Erkennung semimyonischer Zerfälle schwerer Quarks durch Neuronale Netze am H1–Detektor bei HERA

Diplomarbeit

vorgelegt von

Stefan Lüders

II. Institut für Experimentalphysik
Universität Hamburg

Hamburg,
September 1997

Inhaltsverzeichnis

Kapitel 1

Einleitung

> Das den Raum ausfüllende Volle besteht aus unzähligen,
> mit bloßen Augen nicht wahrnehmbaren Atomen, also den Unteilbaren.
> *Demokrit, 5. Jh. v. Chr.*

Mehr als 2300 Jahre sollte die oben angeführte Erkenntnis von Demokrit über die Struktur der Materie bestehen bleiben, bis die vielfältigen Erscheinungsformen der Natur von Mendelejew 1869 im Periodensystem der Elemente katalogisiert wurden. Mit der Entdeckung des Elektrons durch J.J. Thomson 1897, E. Rutherfords Entdeckung des Protons 1919, der Formulierung der Quantenmechanik durch W.K. Heisenberg, N. Bohr, E. Schrödinger, u.a. Anfang des 20. Jh. und des Nachweises des Neutrons durch J. Chadwick 1932 war das Periodensystem mit den drei Teilchen Proton, Neutron und Elektron vollständig erklärt. Die kosmische Höhenstrahlung und neue Teilchenbeschleuniger führten jedoch zu der Entdeckung einer Vielzahl neuer Teilchen, so daß sich die gefundene Ordnung zu einem schwer zu überschauenden „Teilchenzoo“ entwickelte.
1964 formulierten M. Gell–Mann und G. Zweig mit dem Quark–Modell, das „up“–, „down“– und „strange“–Quarks postulierte, eine neue Grundlage in der Ordnung der Teilchen. Das Standardmodell der Elementarteilchenphysik kennt heute drei Familien mit jeweils zwei Quarks und zwei Leptonen als Materiefelder. Weiterhin gibt es fünf Bosonen für die Wechselwirkungen der Materiefelder untereinander (Tabelle 1.1), wobei sich mit der starken und der elektromagnetischen Wechselwirkung aus den Materiefeldern der „Teilchenzoo“, das Periodensystem der Elemente und Demokrits Atom aufbauen lassen.

Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$
Leptonen	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$

Wechselwirkung	Boson
stark	*Gluon*
schwach	$W^\pm$, Z^0
elektromagnetisch	γ
Gravitation	*Graviton*(?)

Tabelle 1.1: *Die fundamentalen Teilchen–Familien (links) und Wechselwirkungen (rechts).*

Am derzeit einzigen Elektron–Proton–Speicherring HERA am Deutschen Elektronen–Synchrotron DESY in Hamburg werden Elektronen bzw. Positronen mit Protonen zur Kollision gebracht. Die Elektronen (Positronen) werden dabei auf eine Energie von 27.6 GeV und die Protonen auf 820 GeV beschleunigt.
Die Schwerpunktenergie von etwa 300 GeV liegt um eine Größenordnung über dem von Fixed–Target–Experimenten erreichten Wert, so daß mit HERA eine neue kinematische Region zugänglich wird. Damit ist es möglich, das Standardmodell zu testen, sowie einen tieferen Einblick in die Struktur von Proton und Photon zu gewinnen. Einer der Forschungsschwerpunkte ist die Physik der schweren *Beauty*– und *Charm*–Quarks, deren Wirkungsquerschnitte bei der erreichten Schwerpunktsenergie zu ausreichenden Ereignisraten führen.

Die vorliegende Arbeit untersucht Prozesse der Erzeugung schwerer Quarks.
Schwere Quarks entstehen bei HERA vorwiegend als Quark–Antiquark–Paare im Prozeß der Boson-Gluon–Fusion mit einem quasireellen Photon (Photoproduktion). Das Quarkpaar bildet einerseits den gebunden Zustand J/ψ bzw. Υ, zum anderen entstehen B– und D–Mesonen durch die Verbindung des schweren Quarks mit einem leichteren Quark.
Das Ziel ist es, die schweren B– und D–Mesonen an der Signatur ihres semimyonischen Zerfalls untereinander und von anderen Prozessen mit Myonen — dem Untergrund — zu unterscheiden. Dabei liegt der Schwerpunkt in der Erkennung von Beauty–Quarks. Die nötigen Unterscheidungskriterien werden anhand eines Satzes von Ereignisvariablen gebildet, die unter Anwendung von Neuronalen Netzen eine Separation ermöglichen sollen.

In Kapitel 2 wird die Speicherringanlage HERA vorgestellt und die für die vorliegende Analyse relevanten Detektor–Komponenten des H1–Experimentes genauer beschrieben.
Das Konzept der Neuronalen Netze und deren mathematische Realisierung durch Feedforward–Netzwerke mit Back–Propagation–Lernalgorithmus wird in Kapitel 3 behandelt.
Kapitel 4 erläutert die theoretischen Grundlagen der Erzeugung schwerer Quarks bei HERA. Die zur Beschreibung der Kinematik von Ereignissen der ep–Streuung verwendeten Größen werden definiert und der dominierende Prozeß der Boson–Gluon–Fusion wird vorgestellt. Die Zerfallskanäle der schweren Quarks in ein Myon und die Untergrund–Prozesse, die zur Myonen–Produktion beitragen oder Myonen im Detektor vortäuschen, werden diskutiert und deren Simulation durch Ereignisgeneratoren beschrieben.
Eine Einteilung der generierten Ereignisse in Klassen mit semimyonischen Zerfällen von Beauty–Quarks, semimyonischen Zerfällen von Charm–Quarks, sowie den Untergrund–Ereignissen erfolgt in Kapitel 5. Es werden grundlegende Auswahlkriterien an die Ereignisse und speziell an die Myonen gestellt und durch das Myon eine topologische Unterteilung des Ereignisses definiert. Anschließend werden in Kapitel 6 die Ereignisvariablen vorgestellt, die aufgrund der Analyse mit dem empirischen Korrelationskoeffizienten prädestiniert sind, die drei Klassen zu trennen.
Die Trennung der Klassen wird in Kapitel 7 nach dem Lernvorgang der Neuronalen Netze mit den ausgewählten Ereignisvariablen vollzogen. Dabei wird die Signifikanz der einzelnen Eingabegrößen überprüft und die Genauigkeit des Lernvorganges abgeschätzt.
Abschließend gibt Kapitel 8 eine Zusammenfassung.

Kapitel 2

HERA und der H1–Detektor

> Three quarks for Muster Mark !
> Sure he hasn't got much of a bark
> And sure any he has it's all beside the mark.
> But O, Wreneagle Almighty, wouldn't un be a sky of a lark
> To see that old buzzard whooping about for uns shirt in the dark
> And he hunting round for uns speckled trousers around by Palmerstown Park ?
> *James Joyce in „Finnegans Wake", 1971*

Die Speicherringanlage HERA ist das derzeit leistungsfähigste und hochauflösendste Elektronen–Mikroskop für die Untersuchung der Struktur von Protonen und deren Konstituenten, den *Quarks* und *Gluonen*, sowie zur Überprüfung des Standardmodells der Elementarteilchenphysik.

Dieses Kapitel gibt einen kurzen Überblick über den Speicherring HERA (Abschnitt 1) und über das Experiment H1, wobei der Schwerpunkt auf den für die vorliegende Analyse relevanten Komponenten des H1–Detektors liegt (Abschnitt 2).

2.1 Der Speicherring HERA

Die **H**adron–**E**lektron–**R**ing–**A**nlage HERA (Abbildung 2.1) des **D**eutschen **E**lektronen–**Sy**nchrotrons DESY wurde nach etwa sechs Jahren Bauzeit und Baukosten von einer Milliarde Mark im Jahre 1990 fertiggestellt. Der Speicherring liegt in einer Tiefe von 10 m bis 25 m unter der Erdoberfläche und hat einen Umfang von 6 336 m.

In zwei separaten Strahlrohren werden Elektronen oder Positronen (*HERA–e*) und Protonen (*HERA–p*) auf Endenergien von 27.6 GeV bzw. 820 GeV beschleunigt und gespeichert, nachdem sie ein System von Vorbeschleunigern durchlaufen haben. Die wesentlichen Komponenten des Vorbeschleunigersystems mit den jeweils erreichten Impulsen sind in Abbildung 2.2 dargestellt. Dabei stehen die Abkürzungen LINAC und PETRA für **Lin**ear **Ac**celerator und **P**ositron–**E**lektron–**T**andem–**R**ing–**A**ccelerator.

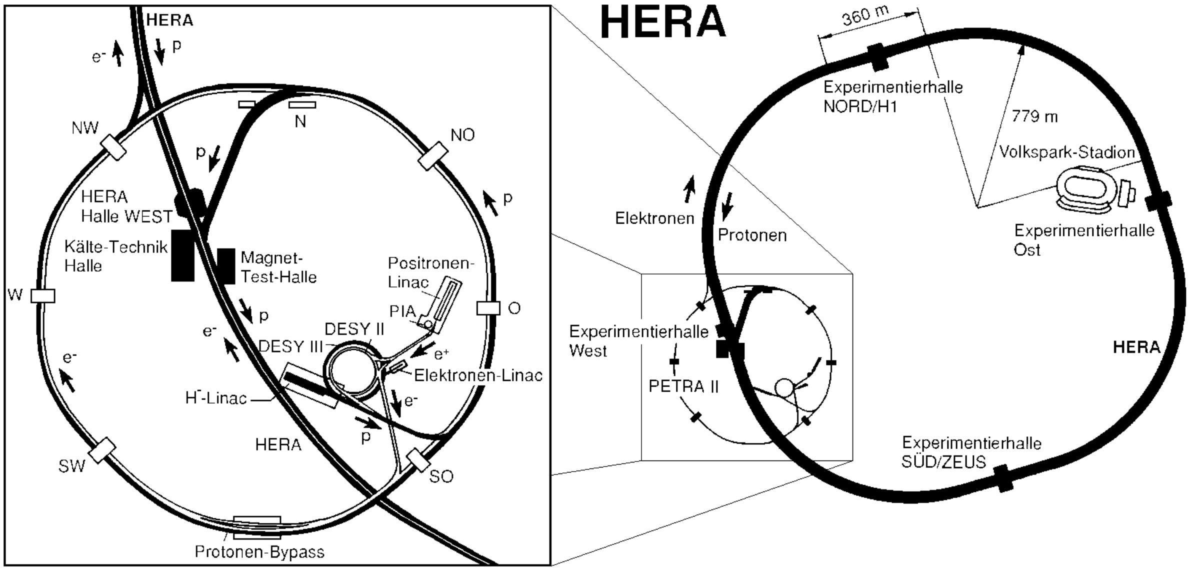

Abbildung 2.1: *Der Speicherring HERA.*

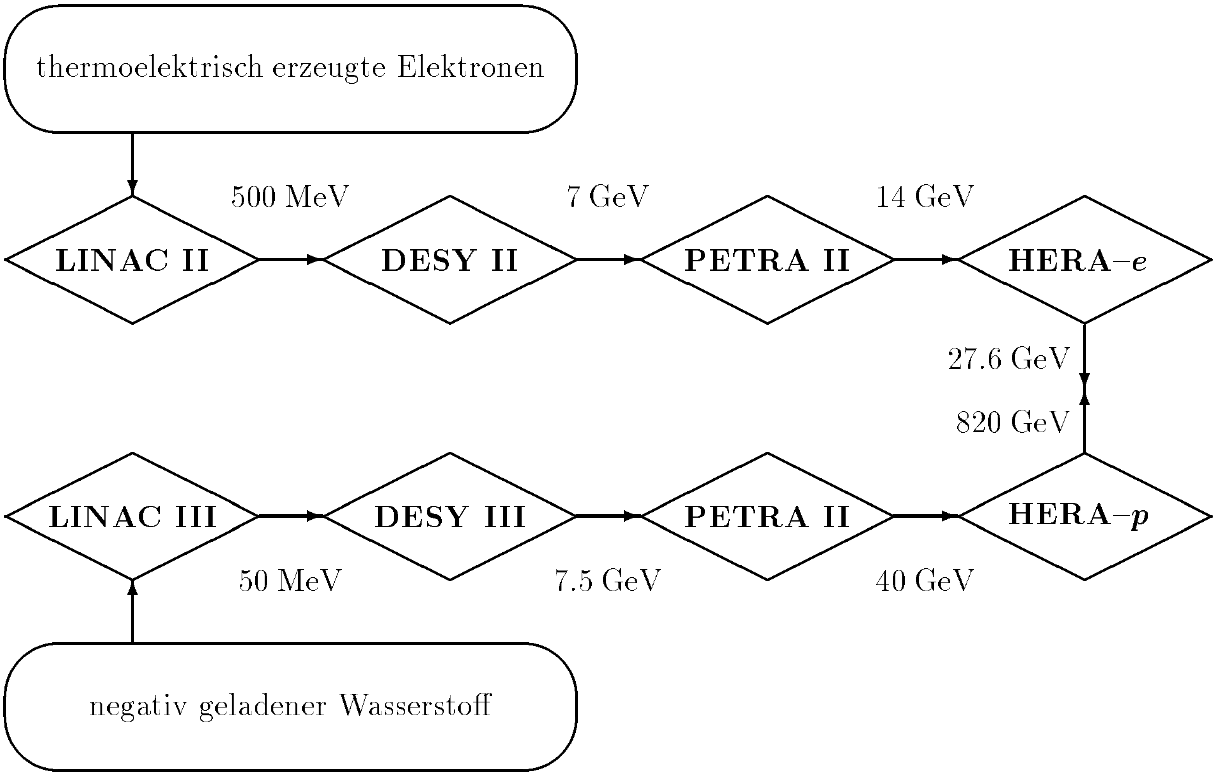

Abbildung 2.2: *Schematische Darstellung der Vorbeschleuniger bei HERA.*

Die Elektronen und Protonen werden an zwei Kreuzungspunkten zur Kollision gebracht. An zwei Wechselwirkungszonen, im Norden und im Süden des HERA–Ringes, wurden die beiden Detektoren H1 und ZEUS errichtet, um die Struktur des Protons aus der ep–Streuung bei höchsten Schwerpunktenergien zu bestimmen. In der Halle West ist das Experiment HERA–B im Aufbau, an dem die CP–Verletzung in B–Meson–Zerfällen untersucht werden soll. Das Experiment HERMES, das die Spinstruktur des Protons und des Neutrons untersucht, befindet sich in der östlichen Halle.

Die erreichbare Schwerpunktenergie des ep–Systems beträgt $\sqrt{s} \simeq 300$ GeV, so daß der maximal mögliche Viererimpulsübertrag $Q^2 \simeq 10^5\,\mathrm{GeV}^2$ erreicht. Damit liegt HERA etwa zwei Größenordnungen über den mit *Fixed–Target*–Experimenten an ruhenden Protonen bzw. Atomkernen zugänglichen Werten.

Für das Erreichen einer hohen *Luminosität* — geliefert wurden 1997 bis Ende August etwa 25 pb^{-1} — werden die Elektronen und Protonen auf je 210 Plätzen im Ring simultan gespeichert. Sind alle Plätze mit Teilchenpaketen (*Bunches*) besetzt, ergibt sich an den Wechselwirkungspunkten ein zeitlicher Abstand zweier *Bunchcrossings* von 96 ns entsprechend einer Frequenz von 10.4 MHz.

2.2 Der H1–Detektor

In der nördlichen Wechselwirkungszone von HERA ist der *Detektor H1* für die Analyse der ep–Streuung aufgebaut (Abbildung 2.3). Er wurde 1992 in Betrieb genommen und hat die Maße $b \times h \times l = 10\,\mathrm{m} \times 12\,\mathrm{m} \times 15\,\mathrm{m}$ bei einem Gewicht von 2800 t.

Der Detektor besteht aus mehreren Komponenten, die z.B. auf die Messung des Impulses von geladenen Teilchen, die Messung der Energie neutraler und geladener Teilchen und Teilchenschauer, die Erkennung des gestreuten Elektrons und die Identifizierung von Myonen spezialisiert sind. Dabei wird, bis auf die Öffnungen für das Strahlrohr, der gesamte 4π–Raumwinkel abgedeckt.

Das H1–Koordinatensystem

Der nominelle Wechselwirkungspunkt im Zentrum des H1–Detektors definiert den Ursprung eines rechtshändigen Koordinatensystems (Abbildung 2.4). Die Flugrichtung der einlaufenden Protonen legt die Richtung der z–Achse fest, die x–Achse zeigt horizontal zum Mittelpunkt des HERA–Ringes und die y–Achse zeigt nach oben. Der Polarwinkel Θ ist der Winkel bezüglich der positiven z–Achse und der Azimutalwinkel Φ ist der Winkel in der xy–Ebene zur x–Achse (Größen, die in der xy–Ebene — auch $R\Phi$–Ebene genannt — definiert sind, werden i.a. mit dem Zusatz „transversal“, abgekürzt mit dem Index t, bezeichnet).

p

e$^+$

$\boxed{1}$ Strahlrohr und Strahlmagnete $\boxed{2}$ Zentrale Spurkammern
$\boxed{3}$ Vorwärts–Spurkammern mit Übergangsstrahlungsmodulen
$\boxed{4}$ Elektromagnetisches Kalorimeter (Blei / Flüssig–Argon)
$\boxed{5}$ Hadronisches Kalorimeter (Edelstahl / Flüssig–Argon)
$\boxed{6}$ Supraleitende Spule ($B = 1.15$ T) $\boxed{7}$ Kompensationsmagnet ($B = 4.83$ T)
$\boxed{8}$ Helium–Kälteanlage $\boxed{9}$ Myon–Kammern
$\boxed{10}$ Instrumentiertes Eisenjoch (Eisenplatten und Streamerrohrkammern)
$\boxed{11}$ Myon–Toroidmagnet ($B = 1.6$ T)
$\boxed{12}$ rückwärtige Spurkammer und warmes Kalorimeter
$\boxed{13}$ Vorwärtskalorimeter $\boxed{14}$ Betonabschirmung
$\boxed{15}$ Flüssig–Argon–Kryostat

Abbildung 2.3: *Der H1–Detektor.*

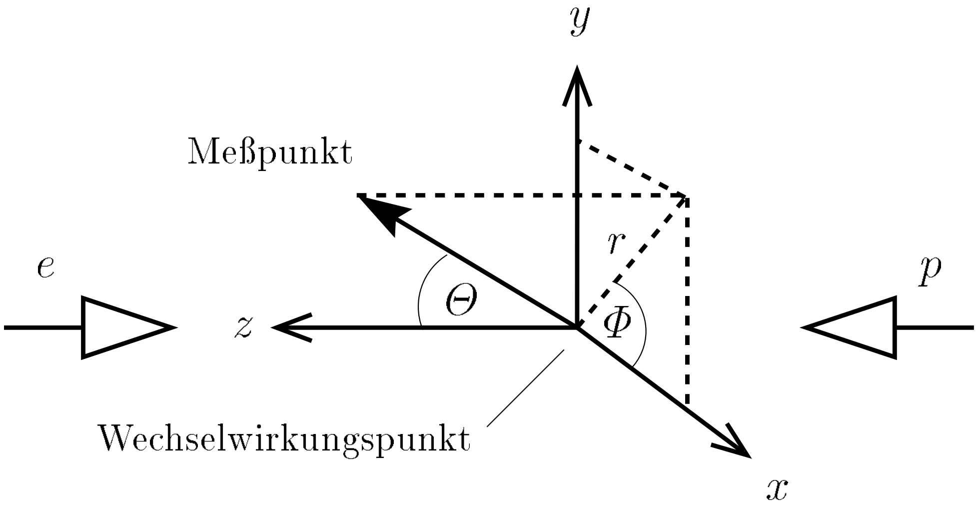

Abbildung 2.4: *Das H1–Koordinatensystem.*

Aufgrund des höheren Protonimpulses folgen die meisten bei der Kollision entstehenden Teilchen der ursprünglichen Protonrichtung. Deshalb ist der Detektor asymmetrisch aufgebaut und unter kleinen Polarwinkeln, im *Vorwärtsbereich*, aufwendiger instrumentiert als im *Rückwärtsbereich*.

Im Folgenden werden die für diese Analyse wichtigen Detektorkomponenten kurz beschrieben, wobei radial vom Wechselwirkungspunkt nach außen vorgegangen wird. Eine detailliertere Beschreibung und die verwendeten Angaben über die Auflösungen finden sich in [H1C97].

2.2.1 Das Spurkammersystem

Zur Rekonstruktion geladener Spuren durch Messung von Ort, Impuls und Ladung und zur Teilchenidentifikation durch Energieverlustmessungen dient das *Spurkammersystem* des H1–Detektors (Abbildung 2.5).

Das Spurkammersystem befindet sich zusammen mit dem Flüssig–Argon–Kalorimeter (Abschnitt 2.2.2) innerhalb einer supraleitenden Solenoid–Spule mit einem Radius von etwa 3 m. Durch ein parallel zum Protonstrahl verlaufendes homogenes Magnetfeld der Stärke $B = 1.15$ T bei $\Delta B/B \leq 2\%$ ermöglicht die Spule eine Impulsmessung von geladenen Teilchen mit einer Genauigkeit von $\sigma(p)/p^2 < 1\%$ GeV^{-1} im zentralen Bereich [H1C97].

Das Spurkammersystem unterteilt sich in drei Bereiche, das Vorwärts–Spurkammersystem, den zentralen Spurendetektor und die rückwärtige Driftkammer.

Das Vorwärts–Spurkammersystem Das *Vorwärts–Spurkammersystem* hat seinen Akzeptanzbereich für Polarwinkel von $7° \lesssim \Theta \lesssim 25°$ und besteht aus drei baugleichen

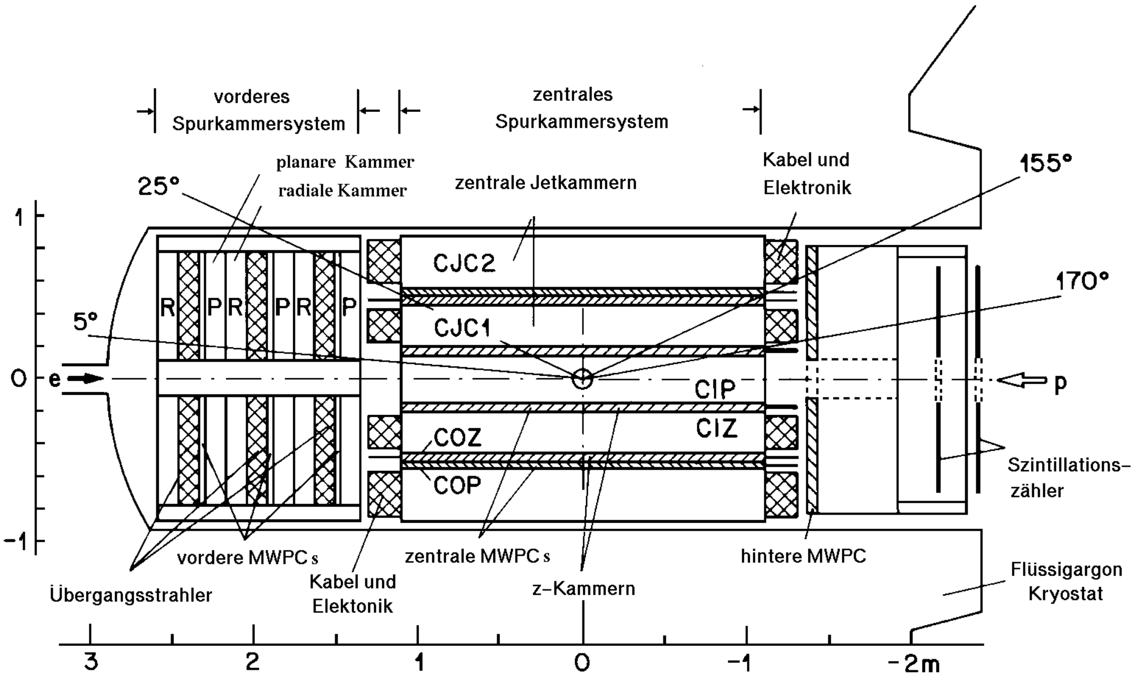

Abbildung 2.5: *Seitenansicht des Spurkammersystems.*

Supermodulen. Jedes Supermodul enthält, vom Wechselwirkungspunkt aus betrachtet, ein *planares Modul* bestehend aus drei, gegeneinander um 60° gedrehten Driftkammern, um Spuren eindeutig in der $R\Phi$–Ebene bestimmen zu können, eine *Vieldrahtproportionalkammer* (MWPC[1]), ein Übergangsstrahlungsmodul sowie eine Driftkammer mit in der $R\Phi$–Ebene radial nach außen verlaufenden Signaldrähten (*radiales Modul*).
Da aufgrund des kleinen Polarwinkels das effektive Magnetfeld schwächer wird, wurden die Module mit einer hohen Signaldrahtdichte ausgestattet. Die Ortsauflösung beträgt für die radialen Module $\sigma_{R\Phi} = 170\,\mu$m bzw. $\sigma_{xy} = 210\,\mu$m für die planaren Module [H1C97].

Zentrale Spurkammern Im zentralen Detektorbereich ($25^\circ \lesssim \Theta \lesssim 155^\circ$) befindet sich die *zentrale Spurkammer*, die sich aus sechs konzentrisch angeordneten Kammern (von innen nach außen : CIP,CIZ,CJC1,COZ,COP,CJC2[2]) zusammensetzt (Abbildung 2.6).

Von primärer Bedeutung für die Spurmessung sind die Driftkammern CJC1 und CJC2. Sie haben eine Ortsauflösung von $\sigma_{R\Phi} = 170\,\mu$m in der $R\Phi$–Ebene und $\sigma_z = 22$ mm entlang der Strahlachse, sowie eine Genauigkeit der Energieverlustmessung zur Teilchenbestimmung von $\sigma(dE)/dE = 10\%$. Die Driftkammern CIZ und COZ dienen zur Messung der z–Koordinate der Spuren und haben eine Ortsauflösung von $\sigma_z \approx 350\,\mu$m, während die Vieldraht–Proportionalkammern CIP und COP ausschließlich schnelle Triggersignale erzeugen (Abschnitt 2.2.5).

---

[1] **M**ulti**w**ire **P**roportional **C**hamber

[2] CIP/COP = **C**entral **I**nner/**O**uter **P**roportional Chamber,
CIZ/COZ = **C**entral **I**nner/**O**uter **Z**–Chamber,
CJC = **C**entral **J**et **C**hamber

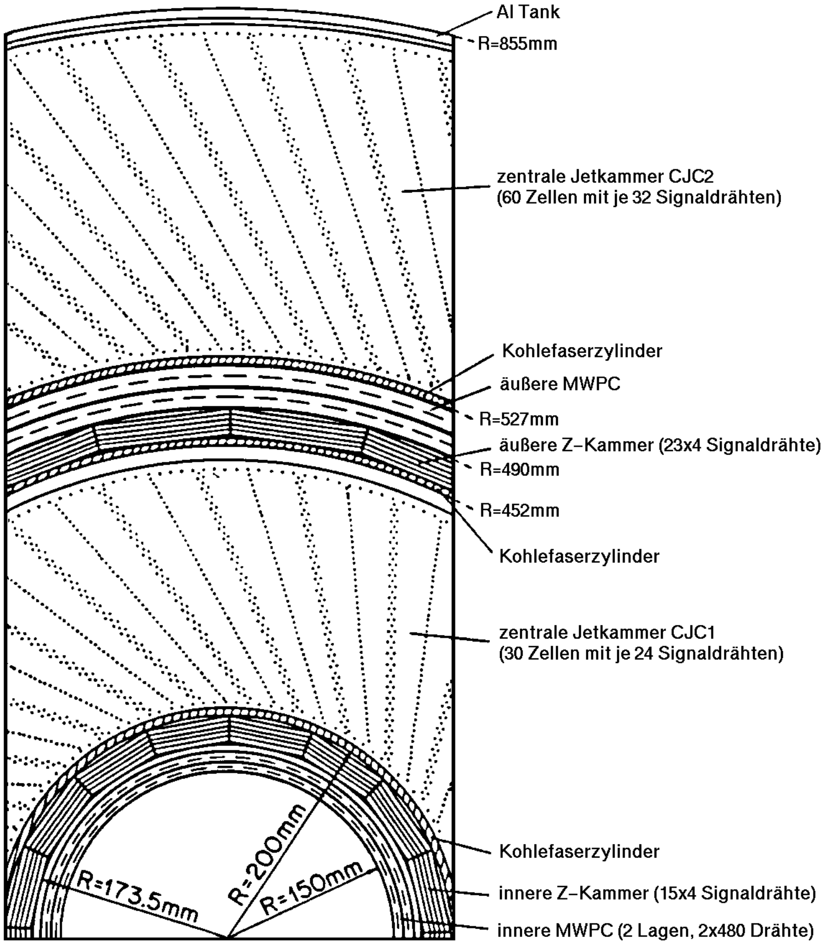

Abbildung 2.6: *Querschnitt durch das Spurkammersystem.*

Rückwärtige Spurkammer Durch die *rückwärtige Driftkammer* wird der dem Spurkammersystem zugängliche Polarwinkelbereich auf $\Theta \simeq 177^\circ$ erweitert, so daß insbesondere Elektronen aus der *ep*–Streuung mit Viererimpulsüberträgen von $1\ \mathrm{GeV}^2 \lesssim Q^2 \lesssim 100\ \mathrm{GeV}^2$ rekonstruiert werden können.

2.2.2 Kalorimetrie

Das Spurkammersystem ist von vier Kalorimetern umgeben, dem Flüssig–Argon–Kalorimeter im zentralen Bereich, dem SPACAL im rückwärtigen Bereich und dem Plug–Kalorimeter in extremer Vorwärtsrichtung. Weiter außerhalb folgt der Tail–Catcher. Sie dienen zur Messung der Energie und Richtung von geladenen und neutralen Teilchen bzw. Teilchenschauern. Durch die Unterteilung in elektromagnetisches und hadronisches Kalorimeter können Elektronen und Photonen von hadronischen Schauern unterschieden werden.

Die Kalorimeter (bis auf den Tail–Catcher) sind noch innerhalb der supraleitenden Spule aufgebaut, um die Dicke des inaktiven Detektormaterials, das Teilchen durchqueren müssen, um zum Kalorimeter zu gelangen, zu minimieren.

Das Flüssig–Argon–Kalorimeter Im zentralen Bereich ($4^\circ \lesssim \Theta \lesssim 154^\circ$) befindet sich in einem Kryostaten das *Flüssig–Argon–Kalorimeter* (LAr–Kalorimeter[3]). Es ist in einen elektromagnetischen und einen hadronischen Teil unterteilt (Abbildung 2.7).

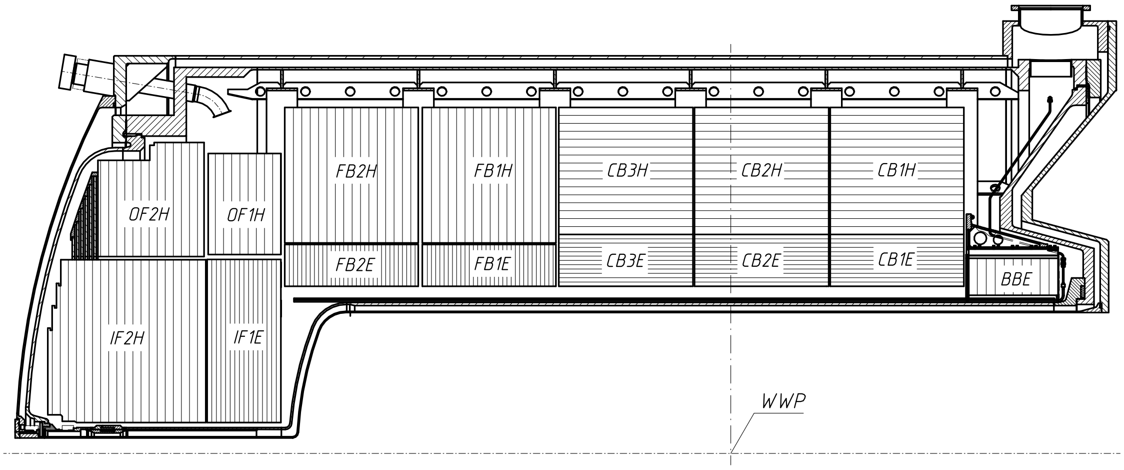

Abbildung 2.7: *Seitenansicht des Flüssig–Argon–Kalorimeters.*

Der elektromagnetische Teil des Flüssig–Argon–Kalorimeters bildet die innere Schicht mit einer Dicke von 20 bis 30 Strahlungslängen X_0. Es setzt sich aus 2.4 mm starken Bleiplatten als Absorbermaterial und dazwischen befindlichen 2.35 mm breiten Spalten mit flüssigem Argon als Ionisationsmaterial zusammen. Die Energieauflösung des elektromagnetischen Teils beträgt $\sigma(E)/E = 11\%/\sqrt{E} \oplus 1\%$ [H1C97].
Nach außen anschließend liegt der hadronische Teil. Als Absorbermaterial werden 15.5mm dicke Edelstahlplatten verwendet, dazwischen liegen ebenfalls 2.35 mm breite Spalten mit flüssigem Argon. Die in Testmessungen mit Pionen bestimmte Energieauflösung beträgt $\sigma(E)/E = 50\%/\sqrt{E} \oplus 2\%$ [H1C97].
Insgesamt besteht das Flüssig–Argon–Kalorimeter aus etwa 45 000 *Zellen*, die eine gute Ortsauflösung ermöglichen. Abhängig vom Polarwinkel hat das gesamte Kalorimeter für ein vom Wechselwirkungspunkt kommendes Teilchen eine Dicke von 5 hadronischen Wechselwirkungslängen λ im Zentralbereich und bis zu 8 Wechselwirkungslängen im Vorwärtsbereich (Abbildung 2.8).

Das Spaghetti–Kalorimeter Im rückwärtigem Bereich wurde Anfang 1995 das „warme“ *Spaghetti–Kalorimeter* SPACAL[4] eingebaut (Abbildung 2.9).

Das SPACAL überdeckt einen Polarwinkelbereich von $151^\circ \lesssim \Theta \lesssim 178^\circ$ und besteht wie das Flüssig–Argon–Kalorimeter aus einem hadronischen und einem elektromagnetischen Teil. Im elektromagnetischen Teil kann das gestreute Elektron in Reaktionen mit $Q^2 \lesssim 100$ GeV^2 nachgewiesen werden.

---

[3] **L**iquid **Ar**gon Calorimeter
[4] **Spa**ghetti–**Cal**orimeter

Abbildung 2.8: *Abhängigkeit der Anzahl der durchquerten Wechselwirkungslängen λ vom Polarwinkel im Flüssig–Argon–Kalorimeter.*

Des weiteren existieren das *Plug–Kalorimeter* für den extremen Vorwärtsbereich ($0.7^\circ \lesssim \Theta \lesssim 3.2^\circ$) und der Tail–Catcher, der sich in dem im nächsten Abschnitt beschriebenen instrumentierten Eisen befindet. Beide Kalorimeter werden in der vorliegenden Analyse nicht benutzt.

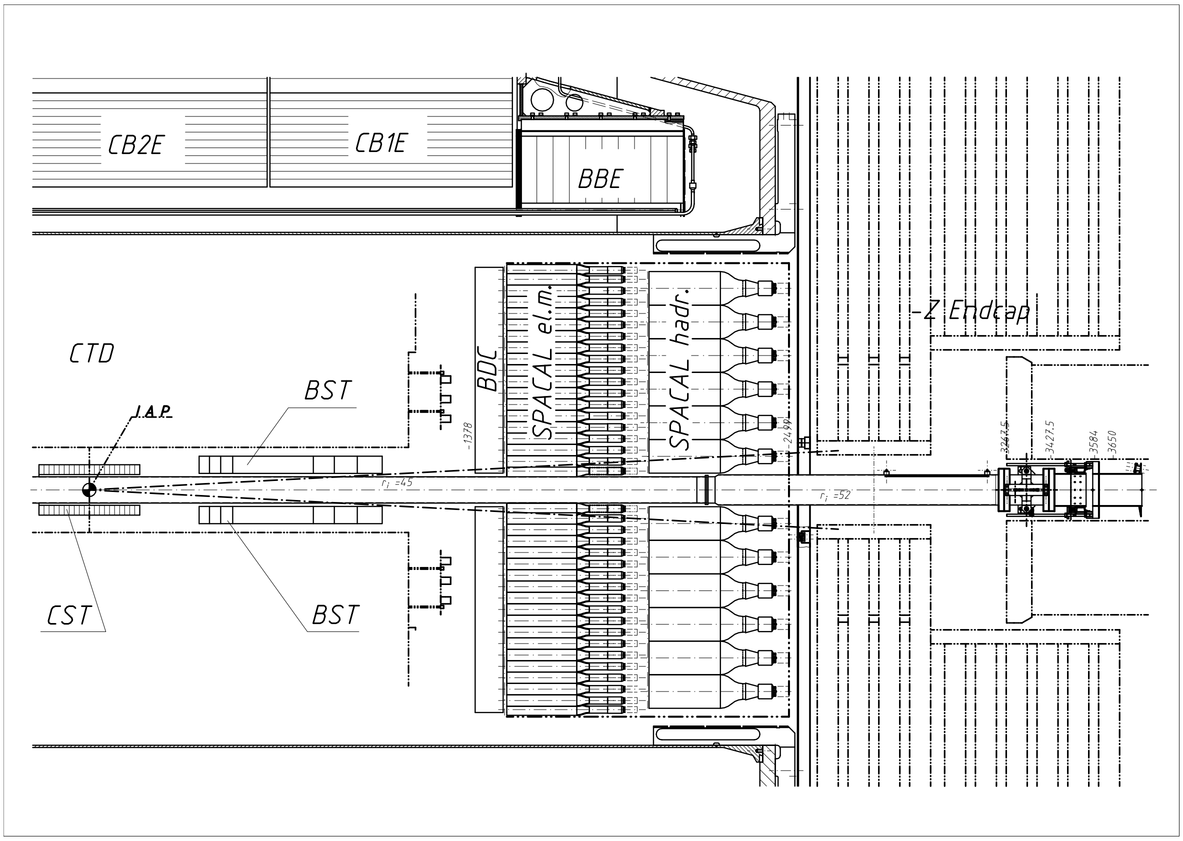

Abbildung 2.9: *Seitenansicht des SPACALs.*

2.2.3 Das instrumentierte Eisen

Außerhalb der Kalorimeter folgt die supraleitende Spule und anschließend das *Eisenjoch* zur Rückführung des magnetischen Flusses. Es gliedert sich in vier Bereiche : den vorderen und hinteren *Zentralbereich* ($35° \lesssim \Theta \lesssim 130°$), der *vorderen Endkappe* (FEC[5], $4° \lesssim \Theta \lesssim 35°$) und der *hinteren Endkappe* (BEC[6], $130° \lesssim \Theta \lesssim 171°$). Gleichzeitig wird das Eisenjoch als aktiver Detektor genutzt (*instrumentiertes Eisen*, Abbildung 2.10).

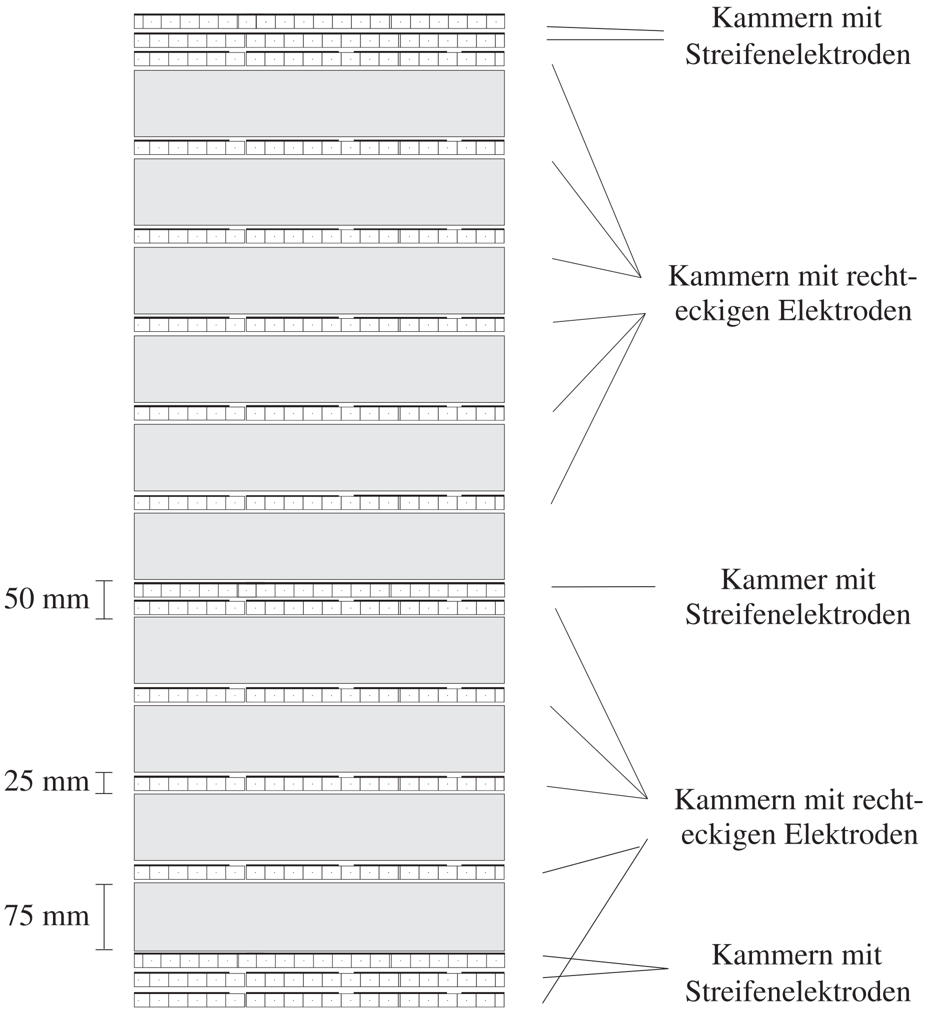

Abbildung 2.10: *Querschnitt durch das instrumentierte Eisen.*

Das Joch ist in zehn jeweils 7.5 cm dicke Eisenplatten unterteilt, die in den dazwischenliegenden Schlitzen von 2.5cm Breite mit *Streamerrohrkammern* instrumentiert sind. Zusätzlich zu diesen zehn Lagen — eine Lage ist doppelt bestückt — sind auf der Innen– und der Außenseite drei weitere Lagen Streamerrohrkammern (*Myonboxen*) angeordnet, so daß ein Teilchen maximal 16 Lagen durchqueren kann.

