Neuronaler Netzwerk-Trigger für Myonen im H1-Detektor bei HERA

Diplomarbeit

vorgelegt von

Susanne Mohrdieck

II. Institut für Experimentalphysik
Universität Hamburg

Hamburg
August 1997

Inhaltsverzeichnis

Kapitel 1

Einleitung

Die Elementarteilchenphysik beschäftigt sich mit der Frage, woraus die Materie aufgebaut ist und was sie zusammen hält. Die Ansichten über die „fundamentalen Teilchen“ der Welt und den zwischen ihnen wirkenden Kräften haben sich im Laufe der Zeit verändert und sind weiterentwickelt worden.
Erstmals wurde ein Modell des Aufbaus der Materie aus nicht weiter teilbaren Bausteinen (Atomen; von gr. atomos $\widehat{=}$ unteilbar) von Demokrit (ca.460-380 v.Chr.) aufgestellt.
Es zeigte sich aber in den Rutherfordschen Streuversuchen (1911 - 1913), daß die Atome weiter in eine aus Elektronen gebildete Hülle und einem positiven Kern zerlegt werden können. Zunächst stellte man sich vor, daß der Kern aus Protonen (1919 entdeckt) und Elektronen gebildet sein müßte. Dieses Bild führt aber zu Widersprüchen. So wurden weitere neutrale Teilchen im Kern gefordert, die Neutronen, die erstmals von Chadwick (1932) nachgewiesen wurden. Die drei Teilchen, Proton, Neutron und Elektron, bildeten von da an die „elementaren Teilchen“ der Materie. Eine theoretische Beschreibung dieses Bildes ist mit Hilfe der Quantentheorie möglich geworden. Als Kräfte zwischen den Teilchen beinhaltet das Modell die elektromagnetische Wechselwirkung und die Kernkraft.
Doch können mit diesem Bild der Materie nicht alle Prozesse in der Natur korrekt erklärt werden, so beispielsweise nicht die Radioaktivität. Wegen der Beschreibung des β-Zerfalls wurde die Einführung der schwachen Wechselwirkung und eines weiteren Teilchens, des Neutrinos, nötig.
Neben der Postulierung des Neutrinos (1956 nachgewiesen) wurden in Experimenten noch weitere Teilchen gefunden (Pionen, Myonen). Schließlich zeigte sich in Streuexperimenten, daß die Protonen und Neutronen selbst aus weiteren Partonen, den Quarks, aufgebaut sind, die im Nukleon durch die starke Wechselwirkung zusammengehalten werden.
Diese Entdeckungen bilden die Grundlage für das heute verwendete Standard-Modell: Die Materie ist aus sechs Leptonen $(e, \nu_e;\ \mu, \nu_\mu;\ \tau, \nu_\tau)$ und sechs Quarks $(u, d;\ c, s;\ t, b)$ aufgebaut. Das Modell beinhaltet zwei Kräfte, die elektroschwache und die starke Wechselwirkung. Die elektroschwache Wechselwirkung ist die Vereinheitlichung von elektromagnetischer und schwacher Wechselwirkung. Sie ist die einzige Wechselwirkung, an die die Leptonen koppeln. Neutrinos wechselwirken sogar nur über die schwache Komponente mit anderen Teilchen.

Mit Streuexperimenten an Speicherringen kann das Standard-Modell getestet werden.
Ein solcher Speicherring ist die Hadron-Elektron-Ring-Anlage HERA am Deutschen Elek-

tronen Synchrotron DESY in Hamburg. Hier werden Positronen und Protonen beschleunigt und bei einer Schwerpunktsenergie von $\sqrt{s} = 300$ GeV zur Kollision gebracht. Es werden mit zwei Experimenten der *ep*-Wechselwirkung (H1 und ZEUS) sowie einem Experiment mit dem Positronenstrahl (HERMES) die Struktur von Proton, Neutron und Photon erforscht. Ebenso wird mit der Anlage nach neuen Teilchen und Aspekten außerhalb des Standard-Modells gesucht (beispielsweise Supersymmetrie und Leptoquarks).

Diese Arbeit wurde am H1-Experiment erstellt. Ein Forschungsschwerpunkt des Experiments ist die Produktion der schweren Quarks *bottom* und *charm*. Schwere Quarks werden im Detektor z.B. durch den leptonischen Zerfall von J/ψ- und Υ-Mesonen oder über semileptonische Zerfälle von Hadronen mit *charm*- oder *bottom*-Quarks in Myonen („inklusive Myonen“) identifiziert.
Neben physikalisch interessanten Prozessen treten im Detektor mit deutlich höherer Rate auch Untergrundereignisse auf, die nicht aus *ep*-Wechselwirkungen stammen. Diese werden beim H1-Experiment über ein vierstufiges Triggersystem von den Physikereignissen getrennt. Die erste Stufe wendet eindimensionale Schnitte auf die vom Detektor bereitgestellten Informationen an. Solche Schnitte reichen jedoch nicht immer aus, um die Ereignisrate für ein Aufzeichnen der Daten genügend zu senken. Häufig wird die Rate nach der ersten Triggerstufe noch vom Untergrund dominiert. Hier wird die zweite Stufe aktiviert, die u.a. mit Hilfe neuronaler Netze den Untergrund weiter reduziert.
Neuronale Netze besitzen den Vorteil, daß sie in hochdimensionalen Räumen Korrelationen zwischen den Ereignissen finden können und über multidimensionale Schnitte in diesem Raum Ereignisse eines bestimmten physikalischen Prozesses vom Untergrund trennen können. Sie werden über Lernalgorithmen auf ein Klassifizierungsproblem optimiert, um dann unbekannte Ereignisse den verschiedenen Klassen zuzuordnen (Training und Generalisierung).

Das Ziel dieser Diplomarbeit ist es, für Ereignisse mit unelastischen J/ψ-Kandidaten und für Daten mit inklusiver Myon-Produktion neuronale Netze zu trainieren, die dann auf der zweiten Ebene des H1-Triggersystems eingesetzt werden können. Dabei stellt sich die Frage, ob ein gemeinsames Netz zur Trennung von Untergrund und den beiden physikalischen Ereignisgruppen effizient arbeiten kann oder ob der Einsatz zweier einzelner Netze erfolgreicher ist.

In dieser Arbeit werden zunächst in Kapitel 2 die wichtigsten Mechanismen zu Produktion und Zerfall von schweren Quarks beschrieben.
Daran schließt sich eine Darstellung des H1-Experiments an (Kapitel 3).
Die theoretischen und mathematischen Grundlagen neuronaler Netze, wie sie in der zweiten Triggerstufe bei H1 verwendet werden, werden in Kapitel 4 gegeben.
In den folgenden Kapiteln dieser Arbeit wird die Erstellung der Netze und ihr Einsatz in der Datennahme behandelt.
Gegenstand des Kapitels 5 ist die Selektion der für das Training benötigten physikalischen Datensätze und der Untergrundereignisse. In Kapitel 6 wird dann die Trainingskette vorgestellt, an deren Ende das Netz fertiggestellt ist. Außerdem wird auf die im Training und Test erreichten Resultate der Netze eingegangen. Am Ende des Kapitels werden die Effizienzen der Netze in der Datennahmeperiode 1997 besprochen.

Kapitel 2

Produktion und Zerfall schwerer Quarks bei HERA

Eines der Forschungsgebiete beim H1-Experiment am Speicherring HERA ist die Produktion schwerer Quarks Q (*bottom* b und *charm* c). Die Produktion kann über mehrere verschiedene Mechanismen erfolgen. Es entstehen dabei meistens schwere Quark-Antiquark-Paare, die in gebundene $Q\overline{Q}$-Zustände (J/ψ- und Υ-Mesonen) oder mit leichten Quarks in D- und B-Mesonen übergehen können.
Die schweren Quarks werden über ihre typischen Zerfallsprodukte identifiziert, insbesondere über die leptonischen und semileptonischen Zerfälle in Elektronen und Myonen.
In der vorliegenden Arbeit werden Ereignisse mit J/ψ-Mesonen im Zerfallskanal $\mu^+\mu^-$ und mit inklusiver Myon-Produktion verwendet. Inklusive Myonen sind Ereignisse mit mindestens einem Myon im Endzustand. Dieses kann z.B. aus dem Zerfall schwerer Quarks oder auch seltener aus dem eines W-Bosons hervorgehen.
In diesem Kapitel werden die wichtigsten Produktionsmechanismen für schwere Quarks sowie ihre Zerfälle beschrieben.

2.1 Produktion schwerer Quarks

Die ep-Wechselwirkung bei HERA läßt sich in zwei kinematische Bereiche unterteilen:

- der Bereich tiefunelastischer Streuung (DIS, **D**eep **I**nelastic **S**cattering) mit großem Viererimpulsübertrag Q^2.
- der Bereich der Photoproduktion mit $Q^2 \approx 0$.

Beide Bereiche liefern Beiträge zur Produktion schwerer Quarks durch den dominierenden Prozeß der Boson-Gluon-Fusion (BGF). In erster Ordnung (α_s, α) koppelt bei der Boson-Gluon-Fusion ein vom Positron ausgesandtes Boson (γ, Z^0, W^+) über ein Quark-Antiquark-Paar an ein Gluon, das aus dem Partoninhalt des Protons stammt. Dabei überwiegt in der *ep*-Wechselwirkung der Austausch eines quasireellen Photons ($Q^2 \approx 0$), d.h. es liegt Photoproduktion vor (Photon-Gluon-Fusion):

$$\gamma + g \rightarrow Q + \overline{Q} \tag{2.1}$$

Durch diese Tatsache ist die Entstehung eines schweren Quarks zusammen mit einem leichteren Quark, die nur über geladene Ströme mit $W^{\pm}$ möglich ist, gegenüber der Bildung eines schweren Quark-Antiquark-Paares stark unterdrückt. In Abbildung 2.1 sind die Feynmangraphen für die Photon-Gluon-Fusion gezeigt.

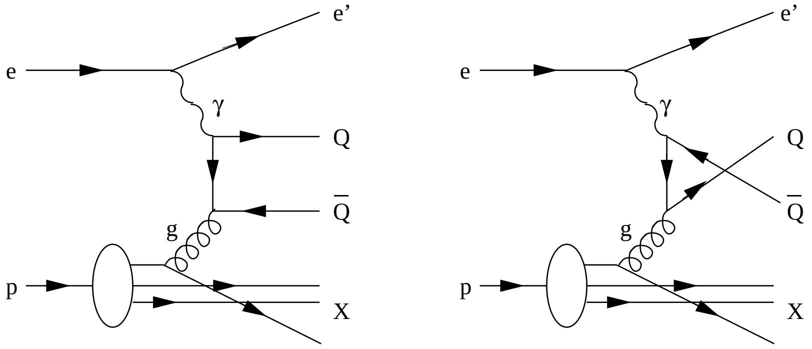


Abbildung 2.1: *Produktion schwerer Quarks mit Photon-Gluon-Fusion.*

Neben diesem Vorgang erster Ordnung können noch weitere Prozesse höherer Ordnung in (α_s, α) auftreten.
Zur theoretischen Berechnung des Wirkungsquerschnittes der Photon-Gluon-Fusion wird die *ep*-Wechselwirkung in zwei unabhängige Vorgänge unterteilt: die Emission eines Photons vom Positron und die Wechselwirkung des Photons mit dem Proton. Die Emission des Photons wird dabei durch die Weizsäcker-Williams-Näherung [Weiz34] beschrieben, die das elektromagnetische Feld eines sich schnell am Nukleon vorbei bewegenden Positrons als Fluß virtueller Photonen betrachtet. Hieraus wird für die $b\overline{b}$-Produktion ein Wirkungsquerschnitt von $\sigma(bottom) = 6.0 \pm 2$ nb abgeschätzt [Eich88, Frix95]. Für die $c\overline{c}$-Erzeugung wurde $\sigma(charm) = 941 \pm 160^{+142}_{-120}$ nb gemessen [Aid96].
Im folgenden werden die Prozesse höherer Ordnung aufgelistet (nach [Krüg94]):

- Prozeß mit aufgelöstem Photon: Bei Prozessen höherer Ordnung in α_s kann die hadronische Komponente des Photons aufgelöst werden. Es reagiert ein Gluon oder Quark (q) aus dem Photon mit einem Gluon oder Quark (q) des Protons. Es entsteht im Endzustand neben dem Protonrest und dem $Q\overline{Q}$-Paar ein vom Photon stammender hadronischer Rest (*photon remnant*).

$$q + \overline{q} \rightarrow Q + \overline{Q} \tag{2.2}$$
$$g + g \rightarrow Q + \overline{Q} \tag{2.3}$$

- Prozesse höherer Ordnung mit direktem Photon: Dieser Prozeß ähnelt dem in Abbildung 2.1 dargestellten Vorgang. Das Gluon strahlt, bevor es über ein $Q\overline{Q}$-Paar an das Photon koppelt, ein weiteres Gluon ab.

$$\gamma + g \rightarrow Q + \overline{Q} + g \tag{2.4}$$

- QCD-Compton-Prozesse: das Photon reagiert direkt mit einem leichten Quark (q) oder Antiquark ($\overline{q}$) aus dem Proton. Das betroffene Quark sendet ein Gluon aus, das in ein $Q\overline{Q}$-Paar übergeht.

$$\gamma + q(\overline{q}) \rightarrow Q + \overline{Q} + q(\overline{q}) \tag{2.5}$$

Die tiefunelastische Streuung kann außer durch die Boson-Gluon-Fusion auch durch die in Abbildung 2.2 dargestellten Prozesse zur Produktion schwerer Quarks beitragen.

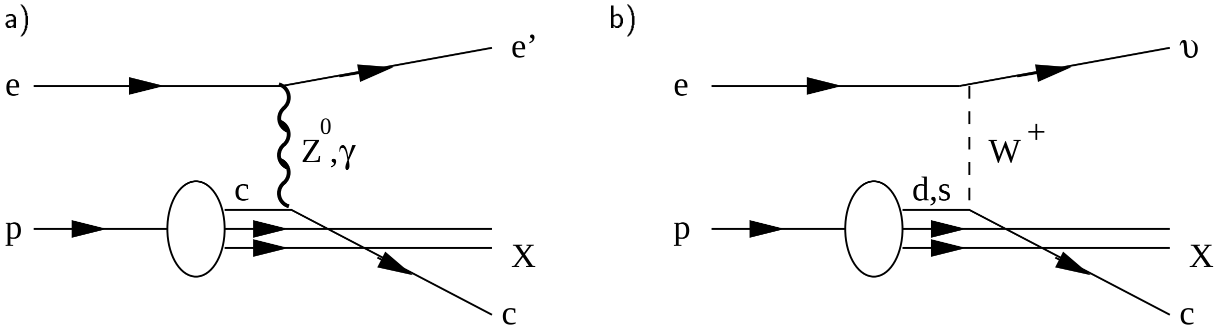


Abbildung 2.2: *Tiefunelastische Streuung mit Bildung schwerer Quarks (am Beispiel des c-Quarks): a) neutraler Strom mit Austausch der Bosonen γ, Z^0, b) geladener Strom mit Austausch eines W-Bosons.*

Dabei strahlt das Positron ein Eichboson (γ, Z^0, W^+) ab, das mit einem Quark des Protons in Wechselwirkung tritt. Es entsteht ein hadronischer Endzustand X.
Die tiefunelastische Streuung ist ein rein elektroschwacher Vorgang, der durch das Quark-Parton-Modell beschrieben wird. Er setzt sich aus einem Anteil mit geladenen Strömen und einem mit neutralen Strömen zusammen. Der Beitrag der neutralen Ströme zum Wirkungsquerschnitt ist jedoch aufgrund des geringen *charm*- und *bottom*-Anteils im Proton klein. Die Existenz von *charm* und *bottom* im Proton ist noch nicht nachgewiesen, wird aber vermutet. Mit einer Schwerpunktsenergie von $\sqrt{s} = 314$ GeV lassen sich als Wirkungsquerschnitte für tiefunelastische Streuung $\sigma(charm) = 4.1$ pb und $\sigma(bottom) = 0.12 \cdot 10^{-2}$ pb abschätzten [Ali87]. Diese sind gegenüber denen der Photoproduktion vernachlässigbar.

	Wirkungsquerschnitt [nb]	
	bottom	*charm*
Photoproduktion	6.0	941
DIS	$0.12 \cdot 10^{-5}$	$4.1 \cdot 10^{-3}$

Tabelle 2.1: *Produktionswirkungsquerschnitte*

2.2 Zerfall schwerer Quarks

Schwere Quarks zerfallen über schwache Wechselwirkung mit geladenen Strömen:

$$
\begin{aligned}
b &\rightarrow c + W^- && (2.6)\\
b &\rightarrow u + W^- && (2.7)\\
c &\rightarrow s + W^+ && (2.8)
\end{aligned}
$$

Die Zerfallsbreite für diese Übergänge ist proportional zum Quadrat des Matrixelements in der Cabbibo-Kobayashi-Maskawa-Matrix. Wegen eines sehr kleinen Matrixelements ist der Zerfall (2.7) stark unterdrückt.

Da die Massen der Quarks kleiner als die Masse des W-Bosons sind, sind die Bosonen virtuell. Sie zerfallen zum einen hadronisch unter Bildung eines $q\overline{q}$-Paares oder semileptonisch in ein Lepton und ein Neutrino (Abbildung 2.3).

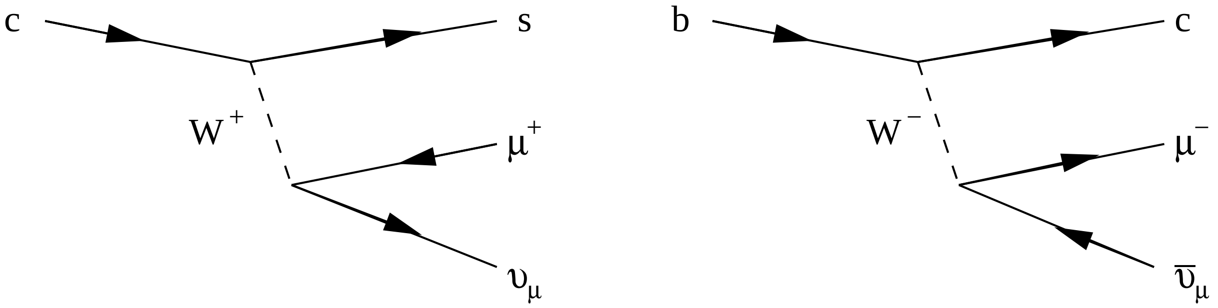


Abbildung 2.3: *Semileptonischer Zerfall schwerer Quarks in Myonen.*

Neben Myonen sind als Zerfallsprodukte Elektronen und Tauonen (nur für das b-Quark) möglich. Die Wahrscheinlichkeit der einzelnen Endzustände kann im Zuschauermodell berechnet werden. Dabei wird das schwere Quark als „frei“ unter Vernachlässigung der Wechselwirkung mit anderen Quarks und Gluonen betrachtet.

Die theoretisch berechneten Verzweigungsverhältnisse für die einzelnen semileptonischen Zerfälle sind in Tabelle 2.2 aufgelistet.

Quark	Verzweigungsverhältnis		
	$\rightarrow e$	$\rightarrow \mu$	$\rightarrow \tau$
b	0.11	0.11	0.026
c	0.11	0.11	0

Tabelle 2.2: *Theoretisch bestimmte Verzweigungsverhältnisse für den Zerfall von* b- *und* c-*Quarks in Leptonen [Eich88]*

2.3 Produktion und Zerfall von J/ψ-Mesonen

J/ψ-Mesonen sind gebundene $c\bar{c}$-Zustände. Man unterscheidet bei der Produktion in ep-Wechselwirkung zwei Entstehungsarten: die Diffraktion und die unelastische Erzeugung. Wenn nur Impuls übertragen wird und keine Quantenzahlen ausgetauscht werden, spricht man von Diffraktion. Sie unterteilt sich weiter in einen elastischen und einen protondissoziativen Anteil. Der elastische Prozeß zeichnet sich dadurch aus, daß im Detektor nur die Zerfallsprodukte des J/ψ-Mesons nachgewiesen werden.
J/ψ-Mesonen werden unelastisch vorwiegend in Photon-Gluon-Fusion erzeugt. Dabei ist zu beachten, daß bei der unter Abschnitt 2.1 beschriebenen Produktion das Quark-Antiquark-Paar in einem farbgeladenen Zustand vorliegt. Um zu den farblosen J/ψ-Mesonen zu gelangen, muß mindestens ein Gluon abgestrahlt werden (Abbildung 2.4 a)).

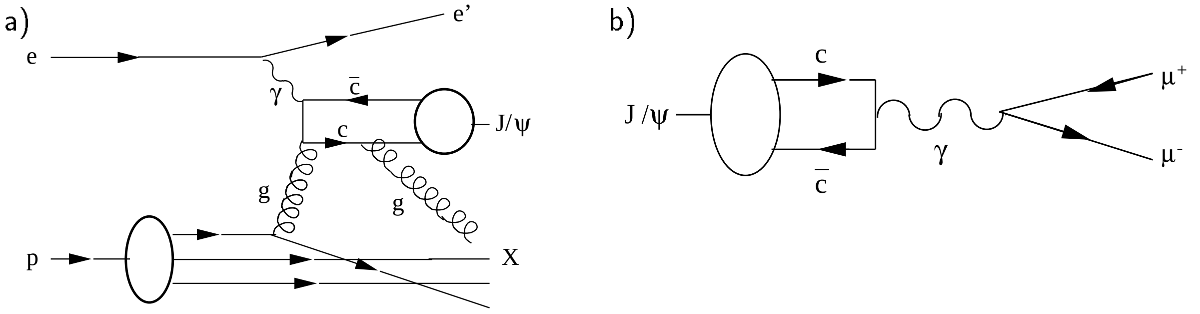


Abbildung 2.4: *a) Erzeugung eines J/ψ-Mesons, b) Zerfall eines J/ψ-Mesons in zwei Myonen.*

Neben der direkten Erzeugung von J/ψ-Mesonen über die Produktion eines schweren Quark-Antiquark-Paares mit anschließender Aussendung eines Gluons können J/ψ-Mesonen auch über die Fragmentation des Quark-Antiquark-Paares oder eines in der γp-Wechselwirkung gebildeten Gluons produziert werden. Größere Beiträge durch diesen Vorgang werden bei J/ψ-Mesonen mit großem Transversalimpuls erwartet.
Der Übergang des Quark-Antiquark-Paares in den farbneutralen gebundenen Zustand wird im Farb-Singulett-Modell (*color singlet model*) [Berg81, Baie81] beschrieben. Bei der Berechnung des Wirkungsquerschnittes wird für das $c\bar{c}$-System nur der farblose $^3S_1$-Zustand berücksichtigt.
Ein neueres Modell bezieht neben diesem Zustand auch farbgeladene Zustände und höhere Anregungen im Fock-Raum ein, die Farb-Oktett-Beiträge [Bodw95]. Über Abstrahlung weicher Gluonen erfolgt der Übergang aus diesen anderen Anregungen in den $^3S_1$-Zustand des J/ψ-Mesons. Nach theoretischen Voraussagen sind Prozesse dieser Art für die Erzeugung von J/ψ-Mesonen über aufgelöste Photonen von Bedeutung. Dadurch würde der Beitrag der Produktion mit aufgelösten Photonen um einen Faktor 10 steigen.

J/ψ-Mesonen zerfallen zu 87% in Hadronen und zu je 6% über elektromagnetische Wechselwirkung in ein Elektron und ein Positron bzw. zwei Myonen [PDG]. In dieser Arbeit wird der Zerfall in zwei Myonen zur Identifizierung verwendet (Abbildung 2.4 b)).

Kapitel 3

Das H1-Experiment

In diesem Kapitel wird ein Überblick über die **H**adron-**E**lektron-**R**ing-**A**nlage (HERA) und das Experiment H1 gegeben. Dabei werden bei der Beschreibung des H1-Detektors nur die für diese Arbeit relevanten Bestandteile genauer erläutert. Gegenstand des letzten Teils des Kapitels ist die Entstehung von Untergrundereignissen und ihre Reduzierung in der Datennahme.

3.1 Der Speicherring HERA

In dem 6336 m langen Speicherring HERA werden seit 1992 Elektronen (seit Juli 1994 Positronen) und Protonen beschleunigt, gespeichert und zur Kollision gebracht. Durch den Übergang von e^- zu e^+ erhöht sich die Lebensdauer des Leptonenstrahls von 3 auf 11 Stunden.
In einem System aus Vorbeschleunigern werden die Teilchen zunächst auf Energien von 12 GeV (e) bzw. 40 GeV (p) beschleunigt, bevor sie im HERA-Ring in zwei separaten Strahlrohren auf ihre Endenergien gebracht und gespeichert werden (Abbildung 3.1). Die Endenergie beträgt für die Positronen 27.5 GeV und für die Protonen 820 GeV. Damit ergibt sich eine Schwerpunktsenergie des ep-Systems von $\sqrt{s}$ =300 GeV, so daß ein maximaler Viererimpulsübertrag Q^2 von ungefähr 10^5 GeV^2 erreicht wird. Dieser Wert liegt zwei Größenordnungen über dem von Experimenten am ruhenden Proton (*fixed target*-Experimente).
Die von HERA gelieferte integrierte Luminosität für die letzten Jahre ist in Tabelle 3.1 zusammengefaßt. Positronen und Protonen werden in dem Speicherring simultan auf maximal je 220 mit Teilchenpaketen (*bunches*) besetzten Plätzen gespeichert. Dies führt an

	1995	1996	1997
integrierte Luminosität	5 pb^{-1}	10 pb^{-1}	>15 pb^{-1}

Tabelle 3.1: *Integrierte Luminositäten der letzten Jahre.*

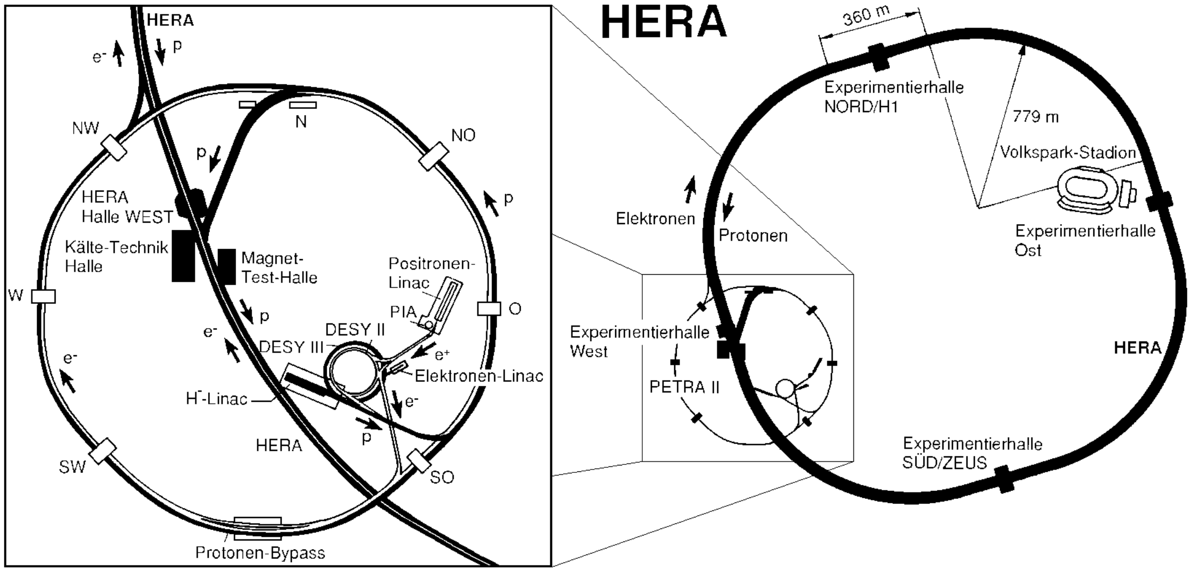


Abbildung 3.1: *Der Speichering HERA und seine Vorbeschleuniger.*

den Wechselwirkungspunkten zu Abständen von 96 ns zwischen zwei Kollisionen von Paketen (*bunch crossings*) entsprechend einer Umlauffrequenz von 10.4 MHz. Die Zeit zwischen zwei Kollisionen von Paketen wird auch als Zeiteinheit verwendet, sie wird im folgenden mit 'Zeitscheibe' bezeichnet. Beim Auffüllen der Plätze mit Paketen wird eine bestimmte Anzahl von Plätzen nur mit Positronen oder Protonen gefüllt, so daß sie nicht miteinander kollidieren. Mit diesen Paketen kann der nicht aus ep-Wechselwirkungen stammende Untergrund untersucht werden. Die wichtigsten von Protonen induzierte Untergrundquellen sind Kollisionen mit dem Strahlrohr und Wechselwirkungen mit dem darin befindlichen Restgas. Für die Positronen ist die Synchrotronstrahlung anzuführen.

Die beiden Teilchenstrahlen werden in zwei Wechselwirkungszonen zur Kollision gebracht, an denen die Detektoren der Experimente H1 und ZEUS aufgebaut sind. Neben diesen beiden Experimenten arbeiten zwei *fixed target*-Experimente am Speicherring HERA. HERMES untersucht unter Verwendung longitudinal polarisierter Positronen und polarisierter Protonen und Neutronen die Spinstrukturen von Proton und Neutron. Noch im Aufbau befindlich ist das Experiment HERA B, mit dem durch ein nahe an den Protonenstrahl herangeführten Draht B-Mesonen mit hoher Rate erzeugt werden sollen und die CP-Verletzung beim Zerfall der B-Mesonen untersucht werden soll.

3.2 Der H1-Detektor

Der Universaldetektor des H1-Experiments ist in der Experimentierhalle Nord des Speicherrings HERA aufgebaut. Der Detektor umfaßt nahezu den gesamten 4π-Raumwinkel um den nominellen Wechselwirkungspunkt und ermöglicht bei guter Orts- und Impulsauflösung eine Identifizierung der Teilchen und die Messung ihrer Energie.

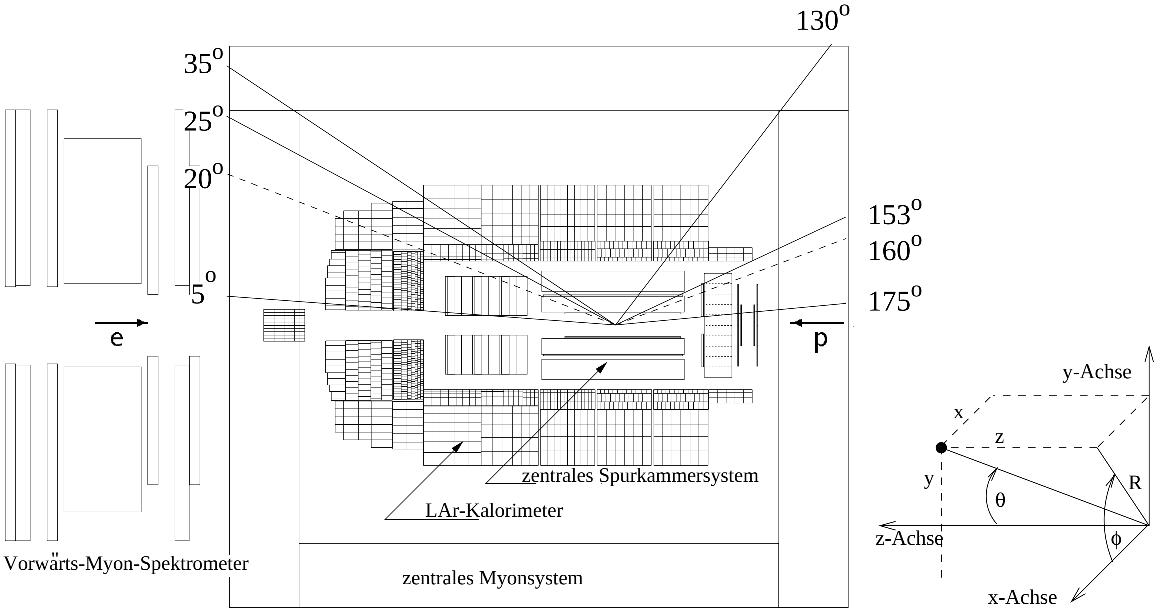
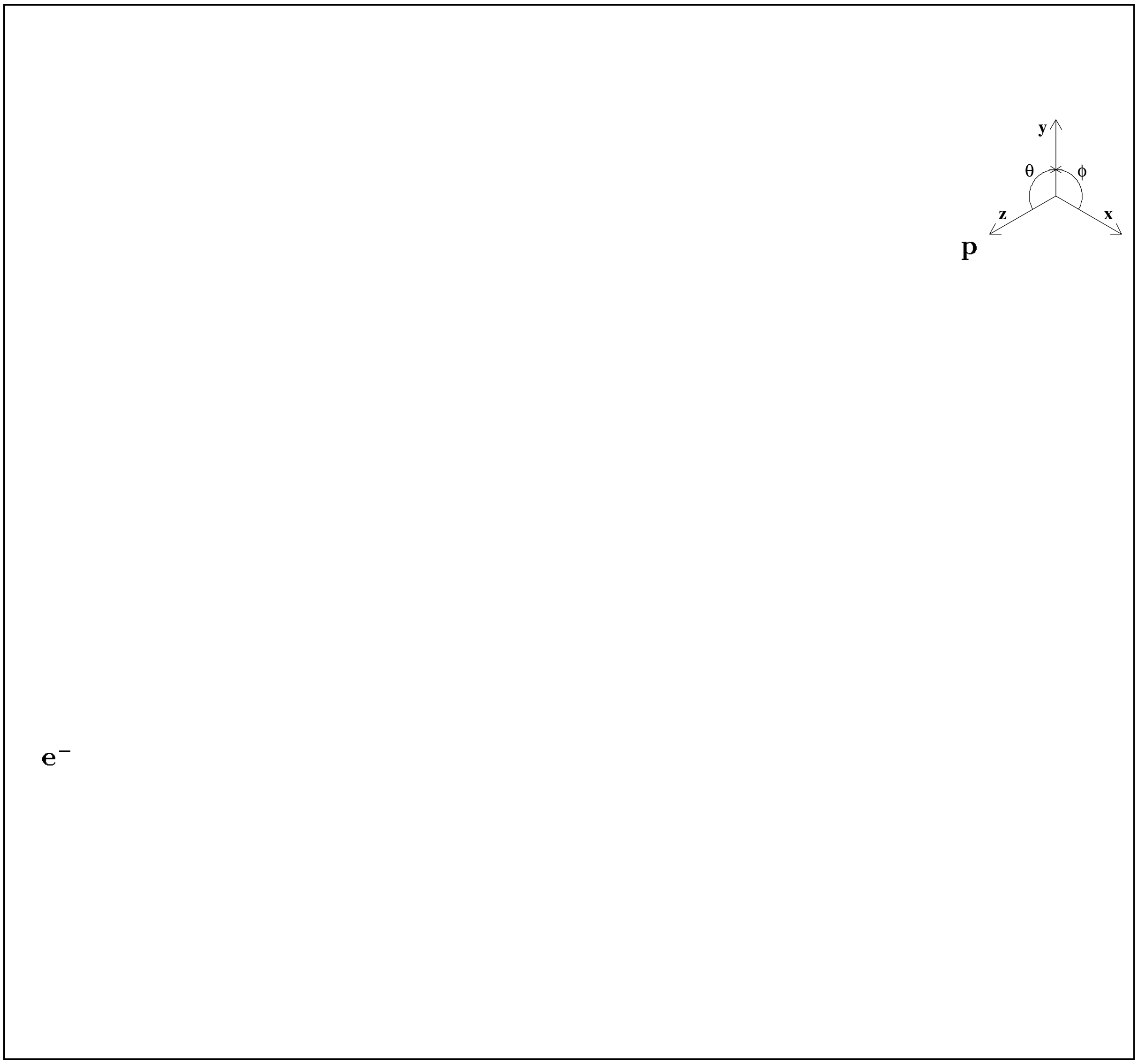


Abbildung 3.2: *Polarwinkelbereiche des H1-Detektors: Eingezeichnet sind die Akzeptanzbereiche des zentralen Myonsystems, des Flüssig-Argon-Kalorimeters sowie des zentralen Spurkammersystems. Rechts ist das H1-Koordinatensystem dargestellt.*

Das bei H1 verwendete Koordinatensystem ist ein rechtshändiges Koordinatensystem mit Ursprung im nominellen Wechselwirkungspunkt. Die z-Achse ist durch die Richtung des einlaufenden Protons definiert, die x-Achse zeigt zum Mittelpunkt des Speicherrings, y weist nach oben (Abbildung 3.2). Der Bereich kleiner Polarwinkel θ wird mit Vorwärtsrichtung bezeichnet.
Man erwartet, daß die Ereignisse eine Symmetrie im Azimutwinkel ϕ aufweisen und daß wegen des Lorentz-Boosts ein Großteil der Reaktionsprodukte in Vorwärtsrichtung geht. Deshalb ist der Detektor im vorderen Bereich aufwendiger gestaltet. Außerdem sind die Detektorkomponenten zylindersymmetrisch um das Strahlrohr angeordnet.
In Abbildung 3.3 ist ein Überblick über den Detektor zu sehen, Abbildung 3.2 zeigt die Polarwinkelbereiche einzelner Detektorkomponenten. Eine detaillierte Beschreibung der Detektorbestandteile befindet sich in [H1-96-01, H1-97]. Hier werden nur die wichtigsten Komponenten vorgestellt.

