal processes. To be more specific, the cross-section is calculated as

$$\sigma = \frac{N_{obs} - N_{exp}^{background}}{L \cdot \mathcal{A}}, \tag{8.3}$$

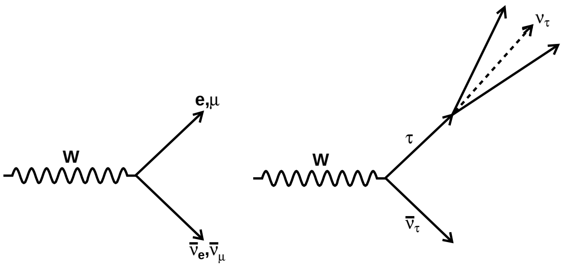
where N_{obs} denotes the number of events observed in the data, $N_{exp}^{background}$ the expected contribution from background processes, L is the integrated luminosity of the analysed datasets, and $\mathcal{A}$ the acceptance of the event selection for signal processes. The acceptance $\mathcal{A}$ is estimated using simulated samples of signal events. In each bin in which the cross-section is measured, the acceptance $\mathcal{A}$ is calculated as the ratio of the number of reconstructed to the number of generated signal events in that bin:

$$\mathcal{A} = \frac{N_{rec}}{N_{gen}}.$$

The events contributing to the numerator and denominator are required to pass the criteria defining the phase-space in which the cross-section is measured (at the generator level). In addition, the events contributing to the numerator are required to pass the final event selection (at the reconstructed level).

The cross-sections for the production of events with isolated electrons, muons and tau leptons and large missing transverse momentum are measured in the phase-space defined in the table on the right of figure 8.8. The acceptance of the event selections defined by the criteria listed in tables 7.2, 7.4, 7.6, and 7.8 in the phase-space defined in table 8.8 are shown in table 8.12 [1].

[1] The acceptances shown are estimated for Standard Model W production. The possibly different



Quantity	Requirement
Lepton	$p_T^\ell > 10\ GeV$ $5^\circ < \theta_\ell < 140^\circ$ $D_{jet} > 1.0$
Missing momentum	$P_T^{miss} > 12\ GeV$

Figure 8.8: *Illustration of W decays to electrons, muons and tau leptons (left) and definition of the phase-space in which the cross-sections for the production of events with isolated leptons and large missing transverse momentum are measured (right).*

P_T^X $[GeV]$	$\mathcal{A}_e$ $[\%]$	$\mathcal{A}_\mu$ $[\%]$	$\mathcal{A}_\tau$ $[\%]$
< 25	59.2	28.0	7.8
> 25	59.5	61.0	7.7
All P_T^X	59.3	34.1	7.8

Table 8.12: *Estimated acceptances of the event selections defined by the criteria listed in tables 7.2, 7.4, 7.6, and 7.8 for events with isolated electrons, muons and tau leptons in the phase-space defined in table 8.8.*

The extracted cross-sections are shown in table 8.13. The measured cross-sections are compared to the Standard Model expectations. Within the Standard Model, the cross-sections are expected to be dominated by the contribution from W production, to which is attributed a theoretical uncertainty of 15% (*cf.* section 2.6.1). In comparison to the cross-sections expected for the production of events with isolated electrons or muons, smaller cross-sections are expected for the production of events with isolated tau leptons: although W bosons decay into electrons, muons and tau leptons with equal branching fractions, the transverse momentum of the additional neutrino produced in the tau decay is likely to cancel that of the neutrino produced in the decay of the W boson, causing fewer of the events with isolated tau leptons to pass the $P_T^{miss} > 12\,GeV$ criterion defining the phase-space in which the cross-sections are measured (*cf.* table 8.8). The measured cross-sections are in agreement with the Standard Model expectation for $P_T^X < 25\ GeV$. For $P_T^X > 25\ GeV$, the cross-sections measured for the production of events with isolated electrons or muons significantly exceed the expectation, reflecting the excess of events with isolated electrons or muons observed at large hadronic transverse momenta. The cross-sections measured for the production of events with isolated tau leptons are compatible with the Standard Model expectation within the estimated statistical and systematic uncertainties. This compatibility is not surprising, as the statistical uncertainties (resulting from the low selection efficiency) of the cross-sections measured in the tau channel are too large to be sensitive to an excess of the size observed in the electron and muon channels.

acceptances for signal processes beyond the Standard Model are accounted for by attributing a model uncertainty of 10% to the estimated acceptance (*cf.* footnote 3 on page 183).

P_T^X [GeV]	σ_{e^+p} [pb^{-1}]	σ_{e^-p} [pb^{-1}]	σ_{ep} [pb^{-1}]	SM Expectation [pb^{-1}]
Events with Isolated Electrons				
< 25	$0.068^{+0.063}_{-0.031}$	< 0.366	$0.065^{+0.059}_{-0.030}$	$0.089^{+0.013}_{-0.013}$
> 25	$0.043^{+0.048}_{-0.015}$	< 0.235	$0.037^{+0.043}_{-0.014}$	$0.020^{+0.003}_{-0.003}$
All P_T^X	$0.111^{+0.074}_{-0.040}$	< 0.360	$0.101^{+0.066}_{-0.039}$	$0.109^{+0.016}_{-0.016}$
Events with Isolated Muons (in LAr and muon triggered phase-space)				
< 25	$0.055^{+0.108}_{-0.032}$	< 0.533	$0.043^{+0.097}_{-0.025}$	$0.088^{+0.013}_{-0.013}$
> 25	$0.076^{+0.070}_{-0.012}$	< 0.239	$0.067^{+0.033}_{-0.025}$	$0.020^{+0.003}_{-0.003}$
All P_T^X	$0.180^{+0.122}_{-0.090}$	< 0.431	$0.156^{+0.110}_{-0.073}$	$0.108^{+0.016}_{-0.016}$
Events with Isolated Tau Leptons				
< 25	< 0.262	< 2.325	< 0.236	$0.065^{+0.010}_{-0.010}$
> 25	< 0.266	< 2.285	< 0.239	$0.021^{+0.003}_{-0.003}$
All P_T^X	< 0.263	< 2.314	< 0.237	$0.087^{+0.013}_{-0.013}$
Events with Isolated Electrons, Muons or Tau Leptons				
< 25	$0.130^{+0.130}_{-0.070}$	< 0.608	$0.111^{+0.121}_{-0.062}$	$0.243^{+0.036}_{-0.036}$
> 25	$0.170^{+0.212}_{-0.086}$	< 0.344	$0.149^{+0.071}_{-0.077}$	$0.062^{+0.009}_{-0.009}$
All P_T^X	$0.325^{+0.203}_{-0.133}$	< 0.574	$0.282^{+0.167}_{-0.122}$	$0.304^{+0.046}_{-0.046}$

Table 8.13: *Measured cross-sections for events with isolated electrons, muons and tau leptons in e^+p and e^-p collisions compared with the Standard Model expectations. The measured e^-p cross-sections are quoted for a centre-of-mass energy of $\sqrt{s} = 318$ GeV. The measured e^+p and ep cross-sections are quoted for effective centre-of-mass energies of $\sqrt{s}_{eff} = 312$ GeV and $\sqrt{s}_{eff} = 313$ GeV, respectively. The uncertainties quoted on the measured cross-sections are calculated by adding the statistical and systematic uncertainties in quadrature. The cross-sections measured at the different (effective) centre-of-mass energies and in e^+p and e^-p collisions are compared with a Standard Model expectation for e^+p collisions and an effective centre-of-mass energy of $\sqrt{s}_{eff} = 313$ GeV. The differences between the cross-sections expected for the different (effective) centre-of-mass energies and in e^+p and e^-p collisions are negligible within the current statistical and systematic uncertainties. For bins in which the number of events selected in the data is below the expectation for background processes, an upper limit on the cross-section is calculated at a confidence level corresponding to one Gaussian standard deviation (in this case, the difference $N_{obs} - N_{exp}^{background}$ in the numerator of equation 8.3 is substituted by a limit on the number of signal events estimated by a technique described in detail in reference [164]).*

Chapter 9

Discussion

The results of the search for events with isolated leptons and large missing transverse momentum are presented in the previous chapter. In total, 19 events with isolated electrons, muons or tau leptons are selected, of which 11 are observed in the region $P_T^X < 25\ GeV$ and 8 in the region $P_T^X > 25\ GeV$. While the number of events observed at small hadronic transverse momenta is compatible with the Standard Model expectation, an excess of events is observed at large hadronic transverse momenta. The excess is most significant for $P_T^X > 40\ GeV$, where 5 events (2 with isolated electrons and 3 with isolated muons) are observed, while only 1.4 ± 0.4 are expected from all Standard Model processes.

9.1 Possible Interpretations

In principle, there exist three possible interpretations for the excess of events observed at large hadronic transverse momenta: within the Standard Model, the observed excess may be explained as arising firstly from an underestimation of the contribution from background processes or secondly as a statistical fluctuation. The third - and most intriguing - interpretation of the observed excess is that is due to beyond the Standard Model processes. Each of the three possible interpretations is discussed in turn in the following:

Underestimated Background Contribution

The interpretation of the observed excess as being due to an underestimated background contribution is an unlikely explanation. Firstly, a comparatively large additional contribution of background processes is necessary to explain the observed excess, and an underestimation of the background contribution by a large factor seems unlikely given the agreement between the Monte Carlo simulation and the data in the background enriched control samples discussed in section 8.1.2. Secondly, no evidence for unexpected background contributions is seen in the distributions of the kinematic and topological quantities shown in figure 8.7. Finally (and most importantly), the contribution from background processes to the selected event sample is expected to be concentrated at small hadronic transverse momenta (*cf.* table 8.10) - if the contribution from background processes were underestimated, an excess of events would be expected in the region of small, but not in the region of large hadronic transverse momenta.

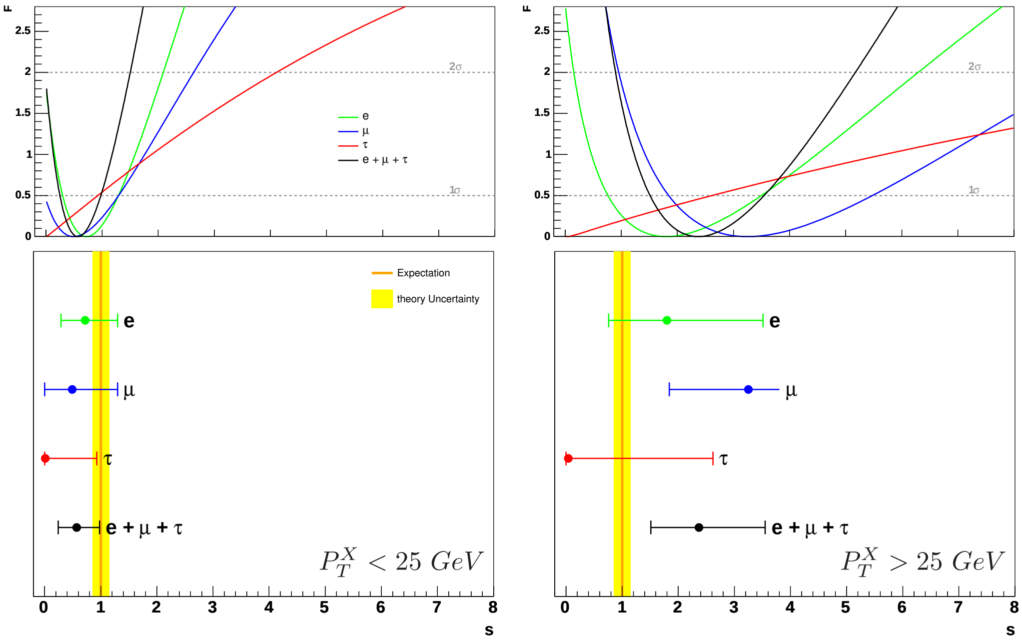


Figure 9.1: *Likelihood $F \equiv -\ln \hat{\mathcal{P}}_s$ of the observation in the electron, muon and tau lepton channels and their combination (top) as a function of the scale factor s that parametrizes the contribution from possible BSM processes. The agreement between data and expectation at small hadronic transverse momenta ($P_T^X < 25\ GeV$) is illustrated on the left, the excess over the Standard Model expectation observed in the data at large hadronic transverse momenta ($P_T^X > 25\ GeV$) is illustrated on the right. From the intersection of F with the line of constant $F = 0.5$, limits on the scale factor s at the level of one Gaussian standard deviation (bottom) are derived.*

Statistical Fluctuation

The interpretation of the observed excess as being merely a statistical fluctuation of the data is not very likely: The Poisson probability for the signal and background expectation of 1.4 ± 0.4 events to fluctuate up to the 5 events observed in the region $P_T^X > 40\ GeV$ amounts to only 2.2% [1].

The significance of the excess observed in the region $P_T^X > 25\ GeV$ and the agreement with the expectation in the region $P_T^X < 25\ GeV$ is graphically illustrated in figure 9.1. In the figure, $F \equiv -\ln \hat{\mathcal{P}}_s$ is shown as a function of the scale factor s that parametrizes the contribution from possible BSM processes (*cf.* section 8.5) [2] for the individual electron, muon and tau lepton channels and their combination. From the intersection of the graphs with the line of constant abscissa $s = 1$ representing no

[1]The probability is calculated as

$$P = \sum_{i=N_{obs}}^{\infty} P\left(i|N_{exp}, \sigma_{exp}\right),$$

where $P\left(i|N_{exp}, \sigma_{exp}\right)$ represents the probability to observe N_{obs} events when $N_{exp} \pm \sigma_{exp}$ are expected as given by equation 8.1.

[2]The function F is termed the *log-likelihood* function [165]. The absolute value of F is of no physical meaning - only the difference between $F(s)$ and $\min_{x\epsilon\mathbb{R}} F(x)$. For convenience, an additive constant is added to F, such that $\min_{x\epsilon\mathbb{R}} F(x) = 0$.

contribution from possible BSM processes, the significance of the excess observed in the individual channels can be read off: the number n of equivalent Gaussian standard deviations σ is related to the ordinate of the intersection by $n = \sqrt{2F}$ [165]. Numerically, the significance of the excess observed in the region $P_T^X > 25\ GeV$ in the combined electron, muon and tau lepton channels is equivalent to 2.3σ.

Contribution from BSM Processes

An alternative explanation for the observed excess is that it is due to beyond the Standard Model processes. The theories considered in this context are mentioned in section 2.7. They comprise the production of lepto-quarks, of excited fermions, of supersymmetric particles, and of single top quarks:

In theories predicting lepto-quark production, the existence of particles with the quantum numbers of both quarks and leptons - the lepto-quarks - is postulated [111]. In *ep* interactions at HERA, lepto-quarks may be produced by the fusion of the beam electron with one of the quarks within the proton. In the subsequent lepto-quark decays, an isolated lepton and a jet of high transverse momentum, but no undetected particles may be produced [3]. The production of lepto-quarks is therefore an unlikely explanation for the observed excess, as lepto-quark events are intrinsically balanced and the isolated leptons produced in the lepto-quark decays are typically observed in a "back-to-back" configuration with the hadronic final state - in contrast to the observed acoplanarity angle distribution shown in figure 8.7.

In theories predicting the production of excited fermions, excited states of the Standard Model electron, muon and tau lepton are postulated [114]. At HERA, excited fermions may be produced by the excitation of beam electrons in their interaction with the proton beam. The excited electron may subsequently revert back to its ground state by radiating off a gauge boson; events with isolated leptons and large missing transverse momentum may be produced if W bosons are radiated. In these cases, final states similar to the observed events may be produced. The production of excited fermions is an unlikely explanation of the observed excess, however, as no evidence for excited fermions has been found in a dedicated search [115].

In supersymmetric theories, events with isolated leptons and large missing transverse momentum may be produced by decays of supersymmetric particles (e.g. stop quarks) to W bosons or to leptons and neutralinos. For supersymmetric particles to be produced at HERA, R-parity has to be violated [4], as the pair-production of supersymmetric particles is excluded for HERA centre-of-mass energies by the LEP mass limits. The production of single supersymmetric particles is an unlikely explanation for the observed excess of events with isolated leptons and large missing transverse momentum, as in R-parity violating supersymmetric theories other exotic final states are predicted in addition, which have not been found in the H1 data [118, 120].

[3] Note that for events with isolated muons or tau leptons to be produced in lepto-quark decays, lepton flavour conservation has to be violated.

[4] The R-parity is defined as a multiplicative quantum number in supersymmetric theories. Its value is $+1$ for all Standard Model particles and -1 for their supersymmetric partners. In R-parity conserving supersymmetric theories, supersymmetric particles can be produced pair-wise only, while in R-parity violating theories single supersymmetric particles may be produced.

The most likely of the above BSM explanations for the observed excess is the production of single top quarks by flavour-changing neutral current interactions as described in detail in section 2.7.1. The decay of single top quarks to W bosons and bottom quarks would simultaneously explain the presence of the isolated leptons, large missing transverse momentum and the large transverse momenta of the hadronic final states in a sizeable fraction of the observed events.

Production of Single Top Quarks

The compatibility of the observed excess with the production of single top quarks is checked by examining the distribution of charges and reconstructed masses. For single top events, a positively charged lepton and a reconstructed invariant mass $M^{\ell\nu X}$ of the lepton-neutrino-hadron system at approximately the nominal top quark mass is expected [5]. The correlation between the charges of the isolated leptons, the hadronic transverse momenta, and the invariant masses $M^{\ell\nu X}$ and $M_T^{\ell\nu}$ reconstructed in the observed events with isolated electrons or muons is shown in comparison to the expectation for Standard Model W production and single top production processes in figure 9.2. In the figure, several events with positively charged leptons or leptons of undetermined charge can be seen in the region of large hadronic transverse momenta P_T^X and high invariant masses $M^{\ell\nu X}$ of the lepton-neutrino-hadron system. In a dedicated analysis, several of the events observed in the region of large hadronic transverse momenta and high invariant masses are found to be compatible with the expectation for single top production [10, 11].

In the dedicated analysis, the contribution from Standard Model W production is suppressed by requiring events with isolated electrons or muons of either positive or undetermined charge to have $P_T^X > 30\ GeV$ and $M^{\ell\nu X} > 140\ GeV$. Using these criteria, 5 events (3 with isolated electrons and 2 with isolated muons) are selected in the 1994-2000 e^+p and e^-p datasets, exceeding an expectation of 1.3 ± 0.2 from all Standard Model processes (of which 1.2 ± 0.2 are expected from W production) [11].

The dedicated analysis also searched for a signal for single top production in the dominant hadronic decay channel

$$t \rightarrow bW^+ \rightarrow bq\bar{q}',$$

which has a branching fraction of about 70% [82] [6]. The events are selected by requiring them to contain three jets of large transverse momentum, with two of these having an invariant di-jet mass compatible with the nominal W boson mass (reconstructed for any combination of jets). Further criteria designed to suppress

[5] For heavy top (anti-top) quarks to be produced by FCNC interactions in ep collisions, an up (anti-up) quark within the proton is required with high Bjorken $x \gtrsim 0.3$. Whereas up quarks exist as valence and "sea" quarks within the proton, anti-up quarks exist within the proton as "sea" quarks only. The probability for "sea" quarks to be found at the high x required for the production of top quarks is very small. (The production of anti-top quarks is suppressed by factor of $\mathcal{O}(80)$ in relation to the production of top quarks [11].) In the decays of top quarks,

$$t \rightarrow bW^+ \rightarrow b\ell^+\nu,$$

only positively charged leptons are produced, due to charge conservation.

[6] A search has also been made for evidence for Standard Model W production in the hadronic channel, but the background (mainly from photoproduction processes) was found to be too high for a signal to be observable [166].

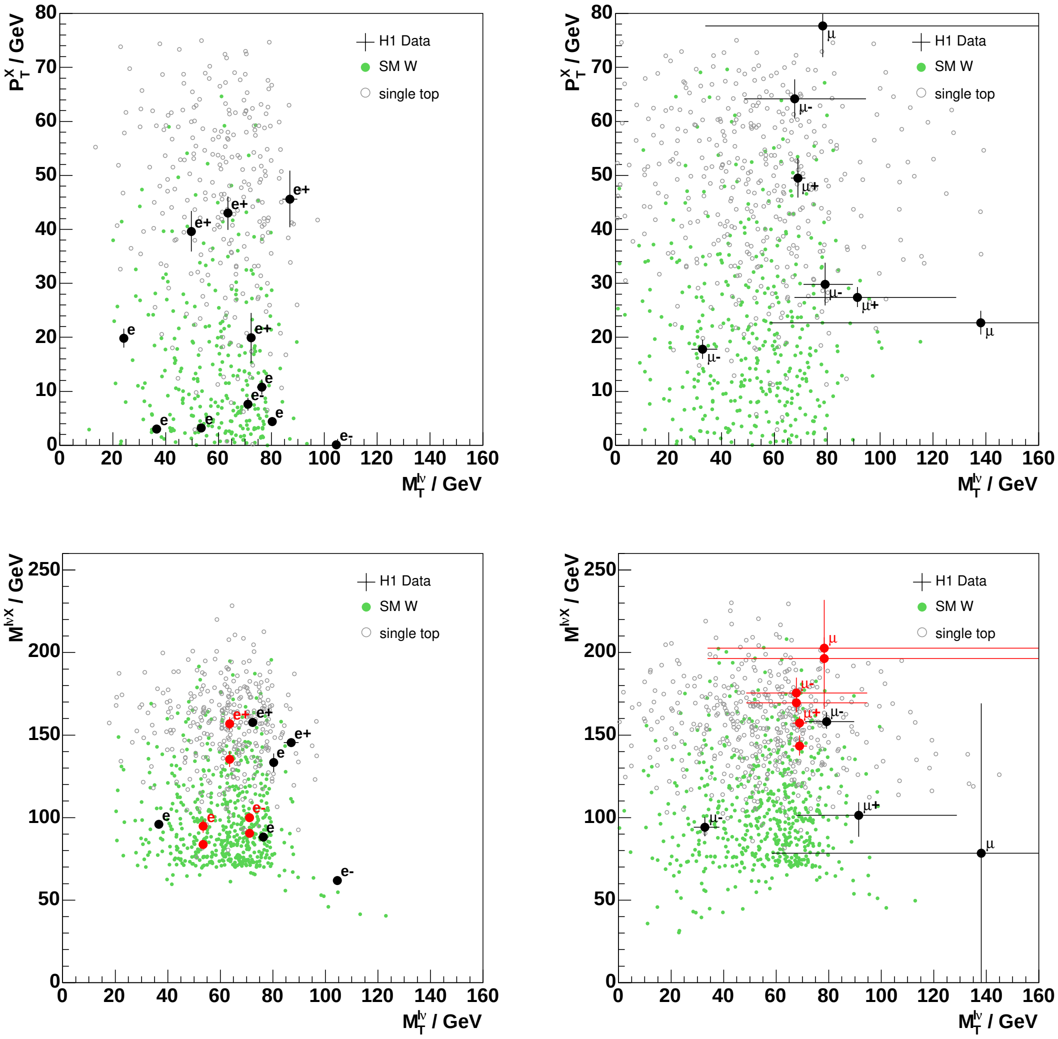


Figure 9.2: *Correlation between the transverse mass of the hypothetical lepton-neutrino system and the hadronic transverse momentum (top) and between the transverse mass and the invariant mass of the lepton-neutrino-hadron system (bottom) reconstructed in the observed events with isolated electrons (left) and muons (right) compared with the expectation for Standard Model W production and single top production processes. The observed events are represented by crosses, the size of which indicate the estimated uncertainties on the reconstructed invariant masses and transverse momenta. If multiple solutions are possible for the invariant mass $M^{\ell\nu X}$ of the lepton-neutrino-hadron system (cf. section 5.4.2), both solutions are shown in red. The charge of leptons is considered as undetermined if the significance of the charge measurement in terms of its estimated uncertainties amounts to less than two Gaussian standard deviations.*

various backgrounds are also required (see reference [10] for details of the event selection). No evidence for the production of single top quarks is found in the hadronic channel. In the 1994-2000 e^+p and e^-p datasets, 18 events are selected, in agreement with an expectation of 20.2 ± 3.6 from all Standard Model processes (dominated by the contribution from photoproduction processes) [11].

The results of the search for single top production in the hadronic channel do not exclude the interpretation of the excess of events with isolated leptons and large missing transverse momentum as arising from single top production. In combination with the results of a similar analysis performed by the ZEUS collaboration, in which no evidence for the production of single top quarks was found in the hadronic channel [7], an upper limit of

$$\sigma\,(ep \rightarrow etX) < 0.22\;pb^{-1}$$

(at the 95% CL) can be set on the cross-section for the production of single top quarks in ep collisions. (The combined limit is calculated by likelihood methods similar to those described in reference [163].) For branching fractions of top quarks to decay leptonically via

$$t \rightarrow bW^{+} \rightarrow b\ell^{+}\nu$$

of about 10% (per lepton channel) and an estimated selection efficiency of 32% in the electron and muon channels [8], the upper limit on the production cross-section set by the results of the hadronic channel corresponds to a possible contribution of about 2 events (one each in the electron and muon channels) to the sample of events with isolated leptons and large missing transverse momentum. A contribution of this size is just about sufficient to explain the excess of events observed at large hadronic transverse momenta: if a contribution of 2 events from single top production is added to the Standard Model expectation of 1.4 ± 0.4, the Poisson probability for the signal and background expectation to fluctuate up to the 5 events observed in the region $P_T^X > 40\;GeV$ increases, from 2.2% to 43% [9].

9.2 Comparison to Previous H1 Analyses

The results of the analysis presented in this thesis are compatible with previous searches for isolated electron and muon events made using the same datasets. The results of previous searches for events with isolated leptons and large missing transverse momentum are summarized in tables 9.1 and 9.2 [3,4] [1]. The numbers quoted in these tables can be directly compared with those quoted in tables 8.5 and 8.6. In

[7] In the ZEUS analysis, 14 events are selected in the 1994-2000 e^+p and e^-p datasets, in agreement with an expectation of $17.6^{+1.8}_{-1.2}$ from all Standard Model processes.

[8] The quoted selection efficiency for single top decays include the requirement $P_T^X > 40\;GeV$ in addition to the selection criteria listed in tables 7.2 and 7.4. In most single top production events, the transverse momentum of the hadronic system P_T^X is above 40 GeV (*cf.* figure 2.15).

[9] The Poisson probability is calculated as described in footnote 1 on page 193. The compatibility of the observed excess of events with isolated leptons and large missing transverse momentum with the interpretation that it contains a contribution from single top production processes is additionally checked by the method described in section 8.5. The contribution from all Standard Model processes is considered to be background and the production of single top quarks as signal. The signal expectation in the leptonic and hadronic channels is corrected for the different branching fractions and selection efficiencies and allowed to vary by a common scale factor that represents the cross-section for single top production. For the combined H1 and ZEUS analyses, the level of compatibility between the results in the leptonic and hadronic channels is estimated to be 4.7%.

[1] The results of the analyses presented in references [3], [4] and [5] are very similar. For clarity, the results of the analysis presented in this thesis are compared with those of only one of the three references: the results for e^+p are compared with those of reference [4] and the results for e^-p with those of reference [3].

P_T^X [GeV]	Data	SM Expectation	SM Signal	Other SM Processes
e^+p collisions				
< 12	5	6.40 ± 0.79	4.45 ± 0.70	1.95 ± 0.36
12 - 25	1	1.96 ± 0.27	1.45 ± 0.24	0.51 ± 0.12
25 - 40	1	0.95 ± 0.14	0.82 ± 0.13	0.13 ± 0.04
> 40	3	0.54 ± 0.11	0.45 ± 0.11	0.09 ± 0.04
All P_T^X	10	9.85 ± 1.31	7.17 ± 1.18	2.68 ± 0.56
e^-p collisions				
< 12	0	1.46 ± 0.30	0.86 ± 0.26	0.60 ± 0.16
12 - 25	0	0.50 ± 0.11	0.31 ± 0.09	0.19 ± 0.06
25 - 40	1[1]	0.26 ± 0.06	0.16 ± 0.05	0.10 ± 0.04
> 40	0	0.12 ± 0.03	0.06 ± 0.02	0.06 ± 0.03
All P_T^X	1	2.34 ± 0.50	1.39 ± 0.43	0.95 ± 0.29

[1] this event failed the selection criteria in reference [3], but was selected in the later publication [4]

Table 9.1: *Number of events with isolated electrons and large missing transverse momentum selected in previous H1 analyses of the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background contributions. The uncertainties quoted on the expectation are calculated by adding statistical and systematic uncertainties in quadrature.*

P_T^X [GeV]	Data	SM Expectation	SM Signal	Other SM Processes
e^+p collisions				
< 12	-	-	-	-
12 - 25	2	1.11 ± 0.19	0.94 ± 0.18	0.17 ± 0.05
25 - 40	3	0.89 ± 0.14	0.77 ± 0.14	0.12 ± 0.03
> 40	3	0.55 ± 0.12	0.51 ± 0.12	0.04 ± 0.01
All P_T^X	8	2.55 ± 0.45	2.22 ± 0.44	0.33 ± 0.09
e^-p collisions				
< 12	-	-	-	-
12 - 25	0	0.32 ± 0.09	0.28 ± 0.08	0.04 ± 0.03
25 - 40	0	0.19 ± 0.06	0.18 ± 0.05	0.02 ± 0.01
> 40	0	0.08 ± 0.02	0.08 ± 0.02	0.00 ± 0.00
All P_T^X	0	0.59 ± 0.17	0.54 ± 0.15	0.06 ± 0.04

Table 9.2: *Number of events with isolated muons and large missing transverse momentum selected in the LAr triggered phase-space by previous H1 analyses of the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes, and its signal and background contributions. The uncertainties quoted on the expectation are calculated by adding statistical and systematic uncertainties in quadrature.*

total, 19 events with isolated electrons or muons were selected in the analyses presented in references [3] and [4]. All these 19 events are also selected in the analysis presented in this thesis. The total signal and background expectations of 11.3 ± 2.2 and 4.0 ± 1.0 events in previous analyses are in agreement with those of the analysis presented in this thesis for events with isolated electrons or isolated muons in the LAr triggered phase-space [2].

In comparison to previous analyses, the excess of events observed at large hadronic transverse momenta is less significant in the analysis presented in this thesis. In previous analyses, the most significant excess was observed for $P_T^X > 25\ GeV$, where 11 events were selected in the data, while only 3.6 ± 0.7 were expected from all Standard Model processes (*cf.* tables 9.1 and 9.2). The Poisson probability for the excess observed in previous analyses to be compatible with a statistical fluctuation is only 0.16%, significantly smaller than the value of 2.2% calculated for the excess observed in the region $P_T^X > 40\ GeV$ in the analysis presented in this thesis (*cf.* section 9.1).

The lower significance of the excess observed in the analysis presented in this thesis has two reasons. The first is the extension of the analysis to the tau channel, in which no excess of events over the Standard Model expectation is observed [3]. The second reason are differences in the hadronic transverse momenta reconstructed in the 19 selected events.

The effect of not observing an excess in the tau channel on the significance of the excess observed in the combined electron, muon and tau lepton channels is small, as only a small number of events is expected in the tau channel in comparison to the number of events expected in the electron and muon channels. For $P_T^X > 25\ GeV$, where the excess in the electron and LAr triggered phase-space of the muon channel was observed in previous analyses, only 0.2 ± 0.1 events are expected from all Standard Model processes in the tau channel.

The main reason for the lower significance of the excess observed in the analysis presented in this thesis are the differences in the hadronic transverse momenta in the selected events in comparison to previous analyses. In the analysis presented in this thesis, 3 of the 11 events reconstructed in the region $P_T^X > 25\ GeV$ in

[2]In comparison to previous analyses, an approximately 5% smaller contribution from signal processes and an approximately 20% higher contribution from background processes is expected in the analysis presented in this thesis. The differences mainly arise from the electron channel. The smaller signal expectation is mainly due to a complementary electron identification algorithm that is used in addition to that described in section 4.3 in previous analyses which is not implemented in the H1OO framework. The higher background expectation is due to the contributions from NC and CC DIS processes. The higher contribution from NC DIS processes is compatible with being a statistical fluctuation of the number of selected Monte Carlo events from which the background contribution is estimated. (Although at least ten times the data luminosity is simulated for all background processes, statistical fluctuations are noticeable in some cases.) The higher contribution from CC DIS processes is due to slightly different track-link criteria required for the electron candidates in the analysis presented in this thesis in comparison to previous analyses and to different Monte Carlo datasets used to estimate the background contribution. (In the analysis presented in this thesis, a set of higher simulated luminosity is used.)

[3]The probability that the results from the electron, muon and tau lepton channels are compatible is discussed in section 8.5. Note that the muon triggered phase-space for the muon channel, in which no excess of events over the Standard Model expectation is observed, hardly contributes to the region $P_T^X > 25\ GeV$ of large hadronic transverse momenta (*cf.* table 8.7).