---

[5] **F**orward **E**ndcap
[6] **B**ackward **E**ndcap

Die Streamerrohrkammern dienen einerseits zur Identifizierung und Impulsbestimmung von Myonen und andererseits zur Energiemessung des Anteils von hadronischen Schauern, der aus dem Flüssig–Argon–Kalorimeter entweicht und ins Eisen gelangt (*Tail Catcher*). Die Kammern sind mit Gas gefüllt und haben einen Querschnitt von 1 cm × 1 cm. In der Mitte verläuft ein Kupfer–Beryllium–Draht, der im Zentralbereich parallel zur Protonrichtung und in den Endkappen entlang der x–Achse gespannt ist. Jeweils acht Stramerrohrkammern bilden ein *Profil*, zwei Profile werden zu einem gasdicht verschlossenen *Element* zusammengefaßt.
Zusätzlich zu den Streamerrohrkammern sind auf jeweils zwei Lagen der Myonboxen und auf eine der Doppellagen streifenförmige Influenzelektroden im Abstand von 20mm aufgeklebt. Diese ermöglichen die Messung einer zweiten Ortskoordinate. Die anderen elf Lagen tragen rechteckige Influenzelektroden für den Tail–Catcher.
Die Ortsauflösung beträgt für die Drähte $\sigma_{wire} = 3\ldots4\,\mathrm{mm}$ und $\sigma_{strip} = 10\ldots15\,\mathrm{mm}$ für die Streifen; die Impulsauflösung beträgt $\sigma(p)/p \approx 0.35$ und die Effizienz der Myonerkennung beträgt für die einzelnen Lagen 80% [H1C97].

Für die vorliegende Analyse werden nur Informationen aus dem Zentralbereich verwendet. Die Endkappen werden nicht weiter berücksichtigt.

2.2.4 Weitere Komponenten des H1–Detektors

Das Vorwärts–Myon–System Neben dem instrumentierten Eisen steht im Vorwärtsbereich ($3° \lesssim \Theta \lesssim 17°$), vor dem instrumentierten Eisen, zusätzlich der *Vorwärts–Myon–Detektor* FMD[7] zur Erkennung von Myonen mit einem Impuls oberhalb von 5 GeV zur Verfügung. Er besteht aus drei Driftkammern vor und hinter einem Eisen–Toroidmagneten mit einem Magnetfeld der Stärke $B(r) \approx 1.5\ldots3.5$ T. Die Impulsauflösung beträgt $0.25 < \sigma(p)/p < 0.32$ [H1C97].
Informationen aus dem Vorwärts–Myonspektrometer wurde in dieser Analyse nicht benutzt.

Proton–Tagger und Vorwärts–Neutron–Kalorimeter Protonfragmente, die die Wechselwirkungszone unter so kleinen Polarwinkeln verlassen, daß sie im H1–Detektor nicht mehr nachgewiesen werden können, werden zum Teil mit Hilfe des *Proton–Taggers* bei $z = 24$ m nachgewiesen.
Zum Nachweis von Neutronen befindet sich bei $z = 106$ m, unterhalb des Protonstrahlrohres, ein *Vorwärts–Neutron–Kalorimeter*.

Das Vorwärts–Proton–Spektrometer Zum Nachweis von Protonen aus der diffraktiven Streuung befindet sich weit vor dem H1–Detektor das *Vorwärts–Proton–Spektrometer*. Es besteht aus den Ablenkmagneten von HERA–p und vier Detektoren, die in beweglichen Tauchgefäßen („*römischen Töpfen*“) vertikal (bei $z = 81$ m und $z = 90$ m) bzw. horizontal (bei $z = 64$ m und $z = 80$ m) an den Protonstrahl herangefahren werden.

---

[7] **F**orward **M**uon **D**etector

Das Luminositätssystem Die Bestimmung der absoluten Luminosität in der Wechselwirkungszone geschieht über den Bethe–Heitler–Bremsstrahlungsprozeß $e\ p \rightarrow e\ p\ \gamma$, dessen Wirkungsquerschnitt bekannt ist. Zu diesem Zweck sind bei $z = -33.4$ m der *Elektron–Tagger* und bei $z = -102.9$ m der *Photon–Tagger* installiert.
Der Elektron–Tagger erlaubt zusätzlich die Erkennung des gestreuten Elektrons bei sehr kleinen Viererimpulsüberträgen von $Q^2 < 0.01$ GeV^2.

Flugzeitdetektoren und Vetowände Zur Identifikation von protoninduzierten Untergrund befinden sich innerhalb des Toroid–Magneten, innnerhalb des Plug–Kalorimeters und zwischen SPACAL und der hinteren Endkappe die *Flugzeitdetektoren*. In rückwärtiger Richtung stehen bei $z = -6.5$ m und $z = -8.1$ m die *Vetowände*.

2.2.5 Das Trigger– und Datennahmesystem

Bei HERA treffen in den Wechselwirkungszonen alle 96 ns Elektronen– und Protonen–Pakete aufeinander. Die mögliche Rate von *Ereignissen* liegt damit bei 10.4 MHz. Da jedoch die Ausleseelektronik, der Datentransfer und die Speicherkapazitäten Ereignisraten von höchstens 10 Hz aufnehmen können, ist es Aufgabe des *Triggersystems*, physikalisch interessante Ereignisse zu behalten und Untergrundereignisse zu verwerfen. Die bei HERA dominierenden Untergrundprozesse sind :

- Streuung von Strahlprotonen am Strahlrohr selbst oder an anderen Teilen des Speicherringes (Magneten, Kollimatoren) die zu einer Ereignisrate von etwa 50 kHz führen [Kru94].
- Streuung von Strahlprotonen an Molekülen des Restgases im Strahlrohr im Bereich des H1–Detektors. Diese ist um ein Vielfaches häufiger als die *ep*–Streuung.
- Halo–Myonen, die parallel zum Strahl fliegen.
- Synchrotronstrahlung von Elektronen, die zu Treffern in der zentralen Spurkammer führen kann.
- Myonen aus der kosmischen Höhenstrahlung, die den Detektor durchqueren. Die Rate liegt bei etwa 2 kHz.

Zur Reduktion des Untergrundes wird bei H1 ein vierstufiges Triggersystem angewendet, dessen einzelne Stufen mit ***Level 1*** bis *Level 4* (L1 bis L4) bezeichnet werden. Akzeptiert eine Stufe ein Ereignis, so wird die nachfolgende Triggerstufe gestartet, wobei nachfolgende Stufen vorherige Entscheidungen verifizieren oder verwerfen.

L1: Die erste Triggerstufe L1 bildet logische Verknüpfungen (*Subtrigger*) aus den Triggersignalen der einzelnen Detektorkomponenten. Während der Entscheidungszeit (maximal 2.3 μs) werden die Daten in eine *Pipeline* gefüllt, so daß L1 totzeitfrei arbeitet. Akzeptiert L1 ein Ereignis (*L1–Keep*), wird die Pipeline gestoppt und die Signalauslese und damit die Totzeit der Datennahme beginnt. Die Ereignisrate soll mit L1 auf höchstens 1 kHz reduziert werden.

L2: Der L2–Hardwaretrigger führt während der beginnenden Auslese mit Hilfe topologischer Korrelationen und Neuronaler Netzwerke eine Überprüfung der L1–Subtrigger durch. Die Triggerentscheidung erfolgt nach etwa 20 μs. Eine negative Entscheidung bricht die Datenauslese ab und leitet einen Neustart der Datennahme ein. Die Ereignisrate wird durch L2 auf 50 Hz verringert.

L3: Der Softwaretrigger L3 soll innerhalb von etwa 800 μs durch genauere Analyse der Triggerdaten und mit Hilfe zusätzlicher Informationen aus den Detektorkomponenten die Ereignisrate auf 50 Hz reduzieren. L3 war bis 1997 nicht implementiert.

L4: Im Gegensatz zu den ersten drei Triggerstufen stehen dem L4–Trigger die vollständige Information aller Detektorkomponenten zur Verfügung. Mit Hilfe eines Multiprozessorsystems (*Filter–Farm*) wird durch schnelle Filteralgorithmen eine vereinfachte Ereignisrekonstruktion und weitere Datenreduktion (10 Hz Ereignisrate) durchgeführt. Die Entscheidung des L4–Triggers liegt nach etwa 1 s vor.

Eine umfassende Beschreibung der Datenahme und des H1–Triggersystems befindet sich in [Els93].

Rekonstruktion

Die Ereignisse, die von den vier Triggerstufen akzeptiert wurden, werden als *Rohdaten* auf Magnetbändern zwischengespeichert. Erst dann erfolgt, unabhängig von der Datennahme, die vollständige *Rekonstruktion* durch das Programmpaket `H1REC`, so daß die Daten innerhalb weniger Stunden zur Verfügung stehen. Die Rekonstruktion kann wie folgt grob gegliedert werden :

- Diejenigen Treffer in einem Spurkammersystem, die sich gemeinsam durch eine *Helix* parametrisieren lassen, werden zu einer Spur verbunden. Danach werden Spuren in der CJC1 und CJC2 sowie aus den Spurmodulen des vorderen Spurkammersystems verbunden. Aus der Krümmung und der Richtung der Spur wird der Impuls geladener Teilchen bestimmt.
- Die Position des primären Vertex in der xy–Ebene wird durch geladene Teilchen mit hohem Impuls aus mehreren Ereignissen errechnet. Die z–Koordinate errechnet sich für jedes Ereignis aus den Impulsen aller geladenen Teilchen.
- Die Signale der Kalorimeterzellen werden zu *Clustern* gruppiert.
- Die Signale der Myon–Detektoren werden zu Myonspuren verbunden.
- Treffer im instrumentierten Eisen werden mit den Spuren des Spurkammersystems verbunden. Dazu wird die Spurkammer–Spur bis in das Eisen extrapoliert und die Wahrscheinlichkeit $P(\chi^2)$, daß die Treffer in den Lagen des instrumentierten Eisens und die im Spurkammersystem gemessene Spur vom gleichen Teilchen kommen, errechnet. Damit die Spuren aneinander angepaßt werden, muß die Wahrscheinlichkeit größer als 10^{-3} sein. Dabei sind Anpassungen mehrerer innerer Spuren an eine Eisenspur möglich und umgekehrt.

- Analog werden Spurkammerspuren mit Spuren im Vorwärts–Myon–System sowie mit Kalorimeter–Clustern verbunden.

Eine nachfolgende Vorselektion (*L5–Klassifikation*) ordnet die rekonstruierten Ereignisse anhand einfacher Bedingungen in physikalische Klassen ein, die anschließend von der *Ereignisklassifikation* einer verschärften Klassifikation und Reduktion unterworfen werden. Die so selektierten Ereignisse stehen zur physikalischen Analyse auf Festplatten zur Verfügung.

Kapitel 3

Theorie der Neuronalen Netze

> Cogito ergo sum.
> Ich denke, also bin ich.
> *René Descartes*

Seit Computer existieren, ist die Künstliche Intelligenz, die Kunst analog zum menschlichen Hirn zu denken, selbstständig Muster zu erkennen und zu lernen, Ziel vieler Programmierer. Aber solange das Gehirn mit seinen abermilliarden Verbindungen, seinen Funktionsabläufen und seiner Informationsspeicherung unverstanden ist, bleiben nur einfachere Konzepte als Ansatz. Das künstliche ***N**euronale* ***N**etz* (NN) ist eines dieser Konzepte. Basierend auf der Struktur des Hirns, kopiert es seine kleinsten Bestandteile, die Neuronen, und faßt sie in einem mathematischen Konstrukt zusammen. Die Determiniertheit der Mathematik schließt Künstliche Intelligenz aus, gleichzeitig bedingt sie aber die Musterfindung nach intensivem Lernvorgang und erlaubt so eine (jedoch weit geringere) Nachahmung des Gehirns.

Abschnitt 1 behandelt die Grundlagen im Aufbau des menschlichen Hirns, Abschnitt 2 überträgt das biologische Konzept auf Neuronale Netze und faßt die gewonnenen Erkenntnisse in einem mathematischen Konstrukt zusammen. Das Konzept von Neuronalen Netzen soll an einem einfachen Beispiel, der Erkennung des Ausgabewertes der XOR–Funktion, verdeutlicht werden (Abschnitt 3). Zum Schluß erfolgt die Beschreibung der programmtechnischen Umsetzung mit dem Programmpaket `JETNET` (Abschnitt 4).

3.1 Biologische Nervenzellen

Der menschliche Denkapperat, das Gehirn, besteht aus 10^{10} bis 10^{11} Nervenzellen, den *Neuronen*, von denen jede mit hunderten bis tausenden anderer Neuronen in Verbindung steht, und die sich gegenseitig über elektrische Signale beeinflussen. Ein Neuron besteht aus einem Zellkörper (*Soma*) mit vielfach verzweigten kurzen Fortsätzen (*Dendriten*) und einer langen Nervenfaser (*Axon*) (Abbildung 3.1).

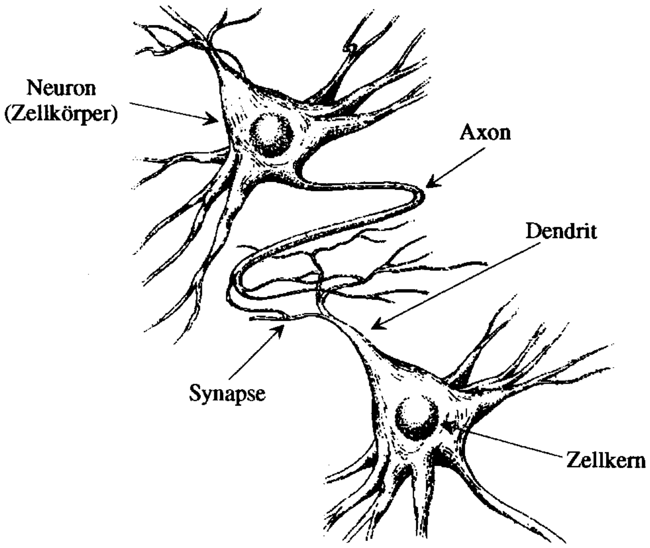

Abbildung 3.1: *Aufbau und Verknüpfung zweier biologischer Neuronen.*

Im Ruhezustand hält es auf seiner Zellmembran eine Potentialdifferenz aufrecht. Dies liegt an der erhöhten Konzentration von Natrium–Ionen außerhalb des Zellkerns. Wird das Neuron über seine Dendriten erregt, ändert sich die Durchlässigkeit der Membran und die Natrium–Ionen strömen in den Zellkern. Überschreiten sie dort eine bestimmte Konzentration, baut sich am *Axonhügel*, dem Ursprung der Nervenfaser am Zellkörper ein Aktionspotential auf. Das Zellinnere wird kurzzeitig positiv gegenüber dem Zelläußeren, das Neuron „feuert". Innerhalb einer Millisekunde kehrt das Potential wieder in seinen Ruhezustand zurück, indem es unter Aufwand von Energie die Natrium–Ionen aus dem Zellkern entfernt werden (Ionen–Pumpe).
Das Aktionspotential baut sich entlang des Axons in ähnlicher Weise auf und gelangt so zu den Verbindungen mit anderen Neuronen, den *Synapsen*. Allerdings verhindert der *synaptische Spalt* ein Überspringen des Potentials auf deren Dendriten. Vielmehr werden durch das Aktionspotential chemische Botenstoffe, *Neurotransmitter*, in den synaptischen Spalt ausgeschüttet, die zum Dendriten hinüberdiffundieren und sich dort an spezifische *Rezeptoren* heften. Auf diese Weise können sich Ionenkanäle öffnen. Ob ein Transmitter das Neuron hemmt oder erregt, hängt von der Art des Rezeptors ab.

Die Natur der Vorgänge, die sich im Neuron abspielen, bringt es mit sich, daß eine Nervenzelle nicht öfter als 200 mal in der Sekunde feuern kann. Die Ausbreitungsgeschwindigkeit des Aktionspotientials über das Axon beträgt maximal 100 m/s. Diese Betrachtung und die Erkenntnis, daß sich wesentliche Leistungen des menschlichen Gehirns in weniger als einer zehntel Sekunde abspielen, lassen auf eine massive Parallelität von Neuronenfunktionen schließen [Nau94].

Die Leistung des Menschen zu *lernen* ist mit der Struktur und Funktionsweise biologischer Neuronennetze allein nicht zu erklären. Die ersten Ideen zu diesem Problem formulierte

D.O. Hebb [Heb49]. Hebb postulierte, daß ein *Lernvorgang* die Verbindung zweier Neuronen verändert und damit die Gewichtung der Synapse wachsen oder abnehmen läßt, je nachdem wie häufig diese Verbindung zur Signalübertragung Verwendung findet. Dieser Lernvorgang wird als *Hebb'sche Lernregel* bezeichnet.

Die gewonnene Erkenntnis über den Aufbau und die Funktionsweise der Neuronennetze kann nun auf künstliche Strukturen übertragen werden.

3.2 Das mathematische Konstrukt

Aus der Biophysik der Neuronen lassen sich folgende Prinzipien gewinnen :

- Ein Neuron ist ein binäres Schaltelement, das entweder aktiv oder inaktiv ist.
- Jedes Neuron besitzt einen festen Schwellenwert.
- Ein Neuron empfängt Eingaben von erregenden bzw. hemmenden Synapsen. Diese sind nach ihrer Signifikanz gewichtet.
- Es gibt einen Zeitpunkt für die Integration aller synaptischen Eingaben. Das Neuron wird aktiv, wenn die Summe aller Eingaben seinen *Schwellenwert* überschreitet.
- Der Lernvorgang variiert die Gewichte der Synapsen in Abhängigkeit ihrer Beanspruchung.

Dieses Modellneuron entspricht nur noch eingeschränkt seinem biologischen Äquivalent. Seine wesentlichen Grundzüge entspringen der theoretischen Vorstellung von W. Pitts und W.S. McCulloch [McC43]. Es wird daher auch *McCulloch–Pitts–Neuron* genannt.

3.2.1 Übertragung auf ein künstliches Neuron

Die grundlegenste Form eines Neuronalen Netzes ist die mathematische Implementierung eines einzelnen McCulloch–Pitts–Neurons, das *Perceptron* [Ros58]. Dazu werden die oben aufgezählten biophysikalischen Prinzipien direkt auf die mathematische Ebene übertragen.

Die Binarität läßt sich mathematisch durch die Heaviside'sche Sprungfunktion

$$g(z) = \begin{cases} 1 & \text{für} \quad z > 0 \\ 0 & \text{für} \quad z < 0 \end{cases} \tag{3.1}$$

darstellen. Für den Lernvorgang wird es sich als notwendig erweisen, $g(z)$ als stetig differenzierbar zu wählen. Sie wird als *Aktivierungsfunktion* bezeichnet und gewöhnlich wie folgt gewählt :

$$g(z) = \left(1 + e^{-2z}\right)^{-1} \tag{3.2}$$

Die synaptischen Eingaben werden durch die *Eingabegrößen* $x_k \in I\!R$ realisiert, der Index bezeichnet dabei die k–te Größe; die Gesamtanzahl aller Eingaben beträgt N_x. Die Gewichtung der Synapse geschieht durch einen Faktor $w_k \in I\!R$, wobei $w_k < 0$ hemmend und $w_k > 0$ anregend ist. Durch Vergleich der Summe aller Eingabegrößen des Perceptrons mit seinem Schwellenwert θ, ergibt sich der *Ausgabewert* o des Perceptrons :

$$o := g(z) = g\left(\sum_k w_k \, x_k - \theta\right) \tag{3.3}$$

Der Ausgabewert $o \in I\!R$ nimmt entsprechend der Aktivierungsfunktion Werte zwischen 0 und 1 an. Mit dieser Funktion läßt sich ein gegebener Satz von Eingabegrößen in die zwei *Klassen* „0“ und „1“ einordnen. Das Schema des Perceptrons ist in Abbildung 3.2 dargestellt.

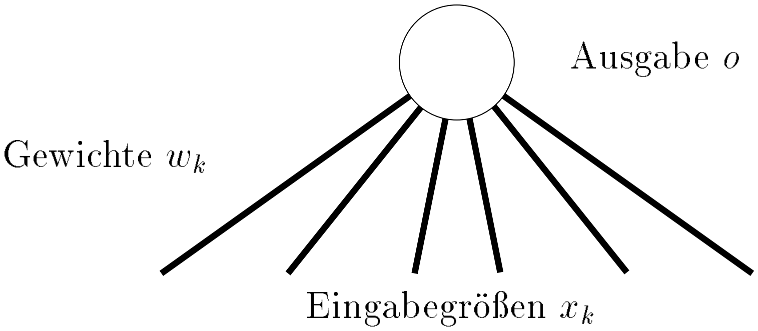

Abbildung 3.2: *Schematische Darstellung eines Perceptrons.*

Werden alle w_k und x_k durch entsprechende Vektoren $\vec{w}$ und $\vec{x}$ ausgedrückt, so ergibt $\vec{w} \cdot \vec{x} - \theta = 0$ eine Ebene im Hyperraum der Eingabegrößen $I\!R^{N_x}$. Diese *Hyperebene* trennt die beiden Klassen. Allerdings liefert ein Perceptron nur eine Hyperebene, weshalb die Funktion 3.3 nur linear trennbare Klassen teilen kann (Abbildung 3.3).

3.2.2 Feedforward–Netzwerke

Um die Trennung komplexerer, nicht linear separabler Klassen zu ermöglichen, wird das Perceptron–Modell zu mehrschichtigen Neuronalen Netzwerken erweitert (Abbildung 3.4).

Die Eingabegrößen x_k durchlaufen hier jeweils ein Perceptron, bei mehrschichtigen Neuronalen Netzen mit *Knoten* bezeichnet, mit der Schwelle $\theta_k = 0$ und der linearen Aktivierungsfunktion $g(z) = z$. Die Schicht der Eingabegrößen wird *Eingabelage* genannt. Die Knoten des Ausgabewertes o_i (**o**utput nodes) heißen *Ausgabelage*, wobei N_o die Gesamtanzahl aller Ausgabeknoten angibt. Es genügt $N_o = 1$ zur Unterscheidung von zwei Klassen. Mehrere Ausgabeknoten ($N_o > 1$) erlauben die Trennung von mehr als zwei Klassen (z.B. zur Datenkodierung oder Datenkompression [Roj91]). Die möglichen Lagen zwischen Eingabe– und Ausgabelage werden *versteckte Lagen* genannt. Prinzipiell

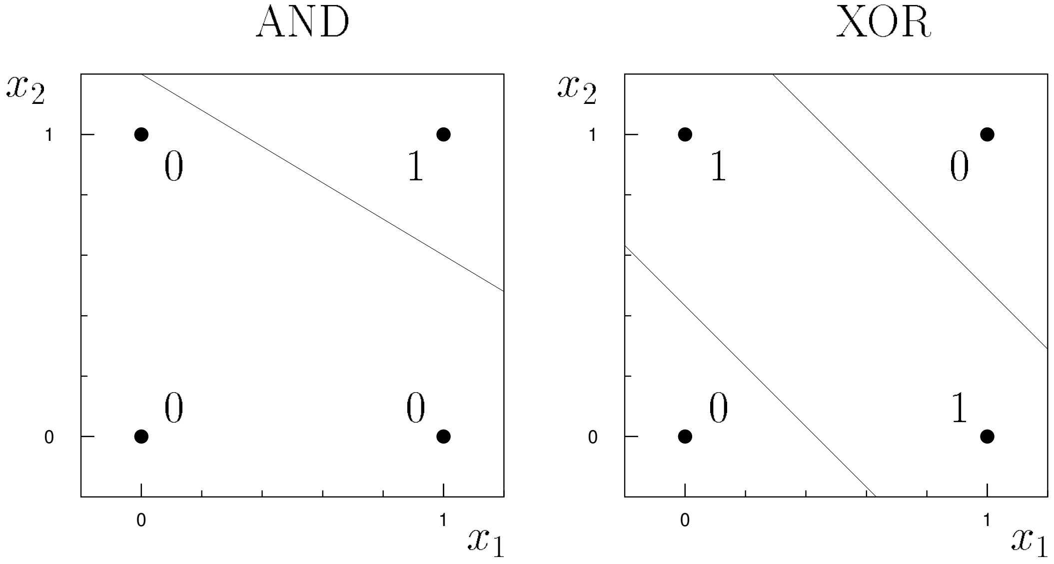

Abbildung 3.3: *Hyperebenen zur Trennung der AND– und XOR–Antwortklassen. Die Klassen „0“ und „1“ der AND–Funktion sind linear separierbar (links), die Klassen der XOR–Funktion hingegen nicht (rechts). Es werden zwei Hyperebenen zur Trennung benötigt.*

reicht eine versteckte Lage für jedes Mustererkennungsproblem aus (*Kolmogorovs Theorem* [Kol57]). Dies gilt speziell für Probleme der Hochenergie–Physik [Lon93]. Mehrere Ebenen können aber die Trennung der Klassen vereinfachen bzw. verbessern. Hier betrachtete Neuronale Netze besitzen von nun an *genau eine* versteckte Lage, ihre Knoten werden mit h_j (**h**idden nodes) abgekürzt. Die Anzahl der Knoten wird mit N_h bezeichnet.

Die Knoten der verschiedenen Lagen sind untereinander mit gewichteten Verbindungen verknüpft. Es wird zwischen *Feedforward–* und *Feedback–Netzen* unterschieden. Im Falle eines Feedforward–Netzes erhält jeder Knoten seine Eingabe aus der untergeordneten La-

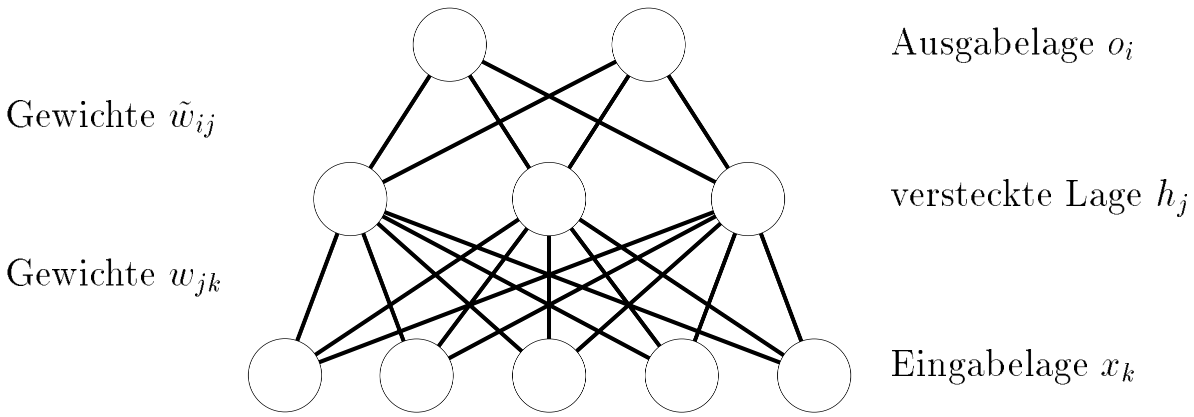

Abbildung 3.4: *Ein mehrlagiges Neuronales Netzwerk mit Feedforward–Architektur.*

ge und gibt seine Ausgabe an die übergeordnete Lage weiter. Es gibt keine Rückkopplung zwischen übergeordneter und untergeordneter Lage, was gerade in Feedback–Netzen realisiert ist (Abbildung 3.4 zeigt demnach ein Feedforward–Netz).

Mit der Einschränkung auf eine versteckte Lage erweitert sich Gleichung 3.3 zu

$$o_i = g\left(\sum_j \tilde{w}_{ij}\, h_j - \tilde{\theta}_i\right) \tag{3.4}$$

für die Ausgabeknoten und

$$h_j = g\left(\sum_k w_{jk}\, x_k - \theta_j\right) \tag{3.5}$$

für die einzelnen versteckten Knoten, mit der bekannten Aktivierungsfunktion $g(z)$ (Gleichung 3.2). $\tilde{w}_{ij}$ bzw. w_{jk} sind die gewichteten Verbindungen der einzelnen Knoten zwischen versteckter Lage und Ausgabelage bzw. Eingabelage und versteckter Lage. Die Summen erstrecken sich immer über alle vorhandene Knoten.

3.2.3 Lernvorgang, Back–Propagation und die Minimierung der Fehlerfunktion

Nachdem die Statik Neuronaler Netze erläutert wurde, wird jetzt auf die Dynamik, d.h. auf die Implementierung des Lernvorganges eingegangen.

Die Eingabewerte $x_k^{(p)}$ und die gewünschten Ausgaben (Sollgrößen, **t**rue values) $t_i^{(p)}$ bilden zusammengenommen ein *Muster* (**p**attern) p. Die Menge aller Muster wird mit $\mathbb{T}$ bezeichnet. Zu jedem Muster p errechnet das Neuronale Netz mit den Formeln 3.4 und 3.5 einen Ausgabewert $o_i^{(p)}$. Dabei ist zu beachten, daß $t_i^{(p)}$ den *wahren* und $o_i^{(p)}$ den *errechneten* Ausgabewert eines Musters p angibt. Ziel des *Trainings* ist es, die Gewichte $\tilde{w}_{ij}$ und w_{jk} so zu bestimmen, daß die Korrelation zwischen $t_i^{(p)}$ und $o_i^{(p)}$ maximal wird. Diese Optimierung geschieht durch Minimieren der *Fehlerfunktion*

$$E_\chi = \frac{1}{2N_{\mathbb{M}}} \sum_{p\,\in\,\mathbb{M}} \sum_i \left(o_i^{(p)} - t_i^{(p)}\right)^2 \tag{3.6}$$

mit dem *Gradientenabstiegsverfahren* [Bro91], nachdem über alle Muster einer Teil**m**enge $\mathbb{M} \subset \mathbb{T}$ summiert wurde.

Das Gradientenabstiegsverfahren soll nun kurz vorgestellt werden. Dazu werden die Abkürzungen

$$\tilde{z}_i := \sum_j \tilde{w}_{ij}\; h_j - \tilde{\theta}_i \tag{3.7}$$

und

$$z_j := \sum_k w_{jk}\; x_k - \theta_j \tag{3.8}$$

definiert. Außerdem soll $\mathbb{M}$ ohne Einschränkung der Allgemeinheit genau ein Muster ($N_{\mathbb{M}} = 1$) enthalten.

Um die Fehlerfunktion 3.6 durch Optimierung der Gewichte zu minimieren, werden die *Anpassungen* der Gewichte $\Delta\tilde{w}_{ij}$ bzw. Δw_{jk} proportional zur Änderung der Fehlerfunktion bezüglich einer Änderung der Gewichte gewählt :

$$\Delta\tilde{w}_{ij} = -\eta\, \frac{\partial E_\chi}{\partial \tilde{w}_{ij}} = -\eta\, \frac{\partial E_\chi}{\partial o_i}\, \frac{\partial o_i}{\partial \tilde{w}_{ij}} \tag{3.9}$$

$$\Delta w_{jk} = -\eta\, \frac{\partial E_\chi}{\partial w_{jk}} = -\eta\, \frac{\partial E_\chi}{\partial h_j}\, \frac{\partial h_j}{\partial w_{jk}} \tag{3.10}$$

Der Proportionalitätsfaktor wird als *Lernparameter* η eingeführt, $0 \leq \eta \leq 1$, und sein Wert dem Neuronalen Netz vorgegeben. Für die Anpassung der Gewichte zwischen versteckter und Ausgabelage ergibt sich unter Ausnutzung der Definition von E_χ (Gleichung 3.6) und der Gleichung 3.5 die *Delta–Regel* [Wid60]

$$\Delta\tilde{w}_{ij} = -\eta\; \delta_i\; h_j \tag{3.11}$$

mit dem *Fehlersignal*

$$\delta_i = (o_i - t_i)\;\, g'\,(\tilde{z}_i)\,. \tag{3.12}$$

Hierbei bezeichnet $g'\,(\tilde{z}_i)$ die erste Ableitung der Aktivierungsfunktion $g\,(\tilde{z}_i)$. Um die Anpassung der Gewichte zwischen Eingabelage und versteckter Lage zu bestimmen, wird die Kettenregel auf $\frac{\partial E_\chi}{\partial h_j}$ in Gleichung 3.10 angewendet :

$$\Delta w_{jk} = -\eta\, \sum_i \frac{\partial E_\chi}{\partial o_i}\, \frac{\partial o_i}{\partial h_j}\, \frac{\partial h_j}{\partial w_{jk}} \tag{3.13}$$

Die einzelnen Ableitungen ergeben sich der Reihe nach aus den Gleichungen 3.6, 3.5 und 3.4, sowie mit der Delta–Regel :

$$\Delta w_{jk} = -\eta\, \sum_i (o_i - t_i)\;\, g'\,(\tilde{z}_i)\;\, \tilde{w}_{ij}\; x_k\; g'\,(z_j) = -\eta\, \sum_i \delta_i\; \tilde{w}_{ij}\; x_k\; g'\,(z_j) \tag{3.14}$$

Diese Gleichung heißt *verallgemeinerte Delta–Regel* [Rum86]. Hierbei wird das Fehlersignal δ_i, das in der versteckten Lage und Ausgabelage gebildet wurde, in die Eingabelage zurückgereicht, wodurch dieser Algorithmus den Namen *Back–Propagation–Learning* erhalten hat.

Anschaulich bewegt sich der Back–Propagation–Algorithmus im Raum der Gewichte $\tilde{w}_{ij}$ und w_{jk} auf der Fehlerfläche E_χ und versucht, ein Minimum zu finden. Da er aber seine Informationen nur aus der Umgebung um $\tilde{w}_{ij}$ und w_{jk} bezieht, gibt es eine Reihe von Schwierigkeiten, das globale Minimum zu erreichen. Für die meisten Anwendungen reicht ein gutes lokales Minimum, um die Fehlerfläche für die Eingabegrößen zu optimieren.

Einige Probleme sind :

- Flache Plateaus der Fehlerfläche :

  Da die Anpassungen vom Betrag des Gradienten abhängig sind, kann der Back–Propagation–Algorithmus auf flachen Plateaus diese nicht unbedingt wieder verlassen. Er läuft quasi im „Kreis“. Problematisch ist dies, da nicht unterschieden werden kann, ob der Grund für die Stagnation des Lernprozesses ein flaches Plateau oder ein Minimum ist.

- Oszillationen in steilen Schluchten der Fehlerfläche :

  Oszillationen in steilen Schluchten können auftreten, wenn der Gradient am Rande der Schlucht so groß ist, daß durch die Gewichtsänderung ein Sprung auf die gegenüberliegende Seite der Schlucht erfolgt.

- Verlassen guter Minima :

  Als Erweiterung zur Oszillation kann es vorkommen, daß in engen Tälern der Fehlerfläche der Gradient so groß ist, daß die Anpassung der Gewichte aus dem guten Mimimum heraus in ein weniger gutes Mimimum führt.

Aus diesem Gründen wird den Anpassungen $\Delta\tilde{w}_{ij}$ und Δw_{jk} einen Trägheitsterm α, $0 \leq \alpha \leq 1$, hinzugefügt :

$$\Delta\tilde{w}_{ij}^{n+1} = -\eta\ \delta_i\ h_j\ +\ \alpha\ \Delta\tilde{w}_{ij}^{n} \tag{3.15}$$

$$\Delta w_{jk}^{n+1} = -\eta \sum_i \delta_i\ \tilde{w}_{ij}\ x_k\ g'(z_j)\ +\ \alpha\ \Delta w_{jk}^{n} \tag{3.16}$$

Der Trägheitsterm bindet die vorherige Anpassung in die Neuberechnung mit ein. Schritte in die gleiche Richtung (gleiches Vorzeichen von $\Delta\tilde{w}_{ij}^{n+1}$ und $\Delta\tilde{w}_{ij}^{n}$ bzw. Δw_{jk}^{n+1} und Δw_{jk}^{n}) werden verstärkt und Plateaus lassen sich zügig überqueren. Oszillationen hingegen sind Rückschritte (entgegengesetztes Vorzeichen) und werden gedämpft. Daher heißt α auch *Dämpfung* und ergibt einen weiteren einstellbaren Parameter.

Neben dem Gradientenverfahren gibt es noch weitere Methoden [Pet89, Rog94], ein globales Minimum zu finden. Auf sie soll hier nicht weiter eingegangen werden.

Insgesamt ergibt sich folgendes Lernverfahren :

1. Initialisierung der Gewichte $\tilde{w}_{ij}$ und w_{jk} mit einer Zufallszahl $\in [-1, 1]$.

2. Beginn des ersten Lernzykluses. Dazu wird eine Teilmenge $\mathbb{M}_n$ aus der Menge aller **T**rainingsmuster $\mathbb{T}$ ausgewählt, wobei für den Durchschnitt aller Teilmengen $\bigcap\limits_n \mathbb{M}_n = \emptyset$ gelten muß und für deren Vereinigung $\bigcup\limits_n \mathbb{M}_n = \mathbb{T}$.
3. Berechnung der Ausgabewerte des Netzes $o_i^{(p)}$ aus den Eingabegrößen $x_k^{(p)}$ sukzessiv für alle Muster in $\mathbb{M}_n$.
4. Berechnung der Fehlerfunktion aus $o_i^{(p)}$ und $t_i^{(p)}$ aller Muster in $\mathbb{M}_n$.
5. Optimierung der Gewichte nach dem Back–Propagation–Algorithmus.
6. Hat sich ein Minimum gefunden, d.h. die Gewichte sind *konvergiert*, oder wurde eine bestimmte Anzahl von Lernzyklen erreicht, wird der Lernvorgang beendet. Ansonsten wird ein neuer Lernzyklus gestartet, erneut eine Teilmenge $\mathbb{M}$ ausgewählt und mit Punkt 3 fortgefahren.

3.2.4 Testphase und Ausführungsphase

Wurde eine Konvergenz aller Gewichte erreicht, können in der Test– und Ausführungsphase dem Neuronalen Netz unbekannte Muster vorgegeben werden. Das Netz errechnet für jedes Muster den Ausgabewert $o_i^{(p)}$ und klassifiziert das für ihn unbekannten Muster anhand der Merkmale gelernter Trainingsmuster. Diese Verallgemeinerung des Lernvorganges heißt *Generalisierung*.
Im Falle der Testphase wird eine Menge von Testmustern (**V**erification sample) $\mathbb{V}$ gebildet, die nicht im Training verwendet wurden ($\mathbb{T} \cap \mathbb{V} = \emptyset$), von denen aber der gewünschte Ausgabewert $t_i^{(p)}$ bekannt ist. Die Übereinstimmung von $o_i^{(p)}$ und $t_i^{(p)}$ für alle Testmuster gibt Auskunft über die Güte des Lernvorganges.
Ist die Testphase erfolgreich, kann das Neuronale Netz in der Ausführungsphase auf nicht klassifizierte Muster, bei denen $t_i^{(p)}$ unbestimmt ist, angewendet werden.

3.3 Das XOR–Problem

Die klassische Lernaufgabe für Neuronale Netze ist das sogenannte XOR–Problem. XOR (exklusives Oder, e**X**clusive **OR**) ist die logische Verknüpfung zweier Eingänge zu einem Ausgabesignal nach Tabelle 3.1.

Ziel ist es, das Neuronale Netz so anzulernen, daß es anhand der Eingabewerte selbstständig den Ausgabewert ermitteln kann. Wie in Abschnitt 3.2.1 schon erwähnt wurde, können die zwei Ausgabeklassen nicht mit Hilfe einer Hyperebene getrennt werden (Abbildung 3.3). Zwei Hyperebenen entsprechen zwei versteckten Knoten, so daß sich hier das einfachste aller mehrschichtigen Neuronalen–Netzwerk–Probleme ergibt.

Ein kurzes Training des Neuronalen Netzes mit zwei Eingabeknoten, zwei versteckten Knoten und einem Ausgabeknoten bestimmte die Gewichte und Schwellenwerte wie in Abbildung 3.5 angegeben.

p	$x_1^{(p)}$	$x_2^{(p)}$	$t_1^{(p)}$
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	0

Tabelle 3.1: *Logische Verknüpfung zweier Eingaben x_1 und x_2 nach der XOR–Funktion zu einem wahren Ausgabewert t_1.*

Werden die Werte für die Gewichte und Schwellen in das Argument der Gleichung 3.5 eingesetzt, so ergeben sich zwei, die Ausgabeklassen teilende Geraden :

$$w_{11} \cdot x_1 + w_{12} \cdot x_2 - \theta_1 = -2.990 \cdot x_1 - 3.007 \cdot x_2 + 4.476 = 0 \tag{3.17}$$

$$w_{21} \cdot x_1 + w_{22} \cdot x_2 - \theta_2 = -4.086 \cdot x_1 - 4.064 \cdot x_2 + 1.759 = 0 \tag{3.18}$$

Die Gewichte $\tilde{w}_{11}$ und $\tilde{w}_{12}$ zwischen verstecktem und Ausgabe–Knoten (Gleichung 3.4) werden nur zur Normierung der Ausgabewerte benötigt. Zur Veranschaulichung wurden in Abbildung 3.6 die Eingabevariablen x_1 und x_2 über das Intervall $[0, 1]$ variiert. Deutlich sind die beiden Hyperebenen (Gleichung 3.17 und 3.18) an den Flanken der Ausgabegröße o_1 zu erkennen (vgl. Abbildung 3.3).