Das Spurkammersystem

Das Spurkammersystem besteht aus Driftkammern zur präzisen Spurrekonstruktion und Impulsmessung geladener Teilchen sowie Proportionalkammern, die nur grobe Information über Spurverlauf und Ereignisvertex liefern, aber aufgrund ihrer größeren Geschwindigkeit zu Triggerzwecken verwendet werden. Es setzt sich aus einem zentralen Spurdetektor (CTD, **C**entral **T**racking **D**etector), einem Vorwärtsspurdetektor (FTD, **F**orward **T**racking **D**etector) und einem rückwärtigen Driftkammersystem zusammen. Spuren, die im zentralen Spurdetektor rekonstruiert werden, heißen zentrale Spuren, die im Vorwärtsdetektor rekonstruierten Vorwärtsspuren. Werden für ein Teilchen in beiden Detektoren



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

Abbildung 3.3: *Der H1–Detektor.*

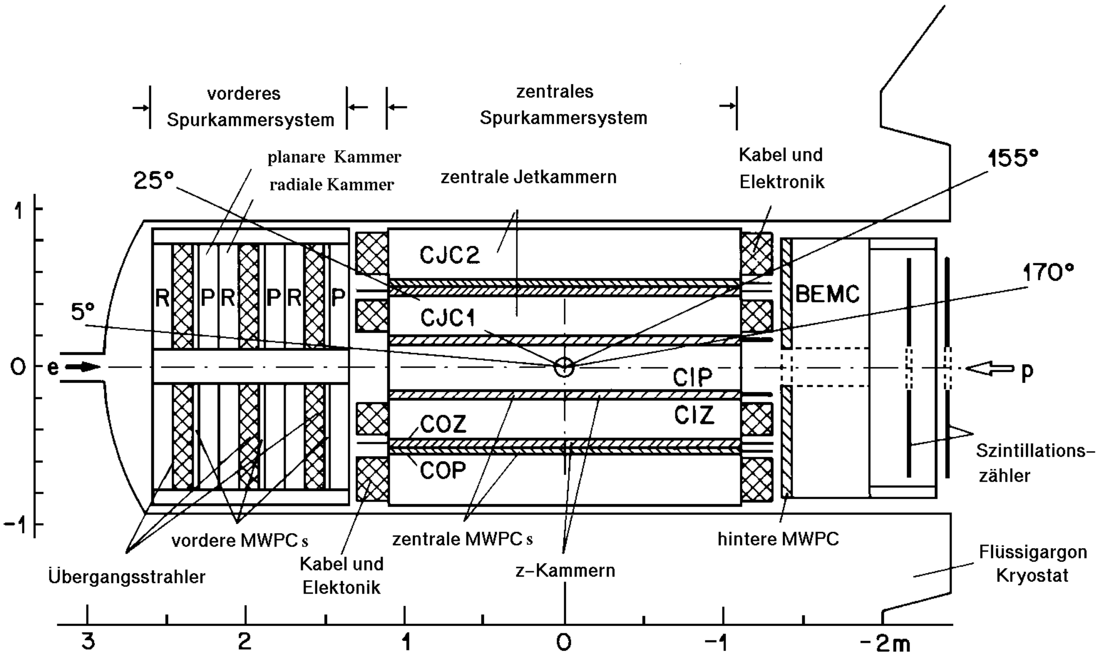


Abbildung 3.4: *Seitenansicht des Spurkammersystems.*

Spuren gefunden und können diese miteinander verbunden werden, so spricht man von kombinierten Spuren. Das Magnetfeld von 1.2 Tesla wird von einer supraleitenden Spule erzeugt, die das Kalorimeter umgibt.

Der zentrale Spurkammerdetektor deckt einen Winkelbereich von $20^\circ \lesssim \theta \lesssim 160^\circ$ ab. Er setzt sich aus zwei Jetkammern CJC1 und CJC2 (**C**entral **J**et **C**hamber), zwei weiteren Driftkammern CIZ und COZ (**C**entral **I**nner/**O**uter $\boldsymbol{z}$-Chambers) und den beiden Proportionalkammern CIP und COP (**C**entral **I**nner/**O**uter **P**roportionalchamber) zusammen (Abbildung 3.4). Die Jetkammern bestehen aus 2640 parallel zur Strahlachse gespannten Drähten (in 24 bzw. 32 Lagen), die in 30 bzw. 60 Driftzellen angeordnet sind. Geladene Teilchen hinterlassen beim Durchqueren des Detektors in diesen Zellen Treffer. In den Zellen werden Spursegmente rekonstruiert. Sind diese Parameter für zwei benachbarte Zellen ähnlich, so werden sie zu einer Spur zusammengefaßt. Danach werden Verbindungen zwischen CJC1 und CJC2 und den Spuren im Vorwärtsspurdetektor gesucht. Bei der Spurrekonstruktion treten folgende Ambiguitäten auf:

- Wenn ein Teilchen in der zentralen und der vorderen Spurkammer eine Spur hinterläßt, erhält es eine zentrale, eine vordere und eine kombinierte Spur (wenn die zentrale und vordere Spur in Einklang zu bringen sind).

- Können die Spurteile aus der CJC1- und der CJC2-Spurkammer aufgrund eines zu großen Winkels zwischen den Spursegmenten nicht miteinander verbunden werden, so bestehen beide Spuren für das Teilchen weiter.

- Durch schlechte Rekonstruktion können in den Spurkammern aus den angesprochenen Drähten mehrere Spurhypothesen für ein Teilchen aufgestellt werden.

Mit den Jetkammern wird eine gute Auflösung in der $r\phi$-Ebene erreicht, während die Genauigkeit in z sehr viel geringer ist. Eine höhere Präzision wird mit Hilfe der beiden z-Kammern CIZ und COZ erreicht.
Die Proportionalkammern CIP und COP bestehen jeweils aus einer Doppelkammer mit einer Drahtlage in jeder Kammer. Die Auslese dieser Kammern wird mit rechteckigen Elektroden (*pads*) durchgeführt. Für die innen liegende Kammer CIP werden 480 *pads* verwendet, wobei eine 60-fache Unterteilung in z und eine 8-fache Segmentierung in ϕ vorliegt. COP ist in z 18 mal und in ϕ 16 mal geteilt, so daß 288 *pads* ausgelesen werden.

In dem Bereich von $5^\circ \lesssim \theta \lesssim 25^\circ$ dient der Vorwärtsspurdetektor zur Spurrekonstruktion. Er besteht aus drei Supermodulen, die sich jeweils aus drei gegeneinander um 60° gedrehten planaren Driftkammern, einer Proportionalkammer, einer Übergangsstrahlungskammer und einer radialen Driftkammer zusammensetzen. Mit den planaren Kammern werden die x- und y-Koordinaten bestimmt, die Übergangsstrahlungskammer in Kombination mit der radialen Kammer trägt zur Teilchenidentifizierung durch Übergangsstrahlung bei. Die radiale Kammer liefert einen groben Wert für die radiale Koordinate sowie eine sehr genaue Messung von ϕ.
Die erste Proportionalkammer dient zusammen mit den beiden Proportionalkammern des zentralen Spurkammersystems als Grundlage für den z-Vertex-Trigger. Sie besteht aus zwei Drahtlagen und drei Kathodenebenen. Auf die beiden äußeren Ebenen sind zur Auslese der Kammer 384 *pads* aufgesetzt. Sie sind radial in 20 Ringe gegliedert. In ϕ liegt für die inneren 16 Ringe eine 8-fache Segmentierung vor, die äußeren vier Ringe sind 16 mal unterteilt.

Vervollständigt wird das Spurkammersystem durch acht im rückwärtigen Bereich des Detektors befindliche Driftkammerlagen. In ihnen werden vorwiegend die Spuren der wenig gestreuten Positronen ($155^\circ \leq \theta \leq 175^\circ$) nachgewiesen.

Das Kalorimeter

Das Spurkammersystem ist von dem Flüssig-Argon-Kalorimeter (LAr-Kalorimeter) umgeben, das einen Winkelbereich von 4° bis 153° abdeckt. Es ist in z-Richtung in acht konzentrische Räder unterteilt, die wiederum in ϕ in acht Oktanten gegliedert sind.
Das LAr-Kalorimeter setzt sich aus einem elektromagnetischen Kalorimeter mit einer Tiefe von 20 Strahlungslängen (innen) und einem hadronischen Anteil mit einer Tiefe von 4-8 hadronischen Wechselwirkungslängen (außen) zusammen. Es dient zur Energieflußmessung der Ereignisse und zur Identifizierung von Teilchen. In dieser Arbeit wird es zur Myonerkennung eingesetzt.
Die im elektromagnetischen und im hadronischen Teil gefundenen Energien dienen als Grundlage für den Kalorimetertrigger. Er kann nicht durch minimalionisierende Teilchen, also auch nicht durch Myonen, ausgelöst werden.

Neben dem LAr-Kalorimeter sei hier noch das Spaghetti-Kalorimeter (SpaCal) erwähnt. Es befindet sich im rückwärtigen Teil des Detektors und weist das gestreute Positron bei großen Winkeln ($155^\circ \lesssim \theta \lesssim 178^\circ$) nach. In der etwas älteren Abbildung 3.3 ist an seiner Stelle ein elektromagnetisches Kalorimeter (BEMC) eingezeichnet, das 1995 gegen das SpaCal ausgetauscht wurde.

Das instrumentierte Eisen

Neben der Rückführung des magnetischen Flusses der supraleitenden Spule wird das Eisenjoch zum Nachweis minimalionisierender Teilchen und zur Energiemessung verwendet. Es ist in vier Teile unterteilt: den vorderen und hinteren Zentralbereich (*Barrel*, $35^\circ \lesssim \theta \lesssim 130^\circ$), die hintere Endkappe (*Backward Endcap* BEC, $130^\circ \lesssim \theta \lesssim 171^\circ$) und die vordere Endkappe (*Forward Endcap* FEC, $4^\circ \lesssim \theta \lesssim 35^\circ$). Das instrumentierte Eisen ist im Zentralbereich in ϕ und z segmentiert, in den Endkappen in x und y. Dies führt zu 32 Modulen im Zentralbereich und 16 Modulen in jeder Endkappe.

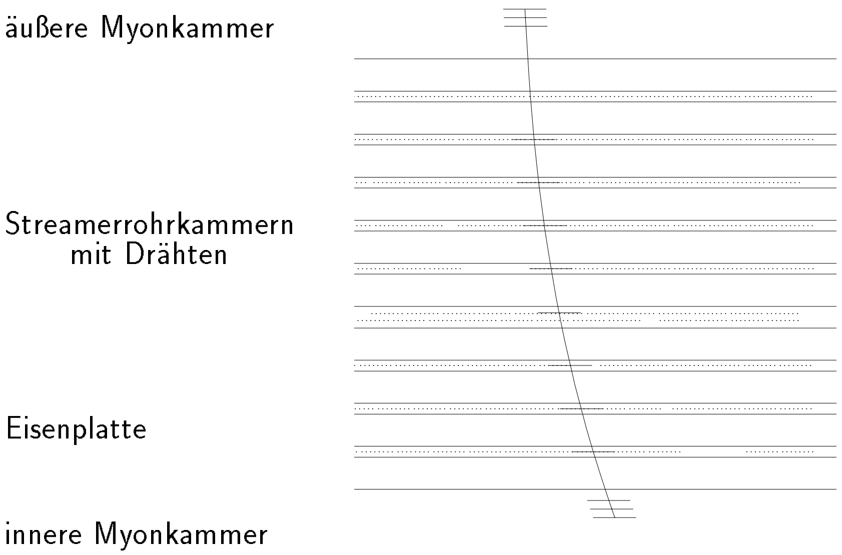


Abbildung 3.5: *Struktur des instrumentierten Eisens. Das hier gezeigte Myon hat in 14 von den 16 Lagen ein Signal ausgelöst (angedeutet durch die kurzen Querbalken).*

Um den Nachweis von Myonen zu ermöglichen, sind in jedem Modul zwischen 7.5 cm dicken Eisenplatten Streamerrohrkammern eingebaut worden, und zwar im Zentralbereich mit den Drähten parallel zur z-Richtung, in den Endkappen parallel zur x-Achse. Es entstehen so 10 Eisen-Streamerrohr-Lagen (vgl. Abbildung 3.5). Eine dieser Lagen ist mit zwei Streamerrohrkammern versehen. Zusätzlich zu diesen 10 Lagen sind an der inneren und äußeren Seite des Jochs die sogenannten Myonkammern mit jeweils drei weiteren Steamerrohrkammern angebracht. So werden von einem Teilchen bis zu 16 Lagen durchquert.
Die Streamerrohrkammern haben einen Querschnitt von 1cm × 1cm. Acht der Kammern werden jeweils zu einem Profil zusammengefaßt, zwei Profile zu einem Element. Auf diese

Elemente sind in fünf Lagen senkrecht zu den Drähten verlaufende Streifen-Kathoden befestigt, wodurch eine Spurrekonstruktion in drei Dimensionen möglich ist. In den restlichen Lagen werden rechteckige Kathoden (*pads*) verwendet, die zur Vervollständigung der Kalorimeterinformation dienen.

Nach vorne auslaufende Myonen werden im H1-Experiment mit dem Vorwärts-Myon-Spektrometer identifiziert. Da so erkannte Myonen in dieser Arbeit nicht berücksichtigt worden sind, wird hier von einer weiteren Beschreibung abgesehen.

3.3 Untergrundquellen

Neben den physikalisch interessanten ep-Reaktionen treten im Detektor Untergrundereignisse mit weitaus größerer Rate auf (vgl. Tabelle 3.2). Vier Quellen des Untergrundes werden im folgenden beschrieben.

Untergrund	Strahl-Gas-Wechselwirkungen	50 kHz
	kosmische Myonen im *Barrel*	700 Hz
Physikprozesse	$c\bar{c}$-Produktion	15 Hz
	DIS mit niedrigem Q^2	2.2 Hz
	DIS mit hohem Q^2	1.4 min^{-1}
	Produktion von W-Bosonen	0.5 d^{-1}

Tabelle 3.2: *Raten für verschiedene Prozesse bei einer Luminosität von* $1.5 \cdot 10^{31}$ cm^2 s^{-1} *[H1-96-01].*

Strahl-Gas-Reaktionen

Als Strahl-Gas-Reaktionen werden Wechselwirkungen der Strahlprotonen mit dem im Strahlrohr verbliebenen Restgas bezeichnet. Sie können überwiegend durch einen entlang z vom nominellen Vertex verschobenen Ereignisvertex von ep-Wechselwirkungen unterschieden werden. Durch die hohe Energie der Protonen sind die Spuren der Reaktionsprodukte sehr stark nach vorne gerichtet.

Strahl-Wand-Wechselwirkungen

Andere von Protonen induzierte Untergrundereignisse sind die Strahl-Wand-Wechselwirkungen. Sie umfassen Reaktionen von Haloprotonen mit den Strahlführungselementen, insbesondere mit der Strahlrohrwand. Die Wechselwirkungen, die ein Detektorsignal auslösen, finden weit vor dem Detektor statt. Die entstehenden Spuren treffen den Detektor unter kleinen Winkeln. Aus ihnen konstruiert sich ein Ereignisvertex, der in x und y weit von der Strahlmitte entfernt liegt. Die Rate dieser Reaktionen beträgt 40-250 kHz [MeyA95].

Kosmische Strahlung

Kosmische Strahlung besteht beim Erreichen der Erde u.a. aus Myonen, die den Detektor mit einer Rate von einigen kHz treffen. Durchqueren sie die Wechselwirkungszone, so ähnelt die hinterlassene Signatur den Zerfällen schwerer Quarks, beispielsweise elastisch erzeugter J/ψ-Mesonen im Zerfallskanal $\mu^+\mu^-$.

Synchrotronstrahlung

Von beschleunigten geladenen Teilchen werden Photonen ausgesandt. Im Falle eines Kreisbeschleunigers spricht man von Synchrotronstrahlung. Die ausgesandte Strahlung hängt mit m^{-4} von der Masse der beschleunigten Teilchen ab. Sie ist für die Protonen irrelevant. Die von den Positronen ausgesandte primäre Strahlung wird vollständig durch Kollimatoren aus dem Detektor ferngehalten, die sekundäre nur teilweise.

3.4 Das H1-Triggersystem

Die oben beschriebenen Untergrundreaktionen müssen im Experiment von den physikalisch interessanten Ereignissen aus *ep*-Wechselwirkungen getrennt werden, da bei der hohen Rate von Strahlkreuzungen (ca. 10MHz) nicht die gesamte vom Detektor kommende Ereignismenge gespeichert werden kann. Zur Trennung von Physik und Untergrund wurde beim H1-Experiment ein vierstufiges Triggersystem installiert (Abbildung 3.6). Die Stufen, mit L1 bis L4 (L steht für *level*) bezeichnet, werden im folgenden vorgestellt.

3.4.1 Erste Triggerstufe (L1)

Die Kollisionsrate zwischen Proton- und Positronpaketen beträgt etwa 10 MHz. Diese hohe Eingangsrate, die sehr viel Untergrund enthält, muß vom H1-Triggersystem auf eine Rate von etwa 10 Hz reduziert werden. Dies ist die maximale Rate, die auf Band geschrieben werden kann.

Die erste Stufe des Systems senkt auf Grundlage von 192 Triggerelementen die Eingangsrate auf unter 5 kHz. Die Triggerelemente werden in der Hardware gebildet und sind in neun Bereiche unterteilt, die jeweils auf den verschiedenen Subsystemen des Detektors basieren. Sie werden durch logische Verknüpfungen in der zentralen Triggerlogik (CTL, **C**entral **T**rigger **L**ogic) zu 128 sogenannten Subtriggern kombiniert. Sie werden mit Sx abgekürzt ($x = 0, \ldots, 127$). Das Akzeptieren durch einen dieser 128 Subtrigger ist hinreichend, um ein L1*keep* auszulösen, d.h. das Ereignis wird von den nachfolgenden Triggerstufen weiter geprüft.

Jeder der Subtrigger hat eine spezifische Rate. Ist diese Rate zu hoch, so wird der Subtrigger mit einem Untersetzungsfaktor (*prescale*-Faktor) f versehen. Das bedeutet, daß von den Ereignissen, die der Subtrigger akzeptiert hat, nur jedes f-te weiter untersucht wird.

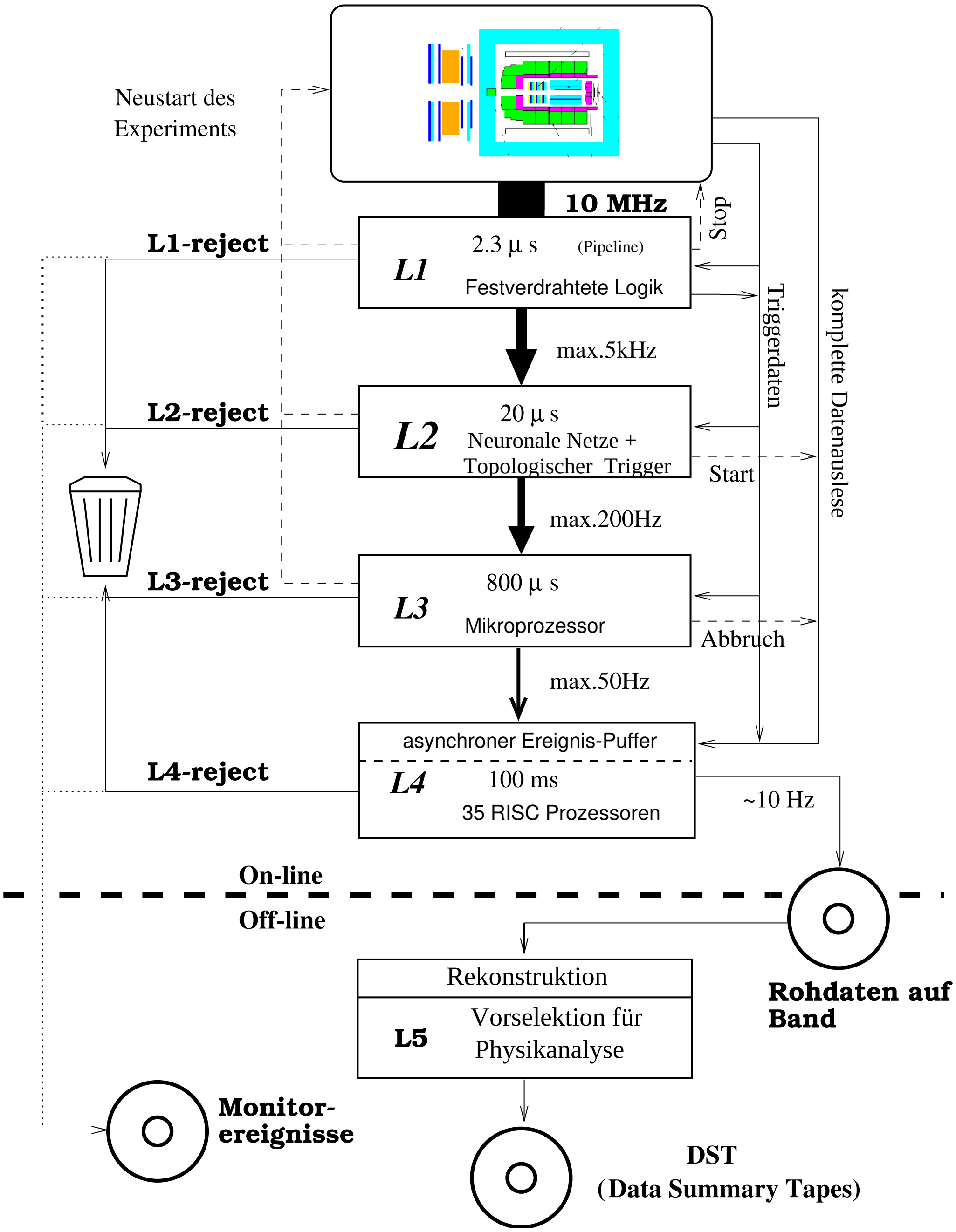


Abbildung 3.6: *Die verschiedenen Stufen des H1-Triggersystems: Das Triggersystem dient der Ratenreduktion im Experiment durch Verwerfen von Untergrund (reject) und Akzeptieren physikalisch interessanter Ereignisse (keep). Die einzelnen Stufen werden nacheinander durchlaufen. Nur wenn ein Ereignis von allen Stufen akzeptiert wird, wird es als „Physik“ auf Band gespeichert. Ein Bruchteil (für L4 1%) der verworfenen Untergrundereignisse wird für Monitorzwecke ebenfalls gespeichert (weitere Erläuterungen im Text).*

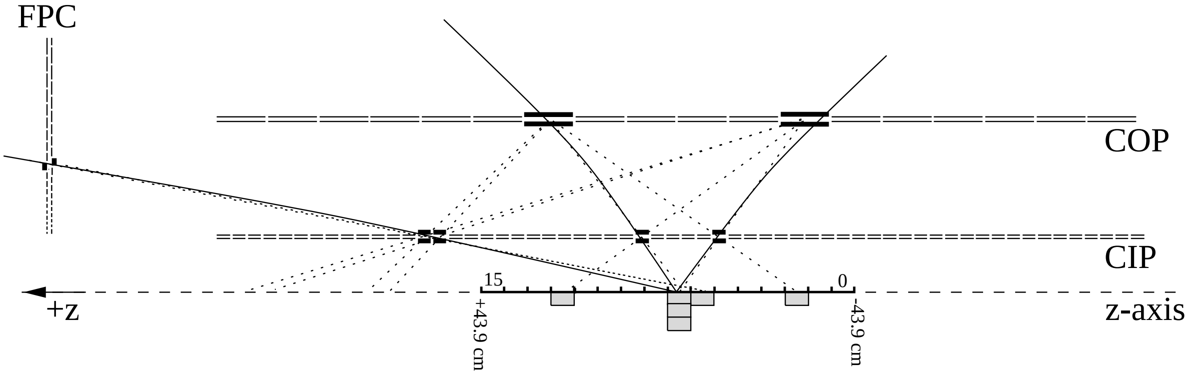


Abbildung 3.7: *Entstehung des z-Vertex-Histogramms: Vier angesprochene* pads *(dick) werden zu Strahlen (*rays*) kombiniert (gestrichelte Linien). Die* rays *treffen die in Abschnitte unterteilte Wechselwirkungszone. In den Abschnitten hinterlassen sie einen Eintrag. Die Gesamtzahlen der Einträge in jedem Abschnitt ergeben das z-Vertex-Histogramm. Man erkennt weiter, daß wirkliche Spuren (durchgezogene Linien) beschreibende* rays *alle auf einen Abschnitt zeigen, während „schlechte“* rays *in anderen Abschnitten enden oder keine Einträge hinterlassen.*

Die anderen $f-1$ werden verworfen, solange sie nicht noch von einem weiteren Subtrigger als „Physik“ erkannt worden sind.
Für die Zeit von 24 Zeitscheiben werden die vom Detektor kommenden Daten in einem Schieberegister (*pipeline*) zwischengespeichert. So hat die erste Triggerstufe des H1-Triggersystems für ihre Entscheidung pro Ereignis 2.3 μs Zeit und läuft dennoch totzeitfrei. Bei erfolgtem L1*keep*-Signal wird das Schieberegister gestoppt und die Totzeit des Systems beginnt.
Drei für diese Arbeit wichtige L1-Triggersubsysteme, der z-Vertex-Trigger, der $DCR\phi$-Trigger und der Myontrigger, werden im folgenden erläutert.

3.4.1.1 z-Vertex-Trigger

Der z-Vertex-Trigger [H1-96-01, Beck96] basiert auf den von den zentralen Proportionalkammern CIP und COP und der ersten vorderen Proportionalkammer (FPC, **F**orward **P**roportional **C**hamber) (Abschnitt 3.2) gelieferten Informationen. Ein den Detektor durchquerendes Teilchen trifft vier Lagen in diesen Kammern, jeweils zwei in CIP und COP oder je zwei in CIP und FPC. Dies führt zu vier *pad*-Signalen, aus denen durch geradlinige Strahlen sogenannte *rays* gebildet werden (Abbildung 3.7). Dabei wird eine 16-fache Segmentierung in ϕ verwendet. Es können maximal 34400 *rays* berechnet werden, aktive *rays* sind dabei wirklich im Ereignis gefundene *pad*-Kombinationen.
Der 87.84 cm große Bereich um den nominellen Wechselwirkungspunkt wird in 16 gleichlange Abschnitte entlang der z-Achse unterteilt. Alle aktiven *rays*, die die Wechselwirkungszone treffen, geben in diesen Abschnitten einen Eintrag. So entsteht das z-Vertex-Histogramm. Der Abschnitt mit den meisten Einträgen wird als Ereignisvertex für den Trigger angenommen.

„Gute“ *rays*, die einer von einer ep-Wechselwirkung stammenden Spur entsprechen, schneiden alle im gleichen Abschnitt die z-Achse. So entsteht ein Maximum über den durch „schlechte“ *rays* beschriebenen Untergrund (vgl. Abbildung 3.7). Ereignisse mit einem Vertex weit außerhalb der Wechselwirkungszone haben in ihrem z-Vertex-Histogramm nur die zufällig verteilten Untergrund-Einträge.
Aus dem z-Vertex-Histogramm werden verschiedene Triggerelemente gebildet. Als Eingangsgrößen für die Elemente stehen die totale Summe s der Einträge im Histogramm, die maximale Anzahl p der Einträge in einem Abschnitt, die Position des Maximums sowie zwei Qualitätsbits zur Verfügung. Das erste Qualitätsbit Q_1 ist gesetzt, wenn alle Einträge im Histogramm in vier benachbarten Abschnitten liegen. Wird das Maximum des Histogramms in einem der vier Randabschnitte angenommen, so ist Q_2 gesetzt. Einige der Triggerelemente sind in Tabelle 3.3 aufgeführt. Für das letzte Triggerelement, $zVtx\_sig$, werden folgende Bedingungen unterschieden:

$$zVtx\_sig1 \quad : \quad (\frac{p-b}{\sqrt{p}} > 1.5) \cdot \overline{Q_2} \tag{3.1}$$

$$zVtx\_sig\_mu\_D \quad : \quad (\frac{p-b}{\sqrt{p}} > 1.5) \cdot \overline{Q_2} \text{ oder } ((s+3) > 6 \cdot p) \cdot \overline{Q_2} \tag{3.2}$$

$$zVtx\_sig\_net \quad : \quad \text{sehr signifikantes Maximum} \tag{3.3}$$

Die Größe b ist gegeben durch $\frac{s-p}{15}$. Die erste Gleichung wird von Ereignissen mit einem signifikant über dem Untergrund liegenden Maximum erfüllt. Diese Bedingung wird als $zVtx\_sig1$ bezeichnet. Ist die zweite Gleichung erfüllt, so liegt die Bedingung $zVtx\_mu\_D$ vor. Die dritte Bedingung, $zVtx\_sig\_net$, beinhaltet Schnitte, die mit Hilfe eines neuronalen Netzes bestimmt wurden. Erfüllt ein Ereignis diese Schnitte, so besitzt es ein sehr signifikantes Maximum.

T_0	mindestens ein Eintrag im Histogramm gefordert, $s \geq 1$
mul	kodiert in Zahlen von 0 bis 7 die Summe s
small	$1 < p < 5$ und $s - p < 3$
cluster	alle Einträge liegen in vier benachbarten Abschnitten
sig	kodiert die Bedingungen (3.1) bis (3.3) in Zahlen von 0 bis 3: 0 $\widehat{=}$ kein signifikantes Maximum 1 $\widehat{=}$ sig_mu_D 2 $\widehat{=}$ sig_mu_D und $sig1$ 3 $\widehat{=}$ sig_mu_D und $sig1$ und sig_net

Tabelle 3.3: *Beispiele für die aus dem z-Vertex-Histogramm gebildeten Triggerbedingungen. Weitere Erläuterungen im Text.*

Die oben beschriebenen aktiven *rays* werden für weitere Triggerzwecke in maximal 224 *big ray*-Kandidaten gruppiert. Ein *big ray*-Kandidat umfaßt *rays*, die zum gleichen ϕ-Segment gehören und auf den gleichen Abschnitt des z-Vertex-Histogramms zeigen. Im Polarwinkel wird für die *big ray*-Kandidaten eine Unterteilung in 14 Abschnitte entsprechend der Geometrie im LAr-Trigger vorgenommen. Die in einem Ereignis gefundenen *big ray*-Kandidaten werden im weiteren Verlauf der Arbeit mit *big rays* bezeichnet.

3.4.1.2 $DCR\phi$-Trigger

Der $DCR\phi$-Trigger [H1-96-01, Ried95] soll Spuren in den zentralen Jetkammern (CJCs) mit einem DCA (**D**istance of **C**losest **A**pproach) von weniger als 2 cm finden. Der DCA ist der kleinste Abstand einer Spur zum nominellen Vertex. Durch die Grenze von 2 cm unterdrückt der Trigger Untergrundereignisse aus Strahl-Gas-Reaktionen, kosmischen Myonen und Synchrotronstrahlung, weil sie zu Spuren mit großem DCA führen.

Von den 54 Signaldrahtlagen der CJCs werden zehn für den $DCR\phi$-Trigger verwendet, sieben in der CJC1 (Lage 2, 4, 8, 12, 16, 20 und 22) und 3 aus der CJC2 (Lage 27, 29 und 31). Jede Lage ist in sogenannte Driftzeit-*pads* unterteilt. Die Größe dieser *pads* beträgt für die innersten zwei Triggerlagen 48 ns, für die folgenden sieben 96 ns und für die letzte 144 ns.

Eine Spur mit einem Transversalimpuls größer als 400 MeV wird in jedem Fall eine der CJC1-Triggerlagen in der Nähe des Signaldrahtes treffen, d.h. daß ein *pad* mit einer Driftzeit kleiner als eine Zeitscheibe gefunden wird. Dieses *pad* wird „promptes Signal“ genannt und mit den t_0-Masken verglichen. t_0-Masken sind die Driftzeit-*pads*, für die die Driftzeit kleiner als eine Zeitscheibe ist. Durch die „prompten Signale“ wird so der Zeitpunkt des Ereignisses festgelegt.

Neben den t_0-Masken werden für den $DCR\phi$-Trigger Treffer-Masken definiert, die durch die Spurlage im Driftraum und die Krümmung der Spur im Magnetfeld beschrieben werden. 10^4 solcher Masken decken den gesamten Raum aller möglichen Spuren ab. Durch Vergleichen der im Ereignis gefundenen Spurmuster mit diesen Masken werden die aktiven Masken herausgefunden.

Bei den Spuren ist eine Unterscheidung nach Ladung ($+/-$) und für zwei Transversalimpulsschwellen (niedrig: >450 MeV und nicht hoch; hoch: >800 MeV) möglich. Schließlich wird die Anzahl der aktiven Masken in 15 ϕ-Sektoren getrennt für positive und negative Ladung sowie für die niedrige und die höhere Transversalimpulsschwelle gezählt.

Als Triggerelemente werden beispielsweise Schwellenwerte für die Gesamtanzahl der aktiven Masken gefordert (Tabelle 3.4).

