Analyse der Erzeugung schwerer Quarks in ep-Wechselwirkung mit Neuronalen Netzen

Diplomarbeit
am II. Institut für Experimentalphysik
der Universität Hamburg

vorgelegt von
Marcel Propp

Hamburg
Juni 2000

Inhaltsverzeichnis

Einleitung

Bei der 1992 in Betrieb genommenen Speicherringanlage HERA werden Elektronen und Protonen bei einer Schwerpunktsenergie von $\sqrt{s} \approx 300$ GeV zur Wechselwirkung gebracht. Bei dieser Energie können u.a. auch schwere Charm(c)- und Beauty(b)-Quark-Antiquarkpaare erzeugt werden. Die Produktion schwerer Quarks wird bei den HERA-Energien von der Photon-Gluon-Fusion dominiert. Bei diesem Vorgang wechselwirkt ein vom Elektron abgestrahltes Photon unter Erzeugung eines Quark und Antiquarks mit einem Gluon des Protons. Aufgrund der schwereren Quarkmasse ist die b-Produktion um ca. einen Faktor 200 gegenüber der c-Produktion unterdrückt.

Der Wirkungsquerschnitt (WQ) für die Erzeugung von b- und c-Quarks in der ep-Wechselwirkung kann mit störungstheoretischen Rechnungen der Quanten-Chromo-Dynamik (QCD) berechnet werden. Die theoretische Ungenauigkeit des berechneten c-Wirkungsquerschnittes ist wesentlich gröser als die des b-Wirkungsquerschnittes. Dies resultiert aus der für die Berechnung verwendeten Energieskala, die im Bereich der Photoproduktion ($Q^2 \approx 0$) in etwa der Quarkmasse gleichgesetzt wird. Aufgrund der niedrigeren c-Masse ergeben sich also die größeren Ungenauigkeiten bei der Berechnung des Wirkungsquerschnittes.

Es gibt mehrere Veröffentlichungen zur Erzeugung schwerer Quarks von H1 und ZEUS. Die Ereignisse mit schweren Quarks werden dabei i.a. über das D^*-Meson identifiziert. Dieses Zerfällt u.a. in dem folgenden Zerfallskanal:

$$D^* \longrightarrow D^0\pi_s \longrightarrow (K\pi)\pi_s$$

der in einem sehr engen kinematischen Bereich abläuft, da die Ruhemasse des D^* nur geringfügig größer ist als die des D^0. Dieser Drei-Teilchen-Zerfall hat sich als relativ gut zu rekonstruieren erwiesen und bietet eine gute Identifikationsmöglichkeit von Ereignissen mit schweren Quarks und ermöglicht außerdem eine nahezu untergrundfreie Analyse. Die aus den Analysen gewonnenen c-Wirkungsquerschnitte liegen etwas über den theoretischen Erwartungen, aber immer noch im Rahmen der meßtechnischen und theoretischen Ungenauigkeiten. Erste Messungen, anhand der Analyse von $p_{T,rel}$-Verteilungen der Myonen aus semimyonischen Zerfällen schwerer Quarks, ergaben signifikant höhere b-Wirkungsquerschnitte als der theoretisch erwartete. So ergab die H1-Analyse einen um einen Faktor 4.6 und die ZEUS-Analyse einen um einen Faktor 1.2-3 größeren b-Wirkungsquerschnitt. Auch die Messung am Tevatron ergab einen um 2.5 größeren b-Wirkungsquerschnitt. Es ist deshalb wichtig, diese Messungen über unabhängige Analysen zu verifizieren.

Ziel dieser Arbeit ist die Untersuchung, wie b- und c-Ereignisse bei Messungen mit dem H1-Detektor unterschieden werden können. Dazu werden u.a. mehrere geeigneten kinematischen-Variablen wie z.B. die rekonstruierte D^0-Zerfallslänge betrachtet. Diese ist aufgrund des zusätzlichen b $\rightarrow$ c Zerfalls im Durchschnitt für b-Ereignisse größer als für c-Ereignisse. Die rekonstruierte D^0-Zerfallslänge kann nur mit den Zusatzinformationen des 1997 eingebauten Central-Silicon-Tracker (CST) gemessen werden, da mit diesen eine Verbesserung der Ortsauflösug der Teilchenspuren erreicht wird. Auch verschiedene Shape-Variablen, die die unterschiedlichen Jetstrukturen von b- und c-Ereignissen beschreiben, bieten gute Aussichten auf eine Unterscheidung der Ereignisse. Neben einzelnen kinematischen Variablen werden auch multidimensionale Analysen betrachtet. Eine solche kann z.B. mit einem Neuronalen Netz durchgeführt werden. Neuronale Netze wurden schon mehrfach mit Erfolg in der Hochenergiephysik eingesetzt und sollen deshalb auch in dieser Studie betrachtet werden.

Kapitel 1

Das H1-Experiment

In diesem Kapitel werden der ep-Speicherring HERA sowie der H1-Detektor beschrieben. Dabei werden insbesondere die Komponenten des Detektors betrachtet, die für die Spurrekonstruktion zuständig sind.

1.1 Der HERA Speicherring

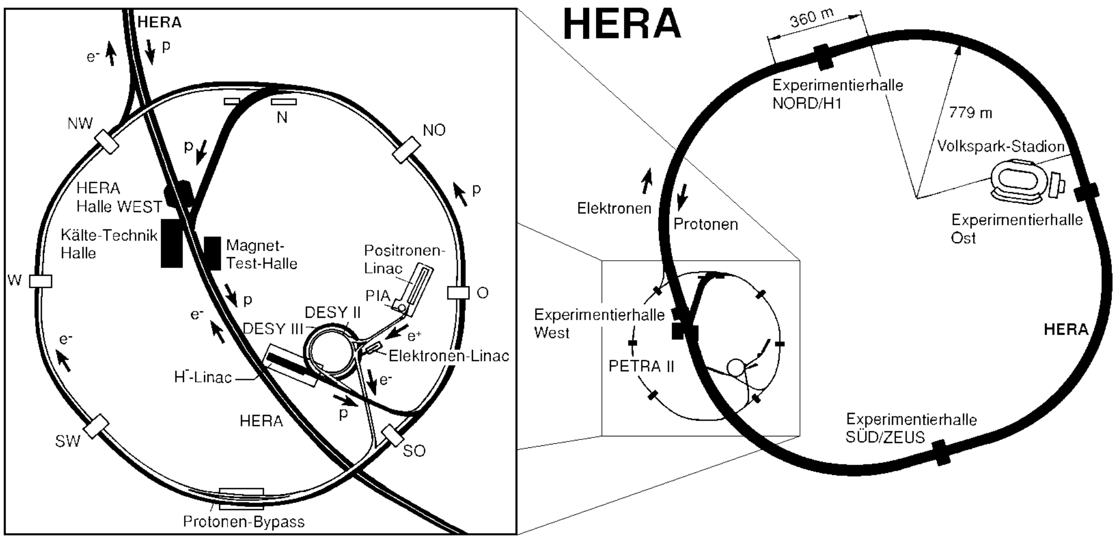

Abbildung 1.1: *Der HERA Speicherring mit seinen Vorbeschleunigern.*

Die Hadron-Elektron-Ring-Anlage (HERA) am Deutschen-Elektronen-Synchrotron (DESY) in Hamburg ist die weltweit einzige Anlage, in der Elektronen[1] und Protonen zur Kollision gebracht werden. Dazu werden Pakete dieser beiden Teilchen in zwei in sich

[1] Von 1995 bis 1997 wurden Positronen statt Elektronen gespeichert. In diesem Kapitel wird der Begriff Elektron synonym für beide Teilchensorten verwendet.

geschlossenen ringförmigen Vakuumröhren, die den unterirdischen 6.3 km langen HERA-Ring bilden, in entgegengesetzter Richtung beschleunigt und an den nördlichen und südlichen Wechselwirkungspunkten aufeinandergeschossen. An diesen Orten befinden sich die Detektoren H1 und ZEUS, während die Strahl-Target-Experimente HERMES und HERA-B in den Hallen Ost und West aufgebaut sind (Abbildung 1.1).

Vor der Einspeisung in den HERA-Ring werden die Elektronen und Protonen in den Linearbeschleunigern LINEAC II und III sowie den dafür umgebauten Speicherringen DESY und PETRA vorbeschleunigt. In den Tabellen 1.1 und 1.2 sind die beiden Parameter Länge und maximale Strahlenergie der einzelnen Beschleuniger angegeben. Im HERA-Ring selber werden die Teilchen in den Hochfrequenzbeschleunigerstrecken, die sich in den geraden Abschnitten des Rings befinden, auf die Energien von 27.5 GeV für die Elektronen und 820 GeV (920 GeV seit 1998) für die Protonen beschleunigt. Dies ergibt eine Schwerpunktsenergie von $\sqrt{s} \approx 300$ GeV bei der Kollision der beiden Teilchenpakete.

Beschleuniger	Länge/Umfang	max. Strahlenergie
LINEAC II	70 m	450 MeV
DESY II	293 m	7.5 GeV
PETRA II	2304 m	14 GeV
HERA-e	6336 m	27.5 GeV

Tabelle 1.1: *Beschleuniger für Elektronen.*

Beschleuniger	Länge/Umfang	max. Strahlenergie
LINEAC III	32 m	50 MeV
DESY III	317 m	7.5 GeV
PETRA II	2304 m	40 GeV
HERA-p	6336 m	820 GeV

Tabelle 1.2: *Beschleuniger für Protonen.*

Um die Teilchen auf der Kreisbahn zu halten, werden sie in den Kurven des HERA-Rings durch senkrecht zu ihrer Flugrichtung angelegte Magnetfelder abgelenkt. Die dafür benötigten hohen Feldstärken werden mit supraleitende Dipolmagneten erzeugt. Diese müssen mit flüssigem Helium auf 4.3 Kelvin gekühlt werden, um ihre supraleitenden Eigenschaften entfalten zu können. Zur Fokussierung der jeweiligen Teilchenstrahlen werden alternierend angeordnete Quadrupolmagnete verwendet.

		Design		1997	
		e	p	e	p
Strahlenergie	[GeV]	30.0	820	27.6	820
mittlerer Strahlstrom	[mA]	58	163	28	74
Anzahl der Teilchenpakete		220	220	175	175
mittlere Luminosität	[$cm^{-2}s^{-1}$]	1.5×10^{31}		4.3×10^{30}	
spezifische Luminosität	[$cm^{-2}s^{-1}mA^{-2}$]	4.0×10^{29}		4.8×10^{29}	
integrierte Luminosität	[$nb^{-1}Jahr^{-1}$]			32400	

Tabelle 1.3: Einige Parameter des HERA-Speicherrings.

In Tabelle 1.3 sind einige der wichtigsten Parameter des HERA-Speicherrings angegeben. Bei einer Anzahl von jeweils 220 Elektronen- und Protonen-Paketen ergibt sich für die Kollisionsrate (BC, bunch crossing) der Pakete eine Frequenz von 10.4 MHz. Dies entspricht einem Abstand von 96 ns.

1.2 Der H1-Detektor

Bei dem H1 Detektor handelt es sich um einen Universal-Teilchendetektor, der dazu dient, die Spuren und Energien der bei der ep-Wechselwirkung enstehenden bzw. gestreuten Teilchen zu messen. Um eine möglichst vollständige Rekonstruktion eines solchen Ereignisses zu gewährleisten, deckt der Detektor mit seinen einzelnen Komponenten nahezu den gesamten Raumwinkelbereich ab (4π-Detektor). Der asymmetrische Aufbau des Detektors berücksichtigt die unterschiedlichen Energien der beiden Teilchenstrahlen, die bei der Kollision zu einem Gesamtimpuls, im Laborsystem, in Richtung des Protons führen. Dies bewirkt, daß in diesem als vorwärts definierten Bereich ein Großteil der bei der Kollision erzeugten Teilchen fliegen. Der H1 Detektor wiegt 2800 Tonnen und besitzt die Ausmaße 12m $\times$ 15m $\times$ 10m (Abbildung 1.2). Im folgenden werden die wichtigsten Detektorkomponenten erläutert. Eine detaillierte Beschreibung des H1 Detektors findet sich bei [H1C96a].

Das H1-Koordinatensystem

Der nominelle Wechselwirkungspunkt bildet den Ursprung des bei H1 verwendeten Koordinatensystems. Die gewählten kartesischen Koordinaten bilden ein Rechtssystem und sind wie folgt definiert:

x : in der HERA-Ringebene, zum Mittelpunkt zeigend
y : senkrecht auf die HERA-Ringebene, nach oben zeigend
z : in Protonrichtung

In vielen Fällen ist es jedoch praktischer leicht modifizierte Kugelkoordinaten zu verwenden. Diese transformieren sich aus den kartesischen Koordinaten wie folgt:

r : radialer Abstand zur z-Achse, $r = \sqrt{x^2 + y^2}$
ϕ : Winkel zur x-Achse in der xy-Ebene, Azimutalwinkel $\phi = \arctan(y/x)$
θ : Winkel zur z-Achse, Polarwinkel $\theta = \arctan(r/z)$

In Abbildung 1.3 wird das H1-Koordinatensystem verdeutlicht.

Die Silizium-Vertex-Detektoren

Der zentrale Silizium-Vertex-Detektor (Central Silicon Tracker, CST) besteht aus 32 Siliziumstreifendetektoren, die in zwei Lagen a zwölf und zwanzig zylinderförmig um die Strahlachse am nominellen Wechselwirkungspunkt angeordnet sind. Es wird also mit dem CST der gesamte ϕ-Bereich sowie der Polarwinkelbereich $30° \leq \theta \leq 150°$ abgedeckt. Aufgrund der hohen Ortsauflösung des CST ist eine wesentlich präzisere Spurmessung und

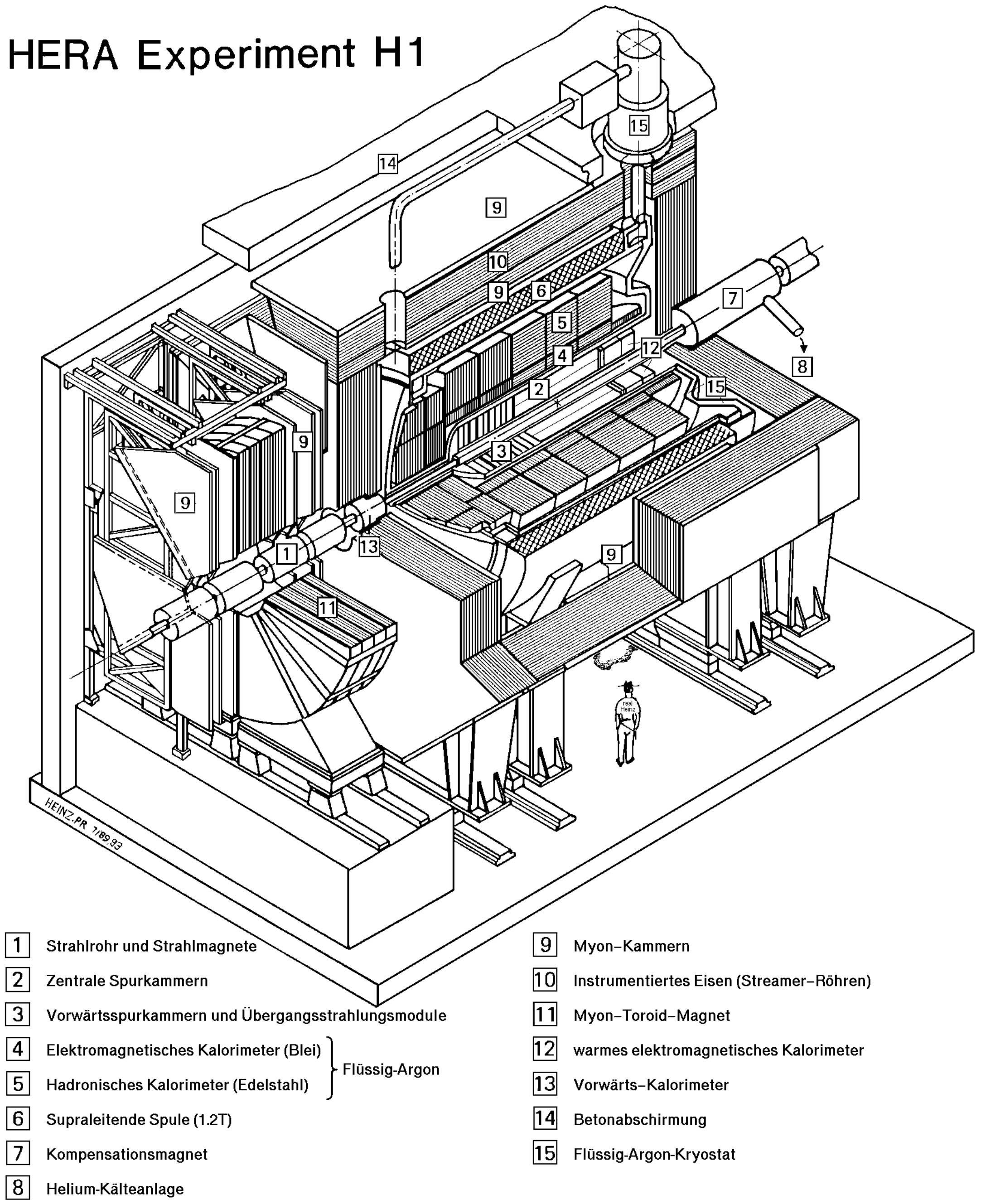

Abbildung 1.2: *Perspektivische Darstellung des H1 Detektors. Die Protonen durchqueren den Detektor von rechts nach links, der Vorwärtsbereich ist also links im Bild dargestellt.*

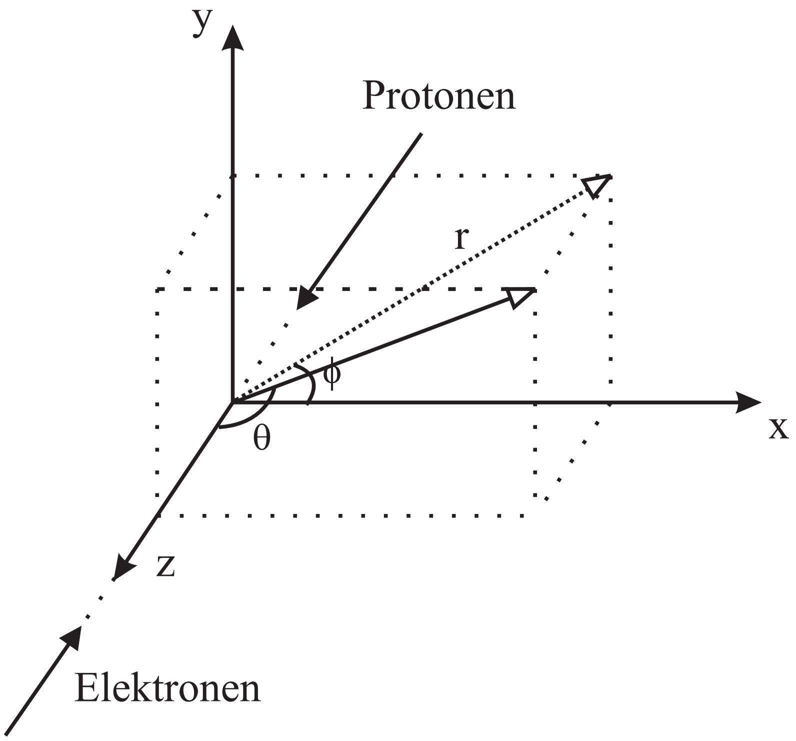

Abbildung 1.3: *Das H1-Koordiantensystem.*

eine daraus resultierende genauere Bestimmung des Primär- und der Sekundärvertices möglich. Er wurde 1996 abschließend eingebaut und erstmals 1997 für die Datennahme verwendet. Der CST ist ausführlich in [Pre96] beschrieben.

In z-Richtung hinter dem CST befindet sich ein weiterer Silizium-Detektor, der BST (Backward Silicon Tracker). Er dient zur Erfassung des gestreuten Elektrons in dem Polarwinkelbereich $162° \leq \theta \leq 176°$.

Die Spurdetektoren

Mit den drei im H1-Detektor eingebauten Spurdetektoren für jeweils den Vorwärts-, Zentral- und Rückwärtsbereich (s. Abb. 1.4) wird nahezu der gesamte Polarwinkelbereich θ abgedeckt. Sie dienen dazu, die Flugbahnen der geladenen Teilchen zu messen.

Der zentrale Spurdetektor besteht aus den folgenden Komponenten:

- Die beiden Spurkammern CJC1 (Central Jet Chamber) und CJC2 sind die wichtigsten Komponenten der Spurrekonstruktion. Mit ihnen werden die Polarwinkelbereiche $11° \leq \theta \leq 169°$ (CJC1) bzw. $26° \leq \theta \leq 154°$ (CJC2) abgedeckt. Mittels der Messung der Driftzeit können die $r\phi$-Koordinaten der geladenen Teilchen mit einer Auflösung von $\sigma_{r\phi} = 150$ μm bestimmt werden. Die z-Koordinaten erhält man durch Ladungsteilung, jedoch nur mit einer Genauigkeit von $\sigma_z = 2.2$ cm. Die wesentlich schlechtere Auflösung in der z-Richtung resultiert aus der Anordnung der Signaldräte der CJC, die parallel zur z-Achse verlaufen.