3.4 Das Programmpaket `JETNET`

Die Implementierung des Neuronalen Netzes erfolgt mit dem Programm–Paket `JETNET 3.0` [Lon93] der Universität Lund, Schweden. Es handelt sich hierbei um ein mehrlagiges

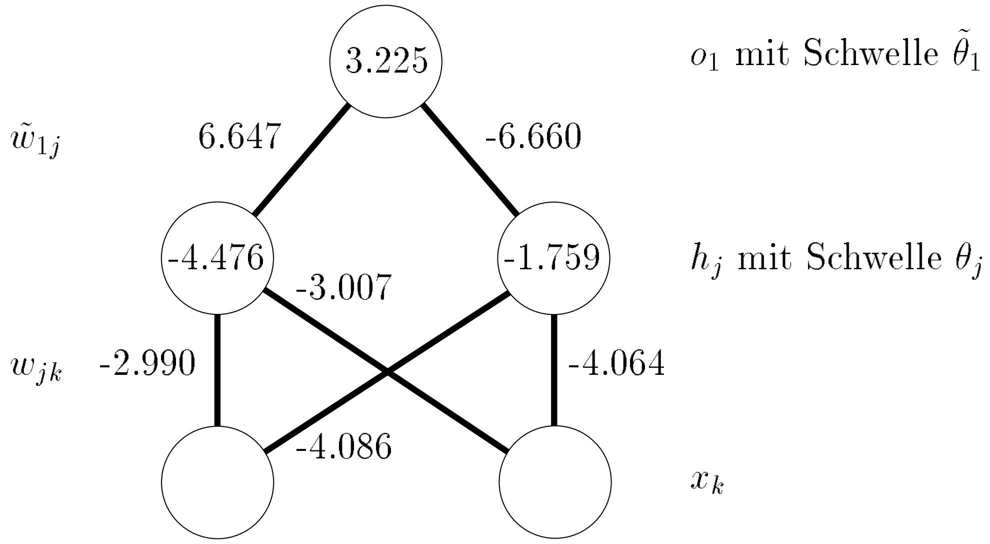

Abbildung 3.5: *Einfaches Neuronales Netz für das XOR–Problem. Die angegebenen Werte für $\tilde{w}_{1j}$, w_{jk}, $\tilde{\theta}_1$ und θ_j stammen aus einem hier durchgeführten Lernvorgang.*

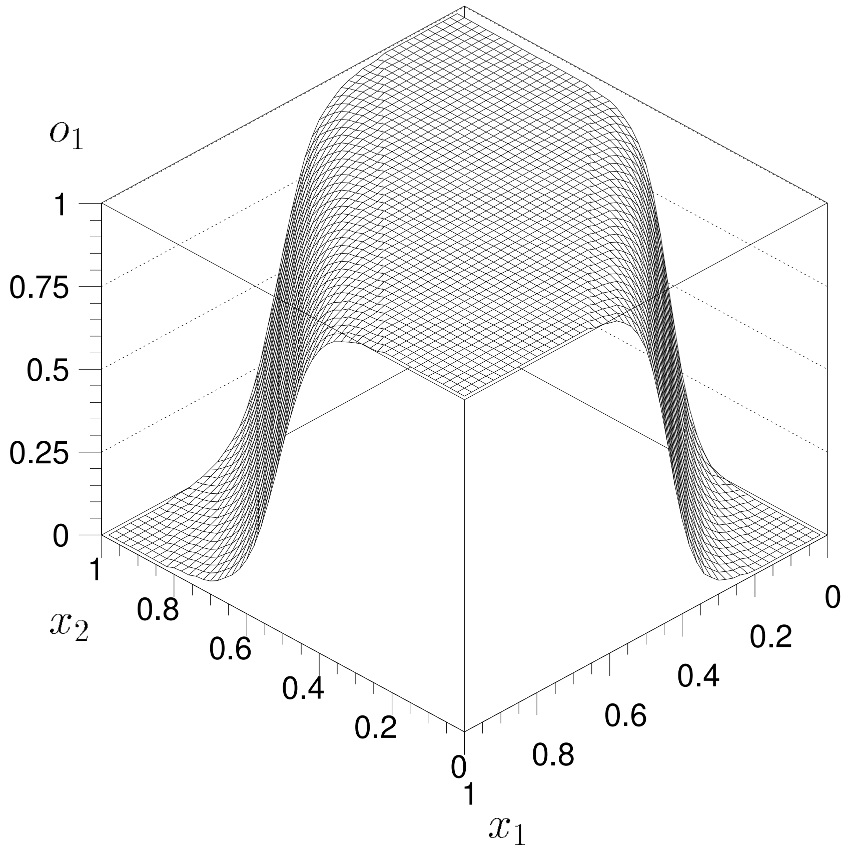

Abbildung 3.6: *Antwort o_1 des Neuronalen Netzes auf kontinuierliche XOR–Eingaben x_1 und x_2. Die Flanken der Ausgabe bilden die klassentrennenden Hyperebenen.*

Neuronales Netz mit Feedforward–Architektur und Back–Propagation–Lernalgorithmus mit Routinen für die Trainings–, Test– und Generalisierungsphase.
Initialisieren läßt sich ein Neuronales Netz mit maximal 1000 Knoten pro Lage und maximal 10 versteckten Lagen. Die wichtigsten einstellbaren Parameter sind :

- die Anzahl der Lagen im Neuronalen Netz
- die Anzahl der Knoten N_x, N_h und N_o in jeder Lage
- die Anzahl $N_{\mathbb{M}}$ der Muster in der Menge $\mathbb{M}$
- die Aktivierungsfunktion, z.B.
  - $g(z) = \tanh z$
  - Sigmoidfunktion (Gleichung 3.2) :

$$g(z) = \left(1 + e^{-2z}\right)^{-1}$$

- die Fehler–Funktion, z.B.
  - Log–Likelihood–Fehler :

$$E_{LLF} = -\frac{1}{2N_{\mathbb{M}}} \sum_{p \,\in\, \mathbb{M}} \sum_i \ln\left(1 - \left(o_i^{(p)} - t_i^{(p)}\right)^2\right)$$

  - Entropie–Fehler :

$$E_{EF} = -\frac{1}{2N_{\mathbb{M}}} \sum_{p \,\in\, \mathbb{M}} \sum_i \left(1 - t_i^{(p)}\right) \ln\left(1 - o_i^{(p)}\right) - t_i^{(p)} \ln\left(o_i^{(p)}\right)$$

  - χ^2–Fehler (Gleichung 3.6) :

$$E_{\chi} = \frac{1}{2N_{\mathbb{M}}} \sum_{p \,\in\, \mathbb{M}} \sum_i \left(o_i^{(p)} - t_i^{(p)}\right)^2$$

- der Lern–Algorithmus
  - Gradienten–Verfahren
  - Manhattan–Verfahren [Pet89]
  - Langevin–Verfahren [Rog94]
- die Lernrate η
- die Dämpfung α

Daneben stehen noch eine ganze Reihe von weiteren Parametern, speziell zur Verbesserung der *Güte* des Netzes, zur Verfügung [Lon93]. Die Güte macht eine Aussage über die Leistung des Neuronalen Netzes, wie schnell und exakt es ein Minimum der Fehler–Funktion findet.

Kapitel 4

Die Produktion schwerer Quarks bei HERA

> Kein Ding entsteht planlos,
> sondern alles aus Sinn und unter Notwendigkeit.
> *Leukippos von Milet, 5. Jh. v. Chr.*

Die konzeptionelle Auslegung des H1–Detektors mit dem instrumentierten Eisen und dem Vorwärts–Myon–System ermöglicht die Untersuchung von Prozessen mit einem Myon im Endzustand. In diesem Kapitel werden die physikalischen Prozesse diskutiert, die zu der Produktion schwerer Quarks und deren semimyonischem Zerfall führen. Anhand dieses Zerfalls und unter Nachweis des Myons im instrumentierten Eisen sollen später das schwere Beauty– und Charm–Quark voneinander und von anderen Ereignissen mit Myonen — dem Untergrund — unterschieden werden.

Zunächst werden die wichtigsten kinematischen Größen der Elektron–Proton–Streuung definiert. Anschließend wird die tiefinelastische Streuung und ihr Grenzfall für kleine Viererimpulsüberträge, die Photoproduktion, beschrieben (Abschnitt 1). Der bei der Erzeugung schwerer Quarks dominante Prozeß, die Boson–Gluon–Fusion, wird in Abschnitt 2 betrachtet und der semimyonische Zerfall der schweren Quarks diskutiert. Neben diesen Zerfallsmyonen werden Untergrundquellen zur Erzeugung von Myonen aufgezeigt und Prozesse erläutert, die Myonen vortäuschen (Abschnitt 3). In Abschnitt 4 wird abschließend die Berechnung der theoretischen Modelle und die daraus folgende Ereignis–Generierung mit Hilfe der Monte–Carlo–Methode beschrieben.

4.1 Kinematik der *ep*–Streuung

Die Bezeichnungen der Vierervektoren der an der Lepton–Proton–Streuung beteiligten Teilchen sind in Abbildung 4.1 abzulesen :

$P = (E_P, \vec{p}_P)$: einlaufendes Proton

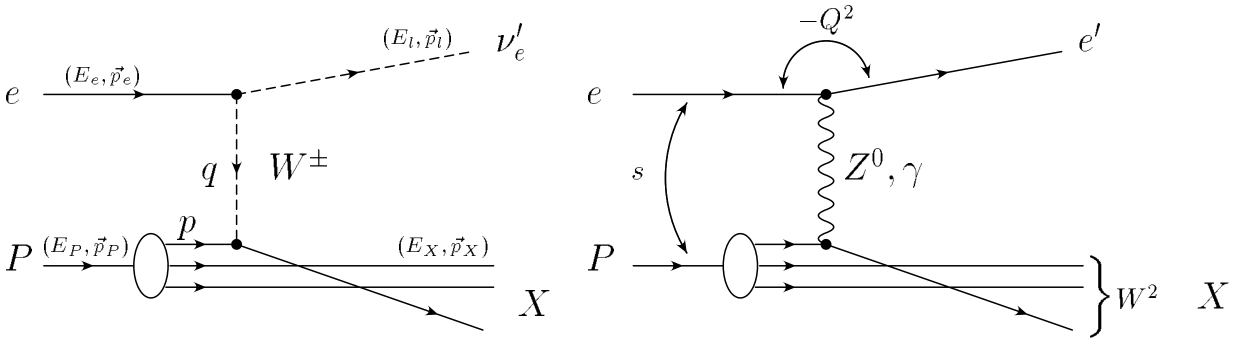

Abbildung 4.1: *Feynman–Graphen für Prozesse der tiefinelastischen Streuung, links unter Austausch eines geladenen Bosons, rechts mit einem neutralen Boson.*

p : Parton

$e = (E_e, \vec{p}_e)$: einlaufendes Elektron

q : Austauschboson

$l = (E_l, \vec{p}_l)$: gestreutes Lepton

$X = (E_X, \vec{p}_X)$: hadronischer Endzustand

Im Schwerpunktsystem besitzt der *ep*–Prozeß drei Freiheitsgrade, wovon einer durch die Impulse von einlaufendem Elektron und Proton bei HERA festgelegt ist. Damit ist der Prozeß in seiner Kinematik vollständig bestimmt, wenn beispielsweise die Energie E_l und der Polarwinkel Θ_l des gestreuten Elektrons nachweisbar sind.

Zur Beschreibung der Streuprozesse werden bevorzugt lorentzinvariante Variablen benutzt, die sich folgendermaßen aus den Viererimpulsen der beteiligten Teilchen berechnen lassen :

- das *Quadrat des Viererimpulsübertrages* vom Elektron auf das abgestrahlte Boson, auch mit *Virtualität des Bosons* bezeichnet :

  $$Q^2 \equiv -q^2 = -\left(e - l\right)^2 \simeq 2E_e E_l \left(1 + \cos\Theta_l\right) \tag{4.1}$$

  Q^2 entspricht dem negativen Quadrat des Viererimpulses des Austauschbosons und hat für virtuelle Bosonen einen positiven Wert. Für reelle masselose Teilchen hingegen ist das Quadrat des Viererimpulses und damit auch das Quadrat des Viererimpulsübertrages immer Null.

- das *Quadrat der Schwerpunktsenergie im ep–System*

  $$s = \left(P + e\right)^2 \simeq 4E_e E_P. \tag{4.2}$$

- das *Quadrat der Schwerpunktsenergie des* γp*–Systems*, entsprechend der *invarianten Masse des hadronischen Endzustandes*

$$W_{\gamma p}^2 = (q+P)^2 = m_P^2 + 2qP - Q^2. \tag{4.3}$$

Die angegebenen Nährungen gelten unter Vernachlässigung der Elektron– und Protonmassen.

Aus diesen Größen lassen sich die häufig benutzte *Bjorkensche Skalenvariable* x und die *Inelastizität* y ableiten :

$$x = \frac{Q^2}{2\,(q \cdot P)} \quad \text{und} \quad y = \frac{q \cdot P}{e \cdot P} \tag{4.4}$$

Im *Quark–Parton–Modell* wird das Proton als System freier Konstituenten, den *Partonen*, behandelt. Das Quark, das mit dem Impuls xP in die Streuung eingeht, hat demnach keine zur Strahlrichtung transversale Komponente. Unter Vernachlässigung der Masse der leichten Quarks steht die Bjorkensche Skalenvariable x näherungsweise für den Impulsanteil des wechselwirkenden Partons am Gesamtimpuls des Protons. Die Größe y gibt die Inelastizität des Elektron–Proton–Systems an und kann als relativer Energieverlust des gestreuten Elektrons im Ruhesystem des Protons interpretiert werden.
Das Quark–Parton–Modell ist eine Vereinfachung. Nach dem *Standardmodell der Elementarteilchenphysik* setzt sich das Proton aus drei Arten von Partonen zusammen : den *Valenzquarks*, den *Seequarks* und den Austauschteilchen der starken Wechselwirkung, den *Gluonen.*

In der *ep*–Streuung wird zwischen Ereignissen der tiefinelastischen Streuung und ihrem Grenzwert für Viererimpulsüberträge nahe Null, den Ereignissen der *Photoproduktion*, unterschieden.
Bei der *tiefinelastischen Streuung* besitzt das vom Elektron abgestrahlte Boson einen so hohen Viererimpuls, daß die Struktur des Protons aufgelöst werden kann. Bei H1 werden üblicherweise Ereignisse dann als tiefinelastisch definiert, wenn das gestreute Elektron im zentralen Detektor nachgewiesen werden kann, was bei einem Viererimpulsübertrag Q^2 ab etwa $1\,\mathrm{GeV}^2$ der Fall ist — andernfalls wird das Ereignis als Photoproduktionsereignis betrachtet[1].

4.1.1 Tiefinelastische Streuung

Prozesse mit hohem Viererimpulsübertrag Q^2, Prozesse der tiefinelastischen Streuung, lassen sich im Rahmen der **Q**uanten**c**hromo**d**ynamik (QCD) berechnen, denn dann ist die Kopplungskonstante α_s für störungstheoretische Näherungen hinreichend klein. In der niedrigsten Ordnung des elektroschwachen Standardmodells läuft die tiefinelastische Streuung über den *geladenen* bzw. den *neutralen Strom* ab :

$$e + p \rightarrow \nu_e + X \quad \text{bzw.} \quad e + p \rightarrow e + X \tag{4.5}$$

---

[1] Vor 1995 wurde diese Trennung bei $Q^2 < 4\,\mathrm{GeV}^2$ gemacht, da der Vorgänger des SPACALs einen kleineren Polarwinkelbereich abdeckte.

Das einlaufende Elektron wechselwirkt nach dem Quark–Parton–Modell mit einem der Partonen im Proton. Die Wechselwirkung wird durch ein Eichboson vermittelt, welches im Falle des neutralen Stromes das Photon oder das Z^0 ist und beim geladenen Strom das $W^{\pm}$–Boson. Nach der anschließenden Fragmentation entsteht zum einen ein hadronischer Endzustand X, zum anderen produziert der Streuprozeß ein gestreutes Lepton (Abbildung 4.1).

4.1.2 Photoproduktion

Mit sinkendem Quadrat des Viererimpulsübertrages kann der Beitrag des geladenen Stromes und die Beimischung des Z^0–Bosons am neutralen Strom aufgrund der hohen Massen von $W^{\pm}$ und Z^0 ($M_{Z^0} \simeq 91$ GeV^2 und $M_{W^{\pm}} \simeq 80$ GeV^2 [PDG96]) und des damit kleiner werdenden Propagators $1/(q^2 + M^2_{Z^0,W^{\pm}})$ vernachlässigt werden. Das verbleibende Austauschteilchen, das Photon, wird zunehmend „reeller“. Der Grenzfall $Q^2 \rightarrow 0$ wird als Photoproduktion bezeichnet.
In der Photoproduktion wird die ep–Streuung in zwei unabhängige Prozesse unterteilt : Emission eines Photons vom einlaufenden Elektron und Wechselwirkung des Photons mit dem Proton. Die Photonemission wird durch die Weizsäcker–Williams–Näherung [Wei34, Wil34] beschrieben. Bei dieser Methode wird das elektromagnetische Feld eines sich schnell an einem Nukleon vorbei bewegenden Elektrons als Strom virtueller Photonen betrachtet. Damit läßt sich der Wirkungsquerschnitt der ep–Streuung auf den γp–Wirkungsquerschnitt reduzieren, gefaltet mit einem Faktor für den Photonfluß [Che76].

4.2 Prozesse zur Erzeugung schwerer Quarks

4.2.1 Boson–Gluon–Fusion

Der dominierende Prozeß zur Erzeugung schwerer Quarks bei HERA ist die *Boson–Gluon–Fusion*. In niedrigster Ordnung $\mathcal{O}(\alpha, \alpha_s)$ koppelt das vom Elektron entsandte Boson unter Bildung eines Quark–Antiquark–Paares an ein Gluon des Protons (Abbildung 4.2). In diesem störungstheoretisch berechenbaren *harten Prozeß*, an dem das Boson, das Gluon und das Quarkpaar teilnehmen, können im neutralen Strom immer nur Quark und Antiquark gleicher Quarksorte und im geladenen Strom nur verschiedener Quarksorte produziert werden. Bei HERA überwiegen Ereignisse mit quasireellen Photonen ($Q^2 \approx 0$), so daß der geladene Strom wie auch der Z^0–Anteil am neutralen Strom vernachlässigt werden können. Es genügt deshalb im Folgenden nur die *Photon–Gluon–Fusion* zu betrachten.

Neben den bisher definierten Vierervektoren treten in der Boson–Gluon–Fusion drei weitere Teilchen auf, deren Vierervektoren folgendermaßen bezeichnet werden :

g : Gluon

$p_Q = (E_Q, \vec{p}_Q)$: erzeugtes Quark

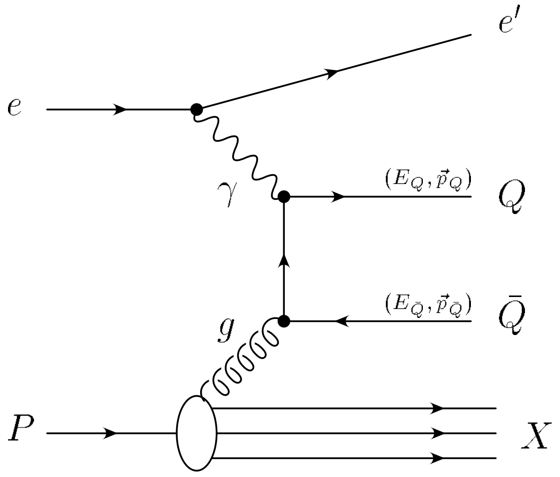

Abbildung 4.2: *Feynman–Graph für den Prozeß der Photon–Gluon–Fusion in erster Ordnung in* α_s *mit der Erzeugung eines Quarkpaares* $Q\bar{Q}$.

$p_{\bar{Q}} = (E_{\bar{Q}}, \vec{p}_{\bar{Q}})$: erzeugtes Antiquark

Durch die neu hinzukommenden Teilchen ergeben sich drei weitere Freiheitsgrade, so daß zur vollständigen Beschreibung der Kinematik mindestens drei weitere unabhängige Variablen notwendig sind:

- der *Impulsanteil* x_g *des Gluons am Protonimpuls* definiert durch

$$g = x_g P. \tag{4.6}$$

- das *Quadrat der invarianten Masse des erzeugten Quarkpaares*, auch *invariante Masse des Photon–Gluon–Systems*

$$\hat{s} = \left(p_Q + p_{\bar{Q}}\right)^2 . \tag{4.7}$$

  Es gilt die Beziehung

$$\hat{s} = Q^2 \, \frac{x_g - x}{x}. \tag{4.8}$$

- das Verhältnis vom tatsächlichen zum maximal möglichen Impuls eines schweren Quarks

$$z = \frac{P \cdot p_Q}{P \cdot q}. \tag{4.9}$$

- der *azimutale Winkel zwischen der Hadronen– und Leptonenebene im Boson–Gluon–Schwerpunktsystem*

$$\cos\Phi = \frac{(\vec{p}_p \times \vec{p}_e) \cdot (\vec{p}_p \times \vec{p}_Q)}{|\vec{p}_p \times \vec{p}_e| \cdot |\vec{p}_p \times \vec{p}_Q|}. \tag{4.10}$$

Es ist zu beachten, daß die Skalenvariable x nicht mehr die anschauliche Bedeutung wie im Quark–Parton–Modell besitzt.

Schwere Quarkpaare entstehen, wenn die invariante Masse des Photon–Gluon–Systems $\hat{s}$ genügend groß ist. Der theoretisch errechnete Wirkungsquerschnitt hängt von den zugrundegelegten Quarkmassen ab. Unter Annahme einer Charm–Quarkmasse von $m_c = 1.5\text{GeV}$ und einer Beauty-Quarkmasse von $m_b = (4.75 \pm 0.25)$ GeV ergibt sich bei einer Schwerpunktsenergie $\sqrt{s} = 314$ GeV bis zur Ordnung (α^2, α_s) [Eli89] :

$$\sigma(ep \to c\bar{c}\,X) = 680^{+260}_{-180}\,\text{nb}$$

$$\sigma(ep \to b\bar{b}\,X) = 6.0 \pm 1.1\,\text{nb}$$

Am H1–Detektor wurde der totale Wirkungsquerschnitt aus der Analyse von $D^{*\pm}$–Photoproduktionsereignissen bestimmt :

$$\sigma(ep \to Q\bar{Q}\ X) = 941 \pm 160^{+142}_{-120}\,\text{nb [H1C96]}$$

$$\sigma(ep \to c\bar{c}\,X) = 807 \pm 133^{+163}_{-146}{}^{+711}_{-202}\,\text{nb [Kat97]}$$

4.2.2 Prozesse mit aufgelöstem Photon

Im Gegensatz zu der bisher beschriebenen Photon–Gluon–Fusion in niedrigster Ordnung in α_s (Abbildung 4.2) kann in Prozessen höherer Ordnung die hadronische Komponente des Photons aufgelöst werden. Dies wird als Prozeß mit *aufgelöstem Photon* bezeichnet (Abbildung 4.3), der Prozeß niedrigster Ordnung heißt Prozeß mit *direktem Photon*.

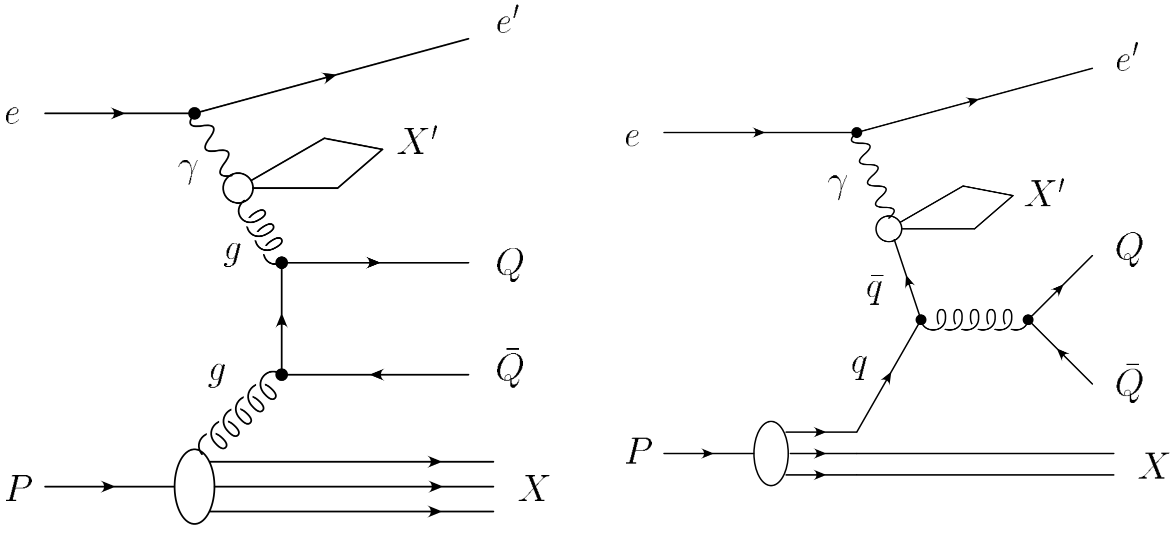

Abbildung 4.3: *Feynman–Graphen für die Erzeugung schwerer Quarks mit einem aufgelösten Photon. In der linken Abbildung ist der Prozeß $gg \to Q\bar{Q}$ (Gluon–Gluon–Fusion) gezeigt, in der rechten Abbildung der Prozeß $q\bar{q} \to Q\bar{Q}$.*

Abbildung 4.4 zeigt die Abhängigkeit des Wirkungsquerschnittes von der Quarkmasse. Neben dem Prozeß mit direktem Photon (γg) sind auch Beiträge höherer Ordnung in α_s dargestellt und es wird deutlich, daß der „direkte“ Prozeß bei den Massen der schweren Quarks ($m_c \simeq 1.5\,\mathrm{GeV}$, $m_b \simeq 4.5\,\mathrm{GeV}$ [PDG96]) dominiert. Der Beitrag der aufgelösten Komponente am Gesamtwirkungsquerschnitt für die Produktion schwerer Quarks wird in der Größenordnung $\mathcal{O}(20\%)$ erwartet [Fri95]. Für kleinere Massen kann nach Abbildung 4.4 der aufgelöste Anteil überwiegen. In der Photoproduktion leichter Quarks dominiert die aufgelöste Komponente den Gesamtwirkungsquerschnitt mit 80% [Kat97].

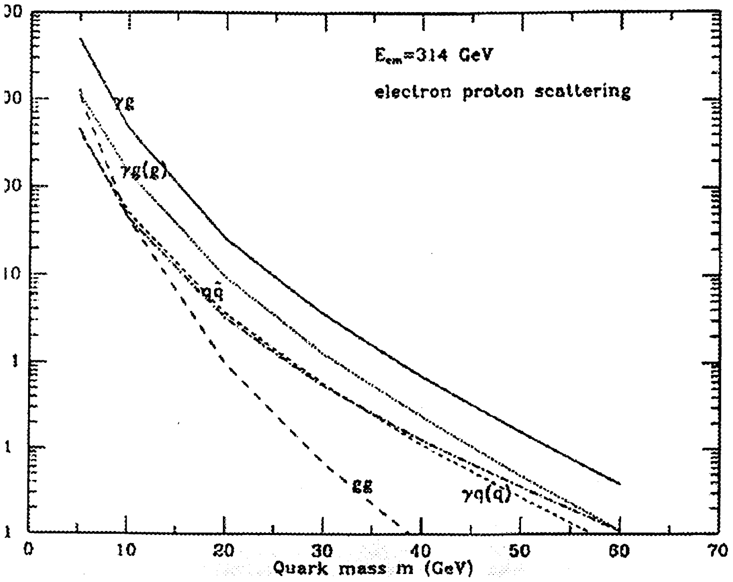

Abbildung 4.4: *Wirkungsquerschnitte für die Produktion schwerer Quarks als Funktion der Quarkmasse [Kun89]. Die durchgezogene Linie beschreibt den „direkten“ Prozeß (γg).*

4.2.3 Partonschauer, Fragmentation und Jetbildung

Die bei der Boson–Gluon–Fusion entstandenen Quarks besitzten eine hohe Virtualität. Die Partonen werden zunächst als quasifreie Teilchen behandelt, die die Möglichkeit haben, weitere Partonen durch Gluonabstrahlung zu bilden. Der sich so ausbildende Schauer ist ein Prozeß höherer Ordnung in α_s. Er läßt sich durch das *Partonschauermodell* beschreiben [Ing88a]. In diesem Modell senden die quasifreien Partonen virtuelle Partonen aus, die durch weiteres Aussenden virtueller Partonen auf ihre Massenschale gehoben werden. Das ursprüngliche Parton wird dadurch zunehmend zeitartiger. Die Wahrscheinlichkeit der Ausstrahlung eines Partons von einem anderen Parton und die damit verbunden Abnahme der Virtualität wird durch die *Altarelli–Parisi–Gleichungen* beschrieben, wobei der Unterschied zwischen Gluon und Quark Berücksichtigung findet [Ing88b].

Das Partonschauermodell beschreibt die Schauerentwicklung bis hinunter zu Partonen einer Virtualität von etwa 1 GeV^2. Bei dieser Energie können die Partonen nicht mehr als quasifreie Teilchen betrachtet werden, so daß die weitere Hadronisierung durch ein *Frag-*

mentationsmodell beschrieben wird. Das dafür verwendete *Lund–String–Modell* [Sjo88] verbindet die im Partonschauer entstandenen Partonen untereinander und zusammen mit den Protonquarks durch Farbfelder (*Strings*) zu farbneutralen Paaren von Quarks und Antiquarks oder Quarks und Diquarks. Das Potential eines Strings wird durch

$$V(r) = \kappa r \quad \text{wobei} \quad \kappa \approx 1\ \mathrm{GeV/fm}$$

beschrieben.
Durch die unterschiedlichen Impulse fliegen die Quarks auseinander, wobei die kinetische Energie in potentielle Energie des Strings umgewandelt wird. Erreicht die deponierte Energie einen bestimmten Wert, bricht der String auf und die potentielle Energie erzeugt ein neues (Di–)Quarkpaar. Aus dem ursprünglichen String werden somit zwei mit jeweils einem ursprünglichen und einem neuen (Di–)Quark an den Enden. Nun setzt sich der Vorgang wiederholt fort, bis die potentielle Energie eines Strings nicht mehr zur Bildung eines (Di–)Quarkpaares ausreicht. Die übrigbleibenden, mit einem String verbundenen Quarks sind Mesonen und Baryonen, die den hadronischen Endzustand X bilden. Durch die ursprünglichen Inpulse sind sie in ihrer Flugrichtung stark korreliert und bilden Bündel von Teilchen, *Jets*.

4.2.4 Semileptonische Zerfälle schwerer Quarks

Im Rahmen des Standardmodells wird der Zerfall schwerer Quarks mit Hilfe der schwachen Wechselwirkung beschrieben. Die Zerfallsbreite Γ für diesen Prozeß $q^a \to q^b\ X$ ist proportional zu den Quadraten der Elemente der ***C**abbibo–**K**obayashi–**M**askawa–Matrix* (CKM) [Cab63, Kob73] :

$$\Gamma\left(q^a \to q^b\right) \sim |V_{q^a q^b}|^2 \tag{4.11}$$

Wegen des kleineren CKM–Matrixelementes ist der Zerfall $b \to u\ W^-$ gegenüber dem Zerfall $b \to c\ W^-$ stark unterdrückt, $|V_{bu}|/|V_{bc}| = 0.08 \pm 0.02$ [ScP95]. Das erzeugte $W^\pm$–Boson ist virtuell, da seine Masse größer als die des Quarks ist. Das Boson zerfällt entweder semileptonisch in ein $l\nu_l$–Paar oder hadronisch in ein $q\bar{q}$–Paar. In Abbildung 4.5 sind zwei Zerfallskanäle dargestellt.

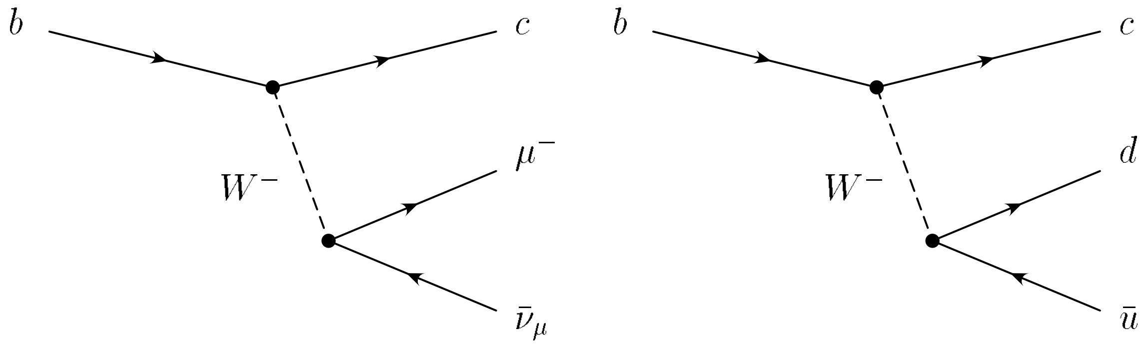

Abbildung 4.5: *Semileptonischer (links) und hadronischer (rechts) Zerfall eines b–Quarks.*

Das CKM–Matrixelement für den Übergang $c \to d\, W^+$ beträgt ungefähr 20% des Überganges $c \to s\ W^+$. Zusammenfassend zerfallen aufgrund der CKM–Matrixelemente b–Quarks fast ausschließlich in c–Quarks und c–Quarks bevorzugt in s–Quarks.

In der vorliegenden Analyse wird der semileptonische Dreikörperzerfall schwerer Hadronen, bei denen ein Myon entsteht (auch *semimyonischer Zerfall* genannt), zur Identifikation schwerer Quarks ausgenutzt. Im *Zuschauer–Modell* wird angenommen, daß das leichte Quark völlig unbeteiligt am Zerfall des schweren Quarks ist.
Ohne Berücksichtigung der unterschiedlichen Phasenräume aufgrund der unterschiedlichen Quarkmassen und der unterschiedlichen Matrixelemente können die semimyonischen Zerfallswahrscheinlichkeiten durch Aufzählen aller möglichen Zerfallskanäle bestimmt werden. Für das c–Quark lauten sie folgendermaßen :

$$c \to s + \mu^+ + \nu_\mu$$
$$c \to s + e^+ + \nu_e$$
$$c \to s + \bar{d} + u$$

Der letzte Übergang geht wegen der drei Möglichkeiten für das $W^\pm$–Bosons, in einen farbneutralen $q\bar{q}$–Zustand zu zerfallen, mit der dreifachen Gewichtung ein. Somit ergibt sich :

$$\frac{\Gamma(c \to \mu\ X)}{\Gamma(c \to X')} \simeq 20\%. \tag{4.12}$$

Wird dieselbe Abschätzung auf den Zerfall des b–Quarks angewendet, ergibt sich eine semimyonische Zerfallswahrscheinlichkeit von 11%, da aufgrund der höheren Masse des b–Quarks zusätzlich zu den Zerfällen

$$b \to c + \mu^- + \nu_\mu$$
$$b \to c + e^- + \nu_e$$
$$b \to c + d + \bar{u}$$

der Zerfall in ein τ–Lepton und der Zerfallskanal $W^- \to \bar{c}\ s$ kinematisch erlaubt sind.

Um die Zerfallswahrscheinlichkeiten der schweren Quarks vollständig zu berechnen, muß jedoch noch der Einfluß unterschiedlicher Phasenraumfaktoren und QCD–Korrekturen berücksichtigt werden.
Die theoretischen semimyonischen Verzweigungsverhältnisse betragen für das b– und c–Quark jeweils 11.0 % [Eic88]. Die Mittelung über experimentelle Verzweigungsverhältnisse ergibt für $b \to \mu\ X$ (10.7 ± 0.7) % und für $c \to \mu\ X$ $(8.1^{+1}_{-0.9})$ % [PDG96].

Neben den Myonen, die direkt aus dem semimyonischen Zerfall des b–Quarks entstehen, werden indirekt auch Myonen durch den *Kaskaden-Zerfall* $b \to c\ X' \to \mu\ X$ produziert (Abbildung 4.6). Da b–Quarks fast ausschließlich in c–Quarks zerfallen, entsteht so ein weiterer, nicht zu vernachlässigender Kanal zur Produktion von Myonen. Das mittlere experimentelle Verzweigungsverhältnis für den Kaskaden–Zerfall beträgt (7.9 ± 0.8) % [PDG96].

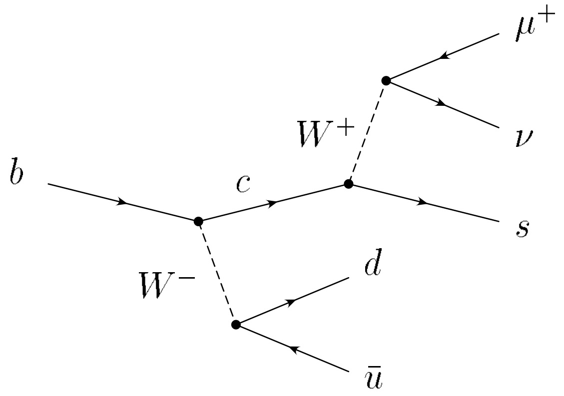

Abbildung 4.6: *Der Kaskaden–Zerfall.*

4.2.5 Kinematik der Zerfälle schwerer Quarks

Die Kinematik des Ereignisses wird durch den Zerfall des ursprünglichen Quarks charakterisiert und hängt von dessen Masse ab. Im Quark–Schwerpunktsystem, unter Erhaltung der Gesamtenergie, bekommen Zerfallsteilchen aus b–Quarks wegen dessen größerer Masse im Mittel einen höheren Impuls als Teilchen aus dem Zerfall eines c–Quarks. Damit steht im Laborsystem den Zerfallsteilchen eines schweren Quarks mehr Energie für Impulse transversal zur ursprünglichen Quarkrichtung zur Verfügung als den Teilchen aus dem Zerfall eines leichteren Quarks. Dies wird im Folgenden auch als *Massenargument* bezeichnet.

Abbildung 4.7 zeigt den transversalen Impuls des Myons bezüglich der Flugrichtung des Quarks, aus dessen Zerfall es entstanden ist, betrachtet in der $R\Phi$–Ebene :

$$p_{\perp R\Phi} := \sqrt{(\vec{p}_{t,\mu})^2 - (\vec{n}_{t,Q} \cdot \vec{p}_{t,\mu})^2} \tag{4.13}$$

Dabei gibt $\vec{n}_{t,Q} = \frac{\vec{p}_{t,Q}}{|\vec{p}_{t,Q}|}$ die Flugrichtung des Quarks an.
Der Unterschied in den Verteilungen aufgrund des Massenargumentes zeigt sich deutlich.

4.3 Myonen aus anderen Prozessen

Neben den semimyonischen Zerfällen schwerer Quarks gibt es noch eine Reihe anderer Prozesse, bei denen Myonen entstehen können oder die im instrumentierten Eisen myonähnliche Signaturen hinterlassen und damit Myonen *vortäuschen*. Alle diese *Untergrund*–Prozesse *dominieren* die Ereignisse mit Myonen.
Die Analyse von semimyonischen Zerfällen schwerer Quarks für den Zentralbereich des H1–Detektors zeigt, daß Myonen aus Zerfällen leichter Hadronen und fehlidentifizierte Pionen den größten Anteil am Untergrund bilden [Kru94].

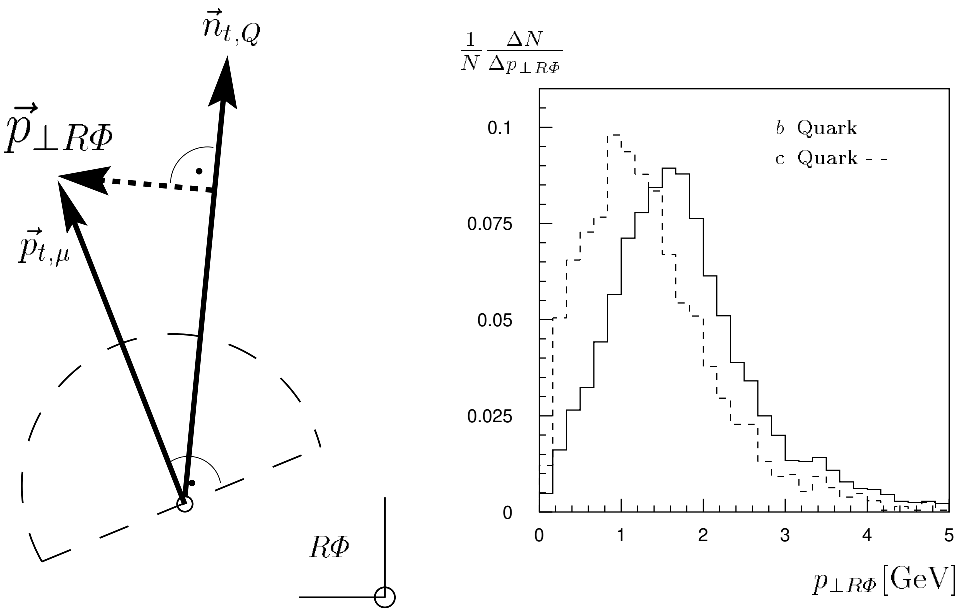

Abbildung 4.7: *Definition des transversalen Impulses $p_{\perp R\Phi}$ des Myons bezüglich der Quarkrichtung (links) und seine Verteilung für b– und c–Quarks (rechts).*

Zerfallsmyonen Die Hauptbeiträge aus Zerfällen leichter Hadronen stammen von geladenen Pionen und Kaonen, die *vor* dem Erreichen des Kalorimeters in Myonen zerfallen. Die Verzweigungsverhältnisse BR (**B**ranching **R**atio) sind für den semimyonischen Kanal $\pi \to \mu\nu$ nahezu 100% und für den Kanal $K \to \mu\nu$ 67% [PDG96].
Die Wahrscheinlichkeit, daß dieser Zerfall vor Erreichen des Kalorimeters stattfindet, ergibt sich aus

$$P(\lambda) = BR \cdot \left(1 - e^{-\frac{\lambda}{\gamma\beta c\tau}}\right) = BR \cdot \left(1 - e^{-\frac{\lambda m}{p c\tau}}\right) \quad \text{wobei} \quad BR = \begin{cases} 1 & \text{für Pionen} \\ 0.67 & \text{für Kaonen} \end{cases} \tag{4.14}$$

mit den gebräuchlichen Abkürzungen $\gamma = \frac{E}{m}$ und $\beta = \frac{p}{E}$. Es bezeichnet λ die Weglänge vom Wechselwirkungspunkt bis zum Kalorimeter. Die Masse, der Impuls und die Lebensdauer des jeweiligen Hadrons werden durch m, p und τ angegeben.