Ta	mindestens eine Spur mit $p_t > 450$ MeV
Tb	mindestens zwei Spuren mit $p_t > 450$ MeV
Tc	mindestens drei Spuren mit $p_t > 450$ MeV
Thig	mindestens eine Spur mit $p_t > 800$ MeV
Tneg	mindestens eine negative Spur mit $p_t > 450$ MeV

Tabelle 3.4: *Beispiele für Triggerelemente des DCRϕ-Triggers*

3.4.1.3 Myontrigger des instrumentierten Eisens

Für Triggerzwecke werden von den 16 Streamerrohrlagen des instrumentierten Eisens fünf Lagen verwendet (Lage 3, 4, 5, 8 und 12). Für jedes Modul können verschiedene Koinzidenzbedingungen von angesprochenen Lagen verlangt werden, beispielsweise: 3 aus 5 Lagen. Für die unterschiedlichen Bereiche des instrumentierten Eisens können verschiedene Bedingungen in den Modulen gefordert werden. Für die hintere Endkappe und die vordere äußere Endkappe werden drei angesprochene Lagen verlangt, für den Zentralbereich zwei und für die vordere innere Endkappe vier.
Die aus der Koinzidenzbedingung gebildeten Signale, die in Tabelle 3.5 mit 'μ' bezeichnet sind, werden an den L1-Trigger weitergegeben. Aus den Signalen werden acht Triggerelemente gebildet, die entweder mindestens ein oder mehr als ein angesprochenes Modul in verschiedenen Polarwinkelbereichen fordern (Tabelle 3.5) [H1-96-01].

Mu_BIEC	$> 0\ \mu$ in der inneren hinteren Endkappe
Mu_BOEC	$> 0\ \mu$ in der äußeren hinteren Endkappe
Mu_2_BIoOEC	$> 1\ \mu$ in der inneren oder $> 1\ \mu$ in der äußeren hinteren Endkappe
Mu_Bar	$> 0\ \mu$ im *Barrel*
Mu_FOEC	$> 0\ \mu$ in der äußeren vorderen Endkappe

Tabelle 3.5: *Beispiele für Triggerelemente des Myontriggers im instrumentierten Eisen. Die inneren Endkappen umfassen die sechs um das Strahlrohr liegenden Module der Endkappen, die äußeren Endkappen die weiter außen gelegenen.*

3.4.2 Zweite Triggerstufe (L2)

Nach einem L1*keep*-Signal wird die zweite Triggerstufe des Triggersystems aktiv. Der L2-Trigger ist ebenso wie L1 ein Hardware-Trigger. Er soll in einer Zeit von 20 μs die von L1 kommende Ereignisrate auf eine Rate von maximal 200 Hz reduzieren. Da jedoch die dritte Triggerstufe des Systems noch nicht benutzt wird (vgl. Abschnitt 3.4.3), muß L2 die Rate auf unter 50 Hz senken. Dazu werden gezielt einzelne (in der dieser Arbeit zwei)

L1-Subtrigger mit hoher Rate (>2Hz) in ihrer Rate reduziert.
Grundlage für die Entscheidung der zweiten Triggerstufe bilden die Triggerinformationen der einzelnen Detektorkomponenten, die L1-Entscheidungen sowie statistische Angaben wie beispielsweise das z-Vertex-Histogramm und die Anzahl von angesprochenen Modulen im Myondetektor. Diese Daten werden im L2-Trigger durch zwei verschiedene Systeme, L2NN (**L2** **N**euraler **N**etzwerktrigger) und L2TT (**L2** **T**opologischer **T**rigger), weiter verarbeitet. Diese beiden Triggerkonzepte werden in den Abschnitten 3.4.2.1 und 3.4.2.2 erläutert.
Wird ein Ereignis durch die zweite Triggerstufe angenommen (L2*keep*), so bleibt das Schieberegister gestoppt und die totale Auslese der Detektordaten beginnt. Bei einer negativen Entscheidung (L2*reject*) wird die Auswertung abgebrochen und das Schieberegister wieder gestartet.

3.4.2.1 L2TT

Der L2 Topologische Trigger [Bizo91, Bizo97] projiziert die vom Detektor kommenden Daten für jedes Subsystem auf eine 16×16 Matrix in θ und ϕ. Für die Projektionen werden bis zu 16 Familien definiert. Es können dabei reine (z.B. nur Kalorimeterinformation) oder gemischte (z.B. Kalorimeter und $DCR\phi$-Spuren) Familien gebildet werden. Bei den gemischten Familien werden die Informationen der verschiedenen Subsysteme in den verschiedenen $\theta\phi$-Bereichen durch beliebige logische Bedingungen verknüpft.
Die Projektionen einer Kalorimeterfamilie werden durch einen Energieindex beschrieben. Der Energieindex wird durch das maximale W mit zum Beispiel $W = \log_2 E$ bestimmt. Dabei ist E die Energie eines *big towers* im Kalorimeter. *Big towers* sind an Strahlen orientierte Teilbereiche des Kalorimeters, ähnlich den *big rays* in den Proportionalkammern. Sie haben dieselbe Unterteilung in θ und ϕ wie diese.
Außerdem erhält jede Projektion vier topologische Indizes $\phi, \theta\text{-}\phi, \theta\text{-}\theta$ und θ. Die topologischen Indizes können die gesamte Anzahl aktiver Zellen der Matrix oder die relative Lage von Zellen, von kleinen oder großen *clustern* (Bereich von n×m aktiven Zellen, z.B. 2×2 für kleine *cluster*) wiedergeben. Sie werden bezüglich der Projektion auf ϕ und auf θ sowie bezüglich charakteristischer Korrelationen zwischen θ und ϕ bzw. zwischen den θ-Projektionen zweier verschiedener Familien klassifiziert. Die verschiedenen topologischen Indizes werden verschiedenen Maschinen zugeordnet, die Triggerelementen entsprechen.
Für jeden topologischen Index und jede Familie wird in der W-versus-Index-Ebene für jedes physikalische Ereignis ein „Abstand“ zum Untergrund berechnet. In jeder Maschine wird dann über die „Abstände“ summiert. Ist das Ergebnis größer als ein für die Maschine spezifischer Grenzwert, so wird das Triggerbit der Maschine gesetzt, und das Ereignis wird von L2TT akzeptiert.
Es wird im L2TT ein grober Überblick über die Ereignistopologie gewonnen und topologische Korrelationen gefunden, durch die sich physikalische Ereignisse von Untergrund trennen lassen.
Beispiele für die Anwendung von L2TT sind ein Trigger für geladene Ströme auf dem L1-Subtrigger S71 und ein Trigger für QED-Compton-Ereignisse, der die Rate des L1-Subtriggers S13 reduziert.

3.4.2.2 L2NN

Ein anderes Konzept wird mit dem L2 Neuronalen Netzwerktrigger [Fent95, Köhn97] verfolgt. Mit Hilfe von 2-lagigen neuronalen *feed forward*-Netzen (Abschnitt 4) werden hochdimensionale Korrelationen und Schnitte im multidimensionalen Raum der Triggerdaten gesucht, um Physikereignisse von Untergrund zu unterscheiden. Die Netze werden vorher mit selektierten Physik- und Untergrunddatensätzen trainiert (vgl. Abschnitte 4.3, 6.1 und 6.3.1) und dann auf die Triggerhardware übertragen. Die Netze zeichnen sich durch eine genauere Trennung von Physik und Untergrund als in zwei Dimensionen und ihre schnelle parallele Struktur aus. Die in die Triggerbedingung eingehenden Detektordaten können für jedes Netz aus der Gesamtmenge unterschiedlich gewählt werden.
Der L2NN-Trigger setzt sich in der Hardware aus einer Reihe von CNAPS/VME-Karten (**C**onnective **N**etwork of **A**daptive **P**rocessor**S**) zusammen, die zusammen mit jeweils einem **D**ata **D**istribution **B**oard (DDB) eine Triggerbox bilden, die einem Triggerelement entspricht. Auf jeder dieser Triggerboxen ist ein Netz installiert.
Das DDB [Gold96] bereitet die Daten der Subdetektoren für die weitere Verwendung vor, beispielsweise bildet es die Summe über die gefundenen *big rays*. Außerdem liest es die Resultate des Netzes aus und leitet sie an die zentrale Triggerlogik weiter.
Die CNAPS/VME-Karten bestehen aus einer Matrix von 64 bis 512 Prozessorknoten, die parallel aus den vorverarbeiteten Detektorgrößen die Ausgabewerte der Knoten der ersten Lage des Netzes berechnen. Aus diesen Ergebnissen wird wiederum mit den Prozessorknoten der Wert am einzigen Knoten der Ausgangslage bestimmt. Das Triggerbit des L2NN-Triggerelements wird gesetzt, wenn das Ergebnis einen für das Netz spezifischen Schwellenwert überschreitet. Das Ereignis wird dann genommen und ein L2*keep* ausgelöst.
Beispiele für im Einsatz befindliche neuronale Netze sind auf S52/S54 laufende Trigger für elastische $J/\psi \to \mu^+\mu^-$ und $J/\psi \to e^+e^-$ sowie auf S68 ein Netz für $\phi \to K^+K^-$.

3.4.3 Dritte Triggerstufe (L3)

Die dritte Stufe des H1-Triggersystems soll bei einem L2*keep*-Signal gestartet werden. In einer Zeit von 800 μs soll sie auf Grundlage noch eingeschränkter Ereignisinformation ihre Entscheidung treffen. Wird ein Ereignis angenommen (L3*keep*), bleibt das Schieberegister gestoppt, bei einem negativen Ergebnis (L3*reject*) wird die Auslese gestoppt und das Schieberegister wieder gestartet.
Mit L3 soll die Rate weiter auf ungefähr 50 Hz gesenkt werden.
In der Planung des Experiments ist der L3-Trigger als aus Mikroprozessoren aufgebauter Softwaretrigger vorgesehen, jedoch wird er bis jetzt noch nicht eingesetzt. Damit muß die Ratenreduktion auf 50 Hz schon von der zweiten Triggerstufe geleistet werden.

3.4.4 Vierte Triggerstufe (L4)

Haben alle vorhergehenden Triggerstufen ein Ereignis akzeptiert, so gelangt es zur vierten Triggerstufe. L4 ist ein Multiprozessorsystem (*filterfarm*), das mit einer vereinfachten Er-

eignisrekonstruktion asynchron seine Entscheidung trifft. Hierfür steht die volle Detektorinformation zur Verfügung. Die akzeptierten Ereignisse werden als sogenannte Rohdaten auf Band geschrieben.
Ist die Auslese eines Ereignisses abgeschlossen, wird das Schieberegister wieder gestartet. Das Triggerergebnis der vierten Triggerstufe liegt nach 100 ms vor. Der Trigger reduziert die Rate um einen Faktor fünf auf ungefähr 10 Hz.

3.4.5 Weitere Datenverarbeitung

Auf einer weiteren fünften Stufe wird eine vollständige Ereignisrekonstruktion durchgeführt und eine Vorauswahl für physikalische Analysen getroffen. Die Ereignisse werden dabei Ereignisklassen zugeordnet und auf **D***ata* **S***ummary* **T***apes* (DST) gespeichert. Diese fünfte Stufe wird nach der Datennahme (*offline*) durchgeführt und trägt nicht zur Totzeit des Experiments bei.

Kapitel 4

Neuronale Netze

Gegenstand dieses Kapitels ist eine kurze Einführung in die Welt der neuronalen Netze. Neuronale Netze sind Gebilde der Mathematik und Informatik, die dem menschlichen Nervensystem nachempfunden sind. Sie werden vorwiegend zur Mustererkennung und in Klassifizierungsproblemen eingesetzt. Durch wiederholtes Anbieten von Lernmustern können neuronale Netze bestimmte Strukturen der Muster erlernen und optimal an ein Problem angepaßt werden (Training). Auf der Grundlage des Gelernten können sie andere unbekannte Muster klassifizieren.

Es sind verschiedene Arten von Netzen und Lernalgorithmen entwickelt worden. Hier wird nur auf die für den L2NN-Trigger relevanten Aspekte eingegangen. Im ersten Abschnitt werden die Parallelen zu biologischen Netzen erläutert. Anschließend wird der mathematische Formalismus der Netze und des Netztrainings dargestellt.

4.1 Von der Biologie zum Netz

Das Nervensystem von Tieren ist von paralleler Struktur. Es ist aus Neuronen aufgebaut, die die Information durch elektrische und elektrochemische Vorgänge übertragen [dtv84]. Ein Neuron besteht aus drei Komponenten: den Dendriten mit Rezeptorzellen, dem Zellkörper und dem Axon mit Synapsen (Abbildung 4.1).

Jede Nervenzelle hat mehrere Fortsätze (Dendriten), die die Eingangsinformationen sammeln. Die Dendriten nehmen über Rezeptorzellen die Information auf, die von den Synapsen benachbarter Zellen ausgesandt wird. Das Signal verschiedener Synapsen kann verschieden groß sein. Auch werden von unterschiedlichen Synapsenarten hemmende oder anregende Signale an die Rezeptorzellen übermittelt. Die Dendriten leiten die eingehenden Signale an den Zellkörper weiter. Hier werden die ankommenden Reize zur Gesamtantwort der Nervenzelle zusammengefügt. Ist die gesamte Summe aller Signale größer als eine durch den Axonhügel festgelegte Reizschwelle, so wird ein Ausgangssignal über das Axon an die Synapsen des Neurons weitergegeben. Die Größe und Art des Ausgangssignals hängt nicht von den eingehenden Signalen und ihrer Summe ab. Durch chemische und elektrische Vorgänge erfolgt dann die weitere Informationsübertragung an die Rezeptorzellen benachbarter Nervenzellen.

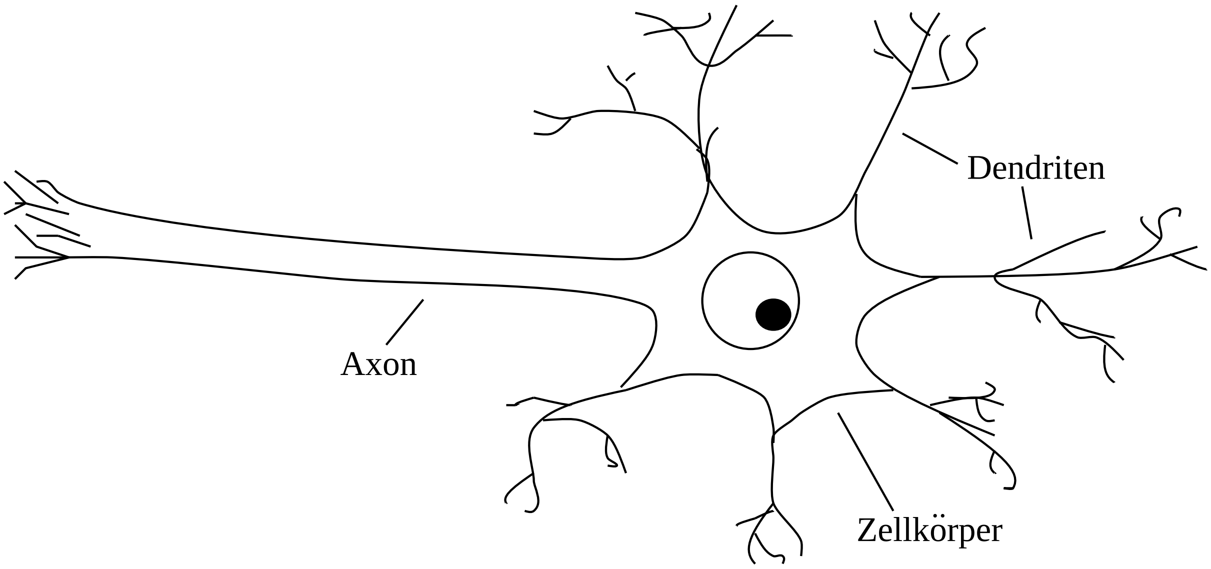


Abbildung 4.1: *Nervenzelle*

Die oben beschriebene biologische Struktur einer Nervenzelle wird für die Bestandteile neuronaler Netze übernommen:

- Eingabeleitungen mit verschieden großen Gewichten w_i, $i = 1, .., n$ analog zu den Dendriten und der Stärke und Art der von den Synapsen ausgehenden Signale
- ein Berechnungskörper, der Knoten genannt wird, mit einer gewissen Schwelle θ analog zu Zellkörper und Axonhügel
- eine Ausgabeleitung analog zum Axon

Das erste Modell eines Netzknotens ist von McCulloch und Pitts 1943 entworfen worden [McCu43].

4.2 Mathematik Neuronaler Netze

Das biologische Modell kann weiter abstrahiert und in mathematische Formeln gefaßt werden. Aus den einzelnen Berechnungskörpern können ganze Netzwerke paralleler Struktur unterschiedlichster Art aufgebaut werden, die neuronalen Netze.

4.2.1 Aufbau und mathematische Beschreibung eines Netzes

Jeder Knoten eines Netzes nimmt Signale auf, die entsprechend ihrer „hemmenden“ und „anregenden“ Art und ihrer Stärke gewichtet sind. Er summiert sie und vergleicht die Summe mit einem für den Knoten spezifischen Schwellenwert. Ist der Wert der Summe größer als die Schwelle, so „feuert“ der Knoten ein Ausgabesignal. Dieses läßt sich wie folgt mathematisch in Formeln fassen:

$$h_j \;=\; g(\sum_{i=1}^{n} w_{ij} \cdot x_i \;-\; \theta_j) \tag{4.1}$$

Dabei werden mit θ_j die Schwelle des Knotens j und mit w_{ij} seine Gewichtsfaktoren bezüglich der n Eingabewerte x_i bezeichnet. Die Funktion g wird Aktivierungsfunktion genannt. Sie berechnet den Ausgabewert h_j des Knotens. Der Ausgabewert eines Knotens kann zusammen mit Signalen von anderen Knoten als Eingabe für weitere Knoten verwendet werden. So bauen sich neuronale Netze aus mehreren Lagen auf. Sie besitzen allgemein mehrere Eingabe- und Ausgabeknoten und verschiedenste Verbindungen zwischen den Knoten.

In einem *feed forward*-Netz wird die Information nur in eine Richtung weitergegeben. Es liegen im Netz keine rekursiven Schleifen vor. Die Ausgabewerte einer Schicht werden direkt als Eingabewerte der nachfolgenden Schicht verwendet.

In Abbildung 4.2 ist eine Skizze eines neuronalen *feed forward*-Netzes dargestellt, wie es auf der Triggerstufe L2 benutzt wird. Das Netz besteht aus einer Eingabeschicht, einer

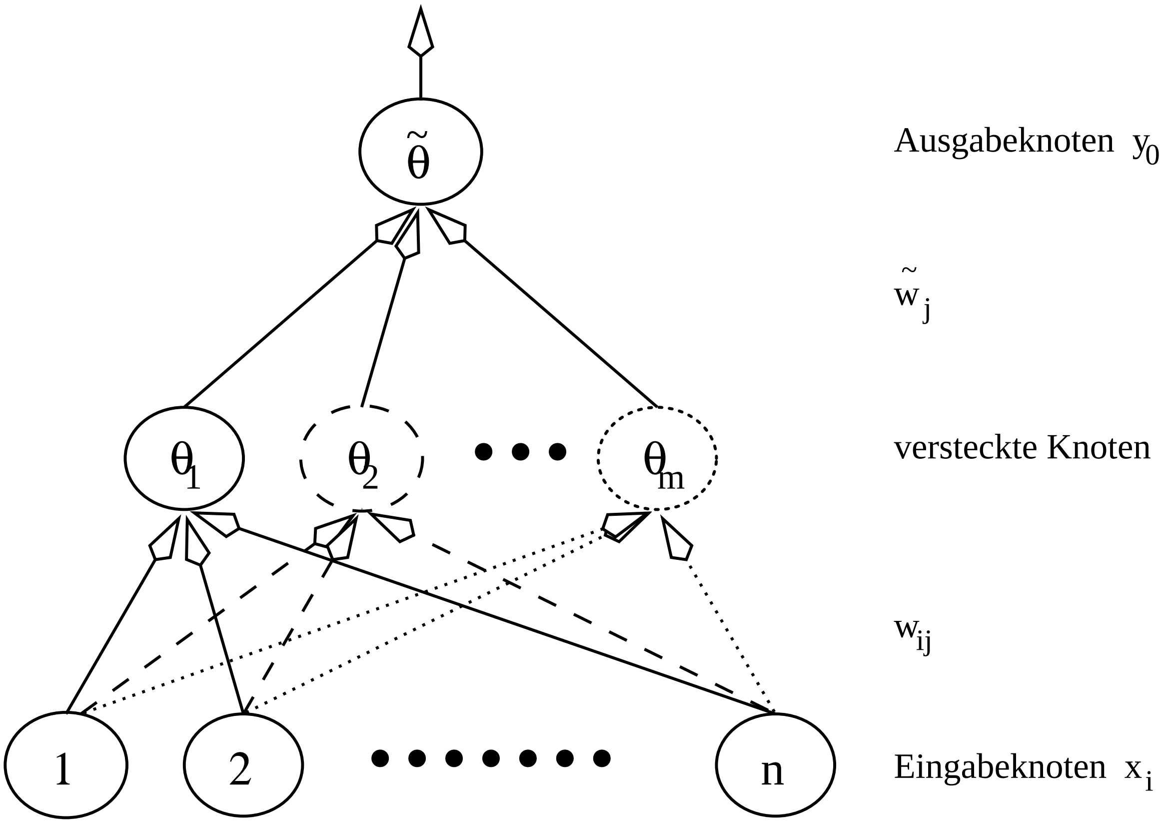


Abbildung 4.2: *Künstliches neuronales Netz mit einer versteckten Lage: Es besteht aus n Eingabeknoten, m Knoten in der versteckten Lage und einem Ausgabeknoten. Die Eingabewerte x_i werden mit Gewichten w_{ij} multipliziert. Der Knotenwert h_j ergibt sich aus der Summe der gewichteten Eingabewerte minus einem Schwellenwert.*

versteckten Schicht mit m Knoten (vgl. Abschnitt 4.2.2) und einer Ausgabeschicht. Die n Werte der Eingabeschicht entsprechen der vom Detektor gesendeten Information. Die für ein Netz aus dieser Information ausgewählten Größen heißen Eingabegrößen des Netzes. Nicht alle Netze des Triggers müssen auf Grundlage der gleichen Größen arbeiten.
Beide Lagen, versteckte Schicht und Ausgabeschicht, werden durch die Formel 4.1 beschrieben. Dabei wird in beiden Fällen dieselbe Aktivierungsfunktion g verwendet. Insgesamt ergibt sich für den Ausgabewert y_0 des Netzes:

$$y_0 \;=\; g(\sum_{j=1}^{m} \tilde{w}_j \cdot g(\sum_{i=1}^{n} w_{ij} \cdot x_i \;-\; \theta_j) \;-\; \tilde{\theta}) \tag{4.2}$$

Für die Verwendung eines neuronalen Netzes als Trigger t reicht ein Ausgabeknoten aus. Um eine Triggerentscheidung zu ermöglichen, wird ein Schnittwert $y_{Schnitt}$ je nach Art des Problems definiert. Ist der Ausgabewert y_0 des Netzes größer als der festgelegte Schnitt, so wird ein Ereignis der Klasse der physikalischen Ereignisse („Physik“) zugeordnet, andernfalls als Untergrund klassifiziert:

$$\text{Triggerbit } t \;=\; \begin{cases} 1 & : \; y_0 \geq y_{Schnitt} \\ 0 & : \; y_0 < y_{Schnitt} \end{cases} \tag{4.3}$$

4.2.2 Die versteckte Lage

Mit einem einzigen Knoten (Perzeptron) [Rose58] können nicht alle Probleme gelöst werden, bei denen zwei verschiedene Arten von Ereignissen klassifiziert werden sollen. Gleichung 4.1 beschreibt geometrisch die Lage einer Hyperebene im n-dimensionalen Raum der Eingabegrößen. Somit werden mit einem Knoten nur linear separable Klassen getrennt.
Ein mit einem einzigen Knoten nicht lösbares Beispiel ist das XOR-Problem [Roja93], das im folgenden skizziert wird:
Seien x_1 und x_2 zwei Variable im zweidimensionalen Raum, die jeweils die Werte Null oder Eins annehmen können. Von einem Netz sollen nun die beiden Fälle unterschieden werden, daß beide Eingabewerte den gleichen Wert besitzen (Ausgabewert = 0) bzw. daß beide Eingabewerte ungleich sind (Ausgabewert = 1). Wie man an Abbildung 4.3 erkennt, ist eine Separation der beiden Klassen im 2-dimensionalen Raum mit nur einer Gerade nicht möglich. Für eine Trennung wird eine weitere Gerade benötigt. Eine zweite Gerade bedeutet aber einen zweiten Knoten in der gleichen Schicht. Die beiden Knoten zerlegen den 2-dimensionalen Raum in drei Regionen, die je nach den Entscheidungen der beiden Knoten (0 oder 1) kodiert sind. Es treten dabei die Kombinationen 10, 01 und 00 auf.

x_1	0	1	0	1
x_2	0	0	1	1
y_0	0	1	1	0

Tabelle 4.1: *Das XOR-Problem im 2-dimensionalen Raum der Variable x_1 und x_2. y_0 ist der Ausgabewert des Netzes.*

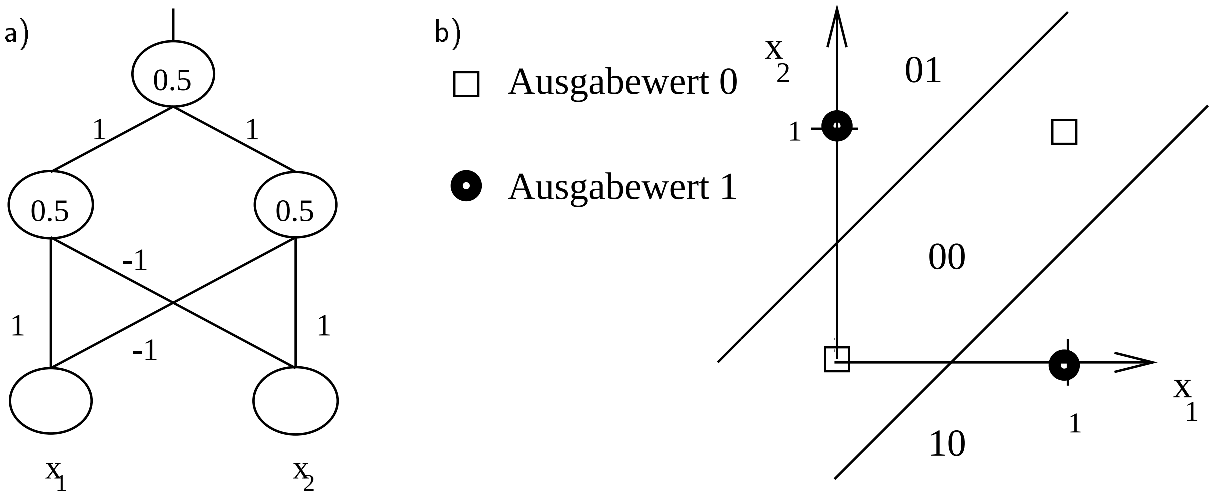


Abbildung 4.3: *a) Netz zur Lösung des XOR-Problems. An den Verbindungslinien sind die Gewichte eingetragen, in den Knoten die jeweiligen Schwellen. Ist der Wert der Summe der Eingangsgrößen größer als die Schwelle, so „feuert“ der Knoten eine 1 sonst eine 0. b) Geometrische Darstellung in der 2-dimensionalen Ebene. Die Zahlenpaare für die drei Teilbereiche ergeben sich aus den Ausgabewerte der beiden versteckten Knoten. Der Ausgabeknoten trennt linear diese Zahlenpaare.*

Ein weiterer Knoten in der Ausgabeschicht trennt die Kombinationen linear und gibt den Ausgabewert Null oder Eins heraus. Eine Lösung des Problems ist in Abbildung 4.3 dargestellt [Roja93].
Auf der anderen Seite reichen nach dem Theorem von Kolmogorov Netze mit einer Eingabeschicht, einer versteckten Lage und einer Ausgabelage aus, um beliebig komplexe Klassen im n-dimensionalen Raum zu trennen [Kolm57, Hech87]. Im Falle der Hochenergiephysik liegen zwischen den beiden Klassen meist sogar einfache Trennflächen vor [Lönn93].

4.2.3 Die Aktivierungsfunktion

Die Aktivierungsfunktion $g(z)$ eines mehrschichtigen Netzes ist i.a. eine nicht-lineare beschränkte Funktion. Wäre sie linear, so würde Gleichung 4.2 sich auf ein einziges lineares Gleichungssystem reduzieren. In entsprechender Weise würde sich das Netz auf eine Eingabe- und eine Ausgabeschicht verringern. Das ist gleichbedeutend damit, daß die beiden Klassen wieder linear trennbar sein müßten. Deshalb muß die Aktivierungsfunktion nicht-linear sein.
Ein Beispiel für eine häufig verwendete Aktivierungfunktion ist die *sigmoid*-Funktion s (Abbildung 4.4):

$$s(z) \;=\; \frac{1}{1+\exp(-\,z/T)} \quad ;\, T \neq 0 \;,\; z \,\epsilon\, R \tag{4.4}$$

Für $T \to 0$ geht die *sigmoid*-Funktion in die *Heavyside*-Funktion über, die bei der Darstellung des XOR-Problems verwendet worden ist.

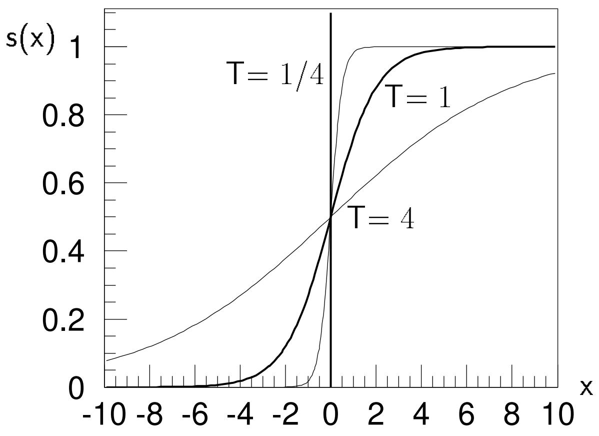


Abbildung 4.4: *Die* sigmoid*-Funktion: Sie liefert als Netzergebnis Werte zwischen 1 und 0. Für $T \rightarrow 0$ nähert sie sich der* Heavyside*-Funktion. Die Ableitung der* sigmoid*-Funktion geht direkt in den Lernalgorithmus (Abschnitt 4.3) ein.*

Die *sigmoid*-Funktion bildet das Netzergebnis auf das Intervall $[0, 1]$ ab. Die beiden asymptotischen Ausgabewerte Null und Eins lassen sich als Idealfall direkt dem Triggerbit zuordnen:

0 entspricht Untergrund
1 entspricht „Physik“

Durch die *sigmoid*-Funktion wird zwischen den beiden Idealfällen ein verschmierter weicher Übergangsbereich an der trennenden Hyperebene geschaffen, in den nicht eindeutig zuzuordnende Grenzfälle zwischen beiden Klassen abgebildet werden. Je nach Lage des Schnittwerts $y_{Schnitt}$ werden sie der einen oder anderen Klasse zugeordnet.

Ein weiterer Vorteil der *sigmoid*-Funktion ist ihre Differenzierbarkeit, die für den Lernalgorithmus (Abschnitt 4.3) von Bedeutung ist.

4.3 Lernalgorithmus

Neuronale Netze besitzen die Fähigkeit zu Lernen. Lernen bedeutet dabei, daß die Gewichts- und Schwellenfaktoren der einzelnen Knoten an ein vorliegendes Klassifizierungsproblem angepaßt werden.

Es gibt verschiedene Lernalgorithmen. Hier wird der *backpropagation*-Algorithmus vorgestellt, der für das Training der auf L2 benutzten Netze verwendet wird. Das Verfahren wurde erstmals 1985 von Rumelhart veröffentlicht [Rume86] und gehört zu den Algorithmen des überwachten Lernens.

Zum Lernen werden dem Netz wiederholt Lernmuster mit bekanntem Ausgabewerten y_0^{soll} angeboten. Das sequentielle Anbieten aller Lernmuster wird mit Lernepoche bezeichnet. Die Menge der Lernmuster wird Trainingsdatensatz genannt. Er besteht aus N Ereignissen der beiden zu trennenden Klassen (in der vorliegenden Arbeit „Physik“ und Untergrund genannt).

Das Verfahren sucht numerisch das globale Minimum der Fehlerfunktion E (Gleichung 4.5) durch iterative Veränderung der Gewichte in Gradientenrichtung $-\vec{\nabla} E$ (Gradientenabstiegsverfahren). Für den Ausgabeknoten lautet die Fehlerfunktion E:

$$E \;=\; \sum_{n=1}^{N} (y_{0,n}^{ist}(x_{n,i}, w_i) \;-\; y_{0,n}^{soll})^2 \tag{4.5}$$

Dabei sind $y_{0,n}^{ist}$ die Ausgabewerte des Netzes mit den gegenwärtigen Gewichten w_i und $x_{n,i}$ die Eingabewerte für das Lernmuster n. Zum Beginn des Trainingsprozesses werden die Gewichte des Netzes zufällig vorgegeben.
Ausgehend von der Ausgabelage und E werden die Gewichte des Knotens angepaßt. Die Korrektur der Gewichte ist gegeben durch:

$$\Delta w_i \;=\; -\gamma \cdot \frac{\partial E}{\partial w_i} \;. \tag{4.6}$$

γ gibt die Konvergenzgeschwindigkeit der Anpassung an und wird als Lernparameter bezeichnet. γ ist ein frei wählbarer Parameter des Trainings und muß für jedes Problem neu empirisch festgestellt werden. Zu große Werte für γ ($\gamma \gg 1$) können zu Schwierigkeiten beim Erreichen des Minimums führen, während bei kleinen Werten ($\gamma \ll 1$) die Anpassung der Gewichte in einem lokalen Minimum für E enden kann. An Gleichung 4.6 ist auch zu erkennen, daß eine differenzierbare Aktivierungsfunktion Voraussetzung für das Lernen der Netze ist.
Für mehrlagige Netze wird mit Hilfe der Kettenregel ausgehend von einer Lage der Fehler E zurückpropagiert und die Gewichtsänderungen der darunterliegenden Lage nach Gleichung 4.6 bestimmt. Die Berechnung der Gewichtsänderungen erfolgt also in entgegengesetzter Richtung zur Ermittlung der Ausgabewerte.
Bei starken Fluktuationen der Fehlerfunktion, beispielsweise wenn das Minimum in einem steilen Tal liegt, kann es bei der Gewichtsanpassung zu unerwünschten Oszillationen kommen. Diese können mit einem Momentumfaktor α durch Einbeziehen der Gewichtsänderungen im vorhergehenden Lernschritt (k-1) gedämpft werden:

$$\Delta^{(k)} w_i \;=\; -\gamma \cdot \frac{\partial E^{(k)}}{\partial w_i^{(k)}} \;+\; \alpha \cdot \Delta^{(k-1)} w_i \;. \tag{4.7}$$

α ist kleiner als Eins und muß ebenfalls empirisch bestimmt werden.
Um die Konvergenzgeschwindigkeit zu erhöhen, kann nach jedem Lernmuster statt nach einem gesamten Durchlauf des Trainingsdatensatzes eine Anpassung der Gewichte erfolgen. Diese Art von Algorithmus wird *online*-Training genannt und für das Training der Netze für den L2NN-Trigger verwendet.

