Untersuchungen zu neuronalen Triggern für J/ψ-Erzeugung in Elektron-Proton-Stößen bei HERA

Diplomarbeit an der Fakultät für Physik
der
Ludwig-Maximilians-Universität München

vorgelegt von
Enrico Concetti

angefertigt am

Max–Planck–Institut für Physik
– Werner–Heisenberg–Institut –

München
April 1998

Erstgutachter: Prof. Dr. Christian Kiesling

Zweitgutachterin: Prof. Dr. Dorothee Schaile

Inhaltsverzeichnis

1 Einleitung

Der Mensch hatte sich schon in antiken und untergegangenen Zivilisationen, wie Babyloniern und Ägyptern, gefragt, woraus die Welt und die Materie besteht und, ob es Grundbausteine gibt, aus denen man die ganze Vielfalt der Natur herleiten kann. Eine erste naturphilosophische Auffassung über diese Frage lieferten jedoch die Vorsokratiker Thales, Anaximandros und Anaximenes im 6. Jahrhundert v. Chr., die aus Milet in dem ionischen Griechenland stammten. Insbesondere Anaximandros war überzeugt, daß Wasser, Luft, Erde und Feuer die vier Grundelemente sind [1].
In der modernen Physik spricht man nicht mehr von Wasser, Luft, Erde und Feuer, sondern von Leptonen und Quarks. Sie werden heute als die Elementarteilchen gesehen, die nicht aus anderen Teilchen zusammengesetzt sind. Aus der folgenden Tabelle kann man die Größenordnung dieser Objekten entnehmen.

Atom		10^{-10} m
Nukleus		10^{-14} m
Proton		10^{-15} m
Lepton / Quark	$\leq$	10^{-18} m

Tabelle 1.1: Größenordnung von einigen Objekten .

Es gibt sechs Arten von Leptonen und sechs Arten von Quarks, aber innerhalb dieser Gruppen sind die Elemente in drei aus jeweils zwei gekoppelten Teilchen zusammengesetzten Klassen unterteilt, die sich voneinander hauptsächlich durch die Masse unterscheiden (siehe Tabelle 1.2). Während die Leptonen, unter denen das Elektron lange, sehr gut bekannt ist, nachgewiesen und beobachtet worden sind, können die Quarks einzeln nicht beobachtet werden, weil sie eher eine notwendige theoretische Schöpfung sind und treten in der Natur nur untrennbar in anderen Gebilden auf: Baryonen und Mesonen, die aus drei bzw. zwei Quarks bestehen [2].
Jedem von den oben erwähnten Elementarteilchen wird ein Anti-Teilchen zugewiesen, das das entsprechende Elementarteilchen der Anti-Materie ist. Um das Bild zu vervollständigen, braucht man noch eine weitere Klasse von Elementarteilchen, die Bosonen, von denen einige für die Wechselwirkungen in der Materie verantwortlich sind, wie z.B. das Photon für die elektromagnetische Kraft.
Was bisher kurz und grob geschildert worden ist, fällt es unter dem Namen "**Standard Modell** der Elementarteilchenphysik". Wie jede physikalische

Theorie ist das **Standard Modell** keine vollständige und abgeschlossene Darstellung der Welt, sondern viel mehr eine dynamische Vorstellung, die aus der Zusammenlegung von verschiedenen und noch erforschten Theorien wie **QCD** (Quantenchromodynamik) und **QED** (Quantenelektrodynamik) u.a. zustandegekommen ist [3].

Leptonen			**Quarks**		
Flavor	Masse [GeV/c^2]	Ladung	Flavor	Masse [GeV/c^2]	Ladung
ν_e	$< 7 \text{ x } 10^{-9}$	0	u	0.005	2/3
e	0.000511	-1	d	0.01	-1/3
ν_μ	< 0.0003	0	c	1.5	2/3
μ	0.106	-1	s	0.2	-1/3
ν_τ	< 0.03	0	t	170	2/3
τ	1.7771	-1	b	4.7	-1/3

Tabelle 1.2: Die Elementarteilchen von dem Standard Modell .

Ein Pionier in der experimentellen Forschung in der Teilchenphysik war Ernest Rutherford (1871-1937). Der englische Wissenschaftler begann 1910 dünne Goldfolien mit α-Teilchen, d.h. Nuklei von Helium, zu beschießen [4].
Wegen des Energiegehaltes der geschossenen α-Teilchen beschränkten sich die Untersuchungen von Rutherford auf die Größenordnung des Atomkerns (siehe Tabelle 1.1). Seitdem ist die Art und Weise, wie solche Experimente ausgeführt werden, nicht viel geändert worden; die Energie der geschossenen Teilchen ist aber um einen Faktor 1000 erhöht worden. Das erlaubt die Erforschung auf der Skala der Nukleuskomponenten, wie z.B. das Proton, um Schlußfolgerungen über die Bausteine seiner inneren Struktur, die Quarks, ziehen zu können. Außer den Experimenten mit einem im Labor ruhenden Zielobjekt (*Fixed Target*), gegen das Teilchen geschossen werden, führt man auch solche heute aus, in denen zwei Teilchen zuerst beschleunigt und dann kollidiert gelassen werden, so daß man eine noch höhere Gesamtschwerpunktenergie erreicht. Man unterscheidet deshalb grundsätzlich zwischen zwei verschiedenen Arten von Laboranlagen, in denen die Zusammenstößen stattfinden: Beschleuniger und Speicherringe [5].
In den Beschleunigern werden die Teilchen (üblicherweise Protonen, Elektronen oder Ionen) durch elektrische Felder auf das gewünschte Energieniveau gebracht; daher müssen die Teilchen geladen sein. Danach werden sie auf das Zielobjekt gerichtet und dagegen geschickt, eventuell an mehrere verschiedene gleichzeitig, wo die Detektoren aufgebaut sind und die Daten aufgenommen werden.

In den Speicherringen laufen zwei unterschiedliche Strahlenbündeln in entgegengesetzte Richtungen in einem ringförmigen Vakuumrohr auf verschiedene Höhen. An bestimmten Stellen des Ringes, an denen sich die Detektoren befinden, werden die Strahlenbündeln auf dieselbe Höhe gesetzt und die Teilchen kollidieren gegeneinander. Der resultierrende Teilchenschauer durchdringt die Detektoren und hinterläßt Spuren, aus denen man nachträglich bestimmen kann, um welches Teilchen es sich gehandelt hat.
Es reicht aber nicht, nachträglich festzustellen, welche Teilchen erzeugt worden sind, weil die Datenmenge so riesig ist, daß es in der Praxis unmöglich ist, alle Daten aufzunehmen. Es muß eine Selektion während des Experimentes stattfinden, um nur die echte relevante physikalische Ereignisse aufzuzeichnen und diejenigen auszuschliessen, die sich aus den Zusammenstößen mit dem nicht vermeidbaren Restgas im Vakuumsrohr ergeben.
Dafür werden sogenannte Trigger eingesetzt, die die Aufgabe haben, den Spreu vom Weizen, d.h. die Physik von dem Untergrund zu unterscheiden. Sie müssen eine Entscheidung meistens in wenigen Nanosekunden treffen und diese dem Datennahmeapparat mitteilen, der sich dann erst einschaltet, um die Daten zu speichern, wenn es nötig ist.
An der **H**adron-**E**lektron-**R**ing-**A**nlage (**HERA**) in Hamburg werden bei dem **H1**-Experiment seit 1994 neuronale Netze als intelligente Trigger in einem mehrstufigen Triggerkonzept mit Erfolg eingesetzt. An diesem Lepton-Proton-Speicherring werden verschiedene Experimente betrieben. In dem **H1**-Experiment interessiert man sich besonders für die Struktur des Protons. Mittels des **H1**-Detektors werden die Ereignisse der Positron-Proton-Streuung untersucht. Die von dem Detektor aufgesammelte Informationen werden verarbeitet und einige so berechnete Größen werden in den neuronalen Netzen verwendet, die anhand dieser Informationen die Entscheidung treffen müssen, ob das aktuelle Ereignis physikalisch relevant ist oder nicht.
Diese Arbeit beschäftigt sich mit einer Studie über neue mögliche Größen, die die Arbeit der Netze erleichtern und ihre Leistung steigen sollen. Insbesondere wird die Aufmerksamkeit auf eine bestimmte Reaktion in der Physikdaten gesetzt: die elastische Produktion des Vektormesons J/ψ mit einer invarianten Masse von $m = 3,097\ GeV$, die bei dem Zerfall in einem Positron und einem Elektron nachgewiesen wird:

$$e^+p \longrightarrow e^+p\,(J/\psi \longrightarrow e^+e^-). \tag{1.1}$$

In Kapitel 2 wird eine Übersicht des **H1**-Experimentes und des dazugehörigen **H1**-Detektors angeboten. In Kapitel 3 wird man kurz in die Physik bei **H1** eingeführt und es wird die diffraktive elastische Erzeugung des J/ψ behandelt. In Kapitel 4 werde ich auf das Triggersystem eingehen und den mit neuronalen Netzen arbeitenden Teil hervorheben.
Im Kapitel 5 werden allgemeine Eigenschaften neuronaler Netze beschrieben, so daß man sich eine Vorstellung machen kann, wie sie trainiert werden können und welche mathematische Überlegungen im Hintergrund stehen.
Die Kriterien der Datenselektion werden im Kapitel 6 dargestellt. Das Kapitel

7 stellt die Clusteralgorithmen vor, mit derer Hilfe neue einzusetztende Größen herleitet werden. Die Ausführung des Netzwerktrainings, womit ein Netz für die richtigen Entscheidungen vorbereitet wird, wird im Kapitel 8 besprochen. In Kapitel 9 werden die Ergebnisse und die daraus geleitete Schlüße diskutiert. Das Kapitel 10 enthält eine Zusammenfassung.

2 Das H1-Experiment

2.1 Der Speicherring HERA

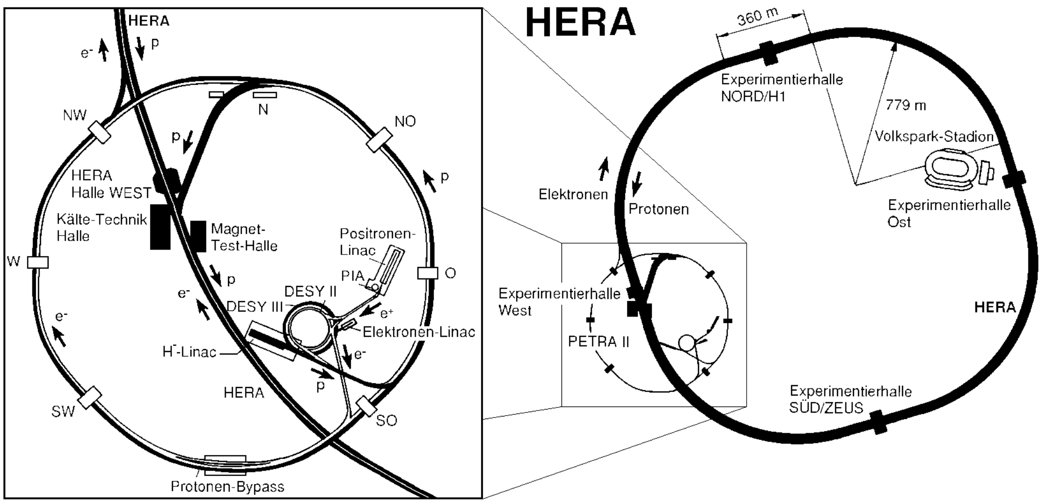


Abbildung 2.1: Der HERA Speicherring (rechts) mit dem vergrößerten Vorbeschleunigersystem (links).

Die **H**adron-**E**lektron-**R**ing-**A**nlage (**HERA**) in Hamburg verfügt über zwei getrennte Speicherringe und ein System von Vorbeschleunigern (siehe Abbildung 2.1), die die Teilchen auf die gewünschte Einschußenergie bringen, bevor sie in den **HERA**-Ring injiziert werden. Es handelt sich um Positronen und H^--Ionen; aus diesen letzten werden die beiden Elektronen vor der Übergabe abgestreift, so daß nach **HERA** Protonen extrahiert werden [6].
Die Speicherringe befinden sich in einem ringförmigen Tunnel, 10-15 m unter der Erde und haben einen Umfang von 6,3 km. Dort werden die Positronen durch herkömmliche Dipolmagnete gelenkt und in den Rohren gehalten, während supraleitende Hohlraumresonatoren sie von 7 GeV auf die Energie von 27,5 GeV beschleunigen, indem die Synchrotronstrahlungsverluste durch die von den Resonatoren erbrachte Leistung von 13,2 MW ausgeglichen werden. Daß die Protonen ihre endgültige Energie von 820 GeV erreichen, wird es durch supraleitende Magnete bewerkstelligt, die ein Feld von 4,68 T erzeugen. Die Teilchen werden in Bündeln zusammengefaßt, die in zwei Experimentierhallen aufeinander gelenkt werden und sich während des Laufes einer Messungsperiode (*Run*) an den Detektorenstellen alle 96 ns (*Bunch Crossing*, *BC*) durchdringen.

HERA wird von der Stiftung **D**eutsches **E**lektronen **Sy**nchrotron (**DESY**) verwaltet, die dadurch internationale Forschung fördert [7].

2.2 Der H1-Detektor

Um die von den Positron-Proton-Kollisionen erzeugten Teilchenschauer nachzuweisen und die teilnehmenden Teilchen zu identifizieren, wird der **H1**-Detektor (siehe Abbildung 2.2) benützt [8].
Man beachte das angegebene Koordinatensystem. In dem aus kohlefaserverstärktem Aluminium bestehenden Strahlrohr $\boxed{1}$ laufen die Protonen in die Richtung der z-Achse und die Positronen in die entgegengesetzte. Der ganze Detektor ist in verschiedenen Schichten um dieses Strahlrohr aufgebaut. Um die Wechselwirkungszone befinden sich zuerst die zentralen Spurkammern $\boxed{2}$, die eine Erweiterung $\boxed{3}$ in der Vorwärtsrichtung besitzen. Zur Energiemessung dient das Flüssig-Argon-Kalorimeter mit seinem elektromagnetischen- $\boxed{4}$ und hadronischen $\boxed{5}$ Teil.
Das Kalorimeter deckt den größten Winkelbereich $4° \leq \theta \leq 154°$ ab, während der vorwärtige Winkelbereich $0,6° \leq \theta \leq 3,5°$ von einem Kupfer-Silizium-Kalorimeter (**PLUG**) $\boxed{13}$ und der rückwärtige Winkelbereich $151° \leq \theta \leq 176°$ von einem Blei-Szintillator-Kalorimeter (**SPACAL**) $\boxed{12}$ erfaßt werden.
Die supraleitende Solenoid-Spule $\boxed{6}$ erzeugt ein starkes Magnetfeld (1,15 T), mit dessen Wirkung Transversalimpulsen bis zu 50 GeV/c bestimmt werden können. Die Solenoid-Spule ist von Myonkammern $\boxed{9}$ und ein Rückflußjoch aus Eisen mit Streamerröhren $\boxed{10}$ umgeben; die ersten sollen die Myonen nachweisen und das zweite mißt die Energie von hadronischen Teilchen, die das ganze Kalorimeter durchdrungen hatten. Weitere Myonkammern $\boxed{9}$ schließen den Detektor im Vorwärtsbereich und werden von einem eigenen Myon-Toroid-Magnet bedient.
Der **H1**-Detektor ist in der Vorwärtsrichtung stärker instrumentiert, weil man eher in dieser Richtung auslaufende Teilchen erwartet, da das Proton mit einem größeren Anteil zu der Schwerpunktsenergie beiträgt.

2.2.1 Das Flüssig-Argon-Kalorimeter

Das Flüssig-Argon-Kalorimeter dient zur Energiemessung zum einen von elektrisch geladenen Teilchen in seinem inneren elektromagnetischen Schale und zum anderen von hadronischen Schauern in seinem äußeren hadronischen Schale. Ein Kryostat sorgt dafür, daß bei dem flüssigen Argon eine konstante Temperatur von 90 K herrscht. Im elektromagnetischen Teil sind 2,4 mm dicke Bleiplatten als Absorber zwischen 2,35 mm dicken Argonschichten gestellt. Für den hadronischen Teil werden 16 mm starken Edelstahlplatten verwendet; die Dicke der Argonschicht beträgt hier 2,4 mm [9].

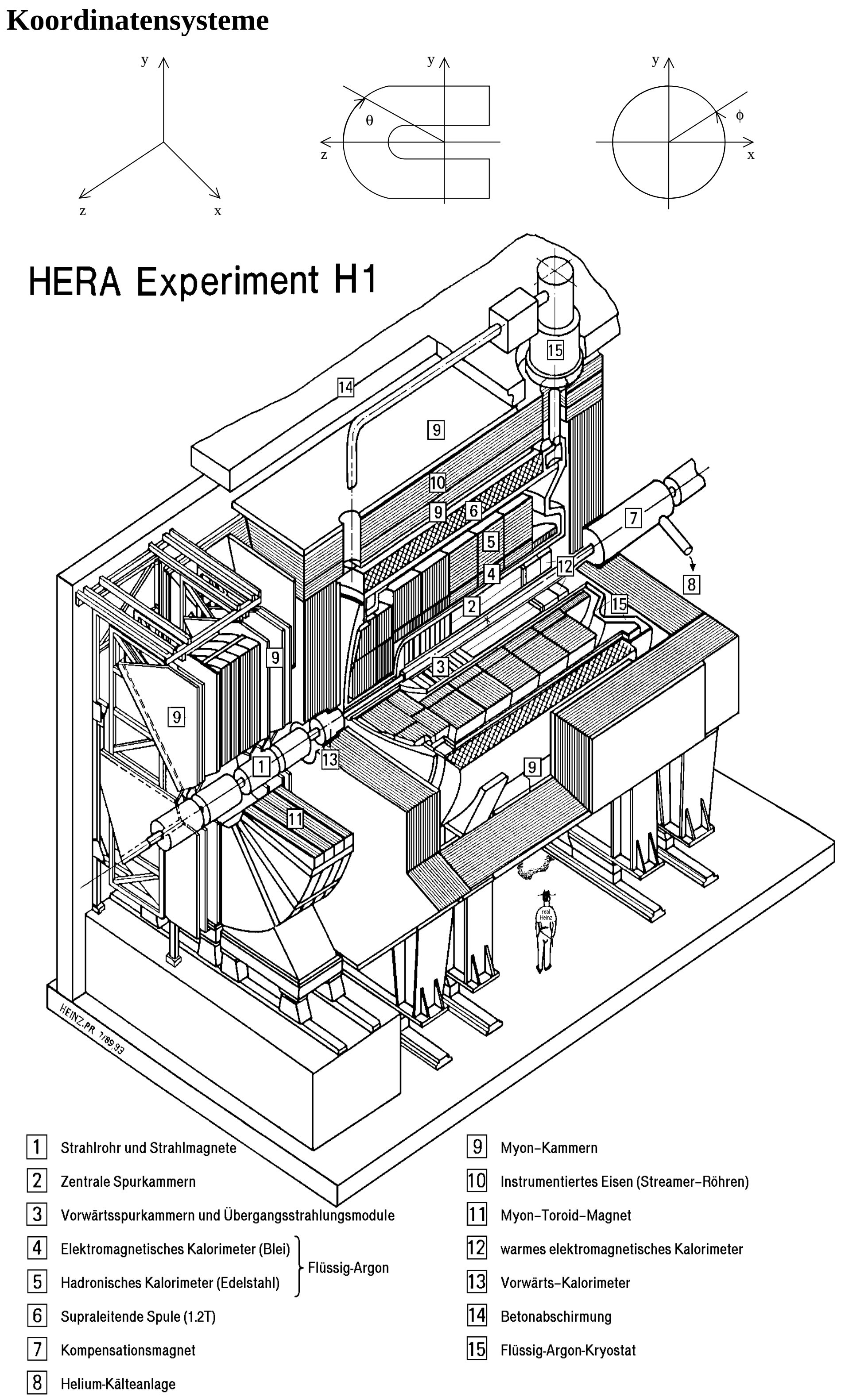


Abbildung 2.2: Koordinatensysteme und schematische Ansicht des H1-Detektors.

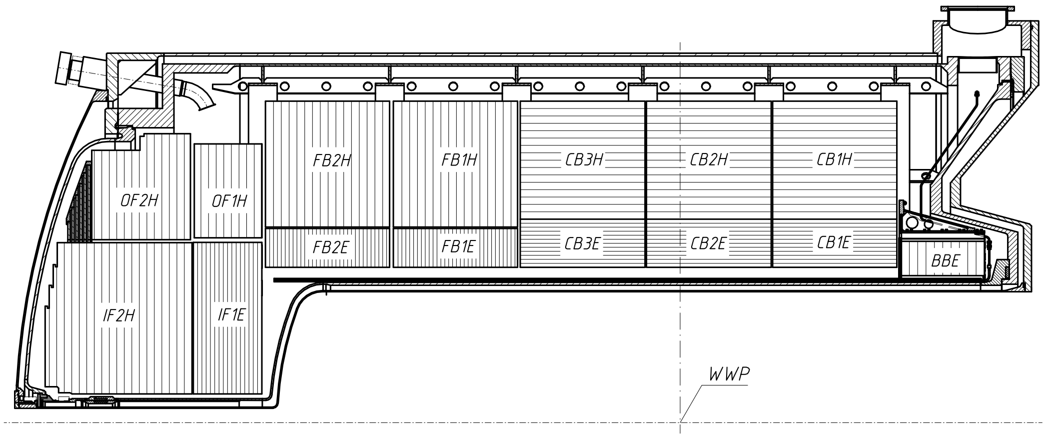


Abbildung 2.3: Schematischer Längsschnitt durch das Flüssig-Argon-Kalorimeter. Zu sehen ist die Ausrichtung der Absorberplatten gegen den nominellen Wechselwirkungspunkt (WWP).

Die elektrisch geladene Teilchen und die Photonen verlieren ihre Energie durch Bremsstrahlung und e^+e^--Paarbildung, während hadronische Teilchen durch inelastische Stöße mit den Atomkernen; in diesen Fällen entstehen entweder elektromagnetische oder hadronische Schauer. Diese sekundäre Teilchenschauer ionisieren das flüssige Argon. Die so freigesetzte Elektronen werden von Ausleseflächen registriert. In diesen Ausleseflächen nehmen die Auslesezellen (**G**eometric **P**ads, **GP**) nur eine Stichprobe des Gesamtschauers auf. Die statistischen Schwankungen dieser Stichprobe beeinflußen und bestimmen die Energieauflösung des Kalorimeters nach der folgenden Formel:

$$\frac{\sigma(E)}{E} \propto \frac{1}{\sqrt{E}}. \tag{2.1}$$

Damit die Energieauflösung möglichst unabhängig von der Richtung der vom Wechselwirkungspunkt kommenden Teilchen ausgemessen wird, sind die Absorberplatten im Zentralteil parallel und im Vorwärts- und extremen Rückwärtsbereich senkrecht zur Strahlachse angeordnet, wie aus der Abbildung 2.3 zu entnehmen ist, so daß die Teilchen unter einem möglichst großen Winkel auf sie treffen. Die echten Energien von hadronischen Teilchen müssen zudem von einem Programm durch ein Gewichtungsverfahren berechnet werden, weil das Kalorimeter nicht kompensierend ist. Das heißt, daß ein elektromagnetisch wechselwirkendes Teilchen ein höheres Signal als ein hadronisches Teilchen erzeugt und von diesem Signal müssen die hadronischen Teilchen erkannt und rekonstruiert werden [10].
Das Kalorimeter verfügt deswegen über eine genug feine Segmentierung, die auch eine gute Ortsauflösung ermöglicht. Die Ausdehnung der Auslesezellen in der longitudinalen Richtung beträgt nämlich nur einige Strahlungslänge und die Fläche entspricht etwa dem Molière-Radius eines Elektrons.

2.2.2 Das Spurkammersystem

Das Spurkammersystem umgibt schalenförmig die Wechselwirkungszone. Es wird ein Winkelbereich von $5° \leq \theta \leq 170°$ abgedeckt (s. Abb. 2.4).
Die geladenen Teilchen, die eine Kammer überqueren, hinterlassen eine Spur von ionisierten Molekülen in dem darin enthaltene Gas. Daher wird ein Strompuls in den inneren Drähten vermessen. Außer der Ausrichtung dieser Spur können dank der vom Magnetfeld gekrümmten Flugbahn (siehe Paragraph 2.2) auch der Impuls und die Ladung des Teilchen bestimmt werden.
Es folgt eine kurze Darstellung der beteiligten Komponenten [8].

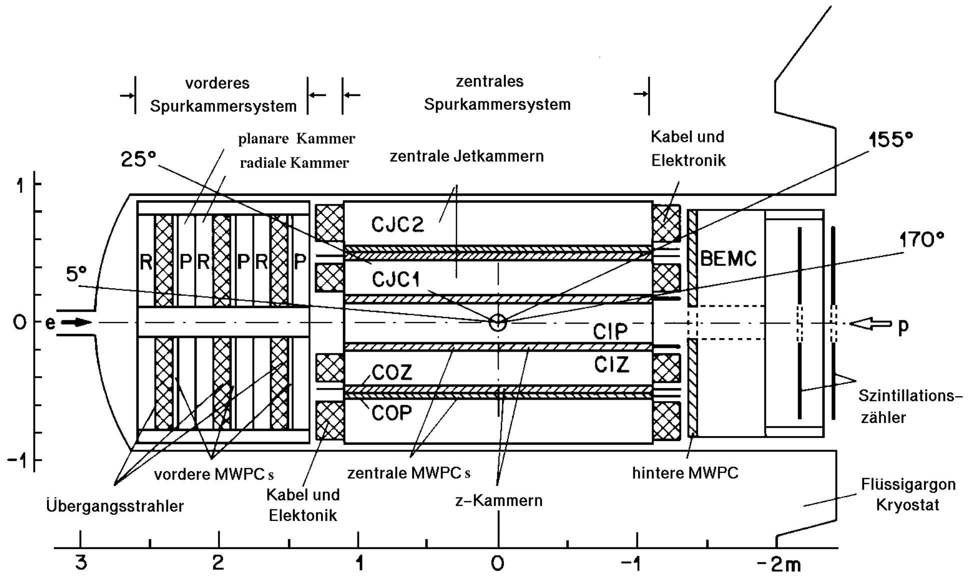


Abbildung 2.4: Seitenansicht des H1-Spurkammersystems.