Myonen aus der kosmischen Höhenstrahlung Myonen aus der kosmischen Höhenstrahlung durchqueren den H1–Detektor mit einer Rate von ungefähr 2 kHz. Die dabei hinterlassenen Signaturen im Spurkammersystem, Kalorimeter und instrumentierten Eisen können Reaktionen mit Myon–Paaren ähneln, so daß die kosmischen Myonen auch von dem H1–Triggersystem und der L5–Klassifikation erfaßt werden können. Daneben können kosmische Myonen, die in zufälliger Koinzidenz mit einem *ep*–Ereignis im H1–Detektor eintreffen, dieses Ereignis überlagern.

Strahl–Gas– und Strahl–Wand–Ereignisse Als weiteres können Wechselwirkungen der einlaufenden Protonen mit dem Restgas oder mit der Strahlrohrwand Myonen produzieren. Ereignisse dieser Art ähneln in der Topologie Photoproduktionsereignissen mit kleinem Transversalimpuls und werden somit ebenfalls nicht unbedingt von dem Triggersystem und der L5–Klassifikation erkannt.

Fehlidentifikation von Hadronen Neben den Myonen aus der kosmischen Strahlung, den Strahl–Gas– und Strahl–Wand–Ereignissen und aus den Zerfällen von leichten Hadronen können Pionen oder andere Hadronen auch selbst Signaturen im instrumentierten Eisen hinterlassen und somit Myonen vortäuschen.

Während Pionen oder Kaonen das Kalorimeter durchqueren, erzeugen sie dort in der Regel einen hadronischen Schauer, indem sie mit dem Material des Kalorimeters stark wechselwirken. Die erzeugten sekundären Teilchen können teilweise bis in das instrumentierte Eisen gelangen und dort ihre restliche Energie deponieren. Das ursprüngliche Pion wird deshalb auch als *Punch–Through–Pion* bezeichnet.

Die Wahrscheinlichkeit P, daß ein Pion oder ein anderes Hadron Material der Dicke d durchquert, *ohne* hadronisch zu wechselwirken, beträgt

$$P(d) = e^{-\frac{d}{\lambda}} \tag{4.15}$$

wobei λ die hadronische Wechselwirkungslänge des Materials ist. Ein solches sogenanntes *Sail–Through–Pion* verhält sich wie ein minimalionisierendes Teilchen und täuscht im instrumentierten Eisen ein Myon vor.
Die Dicke des Flüssig–Argon–Kalorimeters beträgt im Zentralbereich vier bis sechs hadronische Wechselwirkungslängen, so daß die Wahrscheinlichkeit eines Pions für Sail–Through fast 1% beträgt [Sel95] (vgl. Abbildung 2.8) und damit eine große Quelle für Myonen bildet.

4.4 Ereignisgeneratoren

Der Abschätzung der Reaktion des H1–Detektors auf physikalische Prozesse sowie dem Vergleich der theoretischen Modelle mit den Messungen dienen *Ereignisgeneratoren*.

Zunächst werden die Vierervektoren der Partonen die in einer möglichen Elementarteilchenreaktion entstehen von dem Generatorprogramm anhand der theoretischen Wirkungsquerschnitte (zufällig) erzeugt (*Monte–Carlo–Methode*). Nach dem Partonschauer und der Fragmentation werden daraus die stabilen Teilchen der Hadronisation *generiert*[2]. Nachfolgend werden die Zerfälle der kurzlebigen Teilchen und die vollständige Detektorantwort auf einzelne Teilchen oder Ereignisse mit dem Programmpaket `H1SIM` [MeyJ91] unter Berücksichtigung des Energieverlustes, der Vielfachstreuung und des Magnetfeldes *simuliert*. Lediglich der Zerfall von Teilchen mit einer Lebensdauer kleiner als 8 ns, speziell der Zerfall der schweren B– und D–Mesonen, wird bereits im Generatorschritt vollzogen.

---

[2] Diese Teilchen werden auch mit Teilchen auf Generator–Niveau bezeichnet, abgekürzt GTR für **G**enerator **Tr**acks.

Nach der Simulation stehen idealerweise die gleichen Informationen, wie bei experimentell aufgenommenen Daten zur Verfügung. Zuletzt werden die generierten und simulierten Ereignisse mit dem gleichen Verfahren, welches auf die Daten angewendet wird, rekonstruiert (Abschnitt 2.2.5).

Die für diese Analyse maßgebenden Monte–Carlo–Generatoren sind `AROMA 2.1` und `2.2`, für die Erzeugung von Beauty oder Charm, und `IJRAY/PYTHIA`, für die Erzeugung der leichten Quarks u, d und s.

4.4.1 Ereignisgenerator `AROMA`

Mit dem Monte–Carlo–Generator `AROMA` (Version `2.1` [Ing95] und `2.2` [Ing96]) werden Ereignisse mit schweren Quarks durch die Boson–Gluon–Fusion generiert, wobei die Massen der Quarks berücksichtigt werden. QCD–Strahlungskorrekturen höherer Ordnungen in α_s werden durch Partonschauer im Anfangs– und Endzustand genähert. Die Berechnung des Partonschauers, der Fragmentation und die Zerfälle der schweren Quarks wird von dem Programm `JETSET 7.4` [Sjo93] durchgeführt.

Die für die vorliegende Analyse zur Verfügung stehenden Monte–Carlo–Ereignisse wurden mit einem schweren Quarkpaar ($b\bar{b}$ oder $c\bar{c}$) in der Boson–Gluon–Fusion unter Austausch eines direkten Photons generiert. Die Protonstrukturfunktion folgt der Parametrisierung MRS(G) [Mar95]. Die theoretischen Wirkungsquerschnitte sind

$$\sigma(ep \to b\bar{b}\,X) = 3.2\,\mathrm{nb}$$

$$\sigma(ep \to c\bar{c}\,X) = 612.5\,\mathrm{nb}.$$

Es wurde bereits vor der Simulation der Detektorantwort mindestens ein Myon mit einem minimalen Transversalimpuls im Polarwinkelbereich $4.6^\circ < \Theta_\mu < 174.8^\circ$ verlangt. Die genaue Auflistung aller Schnitte, die Anzahl der erzeugten Ereignisse und die generierte Luminosität $\mathcal{L}$ findet sich in Tabelle 4.2.

4.4.2 Ereignisgenerator `IJRAY / PYTHIA`

Zur Untersuchung des Beitrages der leichten Quarks, bei denen die Myonen im wesentlichen aus dem Zerfall leichter Hadronen entstehen oder durch Punch–/Sail–Through–Prozesse vorgetäuscht werden, wird das Monte–Carlo–Programm `PYTHIA 5.7` [Sjo93] benutzt. Der Photonfluß wird von dem Programm `IJRAY` [Abt93] generiert. Die Hadronisierung der erzeugten leichten Quarks wird mit dem Programm `JETSET 7.4` nach dem Modell der String–Fragmentation durchgeführt. Der Zerfall der leichten Hadronen z.B. in ein Myon findet im Simulationsschritt statt und die Simulation des Punch–/Sail–Through erfolgt im Rekonstruktionsschritt.

Mit `IJRAY/PYTHIA` wurden Ereignisse mit leichten Quarks unter Austausch eines aufgelösten Photons mit $Q^2 < 4\,\mathrm{GeV}^2$ generiert. Für die Protonstrukturfunktion wurde GRV

LO [Glu90] benützt, für die Photonstrukturfunktion GRV–G LO [Glu92]. Bei einem theoretischen Wirkungsquerschnitt von 10 609 nb beträgt die generierte Luminosität 3.76 pb^{-1}. Damit diese hohe Luminosität erreicht werden kann, ohne den zeitaufwendigen Schritt der Simulation zu häufig ausführen zu müssen, wurde mindestens ein Jet mit einer transversalen Energie von $E_{t,Jet} > 6$ GeV auf Generator–Niveau gefordert[3]. Die Jet–Findung geschieht mit Hilfe des Such–Algorithmus `QJCONE` [H1C94], die verwendete Parametereinstellung ist der Vollständigkeit halber in Tabelle 4.1 angegeben. Um den Untergrund durch Punch–/Sail–Through–Pionen zu forcieren, wurde mindestens eine Spur mit $p > 2$ GeV im Polarwinkelbereich $35° < \Theta < 127°$ gefordert.

Rapiditätsbereich (η–Bereich)	2.5 bis -2.5
Gittereinteilung in η	25
Gittereinteilung in Φ	24
Energie der Initiatorzelle	0.3 GeV
Gesamtenergie	6.0 GeV
halber Jet–Öffnungswinkel	1.0

Tabelle 4.1: *Verwendete Parametereinstellung für den Jetalgorithmus* `QJCONE`.

Abschließend zeigt Tabelle 4.2 eine Zusammenfassung der in der vorliegenden Analyse benutzten Ereignisgeneratoren, die Generatorschnitte und die zu Verfügung stehenden Ereignisssätze.

Generator	$\mathcal{L}$	Schnitte auf Generator–Niveau	generierte Ereignisse nach Generator–Schnitten
		mindestens ein Myon in $4.6° < \Theta_\mu < 174°$ mit	
`AROMA 2.1` $b\bar{b}$	$\sim 32\,\mathrm{pb}^{-1}$	$p_{t,\mu} > 1$ GeV	15 755
`AROMA 2.2` $b\bar{b}$	$\sim 64\,\mathrm{pb}^{-1}$	$p_{t,\mu} > 0.9$ GeV	34 450
`AROMA 2.1` $c\bar{c}$	$\sim 10\,\mathrm{pb}^{-1}$	$p_{t,\mu} > 1$ GeV	59 947
`AROMA 2.2` $c\bar{c}$	$\sim 33\,\mathrm{pb}^{-1}$	$p_{t,\mu} > 0.9$ GeV	145 338
`IJRAY/PYTHIA`	$\sim 3.76\,\mathrm{pb}^{-1}$	mindestens ein Teilchen in $35° < \Theta_\mu < 127°$ mit $p > 2$ GeV und ein Jet mit $E_{t,Jet} > 6$ GeV	339 136

Tabelle 4.2: *Zusammenfassung der verwendeten Monte–Carlo–Generatoren. Die Luminosität ergibt sich aus den bei der Generierung verwendeten theoretischen Wirkungsquerschnitten.*

---

[3] Die Generierung von $\mathcal{L} \simeq 1\,\mathrm{pb}^{-1}$ benötigte etwa 2 Monate. Ohne die Forderung nach einem Jet würde sich die benötigte Zeit für die gleiche Luminosität um einen Faktor 100 erhöhen.

Kapitel 5

Ereignis–Selektion

> Die Entscheidung, die wir zu fällen haben,
> ist eine Entscheidung unter Physikern.
> Wir müssen wissenschaftlich vorgehen.
> Wir dürfen uns nicht von Meinungen bestimmen lassen,
> sondern von logischen Schlüssen.
> Wir müssen versuchen, das Vernünftigste zu finden.
> *Friedrich Dürrenmatt in „Die Physiker“, 1962*

Mit den in Kapitel 4 beschriebenen generierten, simulierten und rekonstruierten Monte–Carlo–Ereignissen soll ein Neuronales Netz zur Unterscheidung der Ereignisse mit semimyonischen Zerfällen der Beauty– und Charm–Quarks untereinander und von den Untergrund–Ereignissen trainiert werden. In diesem Kapitel erfolgt eine Vorselektion um einerseits Myonen anzureichern und um andererseits, basierend auf den anzuwendenden Schnitten, dem Neuronalen Netz einheitlich simulierte Ereignisse aus den schweren Quark–Zerfällen und dem Untergrund präsentieren zu können.

Um die zu trennenden Ereignisse genauer zu definieren, erfolgt in Abschnitt 1 eine Einteilung der Ereignisse mit Myonen in drei Klassen, entsprechend der Herkunft des Myons. Maßgeblich bestimmend für das Neuronale Netz sind die mit dem Myon zusammenhängenden Ereignisvariablen (An das Myon werden deshalb grundlegende Bedingungen gestellt). Um das Training des Neuronalen Netzes nicht zu beeinflussen, dürfen keine Unterschiede, die auf verschiedenen Schnitten in der Generierung basieren, in den Ereignissätzen bestehen. Die mit `IJRAY/PYTHIA` produzierten Ereignisse wurden nur für kleine Viererimpulsüberträge generiert. Diese Forderung muß auch den generierten Beauty– und Charm–Ereignissen auferlegt werden (Abschnitt 2).
Abschnitt 3 definiert durch das Myon eine topologische Unterteilung des Ereignisses in zwei Hemisphären. Ein Schnitt in der Hemisphärenenergie berücksichtigt die Forderung nach mindestens einem Jet mit einer Schwellenenergie bei der Generierung der `IJRAY/PYTHIA`–Ereignisse.
In Abschnitt 4 werden die Schnitte in der Anzahl der Spuren erläutert.
Abschließend folgt in Abschnitt 5, Tabelle 5.9, eine Zusammenfassung aller Schnitte.

Selektionskriterien der Myonklasse
• ≥ 1 Myon (identifiziert im instrumentierten Eisen oder FMD) oder • ≥ 2 Myonen (identifiziert im instr. Eisen, FMD oder LAr–Kalorimeter) oder • 1 Myon (instr. Eisen, FMD oder LAr–Kalorimeter) und ≥ 1 Elektron oder • ≥ 1 Myon (instr. Eisen, FMD oder LAr–Kalorimeter) und ≤ 5 weitere Spuren im Spurkammersystem
Spuren im Spurkammersystem
• Nur vertexangepaßte Spuren mit $\mid z_0 \mid \ \leq \ 40$ cm und • Zentrale Spuren: Treffer ≥ 10 in den Spurkammern und $R_{Start} < 30$ cm und • Vorwärtsspuren: $N_{planare\,Module} \geq 1$ Vertexanpassung $p \geq 1.0$ GeV $N_{Module} \geq 2$ für $\Theta_{Vertex} \leq 20°$ $\chi^2_{Spuranpassung}\ /\ N_f \leq 10$ $\chi^2_{Vertexanpassung}\ /\ N_f \leq 10$ $R_0 \leq 5.0$ cm und • Kombinierte Spuren: keine weiteren Schnitte
Spuren im Myonsystem
• Spur im instrumentierten Eisen mit $\rho \leq 100.0$ cm im Zentralbereich Drahtlagen (FEC/Zentralbereich/BEC) $\geq 6/2/3$ (ohne Myonboxen) oder • Spur im FMD mit $-400.0\ cm \leq z_{Vertex} \leq 300.0$ cm $\chi^2_{Anpassung} \leq 20.0$ $\mid \Phi_{erster\ Treffer} - \Phi_{letzter\ Treffer} \mid \leq 1.0$ rad
Verbindung zwischen Spur in Spurkammern und Myonsystem
• $P(\chi^2) \geq 0.001$
Leptonidentifikation im LAr–Kalorimeter
• Myon: mindestens *normale* Qualität [ScG94]
Reduktion kosmischer Myonen

Tabelle 5.1: *Selektion der Myonklasse (Ereignisklasse 24).*

5.1 Ereignisse mit Myonen

5.1.1 Ereignisklasse 24

Die Basis für diese Analyse bilden Ereignisse der *Ereignisklasse 24*, die mit den Monte–Carlo–Generatoren `AROMA` oder `IJRAY/PYTHIA` generiert, simuliert und anschließend rekonstruiert wurden (Abschnitt 2.2.5 und 4.4). Ereignisse mit Myon–Kandidaten werden nach der Rekonstruktion in die Klasse 24 eingestuft, wenn sie den Kriterien in Tabelle 5.1 genügen [MeyA95]. Die einzelnen Anforderungen sind hier nur der Vollständigkeit halber angegeben und sollen nicht weiter erläutert werden.
Tabelle 5.2 gibt die Anzahl aller generierten Ereignisse sowie die Anzahl der Ereignisse nach Forderung der Myonklasse an (vgl. Tabelle 4.2).

	`AROMA` $b\bar{b}$	`AROMA` $c\bar{c}$	`IJRAY/PYTHIA` uds
Generiert (nach Generatorschnitten)	50 205	205 285	339 136
Klasse 24	35 053	98 798	106 079

Tabelle 5.2: *Anzahl von Ereignissen der Myonklasse getrennt nach Monte–Carlo–Generatoren.*

5.1.2 Klassifikation der Ereignisse

Um die Unterscheidung zweier Klassen mit dem Neuronalen Netz trainieren zu können, müssen diese Klassen genau definierte Muster enthalten. Die mit `AROMA` generierten Ereignisse enthalten neben den Myonen aus b– oder c–Quark–Zerfällen auch Myonen aus den Zerfällen leichter Hadronen und vorgetäuschte Myonen. Mit `IJRAY/PYTHIA` wurden hauptsächlich leichte Quarkpaare generiert, die Erzeugung schwerer Quarks jedoch nicht explizit unterbunden. Deshalb werden die generierten Ereignisse aus `AROMA` und `IJRAY/PYTHIA` in drei Klassen bezüglich der Herkunft des Myons aufgeteilt.
Ereignisse aus `AROMA`, in denen ein b– oder $\bar{b}$–Quark semimyonisch zerfällt, $b \to \mu\ X$, ergeben zusammen mit dem Kaskaden–Zerfall $b \to c\ X' \to \mu\ X$ die Muster der Klasse „$b\bar{b}$“. Die semimyonischen Zerfälle $c \to \mu\ X$, in denen das charmed Quark nicht aus einem schwachen b– oder $\bar{b}$–Zerfall entsteht, kommen in die Klasse „$c\bar{c}$“ (Abschnitt 4.2.4).
Die mit `IJRAY/PYTHIA` generierten Ereignisse mit Myonen aus dem Zerfall leichter Hadronen und den vorgetäuschen Myonen (Abschnitt 4.3) werden der Klasse der Untergrund–Ereignisse „uds“ zugeordnet. Die in Kapitel 3 eingeführten Klassen „0“ und „1“ wurden hiermit auf die Beauty–, Charm– und Untergrund–Klasse $b\bar{b}$, $c\bar{c}$ und uds spezifiziert[1].

Die Klassifikation erfolgt mit Hilfe des Myons mit dem größten Impuls auf Generator–Niveau ($p_{max,\mu_{GTR}}$), da die Zerfälle der schweren Quarks bereits in der Generierung erfolgen. Der Teilchenkode [PDG96] seines Mutterteilchens zeigt, ob es sich dabei um ein

---

[1] Eine genauere Einteilung folgt in Kapitel 7.

Teilchen mit Beauty oder Charm und damit um ein Myon aus den Klassen $b\bar{b}$ oder $c\bar{c}$ handelt. Die Klasse *uds* soll frei von b– und c–Quarks ein, deshalb werden Ereignisse nur dann akzeptiert, wenn *kein* Mutterteilchen der Myonen auf Generator–Niveau aus schweren Quarks besteht[2]. Damit ist sichergestellt, daß die Myonen des Untergrundes nur aus den Zerfällen leichter Hadronen entstehen bzw. von leichten Hadronen vorgetäuscht wurden, da dies erst in der Simulation bzw. Rekonstruktion geschieht.

Es ergeben sich folgende Bedingungen :

- `AROMA` $b\bar{b}$:
  Mutterteilchen des Myons auf Generator–Niveau mit $p_{max,\mu_{GTR}}$ enthält Beauty ($b \rightarrow \mu\, X$) oder Charm ($b \rightarrow c\, X' \rightarrow \mu\, X$)
- `AROMA` $c\bar{c}$:
  Mutterteilchen des Myons auf Generator–Niveau mit $p_{max,\mu_{GTR}}$ enthält Charm ($c \rightarrow \mu\, X$)
- `PYTHIA/IJRAY` *uds* :
  Mutterteilchen aller Myonen auf *Generator–Niveau* enthalten weder Beauty noch Charm ($b \nrightarrow \mu\, X$ und $c \nrightarrow \mu\, X$)

Tabelle 5.3 listet die Mutterteilchen der Myonen getrennt nach dieser Klassifikation auf. Ereignisse, in denen das Mutterteilchen der Klassifikation genügt, werden akzeptiert (im folgenden mit „GTR–μ O.K.“ bezeichnet). Es verbleiben die in unter Punkt 1 der Tabelle 5.9 angegebene Anzahl von Ereignissen.

5.1.3 Die Spur– und Myonselektion

Die Spurselektion

Diese Analyse basiert auf Spuren, die im Spurkammersystem und im Zentralbereich des instrumentierten Eisens rekonstruiert wurden und an den Ereignisvertex angepaßt wurden. Die Spurselektion nach L. West [Wes96] stellt Anforderungen an die Qualität der Spuren und löst Mehrdeutigkeiten in der Zuordnung von Spuren und Teilchen auf (Tabelle 5.4).

Es folgen Erläuterungen zu Tabelle 5.4 :

- Bedingungen für Vorwärtspuren
  - R_0 ist der in der $R\Phi$–Ebene gemessene Abstand der Spur von dem Ereignisvertex. Zur Berechnung wird die Spur anhand ihrer Parametrisierung in Richtung des Vertex extrapoliert, ohne den Vertex als Ursprung der Spur anzunehmen.

---

[2] Dabei ist zu beachten, daß bei der Generierung der Untergrund–Klasse *keine* Forderung nach einem Myon auf Generator–Niveau gestellt wurde (Abschnitt 4.4.2).

Mutterteilchen	`AROMA` $b\bar{b}$	`AROMA` $c\bar{c}$	`IJRAY/PYTHIA` uds
	$b\bar{b}$	$c\bar{c}$	uds
B^+	**10 648**		99
B^0	**10 753**		93
B_s^0	**3 183**		22
B_c^+	**2**		
Λ_b^+	**2 046**		17
Ξ_b^+	**147**		1
Ξ_b^0	**132**		
Ω_b^0	**3**		
D^+	**2 513**	**36 632**	1 214
D^0	**3 261**	**49 905**	1 646
D_s^+	**964**	**8 560**	333
J/ψ	**300**		9
Λ_c^+	**199**	**2 424**	144
Ξ_c^+	**28**	**337**	13
Ξ_c^0	**23**	**355**	14
Ω_c^0		**6**	
η	4	27	**102**
ρ	6	8	**30**
τ	841	544	**29**

Tabelle 5.3: *Mutterteilchen des Myons auf Generator–Niveau. Für die Klassen $b\bar{b}$ und $c\bar{c}$ wurde das Mutterteilchen mit dem größten Impuls betrachtet, für die Klasse uds die Mutterteilchen aller Generator–Niveau–Myonen. Fett unterlegt sind die akzeptierten Mutterteilchen. Zusätzlich zu den akzeptierten Myonen auf Generator–Niveau der Klasse uds werden auch die in der Simulation entstehenden Myonen und die vorgetäuschten Myonen akzeptiert (hier nicht aufgelistet).*

Spuren die außerhalb von R_0 um den Vertex liegen, stammen eher aus Zerfällen langlebiger Teilchen, von kosmischen Myonen oder aus Reaktionen von Protonen oder Elektronen mit den Atomen und Molekülen des Restgases im Strahlrohr oder seinen Außenwänden.

- χ^2_{Spur} / N_f und χ^2_{Vertex} / N_f sind die χ^2–Werte der Spur– bzw. Vertexanpassung pro Freiheitsgrad N_f.
- In den drei Modulen der vorderen Spurkammern können Segmente der Vorwärtsspuren erkannt werden. Mit den Bedingungen für $N_{plan.\&\,rad.\,Segm.}$ und $N_{1.\&\,2.plan.\,Segm.}$ lassen sich diese Spursegmente in drei Qualitätsstufen einordnen :
  1. Primäre Segmente haben je mindestens drei Treffer in allen drei Orientierungen eines planaren Moduls.

Vorwärtsspuren		
p_t	$>$	0.150 GeV
Θ	$>$	7.0°
	$<$	25.0°
R_0	$<$	10.0 cm
χ^2_{Spur} / N_f	$<$	10.0
χ^2_{Vertex} / N_f	$<$	25.0
$N_{1.\&\,2.plan.Segm.}$	$\geq$	1
$N_{plan.\&\,rad.Segm.}$	$\geq$	2
σ_p / p	$<$	1.0
Zentrale Spuren		
p_t	$>$	0.150 GeV
Θ	$>$	20.0°
	$<$	160.0°
$\|DCA'\|$	$<$	2.0 cm
R_{Start}	$<$	50.0 cm
Spurlänge	$>$	10.0 cm für $\Theta < 150°$
Spurlänge	$>$	5.0 cm für $\Theta > 150°$
Kombinierte Spuren		
p_t	$>$	0.150 GeV
Θ	$>$	10.0°
	$<$	27.0°
$\|DCA'\|$	$<$	5.0 cm
R_{Start}	$<$	50.0 cm

Tabelle 5.4: *Bedingungen für die Spurselektion.*

2. Sekundäre Segmente haben je mindestens drei Treffer in zwei Orientierungen und mindestens einen Treffer in der dritten Orientierung eines planaren Moduls.
3. Tertiäre Segmente haben je mindestens drei Treffer in zwei Orientierungen des planaren Moduls.

– σ_p / p gibt den Fehler des Spurimpulses an.

- Bedingungen für zentrale Spuren
  - | *DCA'* | (**D**istance of **C**losest **A**pproach) bildet den geringsten Abstand der nicht vertexangepaßten Spur zum Primärvertex in der $R\Phi$–Ebene. Für Spuren schnell zerfallender Teilchen ist der | *DCA'* | klein.

  - R_{Start} gibt den am dichtesten am Strahlrohr liegenden Treffer im Spurkammersystem an.
  - Die Spurlänge ist definiert als radialer Abstand zwischen Anfangs– und Endpunkt der Spur. Bedingt durch die Konstruktion des zentralen Spurkammersystems können Teilchen mit $\Theta > 150°$ nicht mehr alle radialen Lagen durchqueren.

- Kombinierte Spuren setzen sich aus einer zentralen und einer Vorwärtsspur zusammen. Für beide Spuren wird eine gemeinsame Parametrisierung berechnet. Aufgrund des Aufbaus des Spurkammersystems können kombinierte Spuren nur für $10° < \Theta < 27°$ vorkommen.

Bei der Rekonstruktion der einzelnen Spursegmente, beim Zusammenfügen zu einer kombinierten Spur und bei der Anpassung an einen der Vertizes bestehen häufig mehrfache Möglichkeiten. Dies führt zu mehreren rekonstruierten Spuren für ein Teilchen, *Ambiguität* genannt. Um Ambiguitäten aufzulösen, werden nur an den primären Vertex angepaßte Spuren betrachtet. Liegen Ambiguitäten durch das Auftreten von

kombinierten Spuren, zentralen Spuren, Vorwärtsspuren

vor, so wird eine der Spuren in der angegebenen Reihenfolge ausgewählt.

Die Myonselektion

Die Myonenselektion stellt Anforderungen an die Qualität der Spur im instrumentierten Eisen und seiner Verbindung zum Spurkammersystem. Da diese Analyse nur Spuren im Zentralbereich betrachtet, sind andere Bedingungen (Kalorimeter, vordere und hintere Endkappe) nicht aufgelistet (Tabelle 5.5).

ρ	$<$	100 cm
$\lvert z_0 \rvert$	$<$	100 cm
N_{Lagen}	$\geq$	2
i_{Start}	$\leq$	5
i_{Ende}	$\geq$	2

Tabelle 5.5: *Bedingungen für die Myonselektion für Spuren im Zentralbereich.*

Es folgen Erläuterungen zu Tabelle 5.5 :

- ρ ist der geringste Abstand in $R\Phi$ der extrapolierten Spur aus dem instrumentierten Eisen zum Primärvertex (entspricht dem *DCA'* der Spurkammern).
- z_0 ist die z–Koordinate des Schnittpunktes der extrapolierten Spur mit der Strahlachse.

- N_{Lagen} ist die Anzahl der im instrumentierten Eisen angesprochenen Lagen (ohne die Myonboxen).
- i_{Start} bzw. i_{Ende} ist die Nummer der ersten bzw. letzten angesprochenen Lage (vom Vertex aus betrachtet).

Diese Analyse wird sich auf Myonen beschränken, die im instrumentierten Eisen nachgewiesen wurden. Dazu müssen Spuren im instrumentierten Eisen existieren (Einträge in der `DMUO`–Bank), die die Bedingungen der Myonselektion (Tabelle 5.5) erfüllen. Außerdem muß eine Verbindung der Myonspur mit Spuren des Spurkammersystems existieren, deren Anpassungswahrscheinlichkeit größer als 10^{-3} ist (Abschnitt 2.2.5).
Die Anzahl der verbleibenden Ereignisse ist in Tabelle 5.9, Punkt 2, aufgeführt.

5.1.4 Impuls– und Winkelbereiche der Myonen

Impulsschnitte Um ein Myon im Zentralbereich des instrumentierten Eisens nachweisen zu können, muß es zuerst die Spurkammern und das Flüssig–Argon–Kalorimeter durchqueren. Die Verbindung der Spuren im Spurkammersystem und im instrumentierten Eisen kann außerdem bei kleinen Impulsen zu Fehl–Anpassungen führen. Aus diesen Gründen wird dem Myon ein Impuls von $p_\mu > 2$ GeV abverlangt (Abbildung 5.1).

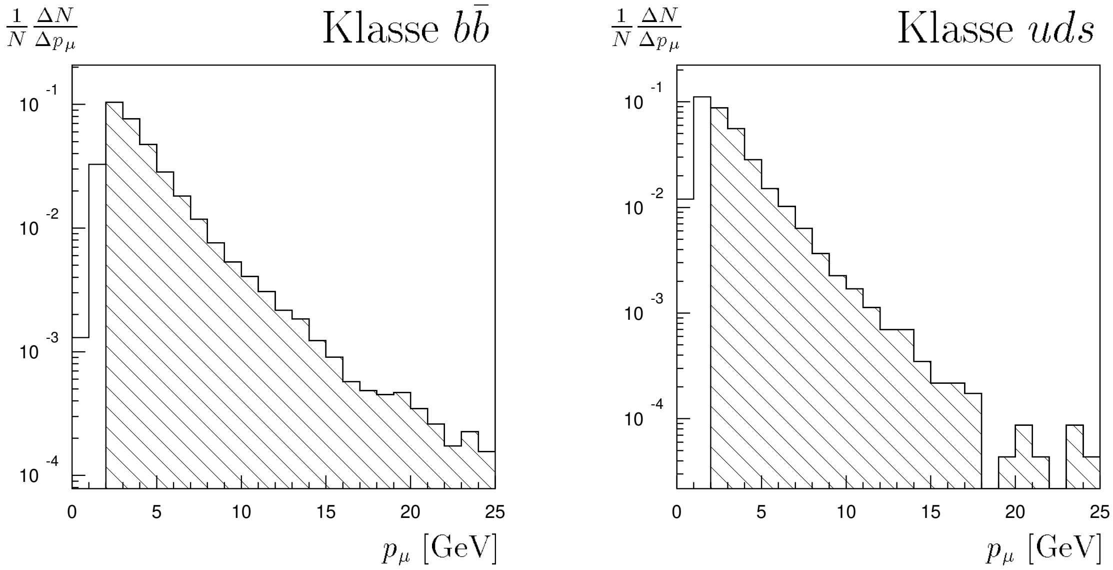

Abbildung 5.1: *Die Verteilung des Impulses aller im Eisen nachgewiesenen Myonen. Die schraffierte Fläche gibt die verbleibenden Ereignisse nach dem Schnitt $p_\mu > 2$ GeV an.*

Begrenzung des Polarwinkels Die Definition einer der Eingabegößen für das Neuronale Netz, die Aktivität, erfordert einen Schnitt auf den Polarwinkel des Myons von $35° < \Theta_\mu < 127°$ (Abschnitt 6.2), da zum einen das Vorwärts–Spurkammersystem bei etwa 7° anfängt und zum anderen das Flüssig–Argon–Kalorimeter, dessen Energiedeposition

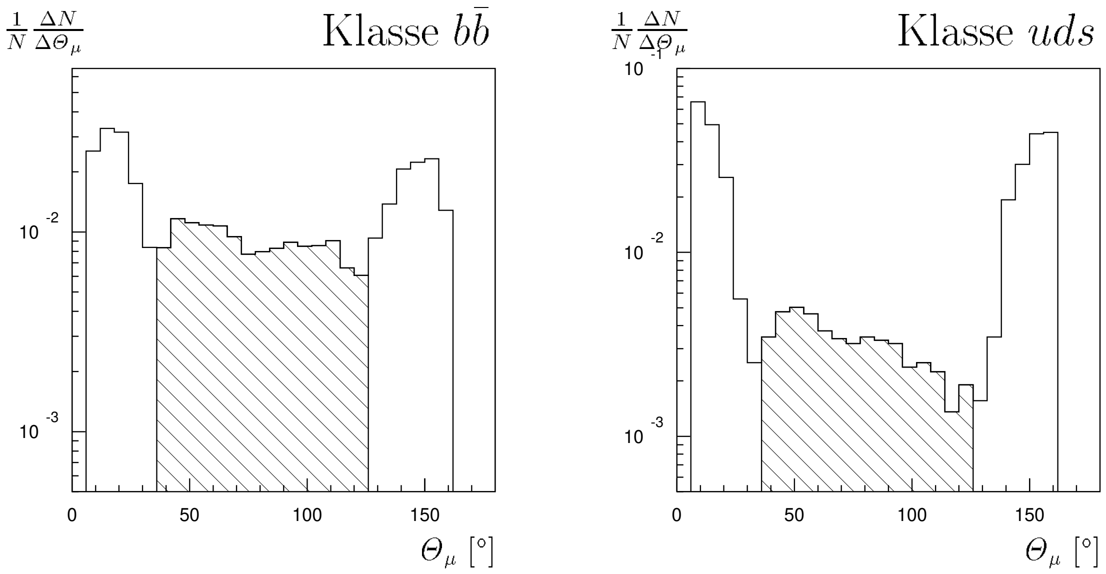

Abbildung 5.2: *Polarwinkelverteilung aller im Eisen nachgewiesenen Myonen mit* $p_\mu > 2\,GeV$ *vor und nach (schraffiert) dem Schnitt* $35° < \Theta_\mu < 127°$ *für die Klassen* $b\bar{b}$ *und* uds*. Ebenfalls erkennbar ist der Akzeptanzbereich der Spurselektion für Vorwärtsspuren, kombinierte und zentrale Spuren von* $7° < \Theta < 160°$ *(vgl. Tabelle 5.4).*

in die Berechnung der Aktivität eingeht, bei $\Theta \simeq 154°$ endet. Im Rückwärtsbereich nimmt außerdem mit wachsendem Polarwinkel das Detektormaterial zwischen Wechselwirkungszone und instrumentiertem Eisen ab und damit die Kontamination mit vorgetäuschten Myonen zu. Deshalb wird für alle Klassen $35° < \Theta_\mu < 127°$ gefordert (Abbildung 5.2).

Das Myon mit dem größten Impuls, welches das Impuls– und Polarwinkelfenster passiert und im instrumentierten Eisen gefunden wurde, wird im folgenden mit „impulsstärkstes“ oder einfach nur „das“ Myon genannt. Andere Myonen werden in dieser Analyse nicht weiter berücksichtigt.
Nachdem alle Anforderungen an das Myon gestellt wurden, verbleiben die in Tabelle 5.9, Punkt 3 und 4, angegebenen Ereignisse.

5.2 Ereignisse der Photoproduktion

Eine weitere Einschränkung bedingt die Generierung der Klasse *uds*. Für diese Klasse wurde bei der Generierung $Q^2 < 4\ \text{GeV}^2$ gefordert (Abschnitt 4.4.2). Damit die drei Klassen miteinander verglichen werden können, dürfen die Beauty– und Charm–Klassen ebenfalls nur Ereignisse mit $Q^2 < 4\,\text{GeV}^2$ enthalten. Um der Ungenauigkeit in der Q^2–Bestimmung zu entgehen, werden entsprechend der Definition in Abschnitt 4.1 nur Ereignisse der Photoproduktion durch die Forderung

$$Q^2 < 1\ \text{GeV}^2 \tag{5.1}$$

selektiert.

5.2.1 Identifikation des gestreuten Elektrons

Zu einer Identifikation des gestreuten Elektrons wird im Flüssig–Argon–Kalorimeter und im SPACAL nach einer lokal begrenzten Energiedeposition (Cluster) gesucht. Die Suche erfolgt einerseits mit Hilfe der Einträge von Clustern in der `DELE`–Bank [Bas95], andererseits über den Suchalgorithmus für Elektronen `XASELE` [ScA96]. Die genaueren Bedingungen sind in den Tabellen 5.6 und 5.7 angeben und sollen hier nicht weiter erläutert werden.

Energie des Clusters mit der größten Energie	E_1	$>$	4 GeV
Energie der vier weiteren Cluster (nach ihrer Energie sortiert)	$E_2 \ldots E_5$	$>$	2 GeV

Tabelle 5.6: *1995er Bedingungen für die Elektronfindung mit der* `DELE`*–Bank.*

Energie des Elektron–Kandidaten	$ELPMIN$	$>$	1.0 GeV
transversaler Impuls des Elektron–Kandidaten	$PTCUT$	$>$	5.0 GeV
Abstand Cluster–Spur	$CLTRC$	$<$	12.0 cm
Konus zur Isolationsberechnung	$ALPHC$	$=$	-0.35
Zylinderradius um das Elektron auszuschließen	$CYLCUT$	$=$	15.0 cm
hadronische Energieskala	$HSCAL$	$=$	2
elektromagnetische Energieskala	$SCALE$	$=$	1
transversale Größe des Clusters	$TVOLC$	$<$	3.1
transversale Cluster–Energie	$TMASC$	$<$	3.75 GeV
Konus zur Bestimmung der Clustergröße	$SCONE$	$<$	0.4
Isolationskriterium gegen hadronische Energiedeposition	$FISOL$	$<$	0.05
Radius des inneren Zylinders	$CYLIN$	$=$	10.0
Radius des äußeren Zylinders	$CYLOU$	$=$	25.0
elektromagnetischer Energieanteil	$EFRAC$	$>$	0.9
Korrektur auf totes Material	$IDEAD$	$=$	1

Tabelle 5.7: *Bedingungen für den Suchalgorithmus für Elektronen* `XASELE`.

Findet sich ein Elektron mit einer Energie $E_e > 7$ GeV, so wird nach Gleichung 4.1 der Viererimpulsübertrag berechnet. Läßt sich kein Elektron finden, wird angenommen, daß es durch das Strahlrohr entkommen ist, damit nur geringen Viererimpulsübertrag haben kann und zur Photoproduktion gehört.

5.2.2 Schnitt in Q^2

Mit identifiziertem Elektron und damit nach Gleichung 4.1 errechnetem Q^2 werden nun Ereignisse der tiefinelastischen Streuung mit $Q^2 > 1\ \mathrm{GeV}^2$ verworfen, Photoproduktionsereignisse mit $Q^2 < 1\ \mathrm{GeV}^2$ akzeptiert. Die Übereinstimmung von errechnetem (Q^2) und generiertem (Q^2_{MC}) Viererimpulsübertrag zeigen Tabelle 5.8 und Abbildung 5.3.