Die Fähigkeit des austrainierten Netzes zur Übertragung des Gelernten auf andere Muster wird mit unbekannten nicht im Training verwendeten Ereignissen überprüft. Dabei erkennt das Netz Ähnlichkeiten zwischen den Lernmustern und dem Testereignis und ordnet dieses daraufhin einer der Klassen zu. Diese Fähigkeit des Netzes wird auch Generalisierung genannt.

Kapitel 5

Selektion der Trainingsdaten

Aufgabe dieser Arbeit ist es, neuronale Netze für Myonen im Zentralbereich des Detektors zu Triggerzwecken zu trainieren. Dabei sind zwei physikalische Kanäle zu beachten, Ereignisse mit unelastischen J/ψ-Kandidaten und Ereignisse mit inklusiver Myon-Produktion. Es stellt sich die Frage, ob beide nur grob ähnliche Ereignisdatensätze (mindestens ein Myon und drei Spuren) durch ein gemeinsames Netz effizient vom Untergrund getrennt werden können oder ob zwei getrennte Netze für den L2-Trigger entwickelt werden müssen. Die Selektion der beiden Ereignisgruppen wird im zweiten Teil dieses Kapitels vorgestellt. Neben den Physikdatensätzen werden für das Training eines Netzes Untergrundereignisse benötigt. Deren Auswahl wird am Ende des Kapitels beschrieben.
Das Kapitel beginnt mit der Vorstellung der L1-Subtrigger, deren Rate die neuronalen Netze reduzieren sollen. Hierbei wird der Aufbau der Subtrigger aus den Triggerelementen, sowie ihre Ratenentwicklung im Jahr 1996 und zu Beginn der Datennahmeperiode 1997 betrachtet.

5.1 L1-Subtrigger in der Datennahme

In dem L2-Triggerkonzept werden ein oder wenige L1-Subtrigger durch neuronale Netze in ihrer Rate verringert (Abschnitt 3.4.2).
In der vorliegenden Arbeit sollen neuronale Netzwerktrigger für Myonen im Zentralbereich des H1-Detektors ($20^\circ < \Theta < 160^\circ$) erarbeitet werden. Es werden daher für die L1-Subtrigger die Myon-Trigger S19 und S22 ausgewählt. Diese Subtrigger decken gemeinsam den Zentralbereich des Detektor ab und akzeptieren Ereignisse mit inklusiver Myon-Produktion und mit unelastischen J/ψ-Kandidaten ($J/\psi \rightarrow \mu^+\mu^-$). Die beiden Subtrigger sind wie folgt aus den Triggerelementen zusammengesetzt:

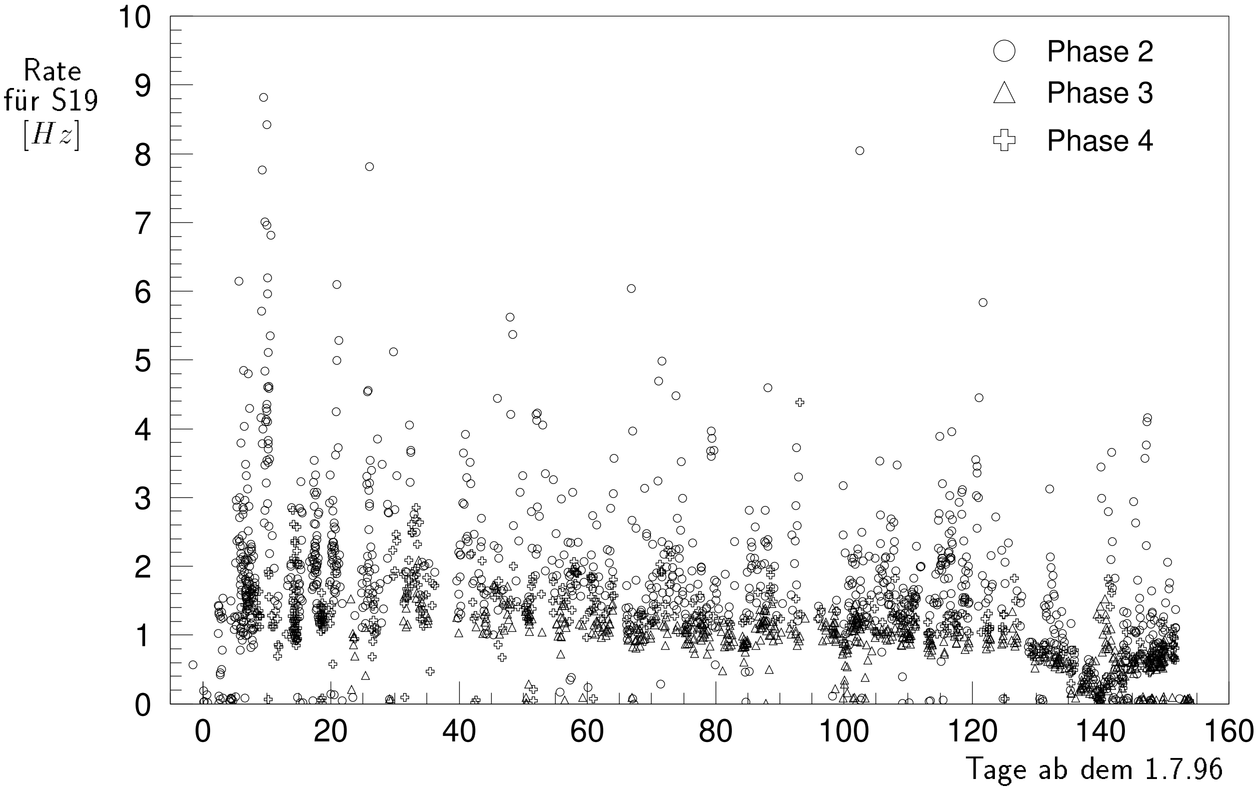
$$\begin{array}{rcl}
S19 & : & Mu\_Bar \,\&\&\, DCRPh\_CNH \,\&\&\, zVtx\_sig1 \\
S22 & : & Mu\_ECQ \,\&\&\, DCRPh\_CNH \,\&\&\, zVtx\_sig1 \\
\text{mit } DCRPh\_CNH & : & DCRPh\_Tc \,\&\&\, DCRPh\_TNeg \,\&\&\, DCRPh\_THig \\
 & & (\|\ \text{bedeutet dabei „oder“, \&\& bedeutet „und“})
\end{array}$$

Triggerelemente	Bedeutung
Mu_Bar	mindestens ein μ im Barrel des Detektors
Mu_ECQ	$Mu\_BOEC \,\|\|\, Mu\_2\_BIoOEC \,\|\|\, Mu\_FOEC$ mindestens ein μ in der äußeren vorderen oder hinteren Endkappe oder mindestens 2 μ in den hinteren Endkappen
$zVtx\_sig1$	signifikantes Maximum im z-Vertex-Histogramm
$DCRPh\_Tc$	mindestens 3 Spuren in den Driftkammern mit $p_t > 450$MeV
$DCRPh\_TNeg$	mindestens eine negative Spur $p_t > 450$ MeV
$DCRPh\_THig$	mindestens eine Spur mit $p_t > 800$ MeV

Tabelle 5.1: *Die verwendeten Triggerelemente. Vgl. dazu auch die Tabellen 3.3 bis 3.5 und die Abschnitte 3.4.1.1 bis 3.4.1.3.*

Die Bedeutung der Triggerelemente ist in Tabelle 5.1 aufgeführt (vgl. auch Abschnitt 3.4.1). Beide Subtrigger verlangen einen Myonkandidaten im *Barrel* (S19) oder in den Endkappen (S22) und mindestens drei Spuren in den zentralen Driftkammern sowie einen signifikanten Vertex in der Region der Strahlkreuzung.
Die Raten dieser Subtrigger im Jahre 1996 sind in Abbildung 5.1 dargestellt. Für Subtrigger S19 schwankt sie zwischen 1.5 und 5 Hz, S22 hat eine Rate von 0.3 bis 3 Hz. Außerdem zeigt sich in der Abbildung, daß die Raten, besonders für Subtrigger S19, nicht sehr stabil sind, sie steigen bis 10 Hz an. Der Untersetzungsfaktor (Abschnitt 3.4.1) ist für beide Subtrigger Eins. Daher erschien es notwendig, diese Subtrigger weiter zu reduzieren und zu stabilisieren. Der Subtrigger S22 wurde dabei in das Training mit einbezogen, um eine einheitliche Behandlung des gesamten Zentralbereichs zu erhalten. Eine Reduktion von 70 bis 80% reicht aus, um die in 1996 gemessenen Raten auf etwa 1 Hz zu senken.
Ein weiterer Grund für die Senkung der Rate dieser Subtrigger auf L2 ist, daß sich so die Möglichkeit eröffnet, die Effizienz der L1-Bedingung der Subtrigger zu verbessern:

- Es könnte eine schwächere Anforderung an das z-Vertex-Histogramm gestellt werden, nämlich $zVtx\_sig\_mu\_D$ statt $zVtx\_sig1$ (Abschnitt 3.4.1). Der Vorteil der neuen Bedingung liegt im folgenden:
  Das Triggerelement $zVtx\_sig1$ erreicht seine maximale Effizienz erst bei hoher Spuranzahl (ca. 15-20) im Ereignis. Andererseits ist die Effizienz der anderen z-Vertex-Triggerbedingungen $zVtx\_small$ und $zVtx\_cluster$, die von anderen Myontrigger benutzt werden, bei 5-spurigen Ereignissen bereits nahezu Null. Dadurch gehen Ereignisse mit mittlerer Multiplizität verloren.
  Das neue Triggerelement erreicht seine maximale Effizienz deutlich früher, bei ca. 3-5 Spuren. Damit können die vorher verlorengegangenen Ereignisse gerettet werden. Ebenso werden Ereignisse akzeptiert, die wenige Spuren haben, aber dennoch die Bedingung von $zVtx\_cluster$ verfehlen.
- Es könnte auch die strenge $DCR\phi$-Bedingung abgeschwächt werden. Z.B. ließe sich statt $DCRPh\_CNH$ nur eine Spur mit großem Impuls ($DCRPh\_THig$) fordern.



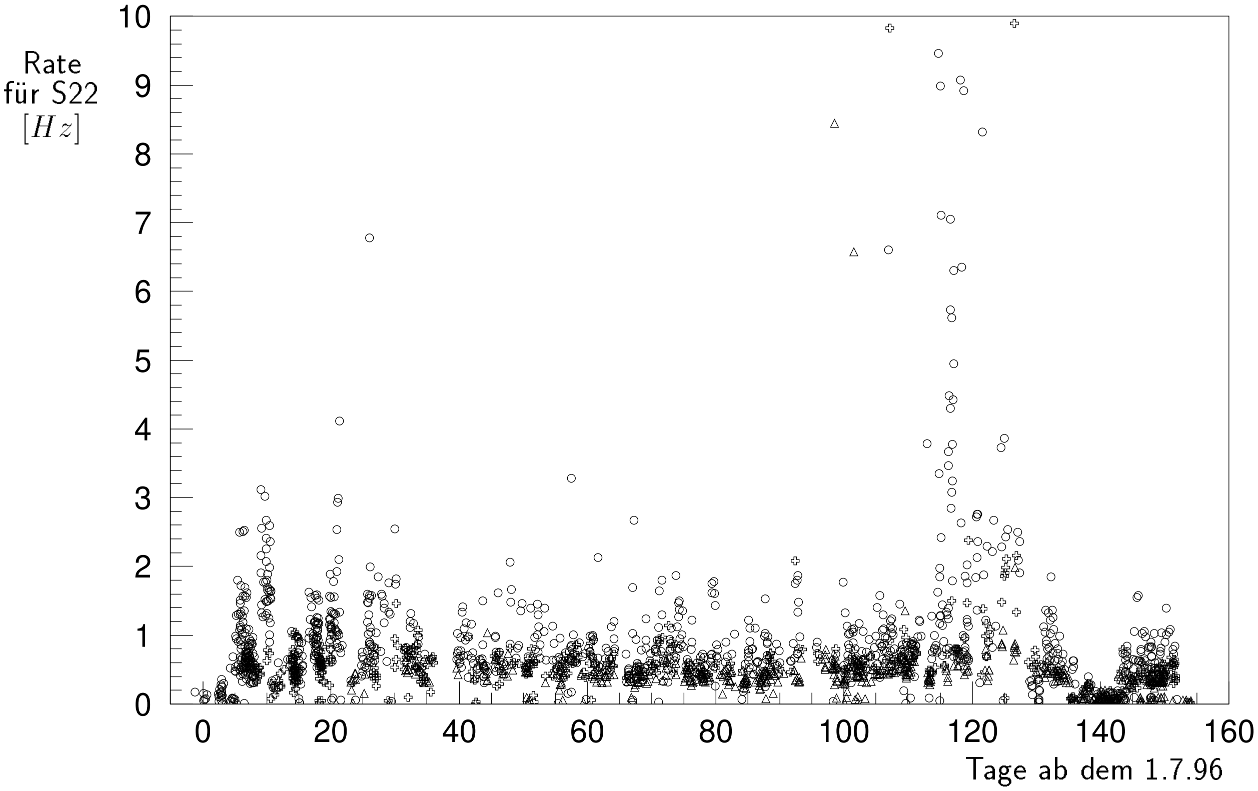


Abbildung 5.1: *Raten der Subtrigger S19 und S22 als Funktion der Zeit in der Datennahmeperiode 1996. Die Datennahme wird in Phasen aufgeteilt: Phase 2 beginnt kurz nach der Beschleunigung des Positronenstrahls und hat noch hohen Untergrund, Phase 3 und 4 folgen später.*

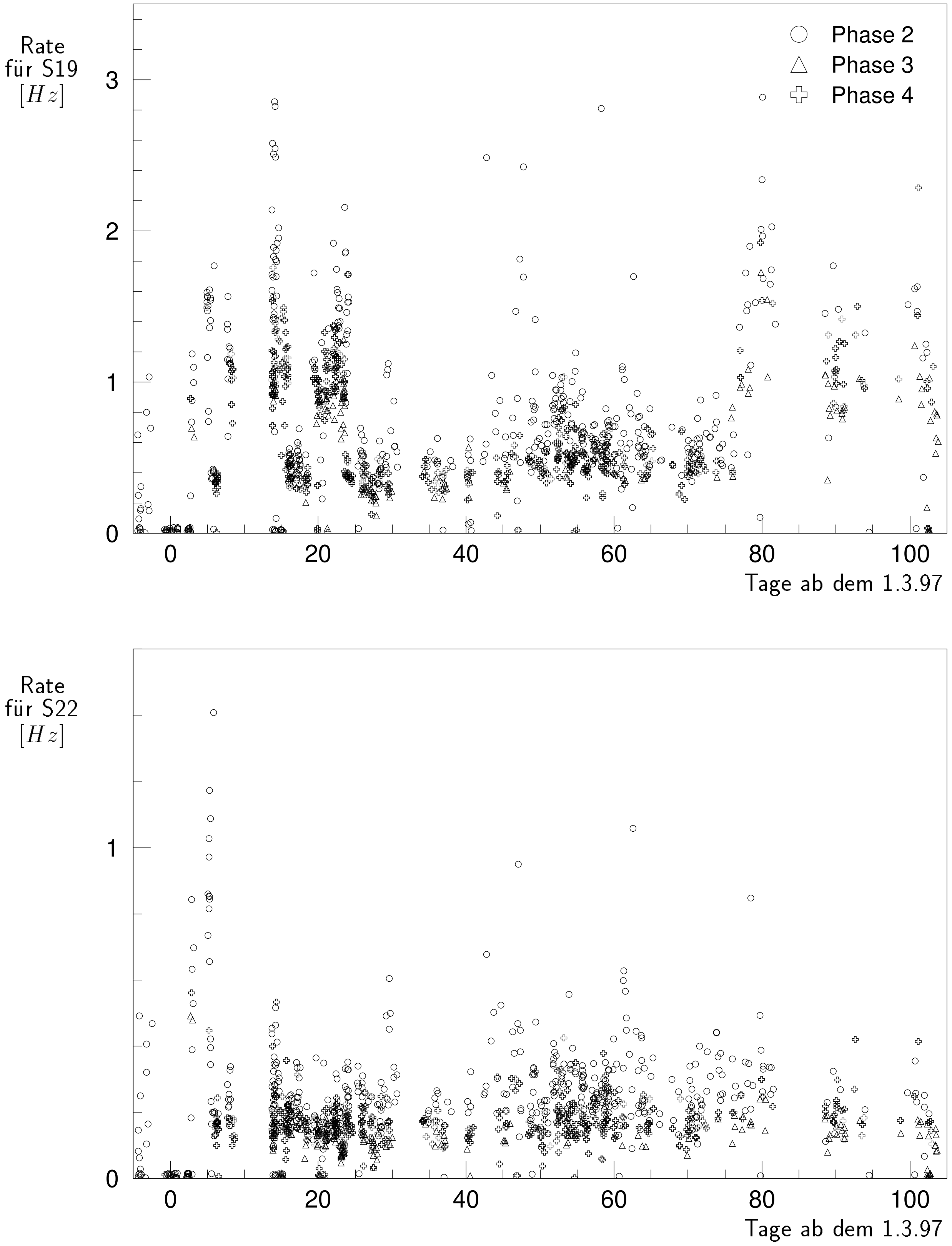


Abbildung 5.2: *Raten der beiden Subtrigger S19 und S22 in der Datennahmeperiode 1997. Die Raten sind gegen die Zeit aufgetragen.*

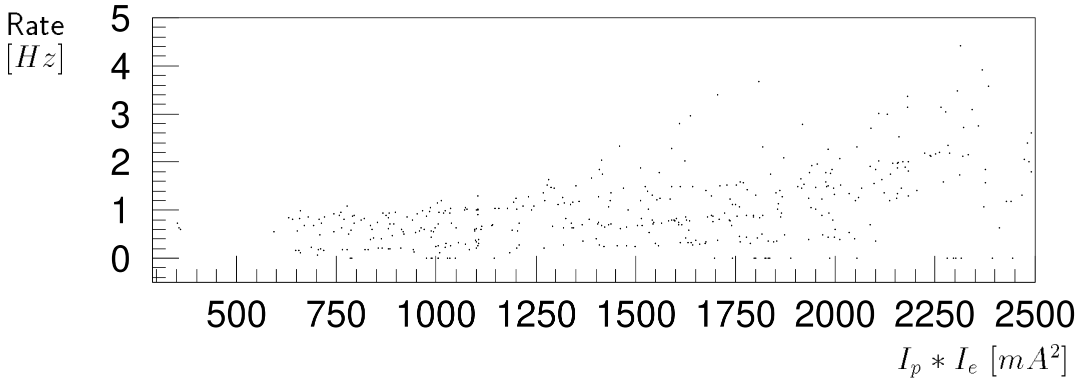


Abbildung 5.3: *Rate des Subtriggers S19 als Funktion des Produkts der Strahlströme.*

Dadurch gewinnt man Ereignisse, bei denen viele Spuren nach vorne gehen und nicht mehr im Akzeptanzbereich des $DCR\phi$-Triggers liegen. Somit ist die Effizienz der Triggerbedingung $DCRPh\_Tc$ für diese gering.

Diese Abschwächungen der Bedingungen von S19 und S22 sind ab Mitte Mai 1997 testweise als Subtrigger S15 $((Mu\_Bar||Mu\_ECQ)\,\&\&\,DCRPh\_THig\,\&\&\,zVtx\_sig\_mu\_D)$ implementiert worden. Das Verhalten der Netze für diesen Subtrigger wird in Abschnitt 6.3.7 untersucht.

Nun wird auf die Entwicklung der Raten von S19 und S22 in der ersten Hälfte des Jahres 1997 eingegangen. Diese sind in Abbildung 5.2 gegen die Zeit aufgetragen. Wie zu erkennen ist, liegen die Raten der beiden Subtrigger 1997 deutlich unter denen des Vorjahres. Beispielsweise hat S19 jetzt eine stabile Rate von ca. 0.5 bis 1.5 Hz. Für die Abnahme der Rate gibt es folgende Gründe:

- Es wurde eine neue Vetobedingung für die Subtrigger eingeführt, die Untergrund besser verwirft $(not(CIPB\_noSPCle\_T\_E1))$.
- Die Auslesezeit der Myontriggerelemente wurde auf im Mittel zwei Zeitscheiben verkürzt, wodurch die Zahl der zufälligen Koinzidenzen abgenommen hat.

Durch die nun niedrigeren Raten ist es vorläufig nicht mehr notwendig, eine weitere Reduktion durchzuführen. Dennoch bleibt die Möglichkeit einer Abschwächung der Triggerbedingungen bestehen und wird mit S15 erprobt. Für diesen Subtrigger können die Netze eingesetzt werden. Außerdem ist geplant, die Strahlströme von HERA zu erhöhen. Dadurch kommt es zu höheren Ereignisraten (Abbildung 5.3), so daß ein Einsatz von neuronalen Netzen nötig werden kann.

5.2 Auswahl der Ereignisse für Training und Test

In diesem Abschnitt wird die Auswahl des Physik- und des Untergrunddatensatzes für das Training der Netze beschrieben. Grundlage für die Selektion der Physikklassen sind von S19 oder S22 akzeptierte Ereignisse der Ereignisklasse 24. Diese Ereignisklasse umfaßt Ereignisse, in denen mindestens ein Myon im instrumentierten Eisen oder Vorwärtsmyonsystem oder im Kalorimeter mit mindestens mittlerer Qualität gefunden worden ist. Im ersten Teil des Abschnitts wird die Identifikation der Myonen und die Auswahl einer Spur aus verschiedenen Spurhypothesen dargestellt.
Auf dieser Vorselektion basierend wird die Selektion der beiden physikalischen Kanäle mit den in den Absätzen 5.2.2 und 5.2.3 angegebenen Bedingungen genauer beschrieben.
Die Auswahl des Untergrundes ist Gegenstand des letzten Teils dieses Abschnittes.

5.2.1 Spur- und Myonidentifikation

Die Spur- und Myonselektion basiert auf den rekonstruierten Spuren in den Spurkammern und im instrumentierten Eisen. Sie wird in der vorliegenden Arbeit mit von Lee West entwickelten Programmen [West96] durchgeführt. Die Selektion dient dazu, sogenannte „gute“ Spuren auszuwählen, Myonkandidaten zu selektieren und einem Teilchen eine eindeutige Spur im Detektor zuzuordnen, da bei der Spurrekonstruktion (Kapitel 3.2) häufig mehrere vom gleichen Teilchen stammende Spuren berechnet werden.
Bei der Spurselektion werden sowohl Bedingungen an die Qualität der Spuren gestellt als auch Ambiguitäten beseitigt. Als „gute“ Spuren gelten Spuren, die den im Anhang A in der Tabelle A.1 zusammengefaßten Schnitten und Kriterien genügen. Hierbei werden getrennt Spuren aus dem zentralen Spurdetektor, den Vorwärtsspurkammern und Kombinationen aus beiden betrachtet. Eine genauere Beschreibung der Anforderungen befindet sich z.B. bei [Dres97].
Ein Myon kann im Detektor durch das instrumentierte Eisen, das vordere Myonsystem oder auch im Kalorimeter identifiziert werden. In dieser Arbeit werden nur im instrumentierten Eisen identifizierte Myonkandidaten für die inklusiven Myonen berücksichtigt. Für J/ψ-Mesonen werden im instrumentierten Eisen oder im Kalorimeter erkannte Myonkandidaten verwendet. Bei der Selektion werden zur Identifizierung der Myonkandidaten verschiedene Bedingungen an die Spuren im instrumentierten Eisen, die Qualität der Verbindung von Spuren im Eisen und im Spurkammersystem sowie an die Spuren im Spurkammersystem für die im Kalorimeter gefundenen Myonen gestellt. Diese Anforderungen sind für die einzelnen Bereiche des instrumentierten Eisens (*Barrel*, FEC und BEC, Abschnitt 3.2) unterschiedlich und werden im Anhang A in Tabelle A.2 zusammengefaßt. Sie werden für die im Eisen erkannten Myonen z.B. genauer bei [Dres97] ausgeführt.
Auch bei der Myonerkennung kann es zu Ambiguitäten kommen. So ist es möglich, daß eine Spur im Eisen mit mehreren Spuren der inneren Spurkammern verbunden werden kann und umgekehrt. Für jede Verbindung wird eine Wahrscheinlichkeit berechnet, daß es sich tatsächlich um das gleiche Teilchen handelt. Es wird die Kombination mit der größten Wahrscheinlichkeit ausgewählt. Es werden dabei nur Verbindungen mit Wahrscheinlichkeiten größer als 10^{-4} berücksichtigt.

5.2.2 Selektion der Ereignisse mit inklusiven Myonen

Bei inklusiven Myonen handelt es sich um Ereignisse, in denen mindestens ein Myon im Detektor nachgewiesen wird. Diese Ereignisse sollen für die beiden Subtrigger, die den Zentralbereich ($20° < \theta < 160°$) des Detektors abdecken, selektiert werden. Dazu werden folgende Schnitte angewandt:

- Es wird mindestens ein im instrumentierten Eisen identifiziertes Myon verlangt.
- Dieses Myon soll im Zentralbereich nachgewiesen sein, d.h. sein Polarwinkel muß zwischen 20° und 160° liegen.
- Sein Transversalimpuls soll größer als 1 GeV sein.
- In dem Ereignis müssen mindestens drei nach der Spurselektion als „gut“ bezeichnete Spuren im Zentralbereich ($20° < \theta < 160°$) vorhanden sein. Damit wird schon ein Großteil des Untergrundes aus kosmischer Strahlung verworfen.

Mit diesen Bedingungen werden aus einem Teil der im Jahre 1996 genommenen Daten 5685 Ereignisse selektiert (vgl. Tabelle 5.3). Ein Beispiel für ein so selektiertes Ereignis ist in Abbildung 5.4 dargestellt.

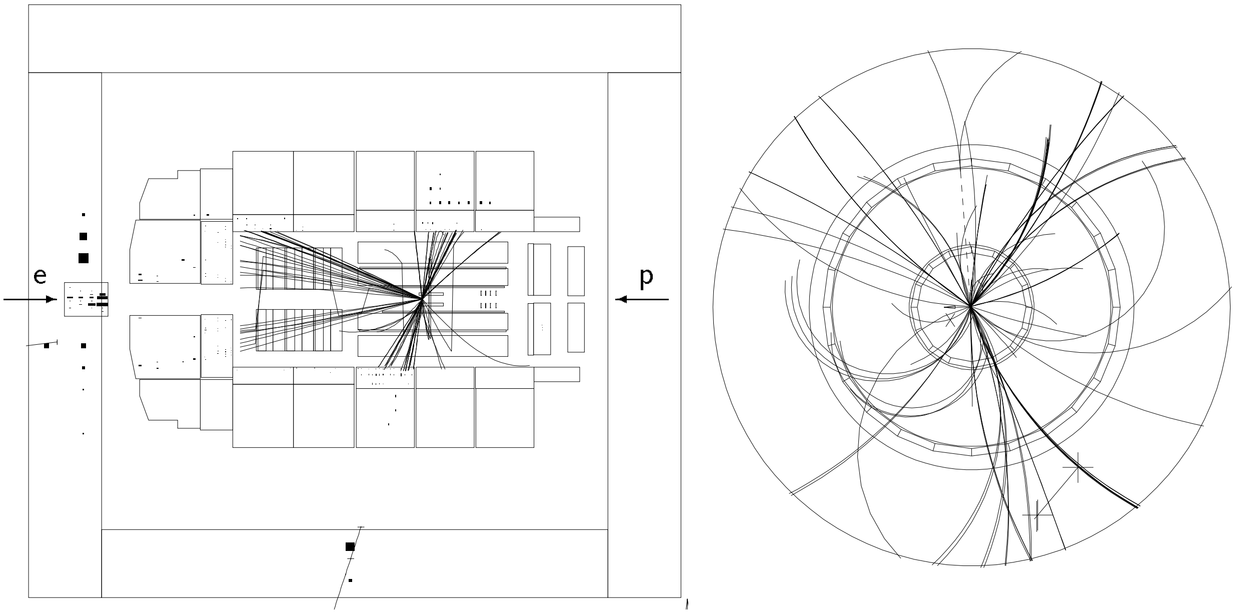


Abbildung 5.4: *Ereignis mit inklusivem Myon: Dargestellt ist links die Seitenansicht des H1-Detektors (ohne vorderes Myonsystem), sowie rechts eine radiale Ansicht der Spurkammern. Das Myon hat das Eisen unter einen Polarwinkel von ca. 75° erreicht und zu einer Annahme des Ereignisses durch Subtrigger S19 geführt. Außerdem sind 3 bis 4 Jets zu erkennen.*

5.2.3 Selektion der Ereignisse mit unelastischen J/ψ-Mesonen

Neben dem Datensatz mit inklusiver Myon-Produktion werden zum Training der Netze Ereignisse mit Kandidaten für unelastisch erzeugte J/ψ-Mesonen verwendet. Solche Ereignisse werden durch die folgenden Schnitte selektiert:

- Es werden mindestens zwei Myonen mit entgegengesetzter Ladung verlangt. Dabei ist es nicht von Bedeutung, ob sie im Kalorimeter oder im instrumentierten Eisen identifiziert werden.
- Mindestens eines dieser Myonen soll im Zentralbereich des Detektors liegen.
- Es werden mindestens drei nach den Bedingungen der Spurselektion „gute“ Spuren im Zentralbereich des Detektors verlangt. Diese Anforderung unterdrückt einen Großteil der Ereignisse mit elastischer J/ψ-Produktion, die genau zwei gute Spuren, nämlich die Myonen, enthalten. Ebenso wird hierdurch ein Teil des Untergrundes aus kosmischen Myonen herausgefiltert.
- Es wird zusätzlich noch ein Schnitt in $\sum\theta$ und $\Delta\phi$ zweier Myonen gegen kosmische Strahlung angewendet [Merk97].
- Es werden nur Ereignisse mit einer invarianten Masse $M_{\mu\mu} > 2$ GeV für zwei Myonen akzeptiert. Hierdurch werden nicht interessierende Ereignisse aus dem Datensatz entfernt, wie z.B. Myonen aus dem Zerfall leichterer Quarks.
  Wenn mehr als zwei Myonen im Ereignis gefunden werden, können jedoch neben dem J/ψ-Kandidat weitere Leptonpaare mit einer Masse kleiner als 2 GeV kombiniert werden.

Nach diesen Schnitten erhält man 677 Ereignisse (vgl. Tabelle 5.3). Da diese Anzahl für das Training eines Netzes sehr wenig ist, wird der Physikdatensatz für die unelastischen J/ψ-Kandidaten mit Ereignissen aus den Daten von 1995 ergänzt. Mit den gleichen Bedingungen wie für den aus 1996 selektierten Datensatz werden 1106 Ereignisse

1996	1995 vor 6.10.	1995 nach 6.10.
S19	S19' && $DCRPh\_Tc$ && $DCRPh\_TNeg$ && $DCRPh\_THig$ && $zVtx\_sig1$	S19' && $DCRPh\_Tc$
S22	S18' && $DCRPh\_Tc$ && $DCRPh\_TNeg$ && $DCRPh\_THig$ && $zVtx\_sig1$	S22' && $DCRPh\_TNeg$ && $zVtx\_sig1$

Tabelle 5.2: *Verwendete Triggerbedingungen für Ereignisse aus der Datennahmeperiode 1995. Die Zusammensetzung der Subtrigger wurde innerhalb von 1995 und zwischen '95 und '96 geändert. Deshalb ist eine Verschärfung der Subtrigger in 1995 notwendig, um für 1996 und 1995 gleiche Triggerbedingungen zu erlangen. Mit S19', S22' und S18' werden die Versionen der Subtrigger in 1995 bezeichnet.*

mit unelastischer J/ψ-Produktion gefunden. Dabei muß berücksichtigt werden, daß sich die Zusammensetzung der Subtrigger aus den Triggerelementen (vgl. Abschnitte 5.1 und 3.4.1) von '95 zu '96 sowie innerhalb '95 verändert hat [MeyA97]. Um die Triggerbedingungen in 1995 so anzupassen, daß sie denen von 1996 entsprechen, werden anstelle der Subtrigger S19 und S22 die in Tabelle 5.2 zusammengefaßten Bedingungen verlangt.
Die Selektionskriterien für die Physikdatensätze sind nochmal in Tabelle 5.3 aufgelistet. Für die unelastischen J/ψ-Kandidaten ist ein typisches Ereignis in Abbildung 5.5 gezeigt. Die invariante Masse der beiden Myonen in dem Ereignis beträgt $M_{\mu\mu} = 3.1$ GeV.

