Entwicklung eines Triggers für diffraktiv erzeugte J/ψ-Mesonen mit Hilfe Neuronaler Netze bei H1

Diplomarbeit

vorgelegt von
Ronald Deckert
aus Bad Oldesloe

II. Institut für Experimentalphysik
Universität Hamburg

Hamburg
November 1997

Erklärung

Hiermit versichere ich,
diese Arbeit selbständig
und nur unter Verwendung
der angegebenen Quellen
verfaßt zu haben.

Gewidmet meiner Familie, insbesondere meinem Vater, Marie-Luise und Dirk

Inhaltsverzeichnis

Kapitel 1

Einleitung

Seit Menschengedenken stehen Fragen nach den Hintergründen des menschlichen Daseins und der Existenz der Dinge im Raum. Menschen gehen diese Fragen auf unterschiedlichste Art und Weise an, ob religiös, philosophisch oder mit Hilfe der Naturwissenschaften, wie der Physik. Früh wurde in der Physik die Vorstellung von unteilbaren Konstituenten der Materie entwickelt. Die intellektuelle Leistung besteht darin, den Eindruck, den man mit dem bloßen Auge von jeder Art von Materie erlangt, gedanklich zu überwinden und etwas *Unsichtbares* zu postulieren und nachzuweisen.

Die mikroskopische Welt wird nach den bahnbrechenden Entdeckungen unseres Jahrhunderts durch das Standardmodell der Elementarteilchenphysik beschrieben. Dies beinhaltet sechs Quarks und sechs Leptonen als elementare Teilchen, zu denen sich die Austauschfeldquanten γ, Z^0, $W^\pm$ und die acht Gluonen gesellen, welche die elektromagnetische, die schwache und die starke Wechselwirkung vermitteln. Ein Austauschteilchen für die Gravitation, das Graviton, ist noch nicht nachgewiesen worden.

Mit dem sich am Speicherring HERA befindlichen H1-Detektor werden Positron-Proton-Reaktionen vermessen und analysiert und insbesondere die Wechselwirkung der elementaren Bestandteile von Protonen - Quarks und Gluonen - untersucht. Als Grundlage dieser Arbeit dienen in Positron-Proton-Reaktionen erzeugte J/ψ-Mesonen, welche in zwei Myonen zerfallen. Mit Hilfe dieser Reaktionen ist es möglich, die Gluondichte des Protons zu messen. Je mehr J/ψ-Mesonen man im Detektor aufzeichnet, desto besser wird man die Verteilung der Gluonen im Proton verstehen.

Die Identifikation der J/ψ-Mesonen erfolgt auf elektronischem Wege mit Detektorinformationen aus dem H1-Detektor. Schwierig gestaltet sich mitunter die Trennung der Reaktionen aus Zerfällen der J/ψ-Mesonen von Untergrundreaktionen, die nicht von ep-Reaktionen stammen. Mit über die Jahre zunehmender Luminosität und damit auch mit der Zunahme sowohl der Untergrund- wie auch der ep-Reaktionen ist es von Interesse die elektronischen Verfahren zur Trennung dieser beiden Ereignisklassen zu verbessern. Dadurch erhöht man bei annähernd konstanter Rate an aufgezeichneten Ereignissen insgesamt den Anteil insbesondere von J/ψ-Mesonen an der Menge aller aufgezeichneten Daten.

Zur Verbesserung dieser Verfahren bedient man sich bei H1 auf einer Stufe des mehrstufigen Triggers der Technik der Neuronalen Netze. Diese computerbasierte Methode gründet sich auf Algorithmen, wie man sie im Zusammenspiel von Nervenzellen im Gehirn erforscht hat. Dabei nutzen diese lernfähigen Rechenverfahren hochdimensionale Korrelationen zwischen Detektorinformationen, welche die zu unterscheidenden Ereignisse charakterisieren. Diese Informationen können damit möglicherweise effektiver genutzt werden als wenn man Schnitte in wenigen Informationen anbringt, wie es traditionell geschieht.

Diese Arbeit beschäftigt sich mit der Entwicklung von neuronalen Netzen zur Erkennung von J/ψ-Mesonen im H1-Detektor mit dem Ziel, einen Beitrag zur Schaffung einer Datengrundlage zur vertieften Analyse des Protons zu leisten.

Gliederung der Arbeit

Im **zweiten Kapitel** erfolgt eine Darstellung vom Aufbau des Speicherringes HERA und des H1-Detektors. Nach einem Abschnitt über allgemeine Ereigniskinematik wird die Physik der J/ψ-Mesonen behandelt. Das zweite Kapitel abschließend folgt ein Abschnitt zur Einführung in das Triggerkonzept von H1. Das **dritte Kapitel** behandelt die benutzte Methodik der Neuronalen Netze vor dem Hintergrund biologischer Analogien, sowie der speziellen Anwendung bei H1. Der Kern der Arbeit findet im **vierten Kapitel** mit der Beschreibung der Durchführung der Untersuchungen und den Ergebnissen des Netztrainings seine Darstellung. Die Zusammenfassung der Ergebnisse und der Ausblick befinden sich im **fünften Kapitel**.

Kapitel 2

Das H1-Experiment

Das H1-Experiment ist eines der Hochenergie-Experimente am Deutschen Elektronensynchrotron DESY in Hamburg. Das DESY wurde 1959 als Gesellschaft bürgerlichen Rechts in Hamburg gegründet. Seit 1992 ist das Institut für Hochenergiephysik Zeuthen dem DESY angegliedert. Neben Positron-Proton Reaktionen, die an den Experimenten H1 und ZEUS am Speicherring HERA vermessen werden, wird am Desy das Experiment HERMES, zur Messung der Spinstruktur von Proton und Neutron, betrieben. Das Experiment HERA B, zur Messung der CP-Verletzung im b-Quark-System, befindet sich im Aufbau. Weiterhin ist am DESY mit dem HASYLAB ein Institut für Experimente mit Synchrotronstrahlung geschaffen worden.

2.1 Der Speicherring HERA

Die Hadron-Elektron-Ring-Anlage HERA wurde zur Speicherung von Elektronen (Positronen) und Protonen in zwei getrennten Speicherringen gebaut. Durch die Analyse der Kollisionen dieser Teilchen kann die Struktur von Proton und Photon im Rahmen des heute gültigen Standardmodells aufgeklärt, und dieses Modell selbst überprüft werden. HERA wurde nach sechs Jahren Bauzeit 1990 fertiggestellt, die erste Elektron-Proton-Kollision erfolgte jedoch erst am 19.Oktober 1991. Die Detektoren H1 und ZEUS haben den Betrieb im Jahre 1992 aufgenommen.
In einem Speicherring läßt man zwei einander entgegengesetzt umlaufende Teilchenstrahlen an verschiedenen Wechselwirkungspunkten im Ring kollidieren. Hierzu benötigt man bei Reaktionen von Teilchen mit deren Antiteilchen nur ein Strahlrohr, da die im Speicherring bahnbestimmend und beschleunigend wirkenden elektromagnetischen Felder invariant gegenüber der CP-Transformation sind. Im selben elektromagnetischen Feld kehrt sich die Bewegungsrichtung des Antiteilchens gegenüber der des dazugehörigen Teilchens um. Beispiele hierfür sind die e^+-e^--Speicherringe DORIS und PETRA am DESY und der p-$\bar{p}$-Speicherring TEVATRON am FERMILAB bei Chicago.

Zur Speicherung von Positronen und Protonen bei HERA, deren Kollisionen bei den Experimenten H1 und ZEUS untersucht werden, mußten zwei separate Strahlrohre, eines für Positronen und eines für Protonen, installiert werden. Seit Juli 1994 wird HERA statt mit Elektronen mit Positronen betrieben, um die Lebensdauer des Leptonenstrahles zu erhöhen. Die Positronen werden über die HERA-Vorbeschleuniger LINAC II, DESY II und PETRA II auf 14 GeV vorbeschleunigt und bei HERA injiziert. Sie erreichen hier ihre Endenergie von 27.5 GeV. Die Protonen werden durch LINAC III, DESY III und PETRA II auf eine Energie von 40 GeV gebracht und bei HERA weiter bis hin zu 820 GeV beschleunigt. HERA läuft somit bei einer Schwerpunktenergie des Proton-Positron-Systems von $\sqrt{s} \approx 300\ GeV$.

Speicherringe haben gegenüber Festtargetexperimenten den entscheidenden Vorteil, daß höhere Schwerpunktenergien erreicht werden können, da die Schwerpunktenergie für hohe Strahlenergien näherungsweise mit der Wurzel des Produktes der Strahlenergien skaliert. Ruht eines der Teilchen, so ist diese Abhängikeit nur noch durch die Wurzel der Strahlenergie des nicht ruhenden Teilchens gegeben. Es wäre z.B. die bei ruhendem Proton für gleiche Schwerpunktenergie $\sqrt{s} \approx 300\ GeV$ nötige Positronenenergie $E_e(\vec{P}_p = 0) \simeq 50\ TeV$. Man kann daher bei einem Speicherring mit weit weniger technischem Aufwand, und damit kostengünstiger, Teilchenreaktionen hoher Schwerpunktenergie erzeugen als bei einem Festtargetexperiment.

Das maximale Viererimpulsübertragsquadrat beträgt bei HERA $Q^2 \approx 10^5\, GeV^2$. Diese Größe wird uns in Abschnitt 2.3 wieder begegnen. Sie liegt bei großer Bjørken-Variable x (siehe Abschnitt 2.3) zwei Größenordnungen über den Werten der hierbei vor HERA richtungsweisenden Experimente.

Eine Positronen-Füllung von HERA hat eine Lebensdauer von bis zu 12 Stunden. Durch die beschleunigenden Hochfrequenz-Kavitäten ist der Energieverlust der Positronen auf Grund der Synchrotronstrahlung auszugleichen. Dies ist für die Positronen der energiebegrenzende Effekt. Die Lebensdauer der Protonen ist mit über 100 Stunden sehr viel länger als die der Positronen. Die Energie der Protonen ist durch das zur Ablenkung notwendige Magnetfeld der supraleitenden Magnete von 4.7 T beschränkt. Um zu akzeptablen Raten von physikalischen Ereignissen zu gelangen, ist HERA für gleichzeitig umlaufende 210 Teilchenbündel (*bunches*) bei jeweils einem Protonenstrom pro Teilchenbündel von 760 μA und einem Positronenstrom pro Teilchenbündel von 290 μA ausgelegt. Man erhält eine Kreuzungsrate der Teilchenbündel von 10.4 MHz. Das bedeutet, daß sich alle 96 ns zwei Teilchenbündel im Bereich des nominellen Wechselwirkungspunktes durchqueren. Diese Zeitspanne von 96 ns ist die Grundzeiteinheit der sogenannten HERA-Uhr, welche sämtliche am Experiment zur Datenauslese installierte Elektronik synchronisiert. Die Ausdehnung der Teilchenbündel senkrecht zur Bewegungsrichtung der Teilchen ist klein, so daß hier die Wechselwirkungszone gut determiniert erscheint. Auf Grund der Ausdehnung der Teilchenbündel in Richtung des Strahlrohres liegt die Wechselwirkungszone in dieser Richtung in einem Bereich von ± 50 cm um den nominellen Wechselwirkungspunkt [1].

2.2 Der H1-Detektor

Der H1-Detektor ist neben dem ZEUS-Detektor für die Vermessung von Positron-Proton-Reaktionen konzipiert worden. In Abbildung 2.1 ist ein Längsschnitt des Detektors entlang der Strahlröhre gezeigt. Das bei der H1-Kollaboration verwendete Koordinatensystem zur Analyse der Messungen mit dem H1-Detektor ist in Abbildung 2.2 skizziert.

Nach allgemeiner Beschreibung des Detektors wird auf die einzelnen Detektorkomponenten näher eingegangen. Dabei liegt der Schwerpunkt der Darstellung bei den Detektorteilen, die für diese Arbeit von Bedeutung sind.

2.2.1 Überblick

H1 ist ausgelegt für die kinematische Rekonstruktion von Positron-Proton-Reaktionen bei Ereignisaufzeichnungsraten im Bereich von 10 Hz bis 15 Hz. Die Konstruktionsanforderungen lagen insbesondere bei guter Impuls- und Ortsauflösung, in der Teilchenidentifikation und Energiemessung, sowie einer möglichst guten Ausnutzung des gesamten Raumwinkels [1]. Die Ausmaße des Detektors betragen ungefähr 12x15x10 m^3 bei einem Gewicht von 2800 t.

Die asymmetrische Instrumentierung des Detektors geschah mit Hinblick auf die Asymmetrie der zu vermessenden Reaktionen. Das Ruhesystem der kollidierenden Leptonen und Hadronen bewegt sich in Protonflugrichtung. Aus diesem Grunde fällt insbesondere bei inelastischen Reaktionen ein Großteil der Teilchen in diesen sogenannten vorderen Detektorbereich.

Die folgende Beschreibung wird anhand von Abbildung 2.1 veranschaulicht. In dieser Abbildung gelangen die Protonen von rechts oben, die Positronen von links unten in den Detektor. Im Zentralbereich des Detektors befinden sich um die Strahlröhre herum die zentralen Spurkammern, mit Hilfe derer Spuren von geladenen Teilchen rekonstruiert, sowie die Impulse gemessen werden können. Durch die auf die geladenen Teilchen wirkende Lorentzkraft, verursacht durch das anliegende Magnetfeld, werden die Teilchentrajektorien innerhalb des Magneten gekrümmt. Das Magnetfeld wird von der sich an das Flüssig-Argon-Kalorimeter anschließenden Spule geliefert.

Das zwischen Spurkammern und Spule gelegene Flüssig-Argon-Kalorimeter wird mit seinem elektromagnetischen Teil zur Identifikation und Energiemessung von Elektronen benutzt, mit seinem hadronischen Teil zur Identifikation und Energiemessung von Hadronen, insbesondere Pionen und Kaonen. Auch Myonen können als minimalionisierende Teilchen identifiziert werden.

Zur Rückführung des magnetischen Flusses befindet sich an die Magnetspule nach außen anschließend das instrumentierte Eisenjoch. Dieses Eisenjoch enthält in Lagen angeordnete Streamerrohrkammern zur Identifikation von Myonen. Die Myonen gelangen auf Grund ihrer im Vergleich zu Elektronen hohen Masse minimalionisierend durch das Kalorimeter und werden danach mit hoher Wahrscheinlichkeit im Eisen detektiert.

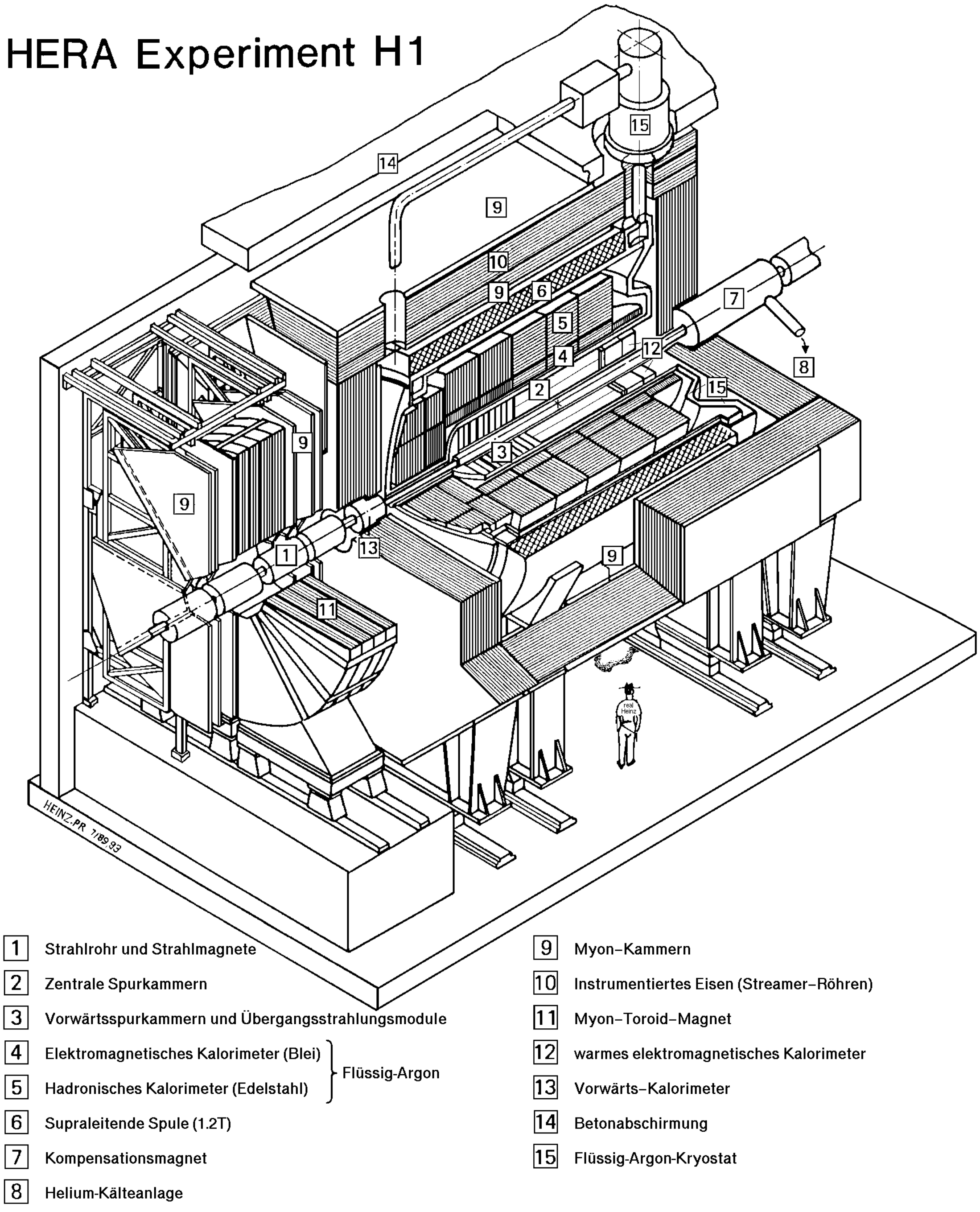


Abbildung 2.1: *Der H1-Detektor.*

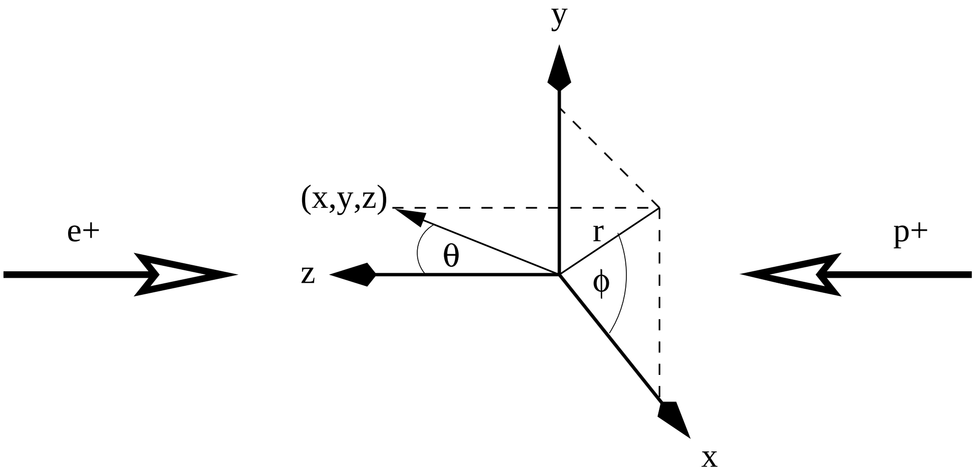


Abbildung 2.2: *Von der H1-Kollaboration verwendetes Koordinatensystem bei H1.*

Eine weitere Aufgabe des Eisenjochs liegt in der Schauerrestmessung von hadronischen Schauern, die nicht auf das Kalorimeter beschränkt bleiben.
Die übrigen, kleineren Detektorkomponenten und weitere Feinheiten werden in den einzelnen Unterkapiteln abgehandelt. Bei der nun folgenden Beschreibung der einzelnen Detektorkomponenten wird insbesondere auf die Spurkammersysteme, das Flüssig-Argon(LAr)-Kalorimeter und das zur Myonendetektion vorgesehene instrumentierte Eisen eingegangen. Das Spurkammersystem und der Myondetektor liefern die Eingabedaten, die für die im späteren vorgestellten Neuronalen Netze benutzt werden.

2.2.2 Das Spurkammersystem

In Darstellung 2.3 sind das zentrale und das vordere Spurkammersystem CTD und FTD gezeigt, von dem jedes für sich getrennt in dem jeweils zu vermessenden Raumwinkelbereich in bezug auf Spurrekonstruktion und Triggereigenschaften optimiert wurde [1].

Der Großteil der Spurvermessung aus Daten des CTD wird von den beiden konzentrischen Driftkammern CJC1 und CJC2 geleistet, welche zylinderförmig um das Strahlrohr angeordnet sind (siehe auch Querschnitt Abbildung 2.4). Diese Kammern sind mit Argon-Kohlendioxid-Methan- und Argon-Ethan-Gemischen gefüllt und besitzen Fassungsvermögen von 1140 bzw. 3080 Litern. Die Kammern sind radial in 30 (CJC1) bzw. 60 (CJC2) Zellen aufgeteilt, wobei jede Zelle von 24 (CJC1) bzw. 32 (CJC2) Signaldrähten durchzogen wird, welche über die gesamte Länge der Kammern von 2.2 m gespannt sind. Wie in Abbildung 2.4 in einem Querschnitt gezeigt wird, sind die Zellen nicht radialsymmetrisch angeordnet, sondern um 30° gegen einen gedachten Mittelpunktstrahl gekippt. Dies geschah zur Optimierung des Spurauflösungsvermögens und zur Verminderung von Doppeldeutigkeiten in bezug auf den Ort der primären Ionisationen bei der Rekonstruktion. Durch die Schrägstellung der Zellen werden auch Spuren von hochenergetischen Teilchen aus Richtung des nominellen Wechselwirkungspunktes

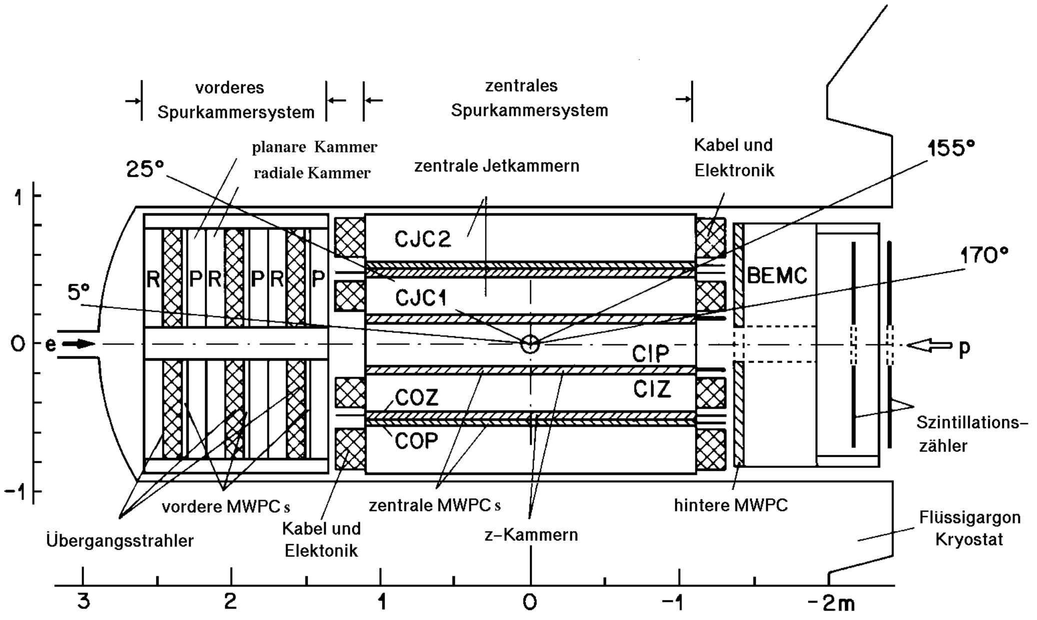


Abbildung 2.3: *Das Spurkammersystem im Längsschnitt.*

an einigen Drähten nahe vorbei gelangen. Dies erhöht die Genauigkeit der Spurvermessung. Die Spuren der Teilchen können bei einer Zeitauflösung von 0.5 ns und einer Ortsauflösung der Spurrekonstruktion von 170 μm in der radialen Ebene und 22 mm in z-Richtung vermessen werden.

Die relativ schlechte Auflösung in der z-Richtung wird verbessert durch die zusätzlich vorhandenen inneren und äußeren Z-Kammern CIZ und COZ, die in den Abbildungen 2.3 und 2.4 dargestellt sind. Man verbessert dadurch die Ortsauflösung der Spurvermessung in z-Richtung auf 300 μm. Die Z-Kammern sind mit 15 (CIZ) bzw. 24 (COZ) Zellen mit je 4 Drahtlagen ausgestattet.

Durch die kombinierte Vermessung der Spuren durch zentrale Spurkammern und Z-Kammern erhält man eine Genauigkeit der Impulsmessung von

$$\frac{\sigma_p}{p^2} < 0.01\,GeV^{-1}. \tag{2.1}$$

In der Auflösung des zur Teilchenidentifikation gemessenen differentiellen Energieverlustes dE/dx werden Werte um $\sigma(dE)$/dE = 10% erreicht.

Insbesondere für Triggerzwecke wurden noch die Vieldrahtproportionalkammern CIP und COP installiert (siehe Bezeichnung MWPC (*Multi-Wire-Proportional-Chamber*) in Abbildung 2.4). Diese bestehen aus jeweils zwei Lagen. Die insgesamt vier Proportionalkammerlagen sind von innen nach außen mit 480, 480, 288 und 288 Auslese-Elektroden ausgestattet. Zur Ermittlung der Triggersignale, die für diese Arbeit von Bedeutung sind, werden die zentralen Proportionalkammern Elektrode für Elektrode

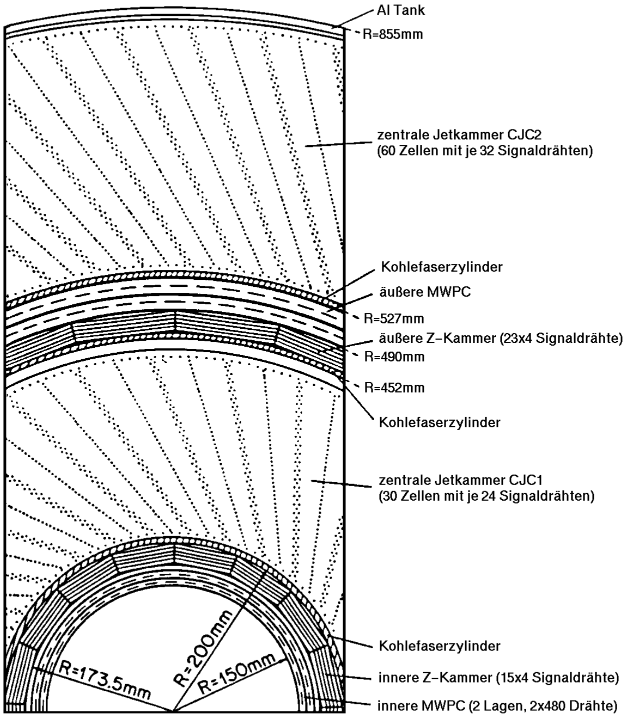


Abbildung 2.4: *Das zentrale Spurkammersystem im Querschnitt.*

ausgelesen. Diese Signale werden mit den Signalen der vorderen Proportionalkammern kombiniert. Hierauf wird in Abschnitt 2.4 näher eingegangen werden.
Das vordere Spurkammersystem FTD besteht aus drei Supermodulen. Jedes dieser Module ist aufgebaut aus einem planaren Modul mit drei Driftkammern zur Impuls- und Spurvermessung, einer Vieldrahtproportionalkammer (FPC) zu Triggerzwecken, einem Übergangsstrahlungsdetektor zur Diskrimination von Elektronen und Pionen und einem radialen Driftkammermodul. Die radialen Module ergänzen die planaren in bezug auf die Ortsauflösung durch eine radiale Signaldrahtspannungsrichtung. Man erhält für die Spurvermessung im vorderen Spurkammersystem eine Ortsauflösung in der radialen Ebene in der gleichen Größenordnung wie im zentralen Spurkammersystem.
Die einzelnen, oben beschriebenen Detektorkomponenten besitzen getrennte Gasfüllungen (für die Driftkammern $Ar-CO_2-CH_4$ im Verhältnis $89.5 : 9.5 : 1.0$) und getrennte elektrostatische Schirmung. Diese Schirmung wird durch eine auf die aus Kohlenstoffasern bestehenden Trennwände aufgedampfte Aluminiumbeschichtung erreicht. Diese

Beschichtung ist zur Vermeidung von Konversionsphotonen aus zu detektierenden Elektronen, die das Aluminium durchdringen, möglichst dünn zu halten (hier 50 μm) [2]. Die vorverstärkten, analogen Drahtsignale werden aus den Spurkammern mit FADC-Karten (Flash Analog Digital Converter) digitalisiert, und der zentralen Triggerlogik (CTL) zur Weiterverarbeitung und Fällung der Triggerentscheidung zur Verfügung gestellt.

Driftkammer	*FTD*		*CTD*		
	Radial	Planar	CJC1 und CJC2	COZ	CIZ
Anzahl der Kanäle	1728	1152	2640	96	60
Datenmenge [Kbyte]	20.5	28.5	33.2	3.4	5.1
Datenmenge nach (Q,t)-Analyse [Kbyte]	2.4	4.1	5.4	0.5	0.9

Tabelle 2.1: *Typischerweise ausgelesene Datenmengen aus den Driftkammern [2]. Die (Q,t)-Analyse bedeutet eine Analyse der Signalzeitdauer t und der detektierten zeitabhängigen Ladung Q.*

Da in Abschnitt 4.2.4 eine Abschätzung der Datenmenge für Ereignisse anhand der Anzahl der bei einem Ereignis angesprochenen Driftkammerdrähte stattfindet, ist in Tabelle 2.1 eine Aufstellung der typischen Datenmengen in der primären Datenauslese für verschiedene Driftkammern zusammengestellt. Man sieht in dieser Tabelle, daß die größten Datenmengen der Driftkammern für ein Ereignis aus den radialen und planaren Vorwärtsdriftkammern und aus CJC1 und CJC2 stammen.

2.2.3 Das Flüssig-Argon-Kalorimeter

Das Flüssig-Argon-Kalorimeter ist insbesondere unter Aspekten guter Teilchenidentifikation und Energiemessung geladener Teilchen und γ-Quanten in die meßtechnische Umgebung des H1-Detektors eingepaßt worden. Es weist eine gute Betriebsstabilität auf und besitzt eine feine Granularität, die der Unterscheidung von Elektronen und Pionen angepaßt wurde. Eine Konstruktionsanforderung war weiterhin die Vermessung von *Jets*, das sind kollimierte Teilchenbündel mit hohen Teilchendichten [2].

Die gegenseitige Lage von Kalorimeter und Spule bedingt, daß die Teilchentrajektorien im Kalorimeter wie in der Spurkammer gekrümmt sind. Hier steht jedoch der Vorteil im Vordergrund, daß zu detektierende Teilchen nicht das Material der Spule durchdringen müssen, bevor sie zur Energiemessung in das Flüssig-Argon-Kalorimeter gelangen.

In Abbildung 2.5 ist ein Längsschnitt des Flüssig-Argon-Kalorimeters zu sehen, der die Aufteilung in Module zeigt. Das Flüssig-Argon-Kalorimeter überdeckt bezogen auf den nominellen Wechselwirkungspunkt einen Polarwinkelbereich von 4° bis 150° [1]. Die Module des Kalorimeters (siehe Abbildung 2.5) wiegen in der Summe 450 t und

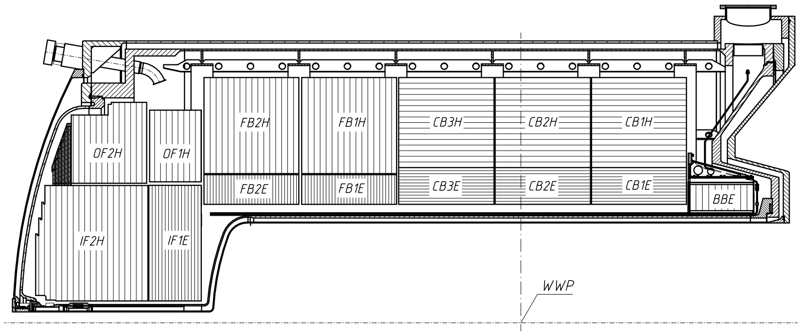


Abbildung 2.5: *Das Flüssig-Argon-Kalorimeter im Längsschnitt.*

Abbildung 2.6: *Darstellung der richtungsabhängigen hadronischen Wechselwirkungslänge λ und der zwanzigfachen Strahlungslänge X_0 für Elektronen im Kalorimeter durch Isolinien.*

sind mit einer Menge von 53 m^3 flüssigem Argon gefüllt. Das Argon wird durch ein Heliumgas-Kühlsystem flüssig gehalten.

Wie in Abbildung 2.5 zu erkennen ist, besitzt das Kalorimeter elektromagnetische und hadronische Komponenten. Man erkennt diese in der Abbildung an 'E' bzw. 'H' am Ende der vierstelligen Modulbezeichnung. Beide funktionalen Einheiten des Kalorimeters sind als nicht-kompensierende *Sampling*-Kalorimeter ausgelegt.

Im elektromagnetischen Teil wurde Blei als Schauermaterial, in dem die Elektronen einen elektromagnetischen Schauer ausbilden, verwendet. Da den Elektronen hier etwa 20 Strahlungslängen X_0 (siehe Abbildung 2.6) entgegenstehen, wird ein Großteil der Energie der Elektronen hier absorbiert und durch das sich als Ionisationsmaterial zwischen den Bleiplatten befindliche flüssige Argon detektiert. Die Bleischichten sind 2.4 mm dick und haben einen Abstand von 2.35 mm. Der relative Fehler der Energiemessung nimmt mit zunehmender Teilchenenergie nach folgender Gleichung ab:

$$\frac{\sigma_E}{E} = \frac{0.11}{\sqrt{E}} \oplus 0.01. \qquad (2.2)$$

$\oplus$ bedeutet quadratische Addition.
Die hadronische Sektion ist unter Verwendung etwas dünnerer Schichtung ähnlich aufgebaut wie die elektronische. Als Schauermaterial wurde Stahl anstatt Blei benutzt. Je nach Polarwinkel der Flugrichtung der Hadronen sind diese zwischen 4 und 8 Wechselwirkungslängen λ ausgesetzt. Bei einem nichtkompensierenden Kalorimeter gehen bis zu 30% der Energie ungemessen verloren. Hierfür sind Spaltung und Spallation von Kernen, sowie Anregung von Kernen und Verdampfen von Neutronen verantwortlich. Es sind daher rechnerische Korrekturen notwendig. Die relative Energieauflösung beträgt im hadronischen Kalorimeterteil

$$\frac{\sigma_E}{E} = \frac{0.5}{\sqrt{E}} \oplus 0.02. \qquad (2.3)$$

Zur elektronischen Auslese des Detektorzustandes während eines Ereignisses stehen im elektromagnetischen Teil 30784 Kanäle und im hadronischen Teil 13568 Kanäle zur Verfügung. Zu Triggerzwecken werden diese Signale in Gruppen zu Triggertürmen addiert. Man erhält durch weitere Signaladdition von bis zu 4 solcher Triggertürme je 240 elektromagnetische und hadronische *Bigtower* [19].

2.2.4 Das Instrumentierte Eisen

Das instrumentierte Eisen schließt sich vom Wechselwirkungspunkt aus gesehen an die Spule an und erfüllt eine Doppelfunktion. Es ist in erster Linie dafür da, das von der supraleitenden Spule erzeugte Magnetfeld induktiv abzuschirmen, damit dieses mit ungefähr 1.2 T sehr starke Feld nur den Raum der Spurkammern, die man zur Impulsmessung benutzt, und den Raum des Flüssig-Argon-Kalorimeters erfüllt. Die sich innerhalb des Eisens befindliche Magnetfeldspule hat einen Durchmesser von 6 m und eine Länge von 5.75 m.
In das Eisenjoch sind zur Vermessung von Myonen 10 Streamerkammerlagen eingebaut worden (Abbildung 2.8). Es befinden sich weiterhin je 3 Streamerkammerlagen innerhalb und außerhalb des Eisens. Eine weitere Funktion der Streamerkammern ist die Vermessung von Schauerresten von Schauern, die sich über den Bereich des Kalorimeters hinaus erstrecken. Die Streamerkammern werden mit einer gesamten Gasmenge von 36 m^3 in der Mischung von $CO_2 : Ar : Isobutan = 0.88 : 0.025 : 0.095$ betrieben [2].

Wie in Abbildung 2.7 ersichtlich ist, unterteilt man den Eisendetektor in 64 Module. Es wird der Bereich der Module 54 bis 59 vordere, innere Endkappe genannt. Die Module 48 bis 53 und 60 bis 63 gehören zur vorderen, äußeren Endkappe. Die hintere, innere Endkappe ist im Bereich der Module 6 bis 11. Zur hinteren, äußeren Endkappe

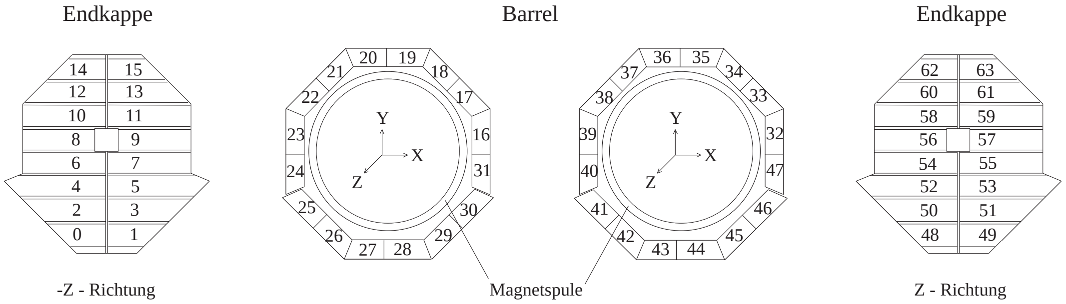


Abbildung 2.7: *Die Module der vier Bereiche des Eisenmyondetektors.*

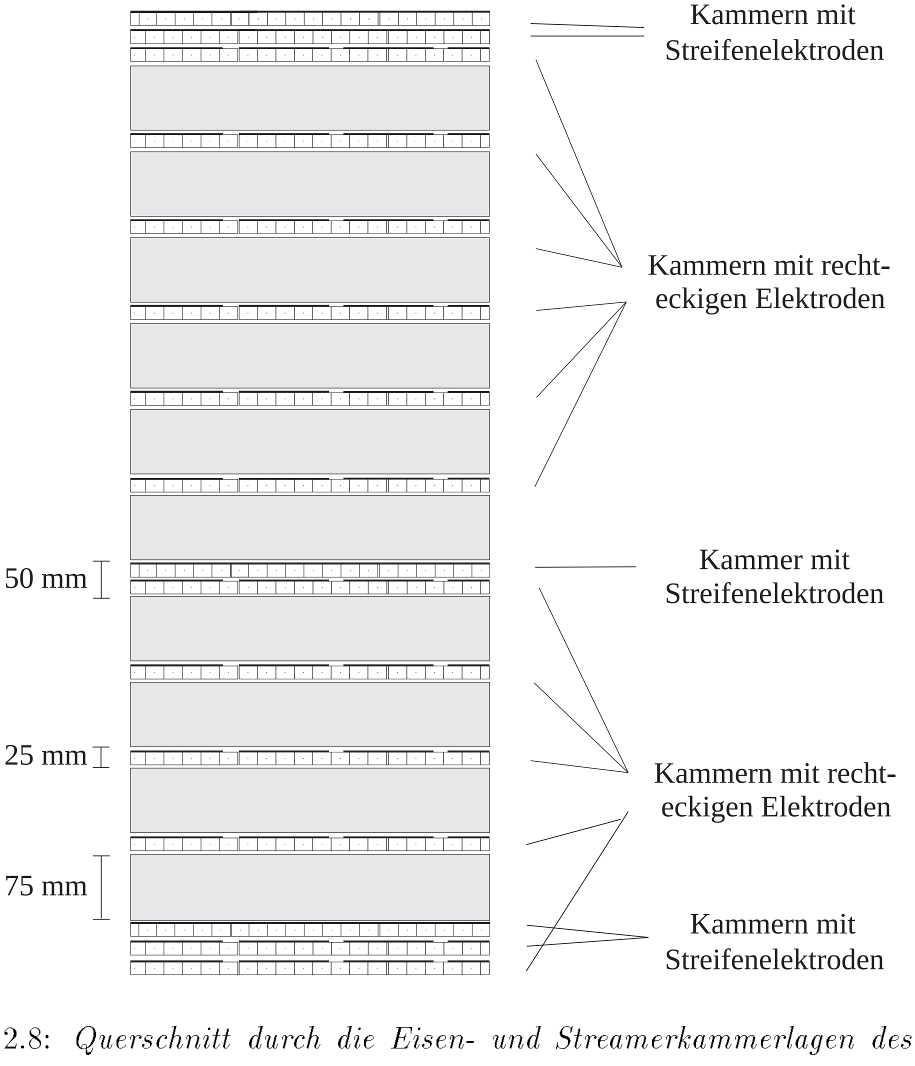


Abbildung 2.8: *Querschnitt durch die Eisen- und Streamerkammerlagen des instrumentierten Eisens.*

gehören die Module 0 bis 5 und 12 bis 15.
Von den 16 Streamerkammerlagen im Eisendetektor sind 5 mit Streifen- und 11 mit rechteckigen Elektroden versehen (siehe Abbildung 2.8). Zu Triggerzwecken werden jedoch nur die Drähte von 5 der 16 Streamerkammerlagen benutzt. Es sind vom nominellen Wechselwirkungspunkt aus betrachtet die Lagen 3, 4, 5, 8 und 12. Für die Triggerung von Myonen im Eisen werden für die Bereiche vordere, innere und äußere Endkappe, Zentralbereich (Barrel), und hintere Endkappe (Abbildung 2.7) verschiedene Bedingungen an die Anzahl der Lagen mit Signal gestellt. Hierauf wird in Abschnitt 2.4 noch näher eingegangen werden.

2.2.5 Sonstige Detektorkomponenten

Im folgenden werden zusätzlich zu den oben genannten, für diese Arbeit nicht im Vordergrund stehende Detektoren genannt und deren Funktion erklärt.