Klasse $b\bar{b}$	$Q^2_{MC} < 1\ \mathrm{GeV}^2$	$Q^2_{MC} > 1\ \mathrm{GeV}^2$
$Q^2 > 1\ \mathrm{GeV}^2$	122 (1.7%)	1 229 (17.4%)
$Q^2 < 1\ \mathrm{GeV}^2$	**5 383 (76.0%)**	**346 (4.9%)**

Klasse *uds*	$Q^2_{MC} < 1\ \mathrm{GeV}^2$	$Q^2_{MC} > 1\ \mathrm{GeV}^2$
$Q^2 > 1\ \mathrm{GeV}^2$	20 (2.7%)	43 (5.9%)
$Q^2 < 1\ \mathrm{GeV}^2$	**654 (89.6%)**	**13 (1.8%)**

Tabelle 5.8: *Korrelation zwischen gemessenem Q^2 und generiertem Q^2_{MC} für die Klassen $b\bar{b}$ und uds. Die Prozentzahlen geben den Anteil an allen Ereignisse der Klasse $b\bar{b}$ bzw. uds an, fett unterlegt sind die akzeptierten Ereignisse.*

Die Tabellenwerte links unten entsprechen der Anzahl der richtig identifizierten Ereignisse mit $Q^2 < 1\ \mathrm{GeV}^2$ bzw. $Q^2 > 1\ \mathrm{GeV}^2$ (rechts oben). Der Wert rechts unten ist die Anzahl von tiefinelastischen Ereignissen, bei denen kein Elektron gefunden wurde und die fälschlicherweise der Photoproduktion zugeordnet wurden, links oben genau entgegengesetzt : Obwohl es sich um ein Ereignis der Photoproduktion handelt, wurde ein Teilchen mit hoher Energiedeposition gefunden und als gestreutes Elektron eingestuft. Bei diesem Teil-

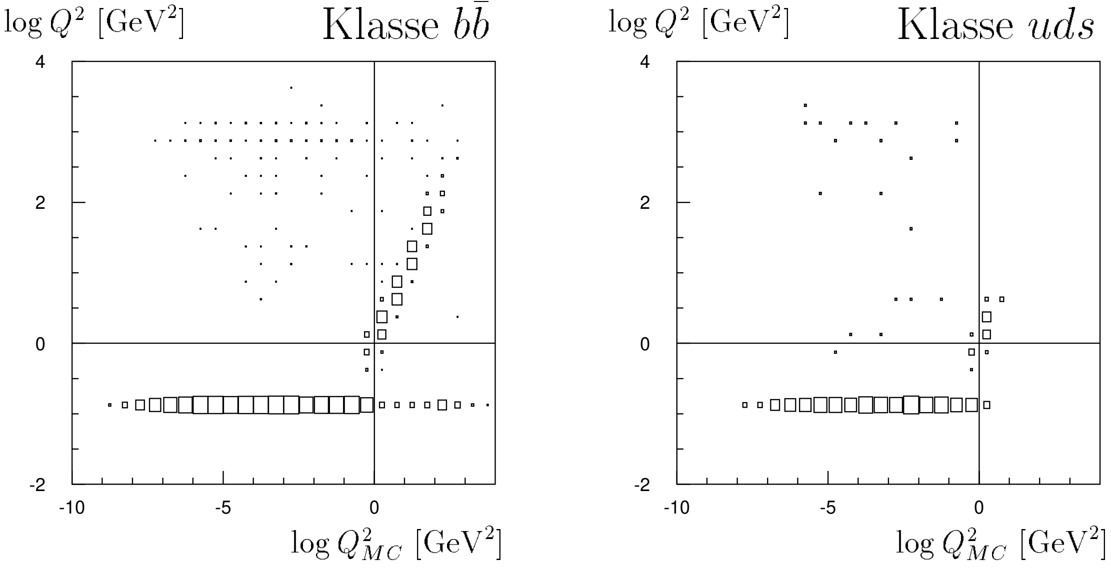

Abbildung 5.3: *Korrelation zwischen gemessenem Q^2 und generiertem Q^2_{MC} für die Klassen $b\bar{b}$ und uds.*

chen wird es sich um ein Photon oder Pion mit einer, dem Elektron ähnelnden Signatur oder um ein Zerfalls–Elektron handeln.

Die Differenz von Energie und Longitudinalimpuls

$$E - p_z = \sum_{i \,\in\, Spuren\ und\ Zellen} E_i - p_{i,z} \approx E_e + E_p - (E_p - E_e) \tag{5.2}$$

ist unempfindlich auf Verluste in Vorwärtsrichtung und gibt eine Erhaltungsgröße, mit der der Q^2–Schnitt überprüft werden kann. In tiefinelastischen Ereignissen kann das gestreute Elektron und seine Energie nachgewiesen werden, die Bilanz ergibt damit

$$E - p_z \approx 27.6\ \mathrm{GeV} + 27.6\ \mathrm{GeV} + 820\ \mathrm{GeV} - 820\ \mathrm{GeV} \simeq 55\ \mathrm{GeV}.$$

In Photoproduktionsereignissen entkommt das auslaufende Elektron im Strahlrohr und kann deshalb nicht identifiziert werden. Seine Energie und sein Impuls fehlen in der Bilanz, d.h. $E - p_z < 55$ GeV (Abbildung 5.4). Tabelle 5.9, Punkt 5, gibt die verbleibenden Ereignisse nach dem Schnitt $Q^2 < 1\ \mathrm{GeV}^2$ an.

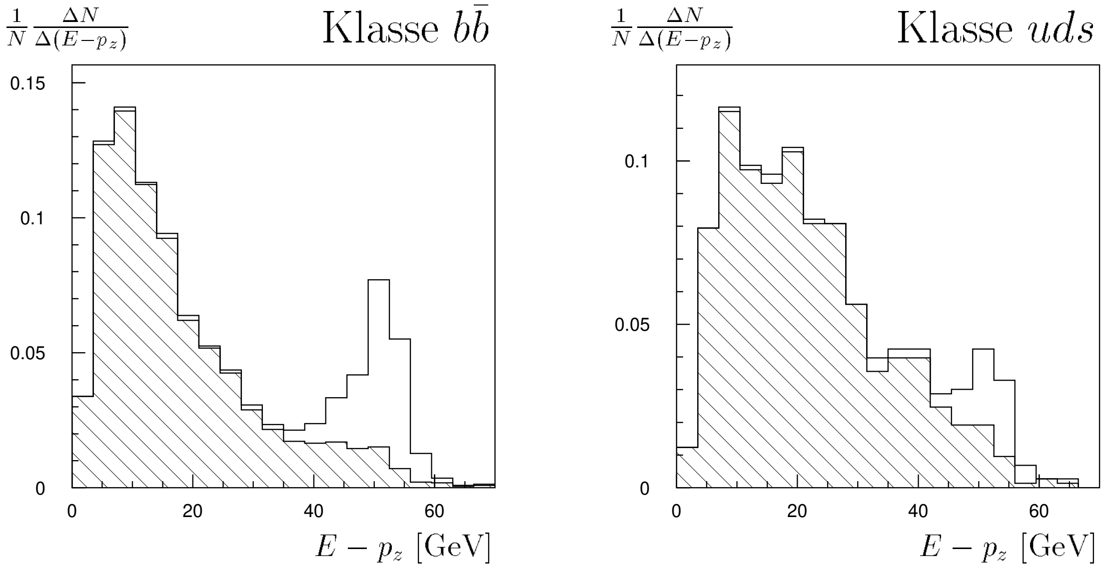

Abbildung 5.4: *$E-p_z$ der Klassen $b\bar{b}$ und uds vor und nach (schraffiert) dem Schnitt $Q^2 < 1$. Tiefinelastische Ereignisse (bei $E - p_z \approx 55\,GeV$) werden mit diesem Q^2-Schnitt unterdrückt.*

5.3 Die Hemisphären

Neben der eben berücksichtigten Forderung $Q^2 < 4\ \mathrm{GeV}^2$ wurde für die Generierung der Klasse *uds* mindestens ein Jet auf Generator–Niveau mit einer transversalen Energie von $E_{t,GTR-Jet} \geq 6$ GeV gefordert (Abschnitt 4.4.2). Dieser Schnitt wird durch die Definition der Hemisphären und der Aufsummierung der Hemisphären–Energie in den Beauty– und Charm–Klassen Berücksichtigung finden.

5.3.1 Definition der Hemisphären

Aus der Kinematik (Abschnitt 4.2.3) folgt, daß die bei der Boson–Gluon–Fusion entstehenden Quarkpaare $b\bar{b}$ oder $c\bar{c}$ in der $R\Phi$–Ebene entgegengesetzt auseinander fliegen. Ihre Zerfallsprodukte sind dann in der Flugrichtung stark korreliert und bilden jeweils einen mehr oder weniger ausgebreiteten Teilchenschauer. Das Ereignis kann damit in zwei *Hemisphären* eingeteilt werden, wovon jede Hemisphäre ein Quark oder ein Antiquark und seine Zerfallsprodukte enthalten sollte (Abbildung 5.5).

Die Korrelation der Flugrichtung des Myons mit der des ursprünglichen Quarks (Abschnitt 4.2.5) zeichnet die Richtung des Myons als eine Hemisphärenachse aus. Bei Identifikation eines Myons läßt sich damit die *μ–Hemisphäre* definieren, der alle Spuren mit

$$\Delta\Phi = |\Phi_{Spur} - \Phi_{\mu}| = \cos^{-1}\left(\frac{|\vec{p}_{t,Spur} \cdot \vec{p}_{t,\mu}|}{|\vec{p}_{t,Spur}| \cdot |\vec{p}_{t,\mu}|}\right) \leq 90^\circ \tag{5.3}$$

zugeordnet werden. Teilchen mit $\Delta\Phi > 90^\circ$ bilden die *gegenüberliegende Hemisphäre.*

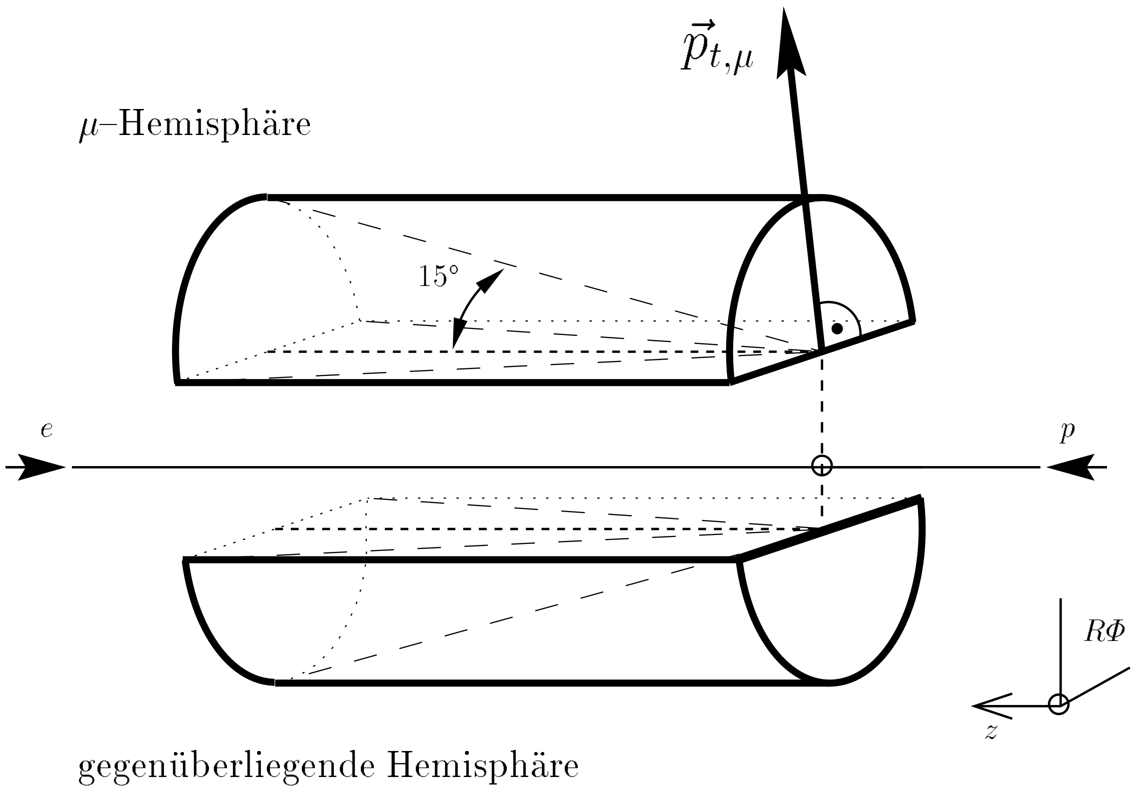

Abbildung 5.5: *Definition der Hemisphären.*

5.3.2 Schnitt in der Hemisphären–Energie

Wird über die transversale Energie aller Spuren einer Hemisphäre unter Annahme der Pionmasse m_π aufsummiert,

$$E_{t,H} = \sum_{\substack{i \in Spuren \\ in\ Hemisphäre \\ mit\ \Theta_i > 15^\circ}} \sqrt{(p_i^2 + m_\pi^2)}\ \sin \Theta_i, \tag{5.4}$$

ergibt sich eine der Forderung bei der Untergrund–Generierung $E_{t,GTR-Jet} \geq 6$ GeV entsprechende Schnittmöglichkeit. Die Begrenzung $\Theta_i > 15^\circ$ soll eine Beeinflussung der Hemisphären durch den Protonrest verhindern.

Um einen Energieschnitt anzubringen wird die Korrelation zwischen der transversalen Jet–Energie $E_{t,GTR-Jet}$ auf Generator–Niveau und der transversalen Energie einer Hemisphäre $E_{t,H}$ nach der Rekonstruktion betrachtet. Durch das impulsstärkste Myon[3] *in* einem Jet auf Generator–Niveau kann dieser *μ–Jet* der entsprechenden μ–Hemisphäre zugeordnet werden. Die Korrelation in der transversalen Energie zwischen μ–Jet und μ–Hemisphäre ist jedoch nur mäßig ausgeprägt (Abbildung 5.6).

Dies liegt zum größten Teil an den Spuren, die zwar aufgrund der Definition 5.3 der Hemisphäre zugeordnet wurden, aber nicht dem Jet auf Generator–Niveau angehören. Migration von Spuren aus dem μ–Jet in die gegenüberliegende Hemisphäre tragen dagegen wenig bei.

---

[3] Abweichend von der obigen Definition ist hier das Myon mit dem größten Impuls auf *Generator–*Niveau gemeint.

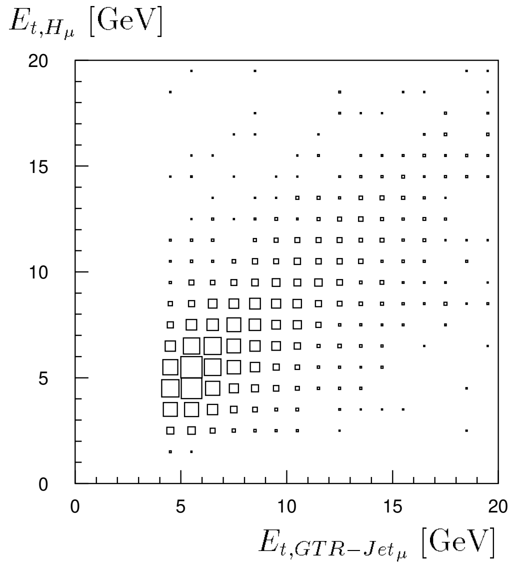

Abbildung 5.6: *Korrelation zwischen der transversalen Energie E_{t,H_μ} der μ–Hemisphäre und der transversalen Energie E_{t,Jet_μ} des Jets, der das Myon beinhaltet.*

Es kann demnach nicht behauptet werden, daß der Jet die Hemisphäre dominiert. Eine direkte Anwendung des Schnittes $E_{t,Jet} \geq 6$ GeV auf die Hemisphären–Energie scheint nicht ratsam. Anstatt dessen erfolgt der Schnitt am Maximum seiner Verteilung der Klasse *uds* durch die Forderung

$$E_{t,H} > \max E_{t,H}(uds) \tag{5.5}$$

für mindestens eine der beiden Hemisphären (Abbildung 5.7). Das Maximum der $E_{t,H}$–Verteilung liegt bei 5 GeV. Die Anzahl von Ereignissen nach dem Schnitt $E_{t,H} > 5$ GeV für jede Klasse liefert Tabelle 5.9, Punkt 6.

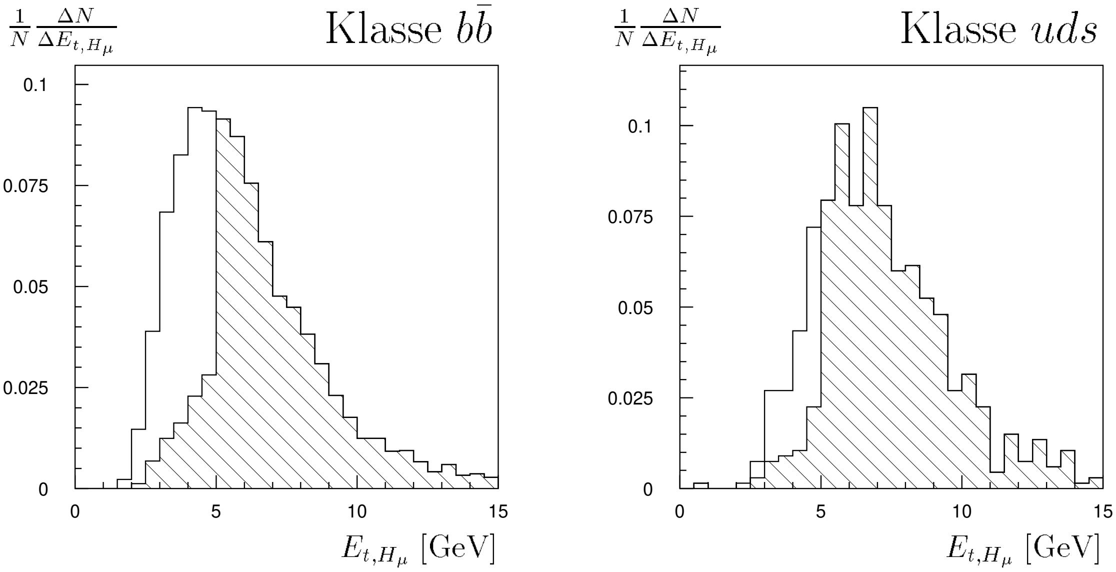

Abbildung 5.7: *Transversale Energie der μ–Hemisphäre E_{t,H_μ} für die Klassen $b\bar{b}$ und uds. Die Ereignisse nach dem Schnitt $E_{t,H_\mu} > 5$ GeV für eine Hemisphäre sind schraffiert dargestellt. In Ereignissen mit $E_{t,H_\mu} \leq 5$ GeV erfüllt die gegenüberliegende Hemisphäre den Schnitt.*

5.4 Schnitte auf die Anzahl von Spuren

Mit der Bedingung von mehr als drei Spuren im Ereignis werden inelastische Ereignisse selektiert. Gleichzeitig wird der Untergrund durch Myonen aus der kosmischen Strahlung unterdrückt, bei der im Detektor häufig nur zwei Spuren gemessen werden.

Um physikalische Ereignisgrößen sinnvoll definieren zu können, werden zusätzlich mindestens zwei Spuren in der μ–Hemispäre benötigt.

In Tabelle 5.9, Punkt 7 und 8, ist erkennbar, daß die Schnitte auf die Anzahl der Spuren keinen großen Einfluß haben.

5.5 Zusammenfassung aller Schnitte

Abschließend wird eine Zusammenfassung der Selektions–Schnitte und der Anzahl der verbleibenden Ereignisse nach jedem Schnitt gegeben (Tabelle 5.9, Punkt 1 bis 8). Vorangestellt sind die theoretischen Wirkungsquerschnitte σ und die generierten Luminositäten $\mathcal{L}$ aus Abschnitt 4.4, sowie die Anzahl der Ereignisse, die nach den Generator–Schnitten und nach der Forderung der Myonklasse verbleiben. Die ausgewählten Ereignisse N (Punkt 9) bilden die drei Klassen mit denen das Neuronale Netz zur gegenseitigen Unterscheidung trainiert werden soll. Hierbei enthält gerade die Klasse uds die geringste Statistik, obwohl, wie aus den Angaben der auf Luminosität normierten Ereignisanzahl $\mathcal{N}$ einsehbar ist, der Untergrund die Ereignisse mit Myonen bei weitem dominiert.

		$b\bar{b}$	$c\bar{c}$	uds
	theoretischer Wirkungsquerschnitt σ in nb	3.2	612.5	10 609
	generierte Luminosität $\mathcal{L}$ in pb^{-1}	$\sim$ 96	$\sim$ 43	$\sim$ 3.76
	nach Generatorschnitten	50 205	205 285	330 136
	Klasse 24	35 053	98 798	106 079
1	GTR–μ O.K.	34 202	98 219	102 724
2	Eisen–Myon	19 210	52 874	7 662
3	$p_\mu > 2$ GeV	7 080	5 470	730
4	$35° < \Theta_\mu < 127°$			
5	$Q^2 < 1$ GeV^2	5 729	4 148	667
6	$E_{t,H} > 5$ GeV für eine Hemisphäre	3 971	2 085	583
7	$N_{Spur,\,Ereignis} > 3$	3 960	2 062	582
8	$N_{Spur,\,\mu-Hemisphäre} > 1$			
9	ausgewählte Ereignisse N	3 960	2 062	582
	ausgewählte Ereignisse $\mathcal{N}$ pro 1 pb^{-1}	41.3	48.0	154.8
	Effizienz der Schnitte ε_{ES}	0.0129	$7.829 \cdot 10^{-5}$	$1.459 \cdot 10^{-5}$

Tabelle 5.9: *Zusammenfassung aller Schnitte.*

Die letzte Zeile gibt die Effizienz der Ereignis–Selektion an. Sie wird wie folgt definiert :

$$\varepsilon_{ES}^{(q\bar{q})} = \frac{N_{q\bar{q}}}{\sigma(ep \to q\bar{q}\,X)\,\mathcal{L}_{q\bar{q}}} \tag{5.6}$$

Hierbei bezeichnet $q\bar{q}$ die Klassen $b\bar{b}$, $c\bar{c}$ oder uds. Die Werte für $N_{q\bar{q}}$, σ und $\mathcal{L}_{q\bar{q}}$ entstammen für die entsprechende Klasse der Tabelle 5.9. Bei diesem Vergleich ist angenommen, daß die Trigger keine Ineffizienz zeigen.

Durch die angewendeten Schnitte wurde das auf die Luminosität normierte Verhältnis der Ereignisanzahlen zugunsten der schweren Quarks erhöht :

$$\left.\frac{\mathcal{N}^{(b\bar{b})}}{\mathcal{N}^{(uds)}}\right|_{ES} = \frac{N_{b\bar{b}}}{\mathcal{L}_{b\bar{b}}} \cdot \frac{\mathcal{L}_{uds}}{N_{uds}} = 0.2665 \tag{5.7}$$

bzw.

$$\left.\frac{\mathcal{N}^{(c\bar{c})}}{\mathcal{N}^{(uds)}}\right|_{ES} = 0.3098 \tag{5.8}$$

Für das Verhältnis der Wirkungsquerschnitte der Beauty– und Charm–Klassen untereinander ergibt sich :

$$\left.\frac{\mathcal{N}^{(b\bar{b})}}{\mathcal{N}^{(c\bar{c})}}\right|_{ES} = 0.8602 \tag{5.9}$$

Tabelle 5.10 zeigt die Herkunft der *rekonstruierten* Myonen aller ausgewählten Ereignisse der Klassen $b\bar{b}$ und $c\bar{c}$. Der Anteil von gebundenen schweren Quarkpaaren (Υ oder J/ψ) in den Ereignissen der Beauty– und Charm–Klassen wird durch die Schnitte stark unterdrückt. Die verbleibenden 55 J/ψ ($\sim$ 1.4% aller Ereignisse) sind vernachlässigbar.

Mutterteilchen des Myons	$b\bar{b}$	$c\bar{c}$
B^+	1 367	
B^0	1 399	
B_s^0	393	
Λ_b^+	269	
Ξ_b^+	20	
Ξ_b^0	6	
Ω_b^0	1	
D^+	150	708
D^0	225	1 121
D_s^+	61	163
J/ψ	55	
Λ_c^+	12	58
Ξ_c^+	1	4
Ξ_c^0	1	7
Ω_c^0		1

Tabelle 5.10: *Mutterteilchen des Myons aller ausgewählten Ereignisse der Klassen $b\bar{b}$ und $c\bar{c}$.*

Für die Klasse *uds* sind in Tabelle 5.11 die als Myon identifizierten Teilchen angegeben. Es ist zu erkennen, daß die als Myon fehlidentifizierten Pionen und Kaonen den größten Anteil bilden (oberer Abschnitt). Für die Zerfalls–Myonen wurde das Mutterteilchen separat aufgeführt (unterer Abschnitt).

Teilchen	*uds*		
π^+	210		
K^+	239		
p	12		
μ	121		
	Mutterteilchen :	η	4
		ρ	2
		τ	2
		π^+	83
		K^+	30

Tabelle 5.11: *Als Myon identifizierte Teilchen aller ausgewählten Ereignisse der Klasse uds. Für die Zerfalls–Myonen wurden zusätzlich das Mutterteilchen angegeben (unterer Abschnitt).*

Kapitel 6

Eingabegrößen für das Netz–Training

> Verhaften Sie die üblichen Verdächtigen.
> *Aus dem Film „Casablanca“, 1943*

Dieses Kapitel beschäftigt sich mit den Auswahlkriterien und der Auswahl der Eingabegrößen des Neuronalen Netzes. Diese Variablen werden aus den in Kapitel 5 vorselektierten Ereignissen bestimmt und sollen dem Neuronalen Netz eine möglichst gute Separation der Beauty–, Charm– und Untergrund–Klassen $b\bar{b}$, $c\bar{c}$ und uds voneinander ermöglichen.

Abschnitt 1 stellt die Auswahlkriterien vor, anhand derer die Eingabevariablen selektiert werden. Die Abschnitte 2, 3 und 4 definieren die topologischen Größen Aktivität, transversaler Thrust und Aplanarität. Der Einfluß des Myons wird in Abschnitt 5 durch seinen Polarwinkel, seine Anzahl der durchquerten Wechselwirkungslänge im instrumentierten Eisen und mit Hilfe der Definition des relativen Impulses bezüglich der transversalen Thrustachse beschrieben. Abschnitt 6 behandelt die Unterscheidungsmöglichkeit durch die Gesamtenergie des Ereignisses.
Um das Netz–Training zu erleichtern, wird in Abschnitt 7 eine Skalierung der Eingabevariablen vorgenommen. Abschnitt 8 gibt abschließend eine Zusammenfassung aller Eingabegrößen.

6.1 Auswahl der Eingabegrößen

Aufgabe des Neuronalen Netzes soll es sein, im mehrdimensionalen Raum eine, jeweils zwei der Klassen $b\bar{b}$, $c\bar{c}$ und uds trennende Hyperfläche zu finden. Dabei liegt der Schwerpunkt auf der b–Erkennung, so daß die Ereignisgrößen vornehmlich danach ausgesucht werden sollten.
Einfacherweise könnte das Neuronale Netz mit allen irgendwie relevanten Ereignisvariablen trainiert werden, wobei die unbedeutenden Variablen nach dem Lernvorgang vernachlässigbar kleine Gewichte bekommen sollten. Diese Methode bereitet jedoch einige Probleme. Zwar wird die Festlegung der trennenden Hyperfläche mit jeder weiteren

vernünftigen Eingabegröße erleichtert, jedoch wächst so auch die Anzahl der Eingabeknoten N_x und damit die Anzahl der Verbindungen zwischen den Lagen, also die zu bestimmenden Parameter w_{jk}. Nach der Faustregel [ScG96]

$$N_{\mathbb{T}} \simeq 10 \cdot N_x \cdot N_h \tag{6.1}$$

werden mit wachsendem N_x bedeutend mehr Muster in der Trainingsmenge $\mathbb{T}$ benötigt, welches aber schon durch die mangelnde Statistik der Klasse der Untergrund–Ereignisse limitiert ist (Kapitel 5).

Der empirische Korrelationskoeffizient

Alternativ zu der oben beschriebenen Methode wurde zuerst die Korrelation jeder einzelnen Variable x_k mit dem Soll-Ausgabewert t_i über den *empirischen Korrelationskoeffizienten* [Bro91] errechnet. Für diese Betrachtung genügt ein einziger Ausgabewert $t \equiv t_1$.

Der empirische Korrelationskoeffizient definiert sich wie folgt :

$$r_k := \frac{\sum\limits_{p\,\in\,(\mathbb{X}\cup\mathbb{Y})} \left(x_k^{(p)} - \langle x_k\rangle\right) \cdot \left(t^{(p)} - \langle t\rangle\right)}{\sqrt{\left[\sum\limits_{p\,\in\,(\mathbb{X}\cup\mathbb{Y})} \left(x_k^{(p)} - \langle x_k\rangle\right)^2\right] \cdot \left[\sum\limits_{p\,\in\,(\mathbb{X}\cup\mathbb{Y})} \left(t^{(p)} - \langle t\rangle\right)^2\right]}} \tag{6.2}$$

Hierbei gehen die Summe über alle vorhandenen Muster p zweier Klassen $\mathbb{X} \cup \mathbb{Y}$, $\langle x_k\rangle$ bzw. $\langle t\rangle$ bezeichnen jeweils das Mittel über alle Muster.
Unter der Voraussetzung eines linearen Zusammenhanges ist r_k ein Maß für die Stärke der Korrelation der Variablen x_k und dem wahren Ausgabewert t. $|r_k|$ nahe eins bedeutet einen sehr engen Zusammenhang, während $|r_k|$ nahe null einen sehr schwachen oder keinen Zusammenhang anzeigt.

Für einen Satz von 114 aus rekonstruierten Spuren und Kalorimeter–Informationen bestimmten Ereignisvariablen x_k wurde r_k für die Unterscheidung der Klassen $b\bar{b}$ und $c\bar{c}$ bzw. $b\bar{b}$ und uds errechnet, wobei für die Klasse $b\bar{b}$ $t = 1$ und jeweils für die beiden anderen Klassen $t = 0$ gesetzt wurde. Größen mit einem empirischen Korrelationkoeffizienten $|r_k| < 0.20$ für die Unterscheidung der Klassen $b\bar{b}$ und $c\bar{c}$ und $|r_k(b\bar{b}\text{ vs. }uds)| < 0.05$ wurden verworfen. Bei Größen mit gleichem physikalischer Hintergrund wurde diejenige genommen, die den besseren Korrelationskoeffizienten besitzen.
Die gesamte Auflistung für alle Variablen, der Vollständigkeit halber mit den Werten für $|r_k(c\bar{c}\text{ vs. }uds)|$, findet sich in Anhang A.

Tabelle 6.1 enthält die Werte des empirischen Korrelationskoeffizienten für die verbleibenden sieben Variablen. Die Unterscheidung der Klassen $b\bar{b}$ und $c\bar{c}$ ergibt im Mittel eine größere Korrelation als die der Klassen $b\bar{b}$ und uds. Die Werte für $|r_k(c\bar{c}\text{ vs. }uds)|$ sind nur zur Vollständigkeit angegeben.

		$\lvert r_k \rvert$		
		$b\bar{b}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$	$c\bar{c}$ vs. uds
1.	Aktivität	0.097	0.220	0.078
2.	transv. Thrust der μ-Hemisphäre	0.050	0.230	0.151
3.	Aplanarität des Ereignisses	0.040	0.213	0.160
4.	Polarwinkel des Myons	0.074	0.021	0.109
5.	Myon–Impuls bzgl. der transv. Thrustachse (μ-Hemisphäre, Thrustberechnung *exklusive* μ)	0.089	0.310	0.172
6.	Anzahl der durchquerten Wechselwirkungs–längen im instrumentierten Eisen	0.173	0.071	0.155
7.	Gesamtenergie	0.094	0.042	0.177

Tabelle 6.1: *Eingabegrößen für das Neuronale Netz mit den Werten des empirische Korrelationskoeffizienten* $\lvert r_k \rvert$ *für die Unterscheidung von* $b\bar{b}$ *und* uds, $b\bar{b}$ *und* $c\bar{c}$ *bzw.* $c\bar{c}$ *und* uds.

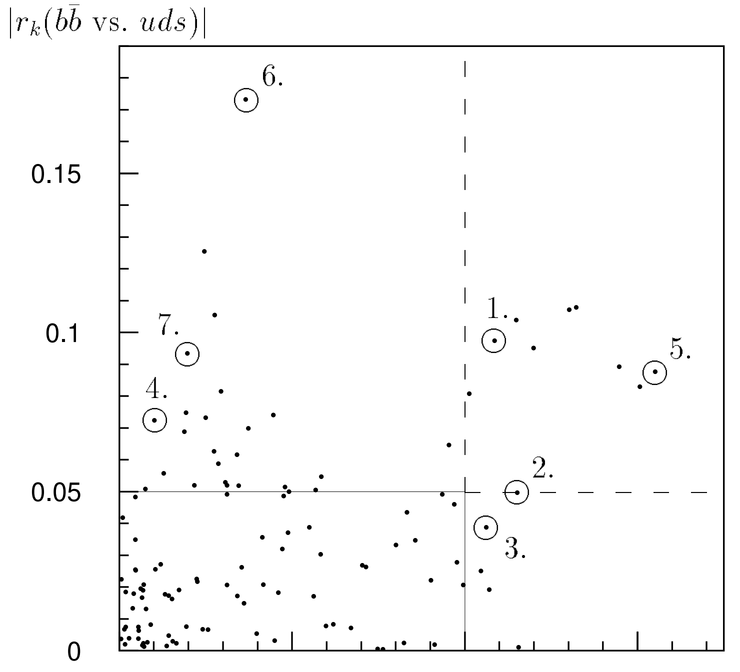

Abbildung 6.1: *Korrelation der empirischen Korrelationskoeffizienten* $\lvert r_k(b\bar{b}$ *vs.* $uds)\rvert$ *und* $\lvert r_k(b\bar{b}$ *vs.* $c\bar{c})\rvert$ *für alle 114 Variablen. Die Linien umschließen die verworfenen Variablen mit* $\lvert r_k < 0.05\rvert$ *für die Unterscheidung der Klassen* $b\bar{b}$ *und* uds *bzw.* $\lvert r_k < 0.20\rvert$ *für die Unterscheidung von* $b\bar{b}$ *und* $c\bar{c}$*. Die sieben ausgewählten Variablen sind mit einem Kreis markiert und nach Tabelle 6.1 nummeriert.*

Die genauere Betrachtung der Abbildung 6.1, in der $|r_k(b\bar{b}\text{ vs. }uds)|$ gegen $|r_k(b\bar{b}\text{ vs. }c\bar{c})|$ für alle 114 Variablen aufgetragen ist, zeigt deutlich das differenzierte Trennungsvermögen der sieben Variablen : Der Thrust (2.) und die Aplanarität (3.) sind besonders sensitiv auf die Unterscheidung der Klassen $b\bar{b}$ und $c\bar{c}$. Der Polarwinkel des Myons (4.), die Anzahl der durchquerten Wechselwirkungslänge im Eisen (6.) und die Gesamtenergie (7.) separieren die Klassen $b\bar{b}$ und uds besonders gut, während die Aktivität (1.) und der relative Impuls des Myons (5.) für beide Unterscheidungen gut sind.

Diese sieben Größen werden in den folgenden Abschnitten erläutert.

6.2 Die Aktivität

Die *Aktivität* ist ein Maß für die Energiedeposition in einem Kegel um das Myon [Alb88]. Diese Energiedeposition bestimmt sich aus dem transversalen Impuls der Spuren im Spurkammersystem und der transversalen Energie der Zellen im Flüssig–Argon–Kalorimeter mit $E_{Zelle} > 10$ MeV. Es werden nur Spuren und Zellen mitgerechnet, die innerhalb des Kegels mit dem halben Öffnungswinkel δ_{max} liegen, d.h. für den Winkelabstand δ_i, den eine Spur bzw. Zelle mit dem Myon bildet, muß gelten :

$$\delta_i = \sqrt{(\Theta_i - \Theta_\mu)^2 + \left(\frac{\Phi_i - \Phi_\mu}{2}\right)^2} < \delta_{max} \tag{6.3}$$

Der halbe Öffnungswinkel des Kegels wurde auf die Unterscheidung der Klassen durch die Aktivität hin optimiert [Kan97]. Das Optimum liegt bei $\delta_{max} = 0.5$ rad, was einer maximalen Azimutalwinkeldifferenz von 14° bzw. einer maximalen Polarwinkeldifferenz von 28° entspricht. Der Faktor $\frac{1}{2}$ trägt der genaueren Bestimmung des Azimutalwinkels Rechnung.
In den Vorwärtsspurkammern werden Spuren ab etwa 7° gemessen, im Rückwärtsbereich endet das Flüssig–Argon–Kalorimeter ungefähr bei 154°. Daher wird der Polarwinkel des Myons auf $35° < \Theta_\mu < 127°$ eingeschränkt. Dies ist genau der in Abschnitt 5.1.4 bereits angebrachte Schnitt auf das Myon.

Die Aktivität wird folgendermaßen definiert :

$$A := \sqrt{\left(\sum_{\substack{i\,\in\,Spuren\\ mit\ \delta_i<0.5\ rad}} p_{t,i}\right)^2 + \left(\sum_{\substack{i\,\in\,Zellen\\ mit\ \delta_i<0.5\ rad}} E_i \cdot \sin\Theta_i\right)^2} \tag{6.4}$$

Um die Energiedeposition der neutralen Teilchen zu berücksichtigen, werden Teilchen die sowohl im Spurkammersystem als auch im Kalorimeter nachgewiesen werden, in ihrer Energie *doppelt gezählt.* Außerdem wird das Myon nicht in den Summen berücksichtigt.

Die Aktivität wird durch die Anzahl und Energie der Spuren in der Nähe des Myons bestimmt. Nach dem Massenargument erhalten beim Zerfall eines Quarks $Q \to \mu\ X$ die

Zerfallsteilchen mehr Energie für Impulse transversal zur ursprünglichen Quarkrichtung und damit auch zueinander (Abschnitt 4.2.5) — die Aktivität sinkt, da die Teilchen dem Aktivitätskegel „entkommen“ können. Im Falle der Klasse *uds* stammt das Myon aus dem Zerfall leichter Hadronen bzw. wird vorgetäuscht und befindet sich somit im Strom der Zerfallsteilchen, so daß die Aktivität im Kegel um das vorgetäuschte Myon groß ist.
Die Verteilung der Aktivität ist in Abbildung 6.2 für die Unterscheidung der einzelnen Klassen dargestellt.

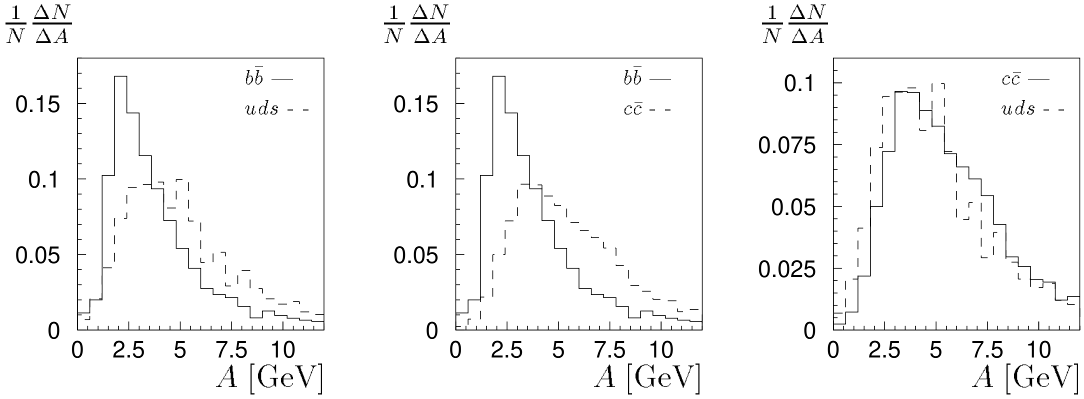

Abbildung 6.2: *Die Aktivität im Vergleich für jeweils zwei Klassen.*

6.3 Der Thrust

Eine weitere topologische Trennungsgröße ist der *Thrust*. Diese Variable beschreibt die Jetartigkeit des Ereignisses und berechnet sich allgemein wie folgt :

$$T := \frac{\max\limits_{\vec{n}_T} \sum\limits_{i \,\in\, Spuren} |\vec{n}_T \cdot \vec{p}_i|}{\sum\limits_{i \,\in\, Spuren} |\vec{p}_i|} \tag{6.5}$$

Die ungerichtete Thrustachse $\vec{n}_T$ ist so definiert, daß die Summe über alle longitudinalen Komponenten der Teilchenimpulse bezüglich dieser Achse maximal wird. Der Thrustwert T ändert sich von 0.5 auf 1 bei einem Übergang von isotropen zu vollständig kollinearen Teilchenverteilungen.

Wegen der Vorwärts–Rückwärts–Asymmetrie des Detektors und der größeren Meßgenauigkeit des Azimutalwinkels wird der Thrust nur noch in der $R\Phi$–Ebene betrachtet.

Der Transversaler Thrust der Hemisphäre

Analog der Thrust–Definition folgt die Definition des *transversalen Thrust*, wobei die Summe nur die Impulskomponenten senkrecht zur Strahlachse berücksichtigt :

$$T_t := \frac{\max\limits_{\vec{n}_{t,T}} \sum\limits_{\substack{i\,\in\,Spuren\\ mit\ \Theta_i>15^\circ}} |\vec{n}_{t,T}\cdot\vec{p}_{t,i}|}{\sum\limits_{\substack{i\,\in\,Spuren\\ mit\ \Theta_i>15^\circ}} |\vec{p}_{t,i}|} \tag{6.6}$$

In der $R\Phi$–Ebene isotrope Teilchenverteilungen besitzen ebenfalls Werte nahe $T_t = 0.5$ und Werte nahe 1 bedeuten vollständig kollineare Verteilungen. Die Einschränkung auf Spuren mit einem Polarwinkel $\Theta_i > 15^\circ$ soll hier und auch für die folgenden topologischen Variablen eine Verschmierung durch den Protonrest verhindern.

Um eine Aussage über die Jetstruktur einer Hemisphäre zu erhalten, wird der transversale Thrust aller in einer Hemisphäre liegenden Spuren $T_{t,H}$ eingeführt.

Das Massenargument ergibt hohe transversale Impulse der Teilchen für die Zerfälle mit schweren Quarks. Der Thrustwert ist demnach für schwere Quarks niedriger als für leichte. Abbildung 6.3 zeigt die Verteilung von T_{t,H_μ} der μ–Hemisphäre unter Mitnahme des Myons im Vergleich der einzelnen Klassen. Dieser Thrust geht als eine Eingabegröße in das Neuronale Netz ein.

