

Interner Bericht
DESY F35D-96-09
April 1996



Identifikation von
neutralen Pionen im ZEUS-Detektor
durch feed forward-neuronale Netzwerke

von

U. Fricke

162

Eigentum der **DESY** Bibliothek
property of **DESY** library
Zugang **22. APR. 1996**
received
Loan period **7** days

DESY behält sich alle Rechte für den Fall der Schutzrechtserteilung und für die wirtschaftliche Verwertung der in diesem Bericht enthaltenen Informationen vor.

DESY reserves all rights for commercial use of information included in this report, especially in case of filing application for or grant of patents.

**"Die Verantwortung für den Inhalt dieses
Internen Berichtes liegt ausschließlich beim Verfasser"**

Identifikation von neutralen Pionen im ZEUS-Detektor durch feed forward-neuronale Netzwerke

Experimentelle physikalische Diplomarbeit
am Fachbereich Physik
der Universität Hamburg

vorgelegt von

Ulrich Fricke

Hamburg 1996

Kurzfassung

In dieser Arbeit wird die Identifikation von π^0 -Mesonen über den Zerfall $\pi^0 \rightarrow \gamma\gamma$ untersucht, bei dem die Signale der beiden Photonen im Kalorimeter des ZEUS-Detektors nicht mehr in zwei getrennte Objekte aufgelöst werden können, sondern zu einem Objekt verschmelzen. Monte Carlo-Simulationen der tiefinelastischen Positron-Proton-Streuung werden zum Training von *feed forward*-neuronalen Netzwerken verwendet, mit denen die π^0 -Identifikation durchgeführt wird. Für das Training werden gleiche Anzahlen von Untergrund- und Signalobjekten verwendet. Unter Berücksichtigung der Informationen des Kalorimeters und des Hadron-Elektron-Separators wird eine Effektivität von 76% bei gleichzeitiger Reinheit von 76% erreicht. Verwendet man nur die Daten des Kalorimeters, so reduziert sich die Effektivität im Energiebereich von 1-9 GeV bei festgehaltener Reinheit von 76% um 34,4%; bei festgehaltener Effektivität von 76% reduziert sich die Reinheit in diesem Energiebereich um 7,7%. Die Anzahl der zur π^0 -Identifikation notwendigen Eingabeinformationen der neuronalen Netzwerke wird auf 12 reduziert. Mit diesen 12 Eingabeinformationen werden die mit den 55 Informationen des Kalorimeters und 22 Informationen des Hadron-Elektron-Separators erreichten Werte der Effektivität und Reinheit reproduziert.

Inhaltsverzeichnis

| | |
|---|-----------|
| Inhaltsverzeichnis | i |
| Abbildungsverzeichnis | iv |
| Tabellenverzeichnis | vi |
| 1 Einleitung | 1 |
| 2 HERA | 4 |
| 2.1 Der Speicherring HERA | 4 |
| 2.2 Tiefunelastische Positron-Proton-Streuung | 5 |
| 3 Der ZEUS-Detektor | 8 |
| 3.1 Die Spurkammern | 10 |
| 3.2 Das Kalorimeter | 10 |
| 3.3 Eisenjoch und Außenkalorimeter | 10 |
| 3.4 Die Myonenkammern | 11 |
| 3.5 Die Vetowand | 11 |
| 3.6 Der Luminositätsdetektor | 11 |
| 4 Das hochauflösende Kalorimeter | 13 |
| 4.1 Aufbau und Funktionsweise des Kalorimeters | 13 |
| 4.2 Entwicklung von Teilchenschauern im Kalorimeter | 17 |
| 4.3 Auswertung der Kalorimeterdaten | 19 |

| | |
|---|-----------|
| 5 Der Hadron-Elektron-Separator | 23 |
| 5.1 Aufbau des Hadron-Elektron-Separators | 23 |
| 5.2 Funktionsweise des HES | 24 |
| 5.3 Auswertung der HES-Dioden | 25 |
| 6 Identifikation von neutralen Pionen | 27 |
| 6.1 Kinematik des Zerfalls $\pi^0 \rightarrow \gamma\gamma$ | 27 |
| 6.2 π^0 -Identifikation durch die invariante Masse | 28 |
| 6.3 Motivation der Nutzung neuronaler Netzwerke | 30 |
| 7 Neuronale Netzwerke | 31 |
| 7.1 Grundlagen neuronaler Netzwerke | 31 |
| 7.2 Aufbau und Funktion neuronaler Netzwerke | 34 |
| 7.3 Feed forward-Netzwerke | 36 |
| 7.4 Training der feed forward-Netzwerke | 38 |
| 7.5 Abgewandeltes feed forward-Netzwerk | 39 |
| 7.6 Der back propagation-Algorithmus | 40 |
| 7.7 Interpretation der Ausgabewerte | 42 |
| 7.8 Grenzen der Klassifikation | 43 |
| 7.9 Auswertung der Ergebnisse des Netzwerkes | 44 |
| 7.10 Kontrolle des Trainings | 45 |
| 7.11 Beispiel eines Klassifikationsproblems | 47 |
| 8 Datenauswahl | 50 |
| 8.1 Monte Carlo-Simulation | 50 |
| 8.2 Vorauswahl der Ereignisse | 50 |
| 8.3 Konstruktion der Eingabeinformationen | 53 |
| 8.4 Konstruktion der Trainings- und Testdatensätze | 55 |
| 9 Identifikation von isolierten π^0-Mesonen | 57 |
| 9.1 Konstruktion der Trainingsdaten | 57 |
| 9.2 Trainingsparameter der neuronalen Netzwerke | 58 |
| 9.3 Trainingsverlauf | 59 |
| 9.4 Klassifikation der Islands | 61 |
| 9.5 Auswirkung der verwendeten Informationen des CAL | 64 |

| | |
|--|-----------|
| 10 Identifikation von nicht-isolierten π^0-Mesonen | 66 |
| 10.1 Konstruktion der Trainingsdaten | 66 |
| 10.2 Trainingsverlauf | 67 |
| 10.3 Klassifikation der Islands | 69 |
| 10.4 Effizienz der π^0 -Identifikation | 72 |
| 10.5 Analyse der Eingabeinformationen | 76 |
| 10.6 Ausblick | 79 |
| 11 Zusammenfassung | 81 |

Abbildungsverzeichnis

| | |
|--|----|
| 2.1 Der Speicherring HERA und seine Vorbeschleuniger. | 4 |
| 2.2 Feynmangraphen niedrigster Ordnung der e^+p -Streuung. | 6 |
| 3.1 Der ZEUS-Detektor im Querschnitt. | 9 |
| 3.2 Der ZEUS-Detektor im Längsschnitt. | 9 |
| 4.1 Gliederung des ZEUS-Kalorimeters. | 14 |
| 4.2 Aufbau eines FCAL-Modules. | 15 |
| 4.3 Aufbau des RCAL. | 16 |
| 4.4 Entwicklung von Teilchenschauern im Kalorimeter. | 18 |
| 4.5 Datenrekonstruktion im Kalorimeter. | 21 |
| 5.1 Position des RHES. | 24 |
| 5.2 Datenrekonstruktion im HES. | 26 |
| 6.1 Minimaler Öffnungswinkel der Photonen des π^0 -Zerfalls. | 28 |
| 6.2 Energiedeposition von π^0 -Zerfällen | 29 |
| 7.1 Funktionsweise einer Nervenzelle. | 31 |
| 7.2 Knoten eines neuronalen Netzwerkes. | 32 |
| 7.3 Transferfunktionen neuronaler Netzwerke. | 33 |
| 7.4 Beispiel eines neuronalen Netzwerkes. | 34 |
| 7.5 Aufbau eines <i>feed forward</i> -Netzwerkes. | 36 |
| 7.6 Geometrische Deutung der Funktionsweise eines Knotens. | 42 |
| 7.7 Bayes'sche Wahrscheinlichkeit. | 43 |
| 7.8 Beispiel eines Klassifikationsproblems | 48 |
| 7.9 Trainingsverlauf im Beispielproblem. | 49 |

| | | |
|-------|---|----|
| 8.1 | Energiespektrum der CAL-Islands nach Anwendung des Boxcuts | 51 |
| 8.2 | Verhältnis von Signal zum Untergrund. | 52 |
| 8.3 | Konstruktion der Kalorimeterinformationen. | 53 |
| 8.4 | Konstruktion der HES-Informationen. | 54 |
| 9.1 | Energiespektrum der isolierten π^0 -Mesonen. | 58 |
| 9.2 | Trainingsverlauf zur Identifikation isolierter π^0 -Mesonen. | 60 |
| 9.3 | Relativer Fehler beim Training zur Identifikation isolierter π^0 -Mesonen | 60 |
| 9.4 | Klassifikation isolierter π^0 -Mesonen. | 62 |
| 9.5 | Effektivität und Reinheit der Identifikation isolierter π^0 -Mesonen. | 62 |
| 9.6 | Energieabhängigkeit der π^0 -Identifikation. | 64 |
| 9.7 | Kontrolle der Konstruktion der CAL-Informationen. | 65 |
| 10.1 | Energiespektrum der nicht-isolierten π^0 -Mesonen. | 67 |
| 10.2 | Trainingsverlauf zur Identifikation nicht-isolierter π^0 -Mesonen. | 68 |
| 10.3 | Relativer Fehler beim Training zur Identifikation nicht-isolierter π^0 -Mesonen | 68 |
| 10.4 | Klassifikation isolierter π^0 -Mesonen. | 70 |
| 10.5 | Analyse der nicht eindeutig identifizierten Signalislands | 71 |
| 10.6 | Effektivität und Reinheit der Identifikation nicht-isolierter π^0 -Mesonen | 72 |
| 10.7 | Approximation der Schnittwerte | 75 |
| 10.8 | Energieabhängigkeit der π^0 -Identifikation. | 75 |
| 10.9 | Hauptachsentransformation der Eingabedaten. | 78 |
| 10.10 | Mittelwert von Effektivität und Reinheit in Abhängigkeit der verwendeten Eingabeinformationen | 78 |

Tabellenverzeichnis

| | | |
|------|---|----|
| 2.1 | Wichtige HERA-Parameter | 5 |
| 4.1 | Winkelbereiche der verschiedenen Kalorimeterteile. | 13 |
| 7.1 | Bezeichnungen zur Beschreibung eines <i>feed forward</i> -Netzwerkes. | 37 |
| 7.2 | Trainingsparameter des abgewandelten <i>feed forward</i> -Netzwerkes. | 40 |
| 8.1 | Vorauswahlkriterien. | 51 |
| 8.2 | Eingabeinformationen der neuronalen Netzwerke. | 55 |
| 9.1 | Konfiguration der verwendeten <i>feed forward</i> -Netzwerke. | 59 |
| 10.1 | Effektivität und Reinheit der π^0 -Identifikation durch neuronale Netzwerke | 73 |

Kapitel 1

Einleitung

Viele der im ZEUS-Detektor durch Positron-Proton-Streuung produzierten Hadronen können nur über ihre Zerfallsprodukte identifiziert werden, da sie aufgrund ihrer kurzen Lebensdauer noch in der Nähe des Wechselwirkungspunktes zerfallen. Aus den mit dem Detektor gemessenen Daten werden die einzelnen Zerfallsprodukte identifiziert. Mit den ermittelten Informationen dieser Zerfallsprodukte können Aussagen über die Identität der zerfallenen Teilchen gemacht werden. Pionen (π^0 , π^+ , π^-) treten bei vielen Prozessen als Zerfallsprodukte von Hadronen auf, z.B. bei dem Zerfall $\omega \rightarrow \pi^+\pi^-\pi^0$ [1]. Die Identifikation der Pionen im Detektor ist daher zum Nachweis vieler Reaktionen und der Berechnung ihrer Wirkungsquerschnitte unerlässlich. Viele Zerfallskanäle von Teilchen mit schweren Quarks konnten bisher im ZEUS-Experiment nicht zu ihrer Identifikation verwendet werden, da die Identifikation der π^0 -Mesonen eine zu starke Unsicherheit und damit einen zu hohen Untergrund aufwies.

Die Lebensdauer der neutralen π^0 -Mesonen beträgt $8 \cdot 10^{-17}$ Sekunden [2] und ist damit um sieben Größenordnungen kleiner als die der geladenen π^+ - und π^- -Mesonen. π^0 -Mesonen werden daher nicht direkt, sondern nur über ihre Zerfallsprodukte nachgewiesen. Sie zerfallen zu 98.8% in zwei Photonen [2], die ihre Energie im segmentierten Uran-Kalorimeter (CAL) des Detektors deponieren. Durch die geringe Segmentierung des Kalorimeters ist es bei vielen π^0 -Zerfällen nicht möglich, die Signale der beiden Photonen in zwei getrennte Objekte aufzulösen. Beide Signale werden als ein Objekt gemessen. Durch die bessere Granularität des im Kalorimeter integrierten Hadron-Elektron-Separators (HES) werden viele dieser Photonenpaare dort als zwei getrennte Objekte nachgewiesen. Zur Identifikation dieser π^0 -Zerfälle muß für ein gegebenes Objekt im Kalorimeter und Hadron-Elektron-Separator entschieden werden, ob dieses von den beiden Photonen eines π^0 -Zerfalls stammt. Der Untergrund besteht dabei aus einzelnen elektromagnetisch aufschauерnden Teilchen (Photonen, Elektronen, Positronen) sowie aus Hadronen. Zur Identifikation der π^0 -Mesonen wird die invariante Masse verwendet. Sind im HES zwei Signale vorhanden, so wird die invariante Masse aus der Position dieser Signale und der Energiedeposition im Kalorimeter berechnet. Anschließend wird die invariante Masse mit der π^0 -Masse von $135 \text{ MeV}/c^2$ [2] verglichen. Ein Problem dieses Verfahrens ist, daß nur die Gesamtenergie eines Objektes im Kalorimeter bekannt ist. Zur Berechnung der invarianten Masse müssen Annahmen über die Anteile der beiden Photonen an der Energiedeposition im betrachteten Objekt gemacht werden. Aus diesen

Annahmen ergeben sich systematische Fehler [3]. Das Verfahren kann nicht mehr angewendet werden, wenn die beiden Photonen auch im HES nicht mehr getrennt werden können. Die Form der Energieverteilung im CAL und HES geht nur über die Energie- und Ortsrekonstruktion in die Berechnung der invarianten Masse und damit in die π^0 -Identifikation ein. Daher werden unter Umständen nicht alle vorhandenen Informationen zur π^0 -Identifikation verwendet.

Neuronale Netzwerke bieten die Möglichkeit, bei der Lösung eines Problems wie der beschriebenen π^0 -Identifikation auch bisher ungenutzte Zusammenhänge der verfügbaren Informationen zu berücksichtigen. Dazu berechnen sie aus den ihnen zugänglichen Informationen je nach ihrer Struktur einen oder mehrere Ausgabewerte, die dann zur π^0 -Identifikation genutzt werden. Die Berechnung dieser Ausgabewerte wird durch die Parameter des Netzwerkes bestimmt. Die zur Lösung eines Problems optimalen Parameter werden iterativ in der sogenannten Trainingsphase ermittelt. Obwohl neuronale Netzwerke keine analytischen Lösungen liefern, sind ihre Resultate häufig eine gute Näherung an die optimale Lösung. Sie können entweder elektronisch in Form von Chips oder als Computersimulation eingesetzt werden. Der Vorteil der elektronischen Variante ist die schnelle Auswertung der eingegebenen Daten, so daß sie bei Problemen mit hohen Ereignisraten, die eine schnelle Verarbeitung der einzelnen Daten benötigen, besser geeignet sind als die simulierten Netze. Simulierte neuronale Netzwerke können dagegen leichter an Veränderungen der Datenstruktur angepaßt werden. Da für ein gegebenes Problem Aufbau und Größe der benötigten neuronalen Netzwerke nicht analytisch ermittelbar sind, muß jedes neuronale Netzwerk zunächst als Computersimulation entwickelt werden. In der Hochenergiephysik werden neuronale Netzwerke z.B. zur Identifikation einzelner Teilchen oder Endzustände und zur Spurenrekonstruktion verwendet. Eine Übersicht von verschiedenen Anwendungen neuronaler Netzwerke in der Hochenergiephysik ist unter [4] gegeben.

In dieser Arbeit werden die Möglichkeiten und Grenzen der Identifikation der oben beschriebenen π^0 -Zerfälle mit neuronalen Netzwerken untersucht. Dazu werden die einzelnen Kalorimeterzellen mit Energiedeposition durch Rekonstruktionsalgorithmen zu mehreren Objekten zusammengefaßt. *Feed forward*-Netzwerke verschiedener Architektur werden mit simulierten Ereignissen der tiefinelastischen Positron-Proton-Streuung zur Identifikation der Objekte verwendet, in denen Energie von beiden Photonen eines π^0 -Zerfalls deponiert wurde. Die Informationen der Netzwerke wurden aus den Daten des Kalorimeters und des Hadron-Elektron-Separators berechnet. In der Trainingsphase werden die neuronalen Netzwerke mit simulierten Ereignissen auf die π^0 -Identifikation trainiert. Die Generalisierung, d.h. die Fähigkeit, mit den so bestimmten Kriterien nicht zum Training verwendete Objekte richtig zu klassifizieren, wird an unabhängigen simulierten Ereignissen getestet.

Im ersten Teil der Analyse wird ein Spezialfall der π^0 -Identifikation untersucht. Alle Objekte im Kalorimeter, die Energie von den beiden Photonen eines π^0 -Zerfalls und zusätzlich Energie von anderen Teilchen enthalten, werden von der Analyse ausgeschlossen. Die neuronalen Netzwerke werden zur Identifikation der Objekte verwendet, deren Energie von den beiden Photonen mindestens eines π^0 -Zerfalls deponiert worden ist. Diese Objekte werden im folgenden als isolierte π^0 -Mesonen bezeichnet. Dieses stark vereinfachte Problem der π^0 -Identifikation wird dazu genutzt, die den Trainingsverlauf bestimmenden Größen der verwendeten neuronalen Netzwerke zu ermitteln. Außerdem werden Aussagen über die Effizienz der π^0 -Identifikation gemacht, wenn neben Informationen des Kalori-

meters auch Informationen des Hadron-Elektron-Separators verwendet werden. Im zweiten Teil der Analyse wird der allgemeine Fall der π^0 -Identifikation behandelt. Auch diejenigen Objekte, die Energie von den beiden Photonen eines π^0 -Zerfalls und anderen Teilchen enthalten, sind in die Analyse eingeschlossen. Mit den im ersten Teil der Analyse bestimmten Werten der Trainingsgrößen werden neuronale Netzwerke zur Identifikation der Objekte verwendet, die Energie von den beiden Photonen mindestens eines π^0 -Zerfalls enthalten. Diese Objekte werden im folgenden als nicht-isolierte π^0 -Mesonen bezeichnet. Dabei wird vor allem auf den Einfluß der Informationen des HES auf die Effizienz der π^0 -Identifikation und auf die Abhängigkeit der π^0 -Identifikation von der Gesamtenergie der betrachteten Objekte im CAL eingegangen. Außerdem wird untersucht, ob die Anzahl der zur π^0 -Identifikation verwendeten Informationen durch geeignete Kombination reduziert werden kann, ohne die Effizienz der π^0 -Identifikation zu verschlechtern. Die Arbeit gliedert sich wie folgt: Nach einer Beschreibung des HERA-Speicherringes (Kapitel 2) und des ZEUS-Detektors (Kapitel 3) werden Aufbau und Funktion des Kalorimeters bzw. des Hadron-Elektron-Separators in Kapitel 4 bzw. 5 erläutert. Anschließend werden Kinematik und Identifikation des π^0 -Zerfalls beschrieben (Kapitel 6). In Kapitel 7 werden Aufbau und Funktionsweise neuronaler Netzwerke erläutert und die in dieser Analyse verwendeten *feed forward*-Netzwerke beschrieben. Nach Erläuterung der Datenselektion (Kapitel 8) werden die Ergebnisse der Identifikation von isolierten bzw. nicht-isolierten π^0 -Mesonen in Kapitel 9 bzw. 10 dargestellt. Eine Zusammenstellung der Ergebnisse wird in Kapitel 11 gegeben.

Kapitel 2

HERA

2.1 Der Speicherring HERA

Der Speicherring HERA (Hadron-Elektron-Ring-Anlage) am Deutschen Elektronen-Synchrotron (DESY) in Hamburg hat einen Umfang von 6336 m und verläuft 10-25 m unter der Erdoberfläche. Er besteht aus zwei separaten Beschleunigersystemen, mit denen in vier Experimentierhallen Untersuchungen durchgeführt werden. Abb. 2.1 zeigt den Speicherring HERA mit seinen vier Experimentierhallen und das System seiner Vorbeschleuniger. In den Beschleunigersystemen werden Elektronen (seit August 1994 statt dessen Positronen) und Protonen auf 27,5 GeV bzw. 820 GeV beschleunigt und zur Kollision gebracht.

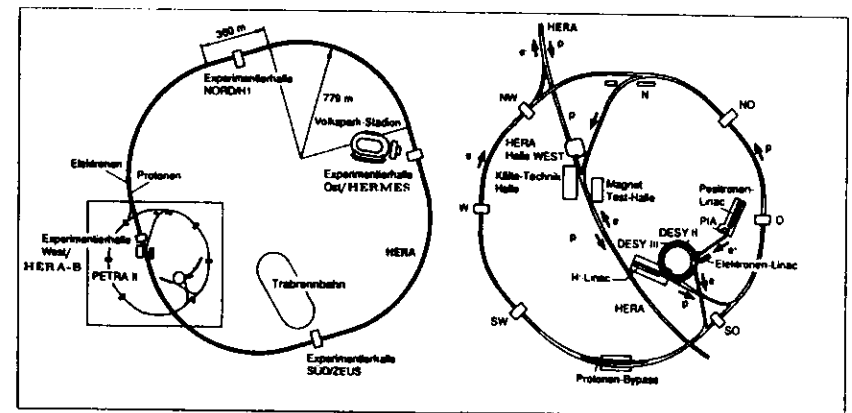


Abbildung 2.1: Der Speicherring HERA am Deutschen Elektronen-Synchrotron (DESY) in Hamburg mit seinen vier Experimentierhallen (links) und dem System seiner Vorbeschleuniger (rechts). In den vier Experimentierhallen sind die Experimente ZEUS (Halle Süd), HERA-B (Halle West), H1 (Halle Nord) und HERMES (Halle Ost) aufgebaut bzw. im Aufbau.

Die einzelnen beschleunigten Teilchenpakete bestehen aus ca. $4 \cdot 10^{10}$ Einzelteilchen. Die mit einer Schwerpunktsenergie von ca. 300 GeV stattfindenden Reaktionen werden mit den beiden Detektoren ZEUS (Halle Süd) und H1 (Halle Nord) gemessen. 1995 hat das Experiment HERMES [5] in der Halle Ost mit der Datennahme begonnen. Ein viertes Experiment, HERA-B [6], befindet sich in der Halle West im Aufbau. Im Gegensatz zu ZEUS und H1 werden bei HERMES und HERA-B keine Kollisionen der Positronen- und Protonenpakete untersucht, sondern Experimente mit feststehenden Zielen am Positronenring (HERMES) bzw. Protonenring (HERA-B) durchgeführt.

Seit dem Anfang der Datennahme der Experimente ZEUS und H1 im Jahre 1992 konnte die Leistung von HERA mit jeder Meßperiode gesteigert werden. In Tabelle 2.1 sind die Werte einiger wichtiger HERA-Parameter für die Meßperioden 1993-1995 und ihre Design-Werte dargestellt. Die Steigerung der Luminosität L hat für die Experimente große Bedeutung. Aus den mit den Detektoren gemessenen Zählraten \dot{N} der verschiedenen identifizierten Prozesse werden mit der Luminosität ihre Wirkungsquerschnitte σ über die Beziehung $\dot{N} = L \cdot \sigma$ berechnet. Eine weiterführende Beschreibung der beiden HERA-Speicherringe ist in [7] gegeben.

| Parameter | 1993-Werte | | 1994-Werte | | 1995-Werte | | Design-Werte | |
|--|------------|------|------------|-----|------------|---------|--------------|-----|
| | e^- | p | e^\pm | p | e^+ | p | e^\pm | p |
| Teilchenenergie (GeV) | 26,6 | 820 | 27,5 | 820 | 27,5 | 820 | 30 | 820 |
| Strahlstrom (mA) | 15,0 | 15,0 | 25,0 | 50 | 30-41 | 55-73 | 58 | 160 |
| Anzahl der Teilchenpakete | 90 | 90 | 168 | 170 | 189 | 180 | 210 | 210 |
| Anzahl der Teilchen pro Paket (10^{10}) | 2,0 | 2,0 | 3,6 | 4,0 | 3,6 | 4,2-5,3 | 3,6 | 10 |
| Luminosität L ($10^{30} \text{cm}^{-2} \text{s}^{-1}$) | 1,0 | | 2,5 | | 3,9-6,3 | | 15 | |
| Integrierte Luminosität (nb^{-1}) | 1088 | | 6216 | | 12150 | | 10^5 | |
| Schwerpunktsenergie (GeV) | 295 | | 300 | | 300 | | 314 | |
| Abstand der Pakete (ns) | 96 | | | | | | | |

Tabelle 2.1: Wichtige HERA-Parameter. Bei den angegebenen Werten von 1993-1995 handelt es sich um Mittel-, bei der integrierten Luminosität um die jeweiligen Absolutwerte.

2.2 Tiefunelastische Positron-Proton-Streuung

HERA ermöglicht das Studium einer Vielzahl verschiedener Reaktionen mit sehr unterschiedlichen Energieüberträgen. Eine detaillierte Beschreibung der bei HERA möglichen physikalischen Prozesse ist unter [8] gegeben. Da in dieser Analyse nur Daten der tiefunelastischen Positron-Proton-Streuung verwendet werden, wird sich auf eine genauere Beschreibung dieses Prozesses beschränkt.

Bei der tiefunelastischen Positron-Proton-Streuung wechselwirkt das Positron über den Austausch eines Eichbosons mit den Konstituenten des Protons. In Abb. 2.2 sind die Feynmandiagramme in niedrigster Ordnung für die beiden möglichen Reaktionstypen dargestellt. Während bei einer Wechselwirkung über einen neutralen Strom (kurz NC für „neutral current“, vgl. Abb. 2.2 links) das Positron über den Austausch eines virtuellen Photons oder eines neutralen Z^0 -Bosons mit einem Valenzquark des Protons wechselwirkt,

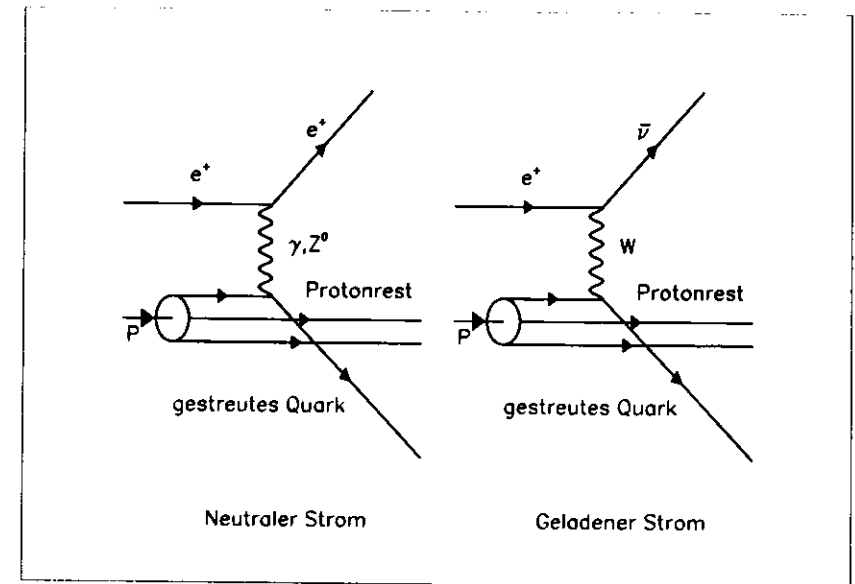


Abbildung 2.2: Tiefunelastische Positron-Proton-Streuung bei HERA über neutrale Ströme (NC, links) und geladene Ströme (CC, rechts).

erfolgt die Wechselwirkung bei einem geladenen Strom (kurz CC für „charged current“, vgl. Abb. 2.2 rechts) über den Austausch eines W^+ - bzw. W^- -Bosons.

Das gestreute Quark und die nicht an der Wechselwirkung beteiligten Konstituenten des Protons fragmentieren in hadronische Jets, den Stromjet und den Protonjet. Die Jets bestehen aus stark wechselwirkenden Teilchen wie Pionen, Nukleonen und Kaonen. Die Teilchen des Protonjets werden in einem Kegel von $\theta \approx 10$ mrad um die Strahlachse emittiert und verlassen das Strahlrohr erst in einiger Entfernung vom Wechselwirkungspunkt. Sie können daher von den Detektoren i.a. nicht nachgewiesen werden. Der Stromjet und das gestreute Positron (NC) bzw. Antineutrino (CC) werden aufgrund der Impulserhaltung zu verschiedenen Seiten der Strahlachse emittiert und tragen entgegengesetzten Transversalimpuls. Der Endzustand der tiefunelastischen Positron-Proton-Streuung besteht in niedrigster Ordnung aus einem Lepton (Positron bzw. Antineutrino), dem Stromjet und dem i.a. nicht meßbaren Protonjet. Neben den in Abb. 2.2 dargestellten Reaktionen sind auch Reaktionen höherer Ordnung möglich.

Zur Beschreibung der Kinematik der tiefunelastischen Positron-Proton-Streuung werden drei Größen verwendet: das Quadrat der Schwerpunktsenergie s , das Quadrat des Viererimpulsübertrages q^2 und das Quadrat der Gesamtmasse des hadronischen Endzustandes W'^2 .

Sie sind definiert als:

$$\begin{aligned}s &= (p_p + p_e)^2 \approx 4 \cdot E_p \cdot E_e \\ q^2 &= (p_e - p_e')^2 = -Q^2 \\ W'^2 &= (q + p_p)^2\end{aligned}$$

Dabei sind E_p und p_p die Energie und der Viererimpuls des einlaufenden Protons und E_e bzw. p_e Energie bzw. Viererimpuls des einlaufenden Positrons. p_e' bezeichnet den Viererimpuls des auslaufenden Leptons. Die kinematischen Variablen x , y und ν liefern eine äquivalente Beschreibung. Sie sind definiert als:

$$\begin{aligned}x &= \frac{Q^2}{2 \cdot m_p \cdot \nu} \\ y &= \frac{2 \cdot m_p \cdot \nu}{s} \\ \nu &= \frac{q \cdot p_p}{m_p}\end{aligned}$$

Dabei ist m_p die Protonmasse. Die drei Größen sind nicht unabhängig, sondern über die Beziehung

$$Q^2 = s \cdot x \cdot y$$

verknüpft. Bei einer Wechselwirkung über einen neutralen Strom können die drei Größen aus der Energie E_e' und dem Streuwinkel θ_e des auslaufenden Positrons, aus der Energie E_j und dem Produktionswinkel θ_j des Stromjets oder aus einer Kombination von zwei dieser Größen berechnet werden. Dabei werden θ_e und θ_j gegen die Flugrichtung der einlaufenden Protonen gemessen. Bei einer Wechselwirkung über einen geladenen Strom stehen zur Berechnung nur die E_j und θ_j zur Verfügung, da das auslaufende Antineutrino nicht nachgewiesen werden kann. Eine genauere Darstellung der möglichen Methoden ist unter [9] gegeben.

Im Quark-Parton-Modell wird x mit dem Bruchteil des Protonimpulses des an der Wechselwirkung beteiligten Quarks identifiziert. y ist der relative Energieverlust des Positrons im Ruhesystem des Protons. Durch HERA konnte die Grenze des bisher zugänglichen kinematischen Bereiches in x von 10^{-3} auf 10^{-5} und in Q^2 von 400 GeV^2 auf 40000 GeV^2 erweitert werden. Damit ist es bei HERA möglich, Daten über die innere Struktur des Protons mit einer Auflösung von $3 \cdot 10^{-18} \text{ m}$ zu erhalten.

Kapitel 3

Der ZEUS-Detektor

Der ZEUS-Detektor wurde zur bestmöglichen Vermessung der bei HERA auftretenden Prozesse konzipiert. Eine detaillierte Beschreibung des Detektors findet sich in [10]. An dieser Stelle wird zunächst das ZEUS-Koordinatensystem beschrieben und anschließend ein kurzer Überblick über den Aufbau des Detektors und die Funktion der verschiedenen Detektorkomponenten gegeben.

Das rechtshändige ZEUS-Koordinatensystem wurde so gewählt, daß die positive Z-Achse identisch mit der Flugrichtung der Protonen im HERA-Ring ist, während die X-Achse in der Ebene des Ringes liegt und auf sein Zentrum gerichtet ist. Die Y-Achse zeigt nach oben. Der Azimutwinkel wird bezüglich der positiven Z-Achse gemessen. Daher beträgt er für die einlaufenden Positronen 180° . Der Nullpunkt des Koordinatensystems ist der nominelle Wechselwirkungspunkt im Zentrum der Detektors. Die Abb. 3.1 und 3.2 zeigen zwei Schnitte durch den ZEUS-Detektor: Abb. 3.1 senkrecht zur Strahlachse, Abb. 3.2 parallel dazu. Der Detektor ist bezüglich des Azimutwinkels nahezu symmetrisch (vgl. Abb. 3.1). Abweichungen von der Symmetrie vor allem im unteren Bereich des Detektors sind durch die technischen Probleme der Konstruktion bedingt. In Abb. 3.2 ist der asymmetrische Aufbau des Detektors bezüglich der Z-Achse zu erkennen. Durch die hohe Energiedifferenz zwischen Protonen (820 GeV) und Positronen ($27,5 \text{ GeV}$) sind die Teilchendichte und die mittlere Energie pro Teilchen in Flugrichtung der Protonen größer als in der Gegenrichtung. Um auch dort effizient zu messen, wurde der Detektor in dieser Richtung tiefer ausgebaut.

Der Detektor besteht aus mehreren Komponenten, die um den nominellen Wechselwirkungspunkt angeordnet sind. Im inneren Bereich befinden sich Spurkammern (VXD, CTD, FDET, RTD) und das hochauflösende Kalorimeter (FCAL, BCAL, RCAL). Zwischen BCAL und CTD ist eine supraleitende Spule installiert, die in der CTD ein homogenes Magnetfeld von $1,54 \text{ Tesla}$ parallel zum Strahlrohr erzeugt. Außerhalb des Kalorimeters befinden sich das Eisenjoch (YOKE), das Außenkalorimeter (BAC) und die Myonendetektoren (FMUI, BMUI, RMUI bzw. FMUON, BMUON, RMUON). Der Detektor ist von einem Betonmantel zur Strahlenabschirmung umgeben. Weitere spezielle Detektorkomponenten sind in weiterer Entfernung zum Wechselwirkungspunkt am Strahlrohr aufgebaut. Sie dienen dem Nachweis von Teilchen, die unter kleinen Winkeln gegen die Strahlachse emittiert werden. Diese Teilchen verlassen das Strahlrohr erst außerhalb der in Abb. 3.1 bzw. 3.2 dargestellten Detektorkomponenten und können daher von diesen nicht nachgewiesen werden. So dienen z.B. das Vorwärts-Proton-Spektrometer [11]

bzw. das Vorwärts-Kalorimeter [12] zum Nachweis von unter kleinen Winkeln in Flugrichtung der Protonen auslaufenden Protonen bzw. Neutronen.