Die zentralen Jetkammern

Die zwei Driftkammern **CJC1** (**C**entral **J**et **C**hamber **1**) und **CJC2** decken einen Winkelbereich von $25° \leq \theta \leq 155°$ und besitzen pro Ebene 56 Signaldrähte, die parallel zur Strahlachse verlaufen, während die Drahtebenen spiralförmig angeordnet sind (siehe Abbildung 2.5).
Mit dieser Anordnung kann man die r- und ϕ-Koordinate und unpräzise auch die z-Koordinate bestimmen, aber die Spuren können vollständig rekonstruiert werden. Die Ortsauflösung in der $r\phi$-Ebene beträgt $\sigma_{r\phi} = 170\ \mu m$, während die z-Koordinate nur mit einer Genauigkeit von $\sigma_z = 22\ mm$ ermittelt werden kann. Zur Identifikation des Teilchen dient schließlich seinen Energieverlust dE/dx.

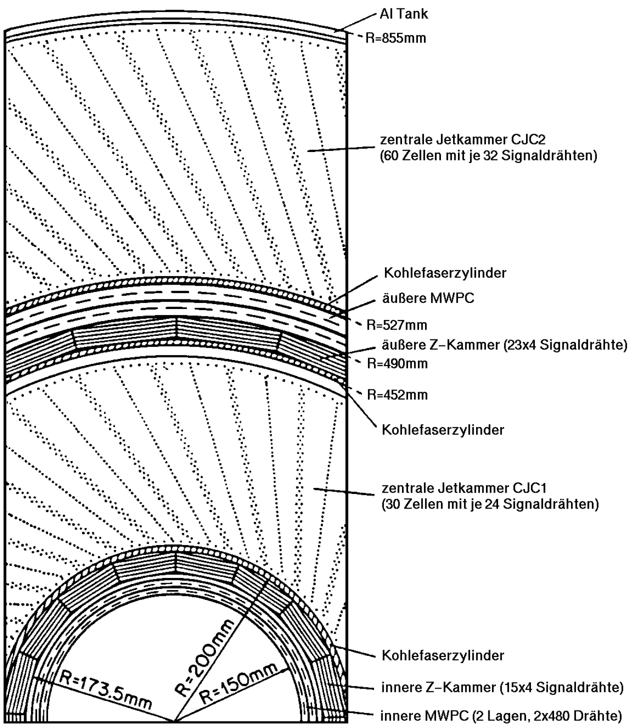


Abbildung 2.5: Radialansicht der zentralen Jetkammern.

Die z-Kammern

Die innere (**C**entral **I**nner **Z**-Chamber, **CIZ**) und äußere (**C**entral **O**uter **Z**-Chamber, **COZ**) z-Kammern bestehen jeweils aus 15 und 24 Auslesezellen mit je vier Signaldrähten. Da diese Drähte konzentrisch zur Strahlachse aufgespannt sind, erlauben sie eine bessere Auflösung in z als es mit den Jetkammern möglich ist. Daher der Name z-Kammern. Die Genauigkeit für die z-Koordinate ist 0,26 mm (**CIZ**) bzw. 0,20 mm (**COZ**).

Die Vieldrahtproportionalkammern

Das Spurkammersystem ist außerdem mit einer Reihe von verschiedenen Vieldrahtproportionalkammern (**M**ulti-**W**ire-**P**roportional-**C**hamber, **MWPC**) ausgestattet.
Die zwei Vieldrahtproportionalkammern (**C**entral **I**nner **P**roportional, **CIP**) und (**C**entral **O**uter **P**roportional, **COP**) bestehen aus drei Kathodenebenen, die zwei Drahtebenen umgeben, deren Lagen um $\pi/16$ gegeneinander verschoben sind, um eine 16fache Segmentierung in ϕ zu realisieren.
Die drei Kathodenebenen sind für die **CIP** in $8\phi \times 60z$ und für die **COP** in $32\phi \times 60z$ sogenannte *Pads* unterteilt. Aus der Koinzidenz von Pad-Signalen in der **CIP** und **COP** läßt sich die z-Komponente des Wechselwirkungsvertex bestimmen.

2.2.3 Die Myonkammern

Myonkammern sind in dem Vorwärtsbereich des Detektors und als nächstliegende Schicht auf der supraleitende Spule zu finden. Sie sind in Front- und Endkappe, vorderen und hinteren Rumpfteil untergliedert. Jeder dieser vier Bereiche ist in 16 Module unterteilt, die wiederum aus 16 Lagen von in das Eisenrückflußjoch integrierten Streamerkammern bestehen (siehe Abbildung 2.2). Es ist unwahrscheinlich, daß Myonen eine Spur im Kalorimeter hinterlassen, weil sie das Gas nur gering ionisieren und wenig Energie dort deponieren. So können sie das Kalorimeter durchdringen und eine eindeutige Aufzeichnung in den Myonkammern hervorrufen. Auch Myonen aus der kosmischen Strahlung können nachgewiesen werden.

3 Die Physik bei H1

Das Hauptziel des **H1**-Experimentes ist, die Protonstruktur zu untersuchen. Damit ist auch eine Behandlung der **QCD** verbunden (s. Kap. 1).
Das theoretische Modell wird somit erforscht und kann von den experimentellen Beobachtungen bestätigt oder auch in Frage gestellt werden.
Dank der erreichten Schwerpunktsenergie von 300 GeV in **HERA** und der guten Eigenschaften des Detektors eignet sich das **H1**-Experiment auch für andere Forschungsgebiete, wie die Suche nach neuen Teilchen, die Photoproduktion und die Erzeugung von schweren Quarks [11].

3.1 Die Kinematik

In der Hochenergiephysik hat man mit solchen Geschwindigkeiten der Teilchen zu tun, daß die Darstellung und die Verfahren der klassischen Mechanik nicht ausreichend sind, um die kinematischen Variablen des Prozesses richtig zu berechnen. Unerläßlich erweist sich die relativistische Mechanik, wo die kinematischen Größen mathematisch als Vierervektoren dargestellt werden und für die die Minkowski-Metrik angewandt wird. Wenn man dieselben Größen in einem anderen Bezugssystem berechnen will, muß man die Lorentz-Transformationen benützen.
Zur Beschreibung eines Streuprozesses ist es deshalb vorteilhaft, solche Größen zu wählen, die Lorentz-Invarianten sind, d.h. sie nehmen in jedem Bezugssystem denselben Wert an. Im **H1**-Experiment wurde das negative Quadrat des Impulsübertrags vom Lepton auf das Quark ausgewählt:

$$Q^2 = -q^2 = -(l'_\mu - l_\mu)^2 \tag{3.1}$$

Die Variablen l'_μ und l_μ sind der Viererimpuls des auslaufenden, bzw. des einlaufenden Leptons. Wenn die Schwerpunktsenergie

$$s = (l + p)^2 \tag{3.2}$$

bekannt ist und der Stoß vollständig elastisch ist, dann beschreibt allein diese Variable den Prozeß und es gilt:

$$Q^2 = 2ME_u \quad \text{mit} \quad E_u = E_{p'} - M \tag{3.3}$$

wobei M die Ruhemasse des hadronischen Streuteilchens und E_u die durch den Stoß übertragene Energie in dessen Ruhesystem bezeichnen (s. Abb. 3.1).
Wenn der Stoß sich inelastisch ereignet, verliert die obige Gleichung ihre Gültigkeit und man findet dagegen:

$$Q^2 < 2ME_u. \tag{3.4}$$

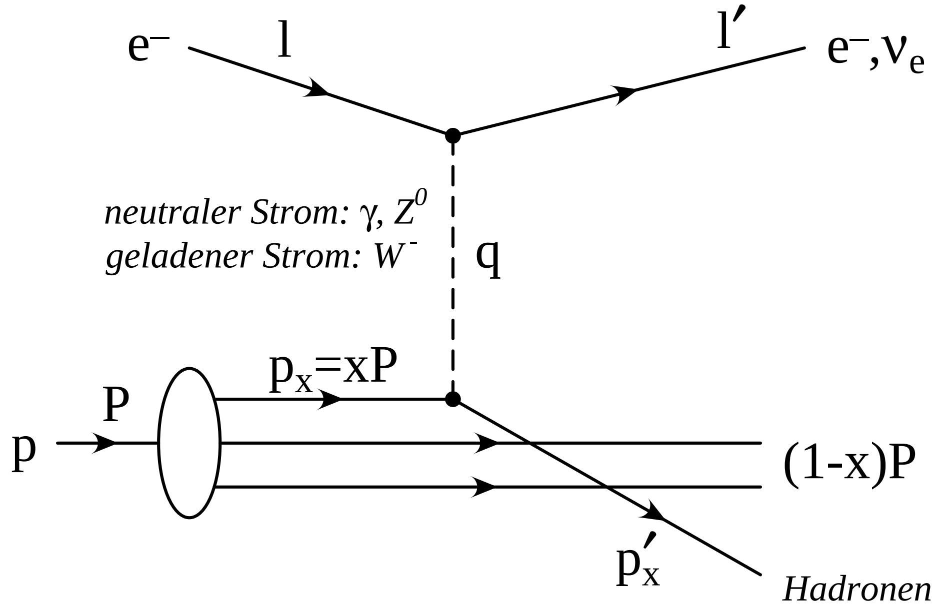


Abbildung 3.1: Feynman-Graph einer Lepton-Quark-Streuung.

Man benötigt deshalb einen weiteren Lorentzskalar (invariante Größe gegenüber Lorentz-Transformationen), um den inelastischen Fall zu beschreiben. Am einfachsten bildet man die Division der obigen Glieder um den folgenden dimensionlosen Wert zu bekommen:

$$x = \frac{Q^2}{2ME_u} \qquad (0 \leq x \leq 1). \tag{3.5}$$

Wenn die inelastische Streuung als elastischer Stoß des Leptons an einem der Konstituenten des Protons interpretiert wird, dann gibt X genau den Impulsanteil des wechselwirkenden Konstituenten am Gesamtimpuls des Protons wider. Analog dazu kann man den normierten Energieübertrag im Ruhesystem des Protons bilden:

$$y = \frac{2ME_u}{s}. \tag{3.6}$$

Die beiden Größen sind mit der folgenden Beziehung verknüpft:

$$Q^2 = sxy. \tag{3.7}$$

Im Experiment werden der Polarwinkel Θ_l' und die Energie E_l' des gestreuten Leptons gemessen. Aus diesen Werten und der Energie des Leptons vor der Streuung E_l lassen sich zwei von den oben erwähnten physikalischen Größen folgenderweise berechnen:

$$Q^2 = 4E_l E_l' \cos^2\left(\frac{\Theta_l'}{2}\right) \tag{3.8}$$

$$y = 1 - \frac{E_l'}{E_l} \sin^2\left(\frac{\Theta_l'}{2}\right). \tag{3.9}$$

Mit der Gleichung (3.7) kann man dann auch X herleiten.

3.2 Die Ereignisklassen

In der niedrigsten Annährung (Bornstreuung) erfolgt der Streuprozeß mit dem Austausch von elektroschwachen Eichbosonen (γ, W^+, W^-, Z^0) [12].
Im allgemeinen können sich zwei Fälle ereignen, die unter den Kathegorien von sogenanntem neutralem Strom (*Neutral Current, NC*) oder geladenem Strom (*Charged Current, CC*) klassifiziert werden (s. Abb. 3.1).

Neutraler Strom

Man spricht von neutralem Strom, wenn kein Ladungsaustausch stattfindet. Das einlaufende Lepton behält seine Ladung, indem es mit dem Proton über ein γ oder ein Z^0 wechselwirkt. Man kann die Spur des Leptons im Detektor nachvollziehen und die Summe der transversalen Impulse der erzeugten Teilchen ist gleich Null (s. Abb. 3.2).

Geladener Strom

Wenn die Wechselwirkung über ein geladenes Boson, wie W^+ oder W^-, stattfindet, handelt es sich um geladenen Strom. Das Lepton wird in ein Neutrino oder ein Antineutrino umgewandelt, das den Detektor durchdringt, ohne Spuren zu hinterlassen. Die Summe der transversalen Impulse bleibt nicht erhalten und aus dem fehlenden Betrag kann das Neutrino erkannt werden.

3.2.1 Physikalische Ereignisse

Für jede Reaktion gibt der Wirkungsquerschnitt σ, der als Übergangsrate von den Anfangszuständen zu den Endzuständen pro Einheitsfluß der Strahlenbündel angesehen werden kann, eine Maß dafür, welches Gewicht eine bestimmte Reaktion in dem Experiment annimmt. Die Stoßrate $\mathcal{R}$ kann dann mit

$$\mathcal{R} = \sigma\mathcal{L}, \tag{3.10}$$

geschrieben werden, wo die Luminosität $\mathcal{L}$ die Beschleunigerparameter enthält.

Photoproduktion

Eine sehr häufige Reaktion, d.h. mit einem großen Wirkungsquerschnitt, repräsentiert die sogenannte Photoproduktion. Es tritt ein virtuelles Photon auf, das als Untergrund für andere Reaktionen störend sein kann. Es kann aber auch mit einem Quark des Protons wechselwirken und Reaktionen höherer Ordnung einleiten. Da das Elektron dabei nur geringfügig abgelenkt wird, ist $Q^2 \approx 0$ [13].

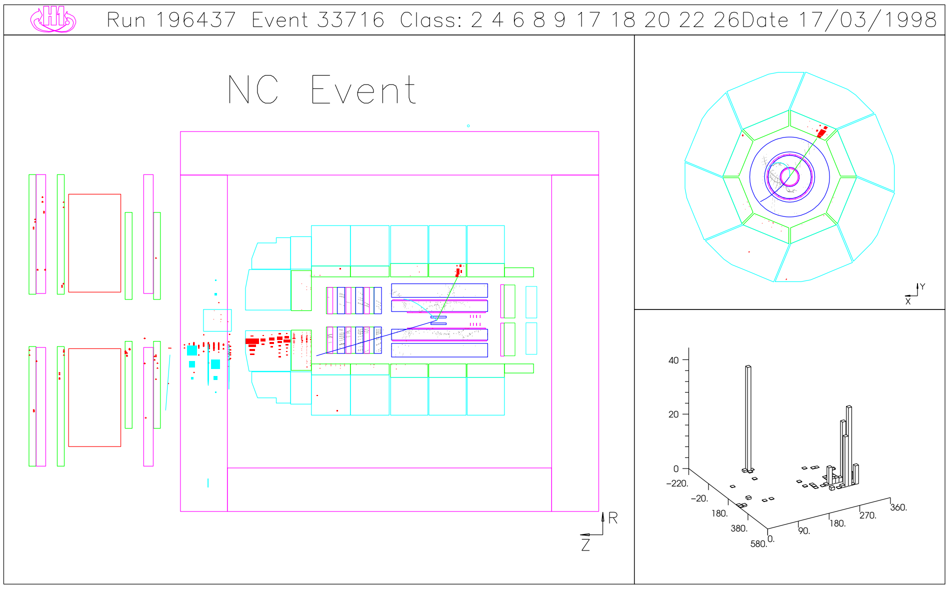


Abbildung 3.2: Ein neutrales Stromereignis im H1-Detektor.

QCD bei kleinem x

Wenn die Variable x sehr klein ($< 10^{-2}$) und Q^2 dagegen groß ist, beginnt die störungstheoretische Behandlung der QCD, nicht mehr zu gelten [14].
Das **H1**-Experiment bietet die Möglichkeit an, diesen Übergangsbereich zu einer nicht störungstheoretischen Behandlung zu erforschen und bei derartigen Reaktionen das Partonmodell durch die Bestimmung der Protonstrukturfunktion zu überprüfen.

Erzeugung von Quarks

Im Inneren des Protons halten die Gluonen die geladenen Quarks zusammen. Wenn ein virtuelles Gluon vom Proton bei dem Stoß abgestrahlt wird, kann es an einem elektroschwachen Boson, meistens dem Photon, durch die Entstehung von einem Quark-Antiquark-Paar koppeln [15].
Auf diese Weise werden leichte Quarks wie *up* (u), *down* (d), *strange* (s) oder die schwere Quarks *charm* (c) und *bottom* (b) produziert. Für die *top* (t) Quarks reicht leider die Schwerpunktsenergie nicht aus (s. Tab. 1.2).

Neue Teilchen

Man erwartet normalerweise neue Teilchen mit hohen Massen, die durch eine große Energiedeposition im Kalorimeter gekennzeichnet sein können.
Ein weiterer Hinweis für die Entdeckung eines neuen Teilchen kann einem hohen fehlenden Transversalimpuls zugeschrieben werden. Die Verwirklichung einer Entdeckung dieser Art würde nicht wenige Schwierigkeiten für das bis heute ziemlich gut bewährte Standardmodell darstellen [16].

3.2.2 Untergrundereignisse

Die Detektoren werden auch von den Produkten nicht gewünschter Reaktionen erreicht, weil Restgas in dem nicht perfekten Vakuum des Strahlrohres immer vorhanden ist und an den Zusammenstößen teilnehmen kann. Solcher sogenannter Untergrund wird von dem Triggersystem (s. Kap. 4) gefiltert und kann meistens anhand des rekonstruierten z-Vertex erkannt werden, der dann üblicherweise von der Wechselwirkungszone weit verschoben ist.
Ein Beispiel davon ist in der Abbildung 3.3 gezeigt, wo man auch die typische Energiedeposition im vorderen Teil des Kalorimeters sehen kann.
In diesem Abschnitt wird auf diese und auf andere mögliche Quellen für Untergrundereignisse eingegangen.

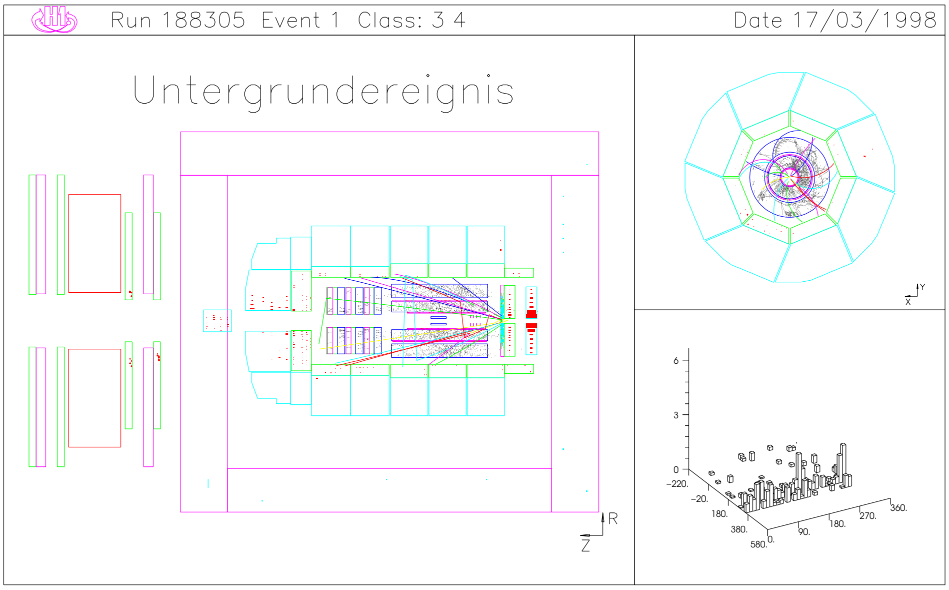


Abbildung 3.3: Ein Untergrundereignis im H1-Detektor.

Reaktionen mit dem Restgas

In dem Strahlrohr herrscht ein sehr niedriger Druck $p \sim 10^{-9} mbar$.
Trotzdem kollidieren nicht selten Protonen mit den Restgasatomen und bilden somit eine große Quelle für Untergrundereignisse, während die analoge Reaktion mit den Elektronen nicht von großer Bedeutung ist.
Da die chemische Zusammensetzung des Restgases nicht genau bekannt ist, sind Abschätzungen der Frequenz solcher Kollisionen schwer durchführbar. Aus Simulationen hat man eine obere Grenze für die Rate gewinnen können, die sich auf $45\,kHz$ beläuft. Dieser Wert ist allerdings mit einem Fehler behaftet, der einen Faktor 5 erreichen kann [17].

Strahl-Wand-Reaktionen

In dem Protonenstrahl fliegen die Protonen in Paketen gebündelt, deren Breite mit der Zeit nicht konstant bleibt. In dem Strahlrohr befinden sich Kollimatoren, die die ausweichenden Protonen entfernen. Bei einer Lebensdauer des Strahls von 20 Stunden kann das eine Verlustrate von $10^5 s^{-1} m^{-1}$ bedeuten. Dennoch stoßen Protonen an Stellen mit kleinerer Apertur gegen die Wand des Rohres und die verursachten Reaktionen spielen die Rolle einer zweiten Untergrundquelle mit einer Rate derselben Größenordnung der Reaktionen mit dem Restgas.

Synchrotronstrahlung

Ein beschleunigtes geladenes Teilchen verliert Energie in Form von elektromagnetischer Strahlung, sogenannter *Synchrotronstrahlung*.
Der Verlust pro Umkreis eines Ringes bei gegebener Geschwindigkeit und Ladung hängt überwiegend von dem Verhältnis zwischen der Energie und der Ruhemasse des Teilchens ab:

$$\delta E \propto \left(\frac{E}{mc^2}\right)^4 . \tag{3.11}$$

In dem **H1**-Experiment ist dieses Verhältnis für Protonen ≈ 900 und für Positronen ≈ 55000, deshalb ist nur die Strahlung der Positronen von Bedeutung. Sie kann in verschiedenen Detektorkomponenten Signale erzeugen.

Kosmische Strahlung

Die kosmische Strahlung dringt ständig in die Erdatmosphäre ein und erzeugt dort durch pN-Reaktionen hochenergetische Myonen, die auf den Detektor zukommen und Signale auslösen.
Wie bei den Reaktionen mit dem Restgas verwirft man den Hauptteil, indem man einen z-Vertex in der Wechselwirkungszone verlangt. Einige Ereignisse mit einer Rate von $0.1\,Hz$ sehen jedoch den *NC*-Ereignissen mit hohem Q^2 ähnlich und bedürfen der Vervollständigung mit detaillierteren Informationen, um erkannt zu werden.

Rauschen

Unvermeidbar ist das weiße Rauschen in der Ausleseelektronik, das besonders das Flüssig-Argon-Kalorimeter betrifft. Mit einer passenden Energieschwelle kann aber das Problem gelöst werden.
Ein kohärentes Rauschen ist dagegen nicht so einfach zu unterdrücken und raffiniertere Techniken müssen angewandt werden.

3.3 Die elastische J/ψ-Produktion

Das Vektormeson J/ψ hat diesen doppelten Namen bekommen, weil es 1974 praktisch gleichzeitig in zwei verschiedenen Experimenten zum ersten Mal in einer Reaktion beobachtet wurde. Im **B**rookhaven **N**ational **L**aboratory (**BNL**) wurde als eine schmale Vektorresonanz, das $J/\psi(3097)$, in der Reaktion

$$pBe \longrightarrow (J/\psi \longrightarrow e^+e^-) + X \tag{3.12}$$

nachgewiesen und am **S**tanford **L**inear **A**ccelerator **C**enter (**SLAC**) als Resonanz im totalen e^+e^--Wirkungsquerschnitt [18].
Im Vergleich zu typischen hadronischen Zerfallsbreiten von $\mathcal{O}(10-200MeV)$ ist seine Breite von $87keV$ sehr schmal. Seine invariante Masse liegt bei $3,097GeV$ und seine Quantenzahlen ($J^{PC} = 1^{--}$) gleichen denen des Photons [19].
Im **H1**-Experiment werden bei der Photoproduktion J/ψ erzeugt, die in ein Leptonpaar zerfallen. Die üblichen Zerfallskanäle sind $J/\psi \longrightarrow \mu^+\mu^-$ und $J/\psi \longrightarrow e^+e^-$. Von dem letzten ist ein Beispiel in der folgenden Figur abgebildet.

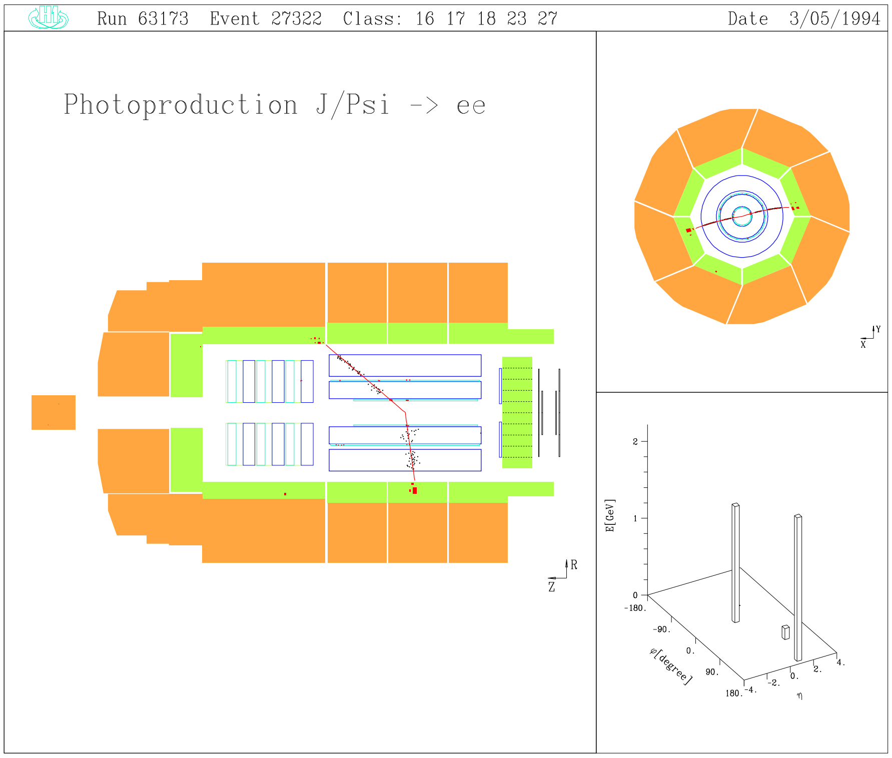


Abbildung 3.4: Ein J/ψ zerfällt in ein e^+e^- Paar.