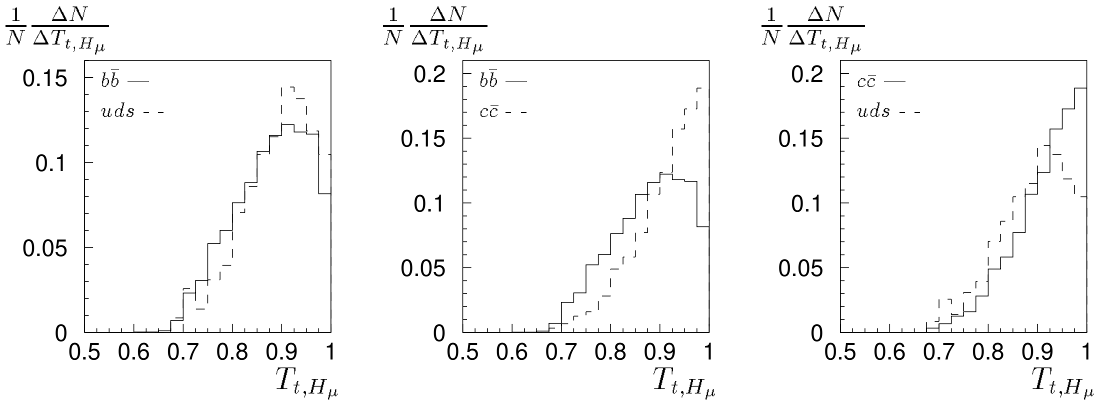

Abbildung 6.3: *Der transversale Thrust der Hemisphäre im Vergleich für jeweils zwei Klassen.*

6.4 Die Aplanarität

Die Struktur eines Ereignisses im Impulsraum läßt sich global durch die Eigenwerte des Impulstensors $\hat{P}_{\alpha\beta}$ $(\alpha, \beta \in x, y, z)$ beschreiben :

$$\hat{P}_{\alpha\beta} = \frac{\sum\limits_{\substack{i\,\in\,Spuren\\ mit\ \Theta_i>15^\circ}} p_{i,\alpha}\cdot p_{i,\beta}}{\sum\limits_{\substack{i\,\in\,Spuren\\ mit\ \Theta_i>15^\circ}} |p_i^2|} \tag{6.7}$$

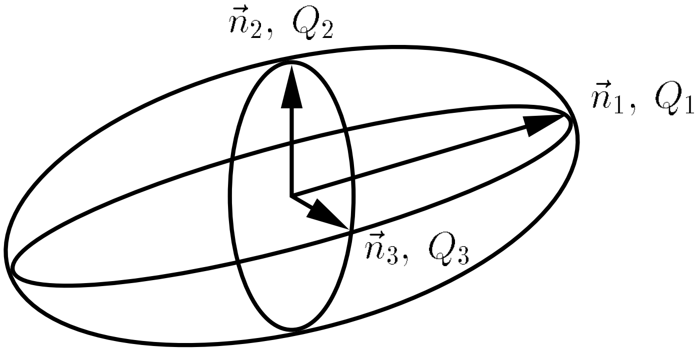

Abbildung 6.4: *Definition des Impulstensors mit seinen drei Hauptachsen. Die Eigenwerte entsprechen den Längen der Halbachsen des Ellipsoids.*

Mit Hilfe einer Hauptachsentransformation und den damit verbundenen Eigenwerten ergeben sich drei weitere topologische Variablen, die die Form des Impulstensors beschreiben und von denen die *Aplanarität* in das Neuronale Netz eingeht.

Sind Q_1, Q_2 und Q_3 mit $Q_1 > Q_2 > Q_3$ und $Q_1 + Q_2 + Q_3 = 1$ die Eigenwerte des Impulstensors und $\vec{n}_3$ der kleinste Eigenvektor (Abbildung 6.4), so berechnet sich die Aplanarität aus

$$Apl = \frac{3}{2} \cdot Q_3. \tag{6.8}$$

Liegen die Impulse aller Teilchen innerhalb einer Ebene, so ist die Aplanarität $Apl \simeq 0$, wogegen $Apl \simeq 0.5$ eine isotrope Verteilung der Teilchenimpulse charakterisiert.

Auch für die Aplanarität greift das Massenargument. Schwerere Quarks bilden weiter ausgedehnte Teilchenverteilungen mit höheren transversalen Impulsen zueinander, so daß die Aplanarität größer wird. Die Unterschiede in den Verteilungen der Aplanarität sind in der Abbildung 6.5 dargestellt.

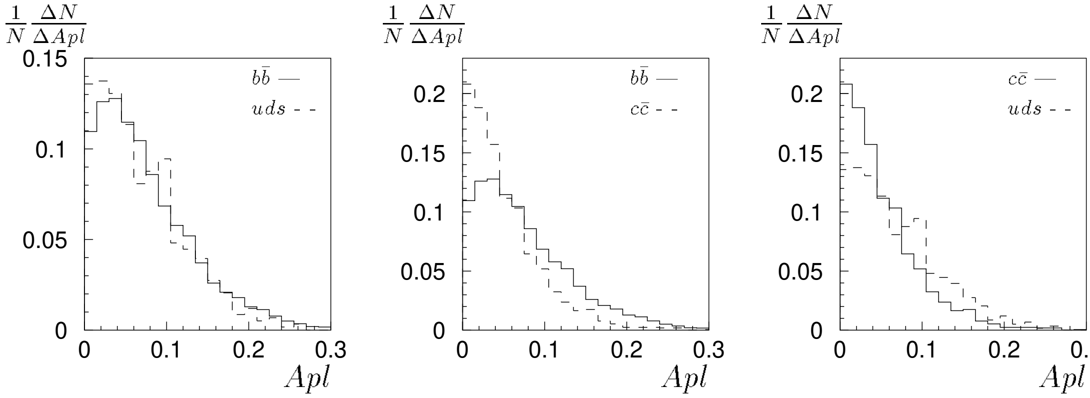

Abbildung 6.5: *Die Aplanarität im Vergleich für jeweils zwei Klassen.*

6.5 Myongrößen

6.5.1 Der Polarwinkel des Myons

Neben den topologischen Größen bestimmt das Myon in den Beauty– und Charm–Klassen und das (vorgetäuschte) Myon der Untergrund–Klasse die Struktur des Ereignisses.

Die Asymmetrie der *ep*–Streuung bewirkt einen Vorwärtsschub[1] aller Teilchen. Leichtere Teilchen werden vom Boost stärker in Vorwärtsrichtung gezwungen als schwere Teilchen. Dadurch erhalten die im harten Prozeß entstandenen Quarks je nach Masse einen unterschiedlichen Vorwärtsschub, der sich über die Fragmentation auch auf den hadronischen Endzustand überträgt.

Desweiteren trägt das in der Boson–Gluon–Fusion der schweren Quarks beteiligte Gluon einen Bruchteil des Protonimpulses in den harten Prozeß. In der Erzeugung leichter Quarks kann sowohl das direkte Photon als auch das aufgelöste Photon mit *jedem* Parton im Proton wechselwirken. Insbesondere tragen die Valenzquarks einen viel höheren Anteil am Protonimpuls als die Gluonen, so daß die resultierenden Reaktionprodukte einen stärkeren Vorwärtsschub erhalten als diejenigen aus der Boson–Gluon–Fusion.

Weiterhin wirkt das Photon im „direkten“ Prozeß dem Protonimpuls stärker entgegen als im Prozeß mit aufgelöstem Photon.

Aus diesen Gründen erfahren die Myonen bzw. vorgetäuschten Myonen der Klasse *uds* einen stärkeren Boost als die Myonen aus den Zerfällen schwerer Quarks (Abbildung 6.6).

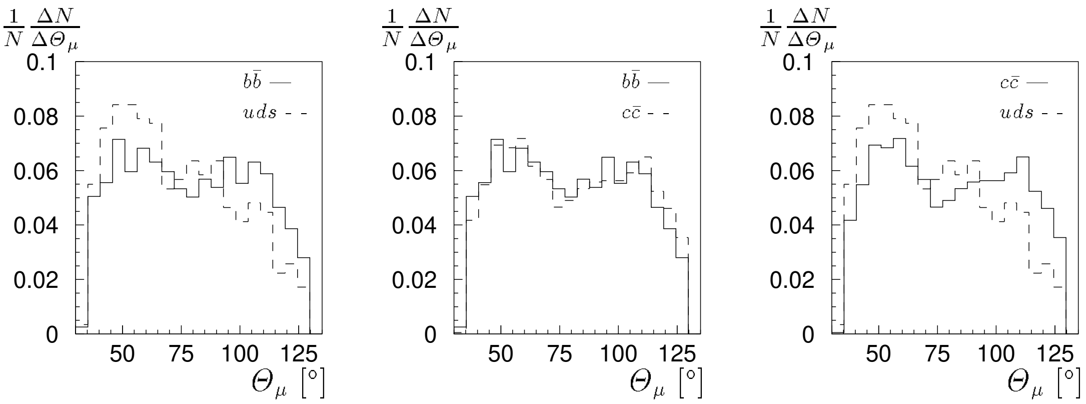

Abbildung 6.6: *Die Polarwinkel des Myons im Vergleich für jeweils zwei Klassen.*

---

[1] Hierfür wird häufig der englische Begriff *Boost* verwendet.

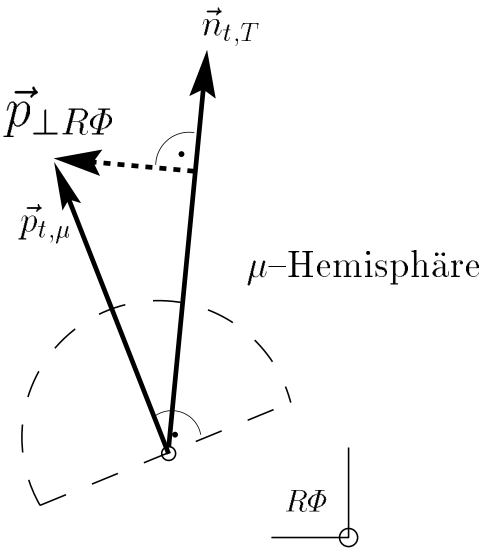

Abbildung 6.7: *Definition des relativen Impulses $p_{\perp R\Phi}$.*

6.5.2 Der Impuls des Myons relativ zur transversalen Thrustachse

Basierend auf der Kinematik des Zerfalls eines schweren Quarks $Q \to \mu\ X$ wird der Thrust, insbesondere wenn das Myon nicht berücksichtigt wird, durch die Impulse der aus X stammenden Teilchen bestimmt. Nach dem Massenargument ist deren transversaler Impuls zur Richtung des Myons umso größer, je größer die Masse des ursprünglichen Quarks ist. Diese Tatsache wird durch die Definition des *relativen Impulses des Myons* zu den restlichen Teilchen, beschrieben durch die Thrustachse der μ–Hemisphäre, berücksichtigt (Abbildung 6.7) :

$$p_{\perp R\Phi} := \sqrt{(\vec{p}_{t,\mu})^2 - (\vec{n}_{t,T} \cdot \vec{p}_{t,\mu})^2} \tag{6.9}$$

Die Thrustberechnung berücksichtigt nur Spuren mit $\Theta > 15°$. Das Myon wird *nicht* mit einbezogen. Abbildung 6.8 zeigt die Verteilungen vergleichend für jeweils zwei Klassen.

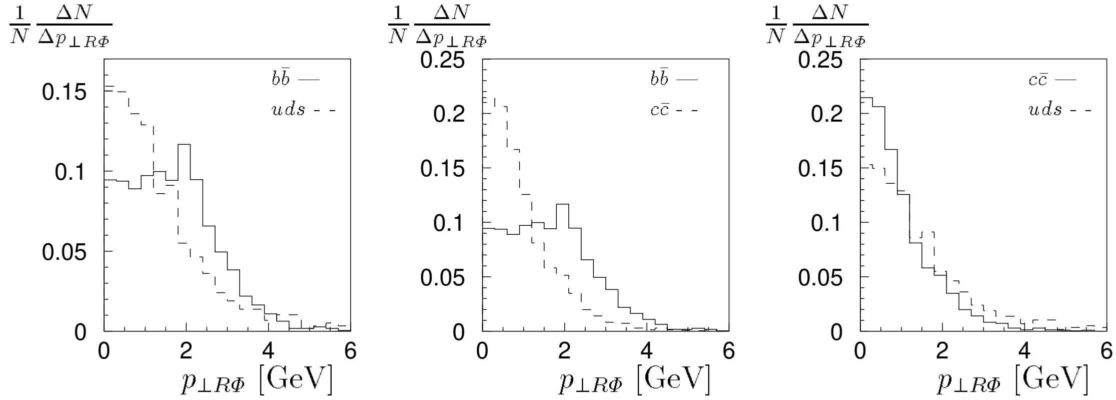

Abbildung 6.8: *$p_{\perp R\Phi}$ im Vergleich für jeweils zwei Klassen. Die transversale Thrustachse $\vec{n}_{\perp,T}$ wurde aus allen Spuren der μ–Hemisphäre berechnet. Das Myon wurde dabei ausgeschlossen.*

6.5.3 Die Anzahl der durchquerten Wechselwirkungslängen im instrumentierten Eisen

Als Kriterium zur Unterscheidung von Myonen und leichten Hadronen, die im instrumentierten Eisen eine Myonsignatur hinterlassen, kann die *Anzahl der durchquerten Wechselwirkungslängen im instrumentierten Eisen* betrachtet werden. Die Hadronen wechselwirken hadronisch mit dem instrumentierten Eisen und werden deshalb schnell absorbiert. Myonen hingegen sind mimimalionisierend und gelangen tiefer in das instrumentierte Eisen (Abbildung 6.9, Unterscheidung von $b\bar{b}$ und uds). Mit den gemessenen Größen Θ_μ und i_{Ende} (Abschnitt 5.1.3), der Dicke der Eisenlagen von 7.5 cm und einer hadronischen Wechselwirkungslänge λ von 131.9 g/cm^2 [PDG96] bei einer Dichte 7.87 g/cm^3 des Eisens ergibt sich

$$WW_{Eisen} = \frac{7.5\,\mathrm{cm} \cdot 7.87\,\mathrm{g/cm}^3}{131.9\,\mathrm{g/cm}^2} \; \frac{i'_{Ende}}{\sin\Theta_\mu}. \tag{6.10}$$

Es ist allerdings zu beachten, daß im instrumentierten Eisen beide Lagen der Doppellage bzw. alle Lagen der Myonenboxen in i_{Ende} berücksichtigt werden (vgl. Abbildung 2.10), aber in die Berechnung von WW_{Eisen} nur einfach (i'_{Ende}) eingehen dürfen.

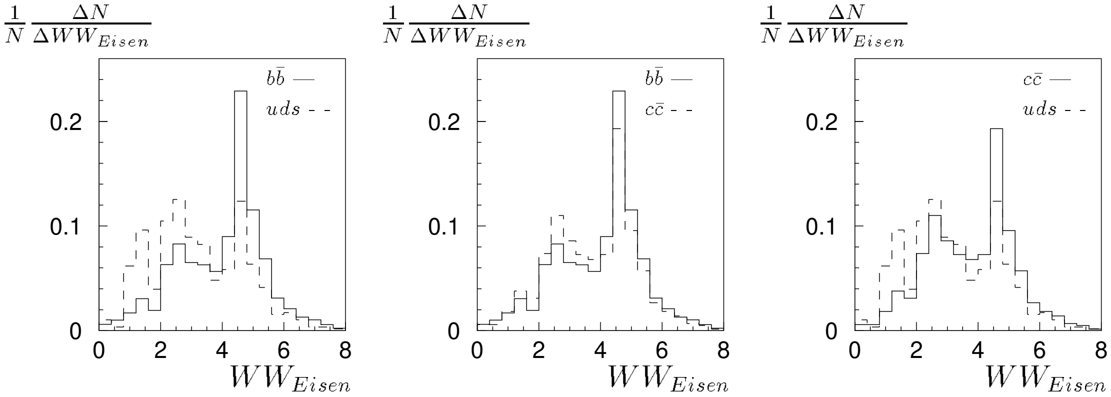

Abbildung 6.9: *Anzahl der durchquerten Wechselwirkungslängen im instrumentierten Eisen im Vergleich für jeweils zwei Klassen.*

6.6 Die Gesamtenergie im Ereignis

Als letzte Variable geht die gesamte Energie des Ereignisses in das Neuronale Netz ein. Sie berechnet sich aus der Energie aller Spuren (*inklusive* des Protonrestes in $\Theta_i < 15°$) unter Annahme der Pionmasse m_π und aller Zellen des Flüssig–Argon–Kalorimeters und des SPACALs :

$$E_{ges} := \sum_{i\,\in\,Spuren} \sqrt{p_i^2 + m_\pi^2} + \sum_{i\,\in\,Zellen} E_i. \tag{6.11}$$

Die Verteilungen sind für die Trennung der Klassen $b\bar{b}$, $c\bar{c}$ und uds in Abbildung 6.10 gezeigt.

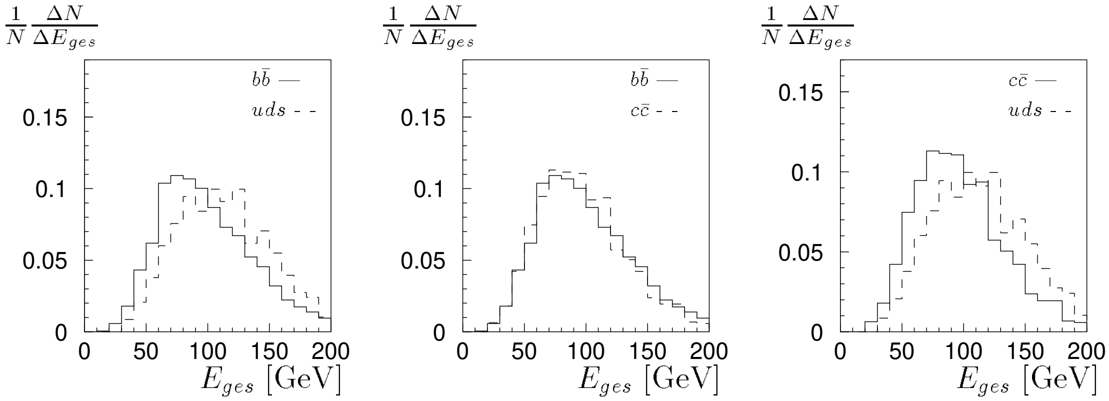

Abbildung 6.10: *Gesamtenergie im Vergleich für jeweils zwei Klassen.*

6.7 Skalierung der Eingabegrößen

Um die Gleichbehandlung aller Eingabegrößen für das Neuronale Netz vorab zu gewährleisten, wird eine *Skalierung* der Eingabegrößen auf Werte im Intervall $[0, 1]$ vorgenommen.

Zu Beginn der Lernprozedur werden die Gewichte mit Zufallszahlen initialisiert (Abschnitt 3.2.3), weshalb Eingabegrößen mit großem Mittelwert anfänglich bevorzugt Einfluß auf die Veränderung der Gewichte ausüben [Moe97]. Die deshalb nötige Eingabeskalierung kann prinzipiell eine zweite versteckte Lage im Neuronalen Netz übernehmen. Aufgrund der geringen Statistik des Untergrundes (Kapitel 5) wird aber darauf verzichtet und per Hand skaliert.

Als Skalierungsfunktion wird die Fermi–Funktion f_F gewählt. Sie bildet den gesamten Wertebereich der Eingabegröße x_k auf das Intervall $[0, 1]$ ab und rückt besonders hohe Werte nahe an die obere Intervallgrenze :

$$f_F(x_k) = \left(1 + \exp\left(-\frac{x_k - \langle x_k \rangle}{\sigma_{x_k}}\right)\right)^{-1} \tag{6.12}$$

Hierbei bezeichnen $\langle x_k \rangle$ den Mittelwert von x_k und σ_{x_k} seine Standardabweichung. Diese Werte werden aus der Verteilung *aller* Klassen bestimmt.

Abbildung 6.11 zeigt die Skalierung für die drei Eingabegrößen Aktivität, transversaler Thrust und Anzahl der durchquerten Wechselwirkungslängen im Eisen für die Unterscheidung der Klassen $b\bar{b}$ und uds (vgl. die unskalierten Verteilungen in den Abbildungen 6.2, 6.3 und 6.9).

6.8 Zusammenfassung aller Eingabegrößen

Abschließend folgt eine Zusammenfassung aller sieben Ereignisvariablen, die zum Trainieren des Neuronalen Netzes verwendet werden, um die drei Klassen $b\bar{b}$, $c\bar{c}$ und uds zu trennen (Tabelle 6.2). Ziel des nächsten Kapitels ist es, mit dem Neuronalen Netz mehrdimensionale Korrelationen der Ereignisvariablen zu finden, die aus den Verteilungen und

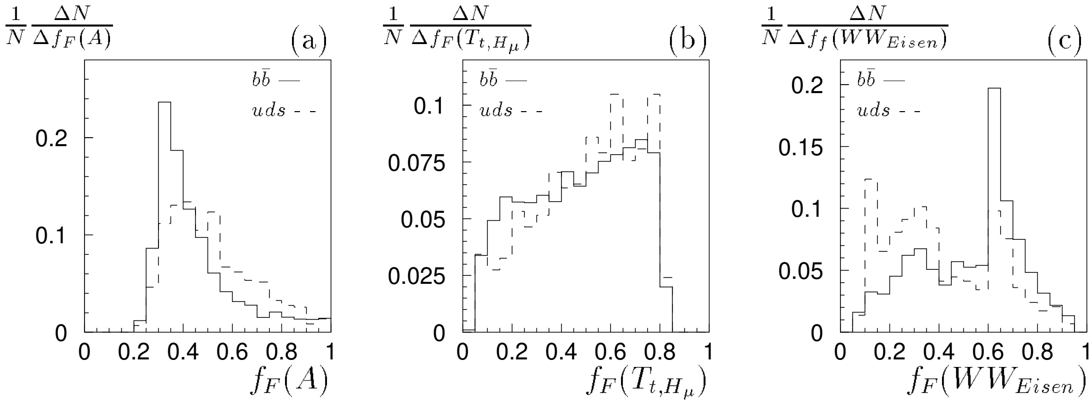

Abbildung 6.11: *Die Verteilung der Aktivität (a), des tranversalen Thrusts (b) und der Anzahl der durchquerten Wechselwirkungslängen im Eisen (c) nach der Skalierung mit der Fermi–Funktion.*

aus der Analyse mit dem empirischen Korrelationskoeffizenten nicht ersichtlich sind. Dadurch soll eine bessere Unterscheidung der Klassen erreicht werden, als es durch Schnitte in den einzelnen Eingabevariablen möglich ist.

Aktivität	A
transversaler Thrust der μ-Hemisphäre	T_{t,H_μ}
Aplanarität des Ereignisses	Apl
Polarwinkel des Myons	Θ_μ
relativer Impuls des Myon bezüglich der transv. Thrustachse (μ-Hemisphäre, Thrustberechnung *exklusive* μ)	$p_{\perp R\Phi}$
Anzahl der durchquerten Wechselwirkungslängen im instr. Eisen	WW_{Eisen}
Gesamtenergie	E_{ges}

Tabelle 6.2: *Zusammenfassung aller Eingabegrößen für das Neuronale Netz.*

Kapitel 7

Anwendung des Neuronalen Netzes

> Vor zwanzig Jahren hattet ihr zwei Prozent.
> Heute habt ihr dreißig Prozent.
> Geht lieber nochmal raus und schaut nach.
> *Stephen W. Hawking über Dunkle Materie*

Nach den Vorbereitungen der Kapitel 5 und 6 wird in diesem Kapitel das Neuronale Netz für die Unterscheidung der Beauty– und Charm–Klassen untereinander, sowie zur Trennung von der Klasse der Untergrund–Ereignisse, durchgeführt.

Abschnitt 1 stellt den Trainingshergang und die Netzstruktur vor. Außerdem werden Kovergenzkriterien definiert, mit denen die Güte der in Abschnitt 2 durchgeführten Trainingsdurchgänge beurteilt werden kann. Dort werden die Resultate für die Trennungen der drei Klassen vorgestellt und die Reproduzierbarkeit der Ergebnisse überprüft. Mit einem Schnitt im Netz–Ausgabewert kann eine Verbesserung der Verhältnisse der Ereignisanzahlen erzielten werden. In Abschnitt 3 wird die Signifikanz der Eingabegrößen untersucht und mit den Ergebnissen aus der Korrelationskoeffizientenanalyse verglichen. Zusammenfassend werden in Abschnitt 4 die Ergebnisse vorgestellt und bewertet.

7.1 Die Trainingsprozedur

Die in Kapitel 5 vorselektierten Ereignisse wurden in drei Klassen eingeteilt von denen nun jeweils zwei Klassen $\mathbb{X}$ und $\mathbb{Y}$, mit $\mathbb{X},\ \mathbb{Y}\ \in\ \left\{b\bar{b},\ c\bar{c},\ uds\right\}$ und $\mathbb{X}\ \neq\ \mathbb{Y}$, verwendet werden, um das Neuronale Netz zu trainieren. Damit ergeben sich drei *Trainingsziele*, die mit „$\mathbb{X}$ vs. $\mathbb{Y}$“ bezeichnet werden (Tabelle 7.1). Für Ereignisse der Klasse $\mathbb{X}$ wird der wirkliche Ausgabewert mit $t \overset{!}{=} 1$ definiert und $t \overset{!}{=} 0$ für die Klasse $\mathbb{Y}$[1]. Es genügt ein Netz mit einem Ausgabeknoten $o \equiv o_1$.

Einen Überblick über die Trainingsprozedur gibt Abbildung 7.1.

---

[1] Damit entspricht $\mathbb{X}$ der „1“–Klasse aus Kapitel 3 und $\mathbb{Y}$ der Klasse „0“ .

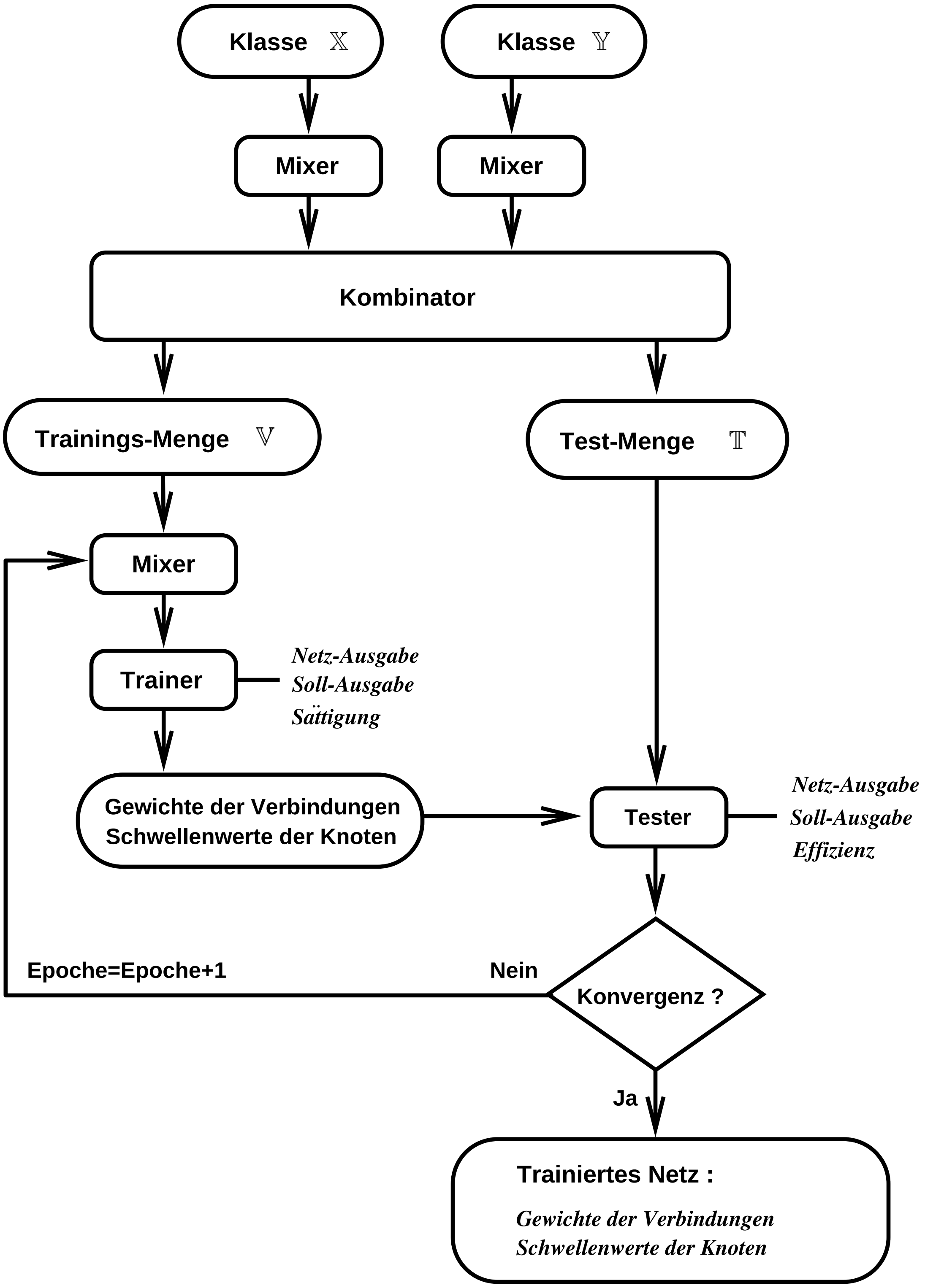

Abbildung 7.1: *Flußdiagramm des Trainings.*

	$\mathbb{X}$ $(t \overset{!}{=} 1)$	$\mathbb{Y}$ $(t \overset{!}{=} 0)$
1.	$b\bar{b}$	uds
2.	$c\bar{c}$	uds
3.	$b\bar{b}$	$c\bar{c}$

Tabelle 7.1: *Die einzelnen Trainingsziele.*

Die Ereignisse beider Klassen werden gemeinsam zu zwei disjunkten Teilmengen, der Menge aller Trainingsmuster $\mathbb{T}$ und der Testmenge $\mathbb{V}$ *kombiniert*. Die Trainingsmenge enthält Ereignisse, mit denen das Neuronale Netz angelernt werden soll. Mit der Testmenge wird der Erfolg des Trainings beurteilt.
Die Aufteilung der Ereignisse geschieht willkürlich, d.h. jedes Ereignis besitzt die gleiche Wahrscheinlichkeit, in eine der beiden Mengen zu kommen, mit der Einschränkung, daß die Anzahl von Mustern aus $\mathbb{X}$ und $\mathbb{Y}$ in $\mathbb{T}$ gleich zu sein hat. Anderenfalls würde eine der Klassen dem Neuronalen Netz häufiger präsentiert werden und damit stärker auf die Gewichte einwirken (im extremen Fall, daß mit einem Ereignis aus $\mathbb{X}$ unter hundert $\mathbb{Y}$–Ereignissen trainiert wird, werden die Gewichte überhaupt nicht auf $\mathbb{X}$ reagieren).

Vor Beginn des Trainings werden die Muster im *Mixer* gemischt, um zu verhindern, daß dem Neuronalen Netz die Muster immer in der gleichen Reihenfolge eingegeben werden und es dadurch Strukturen lernt, die auf der Reihenfolge basieren.
Die Lernprozedur im *Trainer* folgt im wesentlichen der Aufzählung am Ende von Abschnitt 3.2.3. Dem Neuronalen Netz werden disjunkte Teilmengen $\mathbb{M}_n \subset \mathbb{T}$ präsentiert, wobei nach jeder Teilmenge die Gewichte durch Minimierung der Fehlerfunktion angepaßt werden. Wurden alle Muster aus $\mathbb{T}$ trainiert, ist eine *Epoche* beendet. Hat das Neuronale Netz noch keine Konvergenz erreicht, werden die Muster permutiert, die Mengen $\mathbb{M}_n$ neu definiert und eine neue Epoche begonnen.
Information über die Konvergenz des Neuronalen Netzes werden im *Tester* durch Muster in der Testmenge $\mathbb{V}$ erstellt. Mit Hilfe der *Konvergenzkriterien*, den errechneten Ausgabewerten (*Netz–Ausgabewerten*) und den *Soll–Ausgabewerten* findet ein visuelle Beurteilung der Konvergenz statt. Wurde eine Konvergenz erreicht, ergibt sich ein *trainiertes Netz* mit endgültigen Schwellenwerten und Gewichten.
Dieses trainierte Netz wird in der Generalisierungsphase (Abschnitt 3.2.4) auf die Testmenge oder auf unklassifizierte Ereignisse angewendet.

7.1.1 Konvergenzkriterien

Um überprüfen zu können, ob das Neuronale Netz ausreichend gelernt hat, werden nun zwei Kontrollen dargestellt, die Sättigung der Lagen und die Effizienz. Außerdem wird die Größe Reinheit definiert.

Sättigung der Lagen Ein einfaches Maß für die Konvergenz ist die Variation der *Sättigung der einzelnen Lagen* mit zunehmender Epoche.

Für die Sättigung der versteckten Knoten gilt

$$s = \frac{1}{N_{\mathbb{T}}} \sum_{p \,\in\, \mathbb{T}} \sum_{j} \left(1 - 2h_j^{(p)}\right)^2 . \tag{7.1}$$

Hierbei ist h_j der Ausgabewert eines versteckten Knotens j. Die Summe läuft über alle Muster p in der Trainingsmenge, wobei deren Gesamtanzahl $N_{\mathbb{T}}$ beträgt. Werte s nahe eins signalisieren, daß die Ausgaben der Knoten konvergiert haben. Werte nahe null weisen auf ein noch lernendes Netz hin [Lon93].
Analog ergibt sich die Sättigung der Ausgabeknoten, indem h_j durch o ersetzt wird und die Summe über j wegfällt. Hier läßt sich Entwicklung der Sättigung auch am besten verstehen. Die beiden Klassen mit den echten Ausgabewerten $t = 1$ bzw. $t = 0$ sollten gleich oft vorhanden sein, deshalb genügt es, jeweils *ein* Muster jeder Klasse zu betrachten ($N_{\mathbb{T}} = 2$). Für ein perfekt trainiertes Netz sind die Ausgaben $o \simeq 1$ bzw. $o \simeq 0$ und damit

$$s = \frac{1}{N_{\mathbb{T}}} \sum_{p \,\in\, \mathbb{T}} \left(1 - 2o^{(p)}\right)^2 \simeq \frac{1}{2} \left[(1 - 2 \cdot 1)^2 + (1 - 2 \cdot 0)^2\right] = 1. \tag{7.2}$$

Für ein schlecht trainiertes Netz mit $o \simeq 0.5$ für beide Klassen ist

$$s \simeq \frac{1}{2} \left[(1 - 2 \cdot 0.5)^2 + (1 - 2 \cdot 0.5)^2\right] \simeq 0. \tag{7.3}$$

Die Effizienz Die ***Effizienz*** ist eine weitere Größe zur Feststellung der Konvergenz. Sie wird aus den Mustern der Testmenge ermittelt und ist damit, anders als die Sättigung der Lagen, unabhängig vom Training.

Wird angenommen, daß Ereignisse mit $o^{(p)} > o_c$ der Klasse $\mathbb{X}$ zuzuordnen sind und Ereignisse mit $o^{(p)} < o_c$ der Klasse $\mathbb{Y}$, so ist die Wahrscheinlichkeit bei diesem Schnitt ein Ereignis der Klasse $\mathbb{X}$ auch als solches *richtig* identifiziert zu haben

$$P_{\mathit{Eff}}^{\mathbb{X}}(o_c) = \frac{1}{N_{\mathbb{X}}}\, N_{\mathbb{X}}^{(o_c \leq o^{(p)})}. \tag{7.4}$$

Hierbei bezeichnet $N_{\mathbb{X}}^{(o_c \leq o^{(p)})}$ die Anzahl aller Muster p der Testmenge, die gleichzeitig der Klasse $\mathbb{X}$ angehören und einen Ausgabewert größer als o_c haben.
Ein Klasse–$\mathbb{Y}$–Ereignis richtig erkannt zu haben ergibt sich aus

$$P_{\mathit{Eff}}^{\mathbb{Y}}(o_c) = \frac{1}{N_{\mathbb{Y}}}\, N_{\mathbb{Y}}^{(o^{(p)} < o_c)}. \tag{7.5}$$

Das Neuronale Netz ist um so effizienter, je höher die Wahrscheinlichkeit für jede Klasse bei gegebenen Schnitt ist. Um den geeigneten Schnitt o_c zu erhalten, werden beide Wahrscheinlichkeiten in einer *Effizienzkurve* gegeneinander aufgetragen.

Anstelle der Effizienz der Klasse $\mathbb{Y}$ wird häufig der *Unterdrückungsfaktor* definiert :

$$\frac{1}{1 - P_{\mathit{Eff}}^{\mathbb{Y}}(o_c)}$$

Analog zur Sättigung gibt das Verhalten der Effizienz mit wachsender Epoche eine Information über den Fortgang des Lernprozesses. Stagniert die Effizienz $P^{\mathbb{X}}_{\mathit{Eff}}(o_c)$ bei festgelegten Werten $P^{\mathbb{Y}}_{\mathit{Eff}}(o_c)$ (50%, 70%, 80% und 90%), so ist eine Konvergenz des Neuronalen Netzes erreicht und das Training kann beendet werden.

Die Reinheit Als letztes wird die *Reinheit* (**Pur**ity) definiert. Sie gibt den Anteil der *richtig* identifizierten Klasse–$\mathbb{X}$–Ereignisse von allen als $\mathbb{X}$ identifizierten Ereignissen an :

$$P^{\mathbb{X}}_{\mathit{Pur}}(o_c) = \frac{P^{\mathbb{X}}_{\mathit{Eff}}(o_c)}{P^{\mathbb{X}}_{\mathit{Eff}}(o_c) + \left(1 - P^{\mathbb{Y}}_{\mathit{Eff}}(o_c)\right)} \tag{7.6}$$

Der Anteil der *richtig* identifizierten Ereignisse aus Klasse $\mathbb{Y}$ von allen Klasse–$\mathbb{Y}$–Ereignissen ergibt sich analog :

$$P^{\mathbb{Y}}_{\mathit{Pur}}(o_c) = \frac{P^{\mathbb{Y}}_{\mathit{Eff}}(o_c)}{P^{\mathbb{Y}}_{\mathit{Eff}}(o_c) + \left(1 - P^{\mathbb{X}}_{\mathit{Eff}}(o_c)\right)} \tag{7.7}$$

Abbildung 7.2 zeigt die Ausgabewerte aller Muster der Testmenge für ein beispielhaftes Training des Neuronalen Netzwerkes, die Reinheit für beide Klassen, die Effizienz für beide Klassen und die resultierende Effizienzkurve. Die starke Abnahme der Reinheit für Werte gegen eins (Klasse $\mathbb{X}$) bzw. Werte gegen null (Klasse $\mathbb{Y}$) folgt aus der Singularität von Gleichung 7.6 bzw. 7.7 an diesen Grenzwerten. Der beste Schnitt ergibt sich bei $o_c = 0.5$. Damit werden 90.8% aller Klasse–$\mathbb{X}$–Ereignisse richtig identifiziert, während ebenfalls 90.8% der Klasse $\mathbb{Y}$ verworfen werden.

7.1.2 Struktur des Neuronalen Netzes

Das Training wird mit einem Feedforward–Netz mit Gradienten–Lernalgorithmus ausgeführt (Kapitel 3). Es besteht aus drei Lagen, wobei die Eingabelage entsprechend der Anzahl der Ereignis–Variablen sieben Knoten ($N_x = 7$) hat (Kapitel 6), die versteckte Lage aus drei Knoten ($N_h = 3$) besteht und die Ausgabelage einen Knoten ($N_o = 1$) benötigt. Die mathematische Konstruktion benutzt die χ^2–Funktion 3.6 als Fehler und die Sigmoidfunktion (Gleichung 3.2) als Aktivierung. Die Gewichte werden nach je fünf Mustern, $N_{\mathbb{M}} = 5$, neu berechnet und die Konvergenz muß nach maximal 2 500 Epochen erfolgt sein, anderenfalls wird das Training verworfen.

Die in Kapitel 5 vorselektierten Ereignisse (Tabelle 5.9) ergeben die verwendeten Muster. Die Anzahl der Muster in der Trainings– und Testmenge $\mathbb{T}$ und $\mathbb{V}$ sind für jedes Trainingsziel in Tabelle 7.2 angegeben. Die Trainingsziele mit der Klasse *uds* werden wegen der geringen Statistik der Ereignisse aus dieser Klasse auf 450 Muster limitiert. Damit ist der Lernvorgang erschwert und die Aussagekraft der Testmenge stark eingeschränkt.