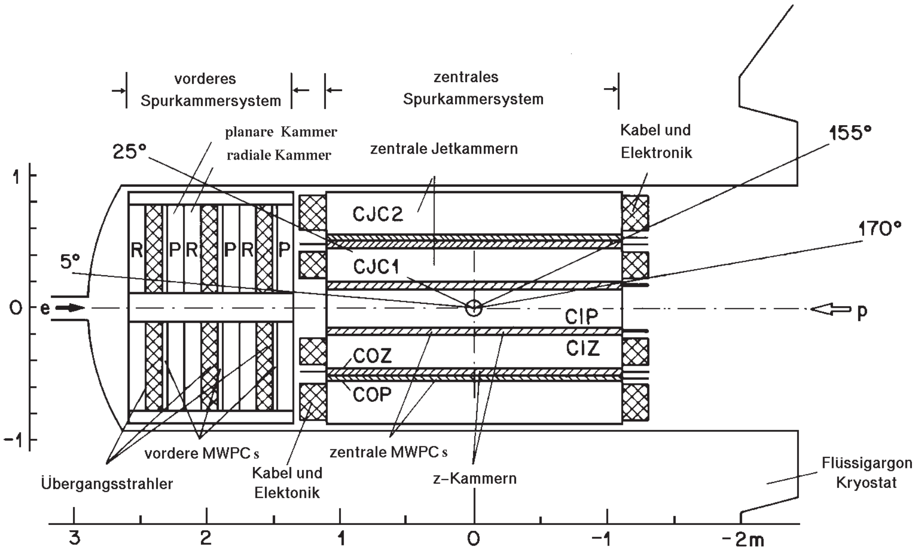

Abbildung 1.4: *Seitenansicht der Spurdetektoren.*

- Die Driftkammern CIZ und COZ (Central Inner bzw. Outer Jet-Chamber) dienen zur Verbesserung der Messung der z-Koordinate. Ihre Signaldräte sind senkrecht zur Strahlachse angeordnet, wodurch eine Verbesserung der z-Auflösung auf $\sigma_z = 350\ \mu$m erreicht wird.

- Die Proportionalkammern CIP und COP (Central Inner bzw. Outer Proportional Chamber) sind ausschließlich für die Eingabe der L1 und L2 Triggersysteme bestimmt.

Mit dem Vorwärts-Spurdetektor (Forward Tracking Detector, FTD), der sich aus drei identischen Modulen aufbaut, wird der Bereich $5° \leq \theta \leq 25°$ abgedeckt. Jedes dieser Module besteht aus je einer dreilagigen planaren Driftkammer, einer Vieldraht-Proportionalkammer, einem Übergangsstrahlungsmodul aus Polypropylen zur Unterscheidung hochenergetischer Pionen und Elektronen und abschließend einer radialen Driftkammer. Die drei planaren Driftkammern bestehen aus jeweils 32 Zellen mit je vier Signaldrähten, die senkrecht zur z-Achse angeordnet sind. Es wird eine Auflösung von $\sigma_{x,y} = 170\ \mu$m erreicht.

Durch das dritte Spurkammer-System, die Rückwärts-Driftkammer (Backward Drift Chamber, BDC), wird der Polarwinkelbereich $151° \leq \theta \leq 177.5°$ abgedeckt. Bestehend aus vier doppellagigen Driftkammern, deren Signaldrähte senkrecht zur z-Achse in oktagonaler Form angeordnet sind, dient sie primär dem Nachweis des gestreuten Elektrons.

Die Kalorimeter

Die Spurdetektoren sind von dem Flüssig-Argon-Kalorimeter (LAr Calorimeter) umgeben. Dieses ist unterteilt in einen elektromagnetischen und hadronischen Teil und dient zur Messung der Energien der in den Zentral- oder Vorwärtsbereich gestreuten Teilchen.

Sein Akzeptanzbereich erstreckt sich von 4° bis 153°. Als Absorbermaterial des elektromagnetischen Kalorimeters wird Blei eingesetzt. Je nach Polarwinkel besitzt es eine Dicke von 20 bis 30 Strahlungslängen. Für das hadronische Kalorimeter mit einer Dicke von 4.5 bis 8 Wechselwirkungslängen wird Edelstahl als Absorbermaterial verwendet. Diese beiden Teile des Flüssig-Argon-Kalorimeter ermöglichen zusätzlich eine Identifizierung von Elektronen und Hadronen.

Für die Erfassung von Hadronen im extremen Vorwärtsbereich 0.6° bis 4° dient das Plug-Kalorimeter, das aus Kupfer Absorbern und Silizium Detektoren besteht.

Der rückwärtige Bereich des H1-Detektors wird von dem SpaCal (Spaghetti Calorimeter) mit einem Akzeptanzbereich von 153° bis 177.8° abgedeckt. Sowohl der elektromagnetische als auch der hadronische Teil des SpaCals verwenden Blei als Absorbermaterial. In dieses sind szintillierende Fasern eingebettet, die mit Photomultiplieren ausgelesen werden.

Der Magnet und das Myonsystem

Eine supraleitende Spule, welche die Spurkammern und das LAr-Kalorimeter umfaßt, erzeugt in diesen Bereichen ein parallel zur Strahlachse in Protonrichtung verlaufendes Magnetfeld, mit einer durchschnittlichen Stärke von 1.15 Tesla. Das Magnetfeld dient zur Ablenkung der geladenen Teilchen. Die Spule besitzt eine Länge von 5.75 m und wird mit flüssigem Helium gekühlt.

Die Spule ist von einem Eisenjoch umgeben, das mit Streamerkammern zur Erfassung von Myonen durchsetzt ist (instrumentiertes Eisen). Zusätzlich befinden sich noch je drei Lagen an Streamerkammern inner- und außerhalb des Eisens. Das Eisenjoch führt den magnetischen Fluß zurück und dient außerdem zur Messung von hadronischen Schauern, die das Kalorimeter gänzlich durchdrungen haben (tail catcher).

Im Vorwärtsbereich befindet sich ein weiterer Myondetektor bestehend aus einem Ringmagneten, der in jeweils drei doppellagige Driftkammern eingebettet ist. Er dient zur Identifizierung von hochenergetischen Myonen in einem Akzeptanzbereich von $3° \leq \theta \leq 17°$.

Das Trigger- und Datennahmesystem

Wie bereits einleitend erwähnt wurde, beträgt die HERA BC-Frequenz 10.4 MHz. Die eigentliche Rate physikalischer Interaktionen liegt jedoch nur bei ungefähr 100 kHz. Diese beinhalten zudem einen großen Teil an Untergrundreaktionen wie etwa Wechselwirkungen der Elektronen oder Protonen mit dem Strahlrohr oder dem Restgas sowie die auf Kosmische- oder Halo-Myonen zurückzuführenden Ereignisse. Zur Trennung dieser Untergrundereignisse von den physikalisch interessanten wird ein mehrstufiges Triggersystem verwendet.

Die erste Triggerstufe (L1, level 1) entscheided, ob ein Ereignis ausgelesen wird. Dazu werden die Informationen, der von den einzelnen Subdetektoren gelieferten 196 Triggerelementen, zu einer logischen Kombination von insgesamt 128 Subtriggern verknüpft. Ist nun eine Subtriggerbedingung erfüllt, wird das L1Keep-Signal gesetzt, und die Daten werden ausgelesen. Diese Vorselektion reduziert die zu erfassende Datenmenge auf ca. 1kHz.

Der L2-Trigger besteht aus zwei parallel laufenden, unabhängigen Systemen, einem auf topologischen Korrelationen und einem auf neuronalen Netzwerken basierenden Trigger. Der L2-Trigger wird durch das L1keep-Signal aktiviert und führt zu einer Datenrate von ungefähr 50 Hz.

Auf der L4 Triggerstufe wird eine vorläufige Ereignisrekonstruktion durchgeführt. Dazu werden verschiedene Filterprogramme verwendet, die weitere Untergrundereignisse verwerfen. Die verbleibende Datenmenge von etwa 10 Hz wird auf Band gespeichert.

Die als L5 bezeichnete offline-Rekonstruktion führt eine umfassende Ereignisrekonstruktion sowie eine Klassifizierung der Ereignisse durch. Alle so rekonstruierten Ereignisse werden auf Band gespeichert. Die klassifizierten Ereignisse werden in einer komprimierten Form, die alle für die weiteren Analysen wichtigen Informationen enthält, auf Festplatte gespeichert.

1.3 Spurmessung

Die für die Spurmessung wichtigste Detektorkomponente ist, wie bereits erwähnt, die zentrale Spurkammer CJC. Mit ihr werden die Richtungen, die Impulse und die Ionisationen dE/dx der geladenen Teilchen gemessen. Mit den Messungen des CST kann die Ortsauflösung der CJC, nicht aber die Impulsauflösung, noch erheblich verbessert werden.

Prinzip der Driftkammer

Durchqueren geladene Teilchen ein Gasgemisch, so ionisieren sie dessen Atome oder Moleküle. Diese Prinzip machen sich Driftkammern zunutze.
Driftkammern bestehen aus mit Drähten durchspannten Gasvolumina. An die Drähte ist eine Spannung angelegt, die elektrische Felder erzeugt. Der Bereich zwischen den Kathoden- und Potentialdrähten wird als Driftbereich bezeichnet. Die einzelnen Drähte sind derart angeordnet, daß innerhalb des Driftbereiches ein homogenes elektrisches Feld herrscht. Dieser Bereich macht den größten Teil der Driftkammer aus. Zwischen den Potentialdrähten befinden sich die Signaldrähte. Dieser Bereich wird als Gasverstärkungsbereich bezeichnet. Das elektrische Feld innerhalb des Gasverstärkungsbereiches ist radialsymmetrisch ($E \propto 1/r$) um die Signaldrähte und wesentlich stärker als im Driftbereich. Der eigentliche Ablauf innerhalb der Driftkammer laßt sich in drei Phasen gliedern:

- Geladene Teilchen ionisieren bei ihrem Durchgang das Gasgemisch und erzeugen freie Elektronen.
- Die so erzeugten freien Elektronen werden zu den Signaldrähten hin beschleunigt. Dabei stoßen sie jedoch immer wieder mit den Gasmolekülen zusammen. Das elektrische Feld und das Gasgemisch sind derart tariert, daß sich im Driftbereich eine konstante Driftgeschwindigkeit einstellt.
- Erreichen die Elektronen den Gasverstärkungsbereich, so werden sie dermaßen stark beschleunigt, daß sie selber anfangen zu ionisieren; es werden somit immer mehr freie Elektronen erzeugt. Diese treffen letztendlich auf einen Signaldraht, der die Ladungsmenge mißt.

Aufbau der CJC

Die CJC besteht aus zwei getrennten hohlzylinderförmigen Driftkammern (CJC1 bzw. CJC2) die parallel zur Strahlachse liegen (s. Abb 1.5). Tabelle 1.4 gibt die Maße der CJC an.

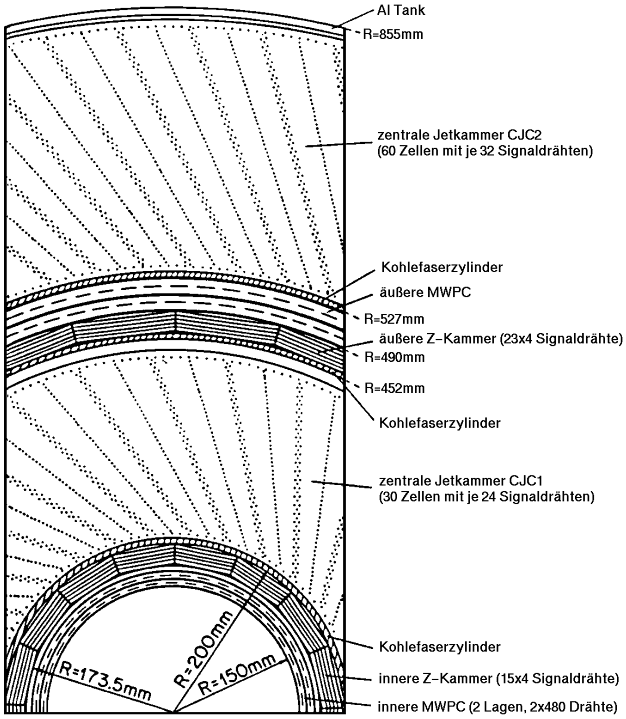

Abbildung 1.5: *Querschnitt der CJC in der $r\phi$-Ebene.*

Die CJC1 ist in 30 Segmente mit je 24 Signaldrähten unterteilt. Die CJC2 hat 60 Zellen mit je 32 Signaldrähten. Die Drähte verlaufen parallel zur z-Achse. Die Signaldrähte sind alternierend um $\pm$ 150 μm aus der nominellen Drahtebene versetzt (Staggering). Diese Anordnung bewirkt, daß die Drähte durch die auftretenden elektromagnetischen Kräfte in die immer gleichen finalen Positionen gedrückt werden, was eine gute Bestimmung der Drahtpositionen ermöglicht.

Maße	CJC1	CJC2
Innenradius r_i [mm]	203	530
Außenradius r_a [mm]	451	844
aktive radiale Länge Δr [mm]	224	296
sensitiver Bereich Δz [mm]	2200	

Tabelle 1.4: *Abmessungen der CJC.*

Die einzelnen Segmente sind um etwa 30° gegenüber der radialen Ausrichtung geneigt. Dies bewirkt eine teilweise Kompensation des Lorentzwinkels, so daß die Driftrichtung annährend senkrecht zur radialen Richtung ist. Dadurch verbessert sich die Auflösung hochenergetischer Teilchen. Außerdem passieren die Teilchen bei diesem Aufbau mindestens eine Signaldrahtebene und durchqueren im Durchschnitt sogar mehrere Zellen, wodurch eine bessere Spurrekonstruktion ermöglicht wird.

Analyse der Signale

In erster Linie mißt die CJC die Driftzeit und die Ladungen Q^+ und Q^- an den beiden Enden des Signaldrahtes. Diese werden kontinuierlich mit einer Frequenz von 104MHz abgelesen.

Aufgrund der annähernd konstanten Driftgeschwindigkeit v_D läßt sich die Driftlänge s wie folgt berechnen:

$$s \approx v_D \cdot (t - t_0) \tag{1.1}$$

dabei ist $t - t_0$ die Zeitdifferenz des Zeitpunktes, an dem das Signal auf dem Draht erzeugt wird, und dem Passieren des Teilchens (BC-Zeit). Die Driftlänge gibt den Abstand der Teilchenspur zum Signaldraht an. Damit kann ein Punkt der Teilchenspur in der $r\phi$-Ebene ermittelt werden (Treffer, Hit). Es muß jedoch beachtet werden, daß man nicht weiß, von welcher Seite die bei der Ionisation erzeugten Elektronen zum Signaldraht gedriftet sind. Es ergeben sich also zwei mögliche Positionen für den Hit. Der richtige dieser beiden Hits läßt sich erst durch die Rekonstruktion der Teilchenspur aus den Messungen mehrerer Signaldrähte ermitteln.

Da der Signaldraht als Ladungsteiler wirkt, läßt sich die z-Koordinate aus dem Verhältnis der an beiden Enden gemessenen Ladungsmenge bestimmen:

$$z \propto \frac{Q^+ - Q^-}{Q^+ + Q^-} \tag{1.2}$$

Desweiteren kann aus der Gesamtladung $Q^+ + Q^-$ der Energieverlust des Teilchens dE/dx berechnet werden. Dieser wird zur Bestimmung der Teilchensorte verwendet.

Der CST

Abbildung 1.6 zeigt die räumliche Anordnung der einzelnen Leitern in den zwei Lagen des CST's. Jede dieser Leitern ist unterteilt in zwei Halbleitern bestehend aus drei Siliziumsensoren und einem Hybrid, auf dem die Ausleseelektronik angebracht ist.

Legt man eine Spannung zwischen der p- und n-Seite der Siliziumsensoren an, so entsteht eine an freien Ladungsträgern verarmte Sperrzone. Geladene Teilchen, die diese Sperrzone durchqueren, erzeugen durch Ionisation Elektron-Loch-Paare. Die Elektronen driften zur n-Seite, die Löcher zur p-Seite, wo die Ladungen gemessen werden.

Auf der p-Seite sind parallel zur z-Achse 1280 Streifen aus mit Bor dotiertem p^+-Implantaten in einem Abstand von 25 μm angebracht. Zur Auslese sind Aluminiumstreifen auf den Implantaten angebracht. Die p-Seite ermöglicht somit die Bestimmung eines Hits in der $r\phi$-Ebene. Die erreichte Auflösung beträgt $\sigma_{r\phi} = 14.4\ \mu$m und ist dadurch in etwa um das Zehnfache genauer als die entsprechende Auflösung eines Hits durch die CJC.

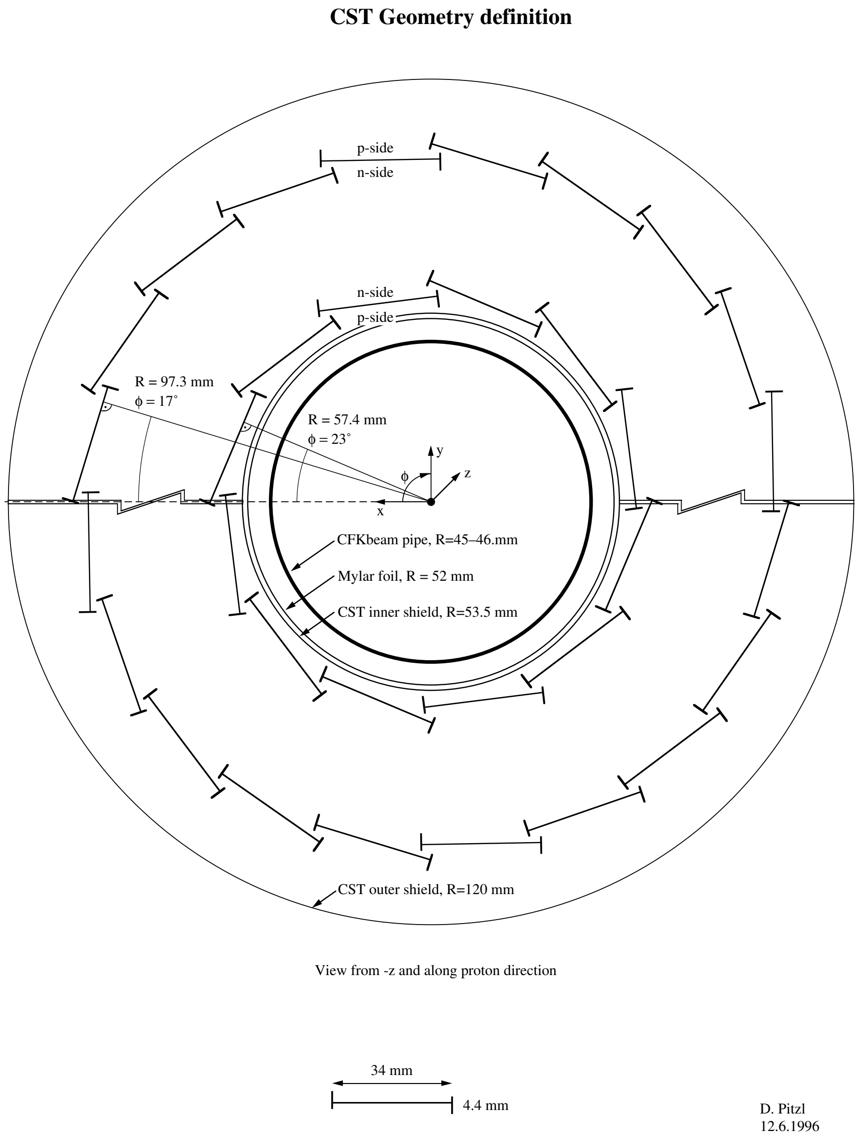

Abbildung 1.6: $r\phi$-*Ansicht des CST.*

Die n-Seite liegt der p-Seite gegenüber. Auf den drei Siliziumsensoren sind 640 mit Phosphor dotierte Streifen (n^+ -Implantate) in einem Abstand von 88 μm senkrecht zur Strahlachse angebracht. Auch hier erfolgt die Auslese über Aluminiumstreifen. Allerdings werden zwei Lagen von Aluminiumstreifen benötigt, die erste Lage zur Auslese der n^+-Streifen und die zweite Lage, die senkrecht zur ersten liegt, um die Signale weiterzuleiten. Die beiden Lagen sind durch eine 5 μm dicke Siliziumdioxidschicht getrennt. Durch diese doppelte Metallisierung erhöht sich jedoch die Kapazität und damit auch das Rauschniveau. Die n-Seite ermöglicht eine Bestimmung der z-Komponente des Hits. Die erreichte Auflösung in $\sigma_z = 25.4$ μm ist gegenüber der CJC-Messung in etwa um einen Faktor 1000 genauer. Abbildung 1.7 zeigt die Anordnungen der Signal- und Auslesetreifen auf den beiden Seiten einer Halbleiter.