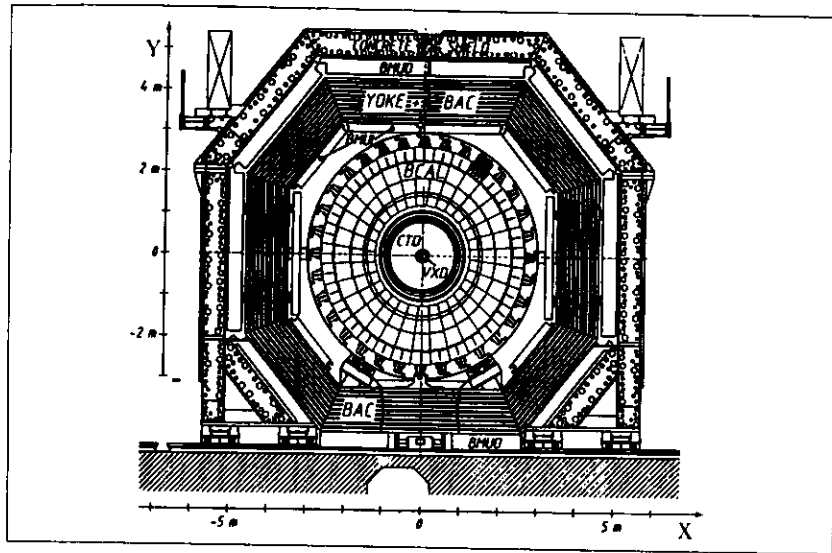


Abbildung 3.1: Der ZEUS-Detektor im Querschnitt.

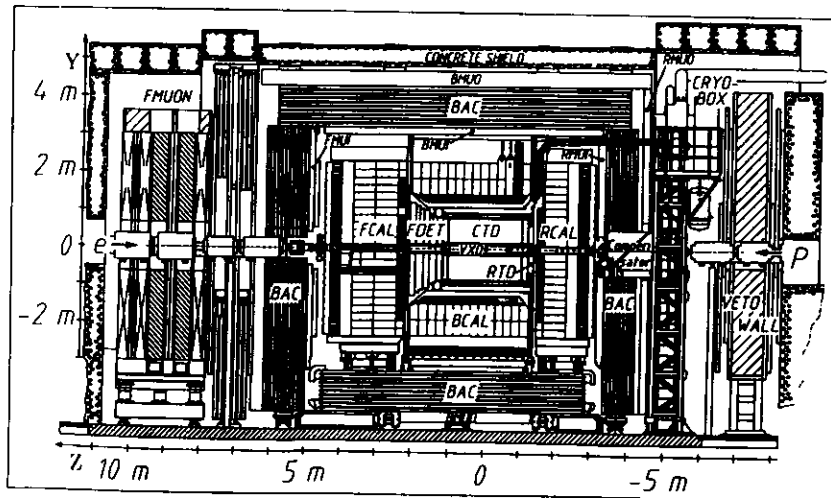


Abbildung 3.2: Der ZEUS-Detektor im Längsschnitt

Um die zur Berechnung der Wirkungsquerschnitte im Detektor nachgewiesener Reaktionen notwendige Luminosität zu messen, sind in Flugrichtung der Positronen zwei Kalorimeter installiert. Sie bilden den Luminositätsdetektor. Im folgenden werden die in den Abbildungen 3.1 und 3.2 dargestellten Detektorkomponenten und der Luminositätsdetektor kurz beschrieben.

3.1 Die Spurkammern

Der innere Bereich des Detektors wird vom Vertexdetektor (VXD), einer Driftkammer zur genauen Bestimmung des Wechselwirkungspunktes, eingenommen. Die zentrale Spurkammer (CTD) umschließt den Vertexdetektor. Durch eine supraleitende Spule wird in der CTD ein homogenes Magnetfeld von 1,54 Tesla erzeugt, das die Bahnen der geladenen Teilchen krümmt. Die Richtung der Krümmung und der Krümmungsradius geben Aufschluß über die Ladung und den Impuls dieser Teilchen. Weitere Spurkammern (FDET, RTD) ergänzen die zentrale Spurkammer nach vorne bzw. hinten.

3.2 Das Kalorimeter

Um das System der Spurkammern ist das hochauflösende Kalorimeter (CAL) eingebaut. Das Kalorimeter dient zur Energiemessung und ist mechanisch in Vorwärts- (FCAL), Rückwärts- (RCAL) und Zentrales-Kalorimeter (BCAL) unterteilt. Es besteht aus abwechselnden Lagen von 3,3 mm angereichertem Uran und 2,6 mm Szintillator. Die Dicke der Lagen ist so gewählt, daß Elektronen und Hadronen gleicher Energie gleich große Signale erzeugen. Das Kalorimeter wird ausführlich in Kapitel 4 beschrieben. Um eine bessere Unterscheidung von Hadronen und Elektronen zu ermöglichen, wurden im Kalorimeter eine (BCAL, RCAL) bzw. zwei (FCAL) 14 mm breite Schlitze nach drei bzw. drei und sechs Uranplatten freigehalten, um dort den Hadron-Elektron-Separator (HES) zu installieren [13]. Da der Hadron-Elektron-Separator neben dem Kalorimeter die für diese Arbeit wichtigste Detektorkomponente ist, wird er ausführlich in Kapitel 5 beschrieben.

3.3 Eisenjoch und Außenkalorimeter

Umschlossen wird das Kalorimeter von einem Eisenjoch (YOKE). Es hat die Form eines an beiden Seiten von Endkappen abgeschlossenen langgestreckten Oktagons und dient zur Rückführung des magnetischen Flusses der supraleitenden Spule. Gleichzeitig wird das Eisenjoch durch die abwechselnde Anordnung von gasgefüllten Proportionalkammern und 7,5 cm dicken Eisenplatten als Außenkalorimeter (BAC) verwendet, um die Energie nachzuweisen, die nicht im CAL deponiert wird. Zur Identifikation und Impulsmessung von Myonen wird im Eisenjoch ein toroidales Magnetfeld von 1,6 Tesla erzeugt.

3.4 Die Myonenkammern

Die Identifikation und Impulsmessung der Myonen erfolgt in allen Myonenkammern nach demselben Prinzip. Die Impulse geladener Teilchen vor und hinter einem dicken Absorber werden verglichen. Im Gegensatz zu Myonen, die im Absorber nur einen geringen Energieverlust erleiden, werden Hadronen im Absorber gestoppt oder stark abgebremst. Die Eisenplatten des BAC dienen als Absorber, während zur Impulsmessung großflächige Zähler, vom Wechselwirkungspunkt aus gesehen, vor (FMUI, BMUI, RMUI) und hinter (FMUON, BMUON, RMUON) dem Eisenjoch aufgebaut sind. Sie bestehen aus Proportionalkammern, die im begrenzten „Streamer-Modus“ betrieben werden. Aus den gemessenen Positionen vor und hinter dem BAC wird der Impuls der durchlaufenden Teilchen berechnet. Zur Myon-Identifikation werden diese mit den in der CTD bestimmten Impulsen verglichen.

Im Vorwärts-Myonspektrometer (FMUON) werden Myonen mit höheren Impulsen (bis ca. 100 GeV/c) als in den zentralen Myonenkammern (BMUON) und Rückwärtsmyonenkammern (RMUON) erwartet. Daher ist das FMUON anders aufgebaut als das BMUON und das RMUON. Außerhalb des Eisenjochs sind zwei Eisentoroide aufgebaut, in denen von 16 normalleitenden Kupferspulen ein Magnetfeld von 1,7 Tesla erzeugt wird. Die Myonenspuren werden durch Positionsmessung in 4 Driftkammern bestimmt, die vor, hinter und zwischen den Eisentoroiden installiert sind.

3.5 Die Vetowand

Durch Streuung der Protonen mit Restgas im Strahlrohr, sowie mit der Wand und den Kollimatoren des Strahlrohres wird der Detektor aus der Richtung der einfliegenden Protonen ständig von einer großen Zahl von Hadronen und Myonen getroffen. Im Detektor werden Energiedepositionen und Spuren dieser Teilchen gemessen. Da die Anfangsbedingungen dieser Reaktionen nicht exakt bekannt sind, ist eine Analyse nicht möglich. Die Vetowand [14] wurde zur Unterscheidung dieser Reaktionen von Wechselwirkungen im Detektor und zum Schutz des Detektors vor ihren Reaktionsprodukten installiert. Sie befindet sich, vom Wechselwirkungspunkt aus gesehen, hinter dem RMUON. Eine senkrecht zur Strahlachse aufgebaute Eisenwand dient als Absorber für den Großteil der Reaktionsprodukte. Vor und hinter der Eisenwand ist je ein Szintillatorhodoskop aufgebaut. Diese beiden Hodoskope werden zur Identifikation der Untergrundreaktionen verwendet.

3.6 Der Luminositätsdetektor

Zur Messung der Luminosität wird der Bremsstrahlungsprozeß



verwendet. Durch den kleinen Energieübertrag vom Positron auf das Photon werden beide unter kleinen Winkeln zur Flugrichtung der Positronen emittiert. Sie können daher von

den in Abb. 3.1 und 3.2 dargestellten Detektorkomponenten nicht nachgewiesen werden. Der Luminositätsdetektor besteht aus zwei separaten Kalorimetern für Positronen und Photonen, die bei $z = -35$ m bzw. $z = -104$ m bis -107 m am Strahlrohr aufgebaut sind. Beide Kalorimeter bestehen aus abwechselnden Schichten von Blei- und Szintillatorplatten. Zur Positionsmessung wird ein Silizium-Streifen-Detektor verwendet.

Die Positronen, die durch den Bremsstrahlungsprozeß (3.1) Energie abgegeben haben, werden durch die Ablenkmagnete des HERA-Ringes aus dem Strahlrohr heraus in das Positronenkalorimeter gelenkt, während die von ihnen abgestrahlten Photonen von den Magneten nicht beeinflusst werden und geradeaus weiterfliegen. Sie gelangen durch ein Fenster im Strahlrohr in das Photonenkalorimeter. Die Luminosität wird durch die Zählrate im Photonenzähler und den mit der Bethe-Heitler-Formel [15] berechenbaren Wirkungsquerschnitt der Reaktion (3.1) bestimmt.

Kapitel 4

Das hochauflösende Kalorimeter

4.1 Aufbau und Funktionsweise des Kalorimeters

Das hochauflösende Kalorimeter des ZEUS-Detektors wird zur Bestimmung der Energien von Teilchen und Teilchenjets betrieben. Ort und Verteilung der Energiedeposition im Kalorimeter geben zusammen mit den Informationen anderer Detektorkomponenten Aufschluß über die stattgefundenen Reaktionen. Das Kalorimeter ist ein sogenanntes „Sandwichkalorimeter“, es besteht aus abwechselnden Lagen von 3,3 mm abgereichertem Uran und 2,6 mm Szintillator. Die Dicke der Lagen ist so gewählt, daß das Kalorimeter voll kompensierend arbeitet, d.h. für Positronen und Hadronen gleicher Energie gleich große Signale liefert. Das Kalorimeter ist in Vorwärtskalorimeter (FCAL), Zentrales-Kalorimeter (BCAL) und Rückwärtskalorimeter (RCAL) unterteilt, deren Lage in Abb. 4.1 dargestellt ist. Die drei Kalorimeterteile sind so angeordnet, daß sich die von ihnen abgedeckten Winkelbereiche überschneiden (vgl. Tabelle 4.1). Zusammen umfassen sie, vom Wechselwirkungspunkt aus gesehen, 99,7% des Raumwinkels.

FCAL und BCAL sind senkrecht zur Strahlachse aufgebaut. Das BCAL bildet einen konzentrisch um das Strahlrohr herum aufgebauten Zylinder. Die Gliederung der drei Kalorimeterteile folgt demselben Prinzip. Sie bestehen aus mehreren Modulen, die mechanisch voneinander getrennt sind. Die einzelnen Module sind in sogenannte Auslesetürme unterteilt. Jeder Turm hat eine Frontfläche von $20 \times 20 \text{ cm}^2$ und ist longitudinal in eine elektromagnetische (EMC-) Sektion und eine (RCAL) bzw. zwei (FCAL, BCAL) hadronische (HAC-) Sektionen unterteilt. Diese Unterteilung dient der Unterscheidung von elektromagnetisch und hadronisch wechselwirkenden Teilchen, da sich von hadronischen Teilchen erzeugte Teilchenschauer i.a. tiefer im Kalorimeter entwickeln als die Schauer

| Kalorimeterteil | Winkelbereich in θ |
|-----------------|-----------------------------|
| FCAL | $2,2^\circ - 39,9^\circ$ |
| BCAL | $36,7^\circ - 129,1^\circ$ |
| RCAL | $128,1^\circ - 176,5^\circ$ |

Tabelle 4.1: Winkelbereiche der verschiedenen Kalorimeterteile

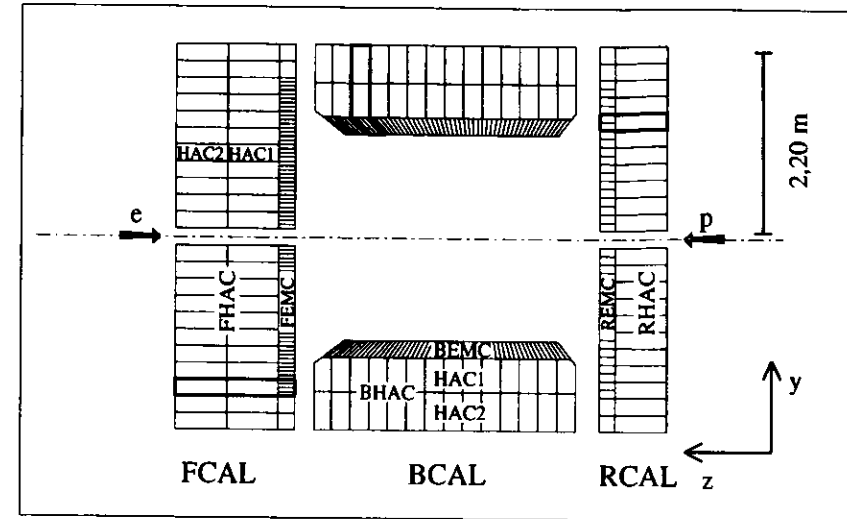


Abbildung 4.1: Darstellung des ZEUS-Kalorimeters in Z-Y-Projektion¹. Das Kalorimeter ist in Vorwärtskalorimeter (FCAL), zentrales Kalorimeter (BCAL) und Rückwärtskalorimeter (RCAL) unterteilt. Die Rechtecke stellen die einzelnen Kalorimeterzellen dar. Die schwarz umrandeten Bereiche deuten jeweils einen Kalorimeterturm an (siehe Text).

elektromagnetisch wechselwirkender Teilchen (siehe Kapitel 4.2). Die EMC-Sektionen sind in vier (FCAL, BCAL) bzw. zwei (RCAL) EMC-Zellen mit einer Frontfläche von $20 \times 5 \text{ cm}^2$ bzw. $20 \times 10 \text{ cm}^2$ unterteilt. Die einzelnen HAC-Sektionen bilden jeweils eine einzige HAC-Zelle. Pro Zelle transportieren zwei Wellenlängenschieber die gemessenen Signale zu den Photovervielfachern. Die Wellenlängenschieber befinden sich an den gegenüberliegenden Seiten der Zellen, während die Photovervielfacher an den vom Wechselwirkungspunkt abgewandten Seiten der Module installiert sind.

Zur besseren Unterscheidung von Hadronen und Positronen befinden sich innerhalb der EMC-Sektionen der drei Kalorimeterteile Lücken für den Einbau einer weiteren Detektorkomponente, des Hadron-Elektron-Separators (HES). Er wird in Kapitel 5 detailliert beschrieben.

Ein komplettes FCAL-Modul mit Wellenlängenschiebern und Photovervielfachern ist in Abb. 4.2 dargestellt. Die Aussparung für die erste Lage des vorderen Hadron-Elektron-Separators („silicium detector“) nach drei Uranplatten ist sichtbar, während die zweite Lage nach sechs Uranplatten nicht eingezeichnet ist. Der im oberen Bereich der Abbildung dargestellte C-Arm dient zur Aufnahme der Elektronik. Die Uran- und Szintillatorplatten der einzelnen Türme des Moduls werden durch Edelstahlbänder zusammengespant („tension strap“).

¹Aus: C. Edelbüttel, Hadronische Energieflußmessungen in der tiefinelastischen ep-Streuung, DESY F35D-96-04, Hamburg, 1996

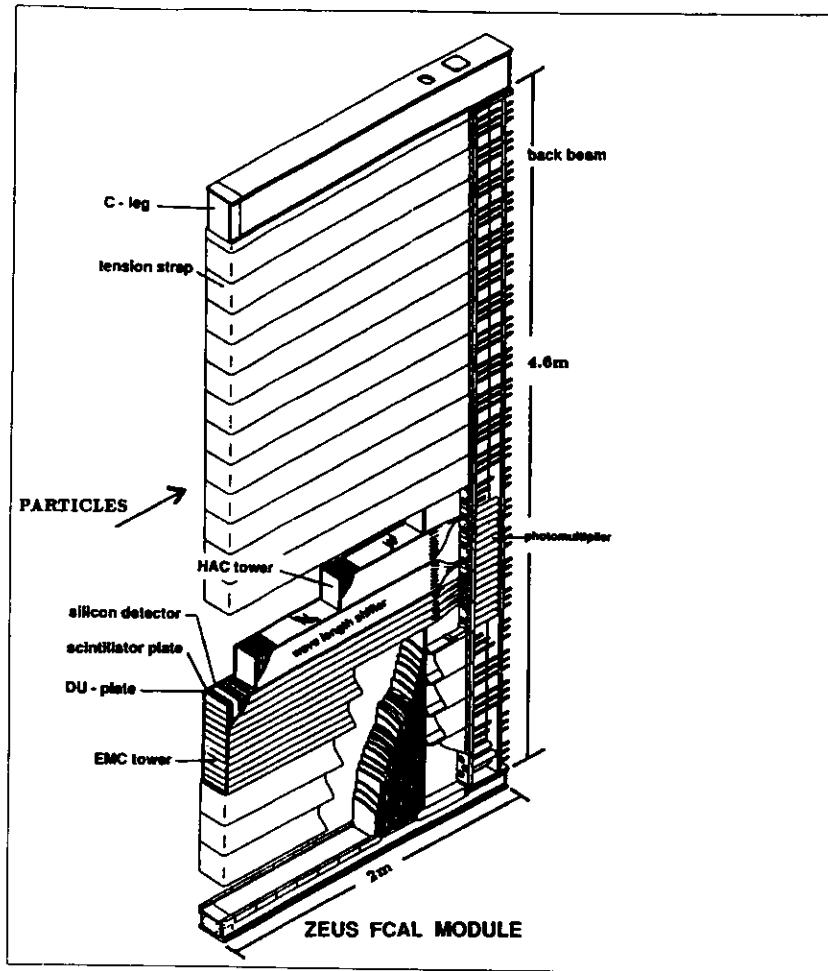


Abbildung 4.2: ZEUS-FCAL-Modul. Die einzelnen Türme bestehen aus abwechselnden Schichten von abgereicherten Uran- und Szintillatorplatten. Ein Turm ist in vier EMC-, Zellen, eine HAC1- und eine HAC2-Zelle unterteilt. Die Signale der einzelnen Zellen werden durch jeweils zwei Wellenlängenschieber zu den im hinteren Bereich des Moduls installierten Photovervielfachern geleitet. Die Aussparung für die erste Lage des vorderen Hadron-Elektron-Separators („silicium detector“) nach drei Uranplatten ist ebenfalls sichtbar, während die zweite Lage nach sechs Uranplatten nicht eingezeichnet ist. Der im oberen Bereich der Abbildung dargestellte C-Arm dient zur Aufnahme der Elektronik.

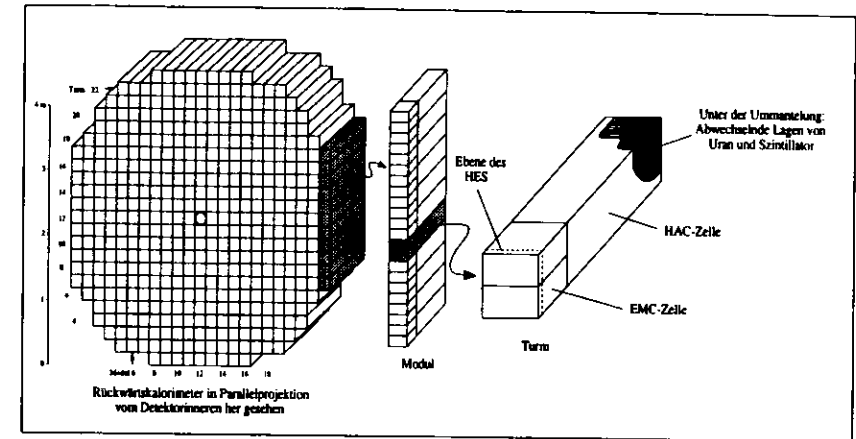


Abbildung 4.3: Das rückwärtige Kalorimeter des ZEUS-Detektors in Parallelprojektion vom Detektorinneren her gesehen². Das RCAL besteht aus 23 Modulen, die wiederum senkrecht in bis zu 23 Türme unterteilt sind. Jeder Turm besteht aus zwei EMC-Zellen und einer HAC-Zelle. Die einzelnen Zellen bestehen aus abwechselnden Lagen von abgereichertem Uran und Szintillator. Der HES ist nach drei Uranplatten im elektromagnetischen Teil des Kalorimeters eingebaut.

Eine weitergehende Beschreibung des Kalorimeters ist unter [16] gegeben. Im folgenden wird speziell der Aufbau des RCAL behandelt, da dort die Daten der durchgeführten Analyse gewonnen wurden. Das RCAL nimmt den Bereich zwischen $z = -148$ cm und $z = -309$ cm ein. Eine schematische Darstellung des RCAL ist in Abb. 4.3 gegeben. Das RCAL besteht aus 23 Modulen, die senkrecht nebeneinander angeordnet sind. Ihre Höhe nimmt von 4,60 m in der Mitte bis zu 2,20 m am Rand ab, während die Tiefe bis auf die etwas kürzeren äußeren Module 1,60 m beträgt. Die einzelnen Türme haben eine Höhe und Breite von je 20 cm. Daher ergibt sich die Höhe bzw. die Breite einer EMC-Zelle eines Turmes mit 10 bzw. 20 cm. Ihre Tiefe beträgt 23,6 cm. Die HAC-Zelle eines Turmes ist $20 \times 20 \times 64,4$ cm³ groß. Hinter den ersten drei Uranplatten innerhalb der EMC-Sektion des RCAL ist der HES installiert.

Durchqueren hochenergetische Teilchen das Kalorimeter, so entwickeln sie überwiegend in den Absorberplatten aus abgereichertem Uran kaskadenförmige Teilchenschauer. In den Szintillatorplatten werden die Moleküle des Szintillators durch Wechselwirkung mit diesen Teilchenschauern angeregt und senden Szintillationslicht aus. Das Szintillationslicht jeder Zelle wird durch zwei Wellenlängenschieber an den Seiten des Moduls zu größeren Wellenlängen hin verschoben und zu den Photovervielfachern geleitet. Die Intensität dieses Lichtes ist ein Maß für die im Kalorimeter deponierte Energie. Die natürliche Radioaktivität des abgereicherten Urans wird zur Kalibration und Überwachung der Energiemes-

²Aus: J.O. Carstens, Identifikation von Elektronen im ZEUS-Hadron-Elektron-Separator mit neuronalen Netzwerken, DESY F35D-94-11, Hamburg, 1994

sung des Kalorimeters benutzt. Die Energieauflösung im ZEUS-Kalorimeter beträgt

$$\sigma_{E,e}^{\text{CAL}} = 18\% \cdot \sqrt{E}, E \text{ in GeV} \quad (4.1)$$

$$\sigma_{E,h}^{\text{CAL}} = 35\% \cdot \sqrt{E}, E \text{ in GeV} \quad (4.2)$$

für Elektronen (e) bzw. Hadronen (h) [17].

4.2 Entwicklung von Teilchenschauern im Kalorimeter

Zur Identifikation der verschiedenen Teilchen im Kalorimeter werden Lage und Form der von ihnen erzeugten Teilchenschauer bestimmt. Die Schauer von Hadronen unterscheiden sich durch ihre Lage und laterale Ausdehnung im Kalorimeter von den durch Elektronen, Positronen oder Photonen erzeugten Teilchenschauern. Myonen bilden i.a. keine Teilchenschauer aus, sondern durchqueren das Kalorimeter als minimal ionisierende Teilchen. Sie werden daher primär in der speziellen Myonendetektoren nachgewiesen (vgl. Kapitel 3). In Abb. 4.4 ist das Schauerverhalten von Hadronen (in diesem Fall geladenen Pionen), Elektronen und Myonen schematisch für einen Teil des RCAL dargestellt.

Bei den im ZEUS-Detektor beobachteten Reaktionen sind die Energien der Teilchen im Endzustand i.a. sehr viel größer als ihre Ruhemasse m_0 . In diesem Fall wird der Energieverlust von Elektronen und Positronen beim Durchgang durch Materie durch die Abgabe von Bremsstrahlung dominiert. Bei diesem Prozeß strahlen sie Energie in Form eines Photons ab. Ist die Energie des abgestrahlten Photons größer als die doppelte Ruheenergie der Elektronen von 0,511 MeV, so kann sich das Photon in ein Elektron-Positron-Paar umwandeln. Diese Paarbildung kann ebenso wie die Abgabe von Bremsstrahlung wegen Impuls- und Energieerhaltung nur im elektromagnetischen Feld eines geladenen Teilchens erfolgen. Bei beiden Prozessen haben die zwei Sekundärteilchen im Mittel die Hälfte der Energie des Primärteilchens. Durch mehrfache Paarbildung und Bremsstrahlung entwickelt sich aus dem ursprünglichen Elektron bzw. Positron ein elektromagnetischer Teilchenschauer aus Elektronen, Positronen und Photonen. Treten Photonen ins Kalorimeter ein, so beginnt die Entwicklung der Teilchenschauer durch Paarbildung. Die Anzahl der Teilchen nimmt zu, bis die mittlere Teilchenenergie der kritischen Energie E_k entspricht. E_k ist die Teilchenenergie, bei der der Energieverlust durch Bremsstrahlung gleich dem Energieverlust durch Ionisation ist. Anschließend geben die Elektronen bzw. Positronen ihre Energie durch Ionisation und die Photonen durch Photoeffekt und Comptonstreuung ab. Dadurch werden die Teilchen abgebremst und der Teilchenschauer stirbt aus.

Die Entwicklung dieser Schauer in verschiedenen Materialien wird durch die sogenannte elektromagnetische Strahlungslänge X_0 beschrieben. Nach Durchqueren einer Materialschicht der Dicke X_0 hat die Energie von Elektronen bzw. Positronen im Mittel auf $\frac{1}{2}$ der Anfangsenergie abgenommen. Für Photonen ist die mittlere freie Weglänge X_p als diejenige Länge definiert, nach der die Zahl der ursprünglichen Photonen durch Paarbildung auf $\frac{1}{2}$ des Anfangswertes abgenommen hat. X_p und X_0 sind durch die Beziehung

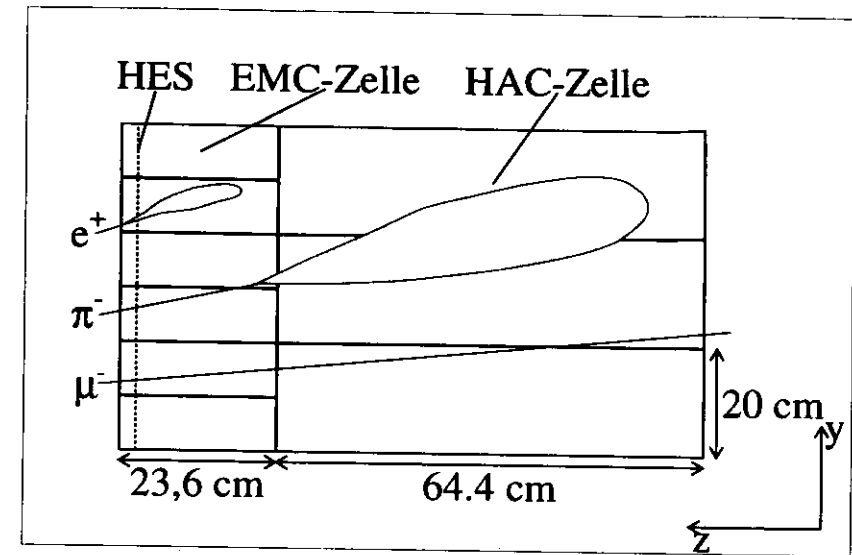


Abbildung 4.4: Illustration der Schauerentwicklung im Kalorimeter. Die grau eingefärbten Bereiche deuten die Ausdehnung der durch die Primärteilchen erzeugten Teilchenschauer an. Die von Hadronen (hier Pionen) erzeugten Teilchenschauer liegen i.a. tiefer im Kalorimeter als elektromagnetische Schauer von Elektronen, Positronen oder Photonen. Diese Teilchen erzeugen in der Regel schon in der Ebene des HES einen Teilchenschauer, Hadronen dagegen i.a. erst dahinter. Sie durchqueren den HES als minimal ionisierende Teilchen. Die transversale Ausdehnung hadronischer Schauer ist in der Regel größer als die elektromagnetischer Schauer. Myonen entwickeln i.a. keine Teilchenschauer im CAL, sondern durchqueren es als minimal ionisierende Teilchen.

$X_0 = \frac{7}{9} \cdot X_p$ miteinander verknüpft. Die elektromagnetische Strahlungslänge im ZEUS-Kalorimeter beträgt

$$X_0 = 0,74 \text{ cm.} \quad (4.3)$$

Zur Beschreibung des Schauerhaltens hochenergetischer Hadronen wird die mittlere freie Weglänge λ_0 verwendet. λ_0 ist die Strecke, die Hadronen im Mittel zwischen zwei Wechselwirkungen zurücklegen. Sie beträgt im ZEUS-Kalorimeter

$$\lambda_0 = 21,0 \text{ cm} \quad (4.4)$$

und ist damit weitaus größer als die elektromagnetische Strahlungslänge X_0 . Bei Hadronen ist der Energieverlust durch Bremsstrahlung vernachlässigbar. Ein hadronischer Teilchenschauer entsteht durch Wechselwirkung der Hadronen mit den Atomkernen des durchflogenen Materials. Eine Darstellung der dabei auftretenden Prozesse ist unter [16]

und [18] gegeben.

Das Schauermaximum t_{max} eines Teilchenschauers ist als die Tiefe des durchflogenen Materials definiert, in der die meisten Sekundärteilchen vorliegen. Sie ist für hadronische ($t_{max,h}$) bzw. elektromagnetische Schauer ($t_{max,e}$) eines Primärteilchens der Energie E durch

$$t_{max,h} = \lambda_0 \cdot (0,2 \cdot \ln \frac{E}{1 \text{ GeV}} + 0,7) \quad (4.5)$$

$$t_{max,e} = X_0 \cdot (\ln \frac{E}{E_k} \pm 0,5) \quad (4.6)$$

gegeben. Dabei gilt das Pluszeichen in (4.6) für Photonen und das Minuszeichen für Elektronen bzw. Positronen. Alle Energien sind in GeV einzusetzen. Die kritische Energie E_k beträgt im ZEUS-Kalorimeter $10,6 \cdot 10^{-3}$ GeV.

Da λ_0 weitaus größer als X_0 ist, folgt aus (4.5) und (4.6), daß $t_{max,h}$ bei Teilchenenergien oberhalb von 35 MeV größer als $t_{max,e}$ ist. Hadronische Schauer entwickeln sich daher i.a. tiefer im Kalorimeter als elektromagnetische Schauer. Das Verhältnis der in den EMC- und HAC-Sektionen deponierten Energien kann aus diesem Grunde zur Unterscheidung der beiden Teilchensorten verwendet werden.

Auch die transversale Ausdehnung hadronischer Schauer ist breiter als die elektromagnetischer Schauer gleicher Energie. 95% der Energie eines hadronischen Schauers werden in einem Zylinder des Radius λ_0 deponiert. Bei elektromagnetischen Schauern ist der Radius dieses Zylinders durch den doppelten Molière-Radius R_M gegeben. Dieser ist definiert als:

$$R_M = X_0 \cdot \frac{\sqrt{\frac{4\pi}{\alpha}} \cdot m_e \cdot c^2}{E_k} \approx X_0 \cdot \frac{21,2 \cdot 10^{-3} \text{ GeV}}{E_k} \quad (4.7)$$

Dabei ist m_e die Elektronenmasse und α die Feinstrukturkonstante. Für das ZEUS-Kalorimeter ergibt sich R_M zu ungefähr $2X_0$. 95% der Energie eines elektromagnetischen Schauers werden daher in einem Zylinder mit dem Radius $4X_0$ deponiert. Damit ist die transversale Ausdehnung eines hadronischen Schauers um etwa eine Größenordnung höher als die eines elektromagnetischen Schauers.

Trotz der unterschiedlichen Schauerprofile elektromagnetischer und hadronischer Teilchenschauer treten bei ihrer Identifikation im Kalorimeter Probleme auf. Vor allem niederenergetische Hadronen schauern bereits in den EMC-Sektionen, und durch Überlagerung der Schauer mehrerer Teilchen können die Profile der einzelnen Schauer nicht mehr exakt rekonstruiert werden.

4.3 Auswertung der Kalorimeterdaten

Der Zweck eines lokalen Zuordnungsalgorithmus ist es, diejenigen Kalorimeterzellen, die zum Schauer einzelner Teilchen bzw. Teilchenjets gehören, zusammenzufassen und damit die Energiedeposition und Position dieser Teilchens bzw. Teilchenjets zu bestimmen.

Der Standard-Algorithmus der Datenrekonstruktion im Kalorimeter ist der Kondensat-Algorithmus [19]. Kalorimeterzellen oberhalb der Mindestenergiedeposition zur Rauschunterdrückung werden von diesem Algorithmus eindeutig jeweils einem Kondensat zugeordnet. Kalorimeterzellen unterhalb der Mindestenergiedeposition werden bei der Konstruktion der Kondensate nicht berücksichtigt. Zur Konstruktion eines Kondensates wird zunächst von allen noch keinem Kondensat zugeordneten Zellen, die Zelle der höchsten Energiedeposition ermittelt. Bei der Konstruktion des ersten Kondensates ist das die Zelle der höchsten Energiedeposition im Kalorimeter. Diese Zelle bildet den Kern des Kondensates. Für alle ihr benachbarten Zellen wird ermittelt, ob sie schon einem Kondensat zugeordnet worden sind. Ist das nicht der Fall, so werden sie dem aktuellen Kondensat angegliedert. Zwei Zellen werden dabei nur dann als benachbart angesehen, wenn sie eine gemeinsame Fläche haben. Für jede dem Kondensat neu angegliederte Zelle werden ebenfalls alle Nachbarzellen überprüft. Werden dem Kondensat keine neuen Zellen mehr zugeordnet, so ist seine Konstruktion abgeschlossen. Die Zelle der höchsten Energiedeposition aller noch keinem Kondensat zugeordneten Zellen bildet die Kernzelle eines weiteren Kondensates. Das Verfahren wird solange durchlaufen, bis alle Zellen oberhalb der Mindestenergiedeposition einem Kondensat zugeordnet worden sind.