Die elastische Vektormesonproduktion ist ein diffraktiver Prozeß, das heißt, daß das Vektormeson zuerst eine Fluktuation des vom Elektron abgestrahlten virtuellen Photons ist, die durch diffraktive Streuung am Proton reell wird.

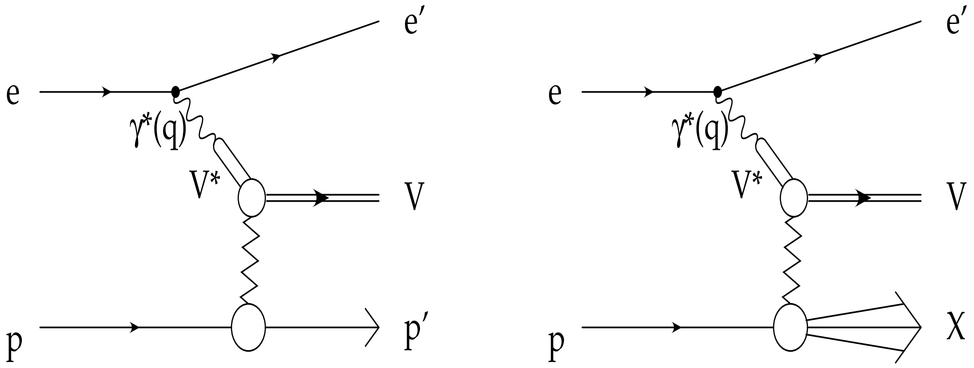


Abbildung 3.5: Diffraktive Vektormeson-Erzeugung: γ^* ist das virtuelle Photon und V^* das virtuelle Vektormeson. Links ist der elastische und rechts der inelastische Fall schematisch dargestellt.

Der Begriff Diffraktion ist zwar aus der Optik entlehnt, ist aber ein nützliches Gedankenmuster in der quantenmechanischen Beschreibung der Teilchen als Wellenpakete für Streuprozesse.

4 Das Triggersystem

Wenn der **HERA**-Speicherring im Betrieb ist, kollidieren alle 96 ns (BC) Positronen und Protonen in der Wechselwirkungszone des **H1**-Detektors (s. Kap. 2).
Wenn man alle Ereignisse aufnehmen möchte, auch diejenigen wo keine Positron-Proton-Streuung stattgefunden hat, ergäbe sich daraus eine Rate von 10,4 MHz. Die gesamten Rohdaten eines einzelnen Ereignisses würden 3 $Mbyte$ Speicherplatz in Anspruch nehmen.
Um den benötigten Speicherplatz auf 50-100 $kbyte$ zu reduzieren, werden die Daten komprimiert und insbesondere wird eine Nullpunktunterdrückung (zero-suppression) ausgeführt. Da aber die maximale Schreibgeschwindigkeit 1,2 $Mbyte/s$ beträgt, können nur bis zu etwa 10 Ereignisse pro Sekunde auf einem Magnetband in digitaler Form geschrieben werden.
Es stellt sich also aus technischen Gründen als unmöglich heraus, die ganze Datenmenge zu speichern. Das wäre aber auch nicht erwünscht, weil eine Vielzahl von Untergrundereignissen neben den interessanten physikalischen Reaktionen auftreten (siehe 3.2.2).
Man erwartet im Durchschnitt 1000 physikalische Ereignisse pro Sekunde, wobei einige Reaktionen wie CC-Ereignisse oder die Erzeugung schwerer Quarks untervertreten sind (siehe 3.2.1). Das soll sich aber mit der Umrüstung der Strahlführungs- und Fokusierungs-Magnete bis zum Jahr 2000 ändern [20].
Die Selektion der Daten wird von einem mehrstufigen Triggersystem vorgenommen (s. Abb. 4.1). Der Trigger reduziert die Eingangsrate von $\sim$10 MHz auf eine Ausgangsrate von 10 Hz, indem er die Untergrundereignisse identifiziert und verwirft. Jede Stufe wirkt wie ein Filter auf die Datenmenge und sobald ein Untergrundereignis an einer Stufe identifiziert ist, werden seine Daten nicht weitergeleitet. Einige davon werden trotzdem gesammelt und werden auf Magnetband gespeichert; andere passieren alle Stufen, weil sie nicht erkannt werden. Wegen der hohen Rate können auch nicht alle Physikereignisse aufgenommen werden und viele gehen wegen vorgegebener Schwellen (Prescale-Faktor) verloren.
Die Auslese der über 270000 elektronischen Kanäle des Detektors benötigt $\sim$1,5 ms, die sogenannte *Totzeit* und während dieser Zeitspanne können keine neuen Daten genommen werden. Die vollständige Auslese wird aber erst durchgeführt, wenn ein Ereignis alle Stufen durchgelaufen hat.
Jede Stufe hat eine bestimmte Zeitbeschränkung, um die Daten zu analysieren, und bewirkt eine spezifische Datenreduktion. Die Zahlen sind in der Abbildung 4.1 angegeben. Die fünfte Stufe gehört nicht zu dem Trigger, sondern dient zur Datenrekonstruktion.
In dem folgenden Abschnitt werden die einzelnen Triggerstufen und das Verfahren, wodurch der Trigger zu einer Entscheidung kommt, beschrieben.

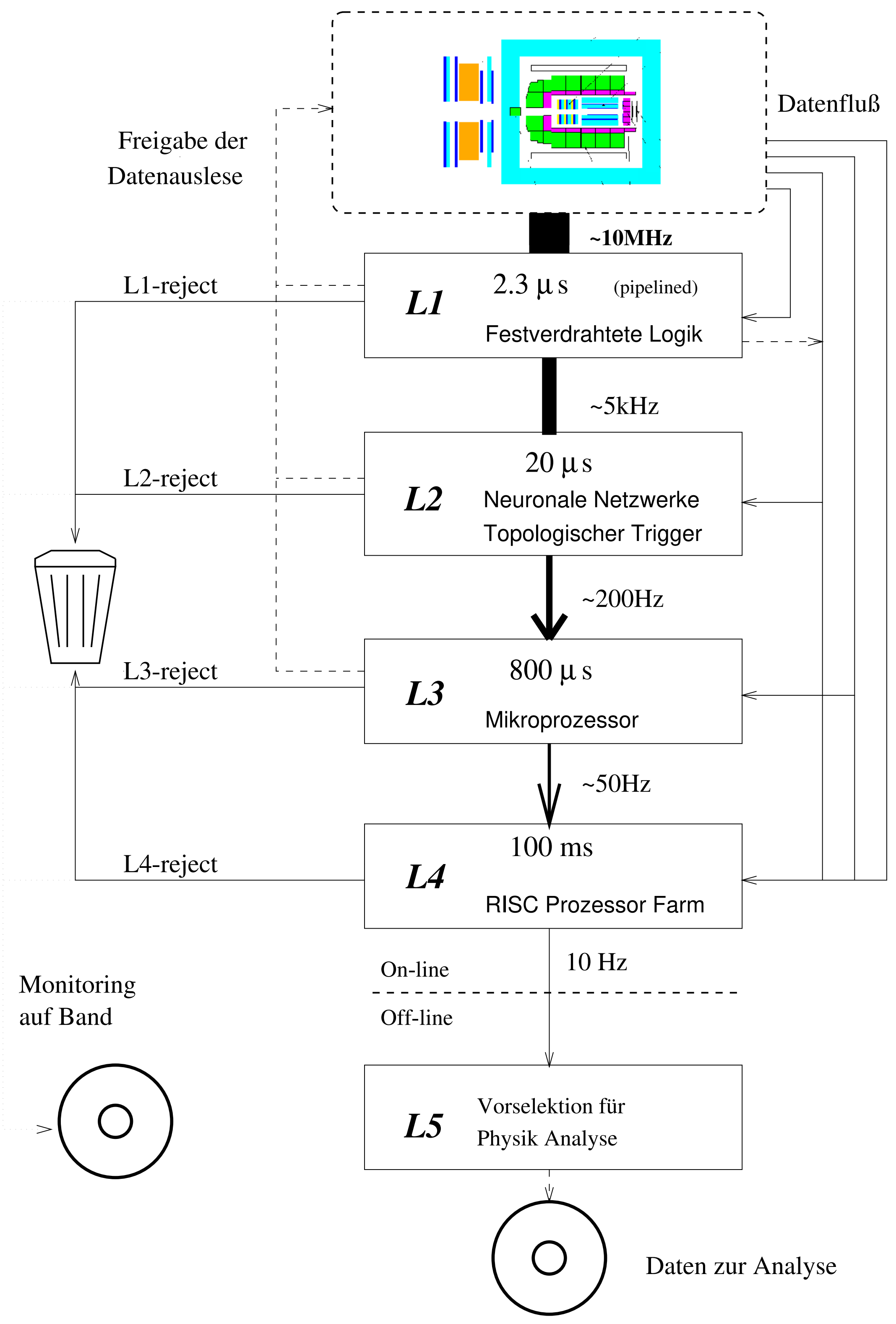


Abbildung 4.1: Schematischer Aufbau des H1-Triggersystems.

4.1 Die Triggerstufen

Die Detektordaten können nicht gleich nach jedem Zusammenstoß dem Triggersystem übergeben werden, weil einige Detektorkomponenten mehrere BC brauchen, um das Signal zu verarbeiten. Beispielweise dauert die Integration der Signale bei den Vorverstärkern des Flüssig-Argon-Kalorimeters 13 BC. Die Triggerinformation wird deshalb erst in einem 2,6 μs langen Pipeline-Zwischenspeicher abgelegt. Wenn die erste Triggerstufe entscheidet, daß das Ereignis behalten werden soll, wird die Pipeline angehalten. Die anderen Triggerstufen fangen inzwischen schon an, das Ereignis auszuwerten. Verwerfen sie es, wird die Auslese abgebrochen und die Pipeline erneut gefüllt.

Triggerstufe 1 (L1)

Die elektrischen Schaltungen der Detektorkomponenten bilden einfache und übergreifende Größen wie Energiesummen oder Spurmultiplizitäten, die jeweils einem Triggerelement (*TE*) zugeordnet sind. Wenn eine Größe eine festgelegte Schwelle überschreitet, wird das entsprechende *TE* aktiviert und diese binäre Information (an oder aus) wird an die zentrale Triggerlogik (*CTL*) übertragen. Hier können bis zu 128 sogenannte Subtrigger durch logische Verknüpfungen zueinanderpassender *TE*s gebildet werden. Wenn ein Subtrigger eine zu hohe Rate erzeugt, wird diese mittels eines Prescale-Faktors F reduziert, indem er nur in jedem F-ten aktivierende Ereignis als tatsächlich an betrachtet wird. Dieser Prozeß erfolgt in der für L1 zur Verfügung stehenden Zeit von 2,3 μs. Wenn einer der Subtrigger gesetzt wird, beginnt die Auslese des Detektors. Gleichzeitig liefern einige Triggerelemente ein Zeitsignal (t0), um den Ereigniszeitpunkt festzulegen und die Driftzeiten der Teilchenspuren zu berechnen.

Triggerstufe 2 (L2)

Die zweite Triggerstufe besteht aus zwei unterschiedlich konzipierten Systemen. Sie bekommt alle in L1 gebildeten Größen. Ein topologischer Trigger trägt sie in θ-ϕ Karten ein und berechnet mit einem eingebauten Algorithmus den „Abstand" zwischen Physik und Untergrund [21].
Ein neuronaler Trigger versucht anhand einer Mustererkennung über die in einem hochdimensionalen Raum geordneten Größen, Physik und Untergrund zu unterscheiden. Er basiert auf künstlichen neuronalen Netzen (s. Kap. 5).
Dieser Trigger wird ausführlicher in dem nächsten Abschnitt beschrieben.
Beide Systeme selektieren ausgewählte Physikklassen, um den Prescale-Faktor auf bestimmten L1-Subtriggern zu verringern. Insgesamt dürfen maximal 20 μs verbraucht werden.

Triggerstufe 3 (L3)

Das Projekt der dritten Triggerstufe ist noch nicht verwirklicht worden.
Geplant ist, daß ein freiprogrammierbares System von AM 29000 RISC Prozessoren außer der vorliegenden L1- und L2-Informationen auch ein Teil der genaueren Detektorinformation benützt, die nach einem von L2 akzeptierten Ereignis (L2-keep) ausgelesen wird. Dafür sind 800 μs vorgesehen worden.

Triggerstufe 4 (L4)

Wenn ein Ereignis von den untergeordneten Triggern angenommen worden ist, steht die gesamte Detektorinformation in Rohdatenformat an dieser Stufe zur Verfügung. In einer Prozessorfarm bewerten 36 parallel arbeitende Prozessoren vom Typ MIPS R3000 diese Daten und führen eine Teilrekonstruktion durch. Nach diesem Schritt werden die rekonstruierten Daten auf Magnetband mit der reduzierten Schreibrate von 10 Hz gespeichert.

Triggerstufe 5 (L5)

Die letzte Triggerstufe läuft *offline* und dient nur zur Ereignisklassifizierung und hat keine Filteraufgabe. Unmittelbar nach Aufzeichnung ordnet sie die auf Band geschriebenen Ereignisse in Physikklassen ein.

4.2 Der L2 Trigger

Die Daten der H1-Subdetektoren erreichen den L2 Trigger über einen 128 *bit* breiten Bus. Der L2 Trigger besteht aus 16 Trigger Boxen, von denen nur 12 aktuell ausgestattet und aktiv sind.
Jede Trigger Box wird mit einem **D**ata **D**istribution **B**oard (**DDB**) und einem Parallelrechner **CNAPS** (**C**onnected **N**etwork of **A**daptive **P**rocessor**S**) gebaut. Auf den Parallelrechnern werden die Ausgabewerte der neuronalen Netze berechnet. Die Aufgabe der **DDB** ist die Aufbereitung der Daten für den zugeordneten Parallelrechner in einem Dateneingabevektor. Noch in diesem Jahr soll ihre Funktionalität mit einer neuen Generation erweitert werden und diese soll mit der Implementierung von neuen Algorithmen in der Lage sein, eine Vorverarbeitung der Daten anzubieten, die die Effizienz der neuronalen Netze verbessert [22].
In dieser Arbeit werden einige Algorithmen zu diesem Zweck simuliert und getestet (s. Kap. 7). Die Steuerung des ganzen Systems erfolgt durch den `h1nnsun`-Rechner in Hamburg oder über Internet. Im Rahmen dieser Arbeit wurde ein *offline* Rate-Display-Programm entwickelt, das die Raten der akzeptierten und der verworfenen Ereignisse inklusiv ihr Verhältnis (Ratio) für jede Trigger Box zeigt (siehe Anhang C). Auf der nächsten Seite wird ein Beispiel geschildert.

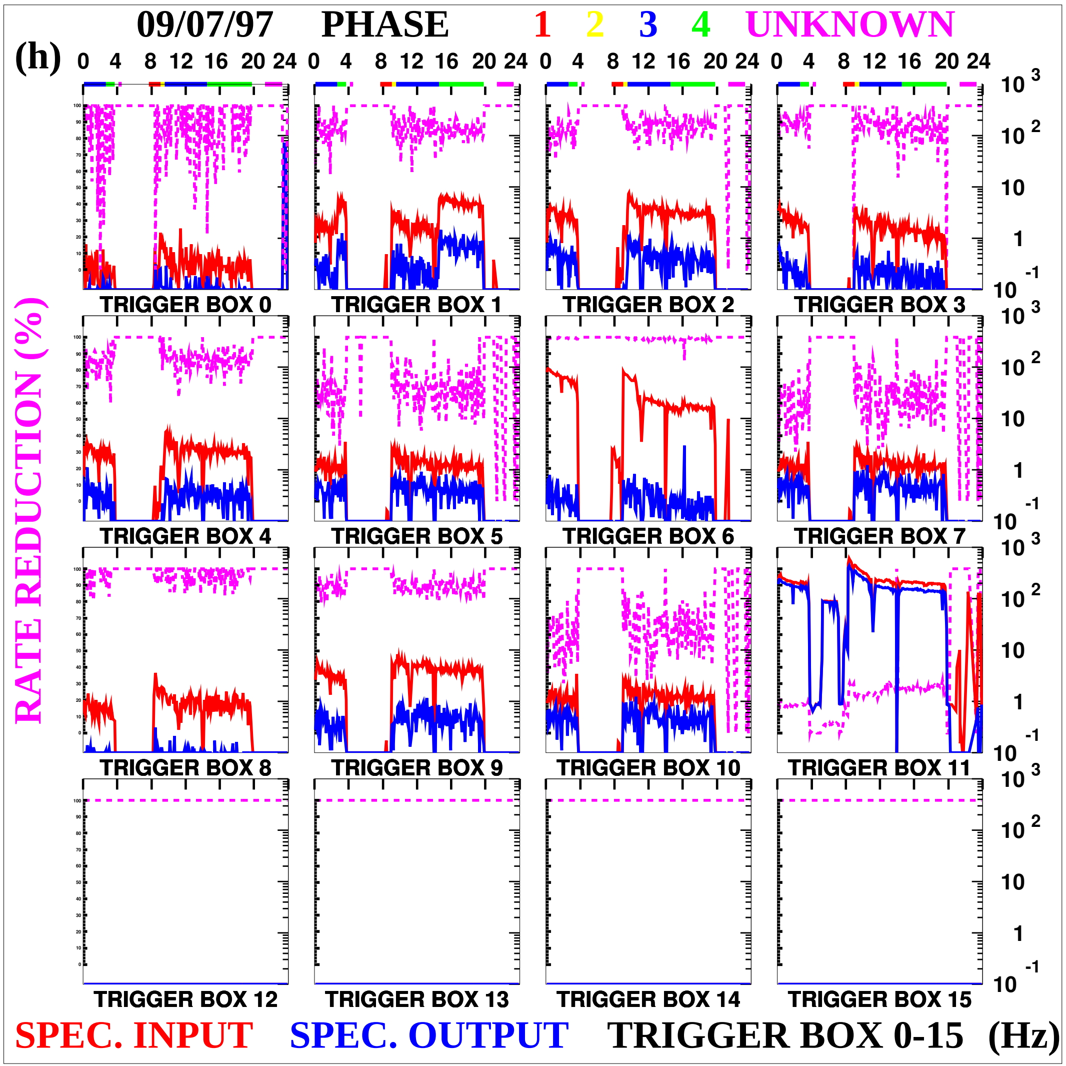


Abbildung 4.2: Die Raten des L2 Triggers vom 9. Juli 1997.

Außerdem sorgt eine Gruppe von anderen Programmen für die *online* Überprüfung und eventuelle Rekonfiguration der einzelnen Trigger Boxen.

5 Neuronale Netze

Die künstlichen neuronalen Netze (**KNN**) bestehen aus mehreren Basiseinheiten, den Neuronen, die miteinander verbunden sind. Diese Neuronen ahmen in ihrer Struktur und Verfahren die biologischen Neuronen im z.B. menschlichen Gehirn nach. Die Stärke der KNN besteht darin, daß sie in der Lage sind, zu lernen und sich so an unerwarteten Ereignissen besser anzupassen. Sie sind im Vergleich zu herkömmlichen Algorithmen besonders in der Mustererkennung erfolgreich und werden dafür in der Hochenergiephysik auch meistens eingesetzt.

5.1 Das Modellneuron

Die Modellbildung des Neurons geht auf W. McCulloch und W. Pitts zurück, die 1943 es in einer ihrer biophysikalischen Arbeiten vorstellten [23].

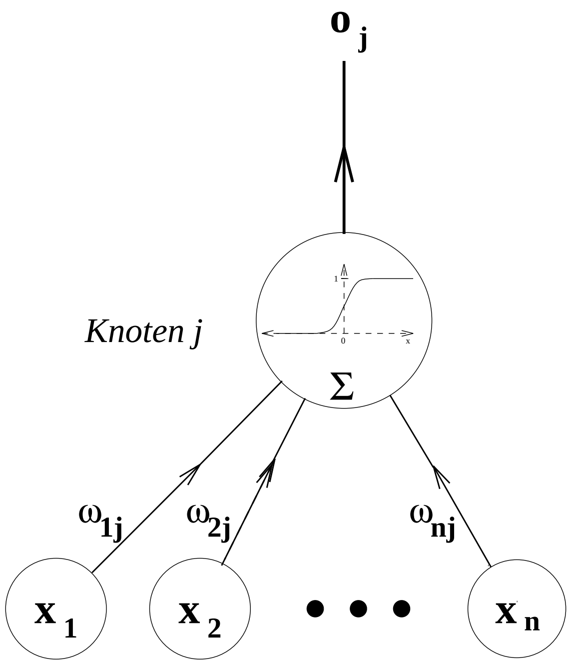


Abbildung 5.1: Mathematisches Modell des Neurons.

Seit damals hat das Modell sich weiter entwickelt, indem auch reelle Werte anstelle binäre Größen benützt werden. Wenn man dann von dem theoretischen Modell zu einer elektronischen Implementierung übergeht, ist man auf

die nährungsweise Repräsentation der Zahlen durch die Schaltung beschränkt. Wie in der Abbildung 5.1, besitzt ein Neuron, das wegen seiner Zugehörigkeit an einem Netz als *Knoten j* gekennzeichnet wird, mehrere Eingangskanäle x_i und einen einzigen Ausgangskanal o_j. Die Eingangssignale können die zu bearbeitende Daten oder die Ausgabewerte von anderen Neuronen sein. Jedes einzelne $x_i \quad (i = 1, ..., n)$ wird mit dem entsprechenden Gewicht w_{ij} multipliziert und danach wird die Summe Σ von den so gewichteten Eingangssignalen gebildet.
Diese Summe wird mit einer vom Neuron spezifischer Schwelle t_j verglichen. Diese Schwelle kann auch als ein festes vorgegebenes Eingangssignal mit einem Gewicht von -1 angesehen werden. Ist die Summe größer als die Schwelle, feuert das Neuron, d.h. es wird ein Ausgangssignal o_j erzeugt und weitergeleitet; ansonsten bleibt das Neuron inaktiv. In einer mathematischen Form entspricht dies der Anwendung einer Aktivierungsfunktion g auf die gewichtete Summe:

$$o_j = g\left(\sum_{i=1}^{N} \omega_{ij} x_i - t_j\right). \tag{5.1}$$

In dem oben beschriebenen Fall handelt es sich um die Heavyside-Stufenfunktion, wie sie auch in dem ersten Modell verwendet wurde. Sie hat aber den Nachteil, nicht stetig und differenzierbar zu sein, wie es für den im weiteren beschriebenen Lernalgorithmus notwendig ist. Es wurde deshalb später die differenzierbare, sigmoidale Fermi-Funktion

$$g(x) = \frac{1}{1 + \exp(-x)} \tag{5.2}$$

bevorzugt, die sich von der Heavyside-Stufenfunktion nur in einem den Schwellenwert umgebenden kleinen Intervall unterscheidet.

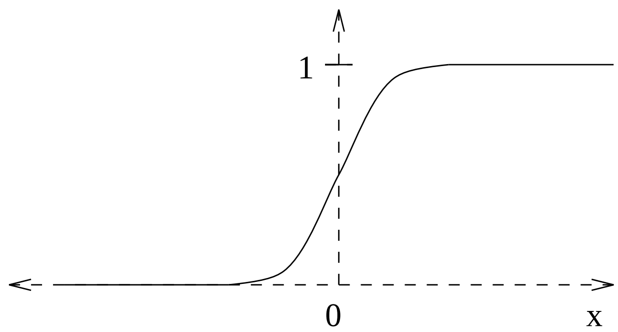


Abbildung 5.2: Die Fermi-Funktion als sigmoidale Aktivierungsfunktion.

Im laufen eines Lernverfahrens werden die Gewichte w_{ij} nach einem vorgegebenen Algorithmus mehrmals verändert. So wird die Antwort des Neurons für die vorgestellten Daten immer mehr verbessert. In der Anpassung der Gewichte widerspiegelt sich der Lernprozeß.

5.2 Das Feedforward-Netz

Was ein neuronales Netz leisten kann, wird in erster Linie durch seine Topologie mitentschieden. Eine bestimmte Architektur dient zu einer besonderen Aufgabe. Außerdem beeinflußen seine Funktionalität die Art der Verknüpfung zwischen den Neuronen und die Anzahl der Neuronen, d.h. der Umfang des Netzes. Für ein Klassifizierungsproblem, wie es sich in der Hochenergiephysik stellt, werden in den meisten Fällen vorwärtsgekoppelte (**Feedforward**) Netze verwendet. In der folgenden Figur ist ein dreilagiges Netz dargestellt.

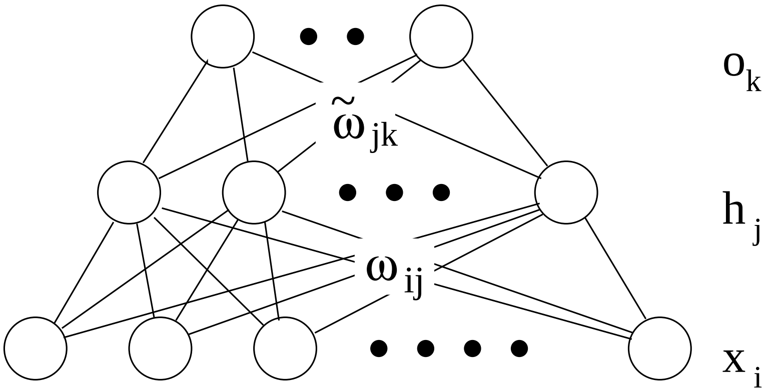


Abbildung 5.3: Ein dreilagiges Feedforward-Netz. $\vec{x}$ ist der Eingabevektor, $\vec{o}$ der Ausgabevektor und $\vec{h}$ der Ausgabevektor der versteckten Schicht.