5.2.4 Auswahl der Untergrundereignisse

Als Untergrund werden in dieser Arbeit sogenannte L4-*transparent runs* verwendet. Dieses sind Daten, die unabhängig von der Entscheidung der vierten Triggerstufe aufgezeichnet werden. Das bedeutet, daß alle von L1 und L2 akzeptierten Ereignisse für weitere Untersuchungen zur Verfügung stehen. Auf den gewählten Subtriggern wird bisher keine L2-Entscheidung getroffen, so daß der selektierte Datensatz alle von L1 angenommenen Ereignisse umfaßt.
Man kann diese als Untergrund verwenden, da nur ca. 0.02 bis 0.2% der von S19/S22 akzeptierten Ereignisse in den Physikdatensatz gelangen. Dieses wird aus der Tabelle 5.4 deutlich. Andere physikalische Prozesse, die nicht zur Produktion von *bottom* und *charm*

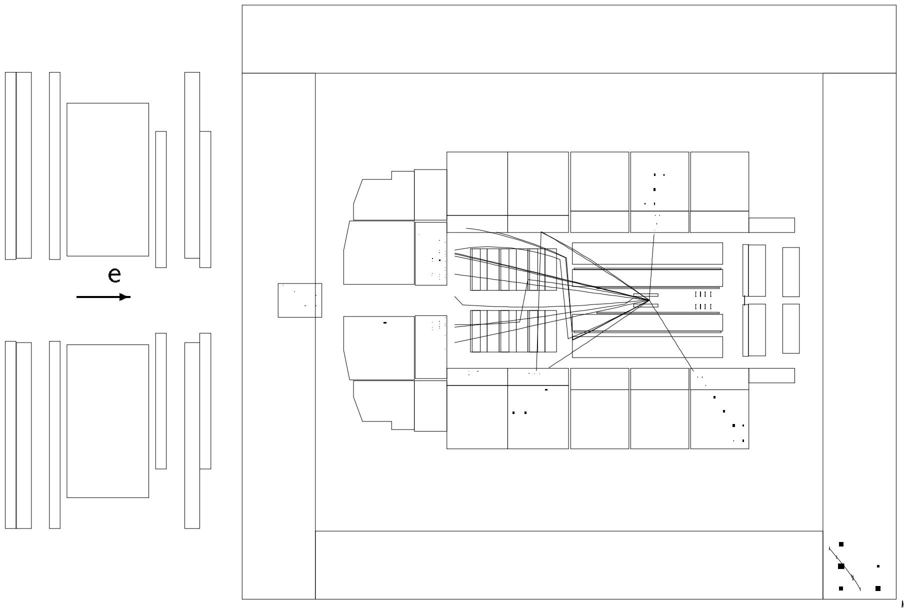


Abbildung 5.5: *Unelastisches J/ψ-Ereignis in der Seitenansicht des H1-Detektors: Das eine Myon wird im Kalorimeter und Eisen erkannt und verläßt den Detektor unter einem Winkel von ca. 125°. Das andere Myon, nur im Kalorimeter identifiziert, weist einen Winkel von ca. 90° auf. Die Masse des aus den beiden Myonen kombinierten J/ψ-Kandidaten beträgt $M_{\mu\mu} = 3.1$ GeV. Das Ereignis wurde von Subtrigger S22 akzeptiert.*

<table>
<tr><th></th><th>J/ψ in 1995</th><th>J/ψ in 1996</th><th>inkl. μ</th></tr>
<tr><td rowspan="6">Schnitte</td><td colspan="3">Klasse 24, S19 oder S22</td></tr>
<tr><td colspan="3">$\geq$ 3 gute Spuren</td></tr>
<tr><td colspan="2">$\geq 2\ \mu$</td><td rowspan="4">$\geq$ 1 „Eisenmyon“
im Zentralbereich
mit $p_t > 1$ GeV</td></tr>
<tr><td colspan="2">$\geq 1\ \mu$ im Zentralbereich</td></tr>
<tr><td colspan="2">$M_{\mu\mu} > 2$ GeV</td></tr>
<tr><td colspan="2">Schnitt gegen kosm. μ</td></tr>
<tr><td>*run*-Nummer</td><td>110946 bis 129464;
130401 bis 131045</td><td>152122 bis 164642</td><td>162081 bis 166214</td></tr>
<tr><td>Datum</td><td>1995</td><td>26.7.96 bis 19.10.96</td><td>5.10.96 bis 30.10.96</td></tr>
<tr><td>integr. Luminosität</td><td>$\sim 5.3pb^{-1}$</td><td>$\sim 5.2pb^{-1}$</td><td>$\sim 2.1pb^{-1}$</td></tr>
<tr><td>Ereigniszahl</td><td>1106</td><td>677</td><td>5685</td></tr>
</table>

Tabelle 5.3: *Übersicht über die Selektion der zum Training verwendeten physikalischen Ereignisse*

	Ereignisse in 1995	Bruchteil bezogen auf Gesamtanzahl
Gesamtanzahl für S19/S22	5-10 Mio.	100%
von L4 genommen	-	S19 4% S22 40-50%
Ereignisklasse 24 && S19/S22	50000	1%
Ereignisklasse 24 && S19/S22 unelastische J/ψ-Kandidaten inklusive Myonen	1000 10000	0.02% 0.2%

Tabelle 5.4: *Abschätzung des Anteils von in dieser Arbeit selektierter „Physik“ an allen von S19/S22 akzeptierten Ereignissen. Da der Anteil gering ist, können Ereignisse aus L4-transparent runs (siehe Text) als Untergrund für das Training genommen werden.*

run-Nummer	163490	163361	156923	167886	166276
N_{Ereig}	506	9	269	198	162
run-Nummer	162604	164621	165687	165688	166278
N_{Ereig}	430	520	978	964	194

Tabelle 5.5: *L4-transparent runs in dem Zeitraum vom 27.8. bis 10.11.96. Aus ihnen ist der Untergrund für das Training der Netze selektiert worden.*

beitragen, werden zum Untergrund gerechnet (z.B. Produktion leichterer Quarks).
Bei der Selektion wird als einzige Bedingung gefordert, daß die Ereignisse durch mindestens einen der beiden Subtrigger S19 und S22 akzeptiert worden sind. Es werden keine Ereignisklasse oder weitergehende Schnitte verlangt. So wurden aus verschiedenen L4-*transparent runs* insgesamt 4230 Ereignisse ausgewählt. Die verwendeten Daten sind in Tabelle 5.5 zusammengestellt. Von den Ereignissen sind 2004 von S19 und 2268 von S22 angenommen worden.
Eines dieser Untergrundereignisse ist in Abbildung 5.6 dargestellt. Das Ereignis ist vom Subtrigger S22 akzeptiert worden. Der Vertex liegt deutlich vor der Wechselwirkungszone des H1-Experiments, und die entstehenden Teilchen sind ausschließlich in die Richtung des einlaufenden Protons nach vorne ausgerichtet. Bei dem Ereignis handelt es sich aller Wahrscheinlichkeit nach um Strahl-Gas-Untergrund.

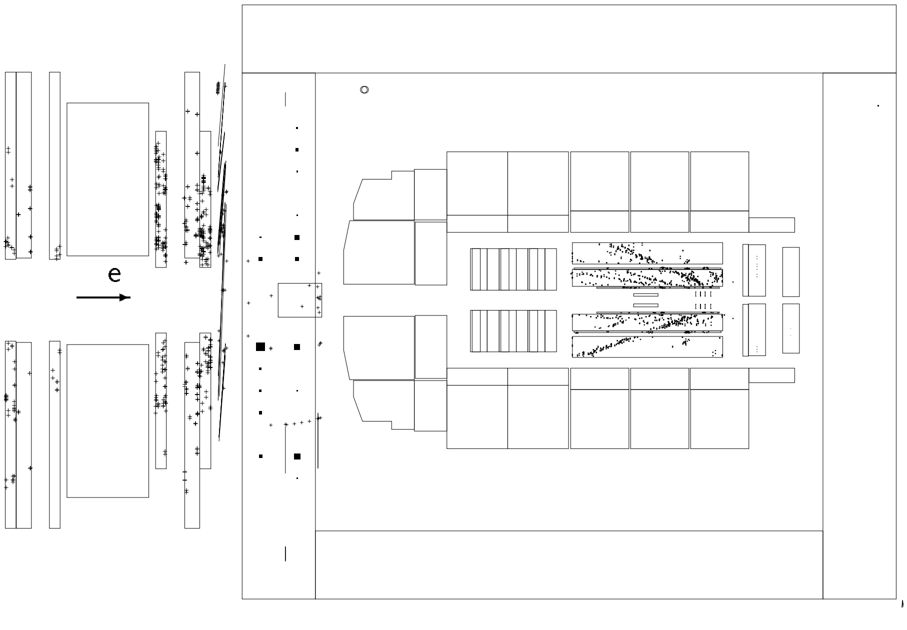


Abbildung 5.6: *Untergrundereignis aus Strahl-Gas-Wechselwirkung in der Seitenansicht des Detektors: Der Ereignisvertex liegt deutlich vor der Wechselwirkungszone. Die Spuren sind stark nach vorne gerichtet.*

Kapitel 6

Das Training und Testen der Netze in der Praxis

Gegenstand dieses Kapitels sind das Training der Netze und die erreichten Resultate.
In Abschnitt 6.1 wird der Vorgang des Lernens der Netze, wie er sich in der Praxis darstellt, beschrieben. Daran anschließend werden die beim Training verwandten Eingabegrößen erläutert. Im Laufe der Bearbeitung der Aufgabenstellung sind neue Eingabegrößen entwickelt worden. Diese werden getrennt von den vorhandenen Größen betrachtet.
Der letzte Teil dieses Kapitels befaßt sich mit dem eigentlichen Training der Netze und den durchgeführten Tests. Dabei werden die erreichten Effizienzen nach Training und Test, sowie das Verhalten der Netze in der Datennahmeperiode 1997, speziell auch für den neuen Subtrigger S15, untersucht. Ebenso wird der Einfluß der Anzahl von versteckten Knoten und Eingabegrößen an einem Beispiel analysiert.

6.1 Überblick über den Trainingsvorgang

Nachdem im Abschnitt 4.3 die theoretischen und mathematischen Grundlagen für das Lernen eines Netzes dargestellt worden sind, wird nun der Trainingsvorgang in der Praxis [Samt96, WesD96, Zobe96] anhand der Abbildung 6.1 erläutert.

Auswahl der Trainingsklassen

Für das Training eines Netzes nach dem Trainingsalgorithmus des *backpropagation*-Algorithmus (Abschnitt 4.3) werden zwei Klassen, „Physik“ und Untergrund, benötigt. Es muß also zu Beginn der Trainingskette eine Auswahl dieser beiden Klassen vorgenommen werden (vgl. Abschnitt 5.2). Dem L2-Triggerkonzept folgend werden dabei nur einige wenige Subtrigger als Grundlage für die Ereignisse gefordert, in dieser Arbeit S19 und S22. Bevor nun das eigentliche Training stattfindet, müssen die selektierten Datensätze einige Aufbereitungsschritte durchlaufen, die schematisch in Abbildung 6.1 dargestellt sind.

Zunächst werden sie vom Netzwerk-Simulator *NETSIM* [Grub95] in ein für die nachfolgenden Programme lesbares Format gebracht, die sogenannten Ereignisvektoren, die u.a. alle möglichen Eingabewerte der Ereignisse beinhalten [EV96].
Außerdem kann mit dem Programm *NETSIM* die Entscheidung eines Netzes simuliert werden, d.h. das Ergebnis des Netzes kann für einen Satz beliebiger Ereignisse berechnet werden. Diese Möglichkeit wurde z.B. für die Berechnung der Effizienz der einzelnen Subtrigger und Netze verwendet (Abschnitt 6.5).

Zusammenstellung der Trainings- und der Testdatensätze

In der Trainingskette eines Netzes folgt nun das Mischen der „Physik“ und des Untergrundes zu einem Trainingsdatensatz durch das Programm *MIXER* [Zobe95]. Hierbei wird die Gesamtanzahl der Trainingsereignisse sowie das Verhältnis von „Physik“ und

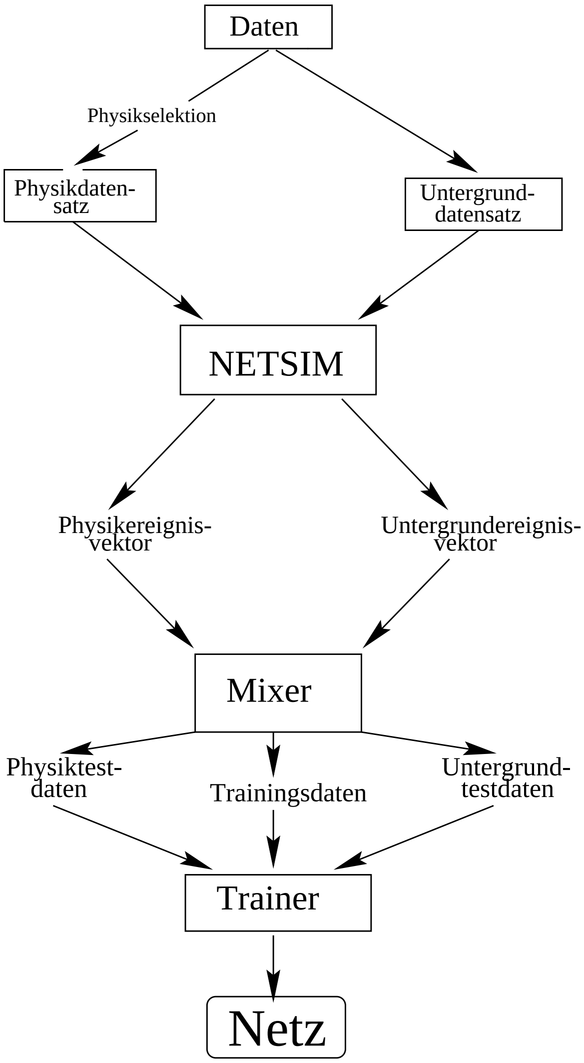


Abbildung 6.1: Vorgang des Trainings in der Praxis.

Untergrund, bzw. auch zwischen verschiedenen Physikklassen, vom Anwender gewählt. Zusätzlich können im *MIXER* Subtriggerbedingungen und physikalische Ereignisklassen verlangt werden. Es werden außerdem die Verteilungen der einzelnen Eingabegrößen in Histogrammen dargestellt.
Neben dem Trainingsdatensatz werden vom *MIXER* Testdatensätze für „Physik" und Untergrund erzeugt. Bei diesen wird darauf geachtet, daß keines der Ereignisse sowohl zum Lernen als auch zum Testen der Netze verwendet wird. Es entstehen vom Trainingsdatensatz unabhängige Datensätze, mit denen die Fähigkeit des Netzes zur Generalisierung überprüft werden kann.

Training und Test der Netze

Das Netz wird nun mit dem Programmpaket *TRAINER* trainiert, der auf Basis des *backpropagation*-Algorithmus (Abschnitt 4.3) arbeitet. Das Netz lernt durch wiederholtes Anbieten der vom *MIXER* gemischten Trainingsereignisse. Nach jedem Lernschritt (Epoche) wird mit dem vorläufigen Netz der Ausgabewert der Trainings- und Testereignisse berechnet und eine Effizienz für das Erkennen der Physikereignisse in den Datensätzen bei bestimmter Untergrundunterdrückung ermittelt.
Vor dem Training werden die vom Netz zu berücksichtigenden Eingabegrößen ausgewählt und die Anzahl der versteckten Knoten festgelegt (vgl. Abschnitt 6.3.2). Ebenso muß man dem TRAINER den Lernparameter γ und den Momentumfaktor α angeben (Abschnitt 4.3). Diese Parameter sind bei verschiedenen Trainingsdurchläufen variiert worden, und für die endgültigen Netze ist eine möglichst gute Kombination ausgewählt worden (vgl. Abschnitt 6.3.1).
Auch muß festgelegt werden, wie oft der Trainingsdatensatz dem Netz während des Lernens angeboten werden soll (Anzahl der Epochen). Erkennt man nach dem Training, daß die Differenz der Effizienz des Testdatensatzes zwischen den letzten zwei Epochen groß ist, so muß in einem neuen Training die Epochenzahl erhöht werden. Zeigt die Effizienz keinen weiteren Anstieg, so ist das Netz fertig trainiert.
Das Testen der Netze erfolgt einmal durch den im TRAINER implementierten Generalisierungsschritt, bei dem die Ausgabewertverteilungen der Testdatensätze für das fertige Netz ermittelt werden. Zum anderen kann ein weiteres Programm, der *TESTER* [WesD95], verwendet werden. Mit diesem kann die Berechnung der Ausgabewerte eines Netzes für einen beliebigen Testdatensatz durchgeführt werden. Außerdem bietet er die Möglichkeit, die Bedeutung der verschiedenen Eingabegrößen zu untersuchen (vgl. Abschnitt 6.3.5).

Beim Training eines Netzes besteht die Gefahr, daß das Netz übertrainiert wird. Dabei handelt es sich um „Auswendiglernen" der Trainingsereignisse im Detail mit zunehmender Lernzeit und Knotenzahl durch das Netz, so daß zwar diese mit steigender Effizienz vom Untergrund getrennt werden, aber die Testereignisse nicht mehr in gleicher Weise erkannt werden können. Die Effizienz des Netzes für die Testereignisse steigt nicht weiter, sondern sinkt häufig wieder ab.

6.2 Eingabegrößen der Netze

Nachdem die verwendeten Datensätze sowie die Trainingskette in der Praxis umrissen worden sind, wird auf die benutzten Eingabegrößen eingegangen. Dabei werden zunächst die bisher im L2NN-Trigger vorhandenen Größen beschrieben und anschließend neuentwickelte Größen vorgestellt.

6.2.1 Vorhandene Eingabegrößen

Die Eingabegrößen stellen die Knoten der Eingangslage des Netzes dar. Ihre Werte sind Bestandteil des Ereignisvektors, der vom *NETSIM* geschrieben wird und vom *MIXER* weiter verarbeitet wird. Von den zur Verfügung stehenden Eingabegrößen werden die folgenden 18 verwendet. Es sind die im Training üblichen Namen [EV96] und eine kurze Erklärung angegeben.

- z-Vertex-Histogramm

– *cpvsum*	Summe der Einträge
– *cpvmax*	maximale Anzahl an Einträgen in einem Abschnitt
– *cpvpos*	Position der maximalen Anzahl

- Driftkammerspuren

– *trhineg*	Anzahl der $DCR\phi$-Spuren mit hohem Impuls und negativem Ladungsvorzeichen
– *trhipos*	Anzahl der $DCR\phi$-Spuren mit hohem Impuls und positivem Ladungvorzeichen
– *trloneg*	Anzahl der $DCR\phi$-Spuren mit niedrigem Impuls und negativem Ladungsvorzeichen
– *trlopos*	Anzahl der $DCR\phi$-Spuren mit niedrigem Impuls und positivem Ladungsvorzeichen
– *trtot*	Anzahl aller $DCR\phi$-Spuren

- Flüssig Argon Kalorimeter (LAr)

– *larife*	Energie im inneren vorderen Teil des LAr
– *larfbe*	Energie im vorderen *Barrel* des LAr
– *larcbe*	Energie im zentralen *Barrel* des LAr
– *laret*	transversale Energie im LAr
– *laretot*	Gesamtenergie im LAr

- Instrumentiertes Eisen

– *ironbe*	Anzahl der angesprochenen Module in der hinteren Endkappe (BEC)
– *ironfe*	Anzahl der angesprochenen Module in der vorderen Endkappe (FEC)
– *ironfb*	Anzahl der angesprochenen Module im vorderen *Barrel*
– *ironbb*	Anzahl der angesprochenen Module im hinteren *Barrel*
– *irontot*	Gesamtanzahl der angesprochenen Module

In Abbildung 6.2 sind für die Verteilungen der Eingabegrößen mehrere Beispiele gezeigt. Dabei sind Ereignisse mit inklusiven Myonen sowie Untergrund dargestellt. Man

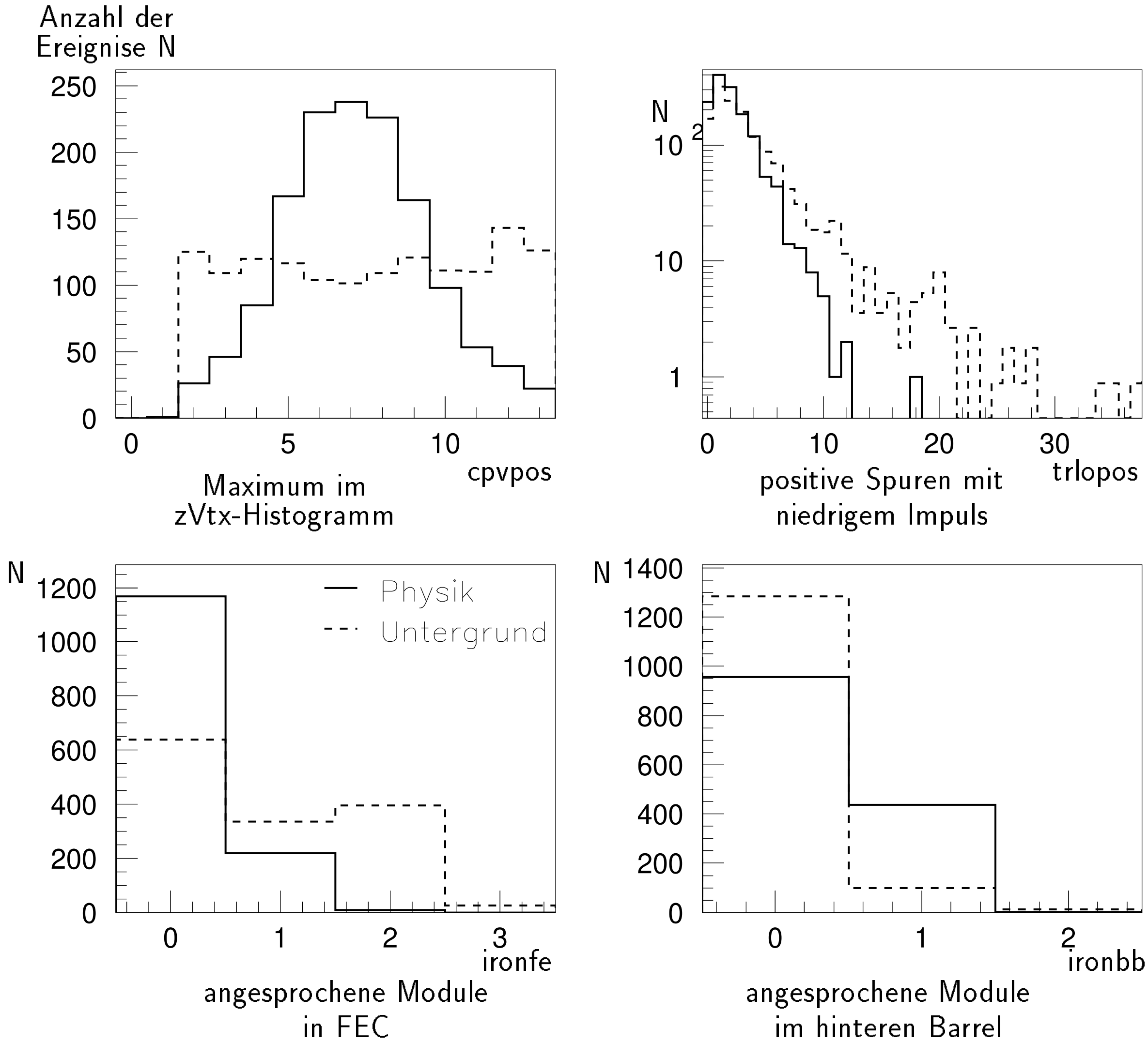


Abbildung 6.2: *Verteilungen von Physik (durchgezogene Linie) und Untergrund (gestrichelt) für die Eingabegrößen* cpvpos, trlopos, ironfe *und* ironbb.

erkennt, daß die Position des Maximums im z-Vertex-Histogramm für Untergrund nahezu gleichmäßig über das gesamte Histogramm verteilt ist, während sich für die Physik ein Maximum in der Mitte des Wechselwirkungsbereiches ausbildet.
Auch enthalten Untergrundereignisse im Mittel mehr Einträge für Spuren in den Driftkammern mit niedrigem Impuls und positiver Ladung.
Aus den beiden dargestellten Größen für das instrumentierte Eisen ist zu sehen, daß für Physik mehr rückwärtige Module angesprochen haben, für Untergrund mehr vordere.

6.2.2 Neuentwickelte Eingabegrößen

Neben den in Abschnitt 6.2.1 erwähnten Größen werden für das Training der Netze weitere neuentwickelte Eingabegrößen benutzt, die eine grobe topologische Verteilung der Spuren im Ereignis berücksichtigen.
Ereignisse aus ep-Wechselwirkung besitzen im Mittel mehr Spuren mit größeren Polarwinkeln θ als der Untergrund. Dieses läßt sich dadurch erklären, daß der Untergrund vorwiegend aus Strahl-Gas-Wechselwirkungen besteht. Die Spuren von Strahl-Gas-Ereignissen gehen in Richtung des Protonimpulses und sind so zum Großteil nach vorne gerichtet. Durch eine Aufteilung des Polarwinkelbereiches in mehrere θ-Bereiche und Summieren der Spuren in diesen Bereichen lassen sich neue Größen zur Unterscheidung von „Physik“ und Untergrund für ein Netz konstruieren.
Für Spurinformationen stehen den neuronalen Netzen des L2-Triggers die *big rays* zur Verfügung. *big rays* werden aus den Proportionalkammern des zentralen und des vorderen Spurkammersystems gewonnen. Die Proportionalkammern liefern dabei eine grobe Spurerkennung (vgl. Abschnitt 3.4.1.1).
big rays werden in 14 θ- und 16 ϕ-Abschnitte unterteilt. Der θ-Abschnitt Null umfaßt dabei den am weitesten nach vorne gerichteten Polarwinkelbereich (vgl. Tabelle B.1 in

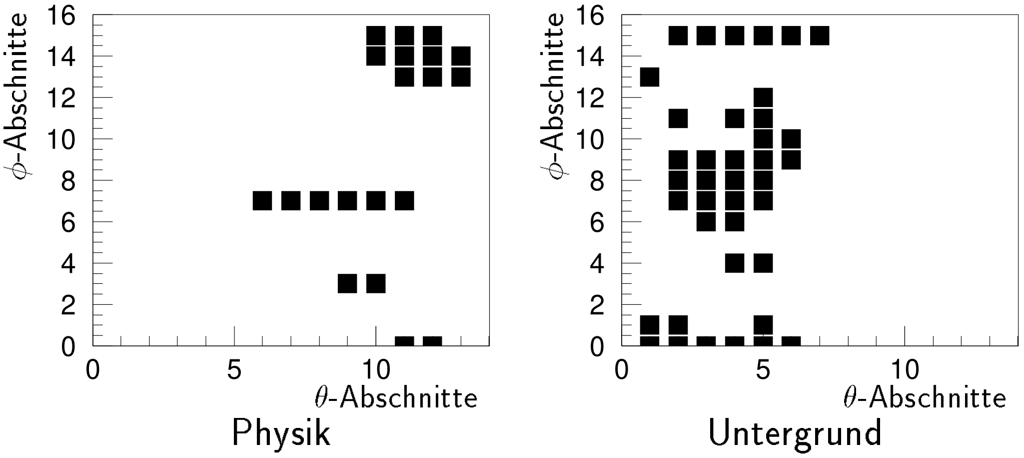


Abbildung 6.3: *Beispiele für die Treffer der* big rays *(Abschnitt 3.4.1.1) in der $\theta\phi$-Ebene; links ein Ereignis mit inklusiven Myonen, rechts eines aus dem Untergrunddatensatz.*

Anhang B). In der Abbildung 6.3 ist die *big ray*-Verteilung in der $\theta\phi$-Ebene für zwei Beispiele dargestellt. Es ist das oben beschriebene Verhalten zu erkennen: Das Physikereignis (links) besitzt mehr *big rays* in Bereichen mit größerem Polarwinkel als das Untergrundereignis (rechts). Dasselbe wird in den Polarwinkelverteilungen der *big rays* in Abbildung 6.4 deutlich. Aufgrund der Unterschiede in den Polarwinkelverteilungen von „Physik“ und Untergrund wird die $\theta\phi$-Ebene in vier θ-Bereiche unterteilt. Die Anzahl der *big rays* wird in den so entstandenen Bereichen gezählt und als Eingabewert für die Netze genutzt (Tabelle 6.1). Bei der Unterteilung im Polarwinkel ist auf die Realisierbarkeit in der Hardware geachtet worden.

Größe	θ-Abschnitte der $\theta\phi$-Ebene
nbigfwd	0 bis 3 , Vorwärtsbereich
nbigfce	4 bis 7 , vorderer Zentralbereich
nbigbce	8 bis 10 , hinterer Zentralbereich
nbigbwd	11 bis 13 , Rückwärtsbereich

Tabelle 6.1: *Neuentwickelte Eingabegrößen und die zugehörigen Abschnitte in der $\theta\phi$-Ebene der* big rays*; für die zugehörigen Winkel sei auf Tabelle B im Anhang verwiesen.*

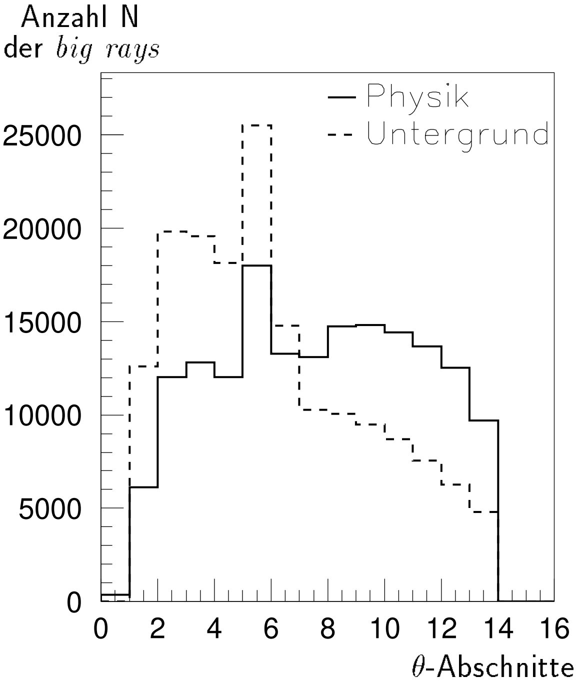


Abbildung 6.4: *Verteilungen der Polarwinkel θ der* big rays *für Physik (durchgezogen) und Untergrund (gestrichelt).*

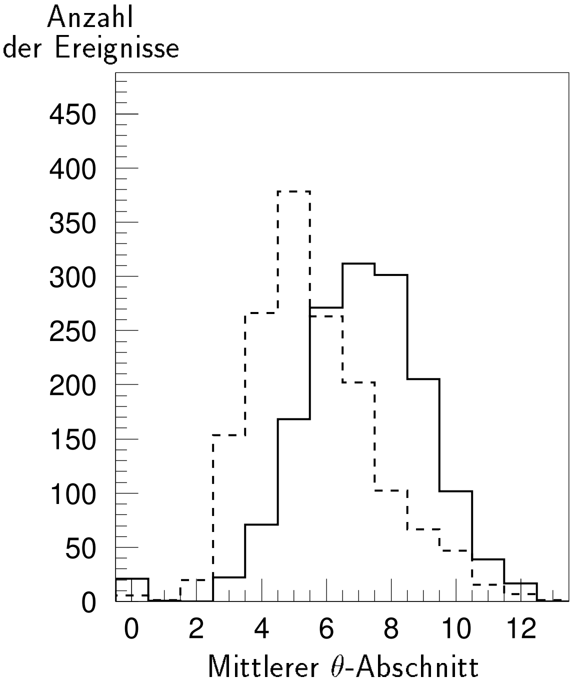


Abbildung 6.5: *Verteilungen des mittleren θ-Abschnitts der* big rays *für Physik und Untergrund.*

In Abbildung 6.6 sind die Verteilungen der neuen Eingabegrößen dargestellt. Wie erwartet zeigt die „Physik" mehr Ereignisse bei einer größeren Anzahl von *big rays* (BR) in den rückwärtigen Bereichen, während Untergrundereignisse im Mittel eine höhere Anzahl bei den das vordere Gebiet abdeckenden Größen besitzen.

Neben den durch die Aufteilung in θ entstandenen Größen ist der mittlere Polarwinkel *mbigthe* der *big rays* eines Ereignisses untersucht worden. Das mittlere θ der *big rays* ist für Untergrund gegenüber dem der Physik zu kleinen Polarwinkeln verschoben (Abbildung 6.5). *mbigthe* konnte in die Programme der Trainingskette implementiert werden, jedoch ist sie in der Triggerhardware auf L2 nicht realisierbar. Somit ist das mittlere θ wieder aus dem Training herausgenommen worden.

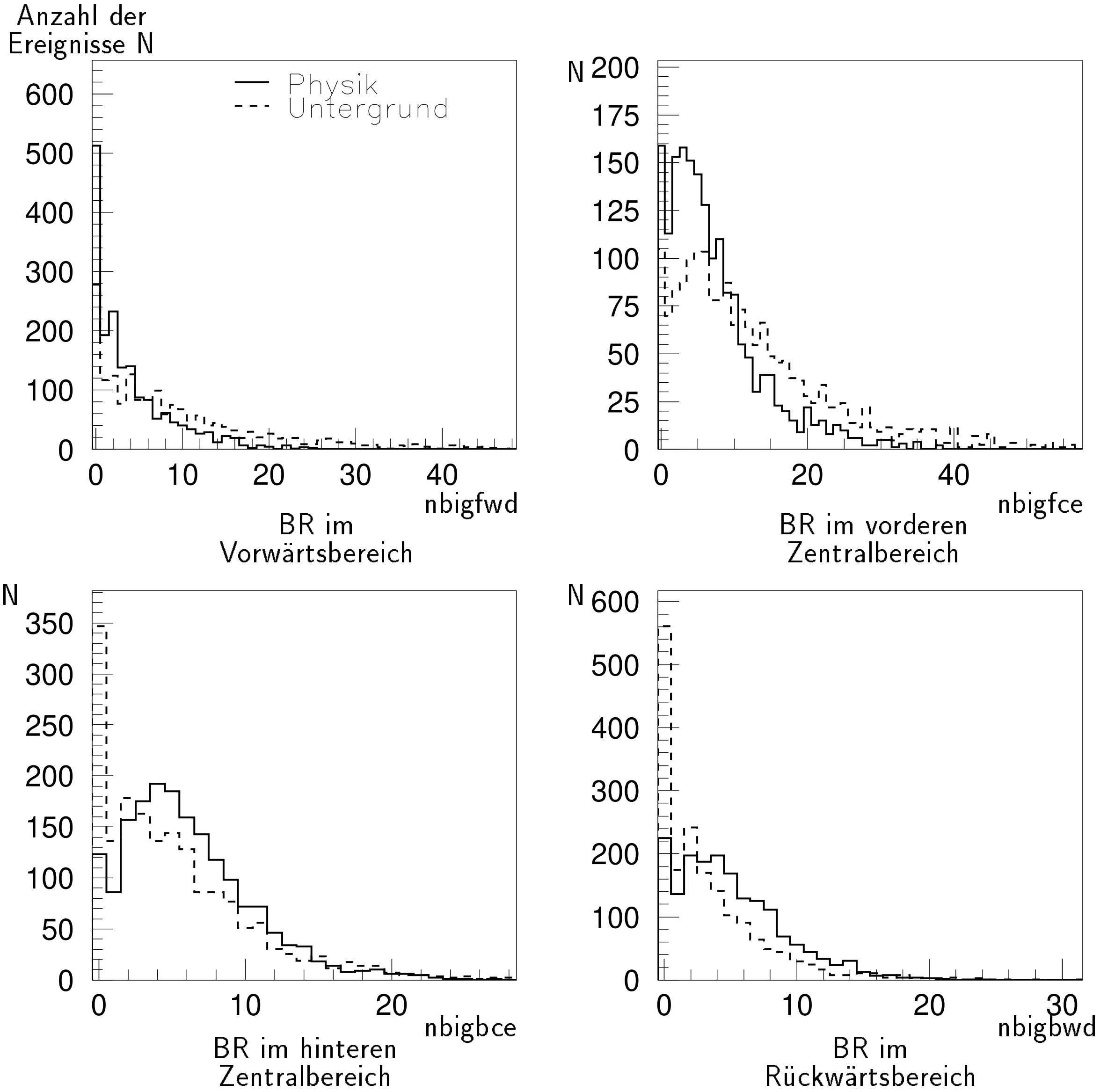


Abbildung 6.6: *Verteilungen von Physik (durchgezogene Linie) und Untergrund (gestrichelt) für die neuentwickelten Eingabegrößen* nbigfwd, nbigfce, nbigbce *und* nbigbwd.

6.3 Training und Tests der Netze

Nachdem im vorhergehenden Abschnitt die Eingabegrößen für die Netze beschrieben worden sind, werden hier das Training der Netze und die Resultate dargestellt. Dabei ist zu beachten, daß insgesamt drei Netze trainiert worden sind:

- je ein Netz für die beiden physikalischen Kanäle: unelastische Produktion von J/ψ-Mesonen im myonischen Zerfallskanal und Erzeugung inklusiver Myonen
- ein kombiniertes Netz für einen Physikdatensatz, der eine Kombination der beiden Kanäle umfaßt

Außerdem bilden zwei L1-Subtrigger die Grundlage der neuronalen Netze. Effizienzen können also für verschiedene Kombinationen der Daten und Subtrigger bestimmt werde, so z.B. eine Effizienz für die Ereignisse mit unelastischen J/ψ-Kandidaten bezüglich des Netzes für Produktion inklusiver Myonen. Auch kann die Kombination der beiden Einzelnetze betrachtet und mit dem kombinierten Netz verglichen werden.