Run Event	P_T^X $[GeV]$ (FSCOMB)	P_T^X $[GeV]$ (HadrooII)
90264 313	$9.3^{+1.1}_{-1.1}$	$7.6^{+1.2}_{-1.2}$
186729 702	$66.7^{+4.9}_{-4.9}$	$77.7^{+5.9}_{-5.8}$
188108 5066	$26.9^{+2.2}_{-2.3}$	$29.8^{+4.0}_{-3.9}$
192227 6208	$60.5^{+5.5}_{-5.4}$	$64.2^{+3.6}_{-3.7}$
195308 16793	$33.3^{+3.6}_{-3.6}$	$27.4^{+1.9}_{-1.8}$
196406 38438	$13.4^{+2.0}_{-1.9}$	$10.8^{+1.4}_{-1.4}$
236176 3849	$25.3^{+2.8}_{-2.8}$	$19.8^{+1.8}_{-1.7}$
248207 32134	$42.7^{+3.9}_{-4.1}$	$43.0^{+3.1}_{-3.1}$
251415 43944	$16.5^{+3.0}_{-3.1}$	$17.8^{+1.8}_{-1.8}$
252020 30485	$44.3^{+3.6}_{-3.6}$	$39.6^{+3.8}_{-3.7}$

Run Event	P_T^X $[GeV]$ (FSCOMB)	P_T^X $[GeV]$ (HadrooII)
253700 90241	$21.2^{+2.6}_{-2.7}$	$17.9^{+1.8}_{-1.8}$
266336 4126	$51.5^{+3.8}_{-4.0}$	$49.5^{+3.7}_{-3.6}$
268338 70014	$46.6^{+3.3}_{-3.3}$	$45.6^{+5.3}_{-5.2}$
269672 66918	$4.6^{+0.7}_{-0.7}$	$3.0^{+0.6}_{-0.6}$
270132 73115	$27.3^{+3.9}_{-3.9}$	$22.7^{+2.2}_{-2.2}$
274357 6157	$6.3^{+1.3}_{-1.3}$	$4.4^{+0.8}_{-0.7}$
275991 29613	$28.4^{+5.7}_{-5.9}$	$19.9^{+4.6}_{-4.6}$
276220 76295	-	-
277699 91265	$3.3^{+0.7}_{-0.7}$	$3.2^{+0.8}_{-0.8}$

Table 9.3: *Hadronic transverse momenta reconstructed by the HadrooII and FSCOMB algorithms for the* 19 *events with isolated electrons or muons and large missing transverse momentum selected by previous analyses in the 1994-2000* e^+p *and* e^-p *datasets.*

previous analyses are reconstructed in the region $P_T^X < 25\ GeV$. As a result, the number of events observed with hadronic transverse momenta between 25 GeV and 40 GeV is found to be in agreement with the Standard Model expectation. The differences are due to different hadronic reconstruction algorithms used in the analysis presented in this thesis and in previous analyses. In previous analyses, the so-called *FSCOMB* algorithm [167] was used for hadronic reconstruction, whereas in the analysis presented in this thesis, the new HadrooII algorithm of the H1OO framework is used. In comparison to the previously used FSCOMB algorithm, the combination of the cluster and track information performed in HadrooII has been demonstrated to yield a better energy resolution for hadrons than does FSCOMB [32]. The correlation between the hadronic transverse momenta reconstructed by the HadrooII and FSCOMB algorithms in the 19 events selected by both the analysis presented in this thesis and previous analyses is shown in table 9.3. The hadronic transverse momenta reconstructed by the two algorithms are very similar in these 19 events, with differences of less than $\mathcal{O}(5\ GeV)$. As is illustrated in figure 9.3, differences of this order are expected from the Monte Carlo simulation. The 3 events reconstructed in the region $P_T^X > 25\ GeV$ in previous analyses and in the region $P_T^X < 25\ GeV$ in the analysis presented in this thesis are all near the boundary between the two regions.

9.3 Comparison to Results of the ZEUS Collaboration

The ZEUS collaboration has also searched for events with isolated electrons, muons or tau leptons and large missing transverse momentum. The results of the ZEUS analysis are summarized in table 9.4 [124, 168]. The ZEUS analysis is restricted to the region $P_T^X > 25\ GeV$ of large hadronic transverse momenta, in which the excess of events with isolated electrons or muons is observed in previous H1 analyses. In the ZEUS analysis, an excess of events over the Standard Model expectation is observed for $P_T^X > 25\ GeV$. The observed excess is not very significant, however:

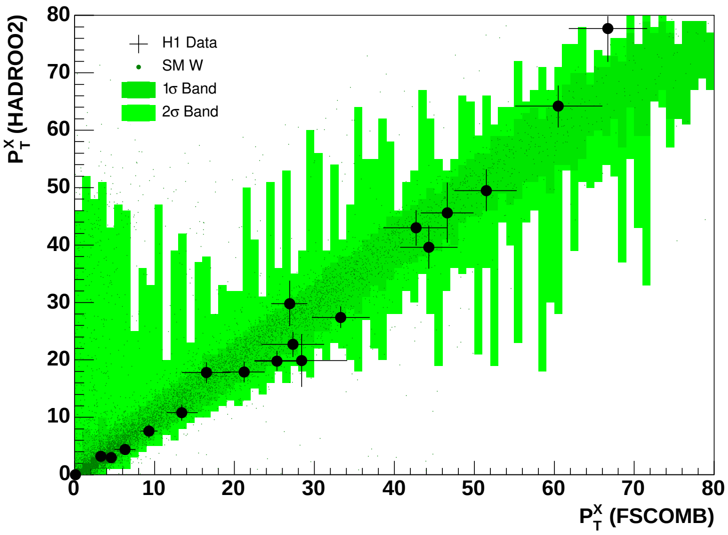


Figure 9.3: *Correlation between the hadronic transverse momenta reconstructed by the HadrooII and FSCOMB algorithms in simulated Standard Model W production events (dots) and in the 19 events selected in the electron channel and the LAr triggered phase-space of the muon channel by both the analysis presented in this thesis and previous analyses (crosses) in the 1994-2000 e^+p and e^-p datasets. Events that are reconstructed in different bins of hadronic transverse momentum in the analysis presented in this thesis in comparison to previous analyses are marked in red. The shaded area indicates the spread of reconstructed events around the diagonal at the level of one Gaussian standard deviation.*

P_T^X [GeV]	Electron obs./exp. (W)	Muon obs./exp. (W)	Tau Lepton obs./exp. (W)
< 12	-	-	-
12 - 25	-	-	-
25 - 40	$2/1.96^{+0.58}_{-0.30}$ (37%)	$5/1.80^{+0.16}_{-0.18}$ (44%)	$1/0.13^{+0.05}_{-0.05}$ (37%)
> 40	$0/0.94^{+0.11}_{-0.10}$ (61%)	$0/0.95^{+0.14}_{-0.10}$ (61%)	$1/0.07^{+0.02}_{-0.02}$ (71%)
all P_T^X	$2/2.90^{+0.59}_{-0.32}$ (45%)	$5/2.75^{+0.21}_{-0.21}$ (50%)	$2/0.20^{+0.05}_{-0.05}$ (49%)

Table 9.4: *Number of events with isolated electrons, muons or tau leptons and large missing transverse momentum observed by the ZEUS collaboration in the 1994-2000 e^+p and e^-p datasets compared with the expectation for all Standard Model processes. The Standard Model expectation is calculated for an integrated luminosity of $L = 130.1\ pb^{-1}$. The uncertainties quoted on the expectation are calculated by adding statistical and systematic uncertainties in quadrature. The contribution from W production to the Standard Model expectation is indicated in percent in brackets. Note that for almost identical signal efficiencies, the signal to background ratio is worse by about a factor of 5 in the ZEUS analysis than in the analysis presented in this thesis (cf. tables 8.5, 8.8 and 8.9).*

the Poisson probability for the expectation of $5.9^{+0.6}_{-0.4}$ events to fluctuate up to the 9 events observed in the combined electron, muon and tau lepton channels amounts to 22% [1]. In the region $P_T^X > 40\ GeV$, only one event is observed, in agreement

[1] The probability is calculated as described in footnote 1 on page 193.

with an expectation of 2.0 ± 0.2.

The results of the analysis presented in this thesis are compatible with those of the ZEUS analysis. The probability $\hat{\mathcal{P}}_s$ (calculated as described in section 8.5) for the number of events with isolated leptons and large missing transverse momentum observed in the H1 and ZEUS analyses in the region $P_T^X > 25\ GeV$ ($P_T^X > 40\ GeV$) to be compatible within the estimated statistical and systematic uncertainties is found to be 68% (6.8 %) [2]. Although the level of compatibility seems a little small in the region $P_T^X > 40\ GeV$, these probabilities are interpreted as the H1 and ZEUS analyses being compatible.

9.4 Conclusion

At present, the significance of the results presented here is limited by the small number of events. It is thus impossible to make a definitive statement on the origin of the observed excess of events with isolated leptons and large missing transverse momentum. On the one hand, the observed excess may be interpreted as a statistical fluctuation. The combination of the results of the analysis presented in this thesis with those of the ZEUS analysis gives 17 (6) events with isolated electrons, muons or tau leptons observed in the region $P_T^X > 25\ GeV$ ($P_T^X > 40\ GeV$), exceeding an expectation of $9.6^{+0.9}_{-0.7}$ ($3.4^{+0.5}_{-0.4}$). In the region $P_T^X > 25\ GeV$, where the excess in the combined results is most significant, the probability for the expectation from Standard Model signal and background processes to fluctuate up to the observed number of events amounts to 3.9% [1]. This probability is not so small that it excludes the interpretation of the observed excess as a statistical fluctuation. On the other hand, the observed events may be interpreted as contributions from beyond the Standard Model processes, in particular through the production of single top quarks by flavour changing neutral current interactions. The interpretation of the excess observed at large hadronic transverse momenta as arising from single top production is supported by the kinematic and topological properties of the observed events and neither excluded by first principles nor ruled out by the results of the search for single top production in the hadronic channel. The significance of the observed excess of events with isolated electrons, muons and tau leptons in comparison to the possible contribution from single top production is graphically illustrated in figure 9.4.

[2] The quoted probabilities are calculated for the compatibility of the *combined* results of the H1 and ZEUS analyses in all three lepton channels. The level of compatibility of the results in the *individual* electron, muon and tau lepton channels is smaller; it is estimated to be 15% (6.4%) in the region $P_T^X > 25\ GeV$ ($P_T^X > 40\ GeV$). The small probabilities calculated for the compatibility of the individual channels are rather indications for intrinsic incompatibilities of the ZEUS results than a measure of the compatibility between the H1 and ZEUS results. The probability $\hat{\mathcal{P}}_s$ for the results of the electron, muon and tau lepton channels of the ZEUS analysis to be compatible with lepton universality amounts to only 2.3% (3.3%) in the region $P_T^X > 25\ GeV$ ($P_T^X > 40\ GeV$) (*cf.* section 8.5 for a comparison with the corresponding probabilities for the results of the analysis presented in this thesis). As has been remarked previously [163], the small level of compatibility of the ZEUS results with lepton universality is due to the observation of a significant excess of events in the tau channel, while no corresponding excess is observed in the electron and muon channels.

[1] The probability is calculated as described in footnote 1 on page 193.

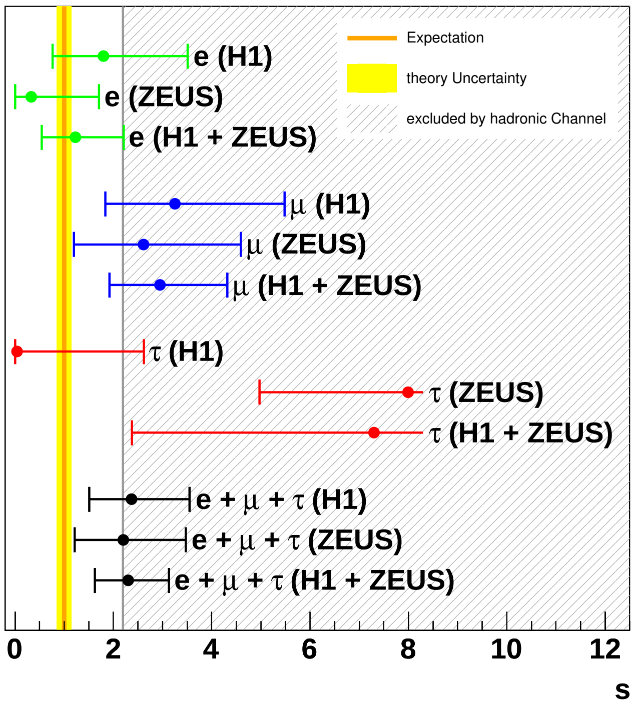


Figure 9.4: *Significance of the observed excess of events with isolated electrons, muons and tau leptons and large missing transverse momentum in the 1994-2000 e^+p and e^-p datasets in the region of hadronic transverse momenta $P_T^X > 25$ GeV. A possible contribution from decays of single top quarks is not ruled out (at the 95% confidence level) by the results of the search for single top production in the hadronic channel as is indicated by the unshaded area.*

The clarification of the origin of the excess observed in the HERA I data remains an exciting challenge for the HERA II running period. The HERA II data is anticipated to significantly increase the presently small numbers of events in the near future: at nominal HERA II running conditions, an integrated luminosity of almost 100 pb^{-1} is expected per year. The prospects for discovering a signal of single top quark production in the HERA II data were studied in reference [163]. It was estimated that 200 pb^{-1} of HERA II luminosity are sufficient to find a 5σ single top production signal in the combined H1 and ZEUS data, if an excess of events with isolated leptons and large missing transverse momentum at the level of that observed by H1 in the HERA I data were seen by both H1 and ZEUS in the HERA II running period [2]. To find signals of other beyond the Standard Model processes, more or less luminosity may be necessary. In any case, the larger event sample expected in the HERA II data will hopefully clarify whether the excess observed in the HERA I data is a statistical fluctuation or a signal for new physics.

First preliminary results for the HERA II e^+p dataset (of 53 pb^{-1} integrated luminosity) indicate that the excess of events with isolated leptons and large missing transverse momentum observed in the HERA I e^+p data continues to be seen by H1 [169] (the ZEUS collaboration has not published preliminary results yet). At

[2]For details of the estimation, see reference [163].

time of writing, the combined HERA I and HERA II e^-p datasets amount to an integrated luminosity of 34 pb^{-1}. A slight excess of events with isolated leptons and large missing transverse momentum may also be seen in the combined e^-p dataset; however, the event sample is not sufficiently large to conclude whether the excess observed in e^+p collisions is also present in the e^-p data. Once more e^-p data is collected, it will be very interesting to compare the e^+p and e^-p results in more detail.

No preliminary results exist yet on single top production in the hadronic channel for the HERA II data. These will be equally interesting, as, if no evidence for the production of single top quarks is seen in the HERA II data, the interpretation of the observed excess of events with isolated leptons and large missing transverse momentum as arising from single top production is likely to be ruled out soon.

Summary

In this thesis, the results of a search for events with isolated leptons and large missing transverse momentum in the e^+p and e^-p datasets recorded by the H1 experiment during the HERA 1994-2000 running period are presented. In comparison to previous H1 analyses, the acceptance is improved by extending the phase-space in which events with isolated muons are searched for and searching for events with isolated tau leptons.

A dedicated algorithm is developed for the identification of hadronic tau decays. The efficiency with which this algorithm identifies hadronic "one-prong" or "three-prong" tau decays is estimated to be 50% and the probability for QCD jets to be misidentified as hadronic "one-prong" ("three-prong") tau decays to be about 0.5% (4%). The true identification capability of the algorithm is demonstrated by finding the first evidence for tau pair production events in the H1 data.

In total, 19 events with isolated electrons, muons, or tau leptons and large missing transverse momentum are observed in the e^+p and e^-p datasets. The total number of observed events is compatible with an expectation of 19.9 ± 3.0 events from all Standard Model processes, which is dominated by the contribution from the production of real W bosons with subsequent leptonic decay. In the region of large hadronic transverse momenta, a significant excess of events over the Standard Model expectation is observed. The observed excess is most significant in the region $P_T^X > 40\ GeV$, in which 5 events are observed, while 1.4 ± 0.4 are expected from all Standard Model processes.

The observed excess may either be interpreted as a statistical fluctuation or as evidence for contributions from beyond the Standard Model processes, in particular the production of single top quarks by flavour-changing neutral current interactions. The probability that the expectation of 1.4 ± 0.4 events fluctuates up to the 5 events observed with $P_T^X > 40\ GeV$ is estimated to be 2.2%. This probability is small, but not so small as to exclude the interpretation of the observed excess as a statistical fluctuation. The kinematic and topological properties of the events observed at large hadronic transverse momenta are compatible with resulting from the decays of single top quarks. The interpretation of the observed excess as single top decays is neither excluded by first principles nor ruled out by current experimental limits on the production cross-section. At present, the limited number of observed events prohibits a definitive statement on the origin of the observed excess.

This will hopefully be clarified in the near future, as the HERA experiments accumulate more data.

Appendix A

Track Reconstruction in Drift Chambers and MWPCs

Both drift chambers and MWPCs are types of *ionisation chambers*. An ionisation chamber is a gas-filled detector that registers the traversal of a charged particle by measuring the ionisation of the gas atoms in the gas-filled detector volume when the charged particle crosses the chamber (see subsection 4.1.1). A high voltage is applied to the system of *electrodes* located inside the chamber and surrounding the gas-filled volume, setting up an electric field. In the electric field, the electron-ion pairs produced by the traversing charged particle become separated and start drifting towards the electrodes. The positively charged ions drift towards the *cathode* and the negatively charged electrons drift towards the *anode*. This charge transport generates a measurable current pulse on the electrodes. In an ionisation chamber, a comparatively small voltage is applied to the electrodes, and hence the drifting electron-ion pairs do not get accelerated enough to ionize further gas molecules. As a result, the current pulse in an ionisation chamber is created entirely by the *primary ionisation* produced by the charged particle traversing the chamber.

For many applications, however, a greater voltage pulse is desirable, because it needs less electronic amplification. Charge amplification by *secondary ionisation processes* is attainable by higher electric field strengths. In this case, the electric field accelerates the drifting electrons so much that the gain in kinetic energy between two collisions with the gas molecules exceeds the ionisation energy of the gas: the drifting electrons produce additional electron-ion pairs. By secondary ionisation processes, the number of electron-ion pairs then increases exponentially, leading to an *avalanche* of electrons drifting to the anode, that leaves behind a nearly stationary region of positive space-charge containing the ions (that drift to the cathode more slowly because of their higher masses). The development in time of such an exponentially increasing avalanche of electrons drifting towards the anode is illustrated in figure A.1. The ratio of the total ionisation (given by the sum of primary ionisation produced by the charged particle traversing the chamber and secondary ionisation produced by charge amplification) to the primary ionisation is called the *charge amplification factor*. For high voltages of a few kV applied to anode wires of typically $\mathcal{O}\left(20\mu m\right)$ in diameter, charge amplification factors of more then 10^{10} may be achieved [24]. For amplification factors up to the order of 10^6, the charge amplification by secondary ionisation processes is independent from the amount of

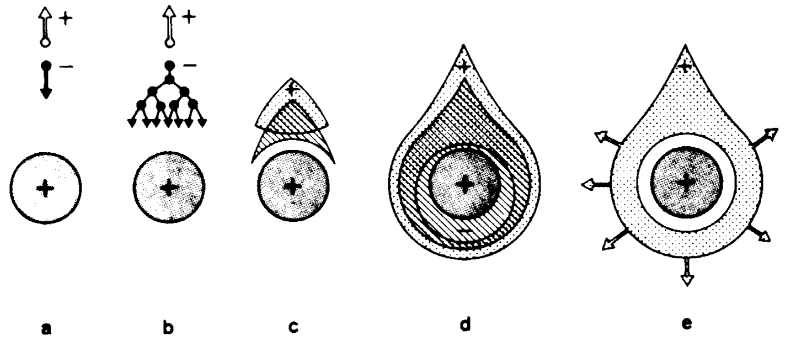


Figure A.1: *Development in time of an avalanche of electrons drifting towards the anode. (a) an electron-ion pair is produced by a traversing charged particle; the electron starts drifting towards the anode, the ion starts drifting towards the cathode. (b) in the accelerating electric field, the primary electron gains enough kinetic energy to ionize further gas molecules, initiating an exponentially increasing avalanche of secondary electrons, while the heavier ion drifts only slowly. (c) the produced secondary electrons all drift towards the anode; the avalanche widens a little, because of diffusion. (d) the avalanche reaches the anode, around which the electrons start to distribute, creating a region of negative space-charge around it. (e) the electrons have been carried off, leaving behind a region of positive space-charge containing the ions, which slowly drift towards the cathode.*

primary ionisation, yielding a total ionisation proportional to the number of primary electron-ion pairs produced by the charged particle traversing the chamber (this operation is termed the *proportional mode*). In higher electric fields, the production of electron-ion pairs by secondary ionisation processes starts to saturate, as the charge distribution of the produced electron-ion pairs starts to shield the electric field of the anode. As a consequence, the proportionality between primary and total ionisation is lost (this operation is termed the *Geiger, or streamer, mode*).

In drift chambers, the timing of the voltage pulses on the anode wires is utilized for a precise track reconstruction. They are usually operated in proportional mode. In order to reconstruct the trajectory of a charged particle traversing the chamber, drift chambers measure the time difference between the traversal of the particle and the arrival of the drift electrons on the anode wire. The distance between the trajectory of the charged particle and the anode wire is then reconstructed from the known drift velocity of the electrons in the chamber gas and the measured time difference. From this distance, the trajectory of the charged particle is determined up to a *left-right ambiguity* in the plane perpendicular to the anode wire (see figure A.2). As is illustrated in the figure, the left-right ambiguity can be resolved by a second layer of anode wires shifted with respect to the first. For the reconstruction of the trajectory in direction of the anode wire three different techniques exist. One possibility is to infer the coordinate along the wire from the ratio of currents measured at both ends of the anode. Since the electrical resistance of wires is proportional to their length and the current is inversely proportional to the resistance, the ratio of currents measured at both ends of the anode wire is inversely proportional to the ratio of the distances between the trajectory of the charged particle and the wire

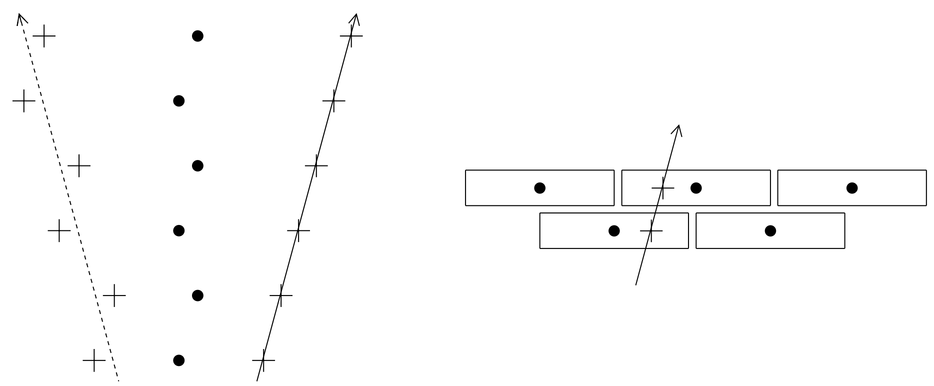


Figure A.2: *Illustration of left-right ambiguity resolved by double layer of anode wires in the same (left) and separate drift volumes (right). The staggering of anode wires illustrated on the left is used in the CJC1 and CJC2; paired layers of drift chambers as illustrated on the right are used to resolve left-right ambiguities in the BDC and FMD.*

ends,

$$\frac{I_1}{I_2} = \frac{d_2}{d_1}.$$

Alternatively, if the arrival times of the current pulses at the wire ends are measured, the coordinate along the wire may be reconstructed by the measured time difference between the pulses at both wire ends. For chambers with segmented cathodes (usually either strips oriented perpendicular to the anode wires or rectangular pads), the coordinate along the anode wire may also be inferred from a measurement of the pulses induced on the cathode segments by electrons drifting to the anode (ions drift to slow). By means of either charge division, time difference or cathode read-out, the starting position of the avalanche of drift electrons may be reconstructed in three dimensions, yielding a three-dimensional space point on the trajectory of the charged particle traversing the chamber, termed a *hit*.

The trajectory of the charged particle traversing the drift chamber, is then reconstructed by a *track-finding algorithm* [170] that searches the hits for track segments. The track segments found are then linked to tracks and the track parameters determined by fitting the hits [1], which provides an estimate for the momentum of the charged particle traversing the chamber. In particular, for drift chambers operated in solenoidal magnetic fields, the transverse momentum of the charged particle may be determined from the radius of curvature R of the track as

$$p_T\ [GeV] = 0.3 \cdot B \cdot R\ [m],$$

where B denotes the magnetic field strength in the chamber. As the spatial resolution of drift chambers is momentum independent, while the curvature of high p_T tracks differs only slightly from an absolutely straight line, the momentum resolution of drift chambers degrades for particles of higher transverse momenta, typically

[1] In the case of drift chambers operated in uniform magnetic fields, the trajectory of charged particles traversing the chamber is a *helix* [133].

as

$$\frac{\sigma_{p_T}}{p_T} \sim const \cdot p_T\ [GeV], \tag{A.1}$$

with a proportionality constant typically of the order of $const \sim 0.1 - 1\%$. The reconstructed tracks are then fitted to a common intersection, the *vertex*, which is the reconstructed position of the primary ep interaction (tracks that are incompatible with a common intersection are excluded from the fit, in order not to bias the reconstruction of the *primary event vertex* with tracks originating from *secondary vertices* that arise from decays of unstable particles). In the fitting procedure, the parameters of tracks included in the fit are modified to match the vertex, so that the transverse momenta, polar and azimuthal angles of the *non-vertex-fitted track* and *vertex-fitted track* hypotheses are in general different. As the vertex-fit normally improves the resolution of the track parameters, the parameters of the vertex-fitted track hypotheses are in general preferred for the momentum reconstruction of charged particles. In general, the reconstructed momenta therefore depend on the vertex reconstruction.

In multi-wire proportional chambers, the drift times to the anode wires are generally not measured. Instead, the trajectories of charged particles traversing the chamber are inferred from the spatial positions of the anode wires detecting the voltage pulses. Accordingly, small distances between anode wires are necessary for a sufficiently precise track reconstruction. As the small distances between anode wires result in small drift times, multi-wire proportional chambers are preferred over drift chambers in trigger applications, where fast signals are necessary.

Appendix B

Observables used for the Identification of Hadronic Tau Decays

For the identification of hadronic tau decays observables are used that are sensitive to the different particle multiplicities and collimation of tau jets and QCD jets. The observables may be grouped into five different categories: the reconstructed multiplicities of charged and neutral particles (*cf.* sections 4.5.2 and 4.5.3), the invariant mass of the jet reconstructed either using the tracks in the tracking detectors or the energy deposits within the tau cone (*cf.* section 4.5.1) in the LAr calorimeter, jet-shape observables (*cf.* section 4.5.4) that describe the distribution of energy deposits in the LAr, the sum of energy deposits in the isolation cone surrounding the tau cone, and estimator variables that are sensitive to specific decay modes of tau leptons.

The precision with which the observables used for tau identification are modelled by the Monte Carlo simulation is checked using a sample of QCD jets in deep-inelastic scattering events selected in the 1999/2000 e^+p data, which is compared with different hadronisation models and detector simulations. The distributions of observables predicted by the "Color Dipole Model" (CDM) are compared with those of the "Matrix Element and Parton Showers" (MEPS) model (*cf.* section 2.4), in order to estimate the uncertainties arising in the Monte Carlo expectation from different hadronisation models. For an estimate of the uncertainties arising from the precision of the detector simulation, the Monte Carlo events generated with the CDM and MEPS hadronisation models are passed through the detector simulation four times, using a fast parametrisation of electromagnetic showers (standard *H1FAST*) [171], a tuned parametrisation (tuned H1FAST) [172] and detailed simulations of electromagnetic showers in the GEANT3 framework (H1SIM) for medium and fine resolutions of the H1 detector geometry; as the simulation of hadronic showers is relatively fast, they are always generated using the fine resolution option and GEANT3.

In the following, the observables used for tau identification are listed with their definitions and their distributions for QCD jets in CC and NC DIS are shown. The uncertainties of the Monte Carlo expectation are illustrated by a shaded band in the figures, indicating the maximum and minimum prediction of the eight different

Monte Carlo sets. In order to access the discriminatory power of the observables, the distributions for tau jets produced in decays of W bosons are shown in addition. The tau jets and QCD jets are reweighted to common energy and polar angle distributions, to make the distributions independent of the different event kinematics of the DIS and W production processes (*cf.* section 4.5.5). The energy and polar angle distributions of the reweighted tau jets and QCD jets are shown in figure B.1.

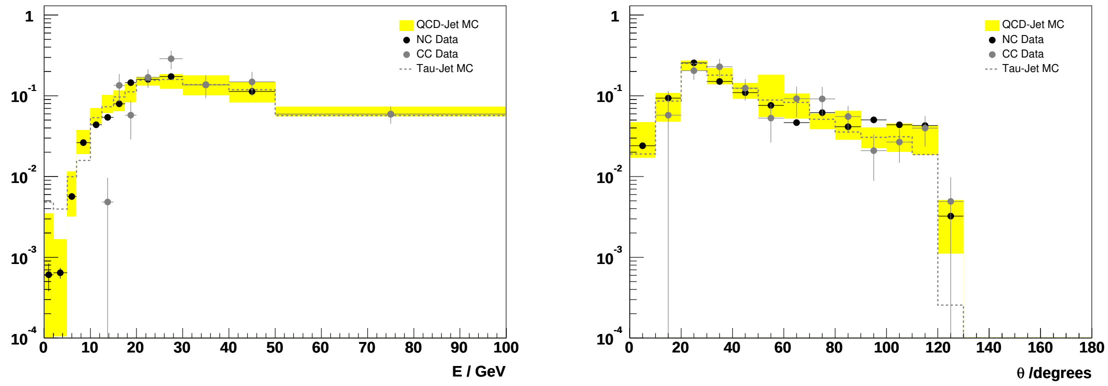


Figure B.1: *Distributions of the energy (left) and polar angle (right) of tau jets and QCD jets after reweighting.*

Particle Multiplicities

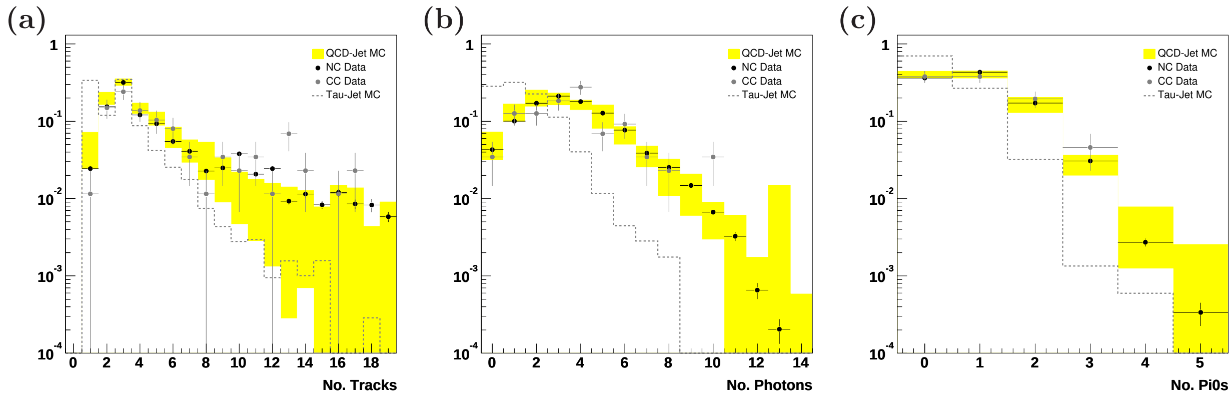


Figure B.2: *Distributions of the number of tracks (a), the number of reconstructed photons (b) and neutral pions (c) for the reweighted tau jets and QCD jets.*

- Number of tracks N_{tracks}
  The number of all vertex-fitted tracks inside the tau cone; note that this number may exceed the multiplicity of charged particles produced in the tau decay, due to photon-conversion.

- Number of photons N_γ
  The number of reconstructed electromagnetic showers inside the tau cone.

- Number of neutral pions N_{π^0}
  The number of neutral pions reconstructed as combinations of electromagnetic showers inside the tau cone.

Invariant Mass of Jet

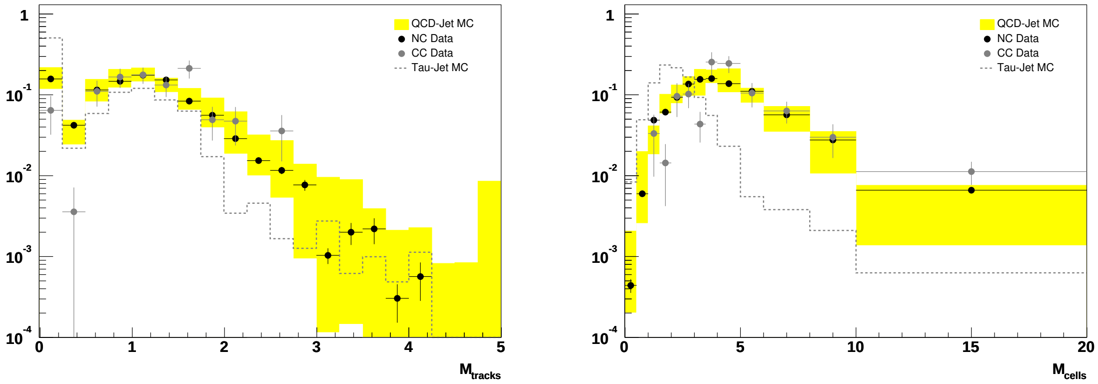


Figure B.3: *Distributions of the invariant mass of jets reconstructed from tracks (left) and energy deposits in the LAr (right) for the reweighted tau jets and QCD jets.*

- Invariant mass M_{tracks}
  The invariant mass of the jet reconstructed using the tracks of charged particles within the tau cone,

$$M_{tracks} = \sqrt{\left(\sum_i E_i\right)^2 - \left(\sum_i p_{x_i}\right)^2 - \left(\sum_i p_{y_i}\right)^2 - \left(\sum_i p_{z_i}\right)^2},$$

  where $E_i, p_{x_i}, p_{y_i}, p_{z_i}$ are the components of the four-vector of charged particle i defined by the measured track momentum and the charged pion mass.