Insgesamt besitzt der CST 81920 Kanäle. Die elektrischen Signale der Halbleitern werden für die Weiterleitung in optische konvertiert und über optische Fasern zu der 40 m

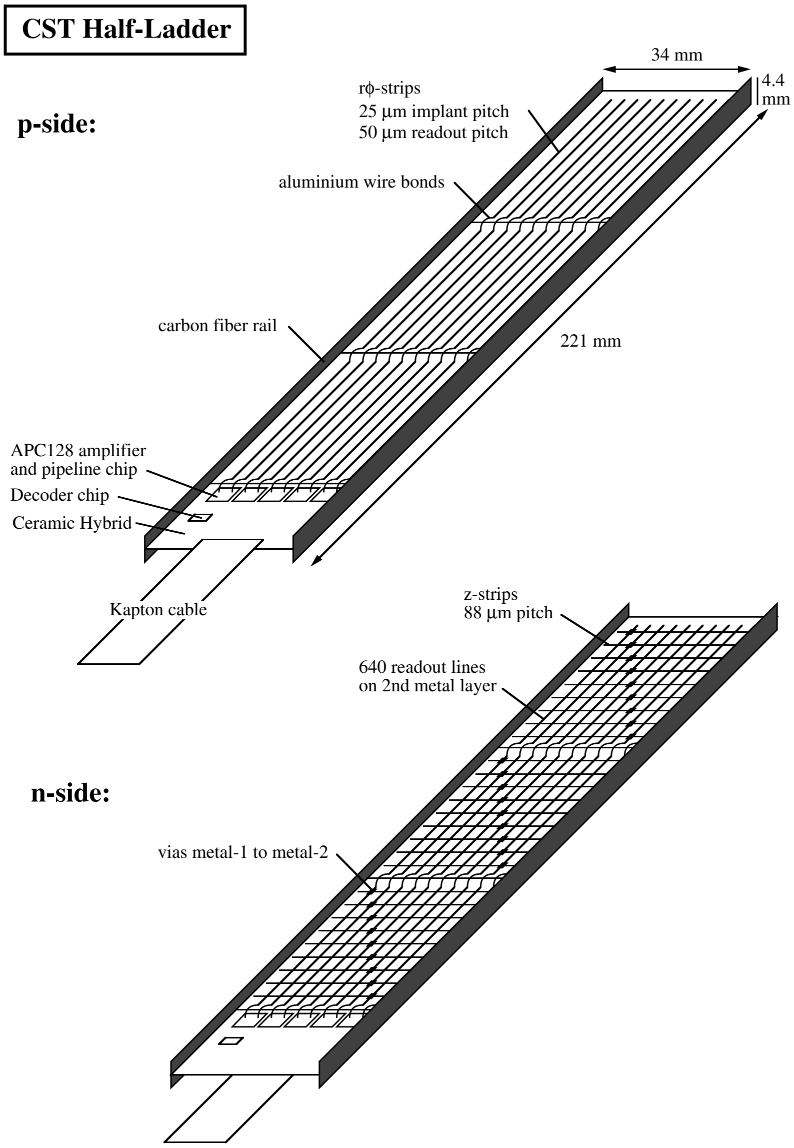

Abbildung 1.7: *Die p- und n-Seite einer Halbleiter.*

entfernten Elektronik übertragen. Dort werden sie von Flash ADCs digitalisiert und von acht Prozessormodulen verarbeitet und gespeichert. Eine ausführliche Beschreibung der Auslese und Steuerung des CST findet sich bei [Kau98].

Kapitel 2

Theoretische Grundlagen

In diesem Kapitel wird ein kurzer Einblick in die Theorie, die den bei HERA beobachteten Prozessen zugrundeliegt, gegeben. Dies geschieht mit einem besonderen Augenmerk auf den wichtigsten Produktionsmechanismus schwerer Quarks, der Boson-Gluon-Fusion, die hauptverantwortlich für die Entstehung der D-Mesonen ist.

2.1 Positron-Proton-Streuung

Die Streuung des Positrons an einem Proton wird durch den Austausch eines Eichbosons der elektroschwachen Wechselwirkung beschrieben. Je nach Ladung des ausgetauschten Teilchens spricht man von einem *neutralen Strom* (neutral current, NC): Austausch eines Photons γ oder eines Z^0-Bosons oder von einem *geladenen Strom* (charged current, CC): Austausch eines $W^{\pm}$-Bosons. Bei letzterem wird das einlaufende Positron in ein Neutrino umgewandelt (s. Abb 2.1).

Die Positron-Proton-Streuung kann zur Auflösung der Struktur des Protons genutzt werden, da das ausgetauschte Boson nicht an das Proton als Ganzes, sondern nur an eines seiner Bestandteile, den sogenannten Partonen koppelt. Bei diesen handelt es sich nach dem *Quark-Parton-Modell* um Quarks und Gluonen. Das Photon koppelt jedoch direkt nur an die Quarks und nicht an die ladungsneutralen Gluonen.

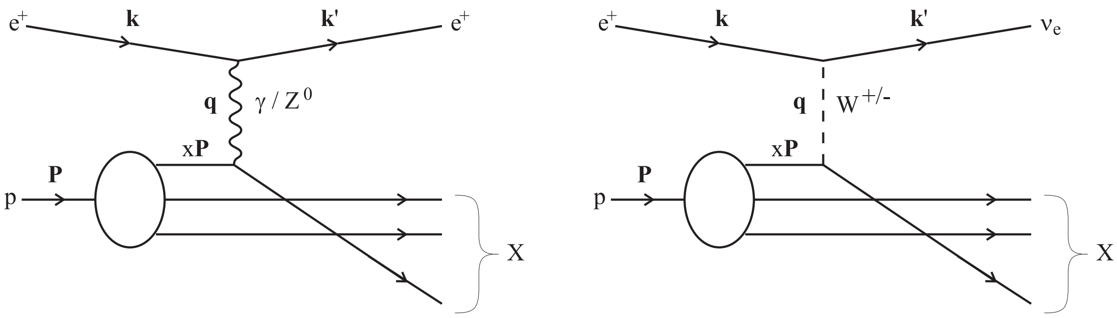

Abbildung 2.1: *Feynman-Diagramme der Positron-Proton-Streuung; links neutraler und rechts geladener Strom*

Wie Abbildung 2.1 zeigt, wird das einlaufende Positron mit dem Vierervektor $\mathbf{k}$ an dem Quark mit dem Anteil x des Gesamtimpulses $\mathbf{P}$ des einlaufenden Protons gestreut. Mit $\mathbf{k}'$ wird der Vierervektor des auslaufenden Leptons und mit X der hadronische Endzustand beschrieben.

Die Schwerpunktsenergie $\sqrt{s}$ ist eine Konstante, die durch die Strahlenergien bestimmt wird:

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

Sie beträgt bei HERA $\sqrt{s} \approx 300$ GeV. Mit ihr und den folgenden kinematischen Variablen läßt sich der gesamte Streuprozeß beschreiben:

$$Q^2 = -\mathbf{q}^2 = -(\mathbf{k} - \mathbf{k}')^2 \tag{2.2}$$

$$x = \frac{Q^2}{2\,\mathbf{P} \cdot \mathbf{k}} \tag{2.3}$$

$$y = \frac{\mathbf{P} \cdot \mathbf{q}}{\mathbf{P} \cdot \mathbf{k}} \tag{2.4}$$

Mit $\mathbf{q}$ wird der Impulsübertrag durch das ausgetauschte Boson beschrieben. Anhand der *Virtualität* Q^2 unterscheidet man zwischen Photoproduktion: $Q^2 \to 0$ und tief unelastischer Streuung (deep inelastic scattering, DIS): $Q^2 \geq 2$ GeV^2.

Die *Bjørkensche Skalenvariable* x gibt den Anteil des Gesamtimpulses $\mathbf{P}$ des Protons an, den das gestreute Parton trägt, während die *Inelastizität* y das Verhältnis zwischen tatsächlichem und maximalem Energieübertrag im Ruhesystem des Protons beschreibt. Sowohl x als auch y sind dimensionslos und liegen im Wertebereich von 0 bis 1.

Unter Vernachlässigung der Teilchenmassen ergibt sich folgender Zusammenhang der oben genannten Variablen:

$$Q^2 \approx s \cdot x \cdot y \tag{2.5}$$

Wie schon erwähnt, kann mit Hilfe der Positron-Proton-Streuung ein Einblick in die Struktur des Protons gewonnen werden. Desweiteren kann man mit ihr sowohl die elektroschwache Wechselwirkung als auch gewisse Aspekte der Quantenchromodynamik (QCD) studieren.

2.2 Photon-Gluon-Fusion

Der bei HERA dominierende Prozeß zur Erzeugung schwerer Quarks[1] ist die Photon-Gluon-Fusion (PGF). Der W oder Z^0 Austausch ist wegen der großen Massen dieser Eichbosonen stark unterdrückt. Wie in Abbildung 2.2 gezeigt wird, wechselwirkt das vom Elektron abgestrahlte Photon, unter Bildung eines Quark-Antiquark-Paares, mit einem Gluon des Protons. Aufgrund der höheren b-Masse sowie der unterschiedlichen Quarkladung werden jedoch wesentlich mehr $c\bar{c}$- als $b\bar{b}$ -Paare erzeugt. Bei der PGF handelt es sich um einen direkten Prozeß in führender Ordnung $\mathcal{O}(\alpha\alpha_s)$, der QCD. Direkte Prozesse sind dadurch gekennzeichnet, daß das Photon als punktförmiges Teilchen elektromagnetisch wechselwirkt. Die Messung des charm(c)- bzw. beauty(b)-Wirkungsquerschnittes bietet

[1] Als schwere Quarks werden in dieser Arbeit b- und c-Quarks bezeichnet.

durch den Vergleich mit den Vorhersagen einen guten experimentellen Test der QCD. Desweiteren läßt sich über diese Messung auch die Gluonendichte des Protons bestimmen. Diese wurde bei HERA unter Verwendung der NLO QCD-Rechnungen aus den Messungen des differentiellen D^*-Wirkungsquerschnittes und den Informationen über das gestreute Elektron berechnet [NPB99]. Das Ergebnis stimmt gut mit dem aus der QCD-Analyse der Proton-Strukturfunktion F_2 überein.

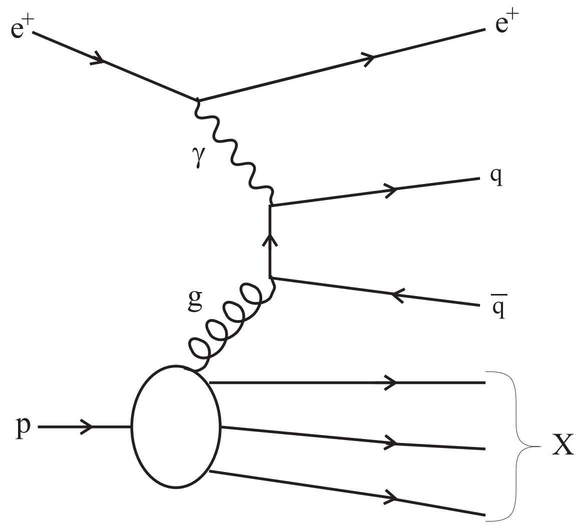

Abbildung 2.2: *Feynman-Diagramm der Photon-Gluon-Fusion*

2.3 Fragmentation schwerer Quarks

Als Fragmentation wird die Umwandlung von Quarks in Hadronen bezeichnet. Zur Beschreibung dieses Vorgangs gibt es verschiedene phänomenologische Modelle (z.B. Field-Feynman-, Lund-Modell), die jedoch nur eine relativ schwache theoretische Rechtfertigung besitzen [Sch95].

Das am weitesten verbreitete Modell ist die Lund-String-Fragmentation: Umso weiter sich das Quark und Antiquark eines erzeugten Paares (z.B. durch die PGF) voneinander entfernen, desto größer wird das Potential zwischen ihnen. Für kleine Abstände ist es, wie das Coulomb-Potential, proportional zu $1/r$. Ähnlich dem elektromagnetischen Feld spannt sich zwischen den beiden Quarks ein Farb-Diplofeld auf. Allerdings verlaufen sämtliche Feldlinien vom Quark zum Antiquark und bilden so einen Flußschlauch (String) von in etwa 1 fm Durchmesser. Übersteigt der Abstand zwischen den beiden Quarks eine Entfernung von ungefähr einem fm, so ist es energetisch günstiger, daß sich ein neues Quark-Antiquark-Paar bildet, welches den Flußschlauch durchtrennt. Nun entsteht ein neuer String zwischen dem ursprünglichen Antiquark und dem Quark des neu entstandenen Paares (analog für die verbleibenden zwei Quarks). Besitzt das neu entstandene Quark-Paar noch genügend Energie, so kann sich dieser Vorgang ein oder sogar mehrere Male wiederholen. Dies ist das Bild, das man sich momentan von der Hadronisierung macht.

Die auf diese Weise gebildeten Hadronen befinden sich in einem relativ kleinen Raumwinkelbereich und werden als Jet bezeichnet. Für die Jetrekonstruktion werden verschiedene Algorithmen verwendet, die auf zwei unterschiedlichen Ideen basieren:

- **Cone-Algorithmus**
  Sucht nach Objekten, die räumlich dicht beieinanderliegen, innerhalb eines Konus (cone).

- **Cluster-Algorithmus**
  Iteratives Anhäufen (clustering) von Objekten aufgrund von Distanzmessungen oder der invarianten Masse z.B. k_t- oder Jade-Algorithmus.

Als Eingabeobjekte werden Kalorimeter- und kombinierte Spur-Kalorimeter-Informationen verwendet. Oftmals sind die Teilchenspuren eines Ereignisses auch räumlich derart gleichmäßig verteilt, daß kein Jet definiert werden kann. Für DIS Analysen wird üblicherweise ein k_t-Algorithmus verwendet.

2.4 Monte-Carlo-Generatoren

Da sich diese Arbeit viel mit Monte-Carlo(MC)-Studien beschäftigt, werden in diesem Abschnitt der dabei verwendete Monte-Carlo-Generator *AROMA* und dessen physikalische Grundlagen kurz vorgestellt.

Monte-Carlo-Generatoren im allgemeinen sind Programme, die physikalische Vorgänge simulieren. Das Bestreben dabei ist, die Realität möglichst genau durch die Simulation wiederzugeben. Dies setzt natürlich ein tiefes Verständnis der zugrundeliegenden Theorien voraus.

Die bei H1 verwendeten MC-Generatoren simulieren die Kinematik eines Ereignisses der ep-Wechselwirkungen bei den bei HERA verwendeten Energien, indem sie zufällige Teilchen in Form von Vierervektoren generieren. Diese simulierten Ereignisse entsprechen den von der Theorie vorhergesagten Wahrscheinlichkeitsdichteverteilungen der ep-Wechselwirkungen.

Neben der korrekten Beschreibung der physikalischen Wechselwirkungen müssen die für H1 generierten MC-Datensätze auch die verschiedenen Detektoreffekte berücksichtigen. Dazu werden sie einer simulierten Messung durch den H1-Detektor unterzogen (*H1SIM*). Anschließend wird eine Ereignisrekonstruktion mit den gleichen Programmpaketen wie für die realen Daten durchgeführt. Die MC- und die gemessenen Datensätze sind somit identisch aufgebaut, und es können dieselben Analyseprogramme für beide verwendet werden. Zusätzlich besitzen die MC-Daten jedoch noch Zeiger (Pointer), die die Verbindung von den generierten zu den rekonstruierten Daten herstellen.

Im Falle des *AROMA*-Generators handelt es sich um eine Simulation der Produktion schwerer Quarks durch die Boson-Gluon-Fusion (BGF):

$$V + g \to Q_f + \bar{Q}_{f'}, \tag{2.6}$$

die ein Unterprozeß, auf Partonen Niveau, der $e^{\pm}p$ Kollision ist:

$$e^{\pm} + p \rightarrow l' + Q_f + \bar{Q}_{f'} + X \tag{2.7}$$

V gibt dabei das austauschende Eichboson, f die Farbe des Quarks, X den hadronischen und l' den leptonischen Endzustand an. Desweiteren wird auch die Masse der schweren Quarks sowie die gesamte Struktur der elektroschwachen Wechselwirkung sowohl bei neutralen- als auch bei geladenen Strömen einschließlich der $\gamma - Z^0$ Interferenz berücksichtigt. Dies geschieht sowohl für den Bereich der Photoproduktion als auch für den der DIS [ARO95].

Desweiteren werden auch die relevanten Korrekturen der Prozesse höherer Ordnung (NLO, next to leading order), die bei den HERA-Energien in der ep-Wechselwirkung stattfinden, berücksichtigt, um eine korrekte Beschreibung der Ereignis-Topologie zu erhalten [MCG99]. Die Gluonbremsstrahlung wird in *AROMA* näherungsweise durch Parton-Schauer-Simulationen berechnet. Ein konstanter Faktor K wurde eingeführt, um die NLO QCD-Korrekturen für die Photoproduktion zu beschreiben, und es werden Abweichungen in der p_t-Balance der produzierten $q\bar{q}$-Paare simuliert.

Weitere Produktionsmechanismen von schweren Quarks, wie CC-Prozesse mit einer Mischung in LO, Streuung an intrinsischem Charm im Proton, Beiträge aus Prozessen mit einem aufgelösten Photon sowie diffraktive Prozesse sind in *AROMA* nicht implementiert, da diese bei den HERA-Energien gegenüber der BGF stark unterdrückt sind.

Kapitel 3

D^*-Rekonstruktion in den H1-Daten

In diesem Kapitel wird die Rekonstruktion des D^*-Zerfalls aus den H1 Daten des Jahres 1997 näher beschrieben. Dieser Zerfall ist gut nachweisbar und die Differenz der D^* und D^0-Masse liefert ein relativ scharfes Signal, das eine gute Identifikation der D-Mesonen ermöglicht.

In Kapitel 4 wird auf die Produktion des D^* eingegangen. Insbesondere werden dabei die D^0-Zerfallslänge sowie einige der bei H1 gebräuchlichen shape Variablen, mit denen sich die Ereignistopologien beschreiben lassen, betrachtet.

Im letzten Kapitel (Kapitel 5) werden dann kurz die Grundlagen von Neuronalen Netzen vorgestellt und das in dieser Studie verwendete Netz beschrieben.

3.1 Der D^*-Zerfall

Das D^*-Meson zerfällt über die starke Wechselwirkung und hat eine entsprechend kurze Lebensdauer. Die Zerfallskanäle dieses Zerfalls werden in Tabelle 3.1 gezeigt. In dieser Studie wird nur der Zerfall des D^* in ein D^0 und ein langsames Pion π_s, $D^{*\pm} \to D^0\pi_s^+$ bzw. $\overline{D^0}\pi_s^-$ (s. Abb. 3.1) betrachtet.

	Zerfallskanal	Anteil (Γ_i/Γ)
Γ_1	$D^0\pi^+$	$(68.3 \pm 1.4)\%$
Γ_2	$D^+\pi^0$	$(30.6 \pm 2.5)\%$
Γ_3	$D^+\gamma$	$(1.1^{+2.1}_{-0.7})\%$

Tabelle 3.1: *Zerfallskanäle des D^{*+} laut [PDG98] (D^{*-}-Zerfallskanäle sind die ladungskonjugierten der oben genannten)*

Das so entstandene D^0 ($\overline{D^0}$) wiederum zerfällt durch die schwache Wechselwirkung (s. Abb. 3.2) mit einer mittleren Zerfallslänge von $c\tau_{D^0} = 124.4\ \mu$m, welche genügend groß ist, um eine Messung der Lebensdauer des D^0 mit dem CST zu ermöglichen. Tabelle 3.2

	Zerfallskanal	Anteil (Γ_i/Γ)
Γ_1	e^+X	$(6.75 \pm 0.29)\%$
Γ_2	μ^+X	$(6.6 \pm 0.8)\%$
Γ_3	K^-X	$(53 \pm 4)\%$
Γ_4	$\overline{K^0}X + K^0X$	$(42 \pm 5)\%$
Γ_{19}	$K^-\pi^+$	$(3.85 \pm 0.09)\%$

Tabelle 3.2: *Die wichtigsten inklusiven Zerfallskanäle des D^0 sowie den in dieser Arbeit betrachtete Zerfallskanal Γ_{19} laut [PDG98] ($\overline{D^0}$-Zerfallskanäle sind die ladungskonjugierten der oben genannten)*

zeigt die wichtigsten Zerfallskanäle des D^0. Auch bei diesem Zerfall wurde wieder nur ein bestimmter Zerfallskanal betrachtet: $D^0 \to K^-\pi^+$.

Der D^*-Zerfall wird also über die beiden Zerfallsreihen

$$D^{*+} \longrightarrow D^0\pi_s^+ \longrightarrow (K^-\pi^+)\pi_s^+$$

$$D^{*-} \longrightarrow \bar{D}^0\pi_s^- \longrightarrow (K^+\pi^-)\pi_s^-$$

bestimmt, die jeweils ein Verzweigungsverhältnis von 2.6% besitzen. Das D^0 wird über seinen Zerfall in ein Kaon und ein Pion rekonstruiert. Beide Zerfallsteilchen sind geladen und können somit direkt in der CJC gemessen werden. Der Zerfall des D^*, über den angegebenen Zerfallskanal, findet in einem sehr engen kinematischen Bereich statt. Die Massendifferenz $\Delta m = m(D^0)\pi_s - m(D^0)$ ist nur geringfügig größer als die Ruhemasse des Pions. Demzufolge besitzt das bei dem D^*-Zerfall erzeugte Pionen einen nur sehr kleinen Impuls[1]. Es ist also zu erwarten, daß sich für die Massendifferenz eine wesentlich bessere Auflösung als für die Signale der Massenverteilungen des D^* oder des D^0 ergibt. Durch einen Schnitt auf die Massendifferenz lassen sich dann die D-Mesonen gut identifizieren.

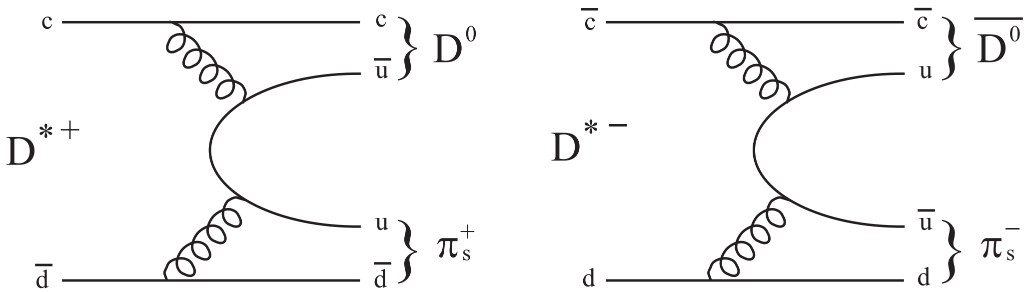

Abbildung 3.1: *D^*-Zerfall im Quarkmodell*

3.2 Der D^*-Algorithmus

Der zur D^*-Suche in dieser Arbeit verwendete Algorithmus rekonstruiert diese aus den meßbaren Spuren des D^*-Zerfalls, nämlich die der geladenen Kaonen und Pionen. In einem ersten Durchlauf bildet der Algorithmus sämtliche Kombinationen zweier Spuren unterschiedlicher Ladung, die jeweils die folgenden Bedingungen erfüllen:

[1] Das entstehende Pion wird aufgrund seines geringen Impulses als langsam bezeichnet und deshalb mit dem Index 's' für slow (Engl. langsam) gekennzeichnet.