6.3.1 Training der Netze

Es sind Netze für Ereignisse mit inklusiver Myonen-Produktion (Netz M), mit unelastischen J/ψ-Kandidaten (Netz J) sowie ein Netz für beide Physikkanäle zusammen (Netz K) trainiert worden.

Beim Training sind für alle drei Netze „Physik“ und Untergrund im Verhältnis 1:1 berücksichtigt worden. Die beiden physikalischen Kanäle sind für das kombinierte Netz K ebenfalls im Verhältnis 1:1 einbezogen worden.

Bei dem Netz M ist darauf geachtet worden, daß am Ende in etwa der gleiche Bruchteil an Ereignissen, die von S19 oder S22 akzeptiert worden sind, vom Netz erkannt wird. Deshalb sind die Anteile von S19 und S22 im Physikdatensatz im Verhältnis 2:1 gemischt worden. Die Mischungsverhältnisse sind in Tabelle 6.2 zusammengefaßt.

Die im *TRAINER* einstellbaren Parameter γ und α (Abschnitt 6.1) sind in verschiedenen Trainingsdurchgängen variiert worden. Als Beispiel für die Wirkung werden einige Ergebnisse aus der Generalisierung (Abschnitt 6.1) des Netzes M mit dem Testdatensatz der von S19 akzeptierten Ereignisse vorgestellt. In Abbildung 6.7 a) ist die Effizienz der Physikerkennung gegen die Untergrundreduktion in der Generalisierung für verschiedene

	inkl.μ		J/ψ 95	J/ψ 96	Untergrund
Netz M	S19: 2	S22: 1	-	-	3
Netz J	-		1	1	2
Netz K	2		1	1	4

Tabelle 6.2: *Gewichtung der Datensätze beim Netztraining.*

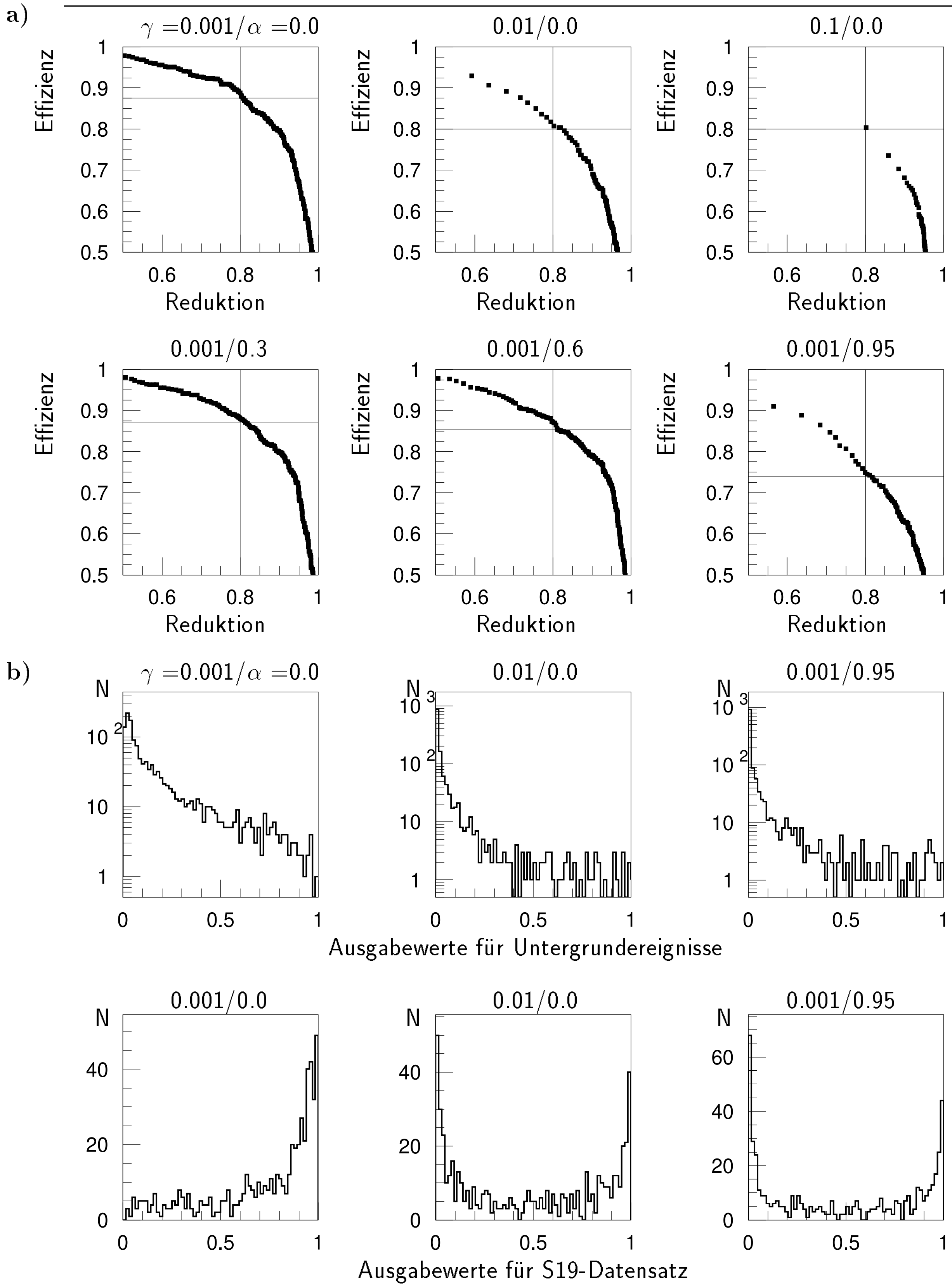


Abbildung 6.7: *a) Effizienz gegen Untergrundreduktion nach der Generalisierung für von S19 akzeptierte Ereignisse mit inklusiver Myon-Produktion bei verschiedenen Kombinationen von Lernparameter γ und Momentumfaktor α (Abschnitt 4.3). Die Effizienz nimmt bei gleicher Reduktion (80%) mit steigendem Momentumfaktor und größerem Lernparameter ab. b) Ausgabewertverteilungen von Untergrund (dritte Zeile) und „Physik“ (unterste Zeile) für die angegebenen Werte γ und α.*

Parameterkombinationen γ und α aufgetragen. Es ist zu erkennen, daß bei gleicher Reduktion (etwa 80%) die Effizienz der „Physik“ mit steigendem Momentumfaktor α sowie mit steigendem Lernparameter γ abnimmt. Auch verschiebt sich der Mittelwert des Ausgabewertes für die Physikereignisse zu Null hin (Abbildung 6.7 b)). Die Erkennung der „Physik“ ist also geringer. Ein ähnliches Verhalten ist auch für die anderen beiden Netze beobachtet worden.

Aufgrund dieser Untersuchungen sind ein Momentumfaktor von 0.0 für alle drei Netze und Lernparameter von 0.001 für das Netz K, 0.002 für das Netz J und 0.0015 für das Netz M ausgewählt worden.

Außerdem ist die Anzahl der Eingabegrößen und der versteckten Knoten in verschiedenen Trainingsdurchläufen verändert worden. Diese Variationen sind Gegenstand der Abschnitte 6.3.2 und 6.3.5. Für weitere Untersuchungen werden die in Tabelle 6.3 aufgeführten Netze verwendet.

	Netz M (inkl.μ)	Netz J (J/ψ)	Netz K (kombiniert)
Anzahl der Eingabeknoten	22	22	22
Anzahl der versteckten Knoten	12	13	14
Anzahl der Ausgabeknoten	1	1	1

Tabelle 6.3: *Netzgeometrien für die drei trainierten Netze*

6.3.2 Einfluß der versteckten Knoten

Um den Einfluß der versteckten Knoten auf die Effizienz zu untersuchen, ist deren Anzahl Knoten um Knoten im Training für das Netz M gesenkt worden. Die anderen Parameter sind dabei nicht verändert worden. Als Vergleichskriterium wird die Physikeffizienz der Generalisierung bei einer Untergrundreduktion von etwa 80% betrachtet (Abbildung 6.8). Die Effizienz verändert sich für die S19-Daten mit der Anzahl der versteckten Knoten nur unwesentlich. Aufgrund der Fehler kann es sich bei dem zu erkennenden leichten Anstieg bei kleiner Knotenzahl um Fluktuationen handeln. Es reicht also ein Netz mit einem Knoten aus, um die von S19 akzeptierten Ereignisse vom Untergrund zu trennen. Dies entspricht einer Teilung des multidimensionalen Raumes der Eingabegrößen in eine „Physik“- und eine Untergrundhälfte mit einer einzigen Hyperebene.

Die von S22 akzeptierten Ereignisse zeigen ein anderes Verhalten: Die Effizienz nimmt mit der Anzahl der versteckten Knoten von 72% auf im Mittel 86% zu. Es wird hier eine Sättigung der Effizienz ab etwa zwölf versteckten Knoten erreicht. Für eine Unterscheidung von Untergrund und von S22 angenommenen „Physik“ sind mehrere Ebenen im Raum der Eingabegrößen notwendig.

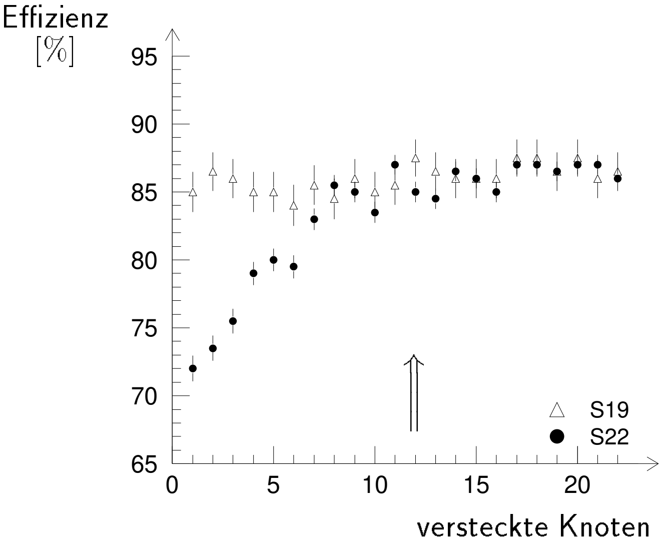


Abbildung 6.8: *Effizienz für von S19 und S22 akzeptierten Daten nach der Generalisierung in Abhängigkeit von der Anzahl der versteckten Knoten für das Netz M . Die Effizienzen des im weiteren verwendeten Netzes sind durch den Pfeil markiert.*

6.3.3 Effizienzen aus Training und Generalisierung

Für die in Tabelle 6.3 aufgeführten Netze sind in den Abbildungen 6.9 bis 6.11 die Ausgabewertverteilungen der verschiedenen Datensätze nach dem Training sowie nach der im *TRAINER* implementierten Generalisierung gezeigt. Die Verteilungen des im Training verwendeten Untergrundes besitzen für alle drei Netze ihr Maximum bei Null und zeigen nur wenige Einträge für Werte größer als 0.4.
Die „Trainings-Physik“ nimmt für alle drei Netze ein Maximum bei dem idealen Ausgabewert Eins an. Somit ist im Training der drei Netze eine Trennung von Physik und Untergrund erreicht worden.
Ob ein Netz das Gelernte auf ihm unbekannte Ereignisse übertragen kann, wird mit den Testdatensätzen im Generalisierungsschritt untersucht. Die Ausgabewerte hiervon sind in den unteren Zeilen der Abbildungen 6.9 bis 6.11 gezeigt und werden im folgenden diskutiert:

- **Netze für Ereignisse mit inklusiver Myon-Produktion**
  In Abbildung 6.9 ist zu erkennen, daß die Physikdatensätze für S19 und S22 auch in der Generalisierung deutlich vom Untergrund getrennt werden. Der Anteil mit einem Ausgabewert kleiner als 0.5 ist jedoch für S22 größer als für S19. Die Erkennung der von S22 akzeptierten Ereignisse ist also schlechter als die der S19-Ereignisse. Dies spiegelt sich auch in einer geringeren Effizienz der von S22 akzeptierten Daten bei einer Reduktion des Untergrundes von 80% wider (vgl. Abbildung 6.12).
  Die Effizienzen liegen etwa 3% unter dem im Training für die gleiche Untergrundreduktion erreichten Ergebnis. Wäre die Effizienz für die Testdaten deutlich geringer als der beim Training erreichte Wert, hätte das Netz die Trainingsereignisse nur „auswendiggelernt“. Dieses Verhalten würde auch in der Verteilung der Ausgabewerte zu erkennen sein. Die Ausgabewerte der Ereignisse würden im Mittel mehr zum Wert Null hin tendieren. Dieses ist hier nicht der Fall.

- **Netz für J/ψ-Ereignisse**
  Auch für das Netz J ist eine deutliche Erkennung von Physik- und Untergrundereignissen bei der Generalisierung gegeben (Abbildung 6.10). Die Ausgabewertverteilung hat aber für die Daten von 1995 ein ausgeprägteres Maximum als für die von 1996. Das führt zu einer um 5% geringeren Effizienz für die aus 1996 selektierten Ereignisse mit J/ψ-Kandidaten bei einer Untergrundreduktion von etwa 80% (Abbildung 6.12).
  Die Effizienz des aus 1995 selektierten Datensatzes hat ungefähr den gleichen Wert

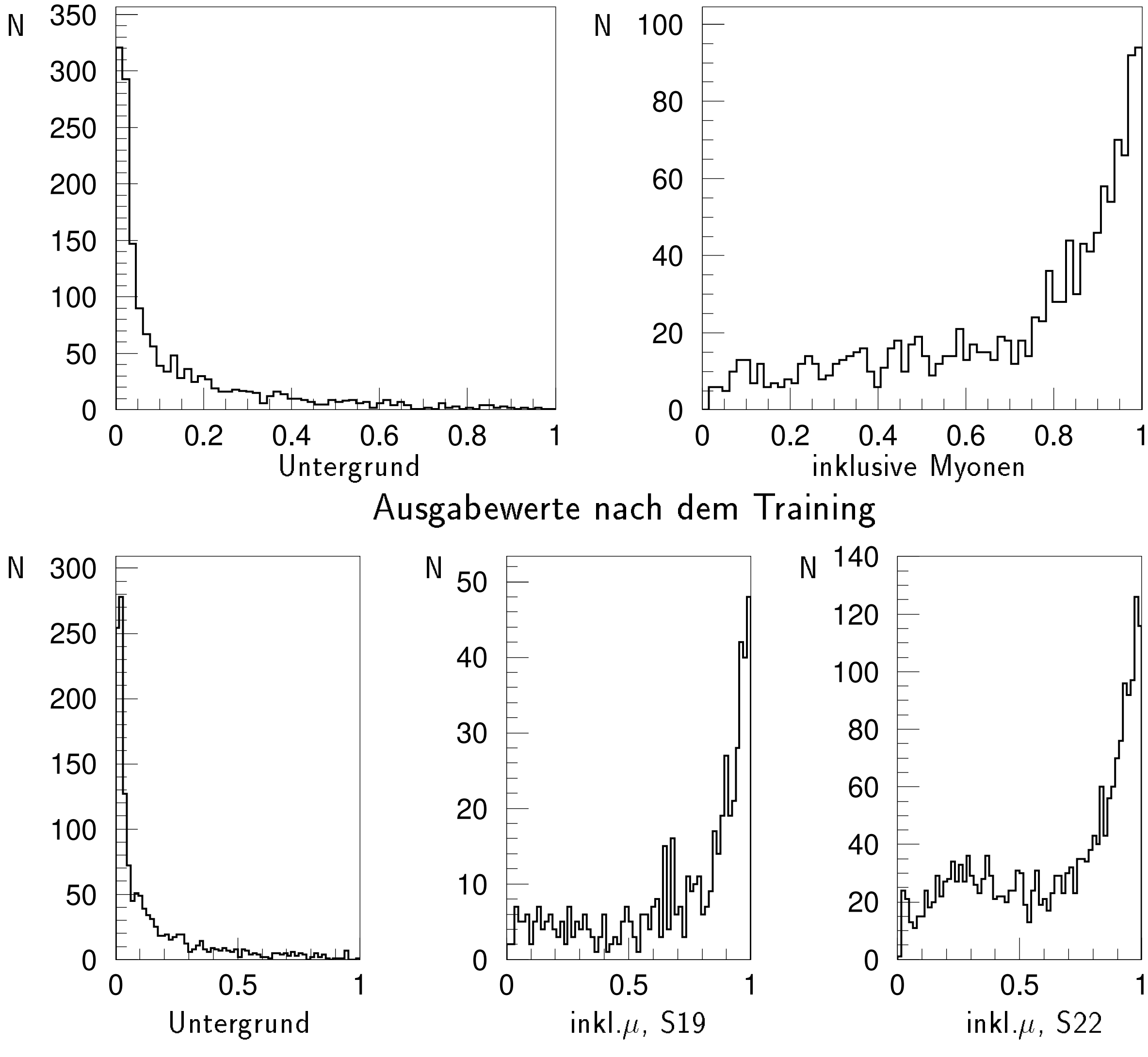
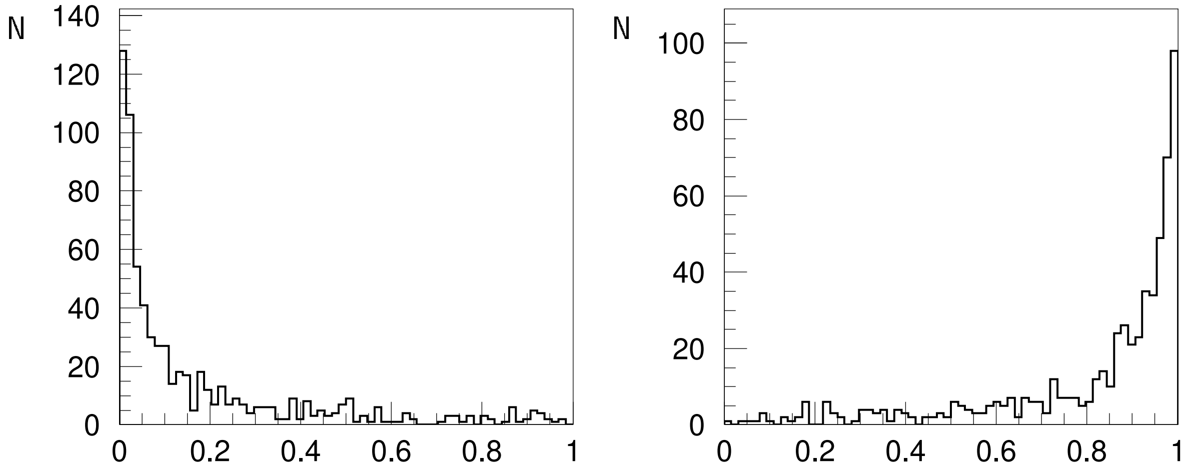


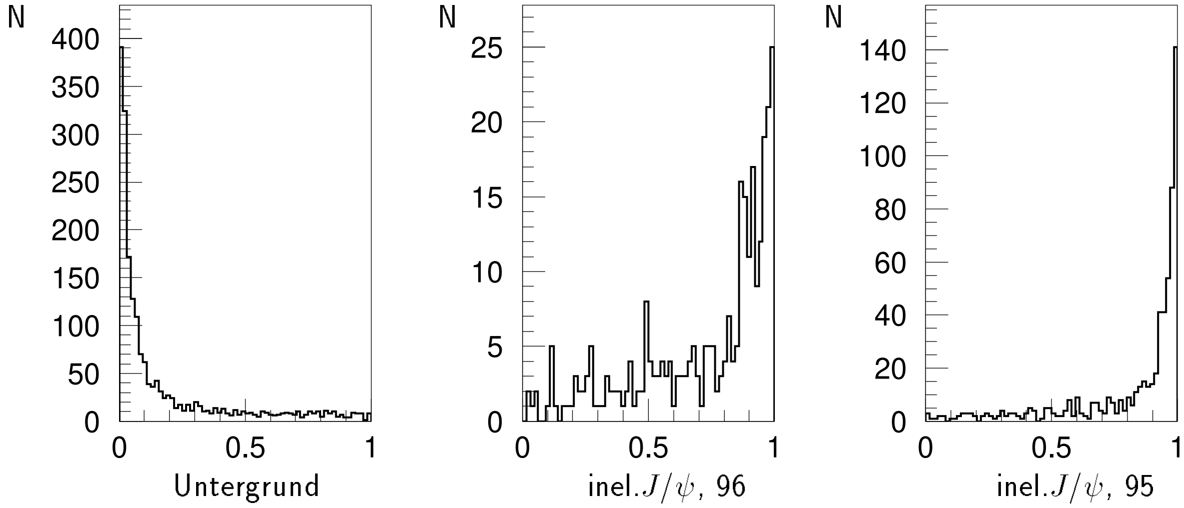
Abbildung 6.9: *Ausgabewert für das* **Netz M**, *oben nach dem Training, unten nach der Generalisierung. Die Verteilungen nehmen ihr Maximum jeweils bei den idealen Werten 0 für Untergrund und 1 für Physik an.*

wie die im Training erzielte. Die im Training gelernten Unterschiede von Physik und Untergrund werden somit vom Netz gut auf unbekannte Ereignisse übertragen.
Eine bessere Erkennung der 1995 genommenen Ereignisse kann darauf zurückgeführt werden, daß die für den Untergrund selektierten L4-*transparent runs* nur aus einem begrenzten Zeitraum von 1996 gewählt worden sind. Damit ist es möglich, daß das Netz für Unterschiede im Detektor und in den Untergrundbedingungen sensibel wird und diese lernt.

- **Kombiniertes Netz**
  Die Trennung der beiden physikalischen Kanäle von den Untergrundereignissen ist

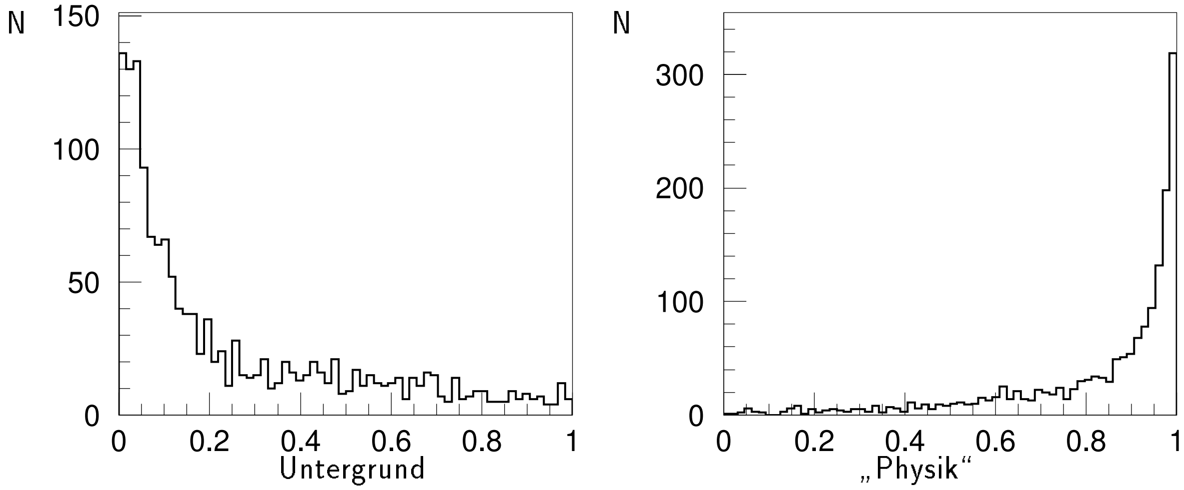


Ausgabewerte nach dem Training

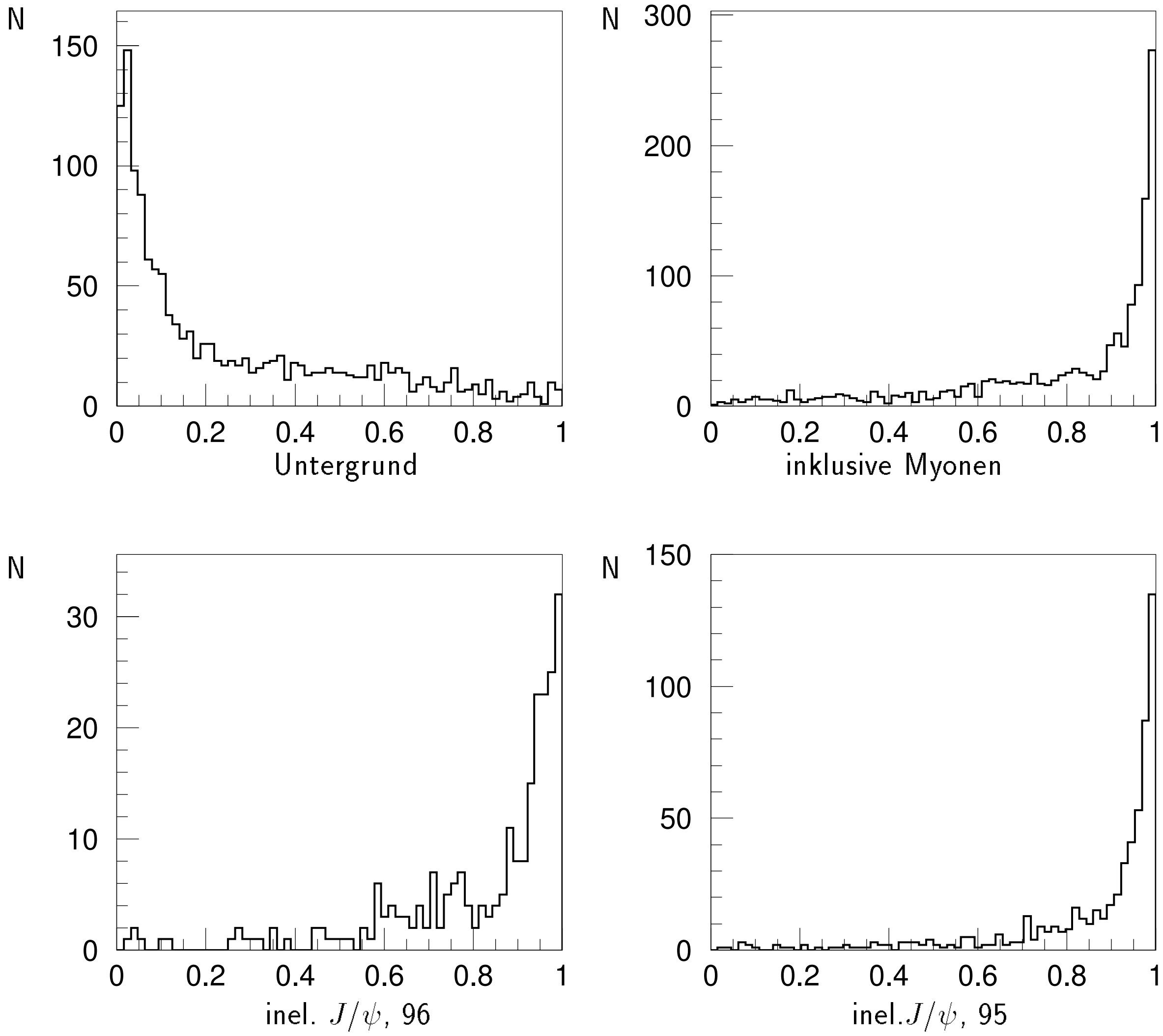


Ausgabewerte nach der Generalisierung

Abbildung 6.10: *Ausgabewertverteilung für das* **Netz J**.



Ausgabewerte nach dem Training



Ausgabewerte nach der Generalisierung

Abbildung 6.11: *Ausgabewertverteilung für das* **Netz K**. *Der mit „Physik“ bezeichnete Datensatz ist eine Mischung aus Ereignissen mit inklusiven Myonen und Ereignissen mit unelastischen J/ψ-Kandidaten im Verhältnis* $1:1$, *wie sie im Training verwendet worden ist.*

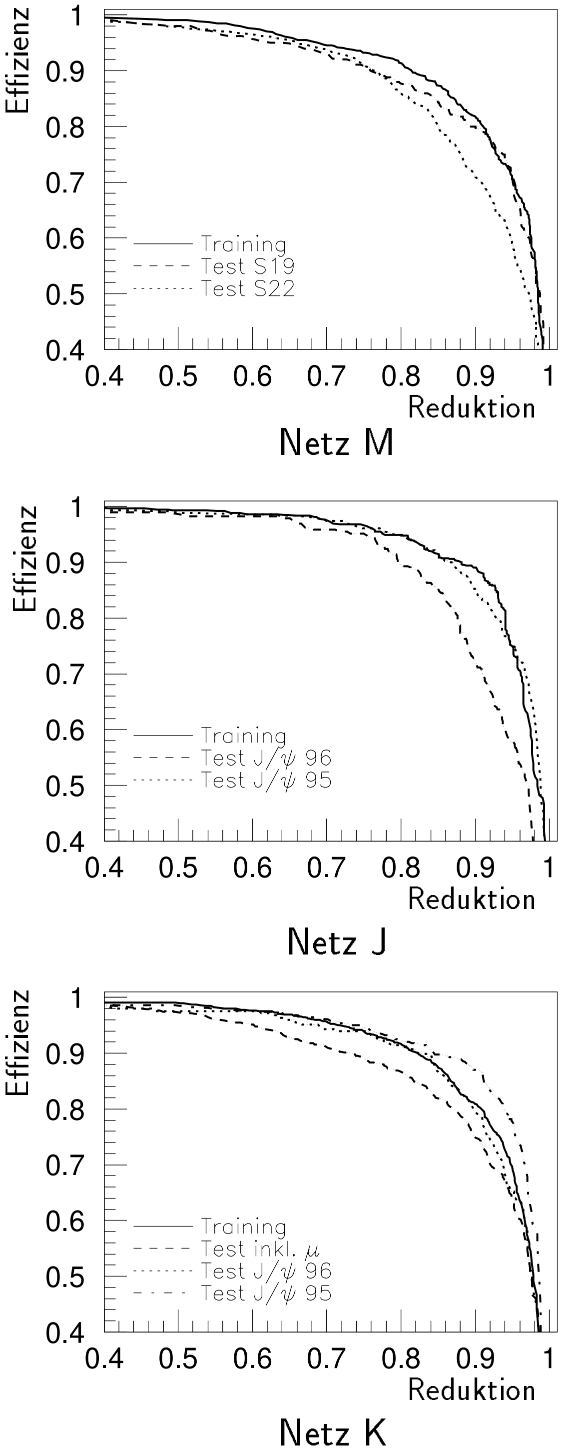


Abbildung 6.12: *Effizienz der „Physik“ aufgetragen gegen die Untergrundreduktion für die drei Netze. Für eine bestimmte Reduktion kann so die erreichte Effizienz abgelesen werden. Eine Reduktion von 0.8 bedeutet, daß der Untergrund um einen Faktor 5, also auf 20%, reduziert wird.*

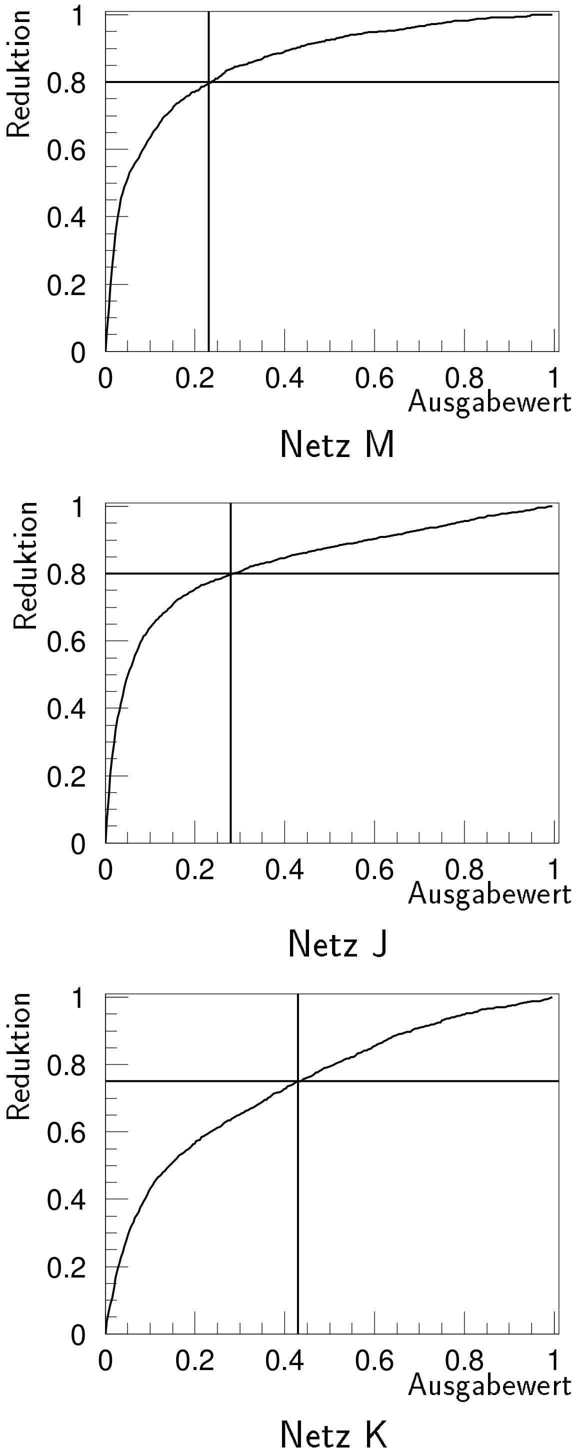


Abbildung 6.13: *Reduktion des Untergrundes als Funktion des Ausgabewerts: Für eine bestimmte Reduktion kann hieraus der Schnitt auf den Ausgabewert des Netzes abgelesen werden. Die Schnitte bei den gewünschten Reduktionen sind für die drei Netze in der Abbildung eingetragen. Oberhalb dieser Schnitte wird das Triggerbit 1 (=Physik), unterhalb 0 (=Untergrund) gesetzt.*

wie für die einzelnen Netze auch durch das kombinierte Netz K möglich (Abbildung 6.11). Die Effizienz der Zuordnung der inklusiven Myonen zur Physik liegt bei einer Untergrundreduktion von 75% um 5 bis 7% unter der der J/ψ-Kandidaten, bei denen die Ereignisse aus 1995 besser erkannt werden (Abbildung 6.12).
Die Werte der mit den Testdatensätzen erreichten Effizienz weichen um weniger als 4% von dem Ergebnis aus dem Training ab.

In Abbildung 6.13 sind die Untergrundreduktionen in der Generalisierung der drei Netze gegen den Ausgabewert des Netzes aufgetragen. Aus diesen lassen sich die zur Verwendung des Netzes als Trigger angebrachten Schnittstellen auf den Ausgabewert ablesen. Nur Ereignisse mit einem Ausgabewert größer als diese Schwelle werden als Physik akzeptiert. Alle Ereignisse mit Werten unterhalb des Schnittes werden verworfen. Für die beiden Netze M und J mit jeweils einem physikalischen Kanal werden die Schnitte so gewählt, daß 80% des Untergrundes verworfen werden. Diese Reduktion reicht aus, um die Raten der Subtrigger in 1996 auf 1 bis 1.5 Hz zu senken (Abschnitt 5.1). Für die gesamte Verringerung des Untergrundes durch beide Netze gemeinsam ergibt sich mit dieser Wahl ungefähr 72%.
Bei dem Netz K wird eine Untergrundunterdrückung auf 1/4 (Reduktion von 75%) zur Festlegung des Schnittwerts verwendet. Damit werden das Netz K und die Kombination der beiden einzelnen Netze M und J vergleichbar.