- Invariant mass M_{cells}
  The invariant mass of the jet reconstructed using the energy deposits within the tau cone in the LAr,

$$M_{cells} = \sqrt{\left(\sum_i E_i\right)^2 - \left(\sum_i p_{x_i}\right)^2 - \left(\sum_i p_{y_i}\right)^2 - \left(\sum_i p_{z_i}\right)^2},$$

  where $E_i, p_{x_i}, p_{y_i}, p_{z_i}$ are the components of a four-vector

$$v_i \equiv (E_i, p_{x_i}, p_{y_i}, p_{z_i}) = (E_i, E_i \sin\theta_i \cos\phi_i, E_i \sin\theta_i \sin\phi_i, E_i \cos\theta_i)$$

  defined by the energy E_i deposited in cell i and the polar and azimuthal angles θ_i and ϕ_i of the position of the i^{th} cell.

Shower-Shape Estimators

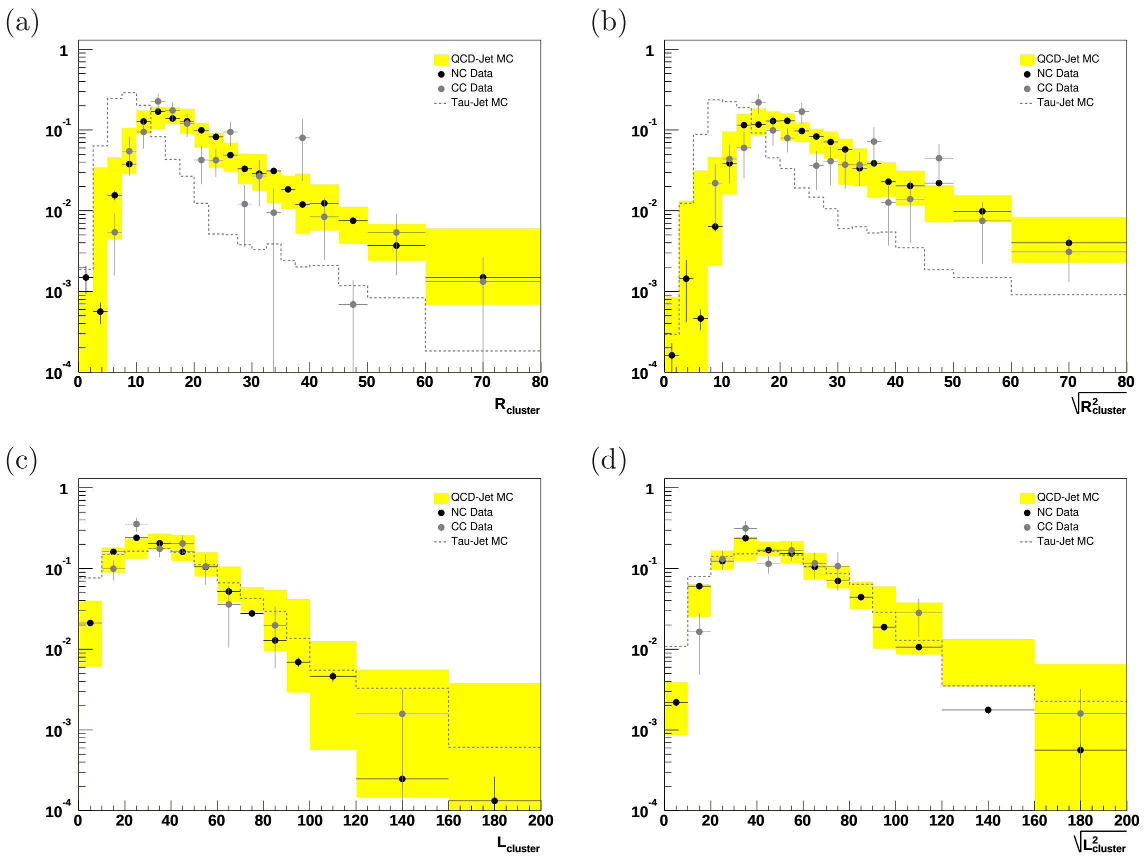


Figure B.4: *Distributions of the first (a) and second (b) radial moments and the first (c) and second (d) longitudinal moments of the energy distribution in the LAr for the reweighted tau jets and QCD jets.*

- First radial moment $< r >$
  The energy-weighted mean distance of closest approach r_i of energy deposits within the tau cone to the axis of the tau jet,

$$< r >= \frac{\sum_i E_i r_i}{\sum_i E_i}.$$

- Second radial moment $< r^2 >$
  The energy-weighted mean squared distance of closest approach r_i^2 of energy deposits within the tau cone to the axis of the tau jet,

$$< r^2 >= \frac{\sum_i E_i r_i^2}{\sum_i E_i}.$$

- First longitudinal moment $< l >$
  The energy-weighted mean distance l_i of the energy deposits within the tau cone to the first electromagnetic layer of the LAr calorimeter

$$< l >= \frac{\sum_i E_i l_i}{\sum_i E_i}.$$

- Second longitudinal moment $< l^2 >$
  The energy-weighted mean squared distance l_i^2 of the energy deposits within the tau cone to the first electromagnetic layer of the LAr calorimeter

$$< l^2 >= \frac{\sum_i E_i l_i^2}{\sum_i E_i}.$$

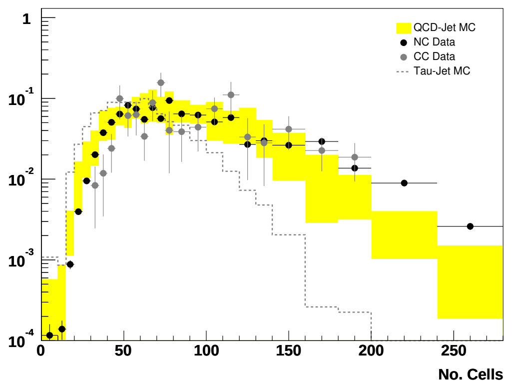


Figure B.5: *Distributions of the number of cells with energy deposits above noise level in the LAr for the reweighted tau jets and QCD jets.*

- Number of cells N_{cells}
  The number of cells with energy deposits above noise level within the tau cone in the LAr.

Isolation Sensitive Estimator Variables

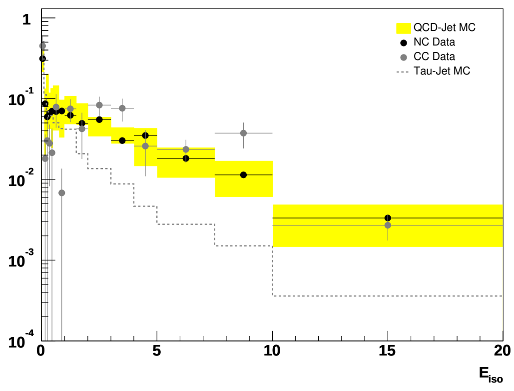


Figure B.6: *Distributions of the sum of energy deposits in the LAr within the isolation cone for the reweighted tau jets and QCD jets.*

- Energy not associated with decay products of the tau lepton E_{iso}
  the sum of energy deposits within the isolation cone that are not contained in the tau cone in the LAr.

Decay-mode Specific Estimator Variables

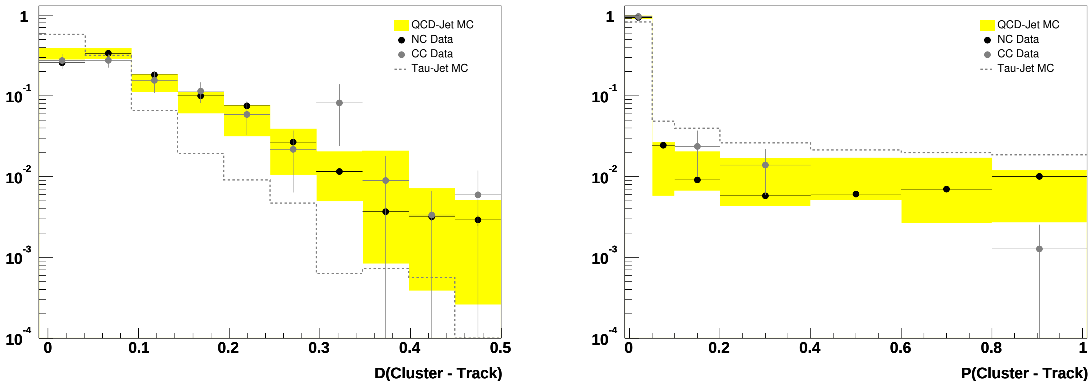


Figure B.7: *Distributions of the cluster-track distance in the η-ϕ metric (left) and the probability for the energy deposits in the LAr to be due to a single charged pion (right) for the reweighted tau jets and QCD jets.*

- Cluster-track distance $D_{cluster-track}$
  The distance in η-ϕ between the directions given by the sum of the track momenta and the energy weighted sum of the positions of the energy deposits within the tau cone in the LAr,

$$D_{cluster-track} = \sqrt{(\eta_{tracks} - \eta_{cells})^2 + (\phi_{tracks} - \phi_{cells})^2}.$$

- Compatibility of the signatures in the tracking detectors and the LAr calorimeter with a single charged pion $P_{cluster-track}$
  The probability for the sum of track momenta and energy deposits within the tau cone in the LAr to be compatible within experimental uncertainties,

$$P_{cluster-track} = 1 - \mathrm{Erf}\,(u)\,, \quad u = \frac{\sum_i E_i - \sum_j P_i}{\sqrt{\left(0.50\sqrt{\sum_i E_i}\right)^2 + \sum_j \sigma^2_{P_j}}},$$

  where the sums extend over all energy deposits i and tracks j within the tau cone and $\mathrm{Erf}\,(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ is the *Gaussian error function*. The uncertainty of the energy deposits in the LAr is estimated from the expected energy resolution of the LAr calorimeter for charged pions (*cf.* section 1.2.2); the uncertainty of the track momenta is determined by the track reconstruction in the central and forward tracking detectors.

- Compatibility of the signatures in the tracking detectors and the LAr calorimeter with an intermediate rho-meson resonance P_ρ
  The probability for the charged pion track and any neutral pion reconstructed in “one-prong” tau jet candidates to be produced by a rho-meson decay. The probability is estimated from the difference between the rho-meson mass $m_\rho \approx 770\ MeV$ and the invariant mass of the reconstructed neutral and charged pions in relation to the decay width $\Gamma_\rho \approx 150\ MeV$ of the rho-meson and the estimated experimental uncertainties,

$$P_\rho = 1 - \mathrm{Erf}\left(\frac{|x|}{\sqrt{2}}\right), \quad x = \frac{M_{\pi^0\pi^\pm} - m_\rho}{\sqrt{\sigma^2_{M_{\pi^0\pi^\pm}} + \Gamma^2_\rho}}.$$

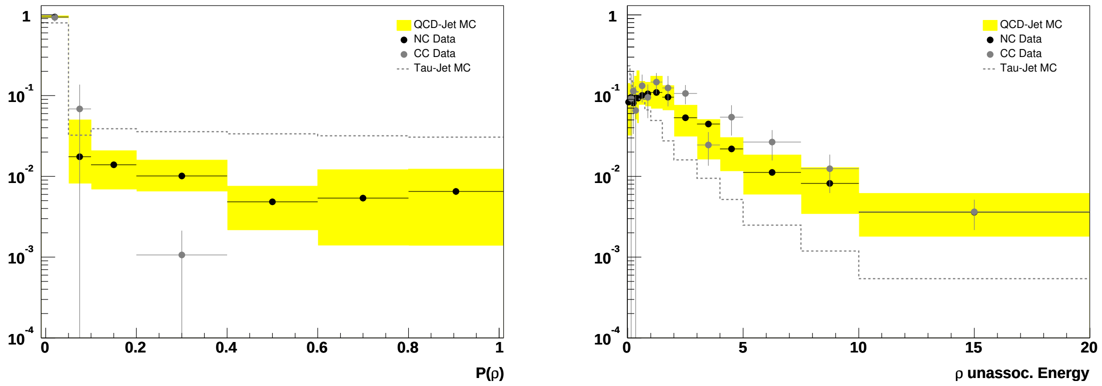


Figure B.8: *Distributions of the probability for an intermediate rho-meson resonance (left) and the sum of energy deposits within the tau cone in the LAr not associated with that resonance (right) for the reweighted tau jets and QCD jets.*

Erf (x) denotes the Gaussian error function. In order to increase the efficiency of the rho reconstruction, an attempt is made to associate singe photons to the charged pion track, if no combination of neutral and charged pions compatible with the rho-meson mass is found.

- Energy not associated with the intermediate rho-meson resonance $E_{\not\rho}$
  The summed energy of the energy deposits within the tau cone that are not attributed to the electromagnetic and hadronic showers induced by the neutral and charged pions associated the rho-meson hypothesis with the highest probability P_{ρ}.

Appendix C

Lepton Identification Studies

The identification of electrons, muons and tau leptons is tested using di-lepton events, produced in lepton-pair production processes or in the decays of J/Ψ mesons into electrons or muons. The cross-sections for these processes are sufficiently large to allow for the lepton identification efficiencies to be checked with sufficient statistical precision for electrons, muons and (with restrictions) for tau leptons. In order to reduce the hadronic background contributing to the samples of di-lepton events used for the estimation of the lepton identification efficiencies, *elastic events*, produced by elastic lepton pair production processes and decays of diffractively produced J/Ψ mesons, are selected. In elastic events, the scattered proton typically escapes undetected through the beam-pipe and only the scattered electron may be detected in addition to the two leptons in the detector. As the scattered electron mostly carries low transverse momentum, the two leptons are typically balanced in the transverse plane, in a characteristic "back-to-back" configuration and are isolated. The elastic events are selected by requiring no tracks or clusters above noise level outside cones of opening angle 0.5 (in the η-ϕ metric) around electrons and muons and of opening angle 1.0 around tau jets. In addition, no signals above noise level are required in the forward muon detector (apart from those associated with muon candidates), the plug calorimeter and the proton remnant tagger. The noise thresholds are those used in the inclusive analysis of diffractive NC DIS processes [60]. The selection of elastic events reduces the combinatorial possibilities for hadrons to be misidentified as leptons and also makes use of the fact that in most hadronic background processes more than two (isolated) hadrons are produced. As a result, the hadronic background is significantly reduced, without requiring tight lepton identification criteria.

C.1 Identification of Electrons

The efficiency with which electrons are identified is checked using elastic di-electron events. These events are selected by requiring two electron candidates in the geometric acceptance of the LAr calorimeter. The first electron candidate, defined as that in the hemisphere with positive azimuthal angle, is required to pass the electron identification criteria listed in table 4.1. The second electron candidate, defined as that in the hemisphere with negative azimuthal angle, is not required to pass any electron identification criteria. In order to determine the electron identification effi-

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within 35 cm around nominal position
Trigger	$ST17 \,\|\, ST33 \,\|\, ST40 \,\|\, ST52 \|$ $ST54 \,\|\, ST61 \,\|\, ST67 \,\|\, ST75$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-ep background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 8$
First electron	$\phi_{e_1} > 0°$ $p_T^{e_1} > 1\ GeV$ "LAr electron" $\|$ "high Kalep Quality" Linked to track in CTD[3]
Second electron	$\phi_{e_2} < 0°$ $p_T^{e_2} > 1\ GeV$ Linked to track in CTD $\|$ FTD[1] Not within $\Delta\phi = \pm 2°$ around a ϕ-crack Not within the z-crack between 15-30 cm
Elastic event	No tracks or clusters above noise threshold[2] No hits in PRT segments $0, 1, 2, 3, 4$ At most one hit in FMD layers θ_1, ϕ_1 At most two hits in FMD layers $\theta_1, \phi_1, \theta_2$ $E_{plug} < 3.5\ GeV$

[1] Passing the selection criteria defined in reference [147]
[2] Not associated with either the first, second or the scattered electron

Table C.1: *Selection requirements for elastic di-electron events. "LAr electrons" are defined to be those passing the identification criteria listed in table 4.1; the "high Kalep Quality" is defined in reference [137].*

ciency, a track in the central or forward tracking detectors is required for the second electron candidate that defines its direction and momentum. The full details of the event selection are listed in table C.1 [1].

With this criteria, 900 events are selected in the 1999/2000 e^+p dataset. The majority of these events are expected to be real di-electron events, produced by elastic electron-pair production processes and decays of diffractively produced J/Ψ mesons into electrons. The distributions of the reconstructed invariant masses and the momenta and polar angles of the electron candidates selected in the data are compared with the Monte Carlo expectation for electron pair and J/Ψ meson production [2] in figure C.1. As can be seen in the figure, the events selected in the

[1] Actually, the event selection is repeated for the complementary assignment of defining the first electron candidate as that in the hemisphere with negative azimuthal angle and the second electron candidate as that in the hemisphere with positive azimuthal angle. In this way, each di-electron event yields up to two estimates of the electron identification efficiency.

[2] The Monte Carlo expectation for electrons produced in decays of J/Ψ mesons is normalized

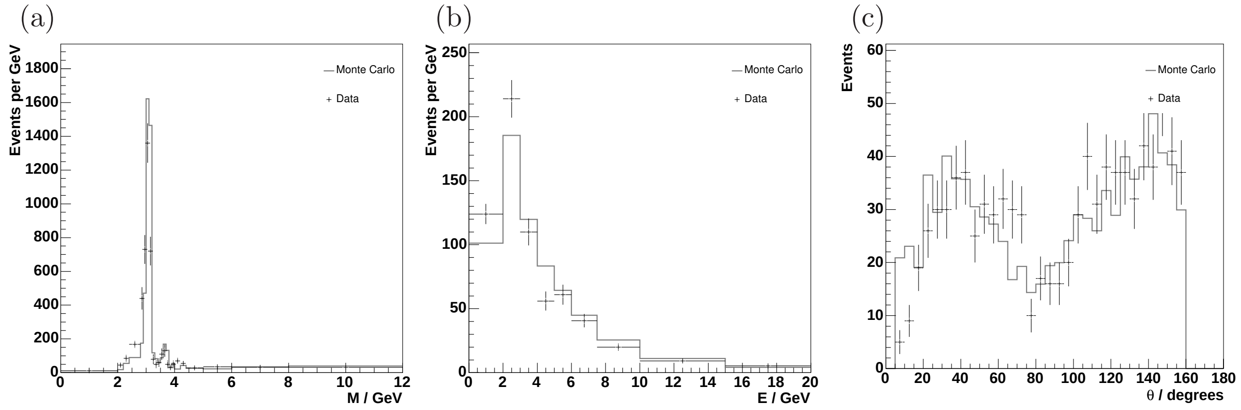


Figure C.1: *Distributions of the invariant mass (a), the momentum (b) and the polar angle (c) of the electron candidates in the events passing the selection criteria listed in table C.1.*

data are compatible with the Monte Carlo expectation for real di-electron events.

Using the event samples selected in the data and the Monte Carlo simulation, the electron identification efficiency may be estimated to be the proportion of tracks in the central or forward tracking detectors that are associated with an electron cluster passing the selection criteria listed in table 4.1 in the LAr calorimeter. The resulting estimates for the electron identification efficiency are shown in figure 4.3 as function of the electron momentum and polar angle.

In the forward region, a difference of about 10% may be seen between the electron identification efficiency in the data and the Monte Carlo simulation. The observed difference may be due to some hadronic background in that region. No obvious background could be found when scanning the events, however, so the observed difference between data and simulation is included in the systematic uncertainties (see chapter 8.1).

C.2 Identification of Muons

The efficiency with which muons are identified is checked using elastic di-muon events. These events are selected by requiring two muon candidates in the polar angle range between 5° and 140°. The first muon candidate, defined as that in the hemisphere with positive azimuthal angle, is required to pass the identification criteria listed in table 4.2. The second muon candidate, defined as that in the hemisphere with negative azimuthal angle, is not required to pass all of the muon identification criteria. In order to determine the muon identification efficiency, a track reconstructed in either the forward muon detector, the instrumented iron or the central or forward tracking detectors is required for the second muon candidate. To ensure that the signature of the second muon candidate is not biased by the trigger conditions, the events are required to be triggered by the first muon. The full details of the event selection are listed in table C.2 [1].

to the data, as the J/Ψ cross-section is predicted by QCD only up to a scale factor, that describes the J/Ψ wave-function [173].

[1] Actually, the event selection is repeated for the complementary assignment of defining the first muon candidate as that in the hemisphere with negative azimuthal angle and the second muon

Quantity	Requirement
Vertex	Reconstructed using CTD tracks z_{vertex} within 35 cm around nominal position
Trigger	$ST14 \parallel ST15 \parallel ST16 \parallel ST18 \parallel ST19 \parallel$ $ST22 \parallel ST34 \parallel ST52 \parallel ST54 \parallel ST56$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system, CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-*ep* background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 6, 7, 8$ all *Ibgfm* bits except $4, 7, 9, 14, 17$ *Ibgam* bits $0, 1, 2, 5, 6, 7, 9$
First muon	$\phi_{\mu_1} > 0^\circ$ "Iron muon"
Second muon	$\phi_{\mu_2} < 0^\circ$ "Iron muon" $\parallel$ track in CTD $\parallel$ FTD[1]
Elastic event	No tracks or clusters above noise threshold[2] No hits in PRT segments $0, 1, 2, 3, 4$ At most one hit in FMD layers θ_1, ϕ_1[3] At most two hits in FMD layers $\theta_1, \phi_1, \theta_2$[3] $E_{plug} < 3.5$ GeV

[1] Passing the selection criteria defined in reference [147]
[2] Not associated with either of the two muons or the scattered electron
[3] Not associated with either of the two muons

Table C.2: *Selection requirements for elastic di-muon events. "Iron muons" are defined to be those passing the identification criteria listed in table 4.2.*

With this criteria, 2712 events are selected in the 1999/2000 e^+p dataset. The vast majority of these events are expected to be real di-muon events, produced by elastic muon-pair production processes and muonic decays of diffractively produced J/Ψ mesons. The distributions of the reconstructed invariant masses and the momenta and polar angles of the muon candidates selected in the data are compared with the Monte Carlo expectation for muon pair and J/Ψ meson production [2] in figure C.2. As can be seen in the figure, the events selected in the data are compatible with the Monte Carlo expectation for real di-muon events.

Using the event samples selected in the data and the Monte Carlo simulation, the efficiencies for the reconstruction of tracks in the forward muon detector, the instrumented iron and the central and forward tracking detectors may be estimated. The track reconstruction efficiency in the forward muon detector is estimated to be the proportion of tracks in the FTD that are associated with a track passing the quality criteria listed in table 4.2 in the FMD. Likewise, the efficiency of the track

candidate as that in the hemisphere with positive azimuthal angle. In this way, each di-muon event yields up to two estimates of the muon identification efficiency.

[2] See footnote 2 on page 219.

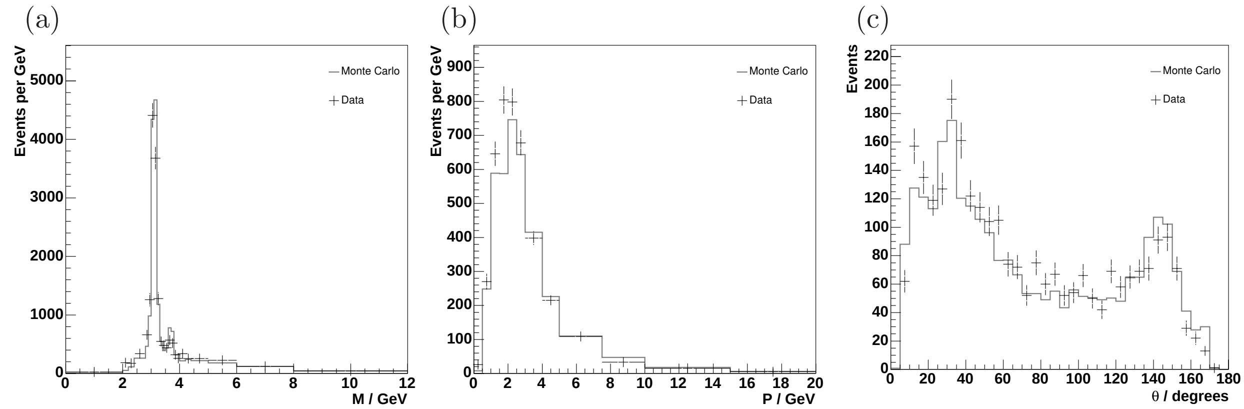


Figure C.2: *Distributions of the invariant mass (a) and the momentum (b) and polar angle (c) of the muon candidates in the events passing the selection criteria listed in table C.2.*

reconstruction in the instrumented iron is estimated to be the proportion of tracks in the central or forward tracking detectors that are associated with a track passing the quality criteria listed in table 4.2 in either the forward or backward endcaps or the barrel region of the instrumented iron. The track reconstruction efficiency in the CTD and FTD is estimated to be the proportion of tracks in the FMD or instrumented iron that are associated with tracks in the central or forward tracking detectors passing the quality criteria given in reference [147]. The estimates for the track reconstruction efficiencies in the forward muon detector, the instrumented iron and the central and forward tracking detectors are shown in figure C.3 as function of the muon momentum and polar angle.

The track reconstruction efficiency in the instrumented iron as function of the muon momentum is seen to differ in Monte Carlo and data. The simulation slightly overestimates the probability that low energy muons reach the iron. Possible reasons for this may be that some detector materials are missing or that rare collision processes with high energy transfers are neglected [140] in the Monte Carlo simulation, so that the energy loss of muons on their way to the instrumented iron is underestimated. For muons resulting from W decays this effect is negligible, however. More important is a potential difference between the data and the Monte Carlo simulation in the forward region. In the efficiencies for track reconstruction in the forward muon detector and the instrumented iron, a difference of the order of $10-20\%$ may be seen. The observed difference may be due to some hadronic background in that region. It seems unlikely that the track reconstruction efficiencies in both detectors are incorrectly modelled by the Monte Carlo simulation to a similar extent, while in the central region the track reconstruction efficiency in the instrumented iron is quite well modelled. As the selected events have not been scanned in order to check for hadronic background, the observed difference between data and simulation is included in the systematic uncertainties (see chapter 8.1).

Given the estimates for the track reconstruction efficiencies in the forward muon detector, the instrumented iron and the central and forward tracking detectors, the muon identification efficiency may be approximated as a function of the momentum

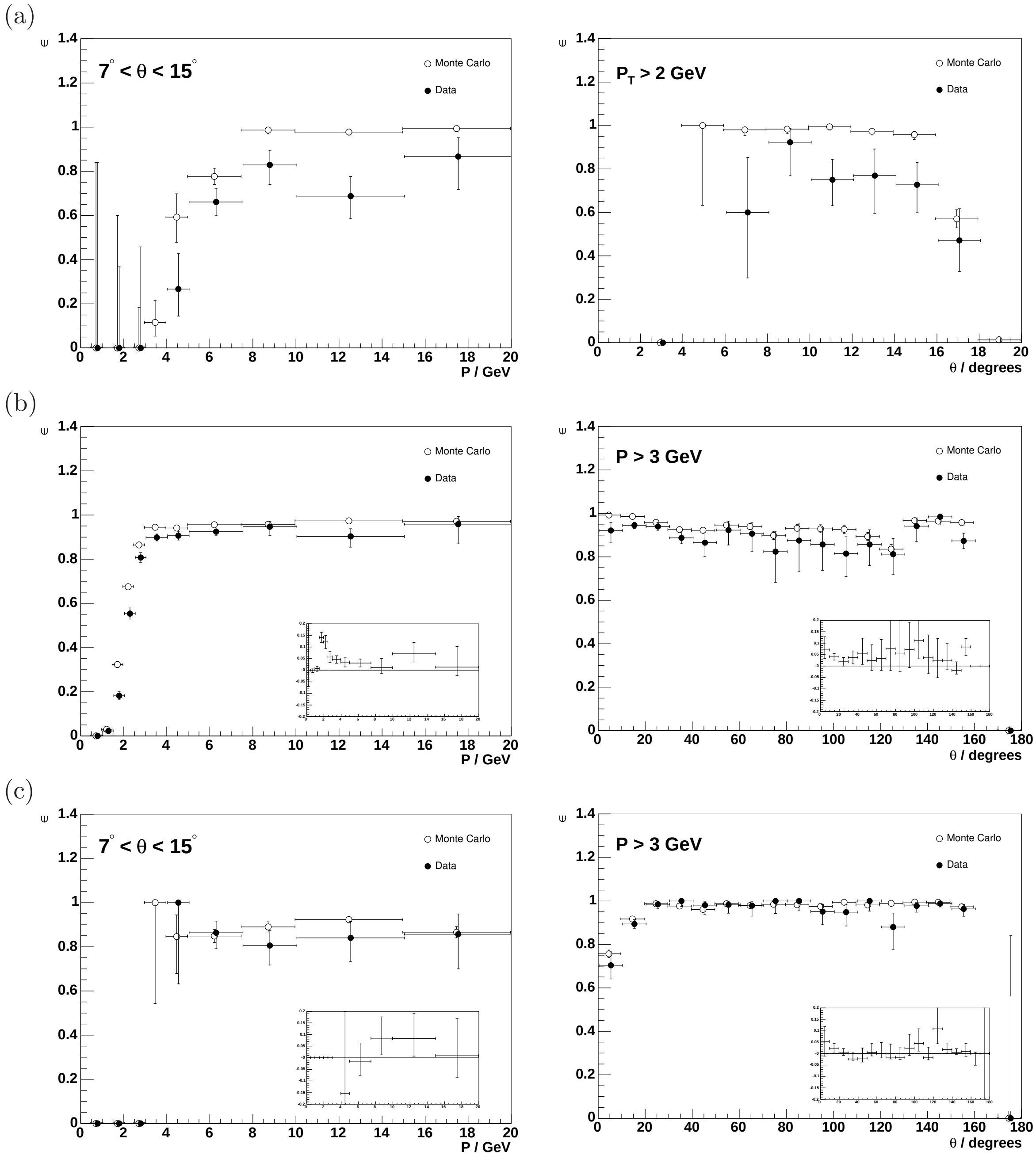


Figure C.3: *Reconstruction efficiencies for tracks in the forward muon detector (a), the instrumented iron (b) and the central and forward tracking detectors (c). The insets show a magnified view of the difference between the data and the Monte Carlo simulation.*

p and the polar angle θ of the muon candidates to be:

$$\varepsilon_{muonId}\left(p/\theta\right) = 1 - \left(1 - \varepsilon_{FMD}\left(p/\theta\right)\right)\left(1 - \varepsilon_{instrIron}\left(p/\theta\right) \cdot \varepsilon_{FTD/CTD}\left(p/\theta\right)\right). \quad \text{(C.1)}$$

Formula C.1 is evaluated using the track reconstruction efficiencies shown in figure C.3 (bin by bin). The resulting estimates for the muon identification efficiency are shown in figure 4.7.

C.3 Identification of hadronic Tau Decays

The ability of the H1 detector to identify hadronic tau decays is checked using di-tau events, produced in elastic tau pair production processes. The strategy for cross-checking the identification of hadronic tau decays is a little different to that used for the cross-checks of the electron and muon identification, due to the significantly smaller cross-section for tau pair production, resulting from the higher mass of the tau lepton [1]. As a consequence of the significantly lower number of signal events and comparatively higher backgrounds, the identification of hadronic tau decays is cross-checked by studying the agreement between data and Monte Carlo simulation in a sample of events containing two identified tau leptons. One of the tau leptons is required to decay leptonically and the other to decay hadronically, in order to reduce the background contribution to the selected event sample. The full details of the event selection are listed in table C.3. In order to avoid the misidentification as hadronic tau decays of electrons (muons) produced by electron (muon) pair production processes, additional neural networks are trained for the rejection of electrons (muons). The same neural networks are used for the identification of tau jets as for the events with isolated tau leptons and large missing transverse momentum.

With this criteria, 27 events are selected in the 1996-2000 e^+p and e^-p datasets, in agreement with the 29.6 expected from the Monte Carlo simulation. Of the 29.6 expected events, 22.3 result from tau pair production processes, 2.2 (2.5) from electron (muon) pair production processes and 2.6 from diffractive neutral current scattering processes. The number of candidate events selected in the individual decay modes of the tau pairs is shown in table C.4, together with the number of events expected for the tau pair production signal and the different background processes. The properties of the events selected in the data are listed in detail in table C.5. The number of candidate events in each of the individual decay modes is found to be in agreement with the Monte Carlo expectation. As an example of the events selected, a candidate for a tau pair decay into a "one-prong" tau jet and a muon is displayed in figure C.4. The properties of all selected events are illustrated in figure C.5, which shows the distributions of $p_T^{lep_\tau}$, θ_{lep_τ}, $p_T^{had_\tau}$, θ_{had_τ}, P_T^{miss} and $\sum E - P_z$.

The precision with which the electron and muon pair production and diffractive NC DIS backgrounds are modelled by the Monte Carlo simulation is cross-checked with background enriched event samples. Based on the selection requirements listed in table C.3, the electron (muon) pair production backgrounds are enhanced by selecting events with "one-prong" hadronic tau decays without cutting on the likelihoods $f_{NN}^{one-prong}$, $f_{NN}^{three-prong}$ and $f_{NN}^{\not e}$ ($f_{NN}^{\not\mu}$). Similarly, the diffractive DIS NC background is enhanced by omitting the cuts on the likelihoods $f_{NN}^{one-prong}$ and $f_{NN}^{three-prong}$ and the missing longitudinal momentum δ^{miss}. The numbers of events obtained using these by the background enriched selections in the 1996-2000 e^+p and e^-p data and the Monte Carlo expectations for the various processes are presented in table

[1] In addition to the significantly smaller production cross-section, the ratio of di-tau to di-electron and di-muon events decreases further once conditions on the detected particles are applied. In contrast to electrons and muons whose momenta are wholly recontructed, the momenta of tau leptons are only partially reconstructed, due to the presence of undetected momentum carried away by the neutrinos that are produced in the tau decays.