In Abb. 4.5 I und II ist in einem Ausschnitt der EMC-Sektionen des RCAL dargestellt, wie der Kondensat-Algorithmus eine gegebene Energieverteilung zu verschiedenen Kondensaten zusammenfaßt. Obwohl die beiden EMC-Zellen der höchsten Energiedeposition nicht nebeneinander liegen, werden sie demselben Kondensat angegliedert. Da der Energiegradient zwischen den einzelnen Kalorimeterzellen nicht berücksichtigt wird, faßt der Kondensat-Algorithmus zwei Schauer zu einem Kondensat zusammen, sobald zwei ihrer Zellen benachbart sind. Daher unterschätzt dieser Algorithmus die Anzahl der Teilchenschauer im Kalorimeter.

Zur Verbesserung der Datenrekonstruktion wurde der Island-Algorithmus entwickelt. Er faßt nicht die einzelnen Kalorimeterzellen zusammen, sondern summiert zuerst für jeden Kalorimeterturm die Energie aller Zellen auf. Anschließend werden die Türme zu sogenannten Islands zusammengefaßt. Bei der Konstruktion der Islands wird der Energiegradient zwischen den verschiedenen Kalorimetertürmen berücksichtigt. Dazu wird für jeden Turm ein Zeiger auf denjenigen Nachbarturm gesetzt, der die höchste Energie besitzt. Im Gegensatz zum Kondensat-Algorithmus werden beim Island-Algorithmus auch solche Kalorimetertürme als benachbart angesehen, die nur eine gemeinsame Kante haben. Hat ein Turm eine höhere Energie als alle seine Nachbartürme, so wird sein Zeiger auf ihn selbst gesetzt. Diese Türme bilden lokale Maxima der Energiedeposition und werden durchnummeriert. Sie bilden die Kerne der einzelnen Islands. Alle Türme, deren Zeiger auf den Kernturm eines Islands gerichtet sind, werden diesem Island zugeordnet. Anschließend wird für alle noch keinem Island zugeordneten Türme überprüft, ob ihre Zeiger auf irgendeinen Turm eines Islands weisen. Ist das der Fall, so werden auch diese Türme den entsprechenden Islands angegliedert. Das Verfahren wird solange durchlaufen, bis alle Türme einem Island zugeordnet sind. In Abb. 4.5 III ist eine Signalverteilung in einzelnen Türmen des RCAL dargestellt. Dabei wurden die Signale der beiden EMC-Zellen eines Turmes aus Abb. 4.5 I zusammengefaßt. In Abb. 4.5 IV sind die Zeiger der einzelnen Kalorimetertürme und die durch den Island-Algorithmus konstruierten Islands dargestellt. Durch die Berücksichtigung des Energiegradienten zwischen den einzelnen Türmen werden statt dem Kondensat 1 (vgl. Abb. 4.5 III) zwei Islands (1 und 3) konstruiert.

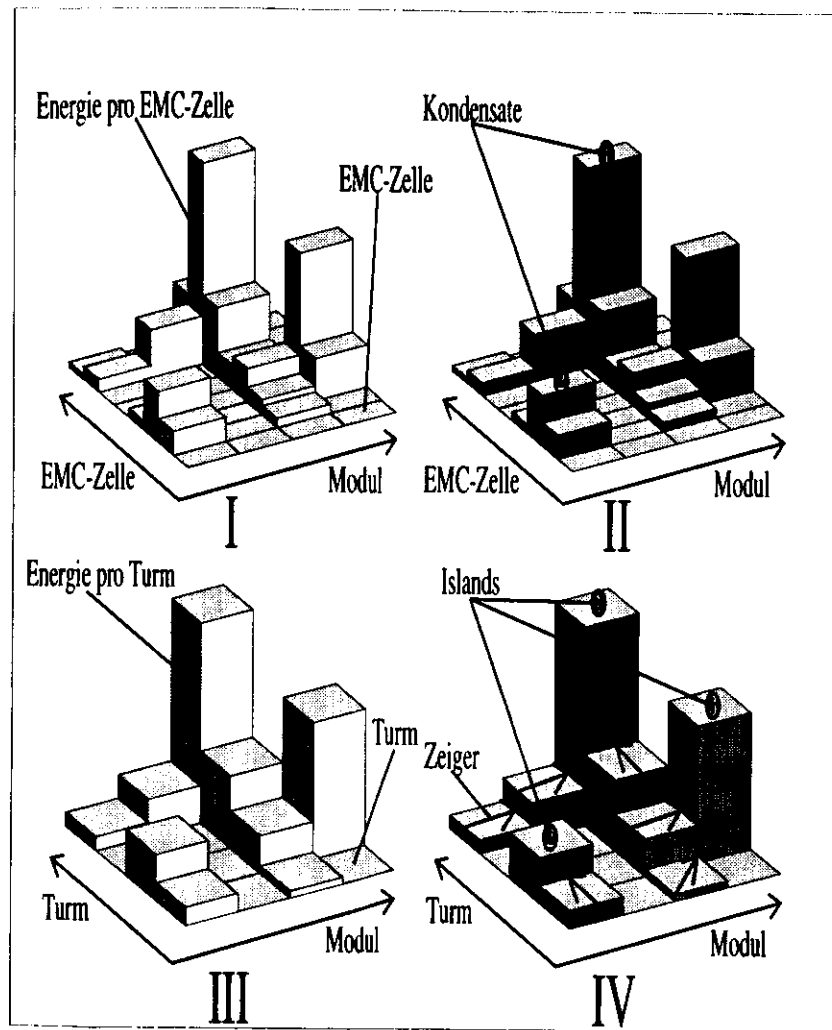


Abbildung 4.5: Datenrekonstruktion im Kalorimeter. Bild I zeigt eine Energieverteilung in den einzelnen EMC-Zellen des RCAL, Bild III die Energieverteilung in den einzelnen Kalorimetertürmen. Die Bilder II bzw. IV zeigen die daraus mit dem Kondensat- bzw. Island-Algorithmus konstruierten Kondensate bzw. Islands. Die unterschiedlichen Graustufen bezeichnen die Zellen bzw. Türme der verschiedenen Kondensate bzw. Islands. Die Pfeile in Bild IV stellen die Zeiger der einzelnen Türme bei der Islandkonstruktion dar (siehe Text).

In dieser Analyse wurde der Island-Algorithmus zur Datenrekonstruktion im Kalorimeter verwendet. Damit können die Signale der beiden Photonen eines π^0 -Zerfalls im Kalorimeter in vielen Fällen auch dann noch in zwei getrennte Islands aufgelöst werden, wenn sie durch den Kondensat-Algorithmus zu einem einzigen Kondensat zusammengefaßt werden. Mit Energie und Position dieser beiden Islands kann die π^0 -Identifikation über die Berechnung der invarianten Masse dieses Systems erfolgen (vgl. Kapitel 6.2). Die Identifikation der π^0 -Zerfälle, bei denen der Island-Algorithmus die Signale der beiden Zerfallsphotonen zu einem einzigen Island zusammenfaßt, ist Gegenstand dieser Analyse.

Kapitel 5

Der Hadron-Elektron-Separator

5.1 Aufbau des Hadron-Elektron-Separators

Um eine bessere Unterscheidung von Hadronen und Elektronen zu ermöglichen, wurden im Kalorimeter eine (BCAL, RCAL) bzw. zwei (FCAL) 14 mm breite Schlitze nach drei bzw. drei und sechs Uranplatten freigehalten, um dort den Hadron-Elektron-Separator (HES) zu installieren. Die Lage des HES im Kalorimeter wurde in der Nähe der maximalen Teilchenzahl elektromagnetischer Schauer gewählt. Für diese Schauer wird daher ein großes Signal beobachtet, während von Hadronen nur eine geringe Energie deponiert wird. Zur Zeit (1995) ist der rückwärtige Hadron-Elektron-Separator (RHES) installiert und in Betrieb. Er wird im folgenden mit HES bezeichnet. Der HES befindet sich im RCAL an der Position der maximalen Teilchenzahl eines von einem 5 GeV Elektron ausgelösten elektromagnetischen Schauers. Da die Lage der maximalen Teilchenzahl elektromagnetischer und hadronischer Schauer nur logarithmisch von der Energie der Primärteilchen abhängt (vgl. Gleichung (4.6)), ermöglicht diese Position auch bei anderen Energien eine Identifikation elektromagnetischer und hadronischer aufschauender Teilchen.

Der HES besteht aus einer Lage Siliziumdioden, die nach drei Uranplatten in den elektromagnetischen Teil des RCAL eingefügt ist. Die Dioden haben eine Grösse von $3,0 \times 3,4 \text{ cm}^2$ und eine Dicke von $400 \mu\text{m}$. Sie werden in sogenannten Skis installiert. In jedem RCAL-Modul werden drei der bis zu 4,3 m langen Skis senkrecht nebeneinander eingebaut. Jeder Ski trägt bis zu 112 Zeilen mit je zwei Dioden nebeneinander. Die Dioden eines Skis sind projektiv angeordnet, so daß die aktive Fläche des HES, vom Wechselwirkungspunkt aus gesehen, in Y-Richtung lückenlos ist. In X-Richtung beträgt der Abstand zwischen den Dioden eines Skis weniger als 1 mm, zwischen den benachbarten Dioden zweier Skis 3 mm und zwischen den Dioden zweier benachbarter Skis in verschiedenen Modulen 16 mm. Die insgesamt 10412 Dioden sind in dem, vom Wechselwirkungspunkt aus gesehen, nicht durch das BCAL verdeckten Bereich des RCAL installiert. Eine Darstellung des Aufbaus der Skis und der Ausleseelektronik des HES ist unter [20] gegeben. Abb. 5.1 zeigt den vollständigen RHES im RCAL, vom Wechselwirkungspunkt aus gesehen.

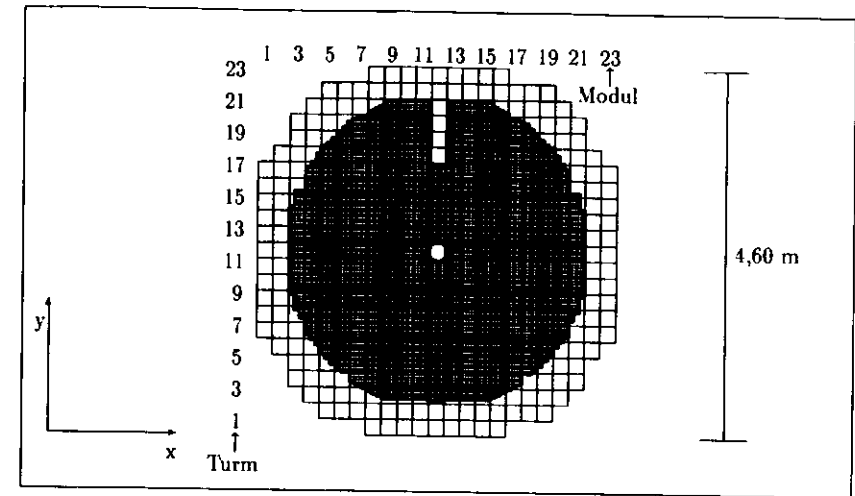


Abbildung 5.1: Schematische Darstellung der aktiven Fläche des RHES (dunkel eingefärbt) auf die Module und Türme des RCAL projiziert. Die Quadrate stellen die $20 \times 20 \text{ cm}^2$ großen Kalorimetertürme dar. Das Strahlrohr nimmt des zentralen Tower des Moduls 12 ein. In den Türmen 18-21 dieses Modules wurde der HES nicht installiert, da dort die Transportleitung des Kühlmittels (flüssiges Helium) der supraleitenden Spule außerhalb der CTD eingebaut ist.

5.2 Funktionsweise des HES

Die Dioden des HES bestehen aus drei unterschiedlich dotierten Siliziumschichten. Die Oberflächenschicht ist stark p-dotiert (p^+) und die unterste Schicht stark n-dotiert (n^+). Die als Detektorschicht betriebene Zwischenschicht besteht aus schwach n-dotiertem Silizium. Die Dioden werden in Sperrichtung betrieben. Dazu wird die unterste Schicht mit Masse und die Oberflächenschicht mit einer Spannung von -100 Volt verbunden. In der Zwischenschicht entsteht die sogenannte Verarmungszone, in der keine freien Ladungsträger mehr vorhanden sind. Durchqueren geladene Teilchen die Verarmungszone, so erzeugen sie Elektron-Loch-Paare, die sich entsprechend ihrer Ladung zur p^+ bzw. n^+ -dotierten Siliziumschicht bewegen. Die erzeugte Ladung ist proportional zur deponierten Teilchenenergie. Ein minimal ionisierendes Teilchen erzeugt ca. 33.000 Elektron-Loch-Paare, was einer deponierten Energie von ca. 120 keV entspricht. Diese Energie definiert die zur Auswertung der HES-Dioden verwendete Energieeinheit 1 mip. Die Ausgangsspannung der ladungsintegrierenden, invertierenden Vorverstärker ist proportional der in den Dioden deponierten Energie.

5.3 Auswertung der HES-Dioden

Die Auswertung der Signale der einzelnen Dioden des HES erfolgt ähnlich wie im Kalorimeter. Die Signale der Dioden werden den lokalen Energiemaxima zugeordnet. Im Gegensatz zum Kalorimeter müssen im HES keine dreidimensionalen Nachbarschaftsbeziehungen berücksichtigt werden, da die HES-Dioden eine Ebene bilden. Da im Durchschnitt 96% der Energie eines elektromagnetischen Schauers in der Diode der maximalen Energiedeposition dieses Schauers und den acht benachbarten Dioden enthalten ist, werden diese neun Dioden in der ZEUS-Standard-Rekonstruktion zu einem Objekt zusammengefaßt [3]. Dabei wird zunächst die Diode der maximalen Energiedeposition des HES ermittelt und mit ihren acht benachbarten Dioden zu einem Objekt zusammengefaßt. Anschließend wird das Energiemaximum aller noch nicht verwendeten Dioden bestimmt und das Verfahren wiederholt. Bei diesem Algorithmus werden Energiedepositionen außerhalb von Gruppen aus 3×3 Dioden als eigenständige Objekte behandelt. Dadurch treten bei der π^0 -Identifikation Probleme auf.

Hat ein Photon eines π^0 -Zerfalls auch außerhalb der acht Dioden um die Diode mit einem lokalen Maximum der Energiedeposition Energie deponiert, so wird das Signal dieses Photons durch die Datenrekonstruktion in zwei Teile aufgespalten. Die Energiedeposition scheint von zwei Teilchen zu stammen. Befinden sich die maximalen Energiedepositionen der beiden Photonen eines π^0 -Zerfalls in benachbarten Dioden, so sind die beiden Teilchen im HES nicht mehr aufzulösen. In diesem Fall wird durch die Konstruktion zusätzlicher Objekte vergetäuscht, daß die beiden Photonen im HES noch als separate Objekte beobachtet werden können.

Um diese beiden Probleme zu vermeiden, wurde in dieser Analyse der Isle-Algorithmus zur Datenrekonstruktion im HES verwendet. Er arbeitet analog zum Island-Algorithmus der Datenrekonstruktion im Kalorimeter, verwendet jedoch die einzelnen HES-Dioden als kleinste Einheiten. Da der Isle-Algorithmus sich bei der Konstruktion der Diodengruppen, den Isles, nicht auf Gruppen von 3×3 Dioden beschränkt und die Energiegradienten zwischen den Dioden berücksichtigt, treten die Probleme des Standard-Algorithmus bei der π^0 -Identifikation nicht auf. Die unterschiedlichen Auswirkungen des Standard-HES-Algorithmus und des Isle-Algorithmus sind in Abb. 5.2 dargestellt. Der Standard-HES-Algorithmus faßt die Dioden der gegebenen Signalverteilung zu vier Gruppen zusammen, wobei die Gruppen 3 und 4 jeweils nur aus einer Diode mit geringer Energiedeposition bestehen. Durch die Berücksichtigung des Energiegradienten zwischen den einzelnen Dioden, konstruiert der Isle-Algorithmus nur zwei Isles. Dabei umfaßt die Isle Nummer 1 alle Dioden der durch den Standard-HES-Algorithmus konstruierten Diodengruppen 1, 3 und 4.

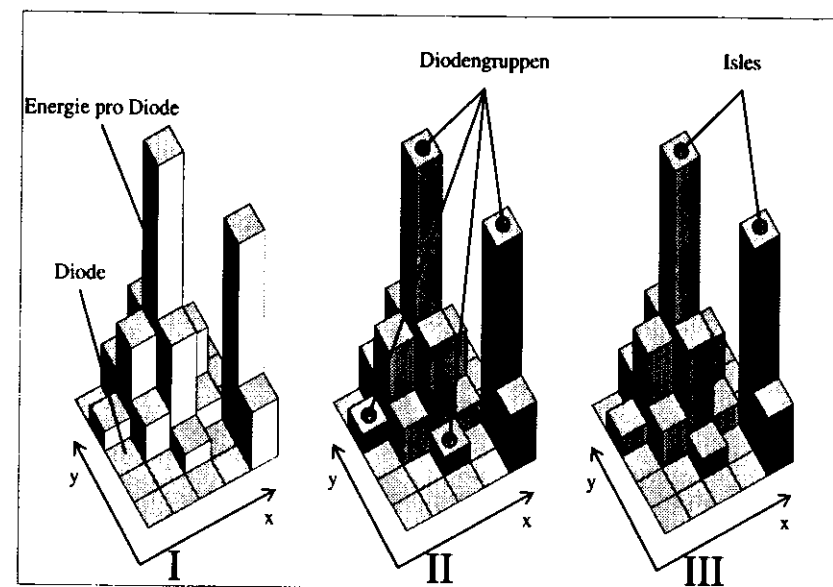


Abbildung 5.2: Datenrekonstruktion im HES. Ein Bin entspricht einer Diode. Die Dioden einer gegebenen Signalverteilung (Bild I) werden durch den Standard-HES-Algorithmus (Bild II) und den Isle-Algorithmus (Bild III) zu verschiedenen Objekten zusammengefaßt. Der Standard-HES-Algorithmus faßt die Dioden zu vier Gruppen zusammen, während der Isle-Algorithmus nur zwei Isles konstruiert. Die unterschiedlichen Graustufen der Energiedeposition deuten die Zuordnung der einzelnen Dioden zu den Diodengruppen bzw. Isles an. Die Zahlen kennzeichnen die zentralen Dioden der konstruierten Objekte.

Kapitel 6

Identifikation von neutralen Pionen

6.1 Kinematik des Zerfalls $\pi^0 \rightarrow \gamma\gamma$

Die Lebensdauer der π^0 -Mesonen beträgt $8,4 \cdot 10^{-17}$ Sekunden [2]. Daher können die im ZEUS-Detektor entstehenden π^0 -Mesonen nicht direkt, sondern nur über ihre Zerfallsprodukte beobachtet werden. Da sie zu 98,8% in zwei Photonen zerfallen [2], wird in dieser Arbeit nur die Identifikation dieses Zerfallkanals betrachtet. Die Effizienz der π^0 -Identifikation hängt vom Öffnungswinkel α zwischen den Impulsvektoren der beiden Photonen im Laborsystem ab. Der minimale Öffnungswinkel α_{\min} ist durch (6.1) gegeben.

$$\tan\left(\frac{\alpha_{\min}}{2}\right) = \frac{m_{\pi^0} c^2}{\sqrt{(E_0)^2 - (m_{\pi^0} c^2)^2}} \quad (6.1)$$

Dabei ist m_{π^0} die Masse der π^0 -Mesonen ($135 \text{ MeV}/c^2$) und E_0 ihre Energie. α_{\min} hängt nur von der Energie der π^0 -Mesonen ab. In Abb. 6.1 ist diese Abhängigkeit dargestellt. Deponieren die beiden Photonen eines π^0 -Zerfalls die Maxima ihrer Energie in benachbarten CAL-Türmen bzw. HES-Dioden, so können ihre Signale im CAL bzw. HES nicht mehr getrennt werden. Die minimalen Öffnungswinkel α_{CAL} und α_{HES} , bei denen die beiden Zerfallsphotonen die Maxima ihre Energien in nicht benachbarten CAL-Türmen bzw. HES-Dioden deponieren, hängen vom Azimutwinkel des zerfallenen π^0 -Mesons ab. Deponieren beide Photonen ihre Energie im RCAL, so sinken α_{CAL} und α_{HES} mit abnehmendem Azimutwinkel. Für die Zerfallsphotonen eines π^0 -Mesons mit einem Azimutwinkel $\theta = 160^\circ$, das am Wechselwirkungspunkt zerfällt, wurden α_{CAL} und α_{HES} berechnet und in Abb. 6.1 eingetragen.

Bei Energien kleiner als 1 GeV ist α_{\min} so groß, daß die beiden Photonen im Kalorimeter als getrennte Objekte beobachtbar sind (vgl. Abb. 6.1). Bei höheren Energien ist die Granularität des Kalorimeters dafür zu groß. Im HES können die beiden Photonen im günstigsten Fall noch bis zu π^0 -Energien von 10 GeV aufgelöst werden. Oberhalb dieser Energie sind die beiden Photonen auch durch den HES nicht mehr als zwei Objekte identifizierbar.

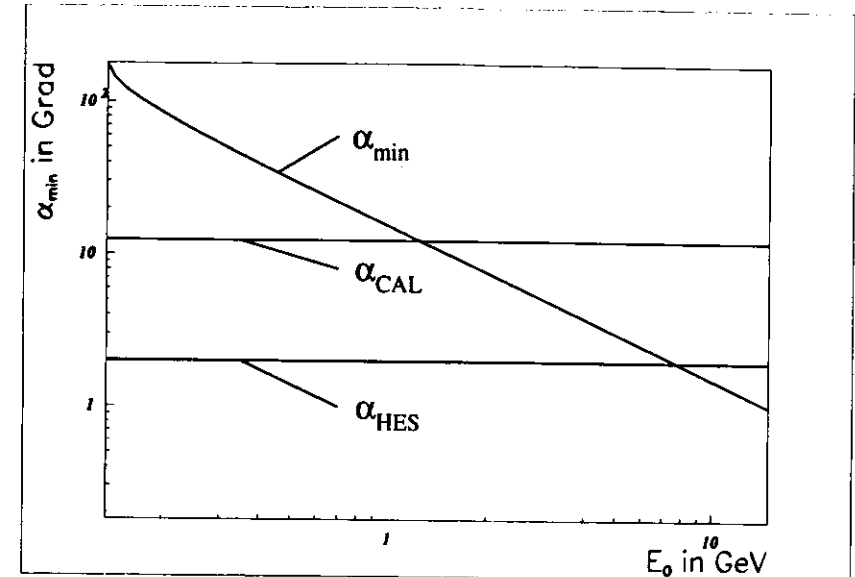


Abbildung 6.1: Minimaler Öffnungswinkel α_{\min} zwischen den Photonen des Zerfalls $\pi^0 \rightarrow \gamma\gamma$ in Abhängigkeit von der Energie E_0 der π^0 -Mesonen. α_{CAL} und α_{HES} sind die minimalen Winkel eines am Wechselwirkungspunkt unter $\theta = 160^\circ$ zerfallenden π^0 -Mesons, bei denen die beiden Zerfallsphotonen im RCAL bzw. HES in nicht benachbarten Türmen bzw. Dioden auftreffen.

6.2 π^0 -Identifikation durch die invariante Masse

Unter der Annahme, daß die π^0 -Mesonen am Wechselwirkungspunkt zerfallen, können die Energiedeposition im CAL und die Position im CAL bzw. HES zur Bestimmung der invarianten Masse M des $\gamma\gamma$ -Systems verwendet werden. Die für eine gegebene Kalorimeterenergie im HES gemessene Energie hat eine breite Verteilung [3], so daß sie nicht zur Bestimmung der Energien der beiden Photonen genutzt werden kann. Die π^0 -Identifikation erfolgt über den Vergleich der invarianten Masse M mit der π^0 -Masse m_{π^0} . Seien E_0 die Energie des π^0 -Mesons und E_1 bzw. E_2 die Energien der beiden Photonen des π^0 -Zerfalls im Laborsystem. Die invariante Masse M ist durch (6.2) gegeben.

$$\begin{aligned} M^2 &= 2 \cdot E_1 \cdot E_2 \cdot (1 - \cos(\alpha)) \\ &= \frac{1}{2} \cdot (E_0^2 - (\Delta E)^2) \cdot (1 - \cos(\alpha)) \end{aligned} \quad (6.2)$$

Dabei ist ΔE die Differenz der Energien beider Photonen und α der Öffnungswinkel zwischen ihren Impulsvektoren. Ist die Energie des π^0 -Mesons klein, so sind die beiden

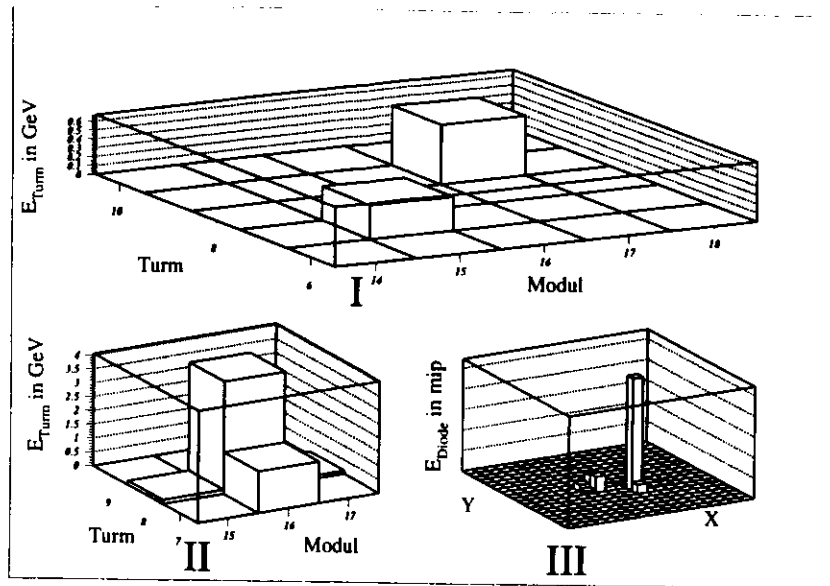


Abbildung 6.2: Energiedeposition im Kalorimeter durch die beiden Photonen eines π^0 -Zerfalls. Bild I zeigt die Energiedeposition eines π^0 -Mesons der Energie von 1 GeV im CAL, Bild II und III die eines π^0 -Mesons von 6,5 GeV im CAL bzw. HES. Ein Bin entspricht einem Kalorimeterturm (Bild I und II) bzw. einer HES-Diode (Bild III). Die beiden Ereignisse stammen aus Monte Carlo-Simulationen.

Photonen im Kalorimeter als zwei getrennte Islands zu erkennen (vgl. Abb. 6.2 I). E_1 und E_2 werden in diesem Fall über die Energien der beiden Islands bestimmt. Der Öffnungswinkel α wird nach Gleichung (6.3) mit den auf den Wechselwirkungspunkt korrigierten Positionen $(x_{1,2}, y_{1,2}, z_{1,2})$ der beiden Islands berechnet:

$$\cos(\alpha) = \frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2} \sqrt{x_2^2 + y_2^2 + z_2^2}} \quad (6.3)$$

Ist E_0 groß gegen die Ruhemasse m_{π^0} , so ist der Öffnungswinkel zwischen den Impulsvektoren der beiden Photonen so klein, daß sich die von ihnen im Kalorimeter ausgelösten Teilchenschauer überschneiden. Liegen die Maxima der beiden Teilchenschauer in aneinandergrenzenden Türmen, so können sie im Kalorimeter nicht mehr getrennt werden; ihre Signale verschmelzen zu einem einzigen Island. Durch die bessere Granularität des HES ist es bei vielen dieser Islands noch möglich, in dieser Detektorkomponente zwei getrennte Signale zu messen. In Abb. 6.2 II und III ist dieser Fall dargestellt. Während im Kalorimeter nur ein Island zu erkennen ist, bilden die Signale der beiden Photonen zwei separate Isles. α wird in diesem Fall aus den auf den Wechselwirkungspunkt korrigierten Positionsmessungen im HES nach Gleichung (6.3) berechnet.

Die Summe der beiden Photonenenergien ist durch die Gesamtenergie des Islands gegeben. Für den Fall, daß die beiden Photonen eines π^0 -Zerfalls im Kalorimeter nicht mehr als zwei getrennt Objekte nachgewiesen werden können, ist der Öffnungswinkel α klein und die Differenz der beiden Photonenenergien vernachlässigbar gering. Daher ist Gleichung (6.4) eine gute Näherung für die Verteilung der π^0 -Energie auf die beiden Photonen:

$$E_1 = E_2 = \frac{1}{2} \cdot E_0 \quad (6.4)$$

Oberhalb von einer π^0 -Energie von 10 GeV ist der Öffnungswinkel der beiden Photonen so klein, daß sie auch im HES nicht mehr separat nachweisbar sind. In diesem Fall kann die invariante Masse M nicht berechnet werden.

6.3 Motivation der Nutzung neuronaler Netzwerke

Bei dem in Kapitel 6.2 beschriebenen Verfahren zur Identifikation von π^0 -Mesonen werden von allen verfügbaren Informationen des CAL und HES nur Energie und Position zur Berechnung der invarianten Masse, und damit zur π^0 -Identifikation, verwendet. Informationen über Lage und Form der durch die beiden Photonen erzeugten Teilchenschauer tragen daher nur indirekt über die Daten- und Positionsrekonstruktion im CAL und HES zur π^0 -Identifikation bei. Durch die Verwendung neuronaler Netzwerke bietet sich die Möglichkeit, diese und andere bisher ungenutzte Informationen aus CAL und HES zur π^0 -Identifikation zu verwenden. Dazu werden neuronale Netzwerke zur Identifikation der Islands verwendet, in denen Energie von beiden Photonen eines π^0 -Zerfalls deponiert worden ist. Für die einzelnen Islands werden den neuronalen Netzwerken Informationen des CAL und HES eingegeben. Die daraus berechneten Ausgabewerte für die einzelnen Islands werden zur π^0 -Identifikation verwendet.

Kapitel 7

Neuronale Netzwerke

7.1 Grundlagen neuronaler Netzwerke

Die Entwicklung künstlicher neuronaler Netzwerke wurde vom Studium des Aufbaus und der Funktionsweise tierischer Nervensysteme, z.B. im menschlichen Gehirn, inspiriert. Neuronale Netzwerke haben nicht den Anspruch, biologische Nervensysteme realistisch zu simulieren, sondern mathematische Modelle mit bestimmten Eigenschaften dieser Systeme zu entwickeln. Einige dieser Eigenschaften sind die hohe Parallelität der durchgeführten Operationen, gute Mustererkennung auch bei unvollständigen oder mit Fremdinformationen behafteten Daten und die Fähigkeit des Lernens. Die Unempfindlichkeit gegenüber Veränderungen des Systems zählt ebenfalls zu diesen Eigenschaften. So wird die Tätigkeit

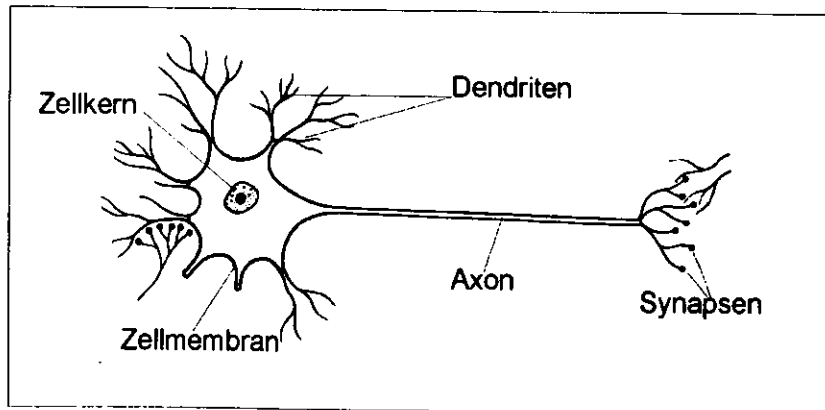


Abbildung 7.1: Schematische Darstellung einer Nervenzelle. Über die Dendriten erhält die Zelle elektrische Signale von anderen Nervenzellen. Übersteigt die Summe der Signale den Schwellenwert der Zelle, so wird über die Synapsen ihres Axons ein Ausgangsimpuls an die Dendriten anderer Zellen weitergegeben.