RÜCKWÄRTIGE DRIFTKAMMER
Dies ist eine Driftkammer vor dem Spacal (siehe unten) zur Vermessung von Teilchenspuren im rückwärtigen Detektorbereich.

SILIZIUMDETEKTOR
Noch innerhalb der Spurkammern wurde ein Halbleiterdetektor mit einem zentralen und einem rückwärtigen Teil realisiert, der sich noch in der Erprobungsphase befindet. Durch diesen soll die Vermessung der Reaktionsvertizes verbessert werden.

SPACAL
Das Spacal ist ein rückwärtiges Kalorimeter, durch welches das noch in Abbildung 2.1 angedeutete, rückwärtige, elektromagnetische Kalorimeter ersetzt wurde. Es besteht wie das Flüssig-Argon-Kalorimeter aus einer elektromagnetischen und einer hadronischen Sektion mit etwa 27 bzw. 29 Strahlungslängen und ungefähr jeweils einer Wechselwirkungslänge. Mit dem Spacal können Teilchen im in bezug auf die Protoneinschußrichtung rückwärtigen Raum, wie zum Beispiel gestreute Positronen aus Ereignissen der tiefinelastischen Streuung, vermessen werden.

VORWÄRTSKALORIMETER
Das Vorwärtskalorimeter befindet sich zwischen vorderem Flüssig-Argon-Kalorimeter und der Strahlröhre und instrumentiert diesen Raum für Teilchenenergiemessungen. Es ist im Gegensatz zum kalten Flüssig-Argon-Kalorimeter ein warmes Kalorimeter mit den aktiven Elementen Kupfer und Silizium.

VORWÄRTSMYONSYSTEM
Diese Detektorkomponente befindet sich außerhalb vom vorderen Eisenjoch und ist daher insbesondere für die Vermessung von Spuren von Myonen geeignet, welche das Eisen

durchdringen. Hierfür benötigen diese Myonen eine Mindestenergie in Vertexnähe von 5 GeV.

PROTON-REMNANT-TAGGER
Mit Hilfe von 7 Szintillationszählern, die über Photomultiplier ausgelesen werden, kann man hier die hadronischen Schauer von dissoziativen Ereignissen messen. Es erfolgt hier die Instrumentierung des Winkelbereiches zwischen Vorwärtskalorimeter und Vorwärtsprotonspektrometer.

VORWÄRTSPROTONSPEKTROMETER
Dieser Detektor besteht aus vier Stationen entlang der Strahlröhre in Protonflugrichtung (63, 80, 81 und 90 Meter vom Wechselwirkungspunkt entfernt) und dient zur Messung vorwärtsgestreuter Protonen aus diffraktiven Proton-Positron-Reaktionen im Hauptdetektor. Dabei wirken Faserhodoskope, die innerhalb der Strahlröhre bis auf wenige Millimeter an den Protonstrahl herangeführt werden, zusammen mit dem Magnetfeld strahlführender Dipolmagnete als magnetische Spektrometer.

ELEKTRON-KLEINWINKELDETEKTOREN
Diese sind bei $z = -33\,m$ und $z = -44\,m$ zu finden und unterstützen durch Detektion von gestreuten Positronen die Vermessung von Photoproduktionsereignissen. Zusammen mit dem Photondetektor bei $z = -103\,m$ wird über den Bremsstrahlungsprozess

$$ep \longrightarrow ep\gamma \tag{2.4}$$

bei H1 die Luminosität gemessen.

2.3 Physik bei H1

Bei H1 werden verschiedene Lepton-Nukleon-Streuereignisse zur Aufklärung insbesondere der Eigenschaften von Protonen und Photonen vermessen. Die Untersuchungen finden größtenteils im Rahmen des heutigen Standardmodells der Elementarteilchenphysik statt. Neue Messungen lassen Spekulationen über darüber hinaus gehende Modelle zu [3].
Die ep-Kollisionen bei HERA können in folgende Ereignisklassen eingeteilt werden. Mit Ereignisraten im Bereich von 100 Hz werden Photoproduktion und diffraktive Ereignisse detektiert. Mit 3 Hz schließen sich tiefinelastische Streuereignisse an. Mit Raten im mHz-Bereich bleiben dann noch die elektroschwachen, tiefinelastischen Streuungen [16].
Das Interesse, sich bei Hera mit der Untersuchung von J/ψ-Mesonen zumeißt in Photoproduktion zu befassen, liegt darin, über den Prozeß der Photon-Gluon-Fusion die Gluonendichte des Protons zu messen. Zum zweiten nähert man sich der Beschreibung diffraktiver Prozesse mit von *weichen* und *harten* Gluonen ausgehenden Modellen an, so daß man etwas über den Übergang dieser Modelle lernen kann [5].

Nach allgemeinen Betrachtungen zur kinematischen Beschreibung von Positron-Proton-Reaktionen im H1-Detektor bei HERA in Abschnitt 2.3.1 wird im darauffolgenden Abschnitt 2.3.2 auf verschiedene Möglichkeiten der Erzeugung von J/ψ-Mesonen bei H1 eingegangen.

2.3.1 H1-Ereigniskinematik

Die möglichen ep-Reaktionen lassen sich allgemein durch die Diagramme in Abbildung 2.9 veranschaulichen. Dabei kann die Wechselwirkung durch Neutrale Ströme (γ, Z^0) oder Geladene Ströme (W^+, W^-) vermittelt werden.

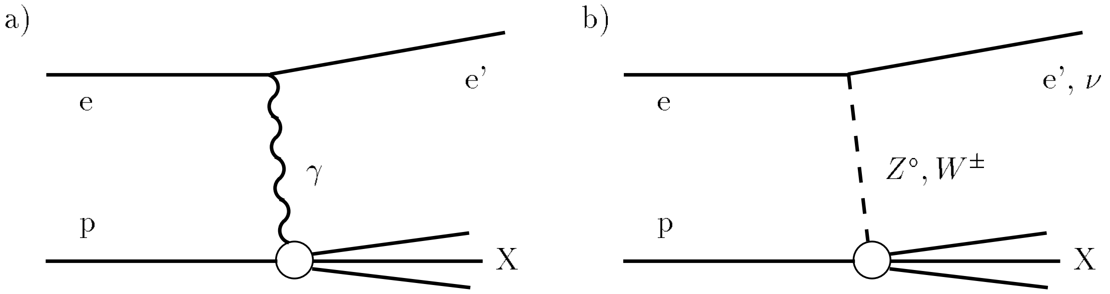


Abbildung 2.9: *Dargestellt sind Graphen für ep-Wechselwirkung der a) elektromagnetischen und b) schwachen Wechselwirkung.*

Der dargestellte hadronische Endzustand X steht für ein gestreutes oder dissoziiertes Proton bei diffraktiven Streuereignissen oder für ein fragmentiertes Proton bei inelastischer Streuung. Die wichtigsten Größen zur kinematischen Beschreibung solcher Reaktionen werden im folgenden eingeführt.
Es gibt allgemeine Eigenschaften von Stoßreaktionen in der Hochenergiephysik, die sich gegenüber Stoßreaktionen, wie sie von der klassischen Mechanik von Massenpunkten her bekannt sind, trotz der hochrelativistischen Kinematik nicht ändern. Dies ist insbesondere die Eigenschaft elastischer Ereignisse bei Stoßpartnern kollinearer Bewegungsrichtung, daß der Stoßprozeß kinematisch bereits durch Angabe einer skalaren Größe, wie zum Beispiel den Streuwinkel oder den Energieübertrag charakterisiert ist. Für inelastische Ereignisse wird eine Größe mehr benötigt, um den Energieverlust der Stoßpartner darzustellen. Benutzt werden dazu bei H1 z.B. das Viererimpulsübertragsquadrat Q^2 und die Bjørkenvariable x.
Bei der elektromagnetischen Wechselwirkung berechnet man die Photon-Proton-Schwerpunktenergie $W_{\gamma p}$. Mit der Protonmasse M, dem Protonimpuls P, dem Impuls des Photons q und der Konvention c=1 ergibt sich für $W_{\gamma p}$ im Protonruhesystem :

$$W_{\gamma p}^2 = (P+q)^2 = M^2 + 2Pq + q^2 = M^2 + 2M\nu - Q^2 \tag{2.5}$$

Dabei ist Q^2 durch

$$Q^2 = -q^2 \tag{2.6}$$

und ν durch den lorentzinvarianten Ausdruck

$$\nu = \frac{P \cdot q}{M} \tag{2.7}$$

definiert, wobei Q^2 im Ruhesystem des Protons das Quadrat des Energieübertrags vom Positron auf das Proton angibt. Bei elastischen Ereignissen ist $W_{\gamma p} = M$, woraus folgt $Q^2 = 2\mathrm{M}\nu$. Bei inelastischen Ereignissen gilt $W_{\gamma p} > M$ und man erhält $Q^2 < 2M\nu$. Diese Unterscheidung kann auch anhand der Bjørkenvariablen x vorgenommen werden.

$$x = \frac{-q^2}{2M\nu} \tag{2.8}$$

Diese Größe ist für das Partonmodell der Nukleonen von Bedeutung. Hierin gibt x den Anteil des Protonimpulses wieder, den das Parton trägt. Dies gilt jedoch nur in einem Bezugsystem, in dem die Partonen schnell sind, so daß der Transversalimpuls und die Ruhemasse des Partons vernachlässigt werden können. Bei Hera ist die Protonstruktur bis in den Bereich $x \approx 10^{-6}$ vermessen worden [4].
Man unterscheidet bei HERA den Bereich der Photoproduktion $Q^2 \simeq 0$ und tiefinelastische Streuung $Q^2 > 1\,GeV^2$. Das maximale Viererimpulsübertragsquadrat bei HERA beträgt $Q^2_{max} = 10^5\,GeV^2$.

2.3.2 Die J/ψ-Produktion

Das J/ψ-Teilchen wurde 1974 bei einer Schwerpunktenergie von $M_{J/\psi} = 3097 MeV$ als extrem schmale Resonanz mit einer Breite von $(87.7 \pm 5)\,keV$ [11] annähernd gleichzeitig in zwei verschiedenen Experimenten entdeckt. Aus diesem Grunde erhielt es seinen Doppelnamen. Das J/ψ-Teilchen ist ein gebundener $c\bar{c}$-Zustand, und zwar der energetisch niedrigste mit Eigendrehimpuls 1. Es existieren auch angeregte Zustände ψ', ψ", usw., die eine höhere Masse als das J/ψ besitzen. Das ψ' entsteht bei einer Masse von $M_{\psi'} = 3685 MeV$ mit einer relativen Rate in ep-Wechselwirkung von etwa 0.2 zur J/ψ-Produktionsrate. Das J/ψ besitzt die gleichen Quantenzahlen wie das Photon $J^{PC} = 1^{--}$ und kann direkt an ein solches koppeln.
Inelastische und diffraktive Produktionsmechanismen können anhand der Elastizität z unterschieden werden. Die Elastizität z ist folgendermaßen definiert:

$$z = \frac{E_\psi}{E_\gamma} \tag{2.9}$$

E_ψ und E_γ sind die Energien des J/ψ-Mesons bzw. des virtuellen Photons betrachtet im Ruhesystem des Protons. Es gilt folgende Unterteilung:

- $z = 1$: Diffraktive, elastische Prozesse.
  Das Proton wird elastisch gestreut. Dieser Prozeß findet ohne vom Proton ausgehenden Farbfluß statt, d.h. die Wechselwirkung kann i.a. nicht durch einzelne Gluonen vermittelt werden.

- $0.95 \lesssim z \lesssim 1$: Diffraktive J/ψ-Produktion mit Proton-Dissoziation.
  Das Proton dissoziiert, wobei die Reaktionprodukte bei H1 im Proton-*Remnant*-Tagger oder Vorwärtskalorimeter nachgewiesen werden können. Auch bei diesen Reaktionen existiert kein Farbfluß vom Proton aus.

- $0 \lesssim z \lesssim 0.9$: Inelastische J/ψ-Produktion.
  Bei diesen Reaktionen fragmentiert das Proton und die austretenden Gluonen und Quarks hadronisieren. Es findet vom Proton ausgehend ein Farbfluß statt.

Diffraktive J/ψ-Produktion

Es gibt zwei Ansätze der theoretischen Beschreibung der diffraktiven J/ψ-Produktion, das Donnachie-Landshoff-Modell und das Ryskin-Modell [8].
Das Donnachie-Landshoff-Modell beschreibt J/ψ-Produktion als einen *weichen* Prozess, der in Regge-Theorie berechnet wird. Der Wirkungsquerschnitt der Emission eines γ-Quants durch das Positron ist durch die Weizsäcker-Williams-Approximation gegeben [6]. Die Fluktuation des vom Positron abgestrahlten virtuellen Photons in das J/ψ-Meson wird durch das Vektor-Meson-Dominanz-Modell beschrieben. Man erhält so drei unabhängige Prozesse: die Abstrahlung des Photons, dessen Fluktuation in ein J/ψ-Meson und die diffraktive Streuung am Proton.

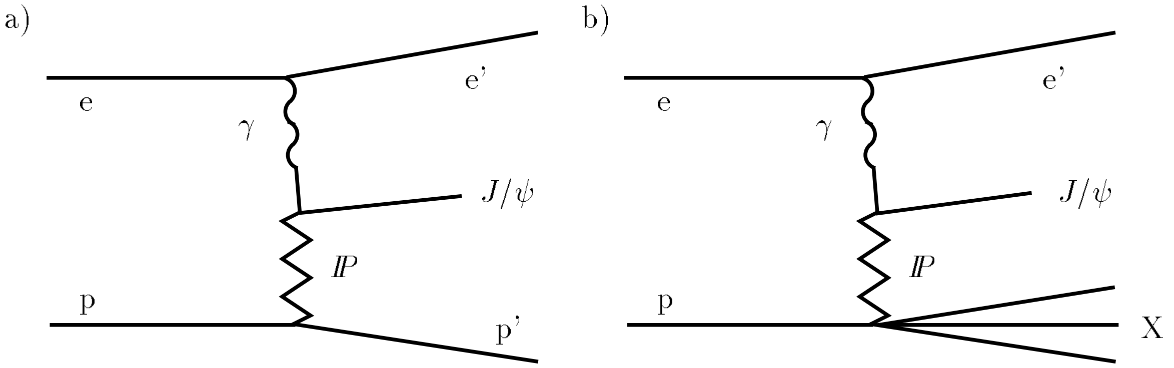


Abbildung 2.10: *Zwei Diagramme zur diffraktiven J/ψ-Produktion, a) elastisch und b) mit Dissoziation des Protons.*

Diese Prozesse sind in den Diagrammen in Abbildung 2.10 schematisch dargestellt. Man postuliert das Pomeron $I\!\!P$ mit den Quantenzahlen des Vakuums, so daß es keine Farbladung trägt.
In Modellen, in denen versucht wird, die diffraktive Streuung im Rahmen der QCD zu beschreiben, muß das Pomeron wegen der Farbladung der Gluonen durch mindestens zwei Gluonen beschrieben werden. Ein solches Modell ist zum Beispiel das Donnachie-Nachtmann-Modell, in welchem nicht störungstheoretisch berechenbare Gluonen wie Photonen an das Proton koppeln [9, 10].

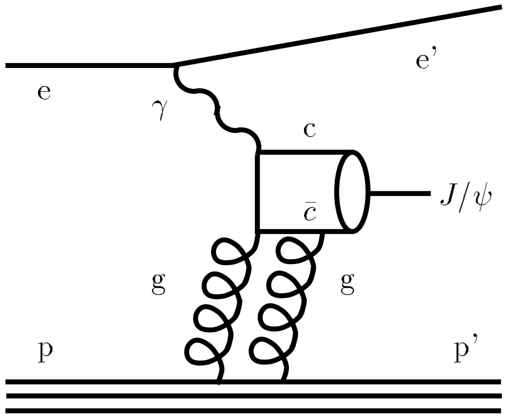


Abbildung 2.11: *Ein Diagramm zum Ryskin-Modell.*

Das zweite Modell für die diffraktive J/ψ-Erzeugung ist das Ryskin-Modell in welchem die elastische J/ψ-Produktion in einer störungstheoretischen Beschreibung im Rahmen der QCD durch den Austausch von Gluonen beschrieben wird. Das Diagramm zu einem der in diesem Modell benutzten Terme der Entwicklung ist in Abbildung 2.11 dargestellt. In diesem Modell wird versucht, den nicht existenten Farbfluß durch eine Gluonenleiter zu beschreiben.

Inelastische J/ψ-Produktion

Die inelastische J/ψ-Produktion wird bei Hera durch die Photon-Gluon-Fusion dominiert [7].

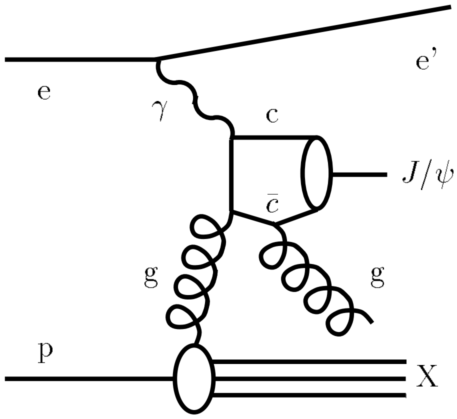


Abbildung 2.12: *Diagramm zum Colour-Singlett-Modell in führender Ordnung.*

Das Feynmandiagramm zum bevorzugten *Colour-Singlett-Modell* in führender Ordnung (α, α_s^2) ist in Abbildung 2.12 gezeigt. Diesem Modell liegt die Abstrahlung eines *harten* Gluons zugrunde, wodurch der Übergang eines $c\bar{c}$-Paares in einen farbneutralen Zustand erreicht wird. Dieser Prozeß ist sensitiv auf die im Proton befindlichen Gluonen und kann in perturbativer QCD berechnet werden.

Neben der hier gezeigten direkten Photonwechselwirkung existieren *Resolved-Photon*-Prozesse, bei denen die hadronische Komponente des Photons in Wechselwirkung mit

einem Gluon oder einem Quark aus dem Proton tritt.
Neben Graphen höherer Ordnung existiert weiterhin das *Colour-Oktett-Modell*, das von einer Abstrahlung von *weichen* Gluonen zur Erreichung eines farblosen $c\bar{c}$-Paares ausgeht. Dabei kann der Übergang des $c\bar{c}$-Paares von einem farbgeladenen in einen farblosen Zustand nicht störungstheoretisch berechnet werden.

Der Zerfall des J/ψ-Mesons

Das J/ψ-Meson kann in Hadronen oder Leptonen zerfallen. In der nachfolgenden Tabelle 2.2 sind die partiellen Zerfallsbreiten angegeben.

Wechselwirkung	Endzustand	Anteil [%]	partielle Breite [keV]
stark	Hadronen	70.7 ± 2	62.0 ± 3.9
elektromagnetisch	Hadronen	17.0 ± 2	14.9 ± 1.9
	e^+e^-	6.02 ± 0.19	5.3 ± 0.4
	$\mu^+\mu^-$	6.01 ± 0.19	5.3 ± 0.4

Tabelle 2.2: *Übersicht über die partiellen Zerfallsbreiten des J/ψ-Mesons [11].*

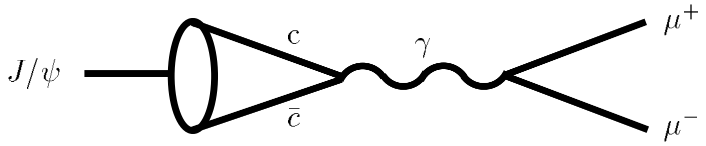


Abbildung 2.13: *Zerfall eines J/ψ-Mesons in ein $\mu^+\mu^-$-Paar.*

Da für diese Arbeit Ereignisse mit dem Zerfallskanal $J/\psi \longrightarrow \mu^+\mu^-$ betrachtet werden, bei denen die beiden Myonen im H1-Detektor gemessen wurden, ist in Abbildung 2.13 der Graph dieses Zerfalls dargestellt. Das in diesem Zerfall intermediäre γ-Quant ist virtuell und konvertiert in ein $\mu^+\mu^-$-Paar.

2.4 Der H1-Trigger

In diesem Abschnitt wird nach einer allgemeinen Beschreibung des Konzeptes, der Funktionsweise und der Eigenschaften des H1-Triggers speziell auf die für das Verständnis der Ergebnisse dieser Arbeit relevanten Triggerstufen eingegangen.

2.4.1 Allgemeine Beschreibung

Der Trigger des H1-Experiments hat die Aufgabe, in möglichst kurzer Zeit eine elektronische Entscheidung darüber zu fällen, ob ein Ereignis von physikalischem Interesse ist oder ob es sich um ein Untergrundereignis handelt, also nicht von einer ep-Reaktion stammt.
Der H1-Trigger ist vierstufig ausgelegt (siehe Abbildung 2.14). Die Stufen heißen L1, L2, L3 und L4, wobei nur die Stufen L1, L2 und L4 genutzt werden. L1 und L2 sind *Hardware*-Trigger, L4 ist *software*basiert und L3 wurde auch als *Software*stufe geplant. Jede Triggerstufe ist durch die Zeitdauer bis zur Triggerentscheidung charakterisiert. Je kürzer die Zeitdauer bis zur Triggerentscheidung, desto gröber ist die Information über das Ereignis bzw. desto einfacher ist deren Verarbeitung auf der jeweiligen Triggerstufe gehalten.

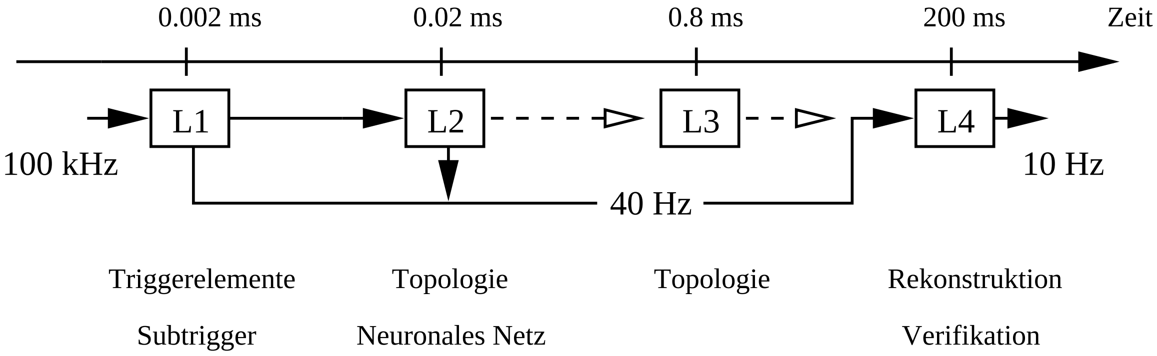


Abbildung 2.14: *Schematische Darstellung der Triggerstufen bei H1.*

L1 trifft zuerst seine Entscheidung, dann kommt L2, hierauf folgt L4. Wenn alle drei Triggerstufen ein Ereignis verifiziert haben, werden die zum betrachteten Ereignis gehörigen Daten auf Magnetband aufgezeichnet. Es ist eine Aufzeichnungsrate an Ereignissen von 10 Hz - 15 Hz vorgesehen. Die Positronen- und Protonenbündel kreuzen sich mit einer Rate von 10.4 MHz. Bei einer primären, am Trigger anliegenden Ereignisrate von ungefähr 100 kHz [16], kommt es nur bei jeder hundertsten Kreuzung der Teilchenbündel zu einer Reaktion. Von diesen Ereignissen kann nur jedes zehntausendste aufgezeichnet werden.
Während der Datenauslese der Subdetektoren bei positiver Triggerentscheidung kann kein neues Ereignis in die Elektronik eingelesen werden. Um diese Totzeit möglichst kurz zu halten, muß die Triggerentscheidung schnell erfolgen. Die gesamte Triggerelektronik ist insbesondere unter Gesichtspunkten der Triggersignalstabilität und guter Unterscheidung zwischen Signalen aus verschiedenen HERA-Uhr-Zeitintervallen (siehe Abschnitt 2.1) konzipiert worden.
Die erste Stufe L1 fällt eine Entscheidung darüber, ob das anliegende Ereignis verworfen werden soll oder ob das Ereignis als *interessant* klassifiziert wird und zur weiteren

Entscheidung zur Stufe L2 oder L4 gelangt. Für diese Entscheidung steht auf L1 ein Zeitraum von 2.5 μs zur Verfügung. Diese Zeitdauer orientiert sich an der Zeitspanne, die man benötigt, um erste aus den integrierten Spannungsverläufen abgeleitete Signale aus dem Flüssig-Argon-Kalorimeter auszulesen und an die zentrale Triggerlogik (CTL) weiterzuleiten [17]. Jedes der Detektorsubsysteme liefert über die jeweilige Ausleseelektronik durch Analog-Digital-Wandler digitalisierte Signale an die zentrale Triggerlogik. Hier werden diese Signale zu 192 Triggerelementen verarbeitet, welche wiederum zu bis zu 128 Subtriggern kombiniert werden. Jeder dieser Subtrigger drückt also eine Entscheidung über das Vorhandensein mehrerer, sich in den Triggerelementen grob äußernder, physikalischer Subdetektorzustände aus. Es gibt Subtrigger, die auf physikalisch interessante Ereignisse ansprechen, und solche, die zur Berechnung von Luminosität und Effizienzen physikalischer Subtrigger dienen. Letztere werden Monitortrigger genannt. Um L1 totzeitfrei betreiben zu können, werden die Daten in einem Schieberegister zwischengespeichert bis die Triggerinformationen aus allen Subdetektoren zur Verfügung stehen und die L1-Entscheidung gefällt ist.

Spricht ein Subtrigger an, wird das Ereignis von L1 akzeptiert. Falls der Subtrigger eine zu hohe Rate aufweist, wird er mit einem Faktor q abskaliert. Dies bedeutet, daß nur jedes q-te Ereignis gespeichert wird. Abskalierung der Rate bedeutet Verlust an Luminosität für die physikalischen Ereignisse.

Die Totzeit des Detektors beginnt, nachdem ein Ereignis von L1 akzeptiert worden ist. Nach positiver L1-Entscheidung wird, sofern kein L2-Triggerelement zur Verfügung steht, die volle Ereignisdatenmenge von nahezu 3 $Mbyte$ aus den Detektorkomponenten ausgelesen und mit 3.3 $Gbyte/s$ an die L4-Triggerstufe übertragen [2]. Ist ein L2-Triggerelement für diesen Subtrigger vorhanden, geschieht dies erst nach positiver Entscheidung durch dieses L2-Triggerelement. Zur Vermeidung von L1-Abskalierungsfaktoren bei zu hohen Triggeraten hat man sich mit der zweiten Triggerstufe die Möglichkeit geschaffen, abhängig vom angesprochenen L1-Subtrigger als Bedingung zur Weiterverarbeitung der Daten, L2-Triggerlemente zu fordern. Der Zeitraum dieser Entscheidungsfindung ist auf 20 μs beschränkt. Dabei steht dem L2-Triggersystem die gleiche Information zur Verfügung, wie dem L1-Triggersystem. Die Informationen können auf L2-Ebene auf Grund der zehnfachen zur Verfügung stehenden Zeit lediglich *intelligenter* genutzt werden.

Auf der zweiten Triggerebene L2 existieren zwei unabhängige, konzeptionell völlig unterschiedliche Systeme. Eines dieser Systeme ist der Topologische Trigger (L2-TT) [21, 22], durch welchen topologische Unterschiede in den Zusammenhängen zwischen den Größen zu triggernder, physikalischer Ereignisse von denen der Untergrundereignisse ausgenutzt werden. Die Komplexität der Berechnung orientiert sich an der Entscheidungszeitspanne von 20 μs.

Das zweite, in dieser Arbeit im Vordergrund stehende, L2-Triggersystem ist das der Neuronalen Netze. Dieses wird auch L2-NN-Trigger gennant. Auch bei diesem Konzept werden Topologien ausgenutzt, allerdings mit Hilfe des problemspezifischen Trainings Neuronaler Netze. Ein solches Netz *lernt* hierbei die Physik und den Untergrund der von einem bestimmten Subtrigger akzeptierten Ereignisse zu unterscheiden, so daß unter Einsatz dieses Netzes und unter Beibehaltung der Subtriggerrate der ereignisunspezifische Abskalierungsfaktor q vermindert werden kann. Das Konzept der Neuronalen Netze in der L2-Triggerstufe wird in Abschnitt 2.4.2 vorgestellt und die Zusammenhänge des Netztrainings in Kapitel 3 erklärt.

Auf Stufe L4 kann mit einer Entscheidungsdauer von 200 ms die gesamte Detektorinformation zur Entscheidung genutzt werden. Ereignisse, welche auch diese Triggerstufe passieren, werden abgespeichert. Nach ungefähr 1-2 ms ist das Auslesen der gesamten Detektordaten abgeschlossen und die beschriebene Triggerprozedur beginnt von vorn.

Mit den auf diese Weise gespeicherten Daten werden unabhängig vom Detektorbetrieb Rekonstruktions- und Klassifikationsprogramme zur weiteren Verarbeitung und Analyse benutzt. Diese Stufe wird L5 genannt.

Eine Datennahmeeinheit (*Run*) ist ein zeitlich zusammenhängender, abgeschlossener Datennahmeabschnitt. Eine Datennahmeeinheit erstreckt sich niemals über mehr als eine Positronen- bzw. Protonen-Füllung von HERA. Zur Berücksichtigung der sich während einer Datennahmeeinheit ständig ändernden Strahl- und damit Reaktions- und Untergrundbedingungen definiert man die Aufzeichnungsphasen 1 bis 4 in Abhängigkeit von den Strahlbedingungen. Für jede dieser Phasen existiert eine definierte, separate Triggerkonfiguration, die den Gegebenheiten dieser Phase angepaßt wurde.

Mit über die Jahre zunehmender Luminosität steigt auch die Rate der Untergrundereignisse. Die Triggerkonfiguration muß daher ständig verbessert werden, um die physikalisch interessanten Ereignisse effizienter filtern zu können.

2.4.2 Der L2-NN-Trigger

Der L2-Neuronale-Netzwerk-Trigger (L2-NN-Trigger) ist ein programmierbarer *Hardware*trigger, der auf einer Neuronalen Netzwerkarchitektur basiert. Er ist für maximal 16 parallel laufende Neuronale Netze ausgelegt.

Man verwendet Neuronale Netze als Trigger zur Ausnutzung hochdimensionaler Korrelationen zwischen den Eingangsgrößen, wegen der hohen Flexibilität, z.B. Training und Einbau neuer Netze, und auf Grund der Schnelligkeit der parallelen Datenverarbeitung von elektronisch implementierten Netzen [18].

Die verwendeten Netze werden in der Elektronik parallel abgearbeitet. Dies spiegelt sich in der schnellen L2-Triggerentscheidung wider. Eine Berechnung der Netzentscheidung bei vorheriger Aufbereitung der Daten ist mit Hilfe eines seriellen Rechners in 20 μs zur Zeit nicht möglich. Die verwendete Architektur der Netze nennt man *Feedforward*-Neuronales-Netz (siehe Abschnitt 3.2).

Jedes dieser Netze läuft auf zumeist einem L1-Subtrigger, d.h. es urteilt über Ereignisse, die von diesem Subtrigger getriggert wurden. Genau für diesen Subtrigger wurde es auch konzipiert. Jedes Netz erhält eine Anzahl von Netzeingabegrößen, auf Grund derer es zu einer Ausgabegröße zwischen -128 und 127, also einer 8-*bit*-Zahl, gelangt. Je nach Netz kann nun ein Schnittwert bestimmt werden, so daß das beurteilte Ereignis bei einer Ausgabegröße unterhalb des Schnittwertes als Untergrund und

oberhalb des Schnittwertes als Physik klassifiziert wird. Die Entscheidung wird an die zentrale Triggerlogik, die für die Koordination der Datenverarbeitung programmiert ist, zurückgeleitet.
Die möglichen, von der zentralen Triggerlogik zur Verfügung gestellten Netzeingabegrößen werden im sogenannten Ereignisvektor zusammengefaßt. Technisch können bis zu 64 verschiedene Eingabegrößen pro Netz von den bei H1 im L2-NN-Trigger verwendeten Netzen benutzt werden. Im folgenden werden die Größen aus dem Ereignisvektor vorgestellt, die beim Training der Netze in dieser Arbeit benutzt werden.

- **z-Vertex-Größen**

  Das z-Vertex-Histogramm stellt die Gesamtheit aller Einträge dar, die dadurch zustande kommen, daß man die getroffenen Auslese-Elektroden (siehe Abschnitt 2.2.2) der vorderen und zentralen Proportionalkammern (CIP, COP und FPC) verbindet und diese Strecken bis zur nominellen Strahlachse verlängert. Liegt ein solcher Strahl in einem Bereich $\pm 40\,cm$ zu beiden Richtungen des nominellen Wechselwirkungspunktes längs der Strahlachse (siehe Abbildung 2.15), so findet ein Eintrag in das z-Vertex-Histogramm statt. Hierbei werden alle möglichen Kombinationen der Auslese-Elektroden gebildet, solange die Verlängerung der Verbindungslinie in den oben genannten Bereich nahe der nominellen Strahlachse fällt. Es finden also auch Einträge in das Histogramm statt, die von Verbindungen von Elektroden stammen, welche nicht auf ein und dasselbe Teilchen angesprochen haben. Dies ist in Abbildung 2.15 verdeutlicht. Das Histogramm ist in 16 Intervalle unterteilt.

  CPVSUM: Diese Größe enthält die Anzahl aller Einträge im z-Vertex-Histogramm.
  CPVMAX: *CPVMAX* ist die Anzahl der Einträge im z-Vertex-Histogramm in dem Intervall mit der maximalen Anzahl an Einträgen aus allen 16 Intervallen.
  CPVPOS: Hier wird die Position des Intervalles mit der maximalen Anzahl an Einträgen, bezeichnet mit einer Zahl zwischen 0 und 15, angegeben.

- **Driftkammergrößen**

  Die Spurkammerelektronik ist mit einer Anzahl von Masken versehen, durch welche sehr schnell Spuren klassifiziert werden können. Es gibt für jede der folgenden Größen jeweils 45 ϕ-Sektoren, welche dadurch aktiviert werden, daß mindestens eine Maske in diesem Sektor angesprochen hat.

  TRHINEG: Diese Größe gibt die Anzahl der ϕ-Sektoren an, in denen mindestens eine der Masken angesprochen hat, die sensitiv ist auf einen transversalen Impuls P_t mit $P_t > 800\,MeV$ und von einem negativ geladenen Teilchen herrührt.
  TRHIPOS: Dies ist das Analogon zur Größe *TRHINEG* für Masken, welche die Krümmungsrichtung positiver Teilchen beschreiben.

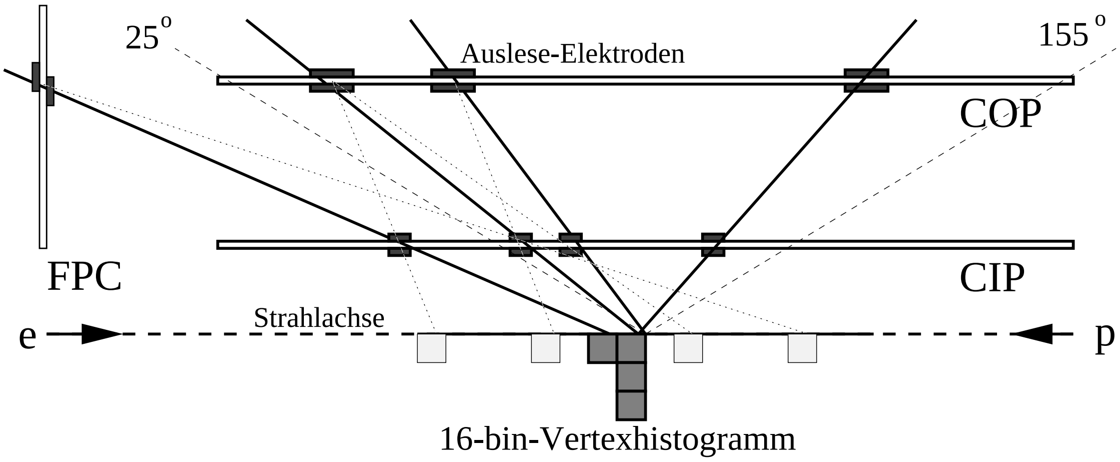


Abbildung 2.15: *Darstellung der Errechnung des z-Vertex-Histogramms. Die durchgezogenen Verbindungen sollen hier zu Elektroden gehören, welche jeweils auf ein und dasselbe Teilchen angesprochen haben. Die gepunkteten Verbindungen verursachen aber auch noch Einträge in das z-Vertex-Histogramm.*

TRLONEG: In Anlehnung an die Größe *TRHINEG* enthält *TRLONEG* die Anzahl der Sektoren, innerhalb derer mindestens eine Maske angesprochen hat, die zur Detektion negativer Teilchen mit $400\,MeV < P_t < 800\,MeV$ ausgelegt ist.
TRLOPOS: Hier stellt sich der Impulsbereich wie bei *TRLONEG* dar, die Masken sind hingegen für positiv geladene Teilchen ausgelegt.

- **Proportionalkammergrößen**

  Das den folgenden Größen zugrunde liegende Konzept der *Bigrays* fußt auf den Verbindungen der angesprochenenen Auslese-Elektroden der vorderen und zentralen Proportionalkammern (CIP, COP und FPC), wie sie auch schon zur Erstellung des z-Vertex-Histogramms herangezogen wurden. *Bigrays* werden die Proportionalkammerelektrodenverbindungen genannt, welche einen Eintrag in der Nähe des Intervalls mit der maximalen Anzahl an Einträgen im z-Vertex-Histogramm verursacht haben. *Bigrays* unterliegen jedoch einer gröberen Einteilung als die zuvor betrachteten Proportionalkammerelektrodenverbindungen und werden auf einer gedachten Karte mit 16 Φ-Intervallen und 14 Θ-Intervallen eingeordnet. Die auf dieser Karte definierten Richtungen stimmen mit den Richtungen der *Bigtower* (siehe Abschnitt 2.2.3) aus dem Flüssig-Argon-Kalorimeter überein, so daß die Basis dafür gelegt ist, die Informationen aus den Spurkammern mit den Informationen aus dem Kalorimeter grob korrelieren zu können. Für den L2-NN-Trigger ist dies jedoch technisch noch nicht realisiert.
  Die *Bigrays* werden auf L2-Trigger-Niveau getrennt in vier Polarwinkelbereichen

gezählt. Die folgenden vier Größen enthalten die Anzahlen an *Bigrays* in den einzelnen Intervallen [15]. Es können auf diese Weise grobe, richtungstopologische Korrelationen und der grobe Richtungsschwerpunkt der Spuren in die Triggerentscheidung einbezogen werden.

NBIGFWD: Anzahl der *Bigrays* in den Θ-Intervallen 0-3 bzw. $\Theta \lesssim 13.6^\circ$.
NBIGFCE: Anzahl der *Bigrays* in den Θ-Intervallen 4-7 bzw. $13.6^\circ \lesssim \Theta \lesssim 28^\circ$.
NBIGFCE: Anzahl der *Bigrays* in den Θ-Intervallen 8-10 bzw. $28^\circ \lesssim \Theta \lesssim 94.8^\circ$.
NBIGBWD: Anzahl der *Bigrays* in den Θ-Intervallen 11-13 bzw. $94.8^\circ \lesssim \Theta \lesssim 152^\circ$.

- **Eingabegrößen aus dem instrumentierten Eisen**

Mit den folgenden fünf Eingabegrößen erhält man jeweils die Anzahl der angesprochenen Module, getrennt nach den vier Bereichen des Eisendetektors (Abbildung 2.7) und aufsummiert in der Größe *IRONTOT*.

IRONBE: Anzahl der angesprochenen Module in der rückwärtigen Endkappe des Eisendetektors.
IRONBB: Anzahl der angesprochenen Module im rückwärtigen Zentralbereich des Eisendetektors.
IRONFB: Anzahl der angesprochenen Module im vorderen Zentralbereich des Eisendetektors.
IRONFE: Anzahl der angesprochenen Module in der Vorwärtsendkappe des Eisendetektors.
IRONTOT: Gesamtanzahl der angesprochenen Module des Eisendetektors.

Bedingung dafür, daß Module aus dem Zentralbereich des Eisendetektors in einer der Größen mitgezählt werden, ist ein Signal aus zwei der inneren vier aus den fünf Triggerlagen. In der vorderen, inneren Endkappe sind es dagegen vier und in der vorderen, äußeren Endkappe 3 aus den fünf Triggerlagen. Für die hintere Endkappe werden drei aus den vorhandenen fünf Triggerlagen verlangt (siehe Abschnitt 2.2.4). Der Grund für diese unterschiedlichen Definitionen in Richtung der Strahlröhre liegt darin, daß vordere Detektorkomponenten bei Untergrundereignissen häufiger ansprechen. Schwache Triggerbedingungen in dieser Region würden zu hohen Triggerraten führen.

- **Kalorimetergrößen**

  Diese Größen werden in dieser Arbeit nicht zum Netztraining verwendet. Mit Ihnen arbeiten jedoch andere Netze, welche in den Abschnitten 4.3, 4.4 und 4.5 mit den Ergebnissen der dort trainierten Netze verglichen werden sollen. Die drei folgenden Größen enthalten Energiesummen aus verschiedenen Teilen des Flüssig-Argon-Kalorimeters. Es werden nachfolgend nur noch die Modulbezeichnungen genannt. Diese können anhand von Abbildung 2.5 nachvollzogen werden.

  LARIFE: Gesamtenergie der Module mit IF zu Beginn der Modulbezeichnung.
  LARFBE: Gesamtenergie der Module mit FB zu Beginn der Modulbezeichnung.
  LARCBE: Gesamtenergie der Module mit CB zu Beginn der Modulbezeichnung.

2.4.3 Die L4-Triggerstufe

In dieser Arbeit wird nach Erstellung eines Neuronalen Netzes für die Triggerstufe L2 auch der Effekte eines Einbaus dieses Netzes in die H1-Triggerumgebung untersucht. Dazu wird die Triggerstufe L4 simuliert. Zur Verdeutlichung der Funktionsweise der Triggerstufe L4 ist in Abbildung 2.16 das sukzessive Zusammenwirken der einzelnen Programmteile der L4-*Software* dargestellt.