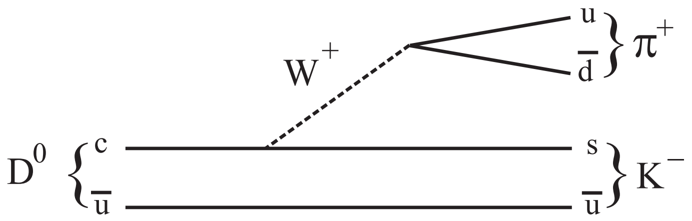

Abbildung 3.2: D^0-*Zerfall im Quarkmodell*

- es handelt sich um eine zentrale Spur
- die Spur kommt vom Primärvertex
- der Transversalimpuls p_T beträgt mehr als 300 MeV/c
- die radiale Spurlänge ist größer als 10cm

Dies dient dazu, aus den gemessenen Spuren eines Ereignisses sämtliche potentiellen Kaon-Pion-Paare zu bestimmen. Es wird dabei die Annahme gemacht, daß es sich bei diesem Paar um das Zerfallsprodukt eines D^0 handelt. Mit den bekannten Massen des Kaons und Pions werden nun die jeweiligen Vierervektoren mit der Formel $E = \sqrt{m^2c^4 + p^2c^2}$ bestimmt und diese anschließend addiert. Aus dem so neu gewonnenen Vierervektor wird die Masse berechnet und überprüft, ob diese in dem Intervall

- $|m_{k\pi} - m_{D^0}| < 100$ MeV/c^2

liegt. Ist dies der Fall, so wird das so rekonstruierte Teilchen als D^0-Kandidat akzeptiert. In einem weiteren Durchlauf sucht der D^*-Algorithmus nach Kandidaten für langsame Pionen π_s, die folgende Bedingungen erfüllen müssen:

- auch hierbei handelt es sich um eine zentrale Spur
- die Spur kommt vom Primärvertex
- der Transversalimpuls p_T beträgt mehr als 120 MeV/c
- die radiale Spurlänge ist größer als 10cm
- Die Ladung des π_s-Kandidaten ist ungleich der Ladung des K -Kandidaten
- der π_s-Kandidaten ist nicht identisch mit dem π-Kandidaten

Jeder so gefundene π_s-Kandidat wird mit den vorher bestimmten D^0-Kandidaten gepaart, und analog zur D^0-Rekonstruktion wird ein D^*-Kandidat aus diesem Paar rekonstruiert, sofern er die vorgegebenen Bedingungen erfüllt:

- $m_{K\pi\pi_s} - m_{K\pi} < 170$ MeV/c^2
- $p_T > 1.5$ GeV/c
- Pseudorapidität $|\eta| < 1.5$

Anstelle des Polarwinkels θ wird auf die dimensionslose, lorentzinvariante Pseudorapidität η des D^*-Kandidaten geschnitten:

$$\eta := \frac{1}{2} \ln \left(\frac{p + p_z}{p - p_z} \right) \tag{3.1}$$

Diese beiden Größen hängen wie folgt zusammen:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right) \tag{3.2}$$

Auf diesem Wege sucht der verwendete Algorithmus ereignisweise nach D^* -Kandidaten, dabei kommt es jedoch auch immer wieder vor, daß Kaonen und Pionen Paare selektiert werden, die nicht von einem D^0-Zerfall herrühren oder daß fälschlicherweise als π_S identifizierte Spuren für die D^*-Rekonstruktion verwendet werden. Dies führt dazu, daß ein großerer kombinatorischer Untergrund in den Daten auftritt. In der weiteren Analyse ist es deshalb wichtig, diesen Untergrund nach Möglichkeit von den wahren D^*-Ereignissen zu trennen. Wie man dies erreichen kann, wird im folgenden gezeigt.

3.3 Das $D^* - D^0$-Massensignal

Für die Analyse werden folgende Index-Files verwendet: `[h1hflav.h1] CDST2.DSTAR_MB.-INDEX, DSTAR_P1.INDEX, DSTAR_P2.INDEX`). Diese Index-Files enthalten die Ereignisse der Daten des Jahres 1997, die den Schnitten in Tabelle 3.3 genügen. Durch diese Vorselektion wird der Umfang der zu prozessierenden Daten erheblich vermindert, ohne dabei jedoch die für die Analyse wichtige Statistik zu verlieren.

Transversalimpuls von K und π	$p_T(K, \pi) > 300$ MeV/c
Transversalimpuls des π_s	$p_T(\pi_s) > 120$ MeV/c
D^0-Masse	1.5146 GeV/$c^2 < m_{K\pi} < 2.0146$ GeV/c^2
Massendifferenz von D^* und D^0	$m_{K\pi\pi_s} - m_{K\pi} < 170$ MeV/c^2
Transversalimpuls des D^*	$p_T(D^*) > 1.5$ GeV/c
Pseudorapidität des D^*	$\lvert\eta(D^*)\rvert < 1.5$

Tabelle 3.3: *Vorselektion der Ereignisse.*

Abbildung 3.3 zeigt die Verteilungen der gemessenen Massendifferenzen vor und nach den angewendeten Analyse-Schnitten. Es ist gut zu erkennen, daß die angewendeten Schnitte das Signal zu Untergrund-Verhältnis S/U wesentlich verbessern. Die einzelnen Schnitte sind in Tabelle 3.4 mit ihren jeweiligen S/U-Werten angegeben. Durch die Schnitte gehen zwar auch Signaldaten verloren, aber es ist immer noch genügend Statistik vorhanden, was sich anhand der hohen Signifikanz von $N/\sigma = 21.5$ klar zeigen läßt. Die einzelnen Schnitte bewirken separat oftmals keine allzugroße Verbesserung des S/U -Verhältnisses; erst im Zusammenspiel mit den anderen Schnitten ergibt sich die doch erhebliche Verbesserung des S/U-Verhältnisses von 0.17 auf 1.5.

Für den Fit wird eine Funktion verwendet, die aus zwei Teilen besteht, einer Gaußfunktion für das Signal und einem Term $a \cdot (\Delta m - m_\pi)^b$ mit variablem Exponenten für den Untergrund. Zur Bestimmung des Untergrundes U wird letzterer Term über den Bereich

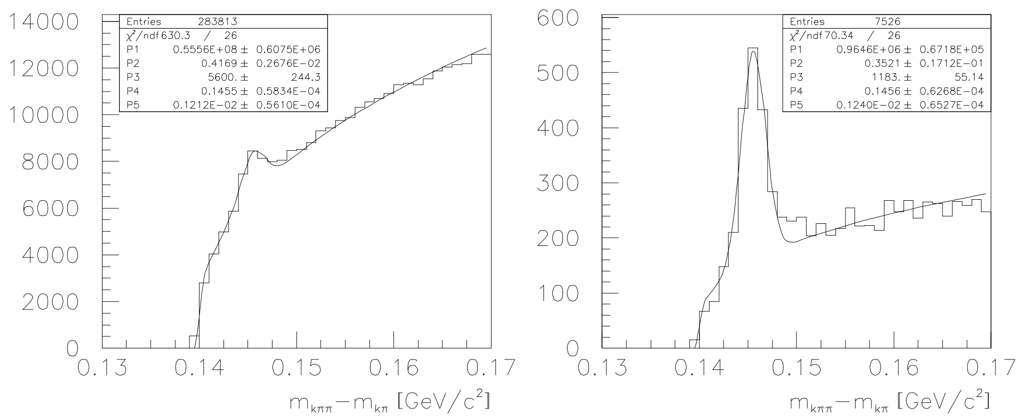

Abbildung 3.3: *Massendifferenz $\Delta m_{D^*D^0}$, links vor und rechts nach den verwendeten Schnitten. Die Fitparameter sind: P1/2 =Faktor/Exponent des Untergrundterms, P3 = Anzahl der Einträge im Signal, P4/5 = Mittelwert und Breite der Gaußkurve.*

der D^*-Signalregion (0.1456 ± 0.0025 GeV/c^2) integriert. Dieser ist so gewählt, daß er in etwa zwei Sigma entspricht. Der Mittelwert der Gaußkurve ist in guter Übereinstimmung mit dem PDG-Wert von $\Delta m_{D^*D^0} = 145.4$ MeV/c^2.

Schnittgröße	Schnittwert	S/U
Keine Schnitte		0.17
rekonstruierte D^0-Masse	$\lvert m_{K\pi} - m_{D^0}\rvert < 40$ MeV/c^2	0.26
Transversalimpuls des D^*	$p_T(D^*) > 2.5$ GeV/c	0.55
radiale Spurlänge des K	$\Delta r > 22$ cm	0.18
radiale Spurlänge des π	$\Delta r > 22$ cm	0.19
normierte K-Testgröße	$LN_K > 0.1$	0.22
normierte π-Testgröße	$LN_\pi > 0.1$	0.19
normierte π_s-Testgröße	$LN_{\pi_s} > 0.1$	0.18
Anzahl der CST-Hits von K und π	CST-Hits(K)$\cdot$ CST-Hits(π) ≥ 2	0.19
Anzahl der Jets	$N_{\text{Jets}} \geq 1$	0.16
Alle Schnitte		1.50

Tabelle 3.4: *Die auf die Daten angewendeten Schnitte, mit den aus den Massendifferenzen bestimmten Verhältnissen von Signal zu Untergrund.*

Die von Jörn Steinhart eingeführte normierte Testgröße LN_i bietet die Möglichkeit, falsche Teilchenhypothesen über einen Vergleich des dE/dx-Wertes eines einzelnen Teilchens mit dem durch die Normfunktion $f_{dE/dx}$ für die entsprechende Massenhypothese gegebenen Erwartungswert zu verwerfen.

Dazu wird als erstes die Likelihood-Wahrscheinlichkeit LH_i für die entsprechende Massenhypothese gebildet [Blo98] und diese dann zur normierten Testgröße LN_i gewichtet. Es werden dabei drei Massenhypothesen berücksichtigt, die des Protons, des Kaons und des

Pions. Die normierte Testgröße LN_i berechnet sich wie folgt:

$$LN_i = \frac{LH_i}{\sum\limits_{j=1}^{3} LH_j} \tag{3.3}$$

Die gewählten Schnitte von $LN_i > 0.1$ bewirken, daß Ereignisse mit offensichtlich falsch identifizierten Kaonen oder Pionen verworfen werden [Ste99].

Die Forderung nach mindestens einem Jet dient nicht der Verbesserung des Verhältnisses von Signal zu Untergrund, sondern ist eine notwendige Bedingung, da im späteren Verlauf der Analyse die Jetdaten benötigt werden.

Kapitel 4

Identifizierung von Charm- und Beauty-Produktion

Im letzten Kapitel (3) wurde erläutert, wieso der D^*-Zerfall für diese Studie verwendet wurde. Im weiteren gilt es nun, aus den selektierten D-Ereignissen denjenigen Anteil zu bestimmen, der aus der Erzeugung von $b\bar{b}$-Paaren herrührt. Dazu werden verschiedene Observablen anhand von MC-Informationen daraufhin getestet, wie gut sie sich zum Trennen von b- und c-Ereignissen eignen.

Da bei den HERA-Energien die Erzeugung von $c\bar{c}$- gegenüber den $b\bar{b}$-Paaren stark dominiert (Faktor 200), entsteht der Großteil der D^* aus der Fragmentation der c-Quarks. Ein wesentlich kleinerer Teil wird über den elektroschwachen Zerfall von B-Mesonen, die durch die Fragmentation der $b\bar{b}$-Paare entstehen, erzeugt. Zur Bestimmung des Beauty-Wirkungsquerschnittes ist es nun wichtig, gerade diesen kleinen Teil möglichst genau zu bestimmen.

Die Ereignistopologien von b- und c-Zerfällen unterscheiden sich in mehreren Aspekten teilweise sehr stark voneinander, wodurch eine Trennung der beiden ermöglicht wird. Die Topologieunterschiede begründen sich aus folgendem:

- Die gemessene Zerfallslänge des D^0 ist für b-Ereignisse größer als für c-Ereignisse. Als Zerfallslänge wird der Abstand des D^0-Zerfallsvertex vom Primärvertex genommen.

- Der Transversalimpuls bezüglich der Jetachse und die Spurmultiplizität sind bei b-Ereignissen größer als bei c-Ereignissen. Dies resultiert aus den unterschiedlichen Massen der beiden Quarks:
  Die Masse des b-Quark $m_b = 4.1 - 4.4$ GeV/c^2 ist nahezu um einen Faktor vier größer als die des c-Quarks $m_c = 1.1 - 1.4$ GeV/c^2 (laut PDG)

Aufgrund der unterschiedlichen Transversalimpulse und Spurmultiplizitäten der b- und c-Ereignisse kommt es zu unterschiedlichen Jetstrukturen. Die Jetstrukturen können über verschiedene Shape-Variablen wie z.B. den Thrust oder die Sphärizität erfaßt werden. Mit Hilfe der Shape-Variablen und den oben genannten Größen ist es möglich, die b- und c-Ereignisse zu trennen.

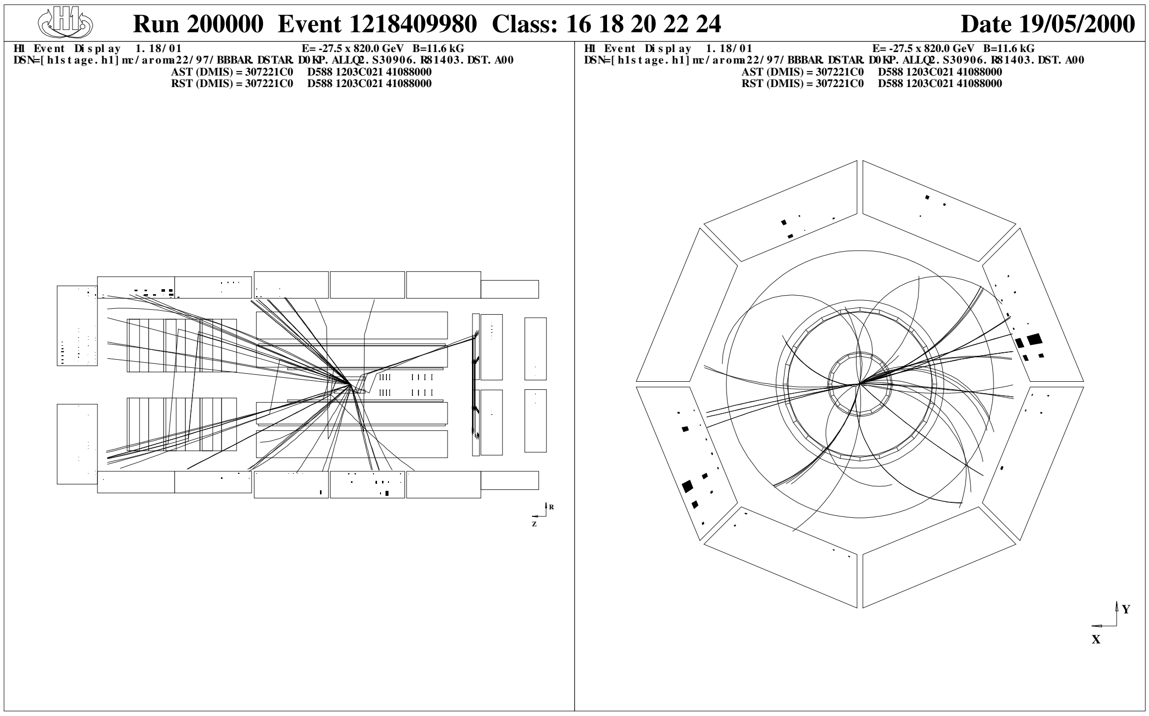

Abbildung 4.1: *Eventdisplay-Ansicht eines MC simulierten b-Ereignisses. Links die Seitenansicht des H1-Detektors (CST, Zentral- und Vorwärtsspurkammersystem, elektromagnetische Kalorimeter), rechts $r\phi$-Ansicht.*

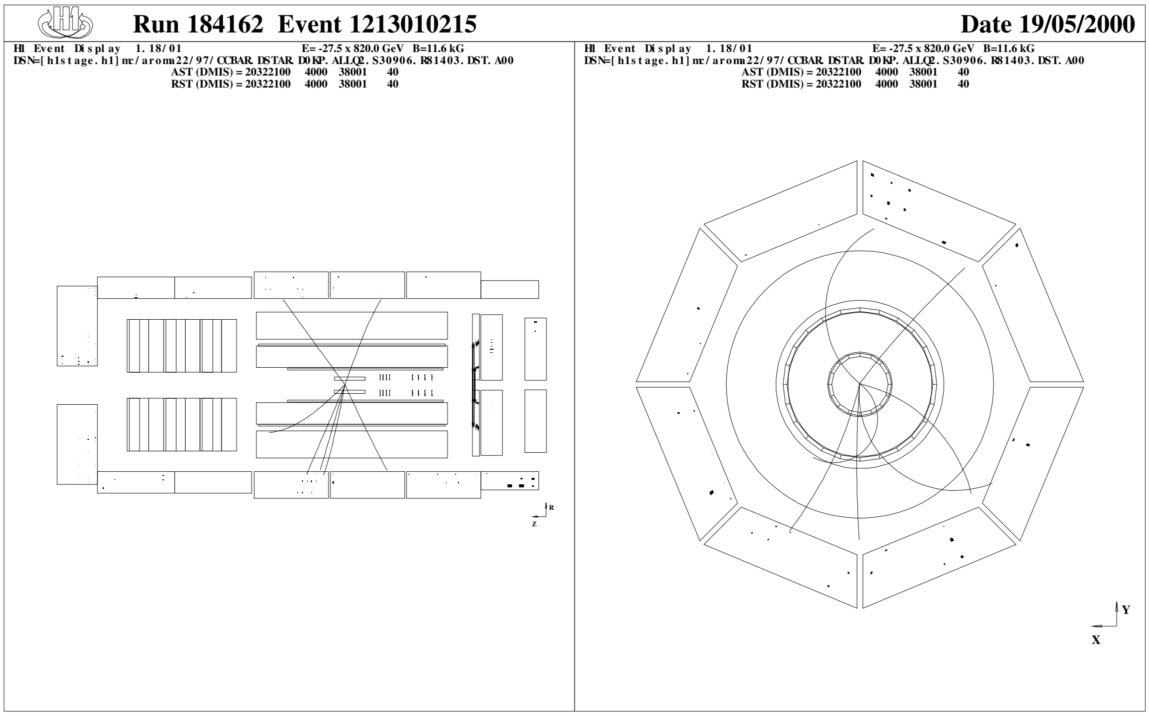

Abbildung 4.2: *Eventdisplay-Ansicht eines MC simulierten c-Ereignisses.*

Abbildungen 4.1 und 4.2 zeigen die Seiten- und die $r\phi$-Ansicht eines simulierten b- bzw c-Ereignisses. Es wurden Ereignisse ausgewählt, die insbesondere die unterschiedlichen Spurmultiplizitäten verdeutlichen. Im Eventdisplay gezeigt ist der CST, das Zentrale- und Vorwärts-Spurkammersystem sowie das elektromagnetische Kalorimeter. In der $r\phi$-Ansicht ist gut zu erkennen, daß die beiden Jets diametral auseinanderfliegen. Tabelle 4.1 gibt einige Werte der gezeigten Ereignisse an.

	sim. b-Ereignis	sim. c-Ereignis
Anzahl der Spuren	16	6
Transversalimpuls des D^0 [GeV/c]	10.1	2.5
Spherizität	0.11	0.02
Aplanarität	0.006	0.010
Masse des Jets [GeV/c^2]	6.1	0.6

Tabelle 4.1: *Einige Werte der simulierten b- und c-Ereignisse.*

Die Größen Spherizität und Aplanarität werden in Kapitel 4.2 erklärt.

4.1 Die D^0-Zerfallslänge

Die mittlere Lebensdauer des D^0 beträgt laut PDG $\tau_{D^0} = 0.415$ ps. Damit ergibt sich ein Wert von $c\tau_{D^0} = 124.4\ \mu$m. Die dreidimensionale sowie die radiale Zerfallslänge l bzw. r sind durch folgende Gleichungen gegeben:

$$l = \beta\gamma c t^* = \frac{p}{m} t^* \tag{4.1}$$

$$r = \frac{p_T}{m} t^* \tag{4.2}$$

mit

$$\beta = \frac{v}{c} \quad \text{und} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{4.3}$$

m, v, p und p_T bezeichnen die Masse, Geschwindigkeit, Impuls bzw. den Transversalimpuls des D^0; t^* entspricht der Lebensdauer des D^0 im Ruhesystem[1].

Zur Bestimmung der Zerfallslänge des D^0 wird die Distanz zwischen dem Primärvertex PV und dem Sekundärvertex SV des D^0-Zerfalls gebildet. Als SV wird der Schnittpunkt der Teilchenspuren des Kaons und des Pions berechnet. Dazu werden die Spuren der CSKR-Bank verwendet, welche die nicht Vertex-gefitteten CST-Spuren enthält. Nur mit Zuhilfenahme der CST ist es möglich, eine genügend hohe Auflösung der Spuren zu erreichen. Die Messungen der CJC alleine sind zu ungenau, um die Zerfallslänge des D^0 zu bestimmen.

Es werden zwei unterschiedlich bestimmte Primärvertices verwendet, der Run-Vertex und der CSPRIM-Vertex (CST Primär).