Das Adjektiv vorwärtsgekoppelt bezieht sich auf die Signalrichtung: in jeder Schicht wird ein Neuron nur mit denen verbunden, die in der nächsten übergeordneten oder untergeordneten Schicht liegen. Die erste Schicht empfängt die Eingangssignale von einer äußeren Datenquelle und rechnet ihre Ausgabewerte, die als Eingangssignale für die zweite Schicht gelten und so weiter, bis die Ausgabewerte der letzten Schicht die Antwort des gesamten Netzes repräsentieren. Die Anzahl der Schichten ist frei wählbar, aber aus theoretischen Erkenntnissen beschränkt man sich gewöhnlich auf dreilagige Netze, die die meisten Klassifizierungsprobleme erfolgreich handeln können.
Diese Behauptung wird von dem Kolmogorov-Theorem abgeleitet, das folgendes besagt: „Jede kontinuierliche m-dimensionale Funktion kann unter nicht sehr strengen Randbedingungen durch ein dreilagiges Netz mit n Eingabe-, $2n+1$ versteckten und m Ausgabeknoten approximiert werden" [24].
Zweilagige Netze reichen dagegen nicht, wenn man nicht lineare Trennungen in der Klassifikation berücksichtigen will. In einem früheren Modell, das *Perzeptron*, fehlte die versteckte Schicht (*hidden layer*). Dieses Netz konnte gut linear separierbare Probleme lösen, aber es scheiterte auch an dem einfachsten nicht linearen Problem wie dem exklusiven Oder. Daran kann man die Bedeutung und die Funktion der versteckten Schicht erkennen. Sie sorgt dafür, daß die nicht lineare Trennbarkeit in eine intern linear separierbare Darstellung abgebildet wird.

In der L2-Gruppe am Max-Planck-Institut für Physik wird derzeit auch mit vierlagigen Netzen experimentiert, um Probleme höherer Komplexität zu lösen. Eine andere Eigenschaft von mehrlagigen Netzen ist ihre Vollvernetzung, wie das Netz in der Abbildung 5.3 aufweist. Jedes Knoten ist in diesem Fall mit allen Knoten der unterliegenden Schicht verbunden.
Man kann absichtlich ausgewählte Verknüpfungen wegfallen lassen, um eine besondere Reaktion des Netzes hervorzurufen oder Symmetrien in den Eingangsdaten wiederzuspiegeln. Für die Leistungsfähigkeit des Netzes ist aber der Lernalgorithmus zweifellos am bedeutendsten, weil er entscheidet, wie sich die Gewichte in den Verknüpfungen ändern sollen.
In dem nächsten Abschnitt wird der Backpropagation-Algorithmus behandelt, der sich als erfolgreich für Klassifizierungsprobleme mit dieser Netztopologie erwiesen hat.

5.3 Der Backpropagation-Algorithmus

Der Lernvorgang für ein neuronales Netz wird *Training* genannt. Während des Trainings wird ein Musterpaar $(\vec{x}, \vec{y})$ in das Netz eingegeben, wo $\vec{x}$ der Eingabevektor und $\vec{y}$ der gewünschte Ausgabevektor ist. Das Netz wird einen Ausgabevektor liefern, der von dem vorgegebenen ausweichen wird.
Man muß den entstandenen Fehler ausrechnen und wie ein Lehrer dem Netz mitteilen, wieviel es falsch lag und in welchem Maß es seine Gewichte ändern muß, um den Fehler zu minimieren. Der Einfachkeit halber wird der Vorgang für ein dreilagiges Netz erklärt, aber es wird auch klar sein, wie er für ein beliebiges mehrlagiges Netz zu verallgemeinern ist (siehe Abbildung 5.3).
Seien N und L die Anzahl der Neuronen in der ersten bzw. zweiten Schicht. Die erste Schicht repräsentiert nur die Elemente des Eingabevektors. Ein Neuron der versteckten Schicht rechnet seinen Ausgabewert nach der Formel 5.1:

$$h_j = g\left(\sum_{i=1}^{N} \omega_{ij} x_i - t_j\right). \tag{5.3}$$

Ein Neuron der letzten Schicht benützt als Eingabevektor den Ausgabevektor der versteckten Schicht und rechnet gleichermaßen seinen Ausgabewert:

$$o_k = g\left(\sum_{j=1}^{L} \tilde{\omega}_{jk} h_j - t_k\right). \tag{5.4}$$

Es wird dann der mittlere quadratische Fehler zwischen dem vorgegebenen Ausgabevektor $\vec{y}$ und dem ausgerechneten $\vec{o}$ ermittelt:

$$E(G) = \frac{1}{2}|\vec{y}(\vec{x}) - \vec{o}(\vec{x})|^2, \tag{5.5}$$

wo G für die Gesamtheit der Gewichte steht. Dieser Fehler hängt von der Wahl der Gewichte ab und das Ziel ist, ihn mit einer passenden Wahl so klein wie möglich zu machen, d.h. sein Minimum zu finden. Die beste Leistung würde das Netz erreichen, wenn das Minimum über alle möglichen Eingabevektoren gesucht wird:

$$E_\infty(G) = \frac{1}{2}\int\limits_A |\vec{y}(\vec{x}) - \vec{o}(\vec{x}, G)|^2 \rho(\vec{x}) d\vec{x}, \tag{5.6}$$

wo A der kompakte n-dimensionale Raum der Eingabedaten und $\rho(\vec{x})$ die Wahrscheinlichkeitsverteilung derselben ist.
Das ist natürlich nicht möglich und man kann nur das Minimum der Funktion mit der vorhandenen endlichen Anzahl P von Eingabevektoren suchen:

$$E_P(G) = \frac{1}{2}\sum_{l=1}^{P} |\vec{y}(\vec{x}_l) - \vec{o}(\vec{x}_l, G)|^2. \tag{5.7}$$

Wenn die beiden Minima sich stark unterscheiden, lernt dann das Netz die vorgestellten Muster auswendig, aber es gelingt ihm nicht, davon zu abstrahieren und scheitert bei der Erkennung von neuen Mustern.
Das Problem ist als *overfitting* bekannt und kann auftreten, wenn die Anzahl der für das Training vorhandenen Muster zu klein ist oder die Muster nicht genug repräsentativ der gesamten Datenmenge sind.
Das Minimum wird mittels einem Gradientenverfahren über schrittweise Annährungen und eine rückwartsrichtete Änderung der Gewichte in den Schichten (daher der Name Backpropagation) gesucht. Am Anfang werden die Gewichte mit zufälligen Werten aus einem Intervall $[-\epsilon, \epsilon]$ initialisiert, um eventuelle Symmetrien zwischen Gewichtsvektoren aufzuheben. Nach der Bearbeitung eines Musters wird der quadratische Fehler (5.5) berechnet.
Die Gewichte zwischen Ausgabe- und versteckter Schicht rücken dann nach der folgenden Formel in Richtung des stärksten Abstiegs auf der Oberfläche der Fehlerfunktion:

$$\Delta\tilde{\omega}_{jk} = -\eta\frac{\partial E}{\partial\tilde{\omega}_{jk}} + \alpha\Delta\tilde{\omega}_{jk}^{alt} = -\eta\delta_k h_j + \alpha\Delta\tilde{\omega}_{jk}^{alt} \tag{5.8}$$

$$\text{mit } \delta_k = (o_k - y_k)\cdot g'(\widetilde{\vec{\omega}_k}\cdot\vec{h}). \tag{5.9}$$

Mit einer Wiederholung der Prozedur ändern sich die Gewichte zwischen Eingabe- und versteckter Schicht:

$$\Delta\omega_{ij} = -\eta\frac{\partial E}{\partial\omega_{ij}} + \alpha\Delta\omega_{ij}^{alt} = -\eta\delta_j x_i + \alpha\Delta\omega_{ij}^{alt} \tag{5.10}$$

$$\text{mit } \delta_j = \sum_{k=1}^{M}\tilde{\omega}_{jk}\delta_k g'(\vec{\omega}_j\cdot\vec{x}), \tag{5.11}$$

wobei M die Anzahl der Neuronen in der Ausgabeschicht ist.
In den obigen Formeln treten zwei Parameter auf, die man nicht a priori, sondern nur durch Versuche mit verschiedenen Werten bestimmen kann.

Die Lernrate η repräsentiert die Schrittweite. Wenn sie zu groß ist, wird das Minimum wahrscheinlich immer wieder übersprungen. Wenn sie zu klein ist, kann es passieren, daß die Suche in einem lokalen Minimum landet und sich dort stabilisiert, ohne das globale Minimum zu finden.
Das Momentum α dämpft die Addition der Gewichtsänderungen $\Delta\omega_{ij}^{alt}$ und $\Delta\tilde{\omega}_{jk}^{alt}$ aus dem vorangegangenen Lernschritt. Diese Maßnahme hilft bei Regionen mit flachem Gradienten und vermeidet das Zustandekommen von Oszillationen.
Dieser Algorithmus wurde 1974 von P. Werbos vorgestellt; berühmt und anwendungstüchtig wurde er aber erst im Jahr 1986 nach der Wiederentdeckung von McClelland und Rumelhart [25].

5.4 Neuronale Netze im L2 Trigger

Die bisher beschriebene Art von neuronalen Netzen wird in dem L2 Trigger eingesetzt. Es handelt sich genau um dreilagige vollvernetze vorwärtsgekoppelte neuronale Netze, die den Backpropagation-Algorithmus als Lernvorgang benützen. Ihre Aufgabe ist, die Eingabedaten in zwei Klassen zu teilen: Physik oder Untergrund.
Dafür benötigen sie ein einziges Neuron in der Ausgabeschicht, das einen Wert zwischen 0 und 1 liefert. Man setzt eine Schwelle t_0, die aus der im Training erreichten Effizienz hergeleitet wird, und bekommt die Triggerfunktion:

$$f_t(\vec{x}) = \begin{cases} 1; & o_0(\vec{x}) - t_0 > 0; \text{ Physik} \\ 0; & o_0(\vec{x}) - t_0 \leq 0; \text{ Untergrund} \end{cases} \tag{5.12}$$

Jedes neuronale Netz ist auf eine bestimmte Physikklasse gerichtet und mit Größen trainiert worden, die diese Klasse am besten charakterisieren.
Die Effizienz ist für jedes Netz unterschiedlich und hängt von der Wahl der Netzparameter η und α während des Trainings ab. Für den L2 Trigger wurde das Netz *offline* trainiert, bis es eine stabile Gewichtenkonfiguration erreicht hatte. So ist die Schwelle t_0 der einzig entscheidende Parameter: je höher er ist, desto mehr Untergrundereignisse werden verworfen und desto mehr Physikereignisse gehen verloren.
Für jede Wahl von t_0 zwischen 0 und 1 ist der Prozentsatz an erkannten Untergrundereignissen und an erkannten Physikereignissen von der Effizienz des Netzes festgelegt. Letztendlich muß man hier einen Kompromiß suchen und sich entschließen, wieviele Physikdaten man bereit ist, zu opfern, um eine bestimmte Quote an Untergrunddaten von der Aufzeichnung auszuschließen.

6 Die Datenselektion

Um die Clusteralgorithmen zu untersuchen, wird eine Datenselektionen der elastischen J/ψ-Produktion aus dem Jahr 1997 verwendet. Da diese Physikklasse schon Gegenstand einer Analyse in der **H1**-Gruppe am **MPI**-München war, kann es auf diese Ergebnisse zurückgegriffen werden und es muß keine neue Analyse dieses Prozesses durchgeführt werden. Die Anzahl der Ereignisse kann aber für größe neuronale Netze zu gering sein. Um diese Beschränkung umzugehen, wurden andere J/ψ-Ereignisse aus einer Monte-Carlo-Simulation genommen. Für den Untergrund wurden **L2/L4**-*Transparent Runs* von 1997 ausgesucht. Es handelt sich um Daten, die von den Triggern **L2** und **L4** nicht gefiltert werden und daher auch Untergrundereignisse enthalten, die an den **L2** Trigger ankommen. Um die Untergrunddaten mit der Ereignisklasse der Physikdaten verträglich zu gestalten, wurden dieselbe Selektionskriterien für beide angewandt. Dieselbe Subtrigger-Bedingungen, die im Abschnitt 6.2 erläutert werden, wurden für beide Datenklassen verlangt.

6.1 L2/L4 Transparent Runs

Um neuronale Netze zu trainieren, muß man sie mit Daten der ausgewählten Physikklasse und des allgemeinen Untergrundes füttern. Die Untergrunddaten wurden aus den **L2/L4**-*Transparent Run* extrahiert, die in der Tabelle 6.1 aufgelistet sind. Für jede Runnummer ist die Anzahl der Ereignisse angegeben.

Runnummer	**Datum**	**Anmerkungen**	**Ereignisse**
180061	16.03.1997	Phase 3, all in, all HV ON	1676
180064	16.03.1997	Phase 3, all in, all HV ON	1345
180068 bis 180070	16.03.1997	Phase 3, all in, all HV ON	1355 1328 1373
180074	16.03.1997	Phase 3, all in r/o	4557
		FPC: OFF evs. 1 - 1700	230
201370	11.10.1997	Phase 2, all in r/o	165
		COZ: OFF evs. 1 - End	(165 + 20)
		FPC: OFF evs. 26308-34541	20

Tabelle 6.1: Zusammenfassung der ausgewählten Runnummer mit der Anzahl der Ereignisse, die die gestellten Subtrigger-bedingungen erfüllt haben.

6.2 Subtrigger-bedingungen

Als Bedingung für die Datenselektion wurde eine logische **ODER**-Verknüpfung zwischen den Subtriggern **S52** und **S54** gestellt. Welche L1-Triggerelemente das beinhaltet, ist es kurz in der folgenden Tabelle geschildert. Die einzelnen Begriffe werden weiter unten erklärt.

Gemeinsame Bedingungen	
Name	*Beschreibung*
zVtx_Cls	all hist. entries within 4 neighbouring bins
DCRPh_TNeg	at least x (=1) negative masks, (x progr'ble)
Subtrigger 52 spezifisch	
Name	*Beschreibung*
LAr_BR	BT above thresh and validated by MWPC track
Topo_Mult < 3	Big Ray multiplicity
Subtrigger 54 spezifisch	
Name	*Beschreibung*
DCRPh_THig	at least x (=1) pt > 800 MeV masks, (x prog.)
Topo_BR	2 and only 2 Rays back to back

Tabelle 6.2: Die L1-Triggerelemente der Subtrigger 52 und 54.

zVtx_Cls

Der Bereich $\pm 43,9\,cm$ um den nominellen Wechselwirkungspunkt wird auf einem *z-Vertex-Histogramm* mit den Strahlen der Proportionalkammern in 16 Intervallen abgebildet. Das **zV***er***t***e***x** **Cl***uster* Element fordert im Histogramm Spuren in vier benachbarten Intervallen.

DCrϕ_Tneg

Das Element **D***rift* **C***hamber* **r**ϕ-**T***rack Trigger for* **neg***ative charges* verlangt mindestens eine Driftkammerspur mit negativer Ladung. In dem untersuchten Zerfallskanal von J/ψ Ereignissen ist es dadurch gerechtfertigt, weil man eine Elektronspur erwartet.

LAR_BR

Das **L***iquid* **A***rgon* & **B***ig***R***ay* Triggerelement bildet eine Korrelation zwischen einer Spur in den Proportionalkammern und einer Energiedeposition im Flüssig-Argon-Kalorimeter, die die Schwelle von $1,5\,GeV$ überschreitet. Die verlängerte Gerade zwischen der Spur und der Energie im Kalorimeter wird *BigRay* genannt. Der Θ-Winkelbereich $5° - 155°$ wird damit eingeschlossen [26].

Topo_Mult < 3

Die **Topo***logical BigRay* **Mult***iplizität* ist die Anzahl der identifizierten *BigRay*. Es wird verlangt, daß nicht mehr als zwei *BigRay* vorhanden sind. Von dem gewählten Zerfall der J/ψ erwartet man nämlich ein Elektron und ein Positron.

DCRϕ_THig

Das Element **D***rift* **C***hamber* **rϕ-T***rack Trigger for* **H***igh transversal impuls* verlangt mindestens eine Spur in der Driftkammer, die einem Teilchen mit einem transversalen Impuls P_t größer als 800 MeV gehört.

Topo_BR

Das **Topo***logical* **B***ig***R***ay* Element fordert die Existenz zweier gegenüberliegender Proportionalkammerstrahlen. Zwischen den Strahlen wird ein Fehler $\Delta\phi = \pm 30°$ in dem Öffnungswinkel geduldet, aber mehr als zwei Strahlen dürfen nicht nachgewiesen worden sein.

Beide Subtrigger fordern also einen z-Vertex in der Wechselwirkungszone und mindestens eine negative Spur in den Kammern.
Der Subtrigger **52** fordert dazu eine Koinzidenz (*BigRay*) von Energiedepositionen im Flüssig-Argon-Kalorimeter und Spuren von geladenen Teilchen in den Proportionalkammern und beschränkt die Anzahl dieser Koinzidenzen auf maximal zwei.
Der Subtrigger **54** verlangt dagegen genau zwei *BigRay* und mindestens ein geladenes Teilchen mit hohem transversalen Impuls.

6.3 Die J/ψ-Daten aus 1997

Die Clusteralgorithmen werden sowohl auf Physikdaten als auch auf Untergrunddaten getestet, aber die Physikdaten bieten die Möglichkeit an, Korrelationen mit den von den Algorithmen herleiteten Größen zu überprüfen.
Zum Schluß dieses Kapitels werden deshalb die wichtigsten Histogramme der Physikselektion gezeigt.

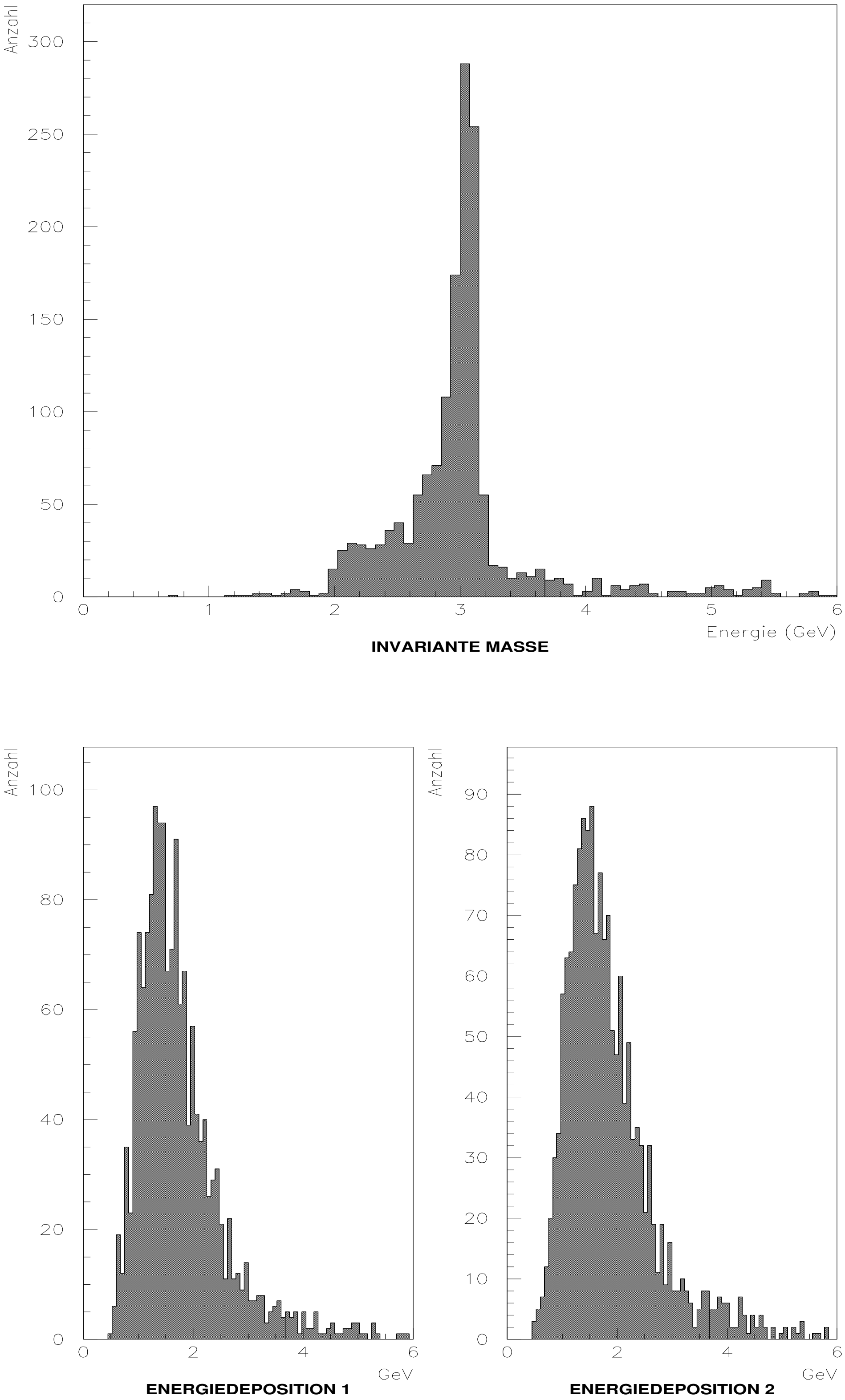


Abbildung 6.1: J/ψ invariante Masse (oben) und Energiedepositionen in dem Flüssig-Argon-Kalorimeter der beiden Zerfallselektronen (unten).

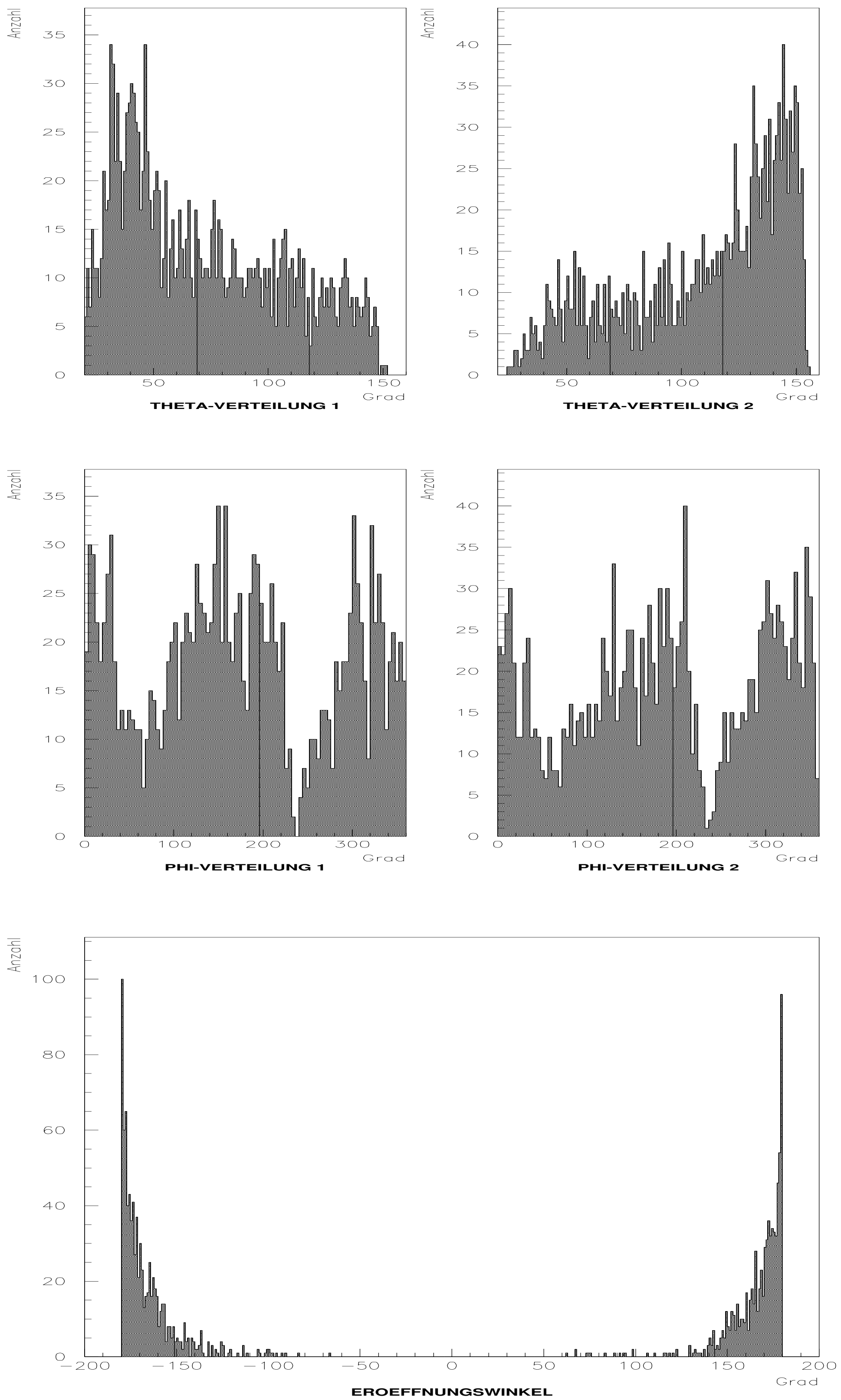


Abbildung 6.2: Die θ-Verteilungen (oben) und die ϕ-Verteilungen (mitte) der beiden Zerfallselektronen, der Eröffnungswinkel $\Delta\phi$ zwischen den beiden (unten).

7 Clusteralgorithmen im L2 Trigger

Der Flüssig-Argon-Kalorimeter ist in jeweils 252 indizierten Bigtower für den elektromagnetischen und hadronischen Bereich geteilt. Ein Bigtower erstreckt sich über mehrere Zellen, wo eine Energiedeposition stattfinden kann (s. Abb. 7.1). Analog dazu wird eine 16×16 Karte der Bigray gebildet, die der Lage der Bigtower entsprechen. Die beiden Karten sind in dem Anhang A dargestellt.