Nach einer positiven Entscheidung durch einen L1-Subtrigger und gegebenenfalls die dazugehörige L2-Entscheidung werden in der ersten Stufe von L4 die Monitortrigger (siehe Abschnitt 2.4.1) verifiziert und das Ereignis bei positiver Entscheidung gespeichert. Im anderen Fall müssen die Ereignisdaten zuerst einen Untergrundfilter in der zweiten Stufe und in dritter Stufe eine Subtriggerverifikation passieren, bevor entweder durch die *Hardscale* auf vierter oder einen der *Finder* auf der fünften Stufe die Speicherung veranlaßt wird. Durch die *Hardscale* werden hauptsächlich Ereignisse mit hohen Werten in einzelnen auf Stufe L4 berechneten Größen wie der transversalen Energie, dem transversalen Impuls von Spuren oder dem Quadrat des Viererimpulsübertrages Q^2 zur Speicherung freigegeben. Die *Hardscale* wirkt auf alle bis zu dieser Stufe gelangten Ereignisse, während die installierten *Finder* nur für Ereignisse, welche bestimmte Subtrigger ausgelöst haben, konzipiert sind. Den *Findern* liegen insbesondere topologische Eigenschaften der Endzustände bestimmter Reaktionen zugrunde. Bei zu hoher Rate nach L4 kommt es auch hier wie bei L1-Subtriggern zu einer Abskalierung in der Anzahl zu speichernder Ereignisse, d.h. es wird nur jedes q-te Ereignis gespeichert (vergleiche Abskalierung auf L1 in Abschnitt 2.4). Der Abskalierungsfaktor q ist vom Ereignis abhängig. (Die Abskalierungsfaktoren auf Stufe L1 und Stufe L4 sind unabhängig.)

Die Entscheidungen der einzelnen, hier beschriebenen Programmstufen beruhen auf einer durch das L4-Programm geleisteten Ereignisrekonstruktion. Für diese Rekonstruktion steht die gesamte Detektordatenmenge des Ereignisses zur Verfügung. Auf spezielle, für diese Arbeit benutzte L4-Bedingungen wird an entsprechender Stelle eingegangen.

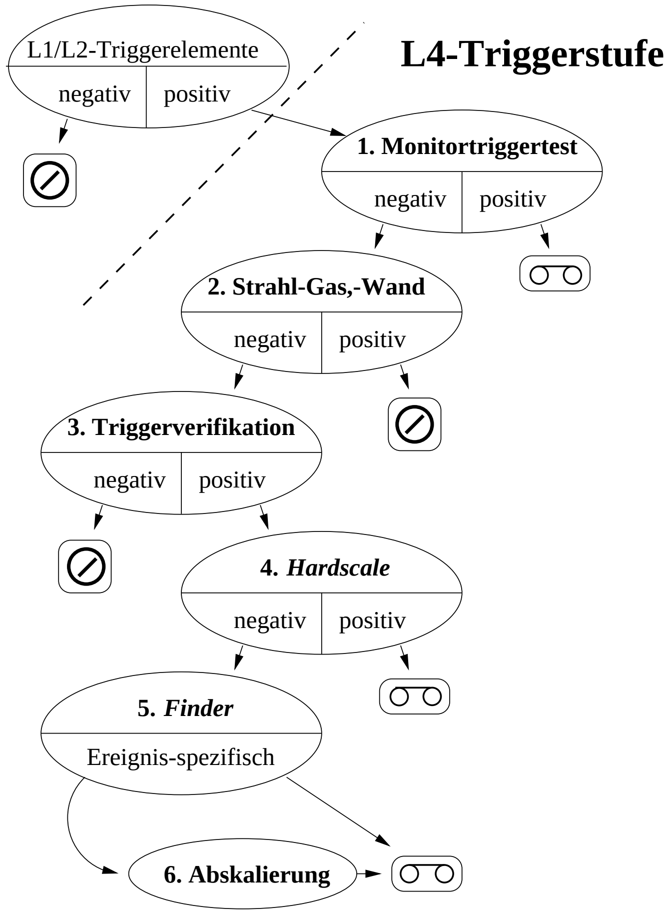


Abbildung 2.16: *Ablaufplan des Programms zur Bestimmung der L4-Triggerentscheidung.*

2.4.4 Die L5-Ereignisklassifikation

Auf der L5-Stufe findet eine Rekonstruktion und Klassifikation der Daten, die von der Triggerstufe L4 getriggert wurden, statt. Nur die Ereignisse, welche einer der L5-Klassen zugeordnet werden können, werden als rekonstruierte Daten gespeichert und weiter verarbeitet. Die in dieser Arbeit beschriebenen Ereignisselektionen basieren auf Ereignissen der Lepton-Klassen 17, 18, und 24. Klasse 17 dient der Selektion niederenergetischer Zerfallselektronen und Klasse 24 der Selektion von Ereignissen mit Myonen. Durch Klasse 18 hingegen werden Ereignisse selektiert mit Elektron-Paaren mit hohen Transversalimpulsen und Ereignisse, in denen keine Leptonen identifiziert wurden, in den Spurkammern jedoch zwei Spuren rekonstruiert wurden, die z.B. von Zerfallsleptonen elastisch erzeugter J/ψ-Mesonen oder Υ-Mesonen herrühren könnten. Die hier nicht explizit dargestellten Schnitte für die obigen Klassifizierungen sind ausnahmslos schwächer als die zur Selektion der Ereignisse zum Training der Neuronalen Netze im späteren dargestellten Schnitte.

Kapitel 3

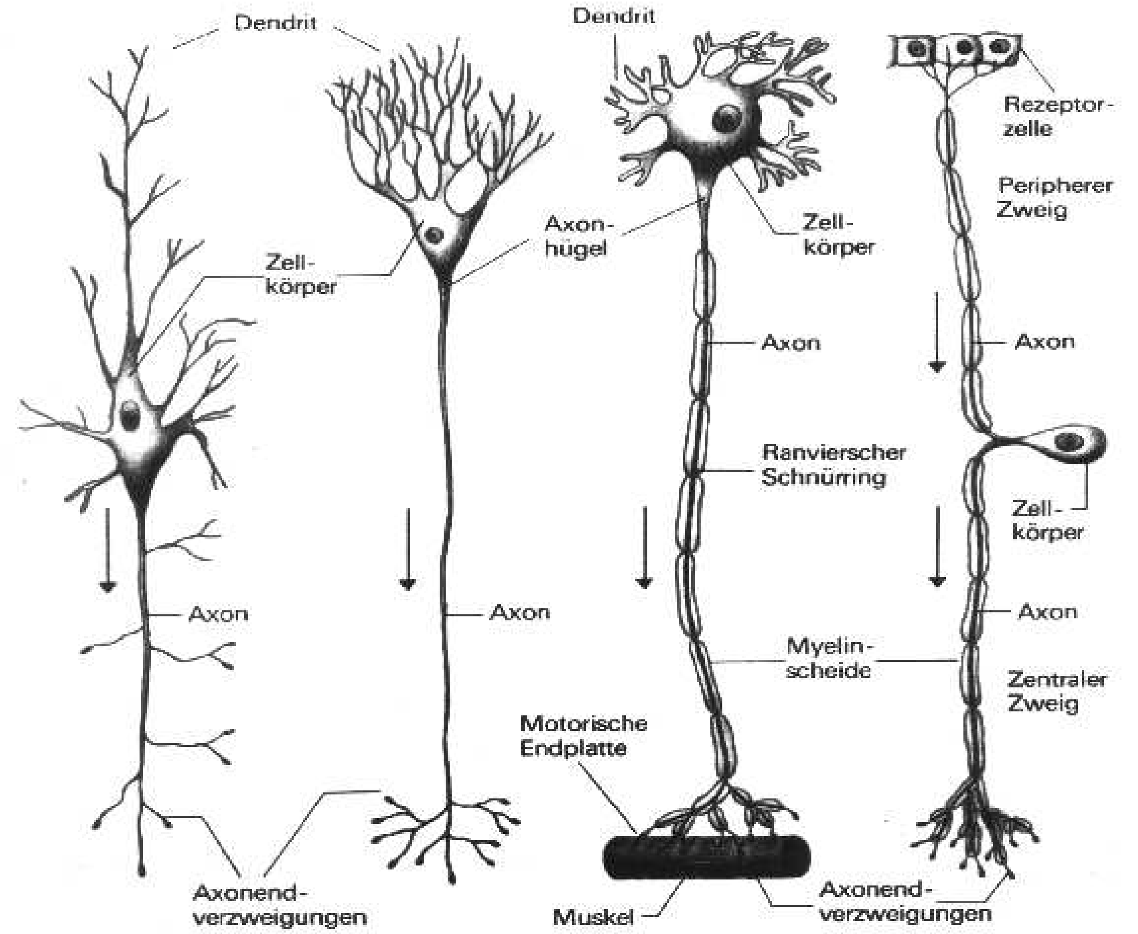
Neuronale Netze

> *Neural Networks are the second best way of doing just about anything.*
> John Denker

Unter künstlichen Neuronalen Netzen versteht man mathematische Modelle, mit denen man in Anlehnung an die Informationsverarbeitung in biologischen Nervensystemen versucht, grundlegende Eigenschaften solcher Systeme zu simulieren. Neuronale Netze werden durch Computerprogramme realisiert und dort eingesetzt, wo zum Beispiel funktionale Zusammenhänge oder hochdimensionale Korrelationen zwischen Größen gesucht werden, welche weder analytisch noch herkömmlich numerisch lösbar sind. Ganz allgemein sind dies Probleme der Klassifikation, der nichtlinearen Optimierung, der Prognose und der Mustererkennung. Dabei ist insbesondere der erste Punkt für diese Arbeit von Bedeutung. Um die bei H1 benutzten Neuronalen Netze sowohl technisch, wie auch vor dem Hintergrund biologischer Grundlagen der Analogiebildung zu Nervensystemen einordnen zu können, wird in diesem Kapitel eine Einführung in diesen Themenkomplex gegeben.

3.1 Physiologische Grundlagen

Komplexe, biologische Organismen existieren oftmals unter lebensfeindlichen und sich ständig ändernden Lebensbedingungen. Dafür maßgeblich ist ihre Fähigkeit, Informationen aus der Umgebung und von ihren eigenen Organen zu sammeln, zu verarbeiten und auf Grundlage angesammelter Erfahrungen und Erbinformationen zu angemessenen Reaktionen zu gelangen. Maßgeblich für diese komplexen Aufgaben der Informationsverarbeitung ist z.B. beim Menschen das Nervensystem. Man unterscheidet hier das zentrale Nervensystem, bestehend aus Gehirn und Rückenmark, sowie das periphere Nervensystem, zu welchem alle übrigen Nervenleitungen gehören. Das zentrale Nervensystem besteht beim Menschen aus ca. 10^{12} Nervenzellen oder auch Neuronen, welche untereinander mittels elektrischer Signale Informationen austauschen können.



a) Interneuronen b) Motoneuron c) sensorisches Neuron

Abbildung 3.1: *Diese Darstellung zeigt Vertreter der drei wesentlichen Arten biologischer Neuronen. Die Pfeile geben die Richtung der Signalleitung an.*

Die dafür benötigten Steuermechanismen sind ständig Veränderungen ausgesetzt; sie unterliegen einem Lernprozess. Diese Lernprozesse sind Gegenstand intensiver Forschung und sind zur Zeit nur in Teilaspekten verstanden [13].

3.1.1 Biologische Neuronen

In Abbildung 3.1 sind die drei grundlegenden Typen von Neuronen dargestellt [12]. Die multipolaren Interneuronen in a) stellen die gewöhnlichen signalleitenden und signalverarbeitenden Neuronen dar. Sie erhalten Spannungsimpulse über die Dendriten, welche bis zum Zellkörper weitergeleitet werden. Übersteigt das am Zellkörper entstehende, aus den Signalen aller Dendriten raum-zeitlich kumulierte Signal einen zellabhängigen Schwellenwert, so wird ein sogenanntes Aktionspotential über das Axon

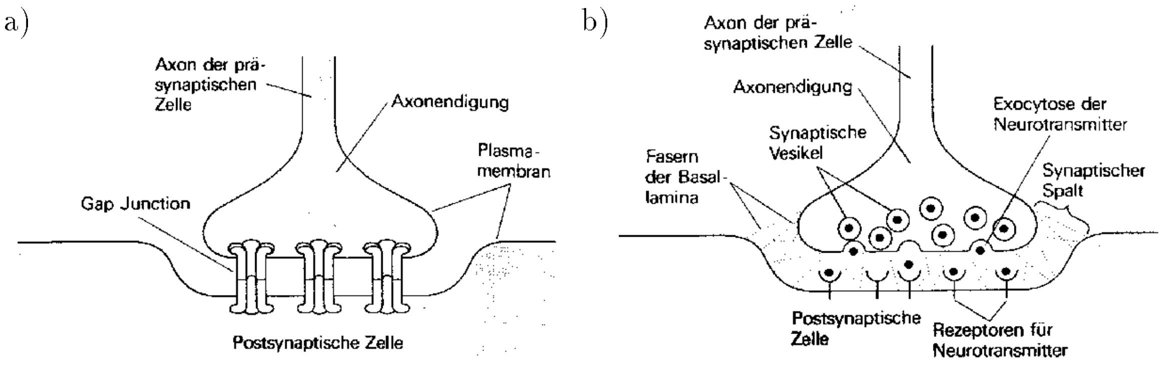


Abbildung 3.2: *Elektrische Synapse a) und chemische Synapse b).*

der Zelle ausgesandt: Die Nervenzelle *feuert*. Ein solches Potential entsteht durch ein Ionenkonzentrationsgefälle zwischen den intra- und extrazellulären wäßrigen Milieus der Zelle, verursacht durch einen Austausch von Na^+- und Ka^+-Ionen durch die Zellwand hindurch, der durch Kanalproteine gesteuert wird. Das Aktionspotential breitet sich ausschließlich entlang der in Abbildung 3.1 angegebenen Pfeilrichtung aus. Axone menschlicher Nervenzellen können bis zu 1 m lang werden. Die Potentiale liegen im Bereich von 10 mV bis 100 mV und werden zwischen Intra- und Extrazellulärraum gemessen.
Ein sogenanntes Motoneuron ist in b) dargestellt. Ein solches Neuron hat die Aufgabe, über den Kontakt seines Axons mit einer Muskelfaser eine Muskelbewegung auszulösen. Ein Vertreter der dritten Klasse von Neuronen ist in c) zu sehen. Ein solches sensorisches Neuron sorgt für die Weiterleitung von Informationen aus angesprochenen Rezeptorzellen, welche je nach Art und Lage, Informationen über den körpereigenen Zustand oder aber die Umgebung liefern.
Die dargestellten drei verschiedenen Neuronenfunktionalitäten werden wir in den Netzwerkarchitekturen der bei H1 verwendeten Netze als Neuronen in versteckter Schicht, Eingabe- und Ausgabeneuronen wiederfinden. Der Begriff Neuron wird in dieser Arbeit ebenfalls im Zeichen biologisch-technischer Analogiebildung je nach Zusammenhang in biologischer oder technischer Bedeutung verwendet.

3.1.2 Interzelluläre Potentialübertragung

Die Potentialübertragung von einer Nervenzelle zur nächsten geschieht über Synapsen. Durch Synapsen wird ein Potential im Axon einer Nervenzelle an die Dendriten einer anderen Nervenzelle übertragen. Es gibt hauptsächlich zwei Arten von Synapsen, elektrische und chemische (siehe Abbildung 3.2 a) und b)) [12].
Bei der elektrischen Synapse werden Spannungsimpulse durch leitende, aus Proteinen bestehende *Gap Junctions* über den synaptischen Spalt hinweg übertragen. Dabei ist die Signalleitungsgeschwindigkeit nahezu gleich der Geschwindigkeit bei der Signalausbreitung in einer Nervenzelle.

In chemischen Synapsen sorgen Neurotransmitter für die interzelluläre Signalübertragung. Neurotransmitter sind Makromoleküle und benötigen etwa 1 ms, um den synaptischen Spalt zu überwinden. Dadurch ist die Signalleitungsgeschwindigkeit über den synaptischen Spalt in chemischen Synapsen um zwei Größenordnungen kleiner als bei elektrischen Synapsen. Chemische Synapsen sind elektrischen Synapsen biologisch dort unterlegen, wo schnelle oder temperaturunabhängige Signalleitung wichtig ist. Außerdem können chemische Synapsen sowohl signalerregend wie auch signalhemmend auf ein Aktionspotential der postsynaptischen Zelle wirken.
Je nach Ausprägung der Synapse, d.h. Anzahl der *Gap Junctions* oder Umfang der Ausschüttung von Neurotransmittern und deren Rezeption, kann eine bestimmte Synapse bei biologischer Signalverarbeitung einer Vielzahl verbundener Neuronen mehr oder weniger im Vordergrund stehen. Der Synapse ist auf diese Weise ein biologisches Gewicht zugeordnet.

3.2 Technische Umsetzung von Neuronalen Netzen

Im Folgenden wird dargestellt, wie die zuvor beschriebenen, biologischen Erkenntnisse zur Lösung technischer Probleme genutzt werden können. Dazu bedient man sich des Modells der künstlichen Neuronalen Netze.

3.2.1 Allgemeine Analogiebildung

Die Entwicklung von Neuronalen Netzen ist durch drei zeitliche Perioden extremen Enthusiasmusses aber auch extremer Skepsis gegenüber den vorliegenden Ergebnissen geprägt. Ich folge in der nachstehenden Darstellung dieser geschichtlichen Entwicklung der Technik [14].

Die McCulloch-Pitts-Zelle

Bei der McCulloch-Pitts-Zelle wird die Eigenschaft einer Nervenzelle simuliert, nach raum-zeitlicher Addition aller von anderen Zellen übertragenen Signale bei Überschreiten eines Schwellenwertes zu *feuern*. Sie stellt die erste Idee einer den Nervenzellen nachgebildeten technischen Funktionseinheit dar und wurde von ihren Namensgebern 1943 entwickelt [14]. Eine solche Zelle ist in Abbildung 3.3 dargestellt. Sie wurde als binäre Schwellenwerteinheit realisiert. Sämtliche Eingabe- und Ausgabegrößen sind binär. Die Summe der Eingangssignale wird mit einem Schwellenwert verglichen. Bei Unterschreiten dieses Wertes ist die Ausgabe 0 sonst 1. Dieser Zusammenhang kann mit der Heavysidefunktion H dargestellt werden:

$$y = H(\sum_{i=1}^{n} x_i - \theta). \tag{3.1}$$

Hierbei sind die x_i die Eingabegrößen, θ ist der Schwellenwert und y die Ausgabegröße. Die Gesamtanzahl der Eingabegrößen ist n. Die Heavysidefunktion stellt hier die sogenannte Aktivierungsfunktion dar.

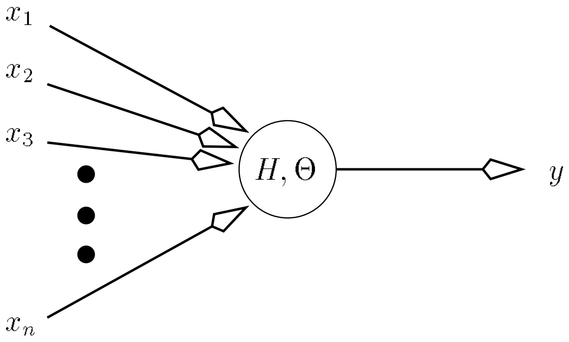


Abbildung 3.3: *Prinzipielle Darstellung einer McCulloch-Pitts-Zelle. Diese ist charakterisiert durch binäre Eingabegrößen x_i, die Heavyside-Funktion H, den Schwellenwert Θ und eine binäre Ausgabegröße y.*

Die einzelnen Eingaben können gegeneinander in beliebigen rationalen Verhältnissen gewichtet werden, wenn man mit der Anzahl der Eingabezweige für eine Größe über eins hinausgeht. Auf diese Weise ist eine einzige solche Zelle dazu in der Lage, im n-dimensionalen Raum der Eingangsgrößen eine lineare Trennung durch den Teilraum der positiven Koordinatenachsen mittels einer planaren Hyperebene zu vollziehen, wobei einer Seite der Ebene der Ausgabewert 0 und der anderen der Ausgabewert 1 zugeordnet wird.

Der Perzeptron-Lernalgorithmus von Rosenblatt

Der erste Lernalgorithmus für Neuronale Netze ist im Jahre 1960 von Rosenblatt veröffentlicht worden [14]. Ein Perzeptron geht aus einer McCulloch-Pitts-Zelle hervor, wenn man die Eingabegrößen vor der Summation mit zu den Eingängen gehörigen Gewichten multipliziert. Da die Gewichte des Perzeptrons reelle Zahlen darstellen, sind die Verhältnisse der Eingangssignalbewertung hierdurch nicht mehr auf rationale Zahlen beschränkt. Auch die Ein- und Ausgabegrößen sind hier reellwertig. Man kann folgende Gleichung angeben:

$$y = H(\sum_{i=1}^{n} w_i * x_i - \theta). \tag{3.2}$$

Hierbei ist w_i das Gewicht des i-ten Eingangs.
Rosenblatt hat nun für das Problem der linearen Trennung zweier Klassen von Eingabevektoren im n-dimensionalen Raum einen konvergenten Lernalgorithmus entwickelt, den Perzeptron-Lernalgorithmus. Auf dessen Details soll hier nicht weiter eingegangen werden. Diese sind in [14] zu finden. Wichtig ist hier der Zugang zur Biologie, da hier die Synapsen zwischen den Nervenzellen unterschiedlich starke Potentialübertragungseigenschaften besitzen und sogar hemmend wirken können, was im Perzeptronmodell negativen Gewichten entspricht. Eine allgemeinere Analogie erhält man, wenn man Perzeptronen mit gewichteten Signalleitungen miteinander verschaltet. Man erhält so ein Mehrschichtperzeptron, ähnlich der Netzarchitektur in Abbildung 3.4.

Das statistische Modell von Hopfield

Wohl einer der für einen Physiker spannendsten gedanklichen Gegenstände aus dem Themengebiet der Neuronalen Netze ergibt sich, wenn man die Leitungen zwischen den Neuronen in beiden Richtungen Informationen übertragen läßt und die oben als Aktivierungsfunktion verwendete Heavysidefunktion durch die Fermifunktion ersetzt. Faßt man nun den gewichteten Informationsfluß als durch interneuronale Wechselwirkung verursacht auf, so erhält man ein Hopfieldnetz, das ein dynamisches System mit temperaturabhängigem Gleichgewichtszustand darstellt. Die Temperatur findet in der Fermifunktion Berücksichtigung.
Man erlangt mit der Fermifunktion einen sanfteren Übergang zwischen den Ausgangszuständen *feuern* und *nicht feuern*, wie die Ausgangszustände der Neuronen auch bezeichnet werden. Es ergibt sich zur Errechnung der Neuronenausgabe folgende Gleichung.

$$y_i = g(\sum_{i=1}^{n} w_i * x_i - \theta) \tag{3.3}$$

Dabei muß für die Funktion g nicht unbedingt die Fermifunktion benutzt werden. Denkbar sind eine Vielzahl sigmoider Funktionen. Die Fermifunktion ist lediglich eine insbesondere für den Physiker wegen der Analogie zur statistischen Mechanik eingängige Möglichkeit, die dem Hopfield-Modell eigen ist. Die Fermifunktion lautet:

$$g(z) = \frac{1}{1 + exp(-\alpha \cdot z)} \tag{3.4}$$

Neben den existierenden Verfahren der statistischen Mechanik, die durch die Verwendung der Fermifunktion zur Verfügung stehen, ergeben sich Vorteile gegenüber anderen sigmoiden Funktionen beim Training von Neuronalen Netzen. Der Faktor α determiniert die Breite des Übergangsbereiches von Funktionswerten nahe bei Null zu Funktionwerten nahe bei eins. Diese Breite geht mit $1/\alpha$. Im Grenzfall $\alpha \to \infty$ erhält man die Heavyside-Funktion (siehe Seite 33). Es gilt $\alpha \propto 1/T$, wenn T die Temperatur darstellt.

3.2.2 Der Backpropagation-Algorithmus

Der Vorgang die richtigen Gewichte und Schwellenwerte für ein Neuronales Netz zu finden wird Lernvorgang oder Training genannt. Die Methodik, mit dem *Backpropagation*-Algorithmus [23, 24, 25, 26] ein Netz zu trainieren, kommt aus der Klasse der Algorithmen des überwachten, korrigierenden Lernens Neuronaler Netze. Diese Lernalgorithmusklasse zeichnet sich gegenüber dem unüberwachten Lernen dadurch aus, daß dem Netz beim Training eine vorgegebene Klasseneinteilung präsentiert wird, so daß die Korrektur der Gewichte auf Grund der Abweichung der aktuellen Ausgabe vom Sollausgabewert stattfindet. Beim unüberwachten Lernen hingegen können Netze auch so trainiert werden, daß diese eine in der Trainingsmenge *versteckte*, mögliche Klassifizierung beim Training selbst herausfinden. Beim korrigierenden Lernen hängt die Stärke der Korrektur linear von der Größe der Abweichung ab. Der *Backpropagation*-Algorithmus ist ein Gradientenabstiegsverfahren.

Die grundlegende Idee des *Backpropagation*-Algorithmus ist, daß der in bezug auf die Ergebnisse eines Netzes für die Trainingsmenge beste Zustand des Netzes erreicht wird, wenn man eine zuvor definierte Fehlerfunktion, welche die Abweichung der aktuellen Netzausgabe y zur Sollausgabe o beschreibt, minimiert. Da hier nur der Spezialfall mit einem Neuron in der Ausgabeschicht behandelt wird, sind y und o skalare Größen. Mit der in Abbildung 3.4 veranschaulichten Nomenklatur kann die benutzte Fehlerfunktion dargestellt werden durch:

$$E(W) = \frac{1}{2} \sum_{q=1}^{N} (y(\vec{x}_q, W) - o(\vec{x}_q))^2. \tag{3.5}$$

Dabei sind $\vec{x}_q$ die Vektoren der Eingabegrößen der Trainingsereignisse und W ist die Menge der Gewichte und Schwellenwerte im Neuronalen Netz. N ist die Anzahl der Ereignisse der Trainingsereignismenge, deren Eingabegrößen dem Netz zwischen zwei Korrekturen der Gewichte präsentiert werden. Diese Anzahl der Präsentationen eines Trainingsereignisses vor erneuter Gewichtekorrektur kann zwischen 1 und der Gesamtzahl aller Trainingsereignisse gewählt werden. Bei dem für diese Arbeit durchgeführten Netztraining wurde nach jedem präsentierten Ereignis eine Korrektur der Gewichte durchgeführt. Ein Abschnitt des Trainings, während dem sämtliche Ereignisse der Trainingsmenge dem Netz präsentiert wurden, wird Epoche genannt.

Idealerweise gilt es, die Menge an Vektoren W_0 zu finden, so daß $E(W_0) \leq E(W)$ für alle $W \neq W_0$ gilt, d.h. das globale Minimum von $E(W)$ zu bestimmen. Mit den so ermittelten Gewichtsvektoren erhält man nach obiger Definition der Fehlerfunktion die geringste quadratische Abweichung der Ausgabe des Netzes y von der Sollausgabe o.

Zuerst werden die Gewichtsvektoren vor der Ausgabeschicht, dann die Vektoren vor den Neuronen der versteckten Schicht korrigiert. Die Korrektur erfolgt in Richtung des negativen Gradienten der Fehlerfunktion bezüglich der jeweils zu korrigierenden Größen. Es wird hier nur der Spezialfall mit einem Neuron in der Ausgangsschicht behandelt, da nur solche Netze für die verwendete Technik relevant sind. Es seien die

Ausgabewerte der versteckten Schicht h_j und der Ausgabeschicht y gegeben durch:

$$h_j = g(\sum_{i=0}^{n} \omega_{ij} x_i) \tag{3.6}$$

und

$$y = g(\sum_{j=0}^{m} \tilde{\omega}_j h_j). \tag{3.7}$$

Dabei ist g die in Abschnitt 3.2.1 beschriebene Fermifunktion. Es gilt $x_0 = -1$ und $h_0 = -1$, so daß die $\omega_{0j}(j = 1, ..., n)$ und $\tilde{\omega}_0$ als Schwellenwerte dienen. Im Einklang mit der Notation aus Abbildung 3.4 gilt $\omega_{0j} = \Theta_j(j = 1, ..., n)$ und $\tilde{\omega}_0 = \tilde{\Theta}$.
Zu Beginn eines Netztrainings werden die Gewichte mit einem Zufallsgenerator erzeugt. Nun kann die Korrektur der Gewichte für die Ausgabeneuronen mit Hilfe der Kettenregel der Differentiation folgendermaßen dargestellt werden:

$$\Delta\tilde{\omega}_j = -\eta \frac{\partial E}{\partial \tilde{\omega}_j} + \alpha \Delta \tilde{\omega}_j^{zuvor} = -\eta (y - o) \cdot g'(\sum_{q=0}^{m} \tilde{\omega}_q h_q) \cdot h_j + \alpha \Delta \tilde{\omega}_j^{zuvor}. \tag{3.8}$$

Dabei ist $g'(z) = \frac{\partial g}{\partial z}$. Die partielle Ableitung der Fehlerfunktion nach dem Gewicht ist an der Stelle der aktuellen Werte der Gewichte und Schwellenwerte zu nehmen. Mit dem Lernparameter η gibt man die Stärke der Korrektur vor. Der Momentumfaktor α gibt einem die Möglichkeit, die Korrektur aus dem vorherigen Korrekturschritt $\tilde{\omega}_j^{zuvor}$ in die aktuelle Korrektur $\tilde{\omega}_j$ miteinzubeziehen. Hierbei trägt man der Tatsache Rechnung, daß das Netz idealer Weise globale Eigenschaften der Ereignisse und nicht Ereignisspezifika erlernen soll. Weiterhin dient der Momentumfaktor, der zusammen mit der vorherigen Korrektur $\Delta\tilde{\omega}_j^{zuvor}$ auch Dämpfungsterm genannt wird, der Vermeidung von Oszillationen. Lange, flache Bereiche der Fehlerfunktion werden mit Hilfe des Momentumfaktors schneller überwunden, und man kann auf diese Weise das Training beschleunigen.
Nach Korrektur der Gewichte vor dem Ausgabeknoten erfolgt auf analoge Weise die Korrektur der Gewichtsvektoren vor der versteckten Schicht mit

$$\Delta\omega_{ij} = -\eta \frac{\partial E}{\partial \omega_{ij}} + \alpha \Delta \omega_{ij}{}^{zuvor} = -\eta \cdot \delta \cdot \tilde{\omega}_j \cdot g'(\sum_{p=0}^{n} \omega_{pj} x_p) \cdot x_i + \alpha \Delta \omega_{ij}^{zuvor} \tag{3.9}$$

mit

$$\delta = (y - o) \cdot g'(\sum_{q=0}^{m} \tilde{\omega}_q h_q). \tag{3.10}$$

Wie schon für Formel 3.8 ist auch hier die partielle Ableitung der Energie an der Stelle der aktuellen Werte der Gewichte und Schwellenwerte zu nehmen.
Die Methodik des Trainierens kann beinhalten, daß man mit relativ großer Lernrate beginnt, welche man von Epoche zu Epoche verringert. Man hält so die Gefahr klein,

daß man bei der Minimierung der Fehlerfunktion in ein lokales Minimum gerät, obwohl möglicherweise ein anderes, z.B. globales, Minimum existiert, das eine bessere Lösung des Problems darstellt. Aus dem gleichen Grunde ist es generell ratsam, ein Netz mit verschiedenen Parameterkonstellationen zu trainieren.
In den obigen Gleichungen wird deutlich, daß bei Anwendung des *Backpropagation*-Algorithmus als Aktivierungsfunktion eine differenzierbare Funktion gewählt werden muß.
Zum *Backpropagation*-Algorithmus existieren weiterführende Verfahren, um die Parameter η und α von vornherein an ein Problem anzupassen [27] und um dem Übertrainieren (siehe Abschnitt 3.2.6) eines Netzes vorzubeugen [28]. Verfahren mit diesen Eigenschaften werden hier nicht verwendet.

3.2.3 Das Konzept Neuronaler Netze für den L2-Trigger

Bei H1 werden *Feedforward*-Neuronale-Netze verwendet. In Abbildung 3.4 ist die dazugehörige Netzarchitektur dargestellt. Die Netze bestehen aus Neuronen oder Knoten, die miteinander durch sogenannte Kanten oder Leitungen verbunden sind.

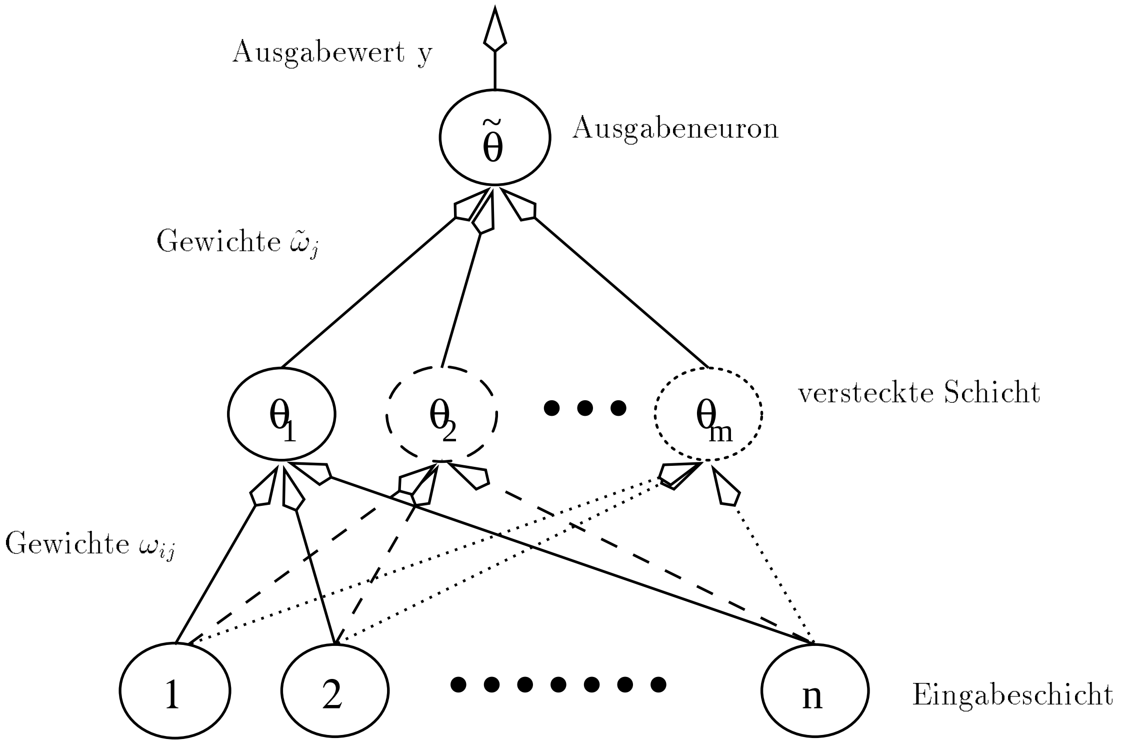


Abbildung 3.4: *Prinzipielle Darstellung einer Neuronalen Feedforward-Netzwerkarchitektur mit der im Text verwendeten Nomenklatur. Die Neuronen der Eingabeschicht sind bis n durchnummeriert, diejenigen der versteckten Schicht sind mit dem Schwellenwert* $\Theta_i (i = 1, ..., m)$ *versehen. Dargestellt ist hier der Spezialfall eines Netzes mit einem Ausgabeknoten, wie sie auch bei H1 verwendet werden.*

Feedforward bedeutet, daß der Informationsfluß ausschließlich aus Richtung der Eingabeschicht in Richtung der Ausgabeschicht verläuft. Jedes der Neuronen liefert nach Summation der gewichteten Eingaben eine schwellenwertabhängige Ausgabe. Die hierbei benutzte Aktivierungsfunktion ist die Fermifunktion.

Nach Wahl der Netzarchitektur und der Aktivierungsfunktion ist ein Netzzustand oder auch eine Netzkonfiguration durch die Angabe der Gewichte und der Schwellenwerte an den Neuronen vollständig charakterisiert. Der Vorgang des *Lernens* bedeutet für ein Neuronales Netz die Einstellung der Gewichte und Schwellenwerte. Das sind in der benutzten Nomenklatur alle ω_{ij} und $\tilde{\omega}_j$ mit $(i = 1, ..., n)$ und $(j = 1, ..., m)$, $\Theta_1, ..., \Theta_m$ und $\tilde{\Theta}$.

Die Einstellung der Gewichte erfolgt durch das *Backpropagation*-Verfahren. Zuvor werden die Gewichte mit Hilfe eines Zufallsgenerators initialisiert. Auf dieses Verfahren und seine Funktionsweise wurde in Abschnitt 3.2.2 näher eingegangen. Ein Netz besitzt für jede der gewählten Eingangsgrößen einen Knoten in der ersten Neuronenschicht, der Eingabeschicht. Nach paralleler Verarbeitung der Multiplikationen und Summationen an den Leitungen zu den Knoten der versteckten Ebene, werden deren mit der Sigmoidfunktion bewerteten Ausgabewerte an den Knoten in der Ausgabeschicht weitergegeben. Die drei hier beschriebenen Arten von Neuronen mit ihren spezifischen Funktionen des Datenempfangs, der Datenweiterleitung und der Datenausgabe zwecks Triggerentscheidung können in direkter Analogie zu den drei in Abschnitt 3.1.1 dargestellten Arten von biologischen Neuronen, den sensorischen, Inter-, und Motoneuronen gesehen werden.

Die Möglichkeit der Anwendung paralleler Datenverarbeitung, die *Feedforward*-Netzarchitekturen ganz allgemein eigen ist, gewährleistet die Einhaltung der Zeitdauer der Triggerentscheidung bei H1. Hier liegt auch der Grund, keine rekursiv arbeitenden Netze, das sind Netze, die Leitungsrückkopplungen enthalten, zu verwenden. Hierbei würde die Entscheidungszeitdauer durch die Konvergenz iterativer Netzalgorithmen determiniert werden.

3.2.4 Die Bedeutung der versteckten Neuronenschicht

Im menschlichen Gehirn findet man sehr viel mehr Schichten von Neuronen als die drei Ebenen der bei H1 verwendeten Netze. Interessant sind die funktionalen Konsequenzen, wenn man von einem Netz ohne versteckte Neuronenschicht zu einem mit versteckter Neuronenschicht übergeht. Ohne die versteckte Schicht würde das Netz die Eingabevektoren nur linear transformieren und entscheiden, auf welcher Seite einer durch die Gewichte bestimmten Ebene im Raum der Eingabevektoren der aktuelle Eingabevektor liegt. Durch die versteckte Schicht ist es möglich, auch sogenannte nicht linear separierbare Probleme zu lösen, d.h. topologisch nicht zusammenhängende Gebiete im Eingaberaum am Ausgabeneuron in einer Klasse zu vereinen [19]. Auf eingängige Weise kann man sagen, der Ausgabeknoten konstruiert ein Volumen aus den Trennebenen, die durch die einzelnen versteckten Knoten dargestellt werden [20].

3.2.5 Hardware für L2

In der konventionellen Computertechnik verwendet man wenige Recheneinheiten, auf denen komplexe Programme laufen. Im Gegensatz dazu stehen die Datenverarbeitung des menschlichen Gehirns und der in *Hardware* realisierten Neuronalen Netze. Hier hat man es mit vielen einfach aufgebauten Neuronen oder Recheneinheiten zu tun. Diese sind technisch als gewichtete Summierverstärker realisiert, es läuft auf einer Recheneinheit also nur ein einfaches Programm ab. Das Besondere der Leistungsfähigkeit Neuronaler Netze liegt in der Verschaltung dieser einfachen Einheiten und in ihrer hohen funktionalen Parallelität.
Der bei H1 benutzte CNAPS(Connective Network of Adaptive Processors)-Prozessor läßt 8 *bit* Ein- und Ausgangsgrößen zu. Dies entspricht einer Ausgabe zwischen 0 und 255, wenn man ohne Vorzeichen*bit* rechnet. Die Gewichte werden durch 16 *bit*- Zahlen dargestellt und die Sigmoidfunktion bildet 16 *bit* auf 8 *bit* Zahlen ab . Durch die Realisierung der Netze in *Feedforward*-Architektur beträgt die Zeit bis zur Entscheidungsfindung wegen der parallelen Verarbeitung der Signale maximal 10 μs, unabhängig davon, wie groß das Netz gewählt wurde. Die *Hardware* für die Neuronalen Netze bei H1 ist für Netzarchitekturen mit 64-64-1 als maximalen Anzahlen an Knoten ausgelegt. Das bedeutet höchstens 64 Eingabe-, 64 versteckte und ein Ausgabeknoten.
Die Datenvorverarbeitung vor dem Netzdurchlauf erfolgt durch die DDB (*Data Distribution Board*). Hier treffen die Daten innerhalb von 5 μs nach positiver L1-Triggerentscheidung ein. In den folgenden 5 μs werden die Daten vorverarbeitet und in den CNAPS-Prozessor geladen. Der Prozessor benötigt 10 μs für die Errechnung der Entscheidung. Auf diese Weise wird die Triggerentscheidung innerhalb einer Entscheidungszeitspanne von 20 μs gewährleistet [17, 18].
Auf Grund der vorgegebenen maximalen Zeitdauer bis zur Triggerentscheidung können als Eingabegrößen nicht beliebig komplizierte Größen benutzt werden. Geeignete Größen sind im Ereignisvektor zusammengefaßt. Eine Darstellung der in dieser Arbeit verwendeten Eingabegrößen wurde bereits in Abschnitt 2.4.2 gegeben.

3.2.6 Software für L2

Die Software für Neuronale Netze in Verbindung mit dem L2-Trigger bei H1 gestattet die Entwicklung und Simulation von Neuronalen Netzen zur späteren Implementation am Experiment.

Das Training eines Neuronalen Netzes bei H1 gestaltet sich wie im Ablaufplan in Abbildung 3.5 dargestellt. Nachdem die für das Training des Netzes gewünschten Ereignisdatensätze vorliegen, werden aus diesen die für das Netztraining relevanten Daten mit dem Programm *NETSIM* in einem komprimierten Datenformat in den Ereignisvektoren abgelegt. Diese Ereignisvektoren werden vom Programm *MIXER* zu Trainings- und Testdateien verarbeitet. Mit diesen Dateien können mit dem *TRAINER*-Programm Neuronale Netze erstellt und erste Tests mit den Testdateien durchgeführt werden.