[1] Größen im Ruhesystem sind durch einen * gekennzeichnet

- Der Run-Vertex ist ein aus vielen Ereignissen eines Runs gemittelter Vertex. Die x- und y-Koordinaten dieses Vertex werden aus den CJC-Spuren eines Runs rekonstruiert. Die Auflösung des Run-Vertex ist durch die Ausdehnung des Strahlflecks, der bunch-crossing-Fläche, gegeben. Sie hat Standardabweichungen von $\sigma_x \approx 155$ μm und $\sigma_y \approx 30$ μm. Die Asymmetrie des Strahlflecks resultiert aus den HERA-Strahlgeometrien. Die z-Koordinate des Run-Vertex wird für jedes Ereignis bestimmt. Die Auflösung in z ist jedoch wesentlich ungenauer als in x und y. Der Primärvertex sowie die Sekundärvertices werden in der DVER Bank gespeichert.

- Bei dem CSPRIM-Vertex handelt es sich um einen ereignisweise bestimmten Vertex. Für jedes Ereignis wird ein PV aus dem Run-Vertex sowie aus den Spurmessungen der CST und CJC rekonstruiert. Es ergeben sich für den CSPRIM-Vertex Auflösungen von $\sigma_x \approx 110$ μm und $\sigma_y \approx 30$ μm, d.h. die Auflösung der x-Koordinate kann verbessert werden. Der Fehler in z beträgt für den CSPRIM-Vertex $\sigma_{z,CSPRIM} = 0.4$ mm und für den CJC-Primär-Vertex $\sigma_{z,CJC} = 1.4$ mm, wie eine uds-MC Studie ergab [Kuh99].

Für die Messung der D^0-Zerfallslänge wird der CSPRIM-Vertex verwendet.

Sämtliche der Mutterteilchen des D^0, wie z.B. das D^*, die über die starke Wechselwirkung zerfallen, besitzen eine derartig kurze Lebensdauer, daß sie keinen meßbaren Beitrag zur Zerfallslänge liefern. Berücksichtigt werden muß jedoch der schwache Zerfall der B-Mesonen. Die PDG gibt eine mittlere Lebensdauer für das B^0 von $\tau_{B^0} = 1.56$ ps und für das B^+ von $\tau_{B^+} = 1.65$ ps an. Daraus ergeben sich die Werte $c\tau_{B^0} = 467.7$ μm bzw. $c\tau_{B^+} = 494.7$ μm. Es ist also zu erwarten, daß bei b-Ereignissen die gemessene D^0-Zerfallslänge größer ist als bei c-Ereignissen, da bei diesen die Zerfallslänge des B-Mesons mit in die Messung der Zerfallslänge des D^0 eingeht. Es wird also im Fall eines b-Ereignisses die Zerfallslänge des D^0 plus die des B-Mesons gemessen.

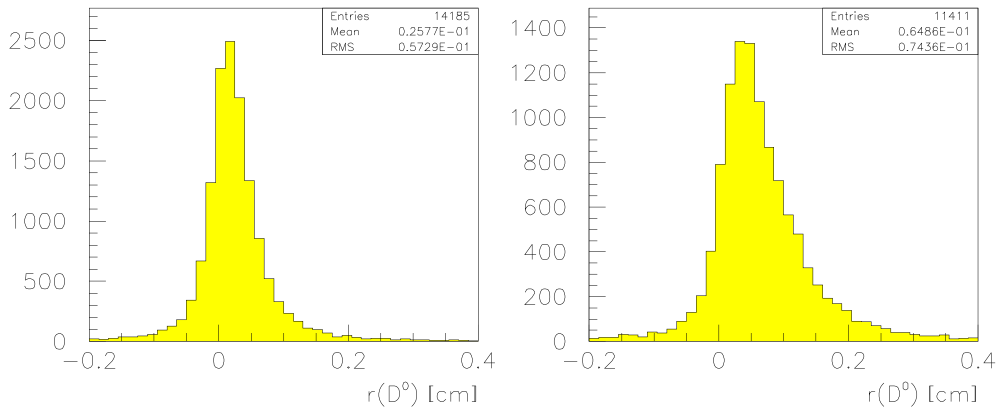

Abbildung 4.3: *Die rekonstruierte radiale Zerfallslänge des D^0 projiziert auf die Impulsrichtung. Links c-Monte-Carlo, rechts b-Monte-Carlo.*

Abbildung 4.3 zeigt deutlich, daß die rekonstruierte radiale Zerfallslänge des D^0 bei b-Ereignissen größer ist als bei c-Ereignissen. Es wird die radiale Zerfallslänge projiziert auf die Impulsrichtung des D^0 verwendet, da dadurch einige der systematischen Meßfehler eliminiert werden. Die negativen Zerfallslängen ergeben sich durch falsch rekonstruierte Vertices. Dies lässt sich kaum vermeiden, da die Ausdehnung des Strahlflecks von der Größenordnung der D^0 Zerfallslänge ist.

Für die Zerfallslängenverteilung sowie für die weitere Analyse wurden die 1997'er Monte-Carlo b- und c-AROMA2.2-Files `BBBAR.DSTAR.D0KP.ALLQ2.S30906.R81403.DST` bzw. `CCBAR.DSTAR.D0KP.ALLQ2.S30906.R81403.DST` verwendet.

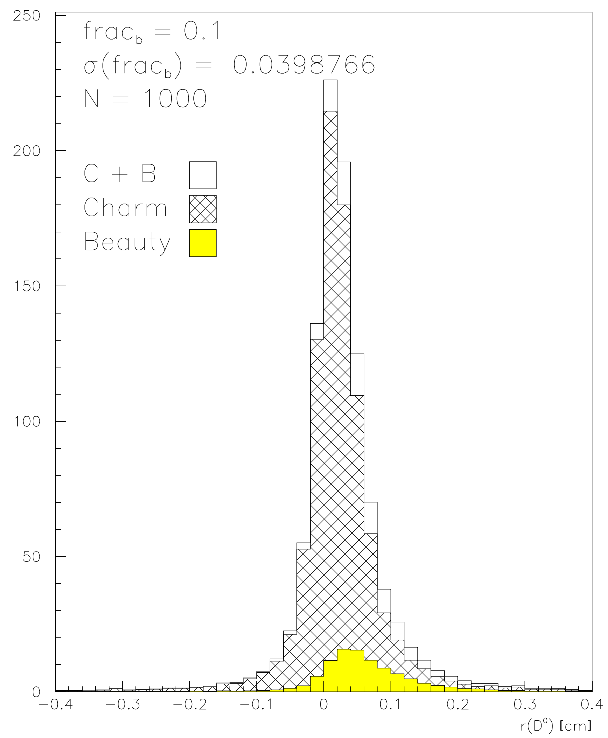

Abbildung 4.4: *Abschätzung des Fehlers für die Bestimmung des b-Anteils über die radialen Zerfallslänge des D^0 projiziert auf den Impuls.*

Für den Vergleich einzelner Observablen ist es wichtig, ein Maß zu definieren, welches angibt, wie gut sich die b- und c-Anteile in den Verteilungen der jeweiligen Observablen trennen lassen. Um dies zu erhalten, wird als Methode ein Maximum-Likelihood-Fit (ML-Fit) verwendet. Mit einen Maximum-Likelihood-Fit lassen sich einzelne Parameter durch eine Anpassung von Referenzverteilungen an die Daten bestimmen [Blo98]. Als zu betrachtender Parameter wird die Anzahl der b-Ereignisse $frac_b$ gewählt. Für diese gilt $frac_b + frac_c = 1$; $frac_c$ gibt natürlich die Anzahl der c-Ereignisse an. Als Referenzverteilungen werden die MC-Verteilungen der jeweiligen Observable verwendet. Diese müssen natürlich entsprechend normiert werden. Es wird jedoch nicht der Maximum-Likelihood-Fit durchgeführt, sondern der statistische Fehler des Fit-Parameters berechnet, was im folgenden als ML-Fehlerabschätzung bezeichnet wird. Dies ist bei einem vorgegebenen Wert des Parameters mit Hilfe der beiden Referenzverteilungen ohne eigentliche Ausführung des Fits möglich (s. Anhang A.1).

Abbildung 4.4 zeigt die mit dieser Methode erhaltenen Referenzverteilungen der rekonstruierten D^0-Zerfallslänge, sowie die über die ML-Fehlerabschätzung bestimmte Standardabweichung des b-Anteils. Diese ergibt sich bei einem geschätzten b-Anteil von $frac_b$ = 0.1 bei 1000 Ereignissen zu $\sigma(frac_b) = 0.040$ mit einer daraus resultierenden Signifikanz (Meßwert/Standardabweichung) von 2.5. Die MC-Referenzverteilungen der rekonstruierten D^0-Zerfallslänge sind also unterschiedlich genug, um eine Trennung der b- und c-Ereignisse mittels dieser Observablen zu ermöglichen.

4.2 Shape-Variablen

In diesem Unterkapitel werden einige topologische Größen betrachtet, die für die Selektion der b-Ereignisse geeignet sind und in der Analyse verwendet werden. Analog zur rekonstruierten D^0-Zerfallslänge wird für jede dieser Größen die ML-Fehlerabschätzung durchgeführt, um ein Maß für die Trennung der b- und c-Ereignisse mit der jeweiligen Observablen zu erhalten.

Spurmultiplizität

Eine weitere Größe, die zur b-Selektion herangezogen werden kann ist die Spurmultiplizität, die Anzahl der geladenen Teilchen pro Ereignis. Aufgrund der höheren Masse des b-Quarks sowie dem b $\to$ c Zerfall entstehen bei einem b-Ereignis mehr Teilchen als bei einem c-Ereignis. Abbildung 4.5 zeigt die Verteilungen der Anzahl der Spuren für b- und c-MC.

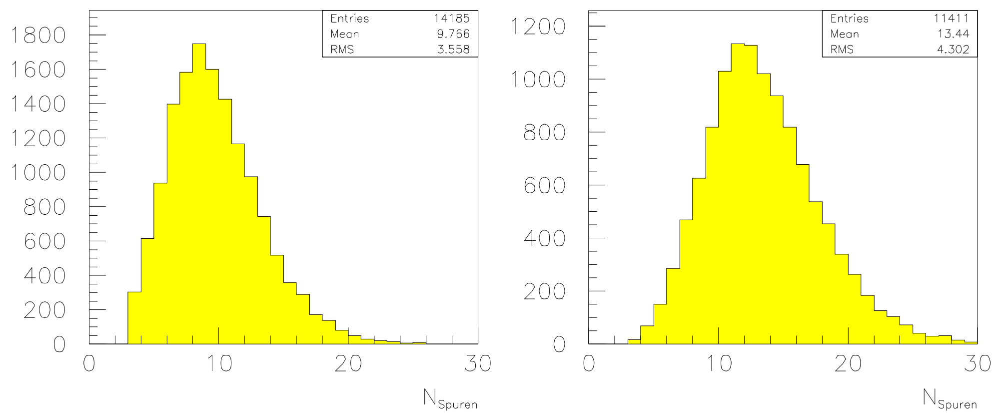

Abbildung 4.5: *Anzahl der Spuren pro Ereignis. Links c-Monte-Carlo, rechts b-Monte-Carlo.*

Es werden nur die in der Zentralen-Spurkammer registrierten Teilchen gezählt, die neutralen Teilchen werden nicht berücksichtigt. Größtenteils werden die gemessenen Spuren durch die langlebigen Pionen und in etwas geringerem Maß durch Kaonen, Myonen und

Elektronen erzeugt. Abbildung 4.6 zeigt die angepaßten MC-Referenzverteilungen. Die ML-Fehlerabschätzung ergibt: $\sigma(frac_b) = 0.031$. Die Spurmultiplizität ist also noch besser zur Trennung der b- und c-Eregnisse geeignet als die rekonstruierte D^0 Zerfallslänge.

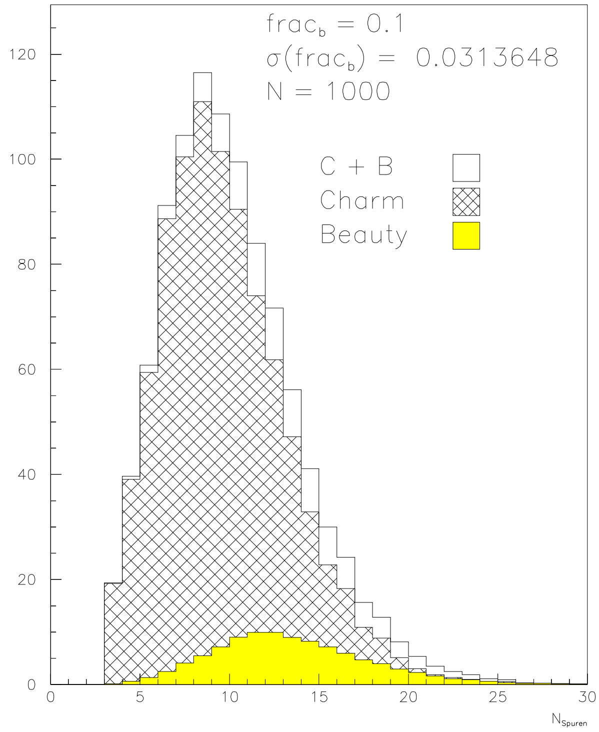

Abbildung 4.6: *Abschätzung des Fehlers für die Bestimmung des b-Anteils über die Anzahl der Spuren pro Ereignis.*

Transversalimpuls des D^0 relativ zur Thrustrichtung

Mit $p_{T,rel}$ wird der Anteil des Impulses des D^0 bezeichnet, der senkrecht zur Jetrichtung steht (Abb. 4.7). Die Jetrichtung ist eine Näherung der ursprüngliche Flugrichtung des schweren Quarks. Bei b-Ereignissen kommt es zu höheren $p_{T,rel}$-Werten als bei c-Ereignissen, da durch den b $\rightarrow$ c Zerfall das so entstandene D^* zusätzlich zur Fragmentation von der ursprünglichen Flugrichtung des b-Quarks abgelenkt wird (Abb. 4.8).

Zur Bestimmung der Jetrichtung wird der Thrust verwendet. Dieser ist wie folgt definiert:

$$T = \max_{\mathbf{n}} T(\mathbf{n}) \tag{4.4}$$

$$T(\mathbf{n}) = 2 \cdot \frac{\sum_i \Theta(\mathbf{p_i} \cdot \mathbf{n})}{\sum_i |\mathbf{p_i}|} \tag{4.5}$$

Θ bezeichnet dabei die Stufenfunktion

$$\Theta(x) = \begin{cases} 1 & : \quad x > 0 \\ 0 & : \quad x \leq 0 \end{cases} \tag{4.6}$$

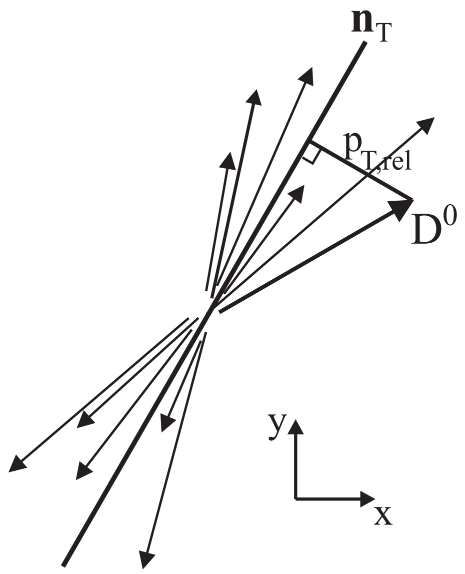

Abbildung 4.7: *Skizze eines Ereignisses in der $r\phi$-Ebene. Eingezeichnet sind die Thrustachse $\mathbf{n}_T$, die D^0 Spur und das dazugehörige $p_{T,rel}$.*

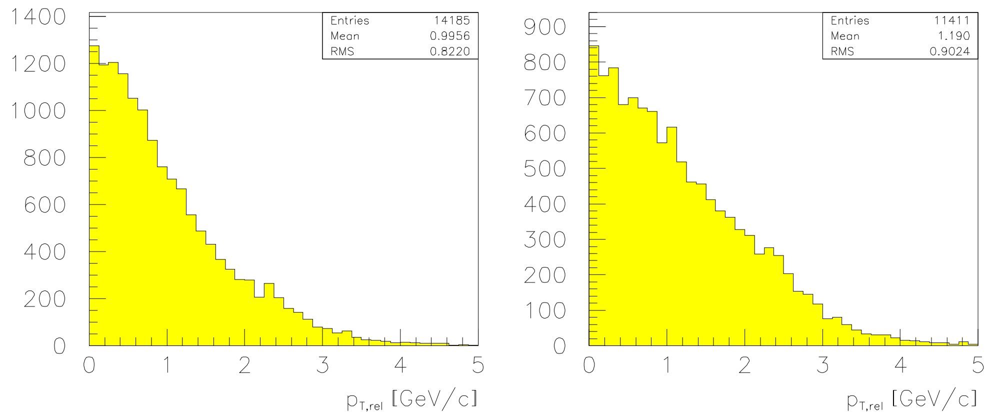

Abbildung 4.8: *Transversalimpuls des D^0 relativ zur Thrustrichtung. Links c-Monte-Carlo, rechts b-Monte-Carlo.*

Derjenige Vektor $\mathbf{n}_T$, der (4.4) erfüllt, wird als Thrustachse bezeichnet [Bra79].

Der Thrust ermöglicht es, einzelne Teilchen mit einem Jet zu assoziieren. Der Thrustwert liegt in dem Intervall von 0.5 bis 1.0 und liefert ein Maß für die Jetartigkeit eines Ereignisses. Bei einem Thrustwert von 1.0 sind alle Teilchenspuren kollinear, während sie bei einem Wert von 0.5 sphärisch verteilt sind und man eigentlich nicht mehr von einem Jet sprechen kann. Mit dem Vektor $\mathbf{n}_T$ wird die Flugrichtung des Jets beschrieben.

Die Größe Thrust wurde zuerst bei e^+e^--Speicheringexperimenten eingeführt, bei denen die Teilchenstrahlen auf gleiche Energien beschleunigt wurden. Das Schwerpunktsystem dieser Experimente entspricht also dem Laborsystem. Demzufolge fliegen die bei diesen Wechselwirkungen entstandenen zwei Jets aufgrund der Impulserhaltung diametral auseinander. Bei den HERA ep-Wechselwirkungen ist dies jedoch nicht mehr der Fall. Das Proton besitzt im Laborsystem einen wesentlich höheren Impuls als das Elektron was dazu führt, daß die Jets bevorzugt in Protonrichtung fliegen (Boost in Protonrichtung).

In der Analyse wird allerdings nur der Thrust in der $r\phi$-Ebene betrachtet, da die Auflösung der Spuren in z zu ungenau ist. In der $r\phi$-Ebene macht sich der Boost nicht bemerkbar und die beiden aus den schweren Quarks resultierenden Jets liegen in guter Näherung auf einer Achse. Diese wird durch die Thrustachse beschreiben.

Die ML-Fehlerabschätzung ergibt $\sigma(frac_b) = 0.123$, dieser Fehler ist wesentlich größer als die Standardabweichung der bisher betrachteten Größen. Die beiden MC-Verteilungen sind also nicht unterschiedlich genug um die b- und c-Ereignisse alleine mit dieser Größe zu trennen. In einer Korrelation unterschiedlicher Observablen kann sie aber dennoch von Nutzen sein und wird deshalb auch in der späteren Analyse mit dem Neuronalen Netz verwendet.

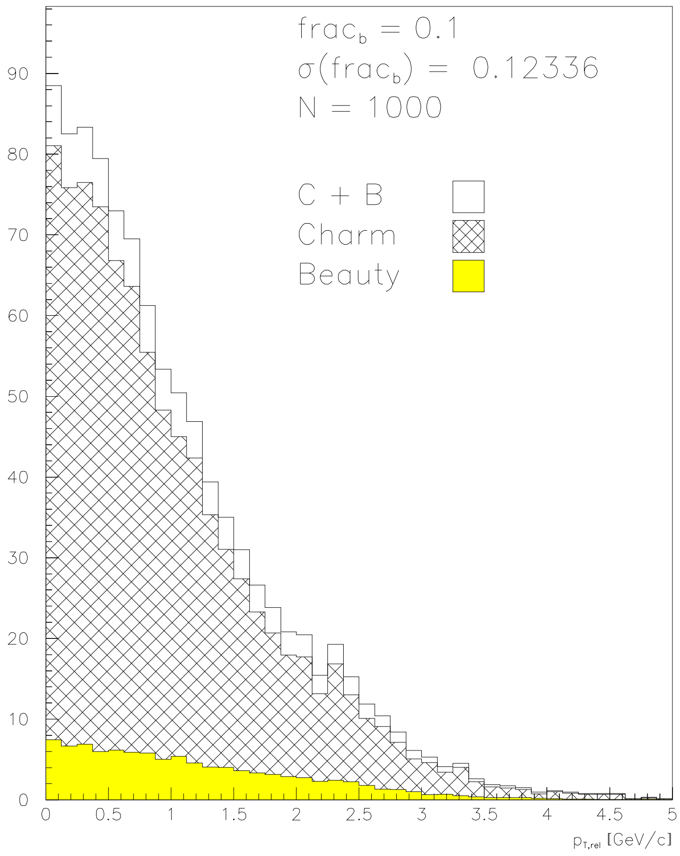

Abbildung 4.9: *Abschätzung des Fehlers für die Bestimmung des b-Anteils mit den $p_{T,rel}$ Verteilungen.*

Sphärizität, Aplanarität, Jetmasse

Die gesamte topologische Struktur eines Ereignisses mit N Teilchenspuren läßt sich im Impulsraum durch den Impulstensor $\hat{P}_{\alpha\beta}(\alpha, \beta \in x, y, z)$ beschreiben:

$$\hat{P}_{\alpha\beta} = \frac{\sum_{i=1}^{N} p_{i,\alpha} \cdot p_{i,\beta}}{\sum_{i=1}^{N} |p_i^2|} \tag{4.7}$$

Die p_i mit $i = 1, ..., N$ geben die Impulse der einzelnen Teilchen an. Durch eine Hauptachsentransformation ergeben sich die drei Eigenwerte Q_1, Q_2, Q_3 des Impulstensors mit den dazugehörigen Eigenvektoren $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$. Gilt nun $Q_1 > Q_2 > Q_3$ und $Q_1 + Q_2 + Q_3 = 1$, so lassen sich die folgenden drei Größen definieren:

- Aplanarität $Apl = \frac{3}{2} \cdot Q_3$
- Planarität $Pl = Q_3/Q_2$
- Sphärizität $S = \frac{3}{2}(1 - Q_1)$

Jede dieser drei Variablen ermöglicht es, die räumliche Verteilung der Spuren eines Ereignisses zu erfassen und dessen Jetartigkeit zu bestimmen. Abbildung 4.10 verdeutlicht die Definition des Impulstensors.