Quantity	Requirement
Vertex	Reconstructed using CTD tracks
	z_{vertex} within 35 cm around nominal position
Trigger	$ST15 \parallel ST17 \parallel ST19 \parallel ST22 \parallel ST33 \parallel$
	$ST34 \parallel ST40 \parallel ST52 \parallel ST54 \parallel ST56 \parallel$
	$ST61 \parallel ST66 \parallel ST67 \parallel ST71 \parallel ST75 \parallel ST77$
High voltage	CJC1, CJC2, LAr, ToF, luminosity system,
	CIP, COP, instr. iron, FMD
Event timing	t_0^{CJC} within 4.8 ns around nominal bunch crossings
Non-ep background finder	*Ibg* bits $0, 1, 2, 3, 4, 5, 7, 8$
	Ibgam bit 7
Had. tau	$\theta_{jet}\ \epsilon [5°, 120°]$
	$p_T^{jet} > 2\ GeV$
	No. tracks either $1 \parallel 3$
	$f_{NN}^{one-prong} > 0.75 \parallel f_{NN}^{three-prong} > 0.75$
	$f_{NN}^{\not e} > 0.75$[1]
	$f_{NN}^{\not \mu} > 0.75$[2]
Lep. tau	$\theta_\ell\ \epsilon [5°, 140°]$
	"LAr electron" $\parallel$ "Iron muon"
Sum of had. + lep. tau charges	0
Elastic event	No tracks or clusters above noise threshold[3]
	No hits in PRT segments $0, 1, 2, 3, 4$
	At most one hit in FMD layers θ_1, ϕ_1[4]
	At most two hits in FMD layers $\theta_1, \phi_1, \theta_2$[4]
	$E_{plug} < 3.5\ GeV$
Diffractive DIS NC rejection	$\delta^{miss} > 5\ GeV$[5]

[1] If event is candidate for "one-prong" + e
[2] If event is candidate for "one-prong" + μ
[3] Not associated with the decay particles of either the hadronic or leptonic tau decay or the scattere
[4] Not associated with muons
[5] If only one electron is detected, which has the same charge as the beam-electron

Table C.3: *Selection requirements for elastic di-tau events in which one tau lepton decays leptonically into either an electron or muon and the other hadronically into either one ("one-prong") or three ("three-prong") charged particles. "LAr electrons" and "Iron muons" are defined to be those passing the identification criteria listed in tables 4.1 and 4.2, respectively.*

C.6. As can be seen in the table, the data are found to be in agreement with the Monte Carlo expectation in the background enriched samples, supporting the assumption that the event sample selected using the requirements listed in table C.3 is indeed dominated by tau pair events. As an additional cross-check, the number of events passing the selection criteria listed in table C.3, but containing an electron or muon of the same charge sign as the tau jet has been checked and is found to be in agreement with the Monte Carlo expectation.

Decay Mode	Data	$\sum$ SM	$\tau\bar{\tau}$	$e\bar{e}$	$\mu\bar{\mu}$	NC
"one-prong" + e	6	5.8	2.9	1.6	0.1	1.2
"three-prong" + e	4	4.3	2.3	0.6	0	1.4
"one-prong" + μ	14	14.3	11.9	0	2.4	0
"three-prong" + μ	2	5.5	5.5	0	0	0

Table C.4: *Number of events in the 1996-2000 e^+p and e^-p data observed in the individual decay modes of the tau pair, compared with the the Monte Carlo expectation for the tau pair production signal and various background processes.*

Run Event	Decay Mode	had. Tau E_{jet} $[GeV]$	had. Tau θ_{jet} [°]	had. Tau ϕ_{jet} [°]	lep. Tau E_ℓ $[GeV]$	lep. Tau θ_ℓ [°]	lep. Tau ϕ_ℓ [°]	scattered Electron E_e $[GeV]$	scattered Electron θ_e [°]	scattered Electron ϕ_e [°]
157475 1459	$\mathbf{1}+\mu$	29.1	14.2	-115.3	9.6	59.9	67.5	19.9	177.0	109.9
162609 35995	$\mathbf{1}+\mu$	3.1	99.1	141.9	3.3	136.7	-27.3		-	
164554 59381	$\mathbf{1}+e$	14.6	52.3	72.0	16.2	21.7	-104.6		-	
192491 36465	$\mathbf{1}+\mu$	12.8	31.9	-179.9	19.9	16.2	-4.5		-	
193050 108609	$\mathbf{1}+e$	24.2	24.2	98.1	78.8	8.3	-84.0		-	
197767 92799	$\mathbf{1}+e$	19.4	14.5	158.6	6.2	101.9	-25.4		-	
247254 18369	$\mathbf{1}+e$	14.5	27.8	-5.3	10.3	52.7	-155.9	22.9	171.1	89.1
252376 102188	$\mathbf{3}+e$	14.1	50.1	76.5	12.5	33.5	-106.8	23.3	173.0	-99.4
256875 67724	$\mathbf{1}+\mu$	3.8	44.6	-114.1	2.4	58.2	71.0		-	
257051 87929	$\mathbf{3}+e$	2.8	73.7	114.3	5.2	91.2	-80.0	20.5	160.7	103.6
257693 172622	$\mathbf{1}+\mu$	4.6	64.9	-70.2	4.7	122.1	88.4		-	
260218 37994	$\mathbf{1}+\mu$	16.0	36.8	-80.6	7.6	87.4	102.0		-	
260299 8124	$\mathbf{1}+e$	7.6	32.1	80.9	6.6	74.1	-66.1		-	
263791 12908	$\mathbf{3}+e$	5.7	43.5	-88.3	9.7	32.1	91.1		-	
264975 46474	$\mathbf{3}+\mu$	5.0	23.4	41.0	6.3	57.9	176.5	23.9	163.6	-23.9
265354 20258	$\mathbf{3}+\mu$	4.7	47.9	-23.4	2.4	95.7	147.4		-	
266010 89149	$\mathbf{1}+\mu$	5.0	38.9	94.8	5.8	117.2	-101.1		-	
267521 18270	$\mathbf{1}+\mu$	3.0	79.3	149.8	7.3	18.2	25.1	26.1	171.6	-137.4
267962 55026	$\mathbf{3}+e$	6.6	36.7	162.3	5.4	132.4	-19.2		-	
270048 12714	$\mathbf{1}+\mu$	2.9	73.3	-67.0	3.1	62.2	113.7		-	
272689 54975	$\mathbf{1}+\mu$	7.1	30.8	-40.0	3.4	68.6	160.7		-	
275724 119702	$\mathbf{1}+\mu$	4.2	102.0	152.1	3.4	111.5	-17.1		-	
276382 12042	$\mathbf{1}+\mu$	8.1	114.2	-100.7	5.8	71.6	70.7		-	
276828 8760	$\mathbf{1}+\mu$	54.8	16.5	35.8	4.6	45.7	-144.3		-	
277029 30906	$\mathbf{1}+e$	9.9	46.9	-76.8	7.8	134.8	106.9		-	
278348 83217	$\mathbf{1}+\mu$	7.7	38.6	112.9	3.5	49.4	-64.0		-	

Table C.5: *Properties of the selected candidate tau pair events.*

All in all, the data are compatible with the Monte Carlo expectation, which demonstrates the hadronic tau decay identification capability of the H1 detector.

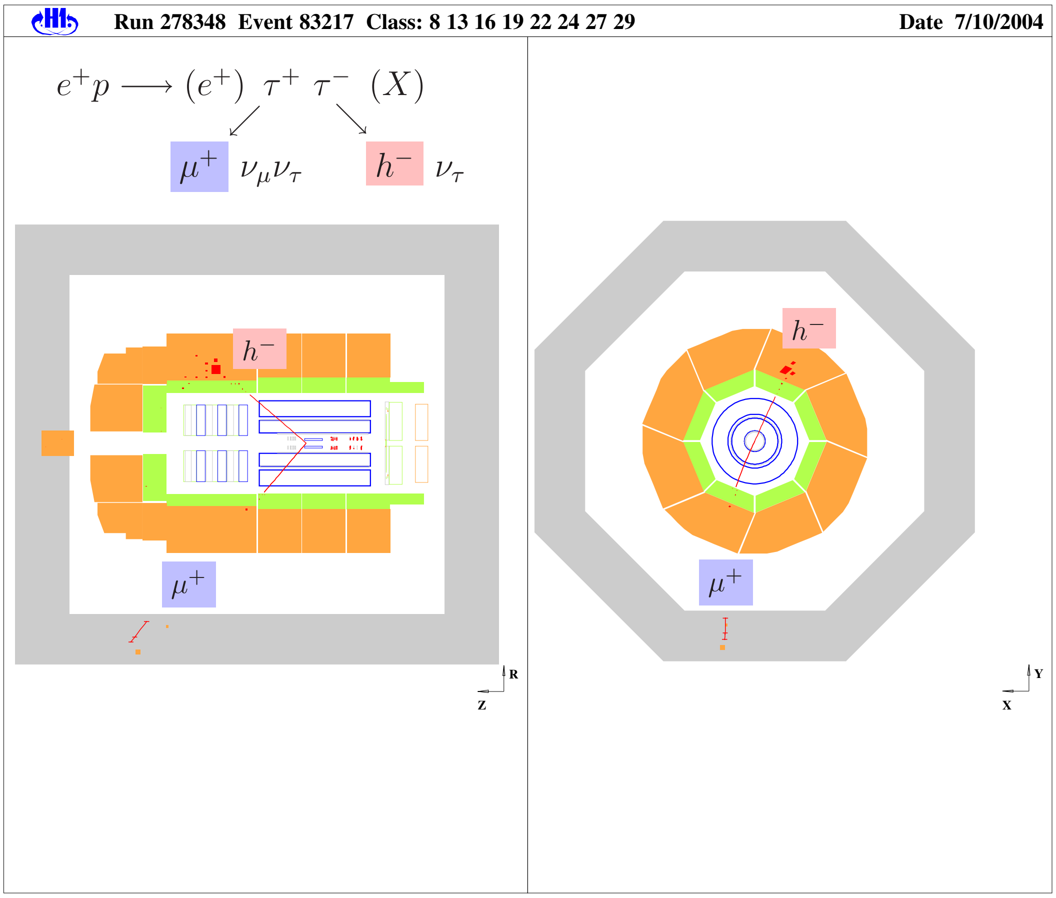


Figure C.4: *Display of one of the selected candidate tau pair events.*

Enriched Selection	Data	$\sum$ SM	$\tau\bar{\tau}$	$e\bar{e}$	$\mu\bar{\mu}$	NC
$e\bar{e}$	170	194.2	15.1	174.3	1.8	3.0
$\mu\bar{\mu}$	27	25.6	10.6	2.3	10.3	2.4
NC	53	41.1	14.7	4.0	1.2	21.1

Table C.6: *Number of events in the background enriched samples in the 1996-2000 e^+p and e^-p data, compared with the Monte Carlo expectation for the tau pair production signal and various background processes.*

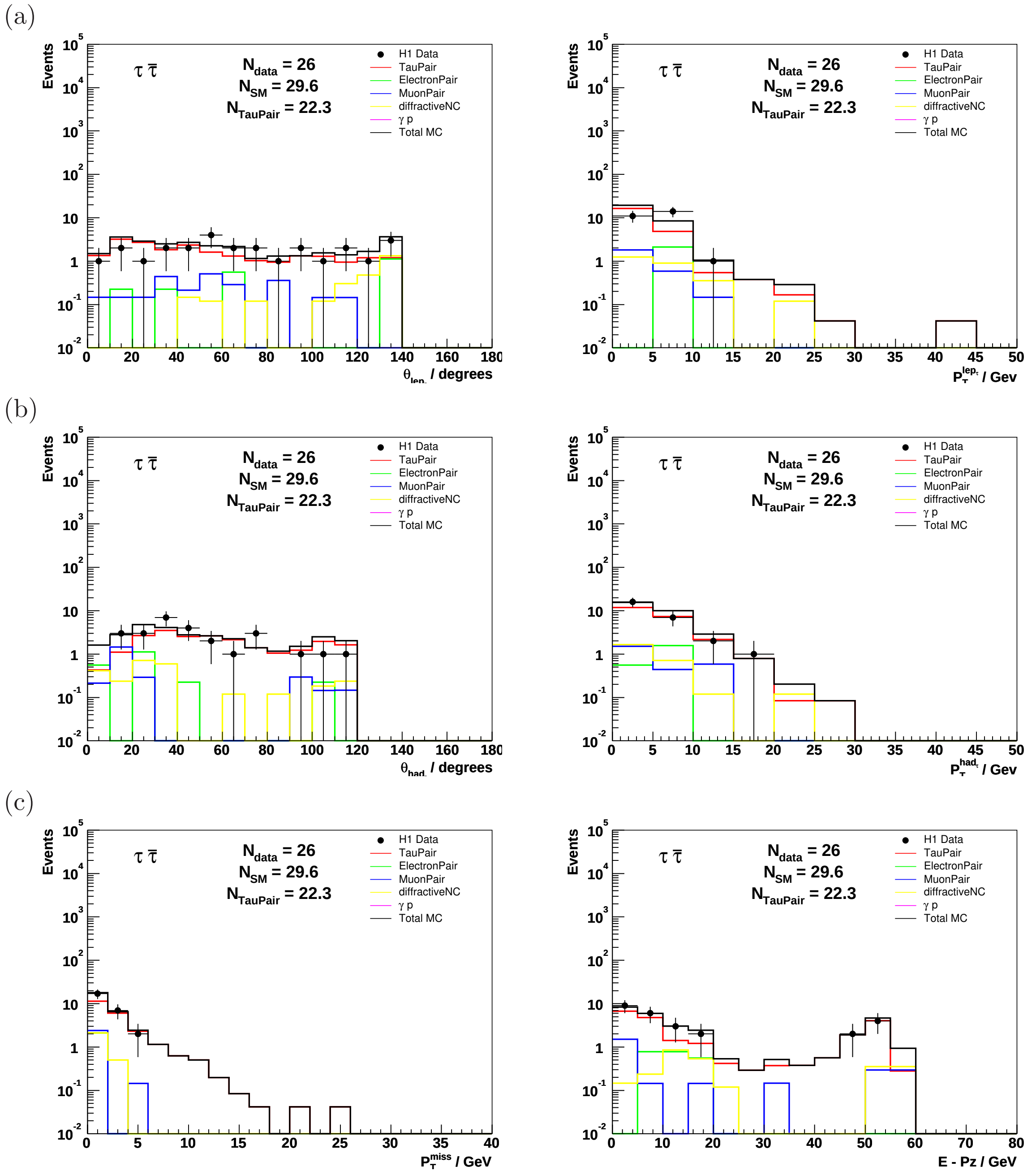


Figure C.5: *Distributions of the polar angles and transverse momenta of the leptons produced in the tau decays (a), the polar angles and transverse momenta of the tau jets (b) and the missing transverse momentum and* $\sum E - P_z$ *(c) of the candidate events for semi-leptonic tau decays.*

Appendix D

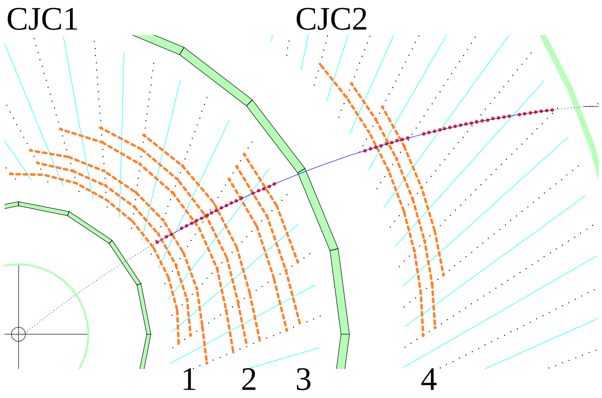
Development of Subtrigger for High p_T Muon Events

The recent HERA upgrade increased the instantenous luminosity of the HERA collider by about a factor of five (design value [12]) with respect to the HERA I value. The rate of *ep* interactions increases by the same factor, resulting in a five times higher trigger rate after the upgrade. As the available bandwidth of the H1 experiment remains unchanged, prescale factors have to be introduced or raised for many subtriggers, effectively diminishing the trigger efficiency of the affected subtriggers. In order to avoid prescaling the subtriggers, their selectivity needs to be improved to reduce the trigger rate. Accordingly, additional trigger systems were installed in the H1 experiment as part of the HERA II upgrade. The most relevant of these for the triggering of events with isolated leptons and large missing transverse momentum are the *fast track trigger* (FTT) [50], the trigger system of the new CIP [18], and possibly the new *jet-trigger* [174] of the LAr calorimeter.

Of the isolated lepton events, the events with isolated muons are most affected by the introduction of prescale factors, especially in the phase-space region of low hadronic transverse momentum P_T^X: unlike the events with isolated electrons and tau leptons, most events with isolated muons are not triggered by the subtriggers $ST66$, $ST67$, $ST71$ [1], $ST75$, and $ST77$ which are based on energy deposits in the LAr (and will be operated without prescales in the future), as muons do not deposit a significant fraction of their energy in the calorimeter. In order to increase the trigger efficiency for events with high p_T muons in the HERA II running period, a new subtrigger was developed, whose rate is expected to be low enough for it to be operated free of prescales. The rate reduction in comparison to existing muon triggers is made possible by the new fast track trigger system.

In the following, the development of the new subtrigger for events with high p_T muons is described. In the first two subsections, the FTT and muon trigger systems are described, before the definition of the new subtrigger is detailed in the third subsection (see reference [18] for a description of the CIP trigger system). Estimates of the trigger efficiency and rate of the new subtrigger expected at nominal HERA II running conditions are presented in the fourth subsection. As the development of the new subtrigger is independent of the analysis of the HERA I data presented in

[1] For technical reasons, the definition of $ST71$ in the HERA II differs slightly from that in the HERA I period.



Trigger Layer	Chamber	Wire Layers	Number of Trigger Groups
1		3, 5, 7	
2	CJC1	10, 12, 14	30 (per Trigger Layer)
3		18, 20, 22	
4	CJC2	4, 5, 8	60

Figure D.1: *Trigger layers of the FTT in CJC1 and CJC2. In total, twelve layers of anode wires are used for track reconstruction by the FTT, grouped into four trigger layers of three wire layers each. The wire layers in CJC1 and CJC2 that are included in the trigger layers are listed in the table.*

the body of this thesis, this section is concluded separately in the fifth subsection.

D.1 Fast Track Trigger

The *fast track trigger* (FTT) is a new track-based trigger system that provides three-dimensional track reconstruction with high precision in real time. It was installed as part of the HERA upgrade programme to enhance the selectivity of the H1 trigger system, as is necessitated by the higher event rates at the increased luminosity. In particular, the fast track trigger was designed to identify rare particle decays by identifying the characteristic mass of the resulting final state. In comparison to the $DCR\phi$ trigger system [175] used for track triggering in the HERA I running period, the FTT provides a significantly improved transverse momentum resolution that makes it useful also for the triggering of rare high p_T processes.

For track reconstruction, the FTT uses the analogue signals from twelve layers of anode wires in the central jet chamber. As is displayed in figure D.1, these are grouped into four *trigger layers*, each containing three layers of anode wires. Of the four trigger layers, three are located in CJC1 and one in CJC2. The trigger layers in CJC1 (CJC2) are divided into 30 (60) *trigger groups* in azimuth, corresponding to the segmentation of the CJC (*cf.* section 1.2.1). The trigger groups contain a total of five anode wires each; in addition to the three wires of the drift cell corresponding to the trigger group, the innermost wire of the adjacent drift cell on the left and the outermost wire of the adjacent drift cell on the right are included in the trigger group, in order to increase the reconstruction efficiency for tracks crossing the boundary between neighbouring drift cells [1].

The analogue signals provided by the twelve wire layers of CJC1 and CJC2 are processed by three consecutive trigger levels in the FTT. The different tasks of the three trigger levels are introduced in figure D.2.

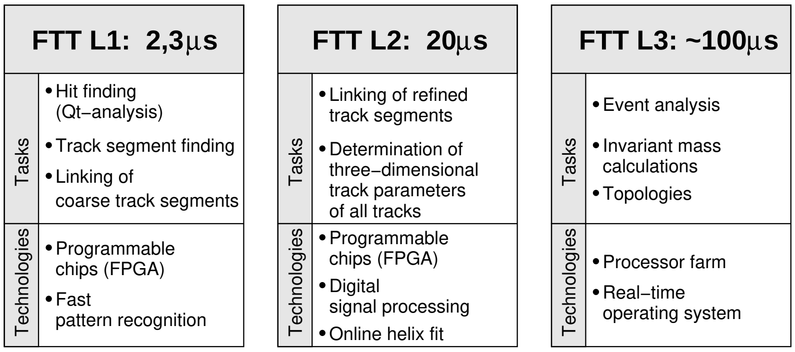


Figure D.2: *The different tasks of the three trigger levels of the FTT (taken from [176]).*

During the 2.3 μs available for the decision of the first trigger level, coarse tracks are reconstructed in the transverse plane, with a transverse momentum resolution of [177]

$$\frac{\sigma_{p_T}^{L_1}}{p_T} \approx 0.40 \cdot p_T \; [GeV]$$

and a resolution in azimuthal angle of

$$\sigma_{\phi}^{L_1} \approx 3^\circ.$$

After each bunch crossing, the track reconstruction algorithm of the FTT starts to search for hits on the anode wires of the CJC. The drift-times to the anode wires and the position of these hits along the wires are estimated by analysing the arrival times of the signal pulses at both wire ends (*Q-t analysis*). From the estimated drift-times, coarse track segments are reconstructed in the $r-\phi$-plane. This is done independently in the various trigger layers by matching the estimated drift-times with a set of expectations that have been calculated in advance for a variety of potential tracks produced at the nominal interaction point. In order to allow the matching to be performed within the short period available for the trigger decision, a dedicated fast pattern recognition algorithm was developed [176]. If at least two track segments with similar curvature and azimuthal angle are found in a combination of the four trigger layers, the track segments are linked to form an *$L1$ track*. As the track reconstruction is restricted to the transverse plane, only the transverse momentum and azimuthal angle of the reconstructed tracks are determined at the first trigger level. Accordingly, the main trigger elements provided by FTT at $L1$ are the multiplicities of tracks above various p_T thresholds. Reconstruction of the vertex position along the beam-pipe using the z-coordinates of the hits is foreseen, but is not yet implemented.

[1] The apparent asymmetry results from the inclination of the anode wire planes with respect to the radial direction and geometrical considerations [45].

During the 20 μs latency period of the second trigger level, the track parameters are re-evaluated with higher precision and a full three-dimensional track reconstruction is performed, yielding resolutions of [176]

$$\frac{\sigma_{p_T}^{L_2}}{p_T} \approx 0.02 \cdot p_T \; [GeV], \qquad \sigma_\theta^{L_2} \approx 4^\circ, \qquad \sigma_\phi^{L_2} \approx 0.1^\circ,$$

in transverse momentum and the polar and azimuthal angles, respectively. These values are close to the intrinsic resolution of the central jet chamber. At the second trigger level, refined track segments are reconstructed with an improved drift-time resolution in the r-ϕ-plane, and a more precise track search is performed by the linking algorithm. The linked track segments are then fitted to a helical trajectory in three dimensions. In the first step of this procedure, the transverse momentum and azimuthal angle of each track are determined by fitting a circle to the track segments in the r-ϕ-plane. In the second step, the polar angle of each track is reconstructed in the r-z-plane, by fitting a straight line to the r- and z-coordinates of the hits associated with the track segments. The polar angles of the resulting *L2 tracks* may be improved by a vertex constraint, reconstructed either by the first trigger level of the FTT or obtained from the CIP trigger via the L2L3 bus [45]. The fit significantly improves the resolution of the track parameters and is the main reason for the improved precision of the track reconstruction at the second trigger level. The trigger elements provided by the FTT $L2$ are again the multiplicities of tracks above various transverse momentum thresholds. In comparison to the first trigger level, however, the trigger rate of the second trigger level may be significantly reduced by applying higher p_T thresholds, using the improved resolution.

The period of 100 μs available for the decision at the third trigger level may be used for partial event reconstruction, in order to identify rare particle decays or final states associated with rare processes. Trigger information from other subdetectors that is available via the L2L3 bus may be used in combination with the tracks reconstructed by the second trigger level of the FTT to enhance the selectivity of the trigger decision.

For trigger studies like that described in the following, the performance of the fast track trigger is simulated using the *FTTEMU* [178] software package.

D.2 Muon Trigger

Triggering on muons which penetrate the instrumented iron is performed using the hits on the anode wires of the limited streamer tubes. Trigger information is provided by five of the ten wire layers in the slits between the iron plates. For the purposes of processing the trigger information and generating the trigger signals, the instrumented iron is segmented into 64 independent *modules*, 16 in each of the forward and backward endcaps and 32 in the barrel. The modules are grouped into five *trigger sectors*, the forward inner (*FIEC*) and outer (*FOEC*) endcaps, the barrel (*Bar*), and the backward inner (*BIEC*) and outer (*BOEC*) endcaps. Of the sixteen modules in each of the two endcaps, the innermost six modules around the beam-pipe are associated with the inner trigger sectors, and the other ten with the outer trigger sectors of the endcaps. The 64 modules in the instrumented iron and the different trigger sectors are shown in figure D.3.

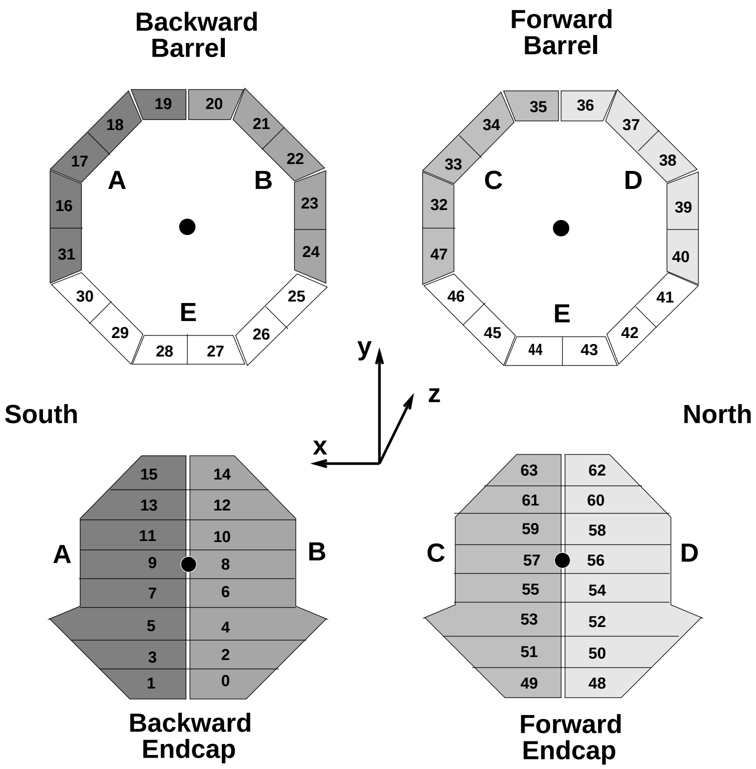


Figure D.3: *Segmentation of the instrumented iron into modules. The forward and backward parts of the barrel are shown separately, as are the forward and backward endcaps. The position of the beam-pipe is indicated by black dots. The innermost six modules around the beam-pipe (module numbers* $54-59$ *and* $6-11$*) are assigned to the* $FIEC$ *and* $BIEC$ *trigger sectors, respectively.*

Trigger Sector	Coincidence Requirement
Backward EndCap	3 hits in layers $0, 1, 2, 5, 9$
Barrel	2 hits in layers $0, 1, 2, 5$
Forward Outer EndCap	3 hits in layers $0, 1, 2, 5, 9$
Forward Inner EndCap	4 hits in layers $0, 1, 2, 5, 9$

Table D.1: *Coincidence requirements for modules in different parts of the instrumented iron. In the barrel, only four layers are used. The quoted layer numbers refer to the first, second, third, sixth and last layer of limited streamer tubes* between *iron plates.*

The trigger signals are generated independently by individual modules. They are defined by coincidences of hits in the five layers of limited streamer tubes that provide trigger information. In each module, all hits on the anode wires of one layer are combined to form a single plane signal, each of the plane signals representing the presence of at least one hit in the limited streamer tubes in the corresponding layer of the module. The trigger signal for individual modules is determined by the number of single plane signals in the five trigger layers of the module. The minimum number of coincidences required to generate a trigger signal is listed in table D.1. In the forward region more coincidences are required in order to suppress the trigger

Trigger System	Trigger Element
	L1
FTT	$p_T > 0.9\ GeV$ track
Muon Trigger	$Mu\_FOEC \parallel Mu\_Bar \parallel Mu\_BOEC$
CIP Trigger	$zVtx\_t0$ $zVtx\_sig > 0$
Veto Walls	$!\,(VETOInnerBg \parallel VETOOuterBg)$
	L2
FTT	$p_T > 5.0\ GeV$ track
	L3
FTT - Muon Trigger (link)	Extrapolated distance to module boundary $< 50\ cm$

Table D.2: *Definition of the new subtrigger for events with high p_T muons by combination of trigger elements at the first, the second, and the third trigger level.*

rate resulting from the hadronic activity in the proton direction.

The trigger signals generated by the individual modules are then combined to generate trigger elements for the first trigger level. Eight trigger elements are defined by forming logical combinations of the trigger signals from the individual modules. Five of the eight trigger elements are defined by the presence of single trigger signals in each of the five trigger sectors, respectively. They can be used to trigger on events with single muons penetrating different regions of the instrumented iron. The other three trigger elements require multiple trigger signals in at least two modules and are reserved to trigger on di-muon and cosmic muon events.

For higher trigger levels, the trigger signals of the individual modules are made available via the L2L3 bus.

D.3 Definition of Subtrigger

The new subtrigger for events with high p_T muons is defined by a combination of trigger elements at the first, the second, and the third trigger level, as listed in table D.2. The subtrigger is designed to efficiently trigger on muons with $p_T^\mu > 10\ GeV$ and is based on a combination of a high p_T track reconstructed by the FTT with a trigger signal in the instrumented iron.

In order to achieve a high trigger efficiency and a tolerable trigger rate, different p_T thresholds are required for the track reconstructed by the FTT at the first and second trigger levels. The thresholds are chosen according to the transverse momentum resolution of the FTT at $L1$ and $L2$. The resolution is best described in terms of the reciprocal of the p_T,

$$\kappa = \frac{1}{p_T}$$

which follows a Gaussian distribution. The resolutions of $\sigma_\kappa^{L_1} \approx 0.40\ GeV^{-1}$ at $L1$ and $\sigma_\kappa^{L_2} \approx 0.02\ GeV^{-1}$ at $L2$ imply that for muons with $p_T^\mu > 10\ GeV$ the inefficiency of the transverse momentum requirements is $\mathcal{O}\,(0.6\%)$ at $L1$ and is negligible at

Trigger Element	Efficiency
Track p_T ($L1$ && $L2$)	98.1%
Muon Trigger	72.8%
$zVtx\_sig$	86.1%
$zVtx\_t0$	91.9%
Track - Muon Trigger link	98.8%

Table D.3: *Efficiencies of the individual trigger elements of the new subtrigger for events with high* p_T *muons.*

$L2$. At the third trigger level, spatial matching between the extrapolated high p_T track and the trigger signal in the instrumented iron is required in order to reject combinatorial background with a high p_T non-muon track and a low p_T muon.

The rejection of non-*ep* background, which contributes significantly to the trigger rate at the first trigger level, is achieved by requiring that the event timing and the vertex position be compatible with originating from an *ep* interaction. As the reconstruction of event timing and vertex position had not yet been implemented in the simulation of the fast track trigger at the time this study was performed, the rejection of non-*ep* background by event timing and vertex position requirements is estimated using data taken with trigger elements provided by the CIP and the Veto walls only.

D.4 Estimation of Trigger Efficiency and Rate

The efficiency of the new subtrigger, defined by the conditions listed in table D.2, for the triggering of events with high p_T muons is estimated using a sample of simulated muonic W decay events passing the selection criteria listed in table 7.6. The trigger efficiency is defined as

$$\frac{\text{number of events passing selection criteria \&\& trigger conditions}}{\text{number of events passing selection criteria}}.$$

With this definition, the estimated trigger efficiency of the new subtrigger for events with high p_T muons in the phase-space defined by the selection criteria in table 7.6 is $\varepsilon \sim 60\%$.