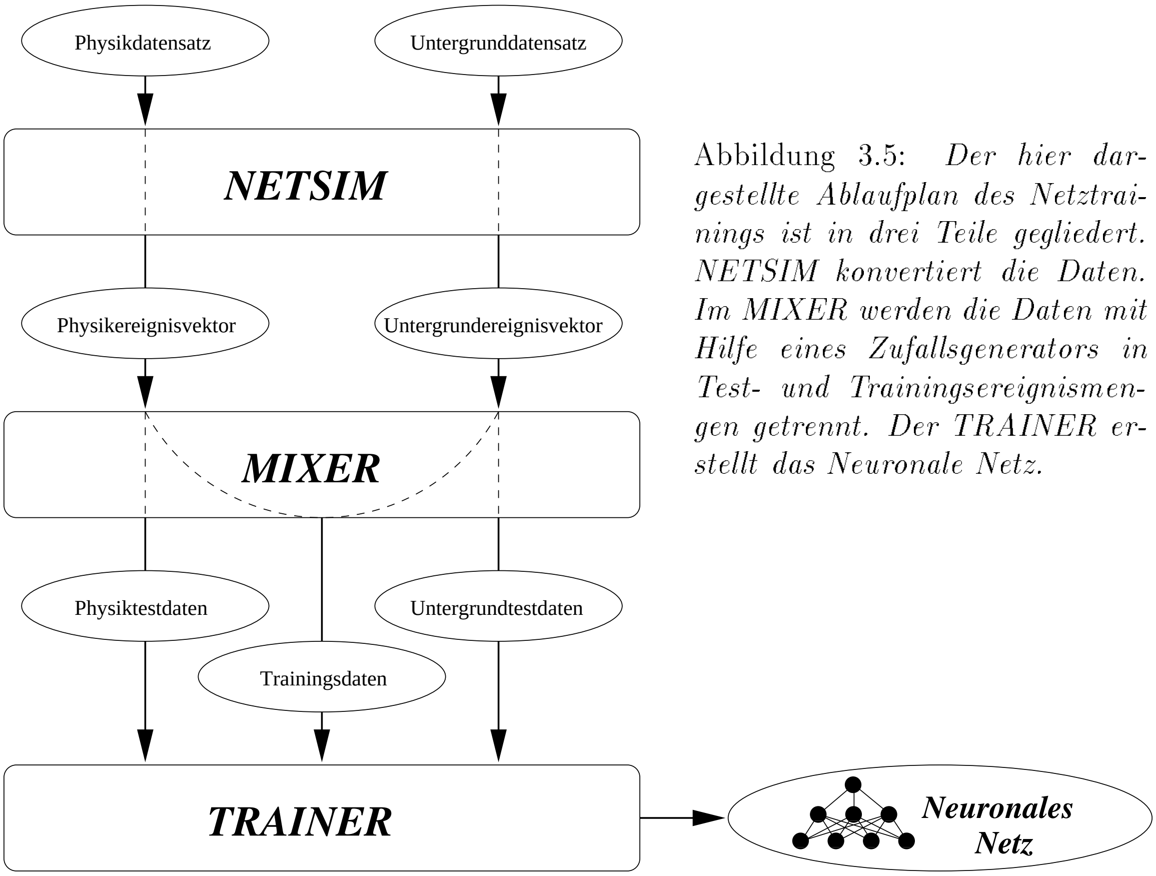


Abbildung 3.5: *Der hier dargestellte Ablaufplan des Netztrainings ist in drei Teile gegliedert. NETSIM konvertiert die Daten. Im MIXER werden die Daten mit Hilfe eines Zufallsgenerators in Test- und Trainingsereignismengen getrennt. Der TRAINER erstellt das Neuronale Netz.*

Die Sollausgabewerte der Netze werden zu 1 für Physikereignisse und 0 für Untergrundereignisse gewählt.

Für die Beurteilung der Güte eines trainierten Netzes ist dessen Fähigkeit entscheidend, Physikereignismenge und Untergrundereignismenge auf Grund allgemeiner, topologischer Eigenschaften dieser Mengen zu unterscheiden. *Lernt* das Netz hingegen spezielle Eigenschaften der Trainingsereignismenge, so nennt man das Netz übertrainiert. Ein übertrainiertes Netz identifiziert man durch ein deutlich schlechteres Ergebnis bei der Simulation des Netzes über der Testereignismenge als dies bei der Trainingsereignismenge der Fall war. Behandelt ein Neuronales Netz beide Mengen in etwa gleich, so hat das Netz eine gute Generalisierungseigenschaft. Über die Generalisierungseigenschaft eines Neuronalen Netzes kann man nur eine Aussage machen, wenn die Testdatenmenge von der Trainingsdatenmenge nahezu statistisch unabhängig ist. Daher werden diese Mengen mit Hilfe eines Zufallsgenerators getrennt.

Anzumerken ist, daß durch dieses Verfahren die zum Training zur Vefügung stehenden Ereignisse um die Testereignisse reduziert werden, wodurch die Statistik des Trainings verschlechtert wird. Hier ist ein Kompromiß zwischen der Statistik des Trainings und der Statistik des Tests zu wählen.

Kapitel 4

Training und Eigenschaften der Netze

Im Rahmen dieser Arbeit wird mit unterschiedlichen Zielsetzungen versucht, für die Subtrigger s34, s52 und s54 Neuronale Netze zu entwickeln. Diese drei Trigger sind für Ereignisse konzipiert, die durch zwei aus dem Zerfall eines diffraktiv erzeugten J/ψ-Mesons stammende Leptonen charakterisiert sind. Ebenso wird erwartet, daß durch diese Subtrigger auch Υ-Mesonen, die in zwei Leptonen zerfallen, getriggert werden. Im folgenden werden nach einer Beschreibung der bei H1 anfallenden Untergrundereignisse die Untersuchungen zu den drei Subtriggern nacheinander abgehandelt.

4.1 Untergrund bei Hera

Die Klasse der Untergrundereignisse ist die Menge an Ereignissen, die getriggert werden, aber nicht von einer ep-Reaktion stammen. Auf die verschiedenen Möglichkeiten der Herkunft des Untergrundes und eine daraus folgende Klassifikation wird hier eingegangen [1].

Im Vordergrund steht der **protoninduzierte Untergrund**, welcher sich in Form von Strahl-Gas-Reaktionen und Strahl-Wand-Reaktionen äußert. Bei den **Strahl-Gas-Reaktionen** handelt es sich um Kollisionen von Protonen mit Restgasatomen in der Strahlröhre insbesondere in der Nähe des nominellen Wechselwirkungspunktes. Die entstehenden Reaktionsprodukte gelangen in den Detektor und lösen die in Abschnitt 2.4 beschriebenen Triggerelemente aus. Die Rate dieser Reaktionen hängt linear vom Protonenstrom, der Flächendichte des Gases und dem Wirkungsquerschnitt der beteiligten, starken Reaktionen ab.

Protonstrahl-Wand-Reaktionen haben ihren Ursprung insbesondere an den Stellen geringer Apertur im Speicherring. Der Protonenstrahl ist hiervon stärker betroffen als der Positronenstrahl, da bei Protonen transversale Ablagen des Strahls von der Sollbahn nicht wie bei Positronen durch Strahlungsdämpfung verringert werden. Die

Sekundärteilchen von Strahl-Wand-Reaktionen können in der Strahlröhre innerhalb des Detektors erneut reagieren und den Trigger auslösen. Zur Reduktion dieses Typs von Untergrund sind auf den geraden Beschleunigungsstrecken von HERA Kollimatoren im Einsatz.

Weiterer Ereignisuntergrund wird durch **kosmische Strahlung** verursacht. Insbesondere besteht dieser aus kosmischen Myonen, welche in der Erdatmosphäre entstehen. Ab einer Energie von ca. 7 GeV ist ein kosmisches Myon in der Lage den gesamten Detektor senkrecht im Bereich der zentralen Spurkammern zu durchqueren [29]. Insbesondere Zerfallsereignisse mit zwei Zerfallsmyonen, wie sie in dieser Arbeit betrachtet werden, sind auf Triggerebene schwer von der Detektorsignatur kosmischer Myonen zu unterscheiden.

Weitere Untergrundquellen sind **elektronisches Rauschen** insbesondere in der Elektronik des Flüssig-Argon-Kalorimeters und **Synchrotronstrahlung** aus dem Bereich der Beschleunigung der Positronen beim letzten Ablenkmagneten vor der Wechselwirkungszone. Auf die letzten beiden Möglichkeiten wird im Rahmen dieser Arbeit nicht näher eingegangen.

Die Signatur und Ausprägung von Untergrundereignissen ist stark von den momentanen Srahlbedingungen bei HERA und den Einstellungen bei H1 abhängig.

Für die weiteren Untersuchungen werden Untergrundereignisse verursacht durch Strahl-Wand-, Strahl-Gas-Reaktionen und Myonen aus Reaktionen der kosmischen Strahlung im Vordergrund stehen.

4.2 Subtrigger s34

Der Subtrigger s34 ist für diffraktiv erzeugte J/ψ- und Υ-Mesonen konzipiert worden. Dies ist an den Triggerelementen, aus denen s34 besteht, zu erkennen. In der folgenden Beschreibung von Subtrigger s34 bedeuten

|| logisches ODER und & logisches UND.

Definition des Subtriggers s34 für die Jahre 1996 und 1997:

- s34: Mu_Bar & DCRPh_Ta & DCRPh_TNeg & DCRPh_THig & (zVtx_small || zVtx_Cls)

Die Triggerelemente sind im einzelnen:

Mu_Bar : Es wird mindestens ein Myonkandidat im Zentralbereich des Eisendetektors (siehe Abbildung 2.7, Module 16-47) getriggert.

DCRPh_Ta : Mindestens in einem Sektor des Spurkammertriggers hat eine Maske angesprochen (siehe Seite 24).

DCRPh_TNeg : Mindestens eine negative Maske des DCrphi-Triggers wurde gesetzt.

DCRPh_THig : Mindestens ein Teilchen hat eine Maske mit einem transversalen

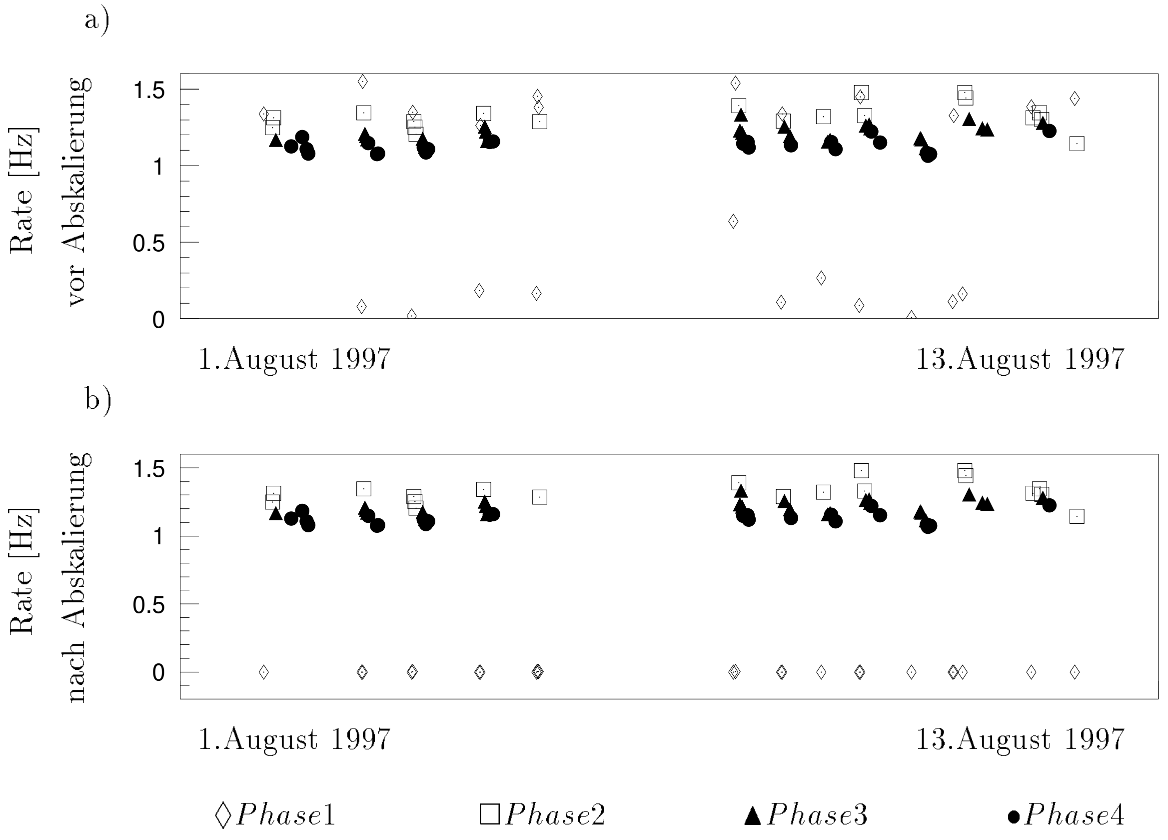
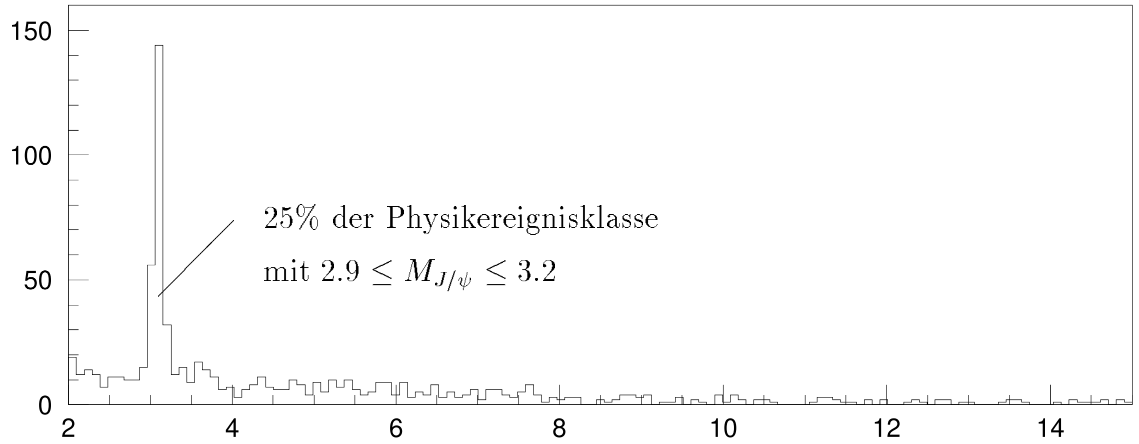


Abbildung 4.1: *In dieser Darstellung finden sich die Raten des Subtriggers s34 a) ohne und b) mit Abskalierungsfaktor. Der Abskalierungsfaktor ist 1000 für Phase 1, weshalb die abskalierte Rate b) in dieser Phase Null beträgt. Für die Phasen 2 bis 4 läuft s34 ohne Abskalierung.*

Impuls von mehr als 800 MeV getriggert.
zVtx_small : Im z-Vertex-Intervall mit den meisten Einträgen sind weniger als 6 Einträge. Die Summe der Einträge aller anderen Intervalle zusammen ist kleiner als 4. Die Summe aller Einträge ist nicht Null.
zVtx_Cls : Alle Einträge im z-Vertex-Histogramm befinden sich innerhalb von höchstens vier benachbarten Intervallen.

Die Einschränkung im z-Vertex-Histogramm bedingt, daß ein Ereignis nur getriggert wird, wenn nur wenige Spuren in Richtung des nominellen Vertex zeigen; auf diese Weise wird die Anzahl solcher Spuren in einem Ereignis begrenzt. Subtrigger s34 wird damit zu einem Trigger für myonische Zerfälle von diffraktiv erzeugten J/ψ-Mesonen, die durch niedrige Spurmultiplizität gekennzeichnet sind.
In Abbildung 4.1 sind die Ereignisraten von Subtrigger s34 mit und ohne Abskalierungsfaktor q (siehe Abschnitt 2.4) für die Datennahme der ersten zwei Wochen vom



Invariante Masse der Myonpaare

Abbildung 4.2: *Invariante Masse der Myonpaare der s34 Physikselektion. Die in dieser Darstellung einzige, statistisch signifikante Resonanz ist bei ungefähr 3 GeV zu erkennen und rührt vom J/ψ-Meson her.*

August 1997 gezeigt.
Es ist aus dieser Abbildung ersichtlich, daß Subtrigger s34 in Phase 1 wegen einer Abskalierung um den Faktor $q = 1000$ nicht zur Ereignisaufzeichnung beiträgt, während in den Phasen 2 bis 4 bei einer Rate von 1-1.5 Hz keine Abskalierung besteht.

4.2.1 Die Selektion der Trainingsdaten

Im folgenden werden die selektierten Daten, welche die interessante Physik enthalten als Physikereignisklasse bezeichnet, obwohl noch ein Anteil an Ereignissen vorhanden ist, der bei einer eingehenderen Analyse als Untergrund verworfen werden würde.
Ebenfalls befinden sich in der Untergrundereignisklasse, die selektiert wird, auch Physikereignisse. Dieser Anteil ist jedoch sehr gering und kann vernachlässigt werden. Vorhandene, kleine Anteile an falsch klassifizierten Ereignissen können beim Training Neuronaler Netze zur Verbesserung der Generalisierungseigenschaften beitragen. Dies wird im Rahmen dieser Arbeit jedoch nicht weiter untersucht.

Selektion der Physikereignismenge

Die mit Hilfe von Subtrigger s34 getriggerten Ereignisse wurden, nachdem sie den L4-Trigger passiert haben und von der L5-Ereignisklassifikation in eine der Klassen 17, 18 oder 24 (siehe Abschnitt 2.4.4) eingeteilt wurden, mit folgenden Schnitten weiter selektiert [30].

- Mindestens ein im Eisendetektor oder im Kalorimeter identifiziertes Myon muß im Ereignis vorhanden sein.
- Es dürfen höchstens vier gute Spuren (siehe Anhang A) in den Spurkammern rekonstruiert worden sein.
- Ein Schnitt gegen kosmische Myonen muß erfüllt werden.
  Dabei muß gelten:
  $$(\frac{180^\circ - \triangle\phi}{4})^2 + (\frac{180^\circ - \triangle\theta}{8})^2 \geq 1. \tag{4.1}$$
  Hierin stellen $\triangle\phi$ und $\triangle\theta$ den azimutalen und polaren Winkelunterschied zwischen den Spuren der beiden Myonen dar.
- Invariante Masse der beiden Myonen $M_{\mu\mu} \geq 2$ GeV.

In Abbildung 4.2 ist die invarianten Masse der beiden Myonen nach obiger Selektion dargestellt. Eine weitere Einschränkung der Daten in der invarianten Masse war auf Grund der zu geringen Statistik zum Training eines Netzes nicht möglich.

Selektion der Untergrundereignismenge

Zum Training der Netze wurden aus Rohdaten aus dem Zeitraum 16.März bis 26.April 1997 aus ingesamt 23 verschiedenen Datennahmeeinheiten (siehe Abschnitt 2.4.1) Ereignisse vor der Abskalierung von s34 für das Netztraining selektiert. Die verwendeten Ereignisse wurden ohne L4-Entscheidung direkt nach der L1/L2-Entscheidung gespeichert (L4-*transparent*). Es standen für das Training 2828 dieser Untergrundereignisse zur Verfügung.

Der große Zeitraum und die große Anzahl an Datennahmeeinheiten wurde gewählt, um zeitabhängige Einflüsse auf die Untergrunddaten zu vermeiden bzw. zu mitteln. Dies erhöht die Wahrscheinlichkeit, aus diesen Daten ein Neuronales Netz zu erhalten, welches den Untergrund gut von der zu triggernden Physik trennen kann. Mit 'gut' ist hier gemeint, daß angestrebt wird, Netze zu trainieren, die den Untergrund bei Hera in seinen typischen Eigenschaften erkennen, auch wenn diese Eigenschaften in einem bestimmten Rahmen zeitabhängig sind.

Der selektierte Untergrund besteht aus zwei topologisch sehr unterschiedlichen Klassen von Ereignissen. Zum einen sind dies kosmische Myonen (siehe Abbildung 4.3), zum anderen Ereignisse mit vielen geladenen Teilchen, die sich zum Teil nahezu parallel zum Strahlrohr bewegen (siehe Abbildung 4.4). Nur auf Grund dieser Parallelität der Spuren zur Strahlröhre werden diese Ereignisse von s34 getriggert, da durch solche Teilchenspuren keine Einträge im z-Vertex-Histogramm entstehen. Subtrigger s34 fordert durch die Bedingungen zVtx_small und zVtx_Cls ein 'kleines' oder 'enges' z-Vertex-Histogramm (siehe Definition auf Seite 43). In Abbildung 4.4 a) ist das z-Vertex-Histogramm für das abgebildete Vielspurereignis gezeigt. Auf Grund der wenigen Einträge in das Histogramm erfüllt dieses Ereignis die oben genannte s34-Bedingung.

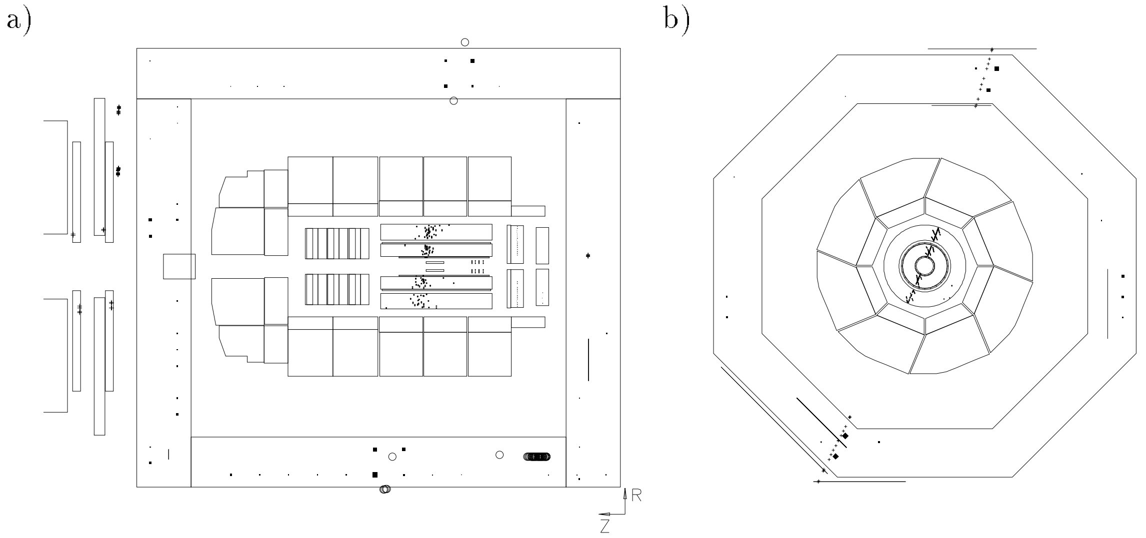


Abbildung 4.3: *Detektorsignatur eines kosmischen Myons, welches durch Subtrigger s34 getriggert wurde. Die Zeichnung zeigt a) einen vertikalen Längsschnitt und b) einen Querschnitt des H1-Detektors.*

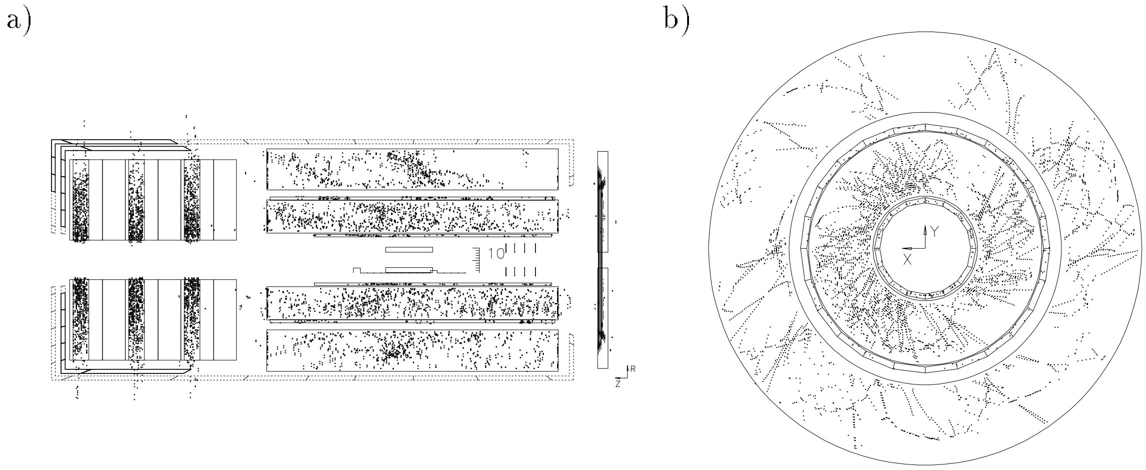


Abbildung 4.4: *Detektorsignatur eines Vielspuruntergrundereignisses, welches durch Subtrigger s34 getriggert wurde. Die Zeichnung zeigt a) einen vertikalen Längsschnitt und b) einen Querschnitt des Spurkammersystems. Charakteristisch für solche Ereignisse ist, daß viele Spuren nahezu parallel zum Strahlrohr liegen. Eine solche Ereignistopologie verursacht nur wenige Einträge im z-Vertex-Histogramm. Dies ist in s34 gefordert. In a) ist das zum Ereignis gehörige z-Vertex-Histogramm abgebildet.*

4.2.2 Netzentwicklung

Im folgenden werden die verschiedenen Trainingsdurchläufe Neuronaler Netze für Subtrigger s34 aus Gründen der Übersichtlichkeit in chronologischer Reihenfolge der Entwicklung dargestellt.
Die Trainings erfolgen nach dem in Abschnitt 3.2.6 dargestellten Ablauf durch Training mit einer Trainingsmenge und Testen der Netze mit einer von dieser statistisch unabhängigen Testmenge. Der Schwerpunkt wird darauf gelegt, daß die trainierten Netze gute Generalisierungseigenschaften zeigen.
Hier nicht dargestellt aber untersucht wurde die Invarianz der Eigenschaften der trainierten Netze bei der Wahl einer anderen Initialisierung des Zufallsgenerators. Die dargestellten Ergebnisse sollten daher nicht von der speziellen Sortierung der Ereignisse abhängen.

Training s34/A

Zum Training wurden 1627 Ereignisse verwendet, davon entstammten 1026 der Untergrund- und 601 der Physikselektion (siehe Tabelle 4.1). Die Anzahl der Epochen belief sich auf 500. Die Korrektur der Gewichte erfolgt nach jedem dem lernenden Netz präsentierten Ereignis.

	Trainingsereignisse	Testereignisse
Untergrund	1026	160
Physik	601	123

Tabelle 4.1: *Anzahlen der benutzten Trainings- und Testereignisse für Training s34/A.*

Die folgenden Größen aus dem Ereignisvektor (siehe Abschnitt 2.4.2) wurden zum Training benutzt: *CPVSUM*, *CPVMAX*, *CPVPOS*, *TRHINEG* , *TRHIPOS*, *TRLONEG*, *TRLOPOS*, *TRTOT*, *NBIGFWD*, *NBIGFCE*, *NBIGBCE* und *NBIGBWD*.
Die Eingangsgrößen dieser Ereignisse sind in Abbildung 4.5 nach Untergrundereignisklasse und Physikereignisklasse getrennt dargestellt. Die auffälligste, statistisch signifikante Abweichung, die festgestellt werden kann, liegt in der Eingabegröße *CPVPOS* vor. Dies kann dadurch erklärt werden, daß bei den als Physik selektierten Ereignissen der Vertex der Reaktion am nominellen Wechselwirkungspunkt liegt. Der Vertex des Untergrundes ist nahezu gleichverteilt. In den Größen *TRLONEG*, *TRLOPOS*, *TRTOT* und *NBIGFCE* erkennt man, daß der Untergrund mehr Ereignisse mit vielen Spuren enthält, als die Physikereignismenge.
In keiner der dargestellten Eingabegrößen ist ein Unterschied zwischen Untergrund und Physik erkennbar, der sich eignen würde einen Schnitt zur Trennung anzubringen. Die beiden Datensätze können sich höchstens durch höherdimensionale Korrelationen voneinander unterscheiden, welche durch das Training des Neuronalen Netzes gefunden werden sollen.

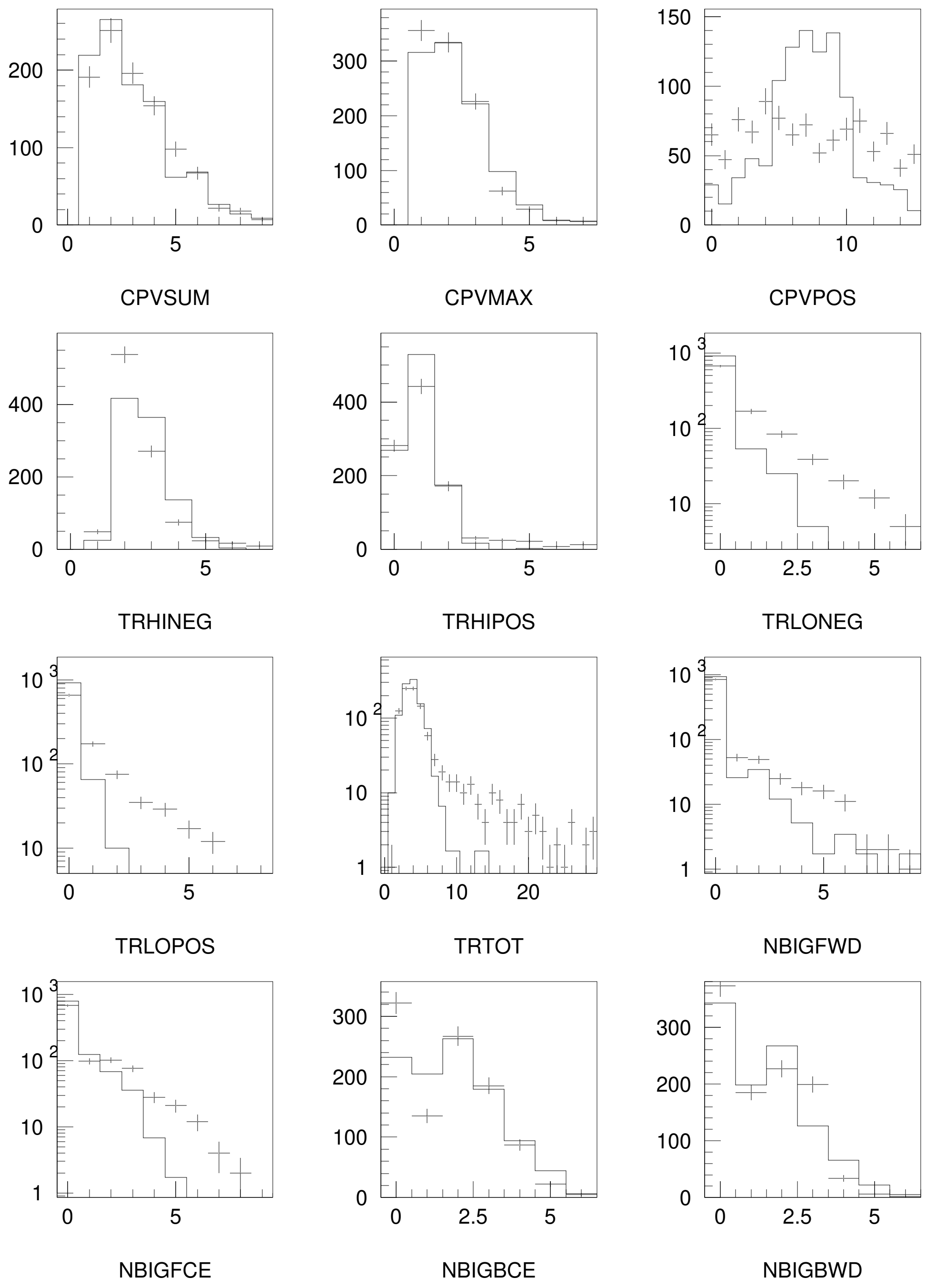


Abbildung 4.5: *Verteilungen der Eingabegrößen der zum Training s34/A benutzten Ereignisse der Untergrund- und Physikereignisklasse. Die Einträge der Physikereignisse wurden in der Summe auf den Untergrund normiert und sind durch die durchgezogenen Linien dargestellt. Zum Vergleich sind die Untergrundverteilungen mit statistischen Fehlern durch Punkte in den gleichen Diagrammen eingetragen.*

Parameter-konfiguration	1	2	3	4	5	6	7	8
Momentumfaktor	0.0	0.3	0.6	0.9	0.0	0.3	0.6	0.9
Lernrate zu Trainingsbeginn	0.001	0.001	0.001	0.001	0.1	0.1	0.1	0.1
Abnahme der Lernrate pro Trainingsepoche	0.0	0.0	0.0	0.0	0.0002	0.0002	0.0002	0.0002
Lernrate zum Trainingsende	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Tabelle 4.2: *Tabelle zur Erklärung der beim Training benutzten Parameterkonfigurationen.*

Das Training wird über einen großen Parameterraum durchgeführt, um einen Überblick über die Stabilität der Trainingsergebnisse bei Variation dieser Parameter zu erhalten. Die Parameterkonfigurationen, die sich hinter den Ziffern 1 bis 8 verbergen, sind in Tabelle 4.2 dargestellt. Es werden Anzahlen an Knoten in der versteckten Schicht von 1-10, 20, 40 und 64 gewählt (siehe Tabelle 4.3).

Knotenindex	1 - 10	11	12	13
zugehörige Anzahl versteckter Knoten	1 - 10	20	40	64

Tabelle 4.3: *Tabelle zur Erklärung des Knotenindex.*

Bei 64 Knoten in der versteckten Schicht haben die Netze $12 \cdot 64 + 64 = 832$ Gewichte, die während des Lernprozesses variiert werden. Diese Anzahl ist halb so groß wie die Anzahl der Trainingsereignisse (siehe Tabelle 4.1). Es ist zu überprüfen, wie sich die Anzahl der Gewichte zur Anzahl der Trainingsereignisse verhält, da bei mehr Gewichten als Trainingsereignissen dem Netz die Möglichkeit eröffnet wird, die Trainingsmenge exakt zu beschreiben. Dies ist hier nicht der Fall. Faustregel ist bei der Art des durchgeführten Netztrainings, daß man etwa 10 mal so viele Trainingsereignisse zum Training benutzt als das Netz Gewichte enthält. Es wird hier auch mit 64 Knoten in der versteckten Schicht trainiert, um die Ergebnisse bis hin zur technischen Grenze der bei H1 installierten Netze zu beobachten.

Die Untergrunderkennung der so trainierten Netze bei einer richtigen Klassifikation von 90% der zum Training verwendeten Physikereignisse ist in Abbildung 4.6 a) dargestellt. Für kleine Knotenanzahlen in der versteckten Knotenschicht erhält man Werte um 45%. Die Werte erhöhen sich mit zunehmender Anzahl an versteckten Knoten und damit mit zunehmender Anzahl an dem Netz zur Verfügung stehenden Netzgewichten, die eingestellt werden können. Bei konstanter, kleiner Knotenanzahl findet mit Benutzung eines größeren Momentumfaktors und einer absteigenden Lernrate eine leichte Verbesserung der Untergrunderkennung über der Trainingsereignismenge statt.

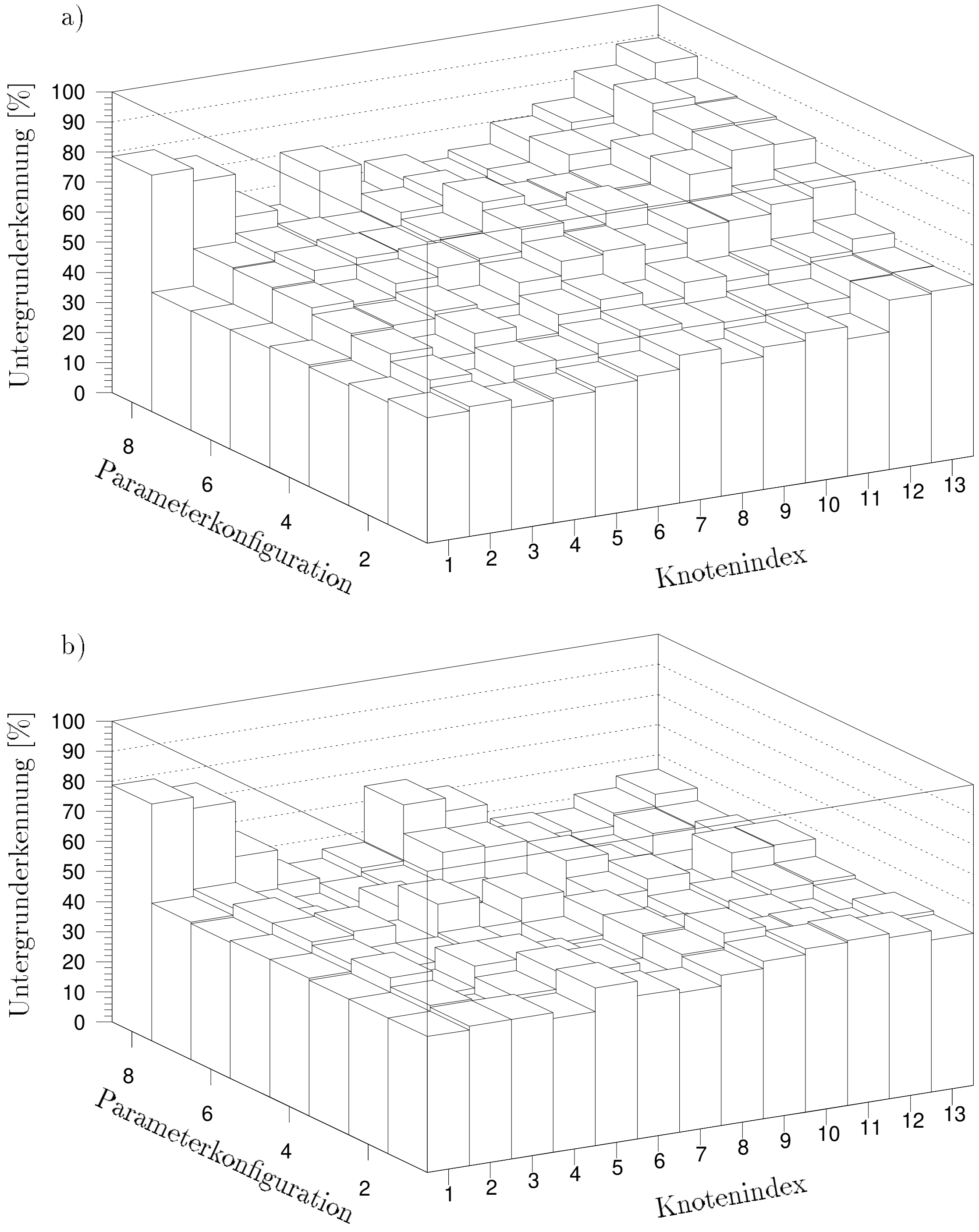


Abbildung 4.6: *a) Trainingsergebnisse und b) Testerbegnisse des Trainings s34/A. Dargestellt ist der Anteil der Untergrundereignisse, den das bei der jeweiligen Parameterkonfiguration und Knotenanzahl trainierte Neuronale Netz als Untergrundereignis klassifiziert, wobei* 90% *der Physikereignisse richtig identifiziert worden sind. Errechnet wurden die Werte bei a) mit 1026 Untergrund- und 601 Physikereignissen, bei b) mit 160 Untergrund- und 123 Physikereignissen. Die exakten Zahlenwerte können in Anhang B (Abbildung B.1 a) und b)) eingesehen werden. (Für die Erklärung von Parameterkonfiguration und Knotenindex siehe die Tabellen 4.2 und 4.3.)*

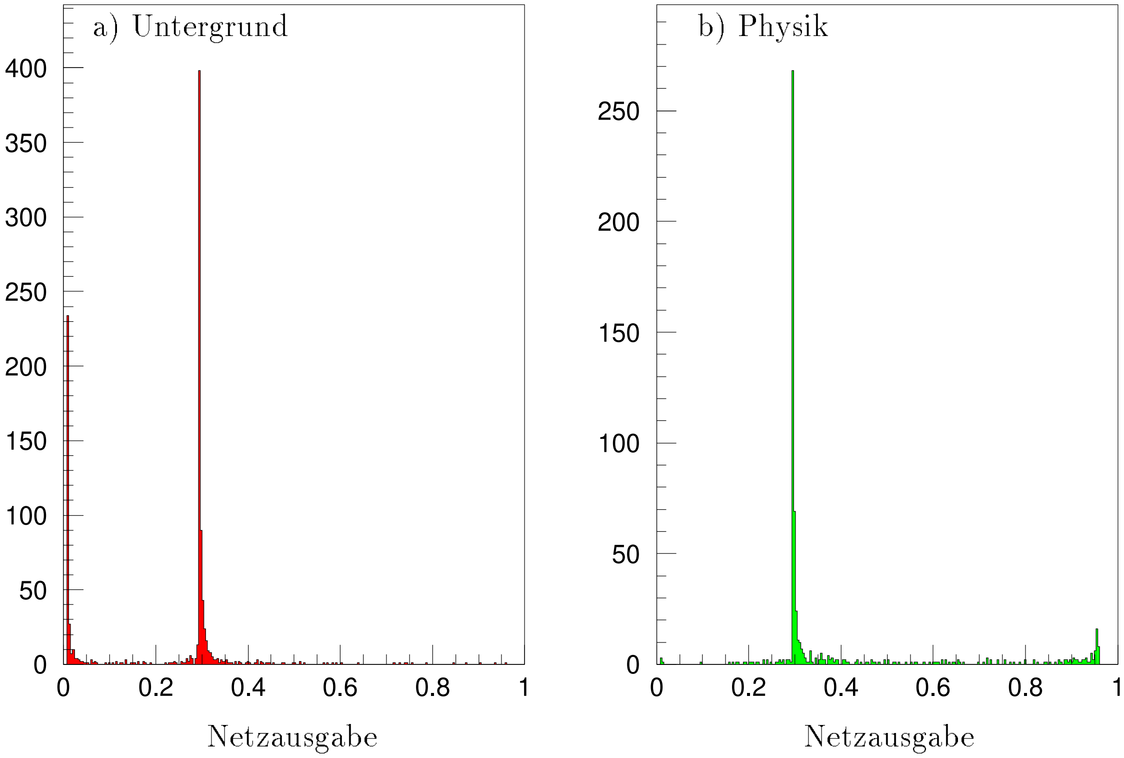


Abbildung 4.7: *Beispiel für ein unbrauchbares Netz (siehe Text). Ein Großteil der Physik- und Untergrundereignisse wird in einem Intervall mit einer Breite kleiner als 0.004 klassifiziert. Ein solches Intervall entspricht ungefähr der Auflösung der Elektronik. Aus Gründen der Übersichtlichkeit ist die für diese Darstellung gewählte Intervallbreite jedoch größer. Dieses Netz stammt aus Training s34/A und ist mit 2 Knoten in der versteckten Schicht, absteigender Lernrate und einem Momentumfaktor von 0.9 trainiert worden.*

Auffällig ist weiterhin, daß für nur einen Knoten in der versteckten Schicht ähnlich gute Ergebnisse wie bei mehr als einem, aber immer noch wenigen Knoten erzielt werden. Bei einem Knoten in der versteckten Schicht findet die Trennung allein durch eine einzige elfdimensionale Ebene im zwölfdimensionalen Raum der Eingabegrößen statt.
Abbildung 4.6 b) stellt in der gleichen Weise wie Abbildung 4.6 a) die Untergrunderkennung nur für Testereignisse dar. Die Testereignisse werden als nahezu statistisch unabhängig von den Trainingsereignissen angenommen (siehe auch Abschnitt 3.2.6). Der bei den Trainingsereignissen in Abbildung 4.6 a) beobachtete Anstieg zu hohen Werten der Parameterkonfiguration und vielen Knoten in der versteckten Schicht der Netze fehlt hier. Die Netze haben hier mit Hilfe der großen Anzahl an einzustellenden Parametern spezielle Eigenschaften der Trainingsereignismenge *erlernt*, welche der Testereignismenge nicht zu eigen sind. Solche Netze nennt man übertrainiert.
Die herausragend hohen Werte in den beiden Abbildungen 4.6 a) und 4.6 b) kommen dadurch zustande, daß die Ausgabewerte der Netze eines großen Anteils der

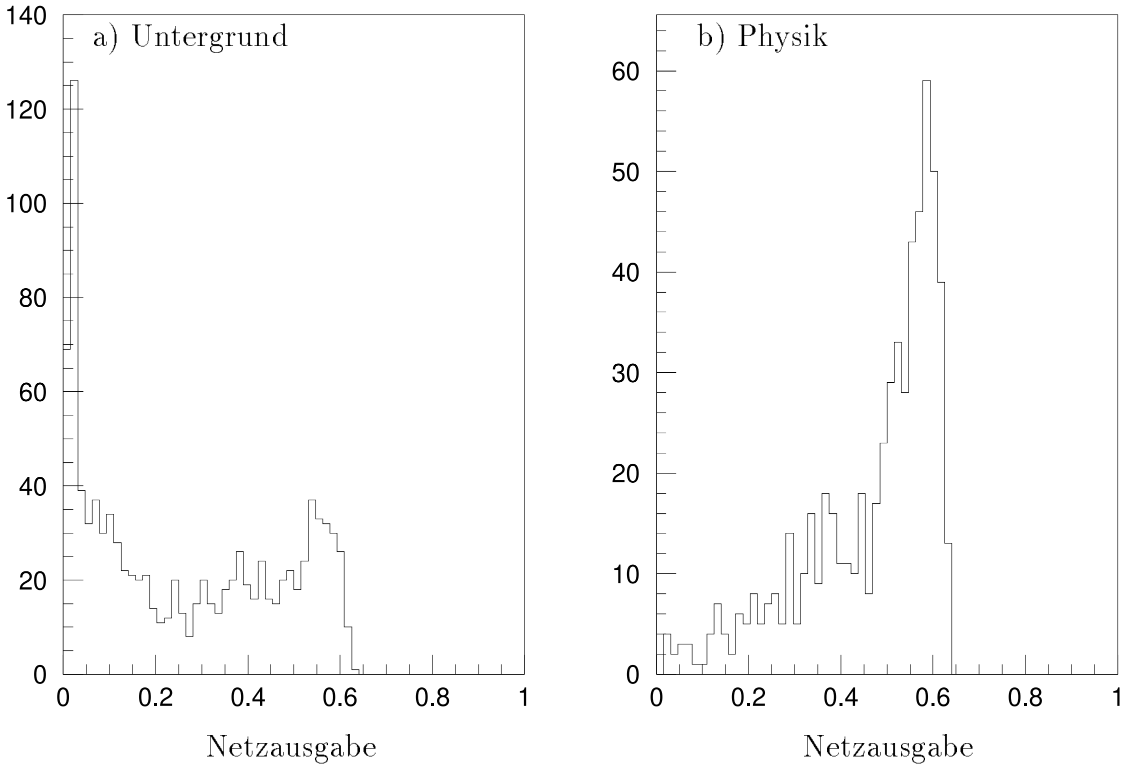


Abbildung 4.8: *Darstellung der Netzausgabewerte für die zum Training verwendeten a) Untergrund- und b) Physikereignisse für das Netz mit 2 Knoten in der versteckten Schicht, absteigender Lernrate und einem Momentumfaktor von 0.6. Dies entspricht der in der Tabelle 4.2 definierten Parameterkonfiguration von 7.*

Untergrund- und Physikereignisse beide in ein und dasselbe Intervall bei 0.3 mit einer Breite von weniger als 0.004 trainiert worden sind. Dieser Sachverhalt wird anhand eines Beispielnetzes mit zwei versteckten Knoten und der Parameterkonfiguration 8 in Abbildung 4.7 verdeutlicht. Der Netzausgabewert innerhalb der neuronalen Elektronik am Experiment wird durch eine 8-*bit* Zahl dargestellt, so daß bei Netzen, die sich wie in Abbildung 4.7 dargestellt verhalten, am Experiment keine Trennung des Netzausgabewertes für Untergrund- und Physikereignisse in Form eines Schnittwertes würde erfolgen können. Mit einer 8-*bit* Zahl können zwischen 0 und 1 256 Intervalle unterschieden werden und es ist $1/256 \approx 0.004$. Solche Artefakte des Trainings sind aufzuspüren und die dazugehörigen Netze von der weiteren Untersuchung und späteren Netzauswahl auszuschließen. In diesem Training tritt dieses Artefakt nur für absteigende Lernrate und hohen Momentumfaktor auf.