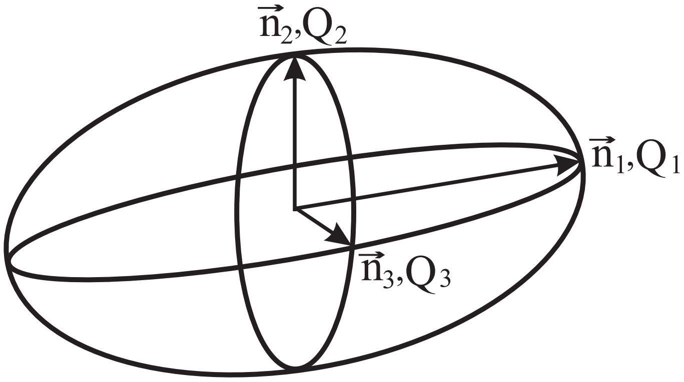

Abbildung 4.10: *Definition des Impulstensors. Die Längen der Halbachsen des Ellipsoid entsprechen den Eigenwerten.*

Die Jetmasse wird über die Addition der einzelnen Vierervektoren der Teilchenspuren eines Jets bestimmt. Für die Jetrekonstruktion wird ein inklusiver k_t-Algorithmus benutzt, der sowohl die CJC- als auch die Kalorimeter Informationen verwendet und den einzelnen Teilchen die Pionmasse zuordnet. Für die Bestimmung der Jetmasse werden die Vierervektoren der Tochterteilchen des jeweiligen Jets addiert. Die derart bestimmte Jetmasse sollte trotz falsch zugeordneter Teilchenmassen, sowie den vernachlässigten neutralen Teilchen, dennoch die unterschiedlichen Massen der jeterzeugenden Quarksorte (b- oder c-Quark) widerspiegeln und sich demzufolge gut zur Trennung der b- und c-Ereignisse eignen.

Abbildung 4.11 zeigt die b- und c-Monte-Carlo-Verteilungen der Sphärizität, Aplanarität und der Jetmasse. Die ML-Fehlerabschätzung ergibt: $\sigma_S(frac_b) = 0.075$, $\sigma_{Pl}(frac_b) =$

0.063 und $\sigma_{m(Jet)}(frac_b) = 0.041$ für die Sphärizität, Aplanarität bzw. Jetmasse (Abb. 4.12). Jede dieser drei Größen unterscheidet sich also genügend in ihren MC-Verteilungen um eine Trennung der b- und c-Eregnisse zu ermöglichen.

Die mit der Methode des Maximum-Likelihood-Fits berechneten Fehler des Fit Parameters, dem b-Anteil, sind in Tabelle 4.2 zusammengefaßt. Die kleinsten Standardabweichungen ergeben sich für die Verteilungen der rekonstruierten radialen D^0-Zerfallslänge, der Spurmultiplizität und der Jetmasse. Aber auch die anderen Größen zeigen Unterschiede in den MC-Referenzverteilungen und können in einer mehrdimensionalen Analyse, wie z.B. mit einem Neuronalen Netz, zur Trennung der b- und c-Ereignisse beitragen.

Meßgröße	ML-Fehlerabschätzung σ
$r(D^0)$	0.040
N_{Spuren}	0.031
$p_{T,rel}$	0.123
Sphärizität	0.075
Aplanarität	0.063
m_{Jet}	0.041

Tabelle 4.2: *Die mit der ML-Fehlerabschätzung bestimmten Standardabweichungen des b-Anteils der einzelnen in der Analyse verwendeten Selektionsgrößen.*

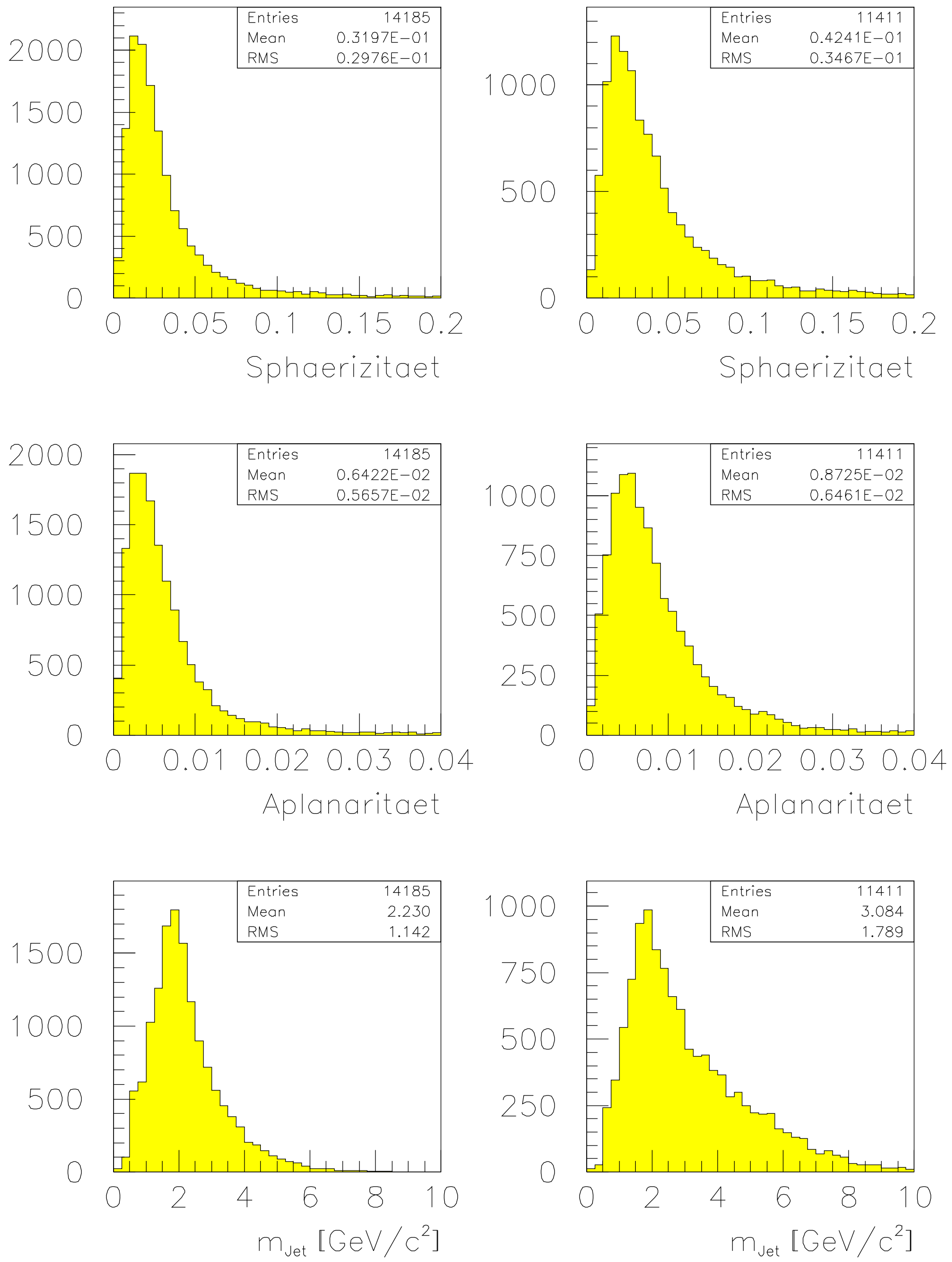

Abbildung 4.11: *Verteilungen der Sphärizität, der Aplanarität und der Jetmasse für c-MC (links) und b-MC (rechts).*

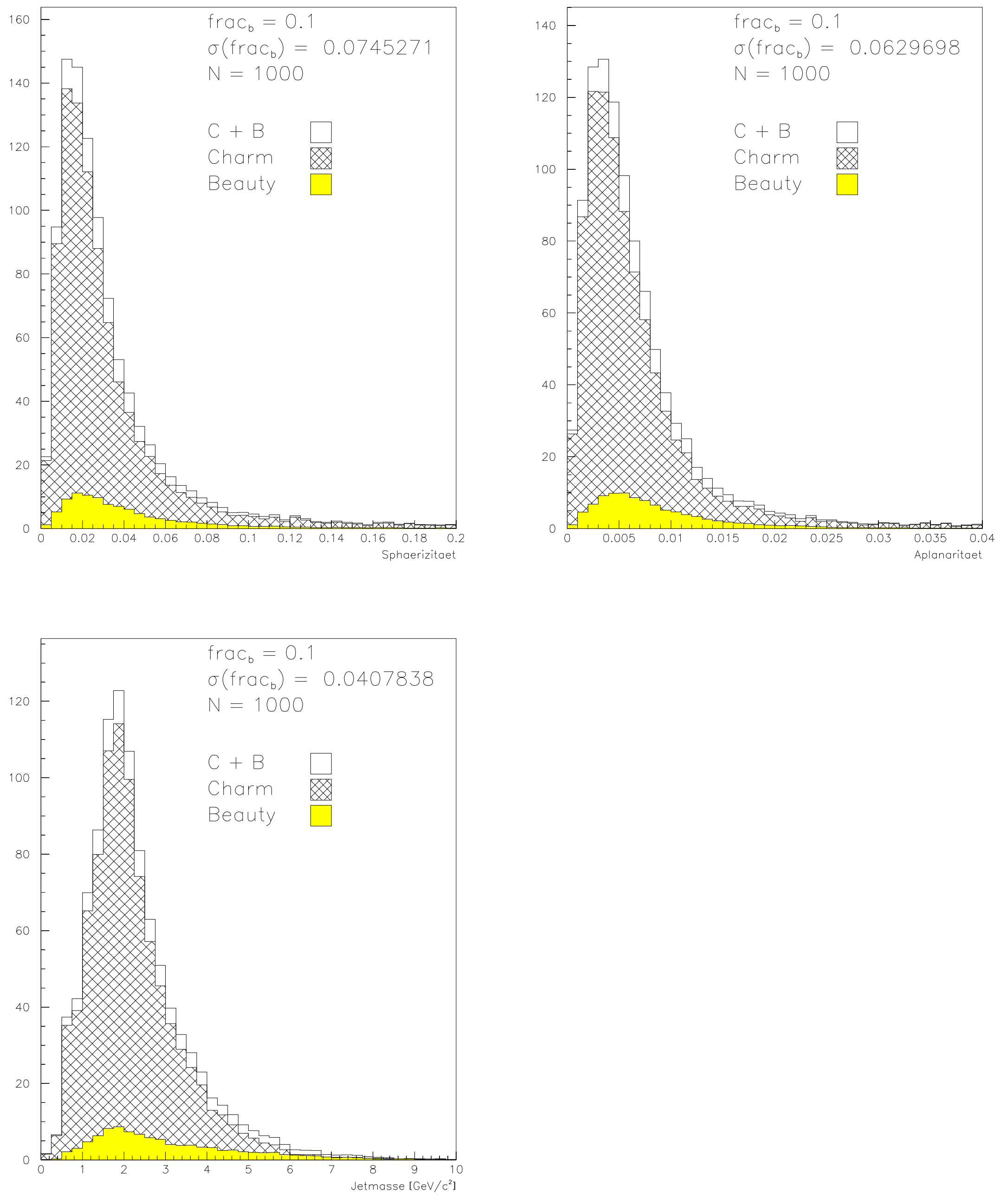

Abbildung 4.12: *Abschätzung des Fehlers für die Bestimmung des b-Anteils mit den Sphärizitäts- (oben links), Aplanaritäts- (oben rechts) und den Jetmassen-Verteilungen (unten links).*

Kapitel 5

Anwendung des Neuronalen Netzes

Im vorigen Kapitel wurden verschiedene Meßgrößen vorgestellt, mit denen eine Anreicherung der b-Ereignisse möglich ist. Im weiteren wird nun geprüft, ob eine Kombination dieser Meßgrößen mittels eines Neuronalen Netzes (NN) eine Verbesserung der b-Anreicherung liefert.

5.1 Grundlagen Neuronaler Netze

Neuronale Netze bieten eine Vielzahl von Anwendungsmöglichkeiten, so können sie u.a. zur Mustererkennung, Kategorisierung, Funktionsapproximation sowie zur Lösung von vielen anderen Problemen verwendet werden. Als wesentliche Stärken des Neuronalen Netzes, die diesen vielseitigen Einsatz ermöglichen, sind dabei die folgenden Punkte zu nennen.

- **Lernfähigkeit**
  Neuronale Netze sind sehr flexibel, da sie sich durch verschiedene anwendungsunabhängige Lernverfahren schnell an neue Umgebungen anpassen können.

- **Robustheit und Fehlertoleranz**
  Selbst unvollständige, verschwommene, unscharfe oder sogar teilweise fehlerhafte Informationen können von einem Neuronalen Netz verarbeitet werden.

- **Generalisierungsfähigkeit**
  Mit den geeigneten Trainingsbeispielen und Lernverfahren können Neuronale Netze Entscheidungsregeln entwickeln, die über die Trainingsdaten hinaus gültig sind.

- **Performanz**
  Trotz oftmals sehr aufwendigen Lernprozessen können Neuronale Netze in der Anwendung eine sehr hohe Performanz erreichen. Dadurch werden sie für Realzeitanwendungen sehr attraktiv.

Aufgrund dieser Eigenschaften sollten sich Neuronale Netze auch gut für den Bereich der Hochenergiephysik eignen. Für die vorliegende Analyse sind insbesondere die Robustheit, Fehlertoleranz und Generalisierungsfähigkeit des Neuronalen Netzes von Interesse. Eine ständige Anwendung bei H1 findet das NN in der Ereignisselektion der L2-Triggerstufe.

Im allgemeinen kann man Neuronale Netze als Abbildungsvorschrift verstehen, die eine Menge von Eingaben in eine Menge von Ausgaben abbildet. Neuronale Netze bestehen aus mehreren *Knoten*, die miteinander verbunden sind. Jede dieser Verbindungen wird durch ein *Gewicht* bewertet. Durch eine Modifikation dieser Gewichte kann das Ausgabeverhalten des Netzes in die gewünschte Form gebracht werden. Überlicherweise sind die einzelnen Knoten in Ebenen angeordnet. Ein Neuronales Netz beinhaltet zumindestens eine Eingabe- und eine Ausgabeebene. Die Eingabedaten (*Muster*) durchlaufen das Netz in Vorwärtsrichtung und werden durch die Knoten mit den entsprechenden Gewichten prozessiert: Die in den Knoten einlaufenden Signale werden summiert, und auf die so erhaltene Summe wird die Aktivierungsfunktion angewendet. Die Ausgabe erfolgt logischerweise an den Knoten der Ausgabeebene. Ein derartiges Netz wird als *Feed-Forward*-Netz bezeichnet.

Ein Feed-Forward-Netz mit nur einer Eingabe- und Ausgabeebene wird als einlagiges[1] Neuronales Netz bezeichnet. Mit einem solchen Netz kann man z.B. die aus der Logik bekannte AND-Funktion lösen. Allerdings zeigt sich auch sehr schnell die Limitation eines solchen Netzes, die gleichfalls bekannte XOR-Funktion kann nicht von einem einlagigen Netz berechnet werden [Her91]. Um diese Problematik zu beheben werden mehrlagige Netze verwendet. Diese enthalten zusätzliche Ebenen, die sogenannten *versteckten Ebenen*. Schon mit einem zweilagigen Neuronalen Netz, also einem Netz mit einer zusätzlichen versteckten Ebene, können sämtliche Boolschen Funktionen dargestellt werden. In Realanwendungen werden deshalb immer mehrlagige Neuronale Netze verwendet.

Das Backpropagation Lernverfahren

In der vorliegenden Analyse wird das Backpropagation (BP) Verfahren verwendet. Dieses ist das am gebräuchlichstende Lernverfahren eines Neuronalen Netzes. Das BP Lernverfahren bietet eine Vorschrift zur Veränderung der Gewichte w_{pq} des Neuronalen Netzes und ermöglicht somit das Lernen anhand der Ein-Ausgabe Paare $\{\xi_k^\mu, \zeta_i^\mu\}$ der Trainigsdaten.

Abbildung 5.1 zeigt den schematischen Aufbau eines zweilagigen Feed-Forward Netzes bestehend aus der Eingabe-, einer versteckten und der Ausgabeebene. Die Ausgabeknoten sind durch O_i, die versteckten Knoten durch V_j und die Eingabeknoten durch ξ_k gekennzeichnet. Mit W_{ij} und w_{jk} sind die Gewichte der Verbindungen zwischen den Ausgabe- und versteckten Knoten bzw. den versteckten und den Eingabeknoten bezeichnet. Die Trainigsdaten ξ_k^μ bestehen aus p Mustern ($\mu = 1, 2, ..., p$) mit jeweils N Eingabewerten ($k = 1, 2, ..., N$).

Ist das Muster ξ^μ gegeben, so erhält der versteckte Knoten j die Netzeingabe

$$h_j^\mu = \sum_k w_{jk} \xi_k^\mu \tag{5.1}$$

[1]Die Eingabeebene wird nicht mitgezählt.

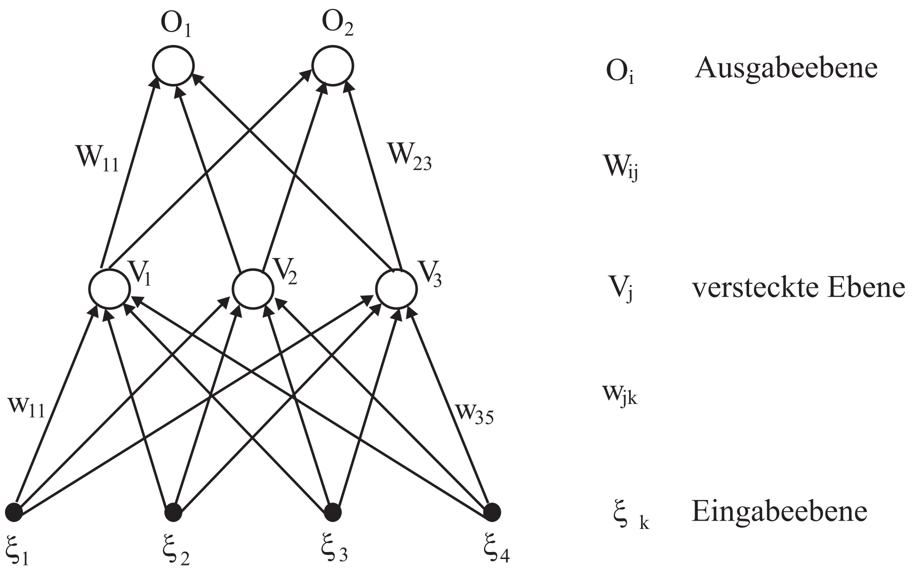

Abbildung 5.1: *Schematischer Aufbau eines zweilagigen Feed-Forward Netzes.*

und produziert die Ausgabe

$$V_j^\mu = g(h_j^\mu) = g\left(\sum_k w_{jk}\xi_k^\mu\right) \tag{5.2}$$

$g(h)$ bezeichnet dabei die Aktivierungsfunktion des Neuronalen Netzes; i.A. wird eine sigmoide Funktion verwendet. Der Ausgabeknoten i empfängt also den Eingabewert

$$h_i^\mu = \sum_j W_{ij}V_j^\mu = \sum_j W_{ij}g\left(\sum_k w_{jk}\xi_k^\mu\right) \tag{5.3}$$

und erzeugt den letztendlichen Ausgabewert

$$O_i^\mu = g(h_i^\mu) = g\left(\sum_j W_{ij}V_j^\mu\right) = g\left(\sum_j W_{ij}g\left(\sum_k w_{jk}\xi_k^\mu\right)\right) \tag{5.4}$$

Um den Trainigserfolg des Netzes zu bestimmen, wird ein Fehler E wie folgt definiert:

$$E(\mathbf{w}) = \frac{1}{2}\sum_{\mu i}[\zeta_i^\mu - O_i^\mu]^2 \tag{5.5}$$

Es wird also über das Quadrat der Differenzen der erwünschten Ausgabe ζ_i^μ und der tatsächlichen Ausgabe O_i^μ des Netzes summiert um den Fehler E zu bestimmen. Gleichung (5.5) wird auch als *Kostenfunktion* des Neuronalen Netzes bezeichnet.

Für das Backpropagation Lernverfahren ergibt sich also die Kostenfunktion zu

$$E(\mathbf{w}) = \frac{1}{2}\sum_{\mu i}\left[\zeta_i^\mu - g\left(\sum_j W_{ij}g\left(\sum_k w_{jk}\xi_k^\mu\right)\right)\right]^2 \tag{5.6}$$

Aufgabe des Training des Neuronalen Netzes ist es, die Gewichte W_{ij} bzw. w_{jk} derart zu verändern, daß sich die Kostenfunktion minimiert. Da diese eine für jedes Gewicht kontinuierlich differenzierbare Funktion ist, bietet sich als geeignetes Verfahren zur Minimierung der Gradientenabstieg an. Wendet man diesen auf die Gewichte der Verbindungen zwischen der versteckten- und der Ausgabebene an, so ergibt sich

$$\begin{aligned}\Delta W_{ij} &= -\eta \frac{\partial E}{\partial W_{ij}} &=& \eta \sum_\mu [\zeta_i^\mu - O_i^\mu] g'(h_i^\mu) V_j^\mu \\ && =& \eta \sum_\mu \delta_i^\mu V_j^\mu \end{aligned} \tag{5.7}$$

mit

$$\delta_i^\mu = g'(h_i^\mu)[\zeta_i^\mu - O_i^\mu] \tag{5.8}$$

η bezeichnet dabei die konstante Lernrate.