Als freie Parameter des Netzes bleiben die Lernrate η (Gleichung 3.9 bzw. 3.10) und

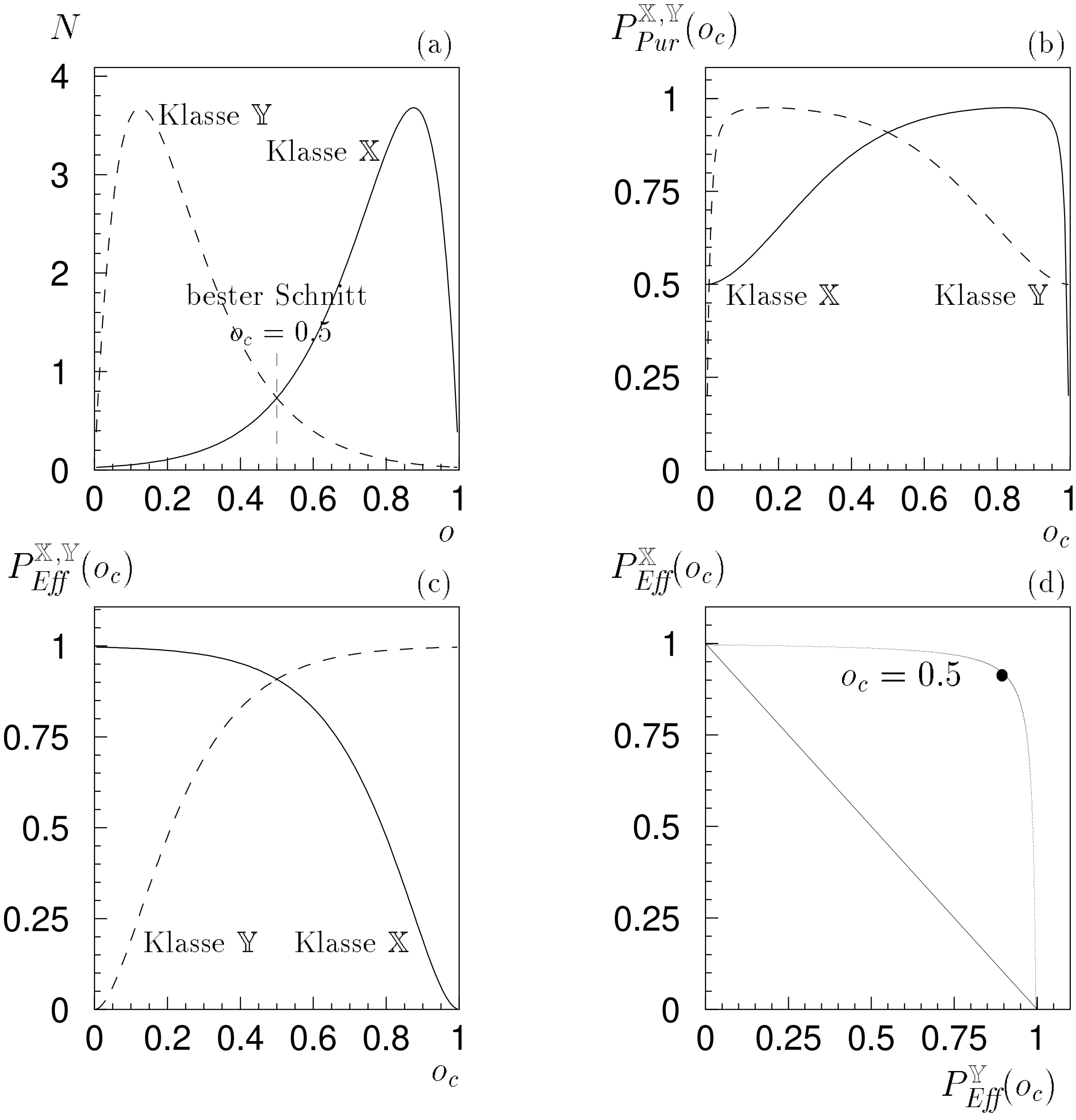

Abbildung 7.2: *Ausgabe eines beispielhaft trainierten Neuronalen Netzes. Abbildung (a): Ausgabewert o des Netzes, (b): Reinheit beider Klassen bei Schnitt in o_c, (c): Effizienz beider Klassen bei Schnitt in o_c und (d): Effizienzkurve.*

die Dämpfung α (Gleichung 3.15 bzw. 3.16). Sie wurden so weit in ihrem Wertebereich variiert, daß sich die maximale Effizienz für die Klasse $\mathbb{X}$ bei festgehaltener Effizienz der Klasse $\mathbb{Y}$ (50%, 70%, 80% und 90%) ergibt.

7.2 Unterscheidung der drei Klassen

Nach Durchführung des Trainings wird die Güte des Neuronalen Netzes in der Generalisierungsphase (Abschnitt 3.2.4) mit Hilfe der Testmenge überprüft. Es zeigt sich, daß die Effizienzen innerhalb der gleich angegebenen Fehlergrenzen für jedes der drei Trainingziele in einem η–Bereich von 10^{-1} bis 10^{-3} nahezu konstant bleiben und dort unabhängig

	Trainingsmenge $\mathbb{T}$	Testmenge $\mathbb{V}$
$b\bar{b}$	450	3 510
uds	450	132
$c\bar{c}$	450	1 612
uds	450	132
$b\bar{b}$	1 500	2 460
$c\bar{c}$	1 500	562

Tabelle 7.2: *Anzahl der Trainings- und Testmuster für die einzelnen Trainingsziele.*

von der Dämpfung α sind. Damit hat sich in allen drei Fällen ein gutes Minimum der Fehlerfunktion finden lassen.
Der jeweilige Trainingsdurchgang eines Trainingszieles, der die besten erreichten Effizienzen ergab, bestimmt die Parameter η_0 und α_0. Mit diesen Parametern werden alle folgenden Trainingsdurchgänge durchgeführt.

Abbildung 7.3 zeigt die Variation der Effizienzen und die Sättigung der Lagen mit zunehmender Epoche für ein typisches Training (Klasse $b\bar{b}$ vs. uds, $\eta_0 = 0.006$, $\alpha_0 = 0.001$). In Abbildung 7.3 a ist zu erkennen, daß die Effizienzen sich mit zunehmender Epoche asymptotisch an feste Werte annähern. Die Sättigung der Lagen strebt mit zunehmender Epoche gegen Werte um $s = 0.15$ (Abbildung 7.3 b). Da es sich hier nicht um perfekt trennbare Klassen handelt, sind Werte nahe eins nicht zugänglich.

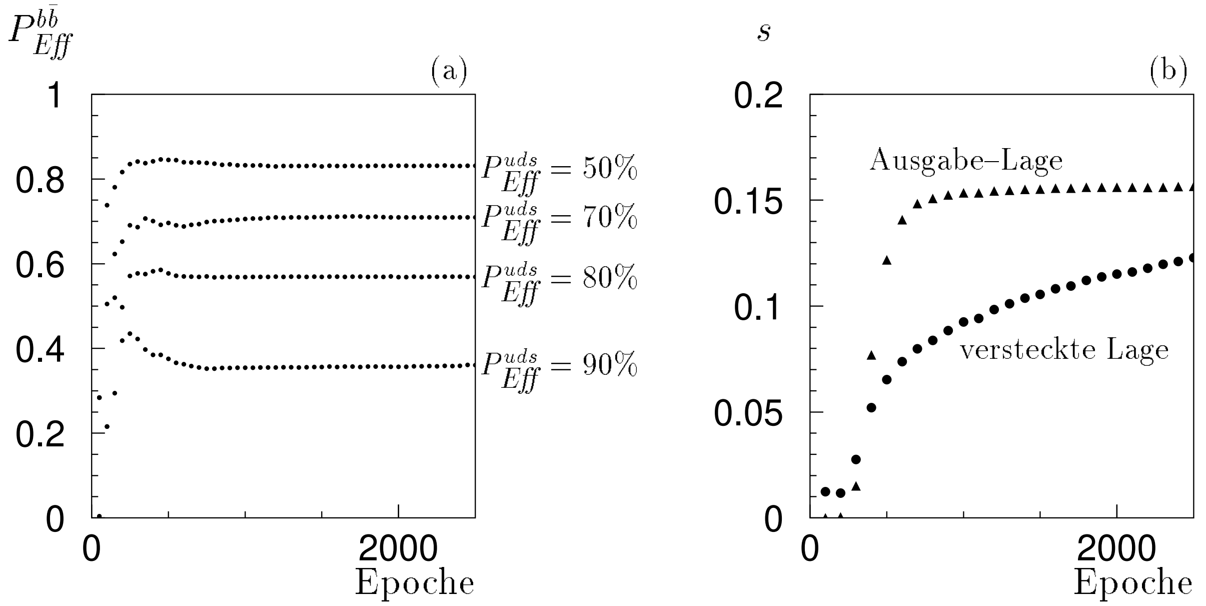

Abbildung 7.3: *Variation der Effizienz (a) und Sättigung der Lagen (b) mit zunehmender Epoche für ein typisches Training (Klasse $b\bar{b}$ vs. uds, $\eta_0 = 0.006$, $\alpha_0 = 0.001$).*

Eine kurze Rechnung zeigt, daß mit $o = 0.5 + \Delta o$ als Ausgabewert für die Klasse $\mathbb{X}$ und $o = 0.5 - \Delta o$ für $\mathbb{Y}$ aus Gleichung 7.1 folgt :

$$s \simeq \frac{1}{2}\left[(1 - 2\cdot(0.5 + \Delta o))^2 + (1 - 2\cdot(0.5 - \Delta o))^2\right] = 4\cdot(\Delta o)^2 \tag{7.8}$$

Mit einer Sättigung bei $s = 0.15$ ergibt sich $\Delta o = 0.19$, so daß die Ausgabe der Klasse $\mathbb{X}$ ihren Mittelwert bei $o = 0.7$ haben sollte bzw. $o = 0.3$ für die Klasse $\mathbb{Y}$. Vorausblickend bestätigen die Abbildungen 7.4 a, 7.5 a und 7.6 a diese einfache Überschlagsrechnung.
Die Überprüfung anhand von Trainingsdurchgängen mit 20 000 Epochen zeigt, daß die Sättigung jeder Lage tatsächlich konvergiert. Die so stärker gesättigte versteckte Lage ergibt aber keine weitere Verbesserung in den Effizienzen. Insgesamt wird somit ein Training mit 2 500 Epochen als ausreichend angesehen.

Um die Reproduzierbarkeit des Netz–Trainings zu untersuchen und damit die Fehlergrenzen des Lernvorganges abschätzen zu können, wurde das Training zwanzigmal mit dem gleichen Lernparameter η_0 und der gleichen Dämpfung α_0 durchgeführt.
Die Ergebnisse für alle drei Trainingsziele werden anhand des Ausgabewertes, der Effizienz und der Reinheit der *jeweiligen Testmengen* in den nächsten Abschnitten präsentiert.

7.2.1 Unterscheidung von $b\bar{b}$ und *uds*

Abbildung 7.4 a zeigt die Ausgabe des Neuronalen Netzes für die Unterscheidung der Klassen $b\bar{b}$ und *uds*. Es konnte eine gute Trennung der Klassen erreicht werden, obwohl die Ausgabewerte des fertig trainierten Netzes die möglichen Intervallgrenzen nicht erreichen. Die Abbildungen 7.4 b und 7.4 c zeigen die Effizienzkurve und die Reinheit in Abhängigkeit vom Schnitt o_c.

Für Effizienzen der Klasse *uds* von 50%, 70%, 80% und 90% (entsprechend den Unterdrückungsfaktoren 2, 3, 5, und 10) sind die Effizienz und die Reinheit der Klasse $b\bar{b}$ in Tabelle 7.3 separat aufgeführt. Außerdem wird der Schnitt im Ausgabewert des Neuronalen Netzes o_c und sein Variationsbereich angegeben.

Mit den trainierten Neuronalen Netzen läßt sich das Verhältnis von $b\bar{b}$ zu *uds* unter den angebrachten Schnitten der Ereignis–Selektion (Kapitel 5) um einen Faktor

$$\frac{P^{b\bar{b}}_{\mathit{Eff}}(o_c)}{1 - P^{uds}_{\mathit{Eff}}(o_c)} = 1.60\ldots 3.13 \tag{7.9}$$

je nach zu erzielender Reinheit der Klasse $b\bar{b}$ verbessern. Bei einer Unterdrückung der Klasse *uds* auf $\frac{1}{10}$ der ursprünglichen Ereignisse verbleiben für die Klasse $b\bar{b}$ 12.9 Ereignisse pro 1 pb^{-1} Luminosität. Dies entspricht einer *gesamten Effizienz* für die Beauty–Klasse $\varepsilon^{(b\bar{b})}_{ges}$ von

$$\varepsilon^{(b\bar{b})}_{ges} = \varepsilon^{(b\bar{b})}_{ES}\cdot P^{b\bar{b}}_{\mathit{Eff}}\Big|_{P^{uds}_{\mathit{Eff}}=90\,\%} = 4.035\cdot 10^{-3}, \tag{7.10}$$

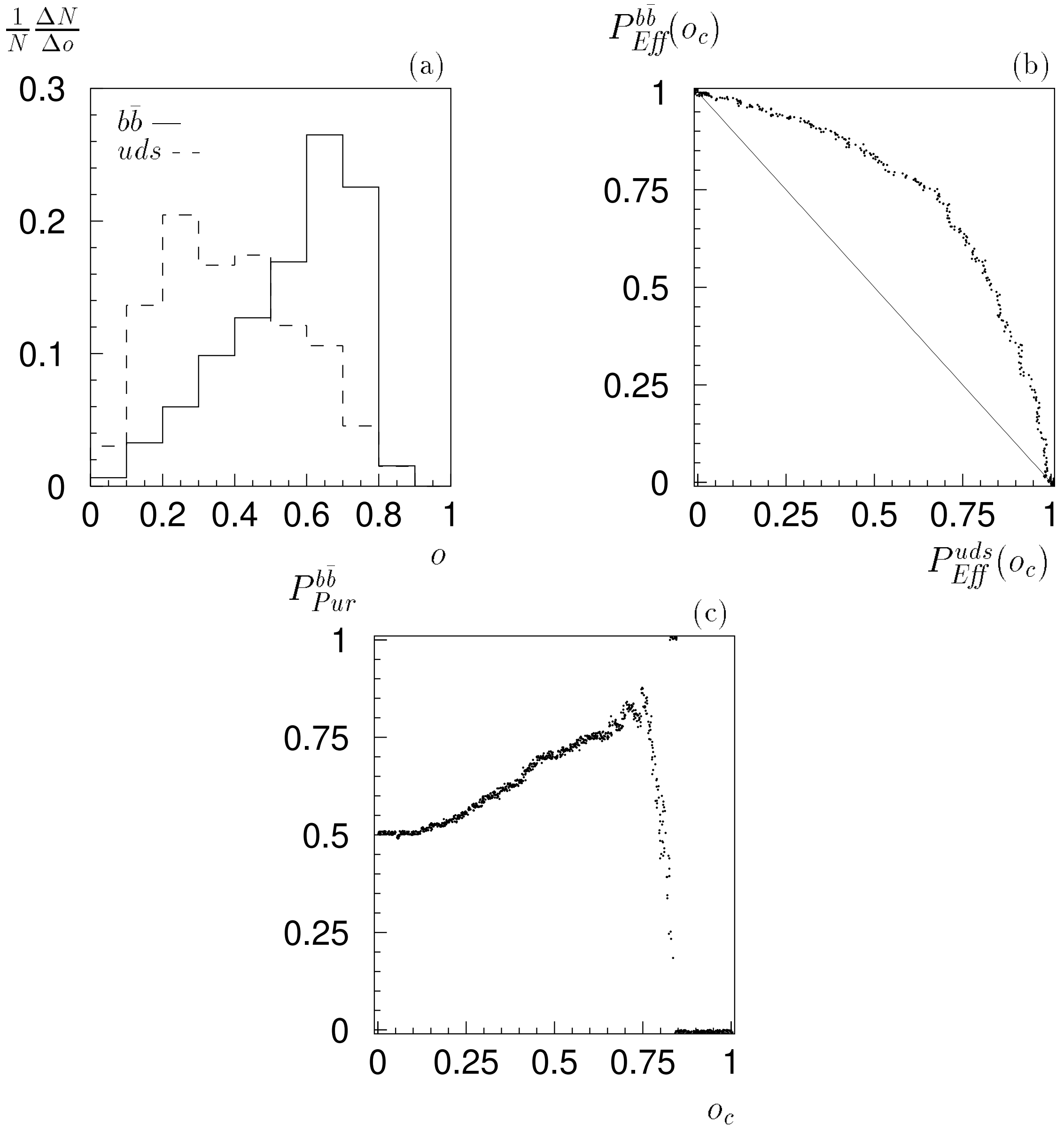

Abbildung 7.4: *Ausgabewert o (a), Effizienzkurve (b) und Reinheit (c) für die Trennung der Klassen $b\bar{b}$ und uds. Dargestellt ist das beste Ergebnis.*

$b\bar{b}$ vs. uds	50%	70%	80%	90%
Effizienz	79.8 ± 2.8	61.1 ± 3.5	48.2 ± 4.2	31.3 ± 4.2
Reinheit	61.5 ± 0.8	67.0 ± 1.3	70.6 ± 1.8	75.5 ± 2.3
Schnitt o_c	0.416 ± 0.023	0.539 ± 0.022	0.606 ± 0.025	0.683 ± 0.026
Bereich von o_c	0.38 ... 0.46	0.48 ... 0.59	0.56 ... 0.66	0.62 ... 0.72

Tabelle 7.3: *Reinheit und Effizienz der Klasse $b\bar{b}$ bei 50%, 70%, 80% und 90% Effizienz der Klasse uds und der entsprechende Schnitt o_c.*

wobei $\varepsilon_{ES}^{(b\bar{b})}$ die Effizienz der Ereignis–Selektion angibt (Kapitel 5). Für die Klasse *uds* ergibt sich $\varepsilon_{ges}^{(uds)} = 1.459 \cdot 10^{-6}$.
Mit 15.5/1 pb^{-1} verbleibenden Ereignissen der Klasse *uds* wird eine Verhältnisverbesserung um den Faktor 3.13 zu Gunsten der Klasse $b\bar{b}$ erreicht. Das auf Luminosität normierte Verhältnis der Ereignisanzahlen beträgt

$$\left.\frac{\mathcal{N}_{b\bar{b}}}{\mathcal{N}_{uds}}\right|_{NN} = 3.13 \cdot \left.\frac{\mathcal{N}_{b\bar{b}}}{\mathcal{N}_{uds}}\right|_{ES} = 0.8341. \tag{7.11}$$

Das Verhältnis $\left.\frac{\mathcal{N}_{b\bar{b}}}{\mathcal{N}_{uds}}\right|_{ES}$ ergibt sich aus dem auf Luminosität normierten Verhältnis der Ereignisanzahlen nach der Ereignis–Selektion.

7.2.2 Unterscheidung von $c\bar{c}$ und *uds*

Analog zum vorigen Abschnitt zeigt Abbildung 7.5 die Ausgabe des Neuronalen Netzes für die Trennung der Klassen $c\bar{c}$ und *uds*. Beide Klassen lassen sich gut trennen. Tabelle 7.4 enthält die dazugehörigen Effizienzen für die Klasse $c\bar{c}$ bei festgelegten Effizienzen der Klasse des Untergrundes. Es sind Verhältnisverbesserungen von 1.62 . . . 3.62 zu erreichen, wobei dies wiederum von der Reinheit der Charm–Klasse bzw. von der gewünschten Unterdrückung der Klasse *uds* abhängt. Für eine Unterdrückung der Klasse *uds* auf $\frac{1}{10}$ verbleiben — auf 1 pb^{-1} Luminosität normiert — 17.4 der Charm–Ereignisse und 15.5 der Untergrund–Ereignisse, entsprechend einer gesamten Effizienz von $\varepsilon_{ges}^{(c\bar{c})} = 2.834 \cdot 10^{-5}$ bzw. $\varepsilon_{ges}^{(uds)} = 1.459 \cdot 10^{-6}$.
Das auf Luminosität normierte Verhältnis der Ereignisanzahlen beträgt

$$\left.\frac{\mathcal{N}^{(c\bar{c})}}{\mathcal{N}^{(uds)}}\right|_{NN} = 3.62 \cdot \left.\frac{\mathcal{N}^{(c\bar{c})}}{\mathcal{N}^{(uds)}}\right|_{ES} = 1.1215. \tag{7.12}$$

7.2.3 Unterscheidung von $b\bar{b}$ und $c\bar{c}$

Als letztes wird das Training zur Trennung der Beauty– und Charm–Klassen betrachtet. Die Ausgabe des Neuronalen Netzes, die Effizienz und die Reinheit sind in Abbildung 7.6 dargestellt. Die ausgesuchten Effizienzen stehen in Tabelle 7.5. Das Verhältnis von $b\bar{b}$ zu $c\bar{c}$ läßt sich um einen Faktor 1.70 . . . 4.44 verbessern.
Mit einem Unterdrückungsfaktor von 10 der Klasse $c\bar{c}$ verbleiben 18.3/1 pb^{-1} Ereignisse der Klasse $b\bar{b}$ und 4.8/1 pb^{-1} der Klasse $c\bar{c}$. Die gesamten Effizienzen betragen $\varepsilon_{ges}^{(b\bar{b})} = 5.723 \cdot 10^{-3}$ bzw. $\varepsilon_{ges}^{(c\bar{c})} = 7.829 \cdot 10^{-6}$ und das auf Luminosität normierte Verhältnis der Beauty– und Charm–Ereignisse ergibt sich zu

$$\left.\frac{\mathcal{N}^{(b\bar{b})}}{\mathcal{N}^{(c\bar{c})}}\right|_{NN} = 4.44 \cdot \left.\frac{\mathcal{N}^{(b\bar{b})}}{\mathcal{N}^{(c\bar{c})}}\right|_{ES} = 3.8192. \tag{7.13}$$

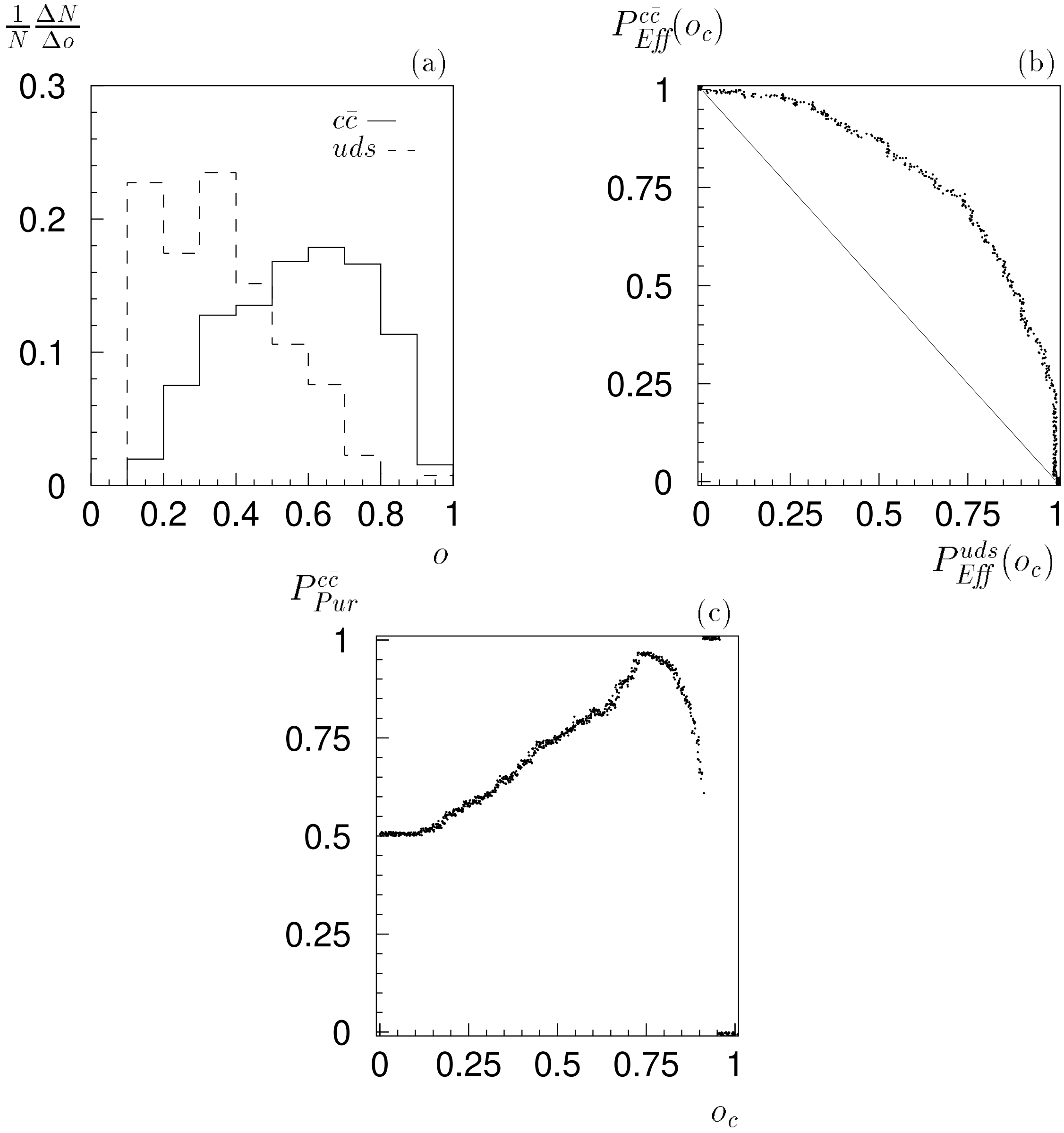

Abbildung 7.5: *Ausgabewert o (a), Effizienzkurve (b) und Reinheit (c) für die Trennung der Klassen $c\bar{c}$ und uds. Dargestellt ist das beste Ergebnis.*

$c\bar{c}$ vs. uds	50%	70%	80%	90%
Effizienz	80.9 ± 2.9	62.9 ± 4.2	51.5 ± 4.7	36.2 ± 4.2
Reinheit	61.8 ± 0.8	67.6 ± 1.4	71.9 ± 1.9	78.2 ± 2.0
Schnitt o_c	0.396 ± 0.032	0.529 ± 0.037	0.599 ± 0.039	0.684 ± 0.039
Bereich von o_c	0.34 … 0.44	0.44 … 0.58	0.52 … 0.66	0.61 … 0.74

Tabelle 7.4: *Reinheit und Effizienz der Klasse $c\bar{c}$ bei 50%, 70%, 80% und 90% Effizienz der Klasse uds und der entsprechende Schnitt o_c.*

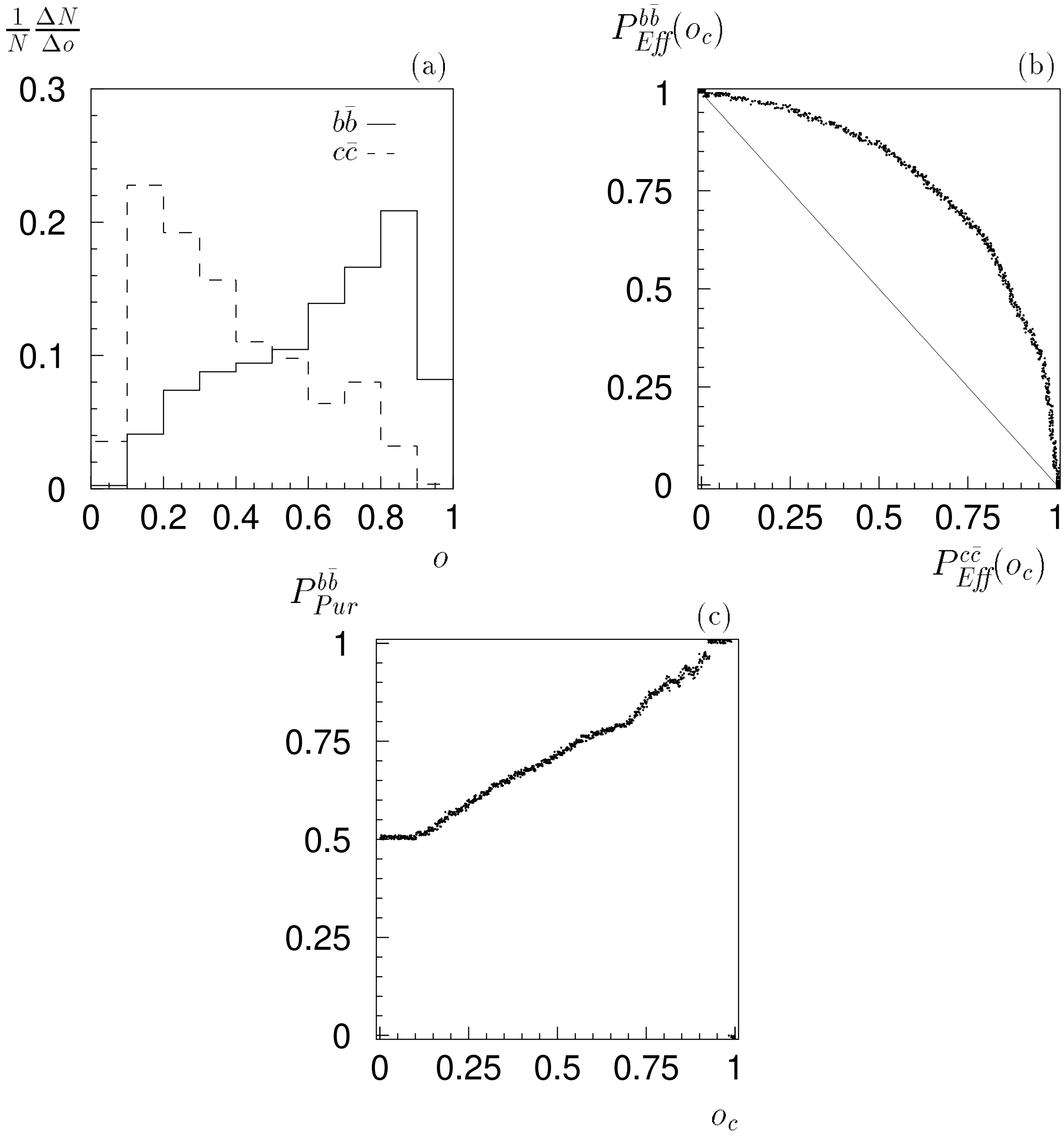

Abbildung 7.6: *Ausgabewert o (a), Effizienzkurve (b) und Reinheit (c) für die Trennung der Klassen $b\bar{b}$ und $c\bar{c}$. Dargestellt ist das beste Ergebnis.*

$b\bar{b}$ vs. $c\bar{c}$	50%	70%	80%	90%
Effizienz	85.1 ± 0.9	71.3 ± 1.4	61.0 ± 2.3	44.4 ± 3.5
Reinheit	63.0 ± 0.2	70.4 ± 0.4	75.3 ± 0.7	81.6 ± 1.2
Schnitt o_c	0.334 ± 0.018	0.474 ± 0.023	0.566 ± 0.028	0.691 ± 0.031
Bereich von o_c	$0.30 \ldots 0.37$	$0.44 \ldots 0.52$	$0.51 \ldots 0.61$	$0.65 \ldots 0.75$

Tabelle 7.5: *Reinheit und Effizienz der Klasse $b\bar{b}$ bei 50%, 70%, 80% und 90% Effizienz der Klasse $c\bar{c}$ und der entsprechende Schnitt o_c.*

Der Vergleich der Fehler der Effizienzen und Reinheiten mit den Fehlern aus den anderen beiden Trainingszielen zeigt, daß die höhere Statistik in der Trainings– und Testmenge (Tabelle 7.2) zu kleineren Fehlern und damit zu einer besseren Reproduzierbarkeit der Ergebnisse führt. Dies bestätigt der stärker eingeschränkte Variationsbereich des Schnittes o_c.

7.3 Signifikanz der Eingabegrößen

Um die *Signifikanz* der Eingabegrößen abschätzen zu können, wurde das Training für alle drei Trainingsziele mit den entsprechenden Parametern η_0 und α_0 erneut durchgeführt, wobei jeweils eine der 7 Eingabevariablen *nicht* mittrainiert wurde. Dies wird im folgenden mit *eingeschränktem Training* bezeichnet. Für die Abschätzung der Reproduzierbarkeit wurde jedes eingeschränkte Training zehnmal wiederholt.

Die erzielten Effizienzen sind, gemittelt über alle Trainingsdurchläufe und unter Angabe der resultierenden Standardabweichungen, in den Tabellen 7.6, 7.7 und 7.8 aufgeführt. Außerdem sind die Werte aus der Analyse mit dem empirischen Korrelationskoeffizienten (Abschnitt 6.1) angegeben.

Bevor eine Aussage über die Signifikanz der einzelnen Variablen gemacht werden kann, muß berücksichtigt werden, daß speziell für die Unterscheidungen von der Untergrund–Klasse ihre geringe Statistik (Tabelle 7.2) zu großen Fehlern in den Effizienzen führt. Somit lassen sich hier höchstens Tendenzen ablesen. Dagegen lassen die Effizienzen für die Unterscheidung der Beauty– und Charm–Klassen eine gute Aussage der Signifikanzen zu. Wegen der höheren Statistik liegt deren Fehler durchschnittlich absolut um 1.8% unter dem Fehler der Trennungen mit der Klasse *uds*.

Die einzelnen Variablen lassen sich grob in vier Signifikanzstufen einordnen :

I. Signifikante Größen :
Mindestens drei Mittelwerte der Effizienzen (bei 50%, 70%, 80% oder 90%) des Trainings aller Variablen und des eingeschränkten Trainings liegen um mehr als die Summe beider Standardabweichungen auseinander.

II. Mäßig signifikante Größen :
Mindestens drei Effizienzen des eingeschränkten Trainings weichen um mehr als eine Standardabweichung von der entsprechenden Effizienz des Trainings mit allen Variablen ab, liegen aber noch innerhalb der Summe beider Standardabweichungen.

III. Schwach signifikante Größen :
Nur zwei Effizienzen erfüllen das Kriterium der mäßigen Signifikanz.

IV. Nicht–signifikante Größen :
Die meisten Mittelwerte stimmen innerhalb der Standardabweichungen überein.

$b\bar{b}$ vs. *uds*	50%	70%	80%	90%	Sign.	$\|r_k\|$
alle Variablen	79.8 ± 2.8	61.1 ± 3.5	48.2 ± 4.2	31.3 ± 4.2		
ohne A	76.6 ± 2.5	58.7 ± 3.8	44.3 ± 6.6	24.6 ± 3.8	III.	0.097
ohne T_{t,H_μ}	80.0 ± 3.1	59.9 ± 4.0	47.4 ± 3.1	28.5 ± 4.0	IV.	0.050
ohne Apl	78.8 ± 3.5	59.2 ± 2.9	45.9 ± 3.3	26.8 ± 5.9	IV.	0.040
ohne Θ_μ	82.1 ± 2.1	60.2 ± 3.2	47.2 ± 3.6	30.8 ± 6.2	IV.	0.074
ohne $p_{\perp R\Phi}$	82.9 ± 2.5	60.0 ± 3.5	45.6 ± 3.1	28.5 ± 3.7	IV.	0.089
ohne WW_{Eisen}	72.0 ± 2.6	51.7 ± 4.2	38.9 ± 2.8	24.7 ± 3.2	I.	0.173
ohne E_{ges}	77.0 ± 2.9	58.2 ± 5.1	44.5 ± 6.5	26.3 ± 5.5	IV.	0.094

Tabelle 7.6: *Effizienz der Klasse $b\bar{b}$ bei 50%, 70%, 80% und 90% Effizienz der Klasse uds und die resultierende Signifikanzstufe, wenn beim Training die angegebene Variable weggelassen wurde. Der obere Abschnitt gibt zum Vergleich den Wert des Trainings mit allen Variablen an. Die rechte Spalte gibt den Wert des empirischen Korrelationskoeffizienten der weggelassenen Variable an.*

$c\bar{c}$ vs. *uds*	50%	70%	80%	90%	Sign.	$\|r_k\|$
alle Variablen	80.9 ± 2.9	62.9 ± 4.2	51.5 ± 4.7	36.2 ± 4.2		
ohne A	80.5 ± 2.4	63.3 ± 3.0	49.4 ± 3.5	33.1 ± 4.0	IV.	0.078
ohne T_{t,H_μ}	80.1 ± 1.4	63.1 ± 2.4	49.0 ± 3.4	33.8 ± 4.3	IV.	0.151
ohne Apl	76.1 ± 2.8	58.9 ± 3.0	46.5 ± 3.4	32.6 ± 5.1	II.	0.160
ohne Θ_μ	75.7 ± 2.9	59.6 ± 4.7	49.2 ± 4.2	31.0 ± 3.7	III.	0.109
ohne $p_{\perp R\Phi}$	81.1 ± 1.9	60.5 ± 4.1	45.9 ± 3.5	30.1 ± 3.5	III.	0.172
ohne WW_{Eisen}	76.7 ± 2.7	58.5 ± 3.4	47.1 ± 4.9	30.2 ± 5.2	II.	0.154
ohne E_{ges}	75.1 ± 3.0	55.6 ± 4.7	46.1 ± 4.6	28.3 ± 4.2	II.	0.177

Tabelle 7.7: *Effizienz der Klasse $c\bar{c}$ bei 50%, 70%, 80% und 90% Effizienz der Klasse uds, analog Tabelle 7.6.*

$b\bar{b}$ vs. $c\bar{c}$	50%	70%	80%	90%	Sign.	$\|r_k\|$
alle Variablen	85.1 ± 0.9	71.3 ± 1.4	61.0 ± 2.3	44.4 ± 3.5		
ohne A	79.5 ± 0.9	64.4 ± 2.4	52.7 ± 2.9	37.4 ± 3.2	I.	0.220
ohne T_{t,H_μ}	84.7 ± 0.9	70.4 ± 1.6	60.3 ± 2.1	43.1 ± 2.6	IV.	0.230
ohne Apl	84.4 ± 1.3	70.9 ± 1.0	59.5 ± 2.2	43.0 ± 2.4	IV.	0.213
ohne Θ_μ	84.5 ± 0.8	70.3 ± 1.4	59.7 ± 2.7	41.7 ± 3.4	IV.	0.021
ohne $p_{\perp R\Phi}$	82.0 ± 1.5	66.1 ± 1.4	55.2 ± 1.9	40.1 ± 2.4	I.	0.310
ohne WW_{Eisen}	84.4 ± 1.0	71.1 ± 2.0	60.9 ± 2.1	44.8 ± 2.5	IV.	0.071
ohne E_{ges}	83.1 ± 0.9	69.7 ± 1.6	58.7 ± 2.1	41.6 ± 1.4	II.	0.042

Tabelle 7.8: *Effizienz der Klasse $b\bar{b}$ bei 50%, 70%, 80% und 90% Effizienz der Klasse $c\bar{c}$, analog Tabelle 7.6.*

In Abbildung 7.7 werden für fünf Variablen die Verteilungen der Effizienzen für das eingeschränkte Training und das Training mit allen Variablen verglichen. Zusammen mit den Tabellen 7.6, 7.7 und 7.8 ergeben sich folgende Ergebnisse :

- Die Aktivität ist für die Trennung von $b\bar{b}$ und $c\bar{c}$ signifikant (Abbildung 7.7 a). Für die Trennung der Klassen $b\bar{b}$ und uds zeigt sich eine schwache Signifikanz.

- Die Anzahl der durchquerten Wechselwirkungslängen im Eisen ist für die Unterscheidung der Klassen $b\bar{b}$ und uds signifikant (Abbildung 7.7 b) und für die Unterscheidung von $c\bar{c}$ und uds mäßig signifikant. Wie erwartet zeigt sich für die Trennung der Beauty– und Charm–Klassen keine Signifikanz, da im Mittel nur die vorgetäuschten Myonen (Klasse uds) im instrumentierten Eisen schnell absorbiert werden (Abschnitt 6.5.3).

- Der relative Impuls ist signifikant für die Trennung von $b\bar{b}$ und $c\bar{c}$ (Abbildung 7.7 c). Schwache Signifikanz zeigt sich für die Unterscheidung von $c\bar{c}$ und uds.

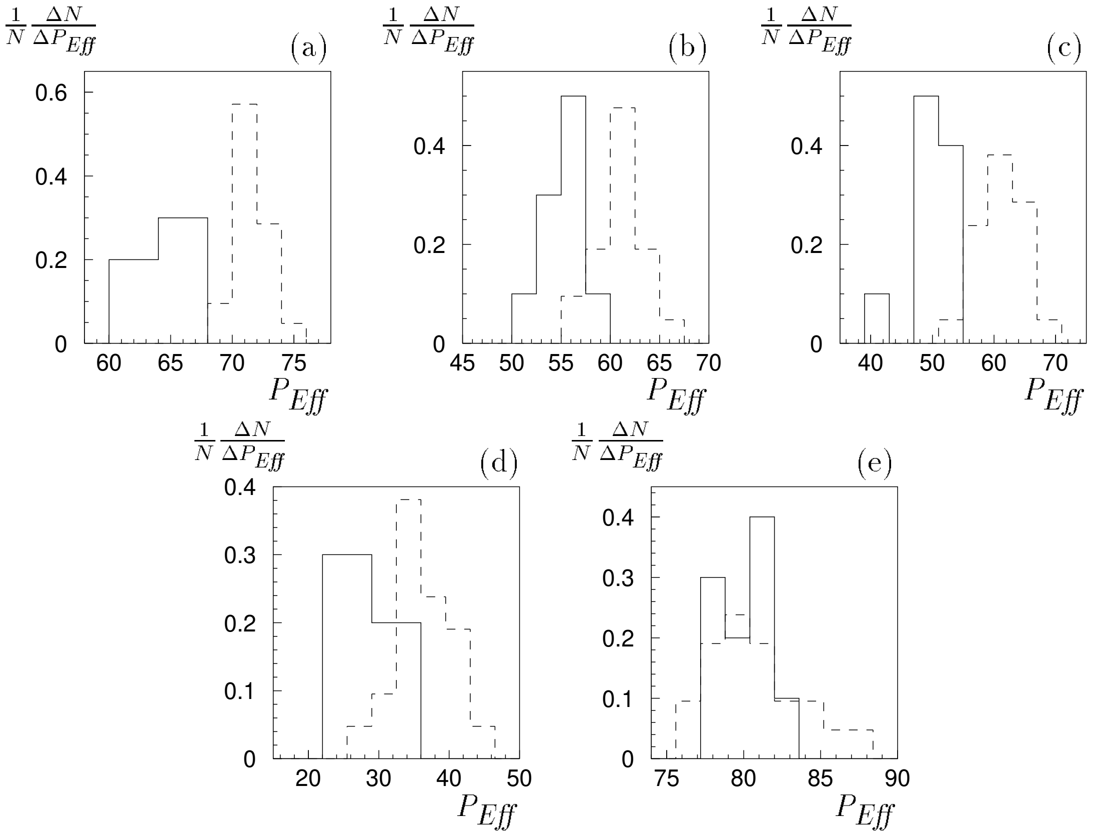

Abbildung 7.7: *Verteilungen der Effizienz für das Training aller Variablen (gestrichelt) und für das eingeschränkte Training (durchgezogen). Die oberen Abbildungen zeigen die Aktivität (a), den relativen Impuls (b) und die Anzahl der durchquerten Wechselwirkungslängen im Eisen (c) als Variablen der Signifikanzstufe I. Abbildung (d) zeigt als Beispiel für die Signifikanzstufen II. und III. die Gesamtenergie und Abbildung (e) den nicht–signifikanten transversalen Thrust (Signifikanzstufe IV).*

- Die Gesamtenergie hat mäßigen Einfluß auf die Trennung der Klassen $c\bar{c}$ und uds (Abbildung 7.7 d) bzw. $b\bar{b}$ und $c\bar{c}$.
- Für die Unterscheidung der Klassen $c\bar{c}$ und uds zeigen die Aplanarität mäßige und der Polarwinkel des Myons schwache Signifikanz.
- Das Training ohne den transversalen Thrust ergibt für kein Trainingsziel eine Verschlechterung der Effizienzen. Der transversale Thrust ist damit als nicht–signifikante Größe einzuordnen (Abbildung 7.7 e).