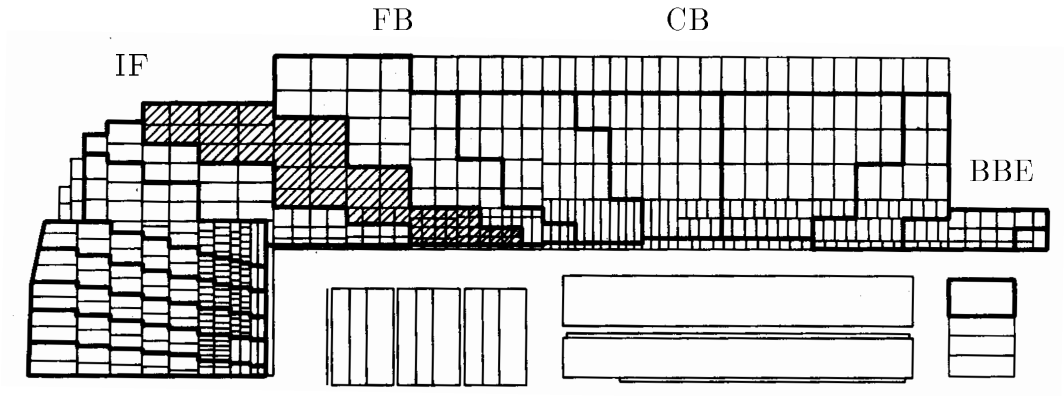


Abbildung 7.1: Flüssig-Argon-Kalorimeter: Der gestrichene Bereich markiert einen der 252 Bigtower.

Die Position einer Bigtower oder Bigray wird durch einen ϕ- und einen θ-Index angegeben. Der ϕ-Index ist proportional zu dem echten ϕ-Winkel, aber er liegt in dem Intervall [0 , 255], weil diese Zahlen in 8 bit kodiert werden können und als Eingabedaten auf dem CNAPS Parallelcomputer handbarer sind [27].
Der θ-Index ist ein reiner Index in dem Intervall [0 , 16]. In Tabelle 7.1 sind die Entsprechenden θ-Winkeln in Grad angegeben.

θ-Index	0	1	2	3	4	5	6	7
θ-Winkel	$1,9°$	$4,6°$	$7,1°$	$9,8°$	$12,4°$	$15,0°$	$18,8°$	$24,2°$
8	9	10	11	12	13	14	15	16
$31,9°$	$42,7°$	$57,9°$	$77,4°$	$107,5°$	$130,3°$	$146,0°$	$160,2°$	$171,8°$

Tabelle 7.1: Liste der θ-Indixen mit den entsprechenden θ-Winkeln.

Während ein Bigray nur die logische Werte 0 und 1 annehmen kann, liegt der Wert einer Bigtower in dem Intervall [0 - 255] und ist proportional zu der aufgetretenen Energiedeposition. Eine Einheit entspricht etwa 125 MeV.
In den nächsten Abschnitten werden ein Clusteralgorithmus für die Bigtower und zwei verschiedene Clusteralgorithmen für die Bigray beschrieben.
Die Grundidee besteht darin, aus den vorhandenen Triggerinformationen Größe mit einer physikalischen Bedeutung zu bilden, die besser helfen können, die Physikdaten von den Untergrunddaten zu unterscheiden.
Bei den Clusteralgorithmen werden benachbarten Bigtower oder Bigray zusammengefaßt und zwar an den Stellen, wo sie konzentriert sind und einer über die anderen dominiert oder als Zentrum des Clusters angesehen werden kann.
Mit der Kenntniss der Energiesumme und den Winkelkoordinaten der Cluster verfügt man über eine topologische Information, wo die wichtigsten Ereignisse im Kalorimeter zusammengestellt sind.

7.1 Bigtower Clusteralgorithmus

Bevor der Bigtower Clusteralgorithmus läuft, werden alle Bigtower mit einem Wert kleiner als fünf auf Null gesetzt. In den Bigtower-Einheiten entspricht es einer Energiedeposition kleiner oder gleich 500 MeV, die meistens vom Rauschen verursacht wird. Der Algorithmus läuft dann sequentiell über alle Indixen in dem Intervall [0 - 255] und prüft die entsprechenden Bigtower.
Wenn keine von den unten aufgezählten Bedingungen eintritt, ist der Bigtower ein Clusterkandidat und die Energiesumme mit den benachbarten Bigtowers wird berechnet. Da die Granularität der Bigtower sich in der ϕ-Richtung ändert, ist die Anzahl der benachbarten Zellen nicht konstant. In einer im Programm kodierten Tabelle (*Lookup-table*) sind die Nachbarschaftbeziehungen anhand der Bigtower-indizen festgelegt.
Der gefundene Cluster wird mit den vorhergehenden verglichen und nach Energie sortiert. Am Ende des ganzen Prozeßes werden die drei höchstenergetischen Cluster ausgewählt. Während des Prozeßes werden die folgenden Bedingungen in der angegebenen Reihenfolge geprüft. Sobald eine wahr ist, wird der Index um eins erhöht und der nächste Bigtower bearbeitet:

1. Der Index verweist auf keinen Bigtower.
2. Die Bigtower-Energie ist gleich null.
3. Es gibt einen benachbarten Bigtower, dessen Energie größer ist.
4. Es gibt einen benachbarten Bigtower mit der gleichen Energie, der schon als Cluster identifiziert wurde.

Die erste Bedingung vermeidet, daß ein nicht existierender Bigtower geprüft wird. Da es 252 Bigtower gibt, müssen vier Indixen keinem Bigtower zugeordnet sein. Diese Indixen haben die Nummer 31, 63, 95 und 127.

Wenn die Bigtower-Energie gleich Null ist, kann der entsprechende Bigtower bestimmt kein Clusterkandidat sein. Dieser Fall wird von der zweiten Bedingung behandelt.
Mit der dritten Bedingung werden die benachbarten Zellen betrachtet. Wenn eine solche eine größere Energiedeposition aufweist, dann kann sie ein Clusterkandidat werden. Aber die aktuell bearbeitete kann gleich ausgeschlossen werden. Die vierte Bedingung dient dazu, daß nur eine von zwei benachbarten Zellen, die die gleiche Energie haben und alle anderen Bedingungen erfüllen, um Clusterkandidat zu werden, tatsächlich auch als Cluster identifiziert wird. Ohne diese letzte Bedingung würde der Algorithmus aus benachbarten Zellen mit der gleichen Energie zwei gleiche Cluster statt einem machen (s. Abb. 7.2).

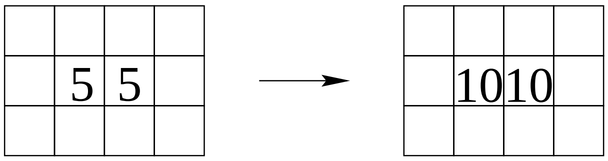


Abbildung 7.2: Ausschnitt aus der Karte: Bigtower (links), Cluster (rechts).

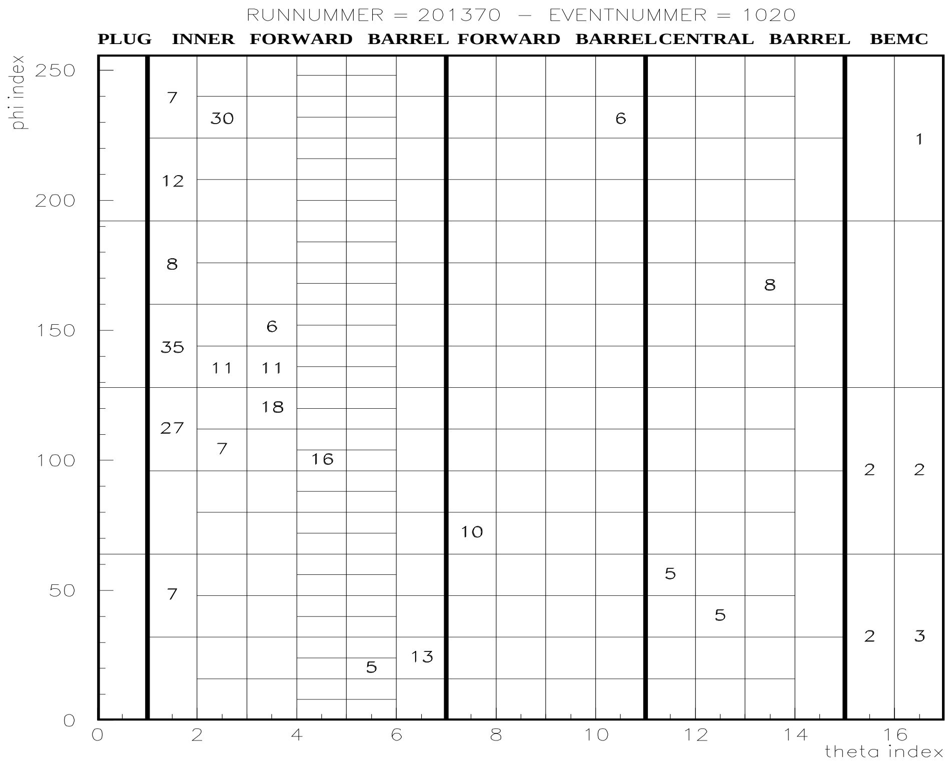
Die Clustersuche wird für den elektromagnetischen und den hadronischen Bereich getrennt durchgeführt. Die totale Anzahl und die Indizes der gefundenen Cluster werden auch gespeichert. Außerdem werden die Winkelunterschiede $\Delta\theta$ und $\Delta\phi$ des zweiten und des dritten Clusters im Bezug auf das erste berechnet und als zusätzliche Variablen geschrieben. Eine Zusammenfassung der erzeugten Größen befindet sich im Anhang B unter DDB Simulation.

Auf den folgenden zwei Seiten befinden sich:

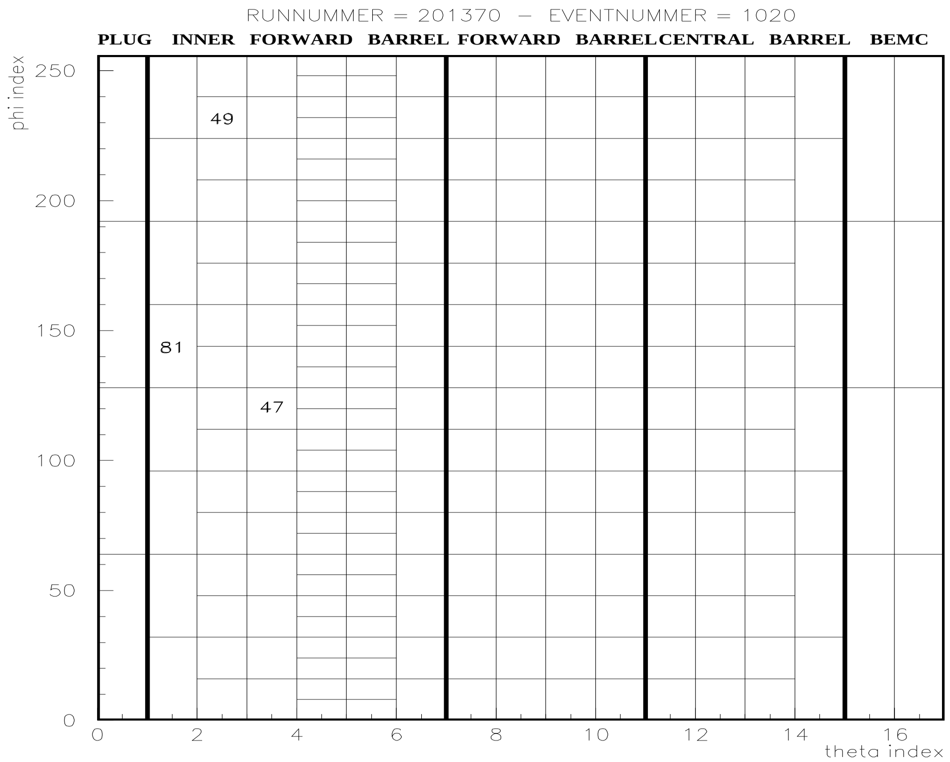
- ein Beispiel für die Clustersuche aus einem reellen Ereignis,
  um zu verdeutlichen, wie der Clusteralgorithmus funktioniert, und

- die Verteilungen der gefundenen Cluster im elektromagnetischen
  und hadronischen Bereich für Physikdaten und Untergrunddaten.

Aus den Cluster-Verteilungen sieht man deutlich die spezifische Merkmale von den selektierten Physikereignissen und Untergrundereignissen.
Bei den Physikereignissen sind die Energiedepositionen im elektromagnetischen Bereich besonders in den **FORWARD BARREL** und **CENTRAL BARREL** ausgeprägt, wo die Zerfallselektronen ihre Spuren hinterlassen. Bei den Untergrundereignissen hat dagegen der **INNER FORWARD BARREL** das größte Gewicht. In dem hadronischen Bereich von beiden Klassen fällt die große Anzahl von Clustern in dem **PLUG**-Kalorimeter auf.



Elektromagnetische Bigtowers



Elektromagnetische Cluster

Abbildung 7.3: Bigtower-Karte (oben) und Cluster-Karte (unten).

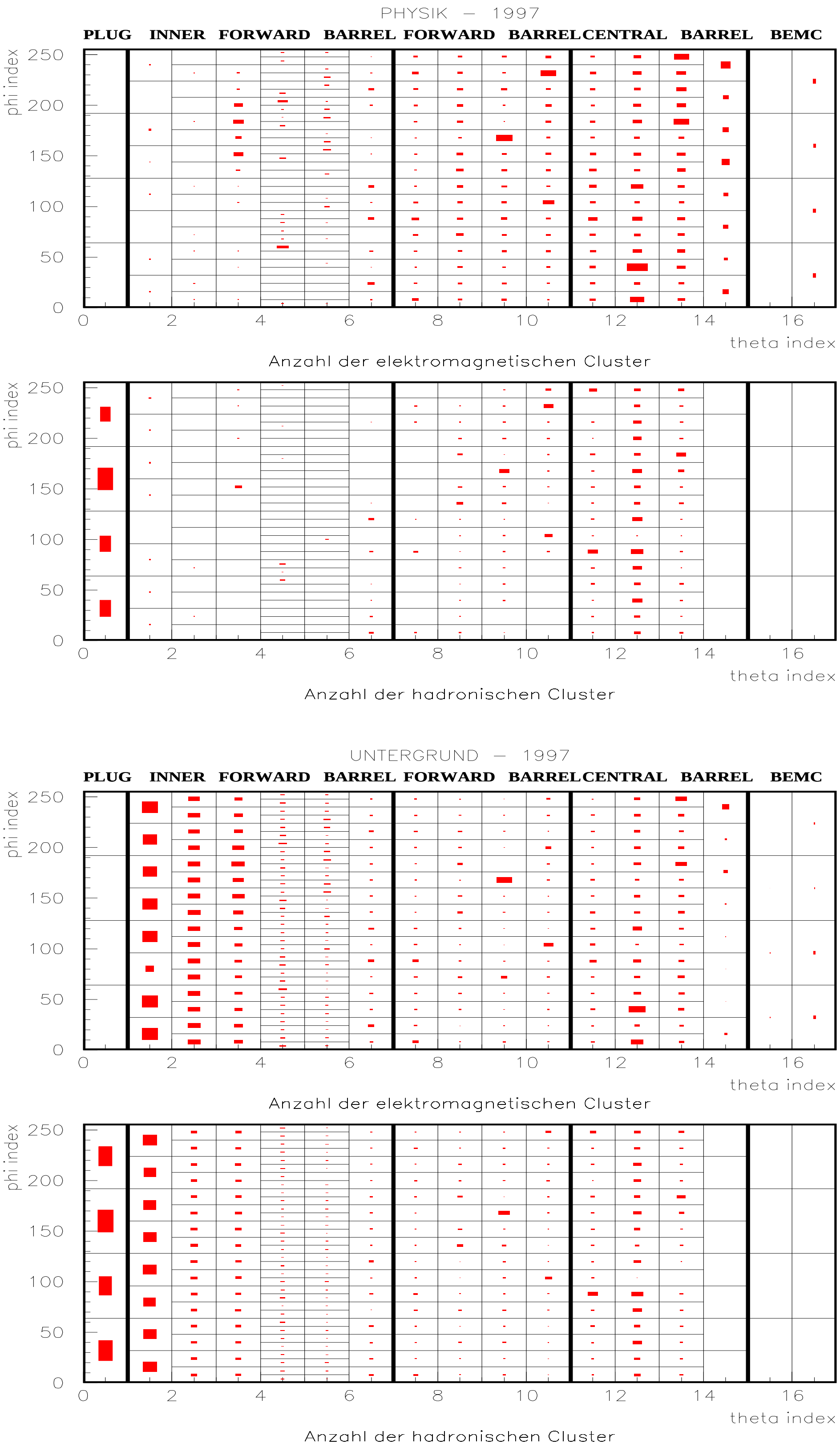


Abbildung 7.4: Cluster-Verteilungen von Physik (oben) und Untergrund (unten).

7.1.1 Korrelation mit der Physik

Es ist interessant, zu untersuchen, ob und wie die elektromagnetischen Cluster der physikalischen Ereignissen mit der Physik korreliert sind.
Da man aber im voraus nicht feststellen kann, welches Cluster eventuell zu welchem Elektron gehört, wurden die $\Delta\phi$ Winkelunterschiede in dem Intervall $[0°\ ,\ 180°]$ von allen Cluster-Elektronen-Kombinationen gebildet.
Das Ergebnis ist in der Abbildung 7.5 gezeigt. Die Ereignisse, bei denen kein elektromagnetischer Cluster gefunden wurde, wurden ausgeschlossen. Sie machen bei dem ersten Cluster etwa $7,3\%$ und bei dem zweiten Cluster etwa $22,8\%$ der gesamten Datenmenge aus.

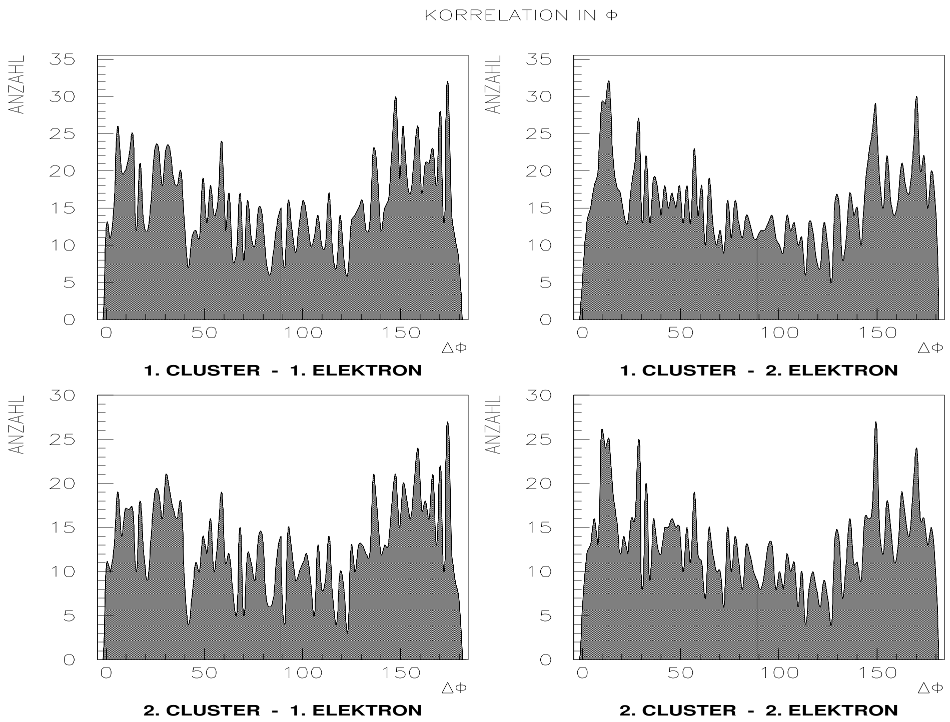


Abbildung 7.5: Winkelunterschiede für die Cluster-Elektronen-Kombinationen.

Wenn ein Cluster einem von den beiden Elektronen entspräche, dann müßte der Winkelunterschied mit diesem Elektron in einer kleiner Umgebung von 0° und mit dem anderen in einer kleiner Umgebung von 180° liegen. Also man würde in den Histogrammen Spitzen bei diesen Werten finden, wenn die meisten Cluster korreliert wären. Da aber diese Spitzen in den Histogrammen nicht besonders ausgeprägt sind, muß man daraus schließen, daß eine schwache Korrelation zwischen den Clustern und den Elektronen besteht.

7.2 Bigray Clusteralgorithmus 1

Der erste Bigray Clusteralgorithmus führt zuerst eine Vorverarbeitung durch und summiert in jeder von Null verschiedenen Zelle die Bits der benachbarten Zellen auf. Auf diese Weise wird aus den Bit-Mustern der Bigray-Karte eine künstliche Karte erzeugt, wo die Bits relativ zu ihrer Konzentration zusammengefaßt sind. Auf dieser Karte läuft dann der Bigtower Clusteralgorithmus (s. Kap. 7.1) und reduziert weiter die Anzahl der Cluster. Am Ende werden die Cluster sich in den Bereichen befinden, wo die meisten benachbarten Bits in der original Bigray-Karte waren. In der Abbildung 7.6 werden die zwei Schritte dieses Prozesses verdeutlicht.

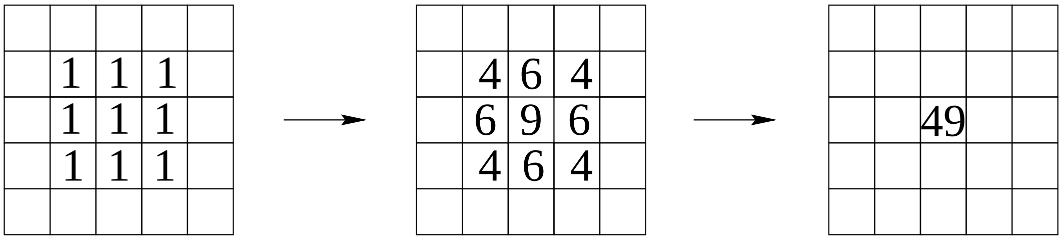


Abbildung 7.6: Clusterreduktion: die künstliche Karte (mitte) und das Ergebnis nach dem zweiten Schritt (rechts) aus der originellen Bits-muster (links).

7.3 Bigray Clusteralgorithmus 2

Der zweite Bigray Clusteralgorithmus kann die Zellen bei konstantem ϕ parallel bearbeiten, obwohl der Durchlauf in dem Programm selbst sequentiell erfolgt. Für jede ϕ-Reihe werden die Zellen in der Reihenfolge überprüft. Wenn die Zelle verschieden von Null ist, wird die Summe mit den Bits aus den benachbarten Zellen berechnet, ansonsten wird der Wert Null gesetzt. Das Ergebnis wird mit dem Ergebnis der in der Reihe vorhergehenden Zelle verglichen. Wenn es kleiner als das vorhergehende ist, wird die vorhergehende Zelle einen Clusterkandidat, ansonsten wird die nächste Zelle bearbeitet. Am Anfang des Prozesses werden die Ergebnisse der ersten Zellen mit dem vorgegebenen Wert Null verglichen. Die gefundenen Cluster werden aufsteigend nach Bits-Summe sortiert. Eine Zwischenkarte, die so aussehen würde, wie bei dem Bigray Clusteralgorithmus 1, wird nicht direkt erzeugt, aber die einzelnen Einträge werden während des Prozesses hintereinander berechnet, wie in der Abbildung 7.7 gezeigt wird.

Die zwei Algorithmen bringen zu unterschiedlichen Cluster-Karten. Ein Beispiel ihrer Anwendung auf einem reellen Ereignis ist in der Abbildung 7.9 dargestellt. Es werden maximal fünf Cluster für den ersten Algorithmus und neun Cluster für den zweiten von dem Programm gespeichert und nur diese werden auch gezeigt. In dem nächsten Abschnitt wird die Korrelation der beiden Algorithmen anhand ihrer Ergebnisse auf einer Ereignismenge von einem **L2/L4** Transparent Run untersucht.

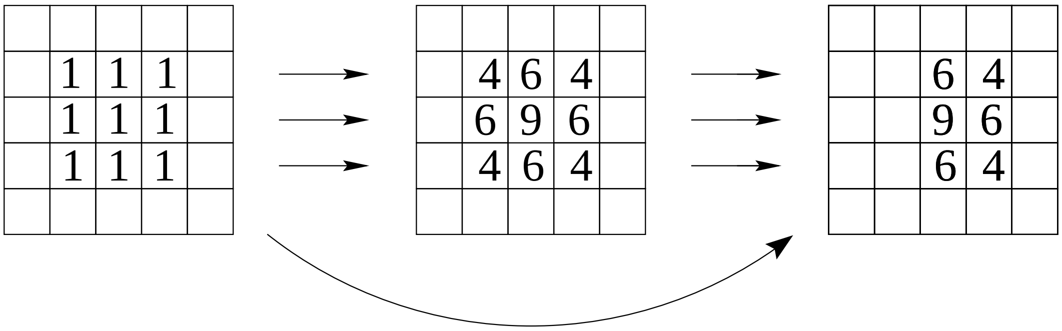


Abbildung 7.7: Zeilenweise Clusterbearbeitung: Bits-Muster (links) und Ergebnis (rechts), die Zwischenkarte (mitte) wird eigentlich in dem Prozeß nicht erzeugt.

7.4 Korrelation der Bigray Algorithmen

Zur Untersuchung der Korrelation zwischen den Bigray Algorithmen wird der $\Delta\phi$ Winkelunterschied zwischen dem ersten bzw. zweiten Cluster von dem ersten Algorithmus und dem ersten bzw. zweiten Cluster von dem zweiten Algorithmus ausgewertet. Da der zweite Algorithmus bei konstanter ϕ ihre Clustersuche durchführt, ist es nach diesem Index am empfindlichsten, weil er in der θ-Richtung schon die Möglichkeit hat, das beste Cluster auszusuchen. Die Abbildung 7.8 zeigt aber eine perfekte Korrelation, was die ersten beiden Cluster betrifft.