Für die Gewichte der Verbindungen zwischen der Eingabe- und der versteckten Ebene erhält man mit dem Gradientenabstiegsverfahren folgende Modifikationen

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = \eta \sum_\mu \delta_j^\mu \xi_k^\mu \tag{5.9}$$

mit

$$\delta_j^\mu = g'(h_j^\mu) \sum_i W_{ij} \delta_i^\mu \tag{5.10}$$

Die δ's geben also die Fehler an, die rückwärts durch das Neuronale Netz laufen, daher auch der Name Backpropagation.

Zusammengefaßt ergibt sich folgender Ablauf für das Backpropagation Optimierungsverfahren:

1. Initialisieren der Gewichte.
2. Die Eingabeschicht wird durch ein Muster ξ_k aktiviert.
3. Das Signal durchläuft in Vorwärtsrichtung das Netz, bis die Ausgabewerte O_i berechnet sind (feed forward).
4. Die Deltas der Ausgabeschicht werden nach (5.8) berechnet und rückläufig durchs Netz an die versteckte Ebene weitergegeben (back propagation).
5. Die Fehler der versteckten Ebene werden mit (5.10) berechnet.
6. Sämtliche Gewichte der Verbindungen werden mit (5.7) bzw. (5.9) neu bestimmt: $w^{\text{new}} = w^{\text{old}} + \Delta w$.
7. Es wird zu Schritt 2 zurückgegangen und das Verfahrens mit einem neuen Muster wiederholt.

Ein Backpropagation Netz führt also ein Gradientenabstiegsverfahren durch, welches eine Optimierung der Parameter in Richtung des steilsten Gefälles der Entscheidungsebene vornimmt. Abbildung 5.2 veranschaulicht diesen Vorgang.

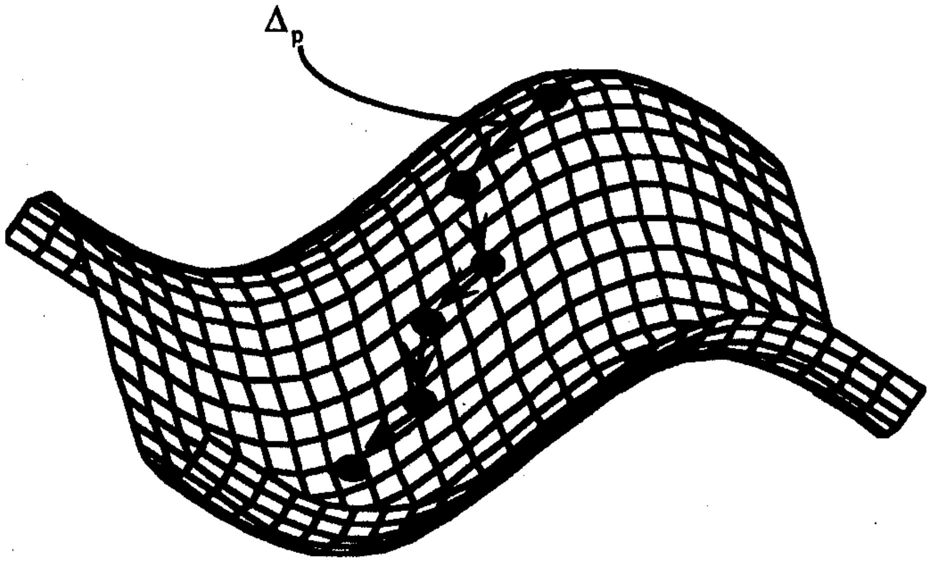

Abbildung 5.2: *Gezeigt ist die sukzessive Optimierung des BP Netzes, die einem Abstieg entlang des steilsten Gefälles der Ereignisebene entspricht. Mit Δ_p wird die Optimierung durch einen Durchlauf des Netzes bezeichnet (aus [She97]).*

5.2 Training und Test des Neuronalen Netzes

Für die Analyse wurde das Programmpaket **Jetnet 3.0** [Pet93] verwendet. Implementiert in diesem Softwarepaket sind verschiedene Strukturen mehrlagiger Neuronaler Netze wie z.B. das Backpropagation Lernverfahren oder selbstorganisierende Netze. Jetnet bietet desweiteren verschiedene Lernalgorithmen sowie eine Vielzahl von Parametern, die es ermöglichen, das Lernverhalten des Neuronalen Netzes abzuschätzen.

In Kapitel 4 wurden verschiedene Observablen für die Trennung der b- und c-Ereignisse getestet. Mit dem Neuronalen Netz als Hilfsmittel können alle diese Observablen gemeinsam zur Ereignisselektion genutzt werden. Es wird dabei erwartet, daß das NN mit diesen kombinierten Größen bessere Resultate erzielt, als es für die jeweiligen einzelnen Observablen der Fall ist. Dabei wird auch untersucht, inwieweit stark korrellierte Größen den Lerneffekt des Neuronalen Netzes verbessern können.

Es werden die sechs bereits in Kapitel 4 erwähnten Observablen, die D^0-Zerfallslänge, die Spurmultiplizität, der $p_{T,rel}$-Wert, die Sphärizität, die Aplanarität und die Masse des Jets als Eingabegrößen des Neuronalen Netzes verwendet. Demzufolge besteht die Eingabeschicht des Netzes aus sechs Knoten, einem für jede der verwendeten Observablen. Für das Training des Netzes werden die Observablen auf eine Größenordnung skaliert. Tabelle 5.1 gibt die Skalierungsfaktorn und den Mittelwert der Verteilungen der einzelnen Obversablen an.

Wie bereits erwähnt, wird für die Analyse das Backpropagation Lernverfahren verwendet. Trainiert wird das Netz mit einem Anteil der im Vorwege untersuchten b- und c-Monte-Carlo-Daten. Insgesamt werden 10000 Ereignisse, die sich aus gleichen Teilen an b- und c-MC zusammensetzen, als Trainingsdaten verwendet. Diese werden zufällig gemischt, damit

Meßgröße	Faktor	b-MC Mittelw.	c-MC Mittelw.
$r(D^0)$	10	0.81	0.47
N_{Spuren}	1/10	1.32	0.94
$p_{T,rel}$	1	1.19	1.00
Sphärizität	50	1.74	1.36
Aplanarität	100	0.91	0.67
m_{Jet}	1/5	0.63	0.45

Tabelle 5.1: *Eingabegrößen des Neuronalen Netzes mit ihren Skalierungsfaktoren und dem Mittelwert der b- und c-MC Verteilungen.*

das Netz mit einer möglichst homogenen Ereignismenge trainiert wird. Ein Durchlauf des Neuronalen Netzes über die gesamten Trainingsdaten wird als *Epoche* bezeichnet. Wünschenswert wäre es, eine derartig große Menge an Trainingsdaten zur Verfügung zu haben, so daß eine einzige Epoche zum Trainieren des Netzes ausreicht. Da dafür aber bei weitem nicht genügend MC-Statistik existiert, wird dieselbe Trainigsmenge wiederholt verwendet. Das Neuronale Netz durchläuft also mehrere Epochen, bis sich kein weiterer nenneswerter Lernerfolg mehr einstellt und das Netz als trainiert angesehen werden kann. Abbildung 5.3 zeigt den Lernerfolg L des verwendeten Netzes. Dieser ist wie folgt definiert:

$$L = \overline{\sum_{\mu} |O^{\mu} - \zeta^{\mu}|} \tag{5.11}$$

Es wurden auch verschiedene Mischverfahren der Trainingsdaten für jede Epoche getestet, dies führte allerdings zu keiner nennenswerten Änderung der Ausgabe des NN.

Die Ausgabe des Neuronalen Netzes erfolgt über einen einzigen Ausgabeknoten. Dieser kann kontinuierliche Werte zwischen 0 und 1 annehmen. Der Sollwert eines c-Ereignisses ist mit 0 und der eines b-Ereignisses mit 1 festgelegt.

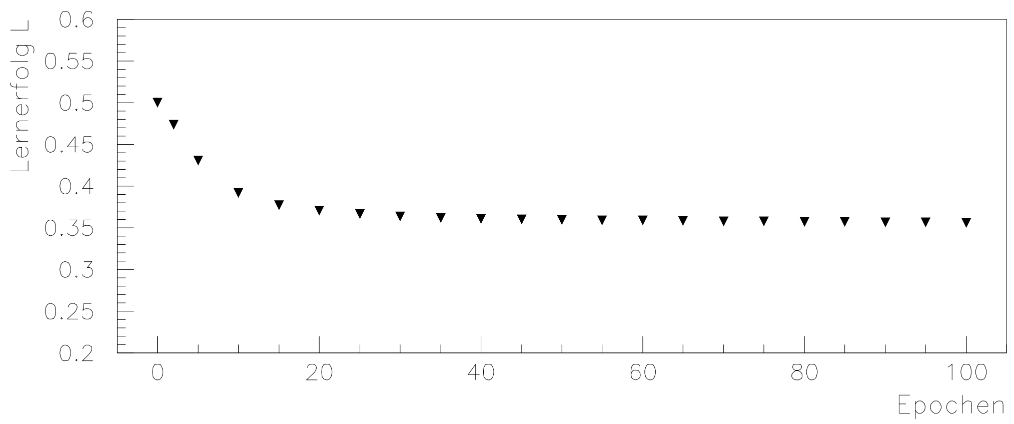

Abbildung 5.3: *Lernerfolg des Neuronalen Netzes.*

Folgende Parameter des Neuronalen Netzes werden für das Training verwendet:

- Lernmethode: Backpropagation
- 2 Ebenen
- 6 Eingabeknoten
- 10 Knoten für die versteckte Ebene
- 1 Ausgabeknoten
- 10000 Trainingsereignisse
- Initialisierung der Gewichte auf 0.5
- Aktivierungsfunktion $g(x) = \frac{1}{1+\exp(-2x)}$
- Lernrate $\eta = 0.2$
- 100 Epochen

Getestet wird das derart trainierte Netz mit den verbleibenden MC-Daten. Dazu wird es nacheinander auf die b- und c-MC-Daten angewendet, allerdings nur auf die Daten, die nicht zum Training des Netzes verwendet wurden. Dies soll zeigen, ob das Netz über die Menge der Trainigsdaten hinaus allgemeingültige Entscheidungsregeln gelernt hat. Abbildung 5.4 zeigt die Verteilungen der Ausgabewerte des Neuronalen Netzes. Mit Hilfe dieser beiden Verteilungen läßt sich auch wieder eine ML-Fehlerabschätzung (Abb. 5.5) durchführen. Diese ergibt einen Fehler von $\sigma(frac_b) = 0.024$ für einen geschätzten b-Anteil von $frac_b = 0.1$. Das Neuronale Netz ist also in der Lage, eine sinnvolle Kombination der Eingabewerte durchzuführen und ermöglicht dadurch eine bessere Trennung der b- und c-Ereignisse als die jeweiligen einzelnen Observablen.

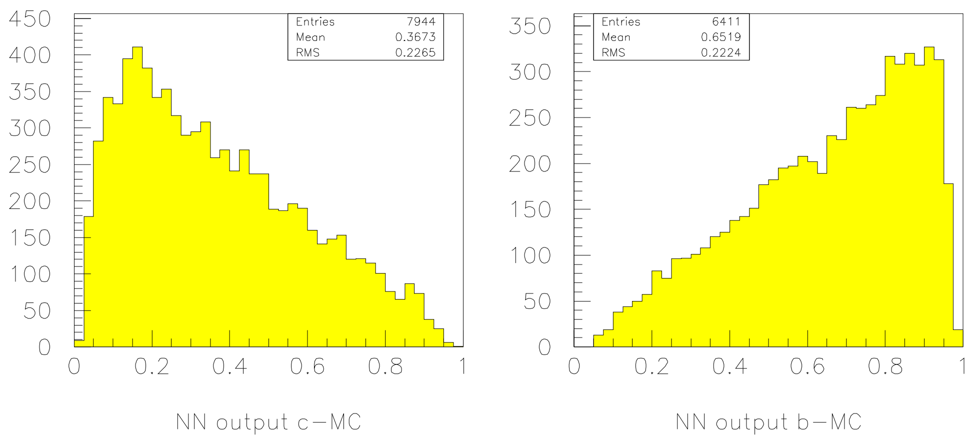

Abbildung 5.4: *Ausgabe des trainierten Neuronalen Netzes angewandt auf c-MC- Daten (links) und b-MC-Daten (rechts).*

Interessant ist es auch, die *Effizienz* des Neuronalen Netzes sowie die dabei noch erhaltene *Reinheit* der Ereignisse zu betrachten. Diese werden aus den Ausgabeverteilungen des

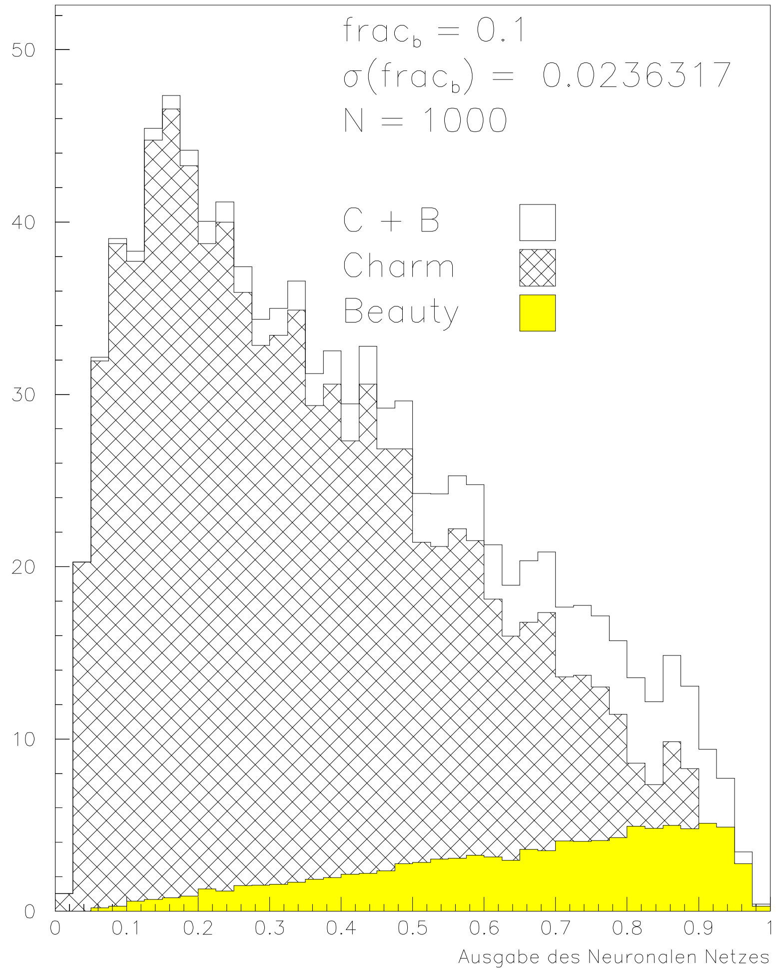

Abbildung 5.5: *Abschätzung des Fehlers für die Bestimmung des b-Anteils über Die Verteilungen der Ausgabewerte des trainierten Neuronalen Netzes.*

Testes des Netzes berechnet. Wählt man einen bestimmten Ausgabewert O_k und legt fest, daß alle Ereignisse mit einem Ausgabewert $O_i > O_k$ als b-Ereignisse gewertet werden, so ergibt sich die Effizienz $E_b(O_k)$ zu

$$E_b(O_k) = \frac{N_b(O_k)}{N_b} \tag{5.12}$$

mit N_b wird die Anzahl aller b-Ereignisse und mit $N_b(O_k)$ die Anzahl aller derjenigen b-Ereignisse, die den Schnitt $O_i > O_k$ erfüllen, bezeichnet. Die Reinheit $R_b(O_k)$, die den Anteil der richtig identifizierten b-Ereignisse von allen Ereignissen angibt, errechnet sich wie folgt

$$R_b(O_k) = \frac{N_b - N_b(O_k)}{(N_b - N_b(O_k)) + N_c(O_k)} \tag{5.13}$$

$N_c(O_k)$ gibt dabei die Anzahl aller c-Ereignisse an, die die Bedingung $O_i > O_k$ erfüllen. Abbildung 5.6 zeigt die gegeneinander aufgetragenen Effizienz- und Reinheits-Werte. Es ist deutlich zu sehen, daß ein Training des Neuronale Netzes mit allen sechs Observablen bessere Ergebnisse liefert als nur mit einer (in diesem Fall die radiale D^0-Zerfallslänge). Hier zeigt sich wieder, daß das NN auch mit stark korrellierte Daten noch wesentliche Lernerfolge erzielen kann.

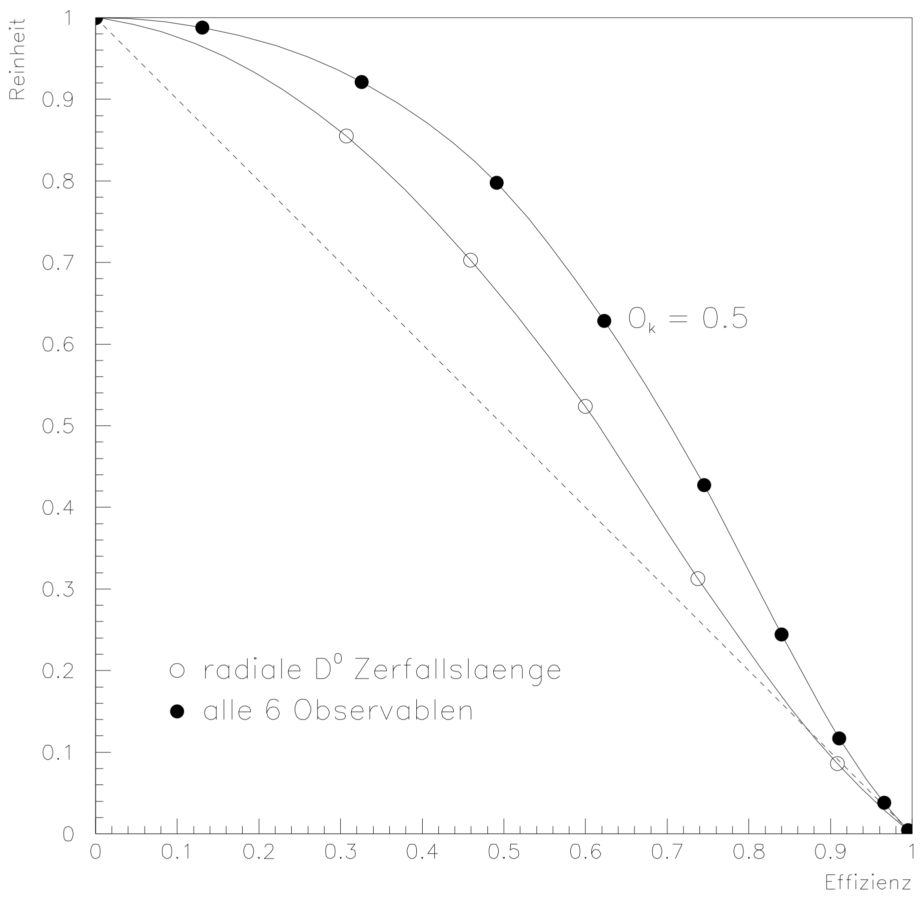

Abbildung 5.6: *Die Reinheit der durch das NN selektierten b-Ereignisse aufgetragen gegen die Effizienz. Die Punkte markieren den gewählten Schnitt O_k der Ausgabewerte des Netzes in Intervallen von 0.1.*

5.3 Analyse der Daten mit dem Neuronalen Netz

Das wie oben beschrieben trainierte und getestete Neuronale Netz wird nun auf die Daten angewendet. Es werden wieder die bereits in Kapitel 3.3 erwähnten Index-Files und Schnitte verwendet. Abbildung 5.7 zeigt die Verteilung der Ausgabewerte des Neuronalen Netzes. Diese Verteilung wird in zehn gleichgroße Intervalle (Bins) des Ausgabewertes des NN unterteilt. Für sämtliche Ereignisse eines Bins wird die Massendifferenz $m_{K\pi\pi_s} - m_{K\pi}$ geplottet und die Anzahl der D^*-Mesonen durch einen Fit bestimmt. Die Fit-Funktion besteht aus zwei Teilen, einem Untergrundterm der Form $N \cdot (\Delta m - m_\pi)^{0.35}$ und einer

Gaußkurve mit einem Mittelwert von 145.4 MeV/c^2 und einer Breite von 1.1 MeV/c^2. Aus der Fläche der Gaußkurve wird die Anzahl der D^*-Mesonen in dem jeweiligen Bin ermittelt. Die einzelnen Verteilungen der Massendifferenz mit den jeweiligen Fit-Ergebnissen sind in Abbildung A.1 in Anhang A.2 dargestellt.