It is instructive to consider the efficiencies of the trigger elements listed in table D.2 individually, as shown in table D.3. The estimated efficiency of the fast track trigger for the reconstruction of a track passing the p_T thresholds of $p_T^{L1} > 0.9\ GeV$ at $L1$ and $p_T^{L2} > 5.0\ GeV$ at $L2$ is about 98%. This high efficiency is due to the high track-finding efficiency of the fast track trigger (determined to be about $\approx 99\%$ for isolated high p_T tracks [176]) and its precise reconstruction of tranverse momenta. The inefficiency introduced by the linking condition between the extrapolated $L2$ track and the muon trigger signal in the instrumented iron is low too, due to the precise reconstructions of polar and azimuthal angles allowed by the FTT. The largest inefficiency is introduced by the muon trigger condition, which limits the trigger efficiency to about 70%. Also, the CIP trigger elements $zVtx$ and $t0$ introduce a sizeable inefficiency. The efficiency of the subtrigger could probably be significantly

Trigger Element	No. Events	Trigger Rate
Mu_FOEC	7393	572 Hz
Mu_Bar	9529	3335 Hz
Mu_BOEC	8872	383 Hz

Table D.4: *Number of events triggered by the trigger elements Mu_FOEC, Mu_Bar and Mu_BOEC in the special run* 341310*; the numbers quoted for the trigger rates correspond to the rates before the application of prescales.*

improved by requiring a logical "OR" of the $zVtx$ and $t0$ conditions of the CIP with the FTT, once the reconstruction of event timing and vertex position is implemented in the fast track trigger. In this case, the trigger efficiency of the subtrigger may get close to the limit $\varepsilon_{optimistic} \approx 70\%$ set by the muon trigger condition.

The expected trigger rate for the new subtrigger is estimated by analysing the events recorded in a special run (run number 341310) taken in February 2003. In this run, events that triggered any muon trigger elements (and passed the prescales) were written to tape without any further conditions (L2 and L4 transparent, no global options). They therefore represent an unbiased sample of events triggered by the muon trigger elements. In total, 25794 events were recorded for the trigger elements Mu_FOEC, Mu_Bar and Mu_BOEC. The number of events triggered by each of these trigger elements and the corresponding trigger rates are summarized in table D.4. Of these 25794 events, 22 pass the conditions listed in table D.2 at the first trigger level, 4 pass the conditions at both of the first two trigger levels, and 1 passes the conditions at all three trigger levels. (As the fast track trigger had not yet been installed at the time the special run was taken, its response is simulated using FTTEMU for all 25794 recorded events). The event numbers of the 22 events that pass at least the conditions at the first trigger level are listed in table D.5. The majority of these are beam-gas background events (the quality of the vacuum in the beam-pipe was bad in February 2003) with many soft tracks, that are rejected by the increased p_T thresholds at the second trigger level. The expected trigger rates at the different trigger levels are determined by weighting each selected event with the prescale factor of the muon trigger element that triggered the event and normalizing to the dead-time corrected run-time ($t_{run} \approx 1621\ s$ for run 341310),

$$f = \frac{1}{t_{run}} \sum{}^{\text{selected}}_{\text{events}} \text{prescale}\,(\text{event})\,. \tag{D.1}$$

The resulting trigger rates expected at the different trigger levels are

$$f^{\mu}_{L_1} \approx 3.7\ Hz, \qquad f^{\mu}_{L_2} \approx 0.5\ Hz, \qquad \text{and } f^{\mu}_{L_3} \approx 0.1\ Hz$$

for conditions as in run 341310.

In order to obtain an estimate of the trigger rate expected in nominal HERA II running conditions, the rates estimated for run 341310 need to be extrapolated to higher beam currents. In comparison to the HERA II design values of $I_e^{nom} = 58\ mA$ and $I_p^{nom} = 140\ mA$ [12], the beam currents in run 341310, $I_e^{\mu} = 6.9\ mA$ and $I_p^{\mu} = 38.3\ mA$, were significantly lower. In the extrapolation of the trigger rates estimated for run 341310 to higher beam currents, the contributions of cosmic muon

Event Number	Muon Trigger Element	passed Trigger Level L1	L2	L3	Event Type
2438	*Mu_BOEC*	×	×		beam-gas
4187	*Mu_Bar*	×			cosmic
6269	*Mu_BOEC*	×			beam-gas
6435	*Mu_Bar*	×			cosmic
11119	*Mu_FOEC*	×			beam-gas
14868	*Mu_BOEC*	×			*ep* (probably noisy)
15389	*Mu_Bar*	×			cosmic
17473	*Mu_FOEC*	×	×		beam-gas
19970	*Mu_BOEC*	×			beam-gas
20145	*Mu_Bar*	×			cosmic
20436	*Mu_Bar*	×			beam-gas
27031	*Mu_FOEC*	×			beam-gas
27607	*Mu_FOEC*	×			*ep*
32453	*Mu_FOEC*	×			beam-gas
33111	*Mu_Bar*	×			cosmic
34689	*Mu_BOEC*	×			cosmic
38189	*Mu_FOEC*	×			beam-gas + beam-halo overlay
39405	*Mu_FOEC*	×	×	×	beam-gas
39945	*Mu_BOEC*	×			beam-halo
40411	*Mu_BOEC*	×			beam-gas
43507	*Mu_Bar*	×	×		cosmic
46579	*Mu_Bar*	×			*ep* (+ beam-gas overlay)

Table D.5: *Event numbers of the* 22 *events in run* 341310 *that pass at least the conditions of the new subtrigger for events with high p_T muons at the first trigger level; the event type has been determined by visual scanning.*

Coefficient	Trigger Level		
	L1	L2	L3
f_{cosmic} $[Hz]$	2.2	0.4	0
$f_{beam-gas/beam-halo}$ $[10^{-2} HzmA^{-1}]$	2.3	0.5	0.2
f_{ep} $[10^{-3} HzmA^{-2}]$	1.9	0	0

Table D.6: *Coefficients for the parametrisation of the rate of cosmic muon, beam-gas/beam-halo and ep events as a function of the beam currents.*

and beam-gas/beam-halo events have to be scaled differently than does the rate of ep interactions that scales with luminosity, i.e. is proportional to $I_e \cdot I_p$. As the beam-gas/beam-halo background is mainly induced by the proton beam, its rate is approximately proportional to I_p, while the rate of cosmic muon events is independent of the beam currents. Accordingly, the trigger rate is parametrized by

$$f\left(I_e, I_p\right) = f_{cosmic} + I_p \cdot f_{beam-gas/beam-halo} + I_e I_p \cdot f_{ep}. \qquad \text{(D.2)}$$

The coefficients f_{cosmic}, $f_{beam-gas/beam-halo}$ and f_{ep} are determined individually for the first, the second, and the third trigger level from the rates of cosmic muon, beam-gas/beam-halo and ep events that pass the conditions at the corresponding trigger level in run 341310. The coefficients determined by converting the numbers of cosmic muon, beam-gas/beam-halo and ep events listed in table D.5 into rates according to equation D.1 are shown in table D.6. For the nominal HERA II beam currents, the estimates resulting from the parametrisation in equation D.2 and the coefficients shown in table D.6 for the trigger rates expected at the different trigger levels are

$$f_{L_1}^{nom} \approx 21\ Hz, \qquad f_{L_2}^{nom} \approx 1.1\ Hz, \qquad \text{and } f_{L_3}^{nom} \approx 0.3\ Hz.$$

In contrast to the trigger rates estimated for run 341310, the trigger rates expected at nominal HERA II beam currents are expected to be dominated by the contribution from ep events.

Since February 2003, the quality of the vacuum in the beam-pipe adjacent to H1 has been significantly improved by the installation of new vacuum pumps in the interaction region of the H1 experiment and as a result of the lower number of gas molecules in the beam-pipe, the rate of beam-gas background has decreased. The reduction in the rate is difficult to quantify, however. Conservatively, the improvement in the vacuum quality is not taken into account in the estimation of the expected trigger rate.

D.5 Conclusion

A new subtrigger for high p_T muons has been developed, based on the recently installed fast track trigger system. The efficiency of the new subtrigger has been estimated using an event sample of simulated muonic W decays. For events with muons with $p_T^\mu > 10\ GeV$ in the geometrical acceptance of the FTT ($25° \lesssim \theta^\mu \lesssim 155°$), the trigger efficiency is estimated to be $\varepsilon \approx 60\%$. The trigger efficiency may possibly be improved to $\varepsilon_{optimistic} \approx 70\%$, once the reconstruction of the event vertex and timing is implemented in the fast track trigger, and the $zVtx$ and $t0$ trigger elements of the FTT are used in disjunction with those of the CIP. The trigger rate of the new subtrigger has been estimated using data recorded in a special run taken in February 2003. For nominal HERA II running conditions, the trigger rate expected at the first, the second, and the third trigger level is estimated to be

$$f_{L_1}^{nom} \approx 21\ Hz, \qquad f_{L_2}^{nom} \approx 1.1\ Hz, \qquad \text{and } f_{L_3}^{nom} \approx 0.3\ Hz.$$

These rates will allow the new subtrigger to be operated without prescales.

Appendix E

Development of a new Beam-Gas Finder

The new finder algorithm identifies beam-gas background by its characteristic isotropic energy distribution. The isotropy of the energy distribution is tested by dividing the central tracking detector and the LAr calorimeter into 16 sectors in azimuth and summing the number of non-vertex fitted tracks in the CTD and energy deposits in the CB and FB wheels of the LAr in each sector. Events are flagged as beam-gas background by the new finder, if tracks are reconstructed in at least 12 sectors of the CTD and the sum of the energy deposits in the LAr exceeds 1 GeV for at least 12 sectors of the CB and FB wheels. Additionally, the energy deposits within a radius of 60 *cm* around the beam-pipe in the IF/OF wheels of the LAr are summed and required to exceed 100 GeV, in order to reduce the inefficiency for events due to *ep* interactions [1]. In table E.1, the conditions for events to be flagged by the new beam-gas finder are summarized.

Quantity	Requirement
$E_{LAr}^{IF/OF}$	$> 100\ GeV$
N_{ϕ}^{CTD1}	≥ 12
$N_{\phi}^{CB/FB2}$	≥ 12

[1] with at least one track in the CTD
[2] with energy deposits of at least 1 GeV in the CB or FB wheels of the LAr

Table E.1: *Selection criteria defining the new beam-gas finder.*

An example of an event exclusively rejected by the new finder algorithm for beam-gas background is shown in figure E.1.

[1]In the nuclear interactions between the beam protons and the gas atoms in the beam-pipe high transverse momenta are often exchanged. Due to the Lorentz boost in the proton direction, the particles produced in beam-gas interactions are typically detected at small polar angles, however.

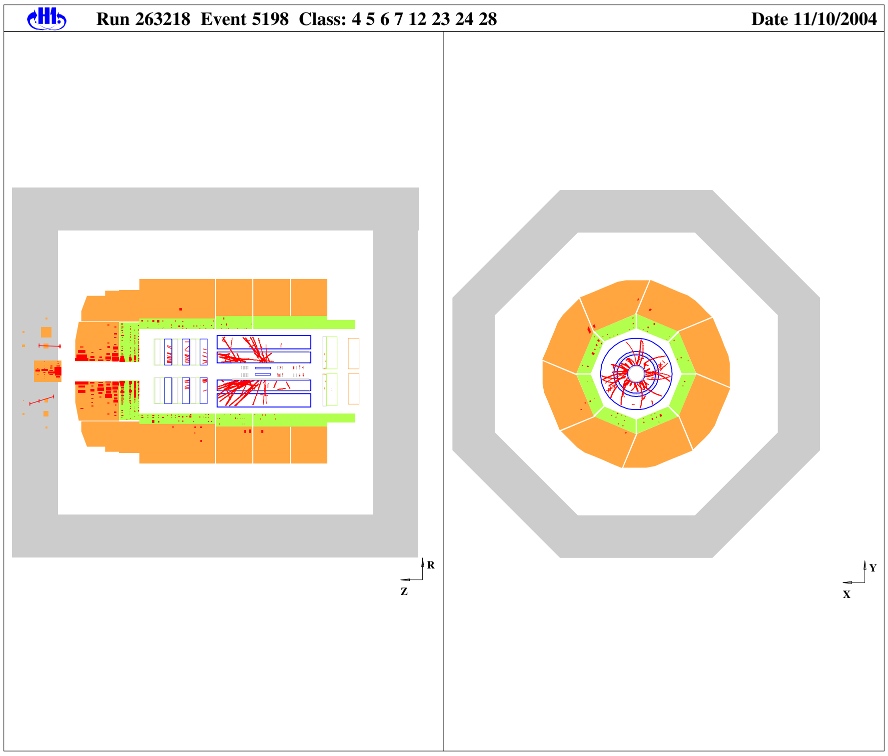


Figure E.1: *Example of background event rejected exclusively by the new finder algorithm for beam-gas background.*

Appendix F

Distributions of Observables in the Background enriched Control Samples

In this section, the distributions of the main kinematic and topological quantities are shown for the events selected in the background enriched samples defined by the selection criteria in table 8.2 in section 8.1.2.

F.1 Events with Isolated Electrons

NC Enriched Control Sample

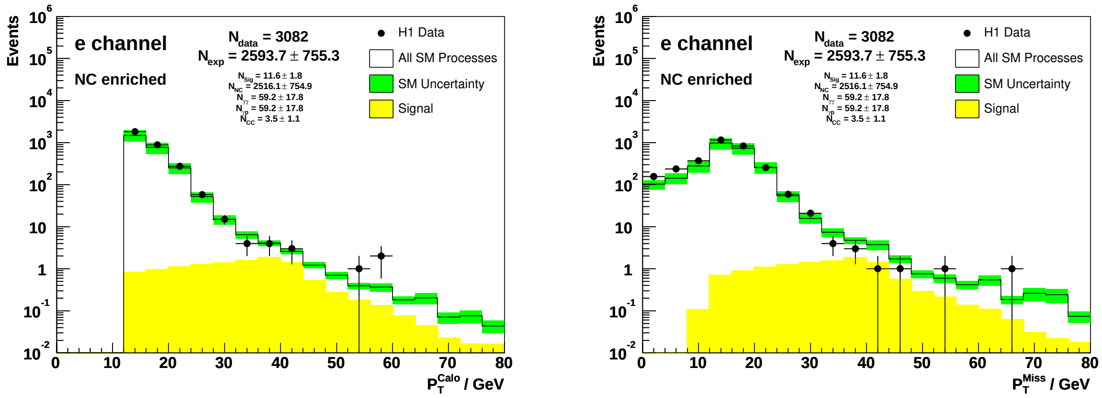


Figure F.1: *Distributions of the imbalance of energy deposits in the calorimeter (a) and the reconstructed missing transverse momentum (b) in events selected in the NC enriched phase-space.*

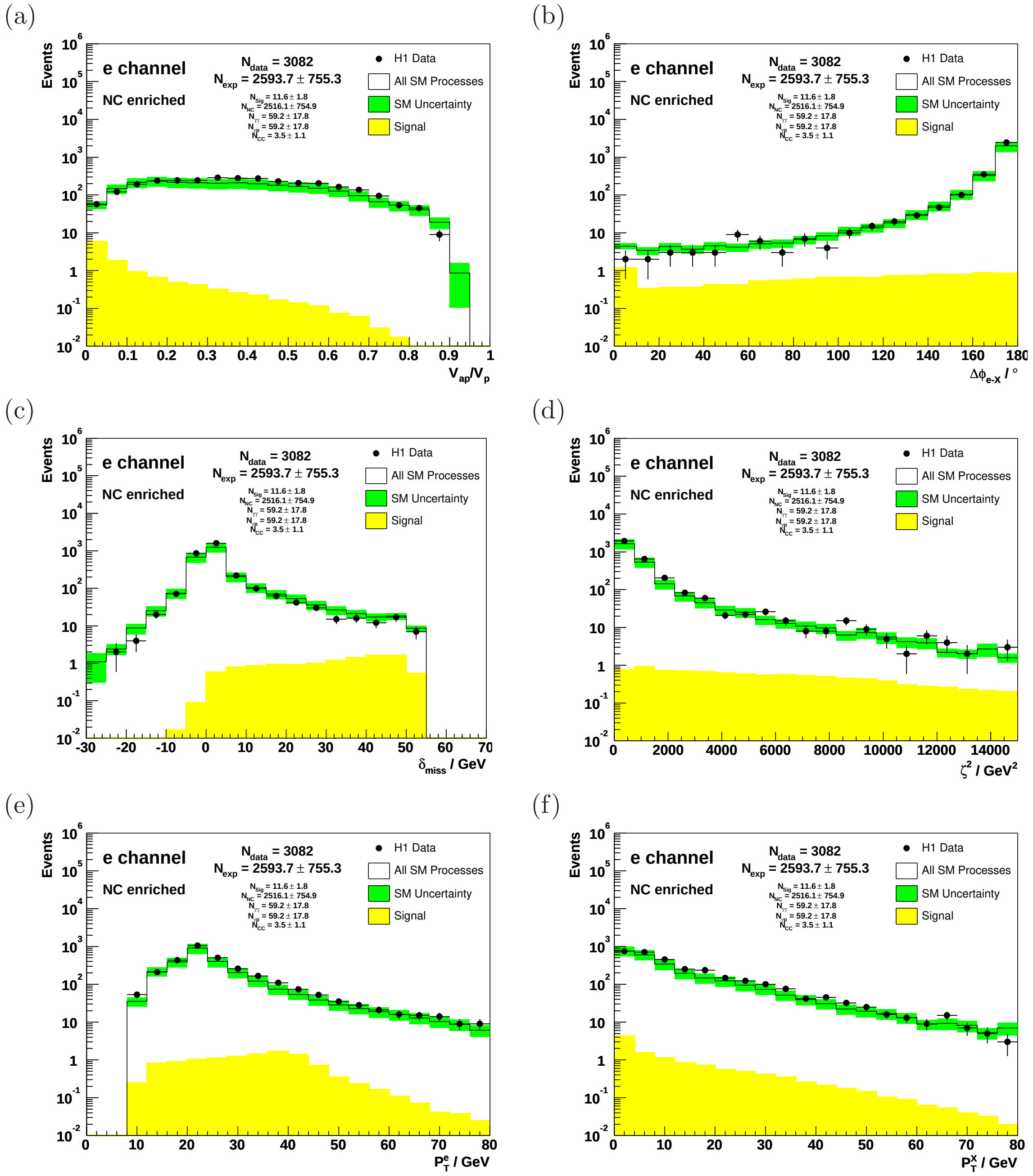


Figure F.2: *Distributions of the ratio V_{ap}/V_p (a), the acoplanarity angle between the electron candidate and the hadronic final state (b), the missing longitudinal momentum (c), the quantity ζ^2 (d), the transverse momentum of the electron candidate (e), and the transverse momentum of the hadronic final state (f) in events selected in the NC enriched phase-space.*

CC Enriched Control Sample

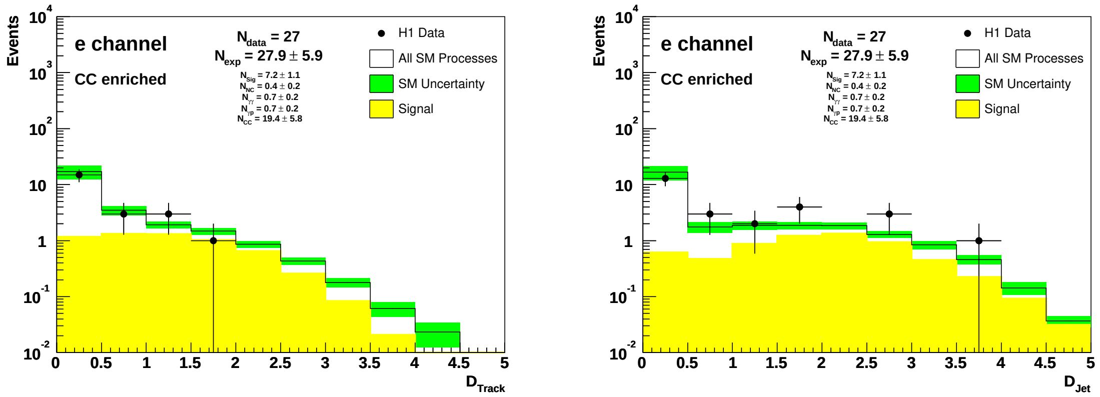


Figure F.3: *Distributions of the distance to the nearest track (left) and the distance to the nearest jet (right) of the electron candidates in events selected in the CC enriched phase-space.*

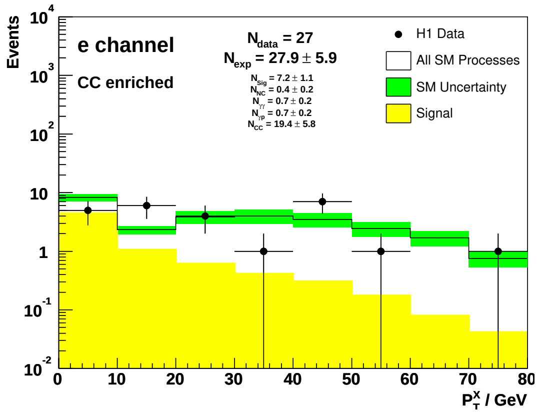


Figure F.4: *Distribution of the transverse momentum of the hadronic system in events selected in the CC enriched phase-space.*

F.2 Events with Isolated Muons

F.2.1 LAr Triggered Phase-Space

NC Enriched Control Sample

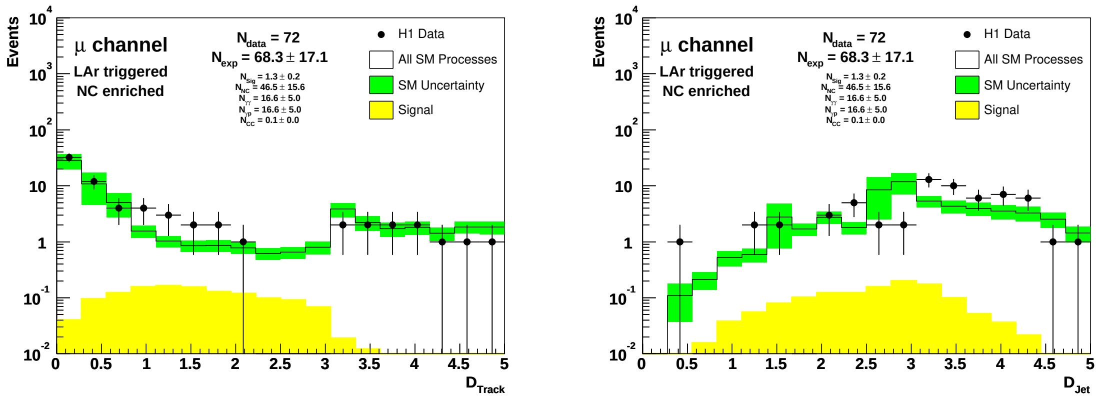


Figure F.5: *Distributions of the distance to the nearest track (left) and the distance to the nearest jet (right) of the muon candidates in events selected in the NC enriched LAr triggered phase-space.*

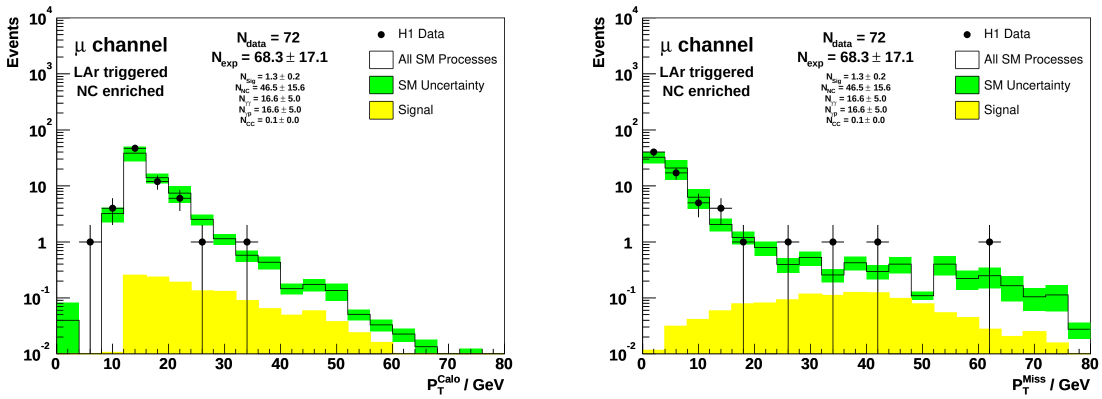


Figure F.6: *Distributions of the imbalance of energy deposits in the calorimeter (left) and the reconstructed missing transverse momentum (right) in events selected in the NC enriched LAr triggered phase-space.*

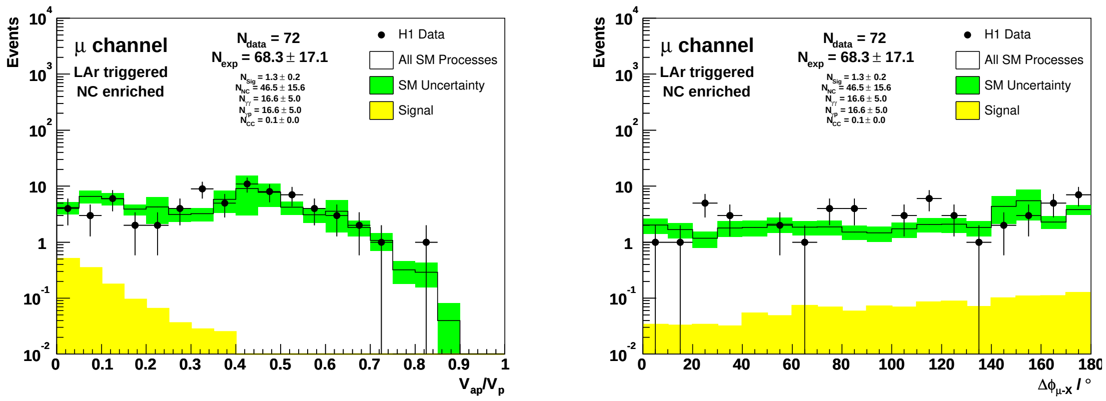


Figure F.7: *Distributions of the ratio V_{ap}/V_p (left) and the acoplanarity angle between the muon candidate and the hadronic final state (right) in events selected in the NC enriched LAr triggered phase-space.*

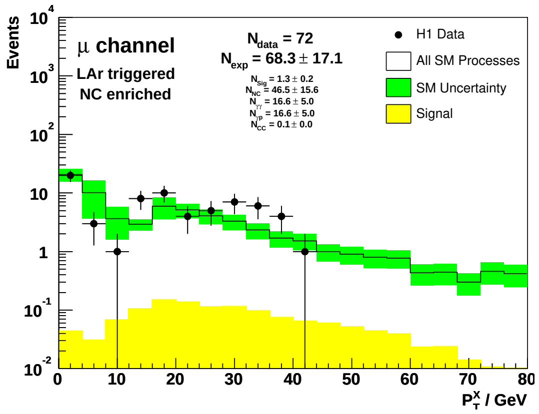


Figure F.8: *Distribution of the transverse momentum of the hadronic system in events selected in the NC enriched LAr triggered phase-space.*

Lepton Pair Enriched Control Sample

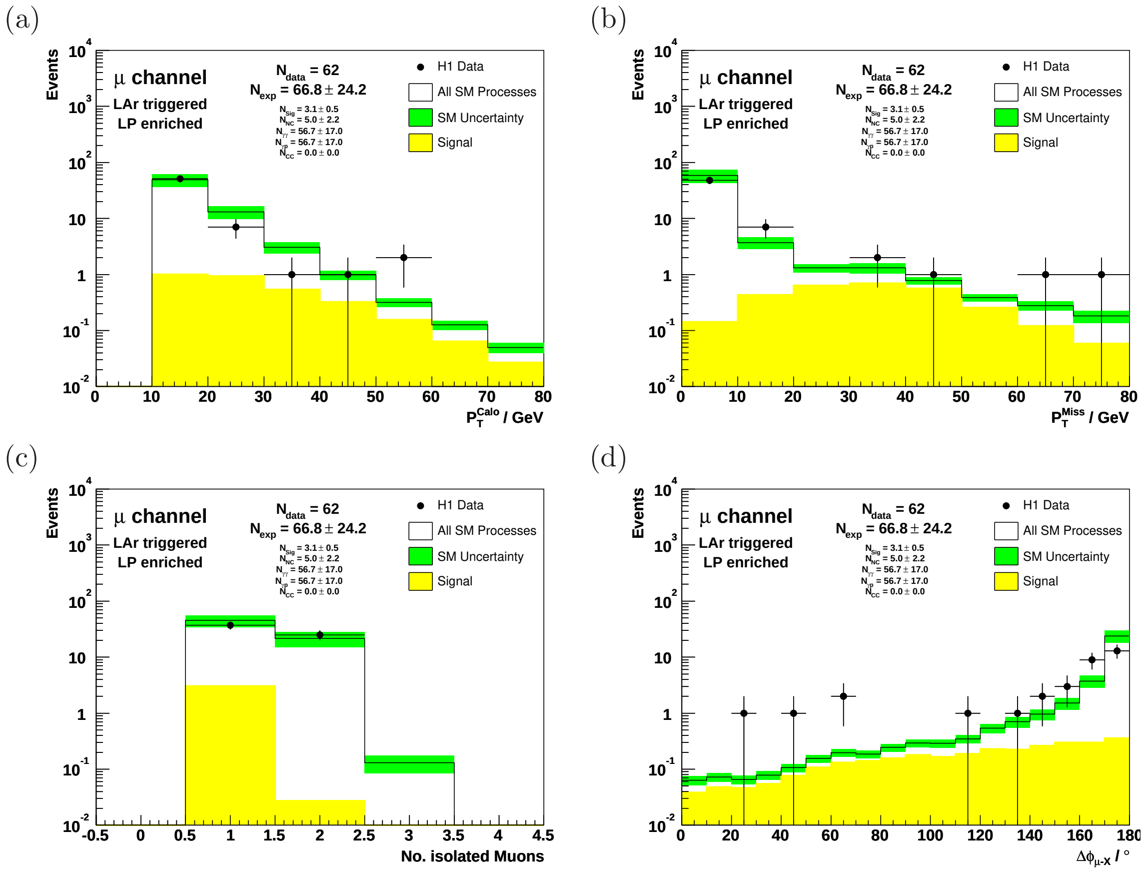


Figure F.9: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the number of isolated muons (c) and the acoplanarity angle between the electron candidate and the hadronic final state (d) in events selected in the $\ell^+\ell^-$ enriched LAr triggered phase-space.*

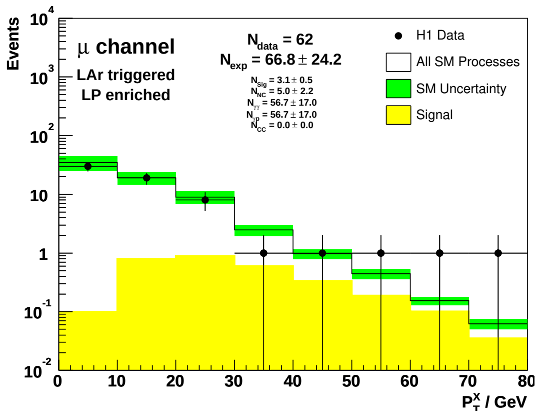


Figure F.10: *Distribution of the transverse momentum of the hadronic system in events selected in the $\ell^+\ell^-$ enriched LAr triggered phase-space.*

Photoproduction Enriched Control Sample

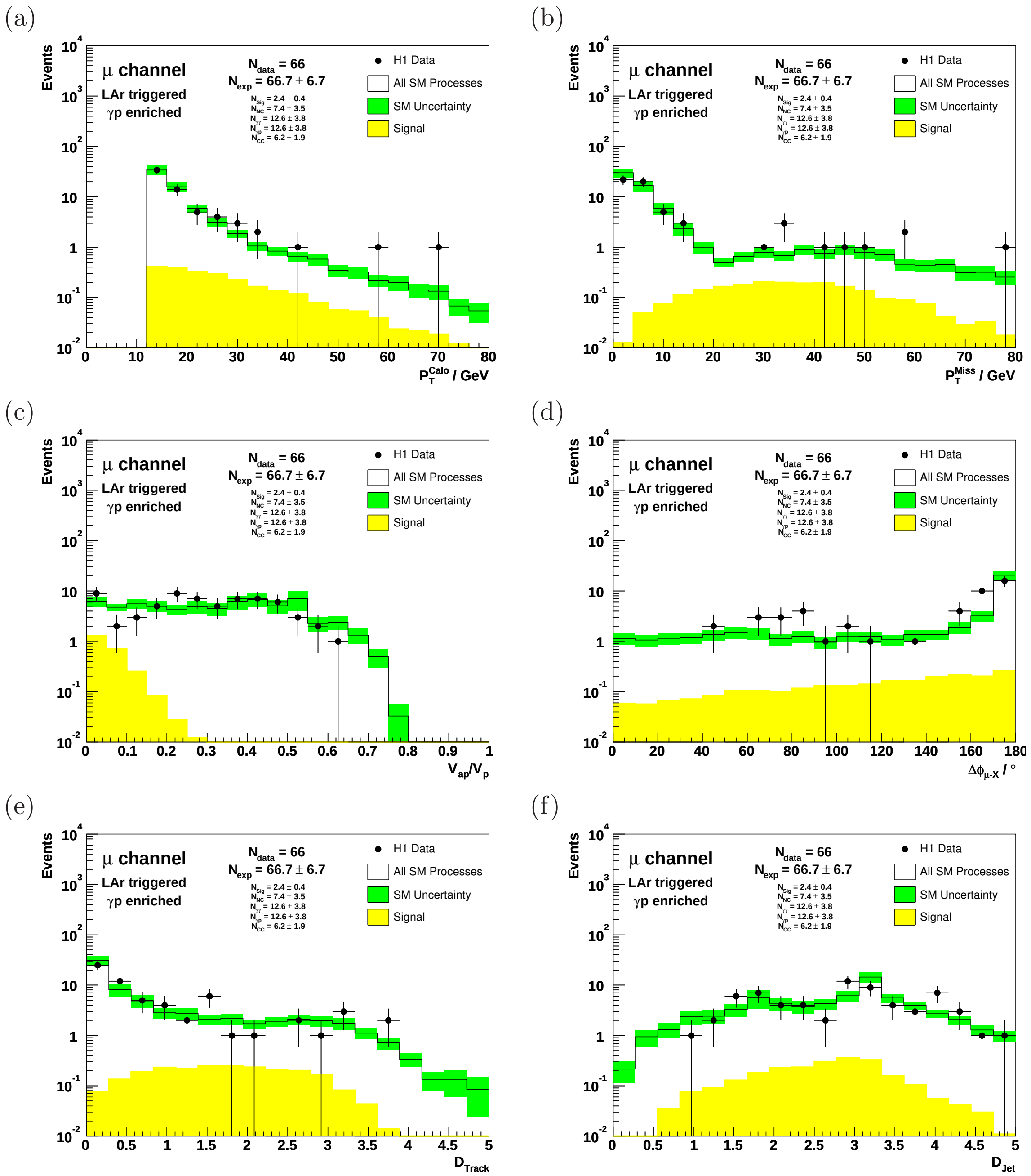


Figure F.11: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the ratio V_{ap}/V_p (c), the acoplanarity angle between the muon candidate and the hadronic final state (d), the distance to the nearest track (e) and the distance to the nearest jet (f) of the muon candidates in events selected in the γp enriched LAr triggered phase-space.*