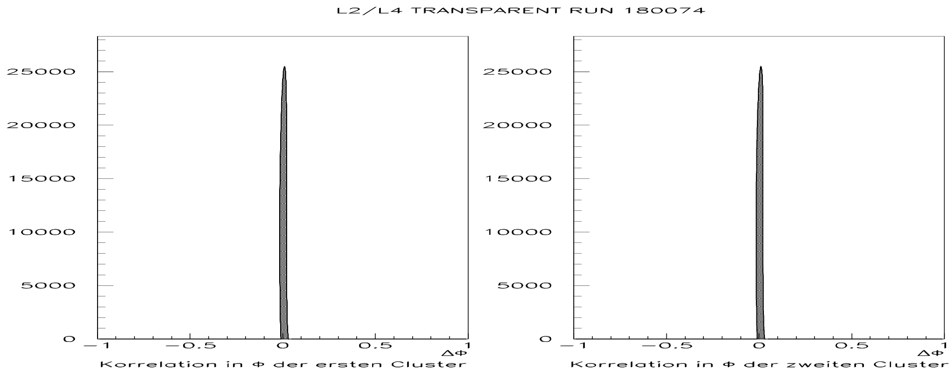


Abbildung 7.8: Korrelation in ϕ der ersten zwei Cluster.

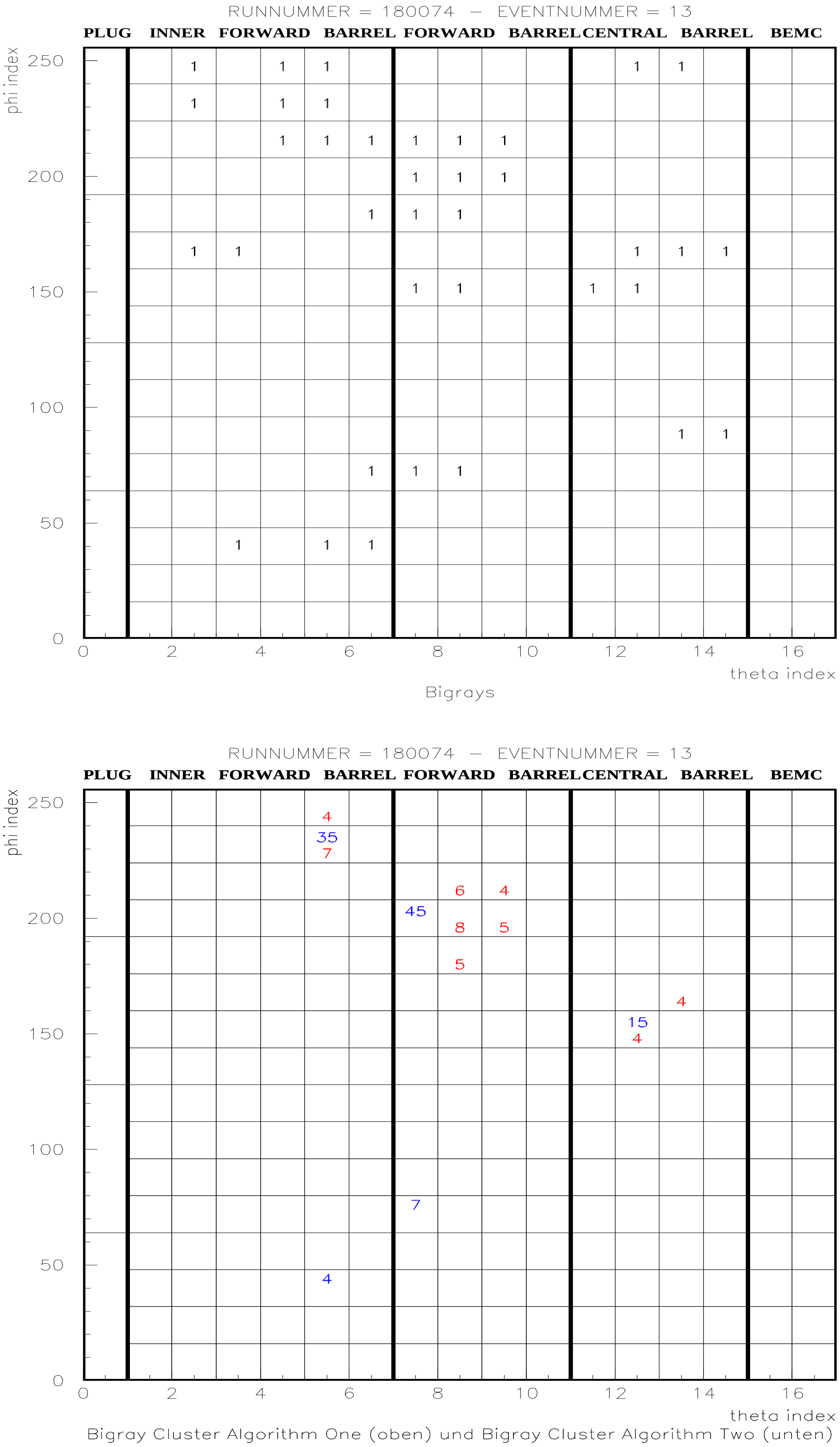


Abbildung 7.9: Bigray-Karte (oben) und Cluster-Karte (unten): in der obenen Reihe sind die Cluster von dem ersten Algorithmus und in der untenen Reihe sind die Cluster von dem zweiten Algorithmus gezeigt.

Worin sich die beiden Algorithmen am meisten unterscheiden, ist in der Anzahl der gefundenen Cluster. In der Abbildung 7.10 ist diese Korrelation gezeigt: die Geraden bezeichnen die Bereiche, wo die Anzahl der Cluster für den zweiten Algorithmus das angegebene Vielfache der Cluster für den ersten Algorithmus ist.

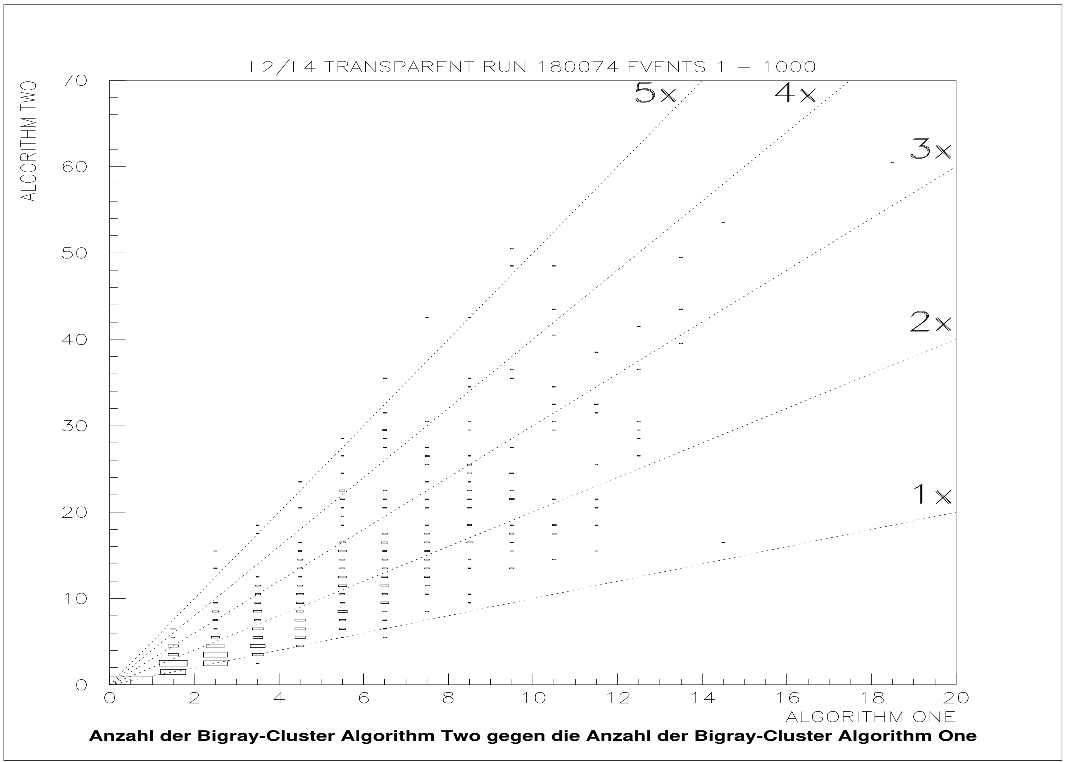


Abbildung 7.10: Korrelation in der Anzahl der gefundenen Cluster.

7.5 Zusammenfassung

Zwischen den beiden Bigray-Algorithmen ist der zweite wegen der parallelen Bearbeitung zu bevorzugen, besonders wenn man nur an den zwei wichtigsten Clustern interessiert ist, wie es oft in physikalischen Zerfallsreaktionen der Fall ist. Im L2 Trigger stehen nur 20 μs für die Verarbeitung der Daten zur Verfügung.

Der Bigtower-Algorithmus kann parallel auf den einzelnen Zellen durchgeführt werden, wenn man auf die vierte Bedingung verzichtet. Sie setzt die Kenntnis der schon identifizierte Cluster voraus. Die Folge wäre eine Art *Doublecounting* bei isolierten Zellen mit der gleichen Energie (s. Abb. 7.2).

Man muß letztendlich auch den eingebauten Sortieralgorithmus für die Cluster-Energien oder Cluster-Bits beachten. Er sortiert einen Cluster, sobald er gefunden wird. Die maximale Anzahl der notwendigen Vergleiche wie auch die der notwendigen Verschiebungen in der Reihenfolge der Cluster beschränkt sich auf:

$$\frac{n(n-1)}{2}, \tag{7.1}$$

wobei n die Anzahl der zu sortierenden Cluster ist. In einer parallelen Verarbeitung kann man diesen Sortieralgorithmus nicht anwenden und man ist gezwungen, die Liste zu sortieren, nachdem alle Cluster gefunden wurden.

8 Das Netzwerktraining

Um ein neuronales Netz zu trainieren und zu testen, muß man eine ganze Reihe von Programmen anwenden, die unterschiedliche Funktionen haben (s. Abb. 8.1).

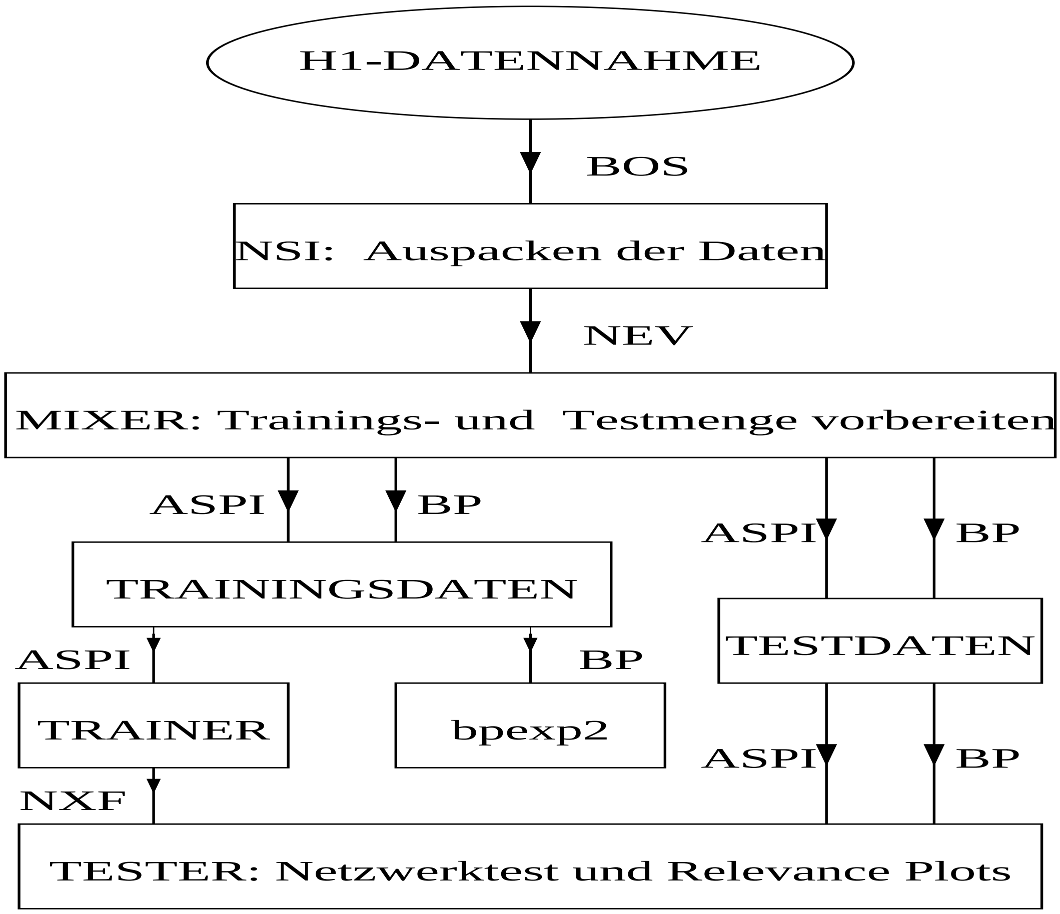


Abbildung 8.1: Kette der Anwendungen für das Netzwerktraining.

Die Triggerdaten werden in **D**ata **S**torage **T**ape (**DST**) oder **P**hysics **O**utput **T**ape (**POT**) in **BOS**-Format gespeichert [28].
Die **POT** sind viel umfangreicher, weil sie auch die rekonstruierte physikalische Daten enthalten. Die Daten werden gruppiert und in **BOS**-Bänken geschrieben. Jede **BOS**-Bank hat ein vierbuchstabiges Kürzel. Die auf dem L2-Bus ankommende Daten befinden sich z.B. in der **BOS**-Bank TL23. Mit dem Kommando `lhb` bekommt man eine kurze Beschreibung des Inhaltes und der Datenformate.

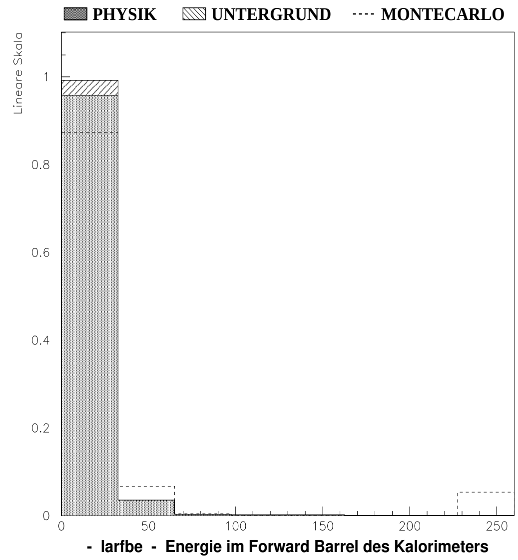
8.1 Die Trainingskette

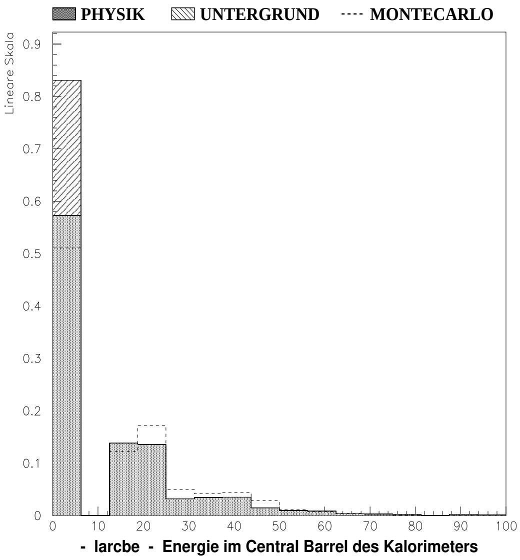
Am Anfang der Trainingskette steht das Programm `nsi`. Es simuliert *offline* neuronale Netzwerke, deren Gewichtskonfiguration vorgegeben ist [29].
Es kann auch dazu benützt werden, die notwendigen **BOS**-Bänke auszupacken, um auf die Triggerinformation zuzugreifen. Die Daten werden in einer **nev**-Datei gespeichert, die das passende Format für das Programm `mixer` hat. Es handelt sich um eine ASCII-Datei, die aus Platzgründen mit dem Programm `gzip` komprimiert wird. Die Datenstruktur ist in einem sogenannten Eventvector festgelegt (siehe Anhang B), wo die Eingabegrößen für die neuronalen Netze nach der Zugehörigkeit zu den Detektorkomponenten gruppiert sind. Zusätzlich werden allgemeine Informationen und die L1-Triggerelemente gespeichert.
Aus der offiziellen Version wurde mit dem Kommando `cvs checkout` eine lokale Kopie in dem Verzeichnis `/home/iws70/concetti/programs` gemacht und nachbearbeitet. Zwei Routinen wurden hinzugefügt, die die Eingabegrößen in der Gruppe DDB-Simulation herstellen. Diese Routinen benützen die schon gespeicherte Information in der Gruppe Detailed-Info und deshalb werden sie in dem Hauptprogramm aufgerufen, nachdem der Eventvector fertiggestellt worden ist. Sie heißen `nsi_ddb2lar` für den Flüssig-Argon-Kalorimeter (Bigtower) und `nsi_ddb2prop` für die Proportionalkammern (Bigray) und der Source-Code befindet sich in der Datei `nsi_ddb2.c` in dem Unterverzeichnis `nsi/src`.
Es werden zwei **nev**-Dateien erzeugt: die eine enthält die Physikdaten und die andere die Untergrunddaten. Die Aufgabe des Programmes `mixer` ist, die ausgewählten Eingabegrößen aus den beiden Dateien zu selektieren und zu mischen [30].
Damit wird es versucht, den Zustand auf dem L2 Trigger nachzuahmen, wo die Reihenfolge der ankommenden Daten zufällig ist bzw. von dem Experiment bestimmt wird. Nur so können die neuronalen Netze richtig lernen, zwischen den zwei Klassen zu unterscheiden, und dann im L2 Trigger eingesetzt werden. Gleichzeitig werden zwei zusätzliche Testmengen aufbereitet, die ausschließlich Physikdaten bzw. Untergrunddaten aus Ereignissen enthalten, die in der gemischten Selektion nicht vorkommen. Sie dienen dazu, die Effizienz der Netze auf unabhängige Datensätze zu testen. Dieses Verfahren heißt Generalisierung. Diese Kontrolle hilft dabei, Fälle von *Overfitting* zu erkennen.
Das Programm `mixer` unterstützt zwei Dateiformate: **ASPI** und **BP**. Der Aufruf von dem Programm muß die Option [-B] enthalten, die die Fixierung von bekannten Fehlern bewirkt. Zusätzlich darf die Option [-c] nicht vergessen werden, um eine funktionsfähige **BP**-Datei zu bekommen. Die **BP**-Format ist für den CNAPS-Parallelrechner vorgesehen [31]. Das Netzwerktraining kann mit dem Programm `bpexp2` durchgeführt werden. Für die untersuchten Netzwerke in dieser Arbeit wurde die **ASPI**-Format benützt. Der Name stammt aus dem zugrundeliegenden Softwarepaket Aspirin/MIGRAINES [32]. Diese Software ist in dem Programm `trainer` eingebunden, das eine Schnittstelle für den Benützer bereitstellt, um Netze zu trainieren [33].

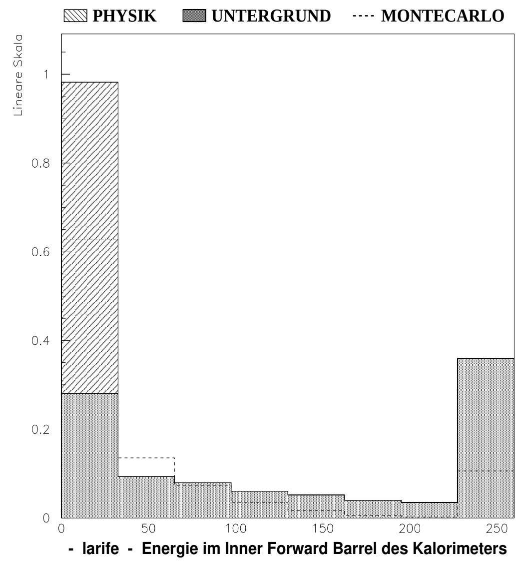
Das Ergebnis sind Gewichtskonfigurationen der untersuchten Netze, die in **nxf**-Dateien gespeichert werden. Diese **nxf**-Dateien können von dem Programm `tester` gelesen werden, das die von dem `mixer` erzeugten Testdaten benützt, um die Generalisierungsfähigkeit der Netze zu testen [34].
Die Testdaten werden auch von dem Programm `trainer` verwendet, um eine Kontrolle der Effizienz des Netzes durchzuführen.
Mehrere Netze mit verschiedenen Lern- und Momentumparametern werden zuerst gleichzeitig trainiert. In jedes Netz werden die Daten aus der Trainingsmenge eingegeben. Nach jeder Eingabe rechnet das Netz seinen Ausgabewert und paßt seine Gewichte dem Backpropagation-Algorithmus entsprechend an. Dieser Prozeß wiederholt sich, bis die gesamte Trainingsmenge verwendet worden ist. An diesem Punkt ist eine sogenannte Epoche abgeschlossen worden. Die Nummer der Epochen ist einstellbar und dem Netz wird in jeder Epoche die gesamte Trainingsmenge vorgestellt.
Nach jeder Epoche wird das Netz mit den Testdaten überprüft und der Fehler wird aufgenommen. Wenn das Netz sich ständig verbessert, sinkt der Fehler bis zu einem asymptotischen Minimum. Wenn das Netz aber übertrainiert wird (*Overfitting*), erreicht der Fehler ein Minimum, bevor das gesamte Training abgeschlossen wurde, und beginnt bei den darauffolgenden Epochen zu steigen. So ist die Effizienz des Netzes am Ende des Trainings die beste für die Trainingsdaten, aber nicht die beste für die Testdaten. Wenn es dann im Betrieb genommen wird, wird seine Effizienz auf neue Daten nicht den Wert haben, den man aus dem Ergebnis des Trainings erwartet hätte. Eine Gegenmaßnahme kann aber anhand der Information über die Testdaten unternommen werden, indem das Training nur bis zu der Epoche wiederholt wird, an der der Fehler von den Testdaten sein Minimum erreicht hatte.
Die Effizienz eines Netzes wird im Bezug auf den weggeworfenen Untergrund berechnet. Es werden Schnitte in der Steuerungsdatei (*steering file*) des Programmes `trainer` gesetzt, mit denen man verlangt, daß ein bestimmter Prozentsatz der Untergrundereignisse verworfen wird. Der Prozentsatz der so behaltenen Physikereignisse gibt die gesuchte Effizienz an.
Um die Effizienz der mit den neu geschaffenen Größen trainierten Netze zu untersuchen, wird ein Bezugsnetz mit den herkömmlichen Größen trainiert. Die enthaltene Kalorimetergrößen werden dann durch die neuen Clustergrößen ersetzt und das so veränderte Netz wird wiederum trainiert und seine Effizienz mit der von dem Bezugsnetz verglichen. Es wird versucht, andere Parameter nicht zu ändern, es sei denn die Effizienz wird dadurch beeinträchtigt. Wenn die Zahl der Eingabegrößen z.B. sich erheblich ändert, sollte auch die Zahl der versteckten Knoten entsprechend angepaßt werden, damit alle eventuelle Korrelationen in den Daten von dem Netz ausgenützt werden können.
Für das Bezugsnetz wurden die Eingabegrößen gewählt, die schon für das Netz Nummer 2 im L2 Trigger mit Erfolg verwendet wurden. Dieses Netz ist auch für den J/ψ-Zerfallskanal in Positron und Elektron abgestimmt worden. Der Vergleich ist so gerechtfertigt, weil er mit derselben Physikklasse durchgeführt wird. Im folgenden werden die Verteilungen der Eingabegrößen gezeigt.

8.2 Analyse der Eingabegrößen

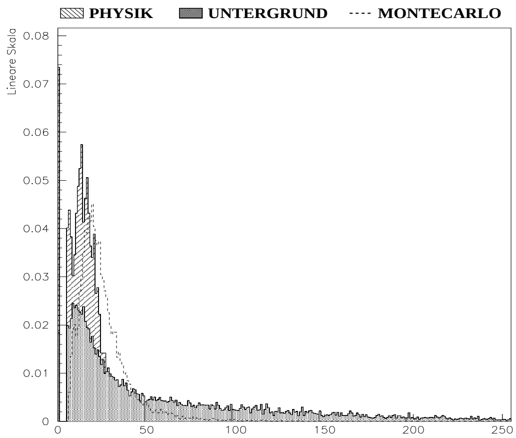
Einige gemeinsame Eingabegrößen wurden in alle Netze eingegeben. Es handelt sich um die Größen `cpvpos`, `cpvmax`, `cpvsum` aus dem z-Vertez Histogramm, `trhineg`, `trhipos`, `trloneg`, `trlopos`, `trtot` aus den Proportionalkammerdaten und `ironfb`, `ironbb`, `irontot` aus den Myonkammerdaten (s. Anh. B). Für das Bezugsnetz wurden außerdem die Kalorimetergrößen `larife`, `larfbe` und `larcbe` verwendet, deren Verteilungen in der Abbildung gezeigt sind.