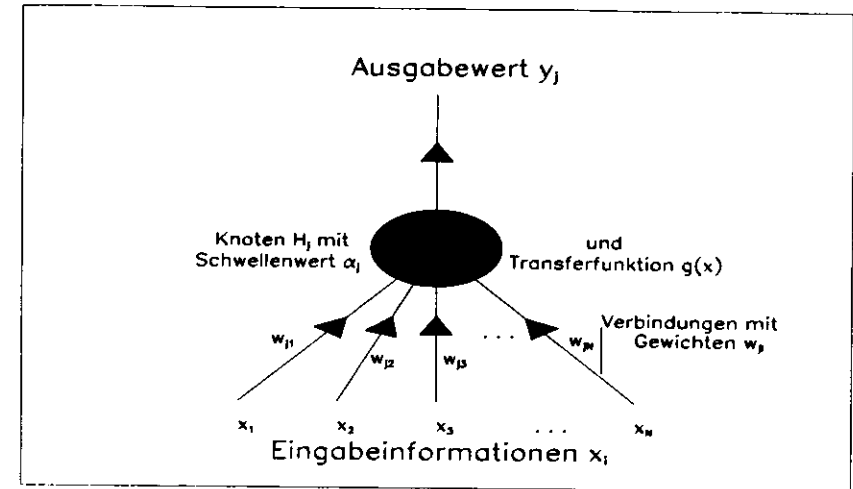


Abbildung 7.2: Funktionsweise der Knoten eines neuronalen Netzwerkes. Die Eingabeinformationen x_i des Knotens H_j werden mit den Gewichten w_{ji} gewichtet und aufsummiert. Mittels der Transferfunktion $g(x)$ und dem Schwellenwert α_j wird der Ausgabewert y_j berechnet. Die Pfeile deuten die Richtung des Datenaustausches an.

des menschlichen Gehirns i.a. nicht merklich davon beeinträchtigt, daß täglich Nervenzellen absterben. Eine schematische Darstellung einer Nervenzelle ist in Abb. 7.1 gegeben. Die einzelnen Nervenzellen sind über ihre Dendriten und Axonen miteinander verbunden. Eine Zelle erhält über ihre Dendriten elektrische Signale von anderen Zellen. Diese Signale bewirken eine Depolarisation der Zellmembran. Übersteigt die Depolarisation einen bestimmten Schwellenwert der Nervenzelle, so wird über ihr Axon ein Ausgangsimpuls an die Synapsen weitergegeben. Dort wird dieser Ausgangsimpuls zu den verschiedenen Dendriten der anderen Zellen durch sogenannte Neurotransmitter gewichtet. Diese chemischen Verbindungen wirken entweder erregend oder hemmend auf die Dendriten der Nachbarzellen [21]. Die Gewichtungen und Schwellenwerte sind i.a. für jede Zelle verschieden. Eine weitergehende Beschreibung der Funktionsweise von Nervenzellen findet sich in [22]. Künstliche neuronale Netzwerke, im folgenden kurz als neuronale Netzwerke bezeichnet, bestehen aus sogenannten Knoten, die untereinander Informationen austauschen. Die Funktionsweise eines Knotens ist schematisch in Abb. 7.2 dargestellt. Der Knoten H_j erhält Eingabeinformationen x_i ($i=1..N$). Die Dendriten werden symbolisch als Verbindungslinien zwischen dem Knoten und den Eingabeinformationen dargestellt. Zur Gewichtung der Eingabeinformationen werden den Verbindungen Gewichte w_{ji} zugeordnet. w_{ji} ist das Gewicht der Verbindung zwischen dem Knoten H_j und der Eingabeinformation x_i . Aus der Summe der gewichteten Eingabeinformationen und dem Schwellenwert α_j des Knotens wird mittels einer mathematischen Transferfunktion $g(x)$ der Ausgabewert des Knotens y_j berechnet. Im einfachsten Modell wird als Transferfunktion eine Stufenfunktion verwendet (vgl. Gleichung (7.1)). Ist die gewichtete Summe der Eingabeinformationen größer als der Schwellenwert des Knotens, so beträgt der Ausgabewert 1,

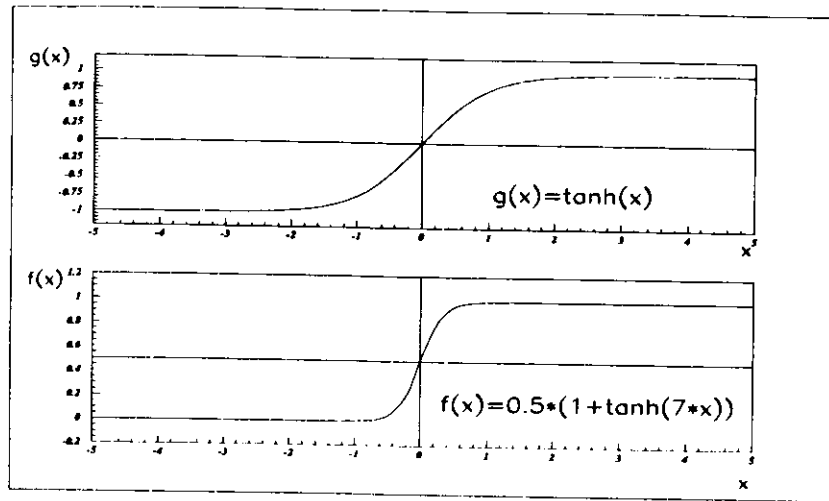


Abbildung 7.3: Beispiele für sigmoide Transferfunktionen neuronaler Netzwerke. Die Funktionen $g(x)$ und $f(x)$ berechnen Ausgabewerte aus dem Intervall $[-1, +1]$ bzw. $[0, +1]$. Außerdem unterscheiden sie sich in der Breite der Übergangszone um $x = 0$, in der die Funktionswerte die größten Abweichungen von den Extremwerten -1 und $+1$ bzw. 0 und 1 haben.

sonst Null:

$$y_j = g(x) = \begin{cases} 1 & \text{für } \sum_{i=1}^N w_{ji} \cdot x_i > \alpha, \\ 0 & \text{sonst} \end{cases} \quad (7.1)$$

Man unterscheidet zwischen linearen und nichtlinearen Transferfunktionen. Durch die Anwendung einer linearen Funktion auf Linearkombinationen der ursprünglichen Eingabeinformationen werden ebenfalls nur lineare Kombinationen dieser Eingabeinformationen berechnet. Daher können neuronale Netzwerke, die nur lineare Transferfunktionen enthalten, nichtlineare Zusammenhänge zwischen der Eingabeinformationen bei der Bearbeitung eines Problems nicht berücksichtigen. Die verwendeten nichtlinearen Transferfunktionen sind i.a. sigmoide Funktionen, deren Funktionswerte in dem Intervall $[-1, +1]$ oder $[0, +1]$ liegen. Solche Funktionen sind z.B. die Signum- oder Tangens-Hyperbolicus-Funktion (vgl. Abb. 7.3).

Im Gegensatz zur Stufenfunktion sind die verwendeten sigmoiden Transferfunktionen in ihrem gesamten Definitionsbereich stetig. Der Übergang von einem Extremwert zum anderen erfolgt kontinuierlich innerhalb der sogenannten Übergangszone um $x = 0$.

Das Argument x der Transferfunktionen stellt allgemein eine Verknüpfung der Eingabeinformationen mit Parametern des Netzwerkes dar. Die Gewichte w_{ji} und Schwellenwerte α_j , sowie alle anderen in die Berechnung von x eingehenden Variablen des Knotens, werden als Parameter des Netzwerkes bezeichnet. Diese Parameter gliedern sich in zwei Gruppen: globale Parameter, die für alle Knoten des Netzwerk identisch sind und lokale

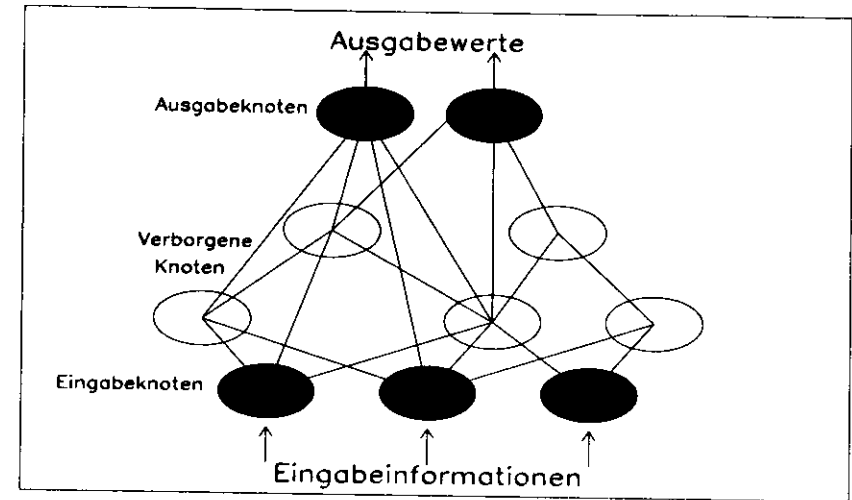


Abbildung 7.4: Allgemeiner Aufbau eines neuronalen Netzwerkes. Die drei Eingabeinformationen jedes Eingabemusters werden durch die drei Eingabeknoten an die verborgenen Knoten bzw. die Ausgabeknoten des Netzwerkes weitergegeben. Die Ausgabeknoten berechnen aus den von den Eingabeknoten und verborgenen Knoten erhaltenen Informationen die beiden Ausgabewerte des Netzwerkes.

Parameter, die sich von Knoten zu Knoten unterscheiden. Zur letzteren Gruppe gehören z.B. die Gewichte w_j und Schwellenwerte α , der einzelnen Knoten H_j .

7.2 Aufbau und Funktion neuronaler Netzwerke

Neuronale Netzwerke berechnen für Gruppen von Eingabeinformationen einen oder mehrere Ausgabewerte. Eine Gruppe von Eingabeinformationen wird im folgenden als Eingabemuster bezeichnet. Die Knoten eines neuronalen Netzwerkes werden in Eingabe-, Ausgabe- und verborgene Knoten eingeteilt. Die einzelnen Eingabeinformationen eines Eingabemusters werden durch jeweils einen Eingabeknoten symbolisiert. Diese Knoten führen keine Rechenoperationen durch, sondern geben die einzelnen Eingabeinformationen an die verborgenen Knoten bzw. die Ausgabeknoten des Netzwerkes weiter. Die Knoten des Netzwerkes, die die Ausgabeinformationen des Netzwerkes berechnen, werden als Ausgabeknoten bezeichnet. Alle Knoten, die nicht unter die Kategorie der Eingabe- oder Ausgabeknoten fallen, werden als verborgene Knoten bezeichnet, da sie keine direkte Verbindung zur „Außenwelt“ haben. In Abb. 7.4 ist ein neuronales Netzwerk mit drei Eingabeknoten, fünf verborgenen Knoten und zwei Ausgabeknoten dargestellt.

Der Datenaustausch zwischen den Knoten eines neuronalen Netzwerkes kann auf zwei Arten erfolgen: rekursiv und nichtrekursiv. Bei nichtrekursiven neuronalen Netzen findet der Datenaustausch zwischen den Knoten nur in einer Richtung statt: von den Eingabeknoten über die verborgenen Knoten zu den Ausgabeknoten. In einem rekursiven Netzwerk

kann der Ausgabewert eines Knotens auch zu einem oder mehreren Knoten, von denen er Eingabeinformationen erhalten hat, oder an sich selbst übermittelt werden. Dort wird er bei der Bearbeitung der nächsten Eingabeinformationen verwendet.

Wird ein neuronales Netzwerk zur Bearbeitung eines bestimmten Problems, beispielsweise eines Klassifikationsproblems (vgl. Kapitel 7.11), verwendet, so müssen die für die Lösung dieses Problem optimalen Parameter des Netzwerkes bestimmt werden. Mit diesen Parametern werden aus den verschiedenen Eingabeinformationen diejenigen Ausgabewerte berechnet, die am wenigsten von den gewünschten Ausgabewerten abweichen. Da diese Parameter nicht analytisch bestimmt werden können, werden bei der Anwendung eines neuronalen Netzwerkes zwei Phasen unterschieden: Trainingsphase und Ausführphase. In der Trainingsphase werden die zur Lösung des Problems optimalen Werte der einzelnen Parameter des Netzwerkes iterativ ermittelt. In der Ausführphase wird das so trainierte Netzwerk zur Bearbeitung des Problems verwendet. Die Fähigkeit, nach dem Training auch bisher unbekannte Muster richtig zu klassifizieren, wird als Generalisierung bezeichnet.

Es werden zwei Arten des Trainings unterschieden: überwachtes und nichtüberwachtes Training. Bei überwachtem Training werden dem neuronalen Netz bei jedem Trainingsmuster auch die gewünschten Ausgabewerte eingegeben. Die Unterschiede zwischen den gewünschten und den durch das neuronale Netz berechneten Ausgabewerten werden dazu verwendet, die Netzwerkparameter iterativ zu verändern, bis die optimalen Parameter erreicht worden sind. Unüberwachtes Training bedeutet dagegen, daß das neuronale Netzwerk keine Kenntnis von der Klassifikation der Trainingsdaten hat. Daher werden bei einem solchen Netzwerk andere Algorithmen als beim überwachtem Training verwendet. Eine Methode ist die des konkurrierenden Lernens [23]. Dabei wird die Lage der Knoten in einem ein- oder mehrdimensionalen Gitter so korrigiert, daß Gebiete kleiner Knotenabstände Eingabemustern mit ähnlichen Eigenschaften entsprechen, während Gebiete großer Knotenabstände die Übergänge zwischen den Musterklassen widerspiegeln. Weitere Informationen über die verschiedenen Formen neuronaler Netzwerke sind unter [24] und [25] gegeben. In den folgenden Abschnitten wird der Aufbau eines *feed forward*-Netzes detailliert dargestellt. Anschließend wird die Funktionsweise eines abgewandelten *feed forward*-Netzwerkes erläutert, da bei der durchgeführten Analyse Netzwerke dieser Art verwendet wurden.

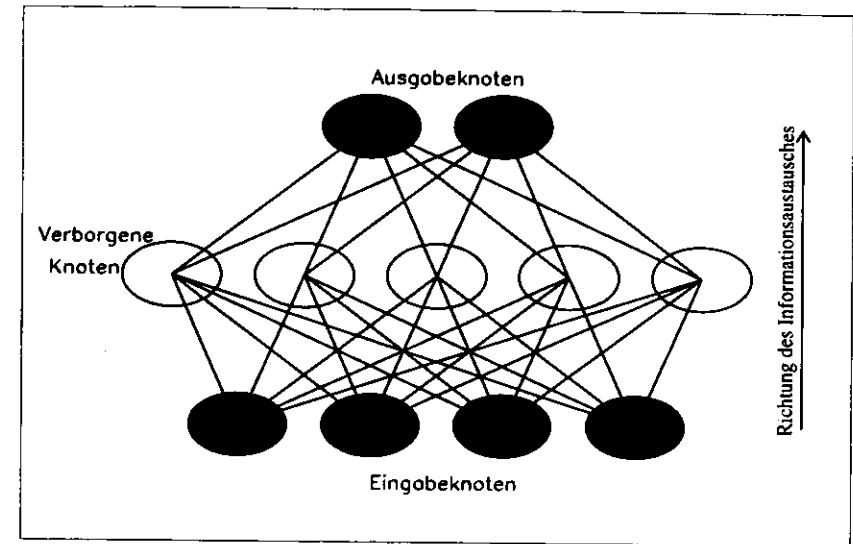


Abbildung 7.5: Aufbau und Funktion eines feed forward-Netzwerkes. Dargestellt ist ein feed forward-Netzwerk mit einer verborgenen Ebene und vier Eingabe-, zwei Ausgabe- und fünf verborgenen Knoten. Der Informationsaustausch erfolgt nur zwischen den Knoten verschiedener Ebenen in der Richtung von der Eingabeebene zur Ausgabebene.

7.3 Feed forward-Netzwerke

In einem *feed forward*-Netzwerk sind die Knoten in verschiedenen Ebenen angeordnet. Ein solches Netzwerk mit drei Ebenen ist in Abb. 7.5 dargestellt. Der Datentransfer ist nichtrekursiv und erfolgt von den Eingabeknoten über die verborgenen Knoten zu den Ausgabeknoten. Die einzelnen Ebenen werden daher von unten nach oben als Eingabeebene, verborgene Ebene und Ausgabebene bezeichnet. Informationsaustausch zwischen den Knoten einer Ebene findet nicht statt.

Die Funktionsweise der *feed forward*-Netze wird an einem Netzwerk mit einer verborgenen Ebene erläutert. Anschließend wird die Verallgemeinerung auf eine beliebige Anzahl verborgener Ebenen diskutiert. Die dazu verwendeten Bezeichnungen sind zur besseren Übersicht in Tabelle 7.1 aufgeführt. Dabei bezeichnen die Indizes i ($i=1..I$) und k ($k=1..K$) die Nummer der Eingabe- bzw. Ausgabeknoten, j ($j=1..J$) die Nummer der verborgenen Knoten und n ($n=1..N$) die Nummer der Eingabemuster.

Für ein gegebenes Eingabemuster M^n ($n=1..N$) bestehend aus den I Einzelinformationen M_i^n ($i=1..I$) werden zunächst die Eingabewerte x_j^n der verborgenen Knoten V_j nach Gleichung (7.2) berechnet:

$$x_j^n = \frac{\sum_{i=1}^I w_{ji} M_i^n - \alpha_j}{t} \quad (7.2)$$

| Symbol | Indexreichweite | Bedeutung |
|---------------------|----------------------|---|
| E_i | $i = 1..I$ | Eingabeknoten |
| V_j | $j = 1..J$ | Verborgener Knoten |
| A_k | $k = 1..K$ | Ausgabeknoten |
| w_{ji} | $i = 1..I, j = 1..J$ | Gewicht zwischen Eingabeknoten i und verborgenem Knoten j |
| w'_{kj} | $j = 1..J, k = 1..K$ | Gewicht zwischen verborgenem Knoten j und Ausgabeknoten k |
| α_j | $j = 1..J$ | Schwellenwert des verborgenen Knotens j |
| α_k | $k = 1..K$ | Schwellenwert des Ausgabeknotens k |
| M^n | $n = 1..N$ | Eingabemuster des Netzwerkes |
| M_i^n | $n = 1..N, i = 1..I$ | Eingabeinformation i des Musters M^n |
| x_j^n | $n = 1..N, j = 1..J$ | Eingabewert des verborgenen Knotens j für Muster M^n |
| V_j^n | $n = 1..N, j = 1..J$ | Ausgabewert des verborgenen Knotens j für Muster M^n |
| x_k^n | $n = 1..N, k = 1..K$ | Eingabewert des Ausgabeknotens k für Muster M^n |
| $A_k^{n,Netz}$ | $n = 1..N, k = 1..K$ | Ausgabewert des Ausgabeknotens k für Muster M^n |
| $A_k^{n,erwünscht}$ | $n = 1..N, k = 1..K$ | Erwünschter Ausgabewert des Ausgabeknotens k für Muster M^n |

Tabelle 7.1: Liste der Bezeichnungen zur Beschreibung der feed forward-Netzwerke

Es wird über alle Eingabeinformationen M_i^n summiert. w_{ji} ist das Gewicht der Verbindung zwischen dem Eingabeknoten E_i ($i = 1..I$) und dem verborgenen Knoten V_j ($j = 1..J$) und α_j der Schwellenwert des verborgenen Knotens V_j . t ist ein konstanter globaler Parameter des Netzwerkes und wird als Temperatur bezeichnet. Die Temperatur hat im Standard *feed forward*-Netzwerk keine Bedeutung, da die Gewichte der einzelnen Knoten nicht normiert sind. Sie wird daher i.a. auf $t = 1$ gesetzt. Im abgewandelten *feed forward*-Netzwerk spielt sie jedoch eine zentrale Rolle beim Training des Netzwerkes (vgl. Kapitel 7.5). Um eine einheitliche Konvention beizubehalten, wurde t schon an dieser Stelle eingeführt.

Aus den nach (7.2) berechneten Eingabeinformationen x_j^n werden mit der sigmoiden Transferfunktion $g(x)$ die Ausgabewerte V_j^n der verborgenen Knoten nach (7.3) berechnet.

$$V_j^n = g(x_j^n) = g\left(\frac{\sum_{i=1}^I w_{ji} M_i^n - \alpha_j}{t}\right) \quad (7.3)$$

Die Berechnung der Eingabewerte x_k^n und Ausgabewerte $A_k^{n,Netz}$ der Ausgabeknoten erfolgt analog zu (7.2) und (7.3):

$$x_k^n = \frac{\sum_{j=1}^J w'_{kj} V_j^n - \alpha_k}{t} \quad (7.4)$$

$$A_k^{n,Netz} = g(x_k^n) \quad (7.5)$$

α_k sind die Schwellenwerte der K Ausgabeknoten des Netzwerkes und w'_{kj} die Gewichte der Verbindungen zwischen den verborgenen Knoten V_j ($j=1..J$) und den Ausgabeknoten A_k ($k=1..K$). Es wird über die Ausgabewerte V_j^n der verborgenen Knoten summiert. Temperatur t und Transferfunktion $g(x)$ sind identisch zu denen in Gleichung (7.3). Sind mehr als eine verborgene Ebene im Netzwerk vorhanden, so müssen die Ein- und

Ausgabewerte der Knoten jeder dieser Ebenen ebenfalls aus den Ausgabewerten der Knoten der jeweils vorangegangenen Ebene berechnet werden. Dabei sind Temperatur und Transferfunktion identisch, während die Gewichte und Schwellenwerte für jeden Knoten verschieden sind. Die Gewichte und Schwellenwerte aller Knoten sind freie Parameter des Netzwerkes. Sie bilden einen Satz von Parametern, der in der Trainingsphase für ein gegebenes Problem optimiert werden muß.

7.4 Training der feed forward-Netzwerke

Zu Beginn des Trainings werden die Anfangswerte der Gewichte und Schwellenwerte des Netzwerkes zufällig aus einem gewissen Intervall initialisiert. Dabei muß darauf geachtet werden, daß die Gewichte der Eingabeinformationen und die Schwellenwerte zweier Knoten einer Ebene nicht gleich sind. Diese Symmetrie kann durch das Training nicht aufgehoben werden. Das Training findet in sogenannten Iterationen v statt. In jeder Iteration werden die einzelnen Gewichte und Schwellenwerte jeweils einmal verändert. Dazu werden dem Netz $N_{training}$ Eingabemuster M^n ($n = 1..N_{training}$) und die erwünschten Ausgabewerte der Ausgabeknoten $A_k^{n,erwünscht}$ eingegeben. Aus der Differenz der erwünschten Ausgabewerte und der Ausgabewerte des Netzes $A_k^{n,Netz}$ wird die Fehlerfunktion $E^{(v)}$ nach der Formel

$$E^{(v)} = \frac{1}{2N_{training}} \sum_{k,n} (A_k^{n,Netz} - A_k^{n,erwünscht})^2 \quad (7.6)$$

berechnet. Dabei wird über die verwendeten $N_{training}$ Eingabemuster und alle K Ausgabeknoten summiert. v bezeichnet die Iteration. Die Fehlerfunktion $E^{(v)}$ ist die mittlere quadratische Abweichung der Ausgabewerte des Netzwerkes von den gewünschten Ausgabewerten. Sie ist eine kontinuierlich differenzierbare Funktion aller Parameter des neuronalen Netzwerkes. Um die zur Lösung des gegebenen Problems optimalen Parameter des Netzwerkes zu erhalten, wird das globale Minimum von E bestimmt. Dazu werden die partiellen Ableitungen der Fehlerfunktion nach den Schwellenwerten und Gewichten berechnet und diese Parameter von Iteration (v) zu Iteration ($v+1$) nach (7.7) und (7.8) verändert:

$$y^{(v+1)} = y^{(v)} - \eta \cdot \Delta y^{(v)} \quad (7.7)$$

$$\Delta y^{(v)} = \frac{\partial E}{\partial y} + \kappa \cdot y^{(v-1)} \quad (7.8)$$

Dabei steht y repräsentativ für die Schwellenwerte α , bzw. α_k und die Gewichte w , bzw. w'_{kj} . Die Schrittweite η ($\eta \in [0, \infty)$) und der Geschwindigkeitsfaktor κ ($0 \leq \kappa < 1$) bleiben während des ganzen Trainings konstant. Sie werden im folgenden als Trainingsparameter bezeichnet. Je größer die Schrittweite η ist, desto stärker wird der Wert des Parameters y pro Trainingsiteration korrigiert. Ist der Gradient $\frac{\partial E}{\partial y}$ im Parameterraum groß, so kann

eine zu hohe Schrittweite eine Oszillation des Wertes von y um den optimalen Wert verursachen. Ist η bei kleinem Gradienten zu klein, sind viele Iterationen nötig, um den optimalen Wert zu erreichen. Der Term $\kappa \cdot y^{(v-1)}$ wurde eingeführt, um das Konvergenzverhalten des Verfahrens zu verbessern. $y^{(v-1)}$ ist der Wert von y der vorangegangenen Iteration. Bei großen Gradienten bewirkt dieser Term eine Abschwächung der Korrektur $\Delta y^{(v)}$, bei kleinem Gradienten eine Verstärkung. Die Werte von η und κ sind problem-spezifisch.

Die Anzahl N_{Training} der dem Netz pro Trainingsiteration vorgeführten Eingabemuster M^n kann kleiner sein als die Gesamtzahl N der Trainingsmuster. Wird die Mittelung in Gleichung (7.6) jedoch über viele Eingabemuster durchgeführt, so werden Oszillationen des Wertes der Fehlerfunktion zwischen den einzelnen Iterationen durch zu starkes Überwiegen einzelner Muster unterdrückt und die Konvergenz des Trainingsprozesses verbessert. Daher wurde bei den für die Analyse benutzten Netzwerken in jeder Iteration über alle Trainingsmuster gemittelt. Durch die zufällige Initialisierung der Parameter und der iterative Natur des Trainings ist es möglich, daß nicht das globale Minimum der Fehlerfunktion erreicht wird, sondern daß die Iteration bei einem lokalen Minimum endet. Zur Verbesserung des Konvergenzverhaltens zum globalen Minimum der Fehlerfunktion wurde für diese Analyse ein abgewandeltes feed forward-Netzwerk benutzt.

7.5 Abgewandeltes feed forward-Netzwerk

Die in dieser Analyse benutzte Methode zur Minimierung der Fehlerfunktion (7.6) nutzt die Abhängigkeit dieser Funktion von der Temperatur t des Netzwerkes aus [26]. Sind die Eingabeinformationen in der Größenordnung von 1 und die Gewichte w_{ji} und w_{kj} der verborgenen Knoten bzw. der Ausgabeknoten in der Form

$$\sum_{i=1}^I (w_{ji})^2 = 1, \forall j \quad ; \quad \sum_{j=1}^J (w'_{kj})^2 = 1, \forall k \quad (7.9)$$

auf 1 normiert, so wird die Zahl der freien Parameter des Netzwerkes um die Summe der Anzahlen der verborgenen Knoten und Ausgabeknoten ($J+K$) reduziert und die Form der Fehlerfunktion von t dominiert. Mit steigender Temperatur sind die lokalen Extrema der Fehlerfunktion immer schwächer ausgeprägt, während bei niedrigen Temperaturen alle ihre Strukturen vorhanden sind. Die Temperatur ist in diesem Modell ebenfalls ein Trainingsparameter des Netzwerkes. Zur Minimierung der Fehlerfunktion wird die Temperatur in derselben Weise verändert wie die Gewichte und Schwellenwerte des neuronalen Netzwerkes (vgl. (7.10) und (7.11)). Die dabei verwendete Schrittweite ist um eine Größenordnung niedriger als die bei der Veränderung der Gewichten und Schwellenwerten benutzte Schrittweite η . Dadurch wird sichergestellt, daß sich die Struktur der Fehlerfunktion (7.6) in Abhängigkeit von t im Laufe der Trainingsiterationen so verändert, daß die Gewichte und Schwellenwerte weiterhin in Richtung auf das globale Minimum der Fehlerfunktion hin optimiert werden.

| Trainingsparameter | Definition | Bezeichnung |
|--------------------|--|----------------------------------|
| t^0 | ≥ 5 | Anfangstemperatur des Netzwerkes |
| γ | Nicht exakt definiert. | Minimalschrittweite |
| $\eta(t)$ | $\eta(t) = 1 + \gamma - \tanh^2 \frac{1}{t}$ | Schrittweite |
| κ | Nicht exakt definiert. | Geschwindigkeitsfaktor |

Tabelle 7.2: Trainingsparameter des abgewandelten feed forward-Netzwerkes.

$$t^{(v+1)} = t^{(v)} - \frac{\eta}{10} \cdot \Delta t^{(v)} \quad (7.10)$$

$$\Delta t^{(v)} = \frac{\partial E}{\partial t} + \kappa \cdot t^{(v-1)} \quad (7.11)$$

Eine Anfangstemperatur t^0 bei Beginn des Trainings von

$$t^0 \geq 5 \quad (7.12)$$

ist notwendig, um die Minimierung der Fehlerfunktion wie beschrieben durchzuführen [26]. Um Oszillationen in der Temperatur bei kleinen Temperaturen zu vermeiden, wird die Schrittweite η mit sinkender Temperatur ebenfalls verändert:

$$\eta(t) = 1 + \gamma - \tanh^2 \frac{1}{t} \quad (7.13)$$

Da bei kleinen Schrittweiten die Anzahl der benötigten Iterationen und damit die Rechenzeit stark ansteigt, wird der Trainingsparameter γ eingeführt. γ ist die Minimalschrittweite und wird je nach Problem zwischen 0.01 und 0 gewählt. In Tabelle 7.2 sind alle Trainingsparameter des Netzwerkes mit ihren Definitionen dargestellt.

7.6 Der back propagation-Algorithmus

Der back propagation-Algorithmus dient zur Berechnung der Änderungen der Parameter eines feed forward-Netzwerkes während der Trainingsphase. Dabei werden die Differenzen $A_k^{n, \text{Netz}} - A_k^{n, \text{erwünscht}}$ für alle Ausgabeknoten A_k und Trainingsmuster M^n dazu verwendet, die Änderungen der Gewichte und Schwellenwerte des Netzwerkes für jede Trainingsiteration nach den Gleichungen (7.7) und (7.8) zu bestimmen. Dazu müssen die partiellen Ableitungen der Fehlerfunktion (7.6) nach den einzelnen Gewichten und Schwellenwerten berechnet werden.

Die Ableitungen der Fehlerfunktion nach den Gewichten der Verbindungen zwischen verborgener Ebene und Ausgabebene w'_{kj} und den Schwellenwerten α_k der Ausgabeknoten haben die Form:

$$\frac{\partial E}{\partial w_{kj}} = \frac{1}{N_{\text{training}}} \sum_n \delta_k^n \cdot V_j^n \quad (7.14)$$

$$\frac{\partial E}{\partial \alpha_k} = \frac{1}{N_{\text{training}}} \sum_n \delta_k^n \quad (7.15)$$

δ_k^n ist dabei als

$$\delta_k^n = \frac{(A_k^{n,\text{Netz}} - A_k^{n,\text{erw\u00fcnscht}})}{t} \cdot g'(x_k^n) \quad (7.16)$$

definiert und x_k^n ist der Eingabewert des Ausgabeknotens A_k (vgl. Tabelle 7.1). $g'(x)$ ist die partielle Ableitung der Transferfunktion $g(x)$ nach ihrem Argument x . F\u00fcr Ableitungen nach den Schwellenwerten α_j der verborgenen Knoten und den Gewichten w_{ji} der Verbindungen zwischen den Knoten der verborgenen Ebene und der Eingabeebene ergeben sich die Gleichungen:

$$\frac{\partial E}{\partial w_{ji}} = \frac{1}{N_{\text{training}}} \sum_n \delta_j^n \cdot M_i^n \quad (7.17)$$

$$\frac{\partial E}{\partial \alpha_j} = \frac{1}{N_{\text{training}}} \sum_n \delta_j^n \quad (7.18)$$

Dabei ist x_j^n der Eingabewert des verborgenen Knotens V_j und δ_j^n definiert als:

$$\delta_j^n = \sum_k \frac{(A_k^{n,\text{Netz}} - A_k^{n,\text{erw\u00fcnscht}})}{t} \cdot g'(x_k^n) \cdot w_{kj} \cdot g'(x_j^n) \quad (7.19)$$

Die δ_j^n werden nach (7.20) aus den gewichteten Summen der δ_k^n berechnet. Die Gewichtung erfolgt mit denselben Gewichten w_{kj} , wie bei der Berechnung der Ausgabewerte des Netzwerkes. Die Richtung des Datenaustausches ist jedoch genau entgegengesetzt, daher die Bezeichnung *back propagation*.

$$\delta_j^n = g'(x_j^n) \cdot \sum_k w_{kj} \cdot \delta_k^n \quad (7.20)$$

Sind die partiellen Ableitungen der Fehlerfunktion (7.6) bekannt, so k\u00f6nnen die \u00c4nderungen der Gewichte und Schwellenwerte nach (7.7) und (7.8) berechnet werden. Die \u00c4nderungen der Temperatur t bzw. der Schrittweite η des abgewandelten *feed forward*-Netzwerkes sind nicht in den *back propagation*-Algorithmus eingebunden und m\u00fcssen separat berechnet werden (vgl. (7.10), (7.11) und (7.13)). F\u00fcr ein *feed forward*-Netzwerk mit einer beliebigen Anzahl verborgener Ebenen kann Gleichung (7.20) entsprechend verallgemeinert werden.

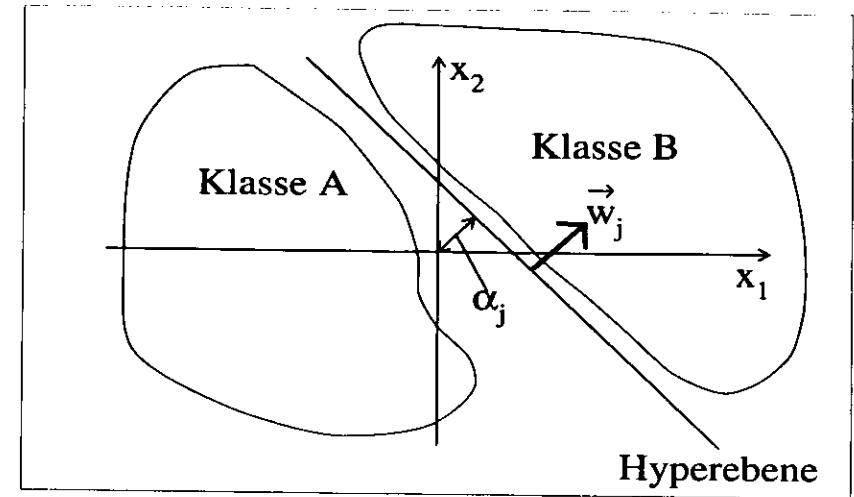


Abbildung 7.6: Geometrische Deutung der Funktionsweise eines Knotens. Jeder Knoten H_j definiert eine Hyperebene im N -dimensionalen Raum seiner N Eingabeinformationen. Der Vektor \vec{w}_j , der Gewichte der Eingabeinformationen definiert die Lage der Hyperebene. Der Schwellenwert α_j ist der euklidische Abstand der Hyperebene zum Ursprung.

7.7 Interpretation der Ausgabewerte

Zur Interpretation der Ausgabewerte des Netzwerkes wird zun\u00e4chst die Funktionsweise eines einzelnen Knotens betrachtet. Anschließend werden die Ausgabewerte des Netzwerkes interpretiert. Die Gewichte w_{ji} eines Knotens H_j definieren die Lage einer Hyperebene im N -dimensionalen Raum seiner Eingabeinformationen x_i ($i=1..N$). Sind die Gewichte nach Gleichung (7.9) auf 1 normiert, so ist der Vektor \vec{w}_j der Gewichte w_{ji} des Knotens der Normalvektor der Hyperebene. Der Schwellenwert α_j des Knotens ist der euklidische Abstand der Hyperebene zum Ursprung (vgl. Abb 7.6). Im einfachsten Fall ist die Transferfunktion des Knotens die Stufenfunktion, so da\u00df allen Punkten auf der einen Seite der Hyperebene der Ausgabewert +1 und allen Punkten auf der anderen Seite der Ausgabewert 0 zugeordnet wird. Bei einer sigmoiden Transferfunktion entsteht beidseitig der Hyperebene eine \u00dcbergangszone. Die Ausgabewerte von Punkten in gro\u00dfem Abstand zur Hyperebene sind noch immer nahe bei 0 bzw. 1, w\u00e4hrend nahe Punkte durch die Transferfunktion Ausgabewerte zwischen 0 und 1 erhalten. Betrachtet man eine Folge von Eingabemustern, die senkrecht auf der Hyperebene steht, so folgen die berechneten Ausgabewerte dieser Eingabemuster dem Verlauf der verwendeten Transferfunktion. Jeder Knoten der ersten verborgenen Ebene definiert eine Hyperebene im Raum der Eingabeinformationen des Netzwerkes. Abstand und Lage der Eingabemuster zur Hyperebene bestimmen den Ausgabewert des Knotens. Die Knoten der folgenden Ebene kombinieren diese Ausgabewerte, indem sie in dem durch sie definierten Raum ihrerseits eine Hyperebene konstruieren. Das Verfahren wiederholt sich in den folgenden Ebenen, bis die Ausgabewerte des Netzwerkes von den einzelnen Ausgabeknoten bestimmt werden. Die

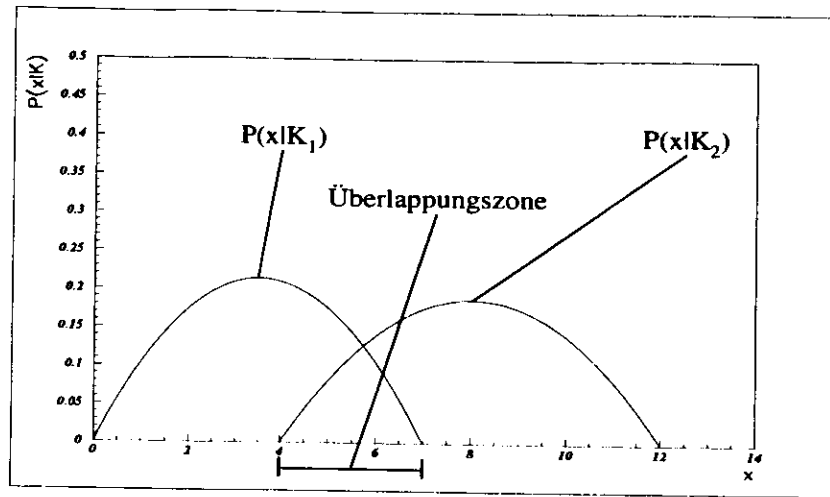


Abbildung 7.7: Beispiel zum Bayes'schen Klassifizierer. $P(x|K_1)$ bzw. $P(x|K_2)$ sind die Wahrscheinlichkeitsverteilungen der Muster x innerhalb der Musterklassen K_1 bzw. K_2 . Die Muster x aus der Überlappungszone müssen statistisch einer Klasse zugeordnet werden. Die geringste Fehlzuordnung wird durch Anwendung des Bayes'schen Klassifizierers erreicht.