Das Netz mit der Parameterkonfiguration 7 und 2 Knoten in der versteckten Schicht wurde ausgewählt, um zu verdeutlichen, welchen Einfluß die auf Seite 45 f. dargestellten Eigenschaften des Ereignisuntergrundes von Subtrigger s34 auf die Netzentscheidung nehmen. Dieses Netz ist als exemplarisch für die Netze mit kleiner Knotenanzahl in der versteckten Schicht und einer Parameterkonfiguration zwischen 1 und 7 anzu-

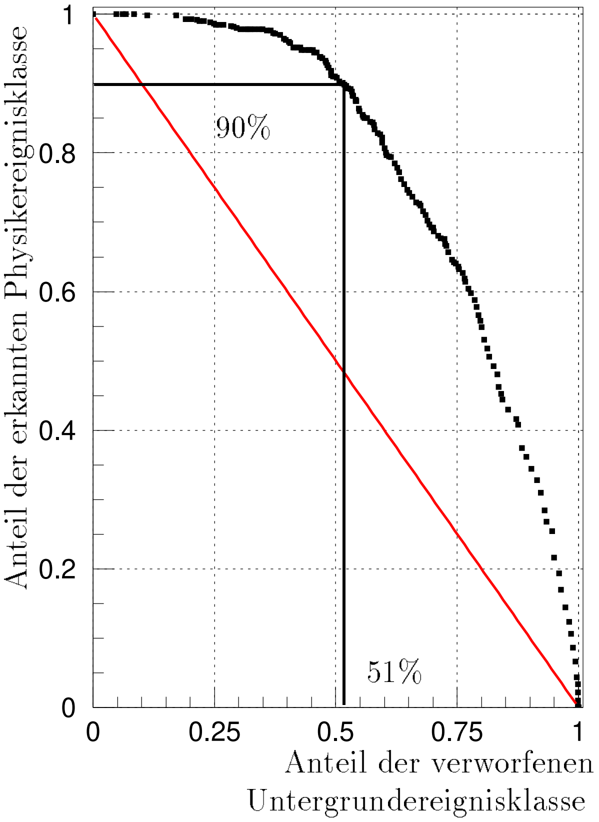


Abbildung 4.9: *Diese Darstellung enthält die zu einem Schnittwert, der auf der x-Achse der beiden Abbildungen 4.8 a) und b) zur Triggerdefinition von Untergrund und Physik gewählt werden muß, gehörenden Werte für getriggerte Physik und zugehörigem verworfenen Untergrund, beide Werte anteilig an der jeweiligen geprüften Gesamtanzahl an Ereignissen. Die Ereignisse der beiden Klassen werden über dem Netzausgabewert bei der Physikereignisklasse von 1 an und bei der Untergrundereignisklasse von 0 an bis zu diesem Wert integriert, zur jeweiligen Gesamtzahl relativiert und das Ergebnis gegeneinander aufgetragen. Dieser Schnittwert stellt den Parameter dieser Kurve dar.*

sehen. Der Ausgabewert dieses Netzes ist in Abbildung 4.8 dargestellt. Die Netzausgabewerte sowohl für die Ereignisse der Untergrund- als auch der Physikereignisklasse streuen über einen großen Bereich, und ein Teil der Untergrundereignisse wird in derselben Region wie die Physikereignisse klassifiziert. Dies deutet darauf hin, daß dieser Teil der Untergrundereignisklasse den Ereignissen der Physikereignisklasse in bezug auf die verwendeten Eingabegrößen ähnlich ist.

In der letzten Abbildung zu Training s34/A (Abbildung 4.9) wird die Untergrunderkennung und die dazu gehörige Erkennung der Ereignisse der Physikereignisklasse für dasselbe Netz wie in Abbildung 4.8 dargestellt. Der Parameter der Kurve aus Abbildung 4.9 ist der zwischen 0 und 1 gewählte Schnittwert, der als Unterscheidungsgrenze zwischen zu triggernden und nicht zu triggernden Ereignissen dienen soll. Die zu einem Schnittwert gehörenden Werte für Untergrunderkennung und Erkennung der Ereignisse der Physikereignisklasse wurden bestimmt, indem beim Untergrund (Abbildung 4.8 a)) von 0 bis zum Schnittwert und bei der Physik (Abbildung 4.8 b)) vom Schnittwert bis zum Wert 1 integriert wird. Diese Integrationsergebnisse werden auf die jeweilige Gesamtzahl von Untergrund- und Physikereignissen normiert und sind in Abbildung 4.9 gegeneinander aufgetragen. Man findet bei einem Anteil von 90% Erkennung der Ereignisse der Physikklasse einen Wert von 51% an erkanntem Untergrundanteil. Dieser Wert kann auch in den Abbildungen 4.6 a) und B.1 a) jeweils beim Knotenindex 2 und der Parameterkonfiguration 7 wiedergefunden werden.

Trennung des Untergrundes von Subtrigger s34 in kosmische Myonen und Ereignisse mit vielen Spuren

Zur weiteren Untersuchung der im voherigen Abschnitt erhaltenen Überlappung der Netzausgabegrößen von Untergrund und Physik (siehe Abbildung 4.8), werden hier die kosmischen Myonen aus dem Untergrund heraussortiert, da diese möglicherweise die Ursache sind.

	Ereignisse mit Signatur eines kosmischen Myons in Durchmusterung	Ereignisse ohne Signatur eines kosmischen Myons in Durchmusterung
Kriterium (DCOS,DTNV) für ein kosmisches Myon erfüllt	424 (99.5%)	18 (4.4%)
Kriterium (DCOS,DTNV) für ein kosmisches Myon nicht erfüllt	2 (0.5%)	387 (95.6%)
Summe	426	405

Tabelle 4.4: *Übereinstimmung des gewählten Kriteriums für kosmische Myonen (DCOS-Bank in den Daten und weniger als 30 Spuren in der DTNV-Bank) mit der Trennung durch visuelle Durchmusterung im H1-Ereignismonitorprogramm.*

Das **Kriterium** zur Trennung der kosmischen Myonen von den übrigen Untergrundereignissen ist die Anwesenheit der DCOS-Bank in den Daten und weniger als 30 Spuren in der DTNV-Bank. In der DTNV-Bank sind die Spuren eines Ereignisses vor deren Anpassung an vorhandene Vertizes gespeichert. Die DCOS-Bank wird für ein Ereignis angelegt, wenn zwei Spuren rekonstruiert werden, die mindestens eine radiale Länge von 8 cm und eine Krümmung von $\kappa < 0.015$ besitzen und nicht bei beiden Spuren der Spurrekonstruktionsstartpunkt vom nominellen Vertex aus gesehen radial mehr als 40 cm vom nominellen Vertex entfernt ist. Die angegebene, maximale Krümmung der Spuren entspricht einem minimalen Transversalimpuls der Spurteilchen von $P_t = 240\,MeV$. Weiterhin darf die Summe der minimalen Abstände der beiden extrapolierten Spuren von der nominellen Strahlachse nicht mehr als 2 cm betragen. Die entscheidende Bedingung, daß eine DCOS-Bank angelegt wird, ist das Ergebnis eines χ^2-Tests, der mit der Krümmung κ, dem Winkel ϕ und dem Abstand nächster Näherung der extrapolierten Spur zur nominellen Strahlachse jeweils beider Teilchenspuren durchgeführt wird. Das beschriebene Selektionskriterium wurde mit visueller Durchmusterung von 831 Ereignissen im H1-Ereignismonitorprogramm überprüft, in dem bei den Ereignissen nach

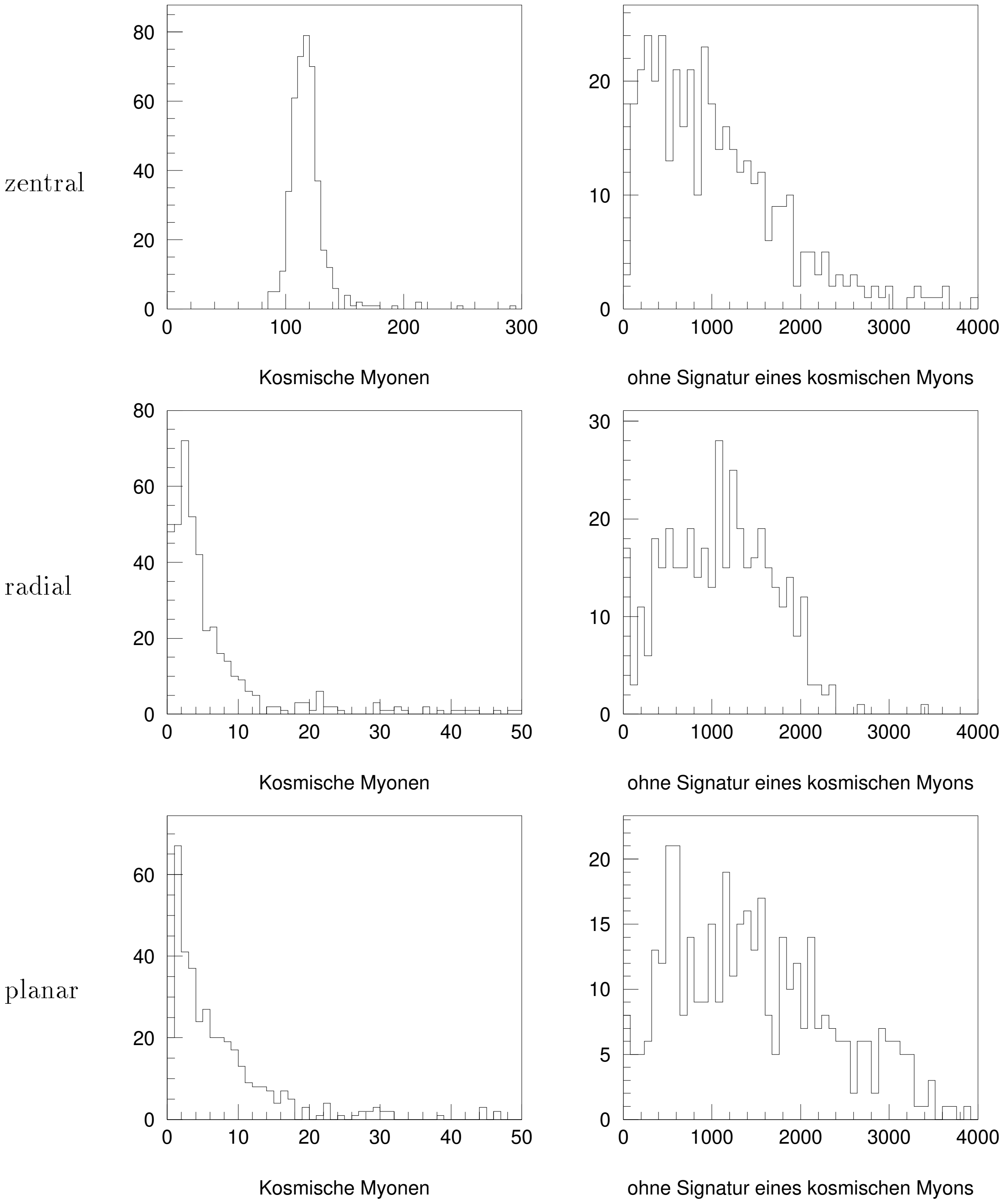


Abbildung 4.10: *Diese Abbildung enthält die Anzahl der signaltragenden Spurkammerdrähte in den zentralen und den vorderen radialen und planaren Spurkammern für 426 selektierte kosmischen Myonen und 405 Ereignisse ohne Signatur eines kosmischen Myons (siehe Tabelle 4.4).*

einer typischen **Signatur** eines kosmischen Myons gesucht wurde. Die Übereinstimmung der beiden beschriebenen Bedingungen an ein kosmisches Myon ist in Tabelle 4.4 dargestellt. Der Unterschied der beiden Klassifizierungen liegt unterhalb von 5%. Daher wird in dieser Arbeit für die Identifikation von kosmischen Myonen das oben beschriebene Kriterium aus den Banken DCOS und DTNV verwendet.
Wie in Tabelle 4.4 dargestellt, weisen von 831 Untergrundereignissen 426 die Signatur eines kosmischen Myons auf, 405 Ereignisse nicht. Für diese Ereignisse ist die Anzahl der getroffenen Spurkammerdrähte in den zentralen und vorderen radialen und planaren Spurkammern (siehe Abschnitt 2.2.2) in Abbildung 4.10 gezeigt. Man erkennt, daß die Ereignisse ohne Signatur eines kosmischen Myons deutlich mehr Drähte ansprechen lassen. Diese sind die speicheraufwendigen Ereignisse aus dem Untergrund, die es hauptsächlich durch ein Netz zu filtern gilt (siehe hierzu auch Tabelle 2.1 und die zugehörigen Ausführungen in Abschnitt 2.2.2).
Aus der gesamten Anzahl der vorhandenen Untergrundereignisse von 2828, wurden mit Hilfe des hier beschriebenen Kriteriums 1593 kosmische Myonen selektiert, so daß sich der Anteil an kosmischen Myonen am verwendeten Untergrund von Subtrigger s34 zu 56% ergibt.

Training s34/B

Nach Abtrennung der kosmischen Myonen aus dem Untergrund können die bestehenden Netze erneut mit den gleichen Einstellungen der Trainingsprogramme wie für Training s34/A trainiert werden. Nur ist diesmal eine getrennte Darstellung der Netzausgabe für kosmische Myonen und Ereignisse ohne Signatur kosmischer Myonen möglich. Dies ist in Abbildung 4.11 exemplarisch für ein Netz gezeigt. Die Gesamtheit der Ereignisse der Physikereignismenge streut genau wie die kosmischen Myonen über einen weiten Bereich, da das Netz mit Hilfe der Informationen der Eingabegrößen nicht Physikereignisse von kosmischen Myonen trennen kann. Im Gegensatz dazu ist die Netzausgabe für die Ereignisse ohne Signatur eines kosmischen Myons näher beim Ausgabewert 0.

Es muß hier festgestellt werden, daß mit den verwendeten Eingabegrößen die Ereignisse der Physikereignisklasse nicht von den kosmischen Myonen getrennt werden können. Eine Verbindung der in vier θ-Intervallen unterteilten *Bigrays* reicht insbesondere zusammen mit der Position des z-Vertex-Maximums *CPVPOS* zur Unterscheidung von kosmischen Myonen und den leptonischen Zerfällen diffraktiv erzeugter J/ψ-Mesonen nicht aus.

Training s34/C

In diesem Training wird auf Grund der vorangegangenen Ergebnisse nur mit Ereignissen der Untergrundereignisklasse trainiert, die kein kosmisches Myon enthalten. Es standen so 1235 Untergrundereignisse zur Verfügung. Die Physikereignisse entstammen derselben Selektion wie in den beiden vorangegangenen Trainings. In Tabelle 4.5

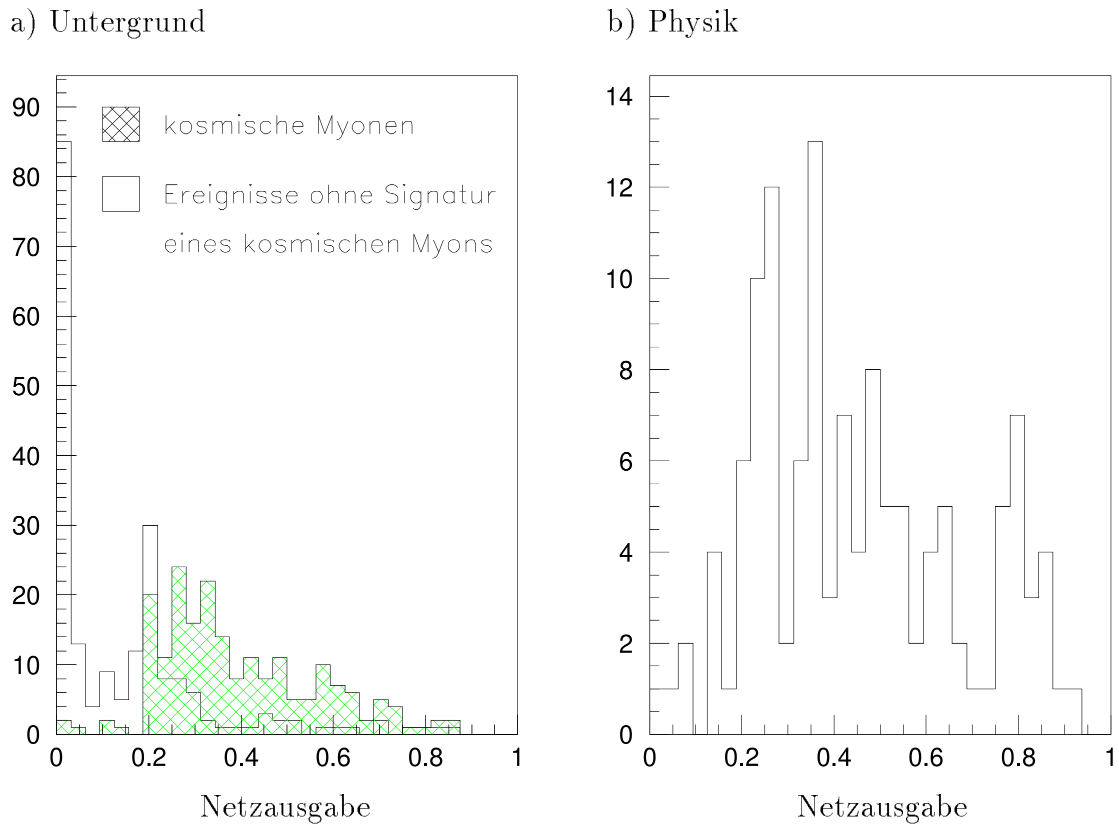


Abbildung 4.11: *Darstellung der Netzausgabewerte der a) Testuntergrundereignismenge, getrennt nach kosmischen Myonen und Ereignissen ohne Signatur eines kosmischen Myons, und der b) Testphysikereignismenge exemplarisch für ein Netz des Trainings s34/B.*

sind die Anzahlen der für Training und Test benutzten Trainings- und Testereignisse dargestellt. Die Physikereignisse stammen aus der Selektion von Seite 44 f..

	Trainingsereignisse	Testereignisse
Untergrund	1065	170
Physik	583	351

Tabelle 4.5: *Anzahlen der benutzten Trainings- und Testereignisse für Training s34/C.*

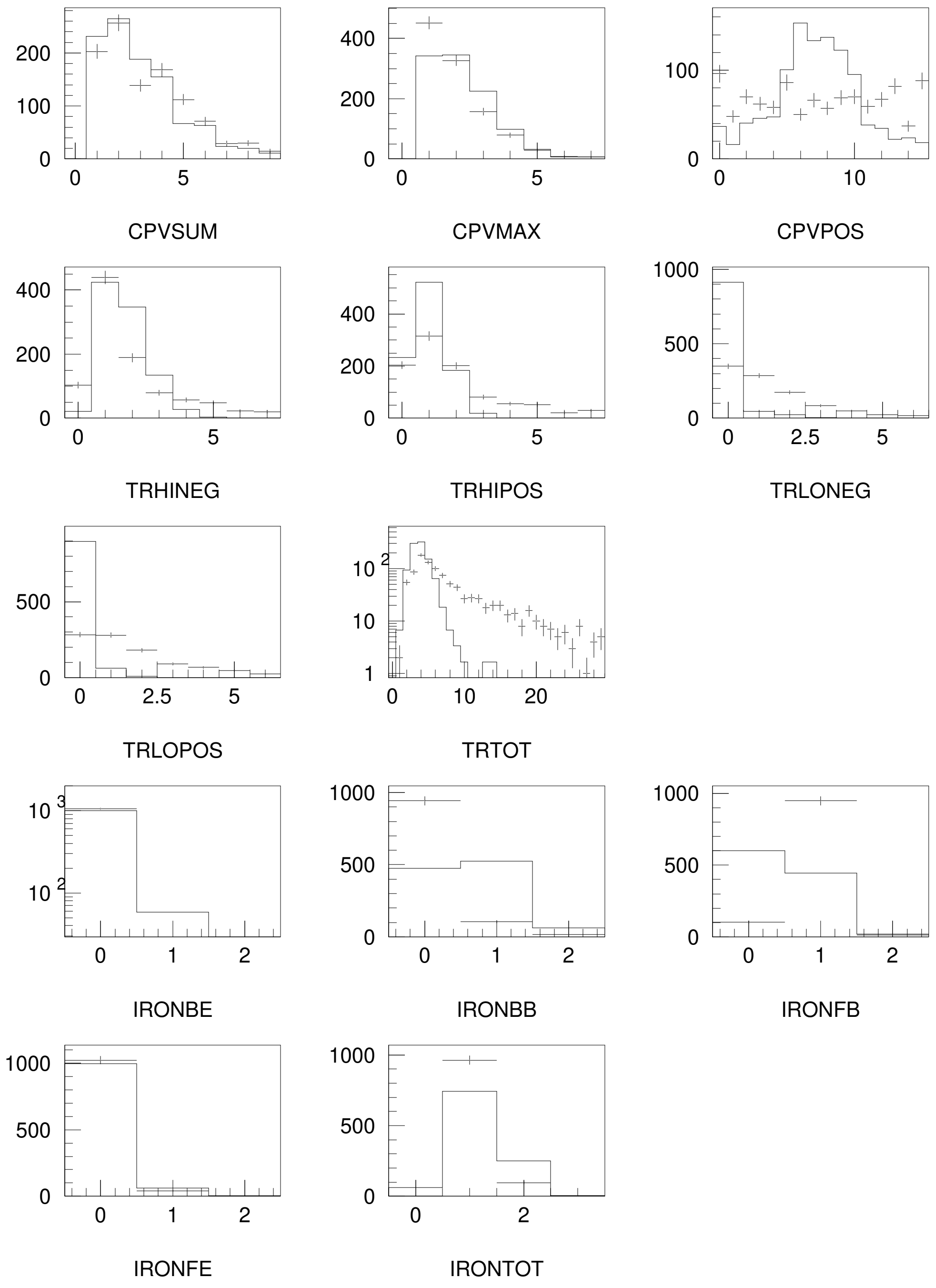


Abbildung 4.12: *Verteilungen der Eingabegrößen der zum Training s34/C benutzten Ereignisse aus der Untergrund- und der Physikereignisklasse. Die Einträge der Physikereignisse wurden in der Summe auf den Untergrund normiert und sind durch die durchgezogenen Linien dargestellt. Zum Vergleich sind die Untergrundverteilungen mit statistischen Fehlern als Punkte eingetragen. (***Die kosmischen Myonen sind aus dem hier dargestellten Untergrund entfernt***).*

Die diesem Training zugrunde liegende Strategie geht dahin, die durch viele Spuren charakterisierten Ereignisse der Untergrundereignisklasse auf Triggerstufe L2 durch ein Neuronales Netz zu verwerfen, damit bei diesen langen Ereignissen nicht der Detektor weiter ausgelesen werden muß und die Daten elektronisch bis zu Stufe L4 transferiert werden müssen. Die in der Datenmenge für ein Ereignis deutlich weniger umfangreichen kosmischen Myonen können dann durch Triggerstufe L4 verworfen werden.

Größen aus	z-Vertex Driftkammer Proprotionalkammer	z-Vertex Driftkammer Eisen	z-Vertex Driftkammer Eisen Proportionalkammer
Untergrunderkennung bei 90% Physikerkennung (Mittelwerte aus den jeweils besten vier Netzen)	(86.3 ± 0.5)%	(92.8 ± 0.5)%	(90.3 ± 0.5)%

Tabelle 4.6: *Testergebnisse der Netze bei Kombinationen von Eingangsgrößen aus verschiedenen Detektorteilen für das Training s34/C.*

Es wurde mit verschiedenen Kombinationen an Eingabegrößen aus den wichtigsten Detektorteilen trainiert. Die Ergebnisse sind in Tabelle 4.6 dargestellt. Die Ergebnisse dieser Untersuchungen waren Grundlage der Entscheidung, das Training mit Spurkammergrößen und Eisengrößen fortzusetzen. Bezüglich der Netzparameterkonfigurationen und der Anzahl der Knoten in der versteckten Schicht wurden darauffolgend die gleichen systematischen Untersuchungen angestellt wie in Training s34/A.

Analog Abbildung 4.5 zu Training s34/A sind die Eingabegrößen der Physik- und der Untergrundereignismenge für Training s34/C in Abbildung 4.12 dargestellt. Im Vergleich zu Abbildung 4.5 sieht man in Abbildung 4.12 Abweichungen der Anzahl von Spuren mit niedriegen transversalen Impulsen $400\,MeV \leq P_t \leq 800\,MeV$, *TRLONEG* und *TRLOPOS*, also sowohl für Spuren von positiv wie auch negativ geladenen Teilchen. Auch ist in den Größen *IRONBB* und *IRONFB* zu erkennen, daß die Detektorsignatur von Ereignissen der Untergrundereignismenge im Eisendetektor im Gegensatz zu Ereignissen der Physikereignismenge durch einen größeren Anteil an Spuren in Vorwärtsrichtung charakterisiert ist.

Analog den Abbildungen 4.6 a) und 4.6 b) aus Training s34/A ist in den Abbildungen 4.13 a) und 4.13 b) für dieses Training die Untergrundunterdrückung bei 90% vom jeweiligen Netz erkannter Physik dargestellt. Die Untergrunderkennung ist sehr viel besser als mit dem ursprünglichen Untergrund von s34, der in Training s34/A benutzt wurde. Dort war sie ca. 45%. Dies bestätigt die Vermutung, daß die Untergrundereignisse ohne Signatur eines kosmischen Myons sehr viel besser von der Physik zu trennen

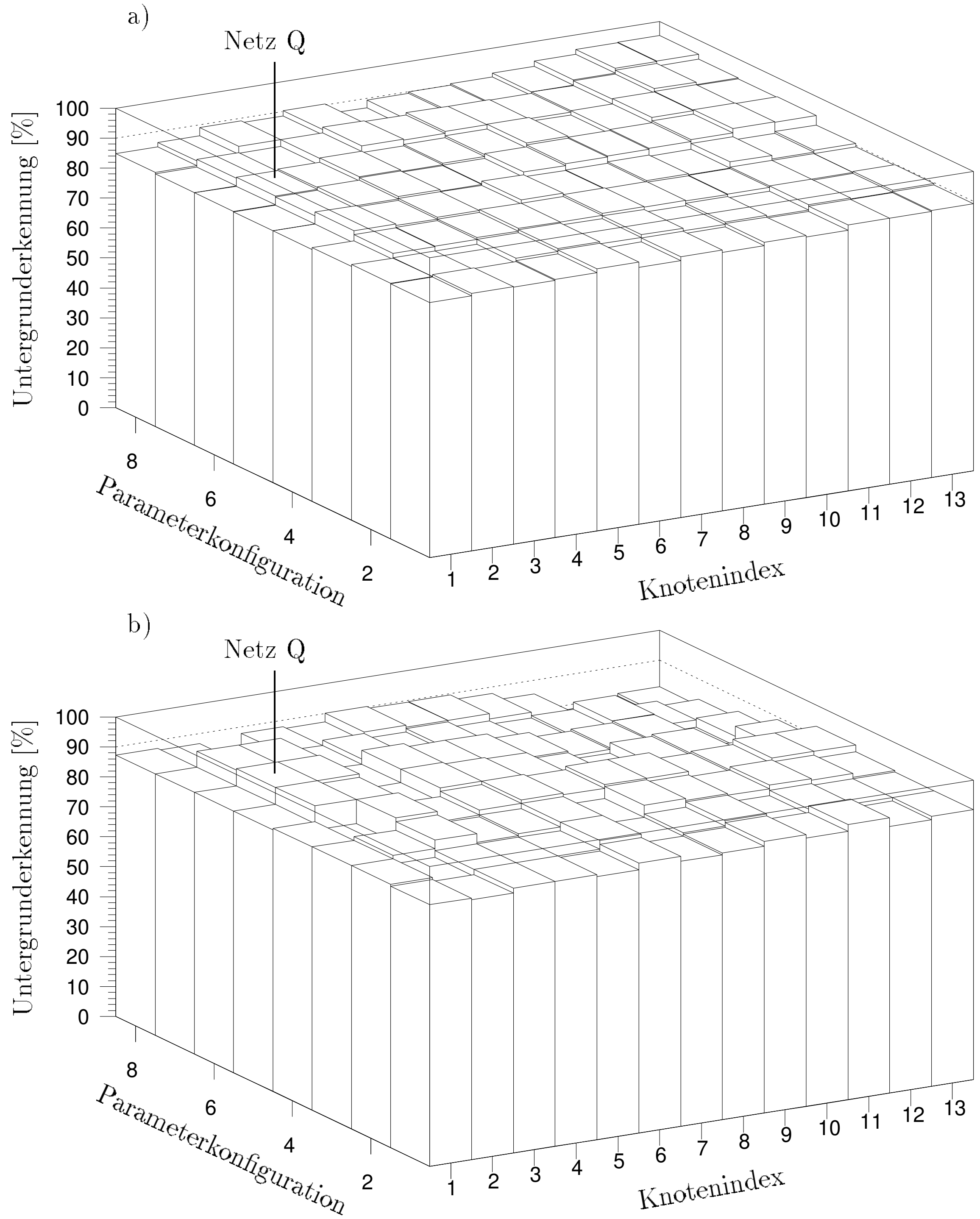


Abbildung 4.13: *a) Trainingsergebnisse und b) Testergebnisse des Trainings s34/C. Dargestellt ist der Anteil der Untergrundergeignisse, den das bei der jeweiligen Parameterkonfiguration und Knotenanzahl trainierte Neuronale Netz als Untergrundereignis klassifiziert, wobei* 90% *der Physikereignisse richtig identifiziert worden sind. Errechnet wurden die Werte bei a) mit 1065 Untergrund- und 583 Physikereignissen und bei b) mit 170 Untergrund- und 351 Physikereignissen. Die exakten Zahlenwerte können in Anhang B Abbildung B.2 eingesehen werden. Eingezeichnet ist weiterhin das letztendlich ausgewählte Netz Q. (Für die Erklärung von Parameterkonfiguration und Knotenindex siehe die Tabellen 4.2 und 4.3.)*

sind als die kosmischen Myonen.
Weiterhin ist wieder der Effekt zu erkennen, daß die Netzergebnisse bei kleinen Anzahlen an Knoten in der versteckten Schicht bis auf statistische Schwankungen unabhängig sowohl von der Parameterkonfiguration wie auch vom Knotenindex sind. Bei hohen Anzahlen von Knoten in der versteckten Schicht und bei großen Werten der Parameterkonfiguration ist wiederum der Effekt des Übertrainierens, also das *Erlernen* von speziellen Eigenschaften der Trainingsereignismenge durch das Netz ersichtlich. Ereignisse der Trainingsereignismenge werden zunehmend besser, solche der Testereignismenge zunehmend schlechter erkannt. Die im Gegensatz dazu für Training und Test nahezu gleichen Ergebnisse der Netze im Bereich kleiner Anzahlen an versteckten Knoten lassen den Schluß zu, daß in diesem Bereich ein Übertrainieren der Netze nicht stattgefunden hat.
Das endgültige Netz der Wahl stammt daher aus diesem Bereich und hat die Parameterkonfiguration 6, also einen Momentumfaktor in Höhe von 0.3 bei 2 versteckten Knoten. Dieses Netz wird im folgenden unter seinem Arbeitsnamen Netz Q untersucht werden.

4.2.3 Untersuchungen an Netz Q

In diesem Abschnitt werden die Eigenschaften von Netz Q untersucht. Insbesondere wird auf die Güte der Trennung von Untergrund- und Physikereignissen mit Hilfe neu selektierter Untergrundereignisse eingegangen.

Bestimmung des Schnittwertes

In den Abbildungen 4.14 sind die Netzausgabewerte für a) Untergrundereignisse und b) Physikereignisse getrennt dargestellt. Ähnliche Verteilungen ergeben sich für die Testereignisse in Abbildung 4.15, was noch einmal bestätigt, daß mit den verwendeten Parametern kein Übertrainieren stattgefunden hat.
Zur Quantifizierung der Ergebnisse wird ein Schnittwert aus dem Intervall [0;1] bestimmt. Unterhalb dieses Wertes wird ein Ereignis als Untergrundereignis, oberhalb dieses Wertes als Physikereignis klassifiziert. Bei der Forderung eines Anteils von 95% richtig klassifizierter Physikereignisse ergibt sich aus den Netzausgabewerten für die Testereignismenge (Abbildung 4.15 b)) ein Schnittwert von 0.2. Für die Abbildungen 4.14 und 4.15 ist der Anteil der Ereignisse mit einem Netzausgabewert oberhalb des Schnittwertes angegeben.
Die Bestimmung des Schnittwertes kann auch anhand der Abbildung 4.16 veranschaulicht werden. Hier wird der Schnittwert zwischen 0 und 1 variiert und dabei jeweils der Anteil an richtig erkannten Physikereignissen und Untergrundereignissen ermittelt. Dies geschieht durch Integration der Einträge der Verteilungen aus den Abbildungen 4.14 und 4.15. Die Ergebnisse sind hier in den Abbildungen 4.16 a) und b) dargestellt.

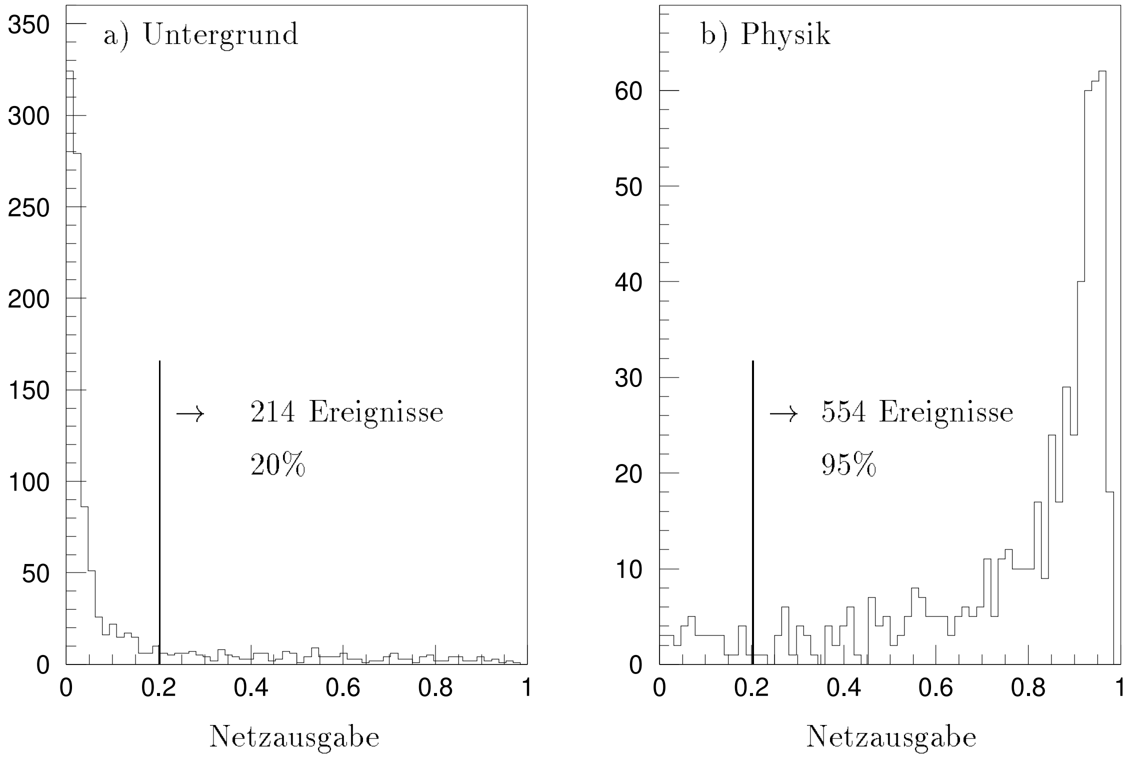


Abbildung 4.14: *Darstellung der Netzausgabewerte für a) Untergrund- und b) Physikereignisse für das Netz Q. Dargestellt sind die Netzausgabewerte für die 1065 Untergrund- und die 583 Physikereignisse der* **Trainings***ereignismenge. Der hier dargestellte Trainingsuntergrund enhält keine kosmischen Myonen.*

Diese Berechnung wurde schon für Abbildung 4.9 durchgeführt und erklärt.
Das Netz kann erst mit Festlegung des Schnittwertes als Trigger verwendet werden. Der hier ermittelte Schnittwert von 0.2 wird in den folgenden Ausführungen im Zusammenhang mit Netz Q benutzt werden.