	Netz M	Netz J	Netz K
Schwelle	0.23	0.3	0.44
Reduktion	$\sim 80\%$	$\sim 80\%$	$\sim 75\%$

Tabelle 6.4: *Schnitte auf die Ausgabewerte für die drei Netze. Ist der Ausgabewert für ein Ereignis größer als dieser Schnittwert, so wird das Ereignis als „Physik" akzeptiert und bei der Verwendung des Netzes als Trigger das Triggerbit gesetzt.*

Die Schnitte auf den Ausgabewert der Netze sind in Tabelle 6.4 zusammengestellt. Mit diesen Schnitten sind die Netzentscheidungen für die selektierten Datensätze durch das Programm *NETSIM* simuliert worden. Tabelle 6.5 listet die dabei erreichten Effizienzen und die Untergrundreduktion für die beiden Subtrigger auf. Eine graphische Darstellung findet sich in Abbildung 6.14. Außerdem kann mit der Simulation untersucht werden, mit welchen Effizienzen Ereignisse mit inklusiver Myon-Produktion von Netz J und Ereignisse mit unelastischen J/ψ-Kandidaten vom Netz M erkannt werden.
Die durch das Netz M erreichte Effizienz der Ereignisse mit unelastischer J/ψ-Produktion ist mit 60 bis 70% für die Daten aus 1995 sehr niedrig. Für eine Erkennung beider von S19 und S22 akzeptierten physikalischen Kanäle mit hoher Effizienz ist also eine Kombination der Netze M und J bzw. die Verwendung des kombinierten Netzes K notwendig. Aus den Zahlen läßt sich auch ablesen, daß mit den beiden auf einen physikalischen Kanal spezialisierten Netzen M und J insgesamt eine höhere Effizienz als mit Netz K erreicht werden kann. Deshalb sind diese beiden Netze und nicht Netz K testweise in die L2-Trigger-Stufe implementiert worden. Das Erreichen höherer Effizienzen führt jedoch zu einem Verlust an Untergrundreduktion gegenüber Netz K von im Mittel 5%, für S19 von 12%.

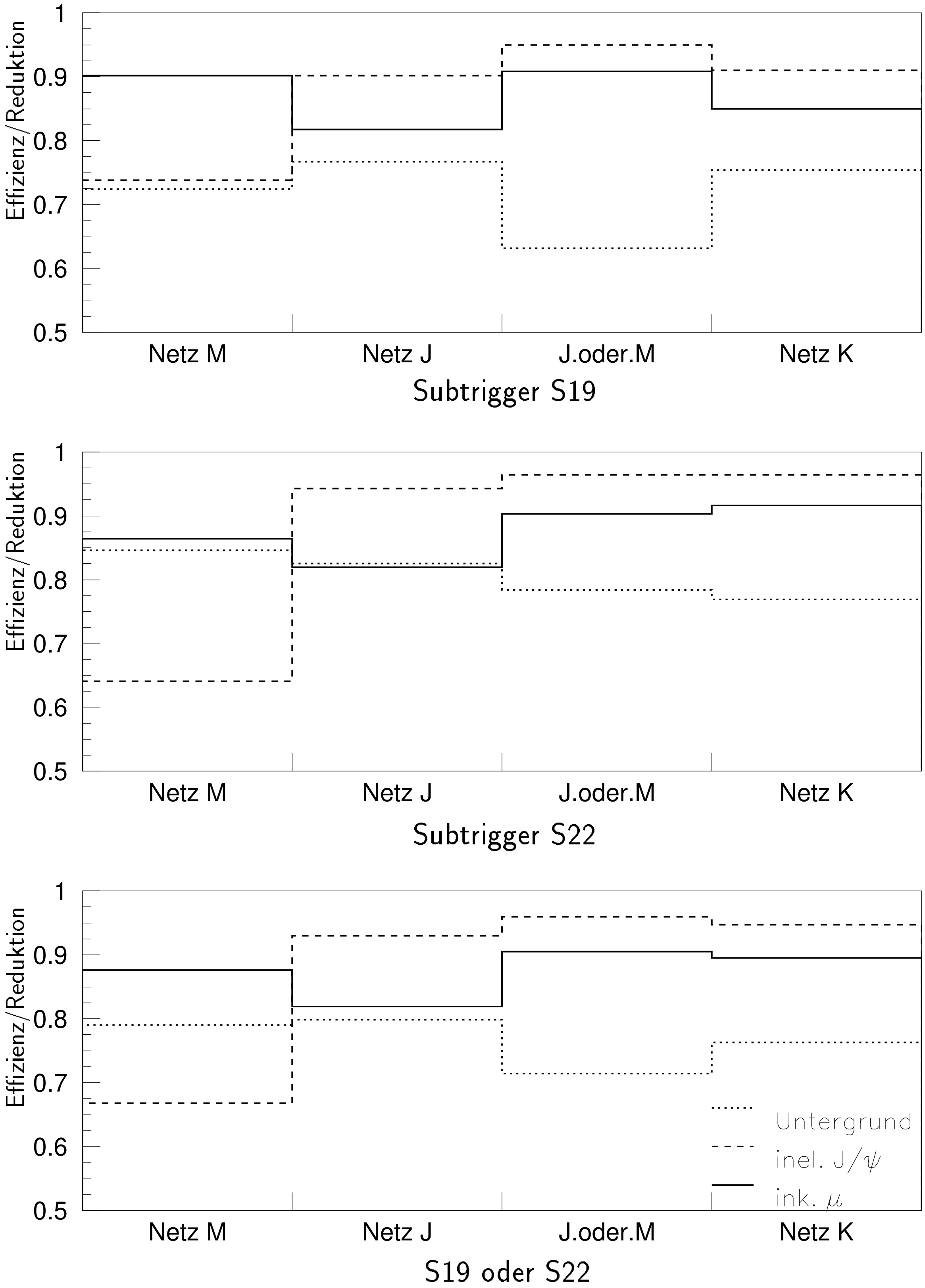


Abbildung 6.14: *Effizienzen der Subtrigger für die Physikdatensätze und Reduktion des Untergrundes (vgl. Tabelle 6.5).*

		Netz M	Netz J	Kombination der Netze M und J	Netz K	Anzahl der Ereignisse
Physikeffizienz in %						
inkl. μ	S19	90.2	81.7	90.9	85.0	1892
	S22	86.4	82.0	90.3	91.6	3806
inel. J/ψ 95	S19	69.1	90.6	94.7	90.6	340
	S22	60.4	95.3	96.7	96.7	780
inel. J/ψ 96	S19	80.4	89.6	95.4	91.7	240
	S22	70.4	92.5	96.3	96.0	453
Untergrundreduktion in %						
Untergrund	S19	72.4	76.7	63.2	75.4	2004
	S22	84.6	82.5	78.4	76.9	2268

Tabelle 6.5: *Effizienz und Reduktion der Netze mit den in Tabelle 6.4 angegebenen Schnitten.*

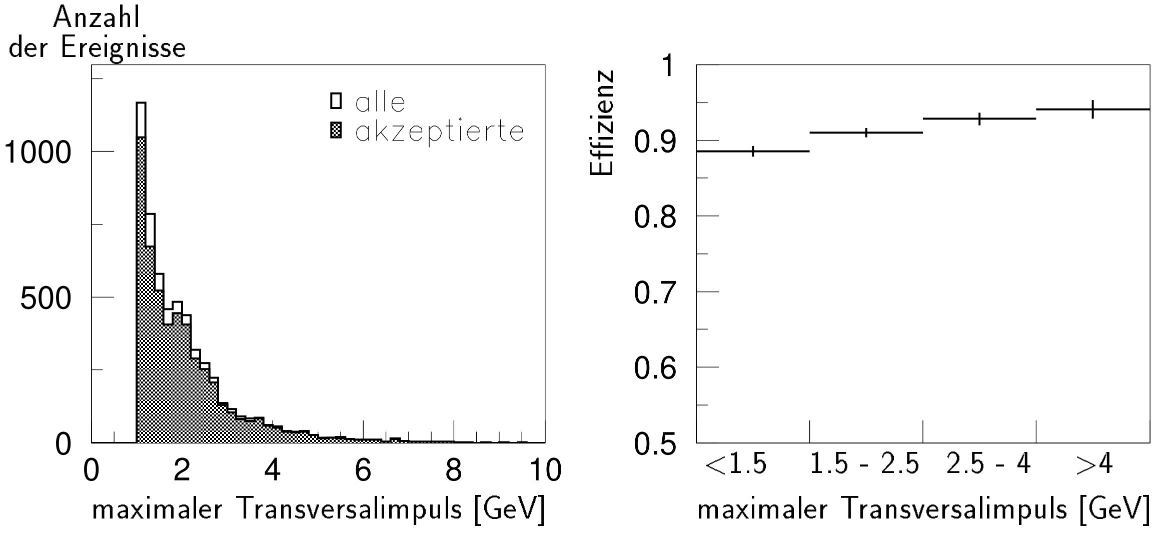


Abbildung 6.15: *Maximaler Transversalimpuls der im Myondetektor erkannten Myonen für die Ereignisse mit inklusiven Myonen vor und nach der Entscheidung der Netze M und J sowie die erreichte Effizienz für die Kombination der Netze M und J in Abhängigkeit von* $p_{t,max}$.

6.3.4 Effizienzen in Abhängigkeit von physikalischen Größen

In diesem Abschnitt werden die Effizienzen in Abhängigkeit von physikalischen Größen untersucht. Für die Ereignisse mit inklusiven Myonen wird der maximale Transversalimpuls der Myonen im Ereignis, für die J/ψ-Kandidaten die invariante Masse und der Transversalimpuls betrachtet.

Für die Ereignisse mit inklusiver Myon-Produktion wird pro Ereignis aus den Transversalimpulsen der im instrumentierten Eisen identifizierten Myonen der maximale Wert $p_{t,max}$ bestimmt. In Abbildung 6.15 ist die Verteilung des maximalen Transversalimpulses $p_{t,max}$ und die Effizienz für die Kombination der Netze M und J in Abhängigkeit von $p_{t,max}$ dargestellt. Sie nimmt mit größerem $p_{t,max}$ um etwa 5% zu.

Die Kombination der beiden Netze M und J akzeptiert mit einer Effizienz von etwa 95%

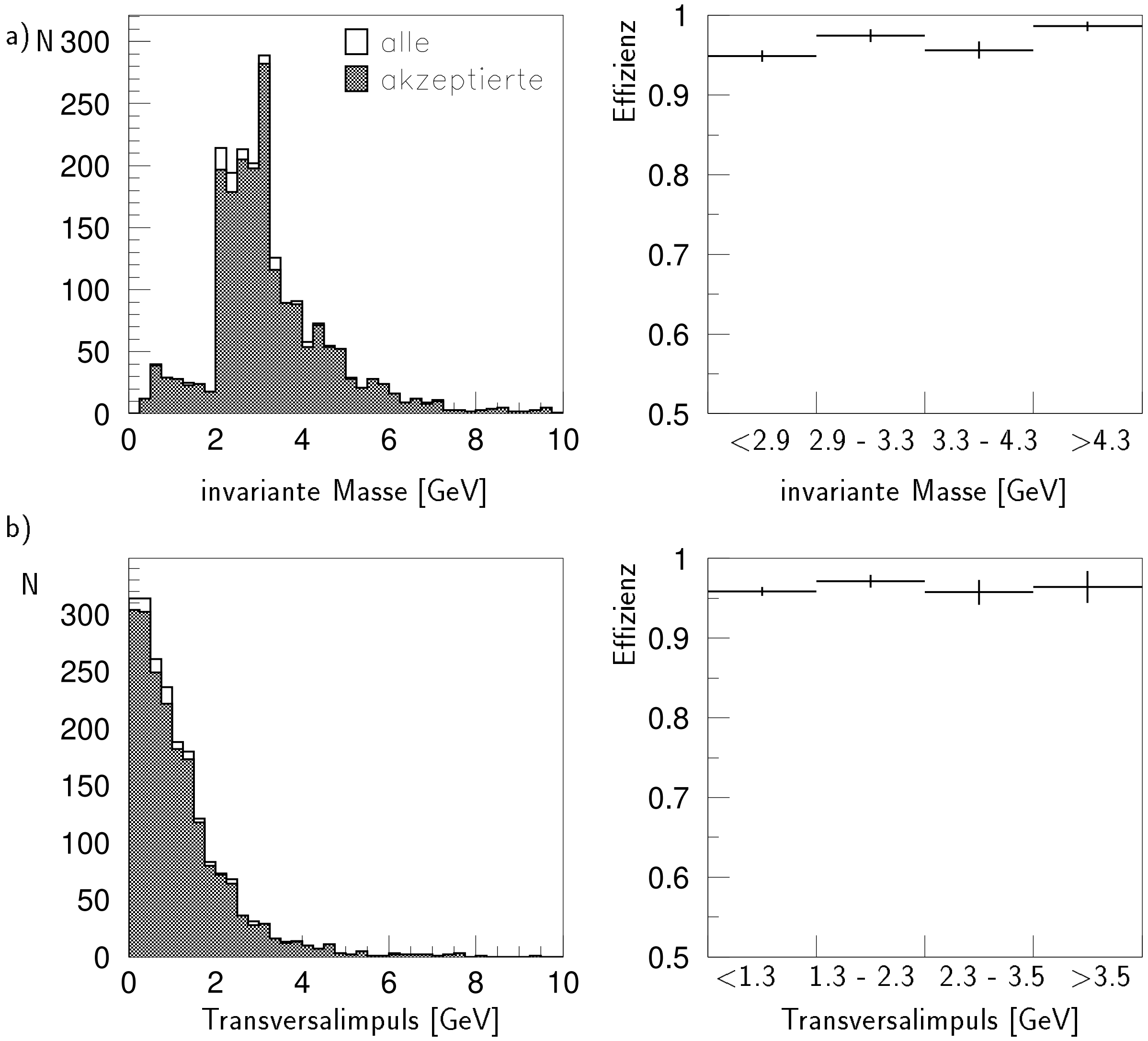


Abbildung 6.16: *a) Invariante Masse und b) transversaler Impuls der J/ψ-Kandidaten vor (weiß) und nach (grau) der Entscheidung der Kombination der Netze M und J sowie die zugehörigen Effizienzen in verschiedenen Massen- und $p_{t,J/\psi}$-Intervallen.*

die Ereignisse mit unelastischen J/ψ-Kandidaten. Abbildung 6.16 zeigt die Verteilungen der invarianten Masse $M_{\mu\mu}$ und des Transversalimpulses $p_{t,J/\psi}$ der J/ψ-Kandidaten für den gesamten Datensatz sowie für die akzeptierten Ereignisse. Man erkennt keine Bereiche, in denen überdurchschnittlich viele Ereignisse verworfen werden. Für die invariante Masse schwankt die Effizienz in einem Bereich von 3%.
Einträge für invariante Massen $M_{\mu\mu}$ kleiner als 2 GeV kommen dadurch zustande, daß bei der Selektion mindestens zwei Myonen mit $M_{\mu\mu}$ oberhalb von 2 GeV verlangt werden. Bei Ereignissen mit drei oder mehr Myonen können weitere Leptonpaare mit Massen $M_{\mu\mu}$ kleiner als 2 GeV gefunden werden (vgl. Abschnitt 5.2.3). Durch mehrere Leptonpaare in einem Ereignis kommt es bei der Bestimmung der Effizienz zu einer Doppelzählung von einigen Ereignissen.
Die Betrachtung der Transversalimpulsverteilung der J/ψ-Kandidaten ist aufgrund verschiedener Modelle über die Herkunft der J/ψ-Mesonen wichtig (vgl. Abschnitt 2.3). Die Beiträge dieser zum Wirkungsquerschnitt der J/ψ-Produktion wird in den Analysen in verschiedenen Transversalimpulsbereichen untersucht. Deshalb ist eine gleichmäßige Verteilung der Effizienz über $p_{t,J/\psi}$ wünschenswert. In Abbildung 6.16 ist keine Abhängigkeit der Effizienz von dem Transversalimpuls $p_{t,J/\psi}$ der J/ψ-Kandidaten zu erkennen. Sie bewegt sich für die einzelnen Impulsintervalle zwischen 95% und 97%.

6.3.5 Die Relevanz der Eingabegrößen

Die Entscheidung neuronaler Netze beruht auf mehreren Eingabegrößen. Mit dem Programm *TESTER* kann die Bedeutung der einzelnen Eingabegrößen untersucht werden.
Hierzu wird das Netzergebnis y_0 für jedes Ereignis eines Datensatzes mit dem auf seinen Mittelwert gesetzten i-ten Eingabewert $\bar{x}_i$ berechnet. Von diesem Ergebnis wird die Differenz zum „normalen“ Ausgabewert, der mit dem tatsächlichen Wert der Eingabegröße x_i bestimmt wird, gebildet. Die normierte quadratische Summe aller N_E im Datensatz verfügbaren Ereignisse ergibt dann die Relevanz R_i der Eingabegröße i:

$$R_i = \frac{1}{N_E} \sum_{j=1}^{N_E} (y_0(\vec{x}_j) - y_0(\vec{x}_j|_{x_{j,i}=\bar{x}_i}))^2 \qquad (6.1)$$

$\vec{x}_j$ ist dabei der Vektor der Eingabegrößen des Ereignisses j. Mit dieser Größe kann der quantitative Einfluß der Eingabegrößen auf das Netz bewertet werden.
Die so definierte Relevanz berücksichtigt weder Korrelationen zwischen den Eingabegrößen noch die Anzahl der Knoten des Netzes. Außerdem hängen die Eingabewerte und ihre Mittelwerte vom verwendeten Datensatz ab. Somit ist ein Vergleich der Relevanzen nur innerhalb eines Netzes sinnvoll, nicht zwischen verschiedenen Netzen [WesD96, Möck94].
In den Abbildungen 6.17, 6.18 und 6.19 sind die Relevanzen der Eingabegrößen für die Trainingsdatensätze der drei Netze dargestellt. In allen drei Fällen sind die Größen *cpvsum, cpvmax* und *ironfe* (vgl. Abschnitt 6.2.1) am wichtigsten. Die neu erarbeiteten Größen, unter ihnen insbesondere *nbigfwd* und *nbigbce* (vgl. Abschnitt 6.2.2), haben mittlere, im Falle des Netzes J großen Einfluß auf die Entscheidung der Netze im Vergleich zu den anderen Größen.

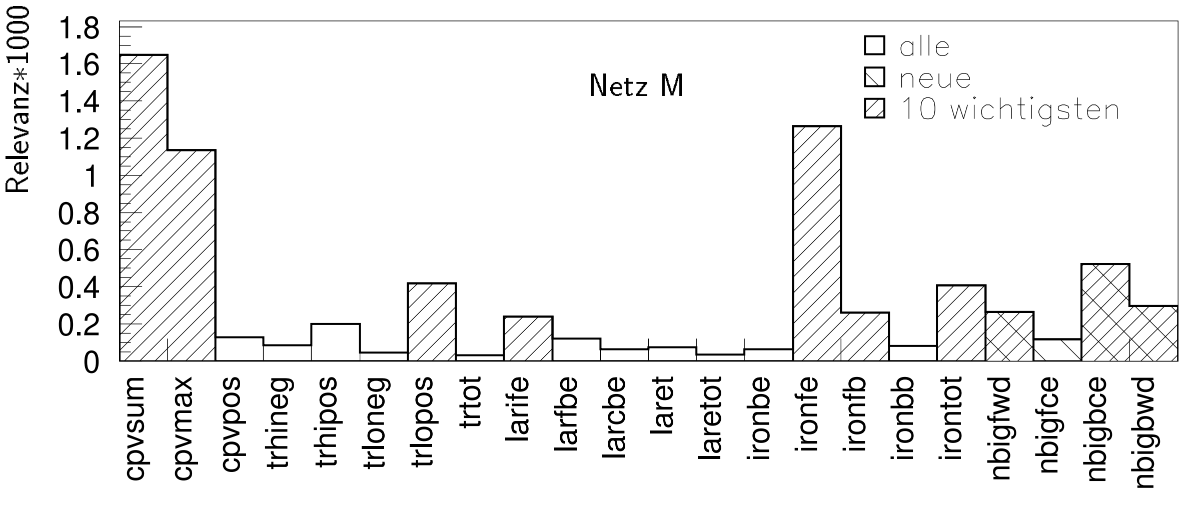


Abbildung 6.17: *Relevanz für den Trainingsdatensatz des Netzes M. Schraffiert sind einmal die neuentwickelten Eingabegrößen (Abschnitt 6.2.2), zum anderen die 10 wichtigsten Größen beim Training des Netzes.*

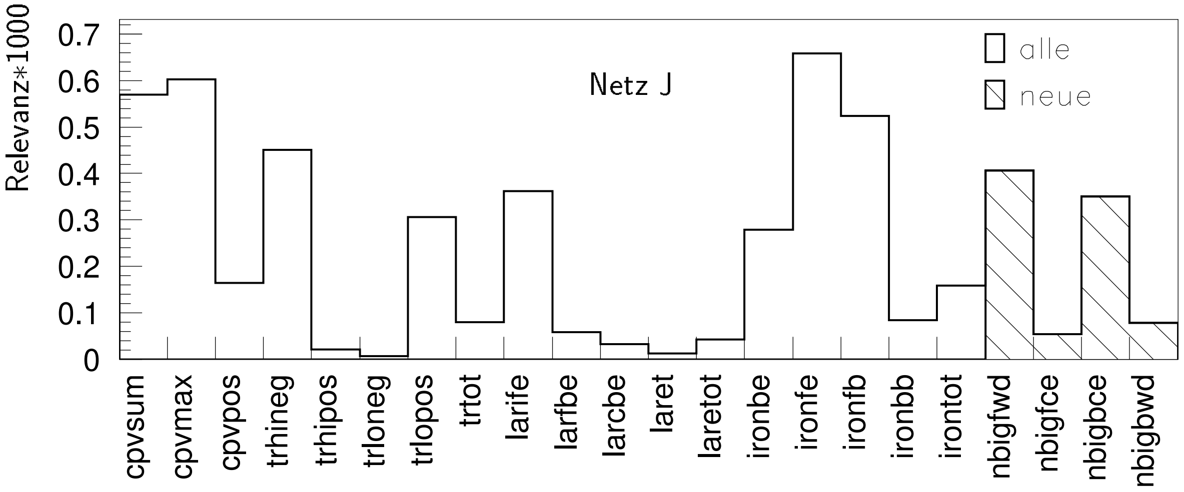


Abbildung 6.18: *Relevanz für den Trainingsdatensatz des Netzes J. Schraffiert sind die neuentwickelten Eingabegrößen.*

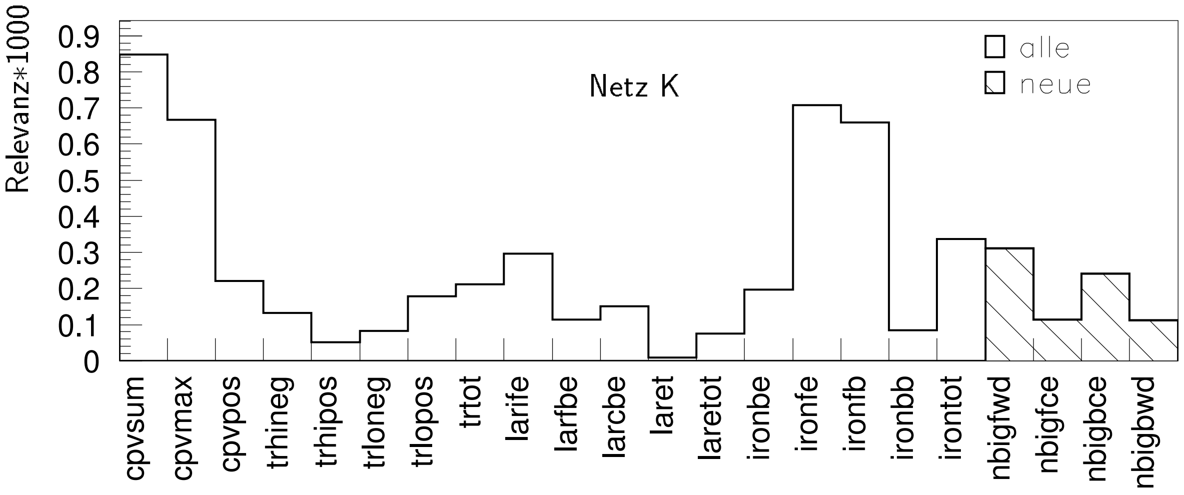


Abbildung 6.19: *Relevanz für den Trainingsdatensatz des Netzes K. Schraffiert sind die neuentwickelten Eingabegrößen.*

Netz $\tilde{\mathrm{M}}$				
	$N_{Eingabegrößen}$	$N_{versteckte}$	Lernparameter γ	Momentum α
	10	9	0.0015	0.0

	Effizienz bei $\sim$ 80% Reduktion			
	$N_{Eingabegrößen}$	Training	S19	S22
Netz $\tilde{\mathrm{M}}$	10	$\sim$ 85%	$\sim$ 84%	$\sim$ 80%
Netz M	22	$\sim$ 91%	$\sim$ 88%	$\sim$ 86%

Tabelle 6.6: *Netz $\tilde{M}$ für inklusive Myon-Produktion mit einer auf 10 Größen reduzierten Anzahl von Eingabeknoten. Angegeben sind die beim Training eingestellten Parameter. Die erreichten Effizienzen werden mit denen des Netzes M mit 22 Eingabegrößen für eine Untergrundreduktion von 80% verglichen.*

Man kann versuchen, weniger Eingabegrößen bei der Netzwerkentscheidung zu verwenden. Dabei müssen der Lernparameter γ, der Momentumfaktor α und die Anzahl der versteckten Knoten neu optimiert werden. Für Ereignisse mit inklusiver Myon-Produktion wurde mit den wichtigsten 10 Eingabegrößen, wie sie sich aus den Relevanzen in Abbildung 6.17 ergeben, ein weiteres Netz $\tilde{\mathrm{M}}$ trainiert. In Tabelle 6.6 wird das Netz $\tilde{\mathrm{M}}$ vorgestellt. Die Effizienzen für das Netz $\tilde{\mathrm{M}}$ gehen in Vergleich zum Netz M um 4-6% zurück. Die Ausgabewertverteilung und die Effizienzen als Funktion der Untergrundreduktion sind in Abbildung 6.20 gezeigt.

Bei dem Training des Netzes ist wiederum die Anzahl der versteckten Knoten von eins bis zehn variiert worden. Es wird dabei das gleiche Verhalten wie bei 22 Eingabeknoten festgestellt (Abschnitt 6.3.2):

- Die Effizienz für S19 steigt mit der Knotenanzahl nur wenig, schwankt in einem Bereich von wenigen Prozent.
- Die Effizienz für S22 fällt mit abnehmender Knotenzahl auf unter 70% ab.

6.3.6 Netze in der Datennahmeperiode 1997

Die beiden Netze M und J, die nur mit Ereignissen mit inklusiver Myon-Produktion bzw. unelastischen J/ψ-Kandidaten trainiert worden sind, wurden Anfang April 1997 in der zweiten Triggerstufe des H1-Triggersystems implementiert[1]. Es wird nur die Netzentscheidung für weitere Tests in den Daten festgehalten, die Netze verwerfen jedoch nicht von den Subtriggern S19 oder S22 akzeptierte Ereignisse.

Mit den gleichen Selektionskriterien (Tabelle 5.3), die bei der Auswahl der Trainings- und Testdatensätze angewandt worden sind, werden in einem begrenzten Zeitraum der 97er Datennahmeperiode 4461 Ereignisse mit inklusiven Myonen und 534 Ereignisse mit unelastischen J/ψ-Kandidaten selektiert. Bei diesen wird die Triggerentscheidung der beiden Netze M und J einzeln und die Kombination beider Entscheidungen ausgewertet. Die Ergebnisse sind in Tabelle 6.7 und in Abbildung 6.21 dargestellt.

---

[1] ab *run* 182760 auf den Triggerboxen 5 (Netz J) und 7 (Netz M)

Bei der Analyse der Untergrundreduktion werden wie zum Training der Netze L4-*transparent runs* verwendet. Aus ihnen werden 921 von S19 oder S22 akzeptierte Ereignisse entnommen.

Die Effizienzen der Ereignisse mit inklusiver Myon-Produktion stimmen gut mit den im Training erreichten überein. Für S22 liegt sie für die Kombination der Netze etwa 3% über dem in Training und Test erreichten Wert.

Die Ereignisse mit unelastischer J/ψ-Produktion werden 1997 im Vergleich zu den Trainings- und Testergebnissen vom Netz M deutlich besser erkannt. Die Effizienz beträgt im Mittel 90% gegenüber 67% im Test. Im Gegensatz dazu sinkt sie für das Netz J von 90% auf 83% für S19 und von 94% auf 91% für S22. Ein Grund für dieses veränderte Verhalten ist möglicherweise, daß sich die Polarwinkelverteilung der *big rays* in den Jahren '95, '96

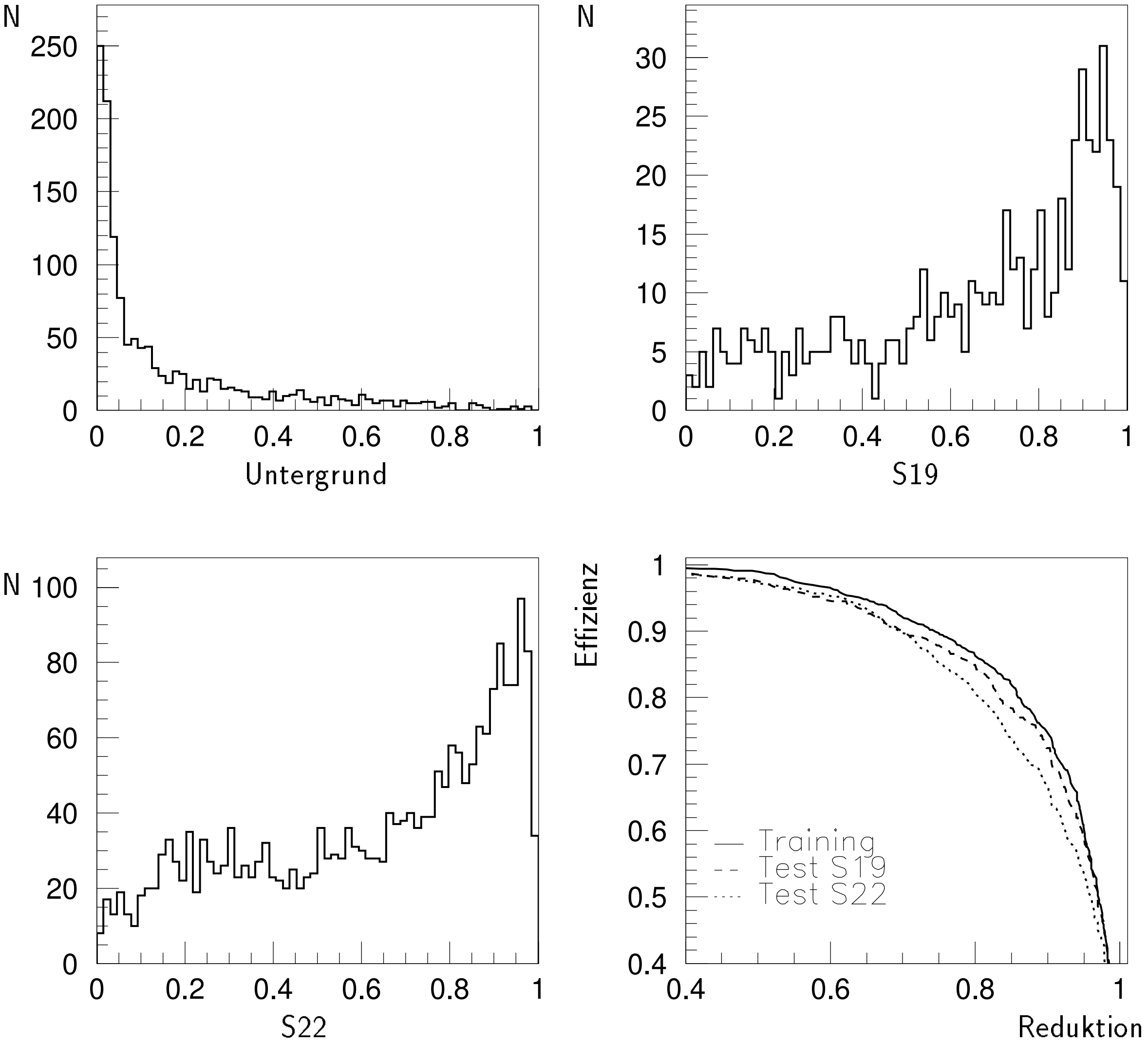


Abbildung 6.20: *Ausgabewertverteilungen nach der Generalisierung für das mit den 10 wichtigsten Eingabegrößen (Abbildung 6.17) trainiertes Netz $\tilde{M}$; im letzten Bild ist die Effizienz als Funktion der Untergrundreduktion aufgetragen.*

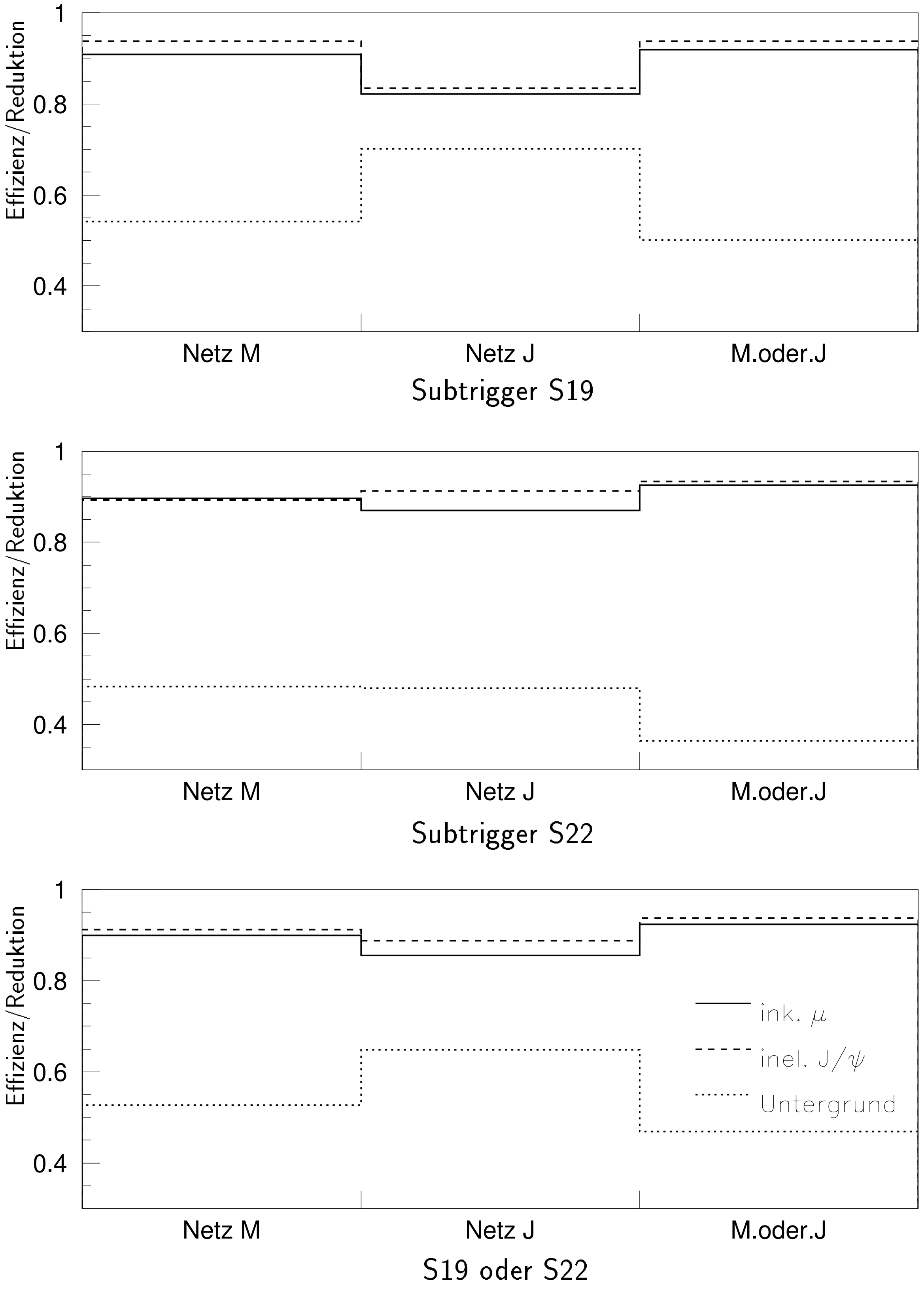


Abbildung 6.21: *Effizienzen der Subtrigger für die Physikdatensätze und Reduktion des Untergrundes für 1997 aufgenommene Daten (zu vergleichen mit Abbildung 6.14 für die Testdatensätze).*