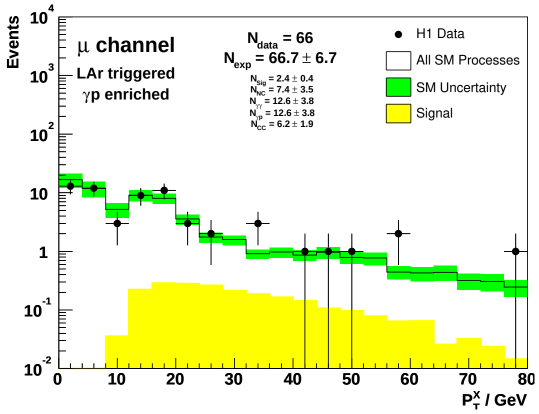


Figure F.12: *Distribution of the transverse momentum of the hadronic system in events selected in the γp enriched LAr triggered phase-space.*

F.2.2 Muon Triggered Phase-Space

NC Enriched Control Sample

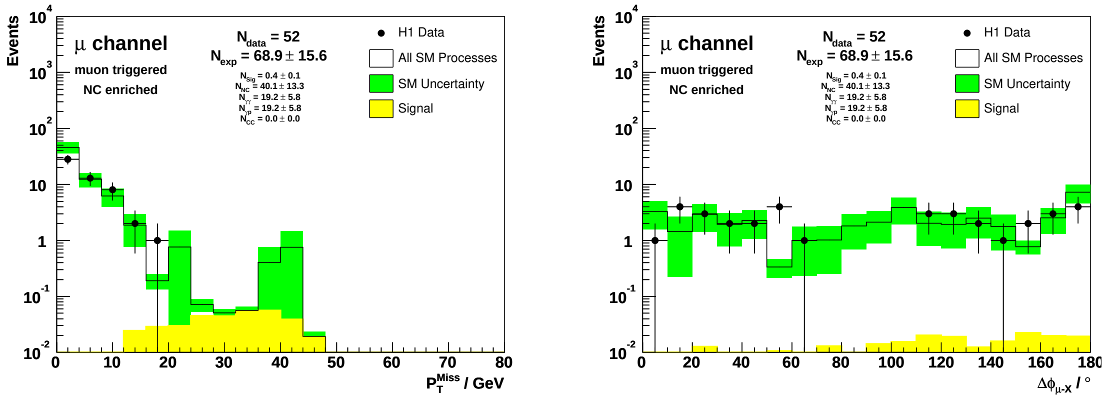


Figure F.13: *Distributions of the reconstructed missing transverse momentum (left) and the acoplanarity angle between the muon candidate and the hadronic final state (right) in events selected in the NC enriched muon triggered phase-space.*

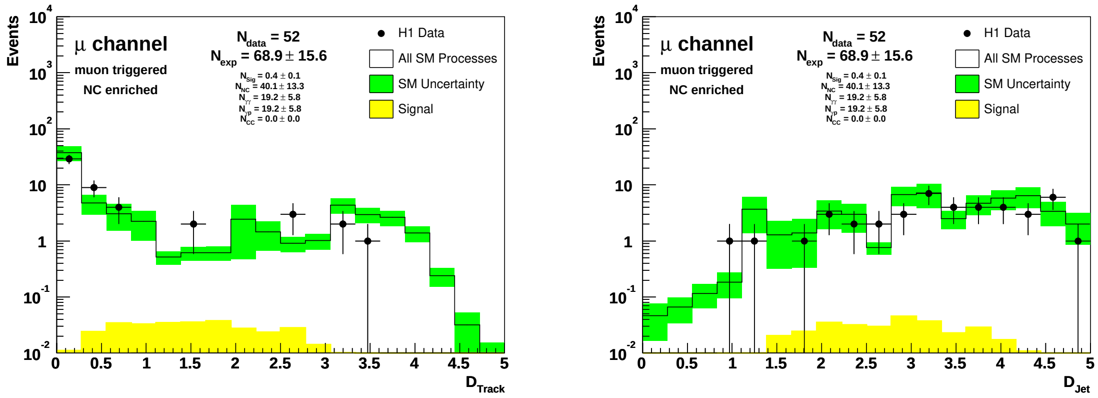


Figure F.14: *Distributions of the distance to the nearest track (left) and the distance to the nearest jet (right) of the muon candidates in events selected in the NC enriched muon triggered phase-space.*

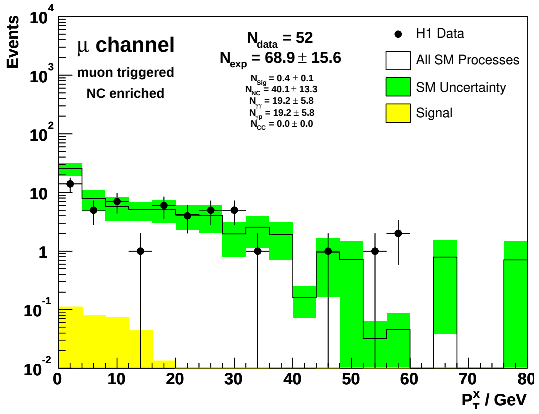


Figure F.15: *Distribution of the transverse momentum of the hadronic system in events selected in the NC enriched muon triggered phase-space.*

Lepton Pair Enriched Control Sample

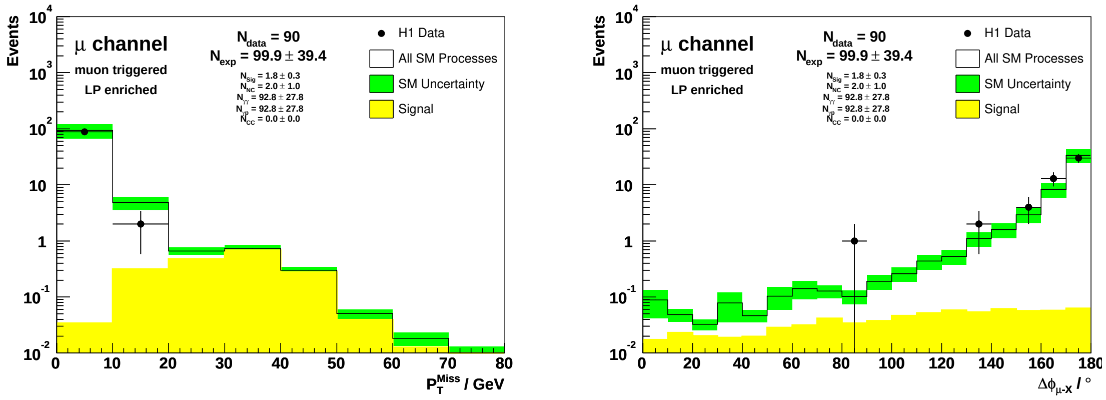


Figure F.16: *Distributions of the the reconstructed missing transverse momentum (left) and the acoplanarity angle between the electron candidate and the hadronic final state (right) in events selected in the $\ell^+\ell^-$ enriched phase-space.*

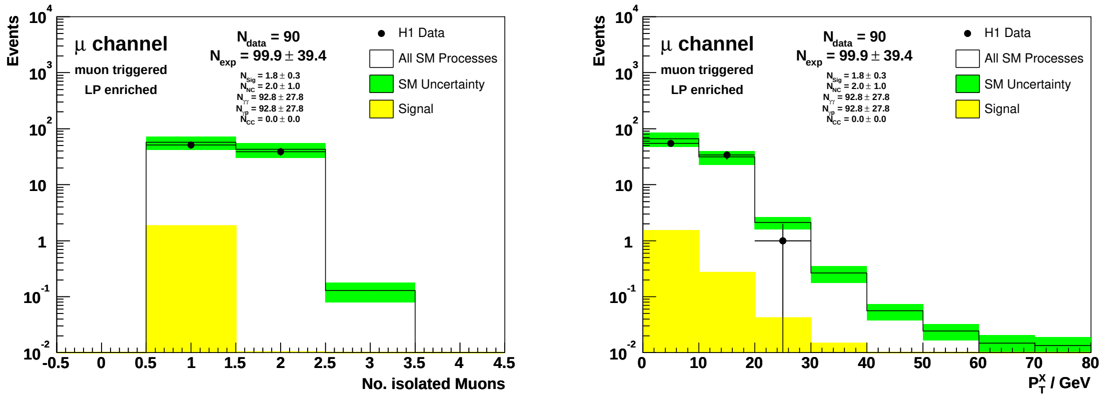


Figure F.17: *Distribution of the number of isolated muons (left) and the transverse momentum of the hadronic system (right) in events selected in the $\ell^+\ell^-$ enriched phase-space.*

Photoproduction Enriched Control Sample

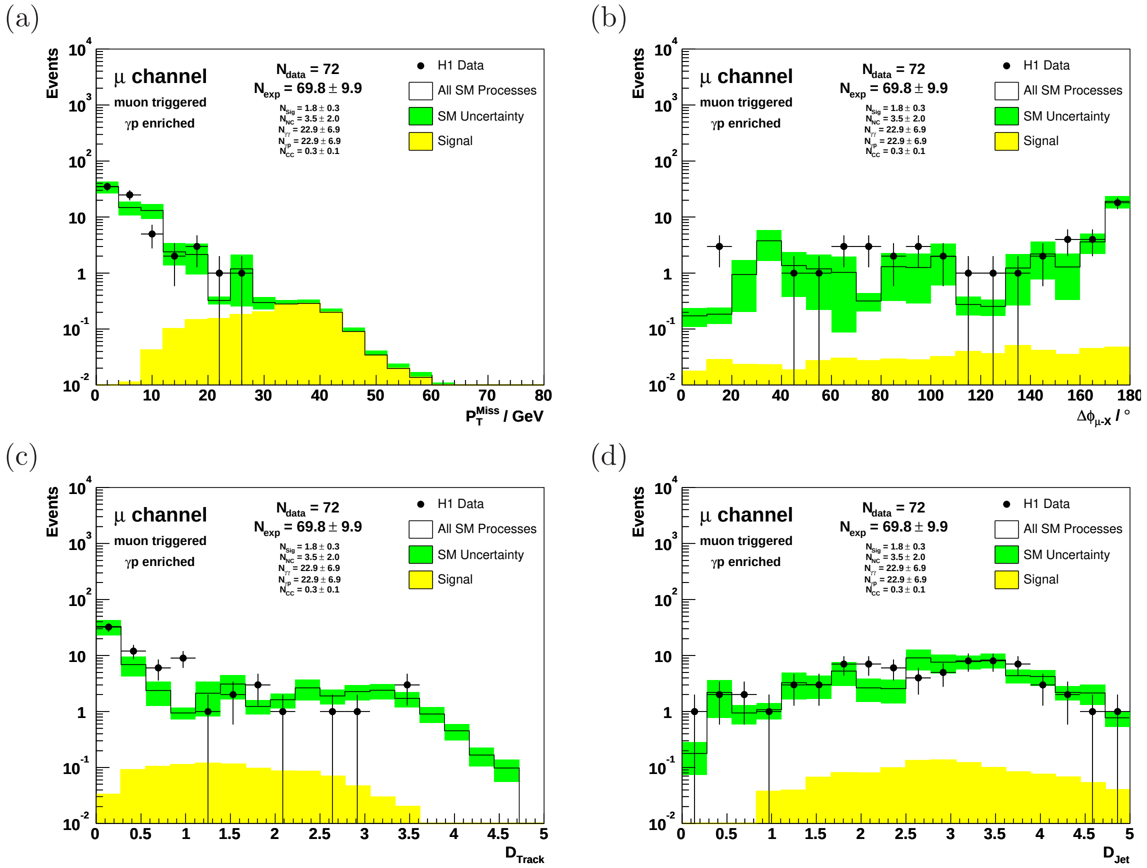


Figure F.18: *Distributions of the reconstructed missing transverse momentum (a), the acoplanarity angle between the muon candidate and the hadronic final state (b), the distance to the nearest track (c) and the distance to the nearest jet (d) of the muon candidates in events selected in the γp enriched muon triggered phase-space.*

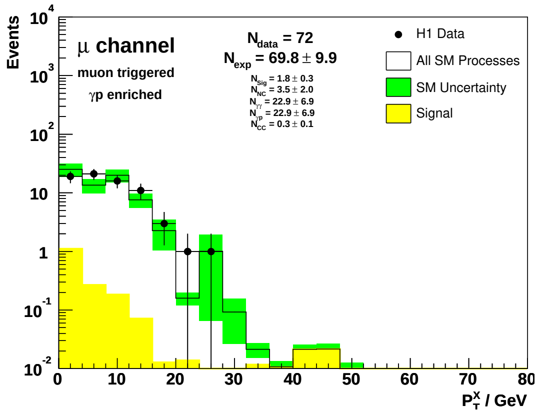


Figure F.19: *Distribution of the transverse momentum of the hadronic system in events selected in the γp enriched phase-space.*

F.3 Events with Isolated Tau Leptons

NC Enriched Control Sample

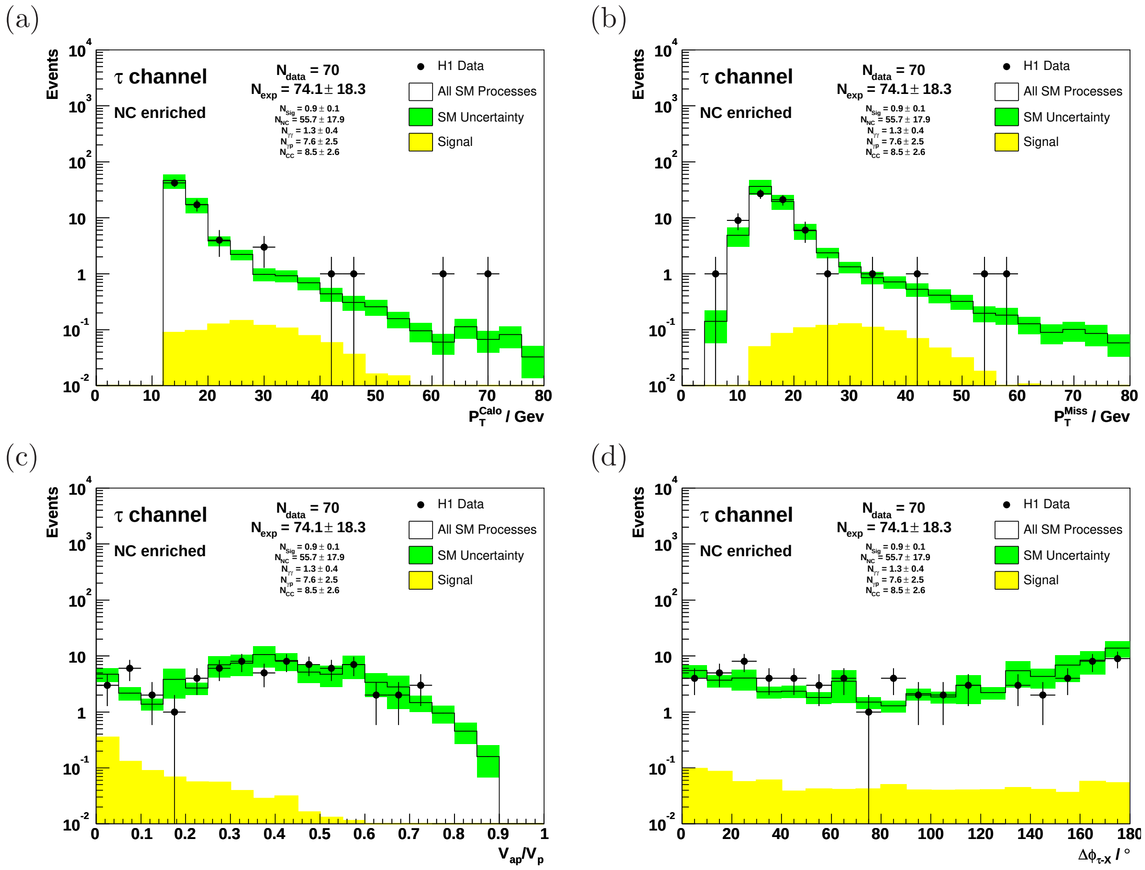


Figure F.20: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the ratio V_{ap}/V_p (c) and the acoplanarity angle between the electron candidate and the hadronic final state (d) in events with tau leptons selected in the NC enriched phase-space.*

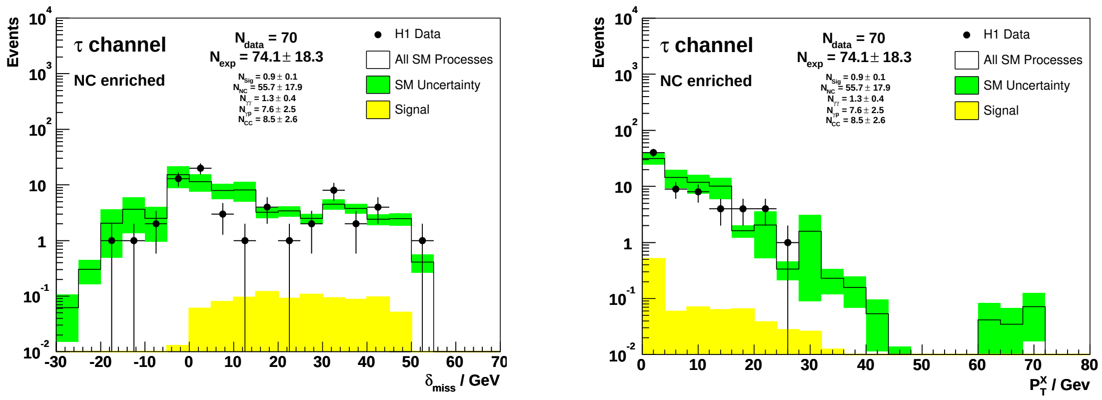


Figure F.21: *Distributions of the the missing longitudinal momentum (left) and the transverse momentum of the hadronic final state (right) in events with tau leptons selected in the NC enriched phase-space.*

CC Enriched Control Sample

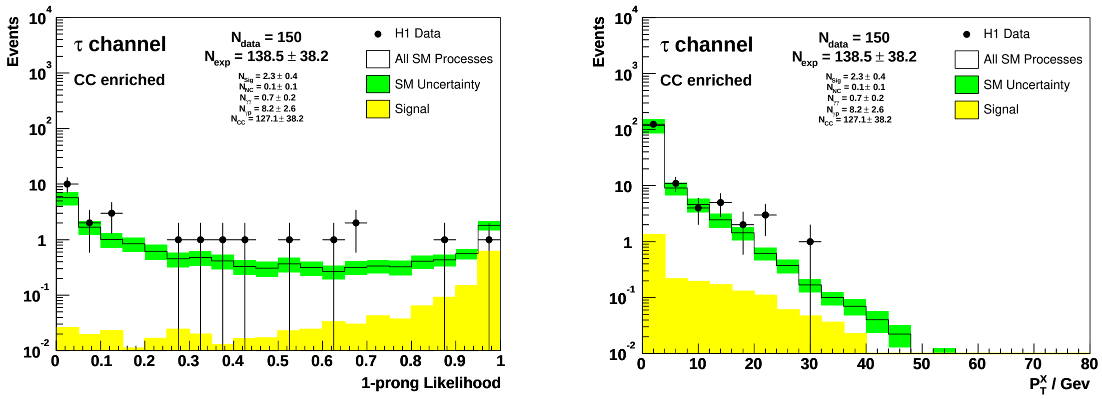


Figure F.22: *Distributions of the classification output of the neural networks described in section 4.5.5 for "one-prong" tau jet candidates (left) and of the hadronic transverse momentum (right) in events selected in the CC enriched phase-space.*

Photoproduction Enriched Control Sample

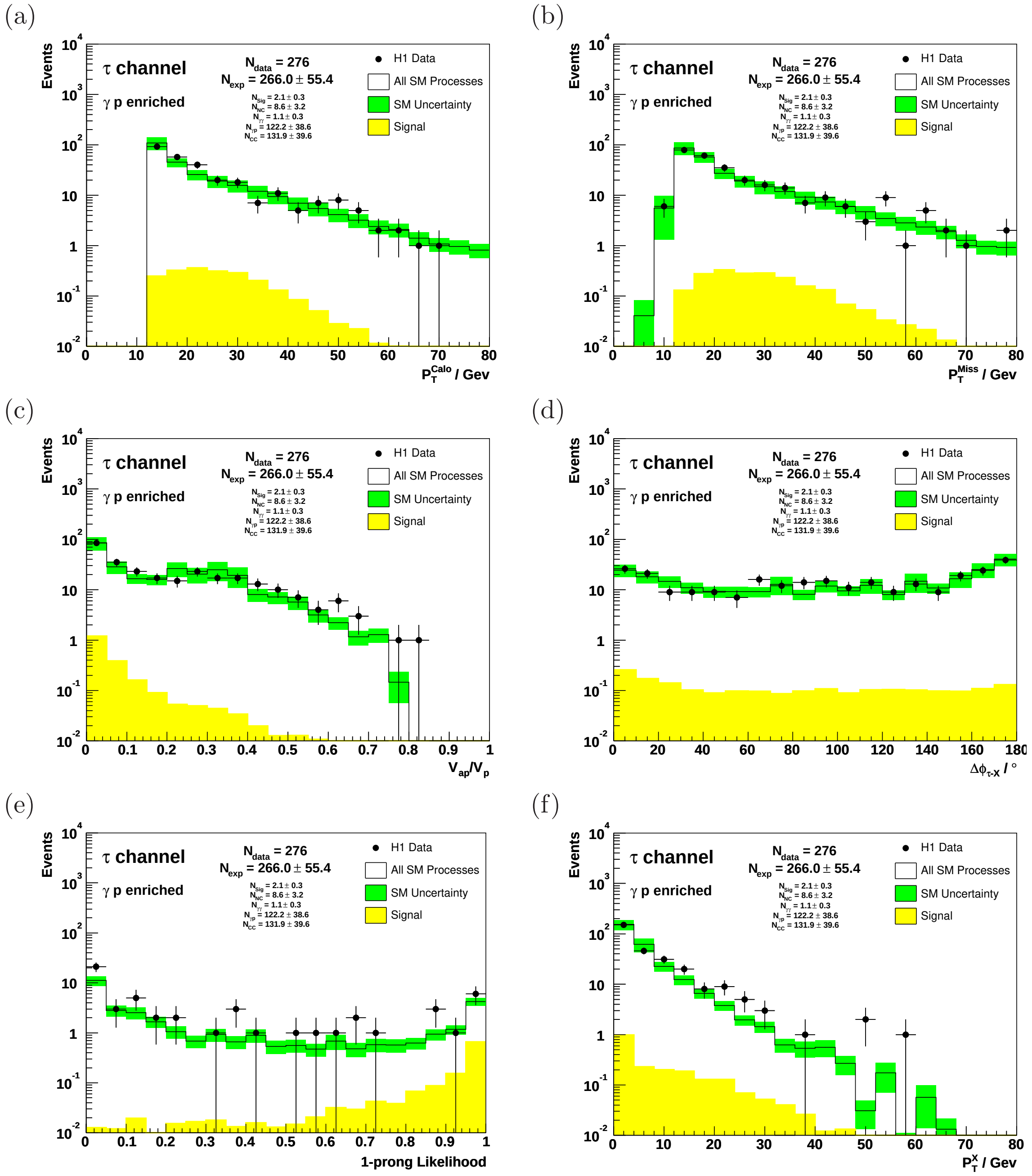


Figure F.23: *Distributions of the imbalance of energy deposits in the calorimeter (a), the reconstructed missing transverse momentum (b), the ratio V_{ap}/V_p (c), the acoplanarity angle between the electron candidate and the hadronic final state (d), the classification output of the neural networks described in section 4.5.5 for "one-prong" tau jet candidates (d) and the transverse momentum of the hadronic final state (e) in events selected in the γp enriched phase-space.*

Appendix G

Event Displays

In this section, event displays of the 21 events with isolated electrons, muons or tau leptons and large missing transverse momentum selected in the 1994-2000 e^+p and e^-p datasets (*cf.* section 8.6) are shown. The events are shown in chronological order.

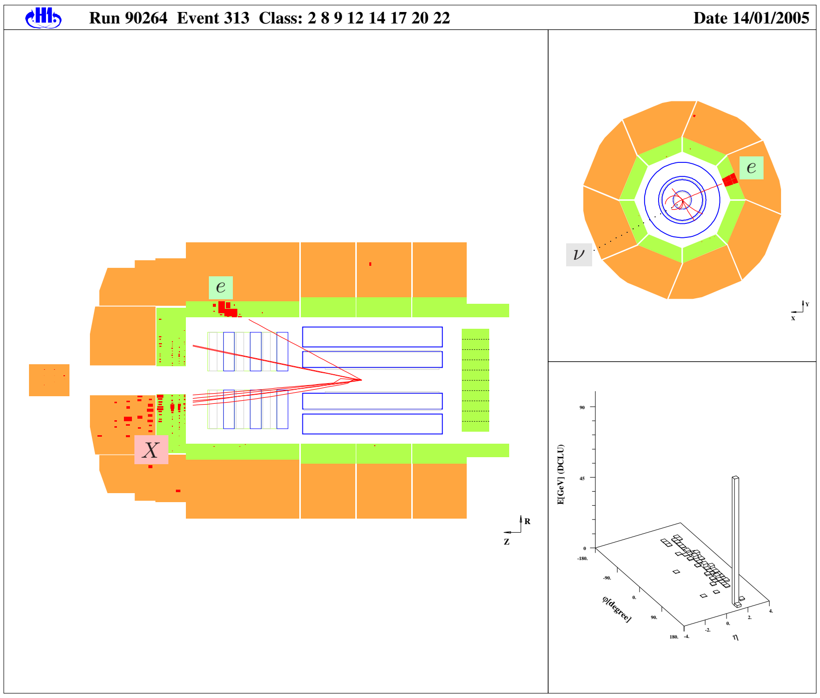


Figure G.1: *Display of event* 313 *in run* 90241*: an event selected in the electron channel in the 1994-97* e^+p *dataset.*

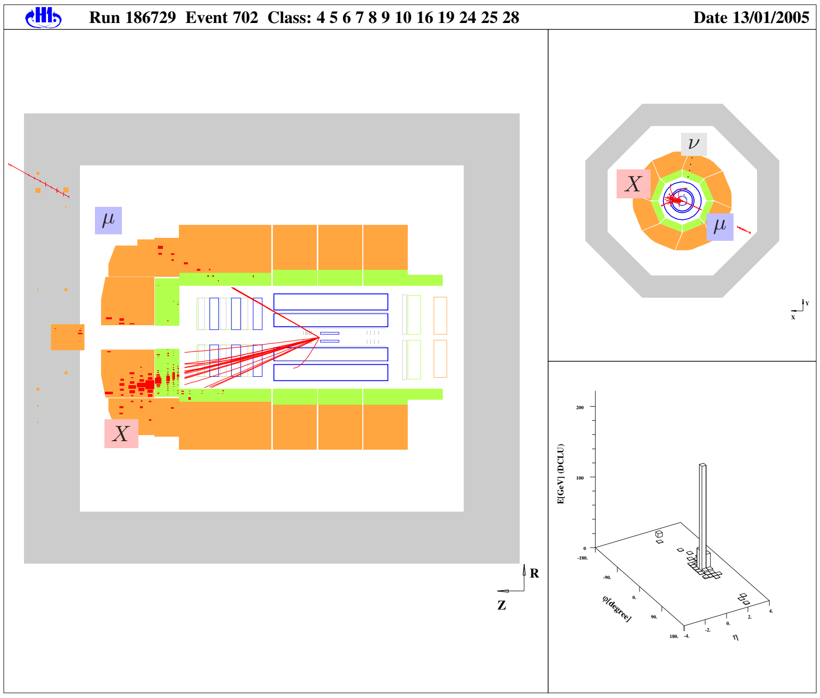


Figure G.2: *Display of event* 702 *in run* 186729*: an event selected in the muon channel in the 1994-97* e^+p *dataset.*

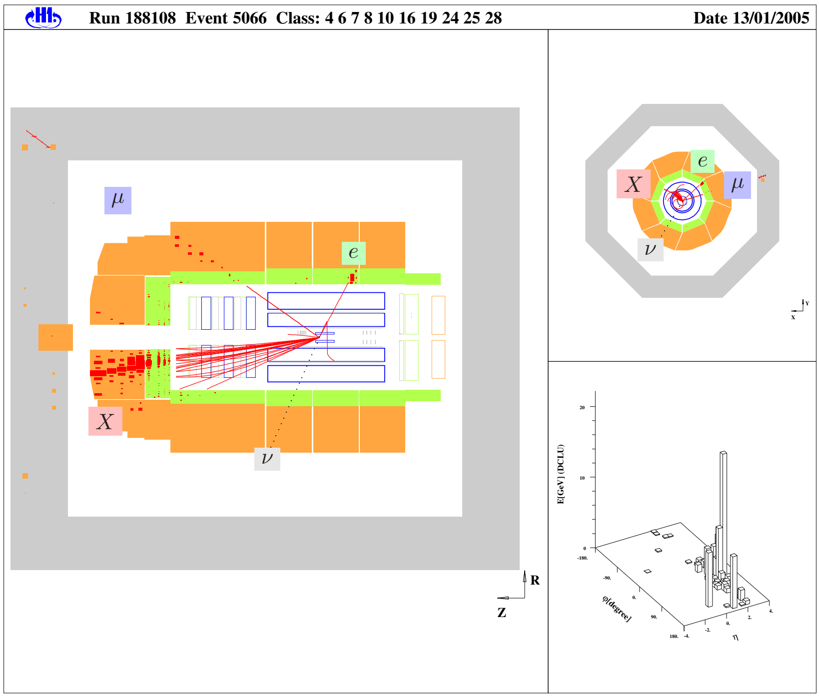


Figure G.3: *Display of event* 5066 *in run* 188108*: an event selected in the muon channel in the 1994-97* e^+p *dataset. In this event, the scattered beam electron is detected in the LAr calorimeter.*

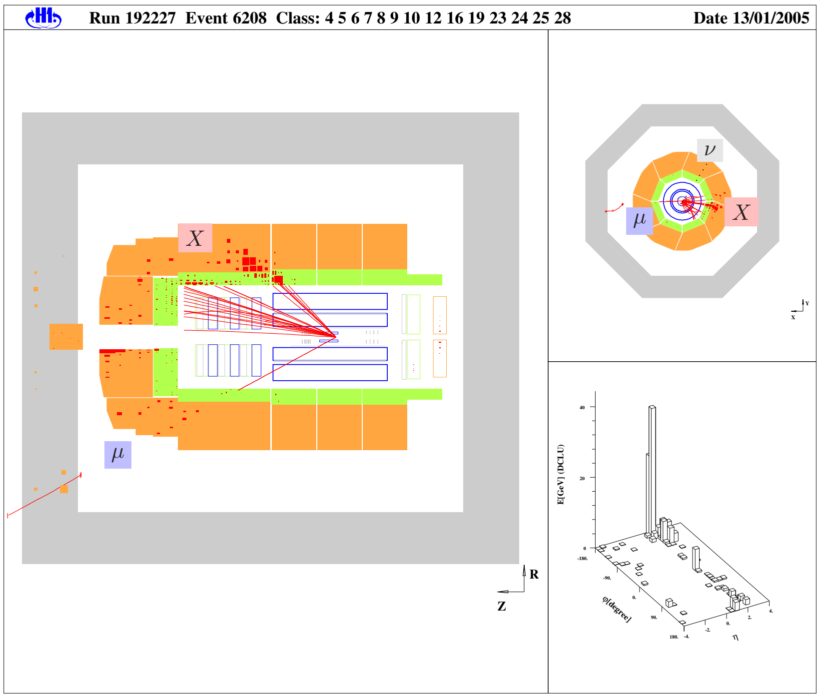


Figure G.4: *Display of event* 6208 *in run* 192227*: an event selected in the muon channel in the 1994-97* e^+p *dataset.*