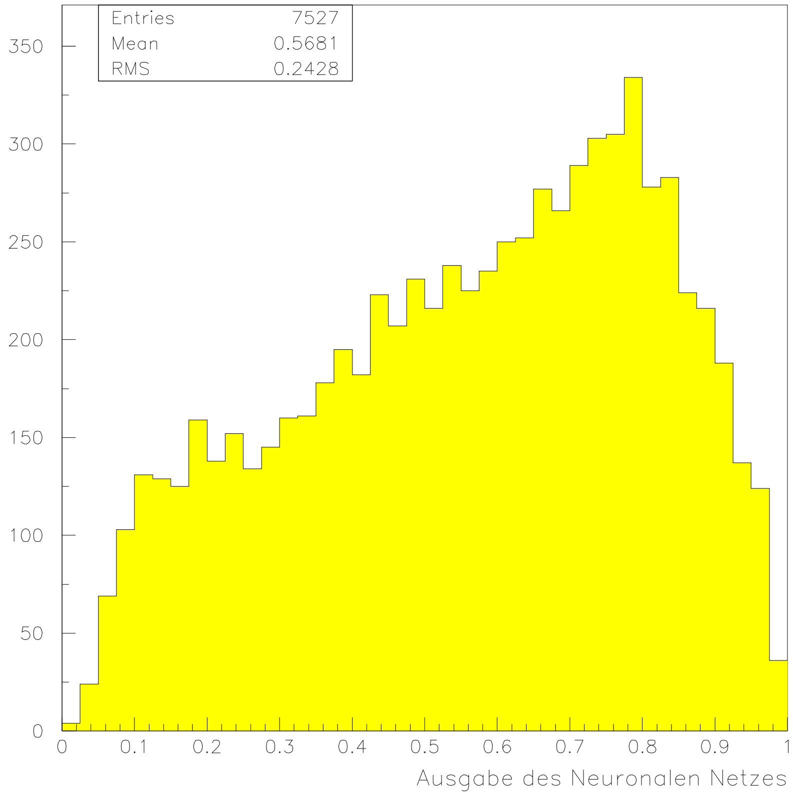

Abbildung 5.7: *Ausgabe des Neuronalen Netzes bei Anwendung auf die H1-Daten des Jahres 1997.*

Die so erhaltene Verteilung der Anzahl der D^*-Mesonen in Intervallen des Ausgabewertes des Neuronalen Netzes ermöglicht es, den Anteil an b- und c-Ereignissen zu bestimmen. Dazu werden die b- und c-MC-Referenzverteilungen (Abb. 5.4) über einen χ^2-Fit an die Daten angepaßt. Als einziger Parameter des χ^2-Fits wird der Anteil der b-Ereignisse ($frac_b$) verwendet. Der Anteil der D^*-Mesonen, die aus $c\bar{c}$-Zerfällen herrühren, berechnet sich dann nach $frac_c = 1 - frac_b$. Abbildung 5.8 zeigt die erhaltene Verteilung der Analyse der 1997'er Daten sowie die über den χ^2-Fit angepaßten MC-Referenzverteilungen.

D^*-Mesonen aus	Anzahl
b-Ereignissen	235 ± 39
c-Ereignissen	876 ± 39

Tabelle 5.2: *Gemessene Anzahl der D^*-Mesonen aus b- und c-Ereignissen.*

Tabelle 5.2 gibt die über den Fit bestimmten Anteile an b- und c-Ereignissen an. Es wurden insgesamt 1111 ± 44 D^*-Mesonen rekonstruiert. Für den Fit ergibt sich $\chi^2 = 20.73$

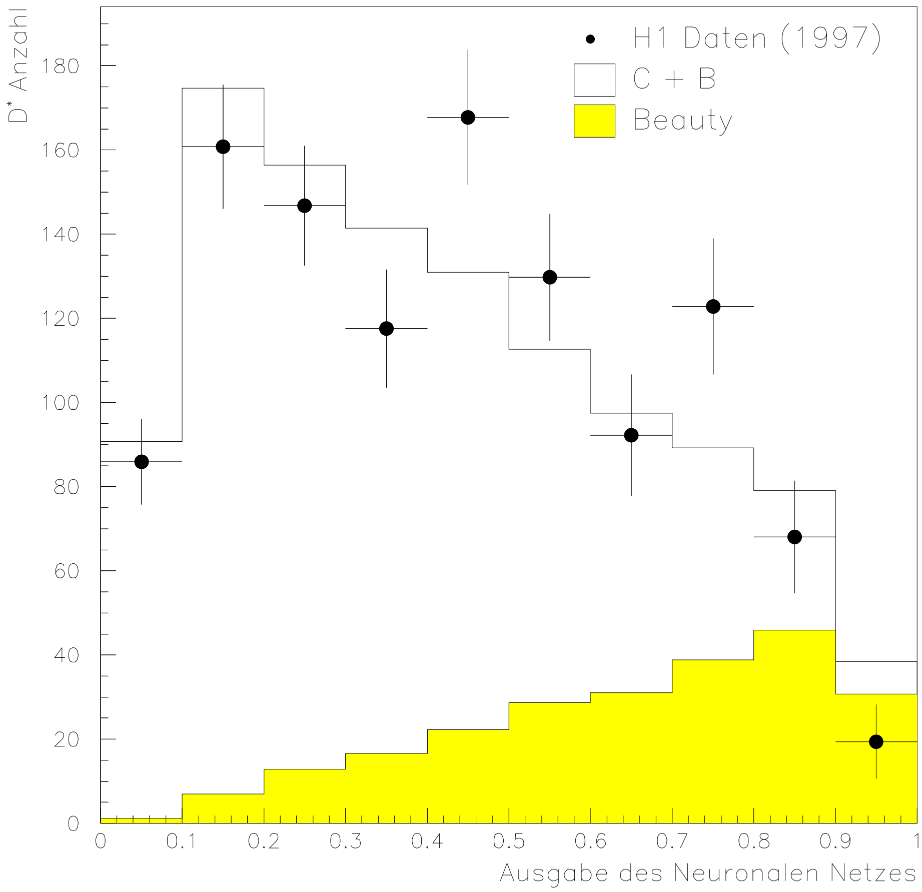

Abbildung 5.8: *Anzahl der D^*-Mesonen in Intervallen des Ausgabewertes des Neuronalen Netzes sowie die über den χ^2-Fit angepaßten MC-Referenzverteilungen.*

bei einer Anzahl von 9 Freiheitsgraden (N_F = 10 Bins - 1 Parameter). Das Verhältnis $\chi^2/N_F = 2.31$ zeigt, daß die Beschreibung der Verteilung der Daten durch die beiden MC-Verteilungen nicht optimal ist. Allerdings gibt es auch keine Anzeichen für starke Abweichungen in einem Intervall der Ausgabegröße. Die Signifikanz $frac_b/\sigma(frac_b) = 6.03$ des b-Anteils ist aber dennoch gut und rechtfertigt die Analyse mit dem Neuronalen Netz.

5.4 Ausblick

Das Ziel weiterführender Analysen sollte die Bestimmung eines b-Wirkungsquerschnittes sein. Dazu müssen geeignete Trigger für die Ereignisselektion gewählt und deren Effizienzen bestimmt werden. Weiterhin ist es notwendig, die Akzeptanz und die Effizienz der D^*-Selektion zu berechnen und eine Unterscheidung der Ereignisse in Photoproduktion und DIS vorzunehmen. Da bei diesen Schritten viel Statistik verlorengeht, sollte die Analyse auf die H1 Daten der Jahre 1998-2000 ausgeweitet werden, wodurch in etwa eine

Vervierfachung der Luminosität erreicht wird. Desweiteren müssen auch noch mögliche systematische Fehler untersucht werden. Es wurde z.B. noch nicht geprüft in wie weit die MC-Verteilungen tatsächlich die Daten beschreiben und wie empfindlich das Ergebnis auf mögliche Abweichungen der für die Analyse verwendeten Verteilungen reagiert.

Zusammenfassung

In dieser Arbeit wird die Erzeugung schwerer Quarks in der ep-Wechselwirkung bei HERA untersucht. Die Ereignisse mit schweren Quarks werden durch den Nachweis des D^*-Mesons im Zerfallskanal $D^* \longrightarrow D^0\pi_s \longrightarrow (K\pi)\pi_s$ identifiziert.

D^*-Mesonen können sowohl bei b- als auch bei c-Ereignissen entstehen. Ziel dieser Arbeit ist es, Ereignisse dieser beiden schweren Quarks möglichst gut zu trennen. Dazu wurden einzelne kinematische Variablen und Methoden untersucht, mit denen eine solche Trennung möglich sein sollte.

Zunächst wurde jede der gewählten kinematischen Variablen einzeln auf ihr Trennvermögen hin untersucht. Um dieses zu bestimmen, wurde der statistische Fehler des b-Anteils aus den Unterschieden der Verteilungen von b- und c-Monte-Carlo abgeschätzt. Insbesondere sechs Variablen erwiesen sich als besonders gut zur Trennung geeignet: Rekonstruierte D^0-Zerfallslänge, Spurmultiplizität, Transversalimpuls des D^0 bezüglich der Thrustachse, Sphärizität, Aplanarität, Jetmasse. Von diesen sechs Variablen erbrachte die Spurmultiplizität den kleinsten abgeschätzten Fehler, bei einem angenommenen b-Anteil von $frac_b$ $= 0.1$, für die Bestimmung des b-Anteils: $\sigma(frac_b) = 0.031$.

Im weiteren wurde untersucht, ob durch eine mehrdimensionale Analyse die Trennung der b- und c-Ereignisse verbessert werden kann. Dazu wurde ein Neuronales Netz benutzt, das nach dem Backpropagation-Lernverfahren trainiert wurde. Für das Training des Netzes wurde ein Teil der b- und c-MC-Daten verwendet. Mit den restlichen MC-Daten wurde das Netz getestet.

Das Neuronale Netz liefert eine normierte Ausgabevariable, die zur Trennung der b- und c-Ereignisse verwendet werden kann. Durch die Wahl eines entsprechenden Schnittes auf die Ausgabevariable kann die Reinheit bzw. die Effizienz der b-Selektion bestimmt werden. So ergibt sich z.B. für eine Untersuchung der MC-Daten eine Reinheit von 78% bei einer Effizienz von 50%. Vergleicht man dies mit Untersuchungen einzelner Variablen, so zeigt sich eine deutliche Verbesserung. Wird für die Selektion z.B. nur die rekonstruierte D^0-Zerfallslänge verwendet, so ergibt sich bei einer Effizienz von 50% nur eine Reinheit von 65%. Es zeigt sich also, daß die mehrdimensionale Analyse mit dem Neuronalen Netz selbst mit stark korrelierten Variablen bessere Ergebnisse als eine b-Selektion mit einzelnen Variablen erzielt.

Das trainierte Neuronale Netz wurde in einem ersten Test auf die H1-Daten des Jahres 1997 angewendet. Die Analyse dieser Daten mit einer integrierten Luminosität von $\mathcal{L}$ $= 18.7$ pb^{-1} ergab eine Anzahl von 1111 ± 44 D^*-Mesonen. Mit dem Neuronalen Netz konnte die Anzahl der D^*-Mesonen aus b-Ereignissen zu 235 ± 39 bestimmt werden. Dieses

Ergebnis zeigt, daß die Methode einen relativ kleinen statistischen Fehler liefert. Allerdings wurde diese Analyse noch nicht auf mögliche systematische Fehler hin untersucht. So muß z.B. noch untersucht werden, in wie weit die experimentellen Verteilungen durch die MC-Simulation beschrieben werden und wie empfindlich das Ergebnis von möglichen Abweichungen abhängt.

Anhang A

Anhang

A.1 Maximum-Likelihood-Fit

Liegen mehrere Messungen der Zufallsvariablen x vor, die der bekannten Wahrscheinlichkeitsdichte $f(x|a)$ genügen, und wird mit a der Parameter, von dem diese Wahrscheinlichkeitsdichte abhängt, beschrieben, so kann mit einem Maximum-Likelihood-Fit die beste Schätzung $\hat{a}$ des Parameters gewonnen werden.

Möchte man z.B. den b-Anteil aus der D^0 Zerfallslängenverteilung bestimmen, so kann man dies mit einem Maximum- Likelihood-Fit erreichen.

In diesem Beispiel ist die Funktion $f(r,a)$, die die Verteilung der Zerfallslänge beschreibt, wie folgt definiert:

$$f(r,a) = a \cdot f_b(r) + (1-a) \cdot f_c(r) \tag{A.1}$$

Wichtig ist dabei, daß die Funktionen $f_b(r)$ und $f_c(r)$, die die Monte-Carlo Verteilungen der D^0 Zerfallslänge darstellen (für b- bzw. c-MC), jeweils einzeln normiert sind. Der Parameter a steht für den über den Fit zu bestimmenden b-Anteil.

Ohne den Fit selbst durchzuführen, läßt sich aus den Eigenschaften der beiden Funktionen $f_b(r)$ und $f_c(r)$ der statistische Fehler des Parameters a berechnen. Dieser ist wie folgt definiert:

$$\sigma(a) \;=\; \frac{1}{\sqrt{N}} \left[\int \frac{1}{f} \left(\frac{\partial f}{\partial a} \right)^2 dr \right]^{-\frac{1}{2}} \;=\; \frac{1}{\sqrt{N}} \cdot H^{-\frac{1}{2}} \tag{A.2}$$

N bezeichnet die Anzahl der Ereignisse der D^0-Zerfallslängenverteilung. Es ergibt sich also

$$\begin{aligned} \frac{\partial f}{\partial a} &= f_b(r) - f_c(r) \\ \Rightarrow H &= \int \frac{(f_b(r) - f_c(r))^2}{a \cdot f_b(r) + (1-a) \cdot f_c(r)}\, dr \end{aligned} \tag{A.3}$$

Dieses Integral wird in der Analyse numerisch berechnet. Je stärker sich die beiden Funktionen $f_b(r)$ und $f_c(r)$ voneinander unterscheiden, desto kleiner wird der Fehler von a.

A.2 Anzahl der D^* in Intervallen des Ausgabewertes des NN

Um eine möglichst untergrundfreie Analyse durchführen zu können, wird die Verteilung der Ausgabewerte des Neuronalen Netzes, die sich über das Intervall [0,1] erstreckt (s. Abb. 5.7), in zehn gleichgroße Bins unterteilt. Für die Ereignisse eines jeden Bins wird die Verteilung der Massendifferenz $m_{K\pi\pi_s} - m_{K\pi}$ aufgetragen und über einen Fit die Anzahl der D^* in dem jeweiligen Bin bestimmt (s. Kapitel 5.3). Abbildung A.1 zeigt die zehn Verteilungen der Massendifferenz und den Fit zur Bestimmung der D^*-Anzahl in den jeweiligen Bins des Ausgabewertes des Neuronalen Netzes.

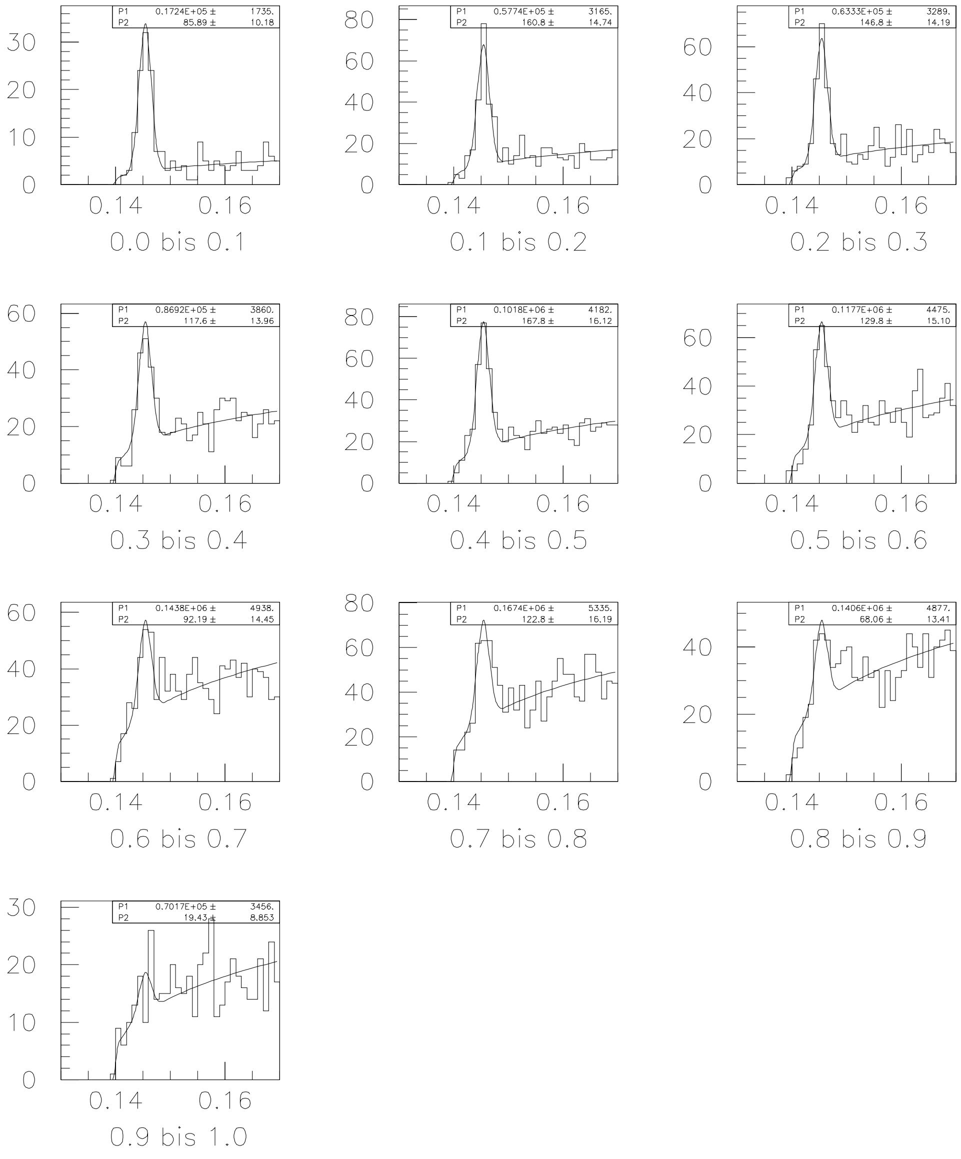

Abbildung A.1: *Verteilung der Massendifferenz* $m_{K\pi\pi_s} - m_{K\pi}$ *[GeV/c²] in Intervallen des Ausgabewertes des Neuronalen Netzes. Die Fitparameter sind: P1 = Faktor des Untergrundterms, P2 = Anzahl der Einträge im Signal.*

Literaturverzeichnis

[ARO95] G. Ingelman, J.Rathsman, G.A. Schuler, AROMA 2.1 - A Monte Carlo Generator for Heavy Flavour Events in *ep* Collisions, DESY (1995)

[Blo98] V. Blobel, E. Lohrmann, Statistische und numerische Methoden der Datenauswertung, Teubner Studienbücher (1998)

[Bra79] S. Brandt und H.D. Dahmen, Axes and Scalar Measures of Two-Jet and Three-Jet Events, Zeitschrift für Physik C (Particles and Fields) 61-70 (1979)

[H1C96a] H1 Collaboration, The H1 Detector at HERA, DESY 96-01 (1996)

[H1C96b] Inclusive D^0 and $D^{*\pm}$ Production in Deep Inelastic *ep* Scattering at HERA, DESY 96-138 (1996)

[Her91] J. Hertz, A. Krogh, R.G. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley Publishing Company (1991)

[Kau98] M. Kausch-Blecken v. Schmeling, The Silicon Microvertex Detector of the H1 Experiment: Readout, Event Reconstruction and Studies on Heavy Quark Decays, Dissertation Hamburg (1998)

[Kuh98] T. Kuhr, Rekonstruktion von V^0s mit dem H1-Silizium- Detektor, Diplomarbeit Hamburg (1998)

[Kuh99] T. Kuhr, Primary Vertex Finding in H1, Crosstalk DESY (1999)

[MCG99] A.T. Doyle, G. Grindhammer, G. Ingelman, H. Jung, Monte Carlo Generators for HERA Physics (Proceedings of the Workshop 1998-1999), DESY-PROC-1999-02 (1999)

[Nar97] B. Naroska, Heavy Flavour Production in Lepton Proton Collision at HERA, Hamburg (1997), H1-01/97-510

[NPB99] H1 Collaboration, Measurement of D^* Meson Cross Sections at HERA and Determination of the Gluon Density in the Proton Using NLO QCD, Nuclear Physics B 545 (1999) 21-44

[Pet93] C. Peterson, T. Rögnvaldsson, JETNET 3.0 - A Versatile Artificial Neural Network Package, CERN-TH. 7135/94 (1993)

[PDG98] Particle Data Group, Review of Particle Physics, The European Physical Journal C (1998)

[Pre96] S. A. Prell, Entwicklung des Steuer- und Auslesemoduls für die Siliziumstreifendetektoren des H1-Detektors, Dissertation Hamburg (1996)

[She97] A. Scherer, Neuronale Netze - Grundlagen und Anwendungen, Vieweg Verlag (1997)

[Sch95] P. Schmüser, Feynman-Graphen und Eichtheorien für Experimentalphysiker (zweite Auflage), Springer Verlag (1995)

[Ste99] J. Steinhart, Die Messung des totalen $c\bar{c}$- Photoproduktions-Wirkungsquerschnittes durch die Rekonstruktion von Λ_c-Baryonen unter Verwendung der verbesserten dE/dx- Teilchenidentifikation am H1 Experiment bei HERA, Dissertation Hamburg (1999)

Danksagung

Ich möchte mich herzlichst bei all den Personen bedanken, die es mir ermöglicht haben, diese Arbeit anzufertigen. Insbesondere gilt mein Dank Herrn Prof. Dr. Volker Blobel für seine gute Betreuung und fortwährende Unterstützung. Desweiteren danke ich Dr. Olaf Behnke, Dr. Marc Weber, Thomas Kuhr, Andy Ziegler und Jeannine Wagner für ihre nie endende Geduld bei der Beantwortung meiner unzähligen Fragen und ihre Hilfestellung bei vielen meiner Probleme. Auch möchte ich mich bei allen hier nicht namentlich genannten Mitgliedern des H1 Experimentes bedanken, die mir während meiner Arbeit weitergeholfen haben. Zum Schluß möchte ich mich noch bei meinen Eltern bedanken, ohne ihre Ermunterungen und Unterstützung während meines gesamten Studiums wäre diese Arbeit nie zustande gekommen.

Ich versichere, daß ich die Arbeit unter Verwendung der angegebenen Quellen selbständig angefertigt habe.