Ausgabewerte eines Eingabemusters sind also ein Maß für die Lage des Musters im Raum der Eingabeinformationen, relativ zu den durch die verborgenen Knoten der ersten Ebene definierten Hyperebenen. Durch das Training des Netzwerkes werden die Hyperebenen so angeordnet, daß eine bestmögliche Unterscheidung der verschiedenen Musterklassen gegeben ist. Schnitte in der Verteilung dieser Ausgabewerte bestimmen die Effizienz der Identifikation der verschiedenen Musterklassen.

7.8 Grenzen der Klassifikation

Auch wenn durch das iterative Training des Netzwerkes das globale Minimum der Fehlerfunktion (7.6) erreicht wird, können trotzdem in vielen Fällen nicht alle Eingabemuster wie gewünscht klassifiziert werden. Die Effizienz eines neuronalen Netzwerkes bei der Klassifikation verschiedener Musterklassen wird durch das sogenannte Bayes'sche Limit begrenzt. Überlappen sich die verschiedenen Musterklassen, so sind bei Mustern aus dem Überlappungsbereich nur statistische Aussagen über ihre Zugehörigkeit zu den verschiedenen Musterklassen möglich. In Abb. 7.7 ist dieses am Beispiel zweier Wahrscheinlichkeitsverteilungen dargestellt. $P(x|K_1)$ bzw. $P(x|K_2)$ sind die Wahrscheinlichkeitsverteilungen der Muster x innerhalb der Klassen K_1 bzw. K_2 . Ein Muster aus dem Überlappungsbereich kann nur über statistische Methoden einer dieser beiden Klassen zugeordnet werden. Die geringste Fehlklassifikation wird durch Anwendung des Bayes'schen Klassifizierers [27] erreicht.

Sind n verschiedene Klassen K_i ($i=1..n$) gegeben, so ist er definiert als:

$$P(K_i|x) = \frac{P(K_i) \cdot P(x|K_i)}{\sum_{j=1}^n P(K_j) \cdot P(x|K_j)} \quad (7.21)$$

Dabei ist $P(K_i|x)$ die a priori Wahrscheinlichkeit, daß das Muster x der Klasse K_i angehört, $P(K_i)$ die a priori Wahrscheinlichkeit der Klassen K_i und $P(x|K_i)$ die Wahrscheinlichkeitsverteilungen der Muster x innerhalb der einzelnen Klassen K_i . Sind $P(K_i)$ und $P(x|K_i)$ für alle Musterklassen K_i bekannt, so kann die Klasse der höchsten a priori Wahrscheinlichkeit $P(K_i|x)$ für ein gegebenes Muster x bestimmt werden. Im allgemeinen sind $P(K_i)$ und $P(x|K_i)$ jedoch nicht für alle Musterklassen bekannt, so daß Annahmen über diese Wahrscheinlichkeiten gemacht werden müssen. Es kann gezeigt werden, daß neuronale Netzwerke mit quadratischer Fehlerfunktion und auf das Intervall $[0, 1]$ normierten Ausgabewerten eine Approximation des optimalen Bayes'schen Klassifizierers bestimmen, ohne Annahmen über die Wahrscheinlichkeitsverteilungen der einzelnen Musterklassen zu machen [28].

7.9 Auswertung der Ergebnisse des Netzwerkes

In dieser Analyse werden feed forward-Netzwerke zur Klassifikation von Eingabemustern aus zwei Musterklassen K_1 und K_2 verwendet. Daher wurde nur ein Ausgabeknoten mit den gewünschten Ausgabewerten von +1 bzw. -1 für Muster der Klassen K_1 bzw. K_2 verwendet. Die Ausgabewerte $A_1^{n, Netz}$ der verschiedenen Eingabemuster M^n wurden über die Beziehung

$$P^n = P(K_1|M^n) = \frac{1}{2} \cdot (A_1^{n, Netz} + 1) \quad (7.22)$$

zur Berechnung der durch die Netzwerke approximierten Bayes'schen Wahrscheinlichkeit P^n verwendet. P^n ist dabei die Wahrscheinlichkeit, daß das Muster M^n aus der Musterklasse K_1 stammt. Damit sind die gewünschten Werte von P^n 1 für Muster der Klasse K_1 und 0 für Muster der Klasse K_2 .

Die Identifikation der Eingabemuster erfolgt durch einen Schnitt in der Verteilung von P^n für alle Eingabemuster. Eingabemuster mit einer Wahrscheinlichkeit P^n größer als der Schnittwert Δ werden als Muster der Klasse K_1 klassifiziert, alle anderen als Muster der Klasse K_2 . Je nach der Lage des Schnittwertes ergibt sich eine unterschiedliche Anzahl von fehlklassifizierten Mustern. Zum Vergleich der verschiedenen Netzwerke werden zwei Größen in Abhängigkeit vom Schnittwert verwendet: Effektivität ε und Reinheit R_1 . Sie sind folgendermaßen definiert:

$$\varepsilon = \frac{N(P^n(K_1) > \Delta)}{N(K_1)} \quad (7.23)$$

$$R_1 = \frac{N(P^n(K_1) > \Delta)}{N(P^n(K_1) > \Delta) + N(P^n(K_2) > \Delta)} \quad (7.24)$$

Dabei sind $N(P^n(K_1) > \Delta)$ bzw. $N(P^n(K_2) > \Delta)$ die Anzahl der Muster der Klassen K_1 bzw. K_2 , deren Wahrscheinlichkeit P^n größer als der Schnittwert Δ ist. $N(K_1)$ ist die Gesamtzahl der Muster der Klasse K_1 . Die Reinheit hängt im Gegensatz zur Effektivität vom Verhältnis $N(K_1) : N(K_2)$ ab. Ihr Index bezeichnet dieses Verhältnis. Bei allen in dieser Analyse verwendeten Datensätzen betrug das Verhältnis 1:1. Mit diesem Wert werden Probleme vermieden, die auftreten, wenn die Anzahl der Muster einer Musterklasse sehr viel größer als die der anderen Musterklasse ist. In diesem Fall wird die Lage der Hyperbenen der neuronalen Netzwerke durch die Minimierung der Fehlklassifikation der Muster der überwiegenden Klasse dominiert. Die Muster der anderen Klasse haben nur wenig Einfluß auf die Lage der Hyperbenen und werden daher zum großen Teil fehlklassifiziert. Die Klassifikation der Muster der beiden Musterklassen mit Hilfe der Ausgabewerte der neuronalen Netzwerke ist daher nur eingeschränkt möglich.

Für ein gegebenes Problem muß entschieden werden, ob einer hohen Reinheit oder Effektivität Priorität gegeben werden soll.

7.10 Kontrolle des Trainings

Während des Trainings werden die Parameter der Netzwerke, wie in den Kapiteln 7.4 bis 7.6 beschrieben, iterativ verändert, um das globale Minimum der Fehlerfunktion (7.6) zu finden. Die dabei verwendeten Eingabemuster bilden den Trainingsdatensatz. Da dieser Datensatz nur eine endliche Zahl von Eingabemustern enthält, besteht die Möglichkeit, daß das neuronale Netzwerk nicht nur auf die Eigenschaften der zu unterscheidenden Musterklassen, sondern auch auf spezielle, nur dem Trainingsdatensatz eigene Merkmale, trainiert wird. Um dieses sogenannte Übertraining zu vermeiden, werden die für das Training verfügbaren Eingabemuster in zwei unabhängige Datensätze aufgeteilt, die in ihren globalen Eigenschaften, wie z.B. Anzahl der Eingabemuster der verschiedenen Musterklassen, übereinstimmen müssen. Während der Trainingsdatensatz wie beschrieben zur Minimierung der Fehlerfunktion (7.6) verwendet wird, hat der Testdatensatz keinen Einfluß auf das Training des Netzwerkes. Zum Vergleich der Ausgabewerte des Netzwerkes für beide Datensätze wird eine neue Größe, die sogenannte Mißidentifikationsrate f eingeführt. Sie ist für einen Datensatz definiert als:

$$f = \frac{N(P^n(K_1) < 0.5) + N(P^n(K_2) > 0.5)}{N(K_1) + N(K_2)} \quad (7.25)$$

Analog zu Gleichung (7.24) sind $N(P^n(K_1) < 0.5)$ und $N(P^n(K_2) > 0.5)$ die Anzahl der Muster aus den Klassen K_1 bzw. K_2 , deren Wahrscheinlichkeiten P^n kleiner, bzw. größer als 0.5 sind. $N(K_1)$ bzw. $N(K_2)$ ist die Gesamtanzahl der Eingabemuster des Datensatzes aus den Klassen K_1 bzw. K_2 .

Zur Kontrolle auf Übertraining wird die Differenz $f_{relativ}$ der Mißidentifikationsraten f_{train} und f_{test} des Trainings- bzw. Testdatensatzes verwendet. $f_{relativ}$ ist definiert als:

$$f_{relativ} = f_{test} - f_{train} \quad (7.26)$$

In den ersten Trainingsiterationen treten aufgrund des iterativen Verfahrens des Trainings und der endlichen Anzahl der Muster in den beiden Datensätzen Schwankungen in $f_{relativ}$

auf. Findet kein Übertraining statt, so klingen diese Schwankungen mit steigender Iterationszahl ab, und der Grenzwert von $f_{relativ}$ für große Iterationen geht gegen Null. Findet Übertraining statt, so beginnt $f_{relativ}$ wieder anzusteigen, da in diesem Fall f_{train} kleiner als f_{test} ist. Zur Identifikation des Übertrainings wird die Standardabweichung $\sigma_{f_{relativ}}$ abgeschätzt. Ist die Anzahl der fehlidentifizierten Muster des Trainings- bzw. Testdatensatzes klein gegen die Gesamtzahl der Muster, so sind σ_{train} bzw. σ_{test} gute Näherungen für die Standardabweichungen von f_{train} bzw. f_{test} .

$$\sigma_{train} = \frac{\sqrt{N_{train}}}{N} = \sqrt{\frac{f_{train}}{N}} \quad (7.27)$$

$$\sigma_{test} = \frac{\sqrt{N_{test}}}{N} = \sqrt{\frac{f_{test}}{N}} \quad (7.28)$$

Dabei ist N_{train} bzw. N_{test} die Anzahl der fehlidentifizierten Muster des Trainings- bzw. Testdatensatzes und N die Gesamtzahl der Muster eines Datensatzes. Findet kein Übertraining statt, so gilt $N_{train} \approx N_{test}$ und $\sigma_{relativ}$ ergibt sich zu

$$\sigma_{relativ} = \sqrt{\frac{f_{train} + f_{test}}{N}} \approx \sqrt{\frac{2 \cdot f_{train}}{N}} \approx \sqrt{\frac{2 \cdot f_{test}}{N}} \quad (7.29)$$

Findet Übertraining statt, so ist N_{test} größer als N_{train} und $\sigma_{relativ}$ übersteigt den in Gleichung (7.29) abgeschätzten Wert. Zur Identifikation des Übertrainings wird daher

$$\sigma_{relativ} = \sqrt{\frac{2 \cdot f_{train}}{N}} \quad (7.30)$$

verwendet. $\sigma_{relativ}$ wird für jede Trainingsiteration berechnet. Ist $f_{relativ}$ im Training ab einer bestimmten Iteration permanent größer als das berechnete $\sigma_{relativ}$, so findet mit hoher Wahrscheinlichkeit Übertraining statt. Durch die iterative Natur des Trainings besteht die Möglichkeit, daß $f_{relativ}$ während der Iterationen, in denen sich f_{train} und f_{test} stark verändern, ansteigt. Dabei kann die Grenze von $\sigma_{relativ}$ für diese Iterationen überschritten werden, ohne daß Übertraining stattfindet. Zur Identifikation des Übertrainings muß stets der Gesamtverlauf von $f_{relativ}$ betrachtet werden. Bei der Entscheidung, nach welcher Anzahl von Iterationen das Training abgebrochen wird, muß außerdem noch der Verlauf von f_{test} berücksichtigt werden. Für ein Problem ist es unter Umständen besser, Übertraining in einem gewissen Rahmen zuzulassen, wenn auch die Mißidentifikationsrate des Testdatensatzes f_{test} weiter verringert werden kann.

Da die Gewichte und Schwellenwerte der neuronalen Netzwerke zufällig in einem gegebenen Intervall initialisiert werden, sind die Werte von t , f_{train} und f_{test} in den ersten Iterationen des Trainings für verschiedene Initialisierungen unterschiedlich. Wenn die Minimierung der Fehlerfunktion zum globalen Minimum führt, wird erwartet, daß t und f nach genügend vielen Iterationen unabhängig von der Anfangsinitialisierung des Netzwerkes werden. Der Vergleich der Änderungen von Temperatur und Mißidentifikationsrate während des Trainings für verschiedene Initialisierungen eines Netzwerkes gibt

Aufschluß über die Güte der Minimierung der Fehlerfunktion. Das Training der verschiedenen neuronalen Netzwerke wurde in dieser Analyse stets mehrmals mit verschiedenen Initialisierungen durchgeführt. Für jede Trainingsiteration wurden die Mittelwerte und Standardabweichungen von t , f_{train} , f_{test} und $f_{relativ}$ nach dem Fehlerfortpflanzungsgesetz berechnet [29].

7.11 Beispiel eines Klassifikationsproblems

Ein multidimensionales Klassifikationsproblem besteht darin, in einem mehrdimensionalen Eingaberaum verschiedene Musterklassen durch geeignete Schnitte zu separieren. Die Dimension des Eingaberaumes wird dabei durch die Anzahl der zur Verfügung stehenden Informationen bestimmt. Überlappen die verschiedenen Muster nicht, so ist eine vollständige Trennung möglich. In Abb. 7.8 I ist ein derartiges Problem in zwei Dimensionen dargestellt. Alle Punkte (X_1, X_2) innerhalb des eingezeichneten Dreiecks gehören der Musterklasse A an, alle Punkte außerhalb der Musterklasse B. Die beiden Musterklassen überschneiden sich nicht, sind jedoch durch einfache Projektion auf eine der beiden Achsen nicht zu separieren.

Zur Trennung der beiden Musterklassen wurden abgewandelte feed forward-Netzwerke verwendet. Dazu wurden als Eingabeinformationen jeweils die Koordinaten X_1 und X_2 der einzelnen Punkte verwendet. Trainings- und Testdatensatz bestanden aus jeweils 2000 Punkten jeder Musterklasse. Pro Trainingsiteration wurden alle 4000 Punkte zur Berechnung der Korrekturen der Parameter des Netzwerkes verwendet. Mit den im Training bestimmten Parametern der Netzwerke wurden die Ausgabewerte aller Punkte des Testdatensatzes berechnet. Die Wahrscheinlichkeit $P(X_1, X_2)$ jedes dieser Punkte, zur Musterklasse A zu gehören, wurde anschließend nach Gleichung (7.22) bestimmt.

Zwei verborgene Knoten waren zur Trennung der beiden Musterklassen ausreichend. In Abb. 7.8 II und III sind für neuronale Netzwerke mit 2 und 3 verborgenen Knoten alle Punkte des Testdatensatzes eingezeichnet, deren Wahrscheinlichkeit $P(X_1, X_2)$ größer als 0.5 war. Außerdem wurden die Lage der zur Trennung der beiden Klassen verwendeten Hyperebenen (in diesem Fall Geraden) eingezeichnet. Mit zwei verborgenen Knoten konnten die beiden Musterklassen nach 2000 Trainingsiterationen nahezu vollständig getrennt werden (siehe Abb 7.8 II). Schon nach 1500 Iterationen wurde die Lage der Hyperebenen pro Trainingsiteration nur noch geringfügig verändert, da nur wenige Punkte in der Nähe des Kreuzungspunktes der beiden Hyperebenen fehlidentifiziert wurden. Zur Beschleunigung des Trainings in einem solchen Fall ist es ratsam, neue Trainings- und Testdatensätze aus dem Bereich der noch fehlidentifizierten Punkte zu konstruieren, um mit diesen das Training fortzusetzen. Im Rahmen dieses Beispiels wurde sich auf 2000 Iterationen beschränkt. Eine weitere Erhöhung der Anzahl der verborgenen Knoten verbesserte die Trennung der beiden Musterklassen nicht weiter (vgl. Abb 7.8 III). Obwohl die dritte Hyperebene in der Nähe der meisten fehlidentifizierten Punkte verlief, wurden die beiden Musterklassen nicht besser als durch das Netzwerk mit 2 verborgenen Knoten getrennt. Auch in diesem Fall hätte die Trennung der beiden Musterklassen durch weiteres Training mit neuen Datensätzen noch verbessert werden können.

Das Netzwerk wurde mit denselben Datensätzen bei fünf verschiedenen Initialisierungen trainiert. Die für jede Iteration berechneten Werte der Temperatur t , der Mißidentifikationsraten f_{train} und f_{test} sowie der relativen Mißidentifikationsrate $f_{relativ}$ wurden aus den

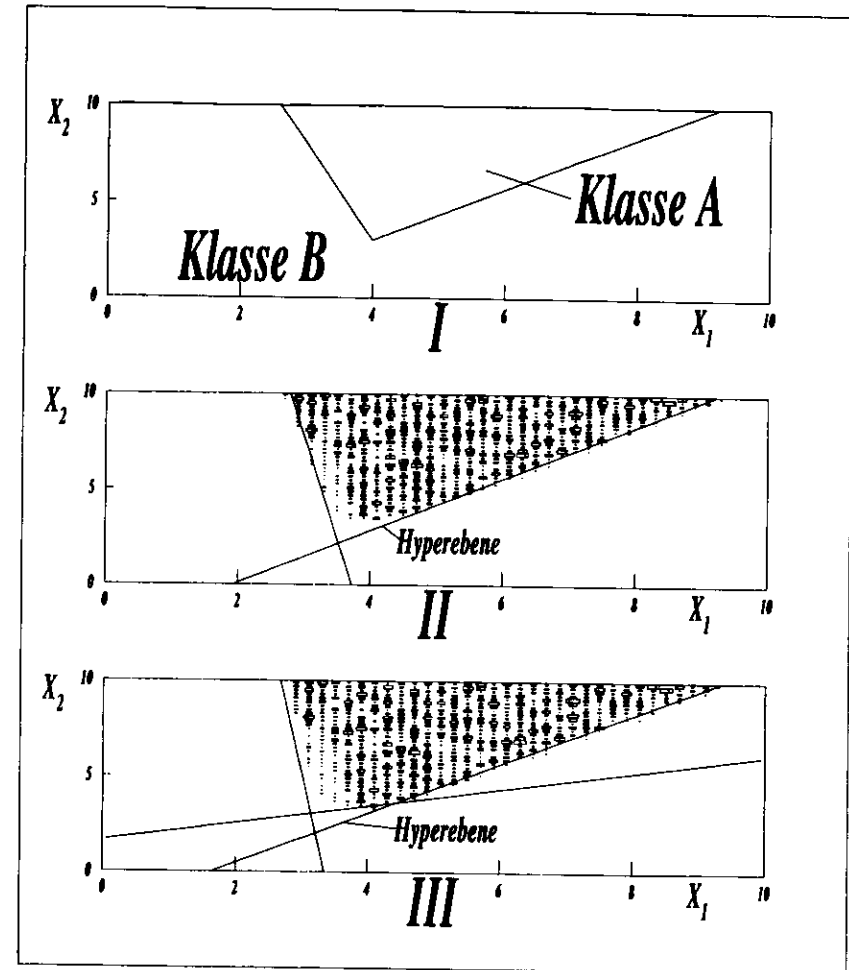


Abbildung 7.8: Bild I zeigt die Verteilung der beiden Musterklassen A und B. Dabei beinhaltet die Musterklasse A alle Punkte innerhalb des eingezeichneten Dreiecks. Alle anderen Punkte der Fläche gehören zur Musterklasse B. Jeder Punkt (X_1, X_2) der dargestellten Fläche soll entweder der Musterklasse A oder der Musterklasse B zugeordnet werden. Bild II bzw. III zeigt für Netzwerke mit zwei bzw. drei verborgenen Knoten alle Punkte des Testdatensatzes mit $P(X_1, X_2)$ größer 0.5. Außerdem ist die Lage der zur Trennung der beiden Musterklassen verwendeten Hyperebenen beider Netzwerke eingezeichnet.

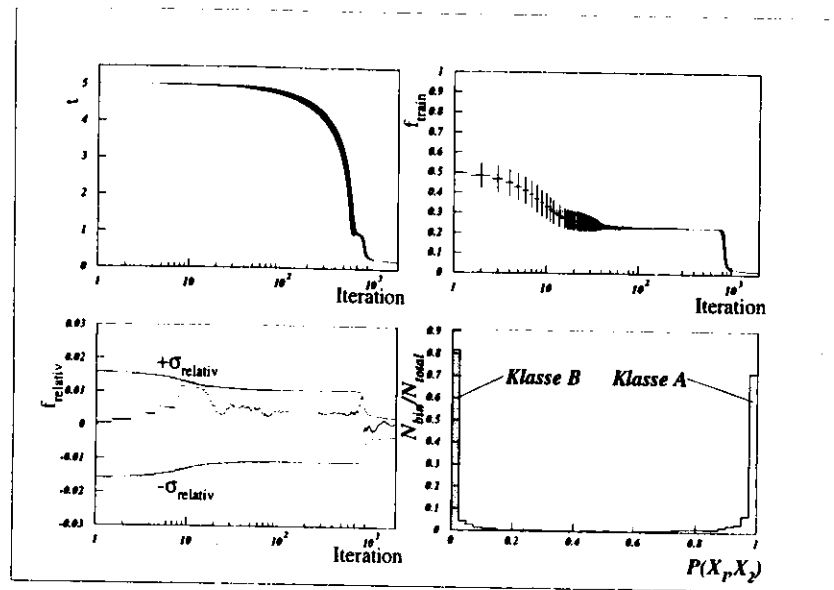


Abbildung 7.9: Trainingsverlauf beim zweidimensionalen Klassifikationsproblem bei Verwendung von zwei verborgenen Knoten. Für die einzelnen Iterationen sind die Mittelwerte der Temperatur t , der Mißidentifikationsrate f_{train} und die relative Mißidentifikationsrate f_{relativ} dargestellt. Die Fehlerbalken dieser Größen entsprechen ihren Standardabweichungen. Bei der Darstellung von f_{relativ} wird aus Gründen der Übersicht auf die Darstellung des Fehlers verzichtet. Außerdem ist die Wahrscheinlichkeitsverteilung $P(X_1, X_2)$ der einzelnen Punkte des Testdatensatzes dargestellt.

fünf Trainingsphasen gemittelt. Der Trainingsverlauf und die berechnete Wahrscheinlichkeitsverteilung der Muster des Testdatensatzes für das Netzwerk mit zwei verborgenen Knoten sind in Abb. 7.9 dargestellt. Die Fehler entsprechen den Standardabweichungen der berechneten Mittelwerte. Der Fehler sowohl von t als auch von f_{train} waren nach 1000 Iterationen vernachlässigbar klein. Nach den 2000 Trainingsiterationen betrug der Wert von f_{train} 0,0159. Zur Identifikation von Übertraining wurde σ_{relativ} nach Gleichung (7.30) berechnet. Zwischen 10 und 20 Iterationen ist f_{relativ} größer als σ_{relativ} . Da f_{relativ} im weiteren Verlauf des Trainings jedoch stets innerhalb von $\pm\sigma_{\text{relativ}}$ war, konnte Übertraining mit hoher Wahrscheinlichkeit ausgeschlossen werden. Die Standardabweichung von f_{relativ} ist nach 1300 Trainingsiterationen um mehr als eine Größenordnung kleiner als f_{relativ} und damit vernachlässigbar klein. Den Punkten der beiden Musterklassen A und B wurden nach 2000 Iterationen Wahrscheinlichkeiten nahe den gewünschten Werten von 1 und 0 zugeordnet.

Kapitel 8

Datenauswahl

8.1 Monte Carlo-Simulation

Zum überwachten Training der *feed forward*-Netzwerke war es notwendig, jedes Island des Trainings- und Testdatensatzes einer der beiden Musterklassen zuzuordnen. Zur Konstruktion der beiden Datensätze wurden daher Monte Carlo-Ereignisse verwendet. Damit konnte die Identität der für die Energiedeposition in den einzelnen Islands verantwortlichen Teilchen und daraus die Musterklasse der Islands bestimmt werden. Es wurden Monte Carlo-Ereignisse der tiefunelastischen Positron-Proton-Streuung verwendet. Die Simulation des ZEUS-Detektors in der Konfiguration von 1994 wurde mit dem Programmpaket MOZART Version 12.1 durchgeführt.

Da zur Zeit der Erstellung dieser Arbeit nur der RHES installiert und in Betrieb war, wurde die Analyse auf Daten des RCAL bzw. RHES beschränkt. Den einzelnen Islands im RCAL wurden nach einem in Kapitel 8.3.2 beschriebenen Verfahren Dioden des HES zugeordnet, die durch den Isle-Algorithmus zu verschiedenen Isles zusammengefaßt wurden. Aus den Informationen der Islands und den zugeordneten Isles wurden die Eingabemuster der neuronalen Netzwerke gewonnen. Die Musterklasse K_1 enthielt die Islands, die vom Zerfall eines π^0 -Mesons stammen, die Musterklasse K_2 alle anderen. Im folgenden werden die Muster der Klasse K_1 als „Signal“, die der Klasse K_2 als „Untergrund“ bezeichnet. Die genauen Definitionen dieser beiden Klassen unterscheiden sich bei der Identifikation der isolierten und nicht-isolierten π^0 -Mesonen und werden in den Kapiteln 9 und 10 gegeben.

8.2 Vorauswahl der Ereignisse

Viele Teilchen, die unter kleinen Winkeln gegen die Strahlachse ins RCAL eintreten, deponieren nur einen Teil ihrer Energie im Kalorimeter, während der Rest der Energie im Strahlrohr nicht nachgewiesen werden kann. Um systematische Fehler zu vermeiden, wurden alle Islands, deren Zentrum innerhalb eines Bereiches von $24 \times 24 \text{ cm}^2$ um das Strahlrohr lag, von der Analyse ausgeschlossen. Dieser Schnitt wird als „Boxcut“ bezeichnet. Zur Positionsbestimmung wurde für alle Islands eine Ortsrekonstruktion mit

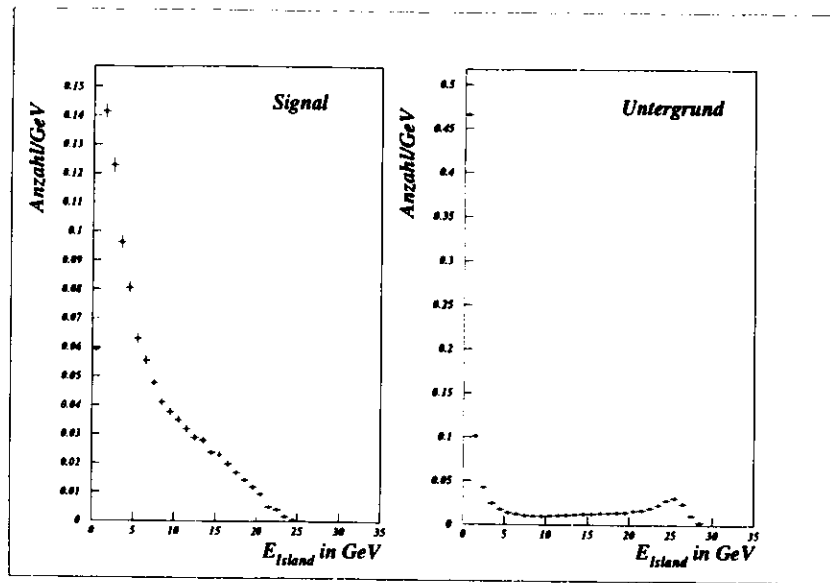


Abbildung 8.1: Energieverteilung der Signal- und Untergrundislands nach Anwendung der Boxcuts. Das linke Bild zeigt die Signalislands, das rechte die Untergrundislands. Beide Verteilungen sind auf 1 normiert.

einer linearen Energiegewichtung der Energiedeposition in den einzelnen Kalorimeterzellen durchgeführt.

Um diejenigen Untergrundeignisse von der Analyse auszuschließen, die einfach als solche zu identifizieren sind, wurden verschiedene Selektionskriterien verwendet. Dabei werden diejenigen Islands als Signal bezeichnet, in denen beide Photonen mindestens eines π^0 -Zerfalls Energie deponiert haben. Alle anderen Islands werden als Untergrund bezeichnet. Deponieren die beiden Photonen eines π^0 -Zerfalls ihre Energie in verschiedenen Islands, so werden diese beiden Islands ebenfalls als Untergrund klassifiziert. Eine Übersicht der einzelnen Kriterien sowie der Prozentsatz der nach ihrer Anwendung verbleibenden Untergrund- und Signalmuster ist in Tabelle 8.1 aufgezeigt. Im folgenden werden die einzelnen Auswahlkriterien kurz erläutert.

In Abb. 8.1 ist die Verteilung der im CAL gemessenen Energie E_{Island} der Untergrund-

| Kriterium | Verbleibende Untergrundislands in % | Verbleibende Signalislands in % |
|---|-------------------------------------|---------------------------------|
| $E_{\text{Island}} \leq 20 \text{ GeV}$ | 95 | 99 |
| HES-Signal hinter Island | 55 | 97 |

Tabelle 8.1: Vorauswahlkriterien und ihre Auswirkungen auf Signal- und Untergrundislands

und Signalislands dargestellt. Die Zusammensetzung des Untergrundes verändert sich mit steigender Energie. Oberhalb von 10 GeV wird der Untergrund von den Positronen der tiefunelastischen Streuung dominiert, unterhalb von 5 GeV dagegen von Hadronen. Da sich die Signalislands fast ausschließlich unterhalb von 20 GeV befinden, wurden alle Islands mit E_{Island} größer als 20 GeV ausgeschlossen.

Jedem Island wurden durch ein in Kapitel 8.3.2 genauer beschriebenes Verfahren Dioden des HES zugeordnet. Eine weitere Reduktion des Untergrundes konnte erreicht werden, indem alle diejenigen Islands von der Analyse ausgeschlossen wurden, bei denen keine der zugeordneten Dioden eine Energiedeposition von mehr als 0,6 mip aufwies. Durch dieses Kriterium wurden 40% der Untergrundislands ausgeschlossen, jedoch nur 2% der Signalislands.

Das Verhältnis V von Signal zu Untergrund nach Anwendung der Vorauswahlkriterien für verschiedene Energien ist in Abb. 8.2 dargestellt. Im gesamten Energiebereich ist der Untergrund größer als das zu identifizierende Signal. Zwischen 6 und 10 GeV ist das Verhältnis von Signal zu Untergrund am größten und beträgt etwa 0,36. Außerhalb dieses Energiebereiches steigt der Untergrund im Verhältnis zum π^0 -Signal stark an.

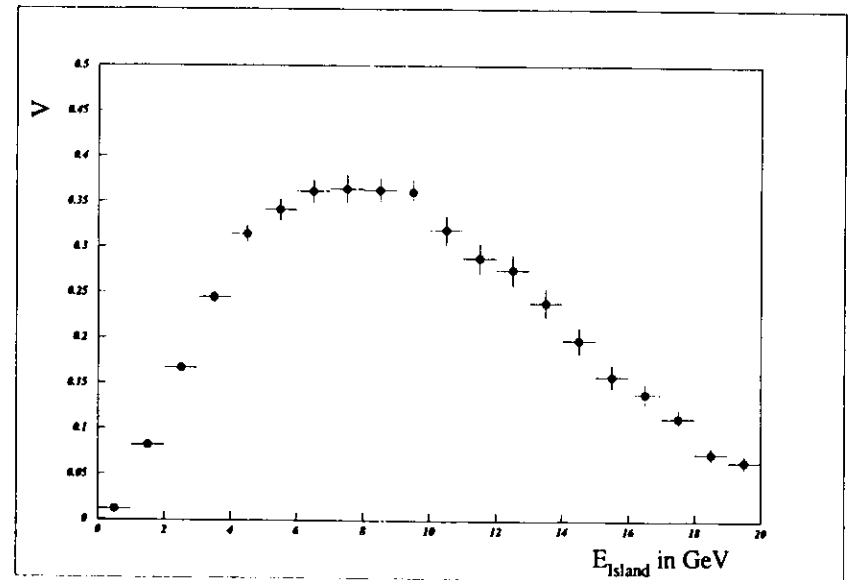


Abbildung 8.2: Verhältnis V von Signal zu Untergrund nach Anwendung der Vorauswahlkriterien.

8.3 Konstruktion der Eingabeinformationen

8.3.1 Informationen des CAL

Die in dieser Analyse für die einzelnen Islands verwendeten Kalorimeterinformationen sind identisch mit den 55 Informationen, die beim Elektronenfinder SINISTRA [30] verwendet werden. Die ersten 54 Eingabeinformationen sind die Meßwerte der einzelnen Photovervielfacher der Kalorimeterzellen. Für ein gegebenes Island werden nur die Photovervielfacher des Turmes der maximalen Energiedeposition und der acht angrenzenden Türme berücksichtigt. Für das RCAL sind die achtzehn EMC- und neun HAC-Zellen dieser neun Türme und ihre Photovervielfacher schematisch in Abb. 8.3 dargestellt. Als weitere Information wird der Kosinus des Einfallwinkels δ relativ zur Oberflächennormalen des Kalorimeters verwendet. Dabei wird angenommen, daß die für die Energiedeposition im Kalorimeter verantwortlichen Teilchen vom Wechselwirkungspunkt kommen. δ wird aus den Positionen des Zentrums des Islands und des rekonstruierten Wechselwirkungspunktes berechnet.