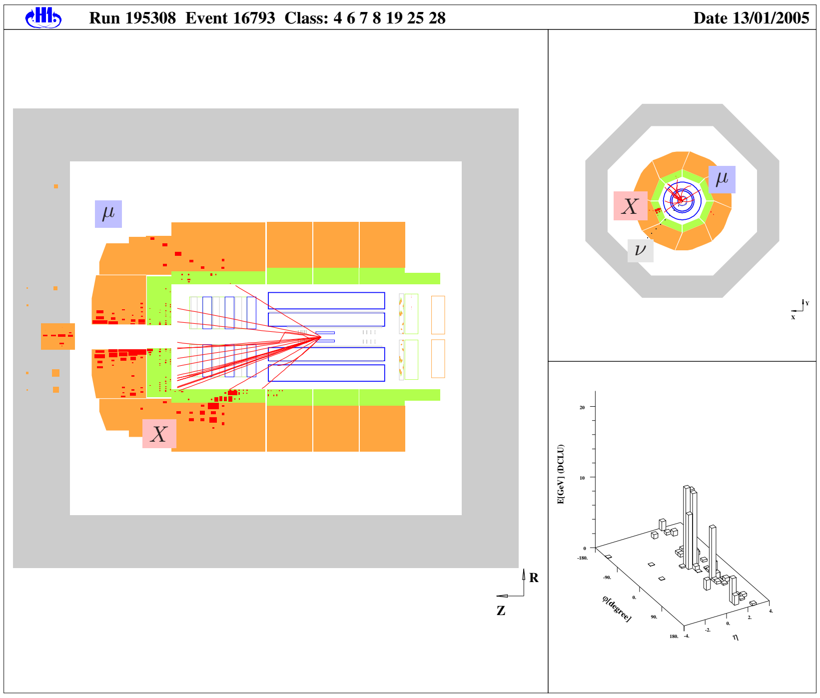


Figure G.5: *Display of event* 16793 *in run* 195308*: an event selected in the muon channel in the 1994-97* e^+p *dataset.*

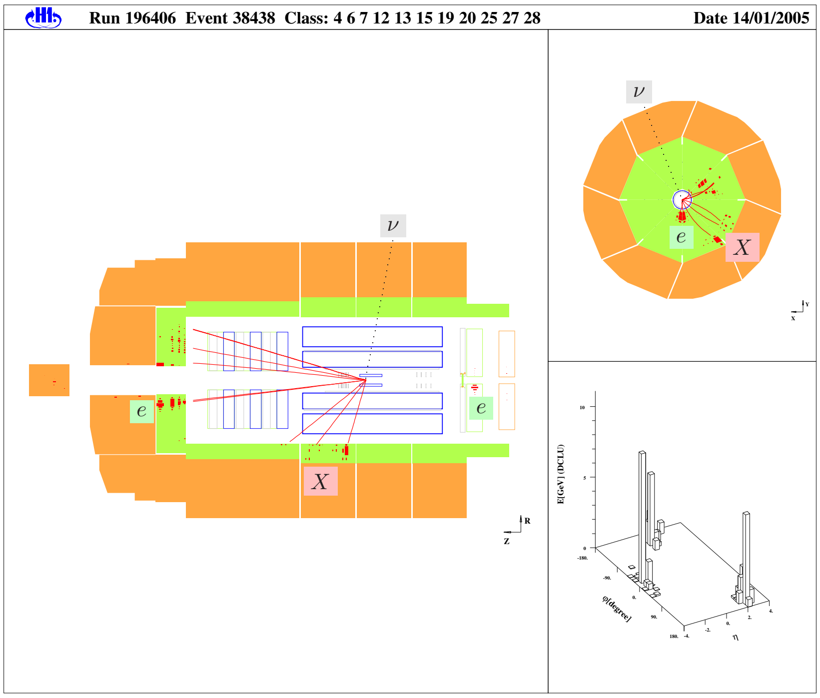


Figure G.6: *Display of event* 38438 *in run* 196406*: an event selected in the electron channel in the 1994-97* e^+p *dataset. In this event, the scattered beam electron is detected in the SpaCal calorimeter.*

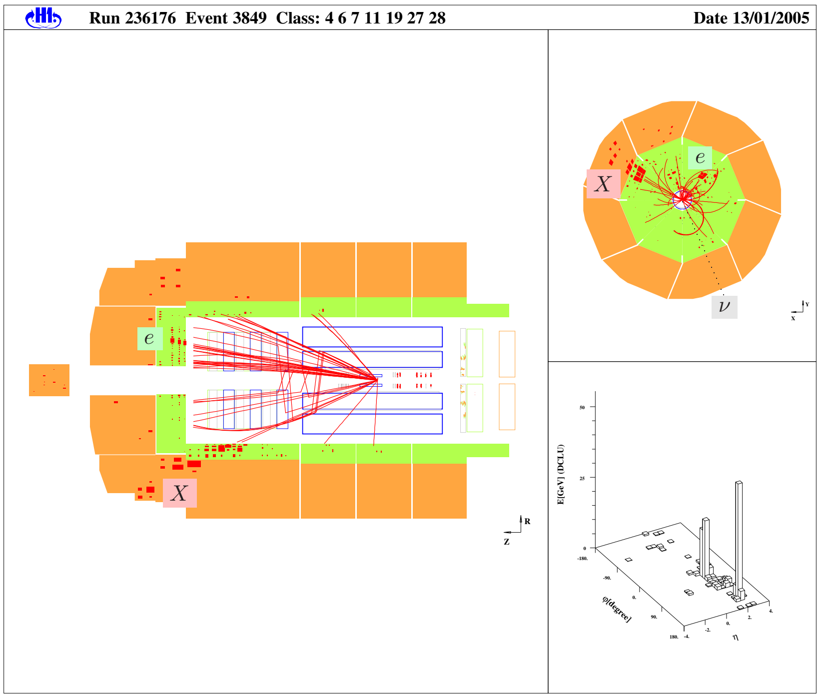


Figure G.7: *Display of event* 3849 *in run* 236176*: an event selected in the electron channel in the 1998/99* e^-p *dataset.*

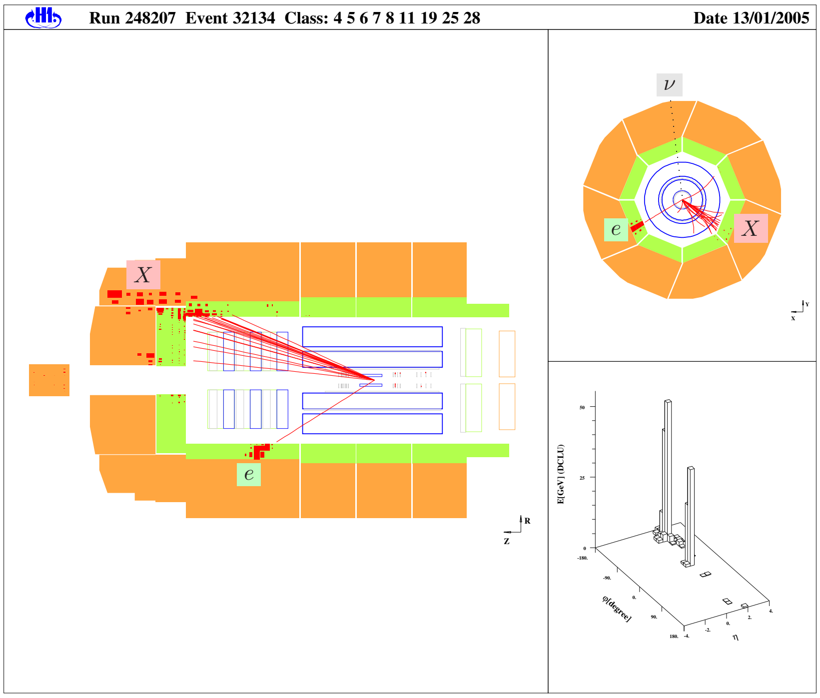


Figure G.8: *Display of event* 32134 *in run* 248207*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

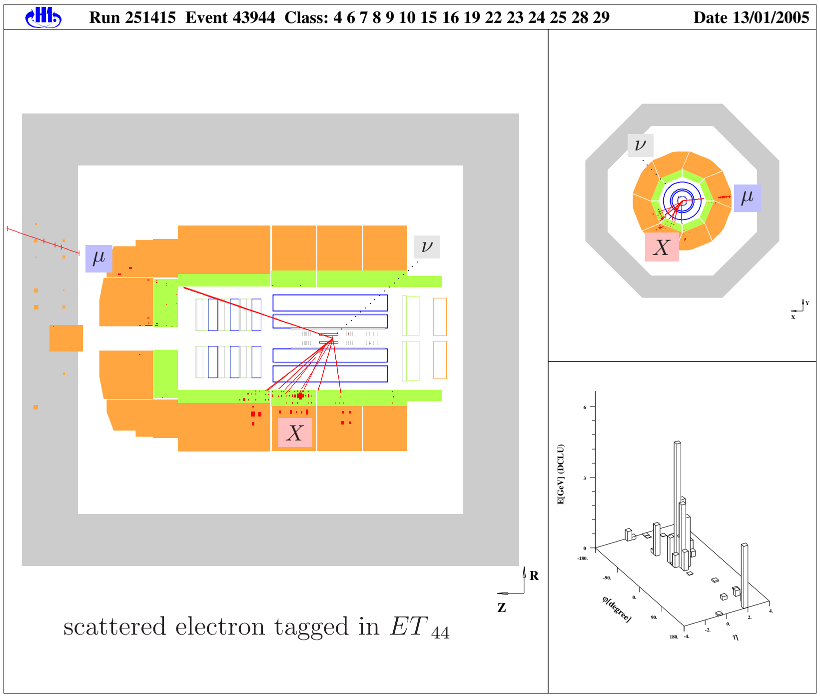


Figure G.9: *Display of event* 43944 *in run* 251415*: an event selected in the muon channel in the 1999/2000* e^+p *dataset. In this event, the scattered beam electron is detected in the* ET_{44} *electron tagger.*

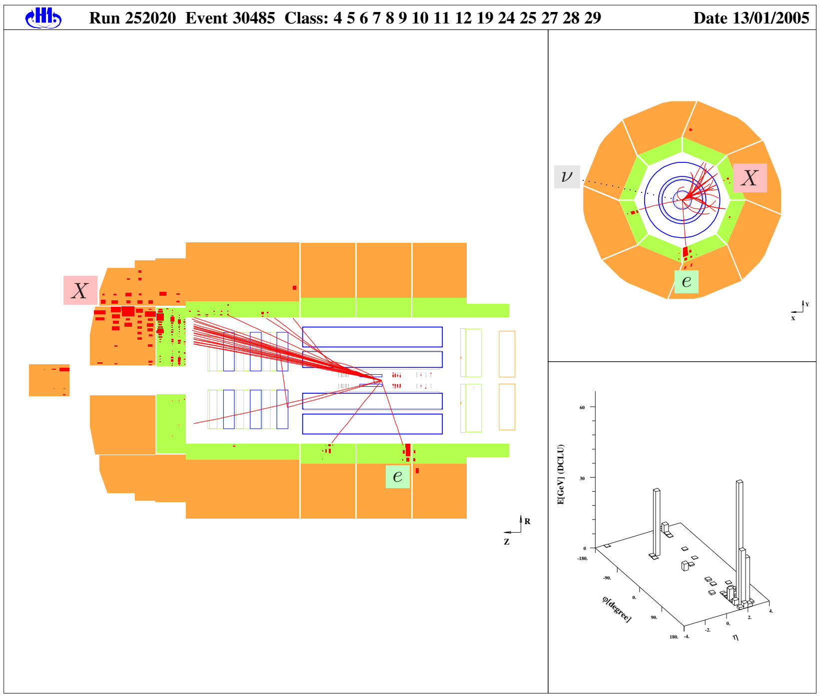


Figure G.10: *Display of event* 30485 *in run* 252020*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

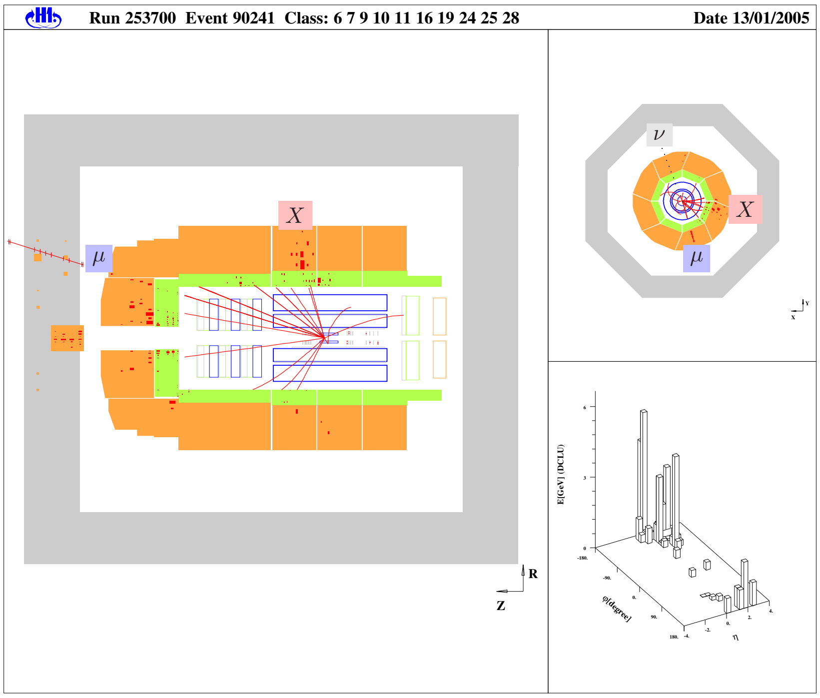


Figure G.11: *Display of event* 90241 *in run* 253700*: an event selected in the muon channel in the 1999/2000* e^+p *dataset.*

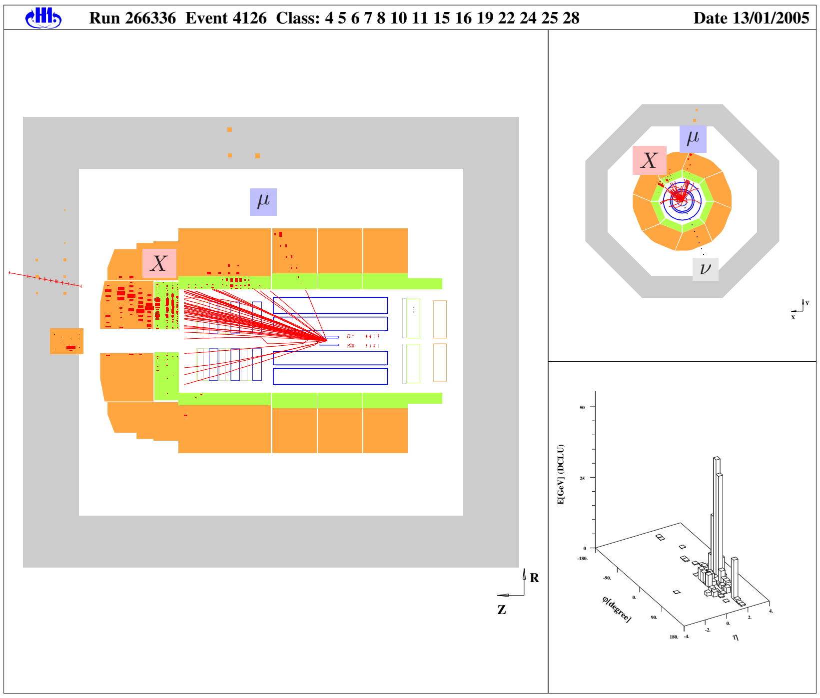


Figure G.12: *Display of event* 4126 *in run* 266336*: an event selected in the muon channel in the 1999/2000* e^+p *dataset.*

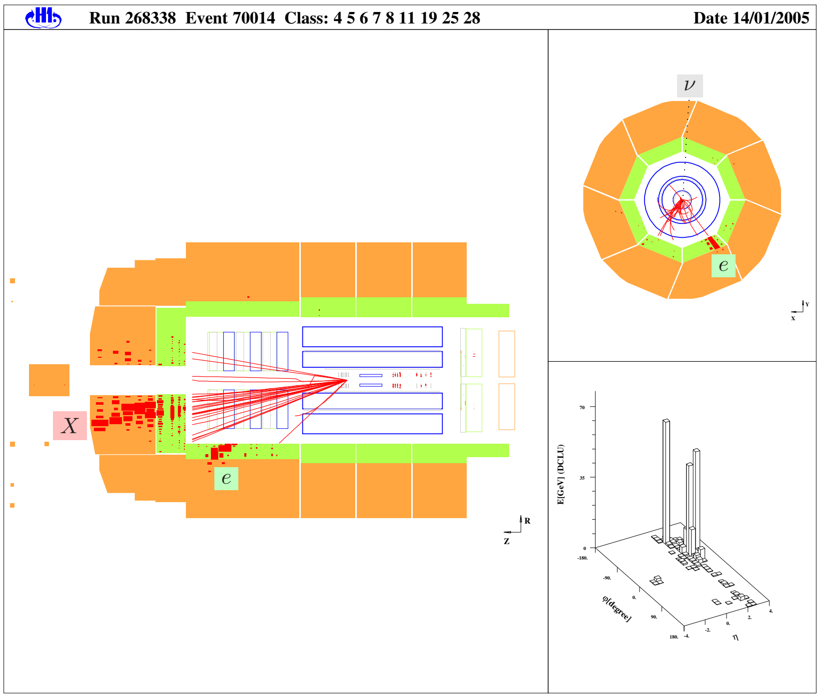


Figure G.13: *Display of event* 70014 *in run* 268338*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

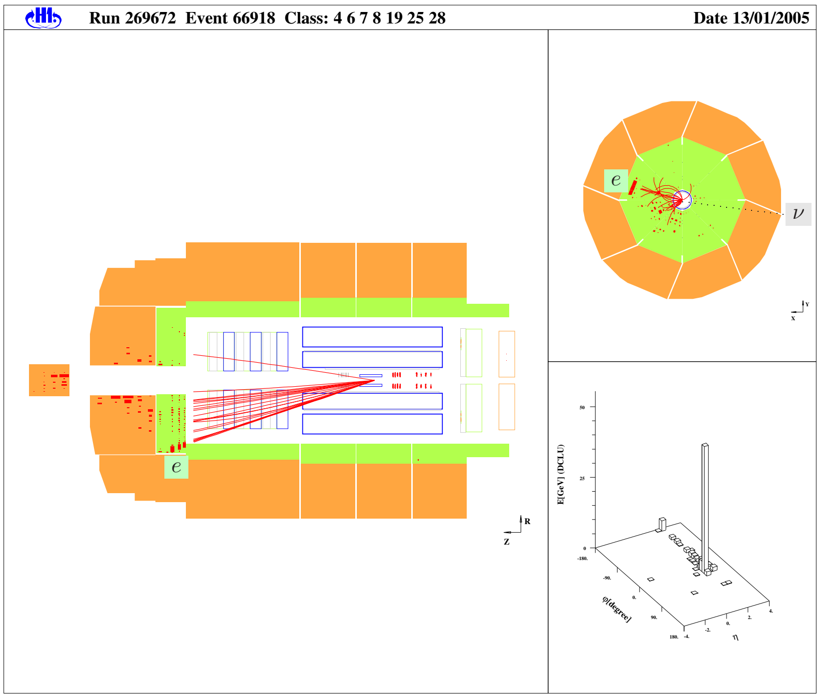


Figure G.14: *Display of event* 66918 *in run* 269672*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

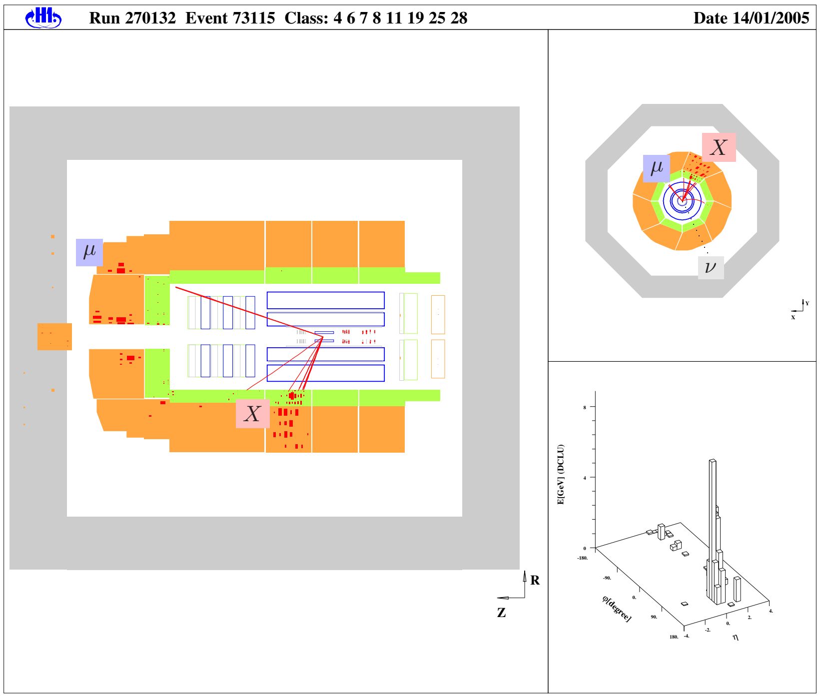


Figure G.15: *Display of event* 73115 *in run* 270132*: an event selected in the muon channel in the 1999/2000* e^+p *dataset.*

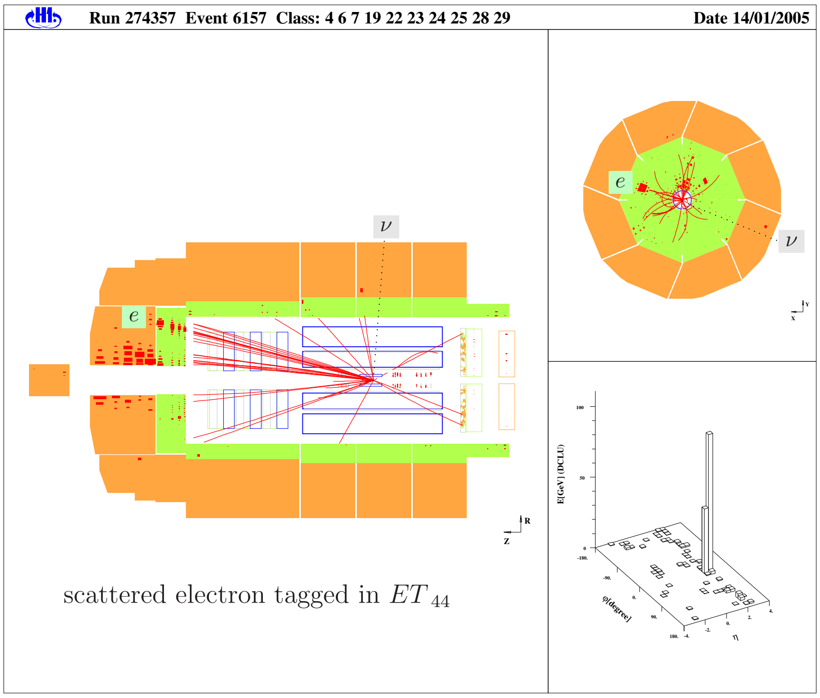


Figure G.16: *Display of event* 6157 *in run* 274357*: an event selected in the electron channel in the 1999/2000* e^+p *dataset. In this event, the scattered beam electron is detected in the* ET_{44} *electron tagger.*

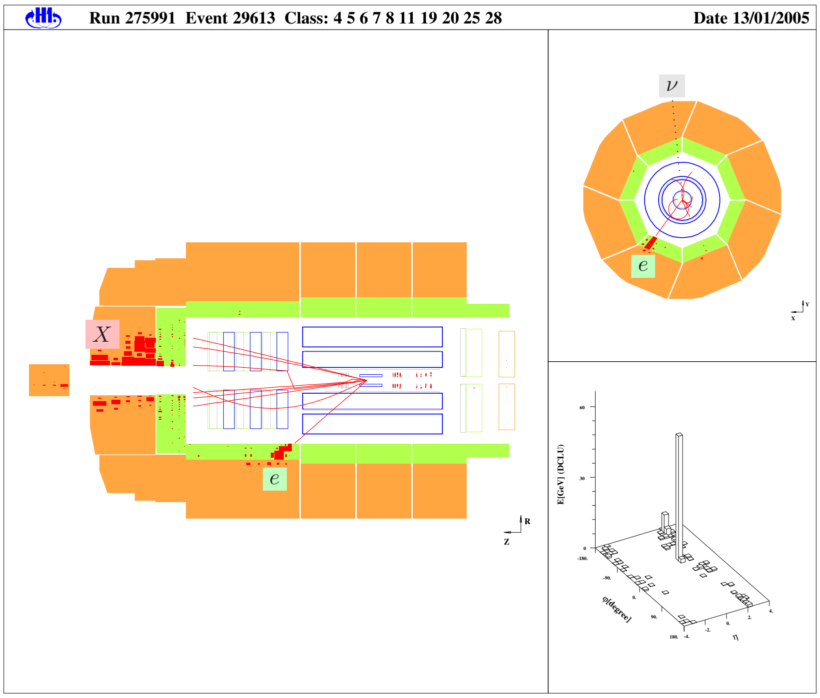


Figure G.17: *Display of event* 29613 *in run* 275991*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

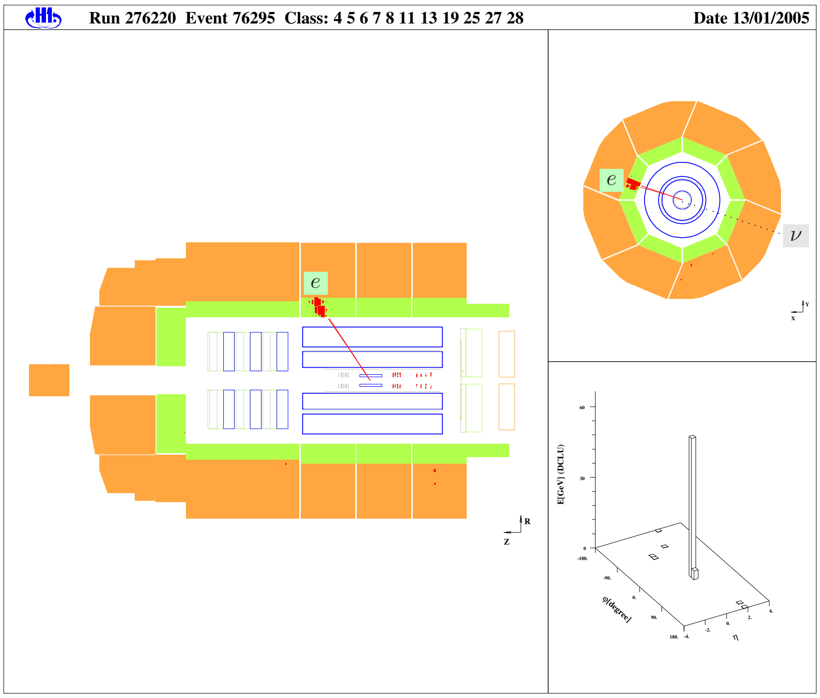


Figure G.18: *Display of event* 76295 *in run* 276220*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

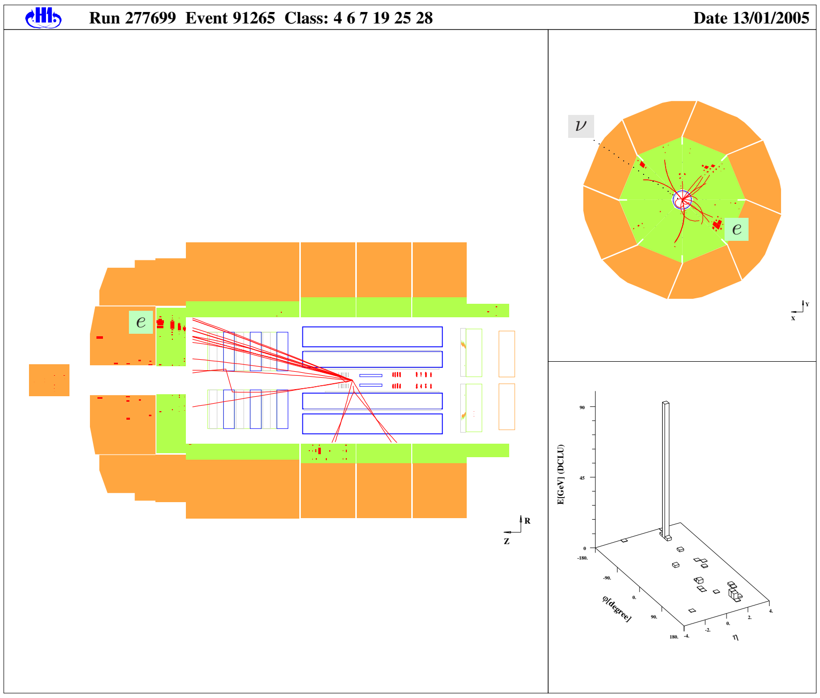


Figure G.19: *Display of event* 91265 *in run* 277699*: an event selected in the electron channel in the 1999/2000* e^+p *dataset.*

Appendix H

Tables

In this section, more detailed tables of the results presented in chapter 8 are shown. The number of events with isolated electrons, muons or tau leptons and large missing transverse momentum selected in the data is compared with the expected contributions from individual Standard Model signal and background processes.