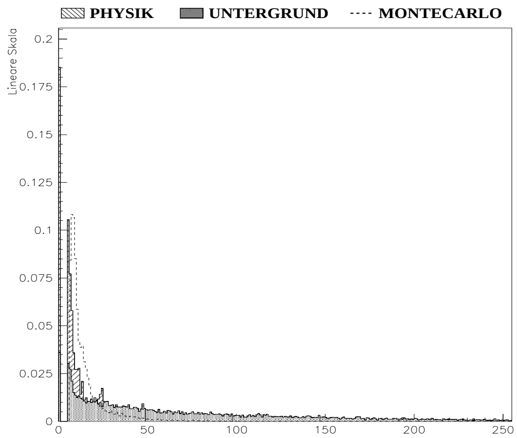




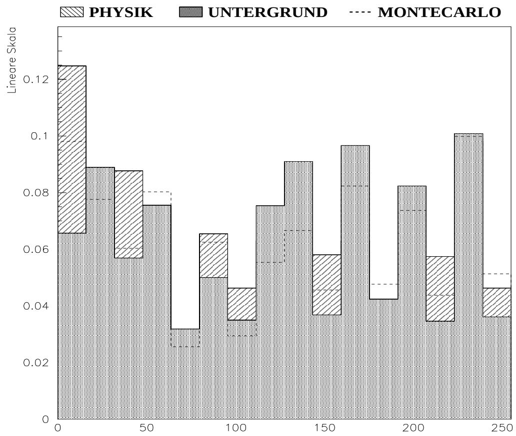
Die Größe `larife`, die die Energiedeposition in dem Inner-Forward-Bereich des Kalorimeters darstellt, bietet eine hervorragende Trennung der zwei Klassen. Fast 90% der physikalischen Datenmenge besitzt den Wert 0, während der Untergrund einen Spitzenwert bei 255 hat.
Auf den nächsten Seiten sind die Verteilungen von Energie, ϕ-Index und θ-Index für die erste drei elektromagnetischen und hadronischen Cluster dargestellt.



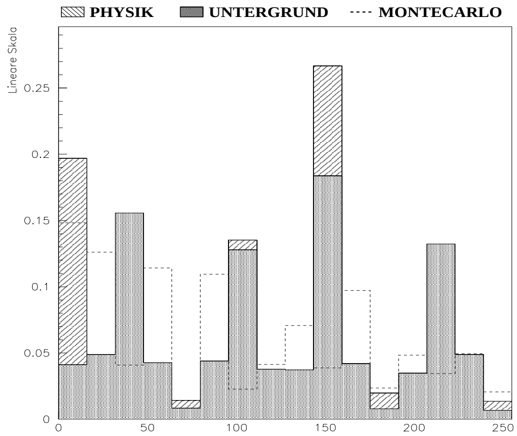
btenej0 **Energie des ersten elektromagnetischen Clusters**



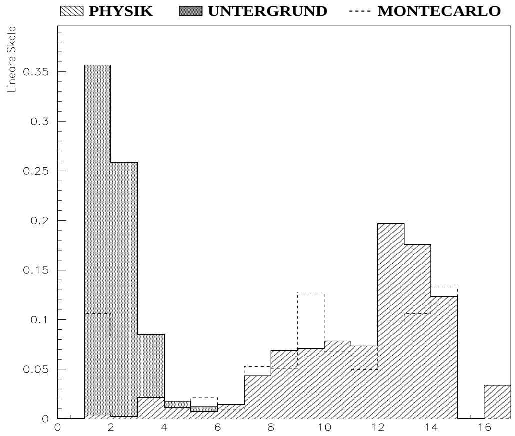
btenhj0 **Energie des ersten hadronischen Clusters**



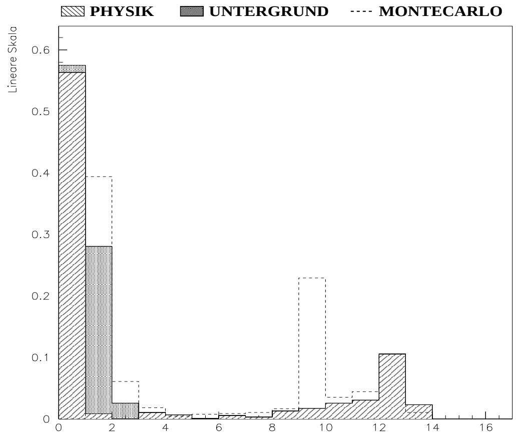
btphej0 **Phi-Index des ersten elektromagnetischen Clusters**



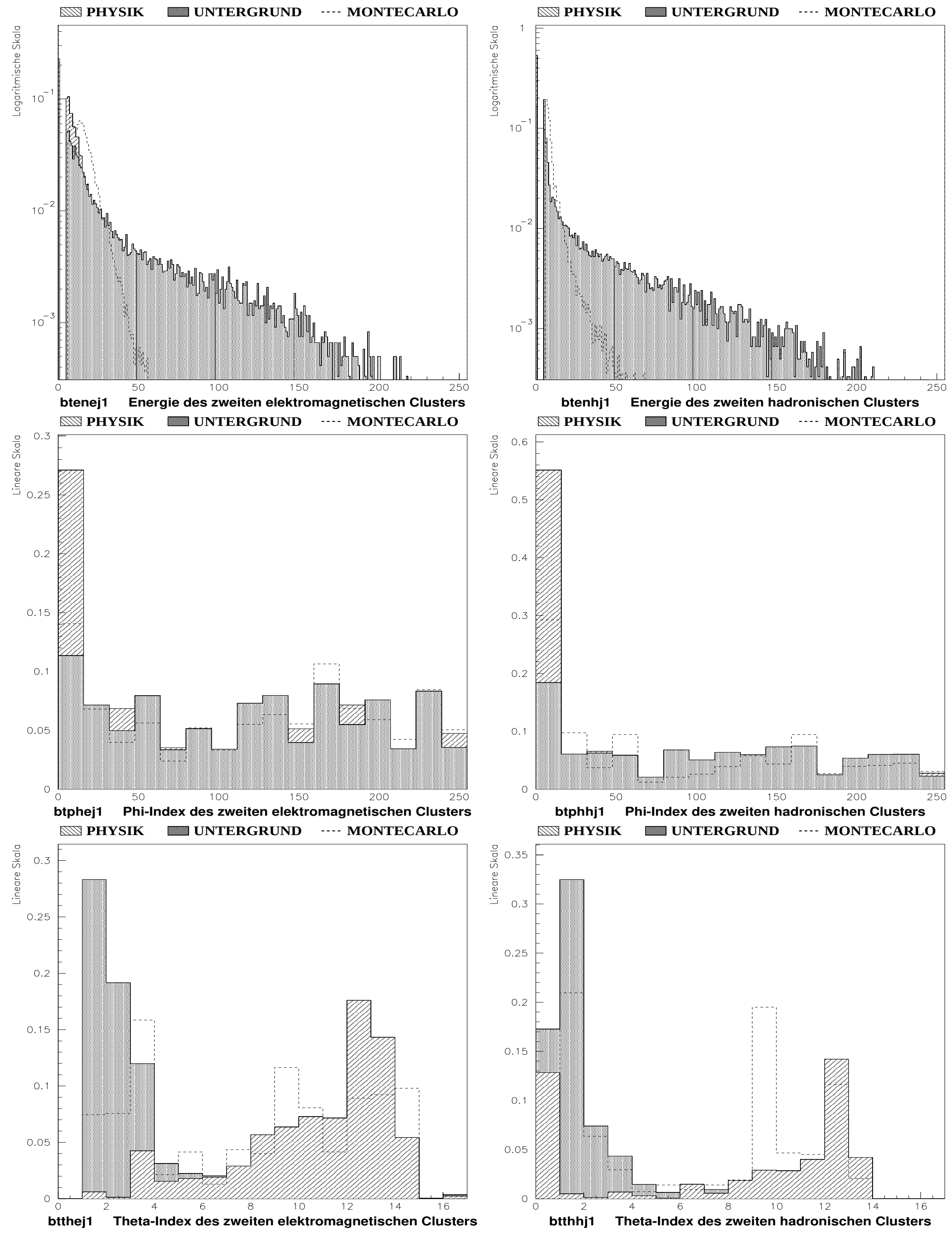
btphhj0 **Phi-Index des ersten hadronischen Clusters**

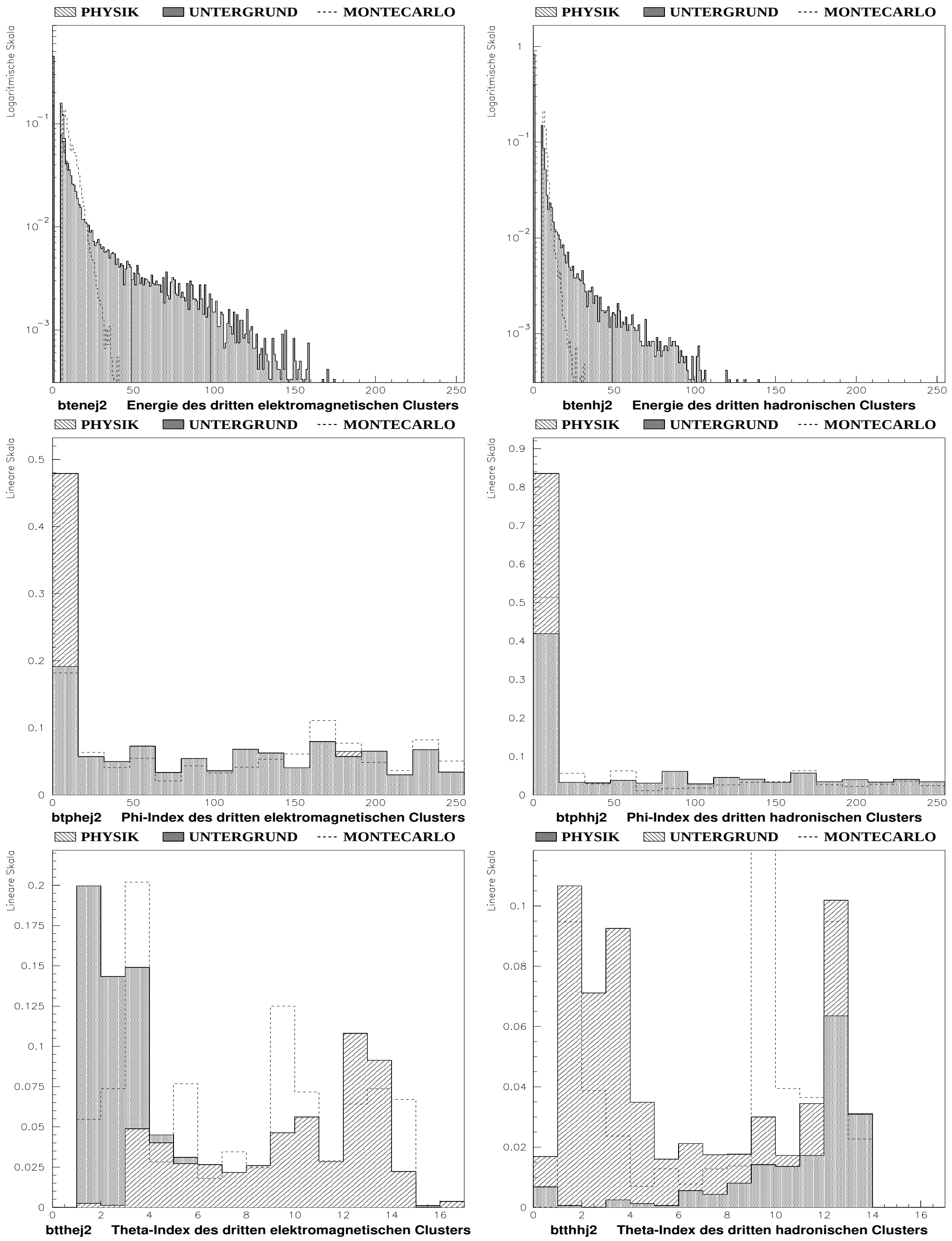


btthej0 **Theta-Index des ersten elektromagnetischen Clusters**



btthhj0 **Theta-Index des ersten hadronischen Clusters**


PHYSIK
UNTERGRUND
MONTECARLO
Logaritmische Skala
btenej1 Energie des zweiten elektromagnetischen Clusters
btenhj1 Energie des zweiten hadronischen Clusters
Lineare Skala
btphej1 Phi-Index des zweiten elektromagnetischen Clusters
btphhj1 Phi-Index des zweiten hadronischen Clusters
btthej1 Theta-Index des zweiten elektromagnetischen Clusters
btthhj1 Theta-Index des zweiten hadronischen Clusters


PHYSIK UNTERGRUND MONTECARLO
Logaritmische Skala
btenej2 Energie des dritten elektromagnetischen Clusters
btenhj2 Energie des dritten hadronischen Clusters
Lineare Skala
btphej2 Phi-Index des dritten elektromagnetischen Clusters
btphhj2 Phi-Index des dritten hadronischen Clusters
btthej2 Theta-Index des dritten elektromagnetischen Clusters
btthhj2 Theta-Index des dritten hadronischen Clusters

Für das Netzwerktraining wurden die elektromagnetische Größen ausgewählt. Die Tabelle 8.1 enthält eine Zusammenfassung der ausgetauschten Eingabegrößen.

Netz	**Eingabegrößen (+ gemeinsame)**	**Versteckte Knoten**
1	larife, larfbe, larcbe	6
2	btenej0, btphej0, btthej0	6
3	btenej0, btphej0, btthej0 btenej1, btphej1, btthej1	7
4	btenej0, btphej0, btthej0 btenej1, btphej1, btthej1 btenej2, btphej2, btthej2	8
5	larife, larfbe, larcbe btenej0, btphej0, btthej0 btenej1, btphej1, btthej1 btenej2, btphej2, btthej2	9

Tabelle 8.1: Zusammenfassung der ausgetauschten Eingabegrößen mit der Anzahl der versteckten Knoten.

Für jedes Netz wurden alle Kombinationen aus den folgenden Lernraten und Momentumparametern probiert: $\eta = (0.0005,\ 0.001,\ 0.002)$ und $\alpha = (0.2,\ 0.5,\ 0.8)$. Die Anzahl der Epochen belief sich auf 3000. Die Netze wurden einmal mit J/ψ-Daten und einmal mit Montecarlo-Daten trainiert.

8.3 Das Trainingsergebnis

Bei dem Ergebnis des Trainings hat sich herausgestellt, daß die besten Parameter für jedes Netz waren: $\eta = (0.0005)$ und $\alpha = (0.2)$. Diese Netze waren nicht übertrainiert und besaßen die beste Effizienz auf den Trainingsdaten. Zwei Beispiele sind in der Abbildung 8.2 angegeben. Um diese Netze richtig zu vergleichen, wurden sie mit dem Programm `tester` einzeln überprüft.
Außer der Effizienz auf den Testdaten, die eine bessere Auswertungsgröße als die Effizienz auf den Trainingsdaten darstellt, wurden Relevanzplots herstellt, die das Thema des nächsten Kapitels sind.

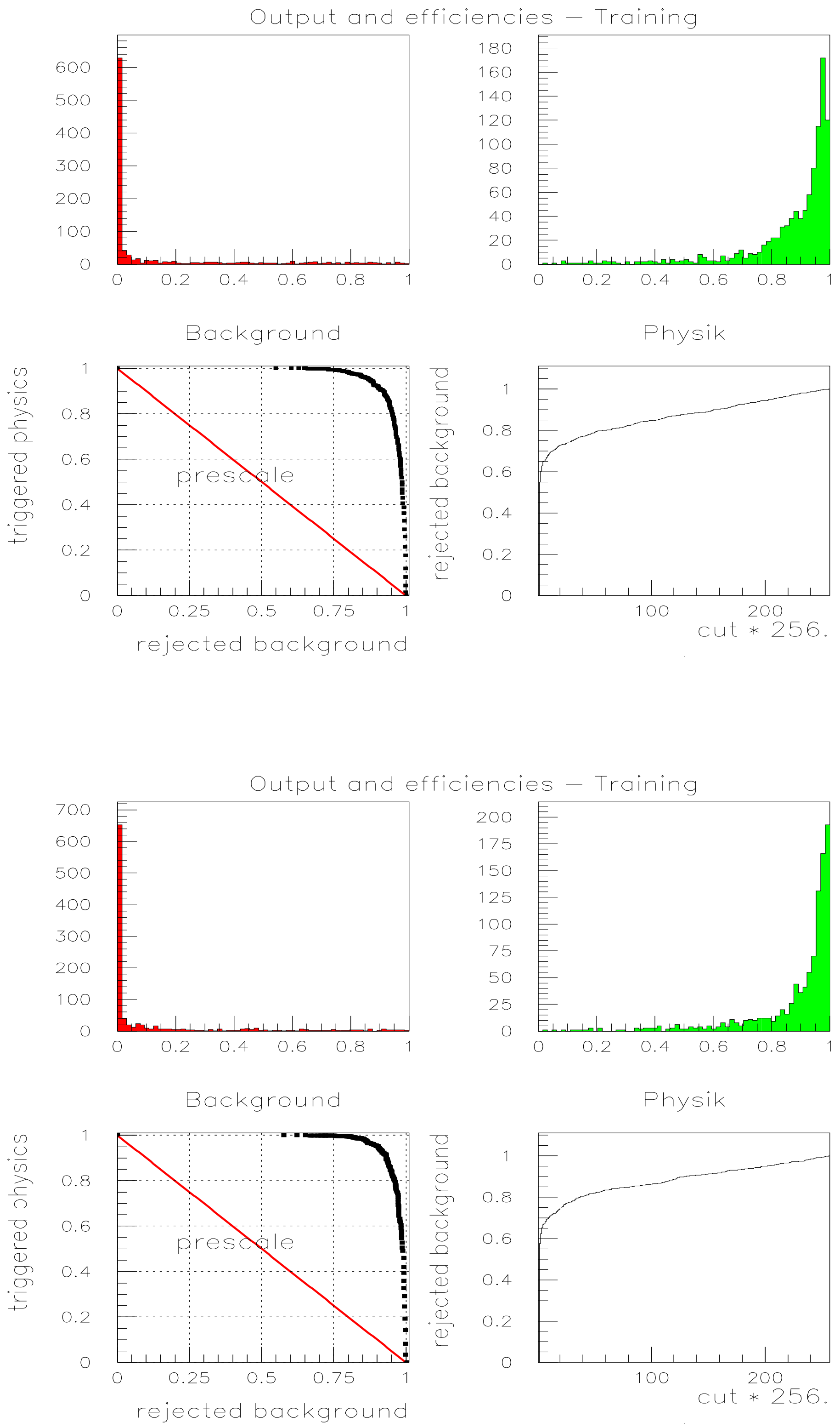
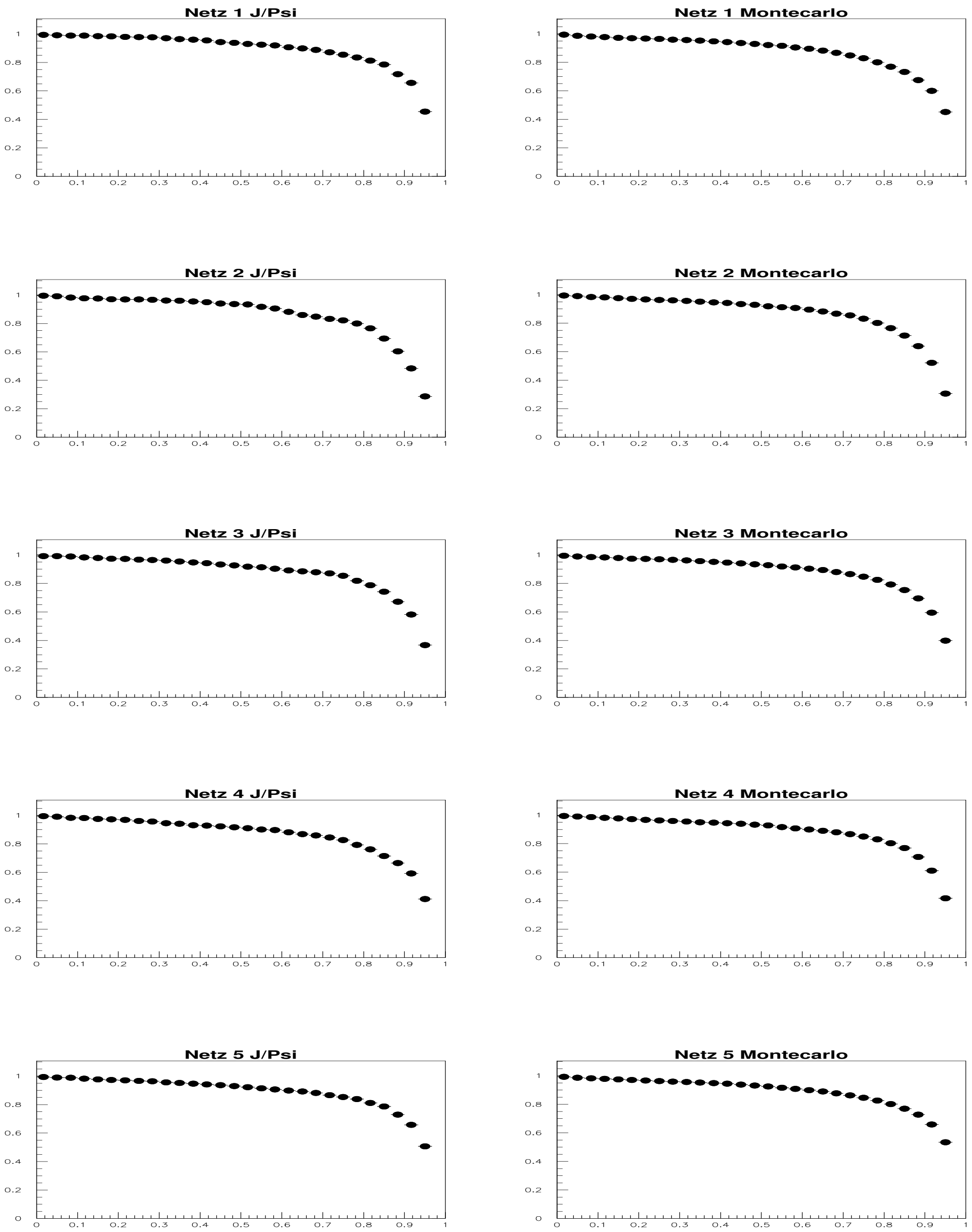


Abbildung 8.2: Die Ausgabewerte und die Effizienzen auf den J/ψ-Trainingsdaten: Netz 2 (oben), Netz 3 (unten). Der *cut* gibt den Schnittwert an, um in der y-Achse gezeigte Physikeffizienz zu bekommen.

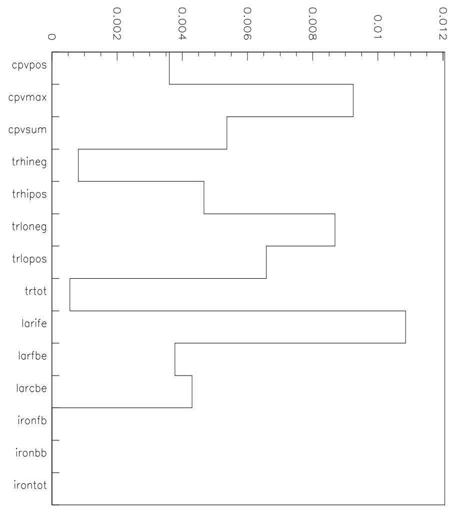
9 Test- und Relevanzplots

Das Relevanzplot ist ein Hilfsmittel, um die Bedeutung der einzelnen Eingabegrößen zu erkennen und auszuwerten. Die *Relevanz* kann als das Maß definiert werden, welchen Anteil eine einzelne Größe an der Klassifikationsentscheidung eines neuronalen Netzes hat [35].
In dem Programm `tester` kann ein Modul eingebunden werden, das diese Relevanzplots während des Tests eines Netzes erzeugt. Auf den folgenden Seiten sind die Effizienzen auf den Testdaten für alle untersuchten Netze und eine Auswahl der aussagekräftigsten Relevanzplots. Die Effizienzen unterscheiden sich nicht sehr viel voneinander. Die Eingabegrößen aus dem Cluster-Algorithmus haben nicht erhebliche Vorteile gebracht. Die Klassifikation wird von der Energiedeposition im Inner Forward Bereich des Kalorimeters dominiert. Unter den neuen Größen schneiden die θ-Indizen der Cluster sehr gut aus, aber sie sind im Prinzip eine Widerspiegelung der Energiedeposition im Inner Forward Bereich, weil die Cluster eine ähnliche Verteilung wie die Energiedepositionen haben. Zu bemerken ist auch, daß das dritte Cluster eine große Rolle in dem Netz spielt, wo es eingegeben wurde. Da der Zerfall in Elektron und Positron erfolgt, sind in den meisten physikalischen Ereignissen nur die erste zwei Cluster bedeutend. Daher unterscheidet sich der dritte in den Untergrunddaten und trägt dazu bei, diese als solche zu erkennen.
Die Effizienzgraphiken sind folgendermaße zu lesen: man wählt die gewünschte Phisikeffizienz auf der y-Achse und liest an der Kurve den entsprechenden x-Wert ab. Dieser besagt, wieviele Untergrunddaten weggeworfen werden.