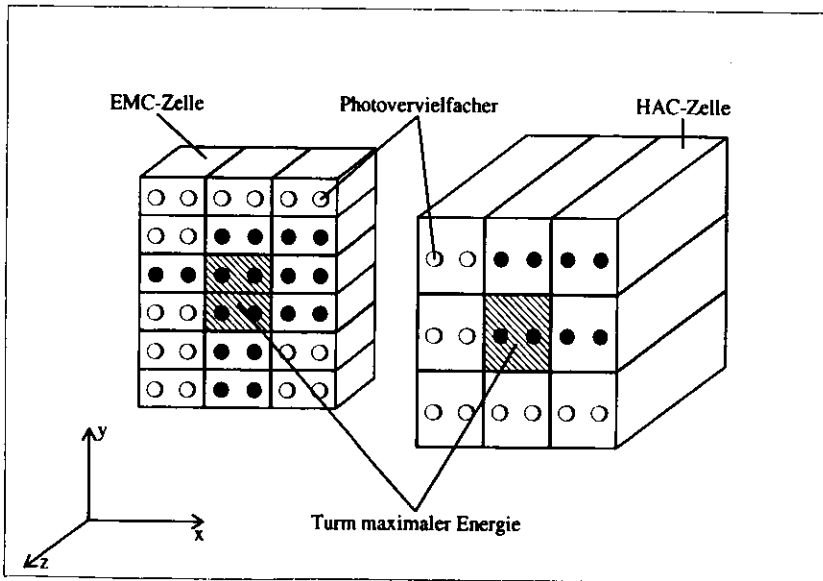


Abbildung 8.3: Schematische Darstellung der Konstruktion der 54 Kalorimeterinformationen für ein gegebenes Island im RCAL. Die verschiedenen Schattierungen deuten unterschiedliche Energiedepositionen an.

8.3.2 Informationen des HES

Zur Berechnung der Eingabeinformationen des HES für ein gegebenes Island werden dem Island zunächst Dioden des HES zugeordnet. Dazu wurden alle HES-Dioden ermittelt, die in den Kalorimetertürmen des Islands liegen. Durch den Aufbau des HES ist es bedingt, daß die Grenze zwischen zwei Kalorimetertürmen in Y-Richtung nicht mit der Grenze zwischen zwei Dioden identisch ist. Sechs Dioden liegen jeweils auf dieser Grenze zwischen zwei Kalorimetertürmen und können daher nicht eindeutig einem Kalorimeterturm zugeordnet werden. Diese Dioden wurden ebenfalls dem Island zugeordnet, wenn mindestens einer der beiden Kalorimetertürme zum Island gehörte. Aus allen zugeordneten Dioden wurden mit dem Isle-Algorithmus Isles konstruiert (vgl. Abb. 8.4). Zur Bestimmung der Position der Isles wurde eine logarithmische Energiegewichtung der Energiedeposition in den einzelnen Dioden verwendet. Aus den so konstruierten Isles wurden die Eingabeinformationen des neuronalen Netzwerkes gewonnen. Es wurden insgesamt 22 Eingabeinformationen verwendet. Deponieren beide Photonen eines π^0 -Zerfalls Energie in den dem Island zugeordneten Dioden, so erwartet man bei nicht zu kleinem Öffnungswinkel zwischen den Photonen zwei separate Isles. Weitere Isles sind durch Fluktuationen in den Teilchenschauern der beiden Photonen und Energiedepositionen anderer Teilchen gegeben. Die Anzahl der konstruierten Isles wurde daher als Eingabeinformation verwendet.

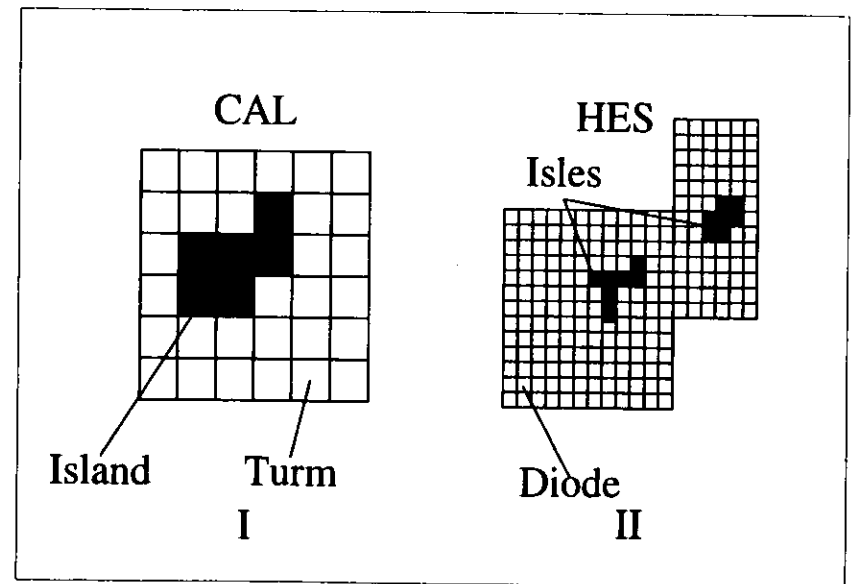


Abbildung 8.4: Schematische Darstellung der Konstruktion der HES-Isles für ein gegebenes Island. Bild I zeigt ein Schema der Kalorimetertürme mit den sechs Türmen eines Islands, Bild II die diesem Island zugeordneten HES-Dioden und die aus ihnen konstruierten Isles.

Neben dieser Information wurden nur Informationen der beiden Isles mit der höchsten Gesamtenergie verwendet. Diese beiden Isles werden im folgenden mit Isle 1 und Isle 2 bezeichnet. Für jede dieser beiden Isles wurden die Energiedepositionen in der Zentrallodiode und den acht umliegenden Dioden als Eingabeinformationen des Netzwerkes verwendet. Da im Durchschnitt 96% der Energie eines elektromagnetischen Teilchens in diesen neun Dioden enthalten ist [3], wurden die weiter entfernten Dioden größerer Isles vernachlässigt. Analog zu dem in Kapitel 8.3.1 beschriebenen Einfallwinkel δ wurden die Einfallwinkel δ_1 bzw. δ_2 relativ zur Oberflächennormalen des HES berechnet und als Eingabeinformationen verwendet. Der Abstand der beiden Isles in der Ebene des HES wurde ebenfalls als Eingabeinformation verwendet. Stammt die Energiedeposition der beiden Isles von den Photonen eines π^0 -Zerfalls, so ist dieser Abstand zusammen mit den Einfallwinkeln δ_1 und δ_2 ein Maß für den Öffnungswinkel zwischen den beiden Photonen. Die insgesamt 77 verwendeten Eingabeinformationen des CAL und HES für jedes Island sind in Tabelle 8.2 dargestellt.

| Nummer | Detektorkomponente | Bedeutung |
|--------|--------------------|---|
| 1-54 | CAL | Signale der Photovervielfacher der EMC- bzw. HAC-Zellen |
| 55 | CAL | $\cos(\delta)$ des Islands |
| 56-64 | HES | Signale der zentralen neun Dioden der Isle 1 |
| 65-73 | HES | Signale der zentralen neun Dioden der Isle 2 |
| 74 | HES | Abstand der beiden HES-Isles 1 und 2 in der Ebene des HES |
| 75 | HES | Anzahl der dem Island zugeordneten HES-Isles |
| 76 | HES | $\cos(\delta_1)$ |
| 77 | HES | $\cos(\delta_2)$ |

Tabelle 8.2: Eingabeinformationen der verwendeten neuronalen Netzwerke

8.4 Konstruktion der Trainings- und Testdatensätze

Die beiden unabhängigen Datensätze für die verschiedenen getesteten neuronalen Netzwerke bestanden aus je 4000 Eingabemustern, aufgeteilt in 2000 Untergrund- und 2000 Signalmuster. Ein Muster enthielt dabei die Informationen eines Islands und der ihm zugeordneten HES-Isles. Zur Untersuchung der Auswirkung der HES-Informationen auf die Effizienz der π^0 -Identifikation wurde eine Gruppe von Datensätzen nur mit Informationen des CAL und eine Gruppe mit Informationen des CAL und HES zusammengestellt. Für beide Gruppen wurden dieselben Islands verwendet. Um Aussagen über die Möglichkeiten der π^0 -Identifikation bei verschiedenen Energien machen zu können, wurden die zur Konstruktion der Datensätze verwendeten Islands nach ihrer Energie unterteilt. Die Breite der zur Unterteilung verwendeten Energiebereiche wurde dabei einheitlich mit 1 GeV festgelegt. Innerhalb der konstruierten Datensätze enthielt jeder Energiebereich die gleiche Anzahl von Islands. Das Verhältnis von Signal- und Untergrundislands wurde in allen Energiebereichen mit 1 gewählt. Die Energieverteilung innerhalb der einzelnen Energiebereiche entsprach dem Energiespektrum der verwendeten simulierten Ereignisse.

Die bei der Verwendung der abgewandelten *feed forward*-Netzwerke notwendige Normierung der Eingabeinformationen (vgl. Kapitel 7.5) wurde für die Signale der Photovervielfacher des CAL und die Dioden des HES separat durchgeführt; damit wurde sichergestellt, daß das Verhältnis der Signale zwischen den einzelnen Photovervielfachern bzw. Dioden erhalten blieb. Die Anzahl der Isles und der Abstand der beiden Isles mit der größten Energie wurden jeweils getrennt normiert. Da der Wertebereich der Kosinusfunktion zwischen -1 und +1 liegt, bestand keine Notwendigkeit, die Eingabeinformationen $\cos(\delta)$, $\cos(\delta_1)$ und $\cos(\delta_2)$ zu normieren.

Kapitel 9

Identifikation von isolierten π^0 -Mesonen

9.1 Konstruktion der Trainingsdaten

In diesem Kapitel wird ein Spezialfall der π^0 -Identifikation behandelt. An diesem vereinfachten Problem wurde das Trainingsverhalten der verwendeten *feed forward*-Netzwerke in Abhängigkeit der Werte der verwendeten Trainingsparameter studiert. Die so ermittelten Werte der einzelnen Trainingsparameter wurden bei der Untersuchung des allgemeinen Problems der π^0 -Identifikation verwendet (vgl. Kapitel 10). Außerdem wurden erste Aussagen über die Auswirkung der HES-Informationen auf die π^0 -Identifikation gemacht.

Bei der Konstruktion der Eingabeinformationen des CAL wurden nur die Signale der Photovervielfacher des Turmes der maximalen Energiedeposition und der acht benachbarten Türme berücksichtigt (vgl. Kapitel 8.3.1). Um die Auswirkungen dieses Verfahrens abzuschätzen, wird in diesem Kapitel untersucht, ob eine Abhängigkeit zwischen den Ausgabewerten der neuronalen Netzwerke und dem Anteil der Gesamtenergie der einzelnen Islands besteht, der in diesen neun Türmen deponiert wurde.

Die π^0 -Identifikation wurde vereinfacht, indem alle diejenigen Islands von der Analyse ausgeschlossen wurden, in denen Energie von den Photonen mindestens eines π^0 -Zerfalls und zusätzlich von anderen Teilchen deponiert wurde. Die verbleibenden Islands, deren Gesamtenergie von den Photonen eines π^0 -Zerfalls deponiert wurde, werden als Signalislands bezeichnet, alle anderen nichtausgeschlossenen Islands als Untergrundislands. Das Energiespektrum der aus den Monte Carlo-Ereignissen gewonnenen Signalislands ist in Abb. 9.1 dargestellt. Mehr als 75% aller Signalislands befinden sich bei Energien unterhalb von 3 GeV. Dabei hat die Verteilung ein deutliches Maximum zwischen 1 und 2 GeV. Bei Energien oberhalb von 3 GeV fällt die Verteilung stark ab, so daß bei Energien oberhalb von 6 GeV kaum noch Islands vorhanden sind. Die Analyse wurde daher auf Islands mit einer Energie von bis zu 6 GeV beschränkt. Die konstruierten Trainings- und Testdatensätze enthielten jeweils 2000 Signal- und 2000 Untergrundislands. Dabei wurde eine flache Energieverteilung der Signal- und Untergrundislands verwendet (vgl. Kapitel 8.4).

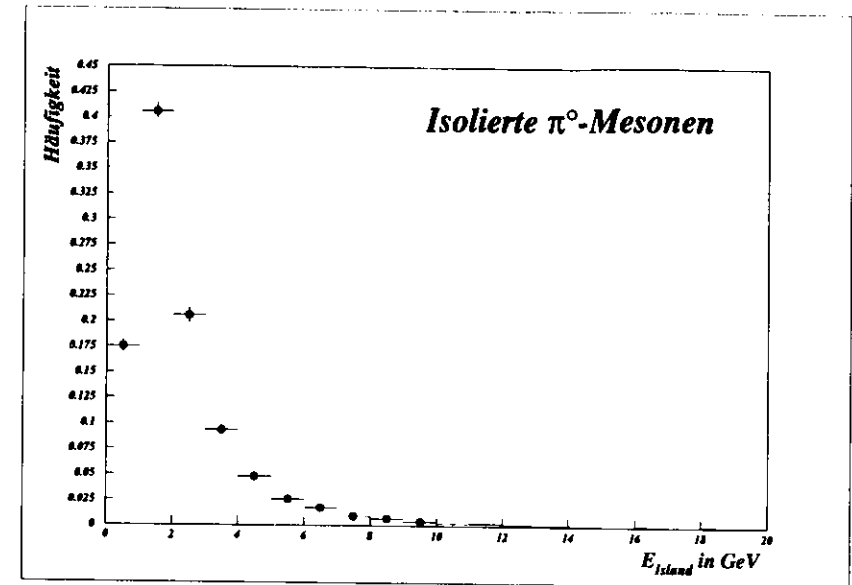


Abbildung 9.1: Energiespektrum der isolierten π^0 -Mesonen. Die Fehlerbalken entsprechen der Breite der gewählten Energiebereiche von 1 GeV bzw. dem statistischen Fehler. Die Verteilung ist auf 1 normiert.

9.2 Trainingsparameter der neuronalen Netzwerke

Alle verwendeten neuronalen Netzwerke entsprachen der in den Kapiteln 7.3 bis 7.5 beschriebenen Form des abgewandelten *feed forward*-Netzwerkes. Ihre Konfigurationen waren bis auf die Variation der Anzahl der Eingabe- und verborgenen Knoten und der Anzahl der Trainingsiterationen identisch. Es wurden zwei Gruppen von *feed forward*-Netzwerken auf die π^0 -Identifikation trainiert. Bei der einen Gruppe wurden als Eingabeinformationen nur die 55 CAL-Informationen verwendet, bei der anderen zusätzlich die 22 Informationen des HES. Zur Unterscheidung der beiden werden im folgenden die Indizes CAL und CAL+HES verwendet. Um die Effizienz der π^0 -Identifikation durch beide Netzwerkgruppen direkt vergleichen zu können, wurden die verwendeten Datensätze in beiden Fällen aus denselben Islands konstruiert. Bei der Beschränkung auf CAL-Informationen wurden jedoch die vorhandenen Informationen des HES nicht als Eingabeinformationen der neuronalen Netzwerke verwendet. Innerhalb der einzelnen Gruppen unterschieden sich die Netzwerke in der Anzahl der verwendeten verborgenen Knoten.

In Tabelle 9.1 sind die die Netzwerke charakterisierenden Größen aufgelistet. Die Anzahl der Eingabeknoten und verborgenen Knoten wurde im angegebenen Bereich variiert. Die Anfangswerte des Geschwindigkeitsfaktors κ und der Mindestschrittweite γ betragen 0,5 und 0,0. Bei größeren Werten bei Temperaturen unterhalb von 0,3 Oszillationen der Temperatur t und der Mißidentifikationsraten f_{train} und f_{test} auf. Dieses Verhalten deutet auf eine Fehlerfunktion mit großem Gradienten bei tiefen Temperaturen hin und

verhindert eindeutige Aussagen über die Effizienz der π^0 -Identifikation. Die Gewichte der Eingabeinformationen jedes Knotens sind nach Gleichung (7.9) normiert. Das Training der verwendeten abgewandelten *feed forward*-Netzwerke wurde mit dem Programm DEXTRA [26] durchgeführt. In jeder Trainingsiteration wurde die Fehlerfunktion (7.6) aus den quadratischen Differenzen der gewünschten und der durch die Netzwerke berechneten Ausgabewerte aller 4000 Muster des Trainingsdatensatzes bestimmt. Anschließend wurden die Parameter des Netzwerkes nach (7.7) und (7.8) verändert.

| Parameter | Wert |
|---|-------------------|
| Eingabeknoten | 55,77 |
| Verborgene Knoten | 2-10 |
| Ausgabeknoten | 1 |
| Anzahl der Signal-/ Untergrundmuster des Trainings- und Testdatensatzes | 2000/2000 |
| Anzahl der pro Trainingsiteration zur Berechnung der Fehlerfunktion verwendeten Trainingsmuster | 4000 |
| Anzahl der Trainingsiterationen | 5000 |
| Transferfunktion | $g(x) = \tanh(x)$ |
| Anfangstemperatur t_0 | 5,0 |
| Initialisierungsintervall der Gewichte | [-1,1] |
| Anfangswert der Schwellenwerte | 0 |
| Geschwindigkeitsfaktor κ | 0,5 |
| Mindestschrittweite γ | 0,0 |

Tabelle 9.1: Konfiguration der *feed forward*-Netzwerke zur Identifikation von isolierten π^0 -Mesonen.

9.3 Trainingsverlauf

Die Effizienz der π^0 -Identifikation der Netzwerke innerhalb der beiden Gruppen wurde verglichen. Es wurde festgestellt, daß vier verborgene Knoten zur π^0 -Identifikation ausreichend waren. Eine weitere Erhöhung der Anzahl brachte keine Verbesserung der Ergebnisse. Alle folgenden dargestellten Ergebnisse wurden mit vier verborgenen Knoten gewonnen.

Zur Kontrolle des Trainings wurde jedes Netzwerk fünfmal mit unterschiedlichen Anfangsinitialisierungen seiner Gewichte trainiert. In Abb. 9.2 sind die Mittelwerte der Temperatur t und der Mißidentifikationsrate f_{train} beider Netzwerke gegen die Trainingsiteration aufgetragen. Die Fehler entsprechen den Standardabweichungen von t bzw. f_{train} . Der Kurvenverlauf ist in beiden Fällen ähnlich. Die Temperatur sank in den ersten 2000 Iterationen auf einen Wert von $t^{\text{CAL}} = 0,23$ bzw. $t^{\text{CAL+HES}} = 0,26$. In den weiteren Iterationen fiel sie langsam weiter bis 0,19 (t^{CAL}) bzw. 0,21 ($t^{\text{CAL+HES}}$). Die Mißidentifikationsraten

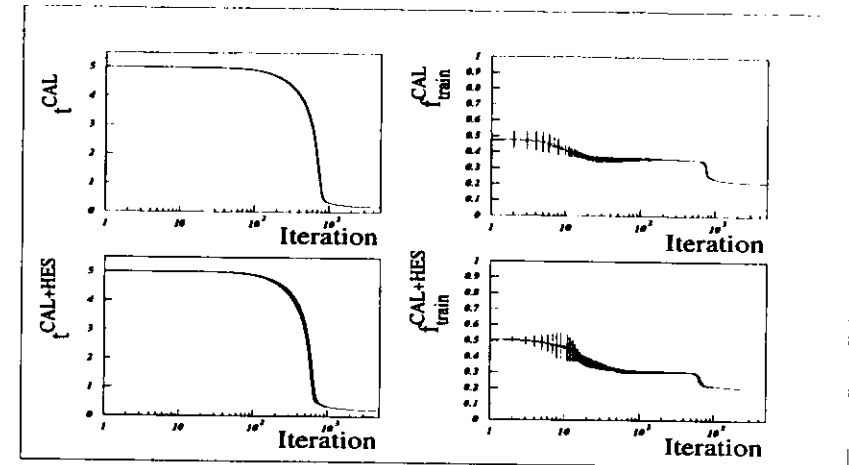


Abbildung 9.2: Trainingsverlauf zur Identifikation isolierter π^0 -Mesonen. Dargestellt sind für alle Iterationen des Trainings Mittelwerte und Standardabweichungen der Temperatur t und der Mißidentifikationsrate f_{train} beim Training mit CAL- (obere Bilder) bzw. mit CAL- und HES-Informationen (untere Bilder).

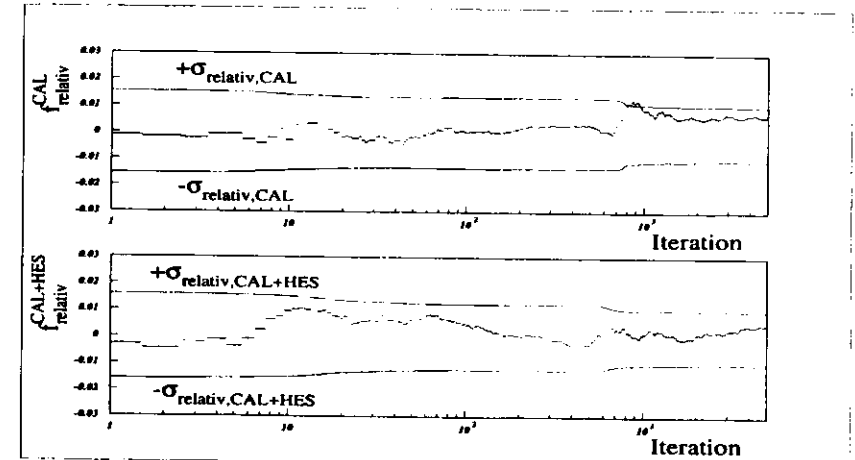


Abbildung 9.3: Relativer Fehler f_{relativ} beim Training zur Identifikation isolierter π^0 -Mesonen. Für jede Iteration sind Mittelwert von f_{relativ} beim Training mit CAL- (oben) bzw. mit CAL- und HES-Informationen (unten) aufgetragen. Die berechneten Vergleichswerte $\pm\sigma_{\text{relativ}}$ zur Identifikation von Übertraining sind ebenfalls eingezeichnet.

der Trainingsdatensätze waren nach 2000 Iterationen ebenfalls nahezu konstant und betragen 0,22 ($f_{\text{train}}^{\text{CAL}}$) bzw. 0,21 ($f_{\text{train}}^{\text{CAL+HES}}$). Im weiteren Verlauf des Trainings sanken sie bis auf $f_{\text{train}}^{\text{CAL}} = 0,21$ bzw. $f_{\text{train}}^{\text{CAL+HES}} = 0,20$. Die Fehler von t und f_{train} waren bei beiden Netzwerken nach 1000 Iterationen um zwei, nach 2000 Iterationen um drei Größenordnungen kleiner als der Mittelwert und somit vernachlässigbar gering.

Trotz unterschiedlicher Anfangsbedingungen konvergierten die Werte von t und f_{train} der fünf verschiedenen Trainingsphasen. Die gewonnenen Ergebnisse hängen damit nicht von den Anfangsbedingungen des Trainings ab, sondern sind durch den Verlauf der Fehlerfunktion (7.6) und der Bestimmung ihres globalen Minimums gegeben.

Zur Kontrolle gegen Übertraining wurde für jede Trainingsiteration der relative Fehler f_{relativ} zwischen dem Trainings- und Testdatensatz nach Gleichung (7.26) berechnet. Der Fehler von f_{relativ} war nach 1000 Iterationen um eine Größenordnung niedriger als f_{relativ} . Auf eine Darstellung wurde verzichtet. f_{relativ} beider Netzwerke ist in Abb. 9.3 für alle Trainingsiterationen dargestellt. Die berechneten Vergleichswerte $\pm\sigma_{\text{relativ}}$ zur Identifikation von Übertraining wurden nach Gleichung (7.30) für jede Trainingsiteration bestimmt und sind in Abb. 9.3 eingezeichnet. Während $f_{\text{relativ}}^{\text{CAL}}$ für ca. 200 Iterationen größer als $+\sigma_{\text{relativ}}$ war, verlief $f_{\text{relativ}}^{\text{CAL+HES}}$ im ganzen Training innerhalb von $\pm\sigma_{\text{relativ}}$. Da auch $f_{\text{relativ}}^{\text{CAL}}$ im weiteren Verlauf des Trainings stets innerhalb dieses Bereiches lag, deutete dieses Verhalten nicht auf Übertraining hin. f_{relativ} betrug bei höheren Iterationen nahezu konstant 0,007, während $f_{\text{relativ}}^{\text{CAL+HES}}$ langsam bis ca. 0,005 anstieg. Dieses Verhalten deutete auf den Beginn des Übertrainings hin. Da der Vergleichswert von $\pm\sigma_{\text{relativ}}$ nicht überschritten wurde und die Mißidentifikationsraten sowohl des Trainings- als auch des Testdatensatzes im weiteren Verlauf des Trainings weiter abnahmen, wurde das Training nicht abgebrochen.

9.4 Klassifikation der Islands

Die aus den Ausgabewerten der einzelnen Islands des Testdatensatzes nach Gleichung (7.22) berechneten Wahrscheinlichkeitsverteilungen sind für beide Netzwerke in Abb. 9.4 dargestellt. Beide Verteilungen haben ähnliche Gestalt. Sowohl bei der Beschränkung auf CAL-Informationen, als auch unter Berücksichtigung der Informationen des HES konnte keine vollständige Separation von Untergrund und π^0 -Signal erreicht werden. Nur 50% der Untergrundislands sind mit $P^n < 0,025$ eindeutig als Untergrund klassifiziert worden. Die Wahrscheinlichkeiten der restlichen Untergrundislands und des Großteils der Signalislands bilden eine breite Verteilung oberhalb von $P^n = 0,45$ (CAL-Informationen) bzw. $P^n = 0,3$ (CAL- und HES-Informationen). Unter Berücksichtigung der Informationen des HES ist das Maximum der Wahrscheinlichkeitsverteilung der π^0 -Signale besser ausgeprägt und liegt näher beim Idealwert von 1. Dieses Verhalten deutet darauf hin, daß der HES zusätzliche Informationen zur Identifikation der π^0 -Signale enthält. Durch verschiedene Schnittwerte in den Wahrscheinlichkeitsverteilungen wurden Untergrund- und π^0 -Islands mit unterschiedlicher Reinheit und Effektivität identifiziert. Zum Vergleich der π^0 -Identifikation durch beide Netzwerke wurden Effektivität und Reinheit für verschiedene Schnittwerte gegeneinander aufgetragen (vgl. Abb. 9.5). Unterhalb einer Reinheit von 70% ist die Effektivität beider Netzwerke innerhalb ihrer statistischen Fehler gleich. Bei der Verwendung von CAL- und HES-Informationen sind für etwa gleich viele Untergrund-

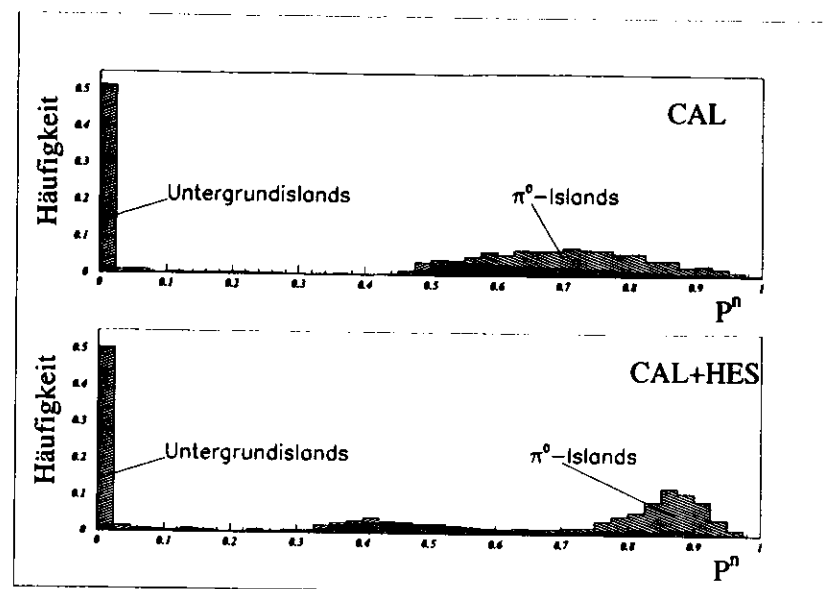


Abbildung 9.4: Zur Klassifikation isolierter π^0 -Mesonen: Die Häufigkeitsverteilung der aus den Ausgabewerten der neuronalen Netzwerke berechneten Wahrscheinlichkeit P^n wurden separat für die Untergrund- und π^0 -Islands aufgetragen. Die obere Abbildung zeigt die Verteilungen bei Training mit CAL-Informationen, die untere Abbildung bei Training mit CAL- und HES-Informationen. Alle Verteilungen sind auf 1 normiert.

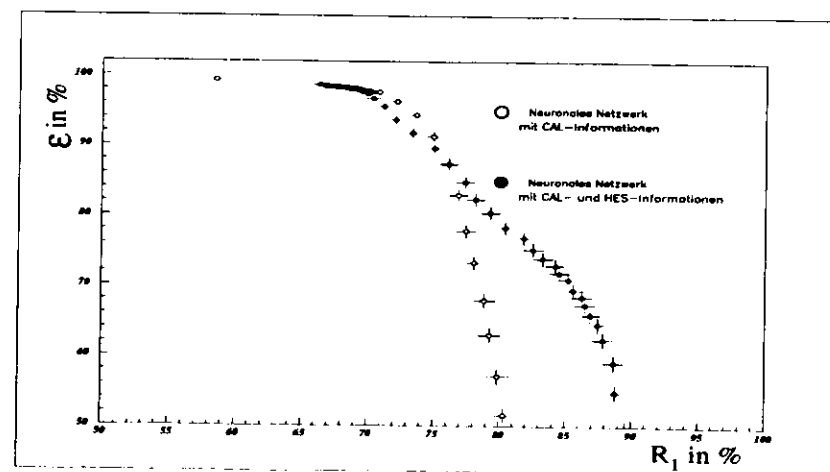


Abbildung 9.5: Effektivität ϵ und Reinheit R_1 für verschiedene Schnittwerte in der Wahrscheinlichkeitsverteilung der neuronalen Netzwerke. Die Fehlerbalken stellen die statistischen Fehler dar.

und Signalmuster Wahrscheinlichkeiten zwischen 0,3 und 0,45 berechnet worden, während dieser Fall bei der Beschränkung auf CAL-Informationen bei weitaus weniger Islands auftrat. Daher gelingt die π^0 -Identifikation bei Schnittwerten zwischen 0,3 und 0,45 nur mit CAL-Informationen mit höherer Reinheit bzw. Effektivität als unter Berücksichtigung der HES-Informationen. In Abb. 9.5 ist das bei einer Reinheit zwischen 70% und 78% der Fall. Oberhalb eines Schnittwertes von 0,5 gelingt die Klassifikation mit zusätzlichen HES-Informationen besser, da weniger Signalislands fehlidentifiziert werden. Daher ist die Effektivität oberhalb einer Reinheit von 78% mit HES-Informationen besser.

Die geringere Effektivität der π^0 -Identifikation unter Verwendung der HES-Informationen bei einer Reinheit von 70% bis 78% ist durch die iterative Natur des Trainings bedingt. Da das Training nach 5000 Iterationen abgebrochen wurde, ist das globale Minimum der Fehlerfunktion (7.6) noch nicht erreicht worden. Aus diesem Grunde werden bei beiden Netzwerken einige Muster noch nicht optimal klassifiziert. Die Effektivität ist bei Schnittwerten zwischen 0,3 und 0,5 unter Berücksichtigung der HES-Informationen geringer als bei Beschränkung auf CAL-Informationen, da mehr Signalmuster mit Wahrscheinlichkeiten in diesem Bereich fehlklassifiziert werden. Bei einer Verlängerung des Trainings ist zu erwarten, daß sich die Effektivität unter Verwendung von CAL- und HES-Informationen bei Reinheiten von 70-78% den Werten bei Beschränkung auf CAL-Informationen angleicht bzw. sie übersteigt. Der Einfluß der HES-Informationen auf die π^0 -Identifikation wird detailliert in Kapitel 10 behandelt. Daher wurde das Training der Netzwerke bei dem in diesem Kapitel behandelten Spezialfall der π^0 -Identifikation nicht über die 5000 Iterationen fortgesetzt.

In den verwendeten Trainings- und Testdatensätzen betrug das Verhältnis von Signal- zu Untergrundislands in Energiebereichen der Breite 1 GeV jeweils 1:1 (siehe Kapitel 8.4). Das in einer Messung erwartete Verhältnis ist dagegen im gesamten Energiebereich kleiner als 2:5 (vgl. Abb. 8.2). Der Schnittwert zur Bestimmung der Effektivität und Reinheit im Testdatensatz sollte so gewählt werden, daß die π^0 -Identifikation mit hoher Reinheit erfolgt. Andernfalls wird bei einem Datensatz, der weitaus mehr Untergrund- als Signalislands enthält, nur noch eine Reinheit unterhalb von 50% erreicht, da mehr Untergrundislands als Signal klassifiziert werden als Signalislands. Gleichzeitig muß jedoch eine möglichst hohe Effektivität gegeben sein. Unter Berücksichtigung von CAL- und HES-Informationen sind Effektivität und Reinheit bei einem Wert von 80% gleich. Zur Untersuchung der Energieabhängigkeit der π^0 -Identifikation wurde die Effektivität für verschiedene Islandenergien bei einer festen Reinheit von 80% berechnet. Die Ergebnisse beider Netzwerke sind in Abb. 9.6 dargestellt. Bei einer Islandenergie zwischen 2 und 5 GeV liegen die Effektivitäten beider Netzwerke oberhalb von 90% und sind nahezu identisch. Bei Energien außerhalb dieses Bereiches ist ihre Effektivität geringer. Unter Berücksichtigung der Informationen des HES ist die Effektivität jedoch deutlich höher als bei der Beschränkung auf CAL-Informationen. Die π^0 -Identifikation erfolgt unter Berücksichtigung der HES-Informationen auch oberhalb von 5 GeV noch mit einer Effektivität von 82%. Bei Beschränkung auf Informationen des CAL sinkt sie im Vergleich zu Islandenergien von 4-5 GeV von 91% auf 33% ab. Dieses Verhalten läßt den Schluß zu, daß die von den Photonen eines π^0 -Zerfalls ausgelösten Teilchenschauer im Kalorimeter stark überlappen. Daher kann mit den Informationen der einzelnen Photovervielfacher bei einer Reinheit von 80% nur noch ca. ein Drittel der π^0 -Mesonen korrekt identifiziert werden. Die höhere Effektivität unter Berücksichtigung der HES-Informationen bei Energien unterhalb von 2 GeV bzw. oberhalb von 5 GeV zeigt, daß der HES bei diesen Energien

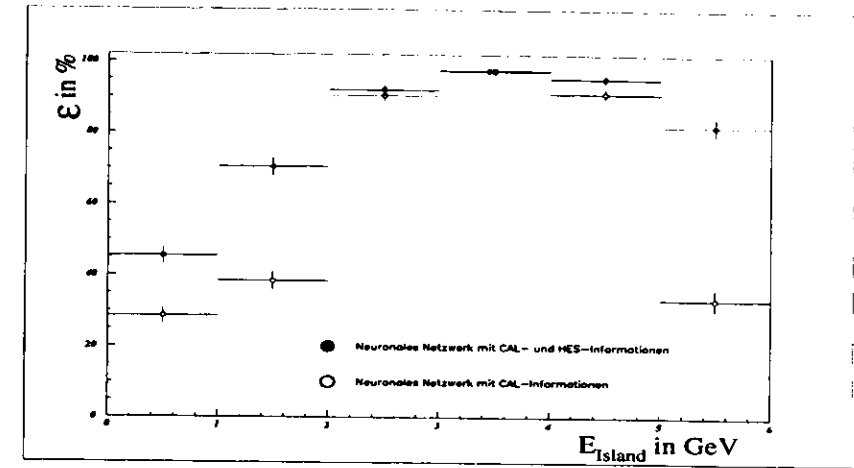


Abbildung 9.6: Effektivität ε für verschiedene Energiebereiche bei einer festen Reinheit von 80%.

zusätzliche Informationen zur π^0 -Identifikation liefert.