H.1 Events with Isolated Electrons

P_T^X [GeV]	Data	W	NC	$\ell\bar{\ell}$	γp	CC
e^+p collisions						
< 12	6	4.12 ± 0.69	0.91 ± 0.33	0.32 ± 0.10	0.17 ± 0.09	1.15 ± 0.36
12 - 25	1	1.31 ± 0.21	0.10 ± 0.07	0.04 ± 0.02	0.01 ± 0.01	0.15 ± 0.05
25 - 40	1	0.75 ± 0.14	0.13 ± 0.08	0.02 ± 0.01	0.00 ± 0.00	0.13 ± 0.05
> 40	2	0.47 ± 0.14	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.09 ± 0.04
all P_T^X	10	6.65 ± 1.18	1.14 ± 0.40	0.39 ± 0.12	0.18 ± 0.09	1.52 ± 0.47
e^-p collisions						
< 12	0	0.55 ± 0.09	0.12 ± 0.05	0.04 ± 0.01	0.02 ± 0.01	0.40 ± 0.14
12 - 25	1	0.18 ± 0.03	0.02 ± 0.01	0.01 ± 0.00	0.00 ± 0.00	0.08 ± 0.04
25 - 40	0	0.10 ± 0.02	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.04 ± 0.02
> 40	0	0.07 ± 0.02	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.03 ± 0.02
all P_T^X	1	0.89 ± 0.16	0.15 ± 0.06	0.05 ± 0.02	0.02 ± 0.01	0.54 ± 0.18
ep collisions						
< 12	6	4.67 ± 0.78	1.02 ± 0.36	0.36 ± 0.12	0.19 ± 0.09	1.54 ± 0.48
12 - 25	2	1.49 ± 0.24	0.12 ± 0.07	0.05 ± 0.02	0.02 ± 0.01	0.23 ± 0.08
25 - 40	1	0.85 ± 0.16	0.14 ± 0.08	0.03 ± 0.01	0.00 ± 0.00	0.17 ± 0.06
> 40	2	0.53 ± 0.16	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.12 ± 0.05
all P_T^X	11	7.54 ± 1.34	1.28 ± 0.44	0.45 ± 0.14	0.20 ± 0.10	2.06 ± 0.64

Table H.1: *Number of events with isolated electrons and large missing transverse momentum selected using the selection criteria listed in table 7.2 in the 1994-2000 e^+p and e^-p datasets compared with the expectations for individual signal and background processes. The uncertainties quoted on the expectations are calculated by adding statistical and systematic uncertainties in quadrature.*

H.2 Events with Isolated Muons

H.2.1 LAr Triggered Phase-Space

P_T^X $[GeV]$	Data	W	NC	$\ell\bar{\ell}$	γp	CC
			e^+p collisions			
< 12	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
12 - 25	3	0.88 ± 0.13	0.00 ± 0.00	0.12 ± 0.04	0.00 ± 0.00	0.01 ± 0.01
25 - 40	2	0.72 ± 0.11	0.00 ± 0.00	0.11 ± 0.03	0.00 ± 0.00	0.01 ± 0.01
> 40	3	0.52 ± 0.08	0.00 ± 0.00	0.06 ± 0.02	0.00 ± 0.00	0.01 ± 0.01
all P_T^X	8	2.12 ± 0.32	0.00 ± 0.00	0.29 ± 0.09	0.00 ± 0.00	0.03 ± 0.01
			e^-p collisions			
< 12	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
12 - 25	0	0.11 ± 0.02	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
25 - 40	0	0.09 ± 0.01	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.07 ± 0.01	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
all P_T^X	0	0.27 ± 0.04	0.00 ± 0.00	0.03 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
			ep collisions			
< 12	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
12 - 25	3	0.99 ± 0.15	0.00 ± 0.00	0.14 ± 0.04	0.00 ± 0.00	0.01 ± 0.01
25 - 40	2	0.81 ± 0.12	0.00 ± 0.00	0.12 ± 0.04	0.00 ± 0.00	0.01 ± 0.01
> 40	3	0.59 ± 0.09	0.00 ± 0.00	0.07 ± 0.02	0.00 ± 0.00	0.01 ± 0.01
all P_T^X	8	2.40 ± 0.36	0.00 ± 0.00	0.32 ± 0.10	0.00 ± 0.00	0.03 ± 0.01

Table H.2: *Number of events with isolated muons and large missing transverse momentum in the LAr triggered phase-space selected using the selection criteria listed in table 7.4 in the 1994-2000 e^+p and e^-p datasets compared with the expectations for individual signal and background processes. The uncertainties quoted on the expectations are calculated by adding statistical and systematic uncertainties in quadrature.*

H.2.2 Muon Triggered Phase-Space

P_T^X $[GeV]$	Data	W	NC	$\ell\bar{\ell}$	γp	CC
			e^+p collisions			
< 12	0	1.51 ± 0.23	0.07 ± 0.05	1.18 ± 0.36	0.00 ± 0.00	0.02 ± 0.01
12 - 25	0	0.15 ± 0.03	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
25 - 40	0	0.03 ± 0.01	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
all P_T^X	0	1.70 ± 0.26	0.07 ± 0.05	1.20 ± 0.37	0.00 ± 0.00	0.02 ± 0.01
			e^-p collisions			
< 12	0	0.20 ± 0.03	0.01 ± 0.01	0.17 ± 0.05	0.00 ± 0.00	0.00 ± 0.00
12 - 25	0	0.02 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
25 - 40	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
all P_T^X	0	0.22 ± 0.03	0.01 ± 0.01	0.17 ± 0.05	0.00 ± 0.00	0.00 ± 0.00
			ep collisions			
< 12	0	1.71 ± 0.26	0.08 ± 0.05	1.35 ± 0.41	0.00 ± 0.00	0.02 ± 0.01
12 - 25	0	0.17 ± 0.03	0.00 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
25 - 40	0	0.03 ± 0.01	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
all P_T^X	0	1.92 ± 0.29	0.08 ± 0.05	1.37 ± 0.42	0.00 ± 0.00	0.02 ± 0.01

Table H.3: *Number of events with isolated muons and large missing transverse momentum in the muon triggered phase-space selected using the selection criteria listed in table 7.6 in the 1994-2000 e^+p and e^-p datasets compared with the expectations for individual signal and background processes. The uncertainties quoted on the expectations are calculated by adding statistical and systematic uncertainties in quadrature.*

H.2.3 Combined Phase-Space

P_T^X [GeV]	Data	W	NC	$\ell\bar{\ell}$	γp	CC
			e^+p collisions			
< 12	0	1.51 ± 0.43	0.07 ± 0.05	1.18 ± 0.46	0.00 ± 0.00	0.02 ± 0.01
12 - 25	3	1.04 ± 0.23	0.00 ± 0.00	0.13 ± 0.05	0.00 ± 0.00	0.01 ± 0.01
25 - 40	2	0.75 ± 0.16	0.00 ± 0.00	0.12 ± 0.04	0.00 ± 0.00	0.01 ± 0.01
> 40	3	0.52 ± 0.17	0.00 ± 0.00	0.06 ± 0.02	0.00 ± 0.00	0.01 ± 0.01
all P_T^X	8	3.82 ± 0.99	0.07 ± 0.05	1.49 ± 0.56	0.00 ± 0.00	0.04 ± 0.02
			e^-p collisions			
< 12	0	0.20 ± 0.06	0.01 ± 0.01	0.17 ± 0.06	0.00 ± 0.00	0.00 ± 0.00
12 - 25	0	0.13 ± 0.03	0.00 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
25 - 40	0	0.10 ± 0.02	0.00 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.07 ± 0.02	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
all P_T^X	0	0.49 ± 0.14	0.01 ± 0.01	0.20 ± 0.08	0.00 ± 0.00	0.00 ± 0.00
			ep collisions			
< 12	0	1.71 ± 0.48	0.08 ± 0.06	1.35 ± 0.52	0.00 ± 0.00	0.02 ± 0.01
12 - 25	3	1.16 ± 0.27	0.00 ± 0.00	0.15 ± 0.05	0.00 ± 0.00	0.01 ± 0.01
25 - 40	2	0.84 ± 0.18	0.00 ± 0.00	0.13 ± 0.04	0.00 ± 0.00	0.01 ± 0.01
> 40	3	0.60 ± 0.19	0.00 ± 0.00	0.07 ± 0.02	0.00 ± 0.00	0.01 ± 0.01
all P_T^X	8	4.31 ± 1.12	0.08 ± 0.06	1.70 ± 0.63	0.00 ± 0.00	0.04 ± 0.02

Table H.4: *Number of events with isolated muons and large missing transverse momentum in the combined LAr and muon triggered phase-space selected in the 1994-2000 e^+p and e^-p datasets compared with the expectations for individual signal and background processes. The uncertainties quoted on the expectations are calculated by adding statistical and systematic uncertainties in quadrature.*

H.3 Events with Isolated Tau Leptons

P_T^X $[GeV]$	Data	W	NC	$\ell\bar{\ell}$	γp	CC
			e^+p collisions			
< 12	0	0.41 ± 0.14	0.00 ± 0.00	0.12 ± 0.05	0.00 ± 0.00	1.01 ± 0.43
12 - 25	0	0.12 ± 0.05	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.03 ± 0.02
25 - 40	0	0.10 ± 0.04	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.02 ± 0.01
> 40	0	0.07 ± 0.05	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.01
all P_T^X	0	0.70 ± 0.28	0.00 ± 0.00	0.14 ± 0.06	0.00 ± 0.00	1.06 ± 0.45
			e^-p collisions			
< 12	0	0.05 ± 0.02	0.00 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.28 ± 0.13
12 - 25	0	0.02 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
25 - 40	0	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
> 40	0	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.01
all P_T^X	0	0.09 ± 0.03	0.00 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.29 ± 0.13
			ep collisions			
< 12	0	0.46 ± 0.16	0.00 ± 0.00	0.13 ± 0.06	0.00 ± 0.00	1.30 ± 0.55
12 - 25	0	0.14 ± 0.05	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.03 ± 0.02
25 - 40	0	0.11 ± 0.05	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.02 ± 0.01
> 40	0	0.08 ± 0.06	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.01
all P_T^X	0	0.79 ± 0.31	0.00 ± 0.00	0.16 ± 0.07	0.00 ± 0.00	1.36 ± 0.58

Table H.5: *Number of events with isolated tau leptons and large missing transverse momentum selected using the selection criteria listed in table 7.8 in the 1994-2000 e^+p and e^-p datasets compared with the expectations for individual signal and background processes. The uncertainties quoted on the expectations are calculated by adding statistical and systematic uncertainties in quadrature.*

Bibliography

[1] T. A. (*and others*), *Observation of an $e^+p \rightarrow mu^+X$ event with high transverse momenta at HERA*, DESY-94-248, 1994.

[2] C. A. (*and others*), *Observation of events with an isolated high energy lepton and missing transverse momentum at HERA*, Eur. Phys. J. **C 5** 575 (1998).

[3] N. Malden, *W production in ep collisions*, PhD thesis, University of Manchester, 2000.

[4] V. A. (*and others*), *Isolated electrons and muons in events with missing transverse momentum at HERA*, Phys. Lett. **B 561** 241 (2003).

[5] D. M. South, *Events with Isolated Leptons and Missing Transverse Momentum in e^+p Collisions at HERA*, PhD thesis, University of Manchester, 2003.

[6] A. A. (*and others*), *A general search for new phenomena in e p scattering at HERA*, Phys. Lett. **B 602** 14 (2004).

[7] U. Baur, J. A. M. Vermaseren, and D. Zeppenfeld, *Electroweak Vector Boson Production in High Energy ep collisions*, Nucl. Phys. **B 375** 3 (1992).

[8] P. Nason and M. Spira, *A Note on W Boson Production at HERA*, J. Phys. **G 25** 1434 (1999).

[9] K. P. O. Diener, C. Schwanenberger, and M. Spira, *Photoproduction of W Bosons at HERA: QCD Corrections*, Eur. Phys. J. **C 25** 405 (2002).

[10] J. Dingfelder, *Search for Anomalous Production of Single Top Quarks with the H1 Experiment at HERA*, Dissertation, Universität Heidelberg, 2003.

[11] A. A. (*and others*), *Search for single top quark production in ep collisions at HERA*, Eur. Phys. J. **C 33** 9 (2004).

[12] G. H. Hoffstaetter, *The luminosity upgrade of HERA*, ICFA Beam Dyn. Newslett. **24** 8 (2001). Also published as DESY-M-01-01A.

[13] P. V. E. (*and others*), *The H1 forward proton spectrometer at HERA*, Nucl. Instr. Meth. **A 446** 409 (2000).

[14] B. List, *Diffraktive J/Ψ-Production in Elektron-Proton Stößen am Speicherring HERA*, Diplomarbeit, Technische Universität Berlin, 1993. Also published as H1-internal note H1-10/93-319.

[15] W. B. (*and others*), *Upgrade of the H1 forward Neutron Calorimeter*, H1-internal note H1-11/99-578, 1999.

[16] I. A. (*and others*), *The H1 detector at HERA*, Nucl. Instrum. Meth. **A 386** 310 (1997).

[17] J. B. (*and others*), *The Central Jet Chamber of the H1 Experiment*, Nucl. Instr. Meth. **A 279** 217 (1989).

[18] M. C. (*and others*), *H1 High Luminosity Upgrade: CIP and Level 1 Vertex Trigger*, H1-internal note H1-01/98-535, 1998.

[19] G. A. B. (*and others*), *Radial Wire Drift Chambers For The H1 Forward Track Detector At Hera: Design, Construction And Performance*, Nucl. Instrum. Meth. **A 283** 471 (1989).

[20] K. Kleinknecht, *Detektoren für Teilchenstrahlung*, Teubner, Stuttgart, 1984.

[21] H. Collaboration, *Proposal for an Upgrade of the H1 Forward Track Detector for HERA 2000*, DESY PRC 98/06, 1998.

[22] B. Schwab, *Das Rückwärtsdriftkammersystem des H1 Experiments: Simulation, Rekonstruktion, Kalibration und erste Datenanalyse*, Dissertation, Universität Heidelberg, 1996.

[23] D. H. Perkins, *Introduction to high energy physics*, Addison-Wesley, Reading, Massachusetts, 1987.

[24] C. Grupen, *Teilchendetektoren*, BI-Wissenschaftsverlag, Mannheim-Leipzig-Wien-Zürich, 1993.

[25] B. A. (*and others*), *Electron/Pion Separation with the H1 LAr Calorimeters*, Nucl. Instr. Meth. **A 344** 492 (1994).

[26] B. A. (*and others*), *The H1 Liquid Argon Calorimeter System*, Nucl. Instr. Meth. **A 336** 460 (1993).

[27] V. Shekelyan, *Simulation and Reconstruction in H1 Liquid Argon Calorimetry*, H1-internal note H1-04/93-288, 1993.

[28] L. Goerlich and H. P. Wellisch, *Documentation of the LAr clustering*, H1-internal note H1-12/91-204, 1991.

[29] C. Issever, K. Borras, and D. Wegener, *An Improved Weighting Algorithm to Achieve Software Compensation in a Fine Grained LAr Calorimeter*, H1-internal note H1-03/03-608, 2003.

[30] H. Wellisch, J. P. Kubenka, H. Oberlack, and P. Schacht, MPI-PhE/94-03, 1994.

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

[32] M. Peez, B. Portheault, and E. Sauvan, *An energy flow algorithm for Hadronic Reconstruction in OO: Hadroo2*, H1-internal note H1-01/05-616, 2005.

[33] B. A. (*and others*), *Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations*, Nucl. Instr. Meth. **A 336** 499 (1993).

[34] H. SpaCal Group: R. D. Appuhn (*and others*), *Hadronic Response and e/π-Separation with the H1 Lead/Fibre Calorimeter*, Nucl. Instr. Meth. **A 382** 395 (1996).

[35] H. SpaCal Group: T. Nicholls (*and others*), *Performance of an Electromagnetic Lead/Scintillating-fibre Calorimeter for the H1 Detector*, Nucl. Instr. Meth. **A 374** 149 (1996).

[36] D. Reyna, *Modifications to SpaCal for H1 High Luminosity Operation and the Effect on Acceptance*, H1-internal note H1-11/98-555, 1998.

[37] *Upgrade of the H1 hadron PLUG calorimeter*, 1998. Available from the H1 Collaboration.

[38] B. Ludwigshafen, Germany.

[39] P. B. (*and others*), *The H1 forward muon spectrometer*, Nucl. Instr. Meth. **A 340** 304 (1994).

[40] H. Bethe and W. Heitler, *On the Stopping of fast Particles and on the Creation of Positive Electrons*, Proc. Roy. Soc. Lond. **A 146** 83 (1934).

[41] H. Collaboration, *Luminosity Measurement in the H1 Experiment at HERA*, Paper submitted to ICHEP 1996.

[42] D. Pieloth, *Experimentelle Überprüfung der Algorithmen zur Luminositätsbestimmung am H1-Detektor*, Diplomarbeit, Universität Dortmund, 1995.

[43] P. A. Čerenkov, *Visible Radiation Produced by Electrons Moving in a Medium with Velocities exceeding that of Light*, Phys. Rev. **52** 378 (1937).

[44] S. A. B. (*and others*), *Addendum to the Proposal "A Fast Track Trigger with High Resolution for H1"*, H1-internal note H1-09/99-576, 1999.

[45] J. Naumann, *Entwicklung und Test der dritten H1-Triggerstufe*, Dissertation, Universität Dortmund, 2002.

[46] J. H. Field, *Triggering the H1 Detector*, H1-internal note H1-12/85-38, 1985.

[47] D. Hoffmann, *Zwei-Elektronen-Ereignisse im H1-Detektor*, Dissertation, Universität Hamburg, 2000.

[48] E. Barrelet, R. J. Ellison, and W. J. H. (*and others*), *A Proposal for the Software L3 Triggers ahead of Event Building in H1*, H1-internal note H1-09/88-92, 1988.

[49] E. B. (*and others*), *The hardware implementation of L3 triggers in H1*, H1-internal note H1-09/88-100, 1988.

[50] A. B. (*and others*), *A fast high resolution track trigger for the H1 experiment*, IEEE Trans. Nucl. Sci. **48** 1276 (2001).

[51] A. C. (*and others*), *Proposal to Merge Level-4 and Level-5 Systems of the H1 Experiment*, H1-internal note H1-12/98-558, 1998. Also published as DESY PRC 99-03.

[52] V. Lemaitre, *L4: Status Report Filtering and Classification*, 1998. Available from http://www-h1.desy.de/itrigger/L4Farm/l4-1998.ps.

[53] P. H.Albrecht, M.Erdmann, *A guide to PHAN - an H1 physics analysis package*, Version 2.03/11. Available from the H1 Collaboration.

[54] A. S. G. CERN, *HBOOK Reference Manual*, Version 4.22, 1993. Available from http://wwwinfo.cern.ch/asdoc/hbook_html3/hboomain.html.

[55] R. Brun, O. Couet, C. E. Vandoni, and P. Zanarini, *PAW: A General Purpose Portable Software Tool for Data Analysis and Presentation*, Comput. Phys. Commun. **57** 432 (1989).

[56] H. Collaboration, *The new object oriented analysis framework for H1*, Talk given at conference for Computing in High-Energy and Nuclear Physics (CHEP) 2003.

[57] R. Brun and F. Rademakers, *ROOT: An Object Oriented Data Analysis Framework*, Nucl. Instr. Meth. **A 389** 81 (1997).

[58] The H1OO Group, *The H1OO Physics Analysis Project*, 2004. Available from the H1 Collaboration.

[59] P. C. (*and others*), *Version Management with CVS*, 1993. Available from http://www.cvshome.org/docs/manual/.

[60] P. Laycock, *A Measurement of the Diffractive Reduced Cross-Section* $sigma_r^D(3)$ *at High* Q^2 *with the H1 Detector at HERA*, PhD thesis, University of Liverpool, 2003.

[61] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*, Wiley, New York, 1981.

[62] *Guide to Simulation Program H1SIM.* Available from the H1 Collaboration.

[63] R. B. (*and others*), *GEANT3 User's Guide.* CERN-DD/EE-84-1.

[64] J. D. Bjorken and E. Paschos, *Inelastic Electron-Proton and* γ*-Proton Scattering and the Structure of the Nucleon*, Phys. Rev. **185** 1975 (1969).

[65] E. Rutherford, *The scattering of a and b particles off Matter and the Structure of the Atom*, Phil. Mag. **21** 669 (1911).

[66] M. B. (*and others*), *Observed Behavior of highly inelastic electron-proton scattering*, Phys. Rev. Lett. **23** 935 (1969).

[67] C. A. (*and others*), *Measurement and QCD analysis of neutral and charged current cross sections at HERA*, Eur. Phys. J. **C 30** 1 (2003).

[68] A. C. B. (*and others*), *Test of QCD and a Measurement of λ from Scaling Violations in the Proton Structure Function $F_2(x, Q^2)$ at high Q^2*, Phys. Lett. **B 223** 485 (1989).

[69] P. A. (*and others*), *Proton and Deuteron F_2 Structure Functions in Deep Inelastic Muon Scattering*, Phys. Lett. **B 295** 159 (1992).

[70] M. A. (*and others*), *Measurement of the Proton and the Deuteron Structure Functions, F_2^p and F_2^d*, Phys. Lett. **B 364** 107 (1995).

[71] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books, Reading Massachusetts, 1995.

[72] W. Buchmüller and G. Ingelman, editors, *DJANGO: The Interface for the Event Generators HERACLES and LEPTO*, 1991, Proceedings of the workshop "Physics at HERA", volume 3, DESY, Hamburg.

[73] H. Jung, *Hard diffractive scattering in high-energy ep collisions and the Monte Carlo generator RAPGAP*, Comput. Phys. Commun. **86** 147 (1995).

[74] C. von Weizsäcker, *Radiation emitted in Collisions of very fast Electrons*, Z. Phys. **88** 612 (1934).

[75] E. J. Williams, *Nature of the high energy Particles of penetrating Radiation and Status of Ionization and Radiation Formulae*, Phys. Ref. **45** 729 (1934).

[76] J. Dainton, *Photon structure and quantum fluctuation*, Phil. Trans. Roy. Soc. Lond. **A 359** 279 (2001). Also published as DESY preprint 00-137.

[77] M. D. (*and others*), *Dijet cross-sections in photoproduction at HERA*, Phys. Lett. **B 348** 665 (1995).

[78] J. J. Sakurai, *Theory Of Strong Interactions*, Annals Phys. **11** 1 (1960).

[79] J. J. Sakurai and D. Schildknecht, *Generalized Vector Dominance And Inelastic Electron - Proton Scattering*, Phys. Lett. **B 40** 121 (1972).

[80] M. Greko, *Deep-Inelastic Processes*, Nucl. Phys. **B 63** 398 (1973).

[81] R. P. Feynman, *Photon-hadron interactions*, Benjamin, Reading, Massachusetts, 1972.

[82] D. E. G. (*and others*), *Review of Particle Physics*, Eur. Phys. J. **C 15** 1 (2000).

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

[84] G. C. (*and others*), *HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** 10 (2001). Also published as hep-ph/0011363.

[85] R. Engel and J. Ranft, *Color singlet exchange between jets and the PHOJET Monte Carlo*, Nucl. Phys. Proc. Suppl. **75A** 272 (1999).

[86] G. Ingelman, J. Rathsman, and G. A. Schuler, *AROMA 2.2: A Monte Carlo Generator for Heavy Flavour Events in ep collisions*, Comput. Phys. Commun. **101** 135 (1997).

[87] C. A. (*and others*), *Measurement of open beauty production at HERA*, Phys. Lett. **B 467** 156 (1999).

[88] D. J. Griffiths, *Introduction to elementary particles*, Wiley, New York, 1987.

[89] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep Scattering in Perturbation Theory*, Sov. J. Nucl. Phys. **15** 438 (1972).

[90] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. **B 126** 298 (1977).

[91] Y. L. Dokshitzer, *Calculation of structure functions of deep-inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, Sov. Phys. JETP **46** 641 (1977).

[92] G. Gustafson, *Dual Description Of A Confined Color Field*, Phys. Lett. **B 175** 453 (1986).

[93] B. A. (*and others*), *Parton Fragmentation And String Dynamics*, Phys. Rept. **97** 31 (1983).

[94] G. C. Fox and S. Wolfram, *A Model For Parton Showers In QCD*, Nucl. Phys. **B 168** 285 (1980).

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

[96] B. Gary, Lecture given at the CTEQ summerschool 2003 in Sant Feliu de Guixols, Spain.

[97] B. Leißner, *Muon Pair Production in Electron-Proton Collisions*, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2002.

[98] N. Arteaga-Romero, C. Carimalo, and P. Kessler, *High P_t lepton pair production at ep colliders: Comparison between various production mechanisms*, Z. Phys. **C 52** 289 (1991).

[99] J. A. M. Vermaseren, *Two Photon Processes at very High Energies*, Nucl. Phys. **B 229** 347 (1983).

[100] W. Buchmüller and G. Ingelman, editors, *LPAIR: A Generator for Lepton Pair Production*, 1991, Proceedings of the workshop "Physics at HERA", volume 3, DESY, Hamburg.

[101] T. Abe, *GRAPE-Dilepton (Version 1.1): A Generator for Dilepton Production in ep Collisions*, Comput. Phys. Commun. **136** 126 (2001).

[102] A. T. Doyle, G. Grindhammer, G. Ingelman, and H. Jung, editors, *GRAPE-Dilepton (Version 1.0)*, 1999, Proceedings of the workshop "Monte Carlo Generators for HERA Physics", DESY, Hamburg.

[103] A. T. Doyle, G. Grindhammer, G. Ingelman, and H. Jung, editors, *Lepton pair Monte Carlo generators for HERA physics*, 1999, Proceedings of the workshop "Monte Carlo Generators for HERA Physics", DESY, Hamburg.

[104] A. A. (*and others*), *Multi-electron production at high transverse momenta in ep collisions at HERA*, Eur. Phys. J. **C 31** 17 (2003).

[105] A. A. (*and others*), *Muon pair production in ep collisions at HERA*, Phys. Lett. **B 583** 28 (2004).

[106] W. Buchmüller and G. Ingelman, editors, *The Monte Carlo Generator EPJPSI for J/Ψ Mesons in High Energy Electron Proton Collisions*, 1991, Proceedings of the workshop "Physics at HERA", volume 3, DESY, Hamburg.

[107] J. Kroseberg, *Untersuchungen zum experimentellen Nachweis von W-Produktion*, Diplomarbeit, Universität Münster, 1998.

[108] P. A. (*and others*), *Measurement of triple gauge boson couplings of the W boson at LEP*, Phys. Lett. **B 586** 151 (2004).

[109] K. J. F. Gaemers and G. J. Gouraris, *Polarisation Amplitudes for $e^+e^- \to W^+W^-$ and $e^+e^- \to ZZ$*, Z. Physics **C 1** 259 (1979).

[110] C. D. (*and others*), *H1EPVEC - $W^{\pm}$ and Z Production Monte Carlo Generator based on EPVEC*, 1999. CPPM Marseille.

[111] W. Buchmüller, R. Rückl, and D. Wyler, *Leptoquarks In Lepton Quark Collisions*, Phys. Lett. **B 191** 442 (1987). Erratum-ibid. B **448** (1999) 320.

[112] C. A. (*and others*), *A Search for Leptoquark Bosons and Lepton Flavor Violation in Positron-Proton Collisions at HERA*, Eur. Phys. J. **C 11** 447 (1999).

[113] H. Collaboration, *Search for Lepton Flavor Violation in ep Collisions at HERA*, Paper submitted to ICHEP 2004.

[114] K. Hagiwara, D. Zeppenfeld, and S. Komamiya, *Excited Lepton Production at LEP and HERA*, Z. Phys. **C 29** 115 (1985).

[115] C. A. (*and others*), *A Search for Excited Fermions at HERA*, Eur. Phys. J. **C 17** 567 (2000).

[116] T. Kon, T. Kobayashi, and S. Kitamura, *Is the High P_T Muon of the $e^+p \rightarrow \mu^+X$ Event observed at HERA a Signature of the Stop ?*, Phys. Lett. **B 376** 227 (1996).

[117] T. K. (*and others*), *Single sbottom/scharm Production at HERA in an R-parity Breaking SuperSymmetric Model*, Phys. Lett. **B 494** 280 (2000).

[118] A. A. (*and others*), *Search for squark production in R-parity violating supersymmetry at HERA*, Eur. Phys. J. **C 36** 227 (2004).

[119] J. Haller, *Search for Squark Production in R-Parity Violating Supersymmetry at HERA*, Dissertation, Universität Heidelberg, 2003.

[120] A. A. (*and others*), *Search for bosonic stop decays in R-parity violating supersymmetry in e+ p collisions at HERA*, Phys. Lett. **B 599** 159 (2004).

[121] A. Vest, *Search for Bosonic Stop Decays in R-Parity Violating Supersymmetry in e^+p Collisions at HERA*, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2004.

[122] H. Fritsch and D. Holtmannspötter, *The Production of Single t Quarks at LEP and HERA*, Phys. Lett. **B 457** 186 (1999).

[123] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions With Lepton - Hadron Symmetry*, Phys. Rev. **D 2** 1285 (1970).

[124] S. C. (*and others*), *Search for single-top production in e p collisions at HERA*, Phys. Lett. **B 559** 153 (2003).

[125] P. A. (*and others*), *Search for Single Top Production at LEP*, Phys. Lett. **B 549** 290 (2002).

[126] F. A. (*and others*), *Search for Flavor Changing Neutral Current Decays of the Top Quark in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8\ TeV$*, Phys. Rev. Lett. **80** 2525 (1998).

[127] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, Wiley, New York, 2001.

[128] J. Rauschenberger, *Prozesse des geladenen Stromes in tief-unelastischer Positron-Proton Streuung bei HERA*, Dissertation, Universität Hamburg, 2002.

[129] P. D. Wilde, *Neural network models : theory and projects*, Springer, London, 1997.

[130] A. Zell, *Simulation Neuronaler Netze*, Addison-Wesley, Bonn-Paris-Reading, Massachusetts, 1994.

[131] R. Rojas, *Neural networks : a systematic introduction*, Springer, Berlin-Heidelberg, 1996.

[132] A. N. Kolmogorov, *On the representation of continous functions of several variables by superposition of continous functions of one variable and addition*, Doklady Akademiia Nauk SSSR **114** 953 (1957).

[133] J. D. Jackson, *Classical Electrodynamics*, Wiley, New York, 1998.

[134] A. A. (*and others*), *Evidence for a narrow anti-charmed baryon state*, Phys. Lett. **B 588** 17 (2004).

[135] P. Bruel, *Recherche d'interactions au-dela du Modele Standard a HERA*, PhD thesis, Universite Paris Orsay, 1998.

[136] A. A. Glazov, *Measurement of the Proton Structure Functions $F_2(x, Q^2)$ and $F_L(x, Q^2)$ with the H1 Detector at HERA*, Dissertation, Humboldt-Universität Berlin, 1998.

[137] B. Naroska, S. Schiek, and G. Schmidt, *Lepton Identification in the H1 Detector at Low Momenta*, H1-internal note H1-05/97-518, 1997.

[138] O. Henshaw, *A First Measurement of the Neutral Current Cross-Section at High Q^2 in Longitudinally Polarised e^+p Scattering with the H1 Detector*, PhD thesis, University of Birmingham, in preparation.

[139] S. Schiek, *Untersuchung der inelastischen Photoproduktion von J/Ψ-Mesonen im H1-Detektor bei HERA*, Dissertation, Universität Hamburg, 1996.

[140] E. Kuntze, *Rekonstruktion von Myonen im H1-Detektor*, Diplomarbeit, Universität Hamburg, 1992.

[141] U. P. Krüger, *Untersuchung der Erzeugung schwerer Quarks durch ihren Zerfall in Myonen im H1-Detektor bei HERA*, Dissertation, Universität Hamburg, 1995.

[142] U. Langenegger, *A Measurement of the Beauty and Charm Production Cross Sections at the ep Collider HERA*, Dissertation, Eidgenössische Technische Hochschule Zürich, 1998.

[143] M. Reece and H. Frisch, *Using Track Impact Parameter To Distiguish Hadronically-Decaying τ's in Top Quark Decays from Light Quark Jets*, CDF/Anal/Top/CDFR/6203.

[144] H. Collaboration, *Measurement of charm and beauty production in deep inelastic e+p scattering at high Q^2 at HERA*, Paper submitted to DIS 2004, 2004.

[145] B. K. Bullock, K. Hagiwara, and A. D. Martin, *Tau polarization and its correlations as a probe of new physics*, Nucl. Phys. **B 395** 499 (1993).

[146] G. I. Brandt, *Suche nach τ-Leptonen mit hohem Transversalimpuls aus Zerfällen einzeln erzeugter top-Quarks bei HERA*, Diplomarbeit, Universität Heidelberg, 2003.

[147] L. West, *How to use the Heavy Flavour Working Group Track, Muon and Electron Selection Code*, 2000. Available from the H1 Collaboration.

[148] J. Cao and Z. Zhang, *Towards an Unbiased Measurement of Kinematic Variables at Low y Region*, H1-internal note H1-12/99-580, 1999.

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

[150] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D 48** 3160 (1993).

[151] A. Mehta, *Z Vertex Determination*, Talk given at the H1 HiP Working Group meeting, January 2002.

[152] V. A. (*and others*), 2002.

[153] A. Pieuchot, J. Marks, and C. Vallee, *Pile-Up Energy Detection Using the Liquid Argon Big Tower Timing Information*, H1-internal note H1-01/95-423, 1995.

[154] I. N. (*and others*), *A Minimal Comprehensive Set of Muon Background Topological Finders for High P_T Physics Analysis*, H1-internal note H1-10/96-498, 1996.

[155] E. C. (*and others*), *QBGFMAR: An Updated Phan Package for Cosmic and Halo Muon Topological Rejection in High P_T Physics Analysis*, H1-internal note H1-11/98-556, 1998.

[156] C. Veelken, *H1NonepBgFinder - Rejection of Cosmic Muon and Beam-Halo Events in the H1OO Framework*, H1-internal note H1-09/02-603, 2002.

[157] F. Keil, *Dijet Production in Charged and Neutral Current e^+p Interactions at High Q^2 at HERA*, Dissertation, Universität Heidelberg, 2001.

[158] S. Mohrdieck, *Neuronaler Netzwerk-Trigger fuer Myonen im H1-Detektor bei HERA*, Diplomarbeit, Universität Hamburg, 1997.

[159] S. E. (*and others*), *Calculating Event Weights in Case of Downscaling on Trigger Levels 1-4*, H1-internal note H1-04/97-517, 1997.

[160] A. Mehta, Personal communication.

[161] I. S. Gradshteyn and I. M. Rhyzhik, *Table of Integrals, Series and Products*, Academic Press, New York, fourth edition, 1980.

[162] W. J. Thompson, *Atlas for computing mathematical functions: an illustrated guidebook for practitioners, with programs in C and Mathematica*, Wiley, New York, 1997.

[163] D. Dannheim, *Search for Events with Isolated Leptons and Large Missing Transverse Momentum in ep Collisions at HERA*, Dissertation, Universität Hamburg, 2003.

[164] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. **A 434** 435 (1999).

[165] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse*, Teubner, Stuttgart-Leipzig, 1998.

[166] G. Frising, *Eine Likelihood-Analyse der H1 Jet-Daten*, Diplomarbeit, Rheinisch-Westfälische Technische Hochschule Aachen, 2000.

[167] C. A. (*and others*), *Diffraction dissociation in photoproduction at HERA*, Z. Phys. **C 74** 221 (1997).

[168] S. C. (*and others*), *Isolated tau leptons in events with large missing transverse momentum at HERA*, Phys. Lett. **B 583** 41 (2004).

[169] C. Veelken, *Search for Events with isolated Leptons and large Missing Transverse Momentum*, Talk presented at DIS 2005.

[170] *H1REC - H1 Reconstruction Program*, Version 9.02/00. Available from the H1 Collaboration.

[171] M. Kuhlen, *The Fast H1 Detector Monte Carlo*, H1-internal note H1-10/92-254, 1992.

[172] C. Schwanenberger, *Improved Description of Electromagnetic Showers in H1FAST*, Talk given at H1 collaboration meeting in Marseille, 2003.

[173] A. Meyer, Personal communication.

[174] T. C. (*and others*), *Proposal to Upgrade the LAr Calorimeter Trigger: The Jet Trigger*, H1-internal note H1-01/99-560, 1999. Also published as DESY PRC 99-02.

[175] T. Wolff, *Entwicklung, Bau und erste Ergebnisse eines totzeitfreien Spurfinders als Trigger fuer das H1-Experiment am Speicherring HERA*, Dissertation, Eidgenössische Technische Hochschule Zürich, 1993.

[176] Y. H. Fleming, *The H1 First Level Fast Track Trigger*, PhD thesis, University of Birmingham, 2003.

[177] Y. H. Fleming, Personal communication.

[178] C. Wissing, *Entwicklung eines Simulationsprogramms und Implementierung schneller Spurfitalgorithmen für den neuen H1-Driftkammertrigger*, Dissertation, Universität Dortmund, 2002.

Acknowledgements

The field of modern high energy physics is a highly collaborative science. The work presented in this thesis would not have been possible without the support of many colleagues in the H1 Collaboration, for which I am grateful.

In particular, I would like to thank:

Cristinel Diaconu, for helping to prepare the paper on the tau results submitted to ICHEP 2005, and finding some time to dig up details of previous Fortran analyses; Joachim Meyer for his advice on the analysis and sharing his expertise on the H1 event display; and my supervisor Timothy Greenshaw for proof-reading the whole thesis and his help with organisational matters.

For her help with organisational and financial matters I would also like to thank the UK liaison officer Susan Ketels, who always made students feel welcome in her office.

Furthermore, I would like to thank:

Paul Laycock for his help with the generic analysis framework; and David South for sharing the code specific to the isolated lepton analysis with me and being such an active referee on the tau lepton results;

Gerhard Brandt and Stefania Xella, for the good team-work on the search for events with isolated tau leptons;

Yves Fleming, Aweys Mohamed and Andrew Rimmer for the friendly atmosphere in the Liverpool office;

and finally John Dainton, Beate Heinemann and Andrew Mehta for inviting me to Liverpool in the first place.