Untersuchung der Relevanz der einzelnen Eingabegrößen am endgültigen Netz Q

Die Relevanz einer Eingabegröße bezüglich eines bestimmten Neuronalen Netzes ist eine Größe, die angibt, wie wichtig die einzelnen Eingabegrößen bei der Entscheidungsfindung relativ zueinander sind. Die Relevanz wird für jede Eingabegröße gesondert berechnet und ist definiert durch

$$R = \frac{1}{N} \sum_{i=1}^{N} (\bar{o}_i - o_i)^2. \tag{4.2}$$

Dabei stellt o_i die gewöhnliche Netzausgabe für Ereignis i dar, welches aus der zur Relevanzbestimmung benutzten Menge der Testereignisse stammt. Die Größe $\bar{o}_i$ ist die

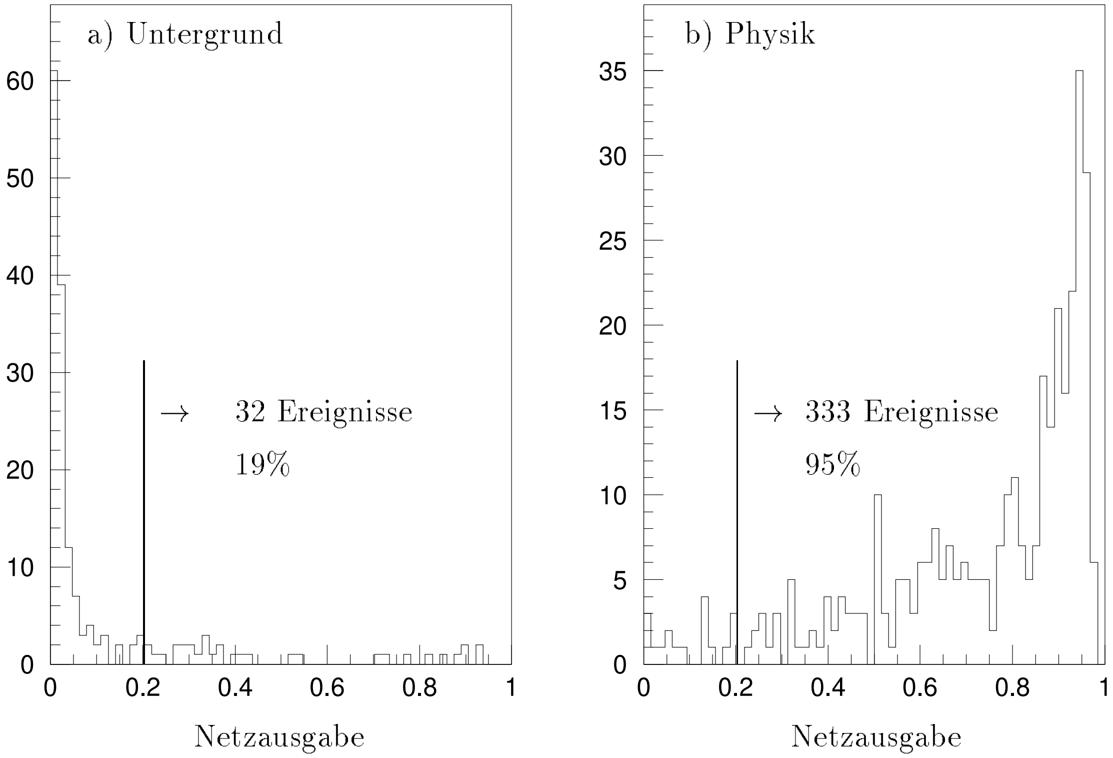


Abbildung 4.15: *Darstellung der Netzausgabewerte für a) Untergrund- und b) Physikereignisse für das Netz Q. Dargestellt sind die Netzausgabewerte für die 170 Untergrund- und die 351 Physikereignisse der* **Test***ereignismenge. Der hier dargestellte Testuntergrund enthält keine kosmischen Myonen.*

Netzausgabe bei Festhalten der betrachteten Eingabegröße auf ihrem Mittelwert aus der Testereignismenge. Beim Vergleich der Relevanzen verschiedener Eingabegrößen ist darauf zu achten, daß diese quadratische Größen darstellen. Schon alleine deswegen kann doppelte Relevanz nicht bedeuten, daß eine Größe doppelt so wichtig bei der Entscheidungsfindung eines Netzes ist, wie eine andere Größe.

In Abbildung 4.17 sind die mit der Testereignismenge ermittelten Relevanzen der Eingabegrößen für Netz Q dargestellt. Man erkennt in dieser Abbildung als wichtigste Größen die drei z-Vertex-Größen *CPVSUM*, *CPVMAX* und *CPVPOS*, die Driftkammergrößen für Spuren mit niedrigem Impuls *TRLONEG* und *TRLOPOS* und die Eisengröße *IRONFB*.

Die hier aufgezeigten Relevanzen sind in Verbindung mit den Eingabegrößenverteilungen aus Abbildung 4.12 zu betrachten. Obwohl hochdimensionale Korrelationen zwischen den Eingabegrößen, auch bei identischen Verteilungen der einzelnen Größen für Untergrund- und Physikereignismenge, zu einer guten Unterscheidung der Klassen durch ein Neuronales Netz führen können, sind hier die meisten Eingabegrößen mit hoher Relevanz auch durch starke Unterschiede in den Verteilungen für Untergrund- und Physikereignismenge charakterisiert.

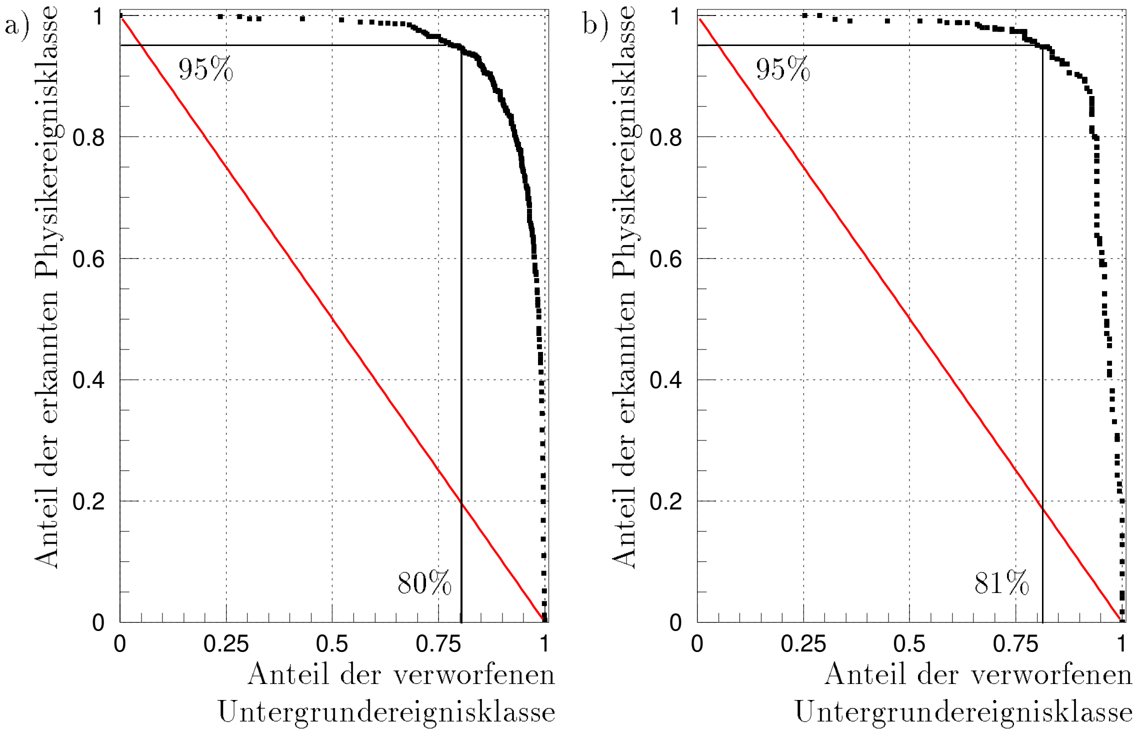


Abbildung 4.16: *Darstellung des Anteils richtig klassifizierter Physikereignisse gegen den Anteil verworfener Untergrundereignisse für a)* **Trainings***ereignisse und b)* **Test***ereignisse mit Variation des Schnittwertes.*

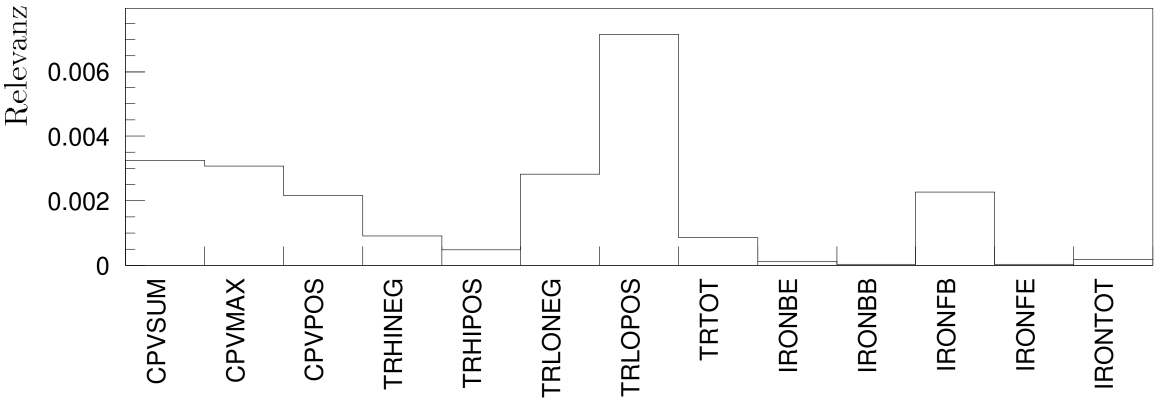


Abbildung 4.17: *In dieser Darstellung findet man für jede Netzeingabegröße getrennt die Relevanz bezüglich Netz Q aufgetragen. Die Werte wurden mit Hilfe der Testereignismenge, die 521 Ereignisse umfaßt, ermittelt.*

Auf der anderen Seite würde man auf Grund der Verteilung von *IRONBB* in Abbildung 4.12 erwarten, daß diese Größe mehr Einfluß auf die Netzentscheidung haben sollte, als sich dies in der Relevanz ausdrückt.

Tests mit kosmischen Myonen und neuer Untergrundselektion

In Abbildung 4.18 a) sind die Netzausgabewerte von Netz Q für 426 kosmische Myonen (nach Kriterium auf Seite 54 *ff.*) dargestellt. Es werden bei einem Schnittwert von 0.2 90% der kosmischen Myonen von Netz Q als zu triggernde Ereignisse klassifiziert. Dies ist annähernd soviel, wie der Anteil der getriggerten Physik im Test, der bei einem Schnittwert von 0.2 95% beträgt (siehe Abbildung 4.15).

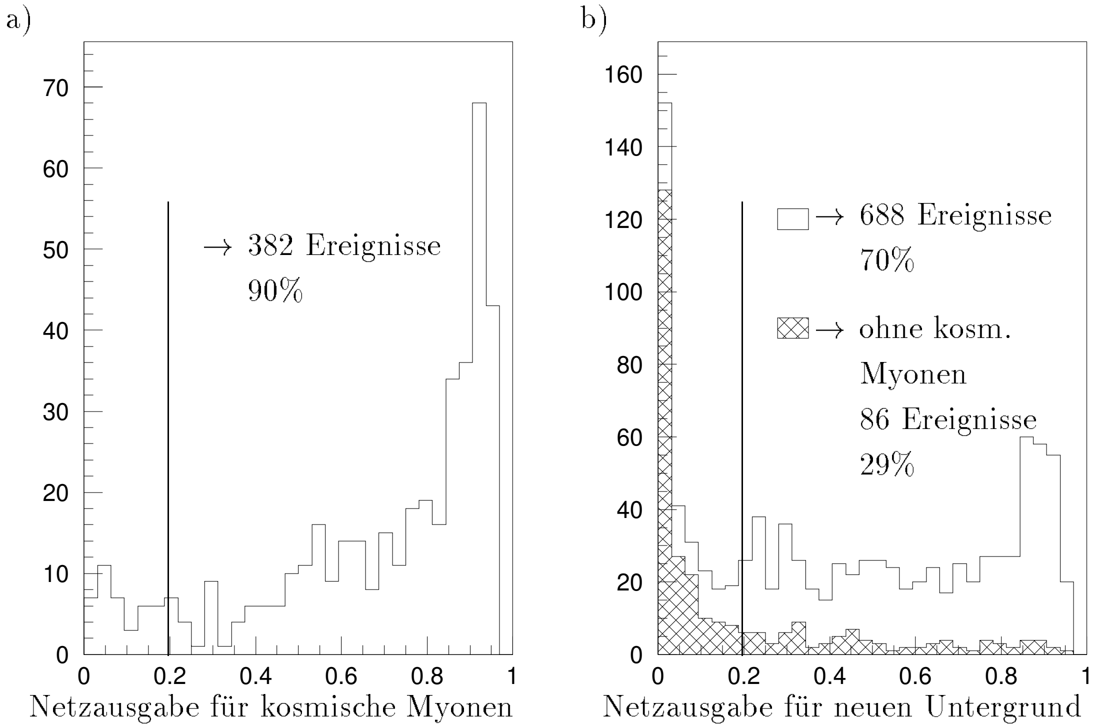


Abbildung 4.18: *Diese Darstellung enthält die Netzausgabewerte von Netz Q für a) 426 kosmische Myonen a), die aus der ursprünglichen Untergrundereignismenge stammen, und b) 982 Untergrundereignisse, die am 27.Juli 1997 bei H1 aufgezeichnet wurden.* 90% *der Myonen und* 70% *des Untergrundes werden von Netz Q mit einem Schnittwert bei 0.2 akzeptiert. In Abbildung b) ist weiterhin kariert dargestellt, daß* 29% *der Untergrundereignisse ohne Signatur eines kosmischen Myons von Netz Q falsch klassifiziert werden. Die Anzahl dieser Ereignisse beträgt insgesamt 293. Dies ist gemessen an der Anzahl aller dargestellten 982 Ereignisse ein Anteil von* 30%*.*

Weiterhin wurde das Netz an anderen als den bisher verwendeten Daten[1] getestet, um zu untersuchen, wie das Neuronale Netz auf anderen Untergrund reagiert. Diese Untergrundereignisse wurden auf gleiche Weise selektiert wie die Untergrundereignisse, die zum Training benutzt wurden. Das Ergebnis ist in 4.18 b) zu sehen. Es handelt sich dabei um 982 Ereignisse, die auf die gleiche Weise selektiert wurden wie die Daten der Untergrundereignismenge für das Netztraining. Die kosmischen Myonen wurden dabei nicht verworfen. Der Anteil an verworfenen Ereignissen der Untergrundereignisklasse von 30% geht in Tabelle 4.7 zur L4-Triggersimulation ein.
Die Tatsache, daß dieser Wert deutlich niedriger ist, als auf Grund der Ergebnisse aus Training s34/A zu erwarten war (siehe Abbildung 4.9), liegt daran, daß der neue Untergrund von 982 Ereignissen mehr kosmische Myonen enthält, nämlich 70%, gegenüber einem Anteil von 56% im Trainingsuntergrund. Eine Erhöhung des Anteils kosmischer Myonen in der Testuntergrundereignismenge führt daher zu einer Verminderung des Anteils der Untergrundereignisse, die vom Netz richtig klassifiziert werden.

Ereignisgröße vor und nach Netz Q

Da die Ereignisse, die zum Training von Netz Q benutzt wurden, keine kosmischen Myonen enthielten, werden vom Netz, wie erwartet, hauptsächlich die Ereignisse mit vielen Einträgen in den zentralen und vorderen (radialen und planaren) Spurkammern

---

[1] aufgezeichnet am 27.Juli 1997

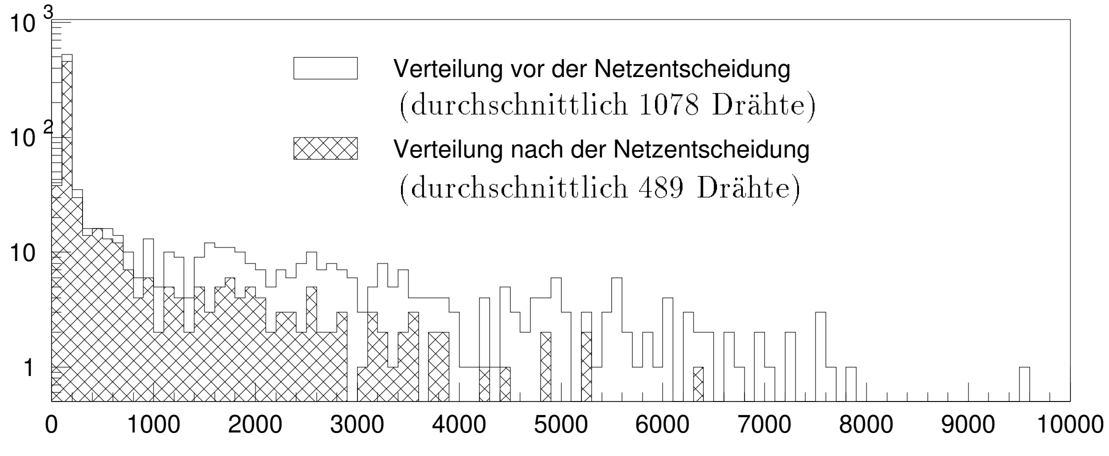


Abbildung 4.19: *Anzahl der angesprochenen Spurkammerdrähte in der Summe der zentralen und der vorderen Spurkammern (radial und planar) für die 1235 für Training s34/C verwendeten Untergrundereignisse.*

als Untergrund identifiziert. Dies ist in Abbildung 4.19 dargestellt. Die durchschnittliche Anzahl angesprochener Drähte der Ereignisse vor Netz Q halbiert sich verglichen mit den Ereignissen, die von Netz Q akzeptiert wurden.

4.2.4 Simulation triggerrelevanter Größen

Um die Triggerrate von Subtrigger s34 nach Triggerstufe L4 bei Einbau des Netzes abzuschätzen, wurde Triggerstufe L4 für die Ereignisse ohne und mit positiver Netzentscheidung simuliert.
Es wurden 521 Ereignisse aus Triggerphase 2 und 982 Ereignisse aus Triggerphase 3 benutzt. In Tabelle 4.7 sind die Anzahlen derjenigen Ereignisse aus dieser Testmenge dargestellt, die die Triggerstufen L2 und L4 passieren.

	Ereignisanzahl vor L2	Ereignisanzahl nach L2	Ereignisanzahl nach L4 () ohne Netz auf L2
Phase 2	521	347	88(106)
Phase 3	982	688	166(219)
		Ratenreduktion durch L2	Ratenreduktion durch L4 () ohne Netz auf L2
Phase 2		0.67	0.25(0.20)
Phase 3		0.70	0.24(0.22)
	Rate vor L2 [Hz] (Werte aus Abb. 4.1)	Rate nach L2 [Hz]	Rate nach L4 [Hz] () ohne Netz auf L2
Phase 2	1.2 – 1.5	0.8 – 1.0	0.20 – 0.25(0.24 – 0.3)
Phase 3	1.1 – 1.3	0.8 – 0.9	0.18 – 0.22(0.24 – 0.29)

Tabelle 4.7: *Ergebnisse der Simulation eines Subtriggers s34 mit Netz Q auf Triggerstufe L2 und der nachfolgenden Triggerstufe L4 sowie der Vergleich zur Simulation ohne Netz Q auf Triggerstufe L2 (Siehe Werte in Klammern ()). Die Raten vor L2 sind Abbildung* 4.1 *entnommen.*

Aus Tabelle 4.7 wird ersichtlich, daß sich die Ratenreduktionen in den Triggerphasen 2 und 3 in dieser Simulation nicht unterscheiden. In beiden Phasen wird durch Multiplikation der Ratenreduktionen der einzelnen Triggerstufen L2 und L4 aus Tabelle 4.7 eine Gesamtratenreduktion der Rate nach L1 auf 17% festgestellt. Daraus folgt eine Gesamtrate nach Triggerstufe L4 phasenabhängig zwischen 0.20 und 0.25 für Phase 2 und 0.17 und 0.22 für Phase 3. Zum Vergleich befinden sich in der Tabelle auch die Ergebnisse einer L4-Triggersimulation ohne Netz Q. Diese Werte für die Rate nach Triggerstufe L4 liegen hier nur leicht höher. Der Vorteil der Installation von Netz Q auf Triggerstufe L2 ist daher im Bereich der Totzeitverringerung durch Verminderung der auszulesenden Ereignisse auf niederer Triggerstufe zu sehen. Zudem sind die durch das Netz verworfenen Ereignisse durch viele Spurkammereinträge charakterisiert. Man spart eine Detektorauslese gerade für diese speicheraufwendigen Ereignisse wie in Abbildung 4.19 gezeigt.

L4-Bedingungen Phase 2			
Ereignis akzeptiert	Ereignisse Anzahl [%]	Ereignis verworfen	Ereignisse Anzahl [%]
Hardscale:		*Triggerverifikation:*	
L4_SC1_PTMAX1	62(70%)	*L4_MCOSMIC*	178(69%)
L4_SC1_PTMAXMU1	24(27%)	*L4_MUON_NOMU*	27(10%)
andere Bedingungen	2(2%)	*andere Bedingungen*	54(21%)
	88		259

Tabelle 4.8: *Die L4-Bedingungen in* **Phase 2**, *die zum Verwurf oder Erhalt eines Ereignisses auf Triggerstufe L4 geführt haben. Es sind nur Bedingungen mit einem Anteil über* 8% *aufgeführt.*

L4-Bedingungen Phase 3			
Ereignis akzeptiert	Ereignisse Anzahl/Anteil	Ereignis verworfen	Ereignisse Anzahl/Anteil
Hardscale:		*Triggerverifikation:*	
L4_SC1_PTMAX1	112(67%)	*L4_MCOSMIC*	379(73%)
L4_SC1_PTMAXMU1	48(29%)	*L4_MUON_NOMU*	47(9%)
andere Bedingungen	6(4%)	*andere Bedingungen*	96(18%)
	166		522

Tabelle 4.9: *Die L4-Bedingungen in* **Phase 3**, *die zum Verwurf oder Erhalt eines Ereignisses auf Triggerstufe L4 geführt haben. Es sind nur Bedingungen mit einem Anteil über* 8% *aufgeführt.*

In den Tabellen 4.8 und 4.9 sind die L4-Bedingungen, die zur Verwerfung oder zum Erhalt der Ereignisse auf dieser Triggerstufe geführt haben, anteilig dargestellt. Hervorzuheben ist insbesondere, daß 69% (Phase 2) und 73% (Phase 3) der durch L4 verworfenen Ereignisse durch die Bedingung *L4_MCOSMIC* verworfen werden, was bedeutet, daß diese Ereignisse durch den L4-Trigger als kosmische Myonen erkannt wurden.

$L4\_MCOSMIC$ setzt voraus: weniger als 3 gute Spuren[1] in den Spurkammern, weniger als 11 Eisenspuren, von denen zwei in unterschiedlichen Hemisphären im Azimutalwinkel Φ liegen, drei Streifenlagen im Eisen, die ein Signal liefern, und höchstens eine Abweichung von 10° davon, daß sich die Spuren vom nominellen Vertex aus gesehen in der Variablen Θ gegenüberliegen. Bei einer Entscheidung der Triggerverifikation durch $L4\_MUON\_NOMU$ konnte ein auf der L1-Ebene detektiertes Myon auf Triggerstufe L4 nicht durch eine Spur im Eisen verifiziert werden.
Ein Anteil von 70% Phase 2 bzw. 67% in Phase 3 der Ereignisse, die getriggert wurden, sind durch die *Hardscale*-Entscheidung $L4\_SC1\_PTMAX1$ charakterisiert, was bedeutet, daß auf L4-Triggerebene eine gute Teilchenspur[1] mit einem transversalen Impuls von $P_t > 2\,GeV$ rekonstruiert wurde. Für die Bedingung $L4\_SC1\_PTMAXMU1$ reicht ein transversaler Impuls von $P_t > 1\,GeV$ an einer guten Spur[1], die außerdem zu einer Eisenspur assoziiert werden kann. Dies geschieht, indem man die Spurkammerspur bis in das Eisen extrapoliert und in einem Winkelbereich von $0.2\,rad$ um den Eintrittspunkt dieser Spur ins Eisen nach dem Anfangspunkt einer Eisenspur sucht. Durch $L4\_SC1\_PTMAXMU1$ sind in Phase 2 27% und in Phase 3 29% der Ereignisse charakterisiert.
Die Untersuchungsergebnisse der Tabellen 4.8 und 4.9 bezüglich der Nutzung des trainierten Netzes im H1-Trigger haben für die Triggerphasen 2 und 3 nur geringe Unterschiede ergeben. Die Anteile der durch verschiedene L4-Bedingungen klassifizierten Ereignisse unterscheiden sich unterhalb von 5 Prozentpunkten.

4.2.5 Das hängende bit

In der Elektronik der Subdetektordatenauslese bis hin zur Bereitstellung der Eingabegrößen für die Neuronalen Hardwarenetze tritt in der Elektronik der Fehler auf, daß ein *bit* für die Berechnung der Größe *TRHINEG* nahezu immer den Wert 1 hat. Dieser Fehler tritt seit etwa Mitte 1995 auf.

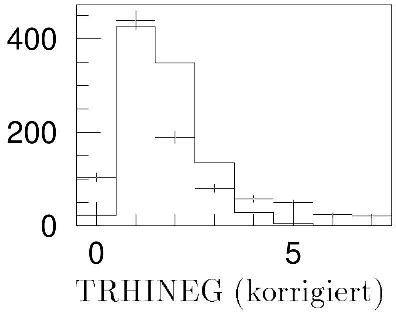


Abbildung 4.20: *Verteilungen von TRHINEG für Untergrund- (Fehlerbalkenkreuze) und Physikereignismenge (Histogramm) nach Abzug von 1. Die Verteilung ist zu vergleichen mit der Verteilung von TRHINEG aus Abbildung 4.12.*

---

[1] Eine gute Spur auf L4 umfaßt Schnitte im Polarwinkel, in der radialen Spurlänge, im radialen Abstand des Spurstartpunktes von der nominellen Strahlachse, in der Anzahl der Punkte zum Rekonstruieren, in der Breite des z-Vertex-Histogramms und im minimalen, radialen Abstand der extrapolierten Spur zur nominellen Strahlachse.

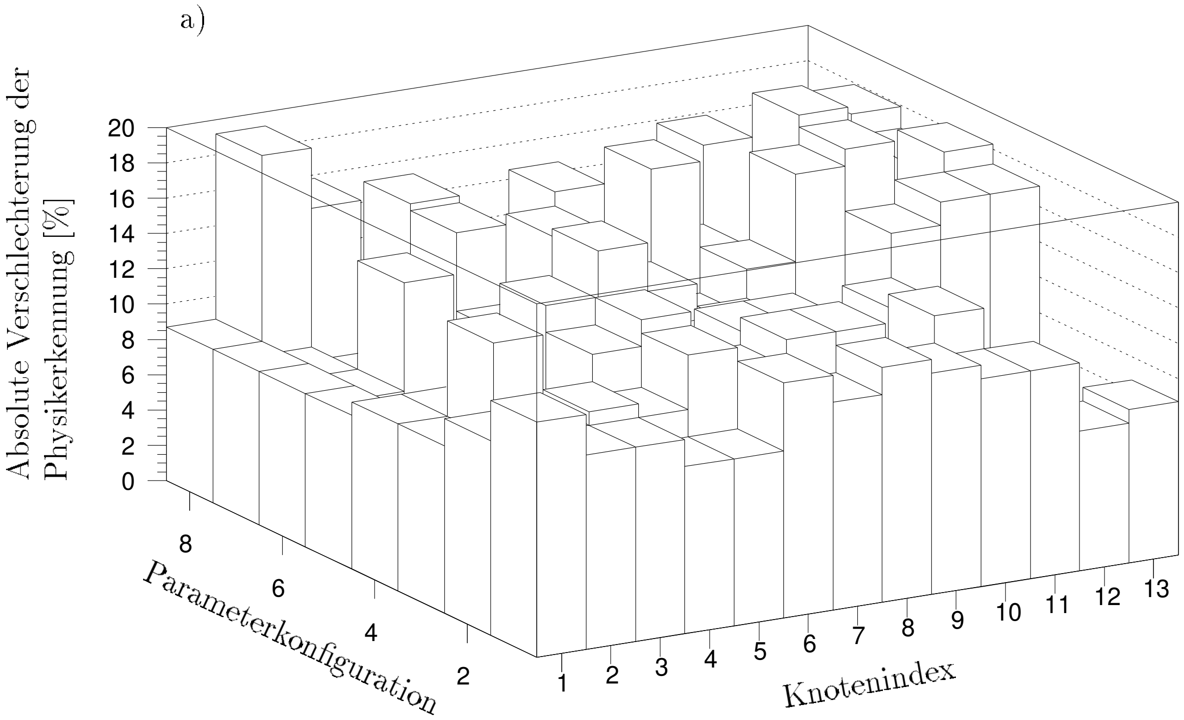
Begründet ist der Fehler darin, daß von den 45 ϕ-Sektoren für *TRHINEG* der Sektor 22 bei jedem Ereignis ein Signal gibt unabhängig davon, ob eine Maske in diesem Sektor angesprochen hat oder nicht. Dies äußert sich in den hier verwendeten Eingabegrößen darin, daß *TRHINEG* um 1 zu groß ist, solange nicht durch mindestens eine Maske aus dem ϕ-Sektor 22 eine Spur eines negativ geladenen Teilchens mit niederem Impuls erkannt wurde. Dann ist nämlich das hängende *bit*, das zu ϕ-Sektor 22 gehört, korrekterweise gesetzt, d.h. für ein solches Ereignis ist richtig, daß dieser Sektor mit 1 zu den Summen angesprochener Sektoren *TRHINEG* und *TRTOT* beiträgt .
Es soll hier simuliert werden, wie Netz Q reagiert, wenn die Elektronik am Experiment wieder richtig arbeitet, das Netz aber mit Daten aus der Periode mit hängendem *bit* trainiert wurde. Dazu wird ein Datensatz erstellt, in dem die Größen *TRHINEG* und damit auch *TRTOT* um 1 erniedrigt sind und das Netz Q über diesen Daten simuliert wird. Man vergleiche das korrigierte *TRHINEG* in Abbildung 4.20 mit *TRHINEG* aus der Abbildung 4.12.

Wert x für TRHINEG	1	2	3	4	5	6	7
Fehler der Simulation x/45 [%]	2.2	4.4	6.7	8.9	11.1	13.3	15.6
TRHINEG-Verteilung für Physik [%] aus Abbildung 4.20	2.2	44.3	36.2	14.1	2.9	0.3	0.0
Fehlerabschätzung der Simulation für Physik : 6.0%							
TRHINEG-Verteilung für Untergrund [%] aus Abbildung 4.20	10.7	45.7	19.7	8.3	5.9	5.1	4.6
Fehlerabschätzung der Simulation für Untergrund : 6.4%							

Tabelle 4.10: *Fehlerberechnung zur Untersuchung des hängenden bits.*

Zuerst wird der Fehler dieses Vorgehens untersucht. Der Anteil an Ereignissen, der auf Grund obiger Argumentation durch diesen Abzug von 1 falsch simuliert ist, wird abgeschätzt zu nahezu 6% für die Physikereignisse und 6.4% für die Untergrundereignisse. Dies ergibt sich wie folgt.
Da es 45 Sektoren für jede Spurklasse gibt (siehe Seite 24), wird angenommen, daß von allen Ereignissen mit $TRHINEG = x$ ein Anteil von $x/45$ durch Verminderung von TRHINEG falsch simuliert wurde. Kompliziertere Anordnungen, wie z.B. das mehrere Spuren innerhalb des Sektors 22 eine Maske auslösen könnten werden hier vernachlässigt.

Zur Gewichtung des so abgeschätzten Anteils an falsch simulierten Ereignissen für verschiedene Multiplizitäten x werden die Verteilungen aus Abbildung 4.20 benutzt. Da die Verteilungen mit hängendem *bit* gemacht wurden, ist dies nicht exakt richtig. Es



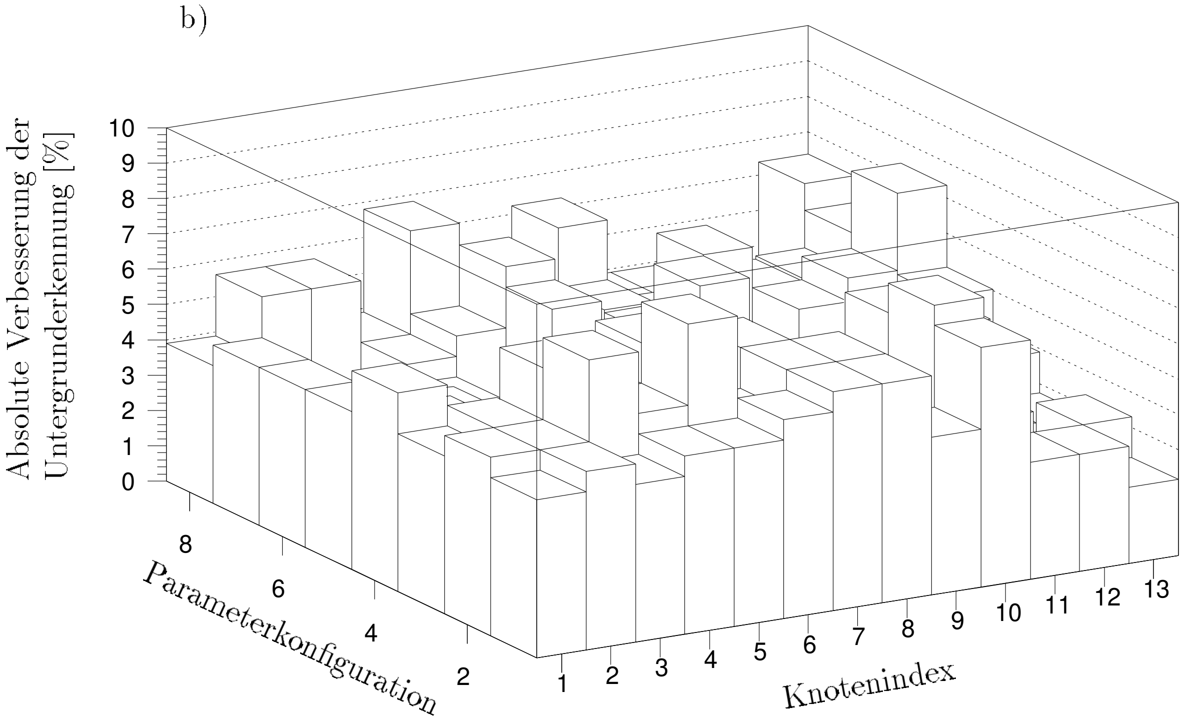


Abbildung 4.21: *a) Verschlechterung der Physikerkennung und b) Verbesserung der Untergrunderkennung, bei Simulation der Situation, daß ein Neuronales Netz, welches mit Daten trainiert wurde, bei deren Aufzeichnung das bit hing, am Experiment installiert ist, in dessen Elektronik das bit nicht mehr hängt. Die Referenz dieser Werte ist die Simulation des Netzes über Testdaten, die wie auch die Trainingsdaten bei hängendem bit aufgezeichnet wurden. Die Zahlenwerte befinden sich in Anhang B Abbildung B.3. (Für die Erklärung von Parameterkonfiguration und Knotenindex siehe die Tabellen 4.2 und 4.3.)*

wird angenommen, daß der Fehler sich durch die Vorgabe dieser Verteilungen nicht stark ändert. Dies bestätigt sich durch die geringe Abweichung der unten stehenden, ermittelten Fehler für Physik und Untergrund, wobei die Verteilungen von Physik und Untergrund aus Abbildung 4.20 deutlich unterschiedlich sind. Die Berechnung ist in Tabelle 4.10 dargestellt.

In Abbildung 4.21 ist zu erkennen, daß sich bei einem am Experiment befindlichen Neuronalen Netz, das mit Daten trainiert wurde, die das hängende *bit* enthalten, die Physikerkennung verschlechert und die Untergrunderkennung verbessert. Vorausgesetzt wurde für die Berechnung der Werte eine Erkennung der Physikereignisse der unkorrigierten Daten von 90%. Für diese Anforderungen wurde der Schnittwert der Netze bestimmt. Mit diesem Schnittwert können dann die Physik- und Untergrunderkennung für die in *TRHINEG* korrigierten Daten ermittelt werden. Die Abweichung dieser Werte von den Werten für die unkorrigierten Daten ist in Abbildung 4.21 dargestellt. Für diese Berechnung wurden vom Training unabhängige Testdaten verwendet. Die Verschlechterung der Physikerkennung beträgt bis zu 19%, gemessen an der gesamten Anzahl an Testereignissen. Beim ausgewählten Netz Q sind es 7%. Beim Anstieg der Untergrunderkennung erhält man bis zu 7% gemessen an der Gesamtzahl an Untergrundereignissen. Für Netz Q steigt diese um 2% an. Diese nicht allzu hohen Werte für das Netz Q reflektieren die relativ geringe Relevanz der Größe *TRHINEG* (siehe hierzu Abbildung 4.17).

An der Schwankungsbreite der Ergebnisse in Abbildung 4.21 ist zu erkennen, daß die Größe *TRHINEG* bei den trainierten Netzen einen sehr unterschiedlichen Einfluß auf die Netzentscheidung hat.

4.3 Subtrigger s52

Der Subtrigger s52 ist wie auch Subtrigger s34 für diffraktiv erzeugte J/ψ- oder Υ-Mesonen konzipiert worden. Wie bei s34 werden zur Sicherstellung der geringen Spurmultiplizität der Ereignisse die Proportionalkammern und der z-Vertex-Trigger verwendet. Weiterhin wird ein Triggerelement aus dem Flüssig-Argon-Kalorimeter benutzt. In der folgenden Triggerelementdarstellung von Subtrigger s52 bedeutet & logisches UND.

Definition des Subtriggers s52 für das Jahr 1997:

- s52: zVtx_Cls & DCRPh_TNeg & LAr_BR & Topo_Mult<3

Die Triggerelemente sind im einzelnen:

zVtx_Cls : Alle Einträge im z-Vertex-Histogramm befinden sich innerhalb von höchstens vier benachbarten Intervallen.

DCRPh_TNeg : Mindestens eine negative Maske des DCrphi-Triggers wurde gesetzt.

LAr_BR : Es gibt mindestens einen *Bigtower* im Flüssig-Argon-Kalorimeter mit Signal

über der Rauschschwelle, der durch einen in dieser Richtung liegenden *Bigray* validiert worden ist.
Topo_Mult<3 : Es sind höchstens 2 *Bigrays* aktiviert.

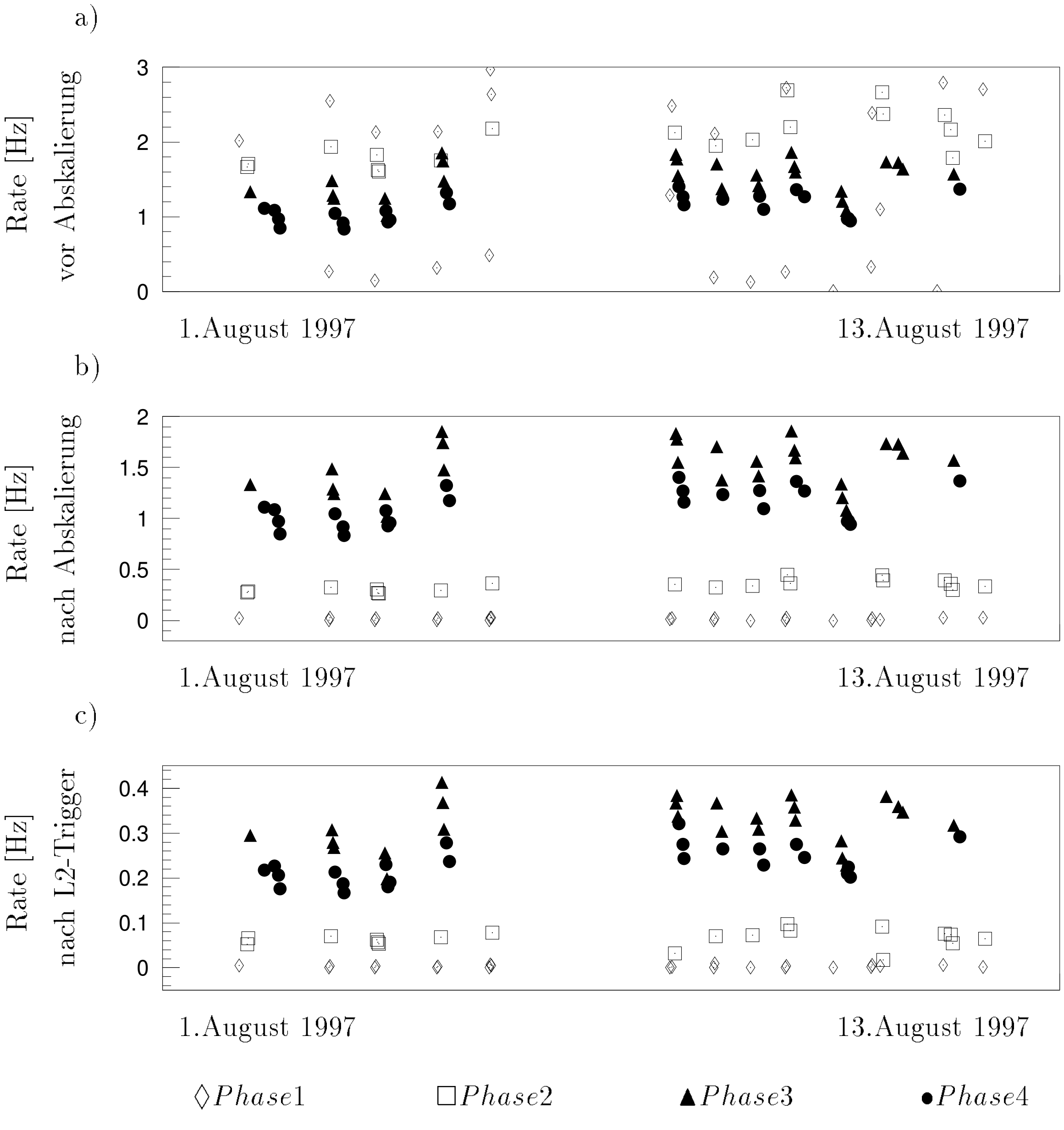


Abbildung 4.22: *Raten des Subtriggers s52 a) ohne, b) mit Abskalierungsfaktor und c) nach positiver L2-Triggerentscheidung. Der Abskalierungsfaktor ist 100 für Phase 1, weshalb die abgestufte Rate in b) in dieser Phase nahezu Null beträgt. Für die Phase 2 ist 6, für die Phasen 3 und 4 ein Abskalierungsfaktor von 1 festgesetzt.*

Durch die Bedingungen zVtx_Cls und Topo_Mult<3 wird erreicht, daß Subtrigger s52 vorwiegend elastische Ereignisse triggert. Eine Rateneinschränkung erfolgt durch LAr_BR als unabhängige Information aus dem Kalorimeter.
Auf Subtrigger s52 ist bereits ein L2-Subtriggerelement installiert. Dieses arbeitet mit den folgenden Größen: *CPVSUM*, *CPVPOS*, *CPVMAX*, *TRHINEG*, *TRHIPOS*, *TRLOWNEG*, *TRLOWPOS*, *TRTOT*, *LARIFE*, *LARFBE*, *LARCBE*, *IRONFB*, *IRONBB* und *IRONTOT* (Erklärung siehe Abschnitt 2.4.2). Es soll hier überprüft werden, welche Ergebnisse mit einem nur mit Hilfe von z-Vertex-, Driftkammer- und Proportionalkammergrößen (siehe Abschnitt 2.4.2) trainierten Netz erreicht werden können.
In Abbildung 4.22 sind die Rate von Subtrigger s52 a) vor und b) nach Abskalierung und c) die Rate nach der Entscheidung des installierten L2-Triggerelementes dargestellt. Die Abskalierungsfaktoren in den Triggerphasen 1, 2, 3 und 4 betragen 100, 6, 1 und 1. Dies zeigt sich in der Ratenreduktion von a) nach b). In b) tragen daher Phase 1 nahezu nicht und Phase 2 nur wenig zur Aufzeichnungsrate für diesen Subtrigger bei. Die Ratenreduktionsfaktoren durch das L2-Triggerelement (Reduktion von b) nach c)), ergeben sich in den Phasen 2, 3 und 4 zu 4, 3.5 und 3.1. Mit dieser Netzleistung sind die hier trainierten Netze im folgenden zu vergleichen.