Unterhalb von 2 GeV ist die Effektivität bei Beschränkung auf CAL-Informationen kleiner als 40%. Auch unter Berücksichtigung der HES-Informationen ist die Effektivität in diesem Energiebereich geringer als bei Energien oberhalb von 2 GeV. Im Vergleich zur Effektivität ohne die Verwendung der Informationen des HES wird die Effektivität um 16,9% (bis 1 GeV) bzw. 32% (1-2 GeV) gesteigert.

9.5 Auswirkung der verwendeten Informationen des CAL

Bei der Konstruktion der Eingabeinformationen des CAL für ein gegebenes Island wurden nur die Energiedepositionen im Turm der maximalen Energie und der acht benachbarten Türme des Islands verwendet. Türme des Islands, die keine gemeinsame Kante mit dem Turm der maximalen Energie hatten, wurden nicht berücksichtigt. Um die Auswirkungen dieses Verfahrens zu untersuchen, wurde für jedes Island der mit CAL- und HES-Informationen konstruierten Trainings- und Testdatensätze das Verhältnis $r = \frac{E_9}{E_{\text{Island}}}$ der Energie E_9 der neun berücksichtigten Türme und der Gesamtenergie E_{Island} des Islands berechnet. Enthalten die vernachlässigten Türme wesentliche Informationen zur π^0 -Erkennung, so muß die berechnete Wahrscheinlichkeit P^n der Signal- und Untergrundislands eine Abhängigkeit von r zeigen.

In Abb. 9.7 I und II ist die Verteilung von r für die Signal- und Untergrundislands des Testdatensatzes dargestellt. Bei mehr als 90% der Signal- und mehr als 80% der Untergrundislands wird mehr als 95% der Energie im zentralen Turm des Islands deponiert. Die

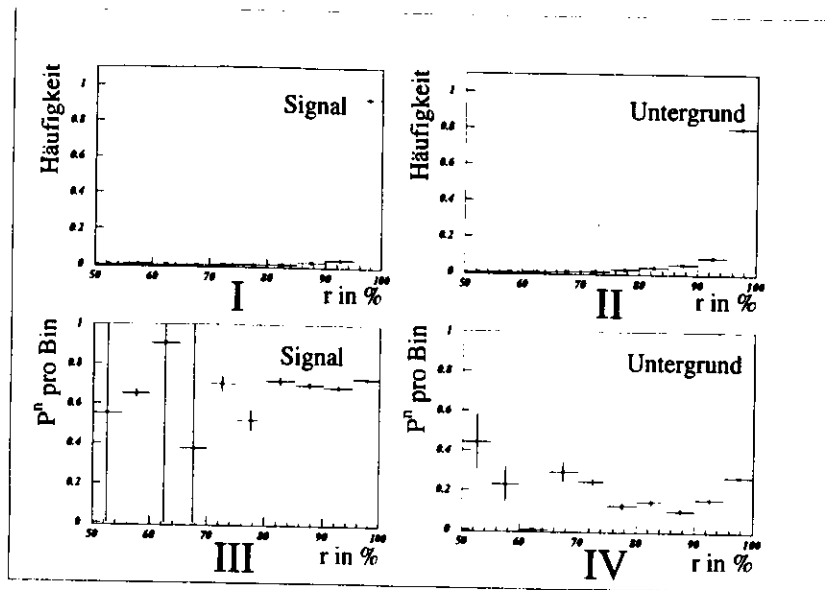


Abbildung 9.7: Kontrolle der Konstruktion der CAL-Informationen. In den oberen Bildern sind die Verteilungen von $r = \frac{E_{\pi^0}}{E_{\text{Island}}}$ für Signal- (links) und Untergrundislands (rechts) dargestellt. Beide Verteilungen sind auf 1 normiert. In den unteren Bildern ist P^n gegen r für Signal- (links) und Untergrundislands (rechts) aufgetragen.

Statistik für r kleiner als 95% ist daher sehr gering. In den Abbildungen 9.7 III und IV ist die mittlere Wahrscheinlichkeit gegen r aufgetragen. Die dargestellten Fehler entsprechen den Fehlern des Mittelwerte. Oberhalb von $r = 80\%$ ist P^n für die Signalislands nahezu konstant. Für die Untergrundislands steigt P^n oberhalb von $r = 90\%$ leicht an. Die Mittelwerte von P^n sind jedoch für Signal- und Untergrundislands deutlich verschieden. Unterhalb von $r = 80\%$ sind die Fehler von P^n durch die geringe Statistik sehr groß. Im Rahmen dieser Fehler zeigt sich auch in diesem Bereich von r keine eindeutige Korrelation zwischen r und P^n der Islands. Daher beschränkt man sich in der weiteren Analyse bei der Konstruktion der CAL-Informationen weiterhin auf die neun zentralen Türme der einzelnen Islands.

Kapitel 10

Identifikation von nicht-isolierten π^0 -Mesonen

10.1 Konstruktion der Trainingsdaten

In diesem Kapitel wird der allgemeine Fall der π^0 -Identifikation behandelt. Nach Anwendung der in Kapitel 8.2 beschriebenen Vorauswahlkriterien werden keine weiteren Islands von der Analyse ausgeschlossen. Alle Islands, in denen Energie von den beiden Photonen mindestens eines π^0 -Zerfalls deponiert wurde, werden als Signalislands, alle anderen Islands als Untergrundislands bezeichnet. Im folgenden bezeichnet E_{Island} die im CAL gemessene Energie eines Islands und E_{π^0} die von den beiden Photonen eines π^0 -Zerfalls im Island deponierte Energie. In Abb. 10.1 I ist die Verteilung der Energie E_{Island} der Signalislands, in Abb. 10.1 II der Quotient aus E_{π^0} und E_{Island} für alle Signalislands in Abhängigkeit von E_{Island} aufgetragen. Beide Verteilungen wurden aus Monte Carlo-Ereignissen gewonnen. Die meisten Signalislands treten bei Energien zwischen 2 und 3 GeV auf und liegen damit um ca. 1 GeV höher als bei den isolierten π^0 -Mesonen (vgl. Abb. 9.1). In diesem Energiebereich beträgt der Anteil der Islandenergie, der von π^0 -Zerfällen stammt, ca. 64%. Die Verschiebung des Maximums der Energieverteilung beruht demnach auf einer Energie von 0,7-1,1 GeV im Island, die nicht von π^0 -Mesonen stammt. Da der Anteil der π^0 -Energie mit steigender Islandenergie sinkt, fällt die Verteilung zu hohen Islandenergien hin nicht so schnell ab wie die der isolierten π^0 -Mesonen. Zum Training der neuronalen Netzwerke wurden Islands mit einer Energie von bis zu 14 GeV verwendet, da bei höheren Energien die Anzahl der vorhandenen Signalislands zu klein wird. Analog zur Konstruktion der Trainings- und Testdatensätze zur Identifikation der isolierten π^0 -Mesonen enthielt jeder Datensatz jeweils 2000 Signal- und 2000 Untergrundislands. Sowohl für die Signal- als auch für die Untergrundislands wurde eine flache Energieverteilung verwendet (vgl. Kapitel 8.4). Der mittlere Anteil der π^0 -Energie an der Gesamtenergie der einzelnen Signalislands entsprach der in Abb. 10.1 II dargestellten Verteilung.

Um die Auswirkungen der HES-Informationen auf die π^0 -Identifikation zu studieren, wurden die *feed forward*-Netzwerke einmal nur mit CAL-Informationen und einmal mit Informationen des CAL und HES trainiert. Zur Konstruktion der entsprechenden Datensätze wurden dieselben Islands verwendet.

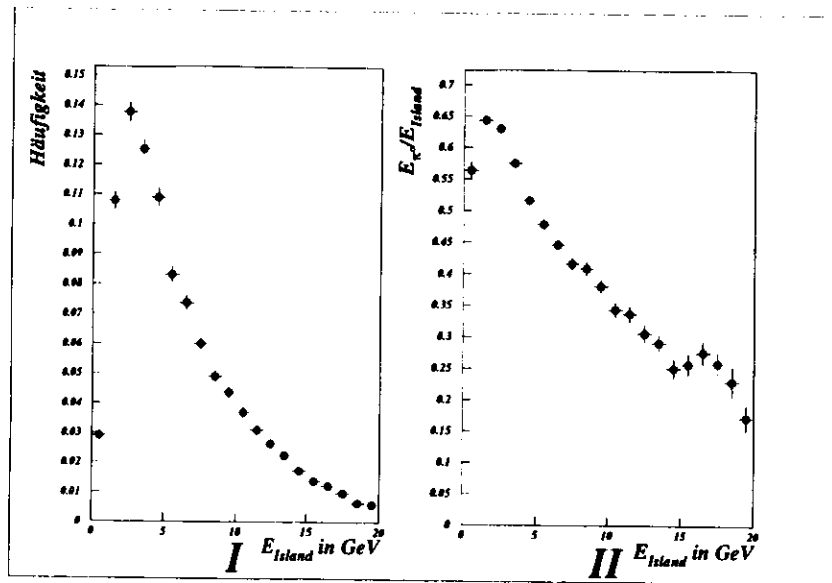


Abbildung 10.1: Energiespektrum der nicht-isolierten π^0 -Mesonen. Bild I zeigt die auf 1 normierte Energieverteilung der Signalislands, Bild II den mittleren Anteil der π^0 -Energie an der Gesamtenergie der einzelnen Signalislands.

10.2 Trainingsverlauf

Das Training der *feed forward*-Netzwerke wurde bis auf die Anzahl der verwendeten verborgenen Knoten mit den bei der Identifikation der isolierten π^0 -Mesonen bestimmten Trainingsparametern durchgeführt (vgl. Tabelle 9.1). In jeder Trainingsiteration wurden alle 4000 Muster des Trainingsdatensatzes zur Berechnung der Fehlerfunktion (7.6) und somit zur Korrektur der Werte der Gewichte und Schwellenwerte der neuronalen Netzwerke verwendet. Durch Vergleich der π^0 -Identifikation durch neuronale Netzwerke mit unterschiedlich vielen verborgenen Knoten wurde ermittelt, daß bei der Beschränkung auf CAL-Informationen sechs verborgene Knoten ausreichend waren, unter Berücksichtigung der HES-Informationen dagegen acht. Eine höhere Anzahl verborgener Knoten verbesserte die π^0 -Identifikation nicht weiter. Die beiden zusätzlichen Knoten deuten darauf hin, daß die Informationen des HES eine Verbesserung der π^0 -Identifikation ermöglichen. Im folgenden werden die Ergebnisse der Netzwerke mit den genannten verborgenen Knoten dargestellt.

Die beiden Netzwerke wurden mit fünf verschiedenen Initialisierungen der Gewichte und Schwellenwerte trainiert. Die Anzahl der Trainingsiterationen betrug 5000. In Abb. 10.2 sind die Mittelwerte und Standardabweichungen der Temperatur t und der Mißidentifikationsrate f_{train} für jede Trainingsiteration dargestellt. Dabei bezeichnen die Indizes CAL bzw. CAL+HES die nur mit CAL- bzw. mit CAL- und HES-Informationen trainierten Netzwerke. Die Fehler der Mißidentifikationsraten $f_{\text{train}}^{\text{CAL}}$ bzw. $f_{\text{train}}^{\text{CAL+HES}}$ waren schon nach

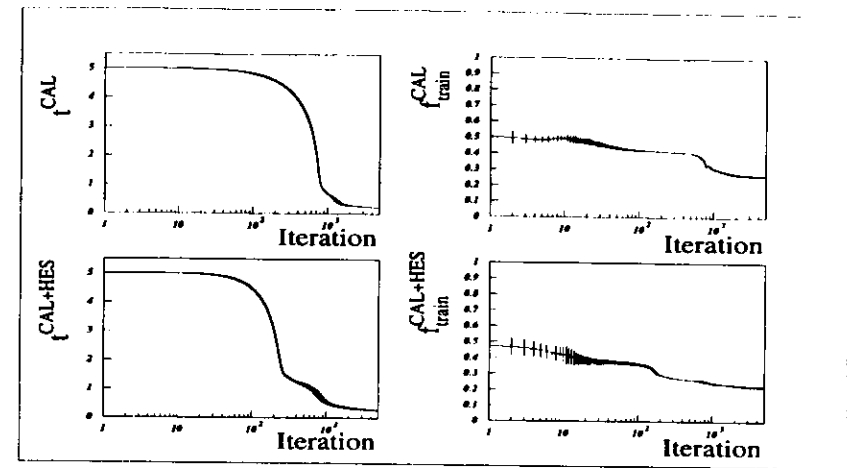


Abbildung 10.2: Trainingsverlauf zur Identifikation nicht-isolierter π^0 -Mesonen. Dargestellt sind für alle Iterationen des Trainings Mittelwerte und Standardabweichungen von t und f_{train} beim Training mit CAL- (oben) bzw. mit CAL- und HES-Informationen (unten).

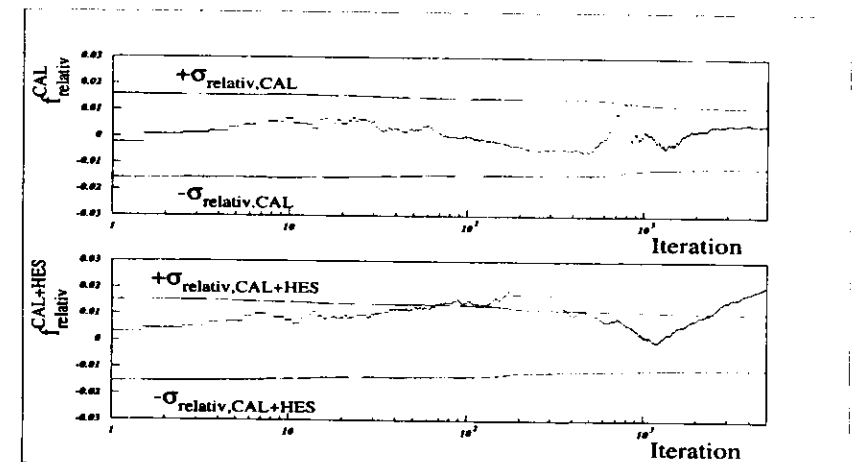


Abbildung 10.3: Relativer Fehler f_{relativ} beim Training zur Identifikation nicht-isolierter π^0 -Mesonen. Für jede Iteration sind Mittelwert und Standardabweichung von f_{relativ} beim Training mit CAL- (oberes Bild) bzw. mit CAL- und HES-Informationen (unteres Bild) aufgetragen. Die zur Kontrolle gegen Übertraining berechneten Vergleichswerte $\pm \sigma_{\text{relativ,CAL}}$ bzw. $\pm \sigma_{\text{relativ,CAL+HES}}$ sind ebenfalls dargestellt.

ca. 200 Iterationen vernachlässigbar klein, während dieser Fall bei der Temperatur erst nach etwa 2000 Iterationen auftrat.

Aus diesem Verhalten kann geschlossen werden, daß die Minimierung der Fehlerfunktion trotz unterschiedlicher Initialisierungen der Gewichte zum selben Ergebnis führt. Im Vergleich zur Identifikation der isolierten π^0 -Mesonen sind die erreichten Minimalwerte von t und f_{train} größer, was zusammen mit dem Anstieg der nötigen Anzahl an verborgenen Knoten auf eine erhöhte Schwierigkeit der π^0 -Identifikation hindeutet.

Zur Kontrolle des Trainings wurde f_{relativ} bei beiden Netzwerken für jede Iteration berechnet. Aus den Werten von $f_{\text{train}}^{\text{CAL}}$ bzw. $f_{\text{train}}^{\text{CAL+HES}}$ wurden $\sigma_{\text{relativ}}^{\text{CAL}}$ bzw. $\sigma_{\text{relativ}}^{\text{CAL+HES}}$ für jede Iteration nach Gleichung (7.30) berechnet. Sie sind zusammen mit den relativen Mißidentifikationsraten beider Netzwerke in Abb. 10.3 dargestellt. Bei der Beschränkung auf CAL-Informationen war $f_{\text{relativ}}^{\text{CAL}}$ bis auf ca. 100 Iterationen ständig innerhalb von $\pm \sigma_{\text{relativ}}^{\text{CAL}}$ und betrug nach 3000 Iterationen nahezu konstant 0,0045. Da kein Übertraining festgestellt wurde, erfolgte kein Abbruch des Trainings.

Beim Training mit CAL- und HES-Informationen wurde Übertraining festgestellt. Ab ca. der 1200sten Iterationen nahm $f_{\text{relativ}}^{\text{CAL+HES}}$ kontinuierlich bis zu einem Wert von 0.015 zu und war nach mehr als 2500 Iterationen permanent außerhalb des Vergleichswertes $\sigma_{\text{relativ}}^{\text{CAL+HES}}$. Das Training wurde daher nach 2000 Iterationen abgebrochen. Obwohl $+\sigma_{\text{relativ}}^{\text{CAL+HES}}$ bereits nach weniger als 400 Iterationen überschritten wurde, fand kein Abbruch des Trainings bei diesen Iterationen statt, da $+\sigma_{\text{relativ}}^{\text{CAL+HES}}$ im weiteren Verlauf des Trainings wieder unterschritten wurde.

Die weiteren Ergebnisse wurden bei Beschränkung auf CAL-Informationen mit 5000 Trainingsiterationen, beim Training mit CAL- und HES-Informationen mit 2000 Trainingsiterationen ermittelt.

10.3 Klassifikation der Islands

Aus den mit den *feed forward*-Netzwerken berechneten Ausgabewerten wurde für jedes Island des Testdatensatzes nach Gleichung (7.22) die Wahrscheinlichkeit P^n berechnet, das Signal mindestens eines π^0 -Zerfalls zu enthalten. Der Index n bezeichnet die Nummer des Islands im Testdatensatz. Die Wahrscheinlichkeitsverteilungen der Untergrund- und Signalislands sind in Abb. 10.4 dargestellt. Die Maxima der Verteilungen sind unter Berücksichtigung der Informationen des HES stärker ausgeprägt und liegen näher an den gewünschten Ausgabewerten von 0 für Untergrund- bzw. 1 für Signalislands. Im Vergleich zur Identifikation der isolierten π^0 -Mesonen sind die Maxima schwächer ausgeprägt. Bei diesem vereinfachten Problem wurden etwa 50% der Untergrundislands mit Wahrscheinlichkeiten kleiner als 0.025 eindeutig als solche identifiziert (vgl. Abb. 9.4). Bei der Identifikation von nicht-isolierten π^0 -Mesonen wurden unter Berücksichtigung der HES-Informationen nur 13% der Untergrundislands eindeutig als solche identifiziert, bei der Beschränkung auf CAL-Informationen weniger als 2%. In beiden Fällen existiert eine breite Überlappungszone der Wahrscheinlichkeitsverteilungen der Signal- und Untergrundislands. Sie zeigt, daß die Identifikation von nicht-isolierten π^0 -Mesonen problematischer als der in Kapitel 9 behandelte Spezialfall der isolierten π^0 -Mesonen ist. Ein Teil der Signalislands wird mit Wahrscheinlichkeiten kleiner als 0,45 nicht eindeutig identifiziert. Zur Analyse dieser Signalislands wurden ihre Energie E_{Island} und die Anzahl ihrer

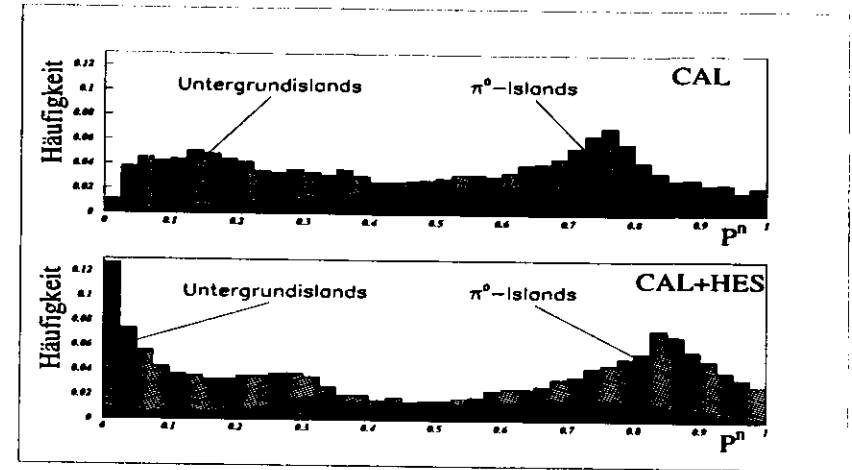


Abbildung 10.4: Zur Klassifikation nicht-isolierter π^0 -Mesonen: Die Häufigkeitsverteilung der aus den Ausgabewerten der neuronalen Netzwerke berechneten Wahrscheinlichkeiten P^n wurden separat für die Untergrund- und π^0 -Islands aufgetragen. Die obere Abbildung zeigt die Verteilungen bei Training mit CAL-Informationen, die untere Abbildung bei Training mit CAL- und HES-Informationen. Alle Verteilungen sind auf 1 normiert.

HES-Isles N_{Isle} mit der Verteilung der beiden Größen aller Signalislands verglichen. Der Schnittwert von $P^n = 0,45$ wurde gewählt, da die Wahrscheinlichkeitsverteilungen der Signal- und Untergrundislands bei diesem Wert identisch sind. Bei kleineren bzw. größeren Wahrscheinlichkeiten überwiegen die Untergrund- bzw. Signalislands. Für die unter Berücksichtigung der HES-Informationen berechnete Wahrscheinlichkeitsverteilung sind die 4 Verteilungen in Abb. 10.5 dargestellt. Bei Beschränkung auf CAL-Informationen ergeben sich im Rahmen der Fehler die gleichen Verteilungen. Sowohl bei der Energieverteilung, als auch bei der Anzahl der HES-Isles pro CAL-Island sind deutliche Unterschiede vorhanden. Die Energieverteilung aller Signalislands ist bis auf eine durch mangelnde Statistik bedingte Abnahme oberhalb von 13 GeV konstant. Die Energieverteilung der Signalislands mit P^n kleiner als 0,45 steigt dagegen unterhalb von 1 GeV stark an. Oberhalb von Islandenergien von 1 GeV ist die Verteilung der nicht eindeutig identifizierten Signalislands flach. Dieses Verhalten deutet darauf hin, daß ein energieunabhängiger Anteil der Signalislands eine ähnliche Signatur wie die Untergrundislands aufweist und daher nur schwierig korrekt als Signal identifiziert werden kann.

Auch die Anzahl der dem Island zugeordneten HES-Isles zeigt deutliche Unterschiede. Während weniger als 11% aller Signalislands nur eine HES-Isle zugeordnet werden konnte, beträgt dieser Anteil bei den Signalislands mit P^n kleiner als 0,45 nahezu 50%, bei Energien unterhalb von 1 GeV sogar 99%. Außerdem ist der Anteil der Islands, denen mehr als zwei HES-Isles zugeordnet wurden, geringer als bei allen Signalislands. Die Häufung der nicht eindeutig identifizierten Signalislands bei Energien unterhalb von 1

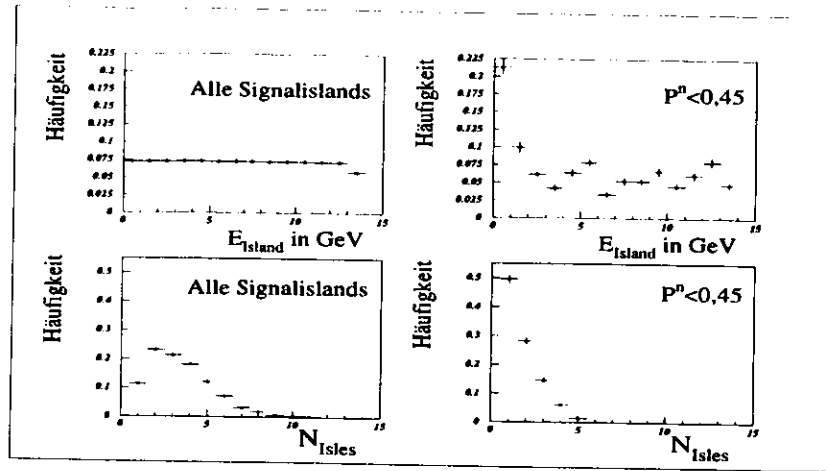


Abbildung 10.5: Analyse der nicht eindeutig identifizierten Signalislands. Für alle Signalislands (linke Seite) und die Signalislands mit einer Wahrscheinlichkeit P^n kleiner als 0,45 (rechte Seite) sind die Verteilungen der Energie E_{Island} und der Anzahl der ihnen zugeordneten HES-Isles N_{Isles} dargestellt. Alle Verteilungen sind auf 1 normiert.

GeV ist durch die Kinematik des π^0 -Zerfalls und die Datenrekonstruktion im Kalorimeter bedingt. Unterhalb einer π^0 -Energie von 1 GeV deponieren die beiden Photonen des π^0 -Zerfalls den größten Teil ihrer Energie i.a. nicht in benachbarten Kalorimetertürmen (vgl. Abb 6.1). Die Signalislands bei diesen Energien sind Spezialfälle mit Energiedeposition von beiden Zerfallsphotonen. Bei dieser Energie tragen die beiden Photonen im Mittel nicht mehr jeweils die Hälfte der π^0 -Energie, sondern die Energie eines Photons kann um ein Vielfaches größer als die des anderen sein. Ist ein Turm mit Energiedeposition des energiereicheren Photons benachbart zum Turm der maximalen Energie des anderen Photons und hat höhere Energie als dieser, so werden die Signale der beiden Photonen zu einem Island zusammengefaßt.

Auch wenn die Türme der maximalen Energiedeposition beider Photonen in verschiedenen Islands liegen, besteht die Möglichkeit, daß Islands konstruiert werden, die den Großteil der Energie eines Photons, jedoch auch geringe Energie des anderen Photons enthalten. Da das Signal des einen Photons in beiden Fällen sehr viel größer als die des anderen Photons ist, kann dieses Island nur schwer von Untergrundislands mit Energiedeposition eines einzelnen Teilchens, insbesondere eines Photons, unterschieden werden.

In beiden Fällen werden durch die geringe Energiedeposition des einen Photons im Island auch in den dem Island zugeordneten HES-Dioden nur schwache Signale gemessen. Der überwiegende Teil der im HES deponierten Energie befindet sich außerhalb der dem Island zugeordneten Dioden. Ist die Energie in den einzelnen dem Island zugeordneten Dioden so klein, daß sie unter der Mindestenergie zur Rauschunterdrückung im HES liegt, so werden diese Dioden bei der Datenrekonstruktion nicht berücksichtigt. Dadurch wird die Zahl

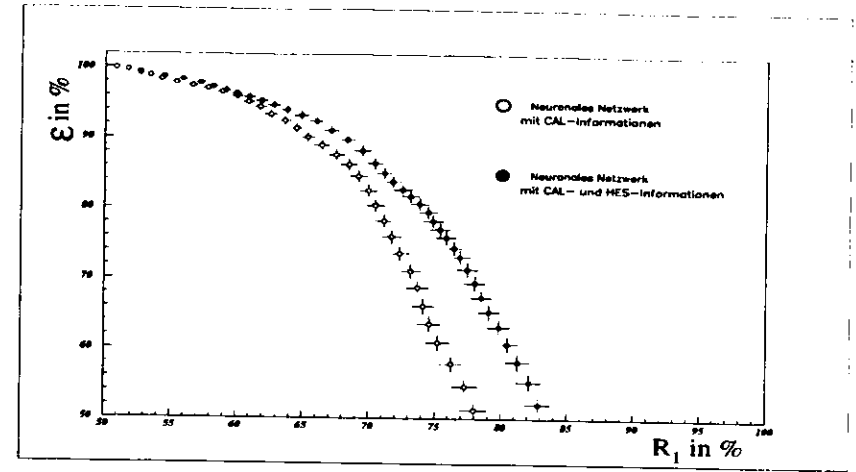


Abbildung 10.6: Effektivität ε und Reinheit R_1 bei der Identifikation nicht-isolierter π^0 -Mesonen für verschiedene Schnittwerte in der Wahrscheinlichkeitsverteilung der neuronalen Netzwerke. Die Fehlerbalken stellen die statistischen Fehler dar.

der Dioden eines Photonensignals reduziert. Im schlimmsten Fall werden alle Dioden mit der Energiedeposition des einen Photons durch die Rauschunterdrückung ausgeschlossen und das Photon im HES nicht nachgewiesen. In diesen Fällen können die π^0 -Zerfälle auch im HES nicht eindeutig identifiziert werden.

10.4 Effizienz der π^0 -Identifikation

Durch unterschiedliche Schnittwerte in der Wahrscheinlichkeitsverteilung wurden Signal- und Untergrundislands mit verschiedener Effektivität und Reinheit identifiziert. Für die einzelnen Schnittwerte sind Effektivität und Reinheit in Abb. 10.6 dargestellt. Unterhalb einer Reinheit von 60% sind die Werte für beide verwendete Netzwerke im Rahmen ihrer Fehler identisch. Oberhalb dieses Wertes ist die Effektivität unter Berücksichtigung der Informationen des HES größer.

Zum Vergleich der π^0 -Identifikation mit und ohne Berücksichtigung der Informationen des HES wurde der Wert der Effektivität bzw. der Reinheit fest vorgegeben und der maximale unter dieser Bedingung erreichte Wert der anderen Größe bestimmt. Die Ergebnisse für konstante Effektivität bzw. Reinheit sind in Tabelle 10.1 dargestellt. Dabei bezeichnen die Indizes CAL bzw. CAL+HES die mit Informationen des CAL bzw. des CAL und des HES erzielten Ergebnisse.

| Effektivität konstant | | | | Reinheit konstant | | | |
|-----------------------|----------------------------|------------------------|------------------|-------------------|------------------------------------|--------------------------------|-------------------------|
| ε [%] | $R_1^{\text{CAL+HES}}$ [%] | R_1^{CAL} [%] | ΔR_1 [%] | R_1 [%] | $\varepsilon^{\text{CAL+HES}}$ [%] | ε^{CAL} [%] | $\Delta\varepsilon$ [%] |
| 50 | 83,3±1,0 | 78,1±1,0 | 5,2±1,4 | 50 | 100,0±0,0 | 100,0±0,0 | 0,0±0,0 |
| 55 | 82,2±1,0 | 77,0±1,0 | 5,2±1,4 | 55 | 98,5±0,3 | 98,0±0,3 | 0,5±0,4 |
| 60 | 80,5±0,9 | 75,5±0,9 | 5,0±1,3 | 60 | 96,3±0,4 | 95,9±0,4 | 0,4±0,6 |
| 65 | 79,0±0,9 | 74,4±0,9 | 4,6±1,3 | 65 | 93,0±0,5 | 90,7±0,7 | 2,3±0,9 |
| 70 | 77,7±0,8 | 73,3±0,8 | 4,4±1,1 | 70 | 87,0±0,6 | 82,5±0,9 | 4,5±1,1 |
| 75 | 76,1±0,8 | 72,0±0,7 | 4,0±1,1 | 75 | 77,8±0,7 | 62,2±1,1 | 15,6±1,3 |
| 80 | 74,1±0,7 | 70,5±0,7 | 3,6±1,0 | 80 | 62,6±0,9 | 11,4±0,7 | 51,2±1,1 |
| 85 | 71,1±0,7 | 69,0±0,6 | 2,1±0,9 | 85 | 43,7±1,0 | — | — |
| 90 | 68,0±0,6 | 65,5±0,6 | 2,5±0,8 | 90 | 9,2±0,6 | — | — |
| 95 | 62,3±0,5 | 60,9±0,4 | 1,4±0,6 | 95 | — | — | — |

Tabelle 10.1: Vergleich von Effektivität und Reinheit bei der Identifikation von nicht-isolierten π^0 -Mesonen mit und ohne Berücksichtigung der Daten des HES. Die Indizes CAL bzw. CAL+HES bezeichnen die mit Informationen des CAL bzw. des CAL und des HES erzielten Ergebnisse. ΔR_1 ist die Differenz der Reinheit bei fester Effektivität, $\Delta\varepsilon$ die Differenz der Effektivität bei fester Reinheit. Wurde die geforderte Reinheit nicht erreicht, so ist in der entsprechenden Spalte ein Strich eingetragen.

Außerdem wurden die Differenzen

$$\Delta\varepsilon = \varepsilon^{\text{CAL+HES}} - \varepsilon^{\text{CAL}} \quad (10.1)$$

$$\Delta R_1 = R_1^{\text{CAL+HES}} - R_1^{\text{CAL}} \quad (10.2)$$

der Ergebnisse beider Netzwerke für konstante Reinheit bzw. Effektivität berechnet und in Tabelle 10.1 eingetragen. Die Reinheit der π^0 -Identifikation durch die neuronalen Netzwerke unterscheidet sich bei fester Reinheit oberhalb von 50% um maximal 5,2±1,4%. Demgegenüber zeigen die bei fester Reinheit von mehr als 70% bestimmten Effektivitäten Differenzen bis zu 51,2±1,1%. Bei Beschränkung auf CAL-Informationen wurde eine maximale Reinheit von 82,6±5,5% erreicht, bei zusätzlicher Verwendung der Informationen des HES betrug dieser Wert 91,1±3,2%.

Da die π^0 -Identifikation durch neuronale Netzwerke im Rahmen dieser Analyse nicht zur Identifikation spezieller physikalischer Prozesse verwendet wurde, bestand kein Anlaß, der Effektivität oder der Reinheit Priorität einzuräumen. Zum Vergleich der π^0 -Identifikation mit und ohne Berücksichtigung der Informationen des HES wurden daher die maximalen Werte der Effektivität und Reinheit bestimmt, bei denen diese beiden Größen innerhalb ihrer Fehler identisch waren. Die größten Werte wurden unter Verwendung von CAL- und HES-Informationen erzielt. Der Mittelwert Ψ der maximalen Effektivität und Reinheit unter der angegebenen Bedingung betrug

$$\Psi_{\text{max}} = 75,8 \pm 0,6\% \quad (10.3)$$

Im Vergleich zur π^0 -Identifikation nur mit CAL-Informationen wurde Ψ_{max} um

$$\Delta\Psi_{\text{max}} = 3,1 \pm 0,8\% \quad (10.4)$$

gesteigert.