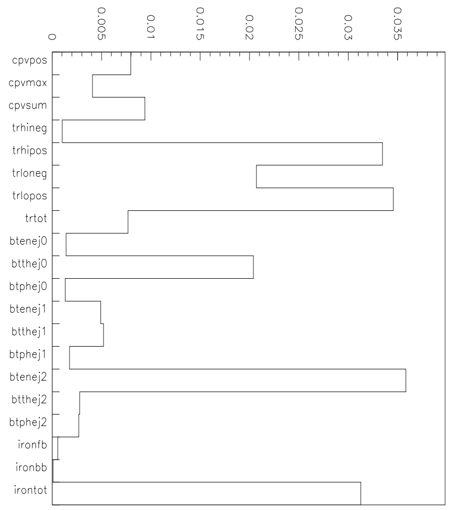


Physikeffizienz gegen Untergrundrejektion

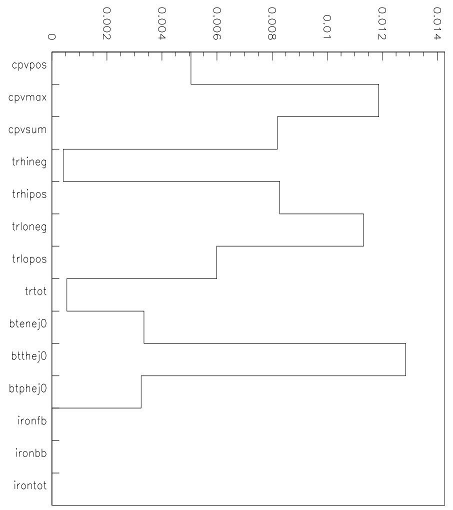
Relevanzplots



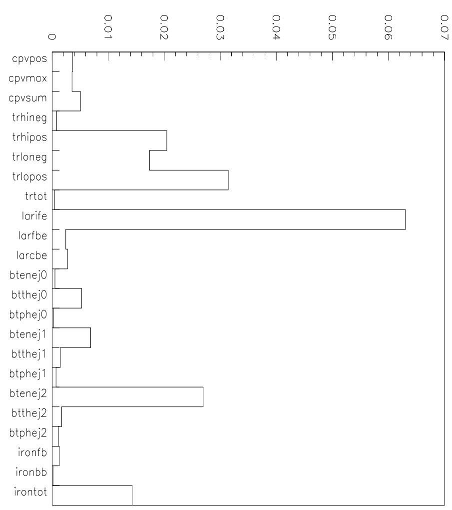
Netz 1 auf Physikdaten getestet



Netz 4 auf Untergrunddaten getestet



Netz 2 auf Physikdaten getestet

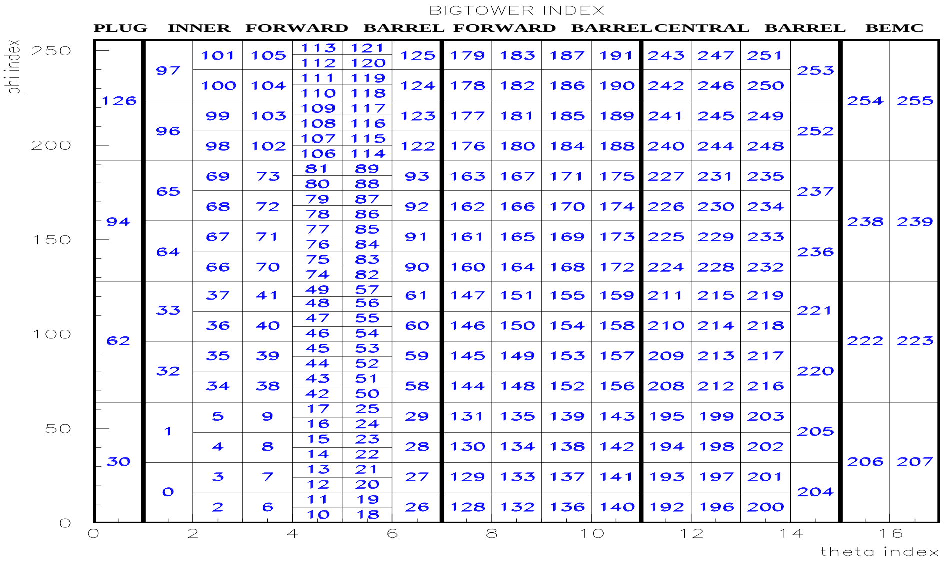


Netz 5 auf Untergrunddaten getestet

10 Zusammenfassung

In dieser Arbeit wurden verschiedene Cluster-Algorithmen entwickelt und verglichen. Der Bigtower-Cluster-Algorithmus summiert die Energiedepositionen in dem Flüssig-Argon-Kalorimeter und sortiert die gefundene Cluster nach Energiegehalt. Der Bigray-Algorithmus 1 und der Bigray-Algorithmus 2 summieren die Bigray-Bits. Obwohl sie unterschiedliche Vorgehensweise haben, stimmen sie bei einer kleinen Anzahl von Clustern überein.
Mit der Clusterinformation, die sich in Energie, θ-Index und ϕ-Index spaltet, wurden neuronale Netze trainiert. Die Grundidee ist eine topologisch geordnete Information anzubieten, die auch physisch relevante Größen enthält.
Der untersuchte Zerfallskanal $J/\psi \longrightarrow e^+e^-$ konnte schon sehr gut von einem vorhandenen Netz selektiert werden. Es ist nicht gelungen, die Ergebnisse von diesem Netz zu übertreffen. Der Lernprozess der Netze wird nämlich von der Energiedeposition in dem Inner Forward Bereich des Kalorimeters beherrscht, weil diese Größe allein erlaubt, eine scharfe Trennung zwischen Physikereignissen und Untergrundereignissen vorzunehmen. Es wäre empfehlenswert, diese Möglichkeit schon auf dem L1 Niveau wahrzunehmen, wo der Entscheidungsprozess über logische Verknüpfungen erfolgt.
Auch wenn die Cluster-Algorithmen für den untersuchten Kanal keine Verbesserung in der Selektion bewirkt haben, stehen sie für Untersuchungen mit anderen Reaktionen zur Verfügung, die sie besser ausnützen können.

Anhang A Bigtower- und Bigray-Karte



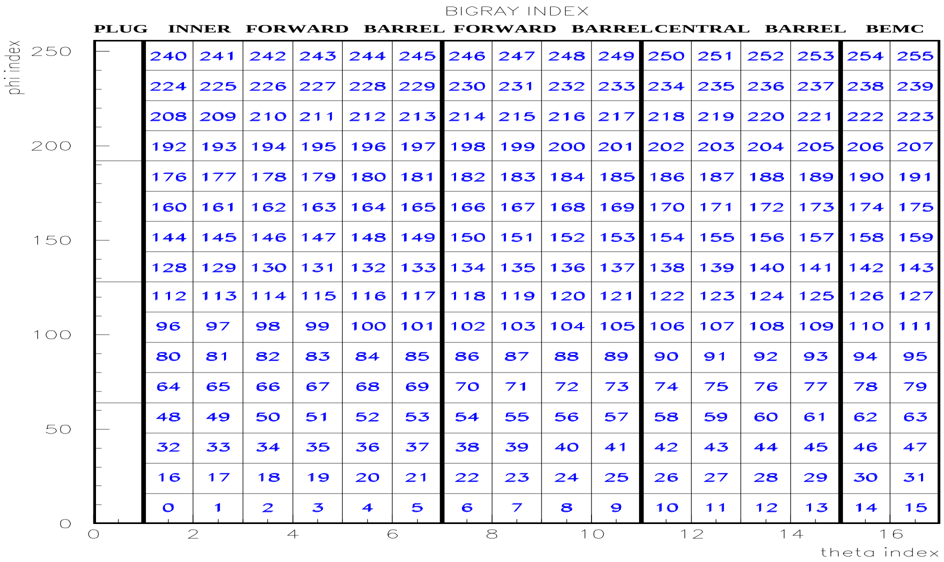


Abbildung A.1: Bigtower-Karte (oben) und Bigray-Karte (unten).

Anhang B L2-Ereignis-Vektor

Identification lables

No.	abbrev.	type	description + comment
0	runnum	%11d	Run number (INT)
1	evtnum	%11d	Event Number (INT)

Analysis Quantities

No.	abbrev.	type	description + comment
2	leptmom	%11d	Lepton momentum (INT)
3	hadmom	%11d	Hadron momentum (INT)
4	x_bj	%7.4f	x Bjorken (FLOAT)
5	y_bj	%7.4f	y Bjorken (FLOAT)
6	Q2	%7.4f	Q2 (FLOAT)
7	gennu	%7.4f	Monte Carlo Quantities (FLOAT)
8	genw	%7.4f	(FLOAT)
9	kti1	%7.4f	(FLOAT)
10	kti2	%7.4f	(FLOAT)
11	wtx1	%7.4f	(FLOAT)
12	wtx2	%7.4f	(FLOAT)
13	pthat2	%7.4f	(FLOAT)
14	shat	%7.4f	(FLOAT)

Trigger Elements

No.	abbrev.	type	description + comment
15-22	tele0-7	%11d	Triggerelements 0-255 (BITFIELD TRELBASE)
23-30	subact0-7	%11d	actual Subtriggerelements 0-255 (BITFIELD)
31-38	subraw0-7	%11d	raw Subtriggerelements 0-255 (BITFIELD L1BASE)
39	l5class	%11d	L5 classification bits (BITFIELD L5BASE)

Validation

No.	abbrev.	type	description + comment
40-45	valid0-5	%11d	Validation bits (BITFIELD VALIDBASE)

L2 Quantities

No.	abbrev.	type	description + comment
46	dcvsum	%3d	sum of entries in rz-histogram (INT)
47	dcvmax	%3d	maximum entry in rz-histogram (INT)
48	dcvpos	%3d	position of maximum in rz-histogram (INT)
49	nbigray	%3d	number of bigrays (INT)
50	nbigfwd	%3d	ïn bin 0 - bin 3 (INT)
51	nbigfce	%3d	ïn bin 4 - bin 6 (INT)
52	nbigbce	%3d	ïn bin 7 - bin 10 (INT)
53	nbigbwd	%3d	ïn bin 11- bin 13 (INT)
54	mbigthe	%3d	mean theta of all bigrays (INT)
55	cpvsum	%3d	sum of entries in zvtx-histogram (INT)
56	cpvmax	%3d	maximum entry in zvtx-histogram (INT)
57	cpvpos	%3d	position of maximum in zvtx-histogram (INT)
58	trhineg	%3d	hits in DCrϕ-chamber neg.charge+high mom. (INT)
59	trhipos	%3d	dito pos.charge+high mom. (INT)
60	trloneg	%3d	dito neg.charge+low mom. (INT)
61	trlopos	%3d	dito pos.charge+low mom. (INT)
62	trtot	%3d	total No. of hits in DCrϕ-chamber (INT)

Calorimeter

No.	abbrev.	type	description + comment
63	pluge	%3d	plug energy (INT)
64	larife	%3d	energy in Inner Forward part of Calo (INT)
65	larfbe	%3d	same in Forward Barrel region (INT)
66	larcbe	%3d	same in Central Barrel region (INT)
67	larex	%3d	x-component of total energy (in L1-format) (INT)
68	larey	%3d	y-component of total energy (in L1-format) (INT)
69	laret	%3d	transversal component of total energy (INT)
70	larele1	%3d	No. of electron candidates(1) in Liquid Argon (INT)
71	larele2	%3d	No. of electron candidates(2) in Liquid Argon (INT)
72	laretot	%3d	Total (weighted) energy in Liquid Argon (INT)
73	calfwde	%3d	summed energy in forward direction (INT)
74	calbare	%3d	summed energy in barrel region (INT)
75	calbwde	%3d	summed energy in backward direction (INT)
76	calbt	%3d	No. of t0-validated Bigtowers (INT)
77	bigtnum	%3d	No. of Bigtowers , em and hadronic (INT)
78	bigterg	%3d	summed energy of all bigtowers (INT)
79	ironfe	%3d	No. of modules hit in Forward Endcap of Iron (INT)
80	ironfb	%3d	No. of modules hit in Forward Barrel of Iron (INT)
81	ironbe	%3d	No. of modules hit in Backward Endcap of Iron (INT)
82	ironbb	%3d	No. of modules hit in Backward Barrel of Iron (INT)
83	irontot	%3d	total no. of hits in Iron (INT)
84	fwmupre	%3d	octant ring hits in pre-toroidal Forward Muon (INT)
85	fwmupos	%3d	octant ring hits in post-toroidal Forward Muon (INT)

No.	**abbrev.**	**type**	**description + comment**
86	spatofha	%3d	No. of hadronic cluster bits for ATOF (INT)
87	spatofem	%3d	No. of electromagnetic cluster bits for ATOF (INT)
88	splatof	%3d	No. of e.m. 'and' hadr. cluster bits for ATOF (INT)
89	spemcl1	%3d	No. of electromagnetic cluster bits for TOF (thresh. 1) (INT)
90	spemcl2	%3d	No. of electromagnetic cluster bits for TOF (thresh. 2) (INT)
91	spemcl3	%3d	No. of electromagnetic cluster bits for TOF (thresh. 3) (INT)
92	spsumcl1	%3d	No. of e.m. and hadr. cluster bits for TOF (thresh. 1) (INT)
93	spsumcl2	%3d	No. of e.m. and hadr. cluster bits for TOF (thresh. 2) (INT)
94	spsumcl3	%3d	No. of e.m. and hadr. cluster bits for TOF (thresh. 3) (INT)
95	spcent1	%3d	No. of cluster bits for inner stack TOF (thresh. 1) (INT)
96	spcent2	%3d	No. of cluster bits for inner stack TOF (thresh. 2) (INT)
97	spcent3	%3d	No. of cluster bits for inner stack TOF (thresh. 3) (INT)
98	spallcl2	%3d	No. of all cluster bits (detailed) for thresh. 2 (INT)
99	spoutcl2	%3d	No. of all outer cluster bits (detailed) for thresh. 2 (INT)

Quadrant Energies

No.	abbrev.	type	description + comment
100	eplq0	%3d	quadrant-energies for Plug (INT)
101	eplq1	%3d	quadrant-energies for Plug (INT)
102	eplq2	%3d	quadrant-energies for Plug (INT)
103	eplq3	%3d	quadrant-energies for Plug (INT)
104	eifq0	%3d	quadrant-energies for IF (INT)
105	eifq1	%3d	quadrant-energies for IF (INT)
106	eifq2	%3d	quadrant-energies for IF (INT)
107	eifq3	%3d	quadrant-energies for IF (INT)
108	efbq0	%3d	quadrant-energies for FB (INT)
109	efbq1	%3d	quadrant-energies for FB (INT)
110	efbq2	%3d	quadrant-energies for FB (INT)
111	efbq3	%3d	quadrant-energies for FB (INT)
112	ecbq0	%3d	quadrant-energies for CB (INT)
113	ecbq1	%3d	quadrant-energies for CB (INT)
114	ecbq2	%3d	quadrant-energies for CB (INT)
115	ecbq3	%3d	quadrant-energies for CB (INT)
116	espq0	%3d	quadrant-energies for SPACAL (INT)
117	espq1	%3d	quadrant-energies for SPACAL (INT)
118	espq2	%3d	quadrant-energies for SPACAL (INT)
119	espq3	%3d	quadrant-energies for SPACAL (INT)
120	ebcq0	%3d	quadrant-energies for BEMC (INT)
121	ebcq1	%3d	quadrant-energies for BEMC (INT)
122	ebcq2	%3d	quadrant-energies for BEMC (INT)
123	ebcq3	%3d	quadrant-energies for BEMC (INT)

Detailed Info

No.	abbrev.	type	description + comment
124-131	bigray0-7	%11d	Bitstream of bigrays (BITFIELD BIGRAYS)
132-133	iron0-1	%11d	one bit for each digi muon module (= 64) (BITFIELD)
134-135	fmupr0-1	%11d	one bit per sector(8) and ring(8) in pre toroid (BITFIELD)
136-137	fmupo0-1	%11d	one bit per sector(8) and ring(8) in post toroid (BITFIELD)
138-169	spcll0-31	%11d	SPACAL cluster bits for TOF, thresh. 2, low byte (INT)
170-201	spclh0-31	%11d	SPACAL cluster bits for TOF, thresh. 2, high byte (INT)
202-297	dcvt0-95	%3d	DC-z-vertex-histogram (96 bins) (INT)
298-313	cpvt0-15	%3d	CIP-COP-z-vertex-histogram(16 bins) (INT)
314-358	cjcs0-44	%3d	tracks in CJC-sectors (45 sectors) (INT)
359-614	btem0-255	%3d	electromagnetic bigtowers in Calo (INT)
615-870	btha0-255	%3d	hadronic bigtowers in Calo (INT)

DDB Simulation

No.	abbrev.	type	description + comment
871-873	btenej0-2	%3d	energy of the BTem jet (INT)
874-876	btthej0-2	%3d	theta of the BTem jet (INT)
877-879	btphej0-2	%3d	phi of the BTem jet (INT)
880-882	seedej0-2	%3d	seeds of the BTem jet (INT)
883-884	dethej0-1	%3d	delta-theta of the BTem jet (INT)
885-886	dephej0-1	%3d	delta-phi of the BTem jet (INT)
887	btnnej	%3d	total number of founded BTem jet (INT)
888-890	btenhj0-2	%3d	energy of the BTha jet (INT)
891-893	btthhj0-2	%3d	theta of the BTha jet (INT)
894-896	btphhj0-2	%3d	phi of the BTha jet (INT)
897-899	seedhj0-2	%3d	seeds of the BTha jet (INT)
900-901	dethhj0-1	%3d	delta-theta of the BTha jet (INT)
902-903	dephhj0-1	%3d	delta-phi of the BTha jet (INT)
904	btnnhj	%3d	total number of founded BTha jet (INT)
905-909	bigena0-4	%3d	clusterbits of the bigray-algorithm one (INT)
910-914	bigtha0-4	%3d	theta of the clusterbits (algorithm one) (INT)
915-919	bigpha0-4	%3d	phi of the clusterbits (algorithm one) (INT)
920-924	bigsea0-4	%3d	seeds of the clusterbits (algorithm one) (INT)
925-928	bigdta0-3	%3d	delta-theta of the clusterbits (alg. one) (INT)
929-932	bigdpa0-3	%3d	delta-phi of the clusterbits (alg. one) (INT)
933	bignna	%3d	total number of founded bits (alg. one) (INT)
934-942	bigenb0-8	%3d	clusterbits of the bigray-algorithm two (INT)
943-951	bigthb0-8	%3d	theta of the clusterbits (algorithm two) (INT)
952-960	bigphb0-8	%3d	phi of the clusterbits (algorithm two) (INT)
961-969	bigseb0-8	%3d	seeds of the clusterbits (algorithm two) (INT)
970-977	bigdtb0-7	%3d	delta-theta of the clusterbits (alg. two) (INT)
978-985	bigdpb0-7	%3d	delta-phi of the clusterbits (alg. two) (INT)
986	bignnb	%3d	total number of founded bits (alg. two) (INT)

Anhang C L2-Raten-Display

Das Macro l2rates.kumac auf dem Rechner h1nnsun.desy.de dient dazu, sich die Raten von den Triggerboxen in dem L2-Trigger-System als Graphen anzuschauen.

Warnung

Das Macro löscht alle Einstellungen und Objekte in dem Arbeitsspeicher von Paw am Anfang und am Ende seiner Ausführung. Deshalb muß man eine getrennte Session von Paw dafür aufrufen, wenn man nicht will, daß die eigene Arbeit gelöscht wird.

Nötige Einstellungen und Aufruf

Um das Macro aufrufen zu können, wie es weiter unten beschrieben wird, braucht man einige Einstellungen, die schon gesetzt sind, wenn man kein eigenes .pawlogon.kumac besitzt. Ansonsten fügt man das Pfad /h1/paw und die Option -auto bei dem Kommando MACRO/DEFAULTS hinzu (z.B. MACRO/DEFAULTS '.,/h1/paw' -auto).
Es reicht dann, l2rates einzugeben und es wird gefragt, welches Datum man möchte. Es ist auch möglich, gleich das Datum einzugeben (z.B. l2rates 30 5 97). Da als Voreinstellung das Systemdatum angenommen wird, kann man das Kommando abkürzen, wie folgt: l2rates 30 5 (im Jahr 1997) oder l2rates 30 (im Mai 1997). Aber l2rates 30 (im Juni 1997) wird als l2rates 30 6 97 interpretiert. Die Bedienung ist dann einfach und braucht keine weitere Erklärung.

Die Graphen

Die Graphen werden so aufgebaut: ein Vektor mit den Zeitstempeln und ein Vektor der gleichen Dimension mit den entsprechenden Raten werden gefüllt und dann wird der zweite gegen den ersten gezeichnet, d.h. die verschiedenen Punkte auf der Ebene x-y (Uhrzeit-Frequenz) werden mit geraden Linien verbunden. Der rote Linie entspricht die Inputrate und der blaue die Outputrate, während die purpure Linie den Ratenunterdrückungsfaktor darstellt, der nach der folgenden Formel berechnet wird:
(1 - output / input) * 100
und gleich 100 gesetzt wird, falls der Input null ist. Die Anzahl und die Zeitabstände der Punkte können unterschiedlich sein und hängen mit dem Programm zusammen, das diese Daten in die Dateien schreibt, die das Macro liest.

Literaturverzeichnis

[1] Stephen F. Mason. Geschichte der Naturwissenschaft. GNT-Verlag, Stuttgart, 1991.

[2] M. Gell-Mann. A schematic model of baryons and mesons. *Phys. Rev. Lett.*, **8**, p. 214, 1964.

[3] G. Pocsik F. Csikor. Standard Model at the Energy of Present and Future Accelerators. Nova Science, 1992.

[4] E. Rutherford. The Scattering of α and β Particles by Matter and the Structure of the Atom. Phil. Mag. 21, 1911.

[5] D. H. Perkins. Hochenergiephysik. Addison Wesley Publishing Company, Bonn, 1990.

[6] B. H. Wiik. HERA Status. In W. Buchmüller und G. Ingelmann, *Proceedings of the Workshop: Physics at HERA*, volume 1, pp. 1–22, Hamburg, 1991. DESY.

[7] P. Waloschek. Das Forschungszentrum DESY. *Praxis d. Naturwiss. Phys.*, **7/38**, 1989.

[8] H1 Collaboration, I. Abt et al. The H1 Detector at HERA. DESY Report DESY H1-96-01, DESY, 1996.

[9] H1 Calorimeter Group, B. Andrieu et al. The H1 Liquid Argon Calorimeter System. *Nucl. Instrum. Meth.*, **A 336**, pp. 460–498, 1993.

[10] H. P. Wellisch et al. Hadronic Calibration of the H1 LAr Calorimeter using Software Weighting Techniques. Technischer Report MPI-PhE/94-03, MPI für Physik, 1994.

[11] H1 Collaboration, . ep Physics beyond 1999. DESY Report DESY H1-97, DESY, 1997.

[12] Steven Weinberg. Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions. *Rev. Mod. Phys.*, **52**, p. 5156, 1980.

[13] H. P. Beck. Measurement of the Total Photoproduktion Cross Section at the Elektron Proton Collider HERA. Diplomarbeit, Universität Zürich, 1996.

[14] H1 Collaboration, S. Aid et al. A Measurement and QCD Analysis of the Proton Structure Function $F_2(x, Q^2)$ at HERA. *Nucl. Phys.*, **B 470**, pp. 3–40, 1996.

[15] Francis Halzen und Alan D. Martin. QUARKS AND LEPTONS: An Indroductory Course in Modern Particle Physics. John Wiley & sons, 1984.

[16] H1 Collaboration, S. Aid et al. A Search for Leptoquarks at HERA. *Phys.Lett.*, **B 369**, pp. 173–185, 1996.

[17] D. Handschuh. Studien zum Protoneninduzierten Untergrund am H1 Detektor und Entwicklung von Filteralgorithmen zu dessen Unterdrückung. Dissertation, Universität Hamburg, 1993.

[18] J. E. Augustin et al. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, **33**, pp. 1406–1408, 1974.

[19] M. Aguilar-Benitez et al Particle Data Group. Review of Particle Properties. *Phys. Rev.*, **D 50**, pp. 530–538, 1996.

[20] R. Eichler. higher lumi and energy of HERA. summary of a discussion with F. Willeke on 21. July 1997.

[21] L. Goerlich et al. Strategy Studies for the H1 Topological L2-Trigger (L2TT). H1–Note H1-01/97-508, DESY, 1997.

[22] C. Kiesling E. Concetti. Upgrade of the Neural Network Trigger: Version II of the Data Distribution Box. Interner Bericht der MPI-Gruppe, November 1997.

[23] W. Pitts W. S. McCulloch. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bull. Math. Biophys.*, **5**, 1943.

[24] A. N. Kolmogorov. *Dokl. Akad. Nauk. USSR*, **115**, pp. 953–956, 1957.

[25] D. E. Rumelhart und J. L. McClelland. Parallel Distributed Processing, volume Vol. 1. MIT Press Cambridge, 1986.

[26] T. Jansen. Untersuchung diffraktiver J/Ψ-Erzeugung bei HERA. Diplomarbeit, Universität Hamburg, 1995.

[27] Adaptive Solutions Inc., Beaverton, OR 97006. CNAPS Data Book, 2.0 Version, 1995.

[28] V. Blobel. The BOS-System Dynamic memory management. Technischer Report third updated printing, Universität Hamburg, 1994.

[29] Andreas Gruber. Documentation for the Network Simulator - nsi -. *H1 / L2, MPI Munich, Germany*, October 1995.

[30] Tobias Zobel. Mixer Documentation. *Max-Planck-Institut fuer Phisik*, Maerz 1995.

[31] Adaptive Solutions Inc., Beaverton, OR 97006. CNAPS/VME Board Reference Manual, 1.2 Version, 1994.

[32] R. R. Leighton. The Aspirin/MIGRAINES Neural Network Software. *MITRE Corporation*, October 1992.

[33] Anselm Kruis. Using the trainer, July 1995.

[34] Dominik Westner. Documentation for the Neural Network Tester. *H1 / L2, MPI Munich, Germany.*

[35] J. Möck. Einsatz neuronaler Netze als intelligenter Trigger im H1-Experiment. Diplomarbeit, TU München, 1994.

Danksagung

Diese Arbeit wurde am Max-Planck-Institut für Physik in München angefertigt. Ich danke allen Mitgliedern der Gruppe für die Hilfsbereitschaft und die ständige Unterstützung.

Namentlich zu nennen ist zu allererst Prof. Dr. Christian Kiesling. Er half mir mit guten Ratschlägen, meine Arbeit zu bewältigen. Dr. Armin Wegner danke ich für die vielen Erklärungen, die er mir geduldig gab. Für die Korrektur der Diplomarbeit bin ich Ludger Janauschek verpflichtet. Eine große Hilfe waren auch Markus Enzenberger für die C-Programmiersprache und Burkhard Reisert für den Flüssig-Argon-Kalorimeter.
Uwe Leupold und Günther Manhart waren immer bereit, die vielen Schwierigkeiten mit den Rechnern zu lösen und Marlene Schaber hatte immer ein Ohr offen bei den kleinen Problemen rund um die Diplomarbeit.

Der Max-Planck-Gesellschaft, insbesondere Prof. Dr. Gerd Buschhorn, verdanke ich die Bereitstellung der erforderlichen Arbeitsmittel.

Erklärung

Hiermit versichere ich, die vorliegende Arbeit unter Angabe aller wesentlichen Quellen und Hilfsmittel selbständig verfaßt zu haben.

München, April 1998

(Enrico Concetti)