Die empirischen Korrelationskoeffzienten bestätigen die Signifikanz der Eingabegrößen. Nur für den transversalen Thrust für die Unterscheidung mit der Charm–Klasse und die Aplanarität zur Trennung der Beauty und Charm–Klassen zeigt sich eine Diskrepanz zwischen Signifikanz und Korrelationskoeffizient. Hierbei muß beachtet werden, daß signifikante Variablen große Korrelationskoeffzienten implizieren, der Umkehrschluß jedoch nicht gilt :
Die Aplanarität und der transversaler Thrust basieren zusammen mit der Aktivität auf ähnlichen Informationen aus dem Ereignis (Kapitel 6). Diese Korrelationen untereinander werden vom Neuronalen Netz berücksichtigt. Das Netz kann ein Fehlen einer Eingabegröße durch die anderen beiden, in Abhängigkeit von der redundanten Information, kompensieren. Der empirische Korrelationskoeffzient hingegen beurteilt die Wichtigkeit *einer einzigen* Variablen allein.

7.4 Bewertung und Ausblick

In diesem Abschnitt werden die erzielten Ergebnisse angegeben und diskutiert, welche Verbesserungen möglich und anstrebenswert sind, um mit den in der vorliegenden Analyse verwendeten Methoden eine noch effizientere Trennung der Klassen und damit ein noch besseres auf Luminosität normiertes Verhältnis der Ereignisanzahlen erzielen zu können.

In Tabelle 7.9 ist die auf Luminosität normierte Anzahl der Ereignisse angegeben, die nach der Ereignis–Selektion und nach der anschließenden Anwendung der Neuronalen Netze verbleibt.

	$b\bar{b}$	$c\bar{c}$	uds
theoretischer Wirkungsquerschnitt σ in nb	3.2	612.5	10 609
generierte Luminosität $\mathcal{L}$ in pb^{-1}	$\sim$ 96	$\sim$ 43	$\sim$ 3.76
Ereignisse pro 1 pb^{-1} nach der Ereignis-Selektion	41.3	48.0	154.8
Ereignisse pro 1 pb^{-1} nach Anwendung Neuronaler Netze			
Unterscheidung mit uds	12.9	17.4	15.5
Unterscheidung mit $c\bar{c}$	18.3	4.8	—

Tabelle 7.9: *Zusammenfassung der auf Luminosität normierten Ereignisanzahl.*

In Tabelle 7.10 sind die Verhältnisse der Wirkungsquerschnitte, die bei der Generierung vorausgesetzt wurden, die auf Luminosität normierten Verhältnisse durch die Ereignis–Selektion und die erzielten Verhältnisse der Ereignisanzahlen nach Anwendung des Neuronalen Netzes mit einem Unterdrückungsfaktor von 10 für die jeweilige Klasse $\mathbb{Y}$ aufgestellt. Das auf Luminosität normierte Verhältnis der Ereignisanzahlen nach Anwendung des Neuronalen Netzes wird gegenüber den durch die Ereignis–Selektion erreichten Verhältnissen (Kapitel 5) um etwa einen Faktor 3 verbessert.

	$b\bar{b}$ vs. uds	$c\bar{c}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$
$\frac{\sigma_{\mathbb{X}}}{\sigma_{\mathbb{Y}}}$	0.0003	0.0577	0.0052
$\left.\frac{\mathcal{N}^{(\mathbb{X})}}{\mathcal{N}^{(\mathbb{Y})}}\right\|_{ES}$	0.2665	0.3098	0.8602
$\left.\frac{\mathcal{N}^{(\mathbb{X})}}{\mathcal{N}^{(\mathbb{Y})}}\right\|_{NN}$	0.8341	1.1215	3.8192

Tabelle 7.10: *Verhältnisse der Wirkungsquerschnitte und die auf Luminosität normierte Verhältnisse der Ereignisanzahlen nach der Ereignis–Selektion und nach dem Training für das Trainingsziel* $\mathbb{X}$ *vs.* $\mathbb{Y}$*, bei einem Unterdrückungsfaktor von 10 für die Klasse* $\mathbb{Y}$*.*

Prinzipiell ist es möglich, mit den erzielten Ergebnissen Beauty–Ereignisse in bei H1 aufgenommene Daten folgendermaßen herauszufiltern. In der Generalisierungsphase werden die Neuronalen Netze auf die unklassifizierte Daten in zwei Schritten angewendet. Im ersten Schritt wird eine Unterscheidung in die Klassen uds und $b\bar{b}$ vollzogen (Trainingsziel 1), wobei vorhandene Charm–Ereignisse nicht beachtet werden und somit in beide Klassen gelangen. Im zweiten Schritt wird das Neuronale Netz aus Trainingsziel 3 zur Herausfilterung von Charm–Ereignissen aus der $b\bar{b}$–Klasse benützt.
Problematisch ist ein solches Vorgehen, weil zwei Neuronale Netze angewendet werden und damit bei *jedem* Schritt ein Teil der Statistik der Klasse $b\bar{b}$ verworfen wird. Deshalb sollte das weitere Vorgehen dahinzielen, Trainingsdurchgänge zur Unterscheidung der Klasse $b\bar{b}$ von der gemeinsamen Klasse $uds \cup c\bar{c}$ durchzuführen.
Analog dazu können Charm–Ereignisse von Untergrund und $b\bar{b}$ bzw. die schweren Quarks gemeinsam von Untergrund–Ereignissen getrennt werden.

Um eine solche Analyse auf Daten anwenden zu können, ist jedoch ein eingehendes Verständnis der Ereignisgeneratoren nötig. Zur Zeit ist noch nicht vollständig geklärt, ob sich die Daten durch die untersuchten Prozesse allein beschreiben lassen. Insbesondere standen für den Anteil der Prozesse mit direktem Photon in der Boson–Gluon–Fusion mit leichten Quarks noch keine generierten, simulierten und rekonstruierten Ereignisse zur Verfügung.
Daneben deutet sich an, daß das Verhalten der Myonen in der Detektor–Simulation noch nicht vollständig verstanden ist.

Die Signifikanzüberprüfung hat gezeigt, daß von den in Kapitel 6 ausgewählten Eingabegrößen die Aktivität, die Anzahl der durchquerten Wechselwirkungslängen im Eisen und der relative Impuls des Myons zur transversalen Thrustachse die Entscheidungen des Neuronalen Neztes dominieren. Mit diesen drei Variablen allein sollten die Trainingsdurchgänge erneut durchgeführt werden und die erreichten Effizienzen mit dem Training aller Variablen verglichen werden. Zusätzliche *unabhängige* Variablen, die noch nicht in Kapitel 6 betrachtet wurden, z.B. Informationen über den DCA' (Abschnitt 5.1.3) mit Hilfe des 1997 in Betrieb gegangenen *Silizium–Vertex–Detektors*[2], bzw. Modifikationen an den betrachteten Variablen könnten zu einer besseren Trennung beitragen.

Eine weitere Untersuchung sollte den Lernalgorithmus variieren, um so eine Steigerung der Güte zu erzielen (Kapitel 3).

Um die Reproduzierbarkeit der Neuronalen Netzausgabe zu verbessern und so der Ungenauigkeit in der Interpretation der Signifikanzen entgegenzuwirken, wird insbesonders für die Klasse *uds* eine mehrfach höhere Statistik nötig sein. Alternativ dazu könnte das Training auch mit Daten durchgeführt werden. Dabei sollte aber sichergestellt sein, daß der darin enthaltenen Anteil von Beauty– und Charm–Ereignisse klein genug ist, um das Ergebnis nicht zu verfälschen.

---

[2] Der Silizium–Vertex–Detektor befindet sich am Wechselwirkungspunkt und umgibt dort das Strahlrohr.

Kapitel 8

Zusammenfassung

> Die Naturwissenschaft gleicht einem gigantischen Kreuzworträtsel,
> dessen Zeilen und Spalten schneller wachsen,
> als sich Lösungen und Antworten auf die Fragen ergeben.
> *Oscar Wilde*

In der vorliegenden Arbeit wurde die Möglichkeit untersucht, in *ep*–Wechselwirkung erzeugte schwere Beauty– und Charm–Quarks anhand ihres semimyonischen Zerfalls untereinander und von Untergrund–Ereignissen unter Anwendung Neuronaler Netze zu trennen. Der Untergrund wird von Myonen aus dem Zerfall leichter Hadronen und von als Myon fehlidentifizierten Hadronen (Punch–Through) dominiert.

Die Untersuchung wurde an simulierten Ereignissen im Bereich der Photoproduktion, $Q^2 < 1$ GeV^2, durchgeführt. Es wurden Ereignisse selektiert, in denen ein Myonkandidat im instrumentierten Eisen des H1–Detektors im Polarwinkelbereich $35° < \Theta_\mu < 127°$ und mit einem Impuls von $p_\mu > 2$ GeV nachgewiesen werden konnte.

Zur Trennung benötigen die Neuronalen Netze eindeutig vorgegebene Ereignisse, weshalb eine Einteilung der vorselektierten Ereignisse in drei Klassen, nämlich in Ereignisse mit Beauty–Quarks, solche mit Charm–Quarks und Untergrund–Ereignisse, durchgeführt wurde.

Die Unterscheidung wird dem Neuronalen Netz anhand von Ereignisgrößen ermöglicht. Durch die Analyse mit einem empirischen Korrelationskoeffizienten wurden sieben Ereignisgrößen ermittelt : Aktivität, transversaler Thrust, Aplanarität, Polarwinkel des Myons, Anzahl der vom Teilchen durchquerten Wechselwirkungslängen, relativer Impuls des Myons bezüglich der transversalen Thrustachse und Gesamtenergie.
Ihre Sensibilität auf die Unterscheidung der drei Klassen basiert, bis auf die Anzahl der Wechselwirkungslängen, auf der Massendifferenz zwischen den Beauty–, Charm– und den leichten Quarks. Die Anzahl der durchquerten Wechselwirkungslängen ist für Myonen und für fehlidentifizierte Hadronen unterschiedlich.

Die Unterscheidung je zweier Klassen wurde mit Hilfe von Neuronalen Netzen in zwei

Schritten durchgeführt. Im Lernvorgang wurde für das Neuronale Netz die größtmögliche Korrelation zwischen den sieben Eingabevariablen und den jeweiligen Klassen errechnet. Mit diesem trainierten Neuronalen Netz wurde anschließend die Trennung von unabhängigen Ereignissen aus den zwei Klassen durchgeführt.

Das durch die Ereignis–Vorselektion erreichte relative Verhältnis der Ereignisanzahlen konnte durch das Neuronale Netz für die Unterscheidung von Beauty– und Untergrund–Ereignissen um einen Faktor 3.1 verbessert werden. Für die Unterscheidung von Charm– und Untergrund–Ereignissen ergab sich ein Faktor 3.6 und für die Trennung der Ereignisse mit schweren Quarks untereinander konnte das Verhältnis um den Faktor 4.4 zu Gunsten der b–Ereignisse verbessert werden. Die mit dieser Anreicherung einhergehende Ereignisreduktion beträgt bei einer Unterdrückung des Untergrundes um 90% für Beauty–Ereignisse 69% und für Charm–Ereignisse 64%. Für die Unterscheidung der Ereignisse mit schweren Quarks untereinander ergab sich bei einer Unterdrückung der Charm–Ereignisse von 90% eine Reduktion der Beauty–Ereignisse von 56%.
Ausgehend von den bei der Generierung angenommenen Wirkungsquerschnitten beträgt die auf eine Luminosität von $1\,\mathrm{pb}^{-1}$ normierte verbleibende Ereignisanzahl nach der Trennung vom Untergrund für die Beauty–Klasse 13, für die Charm–Klasse 17 und für den Untergrund selbst 16. Nach der Trennung der Ereignisse mit schweren Quarks untereinander verbleiben 18 Beauty–Ereignisse pro $1\,\mathrm{pb}^{-1}$ und 5 Charm–Ereignisse.
Die erwartete Luminosität beträgt für den Zeitraum Anfang 1995 bis Ende 1997 etwa $30\,\mathrm{pb}^{-1}$, so daß in den Daten, mit der vorliegenden Methode, etwa 500 b–Ereignisse und 500 c–Ereignisse zu erwarten sind.

Indem im Lernvorgang des Neuronalen Netzes jeweils eine der Eingabevariablen ausgelassen wurde, konnte die Signifikanz der sieben Eingabevariablen untersucht werden. Es wurde gezeigt, daß die Aktivität, die Anzahl der durchquerten Wechselwirkungslängen und der relative Impuls des Myons die Neuronalen Netze in ihrer Entscheidung dominieren.

Um die Unterscheidungsfähigkeit des Neuronalen Netzes weiter zu verbessern und um gezieltere Aussagen über die Signifikanz einzelner Eingabegrößen machen zu können, wird speziell für die Untergund–Ereignisse erheblich mehr Statistik als die generierten $3.76\mathrm{pb}^{-1}$ benötigt.
Die Anwendung auf am H1–Detektor aufgenommene Daten erfordert ein eingehenderes Verständnis, in wie weit sich die Daten allein durch die verwendeten Ereignis–Generatoren beschreiben lassen.

Anhang A

Mögliche Eingabegrößen für das Neuronale Netz

In diesem Anhang sind alle 114 Variablen aufgelistet, die für die Trennung der drei Klassen untersucht worden sind. Sie lassen sich in vier Bereiche aufteilen :

- Variablen, basierend auf der Gesamtenergie
- topologische Variablen
- Variablen, basierend auf dem Myon
- relative Impulse aller Spuren bezüglich einer definierten Achse

In den folgenden Abschnitten werden die Variablen, wenn sie nicht selbsterklärend sind, kurz vorgestellt und deren empirischer Korrelationskoeffizienten angegeben. Dabei sind für die Unterscheidung der drei Klassen $b\bar{b}$, $c\bar{c}$ und uds nur die empirischen Korrelationskoeffizienten $|r_k(b\bar{b}$ vs. $uds)|$ und $|r_k(b\bar{b}$ vs. $c\bar{c})|$ maßgebend. Die Werte für $|r_k(c\bar{c}$ vs. $uds)|$ sind der Vollständigkeit halber angegeben.

A.1 Energiegrößen

Die Energiegrößen basieren auf den Impulsen der Spuren des Spurkammersystems und, für den Fall, daß das Kalorimeter mit betrachtet wurde, den Zellen im Flüssig–Argon–Kalorimeter. Spuren und Zellen im Vorwärtsbereich mit $\Theta < 15^\circ$ wurden zur Unterdrückung des Protonrestes ausgeschlossen.

Tabelle A.1 gibt alle acht Energiegrößen und deren Werte des empirischen Korrelationskoeffizienten für die Trennung der Klassen $b\bar{b}$ und uds sowie $b\bar{b}$ und $c\bar{c}$ an. Dabei wird bei den Energien in einer Hemisphäre unterschieden zwischen

- μ–Hemisphäre, Summe über alle Spuren *inklusive* des Myons,

- μ–Hemisphäre, Summe über alle Spuren *exklusive* des Myons und
- gegenüberliegender Hemisphäre.

Die Zuteilung zu einer Hemisphäre wird nach Abschnitt 5.3.1, Gleichung 5.3, vorgenommen. Für die Gesamtenergie des Ereignisses (Variable 5) wurden *alle* Spuren und Zellen, zusätzlich auch die Zellen des SPACALs, mitgerechnet. Die Berechnung der Energie, entnommen aus der `DMIS`–Bank [Bas95] (Variable 6) konnte nicht ausreichend nachvollzogen werden, so daß sie trotz ihres sehr guten Korrelationskoeffizienten disqualifiziert wurde.

A.2 Topologische Größen

Neben den in Kapitel 6 erläuterten Variablen Thrust, transversaler Thrust und Aplanarität kommen die Größen Sphärizität und Planarität hinzu, deren Definition sich direkt aus dem Impulstensor herleiten lassen.

Energiegrößen					
			$\lvert r_k \rvert$		
			$b\bar{b}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$	$c\bar{c}$ vs. uds
1	$\sum_{Spuren} E_i$		0.004	0.017	0.028
2	$\sum_{\substack{Spuren \\ in\ Hemisphäre}} E_i$	inkl. μ	0.002	0.030	0.088
		exkl. μ	0.001	0.015	0.047
		gegenüber	0.004	0.033	0.024
3	$\sum_{Spuren} E_{t,i}$		0.017	0.011	0.015
4	$\sum_{\substack{Spuren \\ in\ Hemisphäre}} E_{t,i}$	inkl. μ	0.019	0.032	0.090
		exkl. μ	0.019	0.004	0.039
		gegenüber	0.057	0.025	0.080
5	$\mathbf{E_{ges} = \sum_{Spuren\ und\ Zellen} E_i}$		**0.094**	**0.042**	**0.177**
6	$E_{\mathtt{DMIS}}$		0.126	0.052	0.207
7	$\sum_{\substack{Spuren\ und\ Zellen \\ in\ Hemisphäre}} E_i$	inkl. μ	0.009	0.020	0.047
		exkl. μ	0.004	0.025	0.050
		gegenüber	0.005	0.032	0.109
8	$\sum_{\substack{Spuren\ und\ Zellen \\ in\ Hemisphäre}} E_{t,i}$	inkl. μ	0.050	0.017	0.060
		exkl. μ	0.052	0.044	0.037
		gegenüber	0.076	0.041	0.120

Tabelle A.1: *Untersuchte Energiegrößen. Fett unterlegt ist die ausgewählte Eingabegröße.*

Mit den Bezeichnungen in Abbildung 6.4 wird die *Planarität* durch

$$Pla = \frac{Q_3}{Q_2} \tag{A.1}$$

definiert. Liegen die Impulse aller Spuren in einer Ebene ($Apl = 0$), so ist auch $Pla = 0$. Eine isotrope Verteilung der Teilchenimpulse entspricht $Pla = 1$.

Die *Sphärizität* bestimmt sich durch

$$Sph = \frac{3}{2} \cdot (1 - Q_1), \tag{A.2}$$

wobei durch $\vec{n}_1$ die Achse mit dem kleinsten „Trägheitsmoment", die *Sphärizitätsachse*, definiert wird. Die Sphärizität gibt ein weiteres Maß über die Spurenverteilung des Ereignisses an. $S = 0$ entspricht einem idealen 2–Jet–Ereignis und $S = 1$ einer isotropen Verteilung der Teilchenimpulse.

Auch hier wird wieder in den Hemisphären unterschieden (Abschnitt A.1). Für die topologischen Verteilungen des *gesamten* Ereignisses wurde das Myon zum einen miteingerechnet (*inklusive* Myon), zum anderen ausgelassen (*exklusive* Myon).

Die empirischen Korrelationskoeffizienten der zehn topologischen Größen sind in Tabelle A.2 aufgestellt.

A.3 Myongrößen

Die Myongrößen sind durchweg selbsterklärend, wobei die Variable 7 eine andere Definition für die Aktivität darstellt. Die relativen Impulsgrößen 8…11, Tabelle A.3, werden analog der Definition des relativen Impulses des Myons aus Abschnitt 6.5.2 hergeleitet, wobei wiederum zwischen den Hemisphären unterschieden wird (Abschnitt A.1) und sich der Hemisphärenvektor aus der Summe der transversalen Impulse aller in der Hemisphäre enthaltenen Spuren berechnet :

$$\vec{n}_H = \frac{\sum\limits_{\substack{Spuren \\ in\ Hemisphäre \\ mit\ \Theta_i > 15^\circ}} \vec{p}_{t,i}}{\sum\limits_{\substack{Spuren \\ in\ Hemisphäre \\ mit\ \Theta_i > 15^\circ}} |\vec{p}_{t,i}|} \tag{A.3}$$

Die Beschränkung auf Spuren mit $\Theta_i > 15°$ schließt den Protonrest aus.

Die Größen sind in Tabelle A.3 aufgelistet. Die Selektion der Aktivität und der Anzahl der druchquerten Wechselwirkungslängen im Eisen als Eingabevariablen für das Neuronale Netz ist trivial. Die Polarwinkelverteilung des Myons unterscheidet die Klassen $b\bar{b}$ und *uds* viel besser als die Impulsgrößen. Bei der Größe des relativen Impulses ist eine rationale Wahl nicht eindeutig. Speziell die kursiv gedruckten Einträge in Tabelle A.3 stellen gute Alternativen zur gewählten Größe dar.

Topologische Größen					
			$\lvert r_k \rvert$		
			$b\bar{b}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$	$c\bar{c}$ vs. uds
1	Thrust des Ereignisses	inkl. μ	0.002	0.166	0.147
		exkl. μ	0.040	0.110	0.144
2	Thrust einer Hemisphäre	inkl. μ	0.034	0.170	0.101
		exkl. μ	0.052	0.097	0.148
		gegenüber	0.025	0.071	0.088
3	transv. Thrust des Ereignisses	inkl. μ	0.018	0.234	0.211
		exkl. μ	0.065	0.191	0.245
4	**transv. Thrust e. Hemisph.**	**inkl.** μ	**0.050**	**0.227**	**0.151**
		exkl. μ	0.051	0.112	0.161
		gegenüber	0.048	0.063	0.106
5	Sphärizität des Ereignisses	inkl. μ	0.018	0.113	0.077
		exkl. μ	0.035	0.083	0.115
6	Sphärizität einer Hemisphäre	inkl. μ	0.055	0.118	0.038
		exkl. μ	0.036	0.098	0.134
		gegenüber	0.017	0.069	0.080
7	Planarität des Ereignisses	inkl. μ	0.033	0.158	0.103
		exkl. μ	0.025	0.142	0.155
8	Planarität einer Hemisphäre	inkl. μ	0.028	0.021	0.052
		exkl. μ	0.017	0.013	0.001
		gegenüber	0.005	0.025	0.005
9	**Aplanarität des Ereignisses**	**inkl.** μ	**0.040**	**0.213**	**0.160**
		exkl. μ	0.042	0.165	0.209
10	Aplanarität einer Hemisphäre	inkl. μ	0.027	0.140	0.107
		exkl. μ	0.049	0.116	0.163
		gegenüber	0.030	0.100	0.147

Tabelle A.2: *Untersuchte topologische Größen. Fett unterlegt sind die ausgewählten Eingabegrößen.*

Myongrößen					
			$\lvert r_k \rvert$		
			$b\bar{b}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$	$c\bar{c}$ vs. uds
1	**Polarwinkel des Myons**		**0.074**	**0.021**	**0.109**
2	Impuls des Myons		0.008	0.120	0.130
3	transv. Impuls des Myons		0.009	0.125	0.130
4	**Anzahl der durchquerten Wechselwirkungslängen im Eisen**		**0.173**	**0.070**	**0.155**
5	Anzahl der durchquerten Wechselwirkungslängen im gesamten Detektor		0.106	0.054	0.087
6	**Aktivität**		**0.097**	**0.220**	**0.078**
7	$\sqrt{\sum_{\substack{\text{Spuren}\\ \text{mit } \delta_i<0.5\ \text{rad}}} (p_{t,i}^2 + m_\pi^2)\sin^2\Theta_i + \sum_{\substack{\text{Zellen}\\ \text{mit } \delta_i<0.5\ \text{rad}}} E_{t,i}^2}$		0.081	0.203	0.094
8	$p_{\perp R\Phi}$ bzgl. Thrustachse	inkl. μ	*0.095*	*0.240*	0.104
		exkl. μ	*0.083*	*0.301*	0.174
		gegenüber	0.048	0.187	0.233
9	**$\mathbf{p_{\perp R\Phi}}$ bzgl. transv. Thrusta.**	inkl. μ	*0.107*	*0.266*	0.118
		exkl. μ	**0.089**	**0.310**	**0.172**
		gegenüber	0.046	0.196	0.124
10	$p_{\perp R\Phi}$ bzgl. Sphärizitätachse	inkl. μ	0.072	0.090	0.017
		exkl. μ	0.027	0.194	0.156
		gegenüber	*0.088*	*0.289*	0.214
11	$p_{\perp R\Phi}$ bzgl. Hemisphärenvektor	inkl. μ	*0.103*	*0.229*	0.081
		exkl. μ	*0.108*	*0.261*	0.103
		gegenüber	0.003	0.182	0.175

Tabelle A.3: *Untersuchte Myongrößen. Fett unterlegt sind die ausgewählten Eingabegrößen. Die kursiv gedruckten Werte geben Alternativen zum ausgewählten relativen Impuls an.*

A.4 Relative Impulse aller Spuren

Abschließend wurden die longitudinalen und transversalen Komponenten der Impulse der Spuren bezüglich der vier Achsen von Thrust, transversaler Thrust, Sphärizität und Hemisphäre betrachtet. Wie schon in den vorigen Abschnitten wurde auch hier die Unterscheidung zwischen μ–Hemisphäre mit und ohne Myon und gegenüberliegender Hemisphäre gemacht (Abschnitt A.1).
Die allgemeine Berechnung für die relativen Impulse (Tabelle A.4) ist wie folgt :

$$p_{\|} := \sum_{\substack{Spuren \\ in\ Hemisph\ddot{a}re \\ mit\ \Theta_i > 15^\circ}} \bar{n} \cdot \vec{p}_i \tag{A.4}$$

$$p_{\perp} := \sum_{\substack{Spuren \\ in\ Hemisph\ddot{a}re \\ mit\ \Theta_i > 15^\circ}} \sqrt{(\vec{p}_i)^2 - (\bar{n} \cdot \vec{p}_i)^2}, \tag{A.5}$$

wobei $\bar{n}$ die jeweilige Achse bezeichnet. Für die Variablen 2 und 7 der Tabelle A.4 ist das Argument der Summe durch seine Quadratur zu ersetzen.

Analog dazu ergeben sich die relativen Transversalimpulse $p_{\| R\Phi}$ bzw. $p_{\perp R\Phi}$ durch Ersetzung von p_i durch $p_{t,i}$ und $\bar{n}$ durch $\bar{n}_t$ (Tabelle A.5).

Wie sich an den empirischen Korrelationskoeffizienten in den Tabellen A.4 und A.5 erkennen läßt, erreicht keine Variable die Unterscheidungsfähigkeiten, die mit den Energiegrößen, topologischen Größen und Myongrößen möglich waren.

relative Impulse aller Spuren					
			$\lvert r_k \rvert$		
			$b\bar{b}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$	$c\bar{c}$ vs. uds
1	$p_{\|\|}$ bzgl. Thrustachse	inkl. μ	0.002	0.002	0.008
		exkl. μ	0.007	0.005	0.011
		gegenüber	0.025	0.021	0.013
2	$p_{\|\|}$ bzgl. Thrustachse, quadriert	inkl. μ	0.007	0.012	0.001
		exkl. μ	0.007	0.012	0.003
		gegenüber	0.034	0.010	0.002
3	$p_{\|\|}$ bzgl. transv. Thrustachse	inkl. μ	0.026	0.008	0.050
		exkl. μ	0.021	0.045	0.019
		gegenüber	0.053	0.062	0.103
4	$p_{\|\|}$ bzgl. Sphärizitätvektor	inkl. μ	0.002	0.014	0.027
		exkl. μ	0.005	0.005	0.005
		gegenüber	0.048	0.008	0.045
5	$p_{\|\|}$ bzgl. Hemisphärenvektors	inkl. μ	0.001	0.013	0.039
		exkl. μ	0.008	0.002	0.012
		gegenüber	0.008	0.049	0.036
6	$p_{\perp}$ bzgl. Thrustachse	inkl. μ	0.002	0.150	0.137
		exkl. μ	0.069	0.060	0.119
		gegenüber	0.019	0.036	0.031
7	$p_{\perp}$ bzgl. Thrustachse, quadriert	inkl. μ	0.032	0.093	0.050
		exkl. μ	0.018	0.008	0.027
		gegenüber	0.018	0.014	0.034
8	$p_{\perp}$ bzgl. transv. Thrustachse	inkl. μ	0.006	0.042	0.089
		exkl. μ	0.002	0.025	0.067
		gegenüber	0.019	0.031	0.001
9	$p_{\perp}$ bzgl. Sphärizitätvektor	inkl. μ	0.001	0.152	0.136
		exkl. μ	0.073	0.049	0.135
		gegenüber	0.013	0.072	0.049
10	$p_{\perp}$ bzgl. Hemisphärenvektor	inkl. μ	0.006	0.136	0.117
		exkl. μ	0.057	0.059	0.127
		gegenüber	0.007	0.047	0.033

Tabelle A.4: *Untersuchte relative Impulse aller Spuren.*

relative Transversal–Impulse aller Spuren					
			$\lvert r_k \rvert$		
			$b\bar{b}$ vs. uds	$b\bar{b}$ vs. $c\bar{c}$	$c\bar{c}$ vs. uds
1	$p_{\|\|R\Phi}$ bzgl. Thrustachse	inkl. μ	0.021	0.008	0.013
		exkl. μ	0.007	0.000	0.012
		gegenüber	0.021	0.016	0.010
2	$p_{\|\|R\Phi}$ bzgl. transv. Thrustachse	inkl. μ	0.026	0.008	0.050
		exkl. μ	0.021	0.045	0.019
		gegenüber	0.053	0.062	0.103
3	$p_{\|\|R\Phi}$ bzgl. Sphärizitätvektors	inkl. μ	0.020	0.011	0.039
		exkl. μ	0.012	0.015	0.003
		gegenüber	0.040	0.002	0.043
4	$p_{\|\|R\Phi}$ bzgl. Hemisphärenvektors	inkl. μ	0.022	0.001	0.052
		exkl. μ	0.016	0.029	0.006
		gegenüber	0.061	0.054	0.103
5	$p_{\perp R\Phi}$ bzgl. Thrustachse	inkl. μ	0.021	0.199	0.157
		exkl. μ	0.062	0.069	0.141
		gegenüber	0.004	0.080	0.066
6	$p_{\perp R\Phi}$ bzgl. transv. Thrustachse	inkl. μ	0.019	0.215	0.175
		exkl. μ	0.020	0.059	0.152
		gegenüber	0.080	0.083	0.049
7	$p_{\perp R\Phi}$ bzgl. Sphärizitätvektor	inkl. μ	0.023	0.179	0.137
		exkl. μ	0.069	0.077	0.155
		gegenüber	0.003	0.088	0.075
8	$p_{\perp R\Phi}$ bzgl. Hemisphärenvektor	inkl. μ	0.026	0.207	0.161
		exkl. μ	0.049	0.098	0.151
		gegenüber	0.017	0.093	0.064

Tabelle A.5: *Untersuchte relative Transversal–Impulse aller Spuren.*

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[Alb88] C. ALBAJAR ET AL.
Study of Heavy Flavour Production in Events with a Muon Accompanied by Jet(s) at the CERN Proton–Antiproton Collider
Z. Phys. **C37** *(1988) 489-503*

[Abt93] I. ABT
Manual to `IJRAY`, a Monte Carlo Generator Shell to Produce *ep* Events for H1
Internal H1–Note **H1–05/93–290**, *1993*

[Bas95] U. BASSLER, G. BERNADI
`H1KINE` : Reconstruction of Kinematic Variables Inside `H1REC` : User's Guide
Internal H1–Software–Note **51–06/95**

[Bro91] I.N. BRONSTEIN, K.A. SEMENDJAJWEW
Taschenbuch der Mathematik
B.G. Teubner Verlagsgesellschaft, Stuttgart 1991, ISBN 3–8154–2000–8

[Cab63] N. CABBIBO
Unitary Symmetry and Leptonic Decays
Phys. Rev. Lett. **10** *(1963) 531*

[Che76] M.S. CHEN, P. ZERWAS
Equivalent–Particle Approximations in Electron and Photon Processes of Higher–Order QED
Phys. Rev. **D12** *(1976) 187*

[Eic88] R.A. EICHLER, Z. KUNSZT ET AL.
Charm, Bottom and Top Production at High Energy Electron Proton Collisions
Nucl. Phys. **B303** *(1988) 791*

[Eli89] R.K. ELLIS, P. NASON
QCD Radiative Corrections to the Photoproduction of Heavy Quarks
Nucl. Phys. **B312** *(1989) 551*

[Els93] E. ELSEN
The H1 Trigger and Data Acquisition System
Internal H1–Note **H1–01/93–262**, *1993,*
Proc. of the Int. Symp. Electronic Instrumentation in Physics, Dubna 1991

[Fri95] S. Frixione, M.L. Mangano, P. Nason, G. Ridolfi
Phys. Rev. **B348** *(1995) 633*

[Glu90] M. Glück, E. Reya, A. Vogt
Z. Phys. **C48** *(1990) 471, Z. Phys.* **C53** *(1992) 127, Phys. Lett.* **306B** *(1993) 391*

[Glu92] M. Glück, E. Reya, A. Vogt
Phys. Rev. **D46** *(1992) 1973, Phys. Rev.* **D45** *(1992) 3986*

[H1C94] H1 Collaboration
The `QJCONE` Jet Algorithm and Its Implementation in `H1PHAN`
H1-Note **H1–08/94–375**, *1994*

[H1C96] H1 Collaboration, S. Aid et al.
Photoproduction of $D^{*\pm}$ Mesons in Electron–Proton Collisions at HERA
Nucl. Phys. **B472** *(1996) 32–54*

[H1C97] H1 Collaboration, I. Abt et al.
The Tracking, Calorimeter and Muon Detectors of the H1 Experiment at HERA
Nucl. Instr. Meth. **A386** *(1997) 348–396*

[Heb49] D.O. Hebb
The Organization of Behaviour
Wiley, New York 1949

[Ing88a] G. Ingelman, G.A. Schuler
Characteristics of Heavy Flavour Production in *ep* Collisions
Z. Phys. **C40** *(1988)*

[Ing88b] G. Ingelman et al.
Deep Inelastic Physics and Simulation
Proceedings of the HERA Workshop Volume 1, Hamburg 1988

[Ing95] G. Ingelman, J. Rathsman, G.A. Schuler
`AROMA 2.1` — A Monte Carlo Generator for Heavy Flavour in *ep* Collisions
DESY-Draft, June 1995

[Ing96] G. Ingelman, J. Rathsman, G.A. Schuler
`AROMA 2.2` — A Monte Carlo Generator for Heavy Flavour in *ep* Collisions
DESY 96-058, April 1996, ISSN 0418-9833

[Kat97] U. Kathage
Photoproduktion von $D^{*\pm}$–Mesonen bei HERA
Dissertation, Universität Kiel 1997

[Kan97] M. Kander
Persönliche Mitteilung, 1997

[Kob73] M. KOBAYASHI, T. MASKAWA
CP Violation in the Renormalizable Theory of Weak Interaction
Prog. Theor. Phys. **49** *(1973) 652*

[Kol57] A.N. KOLMOGOROV
Dokl. Akad. Nauk. 115, USSR 1957

[Kru94] U.P. KRÜGER
Untersuchung der Erzeugung schwerer Quarks durch ihren Zerfall in Myonen im H1–Detektor bei HERA
Dissertation, Universität Hamburg, Interner Bericht DESY F11/F22–94–02, November 1994

[Kun89] Z. KUNSZT
General Properties of Pertubative QCD
Physics For A New Generation, Proc., 28. Internationale Universitätswochen für Kernphysik, Schladming 1989

[Lon93] L. LÖNNBLAD, C. PETERSON, T. RÖGNVALDSSON
JETNET 3.0 — A Versatile Artificial Neural Network Package
LU TP 93-29, CERN–TH. 7135/94, December 1993

[Mar95] A.D. MARTIN, R.G. ROBERTS, W.J. STIRLING
Phys. Lett. **354B** *(1995) 155–162*

[McC43] W.S. MCCULLOCH, W. PITTS
A Logical Calculus of the Ideas Immanent in Nervous Activity
Bull. Math. Biophysics, 5:115–133, 1943

[MeyA95] A. MEYER
Nachweis von Myonen unter kleinem Polarwinkel im H1–Detektor bei HERA
Diplomarbeit, Universität Hamburg 1995

[MeyJ91] J. MEYER
Guide to Simulation Programm `H1SIM`
Hamburg 1991

[Moe97] J. MÖCK
Persönliche Mitteilung, 1997

[Nau94] D. NAUCK, F. KLAWONN, R. KRUSE
Neuronale Netze und Fuzzy–Systeme
Vieweg, Braunschweig 1994, ISBN 3–528–05265–1

[PDG96] PARTICLE DATA GROUP
Review of Particle Properies
Phys. Rev. **D54** *(1996) 1–720*

[Pet89] C. PETERSON, E. HARTMANN
Explorations of the Mean Field Theory Learning Algorithm
Neural Networks **2** *(1989) 475*

[Rog94] T. RÖGNVALDSSON
On Langevin Updating in Multilayer Perceptrons
Lund Preprint LU TP 93–13, 1994

[Roj91] R. ROJAS
Theorie der Neuronalen Netze
Springer Verlag, Berlin 1991

[Ros58] F. ROSENBLATT
The Perceptron : A Probabilistic Model for Information Storage and Organisation in the Brain
Psychological Review, 65:386–408, 1958

[Rum86] D.E. RUMMELHART, G.E. HINTON, R.J. WILLIAMS
Learning Representations by Back–Propagation Errors
Nature, 323:533-536, 1986

[ScA96] A. SCHÖNING
Untersuchung von Prozessen mit virtuellen und reellen $W^{\pm}$–Bosonen am H1–Detektor bei HERA
Dissertation, Universität Hamburg, Interner Bericht DESY F11/F22–96–02, August 1996

[ScG94] G. SCHMIDT
Nachweis von Myonen im H1–Kalorimeter
Diplomarbeit, Universität Hamburg 1994

[ScG96] G. SCHMIDT
Persönliche Mitteilung, 1996

[ScP95] P. SCHMÜSER
Feynman–Graphen und Eichtheorien für Experimentalphysiker
Springer Verlag, Berlin 1995

[Sel95] R. SELL
Untersuchung der J/ψ–Produktion über den Zerfall $J/\psi \to \mu^+\mu^-$ am ep–Speicherring HERA
Dissertation, Universität Hamburg 1995

[Sjo88] T. SJÖSTRAND
Status of Fragmentation Models
Int. J. Mod. Phys. **A3** *(1988) 751–823*

[Sjo93] T. SJÖSTRAND
`PYTHIA 5.7` and `JETSET 7.4` Physics and Manual
CERN–TH. 7112/93, 1993

[Wei34] C.F. von Weizsäcker
Ausstrahlung bei Stößen sehr schneller Elektronen
Z. Phys. **88** *(1934) 612*

[Wes96] L. West
How to Use the Heavy Flavour Working Group Track, Muon and Electron Selection Code Version 1.00/12
November 1996

[Wid60] B. Widrow, M.E. Hoff
Adaptive Switching Circuits
In IRE WESCON Convention Record, New York 1960

[Wil34] E.J. Williams
Nature of the High Energy Particles of Penetrating Radiation and Status of Ionisation and Radiation Formulae
Phys. Rev. **45** *(1934) 729*

Danksagung

Im Zusammenhang mit dem Entstehen dieser Diplomarbeit möchte ich gerne allen Helfern und Beteiligten danken:

Zunächst bedanke ich mich bei Frau Prof. Naroska für das faszinierende Thema und die gute Betreuung während der Diplomarbeit. Herrn Prof. Heinzelmann danke ich für die Übernahme der Zweitkorrektur.

Desweiteren danke ich besonders Martin und Helge für die freundschaftliche Zusammenarbeit während dieses Jahres, für die konstruktiven Kritiken und die gemeinsamen Diskussionen. Erst mit der Zeit lichten sich die Unklarheiten.
Tim und Arnd sei dank für die computertechnischen Antworten, Ihr gebt ein gutes UCO ab.
Martin, Helge, Arnd und Benno danke ich für die kritische Durchsicht und die vielen Verbesserungsvorschläge, auch wenn ich sie nicht immer berücksichtigt habe. Tut mir leid, daß ich so ein linguistischer Frevler bin.
Außerdem bedanke ich mich bei allen F22–Mitgliedern für die angenehme Arbeitsatmosphäre, letztendlich war ich doch noch mit Euch schwimmen...

Dafür, daß Du Dich so lange in Geduld geübt hast, danke ich Dir Tina. Dito Yo !
Abschließend danke ich meinen Eltern für die Unterstützung während meines Studiums, nun bin ich endlich autark.

Ich versichere, diese Arbeit unter Angabe aller wesentlichen Hilfsmittel und Quellen selbstständig angefertigt zu haben.

Hamburg, September 1997