4.3.1 Die Selektion der Trainingsdaten

In diesem Abschnitt wird auf die Selektion der Physikklasse und der Untergrundklasse eingegangen.

Selektion der Physikereignismenge

Von Ereignissen aus der Physikklasse werden neben der Anforderung, daß diese einer der L5-Ereignisklassen 17, 18 oder 24 (siehe Abschnitt 2.4.4) genügen müssen, folgende Bedingungen verlangt [30]:

- Es sind ein oder zwei Elektronen im Kalorimeter identifiziert worden. Ist nur ein Elektron vorhanden, muß dieses die höchste Qualität besitzen.
- Es gibt höchstens vier in den Spurkammern rekonstruierte, gute Spuren (siehe Anhang A).
- Der transversale Impuls der Zerfallselektronen des J/ψ-Mesons beträgt mindestens $P_t > 0.6\,GeV$.

Aus dieser Selektion standen 1200 Physikereignisse zur Verfügung.

Selektion der Untergrundereignismenge

Die Untergrundereignisse wurden aus 14 Datennahmeeinheiten (siehe Abschnitt 2.4.1) mit dem Attribut L2-L4-*transparent* gewonnen. Die Ereignisse werden dabei allein auf Grund einer positiven L1-Entscheidung gespeichert. Bei der Untergrundselektion ist es im Gegensatz zu Subtrigger s34 bei s52 und s54 von Bedeutung, daß die L2-Triggerstufe bei der Datennahme nicht aktiv war, da auf diesen beiden Subtriggern L2-Triggerelemente am Experiment laufen.
Die verwendeten Datennahmeeinheiten sind über einen Zeitraum von 40 Tagen verteilt, um eine spätere Abhängigkeit der Netze von speziellen Gegebenheiten im Detektor während weniger spezieller Datennahmeeinheiten zu vermeiden. Dies trägt dazu bei, am Ende Neuronale Netze zu erhalten, welche die Physik vom Untergrund an einem speziellen Subtrigger generell gut unterscheiden können. Zum Training wurden 1200 Untergrundereignisse benutzt.
Mit Hilfe des Kriteriums von Seite 54 *ff.* wurden 83 dieser 1200 Ereignisse als kosmische Myonen identifiziert. Dies ist ein Anteil von $(6.9 \pm 0.8)\%$.

4.3.2 Netzentwicklung

Zum Training der Netze für den Subtrigger s52 standen insgesamt 2400 Ereignisse zur Verfügung. Diese teilten sich wie in Tabelle 4.11 gezeigt in Physikereignisse und Untergrundereignisse und in Trainings- und Testereignisse auf.

	Trainingsereignisse	Testereignisse
Untergrund	800	400
Physik	800	400

Tabelle 4.11: *Anzahlen der zur Verfügung stehenden Trainings- und Testereignisse für das Training s52.*

Die Verteilungen der benutzten Eingabegrößen sind in Abbildung 4.23 gezeigt. Der Untergrund ist dabei ähnlich wie bei den Verteilungen der Eingabegrößen zum Training für Subtrigger s34 stärker in Vorwärtsrichtung und zu höheren Spurmultiplizitäten verteilt als die Physik. Die Position des z-Vertexes CPVPOS ist im Untergrund nahezu gleichverteilt, während sie in der Physik enger um die Position des nominellen Wechselwirkungspunktes herum verteilt ist. Keine der Größen eignet sich für einen eindimensionalen Schnitt zur Trennung von Untergrund und Physik.

Training s52

Das Training wurde über den gleichen Parameterbereich wie in dem Training zu Subtrigger s34 durchgeführt. Die Parameter- und Knotenkonfigurationen sind daher schon in den Tabellen 4.2 und 4.3 definiert worden. Es wurde 500 Epochen lang trainiert. Das Ergebnis des Trainings ist in Abbildung 4.24 gezeigt.

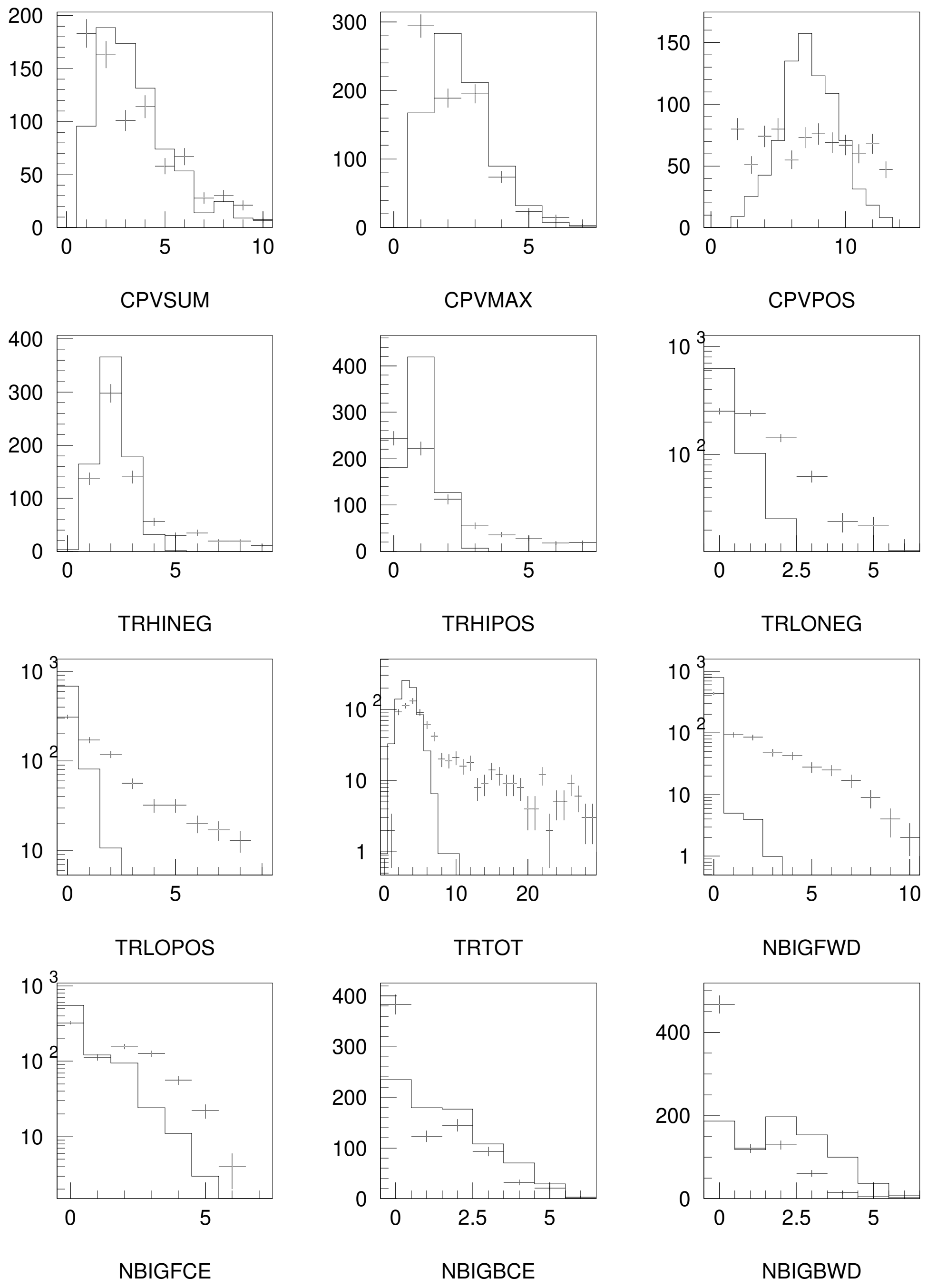


Abbildung 4.23: *Verteilungen der Eingabegrößen der zum Training s52 benutzten Ereignisse der Untergrund- und Physikereignisklasse. Die Einträge der Physikereignisse wurden in der Summe auf den Untergrund normiert und sind durch die durchgezogenen Linien dargestellt. Zum Vergleich sind die Untergrundverteilungen mit statistischen Fehlern durch Punkte in den gleichen Diagrammen eingetragen.*

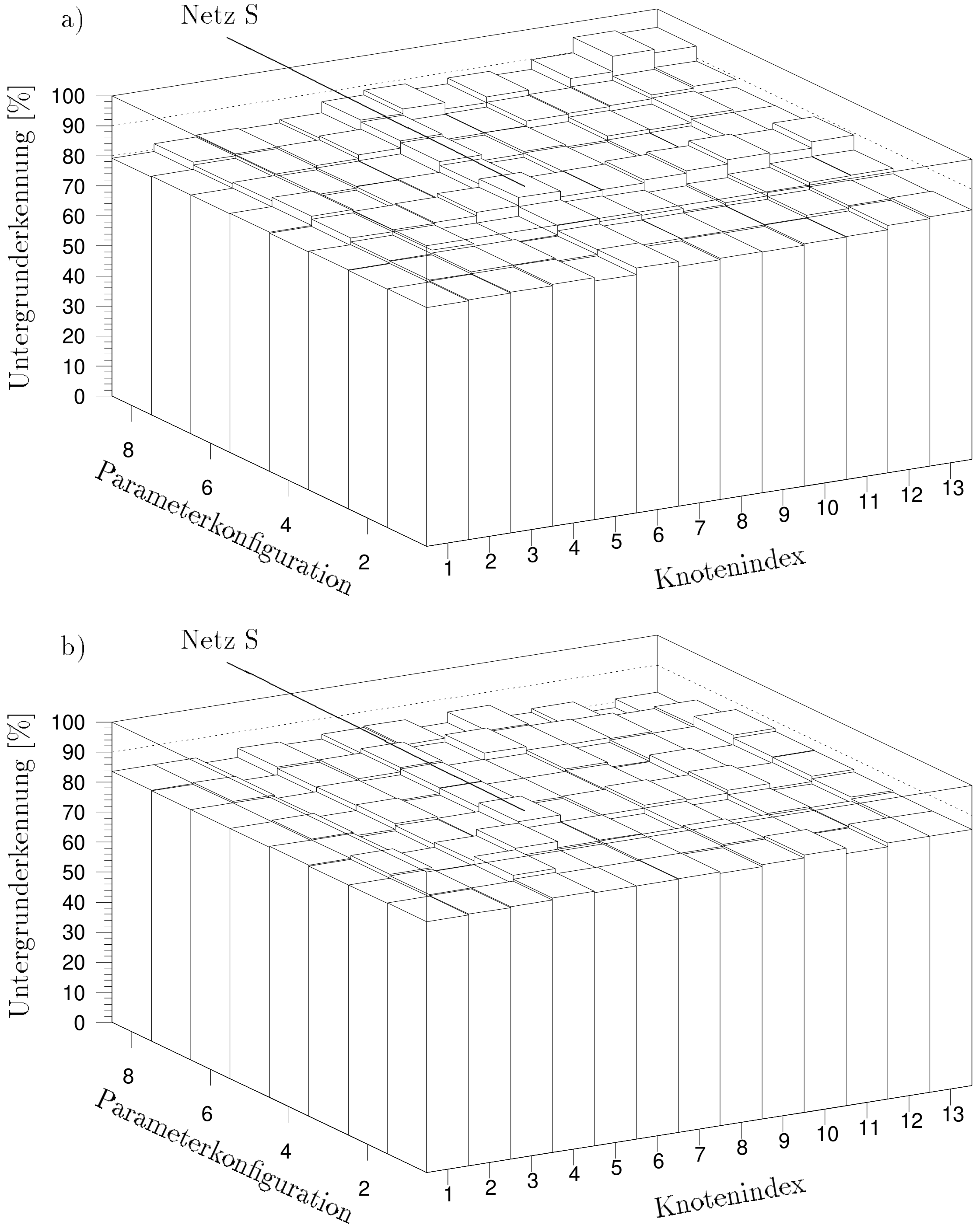


Abbildung 4.24: *a) Trainingsergebnisse und b) Testergebnisse des Trainings s52. Dargestellt ist der Anteil der Untergrundereignisse, den das bei der jeweiligen Parameterkonfiguration und dem Knotenindex (siehe die Tabellen 4.2 und 4.3) trainierte Neuronale Netz als Untergrundereignis klassifiziert, wobei* 90% *der Physikereignisse richtig identifiziert worden sind. Errechnet wurden die Werte bei a) mit 800 Trainingsereignissen und in b) mit 400 Testereignissen (siehe Tabelle 4.11). Die exakten Zahlenwerte können in Anhang B (Abbildung B.4 a) und b)) eingesehen werden.*

Man erkennt in Abbildung 4.24 a), daß die Untergrunderkennung bei den Trainingsereignissen mit größer werdender Parameterkonfiguration und auch leicht mit ansteigender Zahl versteckter Knoten zunimmt. Diese Zunahme ist bis hin zu ungefähr 6 Knoten in der versteckten Schicht auch für die Testereignisse in Abbildung 4.24 b) zu erkennen. Mit höherer Anzahl versteckter Knoten setzt der Effekt des Übertrainierens ein und die Erkennung der Ereignisse aus der Testmenge wird schlechter, die Erkennung der Ereignisse aus der Trainingsmenge besser.
Um ein nicht übertrainiertes Netz zu erhalten, wird dieses bei weniger als sieben Knoten ausgewählt. Als bestes Netz stellte sich das Netz mit 6 versteckten Knoten und einer Parameterkonfiguration von 4, das bedeutet mit einem Momentumfaktor von 0.9 und einer konstante Lernrate von 0.001 heraus (siehe Abbildung 4.24). Dieses Netz wird im folgenden Netz S genannt.

4.3.3 Untersuchungen an Netz S

In Abbildung 4.25 sind die Ausgabewerte von Netz S über den jeweils 400 Testereignissen für Untergrund und Physik gezeigt. Die kosmischen Myonen sind gesondert dargestellt (Abbildung 4.25 c)) und werden von Netz S weit gestreut klassifiziert mit Tendenz in Richtung 1. Diese Abbildungen sind die Grundlage der Berechnung für Abbildung 4.26. Da der Schnittwert (siehe Abschnitt 4.2.2) für das Netz noch nicht feststeht und erst bei Einbau des Netzes, angepaßt an die dann vorherrschenden experimentellen Bedingungen, festgelegt werden kann, sind in Abbildung 4.26 Untergrund- und Physikerkennung als Funktion des Schnittwertes dargestellt. Wie man bei Wahl

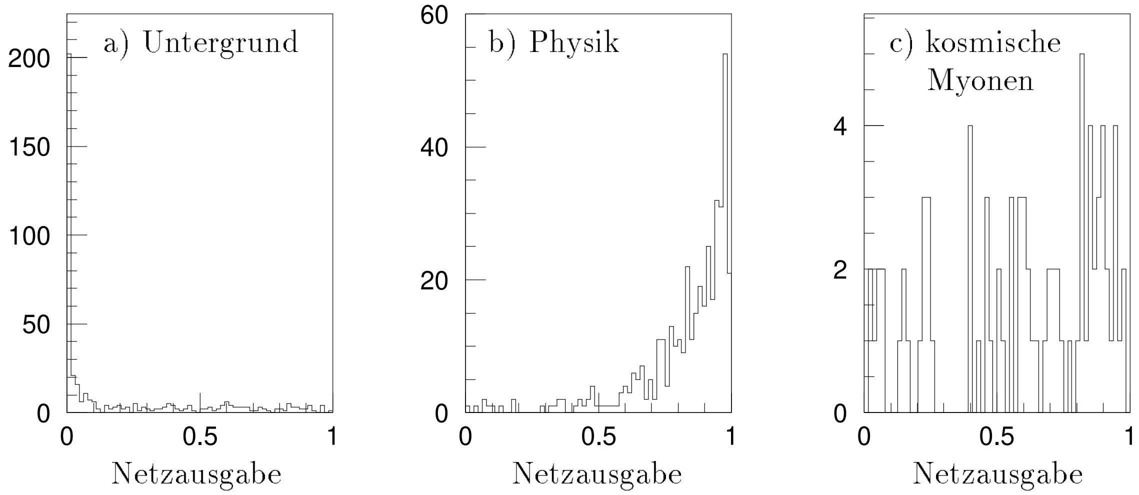


Abbildung 4.25: *Darstellung der Netzausgabewerte für a) Untergrund-, b) Physikereignisse und c) kosmische Myonen für das Netz S. Dargestellt sind die Netzausgabewerte für je 400 Untergrund- und Physikereignisse der Testereignismenge und 83 kosmische Myonen aus dem s52-Untergrund.*

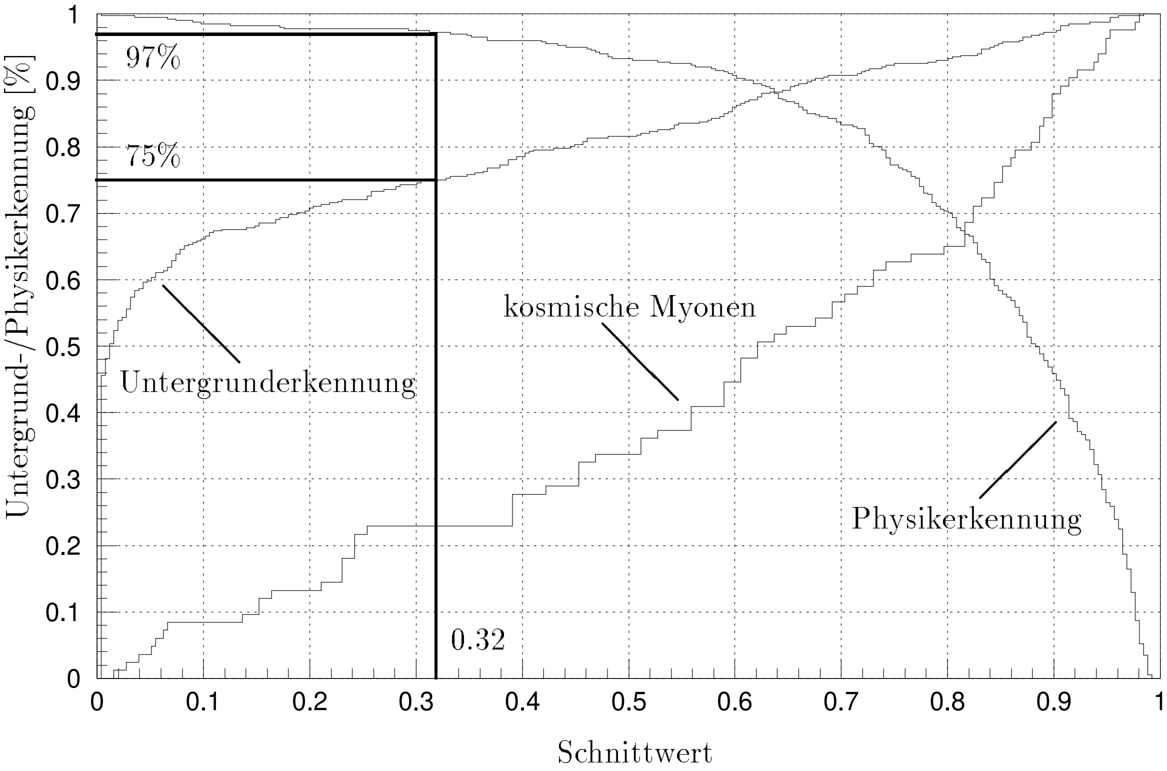


Abbildung 4.26: *Untergrund- und Physikerkennung von Netz S über dem zugehörigen Schnittwert. Zur Berechnung wurden die jeweils 400 Testereignisse für Untergrund und Physik und 83 kosmische Myonen verwendet.*

eines Schnittwertes die Untergrund- und die Physikerkennung berechnet, wurde bereits in Kapitel 4.2.2 dargestellt.

Durch die am Experiment installierte Neuronale L2-Konfiguration wird ein, wie aus Abbildung 4.22 ermittelt wurde, maximaler Reduktionsfaktor von der Rate vor zu der Rate nach den Netzen auf L2 von 4 erreicht. Ein Faktor 4 für die Reduktion durch Stufe L2 bedeutet eine Untergrunderkennung von 75%. Diese wird durch Netz S bei einem Schnittwert von 0.32 erreicht. Dabei beträgt die Physikerkennung ungefähr 97%. Diese Zusammenhänge werden in Abbildung 4.26 verdeutlicht.

Bei einer Effizienz von 97% richtig erkannter Physikereignisse, werden nur 23% der kosmischen Myonen als Untergrund klassifiziert. Eine Trennung von Physik und kosmischen Myonen ist daher bei den verwendeten Daten nicht möglich.

Bemerkt sei hier, daß in der für das hier durchgeführte Training benutzten Physikereignismenge keine Ereignisse vorhanden sind, die weder vom installierten Neuronalen Netz auf Subtrigger s52 noch von irgendeinem anderen Subtrigger getriggert worden sind. Über eben diese Art von Ereignissen fällt Netz S im Falle einer Implementation am Experiment jedoch auch eine Entscheidung, über die hier keine Aussage gemacht werden kann. Der Anteil solcher Ereignisse an allen von Subtrigger s52 getriggerten

Ereignissen ist jedoch klein [30]. Auf dieser Annahme und auf Daten von 1996 beruht auch eine Berechnung der Effizienz von 73% für die Physikerkennung des am Experiment für Subtrigger s52 installierten Netzes. Netz S ist bei gegenüber dem installierten Netz gleicher Untergrunderkennung mit 97% (siehe Abbildung 4.26) richtig klassifizierter Physik deutlich effizienter mit der oben bemerkten Unsicherheit bezüglich der vom installierten Netz verworfenen und von keinem anderen Trigger getriggerten Ereignisse.

4.4 Subtrigger s54

Der Subtrigger s54 ist ebenso wie Subtrigger s34 und s52 für diffraktiv erzeugte J/ψ- oder Υ-Mesonen konzipiert worden; es wird dabei jedoch keine Teilchenidentifikation ausgenutzt. Die Ereignisse werden nur auf Grund ihrer Spurtopologien getriggert. In der folgenden Triggerelementdarstellung von Subtrigger s54 bedeutet & logisches UND.

Definition des Subtriggers s54 für das Jahr 1997:

- s54: zVtx_Cls & DCRPh_TNeg & DCRPh_THig & Topo_BR

Die Triggerelemente sind im einzelnen:
zVtx_Cls : Alle Einträge im z-Vertex-Histogramm befinden sich innerhalb von höchstens vier benachbarten Intervallen.
DCRPh_TNeg : Mindestens eine negative Maske des DCrphi-Triggers wurde gesetzt.
DCRPh_THig : Im DCrphi-Trigger wurde eine Maske mit einem transversalen Impuls von mehr als 800 MeV getriggert.
Topo_BR : Genau 2 *Bigrays*, die sich in der Φ-Projektion der in Abschnitt 2.4.2 beschriebenen Θ-Φ-Karte in Φ-Intervallen befinden, die sich geometrisch gegenüber liegen.

Neben der Bedingung zVtx_Cls für den z-Vertex wird die Bedingung Topo_BR benutzt, um insbesondere diffraktive Ereignisse zu triggern. Dadurch, daß die beiden geforderten Bigrays sich in Φ-Projektion gegenüber liegen müssen, werden elastische $J/\psi \to \mu^+\mu^-, e^+e^-$ mit niedrigem Transversalimpuls des J/ψ bevorzugt.
Wie bei Subtrigger s52 ist am Experiment auch auf Subtrigger s54 schon ein L2-Triggerelement installiert, welches eine Triggerentscheidung mit Hilfe folgender Eingabegrößen trifft: *CPVSUM*, *CPVPOS*, *CPVMAX*, *TRHINEG*, *TRHIPOS*, *TRLOWNEG*, *TRLOWPOS*, *TRTOT*, *LARIFE*, *LARFBE*, *LARCBE*, *IRONFB*, *IRONBB* und *IRONTOT* (Erklärung siehe Abschnitt 2.4.2). Hier soll untersucht werden, was man mit einem nur mit z-Vertex-, Driftkammer- und Proportionalkammergrößen trainierten Netz gegenüber den vorhandenen Netzen erreichen kann.
In Abbildung 4.27 ist die Ratenreduktion durch die sich am Experiment im Einsatz befindlichen Netze zu erkennen. Nach einer Abskalierung der L1-Rate mit den Abskalierungsfaktoren 100, 6, 1 und 1 in den Phasen 1, 2, 3 und 4 wird durch das L2-Triggerelement eine weitere Ratenreduktion der Rate um die Faktoren 2.3, 2.5 und 2.3

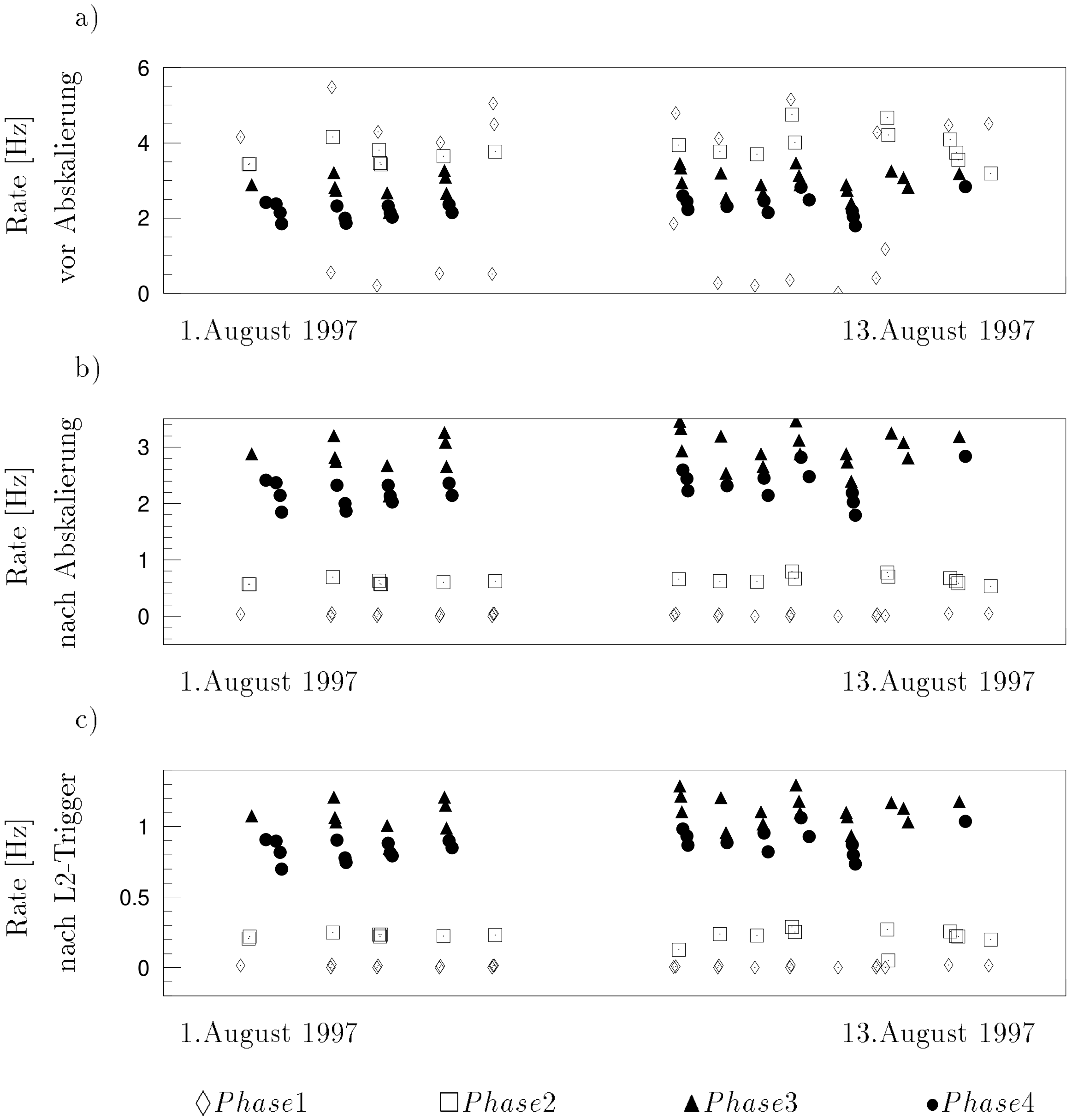


Abbildung 4.27: *Raten des Subtriggers s54 a) ohne, b) mit Abskalierungsfaktor und c) nach positiver L2-Triggerentscheidung. Der Abskalierungsfaktor ist 100 für Phase 1, weshalb die abgestufte Rate in b) in dieser Phase nahezu Null beträgt. Für die Phase 2 ist 6 und für die Phasen 3 und 4 ein Abskalierungsfaktor von 1 festgesetzt.*

in den Phasen 2, 3 und 4 erreicht. Auch hier ist wie bei Subtrigger s52 die Rate in Phase 1 mit dem Abskalierungsfaktor von 100 unterdrückt.

4.4.1 Die Selektion der Trainingsdaten

Selektion der Physikereignismenge

Um die Physikereignismenge für das Training zu erhalten, wurden durch Subtrigger s54 getriggerte Ereignisse nach Durchlaufen der L4-Triggerstufe und der L5-Ereignisklassifikation mit den Klassen 17, 18 oder 24 (siehe Abschnitt 2.4.4) selektiert. Da Subtrigger s54 Ereignisse auf Grund ihrer Spurtopologie im Detektor triggert, sind sowohl in Elektronen als auch in Myonen zerfallene J/ψ-Mesonen unter den selektierten Ereignissen. Diese wurden getrennt weiter selektiert [30].

Schnitte für die Selektion von $J/\psi \rightarrow \mu^+\mu^-$ Kandidaten:

- Es sind ein oder zwei Myonen im Eisendetektor oder im Kalorimeter identifiziert worden. Ist jedoch nur ein Myon im Kalorimeter und keines im Eisen gefunden worden, so soll dieses von höchster Qualität sein.
- Es gibt höchstens vier in den Spurkammern rekonstruierte, gute Spuren (siehe Anhang A).
- Es muß der gleiche Schnitt gegen kosmische Myonen wie schon bei der Selektion für Subtrigger s34 (siehe Seite 45) erfüllt werden.
- Die Zerfallsmyonen des J/ψ müssen einen transversalen Impuls von $P_t > 0.6\,GeV$ besitzen.

Schnitte für die Selektion von $J/\psi \rightarrow e^+e^-$ Kandidaten:

Bis auf eine Verschärfung des Schnittes auf die Anzahl guter Spuren (siehe Anhang A), die hier auf genau 2 beschränkt wird, ist die hier verwendete Selektion identisch mit der Physikselektion zu Subtrigger s52 in Abschnitt 4.3.

Aus den beiden Selektionen der Physikereignisse für Subtrigger s54 standen jeweils 1000 Ereignisse zur Verfügung.

Selektion der Untergrundereignismenge

Wie schon bei der Selektion der Untergrundereignisse für das Netztraining zu Subtrigger s52 wurden Ereignisse aus 14 Datennahmeeinheiten (siehe Abschnitt 2.4.1) mit dem Attribut L2-L4-*transparent* selektiert, die über einen Zeitraum von 40 Tagen verteilt lagen, um zeitliche Unterschiede und spezifische Eigenschaften des Untergrundes aus

verschiedenen Datennahmeeinheiten zu mitteln. Es standen 2000 Untergrundereignisse für das Training zur Verfügung.
Mit Hilfe des Kriteriums von Seite 54 *ff.* konnten 299 dieser 2000 Ereignisse als kosmische Myonen identifiziert werden. Dies ist mit statistischem Fehler ein Anteil von $(15.0 \pm 0.9)\%$.

4.4.2 Netzentwicklung

Für das Training standen insgesamt 4000 Ereignisse zur Verfügung. Diese teilen sich, wie in Tabelle 4.12 dargestellt ist, in Physik- und Untergrundereignisse und in Trainings- und Testereignisse auf.

	Trainingsereignisse	Testereignisse
Untergrund	1400	600
Physik aus myonischer Selektion	700	300
Physik aus elektronischer Selektion	700	300

Tabelle 4.12: *Anzahlen der zur Verfügung stehenden Trainings- und Testereignisse für das Training s54.*

Die Verteilungen der Netzeingabegrößen sind in Abbildung 4.28 getrennt nach Untergrund- und Physikereignissen dargestellt. Auch hier ist wie schon bei dem vorangegangenen Training der Untergrund mehr zu hohen Spurmultiplizitäten verteilt als die Physik, und die z-Vertex-Verteilung der Untergrundereignisse ist flach, während sich die der Physikereignisse um den nominellen Wechselwirkungspunkt verteilt darstellt. Auch hier kann in keiner der einzelnen Größen ein effizienter Schnitt zur Trennung von Untergrund und Physik angebracht werden, so daß man nur versuchen kann, die Trennung mit Hilfe höherdimensionaler Korrelationen zwischen den Größen, wie es z.B. bei einem Neuronalen Netz geschieht, zu erreichen.
Die im Training erzielten Ergebnisse sind in Abbildung 4.29 dargestellt. Auch hier erkennt man wie bei den Untersuchungen zu den beiden vorherigen Subtriggern den Effekt des Übertrainierens bei hohen Anzahlen versteckter Knoten und zu hohem Wert der Parameterkonfigurationen hin. Ausgewählt wurde hier das Netz T mit 4 Knoten in der versteckten Schicht und einer Parameterkonfiguration von 2, da im Bereich um diese Parameterwerte die Erkennung von Test- und Trainingsuntergrund bei 90% erkannter Physik ungefähr gleich groß ist (siehe die Zahlenwerte zu Abbildung 4.29 in Anhang B Abbildung B.5).

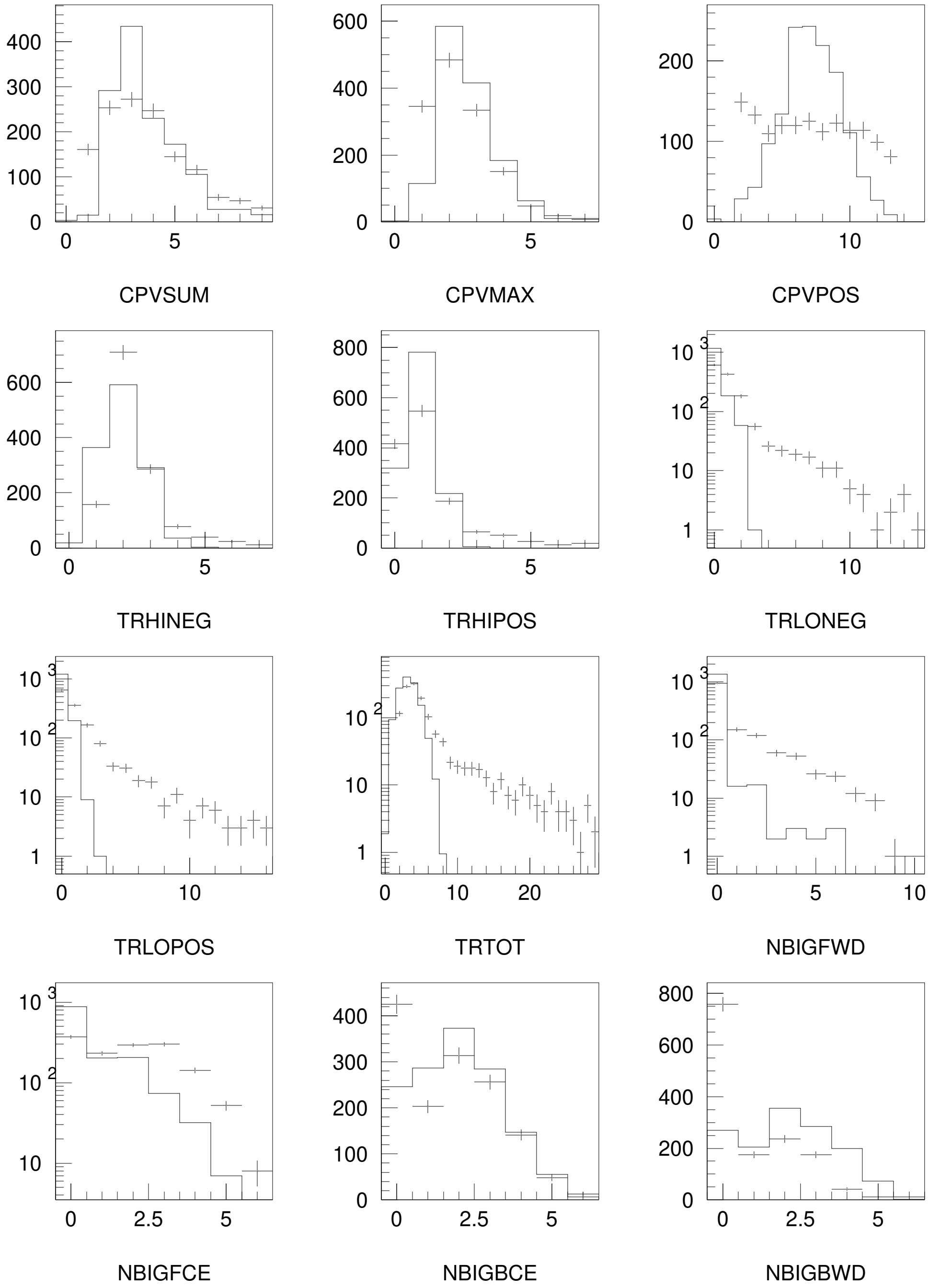


Abbildung 4.28: *Verteilungen der Eingabegrößen der zum Training s54 benutzten Ereignisse der Untergrund- und Physikereignisklasse. Die Einträge der Physikereignisse wurden in der Summe auf den Untergrund normiert und sind durch die durchgezogenen Linien dargestellt. Zum Vergleich sind die Untergrundverteilungen mit statistischen Fehlern durch Punkte in den gleichen Diagrammen eingetragen.*

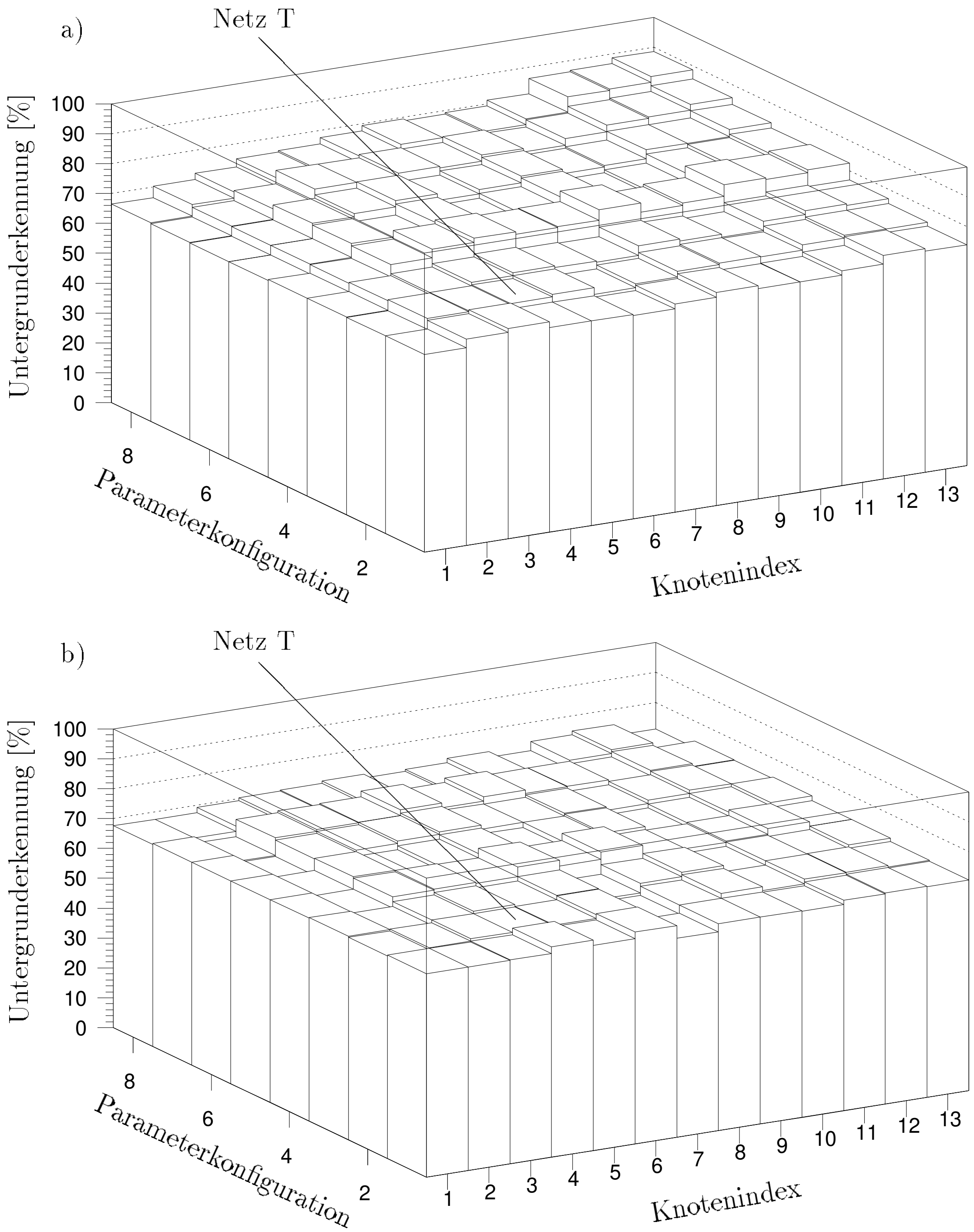


Abbildung 4.29: *a) Trainingsergebnisse und b) Testerbegnisse des Trainings s54. Dargestellt ist der Anteil der Untergrundereignisse, den das bei der jeweiligen Parameterkonfiguration und dem Knotenindex (siehe die Tabellen 4.2 und 4.3.) trainierte Neuronale Netz als Untergrundereignis klassifiziert, wobei* 90% *der Physikereignisse richtig identifiziert worden sind. Errechnet wurden die Werte bei a) mit den jeweils 1400 Trainingsereignissen und in b) mit jeweils 600 Testereignissen (siehe Tabelle 4.12). Die exakten Zahlenwerte können in Anhang B (Abbildung B.5 a) und b)) eingesehen werden.*