		Netz M	Netz J	Kombination der Netze M und J	Anzahl der Ereignisse
Effizienz in %					
inkl. μ	S19	90.8	82.2	91.9	1295
	S22	89.7	86.9	92.6	3178
inel. J/ψ	S19	93.8	83.5	93.8	194
	S22	89.3	91.3	93.4	346
Untergrundreduktion in %					
Untergrund	S19	54.1	70.1	50.1	689
	S22	48.4	47.9	36.4	242

Tabelle 6.7: *Reduktion und Effizienz der in den L2NN-Trigger implementierten Netze (Netz M und Netz J) für die* **Datennahmeperiode 1997** *(zu vergleichen mit Tabelle 6.5 für die Testdatensätze).*

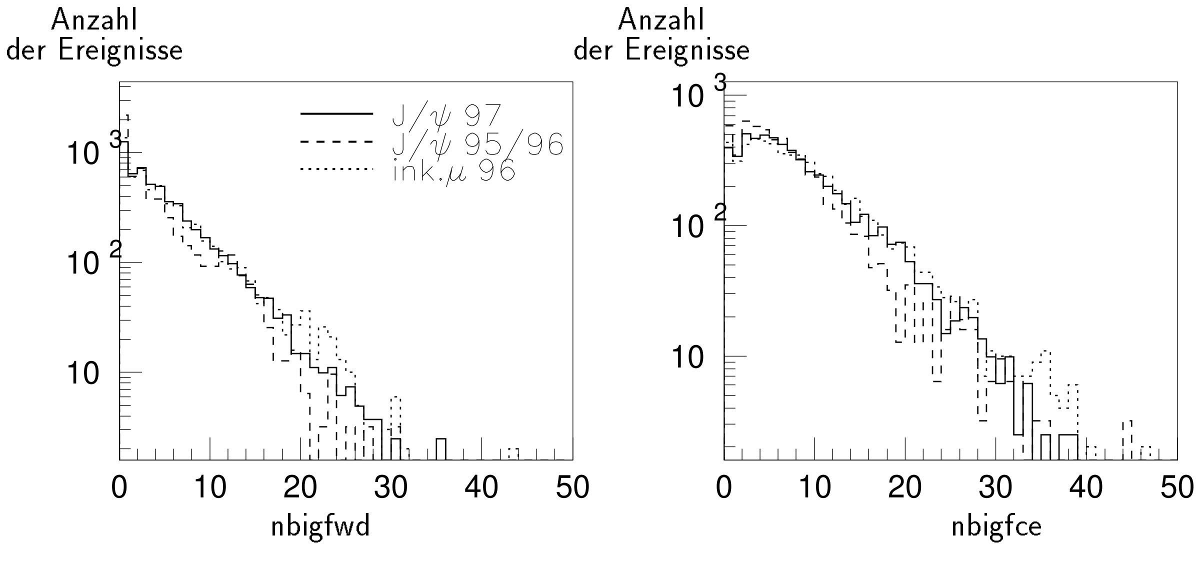


Abbildung 6.22: *Verteilung der nach vorne gerichteten neuentwickelten Eingabegrößen für Ereignisse mit J/ψ-Kandidaten aus 1997 im Vergleich mit den Trainingsereignissen mit unelastischer J/ψ-Produktion und inklusiven Myonen.*

und '97 unterscheiden. Die Verteilungen der aus den *big rays* gebildeten Größen, beispielsweise *nbigfwd* und *nbigfce*, weisen für die in 1997 selektierten Ereignisse mit unelastischen J/ψ-Kandidaten größere Ähnlichkeiten mit in 1996 selektierten Ereignissen mit inklusiver Myon-Produktion auf als mit den im Training verwendeten J/ψ-Daten (Abbildung 6.22). In Abbildung 6.23 ist die Massenverteilung und der Transversalimpuls $p_{t,J/\psi}$ der unelastischen J/ψ-Kandidaten dargestellt. Wie schon im Trainingsdatensatz erkennt man keinen Bereich, in dem von beiden Netzen gemeinsam prozentual deutlich mehr verworfen wird als in anderen. Die Effizienz in Abhängigkeit von der invarianten Masse schwankt zwischen 91% und 96%. Ebenso kann unter Berücksichtigung der Fehler keine drastische Änderung der Effizienz in den verschiedenen $p_{t,J/\psi}$-Intervallen festgestellt werden.

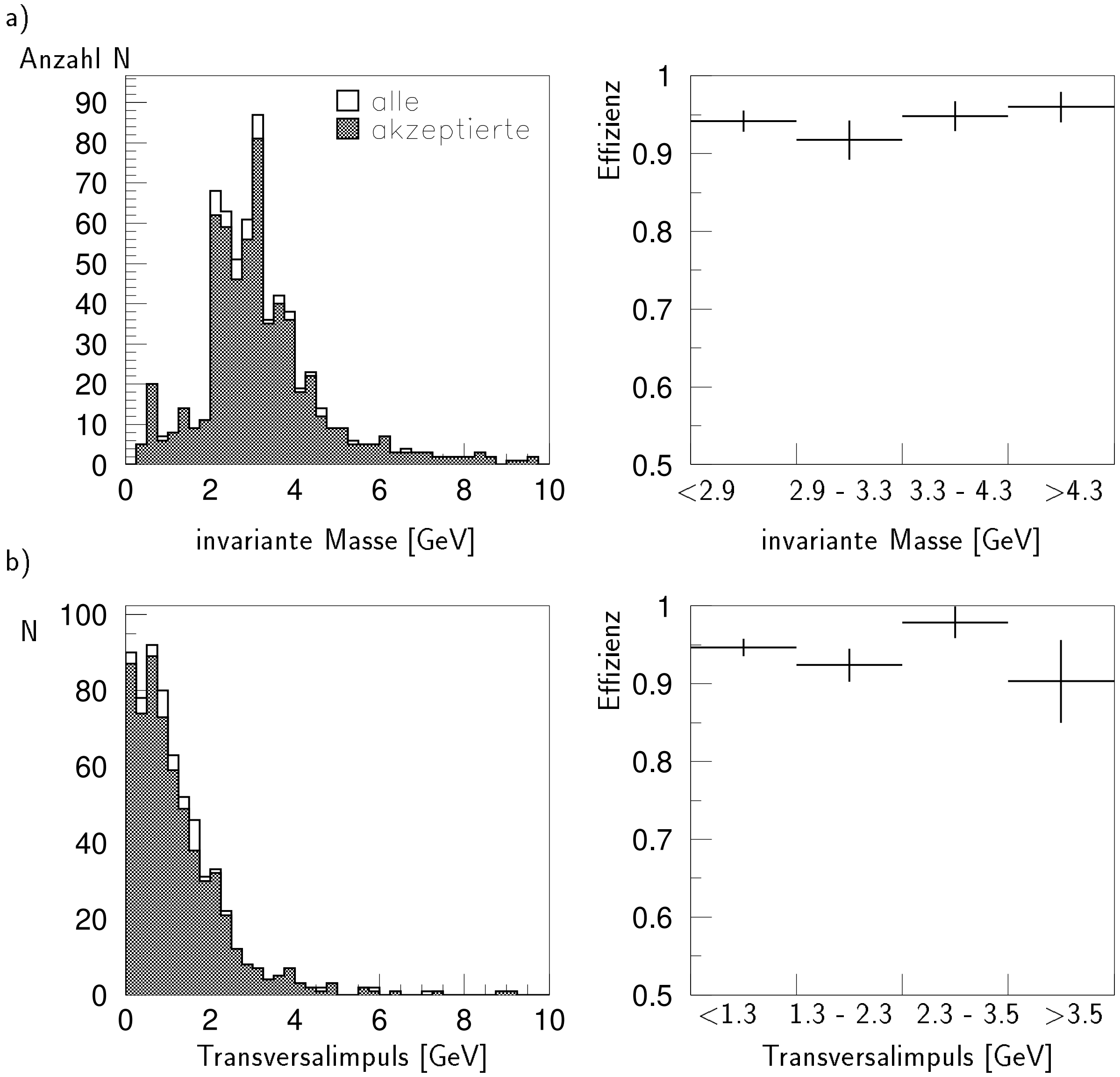


Abbildung 6.23: a) *Invariante Masse und b) transversaler Impuls der J/ψ-Kandidaten vor und nach der Entscheidung durch die Kombination der beiden Triggerelemente in 1997 sowie die zugehörigen Effizienzen in verschiedenen Massen- und $p_{t,J/\psi}$-Intervallen.*

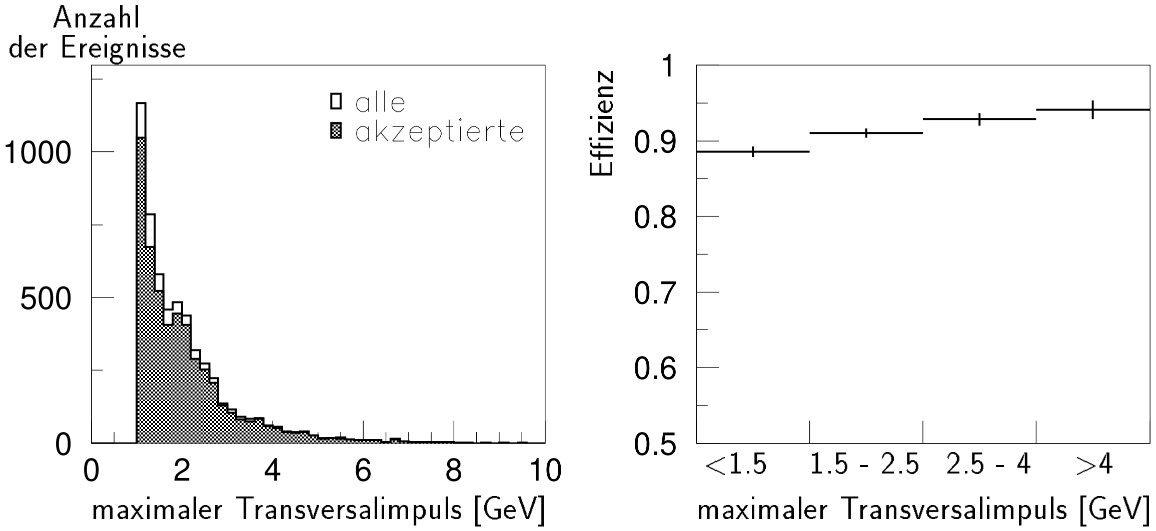


Abbildung 6.24: *Maximaler Transversalimpuls der im Myondetektor erkannten Myonen für alle Ereignisse mit inklusiven Myonen (weiß) und die von der Kombination der beiden Triggerboxen akzeptierten Ereignisse (grau). Außerdem ist rechts die erreichte Effizienz für die Kombination der Triggerboxen in Abhängigkeit von $p_{t,max}$ dargestellt.*

Für die Ereignisse mit inklusiver Myon-Produktion ist in Abbildung 6.24 wieder die Verteilung des maximalen Transversalimpulses $p_{t,max}$ sowie die Effizienz für die Kombination der beiden Triggerboxen in verschiedenen $p_{t,max}$-Intervallen gezeigt. Die Effizienz nimmt für größer werdende $p_{t,max}$ leicht von 92% auf 94% zu.
Die Untergrundreduktion in 1997 hat gegenüber den im Training erreichten Werten abgenommen, besonders für den Subtrigger S22, für den die Reduktion auf ca. 50% zurückgeht. Dies kann auf veränderte Untergrundbedingungen im Jahre 1997 zurückgeführt werden (vgl. Abschnitt 5.1).

6.3.7 Subtrigger S15

Anfang Juni 1997 ist auf der L1-Triggerstufe des H1-Experiments testweise ein neuer Subtrigger S15 für Ereignisse mit Myonen im Zentralbereich des Detektors eingerichtet worden. S15 setzt sich aus den Triggerelementen $(Mu\_Bar \,||\, Mu\_ECQ)$, $DCRPh\_THig$ und $zVtx\_sig\_mu\_D$ zusammen. Somit stellt er eine Abschwächung der Kombination der Subtrigger S19 und S22 dar (vgl. Abschnitt 5.1).
Der Subtrigger hat eine Rate bis 5 Hz (Abbildung 6.25) Sie kann durch eines der implementierten Netze für Ereignisse mit unelastischen J/ψ-Kandidaten oder mit inklusiver Myon-Produktion reduziert werden. Bis zu einer Aktivierung des Netzes hatte der Subtrigger einen Untersetzungsfaktor von 10000.
Das Verhalten des Netzes J ist mit einem L4-*transparent run* untersucht worden [Merk97], der 1548 von S15 akzeptierte Ereignisse enthält. Die Ergebnisse sind in Tabelle 6.8 aufgezeigt. Von diesen 1548 Ereignissen werden 37% ausschließlich von S15 (S15pur) akzeptiert, 63% auch von anderen Myonsubtriggern S18, S19, S22 und S34 (S15+). Von den Ereignissen mit S15pur werden von dem Netz 82% verworfen, von allen S15-Ereignissen 74%.

Wird das Netz auf dem Subtrigger S15 aktiviert, so führen nur die Ereignisse, die das Netz von S15pur akzeptiert, zu einer Erhöhung der Rate nach der zweiten Triggerstufe. Dieser Ratenzuwachs beträgt ungefähr 0.1 bis 0.3 Hz.

	Anzahl der Ereignisse	Anzahl der Ereignisse vom Netz akzeptiert
S15	1548	402≙26%
S15pur	568	104≙18%
S15+	980	298≙70%

Tabelle 6.8: *Reduktion des Untergrundes auf S15 durch das Netz J in der Datennahmeperiode 1997 [Merk97, MeyA97]. S15 umfaßt alle von Subtrigger S15 akzeptierten Ereignisse, S15pur nur die ausschließlich von S15 akzeptierten Ereignisse und S15+ die Ereignisse, die von S15 und einem oder mehreren der Subtrigger S18, S19, S22 oder S34 angenommenen Ereignisse.*

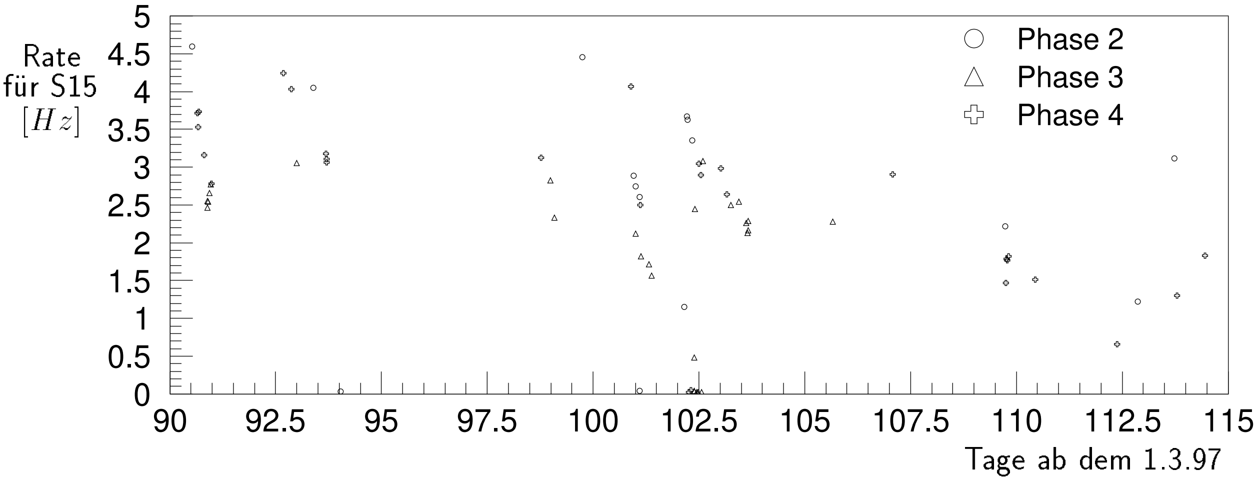


Abbildung 6.25: *Rate von S15 als Funktion der Zeit.*

Kapitel 7

Zusammenfassung und Ausblick

Diese Arbeit wurde beim H1-Experiment am *ep*-Speicherring HERA erstellt. Sie beschäftigt sich mit der Möglichkeit der Trennung von Untergrund und physikalisch interessanten Ereignissen mit Myonen durch neuronale Netze.

Untergrund entsteht im H1-Detektor durch viele verschiedene Prozesse: Strahl-Gas-, Strahl-Wand-Wechselwirkung sowie kosmische Myonen. Die Rate, mit der Untergrundereignisse auftreten, ist wesentlich höher als die physikalischer Prozesse. Deshalb wird durch ein vierstufiges Triggersystem Physik von Untergrund getrennt und nur Kandidaten für physikalische Ereignisse aus *ep*-Wechselwirkungen auf Band gespeichert.

Nach der ersten Stufe dieses Triggersystems zeigten die beiden untersuchten Myontrigger (S19 und S22) in der Datennahmeperiode 1996 ein instabiles Verhalten mit einer Rate bis zu 10 Hz. Diese Rate sollte auf der zweiten Ebene des Triggersystem mit Hilfe neuronaler Netze (L2NN) gesenkt und stabilisiert werden.
Als Eingabegrößen stehen dem Netz die vom Detektor gesendeten Triggerdaten und statistische Angaben, wie das z-Vertex-Histogramm, zur Verfügung. Neben schon vorhandenen Eingabegrößen wurden in dieser Arbeit neue Größen entwickelt und verwendet. Die neuen Größen geben einen groben topologischen Überblick über die Polarwinkelverteilung der Spuren im Detektor.
Die beiden untersuchten Myontrigger überdecken gemeinsam den gesamten Zentralbereich des Detektors ($20^\circ < \theta < 160^\circ$) und akzeptieren Ereignisse mit inklusiver Myon-Produktion sowie mit unelastischen J/ψ-Kandidaten. Es stellt sich dabei die Frage, ob ein gemeinsames Netz effizient den Untergrund von beiden Physikkanälen trennen kann oder ob der Einsatz zweier einzelner Netze erfolgreicher ist.
Für beide Ereignisgruppen wurde ein neuronales Netz trainiert. Als Untergrund wurden dabei Ereignisse aus L4-*transparent runs* verwendet. Neben den beiden Netzen für jeweils eine Physikklasse ist ein weiteres für die beiden physikalischen Ereignisgruppen zusammen erarbeitet worden. Die Trennung von „Physik“ und Untergrund ist in allen drei Fällen möglich.
Für die drei Netze sowie die Kombination der beiden einzelnen Netze können die Effizienzen hinsichtlich der beiden Subtrigger und der beiden Physikkanäle nach Training und

Generalisierung angegeben werden (siehe Tabelle 6.5/Abbildung 6.14). Mit der Kombination der beiden einzelnen Netze wurden höhere Effizienzen für die Physikereignisse erreicht als mit dem kombinierten Netz. Deshalb sind zunächst die beiden einzelnen Netze in den L2-Trigger implementiert worden. Sie verwerfen jedoch keine Ereignisse, die von den untersuchten Myontriggern akzeptiert worden sind. Über die Auswertung von Daten nach der Implementierung konnte das Verhalten der Netze in der Datennahmeperiode 1997 untersucht werden (siehe Tabelle 6.7/Abbildung 6.21). Es zeigten sich in den Effizienzen nur leichte Unterschiede zu den Ergebnissen des Trainings. Die Untergrundreduktion hat sich jedoch verringert. Dies läßt sich vermutlich auf veränderte Untergrundbedingungen im Jahr 1997 zurückführen.
Die Ratenentwicklung der untersuchten Myontrigger im Jahr 1997 zeigt eine deutliche Abnahme gegenüber den Raten von 1996. Sie haben sich auf eine Rate zwischen 0.5 und 1.5 Hz stabilisiert. Damit ist der Einsatz der neuronalen Netze für die beiden Myontrigger (S19 und S22) zur Zeit nicht notwendig.

Das Anwenden der zweiten Triggerstufe des H1-Experiments ermöglicht es, die Bedingungen auf der ersten Stufe abzuschwächen. Mit dem Trigger S15 wurde eine Kombination der beiden untersuchten Myontrigger geschaffen, die weniger scharfe Bedingungen an das z-Vertex-Histogramm und an die Zahl der Spuren stellt. Für diesen neuen Myontrigger ist die Untergrundunterdrückung des Netzes für unelastische J/ψ-Produktion untersucht worden. Das Netz verwirft 74% aller von S15 akzeptierten Ereignisse und 82% der alleine von S15 akzeptierten Ereignisse. Daraus läßt sich ein Ratenzuwachs bzgl. bereits existierender Myontrigger von 0.1 bis 0.3 Hz nach der zweiten Triggerstufe abschätzen.

Das Netz für unelastische J/ψ-Kandidaten wurde Ende Juni 97 auf dem L1-Trigger S15 aktiviert. Das weitere Verhalten des Netzes muß in der Zukunft beobachtet werden: Welche Ereignisse es verwirft und welche Art von Ereignissen es akzeptiert. Von Bedeutung sind dabei Ereignisse mit J/ψ-Kandidaten im Bereich kleiner Schwerpunktsenergien im Photon-Proton-System.

Ebenso wie für den Zentralbereich lassen sich auch für den Vorwärtsbereich neuronale Netze erarbeiten. Im Vorwärtsbereich wird die Ereignisrate noch stärker durch Ereignisse aus Strahl-Gas-Wechselwirkungen dominiert als im Zentralbereich. Diese müssen hier von den physikalischen Prozessen der unelastischen J/ψ-Erzeugung und der inklusiven Myon-Produktion getrennt werden. Dazu müssen solche Untergrundereignisse bei der Selektion der zum Training verwendeten „Physik“ sorgfältig aus dem Datensatz entfernt werden. Dabei liegen bei den Daten mit inklusiven Myonen die Schwierigkeiten in der Trennung von den Strahl-Gas-Wechselwirkungen, bei den Ereignissen mit unelastischen J/ψ-Kandidaten in einer geringen Anzahl von Ereignissen im Vorwärtsbereich. Als Untergrund können eventuell Ereignisse verwendet werden, bei denen die Protonen auf keine Positronen als Kollisionspartner treffen (*pilot bunches*).

Anhang A

Tabellen zur Lee-West-Selektion

In den folgenden Tabellen sind die in der Spur- und Myonselektion gestellten Anforderungen [West96] zusammengefaßt. Dabei werden folgende Abkürzungen verwendet:

- R_0 ist der Abstand in der R-ϕ -Ebene zwischen einer extrapolierten Spur vor der Vertexanpassung und dem Ereignisvertex.
- $\chi^2_{Spuranpassung}$ bzw. $\chi^2_{vertexangepaßteSpuren}$ sind die χ^2-Werte für die Anpassung einer Spur an die getroffenen Drähte bzw. einer auf einen Vertex gezwungenen Spur.
- Die primären und sekundären Spursegmente entsprechen verschiedenen Spurqualitäten in den vorderen Spurkammern, die je nach Anzahl von Treffern in den drei Orientierungen eines planaren Moduls unterschieden werden.
- DCA ist der geringste radiale Abstand der nicht an einen Vertex angepaßten Spur zum Primärvertex des Ereignisses.
- R_{start} ist der Abstand des ersten Treffers in der R-ϕ -Ebene zur z-Achse.
- Unter Spurlänge versteht man hier den Abstand in der R-ϕ-Ebene zwischen Anfangs- und Endpunkt einer Spur.

Vorwärtsspuren		
P_t	$>$	0.150 GeV/c
Θ	$>$	6.0°
	$<$	25.0°
R_0	$<$	10.0 cm
$\chi^2_{Spuranpassung}$	$<$	10.0
$\chi^2_{vertexangepaßteSpuren}$	$<$	25.0
$N_{primäre+sekundäre\ planare\ Segmente}$	$\geq$	1
$N_{planare+radiale\ Segmente}$	$\geq$	2
dp/p	$<$	1
zentrale Spuren		
P_t	$>$	0.150 GeV/c
Θ	$>$	20.0°
	$<$	160.0°
$\|DCA\|$	$<$	2.0 cm
R_{start}	$<$	50.0 cm
Spurlänge	$>$	10.0 cm für $\Theta < 150°$
Spurlänge	$>$	5.0 cm für $\Theta > 150°$
kombinierte Spuren		
P_t	$>$	0.150 GeV/c
Θ	$>$	0.0°
	$<$	180.0°
$\|DCA\|$	$<$	5.0 cm
R_{start}	$<$	50.0 cm

Tabelle A.1: *Anforderungen der Spurselektion*

Nach diesen Kriterien werden die Spuren in dieser Arbeit ausgewählt. Außerdem werden in dieser Arbeit nur Spuren selektiert, die von geladenen Teilchen hervorgerufen worden sind und die vom primären Vertex stammen. Diese Bedingung dient auch zur Auflösung der Ambiguitäten, genauso wie die Bevorzugung von kombinierten vor zentralen Spuren und von zentralen Spuren vor Vorwärtsspuren.

Die Myonselektion greift auf die folgenden Größen und Schnitte zurück:

- ρ, ρ_x ,ρ_y sind der geringste Abstand bzw. seine x- und y-Komponente der aus dem Eisen zurückextrapolierten Spur zum Ereignisvertex in der R-ϕ -Ebene.
- Z_0 ist die z-Koordinate des Anfangspunktes der Eisenspur.
- N_{Lagen}, $i_{erste\,Lage}$, $i_{letzte\,Lage}$ sind die Gesamtanzahl der getroffenen Lagen, die erste bzw. die letzte angesprochene Lage im instrumentierten Eisen.

Barrel		
ρ	$<$	100 cm
Z_0	$<$	100 cm
N_{Lagen}	$\geq$	2
$i_{erste\,Lage}$	$\leq$	5
$i_{letzte\,Lage}$	$\geq$	2
hintere Endkappe		
ρ_x	$<$	100 cm
ρ_y	$<$	100 cm
N_{Lagen}	$\geq$	3
$i_{erste\,Lage}$	$\leq$	8
$i_{letzte\,Lage}$	$\geq$	3
vordere Endkappe		
ρ_x	$<$	100 cm
ρ_y	$<$	100 cm
N_{Lagen}	$\geq$	6
$i_{erste\,Lage}$	$\leq$	5
$i_{letzte\,Lage}$	$\geq$	6
Kalorimetermyonen		
$\mu - Qualität$	$\geq$	2
$Separationswinkel$	$>$	12°

Tabelle A.2: *Bedingungen für die Myonselektion*

Anhang B

Tabelle für $\bar{\theta}$ der $\theta-$Abschnitte der *big rays*

In der folgenden Tabelle sind für die 14 θ-Abschnitte in der $\theta\phi$-Ebene der *big rays* die mittleren θ-Werte angegeben.

θ-Abschnitt	mittlerer Polarwinkel $\bar{\theta}$ [°]
0	4.60
1	6.97
2	9.66
3	12.37
4	14.82
5	18.37
6	24.46
7	31.68
8	42.93
9	58.72
10	80.40
11	109.18
12	131.67
13	146.62

Tabelle B.1: *Mittleres θ der Abschnitte in der* big ray*-Ebene [Müll96]*

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[Ali87] A.Ali et al.
Heavy Quark Physics at HERA
Proceedings of the HERA Workshop, Hamburg (1987)

[Aid96] H1 Collaboration, S.Aid et al.
Nucl. Phys., **B472** (1996) 23

[Baie81] R.Baier, R.Rückl
Hadronic Production of J/ψ and Υ, Transverse Momentum Distribution
Phys.Lett., **102B** (1981) 364

[Beck96] H.Beck
Principles and Operations of the z-Vertex Trigger
Internal note H1-05/96-479 (1996)

[Berg81] E.L.Berger, D.Jones
Inelastic Photoproduction of J/ψ and Υ by Gluons
Phys. Rev., **D23** (1981) 1521

[Bizo91] J.C.Bizot et al.
Proposal for a Topological Level 2 Trigger
Internal note H1-06/91-181 (1991)

[Bizo97] J.C.Bizot et al.
Strategy Studies for the H1 Topological L2-Trigger (L2TT)
Internal note H1-01/97-508 (1997)

[Bodw95] G.T.Bodwin, E. Braaten, G.P.Lepage
Rigorous QCD Analysis of Inclusive Annihilation and Production of Heavy Quarkonia
Phys. Rev., **D51** (1995) 1125

[Dres97] M.Dressel
*Suche nach **b**-Quarks mit Myonen aus dem semileptonischen Zerfall im H1-Detektor bei HERA*
Diplomarbeit am Institut für Experimentelle Kernphysik der Universität Karlsruhe (Juni 1997)

[dtv84] G.Vogel, H.Angermann
dtv-Atlas zur Biologie
erweiterte Ausgabe, 1.Auflage (1984)

[Eich88] R.A.Eichler, Z.Kunszt et al.
Charm, Bottom and Top Produktion in High Energy Electron Proton Collisions
Nucl.Phys., **B308** (1988) 791

[Elli89] R.K.Ellis, P.Nason
QCD Radiative Corrections of the Photoproduction of Heavy Quarks
Nucl. Phys., **B312** (1989) 551

[Frix95] S.Frixione, M.Mangano, P.Nason, G.Ridolfi
Total Cross Section for Heavy Flavour Production at HERA
Phys. Lett., **B348** (1995) 633

[EV96] *New Event Vector*
'wwwh1.mppmu.mpg.de/projects/neuro/software.html'

[Fent95] J.Fent et al.
A Neural Network Second Level Trigger for the H1-Experiment at HERA
Internal note H1-10/95-457 (1995)

[Gold96] D.Goldner
Online-Datenaufbereitung für die Hardwarerealisierung eines künstlichen Neuronalen Netzwerkes als Level-2 Trigger im H1-Experiment
Dissertation am Fachbereich Physik der Universität Dortmund (April 1996)

[Grub95] A.Gruber
Documentation for the Network Simulator -nsi-
(Oktober 1995)

[Hech87] R.Hecht-Nielsen
In Proceedings of the International Conference on Neural Networks, II, S.19-32, IEEE-Press (1982)

[H1-96-01] H1 Collaboration
The H1-Detector at HERA
Internal Report (März 1996)

[H1-97] H1 Collaboration
The Tracking, Calorimeter and Muon Detectors of the H1 Experiment at HERA
Nuclear Instruments and Methods in Physics Research **A386** (1997) 348-396

[Krüg94] U.-P.Krüger
Untersuchung der Erzeugung schwerer Quarks durch ihren Zerfall in Myonen im H1-Detektor bei HERA
Dissertation am Fachbereich Physik der Universität Hamburg (1994)

[Köhn97] 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 (1997)

[Kolm57] A.N.Kolmogorov
Dokl.Akad.Nauk. **115**, USSR (1957)

[Lönn93] L.Lönnblad, C.Peterson, T.Rögnvaldsson
JETNET3.0 - A Versatile Artificial Neural Network Package
LU TP 93-29, CERN-TH.7135/94 (Dezember 1993)

[McCu43] W.McCulloch, W.Pitts
A Logical Calculus of the Ideas Immanent in Nervous Activity
Bull. Math. Biophys., **vol.5** (1943)

[Merk97] P.Merkel
persönliche Mitteilung

[MeyA95] Arnd Meyer
Nachweis von Myonen unter kleinem Polarwinkel im H1-Detektor bei HERA
Diplomarbeit am II. Institut für Experimentalphysik der Universität Hamburg (Februar 1995)

[MeyA97] Arnd Meyer
persönliche Mitteilung

[Möck94] J.Möck
Einsatz neuronaler Netze als intelligente Trigger im H1-Experiment
Diplomarbeit an der TU München (1994)

[Müll96] D.Müller
persönliche Mitteilung

[PDG] Particle Data Group, R.M.Barnett et al.
Particle Physics Booklet, extract from Review of Particle Physics
Phys.Rev, **D54** (1996) 1

[Ried95] J.Riedlberger
The H1 Trigger with Emphasis on Tracking Triggers
Internal note H1-01/95-419 (1995)

[Roja93] R.Rojas
Theorie der neuronalen Netze
Springer Verlag (1993)

[Rose58] F.Rosenblatt
The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain
Psychological Review, **vol.65** (1958)

[Rume86] D.E.Rumelhart, G.E.Hinton, R.J.Williams
Learning Internal Representations by Error Propagation
Parallel Distribution Processing, MIT Press (1986)

[Samt96] D.Samtleben
All you need for training a neural net for L2 (1996)

[Schö96] A.Schöning
Untersuchung von Prozessen mit virtuellen und reellen $W^{\pm}$-Bosonen am H1-Detektor bei HERA
Dissertation am Fachbereich Physik der Universität Hamburg (1996)

[Weiz34] C.F. von Weizsäcker
Ausstrahlung bei Stößen sehr schneller Elektronen
Z. Phys. **88** (1934) 612

[WesD96] D.Westner
Ein neuronaler Netzwerktrigger für die Produktion von J/ψ in ep-Streuung ($J/\psi \rightarrow e^+e^-$)
Diplomarbeit an der Fakultät für Physik der Ludwig-Maximilians-Universität München (März 1996)

[WesD95] D.Westner
Documentation for the Neural Network Tester (Juni 1995)

[West96] Lee West
How to use the Heavy Flavour Working Group Track, Myon and Electron Selection Code
Version 1.00/12 (November 1996)

[Zobe96] T.Zobel
Neuronale Netzwerke als Trigger für die J/ψ-Meson-Produktion in ep-Kollisionen am Beispiel des Zerfalls $J/\psi \rightarrow \mu\mu$
Diplomarbeit an der Fakultät für Physik der Ludwig-Maximilians-Universität München (März 1996)

[Zobe95] T.Zobel
Mixer Documentation (März 1995)