Um Informationen über die π^0 -Identifikation bei verschiedenen Energien zu erhalten, wurden Effektivität und Reinheit bei einem jeweils festen Wert der anderen Größe von 76% für die Islands verschiedener Energie bestimmt. Dazu wurde das Energieintervall von 0-14 GeV in 14 Energiebereiche der Breite 1 GeV unterteilt. Durch die Konstruktion der Datensätze konnten Effektivität und Reinheit für die Islands der verschiedenen Energiebereiche direkt verglichen werden, da die Anzahl und das Verhältnis von Signal- und Untergrundislands in den einzelnen Energiebereichen identisch waren (vgl. Kapitel 8.4). Die durch Vorgabe einer festen Effektivität bzw. Reinheit bestimmten Schnittwerte innerhalb der Wahrscheinlichkeitsverteilung der Islands der einzelnen Energiebereiche wiesen starke Schwankungen auf. Daher traten auch in der mit diesen Schnittwerten berechneten Effektivität bzw. Reinheit große Unterschiede zwischen den einzelnen Energiebereichen auf. Um trotz dieser Schwankungen Aussagen über die Energieabhängigkeit der Effektivität und Reinheit zu erhalten, wurde die Energieabhängigkeit der Schnittwerte für feste Reinheit bzw. Effektivität durch eine stetige Funktion approximiert. Eine gute Approximation wurde sowohl bei Verwendung von CAL- als auch von CAL- und HES-Informationen durch eine Funktion $s(E)$ der Form

$$s(E) = (a_1 E + a_2) \cdot \frac{1 - \tanh[a_3(E - a_4)]}{2} + a_5 \cdot \frac{1 + \tanh[a_3(E - a_4)]}{2} \quad (10.5)$$

erreicht. Dabei sind die a_i ($i=1..5$) bei fester Effektivität bzw. Reinheit konstante reelle Zahlen und E die Islandenergie in GeV. $s(E)$ ist der für die Energie E berechnete Schnittwert aus dem Intervall $[0, 1]$. Die für feste Effektivität bzw. Reinheit bestimmten Schnittwerte und die mit ihnen approximierten Funktionen sind in Abb. 10.7 dargestellt. Da die Steigung der Funktionen bei konstanter Effektivität geringer ist als bei konstanter Reinheit, sind die Schwankungen der nicht konstanten Größe zwischen den einzelnen Energiebereichen in diesem Fall weniger stark. Die Ergebnisse für eine feste Reinheit von 76% sind in Abb. 10.8 I, für eine feste Effektivität von 76% in Abb. 10.8 II dargestellt. Die Fehler entsprechen der Breite der Energiebereiche bzw. dem statistischen Fehler der berechneten Effektivitäten und Reinheiten.

Bei Betrachtung der Effektivität der π^0 -Identifikation sind drei Energiebereiche zu unterscheiden. Unterhalb einer Islandenergie von 1 GeV ist die Effektivität kleiner als 50% und im Rahmen der Fehler mit und ohne Berücksichtigung der HES-Informationen identisch. Die Gründe dieses Verhaltens wurden bereits in Kapitel 10.3 erläutert. Zwischen 1 und 9 GeV ist die Effektivität unter Berücksichtigung der Informationen des HES höher als bei Beschränkung auf CAL-Informationen. Innerhalb dieses Energiebereichs hat die Effektivität der π^0 -Identifikation durch beide Netzwerke bei Energien von 4-6 GeV den niedrigsten Wert. Bei 4-6 GeV Islandenergie tragen Hadronen und Positronen der tiefinelastischen Streuung jeweils zu ca. 45% zum Untergrund bei. Der restliche Untergrund besteht im wesentlichen aus Elektronen und Photonen. Bei höheren bzw. tieferen Energien ist der Untergrund von Positronen bzw. Hadronen dominiert. Die geringere Effektivität bei Islandenergien von 4-6 GeV deutet darauf hin, daß die Mischung der beiden Untergrundklassen mit den iterativ ermittelten Parametern der neuronalen Netzwerke

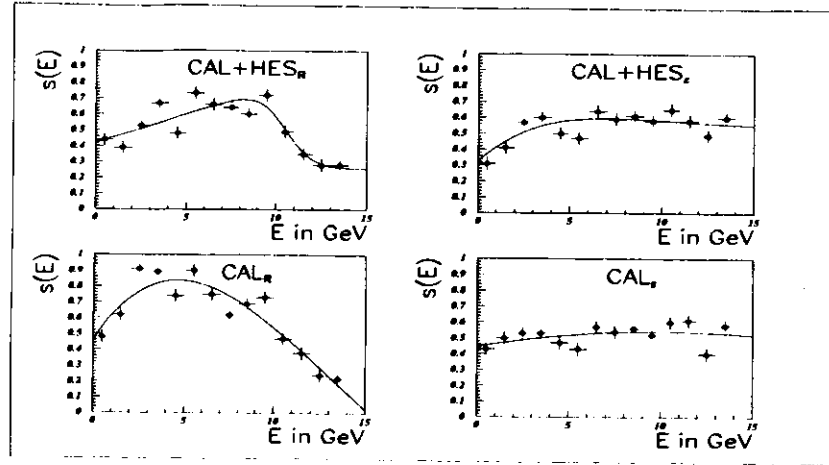


Abbildung 10.7: Approximation der Schnittwerte innerhalb der Wahrscheinlichkeitsverteilungen der neuronalen Netzwerke. Die Indizes ϵ bzw. R bezeichnen die bei einer konstanten Effektivität von 76% bzw. Reinheit von 76% bestimmten Funktionen.

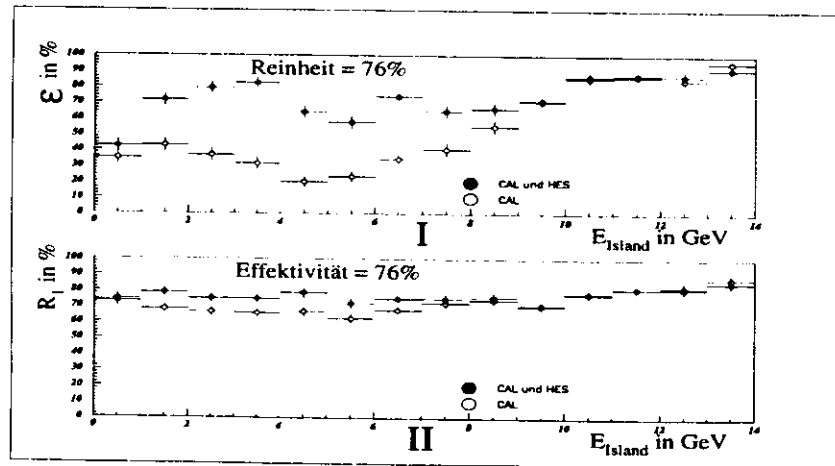


Abbildung 10.8: Energieabhängigkeit der π^0 -Identifikation. Bild I zeigt die Effektivität ϵ in Abhängigkeit der Islandenergie bei einer festen Reinheit von 76%. In Bild II ist die Reinheit R_1 der π^0 -Identifikation gegen die Energie der Islands für eine konstante Effektivität von 76% aufgetragen.

schwieriger zu identifizieren ist als der homogenere Untergrund bei höheren bzw. tieferen Energien. Oberhalb einer Energie von 9 GeV ist die Effektivität beider Netzwerke im Rahmen ihrer Fehler identisch, wobei die größten Werte in beiden Fällen oberhalb von 10 GeV erreicht werden.

Die Reinheit der π^0 -Identifikation zeigt ein ähnliches Verhalten wie die Effektivität. Bei beiden Netzwerken fällt die Reinheit bis zu Energien zwischen 5 und 6 GeV ab und steigt bei höheren Energien bis auf ihren Maximalwert. Bei Energien von 1-7 GeV ist sie unter Berücksichtigung der HES-Informationen größer als bei Beschränkung auf die Informationen des CAL, während sie bei Energien außerhalb dieses Bereiches im Rahmen der Fehler für beide Netzwerke identisch ist. Die Schwankungen zwischen den Werten der einzelnen Energiebereiche sind nicht so stark wie bei der Effektivität, da die verwendeten Schnittwerte nicht so stark mit der Energie variieren.

Die Betrachtung der Energieabhängigkeit der Effektivität und Reinheit zeigt, daß die π^0 -Identifikation im Energiebereich von 1-9 GeV, in dem ca. 75% aller Signalislands liegen (vgl. Abb. 10.8 links), unter Berücksichtigung der Informationen des HES deutlich verbessert werden kann. In diesem Energiebereich konnten Effektivität und Reinheit bei jeweils konstantem Wert der anderen Größe von 76% im Vergleich zur π^0 -Identifikation nur mit CAL-Informationen im Mittel um

$$\Delta\epsilon = 34,4 \pm 3,3\% \tag{10.6}$$

$$\Delta R_1 = 7,7 \pm 3,1\% \tag{10.7}$$

verbessert werden. Der HES liefert in diesem Energiebereich zusätzliche Informationen zur π^0 -Erkennung, da die Signale der beiden Photonen eines π^0 -Zerfalls im Unterschied zum Kalorimeter in vielen Fällen noch getrennt werden können. Oberhalb einer Islandenergie von 9 GeV unterscheiden sich Effektivität und Reinheit beider Netzwerk im Rahmen ihrer Fehler nicht. Bei diesen Energien liefert der HES nur noch in wenigen Fällen zusätzliche Informationen zur π^0 -Identifikation, da die beiden Photonen auch im HES kaum noch als zwei separate Objekte gemessen werden können (vgl. Abb. 6.1).

Während der mittlere Anteil der π^0 -Energie an der Gesamtenergie der einzelnen Islands mit steigender Islandenergie abnimmt (vgl. Abb. 10.8 rechts), steigen Effektivität und Reinheit der π^0 -Identifikation oberhalb von Islandenergien von 5 GeV an. Die höchsten Werte beider Größen werden bei Energien oberhalb von 10 GeV erreicht. Die Unterscheidung der Signale der π^0 -Mesonen von Untergrundeignissen gelingt oberhalb von 10 GeV besser als bei kleineren Energien. Der Untergrund oberhalb von 10 GeV wird zu etwa 80% von Positronen der tiefinelastischen Streuung dominiert. Der Anstieg von Effektivität und Reinheit deutet darauf hin, daß die Signale der Positronen der tiefinelastischen Streuung leichter von denen der π^0 -Mesonen zu unterscheiden sind als die des hadronischen Untergrundes bei tieferen Energien.

10.5 Analyse der Eingabeinformationen

Zum besseren Verständnis der Auswirkung der verschiedenen Eingabeinformationen auf die π^0 -Identifikation wurde eine Hauptachsentransformation der Eingabeinformationen

durchgeführt. Dabei wurden die Eingabeinformationen in ein neues Koordinatensystem transformiert, in dem die neuen Koordinaten nach ihrer Varianz geordnet sind. Die Koordinatentransformation wurde so durchgeführt, daß die Koordinaten kleiner Varianz die kleinstmögliche Varianz besitzen. Sie besteht aus einer Translation zum Schwerpunkt der einzelnen Koordinaten und einer Rotation. Die Rotation ist durch die Matrix der Eigenvektoren der Dispersionsmatrix bestimmt [31]. Die Komponenten jedes Eigenvektors definieren eine Linearkombination der ins Schwerpunktsystem verschobenen ursprünglichen Eingabeinformationen. Diese Linearkombinationen sind die neuen Koordinaten. Der Eigenvektor mit dem höchsten Eigenwert definiert die Linearkombination der ursprünglichen Eingabeinformationen, die am meisten Informationen der Musterklasse enthält. Der Eigenwert des Eigenvektors ist das Quadrat der Varianz dieser neuen Koordinate. Die Eigenwerte σ_{Signal}^2 der durch die Hauptachsentransformation der Eingabeinformationen der Signalislands bestimmten 77 Eigenvektoren sind in Abb. 10.9 I dargestellt. Die so bestimmte Koordinatentransformation wurde anschließend auf die Untergrundislands angewendet. Die Quadrate der berechneten Varianzen der Untergrundislands bzgl. der neuen Koordinaten sind ebenfalls in Abb. 10.9 I eingezeichnet. Für alle 77 Eigenvektoren wurden die Differenz Δ und die relative Differenz δ nach (10.8) und (10.9) berechnet.

$$\Delta = |\sigma_{Signal}^2 - \sigma_{Untergrund}^2| \quad (10.8)$$

$$\delta = \frac{\sigma_{Signal}^2 - \sigma_{Untergrund}^2}{\max(\sigma_{Signal}^2, \sigma_{Untergrund}^2)} \quad (10.9)$$

In Abb. 10.9 II bzw. III ist Δ bzw. δ dargestellt. Die Verteilung der Eigenwerte zeigt für Signal- und Untergrundislands einen ähnlichen Verlauf. In beiden Fällen sind die ersten Eigenwerte in der Größenordnung 1, während der Wert der weiteren Eigenwerte zu einer Größenordnung von 10^{-4} absinkt. Die Differenz Δ der ersten beiden Eigenwerte ist um mindestens eine Größenordnung größer ist bei den anderen Eigenwerten. Dadurch wird nahegelegt, daß nur die ersten beiden Linearkombinationen signifikanten Anteil an der Separation der beiden Musterklassen haben. Die relative Differenz δ hat bei allen neuen Koordinaten dieselbe Größenordnung, was darauf hindeutet, daß trotz der großen Unterschiede in Δ alle Linearkombinationen notwendige Informationen zur π^0 -Identifikation enthalten.

Um Aussagen über den Einfluß der einzelnen Linearkombinationen auf die π^0 -Identifikation machen zu können, wurden *feed forward*-Netzwerke mit einer unterschiedlichen Anzahl dieser Linearkombinationen trainiert und die Ergebnisse verglichen. Dabei wurden beim Training mit l ($l=1..77$) der 77 Linearkombinationen stets die l Linearkombinationen der größten Varianz verwendet. Für unterschiedliche Anzahl der verwendeten Linearkombinationen wurden die größten Werte der Effektivität und Reinheit bestimmt, bei denen diese beiden Größen im Rahmen ihrer statistischen Fehler identisch waren.

In Abbildung 10.10 sind die Mittelwerte Ψ von Effektivität und Reinheit gegen die Anzahl der zum Training der *feed forward*-Netzwerke verwendeten Linearkombinationen aufgetragen. Als Vergleichswert wurde $\Psi_{max} \pm \sigma_{max}$ des mit CAL- und HES-Informationen erreichten Mittelwertes Ψ_{max} von Effektivität und Reinheit (siehe Gleichung (10.3)) eingezeichnet. Bei Verwendung der vier Linearkombinationen der größten Varianz der Signalislands wurde Ψ_{max} innerhalb der statistischen Fehler reproduziert. Diese Linearkombinationen

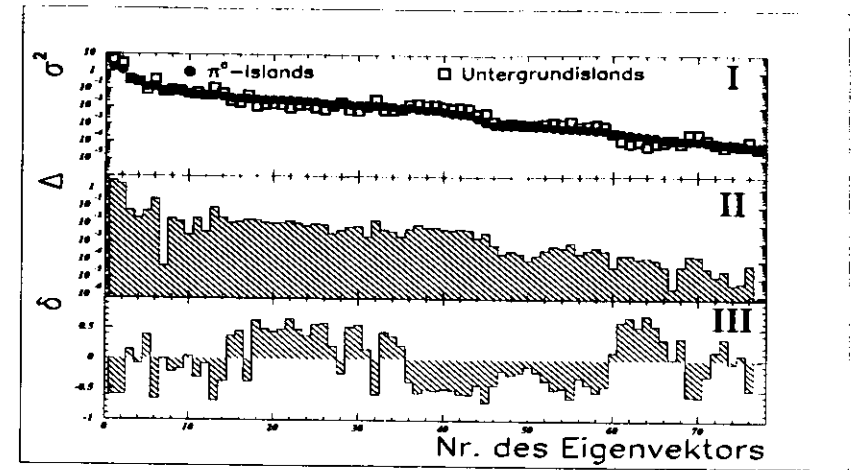


Abbildung 10.9: Ergebnis der Hauptachsentransformation der Eingabedaten aus CAL und HES. Bild I zeigt die Quadrate der Varianzen der Signal- und Untergrundislands bzgl. der aus der Hauptachsentransformation der Signalislands bestimmten Koordinatentransformation. In Bild II bzw. III sind die Differenz Δ bzw. die relative Differenz δ der Quadrate der Signal- und Untergrundislands dargestellt.

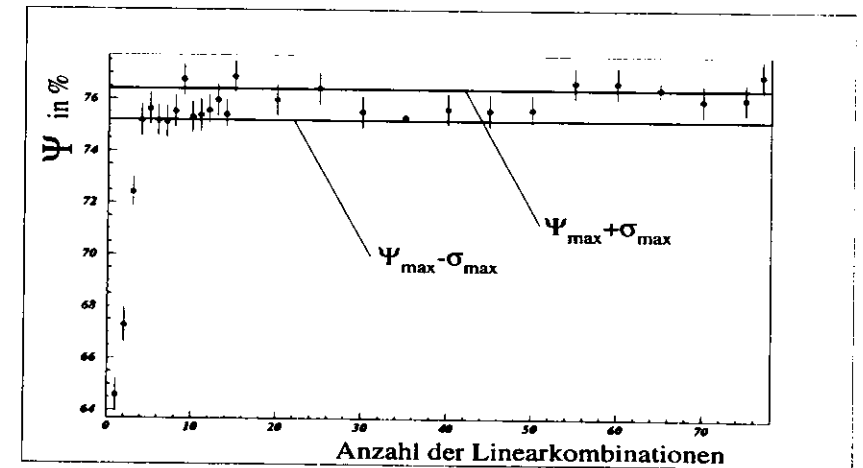


Abbildung 10.10: Mittelwert Ψ von Effektivität ϵ und Reinheit R_i in Abhängigkeit von der Anzahl der verwendeten Linearkombinationen der ursprünglichen Eingabeinformationen. Die dargestellten Fehler entsprechen der Varianz des Mittelwertes von Ψ . Zum Vergleich ist $\Psi_{max} \pm \sigma_{max}$ des mit allen 77 Informationen des CAL und HES erreichten Mittelwertes Ψ_{max} von Effektivität und Reinheit eingezeichnet.

enthalten demnach alle zur π^0 -Identifikation notwendigen Informationen. Nur wenige der Koeffizienten der einzelnen Linearkombinationen hatten einen Wert in der Größenordnung 1, während alle anderen um mindestens eine Größenordnung niedriger waren. Da alle ursprünglichen Eingabeinformationen auf dieselbe Größenordnung normiert wurden (vgl. Kapitel 8.4), wurden die Werte der einzelnen Linearkombinationen nur durch wenige der ursprünglichen Eingabeinformationen dominiert. Über die dominanten Koeffizienten der vier Linearkombinationen der größten Varianz der Signalislands wurden die 12 zur π^0 -Identifikation signifikanten ursprünglichen Eingabeinformationen bestimmt. Diese Eingabeinformationen bestanden für ein Island aus

- den Energiedepositionen in den vier Photovervielfachern der beiden EMC-Zellen des zentralen Turmes des Islands,
- den Energiedepositionen in den zwei Photovervielfachern der HAC-Zelle des zentralen Turmes des Islands,
- der Anzahl der dem Island zugeordneten HES-Isles,
- den Energiedepositionen in den zentralen Dioden der beiden energiereichsten Isles,
- dem Abstand dieser beiden Isles in der Ebene des HES,
- sowie $\cos(\delta_1)$ und $\cos(\delta_2)$.

Der Maximalwert der Effektivität und Reinheit konnte auch bei Verwendung weiterer Linearkombinationen nicht weiter gesteigert werden, sondern stimmte im Rahmen der statistischen Fehler mit dem unter Verwendung aller 77 ursprünglichen Eingabeinformationen des CAL und HES bestimmten Wert von Ψ_{\max} überein.

10.6 Ausblick

Die Identifikation von neutralen Pionen durch *feed forward*-neuronale Netzwerke konnte unter Verwendung von Informationen des CAL und HES mit guter Effizienz durchgeführt werden. Der Hadron-Elektron-Separator lieferte bei Energien von 1-9 GeV zusätzliche Informationen zum Kalorimeter.

In dieser Analyse konnten jedoch noch nicht alle Möglichkeiten des Einsatzes von *feed forward*-neuronalen Netzwerken zur π^0 -Identifikation behandelt werden. Wie die Hauptachsentransformation der verwendeten 77 Eingabeinformationen zeigt, kann die Identifikation von π^0 -Mesonen durch geeignete Kombination dieser Eingabeinformationen vereinfacht werden. Nichtlineare Kombinationen der Eingabeinformationen bieten daher eine weitere Möglichkeit, die Konvergenz des Trainingsprozesses der neuronalen Netzwerke zu beschleunigen. Auch das Training von neuronalen Netzwerken bei unterschiedlichen Energien kann die Effizienz der π^0 -Identifikation verbessern. Die einzelnen Netzwerke sind dadurch unter Umständen besser in der Lage, die Signatur der Untergründereignisse in

den verschiedenen Energiebereichen von denen der π^0 -Mesonen zu unterscheiden. Weitere Möglichkeiten zur Steigerung der erreichten Effizienz liegen in der Verwendung bisher nicht berücksichtigter Informationen des CAL und HES bzw. anderer Detektorkomponenten wie z.B. der Spurkammern. Außerdem kann versucht werden, die Konvergenz des Trainingsprozesses der neuronalen Netzwerke zu verbessern. Ansatzpunkte hierfür bieten die verwendeten Methoden zur Veränderung der Temperatur t , der Minimalschrittweite γ und der Gewichte und Schwellenwerte des Netzwerkes.

Kapitel 11

Zusammenfassung

In dieser Arbeit wurde die Identifikation der Zerfälle $\pi^0 \rightarrow \gamma\gamma$ untersucht, bei denen die beiden Photonen im Kalorimeter des ZEUS-Detektors nicht mehr separat nachgewiesen werden können, sondern nur die Summe ihrer Energiedeposition gemessen wird.

Zur Objektdefinition wurden die einzelnen Kalorimeterzellen durch einen Rekonstruktionsalgorithmus zu sogenannten Islands zusammengefaßt. Dazu wurde zunächst für jeden Kalorimeterturm die Summe der Energiedepositionen in allen Kalorimeterzellen berechnet. Anschließend wurden die Kalorimetertürme mit einem lokalen Maximum der Energie bestimmt und alle anderen Kalorimetertürme diesen Türmen zugeordnet. Dabei wurde der Energiegradient zwischen den einzelnen Kalorimetertürmen berücksichtigt. Es wurden die Kalorimetertürme als benachbart angesehen, die mindestens eine gemeinsame Kante besaßen. Alle Kalorimeterzellen einer so konstruierten Gruppe von Kalorimetertürmen bildeten zusammen ein Island.

Den Islands wurden Dioden des im Kalorimeter integrierten Hadron-Elektron-Separators zugeordnet, da die bessere Granularität dieser Detektorkomponente die Trennung der Signale der beiden Photonen eines π^0 -Zerfalls auch bei höheren π^0 -Energien als im Kalorimeter erlaubt. Diese Dioden wurden nach demselben Prinzip wie die Kalorimetertürme zu verschiedenen Gruppen, den sogenannten Isles zusammengefaßt. Ein Kalorimeterisland und die ihm zugeordneten Isles bildeten zusammen ein Objekt.

Alle Objekte mit Energiedeposition von den beiden Photonen mindestens eines π^0 -Zerfalls wurden als Signal-, alle anderen Objekte als Untergrundobjekte bezeichnet. An simulierten (Monte Carlo-) Ereignissen der tiefunelastischen Positron-Proton-Streuung wurde die Effizienz der Identifikation der beschriebenen π^0 -Zerfälle studiert. Für jedes Objekt wurden 55 Informationen des Kalorimeters und 22 Informationen des Hadron-Elektron-Separators berechnet. Diese Informationen wurden als Eingabeinformationen von *feed forward*-neuronalen Netzwerken verwendet. Die π^0 -Identifikation mit Informationen des Kalorimeters wurde mit der π^0 -Identifikation unter Verwendung der Informationen des Kalorimeters und des Hadron-Elektron-Separators verglichen. In beiden Fällen wurde die Anzahl der verborgenen Knoten der Netzwerke variiert. Zur Analyse der π^0 -Identifikation und zur iterativen Bestimmung der Parameter der neuronalen Netzwerke wurden die jeweils gleichen Anzahlen von Signal- und Untergrundobjekten verwendet. Die maximale im Kalorimeter gemessene Energie der verwendeten Objekte betrug 14 GeV.

Bei der zur Bestimmung der Parameter der *feed forward*-Netzwerke notwendigen Minimierung einer quadratischen Fehlerfunktion wurde die Abhängigkeit dieser Funktion von

der Temperatur t der Netzwerke ausgenutzt. Mit diesem Verfahren wurde bei zufälligen Anfangswerten der einzelnen Parameter eine hohe Übereinstimmung der Ergebnisse erzielt.

Die Effizienz der π^0 -Identifikation durch die verschiedenen Netzwerke wurde durch die Effektivität ε und die Reinheit R_1 verglichen. Dabei wurde ε als der Prozentsatz der korrekt identifizierten Signalobjekte und R_1 als der Anteil der Signalobjekte an den durch die neuronalen Netzwerke als solche klassifizierten Objekten definiert.

Die höchste Effizienz der π^0 -Identifikation wurde bei Verwendung der Eingabeinformationen des Kalorimeters und des Hadron-Elektron-Separators erzielt. Dazu waren acht verborgene Knoten notwendig. Weitere verborgene Knoten verbesserten die π^0 -Identifikation nicht. Bei einer festen Reinheit von 70% wurde eine Effektivität von $87,0 \pm 0,6\%$ erreicht; bei einer festen Reinheit von 80% noch eine Effektivität von $62,6 \pm 0,9\%$. Bei Beschränkung auf die Daten des Kalorimeters reduzierten sich diese Werte um $4,5 \pm 1,1\%$ bzw. $51,2 \pm 1,1\%$ bei fester Reinheit von 70% bzw. 80%.

Unter der Bedingung, daß Effektivität und Reinheit innerhalb ihrer statistischen Fehler identisch waren, wurde unter Berücksichtigung der Informationen des Hadron-Elektron-Separators ein Mittelwert Ψ beider Größen von

$$\Psi = 75,8 \pm 0,6\%$$

erzielt. Gegenüber der Beschränkung auf die Informationen des CAL wurde eine Steigerung um $3,1 \pm 0,8\%$ erreicht.

Im Energiebereich von 1-9 GeV, in dem ca. 75% aller π^0 -Signale auftraten, wurde die π^0 -Identifikation unter Berücksichtigung der Informationen des Hadron-Elektron-Separators im Vergleich zur Beschränkung auf Informationen des Kalorimeters um

$$\Delta\varepsilon = 34,4 \pm 3,3\% \quad \text{bei } R_1 = 76\%$$

$$\Delta R_1 = 7,7 \pm 3,1\% \quad \text{bei } \varepsilon = 76\%$$

verbessert. Außerhalb dieses Energiebereiches war die Effizienz der π^0 -Identifikation mit und ohne Berücksichtigung der Informationen des Hadron-Elektron-Separators im Rahmen der statistischen Fehler identisch.

Zur Untersuchung der Auswirkung der einzelnen Eingabeinformationen auf die π^0 -Identifikation wurde eine Hauptachsentransformation der verwendeten Eingabeinformationen durch Bildung entsprechender Linearkombinationen durchgeführt. Die Effektivität und die Reinheit bei Verwendung von unterschiedlich vielen dieser Linearkombinationen als Eingabeinformationen der neuronalen Netzwerke wurden verglichen. Bereits bei Verwendung von vier Linearkombinationen wurden die mit allen 77 Daten des Kalorimeters und des Hadron-Elektron-Separators erreichten Werte der Effektivität und Reinheit im Rahmen der statistischen Fehler reproduziert. Aus den Koeffizienten dieser vier Linearkombinationen wurden die zwölf dominierenden Informationen des Kalorimeters und des Hadron-Elektron-Separators bestimmt.

Diese 12 Informationen bestanden für ein Island aus

- den Energiedepositionen in den vier Photovervielfachern der beiden EMC-Zellen des Turmes der maximalen Energie des Islands,
- den Energiedepositionen in den zwei Photovervielfachern der HAC-Zelle des Turmes der maximalen Energie des Islands,
- der Anzahl der dem Objekt zugeordneten Isles im Hadron-Elektron-Separator,
- den Energiedepositionen in den zentralen Dioden der beiden energiereichsten Isles,
- dem Abstand dieser beiden Isles in der Ebene des Hadron-Elektron-Separators,
- sowie aus den mit den Positionen dieser beiden Isles und der Position des Wechselwirkungspunktes bestimmten Relativwinkeln zur Flächennormalen des Hadron-Elektron-Separators.

Bei Verwendung weiterer Eingabeinformationen stimmten Effektivität und Reinheit der π^0 -Identifikation innerhalb der statistischen Fehler mit ihren unter Verwendung aller 77 ursprünglichen Informationen des CAL und HES bestimmten Werten überein.

Literaturverzeichnis

- [1] **J. Grosse-Knetter**, *Measurement of Elastic ω Photoproduction*, ZEUS-Note 95-089, Hamburg, 1995
- [2] **L.S. Brown, D.L. Nordstrom**, *Physical Review D Particles And Fields*, Vol. 50, 3rd Series, American Institute of Physics, New York, 1994
- [3] **I. Fleck**, *Nachweis von Elektronen aus der ep-Streuung mit dem Hadron-Elektron-Separator im ZEUS-Experiment*, Interner Bericht DESY F35D-94-09, Hamburg, 1994, Dissertation an der Universität Hamburg
- [4] **M. Kunze**, *Application of artificial neural networks in the analysis of multi-partical data*, Proceedings of the CORINNE II conference, Nantes, World Scientific, 1994
- [5] **HERMES Kollaboration**, *Technical Design Report*, The HERMES Collaboration, July 1993
- [6] **HERA-B Kollaboration**, *HERA-B Proposal*, DESY-Note 94-02, Hamburg, 1994
- [7] **P. Schmüser**, *Zur Fertigstellung von HERA*, Physikalische Blätter, Band 46, Nr.12, Dezember 1990, S.470-474
- [8] **G. Wolf**, *HERA Physics*, DESY-Note 94-022, Hamburg, 1994
- [9] **S. Bentvelsen, J. Engelen, P. Kooijman**, *Reconstruction of (x, Q^2) and extraction of structure functions in neutral current scattering at HERA*, W. Buchmüller, G. Ingelman (Editoren), Physic At HERA, Band I, S. 23 ff, Hamburg, Oktober 1991
- [10] **ZEUS Collaboration**, *The ZEUS detector, Statusreport 1993*, DESY-Note 93-05, Hamburg, 1993.
- [11] **M. Costa, C. Nemoz**, *The ZEUS Leading Proton Spectrometer Automatic Pot Insection Procedure*, ZEUS-Note 95-162, Hamburg, 1995
- [12] **ZEUS FNC GROUP**, *Design of the ZEUS Forward Neutron Calorimeter*, ZEUS-Note 93-123, Hamburg, 1993
- [13] **A. Dwurazny**, *Experimental Study Of Electron-Hadron Separation In Calorimeters Using Silicium Diodes*, Nucl. Instr. and Meth. A277 (1989) 176

- [14] M. Kasprzak et al, *Veto Wall for the ZEUS Detector*, ZEUS-Note 92-072, Hamburg, 1992
- [15] H. Bethe, W. Heitler, *Proc. Roy. Soc.*, A146 (1934) 83
- [16] J. Krüger, *The uranium scintillator calorimeter for the ZEUS detector at the electron-proton collider HERA*, Habilitationsschrift, Hamburg, 1992
- [17] H. Park, *Experience with the ZEUS Calorimeter*, ZEUS-Note 94-159, Hamburg, 1994
- [18] H. Brückmann et al, *Hadron Sampling Calorimetry, A Puzzle of Physics*, Nucl. Instr. and Meth. A263 (1988), 136-149
- [19] P. de Jong, *Status of the Uranium Calorimeter Reconstruction Software*, ZEUS-Note 92-019, Amsterdam, 1992
- [20] U. Wollmer, *Studium der Ortsauflösung im Hadron-Elektron-Separator des ZEUS-Experiments*, Interner Bericht, DESY F35D-95-09, Hamburg, 1995, Diplomarbeit an der Universität Hamburg
- [21] E. Schöneburg, N. Hansen, A. Gawelczyk, *Neuronale Netzwerke - Einführung, Überblick und Anwendungsmöglichkeiten*, Markt und Technik Verlag, 1990
- [22] G. Czihak, H. Langer, H. Ziegler, *Biologie*, Springer-Verlag, 1981, S. 419 ff
- [23] T. Kohonen, *Self-organized Formation of Topologically Correct Feature Maps*, Biological Cybernetics, Vol. 43, S. 59-69, 1982
- [24] John Hertz, Anders Krogh, Richard G. Palmer, *Introduction to the Theory of Neural Computation*, Allan M. Wylde, 1991.
- [25] R. Rojas, *Theorie der neuronalen Netz*, Springer-Verlag, Berlin Heidelberg, 1993
- [26] R. Sinkus, *A novel approach to error function minimization for feedforward neural networks*, DESY-Note 94-182, 1994.
- [27] W.A. Stahel, *Statistische Datenanalyse*, Vieweg, 1995
- [28] M.D. Richard, R.P. Lippmann, *Neural Computation*, 3:461, 1991
- [29] V. Blobel, E. Lohrmann, *Statistische und numerische Methoden der Datenauswertung*, Vorlesungsskript, Universität Hamburg, WS 1993/94
- [30] R. Sinkus, H. Abramowicz, A. Caldwell, *Neural network based electron identification in the ZEUS calorimeter*, DESY-Note 95-054, 1995
- [31] H. Wind, *Principal Component Analysis And Its Applications To Track Finding*, Formulae and Methods in Experimental Data Evaluation, Vol.3, K1-K16, R.K. Bock, 1984

Danksagung

Die Erstellung dieser Arbeit hat mein Interesse an der Physik, insbesondere der Teilchenphysik weiter gesteigert. Das verdanke ich vor allem Prof. Dr. E. Lohrmann, der mich durch den Vorschlag des Themas dieser Arbeit in die faszinierende Welt der Hochenergiephysik und der neuronalen Netzwerke eingeführt hat.

Ich möchte mich herzlich bei allen Mitarbeitern von ZEUS und DESY bedanken, die während der Anfertigung dieser Arbeit stets hilfsbereit auf meine Fragen und Probleme eingegangen sind. Spezieller Dank gebührt dabei P. Göttlicher für die mit ihm geführten Diskussionen über Statistik, die Hilfestellung bei der Handhabung von LATEX und -zusammen mit W. Zeuner - die Anmerkungen über Form und Inhalt meiner Arbeit.

Außerdem danke ich H. Landfermann und U. Rehder für die oft kurzfristig durchgeführten Korrekturen meines Textes, sowie meinem Betreuer R. Sinkus für die Einführungen in die Theorie der neuronalen Netzwerke und in die ZEUS-Rechnersysteme.

In guter Erinnerung werden mir die Arbeit und Mitarbeiter der PR-Abteilung bleiben. Besonders die Informationstreffen und die Arbeitsverteilungen bildeten durch ihre lockere Atmosphäre eine willkommene und interessante Abwechslung.

Zuletzt - damit es in besser in Erinnerung bleibt - danke ich meiner Freundin Michaela und meinen Eltern für die während des Studiums und der Anfertigung dieser Arbeit gewährte Unterstützung. Ohne diese Unterstützung hätte die Diplomarbeit in der vorliegenden Form nicht fertiggestellt werden können.

Erklärung

Hiermit versichere ich, daß ich die vorliegende Arbeit unter Angabe aller verwendeten Quellen und Hilfsmittel selbständig angefertigt habe.

Hamburg, im März 1996



(Ulrich Fricke)