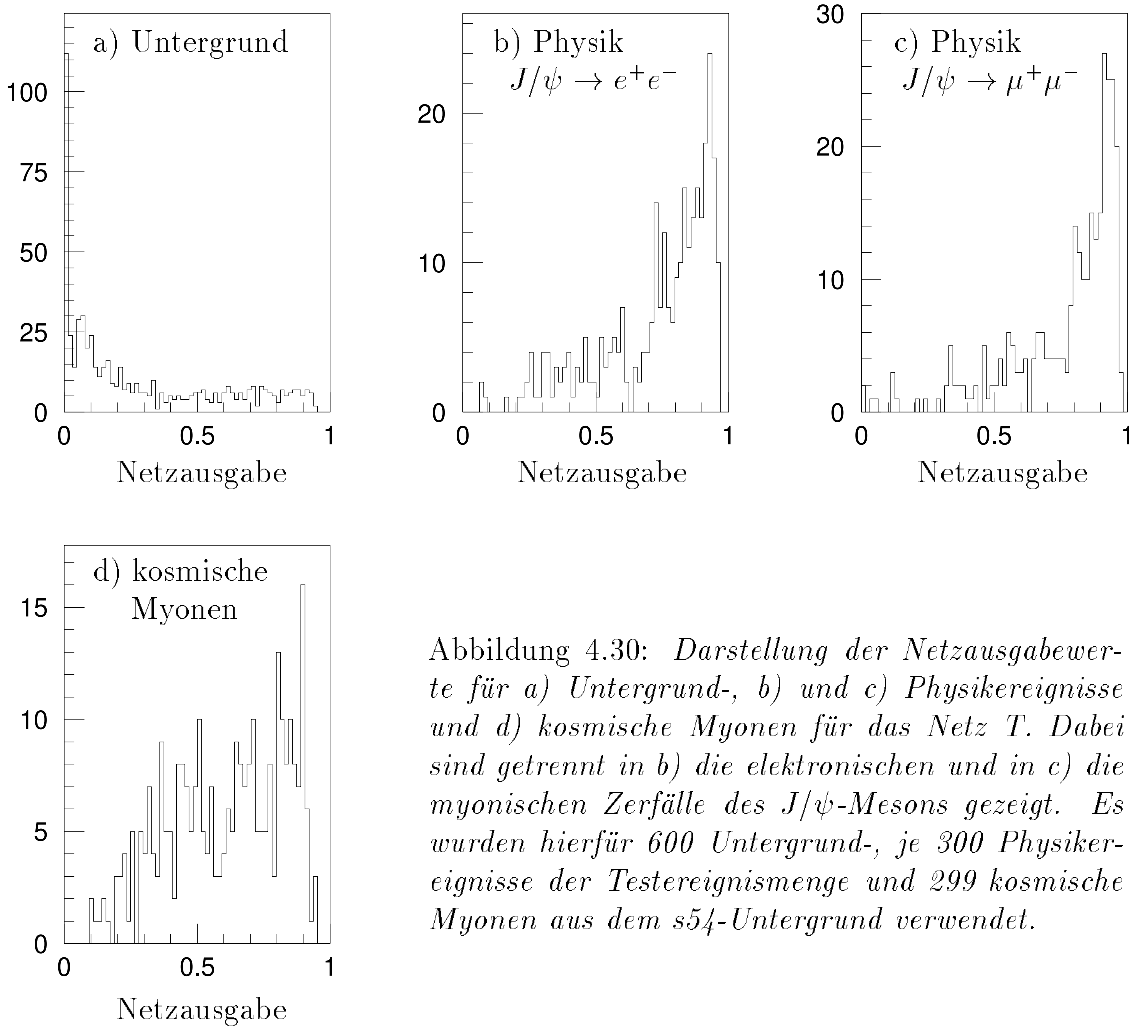


Abbildung 4.30: *Darstellung der Netzausgabewerte für a) Untergrund-, b) und c) Physikereignisse und d) kosmische Myonen für das Netz T. Dabei sind getrennt in b) die elektronischen und in c) die myonischen Zerfälle des J/ψ-Mesons gezeigt. Es wurden hierfür 600 Untergrund-, je 300 Physikereignisse der Testereignismenge und 299 kosmische Myonen aus dem s54-Untergrund verwendet.*

4.4.3 Untersuchungen an Netz T

Nach den Darstellungen der Netzausgabewerte für den Untergrund und die Physikereignisse getrennt nach Zerfällen des J/ψ-Mesons in Elektronen und Myonen in den Abbildungen 4.30 a)-c) kann eine Trennung der Physikereignisse von den Untergrundereignissen für Elektronen und Myonen erreicht werden. In Abbildung 4.31 wird jedoch deutlich, daß in weiten Bereichen des Schnittwertes die myonischen Zerfälle besser als die elektronischen erkannt werden. Die aus den Daten zu Abbildung 4.27 ermittelte Ratenreduktion der am Experiment befindlichen Neuronalen L2-Konfiguration beträgt maximal 2.5. Dies entspricht einer Untergrunderkennung von 60%. Bei diesem Referenzwert sind in Abbildung 4.31 zu Netz T der Schnittwert zu 0.26 und der Anteil der richtig erkannten Physik für beide Physikklassen zu 97% bestimmt worden. Die Ausgabewerte für die kosmischen Myonen sind in Abbildung 4.30 gezeigt und befinden sich ähnlich wie bei Netz S für Subtrigger s52 mit ihren Netzausgabewerten zwischen Untergrund und Physik mit Tendenz zur Physik. Aus der Auswertung der Kurven

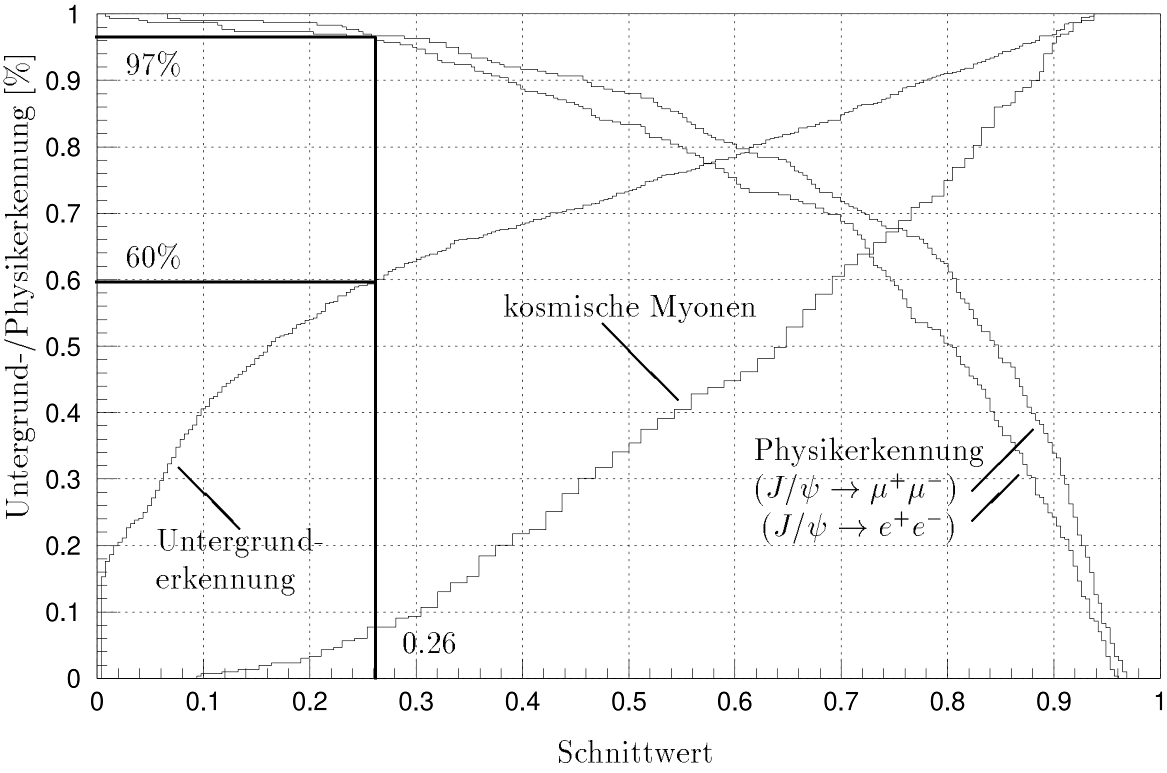


Abbildung 4.31: *Untergrund- und Physikerkennung von Netz T über dem zugehörigen Schnittwert. Zur Berechnung wurden 600 Testereignisse für den Untergrund, je 300 Testereignisse für die Elektronen- und die Myonenselektion und 299 kosmische Myonen verwendet.*

von Abbildung 4.31 geht hervor, daß bei 97% richtig erkannter Physik nur 8% der kosmischen Myonen als Untergrund klassifiziert werden. Eine Trennung von kosmischen Myonen und Physik ist bei den benutzten Daten hier nicht möglich.
Wie schon für Netz S am Ende von Abschnitt 4.3.3 erklärt wurde ist die Effizienz von 97% im Falle einer Implementation von Netz T am Experiment unter Vorbehalt als Vorhersage zu betrachten, da es eine Klasse von Ereignissen gibt, die nicht zum Training benutzt wurde über die aber Netz T am Experiment zu entscheiden hätte (siehe Abschnitt 4.3.3). Die Ereignisanzahl in dieser Klasse von Ereignissen ist jedoch auch bei Subtrigger s54 klein [30]. Nur unter diesem Vorbehalt kann aus Daten von 1996 eine Effizienz des installierten Neuronalen Netztriggers von 99% berechnet werden. Unter Beachtung der beschriebenen Unsicherheit ist Netz T damit bei gleicher Untergrunderkennung mit 97% erkannter Physik (siehe Abbildung 4.31) nahezu gleich gut.

4.5 Ein Netz für Subtrigger s52 und s54 zusammen

Als abschließende Untersuchung wurde für Subtrigger s52 und s54 ein gemeinsames Netz trainiert, da sich die physikalisch interessanten Ereignisse dieser Subtrigger topologisch ähnlich sind. Die zum Training und Testen verwendeten Ereignisse stammen aus den in den Abschnitten 4.3.1 uns 4.4.1 beschriebenen Selektionen.

4.5.1 Netzentwicklung

In Tabelle 4.13 sind die zum Training verwendeten Ereignisanzahlen aufgelistet. Es wurden insgesamt 4000 Ereignisse zum Training und Testen verwendet. Dabei wurden für die Untergrund- und die elektronische Physikselektion zu Subtrigger s54 Ereignisse nicht zugelassen, welche sich schon in einer der Selektionen von Subtrigger s52 befinden, damit sich die trainierten Netze nicht auf die Ereignisschnittmenge von s52 und s54 spezialisieren, sondern die allgemeinen Eigenschaften der durch s52 und s54 getriggerten Ereignisse zur Unterscheidung vom Untergrund nutzen.

	Trainingsereignisse	Testereignisse
Untergrund s52	700	300
Untergrund s54	700	300
Physik s52	700	300
Physik s54 aus elektronischer Selektion	350	150
Physik s54 aus myonischer Selektion	350	150

Tabelle 4.13: *Anzahlen der zur Verfügung stehenden Trainings- und Testereignisse für das Training s52s54.*

Training s52s54

Das Training wurde mit den gleichen Knotenanzahlen in der versteckten Schicht und Parameterkonfigurationen wie zuvor in den Abschnitten 4.3 und 4.4 durchgeführt. Die Anzahl der Epochen betrug 500. Das Netz mit 2 Knoten in der versteckten Schicht und einer Parameterkonfiguration von 2 wurde zur weiteren Untersuchung ausgewählt. Dieses wird im folgenden Netz U genannt.

4.5.2 Untersuchungen an Netz U

In Abbildung 4.32 a)-g) sind die Netzausgabewerte für die fünf beim Training verwendeten Ereignisklassen und die kosmischen Myonen aus den Untergrunddaten darge-

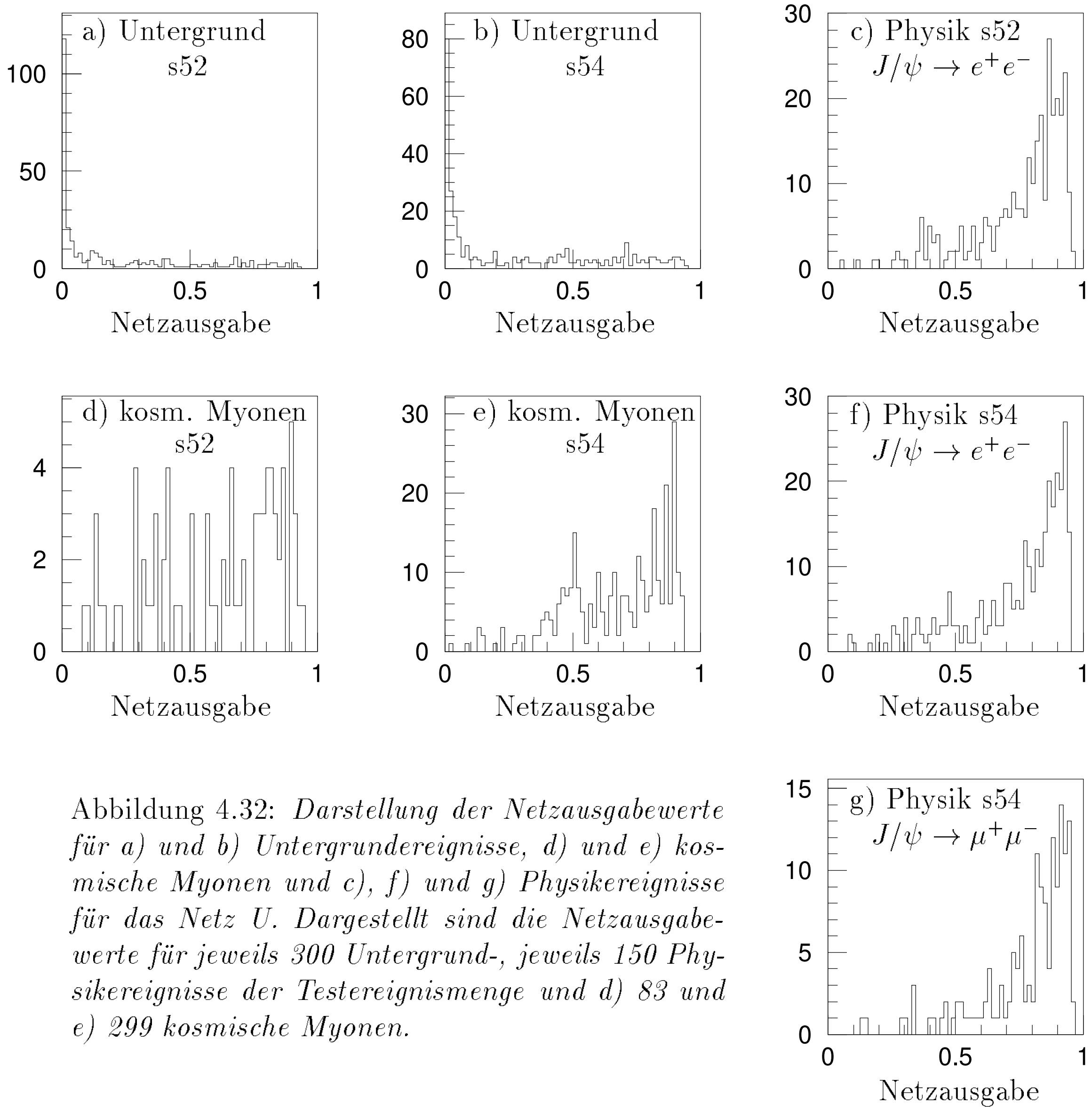


Abbildung 4.32: *Darstellung der Netzausgabewerte für a) und b) Untergrundereignisse, d) und e) kosmische Myonen und c), f) und g) Physikereignisse für das Netz U. Dargestellt sind die Netzausgabewerte für jeweils 300 Untergrund-, jeweils 150 Physikereignisse der Testereignismenge und d) 83 und e) 299 kosmische Myonen.*

stellt. Die elektronische Selektion für Subtrigger s54 wurde bei diesem Test nicht wie zum Training auf die Ereignisse eingeschränkt, welche nicht auch von Subtrigger s52 getriggert wurden. In Abbildung 4.33 sind Untergrund- und Physikerkennung über dem Schnittparameter dargestellt. Über einen großen Schnittwertbereich werden die Physikereignisse der myonischen Selektion von Subtrigger s54 durch Netz U am besten klassifiziert, die Physikereignisse der elektronischen Selektion zu Subtrigger s54 am schlechtesten. Die Physikereignisse zu Subtrigger s52 liegen dazwischen. Der Untergrund von s52 wird deutlich besser erkannt als der Untergrund von s54. Diese Aussage gilt ebenfalls für die kosmischen Myonen aus den Untergrunddaten, wobei auch hier nicht von einer Trennung der kosmischen Myonen von den Physikereignissen gesprochen werden kann (vergleiche mit den Abschnitten 4.3 und 4.4).

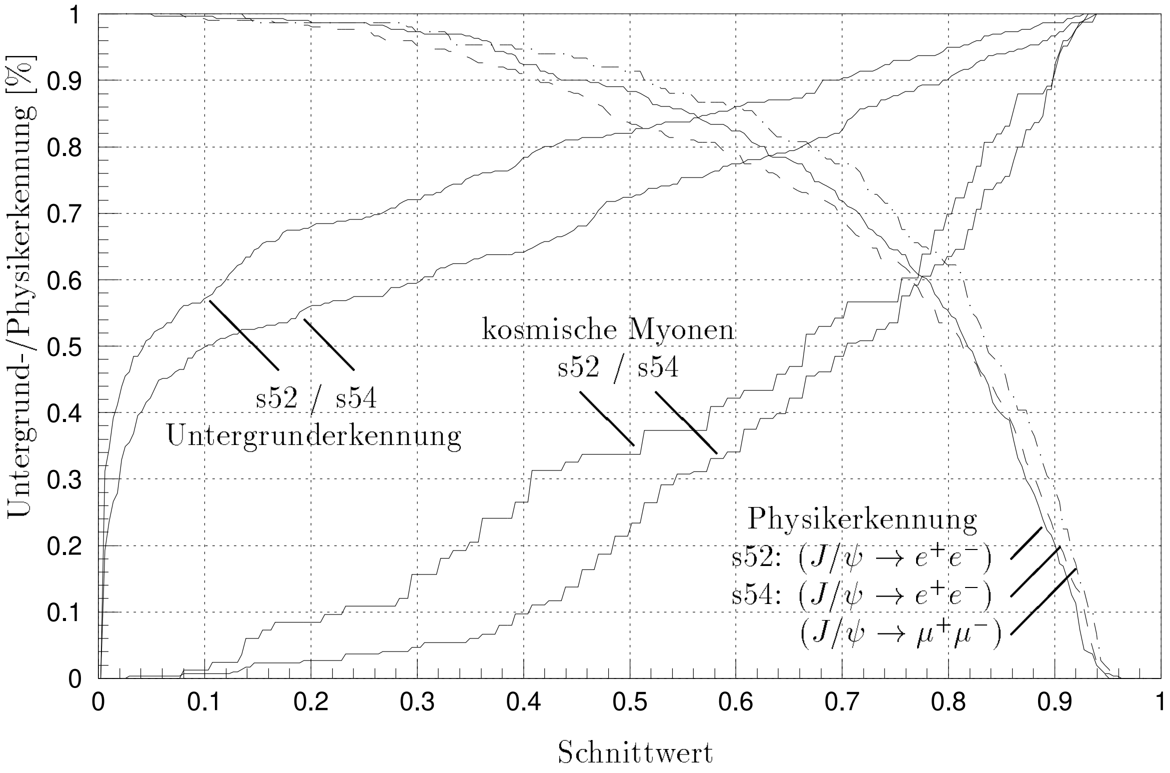


Abbildung 4.33: *Untergrund- und Physikerkennung von Netz U über dem zugehörigen Schnittwert . Zur Berechnung wurden die in Abbildung 4.32 in ihren Netzausgabewerten dargestellten Testereignisse und kosmischen Myonen verwendet.*

Bei einem am Experiment realisierten Ratenreduktionsfaktor von 4 für s52-Untergrund, ist die Effizienz der Physikerkennung von Netz U für Ereignisse von s52 wie bei dem nur für Ereignisse von Subtrigger s52 trainierten Netz S mit 97% gleich geblieben (vergl. Abbildung 4.26). Für einen Ratenreduktionsfaktor von 2.5 für den Untergrund von s54 ist die Physikerkennung von Netz U gegenüber der von Netz T für $J/\psi \to \mu^+\mu^-$-Ereignisse mit 97% ebenfalls gleich groß, während sich die Physikerkennung der $J/\psi \to e^+e^-$-Ereignisse von s54 um 2% auf 95% verschlechtert hat (vergl. Abbildung 4.31).

Kapitel 5

Zusammenfassung der Ergebnisse

Im Rahmen dieser Arbeit wurde die Technik der Neuronalen Netze verwendet, um diffraktiv erzeugte J/ψ-Mesonen, die über den Zerfallskanal in zwei Myonen detektiert werden, auf Triggerebene von Untergrundereignissen zu trennen. Dies geschah speziell für die Subtrigger s34, s52 und s54. Ein Neuronales Netz ist dann für die zweite Triggerstufe im H1-Trigger geeignet, wenn es bei guter Erkennung physikalisch interessanter Ereignisse, die Rate des Triggers vermindert, möglichst die Totzeit verringert und eine Bestimmung der Effizienz des Triggers ermöglicht.
Für jeden der drei oben genannten Subtrigger wurden Untergrundereignisse und physikalisch interessante Ereignisse selektiert. Die Neuronalen Netze sind daraufhin mit zur Verfügung stehenden Detektorinformationen unter der Vorgabe diese Ereignisklassen möglichst gut voneinander unterscheiden zu können trainiert worden. Dabei ergaben sich für Subtrigger s34 die besten Ergebnisse unter Verwendung von Größen aus den Spurkammern und aus dem Myondetektor. Die Netze für die Subtrigger s52 und s54 sind ausschließlich mit Größen aus den Spurkammern trainiert worden.
Ein Teil des Untergrundes bestand bei allen drei Subtriggern aus kosmischen Myonen. Bei keinem der drei Subtrigger konnten die als Untergrund im Detektor anfallenden kosmischen Myonen von den J/ψ-Mesonen getrennt werden. Die zur Verfügung stehenden Netzeingabegrößen ließen dies nicht zu. Das könnte daran liegen, daß diese Größen nur grobe Informationen über die Flugrichtung der Teilchen im Detektor enthalten. Die relativ geringen Anteile kosmischer Myonen von $(6.9 \pm 0.8)\%$ bei s52 und $(15.0 \pm 0.9)\%$ bei s54 wurden jedoch nicht gesondert betrachtet. Der Anteil kosmischer Myonen im Untergrund zu s34 betrug bis zu 70%; diesem mußte also besondere Aufmerksamkeit gewidmet werden.
Betrachtet man nur die Untergrundereignisse zu **s34**, die keine Signatur eines kosmischen Myons trugen, so konnte bei einer Effizienz von 95% erkannter Physikereignisse 80% dieses Untergrundanteils verworfen werden. In einem Test mit Untergrunddaten, die 3 Monate später als die zum Training verwendeten Ereignisse selektiert wurden, ging dieser Anteil auf 71% zurück. Das Netz zeigte eine Ratenreduktionen um bis zu 33%, was einem Ratenreduktionsfaktor von 1.5 als Verhältnis der Rate nach zu der Ra-

te vor der zweiten Triggerstufe entspricht. Eine Simulation der Triggerstufe L4 ergab eine durch die Aufnahme des Netzes in die simulierte Triggerkonfiguration verursachte Ratenreduktion der Rate nach Triggerstufe L4 um maximal 24%. Der hieraus folgende Ratenreduktionsfaktor nach L4 ist mit 1.3 relativ niedrig. Der Vorteil der Verwendung dieses Netzes im Trigger ist eher darin zu sehen, daß die Ereignisse aus dem Untergrund mit negativer Netzentscheidung einen deutlich größeren Datenumfang besitzen als diejenigen, welche durchgelassen werden, so daß Ereignisse mit großem Datenvolumen nicht mehr ausgelesen zu werden brauchen. Dies vermindert die Totzeit des Triggers. Ein erheblicher Anteil kosmischer Myonen, die einen geringen Datenumfang besitzen, wird dann durch Triggerstufe L4 verworfen.

Es wurde weiterhin gezeigt, daß ein einzelnes in der Elektronik fehlerhaft permanent angeschaltetes *bit* großen Einfluß auf die Netzentscheidung haben kann. Dieses *bit* ist wichtig für die Berechung einer der Eingabegrößen der Neuronalen Netze. Es wurde untersucht, wie sich die Netzergebnisse ändern, wenn das hängende *bit* nicht mehr auftritt. Die folgenden Zahlenwerte entstammen einer Simulation, die mit einem systematischen Fehler behaftet ist (siehe Abschnitt 4.2.5). Das Netz zeigte eine Verschlechterung der Physikerkennung von absolut etwa 7%. Die Untergrunderkennung verbessert sich dabei um absolut etwa 2%. Bei mit anderer Parameterwahl trainierten Netzen traten für die Physikerkennung Verschlechterungen von bis zu absolut 19% auf, während sich die Untergrunderkennung um bis zu absolut 8% verbesserte.

Im Gegensatz zu Subtrigger s34 sind für die Subtrigger **s52** und **s54** schon Neuronale Netze im Trigger implementiert. Durch diese werden Reduktionsfaktoren von bis zu einem Faktor 4 bei s52 und 2.5 bei s54 erreicht. Die neu trainierten Netze arbeiten ausschließlich mit Größen aus den Spurkammern im Gegensatz zu den installierten Netzen, die eine Triggerentscheidung auch mit Informationen aus dem Flüssig-Argon-Kalorimeter und dem Myondetektor fällen. Bei Vorgabe der oben genannten Ratenreduktionsfaktoren erreichen die neuen Netze Effizienzen bei der Physikerkennung von 97% sowohl bei s52 als auch bei s54. Dies ist unter den am Ende der Abschnitte 4.3.3 und 4.4.3 genannten Vorbehalten zu vergleichen mit den Effizienzen der im Trigger installierten Neuronalen Netze von 73% für s52 und 99% für s54. Das trainierte Netz für s52 ist daher deutlich besser; das Netz für s54 ist nahezu gleich gut wie die installierte Konfiguration.

Weiterhin wurde mit Untergrund- und Physikereignissen von s52 und s54 zusammen ein Netz trainiert und getestet. Die Effizienzen der Physikerkennungen der einzeln trainierten Netze bei den oben genannten Ratenreduktionsfaktoren werden auch mit diesem gemeinsamen Netz nahezu erreicht. Lediglich die Effizienz der Erkennung der $J/\psi \rightarrow e^+e^-$-Ereignisse für s54 ist mit 95% um 2% niedriger. Unter dieser Einschränkung kann demnach an Stelle der einzelnen Netze auch dieses gemeinsam trainierte Netz am Experiment installiert werden.

Ausblick

Für **s34** wird durch Einbau des Neuronalen Netzes eine geringfügige Ratenreduktion bei gleichzeitiger Verminderung der Totzeit erreicht werden.

Der Einbau der Netze für **s52** und **s54** würde unter Vorbehalten (siehe am Ende der Abschnitte 4.3.3 und 4.4.3) bei gleicher Untergrunderkennung eine deutlich bessere (s52) bzw. nahezu gleich gute (s54) Effizienz der Physikerkennung liefern. Die Berechnung der Effizienz würde sich jedoch sehr erleichtern, da diese Netze nur mit Hilfe von aus den Spurkammern stammenden Größen trainiert wurden, so daß leichter zur Effizienzbestimmung notwendige, von diesen Triggern unabhängige andere Trigger aus der H1-Triggerkonfiguration gefunden werden können, als dies für die sich am Experiment befindlichen Netze der Fall ist. Die schon installierten Netze arbeiten auch mit Größen aus dem Flüssig-Argon-Kalorimeter und dem Myon-Detektor.

Man sollte zukünftig versuchen, die Größen *TRHINEG* und *TRTOT* nicht in das Training von Netzen mit einzubeziehen, um Probleme mit dem hängenden *bit* frühzeitig zu vermeiden.

Anhang A

Bedingungen für gute Spuren

In der folgenden Tabelle sind die genauen Bedingungen, die erfüllt sein müssen, damit in dieser Arbeit eine aus Detektordaten rekonstruierte Spur eine gute Spur genannt wird, gegeben [30]. Es wird dabei zwischen Spuren aus dem zentralen und dem vorderen Spurkammersystem unterschieden. Aus Daten beider Spurkammersysteme rekonstruierte, kombinierte Spuren werden nicht mit zusätzlichen Schnitten versehen.

Spuren aus dem	
zentralen Spurkammersystem	vorderen Spurkammersystem
$\vartheta > 20°$ $R_{start} < 44\,cm$ $R_{end} - R_{start} > 11\,cm\,(\vartheta < 155°)$ $R_{end} - R_{start} > 6\,cm\,(\vartheta \geq 155°)$ $P_t > 0.15\,GeV$ $dca < 10\,cm$	$\chi^2 \leq 50$ (aus der Vertexanpassung) $P \geq 1\,GeV$ (der an den Vertex angepaßten Spur) $P_t \geq 0.15\,GeV$ $r^0 \leq 10\,cm$ $\chi^2 \leq 10$ (aus der Spuranpassung) $N_{Module} \geq 2$ $(\vartheta < 20°)$

Tabelle A.1: *Bedingungen für gute Spuren.*

Erklärung der verwendeten Größen: R_{start} und R_{end} sind die radialen Abstände von der nominellen Strahlachse derjenigen Punkte aus der Menge der Punkte, die zur Spurrekonstruktion verwendet wurden, mit dem kleinsten bzw. größten radialen Abstand. Der dca ist der kleinste, radiale Abstand einer rekonstruierten und extrapolierten Spur zur nominellen Strahlachse. Die χ^2 enstammen χ^2-Tests aus verschiedenen Anpassungen. Die Größe $r°$ gibt den minimalen, radialen Abstand einer aus Daten der vorderen Spurkammern rekonstruierten und in die x-y-Ebene extrapolierten Spur zur nominellen Strahlachse an. N_{Module} ist die Anzahl der Module aus dem vorderen Spurkammersystem mit Signal für das jeweils betrachtete Ereignis.

Anhang B

Zahlenwerte für einige Abbildungen

a)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	78	74	60	53	70	56	63	60	62	69	73	82	89
7	39	51	54	55	55	61	64	59	62	71	71	84	82
6	39	51	54	56	56	59	68	65	65	66	74	81	81
5	39	51	49	54	57	58	65	63	67	65	75	81	80
4	40	48	49	51	49	56	61	62	61	65	66	69	72
3	39	48	44	50	46	52	55	51	56	54	55	57	61
2	40	45	38	45	45	48	51	49	49	51	51	56	57
1	41	43	40	41	42	44	49	44	47	49	45	56	56

Anzahl versteckter Knoten

b)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	78	75	56	45	43	46	65	60	44	47	46	51	55
7	45	48	43	40	43	45	55	54	52	46	44	50	50
6	45	49	44	40	48	45	45	42	54	49	42	43	52
5	45	49	50	43	55	45	52	41	48	52	45	56	57
4	45	48	44	44	45	36	44	51	43	46	48	50	51
3	45	50	42	49	49	45	43	51	47	45	49	45	50
2	45	46	40	45	51	50	43	50	55	50	51	48	51
1	45	46	46	44	52	47	46	50	52	54	54	53	50

Anzahl versteckter Knoten

Abbildung B.1: *Zahlenwerte für die Abbildungen* 4.6 *a) und b).*

a)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	84	86	89	88	90	86	89	89	90	91	93	95	95
7	84	86	86	89	90	90	91	91	91	91	93	94	93
6	84	87	87	89	88	89	89	91	90	91	93	93	93
5	84	86	88	88	88	88	90	90	91	91	92	93	93
4	84	88	87	88	88	89	88	88	89	89	91	89	90
3	84	86	87	87	88	88	89	88	88	88	89	89	89
2	85	85	85	87	88	87	88	87	87	87	89	89	89
1	85	86	85	86	87	85	87	86	87	87	88	88	89

Anzahl versteckter Knoten

b)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	87	86	83	87	87	90	89	89	87	85	78	80	81
7	87	92	92	87	89	87	89	87	85	86	89	84	83
6	87	93	93	87	92	91	91	89	87	88	88	82	85
5	87	93	91	87	92	91	91	89	86	90	90	83	89
4	87	84	93	89	91	92	91	93	93	90	91	93	93
3	87	89	91	89	89	91	88	92	91	91	92	91	90
2	87	86	88	88	88	91	88	89	91	89	90	90	90
1	87	86	88	88	87	90	88	89	90	89	91	89	90

Anzahl versteckter Knoten

Abbildung B.2: *Zahlenwerte für die Abbildungen* 4.13 *a) und b).*

a)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	8	19	15	10	15	6	3	11	9	9	10	1	9
7	8	7	7	8	14	9	16	9	10	17	15	15	17
6	8	7	14	6	10	15	6	18	13	5	20	7	16
5	8	7	7	5	11	15	13	10	13	18	19	9	18
4	9	9	8	14	8	5	11	11	7	8	15	17	17
3	9	14	8	13	14	4	13	13	5	13	6	7	4
2	11	11	11	11	14	12	14	14	11	14	7	7	2
1	13	11	11	9	9	13	11	13	12	11	11	7	8

Anzahl versteckter Knoten

b)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	3	5	5	3	6	4	3	6	3	3	3	0	6
7	4	2	2	3	4	6	3	4	3	5	4	4	5
6	4	2	2	3	2	5	1	4	5	2	4	5	7
5	4	2	3	3	3	3	1	3	3	2	3	3	5
4	5	4	2	5	3	2	1	1	2	3	5	5	3
3	4	5	3	5	6	4	6	2	2	4	1	2	2
2	5	5	7	5	7	5	6	6	3	7	3	2	3
1	4	5	4	5	5	5	6	6	4	6	3	3	2

Anzahl versteckter Knoten

Abbildung B.3: *Zahlenwerte für die Abbildungen* 4.21 *a) und b).*

a)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	79	81	82	82	83	87	88	86	88	86	89	95	93
7	79	81	82	82	84	86	84	84	86	86	87	88	89
6	79	81	81	81	82	85	83	84	83	85	85	87	86
5	79	82	80	81	82	84	82	84	83	84	84	86	86
4	79	81	82	82	83	86	84	84	85	86	87	87	89
3	79	80	81	80	80	83	82	82	81	81	83	84	84
2	79	80	80	80	79	81	80	80	80	80	81	82	82
1	79	79	80	80	79	81	79	80	80	80	81	82	83

Anzahl versteckter Knoten

b)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	83	83	81	84	83	84	85	82	85	81	82	78	81
7	83	83	82	84	84	86	84	86	85	85	84	84	84
6	83	83	82	84	83	85	84	84	85	83	84	84	87
5	83	83	82	84	83	85	84	84	84	83	85	84	86
4	83	83	82	83	84	87	84	84	85	85	86	84	86
3	83	84	82	84	86	84	84	83	84	83	84	84	85
2	83	84	83	86	84	84	84	84	85	84	84	84	85
1	83	83	84	84	84	84	84	84	84	85	84	85	85

Anzahl versteckter Knoten

Abbildung B.4: *Zahlenwerte für die Abbildungen* 4.24 *a) und b).*

a)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	66	70	72	74	75	77	79	78	79	81	86	86	88
7	66	69	72	77	77	78	77	79	78	80	83	83	85
6	66	69	74	75	77	75	76	79	77	78	81	81	83
5	65	69	73	72	76	73	75	78	75	76	79	80	81
4	66	68	73	76	77	76	76	80	77	78	82	81	82
3	66	69	69	71	72	71	71	74	73	72	76	77	76
2	66	69	70	70	71	68	70	72	72	72	74	75	75
1	66	69	70	68	68	68	69	71	70	70	71	74	74

Anzahl versteckter Knoten

b)

Parameterkonfiguration	1	2	3	4	5	6	7	8	9	10	20	40	64
8	67	67	69	70	70	71	70	70	72	70	73	74	71
7	67	65	71	72	72	74	72	74	73	73	72	74	73
6	67	65	70	71	72	73	71	71	72	71	72	73	73
5	67	66	71	69	70	72	70	71	72	70	71	72	73
4	67	66	70	70	70	73	72	74	71	71	71	73	72
3	68	67	69	69	69	70	67	70	71	69	71	70	72
2	67	68	69	69	69	70	66	69	70	69	71	71	70
1	68	68	68	70	68	70	66	69	69	69	70	70	70

Anzahl versteckter Knoten

Abbildung B.5: *Zahlenwerte für die Abbildungen* 4.29 *a) und b).*

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[1] H1-Kollaboration, *The H1 Detector at Hera*
Internal Report, DESY H1-96-01, März 1996

[2] H1-Kollaboration, *The tracking, calorimeter and muon detectors of the H1 experiment at HERA*
Nuclear Instruments Methods in Physics Research A 386 (1997) 348-396, 1997

[3] H1-Kollaboration, *Observation of Events at Very High Q^2 in ep Collisions at HERA*
Desy 97-24, Februar 1997

[4] H1-Kollaboration, *A Measurement of the Proton Structure Function $F_2(x,Q^2)$ at Low x and Low Q^2 at HERA*
Desy 97-042, März 1997

[5] B.Naroska, *Heavy Flavour Production in Lepton Proton Collisions at Hera*
Internal Note H1-01/97-510, Januar 1997

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

[7] S.Schiek, *Untersuchung der inelastischen Photoproduktion von J/ψ-Mesonen im H1-Detektor bei HERA*
II.Inst.Exp.Physik, Univ. Hamburg, Dissertation 1996

[8] G.Schmidt, *Untersuchung der diffraktiven Photoproduktion von J/ψ-Mesonen im H1-Detektor bei HERA*
II.Inst.Exp.Physik, Univ. Hamburg, Dissertation 1996

[9] J.R.Forshaw, *Theory of Vector Meson Production*
Preprint HEP-PH/9506373, 1995

[10] P.V.Landshoff, O.Nachtmann, *Vacuum structure and diffraction scattering*
Z.Phys.C 35, 409-416, 1987

[11] Particle Data Group, *Particles and Fields - Review of Particle Properties*
Physical Review D Vol.54 No.1, 1996

[12] J.Darnell, H.Lodish, D.Baltimore, *Molekulare Zellbiologie*
Walter de Gruyter, Berlin/New York, 1994

[13] W.H.Calvin, G.A.Ojemann, *Einsicht ins Gehirn*
Carl Hanser Verlag, München/Wien, 1995

[14] Rojas, *Theorie der neuronalen Netze - eine systematische Einführung*
Springer-Lehrbuch, 1993

[15] S.Mohrdiek, *Neuronaler Netzwerk-Trigger für Myonen im H1-Detektor bei HERA*
Universität Hamburg, August 1997

[16] J.H.Köhne et al., *Realization of a Second Level Neural Network Trigger for the H1 Experiment at HERA*
Abstract 074, AIHENP'96, Lausanne, 2.-6.September 1996

[17] J.H.Köhne et al., *Realization of a Second Level Neural Network Trigger for the H1 Experiment at HERA*
Internal Note H1-01/97-509, Januar 1997

[18] A.Gruber et al., *A Neural Network Architecture for the Second Level Trigger in the H1-Experiment at the Electron Proton Collider HERA*
Proceedings IEEE conference, New Orleans, Mai 1994

[19] J.Moeck, *Einsatz neuronaler Netze als intelligente Trigger im H1-Experiment*
Technische Universität München, März 1994

[20] P.Ribarics, *L2-L3 Background Rejector Based on the Encapsulation of Background Clusters.*
Internal Note H1-08/93-310, August 1993

[21] J.-C.Bizot et al., *Proposal for a topological level 2 trigger*
Internal note H1-06/91-181

[22] J.-C.Bizot et al., *Status of Simulation for a topological Level 2 trigger*
Internal note H1-02/92-212

[23] P.Werbos, *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*
Ph.D. Thesis, Harvard University, 1974

[24] R.J.Williams, *On the Use of Back-Propagation in Associative Reinforcement Learning*
In IEEE International Conference on Neural Networks (San Diego 1988), New York:IEEE.

[25] D.B.Parker, *Learning Logic*
Technical Report TR-47, Center of Computational Research in Economics and Management Science, Massachusetts Institute of Technology,
Cambridge, MA, 1985.

[26] D.E.Rumelhart, G.E.Hinton, R.J.Williams *Learning Internal Representation by Error Propagation*
Parallel Distibuted Processing I (chap.8, pp.318-362), Cambridge, MA: MIT Press, 1986

[27] Xiao-Hu Yu, Guo-An Chen, *Efficient Backpropagation Learning Using Optimal Learning Rate and Momentum*
Neural Networks, Vol.10 No.3, pp.517-527, April 1997

[28] C.Schittenkopf, G.Deco, W.Brauer, *Two Strategies to Avoid Overfitting in Feedforward Networks*
Neural Networks, Vol.10 No.3, pp.505-516, April 1997

[29] M.Kander, *Persönliche Mitteilung*

[30] P.Merkel, *Persönliche Mitteilung*

An dieser Stelle möchte ich mich bei allen bedanken, die mir im vergangenen Jahr durch fachliche Hilfestellung und persönliche Ermutigung zur Seite gestanden haben:

Frau Prof. Naroska danke ich für die interessante Aufgabenstellung, die gute Betreuung und Ihre verständnisvolle Art mit meinen zahlreichen Interessen umzugehen.

Herrn Prof. Büßer danke ich für die Übernahme der Zweitkorrektur dieser Arbeit.

Allen Mitgliedern der Myon-Gruppe und der Gruppe des FPS danke ich für ein stets offenes Ohr und die lösungsorientierte Zusammenarbeit.